

# An Image Fusion Technique Based on Hadamard Transform and HVS

Radhika Vadhi

Department of ECE,  
University College of Engineering,  
JNTUK, Kakinada,  
Andhra Pradesh, India.  
radhikav139@gmail.com

Veera Swamy Kilari

Department of ECE,  
QIS College of Engineering and  
Technology, Ongole,  
Andhra Pradesh, India.  
kilarivs@yahoo.com

S. Srinivas Kumar

Department of ECE, University  
College of Engineering, JNTU,  
Kakinada,  
Andhra Pradesh, India.  
samay\_ssk2@yahoo.com

**Abstract**—The main endeavor of image fusion is to obtain an image that contains more visual quality information than any one of the source images. In general, the source images may be multi focus, multi modality, multi resolution, multi temporal, panchromatic, satellite images considered for fusion. This paper discusses image fusion using Hadamard Transform (HT). In this work, Human Visual System (HVS) is investigated for image fusion in the HT domain. The proposed fusion process contains three important parts, (1) divide source images into sub images / blocks and transform them into HT domain. (2) multiply transformed coefficients with HVS based weightage matrix of HT and select the highest value from them (3) fuse the corresponding block of selected coefficients from source images in to an empty image. The utility of HVS makes the coefficients more significant. The performance of the proposed method is analyzed and compared with Discrete Wavelet Transform (DWT) based image fusion technique. Implementation in HT domain is simple and time saving when compared with DWT.

**Keywords**—Hadamard Transform; HT; human visual system; HVS; discrete wavelet transform; DWT; image Fusion

## I. INTRODUCTION

Image fusion is a process of combining information from two or more source images of a scene in to a single composite image that is more informative and more suitable for visual perception or machine processing. The concept of image fusion has been used in a wide variety of applications like medicine, remote sensing, machine vision, automatic change detection, biometrics, robotics, microscopic vision etc. An image fusion algorithm [1-6] should require entire salient information contained in the input images. Fusion process should not introduce any artifacts or inconsistencies which can distract images. Image fusion must be reliable, robust, tolerant to noise or misregistrations.

Generally fusion algorithms can be categorized into spatial domain fusion and transform domain fusion. The spatial domain fusion algorithms can use the local spatial features such as mean, variance, gradient, spatial frequency, and local standard deviation. The spatial domain techniques are simple

and less complex, but not robust. Because of this, multi resolution transforms are used to represent the sharpness and edges of an image in the design of fusion algorithms. The transformed coefficients provide the information of an image that can be used to select the blocks from source images to the fusion image. Wavelet Transforms have been successfully used in fusion algorithms due to their advantages such as increased directional information and the absence of blocking artifacts that often occur in Discrete Cosine Transforms (DCT) and Hadamard Transform (HT). However, Wavelet Transform implementation takes more time and the process is more complex. A major problem with DWT is its shift variant nature caused by sub-sampling which occur at each level. A small shift in the input signal results in a completely different distribution of energy between DWT coefficients at different scales [16-21]. To overcome the shift variant problem, block based transforms are preferred. Human Visual System (HVS) is also used to suppress the blocking artifacts. Hence, HT with HVS is explored for image fusion in this manuscript. HVS model is easy to adapt to the specified resolution for viewing. Implementation of HVS models with block based transform is easy than multi-resolution transforms like DWT, Curvelet, and Contourlet.

The elements of the basis vectors of the HT [7] take only the binary values  $\pm 1$  and are, therefore, well suited for digital hardware implementations of image processing algorithms. Hadamard Transform offers a significant advantage in terms of a shorter processing time as the processing involves simple integer manipulation (compared to floating point processing with DCT and DWT) and easier hardware implementation than any common transform techniques. So it is computationally less expensive than many other orthogonal transforms.

This paper is organized as follows. The HVS model and design considerations are discussed in section II. Computation of 2D-HT and its inverse is discussed in section III. The proposed fusion algorithm using HVS model is presented in section IV. The experimental results are given in section V and the conclusions in section VI.

II. MATHEMATICAL ANALYSIS

A. Human Visual System

The HVS has been investigated by several researchers [8-11]. The visual sensitivity and selectivity are the important requirements for the development of HVS. The HVS is based on the psychophysical process that relates psychological phenomena (luminance, contrast and brightness etc.) to physical phenomena (light sensitivity, spatial frequency and wavelength etc.) [9]. The HVS is complicated, as it is a nonlinear and spatial varying system. Putting its multiple characteristics into a single equation, especially one that is linear, is not an easy task. The work in [8] may be first breakthrough to incorporate the HVS in image coding. HVS as a nonlinear point transformation followed by the modulation transfer function (MTF) is given by

$$H(f) = a(b + cf) \exp(-c(f)^d) \quad (1)$$

where,  $f$  is the radial frequency in cycles/degree of the visual angle subtended and  $a, b, c$  and  $d$  are constants. HVS model proposed in [15] is applied for generating the quantization table for image compression applications. This HVS model is a modified version of the one in [8] with  $a=2.2, b=0.192, c=0.114$  and  $d=1.1$  respectively. The MTF of HVS has been successfully applied to the optimal image half toning [9, 10]. The MTF as reported in [15] is

$$W(u,v) = \begin{cases} 2.2 \left[ 0.192 + 0.114 f(\hat{u},v) \right] \exp(-0.114 f(\hat{u},v)^{1.1}) & \text{if } \hat{f}(u,v) > f; \\ 1.0 & \text{otherwise} \end{cases} \quad (2)$$

where  $(u,v)$  is radial spatial frequency in cycles/degree and  $f$  is the frequency of 8 cycles/degree at which the exponential peak. To implement this, it is necessary to convert discrete horizontal and vertical frequencies,  $\{f(u), f(v)\}$  into radial visual frequencies. For a symmetric printing grid, the horizontal and vertical discrete frequencies are periodic and given in terms of the dot pitch  $\Delta$  and the number of frequencies  $N$  by:

$$f(u) = \frac{u-1}{\Delta N}, f(v) = \frac{v-1}{\Delta N} \quad (3)$$

Converting these to radial frequencies, and scaling the result to cycles/visual degree for a viewing distance (dis) in millimeters gives

$$f(u,v) = \frac{\pi}{180 \arcsin \frac{1}{\sqrt{1+dis^2}}} \sqrt{f(u)^2 + f(v)^2} \quad (4)$$

Finally, to account for variations in visual MTF as a function of viewing angle,  $\theta$ , these frequencies are normalized by an angular-dependent function,  $s(\theta(u,v))$ , such that

$$\hat{f}(u,v) = \frac{f(u,v)}{s(\theta(u,v))} \quad (5)$$

Where,  $s(\theta(u,v))$  is given in [15] as:

$$s(\theta(u,v)) = \frac{1-\omega}{2} \cos(4\theta(u,v)) + \frac{1+\omega}{2} \quad (6)$$

B. Hadamard Transform

In series form the transform pair becomes Forward Hadamard Transform for a 2-D signal  $f(m,n)$  is defined as:

$$F(k,l) = \frac{1}{N} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} f(m,n) (-1)^{\sum_{i=0}^{N-1} b_i^{(m)} b_i^{(k)} + b_i^{(n)} b_i^{(l)}} \quad (7)$$

where  $0 \leq m, 0 \leq n, 0 \leq N-1, k, l = 0, 1, \dots, N-1$

The Inverse Hadamard Transform is defined as:

$$f(m,n) = \frac{1}{N} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} F(k,l) (-1)^{\sum_{i=0}^{N-1} b_i^{(m)} b_i^{(k)} + b_i^{(n)} b_i^{(l)}} \quad (8)$$

where  $m, n = 0, 1, \dots, N-1$ .  $b_i(l)$  represents the  $i^{\text{th}}$  (from LSB) bit of the binary value of  $n$  decimal number represented in binary.

The HT [7] has been used in image processing, image compression, and image fusion. Let  $f(m,n)$  represents the original image and  $F(k,l)$  the transformed image. The HT is given by

$$F(k,l) = \frac{H_n[f(m,n)]H_n}{N} \quad (9)$$

Where  $H_n$  represents an  $N \times N$  Hadamard matrix, with element values either +1 or -1. The Hadamard transform is real, symmetric, and orthogonal that is

$$H_n = H_n^* = H_n^t = H_n^{-1} \quad (10)$$

The inverse transformation is

$$f(m,n) = \frac{H_n[F(k,l)]H_n}{N} \quad (11)$$

The Hadamard matrix of the order  $n$  is generated in terms of Hadamard matrix of order  $n-1$  using kronecker product ' $\otimes$ ' given by

$$H_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ H_n = H_{n-1} \otimes H_1 \quad (12)$$

HT matrix has its AC components in a random order. The processing is performed based on the  $8 \times 8$  sub-blocks of the whole image, the third order HT matrix  $H_3$  is used. By applying  $H_3$  becomes:

$$\frac{1}{\sqrt{8}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{bmatrix} \quad (13)$$

III. PROPOSED METHOD

Given Hadamard matrix, the number of sign changes in each row of the Hadamard Transform matrix is given as 0,7, 3, 4, 1, 6, 2 and 5 in the rows 1 to 8 respectively. The number of sign changes in each column of the Hadamard Transform matrix is given as 0, 7, 3, 4, 1, 6, 2 and 5 in the columns 1 to 8 respectively. The number of sign changes is referred to as sequence. The concept of sequence is analogous to frequency for the Fourier transform. Therefore,  $R = [0 \ 7 \ 3 \ 4 \ 1 \ 6 \ 2 \ 5]$ ,  $C=[0 \ 7 \ 3 \ 4 \ 1 \ 6 \ 2 \ 5]$ . The horizontal and vertical discrete frequencies in the Hadamard domain are given in Equation (12)

$$f(u) = \frac{R(u)}{\Delta * 2N} \quad \text{for } u=1,2,\dots,N$$

$$f(v) = \frac{C(v)}{\Delta * 2N} \quad \text{for } v=1,2,\dots,N \quad (14)$$

The dot pitch ( $\Delta$ ) of the high resolution computer display is about 0.25mm. High resolution computer display is about 128 mm height and 128 mm width to display a 512x512 pixel image. The appropriate viewing distance is four times of height. Hence, distance is considered as 512 mm. Constant  $\omega$  is a symmetric parameter, derived from experiments and set to 0.7. Thus, the human visual frequency weighting matrix  $w(u, v)$  of (2) is calculated for HT using (3) and (4) as given in Table I [11].

TABLE I. HVS WEIGHTING MATRIX FOR HT

1.0000	0.6571	1.0000	0.9599	1.0000	0.7684	1.0000	0.8746
0.6571	0.1391	0.4495	0.3393	0.6306	0.1828	0.5558	0.2480
1.0000	0.4495	0.7617	0.6669	1.0000	0.5196	0.8898	0.5912
0.9599	0.3393	0.6669	0.5419	0.9283	0.930	0.8192	0.4564
1.0000	0.6306	1.0000	0.9283	1.0000	0.7371	1.0000	0.8404
0.7684	0.1828	0.5196	0.4940	0.7371	0.2278	0.6471	0.2948
1.0000	0.5558	0.8898	0.8192	1.0000	0.6471	0.9571	0.7371
0.8746	0.2480	0.5912	0.4564	0.8404	0.2948	0.7371	0.3598

The human visual frequency weighting matrix  $w(u, v)$  indicates the perceptual importance of the transform coefficients. After multiplying the 64 Hadamard coefficients with human visual frequency weighting matrix the weighted Hadamard coefficients contribute the same perceptual importance to human observers.

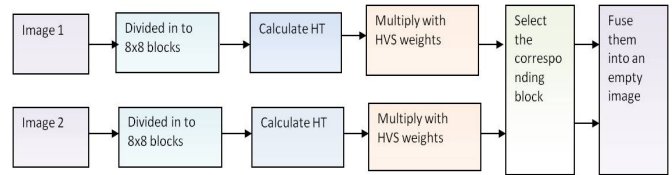


Fig. 1. Experimental results with different images

The fusion algorithm is given as:

- Both the input images are first divided into non overlapping 8 x 8 blocks.
- For each block apply HT using (7) results  $T(u, v)$ .
- All the transform coefficients are multiplied with weighting matrix  $w(u, v)$  results  $R(u, v)$ .

$$R(u, v) = \sum_{i=1}^8 \sum_{j=1}^8 T(u, v)w(u, v) \quad (15)$$

- Select the corresponding block from each source images, according to the highest  $R(u, v)$

$$T_F(u, v) = \begin{cases} T_1(u, v) & R_{B1}(u, v) \geq R_{B2}(u, v) \\ T_2(u, v) & R_{B1}(u, v) < R_{B2}(u, v) \end{cases}$$

Where  $R_{B1}(u, v)$  and  $R_{B2}(u, v)$  are the corresponding blocks with highest  $R(u, v)$  from each source images.  $T_F(u, v)$  is the selected block from source images for fusion.

- Fuse the corresponding blocks in to the empty image.

IV. EXPERIMENTAL RESULTS

Three multi focused images from [16] are used as source images. To compare the results, DWT method with maximum selection rule is tested [3].

$$MSE = \frac{1}{(n-1)(m-1)} \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} [s(i, j) - f(i, j)]^2 \quad (16)$$

where n is the number of total pixels and  $s(i,j)$  and  $f(i,j)$  are the pixel values in the original and fused image.

$$PSNR = 10 \log_{10} \left( \frac{255^2}{MSE} \right) \quad (17)$$

The Mutual information (MI) [13] is also one important measure to test the fused image quality. Mutual information between the source images  $s_1, s_2$  and the fused image  $f$  is

$$M_f^{S_1 S_2} = I_{S_1 f} + I_{S_2 f} \quad (18)$$

Consider  $k$  can be 1 or 2, then  $I_{S_k f}$  is

$$I_{S_k f}(f, s_k) = \sum_{f, s_k} P_{S_k f}(f, s_k) \log \frac{P_{S_k f}(f, s_k)}{P_f(f)P_{s_k}(s_k)}$$

where  $P_{S_k f}$  is the jointly normalized histogram.

The Edge Strength and Orientation Preservation [14] values can be calculated as

$$Q^{S_1 S_2 f} = \frac{\sum_{n=1}^N \sum_{m=1}^M (Q^{S_1 f}(n,m)w^{S_1}(n,m) + Q^{S_2 f}(n,m)w^{S_2}(n,m))}{\sum_{n=1}^N \sum_{m=1}^M (w^{S_1}(n,m) + w^{S_2}(n,m))} \quad (19)$$

$Q^{S_k f}(n,m)$  represents the edge strength and orientation strength,  $w$  represents weight.

Feature Similarity (FSIM) [12, 18] Index can be measured as:

$$FSIM = \frac{\sum_{x \in \Omega} S_L(x) \cdot PC_m(x)}{\sum_{x \in \Omega} PC_m(x)} \quad (20)$$

Where  $\Omega$  means the whole image spatial domain,  $PC_m$  means Phase congruent structure,  $S_L$  is the similarity depends upon the gradient measure.

The usable gray level values range from 0 to 255. The other quality parameter is Normalized Cross Correlation (NCC) is given as:

$$NCC = \frac{\sum_i \sum_j (s(i,j) - \bar{s}(i,j))(f(i,j) - \bar{f}(i,j))}{\sqrt{\left[ \sum_i \sum_j (s(i,j) - \bar{s}(i,j))^2 \right] \left[ \sum_i \sum_j (f(i,j) - \bar{f}(i,j))^2 \right]}} \quad (21)$$

Where,  $\bar{s}(i,j)$  indicates the mean of the original image and  $\bar{f}(i,j)$  indicates the mean of the fused image. PSNR and NCC are commonly used as a measure of quality of fused image with respect to the ground truth image. Typical PSNR values are 20 to 50 dB, where higher is better. NCC is one of the methods used for template matching; a process is used for finding incidences of a pattern or object within an image.

The above mentioned metrics evaluate the amount of information transferred from source images to fused image. From the information point of view mutual information statistically depend on two random variables. The metric  $Q^{S_1 S_2 f}$  evaluates the edge information. With the feature similarity metric phase congruent and edge information between source images and fused image can be obtained.

Other metrics considered to assess the fused images are Mutual Information (MI) [13], Objective image fusion performance [14], and Feature Similarity Index Measure (FSIM) for image quality assessment [12]. Objective quality assessment depends on the Edge Strength and Orientation Preservation (ESOP). Experiments are performed with different images and different metrics. Proposed method gives better performance than DWT in terms of all metrics. Experimental results are presented in Table II. Experimental results clearly indicate that the proposed method gives better results in terms of subjective and objective assessment. A bar chart related to the PSNR values is shown in Figure 2.

Experimental results clearly indicate that the proposed method gives better results in terms of subjective and objective assessment. MI measure is objective, concise, meaningful, and explicit. In the proposed method, considerable improvement in MI values is observed. FSIM concentrate on gradient measure. ESOP concentrates on edges in the fused image. From Table III, it is observed that the proposed method HT+HVS is surpasses the other statistical methods such as HT+Average, HT+Variance [17]. PSNR values for the existing method and the proposed method are shown in Figure 2. More than 7 dB improvement is observed in PSNR values using the proposed method.

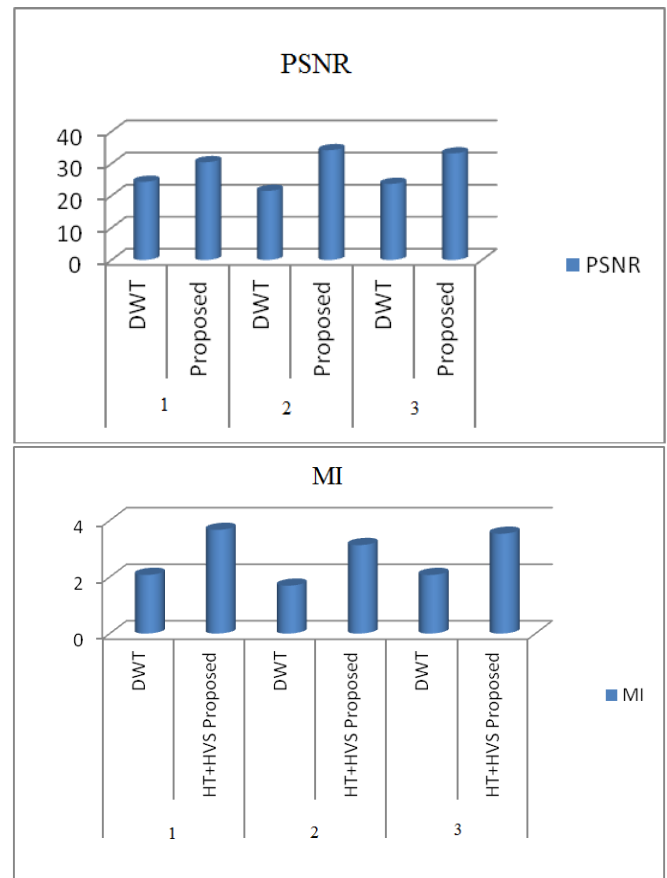


Fig. 2. Graphical representation of PSNR and MI

## V. CONCLUSIONS

In this paper, HT with Human Visual System for image fusion is presented. The proposed algorithm is efficient than DWT fused method. Experimental results indicate that, HVS based fusion gives better quality over traditional fusion rules. Further, more than 7dB improvement is observed in PSNR values with the proposed method. HVS based methods can be extended for image fusion using other transforms also. HT offers a significant advantage in terms of a shorter processing time as the processing involves simpler integer manipulation (compared to floating point processing with DCT and DWT) and the ease of hardware implementation

TABLE II. EXPERIMENTAL RESULTS

Image No	Fusion rule	PSNR	MI	ESOP	NCC	FSIM
1	DWT	24.3543	2.0754	0.4643	0.9226	0.8846
	HT+HVS Proposed	<b>30.4657</b>	<b>3.6681</b>	<b>0.9148</b>	<b>0.9829</b>	<b>0.9693</b>
2	DWT	21.5123	1.6946	0.4467	0.9056	0.8467
	HT+HVS Proposed	<b>34.2161</b>	<b>3.1300</b>	<b>0.9753</b>	<b>0.9970</b>	<b>0.9896</b>
3	DWT	23.7146	2.0732	0.4766	0.9366	0.8967
	HT+HVS Proposed	<b>33.2027</b>	<b>3.5307</b>	<b>0.9753</b>	<b>0.9920</b>	<b>0.9774</b>

TABLE III. EXPERIMENTAL RESULTS

Image No.	Metric	HT+Average	HT+Variance	HT+HVS [Proposed]
1	MI	3.0699	4.4407	<b>4.5343</b>
	ESOP	0.8063	0.9076	<b>0.9148</b>
	FSIM	0.9477	0.9997	<b>0.9997</b>
	NCC	0.9898	0.9989	<b>0.9992</b>
2	MI	2.9045	2.8182	<b>3.1300</b>
	ESOP	0.8623	0.9716	<b>0.9753</b>
	FSIM	0.9597	0.9651	<b>0.9896</b>
	NCC	0.9929	0.9937	<b>0.9970</b>
3	MI	2.8385	3.0513	<b>3.8318</b>
	ESOP	0.9664	0.9644	<b>0.9753</b>
	FSIM	0.9345	0.9376	<b>0.9872</b>
	NCC	0.9706	0.9708	<b>0.9943</b>

## REFERENCES

- [1] A. Garzelli, "Possibilities and limitations of the use of wavelets in image fusion", IEEE Geoscience and Remote Sensing Symposium. Vol. 1, pp. 66-68, 2002
- [2] G. Paella, "A general frame work for multiresolution image fusion: from pixels to regions", Information Fusion, Vol. 4, No. 4, pp. 259-280, 2003
- [3] D. A. Godse, D. S. Bormane, "Wavelet based image fusion using pixel based maximum selection rule", International Journal of Engineering Science and Technology, Vol. 3, No. 7, pp. 5572-5577, 2011
- [4] N. Mitianoudi, T. Stathaki, "Pixel-based and region-based image fusion schemes using ICA bases", Information Fusion, Vol. 8, No. 2, pp. 131-142, 2007
- [5] J. Jayanth, S. Koliwad, "Performance degraded by the sensor noise at pixel level image fusion", International Journal of Computer Applications, Vol. 8, No. 9, pp. 23-28, 2010
- [6] J. Tang, "A contrast based image fusion technique in the DCT domain", Digital Signal Processing, Vol. 14, No. 3, pp. 218-226, 2004
- [7] J. Johnson, M. Puschel, "In search of the optimal walsh-hadamard transform", International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Vol. 6, pp. 3347-3350, 2000, pp. 3347-3350, 2000.
- [8] J. Mannos, D. Sakrison, "The effect of a visual fidelity criterion in the encoding of images", IEEE. Trans. Information Theory, Vol. 20, No. 4, pp. 525-536, 1974
- [9] T. N. Pappas, J. P. Allebach, D. L. Nehhoff, "Model-based digital halftoning", IEEE Signal Processing Magazine, Vol. 20, No. 4, pp. 14-27, 2003
- [10] J. Sullivan, L. Ray, R. Miller, "Design of minimum visual modulation halftone patterns", IEEE Transactions on Systems, Man and Cybernetics, Vol. 21, No. 1, pp. 33-38, 1991
- [11] K. Veeraswamy, S. Srinivaskumar, B. N. Chatterji, "Designing quantization table for Hadamard transform based on human visual system for image compression", Graphics, Vision and Image Processing, Vol. 7, No. 3, pp. 31-38, 2007
- [12] L. Zhang, L. Zhang, X. Mou, D. Zhang, "FSIM: A Feature Similarity Index for Image Quality Assessment", IEEE transactions on Image Processing, Vol. 20, No. 8, pp. 2378-2386, 2011
- [13] G. Qu, D. Zhang, P. Yan, "Information measure for performance of image fusion", Electronic Letters, Vol. 38, No. 7, pp. 313-315, 2002
- [14] C. S. Xydeas, V. Petrovic, "Objective image fusion performance measure", Electronic Letters, Vol. 36, No. 4, pp. 308-309, 2000
- [15] S. Daly, "Subroutine for the generation of a two dimensional human visual contrast sensitivity function", Tech. Rep 2332037, Eastman Kodak, 1987
- [16] <http://www.metapix.de/toolbox.htm>
- [17] M. B. A. Haghighat, A. Aghagolzadeh, H. Seyedarabi, "Multi-focus image fusion for visual sensor networks in DCT domain", Computers & Electrical Engineering, Vol. 37, No. 5, pp. 789-797, 2011
- [18] R. Vadhi, V. Kilari, S. K. Samayamantula, "Uniform based approach for image fusion", in Eco-friendly Computing and Communication Systems, Springer Berlin Heidelberg, pp 186-194, 2012
- [19] D. Srinivas Rao, M. Seetha, M. H. M. Krishna Prasad, "Quality assessment parameters for iterative image fusion using fuzzy and neuro fuzzy logic and applications", Procedia Technology, Vol. 19, pp. 888-894, 2015
- [20] H. Lin, Y. Tian, R. Pu, L. Liang, "Remotely Sensing Image Fusion Based on Wavelet Transform and Human Visual System", International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol. 8, No. 7, pp. 291-298, 2015
- [21] Y. Yang, W. Zheng, S. Huang, "Effective Multifocus Image Fusion Based on HVS and BP Neural Network", The Scientific World Journal, Vol. 2014, Article ID 281073, pp.1-8, 2014