Multi-Sensor Fusion of Electro-Optic and Infrared Signals for High Resolution Visible Images: Part II

Xiaopeng Huang, Ravi Netravali*, Hong Man and Victor Lawrence
Dept. of Electrical and Computer Engineering
Stevens Institute of Technology
Hoboken, New Jersey 07030, United States
{xhuang3, hong.man, victor.lawrence}@stevens.edu

*Dept. of Electrical Engineering
Columbia University
New York City, New York 10027, United States
ran2290@gmail.com

***Abstract*—Electro-Optic (EO) image sensors exhibit the properties of high resolution and low noise level, but they cannot reflect information about the temperature of objects and do not work in dark environments. On the other hand, infrared (IR) image sensors exhibit the properties of low resolution and high noise level, but IR images can reflect information about the temperature of objects all the time. Therefore, in this paper, we propose a novel framework to enhance the resolution of EO images using the information (e.g., temperature) from IR images, which helps distinguish temperature variation of objects in the daytime via high-resolution EO images. The proposed novel framework involves four main steps: (1) select target objects with temperature variation in original IR images; (2) fuse original RGB color (EO) images and IR images based on image fusion algorithms; (3) blend the fused images of target objects in proportion with original gray-scale EO images; (4) superimpose the target objects' temperature information onto original EO images via the modified NTSC color space transformation. Therein, the image fusion step will be conducted by the quantitative (*Yang et al.* proposed adaptive multi-sensor fusion algorithm) approach in this part. Revealing temperature information in EO images for the first time is the most significant contribution of this paper. Simulation results will show the transformed EO images with the targets' temperature information.**

***Index Terms*—Component, formatting, style, styling, insert.** *(key words)*

I. Introduction

Multi-sensor image fusion is a widely employed technique that takes advantage of sensors at different frequency bands and aims at actively combining multi-source images together to maximize meaningful information and reduce redundancy. For example, Electro-Optic (EO) image sensors exhibit the properties of high resolution and low noise level in the daytime, but they cannot reflect information about the temperature of objects and do not work in dark environments. In contrast, infrared (IR) image sensors exhibit the properties of low resolution and high noise level, but IR images can reflect information about the temperature of objects all the time. Therefore, how to reflect information about objects' temperature (e.g., abnormal temperature or temperature variation) onto EO images with help of IR images is a worthy topic that has not been resolved yet. We have completed research on multi-sensor fusion of EO and IR signals for high-resolution night images [1-3], which results in high- resolution, clear-edged IR images of any object and helps distinguish objects at night.

In the past decades, scientists have developed many efficient image fusion algorithms, such as the expectation maximization (EM) fusing algorithm [4], the discrete wavelet transform (DWT) fusing algorithm [5], the Laplacian pyramid fusing algorithm [6], and the principal component analysis (PCA) [7], etc. However, most of the existing fusion algorithms are set for their specific scenarios, so these algorithms may not exhibit the best performance when environments change. To the best of our knowledge, little literature focuses on developing adaptive fusion algorithms [8-10]. Yang et al. [9] proposed an adaptive fusion algorithm, which selects the most discriminative feature from a set of features to determine the fusion rule. The proposed algorithm has been shown to adaptively achieve the best fusion results for different selected target objects.

RGB color space is one of the most popular formats for color images. However, there are other color spaces whose use in some applications may be more convenient and/or meaningful than RGB. These spaces are transformations of the RGB space and include the NTSC, YCbCr, HSV, CMY, CMYK, and HIS color spaces [11]. Therein, the NTSC color system is used in analog television. One of the main advantages of this format is that gray-scale information is separate from color data. Hence, NTSC color space

978-1-4673-0831-1/12/$31.00 ©2012 IEEE

provides a convenient approach to superimposing objects' temperature information onto color images.

In this paper, we propose a novel framework to reveal information about the temperature (e.g., abnormal temperature or temperature variation) of objects in EO images based on image fusion algorithms and modified NTSC color space transformations. The proposed framework will help distinguish and monitor the temperature variation of objects in the daytime via high-resolution EO images rather than traditional IR images. The proposed novel framework involves four main steps: (1) select target objects with temperature variation in original IR images; (2) fuse original RGB color (EO) images and IR images based on image fusion algorithms; (3) blend the fused images of target objects in proportion with original gray-scale EO images; (4) superimpose the target objects' temperature information onto original EO images via the modified NTSC color space transformation. Therein, the image fusion step will be conducted by the quantitative (*Yang et al.* proposed adaptive fusion algorithm) approach in this part. Revealing temperature information in EO images for the first time is the most significant contribution of our proposed framework. Simulation results will show the transformed EO images with the targets' temperature information.

The following questions will be addressed in this paper: (1) Are there some things we can see now that we could not see before by utilizing our framework? (2) In the quantitative approach, why do we adopt an adaptive multi-sensor fusion algorithm rather than conventional low-complexity fusion algorithms (e.g., PCA)? (3) Can our framework adaptively detect target objects? Is it possible for this framework to deal with multiple target objects simultaneously? If so, how should we select the best fusion rule? (4) What are the main applications of our novel framework? Can our proposed framework be deployed in any real-time image processing systems?

The remainder of this paper is organized as follows. Section II presents the quantitatively adaptive fusion algorithm-based image fusion process. Section III includes the modified NTSC color space transformation and simulation results. Finally, Section IV concludes this paper.

II. Adaptive Fusion Algorithm-Based Image Fusion Procedure

Target object selection is a critical step (Step 1) in our proposed framework. Figure 1 shows two original image pairs. Comparing EO and IR images in Pair 1, we find out that the temperature of the left and front part (engine) of the car is higher (temperature variation) than that of other parts, while we cannot see any difference in the corresponding EO image. Therefore, we select the engine part of this car as our target region shown in a red rectangle (186×115 pixels) for image Pair 1. Similarly, comparing EO and IR images in Pair 2, we find out that the temperature of two green color objects (transformers) is high, while we cannot see any temperature difference in the EO image. Therefore, we select the two transformers, shown in a red rectangle (84×23 pixels), as our target objects for image Pair 2.

(a)

(b)

Figure 1. (a) Original image Pair 1; (b) Original image Pair 2

We adopt the *Yang et al.* proposed adaptive image fusion algorithm to complete the quantitative image fusion procedure (Step 2). After determining the target objects in IR images, we apply this algorithm to fuse partially registered (e.g., perfect registration for the target object) EO and IR image pairs. Specifically, we assume four-channel linear combinations of R, G, B (visible channels) and IR (infrared channel) for EO and IR image fusion. Supposing γ_i is the channel i , where $i = 1, \ldots, 4$ represents R, G, B, and IR channels, then the fused image can be described as a linear combination of four different channels:

$$F = \sum_{i=1}^{4} \omega_i \gamma_i, \quad \omega_i = -2, \ldots, 2 \qquad (1)$$

where F represents the four-channel-criterion fused image, and its pixel value is normalized to 0~255 (unit 8). The integer coefficients ω_i are in the range of $[-2, 2]$. In other words, we could choose the following set of feature candidates:

$$\mathcal{F} \equiv \{\omega_1 R + \omega_2 G + \omega_3 B + \omega_4 IR \mid \omega_i \in [-2, -1, 0, 1, 2]\} \qquad (2)$$

Since the integer coefficients of four-channel linear combinations vary between -2 to 2, the total number of the coefficient set is 5^4. By pruning of all the redundant coefficients, such as [-1, -1, -1, -1], [2, 2, 2, 2] and [0, 0, 0, 0], the size of candidates is cut down to 544. The acceptable candidate features should be like R+G-2B+IR, etc.

How to select the best fusion rule from many linear combination candidates in the fusion set is the most critical thing of the proposed algorithm. Here, we first focus on the selected target object and its surrounding background rather than entire images to determine the best fusion rules, and then we apply the determined fusion rule to entire images and complete the image fusion process. We adopt the empirical discriminability measure form [12] to determine the most discriminative linear combination. The target object and its surrounding background are divided into two separate classes. For any fusion candidates in $\mathcal{F}$, the normalized histograms of both the object class and the background class are calculated. The histogram of length 2^b, where b is the number of bits of resolution (we set $b = 5$ in this paper), indicates the frequency of pixel values in certain intervals.

For each feature candidates, we form a normalized histogram to represent the empirical discrete probability distribution, where $p(i)$ is for the object and $q(i)$ is for the background. We create a new "tuned" feature tailored to better discriminate between object and its surrounding background. The tuned feature is formed as the log likelihood ratio [12], which maps object/background distributions into positive values for colors distinctive to the object and negative for colors associated with its surrounding background. The log likelihood ratio of a feature value i is given by [12]

$$L(i) = \log \frac{\max\{p(i), \delta\}}{\max\{q(i), \delta\}} \qquad (3)$$

where δ is a small value set as 0.001 to prevent dividing by zero or taking the log of zero. By using the equality $\mathrm{var}(x) = Ex^2 - (Ex)^2$, we compute the variance of L (i) with respect to object class distribution p(i) as [12]

$$\begin{aligned} \mathrm{var}(L; p) &= E[L^2(i)] - (E[L(i)])^2 \\ &= \textstyle\sum_i p(i) L^2(i) - [\sum_i p(i) L(i)]^2 \end{aligned} \qquad (4)$$

Similarly, the background class distribution $q(i)$ can be written as

$$\begin{aligned} \mathrm{var}(L; q) &= E[L^2(i)] - (E[L(i)])^2 \\ &= \textstyle\sum_i q(i) L^2(i) - [\sum_i q(i) L(i)]^2 \end{aligned} \qquad (5)$$

Therefore, the variance ratio of the log likelihood function can be defined as [12]

$$VR(L; p, q) \equiv \frac{\mathrm{var}(L; \frac{p+q}{2})}{[\mathrm{var}(L; p) + \mathrm{var}(L; q)]} \qquad (7)$$

which is the total variance of L over both object and background class distributions, divided by the sum of the within class variances of L for object and background treated separately.

By far, we could apply the adopted linear combination scheme and the linear discriminative analysis (LDA) evaluation to the target object and it's near surrounding region. After we find out the most discriminative feature among more than five hundred feature candidates, the next step is to use the determined fusion rule to the entire image pair.

In image Pair 1, the red rectangular (186×115 pixels) area is chosen as our target, and we set its local background as 40 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. Similarly, in image Pair 2, the red rectangular (84×23 pixels) area is chosen as our target, and we set its local background as 10 pixels for each side out of the box. The linear combination vector of each image pair with the most discriminative feature (maximum variance ratio in Eq. (7)) is selected from more than five hundred feature candidates. Then, this selected fusion rule will be applied to the entire source image pair. Figure 2 shows two fusion results obtained by utilizing the adaptive image fusion algorithm. Here, we only completed partial registration (perfect registration for target objects) for these two image pairs.

(a)

(b)

Figure 2 (a) Adaptively fused result of the image Pair 1; (b) Adaptively fused result of the image Pair 2

III. Modified NTSC Color Space Transformation

How to superimpose the temperature variation onto color (EO) images is a challenge of our proposed framework (Steps 3 and 4). The NTSC color system is used in analog television. One of the main advantages of this format is that gray-scale information is independent of color data. In the NTSC format, image data consists of three components: luminance (Y), hue (I), and saturation (Q), where the choice of the letters of YIQ is conventional. The luminance component represents gray-scale information of an EO signal, and the other two components represent the color information of an EO signal. Conventionally, the YIQ components are obtained from the RGB components of an EO image using the linear transformation [11]:

$$\begin{bmatrix}\mathbf{Y}\\\mathbf{I}\\\mathbf{Q}\end{bmatrix}=\begin{bmatrix}\mathbf{0.229} & \mathbf{0.587} & \mathbf{0.114}\\\mathbf{0.596} & \mathbf{-0.274} & \mathbf{-0.322}\\\mathbf{0.211} & \mathbf{-0.523} & \mathbf{0.312}\end{bmatrix}\begin{bmatrix}\mathbf{R}\\\mathbf{G}\\\mathbf{B}\end{bmatrix} \tag{8}$$

Similarly, the RGB components are obtained from the YIQ components using the linear transformation [11]:

$$\begin{bmatrix}R\\G\\B\end{bmatrix}=\begin{bmatrix}1.000 & 0.956 & 0.621\\1.000 & -0.272 & -0.647\\1.000 & -1.106 & 1.703\end{bmatrix}\begin{bmatrix}Y\\I\\Q\end{bmatrix} \tag{9}$$

In this paper, we propose a modified NTSC transformation scheme, where we change original Y (luminance) to $Y' = Y + \alpha\,(I)$, where I is the fused image of the target object, and α is the proportion of I and Y for the alpha blending process. The purpose of using α is to adjust the weight of temperature information in the final obtained color images. The concrete steps of this modified transformation can be described as follows:

- Transform the original RGB color image to an NTSC color image, and pick up the gray-scale component (image) Y.
- Blend the fused Image I of the target object with the gray-scale Image Y, which makes Y become $Y' = Y + \alpha\,(I)$. In this step, Image I should be superimposed onto corresponding position in Image Y.
- Transform the modified $Y'IQ$ image to an RGB color image, which shows information about the temperature of the target objects with high resolution.

Figure 3 (a) shows the results of original gray-scale images (Y component) of EO Image 1 and EO Image 2. Figure 3 (b) shows the results of modified gray-scale images (Y' component) of EO Image 1 and EO Image 2 based on quantitatively fused images. Figure 4 shows the results of transformed high-resolution EO images with information about the temperature of target objects.

(a)

(b)

Figure 3. (a) Y (luminance) component of EO Image 1 and EO Image 2; (b) Y' (luminance) component of EO Image 1 and EO Image 2 based on quantitatively fused images, where $\alpha = 0.2$ (Image 1), and $\alpha = 0.6$ (Image 2)

Figure 4. Transformed high-resolution EO Image 1 and EO Image 2 based on quantitatively fused images.

From the obtained results in Fig.4, we could clearly see information about the temperature of target objects in EO images for the first time. We are now able to answer the questions proposed in the introduction part as follows:

(1) By utilizing our novel framework, we could see temperature information about target objects in EO images for the first time. For example, we can see the temperature variation of the engine in EO Image 1 and the temperature of the transformers in EO Image 2.

(2) Most of the existing conventional image fusion algorithms (e.g., PCA) are set for their specific scenarios; when the background or the target object changes at different scenes, the fusion algorithms may not have the best performance anymore. However, adaptive fusion algorithms choose the most discriminative feature from a set of features to determine the fusion rule; on the other hand, the fusion rule will change with different target objects.

(3) In addition to the adopted adaptive fusion algorithm, we utilize other pre-processing algorithms, e.g., the Fuzzy c-means (FCM)) algorithm [13], which clusters the histogram of the original IR image into **c** classes and calculates the center value representing the contrast of targets and background for each class. This allows us to adaptively determine the target objects with larger center values (multi-target objects) or the largest center value (one target object) within the original IR image. For multiple target images, the best fusion rule will be the linear combination with the maximum variance ratio calculated from all targets' feature candidates.

(4) Our proposed novel framework, which utilizes the merits of both EO and IR image sensors and superimposes information about the temperature of objects onto high-resolution EO images, can be widely employed in various surveillance and military applications, especially in applications of recognizing objects using their temperature information. In addition, it is possible for our framework to be embedded in real-time image processing systems.

Although IR image sensors can reflect information about the temperature of objects, they have low resolution and sometimes cannot clearly distinguish objects. For example, in the original EO Image 1, we cannot see any information about the engine's temperature. However, the transformed EO Image 1 provides high-resolution information about the engine's temperature, which is higher than that of other parts (normal temperature) of the car. This temperature information is really helpful to judge the status of this car, such as whether the engine is on, or whether it has been stopped for a long time or a short time. Alternatively, in the original EO Image 2, we see two green color objects and we may not know exactly what they are; correspondingly, in the original IR Image 2, we see that the temperature of these two objects is high, but we cannot distinguish the color and edge of these two objects. However, in the transformed EO Image 2, we can see two high-resolution green-colored objects with high temperature. Therefore, we can recognize that these two objects are transformers.

IV. Conclusion

How to utilize the merits of both EO and IR image sensors in a complementary way to obtain high- resolution EO and IR images is a promising and challenging research topic. We have previously proposed a novel framework to fuse EO and IR signals for high-resolution night images, which helps distinguish objects at night via clear-edged IR images for the first time. However, no literature focuses on revealing information about the temperature in high-resolution EO images. Therefore, in this paper, we proposed a novel framework to superimpose information about the temperature of objects onto EO images via the modified NTSC color space transformation. By employing our proposed framework, we can clearly see the temperature variation of any objects in the transformed EO images, which can be regarded as a breakthrough of multi-sensor image fusion.

References

[1] Huang, X; Netravali, R; Man, H; Lawrence, V. Fusing Electro-Optic and Infrared Signals for High Resolution Night Images. *In Proceedings of SPIE Image Processing: Algorithms and Systems X*, **2012,** 8295

[2] Huang, X; Netravali, R; Man, H; Lawrence, V. Improved Fusing Infrared and Electro-Optic Signals for High-Resolution Night Images. *In Proceedings of SPIE Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XXIII*, **2012**, 8355

[3] Huang, X.; R. Netravali, R.; Man, H.; Lawrence, V. Multi-Sensor Fusion of Electro-Optic and Infrared Signals for High-Resolution Night Images. *Accepted by Sensors,* **2012**

[4] Liu G.; Sun S.; Jing Z. Image fusion based on an expectation maximization algorithm. *Opt. Eng.,* **2005**, volume *44*.

[5] Zheng Y.; Essock E.A.; Hansen B.C. Advanced discrete wavelet transform fusing algorithm and its optimization by using the metric of image quality index. *Opt. Eng.*, **2004**, *44*, pp. 1-12

[6] Zheng, Y. Multi-scale fusion algorithm comparisons: Pyramid, DWT and iterative DWT. *12th International Conference on fusion*, **2009**; pp. 1060–1067.

[7] Pohl, C.; Van Genderen, J. L. Review article multisensory image fusion in remote sensing: concepts, methods and applications. *International Journal of Remote Sensing,* **1998**, 19, pp.823-854

[8] Guo, Y.; Xie, M.; Yang, L. An adaptive image fusion method based on local statistical feature of wavelet coefficients. *International CNMT Symposium*, **2009**; pp.1-4

[9] Yang, G.; Yin, Y.; Man, H.; Desai, S. Adaptive fusion of infrared and visible images in dynamic scene. In *SPIE proceedings of Unmanned/Unattended Sensors and Sensor Networks VIII*, Sep. **2011**

[10] Tian, J.; Chen, Li. Adaptive multi-focus image fusion using a wavelet-based statistical sharpness measure. *ACM Journal of Signal Processing*, Sep. **2012**, 92, pp. 2137-2146

[11] Gonzalez, R. C.; Woods, R. E.; Eddins, S. L. Digital Image Processing Using Matlab, 2nd edition, Gatesmark Publishing, **2009**

[12] Collins, R. T.; Liu, Y.; Leordeanu, M. Online selection of discriminative tracking features. *IEEE Trans. Pattern Anal. Mach. Intell.*, **2005**, 27, pp. 1631-1643

[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. **1990**, 23, pp. 935-952