A Novel Two Step Region Based Multifocus Image Fusion Method

Tanish Zaveri, *Member, IACSIT, Mukesh Zaveri*

***Abstract*—Image fusion is a process of combining multiple input images of the same scene into a single fused image, which preserves relevant information and also retains the important features from each of the original images and makes it more suitable for human and machine perception. In this paper, a novel region based image fusion method is proposed. The literature review shows that region based image fusion algorithm performs better than pixel based fusion method. Proposed algorithm is applied on large number of registered images and results are compared using standard reference and nonreference based fusion parameters. The proposed method is also compared with different methods reported in the recent literature. It has been observed that simulation results of our proposed algorithm is consistent and preserves more information compared to earlier reported pixel based and region based methods.**

***Index Terms*—Multifocus, Multisensor, Normalized Cut, Spatial Frequency**

I. Introduction

Multifocus image fusion has been used widely in the field of image analysis task such as target recognition, remote sensing and medical image processing. The aim of image fusion is to combine relevant information from two or more source images into one single image such that the single image contains most of the information from all the source images. In the application of digital camera, optical lenses suffer from a limited depth of focus. Because of this limitation, it is often not possible to get an image that contains all relevant objects in focus. Part of the image which is out of focus has less depth of field. One possible solution to overcome this problem is to take several pictures with different focus settings and combine them together into a single frame to get all the information from the less focus area using image fusion method. The goal is to enhance the image quality and information so that it provides more detail information than the information available in single image.

Image fusion methods [1] are classified mainly into two categories: (i) pixel based and (ii) region based. Pixel based methods [1, 2, 3] generally deal with pixel level information directly. Pixel level image fusion methods are affected by blurring effect which directly affect on the contrast of the image [1, 4]. These methods are generally time consuming as they require more number of computations. The number of computations can be reduced by adpoting multi-resoultion approach. Multi-resoultion (MR) methods [4, 5] are very useful for image fusion because the real world objects usually consist of structures at different scales and human visual system also processes information in a MR fashion. MR methods are computationally efficient and more robost. Various methods based on the multiscale transforms have been proposed such as contourlet transform based method [3], pyramid-based approach [5] and discrete wavelet-based method [2]. The basic idea is to perform a multiresolution decomposition on each source image, then integrate all these decompositions to form a composite representation, and finally reconstruct the fused image by performing an inverse multiresolution transform. Second category of alogrithms for image fusion use segmentation, namely, known as region based methods [4, 6].

Generally, the source images availabe for image fusion have limited depth of focusing problem. In such a scenario the region based methods are preferred. In this paper we propose a novel region based algorithm which solves this limited focusing depth problem and produces final fused image which contains all the information of source images which is simultaneously not present in single multifocus image. The fundamental idea here is that we human begins interpret images at the region or object level rather than pixel level. Region based algortithm has many advantages over pixle based like it is less sensitve to noise, better contast, less afected by misregistration. Also a region is more meaningful structure in multifocus image. The heart of our algorithm is the segmentation of an image and recently proposed normalized cut based [7] image segmentation method is used in our proposed image fusion algorithm. The proposed algorithm is described in following section. The reference based and nonreference based image fusion evaluation parameters are introudced in section 3. The simulation results and assessment are described in Section 4. It is followed by the conclusion.

II. Proposed Algorithm

The proposed region based algorithm consists of mainly three steps. In first step the given source images, which are to be fused, are segmented. The idea behind using the segmentation in the proposed algorithm is based on the method described in [7]. In [7], the algorithm uses on the perceptual grouping for vision problem. Rather than focusing

Manuscript received June 7, 2009.

Tanish H. Zaveri, is with the Nirma University, EC Dept., Ahmedabad, Gujarat, India (corresponding author phone: +91-2717-241911-15; fax: +91-2717-21916; e-mail: ztanish@ieee.org).

Mukesh A. Zaveri, is Head, Department of Computer Engg., Sardar Vallabhbhai National Institute of Technology, Surat, Gujarat, India. (e-mail: mazaverir@coed.svnit.ac.in)..

on local features, our approach aims at extracting the global information of an image. In the proposed method, the image segmentation process is treated as a graph partitioning problem. A novel global criterion, normalized cut, is used for segmenting the graph [11].

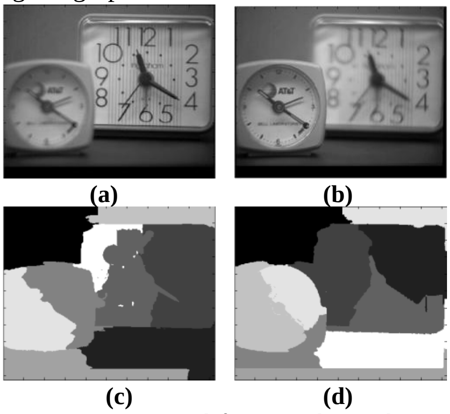


Figure.2 (a) input image with focus on right (IA) (b) input image with focus on left (IB) (c) Segmentation on IA (d) Segmentation on IB

The output of the segmentation step is the heart of the proposed method as the whole implementation is based on the segmented region. Even very small change in segmentation result can bring a huge difference to the final result. For example, the two source images, namely, IA and IB, having different focused points are segmented using normalized cuts algorithm in first step. In source image IA, the right clock is focused where left clock is unfocused. First step gives an ideal segmentation of unfocused part in the image. It can be seen clearly from Fig. 2 (c) and (d). The algorithm perfectly segments the left clock into two regions depending on clock hand position. Similarly, when image IB is segmented, the algorithm perfectly segments right clock into three regions depending on clock hand position. One of the major advantages of normalized cut based segmentation is that even if user increases the number of thresholds in segmentation algorithm, the resultant segmented image contains the same number of region as in unfocused area so it allows user to use more number of region in focused area. Block diagram of proposed method is shown in the Fig. 1. Once the input images are partitioned using normalized cut segmentation algorithm, the second step is to create the first level of fused image using fusion rules-1.

In this step, fusion rule-1 is applied on all the regions extracted from both images and first level fused image Irl1 is reconstructed. The fusion rule-1 is based on spatial frequency (SF) which is explained later in this section. SF of extracted regions of from Image IA and IB is defined as SF_A and SF_B respectively which is computed and compared using following rule. The fusion rule-1 is defined as

$$I_{f1n} = \begin{cases} R_{An} & SF_{An} \geq SF_{Bn} \\ R_{Bn} & SF_{An} \leq SF_{Bn} \end{cases} \tag{1}$$

Here, n represents region number and the range of n depends on the number of segmented regions available after applying segmentation algorithm. If the number of regions is **i** than n = 1, 2, 3….i. I_{f1n} is a resultant intermediate fused image after all nth fused regions are extracted by applying fusion rule-1. Regions extracted from images IA and IB are represented as R_{An} and R_{Bn} respectively. Merge all the regions to generate intermediate fused image Irl1 and Irl2 as shown in the Figure 2. Irl1 is generated using the segmented regions based on image IA where the left clock is unfocused but in the image Irl1 the left clock is focused and moreover, the other parts of image also giving better information than original. Similarly Irl2 will have right clock totally in focus. It is followed by the last step where the second level of fusion rules are applied which gives the final fused image so in order to get better fused image both the energy and SF image quality parameters are used to define fusion rule-2. The images, Irl1 and Irl2, are segmented using the same regions extracted while considering input image IA. Image fusion rule-2 is applied on all the regions of image Irl1 and Irl2. The fusion rule-2 is defined as

$$I_{fuse} = \begin{cases} I_{rl1} & SF_{An} \geq SF_{Bn} \ \& \ E_{An} \geq E_{Bn} \\ I_{rl2} & SF_{An} < SF_{Bn} \ \& \ E_{An} < E_{Bn} \\ (I_{rl1} + I_{rl2})/2 & else \end{cases} \tag{2}$$

Finally, the fused image, namely, $\boldsymbol{I_{fuse}}$ is generated by merging all the extracted regions using fusion rule-2. The fusion rule-1 and fusion rule-2 are formulated based on two parameters, namely, the spatial frequency and energy. The reason to choose these parameters is that the spatial frequency indicates the clarity of image [9] and energy indicates activity level in image. Spatial frequency is an important indication to measure the image details. The higher the image details more are the SF. The spatial frequency is defined as

$$\mathrm{SF} = \sqrt{\mathrm{RF}^2 + \mathrm{CF}^2} \tag{3}$$

Energy is also used to measure the overall activity level of an image, which is also an important indication to measure the image details. The energy for image F of size M x N is defined as

$$E = \sum_{i=1}^{M} \sum_{j=1}^{N} [F(i, j)]^2 \tag{4}$$

As described above, the proposed image fusion method uses only the regions to generate resultant image and it never uses an individual **pixel** and hence, the proposed method is able to overcome all the problems of pixel based image fusion methods.

III. Evolution Criteria For Fused Image

Any image fusion algorithm can be assessed using two categories of performance measurements parameters which are subjective and objective which may further divided into reference [10] and non reference quality assessment parameters [10]. Subjective indices rely on the ability of people's comprehension and are hard to come into application. While objective indices can overcome the influence of human vision, mentality and knowledge, and make machines automatically select a superior algorithm to accomplish the mission of image fusion. Objective indices can be divided into three categories based on subjects reflected. One category reflects the image features, such as entropy, spatial frequency and gradient [3]. The second reflects the relation of the fused image to the source images, such as mutual information. The third reflects the relation of the fused image to the reference image, such as cross entropy, correlation coefficient, Root mean square error (RMSE). We have used each category of fusion parameter to evaluate our final fused image.

A. Reference Based Image Fusion Parameters

Most widely used reference based image fusion performance parameters are Entropy, Structural similarity Matrix (SSIM), Quality Index (QI), Mutual Information (MI), Root mean square error (RMSE).

1) Root mean square error

The Root mean square error (RMSE) is well known parameter to evaluate fused image. It represents amount of deviation present in fused image compared to reference image [10]. The RMSE is calculated between fused image F and standard reference image R which is defined as

$$\mathrm{RMSE} = \sqrt{\frac{\sum\sum[R(i,j)-F(i,j)]^2}{MN}} \tag{5}$$

2) Mutual Information

Mutual information (MI) indices also used to evaluate the correlative performances of the fused image and the reference image as explained in [9]. Let *A* and *B* be two random variables with marginal probability distributions $P_A(a)$ and $P_B(b)$, and joint probability distribution $P_{AB}(a,b)$, mutual information is defined as

$$\mathrm{MIr}_{AB} = \sum P_{AB}(a,b).\log\frac{P_{AB}(a,b)}{P_A(a)P_B(b)} \tag{6}$$

A higher value of mutual information (MI) indicates that the fused image contains fairly good quantity of information presented in fused image compared to reference which is defined as MIr. A higher value of mutual information (MIr) represents more similar the fused image compared to reference image.

The structural similarity index measure (SSIM) proposed by Wang and Bovik [9] is based on the evidence that human visual system is highly adapted to structural information and a loss of structure in fused image is indicating amount of distortion present in fused image. It is designed by modeling any image distortion as a combination of three factors; loss of correlation, radiometric distortion, and contrast distortion as mention in [8, 9]. The dynamic range of SSIM is [-1, 1]. The higher the value of SSIM indicates more similar structures in fused and reference image. If two images are identical, the similarity is maximal and equals 1.

B. Non Reference Based Image Fusion Parameter

The Mutual information (MI), the objective image fusion performance metric $Q^{AB/F}$, spatial frequency (SF) [10] and entropy [10] are important image fusion parameters to evaluate quality of fused image when reference image is not available. MI described in A.2 & in (6) can also be used to evaluate fused images without the reference image by computing the MI between source image IA and fused image IFUSE called as I_{AF} and similarly find I_{BF} using image IB as a source image and calculate total MI as defined by

$$MI = I_{AF} + I_{BF} \tag{7}$$

1) Objective Image Fusion Performance Measure

The goal in pixel level image fusion is to combine and preserve in a single output image all the "important" visual information that is present in a number of input images. Thus an objective fusion measure should (i) extract all the perceptually important information that exists in the input images and (ii) measure the ability of the fusion process to transfer as accurately as possible this information into the output image. In this work we associate important visual information with the "edge" information that is present in each pixel of an image. Notice that this visual to edge information association is supported by Human Visual System [8] studies and is extensively used in image analysis and compression systems.

The objective image fusion performance metric $Q^{AB/F}$ which is proposed by Xydeas and Petrovic [8] reflects the quality of visual information obtained from the fusion of input images and can be used to compare the performance of different image fusion algorithm. Furthermore, by evaluating the amount of edge information that is transferred from input images to the fused image, a measure of fusion performance can be obtained. Consider two input images A and B, and a resulting fused image F. Note that the following methodology can be easily applied to more than two input images. A Sobel edge operator is applied to yield the edge strength g(n,m) and orientation α(n,m) information for each pixel p(n,m), $1 \le n \le N$ and $1 \le m \le M$. Thus for an input image A edge strength g(n,m) and orientation α(n,m) is defined as [8].

$$g_A(n,m)=\sqrt{s_A^x(n,m)^2+s_A^y(n,m)^2} \tag{8}$$

$$\alpha_A(n,m) = \tan^{-1}\left(\frac{s_A^y(n,m)}{s_A^x(n,m)}\right) \tag{9}$$

Where $s_A^x(n,m)$ and $s_A^y(n,m)$ are the output of the horizontal and vertical Sobel templates centered on pixel p(n,m) and convolved with the corresponding pixels of image A. The relative strength and orientation values of $G^{AF}(n,m)$ and $A^{AF}(n,m)$ of an input image A with respect to F are formed as [7]. SF is defined in the proposed algorithm section II. The entropy is also used to evaluate fused image as described below

$$G^{AF}(n,m) = \begin{cases} \dfrac{g_F(n,m)}{g_A(n,m)}, \text{if } g_A(n,m) > g_F(n,m) \\ \dfrac{g_A(n,m)}{g_F(n,m)}, \text{otherwise} \end{cases} \tag{10}$$

$$A^{AF}(n,m) = \frac{\left\| \alpha_A(n,m) - \alpha_A(n,m) \right| - \pi/2 \Big|}{\pi/2} \quad (11)$$

These are used to derive the edge strength and orientation preservation values

$$Q_g^{AF}(n,m) = \frac{\Gamma_g}{1+e^{K_g(G^{AF}(n,m)-\sigma_g)}} \quad (12)$$

$$Q_\alpha^{AF}(n,m) = \frac{\Gamma_g}{1+e^{K\alpha(A^{AF}(n,m)-\sigma_\alpha)}} \quad (13)$$

$Q_g^{AF}(n,m)$ and $Q_\alpha^{AF}(n,m)$ model perceptual loss of information in F, in terms of how well the strength and orientation values of a pixel p(n,m) in A are represented in the fused image. The constants Γ_g, κ_g, σ_g and Γ_α, κ_α, σ_α determine the exact shape of the sigmoid functions used to form the edge strength and orientation preservation values, see equations (12) and (13). Edge information preservation values $Q^{AB/F}$ is then defined as

$$Q^{AB/F}(n,m) = \frac{\sum_{n=1}^{N}\sum_{m=1}^{M} Q^{AF}(n,m)w^A(n,m) + Q^{BF}(n,m)w^B(n,m)}{\sum_{i=1}^{N}\sum_{j=1}^{M}(w^A(i,j)+w^B(i,j))} \quad (14)$$

Where $w^A(i,j)=[g_A(n,m)]^L$ and $w^B(i,j)=[g_B(n,m)]^L$ weights are function of edge strength. We have considered L is constant 1. The range of Q^{AF} is $0 \le Q^{AF}(n,m) \le 1$. A value of 0 corresponds to the complete loss of edge information, at location (n,m), as transferred from source image A to fused image F.

2) Information Entropy

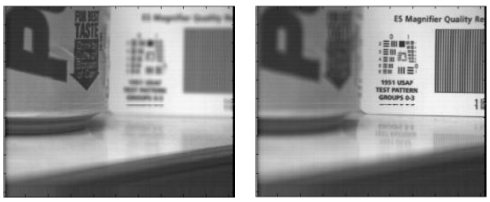
Entropy is an index to evaluate the information quantity contained in an image. The entropy of the fused image F is defined as

$$E = -\sum_{i=0}^{L-1} p_i(f)\log_2 p_i(f) \quad (15)$$

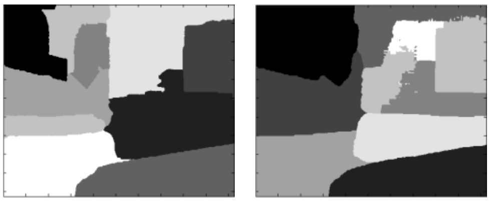
Where p is the normalized histogram of the fused image to be evaluated in our case it is IFUSE, L is maximum value for a pixel in the image which defines the total of grey levels. The entropy issued to measure the overall information in the fused image. The larger the entropy value better the fusion results. The simulation results are discussed in detail in the next section.

IV. Simulation Result And Assessment

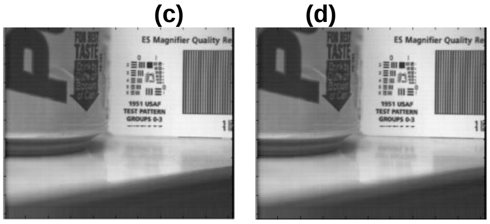
The proposed algorithm has been implemented on Matlab 7 and tested on 20 sets of reference based and 5 sets of nonreference based multifocus images.



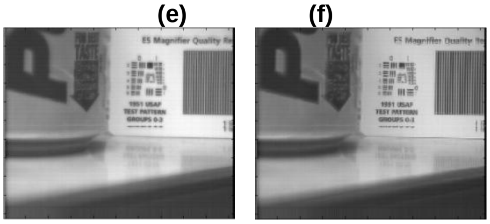
(a) (b)



(c) (d)



(e) (f)

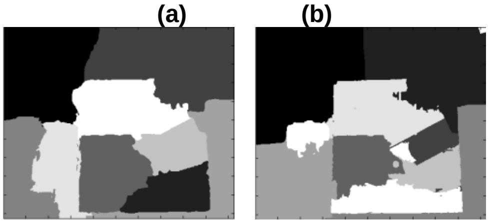


(g) (h)

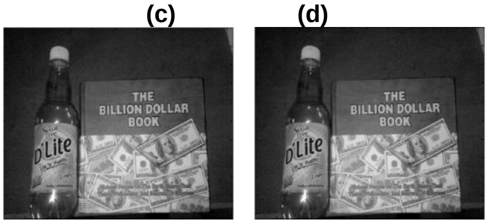
Figure.4 Pepsi Image (a) Input image (IA) (b) Input image (IB) (c) Segmentation on IA (d) Segmentation on IB (e) Final fused image using proposed method (Ifuse) (f) Region based algorithm (g) Fused image using DWT (h) Fused image using CT

Among all set of images results of two tested images are shown in the Figure 3 and 4. First tested image of size 512 x 512 are shown in the Fig.3 (a)-(b) are book image where reference image is available and other tested image of size 512 x 512 are shown in the Fig.4 (a)-(b) Pepsi image where reference image is not available. Proposed algorithm results are compared with three recently proposed algorithms which are simple region based scheme [7], wavelet transform based scheme [2] and contourlet transform based image fusion scheme [3].

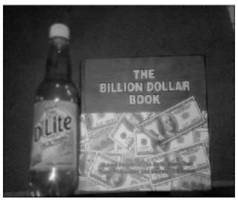
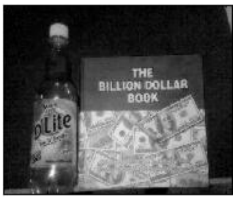
(a) (b)



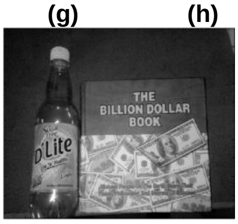
(c) (d)



(e) (f)

(g) **(h)**



(i)

Figure.3 Book image (a) Input image (IA) (b) Input image (IB) (c) Segmentation on IA (d) Segmentation on IB (e) Final fused image using proposed method (Ifuse) (f) Region based algorithm (g) Fused image using DWT (h) Fused image using CT (i) Reference image

To validate the performance of proposed algorithm the results are compared using standard nonreference based and reference based parameters which are depicted in Table I and II respectively. It is clearly seen from simulation results shown in Table I and II that both reference and non reference based evaluation parameters are giving better expected values for proposed algorithm compared to all other three algorithms [2, 3, 6]. It is also important to note that as shown in Table I. RMSE of proposed method is much lower compared to other algorithm which indicates that proposed algorithm fused image is very near to the reference image. These results also indicate that both region based algorithm are preserving more information compared to pixel based algorithm. Our proposed method is producing superior results for all the parameters like MIr, QI, RMSE, SSIM, SF, $Q^{AB/F}$ and MI. The compared methods are analyzed for few parameters in the mention literature [2, 3, 6] from above mention parameters but proposed algorithm is tested for all standard image fusion parameters simultaneously and it is giving better results compared to all other algorithms. It is also tested on many set of images and simulation results are consistent.

TABLE I. REFERENCE BASE PERFORMANCE PARAMETERS

Fusion Parameters	**Book (with reference)**			
	Propose Method	***Region based***	***DWT***	***Contourlet Based***
MIr	7.124	6.710	6.233	6.520
QI	0.997	0.981	0.968	0.949
RMSE	0.983	6.032	7.911	8.178
SSIM	0.991	0.978	0.920	0.935

TABLE II. NON REF. BASED PERFORMANCE PARAMETERS

Fusion Parameters	***Pepsi (without reference)***			
	Proposed Method	**Region based**	**DWT**	**Contourlet Based**
SF	13.770	13.620	11.834	13.150
MI	6.303	6.028	5.110	5.739
QAB/F	0.792	0.779	0.619	0.682

V. CONCLUSION

The proposed method provides powerful two step framework for image fusion which results in good quality fused image for different categories of source images. From the simulation and results, it is evident that our proposed algorithm preserves more detailed fused images compared earlier reported algorithms. The simulation results also show that instead of directly using individual pixels for the fusion process, the region based approach is giving better results. Proposed algorithm also includes other advantages (1) In the method unfocused area of image provides ideal segmentation so user can increase more number of regions to extract more details from the other focus part of the image (2) Proposed method divided into two steps so two different fusion rules are used at two different stages in single algorithm which is helpful to preserve more information in resultant image and increases the robustness of algorithm. (3) It is less sensitive to noise, misregistration and hardly any blurring effect or change of contrast seen. The results can be improved by testing for different fusion rules and parameters. The segmentation algorithm should be applied two times in the algorithm increases computation time is only a disadvantage of proposed scheme. We need to investigate the effect of different segmentation algorithm in resultant fused image. The algorithm can further extend by applying new fusion rules and their combinations using region based image fusion approach.

REFERENCES

[1] G. Piella, "A general framework for multiresolution image fusion: from pixels to regions," Research Report PNA-R0211, CWI, Amsterdam, 2002.
[2] Anna Wang, Jaijining Sun and Yueyang Guan, "The application to wavelet transform to multimodality medical image fusion," IEEE, 2006.
[3] Miao Qiguang and Wang Baoshul,"A novel image fusion method using contourlet transform," Communications, Circuits and Systems Proceedings, 2006 International Conference on, Volume: 1, pp 548-552, 2006.
[4] Gemma Piella, "A region based multiresolution image fusion algorithm," ISIF, IEEE, 2002.
[5] Zhong Zhang and Rick Blum,"A categorization of multiscale decomposition based image fusion schemes with a performance study for digital camera application," Proceedings of IEEE, Aug 1999, Vol. 87, pp 1315-1326.
[6] Shutao Li and Bin Yang,"Multifocus image fusion using region segmentation and spatial frequency," Image and Vision Computing, Elsevier, 26 (2008) 971–979.
[7] J. Shi and J. Malik,"Normalized cuts and image segmentation," IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (8) (2000) 888–905.
[8] C.S. Xydeas and V. Petrovic, "Objective image fusion performance measure," Electronics Letters, Vol. 36, No. 4, February 2000, pp 308-309.
[9] Yin Chen, Rick S. Blum, "A automated quality assessment algorithm for image fusion," Image and vision computing (Elsevier), 2008.
[10] Zheng Liu and Robert Laganiere, "On the use of phase congruency to evaluate image similarity", IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2006, Vol. 2, pp 937-940.
[11] Timothee Cour, Florence Benezit, Jianbo Shi, Multiscale Normalized Cuts Segmentation Toolbox for MATLAB, available at http://www.seas.upenn.edu/ ~timothee.

Tanish Zaveri received his BE degree in Electronics Engineering from Sardar Vallabhbhai Regional College of Engineering, Surat under South Gujarat University in 1998 and obtained his M.Tech. degree in Biomedical Engineering from Indian Institute of Technology, Bombay, India in 2005. He is currently pursuing his Ph.D. in Computer Engineering from Sardar Vallabhbhai National Institute of Technology, Surat.

He is presently working as an assistant professor in Electronics and Communication Engineering Department, Nirma University, Ahmedabad, India. He has more than 10 years of teaching experience. His research interests are mainly focused on multimodality image fusion, biomedical image processing and speech processing. He is a Life member of Indian Society of Technical Education and Institution of Electronics and Telecommunication Engineers. He is also a member of Institute of Electrical and Electronics Engineers and *IACSIT*.

Mukesh A. Zaveri received the B.E. degree in electronics engineering from Sardar Vallabhbhai Regional College of Engineering and Technology, Surat, India, in 1990, the M.E. degree in electrical engineering from Maharaja Sayajirao University, Baroda, India, in 1993, and the Ph.D. degree in electrical engineering from the Indian Institute of Technology, Bombay, Mumbai, in 2005. He is currently an Assistant Professor and the Head of the Computer Engineering Department, Sardar Vallabhbhai National Institute of Technology. His current research interests include the area of signal and image processing, multimedia, computer networks, sensor networks, and wireless communications.

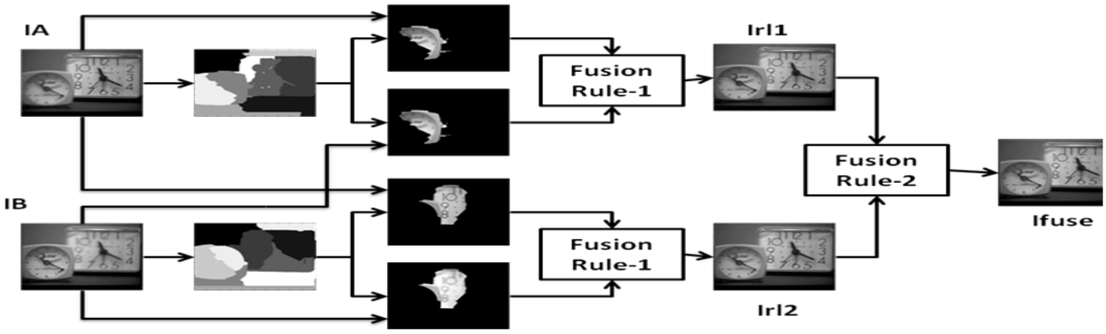


Figure.1 Block Diagram of Proposed Method