



Image Enhancement by Fusion in Contourlet Transform

Melkamu H. Asmare¹, Vijanth S. Asirvadam², Lila Iznita³ and Ahmad Fadzil M. Hani⁴

^{1,2,3,4} Department of Electrical and Electronic Engineering, Universiti Teknologi PETRONAS, Tronoh, Perak, Malaysia

¹melkhun@gmail.com, ²vijanth_sagayan@petronas.com.my, ³lilaiznita@petronas.com.my

Abstract: Most existing image enhancement algorithms work on a single image. Their performance is limited to the capacity of the sensor by which the image is taken. In some cases they completely fail to provide us the necessary enhancements. This paper proposes a composite image approach for enhancing still images. The approach proposed combines the relevant features of the input images and produce a composite image which is rich in information content for human eye. The input images are first decomposed into multiple resolutions by using the contourlet transform which provides a better representation than the conventional transforms. Transformed coefficients are combined with a predefined fusion rules. The resultant image is found by performing inverse contourlet transformation of the composite image. The results found are encouraging and the algorithm does not introduce any distortion for applications in low light and/or non uniform lighting conditions. The composite image also contains almost all of the salient features of the input images

Keywords : Image Enhancement, Single image, Composite Image, Fusing Images, Contoulet Transform.

1. Introduction

Image enhancement is a process of improving the interpretability or perceptibility of information in images for human viewers or to provide better input for other automated image processing techniques. Low visibility in still images is generally presented in as dark shadows, over bright regions and blurred details.

One possible way to solve the above problems is to apply image enhancement original single image. A lot of algorithms [1-4] have been developed in this area and their performance is limited with the performance of the sensors in which the image is taken. Either due to design or observational constraints a single image approach usually fails in providing the necessary enhancements.

The other possible approach is to enhance image features by using the information gathered from multiple images. For example, one can combine image from a night vision camera with an image from a visual camera. In this case, the night vision camera is capable of taking images in low light condition but it cannot capture any color information. On the other hand, the visual camera can take the color information but the image captured will have low contrast and dark shadows. Combining these two images one can successfully capture all the relevant information. This process is called image fusion. Since the fused image generally possesses more scene information than any single input image, image fusion can also be considered as an image enhancement process. The multi sensor image fusion has become less expensive as the price of image sensors has dropped in the last decades.

The combination process can take place at different levels of information representation. A common categorization is to distinguish between pixel and region level [5]. Image combination at pixel level is the integration of low-level information, in most cases physical measurements such as pixel intensity. It generates a composite image in which each pixel is determined from a set of corresponding pixels in the various sources. In this paper a pixel based approach transformed image is used to compared for activity measure.

The first multiresolution based image fusion approach was proposed by Burt [6]. This implementation used a Laplacian pyramid and a “choose max” selection rule that, at each sample position in the pyramid, the source pyramid coefficient with the maximum value is copied to the composite pyramid. Toet [7-8] proposed using a ratio of low pass pyramid and the same selection rule which is used by Burt above to fuse IR and visible image. Burt and Kolczynski [9] proposed another method which introduces the concept of activity measure and match measure using the wavelet transform. Several other methods exist in the literature [10-16] focused on fusing dual (or composite images) with various multi resolution transformation techniques.

This propose a contourlet transform based multiresolution approach to combine (or fused) the two source images for vision clarity. The rest of the paper is divided as follows, in section two(2) the contourlet transform is discussed briefly but in detail. Section three(3) details out the composite images approach proposed in the paper for image enhancement. Section four(4) provides the comparison results and discussion part and finally section five(5) concludes the paper.

2. Introduction to Contourlet Transform

For image enhancement, one needs to improve the visual quality of an image with minimal image distortion. Wavelet bases present some limitations, because they are not well adapted to the detection of highly anisotropic elements such as alignments in an image. Recently Do and Vetterli [12] proposed an efficient directional multi resolution image representation called the contourlet transform. Contourlet transform has better performance in representing the image salient features such as edges, lines, curves and contours than wavelet transform because of its anisotropy and directionality. It is therefore well-suited for multi- scale edge based color image enhancement.

The contourlet transform consists of two steps which is the sub band decomposition and the directional transform. A Laplacian pyramid is first used to capture point discontinuities, then followed by directional filter banks to link point discontinuity into lineal structure. The over all result is an image expansion using basic elements like contour segments, thus the term contourlet transform being coined. Figure 1 shows a flow diagram of the contourlet transform.

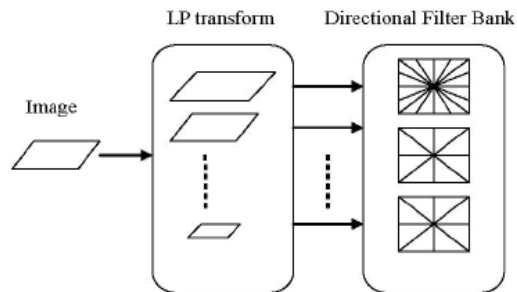


Figure 1. Contourlet transform framework.

Figure 2 below shows the contourlet filter bank. First, multi scale decomposition by the Laplacian pyramid, and then a directional filter bank is applied to each band pass channel.

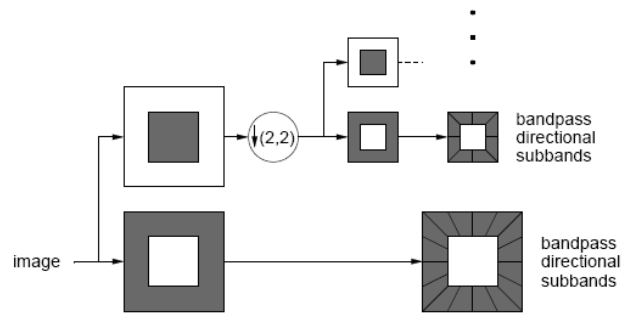
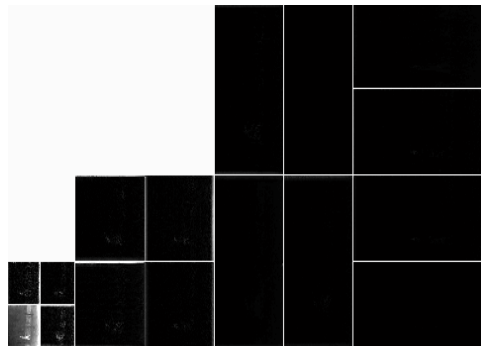
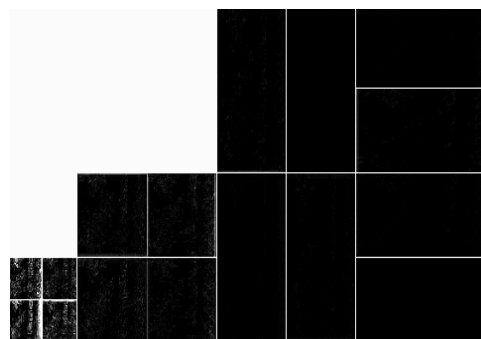


Figure 2. Contourlet filter bank.

Figure 3 shows a sample contourlet transform coefficients of the test images in case study IV (refer to the results sections). This is level three, with 3, 4, and 8 directions from coarse to fine levels respectively.



(a)



(b)

Figure 3. Contourlet transform coefficients of the source images, a) visual image
b) night vision image

3. Image Enhancement using composite images

Today, most composite image approaches employ pixel fusion methods. The advantage of pixel fusion is that the images used contain the original information. Furthermore, the algorithms are rather easy to implement and time efficient. As the present author observed before in one of the findings [13], an important pre-processing step in pixel based fusion methods is image registration, which ensures that the data at each source is referring to the same physical structures. In the remainder part of the paper, it will be assumed that all source images have been registered.

Figure 4 shows a detailed block diagram of the proposed scheme in flow layout. The framework contains eight modules and the description of each of the building blocks is given below in each process block.

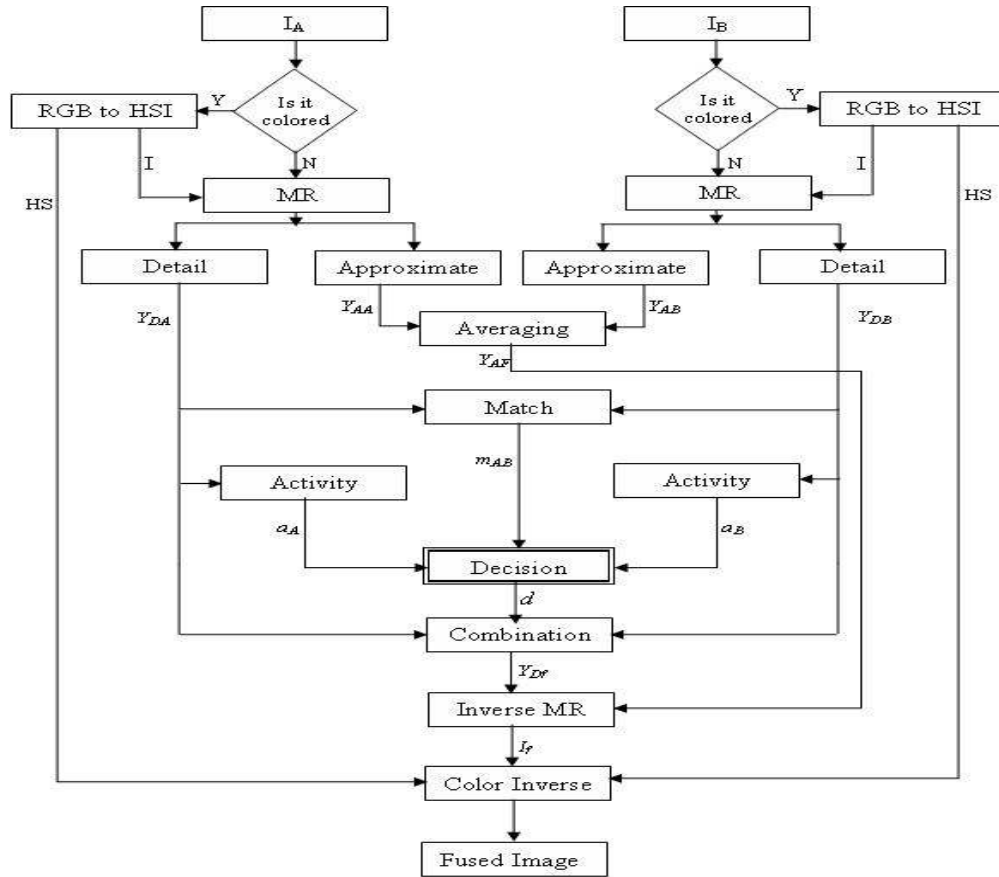


Figure 4. General Composite image enhancement Framework

A. MR Analysis:

This block computes an MR (or multi-resolution) decomposition using the contourlet transforms from the input sources. The contourlet transform generates approximate and detail coefficients. Since both (approximate and detail) of their different physical meanings, thus they are treated by the combination algorithm in a different manner. The term approximate and detail coefficient will be explained in the later stage.

B. Color Transformation

Usually the visual images carry color information which has the important properties. Color images are represented by three spectral bands, the usual or commonly used RGB bands. Performing the analysis on each spectral band is a huge sum of task but instead the color image is first converted to an appropriate color space, in this work, an HSI (Hue, Saturation, and Intensity) being adopted, which successfully separates the color information and the intensity. The research work limits the analysis on the intensity image only.

C. Approximate Coefficients:

In order for approximate images represent a coarse representation of the original image I_A and I_B respectively. They inherit some of its properties such as the mean intensity or some coarse texture information. Thus, coefficients Y_{AA} and Y_{AB} with high magnitudes do not necessarily correspond with salient features. Therefore, the composite image Y_{AF} is a simple averaging of the input images as given in equation 1.

$$Y_{AF} = \frac{Y_{AA} + Y_{AB}}{2} \quad (1)$$

However, there is a technical reason behind this simple averaging, which is based on the assumption that the source images contain additive Gaussian noise and the decomposition level is high enough, a minimum of four levels in out case. Important image features have already been captured by detail coefficients. Thus, the approximate images contain mainly noise and averaging them reduces the variance of the noise while ensuring appropriate mean intensity.

D. Detail Coefficients:

For detail images the relevant perceptual information relates to the edge information that is present in each of the detail coefficients Y_{DA} and Y_{DB} [13-14]. Detail coefficients having large absolute values correspond to sharp intensity changes and hence to salient features in the image such as edges, lines and region boundaries. Thus the fusion is done as follows [7]. Match and activity measures will be performed on the coefficients. From these two measures we will construct a decision matrix which is a key for the combination of the images. The final composite image is found by taking the inverse contourlet transform of the combination output.

E. Activity Measure:

It is the degree of saliency of each coefficient in Y_{DA} and Y_{DB} (i.e. its importance to image fusion at hand). The saliency should increase when features are in focus or it should give emphasis on the contrast difference. In MR decomposition systems the contrast information can be attributed by the magnitude of the detail frequency components. Higher amplitude implies a good contrast. The fact that human visual system is highly sensitive to local contrast changes (edges) lets us to compute the activity as a local energy measure.

$$a_i^k(x, y) = \sum_{\Delta n \in w^k} |Y_i^k(x + \Delta n, y + \Delta n)|^2 \quad (2)$$

Where k is the decomposition level, i is input image in this case images A and B, and W is a finite local window of size 1X1, 3X3, or 5X5.

The activity measures of the source images (Figure. 5(a)) can be visualized as a gray scale image where higher gray values pixel indicates a higher saliency of that pixel. The man in Figure. 5(b) shows a higher gray level thus its saliency is very high.



Figure 5. Visualization of the Activity measures (a) Source Image (b) Visual Image (c) night vision image

F. Match measure:

It is used to quantify the degree of similarity between the sources. Precisely $m_{AB}^k(x, y)$ reflects the resemblance between Y_{DA} and Y_{DB} . The match measure tells where the sources are different and to what extent they differ. We can use this information to appropriately combine them. In our algorithm it is defined as a normalized correlation average over neighborhood of the samples as shown in equation (3). Where W^k being window of size either 1x1 3x3 or 5x5 centred at the origin.

$$m_{AB}^k(x, y) = \frac{2 \sum_{\Delta n \in W^k} Y_A^K(x + \Delta n, y + \Delta n) Y_B^k(x + \Delta n, y + \Delta n)}{a_A^k(x, y) + a_B^k(x, y)} \quad (3)$$

G. Decision Map:

It is the core of the combination algorithm. Its output governs the actual combination of the coefficients of the various sources. It controls the weights to be assigned to each source coefficients. The conventional approach is to assign a weight which is directly proportional to the activity measure. But this has a contrast reduction effect in images which have opposite contrast. It actually assumes that at a point only one of the source images is a valid choice. The ultimate solution is to apply a selection rule which considers the match measure which is proposed by Burt [6]. At points where the source images are different the combination selects the most salient component. At location where they are similar, we average the source components. The averaging provides stability and reduces noise. In this case it is computed as shown in equation 4 where T is decision threshold, T=0.85 for this experiment.

$$d^k(.) = \begin{cases} T & \text{if } m_{AB}^k(.) \leq T \text{ and } a_A^k(.) > a_B^k \\ 1-T & \text{if } m_{AB}^k(.) \leq T \text{ and } a_A^k(.) \leq a_B^k \\ \frac{1}{2} + \frac{1}{2} \left(\frac{1-m_{AB}^k(.)}{1-T} \right) & \text{if } m_{AB}^k(.) > T \text{ and } a_A^k(.) > a_B^k \\ \frac{1}{2} - \frac{1}{2} \left(\frac{1-m_{AB}^k(.)}{1-T} \right) & \text{if } m_{AB}^k(.) > T \text{ and } a_A^k(.) \leq a_B^k \end{cases} \quad (4)$$

H. Combination Map:

It describes the actual combination of the transform coefficients of the sources. In this research we used the weighted average where the weights are determined from the decision map.

$$y_F^K(.) = d^K(.)y_A^K(.) + (1-d^K(.))y_B^K(.) \quad (5)$$

I. MR Synthesis:

Finally, the composite image is obtained by applying the inverse transform on the composite MR decomposition YF,

$$I_F = \Phi^{-1}(Y_F) \quad (6)$$

4. Results and Discussion

The proposed method is applied in several types of images using five case studies for visual comparison.

A. Case Study I

The method is applied in fire Fighting and rescue operation. The image is used by permission from [14]. The image in Figure 5a is taken by a visual camera. It shows all the colour information of the houses and the smoke. The second image in Figure 5b is taken by a night vision camera. It reveals the details of the houses and three men. Using either of the images alone is misleading for a fire Fighter group so we fuse the two images and the composite image in Figure 5c shows all the salient information.

B. Case Study II

The method can also be applied for driving at night or bad weather. It is used by permission from [9]. The man by the side of the car is almost invisible in the visual image of Figure 6a while we can clearly see him in the thermal image in Figure 6b in spite of the fact that there is

no colour information. The composite image is precise and contains almost all relevant information for clear visual perception.

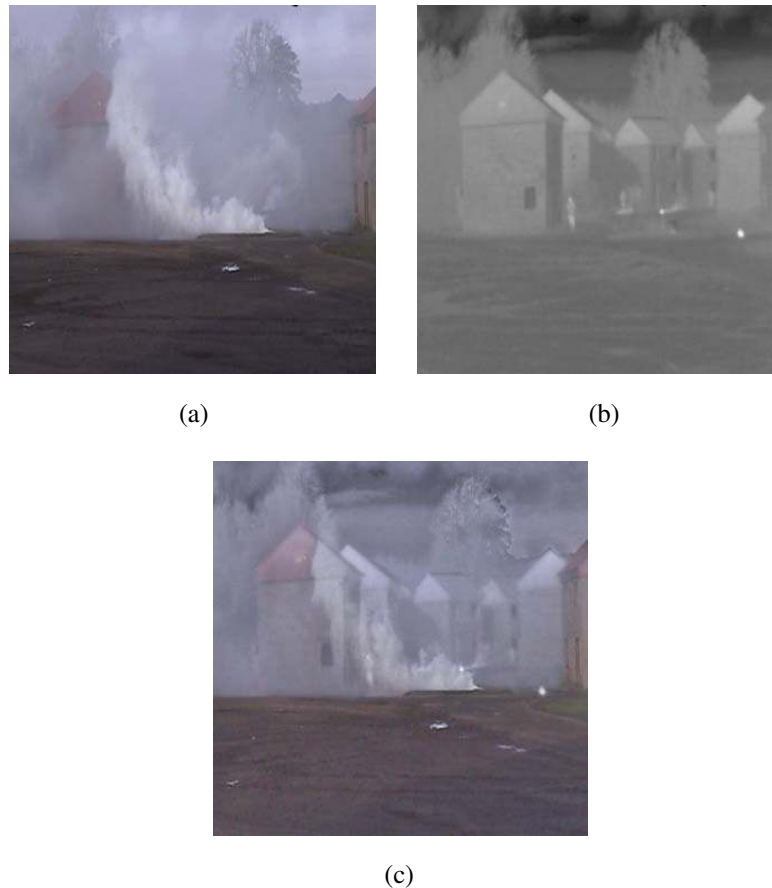


Figure 6. Case Study I, a) visual image b) night vision image c) composite image

C. Case Study III

The sample result on Figure 7 shows a house and trees. It is used by permission from [14]. In the visual image of Figure 7(a) the color is clear and the image is so natural. The infrared image of Figure 7(b) shows the same image but the normal color information is lost but this image shows the details of the house and the electric cable in front of the house which is not visible in the visual image. The composite image in Figure 7c shows more complete information with no color degradation.

D. Case Study IV

This approach can also be used in security surveillance applications in Figure 8. This image is taken in a typical area surveillance scenario for monitoring an outdoor scene in a bad visibility condition. It is used by permission from [8]. A man is walking through the scene behind the bushes. He is barely visible in the visual image of Figure 8a while the night vision camera in Figure 8b captures him clearly but not the details of the vegetation. Thus appropriately combining these two images will generate a complete composite image as shown in Figure 8c.

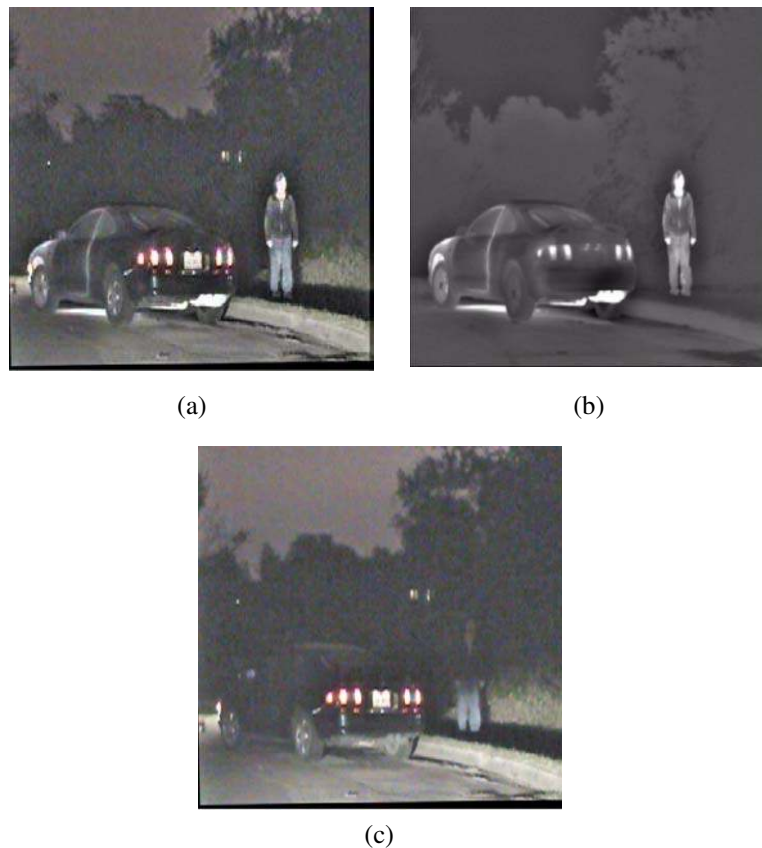


Figure 6. Case Study II, a) visual image b) night vision image c) composite image

E. Case Study V

The algorithm is also applied to the typical multi focal images as shown in Figure 8. The smaller watch is blurred while the bigger one is clear in Figure 8a. The reverse is true in the Figure 8b. The composite image shows both watches with no blurring effect and has good quality with fine details revealed.

F. Red Eye Removal

When capturing photographs in low lighting environments it is a common practice to use the flash mode of the cameras. Doing so will give us a brighter image which has a greater signal-to-noise ratio that reveals the details of the scene that would always be hidden in a photograph acquired under ambient illumination without the flash. However using the flash may introduce artificial artifacts such as a red eye. The optimal solution for this problem is to take a pair of images one with flash and the other with no flash. These two images can be combined using this algorithm and the result is encouraging.

The sample result in Figure 8 shows the application of the fusion method in red eye removal of a girl (courtesy of [17]). Figure 8(a) shows a picture taken with the flash is on and has red eye whereas the picture in Figure 8(b) is the same image taken using no flash mode. This image suffers from low illumination but there is no red eye. Combining these two images with fusion method based on contourlet transform can creates a more complete image without the red eye as shown in Figure 8(c).

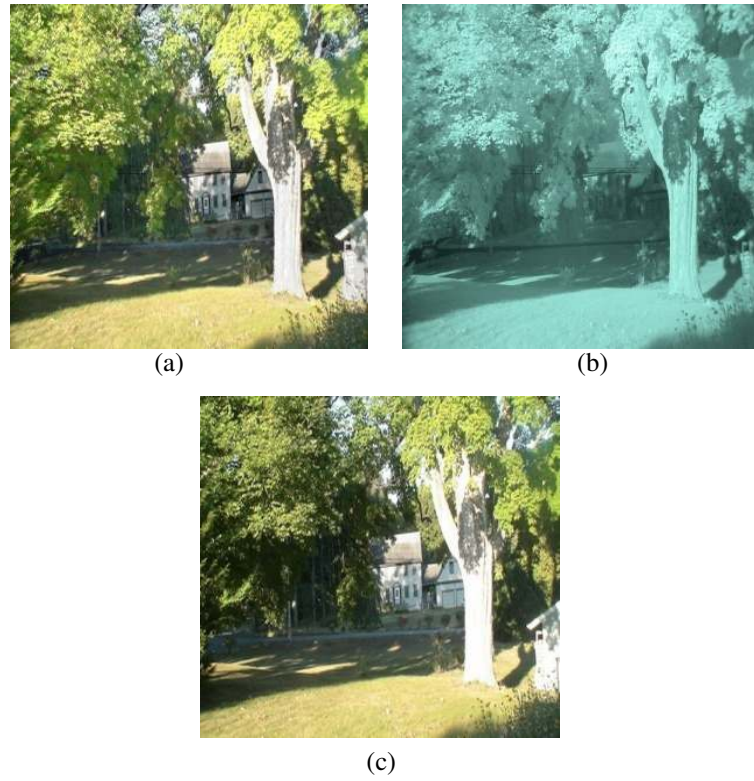


Figure 7. Case Study III, a) visual image b) Infrared image c) composite image

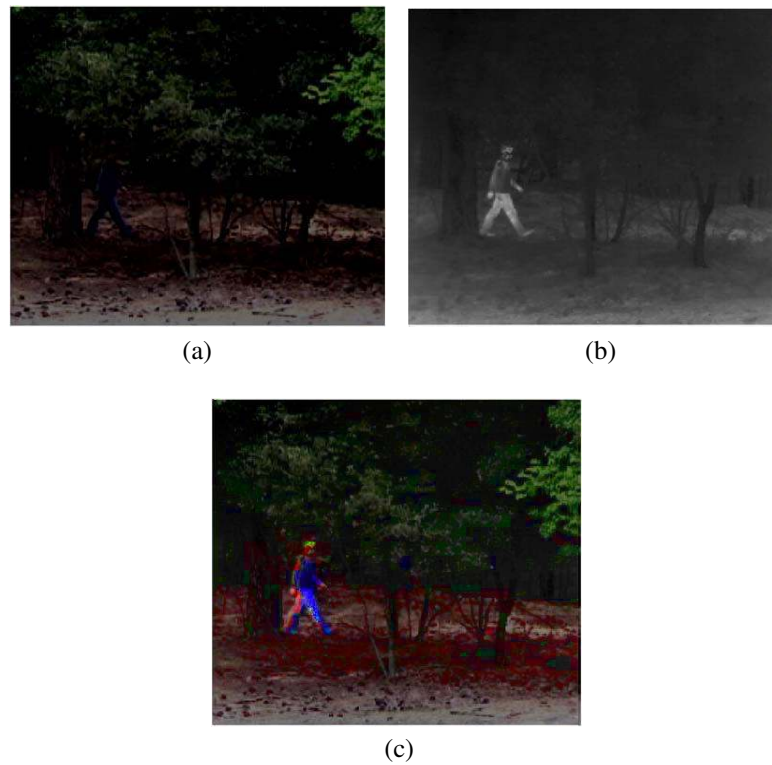


Figure 8. Case Study IV, a) visual image b) night vision image c) composite image

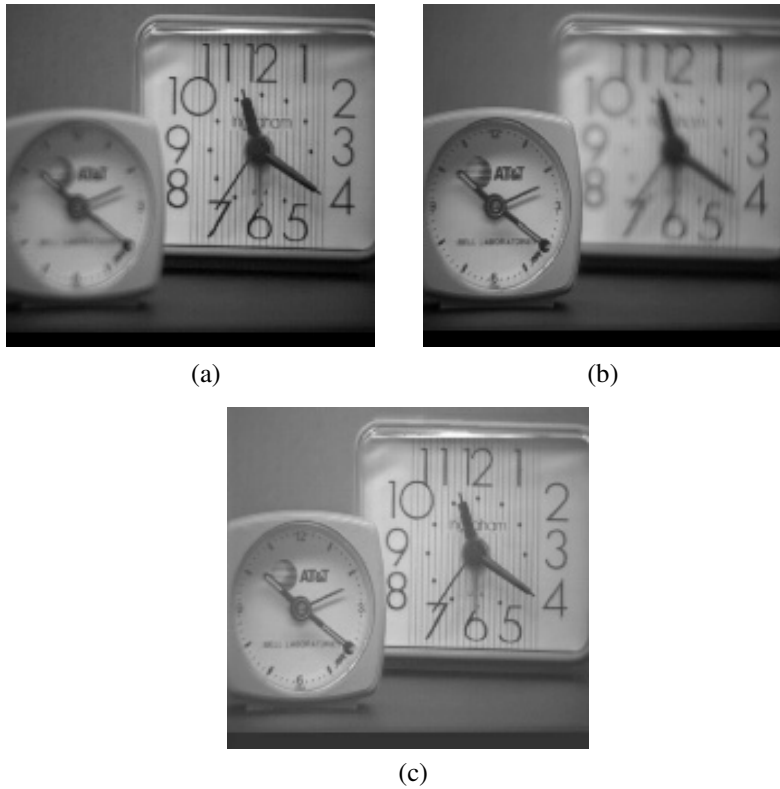


Figure 7. a) Multi focal input image b) multi focal second input image c) fusion result

Table 1 compares the quality of these composite images using the windowed structural similarity index (WSSIM) and PSNR between the original and the generated fused images. The numerical results in the Table 1 show that the contourlet result is way superior to the wavelet method. All the methods gave good indexes as the test images provided are easy to combine. However, the contourlet transform method gives the maximum computation time. This is mainly because the construction of the contourlet transform consists of double filter banks which take higher time to generate the coefficients.

Table 1. Performance evaluation result of the test image on Figure 8

Fusion Method	WSSIM	PSNR	Computational Time (in seconds)
Contourlet Method	0.99649	34.7126	4.171588
Wavelet Method	0.99612	31.9957	4.051072
Simple Averaging	0.93573	26.5733	1.288362

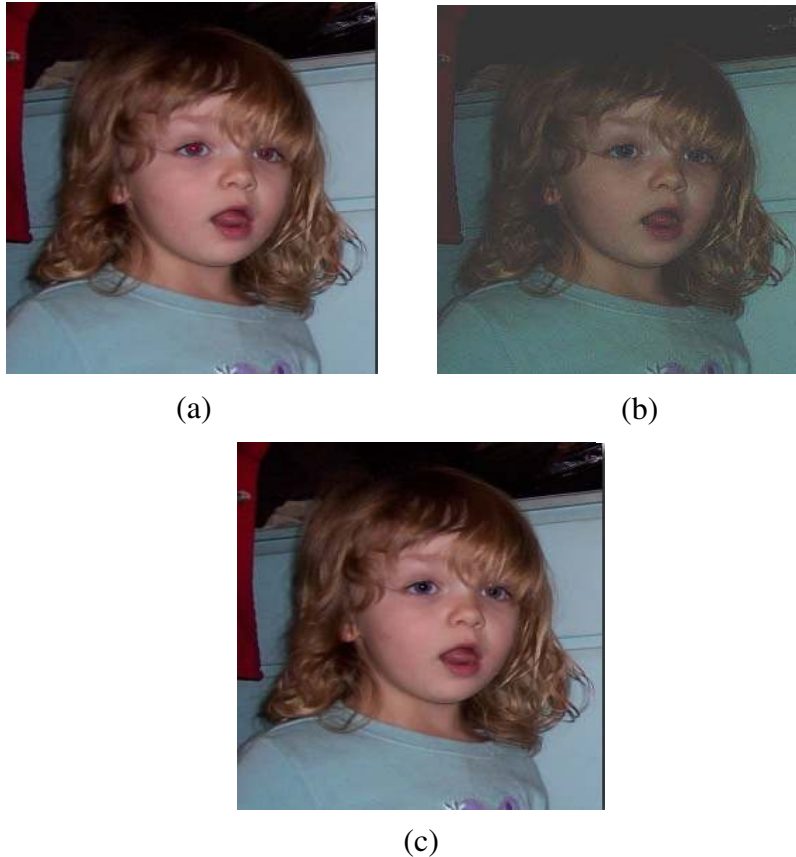


Figure 8. Red eye removal

(a) Image taken when the flash is on (b) image taken when the flash is off (c) fused image

5. Conclusion

In many applications human perception of the fused images is of a fundamental importance and as a result the enhancement results are mostly evaluated by subjective criteria. The experimental results show that the fusion algorithm gives encouraging results for both multi modal and multi focal images. Since the image salient features such as edges lines and contours are well represented using the contourlet transform, the fusion process did not introduce any distortion to the original image.

Directional information introduced by the contourlet transform yields the best description of the all the salient information in the both test images. Thus the composite image is more complete and looks natural and the noise level is minimal. Utilization of the composite image is expected to increase the performance of the subsequent processing tasks. By integrating information, this approach can reduce dimensionality. This results in a more efficient storage and faster interpretation of the output. Using redundant information, the composite image increases the accuracy as well as reliability, and using complementary information, the process also improves interpretation capabilities with respect to subsequent tasks. This leads to more accurate data, increased utility and robust performance.

6. Acknowledgement

We would like to acknowledge Universiti Teknologi PETRONAS for graduate assistantship and providing facilities for the research.

References

- [1] Koen Vande Velde, "Multi-scale Color Image Enhancement", *IEEE Proceedings of International Conference on Image Processing (ICIP 99)*, Vol. 3, pp. 584-587, 1999.
- [2] Z. Rahman, D. D. Jobson, and G. A. Woodell, "Retinex processing for automatic image enhancement" *Journal of Electronic Imaging*, Vol. 13, No. 1, pp. 100-110, January 2004
- [3] Jean-Luc Starck, Fionn Murtagh, Emmanuel J. Candès, and David L. Donoho, "Gray and Color Image Contrast Enhancement by the Curvelet Transform", *IEEE Transactions on Image Processing*, Vol. 12, No 6, , pp. 709-716
- [4] Melkamu. H. Asmare ,Vijanth S. Asirvadam and Lila Iznita, "Multiscaled Image Enhancement using Contourlet Transform", *Platform*, Vol. 6, No. 2, July-Dec Issue 2008
- [5] Melkamu. H. Asmare ,Vijanth S. Asirvadam and Lila Iznita, "Multi-scale color image enhancement using contourlet transform", *International Graduate Conference on Engineering and Science (IGCES 08)*, Johor Bharu, pp. 117-121, December 2008.
- [6] Burt, P. J. "A gradient pyramid basis for pattern selective image fusion", *SID Technical Digest 16* , pp. 467-470, 1985.
- [7] Tao Li, "Multi Modal Enhancement techniques for visibility improvement of digital images" *Ph.D Thesis, Old Dominion University*, 137 pages, 2006.
- [8] Toet, A, "Image fusion by a ratio of low-pass pyramid. *Pattern Recognition*, Vol 9 , pp. 245-253, 1989.
- [9] Burt, P. J., and Kolczynski, R. J., "Enhanced Image Capture Through Fusion". *In Proceedings of the 4th International Conference on Computer Vision* , pp. 173-182, May 1993
- [10] Andrea Masini, Francesco Branchitta, Marco Diani, Giovanni Corsini, "Sight enhancement Through Video Fusion in a surveillance system" *14th International Conference on Image Analysis and Processing*, pp. 554-559., 2007
- [11] Li Tao, Hau Ngo, Ming Zhang, Adam Livingston and Vijayan Asari, "A Multi-sensor Image Fusion and Enhancement System for Assisting Drivers in Poor Lighting Conditions", *Proceedings of the 34th Applied Imagery and Pattern Recognition Workshop (AIPR05)*, 2005.
- [12] Do and Vetterli, "The Contourlet Transform: An efficient Directional Multi Resolution Image Representation", *IEEE Transactions on Image Processing*, Vol. 14, pp: 2091-2106, 2005.
- [13] Melkamu. H. Asmare ,Vijanth S. Asirvadam and Lila Iznita, "Multi-Sensor Image Enhancement and Fusion for Vision Clarity Using Contourlet Transform", *International Conference on Information Management and Engineering (ICIME 2009)*, pp. 352-356, April 2009
- [14] Gang Liu, Xue-qin Lu," Pixel-level image fusion based on fuzzy theory" *Proceedings of the Sixth International Conference on Machine Learning and Cybernetics, Hong Kong*, ,pp.1510-1514, August 2007.
- [15] Wei Huang, Zhongliang Jing, "Multi-Focus Image Fusion using Pulse Coupled Neural Network Source" *Pattern Recognition Letters*, Vol 28, Issue 9, pp. 1123-1132, July 2007.
- [16] G. Piella, Adaptive Wavelets and Their Applications to Image Fusion and Compression, *Ph.D. Thesis, CWI & University of Amsterdam*, 2003.
- [17] G. Petschnigg, M. Agrawala, H. Hoppe, "Digital Photography with Flash and No-Flash Image Pairs", *ACM 2004*, Page(s): 664-672.



Melkamu H. Asmare was born in Dangla, Ethiopia, in 1982. He received B.Sc. Degree in Electrical and Computer Engineering from Addis Ababa University, in 2005. He has just completed his Masters in Electrical Engineering in Universiti Teknologi PETRONAS in July 2009. Currently he took up lectureship position at University of Bahadar Ethiopia. His research interests include Image processing, color science, digital signal processing and image fusion.



Dr. Vijanth S. Asirvadam is from an old mining city of Malaysia called Ipoh. He studied at the University of Putra, Malaysia for the Bachelor Science (Hon) majoring in Statistics and graduated in April 1997 before leaving to Belfast Queen's University Belfast to do his Masters where he received his Master's Science degree in Engineering Computation with a Distinction. He later joined the Intelligent Systems and Control Research Group at Queen's University Belfast in November 1999 where he completed his Doctorate (Ph.D.) researching Online and Constructive Neural Learning methods. He took previous employments as a System Engineer and later as a Lecturer at the Multimedia University Malaysia and also as a Senior Lecturer at the AIMST University. Since November 2006, he serve as a Senior Lecturer at the department of Electrical and Electronics Engineering, Universiti Teknologi PETRONAS. His research interest includes linear and nonlinear system identification, model validation and application of intelligent system in computing and image processing. Dr Vijanth is member of Institute of Electrical and Electronics Engineering (IEEE) and Institute of Engineering Technology (IET). He has over 60 publications in local and international proceedings and journals.



Lila Iznita Izhar obtained her BEng in Electrical and Electronics Engineering from the University of the Ryukyus, Japan in 2002. She later earned her MSc in Electrical and Electronics Engineering from Universiti Teknologi PETRONAS in 2006. She is currently a lecturer with Universiti Teknologi PETRONAS. She has over 20 publications in local and international proceedings and journals and 2 patents being filed. Her main research interests are in the area of Medical Image Analysis, Computer Vision and Image Processing and Biomedical Engineering.