A Projection-Based Image Quality Measure

Jianxin Pang, Rong Zhang, Lu Lu, Jinhui Tang, Zhengkai Liu

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

Received 25 October 2007; revised 31 March 2008; accepted 27 May 2008

ABSTRACT: Objective image quality measure, evaluating the image quality consistently with human perception automatically, could be employed in image and video retrieval. And the measure with high efficiency and low computational complexity plays an important role in numerous image and video processing applications. On the assumption that any image's distortion could be modeled as the difference between the projection-based values (PV) of reference image and the counterpart of distorted image, we propose a new objective quality assessment method based on signal projection for full reference model. The proposed metric is developed by simple parameters to achieve high efficiency and low computational complexity. Experimental results show that the proposed method is well consistent with the subjective quality score. Int J Imaging Syst Technol, 18, 94–100, 2008; Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/ima.20156

Key words: image quality assessment; signal projection; inner product; objective quality

I. INTRODUCTION

In the field of image processing, image quality assessment is a challenging and fundamental task. And it is with many interests in a variety of applications in the procedure of image and video processing system, e.g., acquisition, processing, coding, storage, transmission, and reproduction (Wang et al., 2004), for example, in image and video retrieval most users prefer to receive high-quality images and videos rather than low-quality ones. The image similarity measurement could be a potential candidate for content-based image retrieval. And image quality measurement could also be used in near-duplicate image detection which is an important application of content-based image retrieval. In many cases, quality measure methods with high efficiency and low computational cost are desirable, especially for some real-time or high-performance applications, e.g., image and video retrieval (Tao et al., 2006a,b, 2007, which is worth investigating.

Image quality assessment can be classified as *subjective* and *objective* ones (Eskicioglu and Fisher, 1993). As we all know, the best way to assess an image or a picture is by human beings, because all of the images and pictures are used and viewed just by human beings. The mean option score (MOS) and the difference mean option score (DMOS) are two kinds of subjective assessment which are widely used in the image and video quality assessment. But in practice, it is usually expensive, time-consuming, inconvenient, and environment-limited. Moreover the subjective assessment method may be affected by various factors, such as the mood of the candidate, the testing equipment, the individuality of the candidate, and so on. So it is important to develop an objective image quality metric that can automatically and exactly value the image quality.

This article will focus on the full-reference image QA, which means that the original image is fully known as the reference one, and we take it for granted that the original image is 'perfect' or of 'high quality' and is used as the reference one.

In the past three decades, many objective image quality assessment methods have been put forward (Eskicioglu and Fisher, 1993; Wang et al., 2004).

Among them, mathematically defined metrics are simple and widely used at present, for instance the mean square error (MSE), the root mean square error (RMSE), the signal to noise ratio (SNR), and the peak signal to noise ratio (PSNR). However, all these mathematically defined metrics cannot completely meet the characteristic of human's perception (Teo Heeger, 1994; Wang etal., 2002; Wang et al., 2004; Tao et al., 2006a), although they are independent to the images and easy to calculate. The quality measurement based on human visual systems (HVS) is put forward by Mannos and Sakrison in a famous article (Mannos Sakrison, 1974) in 1974. Some other people also contribute a lot in this field (Watson, 1993; Chou and Li, 1995; Karunasekera and Kingsbury, 1995; Mayache, 1998). Although the image quality metric based on the psychophysical measurement of HVS is mostly accepted, the complexity of the HVS and the finitude of the cognizing of the human beings still keeps this metric from going much further (Wang et al., 2004).

Recently, using structural distortion to measure the image quality is another appropriate candidate for image quality assessment. The most known metric SSIM (Structural Similarity) is brought forward by Wang et al. (2004). On the assumption that "human visual perception is highly adaptive for extracting structural information from a scene," they propose the "SSIM" metric which compares the structural similarity between the original image and the

Correspondence to: Jianxin Pang; e-mail: waltonpang@ustc.edu

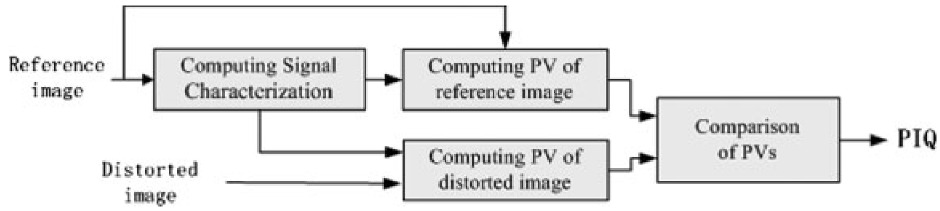

Figure 1. The flowchart of PIQ.

distorted one. The arithmetic of Mean Structural Similarity (MSSIM) is that: the image is divided into small blocks, and the distortion of each block is computed using the information of luminance, contrast and structure, and the final quality score of the image is the mean value of the blocks'. In the work (Shnayderman et al., 2006) Shnayderman et al. propose 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. These metrics above are of simple parameters and low computational complexity, which are potential to replace the role of those mathematically defined metrics. Some new measures are brought out recently (Beghdadi and Popescu, 2003; Zheng et al. 2003; Wang et al., 2006; Venakata et al., 2007; Lu, accepted for publication).

Generally, the developed image quality measure evaluates the quality of images to achieve the agreement with the human perception, also it would be adapted to both individual and cross distortion types, i.e., it is universal and does not depend on testing images, testing environment and the observers individually. Moreover, the image quality measure with high efficiency and low computational complexity are also desired for some image and video processing applications. Although mathematically defined metrics have those properties, they do not have a satisfactory performance to be consistent with human perception. In our previous work (Pang et al., 2007), a quality metric based on matching pursuit is proposed by extracting the important image structure and developing a set of structural characteristics, however, this metric is of high computational complexity and not adaptive to real-time applications. This work is to develop a more universal measure with high efficiency and low computational complexity by using simple parameters and our approach expect to work as an expansion for those mathematically defined ones.

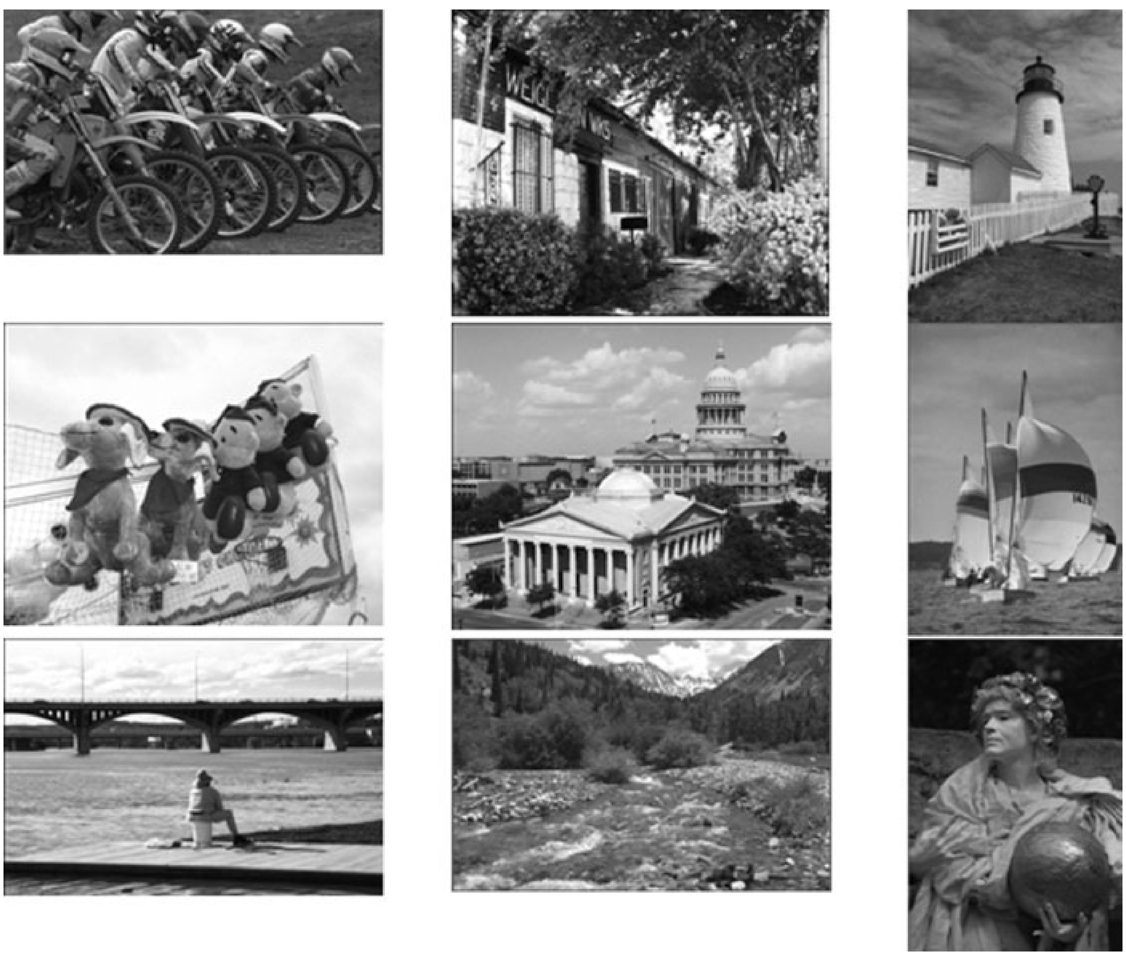

Figure 2. Some example images from the database (all images are resized and converted into grayscale image for visibility).

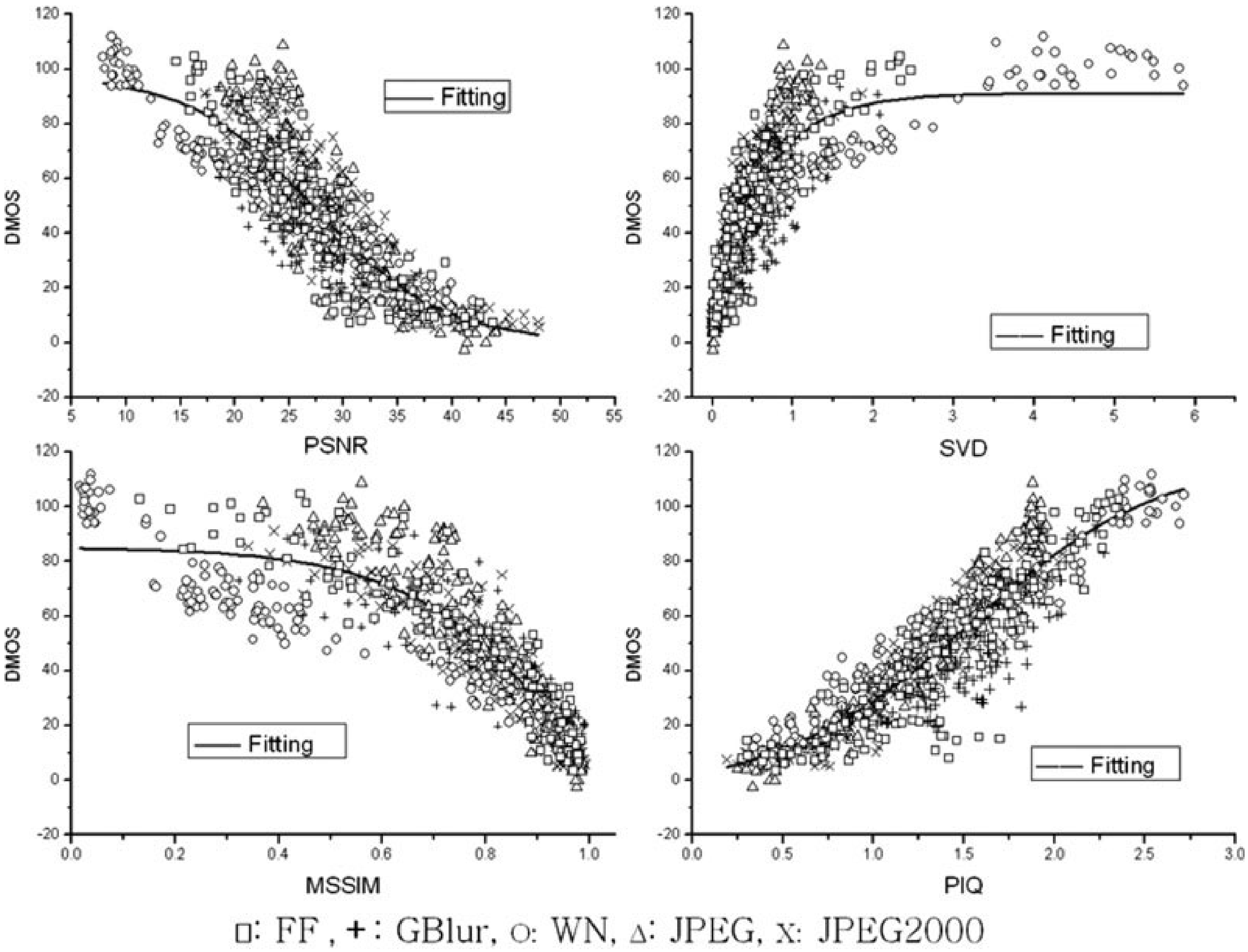

Figure 3. Scatter plots for each of PSNR, SVD, MSSIM, and PIQ for the distortion of five types

In this article, on the assumption that any image's distortion could be modeled as the differences between the projection-based values (PV) of reference image and the counterpart of distorted image, we develop a new image quality assessment metric based on signal projection (PIQ, the Projection-based Image Quality measure), which is to achieve high efficiency and low computational complexity.

The article is organized as follows: Section II introduces the metric of PIQ; In Section III our experimental results are compared with some other metrics. Section IV concludes.

II. THE PROPOSED MEASUREMENT

A. Projection-Based Image Quality Measure: PIQ. The reference image is divided into K blocks with the block size $m \times l$, and the nth block is defined as the vector $\overline{\mathbf{B}_n} \in \mathbf{R}^{m\times l}$. $\overline{\mathbf{b}_n} \in \mathbf{R}^{m\times l}$ is defined for the counterparts of the distortion image. The signal characterization of $\overline{\mathbf{B}_n}$ is defined:

$$\overline{\mathbf{C}_n} = \frac{\overline{\mathbf{B}_n}}{\|\overline{\mathbf{B}_n}\|}, \overline{\mathbf{C}_n} \in \mathbf{R}^{m\times l} \tag{1}$$

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

Let "$\langle\ \rangle$" be the projection operator (also called inner product or mapping). In this work, the projection-based value (PV) is defined as the inner product value between two signals, and we introduce PV into our method. Let $\mathbf{E}_n$ be the PV of $\overline{\mathbf{B}_n}$, and $\mathbf{e}_n$ is the PV of $\overline{\mathbf{b}_n}$. Then $\mathbf{E}_n$ and $\mathbf{e}_n$ can be calculated by projecting $\overline{\mathbf{B}_n}$ and $\overline{\mathbf{b}_n}$ onto $\overline{\mathbf{C}_n}$ individually:

$$\mathbf{E}_n = \langle \overline{\mathbf{B}_n}, \overline{\mathbf{C}_n} \rangle \tag{2}$$

$$\mathbf{e}_n = \langle \overline{\mathbf{B}_n}, \overline{\mathbf{C}_n} \rangle \tag{3}$$

Define the structural distortion intensity:

$$\mathbf{SD} = Sqrt\left[\frac{1}{K} * \left(\sum_{n=1}^{k} (\mathbf{E}_n - \mathbf{e}_n)^2\right)\right]. \tag{4}$$

Here, we carefully propose that our predictive score of objective quality is a logarithmic function of the distortion intensity which should obey the Weber-Fechner law (Levine, 2007) (a constant relative difference in the intensity corresponds to a constant absolute difference in the logarithm of the intensity). Therefore, the predictive quality score of PIQ is defined as:

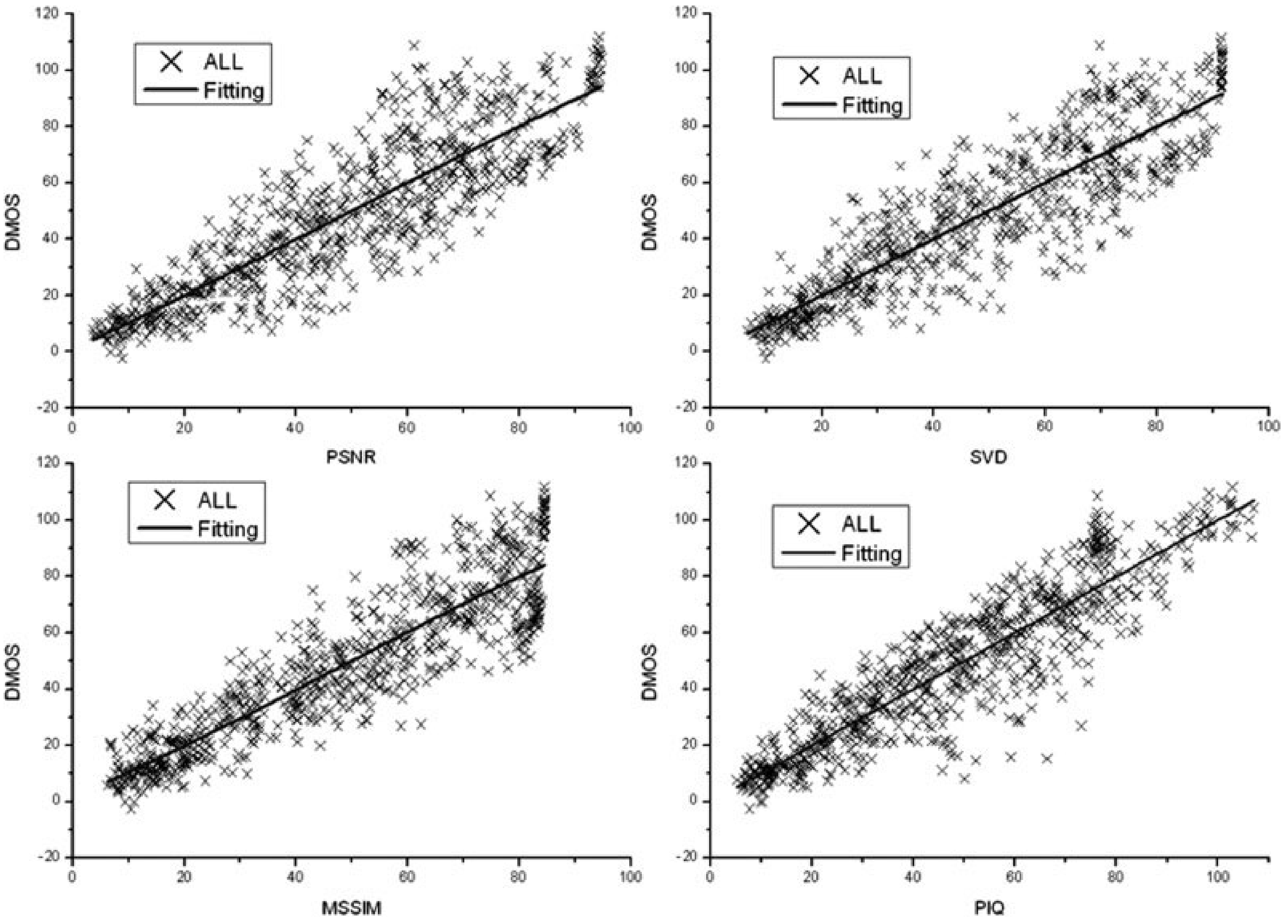

Figure 4. Scatter plots for each of PSNR, SVD, MSSIM, and PIQ for the distortion of five types after a nonlinear fitting.

$$\mathbf{PIQ} = \log(\mathbf{SD}) \tag{5}$$

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

The proposed PIQ models any distortion as the difference of the PVs; and it measure the difference of PVs between the reference and distortion images by the simple projection processing. Here the signal characterization $\overline{\mathbf{C}_n}$ could describe the shape, edge, texture and others, and it works as a structural subspace (Tao et al., 2008). The test images are projected on the same structural subspace. The PVs, which are gained by projecting the distorted images onto $\overline{\mathbf{C}_n}$, are desired to represent the signal contribution along $\overline{\mathbf{C}_n}$, and the variation of the images' degradation will be reflected by their PVs. Therefore we compare the differences of PVs to measure the distortion magnitude. The actual value is meaningless, but the comparison between two values for different distorted images gives one measure of quality. The lower the predicted score of PIQ is, the better the image quality is. When the distortion and the reference images are identical, $\mathbf{SD} = 0$.

B. Implementation With Simple Method of Inner Product. An inner product is a projection or mapping that associates a scalar with every pair of elements in a vector space, which satisfies the properties of symmetry, linearity, and positive definiteness (Taylor and Lay, 1980). Therefore it is reasonable that in any given vector space a number of inner product can be defined. In this article, we select one simple method of inner product as an appropriate implementation for our image quality measure.

Let $\vec{\mathbf{x}}, \vec{\mathbf{y}} \in \mathbf{R}^l$ be two vectors of the same dimension. The chosen method of inner product between them is

$$\left\langle \vec{\mathbf{x}}, \vec{\mathbf{y}} \right\rangle = \left\langle \vec{\mathbf{y}}, \vec{\mathbf{x}} \right\rangle = \vec{\mathbf{x}}^{\mathbf{T}} \vec{\mathbf{y}} = \vec{\mathbf{y}}^{\mathbf{T}} \vec{\mathbf{x}} \tag{6}$$

Then Eq. (4) can be rewritten as

$$\mathbf{SD} = Sprt\left[\frac{1}{K} * \sum_{n=1}^{k} \left(\frac{\overline{\mathbf{B}_n}^{\mathbf{T}} (\overline{\mathbf{B}_n} - \bar{\mathbf{b}}_n)}{\|\mathbf{B}_n\|} \right)^2 \right] \tag{7}$$

The proposed metric is similar with SNR, and PIQ could work as an expansion for SNR.

The flowchart of the PIQ is shown in Figure 1.

III. EXPERIMENTAL RESULTS AND DISCUSSION

The database we use in our experiments is the known "LIVE Image Quality Assessment Database Release 2" (Sheikh, 2005), and the database is composed of color nature images. The subjective score of the images (DMOS, Difference Mean Opinion Score) comes from the latest database (Sheikh and Bovik, 2006; Sheikh et al., 2006). Some images in the database are randomly selected in Figure 2 as an example.

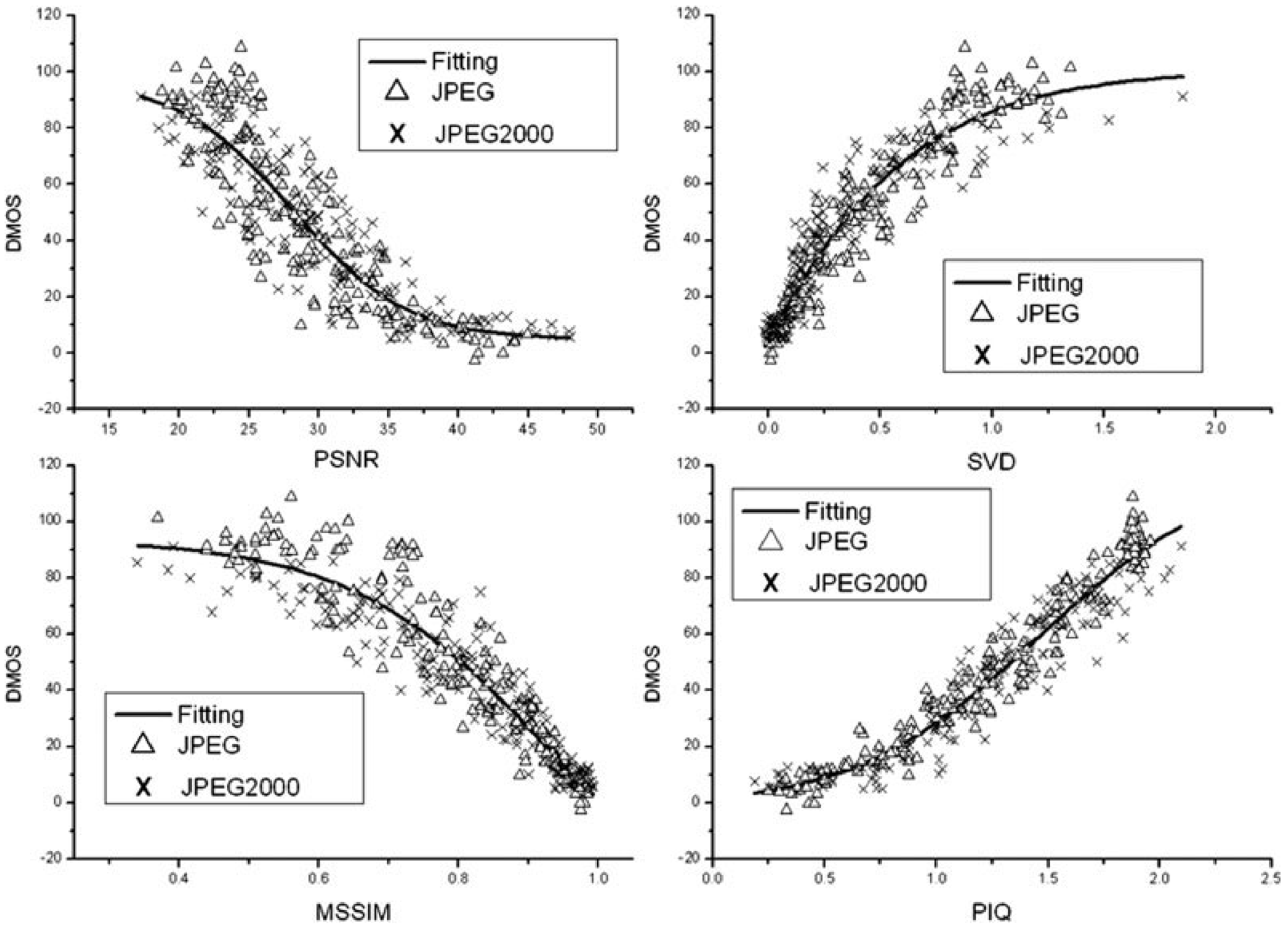

Figure 5. Scatter plots for each of PSNR, SVD, MSSIM, and PIQ for JPEG and JPEG2000.

The database includes 29 reference color images, each of which contains five distortion types (total 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, which could often take place in practical applications, are introduced into studies in this work. FF is a simulation of transmission errors in compressed JPEG2000 bit stream using a Fast fading Rayleigh channel model. The RGB components are blurred using a circular-symmetric 2-D 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, and these kinds of distortion 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 (VQEG, 2000). The widely used and simple metric PSNR, and other two, MSSIM (Wang et al., 2004) and SVD (Shnayderman et al., 2006) are selected to compare with our metric.

In our experiment we choose the block size as $m \times l = 8 \times 8$, just because it is a common size in many image/video processing applications, and both SVD and MSSIM use this window size. The experiments work with the luminance of the images. We convert color images into grayscale ones by separating the luminance information from the color information. SVD and MSSIM also work with the luminance.

A. Experimental Results. In Figures 3–Figure 5, the *X*-axis is the predictive score of each assessment metric and the *Y*-axis is DMOS. The lines in Figures are nonlinear fitting curves which are used for regression or fitting for each of those methods, and the logistic function is with five variables as follows:

$$\mathrm{logistic}(x) = a_1 + \frac{a_2 - a_3}{1 + \exp(\frac{x - a_4}{a_5})} \tag{8}$$

Figure 3 is the result for all images comparing the performances between each of the four metrics and DMOS in cross-type

Table I. PCC-based for each of PSNR, SVD, MSSIM, and PIQ.

	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
PIQ	0.9090	0.8982	0.8536	0.9699	0.9796	0.9555	0.9659

Table II. SORCC-based for each of PSNR, SVD, MSSIM, and PIQ.

	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
PIQ	0.9088	0.8962	0.8340	0.9658	0.9668	0.9508	0.9654

Table III. RMSE-based for each of PSNR, SVD, MSSIM, and PIQ.

	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
PIQ	10.8703	9.9735	9.6224	6.8107	6.4021	7.4449	7.5081

distortions. Figure 4 is for all images after the nonlinear fitting. And Figure 5 is the result for JPEG and JPEG2000 which compare the performances of cross image coding types.

Tables I and II compare the Spearman rank order correlation-coefficient (SROCC) and the Pearson correlation-coefficient (PCC) between each of the four metrics and DMOS. Table III compares RMSE between each of the four metrics and DMOS.

B. Discussion. From experimental figures, we could 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 the best performance in WN, as the structural information of the WN-distortion image is distorted only by WN which are statistically independent, thus PSNR can count these errors or characters more accurately, which distort the images. When the errors or characters are not uncorrelated, PSNR cannot perform well enough simply and accurately. In this case, the other metrics try to overcome some systematic drawbacks of PSNR.

From the top of the *Y*-axis in Figures 3 and 4, we can see that SVD and MSSIM perform not 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 they are not well adaptive to bad quality images and their sensitivity to bad quality images is not satisfactory either. While, we can also see that PIQ shows the stability for all images.

SVD and MSSIM have close performances which are shown in Tables I–III, and MSSIM shows the best of all in FF. In individual distortion type, PIQ performs better than the others in JPEG and JPEG2000, and it also has a good performance in FF, GBlur and WN. PIQ has the best performance in cross-distortion types either, especially in cross coding types.

There are some more issues which are worth investigating. In Eq. (1), the normalization processing may be optimized for some special applications by using some thresholds such as the brightness, contrast and energy distribution. All of the four metrics only work with the luminance of the images. However subjective score DMOS are gained by the observers who evaluating the color images. Therefore, if the distortion of color information happens, and the distortion cannot be detected in luminance channel, it is much difficult to assess those images exactly by only 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. It will not be an easy job to study color distortion for image quality. Moreover, the sensitivity of PIQ to slight distortions in rotation, shift and magnification is not satisfactory, which is to be taken into future study.

PIQ has a less computational complexity compared with SVD and MSSIM. The implementation on a 768×512 image on a Pentium IV 3.0 GHz laptop by using the luminance information takes about 0.1 s. Besides, the typical PIQ values range between 0 and 3 in this implementation.

IV. CONCLUSION AND FUTURE WORK

On the assumption that any image distortion can be modeled as the difference of the PVs, we propose an objective image measure based on signal projection. And the proposed implementation with simple parameters achieves high efficiency and low computational complexity. Besides, we attempt to discuss the relationship between the distortion intensity and the subjective visual quality. The experimental results show that PIQ performs better than PSNR, SVD, and MSSIM. This metric is well adaptable not only in individual distortion type, especially image coding types, but also in cross-distortion types.

There are numerous distortion types for images in practice and only five types of them are introduced in this paper. Thus our future work is to explore into more aspects about the relationship between the image distortion and the subjective visual quality, and we will also focus on the research of extending the proposed metric to color images and video sequences, especially for image and video retrieval.

ACKNOWLEDGMENTS

The authors thank the editors and all anonymous reviewers for their constructive comments on this article.

REFERENCES

A. Beghdadi and B.P. Popescu, A new image distortion measure based on wavelet decomposition, Proc IEEE Int Symp Signal Process Appl (2003), 485–488.

C.H. Chou and Y.C. Li, A perceptually Tuned subband image coder based on the measure of just-noticeable-distortion profile, IEEE Trans Circuits Systems Video Technol 5 (1995), 467–476.

A.M. Eskicioglu and P.S. Fisher, A survey of image quality measures for gray scale image compression, Proceedings of Space and Earth Science Data Compression Workshop, Washington, pp. 49–61, Apr 1993.

S.A. Karunasekera and N.G. Kingsbury, A distortion measure for blocking artifacts in images based on human visual sensitivity, IEEE Trans Image Process 4 (1995), 713–724.

M.W. Levine, Fundamentals of sensation and perception, 3rd ed., Oxford University Press, New York, 2000.

W. Lu, X. Gao, D. Tao, and X. Li, A Wavelet-based image quality assessment method (accepted for publication in Int J Wavelets, Multiresolution Inform Process).

J.L. Mannos and D.J. Sakrison, The effects of a visual fidelity criterion on the encoding of images, IEEE Trans Inform Theory 20 (1974), 525–536.

A. Mayache, T. Eude, and H. Cherifi, A comparison of image quality models and metrics based on human visual sensitivity, Proc IEEE Int Conf Image Process 3 (1998), 409–413.

J. Pang, R. Zhang, L. Lu, and Z. Liu, Quality assessment for image coding based on matching pursuit, Proc IEEE Int Conf Multimedia Expo (2007), Beijing, China, 296–299.

H.R. Sheikh and A.C. Bovik, Image information and visual quality, IEEE Trans Image Process 15 (2006), 430–444.

H.R. Sheikh, M.F. Sabir, and A.C. Bovik, A statistical evaluation of recent full reference image quality assessment algorithms, IEEE Trans Image Process 15 (2006), 3440–3451.

H.R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, LIVE Image Quality Assessment Database Release 2, Available at http://live.ece.utexas.edu/research/quality, 2005.

A. Shnayderman, A. Gusev, and A.M. Eskicioglu, An SVD-based grayscale image quality measure for local and global assessment, IEEE Trans Image Process 15 (2006), 422–429.

D. Tao, X. Li, and S. J. Maybank, Negative samples analysis in relevance feedback, IEEE Trans Knowledge Data Eng 19 (2007), 568–580.

D. Tao, X. Tang, and X. Li, Which components are important for interactive image searching? IEEE Trans Circuits Systems Video Technol 18 (2008), 3–11.

D. Tao, X. Tang, X. Li, and X. Wu, Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval, IEEE Trans Pattern Anal Machine Intelligence 28 (2006a), 1088–1099.

D. Tao, X. Tang, X. Li, and Y. Rui, Kernel direct biased discriminant analysis: A new content-based image retrieval relevance feedback algorithm, IEEE Trans Multimedia 8 (2006b), 716–727.

A.E. Taylor and D.C. Lay, Introduction to functional analysis, Wiley, New York, 1980.

P.C. Teo and D.J. Heeger, Perceptual image distortion, Proc SPIE 2179 (1994), 127–141.

R.D. Venakata, N. Sudhakar, B.L. Ramesh, and R.L. Pratap, Image quality assessment complemented with visual regions of interest, Proc IEEE Int Conf Comput: Theory Appl (2007), Kolkata, India, 681–687.

VQEG, Final Report from the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment, Available at http://www.vqeg.org/, Mar. 2000.

Z. Wang, A.C. Bovik, H. R. Sheikh, and E.P. Simoncelli, Image quality assessment: From error visibility to structural similarity, IEEE Trans Image Process 13 (2004), 600–612.

Z. Wang, A.C. Bovik, and L. Lu, Why is image quality assessment so difficult, Proc IEEE Int Conf Acoustics Speech Signal Process 4 (2002), 3313–3316.

Z. Wang, G. Wu, H.R. Sheikh, E.P. Simoncelli, E. Yang, and A.C. Bovik, Quality-aware images, IEEE Trans Image Process 15 (2006), 1680–1689.

A.B. Watson, DCT quantization matrices visually optimized for individual images, Presented at Human Vision, Visual Processing, and Digital Display IV, Bellingham, WA, 1993.

D. Zheng, J. Zhao, W.J. Tam, and F. Speranza, Image quality measurement by using digital watermarking, Proceedings of IEEE International Workshop on Haptic, Audio and Visual Environments and Their Applications, Ottawa, pp. 65–70 Sept. 2003.