A New Method of Image Quality Assessment

SHUANG LIANG Peking University School of Mathematical Sciences CHINA liangshuang12@pku.edu.cn

GUANXIANG WANG Peking University School of Mathematical Sciences No.5 Yiheyuan Road Haidian District, Beijing No.5 Yiheyuan Road Haidian District, Beijing **CHINA** gxwang@math.pku.edu.cn

SHULI WANG Peking University School of Mathematical Sciences No.5 Yiheyuan Road Haidian District, Beijing No.5 Yiheyuan Road Haidian District, Beijing CHINA shuliw@pku.edu.cn

YU WANG Peking University School of Mathematical Sciences CHINA wangyu amo@pku.edu.cn

Abstract: In image processing, one of the important items is how to measure the quality of an image. This paper attempts to establish an assessment model corresponding to mean human perception, or mean opinion score (MOS). Human vision system has a feature of being more sensitive to the brightness change of an area than to that of discrete points, and being more sensitive to the change of moderate brightness than to the change of very bright one and very dark one. According to such a feature, by defining brightness discrimination and error density, using gradient as well, this paper proposes a new method of image quality assessment. The new method is tested on TID2008 data and the results are compared with those of existing methods.

Key–Words: image quality assessment, brightness discrimination, error density, gradient

1 Introduction

Along with the recent advances in digital imaging and communication technologies, the levels of the image acquisition, processing, compression, storage and transmission have been significantly improved. However, no matter how good equipments are used, the degeneration of image in above processes is unavoidable. As a result, it is necessary to establish a uniform assessment system which can well correspond to the subjective score by human beings [1, 10, 11, 20, 21, 23]. According to the amount of information from real scenes, image quality assessments are classified as full-reference ones, no-reference ones and reducedreference ones [2]. We here study the problem of image quality assessment with full-reference [14].

In present literatures, there are four kinds of models to solve the problem of image quality assessment with full-reference. The first kind are based on principles of statistics, such as mean square error (MSE) [22] and peak value signal to noise ratio (PSNR) [2]. The traditional statistical methods are widely used because of the simplicity and definite physical meaning. The second kind use analysis of the structure information of images, such as weighted-peak signal

to noise ratio (WSNR) [3, 4], noise quality measures (NQM) [4], universal quality index (UQI) [5], visual signal to noise ratio (VSNR) [12]and structural similarity (SSIM) [2, 6, 7, 13, 17]. Such models are obviously superior to the traditional statistical models because they focus on the sensitivity of human vision to, as well as on the structure information composed of, different parts of an image. The third kind pay main attention to the natural scene of images, such as visual information fidelity (VIF) [9], information fidelity criterion (IFC) [8]. These methods mainly use wavelet transform to extract and compare the scene information of the original image and the degraded image. Such methods are more accurate to images degraded by changes of brightness and contrast. The last kind use feature similarity (FSIM) based on image texture analysis and representative methods [15]. Liu et.al have proposed phase congruency [19] and gradient modulus [18], combined such important information as brightness [15]. These methods are comprehensive and of higher accuracy.

In this paper, we propose a new method of image quality assessment based on brightness discrimination, error density and gradient (BDEDG). According to the feature of human vision system, human beings are more sensitive to the brightness change of an area than that of discrete points, so we define the error density of an area in an image. In addition, because the human beings are more sensitive to the change of appropriate brightness than to that of very bright or very dark scenes, we define the brightness discrimination to measure the sensitivity. Making use of error density, brightness discrimination and gradient of an image, we establish the model BDEDG to measure the quality of the image. The proposed BDEDG is tested on TID2008 database [16]. Compared with some current state-of-art methods, our method gives a satisfying performance.

The rest of the paper is organized as follows. In Section 2, the detail of the proposed method BDEDG is presented. The experiments results and comparisons are shown in Section 3. Finally, the conclusion is given in Section 4.

2 The proposed BDEDG

It is a fact that the sensitivity of human vision to different areas with different texture densities and different brightness is different. Therefore we propose a new image quality assessment based on brightness discrimination, error density and gradient (BDEDG).

Denote the original image as $I_{ori} = (I_{ori}(i,j))_{S \times T}$ and the degraded image as $I_{deg} = (I_{deg}(i,j))_{S \times T}$. The absolute difference between I_{ori} and I_{deg} is

$$Error(i, j) = |I_{ori} - I_{deg}|, i = 1, ..., S, j = 1, ..., T.$$

Obviously, it will not cause any significant response of human vision system even if Error(i,j) are very large on some individual points. That is to say, human vision is sensitive to the brightness change of an area instead of individual points. So we define the error density of point (i,j) as

$$DError(i,j) = \frac{1}{(2w+1)^2} \sum_{s=i-w}^{i+w} \sum_{t=j-w}^{j+w} Error(s,t),$$

where the local area of size $(2w+1) \times (2w+1)$ centers at (i,j).

Usually human vision is not sensitive to the brightness change in very bright or very dark areas of an image, but sensitive to the change of appropriate brightness. So we define the brightness discrimination to measure the sensitivity of human vision to brightness as

$$Bd(i,j) = \Phi(\frac{I_{ori}(i,j)}{255}), i = 1,...,S, j = 1,...,T,$$

where $\Phi(x) = x(1 - x)$. The value of Bd is small when the brightness is too high or too low.

Another fact is that human vision is more sensitive to the brightness change in the smooth area than in the area of complicated texture. So it is reasonable to take the gradient modulus $G(i,j) = gradient(I_{ori}(i,j)), i = 1, ..., S, j = 1, ..., T$ into consideration.

Thus we define a new image quality assessment based on brightness discrimination, error density and gradient (BDEDG) as

$$\gamma = \frac{1}{S \times T} \sum_{i,j} \frac{[Bd(i,j) \cdot Error(i,j) \cdot DError(i,j)]^{\lambda_1}}{[G(i,j) + c]^{\lambda_2}},$$
$$BDEDG = exp(-\lambda_3 \gamma).$$

Where $\lambda_1, \lambda_2, \lambda_3, c > 0$. The parameter c is a small value to ensure that the denominator isn't zero.

3 Experimental results

TID2008 database includes totally 1700 distorted images which degenerate from 25 original images and 17 kinds of distortion with 4 levels. There are 100 images distorted by one kind of distortion. Specifically, 17 kinds of distortion are: 1-additive Gaussian noise, 2-additive noise in color components, 3spatially correlated noise, 4-masked noise, 5-high frequency noise, 6-impulse noise, 7-quantization noise, 8-gaussian blur, 9-image denoising, 10-JPEG compression, 11-JPEG2000 compression, 12-JPEG transmission errors, 13-JPEG2000 transmission errors, 14non eccentricity pattern noise, 15-local block-wise distortions of different intensity, 16-mean shift (intensity shift) and 17-contrast change.

When make test on TID2008 database, for one model of quality assessment and one kind of distortion, we calculate the correlation coefficient of the 100 scores gotten by this assessment and the corresponding 100 scores in MOS. So we have 17 correlation coefficients for one model of quality assessment tested on TID2008.

To compare two sets of data, correlation analysis methods are used. Here, correlation analysis aims to evaluate the effectiveness of an objective assessment method by comparing the result with the subjective score given by human beings. Three methods are widely used for correlation analysis: Spearman Rank Order Correlation Coefficient (SROCC), Kendall Rank Order Correlation Coefficient (KROCC) and Pearson Linear Correlation Coefficient (PLCC) [23].

To find the optical optimal values of the parameters in BDEDG, 5 values of λ_1 , 5 values of λ_2 and 15

values of λ_3 are tested. Let $\lambda_1, \lambda_2 = 0.25, 0.5, 1, 2, 3$, and $\lambda_3 = 0.1 : 0.1 : 1.5$. For every set of values of (λ_1, λ_2) , find the best value of λ_3 . The results are shown in table I, where the results in the cell of the *i*th line and *j*th row are calculated with $\lambda_1 = \lambda_1(i), \lambda_2 = \lambda_2(j)$. In every (i,j) cell, the value of λ_3 is the best value (not necessarily unique), where P, S and K represent the numbers of superiors of BD-EDG compared to other 11 models corresponding to the 3 kinds of correlation respectively. Note that, for each kind of correlation, there are 187 correlation coefficients for 11 assessment models and 17 kinds of distortion. In Table 1, computation shows that (1,3,1) is the optimal set of parameter $(\lambda_1, \lambda_2, \lambda_3)$.

The proposed BDEDG model is tested on TID2008 database, and the result is compared with 11 current state-of-art methods of image quality assessment with full-reference. Correlation coefficients computed with BDEDG in sense of SROCC, KROCC and PLCC are displayed in Table 2-4 respectively.

As shown in Table 2, the SROCC values of BDEDG, among the 17 kinds of image distortion, are superior to MSE, PSNR, SSIM, MSSIM, NQM, SNR, UQI, VIF, VIFP, VSNR, WSNR in 11,11,9,8,15,14,13,11,11,15,11 cases respectively.

As shown in Table 3, the KROCC values of BDEDG, among the 17 kinds of image distortion, are superior to MSE, PSNR, SSIM, MSSIM, NQM, SNR, UQI, VIF, VIFP, VSNR, WSNR in 11,11,9,9,13,13,14,10,11,16,12 cases respectively.

As shown in Table 4, he PLCC values of BDEDG, among the 17 kinds of image distortion, are superior to MSE, PSNR, SSIM, MSSIM, NQM, SNR, UQI, VIF, VIFP, VSNR, WSNR in 12,12,11,12,14,14,12,12,16,12 cases separately.

In summary, our method BDEDG gives a satisfying performance on database TID2008.

4 Conclusion

In this paper, we study the problem of image quality assessment with full-reference. We propose a new method of image quality assessment based on brightness discrimination, error density and gradient (BDEDG). We test our proposed method BDEDG on TID2008 database, compared it with 11 current stateof-art methods of image quality assessment with fullreference by computing the correlation coefficients of SROCC, KROCC and PLCC separately. The experiments have shown that our method BDEDG performs very well on TID2008 database.

References:

- [1] W. Zhou and A. C. Bovik, *Modern image quality assessment*, New York: Morgan and Claypool Publishing Company, 20-30, 2006.
- [2] W. Zhou, A. C. Bovik, H. R. Sheikh, et al., Image quality assessment: from error visibility to structure similarity, *IEEE Trans. on Image Processing*, 13(4), 2004, pp. 600–612, .
- [3] T. Mitsa and K. Varkur, Evaluation of contrast sensitivity functions for the formulation of quality measures incorporated in half toning algorithms, *ICASSP* 1993, pp. 301–304.
- [4] N. Damera-Venkata, T. D. Kite, W. S. Geisler, et al., Image quality assessment based on a degradation model, *IEEE Trans. on Image Processing* 9, 2000, pp. 636–650.
- [5] W. Zhou and A. C. Bovik, A universal image quality index, *IEEE Trans on. Signal Processing Letters* 9, 2002, pp. 81–84.
- [6] W. Zhou, P. Simoncelli, A. C. Bovik and H. R. Sheikh, Multiscal structure similarity for image quality assessment, *Proc. of the 37th Asilomar Conference on Signals, Systems and Computers*, 2003, pp. 1398–1402.
- [7] G. Chen, C. Yang and Xie S.,, Gradient-based structure similarity for image quality assessment., *Proc. Of International Conference on Image Processing*, 2006, pp. 2929–2932.
- [8] H. R. Sheikh, A. C. Bovic and De Veciana G., An information fidelity criterion for image quality assessment using natural scene statistics, *IEEE Tans. on Image Processing* 14, 2005, pp. 2117–2128.
- [9] H. R. Sheikh and A. C. Bovic, Image information and visual quality, *IEEE Trans. on Image Processing* 15, 2006.
- [10] A. B. Watson, DCTune: A technique for visual optimization of DCT quantization matrices for individual images, *Society for Information Display Digest of Technical PapersXXIV*, 1993, pp. 946–949.
- [11] S. Lee, M. S. Pattichis and A. C. Bovik, Foveated Video Quality Assessment, *IEEE Trans. on Image Multimedia* 4, 2002.
- [12] D. M. Chandler and S. Hemami, VSNR: A wavelet-based visual signal-to-noise ratio for natural images, *IEEE Tran. on Image Processing* 16, 2007.
- [13] C. Li and A. C. Bovik, Content-partitioned structural similarity index for image quality assessment, *Signal Processing: Image Communication* 25, 2010, pp. 517–526.

- [14] E. C. Larson and D. M. Chandler, Most apparent distortion: full-reference image quality assessment and the role of stratege, *Journal of Electronic Imaging* 19, 2010.
- [15] L. Zhang, X. Mou and D. Zhang, FSIM: A feature similarity index for image quality assessment, *IEEE Tran. on Image Processing* 20, 2011.
- [16] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli and F. Battisti, TID2008-A database for evaluation of fullreference visual quality assessment metrics, *Adv. Modern Radioelectron* 10, 2009, pp. 30–45.
- [17] M. P. Sampat, Z. Wang, S. Gupta, A. C. Bovik and M. K. Markey, Complex wavelet structural similarity: A new image similarity index, *IEEE Trans. on Image Processing* 18, 2009, pp. 2385– 2401.
- [18] D. Marr and E. Hildreth, Theory of edge detection, *Proc. R. Soc. Lond. B* 207, 1980, pp. 187– 217.
- [19] P. Kovesi, Image features from phase congruency, *Videre: J. Comp. Vis. Res.* 1, 1999, pp. 1– 26.
- [20] A. M. Eskicioglu and P. S. Fisher, Image quality measures and their performance, *IEEE Trans. On Comm.* 43, 1995, pp. 2959–2965.
- [21] M. P. Eckert and A. P. Bradley, Perceptual quality metrics applied to still image compression, *Signal Processing* 70, 1998, pp. 177–200.
- [22] B. Girod, What's wrong with mean-squared error. Digital Images and Human Vision, MA: MIT Press, A. B. Watson, Ed. Cambridge, 1993.
- [23] A. Liu, W. Lin and M. Narwaria, Image Quality Assessment Based on Gradient Similarity, *IEEE Trans. on Image Processing* 21, 2012.

1a		eter selectio		0
Lamda_3=0.7	Lamda_3=1	Lamda_3=0.6 0.7	Lamda_3=0.2	Lamda_3=0.1
P=118	P=120	P=115	P=103	P=80
S=106	S=106	S=111	S=90	S=69
K=101	K=105	K=110	K=89	K=73
Total=325	Total=331	Total=336	Total=282	Total=222
Lamda_3=0.8 0.9	Lamda_3=0.8 0.9	Lamda_3=0.2 0.3	Lamda_3=0.3	Lamda_3=0.2
P=126	P=124	P=129	P=135	P=132
S=112	S=122	S=128	S=126	S=125
K=112	K=119	K=129	K=128	K=122
Total=350	Total=365	Total=386	Total=389	Total=379
Lamda_3=0.7 1	Lamda_3=0.6 1	Lamda_3=0.5 1	Lamda_3=0.8 1	Lamda_3=1
P=122	P=124	P=122	P=127	P=133
S=118	S=118	S=121	S=127	S=127
K=115	K=116	K=125	K=129	K=129
Total=355	Total=358	Total=368	Total=383	Total=389
Lamda_3=0.8 0.9	Lamda_3=0.1 1	Lamda_3=0.9 1	Lamda_3=1	Lamda_3=1
P=82	P=80	P=79	P=78	P=77
S=114	S=106	S=108	S=108	S=111
K=101	K=103	K=112	K=115	K=114
Total=287	Total=289	Total=299	Total=301	Total=302
Lamda3=0.1 1	Lamda_3=0.1 1	Lamda_3=0.1 1	Lamda_3=0.1 1	Lamda_3=0.1 1
P=60	P=59	P=62	P=61	P=60
S=100	S=102	S=102	S=102	S=103
K=105	K=106	K=109	K=110	K=109
Total=265	Total=267	Total=273	Total=273	Total=272

Table 1: Parameter selection in BDEDG

	VSNR WSNR	0.7728 0.8714	0.7793 0.8220	0.7665 0.8483	0.7295 0.6118	0.8811 0.9129	0.6471 0.8941	0.8270 0.8648	0.9330 0.9326	0.9286 0.9338	0.9174 0.9218	0.9515 0.9566	0.8055 0.7382	0.7909 0.8335	0.5716 0.6891	0.1926 0.2907	0.3715 0.7588	0.4239 0.5722
	VIFP V	0.8005 0.	0.8357 0.	0.8414 0.	0.8429 0.	0.8811 0.	0.7994 0.	0.7903 0.	0.9449 0	0.9185 0.	0.9150 0.	0.9588 0	0.8386 0.	0.8325 0.	0.7701 0.	0.8352 0.	0.5106 0.	0.8075 0.
	VIF	0.8800	0.8785	0.8703	0.8698	0.9075	0.8331	0.7956	0.9546	0.9189	0.9170	0.9713	0.8582	0.8510	0.7608	0.8320	0.5132	0.8190
of SROCC	IJŊIJ	0.5161	0.4584	0.5359	0.7269	0.6722	0.4950	0.5606	0.8836	0.7754	0.7701	0.9116	0.8348	0.6714	0.7398	0.8070	0.5617	0.5201
Table 2: Comparison of results in the sense of SROCC	SNR	0.8338	0.8513	0.8332	0.6995	0.9050	0.8776	0.8205	0.8173	0.9123	0.8054	0.7837	0.7215	0.7952	0.5022	0.4016	0.7603	0.5948
results in	NQM	0.7679	0.7490	0.7720	0.7067	0.9015	0.7616	0.8209	0.8846	0.9450	0.9075	0.9532	0.7373	0.7262	0.6800	0.2348	0.5245	0.6191
iparison of	MISSIM	0.8094	0.8064	0.8195	0.8155	0.8685	0.6868	0.8537	0.9607	0.9571	0.9348	0.9736	0.8736	0.8525	0.7336	0.7617	0.7374	0.6400
ble 2: Con	SSIM	0.8310	0.8134	0.8438	0.7561	0.8919	0.7072	0.8745	0.9596	0.9595	0.9270	0.9723	0.8668	0.8707	0.7168	0.8529	0.7575	0.6329
Ta	PSNR	0.9327	0.9068	0.9229	0.8487	0.9323	0.9177	0.8699	0.8682	0.9381	0.9011	0.8300	0.7665	0.7765	0.5931	0.5852	0.6974	0.6126
	MSE	0.9115	0.9068	0.9229	0.8487	0.9323	0.9177	0.8699	0.8682	0.9381	0.9011	0.8300	0.7665	0.7765	0.5931	0.5852	0.6974	0.6126
	BDEDG	0.9319	0.8911	0.9356	0.8748	0.9114	0.8518	0.8888	0.8961	0.9465	0.9342	0.9373	0.8211	0.8841	0.7108	0.3918	0.6237	0.5044
	distortion	1	2	3	4	5	9	7	8	6	10	11	12	13	14	15	16	17

distortion BDEDG 1 0.9227 2 0.9053 3 0.9359 4 0.8837		MSE	PSNR									
0.92 0.90 0.93 0.93 0.93 0.93 0.93	_			SSIM	MISSIM	NQM	SNR	IJŊIJ	VIF	VIFP	VSNR	WSNR
0.90 0.93 0.88		0.9028	0.9327	0.7669	0.7477	0.7455	0.8097	0.5238	0.8667	0.7774	0.7448	0.8531
0.93		0.9079	0.9219	0.7853	0.7781	0.7345	0.8430	0.4629	0.8953	0.8291	0.7639	0.8116
0.88		0.9154	0.9523	0.7957	0.7602	0.7573	0.8176	0.5440	0.8585	0.8330	0.7500	0.8346
		0.8635	0.8617	0.7308	0.7873	0.7084	0.6815	0.7555	0.8915	0.8493	0.7530	0.5976
0.9491		0.9138	0.9673	0.8208	0.8220	0.9161	0.9071	0.6898	0.9451	0.8753	0.8832	0.9191
0.8483		0.8762	0.9063	0.6320	0.6250	0.7438	0.8583	0.4794	0.8149	0.7836	0.6242	0.8817
0.8746		0.8438	0.8903	0.7906	0.7567	0.8073	0.8049	0.5505	0.7453	0.6802	0.8130	0.8494
0.8618		0.4870	0.8376	0.8782	0.8774	0.8731	0.8118	0.8749	0.9392	0.9416	0.9160	0.9309
0.9431		0.8318	0.9412	0.9135	0.9148	0.9549	0.9066	0.8011	0.8975	0.8948	0.9194	0.9347
0.9327		0.7679	0.8892	0.9300	0.9310	0.9203	0.7892	0.7885	0.9324	0.9188	0.9058	0.9296
0.9475		0.8522	0.8657	0.9516	0.9386	0.9524	0.7794	0.9189	0.9171	0.9412	0.9343	0.9518
0.7823		0.7247	0.7653	0.8276	0.8241	0.7332	0.5925	0.8378	0.8719	0.8537	0.6466	0.7207
0.8805		0.7244	0.7874	0.8310	0.7876	0.7341	0.8036	0.6790	0.8310	0.8180	0.7610	0.8354
0.7199		0.5506	0.5975	0.6608	0.6646	0.6822	0.5107	0.7212	0.7363	0.7538	0.5658	0.6791
0.4219		0.5250	0.5735	0.8720	0.7962	0.2144	0.4112	0.8430	0.8336	0.8452	0.2727	0.2654
0.6027		0.7004	0.6819	0.7267	0.6690	0.5251	0.7307	0.4242	0.5920	0.5945	0.2469	0.7292
0.4891	161	0.5915	0.6043	0.7004	0.7688	0.6283	0.6049	0.5335	0.8832	0.8492	0.4285	0.5946

			T_a	ible 4: Cor	Table 4: Comparison of results in the sense of PLCC	f results in	the sense	of PLCC				
distortion	BDEDG	MSE	PSNR	SSIM	MISSIM	NQM	SNR	IJŊIJ	VIF	VIFP	VSNR	WSNR
1	0.7600	0.7207	0.7207	0.6321	0.6095	0.5626	0.6325	0.3591	0.6792	0.6019	0.5617	0.6754
2	0.7050	0.7289	0.7289	0.6165	0.6045	0.5524	0.6580	0.3097	0.6942	0.6480	0.5844	0.6184
3	0.7589	0.7298	0.7298	0.6464	0.6160	0.5611	0.6254	0.3774	0.6784	0.6440	0.5611	0.6436
4	0.6791	0.6336	0.6336	0.5647	0.6219	0.5189	0.4930	0.5263	0.6923	0.6479	0.5435	0.4464
5	0.6907	0.7465	0.7465	0.6686	0.6377	0.7064	0.7036	0.4555	0.6741	0.6530	0.6607	0.7250
9	0.6493	0.7512	0.7512	0.5072	0.4846	0.5441	0.6913	0.3311	0.6284	0.5856	0.4523	0.7156
7	0.7057	0.6788	0.6788	0.6924	0.6625	0.6251	0.6150	0.3999	0.6479	0.6412	0.6348	0.6684
8	0.7862	0.7328	0.7328	0.8288	0.8271	0.6783	0.6475	0.6758	0.8205	0.7878	0.7733	0.7717
6	0.7905	0.7804	0.7804	0.8332	0.8304	0.7921	0.7387	0.5775	0.7610	0.7633	0.7638	0.7674
10	0.7736	0.7308	0.7308	0.7334	0.7580	0.7142	0.6035	0.5445	0.7195	0.7404	0.7344	0.7300
11	0.7824	0.6382	0.6382	0.8573	0.8656	0.8034	0.5767	0.7415	0.8515	0.8215	0.8042	0.8152
12	0.6212	0.5788	0.5788	0.6669	0.6796	0.5388	0.5210	0.6332	0.6572	0.6345	0.5978	0.5388
13	0.7139	0.5893	0.5893	0.6876	0.6622	0.5294	0.5965	0.4825	0.6593	0.6369	0.5982	0.6470
14	0.5084	0.4191	0.4191	0.5114	0.5238	0.4737	0.3463	0.5446	0.6001	0.6052	0.3989	0.4890
15	0.2902	0.4189	0.4189	0.6584	0.5515	0.1635	0.2837	0.5947	0.6215	0.6171	0.1350	0.2117
16	0.4336	0.5007	0.5007	0.5678	0.5525	0.3705	0.5658	0.4080	0.3523	0.3522	0.2572	0.5638
17	0.3446	0.4340	0.4340	0.4761	0.4830	0.4178	0.4433	0.3246	0.5798	0.5713	0.2787	0.4105