5-31-2005

Multilevel fusion using enhanced feature detection

Trond M. Ostrem

Samuel Peter Kozaitis

Multilevel Fusion Using Enhanced Feature Detection

T. Ostrem, S. P. Kozaitis, and I. Laszlo*
Florida Institute of Technology
Department of Electrical and Computer Engineering
150 W. University Blvd.
Melbourne, FL 32901

*Eötvös Loránd University
Faculty of Informatics
Budapest, Hungary

ABSTRACT

We combined images from different sensors based on the magnitude of multiscale products of the wavelet transform. Using the product of nondownsampled wavelet coefficients across scales, we formed a fusion rule between two images. In this way, the products represented features than were related to contrast. Therefore, using our approach allowed us to enhance the contrast of two images using a pixel-level fusion approach. In experiments, our approach compared favorably to using only a single level, or other methods using subjective tests.

Keywords: multiresolution image processing, sensor fusion, wavelet transform

1. INTRODUCTION

Current methods for pixel-based image fusion typically attempt to combine input images into a single image without loss of information.[1,2] Often, the energy in a particular region is considered as the information to be retained. A pixel-based fusion method is needed that preserves important features of an image. Such an approach could form the basis of a successful approach to multisensor visualization.

Traditional methods to extract features often involve the use of contrast, usually through gradient of zero-crossings. In one approach, a theory for multispectral contrast has been developed that allows a multispectral image to be viewed as grayscale image.[3] In this approach, the problem of grayscale visualization was interpreted as 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. This formalism seeks the best grayscale image rather than a projections onto a one-dimensional grayscale axis. Therefore, the maximum contrast at each point can be determined.

In an effort to reduce the effect of distortion, fusion was performed on a multiresolution gradient map representation of images.[4] At each resolution, input images were represented as gradient maps and combined to produce new fused gradient maps. The fused gradient maps are processed using gradient filters derived from the difference between adjacent pixels as the Euclidian distance of the contrast components in the horizontal and vertical directions, and wavelet filters. The fused output image is obtained by applying a reconstruction process that is analogous to the DWT. To get a high contrast fused image, the fusion rules favor high absolute contrast values, so

Visual Information Processing XIV, edited by Zia-ur Rahman, Robert A. Schowengerdt,
Stephen E. Reichenbach, Proc. of SPIE Vol. 5817 (SPIE, Bellingham, WA, 2005)
0277-786X/05/$15 · doi: 10.1117/12.603932

the fusion rule simply followed the "choose max" approach. In this approach, the filters were optimized for edges defined between two adjacent pixels.

Another multiresolution approach used an orthogonal representation of images based on the wavelet transfrom.[5] Although not developed specifically for the fusion of different sensor images, 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 if the contrast is considered to be equivalent to the first derivative of the image.

This paper describes a method for pixel-level fusion based on the consistency of features through different scales of the wavelet transform. Using wavelets that respond to the first derivative allow the method to improve the estimate of the local contrast; therefore, improving the overall process. A variety of wavelet filters and scales can be used so the edges detected are not limited in scope. To examine the consistency of scales of an image, we used a nondownsampled version of the wavelet transform so pixels could be compared directly. In addition, we used an orthogonal representation of input images so they could be readily combined from two different sensors. In the next section, we briefly describe how we combined different wavelet scales, and then described some experimental results.

.

2. FEATURE DETECTION USING MULTISCALE PRODUCTS

Using a separable wavelet transform, an image *I(x,y)*, may be decomposed into lower resolution approximation images at j different scales $A_{2^j} I$, and three detail images $D^1_{2^j} I$, $D^2_{2^j} I$, $D^3_{2^j} I$, at each scale where $-1 \leq j \leq -J$,

$$A_{2^j} I = I(x,y) \otimes \phi_{2^j}(-x)\phi_{2^j}(-y), \tag{1}$$

$$D^1_{2^j} I = I(x,y) \otimes \phi_{2^j}(-x)\psi_{2^j}(-y), \tag{2}$$

$$D^2_{2^j} I = I(x,y) \otimes \psi_{2^j}(-x)\phi_{2^j}(-y), \tag{3}$$

$$D^3_{2^j} I = I(x,y) \otimes \psi_{2^j}(-x)\psi_{2^j}(-y). \tag{4}$$

The product of noisy wavelet coefficients between different scales has been shown to detect intensity discontinuities better than traditional methods.[6] Therefore, considering the discontinuities as local variations in contrast, we used an orthogonal wavelet that detects changes in the first derivative. For example, using the first n scales of the wavelet transform (j = -1, -2 . . . -n), a product is formed between the image's nondownsampled wavelet transform coefficients of the n scales,

$$C^k I = \prod_{j=-1}^{-n} D_{2^j}^k I \,. \tag{5}$$

This is done for each input image independently. Then, the two input images are compared and a reconstruction map is formed. In our examples, we used the maximum of corresponding coefficients of the two images to determine the reconstruction map according to,

$$M_1^k(x,y) = \begin{cases} C_1^k I_1(x,y) & if \ |C_1^k I_1(x,y)| > |C_1^k I_2(x,y)| \\ C_1^k I_2(x,y) & if \ |C_1^k I_1(x,y)| < |C_1^k I_2(x,y)| \end{cases}, \tag{6}$$

where (x, y) indicate the pixel locations in an image. The same rule is used to determine the reconstruction map for all subbands. Therefore, a complete orthogonal representation of an image was formed, and the inverse wavelet transform was calculated to produce a final image.

3. EXPERIMENTS

We used the two images in Fig. 1 of a person's face taken with visible and far-infrared image sensors in our experiments. Figure 2 shows the nondownsampled wavelet transform of the far-infrared image at six different scales. In all cases, the approximation image at a particular scale and the detail images are shown. The result of fusion using multiscale products is shown in Fig. 3. It can be seen that visually, the images retained much of their important information, but the images are quite different. For example, the nose in the image obtained using six levels is darker than the others. In addition, the eyes are not as visible in the result with three levels.

The fusion results in Fig. 4 show a visual comparison to other methods. The result using the product method seems to retain most of the important information. The orthogonal representation also seems to retain much information; however, the lamp is dimmer. Generally, these two images are visually more appealing than when using the other two methods.

4. CONCLUSION

We combined images from different sensors based on the magnitude of multiscale products of the wavelet transform. Using different scales gave different results. It would be useful to related performance metrics with different parameters of our approach. However, our approach seemed to compare favorably to using only a single level, or other methods using subjective tests.

REFERENCES

[1] R. C. Luo, C-C, Yih, and K. L. Su., “Multisensor fusion and integration: approaches, applications, and future research directions”, IEEE Sensors Journal, **2**, 107-119,2002.

[2] Z. Zhang, and R. S. Blum, “A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application, Proc. IEEE, **87**, 1315-1326, 1999.

[3] D. A. Socolinsky, and L. B. Wolff, “Multi-spectral Image Visualization through First Order Fusion,” IEEE Transactions on Image Processing, **11**, 923-931, 2002.

[4] V. S. Petrovic, and C. S. Xydeas., “Gradient-Based Multi-resolution Image Fusion”, IEEE Transactions on Image Processing, **13**, 228-237, 2004.

[5] P. Scheunders, “An orthogonal wavelet representation of multivalued images”, IEEE Transactions on Image Processing, , **12**, 718-725, 2003.

[6] Y. Lee, and S. P. Kozaitis, "Multiresolution gradient-based edge detection in noisy images using wavelet domain filters,” *Optical Engineering*, 39, 2405-2412, 2000.

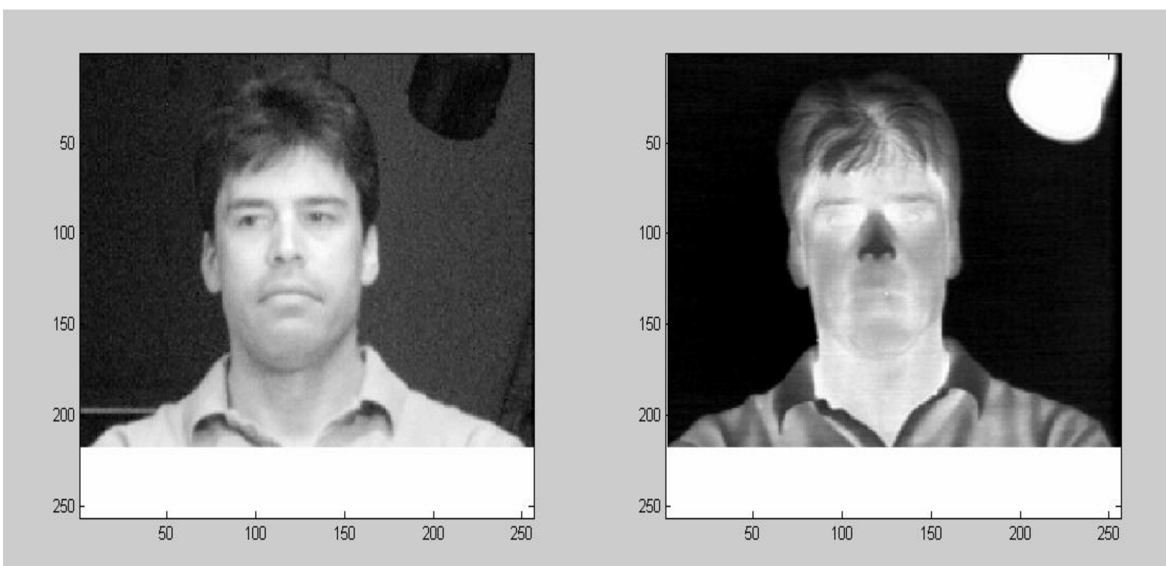

Figure 1 Visible and infrared images used in experiments.

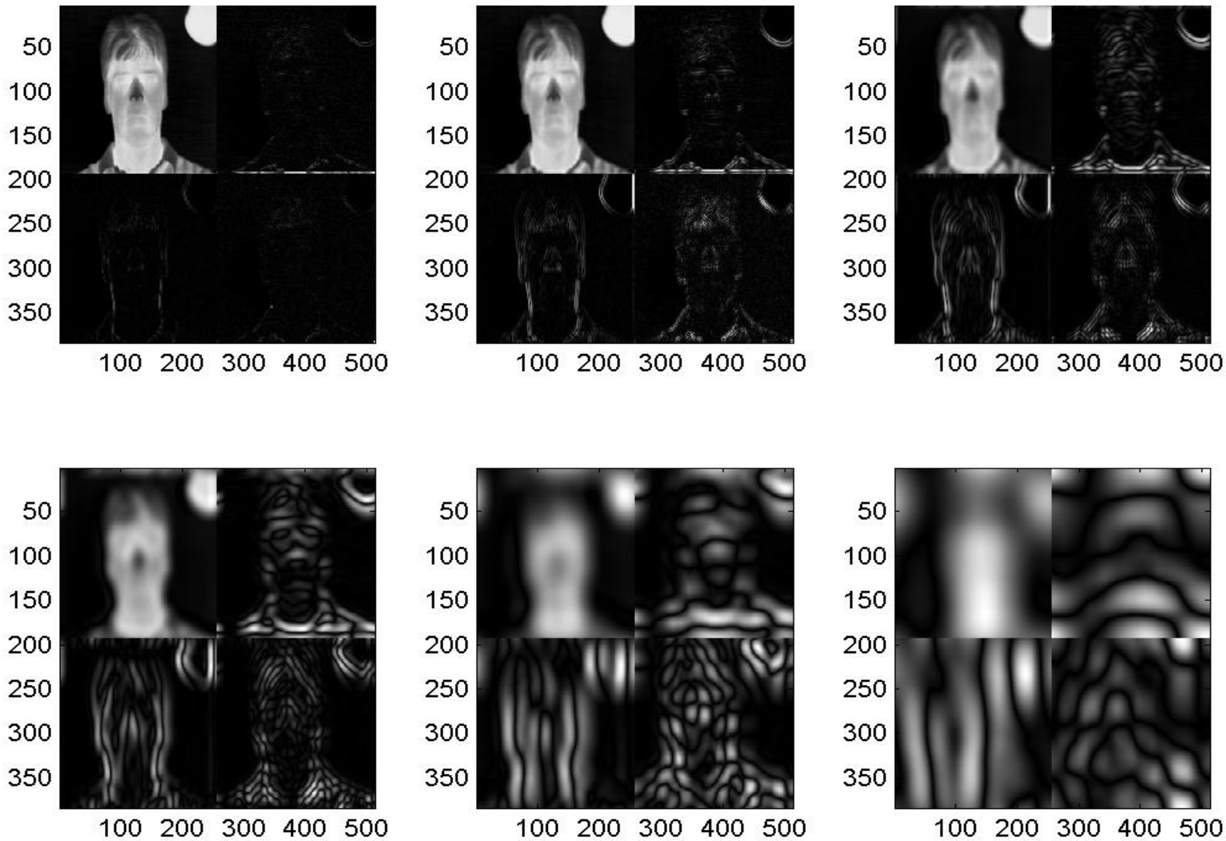

Figure 2 Nondownsampled wavelet transform of an image in Fig. 1 at different scales.

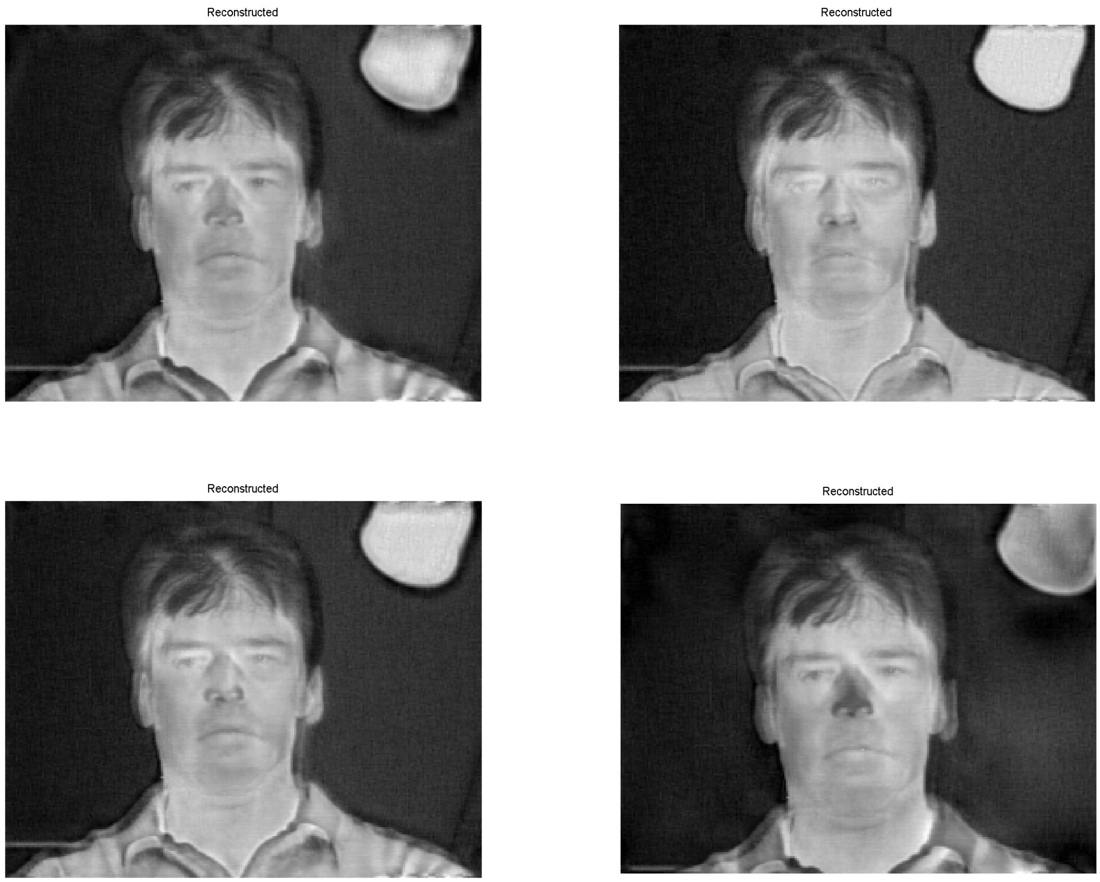

Figure 3 Fused image using different levels (a) 2 upper left (b) 3 upper right (c) 4 lower left (d) 6 lower right

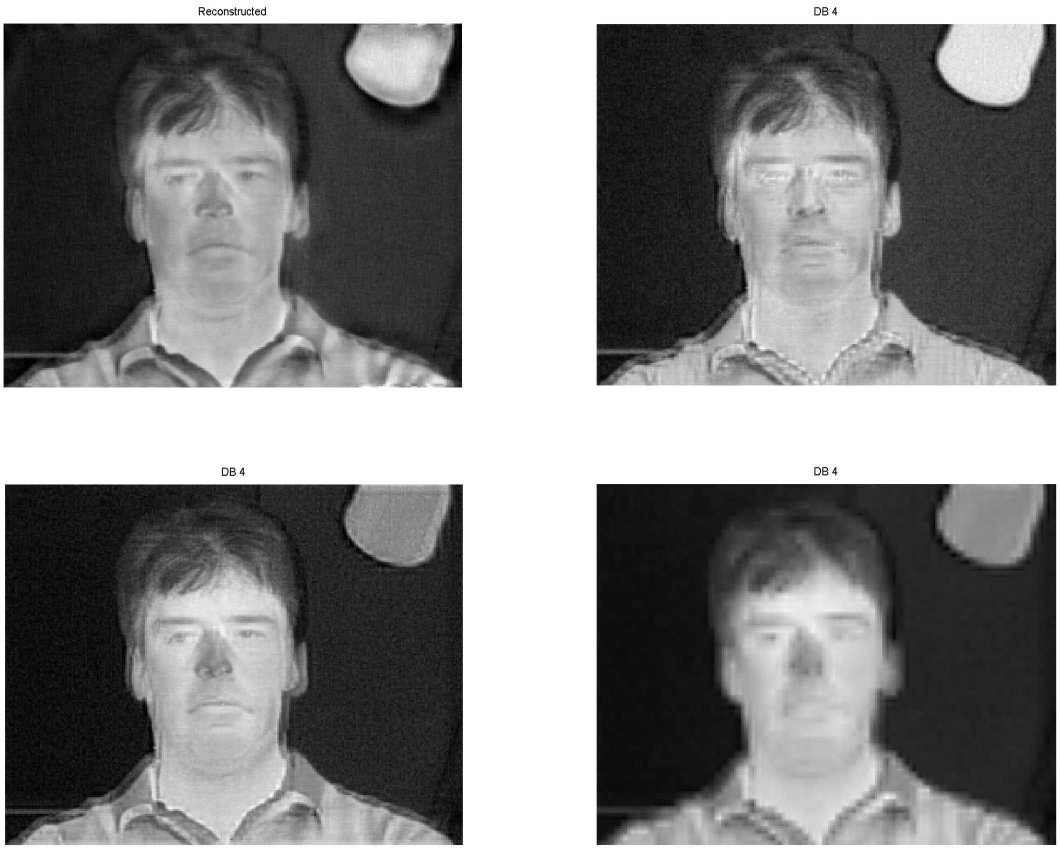

Figure 4 Fused image using different approaches (a) product method using two levels (b) “choose max” same as product method using one level (c) orthogonal representation approach (d) gradient map approach.-