# Multifocus Image Fusion based on Human Visual Perception

Rahul Patil
M.E. Student,
Dept. of Electronics and Telecommunication Engg.,
St. Francis Inst. of Technology,
Mumbai, India

Vaqar Ansari
Assistant Professor
Dept. of Electronics and Telecommunication Engg.,
St. Francis Inst. of Technology,
Mumbai, India

Deepak Jayaswal
Professor
Dept. of Electronics and Telecommunication Engg.,
St. Francis Inst. of Technology,
Mumbai, India

## ABSTRACT
The process by which two or more images are merged into a single image is called image fusion, where important characteristics from each of the original image are revived. As images are acquired from different instrument modalities, in order to combine all the capture techniques fusion of image forms a fundamental process. Multifocus image fusion constructs an combined image from multiple source images having focus on different objects from same scene. To achieve this, a spatial domain algorithm is proposed which divides each source image into blocks of sizes varying adaptively. Edge information is extracted from the image by using edge detection techniques. The quality metrics will be obtained for each block, based on human visual perception instead of simple metrics like MSE and PSNR. For the purpose of testing the proposed work, a readily available database of Laboratory for Image and Video Engineering (LIVE) will be used. To demonstrate the quality of the final fused image, evaluation will be done based on the concepts of human visual perception.

## Keywords
Image Fusion, Principal Component Analysis, Pyramid Methods, Discrete Wavelet Transform, Multifocus.

## 1. INTRODUCTION
Image fusion is a sub-field of image processing in which two or more images of a scene are combined into a single composite image that is more informative and is more suitable for visual perception and for digital processing. Images contain a lot of information contained in them. Clarity of the images is very important if all of the information has to be reciprocated [1]. In general, the problem that image fusion tries to solve is to combine information from several images (sensors) taken from the same scene in order to achieve a new fused image, which contains the best information from the various images. With the development of new imaging sensors arises the need of a meaningful combination of all employed imaging sources.

Fusion is a technique to improve the quality of information from a set of images [2]. By the process of image fusion the good information from each of the given image is fused together to form given resultant image whose quality is superior to any of the input images. The objective is to reduce uncertainty and fully utilize complementary and redundant information from the original images. Moreover the goal is to combine the original multiple images to produce a more precise, comprehensive and reliable image interpretation of the scene. The aim is to integrate multiple images of the same scene into a composite image so that the new image is more suitable for visualization, detection and recognition tasks [3].

Taking this into consideration multifocus image fusion can be defined as a process of combining several images with different focus into one uniformed focused image. In simple words an all in-focus image has to be acquired from different focal planes of the various source images and fusing them together into one single image where all objects in the scene appear to be in focus. A drawback is that multifocus image fusion involves processing and storing of scaled data which are of same size as the original image, which results in a huge amount of memory and time requirement. The various applications include military, medical, machine learning and remote sensing, automatic change detection, biometrics etc. Generally , image fusion is divided into basic three levels: pixel level fusion, feature level fusion, and decision level fusion[4][5].In Pixel fusion is the lowest-level fusion, which analyses and integrates the information before the original information is estimated and recognized. In the feature fusion is done in the middle level, which analyzes and deals with the feature information such as edge, contour, direction obtained by pre-treatment and feature extraction. Decision level is a higher level of fusion where input images are processed individually for information extraction and the information combined by applying decision rules to reinforce common interpretation [6].

Section II discusses the proposed image fusion techniques that have been elaborated in this paper. Section III describes the various performance analysis techniques for image fusion followed by various criteria of checking the performance of result.

## 2. PROPOSED TECHNIQUES
### 2.1 Fixed Block Size Adaptive Threshold Method (FBS-AT)
In this algorithm (FBS -AT), selection is made in three iterations described as follows: The source images are divided into a certain number of blocks. Then, the difference between edge information from the two source images is computed for each block. Next, the mean of all these differences is calculated, this mean is set as the adaptive threshold. (T). The differences are compared with this threshold T and only those blocks for which the difference exceeds the threshold are chosen and incorporated into the final fused image from their corresponding source image [10].

The rest of the blocks are passed on to the next iteration. In the second iteration, a new adaptive threshold is set by calculating the mean of the differences of the regions which were passed over from the last iteration [10]. Repeat again, the difference between the numbers of edge pixels for corresponding image block from different source images, is compared with the threshold, and if the difference is higher than the threshold then the respective block with higher edge information is fuse into the fused image. In the third iteration, all the blocks for which no decision has been made are analyzed and the blocks with relatively higher edge information is selected to be part of the fused image [10].

| | | | |
|---|---|---|---|
| 127 | 715 | 1398 | 1222 |
| 1203 | 1545 | 992 | 1715 |
| 1376 | 1814 | 1219 | 1482 |
| 1111 | 897 | 275 | 429 |

| | | | |
|---|---|---|---|
| 0 | 950 | 2471 | 1525 |
| 646 | 1279 | 3324 | 2233 |
| 712 | 1609 | 3134 | 1910 |
| 521 | 532 | 389 | 434 |

Average threshold=611.15

| | | | |
|---|---|---|---|
| 127 | 235 | 1073 | 303 |
| 557 | 268 | 2332 | 518 |
| 664 | 205 | 1915 | 428 |
| 590 | 365 | 114 | 05 |

**Figure 1. Depicts the process of computing adaptive threshold and getting the average threshold which is the deciding parameter for the incorporation of blocks from the input images into the output image**

## 2.2 Adaptive Block Size Adaptive Threshold Method (ABS-AT)

This algorithm is an enhancement to the FBS-AT algorithm. The basis of this enhancement was that if the images are divided into blocks of different sizes, means reduce the block size it may give different results. The algorithm goes as follows: The first iteration is carried out in the same way as described in FBS – AT. In the next iteration, the image is divided such that each block is subdivided by twice the number of divisions used in last iteration, i.e. each block of last iteration will be considered as 4 separate blocks. The mean of the differences of edge information from the two source images of these blocks is calculated and set as the new threshold [10]. The regions for which the adaptive threshold criteria is met are incorporated into the final fused image and remaining blocks are passed over to the next iteration. The upper bound on maximum number of divisions and/or minimum block size is set as a control parameter to conclude these iterations and move on to the next stage. At the end of all the iterations of step 2, the blocks for which no decision has been made are analyzed simply by comparing number of respective edge pixels, i.e., for each of these left-over regions, information is taken from the source image which contains higher edge information in that area [10].

## 2.3 Image fusion using adaptive thresholding and cross filtering

Select a source image clicked in two ways that is, having foreground in focus in the first image and having the background in focus in the second image. The images are then subtracted row wise and column wise. In the previous described algorithms, the images are cut in a definite way. In the FBS-AT algorithm, the images are cut in a definite number of parts, that is, 4 parts. This process was a bit modified in the second algorithm, the ABS-AT algorithm, where for every next iteration, every section was divided into the (previous number of sections *4) subsections. Cutting the images row wise and column wise and obtaining the minima is a process which hasn't been done until now and it optimizes the entire segmentation process. As a result, get a single matrix consisting of the row values and column values, then by summing column values it can achieve minima for column values only. Then by transposing that single matrix obtained earlier it can also achieve minima for row values by summing up the column values. After this procedure the minima of the row and column matrix obtained is then selected for cutting the two images respectively.

| | | | |
|---|---|---|---|
| 2 | 1 | 5 | 7 |
| 3 | 2 | 0 | 2 |
| 0 | 0 | 1 | 0 |
| 4 | 1 | 6 | 2 |
| **9** | **4** | **12** | **11** |

**Figure 2a**

| | | | |
|---|---|---|---|
| 2 | 3 | 0 | 4 |
| 1 | 2 | 0 | 1 |
| 5 | 0 | 1 | 6 |
| 7 | 2 | 0 | 2 |
| **15** | **7** | **1** | **13** |

**Figure 2b**

**Figure 2a: summing operation for column value**
**Figure 2b: summing operation for row value**

Here it has need to perform summing operation column wise and the resultant is obtained below the fig 2 (a) matrix. As seen in the above fig 2 (a), out of all the 4 values obtained i.e. 9, 4, 12,11 of column number 1,2,3,4 respectively, and getting the minima value (4) for column 2, so select column number 2. Similarly performing the transpose of the matrix in fig 2 (a) and then obtaining the row values by performing the summing operation for the column matrix as in fig 3.10 (b). Now in fig 2 (b), out of all the 4 values obtained i.e. 15, 7,1,13 hence getting the minima value of (1) for column number 3. Now select the cutting parameter as (3, 2), as shown in the below figure by the concept of minima it can achieve smooth blending which can overcome the drawback of $2^{nd}$ algorithm

| | | | |
|---|---|---|---|
| 3 | 4 | 7 | 9 |
| 4 | 3 | 1 | 4 |
| 1 | 3 | 5 | 1 |
| 1 | 2 | 1 | 1 |

| | | | |
|---|---|---|---|
| 1 | 3 | 2 | 2 |
| 1 | 1 | 1 | 2 |
| 1 | 3 | 4 | 1 |
| 5 | 3 | 7 | 3 |

**Figure 3. Depicts the differences obtained after subtracting the source images row wise and column wise. The minima from these is selected and the images are then cut in this way**

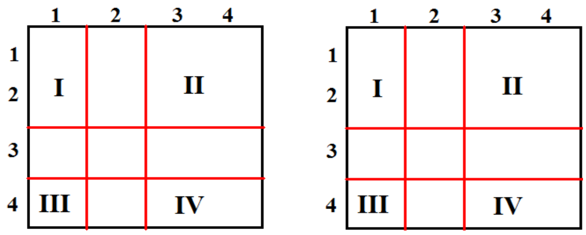


**Fig 4.a** **Fig 4.b**

**Figure 4a: minima section in first source image**
**Figure 4b: minima section in second source image**

As seen in the above our original images are cut into 4 parts ie. I, II, III, IV for respective two original images, take same part of both images ie. first part of fig 4(a) and first part of fig 4(b), used as two different images and then apply Adaptive Block size and Adaptive Threshold (ABS-AT) method and put the decision in final fused image.

| | | | |
|---|---|---|---|
| 2 | 1 | 5 | 7 |
| 3 | 2 | 0 | 2 |
| 0 | 0 | 1 | 0 |
| 4 | 1 | 6 | 2 |
| **9** | **4** | **12** | **11** |

| | | | |
|---|---|---|---|
| 2 | 3 | 0 | 4 |
| 1 | 2 | 0 | 1 |
| 5 | 0 | 1 | 6 |
| 7 | 2 | 0 | 2 |
| **15** | **7** | **1** | **13** |

**Figure 5. Obtaining different maxima and minima**

| | | | |
|---|---|---|---|
| 2 | 255 | 5 | 7 |
| 3 | 255 | 0 | 2 |
| 0 | 255 | 6 | 2 |
| 4 | 255 | 1 | 0 |
| **9** | **1020** | **12** | **11** |

| | | | |
|---|---|---|---|
| 1 | 2 | 255 | 4 |
| 2 | 3 | 255 | 1 |
| 5 | 0 | 255 | 6 |
| 7 | 2 | 255 | 2 |
| **15** | **7** | **1020** | **13** |

**Figure 6. Replacing minima obtained from previous with maxima**

After doing this procedure, the minima are replaced by a maxima and new minima are selected. As shown in fig.6 the all minima column value set as maximum value because of that it will not select again for iterative process. This marks the end of iteration one. This procedure is carried out for II, III, and IV parts respectively for two images and then whichever image is suitable as per decision between two images is taken as the final fused image. This process is carried out up to 100 iterations. Once the final output is obtained, cross filtering is done. The final image is then compared with the two source images. By this we get to know that significant data from which of the two source images is inculcated in the final image. Some regions of the final image are then replaced by the same regions from the source images. This improves the smoothness of the image.Bilateral filtering is a local, nonlinear and no iterative technique which combines a classical low-pass filter with an edge-stopping function that attenuates the filter kernel when the intensity difference between pixels is large. As both gray level similarities and geometric closeness of the neighbouring pixels are considered, the weights of the filter depend not only on Euclidian distance but also on the distance in gray/color space. The advantage of the filter is that it smoothes the image while preserving edges using neighbouring pixels. $A(x,y)$ is the unrecognized details of image A obtained by subtracting image A from fused image by algorithm of minima and $B(x,y)$ is the unrecognized details of image B.

$I(x,y)$ is the output of algorithm minima method and $I_1(x,y)$ original image 1(background in focus)and $I_2(x,y)$ original image2 (foreground in focus).

$$A(x,y) = I_1(x,y) - I(x,y) \quad (1)$$

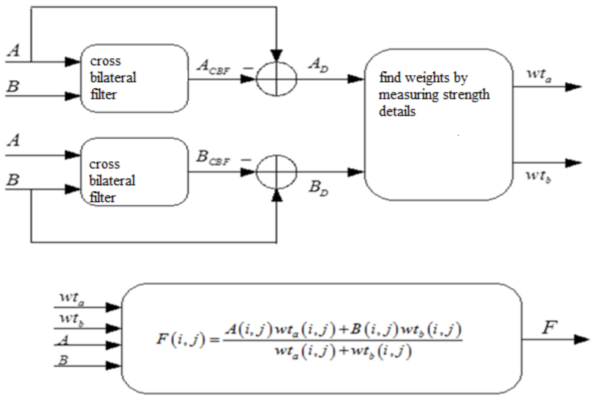


**Figure 7. Depicts a bilateral cross filter [12]**

$$B(x,y) = I_2(x,y) - I(x,y) \quad (2)$$

where, $I(x,y)$ - output of minima method

$I_1(x,y)$ - original image 1(background in focus)

$I_2(x,y)$ - original image2 (foreground in focus).

$$C(x,y) = h_{GHPF}(x,y) * \mathrm{A}(x,y) \quad (3)$$

where, $h_{GHPF}$ is the transfer function of Gaussian high pass filter.

$$C(x,y) = \sum_{i=1}^{x}\sum_{j=1}^{y} h_{GHPF}(x-i, y-j) * \mathrm{A}(\mathrm{i},\mathrm{j})$$

$$|B(x,y) - C(x,y)| = A_D(x,y)$$

where, $A_D(x,y)$ is the detail of image A

$$D(x,y) = h_{GHPF}(x,y) * \mathrm{B}(x,y)$$

$$D(x,y) = \sum_{i=1}^{x}\sum_{j=1}^{y} h_{GHPF}(x-i, y-j) * \mathrm{B}(\mathrm{i},\mathrm{j})$$

(4)

$$|\mathrm{A}(x,y) - D(x,y)| = B_D(x,y)$$

where, $B_D(x,y)$ is the detail of image B

$$A_D(x,y).k_1 = wt_a(x,y)$$

$$B_D(x,y).k_1 = wt_b(x,y)$$

$k_1$ and $k_2$ are the constants

$$h_{GHPF} = 1 - e^{-D^{2(x,y)/2\sigma^2}}$$

$$F(i,j)=\frac{A(x,y)wt_a(x,y)+B(x,y)wt_b(x,y)}{wt_a(x,y)+wt_b(x,y)} \quad (5)$$

$F(i,j)$ is the final output image

# 3. PERFORMANCE EVALUATION

The image quality evaluation methods that already exist can be divided into subjective and objective analysis. Subjective analysis basically involves taking the opinion of human observers involving the perceptual appeal of the fused image. The objective performance parameters are as follow:

1. *Average Pixel Intensity (μ or F):* It is an index of contrast of the images

$$\mu=\bar{F}=\frac{\sum_{i=1}^{m}\sum_{j=1}^{n} f(i,j)}{m \times n} \quad (6)$$

Here $f(i,j)$ is pixel intensity for position $(i,j)$ of image $F$

2. *Average Gradient (G):* It is a measure of Sharpness and Clarity degree of the image.

$$\bar{G}=\frac{\sqrt{\sum_i \sum_j (f(i,j)-f(i+1,j))^2+(f(i,j)-f(i,j+1))^2}}{m \times n} \quad (7)$$

3. *Standard Deviation (SD or σ):* It basically reflects the spread in data in the image. This metric is more efficient in the absence of noise. It measures the contrast in the fused image. ie. An image with high contrast would have a high standard deviation. The standard deviation (SD), which is the square root of variance, reflects the spread in the data. Thus, an image with high contrast would have a high standard deviation, and a low contrast image would have a low standard deviation.

$$\frac{\sum_{i=1}^{m}\sum_{j=1}^{n}\left(f(i,j)-\bar{F}\right)^2}{mn} \quad (8)$$

where m x n is the size of the image.

4. *Entropy (H):* This parameter evaluates the quantity of information present in the image. It is basically an index to evaluate the information quantity contained in an image. If the value of entropy becomes higher after fusing, it indicates that the information increases and the fusion performances are improved. Entropy is represented by the mathematical formula:

$$H=\sum_{f=0}^{G} p(i)\log_2\{p(i)\} \quad (9)$$

where $G$ is the number of gray levels in the image's histogram (255 for a typical 8-bit image), and $p(i)$ is the normalized frequency of occurrence of each gray level.

5. *Mutual Information (MI) or Fusion Factor:* It is the measure of the correlative information content in fused images with respect to source images. This is very similar to mutual information. Suppose let us consider A and B be the source images and let F be the fused image. When no reference images are available, fusion assessment is performed as follows:

$$MI_{AF}=\sum_a\sum_f p_{A,F}(a,f)\,log_2\frac{p_{A,F}(a,f)}{p_A(a)\,p_F(f)} \quad (10)$$

$$MI_{AB}^{F}=MI_{AF}+MI_{BF}$$

where $MI_{AF}$ and $MI_{BF}$ quantify mutual information between source A and fused image F and source image B and fused image F respectively. $MI_{AB}^{F}$ is the overall mutual information between source images and fused image. Larger value of mutual information gives the better fusion results.

6. *Fusion Symmetry (FS):* It is an indication of how symmetric the output image is with the input image. If the final fused image is equally symmetric to both the source images, value of fusion symmetry will be closer to 2 and the quality of fusion will be better

$$FS=2-\left|MI_{AF}/(MI_{AF}+MI_{BF})-0.5\right| \quad (11)$$

7. *Normalized Correlation (CORR):* It is a measure of relevance of fused image to source images.

$$r_{AF}=\frac{\sum_i\sum_j(a(i,j)-\bar{A})(f(i,j)-\bar{F})}{\sqrt{(\sum_i\sum_j(a(i,j)-\bar{A})^2)\sum_i\sum_j(f(i,j)-\bar{F})^2)}} \quad (12)$$

$$CORR=(r_{AF}+r_{BF})/2$$

here raf and rbf represents normalized between source images and fused image and correlation stands for overall average normalized correlation.

8. *Petrovic Metric Parameter QABF:* It is an index of edge information preservation.

9. *Petrovic Metric Parameter LABF:* It is a measure of loss of edge information.

10. *Petrovic Metric Parameter NABF:* It is a measure of noise.

# 4. DISCUSSION

The database used for fusion purpose which has been made freely available for research purposes like laboratory for image and video engineering (LIVE) [11], imagefusion.org. After experimenting several standard test pair of multifocus images, finally calculated the Average Pixel Intensity, standard deviation, entropy, average gradient, mutual information, fusion symmetry, normalized correlation and petrovic metric parameter and all these parameters have been represented in Table 2 also in graphical form shown in Figure 9a, 9b, 9c. In the multifocus image fusion there is various standard test pairs (512 X 512) of multifocus images which were provided by online resource for research in image fusion (http://www.imagefusion.org). As shown in above table, the parameters for performance evaluation have been calculated for five standard test pairs .Average pixel intensity for pen test pair is maximum in comparison with the other test pairs. The result achieved is better in our algorithm than existing ones which are discussed below for fusion symmetry, correlation and petrovic metric parameters other than first five parameter.

**Figure 8a. Foreground in focus Figure 8b. Background in focus**

**Figure 8c. Final fused image**

## 5. CONCLUSION

The proposed fusion technique is well suited for fusion of multifocus images in spatial domain. Through the experiments conducted on standard test pairs of multifocus images, it was found that the proposed method has shown superior performance in most of the cases as compared to previous methods in terms of subjective and objective analysis parameters. The major achievements of the proposed method are minimum artifacts (lowest NABF) and maximum edge preservation (highest QABF). This is a significant achievement, as artifacts may lead to wrong interpretations which can be catastrophic, especially in applications like surveillance where it can result into false alarms.

For future research remaining the performance of proposed method could be improved by exploring entropy, symmetry between source images and fused image, correlation and standard deviation compared to other methods. As per our project objective i.e. aimed to achieve a multifocus image that is clear for human visual perception and that has been achieved.

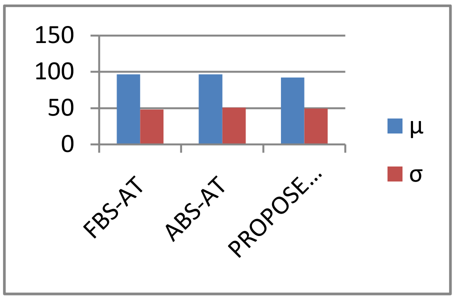


**Figure 9a.**

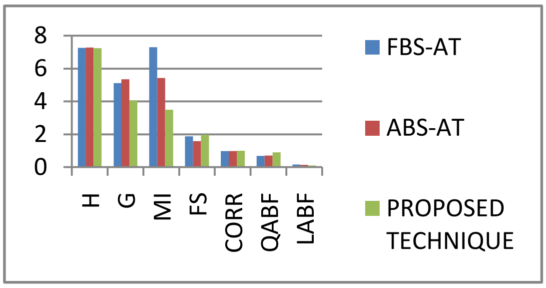


**Figure 9b.**

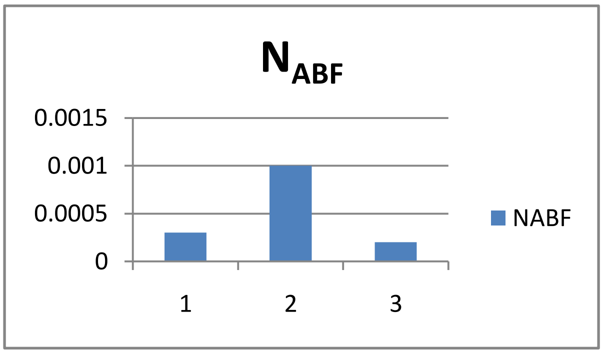


**Figure 9c.**

**Figure 9. Multiple bar graph representing the values for different performance parameters**

**Table 1: Performance comparison of fusion results for 'clock' pair**

| | **μ** | $\sigma$ | **H** | **G** | **MI** | **FS** | **CORR** | $Q_{ABF}$ | $L_{ABF}$ | $N_{ABF}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| **FBS-AT** | 96.646 | 47.941 | 7.263 | 5.126 | 7.3 | 1.872 | 0.978 | 0.69 | 0.161 | 0.0003 |
| **ABS-AT** | 96.671 | 50.739 | 7.278 | 5.34 | 5.425 | 1.582 | 0.978 | 0.705 | 0.141 | 0.001 |
| **Proposed Technique** | 92.40 | 49.25 | 7.24 | 4.06 | 3.50 | 1.9625 | 0.9880 | 0.8968 | 0.1028 | 0.0002 |

# 6. REFERENCES

[1] S. Krishnamoorthy, K. Soman, Implementation and Comparative Study of Image Fusion Algorithms. International Journal of Computer Applications. November 2010, 25-35.

[2] O. Prakash, R. Srivastava, A. Khare, Biorthogonal wavelet transform based image fusion using absolute maximum fusion rule Information & Communication Technologies (ICT). April 201,577-582.

[3] U. Patil, U. Mudengudi. Image fusion using hierarchical PCA," Image Information Processing (ICIIP), 2011 International Conference. November 2011,1-6.

[4] A. Saleem, A. Beghdadi, B. Boashash, Image quality metrics based multifocus image fusion Visual Information Processing (EUVIP), 2011 3rd European Workshop. July 2011, 77-82.

[5] V. Petrovic, C. Xydeas, Gradient-based multiresolution image fusion Image Processing. February 2004, 228-237.

[6] A. Saleem, A. Beghdadi, B. Boashash, Image quality metrics based multifocus image fusion Visual Information Processing (EUVIP), 2011 3rd European Workshop. July 2011,77-82.

[7] Sahu, D.Kumar, and M.Parsa, Different Image Fusion Techniques–A Critical Review. International Journal of Modern Engineering Research (IJMER).October 2012, 4298-4301.

[8] S. Bedi, R. Khandelwal, Comprehensive and Comparative Study of Image Fusion Techniques International Journal of Soft Computing and Engineering (IJCSE). March 2013, 300-304.

[9] H. Kekre, D. Mishra and R. Saboo, Review on image fusion techniques and performance evaluation parameter International Journal of Engineering Science and Technology (IJEST). April 2013, 881-889.

[10] P. Shah, A. Kumar, S. Merchant. U. Desai, Multifocus image fusion algorithm using iterative segmentation based on edge information and adaptive threshold Information Fusion (FUSION), 2012 15th International Conference. July 2012, 1976-1981.

[11] Laboratory for image and video engineering, [Online] Available: http://live.ece.utexas.edu/research/quality/

[12] B. K. Shreyamsha Kumar, Image fusion based on pixel significance using cross bilateral filter Journal on Signal, Image and Video Processing, October 2013, 1-12.