

Image Fusion Algorithm Assessment based on Feature Measurement

Jiying Zhao and Robert Laganière
School of Information Technology and Engineering
University of Ottawa
800 King Edward Ave.
Ottawa, ON K1N6N5 Canada
E-mail: (jyzhao, laganier)@site.uottawa.ca

Zheng Liu
Institute for Aerospace Research
National Research Council Canada
Montreal Road 1200, Bldg M-14
Ottawa, ON K1A0R6 Canada
E-mail: zheng.liu@ieee.org

Abstract

This paper presents a new metric to assess the performance of the image fusion algorithm. The proposed metric is based on image phase congruency, which provides an absolute measurement of image feature with a numeric value ranging from zero to one. By comparing the local cross-correlation of corresponding phase congruence maps of input images and fusion output, the quality of the fused result is assessable without a perfect reference. The preliminary experimental results on multi-focused image pairs demonstrate the efficiency of this approach.

1 Introduction

Pixel-level image fusion has been investigated in various applications and a number of algorithms have been developed and proposed. A typical fusion operation implemented at the pixel level is illustrated in Figure 1. The success of the post-processing or analysis relies largely on the efficiency of the particular fusion algorithm. However, few authors have addressed the problem of how to assess the efficiency of those algorithms and evaluate the resulting fused images objectively and quantitatively.

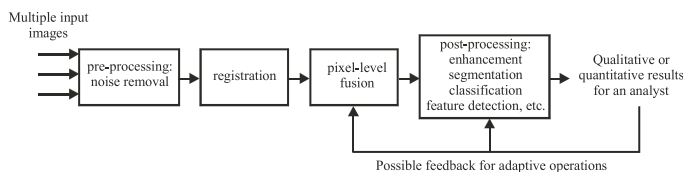


Figure 1. Pixel-level image fusion.

A typical example for pixel-level image fusion is the fusion of multi-focused images from a digital camera [6, 15].

In such case, a cut and paste operation is applied to obtain the full-focused image that will serve as a reference for evaluating the fusion results. However, such operation does not assure a perfect reference image. In some applications, the ground truth reference can be generated from a more precise measurement [3]. Such comparison can only be applied after the acquired images are fully registered, i.e. converted to the same resolution, size, and format. The evaluation metric should be optimized for the image feature. Pixel-by-pixel comparison does not meet the requirement, because in the original image pixels are closely related. Moreover, it would be better if the quantitative evaluation can still be achieved without the presence of reference image. This is the case of most practical applications. The evaluation metric should provide a measurement of how well the information of the inputs is integrated into the output.

In this study, a new fusion quality index is proposed and implemented through the local cross-correlation of the phase congruency maps of the input images. Therefore, the image feature-based measure provides a blind evaluation of the fusion result, i.e. no reference image is needed. This metric takes the advantage of the phase congruency measure which provides a dimensionless contrast- and brightness-invariant representation of image features. The fusion quality index is compared with recently developed blind evaluation metrics. The efficiency of the new metric is verified by the test on the fusion results achieved by several multiresolution pixel-level fusion algorithms.

2 Feature-based Assessment

2.1 Image feature from phase congruency

Gradient-based image feature detection and extraction approaches are sensitive to the variations in illumination, blurring and magnification. The threshold applied needs to be modified appropriately. A model of feature perception named local energy was investigated by Morrone and

Owens [9]. This model postulates that features are perceived at points in an image where the Fourier components are maximally in phase. A wide range of feature types give rise to points of high phase congruency. With the evidence that points of maximum phase congruency can be calculated equivalently by searching for peaks in the local energy function, the relation between the phase congruency and local energy is established, that is [4, 5]:

$$PC(x) = \frac{E(x)}{\sum_n A_n(x) + \varepsilon} \quad (1)$$

$$E(x) = \sqrt{F^2(x) + H^2(x)} \quad (2)$$

where $PC(x)$ is the phase congruency at some location x and $E(x)$ is the local energy function. A_n represents the amplitude of the n^{th} component in the Fourier series expansion. A very small positive constant ε is added to the denominator in case of small Fourier amplitudes. In the expression of local energy, $F(x)$ is the signal with its DC component removed and $H(x)$ is the Hilbert transform of $F(x)$.

To extend the algorithm to images, the one-dimensional analysis is applied to several orientations and the results are combined in different ways. The 2D phase congruency can be expressed as [4]:

$$pc(x) = \frac{\sum_o \sum_n W_o(x) [A_{no}(x) \Delta \Phi_{no}(x) - T_o]}{\sum_o \sum_n A_{no}(x) + \varepsilon} \quad (3)$$

where o denotes the index over orientation. The noise compensation T_o is performed in each orientation independently. By simply applying the Gaussian spreading function across the filter perpendicular to its orientation, the one-dimensional Gabor filter can be extended into two dimensions. The orientation space can be quantified using a step size of $\pi/6$, which results in 6 different orientations. For an extensive discussion of the underlying theory, readers are referred to reference [4, 5].

2.2 Evaluation metric

The comparison of images can be carried out by comparing their corresponding phase congruency features. It is appropriate to evaluate the space-variant features locally and combine them together [13, 14]. Cross-correlation can be one choice to measure the similarity between the feature maps [8]. This value is computed by the zero-mean normalized cross-correlation (ZNCC) [8] in formula 4:

$$ZNCC = \frac{\sum_{m=1}^M \sum_{n=1}^N (A(m, n) - \bar{A})(B(m, n) - \bar{B})}{\sqrt{\sum_{m=1}^M \sum_{n=1}^N (A(m, n) - \bar{A})^2 \cdot \sum_{m=1}^M \sum_{n=1}^N (B(m, n) - \bar{B})^2}} \quad (4)$$

where \bar{A} and \bar{B} are the average value of the two images.

As stated previously, a blind evaluation of the fused image is preferred for practical applications, because a ground truth or a perfect reference is not always available for comparison. The pixel-level fusion is to integrate image features like edges, lines, and region boundaries into one composite image. The success of the fusion algorithm will be assessed by the measure of the image features available in the fused image and those from multiple input sources. Phase congruency is used as the basis for the feature extraction and measurement. The flowchart in Figure 2 presents the procedure to calculate the blind evaluation metric P_{blind} . The phase congruency maps of the input and fused images are firstly calculated. A third feature map M_{pc} is derived by point-by-point maximum selection of the two input maps A_{pc} and B_{pc} , i.e. retaining the larger feature points $M_{pc}(i, j) = \max(A_{pc}(i, j), B_{pc}(i, j))$. The feature map of the fused image, F_{pc} , is then compared to A_{pc} , B_{pc} , and M_{pc} locally. At each sub-block, the cross-correlation values between these maps are computed. The evaluation index P is the average over all the blocks. The flow chart is controlled by three values: A_{pc} , B_{pc} , and F_{pc} , and there is $D_{pc}(k) = A_{pc}(k) + B_{pc}(k)$.

Again, when the local cross-correlation of the phase congruency map is considered, the denominator in Equation 4 might be zero. However, this does not mean a low correlation. The summation of each block is considered in the flowchart of Figure 2: 1) when the input blocks are zero ($D_{pc}(k) = 0$), the correlation result should be one if the fused block is zero too ($F_{pc}(k) = 0$); otherwise the cross-correlation value is set to zero; 2) when the fused block is not zero, the cross-correlation value is computed with either of the input block that is not zero; 3) when both the input and fused blocks are not zero ($D_{pc}(k) \neq 0$), the maximum of the cross-correlation of the fused block with A_{pc} , B_{pc} , and M_{pc} is selected as the result. The overall cross-correlation value is: $P = \frac{1}{K} \sum_{k=1}^K P(k)$, where K is the total number of blocks.

Herein, we use the maximum-selected feature map as part of the reference, because the feature can be compared through the dimensionless measure directly. Unlike pixels which are closely related in the original images, the points in the phase congruency map indicate the salience of image feature. Therefore, the selection of feature points is not equivalent to the selection of pixels with larger value in the original image followed by the computation of the whole phase congruency map. Selecting larger feature points can provide a reference for comparison although this arrangement is not always optimal. For the combinative fusion, especially when heterogeneous sensors are involved, the feature in the fused image may come from input images or a combination of them, as shown in Figure 3. That is why we need both the inputs and maximum-selected map for the

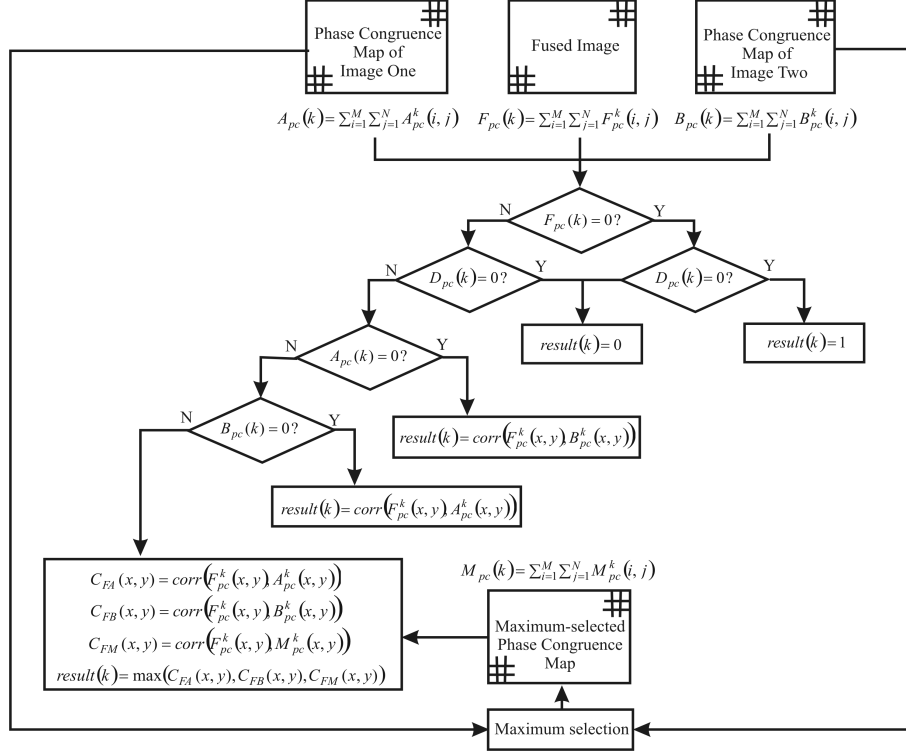


Figure 2. The blind evaluation algorithm by using phase congruency map.

similarity measure and comparison.



Figure 3. Four cases in a combinative fusion. For a small local region in the fused image, the local feature may come from the corresponding block of the input image A or B, or a combination of them.

3 Experimental Results

In the experiment, we use five pairs of multi-focused images for testing¹. The following algorithms are selected for comparison: Laplacian pyramid, gradient pyramid, ratio-of-lowpass (RoLP) pyramid, Daubechies wavelet four, spatially-invariant discrete wavelet transform (SIDWT), and Simoncelli's steerable pyramid. The detailed implementation of these algorithms can be found in refer-

¹Courtesy of SPCR Lab of Lehigh University and Ginfuku at Kyoto.

ences [1, 2, 12, 11, 7]. The basic fusion rule applied is averaging the low-frequency components while selecting the coefficients with larger absolute value in other frequency bands. The decomposition was carried out to level four and four orientational frequency bands were employed in the steerable pyramid implementation.

The fused images were evaluated by Piella's fusion quality index (Q) [10] and blind assessment method. The results are listed in Table 1 and the difference can be much easily identified from the proposed P_{blind} metric.

4 Summary

In this paper, we propose a new metric to assess the performance of multiresolution image fusion algorithms. The metric is based on the local cross-correlation of image feature maps. The measurement of image feature is implemented with Kovessi's phase congruency theory. Both the reference-based evaluation and blind assessment are carried out on the fusion of a group of multi-focused images.

References

[1] E. H. Adelson, C. H. Anderson, J. R. Bergen, P. J. Burt, and J. M. Ogden. Pyramid methods in image processing. *RCA*

Table 1. Evaluation of the fusion result of multi-focused images.

	Assessment metric	Lapacian Pyramid	Gradient pyramid	Ratio-of-lowpass pyramid	Daubechies wavelet four	SIDWT (Haar)	Steerable pyramid
Lab	Q	0.9566	0.9490	0.9178	0.9480	0.9627	0.9574
	P_{blind}	0.8505	0.8293	0.7331	0.8186	0.8466	0.8585
book	Q	0.9474	0.9400	0.9253	0.9365	0.9569	0.9489
	P_{blind}	0.8159	0.7967	0.7094	0.7738	0.8236	0.8268
food	Q	0.9803	0.9506	0.9771	0.9794	0.9829	0.9807
	P_{blind}	0.9098	0.8857	0.8916	0.9045	0.9098	0.9124
Pepsi	Q	0.9667	0.9627	0.9342	0.9607	0.9711	0.9664
	P_{blind}	0.7612	0.7342	0.6691	0.7332	0.7662	0.7936
object	Q	0.9629	0.9450	0.9559	0.9542	0.9693	0.9637
	P_{blind}	0.8534	0.8218	0.8051	0.7986	0.8356	0.8471



Figure 4. The multi-focus images used for the test. From top to bottom: laboratory, books, Japanese food, Pepsi, and object. From left to right: full-focused image, left-focused image, and right-focused image.

Engineer, 29(6):33–41, 1984.

- [2] P. J. Burt and R. J. Kolczynski. Enhanced image capture through fusion. In *Proceedings of 4th International Conference on Image Processing*, pages 248–251, 1993.
- [3] D. S. Forsyth and J. P. Komorowski. NDT data fusion for

improved corrosion detection. In X. E. Gros, editor, *Applications of NDT Data Fusion*. Kluwer Academic Publisher, 2001.

- [4] P. Kovesi. *Invariant Measures of Image Features from Phase Information*. PhD thesis, University of Western Australia, May 1996.
- [5] P. Kovesi. Image features from phase congruency. *Videre: A Journal of Computer Vision Research*, 1(3), 1999.
- [6] H. Li, B. S. Manjunath, and S. K. Mitra. Multisensor image fusion using the wavelet transform. *Graphical Models and Image Processing*, 57(3):235–245, 1995.
- [7] Z. Liu, K. Tsukada, K. Hanasaki, Y. K. Ho, and Y. P. Dai. Image fusion by using steerable pyramid. *Pattern Recognition Letters*, 22:929–939, 2001.
- [8] J. Martin and J. L. Crowley. Experimental comparison of correlation techniques. In *Proceedings of International Conference on Intelligent Autonomous Systems*, 1995.
- [9] M. C. Morrone and R. A. Owens. Feature detection from local energy. *Pattern Recognition Letters*, 6:303–313, 1987.
- [10] G. Piella and H. Heijmans. A new quality metric for image fusion. In *Proceedings of International Conference on Image Processing*, Barcelona, 2003.
- [11] O. Rockinger. Image sequence fusion using a shift-invariant wavelet transform. In *Proceedings of International Conference on Image Processing*, volume 3, pages 288–301, 1997.
- [12] A. Teot. Image fusion by a ratio of low-pass pyramid. *Pattern Recognition Letters*, 9:245–253, 1989.
- [13] Z. Wang and A. C. Bovik. A universal image quality index. *IEEE Signal Processing Letters*, 9(8):81–84, 2002.
- [14] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error measurement to structural similarity. *IEEE Transactions on Image Processing*, 13(1), 2004.
- [15] Z. Zhang and R. S. Blum. Image fusion for a digital camera application. In *Proceedings of 32nd Asilomar Conference on Signals Systems, and Computers*, pages 603–607, Monterey, CA, 1998.