Pattern Recognition 45 (2012) 2193–2204

Contents lists available at SciVerse ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Content-based image quality metric using similarity measure of moment vectors

Kim-Han Thung *, Raveendran Paramesran, Chern-Loon Lim

Department of Electrical Engineering, University of Malaya, Lembah Pantai, 50603 Kuala Lumpur, Malaysia

ARTICLE INFO

Article history:
Received 8 December 2010
Received in revised form
30 September 2011
Accepted 1 December 2011
Available online 9 December 2011

Keywords:
Image quality assessment
Tchebichef moments
Similarity measure
Block classification

ABSTRACT

In this paper, the similarity of moment vectors between the test and the reference image blocks together with the result from the block classification are used in the formulation of an image quality metric (IQM). First, the reference and the test images are divided into non-overlapping 8×8 blocks and transformed into moment domain using Discrete Tchebichef Transform. The moment features are then used in two operations: the local quality index calculation and the image content (block) classification. The local quality index is obtained from the similarity measure of moment vectors between the reference and the test image blocks. Next, the content of each reference image block is classified into three types: "plain", "edge" and "texture", based on its moment energy level and moment energy distribution. The local quality indices obtained from all the image blocks are then averaged based on the block types to obtain three mean quality scores for each test image. The performance of these three mean quality scores and their combinations are studied using the LIVE database. The results show that the performance of the metric is significantly improved by combining the mean quality scores from the edge and texture image region. The best combination (the proposed metric) is then compared with five other IQMs using the LIVE database and four other independent databases. The results show that the proposed metric performs comparatively well for all the databases.

1. Introduction

Image is ubiquitous in modern communication. However, during the process like image acquisition, compression and transmission, different types of distortions will appear on the reproduced image. The acceptability of the reproduced image can be most accurately evaluated by human beings, but the subjective assessment process [1] is too time-consuming and impractical for most applications. Thus, it is crucial to have an objective image quality metric (IQM) that can predict the image quality as closely to human assessment as possible [2]. There are three types of objective IQM based on the availability of the undistorted reference image. Among them, the *full-reference quality assessment* (FRQA) is the most fundamental and popular approach, where the quality of a test image is obtained based on the comparison with the reference image. It is generally believed that the understanding of image quality gained from the FRQA, could benefit the more advanced quality assessment techniques where the reference image is only partially available (reduced reference) or not available (no-reference).

The simplest solution to the FRQA problem was by comparing the pixel intensity between the reference and the test image directly using some kind of mathematical models [3,4], such as mean squared error, peak signal-to-noise ratio (PSNR), Minowski Error, etc. Though this approach is simple, straight forward and mathematically tractable, they do not correlate well with perceived quality measurement [5,6].

Another group of studies address the inadequacy of the mathematical models by incorporate HVS features in their algorithms [7–11]. In these metrics, HVS is modeled using bottom up approach and the quality degradation is measured based on the visibility of distortions. Though these HVS-based IQMs are universally accepted, they leave some unexamined and unanswered questions salient to the development of a viable metric [12]. For example, its definition of quality degradation is questionable as not all visible errors have the same objectionableness, and it is unclear whether the parameters obtained from the simple psychophysical experiments can be generalized to real life images [12].

The modern IQMs on the other hand incorporate the HVS features using the top-down approach, thus largely averted the assumptions made by the early HVS-based IQMs. Structural Similarity image index (SSIM) [12] and its variant multi-scale SSIM (MS-SSIM) [13], for example, define the image degradation as the structural change of the two images, based on the observation that human perception is highly adapted to the highly structured natural

* Corresponding author. Tel.: +60379675253.
E-mail addresses: henrythung@yahoo.com (K.-H. Thung), ravee@um.edu.my (R. Paramesran), limchernloon@gmail.com (C.-L. Lim).

0031-3203/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.patcog.2011.12.001

images. In these metrics, the structural distortion is estimated by local cross-correlations of the two images. However, this definition is not widely accepted [14] and it was shown in [15] that SSIM is directly related to the MSE. Visual fidelity information (VIF) [14,16] is another alternative to assess image quality based on natural scene statistics. However, in [17], it was also demonstrated that SSIM and VIF metrics are equivalent. Nevertheless, experimental study using LIVE database shows that the quality predictions of SSIM and VIF correlate well with the subjective ratings.

Recently, we have introduced a moment correlation index as a new IQM [18]. In this metric, the moment correlation indices are applied to nine low order moment values for every 8×8 non-overlapping image block, and the final quality score is obtained by averaging all the moment correlation indices. Two orthogonal moments (Tchebichef and Krawtchouk) have been investigated in [18], and the result shows that they can be used as alternative features to quantify local structural changes between two images.

Generally, most of the metrics perform well for single type of distortion, but when more distortion types are involved, the performance of the quality metrics drop. This is probably because human perception has different objectionableness for different types of distortion, and this difference cannot be addressed efficiently using the local quality index alone.

To improve the correlation of the metric with human subjectivity, some researchers proposed to assign appropriate weights to the local quality indices. The weighting criteria that have been used (for the SSIM index) including the local energy, local quality/distortion, local information [19], visual fixation [20] and image content [21]. In these metrics, heavier weights are assigned to local quality indices at image areas with more visual importance.

In this paper, we use another approach to improve the prediction of the image quality index. The "weights" are determined from the relationship among the image content, local quality index and the human perception for different types of distortion, rather than from the visual importance aspect of the image content [21]. It is intuitively true that the impact of distortion in an image depends not only on the image content but also on the source of the distortion. Some distortions are more objectionable for certain types of image content. For example, the plain region of the image is more sensitive to Gaussian additive noise than the texture region; the edges are more susceptible to distortion caused by blurring than a plain region; while the complex texture image region is less sensitive to high frequency noise (contrast masking effect). Thus, the performance of the quality metric largely depends on how well the local quality index reflect these perceptual difference at different image regions. This has motivated us to subdivide the image into blocks and to classify them into three types, namely "plain", "edge" and "texture" (PET) based on the types of content. The performances of the local quality index at these three types of image regions are then examined for various types of distortion. Local quality indices from the image regions that perform well (correlate highly with subjective ratings) are then combined to get a final quality metric.

A content-based IQM using similarity measures of moment vectors is proposed in this paper. The main features of this metric are summarized below:

1. *Tchebichef moments are used as image features in local quality index formulation and image content (block) classification.* Tchebichef moments are chosen in this paper due to its desirable properties described in Section 2. First, local quality index is formulated using similarity measure of (Tchebichef) moment vectors. Second, the orientation property of the Tchebichef moments is utilized to classify the image block into three types (PET) [22]. In this way, computation time is saved as no other operators (e.g. sobel operator) are needed to extract edge and texture information about the image as in [21].
2. "*Weighting scheme*" *is driven by the performance of the local quality indices at three image regions rather than the visual importance*. As pointed out earlier, direct averaging the local quality indices of all the image blocks to a final quality score does not correlate satisfactorily well with the human subjectivity [18]. Thus, a "weighting scheme" based on the types of image content of the image block is introduced. First, the local quality indices are averaged into three mean quality scores based on the PET block types. The correlation of these three mean quality scores with the subjective ratings are then examined using the LIVE database. This is to find out how well and how differently the quality index performs at the three (PET) image regions for different types of distortions. The mean quality scores that correlate similarly and highly with the subjective scores are combined to get a final quality metric that works well for various types of distortion.

This paper is organized as follows. Section 1 is the introduction which critically reviews the precursors to this study and highlights the main features of the proposed metric. This section describes the Tchebichef moments, the image features used in this metric. Section 3 describes another theoretical underpinning: block classification using Tchebichef moments. Section 4 describes the design details of the proposed metric, including the similarity measures of moment vectors, the three mean quality scores and how they are combined. Section 5 consists of the analysis and discussion of experimental results including a comparison of the results with those obtained by other state-of-the-art metrics like SSIM and VIF. We conclude our paper in Section 6.

2. Tchebichef moments

Tchebichef moment was chosen for this study owing to some of its desirable properties [18]:

- It is discrete orthogonal, and thus can be used directly on the discrete image without discrete approximation error and coordinate system transformation [23].
- It is an efficient image descriptor, as it can be used to represent an image without information redundancy in the moment set and is able to detect any small variation in the pixel intensity.
- It can be used to effectively extract statistical and structural information about an image based on successful application of moment features in pattern recognition [24,25], object classification [26,27], edge detection [28], image reconstruction [29] and data compression [30,31]. As the moments are computed over every image pixel, the image information obtained from the moments is global. When it operates on an image block, it extracts global information about the image block and thus provides us spatial "global" information about an image.
- It has orientation property. The Tchebichef moment matrix of the edge image block is often oriented, and thus can be utilized as block classifier.

2.1. Computation of Tchebichef moments

The Tchebichef moments of order $(p+q)$ of an image with intensity function $f(x,y), x \in 0,1,\ldots,M-1, y \in 0,1,\ldots N-1$ is defined as [32]:

$$T_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \tilde{t}_p(x;M)\tilde{t}_q(y;N)f(x,y) \quad (1)$$

where $\tilde{t}_p(x;M)$ and $\tilde{t}_q(y;N)$ are the normalized Tchebichef polynomials given by

$$\tilde{t}_p(x;M)=\frac{t_p(x;M)}{\sqrt{\rho(p;M)}},\quad \tilde{t}_q(y;N)=\frac{t_q(y;N)}{\sqrt{\rho(q;N)}} \tag{2}$$

and $t_n(x;N)$ is nth-order N-point Tchebichef polynomial which is defined as

$$t_n(x;N)=n!\sum_{k=0}^{n}(-1)^{n-k}\binom{N-1-k}{n-k}\binom{n+k}{n}\binom{x}{k} \tag{3}$$

For the rest of the paper, we will use $\tilde{t}_n(x)$ to represent $\tilde{t}_n(x;N)$. $\tilde{t}_n(x)$ is the orthonormal version of Tchebichef polynomials and it can be calculated using the recurrence relation below [33]:

$$\tilde{t}_n(x)=\alpha_1(2x+1-N)\tilde{t}_{n-1}(x)+\alpha_2\tilde{t}_{n-2}(x),\quad n=2,3,\ldots N-1;\quad x=0,1,\ldots N-1 \tag{4}$$

where

$$\alpha_1=\frac{1}{n}\sqrt{\frac{4n^2-1}{N^2-n^2}}$$

$$\alpha_2=\frac{1-n}{n}\sqrt{\frac{2n+1}{2n-3}}\sqrt{\frac{N^2-(n-1)^2}{N^2-n^2}} \tag{5}$$

The initial conditions for the above recursion are given as

$$\tilde{t}_0(x)=1/\sqrt{N}$$

$$\tilde{t}_1(x)=(2x+1-N)\sqrt{\frac{3}{N(N^2-1)}} \tag{6}$$

2.2. Matrix form implementation

In Matlab, it is easier to calculate Tchebichef moments (1) in matrix form. The set of Tchebichef moments of a square image $\boldsymbol{I}=\{f(x,y)\}_{x,y=0}^{x,y=N-1}$ up to order $(p+q)$ can be determined by

$$\boldsymbol{T}=\boldsymbol{P}\boldsymbol{I}\boldsymbol{Q}^T \tag{7}$$

where $\boldsymbol{P}$ and $\boldsymbol{Q}$ are transformation matrices which contain Tchebichef polynomials of orders up to p and q respectively,

$$\boldsymbol{P}=\begin{pmatrix}\tilde{t}_0(0) & \ldots & \tilde{t}_0(N-1)\\ \vdots & \ddots & \vdots\\ \tilde{t}_p(0) & \ldots & \tilde{t}_p(N-1)\end{pmatrix},\quad \boldsymbol{Q}=\begin{pmatrix}\tilde{t}_0(0) & \ldots & \tilde{t}_0(N-1)\\ \vdots & \ddots & \vdots\\ \tilde{t}_q(0) & \ldots & \tilde{t}_q(N-1)\end{pmatrix} \tag{8}$$

and $(.)^T$ denotes the transpose of the matrix. The complete set of the Tchebichef moments of image $\boldsymbol{I}$ can be obtained by using (7) and substituting values $p=N-1$ and $q=N-1$ in (8).

2.3. Tchebichef basis images

Tchebichef moments of a square image $\boldsymbol{I}=\{f(x,y)\}_{x,y=0}^{x,y=N-1}$ as in (1) can be viewed as the projection of the image $\boldsymbol{I}$ on the basis image Φ_{pq}, which is the product of vectors $\tilde{t}_p$ and $\tilde{t}_q$:

$$\tilde{\boldsymbol{t}}_{\boldsymbol{p}}=[\tilde{t}_p(0)\ \tilde{t}_p(1)\ \ldots\ \tilde{t}_p(N-1)]$$

$$\tilde{\boldsymbol{t}}_{\boldsymbol{q}}=[\tilde{t}_q(0)\ \tilde{t}_q(1)\ \ldots\ \tilde{t}_q(N-1)]$$

$$\boldsymbol{\Phi}_{\boldsymbol{pq}}=[\tilde{\boldsymbol{t}}_{\boldsymbol{p}}]^T\tilde{\boldsymbol{t}}_{\boldsymbol{q}} \tag{9}$$

In other words, Tchebichef moment T_{pq} measures the correlation between the basis image $\boldsymbol{\Phi}_{\boldsymbol{pq}}$ and image I. It records a high positive value if there is a strong similarity between them, i.e. the higher the value, the greater the similarity. The complete set of Φ_{pq} of an 8×8 image block is shown in Fig. 1, where it can be seen that the basis images are changed from low spatial frequency (around the upper left hand corner of the figure) to high spatial frequency (around the lower right hand corner of the figure) when the orders of the moment are increased.

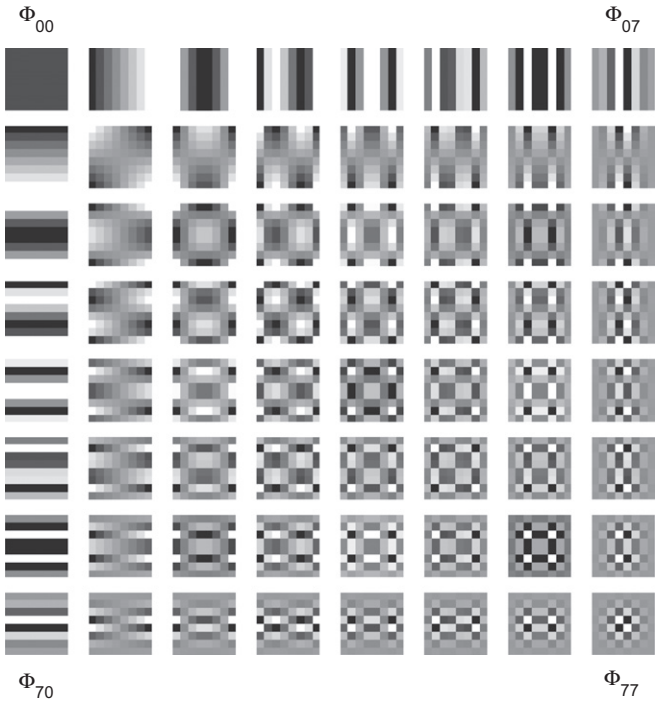

Fig. 1. The 8×8 Tchebichef basis images Φ_{pq}.

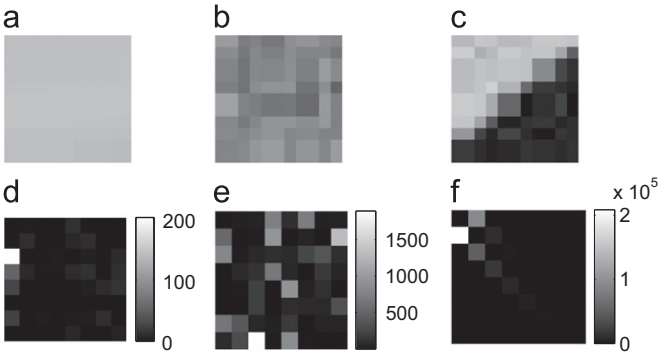

Fig. 2. Sample 8×8 image blocks for (a) "Plain", (b) "Texture" and (c) "Edge", with their corresponding 8×8 moment energies (e.g. moment square values) shown underneath in (d)–(f) respectively.

3. Moment-based image content block classification

Different types of image content have different patterns in their moment blocks. Fig. 2 shows three sample image blocks with their corresponding moment energies (moment square) distribution. T_{00} (DC component) is set to zero so that the other moment components can be seen easily. From the figure, several observations can be made:

- "plain": lower moment energies compared with "edge" and "texture".
- "edge": moments are high in energies and oriented.
- "texture": moments are high in energies and distributed more evenly or randomly in the moment block.

The orientation property of the moment block for edge image can be foreseen by examining the Tchebichef basis images in Fig. 1. In the figure, all the basis images of the first row (except the first one) look like vertical stripes with increasing frequency from left to right. Since the moment value gives us the correlation strength between the Tchebichef basis image and the image block [34], the image block with vertical edges will have high moment value at these locations (first row of the moment block). Similarly, the first column of the moment block is related to the horizontal edge, while the diagonal elements of the moment block are related to the diagonal edge.

3.1. Edge features extraction

The moments distribution of five ideal edge image blocks with various gradient angles are shown in Fig. 3. From the figure, it can be seen that moment energy (T^2_{pq}) block has fewer significant components (visible low order moments) than the absolute moment ($|T_{pq}|$) block. Thus, moment energies are preferable in classification algorithm as fewer components (and thus less computation) are needed to detect an edge image block.

Four regions in the Tchebichef moment block are defined in Fig. 4. They are labeled as "DC", "V", "H" and "D", representing the DC component, the vertical edge, the horizontal edge and the diagonal edge regions respectively. The unlabeled region (around the lower right corner) contains higher order moments and is expected to give us some textural information regarding the image. VE, HE and DE are defined as the sum of square Tchebichef moments at three edge regions ("V", "H" and "D") respectively:

$$HE = \sum_{i=1}^{3} T_{i0}^2, \quad VE = \sum_{j=1}^{3} T_{0j}^2$$

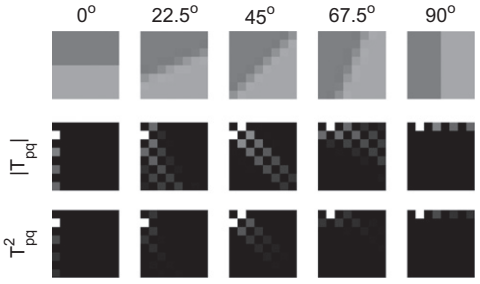

Fig. 3. Ideal edge image block with their corresponding moment block.

$$DE = \sum_{i=1}^{2} T_{i0}^2 + \sum_{i=0}^{2} T_{i1}^2 + \sum_{i=0}^{3} T_{i2}^2 + \sum_{i=2}^{3} T_{i3}^2 \tag{10}$$

where T_{ij} represents the moment value of the i-th row and j-th column of the moment block as shown in Fig. 4. These values are normalized by dividing them by the sum of square of all non-DC moment values (SSM). SSM can be calculated more efficiently by using Parseval's Theorem, which states that the sum of square of a function is equivalent to the sum of square of its transform. Thus, SSM is equivalent to the sum of square of the mean removed luminance intensity of the image block:

$$SSM = \left(\sum_{i,j=0}^{7} T_{ij}^2 \right) - T_{00}^2 \equiv \sum_{i,j=0}^{7} [I_b(i,j) - \overline{I_b}]^2 \tag{11}$$

where I_b is the luminance intensity of an image block and $\overline{I_b}$ is its mean.

A plot of three normalized moment energies (HE/SSM, VE/SSM and DE/SSM) for ideal edge image blocks of different angles are shown in Fig. 5. As shown in the figure, the three normalized moment energies cover different range of edge angle. HE/SSM and VE/SSM can be used to detect horizontal and vertical edge up to $\pm 20^\circ$ respectively, while DE/SSM covers the rest of the edge angles. The maximum of these three values can be used as a single edge detection feature, which is called maximum edge ratio for short. From our experimental study, this value is generally more than 0.8 for an ideal edge.

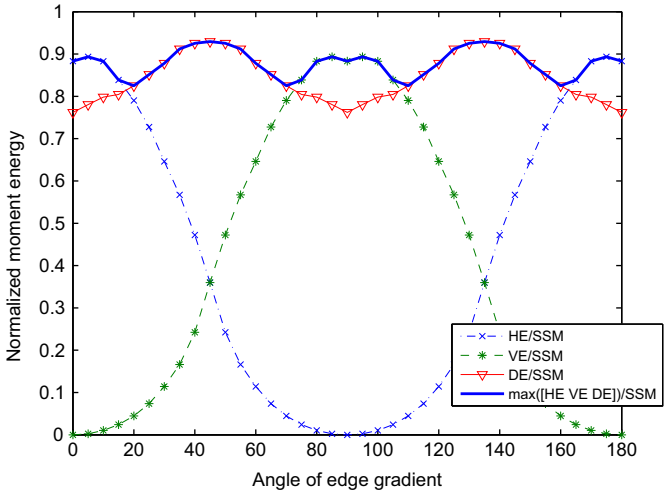

Fig. 5. Normalized moment energy for ideal edge image block at various gradient angle.

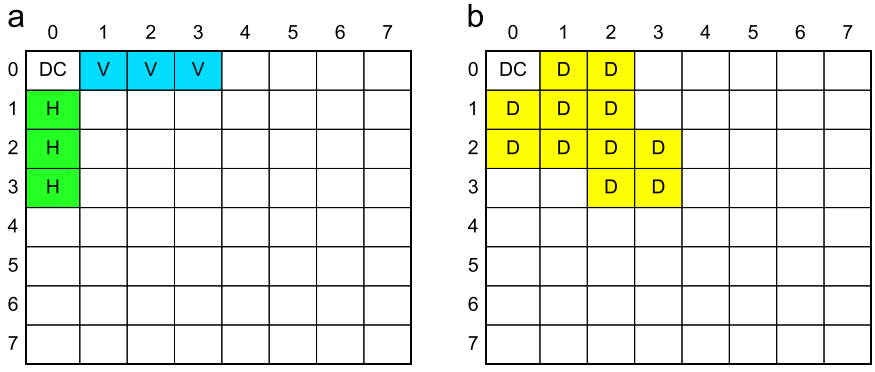

Fig. 4. Four regions of 8×8 Tchebichef Moment Blocks: (1) DC, DC component, (2) V, vertical edge Component, (3) H, horizontal edge component, (4) D, diagonal edge component. Each element in the p-th row and q-th column of the moment block represents moment feature T_{pq}.

3.2. Maximum edge ratio for distorted edge and random image block

The maximum edge ratios are examined when the ideal edges are distorted with additive Gaussian white noise. From our experimental study, when the distortion level is increased, the image block becomes more texture like and the maximum edge ratio will generally drop.

The maximum edge ratios for image blocks with random pixel value and minor smoothing (simulating texture image block) are examined in this study as well. From the experimental study, the maximum edge ratio for this case is generally less than 0.6.

When the edge or texture image block is smoothed, the maximum edge ratio increases as the moment distribution is shifted to the lower order side. Concurrently, the SSM value decreases as there is less detail information in the image block. When the image block is smoothed continuously, it will reach a state where the image block become plain. The SSM value when this happens is observed to be around 4000.

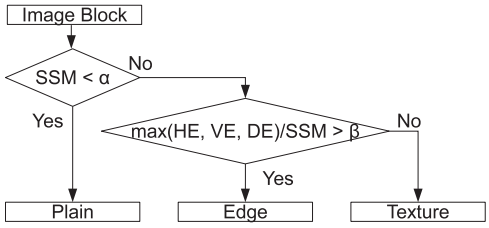

Fig. 6. Block classification using moment values.

Table 1
Block classification confusion matrix.

Test	Classification output			
	P	E	T	Total
P	99	1	0	100
E	0	96	4	100
T	4	9	87	100

3.3. Classification algorithm and its performance

The block classification algorithm is summarized in Fig. 6. The α value is set to 4000. Considering the edge in the real image which has different kind of background and shape, the β threshold value is set to 0.7 to differentiate between edge and texture blocks.

A sample of 300 plain, edge and texture blocks (100 each) were randomly extracted from the plain (e.g. smooth surface, sky, etc.),

Table 2
The pictures from LIVE database and their corresponding ratio of "plain" (r_p), "edge" (r_e) and "texture" (r_t) blocks. The pictures are sorted in ascending order of r_p.

Pictures	r_p	r_e	r_t	Pictures	r_p	r_e	r_t
1 building2	0.03	0.32	0.65	16 sailing4	0.41	0.20	0.39
2 stream	0.04	0.16	0.80	17 carnivaldolls	0.42	0.33	0.25
3 studentsculpture	0.06	0.33	0.61	18 house	0.42	0.21	0.37
4 bikes	0.16	0.42	0.42	19 lighthouse	0.45	0.16	0.39
5 buildings	0.17	0.35	0.48	20 ocean	0.48	0.15	0.37
6 cemetry	0.18	0.34	0.48	21 statue	0.49	0.35	0.16
7 rapids	0.22	0.34	0.44	22 lighthouse2	0.51	0.18	0.31
8 dancers	0.23	0.35	0.42	23 womanhat	0.53	0.23	0.24
9 coinsinfountain	0.24	0.45	0.31	24 sailing3	0.55	0.28	0.17
10 flowersonih35	0.24	0.22	0.54	25 monarch	0.57	0.38	0.05
11 manfishing	0.25	0.33	0.42	26 sailing2	0.60	0.23	0.17
12 woman	0.29	0.27	0.44	27 plane	0.63	0.15	0.22
13 paintedhouse	0.31	0.27	0.42	28 caps	0.68	0.17	0.15
14 sailing1	0.34	0.15	0.51	29 parrots	0.75	0.18	0.07
15 churchandcapitol	0.37	0.30	0.33				

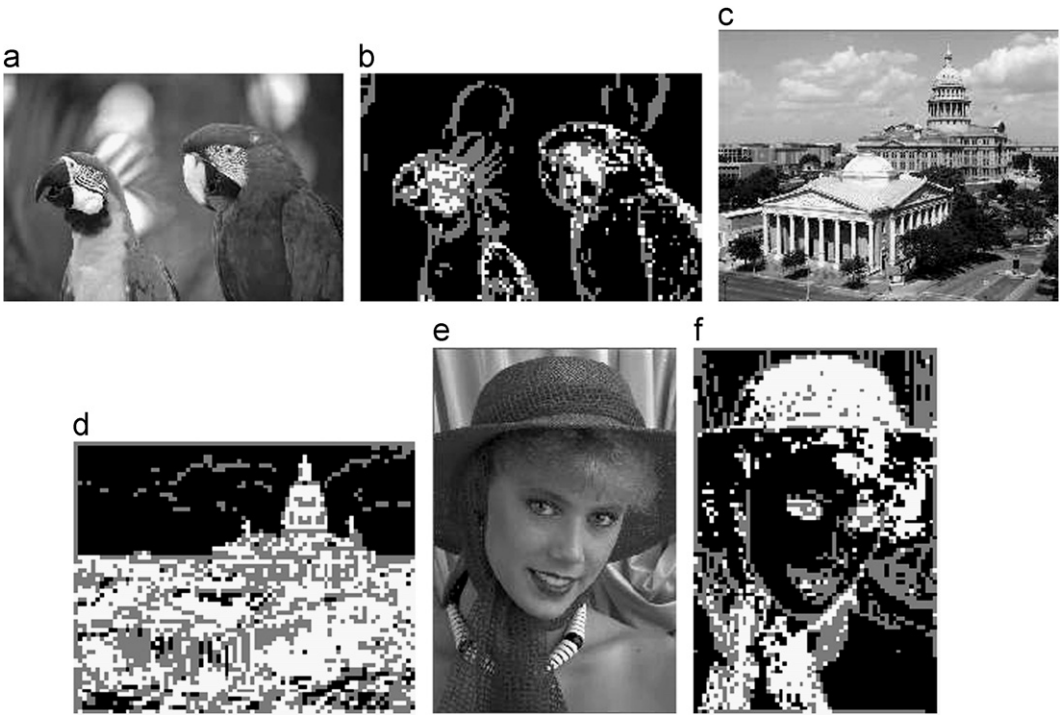

Fig. 7. Block classification output of "parrots", "churchandcapitol" and "womanhat". Image (a), (c) and (e) on the left are the luminance level of the original images, while images (b), (d) and (f) on the right are the corresponding 8×8 block classification output. The plain, edge and texture blocks are represented by black, gray and white colors respectively. The ratios of the plain (P), edge (E) and texture (T) blocks of these images are shown in Table 2.

edge (e.g. boundaries of objects) and texture (e.g. grasses, trees, etc.) regions of 10 different images manually. Most of the images are taken from the LIVE database. The classification confusion matrix using these extracted image blocks is shown in Table 1. The table shows that the algorithm can differentiate most of the image blocks. Some of the edge image blocks are misclassified into texture blocks because it contains textural pattern. Similarly, some of the texture blocks with edge like features were misclassified as edge blocks.

A pretest, which involved the block classification of three images from the LIVE database [35], namely the "parrots", "churchandcapitol" and "womanhat", was run to determine the viability of the algorithm. Fig. 7 shows the results of this pretest. The results validate the algorithm designed showing that it can basically classify the image block into plain, edge and texture. Table 2 shows the ratio of the "plain", "edge" and "texture" blocks of all the 29 images from the LIVE database.

4. Proposed IQM: content-based quality metric using similarity measure of moment vector

The design of the proposed metric is shown in Fig. 8. The process of getting the final quality score for a test image is summarized below:

1. The reference and the test images are first subdivided into non-overlapping 8×8 image blocks.
2. Each image blocks are transformed into moment domain using Discrete Tchebichef Transform (DTT).
3. Moments of the reference image blocks are used to classify the image blocks into either "plain", "edge" or "texture" (details described in Section 3).
4. Local quality indices (S_i) are computed from the similarity measure of moment vectors between the test and the reference image blocks (Eq. (15)).
5. Local quality indices are averaged based on the types of image content of the image block. Three mean quality scores (S_p, S_e and S_t) are obtained for each test image, one from each image regions (Eq. (16)).
6. Mean quality scores are combined to produce the final quality score. The combine function used is driven by the performance of the mean quality scores.

The formulation of the local quality index and mean quality score are discussed in details in the following subsections.

4.1. Similarity measure of moment vector

The moment correlation index in [18] is reformulated in this paper. In [18], the moments values are implicitly assumed to be equally important as all the moment correlation indices of all moment orders are averaged over all the image blocks. However, not all low order moments are equally important at all image blocks. For example, for an image block with a horizontal edge, only certain moment values with the right orientation have significant large value. The other moment values are small and negligible. As each moment correlation index gives a value from -1 to 1, these negligible small moment values will also contribute some significant local quality indices which will be averaged to a final quality score. Using moment vectors at each image block, the dominant moment features will rule out the insignificant moment values automatically and avoid this instability problem. The following shows the detail formulation of the similarity measure of moment vectors as local quality index.

Tchebichef moments of the reference image and the distorted image can be seen as components of vectors **a** and **b** respectively. Let $\{T_{01}, T_{10},\ldots,T_{pq}\}$ denote the (non-DC) moment values computed from a 8×8 block of the reference image, and $\{T'_{01},T'_{10},\ldots,T'_{pq}\}$ the corresponding (non-DC) moment values from the distorted image, then the two moment vectors are given as:

$$\mathbf{a} = (T_{01},T_{10},\ldots,T_{pq})$$

$$\mathbf{b} = (T'_{01},T'_{10},\ldots,T'_{pq}) \tag{12}$$

Any distortion in the image varies its moment features and cause some differences between vector **a** and **b**. The length of vector difference between the moment vector **a** and **b**, can be used to measure the degree of distortion of the test image block. The greater the vector difference **a**–**b** with respect to the length of **a** and **b**, the higher the distortion level of the test image is deemed to be. Thus the possible distortion measure that utilizing the length of the vector difference is $|\mathbf{a}-\mathbf{b}|/|\mathbf{a}|+|\mathbf{b}|$, which has a minimum of 0 when both vectors are the same, and a maximum of 1 when both vectors are opposite to each other. By using the equation $S = 1-D$, with S denotes the similarity measure and D denotes the dissimilarity measure, a corresponding similarity measure utilizing the length of vector difference is defined as:

$$S_{mv} = 1 - \frac{|\mathbf{a}-\mathbf{b}|}{|\mathbf{a}|+|\mathbf{b}|} \tag{13}$$

The range of this similarity measure using moment vector (S_{mv}) is [0 1]. When $\mathbf{a} = \mathbf{b}$ (as in perfect fidelity) $S_{mv} = 1$; when $\mathbf{a} = -\mathbf{b}$ (greatest distortion) $S_{mv} = 0$. When both the **a** and **b** are zero, the S_{mv} is set to 1 to avoid the singularity problem in (13).

Note that DC components (T_{00} and T'_{00}) are not included in (13), as these values are related to the mean intensity level of the image block and provide no structural information about the image block. These values are used as DC component similarity measure:

$$S_{dc} = \frac{2T_{00} \times T'_{00}}{T_{00}^2 + T_{00}'^2} \tag{14}$$

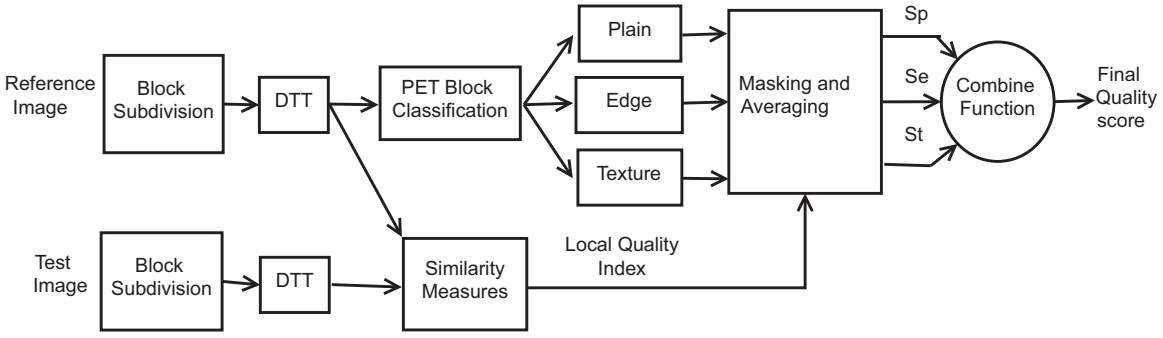

Fig. 8. Design of the IQM using Tchebichef moments in two operation: similarity measure of moment vectors and block classification.

In this equation, the range of the DC component similarity measure (S_{dc}) is [0 1]. When both T_{00} and T_{00} are the same, $S_{dc}=1$, when either one of them is zero, $S_{dc}=0$. S_{dc} is set to 1 when $T_{00}=T'_{00}=0$ to avoid the singularity problem in the equation. Note that this measure is the same as the luminance comparison in SSIM [12]; it gives us information about non-structural distortion (e.g. luminance shift). This measure is combined with S_{mv} using the following equation:

$$S_i=(S_{dc}(i)+S_{mv}(i))/2 \tag{15}$$

where S_i is the similarity measure of i-th block of the image, giving us the local quality index.

The similarity measure above is symmetrical, bounded (in between 0 and 1) and has a unique maximum. However, to be used as a good quality measure, it should be consistent with perceptual quality regardless of the types of distortion involved. This is taken care of through the inclusion of masking and averaging of the local quality indices (based on the output of block classification) as will be shown in the following subsection.

The order of moment used in the computation may also affect the performance of the quality metric. If we use too few moments, there might not be enough information to be used in quality assessment. If the order of the moments used is too high, two problems may arise: (1) there will be more computation involved and (2) the HVS probably is insensitive to the distortion features perceptible through the use of high order moments (as HVS, for example, is insensitive to high spatial frequency [11]). In this study, we decided to use all the moments in the 4×4 moment block, equivalent to a maximum moment order of six ($p+q=6$).

4.2. Mean quality scores of PET blocks

After the local quality indices are obtained from the similarity measures, they are masked and averaged based on the image block types. Three mean quality scores at the plain (S_p), edge (S_e) and texture (S_t) regions of the image are computed using the following equations:

$$S_p=\sum S_i/N_p,\quad i\in \text{ Plain blocks}$$

$$S_e=\sum S_i/N_e,\quad i\in \text{ Edge blocks}$$

$$S_t=\sum S_i/N_t,\quad i\in \text{ Texture blocks} \tag{16}$$

where N_p, N_e and N_t are the total image blocks for plain, edge and texture regions of the reference image.

These three mean quality scores (S_p, S_e and S_t) can be combined in many different ways to obtain a single quality score. In this study, they are combined using a linear function as the experimental study (which will be shown in the result section) shows that two of mean quality scores are correlate highly and similarly with the human subjectivity. The combine functions used in this study are shown in Table 3. The first four metrics are obtained by direct averaging of the mean quality scores of the respective image regions. The next four metrics are obtained by weight averaging the mean quality scores using the ratio of the image regions (from Table 2).

Table 3
Combine functions of the metric used in the study.

Metrics	Combine functions
S_{pe}	$(S_p+S_e)/2$
S_{pt}	$(S_p+S_t)/2$
S_{et}	$(S_e+S_t)/2$
S_{pet}	$(S_p+S_e+S_t)/3$
S_{per}	$(S_p\cdot r_p+S_e\cdot r_e)/(r_p+r_e)$
S_{ptr}	$(S_p\cdot r_p+S_t\cdot r_t)/(r_p+r_t)$
S_{etr}	$(S_e\cdot r_e+S_t\cdot r_t)/(r_e+r_t)$
S_{petr}	$(S_p\cdot r_p+S_e\cdot r_e+S_t\cdot r_t)/(r_p+r_e+r_t)$

5. Experimental results and discussion

The three mean quality scores (S_p, S_e and S_t) were first tested using four images from the LIVE database to find their relationship with the subjective ratings for various types of distortion. The result gives us the information about how well the local quality index performs at PET image regions under various types of distortion and gives us the preliminary clue on how the mean quality scores should be combined. In the second experiment, the three mean quality scores together with the combinations defined in Table 3 were tested on the entire LIVE database. The combine function that performed the best using the entire LIVE database was chosen as the proposed metric and compared with other objective quality metrics. The comparative study was first done using the LIVE image database and then followed by four other independent image databases.

5.1. LIVE database and performance metrics

LIVE database [35] has been used extensively in image quality assessment for cross-image and cross-distortion performance evaluations. The database consists of 29 original 24-bits/pixel color images and 779 distorted images representing five types of distortions, including 175 JPEG compressed, 169 JPEG-2000 compressed, 145 Gaussian white noisy (GWN), 145 Gaussian blurred (GB) and 145 fast fading (FF) Rayleigh channel noisy images. The realigned DMOS subjective scores [36] were used in this study to evaluate the performance the objective metric.

For a fair comparison between different objective IQMs, the images were preprocessed in the same way in this study. Grayscale versions of the images were used by using this pixel-wise transformation: $Y=0.2989R+0.5866G+0.1145B$, where Y, R, G and B denote the 8-bit grayscale (luminance), red, green, and blue intensities of the images respectively. Ideally, quality scores predicted by the objective IQM should be correlated with the subjective scores in a predictable and repeatable fashion. However, this relationship may not necessarily be linear and thus a logistic function is needed in a fitting procedure to provide a nonlinear mapping between the objective and subjective scores. The following logistic function with monotonic constraint was used for nonlinear fitting in this study [37,12]:

$$f(x)=\frac{\tau_1-\tau_2}{1+e^{(x-\tau_3)/\tau_4}}+\tau_2 \tag{17}$$

where the parameters τ_1, τ_2, τ_3 and τ_4 that minimized the sum of squared error between the transformed metric output $\{f(x)\}$ and the corresponding subjective ratings are obtained by using Matlab *nlinfit* function.

Three performance metrics were chosen to evaluate the proposed metric in this paper [2]. The first two metrics are the Pearson linear *correlation coefficient* (CC) and the *root-mean-square error* (RMSE) between DMOS and the objective scores after nonlinear regression, which give us the measure of prediction accuracy and prediction error respectively. The third metric is the *Spearman rank-order correlation coefficient* (SROCC) between the objective and the subjective scores, a measure of prediction monotonicity, which tells us the degree to which the model's predictions agree with the relative magnitudes of subjective quality ratings.

5.2. Performance of the three mean quality scores (S_p, S_e and S_t) using partial LIVE database

To study how the image content affects the quality index for different types of distortion, the correlations between the mean quality scores (S_p, S_e and S_t) and the subjective quality ratings are computed using four images from the LIVE database.

Four images with moderate ratio of PET region (ranked 12–15 in Table 2) were used for this test. Fig. 9 shows the scatter plot between the S_p, S_e and S_t and the subjective DMOS score. Different dot colors are used to represent quality scores from different images. The smooth curve in the figure is the best fitting curve using Eq. (17). The correlation coefficient (CC) after nonlinear fitting is computed and shown at the upper right corner of the plot. From the Fig. 9(a), it is obvious that the S_p has different correlation with the DMOS for different types of distortions (because the fitting curves are different in shape). Though the CC values were high for individual distortions, the performance of S_p dropped when all the distortions were considered. This implies that S_p predicts well for individual distortion but not for overall prediction. This "plain" image region is a typical example of a "visually important" image region [21] which has poor cross-distortion prediction due to its perception sensitivity difference to different types of distortion.

On the other hand, Fig. 9(b) and (c) shows that the similarity measures of moment vectors at the "edge" and the "texture" regions have roughly equal "sensitivity" to various types of distortion. This deduction is supported by the fact that the fitting curve of all the five distortions as well as the fitting curve of overall distortion have a similar shape. In all these curves, their nonlinear correlation coefficient (CC) values are high as well, ranged from 0.951 to 0.989. This suggests that the local quality

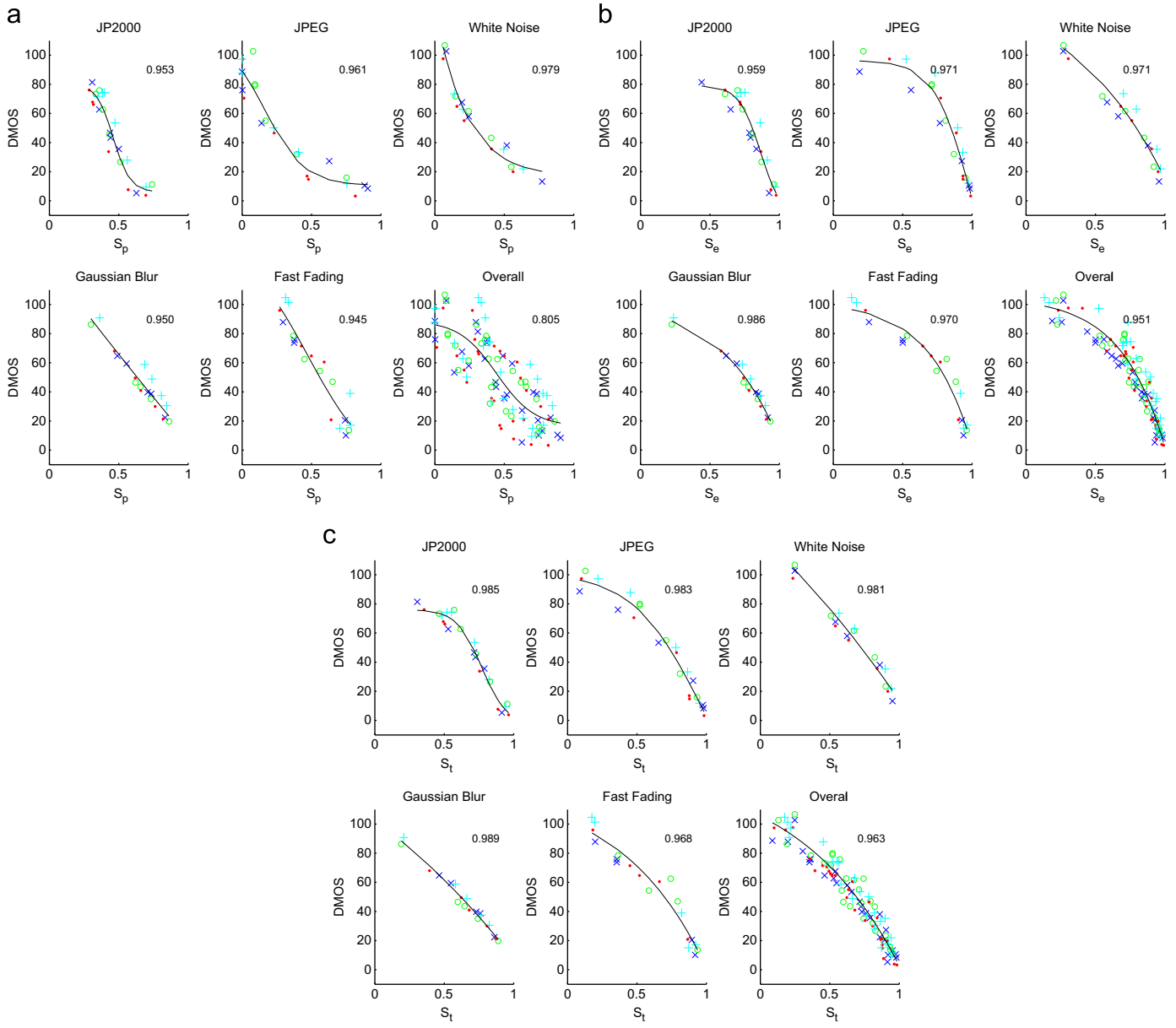

Fig. 9. Scatter plot with fitting curve between the mean similarity measure of moment vector at three image regions and the DMOS. (a) S_p VS DMOS (b) S_e VS DMOS (c) S_t VS DMOS. Four pictures (ranked 12–15 in Table 2) are used and they are shown with different dot colors. The corresponding correlation coefficient (CC) value after nonlinear regression is shown at the upper right hand corner of the plot.

indices at the "edge" and "texture" image regions can be combined linearly for a universal quality metric.

5.3. Performance of the mean quality scores and their combinations for the full LIVE database

From the previous results of the experimental study reported in Section 5.2 above, we know that the similarity measure of moment vectors at the plain region of the image is not suitable for a cross-distortion quality metric. While similarity measures at the edge and texture regions of the image correlate similarly with the subjective ratings for all the five distortion types in the LIVE database. Thus, a combination of mean quality scores from the edge and texture regions is expected to give us a better objective quality metric. To verify this hypotheses, the mean quality scores and their combinations as defined in Table 3 are tested on the entire LIVE database. Table 4 shows the results of the test. As expected, the metrics which did not involve S_p, highlighted in bold (S_e, S_t, S_{et} and S_{etr}) in Table 4, generally performed better than the others that involve S_p. As there are only marginal difference between the highlighted metrics in Table 4, only S_{etr} is chosen for the performance comparison study with other IQMs.

Table 4
Performance of the mean quality scores and their combine functions defined in Table 3. The results are based on the realigned DMOS scores [36] of 792 distorted images in LIVE database.

Metrics	Performance measure		
	CC	RMSE	SROCC
S_p	0.7994	16.42	0.7908
$\boldsymbol{S_e}$	**0.9569**	**7.93**	**0.9557**
$\boldsymbol{S_t}$	**0.9598**	**7.67**	**0.9598**
$\boldsymbol{S_{et}}$	**0.9638**	**7.28**	**0.9630**
S_{pe}	0.9082	11.44	0.8929
S_{pt}	0.9269	10.25	0.9196
S_{pet}	0.9452	8.92	0.9384
$\boldsymbol{S_{etr}}$	**0.9656**	**7.10**	**0.9643**
S_{per}	0.8501	14.39	0.8343
S_{ptr}	0.8884	12.54	0.8799
S_{petr}	0.9077	11.46	0.8972

5.4. Performance comparison of the proposed metric (S_{etr}) with other IQMs

The performance of the proposed metric (S_{etr}) was compared with five existing objective quality metrics, namely the (1) PSNR [37], (2) SSIM [12,38], (3) MS-SSIM [13], (4) VIF [14,35], and (5) IQM model using Tchebichef moments by Wee et al. (Tch-wee) [18]. All the metrics were used following the source code and default setting given by the respective authors.

All the metrics in compared are applied to grayscale version of the images and their performance is reported in Table 5. From the table, it can be seen that the proposed metrics (S_{etr}) performed competitively with other IQMs. VIF has the highest performance score for almost all the single distortion types (except white Gaussian noise) at first glance, but when all distortion types are used, the proposed metric seems to have performed better. A further statistical test was done to check whether these performance difference are statistically significant.

For this study, the Kolmogorov–Smirnov test (KS-test) was chosen to evaluate the statistical significance of the proposed metric (S_{etr}) in comparison with other objective metrics. It was applied to the residuals (prediction errors between the DMOS and the objective metric) of the proposed metric and another IQMs to determine whether they differ significantly. The KS-test, as it was deemed, is advantageous over the F-test used by other researchers [36]. The KS-test makes no assumption about the Gaussianity of the dataset. We had used the Kurtosis and Jarque–Bera test (JB-test) [39] to test the Gaussianity of the residuals of all the metrics and the results on Table 6 show that not all residuals are Gaussian distributed. This has strengthened the rationality of using KS-test instead of F-test.

The residuals of all the metrics were compared with the residuals of S_{etr} by using Matlab function *kstest2* and the results are shown in Table 7. The *null hypothesis* (H_o) said that the residuals from the two metrics come from the same distribution and are statistically indistinguishable (with 95% confidence level). The value 1 on the table means H_o is rejected, and thus implies statistical significance of the difference in metric performance in Table 5. In contrast, a value 0 on the table implies the acceptance of H_o: the performance difference of the objective metric with the proposed metric is statistically insignificant.

Table 5
Performance of the IQMs using the LIVE database based on the realigned DMOS scores [36]. The performance measures used here are Correlation Coefficient (CC) and RMSE after nonlinear regression fitting, and Spearman Rank Order Correlation Coefficient (SROCC).

Measure	Metric	JPEG2000	JPEG	White Noise	Gaussian blur	Fast fading	All
CC	PSNR	0.8996	0.8879	**0.9858**	0.7835	0.8895	0.8701
	SSIM	0.9665	0.9789	0.9701	0.9451	0.9489	0.9384
	MS-SSIM	0.9678	0.9823	0.9734	0.9515	0.9333	0.9402
	VIF	**0.9772**	**0.9864**	0.9839	**0.9743**	**0.9617**	0.9597
	Tch-wee [18]	0.9586	0.9431	0.9726	0.9516	0.9536	0.9187
	S_{etr}	0.9648	0.9812	0.9827	0.9686	0.9553	**0.9656**
RMSE	PSNR	11.017	14.653	**4.703**	11.478	13.015	13.468
	SSIM	6.475	6.503	6.790	6.036	8.991	9.445
	MS-SSIM	6.352	5.972	6.412	5.682	10.232	9.304
	VIF	**5.358**	**5.233**	4.998	**4.157**	**7.808**	7.675
	Tch-wee [18]	7.188	10.588	6.501	5.680	8.580	10.789
	S_{etr}	6.634	6.150	5.175	4.595	8.422	**7.103**
SROCC	PSNR	0.8955	0.8748	0.9854	0.7823	0.8907	0.8742
	SSIM	0.9614	0.9755	0.9694	0.9517	0.9556	0.9473
	MS-SSIM	0.9630	0.9812	0.9773	0.9550	0.9408	0.9516
	VIF	**0.9696**	**0.9840**	**0.9858**	**0.9728**	**0.9650**	0.9632
	Tch-wee [18]	0.9534	0.9384	0.9544	0.9593	0.9582	0.9183
	S_{etr}	0.9577	0.9760	0.9811	0.9643	0.9519	**0.9643**

The best performed metric at first glance is highlighted in boldface.

Table 6
Gaussianity test on the residuals of the metric. The first number is the kurtosis[a] value while the second number following the slash is the result of the Jarque–Bera test (JB-test).[b]

Metric	JPEG2000	JPEG	GWN	GB	FF	All
SSIM	2.71/1	3.12/1	2.14/1	2.92/1	2.83/1	2.95/1
MS-SSIM	2.71/1	3.42/1	2.26/1	2.93/1	4.97/0	3.66/0
VIF	3.03/1	3.89/0	2.65/1	2.63/1	3.28/1	2.83/1
Tch-wee [18]	4.06/0	3.39/1	2.35/1	3.58/1	3.26/1	2.51/0
S_{etr}	2.57/1	4.49/0	2.66/1	2.71/1	5.30/0	4.21/0

[a] Residuals with kurtosis values between 2 and 4 can be assumed as Gaussian distributed.
[b] For the JB-test, the number "1" means that the residuals are Gaussian distributed at 95% of confidence level while number "0" means otherwise.

Table 7
Result of the Kolmogorov–Smirnov test (KS-test) between residuals of metric S_{etr} and other metrics at 95% confidence level.

Metric	JPEG2000	JPEG	GWN	GB	FF	All
PSNR	0	1	0	1	0	1
SSIM	0	0	0	0	1	1
MS-SSIM	0	0	0	0	0	1
VIF	0	0	0	0	0	0
Tch-wee [18]	0	1	0	0	0	1

H_0: Residuals from both metrics are from the same distribution and statistically indistinguishable. H_1: Residuals from Both metrics are different. Number "1" means reject H_0, number "0" means accept H_0.

From Tables 7 and 5, it can be concluded that the proposed metric S_{etr} is statistically shown to be significantly better than the metrics PSNR, SSIM, MS-SSIM and Tch-wee using all distorted images from the LIVE database. However, Table 7 also shows that the performance of metric VIF, and S_{etr} are statistically indistinguishable. The scatter plot between the subjective DMOS and the predictions from the objective metrics (vif, Tch-wee and S_{etr}) are shown in Fig. 10.

5.5. Cross-database validation

The proposed metric S_{etr} was also tested on four other independent databases, including A57 database [40] (54 test images, 6 types of distortion), IVC database [41] (235 test images, 4 types of distortion), MICT database [42] (168 test images, two types of distortion) and TID2008 database [43] (1700 test images, 17 types of distortion). These databases consist of color and monochrome images of different contents, and distorted by different types of distortion such as JPEG, JPEG2000, Gaussian blur, white noise, coding artifacts, transmission error, mean shift, etc. These databases differ in terms of size of database, image contents, distortion types, number of subjects and subjective rating methodologies [44].

The performance of the objective metrics for these databases is shown in Table 8, where the *correlation coefficient* (CC) and the RMSE after nonlinear regression using Eq. (17) are reported together with the *spearman rank order correlation coefficient* (SROCC). The best performing metric for each database at first glance is shown in boldface. From the table, none of the objective models outperforms others for all the four databases. Some of the objective metric perform well for certain databases but poorly on other database. For example, VIF performs well for IVC, MICT and TID2008 databases but correlate poorly with the subjective scores for the A57 database. On the other hand, our proposed metric performs the best for the A57 database and consistently well for other databases, by recording a correlation coefficient (CC after nonlinear regression fitting) of at least 0.80.

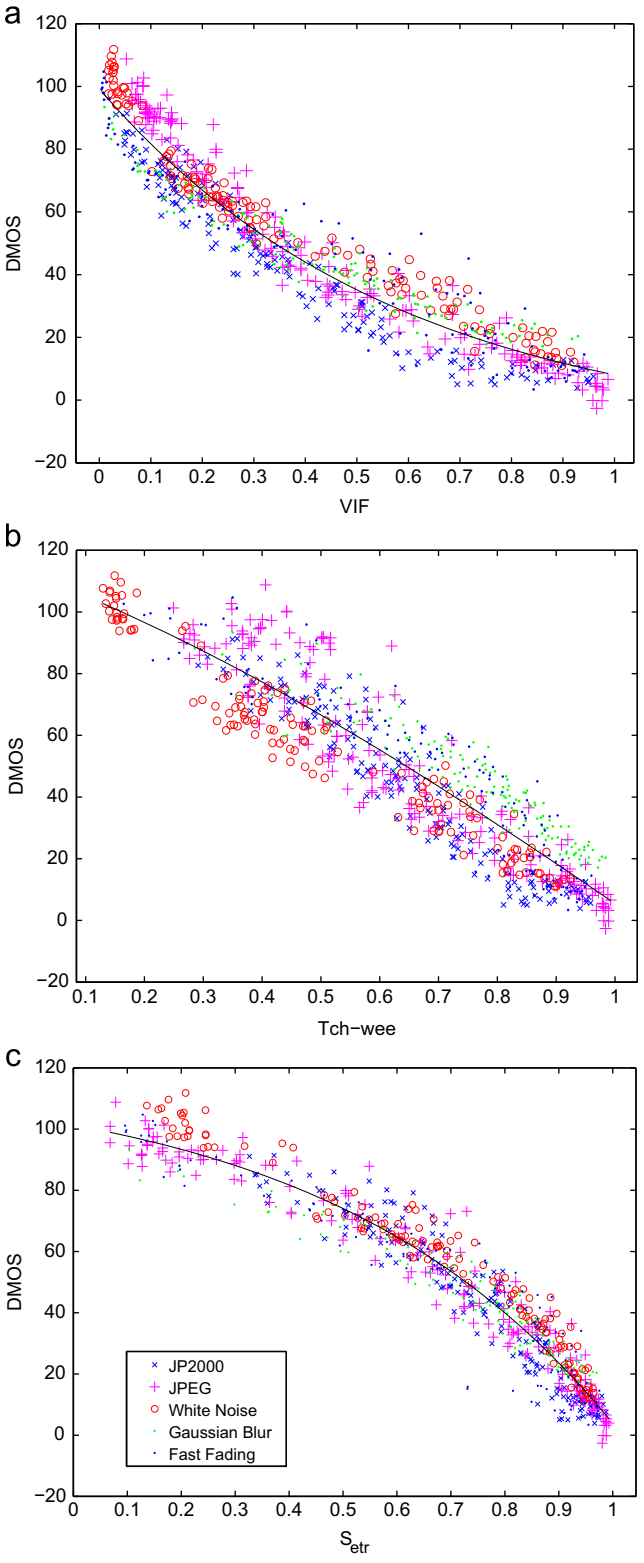

Fig. 10. Scatter plot between realigned DMOS scores [36] and predicted scores from objective IQMs: (a) VIF, (b) Tch-wee [18] and (c) S_{etr}. The datapoints with different distortions are distinguishable by its colors and shapes. The smooth curve is the nonlinear fitting curve using Eq. (17). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 8
Correlation Coefficient (CC), root mean square error (RMSE) and spearman rank-order correlation coefficient (SROCC) between the subjective scores and predicted scores of the objective metric for the A57, IVC, MICT and TID2008 database.

Measure	Database	PSNR	SSIM	MS-SSIM	VIF	Tch-wee	S_{etr}
CC	A57	0.6630	0.8021	0.8745	0.6148	0.5816	**0.9207**
	IVC	0.7191	**0.9116**	0.9108	0.9026	0.8771	0.9040
	MICT	0.6352	0.8883	0.8953	0.9143	**0.9242**	0.8764
	TID2008	0.5640	0.7717	0.7961	**0.8054**	0.7173	0.8024
RMSE	A57	0.1839	0.1467	0.1191	0.1937	0.1998	**0.0959**
	IVC	0.8466	**0.5007**	0.5031	0.5244	0.5853	0.5209
	MICT	1.1609	0.5495	0.5382	0.4808	**0.4468**	0.6027
	TID2008	1.1081	0.8535	0.8120	**0.7953**	0.9350	0.8008
SROCC	A57	0.6250	0.8056	0.8568	0.6207	0.6001	**0.9012**
	IVC	0.6885	**0.9018**	0.8974	0.8964	0.8586	0.8959
	MICT	0.4593	0.9024	0.9093	0.9185	**0.9316**	0.8770
	TID2008	0.5529	0.7752	**0.8091**	0.7496	0.6653	0.7483

The best performed metric at first glance is highlighted in boldface.

6. Conclusion

In this study, we have demonstrated the appropriateness and efficacy of a content-based image quality metric using moment vectors with a "weighting algorithm" that driven by the correlation of local quality indices at three image regions with the subjective ratings. The Tchebichef moments are used to accomplish two tasks in this proposed metric. First, the local quality index is formulated by using similarity measure of moment vectors between the reference and the test image blocks. Second, the image blocks of the reference image are classified into three types, namely "plain" (P), "edge" (E) and "texture" (T) based on their moment energy level and moment energy distribution. An experiment was done to examine how the objective metric at different image regions response to different types of distortion. This was simulated by examined the correlation between the three mean quality scores (S_p, S_e and S_t) from the classified three image regions with the subjective ratings for different types of distortion. It was found that S_p correlates differently with the subjective ratings for different types of distortion and thus is not suitable for a cross-distortion quality metric. On the other hand, S_e and S_t correlate well and similarly with the subjective ratings for various types of distortions. This suggests that both mean quality scores can be combined linearly to get a single final quality score. This was confirmed by the second experimental study, where it was found that a better quality metric is formed by excluding S_p and combining S_e and S_t in proportion to their ratio in the image. In the third experimental study, S_{etr} derived from the proposed quality model were compared with five existing objective metrics using the LIVE database. The results show that the proposed metrics and VIF are statistically indistinguishable but the proposed metrics are better than PSNR, SSIM, MS-SSIM and Tch-Wee for overall distortions. The proposed metric S_{etr} was also tested on four other independent databases, and the results show that the proposed metric performs competitively and consistently well for all the four databases. Since only part of the image is used in the proposed metric, this metric could be seen as a reduced reference quality metric but with a performance of a full-reference metric.

Acknowledgments

The authors would like to thank Dr. Nesam for language and grammar corrections and anonymous reviewers for their valuable and insightful comments which have improved the quality of this manuscript.

References

[1] Methodology for the subjective assessment of the quality of television pictures, Recommendation ITU-R BT.500-10, 1998.
[2] VQEG, Final Report from the Video Quality Experts Group on the Validation of Objective Models of Video Quality assessment, Phase II, Technical Report, 2003 ⟨http://www.vqeg.org/⟩.
[3] A.M. Eskicioglu, P.S. Fisher, Image quality measures and their performance, IEEE Transactions on Communications 43 (1995) 2959–2965.
[4] İsmail Avcıbaş, Image Quality Statistics and Their use in Steganalysis and Compression, Ph.D. thesis, Boğaziçi University, 2001.
[5] B. Girod, What's wrong with mean-squared error, in: A.B. Watson (Ed.), Digital Images and Human Vision, MIT Press, Cambridge, MA, USA, 1993, pp. 207–220.
[6] Z. Wang, A. Bovik, Mean squared error: Love it or leave it? a new look at signal fidelity measures, IEEE Signal Processing Magazine 26 (2009) 98–117.
[7] J. Mannos, D. Sakrison, The effects of a visual fidelity criterion of the encoding of images, IEEE Transactions on Information Theory 20 (1974) 525–536.
[8] J. Lubin, Digital Images and Human Vision, MIT Press, pp. 163–178.
[9] A.B. Watson, DCT quantization matrices visually optimized for individual images, in: B.E. Rogowitz (Ed.), SPIE Proceeding of Human Vision, Visual Processing, and Digital Display IV, 1913, SPIE Press, 1993, pp. 202–216.
[10] J.-B. Martens, L. Meesters, Image dissimilarity, Signal Processing 70 (1998) 155–176.
[11] T.N. Pappas, R.J. Safranek, Perceptual criteria for image quality evaluation, in: Handbook of Image and Video Processing, Academic Press, 2000, pp. 669–684.
[12] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Transactions on Image Processing 13 (2004) 600–612.
[13] Z. Wang, E.P. Simoncelli, A.C. Bovik, Multiscale structural similarity for image quality assessment. in: 37th Asilomar Conference on Signal, Systems and Computers, vol. 2, IEEE, 2003, pp. 1398–1402.
[14] H.R. Sheikh, A.C. Bovik, G. de Veciana, An information fidelity criterion for image quality assessment using natural scene statistics, IEEE Transactions on Image Processing 14 (2005) 2117–2128.
[15] R. Dosselmann, X. Yang, A comprehensive assessment of the structural similarity index, Signal, Image and Video Processing 5 (2011) 81–91.
[16] H.R. Sheikh, A.C. Bovik, Image information and visual quality, IEEE Transactions on Image Processing 15 (2006) 430–444.
[17] K. Seshadrinathan, A. Bovik, Unifying analysis of full reference image quality assessment, in: Fifteenth IEEE International Conference on Image Processing, 2008. ICIP 2008, IEEE, 2008, pp. 1200–1203.
[18] C.-Y. Wee, R. Paramesran, R. Mukundan, X.-D. Jiang, Image quality assessment by discrete orthogonal moments, Pattern Recognition 43 (2010) 4055–4068.
[19] Z. Wang, X.L. Shang, Spatial pooling strategies for perceptual image quality assessment, in: IEEE International Conference on Image Processing, IEEE, GA, USA, 2006, pp. 2945–2948.
[20] A.K. Moorthy, A.C. Bovik, Visual importance pooling for image quality assessment, IEEE Journal of Selected Topics in Signal Processing 3 (2009) 193–201.
[21] C.-F. Li, A.C. Bovik, Content-partitioned structural similarity index for image quality assessment, Signal Processing: Image Communication 25 (2010) 517–526 (Special Issue on Image and Video Quality Assessment).
[22] H.H.Y. Tong, A.N. Venetsanopoulos, A perceptual model for JPEG applications based on block classification, texture masking, and luminance masking. in: International Conference on Image Processing (ICIP), vol. 3, IEEE, 1998, pp. 428–432.
[23] R. Mukundan, S.-H. Ong, P.-A. Lee, Discrete vs. continuous orthogonal moments for image analysis, in: International Conference on Imaging Systems, Science and Technology (CISST), University of Canterbury, Computer Science and Software Engineering, Las Vegas, USA, 2001, pp. 23–29.
[24] J. Flusser, T. Suk, Pattern recognition by affine moment invariants, Pattern Recognition 26 (1993) 167–174.
[25] S.O. Belkasim, M. Shridhar, M. Ahmadi, Pattern recognition with moment invariants: a comparative study and new results, Pattern Recognition 24 (1991) 1117–1138.
[26] M. Tucceryan, Moment-based texture segmentation, Pattern Recognition Letters 15 (1994) 659–668.
[27] C.-C. Chen, Improved moment invariants for shape discrimination, Pattern Recognition 26 (1993) 683–686.
[28] S. Ghosal, R. Mehrotra, Orthogonal moment operators for subpixel edge detection, Pattern Recognition 26 (1993) 295–306.
[29] B. Bayraktar, T. Bernas, J. Paul Robinson, B. Rajwa, A numerical recipe for accurate image reconstruction from discrete orthogonal moments, Pattern Recognition 40 (2007) 659–669.
[30] O. Hunt, R. Mukundan, A Comparison of Discrete Orthogonal Basis Functions for Image Compression, in: Conference on Image and Vision Computing New Zealand (IVCNZ'04), University of Canterbury, 2004, pp. 53–58.
[31] R. Mukundan, Transform coding using discrete Tchebichef polynomials, in: Sixth IASTED International Conference of Visualization Imaging and Image Processing (VIIP 2006), Computer Science and Software Engineering, University of Canterbury, 2006.

[32] R. Mukundan, S.-H. Ong, P.-A. Lee, Image analysis by Tchebichef moments, IEEE Transactions on Image Processing 10 (2001) 1357–1364.
[33] R. Mukundan, Some computational aspects of discrete orthonormal moments, IEEE Transactions on Image Processing 13 (2004) 1055–1059.
[34] P.-T. Yap, P. Raveendran, Image focus measure based on Chebyshev moments, in: IEE Proceedings on Vision, Image and Signal Processing, vol. 151, IET, 2004, pp. 128–136.
[35] H.R. Sheikh, Z. Wang, L. Cormack, A.C. Bovik, Live image quality assessment database release 2, 2010 〈http://live.ece.utexas.edu/research/quality〉.
[36] H.R. Sheikh, M.F. Sabir, A.C. Bovik, A statistical evaluation of recent full reference image quality assessment algorithms, IEEE Transactions on Image Processing 15 (2006) 3440–3451.
[37] VQEG, Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment, Technical Report, 2000 〈http://www.vqeg.org/〉.
[38] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, The SSIM index for image quality assessment, 2011 〈http://www.ece.uwaterloo.ca/~z70wang/research/ssim/〉.
[39] C.M. Jarque, A.K. Bera, A test for normality of observations and regression residuals, International Statistical Review/Revue Internationale de Statistique 55 (1987) 163–172.
[40] D.M. Chandler, S.S. Hemami, A57 image database, 2007 〈http://foulard.ece.cornell.edu/dmc27/vsnr/vsnr.html〉.
[41] P. Le Callet, F. Autrusseau, Subjective quality assessment irccyn/ivc database, 2005 〈http://www.irccyn.ec-nantes.fr/ivcdb/〉.
[42] Z.M.P. Sazzad, Y. Kawayoke, Y. Horita, Mict image quality evaluation database, 2010 〈http://mict.eng.u-toyama.ac.jp/mictdb.html〉.
[43] N.N. Ponomarenko, Tampere image database 2008, version 1.0, 2008 〈http://www.ponomarenko.info/tid2008.htm〉.
[44] K.-H. Thung, P. Raveendran, A survey of image quality measures, in: 2009 International Conference for Technical Postgraduates (TECHPOS), pp. 1–4.

Kim-Han Thung received his B.Eng. and M.Eng.Sc. degrees in electrical engineering from University of Malaya, Malaysia, in 2002 and 2005 respectively. He then worked as software engineer at Motorola Penang and application engineer at a local company. He is currently a doctoral research student in the Department of Electrical Engineering, University of Malaya under supervision of Prof. Dr. P. Raveendran. His research interests lie in the field of image and signal analysis, pattern recognition, image segmentation and statistical analysis.

Raveendran Paramesran received his B.Sc. And M.Sc. Degrees in electrical engineering from South Dakota State University, Brookings, and the Dr.Eng. degree from University of Tokushima, Japan, in 1984, 1985, and 1994 respectively. He is currently a Professor at the Department of Electrical Engineering at University of Malaya. His research interests include image and signal analysis, video coding and Brain Computer-Interface (BCI) technology. He is a Senior Member of IEEE. He is also a reviewer for IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), Systems, Man, and Cybernetics (SMC) Part B: Cybernetics, Image Processing and also several other international journals.

Chern-Loon Lim received his B.Eng. and M.Eng.Sc. degrees in electrical engineering from the University of Malaya, Malaysia, in 2005 and 2007 respectively. He is currently pursuing his Ph.D degree in Department of Electrical Engineering at University of Malaya. His research interests lie in the field of image analysis, video coding and pattern recognition. His present work includes fast computation of orthogonal moments.