Perceived Quality Assessment of Color Images

Mohammed Ahmed Hassan
Assistant Professor, Department of applied and Computer Sciences
Seiyun Community College
Seiyun, Yemen

***Abstract*—Image quality assessment is an important tool in many image-processing applications such as image acquisition, watermarking, compression, transmission, restoration, enhancement, and reproduction. In this paper, we present a metric for assessing the visual quality of color images. The proposed metric is designed to measure the perceivable distortion in the CIELAB color space. The CIELAB just noticeable color difference (JNCD) is used as the visibility threshold of distortion for each color pixel. Simulation results show that the assessment of the proposed image quality metric is more correspondent with the subjective assessment than other state-of-art metrics.**

Keywords—image quality metrics; CIELAB color space; just noticeable color difference (JNCD).

I. INTRODUCTION

Today with the increase concern in research and development with digital imagery, there is a real need for image quality assessment methods that quantify how a distorted image looks compared to the original one as perceived by a human observer. Image quality assessment methods can be classified into two categories: subjective and objective. The subjective image quality assessment methods are accurate in estimating the visual quality of an image because they are carried out by human subjects but are costly process that requires a large number of observers and takes a significant time. On the other hand the objective image quality assessment methods are computer-based methods that can automatically predict the perceived image quality. Hence the objective image quality assessment methods gained more popularity.

Objective image quality assessment methods also may be classified into full reference, reduced reference, and no reference methods based on the availability of the reference image. Full reference image quality assessment requires complete information about the reference image; and partial information about the reference image is required for the reduced reference image quality assessment; while no information about the reference image is needed in no reference image quality assessment. This paper focuses on the full reference image quality assessment methods for color images where both the original and the test images are available.

Many researchers have contributed significant research in the design of objective image quality methods starting from the widely used mean square error (MSE) metric and its correlated peak signal to noise ratio (PSNR). The weighted signal to noise ratio (WSNR) [1] simulates the human visual system properties by filtering both the reference and distorted images with contrast sensitivity function and then compute the SNR. Miyahara [2] proposed a picture quality scale (PQS) based on three distortion factors: the amount, location and structure of error. The perceptual color fidelity metric (S-CIELAB)[3] is a spatial extension to the CIELAB metric for measuring color reproduction errors of digital images. It simulates the spatial sensitivity of the human visual system by spatial filter process on images. Wang and Bovik [4] proposed a new universal image quality index (UQI) and its improved form the single-scale structural similarity index (SSIM) [5] by modeling the image distortion as the combination of loss of luminance, contrast, and correlation. In [6] the single-scale structural similarity index was extended into a multi-scale structural similarity index (MSSIM) that works in multi-scales of an image and achieved a better result than SSIM. Information fidelity criterion (IFC) [7] and visual information fidelity (VIF) [8] both are based on information theory in which the distorted image is modeled as a sequence of passing the reference images through distortion channels and quantify the visual quality as mutual information between the test image and the reference image. Shnayderman et al. [9] explored the feasibility of singular value decomposition (SVD) for quality measurement. In [10] a two a two-staged wavelet based visual signal to noise ratio (VSNR) was proposed based on the low-level and the mid-level properties of human vision. A structural information-based image quality assessment algorithm [11] uses LU factorization for representation of the structural information of an image. An image quality metric using the phase quantization code [12] was proposed and extended to amplitude/phase quantization code [13]. Wang and Li [14] incorporated the idea of information content weighted pooling and applied it to peak signal to noise ratio (PSNR) and structural similarity measure (SSIM) leading to an information content weighted PSNR (IW-PSNR) and an information content weighted SSIM (IW-SSIM). In [15] a feature similarity index (FSIMc) for color image quality assessment is proposed based on the fact that human visual system understands an image mainly according to its low-level features. Specifically, two kinds of features, the phase congruency (PC) and the image gradient magnitude (GM) are used in FSIMc.

II. CIELAB COLOR SPACE

The CIELAB color spaces is considered to be perceptually uniform and referred to as uniform color spaces in which the Euclidean distance between any two different colors in the color space correspond approximately to the difference perceived by the human vision between the two colors. This color space was established by CIE (International Commission on Illumination) as nonlinear transformations of tristimulus XYZ values to overcome the non-uniformity of color spaces that had been discussed by MacAdam [16]. The three coordinates of CIELAB (L, a, and b) represent the lightness of the color, its position between red/magenta and green, and its position between yellow and blue respectively. The CIE recommended using XYZ coordinate system to transform RGB values to CIELAB as following:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.4124564 & 0.3575761 & 0.1804375 \\ 0.2126729 & 0.7151522 & 0.0721750 \\ 0.0193339 & 0.1191920 & 0.9503041 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (1)$$

And then from XYZ to CIELAB color space

$$L^* = \begin{cases} 116(Y/Y_n)^{1/3} - 16, & \text{if } Y/Y_n > 0.008856 \\ 903.3(Y/Y_n), & \text{if } Y/Y_n \le 0.008856 \end{cases} \quad (2)$$

$$a^* = 500(f(X/X_n) - f(Y/Y_n)) \quad (3)$$

$$b^* = 200(f(Y/Y_n) - f(Z/Z_n)) \quad (4)$$

where

$$f(t) = \begin{cases} t^{1/3}, & \text{for } t > 0.008856 \\ 7.787 * t + 16/116, & \text{for } t \le 0.008856 \end{cases} \quad (5)$$

Here $Y_n = 1.0$ is the luminance, and $X_n = 0.950455$, $Z_n = 1.088753$ are the chrominances for the D_{65} white point.

III. The Proposed Method

A useful rule of thumb in CIELAB color space is that any two colors can be distinguished if the Euclidean distance between these two colors is greater than threshold 2.3 [17]:

$$\Delta E_{Lab} = \sqrt{\Delta L^2 + \Delta a^2 + \Delta b^2} > 2.3 \quad (6)$$

This distortion threshold is known as the Just Noticeable Color Difference (JNCD) threshold. Therefore all the colors within a sphere of radius equal to the JNCD threshold are perceptually indistinguishable from each other.

To estimate the perceived quality of a given distorted color image, the original and distorted images are first transformed to CIELAB color space for measuring the associated JNCD threshold and then the proposed image quality metric, termed as Improved CIELAB, is defined as:

$$\Delta IE = \frac{1}{MN} \sum_{i}^{N} \sum_{j}^{M} d(I_{ij} - I'_{ij}) \delta_{ij}$$
$$\delta_{ij} = \begin{cases} 1 & \text{if } d(I_{ij} - I'_{ij}) > JNCD_{Lab} \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

where d is the Euclidean distance between two colors in the CIELAB color space, M and N are the image dimensions.

IV. Resutts and Discussion

In this section, the performance of the proposed image quality measure in terms of the ability of predicting the subjective ratings is analyzed. We used the popular Tampere Image Database (TID2008) [18] to test the performance of proposed quality measure. This database is the most recent and largest database so far available that includes more images and more distortion types for verification of full reference quality metrics. The TID20 08 database contains 1700 distorted images (25 reference images ×17 types of distortions×4 levels of distortions). Mean Opinion Scores for this database have been obtained as a result of 838 subjective experiments. During these tests, observers from three countries (Finland, Italy, and Ukraine) have carried out about 256000 individual human quality judgments.

The proposed quality measure was applied to the set of images used in the TID2008 and the results were compared to the subjective MOS. For comparison, the same set of images were presented to 6 well-known objective image quality measures that are commonly used and their implementations are publicly available on the Internet namely: universal image quality index (UQI) [4], structural similarity index (SSIM) [5], multiscale structural similarity index (MSSIM) [6], information fidelity criterion (IFC) [7], visual information fidelity (VIF) [8], and Visual Signal to Noise Ratio (VSNR) [10].

TABLE I. Pearson's Correlation Coefficient of the Scores Given by Different Image Quality Assessment Metrics Against MOS from TID2008 Image Database

	SSIM	MSSIM	UQI	IFC	VIF	VSNR	ΔIE
Additive Gaussian noise	0.767	0.748	0.532	0.581	0.867	0.745	0.922
Additive noise in color components	0.785	0.778	0.482	0.535	0.895	0.764	0.925
Spatially correlated noise	0.796	0.76	0.551	0.611	0.859	0.75	0.934
Masked noise	0.731	0.787	0.760	0.730	0.892	0.753	0.778
High frequency noise	0.821	0.822	0.685	0.712	0.945	0.883	0.959
Impulse noise	0.632	0.625	0.565	0.493	0.815	0.624	0.818
Quantization noise	0.791	0.757	0.553	0.110	0.745	0.813	0.818
Gaussian blur	0.878	0.877	0.890	0.871	0.939	0.916	0.656
Image denoising	0.914	0.915	0.803	0.712	0.898	0.919	0.902
JPEG compression	0.93	0.931	0.796	0.782	0.932	0.906	0.914
JPEG2000 compression	0.952	0.939	0.914	0.819	0.917	0.934	0.845
JPEG transmission errors	0.828	0.824	0.843	0.775	0.872	0.647	0.658
JPEG2000 transmission errors	0.831	0.788	0.685	0.699	0.831	0.761	0.658
Non eccentricity pattern noise	0.661	0.665	0.734	0.836	0.736	0.566	0.651
Local block-wise distortions of different intensity	0.872	0.796	0.852	0.703	0.834	0.273	0.787
Mean shift (intensity shift)	0.727	0.669	0.563	0.484	0.592	0.247	0.739
Contrast change	0.70	0.769	0.469	0.296	0.883	0.428	0.445

TABLE II. SPEARMAN'S RANK ORDER CORRELATION COEFFICIENT OF THE SCORES GIVEN BY DIFFERENT IMAGE QUALITY ASSESSMENT METRICS AGAINST MOS FROM TID2008 IMAGE DATABASE

	SSIM	*MSSIM*	*UQI*	*IFC*	*VIF*	*VSNR*	*ΔIE*
Additive Gaussian noise	0.831	0.809	0.521	0.582	0.88	0.773	0.915
Additive noise in color components	0.813	0.806	0.474	0.553	0.878	0.779	0.914
Spatially correlated noise	0.844	0.82	0.540	0.598	0.87	0.766	0.922
Masked noise	0.756	0.816	0.737	0.733	0.87	0.729	0.796
High frequency noise	0.892	0.868	0.671	0.736	0.907	0.881	0.932
Impulse noise	0.707	0.687	0.588	0.533	0.833	0.647	0.900
Quantization noise	0.875	0.854	0.548	0.591	0.796	0.827	0.838
Gaussian blur	0.96	0.961	0.895	0.877	0.955	0.933	0.862
Image denoising	0.96	0.957	0.782	0.800	0.919	0.929	0.900
JPEG compression	0.927	0.935	0.773	0.818	0.917	0.917	0.914
JPEG2000 compression	0.972	0.974	0.918	0.944	0.971	0.952	0.853
JPEG transmission errors	0.867	0.874	0.837	0.797	0.858	0.805	0.692
JPEG2000 transmission errors	0.871	0.852	0.675	0.730	0.851	0.791	0.736
Non eccentricity pattern noise	0.717	0.734	0.747	0.841	0.761	0.572	0.702
Local block-wise distortions of different intensity	0.853	0.762	0.808	0.677	0.832	0.193	0.801
Mean shift (intensity shift)	0.757	0.737	0.631	0.438	0.513	0.371	0.757
Contrast change	0.633	0.64	0.489	-0.275	0.819	0.424	0.444

Two criteria were used to evaluate the performance of the image quality metrics. These criteria characterize two attributes related to the prediction of each image quality metric [19]:

1. *Prediction Accuracy:* The ability of an objective image quality metric to predict the subjective MOS with minimum average error. The Pearson's linear correlation coefficient was used to measure the prediction accuracy.
2. *Prediction Monotonicity*: The ability of given by an objective image quality metric to give values that are monotonic in their relationship to the corresponding subjective MOS values. This attribute was measured by the Spearman's rank order correlation coefficient.

Tables I and II show Pearson's correlation coefficient and Spearman's rank order correlation coefficient of scores given by several image quality assessment metrics for individual distortions from TID2008 image database. It is clear from those tables that the proposed metric outperforms the other metrics for the quality assessment of many distortion types such as Additive Gaussian noise, Additive noise in color components, Spatially correlated noise, High frequency noise, Impulse noise, Quantization noise, and Mean shift. While it has a competitive performance with the other metrics for Masked noise, Image denoising, and JPEG compression.

V. CONCLUSION AND FUTURE WORK

In this paper, we presented a mathematically simple but novel metric for the quality assessment of color images. This metric is based on the Just Noticeable Color Difference (JNCD) threshold of the perceptually uniform color space, CIELAB. Results show that the proposed quality metric provides ratings that are more consistent with human perception of color image quality for many distortion types. Future work includes taking into account the different sensitivities of the human visual system (HVS) to the contents of the images where many studies have found that the perceptibility of color difference depends on the contents of the images.

REFERENCES

[1] T. Mitsa and K. Varkur, "Evaluation of contrast sensitivity functions for the formulation of quality measures incorporated in halftoning algorithms," in *Proc. IEEE International Conference on Acoustic, Speech, and Signal processing*, pp. 301–304, 1993.

[2] M. Miyahara, K. Kotani, and V. Algazi, "Objective picture quality scale (PQS) for image coding," *IEEE Trans. Communications,* vol. 46, no. 9, pp.1215–1226, 1998.

[3] X. Zhang and B. Wandell, "A spatial extension of CIELAB for digital color image reproduction," in *Proc. SID International Symposium Digest of Technical Papers*, vol. 27, pp. 731–734, 1996.

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

[5] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Processing*, vol. 13, no. 4, pp. 600-612, 2004.

[6] Z. Wang, E. Simoncelli, and A. Bovik, "Multi-scale structural similarity for image quality assessment," in *Proc. 37th IEEE Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 1398–1402, 2003.

[7] H. Sheikh, A. Bovik, and G. 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, 2005.

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

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

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

[11] H. S. Han, D. O. Kim, and R. H. Park, “Structural information-based image quality assessment using LU factorization, “ *IEEE Trans. on Consumer Electronics*, vol. 55, no. 1, pp. 165–171, 2009.

[12] D. O. Kim and R. H. Park, “Image quality measure using the phase quantization code,” *IEEE Trans. on Consumer Electronics*, vol. 56, no. 2, pp. 937–945, 2010.

[13] D. O. Kim and R. H. Park, “Image quality assessment using the amplitude/phase quantization code,” *IEEE Trans. on Consumer Electronics*, vol. 56, no. 4, pp. 2756–2762, 2010.

[14] Z. Wang and Q. Li, “Information content weighting for perceptual image quality assessment,” *IEEE Trans. on Image Processing*, vol. 20, no. 5, pp. 1185–1198, 2011.

[15] L. Zhang, L. Zhang, X. Mou, and D. Zhang, “FSIM: A feature similarity index for image quality assessment,” *IEEE Trans. on Image Processing*, vol. 20, no. 8, pp. 2378-2386, 2011.

[16] D. L. MacAdam, “Specification of small chromaticity differences in daylight,” Journal of the Optical Society of America, vol. 33, no. 1, pp. 18-26, Jan. 1943.

[17] M. Mahy, L. Van Eyckden, and A. Oosterlinck, “Evaluation of uniform color spaces developed after the adoption of CIELAB and CIELUV,” *Color Res. Appl.*, vol. 19, pp. 105-121, Apr. 1994.

[18] 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,” Advances of Modern Radioelectronics, vol. 10, pp. 30-45, 2009.

[19] ITU-R, “Methodology for the subjective assessment of the quality for television pictures,” *Recommendation ITU-R BT.500-11*, Geneva, 2002.