

Video-Based Stylized Rendering Using Frame Difference

Seung Taek Ryoo[#], Yeong Kwon Kim^{\$}, Hun Joo Lee[%] and Sang Hyun Seo^{*}

[#]School of Computer Engineering, Hanshin University, Osan, Republic of Korea
E-mail: stryoo@hs.ac.kr

^{\$}Department of Computer Information, Osan University, Osan, Republic of Korea
E-mail: ykkim@osan.ac.kr

[%]Electronics and Telecommunications Research Institute (ETRI), 218 Gajeong-ro, Yuseong-gu, Daejeon, 34129, Rep. of Korea
E-mail: hjoo@etri.re.kr

^{*}Division of MediaSoftware, Sungkyul University, Anyang, Republic of Korea
E-mail: shseo75@gmail.com

Abstract— In this paper, we suggest video based stylized rendering using frame difference. Stylized rendering using video frame has a temporal problem that occurs a difference between the previous and current frame. To reduce the temporal problem, we generate reference maps using temporal frame difference in correction and rendering steps. A correction method using reference maps can be reduced flickering effect caused by frame difference between the previous and current frame. We use a background map, an average map, and a quadtree-based summed area table as reference maps. Among these reference maps, the method using quadtree based summed area table can completely remove a flickering and popping effect. Also, a post-blurring method using bilateral filtering can be represented smooth, stylized rendering by removing unnecessary noise. Suggested stylized rendering system can be used in various fields such as visual art, advertisement, game and movie for stylized image contents generation.

Keywords— stylized rendering; non-photorealistic rendering; sketch rendering; temporal coherence; frame difference

I. INTRODUCTION

Stylized rendering is a non-photorealistic rendering (NPR) method that simulates artistic techniques in rendering such as pen-and-ink, sketch, cartoon, watercolor, and painterly rendering, instead of photorealistic rendering (PR) [1]. In recent years, studies have been actively conducted to convert an original frame into a stylized frame in real time by using video streams instead of a still image. Stylized rendering methods using the video stream have problems that temporal coherence cannot be maintained due to the difference between the previous and current frame.

The temporal coherence problems in sketch rendering include both spatial and temporal aspects of the feature line such as flatness, motion coherence and temporal continuity [2]. If motion coherence and temporal continuity are satisfactory but lack flatness, the rendered scene which focused 3D appearance cannot appear similar to traditional hand-drawn animations. Keeping feature lines ensure flatness and temporal continuity but produces shower door effect [3], because the motion of feature lines has no correlation with the motion flow of the scene. Finally, if

flatness and motion coherence are maintained but temporal continuity is neglected, the animated scene produce flickering and popping effects [4], [5] since the position of feature lines varies randomly from frame to frame.

Sketch rendering algorithms can be classified into image space and object space [6]. Image space algorithms [7]-[13] use image processing technique to extract lines such as silhouettes, contours or boundaries. These extracted lines may have flickering due to lack of 3D geometry information.

Object space algorithms [14], [15] use the surface of 3D geometry models to extract feature lines. These feature lines can be easy to transform into various width and painting styles. However, lines without exact control of temporal coherence can be produced a variety of artifacts such as sliding or popping effect.

In this paper, we suggest image space based temporal coherence algorithm using pixel difference from frame to frame. To reduce flickering effect at background part, we propose original image correction method using reference maps in correction step. Also, to improve sketch rendering effect, we apply post-blurring method to the stylized image in rendering step.

This paper is organized as follows: The still image and video-based sketch rendering method are discussed in detail in Section II. The results of suggested video based sketch rendering are presented in Section III. Finally, the conclusion of this paper and further areas to study are illustrated.

II. MATERIAL AND METHOD

A. Still Image Based Sketch Rendering

As shown in Fig. 1, the first step of sketch rendering convert RGB image captured through video file or web camera to grayscale. And then, we invert the grayscale image to obtain a negative image and apply blurring method to the negative image. Finally, we combine the negative image and the blurred image using color dodge algorithm [16].

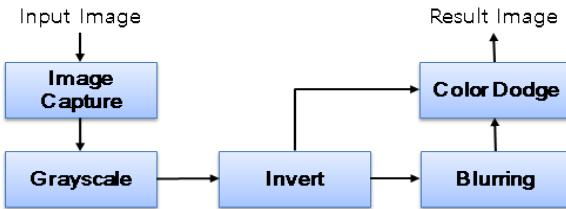


Fig. 1 The flowchart of still-image based sketch rendering

Color dodge algorithm defines as shown in Eq. 1. A means the pixel value of the grayscale image, B the pixel value of the blurred image, C is the final pixel value of the image which is applied color dodge algorithm.

$$C = 255 * B / (255 - A) \quad (1)$$

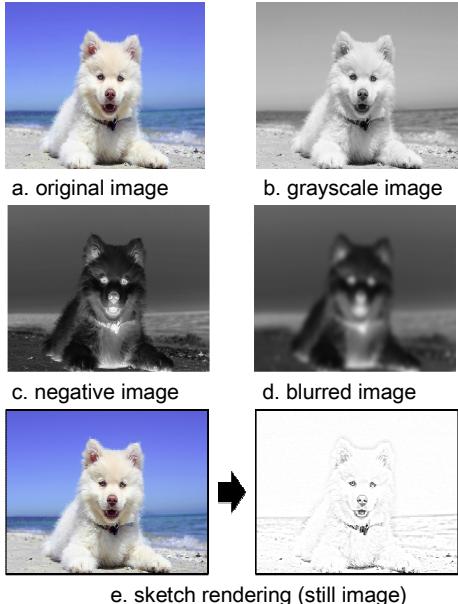


Fig. 2 The result of still-image based sketch rendering

A blurred image can be generated from Gaussian blur or bilateral filtering. The differences between Gaussian blur and bilateral filter will be explained in the next section. Fig.

2 shows the resulting image of still image sketch rendering using color dodge algorithm.

If we use a video instead of a still image, flickering and popping effect occurs in which a pixel value at a certain position is randomly changed due to the difference of color between the previous and current frame. Fig. 3 shows a flickering effect when applying still image based sketch rendering method to video. In order to reduce these problems, an algorithm for correcting an original image is required.

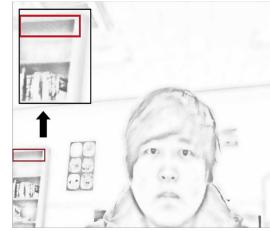


Fig. 3 The problems of temporal coherence in stylized rendering

B. Video-Based Sketch Rendering

Fig. 4 shows the overview of suggested video based sketch rendering system. To maintain temporal coherence by removing flickering effect occurring in the background, we suggest a method of correcting an original image using reference maps in the correction step and a post-blurring method in the rendering step.

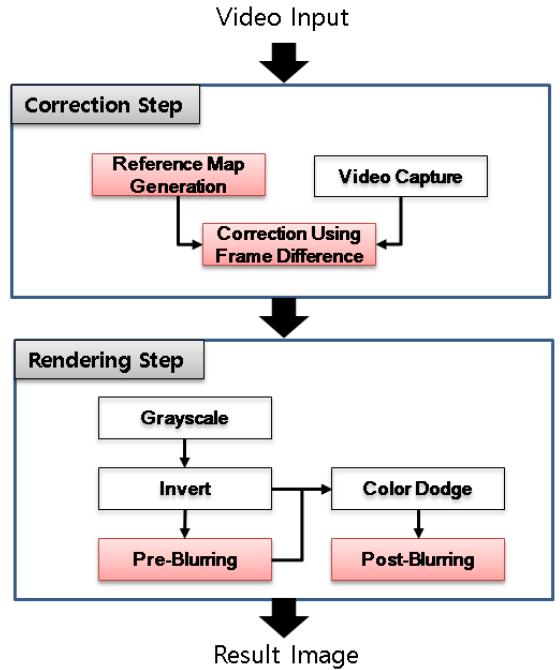


Fig. 4 The overview of suggested video based sketch rendering system.

In correction step, video frames can be input through video file or web camera, and then the original frame can be corrected by using frame difference between original image and reference map. In rendering step, we transform corrected frame into a grayscale image and invert the color. And then we blur the inverted frame using bilateral filtering and apply color dodge algorithm in the blurred frame. To reduce the flickering effect and improve sketching effect, a post-blurring method using bilateral filtering is applied again in

the rendered frame. The detailed explanation of each step will be given in the next section.

C. Correction Step

To reduce the problem of video-based sketch rendering, it is necessary to correct an original image to compensate for differences between frames. In this study, we correct an original image using reference maps such as background map, average map and quadtree based summed area table.

1) *Background Map*: Background map means an image which has removed foreground part and left background part only. Background map can be generated by capturing an image without foreground in advance or by storing only pixels having a minimum color difference between frames from a video recorded during a predetermined time [17], [18].

According to Eq. 2, a corrected image is generated using frame difference between an original image and background map.

$$D_B(x, y, t) = |I(x, y, t) - M_B(x, y)| \quad (2)$$

$$C_B(x, y, t) = \begin{cases} I(x, y, t) & i \# D_B(x, y, t) \geq th_B \\ M_B(x, y) & i \# D_B(x, y, t) < th_B \end{cases}$$

The frame difference D_B refers to absolute values obtained by subtracting corresponding pixel values from an original image I and background map M_B . C_B is corrected image using background map. If the difference between frames is smaller than the threshold value th_B , the corresponding pixel value of the corrected image is replaced with the pixel value of background map. Otherwise, the pixel value of the corrected image is replaced with an original image. Fig. 5 shows a background map, an original image and a corrected image using background map.



Fig. 5 The correction of original image using background map

2) *Average Map*: Since background map needs to capture an image without foreground in advance or record a video for a certain time, there arises a problem of generating a background map again when the viewpoint of the camera changes. To improve this problem, we suggest an average map that divides an image into specific grid areas and stores averages by area. In this study, we divide an image into 2 by 2, 4-by-4 and 8-by-8 pixel area. Also, we use average map for reference frame and average map for the current frame. The average map M_A can be calculated as shown in Eq. 3. d refers to the pixel count per coordinate axis of the divided area.

$$M_A(x, y, t) = \frac{\sum_{a=x \times d}^{(x+1) \times d} \sum_{b=y \times d}^{(y+1) \times d} I(a, b, t)}{d^2} \quad (3)$$

Corrected image using average map can be generated by calculating color difference between average map for reference frame and average map for current frame as shown in Eq. 4.

$$D_A(x, y, t) = |M_A(x, y, t) - M_A(x, y, 1)| \quad (4)$$

$$C_A(x, y, t) = \begin{cases} I(x, y, t) & i \# D_A(x, y, t) \geq th_A \\ I(x, y, 1) & i \# D_A(x, y, t) < th_A \end{cases}$$

The average map difference refers to the absolute values obtained by subtracting the corresponding pixel values of the average map for the current frame $M_A(x, y, t)$ and the average map for the reference frame $M_A(x, y, 1)$. C_A is corrected image using average map. If the difference between frames is smaller than the threshold value th_A , the corresponding pixel value of the corrected image is replaced with the pixel value of reference frame. Otherwise, the pixel value of the corrected image is replaced with the current frame, and the pixel value of the average map for reference frame is also changed the pixel value of the average map for the current frame.

Fig. 6 shows the corrected images using average map according to 2 by 2, 4-by-4 and 8-by-8 pixel areas. As shown in Fig. 6-c, when correcting the image using average map 8x8, the previous and the current frame is changed, but the pixels of the divided area are not properly reflected, resulting in a problem that the seam between the divided areas is noticeable.

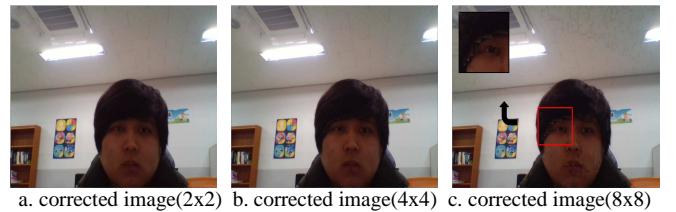


Fig. 6 The correction of original image using average map

3) *Quadtree-Based Summed Area Table*: To resolve the problem in the average map, we suggest image correction using quadtree based summed area table (QSAT). QSAT is a map that is expanded by applying summed area table (SAT) to a quadtree. First, we create an SAT M_{SAT} for reference and current frame as shown in Eq. 5.

$$M_{SAT}(a, b, t) = \sum_{x=1}^a \sum_{y=1}^b I(x, y, t) \quad (5)$$

The average value of each region is obtained by using SAT. If the average difference between corresponding regions in current and reference frame is greater than the threshold value th_S , each region is subdivided into four sub-regions. This process is recursively repeated until the average difference between corresponding regions is less than th_S or until it reaches the predefined smallest sub-region. Fig. 7-a shows the result of image segmentation by quadtree using SAT.

The corrected image using QSAT can be generated by comparing each region divided into quadtree as shown in Fig. 7-b. If the average difference of the corresponding region between current and reference frame is less than th_s , the corresponding region of the corrected image is replaced with the region for the reference frame. Otherwise, it is replaced with the region for the current frame.

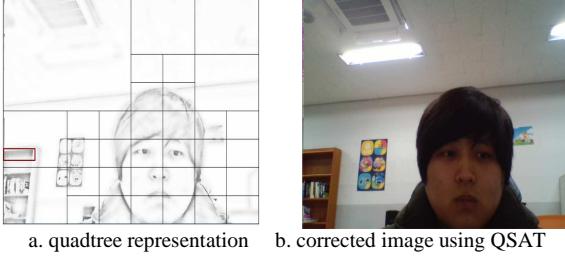


Fig. 7 The correction of original image using QSAT

D. Rendering Step

The rendering step of suggested video based sketch rendering method is similar to the conventional still image based sketch rendering method except for the blurring step [19]. To represent natural sketch effect, we subdivide the blurring step into preprocessing and postprocessing blurring. If only pre-blurring is applied, the extracted feature line does not explicitly reflect the sketch effect.

In this study, we use Gaussian blur and bilateral filter algorithm. When the sketch rendering is performed using Gaussian blur [20], the feature line is expressed correctly but the line is thin, and the object is not clearly displayed. When the bilateral filter [21] is used, the thickness of the line is properly expressed, but unnecessary noise is extracted. Fig. 8 shows a comparison of images using Gaussian blur and bilateral filter in pre-blurring.



Fig. 8 The comparison of Gaussian blur and bilateral filter.

When these blurring methods are applied to a video, the positions of unnecessarily extracted noises are randomly changed every frame. To reduce this problem, we append post-blurring method to sketch rendered image. Fig. 9 shows the results obtained by using Gaussian blurring and bilateral filtering in pre-blurring and post-blurring, respectively.

Fig. 9-a shows the result of applying a Gaussian blur to pre-blurring and bilateral filter to post-blurring. Noises have been removed, but the image is much blurred, so foreground objects and background are not clearly represented. As in the previous case, when Gaussian blur is applied to pre-blurring and post-blurring, the image is very blurry (Fig. 9-b). When applying a bilateral filter to pre-blurring and applying Gaussian blurring to post-blurring, the unnecessary noises remain unremoved as shown in Fig. 9-c. Finally, when the

bilateral filter is used for pre-blurring and post-blurring, the noise is removed, and the resulting image is clearly displayed (Fig. 9-d). In this study, we used bilateral filtering method in pre-blurring and post-blurring because it maximizes sketch effect.

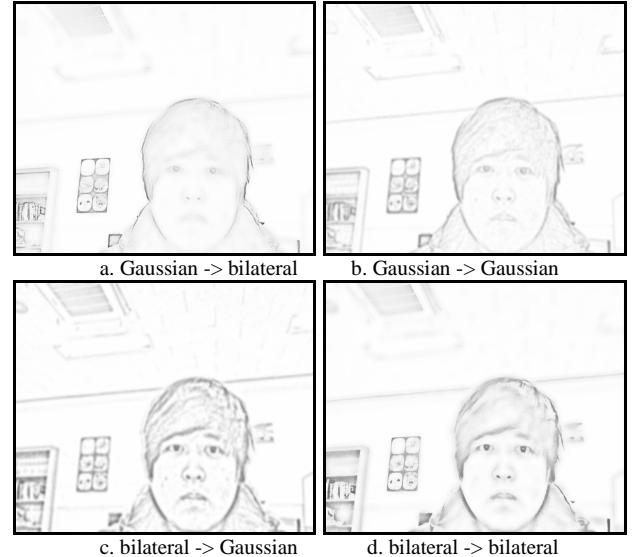


Fig. 9 The comparison of image using pre-blurring and post-blurring

III. RESULTS AND DISCUSSION

Fig. 10 shows the resulting image of applying the original image, the corrected image using the background map, the corrected image using the average map, and the corrected image using QSAT.

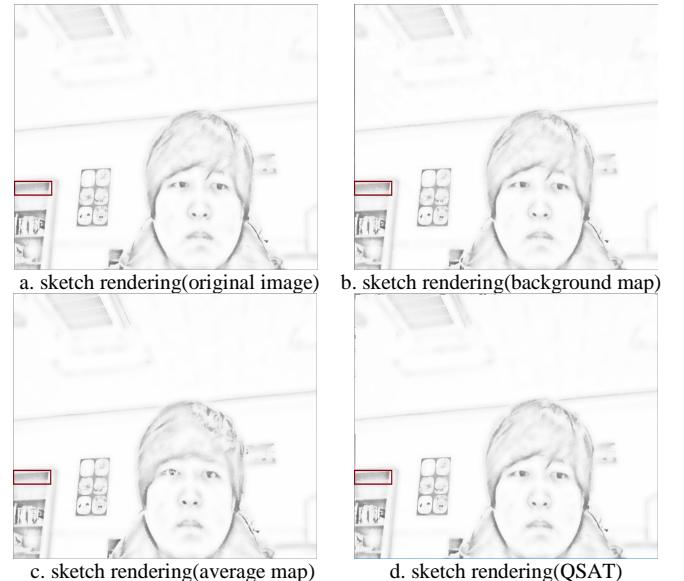


Fig. 10 The comparison of the resulting image of suggested video based sketch rendering

Table 1 shows the ratio of changed pixels by comparing the corresponding pixel difference of background part (the red box area as shown in Figure 10) between neighboring frames. F_i means a final sketch rendered image for the current frame, F_{i-1} for the previous frame and F_{i+1} for next frame. When the original image is used as it is, the ratio of changed pixels in the neighboring frames is highest. The

animated scene produces flickering and popping effects. When the corrected image using background map is applied, the ratio of changed pixels is slightly reduced. Also, when the corrected image using average map is applied, the ratio is improved more than when background map is used. Finally, when applying the corrected image using QSAT, there is no changed pixel in the background part between neighboring frames. The suggested QSAT method can be removed a flickering and popping effect.

TABLE I
THE COMPARISON OF PIXEL DIFFERENCE BETWEEN FRAMES

Frames	The ratio of changed pixels in selected region (Changed pixels/ Total pixels)			
	Original Image	Background Map	Average Map (4x4)	QSAT
$ F_{i-1} - F_i $	 76.8%	 26.4%	 15.4%	 0%
$ F_i - F_{i+1} $	 83.5%	 62.9%	 37.4%	 0%

Fig. 11 shows the resulting image using video-based sketch rendering. It can be seen that our suggested sketch rendering method is well applied not only in indoor but also outdoor scene.

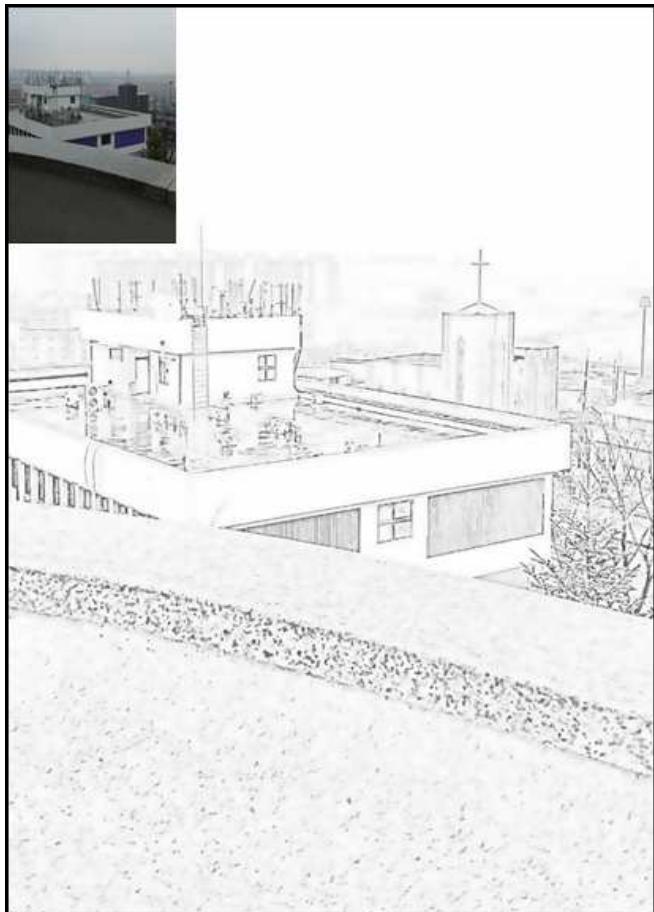


Fig. 11 The result of suggested video based sketch rendering

IV. CONCLUSION

In this paper, we suggest original image correction method using reference maps such as background map, average map, and QSAT to reduce flickering effect at background part in correction step. Also, we append post blurring method using bilateral filtering to reduce the flickering effect and improve sketch effect in rendering step. Suggested rendering system can be used in various fields such as visual art, advertisement, game and movie for stylized image contents generation.

When a corrected image is generated by using QSAT, it sometimes appears that the boundary of the divided area is not smoothly connected. To remove this problem, we need to improve the proposed algorithm in future work.

ACKNOWLEDGMENT

This research was supported by Hanshin University Research Grant, and by Ministry of Culture, Sports and Tourism (MCST) and Korea Creative Content Agency (KOCCA) in the Culture Technology (CT) and Research Development Program 2015.

Direct questions and comments about this article to the corresponding author; Sanghyun Seo, Sungkyul University, shseo75@gmail.com.

REFERENCES

- [1] Bruce Gooch, and Amy Gooch, *Non Photorealistic Rendering*, A K Peters, Ltd. 2001
- [2] Bénard, Pierre, Adrien Bousseau, and Joëlle Thollot. "State-of-the-Art Report on Temporal Coherence for Stylized Animations", Computer Graphics Forum, Vol. 30, No. 8, pp. 2367-2386, 2011.
- [3] Barbara J. Meier, "Painterly rendering for animation", Proceedings of the 23rd annual conference on Computer graphics and interactive techniques, pp. 477-484, 1996.
- [4] Steven Yantis and John Jonides, "Abrupt visual onsets and selective attention: evidence from visual search", Journal of Experimental Psychology: Human perception and performance, Vol 10, No. 5, pp 601, 1984.
- [5] Michael Schwarz and Marc Stamminger. "On predicting visual popping in dynamic scenes", Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization, pp. 93-100, 2009.
- [6] Lou, Liming, Lu Wang, and Xiangyu Meng. "Stylized strokes for coherent line drawings", Computational Visual Media, Vol. 1, No. 1, pp. 79-89, 2015.
- [7] M. McGuire, C. Ekanayake, J. St-Amour, H. Halen, A. Thibault and B. Martel, Stylized rendering in games. In SIGGRAPH 2010 Course, 2010.
- [8] T. Saito and T. Takahashi, Comprehensible rendering of 3-D shapes. ACM SIGGRAPH Computer Graphics, Vol. 24, No. 4, pp. 197-206, 1990.
- [9] Nienhaus, M.; Dollner, J. Blueprint rendering and sketchy drawings, In GPU Gems 2, Addison-Wesley, pp. 235-252, 2005.
- [10] R. Vergne, D. Vanderhaeghe, J. Chen, P. Barla, X. Granier and C. Schlick, "Implicit brushes for stylized line-based rendering", Computer Graphics Forum, Vol. 30, No. 2, pp. 513-522, 2011
- [11] Real-time vehicle tracking mechanism with license plate recognition from road images, JOURNAL OF SUPERCOMPUTING, 7, 2013.
- [12] Doug DeCarlo, "Depicting 3D shape using lines", IS&T/SPIE Electronic Imaging. International Society for Optics and Photonics, 2012.
- [13] Aniza Othman, Tengku Siti Meriam Tengku Wook, Shereena M. Arif, "Quantization Selection of Colour Histogram Bins to Categorize the Colour Appearance of Landscape Paintings for Image Retrieval", International Journal on Advanced Science, Engineering and Information Technology, Vol. 6, No. 6, pp. 930-936, 2016

- [14] Jufriadiif Na`am, Johan Harlan, Sarifuddin Madenda, Eri Prasetyo Wibowo, "The Algorithm of Image Edge Detection on Panoramic Dental X-Ray using Multiple Morphological Gradient (mMG) Method", International Journal on Advanced Science, Engineering and Information Technology, Vol. 6, No. 6, pp. 1012~1018, 2016
- [15] Pierre Bénard et al, "Active strokes: coherent line stylization for animated 3D models." Proceedings of the symposium on non-photorealistic animation and rendering, pp. 37-46, 2012.
- [16] Scott Valentine, *The hidden power of blend modes in Adobe Photoshop*, Adobe Press, 2012.
- [17] Wei Zeng, Wen Gao and Debin Zhao, Automatic Moving Object Extraction in MPEG Video, IEEE Transaction on Signals and Image processing, pp. 524-527, 2003.
- [18] Intan Kartika and Shahrizat Shaik Mohamed, "Frame Differencing with Post Processing Techniques for Moving Object Detection in Outdoor Environment", IEEE 7th International Colloquium on Signal Processing and its Applications, pp. 172-176, 2011.
- [19] David BL Bong and Adeline SL Ng, "Full-Reference and No-reference Image Blur Assessment Based on Edge Information", International Journal on Advanced Science, Engineering and Information Technology, Vol. 2, No.1, pp.90-95, 2012.
- [20] R. A. Haddad, and A. N. Akansu, "A class of fast gaussian binomial filters for speech and image processing", IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. 39, pp. 723-727, Mar. 1991.
- [21] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images", IEEE Transactions on Computer Vision, pp. 839–846, 1998