A Comparative Study of Wavelet-based Image Fusion with a Novel Fusion Rule II

Bo Jiang*†, Xinyuan Wang*‡, Rui Zhang*§, Ziqi Lin*¶, Qunxia Gao*‖, Jianzhong Luo*** and Xiao Zhang*††
*Guangzhou Institute of Biomedicine and Health
Chinese Academy of Sciences, Guangzhou, Science Park
†Email: jiang.bo@gibh.ac.cn
‡Email: wang.xinyuan@gibh.ac.cn
§Email: zhang.rui@gibh.ac.cn
¶Email: lin.ziqi@gibh.ac.cn
‖Email: gao.qunxia@gibh.ac.cn
**Email: luo.jianzhong@gibh.ac.cn
††Email: zhang.xiao@gibh.ac.cn

***Abstract*—Image fusion is the process of combining information of interest in two or more images of a scene into a single highly informative image. Generally, the multi-resolution image fusion based on the wavelet transform performs better on diverse images than traditional methods. Therefore, a comparative study of wavelet-based image fusion with different wavelet families and fusion methods are far-reaching to guide people in their applications. In our previous paper, a platform of wavelet-based image fusion has been built, which includes different wavelet families, image fusion rules, and statistical evaluation criteria. Therefore, it can be used to assess, and more importantly, select appropriate method for a specific fusion task. In this paper, the platform is further extended with more advanced algorithms and objective evaluation criteria. Discussions with these algorithms with multi-focus and multi-spectral images demonstrate that our proposed fusion approach still outperforms traditional methods in most of multi-focus images and multi-spectral images, which also complies with statistical analysis by objective evaluation criteria.**

keywords: mage fusion; fusion rule; wavelet transform; focus measure; local contrast measure; objective evaluation criterion.

I. Introduction

Image fusion has become a widely used tool to integrate different information images like getting an "all-in-focus" image from a set of multi-focus images or multi-spectral images for various applications such as remote sensing, astronomy, medical imaging, security, surveillance, etc. The main objective of the image fusion is to preserve or combine all important visual information from multiple input images without losing any significant information or features and introducing any artifacts. Concerning the various methods developed for fusing various images from diverse areas, it is necessary to give a general assessment and analysis of the wavelet-based fusion methods.

In our previous paper, discussions of the image fusion focus on multi-resolution analysis, especially in the wavelet-based algorithm [4], in which the comparisons of different wavelet families, decomposition levels (DL), image fusion rules, and evaluation criteria are demonstrated. The paper was to facilitate the understanding of wavelet-based fusion methods and to provide a review and comparison of the methods by introducing a selection of wavelet transforms and a group of different fusion rules. These methods were also analyzed and discussed according to a series of evaluation criteria. And some overall comments were given regarding the advantages and limitations of wavelet-based fusion.

Although the results cannot disclose an universal principle for various applications with one specific wavelet function, decomposition level, fusion rule, and evaluation criterion, some comments from previous experiments are listed below:

1) There is not one wavelet family or one specific wavelet function, which could be general for various applications. But, for most cases, the wavelet functions having smaller length work better than these functions with longer length with less blurriness.
2) For the decomposition level, the level 2 is the best choice for most cases. Once the decomposition level is higher than 3, the blurriness and the artifacts in the fused results will be largely increased.
3) The proposed rule based outperforms traditional rules such as averaging rule, gradient-based rules, energy-based rules, statistical rules, etc., in terms of subjective assessment and objective criteria.
4) About evaluation criteria, there is not one rule, whose result is consistent with human subjective evaluation. Unfortunately, human perception still acts as a critical tool for the image fusion area.

Different from our previous paper introduced comparisons of different wavelet families, decomposition levels, image fusion rules, and evaluation criteria, this paper focuses on introducing more advanced methods into former platform and make it more generic. More importantly, discussions of wavelet-based image fusion on some specific area such as multi-spectral, medical imaging, inferred images are shown. Experimental results show that the proposed fusion rule still outperforms than these algorithms in most areas. However, for

978-1-4673-7723-2/15 $31.00 © 2015 IEEE
DOI 10.1109/IMCCC.2015.280

some multi-spectral areas, such as medical diagnosis, other algorithms works better. This phenomenon demonstrates that different algorithms are appropriate for various applications, which reminds people to select proper algorithm for their specific application. Further, we are still exploring whether there is one evaluation criterion consistent with subjective evaluation or not. Therefore, a group of objective evaluation criteria are introduced to assess the quality of fused images and evaluate the performance of different methods.

II. Wavelet-based Fusion Platform

So far, plenty of image fusion algorithms have been proposed by researchers, but there is not an universally accepted standard method emerged for diverse applications or environments. As same as other subareas in image processing, a quantitative study and comparison is significant to help people initially selecting an appropriate method in their specific application. Therefore, a platform assessing the performance of the image fusion algorithm has been built, which includes diverse images, different decomposition methods due to various wavelet functions, various fusion rules, and several evaluation criteria (Fig. 1). The DWT with one of 77 wavelet functions is first applied on each source image to generate a fusion decision map, which is generated based on a set of fusion rules (18 rules in our platform), where the fused wavelet coefficient map can be constructed from the wavelet coefficients of the source images according to the fusion decision map. Finally, the fused image is obtained by performing the inverse DWT. Then, the performance evaluation will be carried out with 12 evaluation criteria.

A. Wavelet Functions

The wavelets families used in the platform are: Haar, Daubechies (dbN, N = 120), Symlets (symN, N = 120), Coiflets (coifN, N = 15), Biorthogonal (bior (M, N), M = 16, N = 19), Reverse Biorthogonal (rbior (M, N), M = 16, N = 19), and Discrete Meyer (Dmey). In total, we have 77 wavelet functions.

B. Fusion Rules

The fusion rule is the most important step in the image fusion. Therefore, various alternatives are included in the platform and compared in this paper:

1) Averaging fusion rule for low-frequency images, and for high-frequency,
   a) maximum-selection rule selects the pixel with larger absolute value [6].
   b) the pixel having larger standard variance based on neighbors is used [16].
   c) the pixel having larger energy based on neighbors is used [14].
   d) the pixel having larger standard deviation based on neighbors is used [2].
   e) the pixel having more number of larger values is used [17].
2) Gradient-based fusion rule [3]: For low-frequency subimages, averaging fusion rule is applied. For high-frequency subimages,
   a) the pixel-gradient-based method will use the gradient value of the pixel.
   b) the region-gradient-based one uses the weight based on neighbors.
   c) the global-gradient-based one is calculated from all high-frequency bands.
3) For low-frequency subimage, the pixel from a source image having lower energy is used, and for high-frequency subimages, the pixel from a source image having larger energy is used [12].
4) For low-frequency subimage, weighted averaging rule based on the energy is used, and for high-frequency subimages, the value of a pixel from a source image having larger absolute is used [13].
5) Activity measurement: it is based on the statistical quantities of local windows of size 3×3 [1].
6) Pixel significance: it considers the coefficient at the current decomposition level and all his children and grandchildren coefficients [5].
7) Our proposed rule based on focus measure and local contrast measure [4].
8) For low-frequency subimage, weighted averaging rule based on the spatial frequency is used, and for high-frequency subimages, the value of a pixelhaving larger absolute is used [10].
9) The modified feature selection algorithm [8] : The variance of each image patch is computed as activity measure.
10) The region entropy is used to measure the amount of information from the approximation images contributing to the fused result [13].
11) A new averaging fusion rule: averaging fusion rule for low-frequency subimages, and for high-frequency subimages, information is fused based on average gradient (AG) and mutual information (MI) [9].
12) Multiscale products (MSP) [7]: a feature contrast measurement of the multi-scale products is used to select coefficients from the sharpness parts of the high-frequency subbands.

In all, there are 18 rules integrated in our platform, where the first 13 rules marked from 1a to 7 were stated in the previous paper [4] and discussed their performance under different images. In this paper, another 5 rules labelled from 8 to 12 are introduced in this new platform.

C. Objective Evaluation Criteria

It is a challenge to assess the performance of image fusion as the ground-truth is not available in most of cases. For exhaustive study, several classical evaluation criteria shown in different literatures [1] [5] are used in this paper, which are listed as follows:

1) Mean [5]: average pixel intensity measures image brightness.

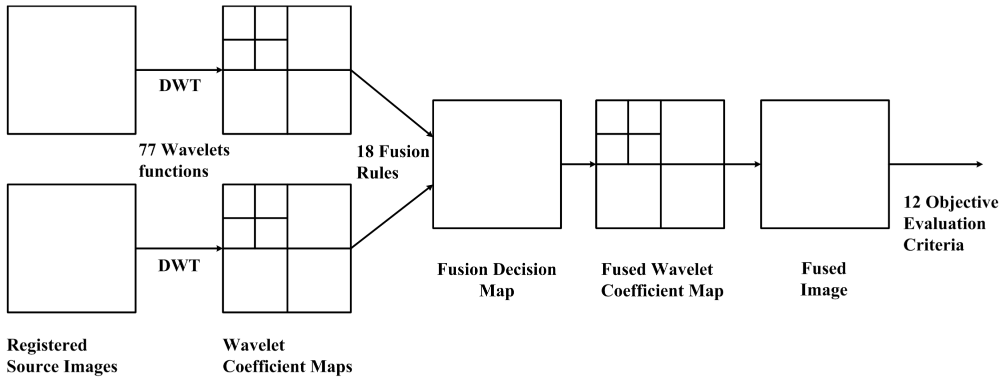

Fig. 1. A wavelet-based image fusion platform.

2) Standard deviation (STD) [5]: reflects image contrast.
3) Root-mean-square (RMSE) [1] measures of the differences between the fused image and the source image.
4) Peak signal to noise ratio (PSNR) [1]: a term for the ratio between the maximum value of a signal and the power of distorting noise.
5) Average gradient (AG) [5] measures a degree of clarity and sharpness.
6) Entropy (H) [5] estimates the amount of information presenting in the image.
7) Mutual information (MI) [5] quantifies the overall mutual information between source images and fused image.
8) Correlation coefficient(CC) [5] measures a relevance of input and output.
9) Spatial frequency(SF) [5] measures the overall information level in the regions (activity level) of an image.
10) A performance measure ρ [8] is defined as the standard deviation of the difference image between the ideal image and the fused image.
11) Structural similarity (SSIM) [15]: an quality assessment metric comparing local patterns of pixel intensities from luminance, contrast, and structure.
12) QABF and LABF measures [11]: measures localized preservation of input gradient information and provides an analysis of fusion performance by quantifying total fusion gain, fusion loss, and fusion artifacts.

III. Experimental Tests

Experimental tests are carried out on various standard test pairs of multi-focus, medical, and multi-spectral images. Due to limited pages, a part of results are shown, where one example represent one kind of areas.

A test strategy is designed according to the following conditions. As we stated in our previous paper, there is not one wavelet family or one specific wavelet function is general for various applications. But, for most cases, the wavelet functions having smaller length work better than these functions with longer length. The smaller length also means fewer calculations during convolution. In our tests, Db2 works better than Db3, whose results are shown in results. For the decomposition level, the results suggest the level 2 may be the best choice. Once the decomposition level is higher than 3, the blurriness and the artifacts in the fused results will be largely increased. A few images work well at level 1 or 3, but the level 2 shows the best result for most images and wavelet functions. Therefore, results with the DL 2 are only demonstrated in this paper, which is superior to the DL 3 during our tests due to less blurriness even for those medical and multi-spectral images.

When comparing with different fusion rules, only 4 rules are selected from these rules stated in the previous paper ranging from rule 1a to 7, for they demonstrated good effects or own better performance in the category to which they belonging, such as 1a belonging to group 1. Therefore, the total 9 algorithms: 1a, 2c, 6, 7, 8, 9, 10, 11, and 12 are compared in our tests, where these algorithms 1a, 2c, 6, and 7 were compared in other tests in the previous paper.

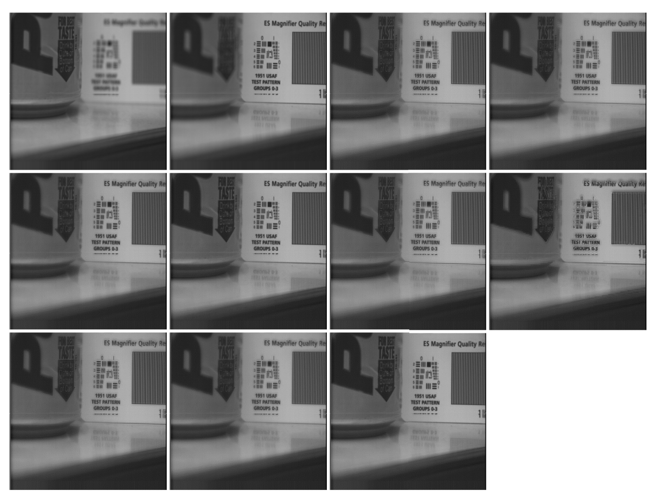

Fig. 2. The results of "Pepsi" (Db2 and DL = 2) with different fusion rules: (top-row-left) first source image; (-second) second source image; (-third) 1a; (-right) 2c; (second-row-left) 6; (-second) 7; (-third) 8 (-right) 9; (bottom-row-left) 10; (-middle) 11; and (-right) 12.

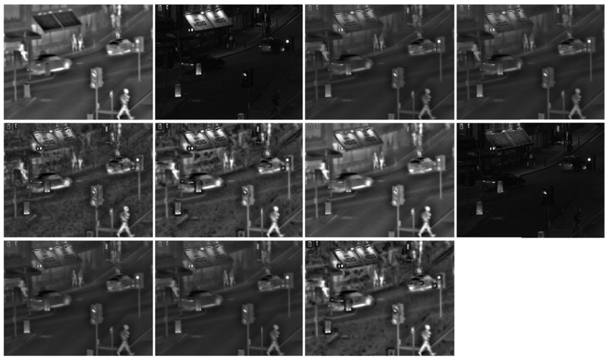

Fig. 3. The results of "Street" (Db2 and DL = 2) with different fusion rules: (top-row-left) infrared image; (-second) visible image; (-third) 1a; (-right) 2c; (second-row-left) 6; (-second) 7; (-third) 8 (-right) 9; (bottom-row-left) 10; (-middle) 11; and (-right) 12.

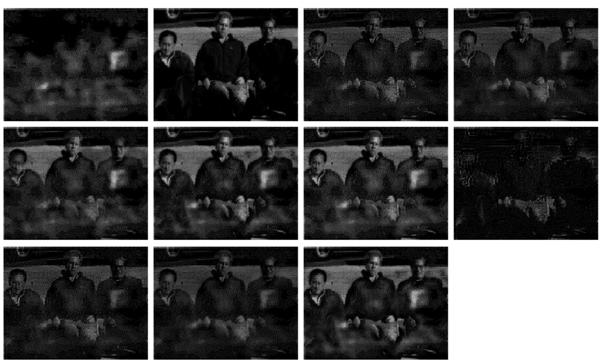

Fig. 4. The results of "Gun" (Db2 and DL = 2) with different fusion rules: (top-row-left) MMW image; (-second) visible image; (-third) 1a; (-right) 2c; (second-row-left) 6; (-second) 7; (-third) 8 (-right) 9; (bottom-row-left) 10; (-middle) 11; and (-right) 12.

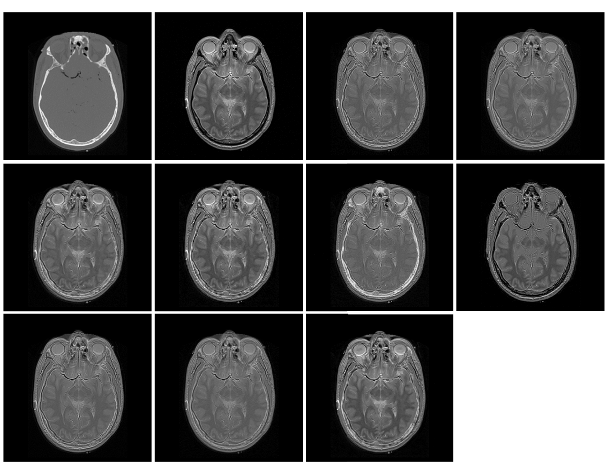

Fig. 5. The results of "medical image" (Db2 and DL = 2) with different fusion rules: (top-row-left) CT image; (-second) MR image; (-third) 1a; (-right) 2c; (second-row-left) 6; (-second) 7; (-third) 8 (-right) 9; (bottom-row-left) 10; (-middle) 11; and (-right) 12.

A. Fusion of Infrared and Visible Images

These algorithms also applied to the infrared (IR) and Visible images. Fig. 3 shows an infrared image, in which the contour of persons or something else could be easily detected, but lights and characters above awnings could not be seen. In visible images, the persons could not be detected, but lights and characters are clearly demonstrated. For these kind of images, the algorithm 8 always performs well by combining different source image together, in which person, lights, or characters are visible with clare edges than other methods. The algorithms 1a, 2c, 10, and 11 have fair results without obvious errors but low brightness and contrast. The algorithms 6, 7, and 12 pay more attention to local gradient values, which introduce lots of errors. The algorithm 9 is the worst one, which makes the bright awnings almost invisible.

B. Fusion of Multi-focus Images

These fusion algorithms are applied to a set of multi-focus images, such as "pepsi" in Fig. 2. The proposed algorithm 7 and E are very close with less blurriness and artifacts, since they are designed based on local contrast measure, but the multi-scale product doesn't demonstrate any performance improvement. The algorithm 6 is the one having a lightly more blurriness than them, but it still works well. The rest images generated by other algorithm have more errors, except the algorithm 9. It brings lots of errors and artifacts.

C. Fusion of Infrared and Visible Images

These algorithms also applied to the infrared (IR) and Visible images. Fig. 3 shows an infrared image, in which the contour of persons or something else could be easily detected, but lights and characters above awnings could not be seen. In visible images, the persons could not be detected, but lights and characters are clearly demonstrated. For these kind of images, the algorithm 8 always performs well by combining different source image together, in which person, lights, or characters are visible with clare edges than other methods. The algorithms 1a, 2c, 10, and 11 have fair results without obvious errors but low brightness and contrast. The algorithms 6, 7, and 12 pay more attention to local gradient values, which introduce lots of errors. The algorithm 9 is the worst one, which makes the bright awnings almost invisible.

D. Fusion of MMW and Visible Images

The effectiveness of these algorithms can be also proved by extending them to the application of millimeter wave (MMW) image and visible image. Fig. 4 shows images with a group of people or a pistol concealed by the third person. Thus, the ideal fused result should make the pistol easily identified in the third person.

As stated in the image fusion test between the infrared and the visible image, the algorithm 8 always performs well by combining different source image together, in which person and pistol are visible, also without introducing any obvious errors. The algorithms 6, 7, and 12 come next with some errors. The algorithms 1a, 2c, 10, and 11 have fair results

Rule	Mean	STD	RMSE	PSNR	AG	H	MI	Cor	SF	ρ	SSIM	QABF	LABF
7	1	4	1	1	1	4	2	1	6	2	1	1	5
8	7	2	2	2	1	2	1	1	1	2	2	1	2
12	0	2	1	0	0	1	2	1	0	1	0	5	0
Total	8	8	4	3	2	7	5	3	7	5	3	7	7
9	0	1	0	1	4	1	2	0	3	3	0	0	0

TABLE I
COMPARISON BETWEEN OBJECTIVE AND SUBJECTIVE EVOLUTIONS

without obvious errors but very low brightness. The algorithm 9 still acts bad with a black image.

E. Fusion of CT and MR Images

For medical images, the computed tomography (CT) data and the magnetic resonance (MR) scan provide different information with bone and dense tissue with less distortions and soft tissue with more distortions, such as shown in Fig. 5.

For these medical images, the algorithm 8 always performs well by keeping more completed dense tissue. However, the algorithms 7 and 12 work better in soft tissues. The algorithm 6 shows a good balance between dense tissues and soft tissues but with low brightness or contrast. The algorithms 1a, 2c, 10, and 11 are better at preserving soft tissues, where dense tissues is not kept well. The algorithm 9 isn't a good option in the medical images.

F. Performance evaluation and analysis

The performance of the fusion rules are also compared with the objective fidelity criteria measures as we stated above. From a subjective point of view, rule 7, 8 and 12 are considered perform better, while rule 9 performs the worst. In the Table II, the first row shows different objective measurements, and the number means the count of each measurements that perform the best in 9 groups . For example, rule 8 has 7 times performs the highest in measurement mean. Whilst, rule 12 never perform the best, so the count of measurement mean is zero. The total row means the sum of rule 7, 8 and 12. Since we have 9 groups, the sample size is 9. From the objective point of view, when measurement mean and STD perform the best, we have 88.89% to say that the rule also perform best from the subjective point of view. In addition, when measurement AG perform the best, we have 44.45% to say that the rule perform the worst.

IV. CONCLUSIONS

Different from our previous paper introduced comparisons of different wavelet families, decomposition levels, image fusion rules, and evaluation criteria, this paper focuses on introducing more advanced fusion algorithms and objective evaluation criteria into former platform and make it more generic. More importantly, discussions of wavelet-based image fusion on some specific area such as multi-spectral, medical imaging, and inferred images demonstrate that our proposed algorithm still has excellent performance under most of various conditions, but some kind of algorithm might be more appropriate for certain application, which reminds people to carefully select algorithm for their specific application.

Due to limited pages, a brief discussion about the consistence between subjective and objective evaluations are demonstrated, which reveals that combining different measures together might be more effective than a single measure. For example, the measures mean and STD have higher probabilities to point out the proper fusion rule according to our tests. More statistical analysis based on objective evaluation criteria will be further discussed in the future.

ACKNOWLEDGEMENTS

The author wishes to thank the Ministry of Finance 'life Science Instrumentation Development Program' managed by CAS with grant No. ZDYZ2012-3 for the funding, which made this work possible.

REFERENCES

[1] S Arivazhagan, L Ganesan, and TG Subash Kumar. A modified statistical approach for image fusion using wavelet transform. *Signal, image and video processing*, 3(2):137–144, 2009.

[2] Liu Bin and Peng Jiaxiong. Image fusion method based on nonseparable wavelets. *Machine Vision and Applications*, 16(3):189–196, 2005.

[3] Li Hui-hui, Guo Lei, and Liu Hang. Research on image fusion using wavelet transform based on gradient selective rule. *Computer Engineering and Application*, 12:76–78, 2005.

[4] Bo Jiang, Rui Zhang, and Xiao Zhang. A comparative study of wavelet-based image fusion with a novel fusion rule. *International Journal of Machine Learning and Computing*, 2014.

[5] BK Shreyamsha Kumar. Multifocus and multispectral image fusion based on pixel significance using discrete cosine harmonic wavelet transform. *Signal, Image and Video Processing*, 7(6):1125–1143, 2013.

[6] John J Lewis, Robert J O'Callaghan, Stavri G Nikolov, David R Bull, and Nishan Canagarajah. Pixel- and region-based image fusion with complex wavelets. *Information fusion*, 8(2):119–130, 2007.

[7] Huafeng Li, Shanbi Wei, and Yi Chai. Multifocus image fusion scheme based on feature contrast in the lifting stationary wavelet domain. *EURASIP Journal on Advances in Signal Processing*, 2012(1):1–16, 2012.

[8] Hui Li, BS Manjunath, and Sanjit K Mitra. Multisensor image fusion using the wavelet transform. *Graphical models and image processing*, 57(3):235–245, 1995.

[9] Zhaodong Liu, Hongpeng Yin, Yi Chai, and Simon X Yang. A novel approach for multimodal medical image fusion. *Expert Systems with Applications*, 41(16):7425–7435, 2014.

[10] Gonzalo Pajares and Jesus Manuel De La Cruz. A wavelet-based image fusion tutorial. *Pattern recognition*, 37(9):1855–1872, 2004.

[11] Vladimir Petrovic and Costas Xydeas. Objective image fusion performance characterisation. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, volume 2, pages 1866–1871. IEEE, 2005.

[12] R Raghavendra and Christoph Busch. Novel image fusion scheme based on dependency measure for robust multispectral palmprint recognition. *Pattern recognition*, 47(6):2205–2221, 2014.

[13] Jing Tian and Li Chen. Adaptive multi-focus image fusion using a wavelet-based statistical sharpness measure. *Signal Processing*, 92(9):2137–2146, 2012.

[14] Hai-Hui Wang. A new multiwavelet-based approach to image fusion. *Journal of Mathematical Imaging and Vision*, 21(2):177–192, 2004.

[15] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *Image Processing, IEEE Transactions on*, 13(4):600–612, 2004.

[16] Chen Ye Wen and Jia Ken Chen. Multi-resolution image fusion technique and its application to forensic science. *Forensic Science International*, 140(2):217–232, 2004.

[17] Kai-Jian Xia, Yu-Feng Yao, Shan Zhong, and Jin-Yi Chang. Image fusion algorithm based on morphology wavelets transform [j]. *Computer Engineering*, 19:081, 2010.