Translation insensitive assessment of image quality based on measuring the homogeneIty of correspondence

Chun-Hsien Chou[1] and Yun-Hsiang Hsu[2]
[1,2]Department of Electrical Engineering, Tatung University, Taipei, Taiwan, R.O.C.
E-mail: chunhsienchou@gmail.com Tel: +886-022182-2928
E-mail: ooshine1031@gmail.com Tel: +886-022182-2928

Abstract—**Image quality assessment plays an important role in the development of many image processing systems. Many full-reference image quality metrics have been proposed and aimed to give the prediction as close as possible to the subjective assessment made by human beings. However, these metrics have a common restriction that pixel-wise correspondence must be established before the evaluation of metric scores. Most of the existing metrics fail to result in accurate prediction even as a test image is differentiated from its original reference merely by one-pixel misalignment. Based on the fact that dissimilar image contents lead to random block correspondence, an image quality assessment method that primarily measures the randomness in the displacement between corresponding blocks from the images in comparison is proposed. The performance of the proposed metric is verified by evaluating the quality of the test images contained in the LIVE and TID2008 databases and the same images translated by various amounts of distance. The correlation between subjective evaluation results and the objective scores evaluated by the proposed metric as well as other five well-know image quality assessment methods is examined. Experimental results indicate that the proposed metric is an effective assessment method that can predict the image quality accurately without the preprocessing of image alignment.**

I. INTRODUCTION

Image quality assessment (IQA) plays an important role in the development of image processing systems, such as image/video compression, content-based image retrieval, image registration and pattern recognition, *etc*. Especially for those systems where processed images are to be viewed by human beings, an objective similarity metric that gives rise to the metric score highly correlated with the subjective assessment of human beings is highly demanded. To achieve this goal, a number of full-reference IQA methods were proposed [1-13], in which VSNR [1], WSNR [2] and NQM [3] exploit characteristics of the human visual system (HVS) to improve the prediction performance. The statistics of the natural scene are utilized in the information fidelity criterion (IFC) metric [8]. The visual information fidelity metric (VIF) [9] assesses the image quality by quantifying the mutual information between distorted and reference images. In [6, 7], the information obtained from the singular value decomposition (SVD) of the image's characteristic matrix is used in the evaluation of metric scores. The structural similarity index (SSIM) [4-5] compares local patterns of pixel intensities as the normalized mean intensity and contrast. The design of SSIM is based on the assumption that the human visual system is highly adapted to the structure information extracted from the perceived scene, and that the change in structural information can provide a good approximation to the perceived image distortion. Many similar studies [10-11] are conducted to modify the SSIM index for more close to human perception. In [12], the phase information embedded in images is taken as the representative features for evaluating the metric score. The ideas behind this metric were motivated from the fact that if an image has some structural distortions, the distortions lead to consistent phase change.

A common shortcoming of the above-mentioned IQA approaches is that they are very sensitive to geometric distortions. Most of these IQA methods are not able to perform properly in predicting image quality if perfect image registration or pixel correspondence is not established before the evaluation of metric scores, and fail to predict the similarity between two identical images which are differentiated merely by a pixel wide of translation. Only few literatures were published on the geometric translation problem in image quality assessment [13-15]. Based on the fact that small geometric image distortions lead to consistent phase changes in local wavelet coefficients, the complex wavelet SSIM (CW-SSIM) [13-14] was found robust to small translations and rotations. In most cases, the process to establish the correspondence between pairs of pixels in the images to be compared can be computationally costly, and the result of the process may lead to an erroneous outcome depending on the contents of the two images.

In this paper, a novel image quality metric called homogeneous correspondence index (HCI) is presented, which is able to accurately predict the similarity between two images without having to establish the pixel correspondence in the first place. The idea behind the proposed metric is that the corresponding features contained in two similar images will not be randomly displaced if simple geometric distortion exists between the two images. In this paper, the correspondence refers to the pair of image blocks located in two different images that are most similar to each other, and

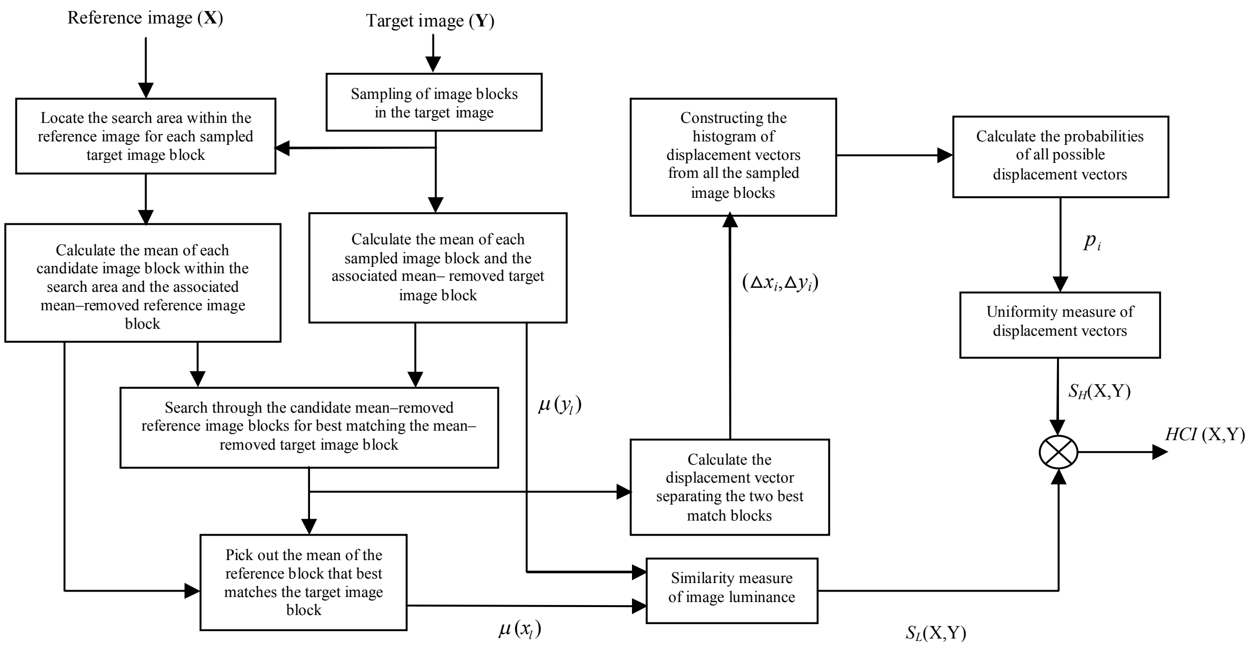


Fig. 1. Block diagram for evaluating the objective score of the proposed metric

associated by the relative displacement in between. The homogeneous correspondence in this case implies that all pairs of corresponding blocks are associated by the same displacement. The probability associated with each possible displacement is used to evaluate the score of the proposed entropy-like metric. The proposed metric has a lower bound of zero and a finite upper bound, the value of which depends upon the number of sampling points for measuring the homogeneous of correspondence. The performance of the proposed metric is verified by evaluating the quality of the test images contained in the LIVE [16] and TID2008 [17] databases and the same images translated by various amounts of distance. The correlation between subjective evaluation results and the objective scores evaluated by the proposed metric as well as other five well-know image quality assessment methods is examined.

II. THE PROPOSED IMAGE QUALITY METRIC

The proposed HCI metric is designed to primarily measure the uniformity of the displacements between the corresponding blocks in both reference and test images. The functional block diagram for evaluating the objective score of the proposed metric is shown in Fig.1. The metric score of HCI can be evaluated from two components as expressed by the following equation

$$HCI\,(\mathbf{X},\mathbf{Y}) = S_H(\mathbf{X},\mathbf{Y}) \cdot S_L(\mathbf{X},\mathbf{Y}) \qquad (1)$$

where **X** and **Y** denote the reference image and the test image to be assessed, respectively. $S_H(\mathbf{X},\mathbf{Y})$ is responsible for measuring the homogeneity of the displacement vectors that associate the corresponding mean-removal blocks in the two images, and can be expressed as

$$S_H(\mathbf{X},\mathbf{Y}) = 1 - R(\mathbf{X},\mathbf{Y}), \qquad (2)$$

where

$$R(\mathbf{X},\mathbf{Y}) = H(\mathbf{X},\mathbf{Y}) \,/\, H_{\mathbf{max}} \qquad (3)$$

is the normalized randomness of the displacement vectors that associate the corresponding mean-removal blocks in the two images. The $H(\mathbf{X},\mathbf{Y})$ in Eq. (3) is an entropy function that measures the randomness of the displacement vectors, as can be derived as

$$H(\mathbf{X},\mathbf{Y}) = -\sum_{i=1}^{N} p_i \log_2 p_i \qquad (4)$$

where N is the number of all possible displacement vectors that associate the corresponding blocks. The value of N depends on the dimension of the search area involved in finding the best matched block in the reference image for each block from the test image. The value of N will be $(2d+1)\times(2d+1)$ if the full search in the reference image is ranged from $-d$ to d along each of the two directions. p_i is the probability associated with the occurrences of the displacement vector indexed by i. That is

$$p_i = \frac{M_i}{M} \qquad \text{for} \quad 1 \le i \le N, \qquad (5)$$

where M denotes the total number of image blocks partitioned from the test image and M_i the number of image blocks that deviate from their corresponding blocks by the displacement vector $(\Delta x, \Delta y)$ indexed by i. $H_{\mathbf{max}}$ implies the maximum randomness, which can be reached as the probabilities of

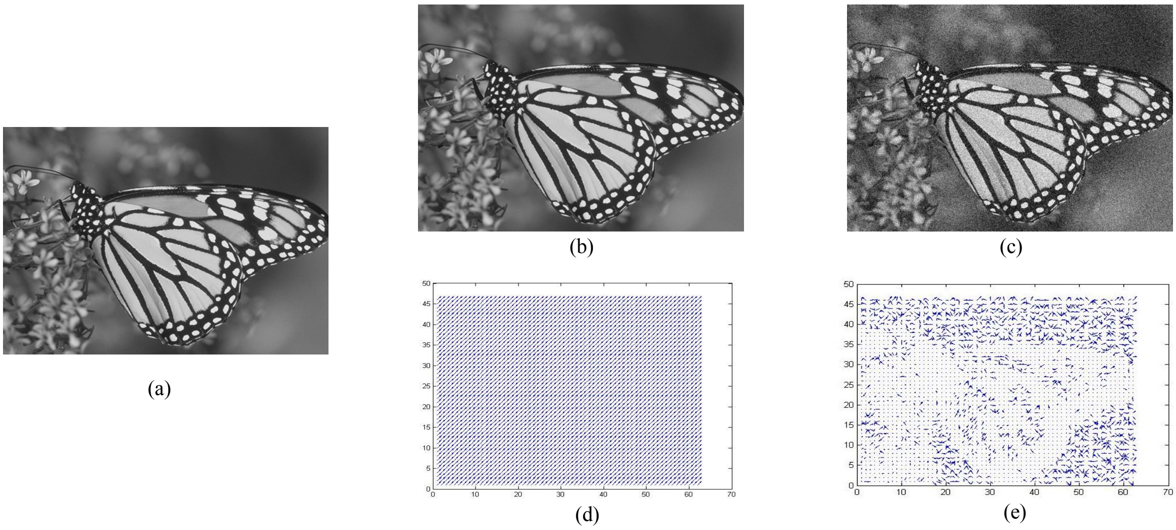


Fig. 2 (a) The original image (monarch); (b) the original image translated by the displacement vector of (1, 1), (MOS=0.999 HCI=1.0, SSIM=0.508); (c) the original image contaminated by AWGN (MOS=0.689 HCI=0.625, SSIM=0.638); (d) the displacement vector field associated with the translated image of (b); (d) the displacement vector field associated with the distorted image of (c).

occurrences of all possible displacement vectors are approximately the same. That is

$$H_{\mathbf{max}} = \log_2 N\,. \tag{6}$$

In the case that two identical images differentiated by simple geometric translation are in comparison, the corresponding blocks from the two images are expected to have the same displacement. The homogeneous correspondence will result in the near-zero value of $R(\mathbf{X},\mathbf{Y})$ and thus the highest value of $S_H(\mathbf{X},\mathbf{Y})$. As a demonstration, Fig. 2(d) displays the uniformity of the displacement vectors that associate the corresponding blocks (of size 8×8) in the reference image (Fig. 2(a)) and the same image translated by a displacement vector of (1, 1) (Fig. 2(b)). Obviously, the slight translation does not affect the subjective quality of the translated image. A high score value is thus expected from the subjective quality assessment, and so should be from an effective objective quality metric. Fig. 2 (e) displays the irregularity of the displacement vectors that associate the corresponding blocks in the reference image (Fig. 2 (a)) and the same image contaminated by AWGN (Fig. 2 (c)). The lower score values evaluated from the subjective and an effective objective quality metrics are therefore expected.

By considering that the human visual system is sensitive to the change in luminance, the similarity measure in luminance is also included in the proposed image quality assessment method. $S_L(\mathbf{X},\mathbf{Y})$ in Eq.(1) functions to measure the similarity in the luminance of the two images, where the luminance correlations of all pairs of corresponding blocks are averaged to contribute to the metric score, as can be expressed by the following equation

$$S_L(\mathbf{X},\mathbf{Y}) = \frac{1}{M}\sum_{l=1}^{M}\frac{2\mu(\mathrm{x}_l)\mu(\mathrm{y}_l)+K}{\mu(\mathrm{x}_l)^2+\mu(\mathrm{y}_l)^2+K} \tag{7}$$

where $\mathbf{x}_l$ denotes the l-th block in the reference image corresponding to block $\mathbf{y}_l$ in the test image. $\mu(\mathbf{x}_l)$ and $\mu(\mathbf{y}_l)$ denote the mean luminance of the image block $\mathbf{x}_l$ and $\mathbf{y}_l$, respectively. The constant K is used to avoid the instability due to zero mean luminance, and control the dynamic range of the score value.

In the process of locating the corresponding block for each block in the test image, the full-search block matching algorithm is applied to the reference image data within the search area defined by the location of the test image block and the parameter of maximum possible displacement. The criterion for determining the best matched block is, in this paper, the mean squared error evaluated from mean-removed image blocks.

III. SIMULATION

To investigate the validity of the proposed metric, the test images contained in the LIVE [16] and TID2008 [17] databases are employed in the simulation. Some of the test images are modified by the translations of various extents for measuring the translation affect. Due to the absence of the subjective evaluation results on the translated images, the subjective experiments that follow the ITU-R procedure [18] were conducted to assess the quality of the translated images, where 22 subjects were involved in the experiment. The mean opinion scores (MOS) obtained from assessing each translated image in the subjective experiment is normalized to the range from 0.0 to 1.0. In evaluating the score of the proposed metric, non-overlapping 8×8 blocks are sampled in the test image for reducing the intensive computational complexity, particularly when search area is enlarged.

The performance of the proposed quality metric is first investigated by the correlation between the subjective evaluation results and the objective evaluation results as the extent of translation varies. Fig. 3 illustrates the behavior of the averaged scores evaluated from five image quality metrics as the extent of translation is increased. The test image, in this case, is identical to the reference image but displaced by a specified number of pixels along the horizontal direction. The curves indicate that the behavior of the proposed metric closely agrees with that of subjective metric. The proposed metric is more correlated with the subjective metric than other metrics. The results also verify that the CW-SSIM metric is only robust to small translations (less than 25 pixels).

The performance of the proposed quality metric is also compared with those of the metrics including SSIM [5], CW-SSIM [13-14], VIF [9], VSNR [1], PSNR in terms of Pearson's correlation coefficients. In this case, non-linear regression is used to fit the experimental data that model the correlation of the objective and subjective metric scores. The comparison results obtained from assessing more than 700 images in the LIVE database are shown in Table I. These images are contaminated by five types of distortion including those due to JPEG and JPEG2000 compression, additive white noises, Gaussian blurring and fast fading. The images distorted by geometric translation are not involved in this comparison. The comparison results shown in Table II are obtained from assessing even wider spectrum of test images including the general test images and translated test images. These test images are obtained from the TID2008 and LIVE databases.

From the experimental results, it can be found that the proposed metric is able to reflect an accurate image quality assessment while the test image occur the translation effect as show in Fig. 4 (e)-(h). The comparison results shown in Tables I and II indicate that the proposed metric is not only insensitive to the distortion due to geometric translation but also has the performance comparable to the existing well-known metrics (such as SSIM and VIF) in predicting the perceptual quality of the test image.

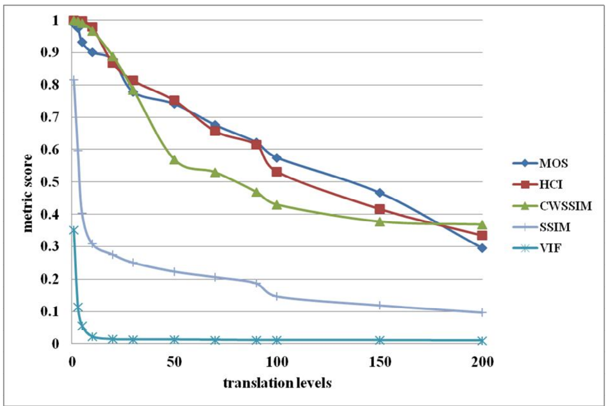


Fig. 3 The behavior of five metric scores as the extent of the translation between two identical images increases.

TABLE I
PERFORMANCE COMPARISON IN PEARSON'S CORRELATION COEFFICIENT (FROM IMAGES IN THE LIVE DATABASE ONLY)

	SSIM	VIF	VSNR	PSNR	CWSSIM	HCI
Jpg2k	0.969	**0.972**	0.894	0.959	0.867	0.941
Jpeg	0.961	**0.977**	0.898	0.956	0.905	0.956
WN	0.989	**0.992**	0.939	0.989	0.878	0.923
Blur	0.963	**0.978**	0.869	0.819	0.848	0.916
Fading	0.925	**0.979**	0.874	0.951	0.875	0.943
Average	0.946	**0.977**	0.883	0.936	0.863	0.941

TABLE II
PERFORMANCE COMPARISON IN PEARSON'S CORRELATION COEFFICIENT (COMBINED DISTORTION/TRANSLATION DATABASE)

	SSIM	VIF	VSNR	PSNR	CWSSIM	HCI
Jpg2k	0.598	0.589	0.232	0.322	0.898	**0.942**
Jpeg	0.087	0.387	0.043	0.191	0.855	**0.951**
WN	0.569	0.359	0.312	0.515	0.747	**0.924**
Blurring	0.241	0.546	0.053	0.065	0.805	**0.901**
Fading	0.139	0.074	0.396	0.231	0.788	**0.910**
Translation	0.706	0.454	0.617	0.722	0.921	**0.987**
Average	0.333	0.293	0.218	0.293	0.869	**0.935**

IV. CONCLUSIONS

In this paper, a novel image quality assessment method that measures the uniformity of the displacements that associate the corresponding blocks from the images to be compared is proposed. The method is verified not only robust to the translation occurred in the image to be assessed but also has the ability to predict the image quality in a quantified level close to that judged by human beings. The proposed metric can cooperate with the existing quality metrics to further enhance the overall performance in predicting image quality

REFERENCES

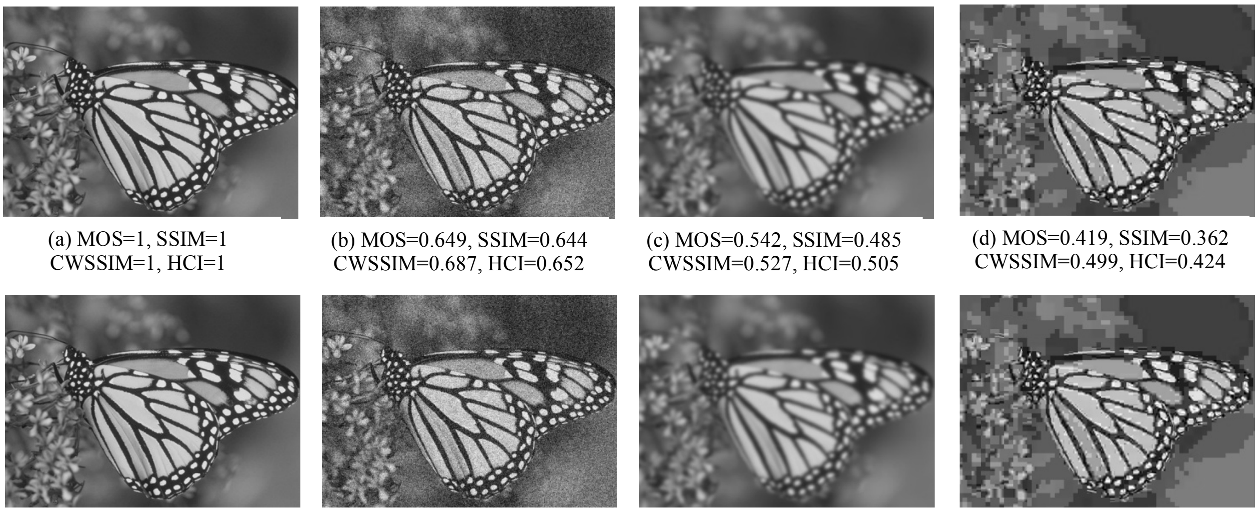
[1] D. M. Chandler and S. S. Hemami, "VSNR: A wavelet-based visual signal-to-noise ratio for natural images," *IEEE Trans. Image Process.*, vol. 16, no. 9, pp. 2284–2298, Sep. 2007.

[2] N. Damera-Venkata, T. D. Kite, W. S. Geisler, B. L. Evans, and A. C. Bovik, "Image quality assessment based on a degradation model," *IEEE Trans. Image Process.*, vol. 9, no. 4, pp. 636–650, Apr. 2000.

[3] N. Damera-Venkata, T. D. Kite, W. S. Geisler, B. L. Evans, and A.C. Bovik, "Image quality assessment based on a degradation model,"*IEEE Trans. Image Process.*, vol. 9, no. 4, pp. 636–650, Apr. 2000.

[4] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Processing Letters*, vol. 9, pp. 81–84, Mar. 2002.

[5] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From Error Visibility to Structural Similarity" *IEEE Trans. Image Processing*, vol. 13, no. 4, pp. 600−612, April. 2004.



(a) MOS=1, SSIM=1 CWSSIM=1, HCI=1

(b) MOS=0.649, SSIM=0.644 CWSSIM=0.687, HCI=0.652

(c) MOS=0.542, SSIM=0.485 CWSSIM=0.527, HCI=0.505

(d) MOS=0.419, SSIM=0.362 CWSSIM=0.499, HCI=0.424

(e) MOS=0.999, SSIM=0.608 CWSSIM=0.987, HCI=0.999

(f) MOS=0.647, SSIM=0.438 CWSSIM=0.671, HCI=0.649

(g) MOS=0.539, SSIM=0.303 CWSSIM=0.515, HCI=0.506

(h) MOS=0.415, SSIM=0.312 CWSSIM=0.487, HCI=0.423

Fig. 4 Comparison of image quality measures for images with different types of distortions and translation. (a)-(d) non-translation images (a) original image , (b) white noise image, (c)blur image, (d)jpeg image; (e)-(h) Add translation insensitive to the test images by 5 pixels to the left

[6] A. Shnayderman, A. Gusev, and A. M. Eskicioglu, "An SVD-based grayscale image quality measure for local and global assessment," *IEEE Trans. Image Processing*, vol. 15, no. 2, pp. 422–429, Feb. 2006.

[7] F. Torkamani-Azar and S. A. Amirshahi, "A new approach for image quality assessment using svd," in *9th International Symposium on Signal Processing and Its Applications (ISSPA 2007)*, Sharjah, United Arab Emirates (UAE)., Feb. 2007, pp. 1 – 4.

[8] H. R. Sheikh, A. C. Bovik, and D. de Veciana, "An information fidelity criterion for image quality assessment using natural scene statistics," *IEEE Trans. Image Processing*, vol. 14, no. 12, pp. 2117 – 2128, Dec. 2005.

[9] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Processing*, vol. 15, no. 2, pp. 430 – 444, Feb. 2006.

[10] C. Li and A. C. Bovik, "Three-component weighted structural similarity index," in *Proc. SPIE*, 2009, vol. 7242, pp. 72420Q-1–72420Q-9.

[11] Z. Wang, L. Qiang, "Information Content Weighting for Perceptual Image Quality Assessment," *IEEE Trans. Image Processing*, vol. 20, pp. 1185–1198, 2011.

[12] Lin Zhang, Lei Zhang, X. Mou and D. Zhang. FSIM: A Feature Similarity Index for Image Quality Assessment. *IEEE Transactions on Image Processing*, 2011.

[13] M. P. Sampat, Z. Wang, S. Gupta, A. C. Bovik, and M. K. Markey, "Complex wavelet structural similarity:A newimage similarity index," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2385–2401, Nov. 2009.

[14] Z. Wang and E. P. Simoncelli, "Translation insensitive image similarity in complex wavelet domain," *IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, vol. II, pp. 573-576, Philadelphia, PA, March (2005) .

[15] M. Prieto and A. Allen, "A similarity metric for edge images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 10, pp. 1265–1273, Oct. 2003.

[16] H. R. Sheikh, K. Seshadrinathan, A. K. Moorthy, Z.Wang, A. C. Bovik, and L. K. Cormack, Image and Video Quality Assessment Research at LIVE 2004 [Online]. Available: http://live.ece.utexas.edu/research/quality

[17] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti, "TID2008—A database for evaluation of full-reference visual quality assessment metrics," *Adv. Modern Radioelectron.*, vol. 10, pp. 30–45, 2009.

[18] VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment phase II," Aug. 2003, http://www.vqeg.org/