DOI: 10.21917/ijivp.2010.0011

OPTIMAL LEVEL OF DECOMPOSITION OF STATIONARY WAVELET TRANSFORM FOR REGION LEVEL FUSION OF MULTI-FOCUSED IMAGES

K. Kannan[1] and S. Arumuga Perumal[2]
[1]*Kamaraj College of Engineering and Technology, Tamil Nadu, India*
E-mail: kannan_kcet@yahoo.co.in
[2]*S.T. Hindu College, Tamil Nadu, India*
E-mail: visvenk@yahoo.co.in

Abstract

In machine vision, due to the limited depth-of-focus of optical lenses in CCD devices, it is not possible to have a single image that contains all the information of objects in the image. To achieve this, image fusion is required which is usually refers to the process of combining two or more different images, each containing different features into a new single image retaining important features from each and every image with extended information content. The approaches to image fusion can be classified into two namely Spatial Fusion and Transform fusion. The most commonly used transform for image fusion at multi scale is Discrete Wavelet Transform since it minimizes structural distortions. But, wavelet transform suffers from lack of shift invariance and this disadvantage is overcome by Stationary Wavelet Transform. This paper describes the optimum level of decomposition of Stationary Wavelet Transform for region based fusion of multi focused images in terms of various performance measures.

Keywords:
Image Fusion, Region Level Fusion, Discrete Wavelet Transform and Stationary Wavelet Transform

1. INTRODUCTION

In machine vision, due to the limited depth-of-focus of optical lenses in CCD devices, it is not possible to have a single image that contains all the information of objects in the image. To achieve this, image fusion is required which is usually refers to the process of combining two or more different images, each containing different features, into a new single image retaining important features from each and every image with extended information content. For example, IR and visible images may be fused as an aid to pilots landing in poor weather or CT and MRI images may be fused as an aid to medical diagnosis or millimeter wave and visual images may be fused for concealed weapon detection or thermal and visual images may be fused for night vision applications. The fusion process should preserve all relevant information in the fused image, should suppress noise and should minimize any artifacts in the fused image. There are two approaches to image fusion, namely Spatial Fusion and Transform fusion. In spatial domain, the pixel values from sources images are taken and average is obtained to form the composite fused image [1].Transform fusion uses pyramid or wavelet transform for representing the source image at multi scale [2]. There are three levels in multi resolution fusion scheme namely Pixel level fusion, feature level fusion and region level fusion [3]. In this paper, it is proposed to find the optimum level of decomposition of Stationary Wavelet Transform (SWT) for region level fusion of multi focused images in terms of various performance measures like Root Mean Square Error (RMSE), Peak to Signal Noise Ratio (PSNR), Quality Index (QI) and Normalized Weighted Performance Metric (NWPM).

2. DISCRETE WAVELET TRANSFORM

Wavelet transforms provide a framework in which a signal is decomposed, with each level corresponding to a coarser resolution, or lower frequency band. There are two types of transforms, continuous and discrete. A continuous wavelet transform is performed by applying an inner product to the signal and the wavelet functions. For a particular dilation a and translation b, the wavelet coefficient $W_f(a,b)$ for a signal f can be calculated as [4],

$$W_f(a,b) = \langle f, \psi_{a,b} \rangle = \int f(x)\psi_{a,b}(x)dx \qquad (1)$$

The original signal can be reconstructed by applying the inverse transform:

$$f(x) = \frac{1}{c_w}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_f(a,b)\psi_{a,b}(x)db\frac{da}{a^2} \qquad (2)$$

where C_ψ is the normalization factor of the mother wavelet. Although the continuous wavelet transform is simple to describe mathematically, both the signal and the wavelet function must have closed forms, making it difficult or impractical to apply. So, the discrete wavelet is used. The term discrete wavelet transform (DWT) is a general term, encompassing several different methods. It is noted that the signal itself is continuous and discrete refers to discrete sets of dilation and translation factors and discrete sampling of the signal. For simplicity, it is assumed that the dilation and translation factors are chosen so as to have dyadic sampling. At a given scale J, a finite number of translations are used in applying multi resolution analysis to obtain a finite number of scaling and wavelet coefficients. The signal can be represented in terms of these coefficients as

$$f(x) = \sum_k C_{Jk}\phi_{Jk}(x) + \sum_{j=1}^{J}\sum_k d_{jk}\psi_{jk}(x) \qquad (3)$$

where $c_{Jk}$are the scaling coefficients and d_{jk} are the wavelet coefficients. The first term in Eq. (3) gives the low-resolution approximation of the signal while the second term gives the detailed information at resolutions from the original down to the current resolution J. The process of applying the DWT can be represented as a bank of filters, as in Fig.1. In case of a 2D image, a single level decomposition can be performed resulting in four different frequency bands namely LL, LH, HL and HH sub band and an N level decomposition can be performed resulting in 3N+1 different frequency bands and it is shown in Fig.1.

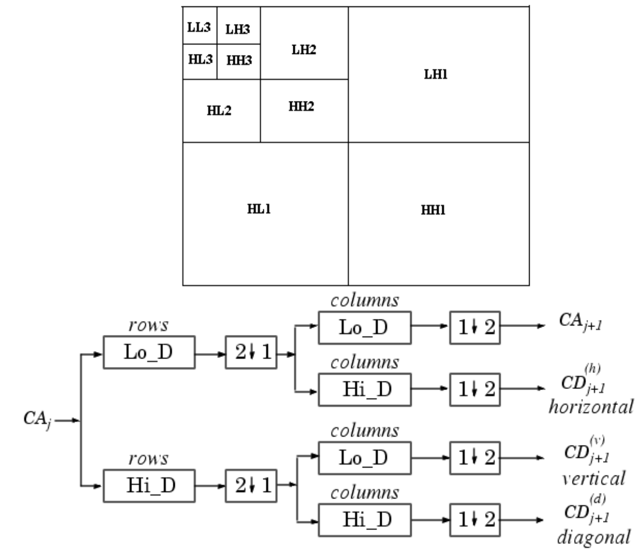

Fig.1. 2D – Discrete Wavelet Transform

At each level of decomposition, the image is split into high frequency and low frequency components; the low frequency components can be further decomposed until the desired resolution is reached. When multiple levels of decomposition are applied, the process is referred to as multi-resolution decomposition. The conventional DWT can be applied using either a decimated or an un-decimated algorithm. In the decimated algorithm, the signal is down sampled after each level of transformation. In the case of a two-dimensional image, down-sampling is performed by keeping one out of every two rows and columns, making the transformed image one quarter of the original size and half the original resolution. The decimated algorithm can be represented visually as a pyramid, where the spatial resolution becomes coarser as the image becomes smaller. The decimated algorithm is shift-variant, which means that it is sensitive to shifts of the input image. The decimation process also has a negative impact on the linear continuity of spatial features that do not have a horizontal or vertical orientation. These two factors tend to introduce artifacts when the algorithm is used in applications such as image fusion.

3. STATIONARY WAVELET TRANSFORM

The Discrete Wavelet Transform is a translation- variant transform. The way to restore the translation invariance is to use some slightly different DWT, called Stationary Wavelet Transform (SWT). 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 trous”, meaning “with holes”. In this case, however, although the four images produced (one approximation and three detail images) are at half the resolution of the original, they are the same size as the original image. 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. This can be visualized in the following Fig.2. 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. A previous level of approximation, resolution J−1, can be reconstructed exactly by applying the inverse transform to all four images at resolution J and combining the resulting images. Essentially, the inverse transform involves the same steps as the forward transform, but they are applied in the reverse order.

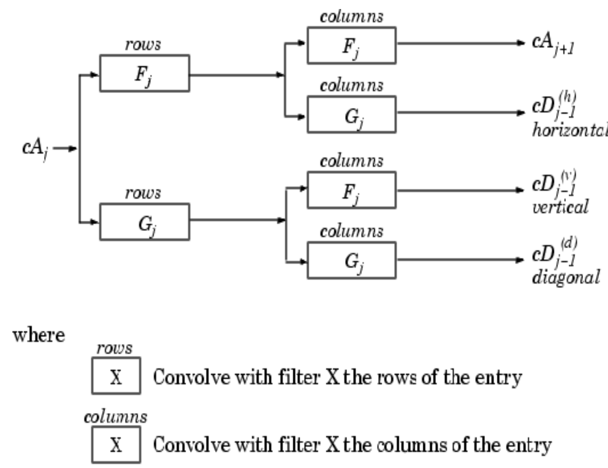

Fig.2. 2D – Stationary Wavelet Transform

4. STATIONARY WAVELET BASED IMAGE FUSION

Stationary wavelet transform is first performed on each source images, and then a fusion decision map is generated based on a set of fusion rules. The fused wavelet coefficient map can be constructed from the wavelet coefficients of the source images according to the fusion decision map. Finally the fused image is obtained by performing the inverse stationary wavelet transform [5]. Let A (x, y) and B (x, y) are images to be fused, the decomposed low frequency sub images of A (x, y) and B (x, y) be respectively lAJ (x, y) and lBJ (x, y) (J is the parameter of resolution) and the decomposed high frequency sub images of A (x,y) and B(x,y) are hAjk (x, y) and hBjk (x, y). (j is the parameter of resolution and j=1,2,3….J for every j, k=1,2,3..). Then, the fused high and low frequency sub-images Fjk (x, y) are given as Fjk (x, y) = Ajk (x, y) if G(Ajk (x, y)) >= G(Bjk (x, y)), else Fjk (x, y) = Bjk (x, y) and FJ (x, y) = lAJ (x, y) if G(AJ (x, y)) >= G(BJ (x, y)), else FJ (x, y) = lBJ (x, y) where G is the activity measure and Fjk (x, y) & FJ (x, y) are used to reconstruct the fused image F′(x, y) using the inverse stationary wavelet transform. The block diagram representing the stationary wavelet based image fusion is shown in Fig.3.

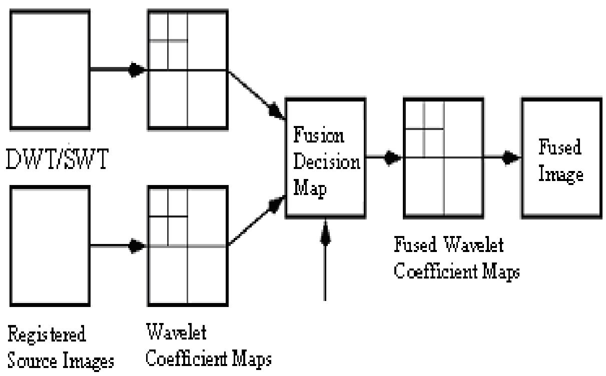

Fig.3. DWT/SWT Based Image Fusion

5. REGION BASED IMAGE FUSION

After creating the pyramid image using a wavelet transform, canny edge detector is applied to the lowest resolution approximation sub band of the image. After the edge detection, region segmentation is performed based on the edge information using region labeling algorithm. In the labeled image, zero corresponds to the edges and other different value represents different regions in the image. The activity level of region k in source image ‘n’, $Al_n(k)$ is given in the Eq.(4) as [6],

$$Al_n(k) = \frac{1}{N_K} \sum_{1 \le j \le N_k} Pj \qquad (4)$$

where N_k is the total number of pixels in region k, P_j is the activity intensity of pixel j in region k, which is the absolute value of pixel j in that region. Next step is to produce the decision map. The size of the decision map is the same as the size of the region image, which is the same size as the approximation band in the wavelet coefficient map. Each pixel in the decision map corresponds to a set of wavelet coefficients in each frequency band of all decomposition levels. Once the decision map is determined the mapping is determined for all the wavelet coefficients. Suppose, there are two registered images A and B to be fused then the decision map will be a binary image. For each pixel in this image, assume that value “1” means image A should be used instead of image B. Likewise the value “0” means image B should be used instead of image A. If a given pixel in the decision map is a “1” the all the wavelet coefficients corresponding to this pixel are taken from image A. If the pixel is “0” all the wavelet coefficients corresponding to this pixel are taken from image B. For a specific pixel of the decision map, P(i,j), this pixel may be:

- In region m of image A, and in region n of image B.
- an edge point in one image, and in certain region in the other image
- an edge point in both image.

The value of each pixel in decision map is assigned according to the following criteria

- Small regions preferred over large regions when comparing activity levels.
- Edge points preferred over non edges points when comparing activity levels.
- High activity-level preferred over low activity level.
- Make decision on non-edge points first and consider their neighbors when making the decision on edge points and avoid isolated points in decision map.

6. EVALUATION CRITERIA

There are four evaluation measures are used in this paper, as follows,

The Root Mean Square Error (RMSE) between the reference image R and fused image F is given as [7],

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}\sum_{j=1}^{N}[R(i,j) - F(i,j)]^2}{N^2}} \qquad (5)$$

The Peak Signal to Noise Ratio (PSNR) between the reference image R and fused image F is given by,

$$\text{PSNR} = 10\log_{10}(255)^2/(\text{RMSE})^2 \text{ (db)} \qquad (6)$$

Quality index of the reference image (R) and fused image (F) is given in the Eq.(7) as [8],

$$\text{QI} = \frac{4\sigma_{ab}\, ab}{(a^2 + b^2)(\sigma_a^2 + \sigma_b^2)} \qquad (7)$$

The maximum value Q=1 is achieved when two images are identical, where a & b are mean of images, σ_{ab} be covariance of R & F, σ_a^2 , σ_b^2 be the variance of image R,F. The Normalized Weighted Performance Metric (NWPM) which is given in the Eq.(8) as [9],

$$NWPM = \frac{\sum \forall_{i,j} Q_{ij}^{AF} W_{ij}^{A} + Q_{ij}^{BF} W_{ij}^{B}}{\sum \forall_{i,j} W_{ij}^{A} + W_{ij}^{B}} \qquad (8)$$

7. EXPERIMENTAL WORK

The method proposed for implementing region level image fusion using stationary wavelet transform takes the following form in general. The two source images to be fused are assumed to be registered spatially. The images are stationary wavelet transformed using the same wavelet, and transformed to the same number of levels. For taking the stationary wavelet transform of the two images, readily available MATLAB routines are taken. In each sub-band, individual pixels of the two images are compared based on the fusion rule that serves as a measure of activity at that particular scale and space. A fused wavelet transform is created by taking pixels from that wavelet transform that shows greater activity at the region level. The inverse stationary wavelet transform is the fused image with clear focus on the whole image.

8. RESULTS

For the above mentioned method, image fusion is performed using stationary wavelet transform and the performance is measured in terms of Root Mean Square Errors, Peak Signal to Noise Ratio, Quality Index & Normalized Weighted Performance Metric and the results are shown in figure 4 and tabulated in table1.

Table.1. Performance Comparison of SWT for various level of decomposition

Level	RMSE		PSNR	
	Lab	Pepsi	Lab	Pepsi
1	2.5717	2.8284	39.9264	39.1001
2	2.526	2.8252	40.0821	39.1097
3	2.4982	2.8099	40.1782	39.1571
4	2.5319	2.8741	40.0619	38.9609
5	2.5492	3.0095	40.0028	38.5608

Level	QI		NWPM	
	Lab	Pepsi	Lab	Pepsi
1	0.9987	0.9981	0.7137	0.7784
2	0.9988	0.9981	0.7128	0.778
3	0.9988	0.9981	0.7048	0.7779
4	0.9987	0.998	0.6999	0.7773
5	0.9987	0.9978	0.6988	0.776

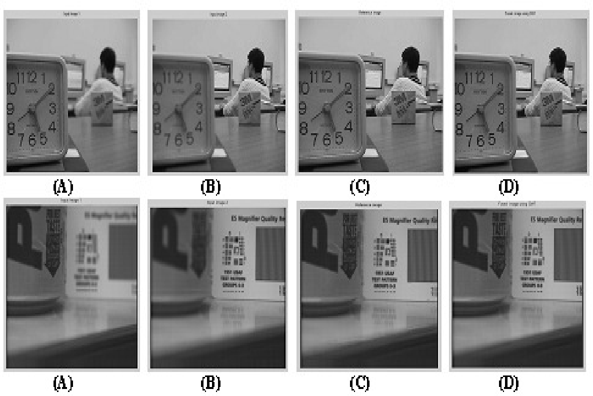

Fig.4. Region Based Image Fusion Using SWT: Row1: Lab Image, Row2: Pepsi Image, a. Input Image 1, b. Input Image 2, c. Reference Image and d. Fused Image using SWT

9. CONCLUSION

This paper presents the optimum level of decomposition of stationary wavelet transform for region based fusion of multi focused images in terms of various performance measures. The third level of decomposition of stationary wavelet transform for region based fusion of multi focused images provides computationally efficient and better qualitative and quantitative results. Hence using these fusion method at the third level of decomposition of stationary wavelet transform, one can enhance the image with high geometric resolution.

REFERENCES

[1] P. J. Burt and R. J. Kolczynski, 1993, "Enhanced image capture through image fusion", in proc.ICCV, pp. 173-182.

[2] P.J. Burt and E. Adelson, 1983. "The Laplcian Pyramid as a Image Codec", IEEE Transactions on Communications, Vol. 31, No.4, pp. 532-540.

[3] H. Li, B.S. Manjunath, and S.K. Mitra, 1995, "Multi sensor image fusion using the wavelet transform", in proc. ICGMIP, pp. 235–245.

[4] S. Mallat, 1998, "Wavelet Tour of Signal Processing", New York, Academic Press.

[5] Rick S. Blum and Yang Jinzhong, 2006, "Image Fusion Methods and Apparatus", US Patent, WO/2006/017233.

[6] Z. Zhang and R.S. Blum, 1997, "Region based image fusion scheme for concealed weapon detection", in Proc. CICC.

[7] Marta Mrak, Sonja Grgic and Mislav Grgic, 2003, "Picture Quality Measures in Image Compression Systems", in Proc. EUROCON'03, pp 233-237.

[8] Zhou Wang and Alan C. Bovik, 2002, "A Universal Image Quality Index", IEEE Signal Processing Letters, Vol. 9, No.3, pp. 81-84.

[9] C.S. Xydeas and V. Petrovic, 2000, "Objective Image Fusion Performance Measure", Electronics Letter, Vol.36, N0.4, pp. 308-309.

[10] K. Kannan and S. Arumuga Perumal, 2007, "Optimal decomposition Level of Discrete Wavelet Transform for pixel based fusion of multi-focused Images", in Proc. ICCIMA'07, Vol.3, pp.314-318.