

Submitted to : **Workshop on Multimedia Technology**

Paper Title : **FUSION OF DEFOCUSED IMAGES USING TRANSFORMS WITH
SEGMENTATION**

Authors : *Ren-Jean Liou*^{*}, *Wen-Hou Hsu*⁺, *Mu-Song Chen*⁺

Affiliations : ^{*}Department of Information Technology, National Ping-Tung Institute of
Commerce, Ping-Tung, 900

⁺Department of Electrical Engineering, Da-Yeh University, Changhua, 515

Contact Author: Ren-Jean Liou

Address: 51, Ming-Shen E. Rd. Ping-Tung, Taiwan 900

Phone: (08) 7238700-6204

Fax: (08) 7238720

e-mail: renjean@npic.edu.tw

ABSTRACT

This paper is concentrated on the improvement of data fusion techniques, which is applied on out-of-focused images. Segmentation is first employed to enhance the image features. An image transform is then applied to fuse the source images. In segmentation processes, quadtree and edge detection techniques are used. In quadtree, regions with image detail will be segmented into blocks with smaller size, and the background of the image will be assigned with larger block size. In edge detection, apparent objects in the image can be extracted. Image feature description is then made easy. After the preprocessing, those salient features in the image such as edges, lines and region boundaries will be selected as part of fused images. The rest of the regions without significant features are further processed using wavelet or DCT transforms. The transform coefficients from different sources are then appropriately combined. The final image is obtained by

taking the inverse transform of the fused coefficients. Simulation results will demonstrate that our approaches will have better quality and efficiency.

1. INTRODUCTION

The purpose of image fusion is to integrate complementary information from multi-sensory data so that the new images are more suitable for human visual perception and object recognition [1][2]. The results will facilitate human inspection, detection and identification. In the process of image fusion, registration is an important step. It assures that information with similar physical structure from different sources can be assigned appropriately [3]. Signals coming from different sources are expected to be combined efficiently so that complementary information can be compared and analyzed.

The most straightforward method of image fusion is sum and averaging. However, it produces images with lower contrast. In order to improve the problem, various approaches are proposed using image pyramid [4]-[8]. The earliest fusion model was proposed in [4] for stereo vision. It utilized the methods of Laplacian-pyramid transform and maximum selection. Gradient pyramid transform and small window measurement is another commonly used method [5]. Until recently, wavelet transform has gained its popularity. Wavelet coefficients are extracted as image features. Area-based maximum selection and consistency verification are utilized for decision making [6].

All the image fusion approaches transfer the input information into a different space. Important features are then selected and merged. Finally an inverse transform is performed to obtain the fused results. The processes are usually performed on the entire image including those evident regions. In many cases, these regions are so apparent that they can be directly extracted without any further consideration. Therefore, the processes in conventional fusion approaches have sub-

stantial redundancies. In this paper, we propose segmentation using quadtree or edge detection as preprocessing schemes in order to find regions with evident features. These apparent regions can bypass the lengthy fusion processes and be used as the results directly. The rest of images are then sent for fusion processing using wavelet or DCT transforms. We propose two new approaches for image fusion in order to compare the performance.. One uses quadtree for segmentation and wavelet transform for image transform and fusion. The other one uses edge detection for segmentation and DCT for image transform and fusion. We will demonstrate that better and more efficient results can be obtained.

This paper is organized as follows. Section 2 describes the segmentation preprocesses using quadtree and edge detection. Section 3 introduces the image fusion scheme using DCT and wavelet transforms. Section 4 defines the performance measure and presents the simulation results. Finally Section 5 is conclusion.

2. SEGMENTATION USING QUADTREE AND EDGE ENHANCEMENT

Many real-life images consist of smooth background and details of high frequency. Since small areas of images tend to have higher correlation, they are easier for detail observation. Certain features are extracted for segmentation into smaller components. After the segmentation, the details of images become evident. In our approach, if only one image source possess such obvious feature, it can be used directly as the final results without any further fusion process. Fusion is only performed on those obscure areas. In such way, the efficiency and accuracy of the system can be greatly improved. Quadtree and edge detection are two of the most popular segmentation schemes.

In Quadtree [9]-[12], images are repeatedly divided into four quadrants based upon the similarity of image content in each block. Image areas with smooth contents are segmented into

bigger blocks. Detailed image areas, which possess apparent variation in content, are divided into smaller blocks.

The idea of Quadtree comes from data structure. A root creates four nodes. Each node can also create up to four sub-nodes. The last partition, which is called leaf, contains only image of similar features. Fig. 1(a) shows an example of Quadtree partition of four generations. Its tree structure is shown in Fig. 1(b). For simplicity and flexibility, this paper uses top-down segmentation style [9]. A 512×512 image is first divided into four 256×256 blocks. Each 256×256 block is then checked for possibility of further partition. The process continues for all new-generated blocks until each block possesses little variation or the smallest preset 32×32 blocks is reached.

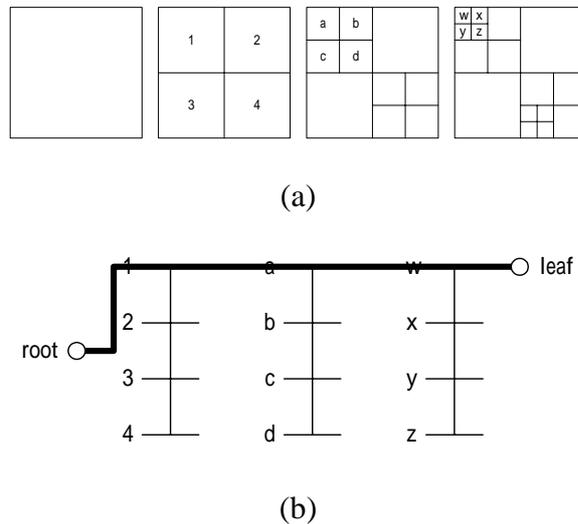


Fig. 1 The division of image using Quadtree and its tree structure.

There are several rules in determining whether further partition is needed. We used the simplest one, which compares the difference between the maximum and the minimum intensities of a block. When the difference is greater than a threshold, it represents that the image is highly variant. The current block is then further partitioned. Otherwise the partition for the current block is stopped.

Edges characterize object boundaries. Hence are useful for segmentation, registration and identification of objects in scenes. There are several approaches for edge detection [13]. Hence the edge information is readily obtained. Edge can be thought of as image contents with abrupt

gray-level change. For a continuous image $f(x, y)$ its derivative assume a local maximum in the direction of the edge. Hence a simple detection technique is to measure the gradient of f in a direction θ , that is

$$D = f_x \cos \theta + f_y \sin \theta \quad (1)$$

The maximum value of D is obtained when θ is the direction of the edge. For digital images, gradient operators, or masks, were introduced. We chose the simplest Roberts method [14]. A pair of masks R_1 and R_2 was defined as

$$R_1 = \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix} \quad R_2 = \begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix} \quad (2)$$

The two masks measure the gradient of the input image in two orthogonal directions. Convolutions are performed between the input images and the two masks. This yields bidirectional gradients, $g_1(x, y)$ and $g_2(x, y)$. Thus we can obtain the magnitude and direction of the two gradient vectors, which is given by

$$g(x, y) = \sqrt{g_1^2(x, y) + g_2^2(x, y)} \quad (3)$$

$$\theta(x, y) = \tan^{-1} \frac{g_2(x, y)}{g_1(x, y)} \quad (4)$$

The values of $g(x, y)$ indicate the strength of edge at the pixel location (x, y) . For hard decision making, a pixel can be declared an edge when $g(x, y)$ is greater than a threshold.

3. THE IMAGE FUSION SCHEME

3.1 The basic algorithm

There are several approaches for image fusion. Discrete wavelet transform (DWT) has become popular due to its nature of multiresolution approach [13]. Owing to its compactness, orthogonal-

ity and the availability of directional information, the wavelet transform can effectively extract salient features at different scales. As a result, wavelet fusion scheme usually produces better results than the traditional Laplacian pyramid based methods [4].

Figure 2 shows the schematic diagram of image fusion using DWT scheme. DWT is first applied to the input images. The transformed images are then separated into the low-high band, the high-low band, the high-high band and the coarser low-low band at different scales. Since larger absolute transform coefficients correspond to sharper brightness changes, i.e., the “salient features,” a good selection rule is to choose the coefficients with higher absolute values in the transform domain as the fusion results. Fusion is performed in such way at all resolution levels. Hence more dominant features at each scale are preserved in the new multi-resolution representation. A new image is then constructed by performing an inverse DWT.

One important criterion for evaluating an image fusion scheme is the stability of inverse transform. The reconstruction of the Laplacian pyramid can be unstable especially in regions where the two images appear significantly different. As a result, artifacts such as blocking effects are often visible. Contrarily, using DWT will not create such problem.

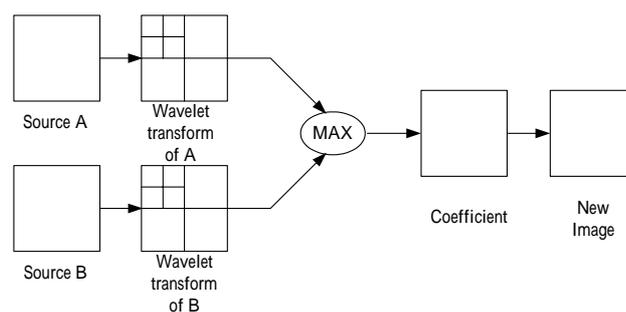


Fig. 2 The block diagram of the image fusion scheme using DWT.

Since most useful image features are larger than one pixel, the pixel-by-pixel maximum selection rule may not be an appropriate method. In [6], an area-based method is proposed which is shown in Fig. 3. The pixel values are compared to that of the center pixel in order to measure the

also increase deviation between the original and reconstructed images. Generally, the process of transformation is the most time-consuming step. If only fractions of image are required to go through the transformations, the system can be more efficient and accurate. Therefore, we propose the use of segmentation as preprocessing in order to improve system performance.

The motivation of image segmentation is to divide image into small blocks so that the statistics of each block will be nearly stationary. In other words, the gray scales in each block will have smaller variation. Thus more compact and related features can be obtained. The results of segmentation can help to determine whether some areas are obvious enough. The corresponding blocks from different sources are compared after the preprocessing. If both of them have similar properties, then they are sent to further processing. Otherwise, source with evident features is used as result directly and the transform processes can be skipped.

Suppose there are two images to be fused, the entire processes proposed are listed in the following steps:

1. Both images are segmented according to the Quadtree or edge detection criterion. Our experiments show that the smallest block with size 32×32 are optimal in the sense of quality and performance.
2. After the Quadtree process is completed, each block is marked by 1 if the smallest size is reached. Blocks with larger size are divided into smaller size of 32×32 and marked by 0. For example, a block with size 64×64 is divided into four 32×32 blocks. And all of them are marked by 0.
3. In edge detection, the blocks with declared edge are marked by 1. Other blocks are marked by 0.
4. The blocks of same position from the two images are compared. If they are both marked by 0 or 1, then these blocks are transformed using DWT or DCT and fused according to the rules in Section 3.1

5. If the blocks of two images are marked differently then the block with mark 1 is used directly as the fusion result.

Steps 4 and 5 are the most important part where the computation can be reduced without affecting the accuracy and performance.

4. SIMULATION RESULTS

There is no objective criterion for effective evaluation of performance under various situations. This is because that there is normally no standard image available for comparison. Nevertheless, somehow we have to use some criterion to evaluate our results. The quality measurement of root mean square error (RMSE) is adopted here. In RMSE, a test image is first chosen as the ideal results. Two corrupted images are created based on this test image. These corrupted images are fused and then compared with the test image to determine the differences. Therefore, RMSE can be defined as

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{x,y} [H(x,y) - F(x,y)]^2} \quad (5)$$

where $H(x, y)$ is the pixel intensity of the test image and $F(x, y)$ is the pixel intensity of the fused image. (x, y) is the pixel coordinates and n is the total number of pixels.

We have applied our approaches to numerous images and obtained superior results. For demonstration, Figs. 4 and 5 show the original images of Airfield and Bridge, respectively. They are both of size 512×512. The left and right parts of Airfield were defocused and are shown in Figs. 6 and 7, respectively. To demonstrate the applicability under various situations, the upper and lower parts of Bridge were defocused and are shown in Figs. 8 and 9, respectively. The fusion result of Airfield and Bridge using quadtree and DWT is shown in Figs. 10 and 12, respectively. The fusion result of Airfield and Bridge using edge detection and DCT is shown in Figs.

11 and 13, respectively. We can see that the final images all look clear and focused.



Fig. 4 Original Airfield.



Fig. 5 Original Airfield.

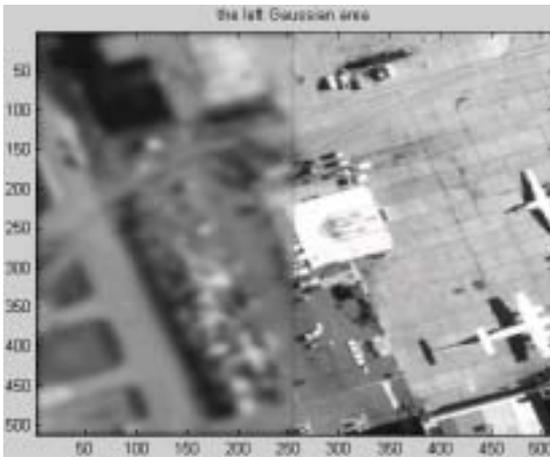


Fig. 6 Airfield with left parts defocused.

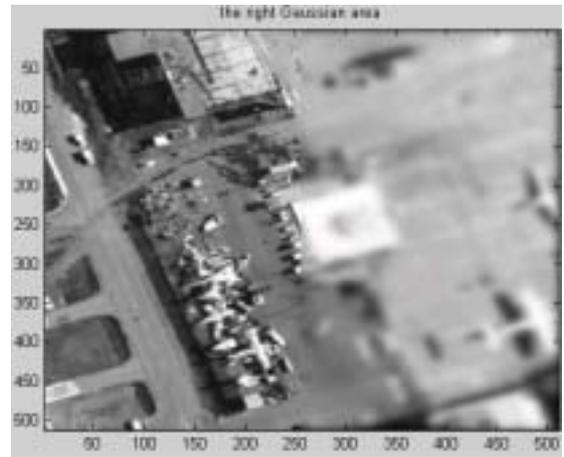


Fig. 7 Airfield with right parts defocused.



Fig. 8 Bridge with upper parts defocused.



Fig. 9 Bridge with lower parts defocused.

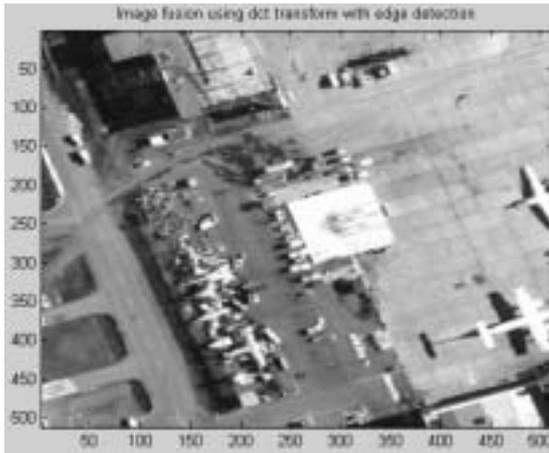


Fig. 10 Fusion results of Airfield using quadtree and DWT.

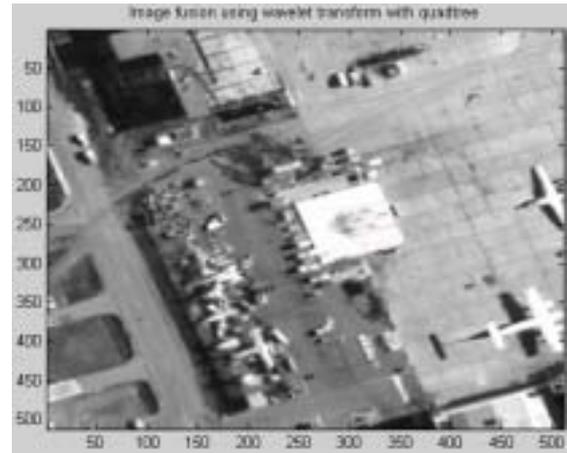


Fig. 11 Fusion results of Airfield using edge detection and DCT.



Fig. 12 Fusion results of Bridge using quadtree and DWT.



Fig. 13 Fusion results of Bridge using edge detection and DCT.

We also compared our method to other conventional approaches: the DWT with region-based methods. Comparisons are made in terms of reconstruction quality (RMSE) and speed (CPU time in seconds). Tables 1 and 2 show the results of applying the three methods on Airfield and Bridge, respectively. The DWT orders were varied from 1 to 3. We can see that the quality and computational speed gains of our approach start immediately from DWT of order 1. The RMSE is far smaller than the conventional method. The benefit in computational speed becomes apparent as the order increases. The computation of DCT has only 1 order. The computation time is at least two times faster than the other approaches. The quality is compatible to the 2nd order DWT with

quadtree. Therefore, our methods perform superior in both quality and speed than the conventional approach.

Table 1 Comparisons of RMSE and CPU time for Airfield.

DWT Orders		DCT with edge detection	DWT with Quadtree	Region-Based DWT
1	RMSE	15.8309	15.9348	21.3265
	TIME	12.208	27.81	54.198
2	RMSE		15.7104	19.1352
	TIME		41.039	77.391
3	RMSE		14.9837	16.896
	TIME		49.331	93.855

Table 2 Comparisons of RMSE and CPU time for Bridge.

DWT Orders		DCT with edge detection	DWT with Quadtree	Region-Based DWT
1	RMSE	15.6384	16.858	18.5994
	TIME	11.717	41.099	51.554
2	RMSE		15.7557	16.9019
	TIME		55.42	71.843
3	RMSE		14.5753	15.2804
	TIME		75.108	92.964

5. CONCLUSION

This paper proposes a new method for image fusion. It combines a preprocessing using Quadtree or edge detection for segmentation. DWT or DCT is then applied for fusion. We have shown that using segmentation can select some obvious areas that do not need go through the transform processes. Hence the computational cost is significantly reduced. The quality of reconstruction is also increased, as fewer disturbances are included. The results will definitely promote the use of image fusion in many new applications.

6. REFERENCES

- [1]R. Luo and M. Kay, "Data fusion and sensor integration: state of the art in 1990s," in *Data Fusion in Robotics and Machine Intelligence*, pp. 7-136, Academic Press, San Diego, 1992.
- [2]M. Pavel, J. Larimer, and A. Ahumada, "Sensor fusion for synthetic vision", in *Proceedings AIAA Conference on Computing in Aerospace*, Baltimore, MD, Oct. 1991.
- [3]H. Li, B. S. Manjunath, and S. K. Mitra, "A contour based approach to multisensor image registration", *IEEE Trans. Image Processing*, vol. 4, no. 3, pp. 320-334, March 1995.
- [4]P. J. Burt, "The pyramid as structure for efficient computation," in *Multiresolution Image Processing and Analysis*, A. Rosenfeld, Ed., pp. 6-35, Springer-Verlag, New York/Berlin, 1984.
- [5]P. J. Burt and R. J. Lolczynski, "Enhanced image capture through fusion", in *Proceedings of the Fourth International Conference on Computer Vision*, Berlin, Germany, pp. 173-182, May 1993.
- [6]Li, H.; Manjunath, B. S. and Mitra, S. K. "Multisensor image fusion using the wavelet transform," *Graphical Models and Image Processing*, vol. 57, no. 3, S. 235-245, 1995.
- [7]A. Toet, "Hierarchical image fusion," *Mach. Vision Appl.* pp. 1-11, Mar. 1990.
- [8]A. Toet, "Multiscale contrast enhancement with application to image fusion," *Opt. Eng.* vol. 31, pp. 1026-1039, 1992.
- [9]J. Vaisey, and A. Gersho, "Image Compression with Variable Block Size Segmentation," *IEEE Trans. on Signal Processing*, vol. 40, no. 8, pp. 2040-2060, Aug 1992.
- [10] C.T. Chen, "Adaptive Transform Coding via QuadTree-Based Variable Blocksize DCT," in *Proceedings of ICASSP'89*, pp. 1854-1857, May 1989.
- [11]R. Distasi, M. Nappi and S. Vitulano, "Image Compression by B-Tree Triangular Coding," *IEEE Trans. on Commu.*, vol. 45, no. 9, pp. 1095-1100, Sep. 1997.

- [12]C.Y. Teng and D.L. Neuhoff, "A new quadtree predictive image coder," in *Proceedings of International Conference on Image Processing* , vol.2 , pp. 73 –76, 1995.
- [13]A. Jain, "Fundamentals of digital image processing," Prentice-Hall, 1989.
- [14]L. Roberts, "Machine perception of three-dimensional solids," in *Computer Methods in Image Analysis*, IEEE Computer Society, 1977.
- [15]Stephane G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674-693, July 1989.