

Multi-source Image Fusion Technology in Different Fields of View

Si Tian^{1*}, Junju Zhang², Yihui Yuan², Benkang Chang²

¹Applied Science and Technology Research Institute, Ningbo Dahongying University, Ningbo 315175, Zhejiang, China;

²Institutes of Electronic Engineering and Optical Technology, Nanjing University of Science and Technology, Nanjing 210094, Jiangsu, China

*Corresponding author, e-mail: tststs1234@163.com

Abstract

There exist a lot of image systems of different features in different field of view (FOV) in our real life. Multi-source fusion technology in different FOV is able to promote the performance of image acquisition system without additional investment in hardware equipment. With regards to multi-image fusion technology in different FOV, in this research a registration method combined with affine transformation and linear interpolation is proposed, which analyzes pre- and post-fused images from the temporal and spatial perspective for comparison, evaluates and compares the fusion result so as to testify the applicable feasibility of the technology.

Keywords: different fields of view, image fusion, multi-source image, registration

Copyright © 2013 Universitas Ahmad Dahlan. All rights reserved.

1. Introduction

Pixel-level fusion [1], a rather widely applied fusion method, is to directly fuse pixel data of source image, making it possible to preserve more original information referring to FOV. It requires for higher source image registration and overlap ratio of viewing field. Generally speaking, under the condition when optical systems of source image are in the same FOV, it'll be much easier for acquired images to realize the registration, helping upgrade the quality of the fused image as well as guaranteeing its information quantity.

However, in practical application, there always appears a big contrast in the FOV of optical systems of certain multi-source acquisition device. Each source image like visible image, low-light image and infrared image has their own features, whose fused image represents details in FOV and highlight the object much better [2]. In this case, there again develop some different characteristics in the application between image registration algorithm and image fusion algorithm. Figure 1 is image fusion diagram in different FOV. Image A and B are acquired by two sets of optical systems from different FOV; wherein, we can see image B is only corresponding partially to image A; image C is the fusion result by image A and B. Actually, only the dotted part in image C can be considered to be the fused image, while the remaining is what image A contains.

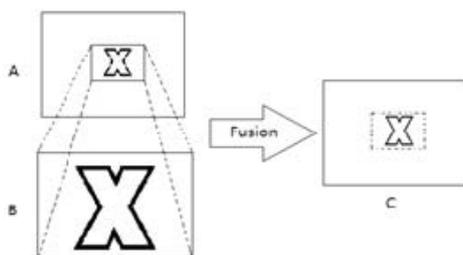


Figure 1. Image fusion in two different FOV

2. Analysis on Frequency Domain

In this part let's firstly take into consideration two different image sensors. Supposing these two image sensors generate modulus $f(x, y)$ and modulus $g(x, y)$ respectively; the former is construed by domain of the wide FOV size $M \times N$ and that of the small FOV size $m \times n$ in $g(x, y)$. A digitalized image can be used to present its grey-level distributed functions by the matrix in different places like:

$$f(x, y) \begin{matrix} x=0,1,2,\dots,M-1 \\ y=0,1,2,\dots,N-1 \end{matrix}, \quad g(x, y) \begin{matrix} x=0,1,2,\dots,m-1 \\ y=0,1,2,\dots,n-1 \end{matrix}$$

Thus, they're represented by the matrix as:

$$[f(x, y)] = \begin{bmatrix} f_{0,0} & f_{0,1} & \cdots & f_{0,N-1} \\ f_{1,0} & f_{1,1} & \cdots & f_{1,N-1} \\ \vdots & \vdots & \ddots & \vdots \\ f_{M-1,0} & f_{M-1,1} & \cdots & f_{M-1,N-1} \end{bmatrix} \quad (1)$$

$$[g(x, y)] = \begin{bmatrix} g_{0,0} & g_{0,1} & \cdots & g_{0,n-1} \\ g_{1,0} & g_{1,1} & \cdots & g_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ g_{m-1,0} & g_{m-1,1} & \cdots & g_{m-1,n-1} \end{bmatrix} \quad (2)$$

Discrete Fourier Transformer (DFT) of frequency domain (u, v) can be accordingly put as:

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(ax/M + by/N)} \quad (3)$$

$$G(u, v) = \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} g(x, y) e^{-j2\pi(ux/m + vy/n)} \quad (4)$$

As for general image, whose main body within two-dimension frequency domain is the low-frequency part, which demonstrates the general figure, featuring contour, contrast-ratio features etc. While those showing edges and waves are the high-frequency part. In Figure 1, the corresponding part in image A to that in image B is always what we're focusing on, as well as the place where frequency energy concentrates. When two images acquired by an image sensor start fusing, the small FOV image is wholly fused, but the non-fused part is preserved. Thus the main features of the image are enhanced, and its details are retained.

Now we discuss a simple weighted average fusion method. The fused image $r(x, y)$ can be shown as:

$$r(x, y) = \begin{cases} f(x, y) & (m < x \leq M, n < y \leq N) \\ af(x, y) + bg(x, y) & (x \leq m, y \leq n) \end{cases} \quad (5)$$

In the above function, $a+b=1$, and a, b are the weighting coefficients of two images. According to the addition theorem of DFT, fused image's DFT turns into:

$$R(u, v) = \begin{cases} F(u, v) & (m < x \leq M, n < y \leq N) \\ aF(u, v) + bG(u, v) & (x \leq m, y \leq n) \end{cases} \quad (6)$$

In formula (6), if $aF(u, v) + bG(u, v)$ is the main body for our attention, showing object's low-frequency components like general shape, featuring contour and contrast-ratio features, the quality of the whole fused image will be secured.

3. Registration Method

As to the multi-source images, since the principles and DPI of image formation are different, and the optical axis may not be parallel, the position of the objects in different images may differ or be distorted, such as shifting, rotation, zooming, and distortion resulting from noises. For the multi-source images in different FOV, based on the affine transformation table, we have designed out a registration method with which the images can be registered rapidly. In accordance with Figure 1, image B is registered based on image A. Matches between the corresponding points of the images are determined so as to decrease or eliminate the difference in the position of the objects and the distortion caused by noise.

Double-source image in the same scene definitely meets with affine transformation model, which we can assume as:

$$X' = RX + T \quad (7)$$

Where R is a rotation matrix, T is translation matrix, a coordinate after translation, rotation and zooming can be expressed by use of affine transformation as:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = k \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \quad (8)$$

Where (x, y) and (x', y') are corresponding coordinates of two images, among which vectors of four parameters-parallel pixels Δx and Δy , rotation angle θ and row zooming multiples k will determine the transformational relation between coordinates of two images.

For the definite information of the same object, the pixels in the fused region of the small FOV image are fewer than that in the wide FOV image, and consequently each pixel in the to-be-fused region of the small FOV image contains the information of at most 4 neighboring pixels in the wide FOV image. Therefore, during the affine transformation from the wide FOV image to the small FOV image, bilinear interpolation calculation, i.e. to calculate the gray value of the pixels in the corresponding wide FOV image with weighted average algorithm, is necessary.

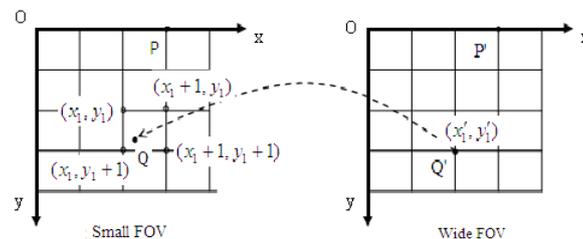


Figure 2. Gray Linear Interpolation Diagram

Use the gray value weighted interpolation of 4 neighboring grid points around point Q as its gray value, i.e. $f(x, y)$ as four values are shown in Figure 2 as (x_1, y_1) , (x_1+1, y_1) , (x_1, y_1+1) , (x_1+1, y_1+1) . For simplicity, let's set $(x_1, y_1)=(0,0)$, and thus $(x_1+1, y_1)=(1,0)$, $(x_1, y_1+1)=(0,1)$, $(x_1+1, y_1+1)=(1,1)$, the distance between point Q and point $(0,0)$ is x , and y towards Y axis, so obviously the value of point Q is $f(x, y)$.

Here comes the result after the first linear interpolation towards X axis as:

$$f(x,0) = f(0,0) + x[f(1,0) - f(0,0)] \quad (9)$$

$$f(x,1) = f(0,1) + x[f(1,1) - f(0,1)] \quad (10)$$

The value of point Q arrives at after the second linear interpolation toward Y axis like:

$$f(x, y) = f(x,0) + y[f(x,1) - f(x,0)] \quad (11)$$

After combination and simplicity we have:

$$f(x, y) = [f(1,0) - f(0,0)]x + [f(0,1) - f(0,0)]y + [f(x,1) + f(0,0) - f(0,1) - f(1,0)]xy + f(0,0) = ax + by + cxy + d \quad (12)$$

Where, a, b, c, d are constants, from which we can learn that bilinear interpolation is actually fitted by hyperbolic paraboloid of 4 known dots. Here to, we can figure out corresponding locations and weighted values of four dots of every pixel in wide FOV after registration to that in small FOV ready for registration; then we can calculate gray values of each pixel of images in small FOV after registration.

4. Analysis of the Experiment Result

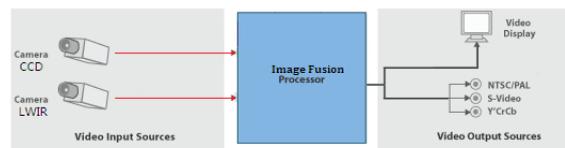


Figure 3. Block Diagram of the Systematic Principle

Image acquisition system consists of visible CCD (small FOV) and long-wave infrared camera (wide FOV). Figure 3 is block diagram of the experimental systematic principle. The processing unit of image fusion is construed by hardware circuit based on DSP chip TMS320DM642 design. Figure 4 & 5 are experimental images fused by visible images and the infrared. The experiment was conducted on the top of laboratory building in Nanjing University of Science and Technology and acquired in July 2007 when the surface temperature is above 40°C, with weighted average algorithm adopted in the experiment i.e. 50% weighting for each image.

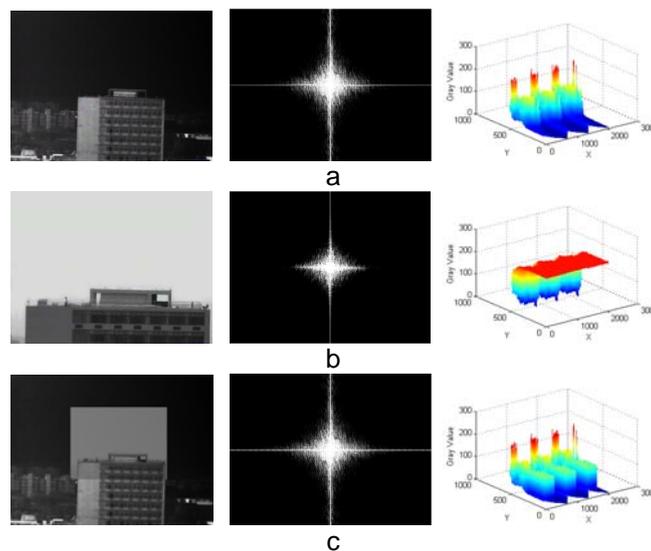


Figure 4. Long-distance Experimental Images (outdoor) Produced by Fusing the Visible Image with the Infrared Image: (a) infrared image and its Fourier spectrum and gray histogram; (b) visible image and its Fourier spectrum and gray histogram; (c) fused image and its Fourier spectrum and gray histogram.

Figure 4 is a long-distance (longer than 2km) image, in which the profile of the objects is obvious and the details are few. Figure 4(a), 4(b) and 4(c) show respectively the acquired infrared image, visible image and fused image and their Fourier spectrum and gray histogram. We can see that in the infrared image, the person can hardly be distinguished from the sky, because the weather is hot, and the temperature of the building surface is high while that of the person is relatively low. Since the field of view is wide, the spectrum graph of the infrared image contains both low-frequency information of the profile and the high-frequency information of some details; what is more, the peak of the gray histogram is obvious. While the visible image contains more low-frequency information, and its gray range is wide. The fused image has the features of the spectrum and gray histogram of the two source images. The peak value is reduced, and the fused part is smoother. In its Fourier spectrum both the low frequency and the high frequency are increased, the reason for which is that the profile and the details of the object are improved.

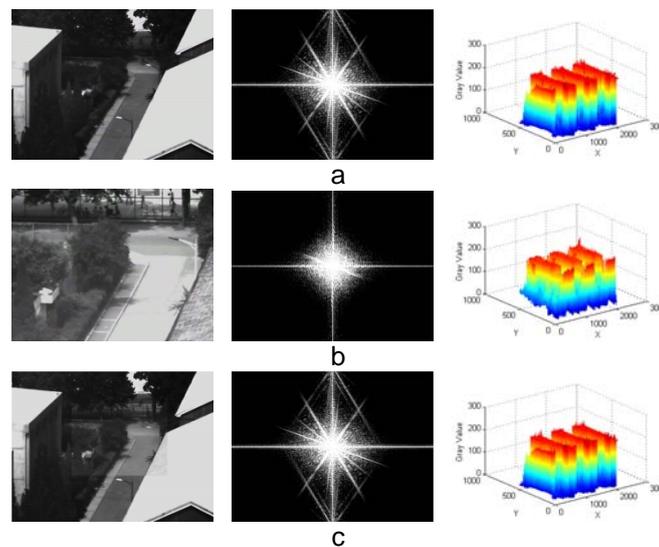


Figure 5. Short-distance Experimental Images (outdoor) Produced by Fusing the Visible Image and the Infrared Image: (a) infrared image and its Fourier spectrum and gray histogram; (b) visible image and its Fourier spectrum and gray histogram; (c) fused image and its Fourier spectrum and gray histogram.

Figure 5 is a short-distance (shorter than 500m) image in which the objects are messy, and it has more details. Figure 5(a), 5(b) and 5(c) show respectively the acquired infrared image, visible image and fused image and their Fourier spectrum and gray histogram. Since the weather is hot, and the temperature of the roof differs a lot from that of the wall, there are a lot of lines and angles in the infrared image, which is reflected as radial lines in the corresponding Fourier spectrum. In this Fourier spectrum, we can also see that some details are presented in a punctate pattern because of the big contrast. The temperature problem also results in the fuzziness of the person in the image and the striking peak in the gray histogram. While the visible image contains more details, making the person and other objects at relatively low temperature easy to distinguish. In its Fourier spectrum the continuity of the high-frequency components is well kept, which reflects that the objects in the image have strong three-dimensional effects. There are also certain peaks in the gray histogram of the visible image caused by the big contrast of the objects, but the histogram is smoother than that of the infrared image. As to the fused image, the information is mostly from the infrared image because the infrared image is partially fused, which can be seen from their spectrum. The gray histogram of the fused image is smoother than that of the infrared image. If we make a comparison between the pre-fused and post-fused images in Figure 4 and Figure 5, we may clearly see that the objects in the fused image are easier to be observed and distinguished than those in the infrared image and visible image.

5. Evaluation of the Experiment Result

Mutual Information (MI) is a concept from the information theory, referring that one variable contains about another variable. Presuming two random variables A&B, whose marginal distribution is PA (a) & PB (b) and joint possibility distribution is PAB (a, b). The fused image shall contain important information of all source images. It's not difficult to learn from what defines MI that it's very suitable to weight advantages and disadvantages of the both quality of fused images and fusion algorithm. As for source images $f(x, y)$, $g(x, y)$ and the fusion image $r(x, y)$, we can conclude how much information of source images F&G that the fusion image R contains:

$$MI_{RF}(r, f) = \sum_{a,b} p_{RF}(r, f) \log \frac{p_{RF}(r, f)}{p_R(r)p_F(f)} \quad (13)$$

$$MI_{RG}(r, g) = \sum_{a,b} p_{RG}(r, g) \log \frac{p_{RG}(r, g)}{p_R(r)p_G(g)} \quad (14)$$

Then the performance index of a fused image can be evaluated by following formula:

$$MI_R^{FG} = MI_{RF}(r, f) + MI_{RG}(r, g) \quad (15)$$

C Ramesh and T Ranjith [3] call MI Fusion Factor (FF). Bigger FF shows that more information is transmitted from source image to the fused image, and thus it represents better quality of the fused image.

$$FS = \text{abs}\left(\frac{MI_{RF}(r, f)}{MI_{RF}(r, f) + MI_{RG}(r, g)} - 0.5\right) \quad (16)$$

The smaller the FS is the better the fusion performs.

Fusion algorithm for our selection for test are weighted average fusion algorithm (AVG), discrete wavelet fusion algorithm (DWT) [4], Laplacian Pyramid fusion algorithm (LAP) [5], translation invariant discrete wavelet fusion algorithm (SiDWT) [6] and contrast pyramid fusion algorithm (CONTR) [7]. Only the fusion part in the infrared image and the fusion image will receive the evaluation. When all source images are of high quality, fusion algorithm of smaller FS should be chosen. We can learn from Table 1 that there's a stronger relativity between the fused image and the visible image.

Table 1. Evaluation Results of the Fusion Performance

| | | AVG | DWT | LAP | SiDWT | CONTR |
|----|----|--------|--------|--------|--------|--------|
| MI | IR | 2.3289 | 2.9598 | 2.4586 | 2.7654 | 2.7652 |
| | VL | 1.5002 | 1.3023 | 1.7304 | 1.3865 | 1.4324 |
| FF | | 3.7871 | 4.2101 | 4.1376 | 4.6732 | 3.9873 |
| FS | | 0.1281 | 0.2039 | 0.0978 | 0.2175 | 0.1146 |

6. Conclusion

This paper studies and explores the methods of registration and fusion for multi-source images in different field of view. Since the field of view involved differs, the registration for such images cannot be conducted as usual. At first the images should be zoomed greatly; thus in order to maintain the information of the source images and conduct the registration precisely, bilinear interpolation calculation is adopted. Through the temporal and spatial analysis of the experiment results, we find that the fused image unites features of the infrared image and that of the visible image, and therefore it is satisfying and easy to be observed.

Image fusion technology is applied in wider and wider fields. Applying areas of multi-source sensor in different FOV tends to be more and more. Fused image results in different field of view with different features make it much easier for people to discern the object in noisy

environment. Owing that image in different FOV adopts the partial fusion approach, the amount of computation in fusion algorithm reduces greatly. With the development of image fusion technology, there'll be more and more new algorithms, whose complexity and information processing capacity will gradually be upgraded. The progress of hardware technology will fail to meet designing requirements of real-time systems. So, there'll be a good prospect for image fusion technology in different field of view to be applied in and especially suitable for such fields requiring for better real time as military field, aeronautics and astronautics in a very long time in the future.

References

- [1] Pajares G, Cruz JM. A wavelet-based image fusion tutorial. *Pattern Recognition*. 2004; 37(9): 1855-1872.
- [2] Si Tian, Junju Zhang, Yihui Yuan, Benkang Chang. Research of Image Fusion Technology Based on Windowing. *Acta Armamentaria*. 2009; 30(4): 438-441.
- [3] C Ramesh, T Ranjith. *Fusion performance measures and a lifting wavelet transform based algorithm for image fusion*. Proc. of the 5th International Conference on Information Fusion. 2002; 1: 317-320.
- [4] SG Nikolov, PR Hill, DR Bull, CN Canagarajah. Wavelets for image fusion. *Wavelets in signal and image analysis*. A. Petrosian and F. Meyer (Eds.), Computational Imaging and Vision Series. 2001; 19: 213-244.
- [5] Peter J Burt, Edward H Adelson. The laplacian pyramid as a compact image code. *IEEE Transaction on Communications*. 1983; 31(4): 532-540.
- [6] A Toet, LJ van Ruyven, JM Valeton. Merging thermal and visual images by a contrast pyramid. *Optical Engineering*. 1989; 28(7): 789-792.
- [7] O Rockinger. *Image sequence fusion using a shift invariant wavelet transform*. Proceedings of the International Conference on Image Processing. 1997; 88-291.