World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:3, No:5, 2009

Region-Based Image Fusion with Artificial Neural Network

Shuo-Li Hsu, Peng-Wei Gau, I-Lin Wu, and Jyh-Horng Jeng

***Abstract*—**For most image fusion algorithms separate relationship by pixels in the image and treat them more or less independently. In addition, they have to be adjusted different parameters in different time or weather. In this paper, we propose a region–based image fusion which combines aspects of feature and pixel-level fusion method to replace only by pixel. The basic idea is to segment far infrared image only and to add information of each region from segmented image to visual image respectively. Then we determine different fused parameters according different region. At last, we adopt artificial neural network to deal with the problems of different time or weather, because the relationship between fused parameters and image features are nonlinear. It render the fused parameters can be produce automatically according different states. The experimental results present the method we proposed indeed have good adaptive capacity with automatic determined fused parameters. And the architecture can be used for lots of applications.

***Keywords*—**Image fusion, Region-based fusion, Segmentation, Neural network, Multi-sensor.

I. INTRODUCTION

MULTI-SENSOR data often exist complementary information, so data fusion provides an effective method to enable comparison and analysis. [1] Image fusion is a sub area of the more general topic of data fusion. The aim of image fusion, apart from reducing the amount of data, is to create new images that are more suitable for the purposes of human/machine perception, and as preceding image-processing. Image fusion [2]-[3] can be roughly defined as the process of combining multiple input images into an image, which contains the 'relevant' information from the inputs. The aim of image fusion is to integrate complementary and redundant information from multiple images to create a composite that contains a better fused image than any of the individual source images. The fused image should increase the performance of the subsequent processing tasks such as segmentation, feature extraction and object recognition. The different images to be fused can come from different sensors of the same basic type or they may come from different types of sensors.

S. L. Hsu is currently with the Department of Information Engineering, I-Shou University, Kaohsiung Taiwan (e-mail: d9203001@stmail.isu.edu.tw).

P. W. Kou is currently with the Department of Information Engineering, I-Shou University, Kaohsiung Taiwan (e-mail: abcdgod@hotmail.com)

I L. Wu is currently with the Department of Information Engineering, I-Shou University, Kaohsiung Taiwan (e-mail: 594510@gmail.com).

J. H. Jeng is a currently with the Department of Information Engineering, I-Shou University, Kaohsiung County, Taiwan (e-mail: jjeng@isu.edu.tw).

This work was supported in part by the National Science Council under Grant NSC96-2221-E-214-044-MY3-2.

There are many papers about image fusion have been published with the emphasis on improving fusion quality and reducing color distortion [4]-[6]. Among many of them, the most popular and effective are, for example [7]-[10], IHS (Intensity, Hue, Saturation), PCA (Principal Components Analysis), arithmetic combinations, and wavelet base fusion [11]. But they have reported the common limitations of existing fusion techniques due to deal with whole image. It often introduces that some part of fused image look like good and others do not. In recent years, there are more and more papers about improved image fusion method with region-based [12]. But it is still no automatic solution for different datasets which means different time or weather [13]. In order to overcome this problem, we propose region-based image fusion architecture with setting fused parameters which can be determined according different image features of regions.

Artificial neural network has good ability to learn from examples and extract the statistical properties of the examples during the training procedure. So it has great advantage at nonlinearly problem. We adopt artificial neural network to solve automatic solution because of the relationship between features and fused parameters are nonlinear.

The rest of this paper is organized as follows: In section II, the basic method survey will be introduced. In section III, we will present modified fusion method and proposed fusion scheme will be described, too. Experimental results will be presented in section IV, and the last section gives some conclusions.

II. IMPROVED IHS-BASED FUSION

The basic idea of IHS fusion method is to convert a color image from the RGB (Red, Green, Blue) color space into the IHS (Intensity, Hue, Saturation) color space. One of them will be replaced by another image when we got the intensive information of both images. Then we convert IHS color space with H and S of being replaced image into RGB color space. See the following procedure:

Step1: Transform the color space from RGB to IHS.

$$\begin{bmatrix} I_v \\ V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{-\sqrt{2}}{6} & \frac{-\sqrt{2}}{6} & \frac{2\sqrt{2}}{6} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

where I_v is intensity of visual image. R,G,B is color information of visual image respectively. V_1 and V_2 are components to calculate hue H and saturation S. $H = \tan^{-1}(V_2 / V_1)$, $S = \sqrt{V_1^2 + V_2^2}$.

Step 2: The intensity component is replaced by intensity of infrared image I_i.

Step 3: Transform the color space from IHS to RGB.

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} 1 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ 1 & \sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} I_i \\ V_1 \\ V_2 \end{bmatrix} \tag{2}$$

where I_i is intensity of infrared image. R', G', B' is color information of fused image respectively.

Because our basic idea is to add useful information of far infrared image to visual image. We set fused parameters in the matrix instead of the intensity of far infrared image I_i to replace the intensity of visual image I_v. The fused parameters will be adjusted according different information of each region. The following formula is modified result.

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} 1 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ 1 & \sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} \alpha I_v + \beta I_i \\ V_1 \\ V_2 \end{bmatrix} \tag{3}$$

where α, β are fused parameters. $0 \le \alpha, \beta \le 1$.

III. Water Region Area

Due to we want to add information of region in infrared image. We only segment the infrared image here. We adopt the method of segmentation according histogram. In accordance with the intuitionist features of the histogram, the peaks of the histogram are considered as watersheds, each valley including two neighboring peaks and a bottom points.

Step 1: Draw image histogram and smooth it to decrease noise influence if necessary.

Step 2: Seek all peaks and bottom points in the histogram.

Step 3: Calculate the water region area from the left bottom point. Here, define θ as a lower limitation ranging from 0.01 to 0.03. The smaller the value of θ is, the more threshold points we will get. When the water region area is larger than θ, the corresponding bottom point will be kept in threshold array T_m. Meanwhile, the corresponding left side peak point will be kept in peak points array P_m. Otherwise, the valley will be taken as invalid. At this situation, compare the two peaks' values located in the valley's two sides: (1) if the left peak point is larger than the right one, it will be treated as the new left peak point. While the next right peak point will be the new right peak one, the smaller between the current and the next bottom point will be regarded as the new bottom point. (2) otherwise, the right peak point, the right bottom point and the next right peak point will be regarded as new left peak point, new bottom point and new right peak point respectively and then the new water region area will be calculated again.

Step 4: Iteratively execute step 3 until all bottom points have been processed and then we can get the threshold array T_m ($m = 1,...,M+1$ and $T_1 < ... < T_m$) and the corresponding peak array P_m ($m = 1,...,M+1$ and $P_1 < ... < P_{m+1}$). Hence, a valid valley V_m includes two neighboring peaks $\{P_m, P_{m+1}\}$ and a threshold T_m ($P_m < T_m < P_{m+1}$).

We can use the multi-threshold to segment the infrared image when finish the above procedure. In our experiments, because there are usually some noises in the infrared image, we do denoise before estimating multi-threshold. There are many good methods about denoising, we don't discuss it in this paper.

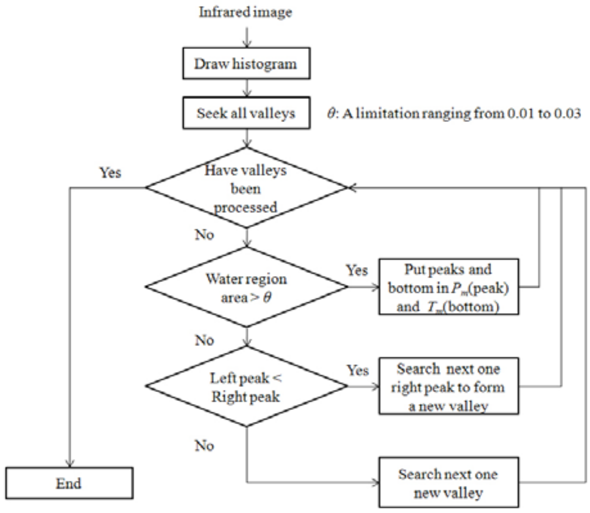


Fig. 1 Diagram of water region area

IV. Artificial Neural Network

Artificial neural network (ANN) has good advantage to estimate the relation between input and output when we could not know the relation of input and output, especially the relation is nonlinear. Generally speaking, ANN is divided into two parts. One is training, another is testing. During the training, we have to define training data and relational parameters. In the testing, we have to define testing data then get fused parameters. It has good ability to learn from examples and extract the statistical properties of the examples during the training procedure. Feature extraction is the important pre-procedure for ANN. In our case, we choice four feature, respectively, average intensity of visual image M_v, average intensity of infrared image M_i, average intensity of region in infrared image M_{ir} and visibility V_i to present as input of ANN. The following is our introduction of features.

The average intensity of visual image M_v:

$$M_v = \frac{1}{H \times W} \sum_{x=1}^{H} \sum_{y=1}^{W} f_v(x, y) \tag{4}$$

where f_v is visual gray image, H and W are height and width of visual image.

Generally speaking, it possible means the content of the image is shot in the daytime when M_v is larger. On the other hand, the content of the image is shot in the night. But it is initial assumption, not accurate.

The average intensity of M_i is defined as follow:

$$M_i = \frac{1}{H \times W} \sum_{x=1}^{H} \sum_{y=1}^{W} f_i(x, y) \tag{5}$$

where f_i is infrared image, H and W are height and width of visual image.

Generally speaking, it possible means the content of the image was shot in the daytime when M_i is larger. On the other hand, the content of the image was shot in the night. If we consider M_v and M_i to assume the shot time, then we can do more assume that the image was shot in the daytime or night when M_v and M_i both are larger or smaller respectively. If M_i is larger and M_v is smaller then we can suppose that the highlight of infrared image could be useful information for us. If M_i is smaller and M_v is larger then we can suppose that it could be no useful information in the infrared image to add to visual image.

The average intensity of region in M_{ir} is defined as follow:

$$M_{ir} = \frac{1}{\sum DB_i} \sum_{(x,y) \in DB_i} f_i(x, y) \tag{6}$$

Where DB_i is the segmented region of infrared image. $\sum DB_i$ is the total number of pixels in the region DB_i.

We can suppose more accurately if we have above three features. For example, the biggest M_{ir} could be not the information what we want if M_i and M_v both are larger. But we have to care about other regions which could be useful information in the same state.

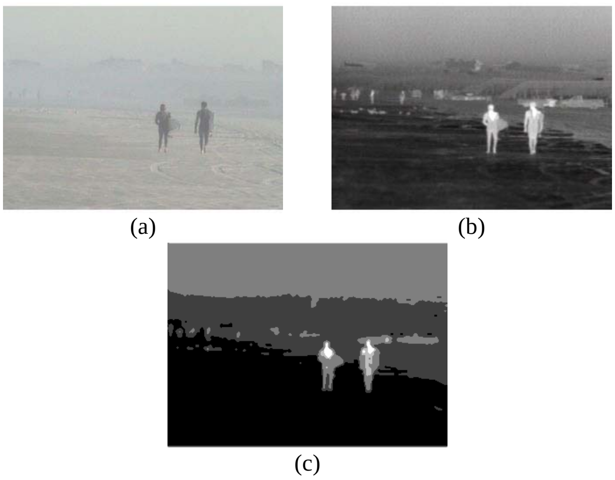


Fig. 2 (a) Visual image (b) Infrared image (c) Segmented infrared image

TABLE I
INTEGRATION OF FEATURES OF EACH REGION FROM SEGMENTED IMAGE

Regions	Features			
	M_v	M_i	M_{ir}	V_i
Region1	181.0579	99.1740	44.2328	0.0164
Region2	181.0579	99.1740	101.02929	0.00655
Region3	181.0579	99.1740	157.25777	0.00401
Region4	181.0579	99.1740	206.11231	0.00074
Region5	181.0579	99.1740	250.57458	0.00104

The visibility V_i is defined as follow:

$$V_i = \frac{1}{\sum DB_i} \sum_{(x,y) \in DB_i} \left(\frac{1}{M_{ir}}\right)^{\alpha} \cdot \frac{| f_i(x, y) - M_{ir} |}{M_{ir}} \tag{7}$$

where DB_i is the segmented region of infrared image. $\sum DB_i$ is the total number of pixels in the region DB_i. M_{ir} is the average intensity of region in infrared image. The α is a visual constant ranging from 0.6 to 0.7. region in infrared image.

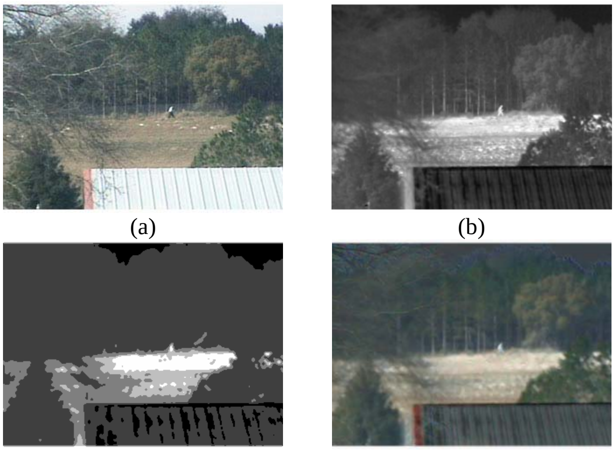
Due to human's eyes sense not only intensity of image but contrast of image. This feature is inspired from the human visual system (HVS), Huang and Li have given its definition. We consider it as the distribution of the region to represent the regions.

We can start to define the training data and testing data when getting the four features. The Fig. 2 is one of our training data, they are visual image, infrared image and segmented infrared image respectively from left to right. We only segment the infrared image here. And we use color depth to represent each region. There are five level to represent five region. Table I is the integration of the features of each region from segmented infrared image. Each region from 1 to 5 is the color level from deep to shallow respectively. One region has four features.

The fused parameters mentioned in section II are the output of ANN. We will adjust them by man-made when define the training data. And we use back-propagation (BP) to be training method of ANN. In practice, the fused parameters will be produced automatically by ANN. The details will be discussed in the following section.

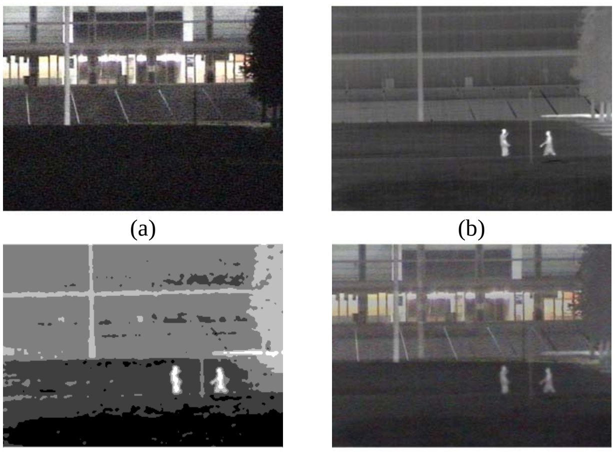
V. EXPERIMENTAL RESULTS

We use more training data like Table I and start training our neural network until weightings converging. Then we can use the trained weightings to get fused parameters when new couple of visual and infrared image in our module. We present two of testing data as following figures. In Fig. 3 and Fig. 4, (a) is visible image, (b) is infrared image, (c) segmentation image (d) fused image (Fig. 3 and Fig. 4). We also compare our method with traditional IHS fusion method in the Fig. 5. Obviously, our method shows not only good result of fused image but more colourful information are hold is better than traditional IHS fusion method.



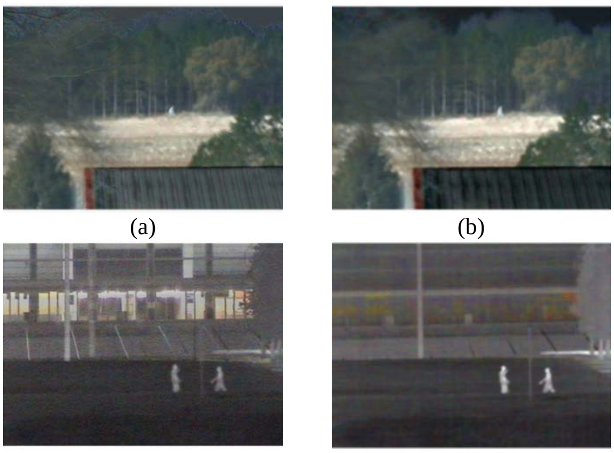
(a) (b) (c) (d)

Fig. 3 (a) Visible image (24bits level, size of 320×240) (b) Infrared image (256 level, size of 320×240) (c) Segmentation of infrared image (d) Fused image



(a) (b) (c) (d)

Fig. 4 (a) Visible image (24bits level, size of 320×240) (b) Infrared image (256 level, size of 320×240) (c) Segmentation of infrared image (d) Fused image



(a) (b) (c) (d)

Fig. 5 (a) Region-based image fused method by ANN (b) Traditional IHS fused method of testing image (c) Region-based image fused method by ANN (d) Traditional IHS fused method

VI. CONCLUSION

This paper describes an application of artificial neural network to multi-sensor image fusion problem. We improve traditional IHS-method and add concept of region-based into image fusion. The aim is that different regions can be used by different parameters in different state about time or weather. Due to the relation between environment and fused parameters is nonlinear. So, we adopt artificial neural network to solve this problem. On the other hand, the fused parameters will be estimated automatically render us to get adaptive appearance in different states. The architecture we proposed is not only can be useful for many applications but also adapted for many kinds of field. We successfully develop the architecture.

REFERENCE

[1] Z. Wang, D. Ziou, C. Armenakis, D. Li, and Q. Li, "A Comparative Analysis of Image Fusion Methods," *Geoscience and Remote Sensing*, vol. 43, no. 6, pp. 1391-1402, June 2006.

[2] J. G. Liu, "Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details," *Int. J. Remote Sensing*, vol. 21, no. 18, pp. 3461-3472, 2000.

[3] M. Li, W. Cai, and Z. Tan, "A region-based multi-sensor image fusion scheme using pulse-coupled neural network," *Pattern Recognition Letters*, vol. 27, pp. 1948-1956, 2006.

[4] L. J. Guo and J. M. Moore, "Pixel block intensity modulation: adding spatial detail to TM band 6 thermal imagery," *Int. J. Remote Sensing.*, vol. 19, no. 13, pp. 2477-2491, 1988.

[5] P. S. Chavez and J. A. Bowell, "Comparison of the spectral information content of Landsat thematic mapper and SPOT for three different sites in the Phoenix, Arizona region," *Photogramm. Eng. Remote Sensing.*, vol. 54, no.12, pp. 1699-1708, 1988.

[6] A. R. Gillespie, A. B. Kahle, and R. E. Walker, "Color enhancement of highly Correlated images- . Channel ratio and chromaticity transformation Techniques," *Remote Sensing Environment*, vol. 22, pp. 343-365, 1987.

[7] J. Sun, J. Li and J. Li, "Multi-source remote sensing image fusion," *INT. J. Remote Sensing*, vol. 2, no. 1, pp. 323-328, Feb. 1998.

[8] W. J. Carper, T. M. Lillesand, and R. W. Kiefer, "The use of Intensity-Hue-Saturation transformation for merging SPOT panchromatic and multispectral image data," *Photogramm. Eng. Remote Sensing*, vol. 56, no. 4, pp. 459-467, 1990.

[9] K. Edwards and P. A. Davis, "The use of Intensity-Hue-Saturation transformation for producing color shaded-relief images," *Photogramm. Eng. Remote Sensing*, vol. 60, no. 11, pp. 1369-1374, 1994.

[10] E. M. Schetselaar, "Fusion by the IHS transform: Should we use cylindrical or Spherical coordinates?," *Int. J. Remote Sensing*, vol. 19, no. 4, pp. 759-765, 1998.

[11] J. Zhou, D. L. Civco, and J. A. Silander, "A wavelet transform method to merge Landsat TM and SPOT panchromatic data," *Int. J. Remote Sensing*, vol. 19, no. 4, pp. 743-757, 1998.

[12] S. Li, J. T. Kwok, Y. Wang, "Multifocus image fusion using artificial neural networks," *Pattern Recognition Letters*, vol. 23, pp. 985-997, 2002.

[13] Q. Yuan, C.Y. Dong, Q. Wang, "An adaptive fusion algorithm based on ANFIS for radar/infrared system," *Expert Systems with Applications*, vol. 36, pp. 111-120, 2009.