4-17-2006

Cross-sensor fusion of imagery for improved information extraction

Samuel Peter Kozaitis

Michel Ouendeno

Cross-sensor fusion of imagery for improved information extraction

S. P. Kozaitis and M. Ouendeno
Florida Institute of Technology
Department of Electrical and Computer Engineering
150 W. University Blvd.
Melbourne, FL 32901

ABSTRACT

We combined cross-sensor data that leads to improved extraction of information from disparate sensors. We presented a new method for signal fusion that uses different transforms for the forward transforms of two images and a common transform for the inverse. When using a fusion rule that selects the maximum value between images, we were able to transfer more energy to the result using our method. Our method could form the basis of a new image fusion approach because it offers a way to transfer more energy to the result not possible with a conventional approach.

Keywords: data fusion, image transform, wavelet, wavelet transform

1. INTRODUCTION

One way to overcome the lack of information in an image is to combine imagery from multiple sensors. Sensors are often suited to the detection of some characteristic that may not be possible with another type of sensor. For example, IR sensors are most sensitive to thermal characteristics, LADAR sensors respond to surface geometry, electro-optical sensors can provide color, and SAR sensors can be very sensitive to metallic objects. A method is needed that combines imagery in such a way as to maximize the useful information into a single viewable image. Such a method could form the basis for many applications such as medical imaging, homeland security, remote sensing, and target recognition.

The use of contrast has been a popular way of representing relevant information in an image. Recent advances include, an optimal grayscale visualization of the contrast of a hyperspectral image[1]. In this approach, the hyperspectral volume was thought of as a surface, and the maximum contrast was developed at each point to determine a contrast vector field. The first derivative in the horizontal direction and vertical directions of an image were found and recorded in separate contrast images. Using differential geometry, a two-dimensional vector field representing the absolute contrast and direction was found. Therefore, the maximum contrast at each point was determined.

A general way of combining information uses a multiresolution approach of orthogonal wavelet representations of images.[2] Although not developed specifically for image fusion, the orthogonal representation accumulates the results of images into a single representation. Using an eigenvector approach the maximum direction of the contrast can be found from the resulting representation.

Another approach was developed on the premise that the location of a feature should be localized. In an effort to localize the contrast, gradient maps of images at multiple resolutions were used as the basis for the fusion decision process.[3] A gradient map is formed from the difference between adjacent pixels as the Euclidian distance of the contrast components in the horizontal and vertical directions. The wavelet transform was used to process the gradient maps. The filters used in the transform were a function of a wavelet and the z-transform of the gradient operator so that a change in intensity in an image was represented locally in the wavelet domain.

Independent Component Analyses, Wavelets, Unsupervised Smart Sensors, and Neural Networks IV,
edited by Harold H. Szu, Proc. of SPIE Vol. 6247, 624707, (2006) · 0277-786X/06/$15 · doi: 10.1117/12.666460

We developed a new method for combining cross-sensor image data that preserves sensor-unique characteristics of each contributing sensor. The basis of our approach is a set of forward wavelet transforms, optimized to preserve the individual information content of each sensor image followed by a common inverse transform that is designed to preserve the most desirable information characteristics of each image after fusion. Our approach has broad applicability to any image fusion problem involving sensor data having different spatial, spectral or temporal characteristics

2. FUSION

2.1 Data Fusion

Data fusion can be thought of combining two signals represented as vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ and somehow extracting the maximum amount of information. Often, transform methods are used because the signals can be represented more compactly in the transform domain, than in the time or spatial domain. In matrix form, the transform of $\mathbf{x}_1$ and $\mathbf{x}_2$ are $\mathbf{Ax}_1$ and $\mathbf{Ax}_2$ respectively, where $\mathbf{A}$ is a matrix and represents a transform. Combining the signals is represented as

$$\mathbf{G} = f(\mathbf{Ax}_1, \mathbf{Ax}_2), \tag{1}$$

where $f(\ )$ represents a fusion function or rule. Then, the resultant image is obtained through the inverse transform as $\mathbf{R} = \mathbf{A}^{-1}\mathbf{G}$. The signal $\mathbf{R}$ represents the fusion of the two signals. Because $\mathbf{A}$ is often a square matrix its inverse is unique if one exists. Therefore, the transforms operating on the two input signals must be the same. If the signals are from disparate sensors, then finding a transform $\mathbf{A}$ that works best for one signal may not be the best for the other signal. If a transform is found that works similarly for both signals, it might not necessarily work well for either signal.

2.2 Data Fusion using wavelet transforms

In our approach, we attempted to use different transforms for signals from disparate sensors, and use a single inverse transform. If such an approach is successful, it could offer many new solutions to data fusion problems because we can use many combinations of transforms rather than just one. The wavelet transform offers a convenient framework for signal fusion. The wavelet transform typically compacts energy efficiently, and has a variety of basis functions that can be used.

The wavelet transform can be represented as a matrix as

$$\mathbf{W} = \begin{bmatrix} \mathbf{L} \\ \mathbf{B} \end{bmatrix} \tag{2}$$

where $\mathbf{A}$ is square matrix, and the rows of $\mathbf{L}$ and $\mathbf{B}$ represent the dual bases of the wavelet transform, the scaling and wavelet functions respectively. We considered different transforms for the two input signals, and for the synthesis transform. A block diagram of the system is shown in Fig. 1 with

$$\mathbf{W}_1 = \begin{bmatrix} \mathbf{L}_1 \\ \mathbf{B}_1 \end{bmatrix},\ \mathbf{W}_2 = \begin{bmatrix} \mathbf{L}_2 \\ \mathbf{B}_2 \end{bmatrix}, \text{and } \mathbf{W}_3 = \begin{bmatrix} \mathbf{L}_3 & \mathbf{B}_3 \end{bmatrix}. \tag{3}$$

Wavelet and scaling functions are usually complicated to describe, and implementation often employs digital filter banks. Therefore, wavelet and scaling functions are usually specified through filter coefficients. The wavelet transform can be efficiently implemented in terms of filter banks consisting of filters and decimators. Usually, a two-channel filter bank is used with low-pass and high-pass filters. The forward transform is determined by low-pass, and high-pass filters, $H_0(z)$ and $H_1(z)$, and the inverse transform is determined by low-pass and high-pass filters, $F_0(z)$ and $F_1(z)$ as indicated in Fig. 2.

If we consider only one level of the transform, then the **L**'s and **B**'s in Eq. 3 can be viewed as low- and high-pass filters, and written in the z-domain as,

$$\begin{array}{lll} \mathbf{L}_1 \rightarrow H_{01}(z), & \mathbf{L}_2 \rightarrow H_{02}(z), & \mathbf{L}_3 \rightarrow F_{03}(z), \\ \mathbf{B}_1 \rightarrow H_{11}(z), & \mathbf{B}_2 \rightarrow H_{12}(z), & \mathbf{B}_3 \rightarrow F_{13}(z). \end{array} \tag{4}$$

We wanted the wavelets in the different forward transforms to be different so they could have different characteristics. Therefore we allowed $H_{11}(z) \neq H_{12}(z)$, and set $H_{01}(z) = H_{02}(z)$ to relate the two transforms.

In a biorthogonal wavelet transform, the high-pass filter of the synthesis transform is related to the low-pass filter of the forward transform.[4] Therefore we set the high-pass filter of $\mathbf{W}_3$ to

$$F_{13}(z) = -H_{01}(-z) = -H_{02}(-z). \tag{5}$$

If we choose $H_{11}(z)$ and $H_{12}(z)$ based on the characteristics of the sensors and the type of features we are considering, then the problem reduces to finding a suitable $F_{03}(z)$.

3. EXPERIMENT

Features of an image are often related to its derivatives. Under the right conditions, the forward wavelet transform can be thought of as a multiscale differential operation If a wavelet has p vanishing moments, it can be interpreted as a differential operator of order p. The value of p can be determined from the low-pass filter $H_0(z)$, and is equivalent to the number of zeros the filter has at $z = -1$ in the z-plane. The inverse transform can be thought of reconstructing a signal with polynomials. Because $H_{01}(z)$ and $H_{02}(z)$ and $F_{13}(z)$ are related as in Eq. (5), when $H_{01}(z)$ or $H_{02}(z)$ has p zeros at z = -1, $F_{13}(z)$ has p vanishing moments. Therefore, a combination of scaling functions with p zeros can reproduce a polynominal up to order p - 1 on any interval.

In our experiments we considered two different transforms for $\mathbf{W}_1$ and $\mathbf{W}_2$. In both transforms, the low-pass filter were the same and the high-pass filters were different, $H_{11}(z) \neq H_{12}(z)$, and set $H_{01}(z) = H_{02}(z)$. The low pass filters have two zeros at z = -1, and the high-pass filters $H_{11}(z)$, and $H_{12}(z)$, had two and four zeros at z = -1 respectively. Their filter coefficients (normalize, and in limited precision) are shown in Table 1.

Table 1 Filter coefficients for transforms $\mathbf{W}_1$ and $\mathbf{W}_2$

$H_{01}(z)$	0.3536	0.7071	0.3536						
$H_{11}(z)$	0.3536	0.3536	-1.0607	0.3536	0.3536				
$H_{02}(z)$	0.3536	0.7071	0.3536						
$H_{12}(z)$	-0.0331	-0.0663	0.1768	0.4198	-0.9944	0.4198	0.1768	-0.0663	-0.0331

We considered the two registered images in Figs. 3(a) and (b) in our experiments. The images were 256 x 256 pixels in size. Transform $\mathbf{W}_2$ was used on image 1, and $\mathbf{W}_1$ was used on image 2. The fusion rule was one where the maximum value of the wavelet coefficients was used on a pixel-by-pixel basis between the two image transform results. For the lower subbands of the image not processed by a wavelet, we used a linear combination between the transforms. The synthesis transform was $\mathbf{W}_1$ and the result for one level of a transform is shown in Fig. 3(c).

Since the fusion rule transfers the maximum amount of energy to the result, we examined the amount of energy in the result for various situations. We considered three different experiments, using one through three levels of the transform. We showed in Table 2 the energy in the output and the value of a quality index using 16 x 16 pixel blocks.[5] Note that our quality index was not normalized so values above 1 are permitted. In addition, we compared the results to when only transform $\mathbf{W}_2$ was used for both input images and synthesis.

Table 2 Image fusion results

		$\mathbf{W}_1 = \mathbf{W}_3$	$\mathbf{W}_1 = \mathbf{W}_2 = \mathbf{W}_3$
Level 1	Quality	1.20	1.18
	Energy	1291	1276
Level 2	Quality	1.47	1.39
	Energy	1464	1430
Level 3	Quality	1.66	1.51
	energy	1636	1592

4 CONCLUSION

We combined cross-sensor data that leads to improved extraction of information from disparate sensors. We expect to reconstruct a fused image that will retain the information content in disparate domains while enhancing the information content of the fused image product. We presented a new method for signal fusion that uses different transforms for the forward transforms of two images and a common transform for the inverse. Using the wavelet transform, we were able to use different high-pass filters for the forward transform. In this way, the transforms can be tailored to specific type of imagery such as infrared and visible. We used the same low-pass filters in the forward transforms to relate the two transforms, and derived the synthesis high-pass filter from this filter. When using a fusion rule that selects the maximum value between images, we were able to transfer more energy to the result using our method. In addition, using a quality index the results also improved using our approach. Therefore, our method could form the basis of a new image fusion approach because it offers a way to transfer more energy to the result not possible with a conventional approach. Anywhere sensors with different characteristics can be used, this method can be applied including, tracking, mulitsource tracking, and handing-off between sensors.

REFERENCES

1. D. A. Socolinsky, and L. B. Wolff, "Multi-spectral image visualization through first order fusion," *IEEE Transactions on Image Processing*, 11, 923-931, 2002.

2. P. Scheunders, "An orthogonal wavelet representation of multivalued images," *IEEE Transactions on Image Processing*, 12(6), 718-725, 2003.

3. V. S. Petrovic, and C. S. Xydeas, "Gradient-based Multiresolution Image Fusion," *IEEE Transactions on Image Processing*, 13(2), 228-237, 2004.

4. G. Strang and T. Nguyen, "Wavelet and filter banks,“ Wellesley-Cambridge Press: Wellesley, MA, p. 105, 1996.

5. G. Piella,and H. Heijmans, "A new quality metric for image fusion," in *Proc. Int. Conf. on Image Processing*, Barcelona, Spain, 2003.

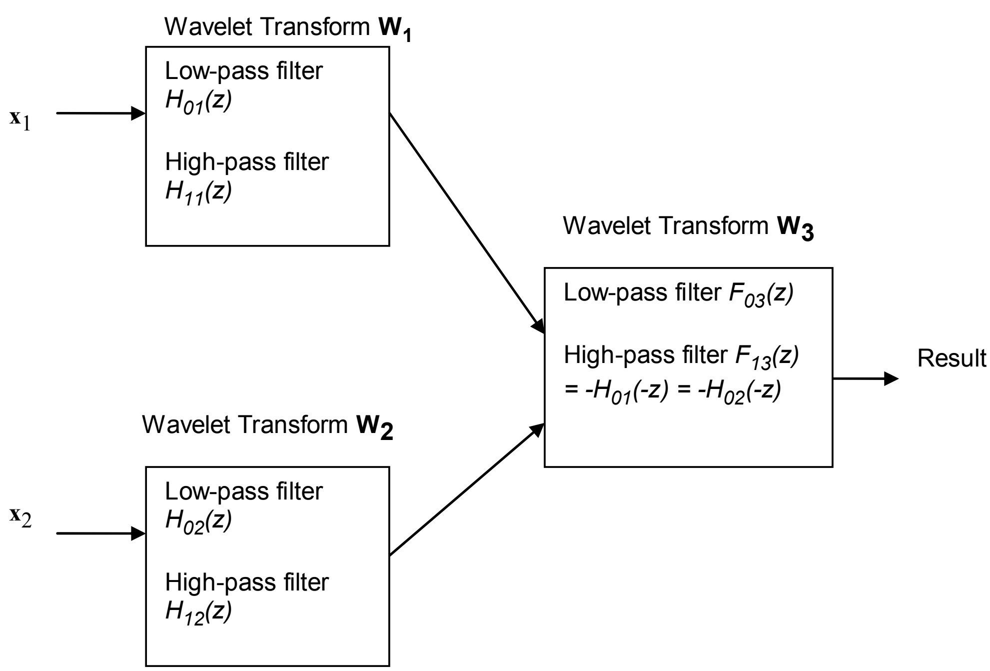


Figure 1 Block diagram of image fusion process.

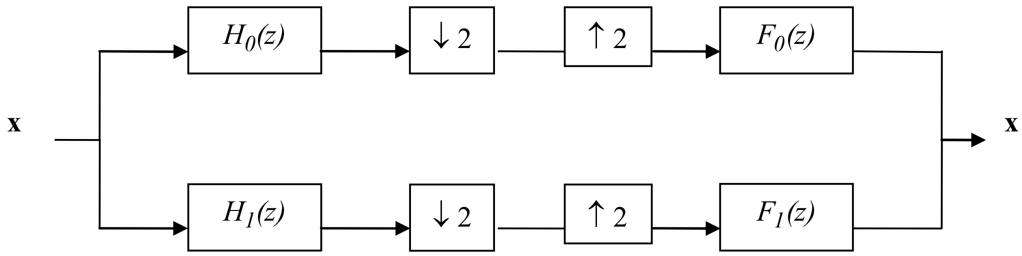
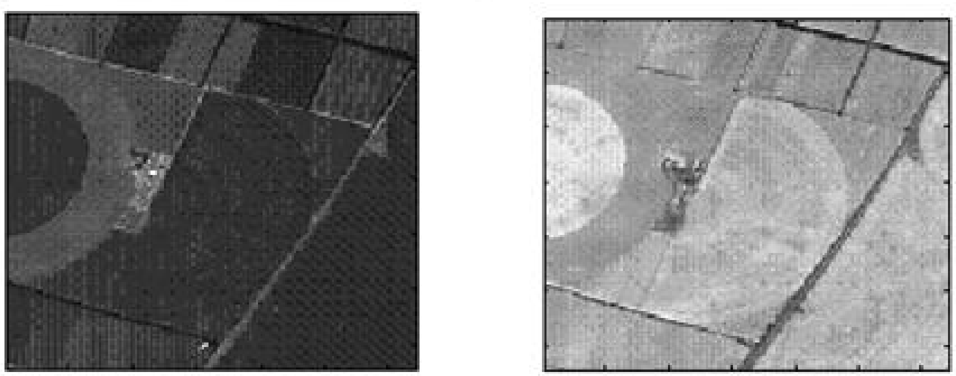
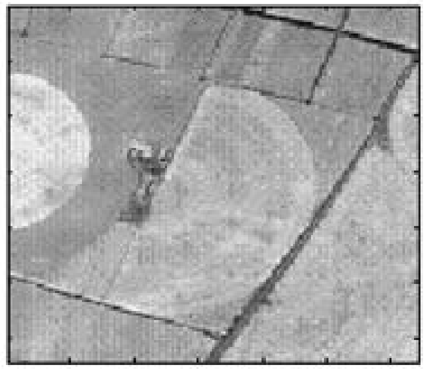


Figure 2 Filter bank and its inverse representing one level of the wavelet transform.



(a) (b)



(c)

Figure 3 Image fusion results (a) input image 1 (b) input image 2 (c) fusion results using $\mathbf{W}_1$ on image 1, $\mathbf{W}_2$ on image 2, and $\mathbf{W}_1$ for synthesis.