ON THE USE OF PHASE CONGRUENCY TO EVALUATE IMAGE SIMILARITY

Zheng Liu and Robert Laganière

School of Information Technology and Engineering (SITE)
University of Ottawa
800 King Edward, Ottawa, Ontario, K1N 6N5 Canada

ABSTRACT

Measuring image similarity is important in many applications. Different algorithms propose to compare images using pixel-based mean square error methods others use structure-based image quality index. We present, here, a new feature-based approach that utilizes image phase congruency measurement to quantify the assessment of the similarities or differences between two images.

1. INTRODUCTION

Measuring image similarity is important in many applications. In the performance assessment of an image processing algorithm, for example, an image is often compared to an available reference. Usually, such comparison is implemented on pixel-based operations, like mean square error (MSE) or root mean square error (RMSE). However, such operations' performance is questionable because the same MSE or RMSE value does not always assure a comparable image similarity under different distortion to perceptually significant features (1).

In practical applications, some post-processing operations largely depends on the availability of image features. Operations, like classification, segmentation, and quantification, are often carried out in a feature space. Therefore, the availability of image features plays an important role in further analysis. This paper proposes a new feature-based method to quantitatively assess image similarity by employing the phase congruency measurement suggested by Kovesi (2; 3). Phase congruency provides an absolute measure of image features such as step edges, lines or corners; it is viewing condition-independent and invariant to changes in illumination and magnification. A local cross-correlation of the phase congruence map can then be calculated between a processed image and a reference image. The averaged cross-correlation value provides a quantitative assessment of the overall image similarity.

2. RELEVANT WORKS

There are a number of metrics available for image comparison. The commonly used approaches include root mean square error (RMSE), normalized least square error (NLSE), peak signal to noise ration (PSNR), and correlation (CORR). The definition of these metrics mentioned above are given in equations 1-4, in which $R(x, y)$ and $I(x, y)$ stand for the reference and target image respectively and L is the maximum pixel value; the size of the images is $M \times N$. These methods are widely used because they represent simple algebraic quality measures that can be easily computed. However, they often do not match very well with human perception of visual quality. For example, images with a similar RMSE value may exhibit a quite different appearance.

$$RMSE = \sqrt{\frac{\sum_{m,n}^{M,N} [R(m,n) - I(m,n)]^2}{MN}} \quad (1)$$

$$NLSE = \sqrt{\frac{\sum_{m,n}^{M,N} [R(m,n) - I(m,n)]^2{}^2}{\sum_{m,n}^{M,N} [R(m,n)]}} \quad (2)$$

$$PSNR = 10 \log_{10} \left(\frac{L^2}{\frac{1}{MN} \sum_{m,n}^{M,N} [R(m,n) - I(m,n)]^2} \right) \quad (3)$$

$$CORR = \frac{2 \sum_{m,n}^{M,N} R(m,n) I(m,n)}{\sum_{m,n}^{M,N} R(m,n)^2 + \sum_{m,n}^{M,N} I(m,n)^2} \quad (4)$$

There are also more sophisticated methods using higher level information to measure image similarity (4; 5). The entropy difference (DE) between two images reflects the difference between the average amount of information they contained. It is defined as:

$$DE = \left| \sum_{g=0}^{L-1} P_R(g) \log_2 P_R(g) - \sum_{g=0}^{L-1} P_I(g) \log_2 P_I(g) \right| \quad (5)$$

where $P_R(g)$ and $P_I(g)$ are the probability of pixel value g for the reference and input image respectively. The mutual information (MI) between the input and reference images are defined on the normalized joint gray level histogram

$h_{RI}(i,j)$ and normalized marginal histogram of the two images i.e. $h_R(i)$ and $h_I(j)$:

$$MI = \sum_{i=1}^{L} \sum_{j=1}^{L} h_{RI}(i,j) \log_2 \frac{h_{RI}(i,j)}{h_R(i) h_I(j)} \quad (6)$$

The structure similarity index metric (SSIM) proposed by Wang (5) is based on the evidence that human visual system is highly adapted to structural information and a measurement of the loss of structural information can provide a good approximation of the perceived image distortion. The definition of the SSIM metric is expressed as:

$$SSIM = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (7)$$

where μ_x and μ_y are the mean values of the two images. σ_x, σ_y, and σ_{xy} are the corresponding variance values. In the above equation, two constant values C_1 and C_2 are defined to avoid the instability when the denominators are very close to zero. These two values are further determined by two subjectively selected values K_1, K_2, and the dynamic range of the pixel values, i.e. $C_1 = (K_1L)^2$ and $C_2 = (K_2L)^2$.

In feature-based image comparison, Morrone and Owens (6) proposed an image feature perception model, which postulated that features were perceived at points in an image where Fourier components were maximally in phase. A wide range of feature types give rise to points of high phase congruency. Kovesi (2) proposed a scheme to calculate the phase congruency with logarithmic Gabor wavelets, which allow arbitrarily large bandwidth filters to be constructed while still maintaining a zero DC component in the even-symmetric filter. The equation of the phase congruency $PC(x)$ at some location x is expressed as the summation over orientation o and scale n:

$$pc(x) = \frac{\sum_o \sum_n W_o(x) \lfloor A_{no}(x)\Delta\Phi_{no}(x) - T_o \rfloor}{\sum_o \sum_n A_{no}(x) + \varepsilon} \quad (8)$$

where $\lfloor \rfloor$ is a floor function expressing the fact that the enclosed quantity is not permitted to be negative. A_n represents the amplitude of the nth component in the Fourier series expansion. A very small positive constant ε is added to the denominator in case of small Fourier amplitudes. T_o compensates for the influence of noise and is estimated empirically. $\Delta\Phi_n(x)$ is a more sensitive phase deviation and defined as:

$$
\begin{aligned}
A_n(x)\Delta\Phi_n(x) =\ & e_n(x)\bar{\phi}_e(x) + o_n(x)\bar{\phi}_o(x) \quad (9)\\
& - \left| e_n(x)\bar{\phi}_o(x) - o_n(x)\bar{\phi}_e(x) \right|
\end{aligned}
$$

where $\bar{\phi}_e(x) = \sum_n e_n(x)/E(x)$ and $\bar{\phi}_o(x) = \sum_n o_n(x)/E(x)$. The term $E(x)$ is the local energy function and is expressed as $\sqrt{(\sum_n e_n(x))^2 + (\sum_n o_n(x))^2}$. The convolution results of the input image $I(x)$ with quadrature pairs of filters at
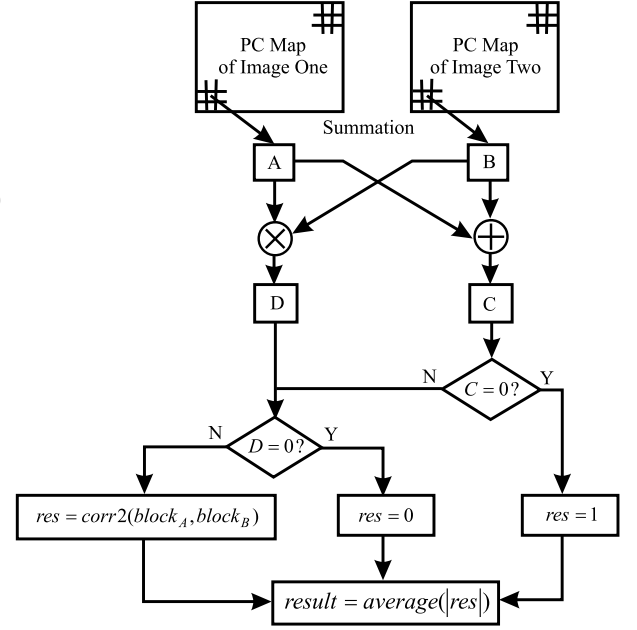

Fig. 1. The procedure for computing local cross-correlation of two phase congruence maps.

scale n, $e_n(x) = I(x) * M_n^e$ and $o_n(x) = I(x) * M_n^o$, consist of the basic components to calculate $PC(x)$. M_n^e and M_n^o denote the even-symmetric and odd-symmetric wavelet at this scale respectively. For detailed information about the theory, readers are referred to reference (3).

3. PROPOSED STRATEGY FOR IMAGE COMPARISON

The comparison of images can be carried out by comparing their corresponding phase congruency feature maps. It is indeed appropriate to evaluate the space-variant features locally and combine them together to obtain a global quality index (7). Cross-correlation can be used to measure the similarity between the phase congruency maps, which are divided into sub-blocks initially. A typical choice for the size of the sub-block window is 5×5.

In a phase congruency map, the sub-block window could be blank, i.e. all the feature points are zero at the locations without any features. Therefore, when either of the sub-block windows is blank, the local cross-correlation value is set to zero. Conversely, if both the sub-block windows are blank, the cross-correlation value is set to one. For the other cases, the value is computed by the zero-mean normalized cross-correlation (ZNCC) (8):

The above procedure is implemented as shown in Figure 1. An overall assessment is achieved by averaging the cross-correlation values from each pre-defined region.

The phase congruency measurement is claimed to be invariant to illumination and contrast changes. This mechanism
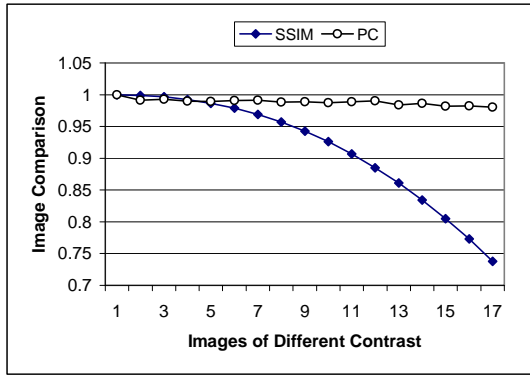
Fig. 2. The contrast-invariant property of the proposed metric (PC) compared to SSIM.

is in accordance with the human visual system that demonstrates good invariance to lighting conditions. The following experiment confirms that invariance property of the phase congruency measure: the Matlab® function "imadjust" was used to adjust an image contrast at different levels. Results obtained with the SSIM algorithm and our proposed method are shown in Figure 2. Compared to SSIM, the PC metric is not sensitive to changes in contrast.

4. EXPERIMENTAL RESULTS

An experiment involving different degraded versions of a source image has been performed. The resulting images were transformed such to result in an identical root mean square error (RMSE) although they differ in appearance. In these various processes, the original image is contaminated by salt-pepper noise, Gaussian noise, speckle noise, mean shifting, contrast stretching, blurring operation, and JPEG compressing respectively. All these images and their corresponding phase congruency maps are presented in Figure 5. Besides the RMSE, a group of metrics are also computed for comparison: normalized least-square error (NLSE), peak signal-to-noise ratio (PSNR), correlation (CORR), difference entropy (DE), mutual information (MI), structure similarity measure (SSIM), and the proposed phase congruence based method (PC). The numerical results are listed in Table 1.

RMSE and NLSE have similar behavior; they cannot distinguish the different levels of distortions that the transformations have imposed. With MI, the larger the value, the more similar are the compared images. This value is however not normalized; the MI value of the source image with itself is required for an absolute measure of similarity. With DE, a value zero indicates a perfect match of two images. This method does not seem to be very discriminative but is very sensitive to compression artifacts.

As illustrated in Figure 3, PC and SSIM seem to behave similarly. Phase congruency is however more sensitive to the
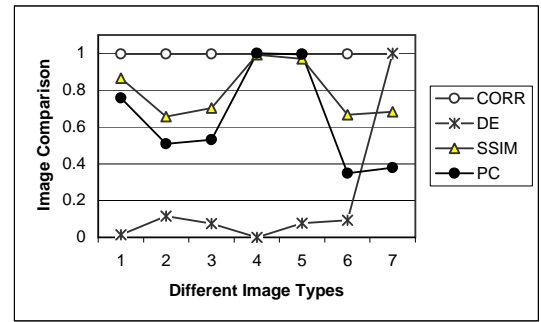

Fig. 3. Comparison of an image before and after processing (Index 1 to 7 corresponds to the different distortion types listed in Table 1).
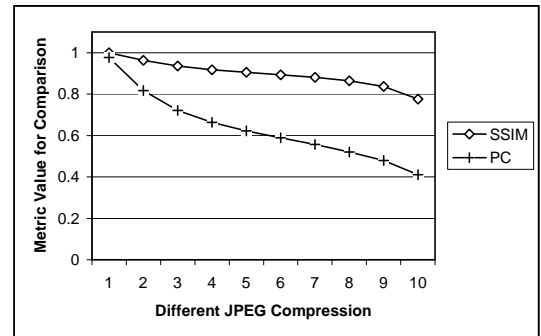

Fig. 4. The comparison of JPEG compressed images.

annoying degradations induced by blurring and compression. Most methods were not able to capture the similarity of the Mean-shifted image. This one is obtained by adding a constant value to the image; such operation does not alter the structural information conveyed by an image. Consequently, the PC metric reaches it maximum value (one) while the DE value is zero for this image.

The sensitivity of the proposed method has been also tested by the following experiment. A group of images has been compressed using JPEG at different level of compression. The compressed images with different quality indices are compared with the reference one by using the SSIM and our PC metric. The average values of the resulting SSIM and PC metrics are plot in Figure 4. The proposed PC metric appears to be more sensitive to the structural alterations and successfully differentiates the different compression factors.

5. CONCLUSION

In this paper, a new feature-based metric for image comparison was proposed. It is based on phase congruency features that are locally correlated and from which a global similarity measure is obtained. The effectiveness of this metric has been investigated through different experiments with standard images. The measure is invariant to changes in image contrast or

	Distortion	RMSE	NLSE $(\times 10^{-2})$	PSNR	CORR	DE	MI	SSIM	PC
(1)	Salt-pepper noise	10.9640	8.9483	25.3140	0.9960	0.0141	3.4111	0.8643	0.7569
(2)	Gaussian noise	11.0060	8.9823	25.4100	0.9960	0.1163	1.4347	0.6556	0.5079
(3)	Speck noise	11.0200	8.9935	24.5020	0.996	0.0754	1.6964	0.7032	0.5311
(4)	Mean shift	11.0000	8.9776	Inf	0.9963	0.0000	3.9347	0.9927	1.0000
(5)	Contrast stretch	10.9950	8.9739	38.6860	0.9961	0.0770	3.4654	0.9698	0.9948
(6)	Blur	11.0430	9.0123	25.4150	0.9959	0.0934	1.5601	0.6671	0.3478
(7)	JPEG compression	10.8390	8.8458	24.9880	0.9961	1.0005	1.4656	0.6824	0.3795

Table 1. The comparison of standard images in Figure 5.

in illumination and exhibits good sensitivity to various image distortion categories.

References

[1] G. Paolo, "Image comparison metrics : A review," May 1998.

[2] P. Kovesi, "Phase congruency: A low-level image invariant," *Psychological Research*, vol. 64, pp. 134–148, 2000.

[3] P. Kovesi, *Invariant Measures of Image Features from Phase Information*, Ph.D. thesis, University of Western Australia, 1996.

[4] Y. Wang and B. Lohmann, "Multisensor image fusion: Concept, method and applications," Tech. Rep., Institut für Automatisierungstechnik, Universität Bremen, German, December 2000.

[5] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, April 2004.

[6] M. C. Morrone and R. A. Owens, "Feature detection from local energy," *Pattern Recognition Letters*, pp. 303–313, June 1987.

[7] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Singal Processing Letters*, vol. 9, no. 3, pp. 81–84, March 2002.

[8] J. Martin and J. L. Crowley, "Experimental comparison of correlation techniques," in *Proceedings of International Conference on Intelligent Autonomous Systems*, 1995.
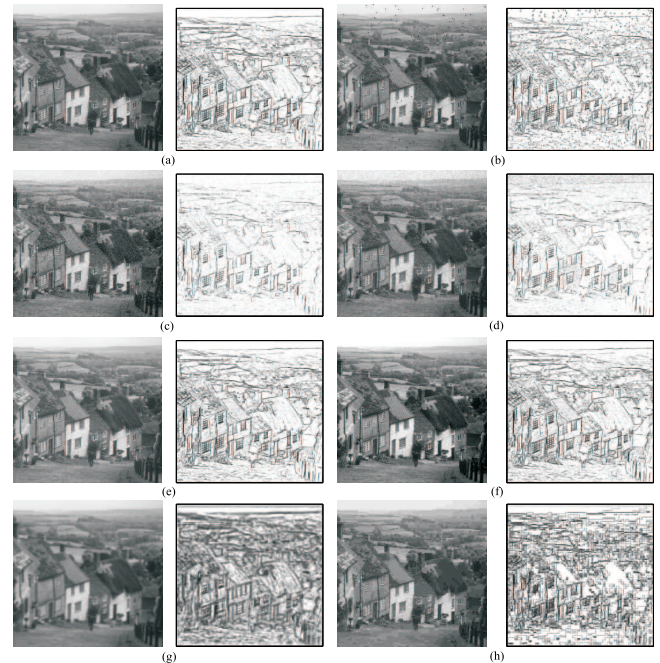
Fig. 5. The "gold hill" image of different distortions: (a)original image (reference); (b)Salt-pepper noise contaminated image; (c) Gaussian noised contaminated image; (d) Speck noise contaminated image; (e) Mean shifted image; (f) Contrast stretch image; (g) Blurred image; and (h) JPEG compressed image. (courtesy of Dr. Z. Wang, New York University).