Wavelet-based fusion approach using unique reconstruction approach

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

ABSTRACT

We used measures based on entropy to evaluate a method designed to fuse imagery from different sensor types. The method uses different forward transforms of input images and a common transform to reconstruct the final result. We attempted to examine the link between the error in a reconstructed result and its associated entropy.

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

1. INTRODUCTION

It has often been difficult to evaluate sensor fusion methods because of the lack of reliable metric. Usually, methods based on contrast, luminance or other factors are used and have shown to be useful.[1-3] These seem to work well for conventional fusion methods. We used a method to fuse imagery that combines that uses a set of forward wavelet transforms, one for each sensor followed by a common inverse transform. Using this approach we hope to combine sensor imagery for true image fusion. We envision a broad applications space for this approach in a rich set of diverse image products that must be quickly and optimally combined. In the next section we briefly discuss our approach followed by some statistical results using measures based on entropy for evaluation.

2. 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 transforms of $\mathbf{x}_1$ and $\mathbf{x}_2$ are $\mathbf{A}\mathbf{x}_1$ and $\mathbf{A}\mathbf{x}_2$ respectively, where $\mathbf{A}$ is a matrix and represents a transform. Combining the signals in the transform domain is represented as

$$\mathbf{g} = f(\mathbf{A}\mathbf{x}_1, \mathbf{A}\mathbf{x}_2), \tag{1}$$

where $f(\ )$ represents a fusion function or rule. Then, the resultant image is obtained through the inverse transform as $\mathbf{r} = \mathbf{A}^{-1}\mathbf{g}$, where the signal $\mathbf{r}$ represents the fusion of the two signals in the spatial domain. However, if the input signals are from disparate sensors, then finding a transform $\mathbf{A}$, that say, maximizes $\|\mathbf{g}\|$ for $\mathbf{x}_1$, may not do so for $\mathbf{x}_2$. If an $\mathbf{A}$ is found that works similarly for both signals, then the value of $\|\mathbf{g}\|$ may not be maximized.

Independent Component Analyses, Wavelets, Unsupervised Nano-Biomimetic Sensors, and Neural Networks V, edited by Harold H. Szu, Jack Agee, Proc. of SPIE Vol. 6576, 65760G, (2007) · 0277-786X/07/$18 · doi: 10.1117/12.718652

In our approach, we used different forward transforms, $\mathbf{A}_1$, $\mathbf{A}_2$ for signals from disparate sensors. If such an approach is successful, it could offer many new solutions to data fusion problems because we can use many combinations of transforms. In this case the fused image, **g** can be represented as

$$\mathbf{g} = f(\mathbf{A}_1\mathbf{x}_1, \mathbf{A}_2\mathbf{x}_2), \tag{2}$$

where the combination of images in the wavelet domain is due to some rule *f*. Because $\mathbf{A}_1$ and $\mathbf{A}_2$ are often square matrices, if their inverses exists; they are unique. Therefore, the transforms operating on the two input signals must be the same if the reconstruction errors $\mathbf{R}\mathbf{A}_1$, and $\mathbf{R}\mathbf{A}_2 = \mathbf{I}$ where **R** represents the inverse transform.

Our goal is to choose $\mathbf{A}_1$, $\mathbf{A}_2$, and **R** to optimizing some metric while keeping the reconstruction errors small enough so they don't adversely affect the result. Our solution to the fusion problem uses the wavelet transforms. Wavelet transforms offer a convenient framework for image reconstruction in this case, and therefore is well-suited to signal fusion. The wavelet transform typically compacts energy efficiently, and may use a variety of basis functions. The wavelet transform can be represented as a matrix as

$$\mathbf{W} = \begin{bmatrix} \mathbf{L} \\ \mathbf{B} \end{bmatrix} \tag{4}$$

where **W** is equivalent to $\mathbf{A}_1$, or $\mathbf{A}_2$, and the sub matrices **L** and **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 (inverse) transform so that,

$$\mathbf{A}_1 = \begin{bmatrix} \mathbf{L}_1 \\ \mathbf{B}_1 \end{bmatrix}, \ \mathbf{A}_2 = \begin{bmatrix} \mathbf{L}_2 \\ \mathbf{B}_2 \end{bmatrix}, \text{ and } \mathbf{R} = \begin{bmatrix} \mathbf{L}_3 & \mathbf{B}_3 \end{bmatrix}. \tag{5}$$

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 low-pass, and high-pass filters of the forward transform are, $LF(z)$ and $HF(z)$, and the low-pass, and high-pass filters of the inverse transform are, $LI(z)$ and $HI(z)$. A block diagram of two filter banks indicating the two forward transforms, where each is equivalent to one level of a wavelet transform is shown Fig. 1, where the subscripts indicate whether the filters belong to $\mathbf{A}_1$ or $\mathbf{A}_2$

If we consider only one level of the transform, then the **L**'s and **B**'s in Eq. 5 can be viewed as low- and high-pass filters, and written in the z-domain as,

$$\begin{array}{lll} \mathbf{L}_1 \rightarrow LF_1(z), & \mathbf{L}_2 \rightarrow LF_2(z), & \mathbf{L}_3 \rightarrow LI_3(z) = LI_1(z) = LI_2(z), \\ \mathbf{B}_1 \rightarrow HF_1(z), & \mathbf{B}_2 \rightarrow HF_2(z), & \mathbf{B}_3 \rightarrow HI_3(z) = LI_1(z) = LI_2(z). \end{array} \tag{6}$$

We wanted the wavelets in the forward transforms to be different so they could have different characteristics to better represent signals from disparate sensors. Therefore, we allowed $HF_1(z) \neq HF_2(z)$, and set $LF_1(z) = LF_2(z)$ to relate the two transforms.

In wavelet transforms, the high-pass filter of the synthesis transform is related to the low-pass filter of the forward transform. Therefore we set the high-pass filter of **R** to

$$HF_3(z) = -LF_1(-z) = -LF_2(-z). \tag{7}$$

If we choose $HF_1(z)$ and $HF_2(z)$ based on the characteristics of the sensors and the type of features we are considering, then the problem reduces to finding a suitable $LI_3(z)$.

3. EXPERIMENT

We primarily considered measures of entropy to evaluate the performance of different approaches because increased entropy has often been associated with better visibility. The entropy of an image can be described as

$$H = -\sum_{i=1}^{G} d(i) \log_2 d(i), \tag{8}$$

where G is the number of gray levels in the image, and $d(i)$ is the normalized histogram.[3] To measure the "distance" between the entropies of an input image and the fused result, we considered the relative entropy,

$$D(d \parallel z_1) = \sum_{i=1}^{G} d(i) \log_2 \frac{d(i)}{z_1(i)}. \tag{9}$$

Because the relative entropy is not symmetric we used the average relative entropy

$$\text{ARE} = (D(d \parallel z_1) + D(z_1 \parallel d)) / 2. \tag{10}$$

In our experiments we considered different biorthogonal transforms for $\mathbf{A}_1$, $\mathbf{A}_2$, and **R,** with different filters for each transform. 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 of the transformed data.

We compared the entropy and ARE for five pairs of infrared/visible images. We used spline biorthogonal wavelets with one vanishing moment in the high-pass filter for visible images and three vanishing moments for infrared images for the forward transforms, and used 5, 7, and 9 zeros in the low-pass filter of the reconstruction transform.

We initially kept the filters constant for the forward transform and varied the zeros of the low-pass filter of the reconstruction transform. We showed the entropy of the five different images as a function of zeros of the reconstruction transform in Fig. 2(a). In all cases the entropy gradually increased as the number of zeros of **R** increased. Although the entropy increased, the error due to the reconstruction transform also increased. Therefore, we examined the ratio of the increase in entropy from the conventional case to the reconstruction error due to the transforms. The conventional case consisted of both forward transforms and reconstruction transform being equal with all filters having three vanishing moments. The results are shown in Fig. 2(b). In all cases the entropy increased over the conventional case. The results showed that although the entropy increased, the ratio of the increase in entropy to the reconstruction error decreased.

We showed the ARE of the five fused images to the visible input images as a function of zeros of the reconstruction transform in Fig. 3(a). In all cases the ARE gradually increased as the number of zeros of **R** increased. Although the ARE increased, the error due to the reconstruction transform also increased. We also examined the ratio of the increase in the ARE from the conventional case to the reconstruction error due to the transforms. The results are shown in Fig. 3(b). In all cases the ARE increased over the conventional case. The results showed that although the ARE increased, the ratio of the increase in entropy to the reconstruction error decreased. Similar results for the ARE between the fused results and the infrared input images are shown in Fig. 4, and for the sum of the AREs in Fig. 5.

4 CONCLUSION

We found that using an unconventional approach for image fusion gave improved results in terms of entropy and average relative entropy when compared to a conventional method. Both the entropy and average relative entropy increased as the number of zeros increased in the reconstruction low-pass filters. However, the error also increased. More investigation in needed to determine the optimum operating point.

REFERENCES

1. G. Piella and H. Heijmans, "A quality metric for image fusion," *in Proc. 2003 Int. Conf. on Image Processing*, Barcelona, Spain (2003).
2. Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.* 9(3), 81-84 (2002).
3. H. Chen and P. K. Varshney, "A perceptual quality metric for image fusion based on regional information," i*n SPIE Proc. 5813, Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications 2005,* 34-45 (2005).

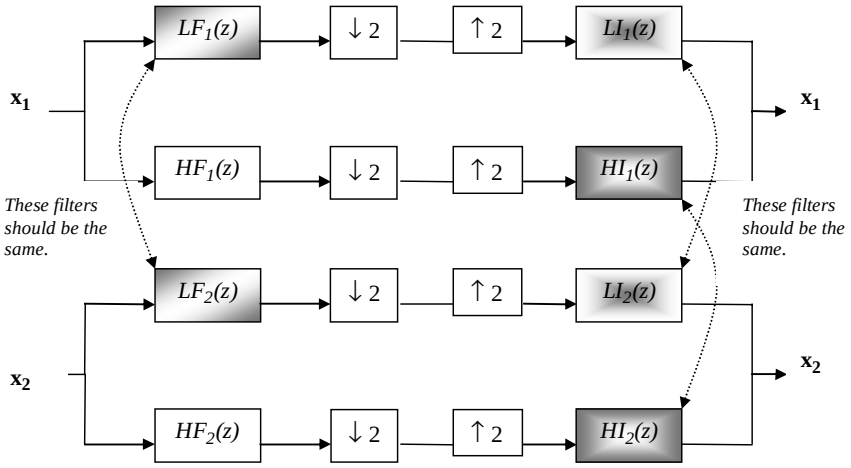


Figure 1 Two filter banks and their inverses each representing one level of the wavelet transform.

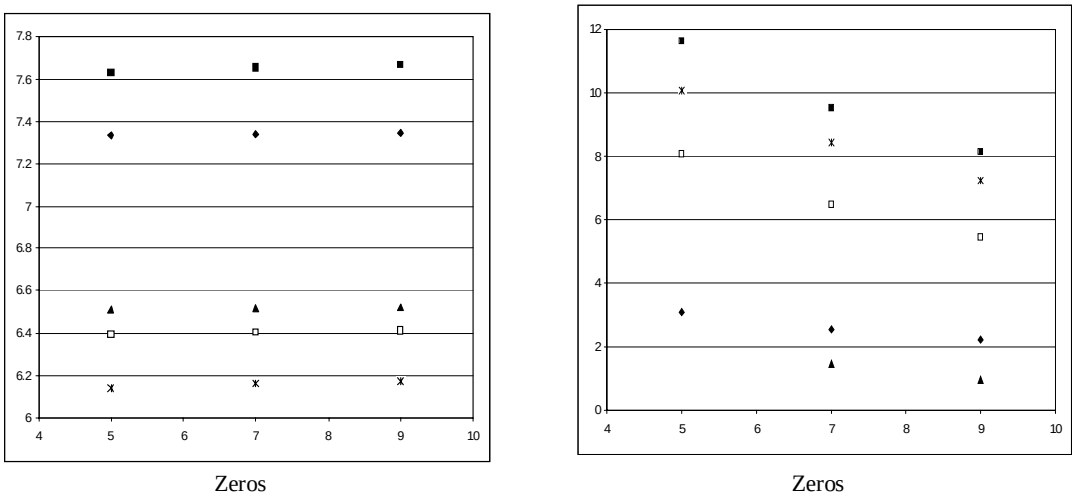
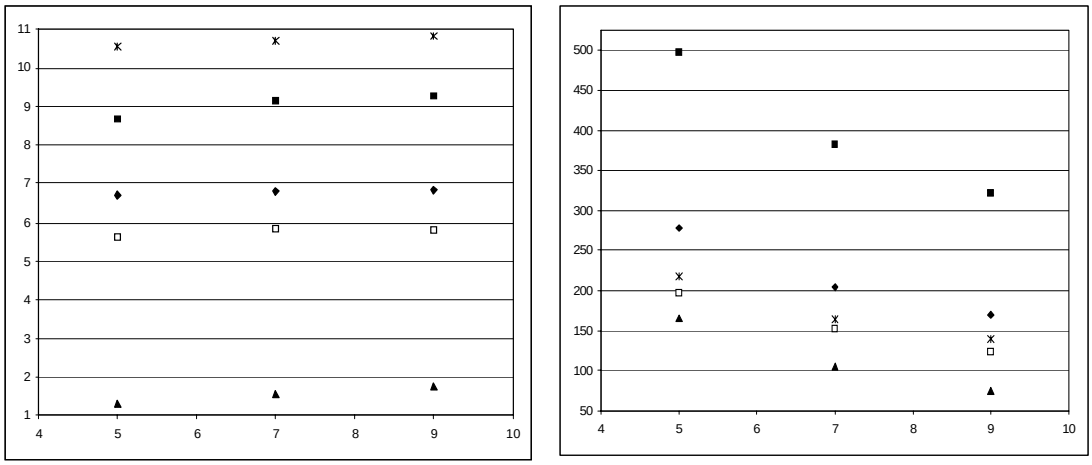
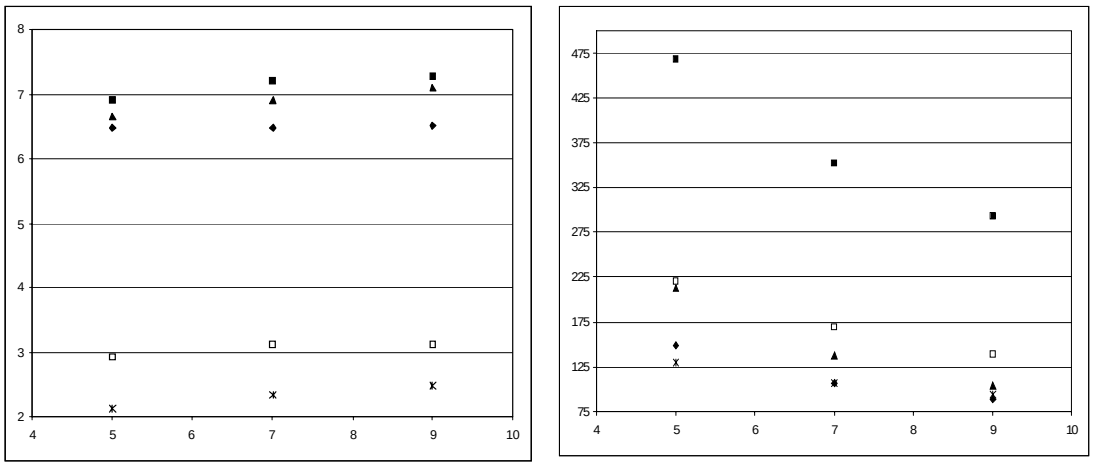


Figure 2 Entropy of fused result as a function of zeros of low-pass reconstruction filter (a) entropy (b) ratio of increase in entropy over conventional case to error in reconstruction transform.



Zeros Zeros

Figure 3 Average relative entropy of fused result to visible image as a function of zeros of low-pass reconstruction filter (a) average relative entropy (b) ratio of increase in average relative entropy over conventional case to error in reconstruction transform.



Zeros Zeros

Figure 4 Average relative entropy of fused result to infrared image as a function of zeros of low-pass reconstruction filter (a) average relative entropy (b) ratio of increase in average relative entropy over conventional case to error in reconstruction transform.

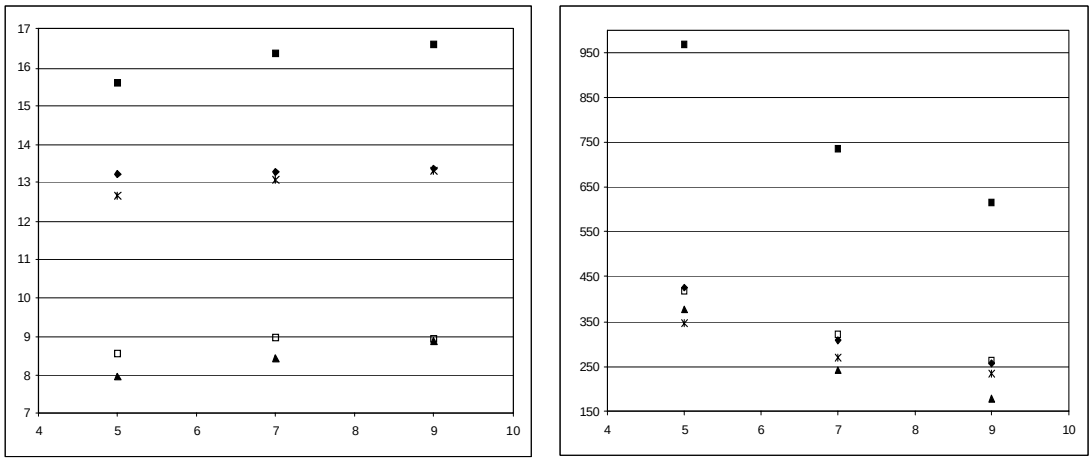


Figure 5 Sum of average relative entropy of fused result to both visible and infrared images as a function of zeros of low-pass reconstruction filter (a) total average relative entropy (b) ratio of increase in average relative entropy over conventional case to error in reconstruction transform.