Abstract—Image quality assessment (IQA) aims to use computational models to measure the image quality consistently with subjective evaluations. The well-known structural similarity index brings IQA from pixel- to structure-based stage. In this paper, a novel feature similarity (FSIM) index for full reference IQA is proposed based on the fact that human visual system (HVS) understands an image mainly according to its low-level features. Specifically, the phase congruency (PC), which is a dimensionless measure of the significance of a local structure, is used as the primary feature in FSIM. Considering that PC is contrast invariant while the contrast information does affect HVS’ perception of image quality, the image gradient magnitude (GM) is employed as the secondary feature in FSIM. PC and GM play complementary roles in characterizing the image local quality. After obtaining the local quality map, we use PC again as a weighting function to derive a single quality score. Extensive experiments performed on six benchmark IQA databases demonstrate that FSIM can achieve much higher consistency with the subjective evaluations than state-of-the-art IQA metrics.

Index Terms—Gradient, image quality assessment (IQA), low-level feature, phase congruency (PC).

I. INTRODUCTION

With the rapid proliferation of digital imaging and communication technologies, image quality assessment (IQA) has been becoming an important issue in numerous applications, such as image acquisition, transmission, compression, restoration, and enhancement. Since the subjective IQA methods cannot be readily and routinely used for many scenarios, e.g., real-time and automated systems, it is necessary to develop objective IQA metrics to automatically and robustly measure the image quality. Meanwhile, it is anticipated that the evaluation results should be statistically consistent with those of the human observers. To this end, the scientific community has developed various IQA methods in the past decades. According to the availability of a reference image, objective IQA metrics can be classified as full reference (FR), no reference (NR), and reduced-reference methods [1]. In this paper, the discussion is confined to FR methods, where the original “distortion-free” image is known as the reference image.

Manuscript received April 13, 2010; revised October 25, 2010; accepted January 19, 2011. Date of publication January 31, 2011; date of current version July 15, 2011. This work was supported by the Hong Kong Research Grants Council (RGC) General Research Fund under Grant PolyU 5330/07E, the Ho Tung Fund under Grant 5-ZH25, and the National Natural Science Foundation of China under Grant 90920003. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Alex ChiChung Kot.

The conventional metrics, such as the peak signal-to-noise ratio (PSNR) and the mean-squared error (MSE) operate directly on the intensity of the image, and they do not correlate well with the subjective fidelity ratings. Thus, many efforts have been made on designing human visual system (HVS) based IQA metrics. Such kinds of models emphasize the importance of HVS’ sensitivity to different visual signals, such as the luminance, the contrast, the frequency content, and the interaction between different signal components [2]–[4]. The noise quality measure (NQM) [2] and the visual SNR (VSNR) [3] are two representatives. Methods, such as the structural similarity (SSIM) index [1], are motivated by the need to capture the loss of structure in the image. SSIM is based on the hypothesis that HVS is highly adapted to extract the structural information from the visual scene; therefore, a measurement of SSIM should provide a good approximation of perceived image quality. The multiscale extension of SSIM, called MS-SSIM [5], produces better results than its single-scale counterpart. In [6], the authors presented a three-component weighted SSIM (3-SSIM) by assigning different weights to the SSIM scores according to the local region type: edge, texture, or smooth area. In [7], Sheikh et al. introduced the information theory into image fidelity measurement, and proposed the information fidelity criterion (IFC) for IQA by quantifying the information shared between the distorted and the reference images. IFC was later extended to the visual information fidelity (VIF) metric in [4]. In [8], Sampat et al. made use of the steerable complex wavelet (CW) transform to measure the SSIM of the two images and proposed the CW-SSIM index.

Recent studies conducted in [9] and [10] have demonstrated that SSIM, MS-SSIM, and VIF could offer statistically much better performance in predicting images’ fidelity than the other IQA metrics. However, SSIM and MS-SSIM share a common deficiency that when pooling a single quality score from the local quality map (or the local distortion measurement map), all positions are considered to have the same importance. In VIF, images are decomposed in different subbands and these subbands can have different weights at the pooling stage [11]; however, within each subband, every position is still given the same importance. Such pooling strategies are not consistent with the intuition that different locations on an image can have very different contributions to HVS’ perception of the image. This is corroborated by a recent study [12], [13], where the authors found that by incorporating appropriate spatially varying weights, the performance of some IQA metrics, e.g., SSIM, VIF, and PSNR, could be improved. But unfortunately, they did not present an automated method to generate such weights.

The great success of SSIM and its extensions owes to the fact that HVS is adapted to the structural information in images. The visual information in an image, however, is often very redundant, while the HVS understands an image mainly based on its low-level features, such as edges and zero crossings [14]–[16]. In other words, the salient low-level features convey crucial information for the HVS to interpret the scene. Accordingly, perceptible image degradations will lead to perceptible changes in image low-level features, and hence, a good IQA metric could be devised by comparing the low-level feature sets between the reference image and the distorted image. Based on the aforementioned analysis, in this paper, we propose a novel low-level feature similarity (FSIM) induced FR IQA metric, namely, FSIM.

One key issue is then what kinds of features could be used in designing FSIM? Based on the physiological and psychophysical evidences, it is found that visually discernable features coincide with those
points, where the Fourier waves at different frequencies have congruent phases [16]–[19], i.e., at points of high phase congruency (PC), we can extract highly informative features. Such a conclusion has been further corroborated by some recent studies in neurobiology using functional magnetic resonance imaging (fMRI) [20]. Therefore, PC is used as the primary feature in computing FSIM. Meanwhile, considering that PC is contrast invariant, but image local contrast does affect HVS’ perception on the image quality, the image gradient magnitude (GM) is computed as the secondary feature to encode contrast information. PC and GM are complementary and they reflect different aspects of the HVS in assessing the local quality of the input image. After computing the local similarity map, PC is utilized again as a weighting function to derive a single similarity score. Although FSIM is designed for gray-scale images (or the luminance components of color images), the chrominance information can be easily incorporated by means of a simple extension of FSIM, and we call this extension FSIMC.

Actually, PC has already been used for IQA in the literature. In [21], Liu and Laganière proposed a PC-based IQA metric. In their method, PC maps are partitioned into subblocks of size 5 × 5. Then, the cross correlation is used to measure the similarity between two corresponding PC subblocks. The overall similarity score is obtained by averaging the cross-correlation values from all block pairs. In [22], PC was extended to phase coherence, which can be used to characterize the image blur. Based on [22], Hassen et al. proposed an NR IQA metric to assess the sharpness of an input image [23].

The proposed FSIM and FSIMC are evaluated on six benchmark IQA databases in comparison with eight state-of-the-art IQA methods. The extensive experimental results show that FSIM and FSIMC can achieve very high consistency with human subjective evaluations, outperforming all the other competitors. Particularly, FSIM and FSIMC work consistently well across all the databases, while other methods may work well only on some specific databases. To facilitate repeatable experimental verifications and comparisons, the MATLAB source code of the proposed FSIM/FSIMC indices and our evaluation results are available online at http://www.comp.polyu.edu.hk/~cszlzhang/IQA/FSIM/FSIM.htm.

The remainder of this paper is organized as follows. Section II discusses the extraction of PC and GM. Section III presents in detail the computation of the FSIM and FSIMC indices. Section IV reports the experimental results. Finally, Section V concludes the paper.

II. EXTRACTION OF PC AND GM

A. Phase Congruency

Rather than defining features directly at points with sharp changes in intensity, the PC model postulates that features are perceived at points, where the Fourier components are maximal in phase. Based on the physiological and psychophysical evidences, the PC theory provides a simple but biologically plausible model of how mammalian visual systems detect and identify features in an image [16]–[20]. PC can be considered as a dimensionless measure for the significance of a local structure.

Under the definition of PC in [17], there can be different implementations to compute the PC map of a given image. In this paper, we adopt the method developed by Kovessi [19], which is widely used in literature. We start from the 1-D signal $g(x)$. Denote by $M_n^e$ and $M_n^o$ the even- and odd-symmetric filters on scale $n$, and they form a quadrature pair. Responses of each quadrature pair to the signal will form a response vector at position $x$ on scale $n$ : $[e_n(x), o_n(x)] = [g(x) * M_n^e, g(x) * M_n^o]$, and the local amplitude on scale $n$ is $A_n(x) = \sqrt{e_n(x)^2 + o_n(x)^2}$. Let $F(x) = \sum_n e_n(x)$ and $H(x) = \sum_n o_n(x)$. The 1-D PC can be computed as follows:

$$PC(x) = \frac{E(x)}{\varepsilon + \sum_n A_n(x)}$$

where $E(x) = \sqrt{F^2(x) + H^2(x)}$, and $\varepsilon$ is a small positive constant.

With respect to the quadrature pair of filters, i.e., $M_n^e$ and $M_n^o$, Gabor filters [24] and log-Gabor filters [25] are two widely used candidates. We adopt the log-Gabor filters because 1) one cannot construct Gabor filters of arbitrarily bandwidth and still maintain a reasonably small DC component in the even-symmetric filter, while log-Gabor filters, by definition, have no DC component, and 2) the transfer function of the log-Gabor filter has an extended tail at the high-frequency end, which makes it more capable to encode natural images than ordinary Gabor filters [19], [25].

The transfer function of a log-Gabor filter in the frequency domain is $G(\omega) = \exp(- (\log(\omega / \omega_0))^2 / 2\sigma^2)$, where $\omega_0$ is the filter’s center frequency, and $\sigma$ controls the filter’s bandwidth.

To compute the PC of 2-D grayscale images, we can apply the 1-D analysis over several orientations and then combine the results using some rule. The 1-D log-Gabor filters described earlier can be extended to 2-D ones by simply applying some spreading function across the filter perpendicular to its orientation. One widely used spreading function is Gaussian [19], [26]–[28]. According to [19], there are some good reasons to choose Gaussian. Particularly, the phase of any function would stay unaffected after being smoothed with Gaussian. Thus, the PC would be preserved. By using Gaussian as the spreading function, the 2-D log-Gabor function has the following transfer function:

$$G_2(\omega, \theta_j) = \exp\left(-\frac{(\log(\omega / \omega_0))^2}{2\sigma^2}\right) \cdot \exp\left(-\frac{(\theta - \theta_j)^2}{2\sigma^2}\right)$$

where $\theta_j = j\pi / J, j = \{0, 1, \ldots, J - 1\}$ is the orientation angle of the filter, $J$ is the number of orientations, and $\sigma_0$ determines the filter’s angular bandwidth. An example of the 2-D log-Gabor filter in the frequency domain, with $\omega_0 = 1/6, \sigma_0 = 0.3$, and $\sigma_0 = 0.4$, is shown in Fig. 1.

By modulating $\omega_0$ and $\theta_j$ and convolving $G_2$ with the 2-D image, we get a set of responses at each point $x$ as $[e_{n,\theta_j}(x), o_{n,\theta_j}(x)]$. The local amplitude on scale $n$ and orientation $\theta_j$ is $A_{n,\theta_j}(x) = \sqrt{e_{n,\theta_j}(x)^2 + o_{n,\theta_j}(x)^2}$, and the local energy along orientation $\theta_j$ is $E_{\theta_j}(x) = \sum_n e_{n,\theta_j}(x)$ and $H_{\theta_j}(x) = \sum_n o_{n,\theta_j}(x)$. The 2-D PC at $x$ is defined as follows:

$$PC_{2D}(x) = \frac{\sum_j E_{\theta_j}(x)}{\varepsilon + \sum_j \sum_n A_{n,\theta_j}(x)}.$$
Fig. 2. Illustration for the FSIM/FSIMc index computation. \( f_1 \) is the reference image, and \( f_2 \) is a distorted version of \( f_1 \).

### Table I

<table>
<thead>
<tr>
<th>Gradient Operator</th>
<th>( G_x(x) )</th>
<th>( G_y(x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sobel</td>
<td>[ \begin{array}{ccc} 1 &amp; 0 &amp; -1 \ 2 &amp; 0 &amp; -2 \ 1 &amp; 0 &amp; -1 \end{array} ] ( \ast f(x) )</td>
<td>[ \begin{array}{ccc} 1 &amp; 2 &amp; 1 \ 0 &amp; 0 &amp; 0 \ -1 &amp; -2 &amp; -1 \end{array} ] ( \ast f(x) )</td>
</tr>
<tr>
<td>Prewitt</td>
<td>[ \begin{array}{ccc} 1 &amp; 0 &amp; -1 \ 1 &amp; 0 &amp; -1 \end{array} ] ( \ast f(x) )</td>
<td>[ \begin{array}{ccc} 1 &amp; 1 &amp; 1 \ 0 &amp; 0 &amp; 0 \end{array} ] ( \ast f(x) )</td>
</tr>
<tr>
<td>Scharr</td>
<td>[ \begin{array}{ccc} 3 &amp; 0 &amp; -3 \ 10 &amp; 0 &amp; -10 \ 3 &amp; 0 &amp; -3 \end{array} ] ( \ast f(x) )</td>
<td>[ \begin{array}{ccc} 3 &amp; 10 &amp; 3 \ 1 &amp; 0 &amp; 0 \end{array} ] ( \ast f(x) )</td>
</tr>
</tbody>
</table>

It should be noted that \( PC_{2D}(x) \) is a real number within \( 0–1 \). Examples of the PC maps of 2-D images can be found in Fig. 2.

### B. Gradient Magnitude

Image gradient computation is a traditional topic in image processing. Gradient operators can be expressed by convolution masks. Three commonly used gradient operators are the Sobel operator [29], the Prewitt operator [29], and the Scharr operator [30]. Their performances will be examined in Section IV. The partial derivatives \( G_x(x) \) and \( G_y(x) \) of the image function \( f(x) \) along horizontal and vertical directions using the three gradient operators are listed in Table I. The GM of \( f(x) \) is then defined as \( \sqrt{G_x^2 + G_y^2} \).

### III. FSIM INDEX

With the extracted PC and GM feature maps, in this section, we present a novel FSIM index for IQA. Suppose that we are going to calculate the similarity between images \( f_1 \) and \( f_2 \). Denote by \( PC_1 \) and \( PC_2 \) the PC maps extracted from \( f_1 \) and \( f_2 \), respectively, and \( G_1 \) and \( G_2 \) the GM maps extracted from them. It should be noted that for color images, PC and GM features are extracted from their luminance channels. FSIM will be defined and computed based on \( PC_1 \), \( PC_2 \), \( G_1 \), and \( G_2 \). Furthermore, by incorporating the image chrominance information into FSIM, an IQA index for color images denoted by FSIMc will be obtained.

#### A. FSIM Index

The computation of FSIM index consists of two stages. In the first stage, the local similarity map is computed, and then in the second stage, we pool the similarity map into a single similarity score.

We separate the FSIM measurement between \( f_1(x) \) and \( f_2(x) \) into two components, each for PC or GM. First, the similarity measure for \( G_1(x) \) and \( G_2(x) \) is defined as follows:

\[
S_{PC}(x) = \frac{2PC_1(x) \cdot PC_2(x) + T_1}{PC^2_1(x) + PC^2_2(x) + T_1}
\]

where \( T_1 \) is a positive constant to increase the stability of \( S_{PC} \) (such a consideration was also included in SSIM [1]). In practice, the determination of \( T_1 \) depends on the dynamic range of PC values. Equation (4) is a commonly used measure to define the similarity of two positive real numbers [1] and its result ranges within \((0, 1]\). Similarly, the GM values \( G_1(x) \) and \( G_2(x) \) are compared, and the similarity measure is defined as follows:

\[
S_{G}(x) = \frac{2G_1(x) \cdot G_2(x) + T_2}{G^2_1(x) + G^2_2(x) + T_2}
\]
Thus, databases so that the proposed FSIM can be conveniently used. Then, significant PC value, it implies that this position intuitively, for a given location than the locations within a smooth area. Since human visual cortex is applications have different contributions to HVS' perception of the image.

\[ \text{GM values. In our experiments, both} \]

\[ \text{and} \]

\[ \text{are parameters used to adjust the relative importance of} \]

\[ \text{and} \]

\[ \text{are combined to get the similarity} \]

\[ \text{and} \]

\[ \text{can be calculated. However, different lo-} \]

\[ \text{are shown in} \]

\[ \text{TABLE IV}\]

\[ \text{TABLE V}\]

\[ \text{where} \]

\[ \text{where} \]

\[ \text{Let} I_1 (I_2) \text{ and} Q_1 (Q_2) \text{ be the I and} Q \text{ chromatic channels of the image} f_1 (f_2), \text{respectively. Similar to the definitions of} S_{pc,c}(x) \text{ and} S_{c,c}(x), \text{we define the similarity between chromatic features as follows:} \]

\[ S_{c}(x) = S_{pc,c}(x) \cdot S_{c,c}(x) \]

Finally, the FSIM index can be extended to FSIM\(_c\) by incorporating the chromatic information in a straightforward manner

\[ \text{where} \]

\[ \text{where} \]

\[ \text{and} \]

\[ \text{IV. EXPERIMENTAL RESULTS AND DISCUSSIONS}\]

\[ \text{A. Databases and Methods for Comparison}\]

To the best of our knowledge, there are six publicly available image databases in the IQA community, including TID2008 [10], CSIQ [32], LIVE [33], IVC [34], MICT [35], and A57 [36]. All of them will be used here for algorithm validation and comparison. The characteristics of these six databases are summarized in Table II.

The performance of the proposed FSIM and FSIM\(_c\) indices will be evaluated and compared with eight representative IQA metrics, including seven state-of-the-arts (SSIM [1], MS-SSIM [5], VIF [4], VSNR [3], IFC [7], NQM [2], and Liu et al.’s method [21]) and the
classical PSNR. For Liu et al.’s method [21], we implemented it by ourselves. For SSIM [1], we used the implementation provided by the author, which is available at [37]. For all the other methods evaluated, we used the public software MeTriX MuX [38]. The MATLAB source code of the proposed FSIM/FSIMc indices is available online at http://www.comp.polyu.edu.hk/~cslzhang/IQA/FSIM/FSIM.htm.

Four commonly used performance metrics are employed to evaluate the competing IQA metrics. The first two are the Spearman rank-order correlation coefficient (SROCC) and the Kendall rank-order correlation coefficient (KROCC), which can measure the prediction monotonicity of an IQA metric. These two metrics operate only on the rank of the data points and ignore the relative distance between data points. To compute the other two metrics, we need to apply a regression analysis, as suggested by the video quality experts group [39], to provide a nonlinear mapping between the objective scores and the subjective mean opinion scores (MOSs). The third metric is the Pearson linear correlation coefficient (PLCC) between MOS and the objective scores after nonlinear regression. The fourth metric is the root MSE (RMSE) between MOS and the objective scores after nonlinear regression. For the nonlinear regression, we used the following mapping function [9]:

$$f(x) = \beta_1 \left( \frac{1}{2} - \frac{1}{1 + e^{\gamma_2(x - \gamma_3)}} \right) + \beta_4 x + \beta_5$$  \hspace{1cm} (12)

where $\beta_i$, $i = 1, 2, \ldots, 5$, are the parameters to be fitted. A better objective IQA measure is expected to have higher SROCC, KROCC, and PLCC while lower RMSE values.

B. Determination of Parameters

There are several parameters need to be determined for FSIM and FSIMc. To this end, we tuned the parameters based on a subdataset of TID2008 database, which contains the first 8 reference images in TID2008 and the associated 544 distorted images. The eight reference images used in the tuning process are shown in Fig. 3. The tuning criterion was that the parameter value leading to a higher SROCC would be chosen. As a result, the parameters required in the proposed methods were set as: $n = 4$, $J = 4$, $\sigma_r = 0.5978$, $\sigma_o = 0.6345$, $T_1 = 0.85$, $T_2 = 160$, $T_3 = T_4 = 200$, and $\lambda = 0.03$. Besides, the center frequencies of the log-Gabor filters at four scales were set as: $1/6$, $1/12$, $1/24$, and $1/48$. These parameters were then fixed for all the following experiments conducted. In fact, we have also used the last 8 reference images (and the associated 544 distorted ones) to tune parameters and obtained very similar parameters to the ones reported here. This may imply that any 8 reference images in the TID2008 database work equally well in tuning parameters for FSIM/FSIMc. However, this conclusion is hard to prove theoretically or even experimentally because there are
Fig. 5. (a)–(f) are PC maps extracted from images Fig. 4(a)–(f), respectively. (a) is the PC map of the reference image while (b)–(f) are the PC maps of the distorted images. (b) and (d) are more similar to (a) than (c), (e), and (f). In (c), (e), and (f), regions with obvious differences to the corresponding regions in (a) are marked by colorful rectangles.

### TABLE VI

PERFORMANCE COMPARISON OF IQA METRICS ON SIX BENCHMARK DATABASES

<table>
<thead>
<tr>
<th>Database</th>
<th>SROCC</th>
<th>FSIMc</th>
<th>MS-SSIM</th>
<th>VFI</th>
<th>SSIM</th>
<th>IFQ</th>
<th>VSNR</th>
<th>NQM</th>
<th>[21]</th>
<th>PSNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>TID2008</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SROCC</td>
<td>0.8805</td>
<td>0.8840</td>
<td>0.8528</td>
<td>0.7496</td>
<td>0.7749</td>
<td>0.5692</td>
<td>0.7046</td>
<td>0.6243</td>
<td>0.7388</td>
<td>0.5245</td>
</tr>
<tr>
<td>KROCC</td>
<td>0.6946</td>
<td>0.6991</td>
<td>0.6543</td>
<td>0.5863</td>
<td>0.5768</td>
<td>0.4261</td>
<td>0.5340</td>
<td>0.4608</td>
<td>0.5414</td>
<td>0.5096</td>
</tr>
<tr>
<td>PLCC</td>
<td>0.8738</td>
<td>0.8762</td>
<td>0.8425</td>
<td>0.8090</td>
<td>0.7732</td>
<td>0.7359</td>
<td>0.6820</td>
<td>0.6135</td>
<td>0.7679</td>
<td>0.5309</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.6525</td>
<td>0.6648</td>
<td>0.7299</td>
<td>0.7888</td>
<td>0.8511</td>
<td>0.9086</td>
<td>0.9815</td>
<td>1.0598</td>
<td>0.8959</td>
<td>1.1372</td>
</tr>
<tr>
<td>CSQ</td>
<td>0.9242</td>
<td>0.9310</td>
<td>0.9138</td>
<td>0.9193</td>
<td>0.8756</td>
<td>0.7482</td>
<td>0.8106</td>
<td>0.7042</td>
<td>0.7642</td>
<td>0.8057</td>
</tr>
<tr>
<td>PLCC</td>
<td>0.9120</td>
<td>0.9192</td>
<td>0.8908</td>
<td>0.9277</td>
<td>0.8613</td>
<td>0.8381</td>
<td>0.8002</td>
<td>0.7423</td>
<td>0.8223</td>
<td>0.8001</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.1077</td>
<td>0.1034</td>
<td>0.1145</td>
<td>0.0980</td>
<td>0.1334</td>
<td>0.1432</td>
<td>0.1575</td>
<td>0.1756</td>
<td>0.1494</td>
<td>0.1575</td>
</tr>
<tr>
<td>LIVE</td>
<td>0.9634</td>
<td>0.9645</td>
<td>0.9445</td>
<td>0.9631</td>
<td>0.9479</td>
<td>0.9224</td>
<td>0.9274</td>
<td>0.9086</td>
<td>0.8650</td>
<td>0.8755</td>
</tr>
<tr>
<td>KROCC</td>
<td>0.8337</td>
<td>0.8363</td>
<td>0.7922</td>
<td>0.8270</td>
<td>0.7963</td>
<td>0.7540</td>
<td>0.7616</td>
<td>0.7413</td>
<td>0.6781</td>
<td>0.6864</td>
</tr>
<tr>
<td>PLCC</td>
<td>0.7957</td>
<td>0.7963</td>
<td>0.7946</td>
<td>0.7950</td>
<td>0.9449</td>
<td>0.8949</td>
<td>0.9231</td>
<td>0.9122</td>
<td>0.8765</td>
<td>0.8721</td>
</tr>
<tr>
<td>IVC</td>
<td>0.9262</td>
<td>0.9293</td>
<td>0.8847</td>
<td>0.8966</td>
<td>0.7224</td>
<td>0.7192</td>
<td>0.6036</td>
<td>0.6342</td>
<td>0.6441</td>
<td>0.5220</td>
</tr>
<tr>
<td>KROCC</td>
<td>0.7564</td>
<td>0.7636</td>
<td>0.7012</td>
<td>0.7165</td>
<td>0.9119</td>
<td>0.9080</td>
<td>0.8032</td>
<td>0.8498</td>
<td>0.8454</td>
<td>0.7199</td>
</tr>
<tr>
<td>PLCC</td>
<td>0.9376</td>
<td>0.9392</td>
<td>0.8934</td>
<td>0.9028</td>
<td>0.4999</td>
<td>0.5105</td>
<td>0.7258</td>
<td>0.6241</td>
<td>0.5697</td>
<td>0.8456</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.4236</td>
<td>0.4183</td>
<td>0.5474</td>
<td>0.5239</td>
<td>0.3782</td>
<td>0.3893</td>
<td>0.3814</td>
<td>0.3814</td>
<td>0.3814</td>
<td>0.3814</td>
</tr>
<tr>
<td>MICt</td>
<td>0.7302</td>
<td>0.7303</td>
<td>0.7029</td>
<td>0.7329</td>
<td>0.6939</td>
<td>0.6413</td>
<td>0.6762</td>
<td>0.7129</td>
<td>0.5152</td>
<td>0.4447</td>
</tr>
<tr>
<td>KROCC</td>
<td>0.9998</td>
<td>0.9967</td>
<td>0.8864</td>
<td>0.9086</td>
<td>0.8794</td>
<td>0.8387</td>
<td>0.8614</td>
<td>0.8911</td>
<td>0.6923</td>
<td>0.6330</td>
</tr>
<tr>
<td>PLCC</td>
<td>0.9998</td>
<td>0.9967</td>
<td>0.8864</td>
<td>0.9086</td>
<td>0.8794</td>
<td>0.8387</td>
<td>0.8614</td>
<td>0.8911</td>
<td>0.6923</td>
<td>0.6330</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.5248</td>
<td>0.5257</td>
<td>0.5625</td>
<td>0.5656</td>
<td>0.5738</td>
<td>0.6723</td>
<td>0.6147</td>
<td>0.5569</td>
<td>0.8674</td>
<td>0.5988</td>
</tr>
<tr>
<td>AS7</td>
<td>0.9781</td>
<td>-</td>
<td>0.8394</td>
<td>0.6223</td>
<td>0.8096</td>
<td>0.3185</td>
<td>0.9355</td>
<td>0.7891</td>
<td>0.7155</td>
<td>0.6389</td>
</tr>
<tr>
<td>PLCC</td>
<td>0.9639</td>
<td>-</td>
<td>0.6478</td>
<td>0.4589</td>
<td>0.6058</td>
<td>0.2378</td>
<td>0.8031</td>
<td>0.5932</td>
<td>0.3755</td>
<td>0.4509</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.9252</td>
<td>-</td>
<td>0.8504</td>
<td>0.6158</td>
<td>0.8017</td>
<td>0.4548</td>
<td>0.9472</td>
<td>0.8020</td>
<td>0.7399</td>
<td>0.6587</td>
</tr>
<tr>
<td>PLCC</td>
<td>0.9393</td>
<td>-</td>
<td>0.1293</td>
<td>0.1936</td>
<td>0.1469</td>
<td>0.2189</td>
<td>0.0781</td>
<td>0.1468</td>
<td>0.1653</td>
<td>0.1849</td>
</tr>
</tbody>
</table>

$C_{25} = 1,081,575$ different ways to select 8 out of the 25 reference images in TID2008.

It should be noted that the FSIM/FSIMc indices will be most effective if used on the appropriate scale. The precisely “right” scale depends on both the image resolution and the viewing distance, and hence, is difficult to be obtained. In practice, we used the following empirical steps proposed by Wang [37] to determine the scale for images viewed from a typical distance: 1) let $F = \max(1, \text{round}(N/256))$, where $N$ is the number of pixels in image height or width; and 2) average local $F \times F$ pixels and then downsample the image by a factor of $F$.

### C. Gradient Operator Selection

In our proposed IQA metrics FSIM/FSIMc, the GM needs to be calculated. To this end, three commonly used gradient operators listed in Table I were examined, and the one providing the best result was selected. Such a gradient operator selection process was carried out by assuming that all the parameters discussed in Section IV-B were fixed. The selection criterion was also that the gradient operator leading to a higher SROCC would be selected. The subdataset used in Section IV-B was used here. The SROCC values obtained by the three gradient operators on the tuning dataset are listed in Table III, from which we can see that the Scharr operator could achieve slightly better performance than the other two. Thus, in all of the following experiments, the Scharr operator was used to calculate the gradient in FSIM/FSIMc.

### D. Example to Demonstrate the Effectiveness of FSIM/FSIMc

In this section, we use an example to demonstrate the effectiveness of FSIM/FSIMc in evaluating the perceptible image quality. Fig. 4(a) is the I17 reference image in the TID2008 database, and Fig. 4(b)–(f) shows five distorted images of I17: I17_01_2, I17_03_2, I17_09_1, I17_11_2, and I17_12_2. Distortion types of Fig. 4(b)–(f) is “additive Gaussian noise,” “spatially correlated noise,” “image denoising,” “JPEG 2000 compression,” and “JPEG transformation errors,” respectively. According to the naming convention of TID2008, the last number (the last digit) of the image’s name represents the distortion degree, and a greater number indicates a severer distortion. We compute the image quality of Fig. 4(b)–(f) using various IQA metrics, and the results are summarized in Table IV. We also list the
subjective scores (extracted from TID2008) of these five images in Table IV. For each IQA metric and the subjective evaluation, higher scores mean higher image quality.

In order to show the correlation of each IQA metric with the subjective evaluation more clearly, in Table V, we rank the images according to their quality scores computed by each metric as well as the subjective evaluation. From Tables IV and V, we can see that the quality scores computed by FSIM/FSIMC correlate with the subjective evaluation much better than the other IQA metrics. From Table V, we can also see that, other than the proposed FSIM/FSIMC metrics, all the other IQA metrics cannot give the same ranking as the subjective evaluations.

The success of FSIM/FSIMC actually owes to the proper use of PC maps. Fig. 5(a)–(f) shows the PC maps of the images in Fig. 4(a)–(f), respectively. We can see that images in Fig. 4(b) and (d) have better perceptible qualities than those in Fig. 4(c), (e), and (f); meanwhile, by visual examination, we can see that maps in Fig. 5(b) and (d) [PC maps of images in Fig. 4(b) and (d)] are more similar to the map in Fig. 5(a) [PC map of the reference image in Fig. 4(a)] than the maps in Fig. 5(c), (e), and (f) [PC maps of images in Fig. 4(c), (e), and (f)]. In order to facilitate visual examination, in Fig. 5(c), (e), and (f), regions with obvious differences to the corresponding regions in Fig. 5(a) are marked by rectangles. For example, in Fig. 5(c), the neck region marked by the yellow rectangle has a perceptible difference to the same region in Fig. 5(a). This example clearly illustrates that images of higher quality will have more similar PC maps to that of the reference image than images of lower quality. Therefore, by properly making use of PC maps in FSIM/FSIMC, we can predict the image quality consistently with human subjective evaluations. More statistically convincing results will be presented in the next two sections.

E. Overall Performance Comparison

In this section, we compare the general performance of the competing IQA metrics. Table VI lists the SROCC, KROCC, PLCC, and RMSE results of FSIM/FSIMC and the other eight IQA algorithms on the TID2008, CSIQ, LIVE, IVC, MICT, and A57 databases. For each performance measure, the three IQA indices producing the best results are highlighted in boldface for each database. It should be noted that except for FSIMC, all the other IQA indices are based on the luminance component of the image. From Table VI, we can see that the proposed FSIM-based IQA metric FSIM or FSIMC performs consistently well across all the databases. In order to demonstrate this consistency more clearly, in Table VII, we list the performance ranking of all the IQA metrics according to their SROCC values. For fairness, the FSIMC index, which also exploits the chrominance information of images, is excluded in Table VII.

From the experimental results summarized in Tables VI and VII, we can see that our methods achieve the best results on almost all the databases, except for MICT and A57. Even on these two databases, however, the proposed FSIM (or FSIMC) is only slightly worse than the best results. Moreover, considering the scales of the databases, including the number of images, the number of distortion types, and the number of observers, we think that the results obtained on
TID2008, CSIQ, LIVE, and IVC are much more convincing than those obtained on MICT and A57. Overall speaking, FSIM and FSI\textsubscript{M}C achieve the most consistent and stable performance across all the six databases. By contrast, for the other methods, they may work well on some databases but fail to provide good results on other databases. For example, although VIF can get very pleasing results on LIVE, it performs poorly on TID2008 and A57. The experimental results also demonstrate that the chromatic information of an image does affect its perceptible quality since FSI\textsubscript{M}C has better performance than FSIM on all color image databases. Fig. 6 shows the scatter distributions of subjective MOS versus the predicted scores by FSIM and the other eight IQA indices on the TID2008 database. The curves shown in Fig. 6 were obtained by a nonlinear fitting according to (12). From Fig. 6, one can see that the objective scores predicted by FSIM correlate much more consistently with the subjective evaluations than the other methods.

V. CONCLUSION

In this paper, we proposed a novel low-level feature-based IQA metric, namely FSI\textsubscript{M}C. The underlying principle of FSI\textsubscript{M}C is that HVS perceives an image mainly based on its salient low-level features. Specifically, two kinds of features, the PC and the GM, are used in FSI\textsubscript{M}C, and they represent complementary aspects of the image visual quality. The PC value is also used to weight the contribution of each point to the overall similarity of two images. We then extended FSI\textsubscript{M}C by incorporating the image chromatic features into consideration. The FSI\textsubscript{M}C and FSI\textsubscript{M}L indices were compared with eight representative and prominent IQA metrics on six benchmark databases, and very promising results were obtained by FSI\textsubscript{M}C and FSI\textsubscript{M}L. When the distortion type is known beforehand, FSI\textsubscript{M}L performs the best while FSI\textsubscript{M}C achieves comparable performance with VIF. When all the distortion types are involved (i.e., all the images in a test database are used), FSI\textsubscript{M}C and FSI\textsubscript{M}L outperform all the other IQA metrics used in comparison. Particularly, they perform consistently well across all the test databases, validating that they are very robust IQA metrics.

REFERENCES

Integrer Computation of Lossy JPEG2000 Compression

Eric J. Balster, Member, IEEE,
Benjamin T. Fortener, Student Member, IEEE, and
William F. Turri, Member, IEEE

Abstract—In this paper, an integer-based Cohen–Daubechies–Feauvea (CDF) 9/7 wavelet transform as well as an integer quantization method used in a lossy JPEG2000 compression engine is presented. The junction of both an integer transform and quantization step allows for a complete integer computation of lossy JPEG2000 compression. The lossy method of compression utilizes the CDF 9/7 wavelet filter, which transforms integer input pixel values into floating-point wavelet coefficients that are then quantized back into integers and finally compressed by the embedded block coding with optimal truncation tier-1 encoder. Integer computation of JPEG2000 allows a reduction in computational complexity of the wavelet transform as well as ease of implementation in embedded systems for higher computational performance. The results of the integer computation show an equivalent rate/distortion curve to the JasPer JPEG2000 compression engine, as well as a 30% reduction in computation time of the wavelet transform and a 56% reduction in computation time of the quantization process on an average.

Index Terms—Cohen–Daubechies–Feauvea (CDF) 9/7 wavelet, integer computation, JPEG2000.

I. INTRODUCTION

JPG2000 is the latest image compression standard from the Joint Pictures Expert Group [9]. It was established as an International Standards Organization (ISO) standard in December of 2000 [8], revised

Manuscript received November 03, 2009; revised September 22, 2010; accepted February 03, 2011. Date of publication February 14, 2011; date of current version July 15, 2011. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. A. R. Reibman.

E. J. Balster is with the Department of Electrical and Computer Engineering, Kettering Laboratory, University of Dayton, Dayton, OH 45469 USA (e-mail: balsterje@notes.udayton.edu).
B. T. Fortener and W. F. Turri are with the University of Dayton Research Institute, Dayton, OH 45469 USA (e-mail: fortener@udr.udayton.edu; william.turri@udr.udayton.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2011.2114353