

# Medical image fusion quality evaluation method combined multi-angle and multi-scale information

Manyu Jia<sup>1</sup>, and Feng Yang<sup>1,2,\*</sup>

<sup>1</sup>School of Computer and Electronics and Information, Guangxi University, Nanning, China

<sup>2</sup>Guangxi Key Laboratory of Multimedia Communications Network Technology, Guangxi University, Nanning, China

**Abstract.** At present, the existing evaluation indicators of image fusion algorithms do not in conformity with the standards of human vision system and are not suitable for medical image application scenarios. Therefore, this paper proposes a medical image fusion quality evaluation method combining multi-angle and multi-scale information. We use Gabor Filter to decompose the image and obtain multi-angle texture information, and design a new index, called Multiscale Structural Similarity with Image Resolution (SSIM-IR), to comprehensively evaluate the fused image from edge, structure and sharpness. The experimental results show that this method is more suitable for medical image application scenarios, consistent with subjective evaluation, and can be used for measuring the advantages and disadvantages of medical image fusion algorithm.

**Keywords:** Gabor filter, Medical image fusion, Fusion quality evaluation, Multiscale structural similarity with image resolution.

## 1 Introduction

Multi-modal medical image fusion can roundly and accurately reflect the structure and function of human organs and tissues and integrate the advantages of different modal medical images, so as to help doctors make more accurate diagnosis of the disease. With the development of image fusion technology, how to evaluate the performance of image fusion algorithm has become a new research direction. Since there is no standard image for medical fusion image to compare, the quality evaluation for medical fusion image algorithm is a difficult point.

The early objective evaluation indicators based on the pixel features of the image lack the consideration of source images, and are mainly unreferenced image evaluation indexes. In addition, there are evaluation indicators based on information theory and structural similarity. At present, the most common evaluation indicators are based on Human Visual System (HVS), which evaluate the quality of fused images by simulating the perception of images by Human eyes. However, the existing evaluation indicators based on HVS has the

---

\* Corresponding author: [fyang@foxmail.com](mailto:fyang@foxmail.com)

limitations that its result can not completely consistent with subjective feelings. It is an inevitable trend to improve subjective perception of fusion images to make objective evaluation indexes more consistent with subjective visual perception. Therefore, many researchers are inclined to combine multiple indicators to obtain a comprehensive indicator. Currently, there is no evaluation method for medical image fusion algorithm.

In this paper, a medical image fusion quality evaluation method combining multi-angle and multi-scale information is proposed, which comprehensively considers edge, structure, brightness, contrast and sharpness, and can be used to evaluate gray and color image fusion algorithms. The experimental results show that the proposed evaluation method with good performance can be used to evaluate medical image fusion algorithms, provide results in accordance with human vision standards.

## 2 Multiscale structural similarity with image resolution

In order to better evaluate the quality of fused Image, this paper design a new index called Multiscale Structural Similarity with Image Resolution (SSIM-IR).

The Structural Similarity (SSIM)<sup>[1]</sup> combines the mean value, variance and covariance of images to evaluate the brightness, contrast and structure between a source image and fusion image without considering the sharpness of images, which is a semi-reference index.

The Average Gradient (AG) can measure the clarity of an image. Let  $\alpha_X$  and  $\alpha_Y$  be the average gradient of source image X and fusion image Y respectively. The sharpness influence factor  $R(X, Y)$  can reflect the sharpness change from source image to fusion image:

$$R(X, Y) = \frac{2\alpha_Y - \alpha_X}{\alpha_Y + \alpha_X} \quad (2)$$

The formula of SSIM-IR can be expressed as:

$$SSIMIR(X, Y) = SSIM(X, Y) \times R(X, Y) \quad (3)$$

In order to integrate all source images and fusion images for evaluation, SSIM-IR further combines the image structure information to measure the structural similarity sharpness between the two source images and the generated fusion images.

The principle of SSIM points out that the variance and covariance of image can reflect structural information. According to the definition of Mutual Information (MI)<sup>[2]</sup>, the weight value of structural information between image X and Y is defined as:

$$\lambda_{XY} = \sigma_X + \sigma_Y - \sigma_{XY} \quad (4)$$

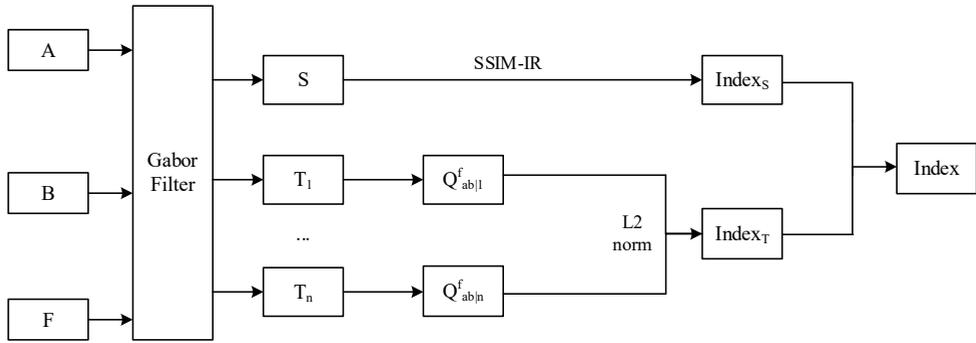
Set source image A, B and fusion image F, and the index formula of SSIM-IR is:

$$SSIM-IR = \frac{\lambda_{AF}}{\lambda_{AF} + \lambda_{BF}} \times SSIMIR(A, F) + \frac{\lambda_{BF}}{\lambda_{AF} + \lambda_{BF}} \times SSIMIR(B, F) \quad (6)$$

## 3 Method

Suppose A and B are medical images obtained by different imaging principles or different parameters of the same imaging principle for the same lesion structure respectively, and F

represents the fusion image of image A and B. The specific construction steps are as follows:



**Fig. 1.** Evaluation index construction framework (S,  $T_1$  and  $T_n$  Represent structure image set and multi-angle texture image set respectively).

### 3.1 Pre-processing

In medical image fusion algorithm, YUV space is frequently YUV used in color image and gray image fusion. When evaluating the fusion algorithm of color and gray medical images, YUV transformation is performed on the color source image and fusion image first, and the y-component obtained is taken as the input.

### 3.2 Image decomposition

Gabor Filter<sup>[3]</sup> is sensitive to the edge of the image and can provide good direction selection and scale selection characteristics. Moreover, the frequency and direction expression of Gabor Filter are similar to that of human vision system. Thus, Gabor Filter can extract texture details from different angles to obtain texture images and decompose structure images to achieve multi-angle and multi-scale decomposition of images.

The operation of image decomposition: decompose source image A, B and fusion image F by Gabor Filter at different angles to obtain texture image  $T_\theta^A, T_\theta^B$  and  $T_\theta^F$ ,  $\theta \in U = \{0^\circ, 22.5^\circ, 45^\circ, 67.5^\circ, 90^\circ, 112.5^\circ, 135^\circ, 157.5^\circ\}$ , and structural image  $S^A, S^B$  and  $S^F$ .

### 3.3 Multi-angle texture information evaluation

$Q^{AB/F}$  [4] is an objective non-reference evaluation index of fusion image quality, and its algorithm uses local edge measurement to estimate the performance degree of significant information from the source image in the fusion image. The multi-angle texture images contain edge information from source images and fusion image.

Thus, use  $Q^{AB/F}$  to calculate the edge gradient index  $Q^f_{ab|\theta}$  of the texture image of the source image from different angles  $T_\theta^A, T_\theta^B$  and the texture image of the fusion image  $T_\theta^F$ , and then obtain the evaluation score  $Index_T$  of the texture information through L2 norm calculation.

L2 norm, also known as Euclidean norm, is defined as the square root of the sum of squares of all elements of a vector. The value of the edge gradient index can be taken as the element of the vector and then integrated through L2 norm, as follows:

$$Index_T = \sqrt{\sum_{\theta \in U} \|Q_{ab|\theta}^f\|^2} \quad (7)$$

### 3.4 Structural information evaluation

Most of the organ structure, brightness and contrast information is retained in the structural image, which can be evaluated by the SSIM-IR designed in this paper.

Use Formula (6) to calculate the SSIM-IR value and get the evaluation score  $Index_s$  of structural information of the structural image.

### 3.5 Index calculation

Calculate the amount of information of the structure image and texture image decomposed by Gabor Filter to get the final quality evaluation score  $Index$ .

$$Index = \frac{EN_T \times Index_T + EN_S \times Index_s}{EN_T + EN_S} \quad (8)$$

Where  $EN_T$  and  $EN_S$  represent the total Information Entropy (EN) of texture image and structure image respectively.

## 4 Analysis of experimental results

In order to ensure the comprehensiveness of the comprehensive indicators proposed in this chapter, the experiment adopted six indicators for comparison, including SF, EN, FMI<sup>[2]</sup>, VIFF<sup>[5]</sup>,  $Q^{AB/F[4]}$  and MS-SSIM<sup>[1]</sup>. The above six indicators used in the experiment are the larger the value is, the better the fusion image quality is.

To verify the performance of the proposed evaluation indicators, 60 pairs of medical images with different morphologies were collected, including CT, MR-T1, MR-T2, MR-PD, MR-Gad, PET and SPECT. For each pair of medical images, 6 fusion images were obtained by the following 6 algorithms: Laplacian Pyramid fusion(LP), CurVelet Transform fusion(CVT), NonSubsampling Contour Wave Transform fusion(NSCT), Discrete Wavelets fusion (DWT), Convolutional Neural Network fusion (CNN)<sup>[6]</sup> and Pulse Coupled Neural Network based on Nonsubsampled Shearlet Transform (NSST-PAPCNN)<sup>[7]</sup>.

### 4.1 Contrast experiment

In order to verify the effectiveness of the proposed quality assessment method, this paper selected Pearson Linear Correlation Coefficient (PLCC) and Spearman Rank order Correlation Coefficient (SROCC) to evaluate the correlation between subjective evaluation and objective indicators methods<sup>[8]</sup>. The value range of SROCC and PLCC is [-1, 1], which represents the relationship between subjective and objective ranking of the fusion algorithm. The higher the value is, the more consistent the quality evaluation method is with the results of subjective evaluation. The following figure shows PLCC and SROCC values of the proposed method and comparison indexes.

From the histogram of figure 2, the proposed method has the highest PLCC and SROCC values, which means that this method is not only in line with the standard of HVS but also

more suitable for medical image fusion scenes.

### 4.2 Case analysis

In order to further verify the performance of the proposed quality assessment method in medical image fusion between different modes, two pairs of medical images were selected from the experimental dataset for analysis. Table 1 shows the evaluation results of these images by comparing indicators and methods proposed in this paper, including the index value and its rank.

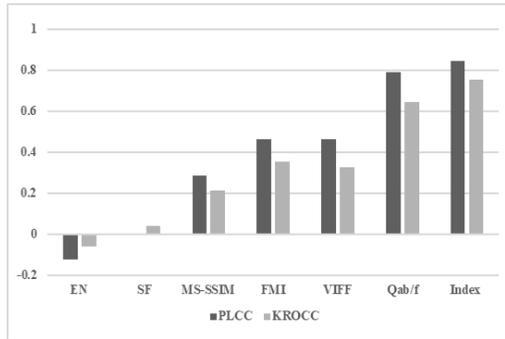


Figure 2. The PLCC and SROCC values chart.

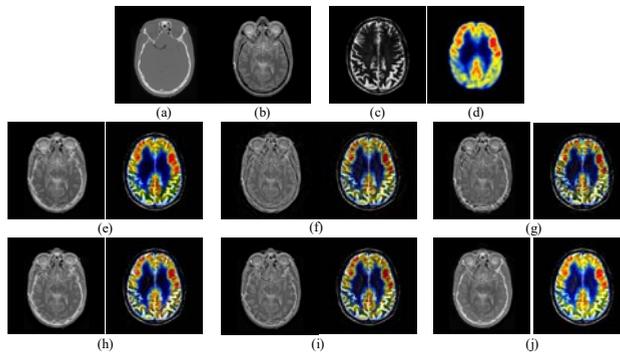


Figure 3. Comparison of image fusion result. (a) and (b), (c) and (d) are CT and MR-T1, MR-T2 and PET images respectively scanned along the axial plane of human brain, and (e) -(j) successively are fusion images obtained by CNN, CVT, DWT, LP, NSCT and NSST\_PAPCNN fusion algorithms

Table 1. Comparison results of different fusion algorithms and evaluation indexes of fusion images.

	Algorithm	rank	SF	EN	FMI	VIFF	$Q^{AB/F}$	MS-SSIM	Proposed
CT and MR-T1 image	CNN	2	20.673(3)	4.9413(6)	0.2803(1)	0.306(3)	0.2439(2)	0.7885(4)	1.1049(3)
	CVT	5	19.325(5)	5.2855(2)	0.2153(6)	0.1878(6)	0.2168(5)	0.786(5)	1.0451(5)
	DWT	6	23.694(1)	5.2708(3)	0.2376(5)	0.245(4)	0.211(6)	0.7705(6)	0.9511(6)
	LP	3	20.807(2)	4.9802(5)	0.2662(2)	0.3307(2)	0.2478(1)	0.7886(3)	1.1183(2)
	NSCT	1	19.866(4)	5.0999(4)	0.248(4)	0.2332(5)	0.242(3)	0.8025(1)	1.1294(1)
	NSST_PAPCNN	4	19.117(6)	5.3329(1)	0.2522(3)	0.343(1)	0.2399(4)	0.7903(2)	1.0642(4)
MR-T2 and PET image	CNN	3	26.118(5)	3.9829(5)	0.7219(1)	0.4784(4)	0.5918(3)	0.7776(4)	1.1287(4)
	CVT	1	27.169(2)	5.1138(1)	0.3161(6)	0.4535(5)	0.575(5)	0.7793(3)	1.1577(1)
	DWT	6	28.308(1)	4.2655(4)	0.6904(3)	0.4387(6)	0.4856(6)	0.7651(6)	1.0223(6)
	LP	4	26.776(4)	3.9447(6)	0.7146(2)	0.5154(2)	0.5975(1)	0.7827(2)	1.1416(3)
	NSCT	2	27.109(3)	4.3003(3)	0.6488(5)	0.4991(3)	0.5972(2)	0.7913(1)	1.1504(2)
	NSST_PAPCNN	5	26.07(6)	4.6291(2)	0.676(4)	0.5488(1)	0.5789(4)	0.7676(5)	1.1286(5)

Through recognition and comparison of fused images in Figure 3, subjective ranking in Table 1 can be obtained, which show in the column of rank. Combining with the value rank of all indicators in Table 1, it can be found that the objective ranking obtained by the proposed method is basically consistent with the subjective evaluation results, while other evaluation indexes deviate from the subjective rank to some extent. It can prove that the proposed method can be used and more suitable for the evaluation of medical image fusion algorithms.

## 5 Conclusion

This paper proposes a medical image fusion quality evaluation method combined multi-angle and multi-scale information. This method utilizes Gabor Filter to achieve multi-angle and multi-scale image decomposition, and design the Multiscale Structural Similarity with Image Resolution (SSIM-IR) index to evaluate structural information. In the experiment, through PLCC and SROCC values and case analysis of multi-modal fusion image results, it can be proved that the proposed evaluation method in this paper performs better than other indicators in the evaluation of medical image fusion algorithm and has more subjective consistency. This method can be used to evaluate multi-modal medical image fusion algorithm and promote the development of medical image fusion technology.

This research is supported by the National Natural Science Foundation of China (61861004, 61762007, 61862006).

## References

1. Zhou W, Simoncelli E P and Bovik A C 2003 Multiscale structural similarity for image quality assessment The Thirty-Seventh Asilomar Conference on Signals (Pacific Grove:IEEE) p 1398-1402
2. Qu G H, Zhang D L and Yan P F 2002 Information measure for performance of image fusion (Electronics Letters vol 38) chapter 7 pp 313-315
3. Daugman J G 1988 Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression (IEEE Transactions on Acoustics, Speech, and Signal Processing vol 36) chapter 7 pp 1169-1179
4. Xydeas C S and Pv V 2000 Objective image fusion performance measure (Military Technical Courier vol 56) chapter 4 pp 181-193
5. Sheikh H R and Bovik A C C 2006 Image information and visual quality (IEEE Transactions on Image Processing vol 15) chapter 2 pp7 430-444
6. Liu Y, Chen X and Cheng J 2017 A medical image fusion method based on convolutional neural networks 2017 20th International Conference on Information Fusion (Xi'an :IEEE) p 1-7
7. Yin M, Liu X N and Liu Y 2018 Medical Image Fusion With Parameter-Adaptive Pulse Coupled Neural Network in Nonsampled Shearlet Transform Domain (IEEE Transactions on Instrumentation and Measurement vol 68) chapter 1 pp 49-64
8. Ma K D, Zeng K and Wang Z 2015 Perceptual quality assessment for multi-exposure image fusion (IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society vol 24) chapter 11 pp 3345.