Image quality assessment metrics by using directional projection

Jianxin Pang (庞建新), Rong Zhang (张 荣), Hui Zhang (张 晖), Xuan Huang (黄 轩), and Zhengkai Liu (刘政凯)

MOE-Microsoft Key Laboratory of Multimedia Computing and Communication, Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei 230027

Received August 23, 2007

Objective image quality measure, which is a fundamental and challenging job in image processing, evaluates the image quality consistently with human perception automatically. On the assumption that any image distortion could be modeled as the difference between the directional projection-based maps of reference and distortion images, we propose a new objective quality assessment method based on directional projection for full reference model. Experimental results show that the proposed metrics are well consistent with the subjective quality score.

OCIS codes: 110.3000, 110.2960, 330.5020.

doi: 10.3788/COL20080607.0491.

In the procedures of image and video processing systems, e.g., acquisition, processing, coding, storage, transmission and reproduction, digital image and video may be in degradation in visual quality. The purpose of the research of image quality assessment (QA) is to develop strategies and algorithms to accurately evaluate the quality consistently with subjective perception, which is a challenging and fundamental job. Image QA is also with many interests in many applications such as dynamically monitoring and adjusting image quality, optimizing algorithms and parameter setting of image processing systems, and benchmarking image processing systems and algorithms[1].

In the past three decades, many objective image QA methods have been put forward[1,2]. Among them, mathematically defined metrics are the simplest and most widely used at present, for instance, the mean square error (MSE), the root MSE (RMSE), the signal-to-noise ratio (SNR), and the peak SNR (PSNR). However, these mathematically defined metrics cannot completely agree with human perception[1,2]. Recently, using structural distortion to measure the image quality is another appropriate candidate. On the assumption that human visual perception is highly adaptive for extracting structural information from a scene, Wang *et al.* presented a mean structural similarity metric (MSSIM) to compare the structural similarity between the reference and the distortion images[1]. And MSSIM has been widely used, for example, for image fusion quality[3]. Shnayderman *et al.* proposed an idea of evaluating the images by computing the distance between the singular values (SVD), which are decomposed from the reference and distortion images individually[4]. Both of these metrics are potential to replace the role of those mathematically defined metrics.

We propose a universal image quality measure, in which the degradation of image quality is modeled as the differences between the directional projection (DP) based maps of reference and distortion images. The DP-based maps are built by using Radon transform[5,6]. Here "universal" means that it does not depend on testing images, testing environment, and the observers individually, and it also should achieve the agreement with the human perception. The proposed approach also tries to be developed with high efficiency and low computational complexity.

The Radon transform can be defined by[6]

$$\Re(s,\theta)\left[\mathbf{f}(x,y)\right] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathbf{f}(x,y)\delta(s - x\cos\theta - y\sin\theta)\mathrm{d}x\mathrm{d}y, \quad (1)$$

where $\mathbf{f}(x,y)$ is a two-dimensional (2D) vector, s is the perpendicular distance from a line to the origin, and θ is the angle formed by the distance vector. Radon transform could be regarded as the projection procedure on different directions, therefore, we name this kind of method the "directional projection". And $\Re(s,\theta)\left[\mathbf{f}(x,y)\right]$ is the vector of the DP-based map.

The reference image is divided into small blocks with the block size $m \times l$, the nth block is defined as the vector $\mathbf{B}_n \in \mathbf{R}^{m\times l}$, and $\mathbf{D}_n \in \mathbf{R}^{m\times l}$ is defined as the vector of the DP-based map. $\mathbf{b}_n \in \mathbf{R}^{m\times l}$ and $\mathbf{d}_n \in \mathbf{R}^{m\times l}$ are defined for the counterparts of the distortion image. $\mathbf{D}_n$ and $\mathbf{d}_n$ are calculated as

$$\mathbf{D}_n = \Re(s,\theta)\left[\mathbf{B}_n\right], \quad (2)$$

$$\mathbf{d}_n = \Re(s,\theta)\left[\mathbf{b}_n\right]. \quad (3)$$

Define the local distortion intensity as SD_n:

$$\mathrm{SD}_n = \left\|\mathbf{D}_n - \mathbf{d}_n\right\|, \quad (4)$$

where $\|\cdot\|$ represents the procedure of calculating the vector norm.

The global distortion intensity SD is simply calculated as the mean of the local distortion intensities:

$$\mathrm{SD} = \mathrm{mean}(\mathrm{SD}_n). \quad (5)$$

Here we carefully propose that our predictive score of objective quality is a logarithmic function of the distortion intensity which obeys the Weber-Fechner law (a constant relative difference in the intensity corresponds to a constant absolute difference in the logarithm of the intensity). Therefore, the image quality measure by using DP is defined as

$$\mathrm{DP} = \log(\mathrm{SD}), \tag{6}$$

and it is clear that $\mathrm{SD} > 0$.

Radon transform is widely used in image processing and analysis[6−8]. In this letter, we use Radon transform to build the DP-based maps for image QA. And the DP-based maps are expected to represent the images' directional characteristics which the pixels, edges and shapes contribute a lot to. The differences of the maps, modeling any image distortion, are desired to represent the variation of the images' degradation. The actual value is meaningless, but the comparison between two values for different distorted images gives a measure of quality. The lower the predicted score of DP is, the better the image quality is. When the distortion and the reference images are identical, $\mathrm{SD} = 0$.

The database used in our experiments is the well-known "LIVE Image Quality Assessment Database Release 2"[9], and the database is composed of color nature images. The subjective score of the images difference mean opinion score (DMOS), comes from the latest database[10]. Some images in the database are randomly selected in Fig. 1 as an example.

The database includes 29 reference color images, each of which contains 5 distortion types (totally 799 images): fast fading Rayleigh (FF, 145 images), Gaussian blur (GBlur, 145 images), white noise (WN, 145 images), JPEG (175 images), and JPEG2000 (169 images). The five distortion types could often happen in practical applications. FF is a simulation of transmission errors in compressed JPEG2000 bit stream by using a fast fading Rayleigh channel model. The RGB (red, green, and blue) components are blurred by using a circular-symmetric 2D Gaussian kernel in GBlur distortion. WN distorts the images by adding white Gaussian noise to RGB components. JPEG and JPEG2000 compress the images at different bit rates, which could often happen in image and video processing applications. We evaluate the performances following the procedures in the Video Quality Experts Group (VQEG) Phase I FR-TV test[11]. And the simple and widely used metrics PSNR, MSSIM[1], and SVD[4] are selected to make a comparison with our metrics.

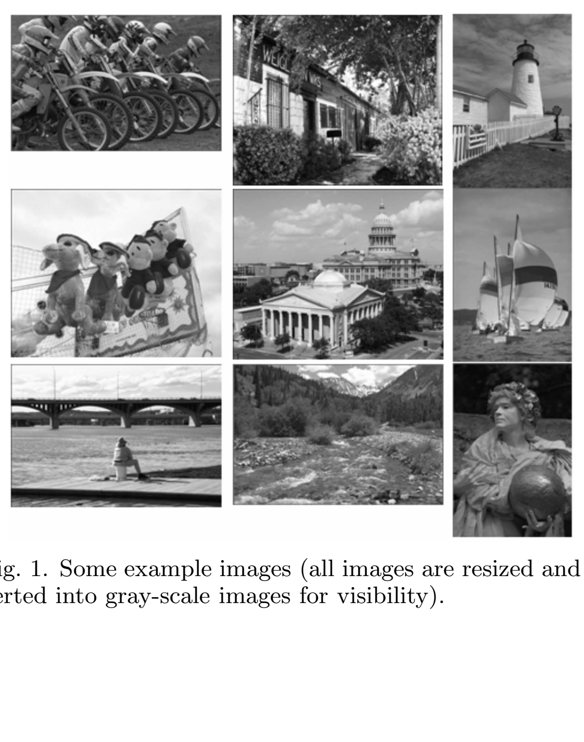

Fig. 1. Some example images (all images are resized and converted into gray-scale images for visibility).

In our experiments, we chose the block size $m \times l = 8 \times 8$, just because it is a common size in many image processing applications and both SVD and MSSIM use this window size. The experiments worked with the luminance of the images. We converted color images into gray-scale ones by separating the luminance information from the color information.

Particularly, we compare three DPs with different parameter θ. DP1 and DP2 are desired to achieve high efficiency and low computational complexity:

DP: $\theta = 0° : 1° : 179°$,
DP1: $\theta = 0° : 45° : 179°$,
DP2: $\theta = 0° : 30° : 179°$.

Figures 2 and 3 show the relationship between the predictive scores of above-mentioned assessment metrics and the DMOS. The lines in figures are nonlinear fitting curves used for regression or fitting for those methods. The logistic function is with five variables as follows:

$$\mathrm{logistic}(x) = a_1 + \frac{a_2 - a_3}{1 + \exp(\frac{x - a_4}{a_5})}. \tag{7}$$

Figure 2 compares the performances among each of the four metrics and DMOS in cross-type distortions. Figure 3 shows the results for JPEG and JPEG2000 images, comparing the performances of cross image coding types. Tables 1 and 2 compare the Pearson correlation-coefficient (PCC) and the Spearman rank order correlation-coefficient (SROCC) among each of the four metrics and DMOS. Table 3 compares RMSE among each of the four metrics and DMOS.

From the above results, we come to a conclusion that PSNR is not well adaptable in all of the distortion types except WN. Meanwhile, it is reasonable that PSNR has a satisfactory performance in WN, because the information of the WN-distortion image is distorted only by WN, and thus PSNR can count these errors more accurately which are statistically independent. When the "errors" or characters which distort the images are not uncorrelated, PSNR cannot work well simply and accurately. In this case, the other metrics try to overcome the systematic drawbacks of PSNR.

From Figs. 2 and 3, we can see that SVD and MSSIM do not perform accurately and sound enough. When the image quality is worse, for example when DMOS is more than 60, SVD and MSSIM show poor performance. This means that they are not well adaptive to low-quality images. We can also see that DP shows the stability for all images.

SVD and MSSIM have close performances, while MSSIM shows the best of all in FF. In individual distortion types, DP outperforms the others in GBlur, JPEG and JPEG2000, and it also has a good performance in FF and WN. DP has the best performance in cross-distortion types, especially in coding types.

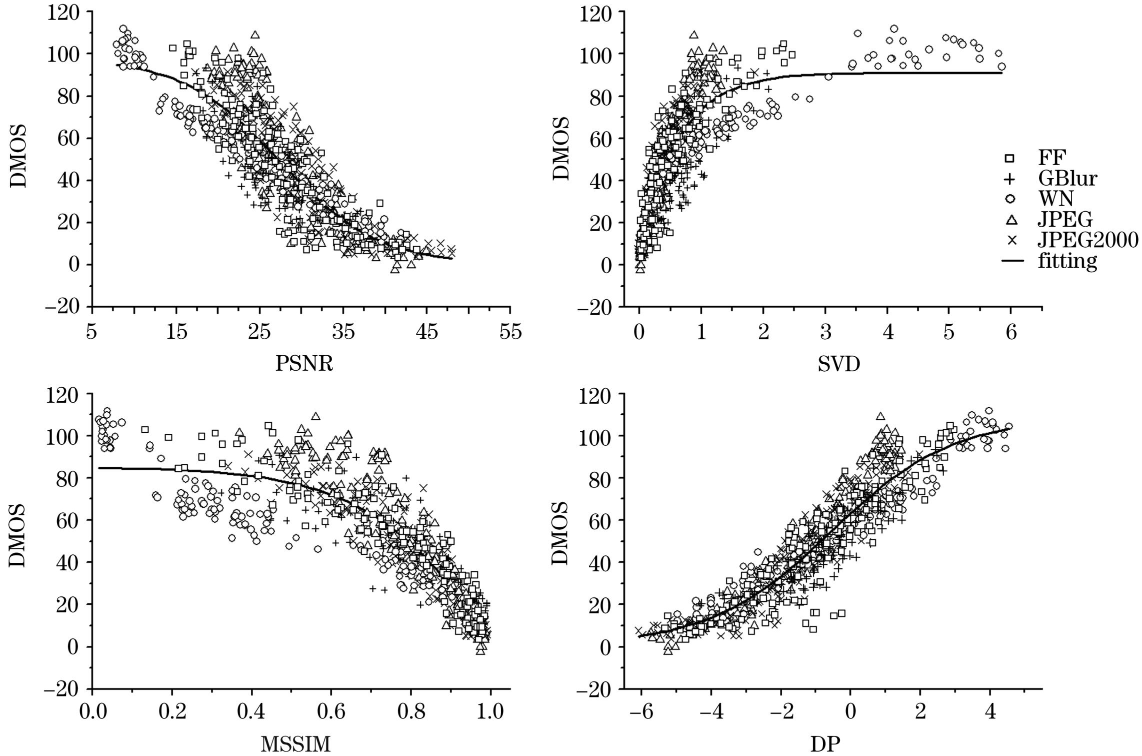

Fig. 2. Scatter plots for PSNR, SVD, MSSIM, and DP for the five types of distorted images.

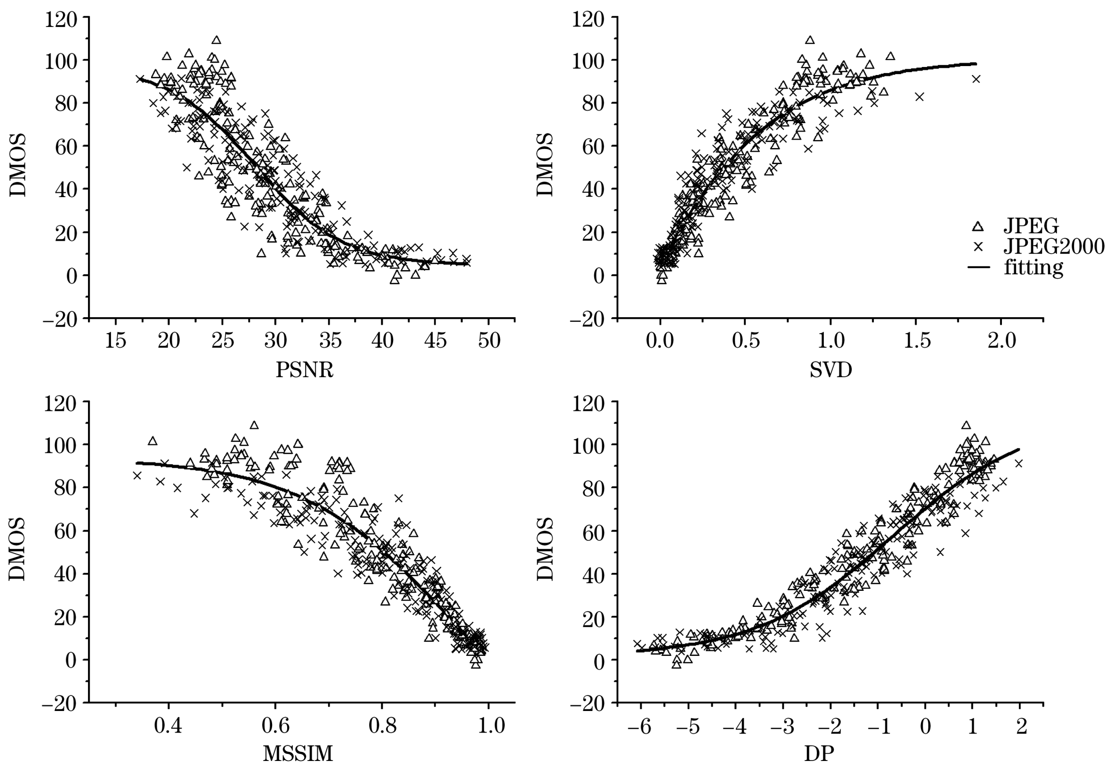

Fig. 3. Scatter plots for PSNR, SVD, MSSIM, and DP for JPEG and JPEG2000 images.

All of the four metrics only work with the luminance of the images. However, subjective DMOSs are gained by the observers evaluating the color images. Therefore, when the distortion of color information cannot be detected in luminance channel, it is very difficult to assess those images exactly only by using luminance information. In this database, FF is the distortion which sometimes degrades the color information, so the plots of four metrics scatter in FF images. It is not easy to study color distortion for image quality. Moreover, the sensitivity of DP to slight distortions in rotation, shift, and magnification is not satisfactory, which is to be taken into future research.

Table 1. PCC Comparison for PSNR, SVD, MSSIM, and PD within Individual Distortion and Cross Types

	All	FF	GBlur	WN	JPEG	JPEG2000	JPEG+JPEG2000
PSNR	0.8693	0.8936	0.7734	0.9844	0.8865	0.8980	0.8863
SVD	0.8822	0.8985	0.7220	0.9786	0.9589	0.9428	0.9466
MSSIM	0.8984	0.9422	0.8465	0.9699	0.9482	0.9407	0.9377
DP	0.9285	0.9029	0.9132	0.9803	0.9734	0.9467	0.9585
DP1	0.9275	0.9036	0.9097	0.9803	0.9731	0.9462	0.9581
DP2	0.9283	0.9033	0.9119	0.9804	0.9731	0.9465	0.9583

Table 2. SROCC Comparison for PSNR, SVD, MSSIM, and PD within Individual Distortion and Cross Types

	All	FF	GBlur	WN	JPEG	JPEG2000	JPEG+JPEG2000
PSNR	0.8744	0.8939	0.7709	0.9831	0.8798	0.8931	0.8890
SVD	0.8872	0.8989	0.7055	0.9835	0.9492	0.9404	0.9474
MSSIM	0.9075	0.9394	0.8595	0.9645	0.9432	0.9357	0.9403
DP	0.9312	0.9018	0.9070	0.9770	0.9602	0.9416	0.9575
DP1	0.9300	0.9014	0.9024	0.9767	0.9600	0.9416	0.9573
DP2	0.9309	0.9013	0.9052	0.9770	0.9603	0.9418	0.9576

Table 3. RMSE Comparison for PSNR, SVD, MSSIM, and PD within Individual Distortion and Cross Types

	All	FF	GBlur	WN	JPEG	JPEG2000	JPEG+JPEG2000
PSNR	13.5029	12.7859	11.7088	4.9192	14.7411	11.1016	13.4290
SVD	12.8636	12.5051	12.7795	5.7595	9.0365	8.4114	9.3492
MSSIM	12.0018	9.5481	9.8349	6.8167	10.1208	8.5588	10.0742
DP	10.1909	12.2479	7.5269	5.5200	7.2996	8.1296	8.4142
DP1	10.2789	12.2046	7.6691	5.5200	7.3427	8.1628	8.4495
DP2	10.2109	12.2211	7.5810	5.5170	7.3366	8.1391	8.4317

DP has a more computational complexity compared with SVD and MSSIM, but the computational complexity of DP1 and DP2 with different parameters has been significantly reduced. The implementation of DP1 and DP2 on a 768×512 image on a Pentium IV, 3.0 GHz laptop using the luminance information takes about 0.2 s, while SVD takes about 1 s and MSSIM takes about 0.1 s in our experiments. Besides, the typical DP values range between -6 and 4 in our implementations.

On the assumption that any image distortion can be modeled as the difference of the DP-based maps, we proposed an objective image measure based on DP by using Radon transform. Besides, we discussed the relationship between the distortion intensity and the subjective visual quality. The experimental results have shown that DP performs better than PSNR, SVD, and MSSIM. This metric is well adaptable not only in individual distortion type, especially in image coding types, but also in cross-distortion types.

There are numerous distortion types for images in practice and we only deal with coding types. Our future work is to explore into more aspects and investigate into the relationship between distortion of structural information and the subjective visual quality, and we will also focus on the research of extending the proposed metric to color images and video sequences.

J. Pang's e-mail address is waltonpang@ustc.edu.

References

1. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, IEEE Trans. Image Process. **13,** 600 (2004).
2. A. M. Eskicioglu and P. S. Fisher, IEEE Trans. Commun. **43,** 2959 (1995).
3. H. Di and X. Liu, Acta Photon. Sin. (in Chinese) **35,** 766 (2006).
4. A. Shnayderman, A. Gusev, and A. M. Eskicioglu, IEEE Trans. Image Process. **15,** 422 (2006).
5. Z. An, S. Zhao, X. Wang, and L. Zhou, Acta Photon. Sin. (in Chinese) **36,** 1176 (2007).
6. S. Xiao and Y. Wu, Chin. Opt. Lett. **5,** 513 (2007).
7. L. Wang, Q. Chang, K. Zhang, and Y. Li, Infrared Laser Eng. (in Chinese) **32,** 163 (2003).
8. Z. Deng and Y. Xiong, Opto-Electron. Eng. (in Chinese) **34,** (10) 98 (2007).
9. H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, "Live image quality assessment database release 2" http://live.ece.utexas.edu/research/quality (Sep. 20, 2006).
10. H. R. Sheikh, M. F. Sabir, and A. C. Bovik, IEEE Trans. Image Process. **15,** 3441 (2006).
11. VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment" http://www.vqeg.org/ (Sep. 21, 2006).