Adaptive fusion of infrared and visible images in dynamic scene

Guang Yang[a], Yafeng Yin[a], Hong Man[a] and Sachi Desai[b]

[a]Stevens Institute of Technology, Castle Point on Hudson, Hoboken, NJ USA 07030
[b]US Army RDECOM, Picatinny Arsenal, NJ USA 07806

ABSTRACT

Visible and infrared image fusion has been widely employed in many surveillance and military applications. A variety of image fusion techniques including PCA, wavelet, curvelet and HSV have been proposed in recent years to improve human visual perception for object detection. One of the main challenges for visible and infrared image fusion is how to automatically determine the optimal fusion strategy for different input scenes with an acceptable computational cost.

In this paper, we proposed a fast and adaptive feature selection based image fusion method to obtain a high contrast image from visible and infrared sensors for targets detection. At first, fuzzy c-means clustering is applied on the infrared image to highlight possible hotspot regions, which will be considered as potential targets' locations. After that, the region surrounding the target area is segmented as the background regions. Then image fusion is locally applied on the selected target and background regions by computing different linear combination of color components from registered visible and infrared images. After obtaining different fused images, histogram distributions are computed on these local fusion images as the fusion feature set. The variance ratio which is based on Linear Discriminative Analysis (LDA) measure is employed to sort the feature set and the most discriminative one is selected for the whole image fusion. As the feature selection is performed over time, the process will dynamically determine the most suitable feature for image fusion in different scenes. Experiment is conducted on the OSU Color-Thermal database, and TNO Human Factor dataset. The fusion results indicate that our proposed method achieved a competitive performance compared with other fusion algorithms at a relatively low computational cost.

Keywords: adaptive fusion, discriminative feature selection, dynamical scenes

1. INTRODUCTION

Image fusion is a popular and important part of data fusion that is aiming at actively combining multiple source images together to maximize meaningful information and reduce redundancy. The idea of multi-modality image fusion is to take advantages of sensors in different frequency bands. It has been widely utilized in many applications such as remote sensing imagery, medical image processing (CT-MRI), and surveillance and military applications (visible-IR). Many techniques, including principal component analysis (PCA)[1] , wavelet transform[2,3] and color space transformation[4] etc., have been well developed in recent 20 years. In most papers, the fusion methods usually consist of two parts, image decomposition and fusion rules.

However, in the state of the art, most fusion rules have been set for their specific scenarios in their papers. When the environment(background) changes at different scenes, the fusion rule perhaps is no longer the best solution anymore for a different scene. Therefore, we proposed an adaptive fusion algorithm which is choosing the most discriminative feature from a set of feature pool to determine the rule. The crucial part of this approach is to calculate features of linear combinations of components in source images with integer coefficients in the range of [-2, 2]. The combination coefficients represent the contributions of each channel in a feature. After that, variance ratio of log likelihood of histograms will be used for measuring all features. The most discriminative feature will be selected and its combination coefficient will be applied to entire fusion system.

Further author information: (Send correspondence to Guang Yang)
Guang Yang: E-mail: gyang1@stevens.edu, Yafeng Yin: E-mail: yyin1@stevens.edu, Telephone: 1 201 216 5621

In addition, to build a more reliable automatic fusion system, pre-detection method is used to choose less biased target candidates. We use Fuzzy c-means clustering method on infrared images to classify the intensities into several clusters. This enhancement makes a big contribution to the next feature selection step.

This paper is organized as follows. Section 2 will introduce the state of the art of image fusion and the related techniques to our fusion system. Section 3 will detail our adaptive fusion system with complete algorithms and procedures. Experimental results, evaluation, and comparisons in Section 4 will demonstrate the advantages of the discriminative feature selection fusion system. Section 5 will provide conclusions and discuss the future work.

2. RELATED WORK

Multi-sensor image fusion has been studied for decades, and many techniques have been developed, such as wavelet transform[2] , pyramid-based fusion, principal component analysis[5] , curvelet[6] , and RGB to IHS color transformation[4] etc. In many cases, images are decomposed into sub-regions or sub-bands, and then applied with different fusion rules to each set of sub-regions correspondingly.

Wavelet was initially used in the field of signal processing since 1989 by Mallat[7] . It has been widely used for decomposing images into sub-domains of coefficient sets. A typical wavelet fusion method takes maximum value at high frequency domain and mean values at low frequency domain of two registered source images. Qin et al.[3] selected local modulus of maximum at edge pixels and its sub-band neighboring pixels. They claimed that their method kept details from source images and suppressed noise. Zhang in[4] employed a statistical fusion rule of determining matching measure and saliency measure in each sub-band.

PCA can be used after wavelet or other decompositions, such as contourlet[8] or curvelet transformation[6, 9] . It usually uses fusion rules to choose the reconstruction coefficients for a sub-band or the complete image window, as seen in[8, 10, 11] .

Another popular fusion idea deals with image color components separately in IHS color space. Usually visible images are constructed in standard RGB space. However IHS color transformation has the advantage that separate spatial (I) and spectral (H, S) information from RGB true color space[1] . Firstly the original visible image is converted into IHS space and represented by I_{vi}, H_{vi} and S_{vi} as Intensity, Hue and Saturation as three components. Then, the infrared image which just has intensity information is fused with I_{vi} component, denoted as I_f. Statistical fusion rules and linear mapping by matching means and variance between I_{vi} and I_f, are applied here. Next I_{vi} is replaced with the enhanced intensity and finally transformed back to RGB representation[4] .

In our system, fuzzy c-means (FCM) clustering method has been adopted in preparing stage. Fuzzy clustering is a classic data partition process to divides data into groups with certain number of natural centers. FCM is a most widely used fuzzy clustering algorithm which was proposed by Bezdek[12] in 1981. FCM have been extended to segment histograms on color components[13] .

3. ADAPTIVE FUSION OF INFRARED AND VISIBLE LIGHT IMAGES

3.1 Algorithm Overview

The proposed fast and adaptive feature selection method aims at obtaining a high contrast image from visible and infrared sensors and consequently passing those more reliable candidates to later targets detection modules.

Preprocessing of visible and infrared image pairs is very helpful to make the system automatic, discriminative as well as adaptive. A low computational algorithm consisting with fuzzy c-means clustering and template matching method is introduced as a pre-screening algorithm. This preprocessing stage searches for possible target candidates at a very low cost, and passes them down to the next feature selection stage. Hence the feature selection algorithm will be applied just on small areas in the image. This approach can save a lot of calculation by avoiding calculating on a whole image at every combination. More over, the LDA feature selection attempts to select the most discriminative feature among all combinations. The target areas are the most representative and meaningful part of the image itself. In another words, if the target areas are effectively highlighted after image fusion, then this image fusion system achieves its purpose. Our preprocessing step ensures that the fusion system is not only running fast but also working on the key points of a image.

To adaptively fuse images from two modalities, we consider our image source as a multi-channel model. The IR image is a grayscale image in which the value of each pixel is a single sample of intensity information. Thus we have four components consisting of one channel from IR gray scale image and three RGB channels from a visible light image. A crucial part of this approach is to calculate features of linear combinations of the four components with integer coefficients in the range of [-2, 2][14] . The combination coefficients directly correspond to the contribution of each channel in features. After that, the variance ratio of log likelihood will measure those fusion results. And the most discriminative feature will be selected and applied its combination to entire fusion system.

3.2 Preprocessing in Fusion System

In our system, given any pair of well aligned visible and infrared images, the first step is to apply the Fuzzy c-means (FCM) clustering to the IR image. In this step, the histogram is clustered into c classes. The class with largest center will give the best contrast of targets and background within this infrared image, as shown in Figure 1. And target candidates are detected by a spatial template.

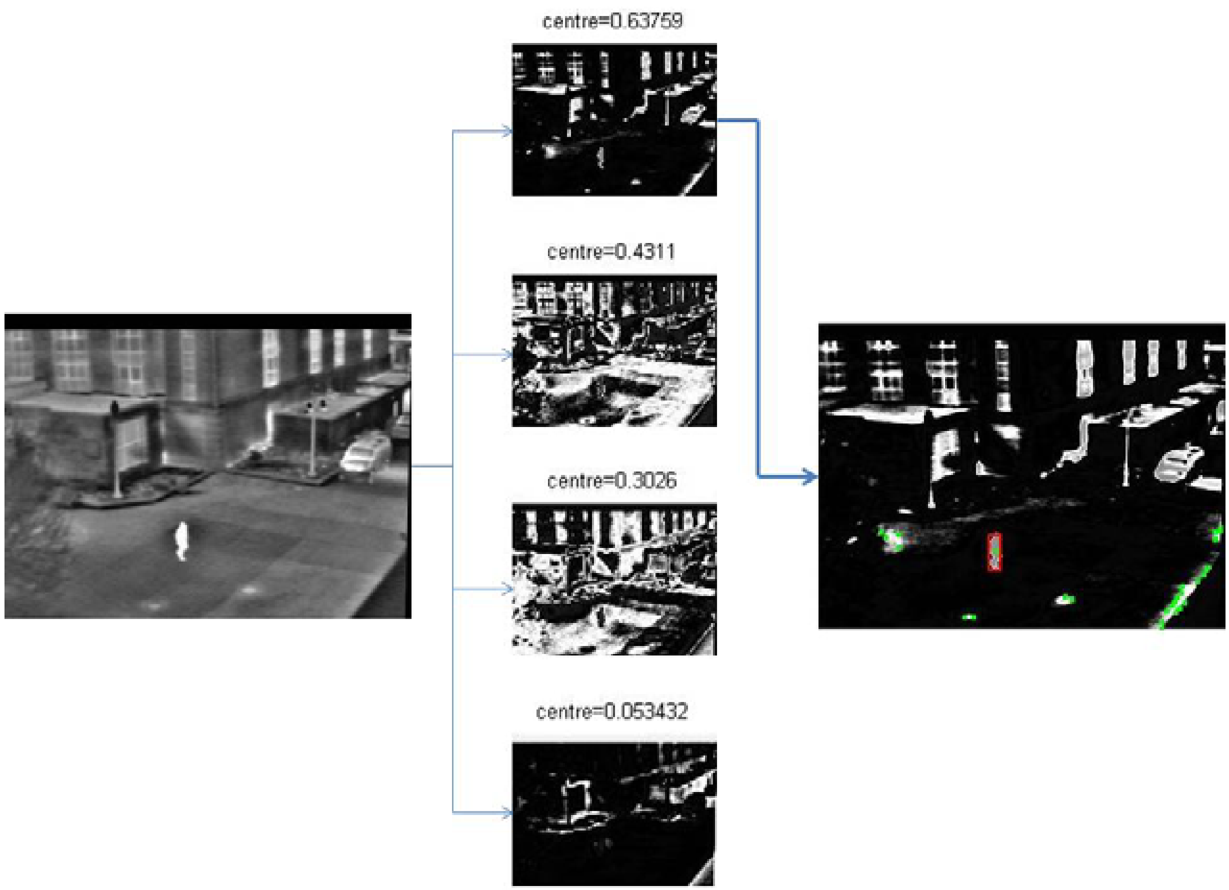


Figure 1: The results of FCM clustering shown on the right image.

In Fig. 1, the first image is the original input IR image and the middle column shows the fuzzy clustering in 4 classes. The last image gives the candidate target region in a red box after filtering out other possible areas in green dots.

3.3 Adaptive Image Fusion Procedure

After determining the region of interest in the IR image, adaptive image fusion is applied to both EO and IR image pairs. Our adaptive fusion algorithm is inspired by the feature selection method proposed by Collins and

Liu[14] . The linear combination of R, G and B channels are utilized in[14] to determine the most discriminative feature and they used it for tracking the object in the scene. While in our system, we combine the infrared channel with R, G, B channel together for IR and visible image fusion. Different with the feature selection method in,[14] we fuse information from multiple spectrum instead of color space. Suppose c_i is channel i, $i = 1, ...4$ stands for R, G, B, and IR channels, then our fused image can be described as linear combination of different channel:

$$F = \sum_{i=1}^{4} \omega_i c_i,\ w_i = -2, \ldots, 2 \tag{1}$$

Where F stands for the fused image from four channels, and its pixel value is normalized to $0 \sim 255$. As the linear combination of four channels using integer coefficients between -2 and 2, the total number in the coefficient set is 5^4. There are some trivial combinations need to be removed, such as [1, 1, 1, 1] and [2, 2, 2, 2]. By pruning all the redundant coefficients, the set size is cut down to 522 . The candidate features will be like R+G+B+IR, R+2G-B-2IR and etc.

As there are many linear combinations in the fusion set, we only apply the fusion locally in the region of interest instead of to the entire image to select the best fusion rules. The regions of interest, which are the object-background regions in the preprocessing step, are segmented from corresponding EO and IR images for fusion separately. To determine the most discriminative linear combination, we adopted the empirical discriminability measure from.[14] The object region and its surrounding background, are treated as two separate classes. For any fusion set in set F, the normalized histograms of the object class and the background class is calculated. A histogram indicates the frequency of pixel values in certain intervals (here we use 32 histogram bins). And the normalized histogram represents the discrete probability density. Here $p(i)$ is the density of the object class and $q(i)$ is one of the background class.

The log likelihood ratio[14] maps the similarity of the object/background distribution into positive and negative values. The positive values for colors can be identified as the object; on the contrary, the negative values are characteristic to the background.

$$L(i) = log \frac{max\{p(i), \delta\}}{max\{q(i), \delta\}} \tag{2}$$

Where δ, is a very small value set as 0.001 in order to prevent dividing by zero or taking the log of zero.

Finally, the variance ratio of likelihood is calculated for the purpose of quantifying the separability of object and background. The variance ratio of the log likelihood function[14] can be defined as:

$$VR(L; p, q) \equiv \frac{var(L; (p + q)/2)}{var(L; p) + var(L; q)} \tag{3}$$

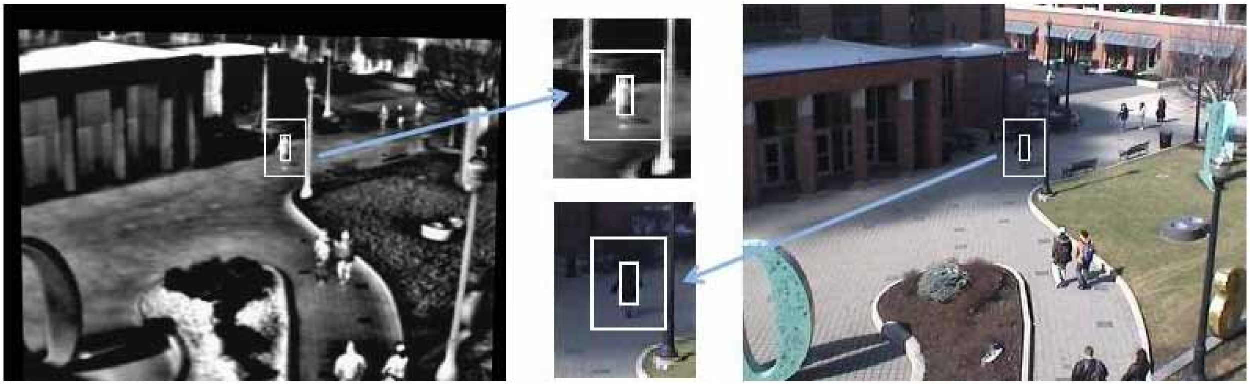
In Equation (3), the variance of $L(i)$ is defined based on $var(x) = Ex^2 - (Ex)^2$. And for any given discrete probability density function $a(i)$, the definition of the variance of $L(i)$ with respect to $a(i)$ is as:

$$var(L; a) = \sum_i [a(i)L(i)]^2 - [\sum_i a(i)L(i)]^2 \tag{4}$$

	Visible	IR	PCA	IHS	Our method
Variance Ratio	0.9125	0.9237	0.5971	0.6027	**1.0291**

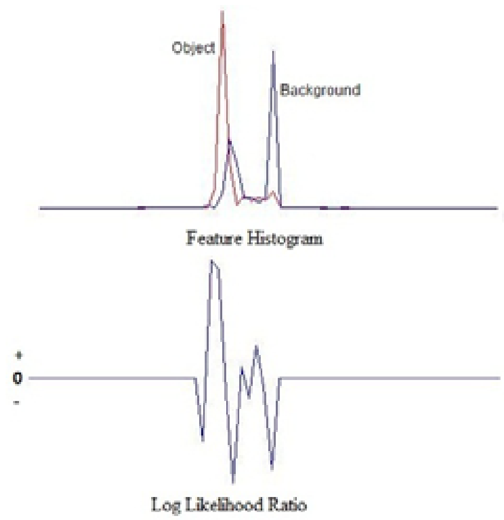
Table 1: Comparisons of PCA and IHS fusion method on same source image pair. The Variance ratio of log likelihood is the feature score that used in our proposed algorithm, according to Equation 3.

Till now, we have applied linear combination fusion scheme and LDA evaluation to the object and its near surrounding regions. And they are applied to a relatively small areas inside the images. After we find out the



(a) Feature regions in source images

(b) Fused image



(c)

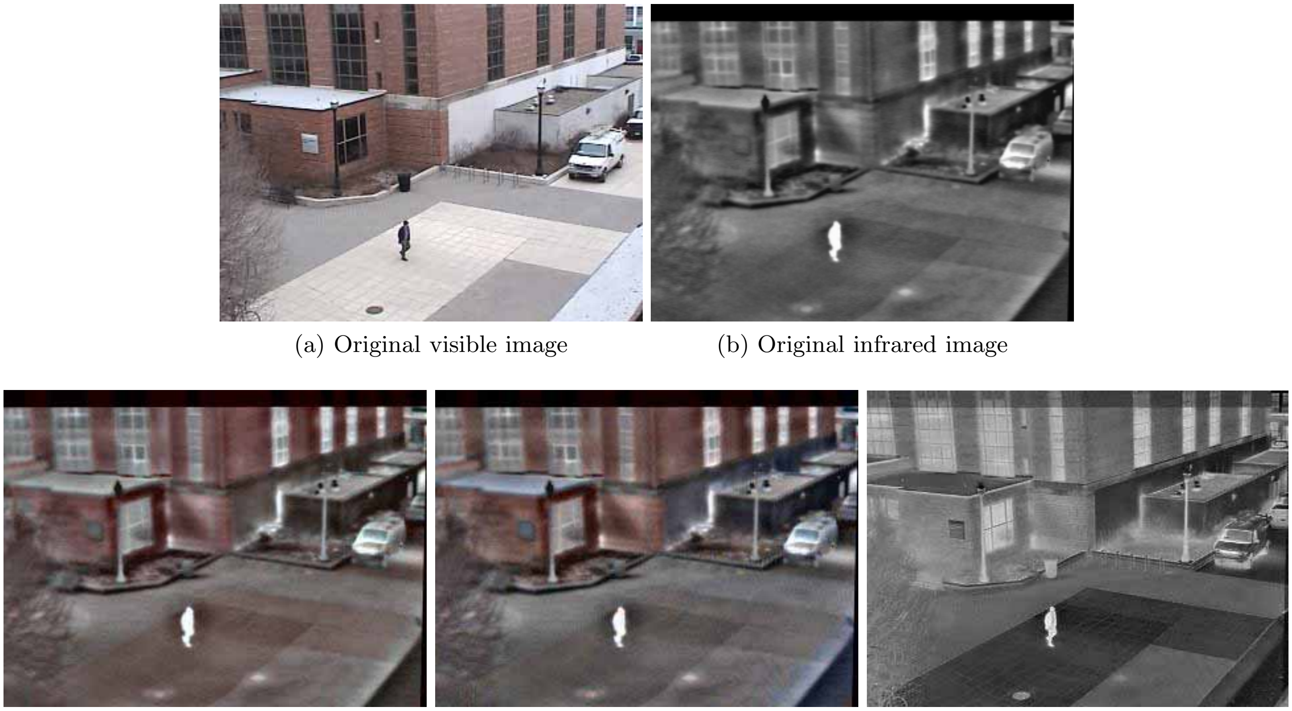
Figure 2: Features are calculated within the region of target and background from the source image. After ranking all of the variance ratio in the feature pool, the selected combination of components is -R+2G+2B-IR, shown in (b). (c) illustrates the histograms of feature regions and the log likelihood ratio in fused image (b).

most discriminative feature among more than five hundreds candidates, the step is to use it as the fusion rule to this image pair. Therefore, the object has its feature best separated from next background and hence help the upcoming detection or tracking work.

4. EXPERIMENTS AND COMPARISONS

We implemented our method by MATLAB running on a laptop of 2.40GHz Intel Core2 Duo PC with 2GB memory. The test images in our experiments are from OSU Color-Thermal database[15] and TNO Human Factors[16] . And we will compare of our method with other fusion approaches, namely PCA, IHS color space transform and fusion by wavelet transform.

Figure. 3 shows a fusion result by our algorithm. The target area is passed by the result shown in red rectangular ($9 \times 28 pixels$), and we set its local backgrounds as 10 pixels for each side out of the box. Linear combinations of four channels and calculations on feature discriminability are then taken place on this target-and-background region. Finally the select feature vector is $[-2, 0, -1, 2]$, which is applied as the fusion rule to the source image pair. Shown in Figure. 3, the final fusion result is the dot product of feature vector and channel components vector: $[-2, 0, -1, 2] \bullet [R, G, B, IR] = -2R - B + 2IR$.



(a) Original visible image (b) Original infrared image

(c) Fused image by PCA method (d) Fused image by IHS transform (e) Fused image by our method

Figure 3: Fusion results comparison

Another test is based on a pair of grayscale images supplied by the TNO Human Factors.[16] The visible image in this test just has one component and denoted as EO component. While implementing the proposed algorithm to this image pair, the trick is that, taking three EO components together as RGB components of an actual color image, and each component still has a coefficient range from -2 to 2. Thus the largest coefficient of EO, could be 6: $F \equiv \{\alpha EO + \beta IR, \alpha \in [-6, 6], \beta \in [-2, 2]\}$. After processing, selected feature is $-3EO + 2IR$ for this pair of source images, as shown in Figure 4. We also compare it with the fusion image by wavelet with 8-level decomposition. Table 2 and Table 3 give the comparisons of the variance ratio of the target-and-background regions in source and fused images.

	Visible	IR	PCA	IHS	Our method
Variance Ratio	0.6584	1.3110	0.5341	0.5364	**1.4195**

Table 2: Evaluation comparisons of Figure 3

	Visible	IR	PCA	IHS	Our method
Variance Ratio	1.7209	0.5334	0.5341	0.6123	**1.8153**

Table 3: Evaluations comparisons of Figure 4

(a) Original visible image (b) Original infrared image

(c) Fused image bu wavelet method (d) Fused image by our method

Figure 4: Fusion result on grayscale image pair

5. CONCLUSION

In this paper, we have introduced an fusion system for infrared and visible images. Based on selecting the most discriminative linear combination in the feature set, this fusion scheme gives a promising fusion result. As we discussed previously, our fusion system can detect the target and let the fusion scheme focus on a small region, consequently looking for the best separation of the object and its background.

REFERENCES

[1] Pohl, C. and Van Genderen, J. L., "Review article multisensor image fusion in remote sensing: concepts, methods and applications," *International Journal of Remote Sensing* **19**, 823–854 (March 1998).

[2] Mitra, S., Manjunath, B., and Li, H., "Multisensor image fusion using the wavelet transform," in [*ICIP94*], I: 51–55 (1994).

[3] Qin, Q., Xu, T., Xiao, M., and Ni, G., "A novel algorithm of target pseudo-color fusion based on image features," 1 –5 (oct. 2008).

[4] Zhang, X., "Infrared and color visible image sequence fusion based on statistical model and image enhancement," in [*Advanced Computer Theory and Engineering, 2008. ICACTE '08. International Conference on*], 934 –937 (dec. 2008).

[5] Shlens, J., "A tutorial on principal component analysis." http://www.snl.salk.edu/ shlens/pub/notes/pca.pdf (December 2005).

[6] Sun, F., Li, S., and Yang, B., "A new color image fusion method for visible and infrared images," *Robotics and Biomimetics, 2008. ROBIO 2007. IEEE International Conference on* , 2043 –2048 (dec. 2007).

[7] Mallat, S., "A theory for multiresolution signal decomposition: the wavelet representation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **11**, 674 –693 (jul 1989).

[8] Zhang, X. and Liu, X., "Pixel level image fusion scheme based on accumulated gradient and pca transform," in [*Computer-Aided Industrial Design and Conceptual Design, 2008. CAID/CD 2008. 9th International Conference on*], 594 –598 (nov. 2008).

[9] Yang, B., Sun, F., and Li, S., "Region-based color fusion method for visible and ir image sequences," in [*Pattern Recognition, 2008. CCPR '08. Chinese Conference on*], 1 –6 (oct. 2008).

[10] Qiu, Y., Wu, J., Huang, H., Wu, H., Liu, J., and Tian, J., "Multi-sensor image data fusion based on pixel-level weights of wavelet and the pca transform," in [*Mechatronics and Automation, 2005 IEEE International Conference*], 653 – 658 Vol. 2 (july-1 aug. 2005).

[11] Metwalli, M., Nasr, A., Farag Allah, O., and El-Rabaie, S., "Image fusion based on principal component analysis and high-pass filter," in [*Computer Engineering Systems, 2009. ICCES 2009. International Conference on*], 63 –70 (14-16 2009).

[12] Bezdek, J., [*Pattern Recognition with Fuzzy Objective Function Algorithms*], Plenum (1981).

[13] Lim, Y. W. and Lee, S. U., "On the color image segmentation algorithm based on the thresholding and the fuzzy c-means techniques," *Pattern Recogn.* **23**(9), 935–952 (1990).

[14] Collins, R. T., Liu, Y., and Leordeanu, M., "Online selection of discriminative tracking features," *IEEE Trans. Pattern Anal. Mach. Intell.* **27**(10), 1631–1643 (2005).

[15] Davis, J. and Sharma, V., "Background-subtraction using contour-based fusion of thermal and visible imagery," *Computer Vision and Image Understanding* **106**(2-3), 162–182 (2007).

[16] Toet, A., Ijspeert, J., Waxman, A., and Aguilar, M., "Fusion of visible and thermal imagery improves situational awareness," *Displays* **18**, 85–95 (December 1997).