Low-Resolution Image Enhancement Assessment

Joel C. De Goma[1], Juan Miguel Aquino[1], Christian Arcelo[1] and Zaliman Sauli[2]

[1]*School of Information Technology, Mapúa University Makati 333 Sen. Gil Puyat Ave., Makati City 1200, Philippines.*

[2]*School of Microelectronic Engineering, Universiti Malaysia Perlis, Pauh Putra Campus, 02600 Arau, Perlis, Malaysia.*

jcdegoma@mapua.edu.ph

***Abstract*—This study aims to address the problem with unrecognisable subject of low-quality images taken from standard resolution web cameras. These images may contain pixelated details, too much noise, and imbalance brightness and contrast. The authors used three algorithms such as Fuzzy Filter Based on Fuzzy Logic for noise reduction, Image Illumination based on Tone Mapping for uneven illumination and Super Resolution Algorithm to reconstruct the facial features of the low-resolution images. After undergoing experiment, results showed that the most acceptable filtering technique among three algorithms is Filtering Fuzzy Filter Based on Fuzzy Logic, Image Illumination Correction based on Tone Mapping for image illumination and with .60-.15-.15 Face Hallucination Super Resolution Parameter significantly improved the quality of face images taken from a low-resolution web camera. Also, results showed that high-resolution versions of low-resolution inputs significantly helped the reconstruction of facial features of low-resolution inputs. 86.67% improvement was recorded from the test images after the processing of images. Thus, the authors concluded that using the combination significantly improved the unprocessed images.**

***Index Terms*—Image Processing; Super Resolution; Image Correction; Filtering; Image Similarity.**

I. INTRODUCTION

Low-resolution image optimisation is an approach to enhance an image and produce an output which is of higher quality and better overall appearance. This optimisation is highly valuable and beneficial for law enforcement agencies and security institutions where identification is made based on recorded images.

In image processing, several approaches and techniques can be utilised. These consist of image processing by enhancement, image processing by reconstruction and image processing by compression. Image processing by enhancement is subdivided into two approaches, Spatial filtering method and Contrast Enhancement method. Spatial filtering method consists of a neighbourhood and a predefined operation that is performed on the image pixels encompassed by the neighbourhood [1]. Contrast enhancements improve the perceptibility of objects in the scene by enhancing the difference in brightness between objects and their backgrounds [1]. Image processing by reconstruction minimises the effect of degradations by filtering the perceived image [2]. Image restoration's effectiveness is dependent on the filter design as well as on the extent and accuracy of the knowledge of degradation process [2]. Image processing by compression involves discrete cosine transform (DCT) that converts data into sets of frequencies and simplifying of bits required and to represent an image [3].

There has been a numerous high-grade image processing software that is being used in different fields of scientific research. However, each image processing has their limitations. The "cubic spline" [4] is a "general image interpolation function, but it suffers from loss of image clarity and edge blurring". The recent attempts on developing on cubic-spline interpolation are yet to be successful. Schreiber [5] suggested a sharpened Gaussian interpolation function to lessen information spillover among pixels and enhance planeness in smooth areas. Stevenson and Schultz [6] put Bayesian's method to use for super-resolution but hypothesised the prior probability. Liu et al. [7] proposed a more structured retargeting method that performs a non-linear warp that accentuates interesting image features.

The research of Kui Jia and Shaogang Gong focused on super-resolution based on learning. When employed to the person's face, this is also widely known as "hallucination" [8]. Zisserman and Capel [9] made use of Eigenface which is chosen from training face database as representation before constraining and super-resolve low- resolution face images. Pasztor [10] approached learning-based super-resolution differently. Learning from several general training images, the authors tried to reconstruct the lost high-frequency information from low-level image primitives. Baker and Kanade [2] proposed a similar method of that with Capel and Zisserman. Based on a set of the pixel training face images by pixel using Gaussian, Laplacian and feature pyramids, they accomplished the same result. Liu and Shum [11] combined the primitive image technique of Freeman and PCA model-based approach.

In this study, the authors will try to incorporate the techniques of image noise reduction and learning based super-resolution and add another illumination algorithm to even out the radiance. A random disparity of colour or pixel information in images is called image noise. Most of the time, it is an aspect of electronic noise. These three different methods will be used together to come up with an optimisation technique. They intend to use Fuzzy Filter for noise reduction, Illumination Correction based on Tone Mapping for image illumination and lastly Hallucination by super-resolution algorithm for recovering the lost high-frequency data that occurs throughout the image formation process. Part of this study will improve existing algorithms since this paper will use different algorithms such as Fuzzy Filer, Tone Mapping based Illumination Correction and Super Resolution Algorithm. These algorithms were combined in order to test its capability to improve the quality of the images. Furthermore, this study will also determine if the combination of the used algorithms enhanced the quality of an image.

II. MATERIAL AND METHODS

This study focuses on image processing of low-resolution images. To support this work, the authors gathered two sets of images to form an image database of Filipino faces. The image database was divided into two sub-datasets which were used as face database and testing set. The face database will be composed of 200 individuals and was taken using a built-in laptop webcam with a resolution of 320x240 at 0.1 megapixels. The testing set was composed of 25 individuals having two images each; low resolution and high resolution. The low-resolution sub-dataset was taken using the same built-in laptop camera, while their high-resolution equivalent was taken using an external webcam camera with a resolution of 640x480 pixels. This dataset was further divided into two sets. Set A will be used for the experimental part, and Set B was used for the actual testing.

Subjects were asked to sit down on a chair and position their faces 12 inches from the webcam. An illustration board was placed at the back of each subject, so all the data gathered would have a standard background. The subjects were seated in front of the laptop and sat straight. Subject's face must be levelled with the level of the camera device. The distance between the subject and the camera must be 12 inches. Surrounding environment was sufficiently lit, and no external lighting device was required. This process was repeated until the database was completed. The process flow of this study is separated into three parts, namely pre-processing, the main process and return output.

A. Pre-Processing

Pre-processing is composed of three steps. The first step is normalising the image which consists of manually cropping of the face and resizing the images to 98x128 pixels Figure 1 shows the normalised image.

Figure 1: Normalized Input Image

The next step is to remove the noise in the image. For the initial step, image noise such as Gaussian noise was filtered using the proposed algorithm by Manglesh Khandelwal et al. [12] which is the Fuzzy Filter based on Fuzzy Logic. Next step is the image illumination correction based on Tone Mapping by Yadong Wu, Zhiqin Liu, Yongguo Han [13]. The image's colour space is extracted using C# application then brightness is adjusted using Tone Mapping rules. The resulting image after ongoing pre-process will be a bitmap image. Figure 2 shows the noise filtered image.

Afterwards, the images' colour space will be extracted using functions to transform and adjust image brightness accordingly. For this step, the algorithm that was used is the Illumination Correction Algorithm based on Tone Mapping. First, colour space decomposition technique was generated to calculate the luminance of a filtered image using the Equation (1).

Figure 2: Noise Filtered Image

$$L = a \times R + b \times G + c \times B \tag{1}$$

where a = 0.27, b = 0.67, c = 0.06. Then, we defined our new colour space by extracting the values of R G B from the filtered image using the Equation (2).

$$R' = R\left(\frac{1}{L}\right): \; G' = G\left(\frac{1}{L}\right): \; B' = B\left(\frac{1}{L}\right) \tag{2}$$

After calculating the luminance value of the image and getting the new colour space definition, tone mapping technology was applied to adjust the brightness of the filtered image using the equation model of tone mapping. Lastly, synthetization of the image between the brightness and colour space was performed. The brightness adjustment signal on the second process and the colour signals R′, G′, B′ which authors obtained on the first process after extracting the value of RGB from the original which was defined as the new value of RGB; will be combined. The resulting output after the conversion will be in PNG format. The converted image will then be used in the succeeding step. Figure 3 shows image illumination corrected image.

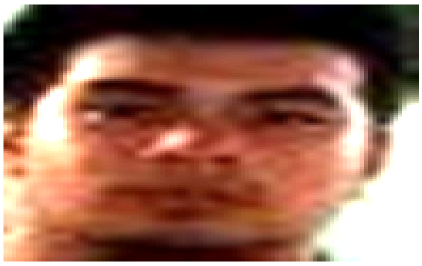

Figure 3: Image Illumination Corrected Image

B. Main processing

Main-processing comprised super-resolution process with two sub-steps which are global linear combination process and local residue compensation process. The algorithms that will be used in this phase will be from the study proposed by Xiang Ma et al. [14]

The pre-processed input image is taken as X_L. The next step is to find the highest image similarity of the input image to the database. The authors used the EyeOpen library to compare image similarities. In this process, if the output is higher than 70%, it means that the images have same structure to the input image [8]. All output that is greater than 70% will be added to a separated folder, while the highest will be XHS. For the pixels to be linearly combined, 60% will come from the low-resolution input image, 20% will come from the highest similarity image, and 20% will come from the output of linearly combined similar images higher than 70%.

In computing the similarity percentage of the images, SIFT algorithm was used. The first step is to recognise the face of the subject by using SIFT features. SIFT features are the invariant features extracted used in matching between two different images. The location of potential interest points is computed by detecting a set of Difference of Gaussian filters (DoG). The difference of Gaussian is identified by comparing a centre pixel from a 3x3 region to its 26 neighbours. After obtaining the interest points, a local feature descriptor is computed at each key point. Each of these key points has an assigned vector of features describing the distribution of local gradients in the nearest neighbourhood of a given key point [8]. The features extracted from the image are compared to the features from each image in the face database.

In comparing the features between two images, a distance between two descriptors must be computed first. After computing for the distance, Simple Graph Matching (SGM) is used in searching for the best matches for each feature vector in the query image. If two points P11 and P12 of image 1 are matched to points P21 and P22 of image 2, then the geometric relation between P11 and P12 and the one between P21 and P22 should also be similar [8]. The procedure for calculating the similarity is as follows:

$$S = 1 - \frac{1}{2\pi N}\sum_{i=1}^{N-1}\left|\left|\Delta\alpha_{i,i-1}\right| - \left|\Delta\beta_{i,i-1}\right|\right| \tag{3}$$

Some key points from the reference set are matched with some key points in the current image. This way two subsets are obtained, subset A and subset B. Each key point in the first subset has an assigned key point in the other. Matched key points are denoted with the same index in the centre. The mass centre is computed for each of the two subsets[8]. The S measure is near 1.0 if two subsets are similar regarding spatial key point distribution if the S measure is greater than 0.5 the structure of image 1 is similar to the structure of image 2 and near 0 if the spatial distributions are different. To find the XS, the proponent must get the linear combination of the similar images using this computation XS= XS1+ XS2 …XSM, where M is the reconstruction weight that must be equal to 1 and XS, is the similar images from the database. Figure 4 shows the SIFT feature image. Figure 5 shows the linear-combined output.

Figure 4: SIFT features image

Figure 5: Linear Combined Output

For the final process of step 1, the authors used the linear combination using this formula $X_{OUT1}=X_L(.6)+X_S(.2)+X_{HS}(.2)$ to obtain the output. Figure 6 shows sample 3x3 matrix value from the image

$$(.5)\underset{X_L}{\begin{pmatrix} 34 & 7 & 70 \\ 150 & 23 & 11 \\ 23 & 11 & 67 \end{pmatrix}} + (.3)\underset{X_{HS}}{\begin{pmatrix} 67 & 74 & 23 \\ 10 & 23 & 28 \\ 51 & 23 & 34 \end{pmatrix}} + (.2)\underset{X_S}{\begin{pmatrix} 65 & 77 & 97 \\ 10 & 63 & 16 \\ 111 & 23 & 27 \end{pmatrix}} = \underset{X_{OUT1}}{\begin{matrix} 50 & 41 & 61 \\ 78 & 31 & 22 \\ 49 & 17 & 49 \end{matrix}}$$

Figure 6: 3 x 3 sample matrix

In step two, the authors used Residue Compensation algorithm to patch up the residue pixels from the result of step 1. To accomplish step two, the output from step one X_{out1} will be estimated, and the final result T_{out1} is obtained using this equation $T_{out1} =(m)T_L+ (m)T_{HS}+(m)T_S$ where m is the reconstruction weight and must be equal to one. To get the T_{OUT1}, the residue of the input image T_L, highest similarity images T_{HS} and Linear combined similar feature images T_S are obtained.

First, the Residue Compensation algorithm used the equation $T_L=(.5)X_{OUT1}- (.5)X_L$ for getting the residue of the input images as shown in Figure 7, where T_L is the residue images of the input.

$$(.5)\underset{X_{OUT1}}{\begin{pmatrix} 126 & 7 & 63 \\ 150 & 22 & 11 \\ 10 & 54 & 67 \end{pmatrix}} - (.5)\underset{X_L}{\begin{pmatrix} 34 & 7 & 45 \\ 150 & 23 & 11 \\ 23 & 11 & 67 \end{pmatrix}} = \underset{T_L}{\begin{matrix} 49 & 0 & 9 \\ 0 & 1 & 0 \\ 4 & 19 & 0 \end{matrix}}$$

Figure 7: 3 x 3 sample matrix for getting the Residue of the Images

If the matrix value is greater than 0, the pixel is understood as a residue. Figure 8 shows the residue Output of T_L.

Figure 8: Residue Output of T_L

After obtaining the residue of the input image, the proponents obtained the residue of the highest similarity images T_{HS} and Linear combined similar feature images T_S. The equation for highest similarity is denoted by $T_H=(.5)X_{out1}- (.5)X_{HS}$. Figure 9 shows the output of T_H.

Figure 9: Residue Output of T_L

The equation for Similar Structure high-resolution images is denoted by $T_S=(.5)X_{out1}-(.5)X_S$ Figure 10 shows the output of T_S.

Figure 10: Residue Output of T_L

After obtaining the three required residues, the image will be linearly combined by using the equation $(.33)T_L+(.33)T_S+(.33)T_{HS}=T_{out1}$. T_{out1} represents the residue which is the output of Step 2 in the Main process. Figure 11 shows the output of Step 2.

Figure 11: Output of Step 2

For the last process, the authors obtained the output image from step one and step two. Final result X_F is obtained by adding the X_{out1} which is the step-one result and T_{out1} which is the step-two result $X_F=(.5)X_{out1}+(.5)T_{out1}$. Figure 12 shows the final output.

Figure 12: Output of Step 2

III. RESULTS AND DISCUSSIONS

The main process of the experimental part was then divided into three parts: Noise Reduction Determination, Image Illumination Correction Determination and Parameter Determination for super-resolution algorithm.

To determine which can give the best result regarding noise reduction, the authors compared three different algorithms for each phase namely: Fuzzy Filter, Gaussian filtering with Edge Preservation and Median Filter. According to the experiment conducted by the researchers, Fuzzy Filter achieved the optimum filtering technique compared to Gaussian and Median Filter because lower mean squared error (MSE), and higher peak signal to noise ratio (PSNR), denotes higher reconstruction quality and the result is shown in Table 1.

Table 1
Average MSE and PSNR of Test Images

Test Item	Noise Populated	Fuzzy Filter	Gaussian Filter	Median Filter
MSE	151.62253	27.05036	115.06407	34.87529
PSNR	26.15668	33.93775	27.52989	32.74427

Then, to determine which can give the best result regarding image illumination correction, the authors compared three different algorithms for each phase namely: Illumination based on Tone Mapping, Multi-Scale Retinex and Single Scale Retinex algorithm. According to the experiment conducted by the researchers, Illumination based on Tone Mapping achieved the excellent illumination technique compared to Multi-Scale Retinex and Single Scale Retinex algorithm because it has the highest similarity percentage to the high-resolution image equivalent. The results are shown in Table 2.

Table 2
An Illumination Similarity Average

Algorithm	Tone Mapping	MS-Retinex	SS-Retinex
Average Similarity Percentage	84.92%	84.523%	82.497%

Then, to determine which can give the best result super-resolution parameters, the authors compared three different parameters: 60%-25%-15%, 50%-10%-40% and 60%-10%-30%. The chosen parameters were based on the study of Xiang Ma et al. [14]. An experiment conducted by the researchers, 60%-25%-15% achieved the best algorithm compared to 50%-10%-40% and 60%-10%-30% because it has the highest value against the two other parameters, results are shown in Table 3.

Table 3
Parameter Similarity Average

Parameter setting	60%-25%-15%	50%-10%-40%	60%-10%-30%
Similarity Percentage	85.131%	83.609%	84.921%

After the training phase, the authors proceeded into the testing phase. In the testing phase, the authors followed the same steps but applying only three best techniques and parameters namely: Fuzzy Filter, Illumination Correction Based on Tone Mapping, and 60%-25%-15% parameter. Testing set is composed of 15 low-resolution images under random conditions considered in the study. Results are shown in Table 4. After undergoing the process, results showed that processed image produced a better similarity percentage of 85.03% over the unprocessed images having 81.87%. This was done by getting the average similarity percentage of unprocessed images and as well as processed images.

Table 4
Parameter Similarity Average

Image Type	Unprocessed Images	Processed Images
Average Similarity Rate:	81.87%	85.03%

Thus overall, this study aims to enhance the quality of low-resolution images by reducing the noise in the image, improving the illumination due to uneven lighting, finding the parameter or pixel ratio which is the most acceptable to improve the quality of the output image.

To overcome this, the authors tested several techniques and algorithms such as Fuzzy Filter, Gaussian Filter and Median filter for noise filtering, Tone Mapping, Multi-scale Retinex and Single-scale Retinex Algorithm for image illumination and Face Hallucination Technique for super-resolution.

In the experimental part of noise filtering, the fuzzy filter was chosen as the most acceptable filtering technique compared to Gaussian filter and Median Filter because of having the lowest MSE average of 27.05036 compared to 115.06407 and 34.87529 of Gaussian Filter and Median Filter (see Table 4)

The fuzzy filter has the least MSE and highest PSNR value (9 out of 10) compared to Gaussian Filter (0 out of 10) and Median Filter (1 out of 10). The lower MSE and higher PSNR denotes higher reconstruction quality. Thus, we concluded that Fuzzy Filter is the optimum filtering technique to be used among the three.

In image illumination, Tone Mapping Technique was chosen because of having the highest similarity percentage when compared to the high-quality equivalent. Having 84.925% compared to 84.523% and 82.497% of Multi-Scale Retinex and Single-Scale Retinex (see Table 4).

For the parameter of Face Hallucination, .60-.15-.15 was chosen because it yielded the highest similarity rate of 83.97% compared to 82.56% and 83.60% of the other parameters

IV. Conclusion

From this study, the experimental results showed that the most acceptable algorithms among the three enhancement techniques were: Fuzzy Filter, Tone Mapping and usage of .60-.15-.15 parameter for super-resolution technique. By combining these three image-enhancement techniques, the authors obtained a higher average similarity rate of 85.03% for the processed images compared to 81.87% of unprocessed images. The authors recorded 86.67% improvement from the test images after the processing of images. Thus, the authors concluded that using the combination significantly improved the unprocessed images.

References

[1] M. S. Nixon and A. S. Aguado, *Feature extraction & image processing for computer vision*. Academic Press, 2012.

[2] T. Morris, *Computer vision and image processing*. Palgrave Macmillan, 2004.

[3] W.-H. Chen and W. Pratt, “Scene adaptive coder,” *IEEE Trans. Commun.*, vol. 32, no. 3, pp. 225–232, 1984.

[4] R. Keys, “Cubic convolution interpolation for digital image processing,” *IEEE Trans. Acoust.*, vol. 29, no. 6, pp. 1153–1160, 1981.

[5] W. F. Schreiber and R. A. Haddad, “Fundamentals of electronic imaging systems,” *Appl. Opt.*, vol. 29, no. 19, 1990.

[6] R. R. Schultz and R. L. Stevenson, “A Bayesian approach to image expansion for improved definition,” *IEEE Trans. Image Process.*, vol. 3, no. 3, pp. 233–242, 1994.

[7] F. Liu and M. Gleicher, “Automatic image retargeting with fisheye-view warping,” in *Proceedings of the 18th annual ACM symposium on User interface software and technology*, 2005, pp. 153–162.

[8] S. Baker and T. Kanade, “Hallucinating faces,” in *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on*, 2000, pp. 83–88.

[9] D. Capel and A. Zisserman, “Super-resolution from multiple views using learnt image models,” in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, 2001, vol. 2, pp. II–II.

[10] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael, “Learning low-level vision,” *Int. J. Comput. Vis.*, vol. 40, no. 1, pp. 25–47, 2000.

[11] C. Liu, H.-Y. Shum, and C.-S. Zhang, “A two-step approach to hallucinating faces: global parametric model and local nonparametric model,” in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, 2001, vol. 1, pp. I–I.

[12] M. Khandelwal, S. Saxena, and P. Bharti, “An efficient algorithm for Image Enhancement,” *Indian J. Comput. Sci. Eng.*, vol. 2, pp. 118–123, 2005.

[13] Y. Wu, Z. Liu, Y. Han, and H. Zhang, “An image illumination correction algorithm based on tone mapping,” in *Image and Signal Processing (CISP), 2010 3rd International Congress on*, 2010, vol. 2, pp. 645–648.

[14] X. Ma, J. Zhang, and C. Qi, “An example-based two-step face Hallucination method through coefficient learning,” *Image Anal. Recognit.*, pp. 471–480, 2009.