REDUCED-REFERENCE PICTURE QUALITY ASSESSMENT USING DISCRIMINATIVE LOCAL HARMONIC STRENGTH

Universitas Multimedia Nusantara
Tangerang - Banten
Irwan Prasetya Gunawan
irwan@unimedia.ac.id

Abstract
This Paper presents a reduced-reference objective picture quality measurement tool for encoded images. We use a discriminative analysis of harmonic strength computed from the edge-detected pictures to create gain and loss information associated within the picture. This information can then be used to indicate typical degradations such as blockiness or blurriness in the processed images. Our simulations show that quality of images encoded with various encoders such as DCT-based JPEG, wavelet-based JPEG 2000, or various processed images can be quantified within a single metric. We have shown that while traditional method of PSNR fails to agree with the subjective quality scores, our method can overcome such limitation.

Keywords: Image Quality Assessment, Objective Measurement, Reduced-Reference, Blockiness, Blur, Harmonic Strength

1. Introduction

The rapid growth of multimedia applications is leading to the awareness of the perceived value of digital video and image quality. In such applications, digital images/videos may have gone through various processes, any of which can introduce a wide range of distortions that may impair the visual quality. It is desirable that the perceived visual quality can be automatically predicted in an objective manner, since subjective evaluation is time-consuming, laborious, and expensive.

Objective quality assessment of the processed signals is usually performed through comparison with an original (unprocessed) signal. According to the availability of the original signal, the assessment can be classified into full-reference (FR) method when the full original signal is present, and no-reference (NR) method when the original signal is completely unknown. Both FR and NR approaches have their own advantages and drawbacks. With the help of the original signal, FR may be the most accurate method of all, but the requirement to have the original signal is sometimes too costly for most practical applications. On the other hand, NR may be more practical than the FR method at the expense of less accuracy because the method is so optimised for a certain type of distortions that it might fail for any others. In between these two extreme methods, a third type of assessment is also available in which a set of side information is utilised to help the assessment process. This is known as a reduced-reference (RR) evaluation method, and the side information is usually comprised of important extracted features derived from the original signal. It is illustrated in Figure 1. This paper concentrates on the reduced-reference image quality assessment.

In the reduced-reference method, we have the flexibility to determine what kind and how much information is used to help the evaluation. The information can be represented as a scalar, vector, or matrix, depending on the amount of bandwidth or storage available to transmit or store the reduced-reference information. It is also driven by the chosen accuracy of the assessment. Reduced-reference framework is then suitable for remote applications such as automatic quality monitoring of transmitted, re-encoded, or transcoded video/image at remote sites, e.g., a relay station in a broadcast network, a router in a computer network, or a head-end in a cable-TV distribution, where computing power is likely to be limited. Reduced-reference method is therefore a practical solution to a wide range of application whilst maintaining the precision of the assessment.

In this work, we have developed a reduced-reference objective picture quality measurement for encoded images based on harmonics analysis of the spatial gradient computed from the picture without using sophisticated human visual system (HVS) characteristics. Harmonics analysis has been used in [1, 2] in a full-reference (FR) quality measurement framework to measure the amount of blockiness distortions typically found in a block-based encoded picture/video. In the full-reference method of [1], the quality metric is expressed as a blockiness index, which is simply the total harmonic amplitude strength (energy) over the blocks that are considered blocky. However, it is well known that different encoders introduce different kind of distortions. Consequently, we may need specific quality measurement method for each encoder to identify and quantify the distortions attributed to that coder. Since harmonic analysis method in the FR model was optimised for measuring blocking artefacts, it may not work properly well for different kind of distortions and for pictures encoded by other than the block-based and DCT-based codecs, such as the wavelet transform.

Our proposed method has been able to mitigate the aforementioned limitations. Firstly, we have removed the requirement to use full-reference picture; instead, we employ the local harmonic strength extracted from pictures as a reduced-
reference. Secondly, we use discriminative analysis of the local harmonic strength to create gain and loss information that correspond to blocking/tiling distortion and blurring/smearing degradations, respectively.

The contribution of our work in this Paper is manifold. First, is the use of a single tool to quantify different distortions. Another feature of our proposed method is the simplicity of its design, achieved without having to resort to complex method such as HVS. HVS may be advantageous to improve the precision of any quality models, but this can only be done at the expense of increased complexity of the overall design. By keeping the model simple, we hope that the method could be easily adopted for real-time video quality assessment model which is more time consuming. We also have improved the discriminative analysis method such that it works not only for block-based DCT for which it was originally designed, but also expand its capability to include quality assessment for wavelet-based transform images. Therefore we have evaluated the performance of our quality measure with images encoded by the two standard codecs: JPEG [3] (which is based on DCT) and JPEG 2000 [4] (based on wavelet). In addition, the proposed method in this Paper is also useful for quality assessment of pictures contaminated by various types of distortions, such as additive, multiplicative, and impulsive noise, as well as low pass filtered images within VHS recorders. More importantly, our method outperforms the shortcomings of the traditional, widely-used PSNR (peak signal-to-noise ratio) measure.

The rest of this Paper is organised as follows. First, we give a brief description of the method in Section 2. Experimental results are given in Section 3. Finally, concluding remarks are given in Section 4.

2. Discriminative Analysis of Local Harmonic Strength

2.1. Typical artefacts

We sought to identify typical artefacts found in the compressed images. We focus on the two most known compression artefacts namely blocking and blurring [5].

Blocking is typically found in the JPEG coded images as well as in most block-based video compression systems such as H.263 or MPEG-2 [6]. It is characterised by the visibility of the underlying block structure in the encoding process [5] as a result of coarse quantization. Such artificial block boundaries appear in the images at low bitrates, for example.

Blurring is the reduced sharpness of edges and spatial detail [5] and is typically found in low bitrates JPEG-2000 coded images [7]. It is usually contributed by coarse quantization of textured areas at the encoder or pre-low-pass filtering of the input image prior to compression [8]. It can also appear as a result of an imperfect image acquisition process; e.g., by uniform linear motion between the image and the sensor [9]. In addition, blurring can also be found in the JPEG coded images [10], especially when high frequency components are filtered out.

Blocking and blurring artifacts as described above usually require different methods of detection. To have a unified metric for these two types of degradations, the two different methods must be combined in a way similar to the multi-metric model was defined in [11]. However, combining two different detection methods is sometimes complicated since each of them may have different output characteristics and dynamic ranges. As we will describe next, a better and simple quality assessment model may be developed as a joint blockiness/blurriness detection method using discriminative local harmonic strength analysis.

2.2. Proposed method

We follow the downstream model[12] of the reduced-reference method in Figure 1. Both feature extraction stages of the original and the processed images use the same algorithm. Note that in Figure 1 feature extraction stage for a coded picture is embedded in the evaluation stage. The extraction stage is illustrated in Figure 2. The detail is explained in the following sub-section.

2.2.1. Feature extraction

In the feature extraction process, first we apply an edge detection algorithm based on a 3 × 3 Sobel operator to the picture to calculate its gradient image. The resulting gradient picture is then modulated according to position and amplitude by changing the visibility of its pixels; i.e., either higher amplitude of gradient or pixels located around the natural edges will lead to a higher factor of reduced visibility. The purpose of modulation is to prevent false identification of natural edges as blocking artefacts later in the harmonic analysis stage. As a result, natural edges (‘contextual details’) are masked to make any distortions in the picture conspicuous. We observed that the magnitudes of gradient for natural edges are mostly of high amplitude that will make them easily noticed and removed accordingly.

On the JPEG compressed images, the identification of contextual details will preserve edges resulting from the block-based DCT operation on the image, provided that these edges are not masked by any nearby contextual details. Weaker (low intensity) edges on both the reference and the decoded images may still be present; however, it is most likely that these edges would appear at about the same location on the pictures and contribute to the same (or at least, adjacent) bin(s) in the frequency spectrum computed from the 2-D Fast Fourier Transform (explained in the next subsection). The later computation in the evaluation stage will cancel out contributions from these contextual edges, giving an overall result of contextual details elimination.

Prior to the harmonics analysis stage, the masked gradient image is subjected to block segmentation. The block size is chosen such that it is sufficiently large enough to account for any vertical and/or horizontal activity within each block. It is also desirable to have the segmented blocks aligned with the DCT block boundary; hence, we include the DCT block boundary detection and use the result of this detection to achieve perfect alignment with the DCT block when performing block.
even when the segmented block is non-aligned with the DCT block boundary, the subsequent process in the model is not affected too much owing to the frequency-domain operation which is insensitive to spatial shift.

Harmonics analysis is then applied to each segmented block in the intended parts of the masked gradient picture. First, 2-D Fast Fourier Transform (FFT) is computed for these blocks. Second, harmonics components of the FFT spectrum are isolated and their magnitudes are accumulated. What makes these harmonics components stand out among other components is the appearance of pseudo-periodic signal coming from the blockiness tiling pattern in the picture. The accumulated magnitudes of these harmonics components within each FFT block are chosen as the local harmonic strength (LHS) feature that we use for the reduced-reference information. Local harmonics strength can be interpreted as the activity in terms of vertical/horizontal edges in the picture.

Each of the local harmonic strength values corresponds with their respective FFT blocks within the picture. Since these blocks are non-overlapped, the local harmonic strength features can be identified by their coordinates. Therefore, all the local harmonic strengths in a picture can be collected as a matrix.

\[
F = \begin{pmatrix}
    f_{11} & f_{12} & \cdots & f_{1m} \\
    f_{21} & f_{22} & \cdots & f_{2m} \\
    \vdots & \vdots & \ddots & \vdots \\
    f_{m1} & f_{m2} & \cdots & f_{mn}
\end{pmatrix}
\]  

(1)

with its element, \(f_{ij}\), represents the local harmonic strength value with \(1 \leq i \leq m, 1 \leq j \leq n\), and \((i, j)\) represents the coordinate of each FFT block. For the reference and the decoded picture we have \(\hat{F}^{(r)}\) and \(\hat{F}^{(d)}\), respectively. We may not need all the elements of these matrices, because we can confine our attention to some parts of the image. If \(R\) represents a set of coordinates of FFT blocks in the region of interest, one may use a more compact representation of the reduced reference information as

\[
\hat{F}^{(r)} = \left\{(i, j, f_{ij}^{(r)}) \mid (i, j) \in R\right\}
\]  

(2)

and

\[
\hat{F}^{(d)} = \left\{(i, j, f_{ij}^{(d)}) \mid (i, j) \in R\right\}
\]  

(3)

for the reference and the processed/decoded pictures, respectively. Note that in order to make a meaningful comparison between these two features in the evaluation stage, the regions of interest in both the reference and the degraded pictures must be the same.

2.2.2. Discriminative analysis

One of the most important stage in the proposed model is the discriminative analysis stage which will be performed once all the features from the reference and the degraded pictures, \(\hat{F}^{(r)}\) and \(\hat{F}^{(d)}\), are collected. Discriminative analysis of the local harmonic strength is expressed through the following equations. Let \(f_{ij}^{(d)}\) and \(f_{ij}^{(r)}\) represent the local harmonic strengths taken from the matrix features \(\hat{F}^{(d)}\) and \(\hat{F}^{(r)}\) for each degraded and original/reference picture, respectively. If we calculate a local difference as \(\Delta_{ij} = f_{ij}^{(d)} - f_{ij}^{(r)}\), then one can define the local harmonic gain, \(e_{ij}^+\), and the local harmonic loss, \(e_{ij}^-\), of the \(ij\)-th block as:

\[
e_{ij}^+ = \begin{cases} 
\Delta_{ij} & \text{when } \Delta_{ij} > \delta_{ij}^+; \\
0 & \text{otherwise}
\end{cases}
\]  

(4)

and

\[
e_{ij}^- = \begin{cases} 
|\Delta_{ij}| & \text{when } \Delta_{ij} < 0 \text{ and } |\Delta_{ij}| > \delta_{ij}^-; \\
0 & \text{otherwise}
\end{cases}
\]  

(5)

respectively. In these equations, \(\delta_{ij}^+\) and \(\delta_{ij}^-\) are the harmonic threshold values below which the perceived harmonic differences are insignificant. The threshold can be applied locally.
within the segmented block by taking into account the local characteristics of each block such as texture, flatness, contrast, etc. Alternatively, we can use a global threshold by using the same value for all the segmented blocks within the picture. Threshold for gain and loss might be chosen differently because they might have different perceived sensitivities. Throughout the paper we have chosen zero value for these thresholds because we have removed intra-block dependency when we apply contextual detail removal during the feature extraction stage. In the areas where $\Delta_{ij} > \delta_{ij}^+$, spatial 'activity gain' is indicated, and it is likely that this gain is proportional to the appearance of the blocking artifacts. On the other hand, regions with $\Delta_{ij} < -\delta_{ij}^-$ signify spatial 'activity loss' that correspond to blurring and/or disappearance of the contextual details.

If $n_x^+$ and $n_x^-$ denote the number of blocks identified as gain and loss, respectively, then the mean harmonic gain, $e^+$, and the mean harmonic loss, $e^-$, can be expressed as

$$e^+ = \frac{1}{n_x^+} \sum_i \sum_j e_{ij}^+$$  \hspace{1cm} (6)$$

and

$$e^- = \frac{1}{n_x^-} \sum_i \sum_j e_{ij}^-,$$  \hspace{1cm} (7)

respectively.

To take into account the non-linearity behaviour of the objective metric, or saturation effect in the vicinity of extreme values, a logarithmic expression can be used for the last two equations. However, we must also consider the possibility when either one of the above means exhibit zero value. Therefore, the logarithmic mean harmonic gain, $G$, and loss, $L$, are calculated as follows:

$$G = \log_{10} \left( 1 + e^+ \right)$$  \hspace{1cm} (8)$$

and

$$L = \log_{10} \left( 1 + e^- \right)$$  \hspace{1cm} (9)$$

where an offset of unity is added to ensure that the resulting values are always positive. The intermediate quality metric $LHS$ based on the discriminative analysis of the local harmonic strength is defined as

$$LHS = \frac{\alpha G + \beta L - \theta}{\gamma},$$  \hspace{1cm} (10)$$

where $\alpha$ and $\beta$ are the weighting coefficients, $\theta$ is an offset value and $\gamma$ is a normalisation factor. According to [12], there is a potential to use non-linear mapping of the objective output defined in Eq. (10) to a subjective rating by using a non-linear logistic function, with the constraint that the function remains monotonic over the full range of data. The function to transform the set of model outputs of Eq. (10) to a set of predicted mean opinion scores is defined as

$$LHS^* = \frac{b_1 - b_2}{1 + \exp \left\{ \frac{b_3 - LHS}{b_4} \right\}} + b_2$$  \hspace{1cm} (11)$$

where $b_1 - b_4$ are the coefficients of the 4-parameter logistic curve. Note that since $G$ and $L$ are defined in logarithmic values, their weighted sum can be considered as a weighted geometric mean of the distortions. The coefficients of Eq. (10) and Eq. (11) can be determined by training the model with any libraries of images. In this Paper, we have used the JPEG and the JPEG-2000 images from the LIVE image quality assessment database [13]. For the given subjective scores in the database, we found that the set of coefficients suitable for assessment in this work are $\alpha = 0.5563, \beta = 1.5, \theta = 2.219, \gamma = 1.565$ for Eq. (10) and $b_1 = 0.845, b_2 = 0.218, b_3 = 0.614,$ and $b_4 = -0.131$ for Eq. (11). By using these coefficients, we can even apply the proposed method to a wide range of images. Note that lower $LHS^*$ values correspond to lower picture quality, either due to the appearance of blockiness/tiling degradations on the picture or due to the loss of information as a result of blurring/smearing. On the other hand, higher $LHS^*$ implies that the picture in question is of higher quality because it contains a relatively small amount of degradations. The performance of the objective model is assessed through computation of the correlation coefficients between the transformed model outputs of Eq. (11) and any subjective ratings.

3. Results

3.1. Application to compressed images

In this section, we demonstrate the results of applying the proposed method to a set of images with various degrees of distortions. We used JPEG and JPEG-2000 encoded images, as well as low-pass-filtered (blurred) images taken from [14] which are different from the image database that we have used to train the model. In order to quantify various visible distortions, we classify the processed images into three classes of BAD, MODERATE, and GOOD. Note that the ratings presented in these images are not the results of subjective tests, but rather being drawn from observation of the images. These are depicted in Figures 3, 4, and 5 for JPEG, JPEG-2000 encoded, and blurred images, respectively. For the objective method, along with our own metric, the $LHS^*$, we have also evaluated these images by the traditional PSNR and the scaled-version of the output of the FR blockiness detector of [1] for comparison.

Inspections of Figures 3 through 5 indicate that the quality metric defined by the local harmonic strength ($LHS^*$) described in Eq. (11) is a more acceptable objective measure than the other two. For any coding (processing) distortion in these figures, $LHS^*$ well quantifies the quality of the images. This is also confirmed by Table 1 which summarises the objective quality comparison among the harmonic strength gain-loss method over the PSNR and the FR method of [1]. The results presented in the table reveal that our method is better than both the PSNR and the FR for subjective quality prediction of all kinds used. There is a strong agreement between our $LHS^*$ values and the visual quality of the images being studied in this work. On the contrary, the same agreement is not apparent with the PSNR and the FR outputs. It is true that the FR model performs relatively well for the JPEG coded images. This is understandable since the FR model of [1] was designed primarily to detect any blockiness on such images.
Unfortunately, other types of distortions may have gone unnoticed by this FR model. This is shown by the output of the FR model for JPEG-2000 images and blurred images, in which the quality score produced by the FR model fails to reflect any significant changes in the visual quality of the images due to severe distortion caused by the loss of sharpness and blurring.

On the PSNR weakness, for example, in Figures 3 and 5, despite the differences in subjective quality, their PSNR values are almost similar (around 26 dB and 25 dB for Figures 3 and 5, respectively). On the other hand in Figure 4, we can see that image with the least PSNR value (around 20 dB for the “Man” image in Figure 4(c)) is in fact the one with the highest visual quality. This is in contrast with the “Woman” image in Figure 4(a) which is the image with the highest PSNR rating of 27 dB for JPEG-2000 coded images, but most notably is the one which is severely distorted. Based on these examples, PSNR as an objective measure is considered less accurate for image quality evaluation.

Table 1 also shows the gain and loss quality sub-indicators $G^*$ and $L^*$, separately. These quality scores are calculated by the non-linear function of Eq. (11), but for each of the logistic function of $G$ and $L$, respectively; i.e., it is the output of the objective model when only one factor is taken into account (either gain or loss, but not both). Inspection of these two factors reveals why the LHS$^*$ gain/loss quality metric behaves well. For example for the JPEG pictures, we see that both gain and loss follow the quality as expected. This is because quality of JPEG images is defined by the quantizer step size, and at high quantizer step sizes, pictures become both blocky (gaining) and blurred (loss of details). For JPEG-2000, where distortion is mainly due to loss of details, but not blocking (due to the overlapping nature of dyadic wavelet filtering), the loss factor is more dominant over the gain factor for the overall quality determination, as agrees with the table. Finally, low pass filter without overlap (Figure 5) has mainly loss factor and the gain might be due to shift of signal level.

The contribution of the gain and loss calculation to the overall objective quality score as explained above can be observed from Figures 3 through 5. One can see that even in the images severely compressed by JPEG, (such as in Figure 3(a)), blockiness may not be the only distortion that appears as a result of the compression, but some losses may also be present in the picture. These losses could come from the flat areas on the image where textual details have been flattened due to coarse quantization. These flat areas can be found on the face, hands, some parts of the hair, and the background area of the “Tiffany” image in Figure 3(a). The amount of loss is significantly reduced in a moderately compressed JPEG image (Figure 3(b)) although the amount of blockiness can still be relatively high. All in all, in JPEG compressed images the amount of gain surpasses the amount of loss. For JPEG-2000 compressed images and also in the heavily filtered images, the situation is the other way around. In such images, the amount of loss is more dominant than the amount of gain. Note that some gain can still be present on the images; this may come from false edges (in Figure 4(a)), false texture (in Figure 4(b)), or probably edge thickening (in Figure 5(a)).

The provisional correlation values given in Table 1 are not the “true” performance measure of the objective models presented in this Paper, but included to give an intuition of the relative performance amongst the models. Pearson correlation is used to evaluate the accuracy of a model. For two sets of data, the objective scores and the output of the objective model, Pearson correlation shows the linear association between them. It is computed by pairing the two data sets and calculate their product-moment coefficient. Another performance criteria widely used to evaluate an objective model is the Spearman rank-order correlation (or Spearman correlation for short). Spearman correlation is used to evaluate the prediction monotonicity of a given objective model to show the extent of agreement between the subjective score and the objective model in terms of the sign of change in picture quality. For each datum in the two sets of data, a number according to the ranking of each datum in its own data set is assigned. From this, the difference in statistical ranking of the corresponding datum can be computed. A high degree of agreement expressed in the Spearman correlation means that the difference in statistical rank is small; i.e., the order of a large number of data in the two data sets are the same.

Even though the images in Figure 3–Figure 5 do not have their MOS (mean opinion score) values, we can see that the differentiation in terms of their visual quality is very obvious. The quality is clearly expressed as three distinctive qualitative attributes (e.g., BAD, MODERATE, and GOOD) such as presented in Table 1. If one can assign a certain quality value in an ascending order for each of these attributes, and use these values as the provisional subjective scores, then one may be able to compute the provisional correlation for each model. It is clear from the correlation presented in Table 1 that the proposed model using the LHS$^*$ outperforms the others by a significant margin. While PSNR and FR methods have weak Pearson correlation coefficients of around 0.5, LHS$^*$ has a high Pearson correlation of more than 0.9.

It is interesting to note that the Pearson correlation coefficient for gain alone is very similar to that of FR. This indicates that the gain parameter can identify blockiness as well as the FR does. On the other hand the loss parameter alone has a very high correlation coefficient of 0.92. This means in image quality assessment, the loss plays more role than the gain. However, combining the loss and the gain into a single value of LHS$^*$ brings the validity of the quality metric to a Pearson correlation coefficient of 0.91. Although this value is slightly less than that of the loss factor, we can see that the differentiation in quality shown by the LHS$^*$ is much better than that of the loss factor. Taking into account both the gain and loss factors gives the objective model LHS$^*$ superior ability to differentiate the visual quality.

Table 1 also shows the provisional Spearman correlation test for various quality metrics. The Spearman correlation for the PSNR quality metric is −0.22 which is very low. It tells us that PSNR does not indicate quality at all. The Spearman
Figure 3. Different perceived quality of block-coded images with similar PSNR. Images taken from [13].

Figure 4. Different perceived quality of images encoded by JPEG 2000 with similar PSNR. Images taken from [13].

Figure 5. Different perceived quality of blurred images with similar PSNR. Images taken from [13].
Table 1. Descriptive quality, PSNR (in decibel), blockiness detector (FR), and proposed discriminative local harmonic strength (LHS) measure for images encoded by JPEG, JPEG-2000, and low-pass filtered/blurred. The relative figure of merits of these metrics are given in terms of provisional Pearson and Spearman correlation for comparison purposes.

<table>
<thead>
<tr>
<th>Fig.</th>
<th>Process</th>
<th>Picture</th>
<th>Visual quality</th>
<th>PSNR (dB)</th>
<th>FR</th>
<th>Proposed RR Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>3(a)</td>
<td>JPEG</td>
<td>Tiffany</td>
<td>BAD</td>
<td>25.96</td>
<td>0.2653</td>
<td>G&lt;sup&gt;+&lt;/sup&gt; 0.8046</td>
</tr>
<tr>
<td>3(b)</td>
<td>Lake</td>
<td>MODERATE</td>
<td></td>
<td>25.89</td>
<td>0.3398</td>
<td>L&lt;sup&gt;+&lt;/sup&gt; 0.5946</td>
</tr>
<tr>
<td>3(c)</td>
<td>Mandril</td>
<td>GOOD</td>
<td></td>
<td>26.01</td>
<td>0.7128</td>
<td>LHS&lt;sup&gt;+&lt;/sup&gt; 0.2464</td>
</tr>
<tr>
<td>4(a)</td>
<td>JPEG</td>
<td>Woman</td>
<td>BAD</td>
<td>27.23</td>
<td>0.6328</td>
<td>G&lt;sup&gt;+&lt;/sup&gt; 0.8353</td>
</tr>
<tr>
<td>4(b)</td>
<td>2000</td>
<td>Barbara</td>
<td>MODERATE</td>
<td>25.87</td>
<td>0.6466</td>
<td>L&lt;sup&gt;+&lt;/sup&gt; 0.5197</td>
</tr>
<tr>
<td>4(c)</td>
<td>Man</td>
<td>GOOD</td>
<td></td>
<td>20.34</td>
<td>0.6730</td>
<td>LHS&lt;sup&gt;+&lt;/sup&gt; 0.3075</td>
</tr>
<tr>
<td>5(a)</td>
<td>Blurred</td>
<td>Woman</td>
<td>BAD</td>
<td>25.11</td>
<td>0.7535</td>
<td>G&lt;sup&gt;+&lt;/sup&gt; 0.8368</td>
</tr>
<tr>
<td>5(b)</td>
<td>Man</td>
<td>MODERATE</td>
<td></td>
<td>25.11</td>
<td>0.7031</td>
<td>L&lt;sup&gt;+&lt;/sup&gt; 0.5258</td>
</tr>
<tr>
<td>5(c)</td>
<td>Barbara</td>
<td>GOOD</td>
<td></td>
<td>25.12</td>
<td>0.9070</td>
<td>LHS&lt;sup&gt;+&lt;/sup&gt; 0.2922</td>
</tr>
</tbody>
</table>

Provisional Pearson Correlation: -0.51, 0.46, 0.42, 0.92, 0.91

Provisional Spearman Correlation: -0.22, 0.50, 0.75, 0.95, 0.95

The graph illustrates the scatter plot profile for images with different visual quality but exhibit equal PSNR values.

3.2. Application to images with other types of distortions

We also tested the proposed method against distorted images with other types of distortions than those mentioned in Sec. 3.1. These test images are taken from [14], from which we also use their subjective scores in terms of the mean subjective rank (MSR). The mean subjective rank is the average of visual quality differences between the original and the processed images, rated by a group of subjects. Hence the low value of subjective rank indicates higher visual quality, and lower visual quality is represented by higher value of the subjective rank. The distortions include mean shift, contrast stretching, impulse salt-pepper noise, multiplicative speckle noise, additive Gaussian noise, as well as blurring and JPEG compression that were tested before. These pictures are illustrated in Figure 7 along with various quality indicators. The results of the assessment, including the output of the full-reference model Universal Quality Index (UQI) given in [14], are also tabulated in Table 2. Note that the data in Table 2 have been presented in a descending order of their MSR; i.e., from good quality picture to the worst one.

As the quality of pictures in Figure 7 show as well as...
4. Conclusions and future works

This Paper has presented a framework to develop a reduced-reference evaluation model using local harmonic strength as a reduced-reference information. By using our proposed method, we were able to extract the gain and the loss information associated with the picture that can be used to indicate typical degradations such as blockiness or blurriness in an image. Although originally designed for detecting blockiness, we have shown that the method can also be used to detect blurriness and other types of distortions. The results indicate that the approach presented here is promising and shows a significant advantage over the traditional PSNR. Therefore, the $LHS^*$ method has a great potential to replace the PSNR for image quality evaluation. We have presented that the $LHS^*$ measure is useful for quality assessment of a wide range of images contaminated by various degradations, including additive, multiplicative, and impulsive noise and also by coding distortion in the block-based as well as the wavelet-based coded images; e.g., JPEG and JPEG 2000 images. The extension of the method for quality measurement of compressed digital video can also be developed by applying the same evaluation method presented here to considering video frames as individual digital images. In fact, the same kind of approach is also used by all the recommended models of objective video quality assessment described in the recent ITU-T Recommendation J.144 [16] which are based on frame quality measurement in one way or another. However, in video we also have to take into account factors such as movement that might have given different perceptual response to the gain and loss factors. Due to the simple approach we have taken in designing the quality model, a variety of services and applications for quality monitoring in the secondary distribution of video signals [12] can benefit from the method proposed in this Paper.

Acknowledgements

The authors wish to acknowledge financial support from the BT Exact, UK.

5. References


Table 2: Assessment of “Lena” image. Note: MSR = Mean Subjective Rank. Correlations values are calculated with respect to the subjective score of MSR.

<table>
<thead>
<tr>
<th>Fig.</th>
<th>Distortion Type</th>
<th>MSR from [14]</th>
<th>UQI from [14]</th>
<th>PSNR (dB)</th>
<th>FR</th>
<th>Proposed RR Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>7(b)</td>
<td>Mean Shift</td>
<td>1.59</td>
<td>0.9894</td>
<td>24.61</td>
<td>1.0000</td>
<td>0.8432</td>
</tr>
<tr>
<td>7(c)</td>
<td>Contrast Stretching</td>
<td>1.64</td>
<td>0.9372</td>
<td>24.61</td>
<td>0.8371</td>
<td>0.8297</td>
</tr>
<tr>
<td>7(d)</td>
<td>Salt-Pepper Noise</td>
<td>3.32</td>
<td>0.6494</td>
<td>24.60</td>
<td>0.2316</td>
<td>0.8400</td>
</tr>
<tr>
<td>7(e)</td>
<td>Speckle Noise</td>
<td>4.18</td>
<td>0.4408</td>
<td>24.61</td>
<td>0.3811</td>
<td>0.8375</td>
</tr>
<tr>
<td>7(f)</td>
<td>Gaussian Noise</td>
<td>4.27</td>
<td>0.3891</td>
<td>24.61</td>
<td>0.3707</td>
<td>0.8375</td>
</tr>
<tr>
<td>7(g)</td>
<td>Blurring</td>
<td>6.32</td>
<td>0.3461</td>
<td>24.63</td>
<td>0.7380</td>
<td>0.8359</td>
</tr>
<tr>
<td>7(h)</td>
<td>JPEG Compression</td>
<td>6.68</td>
<td>0.2876</td>
<td>24.80</td>
<td>0.1969</td>
<td>0.8358</td>
</tr>
</tbody>
</table>

Pearson Correlation: -0.94, 0.65, -0.61; Spearman Correlation: -1.00, 0.73, -0.68.
(e) Multiplicative speckle noise contaminated image, $PSNR = 24.61$, $LHS^* = 0.5234$, $MSR = 4.18$

(f) Additive Gaussian noise contaminated image, $PSNR = 24.61$, $LHS^* = 0.9959$, $MSR = 4.27$

(g) Blurred image, $PSNR = 24.63$, $LHS^* = 0.3433$, $MSR = 6.32$

(h) JPEG compressed image, $PSNR = 24.80$, $LHS^* = 0.3168$, $MSR = 6.98$

**Figure 7:** (continued)


