PROCEEDINGS OF SCIENCE

Cognition-based Fusion Method for Infrared and Visible Image

Junyong Ma
Science and Technology on Electro-Optical Control Laboratory
Luoyang, 471000, China
E-mail: `majunyongzhangduo@126.com`

Duo Zhang
Luoyang Institute of Electro-optical Equipment, AVIC
Luoyang, 471000, China
E-mail: `zhangduomajunyong@126.com`

We proposed an image fusion method based on cognition for infrared and visible images. Multi-source image fusion is guided top-down by using the prior knowledge as acquired from deep understanding of multi-source images. As different fusion methods and fusion rules are used for contents of different importance, it makes the image fusion process more free and flexible. Experiments for the presented method are conducted by using infrared and visible image. The fusion results and the image fusion quality indexes are compared with those of the traditional region-based fusion methods. The comparison show the proposed method is far superior to the traditional region-based fusion methods.

ISCC2017
16-17 December 2017
Guangzhou, China

PoS(ISCC 2017)022

1. Introduction

The image fusion is an information fusion technology concerning multi-source image procession. A fused image can be obtained from multiple images as acquired from different sensors or from the same sensor in different imaging manners. The fused image can reflect multiple information from the original images, and it is favored to achieve a comprehensive description and precise analysis judgment of the targets and scene [1, 2]. While the visible images are rich in contour, texture, color information, the infrared images can provide the temperature information for the targets and show their states. Take advantage of the complementary information, the fusion result from the multi-source images can more clearly describe the states of the targets [3, 4].

Usually, after the image fusion system has collected the infrared and visible images, the image preprocessing is done firstly. Then, the feature extraction and feature matching are completed for image registration. By using coordinate transformation acquired from image registration, the infrared and visible images are mapped into the same coordinate plane. Then determine the common image area of the wide field of view corresponding to the image of the small field of view. Consequently, interpolate infrared image or visible image in order to obtain images of the same resolution and realize the pixel correspondence. Finally, the fusion processing for infrared and visible images is conducted by specific fusion rules.

The traditional fusion method chooses a certain fusion rule to obtain the final fusion results according to the multi-source image characteristic. Burt[5] proposed a pixel-based fusion rule by setting some measurement indicators between corresponding pixels in source images in order to obtain the fused image. As the rule isolates the correlation between the pixels, the fusion result is not very well. Burt P. J and Kolczynski R. J [6] put forward a window-based fusion rule, in which the fusion coefficient was selected by comparing the statistical properties of a window in the source images. Due to its consideration for the pixel correlation in the window, the effect is superior to the pixel-based fusion rule and it has been widely applied. Zhang. Z and Blurn. R. S [7] and Piella G [8] have proposed region-based fusion rule. Each pixel in the image is considered as part of a region or an edge while the regional and the border information of an image is used to guide the fusion rule selection. The region-based fusion rule can acquire better fusion effect, but it is more complicated than other fusion rules. Furthermore, the implementation of this rule is not easy for complex images.

McNeese proposed multi-sensor image cognition model [9]. He built an image fusion system framework for the airborne infrared and visible images from human's understanding of multi-source images. Then, the framework of the system was improved by Amanda c. Muller and s. Narayanan [10]. This further enriched and perfected the cognitive model of thermal infrared image and visible image fusion. In 2010, Toet proposed an image fusion principle based on cognition, in which the reference contour was used to guide the implementation of the image fusion rule [11].

The cognition-based image fusion method is thus proposed on the basis of human’s deep understanding of multi-source images. It is a top-down integration, with the preconceived "fusion result" of the brain are used to guide multi-source image fusion [12,13]. The “fusion result” is the final abstract representations of the image, acquired by human according to their prior knowledge.

The images from different sensors convey different contents for the same scene. In the comprehensive understanding of multi-source images, the human brain often unconsciously integrates the optimal reflections of the different contents in multi-source images to form a "fusion result". It consists of important content, sub-important content and general content of the scene. Actually, the cognitive process embodies the basic idea of multi-source image fusion. According to the principle, we apply the cognition idea into thermal infrared and visible image fusion and propose a cognition-based image fusion framework. Different fusion methods and rules are performed for different types or contents of different importance degree. Compared with the traditional image fusion methods, it breaks through the limit of only by using a single method or a fusion rule. As a result, it has more flexibility and freedom, embodying the free and flexible characteristics of image fusion.

2. Fusion Framework

The framework of image fusion model based on content cognition is shown in Fig. 1.

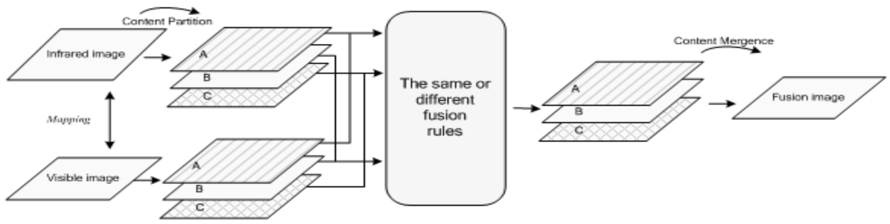


Figure. 1: Framework Model Based on Content Cognition

The framework model can be described as follows. For a given scene of multi-source images, firstly we would clarify the multi-source images into different goal contents according to our different interesting degrees; then, select the appropriate fusion method and fusion rules according to the characteristics of different image contents in multi-source images; finally, different contents in the fusion result are merged to get the final fusion image.

After the importance degree of the contents and the fusion principle is defined, the first step is to realize the goal contents division for the thermal infrared image and visible image. Generally, what we are interested in is the hot target information, mainly coming from the thermal infrared image. As a result, the thermal infrared image is regarded as the reference image, in which different goal contents are classified. The thermal infrared image can be divided into the target region, the target-surrounding region and background region, all of which are mapped into the visible image to obtain their corresponding content regions.

3. Target Segmentation

Currently, common target segmentation methods mainly include threshold segmentation methods, morphological segmentation, region growing methods and clustering segmentation etc. [14]. k-means clustering segmentation algorithm is utilized here. Of course, the direct threshold segmentation method can also be used for these images, in which the gray distribution difference between object and background is larger.

K-means algorithm regards the similarity between the input variables as the rule. Firstly, K different grayscale pixels are determined in the image classification. Then for a pixel in positionwith its grayscale value is, the gray similarity criteria is set as

$$\delta(x, y) = |(f(x, y) - c^k)| \le \epsilon \tag{3.1}$$

where c^k refers to the kth initial clustering centers,refers to an user selected threshold. Secondly, calculate the similarity of the rest of the pixels with the K clustering centers respectively, and the pixels are classified into the category with which has the highest similarity measure. Then, update the average gray level of every cluster. Finally, the image can be divided into different goal content areas after repetition or iteration.

4. Fusion Rules

In general, a scene imaging with multi-source images of thermal infrared and visible images can be divided into target regions, target-surrounding regions and background regions. The target signature mainly comes from the thermal infrared images, which should try to be reserved in the fusion result. The target-surrounding regions mainly comes from the visible image. So, the space texture characteristics of the target-surrounding regions in the visible image should be retained as much as possible; on the other hand, the brightness and contrast of the region should be improved as far as possible. The background region between thermal infrared image and visible image not only has similar parts, but also dissimilar parts. The fusion of background region should mainly take the dissimilar parts into consideration. The fusion rule is described as follows:

For the target content area, let I, V and F represent respectively the thermal infrared image, the visible image and the fusion image. And $C^I(x,y)$ and $C^V(x,y)$ are respectively the pixel values of target content in thermal infrared image and visible image. Then the space domain weighted fusion rules are applied and the target pixel value of the fusion image $C^F(x,y)$ can be represented as

$$C^F(x,y)=\omega^I C^I(x,y)+\omega^V C^V(x,y) \tag{4.1}$$

where ω^I and ω^V are respectively the weighted coefficient of thermal infrared image and visible image.

There are a lot of space texture features in the target-surrounding content region of the visible image. In order to capture the characteristic details of visible image in all directions, multi-resolution method is applied and the decomposition layers can be chosen as 4 layers. In order to minimize influence of the selection of fusion rule to the effectiveness of the proposed fusing framework, the window variance weighting fusion rule is chosen for low frequency of the multi-scale decomposition in target-surrounding content region, as

$$L^F(x,y)=\omega^I L^I(x,y)+\omega^V L^V(x,y) \tag{4.2}$$

$$\omega^I=\sigma^I(x,y)/(\sigma^I(x,y)+\sigma^V(x,y)) \tag{4.3}$$

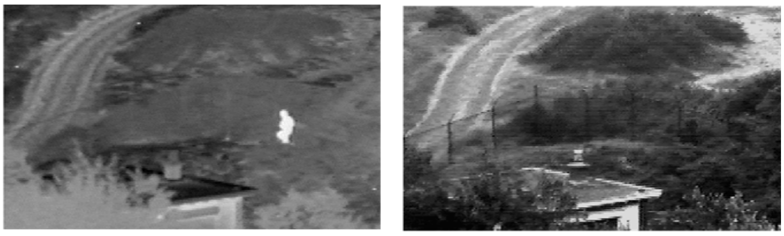
$$\omega^V=\sigma^V(x,y)/(\sigma^I(x,y)+\sigma^V(x,y)) \tag{4.4}$$

where $\sigma^V(x,y)$ and $\sigma^I(x,y)$ are respectively variance of the 3×3 window in the target-surrounding content regions for visible image and infrared image. While the maximum absolute value fusion rule is selected for high frequency part.

In general, there are common parts between background region of the infrared image and that of the visible image. Also, there are dissimilar parts which diverge greatly from each other. Compared with the target-surrounding region, the background region of the visible image contains fewer texture details. It is not an optimal choice to extract the fewer details in the frequency domain of the background region because of limited texture details, moreover some brightness information may be lost during the frequency domain transformation; therefore, the space weighted fusion method and fusion rules are selected for background region in this paper.

5. Experimental Results and Analysis

As shown in Fig. 2, we have adopted the UN Camp thermal infrared image and visible image to illustrate our fusion process and results.



(a) UN Camp infrared image (b) UN Camp visible image

Figure 2: UN Camp Thermal Infrared Image and Visible Image

In our cognition-based fusion experiments, primary focus is the target of information, and the second important focus is the target-surround scene information and the least important focus is the background of the scene.

After we have defined the content level of importance and fusion rules, we begin to implement the content division for the thermal infrared image and visible image. Because we are interested in the hot target information, and the hot target information mainly comes from thermal infrared images, so we use thermal infrared image as the reference image and classify it into different contents. Firstly, the thermal infrared image is divided into the target area, the target-surround region and the background region by using k-means clustering. Then they are mapped to the visible image, as shown in Fig. 3.

The principle of fusion is to retain as far as possible the target features in the thermal infrared image and texture details in visible image. As stated earlier, we convey different fusion rules to different content areas in the image fusion. The fusion result and comparison with traditional region-based method results, as seen in Fig. 4.

The image fusion experiment is executed for infrared and visible image fusion. Meanwhile, two objective evaluation indexes of Edge Mutual Information [15] and Pixel Mutual Information [16] are calculated for fusion performance evaluation. Edge Mutual Information (EMI) reflects "inheritance" of the fusion image from the input image. And the greater this value is, the more edge information reserved in the fusion image. Pixel Mutual Information (PMI) is applied to evaluate quantitatively the overall effect of the fused image. It can describe the information similarity degree between two images. Generally, the bigger the value of MI is, the more mutual information will be included in two images.

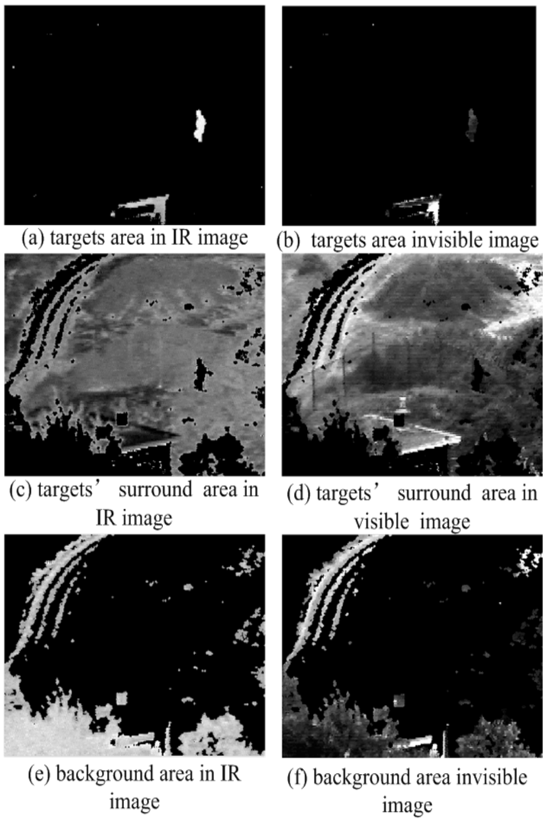


Figure 3: Content Division of UN Camp Thermal Infrared Image and Visible Image

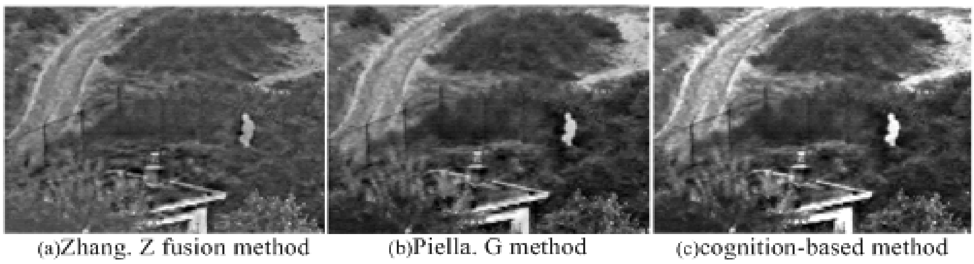


Figure 4: Comparisons of the usion result, in which (a), (b), (c) are respectively fusion results of Zhang Z fusion method, fusion result of Piella. G fusion method, and fusion result of image fusion method based on content cognition.

We can clearly notice from the fusion results that the fusion results of the presented method are superior to two other traditional fusion methods based on region. We can draw the same conclusion from the qualitative evaluation index in Table 1.

Image Fusion Methods	EMI	PMI
Piella. G method	0.51	2.65
Zhang. Z method	0.53	2.67
Proposed method	0.72	2.81

Table 1: Image Fusion Quality Evaluation Index

6. Conclusion

This paper proposes an image fusion framework based on content cognition. Firstly, the thermal infrared image would be divided into the target region, the target-surrounding region and the background region, all of which are mapped into the visible image to obtain their corresponding content regions.

According to the characteristics of different image contents in multi-source images, appropriate time domain or frequency domain fusion method and fusion rules are selected. Finally, different contents of the fusion are merged to get the final fusion image. The limit of using a single fusion method for traditional image fusion framework is broken, which makes the image fusion algorithm and the selection of image fusion rule more free and flexible. The experimental results show that the image fusion method based on content cognition can acquire better fusion effects.

References

[1] Goshtasby A A, Nikolov S. *Image fusion: Advances in the state of the art* [J]. Information fusion, 2007, 8(2): 114-118.

[2]Yang B, Li S T. *Pixel-level image fusion with simultaneous orthogonal matching pursuit* [J]. Information fusion, 2012, 13(1): 10-19.

[3]Piella G. *A general framework for multiresolution image fusion: from pixels to regions* [J]. Information fusion, 2003, 4(4): 259-280.

[4]Li H F, Yu Z T, Mao C L. *Fractional differential and variational method for image fusion and super-resolution* [J]. Neurocomputing, 2016, 171: 138-148.

[5]Burt. P J． *The pyramid as a structure for efficient computation*[C]． In：Multiresolution Image Processing and Analysis, London: Springer Verlag, 1984: 6-35．

[6]Burt P.J, Kplczynski R J. *Enhanced image capture through fusion*[C]. Proc. 4th Int. conf. on Computer Vision, Berlin, Germany, may, 1993:173-182.

[7]Zhang Z., Blum R. S. *A Hybrid Image Registration Technique for A Digital Camera Image Fusion Application* [J]. Information Fusion, 2001, (2): 135-149.

[8]G. Piella, H. Heijmans, *A new quality metric for image fusion* [C]. proceeding of IEEE International Conference on Image Processing, Barcelona, Spain, 2003. 173-176.

[9]M.D. McNeese, H.S. Bautsch, S. Narayanan, *A framework for cognitive field studies*[J], International Journal of Cognitive Ergonomics, 1999, 3: 307–331.

[10]A. C. Muller. S Narayanan, *Cognitively-engineered multi-sensor image fusion for military applications*. Information Fusion, 2009, 10(2): 137-149.

[11]A. Toet, M.A. Hogervorst, S.G. Nikolov, J.J. Lewis, T.D. Dixon, D.R. Bull, and C.N. Canagarajah, *Towards Cognitive Image Fusion*, Information Fusion, 2009,11(2): 95-113.

[12]Pinker S. *Visual cognition: An introduction* [J]. Cognition, 1984, 18(1): 1-63.

[13]A. C. Muller. S Narayanan, *Cognitively-engineered multi-sensor image fusion for military applications*. Information Fusion, 2009, 10(2): 137-149.

[14]Hao X X, Zhao H, Liu J. *Multifocus color image sequence fusion based on mean shift segmentation* [J]. Applied Optics, 2015, 54(30): 8982-8989.

[15]Petrovic V, Xydeas C. *Cross band pixel selection in multi-resolution image fusion.* Proc. of SPIE,

1999, Vol.3719: 319-326.

[16]Guihong, Z. Dali, Y. Pingfan; *'Information measure for performance of image fusion'*, Electronics Letters, Mar 28, 2002, 38(7): 313-315