

# Multi-Focus Image Fusion Based on Spatial Frequency under DWT Analysis and SWT Analysis

P.Duraipandiyani<sup>1</sup>, Dr.S.Ganesan<sup>2</sup>

<sup>1</sup>PG Student, Mailam Engineering College

<sup>2</sup>Head of the Department, Mailam Engineering College

**Abstract-**The project presents multi focus image fusion using stationary wavelet transform with local directional pattern and spatial frequency analysis. In this project, the proposed model utilizes the multi scale decomposition done by stationary wavelet transform for fusing the images in its frequency domain. It decomposes an image into two different components like structural and textural information. It is used to reduce the problems like blocking, ringing artifacts occurs because of DCT and DWT. The low frequency sub band coefficients are fused by selecting coefficient having maximum spatial frequency. The high frequency sub band coefficients are fused by selecting coefficients having maximum LDP code value. The finest details of two images are characterized by local directional pattern descriptors before fusion and it describes local primitives including different types of curves, corners and junctions. Finally, fused two different frequency sub bands are inverse transformed to reconstruct fused image.

**Index Terms-** Discrete Cosine Transform, Discrete Wavelet transform, Local directional Pattern, Principle Component Analysis, Stationary Wavelet Transform.

## I. INTRODUCTION

The result of the limited depth of focus in optical lenses, it is difficult to describe the complex situation with a single image accurately. In wireless visual sensor networks, multiple sensors are applied to obtain images of the same scene, and a centralized fusion centre combines source images from multiple sensors into a single image, which is more suitable for human visual and machine perception. Then, the fused image will be transmitted to an upper node. So far, a lot of researches have concentrated on image fusion performed in the spatial domain. Methods based on multi-scale transform such as discrete wavelet transform (DWT), shift invariant discrete wavelet transform (SIDWT), and non-sub sampled contour let transform (NSCT) are popular. The fused image is then acquired by performing the inverse wavelet transform. The fusion rule plays a very essential rule in fusion process.

In WVSN, images are compressed before transmission to the other nodes. When the source images are saved or transmitted in DCT based standards, the methods applied in DCT domain will reduce computation complexity considerably. Recently, several image fusion techniques in DCT domain have been proposed. Tang *et al.* proposed two methods in DCT domain, namely, DCT+Average and DCT+Contrast. But these methods suffer some undesirable side effects like blurring or blocking artifacts which degrade the image quality. The algorithm proposed in called DCT+AC-Max leads to mistakes in selecting right JPEG coded blocks because the number of higher valued AC coefficients is an invalid criterion when the most of the AC coefficients are quantized to zeros during the quantization. In another approach, variance is considered as a contrast criterion of fusion. However, experiment results in show that variance provides worse performance than other focus measures. The basic process of image fusion based on wavelet transform for the most part includes three steps: decomposition, fusion and reconstruction. Wavelet transform decomposes the input image into pyramidal coefficients of low-pass and high-pass sub-bands and coefficient (approximation and detail) can be calculated. The fused wavelet coefficient map can be build from the wavelet.

## II. DWT BLOCKS ANALYSIS

Wavelets are finite duration oscillatory functions with zero average value. They have finite energy. They are suited for analysis of transient signal the irregularity and good localization properties make them better basis for analysis of signals with discontinuities. Using vector processing, the output matrix of a two dimensional 8x8 DCT for an input matrix is given by:

$$F = C.f.C^t \quad (1)$$

Where  $C$  is an orthogonal matrix consisting of the cosine Coefficients and  $C^t$  are the transpose coefficients.

$$C^{-1} = C^t \quad (2)$$

The inverse DCT (IDCT) is also defined as:

$$f = C^t.F.C \quad (3)$$

According to

$$\text{trace}(ff^t) = \text{trace}(FF^t) \quad (4)$$

where trace  $x$  stands for the trace of  $x$ .

The row frequency (RF) and column frequency (CF) of an  $8 \times 8$  image block are given by:

$$RF^2 = \frac{1}{8 \times 8} \sum_{x=0}^7 \sum_{y=1}^7 (f(x, y) - f(x, y-1))^2 \quad (5)$$

$$CF^2 = \frac{1}{8 \times 8} \sum_{x=1}^7 \sum_{y=0}^7 (f(x, y) - f(x-1, y))^2 \quad (6)$$

The total spatial frequency (SF) of an  $8 \times 8$  block in the spatial domain is calculated as:

$$SF^2 = RF^2 + CF^2 \quad (7)$$

After a small amount of calculation, we can calculate the spatial frequency of the block from the AC coefficients in the DCT domain. We denote  $\Delta x$  and  $\Delta y$  as the difference matrixes of rows and columns respectively:

$$\Delta x = \begin{pmatrix} f(0,1) - f(0,0) & \dots & f(0,7) - f(0,6) & 0 \\ \vdots & \ddots & \vdots & \vdots \\ f(7,1) - f(7,0) & \dots & f(7,7) - f(7,6) & 0 \\ f(1,0) - f(0,0) & \dots & f(1,7) - f(0,7) \\ \vdots & \ddots & \vdots & \vdots \\ f(7,0) - f(6,0) & \dots & f(7,7) - f(6,7) \\ 0 & \dots & 0 & 0 \end{pmatrix}$$

$$\Delta y = \begin{pmatrix} f(0,1) - f(0,0) & \dots & f(0,7) - f(0,6) & 0 \\ \vdots & \ddots & \vdots & \vdots \\ f(7,1) - f(7,0) & \dots & f(7,7) - f(7,6) & 0 \\ f(1,0) - f(0,0) & \dots & f(1,7) - f(0,7) \\ \vdots & \ddots & \vdots & \vdots \\ f(7,0) - f(6,0) & \dots & f(7,7) - f(6,7) \\ 0 & \dots & 0 & 0 \end{pmatrix}$$

The wavelet transform decomposes the image into low-high, high-low, high-high spatial frequency bands at different scales and the low-low band at the coarsest scale. The L-L band contains the average image information whereas the other bands contain directional information due to spatial orientation. Higher absolute values of wavelet coefficients in the high bands correspond to salient features such as edges or lines.

$$\Delta x = fb = C^t FCC^t BC = C^t FBC, \quad (8)$$

$$\Delta y = b^t f = (C^t EC)^t C^t FC = C^t B^t FC, \quad (9)$$

$$\text{where } b = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix},$$

and  $B$  is the DCT presentation of  $b$ .

From (3), (8), and (9), we can find and are the DCT presentations of and, respectively. Then, we get

$$RF^2 = \frac{1}{8 \times 8} \sum_{x=0}^7 \sum_{y=0}^7 \Delta x^2(x, y) = \frac{1}{8 \times 8} \text{trace}(\Delta x (\Delta x)^t)$$

$$= \frac{1}{8 \times 8} \text{trace}(FB(FB)^t) = \frac{1}{8 \times 8} \text{trace}(FBB^t F^t) \quad (10)$$

$$CF^2 = \frac{1}{8 \times 8} \sum_{x=0}^7 \sum_{y=0}^7 \Delta y^2(x, y) = \frac{1}{8 \times 8} \text{trace}((\Delta y)^t \Delta y)$$

$$= \frac{1}{8 \times 8} \text{trace}((B^t F)^t B^t F) = \frac{1}{8 \times 8} \text{trace}(F^t B B^t F) \quad (11)$$

Let  $D$  be the product of  $B$  and  $B^t$ . We can find  $D$  is a diagonal matrix shown at the bottom of the page. Then, we will have:

$$SF^2 = RF^2 + CF^2 = \frac{1}{8 \times 8} [\text{trace}(DF^t F) + \text{trace}(DF F^t)]$$

$$= \frac{1}{8 \times 8} \sum_{u=0}^7 \sum_{v=0}^7 [(D(u, u) + D(v, v)) \times F^2(u, v)]$$

$$= \frac{1}{8 \times 8} \sum_{u=0}^7 \sum_{v=0}^7 [E(u, v) \times F^2(u, v)] \quad (12)$$

where (see the second matrix at the top of the next page). In conclusion, the spatial frequency of a block of pixels can be accurately calculated by the weighted sum of squares of AC coefficients in the DCT block.

### III. PROPOSED METHOD

The fused wavelet coefficient map can be build from the wavelet coefficients of the source images according to the fusion decision map. The fused image is then acquired by performing the inverse wavelet transform. The fusion rule plays a very essential rule in fusion process. Hence, we can use the spatial frequency value as the contrast measure of the blocks of the source images. We only consider two source images  $A$  and  $B$ , but the method can be extended for more than two source images. The fusion process consists of the following steps:

1) Decode and de-quantize the source images, and then divide them into blocks of size  $8 \times 8$ . Denote the block pair at location  $(i, j)$  by  $A_{ij}$  and  $B_{ij}$  respectively.

2) Compute the spatial frequency of each block by, and denote the results of and by and, respectively.

3) Compare the spatial frequencies of two corresponding blocks to decide which should be used to construct the fused image. Create a decision map to record the feature comparison results according to a selection rule:

$$W_{i,j} = \begin{cases} 1 & SFA_{i,j} > SFB_{i,j} + T \\ -1 & SFA_{i,j} < SFB_{i,j} - T \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

Here, T is a user-defined threshold.

4) Apply a consistency verification process to improve quality of the output image. Use a majority filter to obtain a refined decision map R.

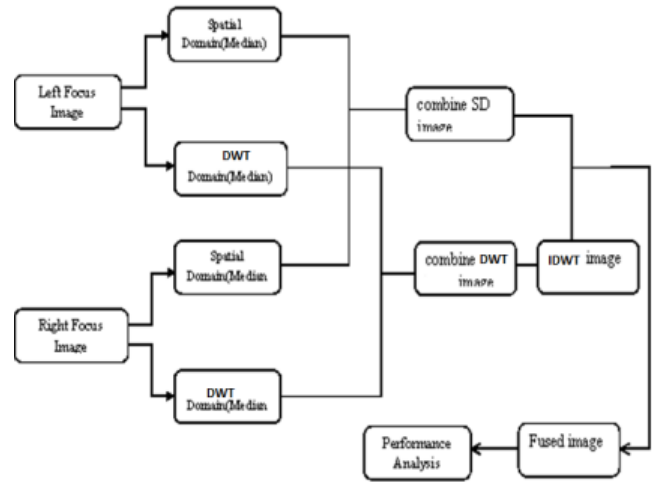
$$R_{i,j} = \sum_{x=i-1}^{i+1} \sum_{y=j-1}^{j+1} w_{x,y} \quad (14)$$

Then, obtain the DCT representation of the fused image based on as:

$$F_{i,j} = \begin{cases} A_{i,j} & R_{i,j} > 0 \\ B_{i,j} & R_{i,j} < 0 \\ (A_{i,j} + B_{i,j})/2 & R_{i,j} = 0 \end{cases} \quad (15)$$

- 4) Quantize the resulting DCT coefficients with a standard quantization table in the standard JPEG coder and then use entropy coding to produce the output bit stream.

In DWT disintegration, the filters are purposely designed so that the consecutive layers of the pyramid only consist of facts which were not previously obtainable at the earlier levels. Christo Ananth et al. [12] proposed a system which uses intermediate features of maximum overlap wavelet transform (IMOWT) as a pre-processing step. The coefficients derived from IMOWT are subjected to 2D histogram Grouping. This method is simple, fast and unsupervised. 2D histograms are used to obtain Grouping of color image. This Grouping output gives three segmentation maps which are fused together to get the final segmented output. This method produces good segmentation results when compared to the direct application of 2D Histogram Grouping. IMOWT is the efficient transform in which a set of wavelet features of the same size of various levels of resolutions and different local window sizes for different levels are used. IMOWT is efficient because of its time effectiveness, flexibility and translation invariance which are useful for good segmentation results. For the NSCT, we use 2, 4, 8 directions in the scales from coarser to finer. For the proposed method, we obtain the results with the threshold of 2. All the images in the simulations are converted to JPEG files.



The wavelet transform can be performed for multiple levels. The next level of decomposition is executed using only the LL image. The outcome is four sub-images each of size identical to half the LL image size.

#### IV. EXPERIMENTAL RESULTS AND ANALYSIS

The simulations of the fusion methods have been conducted with an Intel i5 4570 processor with 4 GB RAM. For the wavelet based methods, the DWT with DBSS (2, 2) and the SIDWT for

TABLE I  
OBJECTIVE EVALUATION OF IMAGE FUSION

Method	SSIM	RMSE	T(μs)
DCT + aver	0.9555	7.213	10.99
DCT + Contrast	0.9792	5.103	511.0
DCT + AC-Max	0.9227	9.326	24.70
DCT + Variance	0.9866	4.541	13.65
DCT + Variance + CV	0.9884	4.260	25.23
DWT	0.9833	4.721	-
SIDWT	0.9871	5.032	-
NSCT	0.9894	4.068	-
DCT + SF(proposed)	0.9888	4.220	13.96
DCT + SF + CV(proposed)	<b>0.9902</b>	<b>4.037</b>	25.54

#### V. STATIONARY WAVELET TRANSFORM

The basic idea is extremely simple. We simply apply appropriate high and low pass filters to the data at each level to produce two sequences at the next level. We do not decimate, and the two new sequences each have the same length as the original sequence. Instead, we modify the filters at each level, by padding them out with zeroes, in a way that we now done, in this section, we compare the performance of our technique with the existing image fusion methods in th

In the first experiment, the performance of the proposed fusion method is demonstrated by fusing 30 pairs of blurred images which are generated by filtering the standard grayscale images. The 2D Stationary Wavelet Transform (SWT) is based on the idea of no decimation. It applies the Discrete Wavelet Transform (DWT) and omits both down-sampling in the forward and up-sampling in the inverse transform. More precisely, it applies the transform at each point of the image and saves the detail coefficients and uses the low frequency information at each level. The Stationary Wavelet Transform decomposition schemes, where  $G_i$  and  $H_i$  are a source image, low pass filter and high-pass filter, respectively.

## VI. ALGORITHM

1. Decompose the two source images using SWT at one level resulting in three details sub bands and one approximation sub band (HL, LH, HH and LL bands).

2. Then take the average of approximate parts of images.

3. Take the absolute values of horizontal details of the image and subtract the second part of image from first

$$D = (\text{abs}(H1L2) - \text{abs}(H2L2)) \gg 0$$

4. For fused horizontal part make element wise multiplication of  $D$  and horizontal detail of first image. And then subtract another horizontal detail of second image multiplied by logical not of  $D$  from first.

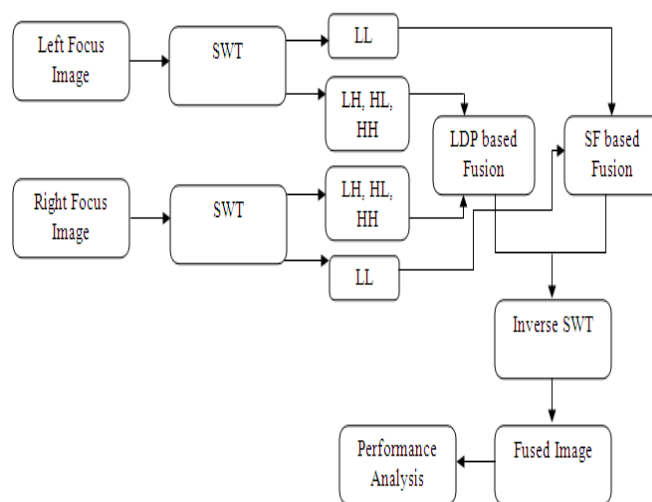
5. Find  $D$  for vertical and diagonal parts and obtain the fused vertical and details of image.

6. Same process is repeated for fusion at first level.

7. Fused image is obtained by taking inverse stationary wavelet transform.

The Discrete Wavelet Transform is not a time invariant transform. The way to restore the translation invariance is to average some slightly different DWT, called un-decimated DWT, to define the stationary wavelet transform (SWT). SWT is a wavelet transform algorithm designed to overcome the lack of translation-invariance of the discrete wavelet transform (DWT). Translation-invariance is achieved by removing the down samplers and up samplers in the DWT and up sampling the filter coefficients by a factor of in the 4th level of the algorithm. The un decimated algorithm is redundant, meaning some detail information may be retained in adjacent levels of transformation. It also requires more space to store the results of each level of transformation and, although it is shift-invariant, it does not resolve the problem of feature orientation.

The SWT is an inherently redundant scheme as the output of each level of SWT contains the same number of samples as the input – so for a decomposition of  $N$  levels there is a redundancy of  $N$  in the wavelet coefficients. It does so by suppressing the down-sampling step of the decimated algorithm and instead up-sampling the filters by inserting zeros between the filter coefficients. Algorithms in which the filter is up sampled are called “à torus”, meaning18



In the decimated case, this means up-sampling the approximation and detail images and applying reconstruction filters, which are inverses of the decomposition scaling and wavelet filters, first by columns and then by rows. The approximation images from the UN decimated algorithm are therefore represented as levels in a parallelepiped, with the spatial resolution becoming coarser at each higher level and the size remaining the same

## VII. LOCAL DIRECTIONAL PATTERN

The high frequency sub-bands contain the detail coefficients of an image, which usually have large absolute values correspond to sharp intensity changes and preserve salient information in the image. So, if the rule mentioned above for low frequency sub-bands is adopted here, the fused results will be blocked. According to the wavelet transform theory, we know that the energy of the high frequency coefficients of a clear image is much larger than that of a blurred one. Based on this analysis and considering that the wavelet coefficient is related to its neighboring region, we propose a fusion scheme by computing the neighboring energy maximum to select the high frequency coefficients.

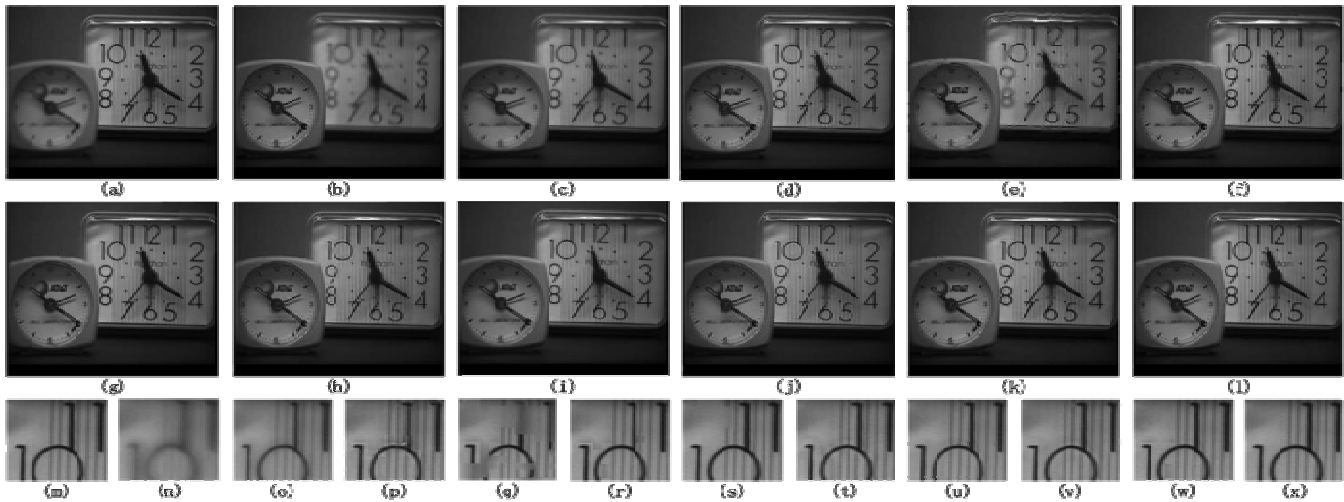


Fig. 1. Source images "Clock" and the fusion results. (a) The first source image with focus on the right. (b) The second source image with focus on the left. (c) Average result. (d) Contrast result. (e) Max result. (f) Variance result. (g) CV result. (h) DWT

Result. (i) SIDWT result. (j) NSCT result. (k) Result of the proposed algorithm. (h) Result of the proposed algorithm with consistency verification. (m)-(x) are the local magnified version of (a)-(l), respectively.

## VIII. CONCLUSION

The project presented a multi focus image fusion method that is based on discrete wavelet and spatial domain coefficients. Here, the discrete wavelet transformation was applied to wavelet sub bands to extract the detailed components present at curved edges to preserve the image quality by reducing loss of information due to fusion process. The pixel level averaging method was used effectively for fusion all spatial domain and wavelet sub bands. At the same time, the experiment results show that the novel method gives more encouraging and effective performance than other existing image fusion methods. In future, Image fusion can be carried out using stationary wavelet transform and the results can be compared with the present outputs.

## REFERENCES

- [1] B. K. Shreyamsha Kumar, M. N. S. Swamy, and M. Omair Ahmad, "Multiresolution DCT decomposition for multifocus image fusion," in *2013 26th Annu. IEEE Can. Conf. Electrical and Computer Engineering (CCECE)*, 2013, IEEE.
- [2] D. Egfin Nirmala, R. K. Vignesh, and V. Vaidehi, "Multimodal image fusion in visual sensor networks," in *2013 IEEE Int. Conf. Electronics, Computing and Communication Technologies (CONECCT)*, 2013, IEEE.
- [3] H. Li, B. S. Manjunath, and S. K. Mitra, "Multisensor image fusion using the wavelet transform," *Graph. Models Image Process.*, vol. 57.3, pp. 235–245, 1995.
- [4] O. Rockinger, "Image sequence fusion using a shift-invariant wavelet transform," in *1997 Proc. Int. Conf. Image Processing*, 1997, vol. 3, IEEE.

- [5] Q. Zhang and B.-L. Guo, "Multifocus image fusion using the nonsubsampling contourlet transform," *Signal Process.*, vol. 89.7, pp. 1334–1346, 2009.

- [6] M. B. A. Haghighat, A. Aghagolzadeh, and H. Seyedarabi, "Multi-focus image fusion for visual sensor networks in DCT domain," *Comput. Electr. Eng.*, vol. 37.5, pp. 789–797, 2011.

- [7] T. Jinshan, "A contrast based image fusion technique in the DCT domain," *Digital Signal Process.*, vol. 14.3, pp. 218–226, 2004.

- [8] Y. Phamila and R. Amutha, "Discrete cosine transform based fusion of multi-focus images for visual sensor networks," *Signal Process.*, pp. 161–170, 1995.

- [9] M. B. A. Haghighat, A. Aghagolzadeh, and H. Seyedarabi, "Real-time fusion of multi-focus images for visual sensor networks," in *6th IEEE Iranian Machine Vision and Image Processing (MVIP)*, 2010.

- [10] W. Huang and Z. Jing, "Evaluation of focus measures in multi-focus image fusion," *Patt. Recognit. Lett.*, vol. 28.4, pp. 493–500, 2007.

- [11] S. Djamel and B. Mouldi, "Image compression via embedded coder in the transform domain," *Asian J. Inf. Tech.*, vol. 5.6, pp. 633–639, 2006.

- [12] Christo Ananth, A. S. Senthilkani, S. Kamala Gomathy, J. Arockia Renilda, G. Blesslin Jebitha, Sankari @ Saranya. S., "Color Image Segmentation using IMOWT with 2D Histogram Grouping", *International Journal of Computer Science and Mobile Computing (IJCSMC)*, Vol. 3, Issue. 5, May 2014, pp-1 – 7

- [13] Y. Phamila and R. Amutha, "Discrete cosine transform based fusion of multi-focus images for visual sensor networks," *Signal Process.*, pp. 161–170, 1995.

- [14] M. B. A. Haghighat, A. Aghagolzadeh, and H. Seyedarabi, "Real-time fusion of multi-focus images for visual sensor networks," in *6th IEEE Iranian Machine Vision and Image Processing (MVIP)*, 2010.

- [15] W. Huang and Z. Jing, "Evaluation of focus measures in multi-focus image fusion," *Patt. Recognit. Lett.*, vol. 28.4, pp. 493–500, 2007.

- [16] S. Djamel and B. Mouldi, "Image compression via embedded coder in the transform domain," *Asian J. Inf. Tech.*, vol. 5.6, pp. 633–639, 2006.