



Laporan Akhir Projek Penyelidikan Jangka Pendek

**Development of Computational Intelligent
Infertility Detection System Based on
Sperm Motility Analysis**

**By
Assoc. Prof. Dr. Nor Ashidi Mat Isa**

Development of Computational Intelligent Infertility Detection System based on Sperm Motility Analysis

Progress Report

Assoc. Prof Dr Nor Ashidi Mat Isa



**Development of Computational Intelligent
Infertility Detection System Based on Sperm
Motility Analysis
(01-01-05-SF0493)
-Report-**

Assoc. Prof. Dr. Nor Ashidi Mat Isa

<u>Description</u>	<u>Page Number</u>
1 Introduction	2
2 Data Collection	2
3 Video Pre-Processing	2
3.1 Fuzzification Process	3
3.2 Optimization of Fuzzification Factor	6
3.3 Modification of Membership Function	8
3.4 Results of Pre-processing Process	9
4 Video Processing	16
4.1 Conversion of Image Sequence Frame	16
4.1.1 Multi Stage Local Adaptive Thresholding	16
4.1.2 Results of Thresholding Process	19
4.2 Debris Removal Process	22
4.3 Frame Differencing	26
4.4 Analysis of Disconnected Object	29
4.4.1 Initial Tracker of Motile Sperm	29
4.4.2 Case 1	36
4.4.3 Case 2	37
4.4.4 Case 3	38
4.4.5 Case 4	40
4.4.6 Sperm Motility Detection	44
5 Development of New Intelligent Classifier Tool	54
5.1 Information Granule and Semantics	54
5.2 Consistency Model of Conflict Decision	56
5.3 Termination Index	56
5.4 Application in Real World Data	56
6 Verification with Human Sperm Sample	57
7 List of Publications	58
8 Appendix	61

1. Introduction

This report presents the findings obtained from the Development of Computational Intelligent Infertility Detection System Based on Sperm Motility Analysis Project (Project No: 01-01-05-SF0493). This report is divided into six main sections namely introduction, data collection, video pre-processing, video processing development of new intelligent classifier tool and verification with human sperm sample. Thorough analysis on each section is presented and discussed. The sperm trajectory analysis is performed on 100 rat sperm motility videos and the proposed Automated Feature-Based Sperm Motility Analyzer has achieved prominent results with accuracy, sensitivity and specificity of 87.52%, 89.74%, and 90.51% respectively. Therefore, the proposed system is significantly reliable and can be possibly been applied to infertility detection. A new intelligent classifier tool is developed in this project namely evolving information granule (EIG). Findings indicate that the proposed EIG achieves reasonable accuracy, high interpretability in terms of distinct information granules and also that it is reliable with the consistency model. Based on these promising results the proposed system is extended to be verified with human sperm sample.

2. Data Collection

The data acquisition process involves collecting samples of sperm from Sprague Dawley rat. The Sprague Dawley is the most widely used outbred albino rat in biomedical and reproduction researches [1]. It became popular choice among researchers due to its small size, easily maintained and managed as well as its capability to adapt well to new surroundings. Besides that, this species are used as models in this research because their genetic, biological and behaviour characteristics closely resemble those of humans and many symptoms of human conditions can be replicated in rats [1].

In this study, the sperm samples are collected from healthy male Sprague Dawley rats, aged eight to ten weeks. The rats are obtained from Laboratory Animal Research Unit, Health Campus Universiti Sains Malaysia. These rats are housed in polycarbonate cages and maintained on a 12 hours light/dark cycle at 22 ± 2 °C and acclimatized to the environment for seven days at physiology laboratory before sperm motility video can be recorded. The rats are anesthetized with anesthetic shot of xylazil and ketamine cocktail and the rats are left with the effect of the anesthetic agents within 15 to 20 minutes as shown in Figure 1 (a). The xylazil and ketamine cocktail is prepared based on the combination of 100 mg/kg ketamine and 10 mg/kg xylazine for every rat's body mass [2].

Laparotomy is conducted in day eight to extract vas deferens where the sperm samples are extricated from it (Figure 1 (b)). The vas deferens is part of the male anatomy that transports sperm from the epididymis to the ejaculatory ducts. The sperm sample is diluted with formal saline as shown in Figure 1 (c). Then, the sperm sample is placed in 10 μ m deep Makler counting chamber and its motility is observed through microscope and recorded using Pixelink™ P1B68 microscope camera. The data collection procedures are supervised by experienced doctor at Universiti Sains Malaysia Hospital (HUSM) to ensure the collected data are validated and suitable to be used in the sperm motility analysis.

3. Video Pre-Processing

The recorded sperm motility video obtained from the data acquisition process is pre-processed before conducting a segmentation process. In this stage, the sperm motility video is converted into sequence of image frames. Then, the image sequence frame is processed to improve the visibility of the image where significant features of the sperms are enhanced. This is because the recorded microscopic images are susceptible to all kinds of factors such as improper alignment of microscope and inadequate cleanliness of the optical elements which resulted in decrement in the image quality and visual effect. The needs to capture and

analyze poor quality, low contrast microscopy images are almost inevitable which became a frequent problem that confronting scientist. These limitations are particularly severe in this study since the microscope is used as a tool for actual experimentation.

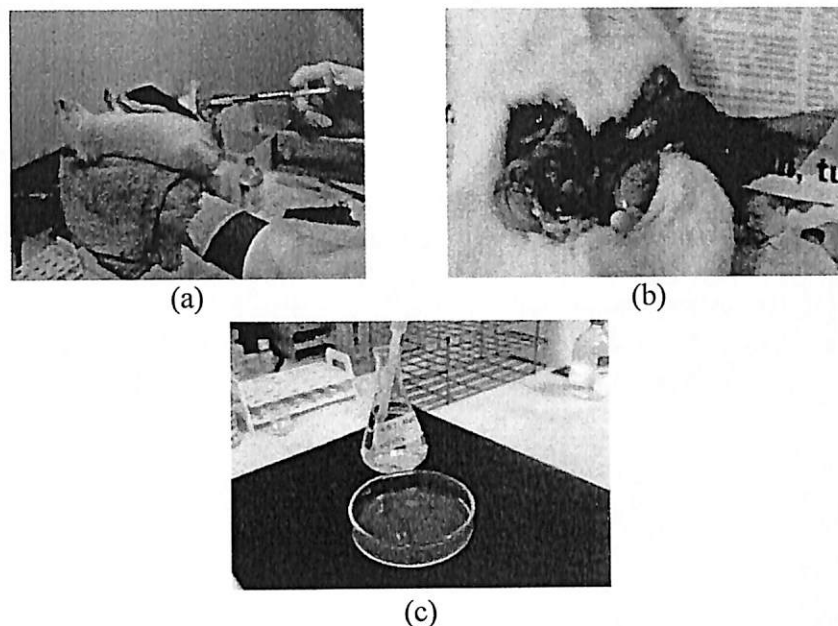


Figure 1 Data acquisition process (a) anesthetic shot with xylazil and ketamine cocktail (b) laparotomy to extract vas deferens (c) dilution with formal saline

The degraded low contrast image sequences in their raw form may still contain useful information that is unseen due to perceptual limitation of the human eye. The brightness needs to be below the detection saturation limit to avoid streaks, smears or over-saturated images in image acquisition stage. Meanwhile, the image intensity also needs to be above a threshold level or the camera aperture needs to be properly adjusted to avoid snow or granulation of the image. Additional electrical fixtures such as light sources, amplitudes, lasers and motors can generate spikes or noise pulses that can appear as spikes across the image. These factors may contribute to the degradation of visual quality of the images and eventually lead to misdiagnose and false detection [3].

Therefore a new enhancement technique namely Adaptive Fuzzy Intensity Measure (AFIM) is proposed. The visibility of the image sequence frame is improved and hidden significant information is revealed. The dominance of sperm features is increased and additional information that are not apparent in the original image whereby regions that are hardly detected by the eyes are retrieved. The proposed AFIM technique adopted the fuzzy approach in enhancing the image sequence frames. This technique consists of three main stages namely the fuzzification process, modification of membership function and defuzzification process as shown in Figure 2.9. Each stage will be elaborated in details in the succeeding subsections.

3.1 Fuzzification Process

In this stage, the sperm motility video that is recorded using microscope camera is converted into sequence of image frame. This image sequence is processed to improve the clarity of the image and to enhance significant features of the sperms. Due to the advent and recent development in imaging technology, the recorded microscopic videos are analysed and interpreted to attain valuable information that could help in sperm motility detection. Video

imaging is currently a popular image acquisition choice for most modern light microscope that has been widely used in medical applications. The video imaging enables faster image acquisition which makes real time imaging possible.

However, video imaging can induce electronic noises in the microscopic images. The microscopic images are susceptible to all kinds of factors such as improper alignment of the microscope and inadequate cleanliness of the optical elements which resulted in decrement in the image quality and visual clarity. The need to capture and analyse poor quality, low contrast microscopy images are a frequent problem confronting scientist/physician. These limitations are particularly severe when the microscope is used as a tool for actual experimentation as in this sperm motility analysis. The degraded low contrast image sequences in their raw form may still contain useful information that is unseen due to perceptual limitation of the human eye. The brightness needs to be below the saturation limit to avoid streaks, smears or over-saturated images during image acquisition stage. Meanwhile, the image intensity also needs to be above a threshold level or the camera aperture needs to be properly adjusted to avoid snow or granulation of the image. Additional electrical fixtures such as light sources, lasers and motors can generate spikes or noise pulses that can appear as spikes across the image. These factors may contribute to the degradation of visual quality of the images and eventually lead to misdiagnose and false detection [3].

Therefore, in the Automated Feature-Based Sperm Motility Analyser System, a new enhancement process is proposed to improve the visibility of the image sequence. This hidden significant information is revealed by the proposed Adaptive Fuzzy Intensity Measure (AFIM) enhancement technique [4]. The dominance of sperm features is increased and additional information that is not apparent in the original image as well as regions that are hardly detected by the eyes are retrieved.

Since the microscopic image sequences are low contrast, most of the pixels possess uneven brightness and their intensity values are vague. The problem in the intensity vagueness occurs when the system tries to determine whether the pixels are bright or dark. Therefore, the fuzzy approach is employed in the AFIM. The fuzzy enhancement process is depicted in Figure 2.

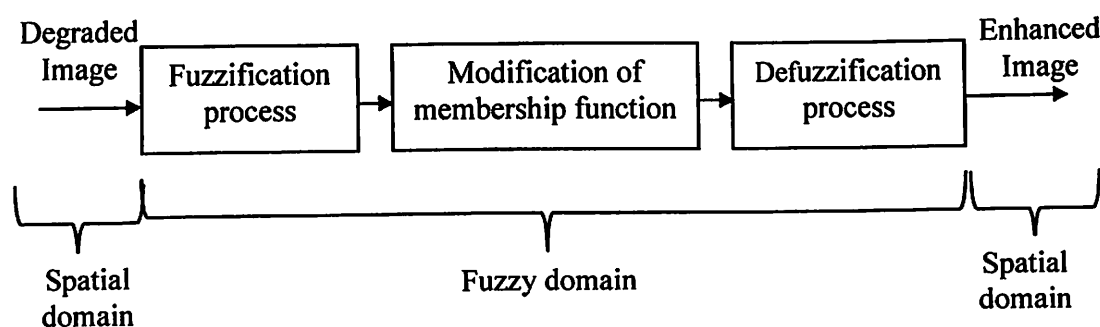


Figure 2 Fuzzy image enhancement process

The enhancement process begins with the fuzzification process. In this stage, the image sequence is divided into bright and dark regions by specifying fuzzy intensity measure which is thresholded using equation (1). The dark region is clustered in the range of $[0, T-1]$, whereas the bright region is clustered in the range of $[T, L-1]$.

$$T_f = L \left[\frac{g_d}{g_a} \right] \quad (1)$$

where

$$g_a = \frac{\sum_{m=0}^{L-1} m \times p(m)}{\sum_{m=0}^{L-1} p(m)} \quad (2)$$

$$g_d = \left[\frac{\sum_{m=0}^{L-1} [(m - g_a)^2 p(m)]}{\sum_{m=0}^{L-1} p(m)} \right]^{1/2} \quad (3)$$

$$\text{fuzzy intensity measure} = \frac{g_d}{g_a} \quad (4)$$

$p(m)$ represents the number of pixels in the histogram of the entire image. g_d and g_a are deviation and mean intensity distributions, respectively. m is the intensity value at pixel (i, j) .

Equations (2) to (4) are derived by considering the mean and deviation of histogram intensity distributions. These equations are calculated to determine the non-homogeneous intensity distributions of the image. This new parameter of fuzzy intensity measure (*i.e.* equation (4)) is proposed which plays a major role in clustering the image into bright or dark regions.

After the image sequence frame is divided into two regions (dark and bright regions) based on the value of T , fuzzification is performed in each region separately. The modified Gaussian membership function is utilized for the fuzzification of the dark region as follows: where $\mu_d(m)$ is the membership function in the dark region and m is the intensity value in the dark region in the range of $[0 T-1]$. m_{avg} and m_{max} are the average intensity and maximum intensity of the image, respectively. ζ_d the fuzzifier function of the dark region, is provided by:

$$\mu_d(m) = \exp \left[- \frac{(m_{max} - (m_{avg} - m))^2}{\zeta_d} \right] \quad \text{for } m < T \quad (5)$$

$$\zeta_d = \alpha \frac{\sum_{m=0}^{L-1} [(m_d - m_{d,avg}) - \sigma_m]^4 p(m_d)}{\sum_{m=0}^{L-1} [(m_d - m_{d,avg}) - \sigma_m]^2 p(m_d)} \quad (5)$$

where σ_m is the standard deviation of intensity of the entire image, $m_{d_{avg}}$ is the average intensity of the dark pixels and m_d and $p(m_d)$ are the intensities and histogram of the dark region, respectively.

The mirror function of the aforementioned Gaussian membership function is utilized to fuzzify the bright region of the image for $m \geq T_f$ as follows:

$$\mu_b(m) = \exp\left[-\frac{(m_{\max} - (m_{avg} - (L - m)))^2}{\zeta_b}\right] \quad \text{for } m \geq T \quad (6)$$

where $\mu_b(m)$ is the membership function of bright region. ζ_b is the fuzzifier function in the bright region.

$$\zeta_b = \alpha \frac{\sum_{m=0}^{L-1} [(m_b - m_{b_{avg}}) - \sigma_m]^4 p(m_b)}{\sum_{m=0}^{L-1} [(m_b - m_{b_{avg}}) - \sigma_m]^2 p(m_b)} \quad (7)$$

where $m_{b_{avg}}$ is the average intensity of the bright pixels, m_b is the intensity of the bright region, and $p(m_b)$ is the histogram of the bright pixels.

The fuzzifier functions of ζ_d and ζ_b calculate the intensity deviation in the dark and bright regions, respectively. The fuzzified image in fuzzy domain is obtained by combining membership functions for both regions (*i.e.* dark and bright regions). α is the fuzzification factor that depends on the intensity values of the input image which will be explained in details in the next subsection.

3.2 Optimization of Fuzzification Factor

The fuzzification factor differs for different input images or image sequence frames as discussed in the previous section. Instead of manually selecting α , this study proposes the automated approach on finding the optimum α . This procedure is performed to ensure the selected α will produce the most pleasant fuzzified image where the illumination are uniformly distributed.

α is selected based on the parameter value that yields the maximum image quality index, Q . Q is a new universal objective image quality index proposed by Zhou and Bovik [5]. This quality index measures any distortions as a combination of three different factors namely loss of correlation, luminance distortions and contrast distortions. These three components are derived as in equation (9) where the first component is the correlation coefficient while second and third components are luminance and contrast distortions respectively.

$$Q = \frac{\alpha_{mF}}{\alpha_m \alpha_F} \cdot \frac{2\overline{m(F)}}{(m)^2 + (F)^2} \cdot \frac{2\tau_m \tau_F}{\tau_m^2 + \tau_F^2} \quad (9)$$

where

$$\bar{m} = \frac{1}{Z} \sum_{y=1}^Z m_y \quad (10)$$

$$\bar{F} = \frac{1}{Z} \sum_{y=1}^Z F_y \quad (11)$$

$$\tau_m^2 = \frac{1}{Z-1} \sum_{y=1}^Z (m_y - \bar{m})^2 \quad (12)$$

$$\tau_M^2 = \frac{1}{Z-1} \sum_{y=1}^Z (F_y - \bar{F})^2 \quad (13)$$

$$\alpha_{mF} = \frac{1}{Z-1} \sum_{y=1}^Z (m_y - \bar{m})(F_y - \bar{F}) \quad (14)$$

In this study, the original and fuzzified images are assumed to have signals of $m = [m_y | y = 1, 2 \dots Z]$ and $F = [F_y | y = 1, 2 \dots Z]$ respectively. m_y and F_y are the intensity levels of the original and enhanced images, respectively. Q is computed by measuring local statistical features and then using a sliding window approach. It starts with the top-left corner of the image, a sliding window of size 3×3 ¹ moves pixel by pixel horizontally and vertically through all the rows and columns of the image until the bottom-right corner is reached. If there are total of H steps to reach the bottom-right corner, then the overall quality index is given by:

$$Q = \frac{1}{H} \sum_{h=1}^H Q_h \quad (15)$$

where Q_h is the local quality index.

Q is in the range of $[-1 \ 1]$ and the best value of Q (i.e. '1') is achieved if and only if m is equal to F . Thus, it is concluded that the optimum α must be selected if it can produce image with highest Q that closest to '1'. Q that closest to '1' indicates the fuzzified image attains homogeneous intensity distribution where the mean luminance error between fuzzified and original images is minimum. It demonstrates that the non-uniform illumination issue in the original image is solved by obtaining more homogeneous intensity distribution without causing any distortion in terms of contrast and luminance.

The optimal procedure for selecting α is described as follows. For a given input image (i.e., original image sequence frame), the value of α is varied from a minimum of 1 to a maximum of 30. For each value of α , the following automated tuning procedures are performed:

- i. Apply the algorithm presented in previous subsection to generate the fuzzified images

¹The smallest sliding window of 3×3 is chosen in this study to ensure more accurate local statistical features can be measured.

- ii. Calculate Q using equation (9)
- iii. Select the parameter value that produces the maximum Q as the optimum value of α , after the two aforementioned steps.

Final output of fuzzified image is generated by adopting the optimum α . Simulations are performed on 100 standard images and 100 sperm motility videos that suffer from non-uniform illumination and low contrast problems to attest this automated optimization of fuzzification factor procedure.

Examples of fuzzification factor optimization on three non-uniform illumination images are shown in Figure 3. The plots of Q in Figures 3 (a), (b), and (c) illustrate the changes in Q as α is varied from 0 to 30. The automated tuning is conducted until a homogeneous image is obtained. The homogeneous image is attained when Q reaches its maximum value. Figure 3 shows that Q reaches its highest value when alpha is 8, 5, and 4 as circled in Figures 3.5 (a), (b), and (c), respectively.

3.3 Modification of Membership Function

Once fuzzification is completed, the original input pixels that exhibit non-uniform illumination and low contrast are transformed into Gaussian distributed pixels. The local contrast of the image is based on intensity difference in a small region and it is computed to preserve the details of the image. Local contrasts are defined for the dark and bright regions as:

$$C_{L_d}(i, j) = \sum_{(i,j) \in W_{i,j}} [\max(\mu_d(i, j)) - \min(\mu_d(i, j))] \quad (16)$$

$$C_{L_b}(i, j) = \sum_{(i,j) \in W_{i,j}} [\max(\mu_b(i, j)) - \min(\mu_b(i, j))] \quad (17)$$

where $\mu_d(i, j)$ and $\mu_b(i, j)$ represent the 3x3 local fuzzified image (*i.e.*, output image obtained after fuzzification process) of μ_d and μ_b , respectively which are centered at position (i, j) . $\max(\mu_d(i, j))$ and $\max(\mu_b(i, j))$ represent the maximum gray level values of the local fuzzified image for dark and bright regions, respectively. $\min(\mu_d(i, j))$ and $\min(\mu_b(i, j))$ denote the minimum gray level values of the local fuzzified image for dark and bright regions, respectively.

Modification of the fuzzified image is performed once the aforementioned steps are executed. Modification is performed to enhance the fuzzified image based on the dark and bright regions, which include the local contrast of the image as shown in equations (18) and (19) respectively.

$$\mu'_d(m) = \frac{1}{1 + e^{\{-C_{L_d}[\mu_d(m) - m_{d,avg}]\}}} \quad \text{for } m < T \quad (18)$$

$$\mu'_b(m) = \frac{1}{1 + e^{\{-C_{L_b}[\mu_b(m) - m_{b,avg}]\}}} \quad \text{For } m \geq T \quad (19)$$

where μ'_d and μ'_b are the modified membership functions in the dark and bright regions, respectively. C_{L_d} and C_{L_b} are the local contrast of dark and bright regions, respectively which are computed to preserve the details in the image.

The above functions modify the original membership functions of $\mu_d(m)$ and $\mu_b(m)$. The modified functions are then defuzzified with the respective inverse membership functions as shown in equation (20). Both regions are combined to obtain the enhanced image. The pixels in the dark region are scaled back to the range $[0 T-1]$, whereas the bright region is translated and scaled back to the region $[T L-1]$.

$$M = \begin{cases} \mu_d'^{-1}(m) & \forall m \leq T \\ \mu_b'^{-1}(m) & \forall m > T \end{cases} \quad (20)$$

where M is the enhanced image obtained from the defuzzification process.

3.4 Results of Pre-processing Process

The results of pre-processing stage are presented in this section. The comparison analysis on sperm images are presented in Figures 3, 5 and 7. These figures show the example of extracted image sequence frames from three different sperm samples namely *sample 1*, *sample 2* and *sample 3* respectively. Figure 3 shows that the example of low contrast image sequence frames where the objects of interest (*i.e.* sperms) are hardly interpreted and undistinguished from the background. Results presented in this figure show that all the enhancement techniques attain better contrast than the original image frame with the contrasts of 16.52, 14.36, 21.47, 15.24 and 12.78 for AFIM, FQM, FHE, FRB and FL respectively. Although, the contrasts are higher than the original image, enhanced image produced by FL has caused over-enhancement. The resulted frame (as shown in Figure 3 (f)) became unnaturally bright where significant features of the sperm are hardly highlighted. The over-enhancement problem has resulted in noise increments since the existing noises in the image are amplified.

Meanwhile, the enhanced images by FQM (Figure 3 (c)) and FHE (Figure 3 (d)) suffer from non-uniform illumination where the bottom left of the image appear brighter than the other regions of the image. The FHE also shows the unnatural enhanced image where saturation can be observed in the top and bottom of the image. The saturated intensities will induce problem during segmentation process where the foreground and the background are difficult to be identified. This problem worsen since the FHE has caused the intensity of the foreground to be saturated (which appears darker) thus makes its illumination become almost similar to the background.

The over-saturated case can also be observed on the non-sperm cells in the enhanced image by FRB. Certain regions in the non-sperm cells are darker than the other cells whereby the significant features of sperm are not highlighted. The enhanced image by FQM (Figure 3 (c)) also fails to improve the clarity of the image and preserve image information as can be depicted in the low entropy value of 5.35.

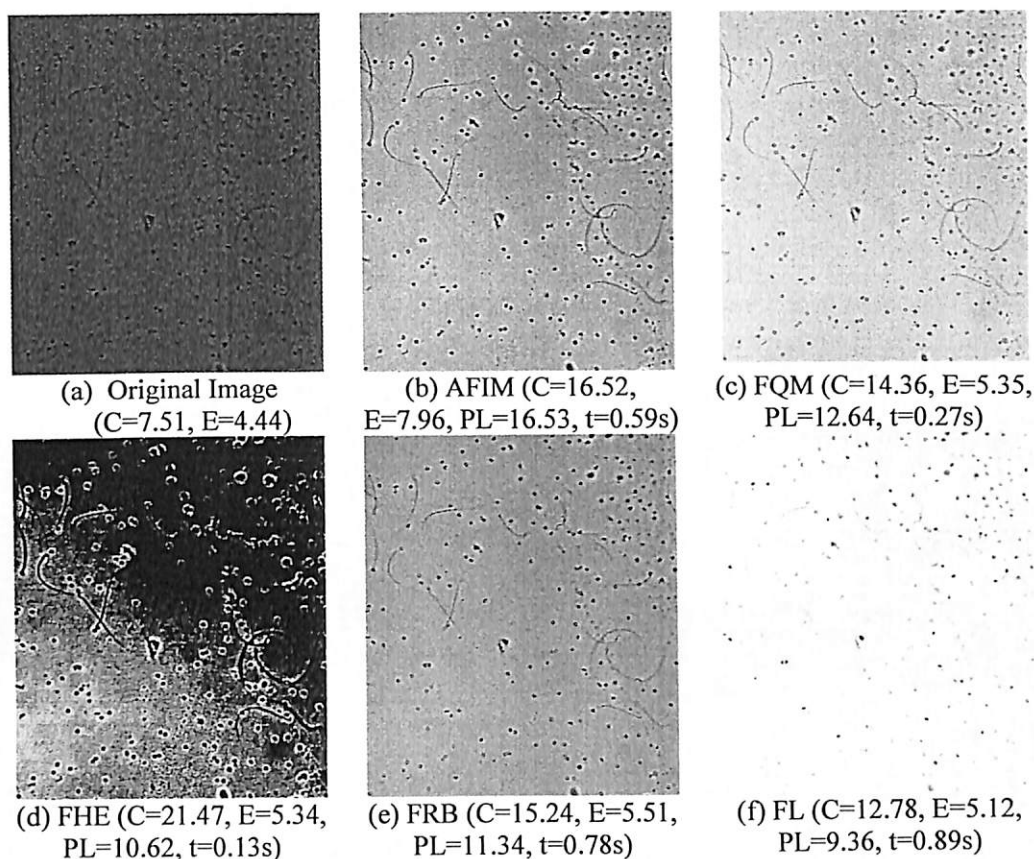


Figure 3 Comparison analysis of 'sample 1' sperm image

Since the sperm images are extracted from a sequence of sperm motility video, analysing quantitative measurement can be done by plotting the E , PL and C for all sequence frames. Since the sperm images are in sequence order, all frames require similar computational time to enhance the images; no significant changes in t can be observed from the first to the last frames. Therefore only average t is considered in the comparison analysis.

This figure shows that the AFIM attains the best results in terms of E and PL where both achieve highest E and PL for most image sequence frames as plotted in Figures 4 (a) and (b). However, as for C , the FRB attains the best contrast where the AFIM is ranked second followed by FHE, FL and FQM. This is because the FRB focuses on improving the overall contrast of the image by relying on the stretching parameter, K . This parameter works well in the uniform brightness image, however when low contrast image and non-uniform illumination problems are observed in one image, the FRB fails to preserve the image details as attested by its lower entropy (Figure 4(a)).

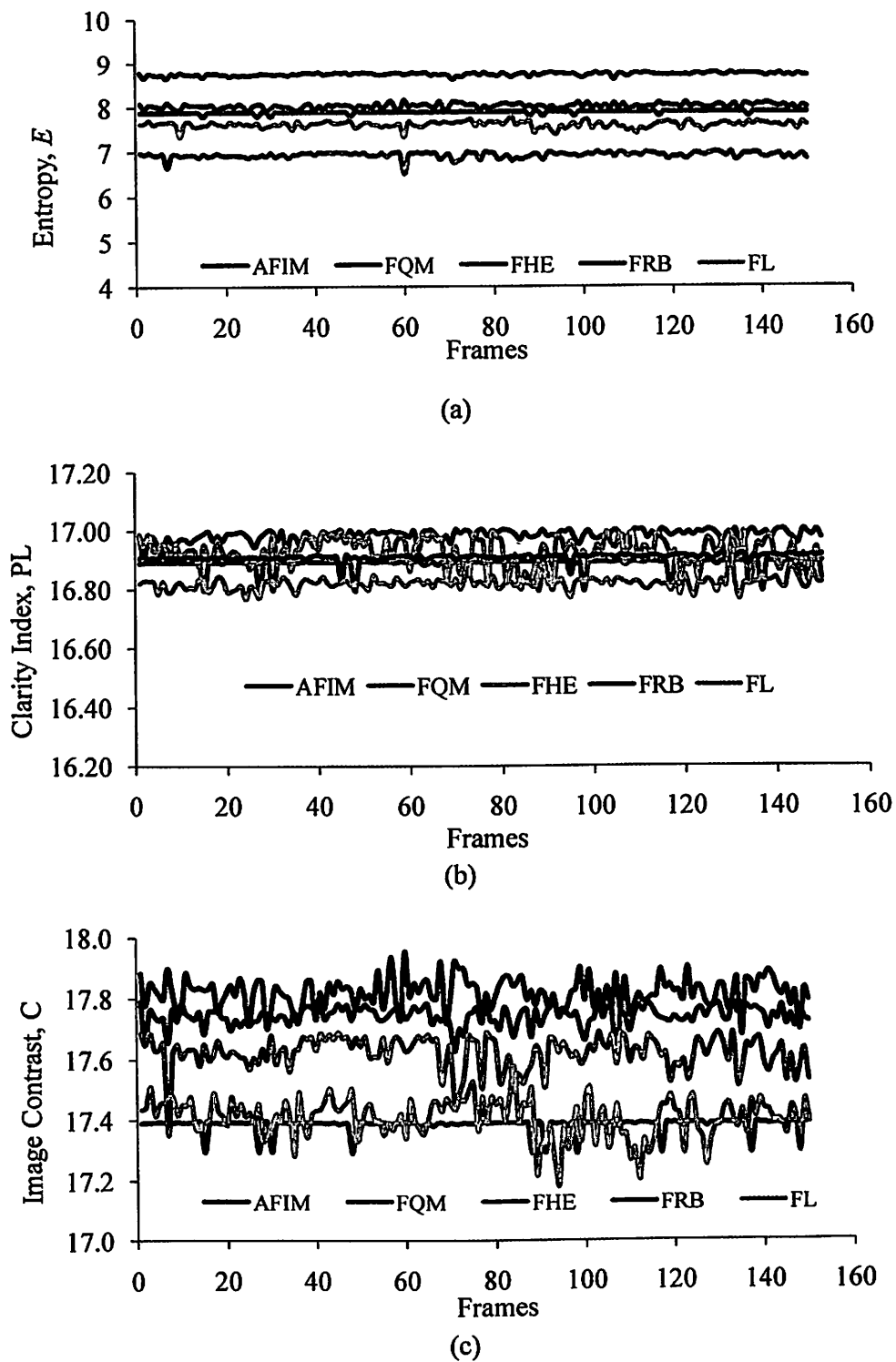


Figure 4 E , PL and C for *sample 1* sperm motility video

The performance of AFIM is also presented in another example of *sample 2* of sperm motility video as shown in Figure 5. The image sequence frame in this sample shows one of the examples of non-uniform illumination and low contrast issues encountered in the recorded image. It can be observed in this figure that the middle region of the image is

brighter than the region at the border of the image. The enhanced image by AFIM in Figure 5 (b) shows that uniform illumination is obtained while the details of the image are preserved.

The dark region accumulated at the border of the image is enhanced by specific function presented in equation (5) whereby illumination of the bright region is improved using equation (7). Both equations have ensured that even brightness is achieved without any under-enhancement or over-enhancement problems. This is because specific functions are utilized to enhance dark region while existing bright region is de-enhanced. In addition, local contrast of the image is intensified using equations (16) and (17); therefore the significant information in the image is sustained.

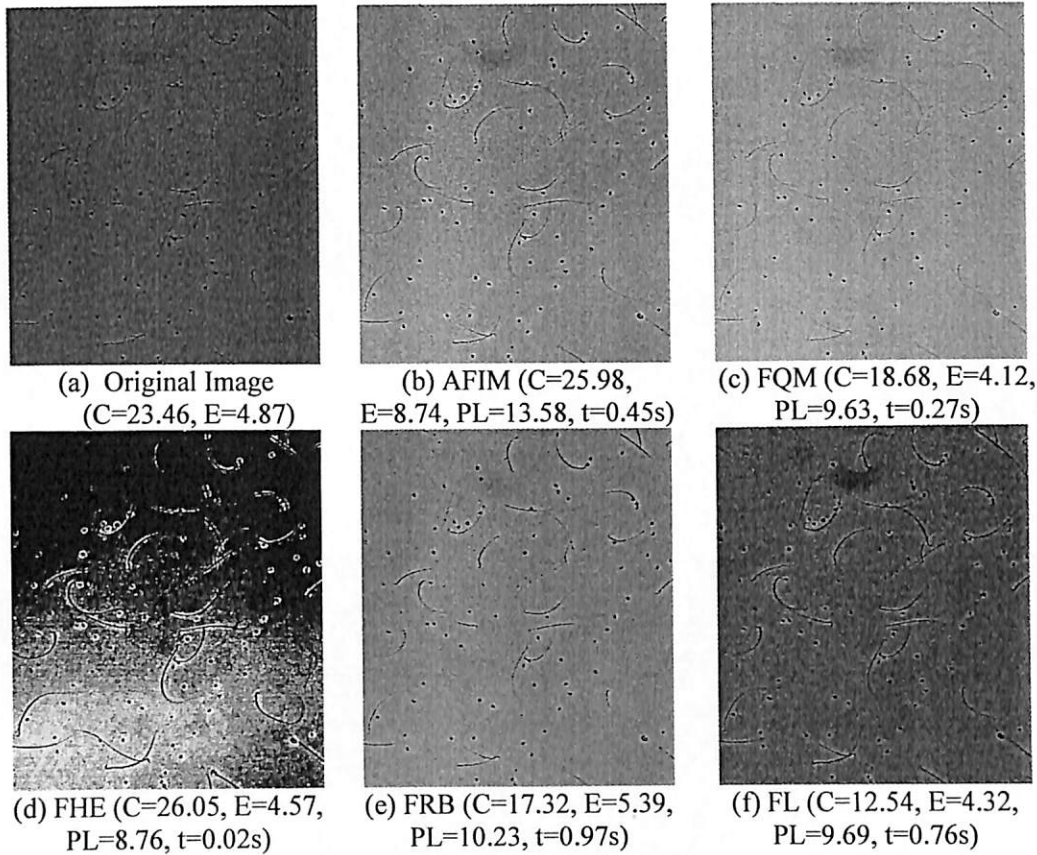
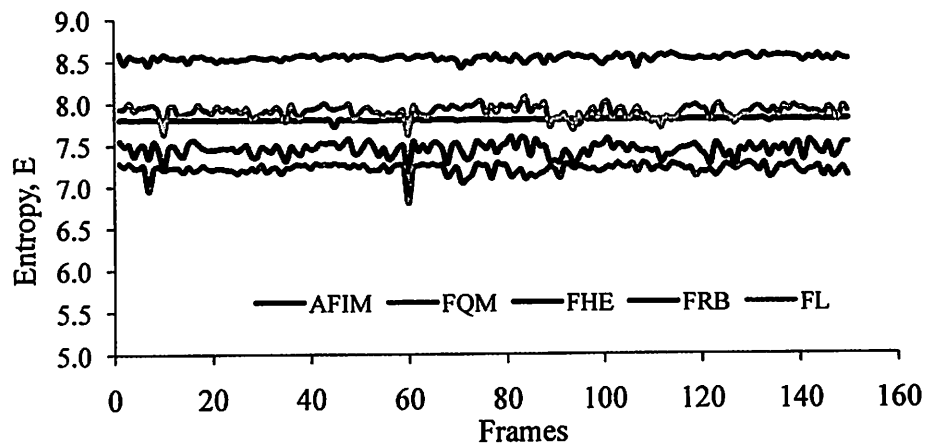


Figure 5 Comparison analysis of *sample 2* sperm image

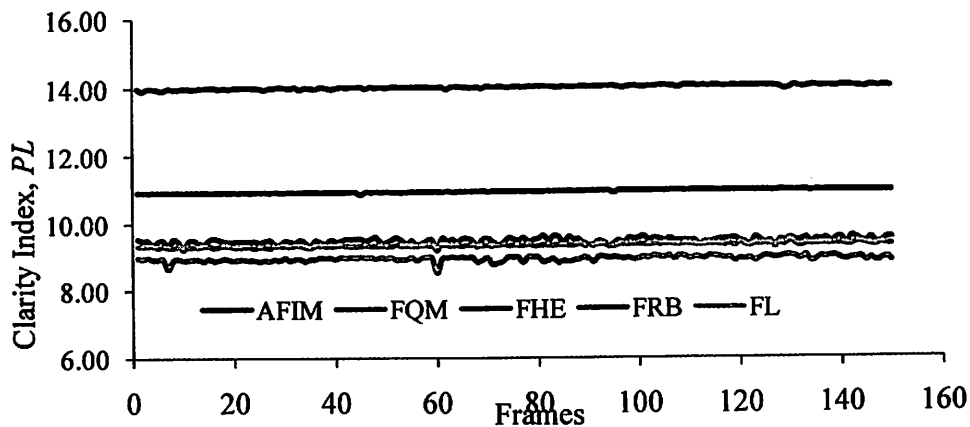
The over-saturated cases are depicted in Figures 5 (d) and (e) where the enhanced images by FHE and FRB seem unnaturally illuminated. The left sides of the images are saturated which causes the details of that region become non-interpretable. The existing noises in both images (*i.e.* FHE and FRB) are amplified thus resulted in lower *PL* of 8.76 and 10.23 respectively. Similar situation with Figure 3, the enhanced image by FL are over-enhanced where the existing bright regions are enhanced. Although the non-uniform illumination issue is handled by FQM, significant sperm features are not enhanced. Therefore, the sperm features are hardly distinguished from their background. The FQM obtains lowest entropy followed by the over-brightened image produced by FL.

Quantitative results of the remaining image sequence frames of *sample 2* are illustrated in Figure 6. The *E* and *PL* analyses shows that the AFIM attains the best results where both are plotted the highest in the Figures 6 (a) and (b). Although the AFIM obtains second highest results in terms of *C*, however the resulted image shown in Figure 5 is

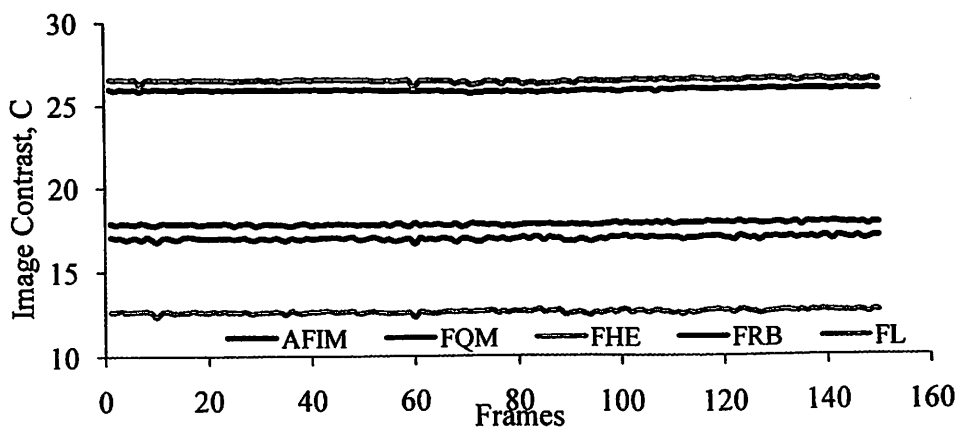
unaffected by the saturation problem. Appropriate enhancement is performed and unnatural appearance of the image is avoided. The saturation problem is significantly affected the details of the image which causes lower E in FHE and FRB (as shown in Figure 5 (d) and (e) as well as plotted graph in Figure 6 (a)).



(a)



(b)



(c)

Figure 6 E , PL and C for 'sample 2' sperm motility video

The comparison analysis is presented in another example of sperm motility video (*i.e. sample 3*) in Figure 7, whereby its supported quantitative results are presented in Figure 8. The AFIM shows the best result where the enhanced image is uniform illuminated and natural-looking image is observed. Though the enhanced image FRB appears well-illuminated, intensity saturation can be observed at the edges of the non-sperm cells. Similar situation can be observed at the edges of the object in Figure 7 (d) where the intensities of that region are over-saturated. This situation could cause problem in segmentation stage since homogeneous region is hardly obtained.

The FL shows the worst result in the qualitative and quantitative analysis presented in Figures 7 and 8 respectively. The FL technique fails to increase the original contrast of the image due to the over-enhancement problem. The over-enhancement problem causes the significant decrement in E and PL which proves that the details of the image are hardly preserved.

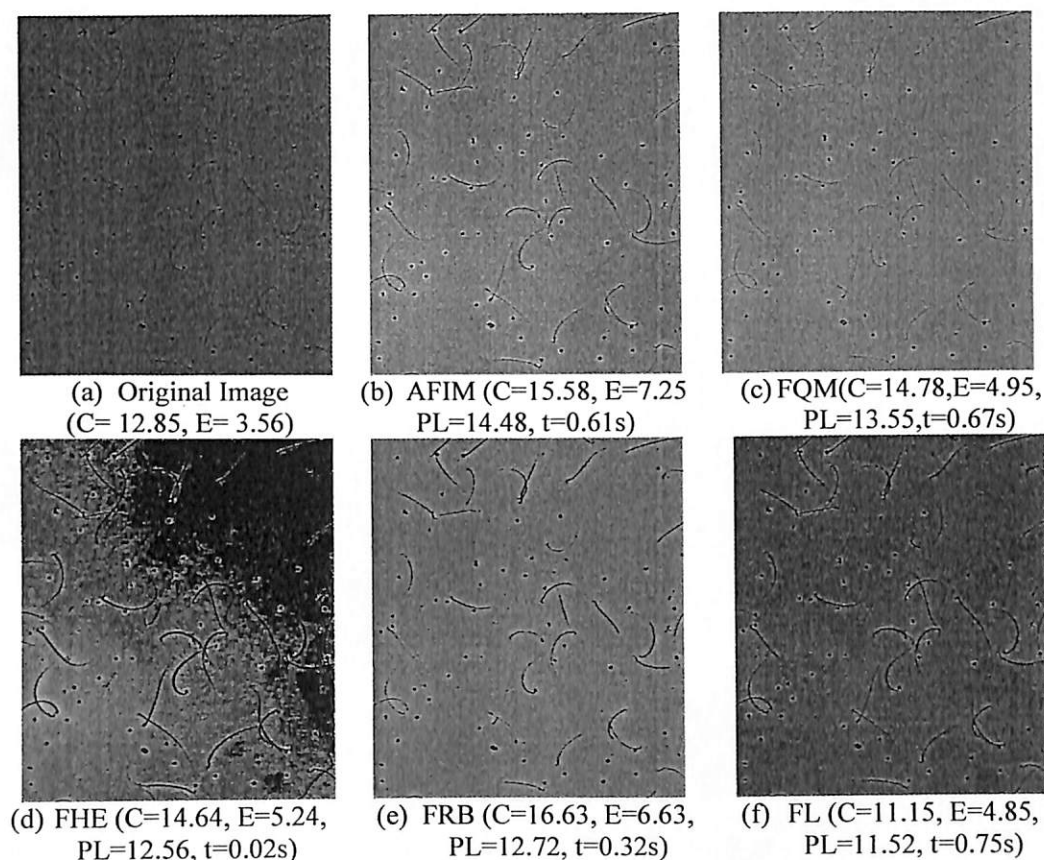


Figure 7 Comparison analysis of *sample 3* sperm image

The average quantitative analyses on 100 sperm motility video are computed and their results are tabulated in Table 1. Each of the sperm motility video consists of 150 image sequence frames where the best results from these analyses are made bold. It is concluded from the table that the AFIM is the most suitable enhancement process to enhance the sperm motility videos since it achieves the best results for C , E and PL . The visual representation of the enhanced image is also improved where the enhanced image sequence frame accomplished more uniform illuminated image and existing noises are kept attenuated.

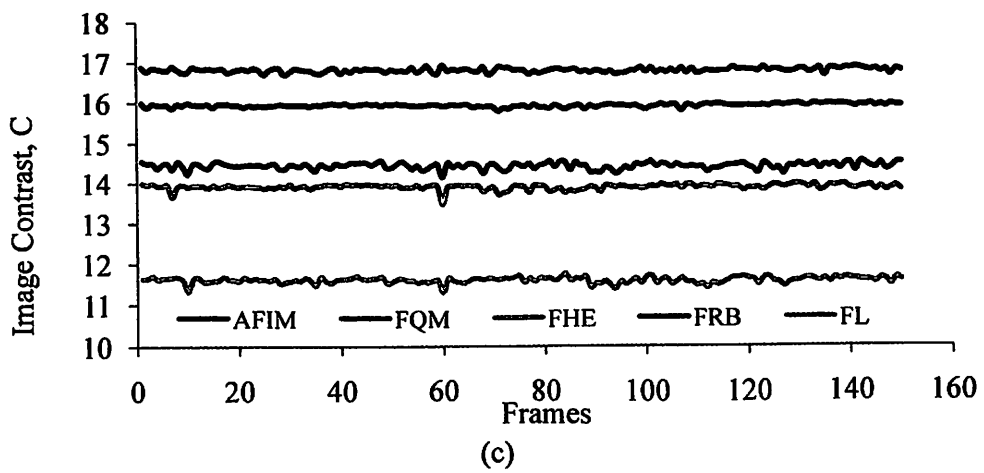
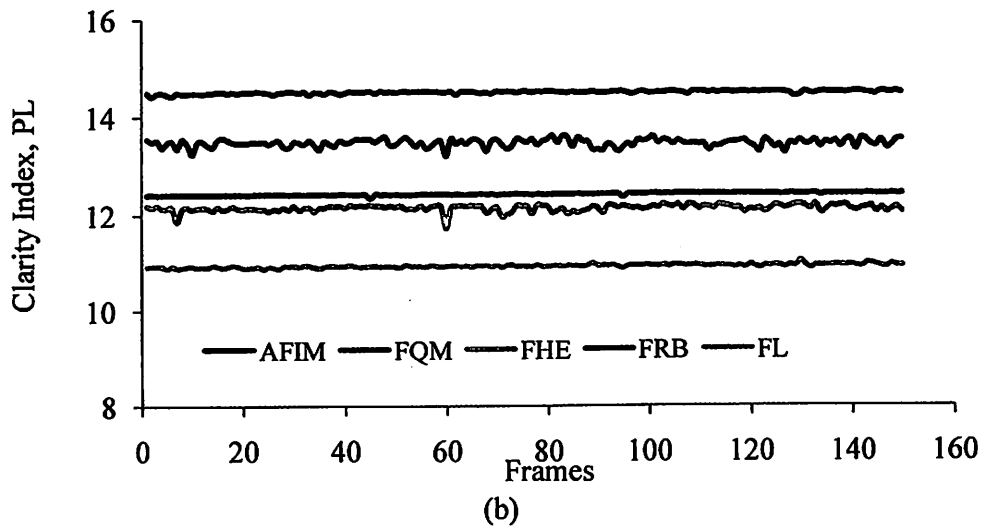
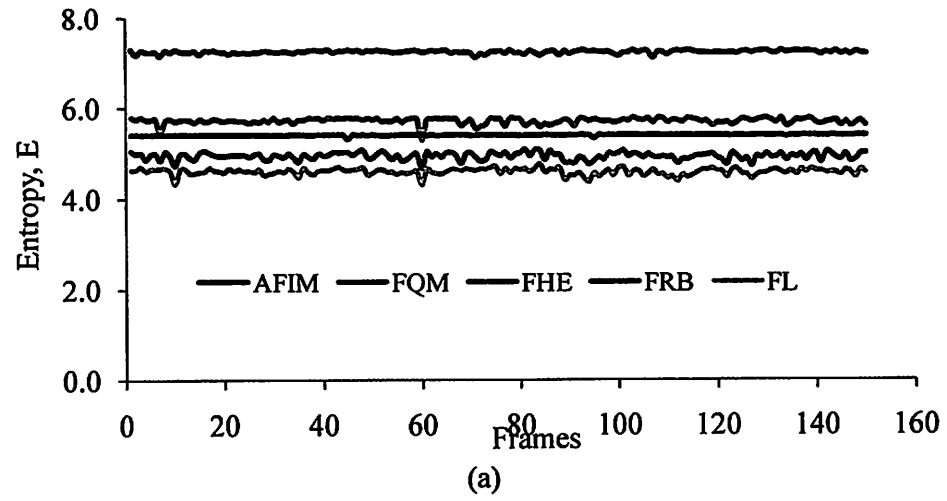


Figure 8 *E*, *PL* and *C* for 'sample 3' sperm motility video

Table 1 Quantitative analyses for the average of the 100 sperm motility videos (*i.e.* each video consist of 150 image sequence frames)

	AFIM	FQM	FHE	FRB	FL
<i>C</i>	16.85±0.03	13.58±0.04	14.63±0.01	15.69±0.02	11.45±0.01
<i>E</i>	7.87±0.12	5.36±0.08	4.62±0.32	6.02±0.35	5.62±0.38
<i>PL</i>	16.32±0.21	15.39±0.05	13.62±0.01	14.32±0.12	10.96±0.03
<i>t (s)</i>	0.73±0.01	0.70±0.10	0.21±0.12	0.64±0.02	0.74±0.08

*the best results are made bold

The higher *PL* verifies that the clarity of the image is improved since its fuzziness degree is decreased as can be supported in equation (22). In other words, the ambiguous pixels that normally occurred at edges and lines of the image are successfully addressed and highlighted to ease further processing task. The enhancement process by the proposed AFIM has proven to provide prominent results in standard and sperm images that could offer an advantage in the segmentation process.

4. Video Processing

4.1.1 Conversion of Image Sequence Frame

The significant features enhanced by AFIM are segmented in this stage. A Multi Stage Local Adaptive Region Based Thresholding (MSLART) technique is proposed to distinguish the objects of interest (*i.e.* sperms) from its background. This process is performed to convert the grayscale image sequence frame to its binary form where white pixels represent the foreground while the black pixels denote as background.

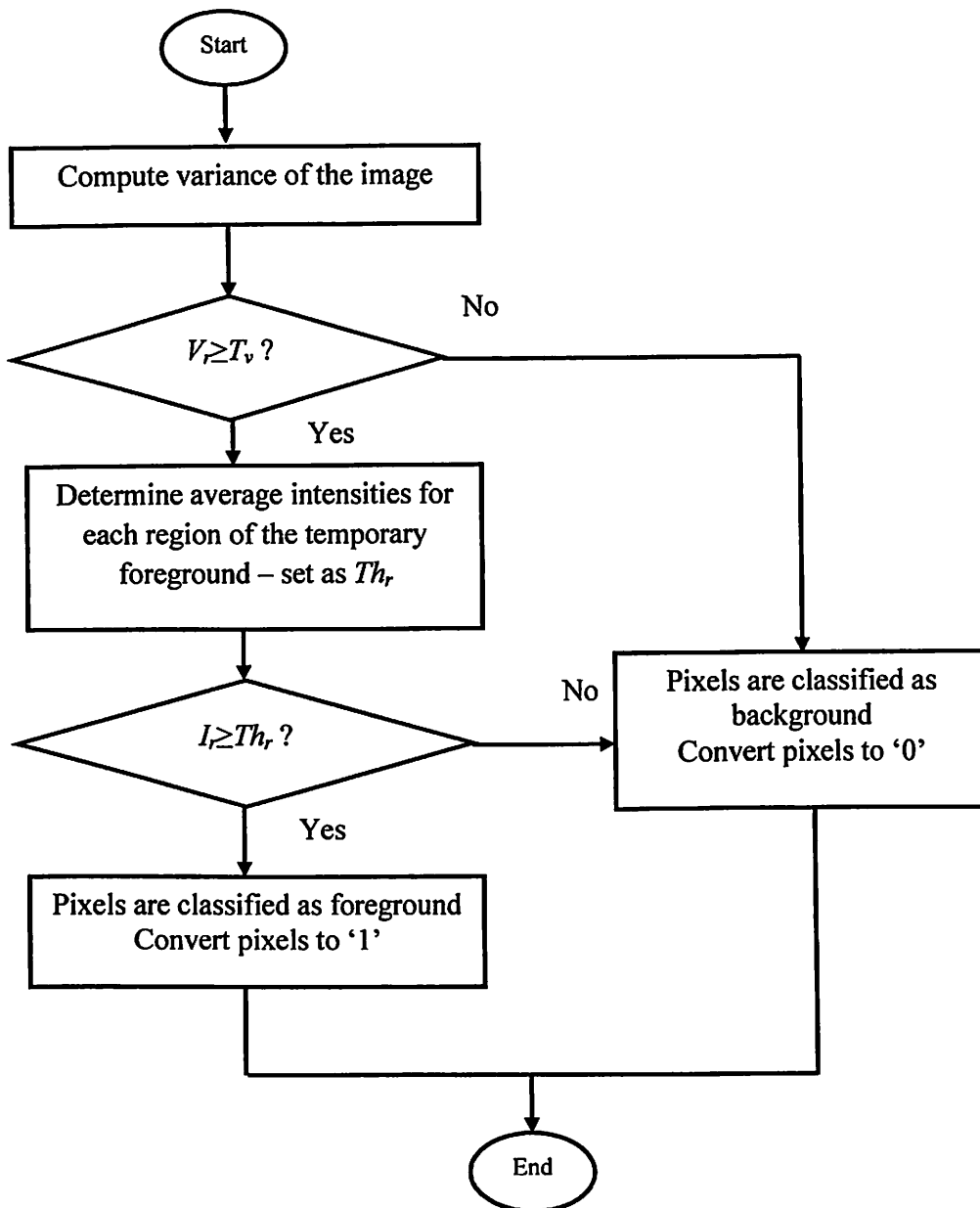
Figure 9 shows the steps involved in the proposed technique, where it consists of multi stages thresholding process to segment the image sequence frame. The sperms are thresholded adaptively by incorporating local statistical parameter (*i.e.* variance) and region intensity of the object.

The process begins by computing the variance image to appraise homogeneity of intensity as given in the equation (29).

$$V_r = \frac{1}{(R \times C) - 1} \sum_{i=1}^{M \in W_M} (M_i - \bar{M})^2 \quad (29)$$

where M_i is the intensity of the enhanced image on the local window and \bar{M} is the average intensity local window.

This process is an initial step to categorize the pixels into two groups of foreground and background. It is conducted locally where the region within the window size is assumed to be uniform. Selection of the window size is made by the observation on the input image. Smaller window size must be chosen if the objects to be segmented are relatively small. Likewise, for the bigger foreground, larger window size must be applied to attain the variance image. Therefore based on the simulation made on 100 sperm motility videos the optimum window size is selected to be 7x7.



Note:

V_r = intensity of the variance image

T_v = variance threshold

Th_r = region threshold

I_r = intensity value of the foreground region

Figure 9 Multi Stage Local Adaptive Region Based Thresholding (MSLART) technique

Examples of variance images are shown in Figure 10. This figure shows the implementation of variance image using the three window sizes of 3x3, 7x7 and 11x11. The variance image obtained from 3x3 window resulted in non-homogeneous segmented region of the sperm. Meanwhile the 11x11 window size has caused dilated object thus resulted the segmented regions appear unnatural. The best result is obtained by the window size 7x7 where the sperm region in the variance image is more homogeneous where the connectivity

of the sperm region can be seen. The sperm region appears naturally segmented compared to the sperm regions produced by window sizes 3x3 and 11x11 respectively.

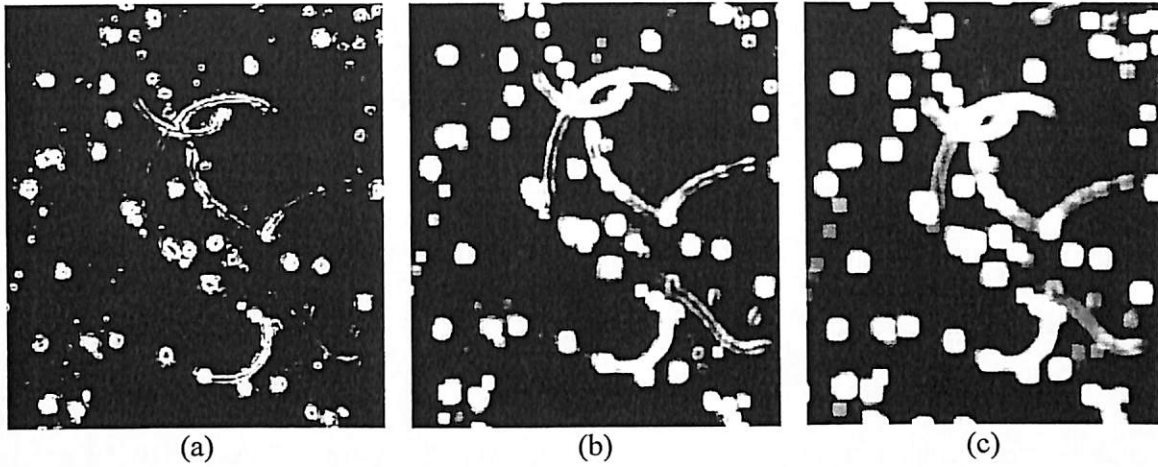


Figure 10 Simulated results of variance image using three different window size (a) 3x3 (b) 7x7 (c) 11x11

When the variance image is obtained, a 'temporary state' of binary image is generated using equation(30). All pixels with intensity values that less than the predetermined variance threshold are considered as unambiguous pixels and denoted as background where their intensity values are converted to pixel '0'. However, if the intensity values of the pixels are more than the T_v , their intensity values are remain unchanged at this stage. This is because these intensity values are the uncertainty pixels that normally exist in the edge of the sperm region:

$$b_t = \begin{cases} V_r & \text{if } V_r \geq T_v \text{ - (i.e. ambiguous pixels)} \\ 0 & \text{if } V_r < T_v \text{ - (i.e. unambiguous pixels)} \end{cases} \quad (30)$$

where b_t is the temporary binary image, V_r is the intensity of the variance image and T_v is the variance threshold.

The variance threshold is selected based on the intensity distribution of the variance image and presented in equation (31). The variance constant, γ is chosen to ensure that T_v lies within the lower range of the intensity value (i.e. normally in the range of 10% to 15% of the intensity variation). This range is chosen to filter out the unambiguous pixels that belong to the background and to ensure only ambiguous pixels are processed in the second stage of MSLART.

$$T_v = \gamma [\max(V_r) - \min(V_r)] \quad (31)$$

In the second stage, a new threshold is set to accurately segment the foreground region. The threshold is determined by computing an average intensity value in each foreground's region:

$$Th_r = \frac{\sum(I_{r1}, I_{r2}, I_{r3}, \dots, I_{rn})}{n} \quad (32)$$

where I_r is the intensity value in the foreground region and n is the number of pixels in each region.

Thus, the uncertainty pixels (*i.e.* from the first stage) are processed in this stage based on the threshold value computed for each temporary foreground region. The pixels in the region are considered as object if their intensity values are more than the region threshold, Th_r . In the meantime, pixels in the region with lower intensity values than Th_r are transformed to black pixel and considered as background.

$$b_r = \begin{cases} 1 & \text{if } I_r \geq Th_r \text{ - (i.e. foreground)} \\ 0 & \text{if } I_r < Th_r \text{ - (i.e. background)} \end{cases} \quad (33)$$

The multi stage approach of segmenting the image sequence frame has ensured that the foreground is properly segmented and distinguished from its background. This process is crucial since the detection of motile sperm is highly dependent on the segmented sperm features. In addition, this approach has provided an advantage of producing homogeneous segmented sperm region by proposing a region based segmentation procedure.

The performance of the segmented image is evaluated quantitatively by using a region non-uniformity (NU) measure. This NU is chosen since it does not require ground-truth image thus reduces the complication of the quantitative analysis. The NU is calculated using:

$$NU = \frac{|F_r| \sigma_f^2}{|F_r + B_r| \sigma^2} \quad (34)$$

where σ^2 represents the variance of the whole image, and σ_f^2 represents the foreground variance. For a well-segmented image, it is expected that NU will be closed to '0' while the worst case of NU will be closed to '1' corresponds to the segmented image where the background and foreground are indistinguishable.

The segmented image in this stage is then processed in the next stage where the unwanted non-sperm cells are removed. This process is elaborated in details in the next subsection.

4.1.2 Results of Thresholding Process

The performance of the thresholding techniques are also investigated for sperm images. The enhanced image sequence frames obtained from pre-processing stage are segmented to distinguish the foreground from its background. The thresholding process is crucial since it will determine the detection of the sperm. The results of the thresholding process are presented in Figures 11 to 13.

Figure 11 (a) shows the enhanced image as reproduce from the Figure 3(b) where the segmented images by MSLART, Otsu, MiniMax and GR are presented in Figures 11 (b), (c), (d) and (e) respectively. The MiniMax produces the worst segmented images where the sperms are hardly distinguished from the background. The MiniMax is too sensitive where the generated energy function causes the foreground misinterpretation when the background is considered as the foreground and converted to white pixels. Meanwhile the segmented images produced by GR and Otsu are homogeneous where the edges of the sperms are smooth. However since the Otsu only uses single threshold value to segment the image, certain regions of the image are not properly segmented.

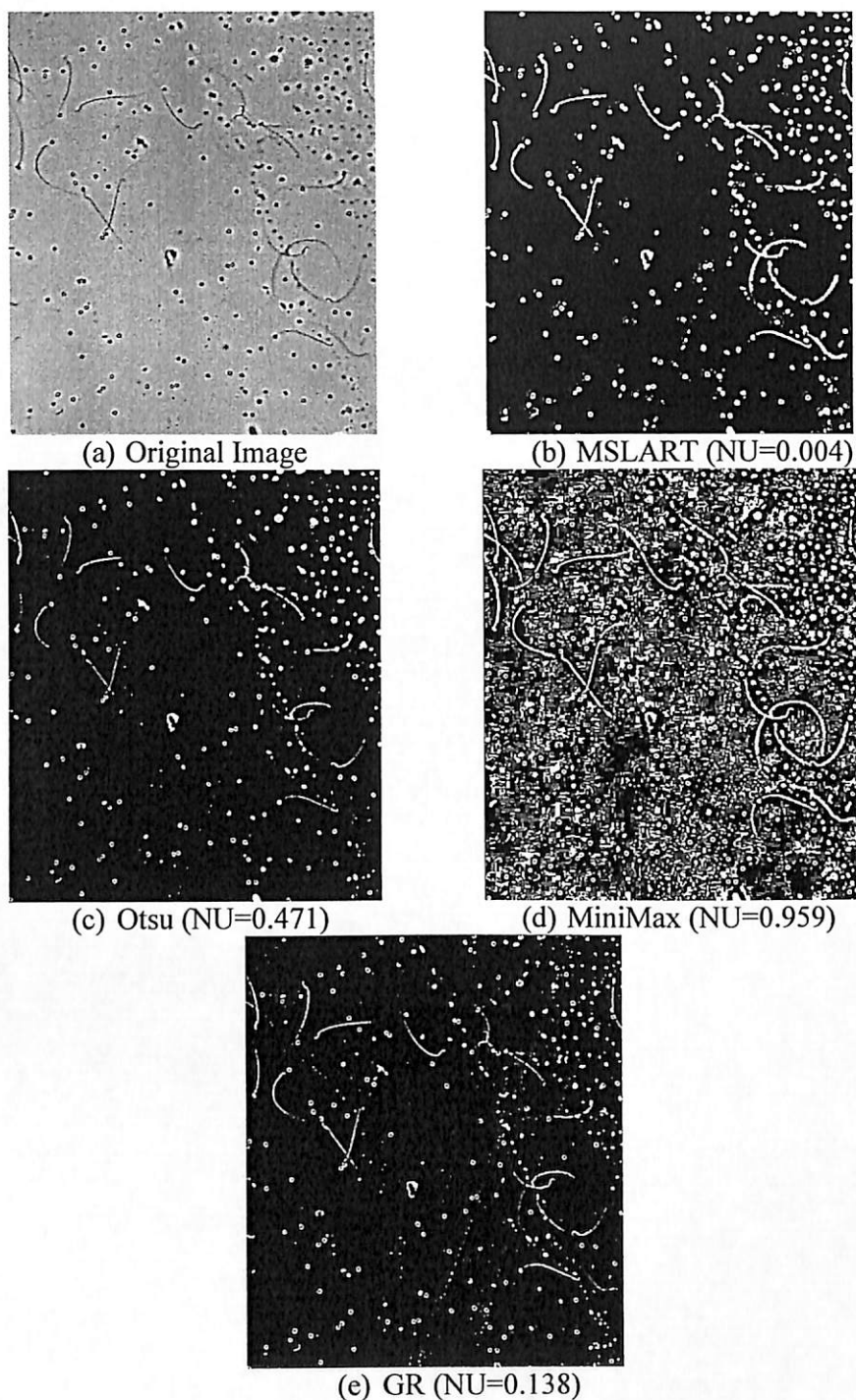


Figure 11 Comparison of thresholding techniques on 'sample 1' image

It can be concluded from the figures presented in Figures 11 to 13 that the MiniMax technique is not suitable to be used to segment the sperm images. This is because the energy function of the MiniMax is too sensitive to noise and intensity deviation in the local window. Therefore, the significant features are hardly extracted from these segmented images. Problems with these noisy images are increased when frame differencing is performed. The residual pixels that normally left after the frame differencing process are now multiplied with the noises from segmented images by MiniMax. Although the Otsu's method is successfully

segmented the image in *sample 1*, this method fails to segment *sample 2* and *sample 3* images. The Otsu's method hardly finds the optimum threshold to adequately segment the images without causing any distortion or unnatural appearance of the segmented images. Some regions in the background are wrongly considered as the foreground as shown in Figure 12 (c) and Figure 13(c) which resulted in second highest NU of 0.445 and 0.392 respectively (the best result of segmented image will produce the lowest NU).

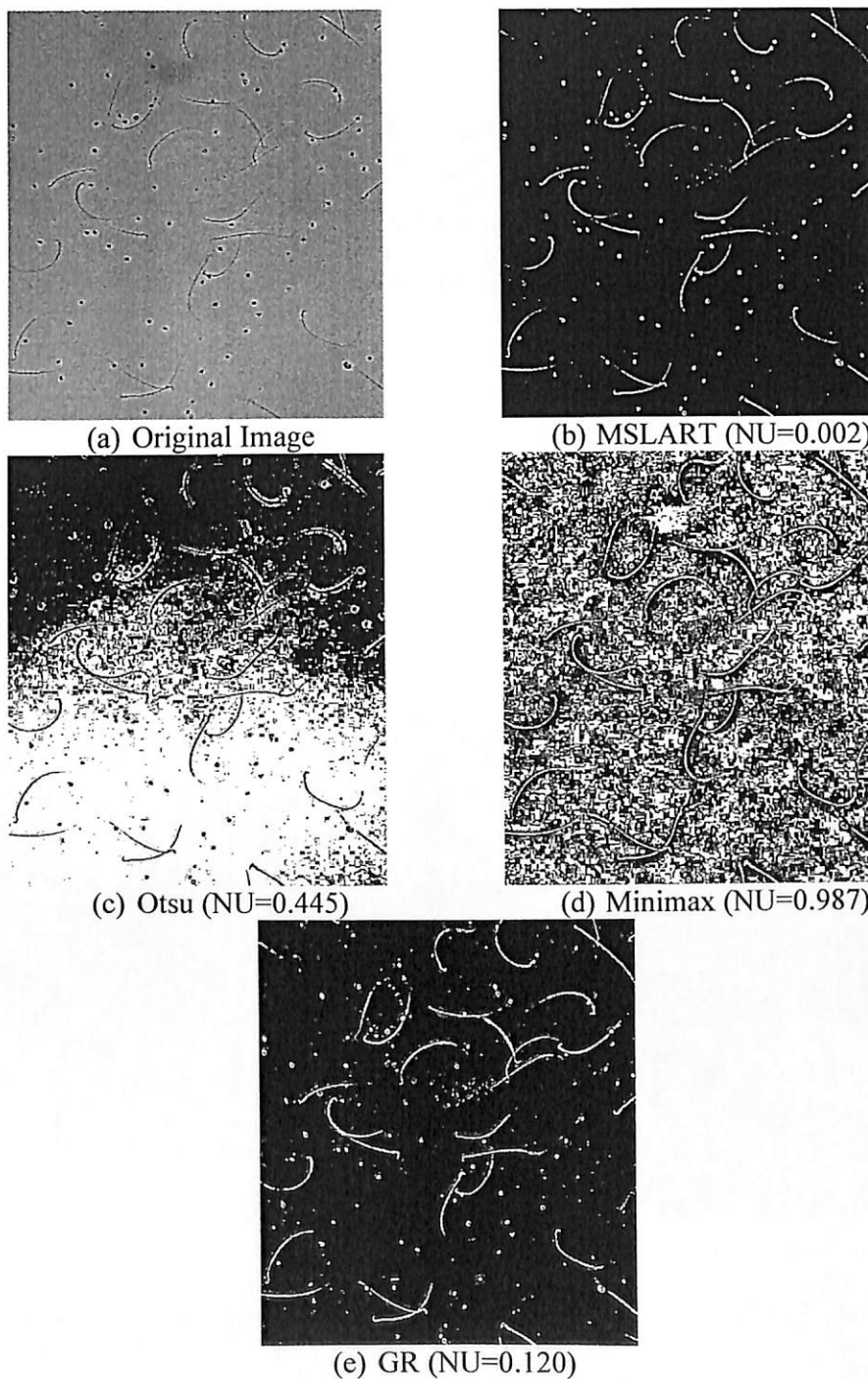


Figure 12 Comparison of thresholding techniques on *sample 2* image

The segmented images produced by GR also show satisfactory results where the sperm and non-sperm cells are well-segmented. Edges of the object are smooth where the foreground can be easily distinguished from the background. The segmented images by GR attain second best results in terms of their qualitative representation for all these three sperm images as well as average NU value for 100 sperm motility videos as presented in Table 2. The MiniMax produces the worst segmented image either in qualitative or quantitative results with the average NU of 0.932. This value is the closest to '1' which means that the foreground and background are indistinguishable.

The adaptive local thresholding approach in the proposed MSLART and GR has ensured that significant details of the image are segmented and thus make them easily to be extracted in the next stage of processing. However, for relatively small object, the GR might introduce problem in obtaining homogeneous segmented region as can be seen in Figure 13 (e).

MSLART attains the best results with smallest NU which indicates that the proposed MSLART technique has successfully obtained homogeneous segmented regions. This is because the MSLART process the unambiguous pixels (normally located at the edge of the sperm's region) separately where a new threshold is set based on the average intensity of the region with unambiguous pixels. By implementing this approach, smooth edges of the sperm region could be obtained. The segmented region of objects of interest (*i.e.* sperms) from MSLART technique are easy to distinguish from their background which is proven by its lowest average NU computed for 100 sperm motility videos.

The local adaptive thresholding provide an advantage over the global approach of Otsu even with the bad illuminated image frame as presented in Figure 12. The smooth segmented regions are important since it will help in the detection process as their results will be presented in the next Section.

Table 2 Average quantitative analysis of NU for 100 sperm motility videos (*i.e.* 150 frames per video)

Techniques	NU
MSLART	0.045
Otsu [6]	0.324
MiniMax[7]	0.932
GR [8]	0.142

4.2 Debris Removal Process

Non-sperm cells that are mostly occupied most of the regions in the image sequence frames are removed in this stage. This process is conducted to ensure that the proposed system only detects all objects that assemble sperm shape. Figure 14 depicts the example of non-sperm cells also known as debris. As shown in this figure, the shape of debris and sperms are apparently different from each other. The debris assembles circular shape while sperms are more likely in cylindrical shape.

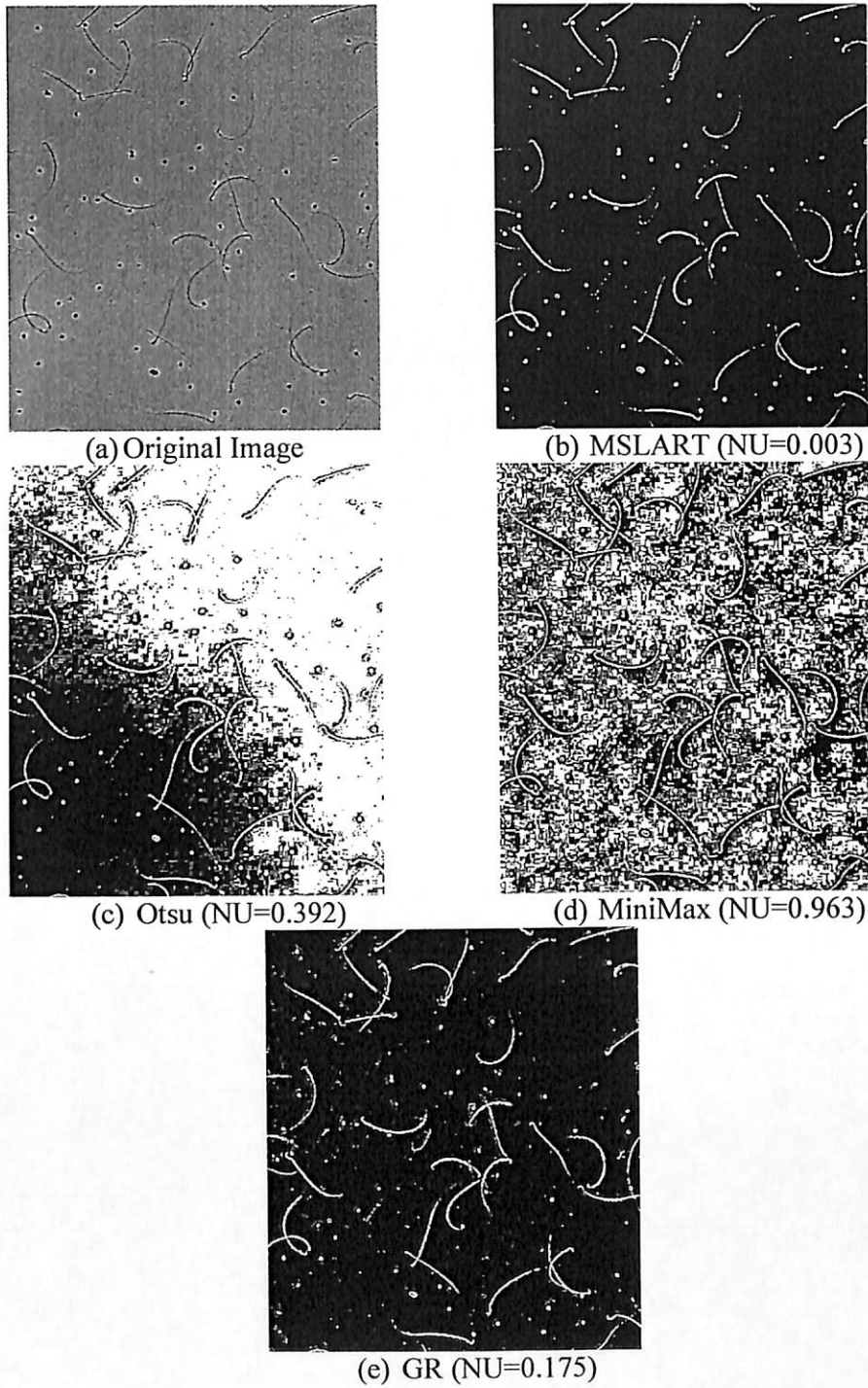


Figure 13 Comparison of thresholding techniques on *sample 3* image

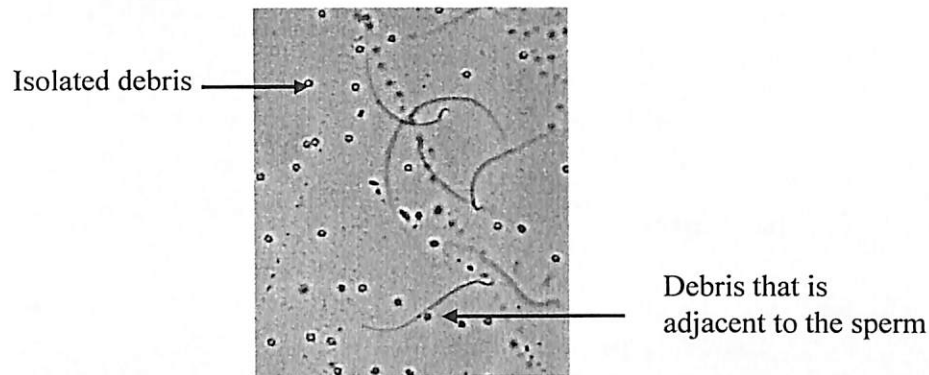


Figure 14 Example of extracted image sequence frame from sperm motility video where isolated debris and debris that is adjacent to the sperm can be observed

Therefore, to discriminate the sperm and the debris, coordinates of the center of all objects are determined. The center of the object denoted as (x_c, y_c) is specified by computing central moment of the object [9]:

$$x_c = \frac{n_{1,0}}{n_{0,0}} \quad (35)$$

$$y_c = \frac{n_{0,1}}{n_{0,0}} \quad (36)$$

where $n_{0,0}$ is the zero order moment which describes the area of the object and $n_{1,0}$ and $n_{0,1}$ are specified in equations (37) and (38) respectively.

$$n_{1,0} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx dy x b(x, y) \quad (37)$$

$$n_{0,1} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx dy y b(x, y) \quad (38)$$

where $b(x, y)$ is the binary image obtained from the image sequence conversion process as presented in previous subsection.

Then, circularity or roundness of the objects is calculated by measuring two parameters namely object's area and perimeter [10] as shown in equation (39). The object's area and perimeter are derived using equations (40) and (41) respectively where radius and angle of the object are measured with respect to the pre-determined center coordinate in equations (35) and (36).

$$circularity = \frac{4\pi A}{P^2} \quad (39)$$

where

$$A = \int_0^{2\pi} \frac{1}{2} r^2 d\phi = \pi \sum_{-\infty}^{\infty} |r_z|^2 \quad (40)$$

$$P = \int_0^{2\pi} \left[(r(\phi)d\phi)^2 + (dr(\phi))^2 \right]^{1/2} = \int_0^{2\pi} d\phi \left[r^2 + \left(\frac{dr}{d\phi} \right)^2 \right]^{1/2} \quad (41)$$

where A is the area of the object, r is the radius of the object, ϕ is the angle of object with respect to its centre, and P is the perimeter of the object.

Figure 15 shows the example of the angle and radius that are extracted from the debris and sperm with respect to their centres. The object is categorized as debris if the circularity is closed to '1' while the object is assumed to be sperm if the object has circularity that is not closed to '1'.

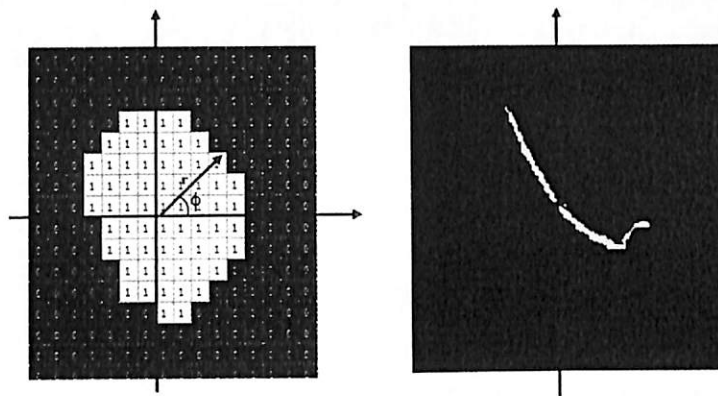


Figure 15 Center, radius and angle of the object in image sequence frame

For all objects that are classified as sperm, pixels belong to these objects are left untouched. However, if the object is classified as debris, coordinate of the object's centre is recorded and coordinates of the mask border are specified. A mask size of 7x7 (Figure 16) is used to indicate a border that determines which region needs to be processed. The 7x7 mask size is chosen in this process to represent the area that needs to be processed since the shape of the debris is very small. In addition, selection of bigger size window might interrupt the adjacent pixels that belong to the sperm's object (as shown in Figure 17). Thus, smallest window size is chosen to remove the unwanted debris while adjacent sperm is kept unchanged for further processing.

The coordinates of the border; top left, top right, bottom right and bottom left are identified as shown in Figure 15. White pixels that lie within this border are converted to black pixels while existing black pixels in this region (if any) are kept unchanged. The residue pixels as shown in Figure 17 are removed after AND Operator which will be explained in the next subsection. The debris removal processed is summarised and presented in the flow chart presented in Figure 18.

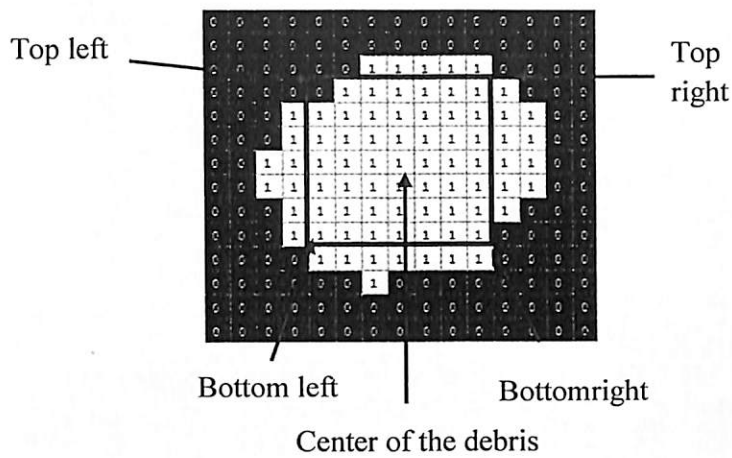
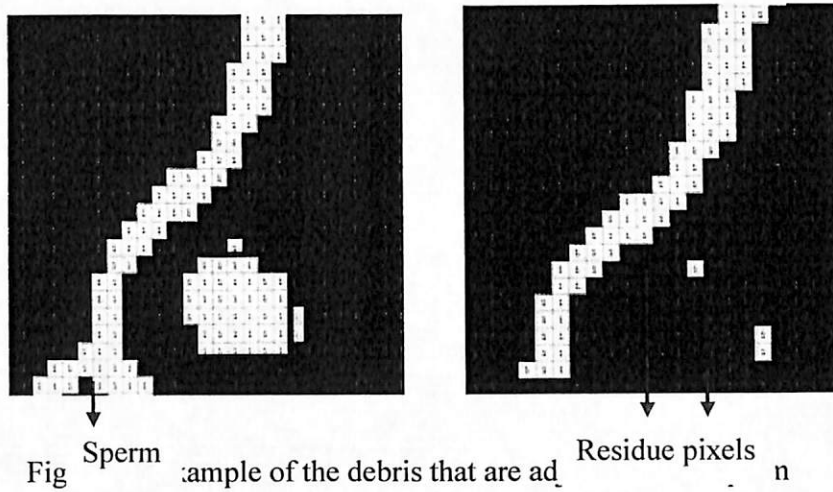


Figure 16 Mask size of 7x7 that is utilized to remove the debris. Pixels in the debris region will be eliminated



4.3 Frame Differencing

The motion features are extracted using frame differencing process that is computed as in equation (42). The frame differencing is calculated by subtracting the current frame with its previous frame to attain the motion features. As given in this equation, if $D(t)$ is equal to '1', the corresponding region is defined as the moving region. Otherwise, the region is considered as background. T in this case is chosen based on the median value of the difference image between current and next frames.

$$D(t) = \begin{cases} 1 & \text{if } \|F(t-1) - F(t)\| \geq T \\ 0 & \text{otherwise} \end{cases} \quad (42)$$

where $D(t)$ represents the existence of difference between background and the input image frame. Meanwhile $F(t-1)$ is the previous frame and $F(t)$ is the current frame. The example of frame differencing implementation is depicted in Figure 18.

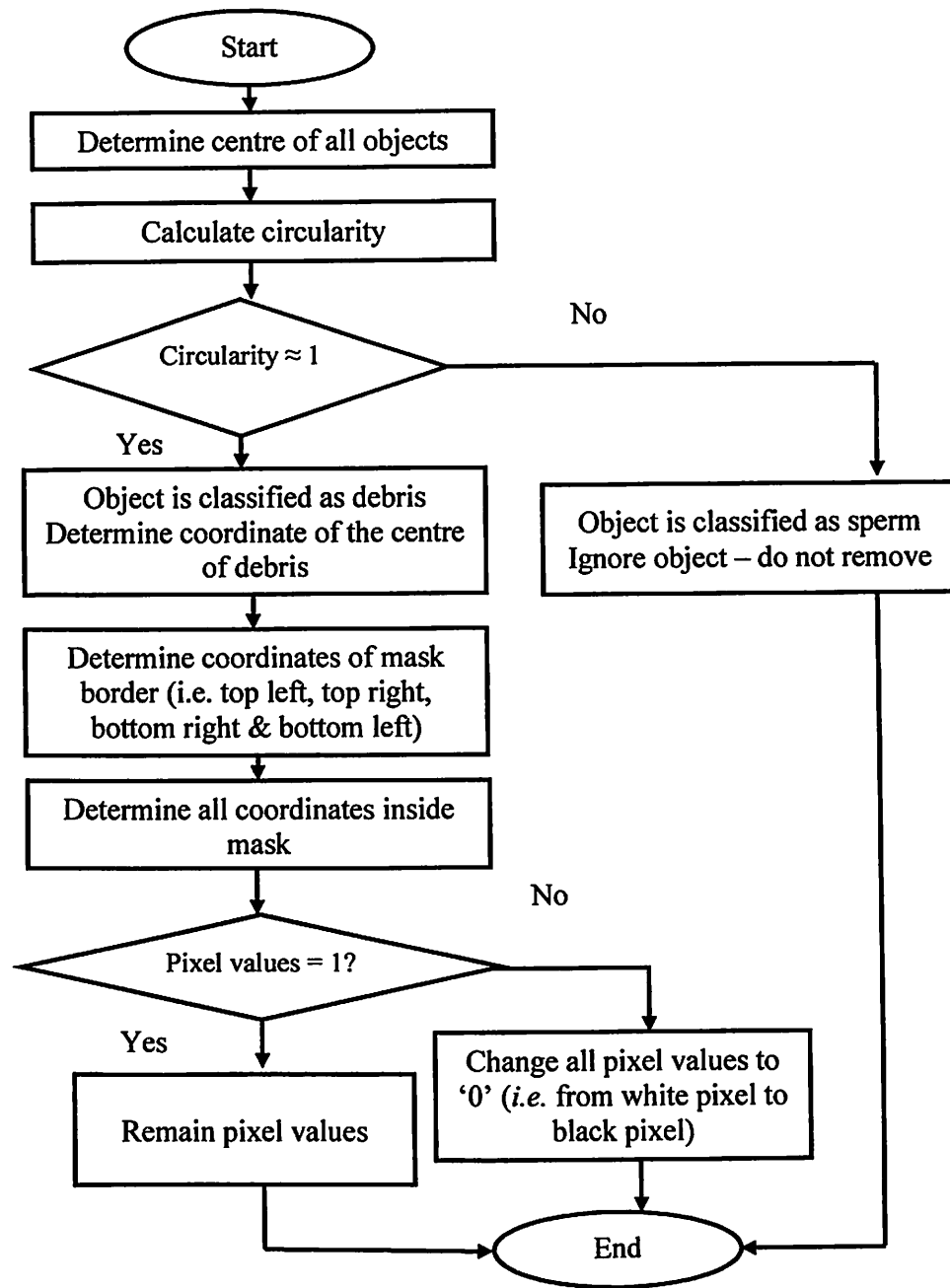


Figure 18 Flow chart of debris removal process

Since the simple frame differencing method often fails because of misidentification of the moving object [11], it is necessary to incorporate sperm features with its motion features to provide initial tracker of the moving sperms. Results obtained from debris removal and frame differencing processes are merged using AND Operator.

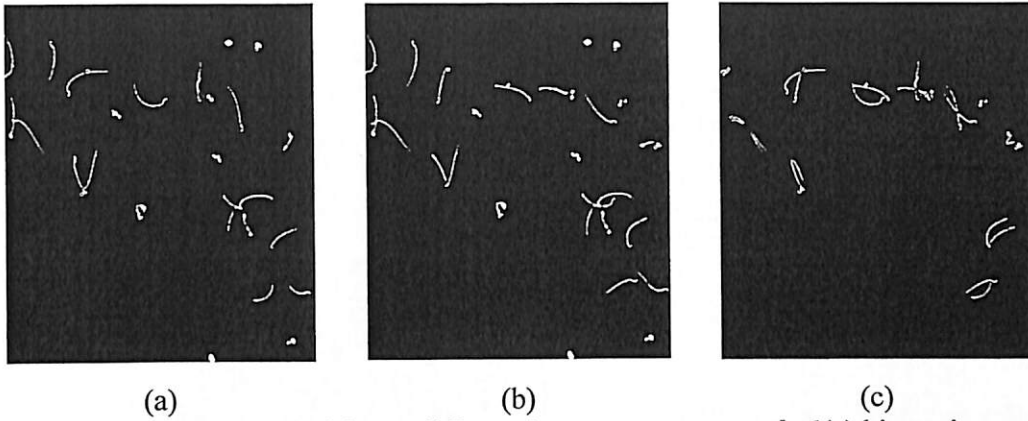


Figure 18 Implementation of frame differencing process on *sample 1* (a) binary image of frame 2 (b) binary image of frame 3 (c) resulted image from frame differencing process

The residual pixels that are left on the merged image frame are removed using morphology filtering. The binary image attained from the merging process of AND Operator is filtered at a scale defined by the size of structuring element. Regions that fit the structuring element are passed by the filter meanwhile smaller structures are blocked and excluded from the output image as depicted in Figure 19.

In this study, structure element of size 2×2 is utilized to remove the residue pixels on the binary image obtained from the AND Operator. Sperm with region size bigger than the structure element is left untouched. By implementing the proposed system, the debris removal process in addition to the frame differencing process, the segmentation errors that could occur after frame differencing are eliminated. The motile sperms are successfully extracted and false detection could be avoided.

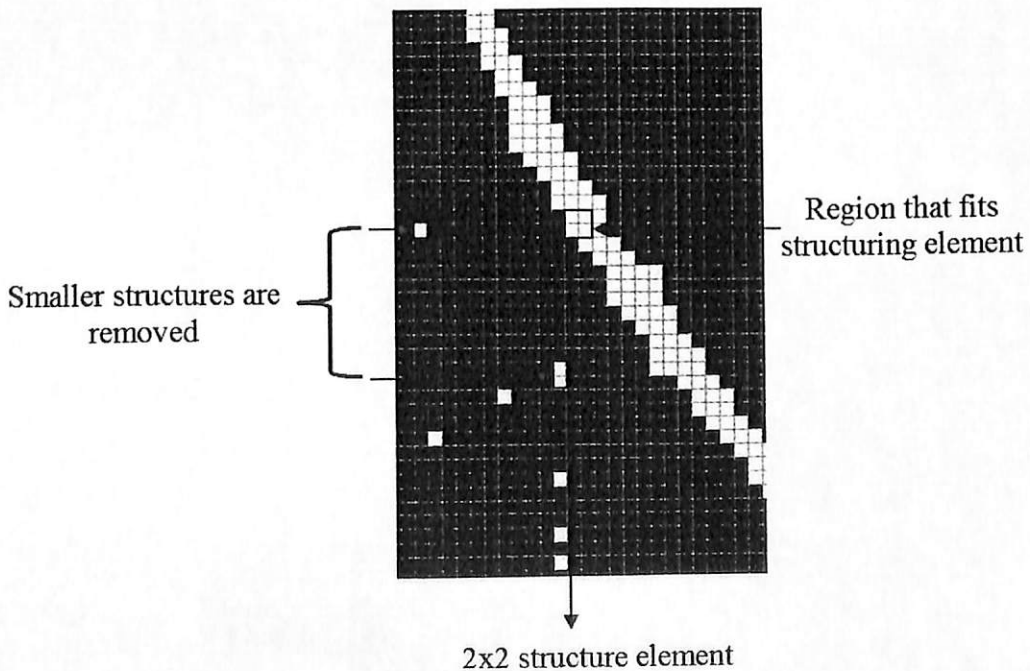


Figure 19 Morphological operation to remove residue pixels

4.4 Analysis of Disconnected Objects

4.4.1 Initial Tracker of Motile Sperm

After the motion features and sperm shape are successfully extracted, initial trackers of the sperm are determined. The center coordinates of the objects in the cleaned binary image are extracted using aforementioned equations (35) and (36). The implementation of the overall sperm feature detection is simplified in Figure 20.

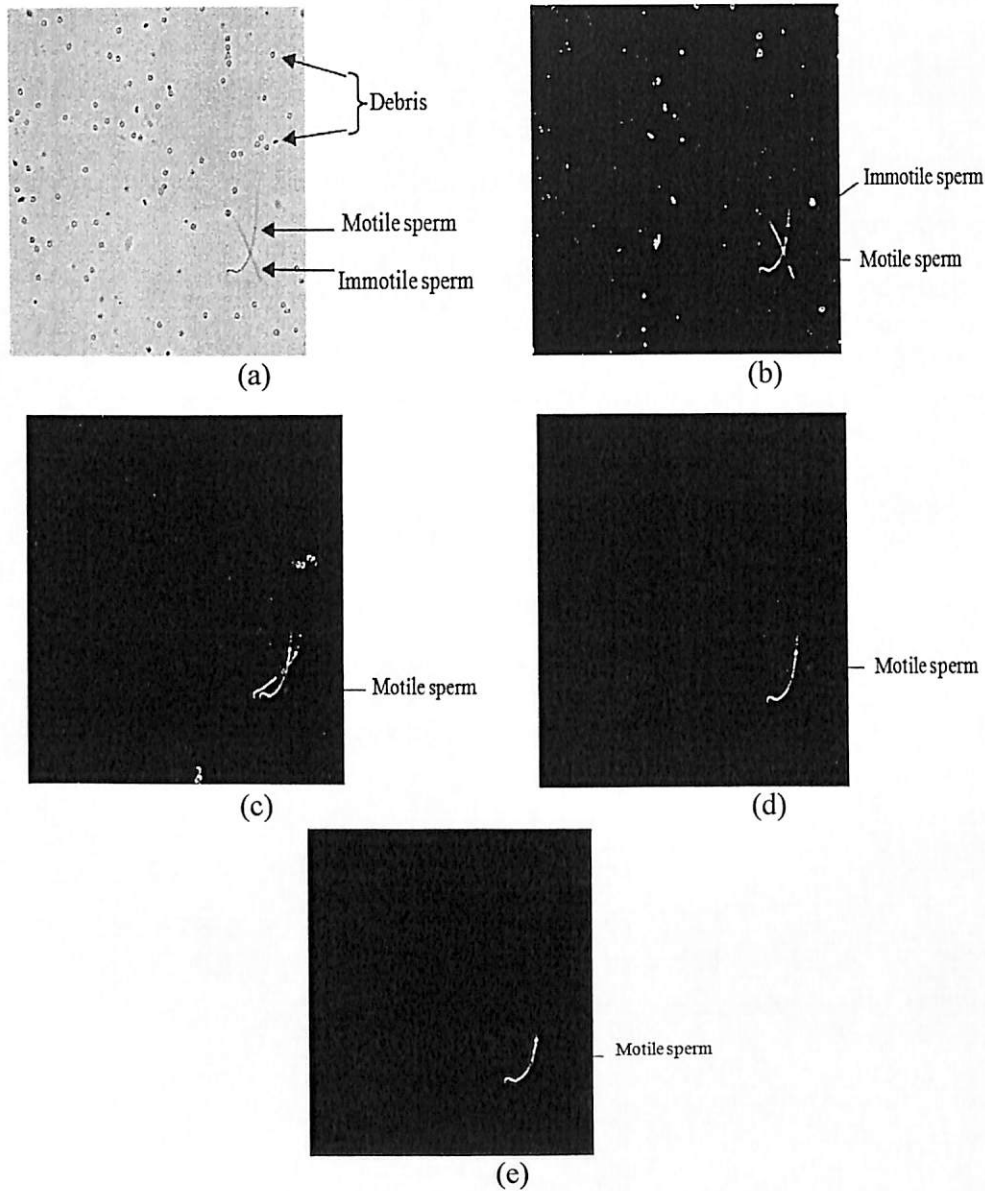


Figure 20 Implementation of the sperm feature detection process (a) enhanced image frame (b) processed image sequence frame from debris removal process (c) processed image sequence frame from frame differencing (d) processed image from AND Operator (e) cleaned binary image sequence frame with initial tracker of motile sperm

Although motion features and sperm shape are successfully extracted, the segmented object tends to be disconnected. This is because some motile sperms are moving too fast and

thus caused the visibility of these motile sperms to be reduced(Figure 21(a)). As a result, the segmented object in binary image (Figure 21 (b)) is disconnected and the object is identified as two different objects which could cause false detection.

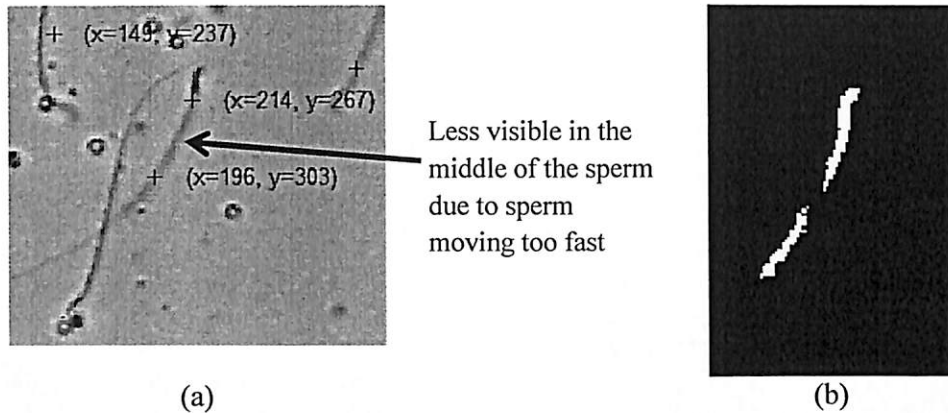


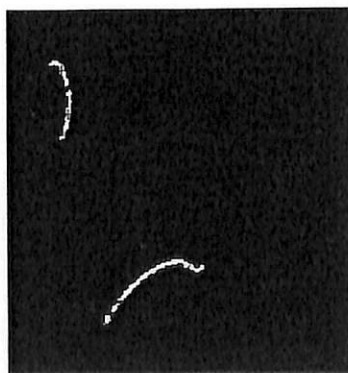
Figure 21 Example of disconnected object that occurs due to the fast moving sperm in (a) grayscale image (b) binary image

Thus, the center coordinates that are determined in this stage will act as the initial trackers of the motile sperms. The disconnected object needs to be attached to their corresponding object that possesses similar orientation. As an example, if the object is in horizontal orientation, the object needs to be attached to its corresponding object that also has similar orientation (*i.e.* horizontal in this case). Since most sperms are adjacent to each other and moving in uncontrolled environment, there might be more than one possibility of coordinate combinations between two disconnected objects.

Figure 22 shows an example of four disconnected objects where the initial tracker of motile sperms has detected four center coordinates. These initial coordinates are utilized to determine number of possible coordinate combinations between disconnected objects using equation (43). As depicted in this figure since k is equivalent to 4, therefore there are six possible coordinate combinations for the disconnected objects as tabulated in Table 3.

$$\text{No. of possible coordinate combinations} = \frac{k(k-1)}{2} \quad (43)$$

$k = \text{initial total number of detected sperms}$



$$\text{No. of possible coordinate combinations} = \frac{k(k-1)}{2}$$

since $k = 4$

\therefore No. of possible coordinate combinations = 6

Figure 22 Example of possible combinations of disconnected object from initial tracker of motile sperm

Table 3 Possible coordinate combination for Figure 22

Possible Coordinate Combination	Coordinates of Candidate 1	Coordinates of Candidate 2	Δx_i	Δy_i	$\Delta x_i + \Delta y_i$
1 [S1-S2]	(261,145)	(265,173)	$261-265 =4$	$145-173 =28$	32
2 [S1-S3]	(261,145)	(321,278)	$ 261-321 =60$	$ 145-278 =133$	193
3 [S1-S4]	(261,145)	(291,305)	$261-291 =30$	$145-305 =160$	190
4 [S2-S3]	(265,173)	(321,278)	$ 265-321 =56$	$ 173-278 =105$	161
5 [S2-S4]	(265,173)	(291,305)	$265-291 =26$	$173-305 =132$	158
6 [S3-S4]	(321,278)	(291,305)	$ 321-291 =30$	$ 278-305 =27$	57

*The significant coordinate combinations are made bold

The extracted coordinate combinations include combination between objects S1 and S2, S1 and S3, S1 and S4, S2 and S3, S2 and S4, as well as S3 and S4. Based on these possible coordinate combinations, the objects are categorized into two groups of *Candidate 1* and *Candidate 2*. For each possible coordinate combination (i.e. 2 disconnected objects), object with smaller x -coordinate is considered as *Candidate 1* whereby the other disconnected object is denoted as *Candidate 2*. As tabulated in Table 3, for the first coordinate combination, object S1 has smaller x -coordinate than x -coordinate of object S2, therefore, object S1 and S2 are categorized as *Candidate 1* and *Candidate 2* groups respectively. Similar situation can be observed with sixth coordinate combination where x -coordinate of object S3 is smaller than x -coordinate of object S4. Therefore S3 is grouped in *Candidate 1* while S4 is categorized as *Candidate 2*.

Each possible coordinate combination is analysed and only correct combination will be extracted to attach the disconnected objects. The disconnected objects must be connected only to their own combination to avoid any deformation of the sperm shape which will then affect the detection of the motile sperm process. However, analysing each coordinate combination requires high computational time and increases complexity of the detection process. Therefore, only significant coordinate combinations will be considered to be processed in order to connect the disconnected objects. To solve this issue, absolute differences of x -coordinates (Δx_i), absolute differences of y -coordinates (Δy_i) and total absolute difference of coordinates between *Candidate 1* and *Candidate 2* ($\Delta x_i + \Delta y_i$) are calculated for every possible coordinate combination as can be seen in the example tabulated in Table 4.

A threshold of 100 is set to determine significant coordinate combinations to be processed. This threshold value is selected based on the observation made on the distance between objects on image sequence frames extracted from 100 sperm motility videos. The significant coordinate combinations are identified if the combinations have the value of $\Delta x_i + \Delta y_i$ is less than 100. These coordinate combinations consist of the disconnected objects that are closed to each other. They are more likely to be attached to form a sperm shape. Therefore, only coordinate combinations that produced $\Delta x_i + \Delta y_i \leq 100$, are considered to be evaluated.

As depicted in Figure 23 and Table 3, from six possible coordinate combinations, only two combinations namely the first and the sixth coordinate combinations are relevant to be considered as disconnected objects. The combinations are for objects S1 and S2, as well as for objects S3 and S4 with $\Delta x_i + \Delta y_i$ of 34 and 57 respectively. It can be also observed in this table that this case only deals with unique coordinates of *Candidate 1* and *Candidate 2* where the sperms are not adjacent to each other. These coordinates are considered as unique since no repetitive coordinates appear in *Candidate 1* and *Candidate 2* as well as no repetitive coordinates can be seen between combination of *Candidate 1* and *Candidate 2*.

As an example in this table, *Candidate 1* consists of objects S1 (*i.e.* coordinate of (264,145)) and S3 (*i.e.* coordinate of (321,278)) while *Candidate 2* consists of objects S2 (*i.e.* coordinate of (265,173)) and S4 (*i.e.* coordinate of (291,305)). It can be observed from this example there are unique coordinates between intra-group *Candidate 1* and also unique coordinates between intra-group *Candidate 2*. In addition, this example shows that the coordinates are unique between inter-group *Candidate 1* and *Candidate 2*.

However, most of the image sequence frames consist of frames with sperms that are moving closely with other sperms. Thus, their disconnected objects are adjacent and non-unique coordinates of *Candidate 1* and *Candidate 2* can be observed. Second example of Table 4 shows that among three possible coordinate combinations, only the first and second combinations are considered in this analysis. This is because the third combination produces $\Delta x_i + \Delta y_i$ more than 100 and therefore is excluded in the analysis of disconnected objects. Object J (*i.e.* coordinate (194,207)) is considered as a non-unique coordinate between combination of *Candidate 1* and *Candidate 2*. Object J (*i.e.* coordinate (194,207)) belongs to *Candidate 2* in the first combination while in the second combination it is assigned as *Candidate 1*. As it could be candidate for more than one combination, it is said to be non-unique coordinates. This combination is made bold in the first example of Table 4.

There are two types of non-unique coordinate as follow where the first type is belong to the coordinate that could possibly belong to more than one combination but only assigned to one candidate (*i.e.* either *Candidate 1* or *Candidate 2*). Meanwhile second type of non-unique coordinate is referred to the coordinate that could possibly belong to more than one combination and assigned as two candidates (*i.e.* both *Candidate 1* and *Candidate 2*).

For clarification of the aforementioned situations of non-unique coordinate can be referred to the second example of Table 4 (*i.e.* second row). In this example, the third and fourth coordinate combinations are neglected since objects L and O as well as objects L and P are far from each other as both combinations attain $\Delta x_i + \Delta y_i$ more than 100. For the first type of non-unique coordinate, it can be observed that coordinates (311,141) (for object L) are repeated in first and second combinations.

This coordinate is considered as the first type of non-unique coordinate as it is assigned as *Candidate 1* for more than one combinations. The first type of non-unique coordinate can also be observed for object P (*i.e.* coordinate (346,241)) where it is assigned as *Candidate 2* in the seventh, ninth and tenth combinations. The second type of non-unique coordinates can also be observed in this example. Coordinate (326,188) (*i.e.* Object M) appears as *Candidate 2* at the first combination and as *Candidate 1* at the fifth, sixth and seventh combinations. In addition, object N (*i.e.* coordinate (367,170)) also shows the example of the second type of non-unique coordinate where it is assigned as *Candidate 2* for the second and fifth combinations as well as *Candidate 1* for eighth and ninth combinations. Similar situation can be observed in object O (*i.e.* coordinate (380,213)) where it became *Candidate 2* for sixth and eighth combinations while it appears as *Candidate 1* in the tenth combination.

Therefore, each non-unique coordinate needs to be analysed and correct coordinate combination needs to be chosen to solve the disconnected object issues. Four cases of disconnected objects are found in this study and successfully identified using the process presented in Figure 23. Each condition to determine all cases of disconnected objects is presented in Table 5.

Table 4 Unique and non-unique combinations of *Candidate 1* and *Candidate 2*

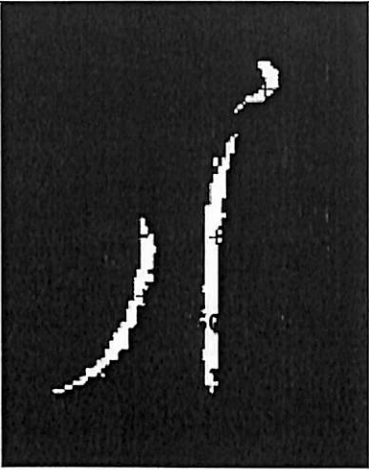
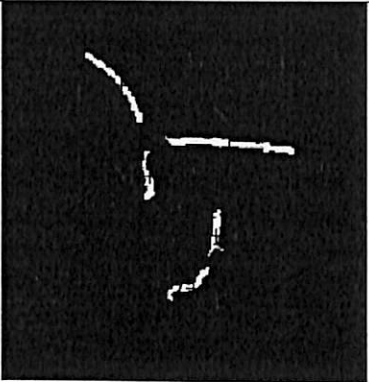
Example	Coordinate Combination																											
	<p>Coordinate Combinations:</p> <ol style="list-style-type: none"> I-J; $\Delta x_i + \Delta y_i = 55$ J-K; $\Delta x_i + \Delta y_i = 37$ I-K; $\Delta x_i + \Delta y_i = 102$ – neglected in analysis since $\Delta x_i + \Delta y_i \geq 100$ <p>Significant Coordinate Combinations to be analyzed:</p> <table border="1" data-bbox="683 637 1305 784"> <thead> <tr> <th>Coordinate Combination</th> <th><i>Candidate 1</i></th> <th><i>Candidate 2</i></th> </tr> </thead> <tbody> <tr> <td>1</td> <td>(205,163)</td> <td>(194,207)</td> </tr> <tr> <td>2</td> <td>(194,207)</td> <td>(160,220)</td> </tr> </tbody> </table> <p>*non-unique coordinate between combination of <i>Candidate 1</i> and <i>Candidate 2</i> is made bold</p>	Coordinate Combination	<i>Candidate 1</i>	<i>Candidate 2</i>	1	(205,163)	(194,207)	2	(194,207)	(160,220)																		
Coordinate Combination	<i>Candidate 1</i>	<i>Candidate 2</i>																										
1	(205,163)	(194,207)																										
2	(194,207)	(160,220)																										
	<p>Coordinate Combinations:</p> <ol style="list-style-type: none"> L-M; $\Delta x_i + \Delta y_i = 62$ L-N; $\Delta x_i + \Delta y_i = 85$ L-O; $\Delta x_i + \Delta y_i = 121$ - neglected in analysis since $\Delta x_i + \Delta y_i \geq 100$ L-P; $\Delta x_i + \Delta y_i = 135$ – neglected in analysis since $\Delta x_i + \Delta y_i \geq 100$ M-N; $\Delta x_i + \Delta y_i = 59$ M-O; $\Delta x_i + \Delta y_i = 59$ M-P; $\Delta x_i + \Delta y_i = 73$ N-O; $\Delta x_i + \Delta y_i = 50$ N-P; $\Delta x_i + \Delta y_i = 92$ O-P; $\Delta x_i + \Delta y_i = 42$ <p>Significant Coordinate Combinations to be analyzed:</p> <table border="1" data-bbox="671 1412 1297 1775"> <thead> <tr> <th>Coordinate Combination</th> <th><i>Candidate 1</i></th> <th><i>Candidate 2</i></th> </tr> </thead> <tbody> <tr> <td>1</td> <td>(311,141)</td> <td>(326,188)</td> </tr> <tr> <td>2</td> <td>(311,141)</td> <td>(367,170)</td> </tr> <tr> <td>5</td> <td>(326,188)</td> <td>(367,170)</td> </tr> <tr> <td>6</td> <td>(326,188)</td> <td>(380,213)</td> </tr> <tr> <td>7</td> <td>(326,188)</td> <td>(346,241)</td> </tr> <tr> <td>8</td> <td>(367,170)</td> <td>(380,213)</td> </tr> <tr> <td>9</td> <td>(367,170)</td> <td>(346,241)</td> </tr> <tr> <td>10</td> <td>(380,213)</td> <td>(346,241)</td> </tr> </tbody> </table> <p>**non-unique coordinates between combination of <i>Candidate 1</i> and <i>Candidate 2</i> are made bold</p>	Coordinate Combination	<i>Candidate 1</i>	<i>Candidate 2</i>	1	(311,141)	(326,188)	2	(311,141)	(367,170)	5	(326,188)	(367,170)	6	(326,188)	(380,213)	7	(326,188)	(346,241)	8	(367,170)	(380,213)	9	(367,170)	(346,241)	10	(380,213)	(346,241)
Coordinate Combination	<i>Candidate 1</i>	<i>Candidate 2</i>																										
1	(311,141)	(326,188)																										
2	(311,141)	(367,170)																										
5	(326,188)	(367,170)																										
6	(326,188)	(380,213)																										
7	(326,188)	(346,241)																										
8	(367,170)	(380,213)																										
9	(367,170)	(346,241)																										
10	(380,213)	(346,241)																										

Table 5 Conditions to be considered to identified disconnected object cases

Condition	Description
1	$\Delta x_i + \Delta y_i \leq 100$
2	i) Unique coordinates of <i>Candidate 1</i> AND ii) Unique coordinates of <i>Candidate 2</i>
3	*Not a type-1 non-unique coordinate Unique coordinates of combination of both <i>Candidates</i>
4	*Not a type-2 non-unique coordinate Total length of <i>Candidate 1</i> and <i>Candidate 2</i> ≥ 100

The first case of disconnected object is defined if the objects produced unique coordinates of *Candidate 1* and *Candidate 2* (i.e. or the objects are not in type-1 of non-unique coordinate) and unique coordinate of combination of both *Candidate 1* and *Candidate 2* (i.e. or the objects are not a type-2 of non-unique coordinate). This case can be represented by the example in Figure 11. Meanwhile, the second case is slightly different from case 1 where the type-2 of non-unique coordinate occurs as shown by the first example of For clarification of the aforementioned situations of non-unique coordinate can be referred to the second example of Table 4 (i.e. second row). In this example, the third and fourth coordinate combinations are neglected since objects L and O as well as objects L and P are far from each other as both combinations attain $\Delta x_i + \Delta y_i$ more than 100. For the first type of non-unique coordinate, it can be observed that coordinates (311,141) (for object L) are repeated in first and second combinations.

This coordinate is considered as the first type of non-unique coordinate as it is assigned as *Candidate 1* for more than one combinations. The first type of non-unique coordinate can also be observed for object P (i.e. coordinate (346,241)) where it is assigned as *Candidate 2* in the seventh, ninth and tenth combinations. The second type of non-unique coordinates can also be observed in this example. Coordinate (326,188) (i.e. Object M) appears as *Candidate 2* at the first combination and as *Candidate 1* at the fifth, sixth and seventh combinations. In addition, object N (i.e. coordinate (367,170)) also shows the example of the second type of non-unique coordinate where it is assigned as *Candidate 2* for the second and fifth combinations as well as *Candidate 1* for eighth and ninth combinations. Similar situation can be observed in object O (i.e. coordinate (380,213)) where it became *Candidate 2* for sixth and eighth combinations while it appears as *Candidate 1* in the tenth combination.

Therefore, each non-unique coordinate needs to be analysed and correct coordinate combination needs to be chosen to solve the disconnected object issues. Four cases of disconnected objects are found in this study and successfully identified using the process presented in Figure 23. Each condition to determine all cases of disconnected objects is presented in Table 5.

If the total objects' length is more than 100, the disconnected objects are more likely occluded with other connected objects (*i.e.* other sperm) which can be classified as Case 3. Meanwhile if the total objects' length is less than 100, the disconnected objects are classified as Case 4. The identification of four cases of disconnected objects is presented in Figure 23.

All cases of the disconnected objects must be attached to its own pair in order to ensure accurate detection. Failure to determine correct coordinate combination will cause wrongly connected objects and thus jeopardize the detection of the motile sperms.

When the correct coordinate combinations are identified, smooth line is plotted. This line is plotted to attach two disconnected objects based on the coordinates that are generated from equations (46) and (47).

$$x_{plot} = x_1 + \left[\frac{0:L_{o(p)}}{50:L_{o(p)}} \right] (\cos\theta) \quad (46)$$

$$y_{plot} = y_1 + \left[\frac{0:L_{o(p)}}{50:L_{o(p)}} \right] (\sin\theta) \quad (47)$$

where *distance* and θ are the distance and angle between two end coordinates and calculated using equations (45) and (48) respectively.

$$\theta = \tan^{-1} \left[\frac{(y_2 - y_1)}{(x_2 - x_1)} \right] \quad (48)$$

wherein this case, x_1 and y_1 are the x and y end coordinates of *Candidate 1*. x_2 and y_2 are the x and y end coordinates of *Candidate 2*. x_{plot} and y_{plot} generate multiple coordinates that connect from end coordinate of *Candidate 1* to end coordinate of *Candidate 2*. Techniques involved in solving four cases of disconnected objects are presented in the next subsections.

4.4.2 Case 1

Case 1 of disconnected object is the simplest case to be solved. Similar to the other cases, this type of disconnected object is separated where $\Delta x_i + \Delta y_i$ is less than 100. In addition, this type of disconnected objects has unique coordinates between intra-group *Candidate 1*, unique coordinates between intra-group *Candidate 2* and unique coordinates between inter-group of *Candidate 1* and *Candidate 2*. However, this case is different from other cases since the disconnected objects resemblance of an isolated sperm. The objects are far from other sperms where no occlusion scenario is observed. This case is easier to be solved since no other disconnected objects need to be considered. The disconnected object can be directly connected to its combination without any measurement and criteria to be fulfilled as shown in **Error! Reference source not found.** The disconnected objects of Q and R are attached to each other using the line plotted which their coordinates are generated using equations (46) and (47).

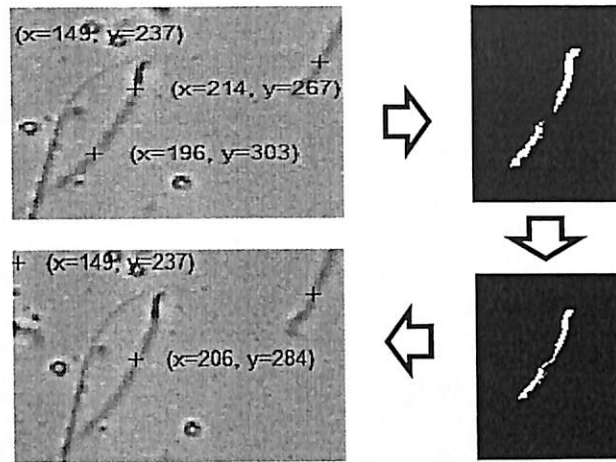


Figure 24 Implementation of Case 1 of the disconnected object

4.4.3 Case 2

As for Case 2 of the disconnected object, sperms are moving in parallel and closed to each other as shown in Figure 25. Objects I and J are disconnected from a single object of motile sperm while object K is another adjacent motile sperm. Initial tracker of motile sperms inaccurately detected three disconnected objects that caused false detection. There are three possible coordinate combinations between two disconnected objects: i) combination between objects I and J, ii) combination between objects J and K and, iii) combination between objects I and K. However, combination between objects I and K is neglected in the analysis since its $\Delta x_i + \Delta y_i$ is more than 100 as shown in the first example of Table 4.

Figure 25 depicts that by relying on the distance between objects or $\Delta x + \Delta y$, object J appears to be closer to K than its correct combination of object I. Therefore, another criterion is considered to solve this case, where the angle between centres of the disconnected objects is calculated (Figure 25 (c)). The angles between centers are calculated in clockwise direction and it is observed that angle between centers of objects I and K is higher than the angle between centers of objects I and J. With the assumption that the same orientation of disconnected objects indicates that they belong to one object, the correct coordinate combination is chosen.

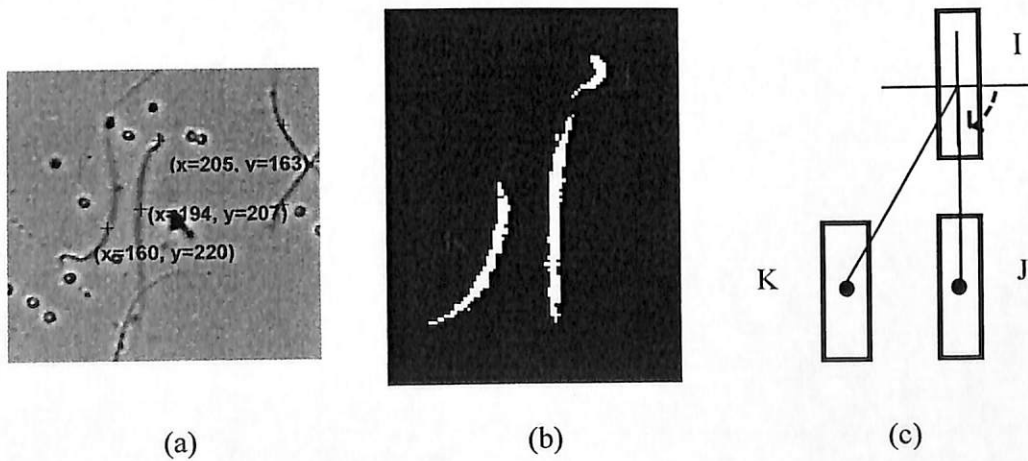


Figure 25 Parallel disconnected object in (a) grayscale (b) binary (c) model of parallel disconnected object where angles between centers are measured

The accurate coordinate combination should attain minimum angle between centers as compared to the other pair combinations as shown in Figure 25 (c). The chosen coordinate

combination is connected using the similar equations (equations (46) and (47) as Case 1 disconnected object and its implementation is presented in Figure 26.

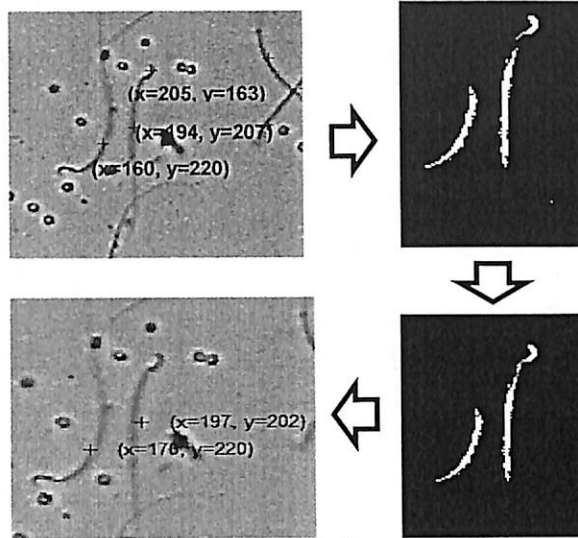


Figure 26 Implementation of Case 2 disconnected object case

4.4.4 Case 3

In this section, the most complicated disconnected object case is presented. This case is prominent to be solved since it is associated with clumped or occlusion between sperms. Figure 27 shows the Case 3 disconnected object in grayscale and binary images as well as model of intersected disconnected object. This figure shows that in the initialization of motile sperm tracker stage, five coordinates are detected and defined as sperm. However, only four of the detected coordinates are the disconnected objects (i.e. objects L and M as well as objects O and P) that should be attached to represent motile sperms.

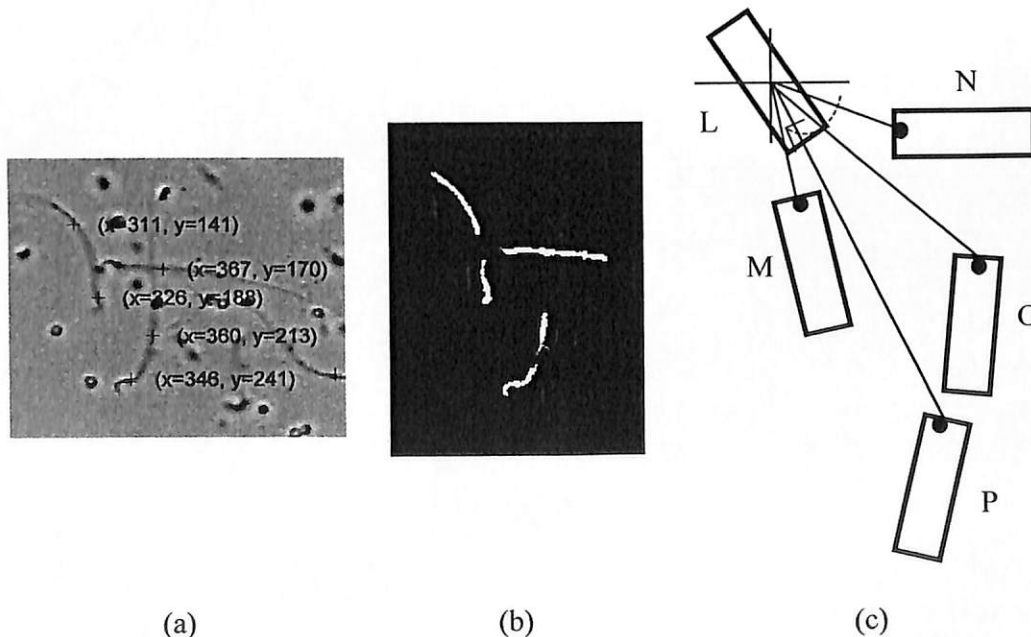


Figure 27 Case 3 disconnected object in (a) grayscale (b) binary (c) model of intersected disconnected object where angles between center and end objects are measured

There are 10 possible coordinate combinations as shown in second example of For clarification of the aforementioned situations of non-unique coordinate can be referred to the second example of Table 4 (*i.e.* second row). In this example, the third and fourth coordinate combinations are neglected since objects L and O as well as objects L and P are far from each other as both combinations attain $\Delta x_i + \Delta y_i$ more than 100. For the first type of non-unique coordinate, it can be observed that coordinates (311,141) (for object L) are repeated in first and second combinations.

This coordinate is considered as the first type of non-unique coordinate as it is assigned as *Candidate 1* for more than one combinations. The first type of non-unique coordinate can also be observed for object P (*i.e.* coordinate (346,241)) where it is assigned as *Candidate 2* in the seventh, ninth and tenth combinations. The second type of non-unique coordinates can also be observed in this example. Coordinate (326,188) (*i.e.* Object M) appears as *Candidate 2* at the first combination and as *Candidate 1* at the fifth, sixth and seventh combinations. In addition, object N (*i.e.* coordinate (367,170)) also shows the example of the second type of non-unique coordinate where it is assigned as *Candidate 2* for the second and fifth combinations as well as *Candidate 1* for eighth and ninth combinations. Similar situation can be observed in object O (*i.e.* coordinate (380,213)) where it became *Candidate 2* for sixth and eighth combinations while it appears as *Candidate 1* in the tenth combination.

Therefore, each non-unique coordinate needs to be analysed and correct coordinate combination needs to be chosen to solve the disconnected object issues. Four cases of disconnected objects are found in this study and successfully identified using the process presented in Figure 23. Each condition to determine all cases of disconnected objects is presented in Table 5.

Table . However, combinations between objects L and O as well between objects L and P are eliminated since both have $\Delta x_i + \Delta y_i$ more than 100. For Case 3, two parameters namely angle between center and end coordinates (Figure 27 (c)) and distance between end coordinates (Figure 28) are introduced. Figure 28 shows the model of intersected object where the distances of the end coordinates are calculated. The distances between end coordinates are calculated from the other coordinates (*i.e.* black circle) with respect to its reference end coordinate denoted in red circle. This figure attests that small distance can be observed between objects L and M, objects O and P and objects L and N.

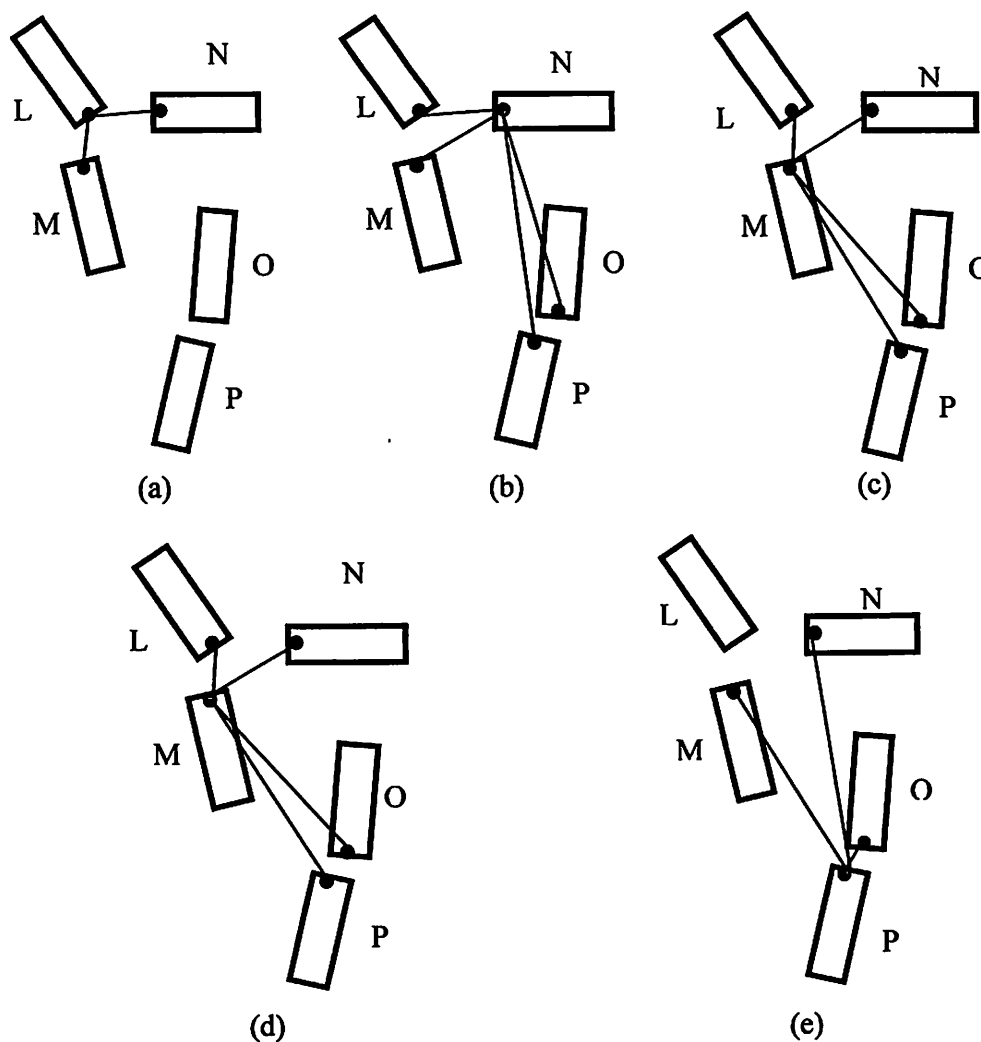


Figure 28 Distance between end coordinates with respect to (a) Object L (b) object N (c) object M (d) object O and (e) object P

However, the disconnected object can only be attached to its own unique combination that has similar orientation. In other words, if object of *Candidate 1* is horizontal, thus its combination of *Candidate 2* must have horizontal orientation or vice versa. Therefore, to avoid non-unique coordinate combinations of objects L and M as well as objects L and N, another parameter is measured. The angle between center and end of the disconnected objects are measured as shown in Figure 27 (c). The intersected disconnected objects are attached to their particular combination only if two conditions are satisfied; i) distance between end coordinates is minimum and ii) maximum angle between center and end coordinates that have similar orientation. If the angle of the centre and end coordinates is not equal to maximum, the coordinate combination is ignored since it is most likely to represent different orientation.

After the accurate coordinate combinations are identified, the disconnected objects are attached to their own combination for detection process. Misidentification of the five coordinates that are initially detected is rectified by connecting these objects using equations (46) and (47). The resulted motile sperms detected coordinates are presented in Figure 29. Connecting disconnected object with correct combination ensures that the intersected or

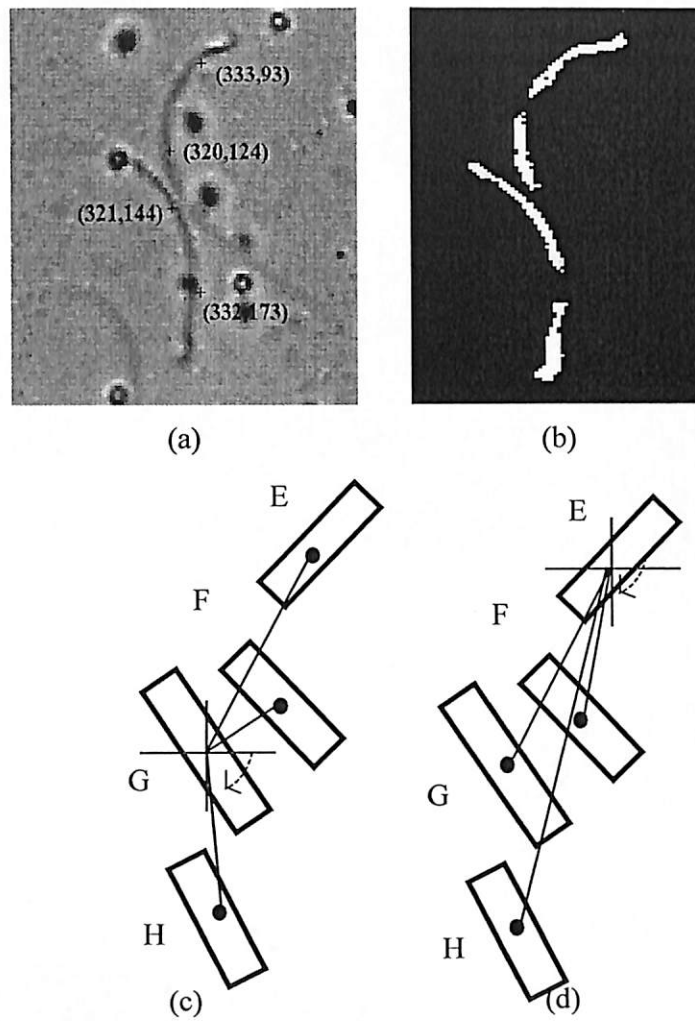


Figure 30 Case 4 disconnected object in (a) gray scale image (b) binary image (c) and (d) model of combined disconnected object where angle between center of the disconnected are measured

Therefore, another parameter is measured and acts as another criterion to be considered to find the correct combination of the disconnected object. Angle between two centers is computed for each possible coordinate combination as shown in Figure 30(c) and (d). This angle is calculated in clockwise direction where the angles between two centers are computed for each possible coordinate combination with respect to object G and E respectively.

As an addendum to the measured angle between two centers, the distance between center and end of object is also computed as shown in Figure 31. This figure shows how the distances between center and end objects are calculated for all possible coordinate combinations. The correct combination of the objects is indicated by having a minimum distance between center and end of the object and minimum angle between two centers. Objects are classified as combination of coordinates to be connected if those criteria are fulfilled as shown in Figures 30 and 31. These criteria are made based on the assumption that two disconnected objects must have similar orientation to be considered as one object (*i.e.* sperm). The smallest angles are obtained from combination of objects G and H (Figure 30(c))

as well as combination of objects E and F (Figure 30(d)). The results are also supported by the measured distance between center of object E and end of object F (Figure 31 (a)) or vice versa as well as distance between center of object G and end of object H or vice versa (Figure 31 (g)).

The correct combination of objects E and F as well as combination of objects G and H are then attached to each other using a line plotted based on the coordinates generated using equations (46) and (47). The implementation of connecting the combined disconnected objects is illustrated in Figure 31.

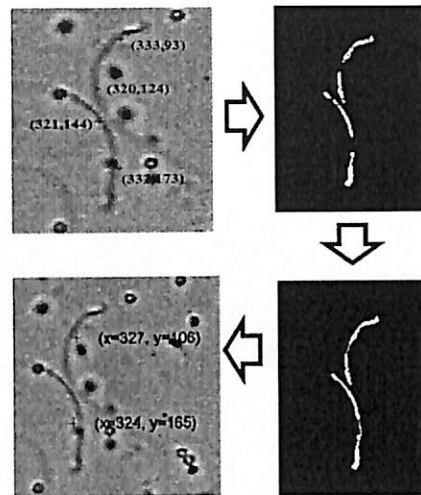


Figure 31 Implementation of Case 4 of disconnected objects

When the disconnected objects cases are solved and objects are successfully attached to their pair coordinate combination, final center coordinates of the sperms are identified. Table 6 summarises utilized criteria to solve each case of the disconnected object.

Table 6 Criteria for each disconnected objects

Case	Parameter to be calculated to solve each case	Parameter chosen to solve each case
Case 1	<ul style="list-style-type: none"> • $\Delta x + \Delta y$ 	<ul style="list-style-type: none"> • No criteria to be chosen since the disconnected objects are far from other objects
Case 2	<ul style="list-style-type: none"> • Angle between two center 	<ul style="list-style-type: none"> • minimum angle between two centres
Case 3	<ul style="list-style-type: none"> • Distance between end coordinates of object • Angle between center and end coordinates of object 	<ul style="list-style-type: none"> • minimum distance between end coordinates of object • maximum angle between center and end coordinates of object
Case 4	<ul style="list-style-type: none"> • Angle between two center • Distance between centre and end coordinates of the object 	<ul style="list-style-type: none"> • minimum angle between two centres • minimum distance between centre and end coordinates of the object

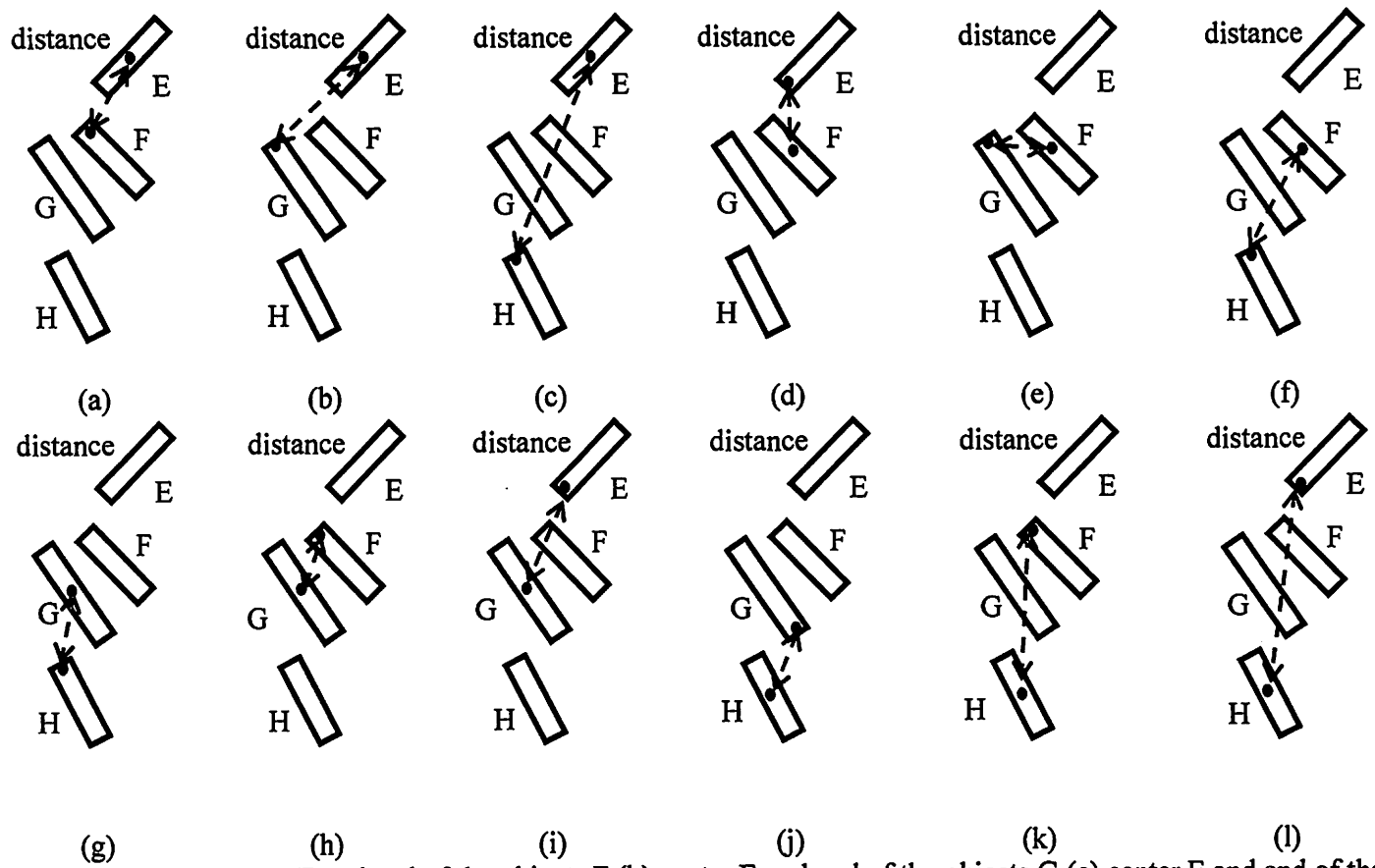


Figure 32 Distance between (a) center E and end of the objects F (b) center E and end of the objects G (c) center E and end of the objects H (d) center F and end of the objects E (e) center F and end of the objects G (f) center F and end of the objects H (g) center G and end of the objects H (h) center G and end of the objects F (i) center G and end of the objects E (j) center H and end of the objects G (k) center H and end of the objects F (l) center H and end of the objects E

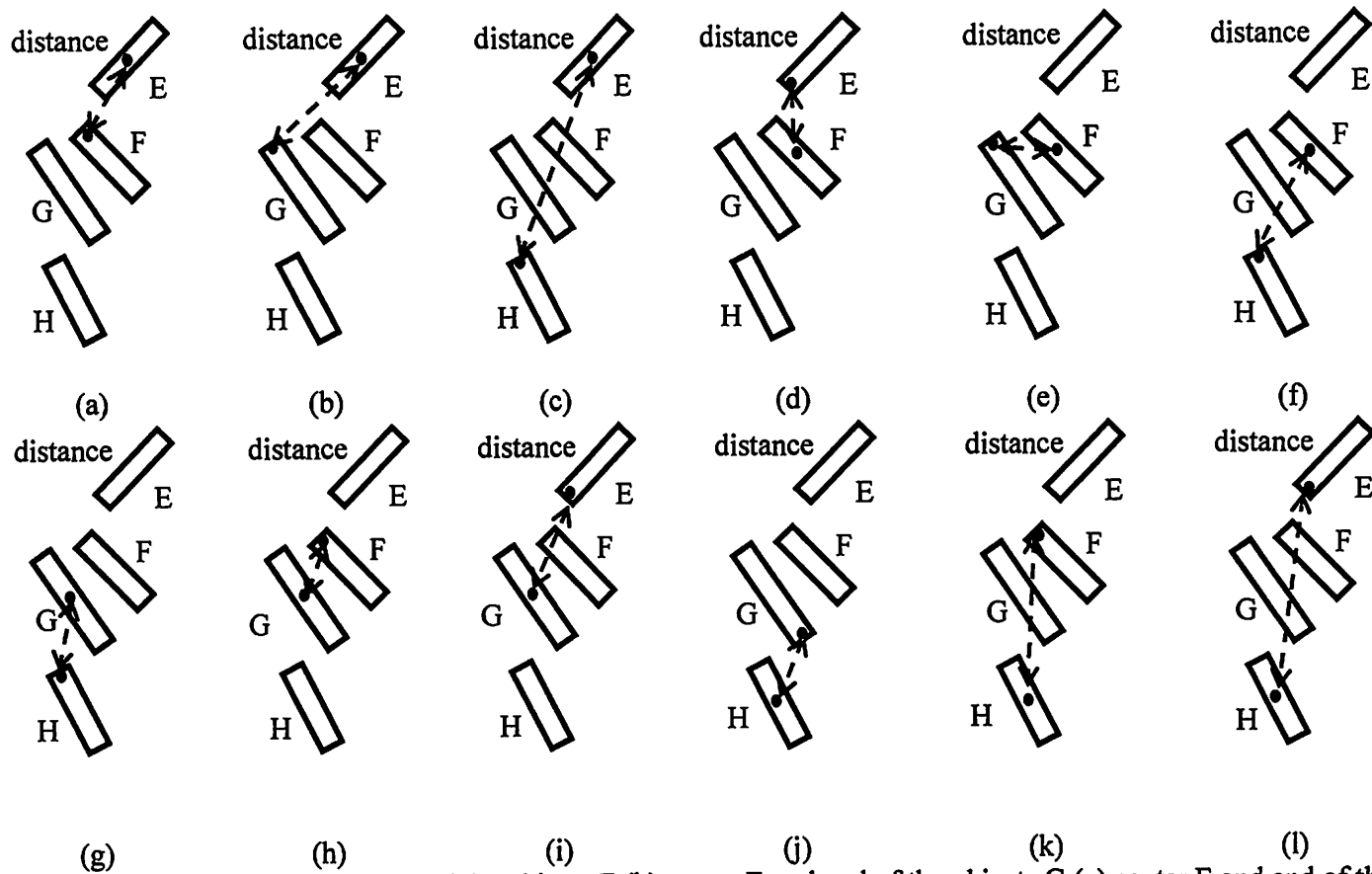


Figure 32 Distance between (a) center E and end of the objects F (b) center E and end of the objects G (c) center E and end of the objects H (d) center F and end of the objects E (e) center F and end of the objects G (f) center F and end of the objects H (g) center G and end of the objects H (h) center G and end of the objects F (i) center G and end of the objects E (j) center H and end of the objects G (k) center H and end of the objects F (l) center H and end of the objects E

4.4.6 Sperm Motility Detection

When issues with the disconnected object cases are addressed and resolved in the previous stage, foreground detection is conducted to locate the center of motile sperms. The proposed system provides an advantage compared to the other CASA instruments where automated initialization of the motile sperms is performed prior to the tracking proses. Many CASA instruments are still relying on the manual or semi-automated detection system in the beginning of the tracking process where initial locations of the motile sperms are manually determined by the physician before sperm trajectory analysis could be performed [12-14].

The centers of the motile sperms are detected by computing the central moment of the connected object (obtained from previous stage) as denoted and discussed previously in Section 3.5. The center coordinates of x and y are given in equations (35) and (36) respectively. These center coordinates are located at the middle region of the sperms as shown in Figure 33. In the proposed system, the foreground is detected in the binary image that contains connected objects obtained from the previous stage.

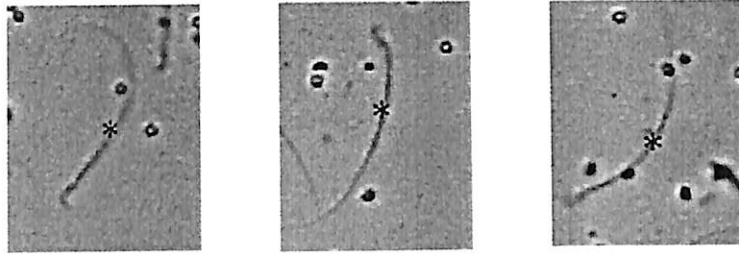


Figure 33 Detected sperms on the first frame

The aim of this stage is to find the sperm in the previous frame that is closest and most similar to the sperm in the current frame. The closeness of the sperm is defined by having a minimum Euclidean distance between centers of two sperms of S_c (*i.e.* sperm in current frame) and S_p (*i.e.* sperm in previous frame) given by:

$$d(S_c, S_p) = \sqrt{\sum_{i=1}^2 (x_c - x_p)^2 + (y_c - y_p)^2} \leq t_{dist} \quad (49)$$

where (x_c, y_c) is the center coordinate of the sperm in current frame while (x_p, y_p) is the center coordinate of the sperm in previous frame. t_{dist} is the distance threshold that is set to reduce the searching area of the tracking system.

The distance threshold is set to 125 based on the observations and tests performed on the 100 sperm motility videos. However, by computing only closeness of the sperms in consecutive frames is not sufficient to track the motile sperms. This is because most sperms are moving close to each other where occlusion occurred in certain cases. Thus, sperms might be matched incorrectly which could cause incorrect sperm trajectory analysis.

Realizing the fact, the similarity between sperms in current and previous frames is determined using equation(50). The similarity is calculated by using the size ratio of the sperms where the S_{z_p} and S_{z_c} are the sizes of the S_p and S_c respectively. t_{size} is the size threshold that is set to 1.2 in this study with the assumption that the object's size is not expected to change significantly between antecedent and subsequent frames. However the appearance of the sperm size might slightly varied due to the uncontrolled sperm's movement as shown in Figure 34.

$$Similarity = \begin{cases} 1 & \text{if } Sz_p > Sz_c \text{ and } \frac{Sz_p}{Sz_c} \leq t_{size} \text{ or } Sz_c > Sz_p \text{ and } \frac{Sz_c}{Sz_p} \leq t_{size} \\ 0 & \text{otherwise} \end{cases} \quad (50)$$

Therefore, there are two conditions need to be considered to match the sperm in the current frame with its corresponding at previous frame. If the Euclidean distance and size ratio between sperm in current and previous frame is less than or equal to the predetermined thresholds, sperm in the previous frame is matched to its corresponding sperm with the sperm in the current frame.

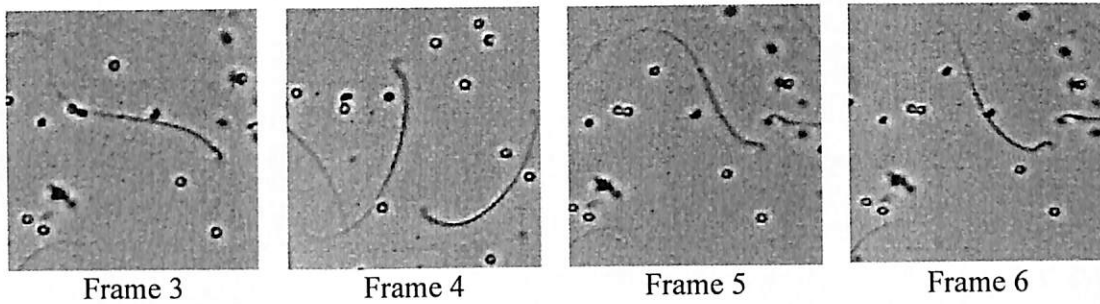


Figure 34 Image sequence frame where the appearance of the sperm size is slightly varied due to its movement

If the closeness and similarity of the objects are not achieved, object is considered as sperm that swims out from the frame. The sperm tracker is reinitialized where the non-matching sperm is removed from the tracker. Figure 35 shows the flow chart of the motile sperm tracker of the proposed system. If the sperm is successfully matched with its corresponding sperm, the process is repeated until all sperms in the frame are processed. This process is continuously conducted until all image sequence frames are analysed.

Since the intersected sperm case is solved in the previous stage, the clumped sperm can be detected individually. Thus the occlusion between sperms can be accurately identified. The distance travelled for each sperm and its velocity is calculated using equation (51). This equation is utilized to calculate curvilinear velocity and straight line velocity where the former velocity is computed for every frame from start to end points of the motile sperms. Meanwhile the straight line velocity is computed by considering only start and end points of the distance travelled by the motile sperm.

$$velocity = \frac{\sqrt{(x_c - x_p)^2 + (y_c - y_p)^2}}{\text{frame rate}} \quad (51)$$

When the velocity of the sperm is calculated, the motile sperms are classified into three categories namely immotile, slow progressive motile and fast progressive motile. The sperms are considered as immotile if their VCL is less than 1.5 pixel/second. For the sperms with VCL more than 1.5 pixel/second are considered normal and can be divided into slow progressive and fast progressive groups. If the velocity of the sperm is more or less than 3 pixel/second, the sperm is considered as fast progressive or slow progressive sperm respectively as shown in Figure 36.

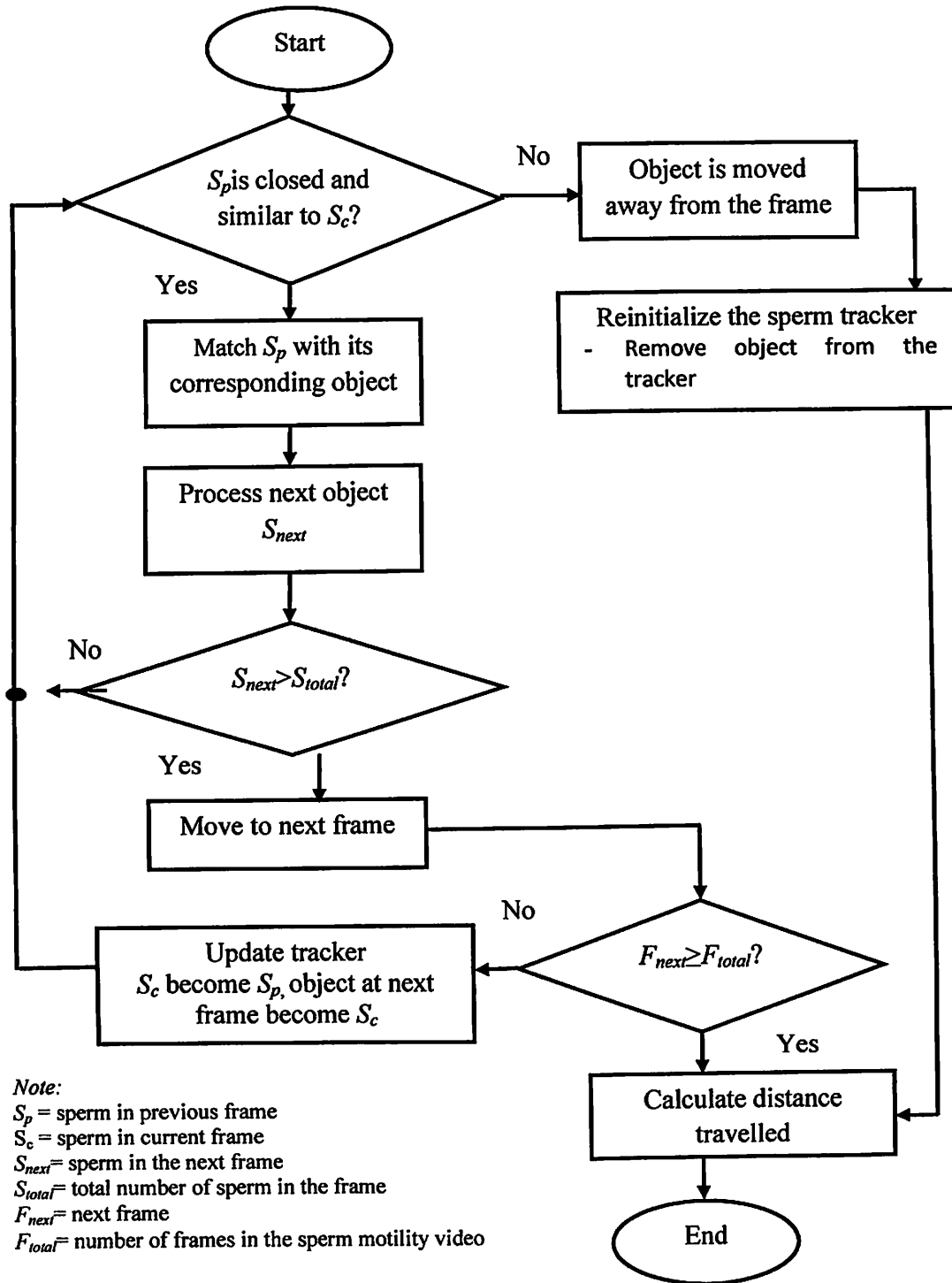


Figure 35 Correspondence-based object matching and re-initialization of the sperm tracker in the proposed system

The performance of the proposed system on sperm detection process is investigated using four metrics which include true positive (TP), true negative (TN), false positive (FP) and false negative (FN). These metrics are employed to address the accuracy, sensitivity, and specificity of the proposed system. All TP, TN, FP and FN values are calculated based on whether the system is able to detect the motile sperms. The detection is classified as TP or

TN if the system correctly defines the immotile sperms or the systems correctly defines the motile sperms respectively. Meanwhile, FN represents the total number of the sperms that should have been classified as immotile sperms according to the physician, however classified wrongly by the system. FP is defined as the total number of immotile sperms which according to the physician does not belong to the motile group however classified falsely into the group. These four metrics are summarised in Table 6. The accuracy, sensitivity and specificity of the system are calculated using equations (52) to (54).

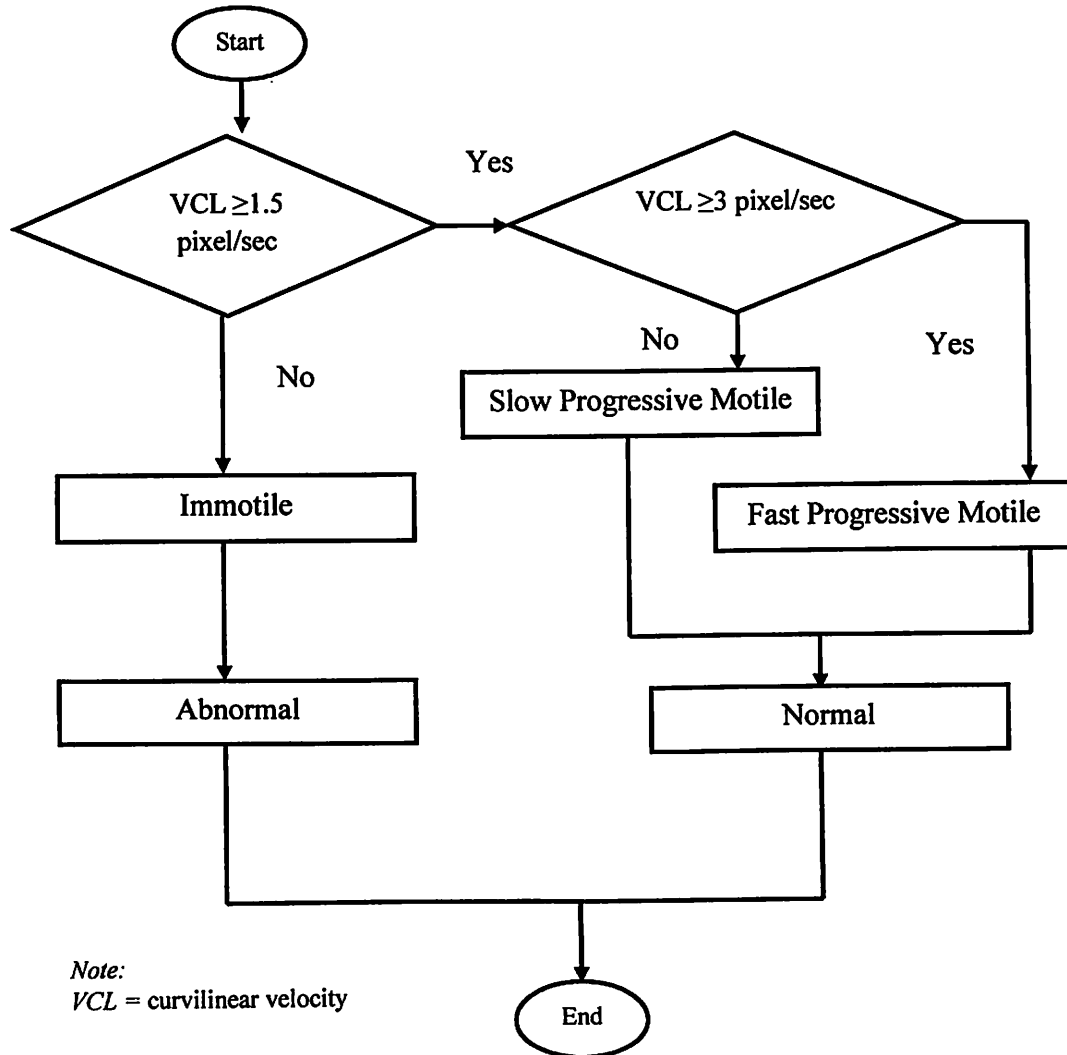


Figure 36 Sperm classification process

$$\text{Accuracy} = \frac{TP + TN}{\text{total data}} \times 100\% \quad (52)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \times 100\% \quad (53)$$

$$\text{Specificity} = \frac{TN}{FP + TN} \times 100\% \quad (54)$$

The accuracy denotes how close the proposed system achieved accurate detection in accordance to the detection manually made by the physician. Meanwhile the sensitivity measures the proportion of positives which the system correctly identified as immotile sperms. Specificity measures the proportion of negatives which the system correctly identified as motile sperms according to physician.

Table 7 Four metrics for accuracy, sensitivity and specificity calculations

Detection by Proposed system	Detection by physician	
	Motile	Immotile
Motile	TN	FN
Immotile	FP	TP

Figures 37, 39 and 41 show the results of trajectory analysis for three motile from different samples of sperm motility videos. These figures show the example of few cropped frames extracted from *samples 1, 2 and 3* sperm motility videos. The trajectory analysis presented in Figure 36 shows the example of slow progressive motile sperm where the movement of the sperms (sperms 1 to 3) are relatively small. It can be observed from 30th frame to 45th frame, sperms 1, 2 and 3 are hardly dislocated from its original place (as shown in 10th frame). All sperms are not moving linearly ('straight line direction') where their non-progressive motile tract resulted in small distance travelled from their original to their last locations. These sperms are non-progressively moving in 'circle direction' which can be supported by their velocities plotted in Figure 38.

Figure 38 shows the plotted curvilinear velocity (VCL) for sperms 1, 2 and 3 as their trajectory previously shown in Figure 37. The VCL is plotted by calculating velocity for each travelled frame using predefined equation (51). As examples, for travelled frame 1, the VCL is calculated based on the distance travelled by sperm from frame 1 to frame 2, whereby for travelled frame 2, the VCL is computed by calculating the distance travelled by sperm from frame 2 to frame 3 and so on.

The average VCL and straight-line velocity (VSL) for the sperms presented in Figures 37 and 38 are tabulated in Table 8. Based on the average VCL presented in this table, the results from the trajectory analysis concluded that the sperms 1, 2 and 3 are classified as normal with slow progressive motile category. Although it can be observed in the Figure 38 that the sperms attained high velocities (in the range of more than 3 pixels/second) for certain travelled frames, the sperms are not progressively moving in linear direction. This problem could reduce the chances of the sperms to move in the proper direction to fertilize the egg.

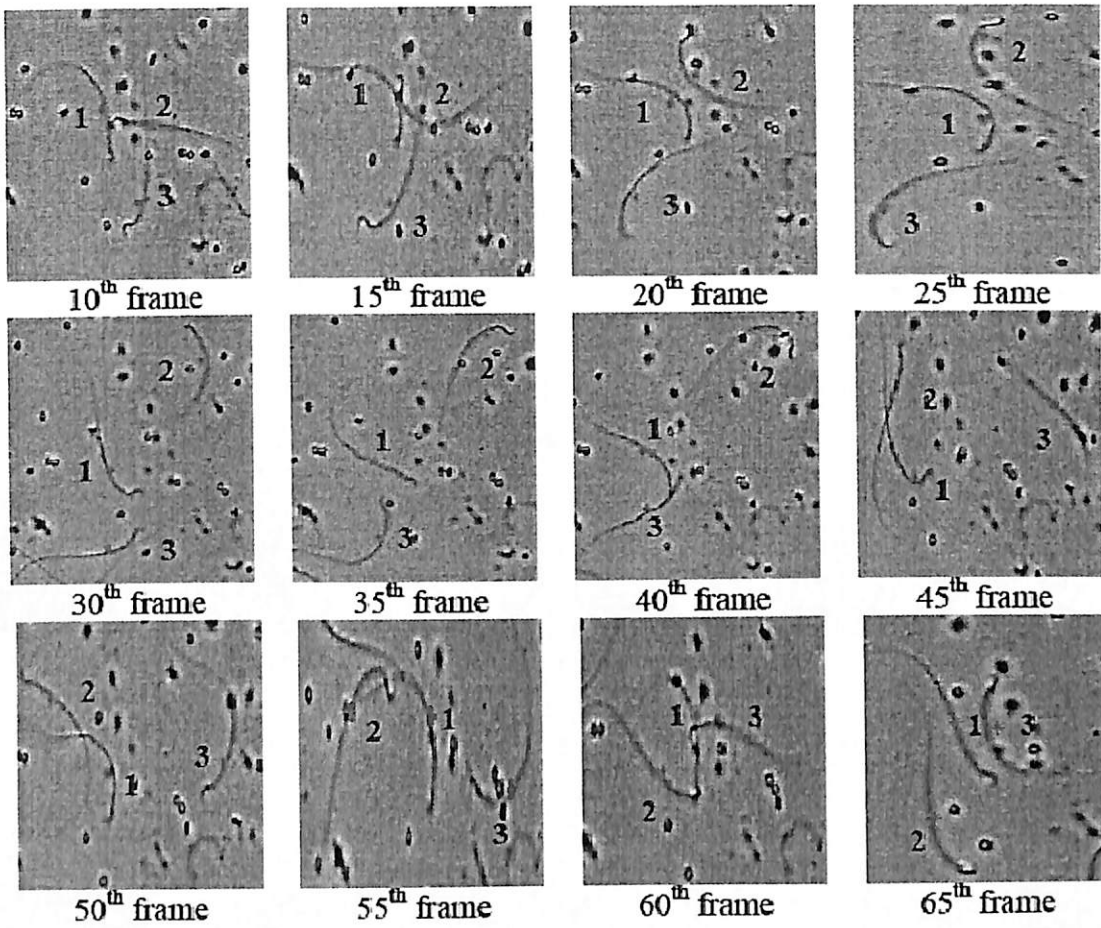


Figure 37 Results of sperm trajectory analysis for cropped frames of *sample 1*

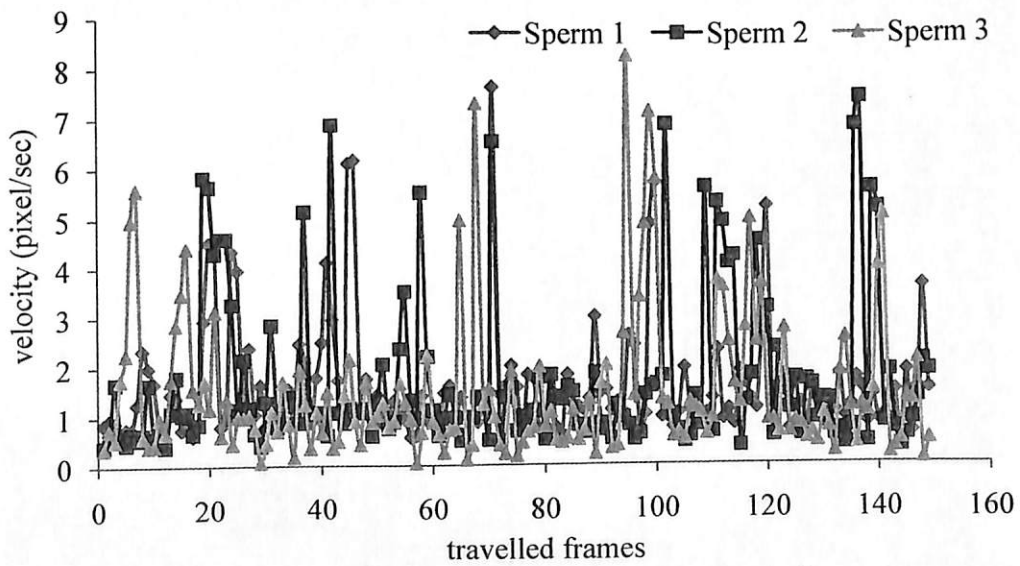


Figure 38 Curvilinear velocity for *sample 1*

Another example of sperm trajectory analysis is presented in Figure 39. The examples of few frames with less debris are presented in this figure. The trajectory analysis is performed for these three sperms and their VCL and VSL are calculated. Sperm 5 is slow-progressively motile where the sperm is moving in same area as its original location. This sperm is also collided with sperms 4 and 6 where they are moving closely together in several frames as can be seen in the 10th, 15th, and 20th frames. The collision between sperms has decreased their velocity which also supported in their plotted graph of VCL in Figure 40 where it is depicted that the VCL is low at the beginning of the travelled frames. Meanwhile, sperm 6 is progressively motile with the average VCL of 3.73 pixel/sec which is classified as fast progressive motile sperms (Table 8).

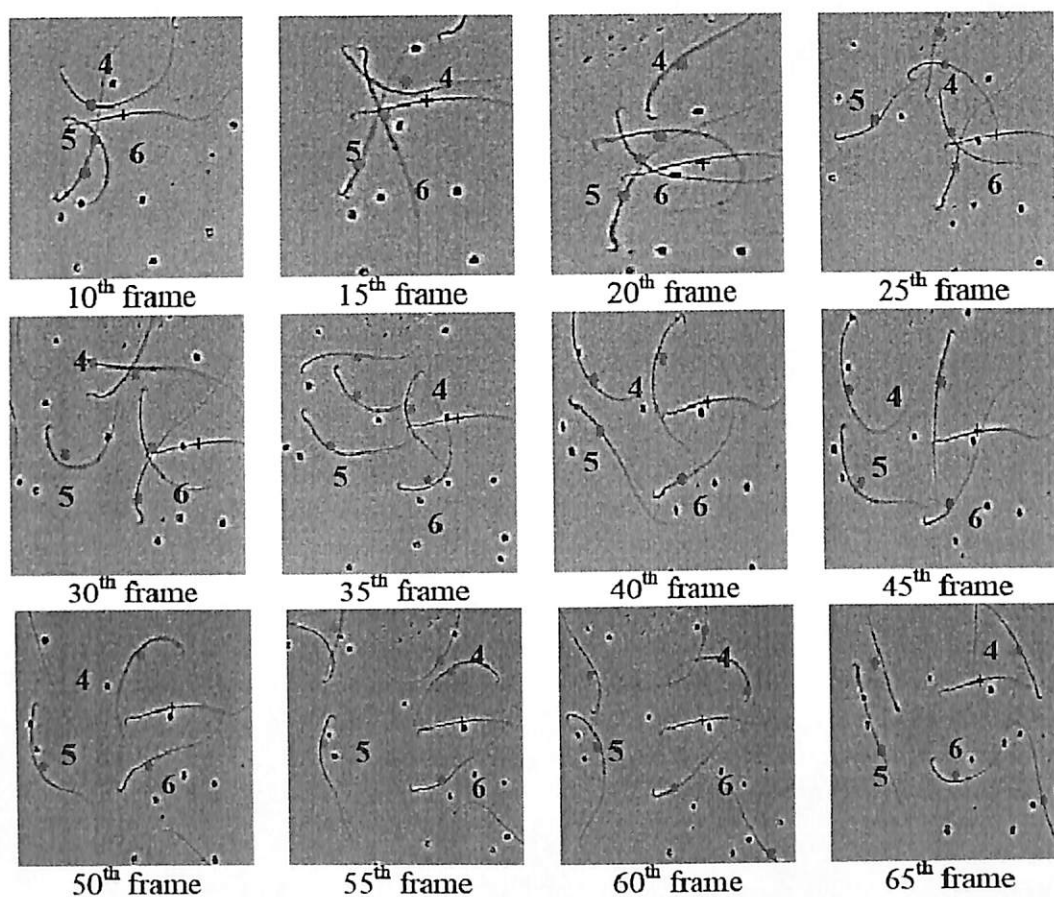


Figure 39 Results of sperm trajectory analysis for cropped frames of *sample 2*

The example of trajectory analysis for *sample 3* and its respective VCL is plotted and presented in Figures 41 and 42 respectively. Sperms 7 and 8 are classified as normal where their trajectory is progressively motile where both sperms are categorized as fast progressive motile with their average VCL of 4.30 pixel/second and 3.87 pixel/second respectively (Table 4.6). Sperm 9 is considered as immotile (*i.e.* abnormal) where its distance travelled is the smallest as compared to sperms 7 and 8. The proximal region of the sperm is non-progressively moved which resulted in lower velocity as can be observed in the nearly zero VCL for most travelled frame as plotted Figure 42.

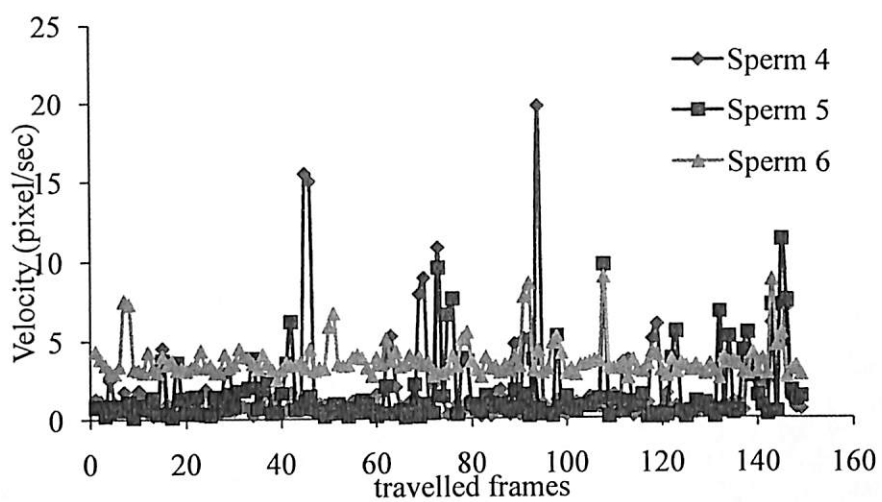


Figure 40 Curvilinear velocity for *sample 2*

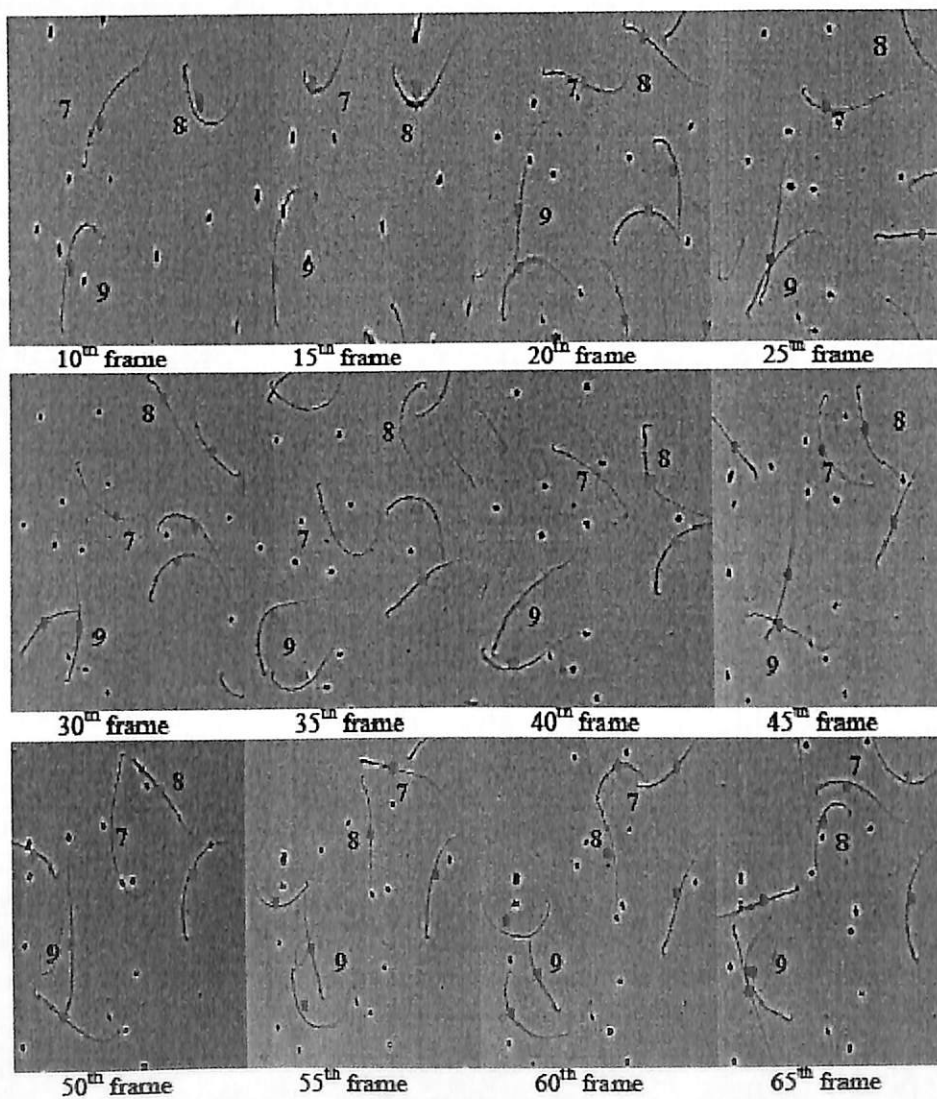


Figure 41 Results of sperm trajectory analysis for cropped frames of *sample 3*

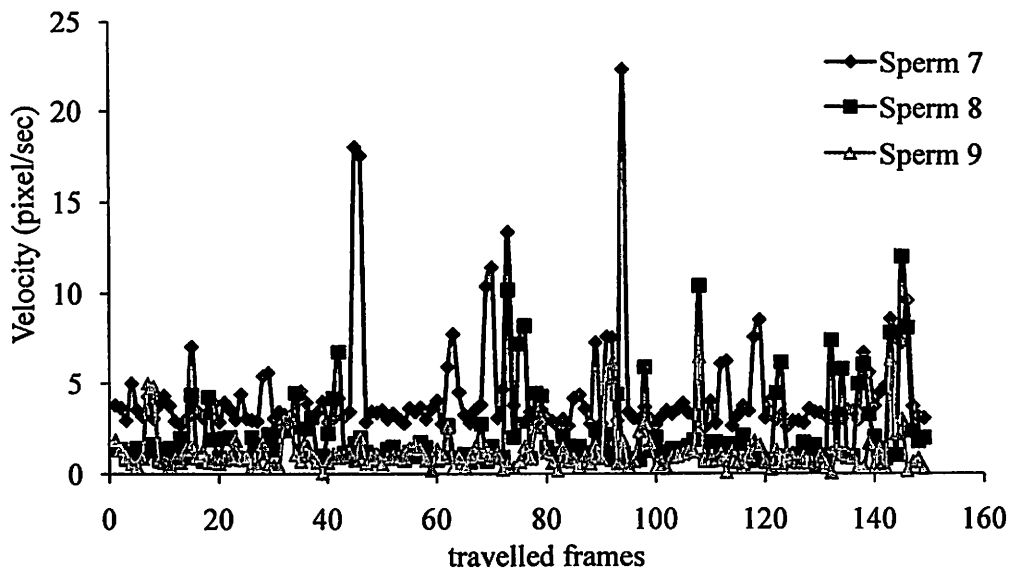


Figure 42 Curvilinear velocity for *sample 3*

The trajectory analyses presented in Figures 37, 39 and 41 show that the motile sperms are successfully identified. The motile sperms are successfully tracked including the occlusion cases between sperms as shown in Figures 37 and 41. Issues with occlusion cases are previously addressed where individual sperms can be identified. The identification of the sperms encountered in the occlusion case will ensure the accuracy of the sperm motile detection. Therefore, the computed velocity can be used to classify the sperm into either three categories of motile sperms (*i.e.* fast progressive motile, slow progressive motile and immotile).

Table 8 summarised the average VCL and VSL for *samples 1 to 3* as presented in Figures 37, 39 and 41 respectively. Results presented in this table indicate that the sperm 6 in *sample 2*, sperm 7 in *sample 3* and sperm 8 in *sample 3* are classified as fast progressive motile sperms with average VCL of 3.73 pixel/second, 4.30 pixel/second and 3.87 pixel/second respectively. Meanwhile, their VSL are 24.12 pixel/second, 25.31 pixel/second and 24.43 pixel/second for sperm 6 (*sample 2*), sperm 7 (*sample 3*) and sperm 8 (*sample 3*) respectively.

Sperm 9 from *sample 3* are classified as immotile with the average VCL of 1.18 pixel/second and VSL of 12.19 pixel/second. The other sperms presented in this table are classified as slow progressive motile since their average VCL is more than 1.5 pixel/second however, their velocity is less than pre-defined threshold of fast progressive motile (*i.e.* 3 pixels/second).

The performance of the proposed system on detecting the motile sperms is further discussed where the accuracy, sensitivity and specificity are calculated. The four metrics are true positive (TP), true negative (TN), false positive (FP), and false negative (FN) as summarized in Table 7. These four metrics are assessed based on whether the system is able to detect the motile sperms. Figures 43 (a), (b) and (c) show the example of the sperm motility detection process for *sample 1*, *sample 2* and *sample 3* respectively. Figure 43 (a) shows the example where total number of debris is more than the total number of sperms whereby in Figures 43 (b) and (c), total number debris and sperms are almost similar.

training dataset. The evolving process continues until the over-fitting condition is realized and therefore, a reasonable fuzzy rule-base is achieved.

(ii) Semantic meaning: Evolving information granule should be specific in a way that well-defined semantics are experienced. Therefore, highly detailed information granule is required for better reflection of the existing experimental data. The formation of the information granule has to adhere to the interpretability constraints at both the fuzzy partition level and rule-base level; these constraints are considered for the interpretability-accuracy tradeoff.

These two requirements are apparently in conflict and interpretability-accuracy tradeoff appeals from the intuitive perspective [15]. Therefore, the operational framework for evolving information granule is formed in the proposed EIG where a sound compromise can be formed between interpretability and accuracy. Having these two requirements in mind, evolving information granule for the output-context fuzzy system is described in detail in the following sections. The proposed EIG evolves the information granule as a self-automated process and fully data driven approach. Figure 44 shows the evolving process and consistency model to ensure the aforementioned requirements. Figure 44(a) shows a flowchart to realize the distinct information granule based on experimental evidence. Termination occurs in the evolving process at an over-fitting state where an effective rule-base is realized and termination index is fully online and estimated from the current and previous evolving states. After the formation of the effective rule-base, the proposed EIG defines the consistency model for rule-base which is depicted in Figure 44(b). Subsequently, the decision matrix is defined for the i th validation input $[x, d]_i^{val}$ which shows a logical view of the rule-base in terms of the validation input. Logical representation explains the momentous value of the validation input within the rule-base and hence a consistent rule is realized. The outcome on the experimental data shows the effectiveness of the proposed system.

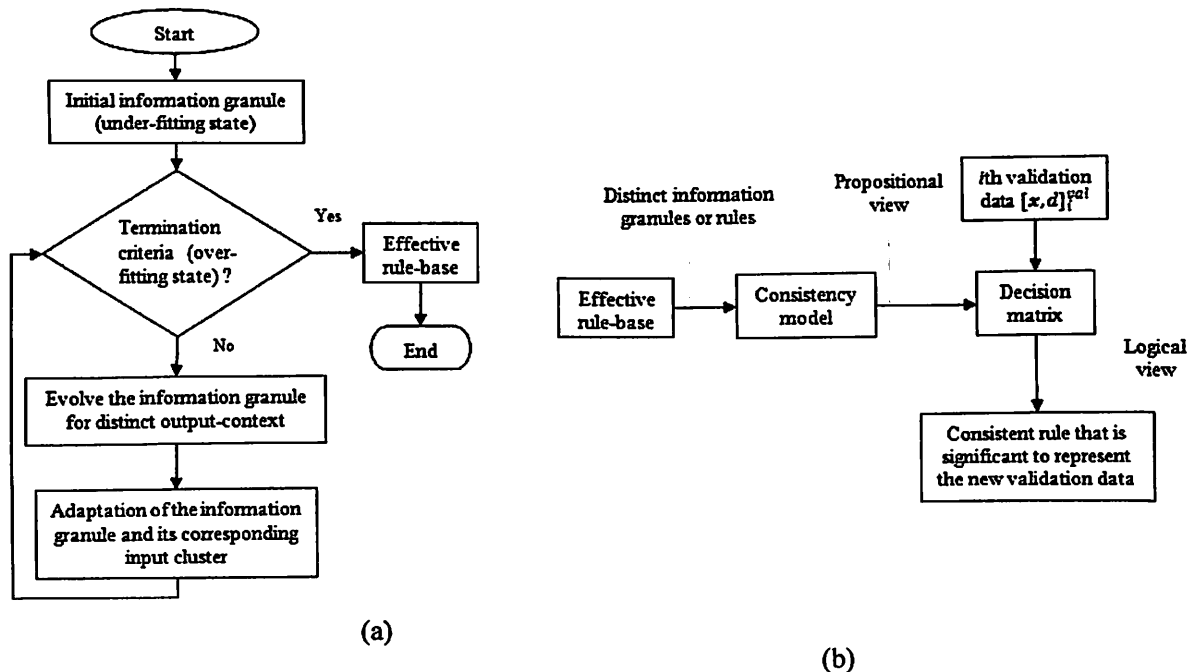


Figure 44 The proposed model adopted in the EIG involves (a) evolving the information granule to realize the effective rule-base and (b) consistency model from effective rule-base and its decision matrix.

5.2 Consistency Model for Conflict Decision

The evolving process in the output domain is considered to obtain the output-contexts as information granules. Each granule is embedded with the corresponding input clusters. Unlike the grid-partitioning approach where a grid-like input partition is established [16], the EIG finds prominent distinct output context and its corresponding input space (Figure 45). The grid-partitioning approach isolates the rule centroids so as to ensure interpretability whereas rules isolate themselves from each other in the EIG that are depicted in Figure 44. Therefore, conflict decisions are observed as the data samples are distributed unevenly over the input domain with low space coverage. To resolve the conflict, studies in fuzzy classification focus on improving the decision boundaries in order to obtain high accuracy as much as possible [15,17]. Another method for improving the interpretability of classifiers is rule compression that discards the less significant antecedent part from the individual rules [15,18,19]. Rule compression may cause conflict situation in the unusual part of the input space and make the system lean more towards the inaccurate classifier for unseen samples [15]. Therefore, the consistency model referred in the EIG has the objective to consider the interpretability-accuracy trade-off.

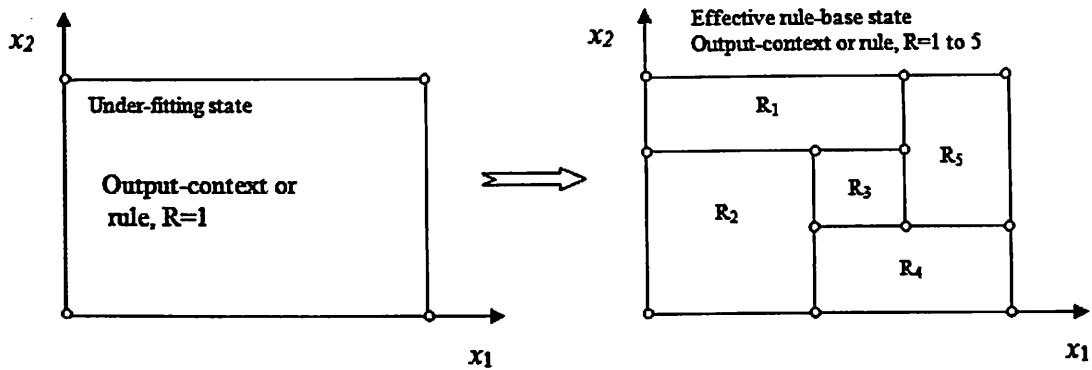


Figure 45 Rule creation for the grid-partitioning approach and EIG. R_1 to R_5 depict the output-context associated with the input cluster (information granule or rule).

5.3 Termination Index

The estimation of the termination index is fully online and is not based on the predefined threshold as in [16], approximated from the previous and the current evolving stage. Evolving granule error (EGE) index is a straightforward index to recognize the over-fitting situation in the evolving process and terminate the algorithm from further evolving.

$$EGE(t) = \frac{E(t)}{E(t-1)}$$

where $E(t)$ and $E(t-1)$ are the approximate training error at t and $(t-1)$ evolving stage, respectively.

5.4 Application in real world data

The proposed EIG has been compared with the existing models. The evaluation of the proposed EIG is carried out for synthetic and real world data.

Example 1: A function approximation problem of one-dimensional data is used to show the performance. We consider 200 data points from the following single-input-single-output nonlinear function:

$$y = 0.6 \sin(\pi x) + 0.3 \sin(3\pi x) + 0.1 \sin(5\pi x)$$

where $x \in [-1,1]$. These 200 data points are then randomly divided from the universe of discourse x : 100 for training and 100 for testing.

Example 2: The function approximation problem involving two-dimensional data is used to show the performance. We consider 200 data points from the following two-input-single-output nonlinear function:

$$y = f(x_1, x_2) = 0.6 + 0.2x_1 + 4x_2 + 0.5x_1x_2 + 25 \sin(0.5x_1x_2)$$

where $x_1 \in [-4,6]$ and $x_2 \in [-2,4]$. These 200 data points are then randomly divided from the universe of discourse x_1 and x_2 : 100 for training and 100 for testing.

Example 3: Considering the following three-dimensional (3D) function approximation:

$$f(x, y, z) = (1 + x^{0.5} + y^{-1} + z^{-1.5})^2$$

with $x \in [1,6]$, $y \in [1,6]$, and $z \in [1,6]$, where 216 training data samples are generated with the step size 1 over this 3D Cartesian product. The test data samples are produced while x , y , and $z \in [1.5, 5.5]$, 125 test samples are taken.

Example 4: A well-known real-world data, Automobile Miles Per Gallon (MPG), is used [20] to evaluate the performance. The output is the fuel consumption of an automobile expressed in miles per gallon; seven input variables are used to distinguish the actual output. For evaluation, we randomly divide the data set into training (60%) and testing (40%) data sets from 316 observations.

Example 5: This data set deals with real estate in the Boston area [21]. It contains 13 input variables with 506 observations; the median value of the house is considered as an output variable. We randomly divide the data set into training (60%) and testing (40%) data sets for evaluation.

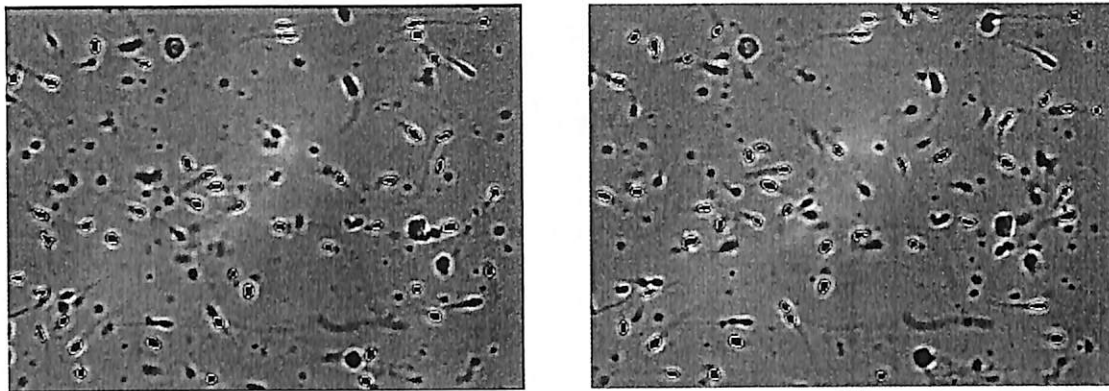
The evaluation results show that the proposed EIG achieves reasonable accuracy, high interpretability in terms of distinct information granules and also that it is reliable with the consistency model. As compared with the existing models [16,18,22,23], this proposed model shows the effectiveness of the operational framework to form the information granules that have a sound compromise between interpretability and accuracy.

The implementation of this new classifier tool on sperm motility analysis exhibit good and comparable results as presented in Section 4. However, this approach requires longer time to be computed and its complexity is higher than the approach presented in previous section. This technique has been successfully applied in another application as discussed in detail in submitted paper no 3 (Section 7: Appendix). Therefore, this technique is currently been improved and submitted for publication (paper no 2, Section 7: Appendix).

6. Verification with Human Sperm Sample

Based on the successfulness of the proposed system on detecting the motile rat sperms, the project is extended to be tested on human sperm samples. However, the project is still in the process of collecting the data since the human ethic approval is expected to be received on the end of March 2014. The detection system on human sperm is conducted on the human sperm sample that the project obtained from the public database.

The proposed system has been modified to suit the detection of the human sperm sample. Although with limited sperm sample obtained from public database, the proposed system has able to detect the sperms and distinguish them from the non-sperm cells (debris) as shown in Figure 46.



(a) (b)
Figure 46 Human sperm sample detection on (a) frame 1 (b) frame 2

The verification on human sperm sample will be conducted in more thorough once the project achieves its human ethic approval.

7. List of Publications

Journals

i) Published/Accepted for Publication

1. K.Hasikin,N.A.M Isa, (2013). Adaptive Fuzzy Intensity Measure Enhancement Technique for Non-Uniform Illumination and Low Contrast Images. *Signal, Image and Video Processing (SIViP)*, DOI: 10.1007/s11760-013-0596-1.
2. S.H. Lim, N.A.M Isa, (2013) A New Histogram Equalization Method for Digital Image Enhancement and Brightness. *Signal, Image and Video Processing*. DOI:10.1007/s11760-013-0500-z
3. Fadzil Ahmad & NorAshidi Mat Isa. (2013). “Intelligent medical disease diagnosis using improved hybrid genetic algorithm – multilayer perceptron network” *SPRINGER Journal of Medical Systems*, DOI 10.1007/s10916-013-9934-7
4. K. Hasikin, N.A.M. Isa, (2012). Adaptive Fuzzy Contrast Factor Enhancement Technique for Low Contrast and Non-Uniform Illumination Images. *Signal, Image and Video Processing*. DOI:10.1007/s11760-012-0398-x

ii) Submitted for Publication

1. K.Hasikin,N.A.M Isa, (2014). Automated Feature-Based Sperm Motility Analyzer for Sperm Motility Detection: Effect of Enhancement on Detection Analysis. *Medical Image Analysis*, submitted on: March 11, 2014.

2. Md. Manjur Ahmed, A.S.N. Huda and NorAshidi Mat Isa. Recursively constructs output-context fuzzy system for infrared thermography based non-destructive characterization of electrical hotspots, *Engineering Applications of Artificial Intelligence*, submitted on March 11, 2014.
3. Md. Manjur Ahmed and NorAshidi Mat Isa. Evolving Output-context Fuzzy System for Effective Rule Base. *Expert Systems With Applications*, submitted on February, 18, 2014.

Conference Proceedings

1. K. Hasikin, N.A.M. Isa: "Enhancement of the Low Contrast Image Using Fuzzy Set Theory". In: Computer Modelling and Simulation (UKSim), 2012 UKSim 14th International Conference on, 28-30 March 2012, pp. 371-376
2. K. Hasikin, N.A.M. Isa: "Fuzzy enhancement for nonuniform illumination of microscopic Sprague Dawley rat sperm image". In: Medical Measurements and Applications Proceedings (MeMeA), 2012 IEEE International Symposium on, 18-19 May 2012, pp. 1-6

References:

- [1] M. A. Suckow, S. H. Weisbroth, and C. L. Franklin, *The Laboratory Rat*: Elsevier Science, 2005.
- [2] P. Flecknell, *Laboratory Animal Anaesthesia*: Elsevier Science, 2009.
- [3] J. Shiraishi, Q. Li, D. Appelbaum, and K. Doi, "Computer-Aided Diagnosis and Artificial Intelligence in Clinical Imaging," *Seminars in Nuclear Medicine*, vol. 41, pp. 449-462, 2011.
- [4] K. Hasikin and N. Mat Isa, "Adaptive Fuzzy Intensity Measure Enhancement Technique for Non-Uniform Illumination and Low Contrast Images " *Signal, Image and Video Processing*, vol. DOI. 10.1007/s11760-013-0596-1, 2013.
- [5] W. Zhou and A. C. Bovik, "A universal image quality index," *IEEE Signal Processing Letters*, vol. 9, pp. 81-84, 2002.
- [6] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 9, pp. 62-66, 1979.
- [7] B. N. Saha and N. Ray, "Image thresholding by variational minimax optimization," *Pattern Recognition*, vol. 42, pp. 843-856, 2009.
- [8] H. Yazid and H. Arof, "Gradient based adaptive thresholding," *Journal of Visual Communication and Image Representation*, vol. 24, pp. 926-936, 2013.
- [9] M. I. Heywood and P. D. Noakes, "Fractional central moment method for movement-invariant object classification," *Vision, Image and Signal Processing, IEE Proceedings -*, vol. 142, pp. 213-219, 1995.
- [10] M. J. Bottema, "Circularity of objects in images," in *Acoustics, Speech, and Signal Processing, 2000. ICASSP '00. Proceedings. 2000 IEEE International Conference on*, 2000, pp. 2247-2250 vol.4.

- [11] C. Beyan and A. Temizel, "Adaptive mean-shift for automated multi object tracking," *Computer Vision, IET*, vol. 6, pp. 1-12, 2012.
- [12] V. R. Nafisi, M. H. Moradi, and M. H. Nasr-Esfahani, "A template matching algorithm for sperm tracking and classification," *Physiol Meas*, vol. 26, pp. 639-51, 2005.
- [13] F. N. Rahatabad, M. H. Moradi, and V. R. Nafisi, "A Multi Steps Algorithm for Sperm Segmentation in Microscopic Image," *World Academy of Science, Engineering and Technology*, vol. 12, pp. 511-513, 2007.
- [14] S. Schäfer-Somi and C. Aurich, "Use of a new computer-assisted sperm analyzer for the assessment of motility and viability of dog spermatozoa and evaluation of four different semen extenders for predilution," *Animal Reproduction Science*, vol. 102, pp. 1-13, 2007.
- [15] M.J. Gacto, R. Alcalá, F. Herrera, Interpretability of linguistic fuzzy rule-based systems: An overview of interpretability measures, *Information Sciences* 181 (2011) 4340–4360.
- [16] D. Wang , X.-J. Zeng,, J. A. Keane, An evolving-construction scheme for fuzzy system, *IEEE Transactions on Fuzzy Systems* 18(2010) 755–770.
- [17] Andri Riid, Ennu Rüstern, Adaptability, interpretability and rule weights in fuzzy rule-based systems, *Information Sciences* (2011), <http://dx.doi.org/10.1016/j.ins.2012.12.048>
- [18] W.L. Tung, C. Quek, "eFSM – A novel online neural-fuzzy semantic memory model," *IEEE Transactions on Neural Networks*, 21 (1), 2010, pp. 136–157.
- [19] C. Mencar, C. Castiello, R Cannone, and A. M. Fanelli, "Design of fuzzy rule-based classifiers with semantic cointension," *Information Sciences*, 181, 2011, pp. 4361-4377.
- [20] <http://archive.ics.uci.edu/ml/datasets/Glass+Identification>, retrieved on 2/10/2013.
- [21] [http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+\(Original\)](http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Original)), retrieved on 2/10/2013.
- [22] W. Pedrycz, K.-C. Kwak, Linguistic models as a framework of user-centric system modeling, *IEEE Transactions on System, Man, and Cybernetics – Part A* 36(4) (2006) 727–745.
- [23] D. Wang , X.-J. Zeng, J. A. Keane, A simplified structure evolving method for Mamdani fuzzy system identification and its application to high-dimensional problems, *Information Sciences* 200(2013) 110–123.

APPENDIX



Adaptive fuzzy intensity measure enhancement technique for non-uniform illumination and low-contrast images

Khairunnisa Hasikin · Nor Ashidi Mat Isa

Received: 5 February 2013 / Revised: 4 December 2013 / Accepted: 4 December 2013
© Springer-Verlag London 2013

Abstract A new enhancement technique based on fuzzy intensity measure is proposed in this study to address problems in non-uniform illumination and low contrast often encountered in recorded images. The proposed algorithm, namely adaptive fuzzy intensity measure, is capable of selectively enhancing dark region without increasing illumination in bright region. A fuzzy intensity measure is calculated to determine the intensity distribution of the original image and distinguish between bright and dark regions. Image illumination is improved, whereas local contrast of the image is increased to ensure detail preservation. Implementation of the proposed technique on grayscale and color images with non-uniform illumination images shows that in most cases (i.e., except for processing time), the proposed technique is superior compared with other state-of-the-art techniques. The proposed technique produces images with homogeneous illumination. In addition, the proposed method is computationally fast (i.e., <1 s) and thus can be utilized in real-time applications.

Keywords Fuzzy enhancement · Fuzzy intensity measure · Non-uniform illumination image · Low contrast

1 Introduction

Advancements in image processing have enabled the analysis of digital images in most computer vision applications [1–4], video surveillance [5–7], and biomedical engineering [8–14]. Digital images are often low in quality and suffer from non-uniform illumination or brightness, loss of details, and poor contrast. These problems become critical when the foreground of interest is difficult to be distinguished from the background, which worsens the segmentation problem and allows false recognition and detection to occur.

The human visual system has far larger dynamic ranges than most commercial cameras and video cameras. These devices have limited dynamic ranges; thus, recorded images obtained from these devices are usually non-homogeneous and low in contrast. Improper lighting condition and external disturbances, which worsen the aforementioned problems, are inevitable during image acquisition.

In this respect, most of the images acquired through commercial cameras and video cameras exhibit problems in non-uniform illumination and low contrast. Although these images contain significant information, such information is not visible because the images suffer from lack of sharpness and are easily influenced by noise. Image enhancement plays an important role as a preprocessing task that can significantly improve image quality. The basic idea of image enhancement is to increase the contrast of the bright and dark regions in order to attain better image quality. The visual information of the image is increased for better interpretation and perception to provide a clear image to the eye or

This project is supported by the Ministry of Science, Technology & Innovation Malaysia through Sciencefund Grant entitled "Development of Computational Intelligent Infertility Detection System based on Sperm Motility Analysis".

K. Hasikin · N. A. Mat Isa (✉)
Imaging and Intelligent System Research Team (ISRT), School of Electrical and Electronic Engineering, Engineering Campus, Universiti Sains Malaysia, 14300 Nibong Tebal, Penang, Malaysia
e-mail: ashidi@eng.usm.my

K. Hasikin
Department of Biomedical Engineering, Faculty of Engineering, University of Malaya, 50603 Lembah Pantai, Kuala Lumpur, Malaysia
e-mail: khairunnisa@um.edu.my

assist in feature extraction processing in computer vision systems [15–18].

Various image enhancement algorithms have been proposed to enhance the degraded images in different applications. Image enhancement can be categorized into three broad types, namely transform, spatial, and fuzzy domains. The related studies on these three enhancement methods are discussed and presented in the succeeding section.

This paper is organized as follows. Related studies on image enhancement based on transform, spatial, and fuzzy domain approaches are elaborated in Sect. 2. Section 3 presents the proposed enhancement algorithm, and Sect. 4 explains the optimization procedure employed to obtain an optimum fuzzification factor. Sections 5 and 6 present the application of the proposed algorithm in color images and image analysis, respectively. The proposed algorithm is tested on non-uniform grayscale and color images in Sect. 7. The test images are compared in terms of visual representation and quantitative measures. Section 8 provides the conclusions of this paper based on the conducted analyses.

2 Related studies

The first method of image enhancement, namely the transform or frequency domain approach, is conducted by modifying the frequency transform of the image. Several enhancement techniques in the transform domain have been reported recently to solve the problem of non-uniform image illumination in face recognition and fingerprint enhancement applications [19–23]. In both applications, images normally exhibit non-uniform illumination; the details in the dark region of the images are less discernible. Enhancement is performed on the frequency transform of the image, and then the inverse transform is computed to obtain the resultant image. The intensities of the image are modified according to the transformation function [24,25].

Although enhancement in the frequency domain produces good results, the low- and high-frequency components in the image are not easily constructed. This is because, the intensity values for low-contrast and non-uniform illumination images are mostly vague and uncertain. As a result, spatial information of the intensity values is insufficient; thus, image representation based on frequency components is not easily constructed. Furthermore, images enhanced by frequency domain methods are normally compressed and result in the loss of valuable information and details. Computing a two-dimensional transform for images with different sizes is very time consuming even with fast transformation techniques; such procedure is not suitable for real-time processing [26].

The second class of image enhancement methods modifies pixels directly. Histogram equalization (HE) represents a prime example of an enhancement technique in the spa-

tial domain. Although HE is suitable for overall contrast enhancement, a few limitations exist. Enhancement by HE causes level saturation (i.e., clipping) effects as a result of pushing intensity values toward the left or right side of the histogram in HE [27]. Saturation effects not only degrade the appearance of the image but also lead to information loss. Furthermore, the excessive change in the brightness level induced through HE leads to the generation of annoying artifacts and unnatural appearance of the enhanced image.

Several brightness and detail-preserving modifications on HE techniques, which include adaptive HE techniques [28–35] as well as histogram specification [30,36,37], have been widely utilized to overcome these limitations in enhancing non-uniform illumination image. Adaptive methods provide better identification of different gray level regions through analysis of histogram in the local neighborhood window of every pixel. One example of modified HE approach is multi-histogram equalization technique [32,38]. In this approach, image histogram is partitioned into multiple segments based on its illumination. The bright and dark regions in each segment are equalized independently. The techniques involve remapping the peaks, which produces perceivable changes in mean image brightness.

Ibrahim and Kong [34] proposed brightness preserving dynamic histogram equalization (BPDHE) to address the peak remapping problem. BPDHE utilizes Gaussian smoothing kernel to smooth peak fluctuations. The valley regions are then segmented, and the dynamic equalization is then performed on each segmented histogram.

Histogram equalization (HE) has furthermore been used in the context of tone mapping (TM) [39] in order to enhance images with non-uniform illumination and low contrast. At first, global histogram adjustment is conducted based on the TM operator. Subsequently, the image is segmented, and adaptive contrast adjustment with the TM operator is performed to increase the local contrast of the image and produce high-quality images.

The retinex approach was introduced by Land [40] to address problem with degraded images that exhibit non-uniform illumination and uneven brightness. This approach compensates for non-uniform illumination by separating illumination from reflectance in the given image.

Enhancement of images with non-uniform illumination can also be possibly conducted through mathematical morphology operation of top hat transform. Top hat transform is a mathematical morphology approach that utilizes structural elements to extract multi-scale bright and dark regions. The image is enhanced by enlarging the extracted bright and dark regions [41].

Another approach that addresses the non-uniform illumination of the image has been proposed by Eschbach [42]. A new parameter “exposure” was introduced and altered by iteratively comparing image intensity with a pair of preset

thresholds of bright and dark regions. The image is processed until the threshold conditions are satisfied.

Although attempts have been made to enhance images by modifying every pixel in the spatial domain, vagueness in intensity values, which are caused by non-uniform lighting, have not been efficiently addressed. Therefore, a fuzzy enhancement technique is employed to overcome the aforementioned problem. Pixels are converted and modified in the fuzzy domain, which is the third category of image enhancement. The fuzzy system tool is adopted in image enhancement because this tool can mimic human reasoning and is beneficial in dealing with ambiguous situations that occur in non-uniform illumination image.

Fuzzy image enhancement was introduced as early as 1981 by Pal and King [43]. The smoothing algorithm of a linear non-recursive filter is employed. This filter acts as defocussing tool in which a part of the intensity of pixels is being distributed to their neighbor. The image is enhanced by optimizing objective parameters, namely index of fuzziness and entropy. Fuzzy set theory concept is widely adopted in image enhancement either globally, locally [44,45], or combined with other approaches such as fuzzy histogram adjustment.

Sheet et al. [46] incorporated fuzzy set theory in histogram modification of digital images, and its performance was compared with the BPDHE approach. This new approach exhibited improved performance compared with BPDHE because the former involves computations employing an appropriate fuzzy membership function. Thus, the imprecision of gray levels is handled well, and histograms appear smoother in the sense that they do not exhibit random fluctuations. The new approach helps obtain meaningful bright and dark regions for brightness preserving equalization.

The fuzzy concept has been adopted by a few researchers [26,47,48]. The "exposure" parameter is further exploited, and its role in fuzzy enhancement is improved. The exposure is calculated and clustered into overexposed and underexposed regions. Two different functions of the modified fuzzy triangular membership function and power-law transformation are utilized to specifically enhance the overexposed and underexposed regions.

The non-uniform illumination problem was further investigated and improved by Verma et al. [48]. The image was categorized into three regions namely, underexposed, overexposed, and mixed regions. Enhancement was performed on color image, which the luminance component was modified with specific functions according to the three aforementioned regions. In this approach, the quantitative measure of exposure is optimized through an iterative procedure to improve image quality [47,49–51]. However, this approach requires a complicated optimization process, which adds to the existing complication of the enhancement process in order to achieve good quality image.

Although numerous studies focus on the development of the enhancement algorithm either locally or globally, the enhancement process that produces images with optimum and best quality remain debatable. An optimally enhanced image refers to a well-illuminated image that with uniform brightness and detail preservation while existing noises are not enhanced.

A new approach in fuzzy enhancement is proposed in this study to address these problems and to efficiently enhance images with the non-uniform illumination and low contrast. The enhancement techniques proposed by the authors in [52,53] successfully enhanced images with non-uniform illumination. However, the details of the image are not well preserved, and significant features are not enhanced and not fully developed which caused significant decrement in clarity of the image. Therefore, the new fuzzy intensity measure proposed in this study involves computations that consider the mean and deviation of histogram intensity distribution. The threshold that distinguishes between dark and bright regions is then determined. The image is clustered into two regions using the fuzzy membership function. The image is enhanced separately in each region to obtain an image with better quality.

3 Proposed algorithm

The proposed algorithm for adaptive fuzzy intensity measure (AFIM) is presented in this section. Considering that image information is vague, the pixel values that constitute the images with non-uniform illumination (i.e., non-uniform intensity and brightness of the image) may not be precise; inherent imprecision is possibly embedded in the images. Determining whether the pixels should be made darker or brighter than their original intensity level during enhancement is difficult. Visual assessment by a human observer is subjective, and quantitative analysis of the image contrast does not represent well the improvement that has been made in the original image. This is because the image contrast is quantitatively calculated by measuring the deviation in the intensity values. This situation justifies the scenario of having high value of image contrast while in terms of qualitative evaluation, the image appears over-enhanced and unnatural. The quantitative measurement of the image contrast only calculates the deviation of the intensity values without considering whether the image is naturally enhanced or unnaturally enhanced. The proposed approach thus adopts the fuzzy approach which addresses vagueness and image uncertainty to enhance the image. The process is performed by associating a degree of belonging to a particular cluster in the fuzzy membership function.

Fuzzy image enhancement has three main stages, namely image fuzzification, modification of membership value for

Fig. 1 Fuzzy image enhancement

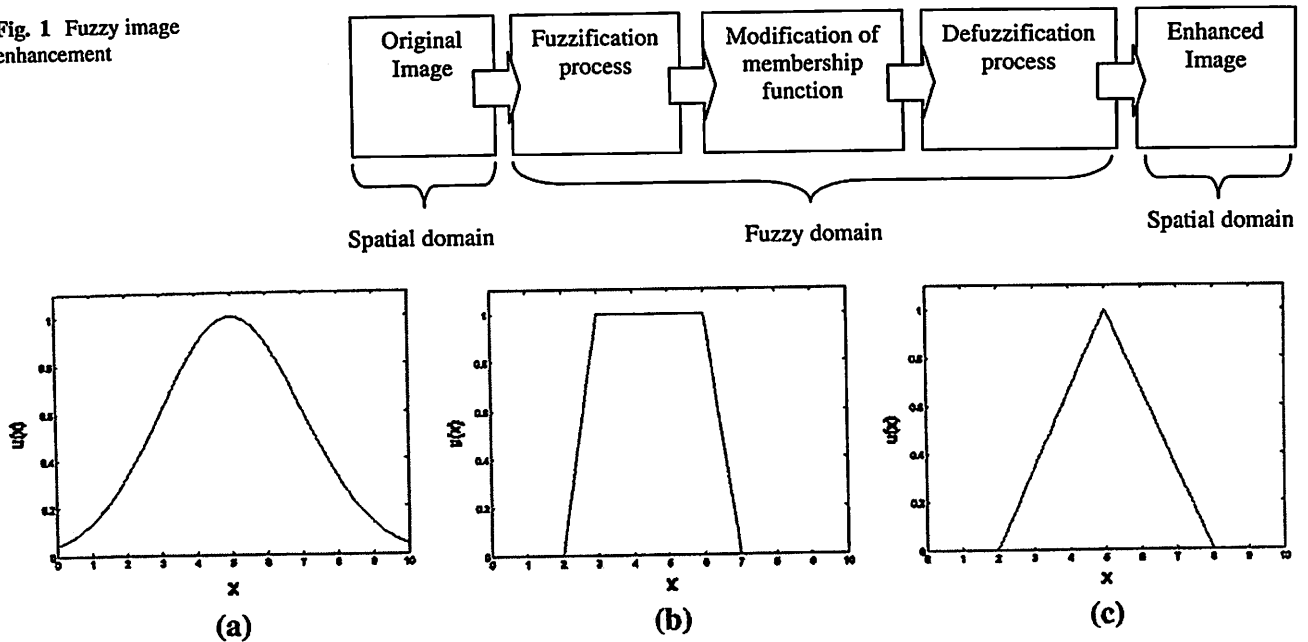


Fig. 2 a Gaussian membership function, b trapezoidal membership function, c triangular membership function

enhancement process, and image defuzzification (Fig. 1). The intensity levels (i.e., pixels values) are converted from spatial to fuzzy domain in the image fuzzification process. Each pixel is assigned either to the dark or bright regions based on a predetermined threshold. The membership values of each pixel are computed.

We consider an image with non-uniform illumination of size $R \times C$ denoted as A with intensity level m at pixel position (i, j) in the range of $[0, L - 1]$ in the image fuzzification stage. R and C are the number of rows and columns in the image, respectively. L is the total number of gray levels in the image. $\mu(m)$ denotes the membership value of the pixels of image A . $\mu(m)$ is calculated for every pixel, and in this case, the $\mu(m)$ is calculated globally to enhance the original image.

For the purpose of fuzzification, the intensity distributions in both regions (i.e., dark and bright regions) are assumed to be Gaussian. This means that the intensity distribution of the image is uniformly distributed in Gaussian shape which the most intensity values are accumulated in the middle of the histogram distribution (i.e., middle region of intensity values). This is because, in the low-contrast and non-uniform illumination images, most of the intensity values are mainly concentrated in the middle of the histogram distribution. This can be observed in Fig. 2a where the histogram has high amplitude at the middle region of the intensity values.

Therefore, a modified Gaussian membership function is utilized to determine the membership values of the pixels in the image that lies in the range $[0, 1]$. The Gaussian member-

ship function is selected in this study because even though separate functions are utilized to enhance the bright and dark regions, smooth transition is required to enhance both regions. The Gaussian membership function with continuous differentiable curves is selected. Other membership functions such as triangular or trapezoidal membership functions do not possess such abilities (Fig. 2).

A certain region in the image with non-uniform illumination appears darker or brighter than the other regions in the image. Thus, a parameter called fuzzy intensity measure is introduced. This parameter considers the mean and deviation of histogram intensity distribution, which is provided by Eqs. (1)–(3). These equations are calculated to determine the non-homogeneous intensity distribution of the image. The calculated fuzzy intensity measure is then utilized to determine a threshold T , which clusters the image into bright and dark regions based on Eq. (4). The dark region is clustered in the range of $[0, T - 1]$, whereas the bright region is clustered in the range of $[T, L - 1]$.

$$g_a = \frac{\sum_{m=0}^{L-1} m \times p(m)}{\sum_{m=0}^{L-1} p(m)} \quad (1)$$

$$g_d = \left[\frac{\sum_{m=0}^{L-1} [(m - g_a)^2 p(m)]}{\sum_{m=0}^{L-1} p(m)} \right]^{1/2} \quad (2)$$

$$\text{fuzzy intensity measure} = \frac{g_d}{g_a} \quad (3)$$

$$T = L \left[\frac{g_d}{g_a} \right] \quad (4)$$

where m is the intensity of the pixel at position (i, j) and $p(m)$ represents the number of pixels in the histogram of the entire image. g_d and g_a are deviation and mean intensity distributions, respectively.

After the image is divided into two regions (dark and bright regions) based on the value of T , fuzzification is performed in each region separately. The modified Gaussian membership function is utilized for the fuzzification of the dark region as follows:

$$\mu_d(m) = \exp \left[-\frac{(m_{\max} - (m_{\text{avg}} - m))^2}{\zeta_d} \right] \text{ for } m < T \tag{5}$$

where $\mu_d(m)$ is the membership function in the dark region and m is the intensity value in the dark region in the range of $[0 T - 1]$. m_{avg} and m_{\max} are the average intensity and maximum intensity of the image, respectively. ζ_d the fuzzifier function of the dark region, is provided by:

$$\zeta_d = \alpha \frac{\sum_{m=0}^{L-1} [(m_d - m_{d\text{avg}}) - \sigma_m]^4 p(m_d)}{\sum_{m=0}^{L-1} [(m_d - m_{d\text{avg}}) - \sigma_m]^2 p(m_d)} \tag{6}$$

where σ_m is the standard deviation of intensity of the entire image, $m_{d\text{avg}}$ is the average intensity of the dark pixels, and m_d and $p(m_d)$ are the intensities and histogram of the dark region, respectively.

The mirror function of the aforementioned Gaussian membership function is utilized to fuzzify the bright region of the image for $m \geq T$ as follows:

$$\mu_b(m) = \exp \left[-\frac{(m_{\max} - (m_{\text{avg}} - (L - m)))^2}{\zeta_b} \right] \text{ for } m \geq T \tag{7}$$

where $\mu_b(m)$ is the membership function of bright region. ζ_b is the fuzzifier function in the bright region.

$$\zeta_b = \alpha \frac{\sum_{m=0}^{L-1} [(m_b - m_{b\text{avg}}) - \sigma_m]^4 p(m_b)}{\sum_{m=0}^{L-1} [(m_b - m_{b\text{avg}}) - \sigma_m]^2 p(m_b)} \tag{8}$$

where $m_{b\text{avg}}$ is the average intensity of the bright pixels, m_b is the intensity of the bright region, and $p(m_b)$ is the histogram of the bright pixels.

The fuzzifier functions of ζ_d and ζ_b calculate the intensity deviation in the dark and bright regions, respectively. α is the fuzzification factor that depends on the intensity values of the input image. The selection of α will be explained in details in the succeeding section.

Once fuzzification is completed, the original input pixels that exhibit non-uniform illumination and low contrast are transformed into Gaussian distributed pixels. The local contrast of the image is based on intensity difference in a small region, and it is computed to preserve the details of the image. Local contrasts are defined for the dark and bright regions as:

$$C_{L_d}(i, j) = \sum_{(i,j) \in W_{i,j}} [\max(\mu_d(i, j)) - \min(\mu_d(i, j))] \tag{9}$$

$$C_{L_b}(i, j) = \sum_{(i,j) \in W_{i,j}} [\max(\mu_b(i, j)) - \min(\mu_b(i, j))] \tag{10}$$

where $\mu_d(i, j)$ and $\mu_b(i, j)$ represents the 3×3 local fuzzified image (i.e., output image obtained after fuzzification process) of μ_d and μ_b , respectively, which are centered at position (i, j) . $\max(\mu_d(i, j))$ and $\max(\mu_b(i, j))$ represent the maximum gray level values of the local fuzzified image for dark and bright regions, respectively. $\min(\mu_d(i, j))$ and $\min(\mu_b(i, j))$ denote the minimum gray level values of the local fuzzified image for dark and bright regions, respectively.

Modification of the fuzzified image is performed once the aforementioned steps are executed. Modification is performed to enhance the fuzzified image based on the dark and bright regions, which include the local contrast of the image as shown in Eqs. (11) and (12), respectively.

$$\mu'_d(m) = \frac{1}{1 + e^{-C_{L_d}[\mu_d(m) - m_{d\text{avg}}]}} \text{ for } m < T \tag{11}$$

$$\mu'_b(m) = \frac{1}{1 + e^{-C_{L_b}[\mu_b(m) - m_{b\text{avg}}]}} \text{ for } m \geq T \tag{12}$$

where μ'_d and μ'_b are the modified membership functions in the dark and bright regions, respectively. C_{L_d} and C_{L_b} are the local contrast of dark and bright regions, respectively, which are computed to preserve the details in the image.

The above functions modify the original membership functions of $\mu_d(m)$ and $\mu_b(m)$. The modified functions are then defuzzified with the respective inverse membership functions as shown in Eq. (13). Both regions are combined to obtain the enhanced image. The pixels in the dark region are scaled back to the range $[0 T - 1]$, whereas the bright region is translated and scaled back to the region $[TL - 1]$.

$$M = \begin{cases} \mu_d'^{-1}(m) & \forall m \leq T \\ \mu_b'^{-1}(m) & \forall m > T \end{cases} \tag{13}$$

where M is the enhanced image obtained from the defuzzification process.

4 Optimization of fuzzification factor

The fuzzification factor differs with different input images as discussed in the previous section. As a result, the optimum parameter value of α must be selected to obtain a pleasant image. Results obtained from simulation on 300 images with non-uniform illumination consisting of 150 grayscale images

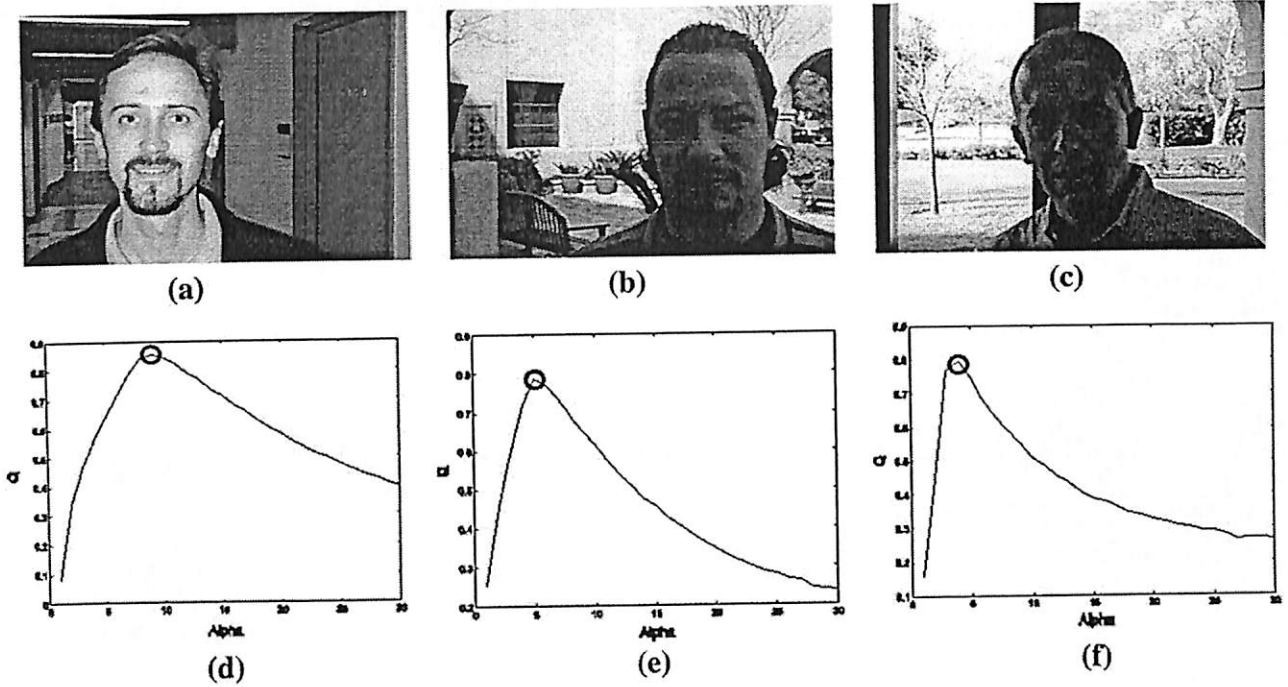


Fig. 3 a–c images with non-uniform illumination, d–f optimization graphs for images with non-uniform illumination (a–c), respectively

and 150 color images show that the optimum value of α is set to the parameter value that yields the maximum image quality index (Q). Q is computed by modeling any image distortion as a combination of three factors, namely loss of correlation, luminance, and contrast distortions as shown in Eq. (14).

The original and enhanced images are assumed to contain $m = \{m_y | y = 1, 2, \dots, Z\}$ and $M = \{M_y | y = 1, 2, \dots, Z\}$, respectively. m_y and M_y are the intensity levels of the original and enhanced images, respectively. The best value of '1' is achieved if and only if $m_y = M_y$. Q is defined as [54]:

$$Q_d = \frac{\alpha_{mM}}{\alpha_m \alpha_M} \cdot \frac{2\bar{m}(\bar{M})}{(\bar{m})^2 + (\bar{M})^2} \cdot \frac{2\tau_m \tau_M}{\tau_m^2 + \tau_M^2} \quad (14)$$

where

$$\bar{m} = \frac{1}{Z} \sum_{y=1}^Z m_y \quad (15)$$

$$\bar{M} = \frac{1}{Z} \sum_{y=1}^Z M_y \quad (16)$$

$$\tau_m^2 = \frac{1}{Z-1} \sum_{y=1}^Z (m_y - \bar{m})^2 \quad (17)$$

$$\tau_M^2 = \frac{1}{Z-1} \sum_{y=1}^Z (M_y - \bar{M})^2 \quad (18)$$

$$\alpha_{mM} = \frac{1}{Z-1} \sum_{y=1}^Z (m_y - \bar{m})(M_y - \bar{M}) \quad (19)$$

Figure 3 shows three non-uniform illumination in which the illumination and intensity distribution of these images are non-homogeneous. The plots of Q in Fig. 3a–c illustrate the changes in Q as fuzzification factor, α varies from 0 to 30. Automated tuning is conducted until a homogeneous image is obtained. The homogeneous image is attained when Q reaches its maximum value. Figure 3 shows that Q reaches its highest value when alpha is 8, 5, and 4 as circled in Fig. 3d–f, respectively.

The optimal procedure for selecting α is described as follows. For a given input image (i.e., original image), the value of α is varied from a minimum of 1 to a maximum of 30. For each value of α , the following automated tuning procedures are performed:

1. Apply the algorithm presented in Sect. 2 to generate an enhanced image
2. Calculate Q with Eq. (14)
3. Select the parameter value that produces the maximum Q as the optimum value of α , after the two aforementioned steps.

The enhanced image is generated by adopting optimum α according to the enhancement process in Sect. 2 to produce the final output. Simulations are performed on 300 test images with non-uniform illumination to validate the automatic selection of α . Examples of the automatic selection of α are presented in Fig. 3.

5 Application in color images

The aforementioned algorithm can also be applied to color images by modifying gray level values. Enhancement for color images is conducted by converting Red, Green, and Blue (RGB) color space into Hue, Saturation, and Intensity (HSI) color space. This conversion is performed because direct enhancement in RGB may produce color artifacts. HSI color space is able to separate chromatic from achromatic information, thus ensuring that the original color of the image is not distorted.

Enhancement is performed by preserving the hue component (H) and transforming the intensity component (I) based on the algorithm presented in Sect. 2. The saturation (S) component is modified with a power transformation function as shown in Eqs. (20)–(22).

$$CF(m) = \frac{\sum_{(m) \in W_{i,j}} (m_{i,j} - \bar{m}_{W_{i,j}})^2}{\sum_{(m) \in W_{i,j}} (m_{i,j} - \sigma_{m_{W_{i,j}}})^2} \quad (20)$$

$$S'_d(m) = \tau_d [S_d(m)]^{(1-CF)} \quad (21)$$

$$S'_b(m) = \tau_b [S_b(m)]^{(1-CF)} \quad (22)$$

where $\bar{m}_{W_{i,j}}$ is the local average gray level value in $W_{i,j}$ window and $\sigma_{m_{W_{i,j}}}$ is the local intensity deviation in the $W_{i,j}$ window. CF is the contrast factor that is calculated to enhance the local contrast of the image. $S'_d(m)$ and $S'_b(m)$ are the modified saturation values of the dark and bright regions in the HSV color space, respectively. $S_d(m)$ and $S_b(m)$ are the corresponding saturation components for the dark and bright regions, respectively. τ_d and τ_b are the saturation intensifier and de-intensifier selected experimentally.

In order to ensure the contrast and details in the local neighborhood window are enhanced, the S component is modified. The modification of S component is conducted by considering the local average gray level value and local intensity deviation as shown in Eqs. (20)–(22). The HSI color space is converted back to RGB color space after the S and I components are adjusted to enhance the image.

6 Image analysis

The quantitative measures for image analysis are presented in this section. Image quality measurement is an impor-

tant research area. Establishing a correct and effective measure to quantify the quality of the enhanced images is difficult. The proposed algorithm as a new enhancement technique is expected to significantly improve the quality of the image while preserving the details. The dark pixels should be enhanced, and noises should not be amplified.

The performances of the proposed algorithm are evaluated and compared based on four quantitative measures, namely the image quality index (Q), contrast (C), clarity index (PL) [41], and computational time (t).

Q is computed with Eq. (14) as discussed in Sect. 3. The image quality index, called color fidelity metric, Q_{color} proposed by [55], is utilized for color images to observe quality improvement during enhancement. The enhanced image, which is in RGB, is transformed to LAB color space. Q_{color} is defined as:

$$Q_{color} = \sqrt{w_l (Q_l)^2 + w_\alpha (Q_\alpha)^2 + w_\beta (Q_\beta)^2} \quad (23)$$

where Q_l , Q_α , and Q_β represent the fidelity factors of l , α , and β channels, respectively. w_l , w_α , and w_β are the corresponding weights attributed to the perceived distortions in each of these channels.

As an addendum to the computed Q , C is employed as the contrast enhancement measurement of the sample images. Large C indicates that the enhancement technique successfully attained appropriate contrast. C is calculated with Eq. (24).

$$C = \sqrt{\sum_{y=0}^{L-1} (M_y - \bar{M}) \times p(M_y)} \quad (24)$$

where M_y , \bar{M} , and $p(M_y)$ are the intensity of the enhanced image, mean intensity of the enhanced image, and histogram of the enhanced image, respectively.

The measure of PL [41] is calculated to measure both noise and clarity in the image. PL is computed by considering the peak signal-to-noise ratio ($PSNR$) and index of fuzziness (γ) in the image, PL is defined as:

$$PL = \frac{PSNR}{\gamma} \quad (25)$$

A large value of $PSNR$ indicates that the corresponding algorithm enhances the image appropriately and produces minimal noise. γ is employed in the analysis because γ is commonly utilized to measure the clarity of the enhanced images. A small value of γ indicates that the enhanced result is clear and that enhancement of the corresponding algorithm produces a good quality image. Dividing the $PSNR$ and γ generates a measure that includes noise condition and image clarity. A large value of PL indicates that the enhanced image contains minimal noise and that the clarity of the image is increased. $PSNR$ and γ are calculated with Eqs. (26) and (28), respectively.

$$PSNR = 10 \log_{10} (L - 1)^2 / MSE \quad (26)$$

where

$$MSE = \sqrt{\frac{\sum_{i=1}^M \sum_{j=1}^N (I_{i,j} - K_{i,j})^2}{M \times N}} \quad (27)$$

$$\gamma = \frac{2}{RC} \sum_{M=0}^{L-1} \min [k_m(1 - k_m)] \quad (28)$$

$$k_m = \sin \left[\frac{\pi}{2} \times \left(1 - \frac{M}{M_{\max}} \right) \right] \quad (29)$$

where M and M_{\max} are the intensity and maximum intensity of the enhanced image, respectively.

Computational time (t) is investigated to measure the computational complexity of the enhancement algorithm. t is defined as the total time required to completely process the input image. It changes dynamically depending on the size of the image which is closely related to the total number of pixels of the image.

7 Results and discussions

The performance of the proposed enhancement technique is presented in this section. Quantitative and qualitative results obtained from the proposed technique are also compared with other fuzzy techniques such as BPDFHE [46] and fuzzy quantitative measure (FQM) presented in [48].

Brightness preserving dynamic histogram equalization (BPDHE) is utilized for comparison in this section because this method considers the crisp histogram of the image, which is beneficial for the enhancement process. FQM is also utilized for comparison because it is related to the proposed method, which computes the quantitative measures of gray levels to enhance the image.

The proposed enhancement technique is also compared with three other non-fuzzy techniques. The techniques include TM presented in [39] which involves the enhancement of non-uniform illumination in high dynamic range image. Discrete cosine transform (DCT), which is conducted in the frequency domain [56], is also included in the analysis. Gamma correction (GC) approach¹ [24] is likewise adopted for comparison.

The experimental results of this study are obtained by implementing and processing the degraded images with Matlab R2010a and Intel(R) Core(TM) i3 2100 3.10GHz and 4 GB RAM. The degraded images utilized for comparative analysis include standard images with non-uniform illumina-

tion (size of 400×264), that are obtained from the California Institute of Technology Computational Vision Database.

Subjective appearance evaluation is performed on several grayscale images with non-uniform illumination as shown in Figs. 4, 5, and 6. Comparative analysis includes observing whether the techniques are able to enhance an entire image without over-enhancing or under-enhancing certain regions in the image. The details of the images are observed to ensure that no information loss occurs in the enhanced image. The quantitative results from each image are also presented in Figs. 4, 5, and 6. The original images presented in Figs. 4, 5, and 6 are having non-uniform illumination. These images suffer from uneven lighting, where the dark regions accumulate on Man 1 and Man 2's faces in Figs. 4 and 5. Meanwhile, the foreground (i.e., Man 3's face) in Fig. 6 appears brighter than the background.

Figure 4 shows that most of the enhancement techniques are able to enhance the image and significantly improve overall brightness/illumination of the image. The performance of the enhancement techniques can be analyzed by observing the brightness of foreground (i.e., big rectangular area) and background (i.e., small rectangular area). TM technique attains the best-enhanced and best-illuminated foreground in which the Man 1's face (Fig. 4f), appear the brightest as compared to the other techniques. However, over-enhancement is apparent in the background of the images enhanced by this method. This is due to the image pixels in the background are clipped to white; thus, the details of the image are not preserved. Similar situations are observed in the enhanced images produced by FQM (Fig. 4c) and GC (Fig. 4g). Although these methods improve the overall brightness of the image, details of the images are loss during the enhancement process. This scenario occurs because although specific functions are utilized and enhancement is conducted separately for bright and dark regions, the local contrast of the image is not considered, resulting in loss of details.

Discrete cosine transform (DCT) and BPDFHE improve image illumination while maintaining the background details. However, both techniques produce dark regions on Man 1's face. BPDFHE amplifies the existing noise during enhancement as exhibited by the lowest value of PL (i.e., 22.63). This finding indicates that BPDFHE does not reduce fuzziness in the original image and the existing noise is de-attenuated.

The enhanced image produced by the proposed AFIM method (Fig. 4b) exhibits appropriate contrast; the edges of the wall and tree in the background are clear. Furthermore, the edges on Man 1's face are clear and smooth as compared to its original image. As discussed in Sects. 3 and 6, Q is computed by considering the luminance (i.e., illumination) and contrast distortion as well as loss of correlation between the original and enhanced images. Therefore, the quantitative result of the proposed method attains the highest Q of 0.97, implies that overall image quality is improved without caus-

¹ In the GC approach, the value of gamma is chosen based on the optimization procedure as presented in Sect. 4. However, for this approach, the gamma values are incremented from 0.1 to 1.0 and gamma value that produces the maximum Q is chosen as the optimum value of gamma.

Fig. 4 Comparison of the enhancement results. **a** Original Man 1 image, **b** the proposed AFIM method, **c** FQM [48], **d** DCT [56], **e** BPDFHE [46], **f** TM [39], **g** GC ($\gamma=0.6$) [24]



ing any saturation effect. Local contrast is enhanced, whereas noise is suppressed.

Although TM technique produces the brightest image (Fig. 4f), unnecessary enhancement is performed to the existing bright region (i.e., background), thus causing over-

brightness at the background. Meanwhile, the proposed AFIM method provides more suitable approach in which the illumination of the image can be performed using fuzzifier and membership functions presented in Eqs. (5) and (6) for dark region as well as Eqs. (7) and (8) for bright region. In

Fig. 5 Comparison of the enhancement results **a** Original Man 2 image, **b** the proposed AFIM method, **c** FQM [48], **d** DCT [56], **e** BPDFHE [46], **f** TM [39], **g** GC ($\gamma=0.8$) [24]



addition, the local contrast parameter in Eqs. (9) and (10) can be modified to preserve the image details.

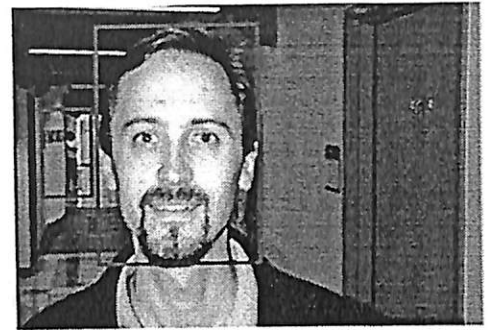
Other examples of non-uniform illumination are presented in Fig. 5. The foreground of the image (i.e., Man 2's face) in Fig. 5 is dark, whereas the background of the image is dom-

inated by the bright region. The brightness on Man 2's face (i.e., dark region) is increased with the function presented in Eq. (5). The proposed method successfully enhances the image, preserves the details of the image, and enhances local contrast as shown in Eq. (9). Figure 5b illustrates the details

Fig. 6 Comparison of the enhancement results **a** Original Man 3 image, **b** the proposed AFIM method, **c** FQM [48], **d** DCT [56], **e** BPDFHE [46], **f** TM [39], **g** GC ($\gamma=0.4$) [24]



(a) $C = 59.21$



(b) $C = 65.48$ $Q=0.95$ $PL=49.85$ $t=0.80s$



(c) $C = 51.41$ $Q=0.65$ $PL=42.61$ $t=1.76s$



(d) $C=58.06$ $Q=0.69$ $PL=36.36$ $t=6.29s$



(e) $C=59.62$ $Q=0.78$ $PL=47.12$ $t=0.26s$



(f) $C = 46.24$ $Q=0.79$ $PL=28.54$ $t=0.17s$



(g) $C = 53.38$ $Q=0.89$ $PL=32.28$ $t=0.13s$

of the tree in the background are enhanced without causing any saturation. The enhanced image produced by the proposed method exhibits an increase in image contrast. Thus, the produced image looks natural.

Fuzzy quantitative measure (FQM) causes saturation in the background. The foreground is darker than the original image, causing the enhanced image to appear unnatural. FQM has the lowest value of Q , which is 0.55. Images

Table 1 Average quantitative analysis of 150 grayscale images with non-uniform illumination

	Proposed method	FQM	DCT	BPDFHE	TM	Gamma
Q	0.97 ± 0.03	0.74 ± 0.14	0.69 ± 0.01	0.68 ± 0.03	0.84 ± 0.05	0.72 ± 0.52
C	69.71 ± 7.78	52.78 ± 3.41	59.34 ± 10.50	72.09 ± 8.84	56.03 ± 11.88	65.49 ± 5.44
PL	75.54 ± 4.18	49.39 ± 17.38	42.38 ± 35.21	44.04 ± 24.31	34.04 ± 12.93	65.86 ± 9.63
$t(s)$	0.93 ± 0.04	0.85 ± 0.16	2.40 ± 0.36	0.21 ± 0.05	0.40 ± 0.04	0.14 ± 0.01

Values tabulated in this table are the average \pm standard deviation values

enhanced by TM and GC over-enhance the existing bright regions (i.e., the background of the image) which causes loss of details at the background. DCT and BPDFHE are able to improve image illumination; however, the foreground is darker, and the edges are less smooth compared with images enhanced by the proposed method.

Other images with non-uniform illumination image are presented in Fig. 6. In contrast to Figs. 4 and 5, the foreground (i.e., Man 3's face) in this figure is brighter than the background. The TM operator over-enhances the existing bright region on Man 3's face. The same effect also occurs in FQM where illumination of the enhanced image is uneven and non-homogeneous. Unwanted intensity saturations are avoided in the proposed method, DCT, and BPDFHE.

Figure 6 also shows that the proposed AFIM method yields the best enhancement result with smooth edges and details as shown in the small rectangle in the figure. Image illumination is enhanced with Eqs. (11) and (12) as refer to its dark and bright regions, respectively. The PL value of the proposed algorithm is bigger than other algorithms, which is 49.85. This result verifies that the proposed AFIM algorithm enhances non-uniform illumination while suppressing existing noise. In addition, the proposed AFIM algorithm improves image quality as exhibited by the highest Q and C values of 0.95 and 65.48, respectively.

Apart from the grayscale images presented in Figs. 4, 5, and 6, the enhanced images produced by the proposed method in comparison with other techniques are presented in Table 3, "Appendix 1". Twenty supplementary images are illustrated, and their respective quantitative analysis is tabulated in Tables 5, 6, 7, and 8, "Appendix 2". In terms of the overall performances, the proposed AFIM method outperforms other enhancement techniques by producing most images with either the highest or second highest C and the highest Q . The capability of the proposed AFIM method to consistently produced high PL values indicates its advantage in producing image with improved clarity and minimal noises. In addition, the proposed AFIM method requires <0.5 s (in most cases) to compute which is comparable with other enhancement techniques.

The performance of the proposed algorithm in enhancing the grayscale images is evaluated by quantitative analysis of 150 grayscale non-uniform illumination images as tabulated

in Table 1. Comparison is performed based on the average and standard deviation values of Q , C , PL , and processing time, t for 150 grayscale images with non-uniform illumination. The best results are presented in bold for each analysis. The error bars of the quantitative analysis of grayscale images presented in Table 1 are also plotted in Figs. 10, 11, 12, and 13 in "Appendix 3". The error bars give better representation of the quantitative analysis presented in Table 1. Each quantitative analysis (i.e., C , Q , PL and t) is plotted in the error bars and compared with other enhancement techniques. Figures 10, 11, and 12 show that the proposed AFIM method attains the highest Q and PL with the lowest standard deviation. Meanwhile the proposed AFIM method obtains second highest rank in terms of C . As elaborated in Sect. 6, high C and Q indicate that the image is successfully enhanced with better contrast and quality of image is improved. In addition, high PL represents that the enhancement technique is capable to enhance the image without enhancing the existing noise.

The comparison presented in Table 1 indicates that in terms of the average and standard deviation values, the proposed AFIM method attains the highest value of Q and PL and outperforms other methods. Meanwhile, the proposed method is able to preserve the details by considering the local contrast of the dark and bright regions in Eqs. (9) and (10), respectively. Image information is retained, and overall contrast is improved. The proposed method achieved the second highest overall contrast after BPDFHE. The proposed AFIM method has successfully reduced the fuzziness of the image, resulting in the highest value of PL as indicated by Eq. (25). Furthermore, the proposed AFIM method involves less complex computations and is comparably easier (i.e., ranked fourth) to execute compared with other enhancement techniques.

Comparison of enhancement performance is also conducted for color images (Figs. 7, 8, and 9) to examine the enhancement effect produced by the proposed method. Figure 7a shows the original image with non-uniform illumination where the foreground (i.e., Man 4's face) is darker than the background. The details of the images are lost, and the intensities in the dark regions are decreased, resulting in the dim region surrounding the man's face. Enhancement process is conducted to improve the quality of the original

Fig. 7 Comparison of the enhancement results a Original Man 4 image, b the proposed AFIM method, c FQM [48], d DCT [56], e BPDFHE [46], f TM [39], g GC ($\gamma=0.3$) [24]

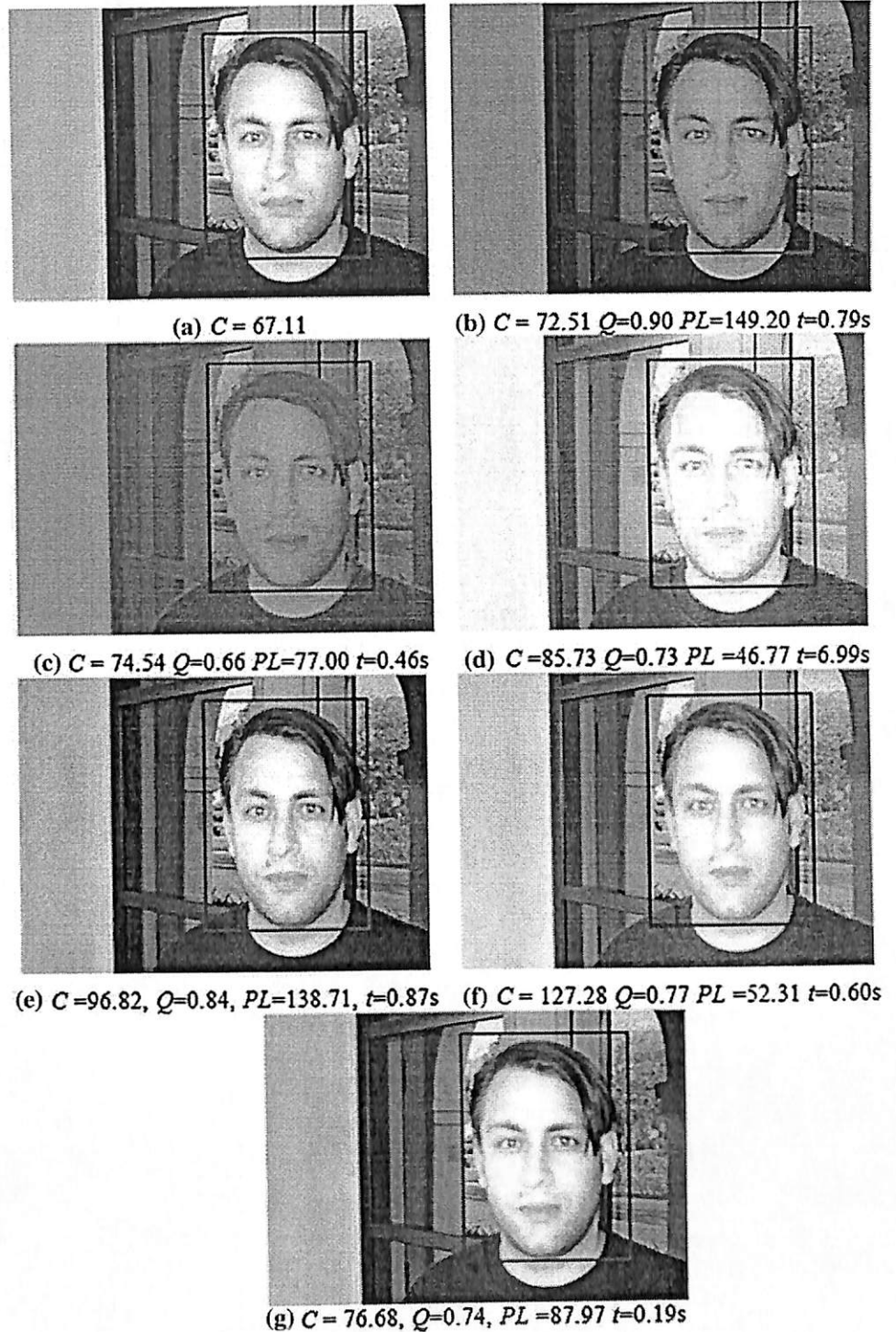


image and to ensure that the processed image exhibits homogeneous illumination.

The proposed AFIM method increases illumination in the dark region (i.e., around the man's face) without causing any intensity saturation as compared to the FQM. The image

enhanced by the proposed method obtained better illumination and improved image details without causing any color saturation. The intensifier and de-intensifier as presented in Eqs. (21) and (22) are utilized in the color enhancement process to ensure that the enhanced image appears natural

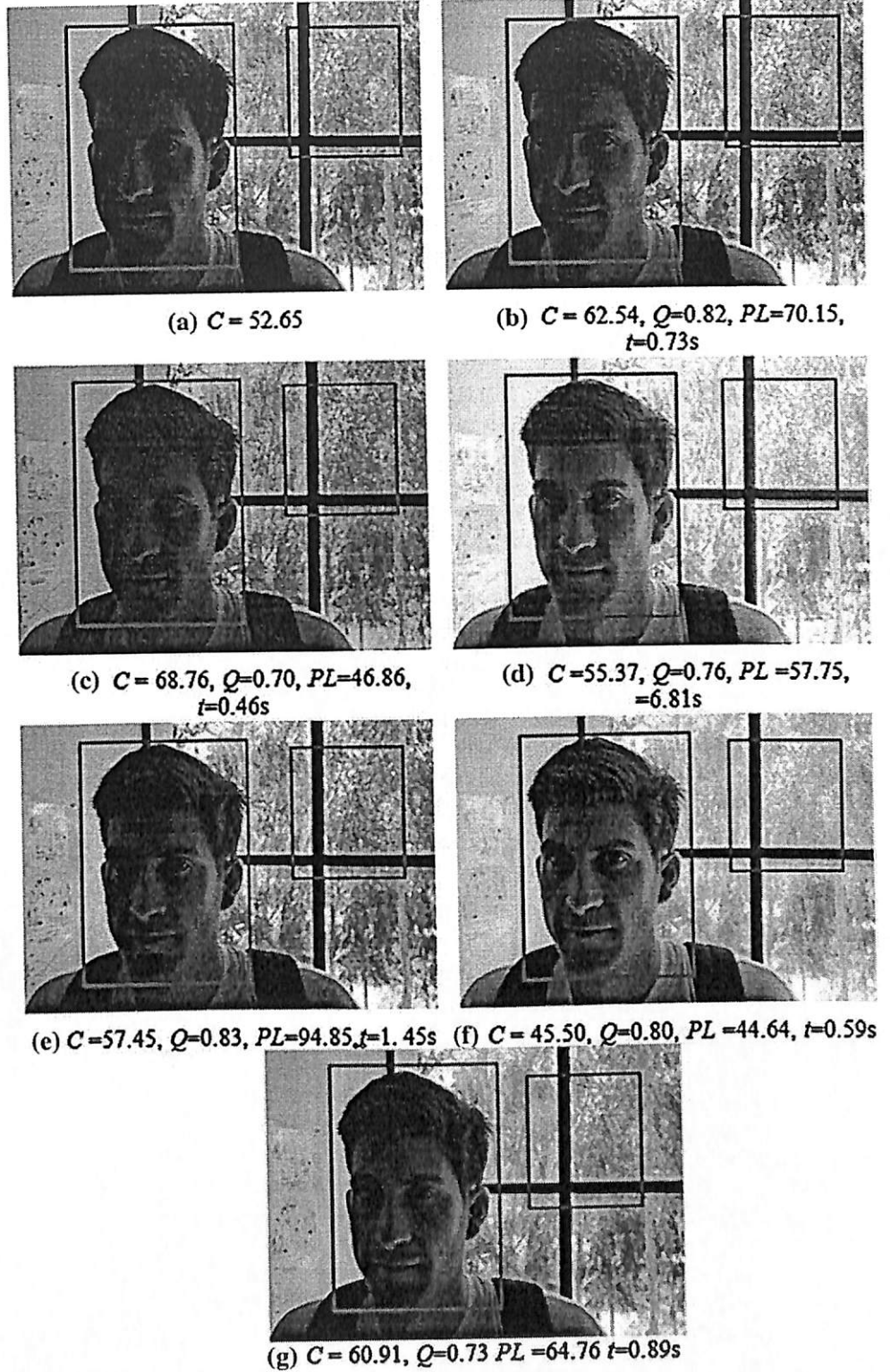
Fig. 8 Comparison of the enhancement results **a** Original Man 5 image, **b** the proposed AFIM method, **c** FQM [48], **d** DCT [56], **e** BPDFHE [46], **f** TM [39], **g** GC ($\gamma=0.8$) [24]



with better homogeneous illumination. In addition, the proposed AFIM method produces enhanced image with the highest Q , C and PL of 0.78, 153.39, and 118.68, respectively. The proposed algorithm outperforms the other enhancement algorithms, obtains uniform illumination image with less noise, and increases image clarity.

Tone mapping (TM) produces a natural-looking enhanced image, similar and comparable to the image enhanced by the proposed AFIM method. However, the TM operator over-enhanced the existing bright region in the image, thereby producing an over-bright region at the background of the image. BPDFHE is able to improve overall image illumina-

Fig. 9 Comparison of the enhancement results **a** Original Man 6 image, **b** the proposed AFIM method, **c** FQM [48], **d** DCT [56], **e** BPDFHE [46], **f** TM [39], **g** GC ($\gamma=0.6$) [24]



tion; however, the foreground is darker than the rest of the regions in the image. Contrast is also increased which attains the second highest of C which is 115.06.

Another example of non-uniform illumination is depicted in Fig. 8. The bright regions are mainly distributed in the foreground (i.e., around Man 5's face), and the background

Table 2 Average quantitative analysis of 150 color images with non-uniform illumination

	Proposed method	FQM	DCT	BPDFHE	TM	Gamma
Q	0.88 ± 0.09	0.74 ± 0.14	0.69 ± 0.11	0.68 ± 0.09	0.84 ± 0.11	0.74 ± 0.10
C	72.31 ± 11.94	52.78 ± 40.85	59.34 ± 15.69	72.09 ± 26.79	56.03 ± 25.08	63.98 ± 30.21
PL	66.95 ± 23.01	49.39 ± 14.75	42.38 ± 20.24	44.04 ± 30.88	34.04 ± 15.17	63.08 ± 44.21
$t(s)$	0.85 ± 0.03	1.85 ± 0.11	6.30 ± 0.58	0.20 ± 0.24	0.21 ± 0.15	0.12 ± 0.02

Values tabulated in this table are the average ± standard deviation values

of the image is dark. Although this figure clearly shows that most of the methods enhance the original image, the problems regarding over-enhancement still occur in the existing bright regions of images produced by DCT, BPDFHE, TM, and GC. Although FQM de-enhanced the bright region, accumulated on Man 5's face, the enhanced image appears unnatural and image appears blurred than the original image presented in Fig. 8a. This finding is supported by the fact that FQM has the lowest value of Q which is only 0.66. The proposed method on the other hand has a Q value of 0.90.

Figure 9 presents other example of images with non-uniform illumination. Visual appearance in this figure indicated that the proposed AFIM method produces a better-enhanced image where the intensity values in the bright region are decreased accordingly and important features are maintained. The saturation intensifier and de-intensifier are applied accordingly with Eqs. (21) and (22) for the dark and bright regions, respectively. A natural-looking enhanced image is attained as supported by the second highest value of Q , which is 0.82.

Discrete cosine transform (DCT), FQM, and GC approaches enhance the existing bright regions in the image background, causing the background to be over-enhanced and blurred. The edges and details of the image appear deteriorated resulting in a lower value of Q and C compared with the enhanced image obtained through the proposed method. Unlike the other enhancement techniques, TM produced the brightest image; however, the image appears unnatural and saturated, especially the foreground. The performance of the proposed AFIM method is further depicted in Table 4, "Appendix 1". Five non-uniform illumination color images are tabulated in this table. It is attested from this table that the proposed AFIM method successfully enhanced those color images without causing loss of details in the image.

The qualitative results presented in Figs. 7, 8, and 9 are supported by the average and standard deviation values presented in Table 2. As highlighted in Table 2, the proposed method attains the highest values of Q , C , and PL . This result indicates that the proposed method successfully enhanced the image while preserving the details and suppressing existing noise. The graphs of 150 color images are plotted in these

figures for different enhancement techniques. The results show that the proposed AFIM method obtains the highest C , Q and PL for most of the color images. In addition, the error bars in Figs. 14, 15, 16, and 17 depict that the proposed method produced the best results that possess the highest average and low standard deviation values for most of the analysis.

8 Conclusions

A novel enhancement technique based on the fuzzy intensity measure was proposed in this study to solve problems regarding degraded images with non-uniform illumination and low contrast. Image illumination is improved by applying a specific Gaussian function for dark and bright regions. The local contrast of the image was enhanced with a sigmoid function to ensure that the details of the image are preserved. The intensifier and de-intensifier were also applied to the saturation level of the color images to avoid over-saturation. Comparative analyses were performed to evaluate the performance of the proposed AFIM technique. The visual quality of the obtained images were compared, and their quantitative measures were computed. Qualitatively, the enhanced images produced by the proposed technique are able to enhance the dark region, and the existing noise was not enhanced. The qualitative findings also supported by quantitative analysis, which indicates that the proposed technique outperformed the other techniques in terms of improving non-uniform image illumination. The findings reveal that the proposed method surpassed the other techniques and obtained the best image quality index. The proposed algorithm was comparably fast because it selectively enhanced the image according to its corresponding bright or dark regions. Existing noise was successfully attenuated compared with the other techniques. Therefore, the proposed technique is suitable for use in real-time applications.

Appendix 1

See Tables 3 and 4.

Table 3 Comparison of enhancement techniques for grayscale images

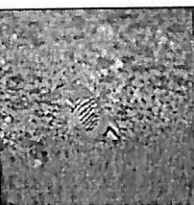



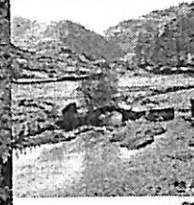
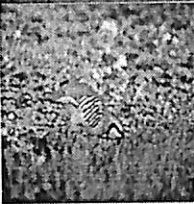









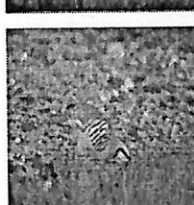




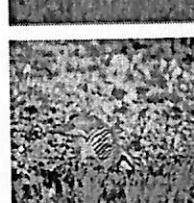



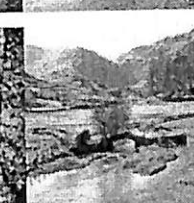
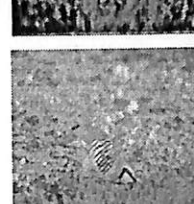



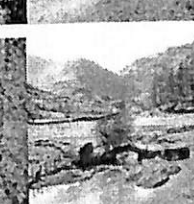
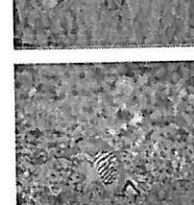



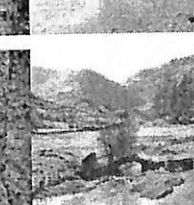
Techniques/ Image	1	2	3	4	5
Original					
AFIM					
FQM					
DCT					
BPDFHE					
TM					
GC					

Table 3 continued




































Techniques/ Image	6	7	8	9	10
Original					
AFIM					
FQM					
DCT					
BPDFHE					
TM					
GC					

Table 4 Comparison of enhancement techniques for color images




































































Techniques/ Image	11	12	13	14	15
Original					
AFIM					
FQM					
DCT					
BPDFHE					
TM					
GC					

Table 4 continued

Techniques/ Image	16	17	18	19	20
Original					
AFIM					
FQM					
DCT					
BPDFHE					
TM					
GC					

Appendix 2

See Tables 5, 6, 7 and 8.

Table 5 Comparison of enhancement results based on image contrast (C)

Image	Enhancement technique						
	Original	AFIM	FQM	DCT	BPDFHE	TM	GC
1	58.06	73.70	51.98	56.59	61.47	59.91	54.74
2	37.07	50.11	46.62	36.12	49.77	42.36	35.21
3	31.31	52.14	48.88	29.95	56.55	29.28	28.33
4	21.73	36.02	22.26	21.21	35.40	30.11	31.43
5	34.38	62.66	52.43	31.93	56.43	35.41	28.48
6	52.53	67.24	65.67	51.69	53.90	44.20	57.19
7	49.22	67.70	66.28	48.21	52.72	41.12	63.59
8	60.19	66.03	61.30	58.46	71.01	54.70	46.95
9	67.53	72.74	72.03	65.44	67.79	62.42	57.88
10	47.31	64.11	62.03	46.11	49.66	40.48	42.81
11	36.79	41.35	25.41	62.21	65.08	64.10	40.58
12	31.40	59.39	26.31	56.74	39.53	31.21	35.21
13	76.79	80.30	58.97	77.95	34.91	25.42	82.14
14	73.39	79.53	56.39	87.09	77.82	26.97	77.83
15	79.42	84.63	59.34	41.79	86.46	23.65	83.97
16	78.22	82.68	42.31	92.35	42.11	28.74	82.67
17	65.35	97.80	56.91	65.08	43.45	36.51	69.99
18	57.15	68.62	55.12	50.53	84.16	41.23	61.05
19	76.21	88.61	51.32	84.02	53.74	45.31	80.76
20	40.41	66.81	24.36	44.19	17.67	28.41	44.35

The best results are made bold
The image contrast is calculated and compared with the original image (i.e., without enhancement) and the enhanced images obtained by different enhancement techniques

Table 6 Comparison of enhancement results based on image quality index (Q)

Image	Enhancement techniques					
	AFIM	FQM	DCT	BPDFHE	TM	GC
1	0.99	0.86	0.67	0.80	0.90	0.93
2	0.94	0.92	0.66	0.96	0.91	0.87
3	0.99	0.93	0.72	0.92	0.89	0.59
4	0.98	0.96	0.78	0.91	0.88	0.96
5	0.96	0.92	0.75	0.92	0.93	0.85
6	0.96	0.87	0.85	0.90	0.90	0.96
7	0.96	0.99	0.76	0.97	0.91	0.94
8	0.95	0.94	0.69	0.92	0.89	0.85
9	0.99	0.96	0.94	0.91	0.88	0.59

Table 6 continued

Image	Enhancement techniques					
	AFIM	FQM	DCT	BPDFHE	TM	GC
10	0.94	0.95	0.91	0.96	0.93	0.86
11	0.72	0.41	0.54	0.69	0.57	0.72
12	0.66	0.38	0.50	0.64	0.56	0.67
13	0.97	0.93	0.90	0.96	0.92	0.96
14	0.93	0.84	0.89	0.83	0.91	0.94
15	0.96	0.93	0.95	0.97	0.85	0.98
16	0.95	0.82	0.93	0.96	0.83	0.97
17	0.94	0.83	0.76	0.84	0.88	0.93
18	0.88	0.67	0.70	0.81	0.73	0.81
19	0.96	0.84	0.93	0.83	0.65	0.78
20	0.66	0.48	0.54	0.76	0.56	0.68

The best results are made bold
The image quality index is computed by considering difference between original and enhanced image as presented in Eqs. (14) and (22) for grayscale and color images, respectively. Thus, the quantitative results attained by the proposed method are tabulated and compared with other enhancement techniques

Table 7 Comparison of enhancement results based on PL

Image	Enhancement techniques					
	AFIM	FQM	DCT	BPDFHE	TM	GC
1	97.40	30.57	48.56	61.09	49.04	74.41
2	80.97	69.44	49.62	57.22	27.55	78.98
3	137.64	52.45	41.23	43.76	20.87	73.64
4	98.67	83.29	50.68	98.44	51.01	88.67
5	92.44	46.55	52.31	48.40	25.68	84.44
6	53.83	30.57	47.69	61.09	49.04	35.41
7	48.54	69.44	51.36	57.22	27.59	36.84
8	68.45	52.45	54.32	43.76	20.86	57.59
9	67.53	103.29	54.37	98.44	51.01	56.18
10	46.97	46.55	37.34	48.40	25.68	49.73
11	96.62	95.63	113.56	137.23	53.34	90.10
12	101.40	87.54	104.21	124.63	74.08	95.55
13	96.05	132.64	91.26	115.32	33.12	81.17
14	206.97	125.21	109.39	75.71	27.92	152.90
15	147.11	124.69	113.21	49.31	27.61	120.70
16	173.18	134.95	99.68	54.92	29.35	239.45
17	185.31	124.36	134.61	93.00	31.23	123.81
18	289.34	123.18	93.21	194.24	45.80	163.44
19	173.12	126.41	104.32	164.84	25.93	171.39
20	506.97	129.74	81.24	165.93	52.83	491.73

The best results are made bold
The PL is computed by considering difference between original and enhanced image as presented in Eq. (25). Thus, the quantitative results attained by the proposed method are tabulated and compared with other enhancement techniques

Table 8 Comparison of enhancement results based on processing time (*t*) (in s)

Image	Enhancement techniques					
	AFIM	FQM	DCT	BPDFHE	TM	GC
1	0.19	1.40	6.72	0.71	0.58	0.49
2	0.09	0.71	5.67	0.18	0.18	0.04
3	0.07	0.66	5.67	0.14	0.14	0.06
4	0.10	0.59	5.61	0.15	0.16	0.07
5	0.06	0.69	5.58	0.14	0.13	0.05
6	0.18	1.30	7.21	0.73	0.57	0.42
7	0.45	0.80	5.72	0.18	0.18	0.15
8	0.35	0.69	5.65	0.15	0.16	0.09
9	0.31	0.62	5.65	0.14	0.14	0.09
10	0.49	0.74	5.61	0.15	0.14	0.19
11	0.51	1.15	7.50	1.22	0.86	0.19
12	0.43	1.09	6.18	0.59	0.50	0.07
13	0.43	1.07	5.94	0.51	0.56	0.06
14	0.44	1.21	5.91	0.51	0.53	0.05
15	0.42	1.17	6.02	0.51	0.52	0.06
16	0.42	1.27	6.04	0.50	0.48	0.05
17	0.41	1.13	6.02	0.50	0.52	0.09
18	0.42	1.44	5.92	0.61	0.48	0.09
19	0.43	1.29	5.99	0.53	0.47	0.07
20	0.44	1.25	5.89	0.52	0.61	0.07

The best results are made bold

Appendix 3

Quantitative analysis for grayscale images

See Figs. 10, 11, 12, and 13.

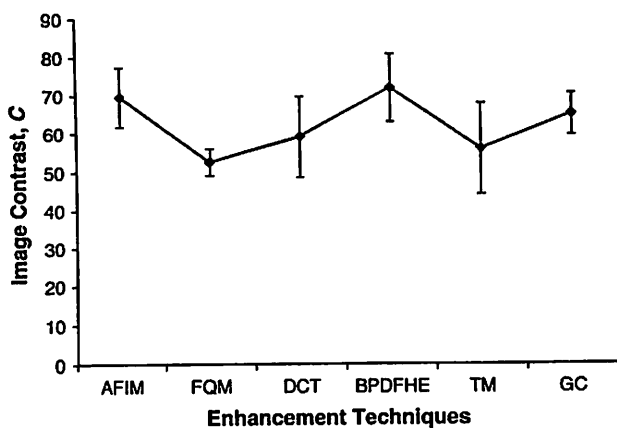


Fig. 10 Error bars of image contrast for different enhancement techniques. Graph is plotted by computing average and standard deviation of image contrast for 150 grayscale images

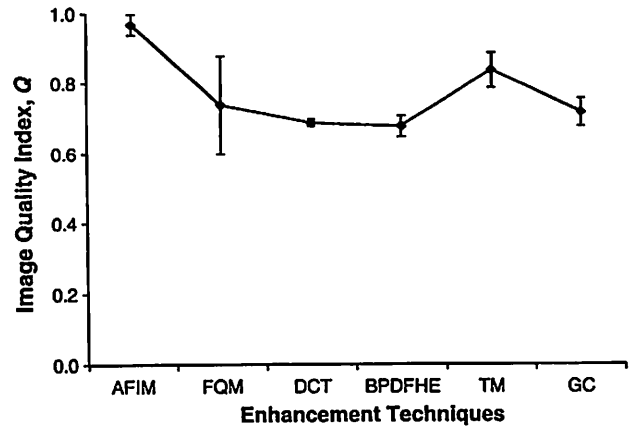


Fig. 11 Error bars of image quality index for different enhancement techniques. Graph is plotted by computing average and standard deviation of image contrast for 150 grayscale images

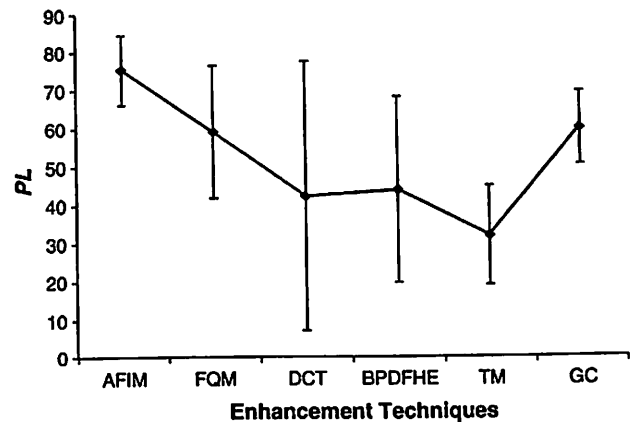


Fig. 12 Error bars of PL for different enhancement techniques. Graph is plotted by computing average and standard deviation of image contrast for 150 grayscale images

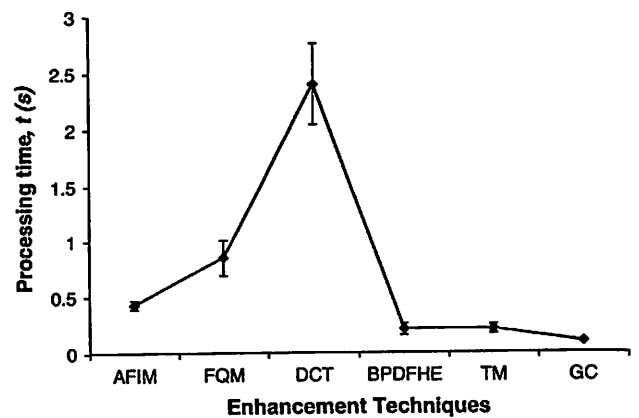


Fig. 13 Error bars of processing time for different enhancement techniques. Graph is plotted by computing average and standard deviation of image contrast for 150 grayscale image

Appendix 4

Quantitative analysis for color images

See Figs. 14, 15, 16, and 17.

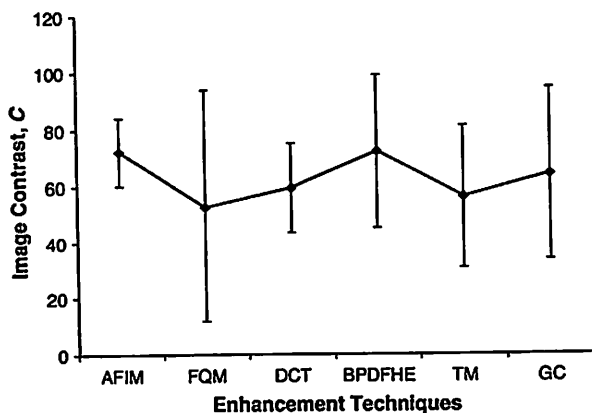


Fig. 14 Error bars of image contrast for different enhancement techniques. Graph is plotted by computing average and standard deviation of image contrast for 150 color images

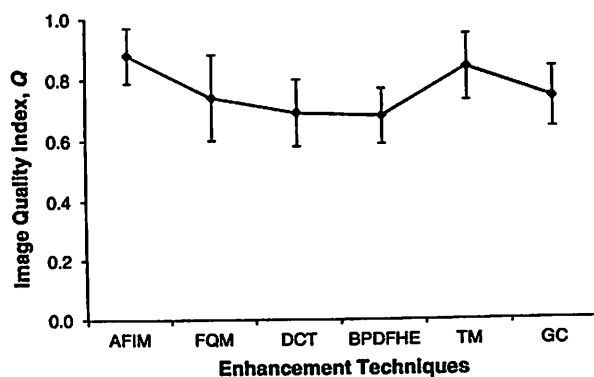


Fig. 15 Error bars of image quality index for different enhancement techniques. Graph is plotted by computing average and standard deviation of image contrast for 150 color images

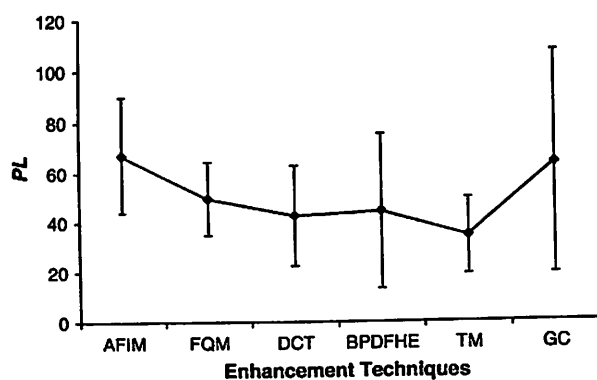


Fig. 16 Error bars of PL for different enhancement techniques. Graph is plotted by computing average and standard deviation of image contrast for 150 color images

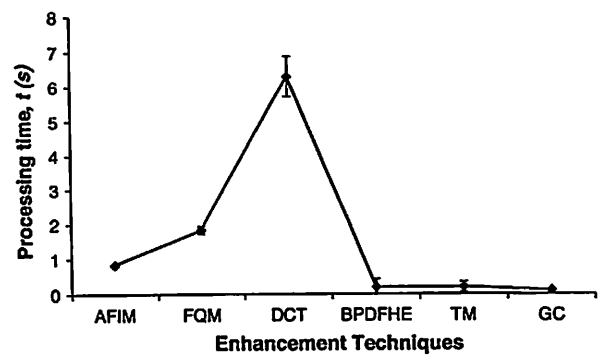


Fig. 17 Error bars of processing time for different enhancement techniques. Graph is plotted by computing average and standard deviation of image

References

- Bumbaca, F., Smith, K.C.: A practical approach to image restoration for computer vision. *Comput. Vis. Graph. Image Process.* **42**(2), 220–233 (1988)
- Huang, K.-Q., Wang, Q., Wu, Z.-Y.: Natural color image enhancement and evaluation algorithm based on human visual system. *Comput. Vis. Image Underst.* **103**(1), 52–63 (2006)
- Chan, F.W.Y.: Measurement sensitivity enhancement by improved reflective computer vision technique for non-destructive evaluation. *NDT & E Int.* **43**(3), 210–215 (2010)
- Xue-bo, J., Hai-jiang, Z., Jing-jing, D., Ya-ming, W.: S-Retinex brightness adjustment for night-vision image. *Procedia Eng.* **15**, 2571–2576 (2011)
- Shieh, W.-Y., Huang, J.-C.: Falling-incident detection and throughput enhancement in a multi-camera video-surveillance system. *Med. Eng. Phys.* **34**(7), 954–963 (2012)
- Kim, S., Shin, J., Paik, J.: Real-time iterative framework of regularized image restoration and its application to video enhancement. *Real-Time Imaging* **9**(1), 61–72 (2003)
- Wu, Y., Sun, Y., Zhang, H.: A fast video illumination enhancement method based on simplified VEC model. *Procedia Eng.* **29**, 3668–3673 (2012)
- Bhaskar, H., Singh, S.: Live cell imaging: a computational perspective. *J. Real-Time Image Process.* **1**(3), 195–212 (2007)
- Feng, J., Xiong, N., Yang, L., Yang, Y.: A low distortion image enhancement scheme based on multi-resolutions analysis in next generation network. *Multimed. Tools Appl.* **56**(2), 227–243 (2012)
- Hasikin, K., Isa, N.A.M.: Fuzzy enhancement for nonuniform illumination of microscopic Sprague Dawley rat sperm image. In: *Medical Measurements and Applications Proceedings (MeMeA), 2012 IEEE International Symposium on*, pp. 1–6 (2012)
- Kimori, Y.: Mathematical morphology-based approach to the enhancement of morphological features in medical images. *J. Clin. Bioinform.* **1**(1), 33 (2011)
- Mittal, D., Kumar, V., Saxena, S., Khandelwal, N., Kalra, N.: Enhancement of the ultrasound images by modified anisotropic diffusion method. *Med. Biol. Eng. Comput.* **48**(12), 1281–1291 (2010)
- Shiraishi, J., Li, Q., Appelbaum, D., Doi, K.: Computer-aided diagnosis and artificial intelligence in clinical imaging. *Semin. Nucl. Med.* **41**(6), 449–462 (2011)
- Shiraishi, J., Li, Q., Appelbaum, D., Pu, Y., Doi, K.: Development of a computer-aided diagnostic scheme for detection of interval changes in successive whole-body bone scans. *Med. Phys.* **34**(1), 25–36 (2007)

15. Lasaponara, R., Masini, N.: Image enhancement, feature extraction and geospatial analysis in an archaeological perspective. In: Lasaponara, R., Masini, N. (eds.) *Satellite Remote Sensing, Remote Sensing and Digital Image Processing*, vol. 16, pp. 17–63. Springer, Netherlands (2012)
16. Pan, X.-B., Brady, M., Bowman, A.K., Crowther, C., Tomlin, R.S.O.: Enhancement and feature extraction for images of incised and ink texts. *Image Vis. Comput.* **22**(6), 443–451 (2004)
17. Ryu, C., Kong, S.G., Kim, H.: Enhancement of feature extraction for low-quality fingerprint images using stochastic resonance. *Pattern Recognit. Lett.* **32**(2), 107–113 (2011)
18. Islam, M.R., Sayeed, M.S., Samraj, A.: Technology review: image enhancement, feature extraction and template protection of a fingerprint authentication system. *J. Appl. Sci.* **10**(14), 1397–1404 (2010)
19. Baradarani, A., Jonathan Wu, Q.M., Ahmadi, M.: An efficient illumination invariant face recognition framework via illumination enhancement and DD-DTWT filtering. *Pattern Recognit.* **46**(1), 57–72 (2013)
20. Chen, T., Wotao, Y., Xiang Sean, Z., Comaniciu, D., Huang, T.S.: Total variation models for variable lighting face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **28**(9), 1519–1524 (2006)
21. Eleyan, A., Ozkaramanli, H., Demirel, H.: Dual-tree and single-tree complex wavelet transform based face recognition. In: *Signal Processing and Communications Applications Conference, 2009. SIU 2009. IEEE 17th, 9–11 April 2009 2009*, pp. 536–539
22. Yun, E.-K., Cho, S.-B.: Adaptive fingerprint image enhancement with fingerprint image quality analysis. *Image Vis. Comput.* **24**(1), 101–110 (2006)
23. Bal, U., Engin, M., Utzinger, U.: A multiresolution approach for enhancement and denoising of microscopy images. *SIViP*. (2013). doi:10.1007/s11760-013-0510-x
24. González, R.C., Woods, R.E.: *Digital Image Processing*. Prentice Hall, Englewood Cliffs, NJ (2002)
25. Cherifi, D., Beghdadi, A., Belbachir, A.H.: Color contrast enhancement method using steerable pyramid transform. *SIViP* **4**(2), 247–262 (2010)
26. Hanmandlu, M., Jha, D.: An optimal fuzzy system for color image enhancement. *IEEE Trans. Image Process* **15**(10), 2956–2966 (2006)
27. Qing, W., Ward, R.K.: Fast image/video contrast enhancement based on weighted thresholded histogram equalization. *IEEE Trans. Consum. Electron* **53**(2), 757–764 (2007)
28. Chen Hee, O., Isa, N.A.M.: Quadrants dynamic histogram equalization for contrast enhancement. *IEEE Trans. Consum. Electron.* **56**(4), 2552–2559 (2010)
29. Chen Hee, O., Isa, N.A.M.: Adaptive contrast enhancement methods with brightness preserving. *IEEE Trans. Consum. Electron.* **56**(4), 2543–2551 (2010)
30. Sen, D., Pal, S.K.: Automatic exact histogram specification for contrast enhancement and visual system based quantitative evaluation. *IEEE Trans. Image Process.* **20**(5), 1211–1220 (2011)
31. Joung-Youn, K., Lee-Sup, K., Seung-Ho, H.: An advanced contrast enhancement using partially overlapped sub-block histogram equalization. *IEEE Trans. Circuits Syst. Video Technol.* **11**(4), 475–484 (2001)
32. Chao, W., Zhongfu, Y.: Brightness preserving histogram equalization with maximum entropy: a variational perspective. *IEEE Trans. Consum. Electron.* **51**(4), 1326–1334 (2005)
33. Sengee, N., Sengee, A., Heung-Kook, C.: Image contrast enhancement using bi-histogram equalization with neighborhood metrics. *IEEE Trans. Consum. Electron.* **56**(4), 2727–2734 (2010)
34. Ibrahim, H., Kong, N.S.P.: Brightness preserving dynamic histogram equalization for image contrast enhancement. *IEEE Trans. Consum. Electron.* **53**(4), 1752–1758 (2007)
35. Lim, S., Isa, N.M., Ooi, C., Toh, K.: A new histogram equalization method for digital image enhancement and brightness preservation. *SIViP*. (2013). doi:10.1007/s11760-013-0500-z
36. Thomas, G., Flores-Tapia, D., Pistorius, S.: Histogram specification: a fast and flexible method to process digital images. *IEEE Trans. Instrum. Meas.* **60**(5), 1565–1578 (2011)
37. Wang, C., Peng, J., Ye, Z.: Flattest histogram specification with accurate brightness preservation. *Image Process. IET* **2**(5), 249–262 (2008)
38. Abdullah-Al-Wadud, M., Kabir, M.H., Dewan, M.A.A., Oksam, C.: A dynamic histogram equalization for image contrast enhancement. *IEEE Trans. Consum. Electron.* **53**(2), 593–600 (2007)
39. Duan, J., Bressan, M., Dance, C., Qiu, G.: Tone-mapping high dynamic range images by novel histogram adjustment. *Pattern Recognit.* **43**(5), 1847–1862 (2010)
40. Land, E.: The retinex. *Am. Sci.* **52**(2), 247–264 (1964)
41. Bai, X., Zhou, F., Xue, B.: Image enhancement using multi scale image features extracted by top-hat transform. *Opt. Laser Technol.* **44**(2), 328–336 (2012)
42. Eschbach, R.: Image-dependent exposure enhancement, Google Patents (1995)
43. Pal, S.K., King, R.A.: Image enhancement using smoothing with fuzzy sets. *IEEE Trans. Syst. Man Cybern.* **11**(7), 494–501 (1981)
44. Chen, H.-C., Wang, W.-J.: Efficient impulse noise reduction via local directional gradients and fuzzy logic. *Fuzzy Sets Syst.* **160**(13), 1841–1857 (2009)
45. Cheng, H.D., Xu, H.: A novel fuzzy logic approach to mammogram contrast enhancement. *Inf. Sci.* **148**(1–4), 167–184 (2002)
46. Sheet, D., Garud, H., Suveer, A., Mahadevappa, M., Chatterjee, J.: Brightness preserving dynamic fuzzy histogram equalization. *IEEE Trans. Consum. Electron.* **56**(4), 2475–2480 (2010)
47. Hanmandlu, M., Verma, O.P., Kumar, N.K., Kulkarni, M.: A novel optimal fuzzy system for color image enhancement using bacterial foraging. *IEEE Trans. Instrum. Meas.* **58**(8), 2867–2879 (2009)
48. Verma, O.P., Kumar, P., Hanmandlu, M., Chhabra, S.: High dynamic range optimal fuzzy color image enhancement using Artificial Ant Colony System. *Appl. Soft Comput.* **12**(1), 394–404 (2012)
49. Vlachos, I., Sergiadis, G., Melin, P., Castillo, O., Aguilar, L., Kacprzyk, J., Pedrycz, W.: The role of entropy in intuitionistic fuzzy contrast enhancement foundations of fuzzy logic and soft computing. In: *Lecture Notes in Computer Science*, vol. 4529, pp. 104–113. Springer, Berlin (2007)
50. Vlachos, I.K., Sergiadis, G.D.: Parametric indices of fuzziness for automated image enhancement. *Fuzzy Sets Syst.* **157**(8), 1126–1138 (2006)
51. Cheng, H.D., Chen, J.-R.: Automatically determine the membership function based on the maximum entropy principle. *Inf. Sci.* **96**(3–4), 163–182 (1997)
52. Hasikin, K., Isa, N.A.M.: Enhancement of the low contrast image using fuzzy set theory. In: *Computer Modelling and Simulation (UKSim)*, 2012 UKSim 14th International Conference on, 28–30 March 2012, pp. 371–376 (2012)
53. Hasikin, K., Isa, N.A.M.: Adaptive fuzzy contrast factor enhancement technique for low contrast and nonuniform illumination images. *SIViP*. (2012). doi:10.1007/s11760-012-0398-x
54. Aarmink, R., Pathak, S.D., de la Rosette, J.J.M.C.H., Debruyne, F.M.J., Kim, Y., Wijkstra, H.: Edge detection in prostatic ultrasound images using integrated edge maps. *Ultrasonics* **36**(1–5), 635–642 (1998)
55. Toet, A., Lucassen, M.P.: A new universal colour image fidelity metric. *Displays* **24**(4–5), 197–207 (2003)
56. Jain, A.K.: *Fundamentals of Digital Image Processing*. Prentice Hall, Englewood Cliffs, NJ (1989)

A new histogram equalization method for digital image enhancement and brightness preservation

Sheng Hoong Lim · Nor Ashidi Mat Isa ·
Chen Hee Ooi · Kenny Kal Vin Toh

Received: 20 December 2012 / Revised: 7 May 2013 / Accepted: 8 May 2013
© Springer-Verlag London 2013

Abstract This paper proposes a new histogram equalization method for effective and efficient mean brightness preservation and contrast enhancement, which prevents intensity saturation and has the ability to preserve image fine details. Basically, the proposed method first separates the test image histogram into two sub-histograms. Then, the plateau limits are calculated from the respective sub-histograms, and they are used to modify those sub-histograms. Histogram equalization is then separately performed on the two sub-histograms to yield a clean and enhanced image. To demonstrate the feasibility of the proposed method, a total of 190 test images are used in simulation and comparison, in which 72 of them are standard test images, while the remainder are made up of real natural images obtained from personal digital camera. The simulation results show that the proposed method outperforms other state-of-the-art methods, both in terms of visual and runtime comparison. Moreover, the simple implementation and fast runtime further underline the importance of the proposed method in consumer electronic products, such as mobile cell-phone, digital camera, and video.

This work was supported by the Ministry of Science, Technology and Innovation, Malaysia under Science Fund Grant entitled 'Development of Computational Intelligent Infertility Detection System Based on Sperm Motility Analysis'.

S. H. Lim · N. A. Mat Isa (✉)
School of electrical & Electronic Engineering,
Imaging and Intelligent Systems Research Team (ISRT),
Universiti Sains Malaysia, Engineering Campus,
14300 Nibong Tebal, Penang, Malaysia
e-mail: ashidi@eng.usm.my

C. H. Ooi · K. K. V. Toh
Intel Technology (M) Sdn. Bhd., Box 121,
Bayan Lepas Free Trade Zone, 11900 Bayan Lepas,
Penang, Malaysia

Keywords Contrast enhancement · Histogram equalization · Brightness preservation · Plateau limit

1 Introduction

Histogram equalization (HE) is widely used due to its simplicity and effectiveness in providing contrast enhancement. In fact, HE-based methods have been used in the field of consumer electronics, medical image processing, image matching and searching, speech recognition, and texture synthesis [1]. Basically, HE transforms the resultant image according to the probability distribution function of the test image. It stretches and flattens the dynamic range of the image. As a result, the overall contrast of the processed image is improved [1]. However, sometimes it can degenerate the quality of the resultant image by introducing the washed-out appearance and by producing unnecessary visual degradation. This is especially true for low-contrast images, (i.e., images with few bins in the histogram contain most of the weight of the input image histogram) where washed-out appearance will occur after applying HE. Furthermore, HE is also known for shifting the mean brightness of the enhanced image, in which the mean brightness of the input image is significantly altered from that of the resultant image. Due to this significant change in mean brightness, unnecessary visual deterioration is unavoidable.

In order to solve the aforementioned problem, many HE-based methods have been proposed. Chen and Ramli [2] introduce minimum mean brightness error bi-histogram (MMBEBHE). The objective of this method is to provide maximum brightness preservation in the enhanced image. MMBEBHE firstly separates the input image histogram into two sub-histograms using a separating point. Then, it independently applies HE to the two sub-histograms. The

separating point is determined by the minimum brightness difference between the input image and the resultant image. However, the MMBEBHE suffers from slow runtime as a result of its highly complex implementation. For example, given an 8-bit image, the MMBEBHE process requires 256 repetitions (i.e., one time per gray level) before the separating point is found. Moreover, the minimal mean shift in the output does not always guarantee the visual quality of the enhanced image.

Another attempt in addressing the problems faced by the HE is known as the recursively separated and weighted histogram equalization (RSWHE) [3], which is introduced to improve the contrast while preserving brightness of the enhanced image. RSWHE consists of three modules. It first performs the histogram segmentation module that recursively separates an image histogram into two or more sub-histograms. The separating point can be either mean or median of the histogram. Then, it proceeds with the histogram weighting module, which modifies the sub-histograms by using a weighting process based on a normalized power law function. Lastly, HE module is applied. In this module, HE is applied to all sub-histograms separately. As claimed in [3], the RSWHE is considered as the leading state-of-the-art method in brightness-preserving image enhancement. However, the RSWHE tends to introduce unnatural enhancement to the processed image under some circumstances. Additionally, the user-defined selection of recursion level prevents the RSWHE to yield optimal enhancement performance.

Next, the weighted clustering histogram equalization (WCHE) [4] is introduced to preserve image brightness, improve contrast, and enhance visualization without producing any undesired artifacts through over-enhancement. The WCHE starts with cluster assigning, and then cluster merging based on three criteria (i.e., cluster weight, weight ratio, and width of two neighboring clusters). Lastly, it performs the cluster transformation function before an improved resultant image is produced. Generally, the WCHE yields satisfactory enhancement results, but the WCHE tends to lose pertinent image details because a cluster with large groups of bins may equalize to a narrow dynamic range [4].

Later on, the bi-histogram equalization with plateau limit (BHEPL) [5] is introduced as one of the options for the system that requires a short processing time image enhancement. The BHEPL first separates the input histogram into two sub-histograms by using the mean of the input histogram. Then, a clipping process is applied by clipping each sub-histogram using the mean of the corresponding occupied intensity in the sub-histogram. Finally, HE is applied to each sub-histogram separately.

Then, the bi-histogram equalization with median plateau limit (BHEPL-D) [6] is introduced to avoid producing unwanted artifacts. The BHEPL-D is similar to the BHEPL,

but unlike the BHEPL, the BHEPL-D clips each sub-histogram using the median of the corresponding occupied intensity in the sub-histogram.

Another approach, namely the dynamic quadrants histogram equalization plateau limit (DQHEPL) [6] is also introduced to avoid producing unwanted artifacts. The DQPLBHE first performs the histogram segmentation that recursively separates the input histogram into four sub-histograms. The separating point used is the median of the input histogram. Then, a clipping process is applied by clipping each sub-histogram using the mean of the corresponding occupied intensity in the sub-histogram. After that a gray level allocation is applied but maintaining the separating point at the first recursion level. Finally, HE is applied to each sub-histogram separately.

In [7], the weighting mean separated sub-histogram equalization (WMSHE) is another method introduced to solve the problem faced by HE, and it aims to achieve complete contrast enhancement for the small-scale detail. WMSHE begins with histogram segmentation by using weighted-mean value, and it then uses a piecewise transformation function to all sub-histograms to accentuate on the enhancement of fine details in the image. However, the effort to bring out fine image details leads to the generation of noise artifacts, largely due to the enhancement noise components in the input image.

More recently, the simple histogram modification scheme (SHMS) is proposed in [8] to solve the washed-out appearance and patchiness effect due to application of HE. SHMS modifies the image histogram before performing HE on the modified histogram. Modification is done by replacing the first nonzero bin to zero-th bin, and then replacing the last nonzero bin with the minimum bin between the last two nonzero bins. This modification is intended to solve the washed-out appearance and patchiness effect. According to [8], SHMS can be used for both single and multiple thresholding cases. The main advantage is that the SHMS can solve the washed-out appearance and patchiness effect if it applies HE. However, for the case of multiple thresholds, it limits brightness to certain extent. As a result, higher degree of brightness preservation is required to avoid generating annoying artifacts in the enhanced images.

Although the mentioned methods can often improve the contrast of the input images, but they do not handle well the common problems in HE, namely generation of washed-out appearance, intensity saturation, and visual degradation due to the enhancement of the noise. Therefore, in this paper, an efficient algorithm in mean brightness preservation and contrast enhancement that prevent intensity saturation is presented. The proposed method is significantly good in preserving the mean brightness and improving the contrast of images. In addition, it is fast and has simple implementation, in which it does not require any time-consuming training or tuning of parameters. These advantages are in-line with the

requirements of many modern consumer electronic products, namely digital camera, camcorder, camera phone, etc.

The rest of this paper is organized as follow. In Sect. 2, the proposed method is explained in detail. In Sect. 3, simulation results using 190 natural images are used to demonstrate the effectiveness of the proposed method as compare to other existing methods. Finally, Sect. 4 concludes this paper.

2 The novel histogram equalization method

In this section, a detailed description of the proposed method will be discussed. Generally, the proposed method can be divided into three stages: histogram segmentation, histogram modification, and histogram transformation. For a given input image stored as an 8-bit integer, its histogram will be segmented into two sub-histograms. The histogram segmentation process is essential because it preserves the original brightness of the input image to a certain extend even after HE is applied. Then, the sub-histograms will be modified using quantization. Quantization is done by selectively clipping the bins in the sub-histogram to one of the plateau limit. The amount of enhancement to be performed on the input image depends on the contents of the image itself. This argument is valid as different regions in a well-exposed captured image have different contrast and local brightness. At bright regions, there should not be much enhancement as it will lead to intensity saturation problem, while at dark regions, it should be enhanced to bring out the image details. The overall brightness should be preserved in order to maintain the natural appearance of the enhanced image. Therefore, the amount of enhancement at each quantized segments will be determined by the plateau limits, and the application of HE (i.e., histogram transformation) stretches the histogram but maintaining the shape of the original histogram. As a result, fine image details can be brought out while the brightness is preserved, and thus, results in a visually pleasant contrasted image. The remainder of this section will be spent on discussing the three stages in greater depth.

2.1 Histogram segmentation

Let consider Fig. 1 which illustrates the hypothetical histogram of an image commonly found in digital photography.

The histogram segmentation stage starts by finding the separating point (SP) (i.e., the mean intensity of the histogram), which can be defined as

$$SP = \frac{\sum_{k=0}^{L-1} l_k \times n_k}{N} \quad (1)$$

where l_k is the k th gray level and n_k is the number of pixels of gray level k , N is the total number of pixels in the test image, and L is the maximum range of the gray

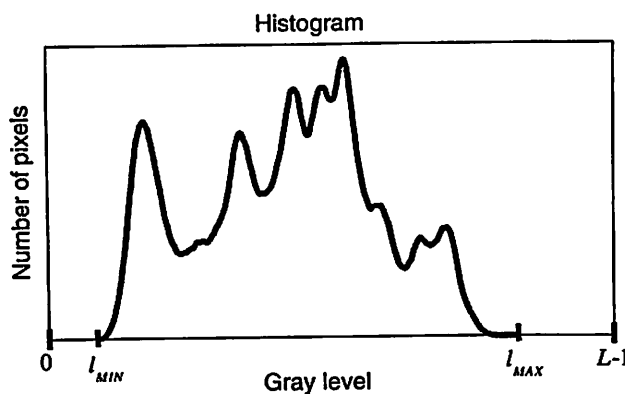


Fig. 1 Histogram of an arbitrary input image

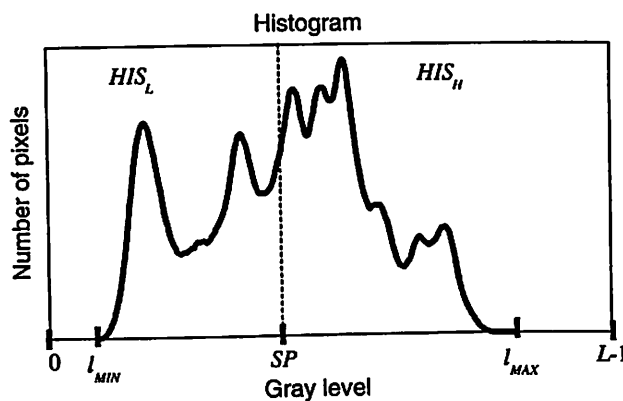


Fig. 2 Histogram segmentation process

level in an image, which is 256 for an 8-bit image. Next, by using the SP value found using (1), the histogram of the test image is separated into two sub-histograms: low sub-histogram, HIS_L , and high sub-histogram, HIS_H . HIS_L consists of pixels that range from minimum gray level l_{MIN} to SP, while HIS_H consists of pixels that range from SP +1 to maximum gray level l_{MAX} . The whole histogram segmentation process is graphically summarized and illustrated in Fig. 2. After the sub-histograms are obtained, they will be modified in the histogram modification stage.

2.2 Histogram modification

After two sub-histograms are segmented, the plateau limits PLs of both sub-histograms are computed. In should be noted that unlike the BHEPL, BHEPL-D, and DQHEPL which use only one plateau limit in each sub-histogram, the proposed method uses three plateau limits. Therefore, there are a total of 6 PLs used in the proposed method, in which three PLs are used in each sub-histogram.

Basically, the PL can be calculated by using a general equation given by:

$$PL = X \times Pk \tag{2}$$

where X is the coefficient with value between 0 and 1, and Pk is the peak bin of the input histogram. In general, the PLs can be obtained by manually setting or by fixing to certain percentage of the Pk . However, these approaches are not user friendly. Therefore, in this work, an automated selection of the PLs is introduced, which suggests that the PLs should be selected based on the local information obtained from the input histogram. Basically, the proposed method uses the local information obtained to automatically compute the X value. One of the ways of extracting information obtained from the input histogram is by using the gray level ratio, GR, in each sub-histogram. It should be noted that the GR is a value between 0 and 1. The GR is used in proposed method as it could reflect the percentage of how much enhancement process should be applied. Lower percentage of enhancement process will be applied to bin with low GR, while for the other bins, the enhancement rate should be increased and vice versa. The calculation of GR will be explained next. By using the GR as the coefficient, the PLs can be calculated as follows

$$PL_{L1} = GR_{L1} \times Pk_L \tag{3}$$

$$PL_{L2} = GR_{L2} \times Pk_L \tag{4}$$

$$PL_{L3} = GR_{L3} \times Pk_L \tag{5}$$

$$PL_{H1} = GR_{H1} \times Pk_H \tag{6}$$

$$PL_{H2} = GR_{H2} \times Pk_H \tag{7}$$

$$PL_{H3} = GR_{H3} \times Pk_H \tag{8}$$

where Pk_L is the peak bin of lower sub-histogram and Pk_H is the peak bin of higher sub-histogram. All the gray level ratios, GRs, are defined as:

$$GR_{L1} = GR_{L2} - D_L \tag{9}$$

$$GR_{L2} = \frac{SP - SP_L}{SP - I_{MIN}} \tag{10}$$

$$GR_{L3} = GR_{L2} + D_L \tag{11}$$

$$GR_{H1} = GR_{H2} - D_H \tag{12}$$

$$GR_{H2} = \frac{I_{MAX} - SP_H}{I_{MAX} - SP} \tag{13}$$

$$GR_{H3} = GR_{H2} + D_H \tag{14}$$

where SP_L and SP_H are the separating points, i.e., means of the lower and higher sub-histograms, respectively, D_L and D_H are the gray level ratio differences of the lower and higher sub-histograms, respectively.

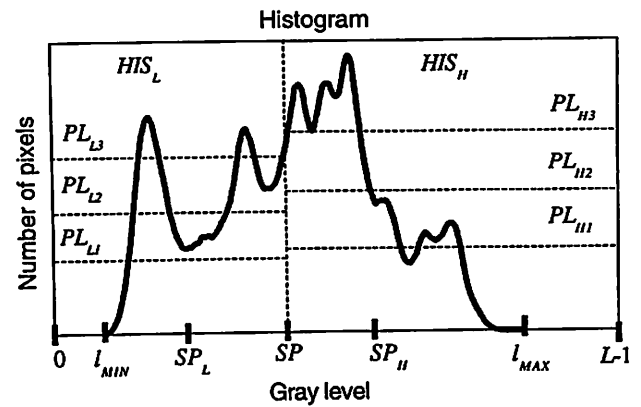


Fig. 3 Histogram with three plateau limits set

SP_L and SP_H can be computed as

$$SP_L = \frac{\sum_{k=I_{MIN}}^{SP} I_k \times n_k}{N_L} \tag{15}$$

$$SP_H = \frac{\sum_{k=SP+1}^{I_{MAX}} I_k \times n_k}{N_H} \tag{16}$$

where N_L and N_H are the total number of pixels in the lower and higher sub-histograms, respectively. As for the gray level ratio differences D_L and D_H , both term are calculated using

$$D_L = \begin{cases} \frac{1-GR_{L2}}{2} & \text{if } GR_{L2} > 0.5 \\ \frac{GR_{L2}}{2} & \text{if } GR_{L2} \leq 0.5 \end{cases} \tag{17}$$

$$D_H = \begin{cases} \frac{1-GR_{H2}}{2} & \text{if } GR_{H2} > 0.5 \\ \frac{GR_{H2}}{2} & \text{if } GR_{H2} \leq 0.5 \end{cases} \tag{18}$$

In Fig. 3, the calculated plateau limits for each sub-histogram is shown. After the PLs values are obtained, the histogram modification can be applied. For reasons to become apparent in a short while, the nonzero bins in the sub-histogram that are lower and equal to the first plateau limit (PL_{L1} and PL_{H1}) are clipped to the first plateau limit. Next, the bins that are located between the first and the second plateau limits will be clipped to the second plateau limit. Then, the bins that are situated between the second and the third plateau limits will be clipped to the second plateau limit. Lastly, for the bins that are greater than the third plateau limit, they will be clipped to the third plateau limit. By doing this, the proposed method can improve the appearance and visual of an image at the bins region below or equal to PL_2 by increasing the enhancement rate. On the other hand, the proposed method decreases the enhancement rate of the bins region greater than the PL_2 to avoid intensity saturation problem. Mathematically, for the lower sub-histogram ($I_{MIN} \leq k \leq SP$), the clipping process can be presented as:

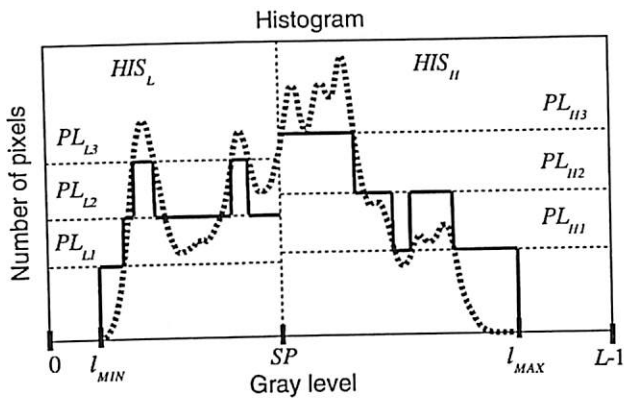


Fig. 4 The quantized histogram

$$HIS_L(k) = \begin{cases} PL_{L1}, & \text{if } HIS_L(k) \leq PL_{L1} \\ PL_{L2}, & \text{if } PL_{L1} < HIS_L(k) \leq PL_{L3} \\ PL_{L3}, & \text{if } HIS_L(k) > PL_{L3} \end{cases} \quad (19)$$

While for the higher sub-histogram ($SP + 1 \leq k \leq l_{MAX}$), the clipping process is given as:

$$HIS_H(k) = \begin{cases} PL_{H1}, & \text{if } HIS_H(k) \leq PL_{H1} \\ PL_{H2}, & \text{if } PL_{H1} < HIS_H(k) \leq PL_{H3} \\ PL_{H3}, & \text{if } HIS_H(k) > PL_{H3} \end{cases} \quad (20)$$

The quantized histogram is shown in Fig. 4. Note that the clipping process locally limits the amount of enhancement. Such strategy circumvents the intensity saturation problems. Once the quantized segments in the sub-histograms are obtained, the processing will proceed to the histogram transformation stage.

2.3 Histogram transformation

Histogram transformation is the final stage in the proposed method. It is done by applying the conventional HE into each modified sub-histogram. However, the probability density function, PDF, and cumulative density function, CDF,

of each modified sub-histogram are needed. PDF and CDF have to be computed first. The PDF and CDF are defined as

For $l_{MIN} \leq k \leq SP$,

$$PDF_L(k) = \frac{n_k}{N_L} \quad (21)$$

$$CDF_L(k) = \sum_{i=l_{MIN}}^{SP} PDF_L(i) \quad (22)$$

For $SP + 1 \leq k \leq l_{MAX}$,

$$PDF_H(k) = \frac{n_k}{N_H} \quad (23)$$

$$CDF_H(k) = \sum_{i=SP+1}^{l_{MAX}} PDF_H(i) \quad (24)$$

where PDF_L and CDF_L are the respective probability density function and cumulative density function of lower sub-histogram, and PDF_H and CDF_H are the probability density function and cumulative density function of higher sub-histogram, respectively. Note that the n_k , N_L , PDF_L , CDF_L , N_H , PDF_H and CDF_H are obtained from the modified histogram. Finally, the enhanced image is obtained with a final round of HE that can be described by:

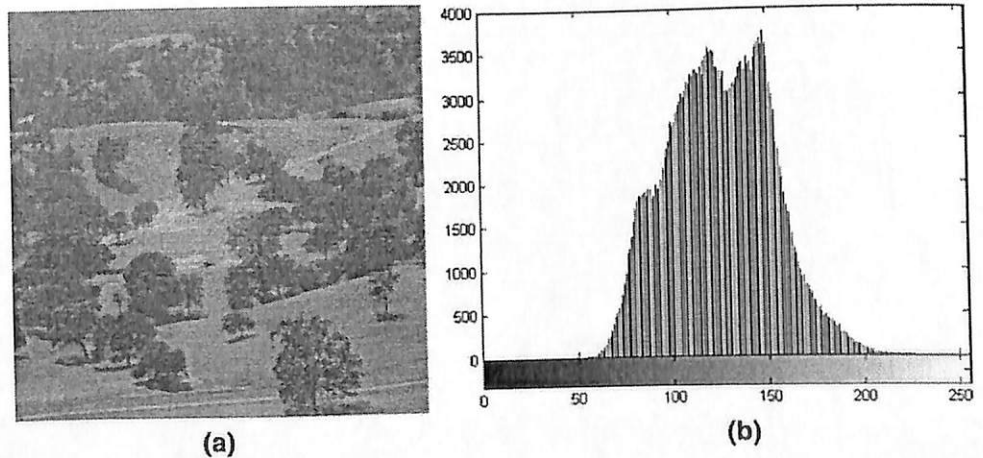
$$T(k) = \begin{cases} SP \times CDF_L(k) & \text{for } l_{MIN} \leq k \leq SP \\ SP + 1 + (255 - SP - 1) \times CDF_H(k) & \text{for } SP < k \leq l_{MAX} \end{cases} \quad (25)$$

2.4 Real image demonstration

In order to further demonstrate the proposed method, a real image is used in the explanation. Fig. 5 illustrates the ‘Hill’ image and its corresponding histogram.

The proposed method starts with the histogram segmentation stage. The histogram segmentation stage begins by finding the SP. The SP value is computed by using Eq. (1), and it is equal to 125. The input histogram is separated into two sub-histograms based on the SP value found. The segmented

Fig. 5 ‘Hill’ image (a) and its corresponding histogram (b)



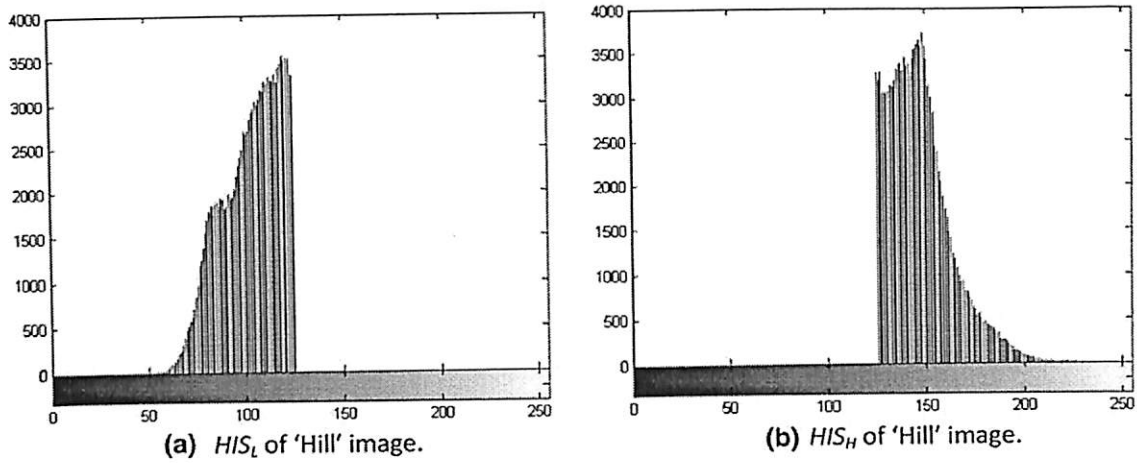


Fig. 6 Segmented histogram HIS_L (a) and HIS_H (b) of 'Hill' image

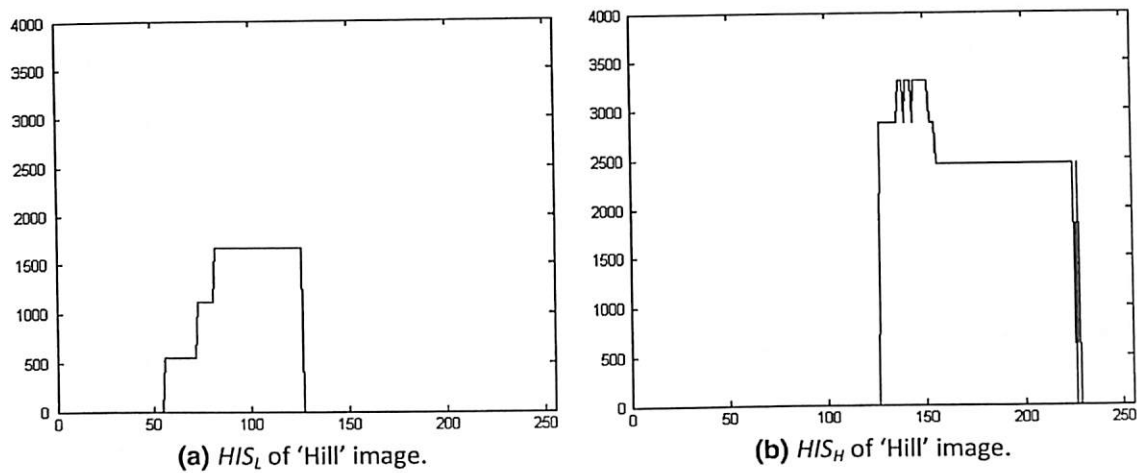


Fig. 7 Quantized histogram HIS_L (a) and HIS_H (b) of 'Hill' image

histogram is as shown in Fig. 6. After the sub-histograms are obtained, they will be modified in the histogram modification stage.

After two sub-histograms are segmented, the PLs of both sub-histograms are computed. Before computing the PLs, several values of local information (i.e., SP_L (Eq. (15)), SP_H (Eq. (16)), l_{MIN} , l_{MAX} , Pk_L , and Pk_H) need to be determined. Based on the local information of the 'Hill' image, the values of SP_L , SP_H , l_{MIN} , l_{MAX} , Pk_L , and Pk_H are obtained automatically to be 103, 148, 55, 226, 3,543, and 3,724, respectively. Then, GR_{L2} and GR_{H2} are calculated by using Eqs. (10), (13), respectively.

$$GR_{L2} = \frac{SP - SP_L}{SP - l_{MIN}} = \frac{125 - 103}{125 - 55} = 0.3143$$

$$GR_{H2} = \frac{l_{MAX} - SP_H}{l_{MAX} - SP} = \frac{226 - 148}{226 - 125} = 0.7723$$

After that D_L and D_H are computed by using Eqs. (17) and (18), respectively.

$$D_L = \begin{cases} \frac{1-GR_{L2}}{2} & \text{if } GR_{L2} > 0.5 \\ \frac{GR_{L2}}{2} & \text{if } GR_{L2} \leq 0.5 \end{cases} = \frac{0.3143}{2} = 0.1571$$

$$D_H = \begin{cases} \frac{1-GR_{H2}}{2} & \text{if } GR_{H2} > 0.5 \\ \frac{GR_{H2}}{2} & \text{if } GR_{H2} \leq 0.5 \end{cases} = \frac{1 - 0.7723}{2} = 0.1139$$

Next, GR_{L1} , GR_{L3} , GR_{H1} , and GR_{H3} are calculated by using Eqs. (9), (11), (12) and (14), respectively.

$$GR_{L1} = GR_{L2} - D_L = 0.3143 - 0.1571 = 0.1572$$

$$GR_{L3} = GR_{L2} + D_L = 0.3143 + 0.1571 = 0.4714$$

$$GR_{H1} = GR_{H2} - D_H = 0.7723 - 0.1139 = 0.6584$$

$$GR_{H3} = GR_{H2} + D_H = 0.7723 + 0.1139 = 0.8862$$

Finally, the PLs are computed according to Eqs. (3) until (8).

$$PL_{L1} = GR_{L1} \times Pk_L = 0.1572 \times 3,543 = 557$$

$$PL_{L2} = GR_{L2} \times Pk_L = 0.3143 \times 3,543 = 1,114$$

$$PL_{L3} = GR_{L3} \times Pk_L = 0.4714 \times 3,543 = 1,670$$

Fig. 8 Enhanced 'Hill' image (a) of the proposed method and its corresponding histogram (b)

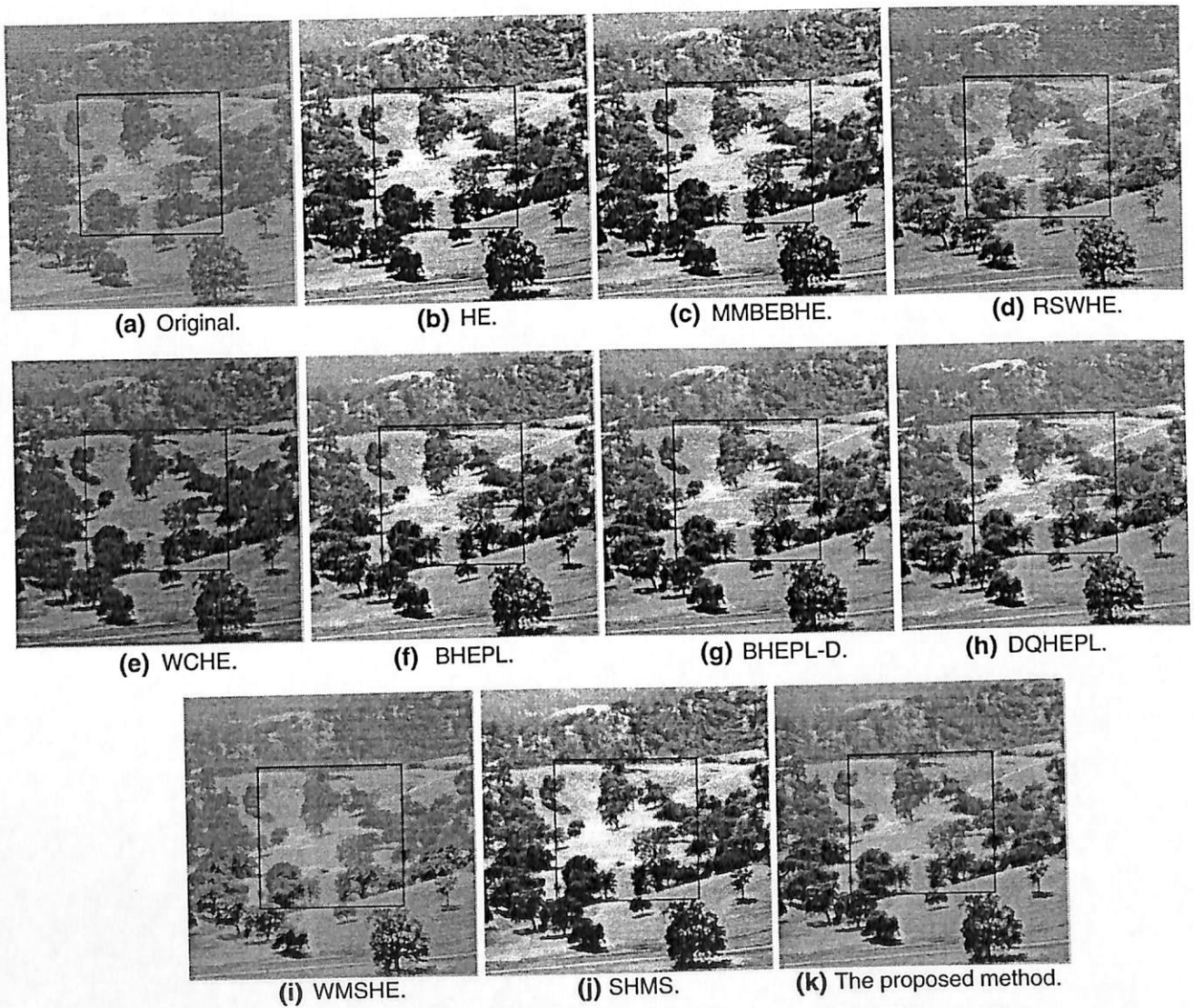
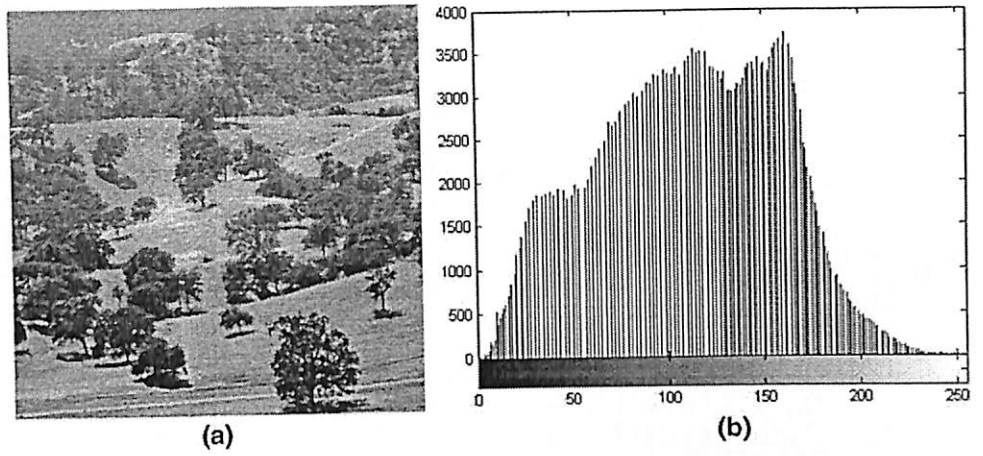


Fig. 9 Simulation results for the 'Hill' image (resolution 512×512)

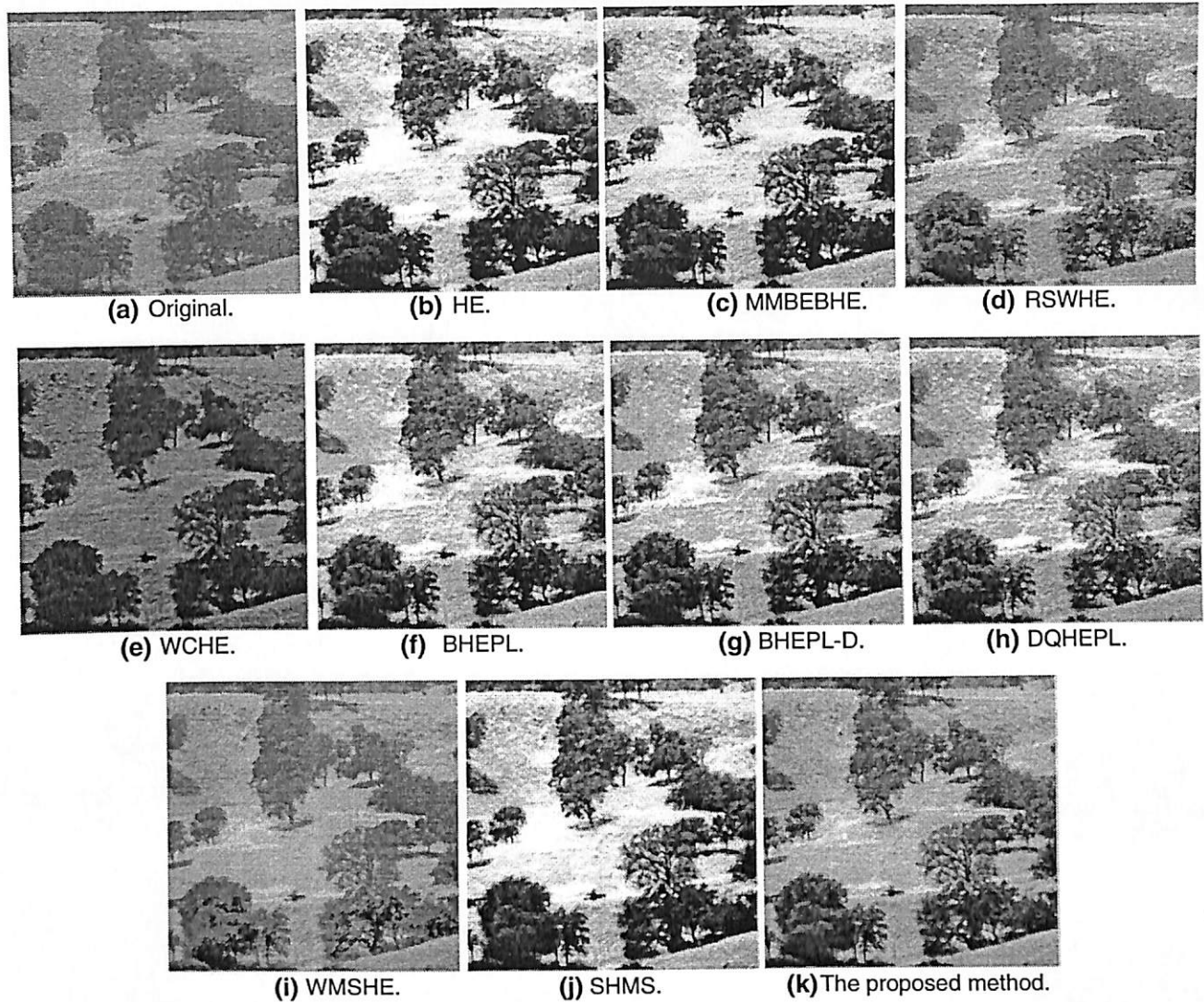


Fig. 10 Simulation results for the highlighted area of 'Hill' image

$$PL_{H1} = GR_{H1} \times Pk_H = 0.6584 \times 3,724 = 2,452$$

$$PL_{H2} = GR_{H2} \times Pk_H = 0.7723 \times 3,724 = 2,876$$

$$PL_{H3} = GR_{H3} \times Pk_H = 0.8862 \times 3,724 = 3,300$$

After the PLs' values are obtained, each sub-histogram is modified by a clipping process as described in Eqs. (19) and (20), respectively. The quantized histogram is as shown in Fig. 7. Once the quantized histograms are obtained, the proposed method will proceed to the histogram transformation stage.

Histogram transformation stage is done by applying HE into each modified sub-histogram. The PDF and CDF need to be calculated (i.e., as described in Eqs. (21)–(24)) before applying HE. The enhanced image is obtained by applying HE as described in Eq. (25). The enhanced image and its corresponding histogram are shown in Fig. 8.

3 Results and discussion

In order to demonstrate the feasibility of the proposed method, few of the state-of-the-art methods previously mentioned in Sect. 1 are implemented for comparison purposes. The 190 natural images are used as the input images to the implemented filters, and their outputs are compared based on two criteria, namely qualitative and quantitative analyses. For simulation purposes, the resolution of the images range from 467×413 to $2,352 \times 1,568$.

In this section, simulation will utilize 190 natural images that can be further divided into two categories: standard images and real images obtained from personal digital camera. The standard images selected are obtained from multiple internet sources that are commonly used for evaluating contrast enhancement method, with their dimensions ranging

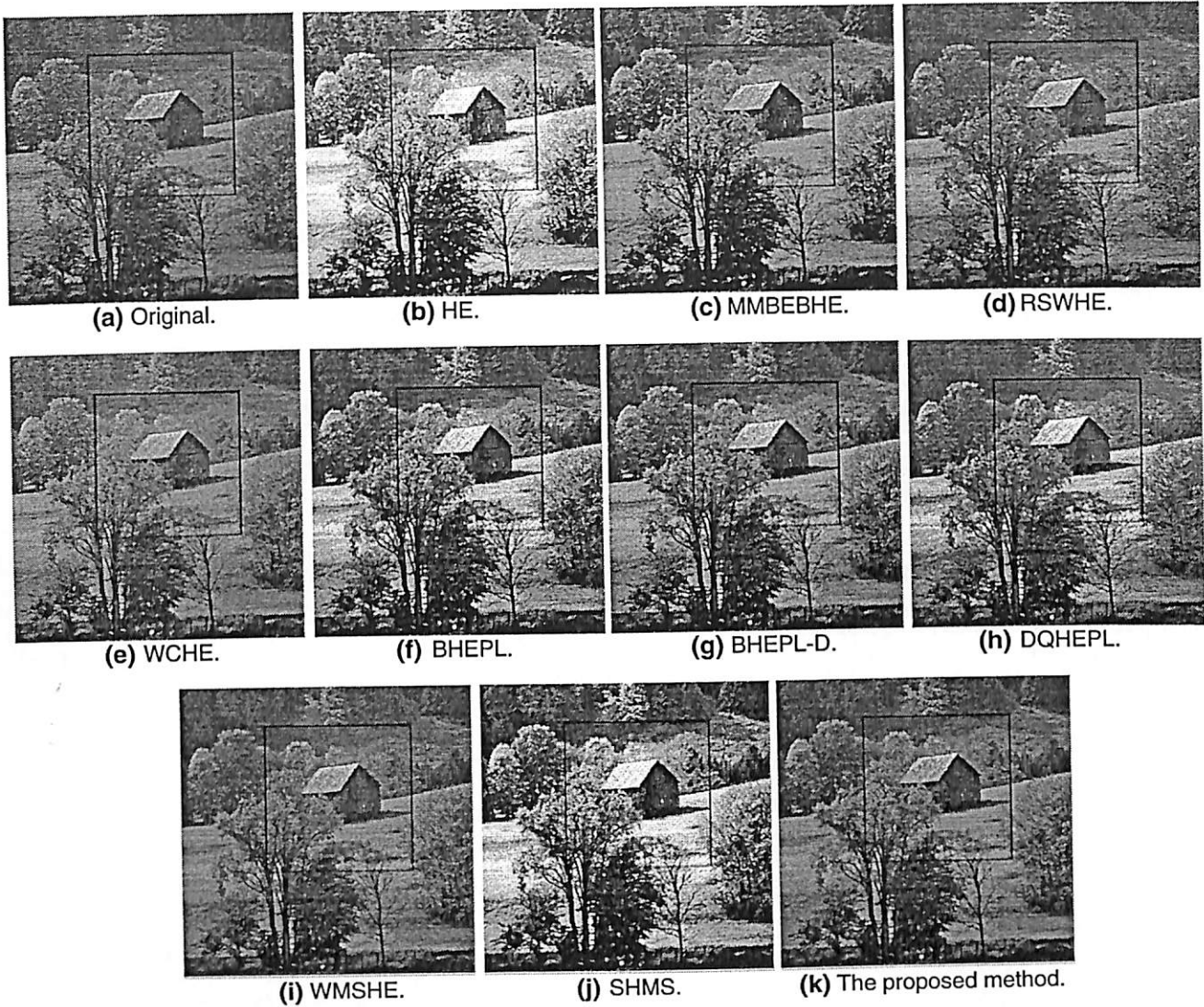


Fig. 11 Simulation results for the 'House' image (resolution 512×512)

from 467×413 to $1,024 \times 1,024$. Most of the standard images selected are obtained from [9] and [10]. On the other hand, 118 real images captured using personal digital camera are used in simulation to demonstrate the enhancement strength of the proposed method when used in real-world application. All the real images are $2,352 \times 1,568$ in dimension. It is worth noting that these 190 images used are 8-bit images.

For comparison, the HE [1], MMBEBHE [2], RSWHE [3], WCHE [4], BHEPL [5], BHEPL-D [6], DQHEPL [6], WMSHE [7], and SHMS [8] are implemented, and their output images are compared to the proposed method. Because some of these methods require the user to select their parameters, a set of values have been selected. For example, the RSWHE is implemented with 4 sub-histograms (recursion level is two) and the separating points used are the mean intensity in the histogram and/or sub-histograms. The SHMS is implemented based on the brightness preserving

bi-histogram equalization (BPBHE) method in [8]. As for the WMSHE, the recursion level is set equal to 6.

Meanwhile, the image analysis and comparison can be divided into two types, namely qualitative and quantitative analysis. The objective of qualitative analysis is to provide visual inspection on whether the enhanced images are visually acceptable to human perception, and whether they have natural appearance upon enhancement, while the quantitative analysis numerically measures the brightness, contrast, entropy, and runtime of the implemented methods.

The quantitative analyses used in this work are the average absolute mean brightness error (AAMBE) [4], average contrast (AC) [11], average entropy (AE) [12], and average processing time (AT). The AAMBE analysis is applied to measure the ability of brightness preservation of the processed images. The smaller value of the AAMBE indicates better mean preservation. The AAMBE is given as

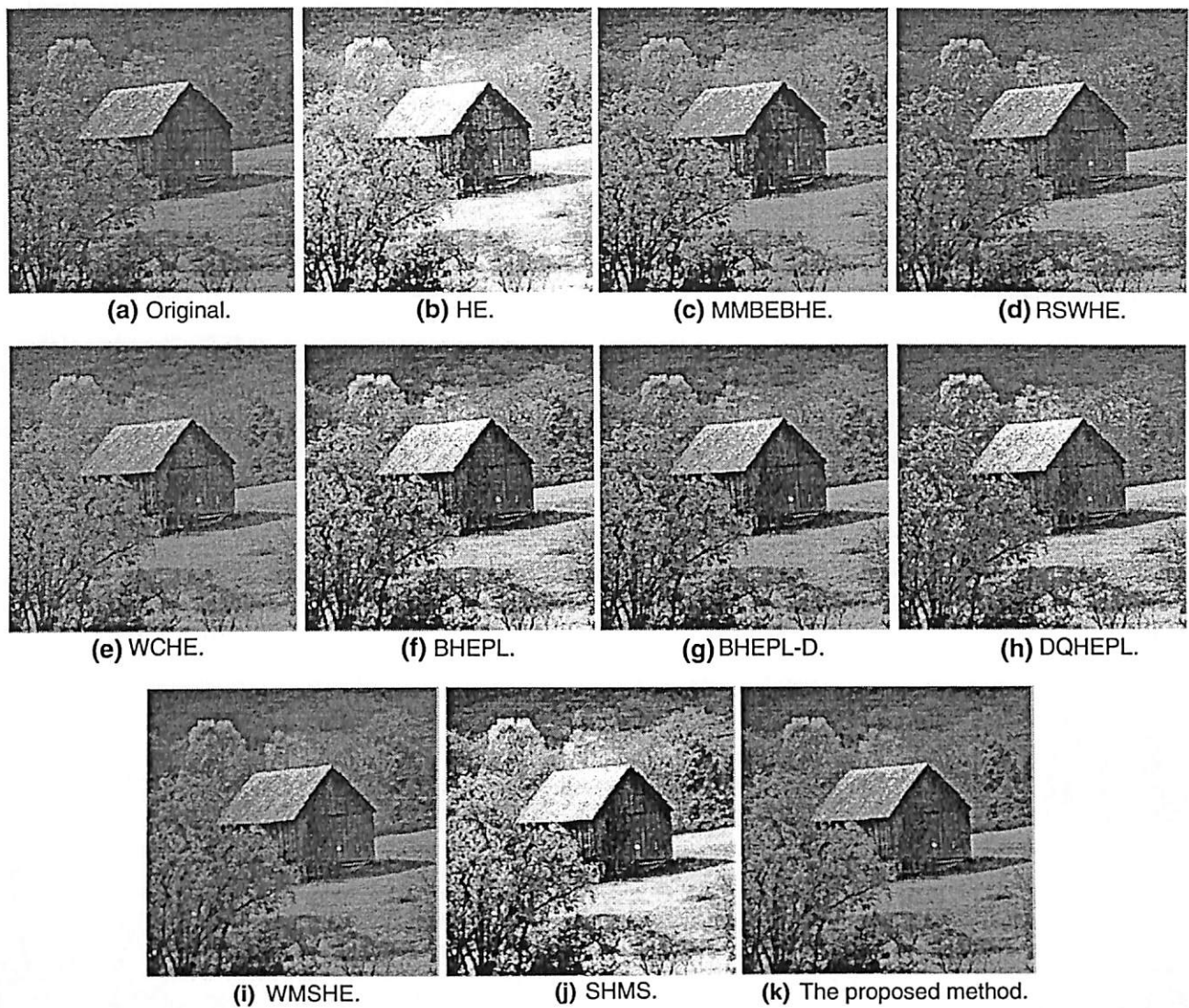


Fig. 12 Simulation results for the highlighted area of 'House' image

Table 1 Quantitative analyses for the average of the 72 standard images

Methods	AAMBE	AC	AE	AT (M/s)
Original		45.2011	6.9772	
HE	32.6774	74.1638	5.7569	0.1427
MMBEBHE	2.2928	63.8465	6.7871	0.2737
RSWHE	2.8056	51.3873	6.9308	0.3401
WCHE	11.1347	47.1228	5.2123	0.2035
BHEPL	10.2056	66.5464	6.9002	0.2434
BHEPL-D	8.3001	62.9752	6.9231	0.2462
DQHEPL	4.3054	57.8736	6.8993	0.7109
WMSHE	3.5278	49.6838	6.7849	0.4043
SHMS	13.6277	73.3359	6.8159	0.3169
Proposed method	3.6559	53.7373	6.9223	0.2999

Bold values indicate the best average results

$$\text{AAMBE} = \frac{1}{M} |E(Y) - E(X)| \quad (26)$$

where M is the total number of input images, which is 72 for standard image and 118 for real image. $E(X)$ and $E(Y)$ are the average intensity values of the input image and the enhanced image, respectively. On the other hand, the AC is used to measure the contrast enhancement of the resultant images. Therefore, the AC value of the well-contrasted image should be higher than the AC value of the test image. The AC is given as:

$$\text{AC} = \frac{1}{M} \sqrt{\sum_{k=0}^{L-1} (I_k - E(Y))^2 \times \text{PDF}(I_k)} \quad (27)$$

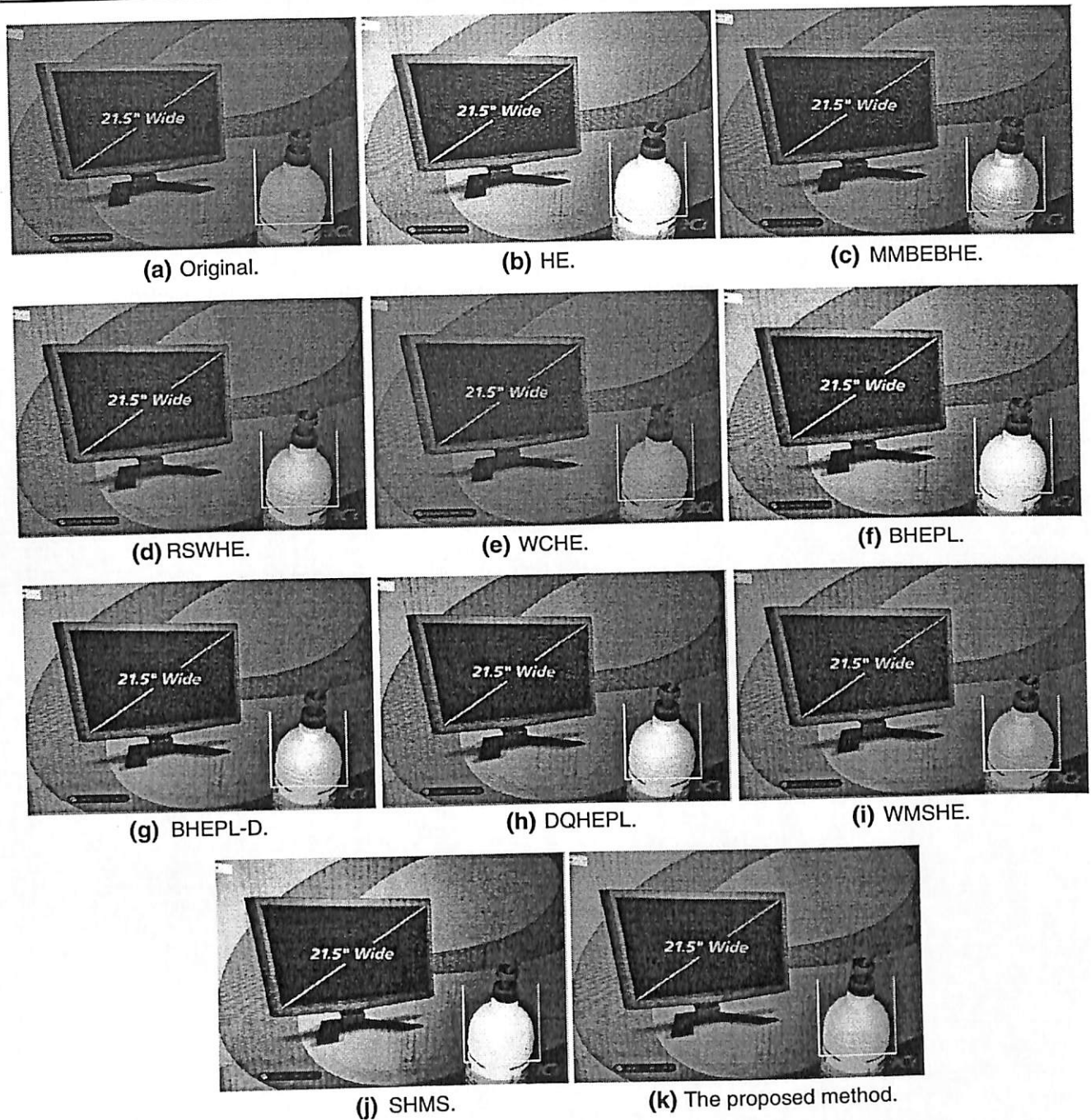


Fig. 13 Simulation results for the 'Box' image (resolution $2,352 \times 1,568$)

where PDF is the probability density function of the resultant image.

The AE is defined as the measurement of information that can be brought out from the enhanced images. As a result, the AE value of the enhanced image should be as close as possible to that of the original image in order to preserve image details. In addition, the closer AE value of ground truth and processed images would indicate that the shape of

the enhanced image histogram is maintained. Meanwhile, intensity saturation problem can be circumvented by maintaining the shape of the input image. Thus, AE is also used to measure the capability of intensity saturation prevention. The AE is defined as

$$AE = -\frac{1}{M} \sum_{k=0}^{L-1} PDF(l_k) \times \log_2(PDF(l_k)) \quad (28)$$

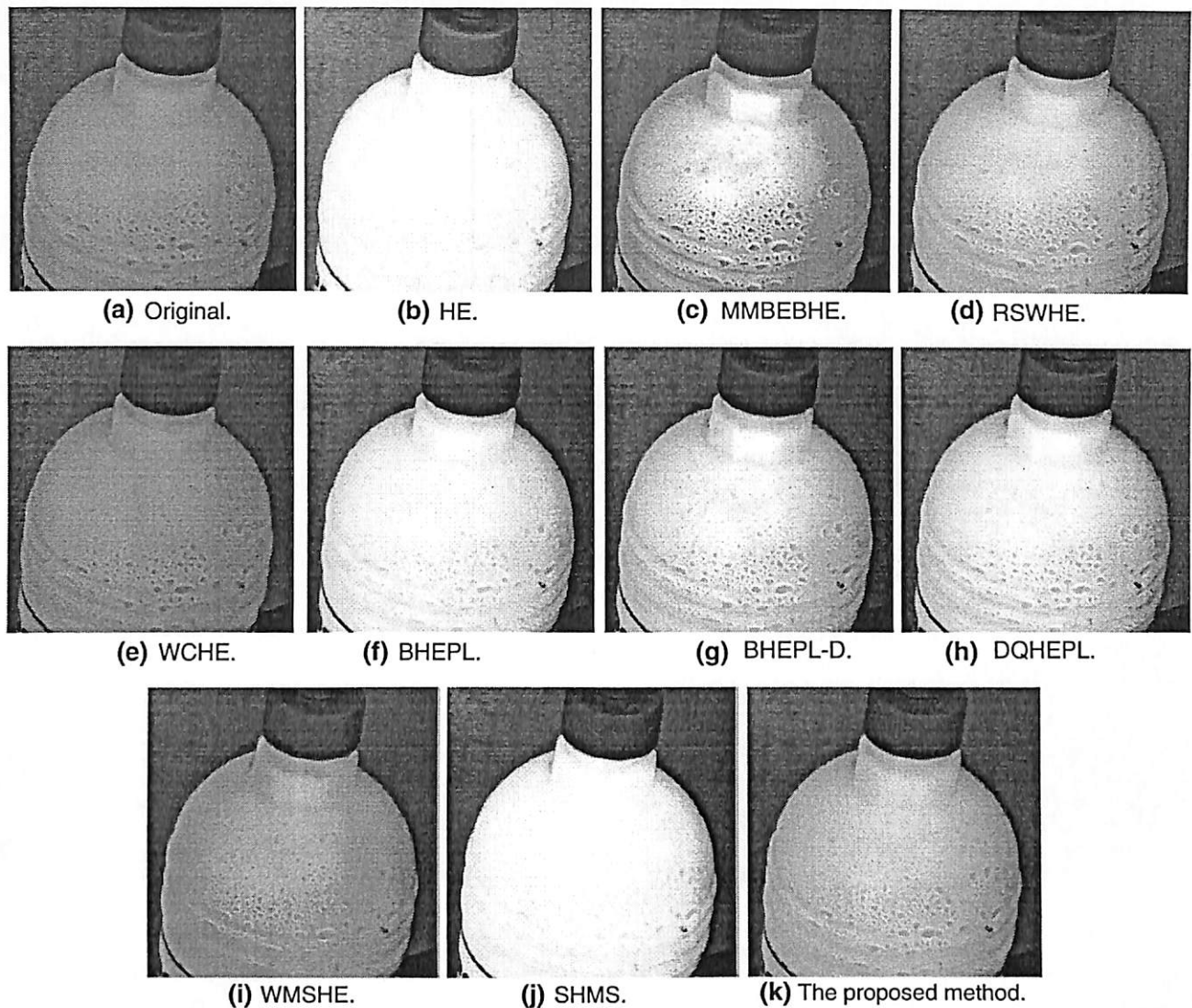


Fig. 14 Simulation results for the highlighted area of 'Box' image

Finally, AT is used to calculate the processing time. In general, the complexity of a method tends to be revealed in its runtime, as a high complexity method would most likely consume longer processing time. Here, the AT is computed in megapixels per second (M/s) and it is given as

$$AT = \frac{1}{M} \sum_{k=1}^M TPM_k \quad (29)$$

where TPM_k is the processing time per megapixels of image k , and it can be defined as

$$TPM_k = \frac{T_k}{N_k} \times 10^6 \quad \text{for } 1 \leq k \leq M \quad (30)$$

where T_k is the runtime of image k , and N_k is total number of pixels of image k .

In this study, the coding of all the implemented methods is written in Matlab 2010b and is run by a personal computer with Pentium dual CPU 4300 1.80 GHz and 1.49 gigabytes of RAM.

3.1 Standard image simulation result

For qualitative analysis, two standard images are used to demonstrate the capability of the proposed method. One of the visual results using standard test images named 'Hill' is shown in Fig. 9. From visual inspection, Fig. 9 shows that all the images become more contrasty; however, only few maintain the natural appearance composure. The enhanced images using the HE, MMBEBHE, BHEPL, BHEPL-D, DQHEPL, and SHMS in Fig. 9b, c, f, g, h, and j, respectively, suffer from saturation problem. This phenomena is clearly demonstrated

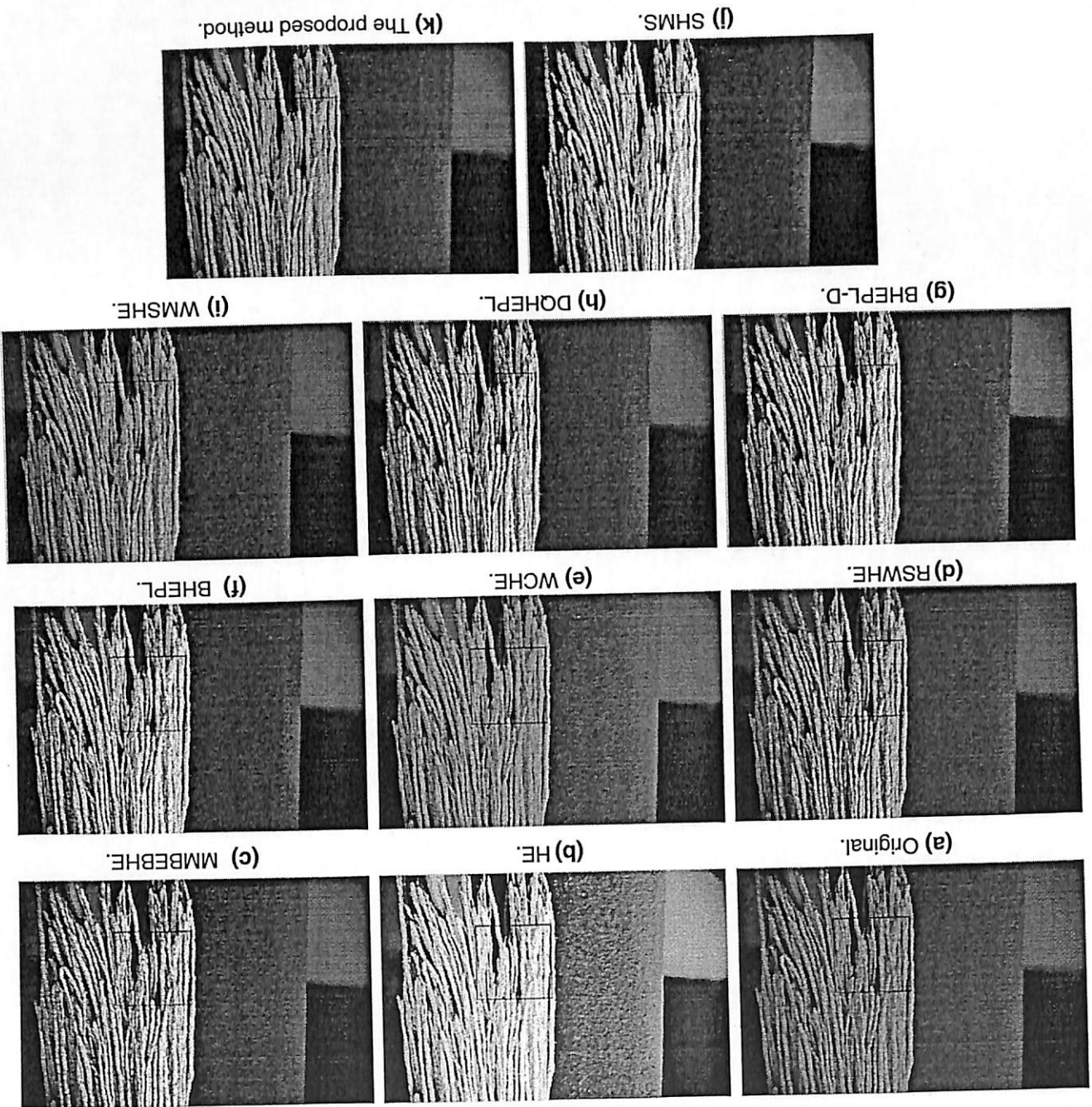


Fig. 15 Simulation results for the 'Mop' image (resolution $2,352 \times 1,568$)

by the over-brighten ground at the corresponding highlighted area as shown in Fig. 10. On the other hand, the enhanced images in Fig. 9d, e, and i for the RSWHE, WCHE, and WMSHE, respectively, produce unnatural contrast enhancement. The WCHE has darken its resultant image as shown in Fig. 9e for the whole 'Hill' image and Fig. 10e for the highlighted area of the 'Hill' image. For the RSWHE and WMSHE, both methods produce unnatural look of resultant images as could clearly be observed at the tree of the highlighted area shown in Fig. 11 similar finding are obtained. There is contrast improvement in all the images; however, only the proposed method in Fig. 11k escapes from the intensity saturation problem in highlighted area (i.e., in Fig. 12k),

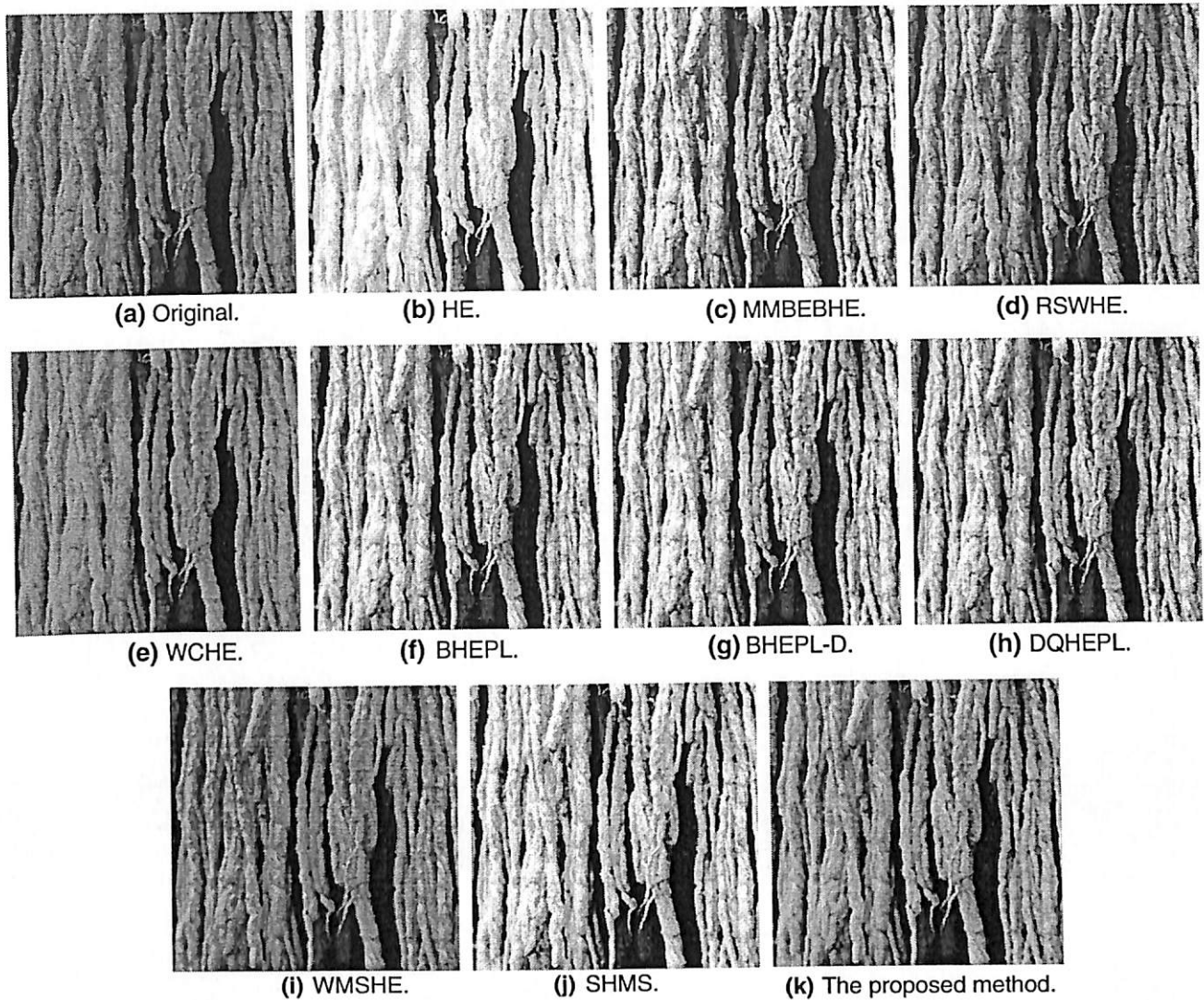


Fig. 16 Simulation results for the highlighted area of 'Mop' image

especially at the ground and the roof. The saturation problem could clearly be seen at both areas in Fig. 12 for other methods.

For the quantitative analysis, the average AAMBE, AC, AE, and AT results presented in Table 1 show that the HE produces the highest average AAMBE. This is not surprising because the HE is not designed for brightness preservation purpose. The MMBEBHE that yields the lowest AAMBE is mainly attributed to its mechanism that produces maximum brightness preservation but does not guarantee the visual quality of the enhanced image. Conversely, the proposed method has the fourth lowest average AAMBE value in this comparison, but its better visual quality as shown in Figs. 9, 11 favors the proposed method as a competitive advantage.

As for the average AC, the results in Table 1 show that all methods produce higher AC value than the input images'

AC value. Thus, all the methods successfully improve the contrast of the test images. In AE analysis, among the implemented methods, the RSWHE produces the highest average AE. Again, the proposed method comes third in terms of the average AE value, which shows its ability to prevent the intensity saturation problem.

In terms of computational complexity that is gauged by the runtime consumed, Table 1 shows that the HE has the fastest runtime, and this is not surprising because its mechanism is straightforward and does not require modification on the histogram. However, the proposed method remains competitive as it records a relatively fast average runtime. Overall, simulation results clearly demonstrate the good performance of the proposed method. It has excellent brightness preservation, good ability in intensity saturation prevention, yield superior detail preservation, and it is relatively fast in its processing.

Table 2 Quantitative analyses for the average of the 118 real images

Methods	AAMBE	AC	AE	AT (M/s)
Original		40.0943	6.8359	
HE	49.0458	74.8233	5.8023	0.0699
MMBEBHE	4.1969	54.8185	6.6435	0.1916
RSWHE	5.5932	49.8982	6.7940	0.3099
WCHE	9.5744	42.2122	5.0790	0.1289
BHEPL	22.4500	69.8726	6.7712	0.2056
BHEPL-D	15.9460	63.6431	6.7930	0.2035
DQHEPL	10.7675	59.4195	6.7718	0.6457
WMSHE	3.8983	44.7528	6.6484	0.3123
SHMS	25.2438	73.9987	6.6911	0.2353
Proposed method	6.9479	50.8093	6.7833	0.2314

Bold values indicate the best average results

3.2 Real image simulation result

To demonstrate the feasibility and usefulness of the proposed method in real-world application, it is used to enhance real images captured using personal digital camera. These images are captured under normal lighting conditions. From visual inspection, it can be clearly seen that for the image called 'Box,' the HE, MMBEBHE, RSWHE, BHEPL, BHEPL-D, DQHEPL, WMSHE, and SHMS methods in Fig. 13b–j, respectively, produce resultant images with intensity saturation problem, especially at the body of the bottle as shown in the corresponding highlighted area in Fig. 14. Meanwhile, as shown in Fig. 13e, the WCHE method fails to improve the contrast of the 'Box' image. However, the proposed method has successfully produced a more natural appearance and artifact-free enhanced image as shown in Figs. 13k and 14k. In addition, for the second real image called 'Mop,' in Fig. 15, Fig. 15k yields the best visual result with natural appearance composure; while others either suffer from intensity saturation problem or unnatural contrast enhancement which could be clearly observed in the highlighted area as shown in Fig. 16. To conclude, the proposed method outperforms the rest of the methods in comparison, both in terms of its natural appearance and artifact-free enhanced image.

Table 2 reports the average results for AAMBE, AC, AE, and AT. The results show that the proposed method has the fourth lowest average AAMBE, which explains the good mean brightness preservation. In addition, the average AC results show that all the methods improve the contrast of the input images, which is numerically indicated by the greater average AC values as compare to that of the input image. Although the RSWHE has the highest average AE, it has only a slight difference (i.e., 0.0107 in difference) with the proposed method which has the third highest AE value. How-

ever, the promising visual results coupled with relatively good and/or comparable quantitative analyses indicate the good performance of the proposed method and accentuate the good performance of the proposed method as compare to some well-known methods in literature.

4 Conclusions

In conclusion, a novel HE-based method is presented. The performance of the proposed method is compared to some well-known HE-based methods in literature. Simulation results using both standard test images and real natural images show that the proposed method has good performance in brightness preservation and image contrast enhancement, which leads to better visual quality. The proposed method has the advantages of being simple, and it does not require any tuning of parameters, which further underline its importance toward image-based application in consumer electronic products.

References

- Gonzalez, R.C., Woods, R.E.: Digital image processing, 3rd edn. Prentice, Upper Saddle River, New Jersey (2008)
- Chen, S.D., Ramli, A.R.: Minimum mean brightness error bi-histogram equalization in contrast enhancement. *IEEE Trans. Consum. Electron.* 49(4), 1310–1319 (Nov. 2003)
- Kim, M., Chung, M.G.: Recursively separated and weighted histogram equalization for brightness preservation and contrast enhancement. *IEEE Trans. Consum. Electron.* 54(3), 1389–1397 (Aug. 2008)
- Sengee, N., Choi, H.K.: Brightness preserving weight clustering histogram equalization. *IEEE Trans. Consum. Electron.* 54(3), 1329–1337 (Aug. 2008)
- Ooi, C.H., Kong, N.S.P., Ibrahim, H.: Bi-histogram equalization with a plateau limit for digital image processing. *IEEE Trans. Consum. Electron.* 55(4), 2072–2080 (Nov. 2009)
- Ooi, C.H., Isa, N.A.M.: Adaptive contrast enhancement methods with brightness preserving. *IEEE Trans. Consum. Electron.* 56(4), 2543–2551 (Nov. 2010)
- Wu, P.C., Cheng, F.C., Chen, Y.K.: A weighting mean-separated sub-histogram equalization for contrast enhancement, Presented at the 2010 international conference on the Biomedical Engineering and Computer Science (ICBECS). Wuhan, China, 23–25 April 2010
- Chang, Y.C., Chang, C.M.: A simple histogram modification scheme for contrast enhancement. *IEEE Trans. Consum. Electron.* 56(2), 737–742 (2010)
- <http://decsai.ugr.es/cvg/CG/base.htm>
- <http://sipi.usc.edu/database/>
- Menotti, D., Naiman, L., Facon, J., Araujo, A.A.: Multi-histogram equalization methods for contrast enhancement and brightness preservation. *IEEE Trans. Consum. Electron.* 53(3), 1186–1194 (Aug. 2007)
- Wang, Y., Chen, Q., Zhang, B.: Image enhancement based on equal area dualistic sub-image histogram equalization method. *IEEE Trans. Consum. Electron.* 45(1), 68–75 (Feb. 1999)

Intelligent Medical Disease Diagnosis Using Improved Hybrid Genetic Algorithm - Multilayer Perceptron Network

Fadzil Ahmad · Nor Ashidi Mat Isa · Zakaria Hussain ·
Muhammad Khusairi Osman

Received: 15 November 2012 / Accepted: 25 February 2013 / Published online: 12 March 2013
© Springer Science+Business Media New York 2013

Abstract An improved genetic algorithm procedure is introduced in this work based on the theory of the most highly fit parents (both male and female) are most likely to produce healthiest offspring. It avoids the destruction of near optimal information and promotes further search around the potential region by encouraging the exchange of highly important information among the fittest solution. A novel crossover technique called Segmented Multi-chromosome Crossover is also introduced. It maintains the information contained in gene segments and allows offspring to inherit information from multiple parent chromosomes. The improved GA is applied for the automatic and simultaneous parameter optimization and feature selection of multi-layer perceptron network in medical disease diagnosis. Compared to the previous works, the average accuracy of the proposed algorithm is the best among all algorithms for diabetes and heart dataset, and the second best for cancer dataset.

Keywords Genetic algorithm · Multi-layer perceptron network · Feature selection and intelligent medical diagnosis

F. Ahmad · N. A. Mat Isa
Imaging and Intelligent Systems Research Team (ISRT), School
of Electrical and Electronic Engineering, Universiti Sains Malaysia
(USM), 14300 Nibong Tebal, Penang, Malaysia

N. A. Mat Isa
e-mail: ashidi@eng.usm.my

F. Ahmad (✉) · Z. Hussain · M. K. Osman
Faculty of Electrical Engineering, Universiti Teknologi
Mara (UiTM), 13500 Permatang Pauh, Penang, Malaysia
e-mail: fam.ld09@student.usm.my

Z. Hussain
e-mail: zakaria183@ppinang.uitm.edu.my

M. K. Osman
e-mail: khusairi@ppinang.uitm.edu.my

Introduction

Research advancement in healthcare technology has significantly improved the quality of human health. Various techniques and engineering principles have been adopted and these attempts have resulted in new invention and innovation in the area related to the healthcare delivery such as medical imaging and instrumentation [1, 2], automatic disease diagnosis [3, 4] and functional electrical stimulation system [5, 6].

In particular, intelligent and automatic disease diagnosis has become an important healthcare research area in supporting pathologist for neoplastic disease classification, especially when dealing with large amount and high dimensional medical database. With the increase number of new reported chronic diseases nationwide, manual diagnosis which relies on pathologist experience to recognize the presence of certain pattern from the database is impractical.

Artificial neural networks (ANN) architecture, the multi-layer perceptron network (MPL) is one of the intelligent based diagnosis systems that have attracted great attention of healthcare research community. This is because of their ability to approximate an arbitrary function mapping [7] and in many cases surpass the conventional statistical technique for the prediction and classification purposes in various fields of applications [8]. However, designing MLP involved optimization of various design parameters and it is a complex task. The works of [9–11] revealed the incorporation of feature selection (FS) into the MLP and simultaneously optimize with design parameters has led to significantly improved performance.

Genetic algorithm (GA) is one of the most widely used optimization tool. It is a computational paradigm based on the survival of the fittest which mimics the natural evolution process. Unlike other search techniques that usually stuck in local convergence, GA can efficiently explore the entire search space owing to the global and multi-directional search characteristic through the iterative optimization of a set

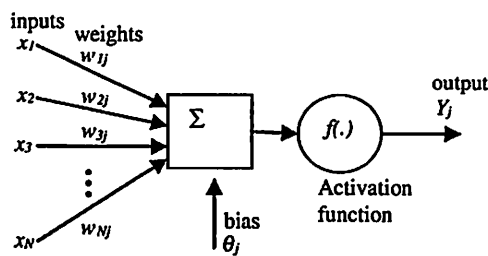


Fig. 1 Model of a node

of potential solution known as population. Unfortunately, GA is unable to fine tune the solution around the potential region due to the incapability of local search. Besides, the pre-mature convergence is another often-observed problem in GA [12]. It is a phenomenon in which the quality of the generated offspring is always being inferior to their parent. Therefore in certain optimization problem where the search space is too complex, conventional GA may not able to produce promising result.

Various attempts have been made to overcome the shortage of the conventional GA. Different techniques are promoted at different phase of GA cycles. Yang [13] introduces chromosome management scheme that divides the chromosome into several segments based on the number of feature groups for local management. Segmented mutation and crossover operator are employed to avoid the production of invalid chromosomes. Solution improvement phase is introduced in [14] where particle swarm optimization (PSO) is used to evolve the individual GA solutions during their existence. Dynamic parameter setting for the mutation and crossover operation is introduced in [15] and genes are ranked according to the frequency of their occurrence in gene subset. In [16], a population manager that preserves the

high-quality and removes the worst chromosome is introduced. In order to avoid the GA from falling at local optima and to expedite the search for global optimum, the sharing concept for crossover and mutation is proposed. Liu [17] proposes new reproduction scheme for the GA which can preserve good genetic information of parent in the new generation.

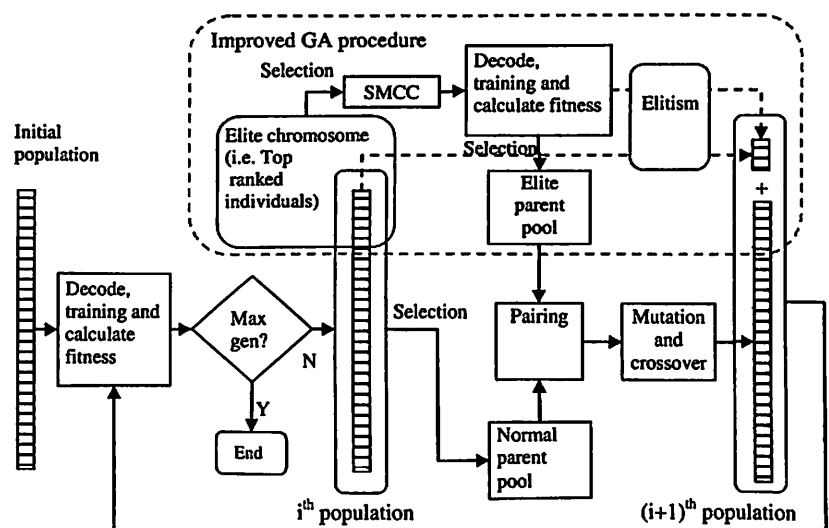
The primary objective of this study is to develop an improved hybrid GA-MLP procedure that optimize the MLP structure and improve the accuracy of the medical diagnostic classification. The procedure involves the selection of suitable feature subset, initial weight and number of hidden node of the MLP. The procedure promotes a local search around potential region instead of destroying it as in conventional GA which some of the near optimal individuals that are not selected as parents are destroyed. The secondary objective is to devise a novel crossover technique namely Segmented multi-chromosome crossover (SMCC). It considers each genetic segment in an elite chromosome and preserves the information contained in the segment. Unlike the conventional crossover that exchange information by a pair of parent chromosome, a chromosome produced by SMCC may carry gene information from more than a pair of parent chromosome.

Basic principle of MLP and GA

Multilayer perceptron network

The structure of a MLP consists of processing nodes or neurons which are grouped in layers. There are an input, an output and hidden layer(s) which size may vary upon the application. However, it has been

Fig. 2 The architecture of the proposed algorithm



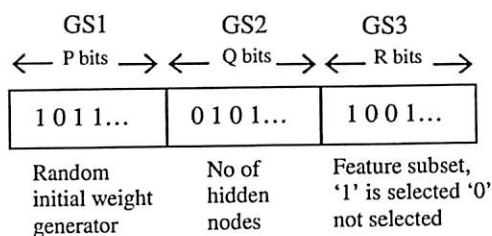


Fig. 3 The chromosome represents the network

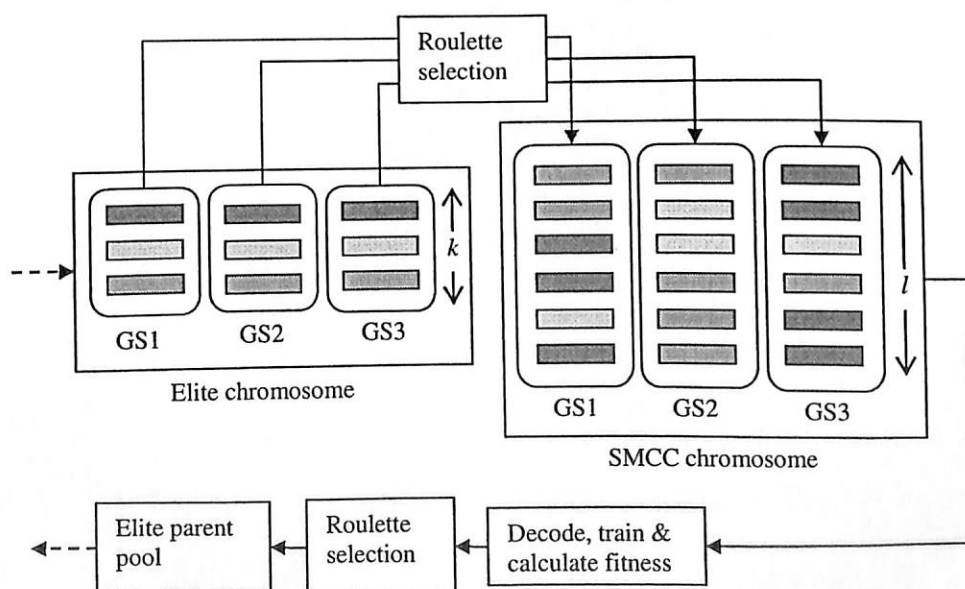
mathematically proved that one hidden layer is sufficient to approximate any function [7]. The number of input nodes is determined by the size of feature subset. Meanwhile output nodes size equals to the number of classification class.

The nodes are interconnected with links. All links have an associated weight that multiples with the propagated information in the network. Forward propagation step begins when input vector are presented at input layer. Each node calculates the activation level and the outcomes are propagated through the following hidden layers until they reach at the output layer. The output of a node is determined by weighted sum of their input and activation function as depicted in Fig. 1. It is mathematically represented in Eq. (1) where the Y_j is the output of the j^{th} node, x_i is the i^{th} input node signal, w_{ij} is the weight between the node i and j , θ_j is the bias value and f is the node's activation function.

$$Y_j = f \left(\sum_{i=1}^N w_{ij} x_i + \theta_j \right) \quad (1)$$

The most common practice is to use tangent sigmoid and linear function for hidden and output nodes respectively. No computation takes place at input nodes.

Fig. 4 The production of SMCC chromosome



This work adapts supervised learning paradigm where the actual output of the network is directly compared with a known desired output pattern to obtain the error. The gradient of the error function is then evaluated and propagated in backwards direction (from output to inner nodes) to adjust the weights. Supervised learning of ANN requires the dataset to be split into at least two set: training set, which is used at learning phase and testing set on which the final selected network is being evaluated.

Genetic algorithm

GA [18, 19] is an optimization technique that mimics the process observed in natural evolution. It is a stochastic search algorithm inspired by Darwin's evolutionary mechanics theory that involves survival of the fittest, reproduction, crossover and mutation. A population of strings known as chromosomes which represent potential solutions (also known as individuals) to an optimization problem evolves toward better solutions. Usually, solutions are encoded in binary sequence of '0' and '1', but other encodings are also possible.

Chromosomes in the initial population are randomly generated and the evolution occurs in generations. For every generation, fitness of each individual in the population is calculated. Based on the fitness value, a group of parents are stochastically selected from the current population. Depending on replacement strategy, a number of top ranking individuals are directly copied to be part of new population. This is referred as elitism. Some will go through crossover and mutation operations to form a new offspring and move to a new population. Same evolution process happens on the newly generated population in the next iteration of the algorithm. Commonly, the iteration is repeated until a predefined quality of fitness or maximum

Table 1 The attributes' label for each datasets

Label	Cancer	Diabetes	Heart
x ₁	Clump Thickness	Number of times pregnant	Age
x ₂	Uniformity of Cell Size	Plasma glucose concentration in an oral glucose tolerance test	Sex
x ₃	Uniformity of Cell Shape	Diastolic blood pressure (mmHg)	Chest pain type
x ₄	Marginal Adhesion	Triceps skin fold thickness (mm)	Resting blood pressure
x ₅	Single Epithelial Cell Size	2-Hour serum insulin (μU/ml)	Serum cholestorl in mg/dl
x ₆	Bare Nuclei	Body mass index	Fasting blood sugar
x ₇	Bland Chromatin	Diabetes pedegree function	Resting electrocardiographic results
x ₈	Normal Nucleoli	Age	Maximum heart rate achieved
x ₉	Mitoses		Exercise induced angina (1 = yes; 0 = no)
x ₁₀			ST depression induced by exercise relative to rest
x ₁₁			The slope of the peak exercise ST segment
x ₁₂			Number of major vessels (0-3) colored by flourosopy
x ₁₃			Thal: 3 = normal; 6 = fixed defect; 7 = reversable defect

generation has been reached. If all happen as expected during the process of simulated evolution, the best chromosome in the final population can turn into a highly evolved and attain optimal or near optimal solution to the problem [20].

Materials and methods

The proposed algorithm

In the proposed algorithm, GA is used to automatically determine and simultaneously optimize the number of hidden nodes, initial weight and feature subsets of MLP through the evolution process. The approach has solved the tedious process of manual determination of hidden node size as the expected end users are healthcare personnel who are unknowledgeable of the MLP design. The architecture of the proposed algorithm is depicted in Fig. 2. The major steps for its implementation are further explained as follow:

Step 1: The algorithm begins with the generation of random initial population. As illustrated in Fig. 3, the chromosome consists of three gene segments (GS) in binary format: 1) the random initial weight generator 2) the number of hidden nodes and 3) the selected feature subset.

Table 2 The LM training parameters

LM training parameter	Value
Initial learning rate	0.001
Learning rate decrease factor	0.1
Learning rate increase factor	10
Maximum Epoch	100

Step 2: Each chromosome in the population is decoded. Hidden node size and random generator in binary are converted to decimal. The unselected features are removed from the training and testing dataset.

Step 3: Information obtained from step 2 is used to initialize the MLP design parameters such as the input, hidden node size and the initial weights. The training process is then executed using the training dataset.

Step 4: The fitness of the entire trained network are calculated using test dataset. The proposed algorithm emphasizes on both the generalization performance and the complexity of the MLP as formulated by the fitness function of the GA,

$$Fitness = w \times ACC + (w - 1) \times \frac{1}{C} \quad 0 < w \leq 1 \tag{2}$$

The contribution of the weighted sum of two parameters: the test accuracy (ACC) and inverse complexity (1/C) towards the overall fitness is governed by w,

Table 3 The improved GA parameters

GA parameter	Size/value
Population	20
Max generation	40
Elite chromosome	5
SMCC chromosome	15
Elite parent pool	9
Normal parent pool	9
Elitism	2
Crossover probability	0.7
Mutation probability	0.08
w	0.75

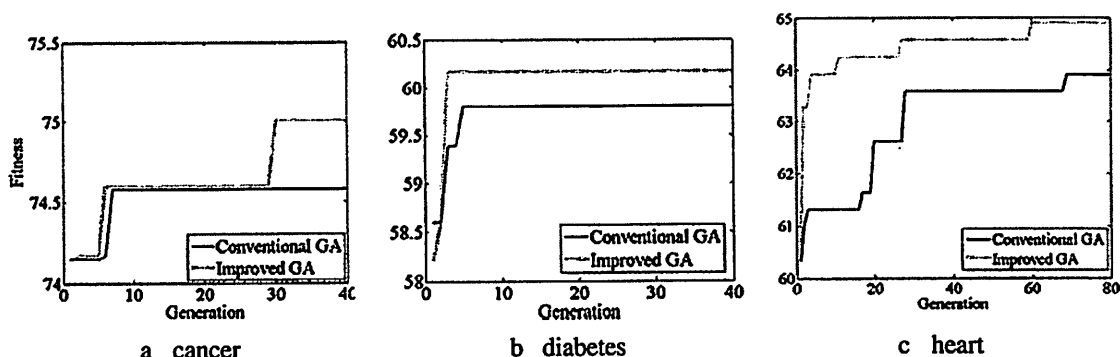


Fig. 5 Comparison of the evolution of the fitness functions for the conventional and improved GA

which is a user define parameter between 0 and 1. The larger the value of w the higher contribution of network accuracy to the overall fitness value, hence the smaller penalty imposed upon the network complexity. The ACC is calculated according to

$$ACC = 100 \times \left(\frac{N_c}{N_t} \right) \tag{3}$$

where N_c is the total number of test examples correctly classified by the network and N_t refers to the total number examples in test dataset. Meanwhile the network complexity (C), i.e. the total number of links in the network is obtained using

$$C = (\alpha \times \beta) + (\beta \times \gamma) + \beta + \gamma \tag{4}$$

where α , β and γ refer to the number of selected feature subset, hidden and output nodes respectively.

Step 5: An elite chromosome group of size k is created from the top ranked chromosomes. SMCC chromosomes (size l) are created from elite chromosomes as shown in Fig. 4. Each gene segment in SMCC is created via roulette selection of respective elite gene segment. It is assumed that the fitness of elite gene segments is equal to the fitness of its respective chromosome. By encouraging the formation and exchange of highly important information among the fittest solution, the possibility of the algorithm to effectively explore the solution space is then improved. Therefore the fitter the chromosome, the greater chance of its gene segment to be selected. SMCC chromosomes form a new set of potential solution. After the re-evaluation of fitness and selection, elite parent pool is formed.

Step 6: The best individual resulted from steps 4 and 5 (i.e. before and after SMCC) are copied directly to the

Table 4 Results for the 10 runs of heart dataset

Run number		1	2	3	4	5	6	7	8	9	10	Average
Selected Attribute, '1' is selected '0' is not selected	x_1	1	1	1	1	0	0	1	0	1	0	
	x_2	1	1	1	1	1	0	1	1	1	1	
	x_3	0	0	1	1	1	1	1	1	1	1	
	x_4	1	1	0	0	0	1	0	0	1	1	
	x_5	1	1	0	1	0	1	1	1	1	1	
	x_6	1	1	1	0	0	0	0	0	0	1	
	x_7	0	0	0	0	0	0	0	0	0	0	
	x_8	1	1	1	0	1	1	1	1	1	1	
	x_9	0	0	0	1	1	1	0	1	0	1	
	x_{10}	0	0	0	0	0	0	0	0	0	0	
	x_{11}	0	0	0	0	0	0	0	0	0	0	
	x_{12}	1	1	1	0	0	0	0	1	1	1	
	x_{13}	1	1	1	1	1	1	1	0	1	0	
% attribute use		61.5	61.5	53.8	46.2	38.5	46.2	46.2	46.2	69.2	84.6	55.4
Hidden node		1	1	1	2	2	3	2	3	4	1	2.0
Connection		13	12	11	18	20	29	20	38	58	13	23.2
Accuracy		85.5	86.8	84.2	85.5	86.8	86.8	86.8	86.8	88.2	85.5	86.3

next generation without being modified by the standard crossover and mutation process (elitism).

Step 7: Members of the elite parent pool are then paired with the members of normal pool for the standard crossover and mutation process as described in [18] and new generation is produced.

Step 8: Steps 2 to 7 are repeated until maximum generation. The individual in the final generation with the best fitness value will be selected as the final solution.

The dataset

The performance of the proposed algorithm is evaluated using three standard medical disease diagnosis datasets obtained from UCI Machine Learning Repository [21]. These include the Wisconsin breast cancer, Cleveland heart disease and Pima Indian diabetes dataset which are widely used in the study related to intelligent medical disease diagnosis. These datasets represent a real-world medical data collected from human patients and the attributes are similar to the one used by the pathologist which can be referred at Table 1. The cancer dataset consist of 699 examples, taken from the Fine Needle Aspirates of human breast tissues requires a correct diagnostic of breast lumps, whether they are either benign or malignant. The diabetes dataset, collected from 768 female Pima Indians patients, aged 21 years or older, is used to predict diabetes negative or positive. The heart dataset is used to predict heart disease based on whether or not at least one of four major vessels is reduced in diameter by more than 50 %. The heart dataset consists of 920 examples.

Experimental setup

The coding for the GA is developed in MATLAB programming environment. The Levenberg Marquart algorithm is used to train the MLP. It is accomplished via the Matlab Neural Network Toolbox [22] using the training parameters shown in Table 2. The hidden and output node's activation functions are Tan-Sigmoid and pure linear transfer functions respectively. The output classification is based on the winner-take-all approach. The P and Q bits (gene segment

Table 6 Best performance over 10 runs

	Cancer		Diabetes		Heart	
	Conv	Imp	Conv	Imp	Conv	Imp
Hidden nodes	1	4	1	1	3	4
Attribute use (%)	33.3	77.8	37.5	37.5	46.2	69.2
Connections	8	42	8	8	29	58
Accuracy (%)	99.4	100	80.7	81.3	86.8	88.2

length, Fig. 3) are set to 15 and 2. Meanwhile the R bit will be automatically set according to the size of the input attribute of the dataset. The improved GA parameters used in the experiments are provided in Table 3.

The datasets obtained from UCI Machine Learning Repository [21] is meant for general machine learning algorithm and cannot be directly learned by the MLP due to inappropriate nominal encoding and missing of attributes values [23]. The size of training and testing examples, which has significant effect on the predictive ability of the MLP are not available in the database. In fact, standard dataset development is vital for ANN learning, for a fair comparison and future replication of the experiment. Therefore, this work follows standard development of dataset as recommended in [23] by randomizing the order of the examples, replacing the missing attributes with the mean of non-missing values and re-scaling the attribute values within the range of 0 to 1 using a linear function. The first 75 % of the examples are used for training and the last 25 % for testing.

For comparison purposes, two set of experiments having common parameters setting is conducted, one with the improved GA and another one with the conventional GA.

Results and discussion

Figure 5a–c show the fitness function of the improved GA that has evolved to a better solution and required less generation to achieve the highest fitness compared to the conventional GA. This is true for all the datasets.

Table 5 Average performance over 10 runs

	Cancer		Diabetes		Heart	
	Conv	Imp	Conv	Imp	Conv	Imp
Hidden nodes	1.4	1.4	1.5	1.9	1.7	2.0
Attribute use (%)	51.1	47.8	41.3	51.3	53.8	55.4
Connections	12.9	12.9	11.4	15.3	18.7	23.2
Accuracy (%)	99.4	99.5	80.1	80.4	85.5	86.3

Table 7 The TP, TN, FN and FP of the best accuracy network

		TP	TN	FN	FP
		Cancer	Conv	640	1090
	Imp	641	1090	9	0
Diabetes	Conv	404	1133	296	87
	Imp	410	1133	290	87
Heart	Conv	318	332	72	38
	Imp	330	326	60	44

Table 8 Average sensitivity, specificity and accuracy over the 10 runs

	Cancer		Diabetes		Heart	
	Conv	Imp	Conv	Imp	Conv	Imp
Specificity	100.0	100.0	82.3	82.5	89.3	88.2
Sensitivity	99.1	99.2	79.3	79.6	82.2	84.5
Accuracy	99.4	99.5	80.1	80.4	85.5	86.3

The structure and performance of the final optimized network obtained by different run of the algorithm may slightly different due to the stochastic nature of the GA, MLP, and the complexity of solution space. This can be observed at Table 4 in a case of heart dataset. The accuracy, defined as the correct classification of the algorithm on unseen test dataset, varies from the lowest 84.2 % to the highest 88.2 %. The variation in the attribute subset and hidden node size selected by different run of the algorithm causes the production of networks with different size of connections (i.e. complexity) according to Eq. (4). Less complex networks are preferred to ease understanding of learning model, improve the predictive accuracy and to minimize the computation and hardware fabrication cost [24]. Run number 9 that achieves the highest accuracy (88.2 %) requires the most complex network (58 connections). The simplest network (11 connections) obtained at run 3 produces the lowest accuracy (84.2 %). Run number 2 that results in the second lowest complexity (12 connections) and second highest accuracy (86.8 %) provides a trade-off between the accuracy and complexity.

Tables 5 and 6 compare the performance of the conventional (conv) and improved GA (imp) in term of the average and best accuracy network over 10 runs respectively. It is obvious that the best and average accuracy of the improved GA outperforms the conventional GA for all the datasets. The highest improvement in the accuracy is observed for heart dataset, followed by the diabetes and cancer dataset. For cancer dataset, the improved GA is capable to enhance the average accuracy with similar network complexity as in conventional GA. For diabetes and heart dataset, higher average accuracy is achieved but requires more complex network.

In medical diagnostic, beside the accuracy, it is important to observe the specificity and sensitivity of the system, defined as the ability to correctly classify those with and without disease respectively and are calculated with the following equations,

$$specificity = \frac{TN}{TN + FP} \times 100\% \tag{5}$$

$$sensitivity = \frac{TP}{TP + FN} \times 100\% \tag{6}$$

The TN, FP, TP and FN are the abbreviation of the number of true negative, false positive, true positive and false negative cases respectively. The comparison of the overall TN, FP, TP and FN for the 10 runs between the improved and conventional GA for all the datasets is shown in Table 7. The average specificity, sensitivity and accuracy are provided in Table 8. In the case of diabetes dataset, the improved GA has enhanced both the specificity and sensitivity of the conventional GA by 0.2 % and 0.3 % respectively. While in the heart dataset, the specificity has significantly improved to 84.5 % from 82.2 %, but the sensitivity slightly dropped from 89.3 % to 88.2 %. Whereas for the cancer dataset, the sensitivity is slightly improved and the specificity maintains at 100 % as achieved by the conventional GA.

Table 9 compares the average classification error percentage (CEP) of this study with Alba [25] and Socha [26]. In [25], weak hybridization technique is applied to combine the GA with backpropagation (GA-BP) and GA with LM (GA-LM) for the training of ANN. Meanwhile in [26] the Ant Colony Optimization (ACO) is used instead of GA. In both works, the hidden node size is manually adjusted to 6 nodes. This is contrary to this study that the hidden node size is automatically optimized and determined by the algorithm. Besides, the FS is not considered in their works. It is obvious that, the improved hybrid GA-MLP produces the best CEP for the diabetes and heart dataset with 19.64 % and 13.65 % respectively. For the cancer dataset, GA-MLP obtained convincing result, the second best among all algorithms. The standard deviation is reasonably low, indicates the stability of the method proposed.

Table 9 Comparison of classification error percentage (CEP) obtained by GA-MLP with previous algorithms

Author Algorithm	Socha [26] ACO-BP	Socha [26] ACO-LM	Alba [25] GA-BP	Alba [25] GA-LM	This Study GA-MLP
Cancer	2.14 (1.09)	2.08 (0.68)	1.43 (4.87)	0.02 (0.11)	0.52 (0.17)
Diabetes	23.80 (1.73)	24.26(1.40)	36.36 (0.00)	28.29 (1.15)	19.64 (0.60)
Heart	18.29 (1.00)	16.53 (1.37)	54.30 (20.03)	22.66 (0.82)	13.65 (0.47)

The values in the bracket indicate the standard deviation. The best results are highlighted in bold

Conclusions

This study presents an improved hybrid GA-MLP procedure that optimizes the MLP structure and improves the accuracy of the intelligent and automatic medical diseases diagnosis system. A novel crossover technique, the SMCC that synthesizes new potential solutions by promoting the information formation and exchange among multiple chromosomes of top ranked solutions is introduced. Compared to the previous works, the average accuracy of the proposed algorithm is the best among all algorithms for diabetes and heart dataset, and the second best for cancer dataset. So far, the algorithm had been tested only on 2-class output classification problem. Further evaluation of the proposed method can be performed on higher output classification using different ANN architecture.

Acknowledgments This project is supported by Ministry of Science, Technology & Innovation Malaysia, Science fund Grant entitle "Development of Computational Intelligent Infertility Detection System based on Sperm Motility Analysis" and Universiti Sains Malaysia Research University-Postgraduate Research Grant Scheme entitled 'Genetic Algorithm-Artificial Neural Network Hybrid Intelligence'.

References

- Lee, S., Lee, T., Jin, G., and Hong, J., An Implementation of Wireless Medical Image Transmission System on Mobile Devices. *J. Med. Syst.* 32(6):471–480, 2008. doi:10.1007/s10916-008-9153-9.
- Qi, H., Kong, L., Wang, C., Miao, L. A., Hand-held Mosaicked Multispectral Imaging Device for Early Stage Pressure Ulcer Detection. *J. Med. Syst.* 35 (5):895–904, 2010. doi:10.1007/s10916-010-9508-x.
- Daliri, M., Automated Diagnosis of Alzheimer Disease using the Scale-Invariant Feature Transforms in Magnetic Resonance Images. *J. Med. Syst.* 36 (2):995–1000, 2012. doi:10.1007/s10916-011-9738-6.
- Waghlikar, K., Mangrulkar, S., Deshpande, A., Sundararajan, V., Evaluation of Fuzzy Relation Method for Medical Decision Support. *J. Med. Syst.* 36 (1):233–239, 2012. doi:10.1007/s10916-010-9472-5.
- Bani Amer, M., Al-Ebbini, L. Fuzzy Approach for Determination the Optimum Therapeutic Parameters in Neuromuscular Stimulation Systems. *J. Med. Syst.* 34 (4):435–443, 2010. doi:10.1007/s10916-009-9256-y.
- Yuan, B., Sun, G., Gomez, J., Ikemoto, Y., Gonzarlez, J., Murai, C., Acharya, U. R., Yu, W., Ino, S. The Effect of an Auxiliary Stimulation on Motor Function Restoration by FES. *J. Med. Syst.* 35 (5):855–861, 2011. doi:10.1007/s10916-010-9517-9.
- Cybenko, G., Approximation by superpositions of a sigmoidal function. *Math Control Signal* 2(4):303–314, 1989. doi:10.1007/bf02551274.
- Paliwal, M., and Kumar, U. A., Neural networks and statistical techniques: A review of applications. *Expert Syst. Appl.* 36(1):2–17, 2009.
- Tian, J., Li, M., and Chen, F., Dual-population based coevolutionary algorithm for designing RBFNN with feature selection. *Expert Syst. Appl.* 37(10):6904–6918, 2010.
- Rudy, S., and Huan, L., Neural-network feature selector. *IEEE T Neural Networ.* 8(3):654–662, 1997.
- Monirul Kabir, M., Monirul Islam, M., and Murase, K., A new wrapper feature selection approach using neural network. *Neurocomputing* 73(16–18):3273–3283, 2010.
- Rocha, M., Neves, J., Imam, I., Kodratoff, Y., El-Dessouki, A., and Ali, M., Preventing Premature Convergence to Local Optima in Genetic Algorithms via Random Offspring Generation. *Multiple Approaches to Intelligent Systems, vol 1611. Lecture Notes in Computer Science.* Springer Berlin, Heidelberg, pp. 127–136, 1999. doi:10.1007/978-3-540-48765-4_16.
- Yang, W., Li, D., and Zhu, L., An improved genetic algorithm for optimal feature subset selection from multi-character feature set. *Expert Syst. Appl.* 38(3):2733–2740, 2011.
- Marinakakis, Y., Marinaki, M., A hybrid genetic - Particle Swarm Optimization Algorithm for the vehicle routing problem. *Expert. Syst. Appl.* 37 (2):1446–1455, 2010. doi:http://dx.doi.org/10.1016/j.eswa.2009.06.085.
- Lee, C.-P., and Leu, Y., A novel hybrid feature selection method for microarray data analysis. *Appl. Soft Comput.* 11(1):208–213, 2011.
- Hsieh, S.-T., Sun, T.-Y., and Liu, C.-C., Potential offspring production strategies: An improved genetic algorithm for global numerical optimization. *Expert Syst. Appl.* 36(8):11088–11098, 2009.
- Liu, Q., Ullah, S., and Zhang, C., An improved genetic algorithm for robust permutation flowshop scheduling. *Int. J. Adv. Manuf. Tech.* 56(1):345–354, 2011. doi:10.1007/s00170-010-3149-6.
- Goldberg, D., *Genetic algorithms in search and optimization*, Istth edition. Addison-wesley, Boston, 1989.
- Whitley, D., Starkweather, T., and Bogart, C., Genetic algorithms and neural networks: optimizing connections and connectivity. *Parallel Comput.* 14(3):347–361, 1990. doi:10.1016/0167-8191(90)90086-O.
- Saxena, A., and Saad, A., Evolving an artificial neural network classifier for condition monitoring of rotating mechanical systems. *Appl. Soft Comput.* 7(1):441–454, 2007. doi:10.1016/j.asoc.2005.10.001.
- http://archive.ics.uci.edu/ml, 2010.
- Demuth, H., and Beale, M., *Neural network toolbox user's guide, version 4.* The MathWorks Inc., Natick, 2003.
- Prechelt, L., Proben1: A set of neural network benchmark problems and benchmarking rules. Technical Report, University of Karlsruhe, Karlsruhe, Germany, 1994.
- Hall, M. A., Correlation-based feature selection for machine learning. Ph.D. Thesis, Department of Computer Science, University of Waikato, Hamilton, New Zealand, 1999.
- Alba, E., and Chicano, J. F., Training Neural Networks with GA Hybrid Algorithms. In: Deb, K. (Ed.), *Genetic and Evolutionary Computation - GECCO 2004, vol 3102. Lecture Notes in Computer Science.* Springer, Berlin, pp. 852–863, 2004. doi:10.1007/978-3-540-24854-5_87.
- Socha, K., and Blum, C., An ant colony optimization algorithm for continuous optimization: application to feed-forward neural network training. *Neural Comput. Appl.* 16(3):235–247, 2007. doi:10.1007/s00521-007-0084-z.

Adaptive fuzzy contrast factor enhancement technique for low contrast and nonuniform illumination images

Khairunnisa Hasikin · Nor Ashidi Mat Isa

Received: 19 June 2012 / Revised: 18 September 2012 / Accepted: 17 October 2012
© Springer-Verlag London 2012

Abstract This paper presents a new enhancement technique using the fuzzy set theory for low contrast and nonuniform illumination images. A new parameter called the contrast factor which will provide information on the difference among the gray-level values in the local neighborhood is proposed. The contrast factor is measured by both local and global information to ensure that the fine details of the degraded image are enhanced. This parameter is used to divide the degraded image into bright and dark regions. The enhancement process is applied on gray-scale images wherein the modified Gaussian membership function is employed. The process is performed separately according to the image's respective regions. The performance of the proposed method is comparable with other state-of-the-art techniques in terms of processing time. The proposed method exhibits the best performance and defeats other methods in terms of preserving brightness and details without amplifying existing noises.

Keywords Nonuniform illumination · Underexposed · Overexposed · Bright · Dark · Contrast factor

This project is supported by Ministry of Science, Technology & Innovation Sciencefund Grant entitle "Development of Computational Intelligent Infertility Detection System based on Sperm Motility Analysis".

K. Hasikin · N. A. Mat Isa (✉)
Imaging and Intelligent System Research Team (ISRT),
School of Electrical and Electronic Engineering,
Engineering Campus, Universiti Sains Malaysia,
14300 Nibong Tebal, Penang, Malaysia
e-mail: ashidi@eng.usm.my

K. Hasikin
Department of Biomedical Engineering, Faculty of Engineering,
University of Malaya, 50603 Lembah Pantai, Kuala Lumpur, Malaysia
e-mail: khairunnisa@um.edu.my

1 Introduction

The visual quality of the most recorded images is inevitably degraded during the image acquisition process because of inadequate lighting and incorrect setting of the aperture or the shutter speed or both. Deficiencies in the image acquisition process often result in low contrast images which normally contain noisy backgrounds. In addition, difficulties in controlling lighting conditions lead to variation in image illumination and thus causing the nonlinear gray-scale intensity mapping. The effects of those defects are replicated on the range and shape of the gray-level histogram of the acquired image.

A low contrast image is characterized by the high amplitudes of its histogram components at one or several locations on the gray-scale, while staying very low in the remaining gray-scale distributions. As a result, the space for foreground histogram components (i.e., an object of interest) is compressed and image contrast is decreased.

Improving image contrast is difficult by simply stretching the histogram of the image or using simple gray-level transformations. Conventional image enhancement techniques generally obtain satisfactory results if the technique and parameters are properly selected. However, conventional techniques often fail to produce satisfactory results in dynamic and broad range of nonuniform illumination images. Therefore, improving the appearance of the image is essential in providing better input images for further image processing tasks.

Numerous studies have been published on image enhancement or also known as contrast enhancement. However, the concept of image contrast lacks a precise definition and has been interpreted as a qualitative rather than a quantitative measure of an image [1, 2]. Uncontrollable deficiencies seldom occurred during the image acquisition process causing

the vagueness and uncertainty in the acquired image. Uncertainty appears in the form of imprecise boundaries and intensities during image digitization.

Therefore, the fuzzy set theory [3–5] has been widely applied by researchers when dealing with the image enhancement. This theory is popular among the researchers because it is a suitable tool for dealing with the uncertainties, and it provides a solution to the problem of precision between classical mathematics and the inherent precision of the real world. The imprecision possessed by the acquired image can be qualitatively perceived by human reasoning. However, no specific quantification can describe imprecision. Thus, a machine may not understand the imprecision. Realizing this limitation, the fuzzy logic tools become a popular choice because it empowers a machine to mimic human reasoning [6] that suitable for image enhancement.

The remainder of this paper is organized as follows. Related studies on the enhancement process are discussed in detail in Sect. 2. The acquired image is categorized based on the new parameter called the contrast factor in Sect. 3. The proposed algorithm for fuzzification and enhancement are presented in Sect. 4. Simulation of the test images and the qualitative and quantitative comparison of the results are discussed in Sect. 5. Finally, this paper is concluded in Sect. 6.

2 Related works

The enhancement process is a preprocessing task needed to obtain a pleasant image. The most popular concept in the enhancement process is based on the modification of the image histogram. This technique is a popular choice among researchers for image enhancement because of its simplicity and capability to produce a good results over variety of images. However, many histogram-based contrast enhancement techniques have normalized image intensity which can produce a washed-out effect [7,8] on the output image or amplify the background noise, or both. Several modified histogram-based contrast enhancement techniques have been proposed to overcome this limitation. These techniques include bi-histogram equalization [9], quadrant dynamic histogram equalization [10], histogram specification [11] and brightness preserving dynamic histogram equalization [12].

These techniques process crisp histograms of images for enhancement purposes. The crisp statistics of digital images suffers from inherent limitation that excludes the inexactness of gray values. Because of this limitation, a dynamic fuzzy histogram equalization with brightness preserving is proposed by Sheet et al. [13] to overcome uncertainty in low contrast images. The fuzzy set theory is employed in this technique to handle the inexactness of gray values wherein

the histogram is partitioned based on local maxima. This equalization method uses a spanning function on the total number of pixels in the partition to perform equalization. However, implementing the fuzzy histogram is challenging on nonuniform illumination images, where bright (overexposed) and dark (underexposed) regions exist in a particular image.

Other approaches which are highly similar to human reasoning that have been applied in image enhancement deal with 'IF-THEN-ELSE' fuzzy rule-based system [14–17]. A set of neighborhood pixels forms the antecedent part of the rule in this method. The pixel to be enhanced is transformed by the consequent part of the rule. This approach incorporates human intuition to make soft decisions on each condition. However, this method suffers from high computational time and difficulty in generating a fuzzy rule. The consequent part of the rule will only be executed if the prior rule is accomplished, thus this technique is difficult to implement in real-time applications.

Besides that, there are other approaches that use several pixel properties such as gray tone or color intensity. These approaches are modeled into a fuzzy set using the membership function. An intensification (INT) operator is applied globally to modify the membership function to reduce fuzziness and increase image contrast [18]. This approach transforms membership values that are above a predetermined threshold to much higher values or modify the membership values that are lower than the threshold value to much lower values in a nonlinear manner. Thus, a dynamic range of images with good contrast can be obtained. However, the INT operator solely depends on the membership function and needs to be continuously applied to the image to attain the desired enhancement.

This limitation is then improved using a Gaussian type of fuzzification function that contains a single fuzzifier and a new intensification operator (NINT) [1]. The fuzzifier is obtained by maximizing the fuzzy contrast. The NINT does not change uniformly because the membership function is marginally changing. Thus computational time is reduced compared with INT.

Although most image enhancement techniques can improve image brightness, several deficiencies are still present in these techniques such as loss of image contrast and details. Thus, various enhancement techniques are proposed by optimizing the information contained in the image which will be used in adjusting gray-level transformation function [19–22]. The optimization of image content includes optimizing entropy [20,23], index of fuzziness [21,24] or a combination of both [25] and intuitionistic fuzzy [20,22].

An objective measure called exposure has been proposed by [26] which will provide an amount of lighting exposure to the image. The image can be divided into underexposed



Fig. 1 Example of underexposed (i.e., black circle) and overexposed (i.e., dotted black circle) regions

and overexposed regions based on this objective measure. Only a few studies address the issue of nonuniform lighting in an image in the context of image enhancement [26–29]. The objective measure is constructed by involving the entropy, the contrast and visual factor of the image. Minimizing this objective measure successfully enhances an image.

However, the image enhancement process performed by optimizing these quantitative measures requires an optimization procedure that is conducted repeatedly to improve the image quality. Therefore, the process needs a more complicated optimization procedure in addition to the enhancement process which can increase computational time.

Attempts have been made to locally enhance the image to increase image details without involving an additional optimization process [30–34]. The local contrast or local entropy in small regions is enhanced in this technique while preventing an increase in global contrast at the same time. Fine edges neglected in global enhancement are enhanced. The clarity of the enhanced image is improved. However, noises and artifacts are enhanced, and multiplex elements of the background are distorted during the enhancement process.

Realizing the fact, a new contrast enhancement technique has been proposed to improve the brightness of the image by considering nonuniform lighting that normally exists in the acquired image (Fig. 1). The major contribution of this paper is the development of the new enhancement technique that preserves image details by calculating differences in gray-level values in the local neighborhood.

3 Image classification based on contrast factor

Most recorded images suffer from low contrast and appear nonhomogeneous in terms of illumination. The image appears darker and brighter when intensity distribution

mainly accumulates at the lower and upper parts of the histogram, respectively. These conditions can be observed in their histograms which do not occupy the entire dynamic range of the intensity distribution. The histogram only occupies a certain area in the lower or upper parts of the total range of the histogram distribution. The image is blurred, poor in contrast and has image details that are hardly interpreted in both cases.

When the recorded image appears dark, its neighborhood pixels are close to the least available dynamic range, and it can be considered as an underexposed image. For a bright image, its neighborhood pixels are found in the highest of available dynamic range and the image is known as an overexposed image.

However, we seldom encounter a solely overexposed (bright) image or a solely underexposed (dark) image. Most of the recorded images are mixed wherein underexposed, overexposed or a combination of both regions are found in one image. As a result, the brightness of image becomes uneven and the image is identified as having nonuniform illumination. This type of image has poor contrast which means that the ratio between the brightest and darkest intensities in the image is reduced. The human perception of the concept of image contrast does not completely correspond to this problem. Therefore, the image contrast is not a preferable reference parameter in enhancing an image.

A new parameter called “contrast factor” is introduced to overcome the aforementioned drawbacks. This parameter indicates the differences among the gray levels for each pixel in the neighborhood window, W_{mn} . The contrast factor is calculated by:

$$CF_{(X_{i,j})} = \frac{\sum_{(i,j) \in W_{i,j}} (X_{i,j} - \bar{X}_{W_{i,j}})^2_{i,j \in W_{i,j}}}{\sum_{(i,j) \in W_{i,j}} \sigma_{\bar{X}_{W_{i,j}}}^2} \quad (1)$$

where $X_{i,j}$ indicates the gray-level values (i.e., intensities) of the image, $\bar{X}_{W_{i,j}}$ represents local average gray level value in the $W_{i,j}$ window and $\sigma_{\bar{X}_{W_{i,j}}}^2$ represents the local deviation in the $W_{i,j}$ window.

The contrast factor is derived based on the difference in luminance (intensity) with the average of local neighborhood luminance to preserve details in the image. The difference is then divided by the standard deviation of the local neighborhood luminance to obtain an overall difference ratio in the local neighborhood window luminance. The value of contrast factor is between [0 1]. The image looks natural and pleasing if the contrast factor is close to 0.5. The image is low in contrast and should be enhanced if the contrast factor deviates far from 0.5.

The image is considered to be a mixed-type image. Thus attempts have been made to divide the image into overexposed and underexposed regions by introducing a new

threshold, T . This threshold is defined to divide the image into two regions where enhancement is conducted separately according to its respective regions as given in Eq. (2).

$$T = L(1 - CF) \quad (2)$$

where L represents the total number of gray levels in an image. The threshold divides the gray levels into two regions namely the dark (i.e., underexposed) region which is in the range $[0, T - 1]$ and bright (i.e., overexposed) region which is in the range $[T, L - 1]$.

4 Image fuzzification, enhancement and defuzzification

Images from California Institute of Technology database [35] are used for the enhancement purposes in this study. As discussed in Sect. 1, that the image pixels are too vague and uncertain to determine whether they are bright or dark pixels. In addition, the human perception of image contrast is subjective and hard to interpret when machines are used. Therefore, the fuzzy set theory has been employed in dealing with the image pixels ambiguity by associating a degree of belonging to a particular property.

The fuzzy enhancement involves three stages namely image fuzzification, modification of membership values for image enhancement and image defuzzification as shown in Fig. 2. In the image fuzzification stage, the recorded image of size $C \times R$ in the image fuzzification stage has intensity levels $X_{i,j}$ in the range of $[0L - 1]$ which is considered as a collection of fuzzy singletons in the fuzzy set notation.

$$I = \cup\{v(X_{i,j})\} = \{v_{i,j}/X_{i,j} \quad i = 1, 2, \dots, C; \\ j = 1, 2, \dots, R \quad (3)$$

where $v(X_{i,j})$ or $v_{i,j}/X_{i,j}$ represents the membership or grade of belonging $v_{i,j}$ of $X_{i,j}$ being the grayscale intensity at the (i, j) th pixel. C and R are the number of rows and columns in the recorded image, respectively.

The original recorded image in the spatial domain will be converted to a fuzzy domain using a specific membership function according to its region (i.e., dark or bright regions). The gray-level histogram in this study is assumed to be distributed in Gaussian where the gray levels are clus-

tered around a single maximum throughout the entire range of the histogram. Therefore, a Gaussian membership function is employed to fuzzify the image to obtain a smooth and differentiable fuzzy model [36].

However, the image is poor in contrast and has nonuniform illumination, thus applying a single Gaussian membership is not suitable to fuzzify the entire region in the image. Different membership functions must be applied accordingly to ensure that they can provide a perfect function to fuzzify the dark and bright regions in the image.

The Gaussian functions are modified to fuzzify dark (i.e., underexposed) and bright (overexposed) regions separately as derived in Eqs. (4) and (5).

$$v(X_u) = \left| X_{\max} - \exp \left[-\frac{(X_{\max} - (X_{\text{avg}} - X_u))^2}{\tau_h^2} \right] \right| \\ \text{for } X < T \quad (4)$$

$$v(X_o) = \left| \exp \left[-\frac{(X_{\max} - (X_{\text{avg}} - (L - X_o)))^2}{\tau_h^2} \right] - X_{\min} \right| \\ \text{for } X \geq T \quad (5)$$

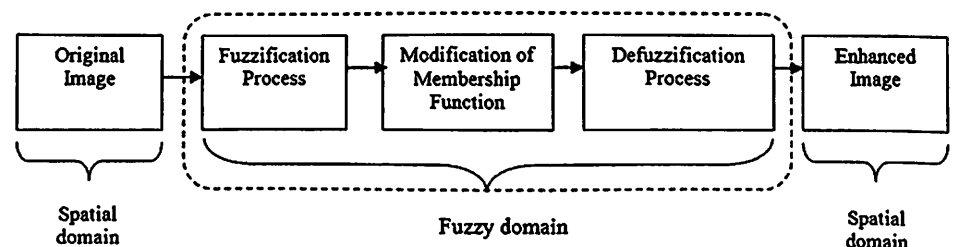
where X_u indicates the gray levels in the underexposed region in the range $[0, T - 1]$, X_o indicates the gray levels in the overexposed region in the range $[T, L - 1]$ and τ_h^2 is the membership factor, calculated using Eq. (6):

$$\tau_h^2 = \alpha \left[\frac{\sum_{X=0}^{L-1} [(X_{\max} - \sigma_X) - X]^4 p(X)}{\sum_{X=0}^{L-1} [(X_{\max} - \sigma_X) - X]^2 p(X)} \right] \quad (6)$$

where σ_X is the standard deviation of the original image. The α and $p(X)$ are the fuzzified factor and histogram of the gray levels, respectively. The membership function in Eq. (4) operates in the region below threshold T , whereas the membership function in Eq. (5) operates in the region above the threshold T .

The membership function is then modified according to its respective region to enhance the image once the image is converted into the fuzzy domain. Different functions will be used for enhancing both regions (i.e., dark and bright regions) as applied in the fuzzification process. The parametric sigmoid functions for enhancing the dark (underexposed) and bright (overexposed) regions are given

Fig. 2 Block diagram of fuzzy enhancement process



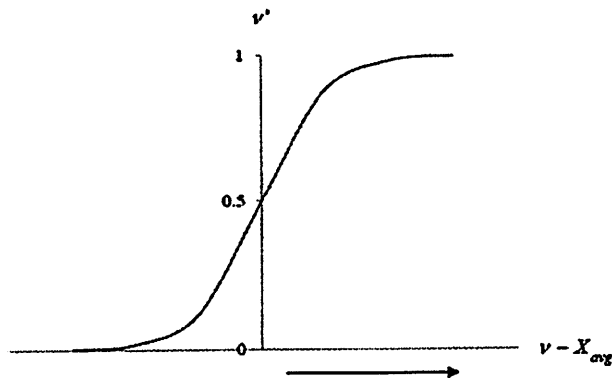


Fig. 3 Example of the sigmoid function

by Eqs. (7) and (8), respectively:

$$\nu/(X_u) = \frac{1}{1 + e^{-[-g(\nu(X_u) - X_{avg})]}} \quad (7)$$

$$\nu/(X_o) = \frac{1}{1 + e^{-[-h(\nu(X_o) - X_{avg})]}} \quad (8)$$

where g and h are the enhanced factors that are used to enhance the fuzzified image. The sigmoid function is chosen because it will achieve a value close to '1' if the exponential parameters presented in Eqs. (7) and (8) are approaching infinity (Fig. 3). Therefore, the unsaturated intensity of the fuzzified image is ensured by this sigmoid function. The aforementioned have modified the membership function and enhancing the original recorded image. Both regions are combined to produce an enhanced image once the fuzzified image is modified,

The modified membership functions are defuzzified using their respective inverse membership functions.

$$Y = \begin{cases} \nu_u^{-1}(X) & \forall X < T \\ \nu_o^{-1}(X) & \forall X \geq T \end{cases} \quad (9)$$

The gray levels in the underexposed region are scaled back in the range between $[0, T - 1]$, whereas the gray levels in the overexposed region are translated and scaled in the range between $[T : L - 1]$. Both regions are then combined into a single region for enhancing the image. The flow chart of the complete enhancement process is illustrated in Fig. 4.

The proposed techniques can be also applied in color image enhancement. A proper color space model that can be used is Hue, Saturation and Intensity color model. This model decouples the chromatic information from the achromatic information. Color artifacts may be introduced if the enhancement is directly conducted in three color spaces of red, green and blue (RGB).

5 Results and discussions

The experimental results using of standard images obtained from the California Institute of Technology Computational Vision database [35] is presented in this section. We demonstrated the performance of the developed algorithm compared with other existing fuzzy gray-scale enhancement techniques.

The performances of the proposed method are compared with the other techniques that also implement the fuzzy theory in image enhancement for fair comparison. Five groups of fuzzy image enhancement techniques exist as discussed in Sect. 1. These techniques include the conventional approach of the (NINT) [1], the application of the fuzzy IF-THEN rules (FRB) [14], the fuzzy objective measure (FOM) [26], the fuzzy local enhancement (FL) [31] and the fuzzy histogram-based equalization technique (FHE) [13].

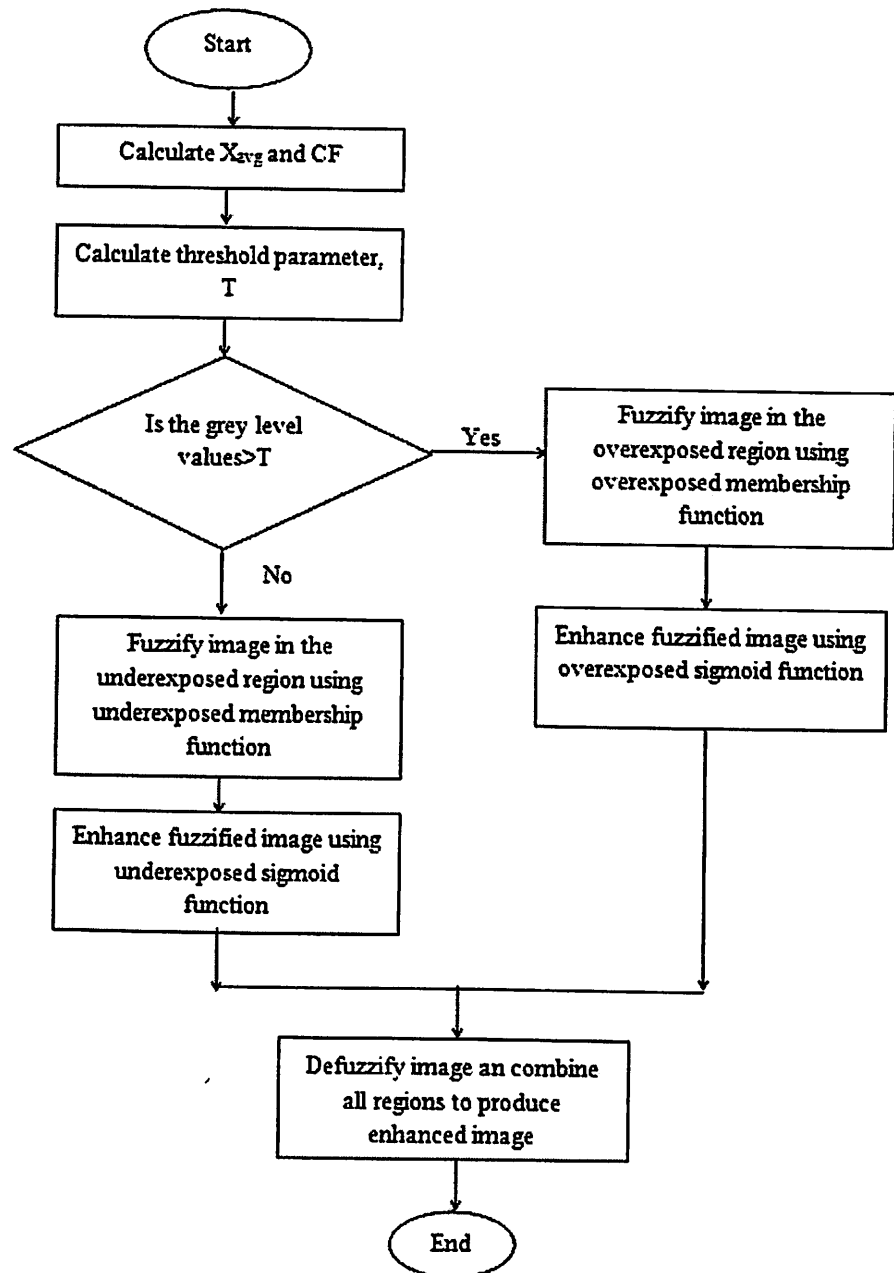
Even though the NINT [1] and the FOM [26] approaches enhanced the image in color model, both are worth consideration because only the gray levels are used without modifying the color components in the image. Hence, we selected these two methods in literature because both are similar approach with our technique, which only adjusts the gray levels to enhance the recorded image.

The FRB [14] and FL [31] methods are relevant compared with our proposed method because these techniques are related to our local contrast factor measurement. Both methods enhance the image by using the information contained in the local neighborhood which can be related to our approach. We also considered the histogram-based approach by choosing the FHE [13] method because it dealt with the crisp histogram of the images that considers the inexactness of the gray-level values to enhance contrast.

The experimental results are compared qualitatively and quantitatively with the aforementioned techniques. The goal of this study is to propose a new computationally fast enhancement technique. In addition, the luminance of the enhanced image is required to be as close to the original image as possible to preserve image brightness and details while suppressing noises. Enhancement is conducted based on global and local information to preserve the details and mean luminance of the image according to the aforementioned requirements. Qualitative analysis is performed by visual comparison, whereas the luminance distortion (LD), entropy (E), peak signal-to-noise ratio (PSNR) and processing time (t) are chosen for quantitative analysis.

LD is selected because it is proposed by a previous study [37] that considered the correlation of mean luminance between the enhanced and the original images. LD is given in Eq.(10).

Fig. 4 Flowchart of the complete enhancement process



$$LD = \frac{2\bar{X}\bar{Y}}{(\bar{X})^2 + (\bar{Y})^2} \quad (10)$$

where

$$\bar{X} = \frac{1}{L} \sum_{i=1}^L x_i \quad (11)$$

$$\bar{Y} = \frac{1}{L} \sum_{i=1}^L y_i \quad (12)$$

X_i and Y_i are the gray-level values for the original and enhanced images, respectively. The LD is in the range [0 1]. The mean luminance of the enhanced image is almost similar to the original image if the LD approaching '1' which means that $\bar{X} \approx \bar{Y}$. Thus the brightness is preserved.

In addition to the brightness preservation, the details of the image also need to be maintained. Therefore, the entropy analysis is used to measure the capability of detail preservation. A higher entropy indicates the higher ability of the

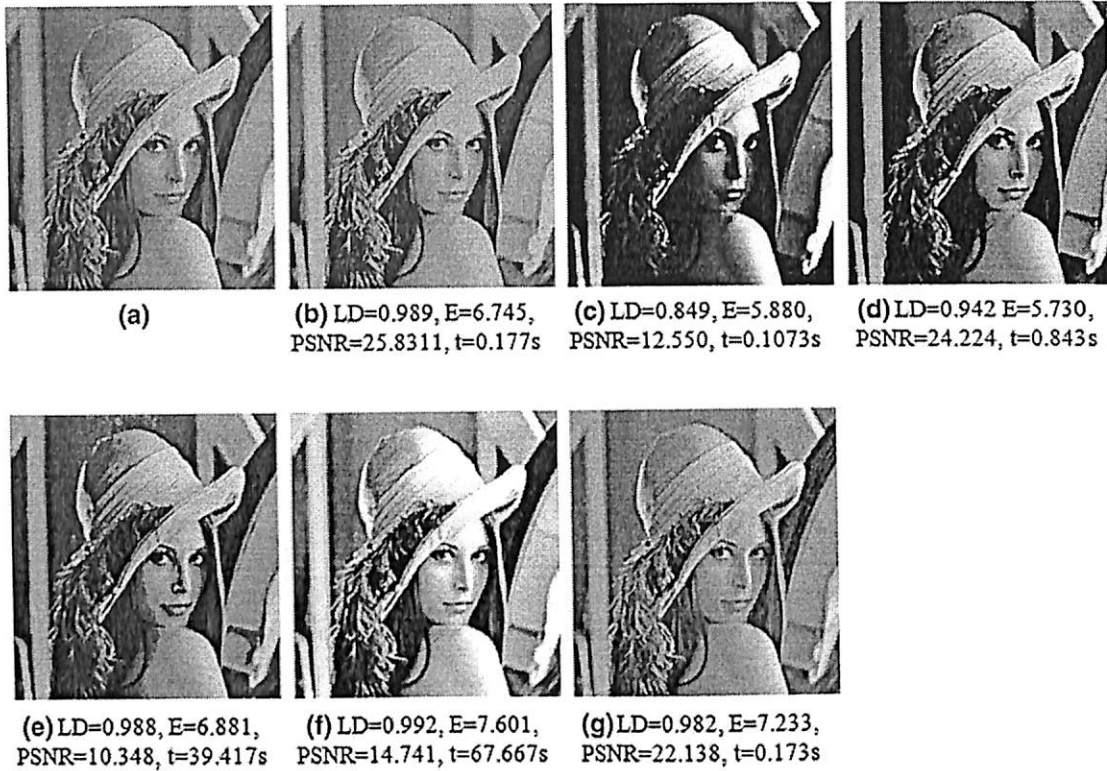


Fig. 5 Comparison of enhancement results a original 'Lena' image, b proposed method, c NINT [1], d FOM [25], e FRB [13], f FL [30], g FHE [12]

proposed method to overcome intensity saturation problems and preserve more details of the image. The entropy of the enhanced image is given by:

$$E = \sum_{i=0}^{L-1} p(Y_i) \times \log_2 p(Y_i) \quad (13)$$

where $p(Y_i)$ is the probability of enhanced gray levels Y_i

The enhancement technique should not significantly amplify the noise level. Thus PSNR analysis is employed. Good enhancement techniques should be able to increase image brightness without enhancing existing noises in the recorded image. The PSNR is calculated using Eq. (14).

$$\text{PSNR} = 10 \log_{10} (L - 1)^2 / \text{MSE} \quad (14)$$

$$\text{MSE} = \sqrt{\frac{\sum_{i=1}^M \sum_{j=1}^N (Y_{i,j} - X_{i,j})^2}{C \times R}} \quad (15)$$

The last analysis is the processing time, which is used to measure the computational complexity introduced by the proposed method. For consistency, 100 standard images size 400×264 pixels are used for each method. The samples are processed in the Matlab R2010a environment using an Intel Core 2 CPU 2 GHz with 2.49 GB RAM.

The proposed method should produce the enhanced image with an LD is close to '1' and have a high entropy to ensure brightness and to preserve details. In addition, the enhanced image is expected to have a high PSNR value and can be executed with short computational time.

Several low contrast images which are actually underexposed were used for qualitative evaluation. These images are "Lena", "Room", and "Fruits", as shown, respectively in Figs. 5, 6, and 7. The figures show the enhanced images from the proposed approach compared with other existing state-of-the-art approaches. The enhanced images from the NINT [(Figs. 5c, 6c, 7c) and FOM (Figs. 5d, 6d, 7d)] approaches have resulted in darker images compared with the original images. Image contrast is decreased. Thus, distinguishing the foreground from the background is difficult. The mean luminance of the enhanced images for both methods deviates from the mean luminance of the original image, thus resulting in lower LD values. The lower LD values indicate that those methods are unable to preserve the mean brightness of the original image.

In terms of detail preservation, the entropy values of the enhanced image using the NINT approach in Figs. 5, 6, and 7 are 5.880, 4.673, and 4.327, respectively. The NINT approach produced the lowest entropy value compared with the other methods. Most of the areas in the enhanced image

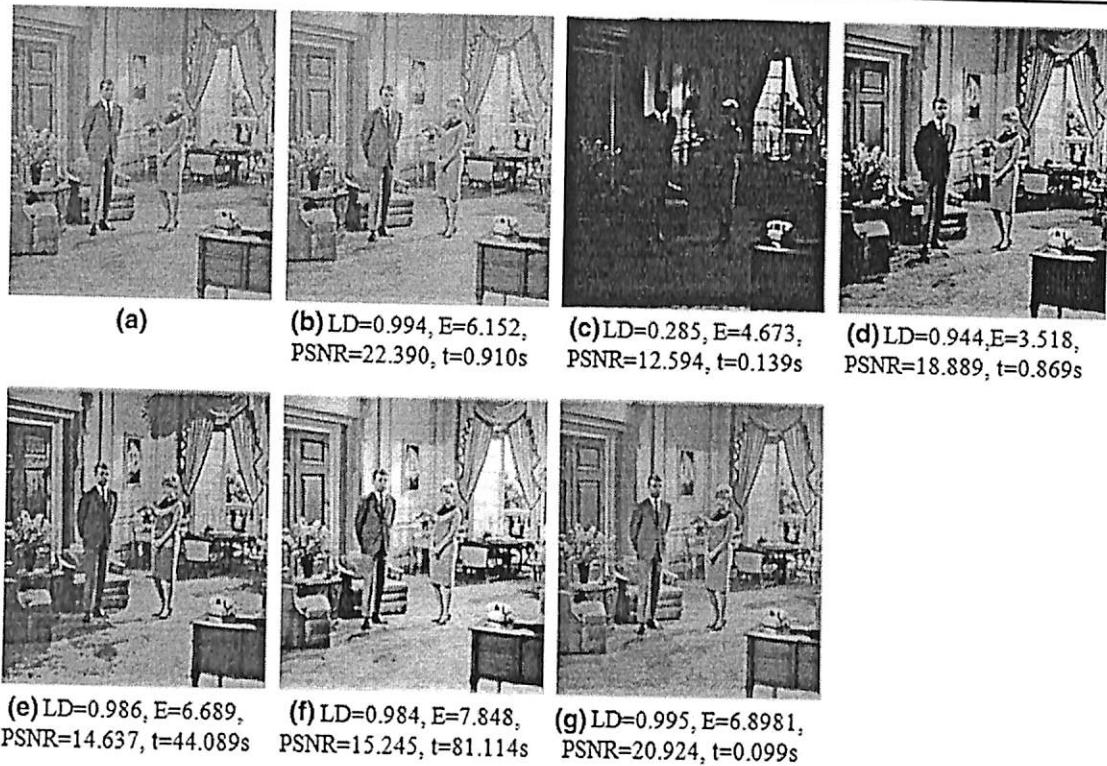


Fig. 6 Comparison of enhancement results a original 'room' image, b proposed method, c NINT [1], d FOM [25], e FRB [13], f FL [30], g FHE [12]

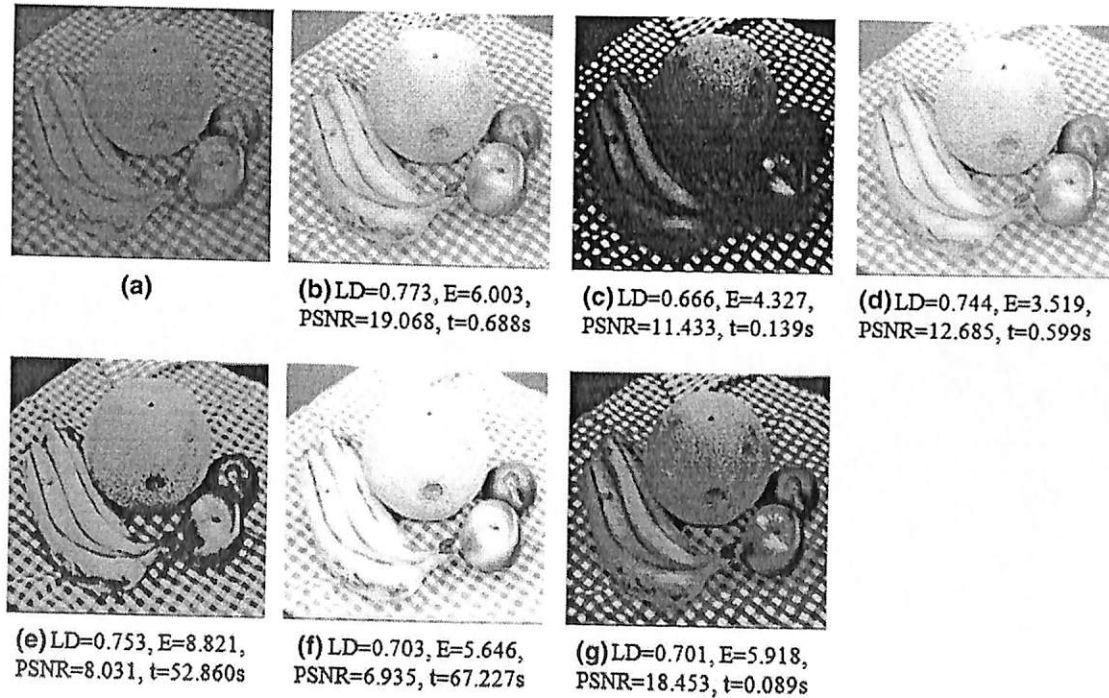


Fig. 7 Comparison of enhancement results a original 'fruits' image, b proposed method, c NINT [1], d FOM [25], e FRB [13], f FL [30], g FHE [12]

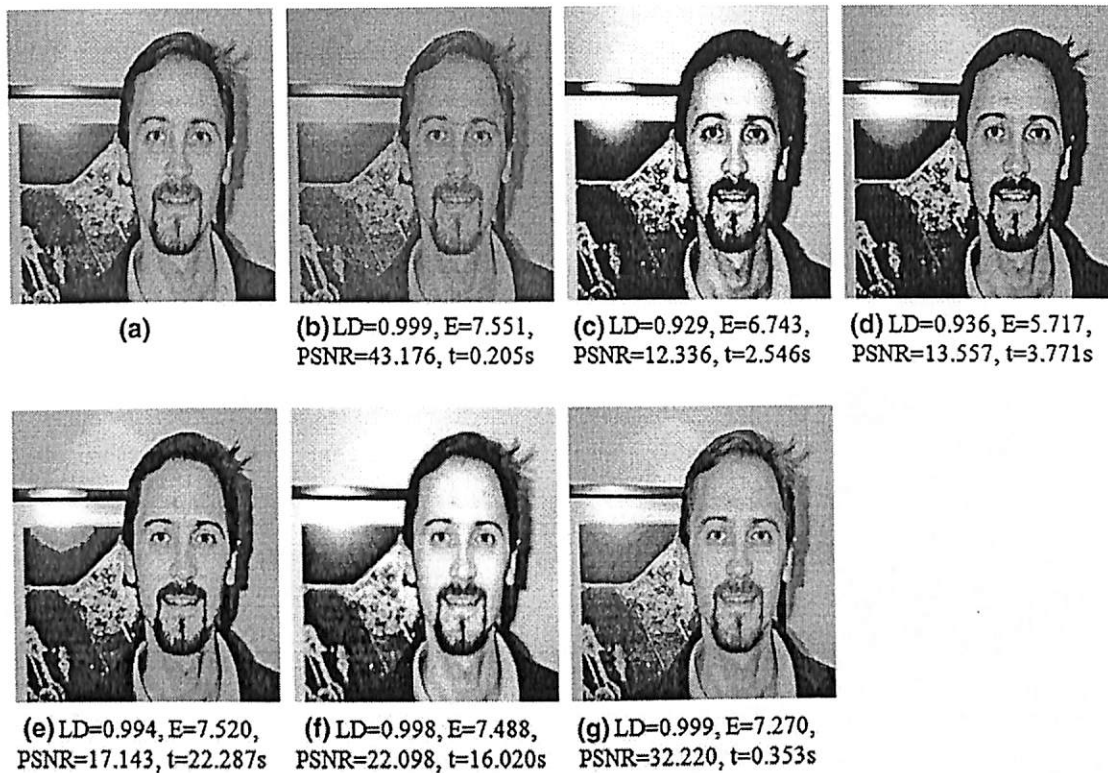


Fig. 8 Comparison of enhancement results a original 'Man 1' image, b proposed method, c NINT [1], d FOM [25], e FRB [13], f FL [30], g FHE [12]

appear dark, with poor brightness. As such, the details of the image are not apparent and cannot be interpreted by the eyes. The NINT technique underenhanced certain regions because the underexposed and overexposed regions of the image were not considered during processing. The NINT technique enhances the image without preserving the details and contrast of the original image. The algorithm also enhances existing noise in the original image, resulting in a lower PSNR value.

The less-enhanced and slightly saturated images are produced through the FRB method, as shown in Figs. 5e, 6e, 7e with PSNR values of 10.348, 14.637, and 8.031, respectively. Conversely, the proposed method enhanced the image while preserving brightness as shown by the highest LD value. The proposed method did not enhance existing noise, as indicated by the highest PSNR value among the different methods.

Figures 5, 6, 7 also show how the FL method overenhanced the original images, resulting in brighter and unnatural images. Our proposed method can be executed quite fast, while preserving details and brightness.

The proposed method is analyzed using mixed-type images to demonstrate its effectiveness in image enhancement. A mixed-type image consists of underexposed (mostly at the image background) and overexposed (mostly at the image foreground) regions in a single image, as shown in

Figs. 8a, 9a, 10a, 11a, and 12a. Our goal is to enhance the image without overenhancing the overexposed region or underenhancing the underexposed region or both. Thus, Figs. 8, 9, 10, 11, 12 are compared with other state-of-the-art methods.

These figures show that the NINT and FOM techniques underenhanced the original image and caused the processed images to look darker than the original image. Both methods enhanced the foreground while producing a darker background. As a result, the enhanced images exhibit improved contrast than the original image, which is also reflected in their respective entropy values. However, both techniques fail to maintain the mean brightness of the original image. Thus, the enhanced images are unpleasant and unnatural to look at. Both methods produced darker images, thus information and details in the images were reduced.

Figures 8, 9, 10, 11, 12 show that the improvement attained with the proposed technique results in a more pleasing enhanced image compared with those produced using the other techniques. The FRB and FL methods overenhanced certain regions (foreground area) of the image because the whole image was enhanced, but the underexposed and overexposed regions of the image were not considered. In addition, the FRB method caused intensity saturation at certain regions of the image, resulting in additional noises caused

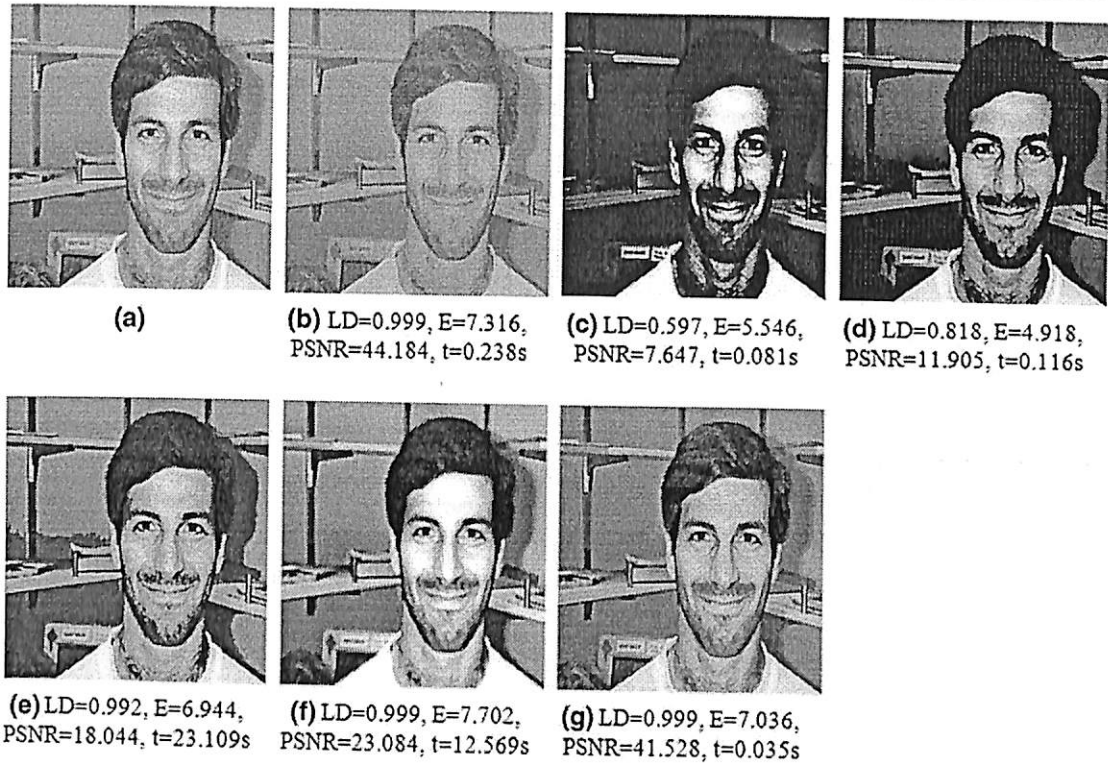


Fig. 9 Comparison of enhancement results a original 'Man 2' image, b proposed method, c NINT [1], d FOM [25], e FRB [13], f FL [30], g FHE [12]



Fig. 10 Comparison of enhancement results a original 'Woman 1' image, b proposed method, c NINT [1], d FOM [25], e FRB [13], f FL [30], g FHE [12]



Fig. 11 Comparison of enhancement results a original 'Woman 2' image, b proposed method, c NINT [1], d FOM [25], e FRB [13], f FL [30], g FHE [12]



Fig. 12 Comparison of enhancement results a original 'Woman 3' image, b proposed method, c NINT [1], d FOM [25], e FRB [13], f FL [30], g FHE [12]

considers poor contrast and nonuniform illumination problems that often occur in a recorded image. A new parameter, called the contrast factor, is proposed based on differences in the gray-level values of pixels in the local neighborhood. Improved image quality is obtained, and the proposed method is able to preserve the details and the mean luminance of the image. The proposed method defeats other methods in terms of LD, detail preservation, and noise suppression. The proposed algorithm has a comparably fast processing time of approximately 181 ms. Thus, the proposed algorithm is a suitable approach for real-time applications.

References

1. Hanmandlu, M., Jha, D., Sharma, R.: Color image enhancement by fuzzy intensification. *Pattern Recogn. Lett.* 24(1-3), 81-87 (2003)
2. Behrman, R., Zamenhof, R., Blazo, K.: Evaluation of a commercial mammography image-enhancement system. *J. Digit. Imaging* 2(3), 163-169 (1989). doi:10.1007/bf03168036
3. Chitra, T., Ray, A.K.: Fuzzy Image Processing and Applications with MATLAB. CRC Press/Taylor & Francis, Boca Raton (2010)
4. Zadeh, L.A.: Outline of a new approach to the analysis of complex systems and decision processes. *Syst. Man Cybern.* In: *IEEE Trans. SMC* 3(1), 28-44 (1973)
5. Gonzalez, R.C., Woods, R.E.: Digital Image Processing. Pearson/Prentice Hall, Englewood Cliffs, NJ (2008)
6. Chen, Q., Xu, X., Sun, Q., Xia, D.: A solution to the deficiencies of image enhancement. *Signal Process.* 90(1), 44-56 (2010). doi:10.1016/j.sigpro.2009.05.015
7. Zhiyu, C., Abidi, B.R., Page, D.L., Abidi, M.A.: Gray-level group-enhancement—part II: the variations. *Image Process. In: IEEE Trans.* 15(8), 2303-2314 (2006)
8. Zhiyu, C., Abidi, B.R., Page, D.L., Abidi, M.A.: Gray-level grouping (GLG): an automatic method for optimized image contrast enhancement—part I: the basic method. *Image Process. In: IEEE Trans.* 15(8), 2290-2302 (2006)
9. Chen, H.O., Kong, N.S.P., Ibrahim, H.: Bi-histogram equalization with a plateau limit for digital image enhancement. *Consum. Electron. In: IEEE Trans.* 55(4), 2072-2080 (2009)
10. Chen, H.O., Malisa, N.A.: Quadrants dynamic histogram equalization for contrast enhancement. *Consum. Electron. In: IEEE Trans.* 56(4), 2552-2559 (2010)
11. Avanaqi, A.: Exact global histogram specification optimized for structural similarity. *Opt. Rev.* 16(6), 613-621 (2009). doi:10.1007/s10043-009-0119-z
12. Ibrahim, H., Kong, N.S.P.: Brightness preserving dynamic histogram equalization for image contrast enhancement. *Consum. Electron. In: IEEE Trans.* 53(4), 1752-1758 (2007)
13. Sheel, D., Garud, H., Suveer, A., Mahadevappa, M., Chatterjee, J.: Brightness preserving dynamic fuzzy histogram equalization. *Consum. Electron. In: IEEE Trans.* 56(4), 2475-2480 (2010)
14. Bhanani, K.R., Baitou, A.: An application of fuzzy relations to image enhancement. *Pattern Recogn. Lett.* 16(9), 901-909 (1995)
15. Choi, Y., Krishnapuram, R.: A fuzzy-rule-based image enhancement method for medical applications. In: *Computer-Based Medical Systems, 1995, Proceedings of the Eighth IEEE Symposium on, 9-10 Jun 1995*, pp. 75-80 (1995)
16. Young, S.C., Krishnapuram, R.: A robust approach to image enhancement based on fuzzy logic. *Image Process. In: IEEE Trans.* 6(6), 808-825 (1997)

Table 1 Comparison of average quantitative analysis for 100 images

Techniques	$t(s)$	PSNR	E	LD
Proposed method	0.181	39.543	7.560	0.999
NINT [1]	0.149	14.170	6.027	0.918
FQM [25]	0.146	17.13	4.778	0.868
FRB [13]	22.169	21.843	7.134	0.99
FL [30]	16.055	19.065	7.472	0.997
FHE [12]	0.035	37.971	7.330	0.991

the low PSNR values of the images. This condition is also reflected in

The proposed technique is faster to execute and produces a more pleasing image compared with the FRB and FL methods. This result happened because brightness was improved accordingly with their respective regions (underexposed and overexposed regions). The intensity in the underexposed regions was increased using a modified sigmoid function in the proposed approach, as presented in Eq. (7). By contrast, the intensity in the overexposed regions was decreased by the respective sigmoid function derived in Eq. (7). Therefore, the dynamic range of the enhanced image is obtained, and the mean luminance or brightness is preserved. The sigmoid functions have avoided the intensity saturation problem, as discussed in the previous section.

The resulting image from FHE approach has an overenhanced foreground, thus causing intensity saturation in certain images as shown in Figs. 10g and 11g. Although the FHE approach is computationally fast, our proposed method produces an enhanced image with better quality, as shown by the quantitative values of LD, E, and PSNR.

The results of the quantitative analysis are given in Table 1, wherein comparison is conducted based on the average values of processing time, PSNR, LD, and E obtained from simulating 100 standard images. The best results obtained for each analysis are presented in bold text. Table 1 indicates that the proposed method gives the best performance, in terms of the highest PSNR, and by exhibiting an LD close to '1'. Therefore, the proposed method is better in preserving brightness. The entropy of the enhanced image resulting from the proposed method also defeated those of other techniques, thus indicating better contrast and detail preservation. However, in terms of processing time, FHE is the fastest to be executed because the image is treated as a mixed region in which overexposed and underexposed regions are not considered.

6 Conclusions

A new fuzzy enhancement technique is proposed for degraded gray-scale images. The new enhancement approach

17. Friedman, M., Schneider, M., Kandel, A.: The use of weighted fuzzy expected value (WFEV) in fuzzy expert systems. *Fuzzy Sets Syst.* **31**(1), 37–45 (1989)
18. Hanmandlu, M., Tandon, S.N., Mir, A.H.: A new fuzzy logic based image enhancement. *Biomed. Sci. Instrum.* **33**, 590–595 (1997)
19. Tizhoosh, H.R., Krell, G., Michaelis, B.: λ -enhancement: contrast adaptation based on optimization of image fuzziness. In: *Fuzzy Systems Proceedings, 1998. IEEE World Congress on Computational Intelligence., The 1998 IEEE International Conference on*, 4–9 May 1998, vol. 1542, pp. 1548–1553 (1998)
20. Vlachos, I., Sergiadis, G., Melin, P., Castillo, O., Aguilar, L., Kacprzyk, J., Pedrycz, W.: The role of entropy in intuitionistic fuzzy contrast enhancement foundations of fuzzy logic and soft computing. In: *Lecture Notes in Computer Science*, vol. 4529, pp. 104–113. Springer, Berlin (2007)
21. Vlachos, I.K., Sergiadis, G.D.: Parametric indices of fuzziness for automated image enhancement. *Fuzzy Sets Syst.* **157**(8), 1126–1138 (2006)
22. Vlachos, I.K., Sergiadis, G.D.: Intuitionistic fuzzy information—applications to pattern recognition. *Pattern Recogn. Lett.* **28**(2), 197–206 (2007)
23. Cheng, H.D., Chen, J.R.: Automatically determine the membership function based on the maximum entropy principle. *Inf. Sci.* **96**(3–4), 163–182 (1997). doi:10.1016/s0020-0255(96)00141-7
24. Pal, S.K.: A note on the quantitative measure of image enhancement through fuzziness. *Pattern Anal. Mach. Intell.* In: *IEEE Trans. PAMI* **4**(2), 204–208 (1982)
25. Nieradka, G., Butkiewicz, B., Melin, P., Castillo, O., Aguilar, L., Kacprzyk, J., Pedrycz, W.: A method for automatic membership function estimation based on fuzzy measures foundations of fuzzy logic and soft computing. In: *Lecture Notes in Computer Science*, vol. 4529, pp. 451–460. Springer, Berlin (2007)
26. Hanmandlu, M., Verma, O.P., Kumar, N.K., Kulkarni, M.: A novel optimal fuzzy system for color image enhancement using bacterial foraging. *Instrum. Meas.* In: *IEEE Trans.* **58**(8), 2867–2879 (2009)
27. Hanmandlu, M., Jha, D.: An optimal fuzzy system for color image enhancement. *Image Process.* In: *IEEE Trans.* **15**(10), 2956–2966 (2006)
28. Verma, O.P., Kumar, P., Hanmandlu, M., Chhabra, S.: High dynamic range optimal fuzzy color image enhancement using artificial ant colony system. *Appl. Soft Comput.* **12**(1), 394–404 (2012). doi:10.1016/j.asoc.2011.08.033
29. Wang, L., Wei, L.-Y., Zhou, K., Guo, B., Shum H.Y.: High dynamic range image hallucination'. In: *EGSR 2007. European Association for Computer Graphics* (2007)
30. Li, G., Tong, Y., Xiao, X.: Adaptive fuzzy enhancement algorithm of surface image based on local discrimination via grey entropy. *Procedia Eng.* **15**, 1590–1594 (2011)
31. Cheng, H.D., Xu, H.: A novel fuzzy logic approach to mammogram contrast enhancement. *Inf. Sci.* **148**(1–4), 167–184 (2002)
32. Vorobel, R., Berehulyak, O., Rutkowski, L., Tadeusiewicz, R., Zadeh, L., Zurada, J.: Gray image contrast enhancement by optimal fuzzy transformation artificial intelligence and soft computing—ICAISC 2006. In: *Lecture Notes in Computer Science*, vol. 4029, pp. 860–869. Springer, Berlin (2006)
33. Farbiz, F., Menhaj, M.B., Motamedi, S.A., Hagan, M.T.: A new fuzzy logic filter for image enhancement. *Syst. Man Cybern. Part B Cybern.* In: *IEEE Trans.* **30**(1), 110–119 (2000)
34. Rongjiang, P., Xiangxu, M.: A method of local enhancement based on fuzzy set theory. In: *Intelligent Control and Automation, 2000. Proceedings of the 3rd World Congress on*, 2000, vol. 1753, pp. 1751–1753 (2000)
35. Computational Vision Group: Computational Vision Archive (Faces). <http://www.vision.caltech.edu/html-files/archive.html> (1999)
36. Piegat, A.: *Fuzzy Modeling and Control*. Physica-Verlag, Wurzburg (Wien) (2001)
37. Zhou, W., Bovik, A.C.: A universal image quality index. *Signal Process. Lett.* In: *IEEE* **9**(3), 81–84 (2002)

Automated Feature-Based Sperm Motility Analyzer for Sperm Motility Detection: Effect of Enhancement on Detection Analysis

Khairunnisa Hasikin^{1,2}, Nor Ashidi Mat Isa¹, Mahaneem Mohamed³, Shahrel Azmin Suandi¹

¹Imaging and Intelligent Research Team (ISRT), School of Electrical and Electronic Engineering, Engineering Campus, Universiti Sains Malaysia, 14300, Nibong Tebal, Penang, Malaysia.

²Department of Biomedical Engineering, Faculty of Engineering, University of Malaya, 50603, Lembah Pantai, Kuala Lumpur, Malaysia

³Department of Physiology, School of Medical Sciences, Health Campus, Universiti Sains Malaysia, 16150, Kubang Kerian, Kelantan, Malaysia.

Email: ashidi@eng.usm.my

Abstract. A new Automated Feature-Based Sperm Motility Analyzer System is presented in this paper. This paper discusses on three major contributions where i) a new enhancement technique is proposed to improve overall clarity of the image sequence frames and significant sperm features are enhanced, ii) efficient debris removal process technique is developed to remove the unwanted debris that accumulates most regions in the image sequence frames, and iii) problems associated with disconnected objects are solved by implementing a new sperm feature detector. Differs from other sperm analyzer instruments that only consider sperms' head or tail, complete sperm features (*i.e.* head, mid-piece and tail) and the occluded sperms are successfully identified. The identification of the complete sperm features makes the individual sperm's detection possible and overestimation case of motile sperms proportion is reduced. The performance of the proposed system is determined by considering the true positive, true negative, false positive and false negative cases on the sperm motility

video. Although problems due to the unwanted debris and disconnected objects are successfully resolved, suitable enhancement approach needs to be selected. This is because, if improper enhancement approach is chosen, significant sperm features are inadequately enhanced which causes the detection rate to be deteriorated. Thus, the detection analysis is conducted by computing a comparative analysis of image sequence enhancement effect on sperm motility detection. The analysis is performed and overall detection rate is analyzed to attest the robustness of the proposed system which incorporated enhancement process. The proposed system achieves promising results in terms of motility detection with the accuracy of 82.67%, sensitivity of 92.34% and specificity of 90.63%.

Keywords: sperm motility; image processing; enhancement; rat sperm

1. Introduction

Semen analysis is an initial and the most essential step of the infertility evaluation which includes physical examination, hormonal evaluation, sperm function testing and genetic analysis. This analysis is considered as a cornerstone of the laboratory evaluation for the infertile male and helps to define the severity of the male factor infertility (Omu, 2013). The diagnosis of infertility is determined when the results of the semen analysis are repeatedly abnormal according to World Health Organization criteria (WHO, 1999). Among many evaluations performed during semen analysis, sperm function testing is crucial and play an important role since it can provide as an indication of a man's health status (Jensen et al., 2009; Omu, 2013).

The sperm count, sperm morphology and sperm motility analyses have been used as biomarkers of male-factor infertility since the middle of the 20th century (Omu, 2013). The former analysis classifies the abnormality of semen samples based on how many million sperm cells in an entire ejaculation are motile. The sperm morphology analysis focuses on the

shape of the sperm (*i.e.* head and tail of the sperm) while sperm motility deals with the movement and kinematic analyses of the sperms. However several studies have demonstrated that the sperm motility is highly contributed to the fertilization rate either in vitro or in vivo than relying only on the sperm count or morphology of the sperm (Guzick et al., 2001; He et al., 2004; Schoevaert-Brossaul, 1984). Such claim is supported by a study of 1,085 sperm samples from infertile men that concluded 815 samples had defects in motility, while 19% had asthenozoosperma (*i.e.* reduced in sperm motility rate) without any defects in sperm number or morphology (Curi et al., 2003). An immotile or abnormally-motile sperm will not fertilize the ovum and thus the assessment of the sperm motility is required (Guzick et al., 2001; He et al., 2004; Schoevaert-Brossaul, 1984).

The sperm motility analysis is one of the common procedures which acts as a critical indicator of the semen quality and fertility potential. The advancement in computerized digital image analysis has made a real-time sperm motility assessment a practical possibility. Due to the growth of sophisticated image processing techniques, computer-aided sperm analysis (CASA) instruments have been developed. These instruments are widely implemented tools that have been used for identification of sperm characteristic (*i.e.* sperm head/mid-piece/ tail) and trajectory analysis of motile sperms.

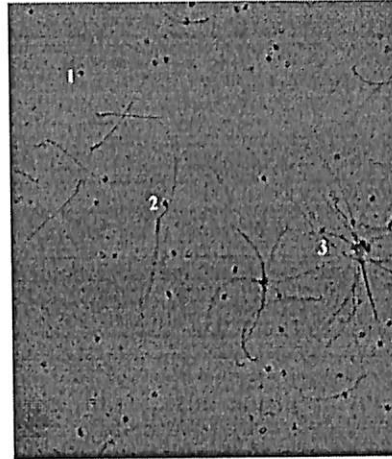
The method used for sperm motility analysis varies between CASA instruments. As an example, CellSoft (Mortimer and Mortimer, 1988; Mortimer, 1997) identifies the sperm head as continuous series of pixels and calculated center of this group of pixels. Meanwhile CellTrak system uses edge detection where only the pixels define the perimeter of an object are digitized and center of sperm head is identified from these coordinates (Mortimer, 1997). The Hamilton Throne CASA instruments identify the point of the sperm to follow by locating the brightest region of each bright image (Yeung et al., 1992). In contrary with the sperm

head detection, Strömberg-Mika Cell Motility Analyzer checks each image identified as a sperm head for the presence of a tail (Neuwinger et al., 1990).

Although CASA instruments are commonly implemented in modern research laboratories, most andrology clinics, especially those in less developed countries are still relying with conventional methods. Measurement of sperm head motion as commonly conducted in most clinics is sensitive to the technique of experimentation (Bayly et al., 2010; Ching-Hei Yeung et al., 1997). Spatial resolution of magnification system (*i.e.* microscope and camera settings) and temporal resolution of the systems (*i.e.* frame rate in relation to the speed of the sperm head) are among crucial considerations in the motility measurement. In addition, human visual system has far larger dynamic ranges than most commercial cameras and video cameras. These devices have limited dynamic ranges; thus recorded images obtained from these devices are usually non-homogeneous and low in contrast. Improper lighting condition and external disturbances which worsen the aforementioned problems are inevitable during image acquisition. When the recorded image frames suffer from non-uniform illumination, certain regions in the image appear brighter than the others. This situation caused the significant features of sperms are not highlighted and as a consequence, misidentification of the sperm head and debris could be occurred. In addition, extraction of the valuable information from the sperm motility video is also hindered by electronic noises acquired during data collection. These problems have made the segmentation between object of interest (*i.e.* motile sperm) and debris become more challenging.

The problems become crucial since the distinction between debris (Figure 1) and sperms plays a major role in determining the sperm's concentration and motility. When the images of clumped or agglutinated sperms (Figure 1) are digitized, individual sperm heads cannot be identified. Since most CASA instruments analyze the sperm trajectory by locating the sperm heads, non-identification of the head caused the whole group is identified as a

piece of debris and thus rejected from the sperm motility analysis. When this occurred, the concentration of sperms is underestimated and the proportion of motile sperm will be overestimated.



¹Debris; ²Debris that is adjacent to the sperm; ³Clumped sperms

Figure 1 Example of extracted image sequence from sperm motility video where debris and clumped sperms are illustrated in the figure

In this paper, a new system namely Automated Feature-Based Sperm Motility Analyzer System is proposed to detect and to perform trajectory analysis of motile sperms. This system is developed based on the detection of motile rat sperm. The samples are obtained from Sprague Dawley rat. The Sprague Dawley is the most widely used outbred albino rat in biomedical and reproduction researches (Suckow et al., 2005). It became popular choice among researchers due to its small size, easily maintained and managed as well as its capability to adapt well to new surroundings. Besides that, this species are used as models in this research because their genetic, biological and behaviour characteristics closely resemble those of humans and many symptoms of human conditions can be replicated in rats (Suckow et al., 2005).

The new system incorporates the enhancement process to improve the visibility of the motile sperm video, connected component and feature-based sperm motility video analysis.

The development of the system is elaborated in details in the succeeding sections. This paper is organized as follows: Section 2 explains thoroughly on the proposed system which includes the enhancement, debris removal and sperm detection processes. Section 3 presents the results of the proposed system as well as their discussion. Section 4 summarizes and concludes the paper.

2. The Proposed System

The block diagram of the proposed system is presented in Figure 2. This figure depicts that the proposed system consists of three main stages. The recorded sperm motility video is converted into sequence of image frames in the first stage where the enhancement process is performed to enhance significant features of the motile sperms. Then, the enhanced features are further processed to extract only the object of interest (*i.e.* motile sperms) in the second stage. Lastly, in stage three, the centers of the motile sperms are identified.

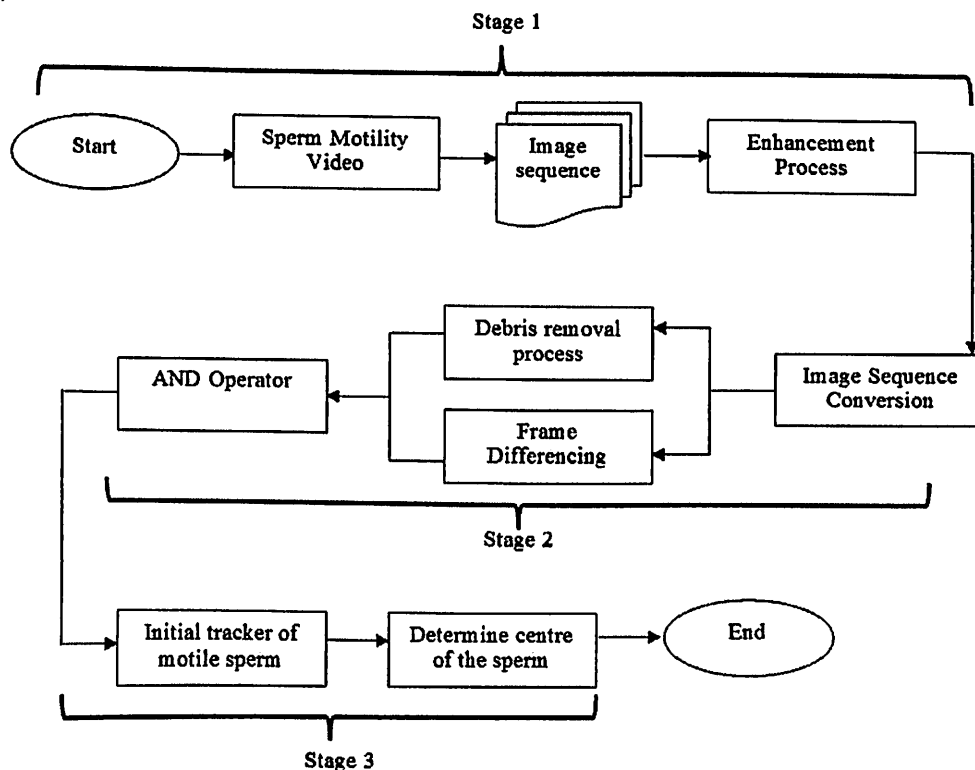


Figure 2 The proposed system

2.1 Stage 1: Pre-processing

In this stage, the sperm motility video that is recorded using microscope camera is converted into sequence of image frame. This image sequence is processed to improve the clarity of the image and to enhance significant features of the sperms. Due to the advent and recent development in imaging technology, the recorded microscopic videos are analysed and interpreted to attain valuable information that could help in sperm motility detection. Video imaging is currently a popular image acquisition choice for most modern light microscope that has been widely used in medical applications. The video imaging enables faster image acquisition which makes real time imaging possible.

However, video imaging can induce electronic noises in the microscopic images. The microscopic images are susceptible to all kinds of factors such as improper alignment of the microscope and inadequate cleanliness of the optical elements which resulted in decrement in the image quality and visual clarity. The need to capture and analyse poor quality, low contrast microscopy images are a frequent problem confronting scientist/physician. These limitations are particularly severe when the microscope is used as a tool for actual experimentation as in this sperm motility analysis. The degraded low contrast image sequences in their raw form may still contain useful information that is unseen due to perceptual limitation of the human eye. The brightness needs to be below the saturation limit to avoid streaks, smears or over-saturated images during image acquisition stage. Meanwhile, the image intensity also needs to be above a threshold level or the camera aperture needs to be properly adjusted to avoid snow or granulation of the image. Additional electrical fixtures such as light sources, lasers and motors can generate spikes or noise pulses that can appear as spikes across the image. These factors may contribute to the degradation of visual quality of the images and eventually lead to misdiagnose and false detection (Shiraishi et al., 2011).

Therefore, in the Automated Feature-Based Sperm Motility Analyzer System, a new enhancement process is proposed to improve the visibility of the image sequence. This hidden significant information is revealed by the proposed Adaptive Fuzzy Intensity Measure (AFIM) enhancement technique (Hasikin and Mat Isa, 2013). The dominance of sperm features is increased and additional information that is not apparent in the original image as well as regions that are hardly detected by the eyes are retrieved.

Since the microscopic image sequences are low contrast, most of the pixels possess uneven brightness and their intensity values are vague. The problem in the intensity vagueness occurs when the system tries to determine whether the pixels are bright or dark. Therefore, the fuzzy approach is employed in the AFIM. The fuzzy enhancement process is depicted in Figure 3.

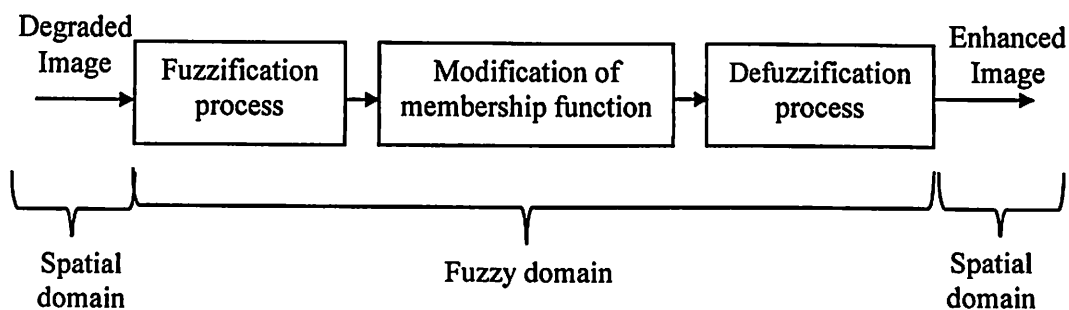


Figure 3 Fuzzy image enhancement process

The enhancement process begins with the fuzzification process. In this stage, the image sequence is divided into bright and dark regions by specifying fuzzy intensity measure which is thresholded using equation (1). The dark region is clustered in the range of $[0, T-1]$, whereas the bright region is clustered in the range of $[T, L-1]$.

$$T_f = L \left[\frac{g_d}{g_a} \right] \quad (1)$$

where

$$g_a = \frac{\sum_{m=0}^{L-1} m \times p(m)}{\sum_{m=0}^{L-1} p(m)} \quad (2)$$

$$g_d = \left[\frac{\sum_{m=0}^{L-1} [(m - g_a)^2 p(m)]}{\sum_{m=0}^{L-1} p(m)} \right]^{1/2} \quad (3)$$

$$\text{fuzzy intensity measure} = \frac{g_d}{g_a} \quad (4)$$

$p(m)$ represents the number of pixels in the histogram of the entire image. g_d and g_a are deviation and mean intensity distributions, respectively. m is the intensity value at pixel (i,j) .

Equations (2) to (4) are derived by considering the mean and deviation of histogram intensity distributions. These equations are calculated to determine the non-homogeneous intensity distributions of the image. This new parameter of fuzzy intensity measure (*i.e.* equation (4)) is proposed which plays a major role in clustering the image into bright or dark regions.

After the image sequence frame is divided into two regions (dark and bright regions) based on the value of T , fuzzification is performed in each region separately. The modified Gaussian membership function is utilized for the fuzzification of the dark region as follows:

$$\mu_d(m) = \exp \left[-\frac{(m_{\max} - (m_{\text{avg}} - m))^2}{\zeta_d} \right] \quad \text{for } m < T \quad (5)$$

where $\mu_d(m)$ is the membership function in the dark region and m is the intensity value in the dark region in the range of $[0 T-1]$. m_{avg} and m_{max} are the average intensity and maximum intensity of the image, respectively. ζ_d the fuzzifier function of the dark region, is provided by:

$$\zeta_d = \alpha \frac{\sum_{m=0}^{L-1} [(m_d - m_{d_{avg}}) - \sigma_m]^4 p(m_d)}{\sum_{m=0}^{L-1} [(m_d - m_{d_{avg}}) - \sigma_m]^2 p(m_d)} \quad (6)$$

where σ_m is the standard deviation of intensity of the entire image, $m_{d_{avg}}$ is the average intensity of the dark pixels and m_d and $p(m_d)$ are the intensities and histogram of the dark region, respectively.

The mirror function of the aforementioned Gaussian membership function is utilized to fuzzify the bright region of the image for $m \geq T_f$ as follows:

$$\mu_b(m) = \exp \left[-\frac{(m_{max} - (m_{avg} - (L - m)))^2}{\zeta_b} \right] \quad \text{for } m \geq T \quad (7)$$

where $\mu_b(m)$ is the membership function of bright region. ζ_b is the fuzzifier function in the bright region.

$$\zeta_b = \alpha \frac{\sum_{m=0}^{L-1} [(m_b - m_{b_{avg}}) - \sigma_m]^4 p(m_b)}{\sum_{m=0}^{L-1} [(m_b - m_{b_{avg}}) - \sigma_m]^2 p(m_b)} \quad (8)$$

where $m_{b_{avg}}$ is the average intensity of the bright pixels, m_b is the intensity of the bright region, and $p(m_b)$ is the histogram of the bright pixels.

The fuzzifier functions of ζ_d and ζ_b calculate the intensity deviation in the dark and bright regions, respectively. The fuzzified image in fuzzy domain is obtained by combining membership functions for both regions (*i.e.* dark and bright regions). α is the fuzzification

factor that depends on the intensity values of the input image which is explained in details in Appendix A.

Once fuzzification is completed, the original input pixels that exhibit non-uniform illumination and low contrast are transformed into Gaussian distributed pixels. The local contrast of the image is based on intensity difference in a small region and it is computed to preserve the details of the image. Local contrasts are defined for the dark and bright regions as:

$$C_{L_d}(i, j) = \sum_{(i,j) \in W_{i,j}} [\max(\mu_d(i, j)) - \min(\mu_d(i, j))] \quad (9)$$

$$C_{L_b}(i, j) = \sum_{(i,j) \in W_{i,j}} [\max(\mu_b(i, j)) - \min(\mu_b(i, j))] \quad (10)$$

where $\mu_d(i, j)$ and $\mu_b(i, j)$ represent the 3x3 local fuzzified image (*i.e.*, output image obtained after fuzzification process) of μ_d and μ_b , respectively which are centered at position (i, j) . $\max(\mu_d(i, j))$ and $\max(\mu_b(i, j))$ represent the maximum gray level values of the local fuzzified image for dark and bright regions, respectively. $\min(\mu_d(i, j))$ and $\min(\mu_b(i, j))$ denote the minimum gray level values of the local fuzzified image for dark and bright regions, respectively.

Modification of the fuzzified image is performed once the aforementioned steps are executed. Modification is performed to enhance the fuzzified image based on the dark and bright regions, which include the local contrast of the image as shown in equations (11) and (12) respectively.

$$\mu_d'(m) = \frac{1}{1 + e^{\{-C_{L_d}[\mu_d(m) - m_{d\text{avg}}]\}}} \quad \text{for } m < T \quad (11)$$

$$\mu'_b(m) = \frac{1}{1 + e^{\{-C_{L_b} [\mu_b(m) - m_{b,avg}]\}}} \quad \text{For } m \geq T \quad (12)$$

where μ'_d and μ'_b are the modified membership functions in the dark and bright regions, respectively. C_{L_d} and C_{L_b} are the local contrast of dark and bright regions, respectively which are computed to preserve the details in the image.

The above functions modify the original membership functions of $\mu_d(m)$ and $\mu_b(m)$. The modified functions are then defuzzified with the respective inverse membership functions as shown in equation (13). Both regions are combined to obtain the enhanced image. The pixels in the dark region are scaled back to the range $[0 T-1]$, whereas the bright region is translated and scaled back to the region $[T L-1]$.

$$M = \begin{cases} \mu'^{-1}_d(m) & \forall m \leq T \\ \mu'^{-1}_b(m) & \forall m > T \end{cases} \quad (13)$$

where M is the enhanced image obtained from the defuzzification process.

2.2 Stage 2: Sperm Feature Detection Process

2.2.1 Conversion of Image Sequence Frame

The enhanced sperm features obtained from the AFIM process is segmented in this stage. An adaptive thresholding is employed in this stage to distinguish the sperm from its background. The image sequences are converted to binary image which white pixels represent the sperms while the black pixels denote as the background. The image sequence conversion is conducted using equation (14).

$$b(x, y) = \begin{cases} 255 & \text{if } M(x, y) > Th \text{ - (i.e. foreground)} \\ 0 & \text{if } M(x, y) \leq Th \text{ - (i.e. background)} \end{cases} \quad (14)$$

where $M(x,y)$ is the enhanced intensity at pixel (x,y) , while $b(x,y)$ is the binary image produced by the thresholding process. Th is the dynamic threshold that is set locally based on the average intensity value (i.e. equation (15)) of the local neighbourhood pixels with window size $(2t+1) \times (2t+1)$ where t is an integer number (Figure 4).

$$Th = \frac{\sum (M_1, M_2, M_3, \dots, M_n)}{n} \quad (15)$$

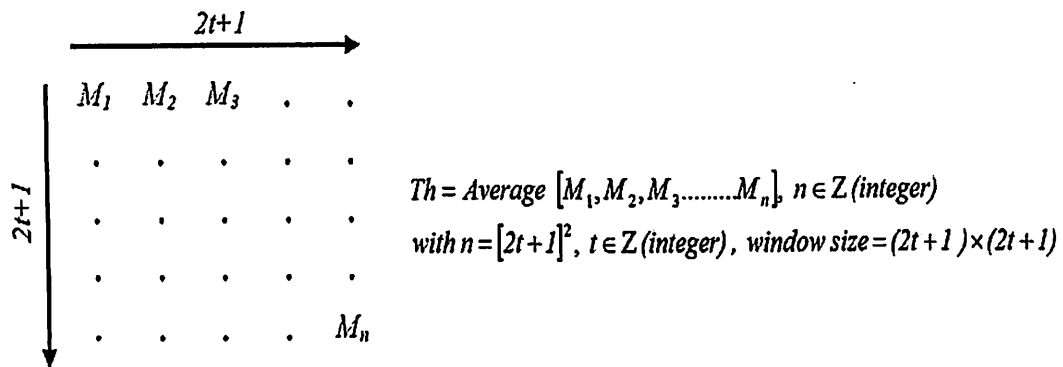


Figure 4 Selection of Th that is computed locally over the entire of image frame¹

2.2.2 Debris Removal Process

The unwanted debris that occupies most regions in the binary image are removed in this process. As depicted in Figure 1, shapes of debris and sperms are apparently different from each other. The debris assembles circular shape while sperms are more likely in cylindrical shape. Therefore, to discriminate the sperm and the debris, coordinates of the center of all objects are determined. The center of the object denoted as (x_c, y_c) is specified by computing central moment of the object:

$$x_c = \frac{n_{1,0}}{n_{0,0}} \quad (16)$$

¹ n is the total number of pixels in the local neighbourhood, n is the odd number which could varies depending on the image. In this study 17×17 window size is utilized to obtain the binary image.

$$y_c = \frac{n_{0,1}}{n_{0,0}} \quad (17)$$

where $n_{0,0}$ is the zero order moment which describes the area of the object and $n_{1,0}$ and $n_{0,1}$ are specified in equations (18) and (19) respectively.

$$n_{1,0} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx dy x b(x, y) \quad (18)$$

$$n_{0,1} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx dy y b(x, y) \quad (19)$$

where $b(x,y)$ is the binary image obtained from the image sequence conversion process as presented in previous subsection.

Then, circularity or roundness of the objects is calculated by measuring two parameters namely object's area and parameter as shown in equation (20). These parameters are derived using equations (21) and (22) respectively where radius and angle of the object are measured with respect to the pre-determined center coordinate in equations (16) and (17).

$$circularity = \frac{4\pi A}{P^2} \quad (20)$$

where

$$A = \int_0^{2\pi} \frac{1}{2} r^2 d\phi = \pi \sum_{-\infty}^{\infty} |r_z|^2 \quad (21)$$

$$P = \int_0^{2\pi} \left[(r(\phi)d\phi)^2 + (dr(\phi))^2 \right]^{1/2} = \int_0^{2\pi} d\phi \left[r^2 + \left(\frac{dr}{d\phi} \right)^2 \right]^{1/2} \quad (22)$$

where A is the area of the object, r is the radius of the object, ϕ is the angle of object with respect to its centre, and P is the perimeter of the object.

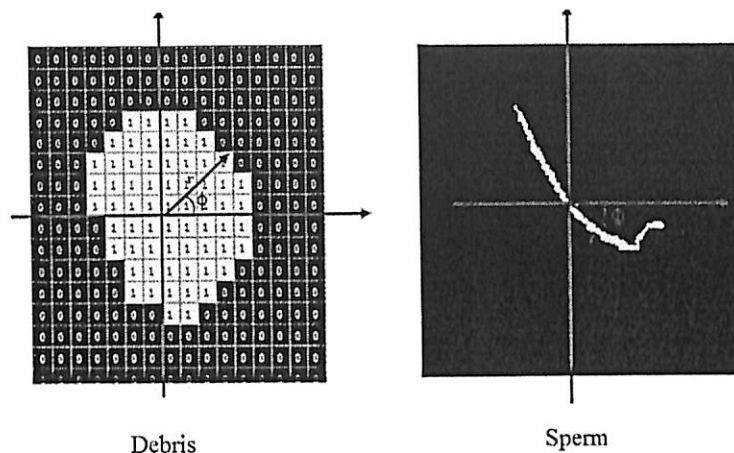


Figure 5 Center, radius and angle of the object in image sequence frame

Figure 5 shows the example of the angle and radius that are extracted from the debris and sperm with respect to their centres. The object is categorized as debris if the circularity is closed to '1' while the object is assumed to be sperm if the object has circularity that is not closed to '1'.

For all objects that are classified as sperm, pixels belong to these objects are left untouched. However, if the object is classified as debris, coordinate of the object's centre is recorded and coordinates of the mask border are specified. A mask size of 7x7 (Figure 6) is used to indicate a border that determines which region needs to be processed. The 7x7 mask size is chosen in this process to represent the area that needs to be processed since the shape of the debris is very small. In addition, selection of bigger size window might interrupt the adjacent pixels that belong to the sperm's object (as shown in Figure 7). Thus, smallest window size is chosen to remove the unwanted debris while adjacent sperm is kept unchanged for further processing.

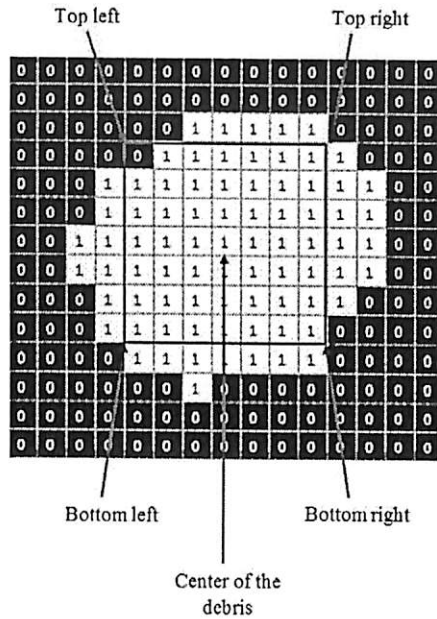


Figure 6 Mask size of 7x7 that is utilized to remove the debris. Pixels in the debris region will be eliminated

The coordinates of the border; top left, top right, bottom right and bottom left are identified as shown in Figure 6. White pixels that lie within this border are converted to black pixels while exiting black pixels in this region (if any) are kept unchanged. The residue pixels as shown in Figure 7 are removed after AND Operator which will be explained in the next subsection. The debris removal processed is summarised and presented in the flow chart presented in Figure 8.

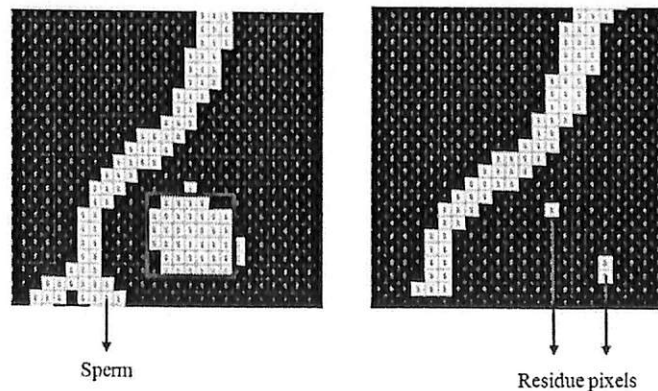


Figure 7 Example of the debris that are adjacent to the sperm

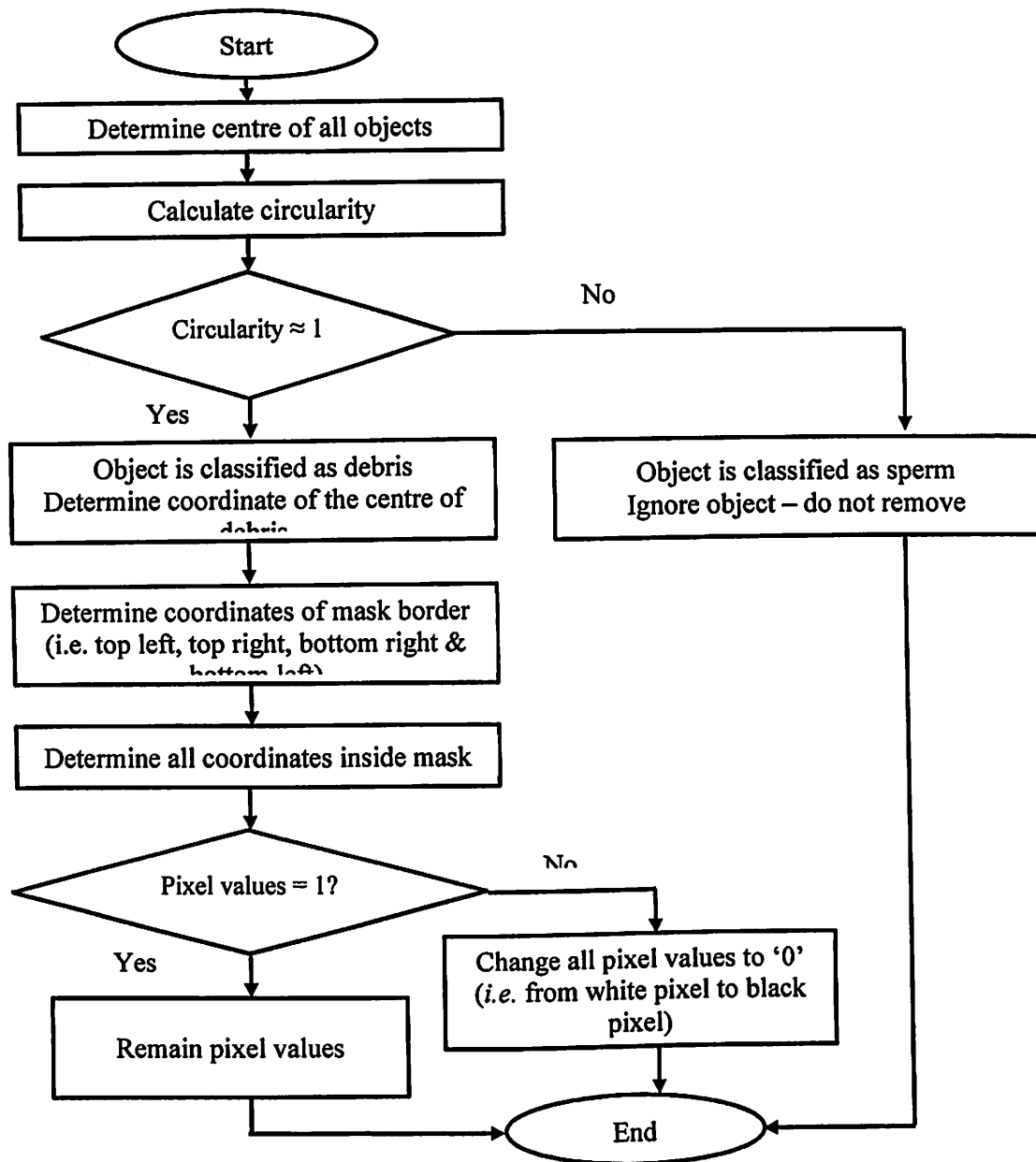


Figure 8 Flow chart of debris removal process

2.2.3 Frame Differencing

The motion features are extracted using frame differencing process that is computed as in equation (23). As given in this equation, if $D(t)$ is equal to '1', the corresponding region is defined as the moving region. Otherwise, the region is considered as background. T in this

case is chosen based on the median value of the difference image between current and next frames.

$$D(t) = \begin{cases} 1 & \text{if } \|F(t-1) - F(t)\| \geq T \\ 0 & \text{otherwise} \end{cases} \quad (23)$$

where $D(t)$ represents the existence of difference between background and the input image frame. Meanwhile $F(t-1)$ is the previous frame and $F(t)$ is the current frame.

Since the simple frame differencing method often fails because of misidentification of the moving object (Beyan and Temizel, 2012), it is necessary to incorporate sperm features with its motion features to provide initial tracker of the moving sperms. Results obtained from debris removal and frame differencing processes are merged using AND Operator. The residual pixels that are left on the merged image frame are removed using morphology filtering. The binary image attained from the merging process of AND Operator is filtered at a scale defined by the size of structuring element. Regions that fit the structuring element are passed by the filter meanwhile smaller structures are blocked and excluded from the output image. In this study, structure element of size 2x2 is utilized to remove the residue pixels on the binary image obtained from the AND Operator. Sperm with region size bigger than the structure element is left untouched.

By implementing the proposed system, the debris removal process in addition to the frame differencing process, the segmentation errors that could occur after frame differencing are eliminated. The motile sperms are successfully extracted and false detection could be avoided.

2.2.4 Initialization of Motile Sperm Tracker

After the motion features and sperm shape are successfully extracted, initial trackers of the sperm are determined. The center coordinates of the objects in the cleaned binary image

are extracted using aforementioned equations (16) and (17). The implementation of the overall sperm feature detection is simplified in Figure 9.

Although motion features and sperm shape are successfully extracted, the segmented object tends to be disconnected. This is because some motile sperms are moving too fast and thus caused the visibility of these motile sperms to be reduced (Figure 10 (a)). As a result, the segmented object in binary image (Figure 10(b)) is disconnected and the object is identified as two different objects which could cause false detection.

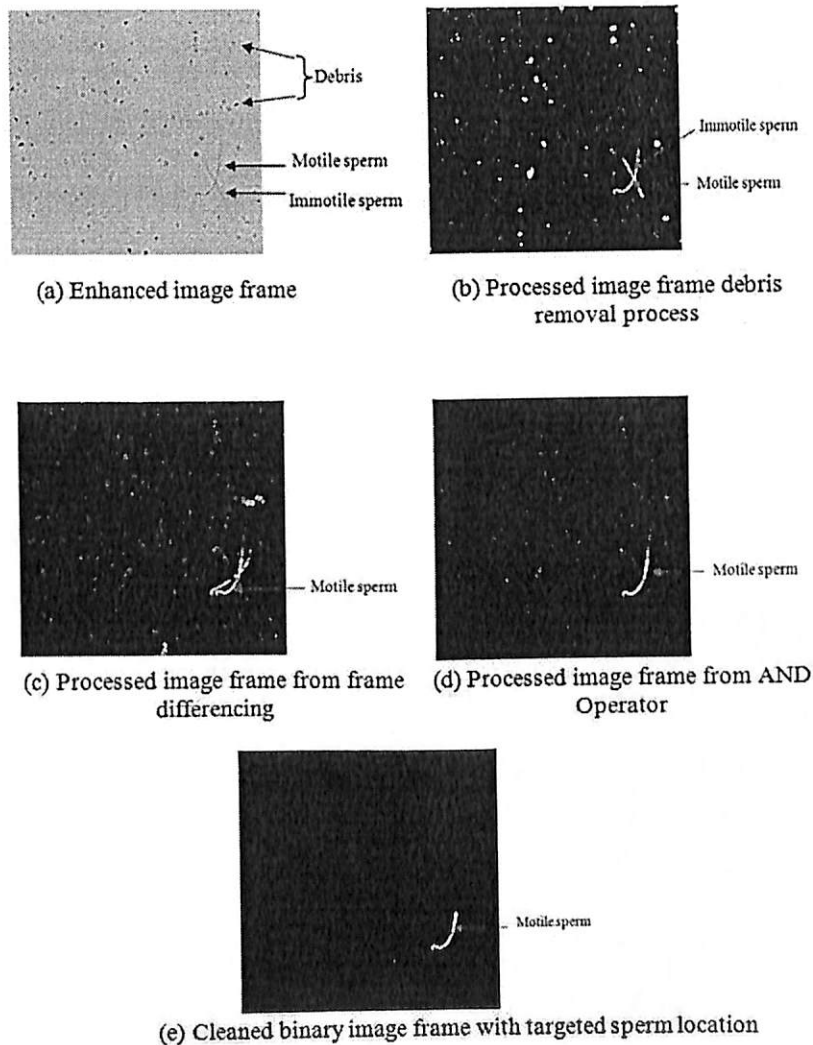


Figure 9 Implementation of the sperm feature detection process

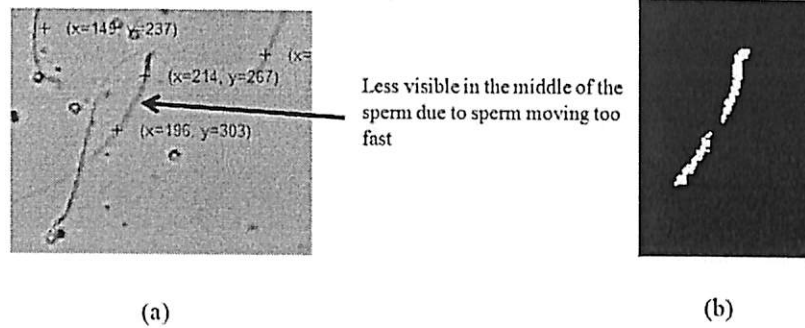


Figure 10 Example of disconnected object that occurs due to the fast moving sperm

Thus, the center coordinates determined in this stage will act as the initial tracker of the motile sperm. The disconnected objects need to be attached to their own pairs that possess similar orientation. Since most sperms are adjacent to one another and moving in uncontrolled environment, there might be more than one possibility of coordinate combinations between two disconnected objects.

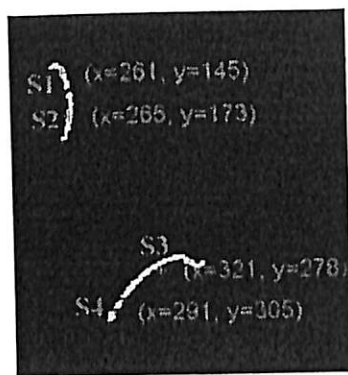
Figure 11 shows an example of possible combinations case between four disconnected objects. The initial tracker of motile sperms detected four center coordinates. These initial coordinates are utilized to determine number of possible coordinate combinations between disconnected objects using equation (24). As depicted in this figure since k is equivalent to 4, therefore there are six possible coordinate combinations for the disconnected objects as tabulated in Table 1.

$$\text{No. possible coordinate combinations} = \frac{k(k-1)}{2} \quad (24)$$

$k = \text{initial total number of detected sperms}$

The extracted coordinate combinations include combination between objects S1 and S2, S1 and S3, S1 and S4, S2 and S3, S2 and S4, as well as S3 and S4. Based on these possible coordinate combinations, the objects are categorized into two groups of *Candidate 1* and *Candidate 2*. For each possible coordinate combination (*i.e.* 2 disconnected objects), object

with smaller x -coordinate is considered as *Candidate 1* whereby the other disconnected object is denoted as *Candidate 2*. As tabulated in Table 1, for the first coordinate combination, object S1 has smaller x -coordinate than x -coordinate of object S2, therefore, object S1 and S2 are categorized as *Candidate 1* and *Candidate 2* respectively. Similar situation can be observed with sixth coordinate combination where x -coordinate of object S3 is smaller than x -coordinate of object S4. Therefore S3 is grouped in *Candidate 1* while S4 is categorized as *Candidate 2*.



$$\text{No. of possible coordinate combination} = \frac{k(k-1)}{2}$$

since $k = 4$

$$\therefore \text{No. of possible coordinate combinations} = 6$$

Figure 11 Example of possible combinations of disconnected object from initial tracker of motile sperm.

Table 1 Possible coordinate combination for Figure 11

Possible Coordinate Combination	Coordinates of Candidate 1	Coordinates of Candidate 2	Δx_i	Δy_i	$\Delta x_i + \Delta y_i$
1 [S1-S2]	(261,145)	(265,173)	$261-265 =4$	$145-173 =28$	32
2 [S1-S3]	(261,145)	(321,278)	$ 261-321 =60$	$ 145-278 =133$	193
3 [S1-S4]	(261,145)	(291,305)	$ 261-291 =30$	$ 145-305 =160$	190
4 [S2-S3]	(265,173)	(321,278)	$ 265-321 =56$	$ 173-278 =105$	161
5 [S2-S4]	(265,173)	(291,305)	$ 265-291 =26$	$ 173-305 =132$	158
6 [S3-S4]	(321,278)	(291,305)	$321-291 =30$	$278-305 =27$	57

*The significant coordinate combinations are made bold

Each possible coordinate combination is analysed and only correct combination will be extracted to attach the disconnected objects. The disconnected objects must be connected only to their own combination to avoid any deformation of the sperm shape which will then affect the detection of the motile sperm process. However, analysing each coordinate

combination requires high computational time and increases complexity of the detection process. Therefore, only significant coordinate combinations will be considered to be processed in order to connect the disconnected objects. Each disconnected object case will be discussed thoroughly in the next subsection.

2.3 Analysis of Disconnected Objects

As previously discussed in the previous section, the complexity of the system can be reduced if only significant disconnected objects are evaluated. Therefore, to solve this issue, absolute differences of x -coordinates (Δx_i), absolute differences of y -coordinates (Δy_i) and total absolute difference of coordinates between *Candidate 1* and *Candidate 2* ($\Delta x_i + \Delta y_i$) are calculated for every possible coordinate combination as can be seen in the example tabulated in Table 1.

A threshold of 100 is set to determine significant coordinate combinations to be processed. This threshold value is selected based on the observation made on the distance between objects on image sequence frames extracted from 100 sperm motility videos. The significant coordinate combinations are identified if the combinations have the value of $\Delta x_i + \Delta y_i$ is less than 100. These coordinate combinations consist of the disconnected objects that are closed to each other. They are more likely to be attached to form a sperm shape. Therefore, only coordinate combinations that produced $\Delta x_i + \Delta y_i \leq 100$, are considered to be evaluated.

As depicted in Figure 11 and Table 1, from six possible coordinate combinations, only two combinations namely the first and the sixth coordinate combinations are relevant to be considered as disconnected objects. The combinations are for objects S1 and S2, as well as for objects S3 and S4 with $\Delta x_i + \Delta y_i$ of 34 and 57 respectively. It can be also observed in this table that this case only deals with unique coordinates of *Candidate 1* and *Candidate 2* where

the sperms are not adjacent to each other. These coordinates are considered as unique since no repetitive coordinates appear in *Candidate 1* and *Candidate 2* as well as no repetitive coordinates can be seen between combination of *Candidate 1* and *Candidate 2*.

As an example in this table, *Candidate 1* consists of objects S1 (*i.e.* coordinate of (264,145)) and S3 (*i.e.* coordinate of (321,278)) while *Candidate 2* consists of objects S2 (*i.e.* coordinate of (265,173)) and S4 (*i.e.* coordinate of (291,305)). It can be observed from this example there are unique coordinates between intra-group *Candidate 1* and also unique coordinates between intra-group *Candidate 2*. In addition, this example shows that the coordinates are unique between inter-group *Candidate 1* and *Candidate 2*.

However, most of the image sequence frames consist of frames with sperms that are moving closely with other sperms. Thus, their disconnected objects are adjacent and non-unique coordinates of *Candidate 1* and *Candidate 2* can be observed. Let consider Table 2, where the first example of Table 2 shows that among three possible coordinate combinations, only the first and second combinations are considered in this analysis. This is because the third combination produces $\Delta x_i + \Delta y_i$ more than 100 and therefore is excluded in the analysis of disconnected objects. Object J (*i.e.* coordinate (194,207)) is considered as a non-unique coordinate between combination of *Candidate 1* and *Candidate 2*. Object J (*i.e.* coordinate (194,207)) belongs to *Candidate 2* in the first combination while in the second combination it is assigned as *Candidate 1*. As it could be candidate for more than one combination, it is said to be non-unique coordinates. This combination is made bold in the first example of Table 2.

There are two types of non-unique coordinate as follow where the first type is belong to the coordinate that could possibly belong to more than one combination but only assigned to one candidate (*i.e.* either *Candidate 1* or *Candidate 2*). Meanwhile second type of non-unique coordinate is referred to the coordinate that could possibly belong to more than one combination and assigned as two candidates (*i.e.* both *Candidate 1* and *Candidate 2*).

For clarification of the aforementioned situations of non-unique coordinate can be referred to the second example of Table 2 (*i.e.* second row). In this example, the third and fourth coordinate combinations are neglected since objects L and O as well as objects L and P are far from each other as both combinations attain $\Delta x_i + \Delta y_i$ more than 100. For the first type of non-unique coordinate, it can be observed that coordinates (311,141) (for object L) are repeated in first and second combinations.

This coordinate is considered as the first type of non-unique coordinate as it is assigned as *Candidate 1* for more than one combinations. The first type of non-unique coordinate can also be observed for object P (*i.e.* coordinate (346,241)) where it is assigned as *Candidate 2* in the seventh, ninth and tenth combinations. The second type of non-unique coordinates can also be observed in this example. Coordinate (326,188) (*i.e.* Object M) appears as *Candidate 2* at the first combination and as *Candidate 1* at the fifth, sixth and seventh combinations. In addition, object N (*i.e.* coordinate (367,170)) also shows the example of the second type of non-unique coordinate where it is assigned as *Candidate 2* for the second and fifth combinations as well as *Candidate 1* for eighth and ninth combinations. Similar situation can be observed in object O (*i.e.* coordinate (380,213)) where it became *Candidate 2* for sixth and eighth combinations while it appears as *Candidate 1* in the tenth combination.

Therefore, each non-unique coordinate needs to be analysed and correct coordinate combination needs to be chosen to solve the disconnected object issues. Four cases of disconnected objects are found in this study and successfully identified using the process presented in Figure 12. Each condition to determine all cases of disconnected objects is presented in Table 3.

Table 2 Unique and non-unique combinations of *Candidate 1* and *Candidate 2*

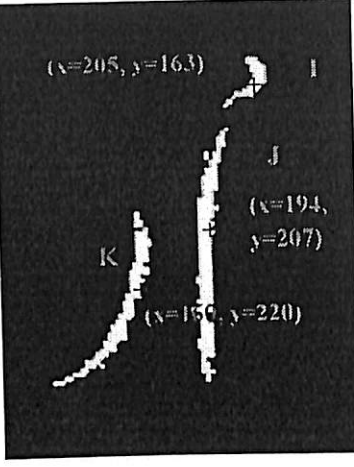
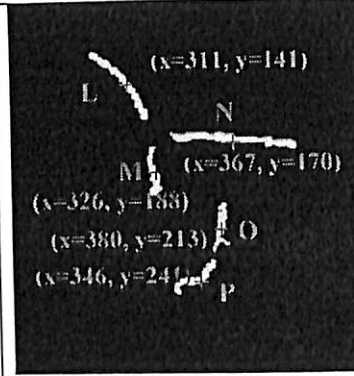
Example	Coordinate Combination																											
 <p>(x=205, y=163) I J (x=194, y=207) K (x=150, y=220)</p>	<p>Coordinate Combinations:</p> <ol style="list-style-type: none"> I-J; $\Delta x_i + \Delta y_i = 55$ J-K; $\Delta x_i + \Delta y_i = 37$ I-K; $\Delta x_i + \Delta y_i = 102$ – neglected in analysis since $\Delta x_i + \Delta y_i \geq 100$ <p>Significant Coordinate Combinations to be analyzed:</p> <table border="1" data-bbox="703 601 1316 737"> <thead> <tr> <th>Coordinate Combination</th> <th><i>Candidate 1</i></th> <th><i>Candidate 2</i></th> </tr> </thead> <tbody> <tr> <td>1</td> <td>(205,163)</td> <td>(194,207)</td> </tr> <tr> <td>2</td> <td>(194,207)</td> <td>(160,220)</td> </tr> </tbody> </table> <p>*non-unique coordinate between combination of <i>Candidate 1</i> and <i>Candidate 2</i> is made bold</p>	Coordinate Combination	<i>Candidate 1</i>	<i>Candidate 2</i>	1	(205,163)	(194,207)	2	(194,207)	(160,220)																		
Coordinate Combination	<i>Candidate 1</i>	<i>Candidate 2</i>																										
1	(205,163)	(194,207)																										
2	(194,207)	(160,220)																										
 <p>(x=311, y=141) L N M (x=326, y=188) (x=367, y=170) (x=380, y=213) O (x=346, y=241) P</p>	<p>Coordinate Combinations:</p> <ol style="list-style-type: none"> L-M; $\Delta x_i + \Delta y_i = 62$ L-N; $\Delta x_i + \Delta y_i = 85$ L-O; $\Delta x_i + \Delta y_i = 121$ - neglected in analysis since $\Delta x_i + \Delta y_i \geq 100$ L-P; $\Delta x_i + \Delta y_i = 135$ – neglected in analysis since $\Delta x_i + \Delta y_i \geq 100$ M-N; $\Delta x_i + \Delta y_i = 59$ M-O; $\Delta x_i + \Delta y_i = 59$ M-P; $\Delta x_i + \Delta y_i = 73$ N-O; $\Delta x_i + \Delta y_i = 50$ N-P; $\Delta x_i + \Delta y_i = 92$ O-P; $\Delta x_i + \Delta y_i = 42$ <p>Significant Coordinate Combinations to be analyzed:</p> <table border="1" data-bbox="703 1349 1316 1712"> <thead> <tr> <th>Coordinate Combination</th> <th><i>Candidate 1</i></th> <th><i>Candidate 2</i></th> </tr> </thead> <tbody> <tr> <td>1</td> <td>(311,141)</td> <td>(326,188)</td> </tr> <tr> <td>2</td> <td>(311,141)</td> <td>(367,170)</td> </tr> <tr> <td>5</td> <td>(326,188)</td> <td>(367,170)</td> </tr> <tr> <td>6</td> <td>(326,188)</td> <td>(380,213)</td> </tr> <tr> <td>7</td> <td>(326,188)</td> <td>(346,241)</td> </tr> <tr> <td>8</td> <td>(367,170)</td> <td>(380,213)</td> </tr> <tr> <td>9</td> <td>(367,170)</td> <td>(346,241)</td> </tr> <tr> <td>10</td> <td>(380,213)</td> <td>(346,241)</td> </tr> </tbody> </table> <p>**non-unique coordinates between combination of <i>Candidate 1</i> and <i>Candidate 2</i> are made bold</p>	Coordinate Combination	<i>Candidate 1</i>	<i>Candidate 2</i>	1	(311,141)	(326,188)	2	(311,141)	(367,170)	5	(326,188)	(367,170)	6	(326,188)	(380,213)	7	(326,188)	(346,241)	8	(367,170)	(380,213)	9	(367,170)	(346,241)	10	(380,213)	(346,241)
Coordinate Combination	<i>Candidate 1</i>	<i>Candidate 2</i>																										
1	(311,141)	(326,188)																										
2	(311,141)	(367,170)																										
5	(326,188)	(367,170)																										
6	(326,188)	(380,213)																										
7	(326,188)	(346,241)																										
8	(367,170)	(380,213)																										
9	(367,170)	(346,241)																										
10	(380,213)	(346,241)																										

Table 3 Conditions to be considered to identified disconnected object cases

Condition	Description
1	$\Delta x_i + \Delta y_i \leq 100$
2	i) Unique coordinates of <i>Candidate 1</i> AND ii) Unique coordinates of <i>Candidate 2</i>
3	*Not a type-1 non-unique coordinate Unique coordinates of combination of both <i>Candidates</i>
4	*Not a type-2 non-unique coordinate Total length of <i>Candidate 1</i> and <i>Candidate 2</i> ≥ 100

As previously discussed, the disconnected object cases are considered to be combined if $\Delta x_i + \Delta y_i$ is less than or equivalent to 100. The disconnected objects are assumed to be distant from each other if $\Delta x_i + \Delta y_i$ is more than 100 and the objects are unlikely belong to one sperm. This condition is the foremost condition needs to be considered to distinguish the disconnected object from other connected object (*i.e.* sperm without disconnected problem).

The first case of disconnected object is defined if the objects produced unique coordinates of *Candidate 1* and *Candidate 2* (*i.e.* or the objects are not in type-1 of non-unique coordinate) and unique coordinate of combination of both *Candidate 1* and *Candidate 2* (*i.e.* or the objects are not a type-2 of non-unique coordinate). This case can be represented by the example in Figure 11. Meanwhile, the second case is slightly different from case 1 where the type-2 of non-unique coordinate occurs as shown by the first example of Table 2.

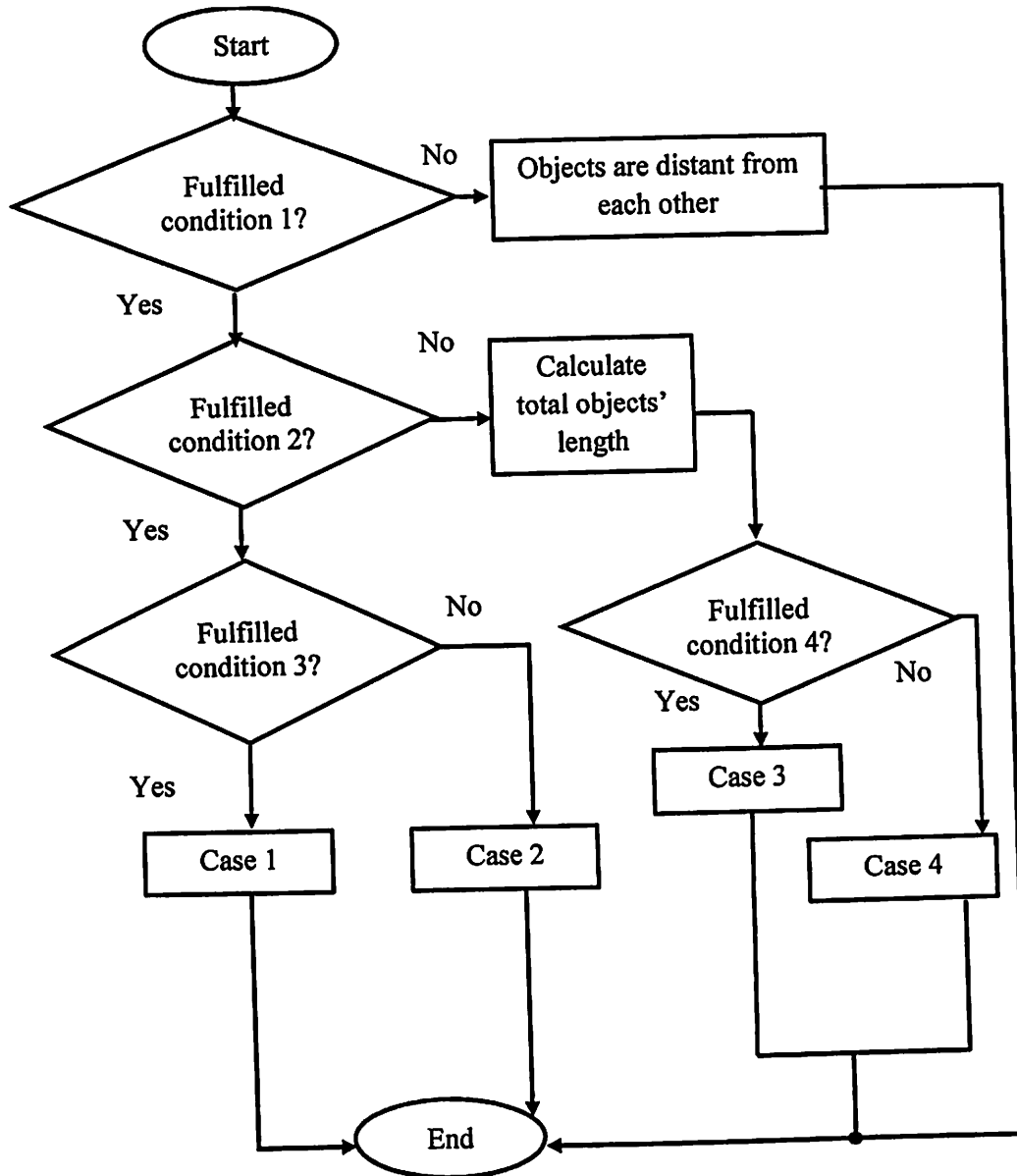


Figure 12 Identification process for four cases of disconnected object

The third and fourth cases of the disconnected objects are defined if type-1 and type-2 non-unique coordinates occur. Table 2 in second row shows the example of these cases where non-unique coordinates can be observed in *Candidate 1*, *Candidate 2* and combination of both. This situation indicates that the disconnected objects are adjacent or

collided with each other and thus another parameter needs to be computed to distinguish between these two cases. Therefore, total objects' length is computed using equation (25):

$$L_T = L_{o1} + L_{o2} \quad (25)$$

where L_T is the total length of the disconnected objects where summation of length of object 1 (L_{o1}) and length of object 2 (L_{o2}). L_{o1} and L_{o2} are computed based on the distance of between end coordinates. The distance is calculated based on equation (26).

$$L_{o(p)} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (26)$$

where x_1 and y_1 , x_2 and y_2 are the x and y coordinates to be processed where in this case they are belong to end coordinates. p is either 1 or 2 that represents objects in *Candidate 1* and *Candidate 2* respectively.

If the total objects' length is more than 100, the disconnected objects are more likely occluded with other connected objects (*i.e.* other sperm) which can be classified as Case 3. Meanwhile if the total objects' length is less than 100, the disconnected objects are classified as Case 4. The identification of four cases of disconnected objects is presented in Figure 12.

All cases of the disconnected objects must be attached to its own pair in order to ensure accurate detection. Failure to determine correct coordinate combination will cause wrongly connected objects and thus jeopardize the detection of the motile sperms.

When the correct coordinate combinations are identified, smooth line is plotted. This line is plotted to attach two disconnected objects based on the coordinates that are generated from equations (27) and (28).

$$x_{plot} = x_1 + \left[\frac{0 : L_{o(p)}}{50 : L_{o(p)}} \right] (\cos \theta) \quad (27)$$

$$y_{plot} = y_1 + \left[\frac{0 : L_{o(p)}}{50 : L_{o(p)}} \right] (\sin \theta) \quad (28)$$

where *distance* and θ are the distance and angle between two end coordinates and calculated using equations (26) and (29) respectively.

$$\theta = \tan^{-1} \left[\frac{(y_2 - y_1)}{(x_2 - x_1)} \right] \quad (29)$$

where in this case, x_1 and y_1 are the x and y end coordinates of *Candidate 1*. x_2 and y_2 are the x and y end coordinates of *Candidate 2*. x_{plot} and y_{plot} generate multiple coordinates that connect from end coordinate of *Candidate 1* to end coordinate of *Candidate 2*. Techniques involved in solving four cases of disconnected objects are presented in the next subsections.

2.3.1 Case 1

Case 1 of disconnected object is the simplest case to be solved. Similar to the other cases, this type of disconnected object is separated where $\Delta x_i + \Delta y_i$ is less than 100. In addition, this type of disconnected objects has unique coordinates between intra-group *Candidate 1*, unique coordinates between intra-group *Candidate 2* and unique coordinates between inter-group of *Candidate 1* and *Candidate 2*. However, this case is different from other cases since the disconnected objects resemble of an isolated sperm. The objects are far from other sperms where no occlusion scenario is observed. This case is easier to be solved since no other disconnected objects need to be considered. The disconnected object can be directly connected to its combination without any measurement and criteria to be fulfilled as shown in Figure 13. The disconnected objects of Q and R are attached to each other using the line plotted which their coordinates are generated using equations (27) and (28).

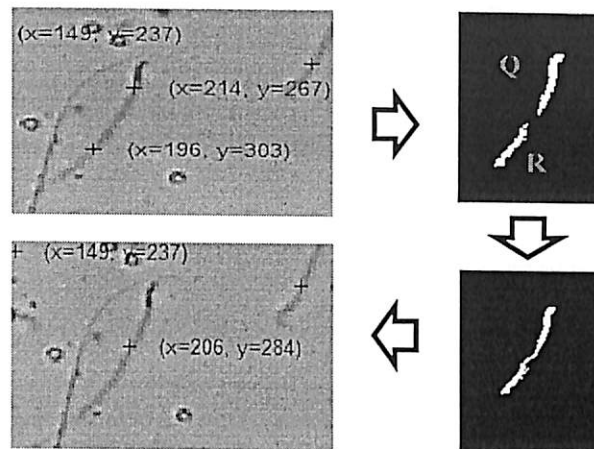


Figure 13 Implementation of Case 1 of the disconnected object

2.3.2 Case 2

As for Case 2 of the disconnected object, sperms are moving in parallel and closed to each other as shown in Figure 14. Objects I and J are disconnected from a single object of motile sperm while object K is another adjacent motile sperm. Initial tracker of motile sperms inaccurately detected three disconnected objects that caused false detection. There are three possible coordinate combinations between two disconnected objects: i) combination between objects I and J, ii) combination between objects J and K and, iii) combination between objects I and K. However, combination between objects I and K is neglected in the analysis since its $\Delta x_i + \Delta y_i$ is more than 100 as shown in the first example of Table 2.

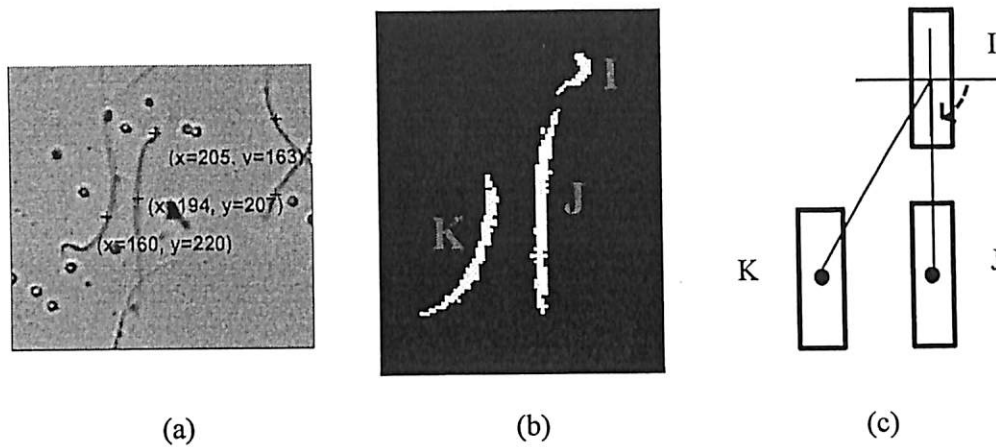


Figure 14 Parallel disconnected object in (a) grayscale (b) binary (c) model of parallel disconnected object where angles between centers are measured

Figure 14 depicts that by relying on the distance between objects or $\Delta x + \Delta y$, object J appears to be closer to K than its correct combination of object I. Therefore, another criterion is considered to solve this case, where the angle between centres of the disconnected objects is calculated (Figure 14 (c)). The angles between centers are calculated in clockwise direction and it is observed that angle between centers of objects I and K is higher than the angle between centers of objects I and J. With the assumption that the same orientation of disconnected objects indicates that they are belong to one object, the correct coordinate combination is chosen.

The accurate coordinate combination should attain minimum angle between centers as compared to the other pair combinations as shown in Figure 14 (c). The chosen coordinate combination is connected using the similar equations (equations (27) and (28) as Case 1 disconnected object and its implementation is presented in Figure 15.

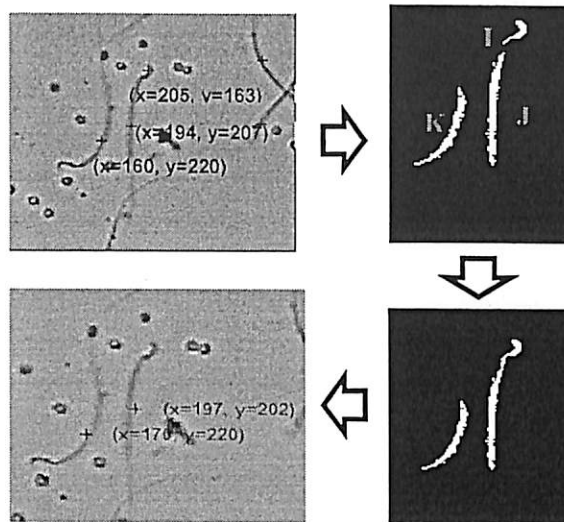


Figure 15 Implementation of Case 2 disconnected object case

2.3.3 Case 3

In this section, the most complicated disconnected object case is presented. This case is prominent to be solved since it is associated with clumped or occlusion between sperms. Figure 16 shows the Case 3 disconnected object in grayscale and binary images as well as model of intersected disconnected object. This figure shows that in the initialization of motile sperm tracker stage, five coordinates are detected and defined as sperm. However, only four of the detected coordinates are the disconnected objects (i.e. objects L and M as well as objects O and P) that should be attached to represent motile sperms.

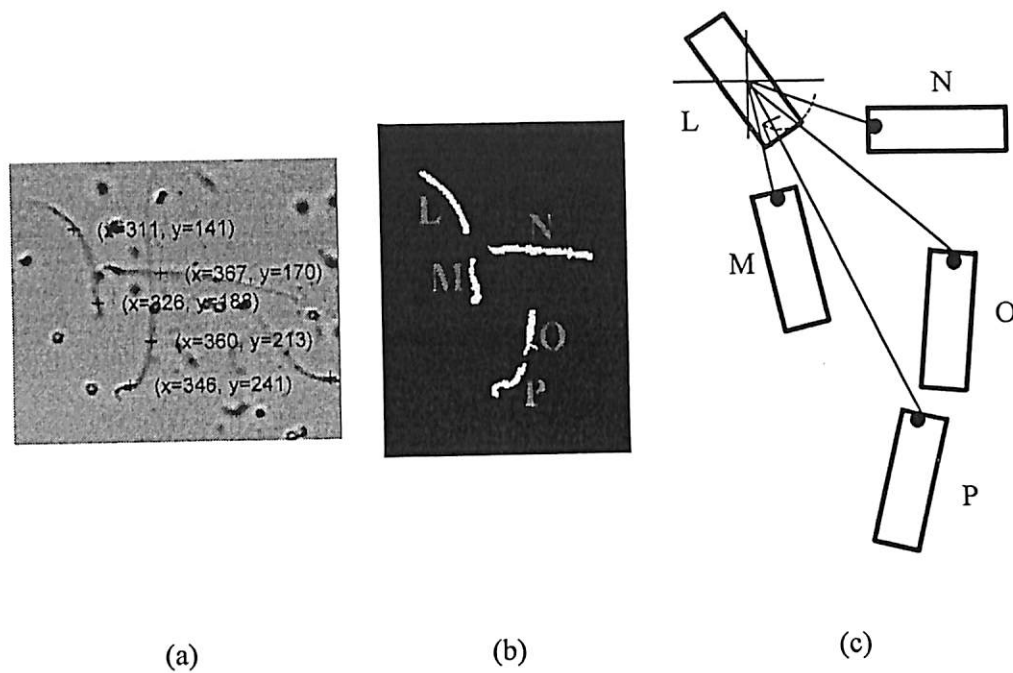


Figure 16 Case 3 disconnected object in (a) grayscale (b) binary (c) model of intersected disconnected object where angles between center and end objects are measured

There are 10 possible coordinate combinations as shown in second example of Table 2. However, combinations between objects L and O as well between objects L and P are eliminated since both have $\Delta x_i + \Delta y_i$ more than 100. For Case 3, two parameters namely angle between center and end coordinates (Figure 16 (c)) and distance between end coordinates (Figure B1, Appendix B) are introduced. Figure B1 shows the model of intersected object where the distances of the end coordinates are calculated. The distances between end coordinates are calculated from the other coordinates (*i.e.* black circle) with respect to its reference end coordinate denoted in red circle. This figure attests that small distance can be observed between objects L and M, objects O and P and objects L and N.

However, the disconnected object can only be attached to its own unique combination that has similar orientation. In other words, if object of *Candidate 1* is horizontal, thus its combination of *Candidate 2* must have horizontal orientation or vice versa. Therefore, to avoid non-unique coordinate combinations of objects L and M as well as objects L and N,

another parameter is measured. The angle between center and end of the disconnected objects are measured as shown in Figure 16 (c). The intersected disconnected objects are attached to their particular combination only if two conditions are satisfied; i) distance between end coordinates is minimum and ii) maximum angle between center and end coordinates that have similar orientation. If the angle of the centre and end coordinates is not equal to maximum, the coordinate combination is ignored since it is most likely to represent different orientation.

After the accurate coordinate combinations are identified, the disconnected objects are attached to their own combination for detection process. Misidentification of the five coordinates that are initially detected is rectified by connecting these objects using equations (27) and (28). The resulted motile sperms detected coordinates are presented in Figure 17. Connecting disconnected object with correct combination ensures that the intersected or collided sperms are successfully identified, thus eliminating the overestimated and underestimated motile sperm detection problems as often encountered in available CASA instruments.

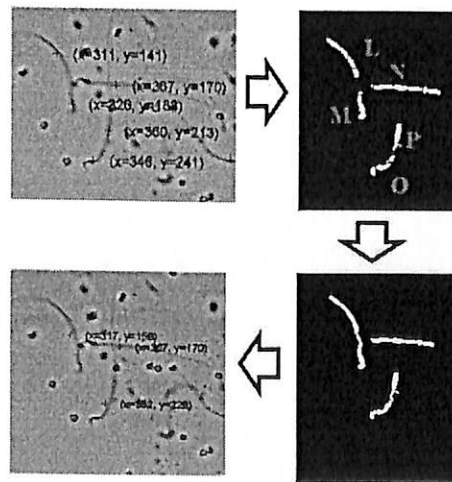


Figure 17 Implementation of Case 3 of disconnected objects

2.3.4 Case 4

Figure 18 shows the example of Case 4 of the disconnected objects in grayscale and binary images. It can be observed from this figure that two sperms are moving very close to each other. However, due to their fast motion of the sperms, proximal region in the midpiece of the sperms are less visible than their heads. This situation caused problem in the segmentation process where the less visible regions are not properly segmented which resulted in the disconnected objects as shown in Figure 18 (b).

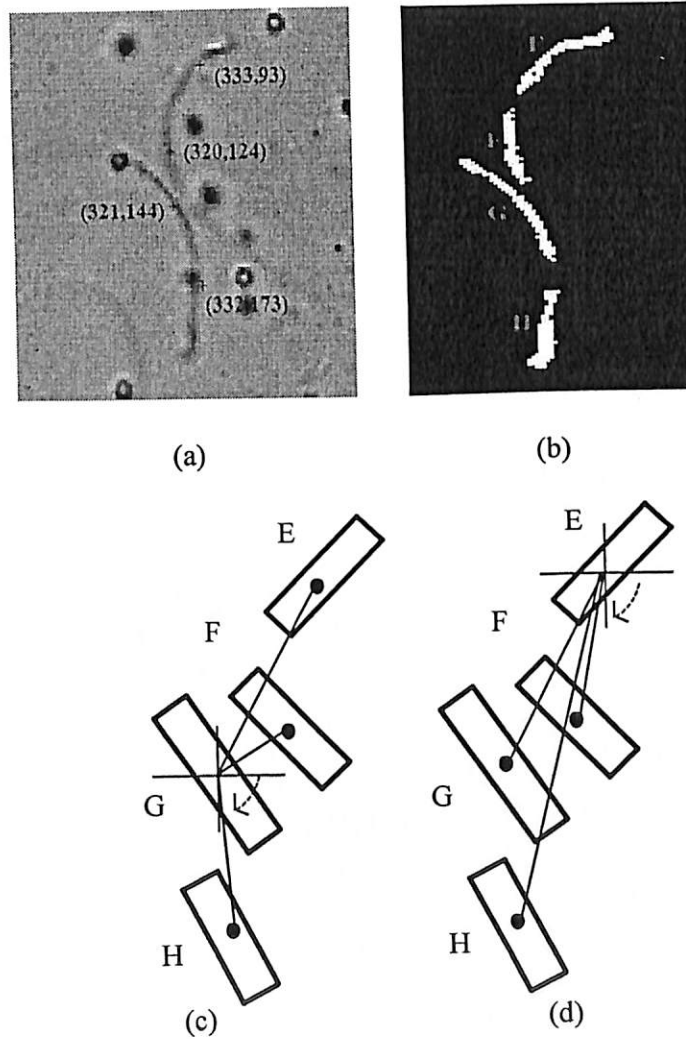


Figure 18 Case 4 disconnected object in (a) gray scale image (b) binary image (c) and (d) model of combined disconnected object where angle between center of the disconnected are measured

This figure shows four disconnected objects namely objects E, F, G and H, thus based on equation (24), there are six possible coordinate combinations of the disconnected objects. The possible coordinate combinations are combination between objects E and F, E and H, E and G, G and H, G and F, and objects F and H. The figure depicts that object F is closer to object G compared to its correct combination (*i.e.* object E), likewise, object G is apparently closer to object F than its correct combination of object H.

Therefore, another parameter is measured and acts as another criterion to be considered to find the correct combination of the disconnected object. Angle between two centers is computed for each possible coordinate combination as shown in Figure 18 (c) and (d). This angle is calculated in clockwise direction where the angles between two centers are computed for each possible coordinate combination with respect to object G and E respectively.

As an addendum to the measured angle between two centers, the distance between center and end of object is also computed as shown in Figure B2, Appendix B. This figure shows how the distances between center and end objects are calculated for all possible coordinate combinations. The correct combination of the objects is indicated by having a minimum distance between center and end of the object and minimum angle between two centers. Objects are classified as combination of coordinates to be connected if those criteria are fulfilled as shown in Figure 18 and Figure B2. These criteria are made based on the assumption that two disconnected objects must have similar orientation to be considered as one object (*i.e.* sperm). The smallest angles are obtained from combination of objects G and H (Figure 18 (c)) as well as combination of objects E and F (Figure 18 (d)). The results are also supported by the measured distance between center of object E and end of object F (Figure B2 (a)) or vice versa as well as distance between center of object G and end of object H or vice versa (Figure B2 (g)).

The correct combination of objects E and F as well as combination of objects G and H are then attached to each other using a line plotted based on the coordinates generated using equations (27) and (28). The implementation of connecting the combined disconnected objects is illustrated in Figure 19.

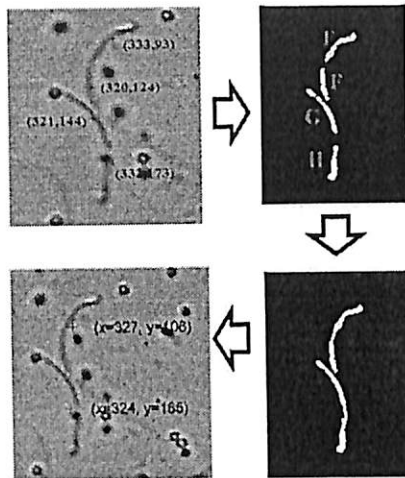


Figure 19 Implementation of Case 4 of disconnected objects

When the disconnected objects cases are solved and objects are successfully attached to their pair coordinate combination, final center coordinates of the sperms are identified. Table 4 summarises utilized criteria to solve each case of the disconnected object.

3 Results and Discussions

The performance and robustness of the proposed system are presented in this section. Table 5 summarises the experimental environment used in this analysis. The proposed system consists of three major contributions which are the pre-processing stage where the visibility of the image sequence frame is improved and sperm features are enhanced. Then, the unwanted debris are removed from the image sequence frame and sperm features are detected. The disconnected objects problems also resolved the occlusion cases that might occur in the sperm motility video.

Table 4 Criteria for each disconnected objects

Case	Parameter to be calculated to solve each case	Parameter chosen to solve each case
Case 1	<ul style="list-style-type: none"> • $\Delta x + \Delta y$ 	<ul style="list-style-type: none"> • No criteria to be chosen since the disconnected objects are far from other objects
Case 2	<ul style="list-style-type: none"> • Angle between two center 	<ul style="list-style-type: none"> • minimum angle between two centres
Case 3	<ul style="list-style-type: none"> • Distance between end coordinates of object • Angle between center and end coordinates of object 	<ul style="list-style-type: none"> • minimum distance between end coordinates of object • maximum angle between center and end coordinates of object
Case 4	<ul style="list-style-type: none"> • Angle between two center • Distance between centre and end coordinates of the object 	<ul style="list-style-type: none"> • minimum angle between two centres • minimum distance between centre and end coordinates of the object

Figures 20 to 22 show the sperm motility detection obtained by the proposed systems. The proposed system I consist of unenhanced image sequence frame while debris is removed and sperm features are detected. The proposed system II enhanced by fuzzy based histogram equalization (FHE) (Sheet et al., 2010) while debris removal and sperm features processes are conducted similar as previous system. Meanwhile in the proposed system III, the image sequence frame is enhanced by fuzzy quantitative measure (FQM) (Hanmandlu et al., 2009). In the proposed system IV, the image sequence frame is enhanced by the proposed AFIM and debris is removed as discussed in Section 2.1. Additional detection analysis is also presented in Appendix C in Table C1 to C2. The comparison with the other enhancement techniques is conducted to assess the enhancement role in sperm detection process. Both techniques are chosen for comparison analysis since they employed the fuzzy approach which is similar to our approach and dealt with the intensity vagueness that occurs in the converted image sequence frames.

Table 5 Experimental Environment

Elements	Environment
CPU	Intel ® Core™ i7-2600 CPU @3.4GHz
Memory	4GB
Software	Matlab R2011a
Camera	Pixelink P1B68 Microscope Camera
Resolution	689x602 (grey scale)
Frame/sec	30
Image Format	JPEG
Number of Frame/video	150frames/video
Number of sperm motility video	100 videos

The results presented in this section focus on the performance of the sperm detection process where the four metrics are employed to address the accuracy, sensitivity and specificity of the proposed system. These metrics include true positive (TP), true negative (TN), false positive (FP) and false negative (FN). All TP, TN, FP and FN values are calculated based on whether the system is able to detect the motile sperms with the aid of enhancement process. The effect of the enhancement process on the sperm detection is analysed. The detection is classified as TP or TN if the system correctly defines the immotile sperms or the systems correctly defines the motile sperms respectively. Meanwhile, FN represents the total number of the sperms that should have been classified as immotile sperms according to the physician, however classified wrongly by the system. FP is defined as the total number of motile sperms which according to the physician does not belong to the motile group however classified falsely into the group. These four classes are summarised in Table 6.

Table 6 TP, TN, FP, and FN cases

Detection by proposed system	Detection by physician	
	Motile	Immotile
Motile	TN	FN
Immotile	FP	TP

The accuracy, sensitivity and specificity of the system are calculated using equations (30) to (32).

$$Accuracy = \frac{TP + TN}{total\ data} \times 100\% \quad (30)$$

$$Sensitivity = \frac{TP}{TP + FN} \times 100\% \quad (31)$$

$$Specificity = \frac{TN}{FP + TN} \times 100\% \quad (32)$$

Accuracy denotes how close the proposed system achieved accurate detection in accordance to the detection manually made by physician. Meanwhile the sensitivity measures the proportion of positives which the system correctly identified as immotile sperms according to physician. Specificity measures the proportion of negatives which the system correctly identified as motile sperms according to physician.

Figure 20 shows an example of sperm detection analysis for *sample 1* video where in this frame, actual numbers of motile and immotile sperms are nine and seven respectively. This video is one of the examples where the debris occupied in most regions of the image frame while significant sperm features are lower in number than debris. Motile sperms are marked in green while the immotile sperms are marked in red. Numbers of TP, TN, FP and FN are also presented in this figure. As it is seen in the figure, the proposed method is successfully improve the overall brightness of the image frame and able to enhance important features in the frame (*i.e.* sperm). Even though the frame is dominantly filled with debris, with the aid of debris removal process, only important features are considered and thus accurate detection can be performed. It has been observed in Figure 20 (c), although the brightness of the original image frame is improved by the proposed system III, the uneven

and non-uniform illumination is still unsolved and center to top right regions appear darker. Thus sperm features are not highlighted properly which caused the false detection.

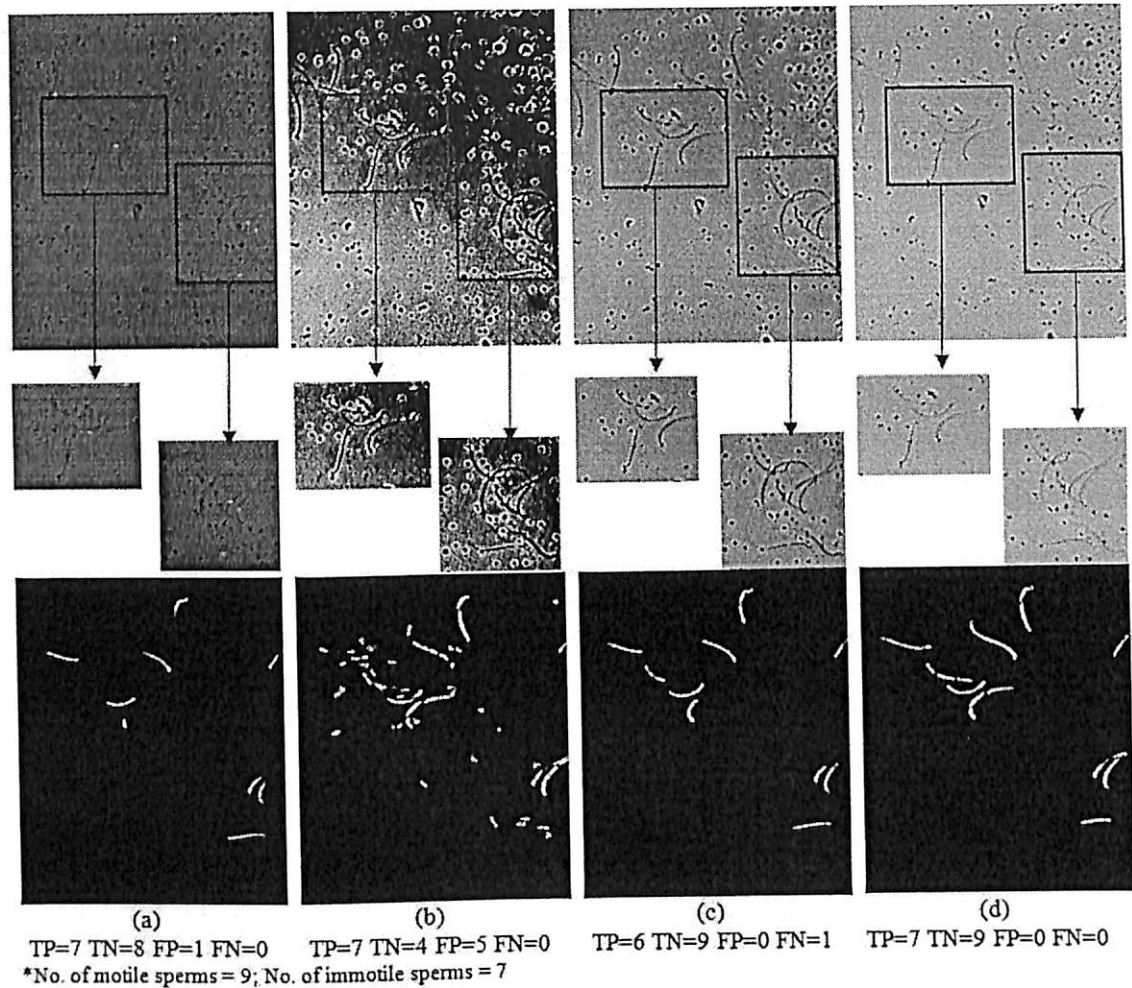


Figure 20 Sperm detection analysis on *sample 1* (a) proposed system I (b) proposed system II (c) proposed system III (d) proposed system IV

Meanwhile the enhanced image frame by the proposed system II is dominated by noise. Insignificant features are highlighted. Although object features (*i.e.* edges of sperm and edges of debris) are highlighted, the enhancement process by the proposed system II enhances the existing noise contains in the image frames and thus causes the over-saturation in left bottom region and under-enhanced in the right region of image frame. The under-

enhanced problem induced by the proposed system II gives significant impact on detection process where only four from nine motile sperms are detected as shown in Figure 20 (b). The intensity values near sperm and debris edges also contributed to the low detection rate by the proposed system II. This is because, the intensity values of sperm and debris edges are oversaturated (which resulted from FHE process) and it is difficult to obtain homogeneous segmented image frame during image conversion process. The non-homogeneous segmented image frame affects the initial tracking process and thus eventually reduces the detection rate.

The percentage of accuracy, sensitivity and specificity rates are plotted for overall frames in the *sample 1* video and presented in Figure C1 in Appendix C. The figure shows that the proposed method's performance in overall is better than other methods in terms of accuracy, sensitivity and specificity. It is attested from this figure that the detection performance is improved if the original image frame is enhanced by the proposed method as compared to enhanced frame by FQM in the proposed system III. The FHE has worsened the detection performance which the accuracy plot is accumulated approximately near 60% as can be seen in the proposed system II.

The proposed system IV in Figure 20 shows that the image sequence frame is successfully enhanced. The significant features of the motile sperms are highlighted and debris are successfully removed. The proposed system IV attains accurate detection of immotile and motile sperms with 7 and 9 sperms respectively.

Figure 21 shows another type of example where the number of debris and number of sperms is almost similar. Number of motile sperms consists in this *sample 2* video is 17 while number of immotile sperms consists in this video is 9. The figure also depicts the occlusion case where two or more sperms are intersected or collided with each other. Since the proposed system addressed this type of issue in the Disconnected-intersect (case 3) case, the occlusion case is solved and the sperms are successfully detected as shown in Figure 21 (b)

and (c). The visibility of the significant sperm feature is enhanced as shown in Figure 21 (b) (i.e. as shown in left red box). By using the fuzzy approach, the vague pixels (the pixels are hardly interpreted as bright or dark pixel) which have low intensity values that accumulated at the sperm shown in left red box of Figure 21 (b) are handled. These pixels are enhanced locally as discussed in subsection 2.1 thus the sperm that is not detected in original frame is successfully detected.

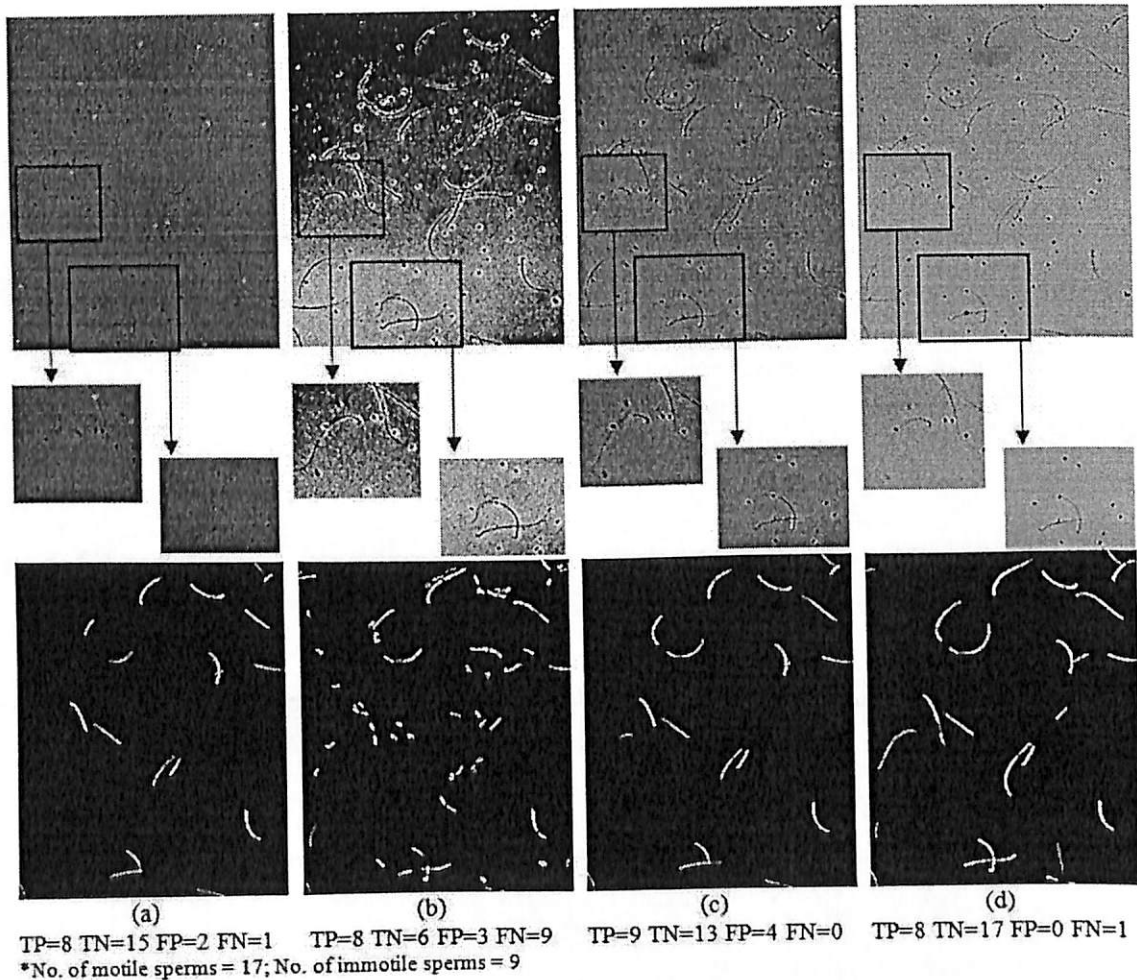


Figure 21 Sperm detection analysis on sample 2 (a) proposed system I (b) proposed system II (c) proposed system III (d) proposed system IV

Although the FQM and FHE are both implementing the fuzzy approach in their enhancement algorithm, both techniques under-enhanced or over-enhanced certain regions which resulted in the false detection. These results are also supported by the graph plotted for

accuracy, sensitivity and specificity of the overall performance of the detection system for sample 2 video in Figure C2, Appendix C. The enhanced images by FHE (*i.e.* the proposed system II) and FQM (*i.e.* the proposed system III) are defeated by the proposed method (*i.e.* AFIM) in overall performance for frame 1 to 150. However the percentage of accuracy for enhanced image frames by the proposed system III is better in most frames than the unenhanced image frames.

The proposed system IV successfully detects the motile and immotile sperms including the occlusion case. The intersected of motile and immotile sperms are detected as shown in the highlighted area (bottom rectangular). This is because, the intersected case is addressed and handled as discussed in Subsection 2.3.3(Case 3).

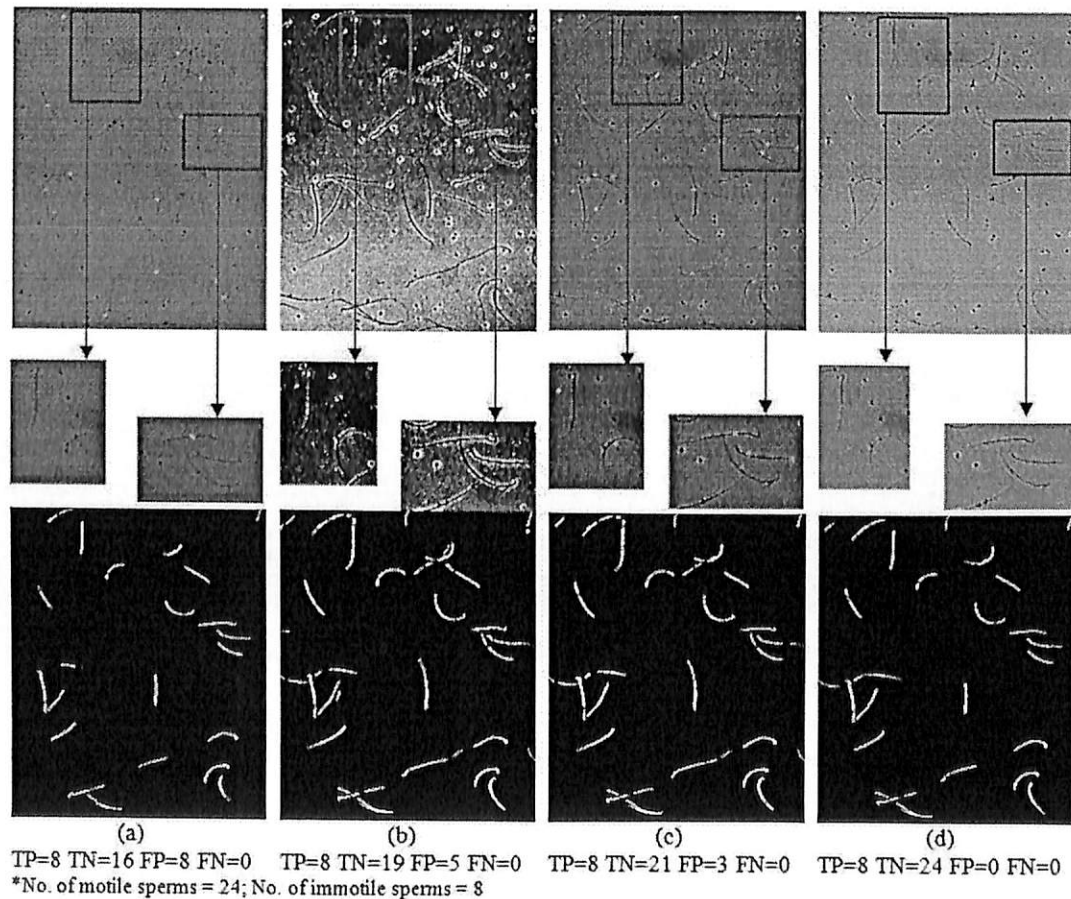


Figure 22 Sperm detection analysis on sample 3 (a) proposed system I (b) proposed system II (c) proposed system III (d) proposed system IV

In addition to the results presented in Figures 20 and 21, another example of detection analysis is presented in Figure 22 and Figure C3 in Appendix C. In Figure 22, total number of sperms is more than the total number of debris. Most regions in the image sequence frame are occupied by the motile sperms. The figure depicts that the debris size in this frame is smaller than those encountered in Figure 20. Thus, the debris are easily eliminated using the process presented in subsection 2.2.1 which can be seen in the majority cases either the image sequence frame is enhanced by proposed AFIM or other methods. However, when the brightness of image sequence frame is decremented, the detection process is bothered, therefore false detection is inevitable as shown in Figure 22 (c) and (d). Nonetheless, results show in Figure 22 attested that the enhancement process either by the proposed or other methods increases the TP rate compared with the unenhanced image sequence frame in Figure 22 (a). The sensitivity and specificity rate obtained by the enhanced image sequence frames by proposed method are higher in most frames as illustrated in Figures C3(b) and C3(c).

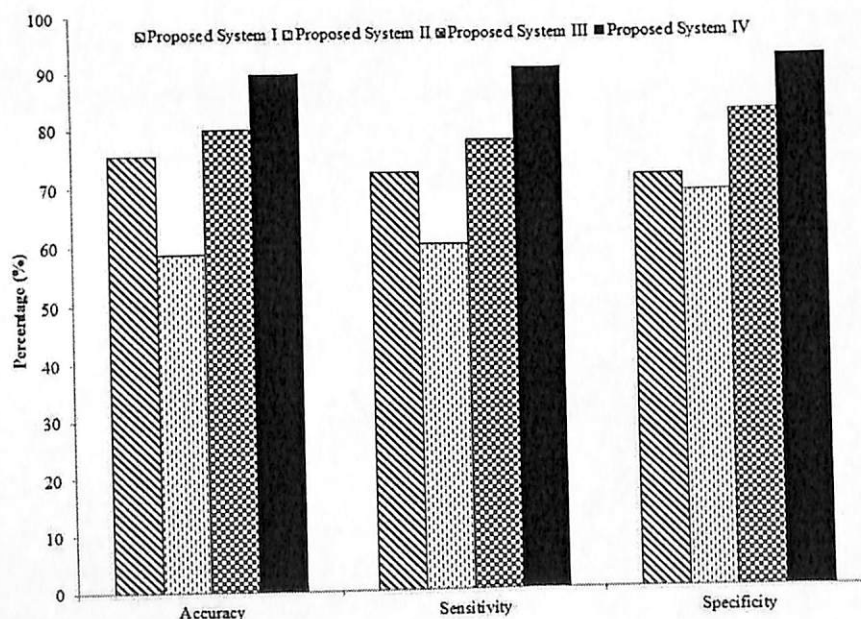


Figure 23 The average percentage of accuracy, sensitivity and specificity for 100 sperm motility video

The performance of the sperm detection analysis is also presented in Figure 23. In this figure, the average percentage of accuracy, sensitivity and specificity are plotted for 100 sperm motility videos. The average and standard deviation results for these 100 sperm motility videos are also tabulated in Table 7. In addition, the statistical paired t-test analysis is also performed to investigate the significant difference between unenhanced sperm motility video and the effect of the proposed enhancement technique implementation. The null hypothesis that has been set in this study is associated with the enhancement process by the proposed or other methods are significantly improved the accuracy of detection. Meanwhile the alternative hypothesis associated with the possibility of the observed difference between enhanced and unenhanced image frames are not likely have difference in terms of their accuracy of detection. The best results presented in Table 7 are made bold.

The proposed enhancement method is successfully enhanced the sperm motility video and attains the highest accuracy, sensitivity and specificity as tabulated in Table 7 and illustrated in Figure 23. The proposed method is also significantly different with p-value of 0.00072 from the unenhanced sperm motility videos. The proposed system III ranked the second highest which the accuracy of 80.06 % which is approximately 5% higher than the unenhanced sperm motility video (p-value of 0.00584). Meanwhile the proposed system II is ranked the lowest with the accuracy, sensitivity and specificity of 68.67%, 61.98% and 68.92% respectively with p-value of 0.00623.

Table 7 The average percentage and standard deviation of accuracy, sensitivity and specificity for 100 sperm motility videos

	Accuracy (%)	Sensitivity (%)	Specificity (%)	p-value (<0.005)
Proposed System I	75.62 ±21.03	71.85±19.63	72.36±20.35	N/A
Proposed System II	68.67±21.65	68.92±34.58	61.98±39.65	√
Proposed System III	80.06±23.58	82.83±29.63	77.94±24.78	√
Proposed System IV	82.67±12.39	92.34±13.62	90.63±14.68	√

*N/A – Not Applicable. The t-test and p-value rows are marked “N/A” because the comparison analysis is only conducted between the original frames with other enhancement methods. A ‘√’ is given in the ‘p-value’ row if the computed p-values are less than 0.005 for significant test.

The results presented in Figures C1 to C3 and in Appendix C prove that the enhancement process plays a major role in detecting the motile sperms. Presented results also proved that although the debris is successfully removed from the image frame and disconnected object issues are solved, the false detection is inevitable if the image frame is not properly enhanced. The capability of the proposed system in removing debris that occupies in most regions of the image frame as shown in Figure 20 is attested by the increment in total number of TN for most cases compared to original image frame.

4 Conclusions

The automated sperm-feature based sperm motility analyzer is proposed in this paper. The proposed system consists of three major contributions in which new enhancement process is proposed to enhance significant sperm features. In addition, the debris removal process is conducted to remove the unwanted debris occupied in most of the regions in the image frame. Problems encountered due to the fast moving sperms are solved where the disconnected objects are correctly attached. The occlusion case and sperms intersected with each others are solved where each individual sperm is identified thus overestimation and underestimation of motile sperms are avoided. The enhanced image sequence from the proposed enhancement technique (*i.e.* proposed system IV) provides an advantage in overall detection performance which the proposed system achieved 82.67% accuracy, 92.34% sensitivity and 90.63% specificity. As an addendum to the aforementioned positive results, the enhanced image by the proposed method is significantly different from the unenhanced image sequence frame which is attested by the statistical paired t-test analysis. The promising results presented in this paper can lead to real time sperm motility tracking where the occlusion cases are successfully handled to ensure accurate sperm motility detection.

Acknowledgement

This project is supported by the Ministry of Science, Technology & Innovation Malaysia through Sciencefund Grant entitled "Development of Computational Intelligent Infertility Detection System based on Sperm Motility Analysis."

References

- Bayly, P.V., Lewis, B.L., Kemp, P.S., Pless, R.B., Dutcher, S.K., 2010. Efficient spatiotemporal analysis of the flagellar waveform of *Chlamydomonas reinhardtii*. *Cytoskeleton* 67, 56-69.
- Beyan, C., Temizel, A., 2012. Adaptive mean-shift for automated multi object tracking. *Computer Vision, IET* 6, 1-12.
- Ching-Hei Yeung, Trevor G. Cooper, Nieschlag, E., 1997. A Technique for Standardization and Quality Control Sperm Motility Assessments in Semen Analysis. *Fertility and Sterility* 67, 1156-1158.
- Curi, S.M., Ariagno, J.I., Chenlo, P.H., Mendeluk, G.R., Pugliese, M.N., Sardi Segovia, L.M., Repetto, H.E., Blanco, A.M., 2003. Asthenozoospermia: analysis of a large population. *Arch Androl* 49, 343-349.
- Guzick, D.S., Overstreet, J.W., Factor-Livak, P., Brazil, C.K., Nakajima, S.t., Coutifaris, C., Carson, S.A., Cisneros, P., Steinkamp, M.P., A.Hill, J., Xu, D., Vogel, D.L., 2001. Sperm Morphology, Motility and Concentration in Fertile and Infertile Men. *The New England Journal of medicine* 345, 1388-1393.
- Hanmandlu, M., Verma, O.P., Kumar, N.K., Kulkarni, M., 2009. A Novel Optimal Fuzzy System for Color Image Enhancement Using Bacterial Foraging. *Instrumentation and Measurement, IEEE Transactions on* 58, 2867-2879.
- Hasikin, K., Mat Isa, N., 2013. Adaptive Fuzzy Intensity Measure Enhancement Technique for Non-Uniform Illumination and Low Contrast Images SIVIP DOI. 10.1007/s11760-013-0596-1.
- He, S., Jenkins-Keeran, K., Woods, L.C., 2004. Activation of Sperm Motility in Striped Bass via a cAMP-independent Pathway. *Theriogenology* 61, 1487-1498.
- Jensen, T.K., Jacobsen, R., Christensen, K., Nielsen, N.C., Bostofte, E., 2009. Good Semen Quality and Life Expectancy: A Cohort Study of 43,277 Men. *American Journal of Epidemiology* 170, 559-565.
- Mortimer, D., Mortimer, S.T., 1988. Influence of system parameter settings on human sperm motility analysis using CellSoft. *Hum Reprod* 3, 621-625.
- Mortimer, S.T., 1997. A critical review of the physiological importance and analysis of sperm movement in mammals. *Hum Reprod Update* 3, 403-439.
- Neuwinger, J., Behre, H.M., Nieschlag, E., 1990. Computerized semen analysis with sperm tail detection. *Hum Reprod* 5, 719-723.
- Omu, A.E., 2013. Sperm Parameters: Paradigmatic Index of Good Health and Longevity. *Medical Principles and Practice* 22(suppl 1).
- Schoevaert-Brossaul, D., 1984. Automated Analysis of Human Sperm Motility Computers and Biomedical Research 17, 362-375.
- Sheet, D., Garud, H., Suveer, A., Mahadevappa, M., Chatterjee, J., 2010. Brightness preserving dynamic fuzzy histogram equalization. *Consumer Electronics, IEEE Transactions on* 56, 2475-2480.
- Shiraishi, J., Li, Q., Appelbaum, D., Doi, K., 2011. Computer-Aided Diagnosis and Artificial Intelligence in Clinical Imaging. *Seminars in Nuclear Medicine* 41, 449-462.
- Suckow, M.A., Weisbroth, S.H., Franklin, C.L., 2005. *The Laboratory Rat*. Elsevier Science.
- WHO, 1999. WHO Laboratory Manual for the Examination of Human Semen and Sperm Cervical Mucus Interaction, In: Organization, W.H. (Ed.), 4th ed. Cambridge University Press, Cambridge, United Kingdom.

Yeung, C.H., Oberländer, G., Cooper, T.G., 1992. Characterization of the motility of maturing rat spermatozoa by computer-aided objective measurement. *Journal of Reproduction and Fertility* 96, 427-441.

Appendix A

Optimization of the Fuzzification Factor

The fuzzification factor differs for different input images or image sequence frames as discussed in the previous section. Instead of manually selecting α , this study proposes the automated approach on finding the optimum α . This procedure is performed to ensure the selected α will produce the most pleasant fuzzified image where the illumination are uniformly distributed.

α is selected based on the parameter value that yields the maximum image quality index, Q . Q is a new universal objective image quality index proposed by Zhou and Bovik (2002). This quality index measures any distortions as a combination of three different factors namely loss of correlation, luminance distortions and contrast distortions. These three components are derived as in equation () where the first component is the correlation coefficient while second and third components are luminance and contrast distortions respectively.

$$Q = \frac{\alpha_{mF}}{\alpha_m \alpha_F} \cdot \frac{2\bar{m}(\bar{F})}{(\bar{m})^2 + (\bar{F})^2} \cdot \frac{2\tau_m \tau_F}{\tau_m^2 + \tau_F^2} \quad (\text{A1})$$

where

$$\bar{m} = \frac{1}{Z} \sum_{y=1}^Z m_y \quad (\text{A2})$$

$$\bar{F} = \frac{1}{Z} \sum_{y=1}^Z F_y \quad (\text{A3})$$

$$\tau_m^2 = \frac{1}{Z-1} \sum_{y=1}^Z (m_y - \bar{m})^2 \quad (\text{A4})$$

$$\tau_M^2 = \frac{1}{Z-1} \sum_{y=1}^Z (F_y - \bar{F})^2 \quad (\text{A5})$$

$$\alpha_{mF} = \frac{1}{Z-1} \sum_{y=1}^Z (m_y - \bar{m})(F_y - \bar{F}) \quad (\text{A6})$$

In this study, the original and fuzzified images are assumed to have signals of $m = [m_y | y = 1, 2, \dots, Z]$ and $F = [F_y | y = 1, 2, \dots, Z]$ respectively. m_y and F_y are the intensity levels of the original and enhanced images, respectively. Q is computed by measuring local statistical features and then using a sliding window approach. It starts with the top-left corner of the image, a sliding window of size 3×3^1 moves pixel by pixel horizontally and vertically through all the rows and columns of the image until the bottom-right corner is reached. If there are total of H steps to reach the bottom-right corner, then the overall quality index is given by:

$$Q = \frac{1}{H} \sum_{h=1}^H Q_h \quad (\text{A7})$$

where Q_h is the local quality index.

Q is in the range of $[-1, 1]$ and the best value of Q (i.e. '1') is achieved if and only if m is equal to F . Thus, it is concluded that the optimum α must be selected if it can produce image with highest Q that closest to '1'. Q that closest to '1' indicates the fuzzified image attains homogeneous intensity distribution where the mean luminance error between fuzzified and original images is minimum. It demonstrates that the non-uniform illumination issue in the original image is solved by obtaining more homogeneous intensity distribution without causing any distortion in terms of contrast and luminance.

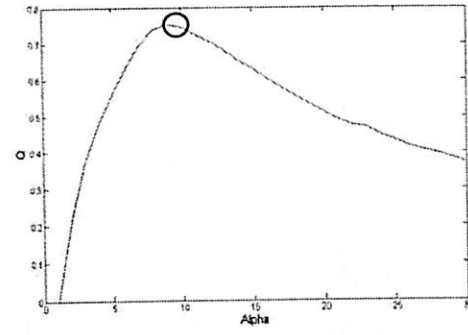
¹ The smallest sliding window of 3×3 is chosen in this study to ensure more accurate local statistical features can be measured.

The optimal procedure for selecting α is described as follows. For a given input image (*i.e.*, original image sequence frame), the value of α is varied from a minimum of 1 to a maximum of 30. For each value of α , the following automated tuning procedures are performed:

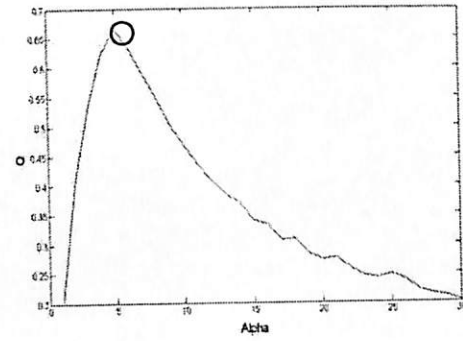
- i. Apply the algorithm presented in Section 2.1 to generate the fuzzified images
- ii. Calculate Q using equation (A1)
- iii. Select the parameter value that produces the maximum Q as the optimum value of α , after the two aforementioned steps.

Final output of fuzzified image is generated by adopting the optimum α . Simulations are performed on 100 standard images and 100 sperm motility videos that suffer from non-uniform illumination and low contrast problems to attest this automated optimization of fuzzification factor procedure.

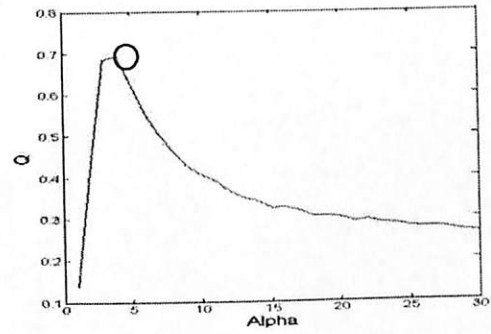
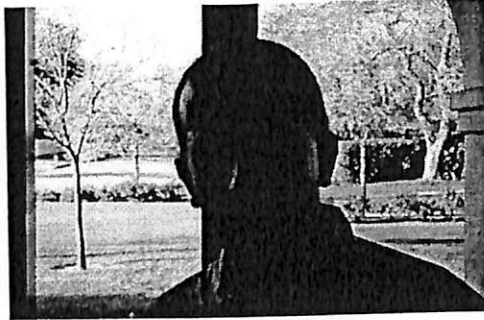
Examples of fuzzification factor optimization on three non-uniform illumination images are shown in Figure A1. The plots of Q in Figures A1 (a), (b), and (c) illustrate the changes in Q as α is varied from 0 to 30. The automated tuning is conducted until a homogeneous image is obtained. The homogeneous image is attained when Q reaches its maximum value. Figure A shows that Q reaches its highest value when alpha is 8, 5, and 4 as circled in Figures A1 (a), (b), and (c), respectively.



(a)



(b)



(c)

Figure A1 (a) – (c) images with non-uniform illumination and their optimization graphs of Q

Appendix B

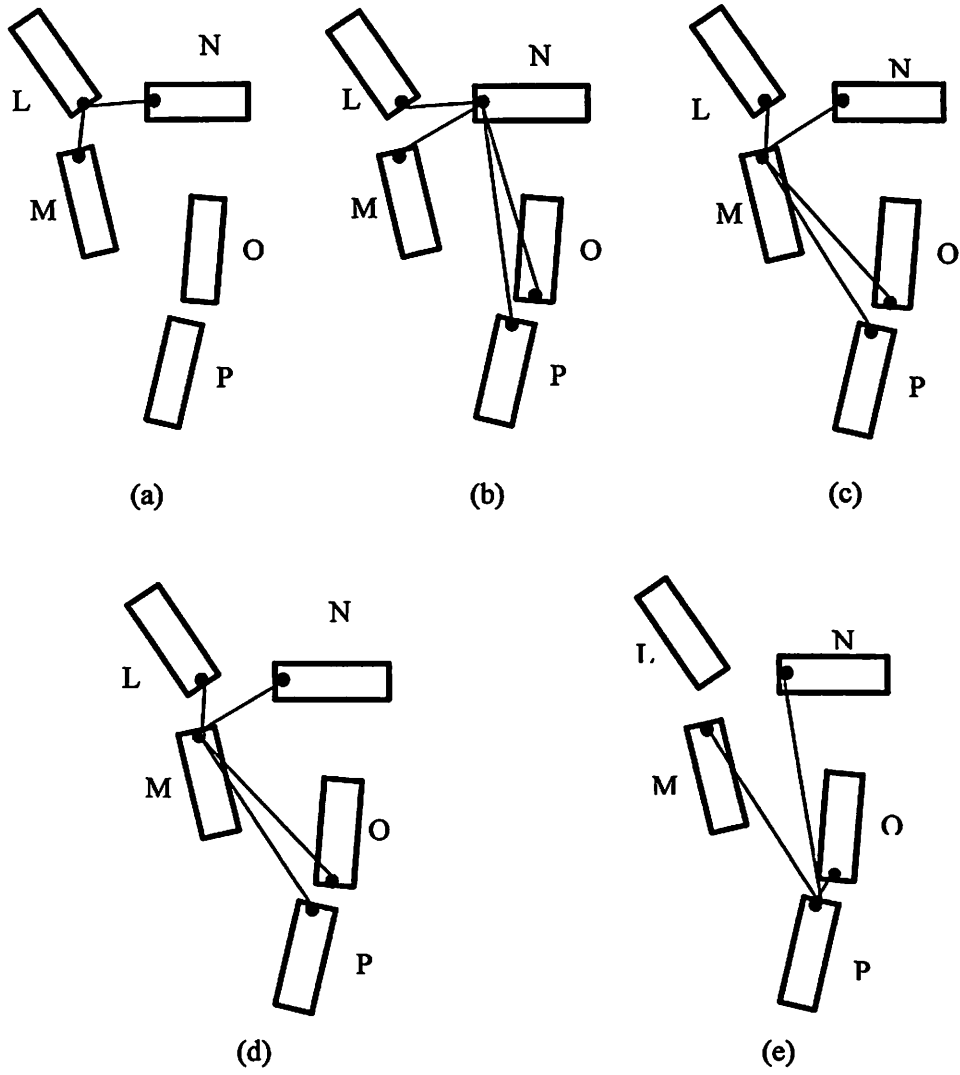


Figure B1 Distance between end coordinates with respect to (a) Object L (b) object N (c) object M (d) object O and (e) object P

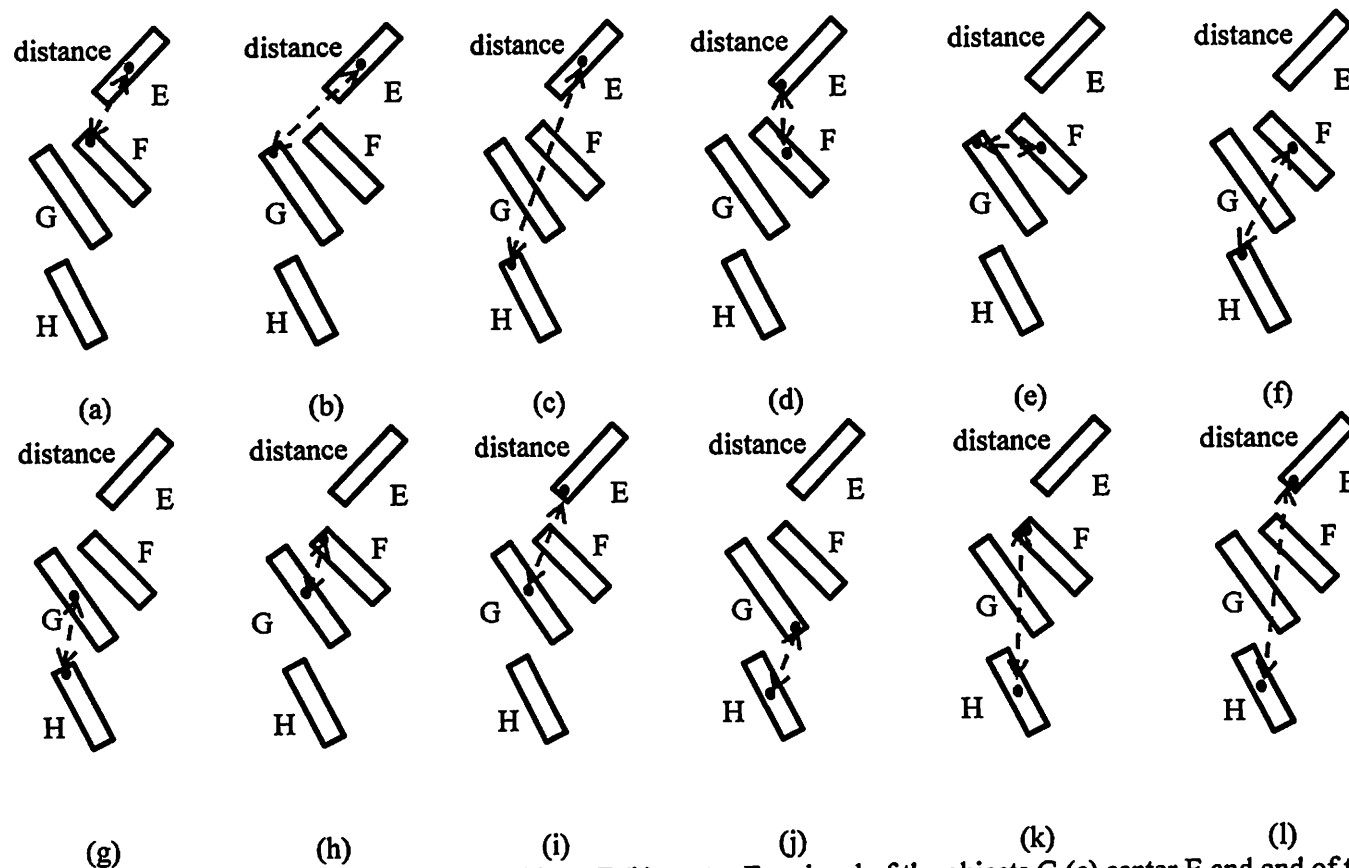
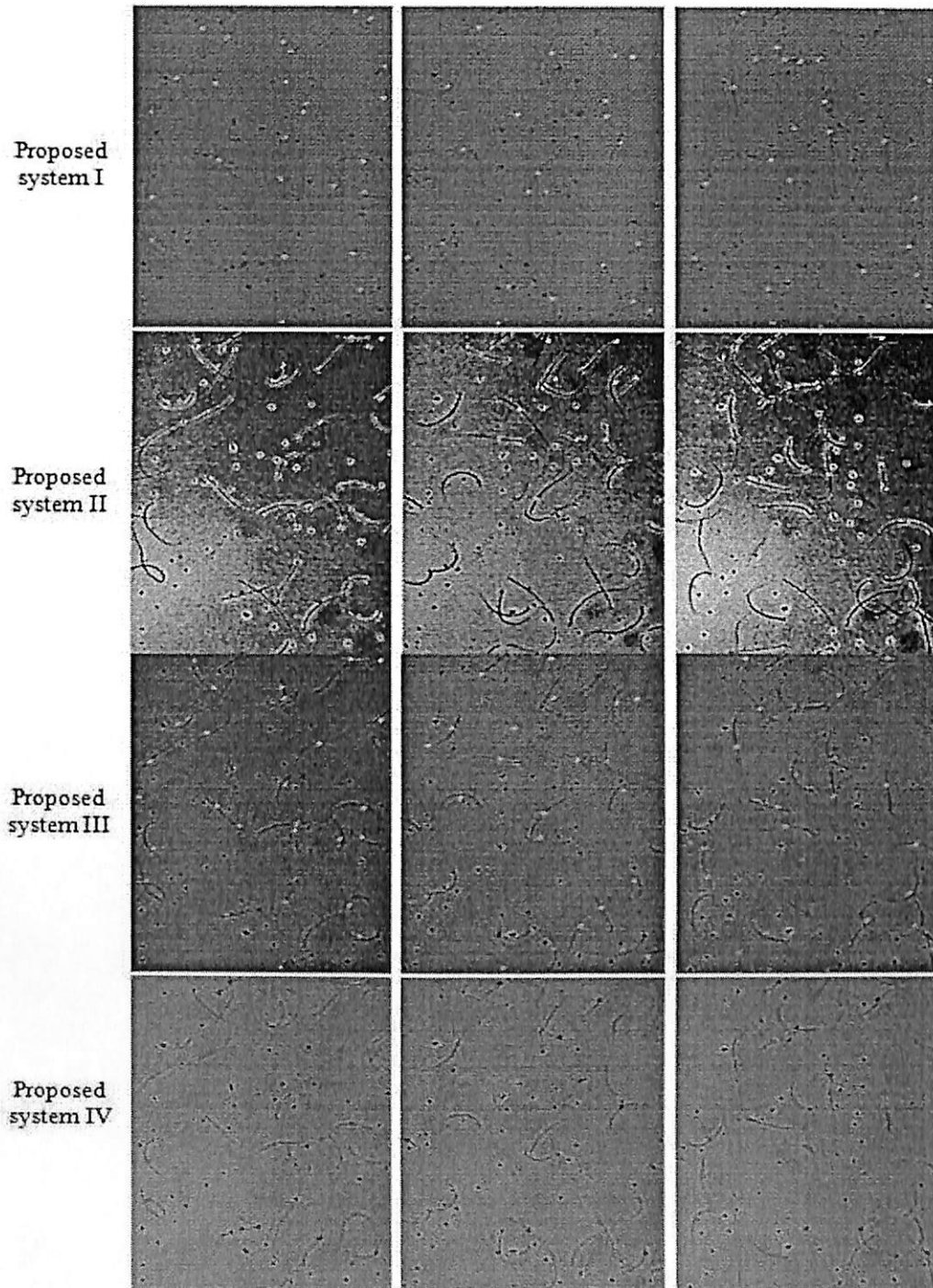


Figure B2 Distance between (a) center E and end of the objects F (b) center E and end of the objects G (c) center E and end of the objects H (d) center F and end of the objects E (e) center F and end of the objects G (f) center F and end of the objects H (g) center G and end of the objects H (h) center G and end of the objects F (i) center G and end of the objects E (j) center H and end of the objects G (k) center H and end of the objects F (l) center H and end of the objects

APPENDIX C

Table C1: Comparison on effect of different types of enhancement process on sperm detection. The video sample used in this table contains almost similar total number of debris and total number of sperms



*The motile sperms are marked in green color while immotile sperms are marked in red color

Table C1 (cont.) : Binary image from previous table

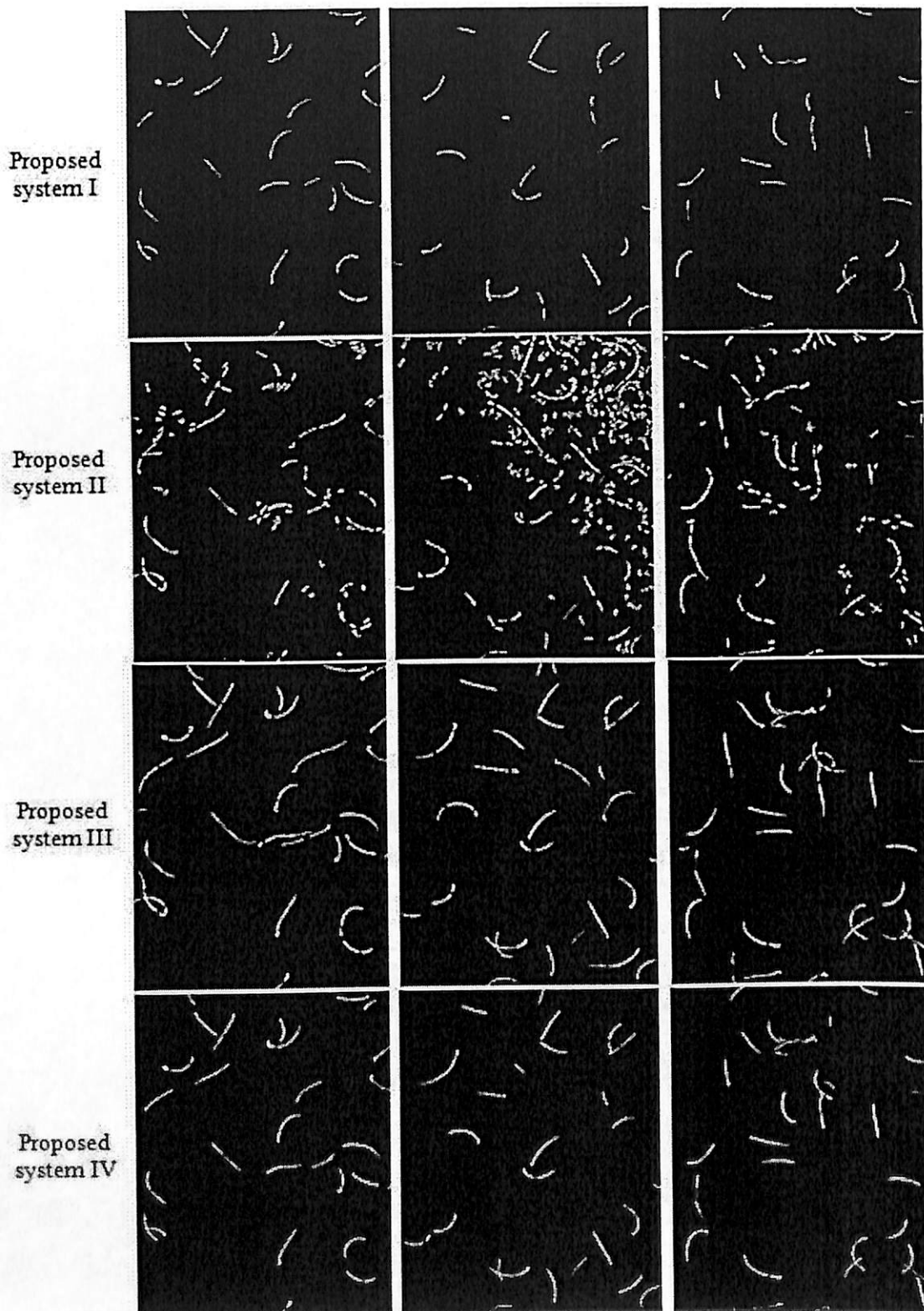
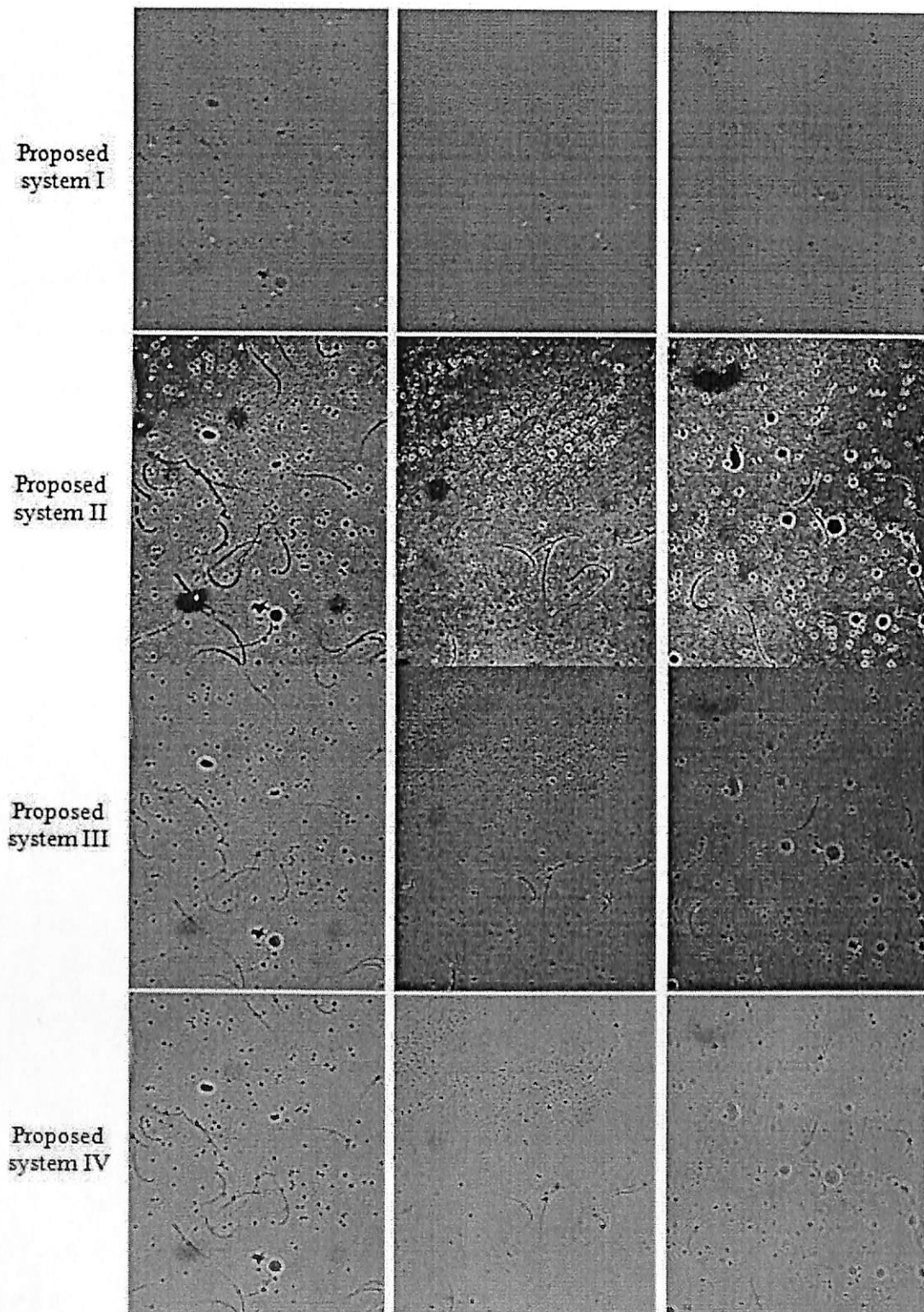
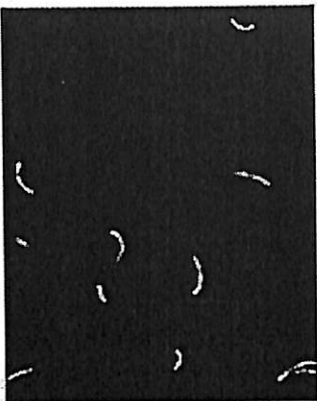













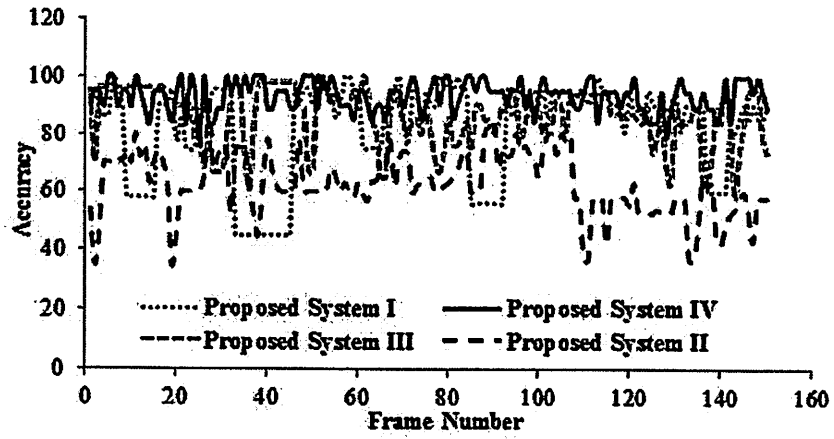
Table C2: Comparison on effect of different types of enhancement process on sperm detection. The video sample used in this table contains has more debris compared to sperms



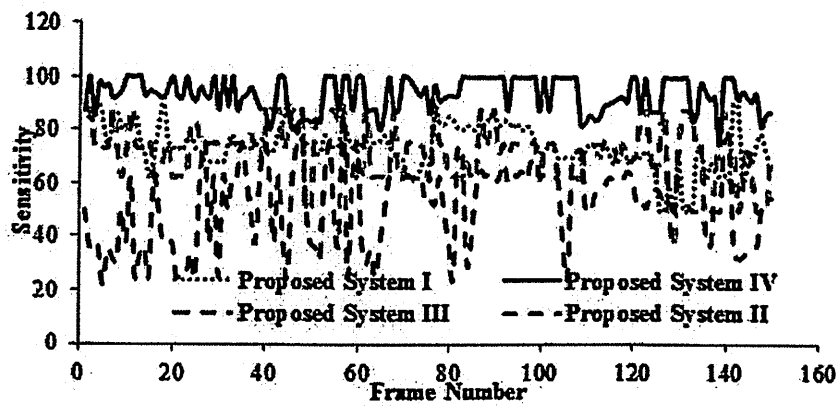
*The motile sperms are marked in green color while immotile sperms are marked in red color

Table C2 (cont.) : Binary image from previous table

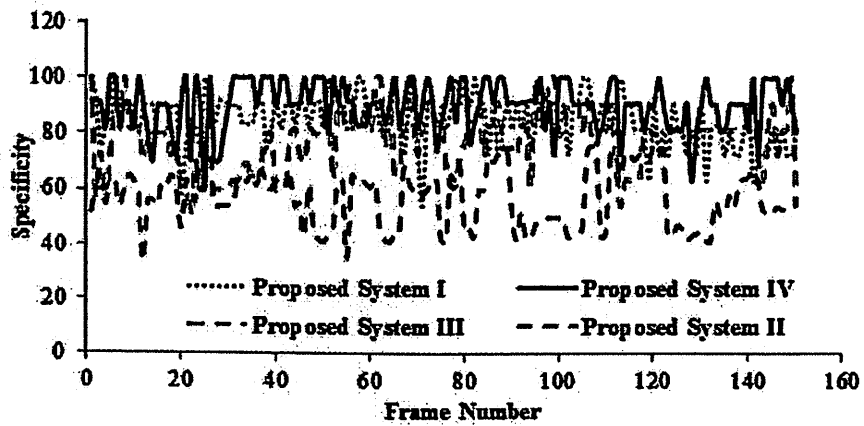
Proposed system I			
Proposed system II			
Proposed system III			
Proposed system IV			



(a)

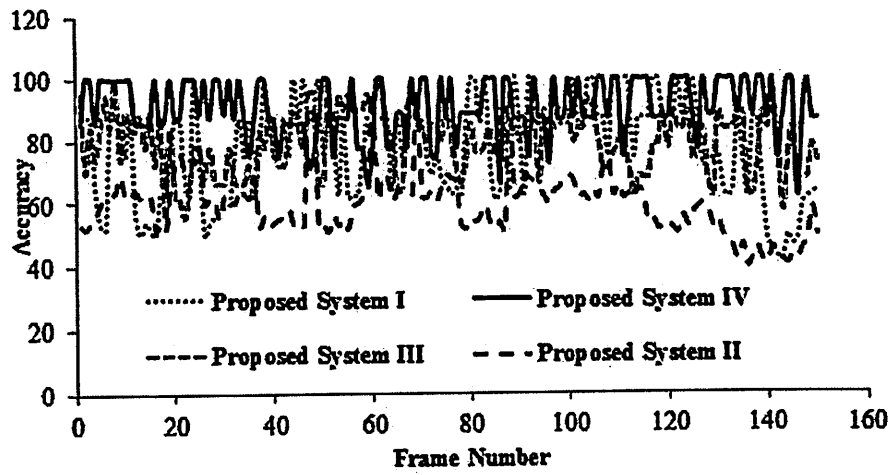


(b)

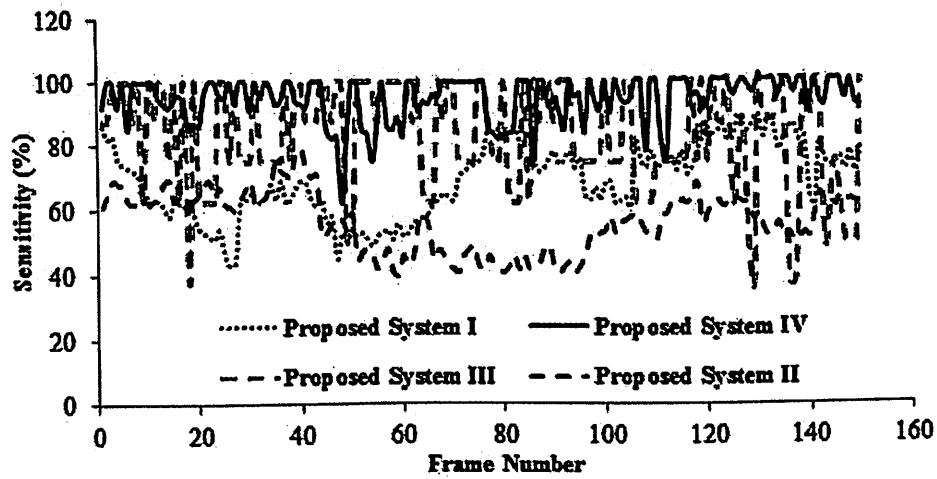


(c)

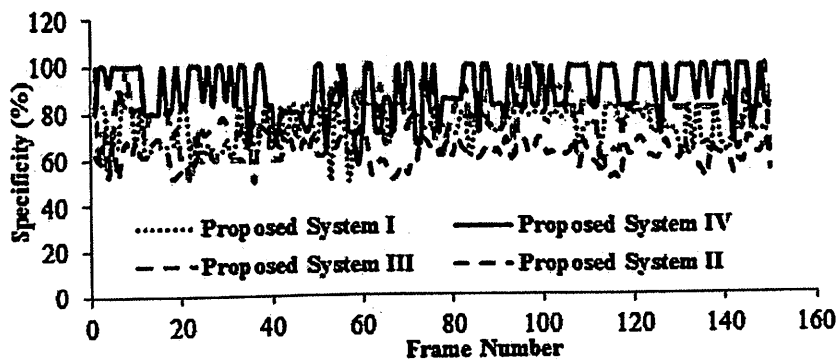
Figure C3: Percentage of detection analysis for video of sample 1 (a) accuracy (b) sensitivity (c) specificity



(a)

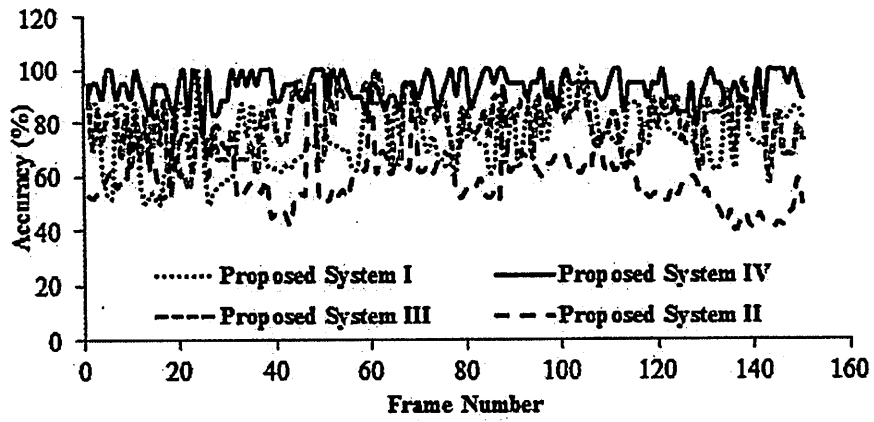


(b)

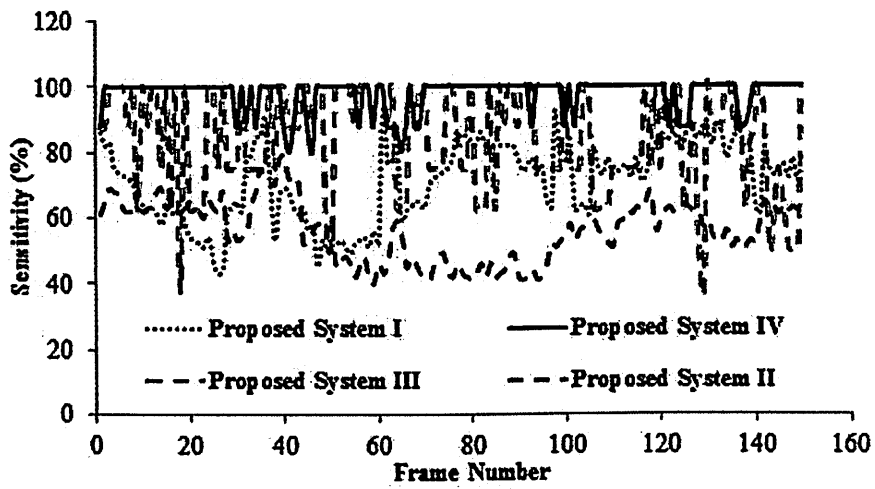


(c)

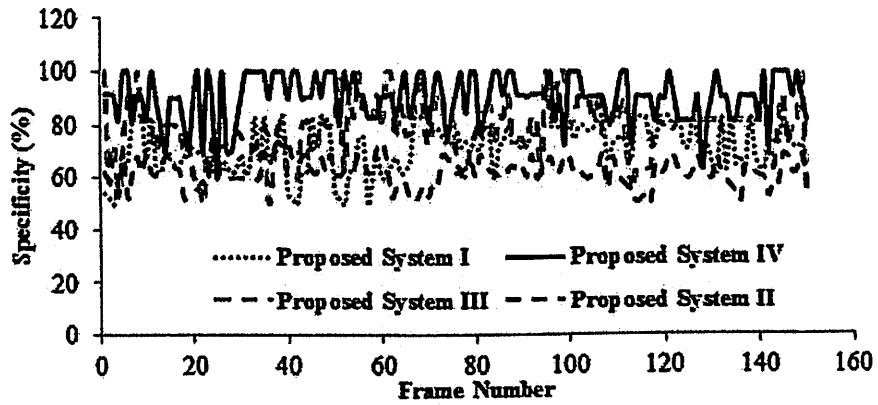
Figure C4: Percentage of detection analysis for video of sample 2 (a) accuracy (b) sensitivity (c) specificity



(a)



(b)



(c)

Figure C5: Percentage of detection analysis for video of sample 3 (a) accuracy (b) sensitivity (c) specificity

Evolving Output-Context Fuzzy System for Effective Rule Base

Md. Manjur Ahmed, Nor Ashidi Mat Isa

Imaging and Intelligent Systems Research Team, School of Electrical and Electronic Engineering, Engineering Campus, Universiti Sains Malaysia, 14300 Nibong Tebal, Penang, Malaysia

E-mail addresses: manjur_39@yahoo.com (M. M. Ahmed), ashidi@eng.usm.my (N.A. Mat Isa)

Abstract—This study develops an evolving output-context fuzzy system (EOCFS) that initiates the evolving process with a single fuzzy rule in which consequent and antecedent parts cover the whole output and input domains respectively. The EOCFS further evolves the output domain and adds more fuzzy rules to achieve an effective rule base. The proposed model is an online self-organizing method that can automatically identify prominent distinct data in the output domain for a new fuzzy rule. Thus, the EOCFS constantly evolves by reducing the model error. In addition, the evolving process is realized on the output domain while the self-adaptive process is achieved on the output domain and its associated input domain. The evolving termination index and uncertainty controller of the self-adaptive process are dynamically attained from past and current knowledge. Therefore, effective rule base is the balanced fuzzy model of the approximated system. To illustrate the effectiveness of the proposed algorithm, synthetic and real-world data are considered with low to high dimension inputs. Results show that the proposed EOCFS achieves better performance than the existing models with regard to accuracy and number of rules.

Index Terms—output-context fuzzy system, information granule, evolving system, dynamic constraint

1. Introduction

Information granularity and its relation to fuzzy systems exhibit the descriptive and functional representations of a global concept [1-2]. The level of abstraction conceived by humans from global knowledge implies interpretability and, hence, information granules. The *interpretability-accuracy tradeoff* of fuzzy models is essential, although this contradicting objective is still an open problem. Changing the level of abstraction conceived by humans enables this tailored approach to address the *interpretability-accuracy tradeoff* of the fuzzy model [3]. However, while the definition of accuracy in a specific application is uncomplicated, the definition of interpretability is relatively problematic.

A fuzzy granular model with conditional or context-based clustering was proposed by Pedrycz [3]–[5]. In earlier research on output-context clustering, the main theme was to define the output-context partition and then cluster the corresponding inputs. In Pedrycz's method, the number of contexts and the number of clusters per context are predefined and fixed. Hence, the computational model of the fuzzy system is manually designed by human experts [6]. Nevertheless, using an arbitrary number of output structure and input cluster is ineffective. The number of output context and its corresponding input clusters are based on the distinct nature of the data and, hence, local distribution. Given the prior knowledge in designing the fuzzy model and considering the limited competence of humans, the result is often highly prejudiced and uncertain. Furthermore, in Pedrycz's method [3], realization of the output context and its corresponding input clusters is fully independent, thus requiring refinements and refocus. Hence, the *stability-plasticity dilemma* [6]–[7] in Pedrycz's method divests the fuzzy system from incorporating past and future knowledge.

Existing models to overcome the aforementioned limitations are well documented in literature. Numerous algorithms to model the adaptive neural fuzzy systems and evolving fuzzy systems have been developed [6] – [20], such as evolving-construction scheme for fuzzy systems (ECSFS) [17], simplified structure evolving method (SSEM) [18], output-constrained cluster approach [19], semantic cointension (SC) [26], self-adaptive fuzzy inference network (SaFIN) [6], self-constructing neural fuzzy inference network (SONFIN) [9], and evolving neural-fuzzy semantic memory (eFSM) [7]. These fuzzy systems were used in attempts to design a consistent

and compact fuzzy rule-based system and thus ensure clear semantic meaning of fuzzy partitions/clusters with reasonable accuracy.

This study focuses on integrating the evolving and self-organizing system in a manner that it can be considered as an *interpretability–accuracy tradeoff* and a *stability–plasticity dilemma*. Unlike the grid partitioning discussed in [3], achieving some distinction points in the output domain to attain reasonable accuracy is significant [23]–[24]. A semantic rule base [6]–[20] is also important to enhance accuracy in a manner that it can be competently handled by humans [26]. In this paper, the evolving output-context fuzzy system (EOCFS) is proposed to solve the aforementioned problem.

The remainder of this paper is organized as follows. Section 2 describes the current trend of evolving fuzzy systems. Section 3 elaborates on the EOCFS methodology and its motivation and algorithms. It also describes the self-adaptive localized learning with some definitions and mathematical explanations. Section 4 covers several experiments to evaluate the performance of the proposed model. Both synthetic and real-world data sets are considered for evaluation. Section 5 discusses the main features of the proposed EOCFS compared with the existing methods. Section 6 summarizes the paper.

2. Current trend of evolving fuzzy rule-based systems

ECSFS [17] and SSEM [19], which are considered as error-reducing evolving methods, were proposed by Wang et al. In ECSFS and SSEM, the structure of fuzzy rule-based systems evolves by reducing the errors to fit the changes within the given systems, and this evolving process continues until it achieves the desired threshold accuracy. The termination criteria of ECSFS and SSEM are achieved through the global threshold accuracy. However, desired threshold and global threshold make these methods unintelligent. In addition, extremum and inflexion points are computed by the least square mean (LSM) method to obtain the highest accuracy in the next step, which leads to higher complexity in the algorithm. In ECSFS and SSEM, learning methods are based on global learning for the rule-consequent parameters and on localized learning for the rule-antecedent parameters. Lack of localized learning in the consequent part may cause under-fitting or over-fitting and, therefore, an unbalanced partition. The existing ECSFS and SSEM models divest the fuzzy system from incorporating past and future knowledge. Therefore, the

stability–plasticity dilemma is not addressed in the existing evolving models, and ECSFS and SSEM may select the noise data as an extremum or inflexion point.

Output-constrained cluster approach [18] and SC [26] were proposed to consider the output domain for partitioning the input data. In the output-constrained cluster approach [18], first, the output space is roughly partitioned by fuzzy c-means, and then data within each output constraint are further refined based on “separability,” which refers to the connectivity of the inputs. Prior knowledge for rough clustering in the output space makes a fuzzy system unintelligent. The results in [18] and [26] are highly subjective and uncertain because prior knowledge for designing the fuzzy system comes from users with limited competence. The output domain is evenly partitioned as in Pedrycz’s method [3] and thus ignores the local distribution of the input data. An evenly partitioned output domain may also cause under-fitting or over-fitting, thus leading to an inaccurate performance.

SONFIN [9], SaFIN [6], and eFSM [7] are self-organizing models that ensure online learning. *Stability–plasticity dilemma* is considered in these models, such that previous knowledge and new information are integrated with each other. Adaptation is made at the consequent and antecedent parts independently. Therefore, structure learning includes pruning the inconsistent or identical rules and deleting the orphaned rules. SONFIN and eFSM use the uniform coverage criterion (i.e., threshold) during structure learning. Rational partitions highly depend on the distribution of the output data. Hence, the threshold causes uncertain partitions in both models. In SaFIN, the clustering approach effectively addresses the *stability–plasticity* condition. SaFIN integrates new and old knowledge, such that a distinct cluster is formed simply by averaging the centers of left and right neighbors. Therefore, SaFIN also uses the uniform coverage criterion of each cluster because the spread and center of a cluster remain constant. Hence, a dynamic uncertainty function is needed to achieve a compact, consistent, and effective rule base.

3. Methodology

3.1 Motivation

The aforementioned limitations of the existing methods (as presented in Section 2) are employed in several propositions that motivated us to design an evolving fuzzy rule-based system. The incorporation of past and future knowledge is effective to avoid the evolving

threshold, which makes the methods more intelligent. A theoretical framework that considers the evolving process and self-organizing method can significantly achieve prominent distinction points on the output domain and hence address the *interpretability-accuracy tradeoff* and *stability-plasticity* conditions of the fuzzy model. Unlike the uniform coverage criterion (i.e., threshold) during structure learning [6]–[7], [9], a dynamic uncertainty function is important to achieve a compact and consistent fuzzy rule base.

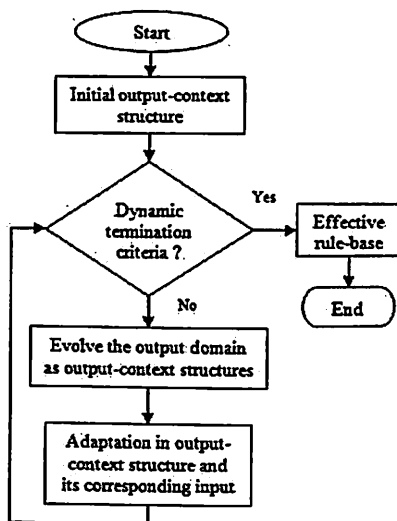


Fig. 1. Flowchart of the proposed EOCFS

The proposed EOCFS is an error-reducing evolving method. The main goal of this method is to achieve prominent distinctions in the output domain. The EOCFS partitions the output domain at a higher error region. Hence, the evolving process proceeds toward a lower error region. Figure 1 shows the flowchart of the proposed system. Each distinction point (or output context) and associated inputs form a structure, which is called the output-context structure. Refinements are made concurrently in the output-context structures and its associated input clusters in a manner that previous knowledge in the system and new knowledge from the training data are incorporated in the system and provide an accurate representation of the fuzzy model. The partitioning and evolving processes are continued until a termination criterion to obtain an effective rule base is fulfilled.

Consider a modeling problem with n input variables and N data samples. In addition, consider the i th training data as $[\mathbf{x}, d]_i$, where \mathbf{x}_i is the input vector and d_i is the corresponding output.

Let $[\mathbf{x}, d]_i$ belong to an output-context structure (s), which encodes an IF–THEN Mamdani-type fuzzy rule at t evolving stage as follows:

$$R^{t,s}: \text{IF } x_1 \text{ is } A_1^{t,(s)}, \text{ and } x_2 \text{ is } A_2^{t,(s)}, \dots \dots \dots, \text{ and } x_p \text{ is } A_p^{t,(s)}, \text{ THEN } y \text{ is } C^{t,(s)}, \quad (1)$$

where $C^{t,(s)}$ is the consequent part associated with the s th output-context structure and $A_p^{t,(s)}$ is the s th antecedent part associated with the p th input variable. A Gaussian membership function is described for each antecedent and consequent part, where c and σ are the center and width respectively, as expressed in Eq. (2):

$$\mu(c, \sigma; x) = e^{-((x-c)^2/\sigma^2)}. \quad (2)$$

The model output of the proposed EOCFS is defined for the s th rule (or s th output-context structure) based on Mamdani-type fuzzy systems. The model output is achieved using the center of averaging method [6,8,17], which is applied as follows:

$$o = \frac{A_{1,2,\dots,p,\dots,n}^{t,(s)}(\mathbf{x}) \times c^{t,(s)}|_{x \in \text{domain}_s}}{A_{1,2,\dots,p,\dots,n}^{t,(s)}(\mathbf{x})|_{x \in \text{domain}_s}}, \quad (3)$$

where $c^{t,(s)}$ is the center of the s th consequent part of the output-context structure ($C^{t,(s)}$) and

$$A_{1,2,\dots,p,\dots,n}^{t,(s)}(\mathbf{x}) = \prod_{\text{domain}_s} (\mu^{t,(s)}(\mathbf{x})) = \prod_{j=1}^n (\mu^{t,(s)}(x_j)). \quad (4)$$

Given that the s th output context and associated inputs form the s th output context structure, the width σ_p^{inp} of the p th input dimension is adapted to cover all input data associated with the s th output structure domain. Therefore,

$$A_{1,2,\dots,p,\dots,n}^{t,(s)}(\mathbf{x})|_{x \in \text{domain}_s} = 1. \quad (5)$$

Consequently, the output model of Eq. (3) can be rewritten as

$$o = A_{1,2,\dots,p,\dots,n}^{t,(s)}(\mathbf{x}) \times c^{t,(s)}|_{x \in \text{domain}_s}. \quad (6)$$

Hence, as shown in Eqs. (5) and (6), the accuracy of the model output depends on the tuning process of the output domain, and a balanced partition of the output context structure is needed to obtain the effective number of fuzzy rules with reasonable accuracy.

3.2 The proposed method

3.2.1 Initial output-context structure

The EOCFS employs the Mamdani-type fuzzy system and starts with an initial domain of output-context structure (or first fuzzy rule) by translating the knowledge of the global output domain. The formation of the first consequent part (when $t = 1, s = 1$) of the output-context structure $C^{t=1,(1)}$ can be described as

$$c^{1,(1)} = d_i \text{ and } \sigma^{1,(1)} = \sigma_{global}, \quad (7)$$

where d_i is the output of the i th training data and σ_{global} is the spread of the output domain for the initial fuzzy rule. d_i and σ_{global} of the Gaussian function are embedded in $C^{t,(1)}$. The formation of the antecedent part of the input cluster associated with the output-context structure $A_p^{t,(1)}$ can be described as

$$A_p^{t,(1)}(x)|_{x \in domain_{s=1}} = \prod_{domain_{s=1}} (\mu^{t,(1)}(x_p)) \text{ and} \\ c_p^{t,(1)} = x_p \text{ and } \sigma_p^{t,(1)} = \sigma_p^{inp}, \quad (8)$$

where σ_p^{inp} is the spread of the input cluster at the p th input dimension and is taken arbitrarily. σ_p^{inp} will be adapted with an uncertainty controller to cover all the input data at the p th dimension and to ensure that the input clusters have clear semantic meaning (which will be explained in Section 3.2.3).

Remark 1: The proposed EOCFS is based on a self-reliant and automated design from each training data. The concept of a self-organizing system [6]–[11] is used in the proposed EOCFS system to integrate the self-organizing numerical methods and the learning methods of the proposed system. A similarity measure (SM) of two consecutive Gaussian-shaped fuzzy levels has been described in [16]. $SM \rightarrow 1$ shows more similarities while $SM \rightarrow 0$ shows more dissimilarities between the two consecutive Gaussian-shaped fuzzy levels. A similarity threshold (β) is considered to identify the similarity between output-context structure and presented value. If the similarity measure between output-context structure and presented value exceeds the

threshold value ($SM > \beta$), then a prominent distinction is observed. Conversely, $SM \leq \beta$ shows the similar linguistic meaning between the output-context structure and presented value.

Global width (σ_{global}) is required to form the initial output-context structure shown in Fig. 2. The step-by-step process for the initial output-context structure is referred to as Algorithm 1.

Algorithm 1: Initial output-context structure
 Requires: Minimum σ_{global} value is arbitrarily taken.
Step 1: Initial fuzzy term for consequent part $C^{1,(1)}$ as (7).
Step 2: FOR each training data $[x, d]_i$
 Define μ for new incoming d_i using (2) and (7).
 IF $SM \leq \beta$ between new data and existing consequent part(s),
 then CONTINUE with *Step 2*.
 ELSE return to *Step 1* by increasing σ_{global} .
Step 3: Global sigma value: Final σ_{global} is required to form the initial fuzzy rule (or initial output-context structure) using (7) and (8).

Fig. 2. Algorithm for the initial output-context structure

3.2.2 Evolving output-context partition and adding fuzzy rules

After defining the initial output-context structure (or first rule), the EOCFS starts evolving the output-context structure by partitioning the output domain. The spread of the output domain of the previous evolving stage (σ_{init}^{t-1}) decreases at a constant rate σ_{evolve} to obtain a new balanced partition. The constant width σ_{evolve} is taken arbitrarily, such that the output partition slowly increases. Figure 3(a) shows the initial rule in which the EOCFS locates the data sample with the largest error. Then, the evolving process starts and the EOCFS splits the output domain and adds rules/output context to reduce the training error, as shown in Figs. 3(b) and 3(c). The spread of the output domain of the previous evolving stage (σ_{init}^{t-1}) is the basis for the next step of the evolving process. All the parameters are locally optimized for each evolving stage. This evolving output-context partition is self-determined from the training data set, in contrast to the greedy algorithm used in [17]. Evolving (described in this section) and adaptation (which will be described in Section 3.2.3) are carried out concurrently for each evolving stage to obtain the effective rule base. Afterward, testing data are applied to this effective rule base to obtain reasonable accuracy.

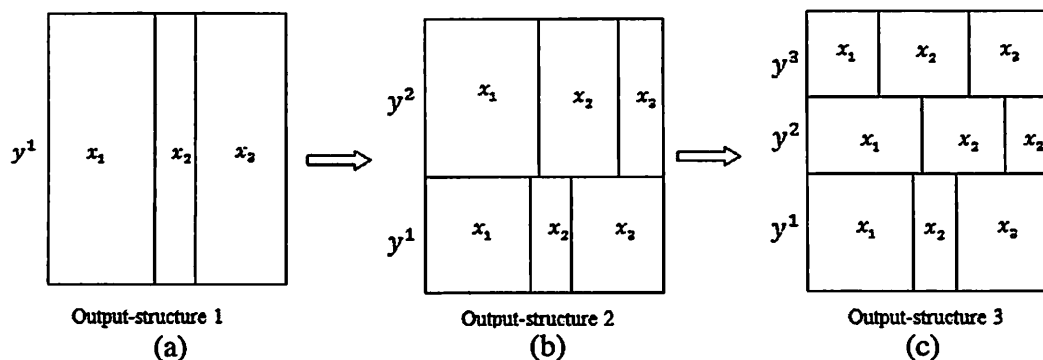


Fig. 3. Output-context structure (or rule) creation for the proposed EOCFS. $y^s = f(\mathbf{x}) = f(x_1, x_2, \dots, x_n)$, where y^s symbolizes the output context of the s th partition ($s = 1, 2, 3, \dots$) and \mathbf{x} represents the corresponding input feature space.

To describe the evolving process, we consider $t = 1$; thus, $s = 1$ means that at the first evolving stage, only a single rule exists and more fuzzy rules and terms should be added. The concept of a self-organizing system [6]–[11] is used for self-determination of the fuzzy levels from each training data. Consider the i th training data $[\mathbf{x}, d]_i$, where \mathbf{x}_i is the input vector and d_i is the corresponding output. At evolving stage $t = 2$, the formation of the first output-context structure ($c^{2,(1)}$) where more than one rule generation ($s > 1$) is considered can be expressed as follows:

$$c^{2,(1)} = d_i \text{ and } \sigma^2 = \sigma_{init}^2 = \sigma_{init}^1 - \sigma_{evolve}. \quad (9)$$

Recall that σ_{evolve} is taken arbitrarily such that the output partition slowly increases. The formation of the antecedent part of the input cluster associated with the output-context structure ($A_p^{2,(1)}$) can be described by Eq. (10). Recall that σ_p^{inp} is the width of the input cluster at the p th dimension and will be adapted to cover all inputs at the p th input domain associated with the output-context structure, expressed as follows:

$$c_p^{2,(1)} = x_p \text{ and } \sigma_p^{2,(1)} = \sigma_p^{inp}. \quad (10)$$

By assuming that an output-context structure exists, the EOCFS then proceeds with computing the similarities between the presented value and existing output-context structures. The similarity match between the output d_i of the i th training data $[\mathbf{x}, d]_i$ and an existing output-

context structure ($C^{t,(s)}$) is expressed as $SM(d_i, C^{t,(s)}) = \mu^{t,(s)}(c^{t,(s)}, \sigma^{t,(s)}; d_i)$. A similarity threshold (β) is considered to identify the best matched output-context structure with the presented value, which is described in Remark 1. If $SM > \beta$, then a prominent distinction is observed and the EOCFS forms a new distinct output-context structure at evolving stage t , which is expressed as follows:

$$c^{t,(s+1)} = d_i \text{ and } \sigma^{t,(s+1)} = \sigma_{init}^t = \sigma_{init}^{t-1} - \sigma_{evolve}. \quad (11)$$

Conversely, if $SM \leq \beta$, then the EOCFS determines the best matched output-context structure denoted as $C^{t,(s^*)}$, where $s^* = \arg \max_{SM \leq \beta} SM(d_i, C^{t,(s)})$. Therefore, the best matched output-context structure can give a satisfactory description of the presented value. Consequently, merging, adjustment, and refinement are applied to the best matched output-context structure ($C^{t,(s^*)}$), as shown in Eq. (12), to adapt the presented value. Therefore, at evolving stage t ,

$$C^{t,(s)} = L(c^{t,(s^*)}, \sigma^{t,(s^*)}; C^{t,(1,2,\dots,S^t)}), \quad (12)$$

where function $L(\cdot)$ defines the linguistic meaning in association with the uncertainty control and will be described in Section 3.2.3. Previous output-context structures are defined as $C^{t,(1,2,\dots,S^t)}$, where S^t is the total number of existing output-context structures at stage t . Figure 4 shows a nonlinear function $f(x)$ in black line that should be approximated, where $f(x, y, z) = (1 + x^{0.5} + y^{-1} + z^{-1.5})^2$. A description of this nonlinear function will be explained in Example 3 in Section 3.3. At the first evolving stage ($t = 1$), the EOCFS starts with a single rule and locates the data sample with the largest error. The evolving process is then continued and is determined to be an effective rule base when the number of rules equals four. The EOCFS determines four distinct center points and forms four new distinct output-context structures. Therefore, four distinct input domains associated with the distinct output structures are formed. References [17], [23], and [24] previously described the selection of splitting points to reduce approximation errors. However, unlike ECSFS [17], where the LSM algorithm is used to select splitting points, the proposed EOCFS evolves and self-determines the distinct output-context structures to obtain an effective rule base.

Remark 2: The terms *adaptation* and *evolving* are used in the proposed EOCFS algorithm. *Adaptation* is used to regulate the new output knowledge and the existing output structures. It is

also used to regulate new input knowledge and existing input clusters. The proposed EOCFS algorithm is a novel self-adaptive fuzzy system that employs the automated formulation of fuzzy rule base. Hence, a linguistic function $L(.)$ with an uncertainty controller is needed to regulate and adapt the output-context structures within the output domain. Similarly, regulation and adaptation are preferable for the input clusters within the input domain. Therefore, a localized learning or *adaptation* of the fuzzy system is conducted to keep the semantic meaning and accuracy of the proposed EOCFS algorithm. A more detailed description of linguistic function, adaptation, and refinement is presented in Section 3.2.3. In contrast, the EOCFS also uses the *evolving* system to evolve the output structures by adding fuzzy rules and terms. The goal is to determine the effective rule base from the training data set through the *evolving* method.

Standard root mean square error (RMSE) of the t th evolving stage is expressed in Eq. (13) by comparing the model output $o(x_i)$ and the desired output des_i of the i th training data:

$$\text{Error function, } E(t) = \sqrt{\frac{1}{N} \sum_{i=1}^N (o(x_i) - des_i)^2}. \quad (13)$$

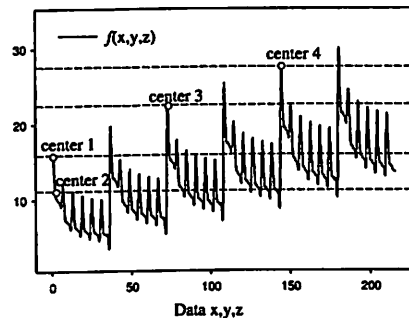


Fig. 4. Evolving selection of the center points shows the distinct output-context structures (rules)

3.2.3 Fuzzy partitions by dynamic constraints: introducing pseudo-sigma

This section discusses localized learning with adaptation and refinement. In the proposed EOCFS, dynamic constraints are considered for each output-context structure and its associated input cluster. Recall that each output-context structure depicts a single rule.

A. Partition of the output-context structure

The EOCFS uses the dynamic constraints for uncertainty control if, and only if, the similarity measure $SM \leq \beta$ is determined between any existing output-context structure and the presented value (see Remark 1 and Section 3.2.1). Assuming that C^{new} is the new consequent fuzzy level for d_i , where training data $[x, d]_i$, the new fuzzy level for d_i is defined as

$$c^{new} = d_i \text{ and } \sigma^{new} = \sigma_{init}^t, \quad (14)$$

where c^{new} and σ^{new} are the center and spread of the new fuzzy level (C^{new}) respectively. We assume that EOCFS determines the best matched output-context structure denoted as $C^{t,(s^*)}$ (see Section 3.2.2). Therefore, dynamic constraints for the output-context structure are used to adapt and refine $C^{t,(s^*)}$ with the presented value C^{new} . Definition 1 describes the conditions that should be considered while adapting between two consecutive fuzzy labels.

Definition 1: Given that $C^{t,(s^*)}$ is the best matched output-context structure for the i th training data $[x, d]_i$, the immediate left and right neighbors of $C^{t,(s^*)}$ are denoted as $C^{t,(s-1)}$ and $C^{t,(s+1)}$ respectively. Let $s_L = (s - 1)$ and $s_R = (s + 1)$ be the left and right indexes of neighbors respectively. The adaptation and refinement can be classified into four cases:

- a. If $C^{t,(s^*)}$ is the first and initial output-context structure (i.e., $s_L = s_R = NULL$), then the best matched output-context structure is updated without any constraint.
- b. If $C^{t,(s^*)}$ has no left neighbor (i.e., $s_L = NULL$), then the uncertainty constraint depends only on the right neighbor $\sigma^{t,(s+1)}$.
- c. If $C^{t,(s^*)}$ has no right neighbor (i.e., $s_R = NULL$), then the uncertainty constraint depends only on the left neighbor $\sigma^{t,(s-1)}$.
- d. If $C^{t,(s^*)}$ has both left and right neighbors, then the uncertainty constraint depends on both left neighbor $\sigma^{t,(s-1)}$ and right neighbor $\sigma^{t,(s+1)}$.

From Definition 1, we may obtain the left neighbor index s_L and the right neighbor index s_R with respect to the best matched output-context structure ($C^{t,(s^*)} < c^{t,(s^*)}, \sigma^{t,(s^*)} >$) as follows:

$$s_L = \begin{cases} NULL & \text{if } c^{t,(s^*)} == \min_{1 \leq s \leq S^t} c^{t,(s)} \\ \arg \min_{c^{t,(s)} < c^{t,(s^*)}} |c^{t,(s^*)} - c^{t,(s)}| & \text{else,} \\ & \text{for } 1 \leq s \leq S^t \end{cases}, \quad (15)$$

$$s_R = \begin{cases} NULL & \text{if } c^{t,(s^*)} == \max_{1 \leq s \leq S^t} c^{t,(s)} \\ \arg \min_{c^{t,(s)} > c^{t,(s^*)}} |c^{t,(s^*)} - c^{t,(s)}| & \text{else,} \\ & \text{for } 1 \leq s \leq S^t \end{cases}. \quad (16)$$

Therefore, the adaptation and refinement of the best matched output context structure ($C^{t,(s^*)}$) with the uncertainty control is expressed in Eq. (17) as follows:

$$\sigma^{t,(s^*)} = \begin{cases} \sigma_{avg}^{t,(s^*)} & \text{if } s_L = s_R = NULL \\ L(P_{\sigma}^R, \sigma_{avg}^{t,(s^*)}) & \text{if } s_L = NULL \\ L(P_{\sigma}^L, \sigma_{avg}^{t,(s^*)}) & \text{if } s_R = NULL \\ L(P_{\sigma}^{LR}, \sigma_{avg}^{t,(s^*)}) & \text{otherwise} \end{cases}, \quad (17)$$

where $\sigma_{avg}^{t,(s^*)}$ is the average width between the best matched output-context structure and the new training data $[x, d]_i$. Function $L(\cdot)$ defines a linguistic function to maintain the fuzzy model as a consistent rule-based system. Assuming that $[a_1, b_1, c_1] \in C^{t,(s^*)}$ and $[a_2, b_2, c_2] \in C^{new}$, where C^{new} is the new fuzzy level for training data $[x, d]_i$, the points $[a_1, b_1, c_1]$ and $[a_2, b_2, c_2]$ denote the left support, right support, and center, respectively. Therefore, the points are $a_1 = c^{t,(s^*)} - \alpha\sigma^{t,(s^*)}\sqrt{\ln 2}$, $b_1 = c^{t,(s^*)} + \alpha\sigma^{t,(s^*)}\sqrt{\ln 2}$, $c_1 = c^{t,(s^*)}$, and $a_2 = c^{new} - \alpha\sigma^{new}\sqrt{\ln 2}$, $b_2 = c^{new} + \alpha\sigma^{new}\sqrt{\ln 2}$, $c_2 = c^{new}$. Here, $\alpha > 0$ is a minimum membership threshold where universe of discourse (UoD) is significantly covered. A detailed description of the minimum membership threshold is presented in [15]. Therefore, the average width $\sigma_{avg}^{t,(s^*)}$ between the best matched output-context structure and the present data d_i can be written as Eq. (18):

$$\sigma_{avg}^{t,(s^*)} = \frac{1}{2\sqrt{\ln 2}} \{\max(b_1, b_2) - \min(a_1, a_2)\}. \quad (18)$$

Linguistic function $L(\cdot)$ controls the uncertainty of the output-context structures, where $L(P_{\sigma}, \sigma_{avg}^{t,(s^*)}) = \min(P_{\sigma}, \sigma_{avg}^{t,(s^*)})$. Pseudo-sigma P_{σ} is used as a parameter in the function $L(\cdot)$ to keep the semantic meaning within the output-context structures and within the input clusters. F_L and F_R are the left and right factors for P_{σ} respectively. These factors are dynamically updated for each set of training data, resulting in a dynamic constraint toward a balanced fuzzy model. Pseudo-sigma in Eq. (17) is calculated using Definition 1, which is expressed below as Eqs. (19) to (27).

(1) Dynamic constraints for P_{σ}^R : At the state of $s_L = NULL$, the dynamic constraint of Pseudo-sigma depends on the right neighbor, and thus the right factor (F_R).

$$F_R = \frac{1}{2} \{c^{t,(s+1)} + (c^{t,(s+1)} - \alpha \sigma^{t,(s+1)} \sqrt{\ln 2})\} \quad (19)$$

$$F_L = \min_{a_1 \in C^{t,(s^*)}, a_2 \in C^{new}}(a_1, a_2) \quad (20)$$

$$\therefore P_\sigma^R = \sqrt{-\frac{(F_L - F_R)^2}{\ln \alpha}} \quad (21)$$

(2) Dynamic constraints for P_σ^L : At the state of $s_R = \text{NULL}$, the dynamic constraint of Pseudo-sigma depends on the left neighbor, and thus the left factor (F_L).

$$F_L = \frac{1}{2} \{c^{t,(s-1)} + (c^{t,(s-1)} + \alpha \sigma^{t,(s-1)} \sqrt{\ln 2})\} \quad (22)$$

$$F_R = \max_{b_1 \in C^{t,(s^*)}, b_2 \in C^{new}}(b_1, b_2) \quad (23)$$

$$\therefore P_\sigma^L = \sqrt{-\frac{(F_L - F_R)^2}{\ln \alpha}} \quad (24)$$

(3) Dynamic constraints for P_σ^{LR} : The states $s_R \neq \text{NULL}$ and $s_L \neq \text{NULL}$ indicate that both left and right neighbors exist. Therefore, the dynamic constraints of Pseudo-sigma depend on both the left and right neighbors.

$$F_L = \frac{1}{2} \{c^{t,(s-1)} + (c^{t,(s-1)} + \alpha \sigma^{t,(s-1)} \sqrt{\ln 2})\} \quad (25)$$

$$F_R = \frac{1}{2} \{c^{t,(s+1)} + (c^{t,(s+1)} - \alpha \sigma^{t,(s+1)} \sqrt{\ln 2})\} \quad (26)$$

$$\therefore P_\sigma^{LR} = \sqrt{-\frac{(F_L - F_R)^2}{\ln \alpha}} \quad (27)$$

Figure 5 shows the dynamic constraints of Pseudo-sigma, which depicts a more semantic meaning and interpretability of the proposed EOCFS. These constraints also maintain a rational partition to obtain an effective rule base. The computation of the dynamic constraints for Pseudo-sigma (P_σ) is shown logically in Eqs. (19) to (27). Taking the condition $s_L = \text{NULL}$ into account, the dynamic constraint depends only on F_R and is computed as half of the center and left supports, as indicated in Eq. (19). Questions may arise regarding this computation which is shown in Theorem 1.

Theorem 1: Lateral positions may differ for F_L and F_R to adjust the adaptation in the EOCFS.

Proof: The EOCFS performs the adaptation in the width ($\sigma^{t,(s^*)}$) of the output-context structure rather than in the center ($c^{t,(s^*)}$), as defined in Eqs. (19) to (27) and Fig. 5. Therefore, considering the width of the output-context structure, and consequently the dynamic constraints, is necessary. If only the center is considered to compute the dynamic constraint, then the EOCFS will not be adaptive. Therefore, the lateral positions of F_L and F_R may vary, but the transparency of fuzzy systems should be considered. This finding ends the proof. \square

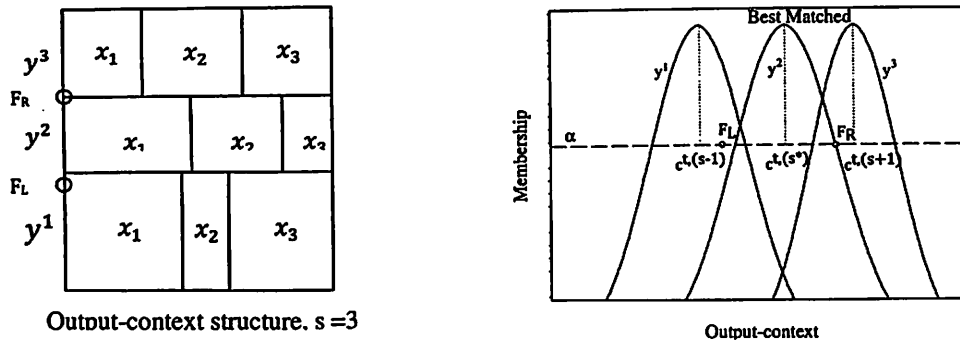


Fig. 5. Dynamic constraints in the output-context structure y^2 (left) and its membership function (right)

B. Input cluster to ensure interpretability

The EOCFS is an output-context fuzzy system, and the clustering of the input data depends on the output context.

Definition 2: $A_p^{t,(s)} < c_p^{t,(s)}, \sigma_p^{t,(s)} >$ is the antecedent part of the p th input variable ($p = [1, 2, \dots, n]$) and is associated with the s th output-context structure. $A_p^{t,(s)}$ is refined using the immediate left ($A_{p-1}^{t,(s-1)}$) and right ($A_{p+1}^{t,(s+1)}$) neighbors. $p_L = (p - 1)$ and $p_R = (p + 1)$ are the left and right indexes of the neighbors respectively. The refinement can be classified into four cases.

- a) If input variable $n = 1$ means that $A_p^{t,(s)}$ is the only input cluster (i.e., $p_L = p_R = NULL$), then $A_p^{t,(s)}$ is adapted without any constraint.
- b) If $A_p^{t,(s)}$ has no left neighbor (i.e., $p_L = NULL$), then the uncertainty constraint only depends on the right neighbor $\sigma_{p+1}^{t,(s)}$.

- c) If $A_p^{t,(s)}$ has no right neighbor (i.e., $p_R = NULL$), then the uncertainty constraint only depends on the left neighbor $\sigma_{p-1}^{t,(s)}$.
- d) If $A_p^{t,(s)}$ has both left and right neighbors, then the uncertainty constraint depends on both the left neighbor $\sigma_{p-1}^{t,(s)}$ and the right neighbor $\sigma_{p+1}^{t,(s)}$.

From Definition 2, we obtain the left neighbor index p_L and the right neighbor index p_R with respect to $A_p^{t,(s)}$, as shown in Eqs. (15) to (16). Therefore, the adaptation and refinement of the p th antecedent part $A_p^{t,(s)}$ using the uncertainty control are expressed in Eq. (28)

$$\sigma_p^{t,(s)}(x_p) = \begin{cases} \sigma_p^{t,(s)}(x_p)|_{x_p \in domain_s} & \text{if } p_L = p_R = NULL \\ L(P_\sigma^R, \sigma_p^{(s)}(x_p)) & \text{if } p_L = NULL \\ L(P_\sigma^L, \sigma_p^{(s)}(x_p)) & \text{if } p_R = NULL \\ L(P_\sigma^{LR}, \sigma_p^{(s)}(x_p)) & \text{otherwise} \end{cases} \quad (28)$$

The computations of Pseudo-sigma (P_σ), F_L , and F_R are the same as indicated in Eqs. (19) to (27), which were described in Section 3.2.3A. Again, the first condition $\sigma_p^{t,(s)}(x_p)|_{x_p \in domain_s}$ is computed from Eq. (18). Therefore, the factors are dynamically updated for each training data as well as each dynamic constraint to ensure interpretability. The adaptation and refinement in the p th input cluster $A_p^{t,(s)}(x_p)$ also show the same properties described in Theorem 1 and Fig. 5.

3.2.4 Algorithm for EOCFS

The proposed EOCFS is an output-context, self-reliant, and fully data-driven approach that is used for the automated formulation of fuzzy rule base. This evolving approach ensures an effective rule base (or model) that can adequately represent the application environment. At first, the EOCFS forms an effective rule base or model from the training data set, and then verifies the model using the test data set. Based on the explanations given in Sections 3.2.1 to 3.2.3, a detailed algorithm for creating an effective rule base from a training data set is illustrated in Fig. 6. The step-by-step process for the proposed EOCFS is referred to as Algorithm 2.

Algorithm 2: EOCFS

Step 1) Initial output-context structure from Algorithm 1 (Section 3.2.1): The EOCFS starts from this initial rule base.

Step 2) Evolving partition of the output domain and adding fuzzy rules: The width of the output domain from the previous stage $\sigma_{init}^t = \sigma_{init}^{t-1} - \sigma_{evolve}$ is updated.

Step 3a) For each training data $[\mathbf{x}, d]_i$, merging, adjustment, and refinement are then applied to the training output (d_i) and the best matched output-context structure ($C^{t,(s^*)}$) according to Eq. (12).

Step 3b) If d_i represents the distinct information, then a new output-context structure ($C^{t,(s+1)}$) is formed according to Eq. (11).

Step 4a) Associated with the output structure (Step 3a), merging, adjustment, and refinement are then applied to the input data \mathbf{x}_i and input cluster ($A_p^{t,(s)}$), as shown in Section 3.2.3B.

Step 4b) If d_i represents the distinct information (Step 3b) associated with the new output-context structure ($C^{t,(s+1)}$), then the new input domain ($A^{t,(s+1)}$) is created.

Step 5) Termination criteria: The evolving process is terminated if the Rational Partition Error (RPE) index exceeds the limit such that $RPE > 1$. Otherwise, Step 2 is repeated and the evolving process and addition of fuzzy rules are continued.

Step 6) Effective rule base: The termination criterion is found at the $(e + 1)$ evolving stage. Therefore, at evolving stage e , the effective rule number S^e represents the fuzzy model for obtaining reasonable accuracy.

Fig. 6. Algorithm for the proposed EOCFS

An important feature of the proposed EOCFS is that it partitions the output domain at a higher error region (distinct point). Therefore, the evolving process proceeds by reducing the model error, and is proven in Theorem 2.

Theorem 2: Let $f(\mathbf{x})$ be a linear/nonlinear continuous function and $o(\mathbf{x}^*)$ be the best matched output model. Then the accuracy at t th evolving stage is higher than the previous stage, therefore $E(t) < E(t - 1)$ and which is true for both the worst and best approximations.

Proof: Let S^t and S^{t-1} be the number of rules (or output-contexts) at evolving stages t and $(t - 1)$ respectively. Therefore, without loss of generality, $S^t \geq S^{t-1}$ (detailed in Fig. 1 and Algorithm 2). The new set of training data $[\mathbf{x}, d]_i$ and the linear/nonlinear continuous function $f(\mathbf{x})$ are then considered. We suppose that at t stage, the best matched output-context structure $C^{t,(k^*)}$ is found for the desired output (d_i). Let $\hat{o}(\mathbf{x}^*)$ be a constant function for the best matched output model $o(\mathbf{x}^*)$ that needs to be approximated using $f(\mathbf{x})$. At the evolving stage t ,

$$\hat{o}(\mathbf{x}^*)^t = c^{t,(k^*)}, \mathbf{x} \in [a_k, b_k], k = 1, 2, \dots, S^t \quad (29)$$

where $a_k = \min(\sigma^{t,(k^*)})$, $b_k = \max(\sigma^{t,(k^*)})$, and $\langle c^{t,(k^*)}, \sigma^{t,(k^*)} \rangle$ are the center and widths of $C^{t,(k^*)}$, respectively. To achieve the worst approximation accuracy, the worst value of parameter $c^{t,(k^*)} = a_k$ (or b_k) means that data are located at the border of the k th output partition. Similarly, if the best approximation accuracy is considered, then $c^{t,(k^*)} = (b_k + a_k)/2$. Therefore, the worst approximation errors in the k th interval and at the t th evolving stage can be given by

$$|f(x) - \hat{\delta}(x^*)^t| = a_k, k = 1, 2, \dots, s^t$$

If the approximation error is intended as $E(t)$ using (13), then the above equation implies that

$$E(t) = \sqrt{(f(x) - \hat{\delta}(x^*)^t)^2} \quad (30)$$

Given that $f(x)$ is a linear/nonlinear continuous function and $a_k = f(a_k) = b_k$, the whole output domain $[a, b]$ is expressed as

$$\sum_{x \in \text{domain}_{s,t}} [a, b] = \sum_{k=1}^N a_k \quad (31)$$

Therefore, from (30) and (31), the worst approximation error in the whole output domain $[a, b]$ is expressed as

$$\sum_{x \in \text{domain}_{s,t}} [a, b] = \sum_{s=1}^{s^t} \sqrt{\frac{1}{N} \sum_{k=1}^N a_k^2} = S^t * E(t) \quad (32)$$

Similarly, for the best approximation errors in the k th interval and at the t th evolving stage, the same approximation for the whole output domain $[a, b]$ is given by

$$|f(x) - \hat{\delta}(x^*)^t| = \frac{b_k + a_k}{2} \quad (33)$$

$$\text{or } \sum_{x \in \text{domain}_{s,t}} [a, b] = \sum_{s=1}^{s^t} \sqrt{\frac{1}{N} \sum_{k=1}^N \left(\frac{b_k + a_k}{2}\right)^2} \quad (34)$$

$$\text{or } \sum_{x \in \text{domain}_{s,t}} [a, b] = S^t * E(t) \quad (35)$$

Therefore, the approximation errors are the same for both the worst and best cases. Again, if the previous $(t - 1)$ evolving stage is considered, then the approximation error for the whole output domain $[a, b]$ is given by

$$\sum_{x \in \text{domain}_t} [a, b] = S^{t-1} * E(t-1) \quad (36)$$

Nevertheless, $S^t \geq S^{t-1}$ for t and $(t-1)$ evolving stages. Furthermore, if $SM > \beta$, then a prominent distinction is observed and the EOCFS forms a new rule (see Algorithm 2), which means that the proposed EOCFS adds a new rule to reach a lower error region while evolving proceeds. Therefore, these two propositions may be written in the form of the inequality $E(t) < E(t-1)$ while the number of rules is $S^t \geq S^{t-1}$. This implication is true for both the worst and best cases. This finding ends the proof. \square

Remark 3: In Algorithm 2, an evolving index, the RPE index, is used for the termination criteria in Step 5. In Theorem 2, as the evolving stage proceeds, training error $E(t)$ is expected to be reduced. However, in practice, two issues arise in the proposed EOCFS that cause the output structure partition to become unbalanced, namely, over-fitting and under-fitting. In the EOCFS, the basic idea is to tune the width of the output structure for each evolving stage according to Eq. (6). We begin from the initial evolving stage where a single rule exists and use the EOCFS to locate the data sample with the largest error. Afterward, evolving proceeds by partitioning the output domain, which means adding more fuzzy rules (detail in Section 3.2.2). This initial rule-based system is in an under-fitting state with a high approximation error. As the evolving stage proceeds, the EOCFS partitions the output structure, reduces the error (Theorem 2), and then adds more fuzzy rules. Evolving continues and will reach a condition wherein $E(t) \geq E(t-1)$, which then contradicts Theorem 2. We define this condition as the over-fitting state and analyze the reason behind this state. As the evolving stage proceeds, the width of the output structure becomes small and the EOCFS leads to very fine partitions in the output domain. At a specific evolving stage, the width of the output structure becomes extremely small, resulting in an unbalanced state, which cannot properly represent the data. Therefore, this unbalanced state leads to a fuzzy system with a large number of rules and causes over-fitting (i.e., the data fit is very close because of the small width of the output structure).

The strategy used in Theorem 2 is discussed in Remark 3. The RPE index is used to handle situations from the under-fitting to the over-fitting. RPE is a straightforward index used to recognize an unbalanced situation in the evolving process and to prevent the algorithm from further evolving.

$$\text{RPE}(t) = \frac{E(t)}{E(t-1)} \quad (37)$$

where $E(t)$ and $E(t - 1)$ are the functions [as in Eq. (13)] of the approximate training error at t and $(t - 1)$ evolving stages respectively. The characteristics of the RPE index are as follows:

- 1) According to Theorem 1 and Eq. (37), $\text{RPE} < 1$ shows a balanced partition in the output domain. The characteristics of the equation for RPE at t evolving stage are as follows:

$$\text{RPE}(t) = \begin{cases} \geq 0 & \text{if } E(t) < E(t - 1) \\ \geq 1 & \text{if } E(t) \geq E(t - 1) \end{cases} \quad (38)$$

The estimation of RPE is fully online, is not a predefined threshold, and is approximated from the current and previous evolving stages. During initialization, $\text{RPE}(1) = 0$ is implied as a base value.

- 2) If $\text{RPE}(t)$ holds the first criteria according to Eq. (38) and Theorem 2, then the evolving process continues. The evolving process partitions the output domain at a higher error region (i.e., distinct data at the center for a new output structure), therefore reducing the error, and the EOCFS proceeds to the rational partition to obtain optimum accuracy. At a specific evolving stage, the RPE index exceeds 1, holds the second criteria according to Eq. (38), and contradicts Theorem 2. Index $\text{RPE} \geq 1$ means that the continuation of the evolving process will cause an over-fitting in the model and should be terminated. A fuzzy system with an effective rule base is thus obtained. Global learning using the RPE index can find an effective rule-based fuzzy system with the lowest approximation error and has reasonable accuracy for testing data. The effective rule base is defined according to Definition 3.

Definition 3: If $\text{RPE} \geq 1$ is found at evolving stage $(e + 1)$, then the termination criteria are fulfilled for the EOCFS algorithm. Therefore, the effective rule base $R^{e,s}$ is found at evolving stage e . Hence, if j denotes the number of evolving stages and s denotes the number of output-context partitions, then

$$e \in R^{e,s}, \quad e = 1, 2, \dots, e, (e + 1), \dots, j \text{ and } E(e) = \min(E(1), E(2), \dots, E(e))$$

3.3 Data samples and analysis

In this paper, similarity threshold (β) represents the similarity between two consecutive fuzzy levels, and $\alpha > 0$ is the minimum membership value that strongly covers the UoD. In the following examples, the threshold $\beta = 0.8$ means that $1 - 0.8 = 0.2$. Hence, a distinct fuzzy level is formed when the similarity among consecutive fuzzy levels is less than 20%. Furthermore, $\alpha = 0.5$ is considered to strongly cover the UoD and to keep the average semantic meaning in the EOCFS model. The evolving constant σ_{evolve} is taken arbitrarily to ensure that the rule number slowly increases.

To evaluate the performance, the proposed EOCFS is compared with existing methods in terms of rule and error; the lower the error, the better the performance. Again, lowering the number of rules enhances model performance, thereby reducing model complexity [13]. This paper compares with the existing models regarding error analysis using Eqs. (13) and (42).

Example 1: A function approximation problem of one-dimensional data is used to show the performance. We consider 200 data points from the following single-input-single-output nonlinear function:

$$y = 0.6 \sin(\pi x) + 0.3 \sin(3\pi x) + 0.1 \sin(5\pi x) \quad (39)$$

where $x \in [-1,1]$. These 200 data points are then randomly divided from the universe of discourse x : 100 for training and 100 for testing.

Example 2: The function approximation problem involving two-dimensional data is used to show the performance. We consider 200 data points from the following two-input-single-output nonlinear function:

$$y = f(x_1, x_2) = 0.6 + 0.2x_1 + 4x_2 + 0.5x_1x_2 + 25 \sin(0.5x_1x_2) \quad (40)$$

where $x_1 \in [-4,6]$ and $x_2 \in [-2,4]$. These 200 data points are then randomly divided from the universe of discourse x_1 and x_2 : 100 for training and 100 for testing.

Example 3: Considering the following three-dimensional (3D) function approximation [20]:

$$f(x, y, z) = (1 + x^{0.5} + y^{-1} + z^{-1.5})^2 \quad (41)$$

with $x \in [1,6]$, $y \in [1,6]$, and $z \in [1,6]$, where 216 training data samples are generated with the step size 1 over this 3D Cartesian product. The test data samples are produced while x , y , and $z \in [1.5, 5.5]$, 125 test samples are taken. Average percentage error (APE) is computed and compared with those of the existing models:

$$\text{APE} = \frac{1}{N_{test}} \sum_{i=1}^{N_{test}} \frac{|O(x_i) - d_i|}{|d_i|} \times 100\% \quad (42)$$

where N_{test} is the number of test data samples, and $O(x_i)$ and d_i are the model and desired outputs respectively for the i th testing data.

Example 4: A well-known real-world data, Automobile Miles Per Gallon (MPG), is used [21] to evaluate the performance. The output is the fuel consumption of an automobile expressed in miles per gallon; seven input variables are used to distinguish the actual output. For evaluation, we randomly divide the data set into training (60%) and testing (40%) data sets from 316 observations.

Example 5: This data set deals with real estate in the Boston area [22]. It contains 13 input variables with 506 observations; the median value of the house is considered as an output variable. We randomly divide the data set into training (60%) and testing (40%) data sets for evaluation.

4. Results and discussion

Tables I and II show the evaluation of Example 1. In Table I, the RPE index shows the effectiveness of each partition, specifically, whether the partition is balanced or over-/under-fitting. The EOCFS evolves the output-context structures (or rules), and $\sigma_{evolve} = 0.005$ is taken arbitrarily to slowly increase the formation of the fuzzy rules. The termination point $\text{RPE} > 1$ is found at the evolving stage, where $\sigma_{init}^f = 0.05$. Therefore, an effective rule base is found at the evolving stage with a minimum approximation error. Hence, in Table I, the EOCFS finds the effective rule base for nonlinear function (39) when the number of fuzzy rules is eight. Testing data are then applied to the effective rule base. Table II shows that the proposed EOCFS is more accurate than ECSFS [17] and linguistic modeling [3]. Only eight effective rules are found, and the testing accuracy for this effective rule base is 0.087. Therefore, the evolving and localized

adaptation in the EOCFS is mathematically significant with respect to the existing models. In Table III, the EOCFS determines the number of effective rules for function (40) to be eight, which is a more accurate determination than those of the existing models.

TABLE I: EVOLVING PROCESS OF THE EOCFS TO OBTAIN AN EFFECTIVE RULE BASE FOR EXAMPLE 1

σ_{init}^t	Number of rules/neurons	Training RMSE	RPE (t)	Testing RMSE
0.080	1	0.4820	0	0.4796
0.075	5	0.2191	0.4545	0.2038
0.070	5	0.2189	0.9990	0.2037
0.065	5	0.2068	0.9447	0.1932
0.060	6	0.1263	0.6107	0.0966
0.055	8, effective	0.1240	0.9817	0.0875
0.050	10	0.1257	1.0137	0.0648
0.045	10	0.1254	0.9976	0.0633
0.040	12	0.1374	1.0956	0.0488

TABLE II: PERFORMANCE COMPARISON IN RMSE OF EXAMPLE 1

Model/Algorithm	Number of fuzzy rules	Training	Testing
Linguistic modeling [3]	6	0.180	0.189
	8	0.150	0.153
ECSFS [17]	4	0.130	0.122
	5	0.116	0.117
	6	0.108	0.117
Proposed EOCFS	6	0.126	0.096
	8, effective	0.124	0.087

TABLE III: PERFORMANCE COMPARISON IN RMSE OF EXAMPLE 2

Model/Algorithm	Number of fuzzy rules	Training	Testing
Linguistic modeling [3]	6	20.69	20.86
	8	15.66	16.08
	10	14.64	15.67
ECSFS [17]	6	16.93	17.77
	8	16.72	17.90
	10	14.46	16.05
	12	13.52	16.13
Proposed EOCFS	7	13.01	17.05
	8, effective	11.42	15.63

The proposed EOCFS performance is better than the Takagi-Sugeno type fuzzy rule in terms of parameter and number of fuzzy rules. In Table IV, the number of fuzzy rules and parameters in Example 3 are significantly reduced. APE is also more reasonable than the existing evolving models [10]–[11], [20]. Although the APE performance index of FLEXFIS [20] is lower than that of EOCFS, the proposed method is reasonably accurate and uses fewer rules and parameters. In Fig. 4, the interpretability issues are addressed in the EOCFS, where four distinct centers are realized automatically. Hence, four fuzzy rules are formed. Figure 7 shows the semantic significance of the proposed EOCFS given that the output-context structures are self-organized by the proposed algorithm. The knowledge embedded in the set of fuzzy rules is highly methodical, that is, four output-contexts (LOW, MED, HIGH, and VERY HIGH) are realized by the EOCFS. Therefore, Fig. 7 shows that the output contexts demonstrate the excellent self-adaptation and self-organizing abilities of the proposed EOCFS model.

TABLE IV: APE PERFORMANCE COMPARISON FOR EXAMPLE 3

Model/Algorithm	Rule type	Number of rules/ neurons	Parameter	Testing APE
GDFNN [10]	Takagi-Sugeno	10	64	1.540
SOFNN [11]	Takagi-Sugeno	9	60	1.124
FLEXFIS Var A [20]	Takagi-Sugeno	31	-	0.257
FLEXFIS Var A [20]	Takagi-Sugeno	43	-	0.209
Proposed EOCFS	Mamdani	1	1	0.479
		4, effective	4	0.479

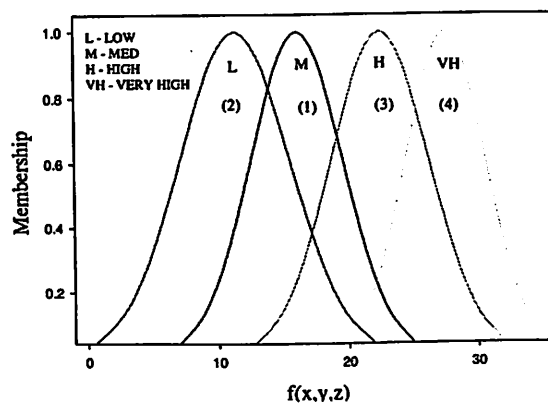


Fig. 7. Membership functions of the output-context structures in Example 3. Numbers 1, 2, 3, and 4 show the sequence of the self-organizing partitions.

Two real-world data sets are considered for evaluation as Examples 4 and 5. For the Automobile MPG data [21], the rational partition on the output domain is valid (i.e., $RPE < 1$) until the number of output-context partition is 11; further partition causes over-fitting. Hence, the fuzzy model cannot represent the data properly if the output-context structure exceeds 11. The EOCFS then establishes an effective rule base for the Automobile MPG data, where the effective number of fuzzy rules is 11. The proposed EOCFS is compared with the existing models in Table V. Again, the EOCFS is more accurate in establishing an effective rule base. The Boston area [22] data are also used to evaluate the EOCFS in Example 5. In Table VI, the results are compared with those of recently published existing methods. The proposed EOCFS is again shown to be more accurate in establishing an effective fuzzy rule base. The effective number of fuzzy rules used to represent the Boston data is 17.

TABLE V: PERFORMANCE COMPARISON IN RMSE OF EXAMPLE 4

Model/Algorithm	Fuzzy rules/ hidden neurons	Training	Testing
Linguistic modeling [3]	36	2.86	3.14
RBFNN_CFCM [25]	15	3.11	3.19
	9	3.07	3.41
RBFNN_FKNN [25]	9	0.21	2.77
Output-constrained [18]	8/12	2.98	2.57
	7/11	2.52	2.78
SSEM [19]	8	2.99	2.79
	16	2.93	2.82
	24	2.86	2.90
Proposed EOCFS	8	3.71	2.75
	11, effective	3.06	2.73

TABLE VI: PERFORMANCE COMPARISON IN RMSE OF EXAMPLE 5

Model/Algorithm	Number of fuzzy rules	Training	Testing
RBFNN	36	5.52	6.91
+ constriction free clustering [3]			
Linguistic modeling [3]	36	4.12	5.32
SSEM [19]	8	3.68	4.64
	16	3.43	4.67
	32	3.18	4.69
Proposed EOCFS	6	3.73	3.77
	12	2.84	3.48
	17, effective	2.76	4.11

5. Features of the proposed EOCFS

Current trends in evolving fuzzy rule-based systems and their limitations are discussed in Section 2. However, unlike existing methods, such as ECSFS [17], SSEM [19], output-constrained cluster approach [18], dynamically constrained GFS [15], SaFIN [6], SONFIN [9], and eFSM [7], the main features of the proposed EOCFS are as follows.

- 1) The proposed EOCFS follows an output context, is self-reliant, and is a fully data-driven approach used to automate the formulation of an effective fuzzy rule base. The proposed EOCFS automatically observes the prominent distinction point on the output domain by measuring the similarity of two consecutive fuzzy levels. Hence, the evolving process partitions the output domain at higher-error regions (distinct points) with the distinct data as the center of the new output-context structure. The evolving process always proceeds by reducing the model error. The proposed EOCFS proceeds to the rational partition until an effective rule base that ensures reasonable accuracy is established.
- 2) The evolving process is terminated depending on the RPE index. RPE estimation is fully online, does not follow a predefined threshold, and is approximated from the current and previous evolving stages. The EOCFS starts the evolving process in an under-fitting condition with a high approximation error; evolving then proceeds toward an over-fitting state by increasing the number of fuzzy rules. The RPE index recognizes the unbalance partition of the output domain during over-fitting and terminates the algorithm from further evolving.
- 3) The proposed EOCFS algorithm is also a novel self-adaptive fuzzy system that can be used for localized learning. Adaptation and refinement are performed in the existing output-context structures and the associated input clusters. A linguistic function $L(.)$ with an uncertainty controller is used to regulate and refine the output-context structures within the output domain. Similarly, adaptation and regulation are desirable for input clusters within the input domain. Therefore, the EOCFS shows more of the semantic meaning of its output-context structures and associated input clusters.
- 4) The EOCFS increases the number of rules by partitioning the output domain, thereby transforming the level of abstraction conceived by humans to achieve a reasonable accuracy. An *interpretability-accuracy tradeoff* is considered to establish an effective rule base that ensures

the linguistic fuzzy model in the proposed EOCFS model. Furthermore, localized learning with uncertainty control is continued such that the previous information in the system and new information from the training data can coexist in the system and provide a more semantic representation of the fuzzy model. This localized learning addresses the *stability-plasticity dilemma* of the proposed EOCFS.

6. Conclusion

EOCFS learning is based on the global learning of rule-consequent parameters and further localized learning in both consequent and antecedent parameters respectively. Hence, the proposed EOCFS learning addresses both *interpretability-accuracy tradeoff* and *stability-plasticity*. The evolving process is also realized in the output domain, whereas the self-adaptive process is realized in both the output domain and its associated input domain. As a result, more semantic rules are formed in the EOCFS. Using the dynamic uncertainty controller $L(.)$ allows the proposed EOCFS model to establish a compact, consistent, and effective rule base.

The EOCFS incorporates new information into past knowledge while the uncertainty controller prunes highly noisy data. However, when dealing with highly noisy data, the EOCFS evolves the system until rational partitions are found. Therefore, the effective rule base is realized at a larger value. Hence, when learning from noise data, the EOCFS achieves reasonable accuracy with the effective rule-base having a larger value.

The features of the proposed EOCFS are described in this paper and compared with those of existing models. The proposed EOCFS is evaluated using both synthetic and real-world data. These evaluation results show that the proposed EOCFS has a more accurate effective rule base than the existing models, and is therefore effective.

Acknowledgements

This work was supported by the Ministry of Higher Education Malaysia under the Commonwealth Scholarship and Fellowship Plan, and in part by the Ministry of Science, Technology and Innovation Malaysia through the Sciencefund Grant titled "Development of Computational Intelligent Infertility Detection System Based on Sperm Motility Analysis."

References

- [1] L. A. Zadeh, Some reflections on soft computing, granular computing and their roles in the conception, design and utilization of information/ intelligent systems, *Soft Comput.* 2(1998) 23–25.
- [2] A. Bargiela, W. Pedrycz, *Granular Computing: An Introduction*, Kluwer Academic Publishers, Boston, London, 2003.
- [3] W. Pedrycz, K.-C. Kwak, Linguistic models as a framework of user-centric system modeling, *IEEE Trans. Syst. Man Cybern. A, Syst., Humans* 36(2006) 727–745.
- [4] W. Pedrycz, Constrained fuzzy c-mean algorithm, *Pattern Recognit. Lett.* 17(1996) 625–631.
- [5] W. Pedrycz, Context fuzzy clustering in the design of radial basis function neural network, *IEEE Trans. Neural Netw.* 9(1998) 601–612.
- [6] S. W. Tung, C. Quek, C. Guan, SaFIN: A self-adaptive fuzzy inference network, *IEEE Trans. Neural Netw.* 22(2011) 1928–1940.
- [7] W. L. Tung C. Quek, eFSM—A novel online neural-fuzzy semantic memory model, *IEEE Trans. Neural Netw.* 21(2012) 136–157.
- [8] J.-S. Wang, C. S. G. Lee, Self-adaptive neuro-fuzzy inference systems for classification applications, *IEEE Trans. Fuzzy Syst.* 10(2002) 790–802.
- [9] C.-F. Juang, C.-T. Lin, An on-line self-constructing neural fuzzy inference network and its applications, *IEEE Trans. Fuzzy Syst.* 6(1998) 12–32.
- [10] S. Wu, M. Er, Y. Gao, A fast approach for automatic generation of fuzzy rules by generalized dynamic fuzzy neural networks, *IEEE Trans. Fuzzy Syst.* 9(2001) 578–594.
- [11] G. Leng, T. McGinnity, G. Prasad, An approach for on-line extraction of fuzzy rules using a self-organising fuzzy neural network, *Fuzzy Sets Syst.* 150(2005) 211–243.
- [12] G.-D. Wu, Z.-W. Zhu, P.-H. Huang, A TS-Type maximizing-discriminability-based recurrent fuzzy network for classification problems, *IEEE Trans. Fuzzy Syst.* 19(2011) 339–352.

- [13] E. G. Mansoori, FRBC: A fuzzy rule-based clustering algorithm, *IEEE Trans. Fuzzy Syst.* 19(2011) 960–971.
- [14] R. Mikut, J. Jäkel, L. Gröll, Interpretability issues in data-based learning of fuzzy systems, *Fuzzy Sets Syst.* 150(2005) 179–197.
- [15] P. Pulkkinen, H. Koivisto, A dynamically constrained multiobjective genetic fuzzy system for regression problems, *IEEE Trans. Fuzzy Syst.* 18(2010) 161–177.
- [16] Y. Jin, W. von Seelen, B. Sendhoff, On generating FC^3 fuzzy rule systems from data using evolution strategies, *IEEE Trans. Syst. Man Cybern. B, Cybern.* 29(1999) 829–845.
- [17] D. Wang, X.-J. Zeng, J. A. Keane, An evolving-construction scheme for fuzzy system, *IEEE Trans. Fuzzy Syst.* 18(2000) 4(2010) 755–770.
- [18] D. Wang, X.-J. Zeng, J. A. Keane, An output-constrained clustering approach for the identification of fuzzy systems and fuzzy granular systems, *IEEE Trans. Fuzzy Syst.* 19(2011) 1127–1140.
- [19] D. Wang, X.-J. Zeng, J. A. Keane, A simplified structure evolving method for Mamdani fuzzy system identification and its application to high-dimensional problems, *Inf. Sci.* 200(2013) 110–123.
- [20] Edwin David Lughofer, FLEXFIS: A robust incremental learning approach for evolving takagi–sugeno fuzzy models, *IEEE Trans. Fuzzy Syst.* 16(2008) 1393–1410.
- [21] UC Irvine Machine Learning Repository, Available from: <http://archive.ics.uci.edu/ml/datasets/Auto+MPG>.
- [22] UC Irvine Machine Learning Repository, Available from: <http://archive.ics.uci.edu/ml/datasets/Housing>.
- [23] B. Kosko, Optimal fuzzy rules cover extrema, *Int. J. Intell. Syst.* 10(1995) 249–255.
- [24] Y. S. Ding, H. Ying, S. H. Shao, Necessary conditions on minimal system configuration for general MISO mamdani fuzzy systems as universal approximators, *IEEE Trans. Syst. Man Cybern. B, Cybern.* 30(2000) 857–864.

[25] S.-B. Roha, T.-C. Ahna, W. Pedrycz, The design methodology of radial basis function neural networks, based on fuzzy K-nearest neighbors approach, *Fuzzy Sets Syst.* 161(2010) 1803–1822.

[26] C. Mencar, C. Castiello, R. Cannone, A. M. Fanelli, Design of fuzzy rule-based classifiers with semantic cointension, *Inf. Sci.* 181(2011) 4361–4377.



Md. Manjur Ahmed received his B.Sc. degree in Computer Science and Engineering from Khulna University of Engineering and Technology in Bangladesh in 2009. He is currently pursuing a Ph.D. in Computational Intelligence and is part of the Imaging and Intelligent Systems Research Team of the School of Electrical and Electronic Engineering, Universiti Sains Malaysia. His current research interests include algorithms, data structure, fuzzy systems, neural networks, and pattern recognition.



Nor Ashidi Mat Isa received his B.Eng. in Electrical and Electronic Engineering (with First Class honors) and his Ph.D. degree in Electronic Engineering (majoring in Artificial Neural Networks) from Universiti Sains Malaysia (USM) in 1999 and 2003 respectively. He is currently an Associate Professor with the School of Electrical and Electronics Engineering, USM. He and his research team (Imaging and Intelligent System Research Team) have published papers nationally and internationally. Their contributions can be found in numerous international and national journals, chapters in books, and in international and national proceedings. His current research interests include intelligent systems, image processing, neural networks, biomedical engineering (i.e., intelligent diagnostic systems), and algorithms.

Recursively constructs output-context fuzzy system for infrared thermography based non-destructive characterization of electrical hotspots

Md. Manjur Ahmed^a, A.S.N. Huda^b, Nor Ashidi Mat Isa^a

^aImaging and Intelligent Systems Research Team, School of Electrical and Electronic Engineering, Engineering Campus, Universiti Sains Malaysia, 14300 Nibong Tebal, Penang, Malaysia

^bPower Electronics & Renewable Energy Laboratory (PEARL), Department of Electrical Engineering, University of Malaya, 59300 Kuala Lumpur, Malaysia

E-mail addresses: manjur_39@yahoo.com (M. M. Ahmed), nhudae0450@gmail.com (A.S.N. Huda), ashidi@eng.usm.my (N.A. Mat Isa)

Abstract

Nowadays infrared thermography technology is being utilized in various applications including faults diagnosis in electrical equipment. The thermal abnormalities are diagnosed through identification and classification of hotspot condition of electrical components. In this article, a new recursively constructs output-context fuzzy system is proposed for characterising the condition of hotspot. An infrared camera is initially utilized to capture the thermal images of components with hotspot and intensity features are extracted from each hotspot. Finally, recursively construct fuzzy system (RCFS) is applied which automatically realizes and formulates the conditions of the thermal abnormalities. According to the priority level, the conditions of hotspots are categorized as normal, warning and critical. The proposed RCFS realizes the prominent distinctions in the output domain using self-organizing method. The termination of the recursive algorithm finds an effective rule base to achieve the well representation of the datasets. The evaluation of the thermographic diagnostic shows the significance of the proposed intelligent system since its automatically identifies the limits of the conditions of hotspots which are rationally acceptable and therefore, establish an intelligent

defect analyzing system. Again, the proposed system realizes less number of fuzzy rules with reasonable accuracy and shows the significance of the system.

Keywords- Electrical hotspots; condition monitoring; infrared thermography; recursive construction; fuzzy system

1. Introduction

Heat energy is an important factor in electrical equipment for increasing operational reliability. Electrical current passes through a resistive component and generates heat. Thermal energy generated from an electrical component is directly in proportional to the square of the current passing through it and resistance (I^2R Loss). Therefore, an increase of resistance results in an increased in heat. Over the time, condition of the electrical components will begin to deteriorate due to various reasons. The reasons include poor or dirty connections, overloading, insulation problems, load imbalances, corrosion and wiring mistakes etc [1]. As components deteriorate, their resistance increases and also generates more heat. The increase in heat energy can cause electrical equipment to fail and also fire may break out. The faults due to the abnormal heating effect are largely preventable if heat is detected at the early stage by an effective screening and then adequately necessary steps are taken.

Infrared thermography (IRT) senses the heat produced in the electrical components. The thermal profiles of different electrical components and connectors are captured by using IR camera. Thermal profile called thermogram consists of heat picture and scale of temperature value of equipment. The different colours of temperature scale represent the different temperature zones in the equipment. Then thermographers analyze the thermal images, classify the condition of hotspots according to the priority level of repairing and comment for further actions. Finally, the equipment are followed, tested and repaired on priority basis [2].

To date, both manual and automatic feature extraction methods are employed for intelligent classification of thermal condition of equipment based on thermography. Almeida et al. proposed an intelligent faults diagnosis system based on thermography for lighting arrestors using two kinds of variables as inputs of neuro-fuzzy network. Thermographic and identification variables were used to

classify faults and the results obtained about 90% of accuracy [3]. In one study, RGB colour scale data and temperature data were used as input features of artificial neural network to detect internal faults [4]. The experiment obtained 99.38% of testing accuracy. Smedberg and Wretman proposed an intelligent classification system based on ANN for diagnosing three-phase fuses and different forms of connections problems [5,6]. The four input parameters of ANN were used which are absolute max temperature, relative max temperature, mean temperature difference compared to the other regions of the image and histogram distance to the other regions of the image. The test error rate using all four feature parameters as input of ANN was 9.5% and the error rate using only the histogram distance as input was 31.2%. The set of data was 74 infrared images. One of the disadvantages of working with a small data set is that the reliability of the results can sometimes be questionable.

On the other hand, several researches have been done based on automatic feature extraction method and intelligent classification system. Nazmul Huda et al. proposed a semi-automatic system for electrical thermography which used 15 statistical features and multilayered perceptron (MLP) network to classify thermal conditions as normal, warning and critical and achieved 78.5% of accuracy [7]. An intelligent system to detect faults of electrical equipment in ground substations based on support vector machine (SVM) as a classifier and 22 image features of Zernike moments was proposed in a research. The diagnosis was able to obtain 68.42% of accuracy [8]. In one study, 10 statistical features and MLP network were employed to differentiate between normal and defected condition. The system achieved 82.40% [9]. In another study, six statistical features and MLP network were used to find the overheated component and it reached to 79.4% of accuracy [10].

In the current study, in order to develop more robust and reliable system, the use of a new recursively construct fuzzy system (RCFS) is introduced to classify the of conditions of hotspots in components. The proposed system employs automatic feature extraction and proposed intelligent classification system. The greyscale images of infrared thermal images of components are segmented using manual thresholding technique. Then automatic feature extraction (AFE) system automatically extracts six intensity features (i.e. maximum, minimum, mean, median, standard deviation and variance). The RCFS is a self-automated system and automatically realizes the conditions of

components which formulate the abnormalities of electrical equipment into classes like normal, warning and critical. In this work, the RCFS is an output-context fuzzy system which recursively form fuzzy rule base by realizing the prominent distinction on output domain. The termination criterion for recursive algorithm is not threshold as like [11], it realizes from previous and present stage of evolving. Hence, the computational model of the fuzzy system is automatically designed by the RCFS rather than human experts.

Neural fuzzy systems are hybrid systems that capitalize on the functionalities of fuzzy systems and neural networks [12]. The black-box nature of a neural network can be resolved by injecting the interpretability of a fuzzy system into the connectionist structure [12-13]. Furthermore, introducing the learning powers of a neural network into a fuzzy system enables the system to automatically refine its parameters [13]. Output-constrained cluster approach [14] and SC [15] were proposed to consider the output domain for partitioning the input data. In the output-constrained cluster approach [14], first, the output space is roughly partitioned by fuzzy c-means, and then data within each output constraint are further refined based on "separability," which refers to the connectivity of the inputs. Prior knowledge for rough clustering in the output space makes a fuzzy system unintelligent. The results in [14] and [15] are highly subjective and uncertain because prior knowledge for designing the fuzzy system comes from users with limited competence. The output domain is evenly partitioned as in Pedrycz's method [16] and thus ignores the local distribution of the input data. An evenly partitioned output domain may also cause under-fitting or over-fitting, thus leading to an inaccurate performance. The aforementioned limitations of the existing models are considered in the RCFS. The RCFS is a self-organizing process and evolving proceeds by considering both input and output space. Evolving process continue until the termination fulfills the criteria and the RCFS realizes an effective rule base.

The remainder of this paper is organized as follows. Section 2 describes the necessity of IRT for fault diagnosis. Section 3 discusses the methodology of thermographic diagnosis. Section 4 covers automatic feature extraction technique. Section 5 elaborates the RCFS and its algorithms. Section 6 evaluates the expert system for thermographic diagnosis of electrical components. Section 7 summarizes the paper.

2. Necessity of IRT for fault diagnosis

The infrared camera is a device to display surface temperature of an object by detecting the infrared energy radiated from the surface of an object. IRT technique is an early internal and external faults diagnosis system of electrical components and provides various advantages over the conventional thermal condition and faults diagnosis tools [17]. Some of the advantages of IRT diagnostic system are described as follows:

a) Preventive /Predictive Maintenance

Typically for maintaining electrical equipment, two types of approaches, either run-to-failure or preventive maintenance are used. The run-to-failure approach is simple and straightforward, where there is no outflow of money on maintenance purpose before equipment eventual failure. The approach waits for equipment failure before any maintenance action is taken. Therefore, the method is more expensive. Thermography based diagnosis system allows preventive/predictive maintenance for early prevention of equipment failure without interrupting running operation and thus saves money. According to the historical data in USA [18], effective use of the preventive/predictive maintenance will reduce about 33% to 50% of the maintenance cost that are wasted by the most manufacturing and production plants.

i) Preventive Maintenance (PM)

Maintaining electrical equipment according to the statistical or historical information of operating capacity, failure history, mean-time-to-failure (MTTF) instead of tracking equipment performance is referred as preventive maintenance. Preventive maintenance program schedules the repairing and rebuilding activities for electrical equipment. Suppose, an electrical equipment normally runs 10 months before needed any repair. Using preventive technique, the equipment will be removed from service and rebuilt after 10 months of operation. In this example, the schedule of repairing of equipment is predefined which is determined by maintenance personnel. Now, if there is no need to rebuild after 10 months. Then, there would be wasted of labor and material. Again, if the equipment fails before 10 months, then it is necessary to fix problems after failure which is usually more

expensive than schedule maintenance. This is a just random decision, not supported by any kind of realistic data.

ii) Predictive Maintenance (PdM)

Maintenance of electrical equipment based on the direct monitoring of the actual operating condition, regularly collecting measurements, efficiency, heat distribution and other indicators instead of depending on statistical or historical data is called predictive maintenance. The program schedules all maintenance activities based on the factual data and if necessary, repairs before the equipment failure [19]. Early detection of imminent equipment failures and preventing them can significantly reduce downtime, costs for maintenance and maximize uptime.

c) Fire Prevention

According to the reports of Fire and Rescue Department of Malaysia about the causes of fire in buildings [20], about 2,317 fire related incidents have occurred between January 2012 and June 2012, making the average number of incidents around 387 a month. In this report, 1049 fires were caused by electrical problems. This was about 46% of the total causes of fire in building and mainly involved the electrical wiring problems (809 cases) and failure of electrical equipment (240 cases). Failure of electrical distribution equipment can potentially produce an ignition and fire. One of the causes of ignition is excessive ohmic heating in electrical distribution. The causes of excessive ohmic heating can be subdivided into gross overloads, excessive thermal insulation, stray currents, ground faults, overvoltage and poor connections [21]. These conditions can occur especially for old buildings with outdated electrical wiring that is deteriorating, inappropriately amended or insufficient for the electrical load. However, the new constructions are also not excluded from these conditions [22, 23].

d) Reduction of Energy Loss

Frequent monitoring of thermal condition of electrical equipment is also necessary for reducing the heat loss which occurs by elevated surface temperatures. The thermal insulation survey of a 460 MW thermal power station in India reveals that about 0.426768 million KJ/hour heat losses occurs due to bare surfaces, inadequate/damaged insulation or open cladding condition in all four units. This is

equivalent to a coal loss of about 1847 MT per annum. Further analysis shows that if these faulty insulated areas are attended by monitoring their thermal condition, there would be financial saving of around 59,052 USD per annum giving a simple payback period of about one month [24].

e) Reduced Maintenance Cost

Identifying fault quickly before critical condition, scheduling the follow-up inspection, repair and diagnosis the fault within appropriate period is the most efficient and cost effective way to increase the reliability of system. By thermographic inspection, it is easy to identify the potential problems, quantify potential energy savings, schedule interventions and set priorities for preventive and predictive maintenance or the need for immediate service to minimize the risk of failure and maintenance cost.

f) Avoiding Unnecessary Repair

Thermographic inspection can see the actual area of defect in equipment. Thus, it reduces disassembling, rebuilding, repairing or unnecessary replacement of good components which is pointless, expensive, time consuming issue. By applying infrared thermographic inspection, the extent of the defect or failure could be identified and this enables repair to be undertaken quickly. Thus maintenance costs are reduced and the revenue is increased.

g) Increased Production & Safety

The diagnosis system allows early avoidance of equipment failure which significantly decreases unscheduled outage and equipment damage. The system can diagnose the faults without interruption or shutdown of the service which results in an increase of production. Additionally the failure of electrical components could be catastrophic, injuring or even killing employees, maintenance personnel or the public.

h) Increased Life Time

The power rating of the equipment indicates the amount of energy the equipment can conduct without being damaged. Before the equipment failure, it is operated at excess power level which resists the electricity flow and generates heat. Therefore, the equipment is overheated and operating

efficiency is decreased. Utilization of thermography can increase the life time and efficiency of equipment by pinpointing the heat earlier.

3. Thermographic diagnosis methodology

3.1 Image capture

An infrared camera is used to capture the thermal image of the targeted electrical equipment. In this work, a total of 253 hotspots from 139 infrared images were captured from the main switch boards (MSB). The Fluke Ti25 thermal camera with fusion technology is used to capture the images. Table 1 shows the specifications of Fluke Ti25. For capturing the image, the thermal imager orientation is directly facing to the target equipment in order to get an accurate measurement. The distance between the target equipment and the thermal imager was in the range of 0.5-1.0 m. The emissivity value was set at 0.95 as recommended for most of electrical equipment thermography. It was noted that the ambient temperature around the equipment is between 30-33°C during the inspection.

Table 1. Fluke Ti24 camera specifications (Source Fluke Ti25 system, 2010)

Temperature range	-20°C to +350°C	Infrared lens type	20 mm F=0.8 lens
Accuracy	±2°C or 2%	Image frequency	9 Hz refresh rate
Visual camera	640 × 480 resolution	Focus	Manual
Detector type	160 × 120 focal plane array (FPA)	Minimum focus distance	Thermal lens: 15 cm (6 in); Visible (light) visual lens: 46 cm (18 in)
Field of view	23° × 17°	Spatial resolution	2.5 mRad
Spectral range	7.5 to 14 μm	Thermal sensitivity	≤0.1°C at 30°C (100mK)

3.2 Typical condition monitoring using IR technology

Manual calculation of component condition is based on comparative temperature analysis between hot and reference spots. This technology is simply called qualitative ΔT factor analysis. After capturing thermal images, the hotspot and reference area are identified visually by analyzing color map. The hotspot supports maximum temperature of faulty component and reference is the minimum temperature of same type, load or same repeated component of the equipment. Then, the difference between the hotspot and reference spot temperature is determined as ΔT factor which is used as the

decision making parameter about the condition of overheated component. ΔT factor can be directly interpreted from RGB data. Several standards for measuring ΔT are found such as InterNational Electrical Testing Association (NETA) [25] and American Society for Testing & Materials (ASTM) [26] etc. The technique is widely used for electrical thermography due to its simplicity and minor influencing by the emissivity. The main drawback of this technique is that it does not work, if in a three phase system, all phases show overheating at the same time which is very infrequent condition in electrical system. Fig. 1 shows the flowchart of IRT inspection of three phase electrical system.

Based on our survey of 253 detect regions 37 hotspots are $\Delta T(^{\circ}\text{C}) \geq 15$, 63 spots are $5 < \Delta T(^{\circ}\text{C}) < 15$ and 153 spots are $\Delta T(^{\circ}\text{C}) \leq 5$. All these conditions of components (i.e. $\Delta T(^{\circ}\text{C})$) were evaluated manually using infrared image analysis software. Some examples of the conditions are illustrated in Fig. 2. Whereas, brighter (more red) regions show the defect compared to the same repeated less bright region.

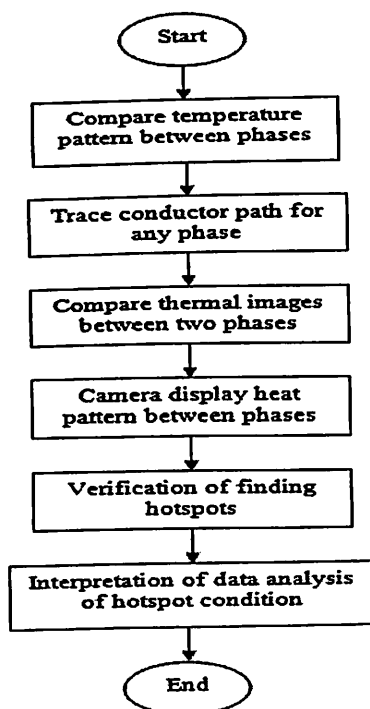


Fig. 1. Flowchart of the infrared thermographic inspection [27]

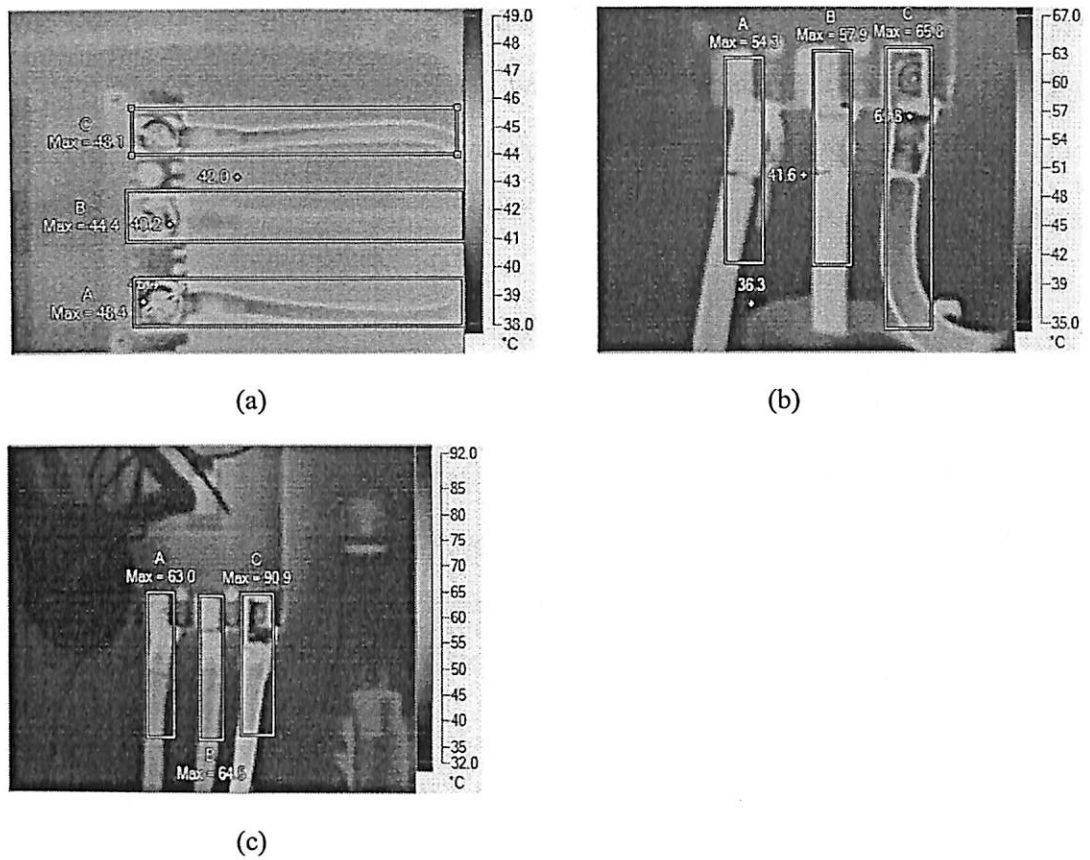


Fig. 2. (a) Hotspots in phase A & C, (b) hotspot (Phase C) seems as warning condition compared with (a), and (c) hotspot (Phase C) appears as critical condition compared with (a) and (b)

4. Hotspot detection and feature extraction

4.1. Hotspot detection

In this section, daime software version 1.3.1 which is a digital image analysis tool was used to detect defects in equipment and to extract quantitative features from defect image [28]. Segmentation is done to find the hotspots of electrical components. Daime can segment the hot region based on the thresholding value. In this study, custom thresholding technique (i.e. manually thresholding) value is set to generate an image of defect. The image of defect was produced by setting the threshold value, T manually. The value of T at which the original grey scale image shows the hotspot clearly, was selected as the threshold value for finding defect of image. After that the image is segmented according to the thresholded image.

In the present work, two cases were considered for selecting the desired hotspots. First, the highest pixel intensity objects and secondly, the object having area of equal to or greater than the half of the maximum area object. As the maximum temperature region of an image carries maximum intensity value so the component of maximum area is generally the highest pixel intensity region and in some cases there is more than one hotspot in the selected image. Some of the fake hotspots are generated due to the reflection from the sun. Using the object editor the hot defects are manually selected while the spurious defect and back ground were removed. Fig. 3 shows the different steps of defect detection in electrical equipment.

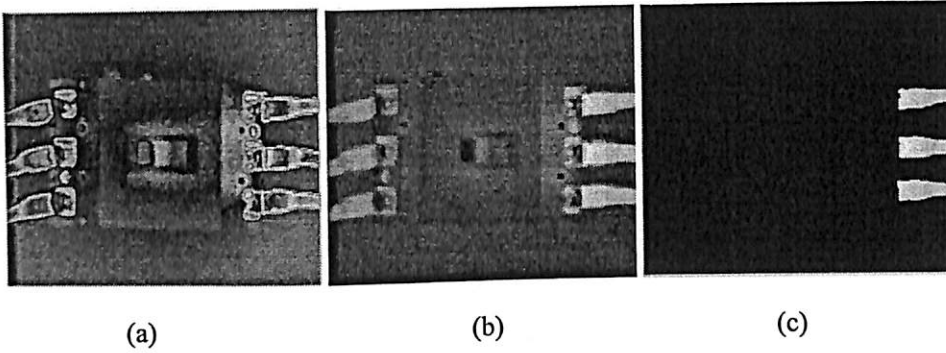


Fig. 3. Typical load imbalance problem (a) Thermal image (b) Greyscale image (c) Segmented image (T=163)

4.2. Automatic feature extraction

A total of 6 intensity features were computed using the pixel intensity values of connected components of the image. The features are maximum intensity, minimum intensity, average intensity, median intensity, standard deviation and variance of intensity values. The extracted features are defined as follows

$$\text{Maximum intensity} = \max \sum_{q=0}^{I-1} qp(q) \quad (1)$$

$$\text{Minimum intensity} = \min \sum_{q=0}^{I-1} qp(q) \quad (2)$$

$$\text{Mean intensity, } \mu = \frac{1}{\sum q} \sum_{q=0}^{I-1} qp(q) \quad (3)$$

$$\text{Variance, } \sigma^2 = \frac{1}{\sum q} \sum_{q=0}^{l-1} (q-\mu)^2 p(q) \quad (4)$$

$$\text{Standard deviation} = \sqrt{\frac{1}{\sum q} \sum_{q=0}^{l-1} (q-\mu)^2 p(q)} \quad (5)$$

where q is the number of distinct grey level in the object image. $p(q)$ is the histogram of the pixel intensity of object and l is the possible intensity level of the object. Mean intensity is the average pixels value which is determines the brightness or darkness of the defined object image. Maximum and minimum intensity define the maximum and minimum pixel intensity value of the object image respectively. Variance determines the dispersion of grey level pixels from the mean and standard deviation of pixel intensities is as like as variance, but different in value. If the intensity values are arranged in ascending order, then middle value is defined as median intensity value.

5. Characterisation of hotspots: Recursively constructs output-context fuzzy approach

5.1 Preliminaries: distinct points and effective rule-base

The main feature of the RCFS is to realize the prominent distinction points in the output domain and construct the effective rule-base. Fig. 4 depicts the realization of the prominent distinct points in the output domain. Authors [11], [29] and [30] previously described about selection of splitting points to reduce the approximation errors. However, unlike ECSFS [11], where LSM algorithm is used to select splitting points, the RCFS is evolved and self-determined the distinct output-context in order to obtain an effective rule-base.

Consider a modeling problem with n input variables and N data samples. Let us assume an input vector of i th training data $[x, d]_i$ where input vector x_i and corresponding output d_i and $x = (x_1, x_2, \dots, x_p, \dots, x_n)$ belongs to an output-context (s) which encodes an IF-THEN Mamdani-type fuzzy rule at t evolving stage as follows

$$R^{t,s}: \text{IF } x_1 \text{ is } A_1^{t,(s)}, \text{ and } x_2 \text{ is } A_2^{t,(s)}, \dots \dots \dots, \text{ and } x_p \text{ is } A_p^{t,(s)} \text{ THEN } y \text{ is } C^{t,(s)} \quad (6)$$

where $C^{t,(s)}$ is the s th consequent part associated with the s th output-context and $A_p^{t,(s)}$ is the s th antecedent part associated with the p th input variable. A Gaussian membership function is described for $C^{t,(s)}$ and its corresponding $A_p^{t,(s)}$.

$$\mu = e^{-((x-c)^2/\sigma)} \quad (7)$$

$$\text{and } \sigma = \sqrt{-\frac{(a_k - b_k)^2}{\ln \alpha}} \quad (8)$$

where, c and σ are the center and width of the membership function. a_k (or b_k) means that data are located at the border of the k th output context. $\alpha > 0$ is a minimum membership value [22] and is also defined as a distinguishability factor that keeps the semantic value for the output-contexts.

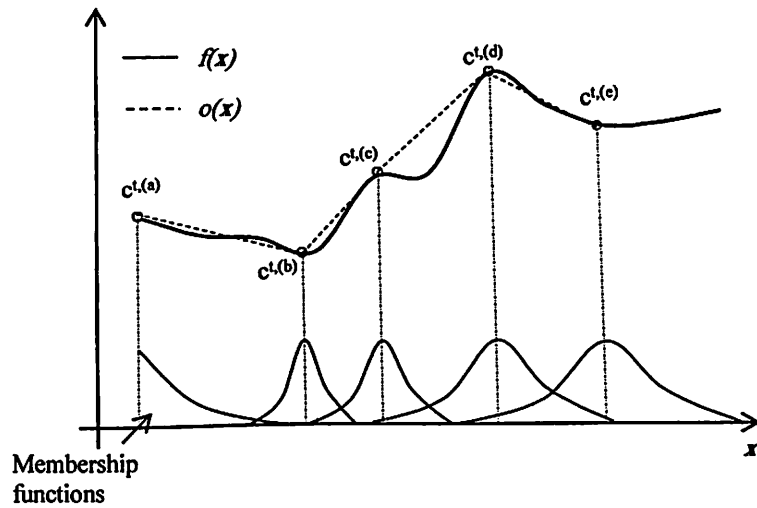


Fig. 4. Realization of the prominent distinct points by the RCFS. $c^{t,(a)}$, $c^{t,(b)}$, $c^{t,(c)}$, $c^{t,(d)}$ and $c^{t,(e)}$ depicts the distinct centers at t evolving stage.

Flowchart of the recursive procedure to construct the output-context fuzzy system depicts in Fig. 6. The RCFS employs the Mamdani-type fuzzy system and starts with an initial domain of whole output space. Then further partitions of the output domain to realize the prominent distinction point(s) and evolves; also adapts its corresponding input clusters. Each output-context

and associated inputs are formed a structure and called output-context structure. Previous knowledge in the system and new knowledge from the training data are incorporated in the system and provide more accurate representation of the fuzzy model. Recursive partitioning and evolving processes are continued until a termination criterion to get an effective rule-base is fulfilled.

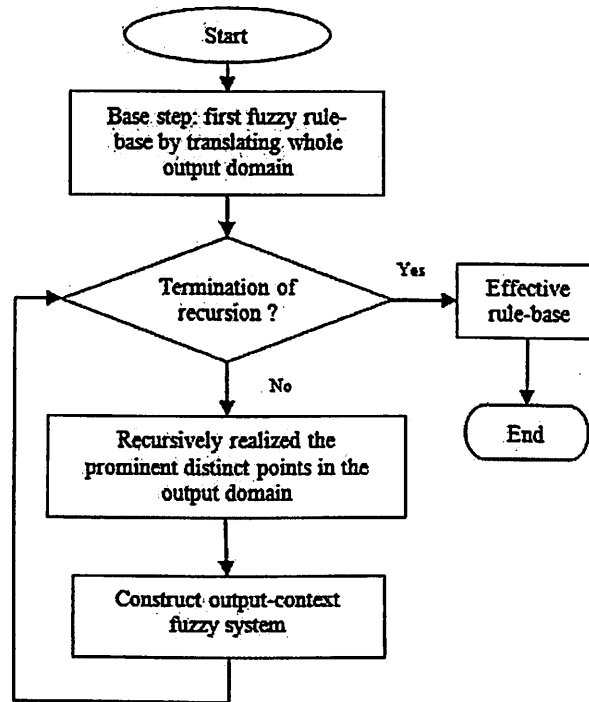


Fig. 6. Flowchart of the RCFS

The output model of s th domain of the RCFS is defined for the s th rule (or s th output-context) based on the Mamdani type fuzzy systems and it uses the center of averaging method [13-14, 32].

$$O = \frac{A_{1,2,\dots,p,\dots,n}^{t,(s)}(x) \times c^{t,(s)} |_{x \in \text{domain } s}}{A_{1,2,\dots,p,\dots,n}^{t,(s)}(x) |_{x \in \text{domain } s}} \quad (9)$$

where $c^{t,(s)}$ is the center of the s th consequent part of the output-context ($C^{t,(s)}$) and

$$A_{1,2,\dots,p,\dots,n}^{t,(s)}(x) = \prod_{\text{domain}_s} (\mu^{t,(s)}(x)) = \prod_{j=1}^n (\mu^{t,(s)}(x_j)) \quad (10)$$

The input cluster is adapted to cover all input data associated with the s th output-context.

Therefore (10) can be represented as

$$A_{1,2,\dots,p,\dots,n}^{t,(s)}(x)|_{x \in \text{domain}_s} = 1 \quad (11)$$

Consequently, the output model of equation (9) can be re-written as

$$o = A_{1,2,\dots,p,\dots,n}^{t,(s)}(x) \times c^{t,(s)}|_{x \in \text{domain}_s} \quad (12)$$

Therefore, accuracy of the model output depends on the tuning process of the output domain, $C^{t,(s)}$ that is depicted in (12). Hence, if proper tuning points (distinct points) are found in the output domain, then a balanced partition of the output domain is realized with an effective rule-base having reasonable accuracy.

5.2 Base step: knowledge of the whole output domain

The RCFS begins with first fuzzy rule by translating the knowledge of the global domain, the formation of the first consequent part (when $t=1, s=1$) of the output-context, $C^{t=1,(s=1)}$ which can be described as

$$c^{t=1,(s=1)} = d_1 \text{ and } \sigma^{t=1,(s=1)} = \sigma_{\text{global}} \quad (13)$$

where d_1 is the output of 1st training data and σ_{global} is the spread of the output domain for initial fuzzy rule. Both d_1 and σ_{global} of the Gaussian function embedded in $C^{1,(1)}$. Again, the first antecedent part associated with the p th input variable can be written as Eq. (14):

$$A_p^{t,(1)}(x)|_{x \in \text{domain}_{s=1}} = \prod_{\text{domain}_{s=1}} (\mu^{t,(1)}(x_p)) \quad (14)$$

Remark 1: The RCFS is a self-organizing method. Similarity measure (SM) of two consecutive Gaussian-shaped fuzzy levels has been described in [31]. $SM \rightarrow 1$ shows the more similarities and $SM \rightarrow 0$ shows the more dissimilarities between the two consecutive Gaussian-shaped fuzzy levels [31]. A similarity threshold (β) is considered to identify the best matched output-context with the presented value. If the similarity measure between the output-context and presented value exceeds the

threshold value (i.e. $SM > \beta$) then a prominent distinction is observed and the RCFS forms a new distinct output-context. Conversely, if the similarity measure between the output-context and presented value not exceed the threshold value ($SM \leq \beta$) then RCFS finds the best matched output-context denoted as $C^{t,(s^*)}$ where $s^* = \arg \max_{SM \leq \beta} SM(d_i, C^{t,(s)})$. Therefore, the best matched output-context structure is being able to give a satisfactory description of the presented value. Merging process is then applied to the presented value and best matched output-context.

The step-by-step process for base step is referred to as Algorithm 1. In base step, the goal is to realize a single fuzzy rule by translating the whole output domain in self-organizing way. Therefore, a prominent distinction is undesirable for base step (Algorithm 1) and hence $SM \not> \beta$ is required to continue the self-organizing process that depicts in step 5.

Algorithm 1: Base step

OUTPUT: Base value σ_{global} that covers whole output domain.

REQUIRES: Minimum σ_{global} value is arbitrarily taken.

1. INITIALIZE the first consequent part, $C^{1,(1)}$ using (13).
2. FOR each training data $[x, d]_i$
 3. Define μ for new incoming d_i using (7).
 4. Similarity measure (SM) between new data (step 3) and existing consequent part (s).
 5. IF $SM \leq \beta$ then CONTINUE with step 2. ELSE continue from the beginning (step 1) by increasing σ_{global} .
6. Base value: Final σ_{global} is taken as base value for the initial fuzzy rule.

5.3 Recursive construction of the output-context fuzzy system

RCFS begins with an initial rule and σ_{global} is the width of the output domain in the first rule that covers whole output domain. After that, evolving process starts by decreasing the σ_{global} at a constant rate σ_{evolve} and therefore adding more fuzzy rules and terms. The constant width σ_{evolve} is taken

arbitrary such that the creation of the output partitions slowly increases. Consequently, this evolving process adds more fuzzy rules and terms in order to find the optimum accuracy at a rational partition of the output-context.

At evolving stage $t=2$, formation of the first output-context ($C^{2,(1)}$) where more than one rule generation ($s > 1$) is considered

$$c^{2,(1)} = d_i \text{ and } \sigma^2 = \sigma_{init}^2 = \sigma_{init}^1 - \sigma_{evolve} \quad (15)$$

Recall that, σ_{evolve} is taken arbitrary such that the creation of the output partitions slowly increases.

By assuming that there is an existing output-context and the RCFS then proceeds with computing the similarities between the presented value and existing output-context structures. The similarity match (see *Remark 1*) between the output d_i of the i th training data $[x, d]_i$ and an existing output-context ($C^{t,(s)}$) is given as

$$SM(d_i, C^{t,(s)}) = \mu^{t,(s)}(c^{t,(s)}, \sigma^{t,(s)}; d_i) \quad (16)$$

If $SM > \beta$, then a prominent distinction is observed and RCFS forms a new distinct output-context at evolving stage, t .

$$c^{t,(s+1)} = d_i \text{ and } \sigma^{t,(s+1)} = \sigma_{init}^t = \sigma_{init}^{t-1} - \sigma_{evolve} \quad (17)$$

Similarity measure for self-organizing system describes in *Remark 1*. Merging process is done only for the sigma (width) between two consecutive membership functions as [32]. Assuming that $[a_1, b_1, c_1] \in C^{exist}$ and $[a_2, b_2, c_2] \in C^{new}$, where C^{exist} and C^{new} are two fuzzy level for training data $[x, d]_i$, respectively. The points $[a_1, b_1, c_1]$ and $[a_2, b_2, c_2]$ denote the left support, right support, and center, respectively. Therefore, the average width $\sigma_{avg}^{t,(s)}$ between the best matched output-context structure and the present data d_i can be written as Eq. (17):

$$\sigma_{avg}^{t,(s)} = \frac{1}{2\sqrt{\ln 2}} \{\max(b_1, b_2) - \min(a_1, a_2)\} \quad (18)$$

Recursive construction of the output-context fuzzy system depict in Algorithm 2. Balanced number of rules (BNR) index is introduced to terminate the recursion and it will be discussed in Remark 2.

Algorithm 2: Recursion procedure

OUTPUT: Prominent distinction points in the output domain which depicts the center of the consequent part of the fuzzy system and its corresponding input cluster.

REQUIRES: Base step (Algorithm 1) to find the base value, $\sigma_{global} \cdot \sigma_{evolve}$, a constant width which taken arbitrarily, ensures the rule number slowly increases.

INITIALIZE: single fuzzy rule by translating the whole output domain using (13) and (14).

Initialize $\sigma_{init}^{t=1} = \sigma_{global}$ and $BNR=0$.

FUNCTION RCFS($\sigma_{init}^t, C^{t,(s)}$)

1. WHILE recursive termination not occur as $BNR \neq 1$

2. FOR each training data $[x, d]_i$

3. Define μ for new incoming d_i using (7).

4. Similarity measure (SM) between new data (step 3) and existing consequent part(s) using (16).

5. IF $SM > \beta$, THEN a prominent distinction is observed as (17).

ELSE $SM \leq \beta$ THEN merge and adapt using (18) for both consequent and antecedent part.

6. Update for the next stage. $\sigma_{init}^t = \sigma_{init}^{t-1} - \sigma_{evolve}$, existing consequent

part(s) $C^{t,(s)}$ and $BNR_t = \frac{(NR_t - NR_{t-1})^2}{NR_t}$

7. FUNCTION RCFS($\sigma_{init}^t, C^{t,(s)}$).

8. IF step 1 not true, RETURN distinct points $(C^{t,(s)})$ and $A_{1,2,\dots,p,\dots,n}^{t,(s)}(x)$.

Remark 2: In Algorithm 2, a recursion index called balanced number of rules (BNR) index is used for the termination criteria. As the evolving stage proceeds, the width of the output-context (or σ_{init}^t) becomes small. At a specific evolving stage, the width of the output-context becomes so small and hence unbalanced that it cannot represent the data properly. Therefore, it leads a fuzzy system with a large number of rules and causes the over-fitting (i.e. fit very close data because of the small width of output structure).

Balanced number of rules (BNR) index is used to handle the situation from the unbalance partitions or over-fitting that clarifies in *Remark 2*. BNR_t is a straightforward index to recognize the unbalance situation in the recursive process at t evolving stage and terminate the algorithm from further evolving.

$$BNR_t = \frac{(NR_t - NR_{t-1})^2}{NR_t} \quad (19)$$

where NR_t and NR_{t-1} are the number of rules at t and $(t - 1)$ evolving stage, respectively. The characteristic of the BNR index as follows:

1) According to (15), $BNR \in [0,1]$ is showed the balanced partition in the output domain.

Characteristics equation for BNR as follows:

$$BNR_t = \begin{cases} \geq 0 & \text{balanced partition} \\ > 1 & \text{unbalanced partition} \end{cases} \quad (20)$$

The estimation of BNR is fully online, not predefined threshold, approximate from the current and previous evolving stage. At initialization, assume that $BNR_{t=1} = 0$ imply as a base value.

2) If BNR_t holds the first criteria from (16), then evolving process will proceed and the RCFS proceeds to the rational partition in order to obtain optimum accuracy. At a specific evolving stage, BNR index becomes greater than 1 and hence it holds the second criteria from (20). Index $BNR_t > 1$ means that further evolving process will cause the unbalanced/over-fitting to the model and it is right time to terminate.

Therefore, we will get a fuzzy system with an effective rule-base. Thus, this Global learning by BNR index find the effective rule-base fuzzy system with lowest approximation error and hence reasonable accuracy for testing data. Standard root mean square error (RMSE) of t th evolving process is given as:

$$E_e = \sqrt{\frac{1}{N} \sum_{i=1}^N (o(x_i) - des_i)^2} \quad (21)$$

where N is the number of training data and, $o(x_i)$ and des_i are the model output and desired output of the i th training data.

If $BNR_t > 1$ is found at evolving stage j , then termination criteria is fulfilled for the RCFS algorithm. Therefore, an effective rule-base is found at evolving stage e as (6) where minimum error is found.

$$e = \{1, 2, \dots, e, \dots, j\} \text{ and } E_e = \min(E_1, E_2, \dots, E_e, \dots, E_j) \quad (22)$$

5.4 Characteristics of the RCFS

The main aspects of the RCFS are as follows:

1. The RCFS automatically observes the prominent distinction point on the output domain.
2. The evolving process is terminated depending on the BNR index which is fully online, does not follow a predefined threshold, and is approximated from the current and previous evolving stages.
3. The RCFS uses number of rules (which means number of distinct points/output-contexts) for termination. Nevertheless, the realization of the effective rule base uses both feature (input) and

output space since it consider minimum error using (21) and (22). Hence, more accurate representation of data can be achieved in the effective rule base.

6. Thermographic diagnostic of electrical components

Automatic feature extraction for thermographic diagnosis of electrical components has been discussed in Section 4 and its actual output data ΔT depicted in Fig. 6. Seven attributes exists in this datasets including six input feature and its corresponding output ΔT . At first, all datasets are used to realize the effective rule base as described in Algorithm 2 (Section 5.3). After that, 10-fold cross validation is used to train and test the data.

In this paper, similarity threshold (β) represents the similarity between two consecutive fuzzy levels, and $\alpha > 0$ is the minimum membership value that strongly covers the UoD [13]. In the following example, the threshold $\beta = 0.8$ means that $1 - 0.8 = 0.2$. Hence, a distinct fuzzy level is formed when the similarity among consecutive fuzzy levels is less than 20%. Furthermore, $\alpha = 0.5$ is considered to strongly cover the UoD [13] and to keep the average semantic meaning in the RCFS model. The evolving constant σ_{evolve} is taken arbitrarily to ensure that the rule number slowly increases.

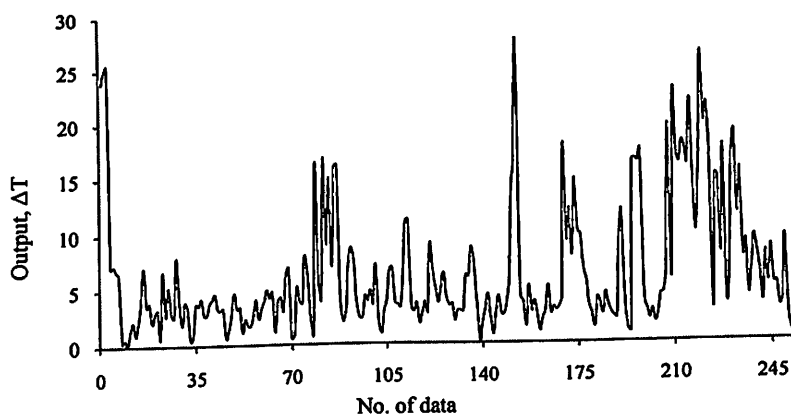


Fig. 6. Actual output data, ΔT (°C) for thermographic diagnostic of electrical components.

The RCFS begins with first fuzzy rule by translating the knowledge of the global domain and $\sigma_{global} = 7.5$ is found from Algorithm 1 (Section 5.2) as base value for the initial fuzzy rule. The RCFS evolves the output-contexts (or rules) as Algorithm 2 (Section 5.3), and $\sigma_{evolve} = 0.5$ is taken

arbitrarily to slowly increase the formation of the fuzzy rules which depicts in Table 2. The termination point $BNR > 1$ is found at the evolving stage, where $\sigma_{init}^f = 2.5$. Therefore, an effective rule base is found at the evolving stage with a minimum error. Hence, in Table 2, the RCFS finds the effective rule base for thermographic diagnostic of electrical components when $\sigma_{init}^f = 6.5$ and the number of fuzzy rules are two. Fig. 7 depicts the consequent parts which are realized by the self-organizing way and distinct points show the semantic interpretability. Distinct points for thermographic diagnostic of electrical components show ΔT values which is used as the decision making parameter about the condition of overheated component. Therefore, Evolving process recursively construct the fuzzy system (Table 2) and finally realize an effective rule base with two distinct points (Fig. 7).

Table 2: Evolving process of the RCFS to obtain an effective rule base for thermographic diagnostic of electrical components

σ_{init}^f	No. of rules	Training error in RMSE	BNR	σ_{init}^f	No. of rules	Training error in RMSE	BNR
7.5	1	18.3974	0	4.5	2	4.3824	0
7.0	2	4.4896	0.5	4.0	2	4.3824	0
6.5	2	4.3824, minimum	0	3.5	2	4.3824	0
6.0	2	4.3824	0	3.0	2	4.3824	0
5.5	2	4.3824	0	2.5	4	3.2737	1.0, termination
5.0	2	4.3824	0				

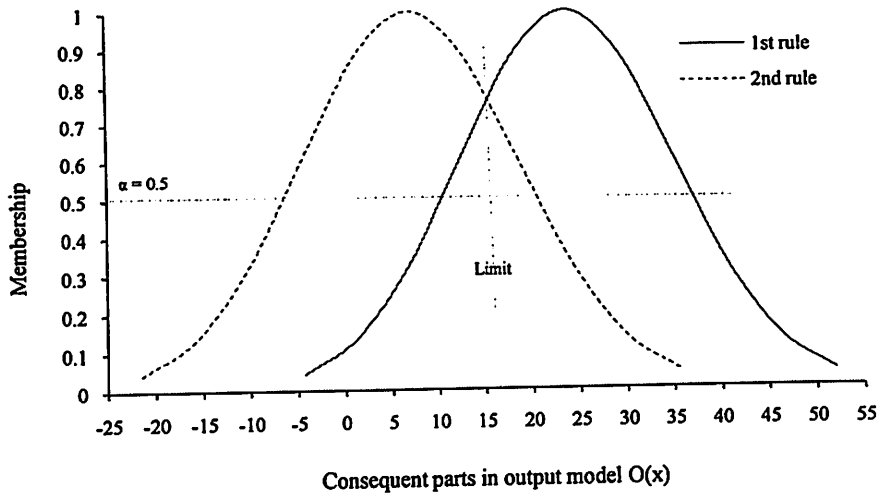


Fig. 7. Two distinct points $\Delta T(^{\circ}\text{C}) = \{23.8, 7.1\}$ as the center of consequent parts shows the semantic interpretability.

The main feature of the RCFS is to realize the prominent distinct points (i.e. ΔT points) in the output domain and consequent parts are formed by considering the distinction points as center (Section 5.1 and Fig. 7). Therefore, two prominent distinct points $\Delta T(^{\circ}\text{C}) = \{7.1, 23.8\}$ are realized by the RCFS. However, the condition of overheated component (i.e. ΔT) might be required the distinction point(s) lower than 7.1 since it is expected to ensure the highly predictive automated system. This highly predictive method will enhance the acute diagnostic of electrical equipment. Therefore, further recursion of Algorithm 2 might be required.

Table 3: Further recursion of Algorithm 2 to obtain an effective rule base

σ_{int}^t	no. of rules	Training error in RMSE	BNR
2	4	2.9472, minimum	0
1.5	4	3.2241	0
1	9	2.9671	2.77, terminate
1.5	15	3.1304	2.4

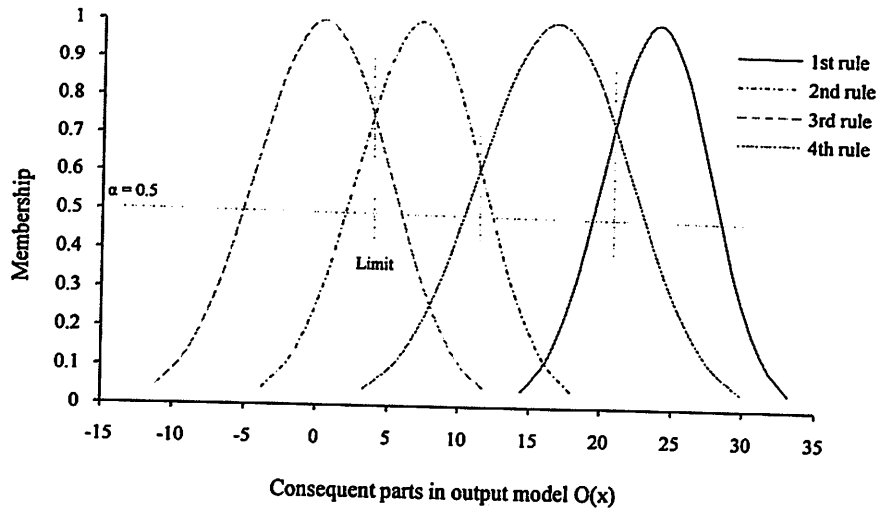


Fig. 8, Further recursion and $\sigma_{init}^f = 2$. Four distinct points $\Delta T(^{\circ}\text{C}) = \{23.8, 7.1, 0.3, 16.6\}$ as the center of consequent parts shows the semantic interpretability.

Further recursion of Algorithm 2 evolves the output-contexts and hence increases the fuzzy rules which depicts in Table 3. The termination point $\text{BNR} > 1$ is found at the evolving stage, where $\sigma_{init}^f = 1$. Therefore, another effective rule base is found at the evolving stage with a minimum error. Hence, in Table 3, the RCFS finds the effective rule base with four distinct points (Fig. 8) for thermographic diagnostic of electrical components when $\sigma_{init}^f = 2$.

We get $\sigma_{init}^f = \{6.5, 2\}$ to realize the effective rule base and now ready for performance analysis in terms error and number of rules. Table 4 shows the performance analysis in RMSE for thermographic diagnostic of electrical components. Recall that, 10-fold cross validation is used to train and test the data.

Table 4: Performance analysis in RMSE for thermographic diagnostic of electrical components

Rule base	No. of rules	Training error	Testing error
Effective rule base, $\sigma_{init}^f = 6.5$	2	4.39	6.72
Effective rule base considering further recursion, $\sigma_{init}^f = 2$	4	3.02	5.97

Since prominent distinction points for ΔT are realized by the RCFS, we might get classification of conditions of electrical components depicts in Table 5, 6 and 7. Limit of the conditions depicts in Fig. 7 and Fig. 8. In Table 5 and Table 6, The RCFS realizes two recommended class namely defect and no defect. Furthermore, three recommended class can be realized by the RCFS (Table 7) namely defect, warning and no defect. These classifications of conditions for ΔT , therefore, an intelligent condition monitoring and an automatic defect analyzing system which improves the defect detection technique of electrical equipment abnormalities. Therefore, this expert system for thermographic diagnosis incorporate the automatic feature extraction (Section 4), classify the condition of hotspots according to the priority level of repairing and comments for further actions.

Classification accuracy for test data reduces (Table 5 and 6) but testing error in RMSE decreases (Table 4) while considering further recursion. Again, in Table 6 and 7, classification accuracy decreases if classification of conditions increases. We analyze the reason behind these two matters as follows: datasets are highly noisy data (ill-defined input space) and Fig. 6 shows the variations on output data, ΔT . Hence, accuracy decreases when classification of conditions are applied to the datasets. Another reason is further recursion ($\sigma_{init}^t = 2$ and rule = 4) after first termination and this shows the effectiveness of BNR index. *Remark 2* and Eq. 20 discusses about the overfitting state means that this situation can not represent the input data properly. We may compare the proposed RCFS with well known classifier MLP in terms of accuracy that depicts in Table 8 (MLP is used because of its availability in MatLab although its uses multilayered algorithm). The purpose of the comparison in Table 8, however, shows the reasonable accuracy that is achieved by the RCFS.

Table 5: Classification of conditions of electrical components realized by the RCFS when number of rules are 2 and $\sigma_{init}^t = 6.5$

Class	$\Delta T(^{\circ}\text{C})$	Recommended actions	Classification accuracy on test data (%)

Defect	$\Delta T > 15.3$ Priority 1	Major discrepancy; repair immediately	92
No defect	$\Delta T \leq 15.3$ Priority 2	Minor overheating; warrants investigation	

Table 6: Classification of conditions of electrical components realized by the RCFS when number of rules are 4 and $\sigma_{init}^t = 2$

Class	$\Delta T(^{\circ}\text{C})$	Recommended actions	Classification accuracy on test data (%)
Defect	$\Delta T > 12.8$ Priority 1	Major discrepancy; repair immediately	84
No defect	$\Delta T \leq 3.8$ Priority 2	Minor overheating; warrants investigation	

Table 7: Classification of conditions of electrical components realized by the RCFS when number of rules are 4 and $\sigma_{init}^t = 2$

Class	$\Delta T(^{\circ}\text{C})$	Recommended actions	Classification accuracy on test data (%)
Defect	$\Delta T > 12.8$ Priority 1	Major discrepancy; repair immediately	80
Warning	$3.8 < \Delta T \leq 12.8$ Priority 2	Probable deficiency; repair as time permits	
No defect	$\Delta T \leq 3.8$ Priority 3	Minor overheating; warrants investigation	

Table 8. Comparison of the classification accuracy on test data between MLP network and the proposed RCFS

Class	Classifier	Accuracy (%)
Defect and No defect. Conditions of ΔT from Table 6	MLP with LM	73.26
	Proposed RCFS	84.00
Normal, warning and critical. Conditions of ΔT from Table 7	MLP with LM	64.35
	Proposed RCFS	80.00

7. Conclusion

In this study, early detection of hotspots and thermal condition monitoring based thermographic fault diagnosis of components is proposed for maintaining reliable and uninterrupted operation of power system which allows predictive and preventive diagnosis before the major failure occurs. Six intensity features (i.e. maximum, minimum, mean, median, standard deviation and variance) were extracted from each hotspot. Then, the extracted features fed to the RCFS. The RCFS automatically realizes the conditions of components which formulate the abnormalities of electrical components into classes like normal, warning and critical. The RCFS is an output-context fuzzy system and recursively construct the fuzzy system by realizing the prominent distinction in the output domain (ΔT). The termination of the recursive algorithm is fully online, not depend on threshold and this termination find an effective rule base for the datasets. Therefore, well representation of the datasets can be achieved by the effective rule base. Hence, effective rule base is the balanced fuzzy model of the diagnostic system since it avoids the over-fitting condition.

The characteristics of the expert system (including automatic feature extraction and the RCFS) are described in this paper and evaluated with the datasets of the thermographic diagnostic of electrical components. The evaluation intelligently realizes the limits of the ΔT (i.e. conditions of components) which are acceptable in the sense and therefore, an automatic defect analyzing system which improves the defect detection technique of electrical equipment abnormalities. Furthermore, the proposed expert system realizes less number of fuzzy rules or distinct ΔT with reasonable accuracy, hence shows the

significance of the expert system. For future work, more efficient image processing method and classifiers will be employed to acquire better performance of the intelligent system.

Acknowledgements

???

References

- [1] Korendo, Z., & Florkowski, M. (2001). Thermography based diagnostic of power equipment. *IEEE Power Engineering Journal*. 15(1), 33-42.
- [2] Huda ASN, Taib S. Suitable features selection for monitoring thermal condition of electrical equipment using infrared thermography. *Infrared Physics & Technology* 2013;61;184-191.
- [3] Almedia, C. A. L., Braga, A. P., Nascimento, S., Paiva, V., Martins, H. J. A., Torres, R., & Caminhas, W. M. (2009). Intelligent thermographic diagnostic applied to surge arresters: a new approach. *IEEE Transactions on Power Delivery*. 24(2), 751-757.
- [4] Shafi'i, M. A. & Hamzah, N. (2010). Internal fault classification using artificial neural network. 4th International Power Engineering and Optimization Conference, 23-24 June, Shah Alam. 352-357.
- [5] Smedberg, M. (2006). Thermographic decision support – detecting and classifying faults in infrared images. Master Thesis, Royal Institute of Technology, Stockholm, Sweden.
- [6] Wretman, D. (2006). Finding regions of Interest in a decision support system for analysis of infrared images. Master Thesis, Royal Institute of Technology, Stockholm, Sweden.
- [7] Nazmul Huda AS, Taib S, Jadin MS, Ishak D. A semi-automatic approach for thermographic inspection of electrical installations within buildings. *Energy and Buildings* 2012; 55:585-591.

- [8] Rahmani, A., Haddadnia, J., & Seryasat, O. (2010). Intelligent fault detection of electrical equipment in ground substations using thermo vision technique. Proceedings of 2nd International Asia Conference on Mechanical and Electronics Engineering, 1-3 August, Kyoto. 150-154.
- [9] Nazmul Huda AS, Taib S. Application of infrared thermography for predictive/preventive maintenance of thermal defect in electrical equipment. *Applied Thermal Engineering* 2013;61(2): 220-227.
- [10] Huda ASN, Taib S, Ghazali KH, Jadin MS. A new thermographic NDT for condition monitoring of electrical components using ANN with confidence level analysis. *ISA Transactions* (2014), <http://dx.doi.org/10.1016/j.isatra.2014.02.003i>.
- [11] D. Wang, X.-J. Zeng,, J. A. Keane, An evolving-construction scheme for fuzzy system, *IEEE Trans. Fuzzy Syst.* 18(2000) 4(2010) 755–770.
- [12] D. Nauck, F. Klawonn, and R. Kruse, *Foundations of Neuro-Fuzzy Systems*. New York: Wiley, 1997.
- [13] S. W. Tung, C. Quek, C. Guan, SaFIN: A self-adaptive fuzzy inference network, *IEEE Trans. Neural Netw.* 22(2011) 1928–1940.
- [14] D. Wang, X.-J. Zeng, J. A. Keane, An output-constrained clustering approach for the identification of fuzzy systems and fuzzy granular systems, *IEEE Trans. Fuzzy Syst.* 19(2011) 1127–1140.
- [15] C. Mencar, C. Castiello, R. Cannone, A. M. Fanelli, Design of fuzzy rule-based classifiers with semantic cointension, *Inf. Sci.* 181(2011) 4361–4377.
- [16] W. Pedrycz, K.-C. Kwak, Linguistic models as a framework of user-centric system modeling, *IEEE Trans. Syst. Man Cybern. A, Syst., Humans* 36(2006) 727–745.
- [17] Kregg, M. A. (2004). Benefits of using infrared thermography in utility substations. Proceedings of SPIE, 12 April, Orlando, Florida. 249-257.

- [18] TBPPM, The basics of predictive / preventive maintenance (2011) [online], [Accessed 02 August 2011], http://www.cedesa.com.mx/pdf/fluke/flukeTi30_serie_nota.pdf
- [19] Epperly, R. A., Heberlein, G. E., & Eads, L. G. (1997). A tool for reliability and safety: predict and prevent equipment failures with thermography. IEEE Conference on Petroleum and Chemical Industry, 15-17 September, Banff, Alta. 59-68.
- [20] Statistics fire causes building fire investigation the fire department and rescue Malaysia, Year 2012, nationwide, from January to June (2012) [online], [Accessed 01 August 2012], Available from World Wide Web: [http://www.bomba.gov.my/fileupload/statistik/ST-PGK%20PK01\(2012\).pdf](http://www.bomba.gov.my/fileupload/statistik/ST-PGK%20PK01(2012).pdf)
- [21] Coutin, M., Plumecocq, W., Melis, S., & Audouin, L. (2012). Energy balance in a confined fire compartment to assess the heat release rate of an electrical cabinet fire. Fire Safety Journal. (52), 34–45.
- [22] Plumecocq, W., M. Coutin, M., Melis, S., & Rigollet, L. (2011). Characterization of closed-doors electrical cabinet fires in compartments. Fire Safety Journal. (46), 243–253.
- [23] Babrauskas, V. (2001) How do electrical wiring faults lead to structure ignitions, 7th international Fire & Materials conference, 22 -24 January, San Francisco, USA. 39-50.
- [24] Garnaik, S. P. Infrared Thermography : A versatile technology for condition monitoring and energy conservation (2011) [online], [Accessed 01 July 2011], Available from World Wide Web: http://www.reliabilityweb.com/art05/Infrared_thermography_a_verseatile_technology.pdf
- [25] Standard for Infrared Inspection of Electrical Systems & Rotating Equipment, Infraspction Institute. <http://www.armco-inspections.com/files/ir/Electrical%20Rotating%20Std.pdf>.2008.
- Accessed 11 June 2011
- [26] ASTM, ASTM E 1934: Standard Guide for Examining Electrical and Mechanical Equipment with Infrared Thermography, West Conshohocken, Pennsylvania, ASTM International, 2005.

- [27] A.H.A. Bakar, H.A. Illias, M.K. Othman, H. Mokhlis, Identification of failure root causes using condition based monitoring data on a 33 kV switchgear, *International Journal of Electrical Power and Energy Systems*, 47 (2013) 305–312.
- [28] Daims H., Lucker S., Wagner M.: daime, a novel image analysis program for microbial ecology and biofilm research. *Environ. Microbiol.* 8, 200-213,2006
- [29] B. Kosko, Optimal fuzzy rules cover extrema, *Int. J. Intell. Syst.* 10(1995) 249–255.
- [30] Y. S. Ding, H. Ying, S. H. Shao, Necessary conditions on minimal system configuration for general MISO mamdani fuzzy systems as universal approximators, *IEEE Trans. Syst. Man Cybern. B, Cybern.* 30(2000) 857–864.
- [31] Y. Jin, W. von Seelen, B. Sendhoff, On generating FC^3 fuzzy rule systems from data using evolution strategies, *IEEE Trans. Syst. Man Cybern. B, Cybern.* 29(1999) 829–845.
- [32] W. L. Tung C. Quek, eFSM—A novel online neural-fuzzy semantic memory model, *IEEE Trans. Neural Netw.* 21(2012) 136–157.

Enhancement of the low contrast image using fuzzy set theory

Khairunnisa Hasikin

Imaging and Intelligent System Research Team (ISRT)
School of Electrical and Electronic Engineering
Engineering Campus, Universiti Sains Malaysia
14300 Nibong Tebal, Penang, Malaysia
eishahasikin@gmail.com

Nor Ashidi Mat Isa

Imaging and Intelligent System Research Team (ISRT)
School of Electrical and Electronic Engineering
Engineering Campus, Universiti Sains Malaysia
14300 Nibong Tebal, Penang, Malaysia
ashidi@eng.usm.my

Abstract—This paper presents a fuzzy grayscale enhancement technique for low contrast image. The degradation of the low contrast image is mainly caused by the inadequate lighting during image capturing and thus eventually resulted in nonuniform illumination in the image. Most of the developed contrast enhancement techniques improved image quality without considering the nonuniform lighting in the image. The fuzzy grayscale image enhancement technique is proposed by maximizing fuzzy measures contained in the image. The membership function is then modified to enhance the image by using power-law transformation and saturation operator. The qualitative and quantitative performances of the proposed method are compared with the other methods. The proposed method produced better quality enhanced image and required minimum processing time than the other methods.

Keywords—nonuniform illumination; fuzzy; grayscale; enhancement; overexposed image; underexposed image

I. INTRODUCTION

Good contrast images with preserving details are required for many important areas namely machine vision, remote sensing, dynamic and traffic scene analysis, biomedical image analysis and autonomous navigation. However most of the recorded images suffer from poor contrast which is due to the inadequate lighting during image acquiring, wrong setting of aperture size and shutter speed as well as nonlinear image intensities mapping.

Difficulties in controlling the lighting conditions during image acquisition process have resulted in variability in image illumination. The captured images turn out to be low contrast and contained underexposed and overexposed regions.

Thus, image enhancement has been employed to increase the quality of the image. Image enhancement is a fundamental task applied in image processing to improve interpretability and appearance of the image. It provides better input image for further image processing task.

Image enhancement can be clustered into two groups namely frequency domain and spatial domain methods. In the frequency domain method, the enhancement is conducted by modifying the frequency transform of the image. Meanwhile in the latter method, image pixels are directly modified to enhance the image. However, computing the enhancement in frequency domain is time consuming

process even with fast transformation technique thus made it unsuitable for real time application [1].

Numerous contrast enhancement techniques normalized the image intensities and often fail to produce satisfactory results for a broad range of non-uniform illumination image. The image is characterized by the fact that the amplitudes of their histogram components are very high at one or several locations on the grayscale, while they are very small, but not zero, in the rest of the grayscale. This makes it difficult to increase the image contrast by simply stretching its histogram. The high amplitude of the histogram components also often prevents the use of the histogram equalization (HE) techniques. Most of the HE techniques could cause a washed-out effect on the appearance of the enhanced image and/or amplify existing noises [2].

In addition, due to the poor and low contrast nature of the acquired image, vagueness and ambiguity are introduced and have led to the increment of uncertainty in the image information. This vagueness in the image appears in the form of imprecise boundaries and intensities during image digitization.

Therefore, fuzzy sets theory [3] has been proposed as a problem solving tool between the precision of classical mathematics and the inherent imprecision of the real world. The imprecision possessed by the acquired image can be perceived qualitatively by human reasoning. However, there is no specific quantification to describe the imprecision and thus machine may not understand them. Realizing this limitation to a great extent, fuzzy logic tools empower a machine to mimic human reasoning.

In the image enhancement field, the fuzzy set theory has been widely utilized by other researchers [1, 4-17]. Pixel property such as gray tone intensity is modeled into a fuzzy set using a membership function. The image is considered as an array of fuzzy singletons having a membership value that denotes the degree of belonging to specific property.

The conventional method of fuzzy enhancement is conducted by using contrast intensification (INT) operator [3]. In this method, dynamic range of image is possible to be obtained since the INT operator will increase and decrease the membership degree above and below threshold value respectively. However, the INT is solely depended on the membership function and it needs to be applied continuously on the image to attain desired enhancement.

This limitation is then improved by [7] using new intensification operator (NINT) which utilized sigmoid function. The NINT does not change uniformly because the membership function is marginally changing, thus computational time can be reduced as compared to the INT.

The other enhancement approach is conducted using fuzzy rule-based technique [18-20] which human intuition is incorporated to make soft decisions on each condition. This method suffers from high computational time and thus made it difficult to automatically generate fuzzy rules. In addition, prior threshold selection needs to be done for each condition and therefore made choosing the optimum threshold could be challenging.

Image enhancement is also done by measuring information contained in the image [11, 13, 21, 22]. The membership function is chosen based on the measured quantity such as image entropy or index of fuzziness. An optimum quantitative measure has to be determined in order to achieve best enhancement quality.

In addition, attempts have been made to enhance the image locally [6, 12, 23]. In this technique, local contrast in small regions is enhanced while at the same time preventing an increase in global contrast. Fine edges which are neglected in global enhancement are enhanced and clarity of the enhanced image is improved. However, noises and artifacts might also be enhanced during the enhancement process.

In this paper, a new contrast enhancement technique has been proposed by minimizing fuzziness in the image without requiring complex procedure and long computational time.

II. THE PROPOSED TECHNIQUE

The original image of size $M \times N$ has intensity levels $m_{i,j}$ in the range of $[0, L-1]$ can be considered as a collection of fuzzy singletons in the fuzzy set notation.

$$I = \cup \{\mu(m_{i,j})\} = \{\mu_{i,j} / m_{i,j}\} \quad (1)$$

$i = 1, 2, \dots, M; j = 1, 2, \dots, N$

where $\mu(m_{i,j})$ or $\mu_{i,j}/m_{i,j}$ represents the membership or grade of belonging $\mu_{i,j}$ of $m_{i,j}$ being the grayscale intensity at the (i,j) th pixel.

A. Fuzzy Measures

Let I be a set with randomly gray level values $\{m_{i,j}$ at (i,j) th pixel $\}$ and $\{p_0, p_1, p_2, \dots, p_{L-1}\}$ are respective probabilities of gray level values. The fuzzification of set I induced two kinds of uncertainties. The first part of uncertainties induced by the random nature of the image given by:

$$H = - \sum_{k=0}^{255} p_k \log(p_k) \quad (2)$$

Meanwhile second uncertainty arises from the fuzziness of the fuzzy set related to the ordinary set given by:

$$J = -\mu_k \log(\mu_k) - (1-\mu_k) \log(1-\mu_k) \quad (3)$$

Therefore, the total entropy, E is expressed by:

$$E = H + J \quad (4)$$

The index of fuzziness (IOF) is calculated using equation (5) as follow:

$$IOF = \frac{2}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \min\{p_{i,j}, (1-p_{i,j})\} \quad (5)$$

where

$$p_{i,j} = \text{sim} \left[\frac{\pi}{2} \times \left(1 - \frac{\mu(m)}{\mu_{\max}} \right) \right] \quad (6)$$

B. Membership Function Calculation

The shape of S-function is commonly used for the representation of the degree of brightness or whiteness of pixels in the grey levels images. The S-function was originally introduced by [2] and the definition of more flexible S-function was proposed by [23]. The flexible S-function has been adapted in this study to fuzzify the original image.

$$\mu(m) = \begin{cases} 0 & \text{for } m \leq a \\ \frac{m-a}{(b-a)(c-a)} & \text{for } a < m \leq b \\ 1 - \frac{(m-c)^2}{(c-b)(c-a)} & \text{for } b < m \leq c \\ 1 & \text{for } m \geq c \end{cases} \quad (7)$$

where m is the intensity of the image and a , b and c are parameters that determined the shape of the S-function. The parameters a , b and c are specified to ensure the membership function maximizes the information contained in the image. It is done by incorporating two fuzzy measures namely fuzzy entropy and index of fuzziness.

Therefore parameters a , b and c are given by equations (8) to (10).

$$a = \alpha E_{\max} \quad (8)$$

$$b = \beta |IOF_{\max} - E_{\max}| \quad (9)$$

$$c = \gamma IOF_{\max} \quad (10)$$

where α , β and γ are the membership factors that are chosen to obtain optimum S-membership function if fuzzified image. IOF_{\max} and E_{\max} are maximum index of fuzziness and maximum entropy respectively.

The calculated membership function transformed the image intensity levels from the spatial domain to fuzzy domain. The original image has been transformed and most of the regions in the image contained mixed region of overexposed and underexposed regions. Therefore, a parameter called 'exposure' is introduced to denote percentage of the image gray levels is underexposed and

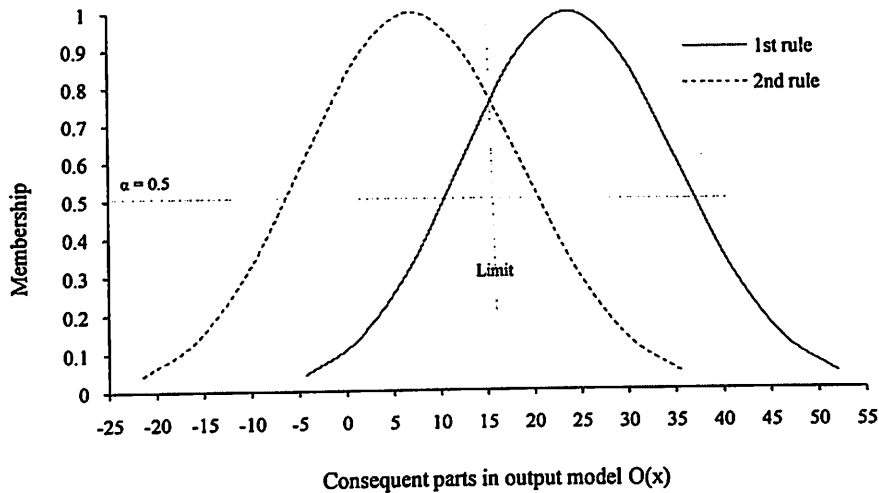


Fig. 7. Two distinct points $\Delta T(^{\circ}\text{C}) = \{23.8, 7.1\}$ as the center of consequent parts shows the semantic interpretability.

The main feature of the RCFS is to realize the prominent distinct points (i.e. ΔT points) in the output domain and consequent parts are formed by considering the distinction points as center (Section 5.1 and Fig. 7). Therefore, two prominent distinct points $\Delta T(^{\circ}\text{C}) = \{7.1, 23.8\}$ are realized by the RCFS. However, the condition of overheated component (i.e. ΔT) might be required the distinction point(s) lower than 7.1 since it is expected to ensure the highly predictive automated system. This highly predictive method will enhance the acute diagnostic of electrical equipment. Therefore, further recursion of Algorithm 2 might be required.

Table 3: Further recursion of Algorithm 2 to obtain an effective rule base

σ_{init}^t	no. of rules	Training error in RMSE	BNR
2	4	2.9472, minimum	0
1.5	4	3.2241	0
1	9	2.9671	2.77, terminate
1.5	15	3.1304	2.4

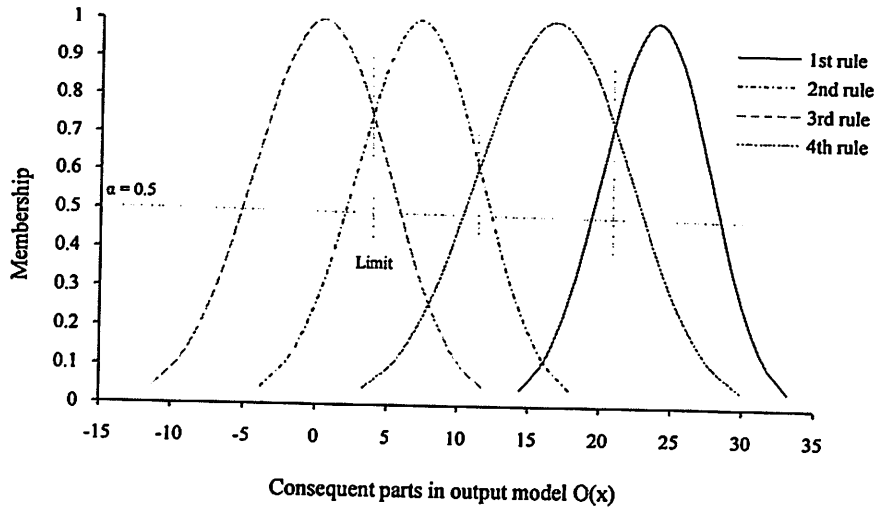


Fig. 8. Further recursion and $\sigma_{init}^t = 2$. Four distinct points $\Delta T(^{\circ}\text{C}) - \{23.8, 7.1, 0.3, 16.6\}$ as the center of consequent parts shows the semantic interpretability.

Further recursion of Algorithm 2 evolves the output-contexts and hence increases the fuzzy rules which depicts in Table 3. The termination point $\text{BNR} > 1$ is found at the evolving stage, where $\sigma_{init}^t = 1$. Therefore, another effective rule base is found at the evolving stage with a minimum error. Hence, in Table 3, the RCFS finds the effective rule base with four distinct points (Fig. 8) for thermographic diagnostic of electrical components when $\sigma_{init}^t = 2$.

We get $\sigma_{init}^t = \{6.5, 2\}$ to realize the effective rule base and now ready for performance analysis in terms error and number of rules. Table 4 shows the performance analysis in RMSE for thermographic diagnostic of electrical components. Recall that, 10-fold cross validation is used to train and test the data.

Table 4: Performance analysis in RMSE for thermographic diagnostic of electrical components

Rule base	No. of rules	Training error	Testing error
Effective rule base, $\sigma_{init}^t = 6.5$	2	4.39	6.72
Effective rule base considering further recursion, $\sigma_{init}^t = 2$	4	3.02	5.97

Since prominent distinction points for ΔT are realized by the RCFS, we might get classification of conditions of electrical components depicts in Table 5, 6 and 7. Limit of the conditions depicts in Fig. 7 and Fig. 8. In Table 5 and Table 6, The RCFS realizes two recommended class namely defect and no defect. Furthermore, three recommended class can be realized by the RCFS (Table 7) namely defect, warning and no defect. These classifications of conditions for ΔT , therefore, an intelligent condition monitoring and an automatic defect analyzing system which improves the defect detection technique of electrical equipment abnormalities. Therefore, this expert system for thermographic diagnosis incorporate the automatic feature extraction (Section 4), classify the condition of hotspots according to the priority level of repairing and comments for further actions.

Classification accuracy for test data reduces (Table 5 and 6) but testing error in RMSE decreases (Table 4) while considering further recursion. Again, in Table 6 and 7, classification accuracy decreases if classification of conditions increases. We analyze the reason behind these two matters as follows: datasets are highly noisy data (ill-defined input space) and Fig. 6 shows the variations on output data, ΔT . Hence, accuracy decreases when classification of conditions are applied to the datasets. Another reason is further recursion ($\sigma_{init}^t = 2$ and rule = 4) after first termination and this shows the effectiveness of BNR index. *Remark 2* and Eq. 20 discusses about the overfitting state means that this situation can not represent the input data properly. We may compare the proposed RCFS with well known classifier MLP in terms of accuracy that depicts in Table 8 (MLP is used because of its availability in MatLab although its uses multilayered algorithm). The purpose of the comparison in Table 8, however, shows the reasonable accuracy that is achieved by the RCFS.

Table 5: Classification of conditions of electrical components realized by the RCFS when number of rules are 2 and $\sigma_{init}^t = 6.5$

Class	$\Delta T(^{\circ}\text{C})$	Recommended actions	Classification accuracy on test data (%)

Defect	$\Delta T > 15.3$ Priority 1	Major discrepancy; repair immediately	92
No defect	$\Delta T \leq 15.3$ Priority 2	Minor overheating; warrants investigation	

Table 6: Classification of conditions of electrical components realized by the RCFS when number of rules are 4 and $\sigma_{init}^f = 2$

Class	$\Delta T(^{\circ}\text{C})$	Recommended actions	Classification accuracy on test data (%)
Defect	$\Delta T > 12.8$ Priority 1	Major discrepancy; repair immediately	84
No defect	$\Delta T \leq 3.8$ Priority 2	Minor overheating; warrants investigation	

Table 7: Classification of conditions of electrical components realized by the RCFS when number of rules are 4 and $\sigma_{init}^f = 2$

Class	$\Delta T(^{\circ}\text{C})$	Recommended actions	Classification accuracy on test data (%)
Defect	$\Delta T > 12.8$ Priority 1	Major discrepancy; repair immediately	80
Warning	$3.8 < \Delta T \leq 12.8$ Priority 2	Probable deficiency; repair as time permits	
No defect	$\Delta T \leq 3.8$ Priority 3	Minor overheating; warrants investigation	

Table 8. Comparison of the classification accuracy on test data between MLP network and the proposed RCFS

Class	Classifier	Accuracy (%)
Defect and No defect. Conditions of ΔT from Table 6	MLP with LM	73.26
	Proposed RCFS	84.00
Normal, warning and critical. Conditions of ΔT from Table 7	MLP with LM	64.35
	Proposed RCFS	80.00

7. Conclusion

In this study, early detection of hotspots and thermal condition monitoring based thermographic fault diagnosis of components is proposed for maintaining reliable and uninterrupted operation of power system which allows predictive and preventive diagnosis before the major failure occurs. Six intensity features (i.e. maximum, minimum, mean, median, standard deviation and variance) were extracted from each hotspot. Then, the extracted features fed to the RCFS. The RCFS automatically realizes the conditions of components which formulate the abnormalities of electrical components into classes like normal, warning and critical. The RCFS is an output-context fuzzy system and recursively construct the fuzzy system by realizing the prominent distinction in the output domain (ΔT). The termination of the recursive algorithm is fully online, not depend on threshold and this termination find an effective rule base for the datasets. Therefore, well representation of the datasets can be achieved by the effective rule base. Hence, effective rule base is the balanced fuzzy model of the diagnostic system since it avoids the over-fitting condition.

The characteristics of the expert system (including automatic feature extraction and the RCFS) are described in this paper and evaluated with the datasets of the thermographic diagnostic of electrical components. The evaluation intelligently realizes the limits of the ΔT (i.e. conditions of components) which are acceptable in the sense and therefore, an automatic defect analyzing system which improves the defect detection technique of electrical equipment abnormalities. Furthermore, the proposed expert system realizes less number of fuzzy rules or distinct ΔT with reasonable accuracy, hence shows the

significance of the expert system. For feature work, more efficient image processing method and classifiers will be employed to acquire better performance of the intelligent system.

Acknowledgements

???

References

- [1] Korendo, Z., & Florkowski, M. (2001). Thermography based diagnostic of power equipment. *IEEE Power Engineering Journal*. 15(1), 33-42.
- [2] Huda ASN, Taib S. Suitable features selection for monitoring thermal condition of electrical equipment using infrared thermography. *Infrared Physics & Technology* 2013;61;184-191.
- [3] Almedia, C. A. L., Braga, A. P., Nascimento, S., Paiva, V., Martins, H. J. A., Torres, R., & Caminhas, W. M. (2009). Intelligent thermographic diagnostic applied to surge arresters: a new approach. *IEEE Transactions on Power Delivery*. 24(2), 751-757.
- [4] Shafi'i, M. A. & Hamzah, N. (2010). Internal fault classification using artificial neural network. *4th International Power Engineering and Optimization Conference*, 23-24 June, Shah Alam. 352-357.
- [5] Smedberg, M. (2006). Thermographic decision support – detecting and classifying faults in infrared images. Master Thesis, Royal Institute of Technology, Stockholm, Sweden.
- [6] Wretman, D. (2006). Finding regions of Interest in a decision support system for analysis of infrared images. Master Thesis, Royal Institute of Technology, Stockholm, Sweden.
- [7] Nazmul Huda AS, Taib S, Jadin MS, Ishak D. A semi-automatic approach for thermographic inspection of electrical installations within buildings. *Energy and Buildings* 2012; 55:585-591.

- [8] Rahmani, A., Haddadnia, J., & Seryasat, O. (2010). Intelligent fault detection of electrical equipment in ground substations using thermo vision technique. Proceedings of 2nd International Asia Conference on Mechanical and Electronics Engineering, 1-3 August, Kyoto. 150-154.
- [9] Nazmul Huda AS, Taib S. Application of infrared thermography for predictive/preventive maintenance of thermal defect in electrical equipment. Applied Thermal Engineering 2013;61(2): 220-227.
- [10] Huda ASN, Taib S, Ghazali KH, Jadin MS. A new thermographic NDT for condition monitoring of electrical components using ANN with confidence level analysis. ISA Transactions (2014), <http://dx.doi.org/10.1016/j.isatra.2014.02.003i>.
- [11] D. Wang, X.-J. Zeng, J. A. Keane, An evolving-construction scheme for fuzzy system, IEEE Trans. Fuzzy Syst. 18(2000) 4(2010) 755–770.
- [12] D. Nauck, F. Klawonn, and R. Kruse, *Foundations of Neuro-Fuzzy Systems*. New York: Wiley, 1997.
- [13] S. W. Tung, C. Quek, C. Guan, SaFIN: A self-adaptive fuzzy inference network, IEEE Trans. Neural Netw. 22(2011) 1928–1940.
- [14] D. Wang, X.-J. Zeng, J. A. Keane, An output-constrained clustering approach for the identification of fuzzy systems and fuzzy granular systems, IEEE Trans. Fuzzy Syst. 19(2011) 1127–1140.
- [15] C. Mencar, C. Castiello, R. Cannone, A. M. Fanelli, Design of fuzzy rule-based classifiers with semantic cointension, Inf. Sci. 181(2011) 4361–4377.
- [16] W. Pedrycz, K.-C. Kwak, Linguistic models as a framework of user-centric system modeling, IEEE Trans. Syst. Man Cybern. A, Syst., Humans 36(2006) 727–745.
- [17] Kregg, M. A. (2004). Benefits of using infrared thermography in utility substations. Proceedings of SPIE, 12 April, Orlando, Florida. 249-257.

- [18] TBPPM, The basics of predictive / preventive maintenance (2011) [online], [Accessed 02 August 2011], http://www.cedesa.com.mx/pdf/fluke/flukeTi30_serie_nota.pdf
- [19] Epperly, R. A., Heberlein, G. E., & Eads, L. G. (1997). A tool for reliability and safety: predict and prevent equipment failures with thermography. IEEE Conference on Petroleum and Chemical Industry, 15-17 September, Banff, Alta. 59-68.
- [20] Statistics fire causes building fire investigation the fire department and rescue Malaysia, Year 2012, nationwide, from January to June (2012) [online], [Accessed 01 August 2012], Available from World Wide Web: [http://www.bomba.gov.my/fileupload/statistik/ST-PGK%20PK01\(2012\).pdf](http://www.bomba.gov.my/fileupload/statistik/ST-PGK%20PK01(2012).pdf)
- [21] Coutin, M., Plumecocq, W., Melis, S., & Audouin, L. (2012). Energy balance in a confined fire compartment to assess the heat release rate of an electrical cabinet fire. Fire Safety Journal. (52), 34–45.
- [22] Plumecocq, W., M. Coutin, M., Melis, S., & Rigollet, L. (2011). Characterization of closed-doors electrical cabinet fires in compartments. Fire Safety Journal. (46), 243–253.
- [23] Babrauskas, V. (2001) How do electrical wiring faults lead to structure ignitions, 7th international Fire & Materials conference, 22 -24 January, San Francisco, USA. 39-50.
- [24] Garnaik, S. P. Infrared Thermography : A versatile technology for condition monitoring and energy conservation (2011) [online], [Accessed 01 July 2011], Available from World Wide Web: http://www.reliabilityweb.com/art05/Infrared_thermography_a_versatile_technology.pdf
- [25] Standard for Infrared Inspection of Electrical Systems & Rotating Equipment, Infrasppection Institute. <http://www.armco-inspections.com/files/ir/Electrical%20Rotating%20Std.pdf>.2008.
- Accessed 11 June 2011
- [26] ASTM, ASTM E 1934: Standard Guide for Examining Electrical and Mechanical Equipment with Infrared Thermography, West Conshohocken, Pennsylvania, ASTM International, 2005.

- [27] A.H.A. Bakar, H.A. Illias, M.K. Othman, H. Mokhlis, Identification of failure root causes using condition based monitoring data on a 33 kV switchgear, *International Journal of Electrical Power and Energy Systems*, 47 (2013) 305–312.
- [28] Daims H., Lucker S., Wagner M.: daime, a novel image analysis program for microbial ecology and biofilm research. *Environ. Microbiol.* 8, 200-213,2006
- [29] B. Kosko, Optimal fuzzy rules cover extrema, *Int. J. Intell. Syst.* 10(1995) 249–255.
- [30] Y. S. Ding, H. Ying, S. H. Shao, Necessary conditions on minimal system configuration for general MISO mamdani fuzzy systems as universal approximators, *IEEE Trans. Syst. Man Cybern. B, Cybern.* 30(2000) 857–864.
- [31] Y. Jin, W. von Seelen, B. Sendhoff, On generating FC^3 fuzzy rule systems from data using evolution strategies, *IEEE Trans. Syst. Man Cybern. B, Cybern.* 29(1999) 829–845.
- [32] W. L. Tung C. Quek, eFSM—A novel online neural-fuzzy semantic memory model, *IEEE Trans. Neural Netw.* 21(2012) 136–157.

overexposed. The exposure denotes an amount of intensity exposition is given by [1]:

$$Exposure = \frac{1}{L} \left[\frac{\sum_{m=1}^L p(m) * m}{\sum_{m=1}^L p(m)} \right] \quad (11)$$

where L is number of gray levels of the image meanwhile $p(m)$ and m are histogram and gray level values of the image respectively. The exposure is normalized in the range of $[0, 1]$. If the value of exposure is less than 0.5 , it denotes that the image contains more underexposed region than overexposed region.

The threshold is determined to divide the image into two parts. The threshold is given by equation (12)

$$T = \theta L(1 - Exposure) \quad (12)$$

where T and θ are threshold and exposure operator respectively. The exposure operator, θ is defined to obtain optimum threshold for enhancement. The threshold, T which is in the range $[0, L-1]$ divides the gray levels into two regions which are $[0, T-1]$ for underexposed region and $[T, L-1]$ for overexposed region.

The membership function (i.e fuzzified image) is then modified to further enhance the fuzzified image.

$$\mu_{enh} = \begin{cases} \sqrt{\tau \mu(m)} & \text{for } \mu(m) < T \\ [\mu(m)]^2 & \text{for } \mu(m) \geq T \end{cases} \quad (13)$$

where τ is the enhancement factor that is used to enhance the image.

It is known that the gray levels of the image are heap near the maximum gray level and minimum gray level for overexposed and underexposed regions respectively. A power-law transformation operator is defined for the improvement of the overexposed region of the image. The intensities of the membership function in overexposed region are improved by modifying their membership function in this region.

Meanwhile the underexposed regions have the exposure values less than 0.5 and thus only need a gradual amount of saturation enhancement. The membership function is modified using saturation operator of square root as given by equation (13). Modification of saturation restores the pleasing nature for such images.

III. DATA ANALYSIS

The proposed method has been implemented on Intel Core 2 CPU 2GHz using Matlab R2010b. 100 standard images (size: 400x264) obtained from California Institute of Technology database which consist of underexposed and overexposed regions are considered as test images.

The enhanced image is analyzed in terms of its output quality and quantitative analysis such as index of fuzziness (IOF), image contrast (C) peak signal to noise ratio (PSNR) and processing time.

In addition, the performance of the proposed algorithm is compared qualitatively and quantitatively with other state of the art methods namely conventional approach of NINT [3], application of fuzzy IF-THEN rules (fuzzy rule-based) [17], fuzzy quantitative measure [10] and fuzzy local enhancement [5] are widely used in image enhancement. The techniques in those literatures are selected since they involved in enhancing image contrast in fuzzy domain. Each of method has been discussed in Introduction part.

For the subjective evaluation in terms of the image quality, the enhanced image is expected to be brighter than the original image without overenhancing overexposed region and/or underenhancing underexposed region.

Furthermore, the applied enhancement method should minimize the uncertainty in image information. Thus, *IOF* is employed to measure the degree of fuzziness (i.e uncertainty) in the image. A smaller *IOF* indicates a better performance of image enhancement. In addition, the method should not significantly amplify the noise level and thus a high value of *PSNR* is required. Furthermore, the enhanced image must obtain optimum image contrast (*C*) to distinguish between the object and the background. The contrast for enhanced image ought to be close to the contrast of the original image to attain good image quality. The *PSNR* and *C* are calculated using equations (14) to (15) respectively. The *IOF* is calculated using equation (5).

$$PSNR = 10 \log_{10} (L-1)^2 / MSE \quad (14)$$

$$C = \sqrt{\sum_{m=0}^{L-1} (m - m_{avg})^2 p(m)} \quad (15)$$

where

$$MSE = \sqrt{\frac{\sum_{i=1}^M \sum_{j=1}^N (\mu_{i,j} - m_{i,j})^2}{M * N}} \quad (16)$$

IV. RESULT AND DISCUSSIONS

The enhanced images produced by the proposed and other methods are presented in Figures 1 to 4. For the subjective qualitative analysis of processed image appearance, the test images namely 'Man 1', 'Woman', 'Man 2' and 'Man 3' are shown in these figures. The original images have poor brightness in the underexposed regions and brightness is higher in the overexposed regions.

The NINT (Figures 1(c), 2(c), 3(c) and 4(c)) and fuzzy quantitative measure (Figures 1(d), 2(d), 3(d) and 4(d)) underenhanced the original image and thus resulted in darker image as compared to the original image. The processed images by both methods tend to underenhanced the region at the center of the image. This is because the both methods processed the whole image without taking into consideration

the information (*i.e* exposure) contained in the image. Therefore it has led to the uneven and unnatural enhanced image

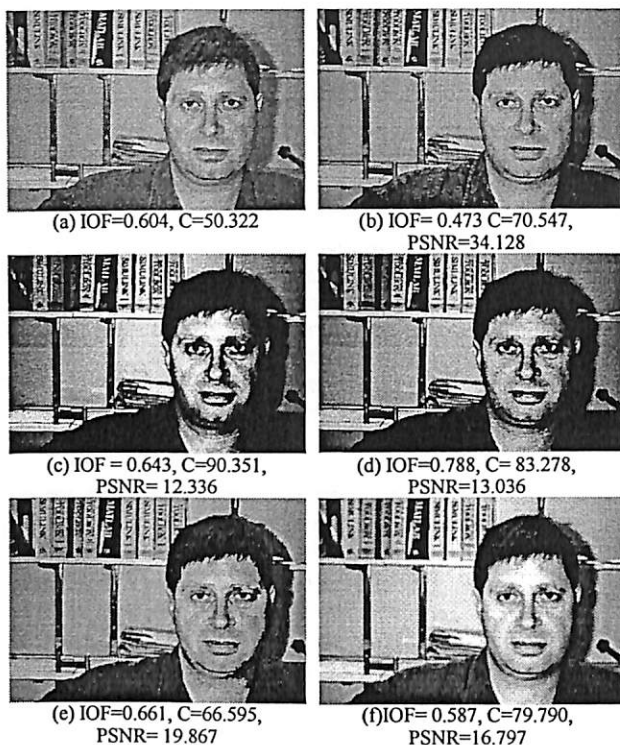


Figure 1: (a) original image (Man 1), enhanced image with (b) proposed method (c) NINT (d) quantitative measure (e) fuzzy rule-based (f) fuzzy local enhancement

Meanwhile the fuzzy rule-based and local fuzzy enhancement overenhanced the original image and thus led to the unnecessary increment in brightness around man's or woman's face.

In addition, the figures depict that the fuzzy-rule based and fuzzy local enhancement produced the overenhanced regions mostly at the center of the images. The fuzzy local enhancement has increased the contrast in local neighbourhoods in the images and revealed the fine details in the image. Furthermore, the fuzzy local enhancement able to avoid intensity saturation as compared to the fuzzy rule-based technique. The fuzzy rule-based technique caused the intensities saturations near Man 1's face and Woman's hair in Figures 1 (e) and 2(e) respectively. Thus, enhancement of those images may contain additional noises due to the unnecessary saturation.

The enhanced images (*i.e* Figures 1(b), 2(b) and 3(b)) obtained by the proposed method are more pleasant and the brightness of the image is improved accordingly with their respective regions. In the proposed method the intensities of the underexposed regions are increased by saturation operator and thus made it brighter than the original image. Meanwhile the intensities of overexposed regions are decreased by power-law transformation and resulted in

decrement in intensity values. Therefore dynamic range of enhanced image is obtained and image contrast is preserved.



Figure 2: (a) original image (Woman), enhanced image with (b) proposed method (c) NINT (d) quantitative measure (e) fuzzy rule-based (f) fuzzy local enhancement

The enhanced images by proposed method are quite similar to the enhanced images by fuzzy rule-based. However, the fuzzy rule-based has overenhanced existing overexposure region at the top corner of the original images as shown in Figures 2 (e) and 3 (e). The proposed method seems able to identify the existing overexposed region and thus reduce the pixel value by using power-law transformation.

The qualitative analysis presented in the Figures 1 to 4 can be supported by quantitative analysis presented in Table 1. The average analysis for 100 standard images of proposed method, NINT, fuzzy rule-based, fuzzy quantitative analysis and fuzzy local enhancement presented in Table 1 are discussed. For each analysis, the best results obtained are made bold.

Table 1 indicates that the proposed method has the best performances in terms of smallest *IOF*, highest *PSNR* and obtained good contrast. However, in terms of the average execution time, NINT has the fastest processing time because NINT is less complex and treated the whole image as mixed region without considering overexposed and underexposed regions.



(a) IOF=0.472, C=64.005

(b) IOF=0.341, C=81.424, PSNR=19.375



(c) IOF=0.443, C=105.632, PSNR=17.211

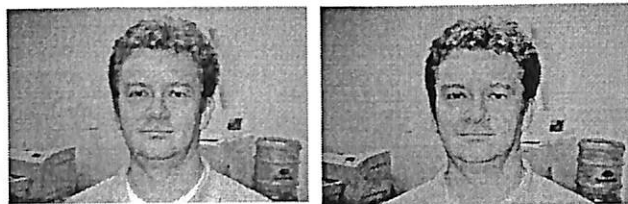
(d) IOF=0.463, C=93.788, PSNR=16.710



(e) IOF=0.398, C=77.286, PSNR=19.789

(f) IOF=0.359, C=92.565, PSNR=18.089

Figure 3: (a) original image (Man 2), enhanced image with (b) proposed method (c) NINT (d) quantitative measure (e) fuzzy rule-based (f) fuzzy local enhancement



(a) IOF=0.686, C=36.100

(b) IOF=0.503, C=48.372, PSNR=28.394



(c) IOF=0.564, C=90.541, PSNR=8.013

(d) IOF=0.675, C=76.819, PSNR=13.758



(e) IOF=0.651, C=49.310, PSNR=22.210

(f) IOF=0.511, C=53.015, PSNR=16.680

Figure 4: (a) original image (Man 4), enhanced image with (b) proposed method (c) NINT (d) quantitative measure (e) fuzzy rule-based (f) fuzzy local enhancement

The fuzzy-rule based needs the longest time to be executed since the decision can only be made if the previous condition is performed. Besides that, the fuzzy local enhancement required longer time to be executed since enhancement process is done locally using overlapping window over entire image.

The proposed method attained smallest *IOF* which indicates that the degree of fuzziness in the enhanced images is lowest as compared to the other enhanced images produced by other methods. The enhanced images by proposed method became more interpretable since the *IOF* of the enhanced images is decreased from the *IOF* of the original images (*i.e* average value of *IOF* obtained from original images = 0.461).

In addition, the proposed method achieved highest *PSNR* among other methods which concludes that the proposed method does not amplify existing noise in the original image. Even though the NINT only requires minimum time to be executed, it has increased the existing noises and artifacts possessed by the original image and thus resulted in the lowest value of *PSNR*.

TABLE I. QUANTITATIVE ENHANCEMENT ANALYSES FOR 100 STANDARD IMAGES (AVERAGE VALUES)

Method\Analysis	Processing Time, t (s)	IOF	PSNR (dB)	C
Proposed Method	0.062	0.349	22.039	71.969
NINT	0.050	0.443	13.947	88.391
Fuzzy rule-based	11.921	0.367	19.096	78.793
Fuzzy Quantitative Measure	0.063	0.410	15.417	82.654
Fuzzy Local Enhancement	11.163	0.584	19.063	81.929

Table 1 also indicates that the image contrast of the proposed method is the lowest and closer to the average contrast of original images which is 55.814. The absolute difference between the average of image contrast by proposed method and the average of those original images contrast is the smallest as compared to the other methods. This is because the produced images by proposed method are better in preserving the contrast of the original images. Thus, it can be concluded, the enhanced image from the proposed method is more natural and image intensities are not saturated.

V. CONCLUSIONS

The new enhancement technique using fuzzy set theory has been developed for grayscale non-uniform illumination image. Findings signified that the proposed method produced better image quality and defeated other methods in terms of image contrast and measure of fuzziness without enhancing existing noise in the image. The proposed algorithm only required minimum processing time (*i.e* approximately 62ms) and thus made it as suitable approach to be used in the real time application.

ACKNOWLEDGMENT

This research was supported by the Universiti Sains Malaysia under a Fundamental Research Grant Scheme, Ministry of Higher Education, Malaysia, titled "Investigation of New Color Image Illumination Estimation Concept for Development of New Color Correction Techniques"

REFERENCES

- [1] M. Hanmandlu and D. Jha, "An Optimal Fuzzy System for Color Image Enhancement," *Image Processing, IEEE Transactions on*, vol. 15, pp. 2956-2966, 2006.
- [2] C. ZhiYu, B. R. Abidi, D. L. Page, and M. A. Abidi, "Gray-level grouping (GLG): an automatic method for optimized image contrast Enhancement-part I: the basic method," *Image Processing, IEEE Transactions on*, vol. 15, pp. 2290-2302, 2006.
- [3] L. A. Zadeh, "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. SMC-3, pp. 28-44, 1973.
- [4] M. Hanmandlu, D. Jha, and R. Sharma, "Color image enhancement by fuzzy intensification," in *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, 2000, pp. 310-313 vol.3.
- [5] E. E. Kerre and M. Nachttegaal, *Fuzzy techniques in image processing*: Physica-Verlag, 2000.
- [6] H. D. Cheng and H. Xu, "A novel fuzzy logic approach to mammogram contrast enhancement," *Information Sciences*, vol. 148, pp. 167-184, 2002.
- [7] M. Hanmandlu, D. Jha, and R. Sharma, "Color image enhancement by fuzzy intensification," *Pattern Recognition Letters*, vol. 24, pp. 81-87, 2003.
- [8] S. Lu, Z. Wang, and J. Shen, "Neuro-fuzzy synergism to the intelligent system for edge detection and enhancement," *Pattern Recognition*, vol. 36, pp. 2395-2409, 2003.
- [9] P. Ensafi, H. Tizhoosh, M. Kamel, and A. Campilho, "Type-2 Fuzzy Image Enhancement Image Analysis and Recognition." vol. 3656: Springer Berlin / Heidelberg, 2005, pp. 159-166.
- [10] D.-l. Peng and A.-k. Xue, "Degraded image enhancement with applications in robot vision," in *Systems, Man and Cybernetics, 2005 IEEE International Conference on*, 2005, pp. 1837-1842 Vol. 2.
- [11] I. K. Vlachos and G. D. Sergiadis, "Parametric indices of fuzziness for automated image enhancement," *Fuzzy Sets and Systems*, vol. 157, pp. 1126-1138, 2006.
- [12] R. Vorobel, O. Berehulyak, L. Rutkowski, R. Tadeusiewicz, L. Zadeh, and J. Zurada, "Gray Image Contrast Enhancement by Optimal Fuzzy Transformation Artificial Intelligence and Soft Computing – ICAISC 2006." vol. 4029: Springer Berlin / Heidelberg, 2006, pp. 860-869.
- [13] I. Vlachos, G. Sergiadis, P. Melin, O. Castillo, L. Aguilar, J. Kacprzyk, and W. Pedrycz, "The Role of Entropy in Intuitionistic Fuzzy Contrast Enhancement Foundations of Fuzzy Logic and Soft Computing." vol. 4529: Springer Berlin / Heidelberg, 2007, pp. 104-113.
- [14] C. Florea, A. Vlaicu, M. Gordan, and B. Orza, "Adaptive Implementation of Nonlinear Fuzzy Image Enhancement Algorithms in the Compressed JPEG Images " in *Advances in Computer Science and IT: InTech*, 2009.
- [15] M. Hanmandlu, O. P. Verma, N. K. Kumar, and M. Kulkarni, "A Novel Optimal Fuzzy System for Color Image Enhancement Using Bacterial Foraging," *Instrumentation and Measurement, IEEE Transactions on*, vol. 58, pp. 2867-2879, 2009.
- [16] P. Melin, O. Mendoza, and O. Castillo, "An improved method for edge detection based on interval type-2 fuzzy logic," *Expert Systems with Applications*, vol. 37, pp. 8527-8535, 2010.
- [17] X.-Y. Wang, T. Wang, and J. Bu, "Color image segmentation using pixel wise support vector machine classification," *Pattern Recognition*, vol. 44, pp. 777-787, 2011.
- [18] K. R. Bhutani and A. Battou, "An application of fuzzy relations to image enhancement," *Pattern Recognition Letters*, vol. 16, pp. 901-909, 1995.
- [19] Y. Choi and R. Krishnapuram, "A fuzzy-rule-based image enhancement method for medical applications," in *Computer-Based Medical Systems, 1995., Proceedings of the Eighth IEEE Symposium on*, 1995, pp. 75-80.
- [20] C. Young Sik and R. Krishnapuram, "A robust approach to image enhancement based on fuzzy logic," *Image Processing, IEEE Transactions on*, vol. 6, pp. 808-825, 1997.
- [21] H. R. Tizhoosh, G. Krell, and B. Michaelis, " λ -enhancement: contrast adaptation based on optimization of image fuzziness," in *Fuzzy Systems Proceedings, 1998. IEEE World Congress on Computational Intelligence., The 1998 IEEE International Conference on*, 1998, pp. 1548-1553 vol.2.
- [22] I. K. Vlachos and G. D. Sergiadis, "Intuitionistic fuzzy information - Applications to pattern recognition," *Pattern Recognition Letters*, vol. 28, pp. 197-206, 2007.
- [23] G. Li, Y. Tong, and X. Xiao, "Adaptive Fuzzy Enhancement Algorithm of Surface Image based on Local Discrimination via Grey Entropy," *Procedia Engineering*, vol. 15, pp. 1590-1594, 2011.

Fuzzy Enhancement for Nonuniform Illumination of Microscopic Sprague Dawley Rat Sperm Image

Khairunnisa Hasikin^{1,2}

¹Department of Biomedical Engineering
Faculty of Engineering,
University of Malaya, Lembah Pantai
50603, Kuala Lumpur, Malaysia
eishahasikin@gmail.com

Nor Ashidi Mat Isa²

²Imaging and intelligent Systems Research Team (ISRT)
School of Electrical and Electronic Engineering,
Engineering Campus, Universiti Sains Malaysia
14300, Nibong Tebal, Penang, Malaysia
ashidi@eng.usm.my

Abstract— This paper presents a fuzzy grayscale enhancement technique for nonuniform illumination of microscopic Sprague Dawley rat sperm image. The microscopic images extracted from the sperm motility analysis video are low in contrast and having nonuniform lighting. Most of the developed techniques enhanced the microscopic image without considering nonuniform brightness in the image. Thus, overenhanced or underenhanced phenomena in the processed image are inevitable. The fuzzy grayscale image enhancement technique is proposed to overcome the aforementioned problems. The enhancement process of sperm images is conducted according to predetermined overexposed and underexposed regions. The proposed method has attained optimum fuzziness measures and the quality of the sperm image is improved. In addition, the proposed method required minimum processing time as compared to the other methods.

Keywords—fuzzy; enhancement; non-uniform illumination; underexposed image; overexposed image; grayscale

I. INTRODUCTION

Male infertility cases have been largely reported in recent years [1-6] and have led to myriad researches in diagnosis and prognosis of these problems. The male infertility are mostly induced by defective sperm functions which include problems in sperm motility, membrane integrity and sperm morphology.

A study of 1085 sperm samples from infertile men conducted by [2] found that 81% of the samples had defects in motility while 19% had asthenozoosperma (*i.e.* reduced in sperm motility) without any malfunction and problem in sperm number and morphology. The sperm motility is crucial during fertilization and thus assessment of sperm motility should be improved.

The sperm motility is a critical indicator of semen quality and fertility potential, because it is required for penetration of cervical mucus. Although there are widely used threshold values for normal semen measurements, published by the World Health Organization (WHO), the available norms for sperm concentration, motility, and morphology fail to meet rigorous clinical, technical, and statistical standards.

The analysis of sperm motility is conducted by recording a sequence of microscopic image frames to trace sperm

movement in a specified period of time. The sperm velocity is then calculated to distinguish between normal and abnormal sperm.

Most commonly, recorded microscopic images suffered from inadequate illumination and thus resulted in low contrast image as well as decrement in image quality. In the past twenty years much effort has been spent in developing the image enhancement algorithm for application such as robot vision, medical scanning system and video surveillance. However, the development of the enhancement algorithm specific to microscopy images has been neglected. Several enhancement algorithms are often applied successfully but they fail to fully exploit the information contained in the microscopic images. In addition, misinterpretation of enhanced image is seldom occurred since the information contained in the image is not fully extracted.

Hence, various image enhancement techniques have been proposed to improve appearance of the image. These techniques are conducted by increasing certain features and decreasing ambiguity between different regions of the image.

Some information might be loss during image recording, due to the mapping from 3-dimensional moving object (*i.e.* sperm) space to 2-dimensional image and thereby uncertainty is occurred in the image. The uncertainty exists due to the variability of gray level image since the contrast and boundary in the image are vague in nature. In addition, human understanding system is never crisp or precise.

Therefore, to defeat the uncertainty and vagueness encountered in the recorded image, the fuzzy set theory has been employed by researchers for grayscale image enhancement [7-12]. The fuzzy set theory has been widely used as a problem solving tool between the precision of classical mathematics and the inherent imprecision of the real world. The imprecision possessed by the acquired image can be perceived qualitatively by human reasoning. However, there is no specific quantification to describe the imprecision and thus machine may not understand them. Realizing this limitation to a great extent, fuzzy logic tools empower a machine to mimic human reasoning.

This project is supported by Ministry of Science, Technology & Innovation Sciencefund Grant title "Development of Computational Intelligent Infertility Detection System based on Sperm Motility Analysis"

The most popular conventional method of fuzzy contrast enhancement is using new intensification operator (NINT) [13] which has been improved from the intensification (INT) operator proposed by [14]. In this technique, a Gaussian type of fuzzification function that contains a single fuzzifier is utilized to enhance the image. The fuzzifier is obtained by maximizing the fuzzy contrast. The NINT does not change uniformly because the membership function is marginally changing, thus computational is reduced as compared to the INT.

Despite the conventional method by NINT, fuzzy rule-based has been utilized by [7, 11, 15-17] to enhance image quality. In fuzzy rule-based, human intuition is incorporated to make soft decisions on each condition. A set of conditions are defined on image pixels for enhancement purposes and these conditions will form the antecedent part of the IF-THEN rules. This approach incorporated with human intuition to make soft decisions on each condition. However, it suffers from high computational time and it is difficult to generate fuzzy rule. In addition, the consequent part of the rule will only be executed if the prior rule is accomplished and thus made it difficult to be implemented in the real time application.

In addition, image enhancement is also conducted by measuring the quantitative information in image quality such as entropy, index of fuzziness, and intuitionistic fuzzy [9, 18-23]. This method can avoid trial and error approach of selecting suitable membership function for enhancement purposes. However, this technique requires complicated optimization procedure that might be conducted repeatedly to obtain more pleasing image.

Apart from global enhancement, fuzzy local enhancement algorithm is also applied in image enhancement by [8, 12, 24, 25]. An $S \times R$ neighbourhood window is selected which is moved at each pixel location and the enhancement is performed locally until the whole image is scanned. In this technique, fine edges which are neglected in global enhancement are enhanced and clarity of the enhanced image is improved. However, noises and artifacts might also be enhanced during the enhancement process.

Since the most acquired microscopic images possessed non-uniform illumination, applying available aforementioned techniques will result in overenhanced or underenhanced regions. Up to date, only few publications address the issue of non-uniform illumination which applied for color image enhancement [26, 27]. Inspired from findings published in [14, 27], an automatic non-uniform illumination image enhancement approach is proposed for grayscale microscopic images. This paper presents the new fuzzy grayscale enhancement technique on non-uniform illumination microscopic images specifically used in sperm motility analysis.

II. THE PROPOSED TECHNIQUE

The recorded video of sperm movement is extracted to sequence of image frames, and each frame is enhanced to improve image quality and illumination. The original image of size $M \times N$ has intensity levels $m_{i,j}$ in the range of $[0 L-1]$ can be considered as a collection of fuzzy singletons in the fuzzy set notation.

$$I = \bigcup \{ \mu(m_{i,j}) \} = \{ \mu_{i,j} / m_{i,j} \} \quad (1)$$

$$i = 1, 2, \dots, M; j = 1, 2, \dots, N$$

where $\mu(m_{i,j})$ or $\mu_{i,j}/m_{i,j}$ represents the membership or grade of belonging $\mu_{i,j}$ of $m_{i,j}$ being the grayscale intensity at the (i,j) th pixel.

Most of the recorded microscopic images suffer from low contrast and the images appear nonuniform in terms of their illumination. These conditions can be seen in their histograms which do not occupy the whole dynamic range of the image. The histogram may occupy only in certain area in lower or upper parts in of the total range. When the intensity distribution is mainly accumulated in the lower part of the histogram, the image appears darker meanwhile if the intensity distribution accumulated in the upper part of the histogram, the image appears brighter. In both cases the image is considered as blurred and having nonuniform illumination. It is often encounter in the image that the brightness is uneven. Due to that, the bright and dark regions can exist in the image and thus make it challenging to linearly stretch the intensity to attain enhanced image.

In the dark region areas, which neighbourhood pixels are close to the least of the available dynamic range is often called as an underexposed region while the bright region areas that occupy in the highest of the available dynamic range is known as overexposed region. In addition there are certain areas in the image that contain both underexposed and overexposed regions in the same image

Therefore intensity exposure has been proposed by [27] to categorize the image into groups of bright and dark regions. The exposure is calculated using equation (2):

$$Exposure = \frac{1}{L} \left[\frac{\sum_{m=1}^L p(m) * m}{\sum_{m=1}^L p(m)} \right] \quad (2)$$

where L is number of gray levels of the image meanwhile $p(m)$ and m are histogram and intensity of the image respectively. The exposure is normalized in the range of $[0 1]$. The threshold is determined to separate the image into two groups namely underexposed and overexposed regions.

$$\beta = \alpha L (1 - Exposure) \quad (3)$$

where β and α are threshold and exposure operator respectively. The exposure operator, α is defined to obtain optimum threshold for enhancement in each region. The threshold, β which is in the range $[0 L-1]$ divides the gray levels into two regions which are $[0, \beta-1]$ for underexposed region and $[\beta, L-1]$ for overexposed region.

In this study, the original image is fuzzified using a modified Gaussian membership function for underexposed and overexposed regions as given by equations (4) and (5) respectively.

Underexposed region:

$$\mu_u(m) = \left| m_{\max} - \exp \left[- \frac{(m_{\max} - (m_{\text{avg}} - m))^2}{2 f_h^2} \right] \right| \text{ for } m \leq \beta \quad (4)$$

Overexposed region:

$$\mu_o(m) = \left| \exp \left[-\frac{(m_{\max} - (m_{\text{avg}} - (L - m)))^2}{2f_h^2} \right] - m_{\min} \right| \text{ for } m > \beta \quad (5)$$

where μ_u and μ_o are the membership function of underexposed and overexposed regions respectively. m_{\max} , m_{\min} and m_{avg} are the maximum, minimum and average intensities of the original image respectively. f_h is a fuzzifier and its value is calculated using equation given by:

$$f_h^2 = 2 \left[\frac{\sum_{m=0}^{L-1} [(m_{\max} - \sigma_m) - m]^4 p(m)}{\sum_{m=0}^{L-1} [(m_{\max} - \sigma_m) - m]^2 p(m)} \right] \quad (6)$$

where σ_m is standard deviation of the original image.

The calculated membership function for underexposed and overexposed regions (*i.e* in equation (4) and (5)) transformed the original image intensity levels from the spatial domain to fuzzy domain. In the fuzzy domain, the fuzzified image consists of three regions namely underexposed, mixed and overexposed regions. All gray levels below $T1$ is assumed to lie in the underexposed region and all gray levels above $T2$ is lie in the overexposed region. The remaining pixels lie in the mixed region.

The membership function (*i.e* fuzzified image) is then modified to further enhance the fuzzified image.

Underexposed region:

$$\mu_{enh} = \gamma \left[\frac{1}{1 + e^{-\mu_o(m)}} \right] \quad \text{for } \mu(m) \leq T1 \quad (7)$$

Mixed region:

$$\mu_{enh} = \mu(m) \quad \text{for } T1 < \mu(m) \leq T2 \quad (8)$$

Overexposed region:

$$\mu_{enh} = \vartheta \left[\frac{1}{1 - e^{-\mu_o(m)}} \right] \quad \text{for } \mu(m) > T2 \quad (9)$$

where

$$T1 = |\mu_{\text{avg}} - \beta_{\mu_m}| \quad (10)$$

$$T2 = |\mu_{\max} - (\mu_{\text{avg}} + \beta_{\mu_m})| \quad (11)$$

From equations (7) to (11), γ and ϑ are enhanced factors that are used to enhance the fuzzified image, μ_{\max} and μ_{avg} are maximum intensity of fuzzified image, and average intensity of fuzzified image respectively. β_{μ_m} is a threshold of fuzzified image calculated using similar equation as (3). The enhanced fuzzy membership function is then defuzzified to transform back to spatial domain.

The enhanced image is analyzed qualitatively and quantitatively and their results are compared with other state of the arts methods namely NINT operator [14], fuzzy rule-based [15], fuzzy quantitative measure [9] and fuzzy local enhancement [8].

Even though the NINT approach [14] enhanced the image in color model, they are worth to be considered since this method is designed to work best in the nonhomogeneous brightness images. Hence, we select this method in literature which improved the image brightness in nonuniform illumination image.

Meanwhile, the fuzzy rule-based [15] and the fuzzy local enhancement [8] are relevant to be compared with our proposed method since both literatures dealt with enhancement of grayscale images. The fuzzy quantitative measure [9] is also considered as comparison with our method since this literature optimizing information from fuzzy measures (*i.e* in this case, index of fuzziness has been used) to enhance the image.

As for the quantitative analysis, the enhanced image is analyzed in terms of index of fuzziness, (*IOF*), peak signal-to-noise ratio (*PSNR*), image contrast (*C*) as well as processing time.

The applied enhancement method in the sperm motility analysis should minimize the uncertainty in image information. Thus, the *IOF* is employed to measure the degree of fuzziness (*i.e* uncertainty) in the image. A smaller *IOF* indicates a better performance of image enhancement. In addition, the method should not significantly amplify the noise level and thus a high value of *PSNR* is required. Furthermore, the enhanced image must obtain optimum value of *C* to distinguish between the sperms and the background. In addition, the contrast for enhanced image ought to be close to the contrast of the original image to attain good image quality. The *IOF*, *PSNR* and *C* are calculated using equations (12) to (14).

$$IOF = \frac{2}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \min\{p_{i,j}, (1 - p_{i,j})\} \quad (12)$$

$$PSNR = 10 \log_{10} (L-1)^2 / MSE \quad (13)$$

$$C = \sqrt{\frac{\sum_{m=0}^{L-1} (m - m_{\text{avg}})^2 p(m)}{M * N}} \quad (14)$$

where

$$p_{i,j} = \text{sim} \left[\frac{\pi}{2} \times \left(1 - \frac{\mu(m)}{\mu_{\max}} \right) \right] \quad (15)$$

$$MSE = \sqrt{\frac{\sum_{i=1}^M \sum_{j=1}^N (\mu_{i,j} - m_{i,j})^2}{M * N}} \quad (16)$$

III. RESULTS AND DISCUSSION

The proposed method has been implemented on Intel Core 2 CPU 2GHz using Matlab R2010b. 100 microscopic images of sperm movement (size: 336x320) which consist of underexposed and overexposed regions are considered as test images.

Figures 1 to 4 show the comparison of enhanced images using the proposed and other methods for the 3rd, 4th, 5th and 6th microscopic sperm image frames respectively. These figures depict that the original image frames (*i.e* Figures 1(a), 2(a) 3(a) and 4 (a)) have non-uniform illumination which center of the image is brighter meanwhile image is darker at the corner.

It is observed in these figures that the enhancement process by other methods namely NINT, fuzzy quantitative measure, and fuzzy local resulted in overenhanced region at the center of the image (*i.e* existing overenhanced region). This is because the other methods processed the whole image without considering the image consists of underexposed and overexposed regions. The illumination of enhanced images from other methods became more uneven and unnatural.

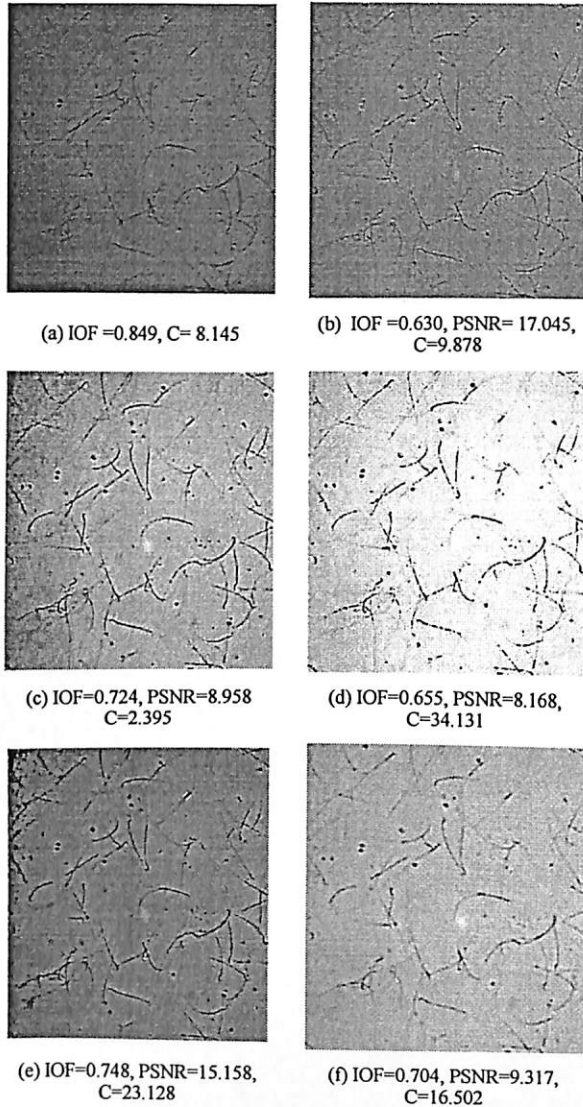


Figure 1: (a) original image (frame 3), enhanced image with (b) proposed method (c) NINT (d) fuzzy quantitative measure (e) fuzzy rule-based (f) fuzzy local enhancement

The fuzzy rule based (*i.e* Figures 1(e), 2(e), 3(e) and 4 (e)) underenhanced the original image frames and thus resulted in darker image as compared to the other methods. The intensities of the fuzzy rule based enhanced image are saturated at the corner of the image and the processed images are unnatural. Thus enhancement of sperms in the image may contain additional noises due to the unnecessary saturation. Meanwhile, the proposed method attained more pleasant image and visibility of the object of interest (*i.e* sperms) is improved than the other aforementioned methods.

The enhanced images obtained from the fuzzy local enhancement (*i.e* Figures 1(f), 2(f), 3(f) and 4 (f)) has increased

the contrast in local neighbourhood in the images and thus revealed the fine details in the image. Furthermore, the fuzzy local enhancement is able to prevent intensity saturation unlike the fuzzy rule based enhancement.

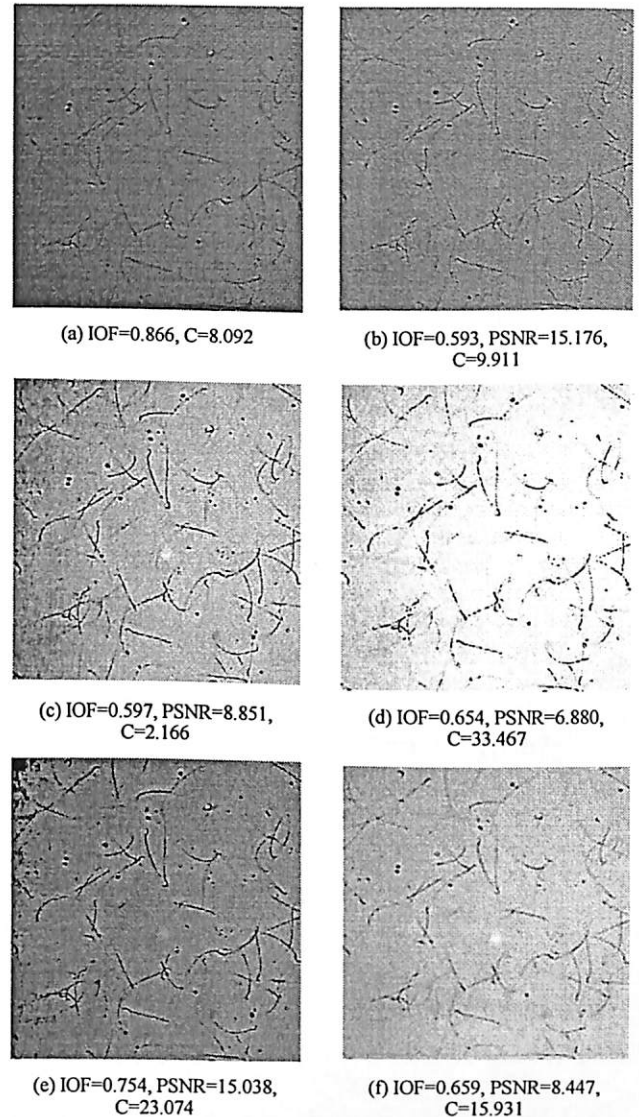


Figure 2: (a) original image (frame 4), enhanced image with (b) proposed method (c) NINT (d) fuzzy quantitative measure (e) fuzzy rule-based (f) fuzzy local enhancement

The enhanced images by the proposed method are quite similar with the enhanced images from fuzzy rule based. However, the fuzzy rule based requires longer computational time and produced saturated regions at the corner of the image.

The average quantitative analysis for 100 sperm images of the proposed method, NINT, fuzzy rule-based, fuzzy quantitative analysis and local fuzzy enhancement are tabulated in Table 1. For each analysis, the best results obtained are made bold. In terms of the average execution time, NINT has the fastest processing time because NINT is less complex and treated the whole image as mixed region without considering overexposed and underexposed regions.

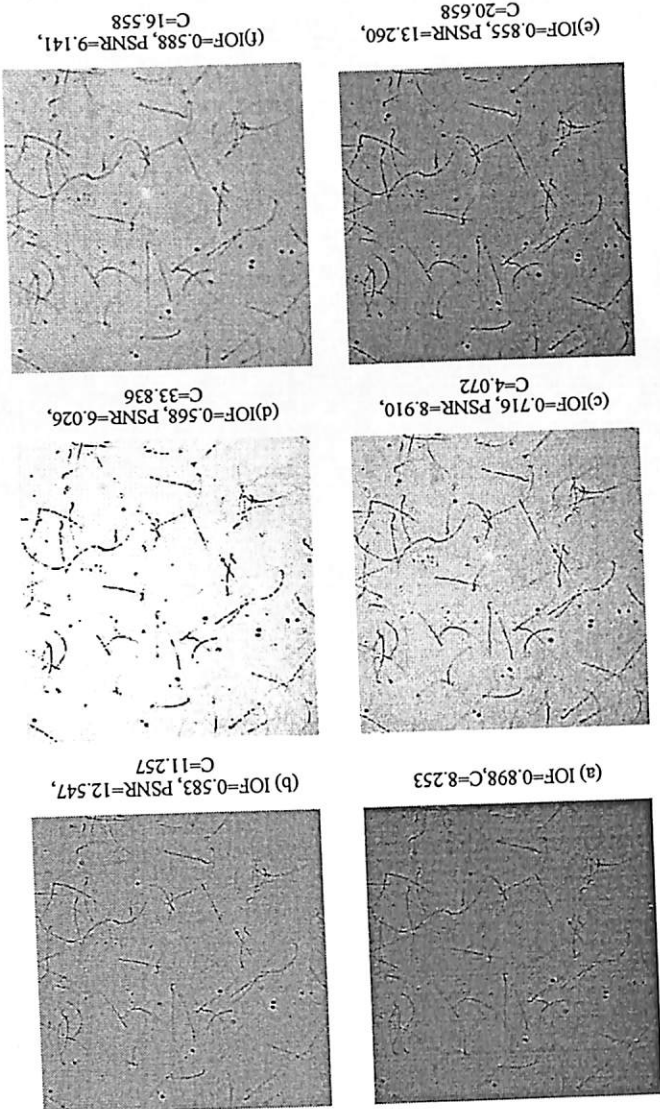
The new fuzzy grayscale enhancement method for non-uniform illumination image has been developed. Findings indicated that the proposed method produced better image quality and outperforms other methods in terms of image contrast and measure of fuzziness without enhancing existing

IV. CONCLUSIONS

Method/Analysis	Processing Time, t (s)	IOF	PSNR (dB)	
			C	PSNR
Proposed Method	0.053	0.489	17.439	9.455
NINT	0.035	0.591	8.618	2.569
Fuzzy rule-based	23.562	0.742	15.207	24.482
Fuzzy Quantitative Measure	0.065	0.587	6.616	32.595
Fuzzy Local Enhancement	11.960	0.547	8.170	15.519

TABLE I. QUANTITATIVE ENHANCEMENT ANALYSES FOR 100 SPERM IMAGES (AVERAGE VALUES)

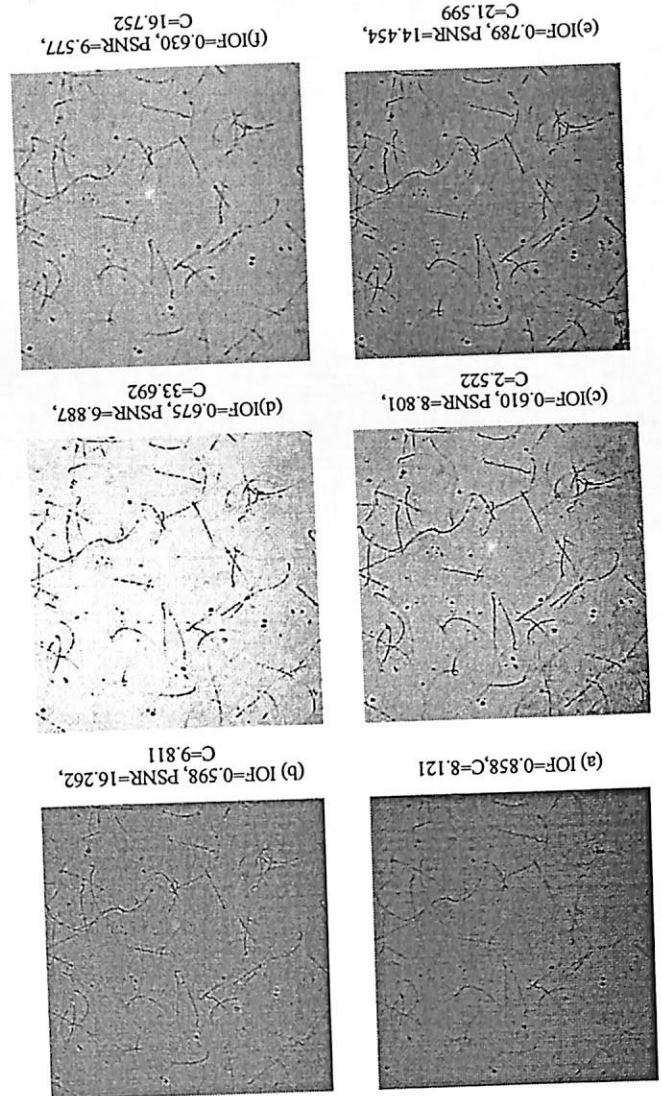
Figure 4: (a) original image (frame 6), enhanced image with (b) proposed method (c) NINT (d) fuzzy quantitative measure (e) fuzzy rule-based (f) fuzzy local enhancement



The fuzzy rule-based needs the longest time to be executed since the enhancement of each stage can only be made if the previous condition is performed. Besides that, the fuzzy local enhancement process is done locally using overlapping window over entire image. Even though the NINT only requires minimum time to enhance the image, this method has enhanced the existing noise and thus resulted in low value of PSNR as shown in Table I and Figures 1 (c), 2 (c), 3(c), and 4(c).

Table I shows that index of fuzziness of the proposed method is the smallest which indicates the performance of the proposed method is the best as compared to the other methods. In addition, the proposed method does not enhance the existing noise. Furthermore, the absolute difference between the average of image contrast by the proposed method and the average of those 100 original images contrast (i.e. 7.94233) is the smallest as compared to the other methods. This is because the produced images by the proposed method are better in preserving the contrast of the original images. Thus, it can be concluded, the enhanced image from the proposed method is more natural and image intensities are not saturated.

Figure 3: (a) original image (frame 5), enhanced image with (b) proposed method (c) NINT (d) fuzzy quantitative measure (e) fuzzy rule-based (f) fuzzy local enhancement



noise in the image. The proposed algorithm only required minimum processing time (*i.e* less than 55ms) and thus made it as suitable approach to be used in the real time sperm motility analysis.

REFERENCES

- [1] I. Sharlip, J. Jarrow, A. Belker, L. Lipshultz, M. Sigman, and A. Thomas, "Best Practice Policies for Male Infertility," *Fertil Steril.*, vol. 77, pp. 873-882, 2002.
- [2] Curi SM, Ariagno JI, Chenlo PH, Mendeluk GR, Pugliese MN, Sardi Segovia LM, Repetto HE, and B. AM., "Asthenozoospermia: Analysis of a Large Population," *Arch Androl.*, vol. 49, pp. 343-349, 2003.
- [3] Derek H.Owen and D. F.Katz, "A Review of the Physical and Chemical Properties of Human Semen and the Formulation of a Semen Simulant," *Journal of Andrology*, vol. 26, pp. 459-469, 2005.
- [4] C.R.Evans, S.Atherton, D.C.Hughes, G.McHale, and M.I.Newton, "Assessing Sperm Motility using Acoustic Plate Mode Devices " in *IEEE International Frequency Control Symposium, 2007 Joint with the 21st European Frequency and Time Forum Geneva IEEE Xplore*, 2007.
- [5] L. Z. Shi, J. M. Nascimento, C. Chandsawangbhuwana, E. L. Botvinick, and M. W. Berns, "An automatic system to study sperm motility and energetics," *Biomed Microdevices*, vol. 10, pp. 573-583, 2008.
- [6] A. Fabbrocini, M. DiStasio, and R. D'Adamo, "Computerized Sperm Motility Analysis Intoxicity Bioassays:A New Approach to Pore Water Quality Assessment," *Ecotoxicology and Environmental Safety*, vol. 73, pp. 1588-1595, 2010.
- [7] Y. Choi and R. Krishnapuram, "A fuzzy-rule-based image enhancement method for medical applications," in *Computer-Based Medical Systems, 1995., Proceedings of the Eighth IEEE Symposium on*, 1995, pp. 75-80.
- [8] H. D. Cheng and H. Xu, "A novel fuzzy logic approach to mammogram contrast enhancement," *Information Sciences*, vol. 148, pp. 167-184, 2002.
- [9] I. K. Vlachos and G. D. Sergiadis, "Parametric indices of fuzziness for automated image enhancement," *Fuzzy Sets and Systems*, vol. 157, pp. 1126-1138, 2006.
- [10] M. Banerjee and M. K. Kundu, "Handling of impreciseness in gray level corner detection using fuzzy set theoretic approach," *Applied Soft Computing*, vol. 8, pp. 1680-1691, 2008.
- [11] T. Chaira and A. K. Ray, *Fuzzy image processing and applications with MATLAB*: CRC Press/Taylor & Francis, 2010.
- [12] G. Li, Y. Tong, and X. Xiao, "Adaptive Fuzzy Enhancement Algorithm of Surface Image based on Local Discrimination via Grey Entropy," *Procedia Engineering*, vol. 15, pp. 1590-1594, 2011.
- [13] M. Hanmandlu, D. Jha, and R. Sharma, "Color image enhancement by fuzzy intensification," *Pattern Recognition Letters*, vol. 24, pp. 81-87, 2003.
- [14] M. Hanmandlu, D. Jha, and R. Sharma, "Color image enhancement by fuzzy intensification," in *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, 2000, pp. 310-313 vol.3.
- [15] K. R. Bhutani and A. Battou, "An application of fuzzy relations to image enhancement," *Pattern Recognition Letters*, vol. 16, pp. 901-909, 1995.
- [16] C. Young Sik and R. Krishnapuram, "A robust approach to image enhancement based on fuzzy logic," *Image Processing, IEEE Transactions on*, vol. 6, pp. 808-825, 1997.
- [17] M. Friedman, M. Schneider, and A. Kandel, "The use of weighted fuzzy expected value (WFEV) in fuzzy expert systems," *Fuzzy Sets and Systems*, vol. 31, pp. 37-45, 1989.
- [18] E. E. Kerre and M. Nachttegaal, *Fuzzy techniques in image processing*: Physica-Verlag, 2000.
- [19] H. R. Tizhoosh, G. Krell, and B. Michaelis, "λ-enhancement: contrast adaptation based on optimization of image fuzziness," in *Fuzzy Systems Proceedings, 1998. IEEE World Congress on Computational Intelligence., The 1998 IEEE International Conference on*, 1998, pp. 1548-1553 vol.2.
- [20] I. K. Vlachos and G. D. Sergiadis, "Intuitionistic fuzzy information - Applications to pattern recognition," *Pattern Recognition Letters*, vol. 28, pp. 197-206, 2007.
- [21] I. Vlachos, G. Sergiadis, P. Melin, O. Castillo, L. Aguilar, J. Kacprzyk, and W. Pedrycz, "The Role of Entropy in Intuitionistic Fuzzy Contrast Enhancement Foundations of Fuzzy Logic and Soft Computing." vol. 4529: Springer Berlin / Heidelberg, 2007, pp. 104-113.
- [22] S. K. Pal, "A Note on the Quantitative Measure of Image Enhancement Through Fuzziness," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. PAMI-4, pp. 204-208, 1982.
- [23] G. Nieradka, B. Butkiewicz, P. Melin, O. Castillo, L. Aguilar, J. Kacprzyk, and W. Pedrycz, "A Method for Automatic Membership Function Estimation Based on Fuzzy Measures Foundations of Fuzzy Logic and Soft Computing." vol. 4529: Springer Berlin / Heidelberg, 2007, pp. 451-460.
- [24] P. Ensafi, H. Tizhoosh, M. Kamel, and A. Campilho, "Type-2 Fuzzy Image Enhancement Image Analysis and Recognition." vol. 3656: Springer Berlin / Heidelberg, 2005, pp. 159-166.
- [25] R. Vorobel, O. Berehulyak, L. Rutkowski, R. Tadeusiewicz, L. Zadeh, and J. Zurada, "Gray Image Contrast Enhancement by Optimal Fuzzy Transformation Artificial Intelligence and Soft Computing - ICAISC 2006." vol. 4029: Springer Berlin / Heidelberg, 2006, pp. 860-869.
- [26] M. Hanmandlu, O. P. Verma, N. K. Kumar, and M. Kulkarni, "A Novel Optimal Fuzzy System for Color Image Enhancement Using Bacterial Foraging," *Instrumentation and Measurement, IEEE Transactions on*, vol. 58, pp. 2867-2879, 2009.
- [27] M. Hanmandlu and D. Jha, "An Optimal Fuzzy System for Color Image Enhancement," *Image Processing, IEEE Transactions on*, vol. 15, pp. 2956-2966, 2006.

Enhancement of the low contrast image using fuzzy set theory

Khairunnisa Hasikin

Imaging and Intelligent System Research Team (ISRT)
School of Electrical and Electronic Engineering
Engineering Campus, Universiti Sains Malaysia
14300 Nibong Tebal, Penang, Malaysia
eishahasikin@gmail.com

Nor Ashidi Mat Isa

Imaging and Intelligent System Research Team (ISRT)
School of Electrical and Electronic Engineering
Engineering Campus, Universiti Sains Malaysia
14300 Nibong Tebal, Penang, Malaysia
ashidi@eng.usm.my

Abstract—This paper presents a fuzzy grayscale enhancement technique for low contrast image. The degradation of the low contrast image is mainly caused by the inadequate lighting during image capturing and thus eventually resulted in nonuniform illumination in the image. Most of the developed contrast enhancement techniques improved image quality without considering the nonuniform lighting in the image. The fuzzy grayscale image enhancement technique is proposed by maximizing fuzzy measures contained in the image. The membership function is then modified to enhance the image by using power-law transformation and saturation operator. The qualitative and quantitative performances of the proposed method are compared with the other methods. The proposed method produced better quality enhanced image and required minimum processing time than the other methods.

Keywords—nonuniform illumination; fuzzy; grayscale; enhancement; overexposed image; underexposed image

I. INTRODUCTION

Good contrast images with preserving details are required for many important areas namely machine vision, remote sensing, dynamic and traffic scene analysis, biomedical image analysis and autonomous navigation. However most of the recorded images suffer from poor contrast which is due to the inadequate lighting during image acquiring, wrong setting of aperture size and shutter speed as well as nonlinear image intensities mapping.

Difficulties in controlling the lighting conditions during image acquisition process have resulted in variability in image illumination. The captured images turn out to be low contrast and contained underexposed and overexposed regions.

Thus, image enhancement has been employed to increase the quality of the image. Image enhancement is a fundamental task applied in image processing to improve interpretability and appearance of the image. It provides better input image for further image processing task.

Image enhancement can be clustered into two groups namely frequency domain and spatial domain methods. In the frequency domain method, the enhancement is conducted by modifying the frequency transform of the image. Meanwhile in the latter method, image pixels are directly modified to enhance the image. However, computing the enhancement in frequency domain is time consuming

process even with fast transformation technique thus made it unsuitable for real time application [1].

Numerous contrast enhancement techniques normalized the image intensities and often fail to produce satisfactory results for a broad range of non-uniform illumination image. The image is characterized by the fact that the amplitudes of their histogram components are very high at one or several locations on the grayscale, while they are very small, but not zero, in the rest of the grayscale. This makes it difficult to increase the image contrast by simply stretching its histogram. The high amplitude of the histogram components also often prevents the use of the histogram equalization (HE) techniques. Most of the HE techniques could cause a washed-out effect on the appearance of the enhanced image and/or amplify existing noises [2].

In addition, due to the poor and low contrast nature of the acquired image, vagueness and ambiguity are introduced and have led to the increment of uncertainty in the image information. This vagueness in the image appears in the form of imprecise boundaries and intensities during image digitization.

Therefore, fuzzy sets theory [3] has been proposed as a problem solving tool between the precision of classical mathematics and the inherent imprecision of the real world. The imprecision possessed by the acquired image can be perceived qualitatively by human reasoning. However, there is no specific quantification to describe the imprecision and thus machine may not understand them. Realizing this limitation to a great extent, fuzzy logic tools empower a machine to mimic human reasoning.

In the image enhancement field, the fuzzy set theory has been widely utilized by other researchers [1, 4-17]. Pixel property such as gray tone intensity is modeled into a fuzzy set using a membership function. The image is considered as an array of fuzzy singletons having a membership value that denotes the degree of belonging to specific property.

The conventional method of fuzzy enhancement is conducted by using contrast intensification (INT) operator [3]. In this method, dynamic range of image is possible to be obtained since the INT operator will increase and decrease the membership degree above and below threshold value respectively. However, the INT is solely depended on the membership function and it needs to be applied continuously on the image to attain desired enhancement.

This limitation is then improved by [7] using new intensification operator (NINT) which utilized sigmoid function. The NINT does not change uniformly because the membership function is marginally changing, thus computational time can be reduced as compared to the INT.

The other enhancement approach is conducted using fuzzy rule-based technique [18-20] which human intuition is incorporated to make soft decisions on each condition. This method suffers from high computational time and thus made it difficult to automatically generate fuzzy rules. In addition, prior threshold selection needs to be done for each condition and therefore made choosing the optimum threshold could be challenging.

Image enhancement is also done by measuring information contained in the image [11, 13, 21, 22]. The membership function is chosen based on the measured quantity such as image entropy or index of fuzziness. An optimum quantitative measure has to be determined in order to achieve best enhancement quality.

In addition, attempts have been made to enhance the image locally [6, 12, 23]. In this technique, local contrast in small regions is enhanced while at the same time preventing an increase in global contrast. Fine edges which are neglected in global enhancement are enhanced and clarity of the enhanced image is improved. However, noises and artifacts might also be enhanced during the enhancement process.

In this paper, a new contrast enhancement technique has been proposed by minimizing fuzziness in the image without requiring complex procedure and long computational time.

II. THE PROPOSED TECHNIQUE

The original image of size $M \times N$ has intensity levels m_{ij} in the range of $[0, L-1]$ can be considered as a collection of fuzzy singletons in the fuzzy set notation.

$$I = \bigcup \{ \mu(m_{i,j}) \} = \{ \mu_{i,j} / m_{i,j} \} \quad (1)$$

$i = 1, 2, \dots, M; j = 1, 2, \dots, N$

where $\mu(m_{ij})$ or μ_{ij}/m_{ij} represents the membership or grade of belonging μ_{ij} of m_{ij} being the grayscale intensity at the (i,j) th pixel.

A. Fuzzy Measures

Let I be a set with randomly gray level values $\{m_{ij}$ at (i,j) th pixel $\}$ and $\{p_0, p_1, p_2, \dots, p_{L-1}\}$ are respective probabilities of gray level values. The fuzzification of set I induced two kinds of uncertainties. The first part of uncertainties induced by the random nature of the image given by:

$$H = - \sum_{k=0}^{255} p_k \log(p_k) \quad (2)$$

Meanwhile second uncertainty arises from the fuzziness of the fuzzy set related to the ordinary set given by:

$$J = -\mu_k \log(\mu_k) - (1-\mu_k) \log(1-\mu_k) \quad (3)$$

Therefore, the total entropy, E is expressed by:

$$E = H + J \quad (4)$$

The index of fuzziness (IOF) is calculated using equation (5) as follow:

$$IOF = \frac{2}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \min\{p_{i,j}, (1-p_{i,j})\} \quad (5)$$

where

$$p_{i,j} = \text{sim} \left[\frac{\pi}{2} \times \left(1 - \frac{\mu(m)}{\mu_{\max}} \right) \right] \quad (6)$$

B. Membership Function Calculation

The shape of S-function is commonly used for the representation of the degree of brightness or whiteness of pixels in the grey levels images. The S-function was originally introduced by [2] and the definition of more flexible S-function was proposed by [23]. The flexible S-function has been adapted in this study to fuzzify the original image.

$$\mu(m) = \begin{cases} 0 & \text{for } m \leq a \\ \frac{m-a}{(b-a)(c-a)} & \text{for } a < m \leq b \\ 1 - \frac{(m-c)^2}{(c-b)(c-a)} & \text{for } b < m \leq c \\ 1 & \text{for } m \geq c \end{cases} \quad (7)$$

where m is the intensity of the image and a , b and c are parameters that determined the shape of the S-function. The parameters a , b and c are specified to ensure the membership function maximizes the information contained in the image. It is done by incorporating two fuzzy measures namely fuzzy entropy and index of fuzziness.

Therefore parameters a , b and c are given by equations (8) to (10).

$$a = \alpha E_{\max} \quad (8)$$

$$b = \beta |IOF_{\max} - E_{\max}| \quad (9)$$

$$c = \gamma IOF_{\max} \quad (10)$$

where α , β and γ are the membership factors that are chosen to obtain optimum S-membership function if fuzzified image. IOF_{\max} and E_{\max} are maximum index of fuzziness and maximum entropy respectively.

The calculated membership function transformed the image intensity levels from the spatial domain to fuzzy domain. The original image has been transformed and most of the regions in the image contained mixed region of overexposed and underexposed regions. Therefore, a parameter called 'exposure' is introduced to denote percentage of the image gray levels is underexposed and