Image Quality Assessment Measure based on Visual Region and displacement with distorted pixels for compressed Images

A. S. Zianou and F. Hachouf

Image quality assessment (QA) plays a major role in a broad range of applications. Evaluating the adequacy of a quality image for a given application (image or video processing applications) is a requisite. However, objective quality assessment is far from being a solved problem. In this paper, a measure for image quality assessment is introduced after a brief review of previous work. A metric based on the distance between blocks of distorted and original image is proposed to overcome some of the limitations of existing approaches. Distorted pixel and visual regions of interest are computed. Experimental results confirm the potential of the proposed measure.

Keywords: Image Quality, Distorted Pixel, Region.

1. INTRODUCTION

Image QA [3,4,6,12,20-22] plays a fundamental role in the design and evaluation of imaging and image processing systems. As an example, QA algorithms can be used to systematically evaluate the performance of different image compression algorithms that attempt to minimize the number of bits required to store an image, while maintaining sufficiently high image quality. Similarly, QA algorithms can be used to evaluate image acquisition and display systems. Communication networks have developed tremendously over the past decade, and images and video are frequently transported over optic fiber, packet switched networks like the Internet, wireless systems, etc. Bandwidth efficiency of applications such as video conferencing and Video on Demand can be improved using QA systems to evaluate the effects of channel errors on the transported images and video. Further, QA algorithms can be used in “perceptually optimal” design of various components of an image communication system. Finally, QA and the psychophysics of human vision are closely related disciplines. Research on image and video QA may lend deep insights into the functioning of the human visual system (HVS), which would be of great scientific value.

Image QA methods can be classified as subjective and objective methods. The first approaches to image quality evaluation are subjective quality testing which is based on observers that evaluate image quality. These tests are time consuming, expensive and have a very strict definition of observational conditions.

The second approaches are the objective image quality testing based on mathematical calculations. It can be roughly divided into three main categories: full reference (FR) criteria [9,16], reduced reference (RR) criteria and no reference (NR) criteria. Obviously,
all these criteria are a function of the distorted image. FR criteria are a function of the original image which is assumed to be free from distortions (called the ‘‘reference image’’). The objective measures used in reference methods include peak signal-to-noise ratio (PSNR) [23], structural similarity (SSIM) [26] and the visual information fidelity (VIF) [15]. Furthermore, in [12], they describe two strategies—visual fixation-based weighting, and quality-based weighting. By contrast with some prior studies they find that these strategies can improve the correlations with subjective judgment significantly. In [13], the relationship between the QA and utility assessment tasks for natural images are used. Two psychophysical experiments are conducted to collect perceived quality scores and perceived utility scores for a collection of test images corresponding to signal-based representations and visual-structure-preserving representations. The results from these experiments provide evidence that any QA algorithm optimized to predict perceived quality scores cannot immediately predict perceived utility scores. In [14], they investigate how the SSIM components contribute to its quality evaluation of common image artifacts. After a nonlinear mapping, the product of the variance and cross-correlation components yields nearly identical linear correlation with subjective ratings as the complete SSIM and mean SSIM computations. In [17], image assessment (IA) is considered as a task-specific evaluation and examines the performance of a variety of objective IA algorithms for three different assessment tasks: fidelity assessment, QA, and utility assessment. In [19], they study two increasingly popular paradigms for image QA - Structural Similarity (SSIM) metrics and Information Fidelity metrics.

RR criteria require a partial knowledge of the reference image (this knowledge is called the ‘‘RR’’). In [24], an image quality criterion is proposed. This criterion, called C4, is fully generic and based on a rather elaborate model of HVS. This model describes the organization and operation of many stages of vision, from the eye to the ventral and dorsal pathways in the visual cortex.

NR criteria do not have any information about the reference image. Wang et al. [27] proposed an effective and efficient QA model for JPEG images. Bovik and Liu [28] modeled the blocking-artifacts to measure the compressed image quality. Mei et al. [29] proposed a novel spatio-temporal QA scheme by using low-level content features for home videos. Luo and Tang [30] evaluated visual quality by using high-level semantic features are utilized. Badu et al. [31] applied the edge amplitude and length, the background luminance, and activity to image QA. Sheikh et al. [11] proposed an NR image QA metric that uses natural scene statistics (NSS) for JPEG2000 compressed images. In [1,2], a new philosophy of image QA model of JPEG2000 based on pixel distortions and edge information is presented. In [5], a blind image QA is proposed. Blind image deconvolution methods are used to determine the metric. Existing direct deconvolution methods based on the cepstrum, bicepstrum and on a spectral subtraction technique are compared across 210 images. A variation of the spectral subtraction method, based on a power spectrum surface of revolution, is proposed. In [7], a no-reference QA metric for images subject is proposed to quantization noise in block-based DCT (discrete cosine transform) domain, as those resulting from JPEG or MPEG encoding. In [8], a novel NR method is presented to assess the quality of JPEG-coded images using a sequential learning algorithm for growing and pruning radial basis function (GAP-RBF) network. In [18], they consider natural scenes statistics and adopt multi-resolution decomposition methods to extract reliable features for QA. In [25], they present a machine learning approach to measure the visual quality of JPEG-coded images. The features for predicting the perceived image quality are extracted by considering key HVS factors such as edge amplitude, edge length, background activity and background luminance.

This paper is organized as follows; some algorithms of image quality measures are
presented in Section 2. The proposed image quality measure is defined next. Performance of the proposed measure will be illustrated by examples involving images with different types of distortion in Section 4.

2. IMAGE QUALITY MEASURES

Full reference QA techniques proposed in the literature can be divided into two major groups: those based on the HVS and those based on arbitrary signal fidelity criteria. Section A reviews conventional Quality Measures: Signal to Noise Ratio (SNR) and PSNR, and section B Structural Similarity Index (SSIM) and Visual region of interest Weighted Quality Index (VroiWQI).

2.1 Conventional Quality Measures: SNR and PSNR

The simplest and most widely used full-reference quality metric are mean squared error (MSE) computed by averaging the squared intensity differences of distorted and reference image pixels, along with the related quantity of PSNR. MSE and PSNR are widely used because they are simple to calculate, have clear physical meanings, and are mathematically easy to deal with for optimization purposes (MSE is differentiable, for example). But they are not very well matched to perceived visual quality. In the last three decades, a great deal of effort has gone into the development of QA methods that take advantage of known characteristics of HVS. The majority of the proposed perceptual QA models have followed a strategy of modifying the MSE measure so that errors are penalized in accordance with their visibility.

2.2 Structural Similarity Index (SSIM) and VroiWQI

This method [26] is based on comparing the structures of the reference and the distorted images. The Structural Similarity Index is defined as a function of luminance comparison \( l(x, y) \), contrast comparison \( c(x, y) \) and structure comparison \( s(x, y) \) represented in equation (1).

\[
SSIM(x, y) = \left[ l(x, y) \right]^\alpha \cdot \left[ c(x, y) \right]^\beta \cdot \left[ s(x, y) \right]^\gamma
\]

where \( \alpha, \beta \) and \( \gamma \) are parameters defining the relative importance of the three components. Specifically these are set to \( \alpha = \beta = \gamma = 1 \) assuming that the three components possess equal importance.

The overall image quality can be evaluated by mean SSIM (MSSIM), which is defined as

\[
MSSIM(x, y) = \frac{1}{M} \sum_{j=1}^{M} SSIM(x_j, y_j)
\]

MSSIM is significantly interesting for its novel theory and better results. However, in [34] it was found that method fails in measuring the badly blurred images. They have proposed an improved QA called Gradient-based Structural Similarity (GSSIM) based on the edge information as the most important image structure information. This measure is very interesting in using edge information with SSIM but, it cannot use for measuring quality of the pixel displacement information. In [33] a visual region of interest weighted quality index is introduced. The index is based on weighted quality indices of local regions that capture structural distortion in the local regions between the test image and the original image. Weights are assigned in accordance with visual regions of interest, characterized by the entropy of the region which emphasizes the texture variance in that region. Equation (3)
gives the expression for \( VroiWQI \)

\[
vroiWQI = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} [E] * [SSIM]}{\sum_{i=1}^{M} \sum_{j=1}^{N} [E]}
\]  \hspace{1cm} (3)

Where \([E]\) is the subset matrix of normalized entropy values \(E\). \(VroiWQI\) not interested in using edge information and the pixel displacement.

Motivated with these studies, in this paper, we proposed a region, distorted and displacement measure which doesn’t exploit a priori knowledge about the distorted image and the types of artifacts.

3. IMAGE QA METHOD

It is easily deducible that most of the distortion in image is due to the block DCT-based (discrete cosine transform) compression. The most popular and widely used image format on Internet and digital cameras happens to be JPEG. Since JPEG uses the block-based DCT transform for coding, to achieve compression, the major artifact that JPEG-compressed images suffer, is blockiness. In the JPEG coding, non-overlapping \(8 \times 8\) pixel blocks are coded independently using DCT transform. The compression ratio and the image quality are mainly determined by the degree of quantization of these DCT coefficients. The undesirable consequences of quantization manifest as blockiness, ringing and blurring artifacts in the JPEG-coded image. It turns out that the subjective data for all these artifacts are highly correlated. Hence, measuring the blockiness in turn indicates the overall image quality.

The proposed image quality metric is designed to take into consideration the various human visual criteria while quantifying image quality. Therefore, the evaluation of a distorted image is based on the calculation of the displacement of pixel belonging to a \(b_{ik}\) of the distorted image in relation to \(b_{ki}\) of the original image. Then, We propose an improved SSIM algorithm; the Region-Displacement with Distorted Pixels measure (RDDM), which compares the position and intensity information between the distorted image block and the original one, and add the displacement and distorted pixels structure comparison, \(pie(b_{ki}, b_{kj})\) in equation (1).

The reference image is denoted by \(Original(mm, nn)\) and the distorted image is denoted by \(Distorted(mm, nn)\). Both images have \(M \times N\) pixels. Our method has four steps as Fig.1 shows. Thus, the images are partitioned into overlapped \(11 \times 11\) blocks, where the overlapping area is on one pixel (see Fig.2). Let \(b_{kj}\) and \(b_{ki}\) be blocks of distorted image and original image respectively. The others steps are described in the following.

3.1 Distorted pixels

Pixel distortions are estimated using the following feature. First, standard deviation (SD) of a central pixel is estimated within \(3 \times 3\) neighborhood pixels which is applied for all available central pixels in the image. SD values are then averaged within a \(3 \times 3\) partially overlapping block (see Fig.2). Let \(X_{ij}\) and \(X_{fij}\) be the centrals pixels of the \(3 \times 3\) pixels ‘neighborhood’ in original and distorted image respectively as shown in Fig. 3 (a)(b); also let \(M_{i}\) be the mean of pixels within the \(3 \times 3\) pixels of the distorted image and \(X_{fij}\) its intensities. All available central pixels of images (original and distorted) are located within the bolded line boundary as shown in Fig.4.
The statistical features (the mean) can be estimated as follows:

\[ M_{fi} = \frac{1}{9} \sum_{i=1}^{3} \sum_{j=3}^{3} X_{fij} \]  

The same procedure is applied on the original image. This one has also a mean \( M_i \) and \( X_{ij} \) its intensities. For each 3×3 partially overlapping block of the original image, the mean is calculated.

\[ M_i = \frac{1}{9} \sum_{i=1}^{3} \sum_{j=3}^{3} X_{ij} \]  

The SD between the two 3×3 partially overlapping blocks of the original image and distorted image can be estimated as follows:

\[ SD = \left( \frac{1}{9} \sum_{i=1}^{3} \sum_{j=3}^{3} \left( M_{fij} - X_{fij} \right)^2 + \left( M_{ij} - X_{ij} \right)^2 \right)^{1/2} \]  

The distorted pixels measurement \( Id \) between the two \( bk_i \) and \( bk_j \) is defined by:

\[ Id = \frac{1}{A} \sum_{i=1}^{A} SD_i \]  

where \( A = 11 \times 11/3 \times 3 \) is the total number of the centrals pixels of the 3×3 pixels “neighborhood” within block of 11×11 pixels. All \( Id \) can be arrayed in a (m×n) matrix \( Id_{global} = [Id_{ij}] \) (1 ≤ i ≤ m=M/11 and 1 ≤ j ≤ n=N/11).

### 3.2 Displacement measure

This criterion calculates the distance between the pixels of the distorted image and the original image. It takes pixels of distorted image within block of 11×11 pixels and computes the pixel displacement in the block of 11×11 pixels of original image. One defines the criterion for each ‘db(x,y)’ pixel of the \( bk_j \) by using equation (8); where \( G_i \) is pixel belongs to \( bk_i \) of the original image, \( f \) can be the intensity of pixel or any other attribute, \( x \) and \( y \) are pixel position of ‘dp’ in \( bk_j \), \( x' \) and \( y' \) are pixel position of \( G_i \) in \( bk_i \) of original image, \( \text{diff}_{ij}(db(x,y),bk_i) \) is distance between ‘db(x,y)’ pixel and \( bk_i \),

\[ d((x,y),(x',y')) = \sqrt{(x-x')^2 + (y-y')^2} \]  

is Euclidian distance between two pixels, \( bk_i = \{G_{i1}(x'_1,y'_1),G_{i2}(x'_2,y'_2),...,\} \), and \( \mid \cdot \mid \) is absolute value.

---

**Fig.1 Image quality assessment method**
Fig. 2. Image pixels and its partitioning into overlapped 11×11 blocks and overlapped 3×3 blocks.

Fig. 3. The 3×3 neighborhood pixels: (a) central pixel $X_{23}$ of the 3×3 pixels of original image (b) central pixel $X'_{23}$ of the 3×3 pixels of distorted image.

Fig. 4. Image pixels: (a) original image ($M \times N$), (b) distorted image ($M \times N$).
\[
\text{diff}_{ij}(db(x,y),bk) = \begin{cases} 
0 & \text{if } x = x' \land y = y' \land f(db) = f(G_i) \\
\min \left( \text{position}, \text{color} \right) & \text{if } x \neq x' \lor y \neq y'
\end{cases}
\] (8)

Where

\[
\text{position} = \min \left( d((x,y),(x',y')), d((x,y),(x'',y'')), \ldots \right),
\]

\[
\text{color} = \min \left( |f(db) - f(G_i)|, |f(db) - f(G_j)|, \ldots \right)
\]

The following formula is used to calculate the displacement pixel measure related to \(bk_j\) of the distorted image:

\[
Dp = \frac{1}{L} \sum_{j=1}^{L} \text{diff}_{ij}(db_j(x,y),bk_j)
\] (9)

Fig.5 (a) Original House image (b) Simulated image showing the Visual regions of interest map

Where \(L (L=11 \times 11)\) is the total number of \(bk_i\).

All \(Dp\) can be arrayed in a \((m \times n)\) matrix \(Dp_{\text{global}} = [Dp_{ij}] \) \((1 \leq i \leq m=M/11\) and \(1 \leq j \leq n=N/11)\).

The overall distorted pixels and displacement evaluation between two blocks \(bk_i\) and \(bk_j\) is obtained by combining the two results defined previously is given by:

\[
MB(\text{Id}_{i,j},Dp_{i,j}) = \max \left( \text{Id}_{i,j}, Dp_{i,j} \right)
\] (10)

3.3 Visual regions of interest

In this stage we have used the same function developed in [33] where the regions in the original image is modeled by computing the entropy “\(e\)” in each block of \(11 \times 11\) pixels of the original image, using equation (11), where \(h_i\) is a random variable indicating intensity, \(p(h_i)\) is the histogram of the intensity levels in a region, \(K\) is the number of possible intensity levels. \(K\) varies from 0 to 255 for gray scale images.

\[
e = - \sum_{i=1}^{K} p(h_i) \log_2 p(h_i)
\] (11)
All $e$ can be arrayed in a $(m \times n)$ matrix $E = [e_{ij}]$ (and $I \leq j \leq n=N/11$), where each element $e_{ij}$ is the entropy of the block.

In this step, the formula for measuring the region structure comparison measures is as follows:

$$s_r(b_{ki}, b_{kj}) = e_{ij}^* s(b_{ki}, b_{kj})$$ (12)

Where $s(b_{ki}, b_{kj}) = (\sigma_{b_{ki}}, b_{ki} + C_3) / (\sigma_{b_{ki}} \sigma_{b_{kj}} + C_3)$, $C_3 = C_2/2$ where $C_2 = (K_2 L_2)$, $K_2 << 1$, $L_2$ is the dynamic range of the pixel values (255 for 8-bit grayscale images), $\sigma_{b_{ki}}$ and $\sigma_{b_{kj}}$ are the standard deviation of blocks $b_{ki}$ and $b_{kj}$ respectively, $\sigma_{b_{ki}, b_{kj}}$ is the covariance of blocks $b_{ki}$ and $b_{kj}$.

Fig.5 displays the original “House” image (512 x 512) and its normalized entropy map $E$, with a value of 1 indicating the region of highest interest represented as white, and a value of 0 indicating the region of lowest interest as black.

The other levels of the image ranging from white to black represent the Visual regions with descending level of interest.

### 3.4 Global error

The Region-Displacement with Distorted Pixels measure ($RDDM$) is described as follows

$$RDDM(b_{ki}, b_{kj}) = [I(b_{ki}, b_{kj})]^\alpha [C(b_{ki}, b_{kj})]^\beta [s_E(b_{ki}, b_{kj})]^\gamma + \omega \cdot pie(b_{ki}, b_{kj})$$ (13)

Where: $pie(b_{ki}, b_{kj}) = MB(Id_{i,j}, Dp_{i,j})$, $s_E(b_{ki}, b_{kj}) = s_r(b_{ki}, b_{kj})$, $\omega << 1$. We set $\alpha=\beta=\gamma=1$.

The overall error measure of the original image and a distorted one are calculated as sum of $RDDM(b_{ki}, b_{kj})$; The normalized measure is defined as

$$RDDM(Original, Distorted) = \frac{1}{2} * Eb \sum_{i=1}^{M} RDDM(b_{ki}, b_{kj})$$ (14)

Where $Eb = (M \times N)/(11 \times 11)$ is the total number of blocks.

This measure is closer to “zero” when the image has the best quality and closer to the “one” in the other case.

Finally, the described algorithm can be summarized in the following steps:

1. All means and displacement error are computed and collected in matrix means ($Id_{\text{global}}$) and matrix displacement error ($Dp_{\text{global}}$) respectively.

2. Compute $MB(Id_{i,j}, Dp_{i,j})$ as distorted and displacement pixel between the two block of $Original(mm,nn)$ and $Distorted(mm,nn)$ images.

3. Compute the Region-Displacement with Distorted Pixels measure ($RDDM$).
4. EXPERIMENTAL RESULTS

To verify the effectiveness of the proposed method, experiments are performed on the Live Image Quality Assess Database Release2 [10] of the Laboratory for Image & Video Engineering in the University of Texas at Austin. The database consists of twenty-nine high resolution 24-bits/pixel RGB color images, distorted using five distortion types: JPEG2000, JPEG, White noise in the RGB components, Gaussian blur in the RGB components, and bit errors in JPEG2000 bit stream using a fast-fading Rayleigh channel model. Each image was distorted with each type, and for each type the perceptual quality covered the entire quality range. We tested the proposed method on available images in the LIVE database. In our experiments, we have used the images with sizes 768 x 512 pixels. These images are the results of the following method: JPEG2000 and JPEG. Fig.6 shows sample images and their respective distorted images. Fig.7 shows reference images of the LIVE database (c), original images (a) and its transformed images (b).

In order to provide quantitative measures on the performance of our proposed RDDM QA method, we follow the standard performance evaluation procedures employed in the video quality experts group (VQEG) FR-TV Phase II test [32], where mainly three evaluation metrics were used (Metric 1, Metric 2, Metric 3, Metric 4 and Metric 5). We performed non-linear mapping between the objective and subjective scores [32]. Five-parameter ($\beta_1, \beta_2, \beta_3, \beta_4$ and $\beta_5$) non-linear mapping is used to transform the set of quality ratings by the objective quality metrics to a set of the predicted Difference Mean Opinion Score ($DMOS_p$) values denoted $DMOS_p$.

\[
DMOS_p = \beta_1 \log \text{logistic}(\beta_2, (VQR - \beta_3)) + \beta_4 + \beta_5
\]  

(15)

\[
\text{logistic}(\tau, VQR) = \frac{1}{2} - \frac{1}{1 + \exp(VQR\tau)}
\]  

(16)

where $VQR$ is the quality rating by the objective method ($RRDM, VroiWQI, GSSIM$ or $PSNR$) and $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ are chosen for best fit. The fitting (optimization of $\beta$...
parameters) was done with MATLAB’s \textit{fminsearch} function. This section lists the evaluation metrics to be calculated on the subjective and objective data.

Once the non-linear transformation has been applied, the objective model's prediction performance is then evaluated by computing various metrics on the actual sets of subjectively measured \textit{DMOS} and the predicted \textit{pDMOS}.

The following metrics are used in calculation:

- **Metric 1:** Pearson Correlation Coefficient (\textit{CC}) between objective (\textit{pDMOS}) and subjective (\textit{DMOS}) scores. It provides an evaluation of \textit{prediction accuracy} and it is defined by:

  \[
  CC = \frac{\sum_{i=1}^{N} (DMOS(i) - \overline{DMOS})(pDMOS(i) - \overline{pDMOS})}{\sqrt{\sum (DMOS(i) - \overline{DMOS})^2} \sqrt{\sum (pDMOS(i) - \overline{pDMOS})^2}}
  \]  

  where the index \(i\) denotes the image sample and \(N\) denotes the number of samples.

- **Metric 2:** Spearman Rank Order Correlation Coefficient (\textit{ROCC}) between objective (\textit{pDMOS}) and subjective (\textit{DMOS}) scores. It is considered as a measure of \textit{prediction monotonicity} and it is defined by:

  \[
  ROCC = 1 - \frac{6 \sum (DMOS(i) - pDMOS(i))^2}{N(N^2 - 1)}
  \]  

  where 6 is a constant (it is always used in the formula).
Fig. 7 (a) Original image (b) simulated image showing the Visual regions of interest map of original image (c) and reference images of the LIVE database.

Fig. 8 Scatter plots of DMOS versus model prediction for JPEG2000, JPEG (RDDM).
Fig. 9 Scatter plots of DMOS versus model prediction for JPEG2000, JPEG (PSNR).

Fig. 10 Scatter plots of DMOS versus model prediction for JPEG2000, JPEG (VroiWQI).

Table 1. Performance comparison of image quality assessment methods (PSNR, VroiWQI, and the RDDM) on JPEG distorted images

<table>
<thead>
<tr>
<th>Model</th>
<th>CC</th>
<th>ROCC</th>
<th>MAE</th>
<th>RMS</th>
<th>OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSNR</td>
<td>0.8921</td>
<td>0.6169</td>
<td>5.991</td>
<td>5.954</td>
<td>0.063</td>
</tr>
<tr>
<td>VroiWQI</td>
<td>0.9009</td>
<td>0.7982</td>
<td>5.097</td>
<td>5.887</td>
<td>0.062</td>
</tr>
<tr>
<td>RDDM</td>
<td>0.9275</td>
<td>0.9061</td>
<td>5.092</td>
<td>5.480</td>
<td>0.060</td>
</tr>
</tbody>
</table>

Table 2. Performance comparison of image quality assessment methods (PSNR, VROIWQI, and the RDDM) for JPEG2000 images

<table>
<thead>
<tr>
<th>Model</th>
<th>CC</th>
<th>ROCC</th>
<th>MAE</th>
<th>RMS</th>
<th>OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSNR</td>
<td>0.8526</td>
<td>0.7541</td>
<td>8,916</td>
<td>8,9169</td>
<td>0.085</td>
</tr>
<tr>
<td>VROIWQI</td>
<td>0.8830</td>
<td>0.8456</td>
<td>4,745</td>
<td>5.7920</td>
<td>0.065</td>
</tr>
<tr>
<td>RDDM</td>
<td>0.9098</td>
<td>0.9168</td>
<td>4,540</td>
<td>5,6055</td>
<td>0.065</td>
</tr>
</tbody>
</table>

- Metric 3: Outlier ratio (OR). This metric evaluates an objective model's ability to provide consistently accurate predictions for all types of images. The model's prediction consistency can be measured by the number of outlier points (defined as having an error greater than some threshold as a fraction of the total number of points). A smaller
outlier fraction means the model’s predictions are more consistent. It is considered as a measure of prediction consistency and it is defined by:

$$OR = \frac{\text{number of outliers}}{N}$$ \hspace{1cm} (19)

where an outlier is a point for which:

$$|DMOS(i) - DMOS_p(i)| > 2 \times \sigma(DMOS(i))$$

where \( \sigma(DMOS(i)) \) represents the standard deviation of the individual scores associated with the image sample \( i \). The individual scores are approximately normally distributed and therefore twice the \( \sigma \) value represents the 95% confidence interval.

Thus, \( 2 \times \sigma(DMOS(i)) \) value represents a good threshold for defining an outlier point.

- Metric 4: Average absolute prediction error (MAE) between objective (\( DMOS_p \)) and subjective (\( DMOS \)) scores is defined by:

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |DMOS(i) - DMOS_p(i)|$$ \hspace{1cm} (20)

- Metric 5: Root Mean Square Error (RMS) between objective (\( DMOS_p \)) and subjective (\( DMOS \)) scores is defined by:

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (DMOS(i) - DMOS_p(i))^2}$$ \hspace{1cm} (21)

Metric 4, 5 are also considered as a measure of prediction accuracy.

The simulation results for all the distorted images are shown in Figs. 8, 9, 10, Tables 1 and 2. Five metrics are used to measure these objective models. The \( CC \) after non-linear regression means the correlation degree between each model and \( DMOS \), and the larger \( CC \) value means the better accuracy. The spearman \( ROCC \), it is considered as a measure of prediction monotonicity, higher \( ROCC \) means the better prediction monotonicity. The \( MAE \), root mean \( RMS \) and \( OR \) are measures of prediction consistency, smaller value means better performance. Figs. 8, 9 and 10 show respectively the scatter plots of \( DMOS \) versus RDDM, PSNR and VroiWQI.

The correlation between \( DMOS \) and RDDM based image quality metric are shown in Fig. 8 (a) and (b). Similar study is carried out using PSNR quality metric (see Fig. 9(a) and (b)) and VroiWQI (Fig. 10(a) and (b)). It is clear that the proposed RDDM is consistent with the subjective scores much better than VroiWQI and PSNR.

The VroiWQI combines three factors: loss of correlation, luminance distortion and contrast distortion. The value computed by that measure is not convenient in subjective evaluation.

In this test, proposed image quality measure produces results that are in good agreement with subjective visual quality of corresponding images. From the validation result listed above, our algorithm RDDM has a better correspondence to human judgment. The results in Tables 1 and 2 show that RDDM gives notably superior values over VroiWQI and PSNR. The results clearly show that the proposed RDDM model predicts the image quality better than VroiWQI metric and comparable with the PSNR. This can also be deduced from the quantitative performance analysis given in Tables 1 and 2.
5. CONCLUSION

In this paper, we have introduced a novel image QA based on distorted pixel, region and displacement. First, the distorted and original images are divided into blocks of 11×11 pixels, then we calculate pixel displacement and distorted pixels, which can be used to compute the whole error. The performance of the proposed metric is found to be better than previously metrics which does not use the displacement pixel in calculation.

We are continuing efforts into improving the RDDM by introducing other attribute in the measurement. The use of new distances may be improving the results. We are hopeful that this new approach will give new insights into visual perception of quality.

Acknowledgment

The authors would like to thank Dr. H.R. Sheikh for providing the LIVE QA Database to test our metric. The authors express their thanks for anonymous reviewers for their constructive comments to improve the quality of this paper.

References


[34] G.H. Chen, C.L. Yang, S.L. Xie, Gradient-Based Structural Similarity for Image Quality Assessment, ICIP06(2929-2932).