RANK LEARNING ON TRAINING SET SELECTION AND IMAGE QUALITY ASSESSMENT

Long Xu[1,2], Weisi Lin[1], Jia Li[3], Xu Wang[4], Yihua Yan[2], Yuming Fang[1]
[1] School of Computer Engineering, Nanyang Technological University, Singapore
[2] National Astronomical Observatories, Chinese Academy of Sciences, Beijing, China
[3] National Engineering Laboratory of Video Technology, School of EE&CS, Peking University, Beijing, China
[4] Department of Computer Science, City University of Hong Kong, Hong Kong
{xulong, wslin}@ntu.edu.sg, jia.li@pku.edu.cn, wangxu.cise@gmail.com, yyh@nao.cas.cn, ymfang@ntu.edu.sg

ABSTRACT

Machine learning (ML) techniques are widely used in recent no-reference visual quality assessment (NR-VQA) metrics by training on subjective image quality databases. In these metrics, the optimization function is constructed based on L^2 norm of the distance between subjective image quality and predicted image quality. There are two problems in these L^2 norm based methods: (1) human's opinion on subjective image quality rating is not reliable at fine-scale level. A small difference between subjective image qualities represented by mean opinion scores (MOSs) of two images may not truly reflect the real quality difference between these two images, but acts as noise. The optimization process should avoid such noise. (2) Generally, human's opinion on pairwise comparison (PC) for image quality is more reliable and believable than MOS. The importance of PC is ignored during the optimization process of existing ML-based studies, which are designed based on the numerical rating system.

In this paper, we introduce image quality ranking concept to establish a new optimization objective instead of L^2 norm optimization, and then a novel NR-VQA is constructed based on ranking learning. The proposed metric firstly suggests a reasonable training set for ML, which is ignored by existing ML-based NR-VQA. The ranking theory is adopted to build optimization function, which reflects the properties of PC over the numerical ranting system used by traditional NR-VQA. By ignoring the small difference between MOSs from two images during the optimization process, the proposed ranking-based NR-VQA can also well address the first problem from the existing related metrics. Experimental results show that the proposed ranking-based NR-VQA can obtain better performance over the state-of-the-art NR-VQA approaches.

Index Terms— Rank learning, image quality assessment, gradient decedent, optimization

1. INTRODUCTION

Since the traditional PSNR/MSE measurement of visual signal (image and video) cannot truly reflects the real human-perceived quality of distorted visual signal, there have been many visual quality assessment (VQA) approaches from the aspects of the human visual system (HVS) and human visual psychology during the last decade. Based on the availability of references, VQA is usually classified into full-reference (FR), no-reference (NR) and reduced-reference (RR) approaches. The FR quality metrics require the whole information of the reference image or video to evaluate the visual quality of the distorted one. One popular category of FR approaches are based on structural similarity (SSIM) of images, which was firstly investigated by Wang et al. in [1][2] and extensively explored subsequently. In [3], Liu et al. claimed that the gradient similarity contribute significantly to structural similarity of images. In [4], Zhang et al. investigated both gradient magnitude and phase congruency for measuring image structural similarity. Ma et al. proposed to incorporate orientation sensitivity and conspicuity of the HVS into SSIM for its improvement in [5]. In [6], Zhang et al. explored anisotropic regularity and irregularity of edges presented in images, and proposed an edge similarity based SSIM (ESSIM). Wang et al studied 3D structural similarity of video from both spatial and temporal directions for video quality assessment by using 3D structure tensor [7].

There have been some well-established FR-VQA, however in many real-world applications, the reference visual signal is not available, e.g., image/video decoding in clients, image/video denoising and super-resolution. In order to evaluate the quality of these processed images or videos, the RR and NR VQA are thus developed. In RR-VQA, information of references is partially extracted and stored into bitsteam along with distorted signals. For example, in video streaming applications, some parameters are extracted from reference video in encoder, and used for quality assessment in decoder since the reference video is unavailable in decoder. In [8], a RR-VQA approach was proposed by exploiting the spatial information loss and the temporal statistical characteristics of the interframe histogram. In [9], the color correlogram was used to analyze the alterations in the color distribution of an image as a consequence of the occurrence of distortions. In [10], Wang et al. modeled the marginal probability distribution of the wavelet coefficients of a natural image by the generalized

Gaussian density (GGD) function [11], and used Kullback-Leibler distance between the marginal probability distributions of wavelet coefficients of the reference image and distorted one to measuring distortion intensity. To decrease the auxiliary information, divisive normalization transform (DNT) [12] was employed to depict the coefficient in the wavelet domain. In [13], the DNT was investigated in Contourlet domain.

As to NR-VQA, three categories of approaches were presented in the literature. The *first* category of approaches take the behavior of specific distortions into consideration. In [19], Sheikh et al. employed wavelet statistical model to capture the distortion introduced by JPEG 2000. Liang et al. [20] combined the sharpness, blurring, and ringing measurements together to evaluate the visual quality of the JPEG 2000 coded image. Brandao et al. [21] proposed an NR-VQA approach based on the DCT domain statistics to evaluate the quality of JPEG coded image. In [22], R. Ferzli et al. integrated the concept of just noticeable blur into probability summation model to measure sharpness/ blurriness. The *second* category of approaches use quality aware clustering. They group the image patches of training set into the given number of classes based on local image features, such as histogram of oriented gradients (HoG), difference of Gaussian (DoG) and Gabor filter. Each cluster center has a quality score which is derived from the qualities of image patches falling into this cluster. Associating cluster centers with their qualities, the researchers established a codebook. Patches of a test image look up codebook to search the most similar codewords and retrieve the associated quality values. In [16], a visual codebook associated Gabor filter based local appearance descriptors with DMOS. The authors of [17] used FSIM [18] instead of DMOS as image patch quality to establish codebook. The *third* category is to utilize machine learning tools to map image features onto image qualities. In [14], Moorthy et al. proposed to use support vector machine (SVM) and support vector regression (SVM) to learn a classifier and an ensemble of regressors for image distortion classification and computing quality of specific distorted image. It deploys summary statistics derived from an on natural scene statistics (NSS) wavelet coefficient model, using a two-step framework for VQA: distortion classification and distortion specific VQA regression. In [15], Tang et al. proposed an approach similar to [14] but with more elaborate features, including distortion texture statistics, blur/noise statistics and histogram of subband. Our method in this paper is closely related to these two papers.

Regarding SVM and SVR used in [14][15], the optimization objective is the minimum of sum of squares of distances between subjective visual quality scores and objective visual quality indexes, i.e., L^2 norm. The objective of L^2 norm optimization is to pursuit the minimum of the sum of absolute numerical difference between observations and predicted/regressed values. However, such an optimization objective cannot address image quality assessment very well. The reason lies in: 1) *the numerical image quality, e.g., with rate of 1-5, is not exactly possessing a strong confidence for measuring real image quality. The small difference of image quality scores may not truly reflect the real difference of image qualities, but acts as noise*; 2) *to assess image quality, pairwise competition is more reliable/reasonable than numerical quality rating. Subjects can accurately select the better one from two pair compared images. However, people cannot rate an image precisely from the single stimulated experiment in image quality rating system*; 3) *the diversity of image content and distortion types also make it difficult to rate image quality numerically under complex scenarios*. Therefore, in this paper, we propose a new optimization objective for VQA based on rank learning. The rank concept is introduced into the optimization.

The rest of this paper is arranged as follow. Section 2 describes the proposed rank learning based VQA in detail. Section 3 presents experimental results. And, the final section concludes this paper.

2. RANK LEARNING BASED NR-VQA

Performance of an objective quality assessment metric can be expressed by how well the objective quality metric predicts the real perceptual visual quality, and specifically, the correlation between objective quality indexes computed from the objective quality metric and subjective quality scores (e.g., MOS or DMOS) provided by subjective quality database is used to indicate such prediction accuracy. For reliable prediction of image quality, the objective quality metrics should be consistent with the behavior of the HVS. In practice, the HVS' perception is too complex to be modeled directly, so machine learning can be employed to mimic it. In machine learning based approaches, such as [14] and [15], the subjective quality scores are used as ground-truth for training a regression model which could be used to predict image quality in VQA tasks. In [14], a VQA model is derived from the SVM regression process which concerns a supervised learning process with the inputs of global image features (NSS) and ground-truth (subjective quality score). In [16] and [17], the local features, DoG and Gabor features which are extracted from local area of each pixel in an image, are clustered to form a redundancy dictionary named codebook to represent the input images. This clustering process can be regarded as an unsupervised learning process. Each codeword is connected to a visual quality value which is deduced from the given DMOS during clustering process.

For the task of objective quality assessment, machine learning approaches are usually used to deduce regression models that approximate perceptual visual quality. The regression model has a specific function form φ_β that depends on a vector of regression parameter β, which is determined by the Least Squares method. Specifically, the minimization of the difference between predicted objective quality indexes and subjective quality scores given by DMOS as

$$\beta^* = \underset{\beta}{\operatorname{argmin}} \left\{ \sum_i \left\| \varphi_\beta(x_i) - g_i \right\|_2 \right\}, \tag{1}$$

where φ represents a learner which computes image quality from each input image. Usually x_i represented by an N-D vector contains the important structural and textural feature of an image. g_i is the ground-truth given by subjective image quality score. For simplicity, the linear learner which can be represented by $\varphi(x_i)=\omega^T x_i$ is used in (1). For fitting more situations, nonlinear relation between input x and $\varphi(x)$ is explored, which can be converted into linear problem by using kernel functions in SVM methods. Observing the optimization objective of (1), the L^2 norm is optimized. However, as mentioned above, people's opinion to pairwise image quality competition is more reasonable than image quality rating. The pairwise competition is more reliable and robust than L^2 norm to model VQA. Thus, in this work, a new optimization objective is established based on image quality ranking instead of L^2 norm, which is formulated as

$$\min_{\varphi} \left\{ \sum_{i \neq j} \left[g_i < g_j \right]_{\mathrm{I}} \left[\varphi(x_i) \geq \varphi(x_j) \right]_{\mathrm{I}} \right\}, \tag{2}$$

where $[x]_{\mathrm{I}}=1$ if the logic decision x holds, otherwise $[x]_{\mathrm{I}}=0$. It focuses on the ranking of image qualities instead of numerical difference between subjective image qualities and predicted objective image qualities. From (2), the false ranking, i.e., the objective quality ranks the inverse order of the ground-truth (subjective image quality), would result in the increase of the cost of optimization objective. Eq. (2) concerns all pairwise comparisons of image qualities among all images. Obviously, an image which has the distinct quality difference from others would contribute more to the optimization objective of (2). Intuitively, if such an image ranks wrong order, i.e., contradictory to ground-truth, the penalty must be large to avoid such a situation. The images with similar image qualities tend to have low weight to the optimization objective (In practice, only the image pairs with the quality difference larger than a predefined threshold, i.e., $|g_i-g_j|>T$ are included in (2)). To make this statement more clear, an illustration is drawn in Fig. 1. From Fig. 1, if the target distinguishes from others significantly, it would have more competitions with others so that having a more important role in optimization objective. For example, the samples in the circle in red should play significantly in determining a regression formula, however the L^2 norm optimization objective would ignore these samples since the overwhelming majority of the samples (95%) are in a straight line. In rank based regression, each sample in the subset in red circle would compare with all samples in other two subsets. Therefore, these 5% samples play significantly in optimization although they account for a small portion of the all.

To seek the solution of optimization problem of (2), the feature mapping function $\varphi(x)$ which describes the relation between image feature and objective image quality index should be predefined. For simplicity, the linear function $\varphi(x)$

Table I: The global feature of image

Feature ID	Feature Description	Computation procedure
f1-f12	Variance of subband (σ)	Fitting a generalized Gaussian to subband coefficients
f12-f24	Shape parameter of subband (γ)	Fitting a generalized Gaussian to subband coefficients
f25-f29	Shape parameter across scale	Fitting a generalized Gaussian to combined subband coefficients (different scale & same orientation, the one among them consisting of all subbands)
f30-f41	Correlations across scales	Computing SSIM between the high pass subband and all band pass subbands
f42-f61	Spatial correlation across subbands	Fitting a polynomial (5 coefficients) to the correlation function
f62-f73	Across orientation statistics	Computing SSIM between adjacent orientations at same scale (for the largest two scales)
f74-f521	Histogram of each subbands	64, 32 and 16 bins for the first, second and third scale respectively

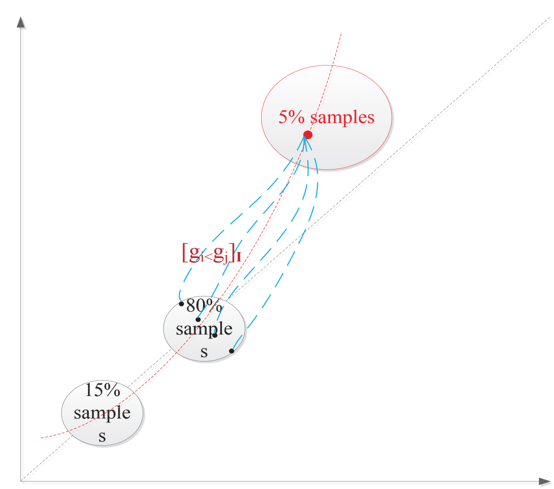


Fig. 1. A flow chart illustrating the disadvantage of L^2 norm optimization

$= \omega^T x$ is assumed in our work following the state-of-the-art works. Thus, the optimization becomes to seek a vector ω which results in the minimum of (1) on training data. In addition, the feature of an image denoted by x is represented by the statistical parameters of pyramid decomposition of an image. Assuming 4 orientations and 3 scales, there would be 12 subbands totally in the pyramid decomposition. According to [14], the feature vector of length 521 is composed of the entries shown in Table I, where variance of each subband, correlation across orientation and scale, shape parameters of subbands and histograms of subbands are included. Here, the feature vector x_i with 521 entries to characterized the global feature of an image. The learning task is to derive ω of the linear model φ(x) on the ground-truth (subjective visual quality scores $\{g_i\}$) which are provided by the subjective image quality database [23]. With the assumption of linear function $\varphi(x)$, the optimization objective of (2) is rewritten as

$$\min_{\omega}\left\{\sum_{i\neq j}\left[g_i < g_j\right]_{\mathrm{I}}\left[\omega^T x_i \geq \omega^T x_j\right]_{\mathrm{I}}\right\}. \qquad (3)$$

Let $L(\omega) = \sum_{i\neq j}\left[g_i < g_j\right]_{\mathrm{I}}\left[\omega^T x_i \geq \omega^T x_j\right]_{\mathrm{I}}$, we call $L(\omega)$ the empirical loss. Since of rank function, we encounter a non-convex optimization problem. As in [24], the Boolean terms related to ω in (3) is replaced by their upper bounds to facilitate the optimization as

$$[\omega^T x_i \leq \omega^T x_j]_I \leq \exp(\omega^T x_i - \omega^T x_j)\,, \qquad (4)$$

where the exponential upper bound is used since it is convex and can facilitate the optimization. After replacing the only one term containing the variable ω in (3), the empirical loss function would turn out to be convex. Then, the gradient decedent method can be employed to solve (3). Note that we have

$$\frac{\partial}{\partial\omega}\exp(\omega^T x_i - \omega^T x_j) = (x_i - x_j)\exp(\omega^T x_i - \omega^T x_j)\,, \quad (5)$$

so the gradient decent direction can be written as

$$\Delta\omega = \lambda\times\sum_{i\neq j}\left[g_i < g_j\right]_{\mathrm{I}}(x_i - x_j)\exp(\omega^T x_i - \omega^T x_j)\,, \quad (6)$$

where λ acts as a iteration step controlling the convergence speed.

From (3), given an initial ω, with the ground-truth $\{g_i\}$, the image features $\{x_i\}$, the empirical loss $L(\omega)$ can be initialized. Next, replacing ω by $\omega+\Delta\omega$, we update $L(\omega)$. By iteratively updating ω and $L(\omega)$, the global minimum objective can be reached. The detailed rank learning process is given in Table II.

It should be pointed that the optimization objective (3) is established intrinsically on image quality ranking instead of image quality rating, so it is suitable for ranking images with respect to their qualities, however, it cannot be directly used to rate image quality, i.e., predict objective image quality to each image. Since the evaluation system of VQA mostly concerns the relative image qualities instead of numerical values, we just use the ranking results on the testing databases as the predicted objective qualities for computing the correlation coefficients for simplicity. Further investigation on conversion of image ranks to image quality would increase the performance of the proposed algorithm.

Table II: The rank learning algorithm

Input: the global attributes $\{x_i\}$,{ $i=0,1,2,\ldots,n$} for training images ; ground-truth subjective quality scores { g_i }, iterative times T, and a threshold for iterative determination Δe.
Output: Model parameter ω
Begin
  Initialization:
  Initialize ω randomly ;
  $e(1) \leftarrow L(\omega)$;
  $t \leftarrow 1$;
  Optimization:
  repeat
    for k from 1 to T
    $\omega \leftarrow \omega+\Delta\omega$;
    $e(1) \leftarrow L(\omega)$;
    $t \leftarrow t+1$;
  until $t \geq T$ or $(e(t)-e(t-1)) < \Delta e$
end

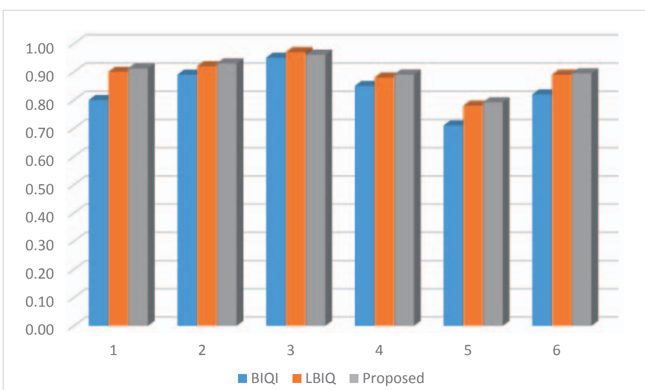
3. EXPERIMENTS AND DISCUSSIONS

The performance of a VQA metric can be evaluated by depicting the correlation between objective quality indexes and subjective quality ratings, specifically the DMOS value of each distorted image. The DMOS value is obtained by subjective viewing tests where many observers participated and provided their opinions to the visual quality of each distorted image. Therefore, it can be regarded as the ground-truth for evaluating the performance of a VQA approach. The subjective image quality databases: LIVE image [23], TID2008 [25], IVC [26], Toyama [27] and A57 [28] can be accessed publicly and widely used as ground-truth by researchers. These image databases contains the most common distortion types, such as blockiness caused by block based compression at low bit rate, ringing caused by wavelet transform, white noise, Gaussian blur and Rayleigh fading channel. In addition, they also contain diverse visual contents.

We perform experiments on LIVE image database which consists of 29 reference images, each image has 5 distortion types (JPEG, JP2K, white noise (WN), Gaussian blur (GB) and fast fading (FF) channel distortions) and 5/6 distortion levels per type. The samples in database are divided into training set and testing set. The 50 percent of samples are randomly selected to form training set, and 100 times training process as in (3) are performed. The average of the computed parameters ω is the final regression model of rank learning based. The effectiveness of a VQA method is usually evaluated by three statistical measurements usually, linear correlation coefficient (LCC), Spearman correlation coefficient (SPRCC) and RMSE, which can measure the correlation between DMOS and predicted image qualities. Specifically, LCC measures prediction accuracy, SROCC evaluates prediction monotonicity, and RMSE indicates the error during the fitting process. Larger LCC and SROCC values indicate better correlation between objective and subjective image qualities, while smaller RMSE values mean smaller error of predictions, therefore a better performance. The LCC, SROCC and RMSE are 0.9017, 0.8947 and 7.6266 respectively for our proposed VQA approach.

Regarding the features, only pyramid decomposition is employed, 4 direction and 3 scales are used in our proposed method. Besides the classical NSS feature, the normalized histogram of each subband is contaminated with the NSS feature. No other elaborate features are explored. As to algorithmic principle, only rank function is used in optimization formulas, and the gradient decedent is used to find the best answer of the optimization problem.

We also performed a comparison to two closely related works, BIQI [14] and LBIQ [15]. As shown in Fig. 2, our proposed algorithm is competitive compared to those two NR-VQA approaches in terms of SROCC. Since only the computation of pyramid decomposition is concerned in the proposed VQA, there is no more computational complexity is introduced compared with these two benchmarks.



	JP2K	JPEG	WN	GB	FF	Overall
BIQI	0.80	0.89	0.95	0.85	0.71	0.82
LBIQ	0.90	0.92	0.97	0.88	0.78	0.89
Proposed	0.9123	0.9301	0.961	0.8906	0.7921	0.8947

Fig. 2. Spearman correlation between subjective quality and predicted quality. The numbers at horizontal axis represent distortion types: JP2K, JPEG, WN, GB, FF and their combination. The vertical axis shows the correlation measurement.

To verify the claimed train set selection, we change the threshold mentioned in Section 2, i.e., $|g_i\text{-}g_j|>T$, to select the training data. The new training set is used in the proposed VQA to check its performance variation. The experimental results are drawn in Fig. 3. From Fig. 3, the performance increases with the threshold as the threshold is less than 8.8. It indicates the noise mentioned above is compressed to result in performance improvement.

It should be noticed that the features used in these three compared methods are different. Since this work focuses on the optimization of rank learning, the effort of exploring image feature does not addressed enough which would be investigated in our feature work. Analyzing the compromised performance of the proposed method, there are three possible reasons: 1) the image feature defined for conventional optimization may not optimal for our task; 2) the gradient descendent algorithm for solving the proposed optimization objective cannot handle non-negative constraints, which results the negative weights for integrating the components of the defined feature in Table I; 3) the connection between image quality index and image quality ranks is not that straightforward. Regarding these reasons, the proposed algorithm can be therefore improved from the efforts to solve the reasons listed above.

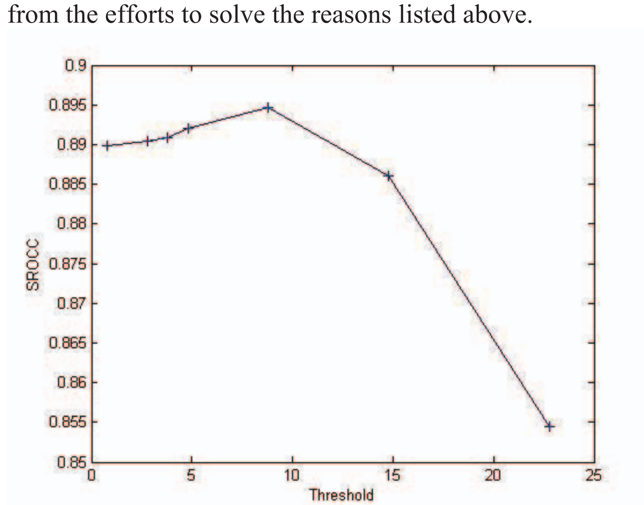


Fig. 3. The SROCC of the proposed NR-VQA with the threshold variation (T= 0.8, 2.8, 3.8, 4.8, 8.8, 14.8, 22.8, and the best performance is at T=8.8 in this experiment).

4. CONCLUSIONS

In this paper, we focus on rank based optimization for image quality assessment. This new optimization objective is based on image quality ranking instead of L^2 norm of numerical prediction error of image quality. It overcomes the vulnerability of L^2 norm optimization to image quality assessment. In addition, it takes advantage of pairwise image quality comparison for performing image quality assessment.

Although there were tons of image quality databases, their subjective quality rating systems are independent and incompatible. Using image quality rank instead of image quality rate, it is possible to co-train VQA model over multiple image quality databases since ranking system only concerns pairwise image quality ranking. The authors would investigate co-training on multiple databases in the near future.

5. ACKNOWLEDGMENT

This work is supported by National Natural Science Foundation of China under Grants 61202242, and 61370113, and Singapore National Research Foundation under its IDM Futures Funding Initiative and administered by the Interactive & Digital Media Programme Office, Media Development Authority.

REFERENCES

[1] A. Bhat, S. Kannangara, Y. F Zhao and I. Richardson, “A full reference quality metric for compressed video based on mean squared error and video content,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 2, pp. 165-173, Feb. 2012.

[2] Z. Wang and A. C. Bovik, “Mean squared error: love it or leave it? - A new look at signal fidelity measures,” IEEE Signal Processing Magazine, vol. 26, no. 1, pp. 98-117, Jan. 2009.

[3] A. Liu, W. Lin, M. Narwaria, “Image Quality Assessment Based on Gradient Similarity,” IEEE Transaction on Image Processing, vol. 21, no. 4, pp. 1500-1512, 2012.

[4] Lin. 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, Aug., 2011.

[5] L. Ma, S. Li, and K. N. Ngan, “Visual horizontal effect for image quality assessment,” IEEE Signal Process. Letters, vol. 17, no. 7, pp. 627-630, Jul. 2010.

[6] X. D. Zhang, X. C. Feng, W. W. Wang, W. F. Xue, “Edge Strength Similarity for Image Quality Assessment,” IEEE Signal Process. Lett. vol. 20, no. 4, pp. 319-322, 2013.

[7] Y. Wang, T. T Jiang, S. W. Ma and W. Gao, “Novel spatio-temporal structural information based video quality metric,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 7, pp. 989-998, July, 2012.

[8] L. Ma, S. Li, and K. N. Ngan, “Reduced-Reference Video Quality Assessment of Compressed Video Sequences,” IEEE Trans. Circuits Syst. Video Technol., vol. 22, no. 10, pp. 1441–1456, Oct. 2012.

[9] J. A. Redi, P. Gastaldo, I. Heynderickx, and R. Zunino, “Color distribution information for reduced-reference assessment of perceived image quality,” IEEE Trans. Circuits Syst. Video Technol., vol. 20, no. 12, pp. 1757–1769, Dec. 2010.

[10] Z. Wang and A. C. Bovik, “Reduced and no-reference image quality assessment: The natural scene statistic model approach,” IEEE Signal Process. Mag., vol. 28, no. 6, pp. 29-40, Nov. 2011.

[11] Z. Wang, G. Wu, H. R. Sheikh, E. P. Simoncelli, E.-H. Yang and A. C. Bovik, "Quality-aware images," IEEE Transactions on Image Processing, vol. 15, no. 6, pp. 1680-1689, June 2006.

[12] Q. Li, 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, pp. 202-211, 2009.

[13] X. Wang, G. Jiang, M. Yu, “Reduced Reference Image Quality Assessment Based on Contourlet Domain and Natural Image Statistics,” Fifth International Conference on image and Graphics (ICIG2009), Xi’an, China, Sep.21-23, 2009.

[14] A.K. Moorthy, A. C. Bovik, “Blind image quality assessment: From natural scene statistics to perceptual quality,” IEEE Trans. Image Processing, vol. 20, no. 12, pp. 3350-3364, 2011.

[15] H. Tang, N. Joshi and A. Kapoor, “Learning a Blind Measure of Perceptual Image Quality,” In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), Colorado Springs, 2011.

[16] P. Ye, D. Doermann, “No-reference image quality assessment using visual codebooks,” IEEE Trans. Image Processing, vol. 21, no. 7, pp. 3129-3138, 2012.

[17] W. Xue, L. Zhang, X. Mou, “Learning without Human Scores for Blind Image Quality Assessment,” In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), pp. 995-1002, 2013.

[18] Lin. 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, Aug., 2011.

[19] H. R. Sheikh, A. C. Bovik, and L. Cormack, “No-reference quality assessment using nature scene statistics: JPEG 2000”, IEEE Trans. Image Process., vol. 14, no. 11, pp. 1918-1927, Nov. 2005.

[20] L. Liang, S. Wang, J. Chen, S. Ma, D. Zhao, and W. Gao, “Noreference perceptual image quality metric using gradient profiles for JPEG 2000”, Signal Processing: Image Communication, vol. 25, no. 7, pp. 502-516, Aug. 2010.

[21] T. Brandao, and M. P. Queluz, “No-reference image quality assessment based on DCT domain statistics”, Signal Processing, vol. 88, no. 4, pp. 822-833, Apr. 2008.

[22] R. Ferzli, L. J. Karam, “A no-Reference objective image sharpness metric based on the notion of just noticeable blur (JNB)”, IEEE Trans. Image Process., vol. 18, no. 4, Apr. 2009.

[23] LIVE Image Quality Database. Available [Online]: http://live.ece.utexas.edu/research/quality/subjective.htm.

[24] J. Li, Y. H. Tian, T. J. Huang and W. Gao, “Cost-Sensitive Rank Learning from Positive and Unlabeled Data for Visual Saliency Estimation,” IEEE Signal Process. Letters, vol. 17, no. 6, pp. 591-594, 2010.

[25] 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,” Adv. Modern Radioelectron., vol. 10, pp. 30–45, 2009.

[26] A. Ninassi, P. Le Callet, and F. Autrusseau, Subjective Quality Assessment-IVC Database 2005 [Online]. Available: http://www2.irccyn.ecnantes.fr/ivcdb

[27] Y. Horita, K. Shibata, Y. Kawayoke, and Z. M. P. Sazzad, MICT Image Quality Evaluation Database 2000 [Online]. Available:http://mict.eng.u-toyama.ac.jp/mict/index2.html

[28] D. M. Chandler and S. S. Hemami, A57 Database 2007 [Online]. Available:http://foulard.ece.cornell.edu/dmc27/vsnr/vsnr.html