

A Gabor Feature-Based Quality Assessment Model for the Screen Content Images

Zhangkai Ni, Huanqiang Zeng^{id}, Senior Member, IEEE, Lin Ma^{id}, Member, IEEE, Junhui Hou^{id}, Member, IEEE, Jing Chen, Member, IEEE, and Kai-Kuang Ma^{id}, Fellow, IEEE

Abstract—In this paper, an accurate and efficient full-reference image quality assessment (IQA) model using the extracted Gabor features, called *Gabor feature-based model* (GFM), is proposed for conducting objective evaluation of *screen content images* (SCIs). It is well-known that the Gabor filters are highly consistent with the response of the *human visual system* (HVS), and the HVS is highly sensitive to the edge information. Based on these facts, the imaginary part of the Gabor filter that has odd symmetry and yields edge detection is exploited to the luminance of the reference and distorted SCI for extracting their Gabor features, respectively. The local similarities of the extracted Gabor features and two chrominance components, recorded in the LMN color space, are then measured independently. Finally, the *Gabor-feature pooling* strategy is employed to combine these measurements and generate the final evaluation score. Experimental simulation results obtained from two large SCI databases have shown that the proposed GFM model not only yields a higher consistency with the human perception on the assessment of SCIs but also requires a lower computational complexity, compared with that of classical and state-of-the-art IQA models.¹

Index Terms—Image quality assessment (IQA), screen content images (SCIs), Gabor feature.

I. INTRODUCTION

IN RECENT years, the *screen content images* (SCIs) have received more and more attentions, since they are widely deployed in various multimedia applications and services, such

Manuscript received July 16, 2017; revised November 20, 2017, May 14, 2018, and May 16, 2018; accepted May 16, 2018. Date of publication May 23, 2018; date of current version June 11, 2018. This work was supported in part by the National Natural Science Foundation of China under Grants 61401167, 61372107, and 61602191, in part by the Natural Science Foundation of Fujian Province under Grants 2016J01308 and 2017J05103, in part by the Fujian-100 Talented People Program, in part by the High-level Talent Innovation Program of Quanzhou City under Grant 2017G027, in part by the Promotion Program for Young and Middle-aged Teacher in Science and Technology Research of Huaqiao University under Grants ZQN-YX403 and ZQN-PY418, and in part by the High-Level Talent Project Foundation of Huaqiao University under Grants 14BS201, 14BS204, and 16BS108. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Patrick Le Callet. (*Corresponding author: Huanqiang Zeng.*)

Z. Ni, H. Zeng, and J. Chen are with the School of Information Science and Engineering, Huaqiao University, Xiamen 361021, China (e-mail: eezkni@gmail.com; zeng0043@hqu.edu.cn; jingzi@hqu.edu.cn).

L. Ma is with Tencent AI Lab, Shenzhen 518057, China (e-mail: forest.linma@gmail.com).

J. Hou is with the Department of Computer Science, City University of Hong Kong, Hong Kong (e-mail: jh.hou@cityu.edu.hk).

K.-K. Ma is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798 (e-mail: ekkma@ntu.edu.sg).

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.2018.2839890

¹The source code for the proposed GFM will be available at <http://smartvplab.org/publications/GFM.html>.

as remote screen sharing, cloud gaming, online news and education, electronic brochures, and so on [1]–[3]. Unlike the natural images that mostly contain natural scenes and are acquired by cameras, the SCIs can be a mixture of natural images with computer-generated content, such as graphics, texts, charts, symbols, and so on. Therefore, the inherited characteristics of the SCIs are quite different from that of the natural images [4]. For example, an advertisement poster image could contain one or multiple images inserted together with text description and plots, and so on. Such image tends to have sharp edges, and high-contrast and vivid few colors in certain regions on the image field.

Since the human eyes are the final receiver of the images, *image quality assessment* (IQA) becomes an important issue in the field of image processing task with the goal of objectively evaluating the image quality in accordance with the *human visual system* (HVS). Generally speaking, the IQA models can be roughly classified into three categories according to the availability of the reference image: *full-reference* (FR), *reduced-reference* (RR), and *no-reference* (NR). This paper focuses on FR-IQA models, which are widely applied in various practical scenarios. For example, they can be used to evaluate the performances of various image processing algorithms [5], guide the image/video coding algorithms for improving the perceptual coding efficiency [6], [7], and assist the development of NR-IQA models [8], [9]. In what follows, the classical and some state-of-the-art FR-IQA models will be succinctly described.

A FR-IQA model is to compute the degree of similarity between the *distorted* image and its reference image. The simplest and commonly-used FR-IQA models are the *peak signal-to-noise ratio* (PSNR) and the *mean square error* (MSE). However, it is quite well-recognized that the evaluation scores computed by these two models are often inconsistent with the judgments made by the HVS, since they only consider the differences incurred at the pixel level [10]. A milestone IQA model, called the *structural similarity* (SSIM) [11], considers the fact that the HVS is highly sensitive to the image's structural information and thus incorporates this thesis into the development. To further improve the SSIM's performance, various FR-IQA models are proposed, such as *visual information fidelity* (VIF) [12], *feature similarity* (FSIM) [13], *gradient similarity* (GSIM) [14], *information weighted SSIM* (IW-SSIM) [15], *gradient magnitude similarity deviation* (GMSD) [16], *visual saliency-based index* (VSI) [17], and *perceptual similarity* (PSIM) [18]. They all exploit various properties of the HVS (e.g., edges [16],

micro- and macro-structures [18], salient regions [19], and so on). However, the above-mentioned FR-IQA models are developed for evaluating the natural images and therefore might not be applicable to the SCIs. This is mainly due to the fact that some fundamental characteristics inherited from the contents of these two types of images are quite different [6], [20]. This motivates the need to design an accurate FR-IQA model for the SCIs.

To investigate the perceptual quality assessment of the SCIs, Yang *et al.* [21] propose an SCI IQA model, called *SCI Perceptual Quality Assessment (SPQA)*, by considering the visual difference of the textual and pictorial regions, while Wang *et al.* [6] employ the visual-field adaptation with content weighting. Gu *et al.* [22] come up with a new IQA model for the SCIs, called the *structure-induced quality metric (SIQM)*, and their further work exploits the *saliency-guided gradient magnitude* to conduct similarity measurement [23]. Fang *et al.* [24] evaluate the quality of SCIs by structure features and uncertainty weighting. Gu *et al.* [25] systematically combines the measurements of variations in global and local structures to predict the quality of SCIs. Recently, Ni *et al.* [26] exploit three salient edge attributes extracted from the SCIs in their developed IQA model, called *edge similarity measurement (ESIM)*, which shows the superior performance among the existing comparable models. It is worth to mention that a new SCI database, denoted as SCID, is also developed and presented in [26].

Despite the above-mentioned efforts, it is our interest to further investigate a more accurate IQA metric or model for the SCIs. In this paper, a novel full-reference Gabor feature-based quality assessment model for the SCIs is proposed, called the *Gabor feature-based model (GFM)*. This is motivated by the fact that an image representation yielded by a set of properly-chosen Gabor filters are highly consistent with the response or judgement as made by the HVS when the human eyes view the image. In our approach, therefore the Gabor features are first extracted from the *luminance* (i.e., the L component recorded in the LMN color space) of the reference and distorted SCI, separately. On this feature-extraction process, a specially-designed Gabor filtering (i.e., the imaginary part with odd symmetry) is conducted on the horizontal and the vertical directions, respectively. The obtained filtering results are combined to form the *Gabor feature map*. The *degree of similarity* measurement is then conducted on these maps for the luminance part and for the chrominance components independently, between the reference and distorted SCIs. Finally, the developed *Gabor-feature pooling* strategy is employed to combine these measurements and generate the final IQA evaluation score for the SCI under evaluation. Experimental simulation results obtained from two large SCI databases have shown that the proposed GFM not only yields a higher consistency with the HVS perception on the evaluation of the SCIs but also requires a lower computational complexity, compared with that of the classical and several state-of-the-art IQA models.

The remaining of this paper is organized as follows. In Section II, the proposed SCI quality assessment model based on the Gabor feature is presented. In Section III,

extensive performance evaluation of the proposed IQA model and multiple state-of-the-art IQA models are conducted and compared. Finally, Section IV draws the conclusion.

II. PROPOSED GABOR FEATURE-BASED MODEL FOR THE SCREEN CONTENT IMAGES

A. Motivation and Overview

Generally speaking, a desirable IQA model or metric should yield the computed evaluation score that is highly consistent with the HVS perception. Furthermore, the computational complexity of the IQA model should be as low as possible from the viewpoint of practical applications. The main objective of our proposed IQA model lies in: how to extract and select the salient features from the SCIs with low computational complexity, and hopefully these features are able to effectively characterize the HVS perception. According to the psychophysical and physiological experimental findings, the profiles of the receptive fields of the simple cells resided in the primary visual cortex of the mammal closely resemble the shapes of the frequency response of the Gabor filter [27], [28]. This reveals that the Gabor filter might be able to well characterize the HVS perception and thus can be used to effectively model the visual features of the image. With this motivation, the *Gabor feature-based model (GFM)* for evaluating the image quality of the SCIs is proposed in this paper with the framework as shown in Fig. 1, which consists of the following stages:

- 1) *Color Space Conversion*: As a full-reference IQA model, both the *reference* SCI (denoted by r in the subscript) and the *distorted* SCI (denoted by d in the subscript) are converted from the RGB color space to the LMN color space for decorrelating the luminance and chrominance components.
- 2) *Extraction of the Gabor Features for the Luminance*: The horizontal- and vertical-oriented imaginary-part Gabor filters are respectively used to convolve with the luminance component $L_r(x, y)$ of the reference SCI and the luminance component $L_d(x, y)$ of the distorted SCI. Note that the imaginary-part Gabor filter can effectively detect edge, which is the salient and sensitive attribute of the HVS. The filtered results are combined to generate the *Gabor feature maps*, $G_r(x, y)$ and $G_d(x, y)$, of the reference and the distorted SCIs, respectively.
- 3) *Similarity Measurements*: The Gabor feature similarity measurement will be conducted by comparing the Gabor feature maps $G_r(x, y)$ and $G_d(x, y)$ to arrive at the *Gabor feature similarity map* $S_G(x, y)$. Likewise, the chrominance component's similarity map $S_C(x, y)$ will be computed based on the chrominance component maps $M_r(x, y)$, $M_d(x, y)$, $N_r(x, y)$, and $N_d(x, y)$.
- 4) *Feature Pooling*: The generated similarity maps, $S_G(x, y)$ and $S_C(x, y)$, will be fused by the proposed Gabor-feature-based pooling strategy for yielding the final IQA evaluation score.

The proposed GFM as outlined above will be successively described in detail in the following subsections.

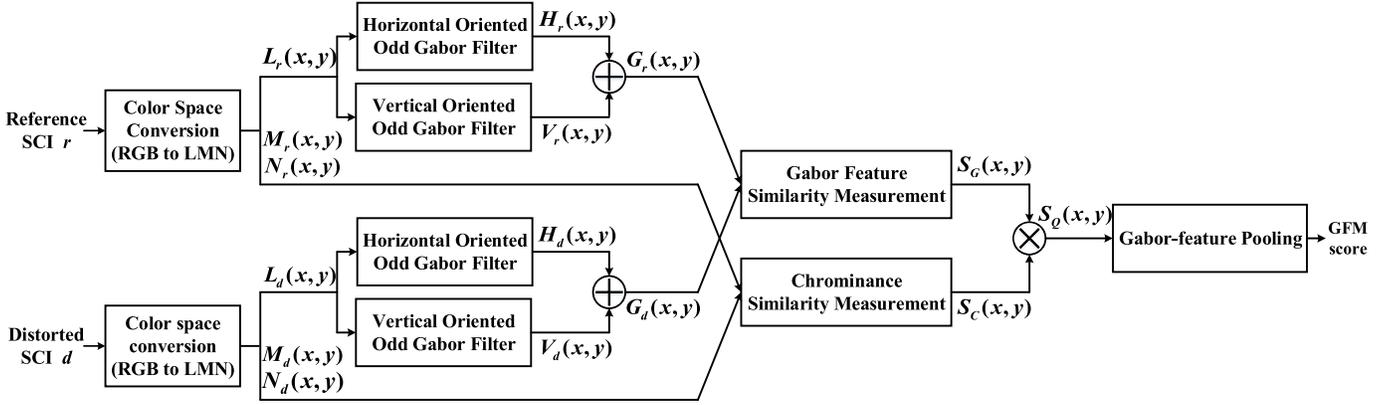


Fig. 1. The framework of the proposed *Gabor feature-based model* (GFM) for objectively evaluating the *screen content images* (SCIs).

B. The Imaginary-Part Gabor Filter

A 2D Gabor filter can be described as a 2D Gaussian kernel function, modulated by an *direction-oriented* complex sinusoidal plane wave, as follows.

$$G(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{1}{2}\left[\left(\frac{x'}{\sigma_x}\right)^2 + \left(\frac{y'}{\sigma_y}\right)^2\right]\right\} \exp(j2\pi fx'),$$

where

$$\begin{aligned} x' &= x\cos\theta + y\sin\theta; \\ y' &= y\cos\theta - x\sin\theta, \end{aligned}$$

where (x, y) denotes the pixel coordinate in the image, f is the frequency of the sinusoidal plane wave (x', y') , and θ is the rotation angle of the sinusoidal plane wave, σ_x and σ_y are the standard deviations of the Gaussian function in the x -direction and y -direction, respectively, which characterize the spatial extent and the bandwidth of the Gabor filter. In this paper, f , σ_x , and σ_y are empirically determined as 0.2, 2.15, and 0.15, respectively.

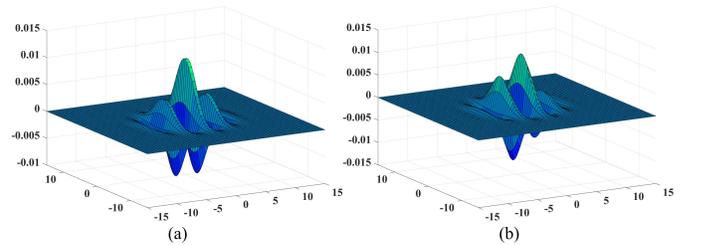
By exploiting the Euler identity,

$$e^{j\theta} = \cos\theta + j\sin\theta, \quad (2)$$

the Gabor filter in (1) can be expressed as a set of two 2D filters—one involves the cosine term (i.e., generating the Gabor's *real* part, equivalently), and the other involves the sine term (i.e., the Gabor's *imaginary* part), respectively. That is,

$$\begin{aligned} G_e(x, y) &= \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{1}{2}\left[\left(\frac{x'}{\sigma_x}\right)^2 + \left(\frac{y'}{\sigma_y}\right)^2\right]\right\} \cos(2\pi fx'), \\ G_o(x, y) &= \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{1}{2}\left[\left(\frac{x'}{\sigma_x}\right)^2 + \left(\frac{y'}{\sigma_y}\right)^2\right]\right\} \sin(2\pi fx'). \end{aligned} \quad (3)$$

Fig. 2 shows a set of 2D Gabor filters, demonstrating two components: the *real*-part Gabor filter, $G_e(x, y)$ and the *imaginary*-part Gabor filter, $G_o(x, y)$. Further note that the



(1) Fig. 2. An illustration of a set of 2D Gabor filters: (a) the *real*-part Gabor filter, $G_e(x, y)$ (with *even* symmetry), and (b) the *imaginary*-part Gabor filter, $G_o(x, y)$ (with *odd* symmetry).

$G_e(x, y)$ is *even* symmetric and often used as an effective blob detector in the field of image processing [29], while the $G_o(x, y)$ is *odd* symmetric and exploited as an efficient and robust solution for conducting edge detection [30].

Since the HVS is highly sensitive to the edge information (e.g., [14], [16], [31]) and a typical SCI contains abundant edge information, this motivates us to exploit the *imaginary-part Gabor filter* (i.e., $G_o(x, y)$) to conduct the perceptual evaluation for the SCIs. Certainly, a set of orientations θ in (1) can be chosen to create a set of $G_o(x, y)$ filters, with each filter corresponding to one specifically chosen orientation θ for capturing the spatial locality of the image content along that direction. Obviously, the more the directions are involved, the higher the computational complexity is required. Considering the tradeoff between the representation effectiveness and the computational efficiency, only two directions are chosen in our work; that is, $\theta = 0$ and $\theta = \pi/2$, as follows: based on the observation that the visual orientation sensitivity tends to be much higher in a cardinal orientation (i.e., a horizontal or a vertical direction) compared with other oblique orientations [32], [33]. Therefore, by individually substituting $\theta = 0$ and $\theta = \pi/2$ into (1), the horizontal- and vertical-oriented odd Gabor filters $G_o(x, y)$ can be obtained and denoted as $G_o^h(x, y)$ and $G_o^v(x, y)$, respectively. That is,

$$\begin{aligned} G_o^h(x, y) &= \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{1}{2}\left[\left(\frac{x}{\sigma_x}\right)^2 + \left(\frac{y}{\sigma_y}\right)^2\right]\right\} \sin(2\pi fx), \end{aligned}$$

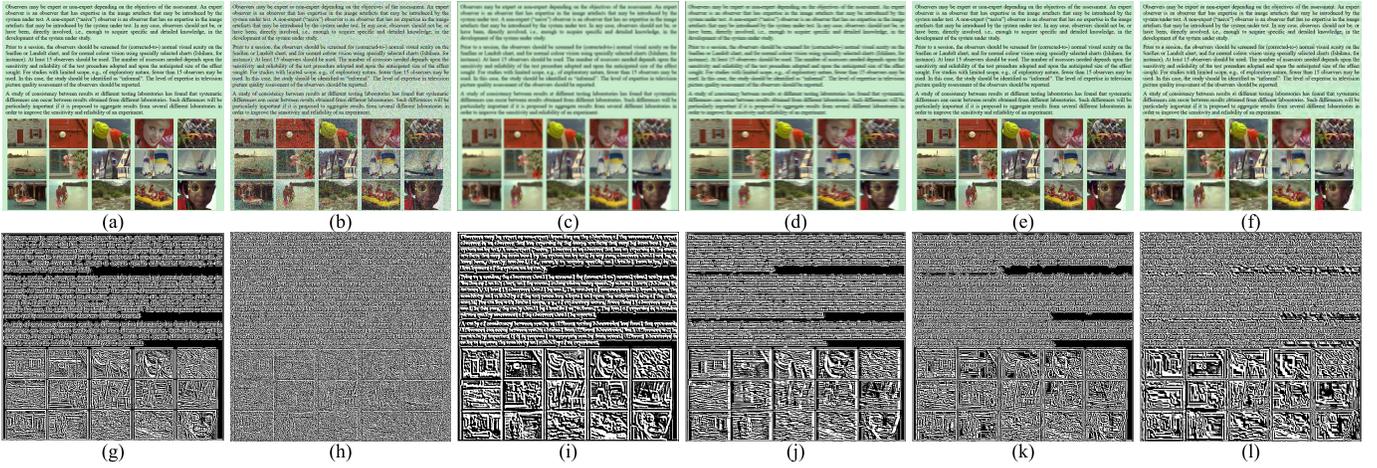


Fig. 3. A illustration of Gabor features of SCIs. [Row 1]: (a) a reference SCI, and the remaining distorted SCIs are caused by: (b) Gaussian noise; (c) Gaussian blur; (d) motion blur; (e) JPEG compression; (f) JPEG2000 compression. [Row 2]: Their corresponding Gabor features, obtained by (7).

$$G_o^v(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{1}{2}\left[\left(\frac{y}{\sigma_x}\right)^2 + \left(\frac{-x}{\sigma_y}\right)^2\right]\right\} \sin(2\pi fy). \quad (4)$$

C. Extraction of the Gabor Features for the Luminance

In our work, an IQA model is developed to assess the image quality of the SCIs by considering both the luminance and the chrominance components, as shown in Fig. 1. First, the proposed approach converts the given reference SCI r and its distorted version d from the RGB color space to the LMN color space individually [34]. The conversion has the effect of decorrelating the luminance component (L) from the chrominance components (M , N) [17]; that is, for all pixels,

$$\begin{bmatrix} L_\zeta(x, y) \\ M_\zeta(x, y) \\ N_\zeta(x, y) \end{bmatrix} = \begin{bmatrix} 0.06 & 0.63 & 0.27 \\ 0.30 & 0.04 & -0.35 \\ 0.34 & -0.60 & 0.17 \end{bmatrix} \begin{bmatrix} R_\zeta(x, y) \\ G_\zeta(x, y) \\ B_\zeta(x, y) \end{bmatrix}, \quad (5)$$

where $L_\zeta(x, y)$, $M_\zeta(x, y)$, and $N_\zeta(x, y)$ denote the luminance and two chrominance components at the pixel position (x, y) of the SCI image ζ (where $\zeta := \{r, d\}$), respectively. Note that the LMN color space is selected in this work, since the weights in the LMN conversion are optimized for the HVS [35].

The Gabor filters as established in (4) are then applied to the *luminance* component of each SCI. That is, the input SCI will be convolved with each filter in (4) to generate the horizontal and vertical responses. That is,

$$\begin{aligned} H_\zeta(x, y) &= G_o^h(x, y) \otimes L_\zeta(x, y) \\ V_\zeta(x, y) &= G_o^v(x, y) \otimes L_\zeta(x, y), \end{aligned} \quad (6)$$

where the symbol “ \otimes ” denotes the convolution operator. A further summation of these obtained responses is defined as the so-called *Gabor features* for the luminance in this paper; that is,

$$G_\zeta(x, y) = H_\zeta(x, y) + V_\zeta(x, y), \quad (7)$$

To demonstrate, the extracted Gabor features of a typical SCI and its distorted versions, involving five different types of distortions, are shown in Fig. 3, respectively.

D. Similarity Measurements

Analogous to the practice exercised in [11], [13], [14], [17], and [36], the Gabor feature similarity, between the luminance components of the reference and distorted SCIs, can be measured as

$$S_G(x, y) = \frac{2G_r(x, y) \cdot G_d(x, y) + C_G}{G_r^2(x, y) + G_d^2(x, y) + C_G}, \quad (8)$$

where $S_G(x, y) \in (0, 1]$, $G_r(x, y)$ and $G_d(x, y)$ are the Gabor feature maps of the r and d , respectively, and C_G is a positive constant that supplies numerical stability when the denominator converges to zero.

Besides the luminance distortion, the chrominance distortion is also frequently encountered in the SCIs. For example, color saturation changes are quite likely resulted from color rendering and screen sharing among different display devices under different settings. Considering that the chrominance components (i.e., $M_\zeta(x, y)$ and $N_\zeta(x, y)$) also affect the perceptual assessment about the visual quality of color images to a large extent [17], [37], the chrominance similarity measurement is measured to describe the quality degradation caused by color distortions, which is defined as

$$S_C(x, y) = \frac{2M_r(x, y) \cdot M_d(x, y) + C_C}{M_r^2(x, y) + M_d^2(x, y) + C_C} \cdot \frac{2N_r(x, y) \cdot N_d(x, y) + C_C}{N_r^2(x, y) + N_d^2(x, y) + C_C}, \quad (9)$$

where $M_r(x, y)$, $N_r(x, y)$ and $M_d(x, y)$, $N_d(x, y)$ are the chrominance component maps of the reference SCI and the distortion SCI, respectively. As the role of C_G played in (8), C_C is a positive constant to ensure the numerical stability.

After performing the similarity measurement on the luminance and two chrominance components for the reference and the distorted SCIs, the proposed GFM combines the luminance’s similarity measurement $S_G(x, y)$ and the chrominance’s similarity measurement $S_C(x, y)$ to obtain the *local quality map* $S_Q(x, y)$:

$$S_Q(x, y) = [S_G(x, y)]^\alpha \cdot [S_C(x, y)]^\beta, \quad (10)$$

where α and β are two positive constants that are used to adjust the relative importance of $S_G(x, y)$ and $S_C(x, y)$. In this paper, these two parameter values are empirically determined as $\alpha = 1$ and $\beta = 0.04$.

E. Gabor-Feature Pooling Strategy

As highlighted in Section II-B, the Gabor filters have shown to fit well the receptive fields of the simple cell in the primary visual cortex [27]. Intuitively, if a pixel has yielded a larger Gabor-feature value, this implies that the HVS will pay more attention on this pixel and result in stronger visual quality perception from this location. Consequently, the *Gabor feature pooling* strategy is designed with such considerations, as follows.

At each pixel location (x, y) , the Gabor feature response from the reference SCI (i.e., $G_r(x, y)$) and the distorted SCI (i.e., $G_d(x, y)$) will be considered, and the larger value of these two quantities will be exploited as the weighting factor to reflect the phenomenon as above-mentioned. Thus, a *weight map* $\omega(x, y)$ can be generated according to:

$$\omega(x, y) = \max\{|G_r(x, y)|, |G_d(x, y)|\}. \quad (11)$$

where $|\cdot|$ denotes the absolute-value function. Therefore, the final IQA evaluation score can be obtained by performing the *weighted* average over all the pixel locations (x, y) on the local quality map $S_Q(x, y)$ as

$$\text{GFM Score} = \frac{\sum_{(x,y)} \omega(x, y) \cdot S_Q(x, y)}{\sum_{(x,y)} \omega(x, y)}. \quad (12)$$

III. EXPERIMENTAL RESULTS

In this section, extensive experimental results are presented and discussed. These results are the performance evaluation of various comparable IQA models, including the proposed *Gabor feature-based model* (GFM). To conduct these experiments, two large SCI databases are used. In what follows, these databases and the evaluation criteria will be described, followed by comparing the performance of the proposed GFM and other state-of-the-art IQA models in terms of their *accuracy* (i.e., the consistency to HVS) and *efficiency* (i.e., the computational complexity).

A. Database and Evaluation Criteria

1) *SCI Database*: Two publicly available SCI databases, which are used to conduct the performance evaluations of the IQA models, are summarized in Table I and succinctly described as follows.

- **SIQAD** [21]: It consists of 20 reference SCIs and 980 distorted SCIs that are generated from the reference SCIs with consideration of using seven types of image distortions and with seven distortion levels created for each distortion type. The distortion types considered in this database are the Gaussian noise (GN), Gaussian blur (GB), motion blur (MB), contrast change (CC), JPEG

TABLE I
TWO PUBLICLY AVAILABLE IQA DATABASES FOR SCI

Database	Reference SCIs No.	Distorted SCIs No.	Distortion Types	Subjects No.
SIQAD	20	980	7	96
SCID	40	1800	9	186

compression (JPEG), JPEG2000 compression (J2K), and the layer segmentation-based coding (LSC).

- **SCID** [26]: It contains 1,840 SCIs, including 40 reference SCIs and 1,800 distorted SCIs created by rendering these reference SCIs using nine different types of image distortions and with five levels of degradations created for each distortion type. Each of these 1,840 SCIs is with a resolution of 1280×720 . The above-mentioned nine distortion types include Gaussian noise (GN), Gaussian blur (GB), motion blur (MB), contrast change (CC), color saturation change (CSC), color quantization with dithering (CQD), JPEG compression (JPEG), JPEG2000 compression (J2K), and HEVC-SCC.

2) *Evaluation Criteria*: By following the standard procedures of conducting performance evaluation as suggested in the *video quality experts group* (VQEG) HDTV test [38], [39], a nonlinear regression formula involving five parameters is employed in this work to map the objective quality predictions to the subjective scores

$$Q_i = \beta_1 \left\{ \frac{1}{2} - \frac{1}{1 + \exp[\beta_2(s_i - \beta_3)]} \right\} + \beta_4 s_i + \beta_5, \quad (13)$$

where s_i is the perceptual quality score of the i -th distorted SCI computed from an IQA model, and Q_i is the corresponding mapped score. The fitting parameters β_1 , β_2 , β_3 , β_4 , and β_5 are to be determined by minimizing the sum of squared errors between the mapped objective score Q_i and the subjective MOS or DMOS values.

After the above-mentioned mapping, four performance indices as commonly-used in the IQA field are exploited to evaluate the proposed GFM model and other IQA models—i.e., *Pearson linear correlation coefficient* (PLCC), *Spearman rank-order correlation coefficient* (SROCC), *Kendall rank-order correlation coefficient* (KROCC), and *Root mean-squared error* (RMSE). The PLCC evaluates the prediction accuracy and is defined as

$$\text{PLCC} = \frac{\sum_{i=1}^n (Q_i - \bar{Q})(m_i - \bar{m})}{\sqrt{\sum_{i=1}^n (Q_i - \bar{Q})^2 (m_i - \bar{m})^2}}, \quad (14)$$

where m_i represents the subjective score (i.e., MOS or DMOS) of the i -th distorted SCI, while \bar{m} and \bar{Q} are the mean values of m_i and Q_i , respectively.

TABLE II
PERFORMANCE COMPARISONS OF DIFFERENT IQA MODELS ON THE SIQAD AND THE SCID DATABASES

Criteria	PSNR	SSIM	IWSSIM	VIF	MAD	FSIM	GSIM	GMSD	VSI	SCQI	SIQM	SPQA	SQI	SQMS	ESIM	SVQI	GFM	
		[11]	[15]	[40]	[41]	[13]	[14]	[16]	[17]	[37]	[22]	[21]	[6]	[23]	[26]	[25]		
SIQAD	PLCC	0.5858	0.7561	0.6519	0.8083	0.5467	0.5888	0.5659	0.7291	0.5543	0.6026	0.8520	0.8584	0.8644	0.8870	0.8788	0.8908	0.8828
	SROCC	0.5570	0.7566	0.6546	0.8069	0.5831	0.5824	0.5483	0.7305	0.5381	0.6113	0.8450	0.8416	0.8548	0.8803	0.8632	0.8836	0.8735
	KROCC	0.4209	0.5583	0.4977	0.6082	0.4486	0.4253	0.4054	0.5488	0.3874	0.4546	0.6521	-	-	0.6936	0.6742	0.6985	0.6876
	RMSE	11.6008	9.3680	10.8550	8.4282	11.9857	11.5700	11.8014	9.7972	11.9152	11.4231	7.4936	7.3421	7.1982	6.6110	6.8310	6.5030	6.7234
SCID	PLCC	0.7622	0.7343	0.7877	0.8200	0.7736	0.7718	0.7042	0.8337	0.7694	0.7489	0.8303	-	-	0.8557	0.8630	0.8604	0.8760
	SROCC	0.7512	0.7146	0.7714	0.7969	0.7576	0.7550	0.6945	0.8138	0.7621	0.7814	0.8086	-	-	0.8320	0.8478	0.8386	0.8759
	KROCC	0.5512	0.5180	0.5726	0.6055	0.5622	0.5541	0.4920	0.6183	0.5648	0.5826	0.6168	-	-	0.6429	0.6516	0.6502	0.6844
	RMSE	9.1682	9.6133	8.7243	8.1069	8.9739	9.0047	10.0552	7.8210	9.0456	9.3846	7.8920	-	-	7.3276	7.1552	7.2178	6.8310
Direct Average	PLCC	0.6740	0.7452	0.7198	0.8142	0.6602	0.6803	0.6351	0.7814	0.6619	0.6758	0.8412	-	-	0.8714	0.8709	0.8756	0.8794
	SROCC	0.6541	0.7356	0.7130	0.8019	0.6704	0.6687	0.6214	0.7722	0.6501	0.6964	0.8268	-	-	0.8562	0.8555	0.8611	0.8747
	KROCC	0.4861	0.5382	0.5352	0.6069	0.5054	0.4897	0.4487	0.5836	0.4761	0.5186	0.6345	-	-	0.6683	0.6629	0.6744	0.6860
Weighted Average	PLCC	0.7051	0.7414	0.7437	0.8162	0.7001	0.7126	0.6594	0.7998	0.6998	0.7015	0.8373	-	-	0.8658	0.8681	0.8702	0.8782
	SROCC	0.6883	0.7282	0.7336	0.8001	0.7011	0.6991	0.6472	0.7868	0.6896	0.7263	0.8204	-	-	0.8476	0.8528	0.8532	0.8751
	KROCC	0.5090	0.5310	0.5484	0.6064	0.5254	0.5124	0.4640	0.5958	0.5074	0.5412	0.6282	-	-	0.6593	0.6589	0.6658	0.6854

The SROCC and KROCC evaluate the prediction monotonicity as

$$\text{SROCC} = 1 - \frac{6 \sum_{i=1}^N d_i^2}{N(N^2 - 1)}, \quad (15)$$

where d_i is the difference between the i -th image's ranks in the subjective and objective evaluations, respectively, and N denotes the total number of samples.

$$\text{KROCC} = \frac{2(N_c - N_d)}{N(N - 1)}, \quad (16)$$

where N_c and N_d denote the number of concordant and discordant pairs found in the database, respectively. The RMSE evaluates the prediction consistency as

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (Q_i - m_i)^2}. \quad (17)$$

It should be pointed out that the higher the values of the PLCC, SROCC, and KROCC, the better the performance of the IQA model, since this indicates that the correlation between the objective and the subjective scores is higher. On the contrary, a lower RMSE value indicates a better performance.

B. Performance Comparison and Analysis

To demonstrate its superiority, the proposed GFM model is compared with multiple IQA models, including PSNR, SSIM [11], IWSSIM [15], VIF [40], MAD [41], FSIM (for grey image) [13], GSIM [14], GMSD [16], VSI [17], SCQI [37], SIQM [22], SPQA [21], SQI [6], SQMS [23], ESIM [26], and SVQI [25], where the last six IQA models are specifically designed for the SCIs (separated by a vertical line in the following Tables II and III), while the rest are all for evaluating the natural images. It is worth to note that four parameters of the proposed GFM model—i.e.,

C_G , C_C , α , and β , are determined based on a subset of SIQAD database, which contains 8 reference SCIs and their rendered 392 distorted SCIs. Following the same practices as suggested in [13], [23], and [37], those parameter values, leading to higher SROCC, will be selected. Through extensive experiments, the values of C_G , C_C , α , and β are empirically determined as 330, 100, 1, and 0.04, respectively.

1) *Overall Performance Comparison*: Table II lists the overall performances of various IQA models on the SIQAD and the SCID databases, respectively. In this table, the first-ranked, the second-ranked, and the third-ranked performance figures of each measurement criterion (i.e., PLCC, SROCC, KROCC, or RMSE) are boldfaced in red, blue, and black, respectively. Note that the program codes of all the models under comparison are downloaded from their original sources, except for two SCI ones (i.e., SPQA and SQI). Therefore, the results of SPQA and SQI on the SCID database and some particular results of SPQA and SQI on the SIQAD database (e.g., KROCC of SPQA, KROCC of SQI, RMSE of SQI on each distortion type) are not available. From Table II, one can see that the proposed GFM yields the best overall performance in terms of PLCC, SROCC, KROCC, and RMSE on SCID database, compared with other state-of-the-art FR-IQA models. On the SIQAD database, the proposed GFM obtains third-place overall performance but almost comparable to the top two models—SQMS and SVQI.

To show the comprehensive performance comparisons over multiple databases, two commonly-used average measurements are exploited to evaluate the average performance of different IQA models over the SIQAD and SCID databases in this work, as suggest in [15], [16], and [42]. These two average measurements can be defined as below:

$$\bar{p} = \frac{\sum_{i=1}^M p_i \cdot w_i}{\sum_{i=1}^M w_i}, \quad (18)$$

where M means the total number of databases ($M = 2$ in this work), p_i indicates the value of the performance index

TABLE III

PLCC, SROCC AND RMSE COMPARISON OF VARIOUS IQA MODELS UNDER DIFFERENT DISTORTION TYPES ON THE SIQAD AND THE SCID DATABASES

	Distortions	PSNR	SSIM	IWSSIM	VIF	MAD	FSIM	GSIM	GMSD	VSI	SCQI	SIQM	SPQA	SQI	SQMS	ESIM	SVQI	GFM		
			[11]	[15]	[40]	[41]	[13]	[14]	[16]	[17]	[37]	[22]	[21]	[6]	[23]	[26]	[25]			
PLCC	SIQAD	GN	0.9053	0.8806	0.8804	0.9011	0.8852	0.7428	0.8448	0.8956	0.8762	0.8807	0.8921	0.8921	0.8829	0.8986	0.8891	0.9031	0.8990	
		GB	0.8603	0.9014	0.9079	0.9102	0.9120	0.7206	0.8831	0.9094	0.8502	0.8535	0.9124	0.9058	0.9202	0.9126	0.9234	0.9132	0.9143	
		MB	0.7044	0.8060	0.8414	0.8490	0.8361	0.6874	0.7711	0.8436	0.6620	0.6949	0.8565	0.8315	0.8789	0.8654	0.8886	0.8722	0.8662	
		CC	0.7401	0.7435	0.8404	0.7076	0.3933	0.7507	0.8077	0.7827	0.7723	0.7119	0.7902	0.7992	0.7724	0.8022	0.7641	0.8087	0.8107	
		JPEG	0.7545	0.7487	0.7998	0.7986	0.7662	0.5566	0.6778	0.7746	0.7124	0.6782	0.7717	0.7696	0.8218	0.7850	0.7999	0.7953	0.8398	
		J2K	0.7893	0.7749	0.8040	0.8205	0.8344	0.6675	0.7242	0.8509	0.7479	0.7225	0.7940	0.8252	0.8271	0.8261	0.7888	0.8342	0.8486	
		LSC	0.7805	0.7307	0.8155	0.8385	0.8184	0.5964	0.7218	0.8559	0.7454	0.7418	0.7204	0.7958	0.8310	0.8119	0.7915	0.8283	0.8288	
		GN	0.9530	0.9354	0.9431	0.9699	0.9315	0.9516	0.9170	0.9273	0.9556	0.9319	0.9269	-	-	0.9298	0.9563	0.9362	0.9497	
		GB	0.7772	0.8711	0.9174	0.8999	0.8559	0.8493	0.8449	0.7348	0.8307	0.8244	0.9266	-	-	0.9081	0.8700	0.9130	0.9156	
	MB	0.7615	0.8794	0.9055	0.8421	0.8362	0.8523	0.8383	0.7954	0.8177	0.8147	0.9152	-	-	0.8968	0.8824	0.8997	0.9023		
	CC	0.7435	0.6903	0.8989	0.8092	0.4987	0.8947	0.8675	0.8041	0.8093	0.8353	0.7821	-	-	0.8441	0.7908	0.8266	0.8787		
	JPEG	0.8393	0.8581	0.9308	0.9418	0.9251	0.9419	0.9373	0.9351	0.9148	0.9036	0.9226	-	-	0.9302	0.9421	0.9356	0.9392		
	J2K	0.9176	0.8586	0.9195	0.9489	0.9381	0.9607	0.9441	0.9422	0.9451	0.9312	0.9076	-	-	0.9468	0.9457	0.9513	0.9226		
	CSC	0.0622	0.0890	0.0527	0.0898	0.1296	0.0966	0.0560	0.0952	0.9119	0.8393	0.0683	-	-	0.0628	0.0694	0.0919	0.8728		
	HEVC-SCC	0.7991	0.7914	0.8883	0.8656	0.8953	0.9228	0.8835	0.9043	0.9035	0.8708	0.8316	-	-	0.8515	0.9108	0.8496	0.8740		
	CQD	0.9210	0.7810	0.8930	0.9085	0.9014	0.9202	0.8974	0.9177	0.8873	0.8823	0.8385	-	-	0.8986	0.9005	0.9047	0.8928		
	SROCC	SIQAD	GN	0.8790	0.8694	0.8743	0.8888	0.8721	0.7373	0.8404	0.8856	0.8655	0.8821	0.8711	0.8823	0.8602	0.8860	0.8757	0.8909	0.8795
			GB	0.8573	0.8921	0.9060	0.9059	0.9087	0.7286	0.8796	0.9119	0.8495	0.8463	0.9102	0.9017	0.9244	0.9119	0.9239	0.9129	0.9132
MB			0.7130	0.8041	0.8421	0.8492	0.8357	0.6641	0.7753	0.8441	0.7658	0.7604	0.8401	0.8255	0.8810	0.8695	0.8938	0.8753	0.8699	
CC			0.6828	0.6405	0.7563	0.6433	0.3907	0.7175	0.7148	0.6378	0.6495	0.5780	0.7055	0.6154	0.6677	0.6949	0.6108	0.7131	0.7038	
JPEG			0.7569	0.7576	0.7978	0.7924	0.7674	0.5879	0.6796	0.7712	0.7196	0.7080	0.7754	0.7673	0.8189	0.7893	0.7989	0.7925	0.8434	
J2K			0.7746	0.7603	0.7998	0.8131	0.8382	0.6363	0.7125	0.8436	0.7299	0.7231	0.7771	0.8152	0.8169	0.8194	0.7827	0.8282	0.8444	
LSC			0.7930	0.7371	0.8214	0.8463	0.8154	0.5979	0.7145	0.8592	0.7419	0.7588	0.7255	0.8003	0.8432	0.8293	0.7958	0.8412	0.8445	
GN			0.9424	0.9171	0.9305	0.9616	0.9262	0.9378	0.9112	0.9341	0.9455	0.9556	0.9133	-	-	0.9155	0.9460	0.9191	0.9370	
GB			0.7702	0.8698	0.9165	0.8954	0.8603	0.8476	0.8420	0.7931	0.8221	0.8638	0.9232	-	-	0.9079	0.8699	0.9079	0.9081	
MB		0.7375	0.8588	0.8918	0.8259	0.8296	0.8370	0.8194	0.8148	0.8013	0.8587	0.9006	-	-	0.8814	0.8608	0.8842	0.8892		
CC		0.7265	0.6564	0.8475	0.6115	0.4784	0.8473	0.8204	0.5672	0.8158	0.7465	0.7435	-	-	0.8027	0.6182	0.7705	0.8225		
JPEG		0.8321	0.8490	0.9275	0.9349	0.9242	0.9403	0.9366	0.9344	0.9142	0.9171	0.9158	-	-	0.9236	0.9455	0.9287	0.9281		
J2K		0.9074	0.8439	0.9067	0.9369	0.9330	0.9484	0.9349	0.9279	0.9307	0.9270	0.8935	-	-	0.9320	0.9359	0.9367	0.9085		
CSC		0.0908	0.0963	0.1336	0.1221	0.1440	0.1182	0.1214	0.1165	0.9141	0.8970	0.0617	-	-	0.0814	0.1037	0.0790	0.8736		
HEVC-SCC		0.8074	0.8263	0.8867	0.8580	0.8771	0.9098	0.8730	0.8958	0.8929	0.8721	0.8517	-	-	0.8667	0.9036	0.8665	0.8712		
CQD		0.9080	0.7766	0.8846	0.8918	0.9024	0.9031	0.8707	0.9047	0.8820	0.9099	0.8301	-	-	0.8913	0.8868	0.8957	0.8907		
RMSE		SIQAD	GN	6.3372	7.0679	7.7044	6.4673	6.9391	9.9860	7.9811	6.6354	7.1890	7.0651	7.0165	6.7394	-	6.5461	6.8272	6.4044	6.6835
			GB	7.7376	6.5701	6.3619	6.2859	6.2269	10.52307.1210	6.3111	7.9900	7.9092	5.8367	6.4301	-	6.2113	5.8270	6.1550	6.1459	
	MB		9.2287	7.6967	7.0600	6.8704	7.1322	9.4432	8.2788	6.9816	9.7450	9.3502	6.0869	7.2223	-	6.5254	5.9639	6.3604	6.5184	
	CC		8.4591	8.4116	6.8184	8.8876	11.56528.3190	7.4160	7.8297	7.9900	8.8342	8.1079	7.6184	-	7.5098	8.1141	7.3996	7.3638		
	JPEG		6.1665	6.2295	5.6406	5.6551	6.0380	7.8072	6.9085	5.9427	6.5950	6.9057	5.6548	6.0000	-	5.8210	5.6401	5.6969	5.1009	
	J2K		6.3819	6.5691	6.1804	5.9412	5.7276	7.7404	7.1675	5.4591	6.8990	7.1859	6.0820	5.8706	-	5.8568	6.3877	5.7309	5.4985	
	LSC		5.3336	5.8253	4.9379	4.6497	4.9025	6.8486	5.9046	4.4121	5.6880	5.7226	5.3576	5.1664	-	4.9813	5.2150	4.7751	4.7736	
	GN		3.8093	4.4458	4.1780	3.0629	4.5714	3.8613	5.0127	4.7044	3.7138	4.5600	4.8222	-	-	4.6250	3.6760	4.4179	3.9378	
	GB		6.6633	5.1998	4.2163	4.6179	5.4775	5.5903	5.6648	7.1821	5.8956	5.9943	4.0989	-	-	4.4336	5.2213	4.3194	4.2566	
	MB	7.0843	5.2044	4.6376	5.8960	5.9947	5.7180	5.9607	6.6249	6.2922	6.3394	4.7388	-	-	4.8352	5.1431	4.7709	4.6121		
	CC	5.9867	6.4767	3.9218	5.2594	7.7590	3.9979	4.4524	5.3211	5.2583	4.9217	6.1281	-	-	4.7995	5.4790	5.0374	4.2732		
	JPEG	8.1718	7.7179	5.4930	5.0536	5.7076	5.0471	5.2369	5.3275	6.0971	6.4390	6.7341	-	-	5.5181	5.0373	5.3053	5.2011		
	J2K	6.3222	8.1562	6.2555	5.0207	5.5103	4.4180	5.2462	5.3283	5.2451	5.8002	7.2951	-	-	5.1191	5.1695	4.9058	6.1385		
	CSC	9.8203	9.8003	9.8257	9.7996	9.7564	9.7933	9.8239	9.7947	4.0392	5.3503	9.8394	-	-	9.8199	9.8156	9.7977	4.8031		
	HEVC-SCC	8.4009	8.5037	6.3904	6.9657	6.1988	5.3583	6.5176	5.9393	5.9628	6.8407	8.197	-	-	7.2938	5.7446	7.3381	6.7590		
	CQD	4.9814	7.9855	5.7530	5.3440	5.5354	5.0054	5.6406	5.0796	5.8964	6.0188	7.1976	-	-	5.6110	5.5607	5.4481	5.7592		

(e.g., PLCC, SROCC, KROCC) on i -th database ($i = 1, 2$ correspond to SIQAD and SCID, respectively), and w_i is the corresponding weight i -th database. For the first average measurement—*Direct Average*, $w_1 = w_2 = 1$. For the second average measurement—*Weighted Average*, w_i is dependent on

the size of database and is set as the number of distorted SCIs in i -th database (i.e., 980 for SIQAD and 1800 for SCID). Note that there is a significant difference in the range of MOS/DMOS in different databases, and the average RMSE is thus not calculated to make the comparison fair.

Table II shows the average performance of different IQA models across these two SCI databases. It can be obviously observed that the proposed GFM yields the highest PLCC, SROCC, and KROCC scores in both Direct Average and Weighted Average performance comparison, meaning that the proposed GFM outperforms all the state-of-the-art models under comparison. It is interesting to see that the VIF model achieves the first place on average performance comparison among those IQA models for natural images. And all the IQA models specifically dedicated to evaluate the quality of SCIs (i.e., SIQM, SPQA, SQI, SQMS, ESIM, SVQI, and proposed GFM) can achieve higher correlations than those models developed for the natural images on SIQAD and SCID databases. This is mainly due to the fact that they consider the special characteristics of the SCI on the design of SCI quality assessment models.

For a clearer illustration, the performance gain of the proposed GFM to other models in terms of SROCC is further computed as follows [23]:

$$\eta = \frac{\mathcal{R}_G - \mathcal{R}_O}{\mathcal{R}_O} \times 100\%, \quad (19)$$

where \mathcal{R}_G and \mathcal{R}_O are the SROCC values of the proposed GFM and other IQA model, respectively. Among the IQA models with first four performances (i.e., GFM, SVQI, ESIM, and SQMS), the performance gain of the proposed GFM to SVQI, ESIM, and SQMS are respectively -1.14% , 1.20% , and -0.77% on SIQAD database, 4.45% , 3.32% , and 5.28% on SCID database, 1.58% , 2.25% , and 2.16% on the direct average comparison, as well as 2.57% , 2.62% , and 3.25% on weighted average comparison. Although the proposed GFM is slightly inferior to SQMS, and SQVI on SIQAD database (less than 1.2%), the proposed GFM is obviously better than SQMS and SVQI on the larger database SCID as well as two average performance comparisons. Moreover, the proposed GFM outperforms ESIM on SIQAD, SCID, and two average performance comparisons. This further reveals that the proposed GFM is superior to all the state-of-the-art models on the evaluation of SCI perceptual quality.

2) *Statistical Significance Test*: To make statistically meaningful comparison among different IQA models, it is essential to perform the statistical significance test to compare the IQA models with each other. For that, a useful statistical significance test method has been recently developed by Krasula *et al.* [43]–[45]. Specifically, this method determines whether the differences in performance are statistically significant based on a critical ration between the *Area Under the receiver operating characteristic Curve* (AUC) for two different models, which is less sensitive to the quality score range in the IQA database. Note that this method requires that the IQA databases provide the MOS with respective *standard deviation* (SD) as the input. This SD means the standard deviation of the observer votes for each particular stimulus. However, both databases (i.e., SIQAD and SCID) used in our experiments do not provide the SD, therefore, this method presented in Krasula *et al.* [43]–[45] can not be employed to perform statistical significance test in this work.

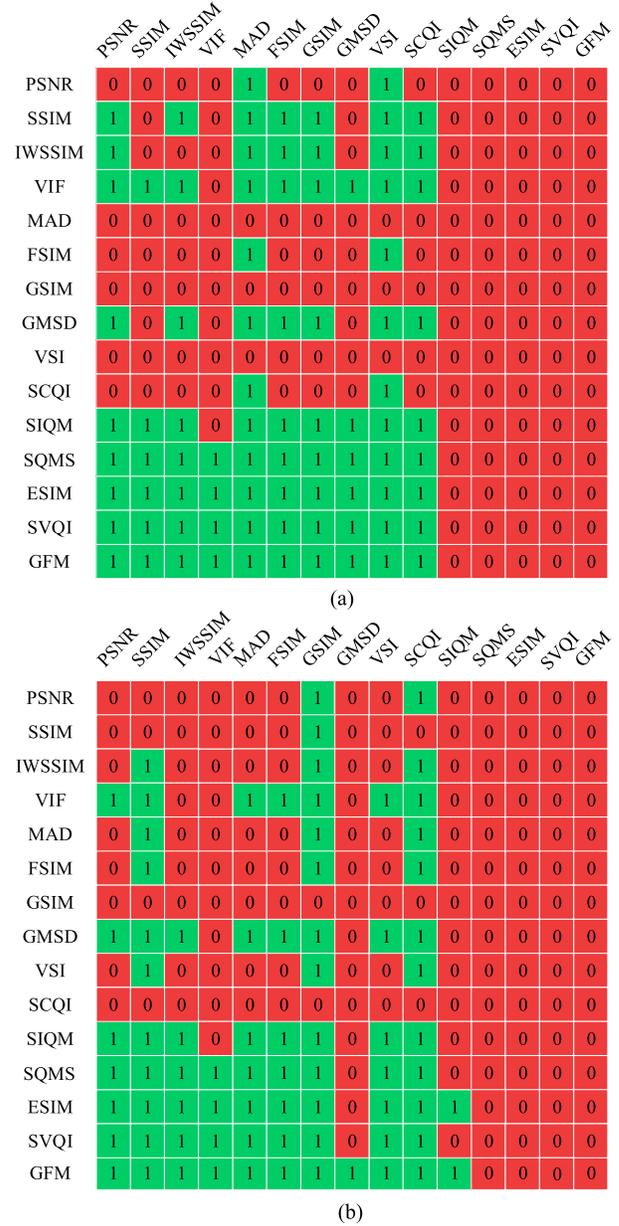


Fig. 4. The results of statistical significance test of the classical and state-of-the-art IQA models on the (a) SIQAD and (b) SCID databases. A value of ‘1’ (highlighted in green) indicates that the model in the row is significantly better than the model in the column. And a value of ‘0’ (highlighted in red) indicates that the model in the row is not significantly better than the one in the column. Note that the proposed GFM is significantly better than most of the competitors on all the two databases (10 out of 14 for SIQAD and 11 out of 14 for SCID), and no IQA model is significantly better than GFM.

Alternatively, the commonly-used F-test as suggested in [16] is performed to assess the statistical significance of the performance difference between any two IQA models under comparison. For that, based on the assumption that the prediction errors of each IQA model follow the Gaussian distribution, the one-side *F-test* with a 95% confidence level is applied to the residuals of every two models to be compared. The corresponding results on the SIQAD and SCID databases are shown in Figure 4, where a symbol “1” with green color means that the IQA model in the row significantly outperforms the

one in the column, a symbol “0” with red color represents that the IQA model is not significantly better than the one in the column. From the results reported in Fig. 4, it can be seen that on SIQAD database, the proposed GFM, SVQI, ESIM, and SQMS perform the best and they do not have significant difference. On the SCID database, the proposed GFM is significantly better than all the other models except for SQMS, ESIM, and SVQI. Moreover, on both SIQAD and SCID databases, there is no IQA model that performs significantly better than the proposed GFM. This statistical significance study further indicates that the proposed GFM has the best performance among various IQA models under comparison, which is in line with the observation from the overall performance comparison in Section III-B 1).

3) *Performance Comparison on Individual Distortion*: To more comprehensively evaluate each IQA model’s ability on assessing image quality’s degradations caused by each distortion type, Table III reports the results that are measured in PLCC, SROCC and RMSE, respectively. And all these experiments are conducted on the SIQAD and the SCID databases. In each row of this table, the first-, second- and third-ranked performance figures are highlighted in red, blue and black bold, respectively, for ease of comparison. It can be observed that the proposed GFM yields the most top-three performances compared with other IQA models. Specifically, in the comparisons in terms of SROCC, it can be seen that the proposed GFM is among the top-three models 8 times, followed by ESIM and FSIM, which are among the top-three performances 6 times and 5 times, respectively. In the comparisons in terms of PLCC, the proposed GFM is among the top-three models 8 times, followed by ESIM (6 times), VIF (5 times), and FSIM (5 times). In the comparisons in terms of RMSE, the proposed GFM is among the top-three models 9 times, followed by ESIM (6 times), VIF (5 times), FSIM (5 times), and IWSSIM (5 times). Moreover, one can see that the proposed GFM is more capable in dealing with the distortions of Gaussian blur (GB), and JPEG compression. This is because the distortions or artifacts yielded by blurring and compression will inevitably degrade the image’s *structure* and make significant changes on the extracted *structural* information. The Gabor features generated in our proposed GFM model can accurately reflect the distortions and therefore make proper assessment.

4) *Scatter Plots*: To visualize the performance yielded by the IQA models under comparison, Fig. 5 further shows a set of scatter plots of the subjective scores (i.e., MOS or DMOS) against the objective scores; the former has been generated offline and stored in the database, while the latter is computed by each IQA model. These are obtained from the following IQA models: PSNR, SSIM, IWSSIM, VIF, FSIM, GSIM, VSI, SIQM, SQMS, ESIM, SVQI and the proposed GFM. Further note that the IQA-model computed scores are the result after mapping by using equation (13).

There is a fitted line (in blue) as presented in each subplot, which is obtained by exploiting a nonlinear curve fitting process according to equation (13). In a way, this line shows the “mean” value of the performance points. That is, for each MOS value (along the vertical axis), hopefully all the

TABLE IV
PERFORMANCE OF LUMINANCE AND CHROMINANCE COMPONENT IN THE PROPOSED GFM MODEL ON THE SIQAD AND SCID DATABASES

Database	Criteria	S_G	S_C	GFM
SIQAD	PLCC	0.8797	0.5414	0.8820
	SROCC	0.8702	0.5296	0.8735
	KROCC	0.6835	0.3659	0.6876
	RMSE	6.8080	12.0349	6.7237
SCID	PLCC	0.8645	0.6976	0.8760
	SROCC	0.8597	0.6871	0.8759
	KROCC	0.6685	0.4986	0.6844
	RMSE	7.1181	10.1476	6.8313

IQA scores (along the horizontal axis) are as close to this MOS value as possible. Equivalently, the closer the performance points gather around this blue line, more accurate the IQA model’s prediction compared with the MOS scores (i.e., ground truth). For example, for the color distortion type—CSC, the scores computed by the proposed GFM are closer to the fitted blue line while that computed by the SQMS, ESIM, and SVQI are far away from the fitted blue line. It means that the proposed GFM is more consistent with the judgement on CSC distortion type made by the HVS than the SQMS, ESIM, and SVQI. Moreover, one can see that the proposed GFM has a ‘tighter’ curve fitting across different distortion types when compare with other IQA models, meaning that the proposed GFM has better overall performance than other IQA models. These can also be observed from Tables II and III. Hence, different IQA models behave quite differently on different distortion types and overall performance, which may be due to the disparity of their quality scores for different distortion types [16], [46].

5) *Performance Comparison on Luminance and Chrominance Components*: To analyze how much of the contribution coming from chrominance component in the proposed GFM, we further examine the performance of only using luminance component (denoted as S_G) or chrominance component (denoted as S_C) for evaluating the SCI perceptual quality on the SIQAD and SCID databases, respectively. This can be conducted by assigning different values to α and β in (10). Specifically, for considering the luminance only (i.e., S_G), $\alpha = 1$ and $\beta = 0$; and for the chrominance only (i.e., S_C), $\beta = 1$ and $\alpha = 0$. The corresponding results are documented in Table IV. One can see that only using luminance component achieves relatively good performance, while only using chrominance component achieves relatively poor performance. The proposed GFM that jointly explores the luminance and chrominance components yields the best performance. This indicates that the HVS is more sensitive to luminance component, while the chrominance component also plays an indispensable role in the perceptual quality assessment.

C. Computational Complexity Comparison

In addition to the accuracy as discussed previously, the *efficiency* (or computational complexity) of the IQA model is

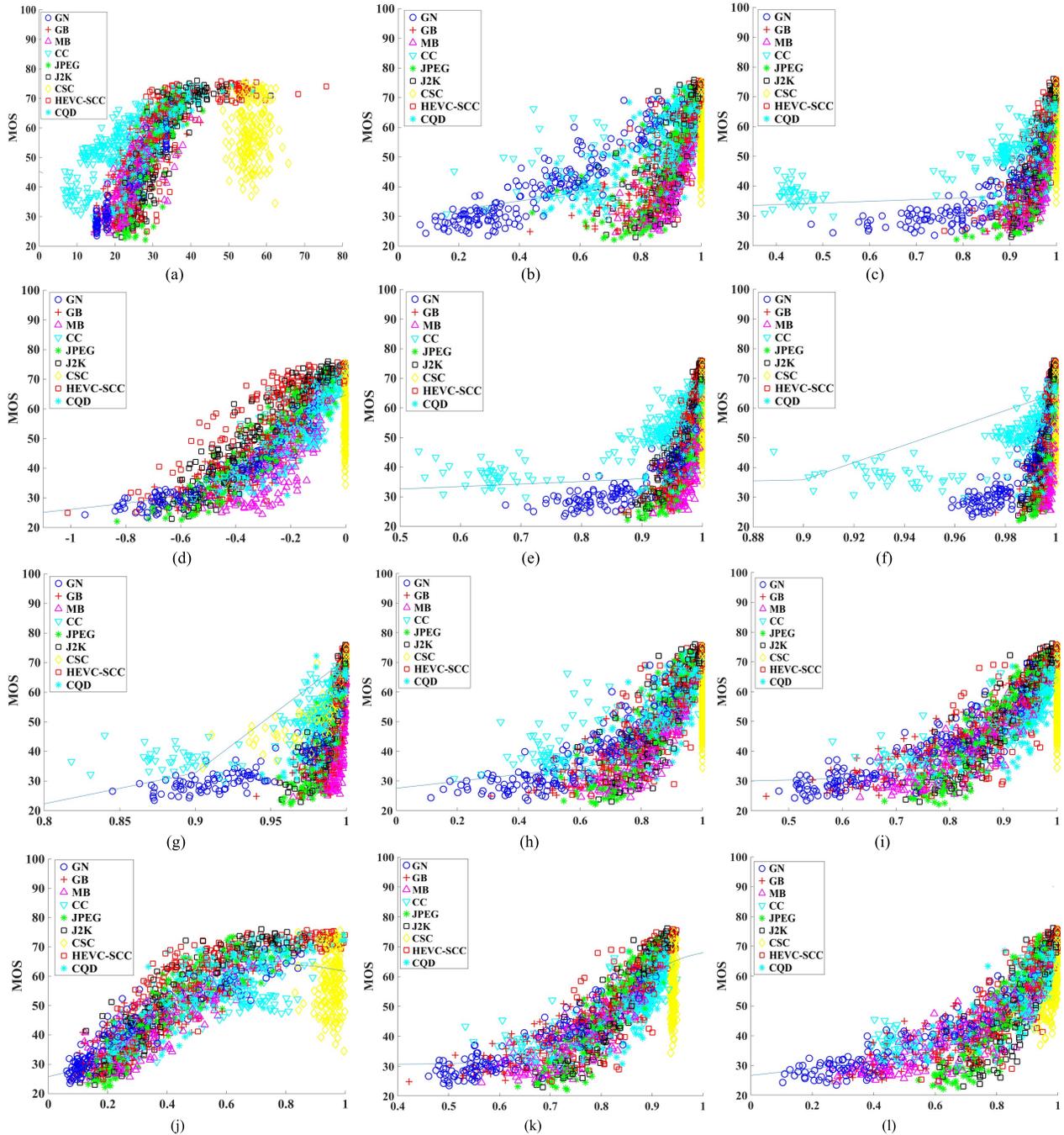


Fig. 5. A set of scatter plots showing the MOS (subjective scores) versus the objective scores computed by the IQA models: (a) PSNR; (b) SSIM; (c) IWSSIM; (d) VIF; (e) FSIM; (f) GSIM; (g) VSI; (h) SIQM; (i) SQMS; (j) ESIM; (k) SVQI; and (l) the proposed GFM, respectively. These experiments are carried out based on the SCID database. Note that there is a blue line presented in each sub-plot, and it is obtained by using a nonlinear curve fitting function.

another figure of merit that needs to be assessed, especially for practical applications. For that, the average running time per image incurred for each IQA model by experimenting on the SCID database (1,800 distorted SCIs with a resolution of 1280×720 for each image) is measured to evaluate its computational complexity. The computer used is equipped with an Intel i5-4590 CPU@3.30GHz with 8GBs of RAM, and the software platform is Matlab R2014b. Note that all the source codes of competing IQA models are obtained from

their authors or websites and are performed under the same test procedure and environment to have a meaningful and fair comparison. The run-time results are documented in Table V. It can be observed that the proposed GFM model requires a relatively low computational complexity. Although the PSNR, SSIM, GSIM, and GMSD are faster than the proposed GFM model, their accuracy measurements are much inferior to ours, as they cannot accurately describe the *perceptual* quality of the SCIs. Among the IQA models with top-four performances

TABLE V
COMPUTATIONAL COMPLEXITY COMPARISON OF DIFFERENT
IQA MODELS, MEASURED IN SECONDS PER FRAME (*spf*)

Model	Time cost (<i>spf</i>)	Model	Time cost (<i>spf</i>)
PSNR	0.0352	VSI [17]	0.2656
SSIM [11]	0.1298	SCQI [37]	0.3681
IWSSIM [15]	1.2629	SIQM [22]	0.2504
VIF [40]	2.5286	SQMS [23]	0.2047
MAD [41]	3.7694	ESIM [26]	3.0126
FSIM [13]	0.4181	SVQI [25]	2.4323
GSIM [14]	0.0855	GFM	0.1668
GMSD [16]	0.0733		

(i.e., GFM, SVQI, ESIM, and SQMS), the proposed GFM requires the least amount of computational time, while delivering fairly high accuracy.

IV. CONCLUSIONS

In this paper, a novel *screen content image* (SCI) quality assessment model, called the *Gabor feature-based model* (GFM), is proposed. The success of our approach is due to the fact that the Gabor features extracted from the SCIs as used in our GFM model are highly consistent with the response or perception of the HVS on the SCIs. Specifically, the Gabor features are extracted from the luminance component of the distorted SCI and of the reference SCI, respectively. The degree of similarity is measured on the luminance's Gabor features and two chrominance components, followed by using a Gabor-feature pooling strategy to generate the final evaluation score. Extensive experiments conducted on two large SCI databases have demonstrated that the proposed GFM model not only outperforms other state-of-the-art IQA models on objectively evaluating the image quality of the SCIs, but also demands a relatively low computational complexity.

Lastly, it is worthwhile of mentioning how to comprehensively evaluate and compare the performances of various IQA models across different IQA databases is still an essential topic in the field of IQA. This stresses the need of new performance evaluation methods and complete report of IQA databases (e.g., including both MOS/DMOS and SD values) in the future.

REFERENCES

- [1] S. Wang, L. Ma, Y. Fang, W. Lin, S. Ma, and W. Gao, "Just noticeable difference estimation for screen content images," *IEEE Trans. Image Process.*, vol. 25, no. 8, pp. 3838–3851, Aug. 2016.
- [2] S. Wang, K. Gu, X. Zhang, W. Lin, S. Ma, and W. Gao, "Reduced-reference quality assessment of screen content images," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 1, pp. 1–14, Jan. 2016.
- [3] S. Wang, K. Gu, S. Ma, and W. Gao, "Joint chroma downsampling and upsampling for screen content image," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 9, pp. 1595–1609, Sep. 2016.
- [4] Z. Ni, L. Ma, H. Zeng, C. Cai, and K.-K. Ma, "Gradient direction for screen content image quality assessment," *IEEE Signal Process. Lett.*, vol. 23, no. 10, pp. 1394–1398, Oct. 2016.
- [5] S. Huang, J. Sun, Y. Yang, Y. Fang, P. Lin, and Y. Que, "Robust single-image super-resolution based on adaptive edge-preserving smoothing regularization," *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 2650–2663, Jun. 2018.
- [6] S. Wang, K. Gu, K. Zeng, Z. Wang, and W. Lin, "Objective quality assessment and perceptual compression of screen content images," *IEEE Comput. Graph. Appl.*, vol. 38, no. 1, pp. 47–58, Jan./Feb. 2018.
- [7] A. Yang, H. Zeng, J. Chen, J. Zhu, and C. Cai, "Perceptual feature guided rate distortion optimization for high efficiency video coding," *Multidimensional Syst. Signal Process.*, vol. 28, no. 4, pp. 1249–1266, Oct. 2017.
- [8] K. Gu, D. Tao, J.-F. Qiao, and W. Lin, "Learning a no-reference quality assessment model of enhanced images with big data," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 4, pp. 1301–1313, Apr. 2018.
- [9] K. Gu, J. Zhou, J.-F. Qiao, G. Zhai, W. Lin, and A. C. Bovik, "No-reference quality assessment of screen content pictures," *IEEE Trans. Image Process.*, vol. 26, no. 8, pp. 4005–4018, Aug. 2017.
- [10] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it? A new look at signal fidelity measures," *IEEE Signal Process. Mag.*, vol. 26, no. 1, pp. 98–117, Jan. 2009.
- [11] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [12] H. R. Sheikh, A. C. Bovik, and G. de Veciana, "An information fidelity criterion for image quality assessment using natural scene statistics," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2117–2128, Dec. 2005.
- [13] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.
- [14] A. Liu, W. Lin, and M. Narwaria, "Image quality assessment based on gradient similarity," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1500–1512, Apr. 2012.
- [15] Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 5, pp. 1185–1198, May 2011.
- [16] W. Xue, L. Zhang, X. Mou, and A. C. Bovik, "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index," *IEEE Trans. Image Process.*, vol. 23, no. 2, pp. 684–695, Feb. 2014.
- [17] L. Zhang, Y. Shen, and H. Li, "VSI: A visual saliency-induced index for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 23, no. 10, pp. 4270–4281, Aug. 2014.
- [18] K. Gu, L. Li, H. Lu, X. Min, and W. Lin, "A fast reliable image quality predictor by fusing micro- and macro-structures," *IEEE Trans. Ind. Electron.*, vol. 64, no. 5, pp. 3903–3912, Jan. 2017.
- [19] W. Zhang, A. Borji, Z. Wang, P. Le Callet, and H. Liu, "The application of visual saliency models in objective image quality assessment: A statistical evaluation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 6, pp. 1266–1278, Jun. 2016.
- [20] S. Wang, K. Gu, K. Zeng, Z. Wang, and W. Lin, "Perceptual screen content image quality assessment and compression," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2015, pp. 1434–1438.
- [21] H. Yang, Y. Fang, and W. Lin, "Perceptual quality assessment of screen content images," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 4408–4421, Aug. 2015.
- [22] K. Gu, S. Wang, G. Zhai, S. Ma, and W. Lin, "Screen image quality assessment incorporating structural degradation measurement," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2015, pp. 125–128.
- [23] K. Gu *et al.*, "Saliency-guided quality assessment of screen content images," *IEEE Trans. Multimedia*, vol. 18, no. 6, pp. 1098–1110, Jun. 2016.
- [24] Y. Fang, J. Yan, J. Liu, S. Wang, Q. Li, and Z. Guo, "Objective quality assessment of screen content images by uncertainty weighting," *IEEE Trans. Image Process.*, vol. 26, no. 4, pp. 2016–2027, Apr. 2017.
- [25] K. Gu, J. Qiao, X. Min, G. Yue, W. Lin, and D. Thalmann, "Evaluating quality of screen content images via structural variation analysis," *IEEE Trans. Vis. Comput. Graphics*, to be published, doi: [10.1109/TVCG.2017.2771284](https://doi.org/10.1109/TVCG.2017.2771284).
- [26] Z. Ni, L. Ma, H. Zeng, J. Chen, C. Cai, and K.-K. Ma, "ESIM: Edge similarity for screen content image quality assessment," *IEEE Trans. Image Process.*, vol. 26, no. 10, pp. 4818–4831, Oct. 2017.
- [27] J. G. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 2, no. 7, pp. 1160–1169, 1985.
- [28] J. P. Jones and L. A. Palmer, "An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex," *J. Neurophysiol.*, vol. 58, no. 6, pp. 1233–1258, Dec. 1987.
- [29] D. P. Casasent and J. S. Smokelin, "Neural net design of macro Gabor wavelet filters for distortion-invariant object detection in clutter," *Opt. Eng.*, vol. 33, no. 7, pp. 2264–2271, Jul. 1994.

- [30] R. Mehrotra, K. R. Namuduri, and N. Ranganathan, "Gabor filter-based edge detection," *Pattern Recognit.*, vol. 25, no. 12, pp. 1479–1494, Dec. 1992.
- [31] Z. Ni, L. Ma, H. Zeng, C. Cai, and K.-K. Ma, "Screen content image quality assessment using edge model," in *Proc. IEEE Int. Conf. Image Process.*, Aug. 2016, pp. 81–85.
- [32] J. Wu, W. Lin, G. Shi, Y. Zhang, W. Dong, and Z. Chen, "Visual orientation selectivity based structure description," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 4602–4613, Nov. 2015.
- [33] R. T. Maloney and C. W. G. Clifford, "Orientation anisotropies in human primary visual cortex depend on contrast," *NeuroImage*, vol. 119, pp. 129–145, Oct. 2015.
- [34] J. M. Geusebroek, R. van den Boomgaard, A. W. M. Smeulders, and H. Geerts, "Color invariance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 12, pp. 1338–1350, Dec. 2001.
- [35] J.-M. Geusebroek, R. van den Boomgaard, A. W. M. Smeulders, and A. Dev, "Color and scale: The spatial structure of color images," in *Proc. Eur. Conf. Comput. Vis.*, Jul. 2000, pp. 331–341.
- [36] L. Xing, L. Cai, H. Zeng, J. Chen, J. Zhu, and J. Hou, "A multi-scale contrast-based image quality assessment model for multi-exposure image fusion," *Signal Process.*, vol. 145, pp. 233–240, Apr. 2018.
- [37] S.-H. Bae and M. Kim, "A novel image quality assessment with globally and locally consistent visual quality perception," *IEEE Trans. Image Process.*, vol. 25, no. 5, pp. 2392–2406, May 2016.
- [38] VQEG. (Aug. 2015). *Final Report From the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment*. [Online]. Available: <http://www.its.bldrdoc.gov/vqeg/vqeg-home.aspx>
- [39] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3440–3451, Nov. 2006.
- [40] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 430–444, Feb. 2006.
- [41] E. C. Larson and D. M. Chandler, "Most apparent distortion: Full-reference image quality assessment and the role of strategy," *J. Electron. Imag.*, vol. 19, no. 1, pp. 011006-1–011006-21, 2010.
- [42] Y. Xu, D. Liu, Y. Quan, and P. Le Callet, "Fractal analysis for reduced reference image quality assessment," *IEEE Trans. Image Process.*, vol. 24, no. 7, pp. 2098–2109, Jul. 2015.
- [43] L. Krasula, K. Fliegel, P. Le Callet, and M. Klíma, "On the accuracy of objective image and video quality models: New methodology for performance evaluation," in *Proc. Quality Multimedia Exper. (QoMEX)*, Jun. 2016, pp. 1–6.
- [44] L. Krasula, P. Le Callet, K. Fliegel, and M. Klíma, "Quality assessment of sharpened images: Challenges, methodology, and objective metrics," *IEEE Trans. Image Process.*, vol. 26, no. 3, pp. 1496–1508, Mar. 2017.
- [45] L. Krasula, M. Narwaria, K. Fliegel, and P. Le Callet, "Preference of experience in image tone-mapping: Dataset and framework for objective measures comparison," *IEEE J. Sel. Topics Signal Process.*, vol. 11, no. 1, pp. 64–74, Feb. 2017.
- [46] R. Soundararajan and A. C. Bovik, "RRED indices: Reduced reference entropic differencing for image quality assessment," *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 517–526, Feb. 2012.



Zhangkai Ni received the M.Eng. degree in communication engineering from the School of Information Science and Engineering, Huaqiao University, Xiamen, China, in 2017. He is currently a Research Engineer with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. His research interests include perceptual image processing, image restoration, and computer vision.



Huanqiang Zeng (S'10–M'13–SM'18) received the B.S. and M.S. degrees from Huaqiao University, Xiamen, China, and the Ph.D. degree from Nanyang Technological University, Singapore, all in electrical engineering.

He was a Post-Doctoral Fellow with the Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong, from 2012 to 2013, and a Research Associate with Temasek Laboratories, Nanyang Technological University, Singapore, in 2008. He is currently a Professor with the School of Information Science and Engineering, Huaqiao University, Xiamen, China.

He has authored over 80 papers in well-known international journals and conferences. His research interests include image processing, video coding, machine learning, pattern recognition, and computer vision. He is a member of the International Steering Committee of International Symposium on Intelligent Signal Processing and Communication Systems. He received the Best Paper Award from the Chinese Conference on Signal Processing 2017 (CCSP2017). He also has been actively serving as the General Co-Chair for the IEEE International Symposium on Intelligent Signal Processing and Communication Systems 2017 (ISPACS2017), the Technical Program Co-Chair for Asia-Pacific Signal and Information Processing Association Annual Summit and Conference 2017 (APSIPA ASC2017), the Area Chair for the IEEE International Conference on Visual Communications and Image Processing (VCIP2015), the Technical Program Committee Member for multiple flagship international conferences. He has been actively serving as an Associate Editor for the *IET Electronics Letters* and the *International Journal of Image and Graphics*, the Guest Editor for multiple international journals, including the *Journal of Visual Communication and Image Representation*, *Multimedia Tools and Applications*, and *Journal of Ambient Intelligence and Humanized Computing*.



Lin Ma (M'13) received the B.E. and M.E. degrees in computer science from the Harbin Institute of Technology, Harbin, China, in 2006 and 2008, respectively, and the Ph.D. degree from the Department of Electronic Engineering, The Chinese University of Hong Kong, in 2013. He was a Researcher with the Huawei Noah's Ark Laboratory, Hong Kong, from 2013 to 2016. He is currently a Principal Researcher with Tencent AI Lab, Shenzhen, China. His current research interests include the areas of computer vision, multimodal deep learning, specifically for image and language, image/video understanding, and quality assessment.

Dr. Ma received the Best Paper Award from the Pacific-Rim Conference on Multimedia in 2008. He was a recipient of the Microsoft Research Asia Fellowship in 2011. He was a finalist in HKIS Young Scientist Award in engineering science in 2012.



Junhui Hou (S'13–M'16) received the B.Eng. degree in information engineering (Talented Students Program) from the South China University of Technology, Guangzhou, China, in 2009, the M.Eng. degree in signal and information processing from Northwestern Polytechnical University, Xi'an, China, in 2012, and the Ph.D. degree from the School of Electrical and Electronic Engineering, NTU, Singapore, in 2016.

He has been an Assistant Professor with the Department of Computer Science, City University of Hong Kong, since 2017. His research interests fall into the general areas of multimedia signal processing, such as adaptive image/video representations and analysis (RGB/depth/light field/hyperspectral), static/dynamic 3D geometry representations and processing (mesh/point cloud/MoCap), and graph signal processing. He is currently serving as an Associate Editor of *The Visual Computer* and the Guest Editor for the *Journal of Visual Communication and Image Representation*. He was a recipient of the prestigious award of the Chinese Government for Outstanding Self-Financed Students Abroad in 2015.



Jing Chen (M'17) received the B.S. and M.S. degrees from Huaqiao University, Xiamen, China, and the Ph.D. degree from Xiamen University, Xiamen, all in computer science.

She is currently an Associate Professor with the School of Information Science and Engineering, Huaqiao University. Her current research interests include image processing and video coding.



Kai-Kuang Ma (S'80–M'84–SM'95–F'13) received the B.E. degree in electronic engineering from Chung Yuan Christian University, Chung-Li, Taiwan, the M.S. degree in electrical engineering from Duke University, Durham, NC, USA, and the Ph.D. degree in electrical engineering from North Carolina State University, Raleigh, NC, USA.

From 1984 to 1992, he was with IBM Corporation at Kingston, Kingston, NY, USA, and Research Triangle Park, NC, USA, focusing on various DSP and VLSI advanced product development. From 1992 to 1995, he was a member of Technical Staff with the Institute of Microelectronics, Singapore, where he is involving in digital video coding and the MPEG standards. He is currently a Full Professor with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. He has published extensively and holds one USA patent on fast motion estimation algorithm. His research interests are in the areas of digital image/video processing and computer vision, including digital image/video coding and standards, image/video segmentation, denoising, enhancement, interpolation, and super resolution. His research interests on computer vision include image matching and registration, scene analysis and recognition, and human-computer interaction.

He was serving as the Singapore MPEG Chairman and the Head of Delegation from 1997 to 2001. On the MPEG contributions, two fast motion estimation algorithms (Diamond Search and MVFAST) produced from his research group have been adopted by the MPEG-4 standard, as the reference core technology for fast motion estimation. He was the General Chair of organizing a series of international standard meetings (MPEG and JPEG), JPEG2000, and MPEG-7 workshops held in Singapore (2001).

He is a member of Sigma Xi and Eta Kappa Nu. He has been serving as the technical program committee member, reviewer, and the session chair of multiple IEEE international conferences. He is an elected member of three IEEE Technical Committees: the Image and Multidimensional Signal Processing Committee, the Multimedia Communications Committee, and the Digital Signal Processing. He is the General Co-Chair of ISPACS2017, ASIPA2017, ACCV2016 Workshop, and VCIP-2013, the Technical Program Co-Chair of ICIP-2004, ISPACS-2007, IHH-MSP-2009, and PSIVT-2010, and the Area Chair of ACCV-2009 and ACCV-2010. He is the Chairman of IEEE Signal Processing Singapore Chapter from 2000 to 2002. He has been serving as an editorial board member for several leading international journals in his research area, such as a Senior Area Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING from 2016 to 2019, an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS from 1997 to 2012, and an Associate Editor for the IEEE TRANSACTIONS ON MULTIMEDIA from 2002 to 2009, the *International Journal of Image and Graphics* from 2003 to 2015, and the *Journal of Visual Communication and Image Representation* from 2005 to 2015, the IEEE TRANSACTIONS ON IMAGE PROCESSING from 2007 to 2010, the IEEE SIGNAL PROCESSING LETTERS from 2014 to 2016, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY since 2015. He was elected as a Distinguished Lecturer of the IEEE Circuits and Systems Society for 2008–2009.