Image Quality Assessment with Global and Local Image Regions for Visual Quality Perception

[1]Syed Shanawaz, [2]G. Amjad Khan

[1]M. Tech (CSP), [2]Assistant Professor (Ph.D)
[1,2]Department of Electronics and Communication Engineering,
[1,2]G. Pulla Reddy Engineering College (Autonomous), Kurnool, India

Abstract: Image Quality Assessment (IQA) by using mathematical methods is offering favorable results in calculating visual quality of distorted images. These methods are developed by examining effective features that are compatible with characteristics of Human Visual System (HVS). But many of those methods are difficult to apply for optimization problems. This paper presents DCT based metric with easy implementation and having mathematical properties like differentiability, convexity and valid distance metricability to overcome the optimization problems. By using this method we are able to calculate the quality of image as a whole as well as the quality of local image regions.

***IndexTerms* - Human Visual System, Image Quality Assessment, Image quality optimization problem.**

I. Introduction

Image quality is a characteristic of the image that tells about how much amount of image is degraded due to different distortions like noise, blurring, fading. Image degradation occurs during image acquisition, transmission, compression, decompression, printing etc. The main aim of many image processing based applications is to improve the visual quality of the image or to reduce the distortions that are present in the image. In this regard lot of work has been going on for developing mathematical Image Quality Assessment (IQA) methods whose results should be compatible with perceived visual quality.

The quality of the image is obtained either by using subjective methods or by using objective methods. In subjective way of quality assessment human observers will rate the image quality. The distorted image is given to image processing experts and they will decide the quality of the image by using Mean Opinion Score (MOS), which is the subjective quality assessment method. In objective methods the image quality is obtained by using computations or algorithms.

Depending on the availability of ideal image or reference image information at the time of quality assessment the objective image quality metrics are classified into three types: i) Full-Reference IQA , ii) Reduced-Reference IQA and iii) No-Reference IQA. In Full-Reference IQA (FR-IQA) the distorted image is compared with original or reference image for obtaining the quality of the image. In Reduced-Reference IQA (RR-IQA) only partial amount of original image is available at the time of assessment to evaluate the quality of degraded image. In No-Reference IQA (NR-IQA) only the distorted image is available and original image is

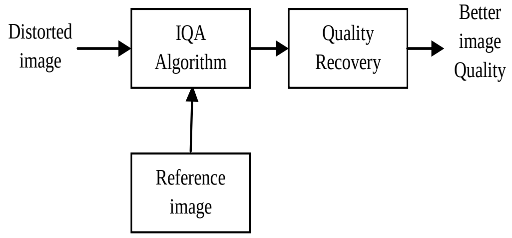


Fig. 1. Full-Reference Image Quality Assessment (FR-IQA)

not available. This paper follows FR-IQA method which is illustrated in Fig.1.

The quality score (Z) of distorted image(Y) for subjective full- reference IQA model in the presence of reference image(X) is given by

$$Z_i = Qo\, F(X_i, Y_i |\, W) + \varepsilon_i \qquad (1)$$

where X_i, Y_i represents the i-th reference and its distorted image. Z_i represents the subjective visual quality score for Y_i. Prediction error is represented by ε_i. F(.) is a full- reference IQA method and W is a set of model parameters. Q(.) is a non linear and a non degenerate monotonic function and is used to map the output of F(.) to Z_i.

One of the most widely used distortion metric is mean squared error (MSE) [1] which serves both finding image quality and optimization of image quality algorithms. It is easy to compute MSE and it also has many useful mathematical properties. Even though it has useful properties like valid distance metricability, differentiability and convexity, the problem is its poor correlation with perceived subjective visual qualities for distortions.

In order to overcome this problem many FRIQA methods [1]-[14] like Structural SIMilarity (SSIM) index [2], Feature SIMilarity (FSIM) index [6], Visual Saliency induced index (VSI) [9] etc., have been proposed. Structural Contrast-Quality Index (SC-QI) and Structural Contrast-Distortion Metric (SC-DM) [10] were also proposed recently in which the features are somewhat complex and non-linear functions in DCT (or pixel) domain. Hence it is being hard for the pixel values to be invertible. These are considered as mathematically objectionable full- reference IQA methods in this paper. If the full- reference IQA methods have the three mathematically desirable properties, it is noticed that they can be used as objective methods for optimization problems of image processing applications effectively.

II. Literature Review

2.1Characteristics of HVS

The Visual system is the part of the central nervous system which gives organisms the ability to process visual detail. There are two important Human Visual System (HVS) characteristics:

i. Spatial Contrast Sensitivity function (CSF) and

ii. Contrast masking (CM) effect.

These characteristics have been applied in some previous FR-IQA methods.

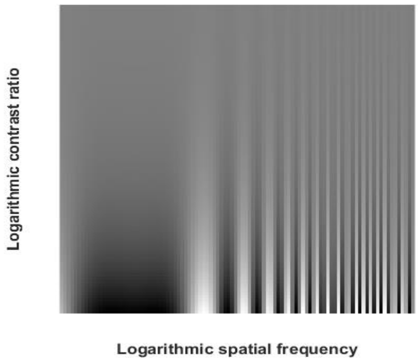


Fig. 2. An example of spatial CSF

According to spatial CSF, Human Visual System (HVS) has different sensitivities to distortions depending on spatial frequency in cycles per degree (cpd). Figure 2 shows the example of spatial CSF. In this example, a sinusoidal grating pattern is introduced into the gray background and it shows the relationship between the spatial frequency and contrast ratio. In Fig. 2, as the spatial frequency of the sinusoidal grating pattern increases along the x-axis the contrast ratio of the sinusoidal grating pattern decreases along the y-axis. Figure 2 shows how the changes in the sinusoidal grating pattern with respect to spatial frequency affect the visual perception of HVS, and it shows the band pass property.

On the other hand, according to contrast masking (CM) effect [13], Human Visual system (HVS) has different sensitivities to distortions depending on background image texture characteristics. There are different image texture regions like homogeneous image texture region, complex image texture region and edge image texture region. HVS responds differently for different noises or distortions. For example, if the original image is distorted with Pseudo additive white Gaussian noise (AWGN) then it will become the distorted image given as Y=X+N, where X, N and Y represents the original image, AWGN distortion and distorted image respectively.

It becomes easy for Human Visual System (HVS) to find that the image is affected with AWGN distortions in homogeneous image texture region where as it becomes difficult to find AWGN distortions in complex image texture region. Similarly this is different for HVS when other distortions were considered. From this we can say that Human Visual System (HVS) perceives different visual quality to different distortions depending on image texture characteristics.

2.2 Some Important FR-IQA metrics

This section deals with overview of some important existing FR-IQA methods. These methods uses different models like similarity measure based models, normalized distance metric based models, internal brain mechanism based models. The following FR-IQA methods include the above mathematical frameworks.

2.2.1 SSIM

Wang *et al.* [2] proposed Structural Similarity (SSIM) index under the assumption that HVS is highly adapted for extracting structural information from the image textures. SSIM is obtained by comparing luminance, contrast and structural information of the reference and distorted images and combining them to get the similarity measure. SSIM is considered as the milestone work in FR-IQA modeling problems. SSIM is further extended to Multi-scale SSIM (M-SSIM) [3] and Information content weighting SSIM (IW-SSIM) [4], etc.

2.2.2 RFSIM

Zhang *et al.* [5] proposed Riesz transform based Feature Similarity (RFSIM) index based on the fact that HVS perceives an image mainly according to its edge regions. RFSIM uses first and second order Riesz transform coefficients as image features. A feature mask is defined as the edge locations of the image and visual quality is calculated only at the edge regions by using canny edge operator. RFSIM is computed by comparing the features at key locations of reference image and distorted image.

2.2.3 FSIM

Zhang *et al.* [6] proposed Feature Similarity (FSIM) index based on the fact that HVS understands an image mainly according to its low-level features. FSIM uses phase congruency (PC) and gradient magnitude (GM) as two main features. Phase congruency is a dimensionless measure of the significance of a local structure. Phase congruency is calculated by using Gabor filtered coefficients and gradient magnitude is calculated by using Scharr gradient operator.

2.2.4 GSM

Liu *et al.* [7] proposed Gradient Similarity (GSM) based method based on the assumption that gradients convey important visual information and crucial to scene understanding. GSM is used to measure the changes in contrast and structure in images. GSM uses four directional high pass filters to capture the change of contrast abd structural information in images. GSM also uses luminance comparison along with mask processing to get the similarity measure between reference and distorted images.

2.2.5 GMSD

Xue *et al.* [8] proposed Gradient Magnitude Similarity Deviation (GMSD) based on the assumption that the image gradients are sensitive to image distortions, while different local structures in distorted image suffer different degrees of degradations. This uses the global variation of gradient based local quality map for overall image quality prediction.

2.2.6 VSI

Zhang *et al.* [9] proposed Visual Saliency induced Index (VSI) which uses gradient magnitude and visual saliency as two main features. Visual saliency is used when computing the local quality map of the distorted image. Visual saliency is calculated by using phase congruency and gradient magnitude is obtained by using the scharr gradient operator. Visual saliency is also employed as a weighting function while pooling the quality score, to reflect the importance of a local region.

Most of the above mentioned FR-IQA methods assume that structural or contrast information in image signals takes a very important role in perceived visual quality. So, in order to characterize such structural or contrast information, finding effective features can be an important factor in modeling an effective full- reference IQA method.

III. Existing Method

3.1Structural Contrast-Quality Index (SC-QI)

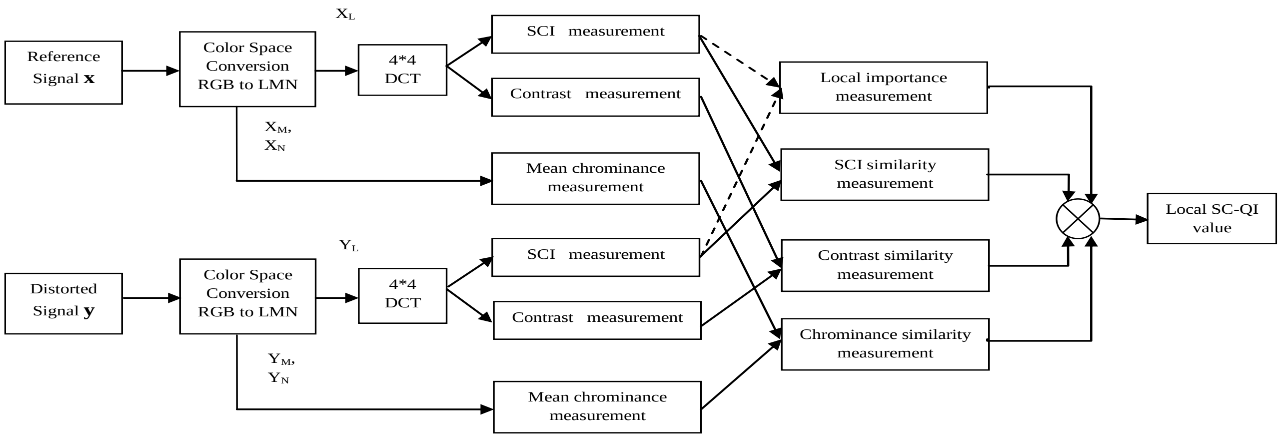


Fig. 3. Block diagram of local SC-QI computation between local image signal x and its local distorted image signal y

Figure 3 shows the block diagram of computation of local SC-QI value between the local reference image signal x ϵ X and local distorted image signal y ϵ Y within the whole reference image X and whole distorted image Y. Local quality value using SC-QI is calculated between x and y for each 4*4 local image block within the whole image. First, the pixel intensity values of x and y are scaled to be in the range of [0, 1]. For the color image signals of x and y which are in the RGB color spaces, they should be decorrelated to luminance (L) and chrominance (M,N) components using LMN color space as

$$\begin{bmatrix} L \\ M \\ N \end{bmatrix} = \begin{bmatrix} 0.06 & 0.63 & 0.27 \\ 0.30 & 0.04 & -0.35 \\ 0.34 & -0.60 & 0.17 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{2}$$

We will then have the local image blocks x_L, x_M and x_N and y_L, y_M and y_N which represents one luminance component and two chrominance components in reference image X and distorted image Y respectively. Only luminance components x_L and y_L are transferred to 4*4 DCT to transform into DCT coefficients. These luminance DCT coefficients are used to calculate the feature values of Structural Contrast Index (SCI) and contrast energies for contrast similarity measurement as shown in Fig. 3. This method follows the similarity measure form which has popularly been used in many FR-IQA methods.

The product of six similarity measures s_k, k=1, 2, . . . , 6 has been used in order to calculate the local quality value between x and y using SC-QI as

$$f(x,y) = \prod_{k=1}^{6} s_k \tag{3}$$

Where each similarity measure s_k, k=1, 2, . . . , 6 has the same form as

$$s_k\left(\phi_{x(k)}, \phi_{y(k)} \middle| \theta_k, v_k\right) = \left(\frac{2\emptyset_{x(k)}\emptyset_{y(k)} + \theta_k}{\emptyset_{x(k)}^2 + \emptyset_{y(k)}^2 + \theta_k}\right)^{v_k} \tag{4}$$

In (4), $\phi_{x(k)}$ represents the feature function of x for the k-th similarity measure and θ_k is used to avoid unstable results when the denominator is close to zero. To control the gradient sensitivity of the similarity measure form θ_k and v_k in (4) serves as model parameters. The product of similarity measures as in (3) gives the meaning that, even if one similarity measure among the six similarity measures is low, then the perceived local visual quality is also low. The first similarity measure (s_1) uses the Structural Contrast Index (SCI) values between x_L and y_L.

As discussed earlier, as per contrast sensitivity function, Human Visual System (HVS) has different sensitivities to distortions depending on spatial frequency. Three similarity measures (s_2, s_3 and s_4) were considered to represent the CSF into the SC-QI design in the form of (4). This is done by comparing contrast energy values in three frequency regions, low frequency (LF), middle frequency (MF) and high frequency (HF) regions respectively in 4*4 DCT blocks for x_L and y_L. The last two similarity measures s_5 and s_6 are calculated by using the chrominance components (M and N signals in (2)), by comparing the two chrominance values. These similarity measures (s_1 to s_6) are obtained in [10].

After the estimation of local quality value $f(.)$ between each pair of $x^{(j)}$ and $y^{(j)}$ for all local image blocks, their weighted average is taken in a pooling stage to calculate the global (whole) quality value of Y compared to X. The global visual quality is defined as

$$F(X,Y) = E[f] = \frac{1}{W}\sum_{j=1}^{J} w(x^j, y^j)\, f(x^j, y^j) \qquad (5)$$

where J in (5) represents the number of 4*4 local image blocks in X and Y, w(x, y) is the local weight with respect to local importance and W is a scaling factor and is equal to sum of all w(x, y) values over all J local image blocks.

3.2 Structural Contrast-Distortion Metric (SC-DM)

SC-DM is the extension work of SC-QI and it uses normalized root mean squared error (NRMSE). By using NRMSE, SC-DM is able to obtain desirable mathematical properties of valid distance metricability with quasi-convexity. The NRMSE is expressed as

$$d(x,y|\theta) = \frac{\|x-y\|_2}{\sqrt{\|x\|_2^2 + \|y\|_2^2 + \theta}} \qquad (6)$$

Instead of using SC-DM as a similarity measure as that of SC-QI, NRMSE is considered for distortion measure as

$$d(\emptyset_x, \emptyset_y|\theta) = NRMSE(\emptyset_x, \emptyset_y|\theta) = \frac{|\emptyset_x - \emptyset_y|}{\sqrt{\emptyset_x^2 + \emptyset_y^2 + \theta}} \qquad (7)$$

The sum of squared six distortion measures (i.e., NRMSEs) gives the local SC-DM value for x and y and is formulated as

$$f(x,y) = \|d\|_2^2 \qquad (8)$$

where $d=\{d_1, d_2,. . . ,d_6\}$. d_1, d_2, d_3 and d_4 have the same features as s_1, s_2, s_3 and s_4 as used in SC-QI. The features in s_5 and s_6 of SC-QI are slightly modified to form d_5 and d_6.

The same weights and pooling methods were employed in SC-DM as that of SC-QI. The only changes of SC-DM as compared to SC-QI are the distortion measure form as in (7) instead of similarity form and the features for d_5 and d_6.

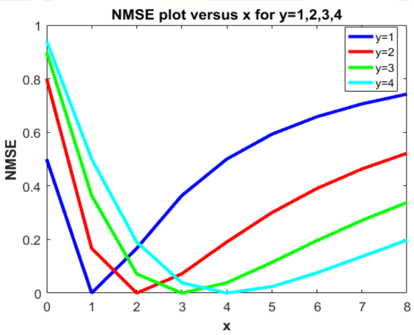


Fig. 4. NMSE plot verses x for fixed y at θ=1

For convenient analysis, to get the local SC-DM form as in (8) for the squared NRMSE, the squared form of NRMSE is defined as normalized mean square error (NMSE) and is given as

$$d^2(x,y|\theta) = NMSE(x,y|\theta) = \frac{(x-y)^2}{(x^2+y^2+\theta)} \qquad (9)$$

where $\theta \geq 0$.An example of NMSE of x for fixed y values (y=1, 2, 3 and 4) and at θ=1 by using (9) is shown in Fig. 4. This shows that at x=y, NMSE has minimal values and is quasi-convex for x.

Even though SC-QI and SC-DM are given very hopeful results, their applications are often limited to objective IQA tasks because they are mathematically hardly able to manage in optimization problems.

IV. PROPOSED METHOD

Many FR-IQA methods recently have been existed based on high-level assumptions such as extracting structural information from image textures for visual quality of Human Visual System (HVS) characteristics, whereas our proposed work is based on low-level assumption i.e., perceived image quality is associated with difference of neural responses between the reference and distorted image signals.

To derive the quality of degraded image, we consider DCT domain as the subband that replicates neural responses in visual cortex. Our proposed method is derived based on probability summation theory and a psychometric function that replicates the low level mechanism of visual cortex. To realize this, we have to consider a psychometric function. The purpose of psychometric function is to serve as an individual detection probability model for distortion at *k*-th DCT coefficient, which is given as

$$(P_k(d_k)) = 1 - exp\{-(v_k \,.\, d_k)^p\} \tag{10}$$

where $d_k = |x_k - y_k|$ is the difference between the original and distorted coefficient values. v_k is the measured sensitivity value for distortion and p is a model parameter to fit the model to the measured detection thresholds of distortions. Even though $(P_k(d_k))$ is detection probability model for distortion visibility it can be interpreted as a detection probability model for quality degradation while considering v_k as a signal importance for perceived visual quality at *k*-th DCT coefficient. Higher the value of v_k, it is more likely to detect the quality degradation with d_k or it is less like to be detected with small v_k [14].

To derive the local quality, an individual detection probability model for quality degradation in (10) is extended by adopting the model for the probability summation theory at each coefficient to a localized probability model for quality degradation in DCT block, which is expressed as

$$P_j(x_j, y_j) = 1 - \prod_{k=1}^{K}\{1 - P_k(d_{j,k})\} \tag{11}$$

Substituting (10) in (11) we get

$$P_j(x_j, y_j) = 1 - exp\{-f(x_j, y_j)\} \tag{12}$$

where

$$f_j(x_j, y_j) = \sum_{k=1}^{K}(v_k \,.\, d_{j,k})^p \tag{13}$$

where $f_j(\,.\,)$ in (13) is the local quality value for *j*-th DCT block. $v_{j,k}$ and $d_{j,k}$ indicates the signal importance and distortion at *k*-th DCT coefficient in *j*-th DCT local block respectively. The whole quality degradation value for the distorted image Y as compared to original image X is defined as the mean of all the local quality values and is given as

$$F(X, Y) = \frac{1}{J}\sum_{j=1}^{J}\sum_{k=1}^{K}(v_{j,k} \,.\, d_{j,k})^p \tag{14}$$

where *J* is the total number of local DCT blocks. We set *p*=2 and assume $v_{j,k}$ to be locally independent (i.e., $v_{j,k} = v_k$, *j*=1, 2, . . . *J*) and is fixed for each subband in (14).
Our proposed method can be simply formulated as

$$F(X, Y) = a^T w \tag{15}$$

$w = [w_1, w_2, \ldots\ldots, w_K]^T$, where $w_k = v_k^2$, *k*=1, 2, ….., *K* and $a = [a_1, a_2, \ldots\ldots, a_K]^T$ is obtained for each subband by using

$$a_k = \frac{1}{J}\sum_{j=1}^{J}(x_{j,k} - y_{j,k})^2 \tag{16}$$

The computation diagram of proposed method is shown in Fig. 5. C_k is the coefficient map consists of *k*-th coefficients of all the $\sqrt{K} * \sqrt{K}$ local DCT blocks.

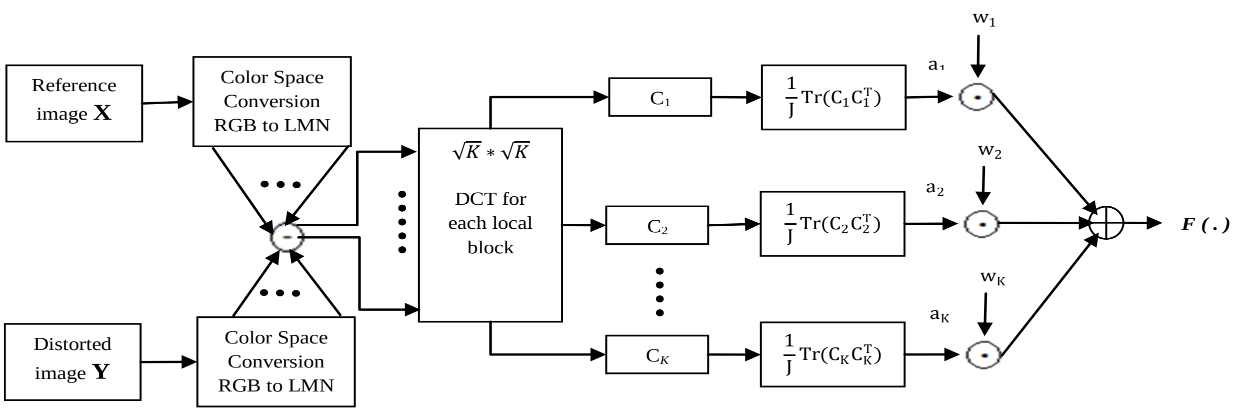


Fig. 5. Computation diagram of proposed method

The final form of proposed metric in (15) is as same as that of [14] and is very simple yet it has some major advantages: (i) it is easy to implement and has desirable mathematical properties as a faithful sign for convex optimization problems; (ii) the simplest form of proposed method can very effectively represent the perceived visual quality characteristics of Human Visual System (HVS) for specific distortion types. These advantages makes this metric can be usable in many image processing applications.

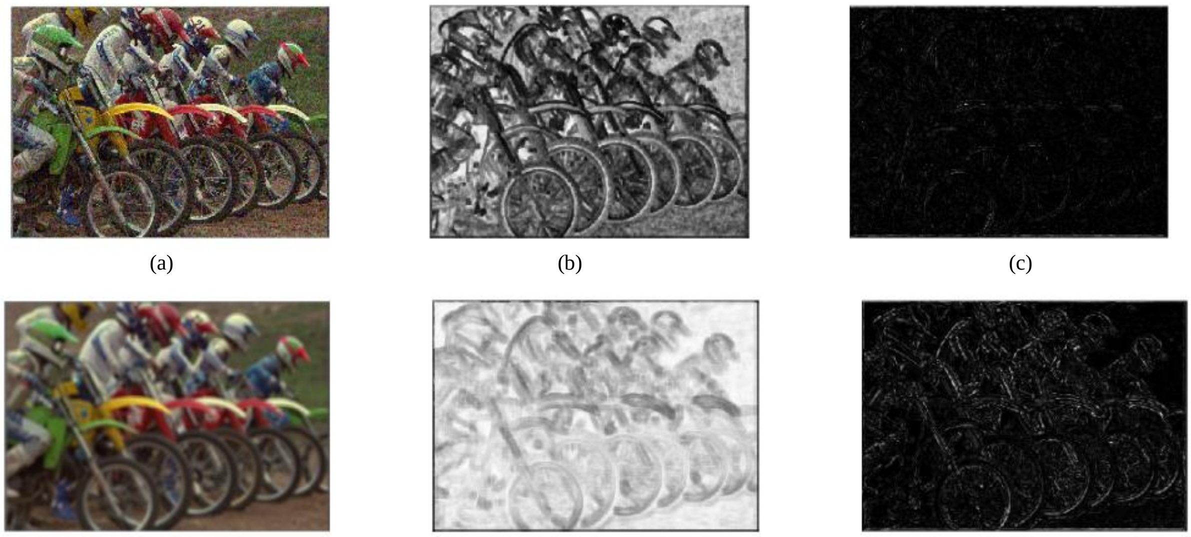
In this method, the number of features (K) is 48 which are obtained from three color channels (LMN), each of the color channel uses 4*4 DCT block having 16 coefficients. In this, we use K=16 for each color channel. As in most existing FR-IQA methods, down-sampling by a factor of 2 using a 2*2 uniform kernel is applied before computing local quality. This reduces the computational complexity and it rarely affects the performance degradation. The local quality values of the proposed method are calculated using a sliding window approach with every one pixel shift, helps in detecting the blockiness artifacts in distorted images.

V. Experimental Results

As discussed earlier Human Visual System (HVS) perceives visual quality depending on different image texture characteristics and various distortion types. In this method we use two types of distortions; pseudo additive Gaussian noise (AGN) and pseudo Gaussian blur (GB).

Figure 6 shows the original image X used to analyze the behavior of some important FR-IQA methods including SC-QI and SC-DM [10]. The response of SC-QI and SC-DM due to these two distortions is also shown in Fig. 7.

Fig. 6. Original image X



(a) (b) (c)

(d) (e) (f)

Fig. 7. Distorted images by additive Gaussian noise (AGN) and Gaussian blur (GB) from the original image X in Fig. 6. SC-QI and SC-DM maps are computed in each row. (a) AGN distortion of X denoted by Y_1. (b) SC-QI map of Y_1. (c) SC-DM map of Y_1. (d) GB distortion of X denoted by Y_2. (e) SC-QI map of Y_2. (f) SC-DM map of Y_2.

In this paper, we not only concentrated on proposed DCT based metric and existing SC-QI and SC-DM, but also analyzed the performance of other important FR-IQA methods. Figure 8 shows the mapping of local image regions of VSI and $FSIM_C$ for both additive Gaussian noise (AGN) and Gaussian blur (GB) distortions.

In Fig. 8, the first column denotes the local maps of Feature Similarity ($FSIM_C$) [6] and second column represents the local quality maps of Visual saliency Induced Index (VSI) [9] respectively. Similarly, the first row in Fig. 8 denotes the local map of $FSIM_C$ and VSI due to pseudo Gaussian blur (GB) distortions and second row denotes the local maps of $FSIM_C$ and VSI due to additive Gaussian noise (AGN) distortions. In our experiments, total eight FR-IQA methods are compared, i.e., SSIM [2], GSM [7], M-SSIM [3], $FSIM_C$ [6], RFSIM [5], VSI [9], and SC-QI (SC-DM) [10] are compared with the proposed quality metric in terms of finding the quality of distorted images.

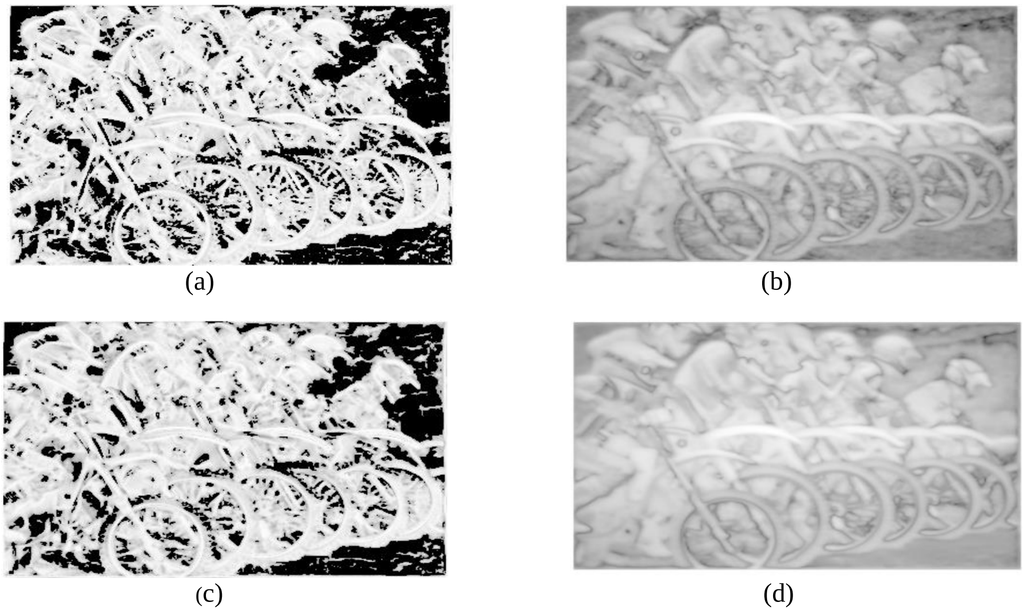


Fig. 8. Mapping of $FSIM_C$ and VSI for GB distortions (first row) and AGN distortions (second row). (a) $FSIM_C$ map of Y_1. (b) VSI map of Y_1. (c) $FSIM_C$ map of Y_2. (d) VSI map of Y_2.

TABLE I

PERFORMANCE COMPARISON OF DIFFERENT FR-IQA METHODS

FR-IQA method	AGN Distortions	GB Distortions
SSIM	0.6209	0.7792
MSSIM	0.8918	0.9255
RFSIM	0.5811	0.8986
FSIM	0.9412	0.9039
GSM	0.9708	0.9558
VSI	0.9722	0.9650
SCQI	0.9727	0.9718
SCDM	0.9728	0.9718
Proposed metric	**0.9735**	**0.9741**

Table I shows the quality values obtained for both additive Gaussian noise(AGN) distortions and pseudo Gaussian blur (GB) distortions for the above mentioned FR-IQA methods. From the performed experiments, it is obvious that the proposed method is easy to implement and it provides valid results for quality measures. It can overcome the image quality optimization problems with its desirable mathematical properties i.e., differentiability, convexity and valid distance metricability.

It is shown in Table I that each method has different performance to different distortions. The proposed DCT based metric produces promising results to different distortions for image quality optimization problems. This indicates that the proposed method can be effectively applied in image compression problems with respect to optimizing perceptual visual quality.

Different color spaces will show effect on the behavior of classification accuracies for some image classification and recognition tasks. The color space conversion used in other image quality metrics [6], [14] is RGB to YIQ. We compare the performance of DCT-based quality metric for YIQ color space with LMN color space.

TABLE II

PERFORMANCE COMPARISON OF PROPOSED DCT-BASED METRIC FOR LMN AND YIQ COLOR SPACES

Color space	AGN Distortions	GB Distortions
LMN	0.9735	0.9741
YIQ	0.9707	0.9721

Table II shows the performance of DCT-based quality metric for LMN and YIQ color spaces. It is noticed that YIQ color space has comparable results against LMN color space, indicating that DCT based metric can effectively be applied for different applications with their dedicated color spaces. These experiments are performed on MATLAB 9.0 (R2016a) software.

VI. CONCLUSION

In this paper, a FR-IQA method based on DCT is presented by considering that HVS has different strategies in perceiving visual quality. The proposed metric has mathematically desirable properties of differentiability, convexity and valid distance metricability and is of less computational complexity as compared to other metrics. This can be effectively used not only for objective IQA tasks, but also for image quality optimization problems with known distortion types. With some modification the proposed metric can also be used with different DCT kernel sizes for improving the perceptual visual quality.

REFERENCES

[1] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it? A new look at signal fidelity measures," *IEEE Signal Process. Mag.*, vol. 26, no. 1, pp. 98–117, Jan. 2009.

[2] 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.

[3] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. Conf. Rec. 37th Asilomar Conf. Signals, Syst., Comput.*, Nov. 2003, pp. 1398–1402.

[4] Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 5. pp. 1185–1198, May 2011.

[5] L. Zhang, D. Zhang, and X. Mou, "RFSIM: A feature based image quality assessment metric using Riesz transforms," in *Proc. 17th IEEE Int. Conf. Image Process.*, Sep. 2010, pp. 321–324.

[6] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.

[7] A. Liu, W. Lin, and M. Narwaria, "Image quality assessment based on gradient similarity," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1500–1512, Apr. 2012.

[8] W. Xue, L. Zhang, X. Mou, and A. C. Bovik, "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index," *IEEE Trans. Image Process.*, vol. 23, no. 2, pp. 684–695, Feb. 2014.

[9] L. Zhang, Y. Shen, and H. Li, "VSI: A visual saliency-induced index for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 23, no. 10, pp. 4270–4281, Oct. 2014.

[10] S.-H. Bae and M. Kim, "A novel image quality assessment with globally and locally consilient visual quality perception," *IEEE Trans. Image Process.*, vol. 25, no. 5, pp. 2392–2406, May 2016.

[11] D. Brunet, E. R. Vrscay, and Z. Wang, "On the mathematical properties of the structural similarity index," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1488–1499, Apr. 2012.

[12] J. Wu, W. Lin, G. Shi, and A. Liu, "Perceptual quality metric with internal generative mechanism," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 43–54, Jan. 2013.

[13] G. E. Legge and J. M. Foley, "Contrast masking in human vision," *J. Opt. Soc. Amer.*, vol. 70, no. 12, pp. 1458–1471, Dec. 1980.

[14] S.-H. Bae and M. Kim, "DCT-QM: A DCT-based quality degradation metric for image quality optimization problems," *IEEE Trans. Image Process.*, vol. 25, no. 10, pp. 4916–4930, Oct. 2016.