Acta Technica 62, No. 3B/2017, 85–94

Comparison and analysis of several pixel-level image fusion algorithms[1]

Qian Chen[2], Zhi-yong Bai[2,3]

Abstract. From the beginning of the 1980s, a worldwide research boom has been triggered by image fusion, which has a wide application prospect in computer vision, remote sensing, robotics, medical image processing and military fields. Pixel-level image fusion is the information fusion at the basic data level. It is high in accuracy, and can provide information in detail that cannot be provided by other levels. Besides, the amount of information to be processed by the pixel-level image fusion is the largest. In this thesis, several commonly used pixel-level image fusion algorithms have been introduced, and their experiments have also been contrasted and analyzed.

Key words. Image fusion, pixel-level, algorithm, contrast, analysis.

1. Introduction

Image fusion is to combine, in a particular scene, an image or image sequence information acquired by two or more sensors at the same time or at different times, so as to generate a new information processing process for this scene interpretation [1]. Pixel-level image fusion is the image fusion at the lowest level but with the highest accuracy, so it can provide the details of the information that other levels of fusion processing do not have, as well as it has to process the largest amount of information and has a high requirement on its equipment [2]. The first multi-scale fusion method, proposed by Burt [3] in 1984, uses the Laplacian pyramid as a multi-scale analysis tool, and the absolute value of the coefficients as the degree of activity, getting the fusion coefficient by choosing the maximum value. A fusion method based on gradient pyramid was proposed by Burt and Kolczynski [4].This method has partially solved the offset effect caused by the opposite contrast ratio of the source coefficients, and has improved the anti-noise property of the algorithm based on the fusion method and the gradient pyramid use. Based on wavelet

[1]This paper was supported by the Fundamental Research Funds for the Central Universities (No. 2017ZY54) and the Nation Key R&D Program of China (2015BAH52F00).

[2]1School of Arts and Design, Beijing Forestry University, Beijing, 100083, China

[3]Corresponding author

transform, Li et al. [5] have proposed an image fusion method, which defines the activity as the largest absolute value in the coefficient window, uses the maximum value selecting mapping method in the decision process and proposes a consistency detection method to remove inaccurate value in the initial decision. According to the frequency response of human visual system, a weighted fusion method based on perceptual information was proposed by Wilson et al. [6]. Compared with Burt and Kolczynski's method, this program has provided a better visual effect. Pu et al. [7], [8] proposed a fusion method based on directional contrast, where the directional contrast is defined as the ratio of the wavelet detail coefficient to the approximate coefficient on the same scale. The absolute value of the direction contrast, in this method, was treated as activity degree, and the maximum value selection strategy was adopted in decision mapping. Based on multi-wavelet transform, Li et al. [9] proposed a multi-sensor image fusion method which uses the maximum selection mapping method. De et al. [10] used morphological wavelet transform to perform multi-scale analysis in order to improve the computational efficiency of the fusion algorithm. This method also uses the same mapping method. There are also some fusion methods based on Curvelet transform [11], Contourlet transform [12], etc. The study of fusion rules is also increasingly biased towards intelligence, such as the combinational use of statistical probability models [13] and stochastic models [14]. An improved image fusion algorithm using the Laplacian pyramid was also proposed by Kumar et al. [15] in 2014. The algorithm begins with applying 2D-DWT to decompose the input images. The lower approximations are subjected to pixel-based Laplacian fusion algorithm. For higher approximations, the SF algorithm needs to be combined with wavelet fusion algorithm. Then, from each image, the new sets of detailed and approximate coefficients are added in order to get the new fused coefficients. Compared with the pixel-based and wavelet-based algorithms, the hybrid model proposed is an improved version of fusion image. Liu et al. [16] proposed a common image fusion framework by combining MST and SR as well as overcoming the inherent flaws of the fusion methods based on MST and SR. In 2016, based on Clustering and NSCT, Xiong et al. [17] proposed a novel Image Fusion Algorithm for Visible and PMMW Images. Aiming at the fusion of visible and Passive Millimeter Wave (PMMW) images, a novel algorithm based on clustering and NSCT (Nonsubsampled Contourlet Transform) was proposed. Experiments demonstrate the superiority of the proposed algorithm for metal target detection compared to wavelet transform and Laplace transform.

2. Common pixel-level image fusion algorithms

2.1. Mean value method

Let $I_1(x,y)$, $I_2(x,y)$ and $I(x,y)$ denote, respectively, the pixel values of the first image, the second image, and the fused image at the point (x,y). Then, in the fusion images, the pixel values at each point are determined by equation

$$\mathrm{I}(x,y)=\begin{cases} I_1(x,y) & (x,y)\in R1 \\ \frac{1}{2}\left(I_1(x,y)+I_2(x,y)\right) & (x,y)\in R2 \\ I_2(x,y) & (x,y)\in R3 \end{cases} \tag{1}$$

Here, $R1$ represents an image region where the first image does not overlap with the second image, $R2$ represents an image area where the first image overlaps the second image, and $R3$ represents an image region in which the second image does not overlap with the first image. To take the average of two images is fast, but generally with unsatisfactory effect, besides in the fused portion, we can sense the tripe obviously, in which the difference is able to be observed. The average value method is used to fuse the image overlap area. The result is that it combined well at the splicing point, but in overlapping area, the simple average method is used, so fuse quilt is obvious.

2.2. Linear transition in overlapping area

In order to eliminate the quilt problem in the overlapping area, a linear transition method was used in the overlapping area recently. The specific method of implementation is to assume that the overlapping area width is L. Take the transition factor σ $(0 \le \sigma \le 1)$. The maximum and minimum values of the x-axis and y-axis of the two images in the overlapping regions are, respectively, $x_{\max}$, $x_{\min}$ and $y_{\max}$, $y_{\min}$, then, the transition factor is

$$\sigma=\frac{(x_{max}-x)}{(x_{max}-x_{min})} \tag{2}$$

and the pixel value of the overlap area is

$$I=\sigma I_1(x,y)+(1-\sigma)I_2(x,y)\ . \tag{3}$$

The pixel values corresponding to the first image and the second image are represented by I_1, I_2, respectively. This method makes the transition part smoother, with no obvious steps. The obvious traces of splicing can be eliminated to perform image fusion by using the linear transition method in overlapping area [19]. However, there are still some splicing traces for two images with large differences in light, and due to the presence of moving objects, the resulting fusion image tends to have double image, and even the same object will appear twice in the splicing image.

2.3. Linear transition in overlapping area

Common pyramids are Gaussian pyramids, Laplacian pyramids, low pass filter pyramids, contrast pyramids, morphological pyramids, gradient pyramids and so on. The Laplacian multi-resolution pyramid structure was adopted in the multiresolution spline technique, which decomposes the image into a set of images at different frequencies, weights the average of the image edges at each decomposed frequencies. Finally, the synthesized images on all frequencies are aggregated into an image. In each frequency band, the coefficients of the weighting function and the size of the color fused region are determined by the difference of the two image characteristics in the frequency band, so this method can propel the images with different intensities having a smooth transition.

Laplacian pyramid transformation is based on the Gaussian pyramid. First the new image A_1* was obtained by enlarging and filtering A_1, so the numbers of rows and columns of A_1*, the same as A_{1-1}, were twice of that in A_1. The formula for A_1* interpolated by A_1 is

$$A_l * (i,j) = 4 \sum_{m=-2}^{2} \sum_{n=-2}^{2} \omega(m,n) A_l \left(\frac{i+m}{2}, \frac{j+n}{2} \right) . \tag{4}$$

It can be summed when the $\frac{i+m}{2}, \frac{j+n}{2}$ are integers. In order to simplify the writing, the Expand operator was introduced in the form

$$\begin{cases} \mathrm{LP}_l = A_l - \mathrm{Expand}\left(G_{l+1}\right) & 0 \leq 1 \leq N-1 \,, \\ \mathrm{LP}_N = A_N & l = N \,, \end{cases} \tag{5}$$

where N is the decomposition layer of Laplacian pyramid and LP_l is the first decomposition layer in the Laplacian pyramid. Thus, except the top layer, each layer is the difference value through Expand between the Gaussian pyramid image of this layer and the higher layer image, so it can be called that LP_0, LP_l, LP_2, ..., LP_N constitute the Laplacian pyramid.

For reconstruction, it can be represented by:

$$A_0 = \mathrm{LP}_0 + \mathrm{Expand}\left(\mathrm{LP}_1 + \mathrm{Expand}\left(\mathrm{LP}_2 + \cdots\cdots + \mathrm{Expand}\left(\mathrm{LP_N}\right)\right)\right) . \tag{6}$$

For sub image in each layer of Laplacian pyramid, the source image A_0 can be accurately reconstructed by using the Expand operator to gradually interpolate and amplify the sub image until it has the same size as the source image and finally adding them up. In general, the construction of Laplacian pyramid decomposition includes low-pass filtering, down-sampling, interpolated value (i.e., magnified size) and band-pass filtering.

2.4. Fusion algorithm based on wavelet decomposition

In computer vision, pixel-based image representation (pixel scale) applies only to some data-level processing, and in more cases, the characteristics of the image on the appropriate scale have to be extracted. As object size in the image is different, an optimal scale is not possible to be defined to analyze the image in advance, so the image content is necessary to be considered from different scales. The first step of wavelet transform divided the signal into low frequency parts (called approximate parts) and high frequency parts (called details). The approximate part represents the main characteristic of the signal. In the second step, the similar operation was performed on the low-frequency part, so the scale changed in turn to the required scale.

Wavelet transform refers to select the appropriate wavelet base and form a series of wavelet through translation and stretching, then project the signal to be analyzed to the signal space constituted by translation and stretching. This translation, stretching and enlarging is one of the characteristics of wavelet transform, which can be analyzed in different frequency and airspace. Wavelet transform can be roughly divided into two types: continuous wavelet transforms and discrete wavelet transform.

For the function $2\varphi(x) \in L^2(R)$, the following equation is satisfied:

$$\int R\varphi(x)\,\mathrm{d}x = 0\,. \tag{7}$$

Then $\varphi(x)$ is called the Basic Wavelet or Mother Wavelet. Then, a function family $\varphi_{a,b}(x)$ can be obtained through translating and stretching the function, where a represents the scaling factor and b represents the translation factor. Then $\varphi_{a,b}(x)$ is called the wavelet continuous dependent on a and b and there holds

$$\varphi_{a,b}(x) = |a|^{-\frac{1}{2}}\varphi\left(\frac{x-b}{a}\right), \quad a, b \in R, \quad a \neq 0\,. \tag{8}$$

3. Evaluation criteria of image fusion

The quality evaluation, after the image fusion, can be divided into subjective quality evaluation and objective quality evaluation. The former one refers to the examines giving the evaluation of the fusion image according to the existing knowledge and evaluation criteria. The examines are easy to be influenced by their experience and subjective psychology, therefore, it is difficult to implement this method, and so is to realize the high real-time scene. The latter one refers to the evaluation using the image parameters and some evaluation indicators to automatically evaluate the image. The objective quality evaluation method can be used to evaluate some properties of the image itself, such as the evaluation pixel intensity, standard deviation, average gradient, entropy, and so on.

3.1. Average Pixel Intensity (API)

Average pixel intensity, also called contrast of the mean value, is the arithmetic mean of all the gray values of the pixels in the image, and the average brightness of the human eye, which can be defined as

$$\mathrm{API} = \overline{F} = \frac{\sum_{i=1}^{M}\sum_{j=1}^{n} f\left(i,j\right)}{mn}\,, \tag{9}$$

where $f(i,j)$ represents the pixel value at (i,j), and $m\,n$ represents the size of the image.

3.2. Standard Deviation (SD)

The standard deviation, reflecting the distribution of the data, is the square root of the variance. It is the dispersion degree measuring the image gray scale relative to the average gray scale. The larger the standard deviation, the more scattered the gray value distribute, which means the image contrast is large, can reflect more information and has a better effect

$$\mathrm{SD} = \sqrt{\frac{\sum_{i=1}^{M} \sum_{j=1}^{n} f\left(i,j\right) - \overline{F}}{mn}}\,. \tag{10}$$

3.3. Average Gradient (AG)

The average gradient, reflecting the clarity of the image, represents the detail information in the edge of an image. It is sensitively reflecting the improvement of image details, and the greater it is, the better performance it has, and the higher the quality is

$$\mathrm{AG} = \frac{\sum_i \sum_j \left(\left(f\left(i,j\right) - f\left(i{+}1,\right)\right)^2 + \left(f\left(i,j\right) - f\left(i{+}1,\right)\right)^2\right)^{1/2}}{mn}\,. \tag{11}$$

3.4. Entropy (H)

Entropy (H) represents the amount of information contained in the image. It is an important indicator expressing the information richness. The larger the entropy, the more information it can fuse in the image and the better the quality becomes

$$\mathrm{H} = -\sum_{k=1}^{255} P_k \log_2 P_k\,, \tag{12}$$

where P_k represents image pixel value if the occurrence probability of k.

4. Experimental results and analysis

In order to compare the performance of various pixel-level image fusion methods, the fusion experiment was made by using images of Figs. 1(a) and 1(b). To guarantee the objectivity, four common objective evaluation indexes, such as Average Pixel Intensity (API), Standard Deviation (SD), Average Gradient (AG) and Entropy (H) were used to evaluate the pixel intensity. The evaluation parameters of each image fusion algorithm used in the above performance evaluation index are summarized in Table 1.

Fig. 1. Test picture

Table 1. Contrast of evaluation parameters for different fusion algorithms

Fusion method	API	SD	AG	H
Mean value method	85.4201	41.5570	4.0299	7.0470
Linear transition in overlapping area	94.0328	44.3115	4.1004	7.0953
Laplacian pyramid	97.5650	45.9082	5.7177	7.1375
Wavelet decomposition	97.5536	45.9065	5.8136	7.1415

According to the table showed above, from the perspective of API and SD, the Laplacian pyramid algorithm based on the tower decomposition, performing little difference with fusion algorithm based on wavelet decomposition, performs best in image fusion. But the API and the SD in the pyramid transformation method and the wavelet decomposition method are significantly greater than that in the average method and the linear transition in overlapping area method, for the API and SD, the highest numerical algorithm for the Laplacian pyramid, the lowest for the average method, and the gap is large, the difference is 12.1449, API value difference is 12.1449, SD value difference is 4.3512. Both the mean value method and the linear transition in overlapping area method have lower relative score on the evaluation method of AG, significantly lower than the latter two. According to H, the fusion algorithm based on wavelet decomposition is relatively good, but there is little entropy difference in these four algorithms, with the difference of only 0.0945 between the maximum and the minimum, the linear transition in overlapping area method is higher than the average method in the four evaluation criteria of API, SD, AG and H. On the whole, the entropy, AG, and SD of the pyramid transformation and the wavelet decomposition method are all larger than that in the other two methods. Therefore, the image fusion algorithm of Laplacian pyramid method and wavelet decomposition method are relatively better, followed by the linear transition in overlapping area method, the last is the average method.

5. Conclusion

With the rapid development of image processing technology and the wide application of various image sensors, the emerging image fusion technology has become an important direction in image processing research. Image fusion makes full use of the complementary information of multiple source images, eliminates redundant messages, so that the fusion image is more in line with human's visual characteristics, and more conducive to the follow-up processing and target detection or tracking of the image. Therefore, image fusion is one of the important and valuable research topics in the field of image processing. In this thesis, several commonly used pixel-level image fusion algorithms are introduced and contrasted by combining them with experiments, then the results obtained were analyzed by using objective quality evaluation system. In recent years, image fusion is one of the most important research topics in this field. At the same time, the research on image fusion technology is still at its initial stage, and has not yet formed a complete theoretical framework and system. Therefore, this thesis has only compared the pixel-level image fusion algorithms, although some conclusions are drawn, there are still many unresolved problems in the field that need to be continued.

References

[1] J. H. Park, K. O. Kim, Y. K. Yang: *Image fusion using multiresolution analysis.* IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 9–13 July 2001, Sydney, NSW, Australia, IEEE Conference Publications *2* (2001), 864–866.

[2] T. Jing, C. Li: *Adaptive multi-focus image fusion using a wavelet-based statistical sharpness measure.* Signal Processing *92* (2012), No. 9, 2137–2146.

[3] P. J. Burt: *The pyramid as a structure for efficient computation.* Springer International Publishing AG, Book Series (SSINF) *12* (1984), 6–35.

[4] P. J. Burt, R. Kolczynski: *Enhanced image capture through fusion.* IEEE International Conference on Computer Vision, 11–14 May 1993, Berlin, Germany, IEEE Conference Publications (1993), 173–182.

[5] H. Li, B. Manjunath, S. Mitra: *Multisensor image fusion using the wavelet transform.* Graphical Models and Image Processing *57* (1995), No. 3, 235–245.

[6] T. Wilson, S. Rogers, J. Myers: *Perceptual-based hyperspectral image fusion using multiresolution analysis.* Optical Engineering *34* (1995), No. 11, 3154–3164.

[7] T. Pu, Q. Fang, G. Ni: *Contrast-based multiresolution image fusion.* Acta Electronica Sinica *12* (2000), No. 28, 116–118.

[8] T. Pu, G. Ni: *Contrast-based image fusion using the discrete wavelet transform.* Optical Engineering *39* (2000), No. 8, 2075–2082.

[9] H. H. Wang: *Multisensor image fusion using discrete multiwavelet transform.* International Conference on Machine Learning and Cybernetics (IEEE Cat. No.04EX826),

26–29 August 2004, Shanghai, China, IEEE Conference Publications *7* (2004), 4331–4336.

[10] I. De, B. Chanda: *A simple and efficient algorithm for multi-focus image fusion using morphological wavelets.* Signal Processing *86* (2006), No. 5, 924–936.

[11] M. G. Choi, R. Y. Kim, M. R. Nam, H. O. Hong: *Fusion of multispectral and panchromatic Satellite images using the curvelet transform.* IEEE Geoscience and Remote Sensing Letters *2* (2005), No. 2, 136–140.

[12] Q. Miao, B. Wang: *A novel image fusion method using contourlet transform.* IEEE International Conference on Communications, Circuits and Systems, 25–28 June 2006, Guilin, China, IEEE Conference Publications *1* (2006) 548–552.

[13] F. Sroubek, J. Flusser: *Image fusion via probabilistic deconvolution.* PRRS International Workshop on Pattern Recognition in Remote Sensing, 27 August 2004, Kingston upon Thames, GB, Proceedings (2004), 1–5.

[14] M. Xu, H. Chen, P. K. Varshney: *An image fusion approach based on Markov random fields.* IEEE Transactions on Geoscience and Remote Sensing *49*(2011), No. 12, 5116–5127.

[15] U. Kumar, B. Vikram, P. Patil: *Enhanced image fusion algorithm using Laplacian pyramid.* International Journal of Engineering & Science Research *4* (2014), No. 7, 525–532.

[16] Y. Liu, S. Liu, Z. Wang: *A general framework for image fusion based on multi-scale transform and sparse representation.* Information Fusion *24* (2015), 147–164.

[17] J. Xiong, W. Xie, J. Yang, Y. Fu, K. Hu, Z. Zhong: *A novel image fusion algorithm for visible and PMMW images based on clustering and NSCT.* International Conference on Computer and Automation Engineering (ICCAE), 3–4 March 2016, Melbourne, Australia, MATEC Web of Conferences *56*(2016), Publications EDP Sciences Journal, Curran Associates, Inc.

[18] B. Tao, Z. Yu, X. Zheng: *Automatic Image Mosaic.* Chinese Journal of Biomedical Engineeringh *16* (1997), No. 4, 316–322.

[19] S. Nizam, Z. Telatar: *Multispectral image fusion based on the Multiwavelet and IHS transforms.* Signal Processing and Communications Applications Conference (SIU), 18–20 April 2012, Mugla, Turkey, IEEE Conference Publications (2012), 1–4.

Received April 30, 2017