Region Based Multiresolution Image Fusion for Visibility Improvement

Li Tao and Vijayan K. Asari; Computational Intelligence and Machine Vision Laboratory, Department of Electrical and Computer Engineering, Old Dominion University; Norfolk, VA 23529, USA

Abstract

This paper presents a new region based multiresolution image fusion technique for visibility improvement of digital images. In the proposed technique, pyramid image segmentation is first employed to achieve the multiresolution segmentation of the match measure image which is computed in spatial domain, and the obtained segmentation map is used to segment the source images. Then the source images are transformed using dual-tree complex wavelet transform for extracting and combining the salient image features. The activity measure of each region is determined by the magnitude of wavelet coefficients. In the third step, the wavelet coefficients of source images are combined in a region-based manner through a double-thresholding scheme governed by the match and activity measures of the corresponding regions of source images. Finally, the fused image is reconstructed via the inverse wavelet transform of the composite wavelet coefficients obtained in the previous step.

Introduction

Multiresolution (MR) image fusion is a useful image processing technique which is able to provide a more complete description of the scene in a single composite image compared to each of the source images. Therefore, it has been investigated and applied in various technical fields, such as medical image system [1], analysis of remote sensed images [2], and targeting system for defense application [3].

MR analysis techniques have been extensively used in image fusion due to its strong capability of representing multiscale image features. Various image pyramid techniques [4-6] and wavelet transform methods [7-9] have been studied for MR image fusion. Recently, dual-tree complex wavelet transform (DT-CWT) [10] shows some advantages over other MR image analysis techniques in terms of image fusion.

Except the MR image analysis, various image fusion schemes at different image feature scale levels (e.g. pixel level and region level) have also been proposed. Compared to pixel level image fusion, region level image fusion is more capable of maintaining the integrity of an image feature. In addition, region based method shows more flexibility in processing different image features in different ways (e.g. emphasizing the objects with respect to the background).

In this paper, a new region level image fusion scheme based on DT-CWT is proposed. Good image fusion results have been obtained with simplified image processing.

Algorithms

The framework of the proposed region-based fusion algorithm is presented in Fig. 1. The major contribution of the proposed algorithm is that pyramid image segmentation is first employed to achieve the MR segmentation of the multilevel match measure images which are computed at pixel level on source images in spatial domain. Then the MR segmentation of source images is conducted using the same segmentation maps obtained from the segmentation of match measure images. In the second step, MR decomposition of source images is performed using DT-CWT. In the third step, the wavelets coefficients of source images are combined in a region-based manner through a double-thresholding scheme governed by the match and activity measures of the corresponding regions. Finally, the fused image is reconstructed via the inverse wavelet transformation of the composite wavelet coefficients obtained in the last step. Each module is described and explained in the remaining part of this section.

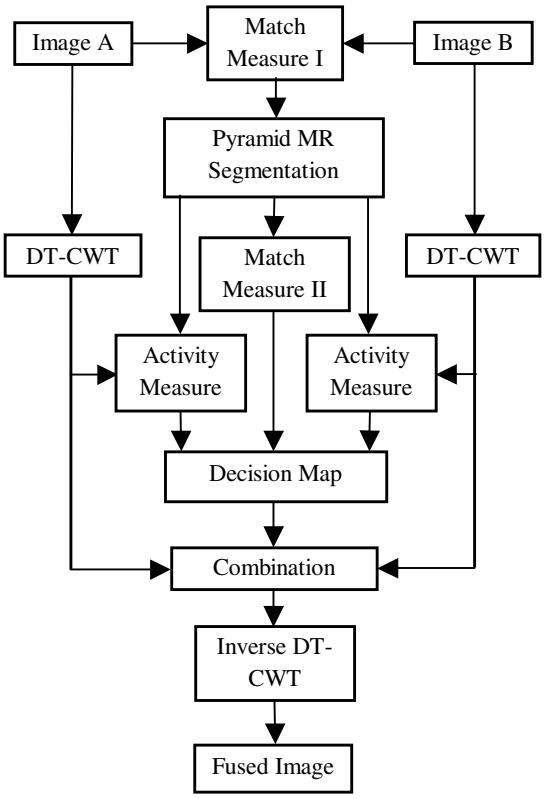

Figure 1. *Region-based MR fusion scheme*

Match Measure I: This module computes the match measure at pixel level. The pixel level match measure result is then used as the input for MR segmentation to classify different regions on the source images in a joint way.

The pixel level match measure computes the similarity between corresponding pixels in both source images. Match

measure is critical in image fusion algorithms because it determines where the source images resemble (or differ) and to which extent so that source images can be combined in an appropriate way. In order to properly compare the two corresponding pixels, neighborhoods surrounding the pixels should also be considered. In our image fusion scheme, pixel level match measure is defined as a normalized correlation averaged over a neighborhood of the samples as in:

$$m_j^{AB}(x,y) = \frac{2\sum_{(x+m,y+n)\in w} I_j^A(x+m,y+n) I_j^B(x+m,y+n)}{\sum_{(x+m,y+n)\in w} (I_j^A(x+m,y+n))^2 + (I_j^B(x+m,y+n))^2} \quad (1)$$

where I_j^A and I_j^B are successively subsampled source images at level j, and w is the 5×5 neighborhood. The subsampled images should have the same matrix sizes as those of wavelet detail coefficients (explained later). The value of m^{AB} is an estimation of the similarity of image features at pixel level, e.g., $m^{AB} = 1$ indicates identical patterns, $m^{AB} < 1$ shows less similarity between features, $m^{AB} = 0$ indicates that the grayscale values of all the pixels in the two neighborhoods are zero.

We propose to compute match measure in the spatial domain instead of the transform domain for more accurate estimation of pattern similarity and less computational cost. The typical approach to calculating match measure is to apply Eq. (1) to all subbands of transformed source images in the wavelet domain. For instance, there are six subbands at each decomposition level produced by DT-CWT, thus Eq. (1) will be used six times at each level in the typical method while the proposed approach only computes Eq. (1) once at each level.

Pyramid MR Segmentation: this module performs standard pyramid multi-resolution segmentation on pixel level match measure images instead of the source images themselves. In region based image fusion, MR segmentation is generally conducted on source images separately or jointly in either spatial domain or transform domain. Compared to those commonly used techniques, the proposed MR segmentation on match measure images has some advantages. First, it is more efficient for it does not segment both source images separately. Match measure images contain the information from both sources. Second, it serves the goal of image fusion better because the fusion of two pixels or two regions is guided by match measure and activity measure. The segmented regions based on match measure have similar match measure values at all pixel locations within a region. Therefore, the pixels in a region can be treated in the same way so that the creation of image artifacts can be controlled. However, the current segmentation methods do not consider match measure when segmenting source images, which may create a region that has different match measures at different pixel locations within the region. Thus image artifacts can be created when two regions are combined with all the pixels in a region are governed by the same rule.

DT-CWT: This module performs MR decomposition of source images using dual-tree complex wavelet transform. The detail coefficients will be used in the next step to compute the activity measure. More importantly, the actual image fusion occurs in wavelet transform domain, which has been indicated in Fig. 1. To obtain optimized image fusion result, the source images are generally decomposed to level 5 in our experiments. If the source image size does not allow the MR decomposition up to 5 levels (i.e. the width or height of the source image is not equal to $a \cdot 2^5$ where a is a positive integer.), the source image will be resampled to the closest size which allows such MR decomposition.

Match measure II: this module computes the similarity between the corresponding regions on both sources. The computation is based on the MR segmentation of source images and the pixel level match measure results obtained in module M*atch Measure I*. In our fusion scheme, the match measure between corresponding regions at level j, $m_{j,R}^{AB}$, is defined as in:

$$m_{j,R}^{AB} = \frac{1}{S_{j,R}} \sum_{z\in R} m_j^{AB}(z) \quad (2)$$

where vector $z = (x, y)$ is the pixel coordinates and $S_{j,R}$ is the area of region R at level j, which is the number of pixels within the region. In fact, the similarity between corresponding regions defined in Eq. (5.4) is an averaged similarity at pixel level.

Activity measure: This module computes the activity measures of regions at all levels in both wavelet transformed source images. This region level activity measure is based on the pixel level activity measure at p and is defined in a similar way to Eq. (2):

$$E_{j,R}^A = \frac{1}{S_{j,R}} \sum_{z\in R} E_j^A(z), \quad E_{j,R}^B = \frac{1}{S_{j,R}} \sum_{z\in R} E_j^B(z) \quad (3)$$

where $E_{j,R}^A$ and $E_{j,R}^B$ are the activity measure of region R at level j in both source images; $E_j^A(z)$ and $E_j^B(z)$ are pixel level activity measures at level j in both source images, which are defined Eq. (4). Thus the activity measure of a region can be considered as the arithmetic average of activity measures of all pixels within the region.

Pixel level activity measure computes the 'saliency' of each pixel in the transform domain. The meaning of saliency depends on the properties of source images and the objective of particular fusion application. Based on the fact that the human vision system (HVS) is primarily sensitive to local contrast changes (e.g. edges), most fusion algorithms compute the activity measure as some sort of energy calculation. In our fusion scheme, the magnitude of the detail coefficients is used to calculate the activity measure as in:

$$E_j^A = \sum_{k=1}^{6} \left|A_j^k\right|, \quad E_j^B = \sum_{k=1}^{6} \left|B_j^k\right| \quad (4)$$

where A_j^k and B_j^k are the complex detail coefficients of the two source images at level j and orientation k; E_j^A and E_j^B are activity measures at level j for both source images, which are the summation of the magnitudes of the detail coefficients of all 6 subbands at each decomposition level. Eq. (4) also has its physical meaning which can be understood by considering that each coefficient of an MR decomposition has a set of related components in other orientation bands and other levels. They represent the image feature(s) at the same (or nearby) spatial location in the original image. Therefore, it is reasonable to take into account the coefficients in all subbands when the image property at one spatial location is being determined.

The pixel level activity measures obtained in Eq. (4) may need to be low-pass filtered to suppress the salient features caused by the impulsive noises in source images. In our fusion algorithm, a spatial convolution with a 3×3 Gaussian mask with a standard deviation of 0.5 is used for this purpose.

Combination: this module combines MR decomposition coefficients of both sources at region level to obtain the composite MR decomposition coefficients of the fused image. The combination of the detail coefficients is performed in a linear way as in:

$$F_j^k(z) = C_{j,R}^A \cdot A_j^k(z) + C_{j,R}^B \cdot B_j^k(z), \quad z \in R \tag{5}$$

where $A_j^k(z)$ and $B_j^k(z)$ are detail coefficients at pixel location z in region R at level j and orientation k while $F_{j,R}^k(z)$ is the composite coefficient at pixel location z in region R at level j and orientation k. $C_{j,R}^A$ and $C_{j,R}^B$ are weight factors for region R at level j in both sources. The Eq. (5) indicates that all pixels within a region will be assigned the same weight factor which is given in the decision map.

The approximation coefficients of the fused image can be obtained directly from those of either source image, or by a linear combination of those of both source images, or by the maximum selection rule (choosing the larger coefficient at each pixel location between both sources). However, to use which method is determined by the properties of both source images and the visual quality of fused images produced by those methods.

Decision map: this module determines the values of the weight factors in Eq. (5) based on the match and activity measures of regions. This is the key point in image fusion. In our image fusion scheme, we propose a double thresholding method to determine the weight factors in Eq. (5) in the following three cases:

Case 1

$$\begin{cases} C_{j,R}^A = \dfrac{E_{j,R}^A}{E_{j,R}^A + E_{j,R}^B}, \\ C_{j,R}^B = \dfrac{E_{j,R}^B}{E_{j,R}^A + E_{j,R}^B}, \end{cases} \quad if \quad m_{j,R}^{AB} \geq 0.9$$

Case 2

$$\begin{cases} C_{j,R}^A = \dfrac{E_{j,R}^A}{E_{j,R}^A + E_{j,R}^B} \cdot T + (1-T) \cdot W_A, \\ C_{j,R}^B = \dfrac{E_{j,R}^B}{E_{j,R}^A + E_{j,R}^B} \cdot T + (1-T) \cdot W_B, \end{cases} \quad if \quad 0.7 < m_{j,R}^{AB} < 0.9$$

$$where \quad T = \frac{m_{j,R}^{AB} - 0.7}{0.9 - 0.7}, \quad and \quad \begin{cases} W_A = 1 & if \quad E_{j,R}^A \geq E_{j,R}^B, \quad otherwise \quad 0 \\ W_B = 1 & if \quad E_{j,R}^A < E_{j,R}^B, \quad otherwise \quad 0 \end{cases}$$

Case 3

$$\begin{cases} C_{j,R}^A = 1 \quad C_{j,R}^B = 0, & if \quad m_{j,R}^{AB} \leq 0.7 \quad and \quad E_{j,R}^A \geq E_{j,R}^B \\ C_{j,R}^A = 0 \quad C_{j,R}^B = 1, & if \quad m_{j,R}^{AB} \leq 0.7 \quad and \quad E_{j,R}^A < E_{j,R}^B \end{cases}$$

In case 1, the match measure is high which represents highly similar patterns, and the fused coefficient is the linear combination of the coefficients of both input images with the weight factors determined by the relative relation between the two activity measures. On the contrary, in case 3, the match measure is low, which represents very dissimilar patterns, and thus only the more salient feature (larger activity measure) is included in the fused image. Between these two extreme cases, case 2, which represents medium similarity, provides a smooth transition with the weight factors set by a linear combination of the two extreme cases. The coefficient T in case 2 modulates the relative importance between those two terms based on the relative position of match measure with respect to the two thresholds. If the match measure is close to 0.9, the weight factors are determined in a way more similar to case 1. Otherwise, if the match measure is close to 0.7, the weight factors are determined in a way more similar to case 3. Finally, the expressions discussed above should ensure that the sum of the two weight factors C_j^A and C_j^B be 1.

Results

A few examples of the source images used in the image fusion experiments are presented in Fig. 1. The left and right images in the upper row are captured by CCD camera and thermal cameras, respectively (for multisensor image fusion), while the left and right image in the bottom row are recorded in daytime and at night, respectively (for multi-illumination image fusion).

Figure 1. *Two sets of Source images (upper row and lower row) used for image fusion experiments*

The image processing and image fusion results of the source images shown in the upper row in Fig. 1 are provided in Fig. 2. The left image in the upper row shows the match measure result of subsampled source images at level 1. The grayscale value at each pixel location represents the similarity between the neighborhoods which are surrounding the corresponding pixels on both subsampled source images at level 1. Higher grayscale value indicates higher degree of resemblance. The right upper row image is the MR segmented result of the match measure image. The images in the middle row demonstrate the pixel level activity measures of the subsampled source images at level 1. Higher grayscale indicates higher saliency which means a significant intensity change in source images. The left image in the bottom row is the decision map for the level 1 sampled source image which is captured by CCD camera. The inverted version of it is

actually the decision map for the level 1 subsampled thermal image. The final image fusion result is presented as the right image in the bottom row.

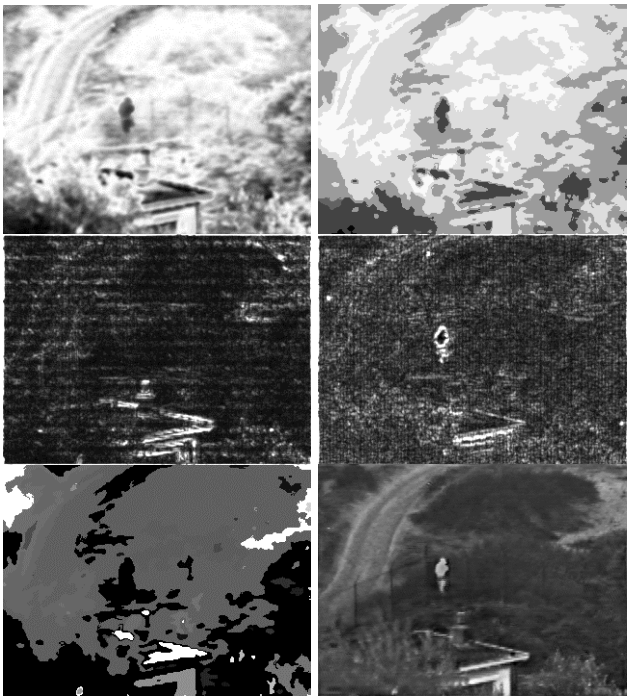

Figure 2. *Image fusion of multisensor images (see the explanation in text).*

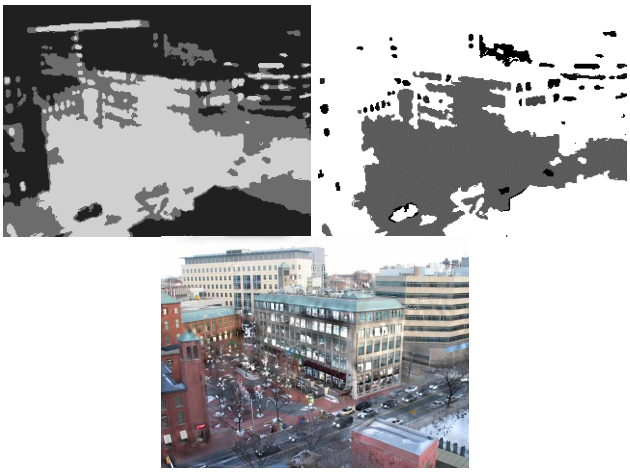

Figure 3. *Image fusion of multi-illumination color images (see the explanation in text).*

The results of the image fusion of the source images shown in the bottom row in Fig. 1 are provided in Fig. 3. Both source images are converted to grayscale images prior to image fusion process. The color information from the daytime image is transferred to the fused grayscale image to restore the fused color image. The left image in the upper row is the MR segmented result of the level 1 pixel level match measure image while the right image is the level 1 decision map for the daytime source image. The final fused color image is the image in the bottom row.

Both image fusion experiments demonstrate good performance and validity of the proposed image fusion algorithm which also exhibits high robustness in other image fusion experiments performed using various types of source images, such as multi-focus images and multi-exposure images.

Conclusion

A new region level MR image fusion algorithm based on DT-CWT has been developed. A new region level image fusion scheme is proposed with simplified image processing. The proposed image fusion algorithm shows good and robust image fusion results.

References

[1] G.K. Matsopoulos, S. Marshall, "Application of morphological pyramids: fusion of MR and CT phantoms," *Journal of Visual Communication and Image Representation*, vol. 6, no. 2, pp. 196–207, 1995.

[2] C. Pohl, J.L. Genderen, "Multisensor image fusion in remotesensing: concepts, methods and applications," *International Journal of Remote Sensing 19 (5)* pp. 823–854, 1998.

[3] B.V. Dasarathy, "Fuzzy evidential reasoning approach to target identity and state fusion in multisensor environments," *Optical Engineering 36 (3)*, pp. 683–699, 1997.

[4] P.J. Burt, "The pyramid as a structure for efficient computation, in: A. Rosenfeld (Ed.)," *Multiresolution Image Processing and Analysis*, Springer-Verlag, Berlin, Germany, pp. 6–35, 1984.

[5] Z. Liu, K. Tsukada, K. Hanasaki, Y.K. Ho, Y.P. Dai, "Image fusion by using steerable pyramid," *Pattern Recognition Letters*, vol. 22, pp. 929–939, 2001.

[6] A. Toet, "Image fusion by a ratio of low-pass pyramid," *Pattern Recognition 9* pp.245–253, 1989.

[7] A. Toet, "Image fusion by a ratio of low-pass pyramid," *Pattern Recognition 9* pp.245–253, 1989.

[8] H. Li, B.S. Manjunath, S.K. Mitra, "Multisensor image fusion using the wavelet transform," *Graphical Models and Image Processing*, vol. 57, no. 3, pp. 235–245, 1995.

[9] H. Li, B.S. Manjunath, S.K. Mitra, "Multisensor image fusion using the wavelet transform," *Graphical Models and Image Processing*, vol. 57, no. 3, pp. 235–245, 1995.

[10] N. G. Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," *Appl. and Comp. Harmon. Analy*," vol. 10, no. pp. 234-253, 2001.

Author Biography

Li Tao received the BS and MS degrees in Electronics and Communication Engineering from Sichuan University, Chengdu, People's Republic of China in 1998 and 2001, respecitvely. She is currently pursuing her PhD degree in the Department of Electrical and Computer Engineering at Old Dominion University, Norfolk, Virginia, USA. Her research interests include digital signal processing, digital image processing and computer vision.

Vijayan K. Asari joined the Department of Electrical and Computer Engineering, Old Dominion University, Virginia as an Associate Professor in August 2000. Dr. Asari received the B.Sc. (Engg.) degree in Electronics and Communication Engineering from the University of Kerala, M.Tech. and Ph.D. degrees in Electrical Engineering from the Indian Institute of Technology, Madras, India.. His research activities are in the areas of signal processing, image processing, neural networks, and digital system architectures for application specific integrated circuits. Dr. Asari is a senior member of the IEEE.