

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

Bo Jiang, Rui Zhang, and Xiao Zhang

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. On the other hand, in this paper, a novel fusion rule based on focus measure and local contrast measure is proposed to analyze the information in images and help selecting appropriate information from different source images to obtain a fused image. Experimental results on various wavelet functions and fusion rules demonstrate that the proposed fusion approach outperforms traditional methods in real images.

Index Terms—Fusion rule, image fusion, local contrast measure, wavelet transform.

I. INTRODUCTION

The computer vision area is to simulate the human vision and the human brain to understand image, then carry out corresponding tasks. Therefore, acquired images from scenes containing all relevant objects in focus are significant for post-processes. However, during imaging, since the camera lens could only focus at a certain distance, some objects in the scene are out of focus and, thus, are blurred and not sharp enough. One solution to overcome this problem is image fusion, which could allow the integration of different information sources with different focus settings to extend the depth of field of an image.

So far, several fusion algorithms have been proposed [1], [2], which are mainly performed at different levels of information representation: pixel-level, feature-level, and decision-level [3]. In pixel-level fusion, the data are combined at the pixel-level of the source images. In the feature-level fusion, the features are extracted from each source data and combined. In the decision-level fusion, the fusion is performed on decisions determined by such as statistical analysis, inference, or classification.

Manuscript received September 30, 2014; revised August 25, 2015. This work is supported by the Ministry of Finance “Life Science Instrumentation Development Program”, managed by Chinese Academy of Sciences with grant No. ZDYZ2012-3.

Bo Jiang is with the Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou 510530, China (e-mail: jiang.bo@gibh.ac.cn).

Rui Zhang is with Sun Yat-sen University, China (e-mail: rayzhangbm@gmail.com).

Xiao Zhang is with the Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou 510530, China (e-mail: zhang_xiao@gibh.ac.cn).

Currently, one widely used framework of image fusion works in multi-resolution analysis, which is generally generated by pyramid transform, wavelet transform, curvelet transform, etc. The discrete wavelet transform (DWT) is the most popular and important multi-scale decomposition method in image fusion, for it could provide more details in spatial domain and has shown a comparative better performance than other fusion methods [4].

A general procedure of the wavelet-based transform includes three steps [5]: first, decompose source images by applying a wavelet transform to obtain wavelet coefficients at different levels; then, combine wavelet coefficients at each decomposition level by a certain fusion rule; and, apply an inverse wavelet transform to obtain a fused image.

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. The results could be, then, used as a reference for selecting the fusion method for image fusion. Therefore, the purpose of this paper is 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 are further analyzed and discussed according to a series of evaluation criteria. And some overall comments are given regarding the advantages and limitations of wavelet-based fusion.

On the other hand, for an image fusion algorithm based on DWT, one of the essential things to improve fusion quality is the selection of fusion rules, which influences the performance of fusion algorithm remarkably [5]. Generally, the concerns of the fusion rules are different for low- and high-frequency subimages. For low-frequency coefficients, whether a coefficient at different decomposition levels in-focus or out-of-focus determines the clarity of a fused image. So, a novel local feature detection method-average energy of Laplacian (AEL) is proposed to be as the focus measure to select the coefficients from the clear regions of the source images.

According to physiological and psychological research, the human vision system (HVS) is highly sensitive to the local image contrast. So, inspired by the local image contrast level measure [6], a novel local contrast measure (LCM) is proposed to select high-frequency coefficients by measuring the ratio of the high-frequency component of image to the local luminance of the background in this paper. The experimental results show that the proposed rule based on AEL and LCM outperforms traditional rules in terms of subjective assessment and objective criteria.

The rest of this paper is organized as follows: Section II

briefly introduce the wavelet-based image fusion. Section III explains the proposed fusion rules. To assess the quality of fused images and evaluate the performance of different methods, a group of objective evaluation criteria are introduced in Section IV. Experimental results and performance analysis are demonstrated in Section V. Section VI presents the conclusion.

II. WAVELET-BASED IMAGE FUSION

A. Multi-resolution Decomposition

The DWT is a tool for multi-resolution analysis which could be extended from one-dimensional (1-D) DWT to two-dimensional (2-D) DWT to handle 2-D image signals, then, used to represent image variations at different scales and orientations. During decomposing, the processes in the DWT include recursive filtering and subsampling. The decomposition will produce four subimages representing four frequency bands: one approximation image Low-Low (LL), and three detail images, such as Low-High (LH), High-Low (HL) and High-High (HH). The LL represents the low-frequency band, which is a smooth subimage of the multi-scale decomposition. Generally, it can be considered as a coarse and sub-sampled approximation of original image. The LH, HL, and HH subimages represent high-frequency bands, which describe the details in the horizontal, vertical, and diagonal orientations of the input image.

After one-level decomposition, the multi-level DWT could further decompose these subimages. Generally, the next level decomposition is only applied to the LL band of the current decomposition stage, which will form a recursive decomposition procedure. Thus, N -level decomposition will finally generate $3N+1$ frequency bands, which include $3N$ high-frequency bands and only one low-frequency band. The 2-D DWT will have a pyramid structure shown in the Fig. 1, which shows the structures of 2-D DWT with three decomposition levels. The frequency bands in higher decomposition levels will have smaller size due to subsampling [7].

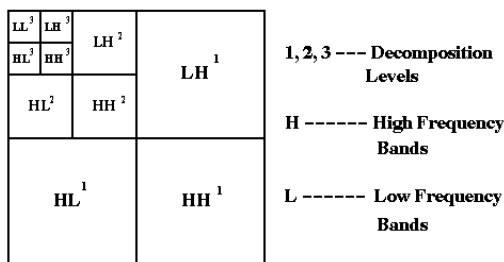


Fig. 1. Pyramid hierarchy of decomposition based on 2-D DWT.

B. Wavelet-Based Image Fusion

A generic wavelet-based image fusion scheme is shown as follows in Fig. 2 [7]. The DWT is first applied on each source image to generate a fusion decision map is generated based on a set of fusion rules, 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.

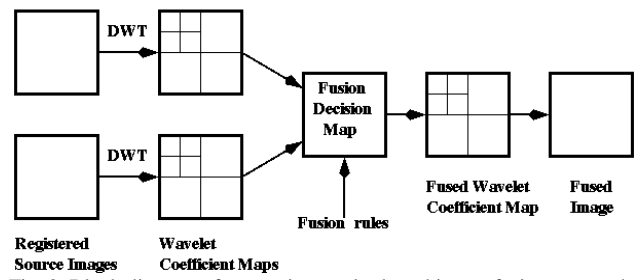


Fig. 2. Block diagram of a generic wavelet-based image fusion approach.

III. PROPOSED FUSION TECHNIQUE

As shown in Fig. 2, the key step in image fusion based on wavelet is that of coefficient combination, namely, the process of merge the coefficients in an appropriate way in order to obtain the best quality in the fused image [5]. An appropriate image fusion rule should preserve all the salient features of the source images and introduces as less artifacts or inconsistency as possible. In addition, the fusion algorithm should be reliable and robust to imperfections such as noise. In this article, we develop a novel multi-focus image scheme to incorporate into the image fusion technique.

As described in Section II, the decomposed coefficients are composed by the low-frequency and high-frequency coefficients, which are generally dealt with separately. For example, an averaging fusion rule tends to merge the mean of the coefficients from source images, but this rule would blur images and reduce the contrast of features appearing in a fused image. Or, a maximum-selection fusion rule selects the largest absolute coefficient at each location from the input images as the coefficient at that location in the fused image. However, an image feature is not only composed and determined by a pixel.

In this paper, we propose the following fusing rules applied to low- and high-frequency subimages:

- 1) For low-frequency subimages, a novel focus measure based on average energy of Laplacian (AEL) is proposed to select clear coefficients.
- 2) For high-frequency subimages, a new local contrast measure (LCM) is proposed to select sharp coefficients.

A. Selection of Low-Frequency Subimage Coefficients

Since the coefficients in the low-frequency subimage represent the approximation component of the source images, a focus measure checking whether a coefficient is in-focus or out-of-focus is necessary. Various focus measures, e. g. gradient-based measure, wavelet-based measure, statistics-based measure, etc., have been proposed to measure the variation of pixels [8], where Laplacian-based measure (LBM) can provide a better performance than others in most cases. In this paper, a new measure-average energy of Laplacian (AEL) based on LBM is proposed to select coefficients from the clear parts of the source images.

The original expression of the energy of Laplacian (EOL) [10] is shown in Equation:

$$EOL(x, y) = (f_{vertical}''(x, y) + f_{horizontal}''(x, y))^2$$

$$f_{vertical}''(x, y) = 2f(x, y) - f(x, y-1) - f(x, y+1) \quad (1)$$

$$f''_{\text{horizontal}}(x, y) = 2f(x, y) - f(x-1, y) - f(x+1, y),$$

where f'' represents the second-order derivative of the pixel at position (x, y) of image f , which could represent different frequency subbands, such as the approximation or the detail of the source images.

However, the EOL only pays attention to vertical and horizontal orientations. From our former work [9], [10], all four orientations including vertical, horizontal, 45 degree, and 135 degree, have to be examined to fully detect and localize image features. Therefore, a modified sum of Laplacian or second-order derivatives of image f is computed as:

$$IEOL(x, y) = (|f''_{\text{vertical}}(x, y)| + |f''_{\text{horizontal}}(x, y)| + |f''_{45}(x, y)| + |f''_{135}(x, y)|)^2. \quad (2)$$

In general, a focus measure will be only performed in the being processed pixel. However, as we stated above, the image feature will be not only decided by a single pixel. Therefore, the region-based AEL are used to be a clarity measure to determine whether a coefficient is in-focus or not. The proposed region-based AEL fusion rule is stated below:

$$AEL(x, y) = \sum_m^M \sum_n^N W(m, n) IEOL(x+m, y+n). \quad (3)$$

where $W(m, n)$ is a weight kernel, and the sum of the kernel should be one to normalize the region area. In this paper, the kernel size is set as 3×3 . To highlight the center pixel of the window, the kernel is set as:

$$W(m, n) = \frac{1}{16} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 8 & 1 \\ 1 & 1 & 1 \end{bmatrix}. \quad (4)$$

The AEL is used to as a focus measure to determine which coefficient is in focus, where pixels with larger values are more possible in-focus. Then, the proposed rule can select in-focus information from source images and put them into the fused image. The proposed AEL-based fusion rule is shown below:

$$f_{A,F}^L(x, y) = w_{A,1} f_{A,1}^L(x, y) + w_{A,2} f_{A,2}^L(x, y), \quad (5)$$

where

$$w_{A,1} = AEL_{A,1}^L(x, y) / (AEL_{A,1}^L(x, y) + AEL_{A,2}^L(x, y)), \quad (6)$$

$$w_{A,2} = AEL_{A,2}^L(x, y) / (AEL_{A,1}^L(x, y) + AEL_{A,2}^L(x, y)), \quad (7)$$

where $f_1(x, y)$ and $f_2(x, y)$ represent different source images. A represents the approximation image of the source image after L th decomposition. L denotes the L th decomposition level. w_1 and w_2 represent different weights for associated source images, respectively. F means the fused image.

B. Selection of High-Frequency Subimage Coefficients

The coefficients in the high-frequency subimages describe the detail component of the source images, which are generally considered as representing salient features such as

edges, boundaries, etc. In general, during fusion, the coefficients in the fused image are the coefficients having larger absolute value from each high-frequency subimage.

But, as we stated above, the image features are determined by a group of pixels [9], [10]. Therefore, the value of a single pixel is not robustly to determine which pixel belongs to the sharp part of the subimage. On the other hand, the physiological study shows that the HVS is highly sensitive to the local image contrast level, which could be defined as the ratio of the local graylevel variation and the local brightness of the background [6].

Therefore, the LCM is proposed based on this physiological study in this paper. The local graylevel variation will be denoted by the region-based AEL, for it can effectively represent the salient features or sharp boundaries of an image. Pixels with larger values of the AEL are more possible in-focus. And, the local brightness of the background is represented as the low-frequency component in the low-frequency subimages. The local brightness of the background will be represented as the mean of a 3×3 area.

$$f_A^L(x, y) = \frac{1}{M+N} \sum_m^M \sum_n^N f_A^L(x+m, y+n), \quad (8)$$

$$LCM_D^L(x, y) = \frac{AEL_D^L(x, y)}{f_A^L(x, y)}. \quad (9)$$

where the local are size $m \times n$ is set as 3×3 . D represents the detail image in the horizontal, vertical, or diagonal orientation of the original image after L th decomposition. A represents the approximation image of the image after L th decomposition. L denotes the L th decomposition level.

Therefore, the proposed selection rule for the high-frequency coefficients in associated subimages is the same as the low-frequency coefficients shown above, except that the values of AEL should be computed based on every detail images, which is shown below:

$$f_{D,F}^L(x, y) = w_{D,1} f_{D,1}^L(x, y) + w_{D,2} f_{D,2}^L(x, y), \quad (10)$$

where

$$w_{D,1} = LCM_{D,1}^L(x, y) / (LCM_{D,1}^L(x, y) + LCM_{D,2}^L(x, y)), \quad (11)$$

$$w_{D,2} = LCM_{D,2}^L(x, y) / (LCM_{D,1}^L(x, y) + LCM_{D,2}^L(x, y)), \quad (12)$$

where a larger value of the local contrast measure means more high frequency information. Consequently, the proposed rule is to extract more useful details from the source image having larger local contrast and put them into the fused image.

IV. 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 [11], [12] are used in this paper, which are listed as follows:

- 1) Mean [12]: average pixel intensity measures image brightness.
- 2) Standard deviation (STD) [12]: square root of the

- graylevel variance reflects image contrast.
- 3) Root-mean-square (RMSE) [11] measures of the differences between the fused image and the source image.
 - 4) Peak signal to noise ratio (PSNR) [11]: a term for the ratio between the maximum value of a signal and the power of distorting noise.
 - 5) Average gradient (AG) [12] measures a degree of clarity and sharpness.
 - 6) Entropy (H) [12] estimates the amount of information presenting in the image.

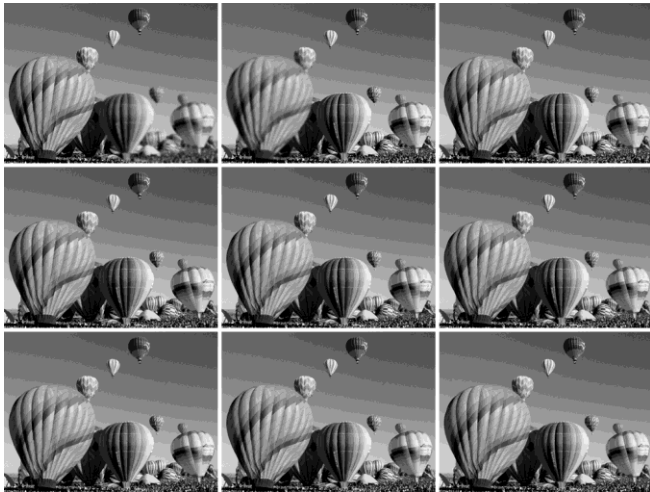


Fig. 3. The best results of “Balloon”: (top-row-left) first source image; (top-row-middle) second source image; (top-row-right) haar wavelet, DL = 1; (second-row-left) db2 wavelet, DL = 1; (second-row-middle) sym3 wavelet, DL = 1; (second-row-right) coif1 wavelet, DL = 1; (bottom-row-left) bior1.3, DL = 1; (bottom-row-middle) rbior1.3 wavelet, DL = 1; and (bottom-row-right) dmey wavelet, DL = 1.



Fig. 4. The best results of “Table”: (top-row-left) first source image; (top-row-middle) second source image; (top-row-right) haar wavelet, DL = 2; (second-row-left) db20 wavelet, DL = 1; (second-row-middle) sym2 wavelet, DL = 1; (second-row-right) coif1 wavelet, DL = 2; (bottom-row-left) bior1.5, DL = 2; (bottom-row-middle) rbior1.5 wavelet, DL = 2; and (bottom-row-right) dmey wavelet, DL = 2.

V. EXPERIMENTAL RESULTS

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, which includes diverse images, different decomposition methods, various fusion rules, and several evaluation criteria, is an exciting research. Experimental tests are carried out on various standard test pairs of multi-focus, medical, and multi-spectral images acquired from imagefusion.org. In this paper, more attention is paid to an ideal image pair “Balloon”, two multi-focus pairs “Table” and “Flower”, and one multi-spectral pair “Gun” due to limited pages. Also, due to the same reason, the rest of the paper mainly focuses on the performance of different wavelet families, decomposition levels (DL), fusion rules, and evaluation criteria in this paper.



Fig. 5. The best results of “Flower”: (top-row-left) first source image; (top-row-middle) second source image; (top-row-right) haar wavelet, DL = 2; (second-row-left) db2 wavelet, DL = 2; (second-row-middle) sym3 wavelet, DL = 2; (second-row-right) coif1 wavelet, DL = 2; (bottom-row-left) bior1.3, DL = 2; (bottom-row-middle) rbior1.3 wavelet, DL = 2; and (bottom-row-right) dmey wavelet, DL = 2.

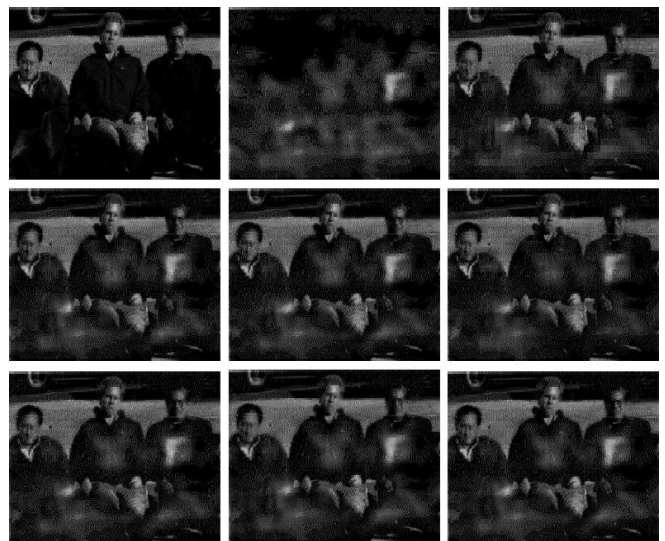


Fig. 6. The best results of “Gun”: (top-row-left) first source image; (top-row-middle) second source image; (top-row-right) haar wavelet, DL = 3; (second-row-left) db2 wavelet, DL = 3; (second-row-middle) sym3 wavelet, DL = 3; (second-row-right) coif1 wavelet, DL = 4; (bottom-row-left) bior2.2, DL = 3; (bottom-row-middle) rbior1.3 wavelet, DL = 3; and (bottom-row-right) dmey wavelet, DL = 3.

A test strategy is designed according to the following steps:

A. Perform a Comparative Study for Different Wavelet Families

The wavelets families used in our study are: Haar, Daubechies (dbN, $N = 1 \dots 20$), Symlets (symN, $N = 1 \dots 20$), Coiflets (coifN, $N = 1 \dots 5$), Biorthogonal (bior (M, N), $M = 1 \dots 6, N = 1 \dots 9$), Reverse Biorthogonal (rbior (M, N), $M = 1 \dots 6, N = 1 \dots 9$), and Discrete Meyer (Dmey). We first select the member of each family with the best results by human perception and this member is used as the representative member for this family. Then, different wavelet families could be compared against other methods. When comparing different wavelet families, the paper only focuses on their performance in our proposed rules. More discussions on their effects for different image fusion rules may be introduced in later work. Besides different wavelet functions, the best decomposition level (DL) for different wavelet functions and families is also examined.

TABLE I: "BALLOON" BEST RESULTS FOR DIFFERENT WAVELET FUNCTION AND DECOMPOSITION LEVEL

Name	DL	Mean	STD	RMSE	PSNR	AG	H
haar	1	113.37	47.93	15.89	60.81	9.494	7.47
db2	1	112.90	48.01	15.95	60.74	9.505	7.48
sym3	1	113.38	47.94	15.88	60.83	9.51	7.47
coif1	1	113.37	47.94	15.90	60.81	9.51	7.47
bior1.3	1	113.37	47.95	15.90	60.80	9.50	7.47
rbior1.3	1	113.37	47.93	15.89	60.82	9.51	7.47
dmey	1	112.95	48.05	15.96	60.74	9.53	7.48

TABLE II: "TABLE" BEST RESULTS FOR DIFFERENT WAVELET FUNCTION AND DECOMPOSITION LEVEL

Name	DL	Mean	STD	RMSE	PSNR	AG	H
haar	2	97.47	46.44	23.06	54.84	7.66	7.29
db20	1	97.35	46.69	23.68	54.52	7.73	7.31
sym2	1	97.49	46.56	23.60	54.46	7.77	7.31
coif1	2	97.45	46.42	23.05	54.85	7.67	7.29
bior1.5	2	97.43	46.43	23.11	54.78	7.70	7.29
rbior1.5	2	97.46	46.41	23.03	54.87	7.66	7.30
dmey	2	97.44	46.39	23.02	54.90	7.66	7.29

TABLE III: "FLOWER" BEST RESULTS FOR DIFFERENT WAVELET FUNCTION AND DECOMPOSITION LEVEL

Name	DL	Mean	STD	RMSE	PSNR	AG	H
haar	2	113.41	51.17	23.74	53.62	6.25	7.42
db2	2	113.40	52.21	23.82	53.54	6.16	7.42
sym3	2	113.36	52.24	23.83	53.55	6.15	7.42
coif1	2	113.38	52.21	23.79	53.59	6.17	7.42
bior1.3	2	113.38	52.26	23.90	53.50	6.27	7.43
rbior1.3	2	113.37	52.17	23.76	53.60	6.13	7.41
dmey	2	113.40	52.27	23.82	53.55	6.13	7.42

TABLE IV: "GUN" BEST RESULTS FOR DIFFERENT WAVELET FUNCTION AND DECOMPOSITION LEVEL

Name	DL	Mean	STD	RMSE	PSNR	AG	H
haar	3	36.55	35.56	82.53	31.70	22.56	6.41
db2	3	36.66	36.19	83.06	31.60	23.63	6.29
sym3	3	36.55	36.27	83.12	31.58	23.88	6.28
coif1	4	34.76	34.87	79.66	32.30	23.06	6.26
bior2.2	3	36.35	36.13	82.84	31.64	23.96	6.26
rbior1.3	3	36.55	35.84	82.93	31.61	22.08	6.38
dmey	3	36.50	36.37	83.32	31.53	23.68	6.30

The results of "Balloon", "Table", "Flower", and "Gun" representing the best result of each wavelet family are shown in Figs. 3, 4, 5, and 6. The subjective (visual) analysis is included in the performance evaluation step, for the goal of image fusion is to enhance comprehensive, accurate, and stable information to make that the fused image is more

suitable for human perception. There are three criteria that are used widely in the literature: (1) information transferred from each individual image to fused image, (2) information lost from the source images, and (3) artifacts introduced due to fusion. Besides subjective analysis, the objective performance measures stated in Section IV are shown in Tables I, II, III, and IV for associated figures. Generally, higher value of these metrics, the quality of the fused image is considered as better, except the RMSE. In these tables, the highest values are bolded, except that the lowest values of the RMSE are bolded.

For these ideal images no matter tested in our experiments or in other researchers' work, they are created by locally smoothing by some filters, such as Gaussian filter based on one reference image. A group of images from one reference image have different, not overlapped, filtered areas. Therefore, they could be used to test the algorithm of the image fusion. But, for these ideal images, the differences among different wavelet families, wavelet lengths, and decomposition levels are hardly to be discerned by human vision (Fig. 3), although there are very small differences in their objective evaluation criteria in Table I. For example, for "Balloon" image, dmey owns the highest value in standard deviation, gradient, and entropy, and sym3 has the best performance in the rest criteria.

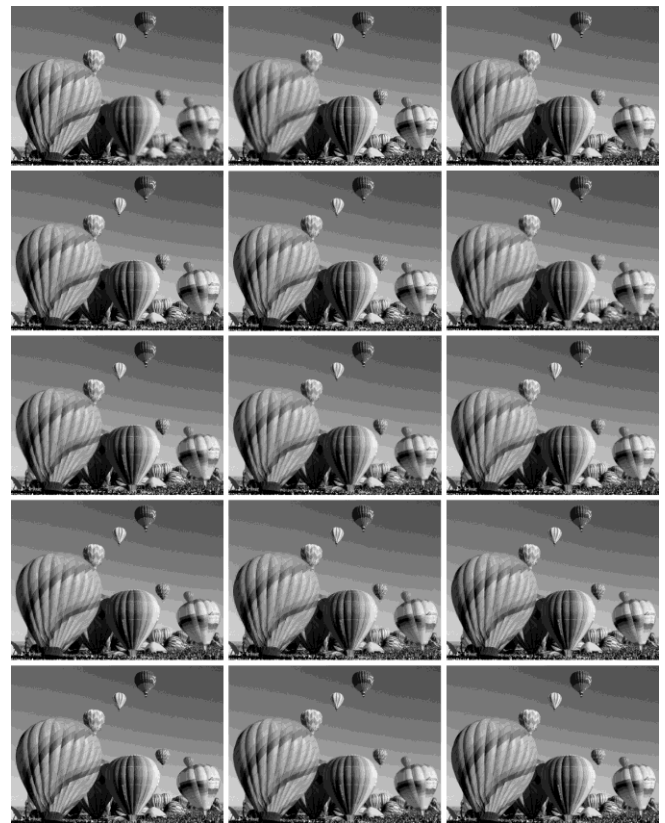


Fig. 7. The best results of "Balloon" (haar wavelet and DL = 2) with different fusion rules: (top-row-left) 1a; (-middle) 1b; (-right) 1c; (second-row-left) 1d; (-middle) 1e; (-left) 2a; (third-row-middle) 2b; (-left) 2c; (-middle) 3; (-right) 4; (bottom-row-left) 5; (-middle) 6; and (-right) 7.

When entering into the area with real images, a challenge here is that the size of objects are changed due to focus varying. Therefore, firstly, the fusion algorithm should carefully work at the boundaries of objects to reduce abrupt changes, such as artifacts, since these boundaries have been

extended from a simple edge, like a line signal, to an intensity ramp profile [10]. Secondly, for those flat areas, the fusion algorithm has to keep the local information intact, without information mingled from other sources.



Fig. 8. The best results of “Table” (haar wavelet and DL = 2) with different fusion rules: (top-row-left) 1a; (-middle) 1b; (-right) 1c; (second-row-left) 1d; (-middle) 1e; (-left) 2a; (third-row-middle) 2b; (-left) 2c; (-middle) 3; (-right) 4; (bottom-row-left) 5; (-middle) 6; and (-right) 7.

Different from ideal image, different wavelet functions and decomposition levels significantly affect the results of the image fusion for real images, such as “Table”, “Flower”, “Gun”, etc.

In Fig. 4, Db20 has the best performance by human perception and the highest STD (Table II). However, sym2 has the best results in Mean, AG, and H criteria. For RMSE and PSNR, dmey owns better performance than others. However, the artifacts in its result are more evident than others.

In Fig. 5, it is hard to determine the better wavelet functions by human perception, although their blurriness and artifacts are different. Also from Table III, there does not exist any wavelet function, which could demonstrate the best performance in all objective criteria.

The image “Gun” is used to test the algorithm for multi-spectral fusion. The result of bior2.2 is the best by human vision in Fig. 6. The db02 also works well. The result of haar wavelet is much worse than bior2.2, but its H is the best in Table IV. The output of bior2.2 has good performance in AG. The result of coif1 is close to the result of bior2.2, which also has best performance shown in RMSE, and PSNR.

The results stated above suggest that all evaluation criteria shown in this paper may work well for some kinds of images, but may not for other kinds. As discussed above, one wavelet function shows good performance judged by human vision,

but its visual quality may not be considered as the best one by statistical results. For different images, the wavelet function having the best performance by subjective evaluation or objective evaluation is different. The worst part is that the wavelet family could be kept as the same for different images. So far, a conclusion could not be drawn that 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. 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.

B. Comparison of Different Fusion Rules

As we stated in the Section III, the fusion rule is an important step in the image fusion. So, various alternatives are compared in this paper, which are as follows:

- 1) Averaging fusion rule for low-frequency subimages, and for high-frequency subimages:
 - a) maximum-selection fusion rule is used to select the pixel with larger absolute value [13].
 - b) The pixel from a source image having larger standard variance calculated based on its 3×3 neighbors is used to be the value in the fused image in the same location [14].
 - c) The pixel from a source image having larger energy calculated based on its 3×3 neighbors is used to be the value in the fused image in the same location [15].
 - d) The pixel from a source image having larger standard deviation calculated based on its 3×3 neighbors with their absolute values is used to be the value in the fused image in the same location [16].
 - e) The pixel from a source image having more number of larger values, e.g. 5, based on its 3×3 neighbors with their absolute values is used to be the value in the fused image in the same location [17].
- 2) Gradient-based fusion rule: it includes different variations such as pixel-, region- and global-gradient-based rule [18]. For low-frequency subimages, averaging fusion rule is applied to combine the information from all source images. For high-frequency subimages,
 - a) The pixel-gradient-based method will use the weighted value based on the gradient value of the pixel in the same location from source image HH bands.
 - b) The only difference in the region-gradient-based method is that it uses the weight based on the 3×3 neighbors of the being processed pixel.
 - c) The global-gradient-based method is almost the same as the region-gradient-based method, except that the weights will be calculated from all three high-frequency bands, not only HH bands.
- 3) For low-frequency subimage, the pixel from a source image having lower energy calculated based on its 3×3 neighbors is used to be the value in the fused image in the same location, and for high-frequency subimages, the pixel from a source image having larger energy

calculated based on its 3×3 neighbors is used to be the value in the fused image in the same location [19].

- 4) For low-frequency subimage, weighted averaging rule based on the energy of the 3×3 neighbors from different source images is used, and for high-frequency subimages, the value of a pixel from a source image having larger absolute is used to be the value in the fused image in the same location [20].
- 5) Activity measurement: it is based on the statistical quantities of local windows of size 3×3 . For low-frequency subimage, the pixel from a source image having larger mean and standard deviation calculated based on its 3×3 neighbors is used to be the value in the fused image in the same location, and for high-frequency subimages, the pixel from a source image having larger standard variance calculated based on its 3×3 neighbors is used to be the value in the fused image in the same location [11].
- 6) Pixel significance: it adds the coefficient at the current decomposition level and all his children and grandchildren coefficients produced during higher-level decompositions in the same location together from different source images to do weight averaging for all low- and high-frequency bands [12].
- 7) The proposed rule based on the AEL and the LCM.

In all, all 13 fusion rules are compared with various images.



Fig. 9. The best results of “Flower” (haar wavelet and DL = 2) with different fusion rules: (top-row-left) 1a; (-middle) 1b; (-right) 1c; (second-row-left) 1d; (-middle) 1e; (-left) 2a; (third-row-middle) 2b; (-left) 2c; (-middle) 3; (-right) 4; (bottom-row-left) 5; (-middle) 6; and (-right) 7.

In our experimental tests of different fusion rules, the wavelet function and the DL are set as the same. For example, the wavelet function will be set as haar, db2, db6, etc. The DL

is set as 2, which is the optimal one achieved in the former section. Then, all the fusion rules are compared based on their results generated from different images.



Fig. 10. The best results of “Gun” (haar wavelet and DL = 3) with different fusion rules: (top-row-left) 1a; (-middle) 1b; (-right) 1c; (second-row-left) 1d; (-middle) 1e; (-left) 2a; (third-row-middle) 2b; (-left) 2c; (-middle) 3; (-right) 4; (bottom-row-left) 5; (-middle) 6; and (-right) 7.

The group one has different variants. The 1a is the most popular one and is considered as one simple and general rule, which could achieve good effects in image fusion for most images. During our tests, all these variants in the group one have almost same performance with good fusion performance except generating artifacts and causing blurriness. However, the rule 1d cannot successfully fuse image and should be not considered as one rule option in the wavelet-based image fusion. The rule 1b should be not considered as an appropriate rule too, since it will bring obvious errors such as artifacts even for ideal fusion images, such as “balloon”. For all other methods, their results are almost the same for ideal fusion images without any obvious artifacts. Therefore, the 1a could be still considered as a representative in this group. Comparing with these methods in the group one, the propose rule could further reduce artifacts and produce more smooth transitions at boundaries but less blurriness.

For those gradient-based fusion rules in the group two, the region- and global-based rules are obviously better than the pixel-based rule due to higher contrast and sharpness, and lower blurriness. The region- and global-based are very close. The global-gradient-based is slightly better in most cases, but it has more blurriness and artifacts than the proposed rule 7.

Both rules 3 and 4 consider energy as a parameter to select an appropriate coefficient. In comparisons, their results have more artifacts than our proposed rule 7. The rule 3 generates

lots of artifacts in some images above average results. Therefore, it is also not recommended as a rule in the fusion.

The advanced rules 5 and 6 are more complicated algorithms than other rules shown above. The rule 5 is more like a combination with some basic rules in group one, but it doesn't further improve the overall performance. It will improve the contrast in the boundaries, but the artifacts are also higher than traditional methods. The rule 6 generally generates little more artifacts and more blurriness than the proposed rule 7. But, for some images, the rule 6 is better than our proposed rule 7 due to less artifacts.

TABLE V: "BALLOON" BEST RESULTS FOR DIFFERENT FUSION RULES WITH HAAR WAVELET AND DECOMPOSITION LEVEL 2

Rule	DL	Mean	STD	RMSE	PSNR	AG	H
1a	2	113.23	47.47	15.74	60.87	9.64	7.47
1b	2	113.23	47.37	15.77	60.79	9.48	7.47
1c	2	113.23	47.42	15.73	60.89	9.53	7.47
1d	2	113.23	46.09	15.67	60.83	6.7	7.45
1e	2	113.23	47.25	15.68	61.00	9.52	7.46
2a	2	113.22	46.68	13.69	62.86	7.92	7.45
2b	2	113.22	47.07	14.26	62.58	8.91	7.46
2c	2	113.22	46.99	14.01	62.87	8.82	7.46
3	2	112.99	46.80	16.84	59.50	9.60	7.45
4	2	113.23	47.50	15.73	60.88	9.64	7.47
5	2	113.40	47.79	16.15	60.35	9.56	7.47
6	2	113.28	47.14	13.74	63.37	8.69	7.46
7	2	113.31	47.67	15.39	61.39	9.40	7.48

TABLE VI: "TABLE" BEST RESULTS FOR DIFFERENT FUSION RULES WITH HAAR WAVELET AND DECOMPOSITION LEVEL 2

Rule	DL	Mean	STD	RMSE	PSNR	AG	H
1a	2	97.90	45.52	22.01	54.90	8.12	7.27
1b	2	97.90	45.31	22.00	54.92	7.56	7.26
1c	2	97.90	45.43	21.96	54.99	7.82	7.27
1d	2	97.90	43.97	22.24	54.55	4.95	7.19
1e	2	97.90	45.28	22.10	54.75	7.83	7.26
2a	2	97.89	44.76	20.33	56.17	6.40	7.23
2b	2	97.88	44.93	20.30	56.33	6.65	7.24
2c	2	97.88	44.99	20.34	56.28	6.76	7.24
3	2	95.63	45.86	25.99	51.85	8.29	7.27
4	2	98.15	45.46	22.12	54.79	8.12	7.27
5	2	97.64	46.10	24.08	53.61	7.85	7.30
6	2	97.90	45.32	20.97	55.96	6.85	7.26
7	2	97.47	46.44	23.06	54.84	7.66	7.29

TABLE VII: "FLOWER" BEST RESULTS FOR DIFFERENT FUSION RULES WITH HAAR WAVELET AND DECOMPOSITION LEVEL 2

Rule	DL	Mean	STD	RMSE	PSNR	AG	H
1a	2	113.22	50.65	20.50	55.88	6.61	7.39
1b	2	113.23	50.38	20.52	55.84	5.89	7.38
1c	2	113.22	50.60	20.49	55.89	6.38	7.38
1d	2	113.23	49.81	20.52	55.84	4.49	7.36
1e	2	113.22	50.47	20.49	55.88	6.36	7.38
2a	2	113.21	50.22	19.22	57.00	5.05	7.37
2b	2	113.21	50.31	19.31	56.93	5.24	7.37
2c	2	113.21	50.38	19.56	56.69	5.53	7.37
3	2	109.58	50.81	25.90	51.78	7.06	7.39
4	2	113.78	50.53	20.75	55.67	6.63	7.39
5	2	113.25	51.79	24.05	53.18	6.42	7.42
6	2	113.25	50.53	20.28	56.20	5.50	7.38
7	2	113.40	52.17	23.74	53.62	6.25	7.41

Partial experimental results are shown in Figs. 7, 8, 9, and 10. Their statistical results are demonstrated in Tables V, VI, VII, and VIII, which show that the best fusion rule 7 doesn't achieve the best performance based on all statistical criteria. In all, during image fusion, a balance has to be always determined among details, contrast, and artifacts during design. So far, the same as other areas in computer vision,

human perception still acts as a critical tool for image fusion. Therefore, a more appropriate criteria based on the local area information to assess the performance of image fusion technologies is an important task to be explored in the future.

TABLE VIII: "GUN" BEST RESULTS FOR DIFFERENT FUSION RULES WITH HAAR WAVELET AND DECOMPOSITION LEVEL 2

Rule	DL	Mean	STD	RMSE	PSNR	AG	H
1a	3	26.78	27.59	61.25	36.85	26.3	5.77
1b	3	26.35	26.05	59.95	37.21	24.5	5.85
1c	3	26.63	27.19	61.09	36.89	25.2	5.78
1d	3	25.31	20.91	55.70	38.47	11.9	6.03
1e	3	26.58	26.88	60.75	36.98	24.8	5.79
2a	3	25.48	23.06	56.52	38.22	14.7	5.98
2b	3	25.89	24.21	56.07	38.38	21.8	5.82
2c	3	26.08	25.06	57.44	37.96	22.9	5.80
3	3	14.50	20.94	51.38	40.13	19.9	4.39
4	3	40.31	36.87	86.59	30.93	28.1	6.13
5	3	29.30	31.97	72.85	33.81	24.4	5.93
6	3	36.35	31.87	75.70	33.27	23.9	6.17
7	3	36.55	35.56	82.53	31.70	22.6	6.41

VI. CONCLUSIONS

The image fusion technique is going to integrate complementary information from different source images together to provide more complete information for post-processing, which has been applied in diverse areas, such as remote sensing, astronomy, medical imaging, security, surveillance, etc. During fusing, a tradeoff between preserving salient features and reducing artifacts has to be carefully balanced. In this paper, more discussions are demonstrated on wavelet-based image fusion.

Since each wavelet-based fusion algorithm has its own set of advantages and limitations, more comprehensive testing in order to fully assess under what conditions each algorithm is more appropriate is significant. Therefore, in this paper, a platform of wavelet-based image fusion has been built, which includes different wavelet families, image fusion rules, and statistical evaluation criteria, which could be used to assess, and more importantly, select appropriate method for a specific fusion task. In this paper, a review and comparison of different wavelet families, decomposition levels, image fusion rules, and evaluation criteria is provided. The future work will include introducing more methods into current platform and make it more generic. More importantly, discussions of wavelet-based image fusion on some specific area might discover some more solid regularity.

Besides the comparison, a new fusion rule for the image fusion based on DWT is proposed to measure the extent of clarity and local contrast, where two proposed measures AEL and LCM are designed based on detecting image features. During fusing, the weighted averaging fusion scheme is used to reduce the artifact occurring at the boundaries of objects. Experimental results show that the proposed fusion rule outperforms than tradition rules. The future work will include developing new fusing scheme instead of the weighted averaging fusion to further improve the local contrast of image, but suppress the artifacts.

ACKNOWLEDGMENT

The author wishes to thank the Ministry of Finance 'Life

Science Instrumentation Development Program' managed by Chinese Academy of Sciences with grant No. ZDYZ2012-3 for the funding, which made this work possible.

REFERENCES

- [1] H. Wonsook, P. H. Gowda, and T. A. Howell, "A review of downscaling methods for remote sensing-based irrigation management: Part I," *Irrigation Science*, vol. 31, no. 4, pp. 831-850, 2013.
- [2] H. Wonsook, P. H. Gowda, and T. A. Howell, "A review of downscaling methods for remote sensing-based irrigation management: Part II," *Irrigation Science*, vol. 31, no. 4, pp. 851-869, 2013.
- [3] X. Y. Luo, J. Zhang, and Q. H. Dai, "A classification-based image fusion scheme using wavelet transform," in *Proc. SPIE 8064 of Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications*, 2011.
- [4] A. Krista, Y. Zhang, and P. Dare, "Wavelet based image fusion techniques — An introduction, review and comparison," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 62, no. 4, pp. 249-263, 2007.
- [5] P. Gonzalo and J. M. Cruz, "A wavelet-based image fusion tutorial," *Pattern Recognition*, vol. 37, no. 9, pp. 1855-1872, 2004.
- [6] T. Alexander, L. J. V. Ruyven, and J. M. Valetton, "Merging thermal and visual images by a contrast pyramid," *Optical Engineering*, vol. 28, no. 7, pp. 287789-287789, 1989.
- [7] Z. Zhang and R. Blum, "Region-based image fusion scheme for concealed weapon detection," in *Proc the 31st Annual Conference on Information Sciences and Systems*, 1997.
- [8] P. Said, D. Puig, and M. A. Garcia, "Analysis of focus measure operators for shape-from-focus," *Pattern Recognition*, vol. 46, no. 5, pp. 1415-1432, 2013.
- [9] B. Jiang, "Real-time multi-resolution edge detection with pattern analysis on GPU," *Journal of Real-Time Image Processing*, 2014.
- [10] B. Jiang, "Multi-resolution edge detection with edge pattern analysis," *Proc. SPIE 8667, Multimedia Content and Mobile Devices*, 86670W, March 7, 2013.
- [11] S. Arivazhagan, L. Ganesan, and T. G. Subash Kumar, "A modified statistical approach for image fusion using wavelet transform," *Signal, Image and Video Processing*, vol. 3, no. 2, pp. 137-144, 2009.
- [12] S. B. K. Kumar, "Multifocus and multispectral image fusion based on pixel significance using discrete cosine harmonic wavelet transform," *Signal, Image and Video Processing*, vol. 7, no. 6, pp. 1125-1143, 2013.
- [13] J. J. Lewis, R. J. O'Callaghan, S. G. Nikolov, D. R. Bull, and N. Canagarajah, "Pixel-and region-based image fusion with complex wavelet," *Information Fusion*, vol. 8, no. 2, pp. 119-130, 2007.
- [14] C. Y. Wen and J. K. Chen, "Multi-resolution image fusion technique and its application to forensic science," *Forensic Science International*, vol. 140, no. 2, pp. 217-232, 2004.
- [15] H. H. Wang, "A new multiwavelet-based approach to image fusion," *Journal of Mathematical Imaging and Vision*, vol. 21, no. 2, pp. 177-192, 2004.
- [16] B. Liu and J. Peng, "Image fusion method based on nonseparable wavelets," *Machine Vision and Applications*, vol. 16, no. 3, pp. 189-196, 2005.

- [17] K. Xia, Y. Yao, S. Zhong, and J. Chang, "Image fusion algorithm based on morphology wavelets transform," *Computer Engineering*, vol. 36, no. 19, pp. 224-226, 2010.
- [18] H. Li, L. Guo, and H. Liu, "Research on image fusion using wavelet transform based on gradient selective rule," *Computer Engineering and Application*, vol. 12, 2005.
- [19] R. Raghavendra and B. Christoph, "Novel image fusion scheme based on dependency measure for robust multispectral palmprint recognition," *Pattern Recognition*, vol. 47, no. 6, pp. 2205-2221, 2014.
- [20] J. Tian and C. Li, "Adaptive multi-focus image fusion using a wavelet-based statistical sharpness measure," *Signal Processing*, vol. 92, no. 9, pp. 2137-2146, 2012.



Bo Jiang received a B.S. degree in information engineering from Chengdu University of Technology, China, in 2004, a M.S. degree in communication and information engineering from the University of Electronic Science and Technology of China, China, in 2007, and a Ph.D. degree in electrical and computer engineering from Old Dominion University, USA, in 2010.

He is currently an associate investigator of the Division of Applied Stem Cell Automation at Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou, Guangdong, China.

His research interests include bioinformatics, generic pattern recognition, and real-time, non-linear image processing.



Rui Zhang was born in Guangzhou, Guangdong, China on Aug. 4, 1992. He is an undergraduate student in Sun Yat-sen University, major in statistics, minor in finance.

He is currently a research assistant in the Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou, China. His research interests include image fusion and wavelet transform.



Xiao Zhang received a B.S. degree in biochemistry from the University of Sheffield, United Kingdom, in 2005, a Ph.D. degree in medical school from the University of Sheffield, United Kingdom, in 2008.

He is currently a principle investigator of the Division of Applied Stem Cell Automation at Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou, Guangdong, China.

His research interests include cell line engineering with instrumentation development for biotherapeutics application in an automated manner under GMP for clinical use hiPSCs application development for HCS/HTS drug screening and clinical application cell signaling in innate immunity and inflammation for cardiovascular application.