A new reduced-reference image quality assessment using structural degradation model

Ke Gu, Guangtao Zhai, Xiaokang Yang, and Wenjun Zhang, *Fellow, IEEE*

Institute of Image Communication and Information Processing, Shanghai Jiao Tong University, Shanghai, China
Shanghai Key Laboratory of Digital Media Processing and Transmissions

***Abstract*—Image quality assessment (IQA) is an important research area in image processing. Reduced-reference (RR) IQA methods contained therein mainly aim to estimate image quality degradations with partial information about the reference image. Following the remarkable achievement of SSIM, structural information has been recognized as one key factor, and has aroused many image quality metrics so far. In this paper, we design a structural degradation model (SDM). Then, the quality score of an image is defined as a nonlinear combination, or SVM based integration, of distance between the structural degradation information of the original and distorted images. Accordingly, a new RR IQA approach using the SDM model is exploited. Experimental results on LIVE database are provided to justify the superior prediction accuracy performance of the proposed method as compared to three significant image quality metrics, PSNR, SSIM and FEDM.**

I. Introduction

Currently, perceptual image quality assessment (IQA) approaches with higher prediction accuracy become more extremely required, in both of the traditional area of image compression and denoising, and the modern domain of medical imaging and user experience. Generally, IQA methods can be divided into two categories: subjective quality assessment and objective quality assessment. The subjective assessment algorithm should be the terminal quality gauge, but it is seldom used because of its time-consuming and expensive shortages. Hence, there has been an increasing interest in developing objective IQA metrics since a decade ago.

According to the availability of reference images to be compared with during the tests, objective IQA method can be further classified into three kinds. Most approaches are known as full-reference (FR) methods, assuming the reference image is completely available. In plenty of practical applications, however, the reference image is not known, and no-reference (NR) image quality metrics are then desirable. As the tradeoff choice, the third type of reduced-reference (RR) IQA approaches are mainly applied to the situation where the reference image is only partially available and some extracted features are made available as side information to help to evaluate quality of the distorted image. From the viewpoint of future application, the RR approach has remarkably high value, not only as an important bridge effect between FR and NR, but also as the possible alternative method of NR. In this work, we focus on RR IQA approach.

Existing RR image quality metrics consist of four types. To model image distortions, the first type of methods [1]-[4] are mostly designed for specific application environments. The second type of approaches resort to model the human visual system (HVS) [5]-[6], where perceptual features motivated from computational models of low level vision were extracted to provide a reduced description of the image. Assuming that most real-world image distortions disturb image statistics and make the distorted image "unnatural", the third type of methods are exploited based on natural image statistics [7]-[8]. Finally, the fourth type of methods are recently proposed free energy based distortion metric (FEDM) [9]. Its underlying idea mainly depends on the free energy principle, proposed by Friston [10]-[11] to explain and unify several brain theories in biological and physical sciences, and the internal generative model to compute the gap between the encountered scene and brains prediction as the estimation of image quality.

It is well-known that the response of HVS is quite different for various spatial frequency [12]. It will be natural to associate that, after low-pass filtering, the spatial frequency response of a white noise image will largely reduce, while the spatial frequency response of a Gaussian blur distorted image will just drop a little. Besides, we also observed that the larger changes of spatial frequency responses generally correspond to white noise images with lower quality, and higher quality images contaminated by other four distortion types. Summarily, it can be concluded that images with various distortion categories and quality ranks, which are processed by low-pass filtering, have different degrees of spatial frequency response decrease. Accordingly, this paper defines structural degradation information of the original and distorted images, referring to some definitions in [13]-[14]. Then, the structural degradation model (SDM) based RR image quality metric can be developed by an effective nonlinear combination, or support vector machine (SVM) [15] based integration, of the distance between the above-mentioned two groups of structural degradation information.

The remainder of this paper is organized as follows. Section II first presents the definition of structural degradation information, and then proposes the SDM model based RR IQA approach. In Section III, experimental results using the laboratory for image and video engineering (LIVE) database [16] are reported and analyzed. Finally, concluding remarks and directions for future research are given in Section IV.

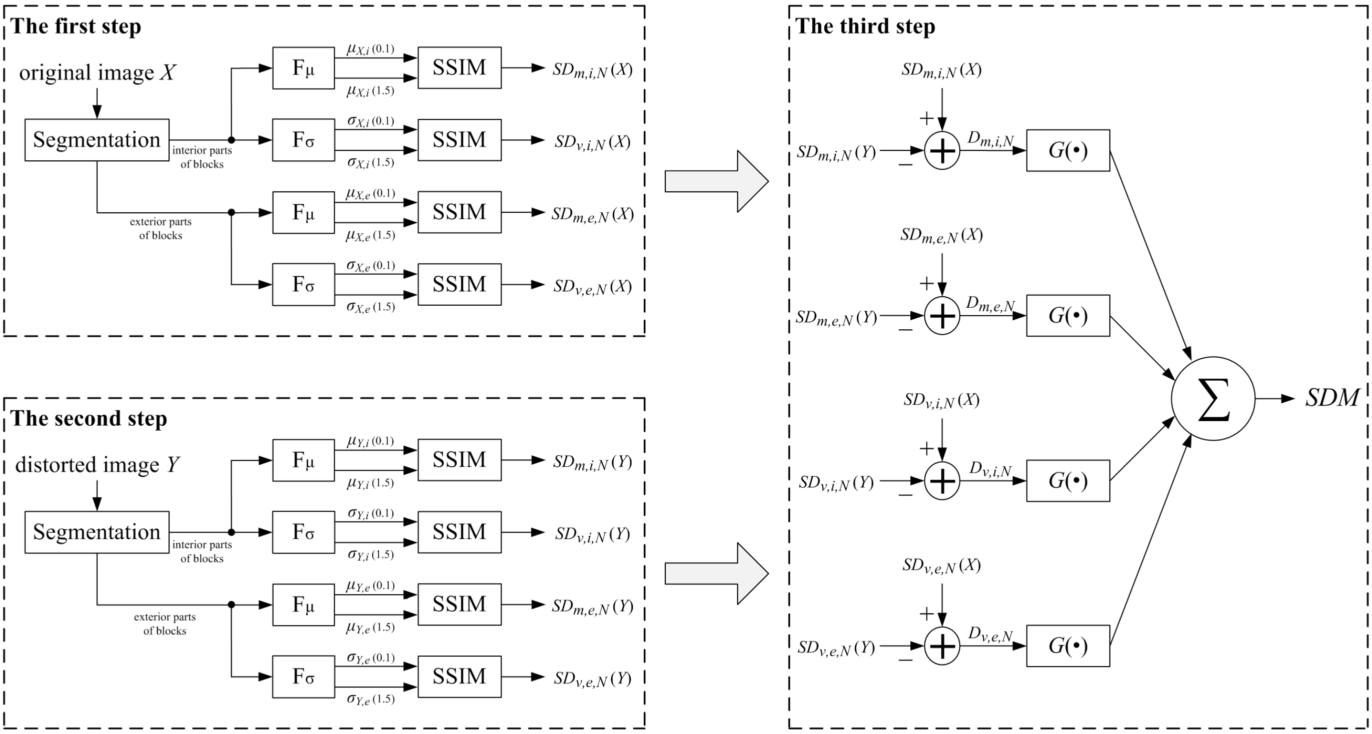


Fig. 1. The framework of the proposed SDM based RR IQA approach.

II. The Prediction Model

The proposed RR IQA algorithm mainly has three steps, and its whole framework is clearly shown in Fig. 1. Among the three dashed frames, the first step is employed to extract structural degradation information from reference images as side information in the encoding ends. The other two steps are the main body of the proposed method. More specific steps are as follows.

First of all, for an original and clear image X, its structural degradation information can be defined as

$$SD_{m,N}(X) = SSIM(\mu_X(0.1), \mu_X(1.5)) \tag{1}$$

$$SD_{v,N}(X) = SSIM(\sigma_X(0.1), \sigma_X(1.5)) \tag{2}$$

where $SSIM(\cdot)$ was given in [13]. Referring to the definitions of μ_X and σ_X in [13]-[14], we redefine $\mu_X(d)$ and $\sigma_X(d)$ as

$$F_\mu(X) = \mu_X(d) = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}x_{ij} \tag{3}$$

$$F_\sigma(X) = \sigma_X^2(d) = \frac{1}{N^2-1}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}(x_{ij}-\mu_X(d))^2 \tag{4}$$

with $\mathbf{w} = \{w_{ij}|i,j=1,\ldots,N\}$, satisfying $Sum(\mathbf{w}) = 1$ and $Var(\mathbf{w}) = d$ ($Sum(\cdot)$ and $Var(\cdot)$ are used to compute the sum and variance values). Thus, the features $SD_{m,N}(X)$ and $SD_{v,N}(X)$ are extracted from original image X.

With the same definitions of Eq. (1)-(2), the structural degradation information for distorted image Y can be evaluated by

$$SD_{m,N}(Y) = SSIM(\mu_Y(0.1), \mu_Y(1.5)), \tag{5}$$

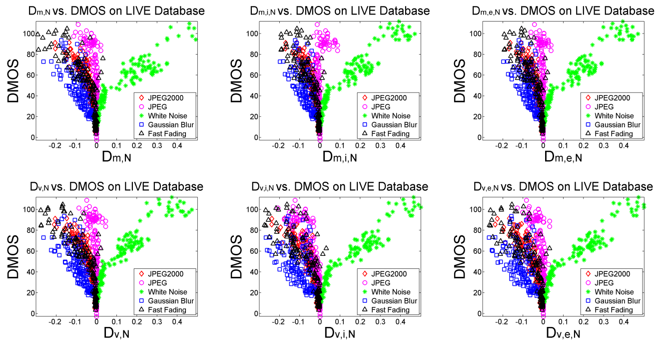
$$SD_{v,N}(Y) = SSIM(\sigma_Y(0.1), \sigma_Y(1.5)). \tag{6}$$

Then, following the definition of D_m and D_v as the distance between the original and distorted images' features:

$$D_{m,N} = SD_{m,N}(X) - SD_{m,N}(Y), \tag{7}$$

$$D_{v,N} = SD_{v,N}(X) - SD_{v,N}(Y), \tag{8}$$

an interesting phenomenon can be viewed in Fig. 2 (a), which shows that both of $D_{m,N}$ and $D_{v,N}$ have the ability of approximately predicting the image quality score for each specific distortion type. However, we noticed that the $D_{m,N}$ and $D_{v,N}$ values of JPEG compressed images almost equals to zero. So, we further classify them as $D_{m,i,N}$, $D_{m,e,N}$ (the $D_{m,N}$ value for interior parts and exterior parts of blocks), and $D_{v,i,N}$, $D_{v,e,N}$ (the D_v value for interior parts and exterior parts of blocks) to overcome the obstacle of no effect to JPEG distortion. As illustrated in Fig. 2 (b)-(c), for JPEG compressed images, the $D_{m,i,N}$ and $D_{m,e,N}$ (or $D_{v,i,N}$ and $D_{v,e,N}$) have quite different values, while they are almost the same for images with other four distortion types. Furthermore, it can be



(a) $D_{m,N}$ and $D_{v,N}$ (b) $D_{m,i,N}$ and $D_{v,i,N}$ (c) $D_{m,e,N}$ and $D_{v,e,N}$

Fig. 2. Scatter plots of six different distance values ((a): $D_{m,N}$ and $D_{v,N}$; (b): $D_{m,i,N}$ and $D_{v,i,N}$; (c): $D_{m,e,N}$ and $D_{v,e,N}$) vs. DMOS on LIVE database for five categories of image distortions.

easily found that the quality levels also have quite important influence, especially for white noise images.

Eventually, based on all the four variables' features, we define an effective nonlinear combination of them to measure the image quality as SDM model:

$$SDM = \sum_{s=\{m,v\},t=\{i,e\}} G(D_{s,t,N}) \tag{9}$$

where $G(D_{s,t,N})$ is defined as

$$\begin{aligned} G(D_{s,t,N}) &= H_p(D_{s,t,N}) \\ &= \begin{cases} H_l\,(D_{s,t,N}) & \text{if } H_l(D_{s,t,N}) \leq \xi_{thr} \\ H_h(D_{s,t,N}) & \text{otherwise} \end{cases} \end{aligned} \tag{10}$$

with ξ_{thr} being the threshold to discriminate high and low quality levels. And $H_p(D_{s,t,N})$ can be evaluated by

$$\begin{aligned} H_p(D_{s,t,N}) &= \lambda_{t,p} \cdot (D_{s,t,N})^{\gamma_{s,t,p}} \\ &+ \kappa_{s,t,p} \begin{cases} (D_{s,t,N})^{\delta_{s,t,p}} & \text{if } D_{s,t,N} \geq 0 \\ (-D_{s,t,N})^{\theta_{s,t,p}} & \text{otherwise} \end{cases} \end{aligned} \tag{11}$$

where $\lambda_{t,p}$, $\kappa_{s,t,p}$, $\gamma_{s,t,p}$, $\delta_{s,t,p}$ and $\theta_{s,t}$ $(s = \{m, v\}, t = \{i, e\}, p = \{l, h\})$ are model parameters. All the parameters are achieved by training as illustrated later. Furthermore, it will be natural to utilize SVM [15] for regression, and the proposed S-SDM can be given by

$$S\text{-}SDM = SVM(D_{s,t,N,f}, m_s) \tag{12}$$

where f represents the factor of down-sampling, following the idea in [17], and m_s is a model trained by SVM on the LIVE database, which consists of 29 source images. Here, the training and testing sets for Eq. (11)-(12) are totally content independent. The training set includes 617 images derived from 23 reference images, and the testing set contains the 162 images derived from the other 6.

III. Experimental Results

Here, we consider a four-parameter logistic function that is suggested by VQEG [18] as the nonlinear regression between the subjective scores and the prediction scores of five metrics PSNR, SSIM, FEDM and the exploited SDM and S-SDM methods:

$$q(x) = \frac{\beta_1 - \beta_2}{1 + exp(-(x - \beta_3)/\beta_4)} + \beta_2 \tag{13}$$

with x and $q(x)$ being the input score and the mapped score, respectively. The free parameters β_1 to β_4 are to be determined during the curve fitting process.

Two commonly used performance metrics, Pearson linear Correlation Coefficient (PLCC) and Spearman Rank-Order Correlation Coefficient (SROCC) as suggested by VQEG [18], are employed to further evaluate two competitive SDM metrics and the other three methods, namely PSNR, SSIM and FEDM, on LIVE database. Their values are tabulated in Table I, and the scatter plots of S-SDM on LIVE database and five data sets of different distortion types are shown in Fig. 3. It can be seen that our proposed RR IQA paradigm has achieved much better results than the most two popular image quality metric PSNR and SSIM, and recently proposed FEDM.

Except superior performance, it is important to clarify that our paradigm has three paramount merits: First, low computational complexity and high execute speed because of parallel processing; Second, strong portability due to the fact that the proposed method can be computed by some basic operations, such as addition and multiplication, and the SSIM algorithm, which has been inserted into a great many systems and softwares as one of benchmark image quality metrics; Third, only four numbers of $SD_{m,i,N}(X)$, $SD_{m,e,N}(X)$, $SD_{v,i,N}(X)$ and $SD_{v,e,N}(X)$ as features extracted from the original image, which is usually negligible as compared to the image files size and can be encoded precisely with only a few bits in the header file.

TABLE I
PEARSON LINEAR CORRELATION COEFFICIENT (PLCC) AND SPEARMAN RANK-ORDER CORRELATION COEFFICIENT (SROCC) VALUES (AFTER NONLINEAR REGRESSION) OF PSNR, SSIM, FEDM AND THE PROPOSED SDM AND S-SDM METHODS ON WHOLE LIVE DATABASE (779 IMAGES), AND FIVE DATA SETS OF DIFFERENT DISTORTION CATEGORIES.

Pearson Linear Correlation Coefficient (PLCC)					
	FEDM	SDM	PSNR	SSIM	S-SDM
JP2K (169)	0.9260	0.9375	0.8996	0.9410	0.9738
JPEG (175)	0.9210	0.9591	0.8878	0.9504	0.9476
White noise (145)	0.9250	0.9718	0.9860	0.9697	0.9911
Gaussian blur (145)	0.9020	0.9240	0.7834	0.8743	0.9339
Fast-fading (145)	0.8750	0.9300	0.8895	0.9428	0.8555
All (779)	–	–	0.8701	0.9014	0.9330

Spearman Rank-Order Correlation Coefficient (SROCC)					
	FEDM	SDM	PSNR	SSIM	S-SDM
JP2K (169)	0.9200	0.9410	0.8954	0.9355	0.9703
JPEG (175)	0.9225	0.9520	0.8809	0.9449	0.9440
White noise (145)	0.9144	0.9697	0.9857	0.9625	0.9897
Gaussian blur (145)	0.9310	0.9332	0.7823	0.8944	0.9257
Fast-fading (145)	0.8520	0.9418	0.8907	0.9413	0.8439
All (779)	–	–	0.8755	0.9103	0.9364

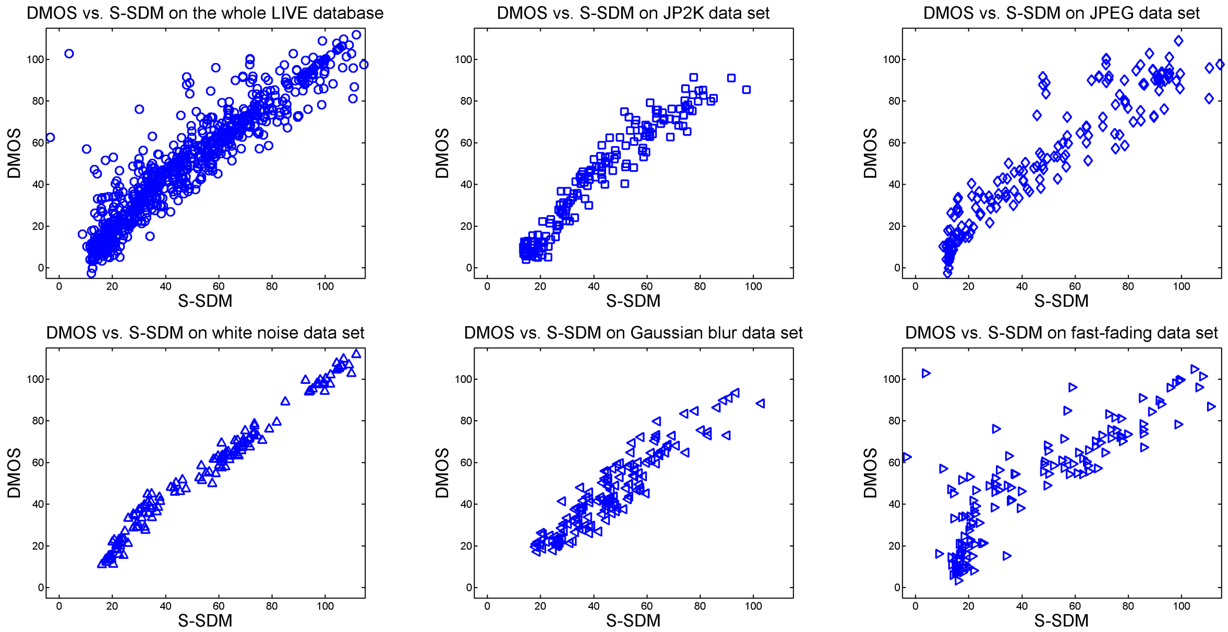


Fig. 3. Scatter plots of DMOS vs. the proposed S-SDM approach on the whole LIVE database and five data sets of different distortion types.

IV. Conclusion

In this paper, we develop a new structural degradation model inspired RR IQA paradigm, which is mainly relying on different spatial frequency response decrease for various distortion types and quality levels. Experimental results on LIVE database verify that the performance of the proposed method is clearly better than full-reference PSNR and SSIM, and recently proposed FEDM algorithms. Moreover, our RR IQA approach also has other inspiring merits, such as low computational complexity, high execute speed, strong portability, and very little reduced-reference information.

Besides, it was also found that the extracted features of FEDM and SDM have strong dependency, which will provide the possible opportunity to extend the proposed method to an effective no-reference image quality assessment approach in the near future.

Acknowledgment

This work was supported in part by NSERC, NSFC (61025005, 60932006, 61001145), SRFDP (20090073110022), postdoctoral foundation of China 20100480603, 201104276, postdoctoral foundation of Shanghai 11R21414200, the 111 Project (B07022), and STCSM (12DZ2272600).

References

[1] I. P. Gunawan and M. Ghanbari, "Reduced reference picture quality estimation by using local harmonic amplitude information," in *Proc. London Commun. Symp.*, Sep. 2003, pp. 137-140.

[2] T. M. Kusuma and H. J. Zepernick, "A reduced-reference perceptual quality metric for in-service image quality assessment," in *Proc. Joint 1st Workshop Mobile Future and Symp. Trends Commun.*, pp. 71-74, Oct. 2003.

[3] S. Wolf and M. Pinson, "Low bandwidth reduced reference video quality monitoring system," in *Proc. Int. Workshop Video Process. Quality Metrics for Consumer Electron.*, Scottsdale, AZ, Jan. 2005, CD-ROM.

[4] P. Le Callet, C. Viard-Gaudin, and D. Barba, "Continuous quality assessment of MPEG2 video with reduced reference," in *Proc. Int. Workshop Video Process. Quality Metrics for Consumer Electron.*, Scottsdale, AZ, Jan. 2005.

[5] M. Carnec, P. Le Callet, and D. Barba, "An image quality assessment method based on perception of structural information," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2003, vol. 3, pp. 185-188.

[6] M. Carnec, P. Le Callet, and D. Barba, "Visual features for image quality assessment with reduced reference," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2005, vol. 1, pp. 421-424.

[7] Z. Wang, G. Wu, H. R. Sheikh, E. P. Simoncelli, E. H. Yang, and A. C. Bovik, "Quality-aware images," *IEEE Trans. Image Process.*, vol. 15, no. 6, pp. 1680-1689, Jun. 2006.

[8] Q. Li and Z. Wang, "Reduced-reference image quality assessment using divisive normalization-based image representation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 2, Apr. 2009.

[9] G. Zhai, X. Wu, X. Yang, W. Lin, and W. Zhang, "A psychovisual quality metric in free-energy principle," *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 41-52, Jan. 2012.

[10] K. Friston, J. Kilner, and L. Harrison, "A free energy principle for the brain," *Journal of Physiology Paris*, vol. 100, pp. 70-87, 2006.

[11] K. Friston, "The free-energy principle: a unified brain theory?" *Nature Reviews Neuroscience*, vol. 11, pp. 127-138, 2010.

[12] Charles F. Hall and Ernest L. Hall, "A nonlinear model for the spatial characteristics of the human visual system," *IEEE Trans. Systems, Man, and Cybernetics*, vol. 7, no. 3, pp. 161-170, March 1977.

[13] 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.*, vol. 13, no. 4, pp. 600-612, Apr. 2004.

[14] K. Gu, G. Zhai, X. Yang, and W. Zhang, "An improved full-reference image quality metric based on structure compensation," in *Proc. APSIPA ASC 2012*, 2012.

[15] C-C. Chang and C-J. Lin, "LIBSVM: a library for support vector machines," 2001, http://www.csie.ntu.edu.tw/∼cjlin/libsvm.

[16] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, "LIVE image quality assessment Database Release 2," [Online]. Available: http://live.ece.utexas.edu/research/quality.

[17] Matthew D. Gaubatz and Sheila S. Hemami, "On the nearly scale-independent rank behavior of image quality metrics," *Proc. IEEE Int. Conf. Image Process.*, 2008.

[18] VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment," March 2000, http://www.vqeg.org/.