# PERFORMANCE ASSESSMENT OF COMBINATIVE PIXEL-LEVEL IMAGE FUSION BASED ON AN ABSOLUTE FEATURE MEASUREMENT

JIYING ZHAO, ROBERT LAGANIÈRE AND ZHENG LIU

School of Information Technology and Engineering University of Ottawa 800 King Edward Ave., Ottawa, Ontario, Canada K1N 6N5 jzhao@uottawa.ca

Received December 2006; revised May 2007

ABSTRACT. In this paper, a new metric for evaluating the performance of the combinative pixel-level image fusion is defined based on an image feature measurement, i.e. phase congruency and its moments, which provide an absolute measurement of image features. By comparing the local cross-correlation of corresponding feature maps of input images and fused output, the quality of the fused result is assessed without a reference. The experimental results on multi-focused image pairs demonstrate the efficiency of the proposed approach.

Keywords: Image fusion, Phase congruency, Feature measurement

1. Introduction. The interest of using information across the electromagnetic spectrum in an image leads to the technique named image fusion. The purposes of image fusion vary with the applications. Some applications require that the image fusion can generate a composite image, which contains the complementary features available in the multiple input images. One of the state-of-the-art techniques is multiresolution analysis (MRA) based image fusion. The input images are first represented (decomposed) in the multiresolution transform domain. Then, the sub-images (components) or transform coefficients are combined based on certain criterion named fusion rule. The fused image is achieved by the reconstruction from the combined sub-images or coefficients. The fusion process is twofold. One is the multiresolution algorithm and the other is the fusion rule, which guides the coefficient combination in the transform domain. The research focuses on an optimal use of these two. A table was presented in [12] to summarize these techniques. A good overview of the MRA-based image fusion technique can be found in reference [2, 14].

While a diverse range of MRA-based fusion algorithms have been proposed, the objective assessment of the fusion results still remains a challenge, although the evaluation process is an application dependent issue. However, if the MRA-based pixel-level fusion is for feature integration, we could ask a general question like, "how are the features fused?" or "how many features are available in the fused image?". The objective and quantitative evaluation of the efficiency of the fusion algorithms has not been fully explored and addressed so far. The assessment of fused image can be carried out in two ways. When a perfect reference image is available, the fused image can be compared with such reference directly. The methods for image comparison can be employed, for example, root

mean square error (RMSE), cross-correlation (CC), peak signal to noise ratio (PSNR), difference entropy (DE), mutual information (MI), etc. [22]. More recently, Wang et al. proposed a structural similarity (SSIM) index to compare images [21]. The SSIM index is a simple and efficient implementation for image quality measurement. In the case of fusing multi-focused image from digital camera [26], a cut and paste operation was applied to obtain the full-focused image that served as a reference. However, such operation does not assure a "perfect" reference. In some applications, the ground truth can be generated from more precise measurement. Unfortunately, such "prefect" reference is not available in most applications. Therefore, the assessment has to be carried out in the absence of a reference image in most situations.

Qu et al. proposed using the summation of the mutual information values of input and the fused images as an assessment metric [18]. However, the value itself does not make any sense if there is not a reference to compare with. Xydeas et al. measured the amount of visual information transferred from the input images to the fused image [17, 23, 24, 25]. The visual information is defined in terms of edge strength and orientation. The Sobel edge detector is employed to extract such information for each pixel. The Sobel edge detection is based on the measurement of the intensity gradient, which depends on image contrast and spatial magnification. Therefore, one does not know in advance what level of edge strength corresponds to a significant feature. Piella et al. defined a Q metric based on the SSIM index [15, 16]. This metric actually calculates the weighted average of the SSIM between the inputs and fused images. The weighting coefficients do not clearly measure the similarity of the input images to the fused result [4]. Cvejic et al. proposed to use the correlation coefficients as the weighting parameters in [4]. For totally complementary features, the weighting average appears to be a good measurement. However, when the sliding window moves to the region with redundant features, this method may become invalid.

In this paper, a new metric based on an absolute image feature measurement is proposed. Image features are quantitatively defined and the availability of the features in the fused image is calculated as the metric to indicate the efficiency of the fusion algorithm. The rest of the paper is organized as follows. The procedure for MRA-based pixel-level fusion is briefly described in section 2. A image fusion assessment metric is proposed in section 3. The metric is based on the image feature measurement, namely, phase congruency. Experimental results are presented in section 4. Discussion of the results and the conclusion of the paper can be found in section 5 and 6 respectively.

2. Pixel-level Image Fusion. Representing an image in the transform domain makes it easier to access the image features such as edges and boundaries, which are usually represented by the larger absolute values (coefficients) in the high-frequency bands. The fusion rule is to retain the coefficients that correspond to the most salient features. The choice of the kernel function that resembles the shape of the signal will result in a larger convolution coefficient. As the input image is represented as a combination of low-, band- and high-pass sub-images (components). The simplest fusion rule is to average the low-pass components and retain the coefficients with larger absolute value in other frequency bands. The selection depends on the individual pixel herein. More sophisticated rules will consider the surroundings of the pixel and its corresponding pixels in other

#### IMAGE FUSION ASSESSMENT

frequency bands. This is known as region-based fusion approach. Readers are referred to reference [14, 2] for the details and we will not repeat it herein.

To illustrate the process, a simple example is given below in Figure 1. Two images with a horizontal and vertical square bar cross the center respectively are fused with six MRA-based fusion algorithms. The six algorithms include Laplacian pyramid (LAP), gradient-based pyramid (GRAD), ratio-of-lowpass pyramid (RoLP), Daubechies wavelet four (DB), shift-invariant discrete wavelet transform (SIDWT), and steerable pyramid (STEER) [1, 3, 9, 19, 11]. The decomposition level is selected as four and the fusion rule is to select the absolute maximum for the high- and band-pass sub-images (components) and average the low-pass sub-images (components).



(a) Horizontal bar. (b) Vertical bar.

FIGURE 1. Two images are used for testing MRA-based image fusion

To clearly show the results, we visualize the fused images in three dimensions in Figure 2. The result depends on how the features are represented by the MRA algorithms. In other words, the same features are treated differently in different MRA algorithms even though the same fusion rule is applied. The purpose of this paper is not a benchmark study for the MRA-based fusion. This simple example demonstrates how the fusion algorithms work and what can be achieved eventually. Readers are referred to the references for detailed implementation and discussion. Rockinger's Matlab<sup>®</sup> toolbox is a good reference for practice as well [28].

3. Feature-based Assessment. The idea of feature-based assessment for image fusion is to count the availability of features from the inputs. The two aspects of the problem consists of how to quantify the features and how to measure the availability of the features in the fused image. In the proposed method, we employ the phase congruency developed by Kovesi [7] to measure image features. A local correlation calculation is implemented to quantify the availability of the features in the fusion result.

3.1. Image feature from phase congruence measurement. 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 [13]. 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 [5, 7]:

$$PC(x) = \frac{E(x)}{\sum_{n} A_{n}(x) + \varepsilon}$$
(1)

$$E(x) = \sqrt{F^{2}(x) + H^{2}(x)}$$
(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  $\varepsilon$  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 [5]:

$$pc(x) = \frac{\sum_{o} \sum_{n} W_{o}(x) \left\lfloor A_{no}(x) \Delta \Phi_{no}(x) - T_{o} \right\rfloor}{\sum_{o} \sum_{n} A_{no}(x) + \varepsilon}$$
(3)

where o denotes the index over orientation. The noise compensation  $T_o$  is performed in each orientation independently. A weighting function W(x) is constructed to devalue phase congruency at locations where the spread of filter response is narrow. 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.

Figure 3 shows the phase congruency maps along six orientations and the final summation, which is the weighted and noise compensated local energy over all directions. Besides, the principal moments of the phase congruency contain the information for the corners and edges. The magnitude of the maximum and minimum moment can be used directly to determine the edge and corner strength [6]. At each point of the image, the following are computed:

$$a = \sum \left( PC\left(\theta\right) \cos\left(\theta\right) \right)^2 \tag{4}$$

$$b = 2\sum \left(PC\left(\theta\right)\cos\left(\theta\right)\right) \cdot \left(PC(\theta)\sin(\theta)\right)$$
(5)

$$c = \sum \left( PC(\theta) \sin(\theta) \right)^2 \tag{6}$$

where  $PC(\theta)$  is the phase congruency value along orientation  $\theta$  and the sum is performed over the six directions. Therefore, the maximum and minimum moments, M and m are given by:

1436

$$M = \frac{1}{2}(c+a+\sqrt{b^2+(a-c)^2})$$
(7)

$$m = \frac{1}{2}(c + a - \sqrt{b^2 + (a - c)^2}) \tag{8}$$

The phase congruence moments of image "Einstein" is shown in Figure 4. The maximum and minimum phase congruency moments are directly related to the significance of edge and/or corner points [6]. This can also be illustrated with Figure 5. For an extensive discussion of the underlying theory, readers are referred to reference [5, 7, 6].

3.2. The feature-based evaluation metric. Figure 6 illustrates the concept of combinative fusion, in which image features are combined. Assume the letter A and B represent the features from two inputs respectively. The MRA-based pixel-level fusion may generate a fused image as shown by the right two blocks. When the features are totally complementary, i.e. the third one, the fusion result can be assessed by comparing with the two inputs respectively. The input images could be the references. However, when some of the features are redundant, i.e. some overlap between these features, such comparison may be invalid. Therefore, we need to create another reference by selecting the points with larger feature measurement value from the two inputs. During the assessment, the features of the fused image will be compared with the features from the inputs respectively as well as the salient features from all the inputs.

Preliminary results of an evaluation metric i.e.  $P_{blind}$  based on phase congruency measurement are presented in [27, 10]. Herein, we propose a new metric, which is implemented by comparing both the phase congruency measurement and its principal moments. This is defined by the product of the three correlation coefficients as:

$$P_{blind}^{\prime} = \left(P_p\right)^{\alpha} \left(P_M\right)^{\beta} \left(P_m\right)^{\gamma} \tag{9}$$

When there are two input images, we will obtain three values, e.g.  $P_p$ ,  $P_M$ , and  $P_m$  respectively, which is defined as the maximum one of  $C_{xy}^k$ :

$$P_{p} = max\left(C_{1f}^{p}, C_{2f}^{p}, C_{mf}^{p}\right)$$
(10)

$$P_{M} = max\left(C_{1f}^{M}, C_{2f}^{M}, C_{mf}^{M}\right)$$
(11)

$$P_m = max \left( C_{1f}^m, C_{2f}^m, C_{mf}^m \right)$$
(12)

and there is:

$$C_{xy}^{k} = \frac{\sigma_{xy}^{k} + C_{k}}{\sigma_{x}^{k} \sigma_{y}^{k} + C_{k}}$$

$$\tag{13}$$

Herein,  $C_{xy}^k$  stands for the correlation coefficients between two sets x and y. The symbol  $\{k|p, M, m\}$  refers to the phase congruency map and its principal moments. The subfix 1, 2, m, and f correspond to the two inputs, their maximum-selected map, and the result derived from the fused image. The exponential parameters  $\alpha$ ,  $\beta$ , and  $\gamma$  can be adjusted based on the importance of the three components. In our experiments, all the three values are set to one and the small constant value  $C_k$  is selected as 0.0001.

	Assessment	Lapacian	Gradient	Ratio-of-	Daubechies	SIDWT	Steerable
	metric	Pyramid	pyramid	lowpass	wavelet	$(\mathrm{Haar})$	pyramid
				pyramid	four		
Lab	SSIM	0.9808	0.9683	0.9253	0.9762	0.9782	0.9855
	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
	$P'_{blind}$	0.7250	0.7110	0.4337	0.5467	0.7067	0.7263
Book	SSIM	0.9556	0.9485	0.9064	0.9503	0.9538	0.9661
	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
	$P'_{blind}$	0.7572	0.7461	0.5536	0.6893	0.7594	0.7634
Food	SSIM	0.9808	0.9369	0.9591	0.9820	0.9800	0.9833
	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
	$P'_{blind}$	0.9473	0.9242	0.8950	0.9470	0.9470	0.9560
Pepsi	SSIM	0.9519	0.9429	0.9177	0.9427	0.9464	0.9502
	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
	$P'_{blind}$	0.8166	0.7853	0.5412	0.7566	0.7853	0.8241
Objects	SSIM	0.9651	0.9437	0.9382	0.9617	0.9658	0.9723
	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
	$P'_{blind}$	0.7973	0.7835	0.7004	0.7385	0.7946	0.8115

TABLE 1. Evaluation of the fusion results of multi-focused images.

To implement a local comparison, each pixel is compared within a 11-by-11 block in the image and only the points with a phase congruency value larger than 0.1 are used in the calculation. Assume there are total K blocks in the image, the final result is obtained by:

$$P'_{blind} = \frac{1}{K} \sum_{k=1}^{K} P'_{blind}(k) \tag{14}$$

4. Experimental Results. In the experiment, we test the metric with five groups of images shown in Figure 7. The first column is the full-focused image while the second and third are the left- and right-focused images respectively. The left- and right-focused images are fused with the MRA-based fusion algorithms as described in section 2. The fused results are evaluated with the SSIM, Q,  $P_{blind}$ , and  $P'_{blind}$  method respectively. The SSIM gives the comparison results of full-focused and partial-focused images. Both Q and  $P_{blind}$  provide a blind assessment for the fusion algorithms.

The evaluation results are listed in Table 1 and plotted in Figure 8. Compared with Q and  $P_{blind}$ , the  $P'_{blind}$  metric highlights the difference between the fusion algorithms and shows the consistency with the SSIM in the experiment.

5. **Discussion.** In the MRA-based pixel-level fusion, the fusion rule is to keep the coefficients with a larger absolute value, which corresponds to the image feature like lines,

#### IMAGE FUSION ASSESSMENT

edges, and boundaries. The input images come with the "features" (noise) that may not be part of the perfect result; however, the coefficient selection process may eventually retains such "feature" in the fused result. We will still get a confirm from our metric that that "feature" is available in the fused image. In that sense, the assessment metric is a tool to evaluate how the information (feature) is transferred to the fused result. This does not assure a perfect result unless the features are totally complementary. That is the limitation of all approaches that are based on feature measurement in the fused image. The only solution to this problem is the optimization of fusion process rather than the assessment metric. An example is given in Figure 9. Image in Figure 9(a) has a twopixel-wide line across the center with a gray scale value of five. The second image is the blurred version of the first image by applying the averaging operation. The maximum gray scale value is around 1.2. The two images can represent a segment from a bigger picture captured by two image modalities. The MRA-based fusion generates a result shown in Figure 9(c) <sup>1</sup>.

If the evaluation metric is to assess whether the features from the two images are fused in the final result. The conclusion could be that the image in Figure 9(a) is "better" than the one in Figure 9(b). A good implementation of fusion will not introduce any noises or artifacts to the result, but the algorithm must be intelligent enough to identify what should be retained. More sophisticated fusion algorithms considers not only an isolated pixel but also its surroundings and correspondences in other frequency bands. However, no suppression operation has been taken into account when the situation in Figure 9 happens. Therefore, the objective evaluation metric cannot be implemented through finding a perfect result as the reference. This is the ultimate goal of fusion. The assessment of the fusion is carried out by evaluating how the features are transferred from the inputs to the fused result.

The averaging of the low-pass components in the fuse rule does degrade the contrast of the image as shown in Figure 2(f). If we simply select the absolute maximum for the low-pass component, the image in Figure 10 can be obtained. For the images used in the example, Figure 10 demonstrates a better result with the steerable pyramid algorithm. However, maximum selection of the low-pass components does not always give a better result. This may depend on the boundary of each region. If the boundary of a region comes from one input image, the interior of this region should use the pixels from the same image.

The metric developed in this paper does not exclude the other metrics, because one metric only measures one characteristic of the image. In most cases, multiple metrics may be needed. Again, the experimental results do not mean one fusion method outperforms the others for all applications. In our case, the multi-focused imaging is studied.

6. Conclusion. In this paper, a new metric for evaluating combinative image fusion is proposed. The assessment is based on the phase congruency measurement of the image and its corresponding moments. The phase congruency measurement provides an absolute measurement of image features with a value between zero and one. The proposed metrics count how the features from input images are fused into the fusion result. A local correlation is calculated in a square block around a pixel. The average over the

<sup>&</sup>lt;sup>1</sup>The grayscale is adjusted ([0, 0.2]) to show the details of the result.

## 1440 J. ZHAO, R. LAGANIÈRE AND Z. LIU

whole image gives the final metric value. Compared with currently available techniques, the proposed metric  $P'_{blind}$  exhibited a larger standard deviation value in the experiment, which illustrated its capability to differentiate the performance of different algorithms. The proposed metric does not exclude other methods, such as root mean square error, because these methods evaluate different properties of an image. Sometimes, a comprehensive analysis of multiple metrics may be needed to accomplish the fusion assessment study.

Acknowledgements. The images used in the experiment are provided by the SPCL at Lehigh University (USA) and Ginfuku at Kyoto (Japan).

#### REFERENCES

- Adelson, E. H., C. H. Anderson, J. R. Bergen, P. J. Burt and J. M. Ogden, Pyramid methods in image processing, *RCA Engineer*, vol.29, no.6, pp.33-41, 1984.
- [2] Blum, R. S., Z. Xue and Z. Zhang, An overview of image fusion, in *Multi-Sensor Image Fusion and Its Applications*, R. S. Blum and Z. liu (eds.), Taylor and Francis, pp.1-35, 2005.
- [3] Burt, P. J. and R. J. Kolczynski, Enhanced image capture through fusion, Proc. of 4th International Conference on Image Processing, pp.248-251, 1993.
- [4] Cvejic, N., A. Loza, D. Bull and N. Canagarajah, A similarity metric for assessment of image fusion algorithms, *International Journal of Signal Processing*, vol.2, no.3, 2005.
- [5] Kovesi, P., Invarian Measures of Image Features from Phase Information, PhD Dissertation, University of Western Australia, 1996.
- [6] Kovei, P., Phase congruency detects corners and edges, Proc. of the Australian Pattern Recognition Society Conference: DICTA 2003, pp.309-318, 2003.
- [7] Kovesi, P., Image features from phase congruency, Videre: A Journal of Computer Vision Research, vol.1, no.3, 1999.
- [8] Kovesi, P., Image features from phase congruency, Tech. Rep., University of Western Australia, 1995.
- [9] Li, H., B. S. Manjunath, and S. K. Mitra, Multisensor image fusion using the wavelet transform, Graphical Models and Image Processing, vol.57,no.3, pp.235-245, 1995.
- [10] Liu, Z. and R. Laganière, Phase congruence measurement for image similarity assessment, *Pattern Recognition Letters*, vol.28, pp.166-172, 2007.
- [11] Liu, Z., K. Tsukada, K. Hanasaki, Y. K. Ho, and Y. P. Dai, Image fusion by using steerable pyramid, Pattern Recognition Letters, vol.22, pp.929-939, 2001.
- [12] Liu, Z., Z. Xue, R. S. Blum, and R. Laganiere, Concealed weapon detection and visualization in a synthesized image, *Pattern Analysis and Applications*, vol.8, no.4, pp.375-389, 2006.
- [13] Morrone, M. C. and R. A. Owens, Feature detection from local energy, Pattern Recogniton Letters, vol.6, pp.303-313, 1987.
- [14] Piella, G., A general framework for multiresolution image fusion: from pixels to regions, *Information Fusion*, vol.4, no.4, pp.259-280, December 2003.
- [15] Piella, G. and H. Heijmans, A new quality metric for image fusion, Proc. of the International Conference on Image Processing, Bacelona, 2003.
- [16] Piella, G., New quality measures for image fusion, Proc. of the International Conference on Information Fusion, Stockholm, Sweden, 2004.
- [17] Qiao, Y. L., Z. M. Lu, J. S. Pan and S. H. Sun, Spline wavelets based texture features for image retrieval, *International Journal of Innovative Computing*, *Information and Control*, vol.2, no.3, pp.653-658, 2006.
- [18] Qu, G., D. Zhang, and P. Yan, Information measure for performance of image fusion, *Electronics Letters*, vol.38, no.7, pp.313-315, 2002.
- [19] Rockinger, O., Image sequence fusion using a shift-invariant wavelet transform, Proc. of the International Conference on Image Processing, vol.3, pp.288-301, 1997.

- [20] Teot, A., Image fusion by a ratio of low-pass pyramid, Pattern Recognition Letters, vol.9, pp.245-253, 1989.
- [21] Wang, Z., A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, Image quality assessment: From error measurement to structural similarity, *IEEE Transactions on Image Processing*, vol.13, no.1, 2004.
- [22] Wang, Y. and B. Lohmann, Multisensor image fusion: Concept, method and applications, *Tech. Rep.*, Institut für Automatisierungstechnik, Universität Bremen, 2000.
- [23] Xu, R. and Y. W. Chen, Wavelet-based multiresolution medical image registration strategy combining mutual information with spatial information, *International Journal of Innovative Computing*, *Information and Control*, vol.3, no.2, pp.285-296, 2007.
- [24] Xydeas, C. S. and V. Petrovic, Objective image fusion performance measure, *Electronics Letters*, vol.36, no.4, pp.308-309, 2000.
- [25] Xydeas, C. S. and V. Petrovic, Objective pixel-level image fusion performance measure, Proc. of the SPIE, vol.4051, pp.89-98, 2000.
- [26] Zhang, Z. and R. S. Blum, Image fusion for a digital camera application, Proc. of the 32nd Asilomar Conference on Signals Systems, and Computers, Monterey, CA, pp.603-607, 1998.
- [27] Zhao, J., R. Laganière, and Z. Liu, Image Fusion Algorithm Assessment based on Feature Measurement, Proc. of the 2006 International Conference on Innovative Computing, Information and Control, Beijing, China, pp.701-704, 2006.
- [28] URL, "http://www.imagefusion.org".



FIGURE 2. The fusion results with different MRA-based fusion algorithms



FIGURE 3. The computation of phase congruency for image "Einstein". The center image is the summed phase congruency map.



FIGURE 4. The principal moments of phase congruency of image "Einstein"



FIGURE 5. The principal moments of phase congruency of the image in Figure 1(a)



FIGURE 6. The combinative fusion



FIGURE 7. 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.



### (a) Laboratory





### (c) Japanese food





(e) Objects

FIGURE 8. The assessment results

1446



(a) Salient feature. (b) Weak feature. (c) The fused result.

FIGURE 9. The example of fusing a strong and a weak feature



FIGURE 10. Fusion implemented by maximum selection of low-pass component