Guided Image Contrast Enhancement Based on Retrieved Images in Cloud

Shiqi Wang, *Member, IEEE*, Ke Gu, Siwei Ma, *Member, IEEE*, Weisi Lin, *Senior Member, IEEE*, Xianming Liu, *Member, IEEE*, and Wen Gao, *Fellow, IEEE*

***Abstract*—We propose a guided image contrast enhancement framework based on cloud images, in which the context-sensitive and context-free contrast is jointly improved via solving a multi-criteria optimization problem. In particular, the context-sensitive contrast is improved by performing advanced unsharp masking on the input and edge-preserving filtered images, while the context-free contrast enhancement is achieved by the sigmoid transfer mapping. To automatically determine the contrast enhancement level, the parameters in the optimization process are estimated by taking advantages of the retrieved images with similar content. For the purpose of automatically avoiding the involvement of low-quality retrieved images as the guidance, a recently developed no-reference image quality metric is adopted to rank the retrieved images from the cloud. The image complexity from the free-energy-based brain theory and the surface quality statistics in salient regions are collaboratively optimized to infer the parameters. Experimental results confirm that the proposed technique can efficiently create visually-pleasing enhanced images which are better than those produced by the classical techniques in both subjective and objective comparisons.**

***Index Terms*—Contrast enhancement, image quality assessment, retrieved images, unsharp masking, sigmoid transfer mapping, free-energy, surface quality.**

I. Introduction

CONTRAST enhancement plays an important role in the restoration of degraded images. Due to poor illumination conditions, low-quality, low-cost imaging sensors and users' inexperience and operational errors, the contrast of captured images may not be satisfactory. To recover proper details for the captured scene, a common applied procedure in low-level computer vision is enhancing the image contrast. Generally, it encompasses the context-free and context-sensitive approaches [1]. The context-sensitive approach aims to enhance the local contrast that is dependent on the rate of change in local intensity. In the literature, various algorithms have been proposed for edge enhancement with linear or nonlinear filtering [2]. However, it is prone to artifacts such as noise and ringing, as enhancement of these undesirable details will very likely introduce annoying distortions [2], [3]. Moreover, it may destroy the rank consistency of pixel levels, leading to inconsistent global contrast. The context-free approach boosts the global contrast by adopting a statistical method such as manipulating the pixel histogram. For example, in the well-known histogram modification (HM) framework, the gray-levels can be spread to generate a more uniform distribution. The limitation of methods in this category falls into the lack of adaptation on various image content, such that the modified histograms of two different images with the same probability distribution may become identical.

As a basic perceptual attribute of an image, contrast makes the representation of objects distinguishable [4]. The most frequently applied context-sensitive enhancement method is unsharp masking, which enhances the details of an image by combining the unsharp mask with the original images. The unsharp mask is generally created by a linear or nonlinear filter that amplifies the high-frequency components of the image signal. Ideally, the filter should preserve sharp edges and be robust to noise, as the enhancement of noise may introduce undesirable details. In the literature, various edge-preserving filters have been applied, such as cubic [5] and weighted median filters [6]–[8]. Moreover, in [3], the contrast enhancement of JPEG compressed images was also studied, in which the image is separated into structure and texture parts, and the texture component that contains the JPEG artifacts is further processed to reduce the compression artifacts.

The histogram equalization (HE) scheme [9], which is the most popular HM method, has been widely adopted in many image processing systems. In HE, the pixel levels are managed according to the probability of input pixel values. The classical HE method suffers from the excessively enhancement due to the existence of dominant gray levels. Moreover, the mean brightness may be no longer preserved. Therefore, various derivatives of HE methods have been proposed, such as brightness preserving bi-histogram equalization (BBHE) [10], dualistic sub-Image histogram equalization (DSIHE) [11], recursive mean-separate histogram equalization (RMSHE) [12], recursive sub-image histogram equalization (RSIHE) [13],

Manuscript received August 13, 2015; revised November 18, 2015; accepted November 20, 2015. Date of publication December 17, 2015; date of current version January 15, 2016. This work was supported in part by the National High-Tech R&D Program of China, 863 Program, under Grant 2015AA015903, in part by the National Natural Science Foundation of China under Grant 61322106, Grant 61571017, Grant 61421062, and Grant 61300110, and in part by the Shenzhen Peacock Plan. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Shu-Ching Chen.

S. Wang, S. Ma, and W. Gao are with the Institute of Digital Media, School of Electronic Engineering and Computer Science, Peking University, Beijing 100871, China (e-mail: sqwang1986@gmail.com; swma@pku.edu.cn; wgao@pku.edu.cn).

K. Gu and W. Lin are with School of Computer Engineering, Nanyang Technological University, Singapore 639798 (e-mail: gukes.doctor@gmail.com; wslin@ntu.edu.sg).

X. Liu is with the School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China (e-mail: xmliu.hit@gmail.com).

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/TMM.2015.2510326

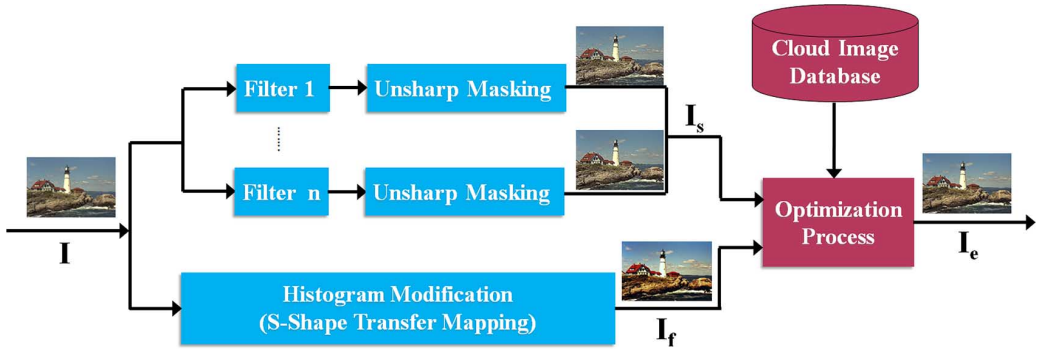


Fig. 1. Illustration of the generalized contrast enhancement framework that unifies the context-free and context-sensitive approaches.

weighted thresholded HE (WTHE) [14], generalized equalization model [15] and histogram modification framework (HMF) [16]. In addition to the HE based methods, the tone-curve adjustment such as sigmoid transfer based brightness preserving (STBP) algorithm [17] was proposed to produce visually pleasing enhanced image according to the close relationship between the third order statistic (skewness) and the surface quality [18]. In most cases, these methods are applied globally to the whole image. Alternatively, they can be easily extended to image regions [19].

One difficulty in traditional contrast enhancement algorithms lies in how to choose the optimal parameters that can produce visually-pleasing quality. Excessive enhancement may destroy the naturalness of images, while insufficient enhancement may not well improve the quality. The commonly-used manual parameter tuning is impractical for most applications as it is labor intensive and time-consuming, and more importantly, only automatic operations are feasible in many meaningful situations. Analogies to image quality assessment (IQA), central to such a problem is finding appropriate reference images that can serve as guidance. Thanks to the cloud, which is characterized by a large quantity of resources, storage and data [20], we can easily get access to high quality guidance images. Recently, with the advance of cloud computing, a huge number of images are uploaded and stored to the cloud every day. There is a high probability of finding very similar images, which are captured at the same location with different views, angles and focal lengths. Inspired by this, cloud based image processing has demonstrated its power in a variety of applications, such as image coding [21] and restoration [22].

In this paper, we attempt to address two issues in contrast enhancement: unifying context-sensitive associated with context-free methods and automatically deriving the proper enhancement level. In the literature, relating different types of contrast enhancement methods has brought more inspirations to the contrast enhancement research. In [23], spatial filters including the bi-lateral, non-local means, and steering regression etc. were unified to improve the image quality. In [15], a joint strategy was proposed by combining the white balancing and contrast enhancement together. Generally, it is widely acknowledged that a good contrast enhancement algorithm should highlight meaningful details properly and suppress visual artifacts simultaneously. By combining the context-free together with context-sensitive methods, we propose a unified contrast enhancement framework, where context-sensitive model tends to enhance the local contrast from the difference of neighbouring pixels, and meanwhile the context-free approach modifies the overall statistical pixel distributions regardless of the local content. More specifically, a multi-criteria optimization strategy is proposed, in which the input image, the enhanced image with unsharp masking, and the sigmoid transformed image are simultaneously considered. Furthermore, following this framework, the best contrast level is inferred by taking advantages of the retrieved images that are selected with the help of a no-reference (NR) IQA method, which predicts the perceived quality of each retrieved image without referencing to its corresponding pristine quality original image [24]. The basic principle behind it is from the reduced reference IQA approach [24], [25], which is based upon the assumption that if the features extracted from the enhanced and guidance images can be better matched, the enhancement quality will be closer to that of the guidance image. Inspired by the recently revealed free-energy theory [26] and surface quality [18], the enhancement level is developed in an automatic manner.

The remainder of this article is organized as follows. Section II presents the unified framework for contrast enhancement, which includes the context-sensitive and context-free approaches. In Section III, we investigate the automatic contrast enhancement level derivation with retrieved images. The proposed scheme is verified in Section IV. Section V discusses the limitations and future work. Finally, Section VI concludes this paper.

II. Unified Contrast Enhancement Framework

In this section, we present the generalized contrast enhancement framework that leverages the context-sensitive and context-free enhancement methods. The advantages of these approaches are incorporated with a joint strategy, targeting at generating a more visually-pleasing images. The generalized framework is illustrated in Fig. 1, where the enhanced images from context-sensitive approach such as unsharp masking and context-free approach such as tone-curve adjustment are fused in

TABLE I
SUMMARY OF NOTATIONS AND DEFINITIONS

Notations	Definitions
$\mathbf{I}$	Input image for contrast enhancement
$\mathbf{I_s}$	Enhanced image by context-sensitive approach
$\mathbf{I_f}$	Enhanced image by context-free approach
$\mathbf{I_g}$	Guidance image
$\mathbf{I_e}$	Enhanced image by the proposed unified contrast enhancement framework
$\boldsymbol{\phi}$	Parameters used in the context-free approach
α & β	Parameters that control the contrast enhancement level
$\mathcal{F}$	Free-energy
$\mathcal{S}$	Surface quality
λ	Parameter that balances the importance of free-energy and surface quality
D	Squared Euclidean distance
DCT2	Two-dimensional discrete cosine transform transform
IDCT2	Two-dimensional inverse discrete cosine transform transform
g	Gaussian kernel

a systematical way to generate the final image. To facilitate understanding, the summary of some notations and definitions that will be used throughout the following paper is shown in Table I.

A. Context-Sensitive Approach

The rationale behind the unsharp masking is to amplify the high-frequency components of the image signal by applying a linear or nonlinear filter [2]. In general, the filtering process can be regarded as fitting a particular model to the input [27], and the residual signal between the original input and the low-pass filtered (e.g., Gaussian smooth) images can be treated as details, which generally contain both image structure and noise. However, in unsharp masking, only image structure should be enhanced as amplification of the noise is usually undesirable. This motivated us to firstly pre-process the image to reduce noise while preserving the edge, followed by a unsharp masking process [2]. Generally, there are various edge-preserving filters, and each of them can generate a unsharp masking version, as illustrated in Fig. 1. The fusion of the processed images is regarded as the context-sensitive enhanced image. Here we demonstrate the case when applying two filters to process the image, including the impulse function and bilateral filter [28]. For bilateral filter, it possesses well edge-preserving ability, and is also easy to construct and calculate [29]. The reason of introducing the impulse function to preserve the input image information is that only applying the edge-preserving filter would give rise to detailed information loss, while the combination of them can make a good balance between the noise robustness and sharpness enhancement. This strategy can also be extended by employing more than two filters to deal with more complex scenarios.

Given the input image $\mathbf{I}$, the unsharp masking can be described as follows:

$$\mathbf{I_s} = \mathbf{I} + \omega_1 \cdot \mathbf{I_{d1}} + \omega_2 \cdot \mathbf{I_{d2}} \tag{1}$$

where $\mathbf{I_{d1}}$ and $\mathbf{I_{d2}}$ represent the high frequency signal generated following the image pre-processing with impulse function and bilateral filter, respectively. More specifically, the Gaussian smoothing is further applied on the pre-processed images and the residuals between the input and smoothened images are treated as the high frequency signal $\mathbf{I_{d1}}$ and $\mathbf{I_{d2}}$. ω_1 and ω_2 are the control factors and here the equal weighting strategy is applied ($\omega_1 = \omega_2 = 0.5$).

B. Context-Free Approach

The context-free enhancement is achieved by the sigmoid transfer mapping [17], [30]. The authors of [18] found that human eyes use skewness or a similar measure of histogram asymmetry in judging the surface quality (e.g., glossiness), and an image with a long positive tail in histogram (namely a positively skewed statistics) tends to appear darker and glossier. This hints the usage of the sigmoid mapping to improve the surface quality [17], such that the quality of the enhanced image better matches the preference of the human visual system (HVS). The context-free enhanced image $\mathbf{I_f}$ is obtained by a four-parameter logistic mapping $M_s(\cdot)$

$$\mathbf{I_f} = f_{\text{clip}}(M_s(\mathbf{I}, \boldsymbol{\phi})) = f_{\text{clip}}\left(\frac{\phi_1 - \phi_2}{1 + \exp(-\frac{(\mathbf{I}-\phi_3)}{\phi_4})} + \phi_2\right) \tag{2}$$

where f_{clip} operation is used to clip the pixel values into the range of $[0, 255]$ and $\boldsymbol{\phi} = \{\phi_1, \phi_2, \phi_3, \phi_4\}$ are parameters to be determined. This function characterizes the mapping curve, and to derive these parameters, four points on the curve, denoted as (x_i, y_i), $i = \{1, 2, 3, 4\}$ should be firstly fixed prior to the transfer process. Here x indicates the input intensity and y indicates the transfer output. As the sigmoid mapping is rolling-symmetry with respect to the straight line $y = x$, three pairs are fixed as follows, $(x_1, y_1) = (0, 0)$, $(x_2, y_2) = (l_{\max}, l_{\max})$, and $(x_3, y_3) = (\frac{l_{\max}}{2}, \frac{l_{\max}}{2})$, where $l_{\max}$ is the maximum intensity value of the input image ($l_{\max} = 255$). Another pair (x_4, y_4) can be set up to control the shape. For example, x_4 can be fixed as a certain number except x_1, x_2 and x_3. Once x_4 is fixed, given a y_4 value, the optimal control parameters $\boldsymbol{\phi}$ can be obtained by searching for the minimization of the following objective function:

$$\boldsymbol{\phi} = \arg\min_{\boldsymbol{\phi}} \sum_{i=1}^{4} |y_i - M_s(x_i, \boldsymbol{\phi})| . \tag{3}$$

Consequently, y_4 is the only control parameter that alters the curvature of the transfer function. In this work, we fix x_4 as 25 and set y_4 to be 3.

(a) (b) (c) (d)

Fig. 2. Comparison of the context-sensitive and context-free enhancement results. (a) Input image "Lighthouse"; (b) $\mathbf{I_s}$; (c) $\mathbf{I_f}$; (d) $\mathbf{I_e}$ with $\alpha = 0.5$ and $\beta = 0.5$.

C. Unified Contrast Enhancement Framework

Both the context-sensitive and context-free approaches have their own advantages in optimizing the contrast quality, and therefore in this paper we formulate the contrast enhancement as a multi-criteria optimization problem. Basically, the goal is to find an image that is close to the enhanced images as desired, but also preserve the structure from the input image $\mathbf{I}$. Therefore, given the parameters α and β that control the contrast enhancement level, the generalized framework is defined as follows:

$$\min\{\mathrm{D}(\mathbf{I_e}, \mathbf{I}) + \alpha \cdot \mathrm{D}(\mathbf{I_e}, \mathbf{I_f}) + \beta \cdot \mathrm{D}(\mathbf{I_e}, \mathbf{I_s})\} \tag{4}$$

where $\mathbf{I_e}$ denotes the enhanced image in the generalized contrast enhancement framework. The enhanced images $\mathbf{I_f}$ and $\mathbf{I_s}$ are generated by the context-free and context-sensitive approaches, respectively. To obtain an analytical solution, D is defined as the squared Euclidean distance. In general, given any two equal-length vectors $\mathbf{x}$ and $\mathbf{y}$, it is formulated as follows:

$$\mathrm{D}(\mathbf{x}, \mathbf{y}) = \sum_i (\mathbf{x}_i - \mathbf{y}_i)^2. \tag{5}$$

Combining (4) and (5), the quadratic optimization problem is derived as follows:

$$\begin{aligned}\mathbf{I_e} &= \mathrm{argmin}_{\mathbf{I_e}}\{\mathrm{D}(\mathbf{I_e}, \mathbf{I}) + \alpha \cdot \mathrm{D}(\mathbf{I_e}, \mathbf{I_f}) + \beta \cdot \mathrm{D}(\mathbf{I_e}, \mathbf{I_s})\} \\ &= \mathrm{argmin}_{\mathbf{I_e}}\{(\mathbf{I_e} - \mathbf{I})^T(\mathbf{I_e} - \mathbf{I}) + \alpha \cdot (\mathbf{I_e} - \mathbf{I_f})^T(\mathbf{I_e} - \mathbf{I_f}) \\ &\quad + \beta \cdot (\mathbf{I_e} - \mathbf{I_s})^T(\mathbf{I_e} - \mathbf{I_s})\}\end{aligned} \tag{6}$$

resulting in the following image fusion process to get the final enhanced image

$$\mathbf{I_e} = \frac{\mathbf{I} + \alpha \cdot \mathbf{I_f} + \beta \cdot \mathbf{I_s}}{1 + \alpha + \beta}. \tag{7}$$

Consequently, different α and β will create different enhancement results. For example, when α goes to infinity $\mathbf{I_e}$ converges to a global enhanced image, and when α and β turn to zero, $\mathbf{I_e}$ preserves the original input image. Therefore, various levels of contrast enhancement can be created by adjusting the two parameters. Ideally, the generation of $\mathbf{I_f}$ and $\mathbf{I_s}$ should be highly dependent on the image content. Fortunately, as α and β are automatically determined based on the retrieved guidance images with feature matching, the enhancement parameters in deriving $\mathbf{I_e}$ and $\mathbf{I_s}$ can be set as constant values that reach the maximum enhancement levels.

In Fig. 2, we demonstrate the contrast enhancement results, including the input image, $\mathbf{I_f}$, $\mathbf{I_s}$ and $\mathbf{I_e}$ with parameters $\alpha = 0.5$ and $\beta = 0.5$. As the proposed scheme incorporates the advantages of both the context-free and context-sensitive approaches, the finally enhanced images $\mathbf{I_e}$ appears more natural and visually-pleasing. It is observed that the unsharp masking of the input image can preserve more details. Moreover, we observe that better surface quality is achieved with the sigmoid transfer, as shown in Fig. 2(c). As a matter of fact, by setting $\alpha \geq 0$ and $\beta \geq 0$ in the optimization process, the enhanced image is upper-bounded by the $\mathbf{I_f}$ and $\mathbf{I_s}$, and lower-bounded by the input image $\mathbf{I}$.

III. Guided Contrast Enhancement Scheme

In the literature, with numerous approaches proposed to enhance image contrast, much less work has been dedicated to automatic determination of contrast enhancement levels. However, improper enhancement level may cause excessive enhancement, leading to unnatural images, which are not desirable in real applications. To avoid manually selection of contrast enhancement parameters, we employ retrieved images to influence the automatic contrast enhancement level derivation. Generally speaking, a large number of near and partial duplicate images in the cloud are captured at the similar location, but with different scales, orientations, and focus lengths. According to the statistics, around 300 million photos are uploaded to Facebook every day. These images are captured with different devices and processed by different softwares, and many of them are highly correlated. For instance, for a given landmark image, it is easy to retrieve many highly correlated images [31]. Another particular example is photo album [32], [33], in which the images have similar content as well as semantic meanings.

In general, automatically enhancing an image to the desired contrast level is difficult, as quality assessment of contrast enhancement is still a non-trivial task [34]. Thanks to the availability of a large number of images from cloud, which make the automatic contrast enhancement from the cognitive point of view possible. Here we make an assumption that the guidance images from cloud could have the perfect enhancement quality, as many of them may have already been manually selected and processed when they were uploaded. To realize the automatic guidance image selection, an NR-IQA method is applied to re-rank the retrieved images such that the one with best quality from the retrieved images is treated as the "guidance". Specifically, given an input image, image retrieval system will return a batch of images with similar content. In additional to the scenario of photo album contrast enhancement, which will be

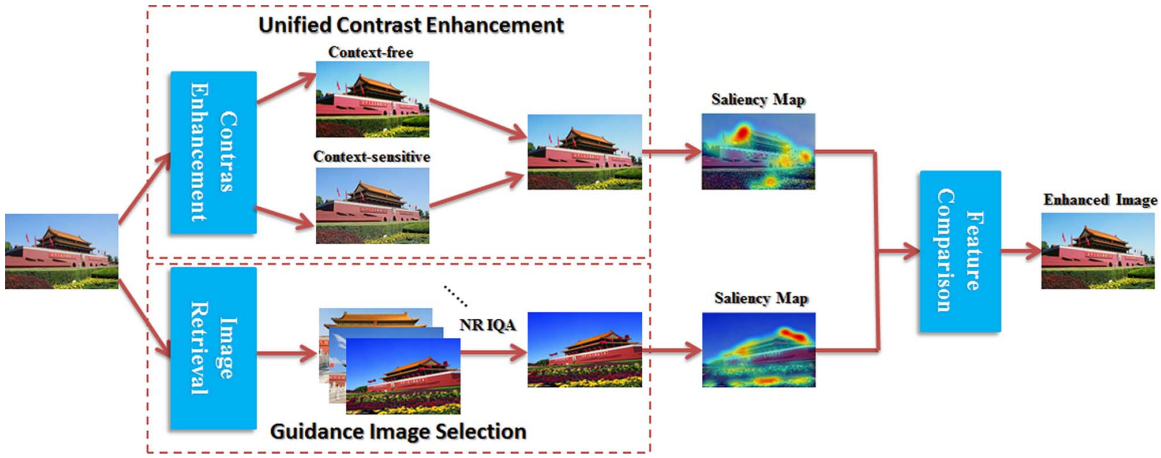


Fig. 3. Flowchart of the automatic contrast enhancement scheme.

discussed in Section III-A, we apply the content-based image retrieval method to obtain the content-similar images. Since content-based method compares the image distances based on the image content such as colors, texture and shapes, it is more appropriate for the task of locating the appropriate guidance images. Subsequently, the NR-IQA is performed to rank the guidance image. As aforementioned, there may exist many high quality retrieved images in cloud, and employing more images as guidance may improve the robustness of the algorithm at the cost of higher computational complexity. In this work, we demonstrate the solution with only one guidance image. It is feasible to extend the scheme by using more guidance images, which will be further discussed in Section IV-B.

The remaining task is to derive the contrast enhancement level that best matches the guidance image. There are various methods for matching the quality of two images. With registered image pairs, this problem can be formulated to be designing full reference IQA methods to compute the visual distance between them. Nevertheless, as the images from cloud may have different orientation and scale, it is difficult to directly apply the full reference IQA. Locating keypoints and extracting descriptors can be an alternative solution, such as computing the well-known Scale-invariant feature transform (SIFT) [35] and Speeded Up Robust Features (SURF) [36] descriptors. However, the design philosophy of these features has little correlation with image quality. Inspired by the reduced reference IQA methods, in this work a few extracted features that can summary the whole image are utilized for contrast enhancement quality comparison and matching. As a matter of fact, contrast is highly revelent to image complexity and surface quality statistics, which inspires us to explore the contrast level derivation with recent findings on brain theory.

A. Guidance Image Selection With NR-IQA

The guidance image selection process is illustrated in Fig. 3. Initially, the input image is used to retrieve highly correlated images from cloud. Since images in the cloud can be either perfect quality or corrupted by various types of distortions, it is desirable to apply an NR-IQA algorithm to rank these images and select the best one. This motivates us to adopt a recently proposed NR-IQA method [37], which achieves state-of-the-art prediction accuracy. It is also worth noting that any other advanced NR-IQA algorithms can be applied here in principle.

In the NR-IQA method, features that are based on the free-energy theory are used to establish the constructive IQA model. Based on the observation that there exists an approximate linear relationship between the structural degradation information and the free-energy of original images, the structural degradation $SDM_\mu(\mathbf{I})$ and $SDM_\sigma(\mathbf{I})$ are compared with the free-energy feature $\mathcal{F}(\mathbf{I})$, and the divergence between them $NRD_1(\mathbf{I}) = \mathcal{F}(\mathbf{I}) - (\xi_1 \cdot SDM_\mu(\mathbf{I}) + \varphi_1)$ and $NRD_2(\mathbf{I}) = \mathcal{F}(\mathbf{I}) - (\xi_2 \cdot SDM_\sigma(\mathbf{I}) + \varphi_2)$ are employed for quality evaluation. The computation of free-energy is described in Section III-B. The structural degradation is evaluated by

$$SDM_\mu(\mathbf{I}) = E\left(\frac{\sigma_{\mu_\mathbf{I}\bar{\mu}_\mathbf{I}} + C_1}{\sigma_{\mu_\mathbf{I}}\sigma_{\bar{\mu}_\mathbf{I}} + C_1}\right)$$
$$SDM_\sigma(\mathbf{I}) = E\left(\frac{\sigma_{\sigma_\mathbf{I}\bar{\sigma}_\mathbf{I}} + C_1}{\sigma_{\sigma_\mathbf{I}}\sigma_{\bar{\sigma}_\mathbf{I}} + C_1}\right) \tag{8}$$

where $E(\cdot)$ denotes the mathematical expectation operator and C_1 is a small positive stability constant that accounts for the saturation effects. Here $\mu_\mathbf{I}$ and $\sigma_\mathbf{I}$ are defined to be local mean and standard deviation of with a 2D circularly-symmetric Gaussian weighting function. By contrast, $\bar{\mu}_\mathbf{I}$ and $\bar{\sigma}_\mathbf{I}$ are the local mean and standard deviation using the impulse function instead of the Gaussian weighting function. $\sigma_{\mu_\mathbf{I}\bar{\mu}_\mathbf{I}}$ denotes the local covariance between two vectors $\mu_\mathbf{I}$ and $\bar{\mu}_\mathbf{I}$, such that the structural degradation information corresponds to the cosine of the angle between the two mean vectors. Analogies to that, $\sigma_{\sigma_\mathbf{I}\bar{\sigma}_\mathbf{I}}$ denotes the local covariance between two vectors $\sigma_\mathbf{I}$ and $\bar{\sigma}_\mathbf{I}$.

The design physiology is due to the fact that both $NRD_1(\mathbf{I})$ and $NRD_2(\mathbf{I})$ values of high-quality images (with very few distortions) are quite close to zero, whereas they will be far from zero when distortions become larger. The parameters ξ_1, φ_1, ξ_2 and φ_2 are trained based on the least square method using the Berkeley database [38]. In additional to these features, image size is also considered as a criterion to exclude low resolution guidance images.

In addition to the single image contrast enhancement with retrieved images from cloud, the framework can be further extended to "photo album contrast enhancement". Photo album

[32], [33] is regarded as a special existence form in cloud storage. For instance, people may want to enhance a batch of images that are taken at similar places. Besides the contrast enhancement for each image in the photo album from the guidance images selected in cloud, alternative strategies based on the proposed framework can be applied as well. For example, they can manually select one guidance image when browsing these images (as the total number of images may not be infinitely large), or manually enhance one image to perfect contrast and treat this image as guidance. Subsequently, other images can be automatically enhanced with the guidance information.

B. Free-Energy-Based Brain Theory

The free-energy theory, which was recently introduced by Friston *et al.* in [26], attempts to explain and unify several brain theories in biological and physical sciences about human action, perception and learning. The basic premise of the free-energy based brain theory is that the cognitive process is manipulated by an internal generative model (IGM). The human brain can actively infer predictions of the meaningful information of input visual signals and avoid the residual uncertainty in a constructive manner. In this work, the free-energy is applied both in NR-IQA as well as the feature matching for contrast enhancement level derivation.

Assuming that the IGM for visual perception is parametric, which explains the scene by adjusting the parameter $\mathbf{v}$. Given the input image $\mathbf{I}$, its "surprise" (determined by entropy) is evaluated by integrating the joint distribution $P(\mathbf{I}, \mathbf{v})$ over the space of model parameters $\mathbf{v}$ [39]

$$-\log P(\mathbf{I}) = -\log \int P(\mathbf{I}, \mathbf{v}) d\mathbf{v}. \tag{9}$$

Since the precise expression of joint distribution $P(\mathbf{I}, \mathbf{v})$ is still well beyond our current level of knowledge about the details of how the brains are working, a dummy term $Q(\mathbf{v}|\mathbf{I})$ is integrated into both the denominator and numerator in (9), which can be rewritten as follows:

$$-\log P(\mathbf{I}) = -\log \int Q(\mathbf{v}|\mathbf{I}) \frac{P(\mathbf{I}, \mathbf{v})}{Q(\mathbf{v}|\mathbf{I})} d\mathbf{v} \tag{10}$$

where $Q(\mathbf{v}|\mathbf{I})$ is an posterior distribution of the model parameters given the input image signal $\mathbf{I}$. This can be regarded as the posterior approximation to the true posterior of the model parameters $P(\mathbf{v}|\mathbf{I})$ in the cognitive process. Another interpretation is that when we perceives the image $\mathbf{I}$, the parameter vector $\mathbf{v}$ of $Q(\mathbf{v}|\mathbf{I})$ is adjusted to obtain the optimal explanation of $\mathbf{I}$, such that the discrepancy between the approximate posterior $Q(\mathbf{v}|\mathbf{I})$ and the true posterior $P(\mathbf{v}|\mathbf{I})$ is minimized [39]. The same technique has been used in ensemble learning or in a variational Bayesian estimation framework. The negative "surprise" can also be interpreted as the log evidence of the image data given the model. In this manner, the minimization of surprise is equivalent with the maximization of the model evidence.

By applying Jensen's inequality, from (10) we derive that

$$-\log P(\mathbf{I}) \leq -\int Q(\mathbf{v}|\mathbf{I}) \log \frac{P(\mathbf{I}, \mathbf{v})}{Q(\mathbf{v}|\mathbf{I})} d\mathbf{v} \tag{11}$$

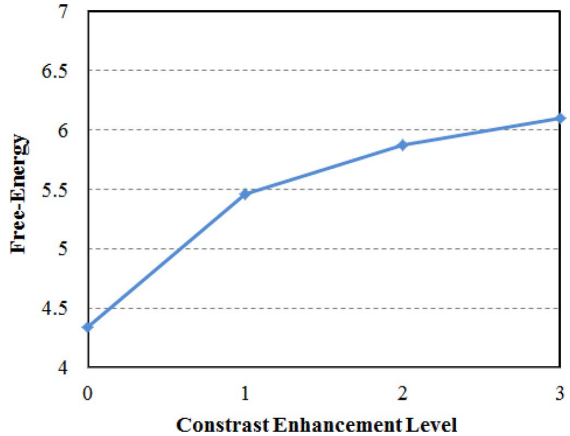


Fig. 4. The relationship between the contrast enhancement level (controlled by parameter β) and free-energy (evaluated in terms of residual entropy) for image "Lighthouse".

and the free-energy is defined as follows:

$$\mathcal{F}(\mathbf{I}) = -\int Q(\mathbf{v}|\mathbf{I}) \log \frac{P(\mathbf{I}, \mathbf{v})}{Q(\mathbf{v}|\mathbf{I})} d\mathbf{v}. \tag{12}$$

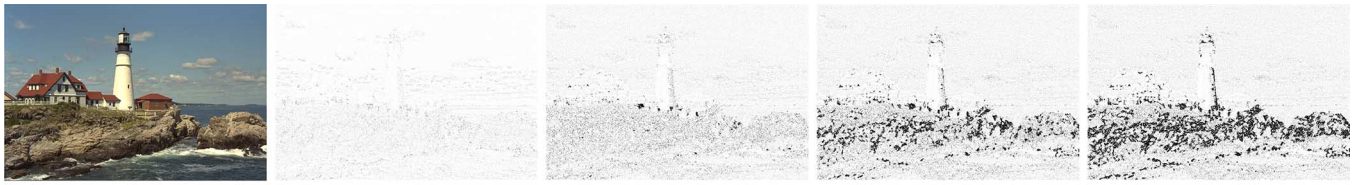
The free-energy $\mathcal{F}(\mathbf{I})$ defines the upper bound of the input image information as $-\log P(\mathbf{I}) \leq \mathcal{F}(\mathbf{I})$. In [39], it is shown that the free-energy can be characterized by the total description length for the kth order autoregressive (AR) model

$$\mathcal{F}(\mathbf{I}) = -\log P(\mathbf{I}|\mathbf{v}) + \frac{k}{2} \log N \quad \text{with} \quad N \to \infty \tag{13}$$

where N denotes the total number of pixels in the images. Thus, the entropy of the prediction residuals between the input and predicted images plus the model cost can be used to estimate $\mathcal{F}(\boldsymbol{I})$. The residuals are also known as the disorderly information that cannot be well explained by the HVS. The derivation process of the AR coefficients can be found in [40]. In this stage, a fixed-model order is chosen, resulting in the ignorance of the second term $\frac{k}{2} \log N$ in comparison.

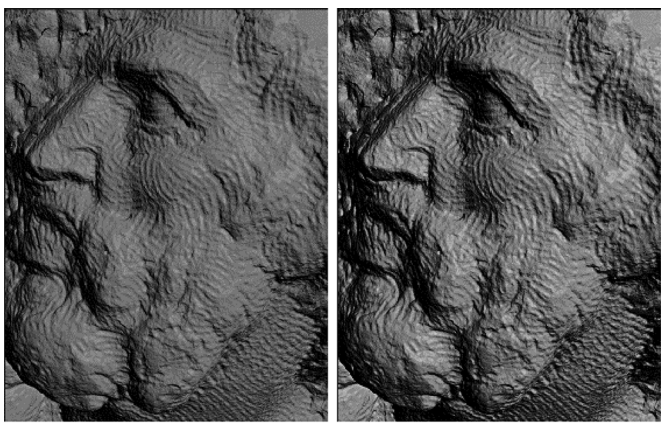
The free-energy based brain theory also reveals that the HVS cannot fully process all of the sensation information and tries to avoid some surprises with uncertainties, which can be regarded as free-energy. In practice, positive contrast change renders high quality images by highlighting the visibility details, which produces more informative content. When perceiving the positive contrast image, the additional informative content will make the image more difficult to describe, as in general the HVS has stronger description ability for low-complexity images than high-complexity versions [41]. This leads to higher free-energy, and vice versa. The prior information from the guidance is able to predict the appropriate free-energy of a visually-pleasing image with a good contrast, which is very efficient in deriving the contrast enhancement levels.

The relationship between the contrast enhancement level and $\mathcal{F}$ is demonstrated in Fig. 4, where the enhancement level is controlled by the context-sensitive parameter β. It is observed that the free-energy $\mathcal{F}$ is increasing monotonously with the enhancement level, which indicates that it has strong description ability for contrast. Moreover, the residual maps between the original and predicted images with AR model are shown in Fig. 5. This



(a) (b) (c) (d) (e)

Fig. 5. Input image and prediction residuals with AR model in terms of different enhancement levels (the residuals are enlarged for visualization and brightness indicates a smaller residual). (a) Image "Lighthouse"; (b) $\beta = 0$; (c) $\beta = 1$; (d) $\beta = 2$; (e) $\beta = 3$.



(a) (b)

Fig. 6. Surface quality comparison of two synthetic images of Michelangelo's St. Matthew sculpture with the same mean luminance [18]. (a) Skewness: -0.62; (b) skewness: -0.13.

also verifies that the energy of prediction residual is highly relevant with the enhancement strength, and high contrast image results in large free-energy.

C. Surface Quality Statistics

It is generally acknowledged that contrast provides a valid cue to surface reflectance [42] and shares more high-level properties with gloss in human perception dimensions [43]. In [18], it is discovered that the human observers use skewness $\mathcal{S}$, or histogram asymmetry to judge the surface quality. In Fig. 6, two renderings of a three-dimensional model of Michelangelos sculpture of St Matthew are demonstrated. It is observed that the right image appears darker and glossier than the left one, and moreover, the skewness of the left image is lower than the right one. Skewness is a measure of the asymmetry of a distribution, and it indicates the balance between the positive and negative tails. As the glossiness is increased regardless of the albedo, the skewness of image histogram tends to increase. To explain the skewness from physiology in human brains, a possible neural mechanism was proposed in [18], which includes on-center and off-center cells and an accelerating nonlinearity to compute the subband skewness.

D. Contrast Level Derivation From Guidance

The flowchart of the contrast enhancement level derivation is illustrated in Fig. 3. Given an input image, the context-free and context-sensitive methods are applied to fuse the enhanced versions. Since human cortical cells are likely to be hard wired to preferentially respond to high contrast stimulus [44], and image saliency is sensitive to noise injection whereas immune to contrast enhancement [30], a saliency region detection algorithm is performed on both the guidance and the fused images. Specifically, the saliency regions are detected by applying a threshold on the saliency maps of the guidance and the fused images, respectively. Subsequently, the features including the free-energy and surface quality within the salient regions are extracted. The parameters that can lead to a minimal feature distance are employed to generate the final enhanced image.

1) Visual Saliency Detection: In the literature, various approaches of visual saliency detection have been proposed [45] and successfully applied in image processing tasks, e.g., IQA metrics design [46], [47] and high efficiency video compression [48]. The recently proposed image signature model [49] makes use of the sign of each DCT component to generate the saliency map. As such, this model just requires a single bit per component, making it efficient with very low cost of computational complexity. Specifically, the image signature is defined as

$$\mathrm{ImgSignature}(\mathbf{I}) = \mathrm{sign}(\mathrm{DCT2}(\mathbf{I})) \tag{14}$$

where $\mathrm{sign}(\cdot)$ is the entrywise sign operator and for each entry with the input value ξ

$$\mathrm{sign}(\xi) = \begin{cases} 1, & \xi > 0 \\ 0, & \xi = 0 \\ -1, & \xi < 0. \end{cases} \tag{15}$$

Subsequently, the reconstructed image is derived by

$$\bar{\mathbf{I}} = \mathrm{IDCT2}(\mathrm{ImgSignature}(\mathbf{I})) \tag{16}$$

where DCT2 and IDCT2 respectively stand for discrete cosine and inverse discrete cosine transforms for the two dimensional image signal. The saliency map is finally obtained by smoothing the squared reconstructed image

$$\mathrm{SaliencyMap} = g * (\bar{\mathbf{I}} \circ \bar{\mathbf{I}}) \tag{17}$$

where g is a Gaussian kernel and '$\circ$' and '$*$' are the entry-wise and convolution product operators respectively. In practical implementations, the saliency map can be converted into a intensity image in the range from 0.0 to 1.0, and with the empirically determined threshold the saliency regions can be classified.

2) Automatic Contrast Enhancement: Based on the analysis of free-energy and surface quality statistics, two features are extracted from the guidance and fused images. Instead of pixel-wisely or patch-wisely comparing image pairs, the global features can achieve high efficiency in dimension reduction,

Fig. 7. One hundred test images.

Fig. 8. One hundred guidance images.

(a) (b) (c) (d)

Fig. 9. Contrast enhancement results comparison. (a) Guidance image; (b) input image; (c) enhanced image with proposed method; (d) HE output.

and also provide good accuracy in summarizing the contrast strength. As such, contrast matching can be converted to the optimization problem based on the guidance image and the fused image as follows:

$$(\alpha^*,\beta^*)=\mathrm{argmin}_{\alpha^*,\beta^*}\left(|\mathcal{F}(\mathbf{I_g})-\mathcal{F}\left(\frac{\mathbf{I}+\alpha\cdot\mathbf{I_f}+\beta\cdot\mathbf{I_s}}{1+\alpha+\beta}\right)| +\lambda|\mathcal{S}(\mathbf{I_g})-\mathcal{S}\left(\frac{\mathbf{I}+\alpha\cdot\mathbf{I_f}+\beta\cdot\mathbf{I_s}}{1+\alpha+\beta}\right)|\right) \tag{18}$$

where the parameter λ balances the magnitude and importance between the complexity measure and skewness measure. Parameters α^* and β^* are the optimized values that lead to an appropriate enhancement level. To facilitate the comparison and reduce the computational complexity, the guidance and fused images are firstly downsampled to the same scale for the feature computation. The final enhanced image is obtained by fusing $\mathbf{I}$, $\mathbf{I_f}$ and $\mathbf{I_s}$ with parameters α^* and β^* according to (7).

It is not straightforward to differentiate the objective function with regards to α and β. Practically, we perform a search to obtain the best enhancement level, as given in [50]. More specifically, the extracted feature is firstly compared in terms of grid within a given range of parameters, followed by a fine search within the reduced ranges.

IV. Experimental Results

In this section, we evaluate the performance of the proposed contrast enhancement scheme in various aspects. Firstly, we present the contrast enhancement results that are evaluated objectively and subjectively in comparison with the methods including the classical HE as well as the more advanced methods such as RSIHE [13], WTHE [14] and HMF [16]. Secondly, the scenario of multiple guidance images is considered, and subjective tests are further conducted to investigate this issue. Thirdly, the robustness of the scheme is evaluated in cases when there is no appropriate guidance image. Finally, the complexity of the proposed scheme is analyzed and demonstrated.

A. Contrast Enhancement Evaluation

As illustrated in Figs. 7 and 8, to evaluate the proposed scheme, in total 100 images are used for testing, which cover a wide range of applications and scenarios including humans, animals, indoor, outdoor, landmarks, products, etc. The corresponding guidance images are obtained based on the similar content retrieved images, which are subsequently ranked via the NR-IQA method. In Fig. 9, we demonstrate the guidance, test, as well as enhanced images with methods including the proposed and classical HE, respectively. As given in Fig. 9(d), the HE usually produces too-dark and too-bright regions, making the image excessively enhanced and unnatural. This results from the large backward-difference of the equalized

(a) (b) (c) (d)

Fig. 10. Demonstration of the blocking artifacts in contrast enhancement (cropped for visualization). (a) Input image; (b) enhanced image with proposed method; (c) HMF output; (d) RSIHE output.

histogram. In comparison, our model not only appropriately enhances the detailed information, but also generates much glossier images for both the natural scenes and objects, being enabled by the sigmoid transfer mapping.

Another interesting observation is that when the input images are corrupted (e.g., the commonly seen JPEG compression), the artifacts after contrast enhancement can be easily boosted by the traditional enhancement methods. However, our method suffers less from this problem thanks to the guided contrast enhancement strategy. As the free-energy increases monotonously with the injected artifacts, such as noise and blocking [39], the contrast enhancement level is suppressed accordingly to achieve matching statistics, resulting in less distortion introduced. One example is demonstrated in Fig. 10, where the cropped images are demonstrated for visualization. Though the input image is in JPEG format, it is observed that the artifacts in the input image are not obvious. However, the variants of HE approach such as HMF and RSIHE clearly amplify the blocking artifacts in contrast enhancement, whereas our approach is able to create compelling images and avoid these apparent artifacts.

To further validate the proposed scheme, we have conducted both objective and subjective experiments. In the subjective test, 25 naive subjects were invited to give an integer score between 0 and 10 for the perceptual quality of each contrast enhanced image, including the proposed, HE, WTHE [14], RSIHE [13] and HMF [16] methods. The subjects participating in this test include 11 females and 14 males, and their ages are between 18 to 31. The principle is to judge the quality based on the basic attributes of natural images such as quality, contrast, as well as naturalness. Moreover, 0 denotes the worst quality and 10 the best. The final quality score of each individual image is computed as the average of subjective scores, termed as mean opinion score (MOS), for all subjects. The subjective testing is conducted with a conventional monitor and each image is randomly played twice. In Fig. 11, we demonstrate the MOSs for each image. It is observed that, the proposed scheme has obtained outstanding results by winning the first place on 82 image sets.

Moreover, objective assessment measures are further employed to evaluate the quality of the contrast changed images. We adopt the analysis of distortion distribution-based (ADD) quality measures, including ADD-SSIM and ADD-GSIM [51], as both of them achieve the state-of-the-art prediction accuracy on contrast-changed images. The average scores over the 100 test images are shown in Table II. Following the methodology of image quality assessment, the scores are obtained by a nonlinear mapping with the logistic regression function, and

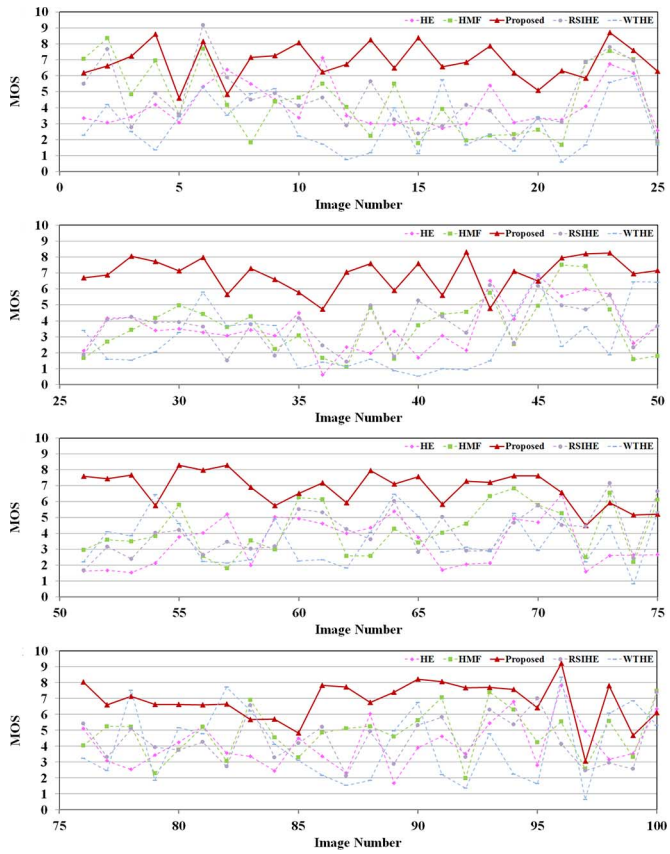


Fig. 11. Mean opinion score for each image.

the mapping parameters are trained using the contrast change image database [52]. It is observed that the proposed scheme produces significantly better images than HE, HMF, RSIHE and WTHE methods in terms of both ADD-SSIM and ADD-GSIM. These results provide proof of the superiority of the proposed scheme in contrast enhancement applications, which originates from both the unified contrast enhancement framework and appropriate enhancement levels derived from retrieved images in similar content.

B. Impact of the Number of Guidance Images

Conceptually, involving more appropriate guidance images may improve the robustness and provide more accurate feature values that represent the perfect image contrast. As such, the robustness of the method can subsequently get improved, leading

TABLE II
OBJECTIVE PERFORMANCE COMPARISONS IN TERMS OF ADD-SSIM AND ADD-GSIM [51]

	HE	HMF	Proposed	RSIHE	WTHE
ADD-SSIM	2.614	2.867	3.329	2.886	2.699
ADD-GSIM	2.514	2.791	3.249	2.798	2.681

Fig. 12. Selected images in the investigation of the impact of NoG and robustness evaluation.

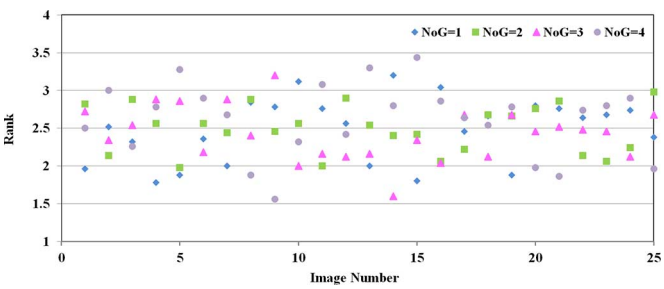


Fig. 13. Quality rank for each enhanced image generated by different NoG images.

to better enhanced images. To validate this assumption, we conduct a subjective test in which 25 images are selected for evaluation, which are shown in Fig. 12. The same 25 subjects as in Section IV-A are invited. The numbers of guidance (NoG) images are from 1 to 4. The multiple guidance images are obtained after using NR-IQA to rank the retrieved images. To distinguish the differences between the enhanced images, we adopt a different strategy in subjective testing, where subjects are suggested to provide the rank among the four images. Specifically, "4" indicates best quality among these four images and "1" indicates the lowest quality. The results are demonstrated in Fig. 13. One can discern that employing more images can improve the contrast enhancement results in most cases. Moreover, in Figs. 14 and 15, we demonstrate two examples of the enhanced images with different NoG, where the original input, enhanced images with one and three guidance images are shown. In Fig. 14, increasing the NoG significantly boost the contrast of the enhanced image, leading to better visual quality. By contrast, in Fig. 15, proper contrast is assigned when increasing the NoG, such that the enhanced image looks more natural. Therefore, it is concluded that the NoG also plays an important role in avoiding the excessive and insufficient enhancement. However, there exists a trade-off when using more guidance images, as more guidance images may impose more computational overhead on the calculation of the features values as well. To validate the concept of the approach, we only use one guidance in other experiments. How to determine the optimal number of guidance images will be studied in the future.

C. Robustness Analysis

It is worth noting that the appropriate guidance images may not always exist in cloud. One reason lies in the fact that retrieved images in similar content can not always be found. Another reason is that the selected guidance image may have improper contrast, especially when there are only low quality images after retrieval. To deal with these cases, in our contrast enhancement, it is not allowed to reduce the image contrast by forcing $\alpha \geq 0$ and $\beta \geq 0$. As such, contrast can only get enhanced with the proposed method, implying that the contrast enhanced image is lower bounded by the input image $\mathbf{I}$, and upper bound by $\mathbf{I_s}$ and $\mathbf{I_f}$. Finally, three possibilities may arise after the contrast enhancement, including insufficient, proper and excessive enhancement. In particular, insufficient enhancement will improve the image quality, but the improvement may not achieve the desired level. The proper enhancement exactly meets the target level, which produces the enhanced image with best quality. By contrast, the excessive enhancement may produce unnatural images.

To further investigate the robustness issue, two additional experiments are conducted, and again images in Fig. 12 are used for testing. In the first experiment, the corresponding guidance images are randomly selected from the 100 guidance images in Fig. 8 to simulate the case when there do not exist retrieved images in similar content. In this manner, the content of input test and guidance are totally different. In the second experiment, the corresponding guidance images that are selected in Section IV-A are further blurred with a Gaussian low-pass filter, targeting at generating the low quality images with improper contrast.

To examine whether the quality of the input image has been improved, we perform a comprehensive subjective study based on the two-alternative forced-choice (2AFC) method. The 2AFC method is widely used in psychophysical studies [53]–[55], where in each trial, a subject is shown a pair of images/videos and is asked (forced) to choose the one he/she thinks to have better quality. Therefore, in each test, 25 pairs of images are evaluated, where each pair is composed of one input image and one enhanced image. The results of the subjective tests are reported in Fig. 16. In each figure, the percentage by which the subjects are in favor of the original input against the enhanced image is shown. We also plot the error bars ($\pm$ one standard deviation between the measurements) over the 25 image pairs. For the first experiment, as can be observed in Fig. 16(a), it turns out that for eight images subjects show obvious preference to the input image against the enhance images (percentage > 0.65). This originates from the reason that these images get excessive enhancement, leading to unnatural

(a) (b) (c)

Fig. 14. Contrast enhancement results comparison (insufficient enhancement case). (a) Test image; (b) enhanced image with $\mathrm{NoG} = 1$; (c) enhanced image with $\mathrm{NoG} = 3$.

(a) (b) (c)

Fig. 15. Contrast enhancement results comparison (excessive enhancement case). (a) Test image; (b) enhanced image with $\mathrm{NoG} = 1$; (c) enhanced image with $\mathrm{NoG} = 3$.

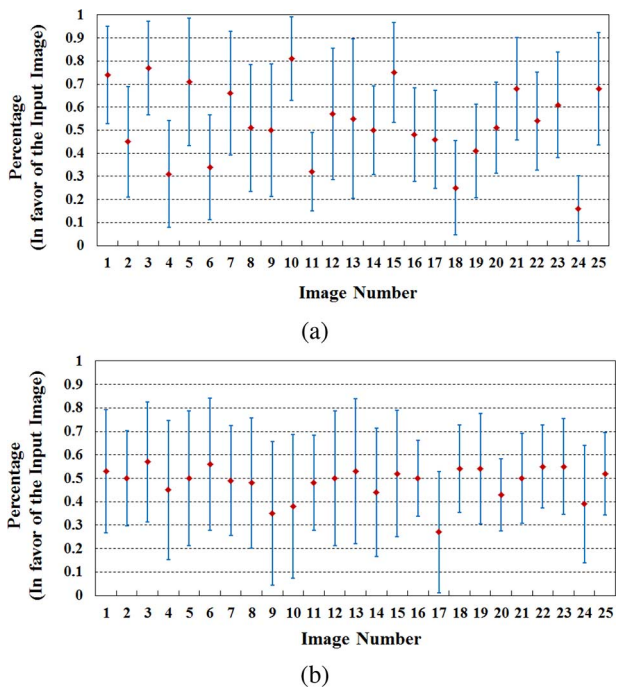


Fig. 16. Subjective tests on the examination of robustness (mean and standard deviation of preference for individual enhanced image). (a) Guidance images with different content case; (b) guidance images with low quality case.

looking. One example is shown in Fig. 17. Moreover, although the guidance images have different content, the quality of five images get significantly improved (percentage < 0.35). This is due to the fact that these guidance images occasionally provide useful features that lead to better quality images, even if the derived contrast level may not be perfect. For the rest of

(a) (b)

Fig. 17. Illustrated of the excessively enhanced image with inappropriate guidance information. (a) Input image; (b) excessively enhanced image.

the images, subjects are difficult to judge whether the enhanced or the input image has better quality. In the second experiment, when the guidance images are blurred, for most of the cases the percentage is still close to 50% and all error bars cross the 50% line. This is because that our proposed approach dose not allow the reduction of image contrast. Even though most of the images cannot get enhanced, their quality has not been degraded either. These results further illustrate the robustness of the proposed scheme.

D. Complexity Analysis

In this subsection, we conduct an experiment to test the complexity of the scheme. In the proposed scheme, both the guidance image selection and contrast enhancement manipulations may introduce computational overhead. Here we focus on the complexity of the contrast enhancement manipulations with guidance images. Given the test image $\mathbf{I}$, the enhanced images $\mathbf{I_s}$ and $\mathbf{I_f}$ can be computed via unsharp masking and tone-curve modification. The remaining task is to derive best parameters (α and β) for image fusion. Here we employ the computation

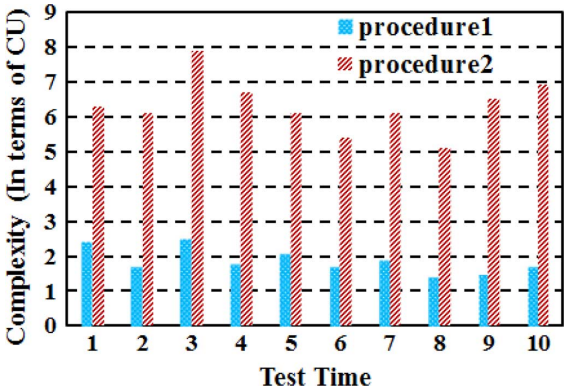


Fig. 18. Running time of procedure1 and procedure2.

uint (CU) to facilitate the quantitative measurement of the computational consumptions. Specifically, we choose the computation time of the context-sensitive unsharp masking as the CU, as this operation is frequently applied in real applications of contrast enhancement. To obtain more accurate complexity comparison, we perform the program in MATLAB for 10 times for each image, and each time the average results over all test images are recorded. The test computer has Intel 3.40 GHz Core processor and 8 GB random access memory. The computational complexity of the context-free and context-sensitivity images ($\mathbf{I_s}$ and $\mathbf{I_f}$) generation (procedure1) and the parameters (α and β) derivation (procedure2) is demonstrated in Fig. 18. One can discern that the complexity of the proposed guidance image contrast enhancement scheme is around 8 times of the unsharp masking, and the computational resources are mostly allocated to the optimal fusion parameters derivation. It is also worth noting that most of the computations (such as AR modeling and Gaussian smoothing) can be performed in parallel within the local image patch. These parallel friendly manipulations also enable its future applications in real scenarios.

V. Discussions

As one of the first attempts on relating the contrast enhancement with "cloud multimedia", our scheme has several limitations that should be improved in the future. First, the current method employs global features such as free-energy and histogram statistics for matching. In practice, users may require enhancement of local areas. In this scenario, after locating the matching position using local descriptors such as SIFT, developing more local feature extraction and statistical modeling work for contrast enhancement is necessary. Second, in the future, how are these statistical features that can describe the quality of contrast enhancement related and what is best way to combine these features may be further exploited. Finally, associating the scale space with image quality and contrast enhancement to gain more adaptability is also meaningful for the guided contrast enhancement scheme.

With the development of large scale image processing, we believe that this area has tremendous potential for future exploration. Here we discuss in detail our vision and hypothesize about possible paths to tread in the future. In the scenario of large scale contrast enhancement, the method must adapt to several stringent requirements. Firstly, the processing must be fast, as light-weighted and power-saving algorithms may greatly benefit the large scale image processing. In the similar fashion of photo album enhancement, one way to achieve this is to exploit the internal correlations within the large scale dataset to further reduce the processing complexity. Secondly, large scale image retrieval requires more accurate retrieval algorithms. Treating the low contrast images as the input for large-scale image retrieval may affect the accuracy of the retrieval performance. Therefore, with the rapid growth of multimedia big data, how to generalize the model to unify the image retrieval with enhancement and select more appropriate guidance images from cloud should be investigated. Thirdly, how to design the sophisticated crowdsourcing system to provide an efficient way in conducting the very large subjective study is very critical, as it can be used to benchmark the large-scale contrast enhancement systems and algorithms. All of these requirements pose new challenges to image quality assessment, retrieval and contrast enhancement research and open up new space for future explorations.

VI. Conclusion

We have proposed a guided image contrast enhancement framework based on the retrieved images from cloud, targeting at automatically generating a visually-pleasing enhanced image. The novelty of this paper lies in the unifying of context-sensitive and context-free contrast enhancement methods, and automatically estimating the enhancement level by matching the extracted features in visual salient region. In particular, the optimization problem is formulated as generating an image that is close to the input, context-free enhanced, as well as the context-sensitive enhanced images. With the utility of the retrieved image, the blind estimation of the contrast enhancement level is performed based on free-energy and surface quality. Experimental results demonstrate the effectiveness of the scheme in image enhancement applications.

Acknowledgment

The authors would like to thank the Associate Editor and anonymous reviewers for their valuable comments that significantly helped them in improving the presentation of the paper.

References

[1] X. Wu, "A linear programming approach for optimal contrast-tone mapping," *IEEE Trans. Image Process.*, vol. 20, no. 5, pp. 1262–1272, May 2011.
[2] G. Deng, "A generalized unsharp masking algorithm," *IEEE Trans. Image Process.*, vol. 20, no. 5, pp. 1249–1261, May 2011.
[3] Y. Li, F. Guo, R. T. Tan, and M. S. Brown, "A contrast enhancement framework with JPEG artifacts suppression," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 174–188.
[4] E. Peli, "Contrast in complex images," *J. Opt. Soc. Amer. A*, vol. 7, no. 10, pp. 2032–2040, 1990.
[5] G. Ramponi, "A cubic unsharp masking technique for contrast enhancement," *Signal Process.*, vol. 67, no. 2, pp. 211–222, 1998.
[6] S.-J. Ko and Y. H. Lee, "Center weighted median filters and their applications to image enhancement," *IEEE Trans. Circuits Syst.*, vol. 38, no. 9, pp. 984–993, Sep. 1991.
[7] R. Lukac, B. Smolka, and K. N. Plataniotis, "Sharpening vector median filters," *Signal Process.*, vol. 87, no. 9, pp. 2085–2099, 2007.
[8] M. Fischer, J. L. Paredes, and G. R. Arce, "Weighted median image sharpeners for the world wide web," *IEEE Trans. Image Process.*, vol. 11, no. 7, pp. 717–727, Jul. 2002.
[9] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 2007.

[10] Y.-T. Kim, "Contrast enhancement using brightness preserving bi-histogram equalization," *IEEE Trans. Consum. Electron.*, vol. 43, no. 1, pp. 1–8, Feb. 1997.
[11] Y. Wang, Q. Chen, and B. Zhang, "Image enhancement based on equal area dualistic sub-image histogram equalization method," *IEEE Trans. Consum. Electron.*, vol. 45, no. 1, pp. 68–75, Feb. 1999.
[12] S.-D. Chen and A. R. Ramli, "Contrast enhancement using recursive mean-separate histogram equalization for scalable brightness preservation," *IEEE Trans. Consum. Electron.*, vol. 49, no. 4, pp. 1301–1309, Nov. 2003.
[13] K. Sim, C. Tso, and Y. Tan, "Recursive sub-image histogram equalization applied to gray scale images," *Pattern Recog. Lett.*, vol. 28, no. 10, pp. 1209–1221, 2007.
[14] Q. Wang and R. K. Ward, "Fast image/video contrast enhancement based on weighted thresholded histogram equalization," *IEEE Trans. Consum. Electron.*, vol. 53, no. 2, pp. 757–764, May 2007.
[15] H. Xu, G. Zhai, X. Wu, and X. Yang, "Generalized equalization model for image enhancement," *IEEE Trans. Multimedia*, vol. 16, no. 1, pp. 68–82, Jan. 2014.
[16] T. Arici, S. Dikbas, and Y. Altunbasak, "A histogram modification framework and its application for image contrast enhancement," *IEEE Trans. Image Process.*, vol. 18, no. 9, pp. 1921–1935, Sep. 2009.
[17] K. Gu *et al.*, "Brightness preserving video contrast enhancement using s-shaped transfer function," in *Proc. Vis. Commun. Image Process.*, Nov. 2013, pp. 1–6.
[18] I. Motoyoshi, S. Nishida, L. Sharan, and E. H. Adelson, "Image statistics and the perception of surface qualities," *Nature*, vol. 447, no. 7141, pp. 206–209, 2007.
[19] J.-Y. Kim, L.-S. Kim, and S.-H. Hwang, "An advanced contrast enhancement using partially overlapped sub-block histogram equalization," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 4, pp. 475–484, Apr. 2001.
[20] M. Armbrust *et al.*, "A view of cloud computing," *Commun. ACM*, vol. 53, no. 4, pp. 50–58, 2010.
[21] H. Yue *et al.*, "Cloud-based image coding for mobile devices-toward thousands to one compression," *IEEE Trans. Multimedia*, vol. 15, no. 4, pp. 845–857, Jun. 2013.
[22] E. Hung, D. Garcia, and R. De Queiroz, "Example-based enhancement of degraded video," *Signal Process. Lett.*, vol. 21, no. 9, pp. 1140–1144, Sep. 2014.
[23] P. Milanfar, "A tour of modern image filtering: New insights and methods, both practical and theoretical," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 106–128, Jan. 2013.
[24] Z. Wang and A. C. Bovik, "Reduced-and no-reference image quality assessment," *IEEE Signal Process. Mag.*, vol. 28, no. 6, pp. 29–40, Nov. 2011.
[25] W. Lin and C.-C. J. Kuo, "Perceptual visual quality metrics: A survey," *J. Vis. Commun. Image Representation*, vol. 22, no. 4, pp. 297–312, 2011.
[26] K. Friston, "The free-energy principle: A unified brain theory?," *Nature Rev. Neurosci.*, vol. 11, no. 2, pp. 127–138, 2010.
[27] J. W. Tukey, *Exploratory Data Analysis*. London, U.K.: Pearson, 1977.
[28] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. 6th Int. Conf. Comput. Vis.*, Jan. 1998, pp. 839–846.
[29] P. Milanfar, "A tour of modern image filtering: New insights and methods, both practical and theoretical," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 106–128, Jan. 2013.
[30] K. Gu, G. Zhai, X. Yang, W. Zhang, and C. Chen, "Automatic contrast enhancement technology with saliency preservation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 9, pp. 1480–1494, Sep. 2015.
[31] H. Yue, X. Sun, J. Yang, and F. Wu, "Landmark image super-resolution by retrieving web images," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 1057–7149, Dec. 2013.
[32] Z. Shi, X. Sun, and F. Wu, "Photo album compression for cloud storage using local features," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 4, no. 1, pp. 17–28, Mar. 2014.
[33] X. Zhang, W. Lin, S. Ma, S. Wang, and W. Gao, "Rate-distortion based sparse coding for image set compression," in *Proc. IEEE Int. Conf. Visual Commun. Image Process.*, to be published.
[34] K. Gu, G. Zhai, X. Yang, W. Zhang, and M. Liu, "Subjective and objective quality assessment for images with contrast change," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2013, pp. 383–387.
[35] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.
[36] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *Proc. Eur. Conf. Comput. Vis.*, 2006, pp. 404–417.
[37] K. Gu, G. Zhai, X. Yang, and W. Zhang, "Using free energy principle for blind image quality assessment," *IEEE Trans.Multimedia*, vol. 17, no. 1, pp. 50–63, Jan. 2015.
[38] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. IEEE Int. Conf. Comput. Vis.*, Jul. 2001, vol. 2, pp. 416–423.
[39] G. Zhai, X. Wu, X. Yang, W. Lin, and W. Zhang, "A psychovisual quality metric in free-energy principle," *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 41–52, Jan. 2012.
[40] J. Wu, G. Shi, W. Lin, A. Liu, and F. Qi, "Just noticeable difference estimation for images with free-energy principle," *IEEE Trans. Multimedia*, vol. 15, no. 7, pp. 1705–1710, Nov. 2013.
[41] S. Ma *et al.*, "Entropy of primitive: From sparse representation to visual information evaluation," *IEEE Trans. Circuits Syst. Video Technol.*, to be published.
[42] S. R. Allred and D. H. Brainard, "Contrast, constancy, and measurements of perceived lightness under parametric manipulation of surface slant and surface reflectance," *J. Opt. Soc. Amer. A*, vol. 26, no. 4, pp. 949–961, 2009.
[43] J. Wang and T. N. Pappas, "Effects of contrast adjustment on visual gloss of natural textures," in *Proc. SPIE 939, Human Vis. Electron. Imaging XX*, Mar. 2015, Art. ID 93940F .
[44] J. H. Reynolds and R. Desimone, "Interacting roles of attention and visual salience in v4," *Neuron*, vol. 37, no. 5, pp. 853–863, 2003.
[45] L. Zhang and W. Lin, *Modeling Selective Visual Attention: Techniques and Applications*. Hoboken, NJ, USA: Wiley, 2013.
[46] H. Liu and I. Heynderickx, "Visual attention in objective image quality assessment: Based on eye-tracking data," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 7, pp. 971–982, Jul. 2011.
[47] X. Min, G. Zhai, Z. Gao, and K. Gu, "Visual attention data for image quality assessment databases," in *Proc. IEEE Int. Symp. Circuits Syst.*, Jun. 2014, pp. 894–897.
[48] H. Hadizadeh and I. V. Bajic, "Saliency-aware video compression," *IEEE Trans. Image Process.*, vol. 23, no. 1, pp. 19–33, Jan. 2014.
[49] X. Hou, J. Harel, and C. Koch, "Image signature: Highlighting sparse salient regions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 1, pp. 194–201, Jan. 2012.
[50] X. Zhang and S. Lyu, "Blind estimation of pixel brightness transform," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2014, pp. 4472–4476.
[51] K. Gu *et al.*, "Analysis of distortion distribution for pooling in visual quality prediction," *IEEE Trans. Broadcast.*, to be published.
[52] K. Gu, G. Zhai, W. Lin, and M. Liu, "The analysis of image contrast: From quality assessment to automatic enhancement," *IEEE Trans. Cybern.*, vol. 46, no. 1, pp. 284–297, Jan. 2016.
[53] S. Winkler, "Analysis of public image and video databases for quality assessment," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 6, pp. 616–625, Oct. 2012.
[54] *Methodology for the Subjective Assessment of the Quality of Television Pictures*, ITU BT. 500-11, ITU, Geneva, Switzerland, 2002.
[55] S. Wang, A. Rehman, Z. Wang, S. Ma, and W. Gao, "Perceptual video coding based on SSIM-inspired divisive normalization," *IEEE Trans. Image Process.*, vol. 22, no. 4, pp. 1418–1429, Apr. 2013.

Shiqi Wang (M'15) received the B.S. degree in computer science from the Harbin Institute of Technology, Harbin, China, in 2008, and the Ph.D. degree in computer application technology from Peking University, Beijing, China, in 2014.

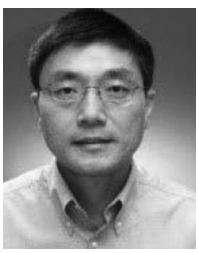
He is currently a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. From April 2011 to August 2011, he was with Microsoft Research Asia, Beijing, China, as an Intern. His current research interests include video compression and image/video quality assessment.

Ke Gu received the B.S. and Ph.D. degrees in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 2009 and 2015, respectively.

He is currently a Research Fellow with the School of Computer Engineering, Nanyang Technological University, Singapore. His research interests include quality assessment and contrast enhancement.

Siwei Ma (M'12) received the B.S. degree from Shandong Normal University, Jinan, China, in 1999, and the Ph.D. degree in computer science from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2005.

From 2005 to 2007, he was a Post-Doctoral Researcher with the University of Southern California, Los Angeles, CA, USA. He then joined the Institute of Digital Media, School of Electronic Engineering and Computer Science, Peking University, Beijing, where he is currently a Professor. He has authored or coauthored over 100 technical articles in refereed journals and proceedings in the areas of image and video coding, video processing, video streaming, and transmission.



Weisi Lin (S'91–M'92–SM'00) received the Ph.D. degree from Kings College, London University, London, U.K., in 1993.

He is currently an Associate Professor with the School of Computer Engineering, Nanyang Technological University, Singapore, and serves as a Lab Head of Visual Processing with the Institute for Infocomm Research, Singapore. He authored or coauthored over 300 scholarly articles, and holds 7 patents. His research interests include image processing, video compression, perceptual visual and audio modeling, computer vision, and multimedia communication.

Prof. Lin is a Chartered Engineer in the U.K., a Fellow of the Institution of Engineering Technology, and an Honorary Fellow of the Singapore Institute of Engineering Technologists. He served as an Associate Editor of the IEEE TRANSACTIONS ON MULTIMEDIA, the IEEE SIGNAL PROCESSING LETTERS, and the *Journal of Visual Communication and Image Representation*. He was the Lead Guest Editor for a special issue on perceptual signal processing of the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING in 2012. He is also on six IEEE Technical Committees as well as the Technical Program Committee of a number of international conferences. He was a Co-Chair of the IEEE MMTC special interest group on quality of experience. He was an Elected Distinguished Lecturer of APSIPA in 2012 and 2013.

Xianming Liu (M'14) received the B.S., M.S., and Ph.D. degree in computer science from the Harbin Institute of Technology (HIT), Harbin, China, in 2006, 2008, and 2012, respectively.

In 2007, he joined the Joint Research and Development Lab (JDL), Chinese Academy of Sciences, Beijing, China, as a Research Assistant. From 2009 to 2012, he was with National Engineering Lab for Video Technology, Peking University, Beijing, China, as a Research Assistant. In 2011, he spent half a year with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada, as a visiting student, where he was a Post-Doctoral Fellow from December 2012 to December 2013. He is currently an Associate Professor with the Department of Computer Science, HIT. He has also been a Project Researcher with the National Institute of Informatics (NII), Tokyo, Japan, since January 2014. He has authored or coauthored papers appearing in over 30 international conference and journal publications, including the IEEE TRANSACTIONS ON IMAGE PROCESSING, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, CVPR, IJCAI, and DCC.

Wen Gao (S'87–M'88–SM'05–F'09) received the Ph.D. degree in electronics engineering from the University of Tokyo, Tokyo, Japan, in 1991.

He is a Professor of Computer Science with Peking University, Beijing, China. Before joining Peking University, he was a Professor of Computer Science with Harbin Institute of Technology, Harbin, China, from 1991 to 1995, and a Professor with the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China. He has authored or coauthored five books and over 600 technical articles in refereed journals and conference proceedings in the areas of image processing, video coding and communication, pattern recognition, multimedia information retrieval, multimodal interface, and bioinformatics.

Dr. Gao served or serves on the Editorial Board for several journals, such as the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the IEEE TRANSACTIONS ON MULTIMEDIA, the IEEE TRANSACTIONS ON IMAGE PROCESSING, the IEEE TRANSACTIONS ON AUTONOMOUS MENTAL DEVELOPMENT, the *EURASIP Journal of Image Communications*, and the *Journal of Visual Communication and Image Representation*. He chaired a number of prestigious international conferences on multimedia and video signal processing, such as IEEE ICME and ACM Multimedia, and also served on the Advisory and Technical Committees of numerous professional organizations.