2012

New structural similarity measure for image comparison

Prashan Premaratne
University of Wollongong, prashan@uow.edu.au

Malin Premaratne
Monash University

Publication Details

P. Premaratne & M. Premaratne, "New structural similarity measure for image comparison," in Emerging Intelligent Computing Technology and Applications, P. Gupta, D. Huang, P. Premaratne & X. Zhang, Ed. Berlin: Springer Berlin Heidelberg, 2012, pp.292-297.

New structural similarity measure for image comparison

Abstract
Subjective quality measures based on Human Visual System for images do not agree well with well-known metrics such as Mean Squared Error and Peak Signal to Noise Ratio. Recently, Structural Similarity Measure (SSIM) has received acclaim due to its ability to produce results on a par with Human Visual System. However, experimental results indicate that noise and blur seriously degrade the performance of the SSIM metric. Furthermore, despite SSIM's popularity, it does not provide adequate insight into how it handles 'structural similarity' of images. We propose a structural similarity measure based on approximation level of a given Discrete Wavelet Decomposition that evaluates moment invariants to capture the structural similarity with superior results over SSIM.

Keywords
comparison, measure, image, similarity, structural

Disciplines
Engineering | Science and Technology Studies

Publication Details
P. Premaratne & M. Premaratne, "New structural similarity measure for image comparison," in Emerging Intelligent Computing Technology and Applications, P. Gupta, D. Huang, P. Premaratne & X. Zhang, Ed. Berlin: Springer Berlin Heidelberg, 2012, pp.292-297.

New Structural Similarity Measure for Image Comparison

Prashan Premaratne and Malin Premaratne*

School of Electrical Computer and Telecommunications Engineering,
University of Wollongong, North Wollongong, NSW, Australia
Department of Electrical and Computer Systems Engineering at Monash University,
Victoria, Australia
malin@ieee.org, prashan@uow.edu.au

Abstract. Subjective quality measures based on Human Visual System for images do not agree well with well-known metrics such as Mean Squared Error and Peak Signal to Noise Ratio. Recently, Structural Similarity Measure (SSIM) has received acclaim due to its ability to produce results on a par with Human Visual System. However, experimental results indicate that noise and blur seriously degrade the performance of the SSIM metric. Furthermore, despite SSIM's popularity, it does not provide adequate insight into how it handles 'structural similarity' of images. We propose a structural similarity measure based on approximation level of a given Discrete Wavelet Decomposition that evaluates moment invariants to capture the structural similarity with superior results over SSIM.

Keywords: Image similarity, structural similarity, moment invariants, SSIM, MISM.

1 Introduction

Comparing two images accurately to ascertain whether there is a match or not is essential for many image processing related tasks such as watermarking, compression and content retrieval. Age-old metrics such as Mean Squared Error (MSE) have been used for decades despite its inability to agree with human subjective analysis [1, 2]. Recently, light has been shed on a new metric that seems to agree with Human Visual System [2]. SSIM has been singled out due to its claim of superiority over the existing metrics [3, 4]. However, it has been observed that SSIM does not perform well with blurred images [4]. Since a blurred version of an image essentially contains the same structure, SSIM's inability to measure the structural similarity of blurred images raise an issue as to whether SSIM does truly look for the structural content. From our research, we have concluded that despite SSIM claim of superiority, its ability to compare similar structures is doubtful as will be demonstrated in the Experimental results section. We have developed a new metric that uses some of the concepts exploited by

* Corresponding author.

D.-S. Huang et al. (Eds.): ICIC 2012, CCIS 304, pp. 292–297, 2012.

SSIM. The new metric demonstrates better performance over SSIM in blurred images and images corrupted by Gaussian and Salt & Pepper noise.

2 Structural Similarity Measure (SSIM)

SSIM attempts to separate the task of similarity measurement of two images into luminance, contrast and structure [2]. Hence, a similarity measure is defined as:

$$\mathrm{SSIM}(\mathbf{P_1},\mathbf{P}_2) = l(\mathbf{P_1},\mathbf{P}_2)\times c(\mathbf{P_1},\mathbf{P}_2)\times s(\mathbf{P}_1,\mathbf{P}_2) \tag{1}$$

Where P1 and P2 are the two images being compared and l, c and s stand for luminosity, contrast and similarity measure. μ and σ are mean and standard deviation of the corresponding images and C1, C2 and C3 are constants used for the stability of equations when μ and σ are extremely small. SSIM defines μ, σ, σP1σP2, l, c, s as follows [3]:

$$\mu_{P1} = \frac{1}{M\times N}\sum_{y=0}^{M-1}\sum_{x=0}^{N-1}\mathbf{P_1}(x,y)$$

$$\sigma_{P1} = \sqrt{\frac{1}{(M\times N-1)}\sum_{y=0}^{M-1}\sum_{x=0}^{N-1}(\mathbf{P_1}(x,y)-\mu_{P1})^2}$$

$$\sigma_{P1}\sigma_{P2} = \frac{1}{M\times N-1}\sum_{y=0}^{M-1}\sum_{x=0}^{N-1}(\mathbf{P}_1(x,y)-\mu_{P1})(\mathbf{P}_2(x,y)-\mu_{P2})$$

$$l(\mathbf{P_1},\mathbf{P_2}) = \frac{2\mu_{P1}\mu_{P2}+C_1}{\mu_{P1}^2+\mu_{P2}^2 C_1}$$

$$c(\mathbf{P}_1,\mathbf{P}_2) = \frac{2\sigma_{P1}\sigma_{P2}+C_2}{\sigma_{P1}^2+\sigma_{P2}^2+C_2}$$

$$s(\mathbf{P}_1,\mathbf{P}_2) = \frac{\sigma_{P1P2}+C_3}{\sigma_{P1}\sigma_{P2}+C_3}$$

$$\mathrm{SSIM}(\mathbf{P}_1,\mathbf{P}_2) = \frac{(2\mu_{P1}\mu_{P2}+C_1)(\sigma_{P1P2}+C_2)}{(\mu_{P1}^2+\mu_{P2}^2 C_1)(\sigma_{P1}^2+\sigma_{P2}^2+C_2)} \tag{2}$$

(2) has been obtained using (1) when C3 = C2/2 for simplicity. However, it is difficult to understand how s(P1,P2) would represent structure as it is simply a function of cross correlation.

3 Moment Invariant Based Structural Similarity Measure (MISM)

The proposed approach here is very well understood as the approximation level of Discrete Wavelet Decomposition of an image results in revealing the structure of the images. The approximation levels remove detail successively and leave the structure intact even at deeper decomposition levels. At each successive level, structure of an image is maintained while removing the texture and detail. Once the image is reduced to an acceptable level, edge detection can be used to further sharpen the structure of the image. If a metric is produced using this structural information, it will truly capture the structural information and will be a valid measure to evaluate the structural integrity thereby making comparing images more meaningful.

Moment Invariants have been used extensively in identifying shapes or outlay of objects for many years [5, 6]. An image reduced to 16x16 or larger using Wavelet decomposition can be used to generate moment invariants to identify the structural makeup of an image. As our research indicates, matching at two such levels will

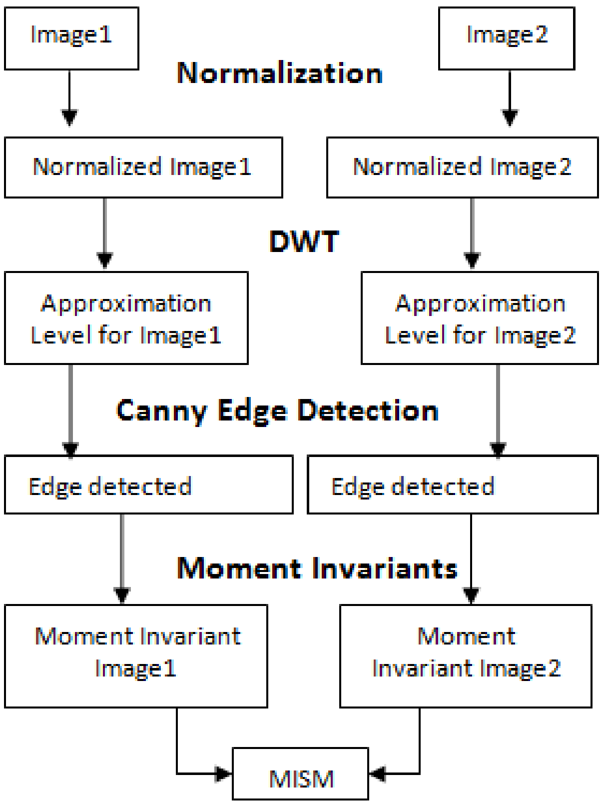

Fig. 1. MISM evaluation using two images

indicate very high similarity for an image undergone blurring or corruption with noise and can be verified visually. Hence the approach complies with Human Visual System and is far superior to MSE estimates.

MISM calculation is outlined in Fig.1. An image is normalized (divided by its own standard deviation) such that the two images being compared have unit standard deviation. An image reduced to an approximation level (usually larger than 16x16) and then edge detected using 'Canny' operator and first moment invariant (ϕ1) is calculated for the entire approximation [5]. Then the approximation level is divided into four quadrants and the first and second moments (ϕi1, ϕi2) are calculated for each

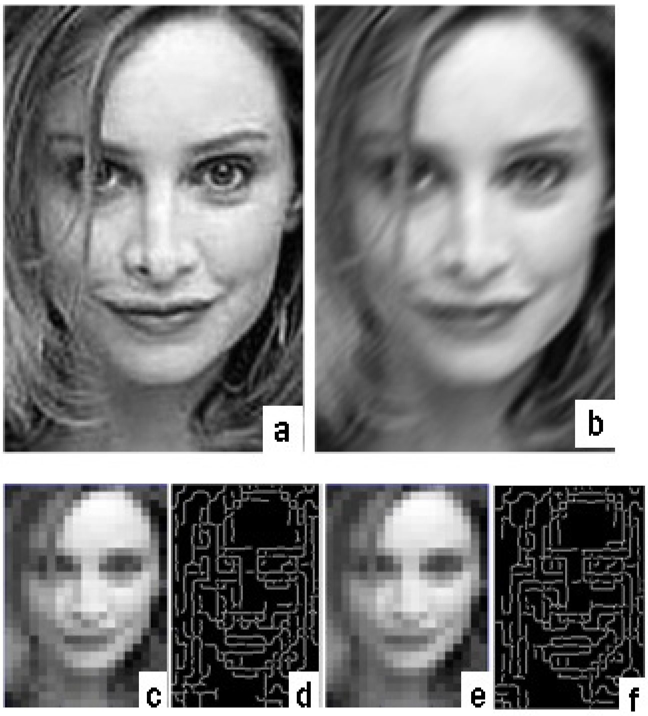

Fig. 2. (a) Original image of Ally, **(b)** Ally with motion blur, **(c)** Level 3 approximation of **(a)**, **(d)** Edge detection of **(c)**, **(e)** Level 3 approximation of **(b)** and **(f)** Edge detection of **(e)**.

Table 1. Comparison of SSIM and MSIM for images

Image	SSIM	MISM
Ally with motion Blur	0.6285	0.8996
Ally with S&P noise	0.6607	0.9315
Ally with Gaussian noise	0.3560	0.9347
Lena	0.1542	0.5525

quadrant. These values are used to calculate the MISM for the entire image using the weights as shown in (3).

$$\text{MISM} = 1-(0.1\frac{|\phi_1-\phi_1'|}{\phi_1}+\sum_{i=1}^{4}0.05\frac{|\phi_{i1}-\phi_{i1}'|}{\phi_{i1}}+\sum_{i=1}^{4}0.15\frac{|\phi_{i2}-\phi_{i2}'|}{\phi_{i2}}) \quad (3)$$

Here ϕ' indicates the moment invariants of the second image. Fig. 2 indicates clearly that the structure is intact at low decomposition levels despite motion blur. This is also true for images corrupted with noise.

4 Experimental Results

MISM shows lot of promise for image similarity based metrics as well as for image matching. As shown in Tab. 1. MISM is developed to be slightly biased towards similarity rather than dissimilarity. Hence, Lena scores 0.1542 compared with Alice using SSIM where as MISM scores 0.5525. On the other hand, when comparing different versions of Ally such as Ally with motion blur, Gaussian noise and Salt & Pepper noise, SSIM measures 0.6285, 0.3560 and 0.6607. If SSIM truly compares structural similarity as the authors claim [3], all these images with the same structure should record a similar SSIM measure. MISM on the other hand, consistently record, 0.8996, 0.9347 and 0.9315 indicating that the proposed measure is certainly measuring the structural similarity.

5 Conclusion

We have evaluated the performance of the SSIM using the programming code made available by the original authors against our MISM and have demonstrated that image structural similarity can be best established accurately using MISM. In our research, we found that MISM is providing more insight to the image structure opposed to SSIM as it does not represent structure as claimed. MISM is very much comparable to SSIM with similar computer processing time.

References

1. Eskicioglu, A.M.: Quality measurement for monochrome compressed images in the past 25 years. In: Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing, vol. 4, pp. 1907–1910 (2000)
2. Girod, B.: What's wrong with mean-squared error. In: Digital Images and Human Vision, pp. 207–220. MIT press (1993)
3. Wang, Z., Bovik, A., Sheik, H.R., Simoncelli, E.P.: Image quality assessment: From error visibility to structural similarity. IEEE Trans. Image Process 4, 13(4), 1–14 (2004)

4. Chen, G.H., Yang, C.L., Po, L.M., Xie, S.L.: Edge-based structural similarity for image quality assessment. In: Proc. ICASSP 2006, vol. 2, pp. 933–936 (2006)
5. Premaratne, P., Nguyen, P., Consumer, Q.: electronics control system based on hand gesture moment invariants. IET Computer Vision 1(1), 35–41 (2007)
6. Premaratne, P., Ajaz, S., Premaratne, M.: Hand Gesture Tracking and Recognition System for Control of Consumer Electronics. In: Huang, D.-S., Gan, Y., Gupta, P., Gromiha, M.M. (eds.) ICIC 2011. LNCS (LNAI), vol. 6839, pp. 588–593. Springer, Heidelberg (2012)