Reduced-Reference Image Quality Assessment Based on DCT Subband Similarity

Amnon Balanov, Arik Schwartz, and Yair Moshe
Signal and Image Processing Laboratory (SIPL)
Andrew and Erna Viterbi Faculty of Electrical Engineering, Technion – Israel Institute of Technology
Technion City, Haifa, 32000, Israel
yair@ee.technion.ac.il, http://sipl.technion.ac.il/

***Abstract*—Reduced-reference image quality measures aim to estimate the visual quality of a distorted image with only partial information about the "perfect quality" reference image. In this paper, we present a reduced-reference image quality assessment (IQA) metric based on DCT Subbands Similarity (RR-DSS). According to the assumption that human visual perception is adapted for extracting structural information, the proposed technique measures change in structural information in subbands in the discrete cosine transform (DCT) domain and weights the quality estimates for these subbands. RR-DSS is simple to implement, incurs low computational complexity, and has a flexible tradeoff between the amount of side information and image quality estimation accuracy. RR-DSS was tested with public image databases and shows excellent correlation with human judgments of quality. It outperforms state-of-the-art RR IQA techniques and even several FR IQA techniques.**

Keywords— image quality assessment (IQA), reduced-reference quality assessment, objective quality measure, Discrete Cosine Transform (DCT).

I. INTRODUCTION

Image quality assessment (IQA) becomes increasingly important due to the many applications involving digital imaging and communication. Most visual media undergoes processing and transmission that introduce distortions and degrade the perceived image quality for the end human observer. It is desirable to evaluate the quality degradation introduced by these distortions in order to benchmark image processing systems and algorithms, for quality control systems or for optimizing parameter values of image processing algorithms. The most reliable evaluation method is by subjective human interpretation and evaluation. However, this method is costly, time-consuming, and in many cases, not practical. Therefore, it is desired to look for an automatic tool that will produce a quantitative assessment that correlates well with human judgment of the quality. There are three approaches for evaluating image quality: full-reference (FR), no-reference (NR) and reduced-reference (RR). The FR approach is the most common, meaning that a reference image is assumed to be known. In many practical applications, however, the reference image is not available, and a no-reference quality assessment approach is desirable. Sometimes it is possible to extract a small amount of side information about the reference image in the form of a set of features made available to help evaluate the quality of the distorted image. In this reduced-reference approach, better results than in the no-reference approach can usually be obtained.

In the FR approach, the peak signal-to-noise ratio (PSNR) has been commonly used for many years but can be a poor predictor of perceived visual quality [1]. Due to the limitations of PSNR, much effort has been spent on designing better visual quality metrics. A popular replacement for PSNR is the structural similarity (SSIM) index [2]. The principal philosophy underlying the SSIM index is that the human visual system (HVS) is adapted to extract structural information from visual scenes. Therefore, the algorithm accounts for distortion in image structure. It computes the mean, standard deviation and covariance of small patches inside an image and combines the measurements into a distortion map. These three terms measure luminance distortion, contrast distortion and the loss of correlation, respectively. A more advanced IQA technique based on the SSIM is the multiscale SSIM (MS-SSIM) [3]. The MS-SSIM index considers the viewing-distance and accounts for the multiscale nature of the HVS. This index calculates multiple SSIM values at multiple image scales and aggregates them. Another FR IQA technique that accounts for structural similarity, and has been proposed by the authors, is DSS [4]. This technique measures the change in structural information in subbands in the discrete cosine transform domain. The quality estimates of each subband are weighted, resulting in a final quality score. For more detailed information on FR IQA techniques, the reader is referred to [5, 6]. As this paper deals with RR IQA, we will now concentrate on recent RR techniques.

When designing an RR quality metric, we need to consider not only its correlation with human judgments of quality but also the amount of side information for representing the extracted features. The extracted features should be relevant to the HVS perception of image quality and sensitive to a variety of image distortions. At the same time, they should not be large, as the extracted features need to be transmitted to the receiver side for the quality analysis. In [7], perceptually motivated features based on models of low-level vision are extracted to provide a reduced description of the image. This method has demonstrated good performance for JPEG and JPEG2000 compression. A natural scene statistics approach for RR IQA is presented in [8]. This approach assumes that the HVS is adapted to the natural visual environment and uses the departure from natural image statistics as a measure of perceptual quality. According to the same approach, a method that utilizes a steerable pyramid for a

wavelet decomposition is proposed in [9]. This wavelet-domain natural image statistic metric (WNISM) models the marginal distribution of the wavelet coefficients of a natural image in each subband by a generalized Gaussian density (GGD) function. The Kullback–Leibler distance (KLD) is used to measure the distribution difference. To avoid the problem of sending the features separate from the reference over the channel, the concept of quality-aware images was proposed in [10]. In quality-aware images, partial reference image information, such as the one proposed in [9], is embedded within the image and can be reliably extracted despite distortions. Therefore, no overhead is introduced and no side information need to be transmitted. The model proposed in [9] was further improved in [11] by employing a nonlinear divisive normalization transform (DNT) after the linear wavelet decomposition. Four features are computed from each wavelet subband: the KLD, and the variance, kurtosis, and skewness of the DNT coefficients from that band. In [12], a multiscale geometric analysis is applied to extract features of the distribution of curvelets, bandlets, wavelets, and contourlets. This method also accounts for the HVS contract sensitivity function (CSF) and for Weber's law of just noticeable difference (JND).

Not all methods for RR IQA are explicitly based on natural scene statistics. A set of reduced reference entropic differencing (RRED) algorithms for IQA based on information theory are proposed in [13]. For those algorithms, the differences between the entropies of wavelet coefficients of the reference and distorted image are measured in order to quantify the image information change, which can indicate the image perceptual quality. A computational intelligence approach for RR IQA using circular extreme learning machine (C-ELM) is presented in [14]. It aims to mimic quality perception instead of designing an explicit model of the HVS. First, a block-level and image-level representation of the visual signal is defined. Then, a learning machine handles the mapping of the feature vector into quality scores.

Several RR IQA techniques follow the FR SSIM index in considering structural changes in images. The RR-SSIM [15] estimates the SSIM index of the distorted image by extracting statistical features from a multi-scale, multi-orientation divisive normalization transform. A fast RR approach for quality assessment of color images by modifying the SSIM index is suggested in [16]. Images are decomposed in the wavelet domain for extracting statistical features. For each wavelet subband, the divergence and standard deviation are extracted. These features are then used to obtain a structural dissimilarity measure sensitive to structural distortions.

As previously mentioned, statistical properties of natural images which are sensitive to the introduced distortions can be employed to develop IQA methods. DCT has been widely adopted and employed for various image and video processing tasks. Therefore, if we can extract features relevant to the HVS based on the DCT coefficients, they can be utilized as the RR features for quality assessment of images. This approach has been applied in [17, 18] by reorganizing the block-based DCT coefficients of an image into a three-level coefficient tree. A generalized Gaussian density is employed to model the coefficient distribution of each reorganized DCT subband. The city-block distance is used to measure the difference between the corresponding reorganized DCT subbands. The visual quality index of the image is obtained by pooling the distances together.

In this paper we follow the same approach and suggest an RR IQA technique that extracts features in block-based DCT subbands. Like in [17, 18], we measure the distances between corresponding DCT subbands and obtain a quality index by pooling those distances together. However, we do not reorganize the coefficients and we extract different features. These features consider structural similarity in each subband by measuring the change of local variance in each subband. The suggested RR DCT subbands similarity (RR-DSS) technique is based on our FR DCT subbands similarity (DSS) technique [4].

The remaining paper is structured as follows. In section II we describe DSS and in the section III we describe the modification required for RR-DSS. Section IV presents and discusses the obtained results. Finally, conclusions are drawn in Section V.

II. Full-Reference DCT Subband Similarity

A well-known property of natural images is that their image power spectra tend to fall off with increasing spatial frequency. The distribution of DCT image coefficients has been approximated by a generalized Gaussian distribution [17-19] and the influence of different distortions on the distribution of these coefficients was analyzed [19]. An insight of this analysis is that one of the statistical features that differentiates the DCT coefficient distributions of distorted and undistorted images is their variance. This property is exploited in the quality metric described below.

The response of the HVS to changes in intensity of the illumination is known to be nonlinear. Due to the frequency and texture masking phenomena (reduction in the visibility of one image component by the presence of another), the HVS is more sensitive to changes in low spatial frequency regions than to changes in high spatial frequency regions [20]. Therefore, when evaluating image quality, it is beneficial to give more weight to distortions in low frequencies. In addition, many modern image processing tasks process each spatial frequency separately and may introduce artifacts with different strength into each subband. Measuring image quality in each spatial frequency subband fits naturally the scheme of those image processing algorithms, so it has the potential to estimate accurately the effect of artifacts introduced by them on the HVS. In the proposed IQA technique, we estimate quality in the DCT domain. The main motivation behind quality assessment in the DCT domain is the observation that the statistics of DCT coefficients change with the degree and type of image distortion [19]. The 2D DCT transform is a natural choice in this case since many image and video processing tasks, such as coding, deblocking, denoising, etc., are based on a block-based DCT transform.

A high-level diagram of the FR DCT subbands similarity (DSS) image quality assessment scheme proposed in [4] is depicted in Fig. 1. First, channel decomposition is performed for both the reference and distorted images by separating each image into subbands that correspond to different 2D DCT spatial frequencies. After channel decomposition, the similarity between each subband in the reference image and its counterpart

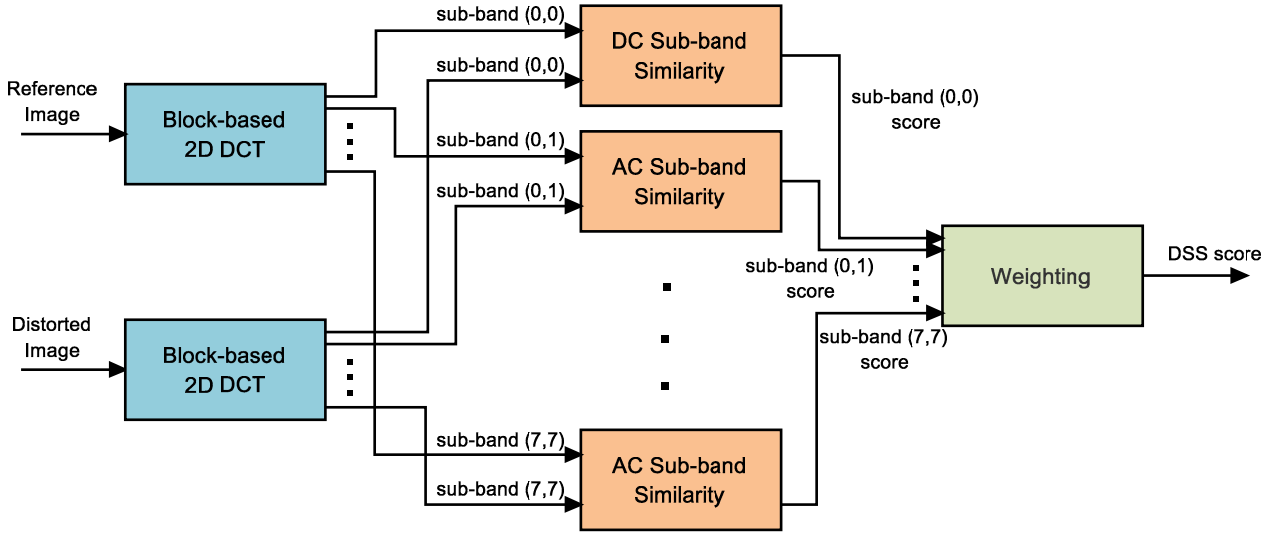

Fig. 1. DCT Sub-bands Similarity (DSS) image quality assessment according to [4]. A similar scheme is used in the proposed RR-DSS technique.

subband in the distorted image is computed, resulting in a subband similarity score. Finally, subband similarity scores are weighted resulting in a scalar DSS quality score.

In the first stage of DSS computation, both images are decomposed into spectral subbands. Each image is divided into small nonoverlapping rectangular blocks of size 8x8 and a 2D DCT transform is applied to each block. The resulting DCT coefficients are decomposed into 64 subbands. The DCT transform is performed locally in the spatial domain in accordance with the fact that the HVS processes the visual space locally. We denote by $X_{m,n}(p,q)$ the coefficient at location (p,q) in the subband (m,n), $0 \le m,n \le 7$, of the 2D DCT transform of the image x. Note that each subband matrix $X_{m,n}$ is 64 times smaller than the image x so $0 \le p < P/8$, $0 \le q < Q/8$ for an image of size $P \times Q$.

After subband decomposition, the similarity between the coefficients of the reference image and the distorted image is computed for each subband. The aim of this stage is to measure the amount of local change in the statistics of coefficients due to various distortions. We compute the local variance $\sigma_{m,n}(p,q)^2$ for patches of size $k \times k$ in each subband (related to pixel patches of size $8k \times 8k$). We use $k = 3$ and measure the change in these variances as follows:

$$DSS_{m,n}(p,q) = \frac{2\sigma^X_{m,n}(p,q)\sigma^Y_{m,n}(p,q) + C}{\sigma^X_{m,n}(p,q)^2 + \sigma^Y_{m,n}(p,q)^2 + C} \quad , \qquad (1)$$

where $\sigma^X_{m,n}(p,q)^2$ is the local variance at location (p,q) in subband (m,n) of the reference image, $\sigma^Y_{m,n}(p,q)^2$ is the local variance at location (p,q) in subband (m,n) of the distorted image, and C is a constant added for numerical stability. Subband similarity score computed according to (1) is normalized so it is ≤ 1. The power spectra of natural images tend to fall off with increasing spatial frequency [19], thus subband variances are smaller for higher spatial frequencies. Normalization allows weighting subband scores while ignoring this fall off. Typical values for the constant C are $100 \le C \le 1000$. The results are robust to the selection of C.

For a natural image, there is a high correlation between adjacent pixels. The correlation between adjacent pixels generates image structure that is important for the perception of image quality. The distortion in image structure can be measured by cross-correlation between pixels of the reference and distorted image [2]. In the DSS IQA metric, we measure the change in image structure by the change in the autocorrelation function of counterpart patches of these images. The autocorrelation function has been suggested as the basis of a texture measure [21] since the locations of peaks in the autocorrelation function relate to the shape and arrangement of texture primitives in the image. Thus, the autocorrelation function captures local image structure. According to the Wiener–Khinchin theorem, the autocorrelation function equals to the inverse Fourier transform of the power spectrum of the image. A similar result can be shown to exist also for the DCT transform. As we decompose the image into spectral subbands, we would like to implicitly consider the autocorrelation of neighboring pixels in these subbands. It can be shown that computing the local variance of a subband of 2D-DCT coefficients is related to computing the autocorrelation of an image patch in the pixel domain. Thus, (1) measures the change in the local autocorrelation function of the image. This is equivalent to measuring image structure of the distorted image relative to the structure of the reference image.

A meticulous reader may notice the similarity between (1) and the equation for computing the contrast component in the SSIM index [2]. In SSIM, the computation is performed on pixels and accounts for the contrast distortion of an image. An important feature of this function is that with the same amount of contrast change $\Delta\sigma = \sigma_x - \sigma_y$, this measure is less sensitive to high base contrast σ_x than to low base contrast. This is consistent with the contrast masking feature of the HVS. In DSS however, the computation is performed on DCT coefficients in a specific subband and accounts for the change of variance of these coefficients. The measure is less sensitive to the same amount of change in variance for high base variance. The base variance is high when there are large changes in the local pixel

region. This is related to high contrast so the masking feature of the HVS is considered here as well.

While AC coefficients give an indication of the amplitude and orientation of edges within an image, the DC coefficient is proportional to the average image amplitude. Hence, it is required to compute the similarity between DC subbands in a way that takes into account also other properties in the DC subbands. It was found that the cross-correlation between the two DC subband images (of the reference and distorted images) plays an important role in estimating the perceived image quality. Thus, we compute the similarity score between the reference and distorted DC subbands as follows:

$$DSS_{0,0}(p,q) = \frac{2\sigma_{0,0}^{X}(p,q)\sigma_{0,0}^{Y}(p,q) + C}{\sigma_{0,0}^{X}(p,q)^2 + \sigma_{0,0}^{Y}(p,q)^2 + C} \cdot \frac{\sigma_{0,0}^{XY}(p,q) + C}{\sigma_{0,0}^{X}(p,q)\sigma_{0,0}^{Y}(p,q) + C} \quad , \tag{2}$$

where $\sigma_{0,0}^{XY}(p,q)$ is the cross-correlation at location (p,q) in the DC subband between the reference and distorted image. The first term in the equation is similar to (1) and measures the change is the local variance while the second term is the local Pearson cross-correlation metric. Like in (1), C is a constant added for numerical stability.

The similarity scores in (1) and (2) are computed point-wise resulting in a similarity score for each coefficient. We would like to use these scores in order to obtain a scalar similarity score for each subband. Humans tend to perceive "poor" regions in an image with more severity than "good" ones, and hence penalize heavily images with even a small number of "poor" regions. Namely, worst quality regions in an image dominate human perception of image quality. This effect can be taken into account by weighting more heavily quality scores from lower scoring regions. One of the simplest and best performing strategies for achieving this goal is percentile scoring [22]. In this pooling technique, the scalar similarity score is the mean of only the lowest w% percentiles of the scores. A typical value of w is 5. We apply percentile scoring to each subband separately, getting 64 subband similarity scores $DSS_{m,n}, 0 \le m,n \le 7$.

After evaluating the distortion in each DCT subband, the 64 similarity scores are weighted to obtain a scalar score:

$$DSS = \sum_{m,n=0}^{7} w_{m,n} DSS_{m,n} \quad , \tag{3}$$

where $w_{m,n}$ is the weight of the score of subband (m,n). As previously mentioned, an important and well-modeled characteristic of the HVS is its higher sensitivity to distortions in low spatial frequencies [20]. It is desired to select a weighting function $f(m,n) = w_{m,n}$ that takes this characteristic into account by giving higher weights to subbands with lower spatial frequencies. Several weighting functions that satisfy this property were tested and a Gaussian weighting function was empirically selected. The standard deviation of the Gaussian determines the proportion between the weight given to low spatial frequencies and to high spatial frequencies. A typical value we use for the standard deviation is $\sigma = \sqrt{6}$.

III. Reduced-Reference DCT Subband Similarity

The full-reference DSS metric described in Section II shows excellent results in estimating human judgment of image quality [4]. This metric is the basis for the RR we suggest in this paper and describe in this section. RR-DSS uses the same scheme as DSS, depicted in Fig. 1. However, in order to reduce drastically the amount of information required about the reference image, we perform three modifications in DSS.

First, we use only a few lowest frequency subbbands for the quality assessment. In DSS, subband similarity scores $DSS_{m,n}$ are weighted with a Gaussian weighting function according to (3). The standard deviation of the Gaussian weighting function is selected empirically according to the HVS. Weights are much larger for low frequency subbands than for mid and high frequency subbands. For example, $w_{0,0}$ is two orders of magnitude larger than $w_{2,2}$ and four orders of magnitude larger than $w_{4,4}$. So, many mid and high frequency subband scores can be neglected with only a minor change in quality assessment results. We sort the subbands by a descending order of their weights $w_{m,n}$ and truncate the list after s subbands. Typical number of subbands s adequate to maintain good IQA results is $3 \le s \le 10$. Since the image is decomposed into 64 DCT subbands, this results in up to one order of magnitude reduction in information size compared with DSS.

A second modification for reducing the amount of information about the reference image is spatial down-sampling. DSS computes the local variance $\sigma_{m,n}(p,q)^2$ for each coefficient in each subband. Since natural images are mostly smooth, the values of these block-based subband coefficients tend to correlate well with spatially neighboring coefficients. Moreover, typical distortions in images, such as blockization, blur, ringing, etc., usually spread over the image. This allows us to uniformly down-sample the local variances while usually not missing important image distortions. Spatial down-sampling coincides with previous results that many image quality metrics can be computed using images decimated by a factor of 6-8 without a substantial degradation in performance [23]. It was also observed in [23] that the effects of down-sampling are generally milder than the effects of filtering. This result implies that aliasing is perhaps beneficial to the assessment process. For DSS, we down-sample the coefficient variances to size $r \times r$ with typical values of $3 \le r \le 10$. These values of r result in only a minor degradation in assessment performance with a large saving in the amount of side information. As an example, for an image of size 640×480, we have 80×60 coefficients in each subband. If we select r=5, we sample only 5×5=25 coefficients variances in each subband so we reduce the amount of side information by two orders of magnitude.

The third modification in RR-DSS compared with DSS is for the subband similarity score of the DC subband $DSS_{0,0}$. In DSS, the similarity score between the reference and distorted DC subbands is computed according to (2). This equation involves computation of the local cross-correlation $\sigma_{0,0}^{XY}(p,q)$ between the DC subband of the reference image and the DC subband of distorted image. For RR-DSS, it is not possible to compute the cross-correlation due to spatial down-sampling. So, in RR-DSS, (1) is used for computing a local similarity score for both the DC and AC coefficients.

TABLE 1. Performance of the proposed RR-DSS and other FR and RR image quality assessment metrics on images from the LIVE, CSIQ, and TID databases. RR-DSS[1] indicated RR-DSS with s=3, r=3. RR-DSS[2] indicates RR-DSS with s=6, r=10.

		PSNR	SSIM	MSSSIM	DSS	WNISM	RRED	RDCT	RRDSS[1]	RRDSS[2]
SROCC	LIVE	0.876	0.910	0.944	0.970	0.750	0.943	0.928	0.924	0.959
	CSIQ	0.806	0.837	0.914	0.953	0.705	0.909	0.804	0.901	0.945
	TID	0.525	0.645	0.853	0.891	0.520	0.823	-	0.710	0.850
LCC	LIVE	0.870	0.938	0.933	0.968	0.738	0.938	0.931	0.925	0.959
	CSIQ	0.800	0.815	0.897	0.950	0.699	0.903	0.820	0.898	0.944
	TID	0.536	0.652	0.839	0.897	0.597	0.825	-	0.750	0.860
RMSE	LIVE	13.368	11.790	8.946	6.830	18.43	9.430	9.965	10.39	7.720
	CSIQ	0.158	0.133	0.115	0.078	0.188	0.112	0.150	0.115	0.086
	TID	1.137	0.851	0.730	0.593	1.070	0.756	-	0.889	0.680
# of scalars		-	-	-	-	18	>5420	30	27	600
running time [ms]		5	95	270	200	2500	900	-	7	22

TABLE 2. Influence of the number of low frequency subbands s on the performance of RR-DSS. For all results, the size of the sampled subbands $r \times r$ is r=4.

		s=3	s=4	s=6	s=8	s=10
SROCC	LIVE	0.938	0.943	0.953	0.955	0.958
	CSIQ	0.920	0.926	0.936	0.938	0.940
	TID	0.796	0.784	0.800	0.818	0.829
LCC	LIVE	0.937	0.944	0.952	0.954	0.956
	CSIQ	0.921	0.925	0.928	0.936	0.937
	TID	0.774	0.808	0.822	0.835	0.845
# of scalars		48	64	96	128	160

TABLE 3. Influence of the size of the sampled subbands $r \times r$ on the performance of RR-DSS. For all results, the number of low frequency subbands s=6.

		r=3	r=5	r=6	r=8	r=10
SROCC	LIVE	0.942	0.953	0.957	0.957	0.959
	CSIQ	0.918	0.937	0.940	0.947	0.948
	TID	0.778	0.830	0.842	0.856	0.850
LCC	LIVE	0.942	0.953	0.957	0.957	0.959
	CSIQ	0.927	0.939	0.941	0.947	0.947
	TID	0.730	0.807	0.827	0.843	0.860
# of scalars		54	150	216	384	600

IV. RESULTS

In order to examine the performance of the proposed image quality assessment metric, we have evaluated it against subjective results on three public image databases – the LIVE Image Quality Assessment Database [24], the Tampere Image Database (TID) [25], and the Categorical Subjective Image Quality (CSIQ) Database [26]. Each database contains hundreds of distorted images and subjective results of few dozen subjects. We use the Spearman rank-order correlation coefficient (SROCC), the Pearson linear correlation coefficient (LCC) and the Root Mean Square Error (RMSE), between the subjective scores and the quality indices to compare the relative performances between the proposed technique and other state-of-the-art FR and RR image quality assessment techniques. The SROCC serves as a measure of prediction monotonicity while the LCC and RMSE serve as measures of prediction accuracy. A better objective IQA measure should have higher SROCC, higher LCC, and lower RMSE. To evaluate the performance of a particular quality assessment technique, we have applied a 5-parameter nonlinear mapping recommended in [27].

TABLE 1 lists the SROCC, LCC and RMSE performance on the three image databases. Four FR IQA metrics - PSNR, SSIM [2], MS-SSIM [3] and DSS [4], and four RR IQA metrics – WNSIM [9], RRED [13], RDCT [18] and RD-DSS, were examined. To obtain high correlation with subjective results, RRED was tested in a configuration that requires a large amount of side information. RR-DSS was tested in two different configurations. RR-DSS[1] indicates s=3, r=3 and RR-DSS[2] indicates s=6, r=10. For the three databases and for the three evaluation criteria, the proposed metric in the RR-DSS[2] configuration yields, on average, very high correlation with subjective results. Its results precede most other metrics examined, both RR and FR, including SSIM and MS-SSIM, and is preceded only by DSS. When evaluating an RR quality metric, we should consider not only its correlation with human judgments of quality but also the amount of side information for representing the extracted features. We see in TABLE 1 that the RR-DSS[1] configuration uses more scalars than WNSIM and substantially outperforms it. It uses slightly fewer scalars than RDCT and slightly outperforms it too (substantially better performance than RDCT for the CSIQ database and slightly worse performance than RDCT for the LIVE database). The performance of RR-DSS[1] is even close to the performance of RRED although RRED uses two orders of magnitude more scalars. Note that the number of scalars used by RRED depends on image resolution so the table indicates a lower bound. The table also lists execution time of each of the metrics in MATLAB on Intel Core i5 processor running at 2.67GHz. RR-DSS has a low computational complexity and is slower only than PSNR. It is two or three orders of magnitude faster than WNSIM and RRED respectively.

We have performed a set of tests to measure the parameter selection tradeoff of RR-DSS. TABLE 2 lists the influence of the number of low frequency subbands s on the performance of RR-DSS with a fixed size of the sampled subbands r=4. As expected, for the three databases and the two evaluation criteria, performance degrades gradually with the removal of more subbands. Note that even with s=3, results are better than PSNR and SSIM (can be seen in TABLE 1). The results of another set of tests for measuring influence of the size of the sampled subbands $r \times r$ on the performance of RR-DSS are listed in TABLE 3. For all results, the number of low frequency subbands

is fixed s=6. Here, again, we see a gradual degradation in performance with the reduction in size of sampled coefficients variances. Even with r=3, results are better than PSNR, SSIM, and on most of the times also MS-SSIM.

The last test we have performed was to select the size of the samples subbands r adaptively. As low frequency subband scores are given large weight values $w_{m,n}$, we will assign them with a larger r. As an example of the efficiency of this approach, we used a fixed number of low frequency subbands s=6 and selected r=6 for the DC subband (0,0) , r=4 for the two lowest frequency AC subbands (1,0) and (0,1), and r=3 for the other 3 subbands. The results we got for this configuration are SROCC equals 0.955, 0.933, and 0.814 for the LIVE, CSIQ and TID databases respectively, and LCC equals 0.955, 0.933, and 0.829 for the same databases respectively. The number of scalars used as a side information is 95 so the amount of side information is somewhere between r=3 (54 scalars) and r=5 (150 scalars) in TABLE 3. However, the resulting performance are comparable to r=5 in this table. We got similar performance to a configuration with r=5 but with a smaller number of scalars. Thus, adaptive sampling size selection can improve the quality assessment performance of RR-DSS.

V. Conclusions

In this paper we have suggested a reduced-reference image quality assessment metric based on DCT subbands similarity (RR-DSS). The proposed metric measures change in structural information in subbands in the discrete cosine transform (DCT) domain and weights the quality estimates for these subbands. RR-DSS shows excellent correlation with human judgments of image quality, outperforming state-of-the-art RR IQA techniques and even several FR IQA techniques. It incurs a low computational complexity and has a flexible tradeoff between the amount of side information and image quality estimation accuracy.

Acknowledgment

The authors would like to thank Prof. David Malah, head of SIPL, for his advice and helpful comments.

References

[1] Z. Wang and A. C. Bovik, "Mean squared error: love it or leave it? A new look at signal fidelity measures," *Signal Processing Magazine, IEEE,* vol. 26, pp. 98-117, 2009.

[2] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *Image Processing, IEEE Transactions on,* vol. 13, pp. 600-612, 2004.

[3] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," presented at the Signals, Systems and Computers, 37th Asilomar Conference on, 2003.

[4] A. Balanov, A. Schwartz, Y. Moshe, and N. Peleg, "Image quality assessment based on DCT subband similarity," in *Image Processing (ICIP), 2015 IEEE International Conference on*, 2015, pp. 2105-2109.

[5] W. Lin and C. C. Jay Kuo, "Perceptual visual quality metrics: A survey," *Journal of Visual Communication and Image Representation,* vol. 22, pp. 297-312, 2011.

[6] S. Chikkerur, V. Sundaram, M. Reisslein, and L. J. Karam, "Objective video quality assessment methods: A classification, review, and performance comparison," *Broadcasting, IEEE Transactions on,* vol. 57, pp. 165-182, 2011.

[7] M. Carnec, P. L. Callet, and D. Barba, "An image quality assessment method based on perception of structural information," in *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, 2003, pp. III-185-8 vol. 2.

[8] Z. Wang and A. C. Bovik, "Reduced-and no-reference image quality assessment," *Signal Processing Magazine, IEEE,* vol. 28, pp. 29-40, 2011.

[9] Z. Wang and E. P. Simoncelli, "Reduced-reference image quality assessment using a wavelet-domain natural image statistic model," in *Electronic Imaging 2005*, 2005, pp. 149-159.

[10] 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, pp. 1680-1689, 2006.

[11] Q. Li and Z. Wang, "Reduced-reference image quality assessment using divisive normalization-based image representation," *Selected Topics in Signal Processing, IEEE Journal of,* vol. 3, pp. 202-211, 2009.

[12] X. Gao, W. Lu, D. Tao, and X. Li, "Image quality assessment based on multiscale geometric analysis," *Image Processing, IEEE Transactions on,* vol. 18, pp. 1409-1423, 2009.

[13] R. Soundararajan and A. C. Bovik, "RRED indices: Reduced reference entropic differencing for image quality assessment," *Image Processing, IEEE Transactions on,* vol. 21, pp. 517-526, 2012.

[14] S. Decherchi, P. Gastaldo, R. Zunino, E. Cambria, and J. Redi, "Circular-ELM for the reduced-reference assessment of perceived image quality," *Neurocomputing,* vol. 102, pp. 78-89, 2013.

[15] A. Rehman and Z. Wang, "Reduced-reference image quality assessment by structural similarity estimation," *Image Processing, IEEE Transactions on,* vol. 21, pp. 3378-3389, 2012.

[16] V. Bhateja, A. Srivastava, and A. Kalsi, "Fast SSIM index for color images employing reduced-reference evaluation," in *Proceedings of the International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA) 2013*, 2014, pp. 451-458.

[17] L. Ma, S. Li, F. Zhang, and K. N. Ngan, "Reduced-reference image quality assessment using reorganized DCT-based image representation," *Multimedia, IEEE Transactions on,* vol. 13, pp. 824-829, 2011.

[18] L. Ma, S. Li, and K. N. Ngan, "Reduced-reference image quality assessment in reorganized DCT domain," *Signal Processing: Image Communication,* vol. 28, pp. 884-902, 2013.

[19] M. A. Saad, A. C. Bovik, and C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the DCT domain," *Image Processing, IEEE Transactions on,* vol. 21, pp. 3339-3352, 2012.

[20] W. K. Pratt, "Digital Image Processing: PIKS Scientific Inside," ed: Wiley-Interscience, 2007.

[21] H.-C. Lin, L.-L. Wang, and S.-N. Yang, "Extracting periodicity of a regular texture based on autocorrelation functions," *Pattern Recognition Letters,* vol. 18, pp. 433-443, 1997.

[22] A. K. Moorthy and A. C. Bovik, "Visual importance pooling for image quality assessment," *Selected Topics in Signal Processing, IEEE Journal of,* vol. 3, pp. 193-201, 2009.

[23] M. D. Gaubatz and S. S. Hemami, "On the nearly scale-independent rank behavior of image quality metrics," in *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, 2008, pp. 701-704.

[24] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, "LIVE image quality assessment database release 2," ed, 2005.

[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," *Advances of Modern Radioelectronics,* vol. 10, pp. 30-45, 2009.

[26] 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, pp. 011006-011006-21, 2010.

[27] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *Image Processing, IEEE Transactions on,* vol. 15, pp. 3440-3451, 2006.