SIViP (2007) 1:215–223
DOI 10.1007/s11760-007-0015-6

ORIGINAL PAPER

Adaptive region-based image fusion using energy evaluation model for fusion decision

Yingjie Zhang

Received: 10 August 2006 / Revised: 4 April 2007 / Accepted: 5 April 2007 / Published online: 9 May 2007

Abstract A new adaptive region-based image fusion approach is proposed. To implement image segmentation, the piecewise smooth Mumford-Shah segmentation algorithm is studied and a fast and simple method is proposed to solve the energy function. Two complementary functions u^+ and u^- of the algorithm, which are respectively looked as objects and background of the image, are extended into the whole image domain, and they are computed by linear or nonlinear diffusion. The key to the algorithm is to make optimal fusion decisions for every segmented region. For this purpose, an evaluation approach has to be given to measure the performances of the available fusion rules. Therefore an energy-based evaluation model, derived from the Total Variation principle, is proposed. By numerical experiment it has been demonstrated that despite an increase in complexity, the new approach has shown a number of advantages over previous ones, for example the ability to preserve all relevant information and remove some of side effects such as reducing contrast and sensitive to error of registration.

Keywords Image fusion · Energy function · Segmentation · Region-based fusion

List of symbols

u	the observed image
u^+	the object regions of an image
u^-	the background regions of an image
u_0	an input image
R^2	2D real domains
$\|\nabla u\|$	gradient of the image u
$\|\nabla u^+\|$	gradient of the image u^+
$\|\nabla u^-\|$	gradient of the image u^-
$\wp$	a bounded open subset in 2-dimensional real domains
ω	an open curve set
Γ	boundary curves
ϕ	level set curves
$\partial\omega$	zero level set curves in 2D space
ϕ_0	zero level set curves in 3D space
d	distance between point and zero level set
E_{TV}	energy of the region of an image from TV model
E_{MTV}	energy of the fused region of an image from the fusion evaluation model
F	speed of evolution of curves
$\delta(\phi)$	delta function in Chan–Vese–Mumford–Shah model
u_{s}	a smooth component of the observed image u_0
s	scale level
$\tilde{u}_{\mathrm{s}}$	resulting image obtained from u_{s} by the four-point averaging method
Ω	segmented regions of an image
u_{R}	a reference image with the least value for each pixel point
λ	a positive parameter to control the fidelity in the energy formulation
μ	a positive parameter to control the fidelity in the energy formulation
k	iteration numbers

Y. Zhang (✉)
School of mechanical engineering, Xi'an Jiaotong University, Xi'an, Shaanxi, 710049, People's Republic of China
e-mail: yjzhang@mail.xjtu.edu.cn

1 Introduction

Multiple sensor modalities provide data at different spatial, temporal and spectral resolutions, which allow for enhanced performance in a wide range of modern military and civilian imaging applications. It is the aim of image fusion to integrate different data in order to obtain more information than can be derived from each of the single sensor data alone. A good example is the fusion of images acquired by sensors sensitive to visible/infrared (VIR) with data from active synthetic aperture radar (SAR). The information contained in VIR imagery depends on the multi-spectral reflectivity of the target illuminated by sunlight. SAR image intensities depend on the characteristics of the illuminated surface target as well as on the signal itself. The fusion of these disparate data contributes to the understanding of the objects observed. In general, image fusion methods can be grouped into two classes: color related techniques and statistical or numerical methods. The first comprises the color composition of three image channels in the RGB color space as well as more sophisticated color transformations like the intensity-hue-saturation (HIS) and hue-saturation-value (HSV). These methods involve the transformation of a three- band combination of a multi-spectral image to an intensity, hue, and saturation color space image. The intensity component of this transformation is replaced with the panchromatic (PAN) image, and a transformation back into an RGB image is performed. Unfortunately, as stated by Liu in [1] that the spectral distortion introduced by these fusion techniques is uncontrolled and not quantified. Another disadvantage of these methods is that they are limited to three band composites.

Statistical approaches are implemented based on channel statistics including correlation and filters like the PCA method [2].

The numerical methods follow arithmetic operations such as image differencing and ratios to add a channel to other image bands. A sophisticated numerical approach uses wavelets in a multi-resolution environment. Wavelet transform is a linear tool in its original from [3], but nonlinear extensions of discrete wavelet transform are possible by various methods like lifting scheme [4]. Wavelet decompression image merger methods address the limitations described above by its ability to be performed on individual bands, and a decrease in the spectral distortion. It shows a good position of a function (here this function is the image) in spatial and frequency spaces.

Image information fusion can be carried out at signal, feature, and symbol levels [5]. Higher-level feature and symbol-level fusion algorithms combine information in the form of feature descriptors and probabilistic variables. The lowest level of signal image fusion fuses image information in its raw image signal representation by combining multiple input

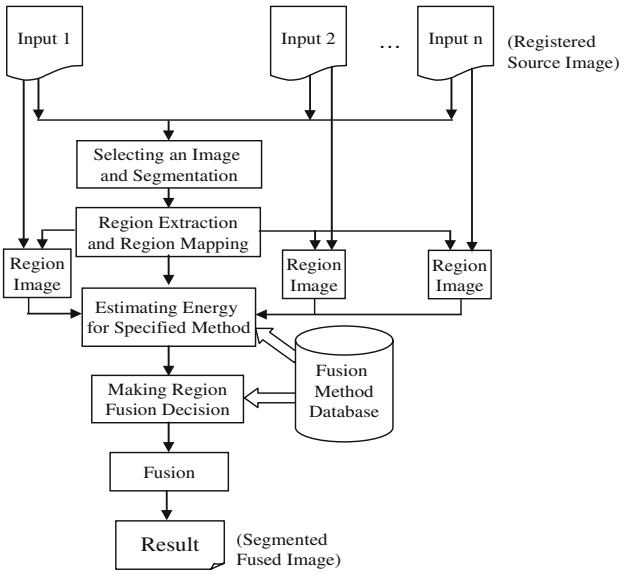

Fig. 1 Block diagram of the image fusion approach

image signals into a single fused image output. The objective of signal-level image fusion is to represent the whole visual information contained in a set of input images as a single fused image without distortion or loss of information. However, it is practically impossible for representation of all the visual information from an input set of images as a single image. A more practical approach is based on the faithful representation in the fused image of which only the most important input information is generally accepted. A signal-level image fusion algorithm is therefore required to identify the most important information in the input set of images and to transfer, without loss, this information into the fused image. One way of achieving feature-level fusion is with a region-based fusion scheme, in which an image is initially segmented in some way to produce a set of regions. Then various properties of the regions can be calculated and applied to determine which features from which images are included in the fused image. Lewis and O'Callaghan et al. in [6] proposed a feature level fusion method, in which region segmentation was performed by a dual-tree complex wavelet transform, and region based fusion is done in the wavelet domain. Yang and Blum [7] implemented the region fusion using a rigorous application of estimation theory, where a statistical image formation model using Gaussian mixture distortion is built for each region and the EM (expectation–maximization) algorithm is used in conjunction with the model to develop the region-level EM fusion algorithm to produce the fused image. The general flow diagram of the region-based fusion algorithm is shown in Fig. 1.

First, an image should be selected for segmentation. In order to obtain better segmentation, the selected image should

be with the most visual information of the input images. Then the image is segmented to produce a set of regions, and the valid regions that are with their areas greater than a given threshold value are located by contour tracing algorithm. The corresponding regions in the other images are computed by linear mapping, and the small-size versions of the input images also are computed. Finally, optimal fusion rules will be made for each region by evaluating all of candidates. In our current scheme, only some simple fusion approaches are included as candidates, but the extension can be performed easily by adding more complex fusion approaches like wavelet-based methods into the library. After that, more computational time needs to be spent in evaluating processes.

In the region-based fusion methods, however, the use of more sources of sensor information in imaging systems leads to "information overload" problems. In response to this, research attention in the past ten years is focused on techniques which aim to deal with the overwhelming increase of input data provided by multi-sensor arrays without canceling out the benefits of additional information. The field of multi-sensor image fusion represents a plethora of such techniques designed to reduce the physical amount of input image data while preserving information.

Note that the previous researches had taken into account relatively simple methods to combining the source images over the whole image domain. It is practically impossible to get the fused image with the best good quality by this way due to non-homogeneous of images. The purpose of this paper is to propose a region-based image fusion scheme for multi-sensor images. As have seen in Fig. 1, an input image with most important visual information is initially segmented to produce a set of regions. Then, the corresponding regions on the other images are obtained by linear mapping. The available fusion rules are stored into a database, and an evaluation model is proposed to make the best fusion decisions for every region. The final fused image is obtained by combining the fused regions in which fusion has been performed with their own fusion rules. Despite an increase in complexity, it has shown a number of advantages over previous methods, for example the ability to preserve all relevant information in the fused image and remove some of side effects such as reducing contrast and sensitive to error of registration.

The rest of this paper is organized as follows. Section 2 introduces related work and background. Section 3 describes a new strategy for implementation of image segmentation. The fusion scheme is described in Sect. 4, and the related modules are also introduced, which is followed by algorithm overview and numeric results given in Sect. 5 and Sect. 6. Finally the conclusions and acknowledgments are given in Sect. 7 and Sect. 8, respectively.

2 Related work and background

In this section, we briefly mention here some of the most related relevant works.

2.1 Rudin–Osher–Fatemi energy model

Known from the existing segmentation and denoising approaches, the energy functional approaches have been given widely attention. Details regarding the interaction and close relations among these approaches can be found in [8–10]. A classical variational denoising algorithm is the total variation (TV) minimizing process of Rudin–Osher–Fatemi (ROF) [11].

Let u be an observed image, u_0 the input image, and $|\nabla u|$ the gradient of u at region boundaries. The ROF algorithm seeks an equilibrium state (minimal energy) of the energy functional comprised of the TV norm of the observed image u and the fidelity of this image to the noisy input image u_0, and a variant of the model is introduced as:

$$E_{\mathrm{TV}} = \int_{\Omega} \left(|\nabla u| + \frac{1}{2}\lambda(u - u_0)^2 \right) \mathrm{d}x, \tag{1}$$

where λ is a scalar controlling the fidelity of the solution to the input image (inversely proportional to the measure of denoising). In the ROF model (1), the energy is defined over the image domain and typically processes local maximum at the intensity edges occurring at object boundaries. The magnitude of the energy reflects the edge intensity distributions. Thus, the image with more visual information has larger energy than the image with few one. Based on this idea, the energy value can be used to evaluate image properties.

2.2 Level sets and Chan–Vese–Mumford–Shah model

Let $\wp$ be a bounded open subset of R^2, with Γ as its boundary. Then a two-dimensional image u_0 can be defined as $u_0 : \wp \to R$. In this case $\wp$ is just a fixed rectangular grid. Now consider the evolving curve Γ in $\wp$, as the boundary of an open subset ω of $\wp$. In other words, $\omega \in \wp$, and Γ is the boundary of ω. This idea is to embed this propagating curve as the zero level set of a higher dimensional function ϕ, which is defined as follows:

$$\phi(x, y, t = 0) = \pm d \tag{2}$$

where d is the distance from (x, y) to $\partial\omega$ at $t = 0$, and the plus or minus sign is chosen if the point (x, y) is outside or inside the subset ω.

Now, the goal is to make an equation for the evolution of the curve. Since the interface Γ may be represented as the

zero iso-contour $\Gamma = \{(x, y)|\phi(x, y, t) = 0\}$. The equation for the motion of the interface is:

$$\begin{aligned} \frac{\partial \phi}{\partial t} &= F|\nabla \phi| \\ \phi(x, y, 0) &= \phi_0(x, y) \end{aligned} \tag{3}$$

where the set $\{(x, y), \phi_0\ (x, y) = 0\}$ defines the initial contour, and F is the speed of propagation.

For certain forms of the speed function F, this reduces to a standard Hamilton-Jacobi equation. There are several major advantages to this formulation. The first is that $\phi(x, y, t)$ always remains a function as long as F is smooth. As the surface ϕ evolves, the curve Γ may break, merge, and change topology.

Chan–Vese [12,13] solved Mumford–Shah energy model [14] using the level set methods. Considering a two-phases image in the piecewise smooth Chan–Vese–Mumford–Shah (CVMS) approach, the object regions are defined as u^+ (maybe multi-connected), the background regions as u^-, and ϕ is the level set function. The pixels in the region of u^+ are with $\phi > 0$, while the pixels in the region of u^- with $\phi < 0$, and the signed distance is applied as the level set function. Generally, an initial contour, in level set-based approaches, has to be interactively given by user for the evolution of the curve. An energy function is defined over the image domain and typically with local maximum at the object's boundaries. According to definition of the energy function, as the curve evolves the noises are continuously removed from both regions of u^+ and u^- with different speeds. As a result, the regions of u^+ and u^- become more and more smooth as time goes by, and the boundaries of the regions are identified as zero level set.

3 Image segmentation and region extraction

Segmentation has been very important for the region-based image fusion approaches. The regions on the image should be exactly extracted so that adaptive fusion rules can be applied on them in turn to improve the fusion quality. However, unlike the previous work in [15,16], each of the regions segmented is extended to the whole image domain for easy to compute. Furthermore, the union of segmented regions not only covers the image domain, but also the range of each region is extended into the whole image domain. For example, if there were only an object on the image to be segmented, the final segmentation would consist of two regions: one covered by the object and another occupied by the background of the image. However each of them has values over the whole image domain with zero on boundaries of the regions, opposite signs in the interior and exterior of regions, respectively. One of advantages of this approach is that objects would never be lost, even if segmentation algorithms missed some of them, since the missed objects would be further processed in the same way with the regions they lie on.

Inspired by the background subtraction approach [17], a new approach is proposed to solve the piecewise smooth CVMS model for region segmentation, which has been demonstrated in [18] detailedly. Unlike definition in the CVMS algorithm, in our scheme, u^+ and u^- are directly given by two smooth versions of the image obtained respectively at different scales or resolutions. Assume that u^+ is regarded as object regions of the image, and u^- as background regions. u^+ is obtained by edge preserving diffusion, and u^- with blurring operation. Consequently, difference between u^+ and u^- is applied to estimate the initial level set.

Note that the gradient terms, i.e. $\left|\nabla u^+\right|$ and $\left|\nabla u^-\right|$ in the CVMS model, induce the blurring effect and the boundary length term affects the strength of the model's rigidity only. Both of them have little contribution to computation of the region boundaries. They can be removed for simplicity, and the energy formulation can be rewritten as:

$$\begin{aligned} E(u^+, u^-, \phi) = &\int\limits_{\text{inside}(\phi)} \left|u^+ - u_0\right|^2 \mathrm{d}x \\ &+ \int\limits_{\text{outside}(\phi)} \left|u^- - u_0\right|^2 \mathrm{d}x. \end{aligned} \tag{4}$$

Then, minimizing (4) described above with respect to ϕ, we can obtain the zero level set as (see Appendix for detail):

$$\begin{aligned} \frac{\partial \phi}{\partial t} &= \delta(\phi)\left(|u^- - u_0|^2 - |u^+ - u_0|^2\right) \\ &= \delta(\phi)(u^- - u^+)\left(\tfrac{u^+ + u^-}{2} - u_0\right) \end{aligned}, \tag{5}$$

With both time step and $\delta(\phi)$ fixed, we can easily compute the values of ϕ for each point. The regions of u^+ or u^- may be respectively located via the sign of ϕ, and their boundaries are determined based on the change of the sign of ϕ.

3.1 Choice of scale level for u^+

For obtaining the u^+ image, the original image should be filtered to remove noises and small texture details, while the object boundaries must be remained. Therefore denoising operators with edge–preserving regularization should be applied to the computation of u^+. Unlike the definition in CVMS model, here the regions of u^+ are extended into the whole image domain with positive value inside, negative outside and zero on its boundaries. However, only the available parts, i.e. the interior of the regions, are used in the segmentation algorithm. For removal of noises and background texture, the image frequently is filtered with some scale level. The choice of scale levels depends on the magnitude of noises and texture details to be removed. The coarser scale should be selected for images with the bigger level of texture details

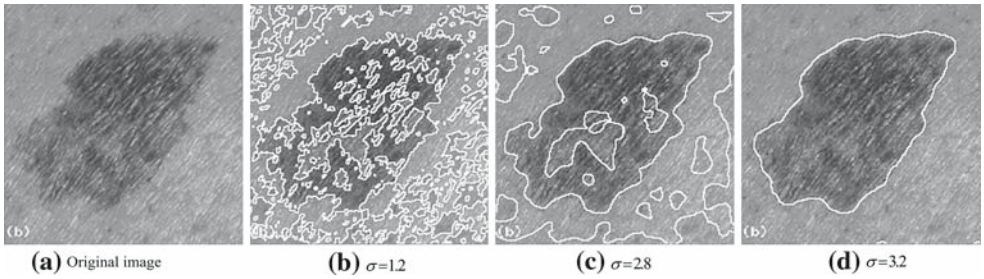

Fig. 2 Original image and u^+ obtained using different filtering parameters

or strong noises, while the finer scale level could be chosen for "clean" images. All kinds of linear or nonlinear methods such as Gaussian filtering, isotropic smoothing, and Perona–Malik regularization can be used for computation of u^+. In our numerical experiment, the four-point averaging has been applied, which is given as:

$$u_{i,j}^{+(k+1)} = \frac{1}{4}(u_{i-1,j}^{+(k)} + u_{i+1,j}^{+(k)} + u_{i,j-1}^{+(k)} + u_{i,j+1}^{+(k)}), \tag{6}$$

where k is iteration numbers, and the number of iteration is equivalent to the scale level. In other words, the results at coarser scale can be obtained from (6) by selecting the more number of iteration, while the one at finer scale by the few number of iteration.

It has been numerically verified that the scale level for u^+ has significant effect on the segmentation results, in particular, for noisy image [18]. Figure 2 shows several results, which are obtained via the convolution of the original image with some normalization 2D Gaussian kernels. The original image is shown in (a), and the corresponding results obtained with different filtering parameters are shown in (b)–(d), respectively.

3.2 Construction of u^- or background

For computation of u^-, the objective is to obtain an approximate background image. The original image has to be blurred enough for the removal of all objects, texture and noises, only the background left behind. It has been demonstrated that if the scale level for u^- were beyond a threshold, the resulting segmentation would keep constant. However it is practically impossible to automatically determine an adaptive threshold for every image. If the threshold were selected too big it would result in expensive computational cost, or else some false objects would be remained as shown in Fig. 2b and c if a smaller threshold were applied. Therefore, one has to select a bigger threshold in practice for better segmentation. To reduce computational cost, a fast multiresolution-based smooth algorithm is studied and applied on the segmentation. Considering multi-resolution images are the approximation of the original image at different scales to form a pyramid, all of them have self-similarity. The similarity is shown in two aspects: one is intra-scale clustering, where the similarity features at the same scale show the clustering property; the other is inter-scale similarity, in which the children inherit the features of their parent and have their own features. Considering the two properties described above, it could be concluded that the children have features similar to their parent plus parent's neighbours across scales and the finer features can be led by the coarser result. On basis of this theory [19], one can do smoothing at coarser scale, then the coarse result is extended to the original image at scale space for speeding up smooth. The algorithm can be described as follows:

1. Let u_s denote a smooth component of the original image u_0, s denotes the selected scale level and a pixel on u_s corresponds to 2^s pixels of u_0.
2. u_s is calculated from u_0 with the bilinear interpolation approach.
3. $\tilde{u}_s$ is obtained by smoothing u_s using the four-point averaging method (6) with a few iterations.
4. Prolongating $\tilde{u}_s$ to the original scale as u^-.

4 Region-based fusion scheme

As have been pointed out before, most previous fusion methods apply only one fusion rule. Thus the actual images are very complex and usually contain a number of non-homogeneous regions. For obtaining better fusion quality, adaptive fusion rules should be respectively applied in each region on the image. Inspired from the scale space approach, the image fusion, in our scheme, is performed not directly on the original image, but on its small-size version, which is obtained by bilinear interpolation to speed up the evaluation process.

4.1 Strategy of fusion evaluation

The goal of this paper is to determine an adaptive fusion rule for each region on the image to get good fusion quality. To do this, the key is to find out a good evaluation approach. It

then is used to make region fusion decisions via evaluating the performances of the available fusion approaches.

Let u denote a fused image, Ω a segmented region, and u^- the approximate background. u_R is defined as a reference image with the least value for each pixel point, which can be calculated by the minimum pixel value fusion algorithm. The energy formula, derived from ROF model, is given by:

$$E_{\mathrm{MTV}} = \int_{\Omega} (|\nabla u| + \lambda(u - u^-)^2 + \mu(u_R - u^-)^2)\,\mathrm{d}x, \quad (7)$$

where λ is a scalar controlling the fidelity of the solution to the input image, and μ is a positive parameter, which controls the effect of approximate background on the fused region Ω.

The first term, i.e. the TV term $|\nabla u|$, is applied to measure the blurring effect of u. It would have larger value if u were with more details or edges. The second term (fidelity term) that measures the closeness to u^- is used to estimate the variance of detail level of the region Ω. The last term is applied to compensate for the variation of the approximate background owing to the change of the total energy. For example, theoretically a larger energy value should be obtained from (7) for a point with higher brightness on the fused image. However a lower energy value could be obtained from the second term if the corresponding point of the approximate background u^- were with higher brightness, too. Therefore the total energy should be adjusted by the last term to maintain a larger value. In which case described above, the point on u^- has higher brightness, whereas the corresponding one on the reference image u_R is with lower brightness. Consequently, a larger value is obtained from the last term, and the larger energy value at the point is maintained as the total energy. On the contrary, if u^- were very close to u_R the last term would have little contribution to the total energy.

Comparing with the TV model, in which the TV term $|\nabla u|$ is included to consider the variation of the details of the image, the fidelity term and the last term in (7) can measure the information remained and the variation of brightness and contrasts, respectively. Therefore it will be more effective to be applied to measure the fusion quality.

4.2 Fusion decision making

In contrast with previous approaches, different fusion rules are chosen and applied to the segmented regions on an image. To select best fusion rule for a specified region, all the available fusion rules in the fusion rule database have to be tried and evaluate one by one in the region. In our current implementation, there are a lots of fusion rules in the library, including average pixel value, PCA, maximum pixel value, minimum pixel value methods, etc.

Generally speaking, the cost of evaluating depends on the number of fusion rules in the fusion rule database. The reason for this is that the fusion rules have to be tried in turn over each of the regions. As more fusion rules are added into the database, it will require much more time in evaluating processes. To overcome the drawback, one way is to reduce the number of available fusion rules, which is implemented by classifying the fusion rules into groups based on applications. First, likely fusion rules are chosen and arranged in a group for a specified application, then the evaluation can be done based on the candidate fusion rules in the group. Practically it is possible to enumerate all available candidate rules for a specified application. Some methods like exemplar-based evaluation or histogram analysis may be considered to exclude some non-available candidates; another is to reduce the evaluating cost by using small-size image based on the philosophy of scale space. That is, evaluation processes are performed on a small-size version of the image. The resulting fusion in each region, however, is carried out on the original size image.

5 Results and comparison

The proposed approach has been implemented in Matlab6.5 on Microsoft Windows XP environment and tested on several images. The weight coefficients in (7) are with $\lambda = 2.8$ and $\mu = 1$ in all our experiments. The results are presented and compared with previous methods. Figure 3 shows the fusion of a visual image (a) and a millimeter wave (94 GHz MMW) image (b), which is employed in concealed weapon detection. It can be seen that the concealed weapons have been exposed on the fused image. The segmentation is performed on the original image (a), and the segmented result is shown in (c) with the number of valid regions 11. The corresponding results obtained by the maximum pixel value and the average pixel value methods are also given in (d) and (e), respectively. The result obtained by proposed method is shown in (f), where different fusion rules have been applied on the segmented regions. For example, in contrast (d) with (f), it is clear that the maximum pixel value method was applied on the region at the center of the image.

Figure 4 shows the fusion of two multi-focus office images. The input images are respectively shown in (a) and (b). The final segmentation, performed on the image (b), is shown in (c). The number of the valid regions is 21. The results by the average pixel value and the PCA approaches are shown in (d) and (e), respectively. The fused image obtained by proposed method is shown in (f). It can be seen that the result with proposed method is slightly better than those of the other ones, since different fusion rules are applied for each of the valid regions.

To demonstrate effectiveness of proposed evaluating model, the results in Figs. 3 and 4 are evaluated by using

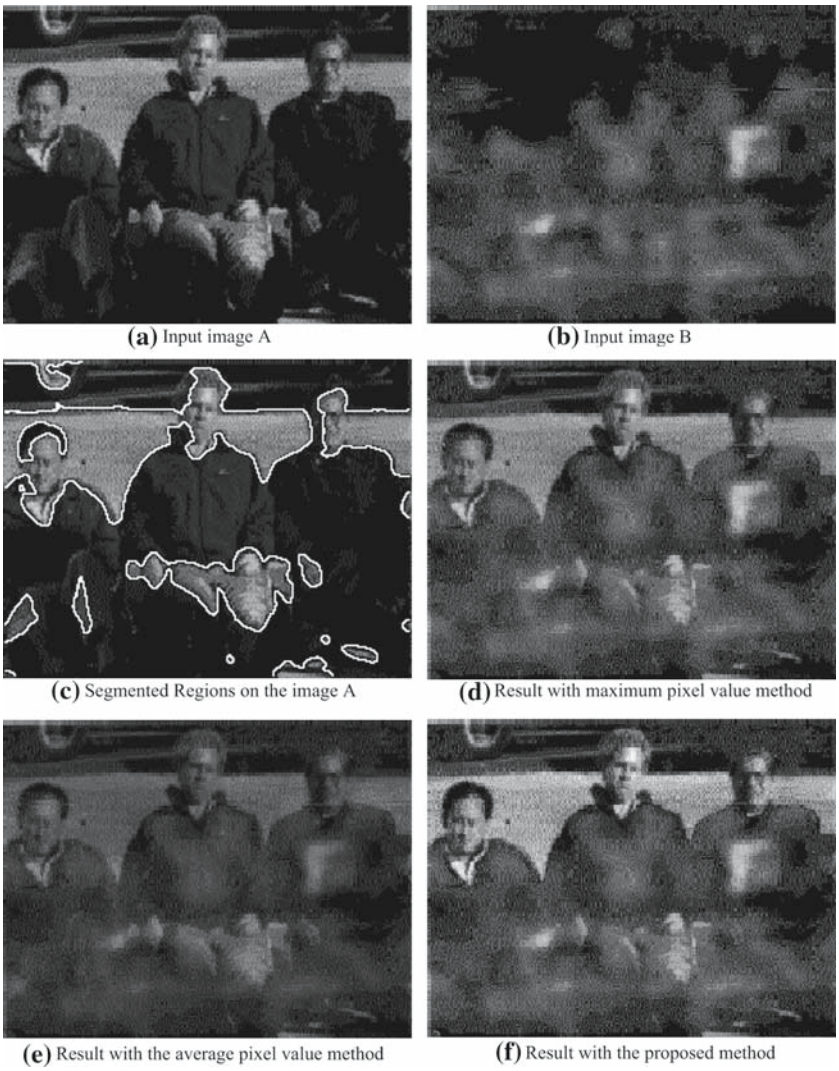

Fig. 3 Fusion results on visual and radiometric images

the proposed algorithm and several classical evaluation approaches, and the evaluating results are analyzed and compared. The evaluation methods include the root mean square error (RMSE) that is used as the evaluation criterion between the exemplar image and the fused image, standard deviation (SD) that is the square root of the variance, which reflects the spread in the data, the entropy which is used as a measure of information content, which is the average number of nits needed to quantize the intensities in the image, the cross entropy that evaluates the information difference between the two images to give the grey distribution of them, and the correlation (CORR) between the fused image and the original image. The evaluated results are listed in Table 1.

It can be clearly seen that the average pixel value method gave the baseline results for the images in Figs. 3 and 4. The PCA method gave a better result for the images in Fig. 3, and an equivalent but a slightly worse result for the images in Fig. 4. In contrast with the average pixel value and PCA methods, the maximum pixel value method gave the best result for the images in Fig. 3. However, these methods have poor results relatively to proposed method.

6 Conclusions

This paper presented an adaptive region-based image fusion approach. First, the whole image domain is separated into a set of regions based on homogeneity of images. The image fusion then is replaced by the fusion on the regions. For obtaining good fusion quality, an energy model is proposed to measure the performances of the candidates in the regions to make optimal fusion rules for each region. Note that available fusion rules come from the fusion method library. Only some simple fusion rules are included now, but more fusion rules can be easily added into for extension. This extension would be future works. Since the fusion rules of the library, however, have to be tried one by one for every region during evaluating, it would be computationally expensive. Several

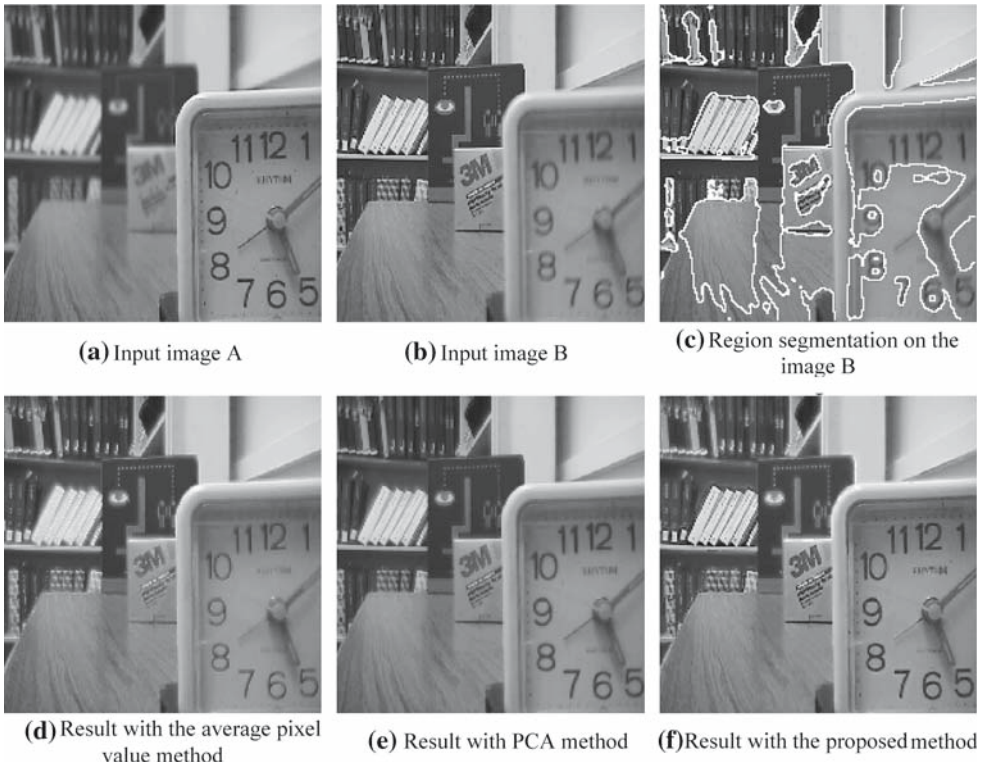

Fig. 4 Fusion of multi-focus office images

Table 1 Results from variant evaluation methods for the fused images in Figs. 3 and 4

Images	Methods	Entropy	Cross entropy	CORR	SD	RMSE	Energy
Figure 3	Average pixel value method	4.7822	1.2894	0.7906	33.3076	25.7100	0.3581
	PCA method	4.5917	2.3752	0.8067	65.1407	35.2116	0.3832
	Maximum pixel value method	3.9715	0.7345	0.7919	56.1162	33.4767	0.3904
	Minimum pixel value method	1.5185	−0.0646	0.3278	14.9178	39.1286	0.1378
	Proposed method	4.4995	2.2482	0.8087	65.4615	35.5667	0.4812
Figure 4	Average pixel value method	7.1828	0.1458	0.9964	107.8074	9.1439	0.3081
	PCA method	7.1849	0.1347	0.9961	107.6955	9.4462	0.3112
	Maximum pixel value method	7.1475	0.0712	0.9935	112.7557	12.5125	0.2947
	Minimum pixel value method	7.1293	0.0707	0.9920	103.5971	13.3671	0.3044
	Proposed method	7.2450	0.1837	0.9889	113.3761	16.4087	0.3103

efforts had been tried to overcome the drawback, but it would be holding a candle to the sun if there were much more fusion rules to be tried. Another difficulty is segmentation. Since one of the input images is selected for segmentation in our algorithm it is not perfect. A few regions may be missed in some applications like the fusion of IR images and visible images. The segmentation based on compounded input images needs to be researched further.

Acknowledgments The author would like to thank the unknown reviewers for their very carefully reading the manuscript, which lead to further improvement of the paper. This work was supported in part by National Hi-Tech Research and Development Program of China (863 program) under Grant no. 2006AA04Z142 and Shaanxi Province Science Foundation under Grant no. 2006K05-G22.

Appendix

Let

$$E_\Omega = \iint_\Omega f(x, y)\,\mathrm{d}x\,\mathrm{d}y \tag{I}$$

The goal is to find out the boundary $\partial\Omega$ of a region Ω for a given function $f : \Omega^2 \to \Omega$ that yields an extreme of the

energy E_Ω, which can be solved by the Green theorem, and described as following:

Let

$$Q = \frac{1}{2}\int_0^{\mathrm{x}} f(t, y)\,\mathrm{d}t, \quad \text{and} \quad P = \frac{1}{2}\int_0^{\mathrm{y}} f(x, t)\,\mathrm{d}t \tag{II}$$

Known from Green theorem,

$$\begin{aligned} E_\Omega &= \iint_\Omega \left(\tfrac{\partial Q}{\partial x} + \tfrac{\partial P}{\partial y}\right)\mathrm{d}x\,\mathrm{d}y = \oint_{\partial\Omega} (Q\,\mathrm{d}y - P\,\mathrm{d}x) \\ &= \oint_{\partial\Omega} \left(Q \cdot \tfrac{dy}{ds} - P \cdot \tfrac{dx}{ds}\right)\mathrm{d}s \\ &= \int_0^L \left(Q \cdot \tfrac{\mathrm{d}y}{\mathrm{d}s} - P \cdot \tfrac{\mathrm{d}x}{\mathrm{d}s}\right)\mathrm{d}s \\ &= \int_0^L F\big[s, x(s), y(s), x'(s), y'(s)\big]\,\mathrm{d}s \end{aligned} \tag{III}$$

where s denotes the arc length and L, the length of $\partial\Omega$. From variational method, the corresponding Euler–Lagrange equation is given as

$$\begin{cases} F_x - \dfrac{\mathrm{d}}{\mathrm{d}s}F_{x'} = 0 \\ F_y - \dfrac{\mathrm{d}}{\mathrm{d}s}F_{y'} = 0 \end{cases}, \tag{IV}$$

and it can derivate the following equations:

$$\begin{cases} y'(s)\dfrac{\partial Q}{\partial x} - x'(s)\dfrac{\partial P}{\partial x} + \dfrac{\mathrm{d}P}{\mathrm{d}s} = 0 \\ y'(s)\dfrac{\partial Q}{\partial y} - x'(s)\dfrac{\partial P}{\partial y} + \dfrac{\mathrm{d}Q}{\mathrm{d}s} = 0 \end{cases} \tag{V}$$

where

$$\begin{cases} \dfrac{\mathrm{d}P}{\mathrm{d}s} = x'(s)\dfrac{\partial P}{\partial x} + y'(s)\dfrac{\partial P}{\partial y} \\ \dfrac{\mathrm{d}Q}{\mathrm{d}s} = x'(s)\dfrac{\partial Q}{\partial y} + y'(s)\dfrac{\partial Q}{\partial y} \end{cases} \tag{VI}$$

Substituting (VI) to (V), one could get

$$\begin{cases} -x'(s)\left(\dfrac{\partial Q}{\partial x} + \dfrac{\partial P}{\partial y}\right) = -x'(s)\cdot f(x, y) = 0 \\ \quad y'(s)\left(\dfrac{\partial Q}{\partial x} + \dfrac{\partial P}{\partial y}\right) = y'(s)\cdot f(x, y) = 0 \end{cases} \tag{VII}$$

Using the gradient descent method, one could obtain the curve evolution equation as:

$$\begin{cases} \dfrac{\partial x}{\partial t} = -f(x, y)\cdot\dfrac{\mathrm{d}y}{\mathrm{d}s} \\ \dfrac{\partial y}{\partial t} = f(x, y)\cdot\dfrac{\mathrm{d}x}{\mathrm{d}s} \end{cases} \tag{VIII}$$

Known from differential geometry, $\vec{\tau} = \left(\frac{\mathrm{d}x}{\mathrm{d}s}, \frac{\mathrm{d}y}{\mathrm{d}s}\right)$ is the unit tangent vector, and $\vec{N} = \left(-\frac{dy}{ds}, \frac{dx}{ds}\right)$ is the unit inward normal vector. Therefore one could obtain:

$$\frac{\partial C}{\partial t} = f(x, y)\cdot\vec{N}. \tag{IX}$$

References

1. Liu, J.G.: Smoothing filter-based modulation: a spectral preserve image fusion technique for improving spatial details. Int. J. Remote Sens. **21**(18), 3461–3472 (2000)
2. Mouat, D.A., Mahin, G.G., Lancaster, J.: Remote sensing techniques in the analysis of change detection. Geocarto Int. **2**, 39–50 (1993)
3. Mallat, S.: A wavelet tour of signal processing. Academic, London (1998)
4. Claypoole, R.L., Davis, G.M., Baraniuk, R.G.: Nonlinear wavelet transforms for image coding via lifting. IEEE Trans. Image Process. **12**(2), 1449–1459 (2003)
5. Klein, L.: Sensor and data fusion concepts and applications. SPIE, Bellingham (1999)
6. Lewis, J.J., O'Callaghan, R.J., et al.: Region-based image fusion using complex wavelets. Proceedings of the 7th International Conference on Information Fusion, Stockholm, pp. 555–562, (2004)
7. Yang, J., Blum, R.S.: A region-based image fusion method using the expectation-maximization algorithm. Proceedings of the Conference on Information Sciences and Systems (CISS), Princeton (2006)
8. Petrovic, V.S., Xydeas, C.S.: Sensor noise effects on signal-level image fusion performance. Inf. Fusion **4**, 167–183 (2003)
9. Weickert, J.: A review of nonlinear diffusion filtering. Scale space theory in computer vision, LNCS 1252. Springer, Berlin, pp. 3–28 (1997)
10. You, Y., Xu, W.: A. Tannenbaum, M. Kaveh, Behavioral analysis of anisotropic diffusion in image processing. IEEE Trans. Image Process **5**(11), 1539–1553 (1996)
11. Rudin, L., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Physica **60**, 259–268 (1992)
12. Chan, T.F., Vese, L.A.: Active contours without edges. IEEE Trans. Image Process. **10**(2), 266–277 (2001)
13. Chan, T.F., Vese, L.A.: A level set algorithm for minimizing the Mumford–Shah functional in image processing. In: Proceedings of the First IEEE Workshop on Variational and Level Set Methods in Computer Vision, pp. 161–168 (2001)
14. Mumford, D., Shah, J.: Optimal approximations by piecewise smooth functions and associated variational problems. Comm. Pure Appl. Math. **42**(5), 577–685 (1989)
15. Zhang, Z., Blum, R.: Region-based image fusion scheme for concealed weapon detection. In: Proceedings of the 31st Annual Conference on Information Sciences and Systems pp. 168–173 (1997)
16. Piella, G.: A region-based multi-resolution image fusion algorithm. In: ISIF Fusion conference, Annapolis, USA, pp. 1557–1564 (2002)
17. Kim, S., Lim, H.: Method of background subtraction for medical image segmentation. In: Proceedings of the 3rd International Conference on Cybernetics and Information Technologies, Systems and Applications, pp. 87–91 (2006)
18. Zhang, Y.J., Ge, L.L.: Automatic generation of initial contour for image segmentation approaches using curve evolving. Electron Lett Comput Vis Image Anal (Accepted and to appear)
19. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. IEEE Trans Pattern Anal Mach Intell **12**(7), 629–639 (1990)