A Novel Image Quality Assessment Metric using Singular Value Decomposition

Syed Salman Ali
University of Regina
syedsalmanali08@gmail.com

Abstract—**Image quality assessment (IQA) plays an important role in many applications such as image compression and transmission. In this paper a full referenced IQA (FR-IQA) model has been proposed which is based upon transformation based technique. Singular value decomposition (SVD) has been used to determine the basis vectors that best describe the input image signal. In contrast to other transformation based techniques such as discrete cosine transformation (DCT) and wavelet transform (WT), SVD does not use predefined basis vectors. In this paper a new methodology has been adopted in which both reference and distorted images are first combined together and then SVD is applied to compute the basis vectors. Projection coefficients of both reference and distorted images when projected onto these basis vectors have been used to calculate the final score. The proposed methodology has been tested on three publicly available image databases. The results of proposed methodology are better than most of the state of the art IQA metrics.**

Index Terms—**Image quality assessment (IQA), Singular value decomposition (SVD), TID2008, CSIQ image database.**

I. Introduction

An automatic image quality assessment (IQA) is of paramount importance in many image processing applications such as image acquisition, compression and transmission etc. An IQA model needs to be efficient enough to assess the quality of image just like humans evaluate it. The traditional IQA model that is mean square error (MSE) or peak signal to noise ratio (PSNR) does not produce results that are consistent with human subjective evaluation. A considerable research has been done in past to prove the deficiency of these simple but weak IQA models [1]. An IQA model can be categorized into three classes 1) Full reference model (FR), which uses original image along with distorted image to compute the score, 2) Reduced reference model, where limited information about original image is present and 3) No reference model where no information of original image is available.

In this paper a FR-IQA model has been proposed to assess the quality of images. FR-IQA models generally employs two approaches to assess the quality of images. The first approach is to design a bottom-up framework to model human visual system (HVS) to assess the quality in the same manner as humans do. This approach requires to develop mathematical model of each biological process that are part of our HVS. To build a bottom-up framework is not a trivial task as it requires complete understanding of biological and psychological processes involved in HVS. Unfortunately such sound and complete understanding of HVS has not been developed yet. Moreover the bottom-up framework becomes way too complex as number of features are added to it. This limitation makes it more difficult to be used in real-world applications. In second approach, a top-down framework is chosen which does not require the complete understanding of HVS. It simulates HVS using some solid assumptions. For instance Zhou wang proposed that structural information in image scene is a key feature and can be used only to model HVS. They proposed structural similarity (SSIM) [2] index, a popular IQA metric, which follows top-down approach. SSIM based model focuses to extract structural information from the images using mean and variance and computes the correlation between their structures. In addition other variants have also been proposed which follows the same paradigm [3]–[5]. Another state of the art framework is Non-Shift Edge Based Ratio (NSER) [6]. It is built upon the assumption that our HVS is tuned to use edge related information and IQA can be performed using this information alone. Some other frameworks that follow top-down approaches are Information Fidelity Criteria (IFC) [7] and Visual Information Fidelity (VIF) [8].

Structural information extraction is a crucial step in top-down approach as it greatly affects the accuracy of overall metric. In addition to above mentioned methodologies, different transformation based techniques such as Discrete Cosine Transform (DCT) [9] and Wavelet transform (WT) [10] have also been used to extract structural information. The goal behind is to transform the image from spatial domain into desired domain in which structural information can be easily extracted. These transformations uses predefined basis functions to transform the input images. These predefined basis functions are not necessarily able to capture structural information in every scenario. On the other hand Singular value decomposition (SVD) [11] provides cutting edge transformation technique that not only transform the image into desired domain but also computes the basis function from the data. In this paper, SVD based metric has been devised to assess the quality of images.

The rest of the paper is organized as follows: Section II presents the proposed SVD based metric. Section III presents the results. Section IV concludes the paper.

II. Methodology

The idea of using SVD in IQA models is not new [12]–[15]. The SVD based IQA models use reference images alone or both reference and distorted images independently to compute the singular values and basis vectors that best describe the

978-1-4799-6560-1/15/$31.00 ©2015 IEEE

input image. Then both reference and distorted images are projected onto these basis vectors to compute the projection coefficients. These coefficients describes the images in the new domain. The coefficients of both reference and distorted images are then used to compute the final score. However in proposed methodology a new paradigm has been opted. Both reference and distorted images are combined together and then singular values and basis vectors are computed by applying SVD on newly formed images.

A. Computation of Basis vector using SVD

The SVD can be applied to images in two ways: 1) Globally and 2) locally. In prior method SVD is applied on complete image without segmenting it into sub images. It is understood that noise is not invariant with respect to different regions of image. Therefore it is more suitable to first segment the image into sub images and then apply SVD on these sub images. The later method follows this routine. In proposed methodology SVD has been applied locally. Both reference and distorted images are first decomposed into m x m sub images. Mean subtraction is performed on each of the local sub image of reference and distorted image. As stated already, both reference and distorted images play their role in computation of basis vectors. Each mean subtracted sub image of reference image is combined side by side with the corresponding mean subtracted sub image of distorted image. It results into new sub image of m x n dimension (n = 2 x m) as shown in Fig. 1. For illustration purpose, global image has been shown in Fig. 1. Finally the SVD is applied on these new sub images.

Let X be the i^{th} new sub image where each column of X is considered as sample vector. The first m sample vectors are from the reference image whereas the last m sample vectors are from distorted images. When SVD is applied on X, it factorizes it into three matrices as mentioned in equation 1.

$$X = U \times S \times V^T \tag{1}$$

$$X = U \times P \tag{2}$$

In equation 1 U is m x m unitary matrix, S is m x n diagonal matrix and V is n x n unitary matrix. SVD transforms the input image into a new domain in which U and V can be viewed as rotational matrices and S can be viewed as scaling matrix. Both U and V are orthogonal matrices. Equation 2 is the refined form of equation 1 in which U matrix can be considered as basis matrix. Each column of U is regarded as m dimensional basis vector. Similarly P is m x n matrix which can be considered as projection matrix. Each column of P is regarded as projection coefficient vector of each sample vector of X. The first m columns of P are the projection coefficient vectors of sample vectors of reference sub image. Whereas the last m columns are the projection coefficient vectors of sample vectors of distorted sub image.

B. Image Quality Assessment

In order to assess the quality of image, projection vectors of each sample vector from reference sub images and corresponding projection vectors of sample vector from distorted sub images are used. The matrix P is decomposed into two matrices that is PR and PD in such a way that PR matrix consists of first m columns of P and PD matrix consists of last m columns of P. Let p_{rj} be the j^{th} projection coefficient vector of PR and p_{dj} be the corresponding j^{th} projection coefficient vector of PD. A scalar value d is computed based upon the difference between the l_3-norm of p_{rj} and p_{dj} which is as follow:

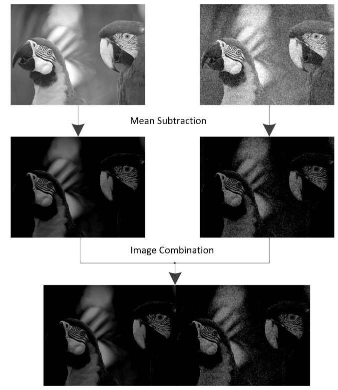

Fig. 1. Image combination

$$ds(j) = \sqrt[3]{\sum p_{rj}^3} - \sqrt[3]{\sum p_{dj}^3} \tag{3}$$

for $j = 1, ..., m$

In equation 3, $ds(j)$ represents the quality score of j^{th} sample vector of local distorted image with respect to corresponding sample vector of local reference image. In order to compute the score for i^{th} sub image, standard deviation based pooling scheme is applied to combine score for all sample vectors which is as follow:

$$dl(i) = \sqrt{\frac{1}{m} \sum_{j=1}^{m} (ds(j) - mds)^2} \tag{4}$$

Where

$$mds = \frac{1}{m} \sum_{j}^{m} ds(j) \tag{5}$$

In equation 4, $dl(i)$ is the quality score for i^{th} local sub image of distorted image with respect to corresponding local sub image of reference image. In order to compute the final score, an overall value is computed using average pooling scheme which uses scores of all local sub images. The final score is as follow:

$$score = \frac{1}{n} \sum_{i}^{n} dl(i) \tag{6}$$

III. Experimental Results

The proposed IQA metric has been evaluated on three publicly available databases: TID2008 database [16], CSIQ [17] and Cornell A57 [18]. The TID2008 comprises of 25 reference

TABLE I
COMPARISON IN TERMS OF SROCC

	PSNR	**UQI**	**NQM**	**VSNR**	**SSIM**	**VIF**	**IFC**	**MAD**	**Proposed**
TID2008	0.552	0.600	0.624	0.704	**0.773**	0.749	0.568	**0.770**	**0.793**
CSIQ	0.800	0.805	0.732	0.810	**0.858**	0.919	0.767	**0.899**	**0.872**
Cornell A57	0.598	0.426	0.798	**0.935**	0.806	0.622	0.318	**0.864**	**0.896**
Direct avg.	0.650	0.610	0.718	**0.816**	0.812	0.763	0.551	**0.844**	**0.854**
Weighted avg.	0.635	0.664	0.663	0.743	0.802	**0.803**	0.627	**0.815**	**0.821**

TABLE II
COMPARISON IN TERMS OF KRCC

	PSNR	**UQI**	**NQM**	**VSNR**	**SSIM**	**VIF**	**IFC**	**MAD**	**Proposed**
TID2008	0.402	0.435	0.461	0.534	**0.576**	0.586	0.424	**0.573**	**0.624**
CSIQ	0.608	0.613	0.591	0.624	0.691	**0.753**	0.590	**0.727**	**0.693**
Cornell A57	0.431	0.333	0.593	**0.803**	0.606	0.459	0.234	**0.670**	**0.723**
Direct avg.	0.480	0.460	0.548	**0.654**	0.624	0.599	0.416	**0.657**	**0.680**
Weighted avg.	0.470	0.492	0.507	0.569	0.615	**0.639**	0.475	**0.626**	**0.649**

images along with 1700 distorted images. Each reference image is distorted with 17 different distortion types including Additive White Noise (AWN) and JPEG compression noise etc. Details of each distortion type can be found at [16]. For each distortion type reference image is distorted using 4 levels of noise. Subjective score comprises of Mean Opinion Score (MOS) that reflects the human perception regarding each distorted image. The CSIQ database consists of 30 reference images. Each reference image is distorted with six types of distortion including AWN, JPEG, JPEG2000, Gaussian Blur (GB), additive pink Gaussian noise (PGN) and Global contrast decrements (CTD). The third database that is Cornell A57 has been developed by Cornell University. It consists of three reference image along with 54 distorted images. Each reference image has been distorted with six distortion types including AWN, JPEG, JPEG2000, GB, JPEG2000 along with dynamic contrast based quantization algorithm and quantization of LH sub bands of 5 levels of discrete wavelet transform keeping the step size such that root mean-squared (RMS) remains constant.

In order to compare the performance of subjective score with objective score, four evaluation measures are used. These are the Spearmans rank order correlation coefficient (SROCC), the Pearson linear correlation coefficient (PLCC), the Kendall rank correlation coefficient (KRCC) and RMS. In this paper SROCC and KRCC score of the proposed methodology has been compared with eight state-of-the-art IQA models: PSNR, UQI [1], NQM [19], VSNR [18], SSIM [2], VIF [8], IFC [7] and MAD [17]

A. Results

Table I shows the SROCC and KRCC score of five other state of the art IQA models along with the proposed methodology. In addition average values are also computed. The average values are computed for two cases: direct average and weighted average where weights are the size of respective database. Bold face font has been used for top 3 models. Table I shows that proposed methodology ranks among top three models but also out performs the other models in case of two database in terms of SROCC comparison. Similarly it is clear from Table II that proposed methodology ranks among top three models in terms of KRCC comparison.

In addition to these results, Table III and Table IV shows the SROCC based comparison of other IQA models with proposed methodology for each type of distortion in TID2008 and CSIQ image database respectively. The top three models have been shown in bold face font. In case of TID2008 database, proposed methodology ranks among top three IQA model for 14 different distortion types. Moreover the proposed methodology outperforms other IQA models for 10 types of distortion including AWN and JPEG. In case of CSIQ database the proposed methodology ranks among top three models for only two types of distortion. Overall the results of proposed methodology are comparable with other state-of-the-art IQA models.

IV. CONCLUSION

In this paper a new methodology based upon SVD for FR-IQA has been proposed. Both reference and distorted images have been used to compute the singular values and basis vectors. Projection coefficient vectors of both reference and distorted images are used to compute the final score. The proposed methodology can be further improved using more advanced pooling strategies.The proposed methodology has been compared with other cutting edge IQA models on TID2008, CSIQ and Cornell A57 image databases. In addition, results for each distortion type are also represented. The results show that proposed methodology is comparable with other IQA models.

REFERENCES

[1] Z. Wang and A. Bovik, “A universal image quality index,” *Signal Processing Letters, IEEE*, vol. 9, no. 3, pp. 81–84, March 2002.

TABLE III
DISTORTION-WISE PERFORMANCE COMPARISON OF SROCC ON TID2008 IMAGE DATABASE

	PSNR	UQI	NQM	VSNR	SSIM	VIF	IFC	MAD	Proposed
AWN	**0.907**	0.516	0.768	0.773	0.811	**0.880**	0.581	0.839	**0.932**
ANMC	**0.899**	0.458	0.749	0.779	0.803	**0.876**	0.546	0.826	**0.917**
SCN	**0.917**	0.536	0.772	0.766	0.815	**0.870**	0.596	0.868	**0.925**
MN	**0.852**	0.727	0.707	0.729	0.779	**0.868**	0.673	0.734	**0.815**
HFN	**0.927**	0.672	0.901	0.881	0.873	**0.907**	0.732	0.886	**0.928**
IMN	**0.872**	0.495	0.762	0.647	0.673	**0.833**	0.534	0.065	**0.816**
QN	**0.870**	0.561	0.821	0.827	**0.853**	0.797	0.586	0.816	**0.901**
GB	0.870	0.884	0.885	0.933	**0.954**	**0.954**	0.856	0.920	**0.941**
DEN	0.942	0.775	**0.945**	0.929	**0.953**	0.916	0.797	0.943	**0.961**
JPEG	0.872	0.770	0.907	0.917	**0.925**	0.917	0.818	**0.927**	**0.962**
JP2K	0.813	0.912	0.953	0.952	**0.962**	**0.971**	0.944	**0.971**	0.964
JGTE	0.752	0.835	0.737	0.805	**0.868**	0.859	0.791	**0.866**	**0.869**
J2TE	0.831	0.671	0.726	0.791	**0.858**	**0.850**	0.730	0.839	**0.914**
NEPN	0.581	0.740	0.680	0.572	0.711	**0.762**	**0.842**	**0.829**	0.730
Block	0.619	0.807	0.235	0.193	**0.846**	**0.832**	0.677	0.797	**0.871**
MS	**0.696**	0.562	0.525	0.371	**0.723**	0.510	0.425	0.516	**0.651**
CTC	**0.586**	0.520	**0.619**	0.424	0.525	**0.819**	0.171	0.272	0.531

TABLE IV
DISTORTION-WISE PERFORMANCE COMPARISON OF SROCC ON CSIQ IMAGE DATABASE

	PSNR	UQI	NQM	VSNR	SSIM	VIF	IFC	MAD	Proposed
AWN	0.936	0.760	0.938	0.925	0.897	**0.957**	0.843	**0.954**	**0.942**
JPEG	0.888	0.891	0.945	0.897	**0.954**	**0.970**	0.941	**0.961**	0.936
JP2K	0.936	0.878	0.955	0.948	**0.960**	**0.967**	0.925	**0.975**	0.953
PGN	0.934	0.716	0.911	0.908	0.892	**0.951**	0.826	**0.957**	**0.942**
GB	0.929	0.936	0.939	0.942	**0.961**	**0.974**	0.953	**0.968**	0.934
CTD	0.862	0.852	**0.938**	0.881	0.793	**0.934**	0.487	**0.921**	0.837

[2] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, no. 4, pp. 600–612, April 2004.

[3] M. Sampat, Z. Wang, S. Gupta, A. Bovik, and M. Markey, "Complex wavelet structural similarity: A new image similarity index," *Image Processing, IEEE Transactions on*, vol. 18, no. 11, pp. 2385–2401, 2009.

[4] Z. Wang, E. Simoncelli, and A. Bovik, "Multiscale structural similarity for image quality assessment," in *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Seventh Asilomar Conference on*, vol. 2, Nov 2003, pp. 1398–1402 Vol.2.

[5] G.-H. Chen, C.-L. Yang, and S.-L. Xie, "Gradient-based structural similarity for image quality assessment," in *Image Processing, 2006 IEEE International Conference on*, Oct 2006, pp. 2929–2932.

[6] M. Zhang, X. Mou, and D. Zhang, "Non-shift edge based ratio (nser): An image quality assessment metric based on early vision features," *Signal Processing Letters, IEEE*, vol. 18, no. 5, pp. 315–318, May 2011.

[7] H. Sheikh, A. Bovik, and G. de Veciana, "An information fidelity criterion for image quality assessment using natural scene statistics," *Image Processing, IEEE Transactions on*, vol. 14, no. 12, pp. 2117–2128, Dec 2005.

[8] H. Sheikh and A. Bovik, "Image information and visual quality," *Image Processing, IEEE Transactions on*, vol. 15, no. 2, pp. 430–444, 2006.

[9] N. Ahmed, T. Natarajan, and K. Rao, "Discrete cosine transform," *Computers, IEEE Transactions on*, vol. C-23, no. 1, pp. 90–93, 1974.

[10] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *Information Theory, IEEE Transactions on*, vol. 36, no. 5, pp. 961–1005, Sep 1990.

[11] G. H. Golub and C. Reinsch, "Singular value decomposition and least squares solutions," *Numerische Mathematik*, vol. 14, no. 5, pp. 403–420, 1970.

[12] M. Narwaria and W. Lin, "Svd-based quality metric for image and video using machine learning," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 42, no. 2, pp. 347–364, 2012.

[13] A. Saha, G. Bhatnagar, and Q. Wu, "Svd filter based multiscale approach for image quality assessment," in *Multimedia and Expo Workshops (ICMEW), 2012 IEEE International Conference on*, 2012, pp. 43–48.

[14] M. Esmaeilpour, A. Mansouri, and A. Mahmoudi-Aznaveh, "A new svd-based image quality assessment," in *Machine Vision and Image Processing (MVIP), 2013 8th Iranian Conference on*, Sept 2013, pp. 370–374.

[15] S. Wang, C. Deng, W. Lin, B. Zhao, and J. Chen, "A novel svd-based image quality assessment metric," in *Image Processing (ICIP), 2013 20th IEEE International Conference on*, Sept 2013, pp. 423–426.

[16] 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," *Advances of Modern Radioelectronics*, vol. 10, no. 4, pp. 30–45, 2009.

[17] E. C. Larson and D. M. Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy," *Journal of Electronic Imaging*, vol. 19, no. 1, p. 011006, 2010. [Online]. Available: http://link.aip.org/link/?JEI/19/011006/1

[18] D. Chandler and S. Hemami, "Vsnr: A wavelet-based visual signal-to-noise ratio for natural images," *Image Processing, IEEE Transactions on*, vol. 16, no. 9, pp. 2284–2298, Sept 2007.

[19] N. Damera-Venkata, T. Kite, W. Geisler, B. Evans, and A. Bovik, "Image quality assessment based on a degradation model," *Image Processing, IEEE Transactions on*, vol. 9, no. 4, pp. 636–650, Apr 2000.