An Image Fusion Approach Based on Segmentation Region

Rong Wang[1], Li-Qun Gao[2], Shu Yang[3], Yu-Hua Chai[4], and Yan-Chun Liu[5]

135# Institute of Information Science & Engineering, Northeastern University,
110004 Shenyang, China
[1] dbdxwangrong@163.com, [2] liqungao2064@yahoo.com,
[3] shuyang024@163.com, [4] yhchai@163.com, [5] liuyanchun@163.com

Abstract

An Image fusion method based on segmentation region is proposed in this paper. First, the source images are decomposed by wavelet to get the approximate and detailed sub-images, and the segmentation by watershed method for these sub-images are used to get the regions of each level, these regions are used to guide fusion process. The activity level and match degree measure of the wavelet coefficients of source images are carried out in these regions, and the maximum value rule and the weighted average rule are respectively used to combine the coefficients of detailed sub-images and approximate sub-image. At last, the combination coefficients are inversely transformed by wavelet to get the final fusion image. The experimental results show that the fusion effect is better.

Key words: Image fusion; image segmentation; wavelet transform.

1 Introduction

With the fast development of the technique of sensor, micro-electronic and communications, the technique of information fusion has been paid more and more attention to by people. In the application field of image processing, image fusion has a wide foreground. The aim of image fusion is to integrate complementary and redundant information from multiple images to create a composite that contains a better description of the scene than any of the individual source images.

The actual fusion process can take place at different levels of information representation[1]. A common categorization is to distinguish between pixel, feature and symbol level. Image fusion at pixel-level means fusion at the lowest processing level referring to the merging of measured physical parameters. Fusion at feature-level requires first the extraction (e.g., by segmentation procedures) of the features can be identified by characteristics such as size, shape, contrast and texture. The fusion is thus based on those extracted features and enables the detection of useful features with higher confidence. Fusion at symbol level allows the information to be effectively combined at the highest level of abstraction. The choice of the appropriate level depends on many different factors such as data sources, application and available tools.

Currently, it seems that most image fusion applications employ pixel-based methods[2], [3]. The advantage of pixel fusion is that the images used contain the original information. Furthermore, the algorithms are rather easy to implement and time efficient. The existing fusion methods based on pixel-level are very sensitive to misregistration, thus the accuracy of image registration is demanded in sub-pixel level. In some examples of image fusion, the fusion method based on the region feature has more meaning than the one based on pixel method. The robusticity of the method based on region feature is better than the one based on pixel, which avoids the problem existing in pixel fusion method, such as the sensitivity to noise and misregistration [4], [5], [6], [7]. The region-based approach have some advantages that the fusion process becomes more robust and avoids some of the well-known problems in pixel-level fusion such as blurring effects and high sensitivity to noise and misregistration.

Image fusion methods based on wavelet transform have been widely used in recent years. The method proposed in reference 5 uses wavelet method to fuse images, but it uses a slip window region with fixed scale to compute the area power to determine which one will be considered in the fused image, so it maybe brings the problem of the inconsistent coefficients selection to the same object. In this paper, a wavelet transform image fusion method based on segmentation regions is proposed, in which each region corresponds to one object in image, and all the fusion procedures are carried out in this regions, thus the abuse of inconsistent coefficients selection to the same object maybe occurred in fixed window region is avoided. It combines effectively the pixel-level and the feature-level method to get a better fusion image. The experimental results show that the performance of our method is better than the one proposed in reference 5.

2 The Fast Algorithm of 2-D Wavelet Transform

There is a quick algorithm for the 2-D orthogonal wavelet transform ---- Malat algorithm, which converts the calculation problem of wavelet transform into the one of the coefficient after wavelet transform. Suppose $H=\{h_n\}$, $G=\{g_n\}$ are the low-pass and the high-pass filters respectively when the image is decomposed, then the process of calculating decomposition coefficients of the tensor product wavelet is,

$$
\begin{aligned}
&c^M(m,n) = f(m,n) \\
&c^{M+1}(m,n) = \sum_{k,l} h(k-2m)h(l-2n)c^M(k,l) \\
&D_1^{M+1}(m,n) = \sum_{k,l} h(k-2m)g(l-2n)c^M(k,l) \\
&D_2^{M+1}(m,n) = \sum_{k,l} g(k-2m)h(l-2n)c^M(k,l) \\
&D_3^{M+1}(m,n) = \sum_{k,l} g(k-2m)g(l-2n)c^M(k,l)
\end{aligned}
\tag{1}
$$

Where $c^M(m, n)$ is the wavelet coefficients at the M level, also the original image data; $c^{M+1}(m, n)$ is the image data at the M+1 level and the low frequency component data after the image data $c^M(m, n)$ at the M level are decomposed by wavelet, and is the general picture of $c^M(m, n)$. The contour of $c^{M+1}(m, n)$ is similar to that of $c^M(m, n)$;

while $D_1^{M+1}(m, n)$, $D_2^{M+1}(m, n)$ and $D_3^{M+1}(m, n)$ are the high-frequency detailed signals after $c^M(m, n)$ is decomposed by wavelet.

3 The Principle of Image Fusion Based on Segmentation Region

3.1 Fusion Method Based on Segmentation Region

The region-based fusion method is shown as Fig.1, there are 7 modules in the whole fusion scheme: image segmentation, wavelet transformation, match measure, activity measure, decision of selecting coefficients, coefficients combination and inverse wavelet transformation.

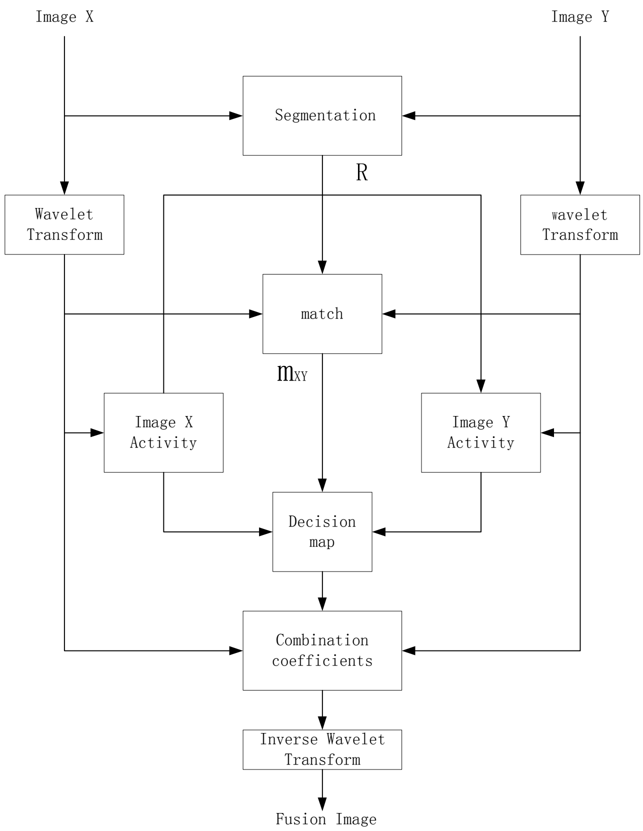

Fig. 1. The scheme of fusion method based on segmentation region

The main task of image segmentation is to obtain a set of segmentation regions R (R^1, R^2,…, R^k), R^k represents the segmentation regions at level k. This representation will guide the other blocks of fusion process: activity measure, match measure and fusion operator are all carried out in them.

In our fusion scheme, the segmentation is a preparatory step toward actual fusion. If the accuracy of segmentation is high, then it is very favorable for improving the fusion performance. In this module, a segmentation method based on watershed and wavelet is used. The concrete steps are as follows.

After creating the pyramid image using a wavelet transform, the lowest-resolution image I^k is segmented through the application of a watershed algorithm and a partition R^k from I^k is generated.

Generally, watershed algorithms produce meaningful image segmentation. However, when an image is degraded by noise, it becomes over-segmented. So it may require further merging of some regions. Our decision on which regions to merge is determined through homogeneity and similarity criteria based on the wavelet coefficients. Each of the segmented regions will have mean, second-order and third-order central moment values of the wavelet coefficients calculated. All the features are computed on the lowest-resolution sub-image. For each region R_i^k of the segmented image R^k of the image segmentation phase, we calculate the mean (M), second-order (μ_2) and third-order (μ_3). They are defined as,

$$\begin{aligned} M &= \frac{1}{num(R_i^k)} \sum\sum R_i^k(x,y) \quad \forall x, y \in R_i^k \\ \mu_2 &= \frac{1}{num(R_i^k)} \sum\sum (R_i^k(x,y) - M)^2 \\ \mu_2 &= \frac{1}{num(R_i^k)} \sum\sum (R_i^k(x,y) - M)^3 \end{aligned} \qquad (\mathbf{2})$$

Where, $num(R_i^k)$ is the number of pixels of segmented region i. To merge the segmented regions using similarity criteria (d), we can use the following equation:

$$\begin{aligned} mv_i &= \frac{1}{N}(R(M_i) + R(\mu_{2i}) + R(\mu_{3i})) \quad i = 1, \ldots, N \\ d(R_i^k, R_j^k) &= (mv_i - mv_j)^2 \quad \forall i, j \in \{1,\ldots,N, \text{ for } i \neq j\} \end{aligned} \qquad (\mathbf{3})$$

Where, mv_i is the similarity value of segmented region i and N is the number of segmented regions. $R(M)$, $R(\mu_2)$ and $R(\mu_3)$ are the mean, second-order and third-order moment values of the segmented region respectively. If the mv values of the adjacent regions satisfy a specified value, two adjacent regions will be merged. The specified values are found by the experiment.

Once the merged image M^k is generated at the image partition R^k, it must be projected down in order to reconstruct the different-resolution image. To project the segmented image with label onto the next level image, with the multi-resolution property to wavelet transformation, we can first obtain a coarse merged image with label on the low-resolution image. Then, the merged image with label can be progressively refined level by level until full resolution is reached. The projection method used to generate M^{k-1} is carried out in following steps:

1. Inverse wavelet transform is applied to M^k, so a M^{k-1} is obtained. The projected image M^{k-1} produces new values because each region of image M^k has an average pixel value.
2. Watershed line image B^{k-1} from the image I^{k-1} is generated by the watershed algorithm.
3. The boundary and labels of the image M^{k-1} are refined according to the image B^{k-1}. To refine the image M^{k-1}, we
   (a) determine the central position at each region of the image B^{k-1}.
   (b) select a label that is the great many included labels in the region of image M^{k-1} including the determined central position in the above step.
   (c) assign the selected label to the region of image Bk-1.
4. The result is the refined segment image at level (k-1) of the pyramid, identified as M^{k-1}. In this step, the image M^{k-1} has many more regions than the number of regions of image M^k. However, the adjacent regions of M^{k-1} are assigned the same label. Therefore, the number of regions of the segmented images M^{k-1} and M^k is equal.

3.2 Fusion Method

Activity Level Measure. The **a**ctivity level of wavelet coefficients reflects the local energy in the space. In general, the points with larger activity level are selected. The activity of each coefficient is defined as,

$$A_I(p) = \sum_{s\in S, t\in T} \varpi(s,t) \left| D_I(m+s, n+t, k, l) \right| \tag{4}$$

Where, $A_I(p)$ is the activity level ($I = X, Y$), $p = (m, n, k, l)$ the wavelet coefficients of image, (m, n) the special position in a given frequency, k the decomposition level, l the frequency band of decomposition, $\varpi(s, t)$ a weight, S and T are sets of horizontal and vertical indexes that describe the current window (typically 3×3 or 5×5), $| D_I(m + s, n + t, k, l) |$ the absolute value of the corresponding coefficient in a small window.

The activity level of each region $R \in R^k$ is defined as,

$$A_I^{(k)}(R \mid p) = \frac{1}{|R|} \sum_{\vec{N}\in R} A_I^{(k)}(\vec{N} \mid p) \tag{5}$$

Where, $\vec{N} = (m, n) \in R$, $R \in R^k$, $|R|$ is the area of region R, $A_I^{(k)}(\vec{N} \mid p)$ the activity level of each coefficient in region R, $A_I^{(k)}(R \mid p)$ is the activity level of region R.

Match Degree measure. The match degree is the similarity measure between the corresponding pixels in source images, it is usually expressed in terms of a local correlation measure, alternatively, the relative amplitude of the coefficients or some other criteria can be used.

A match measure function $M_{XY}(p)$ between the corresponding pixels in source images is defined as,

$$M_{XY}(p) = \frac{\sum_{s\in S, t\in T} \varpi(s,t) D_X(m+s, n+t, k, l) D_Y(m+s, n+t, k, l)}{A_X^2(p) + A_Y^2(p)} \quad \textbf{(6)}$$

Where, $\varpi(s, t)$ is a weight, S and T are sets of horizontal and vertical indexes that describe the current window (typically 3×3 or 5×5), A_X and A_Y the activities calculated by equation (4), $|D_I(m+s, n+t, k, l)|$ the absolute value of the corresponding coefficient in a small window ($I = X, Y$).

The definition of match degree measure of region is,

$$M_{XY}^{(k)}(R \mid p) = \frac{1}{|R|} \sum_{\vec{N}\in R} M_{XY}^{(k)}(\vec{N} \mid p) \quad \textbf{(7)}$$

Where, $M_{XY}^{(k)}(\vec{N} \mid p)$ is the match degree between the corresponding pixels in region R in source images, $M_{XY}^{(k)}(R \mid p)$ the match degree between the corresponding regions in source images.

The Decision of Coefficients' Selection and Combination. The decision of selecting coefficients is the key of combing coefficients, a decision map can be constructed according to the activity measure, match measure and segmentation region obtained by calculating before, which image's coefficients are selected to combine the coefficients are determined by this map. The equation of coefficient combination is as follows,

$$D_F^k(R \mid p) = \varpi_X D_X^k(R \mid p) + \varpi_Y D_Y^k(R \mid p) \quad \textbf{(8)}$$

Where, $D_F^k(R \mid p)$ is the combination coefficient in region R, $D_X^k(R \mid p)$ and $D_Y^k(R \mid p)$ respectively the coefficients of source images X and Y corresponding to R, ϖ_X and ϖ_Y weights, and $\varpi_X + \varpi_Y = 1$.

In fact, the calculation of constructing decision map is equivalent to the one of weights ϖ_X and ϖ_Y, in this paper, the calculation of weights are considered as two instances: maximum value rule and weighted average rule.

Maximum value rule is,

$$\begin{aligned} \varpi_X &= 1, \quad \varpi_Y = 0, \quad A_X^k(R \mid p) > A_Y^k(R \mid p) \\ \varpi_X &= 0, \quad \varpi_Y = 1, \quad A_X^k(R \mid p) < A_Y^k(R \mid p) \end{aligned} \quad \textbf{(9)}$$

Given a matching threshold α, then the weighted average rule is, When $M_{XY}(p) < \alpha$, the calculation of weights are as equation (9), when $M_{XY}(p) > \alpha$, the weights are as follows,

$$\begin{cases} \varpi_X = \frac{1}{2} - \frac{1}{2}(\frac{1-M_{XY}}{1-\alpha}), \quad \varpi_Y = 1 - \varpi_X, \quad A_X^k(R \mid p) < A_Y^k(R \mid p) \\ \varpi_X = \frac{1}{2} + \frac{1}{2}(\frac{1-M_{XY}}{1-\alpha}), \quad \varpi_Y = 1 - \varpi_X, \quad A_X^k(R \mid p) > A_Y^k(R \mid p) \end{cases} \tag{10}$$

Because of their different physical meaning, the approximation image and detailed ones are usually treated by different methods. Detailed coefficients having large absolute values correspond to sharp intensity changes and hence to salient features in the image, such as edges, lines and region boundaries. The nature of the approximation coefficients, however, is different from the detailed ones. The approximation image is a coarse representation of the original image and may have inherited some of the latter's properties, such as the mean intensity or texture information. In this paper, the composite approximation coefficients of the highest decomposition level adopt the weighted average method and the composite detailed coefficients adopt the maximum value method. So the robusticity is improved.

4 The Performance Estimation and Experimental Results

The quality of fusion image is generally estimated by the root mean square error ($RMSE$) between the fusion image and the ideal image, it is defined as,

$$RMSE = \sqrt{\frac{1}{M \times N} \sum_{m=1}^{M} \sum_{n=1}^{N} (G_i(m,n) - G_F(m,n))^2} \tag{11}$$

Where, $G_i(m, n)$ is the gray value of pixel point (m, n) in ideal image, $G_F(m, n)$ the gray value of the corresponding pixel point (m, n) in fusion image, $M \times N$ the image magnitude. The smaller $RMSE$ is, the better the fusion effect is.

Some researchers pointed that the number of the optimal levels of wavelet decomposition for multi-focus image is 3[8]. In our experiments, we use wavelet 'bior 2.2' to decompose the source images to three levels, and a set of segmentation regions are got using the method mentioned in section 3. The combination coefficients of approximate sub-image and detailed sub-images are respectively selected according to the two fusion rules, weighted average rule and maximum value rule. When the threshold $\alpha = 0.85$, the fusion effect is better.

To verify the performance of our algorithm, we use a set of source images (256×256). The comparison with the method proposed in reference 5 is also done. Fig.2 (a) and (b) are the two source images which have different focuses, the aim of fusion is to solve the local blurry problem in image caused by the different focus point. Fig.2 (a) focus on the big plane, it is clear, however, the small one is blurring; Fig.2 (b) focus on the small plane, the big one is blurring and the small one is clear. Fig.2 (c) is the segmentation regions used for guiding fusion (1 level), Fig.2 (d) the ideal image, Fig.2 (e) the fusion image got from the method proposed in reference 5, Fig.2 (f) the fusion image obtained by using our method, the big and small planes are all clear in this image, Fig.2 (g) the difference image between the fusion image got from method in reference 5 and the ideal image, and Fig.2 (h) the difference image between the fusion image got from our method and the ideal image. Seen from the fusion image and the difference image, the performance of our method is superior to the one in reference 5.

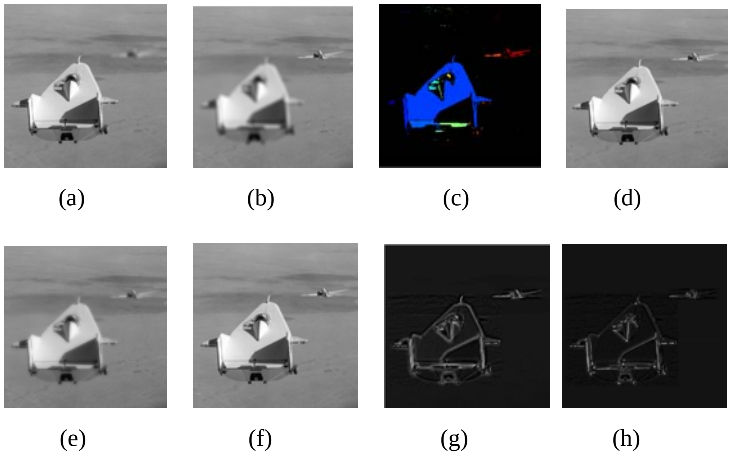

Fig. 2. Source images and experimental results

To verify the fusion algorithm proposed in this paper, we also use equation (11) to estimate the performance of fusion image. The standard deviation of the fusion image got from reference 5 is 13.4203 and the one of our fusion image is 8.9581. Known from the experimental results, the fusion effect of our method is better than the one of reference 5.

5 Conclusions

An image fusion algorithm based on segmentation region is proposed in this paper, it combine effectively the pixel-level and feature-level fusion method. Because it is guided by the segmented regions which represent different object, so the performance is better than the one which is not based on the object regions. First, the source images are decomposed by wavelet to get a set of segmentation regions of each level, and the activity measure, match measure and coefficients combination are carried out in these segmentation regions, in the end, the final fusion image is obtained by the inverse wavelet transformation of combination coefficients. Seen from the results, because our fusion method is carried out based on the segmentation regions and the match degree between source images, activity level and regional features are considered at the same time, the salient information of source image are better reserved, so the fusion performance is better.

References

[1] C. Pohl, J.L. Genderen.: Multi-sensor Image Fusion in Remote Sensing: Concepts, Methods and Applications. International Journal of Remote Sensing, Vol. 19 (5) (1998), pp. 823-854.
[2] C.Y. Wen, J.K. Chen.: Multi-resolution Image Fusion Technique and Its Application to Forensic Science. Forensic Science International. Vol. 140 (2004), pp. 217-232.
[3] Gonzalo Pajares, Jesus Manuel de la Cruz.: A wavelet-based Image Fusion Tutorial. Pattern Recognition. Vol. 37 (2004), pp. 1855-1872.
[4] YongAn Zheng.: The Adaptive Fusion Algorithm of Weighting Regional Features Based on Wavelet Transformation. Computer Engineering and Application. Vol. 6 (2004), pp. 94-96.
[5] Hao Yan.: Multi-scale Image Fusion Based on Wavelet Transformation. Computer Engineering and Applicaition. Vol. 29 (2003), pp. 131-134.
[6] B.J. Matuszewski, L.K. Shark.: Region-based Wavelet Fusion of Ultrasonic, Radiographic and Shearographyc Non-destructive Testing Images. Proceedings of the 15th World Conference on Non-destructive Testing. (2000), pp. 358-363.
[7] Z.Zhang, R. Blum.: A Region-based Image Fusion Scheme for Concealed Weapon Detection. Proceedings of the 31th Annual Conference on Information Sciences and System. (1997), pp. 168-173.
[8] Weiwei Wang.: Multi-focus Image Fusion Algorithm in Wavelet Domain. Systems Engineering and Electronics. Vol. 5 (2004), pp. 668-671.

Rong Wang is a Ph.D. student in Institute of Information Science & Engineering, Northeastern University, the research field is pattern recognition and intelligent system.

Li-Qun Gao is a professor in Institute of Information Science & Engineering, Northeastern University, the research field is pattern recognition and intelligent system.