A Pre-Processing Approach for Efficient Feature Matching Process in Extreme Illumination Scenario

Marizuana Mat Daud[1], Zulaikha Kadim[1], Shang Li Yuen[1], Hon Hock Woon[1] , Ibrahima Faye[2] and Aamir Saeed Malik[2]

[1] MIMOS Berhad, Bukit Jalil, Malaysia

e-mail: marizuana.daud@mimos.my, zulaikha.kadim@mimos.my, yuen.shang@mimos.my, hockwoon.hon@mimos.my

[2] University Technology Petronas, Tronoh, Malaysia

e-mail: ibrahima_faye@petronas.com.my, aamir_saeed@petronas.com.my

***Abstract*— Video or image enhancement is a crucial part in image processing field as it improves the quality of the image before any further processes is applied on the image, which includes feature matching. In this paper, the accuracy of SURF feature descriptors used in feature matching between two input images of extreme illumination levels are evaluated. Based on the evaluation results, a novel pre-processing method to equalize both images intensity with respect to each other while maintaining the image content is proposed. We do so by fusing the cumulative histogram of the input images to compute a new cumulative histogram that will be used to remap both images. From this simple method, the results show that the intensity levels of the images are equalized and accuracy of the feature matching process is improved, in the event of extreme illumination scenario.**

***Keywords*— image enhancement, extreme illumination scenario, cumulative histogram, fusion graph, feature matching.**

I. Introduction

Video enhancement is one of the most important elements especially in surveillance application because surveillance systems monitor the scene from day to night in indoor or outdoor environment. Image enhancement definition in this context is to process a given image so that the resulting image is more suitable than the original as the input to the existing feature matching algorithm. The goal of enhancement is to improve quality of image for further processing such as segmentation, image analysis and feature matching as our intention. However, the performances of most surveillance cameras are not satisfactory at low light or high contrast environments. Low light generates noisy video images, while bright light produces lost of details.

A widely-used approach for finding corresponding points between two images is to detect edges and intensity. These features respond strongly to the corners and intensity gradients. Therefore, different intensity generates different feature keypoints and different descriptors even if it is the same scene. This sort of scheme works well for small changes illumination but will fail if there are large changes between the images. Thus, video enhancement is needed to equalize the illumination that would generate the same features.

In this work, we present a method for enhancement of images that are extremely dark and extremely bright by fusing the cumulative graph of these two images. The output will be input to the existing feature matching algorithm resulting a good match for forward processing such as medical imaging system or surveillance system. In our case, SURF will be used as the descriptor for existing feature matching algorithm.

The remainder of the paper is organized as it follows. Section 2 describes the related works which have been done in the field of image processing. In section 3, the proposed pre-process method for normalizing images is presented. In the same section, the relation between feature matching and the proposed method is also discussed. The experimental results are presented in section 4 and the conclusion is given in Section 5.

II. Related Works

A number of video enhancement methods were proposed in the literature. The most popular enhancement method is histogram equalization [1-3]. The foremost objective of this method is to achieve a uniform distributed histogram by using cumulative density function of the input image. Intensity transformations are among the simplest of all image processing techniques. This type of processing is particularly suited for enhancing white or gray detail embedded in dark regions of an image, especially when the black areas are dominant in size [4]. Though this global approach is suitable for overall enhancement, it fails to adapt if the contrast of images is not uniform, which is common scenario of surveillance images whereby the camera is installed outdoor. Some part of the image appear darker or brighter compared to others. For that reason, there are some local processing approach introduced in the literature, for example adaptive histogram equalization [2]. However, it differs from histogram equalization in the respect that it computes several local histograms and use them to improve the contrast of the overall image.

Yunbo Rao et al. [5] present a technique to fuse a video frames from high quality day-time and night-time backgrounds with low quality night-time video frames. They fused the quality videos by using weighted-average image-fusion algorithm and enhance the low quality night-time video using tone-mapping to separate an image into details and large scale features. Jing Li et al. [6] proposed a method that combining the extracting result of moving and light area, a multi-resolution based fusion to get final enhancement result. They apply illumination histogram equalization to achieve real-time and accurate moving objects segmentation. They present a real time highlight area segmentation algorithm to extract meaningful context which enhance low quality night videos. Melkamu H. Asmare et al. [7] present an approach that combines the relevant features of the input images which are decomposed by using contourlet transform and produce a composite image that provides a better representation than the conventional transforms.

978-1-4673-2584-4/12/$31.00 ©2012 IEEE

III. PROPOSED METHOD

A. Histogram and cumulative histogram revisit

Let x be an image of dimension *m rows* by *n columns* and each pixel has an intensity in the range of 0 to *L-1*, where L is the number of possible intensities in the image, often 256. The histogram (h) of image x can be defined as:

$$h_x(j) = h(x = j) \quad (1)$$

where j=0,1,...,L-1. Then cumulative histogram (H) of x will then be defined as the cumulative count of number of pixels in all histogram bins up to the specified bin, mathematically expressed as below:

$$H(i) = \sum_{j}^{i} h(j) \quad (2)$$

where i = 0,1,...,L-1 and h is the histogram of the image.

B. Proposed pre-processing method

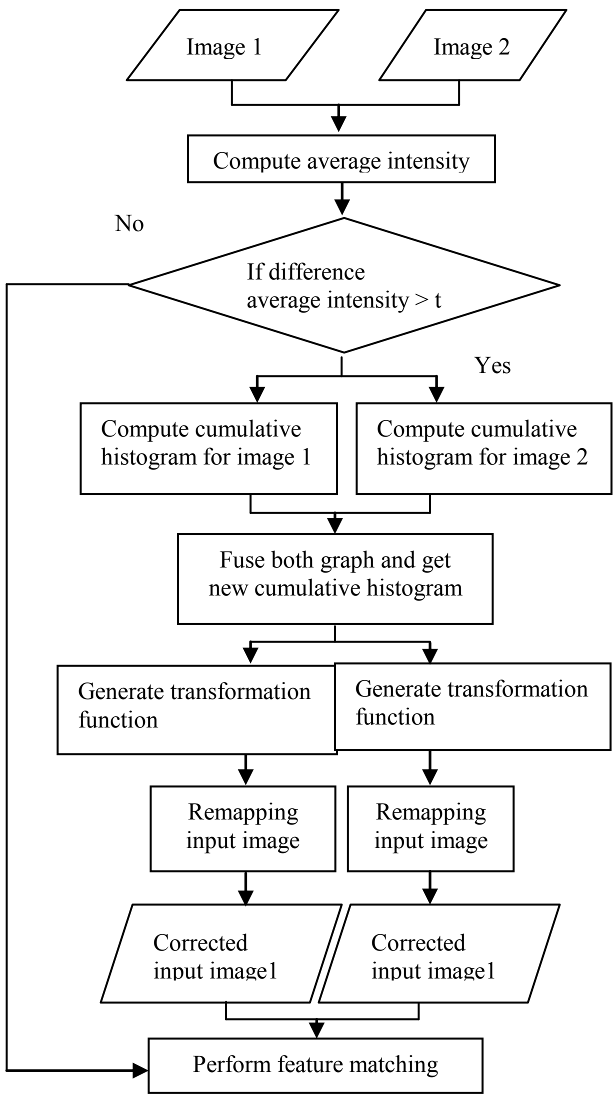


Fig. 1. The overall process flow of the proposed method

The overall process flow is shown in Fig. 1. Referring to Fig. 1, initially, average intensity of both images will be computed. If the difference between both averages is more than threshold (t), then the intensity normalization will take place as follows.

First, a new target cumulative histogram will be computed by taking the midpoint for each bin values as:

$$H_{new}(i) = \arg\min_{i}(H_1(i), H_2(i)) + \frac{|H_1(i) - H_2(i)|}{2} \quad (3)$$

Then, the transformation function is computed by mapping the original cumulative histogram to the new one. The transformation function is implemented using a lookup table. Algorithm for the mapping process is shown in Table 1.

Table 1. Algorithm for generating transformation function

Given: two cumulative histogram of an image x, find the transformation function, *f*.

For each bin *i=0:L-1*
 Get the value of *H(i)*
 Get the intercepting point on H_{new} with the horizontal line of *y=H(i)*
 Get the corresponding bin index, let say j, of the intercepting point found in previous step
 Update the transformation function, *f(i)=j*
End

Then the final corrected image will be computed by remapping the old intensity values in original image to the new ones using the transformation function obtained. The remapping for every intensity value at pixel location *(x,y)* can be expressed as :

$$g(x, y) = f(h(x, y)) \quad (4)$$

where *g(x,y)* is the remapped image at pixel location *(x,y)*, *f()* is the mapping function and *h(x,y)* is the input image at pixel location *(x,y)*. The corrected images will then be the input to the feature point extraction and matching using SURF descriptors. Next the feature point extraction and the matching processes implemented in this research will be discussed.

C. SURF feature point extraction and descriptor revisit

Corrected image 1 and image 2 produced after pre-process stage and perform feature matching using SURF. SURF (Speeded Up Robust Features) is a robust image descriptor. SURF is adaptation of the SIFT descriptor, this algorithm capable of finding correspondence between images but in a fast way due to integral images. The first step is to convert color images to grayscale images. The next step is to localize the interest points. The SURF point detector is based on the Hessian matrix. Given a point **x** = [x,y] in an image I, the Hessian matrix **H**(x,σ) in **x** at scale **σ** is defined as follows

$$H(x,\sigma) = \begin{bmatrix} L_{xx}(x,\sigma) & L_{xy}(x,\sigma) \\ L_{yx}(x,\sigma) & L_{yy}(x,\sigma) \end{bmatrix} \quad (5)$$

where L_{xx} (x, σ) is the convolution of the Gaussian second order derivative $\frac{\partial^2}{\partial x^2} g(\sigma)$ with the image I in point **x,** and similarly for $L_{xy}(x,\sigma)$ and $L_{yy}(x,\sigma)$. The convolution is to approximates second order Gaussian derivatives with simple box filter as shown in Fig. 2 [8].

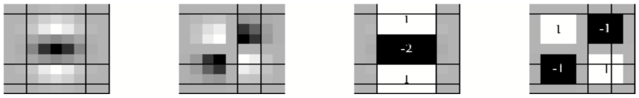


Fig. 2. Gaussian second order partial derivatives and corresponding box filter [10].

The purpose of last step is to build a descriptor that is invariant to view-point changes of the local neighborhood of the point interest. To achieve rotation invariance, a dominant orientation distribution, estimated Haar wavelets. Making use of a spatial localization grid, a 64-dimensional descriptor is then built, corresponding to a local histogram of the Haar wavelet response [9].

D. Feature matching

From the previous steps, a pair of images to be matched is represented by two sets of interest points with their corresponding SURF descriptors. The matching algorithm employed in this research is based on the k nearest neighbor matching. In our case, we set the k equal to 1. First we find the points matches by considering one image as the query and the other as the train image. Then we exchange the query and the train image and find the matches. Only matched point pair that appear in both matching will be considered as a valid match. From this exercise, the result is the list of matching point pairs between first and second image. From these correspondences, a homography matrix can be defined to relate both input images. To further confirm the matching result, a simple evaluation is performed on the matched keypoints, whereby the transformation is performed on the keypoints of the first image. The projected keypoints then will be compared with the matched keypoints in second image. If the norm between projected keypoint and its matched keypoint is less than a threshold, then they are considered as a valid match.

IV. Results

In this section, experiments are conducted to evaluate the SURF descriptor's feature matching accuracy in extreme illumination scenarios without any pre-processing, and with pre-processing using adaptive histogram equalization method and proposed method. Data used for this evaluation is taken from our outdoor camera and its parameters are changed to simulate extreme illumination condition. Also, data from [10]-[13] has been evaluated to prove the capability of the proposed algorithm. The brightness and contrast has been change to simulate extreme illumination scenarios. The results denoted as artificial data as shown in Fig. 7. Fig. 3 shows sample of input images for feature matching, exhibits extreme illumination condition and the results of feature matching without image pre-processing are shown in Fig. 4.

Fig. 5 shows results of proposed pre-processing method on input images of Fig. 4. The pre-processing method successfully equalized the intensity level of both input images with respect to each other and consequently enabled to increase number of correct matching.

In Fig. 6, some results of feature matching performed without pre-processing, with proposed method and adaptive histogram equalization are shown. For this evaluation, the orientation of input images are the same, but the images are extremely illuminated. From the figure, it shows that more feature points can be extracted from pre-processed images; consecutively more number of matching pairs is generated as compared to images without pre-processing step (Fig. 6(a)). Comparing the proposed method (Fig. 6(c)) with the histogram equalization method (Fig. 6(b)), shows that both can generate more matching pairs. However, the proposed method can generate higher percentage matches.

Fig. 3. Sample of input images with extreme illumination condition

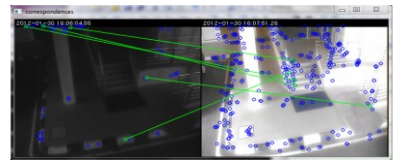
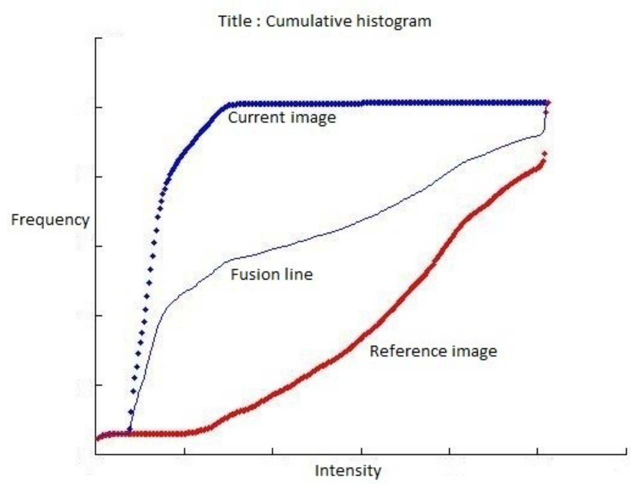
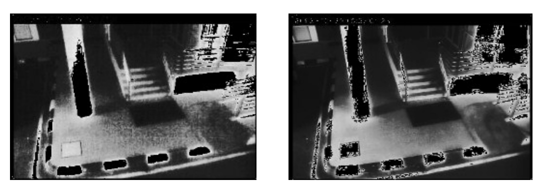


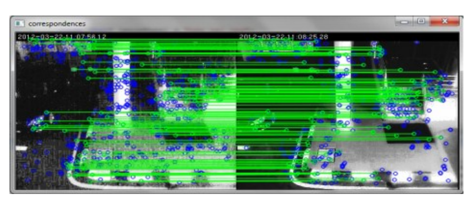
Fig. 4. Results of feature matching using SURF descriptors in extreme illumination condition. Blue circles and green lines indicate keypoints extracted and matching keypoints respectively.



(a)



(b)



(c)

Fig. 5. Results of pre-processing method applied on input images in Fig. 4; (a) cumulative histogram of input images and new equalized cumulative histogram; (b) resultant images after applying proposed pre-processing method; (c) feature matching result on pre-processed images.

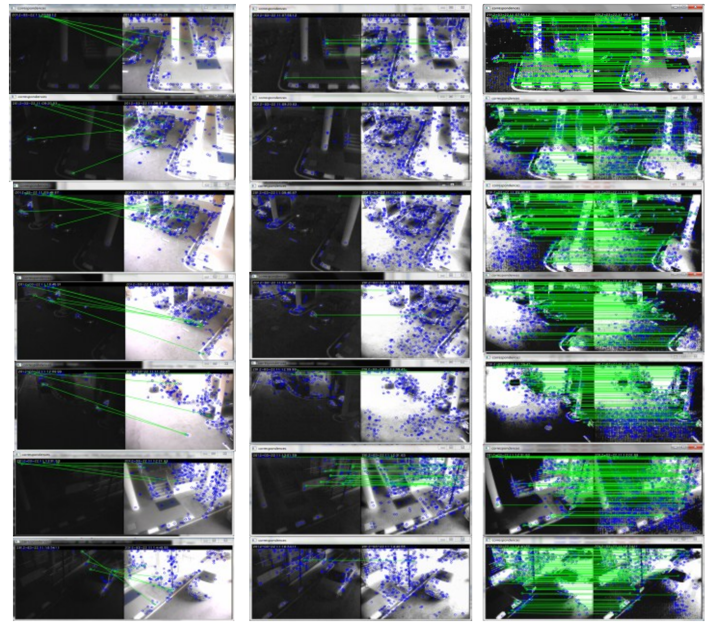
To further evaluate the accuracy of feature matching in extreme illumination condition, a modified quantitative evaluation as described in Fig. 6 is utilized. It is based on the number of correct matches and the number of false matches. The matching pair is obtained using method described in section 3.4. Datasets used for the evaluation is taken from the same camera position, thus the homography between any two images for feature matching is assumed to be an identity matrix. Thus, to characterize whether a match is correct or not, overlap error is computed between all matching points pair. The overlap error measures how well the regions correspond under a transformation, which in our case is an identity matrix homography. The overlap is the ratio between the intercept regions and the union regions of matched descriptors. Then the error is computed as: 1- overlap ratio. If the error is more than 0.5, it is characterized as false matching and otherwise.

Table 2. Quantitative evaluation for comparison between proposed method and adaptive histogram equalization for Fig. 6 (data from local camera).

No.	Adaptive histogram equalization			Proposed method		
	Total of correspondence	No. of correct matches	% correct matches	Total of correspondence	No. of correct matches	% correct matches
1	29	6	17.14	206	141	68.45
2	36	1	2.70	278	137	49.28
3	58	1	1.70	283	166	58.66
4	62	3	4.62	221	120	54.30
5	60	3	4.76	216	112	51.85
6	28	20	41.18	232	125	53.88
7	78	1	1.27	233	130	55.79

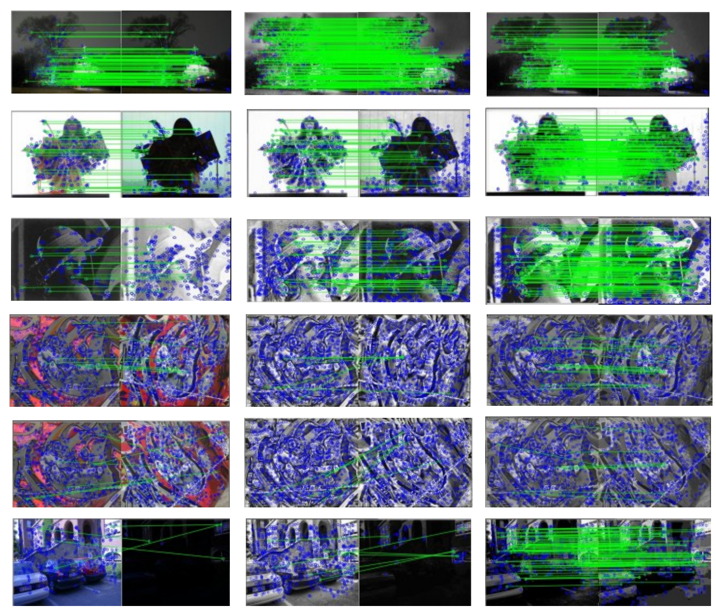
Table 3. Quantitative evaluation for Fig. 7 (artificial data). It shows that percentage of correct matches for proposed method is the highest for all images.

No	% correct matches for original image	% correct matches for adaptive histogram	% correct matches for proposed method
1	91.45	91.88	95.03
2	17.54	25.17	62.32
3	44.28	31.67	77.29
4	2.21	0	5.99
5	0.37	0	2.58
6	0	0	40



(a) (b) (c)

Fig. 6. Results of feature matching performed on (a) without pre-processing; (b) with adaptive histogram equalization method; (c) with proposed pre-processing method. For (b) and (c), only correct matches are displayed.



(a) (b) (c)

Fig. 7. Results of feature matching performed on artificial data in [10] - [13]. (a), (b) and (c) same as illustrate in Fig. 6. The data have been artificially simulated of great illumination changes. Furthermore, it contains of homography transformation to demonstrate competency of the algorithm.

Table 3 shows percentage of correct matches and false matches of feature matching computed on images without pre-processing, with proposed method and adaptive histogram method. The lowest percentage of correct matches is 2.58% due to view point change causes view distortion. However, it might get true homography appropriate to seven correct matches out of 271 descriptors.

Table 4. Comparison of computation time required by adaptive histogram and proposed method. Adaptive histogram is more faster than proposed method about 36%.

Methods	Adaptive Histogram	Proposed Method
Processing time (seconds)	0.0955	Transfer function : 0.1281
		Remapped : 0.1341

V. Conclusion

In this paper, we have proposed a method of pre-processing that is able to improve the quality of the images which are extremely illuminated before feature matching are performed on them. This advantage can be used in any system which has low quality input image such as medical imaging, surveillance and smart car system. Experiment results demonstrate that existing feature matching performs more efficiently with the assistance of pre-process stage.

VI. References

[1] Rajesh Garg, Bhawna Mittal and Sheetal Garg. Histogram equalization techniques for image enhancement. *International Journal of Electronics & Communication Technology (IJECT), 2(1): 107-111 (2011).*

[2] Joung-Youn Kim, Lee-Sup Kim and Seung-Ho Hwang. An advance contrast enhancement using partially overlapped sub-block histogram equalization. *IEEE Trans. on Circuits and Systems for Video Technology, 11(4): 474-484 (2001).*

[3] Chen Soong Der and Manjit Singh Sindhu. Re-evaluation of automatic global histogram equalization-based contrast enhancement method. *Electronic Journal of Computer Science & Information Technology, 1 (1): 13-17(2009).*

[4] Rafael C. Gonzalez and Richard E. Woods. Digital Image Processing. 2nd ed, pp.78-79, Prentice Hall (2002).

[5] Yun Rao, Weiyao Lin and Leiting Chen. Image-based fusion for video enhancement of night-time surveillance. *Optical Engineering 49, 120501 (2010).*

[6] Jing Li, Stan Z.Li, Quan Pan and Tao Yang. Illumination and motion based video enhancement for night surveillance. *IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance, 169-175 (2005).*

[7] Melkamu H.Asmare, Vijanth S. Asirvadam, Lila Iznita and Ahmad Fadzil M. Hani. Image enhancement by fusion in contourlet transform. *International Journal on Electrical Engineering and Informatics 2, 29-42 (2010).*

[8] Azizi Abdullah. Supervised learning algorithms for visual object categorization, 24-27, University Kebangsaan Malaysia (2010).

[9] Edouard Oyallon and Julien Rabin. Speeded-up robust features, Image Processing On Line (IPOL), http://www.ipol.im/pub/algo/or_speeded_up_robust_features/

[10] Tinne Tuytelaars and Krystian Mikolajczyk. Local invariant feature detectors: a survey. *Foudation and Trends in Computer Graphics and Vision, 3(3):177-280 (2007).*

[11] Max Lyons. "Focus and exposure blending software". Internet: http://www.tawbaware.com/forum2/viewtopic.php?p=28414, Feb. 15, 2008 [May. 19, 2011].

[12] Sony Corporation. "Sony bright face technology". Internet: http://sony.stampedeglobal.com/sony-advantage/brightface.asp, 2008 [May. 19, 2011].

[13] Massimiliano Versace. "Fuzzy Logic and memristive hardware". Internet: http://www.neurdon.com/2011/08/09/development-of-massively-parallel-fuzzy-adaptive-neural-algorithms-for-efficient-implementation-on-memristive-hardware/, Aug. 9, 2011 [Dec. 23, 2011].