Comparative Analysis of Wavelet Transform Based Image Fusion Methods

J. M. Kundargi
Assistant Professor

Vinay Shirhatti
B. E. (EXTC)

Department of Electronics and Telecommunicating,
K. J. Somaiya College of Engineering, Mumbai - 77

***Abstract*— Image fusion refers to the technique of integrating information from two or more images to obtain a single image retaining important features from each. The new image obtained after fusion is improved with respect to image quality, more suitable for human and machine perception and various image processing tasks like image analysis, segmentation, object recognition etc. Different fusion methods have been proposed in literature, including multi-resolution analysis. This paper is based on multi-resolution image fusion using wavelet transform. The key step in image fusion based on wavelet transform is that of wavelet coefficient combination in an appropriate way in order to obtain the best quality in the final image. Various coefficient combination methods have been proposed in the literature. This paper presents comparative analysis of a few selected methods , which are commonly used, for the fusion of multi-focus images and medical images.**

Index Terms— Image fusion; Wavelet transform; Multi-sensor fusion

I. INTRODUCTION

Image fusion has become an important topic in image analysis and computer vision [1,2]. With the availability of multi-sensor data in many fields, such as medical imaging, machine vision, it is possible to have several images of the same scene providing different information although the scene is the same. This is because each image has been captured with a different sensor. Image fusion refers to image processing techniques that produce a new, enhanced image by combining images from two or more sensors. The fused image is then more suitable for human/machine perception, and for further image-processing tasks such as segmentation, feature extraction and object recognition. The following two examples clarify these assertions; a detailed explanation of first example is given in Section III.

(1) While capturing any scene using a camera there may be some constraints which hamper the quality of the image and lead to a mismatch between the amount of details desired and those finally obtained in the image. One of such constraints which is inherently presented by the scene being captured is the presence of objects at different depths or distances from the capturing lens. In such a case, if one attempts to capture the details of any particular object by focusing the lens on it then the other details or objects at different distances are obscured and the information at varying depths is not reproduced on a consistent basis. Therefore to obtain details of each and every required object, images at different focusing levels need to be taken. Thus one has to work with multiple number of multi-focus images of the same scene to derive satisfactory details of each of those objects. If, however, one were equipped with a single image which would contain all the important information from such multi-focus images then the job of image analysis would be made much easier. To achieve all objects "in focus", a fusion process is required so that all focused objects are selected.

(2) In medical imaging, we can have CT scan and a magnetic resonance images from the brain of the same patient. The first one is a functional image displaying the brain activity, but without anatomical information. Where as the second provides anatomical information but without functional activity. Moreover, although the two images come from the same brain area, the CT scan has less spatial resolution than the second. The goal of fusion scheme for this examples to achieve a unique image with functional and anatomical information and with the best resolution. In general, the fused image has better quality than any of the original images. A prerequisite for a successful image fusion is that multi-focus images have to be correctly aligned on a pixel- by-pixel basis. Appropriate image registration technique can be used for this purpose. Further discussion assumes that images are perfectly aligned.

The simplest image fusion method is to take the average of two input images. However when this direct method is applied the contrast of the features uniquely presented in either of the images is reduced. In order to solve this problem, several methods based on the multi-scale (pyramid) transforms have been proposed. The basic idea is to perform a multi-resolution 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 multi-resolution transform. Several Laplacian pyramid based (or variants of it) fusion schemes have been developed [3, 4, 5]. Some major advantages of pyramid transform are:

i)It can provide information on the sharp contrast changes, and human visual system is especially sensitive to these sharp contrast changes.

ii)It can provide both spatial and frequency domain localization.

The Laplacian pyramid based image fusion techniques generate fused images with blocking artifacts in the regions where the multi-sensor data are significantly different.

In contrast, the wavelet transform based approach produces more naturally fused images.

By wavelet transform, an image can be represented by a low frequency approximation, which contains the average information of the image, and several high frequency details with different scales and directions, which contain the texture or edge feature of the image. Details about wavelet transform are given in [6, 7]. The use of wavelet transform for multi-sensor fusion is given in [8]. For the multi-sensor images, there are some areas unclear in certain source images which correspond to small wavelet coefficients, and clear in other source images which correspond to large coefficients. Fused image is produced by combining wavelet coefficients from multiple images. The key step in image fusion based on wavelets is that of wavelet coefficient combination in an appropriate way in order to obtain the best quality in the final image. Various coefficient combination methods have been proposed in the literature.

This paper is organized as follows. Section II describes different methods for combining the wavelet coefficients which are used in this paper. In section III, fusion scheme is illustrated with some real examples. A comparative analysis is carried out of the different coefficient combination methods. In section IV, the conclusions are presented.

II. COMBINING WAVELET TRANSFORM COEFFICIENTS

The key step in image fusion based on wavelets is that of wavelet coefficient combination in an appropriate way in order to obtain the best quality in the final image.

A) Basic algorithm

- Let the two original multi-focus images be 'a' and 'b', with focus on different sides.
- Take the wavelet transform of both 'a' and 'b'. The level of decomposition can be arbitrarily chosen. Let the transformed images be A and B.
- Select coefficients from A and B to construct an image C which would represent the transformed image of the final fused image. Use the coefficient values in A and B and an activity level measurement rule to decide whether the coefficient value should be taken from A or B or whether it should be some weighted combination of both.
- Apply inverse wavelet transform to the constructed image C.
- The image 'c' obtained as a result of the previous step represents the final fused image.

The steps 1 and 3 of the proposed algorithm are very straightforward. They serve the purpose of taking us to the transform domain, so that we can work there and then again back to the spatial domain once our work is over. It is, however, this work which decides the performance of the fusion operation. Thus the decision making logic and the subsequent selection criteria form the crux of this application and its success.

Some of the selection criteria which were studied and experimented with are covered ahead and their selection algorithm explained.

B) Pixel Averaging (PA)

In this approach the corresponding coefficients of both A and B are averaged and the resulting value is assigned to the corresponding coefficient of C.

Thus, $C(i,j) = A(i,j) + B(i,j)$

C) Maximum absolute transform coefficient (MA)

The absolute values of corresponding coefficients from A and B are compared and the larger of them is assigned to the corresponding pixel in C [9].

$$\text{Thus, } C(i,j) = A(i,j) \text{ if } \max\{\ \mathrm{abs}(A(i,j)),\ \mathrm{abs}(B(i,j))\ \} = \mathrm{abs}(A(i,j)$$
$$= B(i,j) \text{ if } \max\{\ \mathrm{abs}(A(i,j)),\ \mathrm{abs}(B(i,j))\ \} = \mathrm{abs}(B(i,j)$$

D) Maximum details retaining technique (MD)

The regions which are in focus show a better level of contrast and sharper well defined edges as compared to the blurred regions. Hence, such regions give rise to higher values of coefficients in the horizontal, vertical or diagonal detail images of the transformed image. Therefore to ensure selection of these sharper edges for the final image the maximum absolute coefficient values are chosen from the horizontal, vertical and diagonal detail coefficients and the values are accordingly assigned to the corresponding detail pixels of C [10]. The selection criterion for the approximation coefficients is heuristically decided from the characteristics of the original multi-focus images. If the prime areas are darker in color then the lower coefficient values from the approximation coefficient images of A and B are retained in C.

This can be represented as,

$$C\_LL(i,j) = \min\{\ A\_LL(i,j),\ B\_LL(i,j)\ \}$$
$$C\_LH(i,j) = \max\{\ A\_LH(i,j),\ B\_LH(i,j)\ \}$$
$$C\_HL(i,j) = \max\{\ A\_HL(i,j),\ B\_HL(i,j)\ \}$$
$$C\_HH(i,j) = \max\{\ A\_HH(i,j),\ B\_HH(i,j)\ \}$$

Where LL, LH, HL and HH refer to the corresponding bands in the wavelet transform of 'a' and 'b' i.e. A and B. Here even the maximum value rule can be used for LL band decision in case the important objects are brighter in the image.

E) Area-based (or window-based) maximum selection(AB)

Most of the objects in any image are normally much larger than a few pixels. Hence a decision based on individual pixel activity may not be always an appropriate one. Hence, an area-based approach is employed in this selection method [9]. For this a window is defined of some fixed size and an activity level parameter is decided. This window is then moved across the approximate image of both A and B and at every position i.e. for every area enclosed by the window the predefined activity level is calculated. We have used variance as this parameter for activity level measurement.

III. COMPARATIVE ANALYSIS

Although there have been as many attempts as there have been fusion algorithms, as yet no universally accepted standard has emerged for evaluating image fusion performance. In this work, both qualitative and quantitative methods are used. The qualitative methods are acceptance and verification tests which are accepted or rejected by a possible user, which determine visually the relative perceived image quality based on the contribution that the fusion makes to its specific problem. This is the case for medical image fusion, where the corresponding professional compares the results against other non-imaging data.

The problem with defining a quantitative measure lie in the difficulty of defining an ideal composite image based on multi-focus images or images taken at different times. To illustrate the fusion process, the fusion example of multi-focus visual images is presented.

A) Multi-focus image fusion

Due to the limited depth-of-focus of optical lenses it is often not possible to get an image that contains all relevant objects "in focus". One possibility to overcome this problem is to take several pictures with different focus points and combine them together into a single frame that finally contains the focused regions of all input images.

A test strategy is designed such that a correct focused image is captured and then two images with different sides blurred are generated by applying a low-pass filter to the desired regions. The target image is the original correct image. More than two images can be used but the performance will remain the same.

This paper compares different coefficients merging methods and different resolution levels. As a quality measure the parameter ‘root mean square error’ i.e. RMSE is determined for each case. RMSE is given by:

$$\text{RMSE} = \sqrt{\left\{\sum_{x}\sum_{y}\left(o(x,y) - c(x,y) \right)^2 / (x \,.\, y) \right\}} = \sqrt{\left\{\sum_{x}\sum_{y}\left(d(x,y) \right)^2 / (x \,.\, y) \right\}}$$

The sample input image taken consists of two objects. Figure 1-a is the image with focus on right, figure 1-b is the image with focus on left. DWT is applied to both a and b to obtain approximation and detail coefficients. Such coefficients are selected to obtain a fused multi-scale image. n is the number of levels of decomposition. The results for the the above four fusion selection criteria are presented below.

Fig.1-a Image-a,focus on right

Fig.1-b Image-b ,focus on left

B) Pixel Averaging method (PA)

Figure 1-c is the resultant fused image, using averaging method (PA), for level of decomposition n = 1 as discussed above. Similar fusion process was applied at other values of ‘n’.

Calculated values for RMSE for different levels of n are as shown.

Fig. 1-c , fused image using PA , n=1

Similar fusion process was applied to other image “cameraman.tiff” at different values of ‘n’. The results are presented below. Multi-focus images and fused image using PA method for n=3 for “cameraman.tiff” image are shown in appendix I.

Table 1. Shows **RMSE** using PA method

n	1	2	3	4
sample input image	2.4303	2.4303	2.4303	2.4303
cameraman.tiff	1.4548	1.3610	1.3641	1.3676

Comment on results

The RMSE values indicate that the level of decomposition does not have any effect on the fusion result. Although the fused image helps to reduce the blurring in multifocus of the input images, the technique is found wanting in terms of sharpness since the detail coefficients are also averaged and therefore some edge information is reduced in the final image.

C) Maximum absolute transform coefficient (MA)

Fig. 2 , Fused image for n=1

Fig. 3 , Fused image for n=2

Fig. 4 , Fused image for n=3

Fig. 5 , Fused image for n=4

Table 2. Shows **RMSE** using Maximum absolute transform coefficients.

n	1	2	3	4
sample input image	3.2336	1.7418	1.3052	1.0481
cameraman.tiff	3.2460	1.9981	1.5525	1.4949

Comment on results

Here it can be observed that as the value of 'n' increases the RMSE value reduces initially and then stays fairly constant. As the number of decomposition levels goes on increasing we work at a higher number of bands of frequency and also specifically at higher bands where the maximum absolute selection rule yields better results. However the initial RMSE is somewhat more than the averaging case. Also visually the final image does not seem to carry forward the details very faithfully from the original multi-focus images and some distortions are observed especially at lower values of 'n'.

D) Maximum details retaining technique (MD)

Fig. 6 , Fused image for n=4

Table 3. Shows **RMSE** using Maximum details retaining technique.

n	1	2	3	4
sample input image	2.5592	2.1313	1.0576	0.8377
cameraman.tiff	2.8443	1.6915	1.4542	1.3781

Comment on results

In this case the values of RMSE obtained are slightly lower than for the previous cases. Here again the RMSE value varies with n. The initial RMSE value is however slightly on the upper side and then it reduces with increasing n and stays around that value. Visually the results appear superior to the previous two cases. However this technique is heuristic in nature and its results may vary for different image sets. Depending upon the type of images and the information in them the algorithm to determine the approximation coefficients in C may have to be modified to obtain better and consistent results.

E) Area-based (or window-based) maximum selection (AB)

Fig. 7, Fused image for n=4, w=5*5

Table 4. Shows **RMSE** using Area-based (or window-based) maximum selection , w=5*5

n	1	2	3	4
sample input image	1.8697	1.0759	1.1112	1.0878
cameraman.tiff	1.7760	1.4189	1.3712	1.3545

Table 5. Shows **RMSE** using Area-based (or window-based) maximum selection , n=2

w	5 x 5	7 x 7	9 x 9	11 x 11	13 x 13
sample input image	1.0759	1.0689	1.1478	1.0705	1.0854

Comment on results

This method shows a better overall performance than the previous methods. The RMSE values are very consistent and on the lower side. Besides the window size can be properly chosen according to the image size to further reduce the RMSE value. The biggest positive of this method is the visual clarity and purity with respect to sharper, well defined edges and faithful reproduction of most of the major features of the original multi-focus images.

Table 6. Comparative performance table for the sample input image

n	1	2	3	4
PA	2.4303	2.4303	2.4303	2.4303
MA	3.2336	1.7418	1.3052	1.0481
MD	2.5592	2.1313	1.0576	0.8377
AB (w = 5 x 5)	1.8697	1.0759	1.1112	1.0878

IV. Conclusion

The wavelet transform and its potential use in image fusion applications were studied. The studied methods were implemented using MATLAB and tested for various test images. Suitable and well established statistical parameters were used to evaluate the performance of the proposed fusion methods. The results of the image fusion experiments accompanied by the relevant statistical results and inferences were then finally presented.

The wavelet transform helps to decompose an image into different detail bands without affecting the information content in the image. Hence one can separately work with different detail coefficients thus improving the scope and versatility of fusion techniques.

Apart from the basic platform provided by the wavelet transform various coefficient combining methods and algorithms were studied and tested. In the process of planning out these algorithms, their subsequent implementations and the final result one thing was repeatedly realized that fusion problems vary in their nature due to varying features that come with every problem. Hence, one has to heuristically plan out the basic approach and the algorithms, test them and modify them to improve the results repeatedly till a satisfactory outcome is achieved. The specific application of wavelet based image fusion was studied and implemented namely, the fusion of multi-focus images . It is realized that a number of factors like number of decomposition levels, coefficient selection and combining rules affect the final image in most of the cases. Root mean square error was used to mathematically or statistically judge the techniques used. Besides the final result images were judged by their visual quality which indicated the degree of success achieved in each case. Important conclusions were drawn from these observations which helped in improving the algorithms used.

The results obtained in the experiments performed were quite satisfactory and they indicated the success of the methods used. In spite of all these results it is very difficult to pass a verdict on the absolute superiority of any particular method due to the image specific factors which affect their performance. However a reasonable inference could be drawn in each case such as the consistency and the better results obtained from the area based method in case of fusion of multi-focus images. One important requirement is to try various coefficient combining techniques, which is in turn is related to more experience and a better insight into image related issues to extract the maximum out of this wavelet based approach.

The further applications of this method include its widespread used in fusion of multi-sensor data as in the case of satellite images taken at different wavelengths intended to capture different details. This method is used extensively for fusing panchromatic and multi-spectral images obtained from satellites to obtain a high resolution and spectrally rich image for better image perception and analysis. Wavelet transform based image fusion can be used in case of multi-temporal images to analyze time related characteristics. It can be used to fuse images at different resolution levels to achieve better results as compared to methods which aim at first resizing the images and then combining their information. Besides it can be used in other applications like hidden weapon-detection, navigation aids, remote sensing etc as pointed out in various papers and related literature.

Within every application a comparative study can be further carried out to determine better algorithms, fusion rules, wavelets etc to improve the final result. Algorithms and coding can be optimized for better time performance. Finally, this wavelet based approach is a field where new problem definitions would lead to some new concepts and methods to solve them and hence provides a good scope for research.

References

[1] J.K. Aggarwal, Multi-sensor Fusion for Computer Vision, Springer, Berlin, 1993.

[2] A. Akerman, Pyramid techniques for multi-sensor fusion, Proc. SPIE 2828(1992) 124-131.

[3] A. Toet, “Hierarchial image fusion”, Machine Vision and Applications, pp.1-11, March 1990.

[4] A.Toet, “Multi-scale contrast enhancement with applications to image fusion”, Optical Engineering, Vol.31,pp.1026-1030, May 1992.

[5] P.Burt and R. Lolczynski, ”Enhanced image capture through fusion”, in Proc. The Fourth International Conference on Computer Vision, pp.173-182,1993.

[6] I. Daubechies, Ten lectures on wavelets, SIAM,

CBMS-NSF lecture Notes, vol.61,1992.
[7] Mallat. S., “Wavelets for a vision”, Proc. IEEE, Vol.84,pp.604-614,1996.
[8] Li, H., Manjunath, B. S., and Mitra S.K., “Multi-sensor image fusion using the wavelet transform”, Graphical Models and Image Processing,Vol.57,No.3,1995.
[9] V. Petrovic, C. Xydeas, Multi-resolution image fusion using cross band selection, Proc. SPIE 39 (1999) 319–326
[10] Gonzalo P., Jesus Manual de la Cruz, “A wavelet based image fusion tutorial “,Pattern Recognition 37(2004)1855-1872

Appendix I

Right side blurred image

Left side blurred image

Fused image “cameraman.tiff”