Edge-Preserving Wavelet-Based Multisensor Image Fusion Approach

Lahouari Ghouti, Ahmed Bouridane and Mohammad K. Ibrahim

Abstract— **Images emanating from multiple sensors have been successfully exploited to reduce human and machine errors in practical vision systems. Multiresolution-based schemes have shown interesting potential in the fusion of images obtained from possibly different types of sensors that need to be combined. However, most of the proposed schemes treat all image features equally regardless of their local importance. On the other hand, the human visual system is more sensitive to edges and sharp details. We propose an image fusion scheme where image edges, characterized by wavelet maxima, are considered separately from plain and low activity image regions. This edge-guided fusion offers a trade-off between feature-based and pixel-level fusion schemes. Images are combined in the wavelet domain using a multiresolution representation that is more sensitive to image edges. A comparison of the proposed method with current multiresolution-based fusion schemes shows that the proposed method can achieve better performance in combining and preserving important details in the combined images.**

I. Introduction

In recent years, there has been a growing interest in merging images obtained using multiple sensors in academia, industry, and military due to the important role it plays in the applications related to these fields. Image fusion, a class of data fusion, aims at combining two or more source images from the same scene into an image that retains the most important or salient features present in all the source images according to a specific fusion scheme. The composite image should provide increased interpretation capabilities and significantly reduce both human and machine errors in detection and object recognition. Moreover, image fusion can be performed at three different processing levels according to the stage where the fusion takes place: pixel [1], feature [1] and decision level. In this paper, we are interested in developing a fusion scheme that combines aspects of pixel-level and feature-level fusion approaches.

In pixel-level approach, all or a set of selected pixels in the source images are combined to contribute to each pixel in the fused image. Simple arithmetic rules or more sophisticated combination schemes can be applied to serve this purpose. It is worth noting that the adopted merging procedure should, in essence, contribute to a considerable performance improvement for all posterior processing tasks such as object detection and human/machine vision.

L. Ghouti is with the Information and Computer Science Department. King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia and the School of Computer Science, Queen's University of Belfast, BTN 1NN, UK. Email: ghouti@ccse.kfupm.edu.sa

A. Bouridane is with the School of Computer Science, Queen's University of Belfast, BTN 1NN, UK. Email: A.Bouridane@qub.ac.uk

M. K. Ibrahim is with the Computer Engineering Department. King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia. Email: ibrahimm@ccse.kfupm.edu.sa

Feature extraction plays a major a role in the implementation of feature-level fusion approaches. Prior to the merging of images, salient features, present in all source images, are extracted using an appropriate feature extraction procedure. Then, fusion is performed using these extracted features. The extraction/detection stage should be optimal with respect to image salient features and important regions.

Multiresolution analysis MRA, rooted in computer vision, has received a lot of interest for its attractive features. The application of MRA to image fusion is no exception. The most attractive features are: MRA tools decompose an image into bands that vary in spatial frequency and orientation. This decomposition is very similar to that performed by the human visual system (HVS). Therefore, MRA-based image fusion is considered as one of the most powerful fusion methods. Source images are decomposed into various wavelet subbands of specific orientation and level[1]. All or selected wavelet coefficients are combine to produce wavelet coefficients of the fused image. Finally, the merged wavelet coefficients are reconstructed to produce the fused image. Burt and Adelson [2] are credited to be the first who proposed MRA-based fusion for image coding and binocular fusion in human vision. Their implementation is based on a pyramid structure called Laplacian pyramid.

Toet [3] proposes a contrast pyramid approach to image fusion of of thermal and visual images. Akerman [4] proposes pyramid techniques for the fusion of images emanating from multiple sensors. Burt and Lolczynski [5] suggest the used of image fusion for the enhancement of image capture. Ranchin et al. [6], present an investigatiuon of MRA-based image. They implemented the proposed scheme through wavelet decomposition. The proposed scheme allows enhancing the spatial resolution of a SPOT image by making use of another image obtained from the same satellite using a different band.

Chipman et al. [7] propose an algorithm for the fusion of multispectral aerial photos. The proposed algorithm is based on a set of basic operations on particular sets of wavelet coefficients corresponding to specific wavelet subbands.

In [8], Li et al. develop an algorithm for MRA-based image fusion. The proposed scheme employs an area-based selection rule to pick up the maximum. A consistency verification step is outlined.

Existing MRA-based methods have not investigated the incorporation of an HVS model that is expected to improve the overall performance of any MRA-based fusion scheme.

In this paper, we propose an image fusion scheme based

[1]Usually, it is noted that decomposition beyond level 5 does not provide any improvement in performance.

a special wavelet transform and a particular combination algorithm where wavelet coefficients are considered according to their contribution to the source image saliency. Edges and other high activities regions are first localized through wavelet maxima. Unlike most of the proposed fusion schemes, we focus, in our proposed approach, on edges and salient regions that carry most of the image activity to which the HVS is most sensitive. The proposed fusion scheme is described in Section 2. Some experimental results are presented in Section 3. Conclusions are drawn in Section 4.

II. Edge-Preserving Image Fusion Scheme

One of the main issues in image fusion is image alignment, which refers to pixel-by-pixel alignment of the images. For our algorithm, we assume that the source images are registered, so that the corresponding pixels are aligned. Another important issue is the dynamic range between the source images. Since different sensors provide these images, the dynamic range of the source images must be rescaled to match each other[2]. The non-orthogonal (redundant) wavelet decomposition is applied on the two source (registered) images. The wavelet coefficients are then decomposed into regions according to their ”*edge-importance*”. Then, the wavelet maxima of both images are simply combined to produce the fused wavelet maxima. Wavelet maxima are detected using a detection rule proposed by Mallat and Zhong [9]. Figure 1 shows the proposed fusion algorithm.

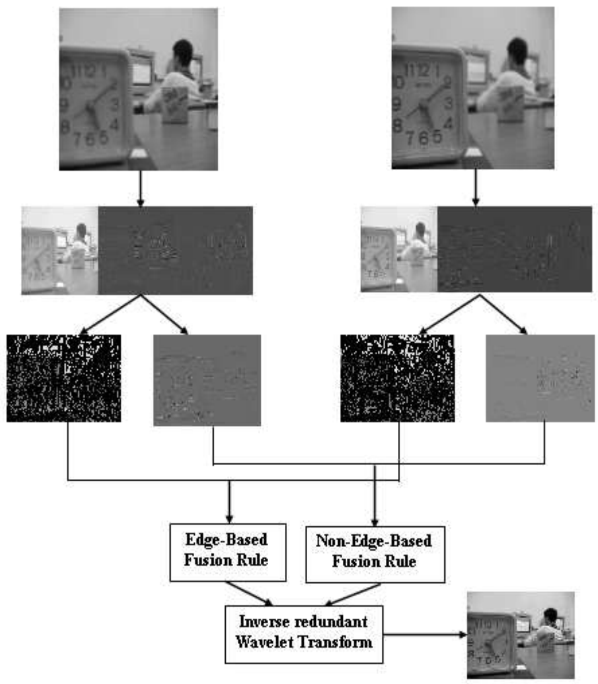

Figure 1: Functional processing of proposed fusion algorithm.

A. Redundant non-orthogonal Wavelet Transform

The orthogonal decomposition is an exact (non-redundant) representation of the analyzed image. It involves subsampling after each decomposition stage as shown in Figure 2. Hence, the approximation image is not smoother than the original image because of the frequency spread after subsampling. If the wavelet coefficients undergo a modification (coefficient merging, quantization, etc.), then the inverse transform preserves this modification because the transform is non-redundant [9].

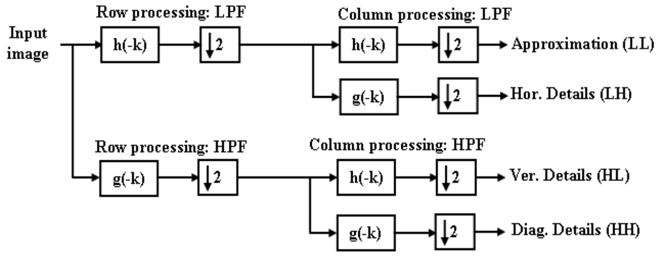

Figure 2: Two-dimensional orthogonal wavelet decomposition: Analysis stage.

On the other hand, non-orthogonal wavelet decomposition does not involve the subsampling stage after filtering at each scale as illustrated in Figures 3, 4. This decomposition has two important features [9]. First, the image dimensions are preserved so that the image dimensions at any scale are in one-to-one correspondence with the image dimensions at the finest scale. The second advantage is that the approximation image at coarser scales is smoother than that at a finer scale [9]. However, non-orthogonal decomposition is redundant, i.e., not every two-dimensional representation is a valid transform [3].

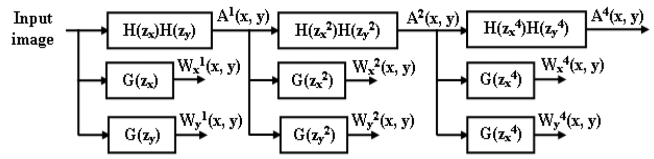

Figure 3: Redundant 2D wavelet decomposition.

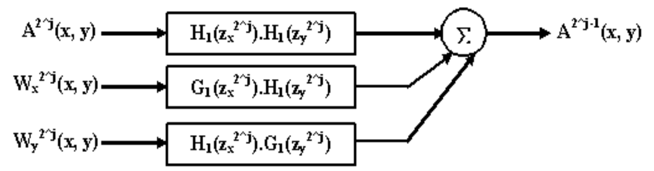

Figure 4: Two-dimensional orthogonal wavelet decomposition: Single-stage reconstruction.

The 2D extension of the redundant wavelet transform is based on tensor product two 1D bases. If the lowpass and the highpass filters of the 1D cubic spline wavelet are $H(z)$ and $G(z)$ respectively, then the 2D non-orthogonal decomposition and reconstruction follow the implementation shown in Figures 3 and 4, respectively. The adopted representation gives three components at each stage, namely the lowpass, the x-channel and the y-channel components. The x-channel and y-channel components give the derivative in the vertical and horizontal directions, respectively [9].

B. Wavelet Maxima

Perceptually important or salient points in each of the source images, that represent the dominant image details such as

[2]We will illustrate our fusion algorithm using two source images. Extension to the case with multiple images is straightforward.

[3]Over-complete representations are characterized by frame theory [9].

edges, should be preserved during the merging process to produce a high fidelity fused image. Specific wavelet basis provide a direct mapping between the image edges and the maxima of the wavelet coefficients. Mallat and Zhong [9] suggest the use of the derivative of the cubic spline as a wavelet basis. This basis is the first derivative of the cubic spline smoothing function. Hence, a local extremum in the wavelet coefficients represents a maximum or a minimum of the derivative of the original signal. These extrema points are the locations of the signal edges [9]. Because the source image will be decomposed in many scales, the wavelet coefficients, at the three decomposition scales, will be correlated to locate the peaks that correspond to signal edges. After correlation, the peaks corresponding to true edges will survive while other peaks will be suppressed through correlation. The principle underlying the detection of "pure" wavelet maxima is illustrated in Figure 5.

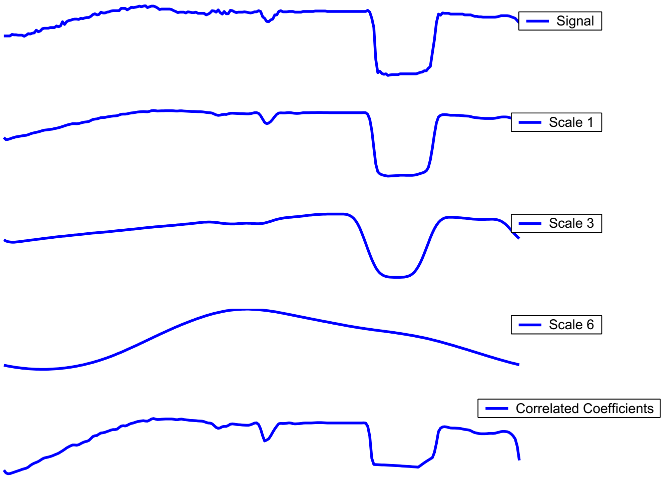

Figure 5: Illustration of wavelet exrema extraction.

For 2D signals, an extrema at point (x_o, y_o) of the x-channel component is defined as:

$$Maximum = W_x\left(s; x_o, y_o\right) > \\ max\left[W_x\left(s; x_o+1, y_o\right), W_x\left(s; x_o-1, y_o\right)\right] \quad (1)$$

$$Minimum = W_x\left(s; x_o, y_o\right) < \\ min\left[W_x\left(s; x_o+1, y_o\right), W_x\left(s; x_o-1, y_o\right)\right] \quad (2)$$

where $W_x\left(s; x_o, y_o\right)$ is the wavelet coefficient at location point (x_o, y_o) in the x-channel component at scale s. Similar relations can be derived for the extrema points in the y-channel component. Importance and properties of the wavelet maxima have been investigated by Berman and Baras [10], Mallat and Zhong [9] and Mallat [9]. Berman and Barras [10] argue that one can completely reconstruct an image from its multiscale edges. This reconstruction can be implemented using the wavelet extrema in a redundant wavelet representation. Cetin and Ansari [11] propose a method for signal recovery from wavelet maxima. Multiple projections onto convex sets, derived from the image, are used for the reconstruction of an image.

C. *Fusion Using Dominant Image Features*

The x-channel and y-channel components, $W_x(s; x, y)$ and $W_y(s; x, y)$, can be viewed as the two components of the gradient vector of the analyzed image, $I(x, y)$, smoothed by a lowpass filter obtained from the cubic spline function defined in [9]. Following the approach outlined in [9], we may define a modulus and angle image at each scale s:

$$Mag\left(s; x, y\right) = \sqrt{\left|W_x\left(s; x, y\right)\right|^2 + \left|W_y\left(s; x, y\right)\right|^2} \quad (3)$$

$$Ang\left(s; x, y\right) = \arctan\left(\frac{W_x\left(s; x, y\right)}{W_y\left(s; x, y\right)}\right) \quad (4)$$

The sharper variation points of $I(x, y)$ smoothed at scale s correspond to the maxima of $Mag(s; x, y)$ along the gradient direction. Mallat and Zhong [9] argue that this maxima detection is essentially equivalent to Canny's non-maxima suppression without requiring the same assumptions[4].

Figures 6 illustrates this idea and its implication in the image fusion algorithm developed in this paper. In Figure 6, we show the source image with left focus and its wavelet maxima representation. It is clear from this figure that the edges at the focus region are more emphasized by this representation.

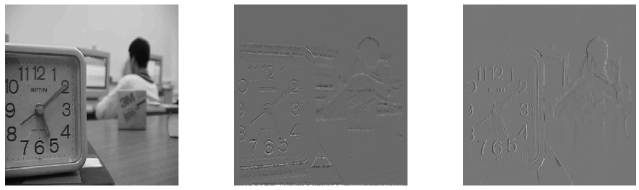

Figure 6: Dominant image features in source image with left focus.

The proposed fusion algorithm performs a "*region labeling*" through the computation of wavelet maxima and then the detection of edges by picking up most dominant maxima in the lowpass, x-channel, and y-channel components by using Equations 1 and 2 defined above. The "*edge*" images, as illustrated in Figure 6, represent "*labeled*" images where nonzero values indicate edges and non-edge regions are indicated by their wavelet coefficients set to zero. Unlike several proposed fusion algorithms, complete processing and region labeling in the wavelet domain represent an attractive characteristic of the proposed scheme that allows simultaneous localization in space and scale [9].

D. *Fusion Procedure*

Once the image details are classified through the amplitude of their wavelet coefficients, the edge regions pertaining to both source images are merged using simple addition operation. Less perceptually image regions, with zero labels, will be merged using more sophisticated operations. For the latter regions, we will first estimate the block activity as the average of the absolute value of the wavelet coefficients in the x-channel and y-channel components. The block activity measure, at location $(x_o, y_o)$is given by [12]:

$$A\left(x_o, y_o\right) = \frac{W}{M}\sum_{m=1}^{M}\frac{1}{3\cdot 2^{2(M-m)}} \\ \sum_{i=1}^{3\cdot 2^{2(M-m)}} \left|W_x\left(s; x, y\right)\right| + \left|W_y\left(s; x, y\right)\right| \quad (5)$$

[4]Canny edge detector is optimal for step edges in the presence of Gaussian noise. Of course, real edges are not simple steps, and real noise is not purely Gaussian in nature. Nevertheless, the canny detector gives good results with real images.

where W is a weight based on the signal-to-noise ratio (*SNR*) of the source image, M is the number of decomposition levels s. The second sum in Equation 5 involves all the wavelet coefficients, corresponding to a specific image pixel at location (x_o, y_o), in the x-channel and y-channel components.

III. Simulation and Experimental Results

Performance analysis of the proposed algorithm is summarized in this section. The test images consist of several pairs of natural images taken using a Sony digital still camera *MVC-FD7*[5]. Figure 7 shows two pairs of test images with opposite side focus. Various merging operations have been applied

Figure 7: Test images with left focus (left column) and right focus (right column).

to merge non-edge regions using Equation 5, while edge regions have been merged such that all the edge information is preserved in the fused image.

To compare fusion effects, several other methods were used to fuse the source images. The fused images using the proposed method are shown in Figure 8. Table I shows a summary of the performance of many fusion schemes to fuse the second pair of images shown in Figure 7 (second row). A subjective fusion quality measure can be provided by comparing the original image[6] with the fused image. For such comparisons, we will use the universal image quality index provided by Wang and Bovic [13]. An index value closer to 1 indicates better perceptual quality of the fused image and its closeness to the original one. These examples clearly illustrate that our algorithm provides fused images with the best visual quality.

Fusion Algorithm	**Universal Quality Index**
Simple Averaging	0.7952
Principal Component Analysis	0.7962
Maximum Selection	0.7439
Minimum Selection	0.7497
Laplacian Pyramid	0.8259
Gradient Pyramid	0.7718
DWT (Daubechies Basis)	0.8178
Shift-Invariant DWT (Haar Basis)	0.8433
Wavelet Maxima	**0.8648**

Table I: Performance of several fusion algorithms: Fusion quality.

[5]The authors would like to thank Prof. Rick Blum, ECE Department, Lehigh University, for kindly providing the test images.

[6]The original image is produced using a simple cut-and-paste technique, taking physically the "*in focus*" areas of each source image.

Figure 8: Fused images using proposed algorithm.

IV. Conclusions

In this paper, we have proposed a novel image fusion scheme that preserves edge details pertaining to source images. Unlike existing fusion methods, the proposed algorithm takes into account in a comprehensive manner image edges, better known as " *high activity*" regions. The basic algorithm idea lies in the discrimination between edge and non-edge regions using wavelet maxima measures. An assessment of the proposed technique has been carried out in terms of the performance achieved against a number of existing methods using universal quality index. The results obtained clearly show the superior performance of the proposed technique.

Acknowledgment: L. Ghouti and M. K. Ibrahim acknowledge the support of KFUPM University.

References

[1] R.C. Luo and M.G. Kay, *Multi-sensor Integration and Fusion for Intelligent Machines and Systems*, Ablex Publishing Corp., 1995.

[2] P. J. Burt and E. Adelson, "The Laplacian pyramid as a compact image code," IEEE Trans. Communications, *IEEE Trans. Communications vol.* COMM-31, No. 4, pp. 532-540, 1983.

[3] A. Toet, "Hierarchical image fusion," Machine Vision App., pp. 1-11, March 1990.

[4] A. Akerman III, "Pyramid techniques for multisensor fusion," SPIE Proceedings on Sensor Fusion V, Vol. 1828, pp. 124131, 1992.

[5] P. J. Burt and R. J. Lolczynski, "Enhanced image capture through fusion," in Proc. of the 4th Intl. Conf. on Computer Vision, pp. 173182, Berlin, Germany, May 1993.

[6] T. Ranchin, L. Wald and M. Mangolini, "Efficient data fusion using wavelet transform: the case of spot satellite images," in SPIE Proc. on Mathematical Imaging: Wavelet App. in Signal and Image Proc., Vol. 2934, pp. 171178, 1993.

[7] L.J. Chipman, T.M. Orr and L.N. Graham, "Wavelets and image fusion," in SPIE Proc. on Wavelet App. in Signal and Image Proc. III, Vol. 2569, pp. 208219, 1995.

[8] H. Li, B. S. Manjunath and S. K. Mitra, "Multisensor image fusion using the wavelet transform," Graphical Models and Image Processing, Vol. 57, No. 3, pp. 235245, May 1995.

[9] S. Mallat, *A Wavelet Tour of Signal Processing*, Acaemic Press, Second Edition, 1998.

[10] Z. Bermand and J. Baras, "Properties of the multiscale maxima and zero-crossings representations," IEEE Trans. on Signal Processing, Vol. 41, No. 12, pp. 3216-3231, Dec. 1993.

[11] A. Cetin and R. Ansari, "Signal recovery from wavelet transform maxima," IEEE Trans. on Signal Processing, Vol. 42, No. 1, pp. 194-196, Jan. 1994.

[12] Z. Zhang and R. S. Blum, "MRegion-based image fusion scheme for concealed weapon detection," in Proc. of 30th Conf. on CISS, March 1997.

[13] Z. Wang and A. C. Bovik, "A Universal image quality index," IEEE Signal Proc. Letters, Vol. 9, No. 3, pp. 8184, March 2002.