Image Sharpness Assessment by Sparse Representation

Leida Li, Dong Wu, Jinjian Wu, Haoliang Li, Weisi Lin, *Fellow, IEEE*, and Alex C. Kot, *Fellow, IEEE*

***Abstract*—Recent advances in sparse representation show that overcomplete dictionaries learned from natural images can capture high-level features for image analysis. Since atoms in the dictionaries are typically edge patterns and image blur is characterized by the spread of edges, an overcomplete dictionary can be used to measure the extent of blur. Motivated by this, this paper presents a no-reference sparse representation-based image sharpness index. An overcomplete dictionary is first learned using natural images. The blurred image is then represented using the dictionary in a block manner, and block energy is computed using the sparse coefficients. The sharpness score is defined as the variance-normalized energy over a set of selected high-variance blocks, which is achieved by normalizing the total block energy using the sum of block variances. The proposed method is not sensitive to training images, so a universal dictionary can be used to evaluate the sharpness of images. Experiments on six public image quality databases demonstrate the advantages of the proposed method.**

***Index Terms*—Dictionary learning, image quality assessment, image sharpness, no-reference, sparse representation.**

I. Introduction

IMAGE quality assessment (IQA) has been an important problem, because the acquisition, processing and transmission of images can introduce various distortions. Although it is possible to judge the quality of images by humans, real-word applications require objective IQA algorithms that can operate automatically without human intervention and meantime keep consistent with human perception [1]–[3].

The most straightforward way to IQA is to compare the distorted image with a distortion-free reference image. With full access to a reference image, this kind of algorithms are often referred to as full-reference (FR) metrics [4]–[6]. Instead of the whole reference image, reduced-reference (RR) metrics only use partial information of the reference [7], [8]. In the past few years, many FR and RR metrics have been proposed. However, the requirement of a reference image (full or partial) restricts their applications in real-world scenarios, where the reference image may not be available. By comparison, no-reference (NR) metrics can predict image quality without any reference, so they are highly desirable in practical applications. NR-IQA metrics can be further classified into distortion-specific and general-purpose methods. Distortion-specific metrics evaluate a specific kind of distortion, typically blocking artifacts [9], [10], ringing effect [11] and blur. General-purpose metrics estimate the quality of an image without knowing the distortion types in advance [12]–[15]. In this paper, we focus on NR image sharpness assessment.

Sharpness is inversely related to blur, which is a key factor in the perception of image quality. While the causes of blur are multifold, such as camera out-of-focus, target motion and image compression, blur is typically characterized by the spread of edges and accordingly the attenuation of high frequency components. Marziliano *et al.* [16] addressed a blur metric in the spatial domain by measuring the spread of image edges. The Sobel operator was first used to detect image edges. Then the width of each edge was computed, and image blur score was defined as the average edge width. Ferzli and Karam [17] introduced the concept of Just Noticeable Blur (JNB). Edge blocks were first selected from the image. Then local contrast and edge width of the blocks were computed and integrated into a probability summation model, producing the sharpness score. An extension of JNB was also proposed by computing the cumulative probability of blur detection (CPBD) [18]. Vu *et al.* [19] proposed the spectral and spatial sharpness (S3) index. The slope of local magnitude spectrum was used to measure the attenuation of high-frequency content, and total variation was used to account for the local contrast. A combination of these two factors was shown effective for image sharpness assessment. Vu and Chandler [20] addressed a fast image sharpness (FISH) index in the wavelet domain. The image was decomposed using the discrete wavelet transform (DWT). A weighted average of the log-energies of the DWT coefficients was computed as the sharpness score. Hassen *et al.* [21] proposed a method using Local Phase Coherence (LPC). It was based on the fact that image blur disrupted the LPC structure, and the strength of LPC can be used to measure image sharpness. Sang *et al.* [22] proposed to use the singular value curve (SVC) for sharpness evaluation. It was based on the observation that blur can cause attenuation of singular values in an image, and the shape of SVC can be used to measure the extent of blur. Bahrami and Kot [23] proposed a method based on the Maximum Local Variation (MLV). The MLV of each pixel was first computed within its 8-pixel

Manuscript received August 03, 2014; revised June 07, 2015 and January 27, 2016; accepted March 20, 2016. Date of publication March 22, 2016; date of current version May 13, 2016. This work was supported by the Fundamental Research Funds for the Central Universities under Grant 2015XKMS032. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Maria Martini.

L. Li and D. Wu are with the School of Information and Electrical Engineering, China University of Mining and Technology, Xuzhou 221116, China (e-mail: reader1104@hotmail.com; starrysky_wd@163.com).

J. Wu is with the School of Electronic Engineering, Xidian University, Xi'an 710071, China (e-mail: jinjian.wu@mail.xidian.edu.cn).

H. Li and A. C. Kot are with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798 (e-mail: hli016@e.ntu.edu.sg; eackot@ntu.edu.sg).

W. Lin is with the School of Computer Engineering, Nanyang Technological University, Singapore 639798 (e-mail: wslin@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/TMM.2016.2545398

neighborhood. Then the MLVs were weighted based on their rankings. Finally, the standard deviation of the weighted MLV distribution was computed as the sharpness score.

The existing image sharpness metrics are mainly based on low-level features, which are usually hand-crafted ones extracted directly from digital representation of an image. Low-level features have little or nothing to do with human perception. High-level features are more preferable in our understanding of visual scenes [24], [25]. From this perspective, low-level features may not be sufficient for IQA. Compared with low-level features, high-level features have more semantic meanings and are believed to be more related with human perception [26]. We believe high-level features can facilitate more effective IQA. Recently, attempts have been made to use sparse representation for IQA. Chang *et al.* [27] addressed the sparse feature fidelity metric for FR-IQA. A feature detector was first trained using Independent Component Analysis (ICA) on natural images. Then the reference and distorted images were represented by sparse features based on the detector. Finally, quality assessment was achieved by comparing the sparse features. Guha *et al.* [28] proposed a sparse representation based FR quality index (SPARQ) using overcomplete dictionaries. A dictionary was first trained using the reference image. Then the dictionary was used to decompose both the reference and distorted images. Finally, the quality score was obtained by comparing the similarity of the sparse coefficients. These metrics are FR ones. To the best of our knowledge, very few work has been done to use sparse representation for NR-IQA. Specifically, He *et al.* [29] proposed an NR image quality metric based on the sparse representation of natural scene statistics (NSS). NSS features were extracted in the wavelet domain for sparse representation. The quality score was generated by weighting difference mean opinion scores (DMOS) using the sparse coefficients. Ding *et al.* [30] addressed a NR quality metric based on Topographic Independent Component Analysis (TICA). TICA was first adopted to extract sparse features. Then the distribution of sparse features was employed to generate the quality score. These two sparse representation based NR quality metrics [29], [30] are general-purpose models, and they do not perform very well on image sharpness. In this paper, we train an overcomplete dictionary and employ it to evaluate image sharpness. In contrast to He *et al.* [29], we do not need to extract hand-crafted features in advance, so image sharpness can be evaluated using the dictionary-based sparse representation directly.

The human visual system (HVS) has the capacity to extract high-level features when understanding an image. Overcomplete dictionaries learned from natural images consist of atoms that are similar to the simple cells in the primary visual cortex [31], [32]. This indicates that high-level features can be extracted using the learned dictionaries. Furthermore, atoms in the dictionaries are typically edge patterns [33]. Since image blur is mainly characterized by the spread of edges, a learned overcomplete dictionary is expected to be able to measure image blur effectively. Motivated by these facts, this paper presents a simple and effective NR SPArse Representation based Image SHarpness (SPARISH) metric. An overcomplete dictionary is first learned using natural images. Then the image blocks are represented using sparse coefficients with respect to the dictionary, based on which the energies of the blocks are computed. Image sharpness score is defined as the variance-normalized energy over a set of high-variance blocks, which is achieved by normalizing the total block energy using the sum of corresponding block variances. It should be noted that while SPARQ [28] and the proposed SPARISH both employ overcomplete dictionary based sparse representation to achieve quality assessment, they are quite different in nature. Firstly and most importantly, SPARQ is a FR image quality index, and SPARISH is an NR image sharpness index. Secondly, SPARQ needs to train different dictionaries for different reference images, and SPARISH uses a universal dictionary. Experiments on public image quality databases demonstrate the advantages of the proposed method.

II. Sparse Representation and Dictionary Learning

A. Sparse Representation

Sparse representation is to approximate a signal as a linear combination of a small number of atoms in a predefined dictionary. In sparse representation, overcomplete dictionaries are commonly used, which means that the number of atoms in the dictionary is greater than the vector space spanned by the input signal [28]. Given an overcomplete dictionary $\mathbf{D} \in \mathbf{R}^{n \times K}$ $(n < K)$ that contains K n-dimensional atoms $\{\mathbf{d}_i\}_{i=1}^{K}$, a signal $\mathbf{y} \in \mathbf{R}^n$ can be represented as

$$\mathbf{y} = \mathbf{D}\mathbf{x} = \sum_{i=1}^{K} x_i \mathbf{d}_i, \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2 \leq \epsilon \tag{1}$$

where $\|\cdot\|_2$ is the ℓ_2 norm, $\mathbf{x} \in \mathbf{R}^K$ is the representation vector, and ϵ is the error tolerance. Since $n < K$, there exists an infinite number of solutions to the above problem. Therefore, some constraints should be imposed on the solutions. For signal representation, we always hope to obtain a representation vector with the fewest number of nonzero elements. As a result, the sparse representation of a signal can be achieved by

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0, \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2 \leq \epsilon \tag{2}$$

where $\|\cdot\|_0$ is the ℓ_0 norm, which denotes the number of nonzero elements. Given an overcomplete dictionary, finding the representation vector $\mathbf{x}$, i.e., the sparse coefficients, can be achieved using pursuit algorithm [34]. In this paper, we employ the Orthogonal Matching Pursuit (OMP) algorithm [35] to obtain the sparse coefficients.

B. Dictionary Learning

Several effective dictionary learning algorithms have been proposed [33], [34]. Generally, the dictionary is learned using a large number of image patches via an iterative update mechanism, with the objective to obtain the best possible representation for each input signal. Given image patches $\{\mathbf{y}_i\}_{i=1}^{N}$, which have been rearranged into 1-D column vectors, the dictionary can be formulated as [36]

$$\begin{aligned} &\mathbf{D} = \arg\min_{\mathbf{D},\mathbf{x}} \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_1 \\ &\text{s.t.} \quad \|\mathbf{D}_i\|_2^2 \leq 1, i = 1, 2, \ldots, K \end{aligned} \tag{3}$$

Fig. 1. Image set used to learn the dictionary.

Fig. 2. Dictionary learned using the images shown in Fig. 1.

where $\|\cdot\|_1$ is the ℓ_1 norm, and the parameter λ is used to balance the sparsity and fidelity of the approximation to $\mathbf{y}$. The iterative update of $\mathbf{D}$ and $\mathbf{x}$ operates as follows [36]:

1) initialize $\mathbf{D}$ as a Gaussian random matrix, and normalize each column to a unit vector;
2) keep $\mathbf{D}$ fixed, update $\mathbf{x}$ using

$$\mathbf{x} = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_1 \tag{4}$$

3) keep $\mathbf{x}$ fixed, update $\mathbf{D}$ using

$$\begin{aligned} &\mathbf{D} = \arg\min_{\mathbf{D}} \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2^2, \\ &\text{s.t.} \quad \|\mathbf{d}_i\|_2^2 \le 1, i = 1, 2, \ldots, K \end{aligned} \tag{5}$$

4) repeat steps (2) and (3) until convergency is achieved.

In this paper, we employ Lee's method [33] to learn the dictionary. A large number of patches are first extracted randomly from a set of natural images. Then the mean value is subtracted to ensure that image structures are learned. The patches are then transformed into column vectors, based on which the dictionary is learned. Fig. 1 shows a set of natural images, and Fig. 2 shows the corresponding learned dictionary. The size of image patch is 8×8, and the size of the dictionary is 64×256. It should be noted that the dictionary shown in Fig. 2 has been visualized for better display. Specifically, each 64×1 atom is reshaped to an 8×8 block, and the 256 atoms are rearranged into 16 rows by 16 columns.

It is observed from Fig. 2 that the atoms in the dictionary are typically edge patterns. These patterns have been shown similar to the receptive fields of neurons in the primary visual cortex [31], [32], and can be used to capture high-level features in images [33]. The proposed method employs this kind of overcomplete dictionary to achieve image sharpness assessment.

III. Sparse Representation-Based Image Sharpness Assessment

Since the atoms in a learned dictionary are typically edge patterns and image blur is characterized by the spread of edges, it is expected that the extent of blur can be measured using the sparse coefficients with respect to the dictionary. The proposed method is based on this observation.

Fig. 3 shows the flowchart of the proposed method. For a blurred image, it is first converted into gray scale and divided into blocks. Then the gradient of the block is computed and the sparse representation of the block gradient is obtained based on the dictionary using the OMP algorithm. The energy of a block is computed using the sparse coefficients. Then the high-variance blocks are selected and their total energy is normalized by the sum of the corresponding block variances, producing the final sharpness score. In order to understand the method better, we will first illustrate the relationship between image blur and the sparse coefficients.

Here, we consider a simplified scenario that a high-quality sharp image $\mathbf{I}_0$ is distorted by pure blur, and the blurred image is denoted by $\mathbf{I}$. Under this scenario, the extent of blur in image $\mathbf{I}$ can be approximately represented by

$$\mathbf{B} = \mathbf{I} - \mathbf{I}_0 \tag{6}$$

where $\mathbf{B}$ is the pure blur signal. By subtracting the mean value $\bar{\mathbf{I}}_0$ from the sharp image (as in the dictionary learning), the equation can be rewritten as

$$\mathbf{B} + \bar{\mathbf{I}}_0 = \mathbf{I} - (\mathbf{I}_0 - \bar{\mathbf{I}}_0). \tag{7}$$

Let $\mathbf{B}' = \mathbf{B} + \bar{\mathbf{I}}_0$, then $\mathbf{B}'$ also denotes the extent of blur in the image. In this paper, $\mathbf{I}$ is represented using the overcomplete dictionary $\mathbf{D}$, namely

$$\mathbf{I} = \mathbf{D}\mathbf{x}, \quad \text{s.t.} \quad \|\mathbf{I} - \mathbf{D}\mathbf{x}\|_2 \le \epsilon \tag{8}$$

where $\mathbf{x}$ is the representation vector. By substituting $\mathbf{B}' = \mathbf{B} + \bar{\mathbf{I}}_0$ and $\mathbf{I} = \mathbf{D}\mathbf{x}$, and then take the mathematical expectations of both sides of (7), it can be rewritten as

$$\begin{aligned} E[\mathbf{B}'] &= E[\mathbf{I}] - E[\mathbf{I}_0 - \bar{\mathbf{I}}_0] \\ &= E[\mathbf{D}\mathbf{x}] - E[\mathbf{I}_0 - \bar{\mathbf{I}}_0]. \end{aligned} \tag{9}$$

It is obvious that $E[\mathbf{I}_0 - \bar{\mathbf{I}}_0] = 0$. So the above equation can be simplified as

$$E[\mathbf{B}'] = E[\mathbf{D}\mathbf{x}]. \tag{10}$$

The above equation indicates that for an image, the extent of blur is related to the sparse coefficients with respect to an overcomplete dictionary.

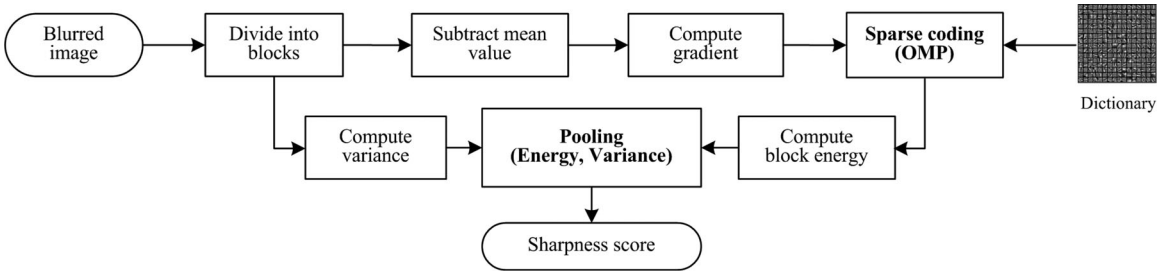


Fig. 3. Flowchart of the proposed image sharpness metric.

It has been shown that image blur is mainly characterized by the spread of edges, and accordingly, the attenuation of energy in high-frequency contents [20]. In other words, a sharp image has higher energy than its blurred counterparts. Inspired by this, we propose to evaluate the extent of blur in an image based on a new energy measure, which can be computed using the sparse coefficients. Specifically, we define the image sharpness score as the variance normalized energy over a set of selected high-variance blocks.

For a blurred image, it is first converted into gray scale. Then the image is divided into blocks. The following operator is employed to compute the gradient of the image blocks:

$$\mathrm{G}_x = [-1\ 0\ 1],\ \mathrm{G}_y = [-1\ 0\ 1]^{\mathrm{T}} \tag{11}$$

where T denotes the transpose. Computing the gradient has the functionality to emphasize edges in images, which can facilitate the subsequent sharpness assessment.

For an image block, the representation vector $\mathbf{x}$ is obtained using the OMP algorithm based on the dictionary $\mathbf{D}$. Then the energy of the block can be computed as the squared ℓ_2 norm of the sparse coefficients. Suppose we have an image block, the gradient is first computed, which is denoted by $\mathbf{b}$ here. Then it is represented using the dictionary $\mathbf{D}$ as

$$\mathbf{b} = \mathbf{D}\mathbf{x} = \sum_{i=1}^{K} x_i \mathbf{d}_i. \tag{12}$$

By taking the squared ℓ_2 norm to both sides of the equation, we have

$$\|\mathbf{b}\|_2^2 = \sum_{i=1}^{K} \langle x_i \mathbf{d}_i, x_i \mathbf{d}_i \rangle = \sum_{i=1}^{K} |x_i|^2 \langle \mathbf{d}_i, \mathbf{d}_i \rangle \tag{13}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product. Since the atoms have been orthonormalized in dictionary learning [33], i.e., $\langle \mathbf{d}_i, \mathbf{d}_i \rangle = 1, i = 1, 2, \ldots, K$, so the equation can be further simplified as

$$\|\mathbf{b}\|_2^2 = \sum_{i=1}^{K} |x_i|^2. \tag{14}$$

This indicates that the energy of the block gradients can be fully represented using the sparse coefficients. If a blurred image is divided into M equal-size blocks, then the average energy of block gradients can be used to measure the extent of blur in this image, which is computed as

$$P = \frac{1}{M} \sum_{i=1}^{M} \|\mathbf{b}_i\|_2^2. \tag{15}$$

(a) Original image (b) Blurred image

Fig. 4. Average energies of block gradients for images with different blur. (a) $P = 3462.70$. (b) $P = 144.09$.

Fig. 4 shows the original image "Caps" and a blurred version, together with their average energies of block gradients calculated using (15). It is observed that due to the blur, the average energy of blurred image reduces accordingly.

In Fig. 5, two images are iteratively blurred using Gaussian low-pass filters with different scales. Then the average energy of block gradients is computed as a function of the blur strength. It is observed that the average energy drops monotonically with respect to the strength of blur. However, when we compare the curves of the two images, we find that even with the same extent of blur, the average energies are different. This is not hard to understand, because the images have different contents. Therefore, the average block energy can only be used to measure the extent of blur for images with the same content. To obtain a sharpness metric that can evaluate image sharpness across different images, the influence of image content should be removed.

With the knowledge that images with complex textures generally have large variations, we propose a simple method to remove the influence of image content. The idea is to normalize the total energy using the total block variances

$$\mathrm{Q} = \frac{\sum_{i=1}^{M} \|\mathbf{b}_i\|_2^2}{\sum_{i=1}^{M} \sigma_i^2} \tag{16}$$

where Q is the sharpness score, and σ_i^2 denotes the variance of the ith block. It should be noted that the variances are computed using the blocks in the blurred image so that they capture the characteristics of image contents.

When employing (16) to compute the sharpness score, another characteristic of the HVS should be considered.

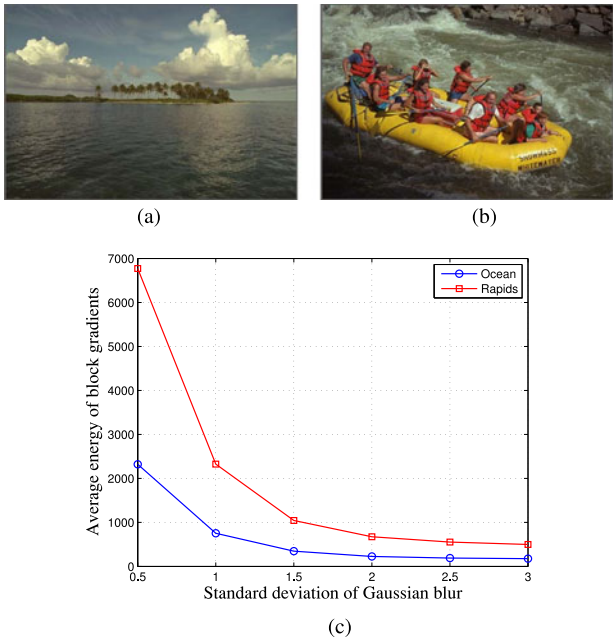


Fig. 5. Plots of the average energies of block gradients with the increase of Gaussian blur on two real images. (a) Ocean. (b) Rapids. (c) Average block energy versus blur strength.

Fig. 6. Image with sharp foreground and blurred background. Humans tend to classify this image as a sharp one.

Humans tend to judge the sharpness of an image according to the sharpest regions [19], [20]. One such example is shown in Fig. 6. Although the image has very small sharp region and very blurred background, we tend to classify it as a sharp image. Furthermore, blur mainly affects the edge regions, and has limited effect on the smooth regions. In [19] and [20], the sharpness score was computed by taking the average sharpness of 1% largest values in the sharpness map. In [37], blur was estimated with sharper edges. In this paper, we only consider sharp blocks in computing the final sharpness score. Here, the sharpness of a block is measured by its variance. In implementation, the blocks are first sorted in descending order according to their variances. Then the top p percent of the high-variance blocks are used to produce the final sharpness score as follows:

$$\mathrm{Q}_{\mathrm{SPARISH}} = \frac{\sum_{i=1}^{Z} \|\mathbf{b}_i\|_2^2}{\sum_{i=1}^{Z} \sigma_i^2} \tag{17}$$

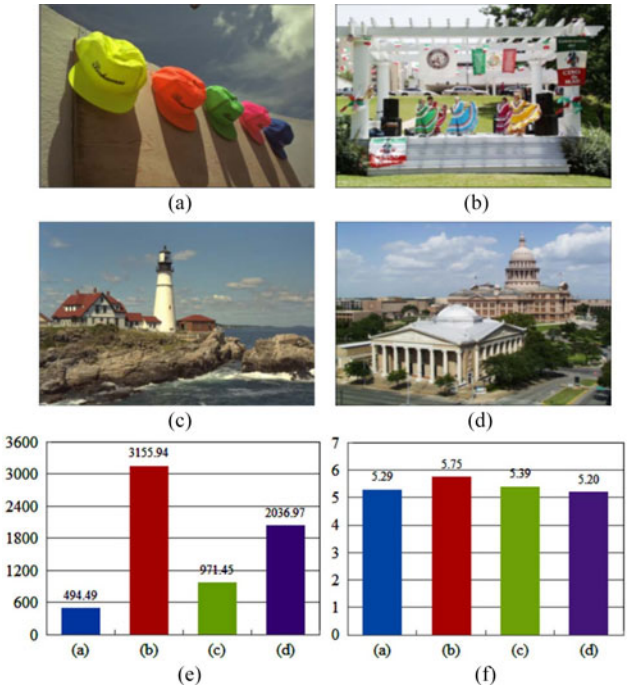


Fig. 7. Images with similar visual quality together with the average block energy before and after normalization. (a) DMOS=57.78. (b) DMOS=58.03. (c) DMOS=59.57. (d) DMOS=58.91. (e) Energy before normalization. (f) Energy after normalization.

where $Z = \lfloor p\% \times M \rfloor$ denotes the number of top $p\%$ high-variance blocks, $\lfloor \cdot \rfloor$ is the *floor* operator.

The variance based normalization is effective for generating sharpness scores insensitive to image content and consistent with human perception. Fig. 7 shows four images with very similar visual quality, together with the average block energies before and after the variance based normalization. In the figure, the subjective scores are measured by DMOS, and high value indicates bad quality.

It is observed from the figure that the four images have quite similar extent of blur. Fig. 7(e) shows the average energy of the top $p\%$ high-energy blocks, and Fig. 7(f) shows the corresponding variance-normalized energy, namely the proposed sharpness score. It is observed that the average block energy fails to measure the sharpness across images, because the average energies shown in Fig. 7(e) differ significantly. However, when the energies are normalized by the corresponding block variances, the produced scores are very similar, which are consistent with the subjective scores. Therefore, the proposed block-variance based normalization is effective in evaluating the sharpness across images.

IV. Experimental Results and Discussions

A. Experimental Settings

Six image quality databases are employed to evaluate the performance of the proposed method, including LIVE [38],[1]

[1]H. R. Sheikh, Z.Wang, L. Cormack, and A. C. Bovik, "LIVE image quality assessment database release 2," [Online]. Available: http://live.ece.utexas.edu/research/quality

CSIQ [40], TID2008 [41], TID2013 [42], LIVE multiply distorted (LIVE MD) image database [47],[2] and Camera Image Database (CID2013) [49], [50]. In LIVE, CSIQ, TID2008 and TID2013, there is a subset of images degraded by pure blur. So they are used to test the performance on pure blur distortion. By contrast, LIVE MD and CID2013 contain images distorted by multiple distortions, such as blur and JPEG, blur and noise. So they are used to test the performance on multiple distortions, which is more likely to occur in real-world scenarios. The numbers of blurred images in the six databases are 145, 150, 100 125, 450 and 474, respectively. So a total of 1444 images are tested in our experiments. In LIVE, CSIQ and LIVE MD, the subjective qualities are measured using DMOS. In TID2008, TID2013 and CID2013, the subjective qualities are measured using mean opinion score (MOS).

Three criteria are adopted to evaluate the performance, including Pearson linear correlation coefficients (PLCC), Spearman rank order correlation coefficients (SRCC) and root mean square error (RMSE). SRCC is used to measure the prediction monotonicity, and it can be computed using the subjective scores and predicted scores directly. PLCC and RMSE are used to measure the prediction accuracy, and they are computed after performing a nonlinear fitting between the subjective and predicted scores. In this paper, the nonlinear fitting is performed using the following four-parameter logistic function:

$$f(x) = \frac{\tau_1 - \tau_2}{1 + e^{(x-\tau_3)/\tau_4}} + \tau_2 \tag{18}$$

where $\tau_1, \tau_2, \tau_3, \tau_4$ are the parameters to be fitted.

The overcomplete dictionary used in this paper is learned using Lee's algorithm [33]. In dictionary learning, 8×8 blocks are used, and the dimension of the dictionary is 64×256. The natural images used to learn the dictionary are shown in Fig. 1, and the learned dictionary is shown in Fig. 2. This dictionary is used in all subsequent experiments. A MATLAB release of SPARISH is available at MATLAB Central.[3]

B. Parameter Tuning

In the proposed method, two parameters need to be determined, namely the sparsity degree and the percentage of high-variance blocks used to generate the sharpness score. Here, sparsity degree denotes the number of atoms in the over-complete dictionary that are used for sparse representation, i.e., number of nonzero elements in the representation vector. In sparse representation, if higher sparsity degree is used, a signal can be represented better, which indicates that the representation error, i.e., $\|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2$ in (1), is smaller. Meantime, higher sparsity will incur higher computational cost, because the pursuit algorithm needs to find more optimal atoms to represent a signal. In this paper, these two parameters are tuned using 70 blurred images randomly taken from the LIVE database. The relations between the PLCC, SRCC values and the sparsity degree L, percentage of high-variance blocks are shown in Fig. 8.

[2][Online]. Available: http://live.ece.utexas.edu/research/quality/live_multidistortedimage.html

[3][Online]. Available: https://cn.mathworks.com/matlabcentral/fileexchange/55106-sparish?s_tid=srchtitle.

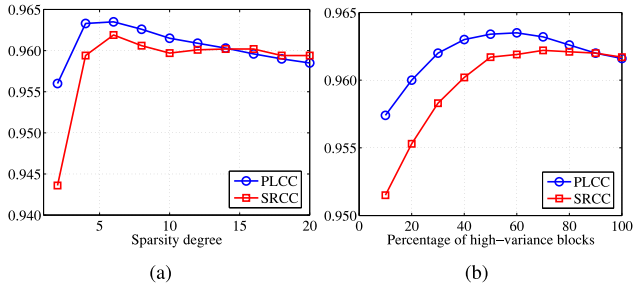


Fig. 8. Plots of PLCC and SRCC values versus sparsity degree and percentage of high-variance blocks based on 70 blurred images in LIVE database. (a) With 60% high-variance blocks used. (b) With sparsity 6.

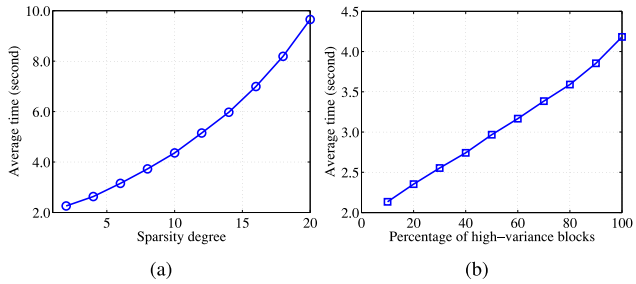


Fig. 9. Plots of average time versus sparsity degree and percentage of high-variance blocks based on images in CSIQ database. (a) With 60% high-variance blocks used. (b) With sparsity 6.

From Fig. 8(a), we know that in the beginning both PLCC and SRCC increase with the sparsity degree. However, when the sparsity exceeds six, they begin to decrease slightly. This may be due to the fact that when the sparsity is low, only the most important atoms are used to represent the image, which is accurate. With the increase of sparsity, the added atoms are less important, so they are limited in characterize the edge changes caused by blur. In this paper, we set the sparsity $L = 6$, which produces the best results for both PLCC and SRCC. Another characteristic observed from Fig. 8(b) is that the PLCC and SRCC values reach the maximum when 60% of the blocks are adopted to generate the quality score.

Fig. 9 further shows the relations between the average computational time (seconds per image) and sparsity degree, percentage of high-variance blocks based on the CSIQ database. This experiment is conducted on a PC with Intel Dual-Core i5 CPU at 3.30 GHz, 4 GB RAM, Windows 10 64-bit, and MATLAB R2013b. It is clear that the computational time increases with the increase of both sparsity and percentage of high-variance blocks. For the sake of efficiency, lower sparsity and percentage is preferred. In this paper, we set the sparsity to $L = 6$ and the top 60% high-variance blocks are employed to generate the sharpness score, which can achieve the best performance and at the same time maintain relatively low computational cost.

C. Performance on Purely Blurred Images

In this section, the performance of the proposed method on pure blur distortion is tested on LIVE, CSIQ, TID2008 and TID2013 databases. For comparison, the results of nine state-of-the-art NR sharpness metrics are included, including Marzil-

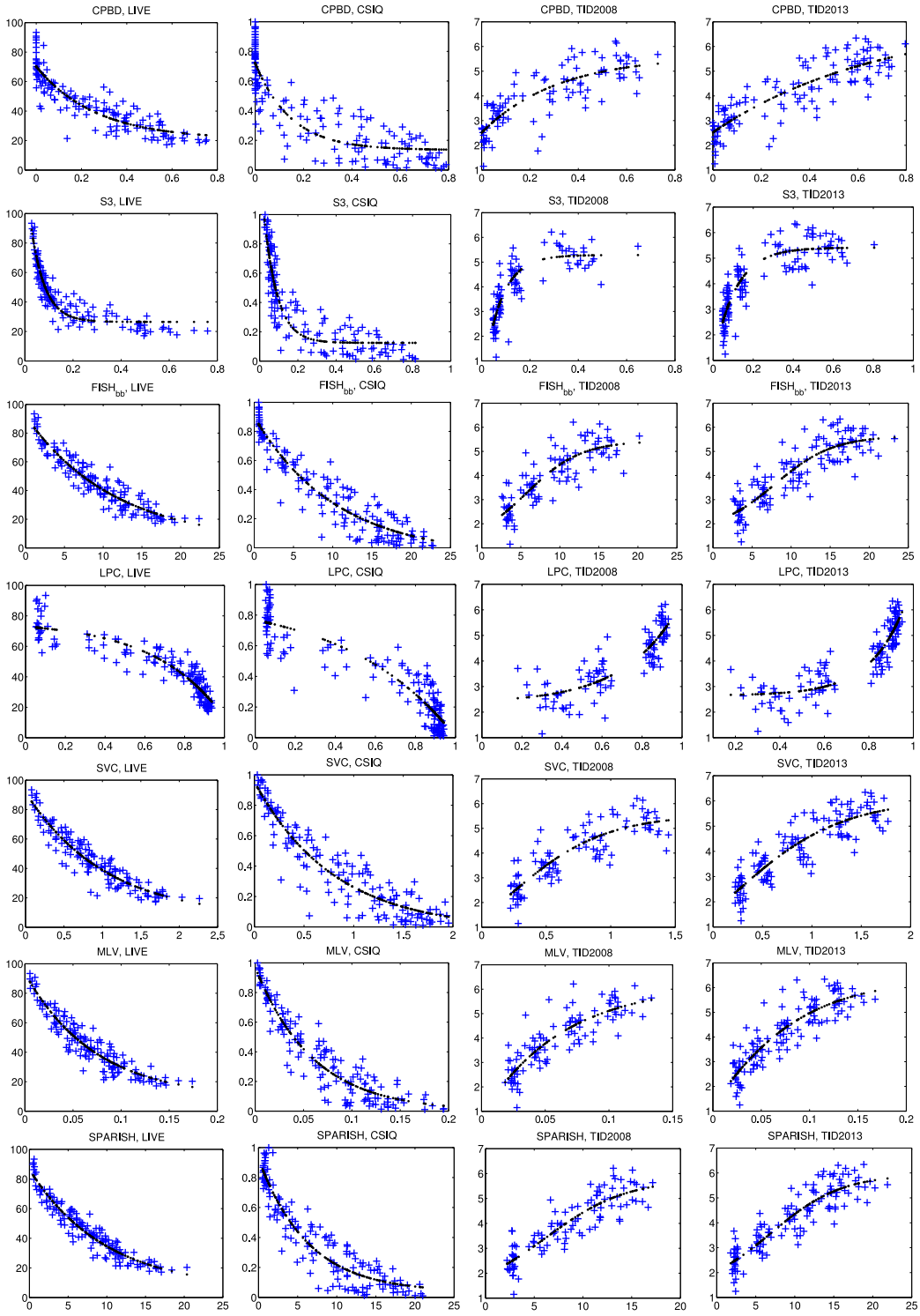


Fig. 10. Scatter plots of the subjective scores versus the predicted sharpness scores generated by SPARISH and state-of-the-art sharpness metrics. The x-axis denotes the metric score and the y-axis denotes the subjective score (DMOS for LIVE and CSIQ, MOS for TID2008 and TID2013).

iano's method [16], JNB [17], CPBD [18], S3 [19], FISH [20], block-based FISH (FISH$_{bb}$) [20], LPC [21], SVC [22] and MLV [23]. The source codes of these state-of-the-art metrics are available at: Marziliano's method, JNB and CPBD [43], S3, FISH and FISH$_{bb}$ [44], LPC [45], MLV [46]. The codes of SVC [22] are provided by the authors.

Fig. 10 shows the scatter plots of the subjective scores versus the predicted scores for the seven most recent sharpness metrics. For the nonlinear fitting, a good metric is expected to produce scatter points that are close to the fitted curve. It can be seen from Fig. 10 that most of the metrics produce good fittings. It is also observed that S3, FISH$_{bb}$, SVC, MLV and SPARISH produce better results than the other metrics. The proposed SPARISH produces the best fitting results in LIVE, TID2008 and TID2013. CPBD, FISH$_{bb}$ and LPC exhibit saturation effect [21], which indicates that for images with very high (or very low) qualities,

TABLE I
PERFORMANCE COMPARISON OF SPARISH AND STATE-OF-THE-ART SHARPNESS METRICS OVER FOUR DATABASES

Database	Criterion	Marz. [16]	JNB [17]	CPBD [18]	S3 [19]	FISH [20]	FISH$_{bb}$ [20]	LPC [21]	SVC [22]	MLV [23]	SPARISH
LIVE (145 images)	PLCC	0.7980	0.8161	0.8955	0.9434	0.9043	0.9440	0.9181	0.9416	0.9429	**0.9595**
	SRCC	0.7977	0.7872	0.9183	0.9435	0.8808	0.9381	0.9389	0.9343	0.9312	**0.9593**
	RMSE	11.1317	10.6754	8.2217	6.1276	7.8854	6.0954	7.3224	6.2201	6.1514	**5.2026**
CSIQ (150 images)	PLCC	0.7936	0.8061	0.8818	0.9107	0.9231	0.9434	0.9158	0.9319	**0.9488**	0.9380
	SRCC	0.7661	0.7624	0.8852	0.9060	0.8941	0.9177	0.9072	0.9055	**0.9247**	0.9141
	RMSE	0.1744	0.1696	0.1351	0.1184	0.1102	0.0950	0.1151	0.1039	**0.0905**	0.0993
TID2008 (100 images)	PLCC	0.6922	0.6931	0.8235	0.8541	0.8079	0.8519	0.8573	0.8556	0.8583	**0.8891**
	SRCC	0.6960	0.6667	0.8412	0.8416	0.7828	0.8378	0.8561	0.8362	0.8552	**0.8869**
	RMSE	0.8367	0.8459	0.6657	0.6104	0.6915	0.6145	0.6040	0.6075	0.6022	**0.5372**
TID2013 (125 images)	PLCC	0.7661	0.7113	0.8552	0.8813	0.8327	0.8756	0.8917	0.8762	0.8818	**0.9004**
	SRCC	0.7620	0.6902	0.8512	0.8608	0.8024	0.8584	0.8889	0.8589	0.8789	**0.8927**
	RMSE	0.8020	0.8771	0.6467	0.5896	0.6910	0.6027	0.5647	0.6013	0.5885	**0.5430**
Direct average	PLCC	0.7625	0.7567	0.8640	0.8974	0.8670	0.9037	0.8957	0.9013	0.9080	**0.9218**
	SRCC	0.7555	0.7266	0.8740	0.8880	0.8400	0.8880	0.8978	0.8837	0.8975	**0.9133**
Weighted average	PLCC	0.7687	0.7644	0.8680	0.9019	0.8740	0.9097	0.8994	0.9065	0.9136	**0.9256**
	SRCC	0.7604	0.7336	0.8778	0.8932	0.8469	0.8938	0.9018	0.8890	0.9021	**0.9163**

the metrics produce very similar sharpness scores. Specifically, CPBD and FISH$_{bb}$ show the saturation effect for very low quality images. LPC shows the saturation effect for both very high and very low quality images (best viewed in LIVE and CSIQ). The proposed method also has the saturation effect for low quality images. This indicates that for heavily blurred images, the proposed method tends to produce very similar sharpness scores.

Based on the nonlinear fittings, the PLCC, SRCC and RMSE values are computed. Table I summarizes the simulation results of the proposed method (SPARISH) and nine state-of-the-art NR image sharpness metrics, where the best result is marked in boldface. To evaluate the overall performance across different databases, we also compute the average results of the four databases, including direct average and weighted average. Direct average is obtained by taking the mean value across the four databases. Weighted average is computed by assigning different weights to different databases, and the weight is defined as $\frac{\#\text{ of blurred images in a database}}{\#\text{ of blurred images in all databases}}$. In this way, bigger database will be assigned bigger weight, and thus has more effect on the weighted average value. Since the RMSE values in different databases have different scales, their average values are not included in the table.

It is known from Table I that the proposed method produces the best results in three of the four databases, namely LIVE, TID2008 and TID2013. In LIVE, SPARISH achieves the highest PLCC and SRCC as well as the smallest RMSE. S3, FISH$_{bb}$, SVC and MLV also produce promising results, and they outperform the remaining metrics. In CSIQ, the performance of the proposed method ranks the third, and it is only slightly worse than MLV and FISH$_{bb}$. In TID2008, the results of SPARISH are much better than the other metrics. In TID2013, SPARISH achieves the best performance. LPC performs slightly worse than SPARISH. S3, FISH$_{bb}$, SVC and MLV also produce promising results and they perform much better than the remaining metrics. From the average results, we know that the proposed method achieves the best overall performance with respect to both prediction accuracy and monotonicity. The overall performances of S3, FISH$_{bb}$, LPC, SVC and MLV are quite similar, which are better than Marziliano's method, JNB, CPBD and FISH.

Based on the experimental results, we further employ *F*-test [51] to estimate the statistical significance of each metric's performance relative to SPARISH. Specifically, *F*-test is performed on the prediction errors between the predicted scores and subjective scores. It can be used to determine if a metric has significantly larger prediction errors than another metric [51]. If the prediction errors of metric A and SPARISH are denoted by σ_A^2 and $\sigma_{\mathrm{SPARISH}}^2$, F score is computed by

$$F = \sigma_A^2 / \sigma_{\mathrm{SPARISH}}^2. \tag{19}$$

If $F > F_{critical}$ (or $F < 1/F_{\mathrm{critical}}$), metric A has significantly larger (or smaller) prediction errors than SPARISH, which indicates that in terms of statistical significance, metric A performs worse (or better) than SPARISH. Otherwise, they are competitive. Here, F_{critical} is a threshold related to the number of prediction errors and a confidence level [52], which can be computed using MATLAB function *finv*. In this paper, the number of prediction errors equals the number of blurred images in each database, and the confidence level is 99%. Accordingly, the F_{critical} values for the LIVE, CSIQ, TID2008 and TID2013 databases are 1.4744, 1.4647, 1.5977 and 1.5197, respectively.

Fig. 11 shows the F statistic results of each metric's prediction error relative to SPARISH, and Table II summarizes the statistical performances. In Table II, "1" indicates that the performance of SPARISH is significantly better than that of the compared metric ($F > F_{\mathrm{critical}}$), while "0" indicates that their performances are competitive ($1/F_{\mathrm{critical}} < F < F_{\mathrm{critical}}$). It is observed from Fig. 11 that the prediction errors of SPARISH are smaller than all the state-of-the-art metrics in LIVE, TID2008 and TID2013 databases. In CSIQ, only FISH$_{bb}$ and MLV produce slightly smaller prediction errors than SPARISH. Table II shows that in terms of statistical significance, the proposed SPARISH outperforms Marziliano's method, JNB, CPBD, FISH

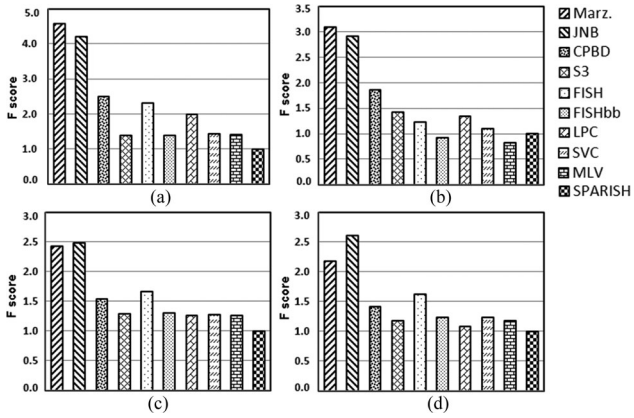


Fig. 11. *F* statistics of the compared metrics against SPARISH. (a) LIVE, (b) CSIQ, (c) TID2008, and (d) TID 2013.

TABLE II
SUMMARY OF STATISTICAL PERFORMANCES BETWEEN SPARISH AND STATE-OF-THE-ART SHARPNESS METRICS IN FOUR DATABASES

Metric	LIVE	CSIQ	TID2008	TID2013
Marz. [16]	1	1	1	1
JNB [17]	1	1	1	1
CPBD [18]	1	1	0	0
S3 [19]	0	0	0	0
FISH [20]	1	0	1	1
$FISH_{bb}$ [20]	0	0	0	0
LPC [21]	1	0	0	0
SVC [22]	0	0	0	0
MLV [23]	0	0	0	0

The value 1 (0) indicates that SPARISH performs better (competitive) than the metric in comparison.

and LPC in at least one database, and meantime S3, $FISH_{bb}$, SVC and MLV are competitive to the proposed SPARISH.

D. Impact of Training Images

The proposed method uses an overcomplete dictionary that is learned on a set of natural images. Therefore, it is important to know how the method performs if different training images are used. To this end, we employ five overcomplete dictionaries (64×256) to test the performance of the proposed method. The characteristics of the five dictionaries are as follows

1) Dictionary I (Dict. I) is trained using 10 000 patches that are randomly selected from ten natural images shown in Fig. 1. These images are taken from the undistorted reference images in the LIVE database.
2) Dictionary II (Dict. II) is trained using 10 000 patches that are randomly selected from ten natural images. These images are taken from the reference images in the TID2008 database. These images are completely different from the images in Fig. 1.
3) Dictionary III (Dict. III) is trained using 50 000 patches that are randomly selected from 59 natural images. These images are the reference images in LIVE (29 images) and CSIQ (30 images) databases. For limited space, the images are not shown here.
4) Dictionary IV (Dict. IV) is trained using 10 000 patches that are randomly selected from four natural images, which are taken from the reference images in TID2008. These images do not overlap with those in Dict. II.
5) Dictionary V (Dict. V) is trained using 10 000 patches that are randomly selected from seven natural images, which are taken from the reference images in TID2008. Furthermore, these images do not overlap with those in Dict. II and Dict. IV.

TABLE III
PERFORMANCE COMPARISON OF SPARISH USING VARIOUS DICTIONARIES ON LIVE, CSIQ, TID2008, AND TID2013 DATABASES

Dictionary	LIVE		CSIQ	
	PLCC	SRCC	PLCC	SRCC
Dict. I	0.9595	0.9593	0.9380	0.9141
Dict. II	0.9556	0.9526	0.9390	0.9144
Dict. III	0.9556	0.9509	0.9396	0.9155
Dict. IV	0.9589	0.9592	0.9314	0.9104
Dict. V	0.9598	0.9606	0.9325	0.9115
Dictionary	TID2008		TID2013	
	PLCC	SRCC	PLCC	SRCC
Dict. I	0.8891	0.8869	0.9004	0.8927
Dict. II	0.8885	0.8860	0.8994	0.8911
Dict. III	0.8847	0.8768	0.8957	0.8834
Dict. IV	0.8903	0.8860	0.9023	0.8926
Dict. V	0.8913	0.8901	0.9020	0.8956

Table III summarizes the experimental results using the five dictionaries on LIVE, CSIQ, TID2008 and TID2013 databases. It is easily observed from the table that when five different dictionaries are used to predict the sharpness scores, the performances are quite similar. Another finding is that although dictionaries I, II, IV and V are trained using very limited number of images, their performances are very similar to that of dictionary III, which uses much more images and patches. These results indicate that the proposed method is not sensitive to training images. Therefore, a universal dictionary can be used to predict the sharpness of common images, and this dictionary can be learned using a small number of randomly selected natural images. This is helpful for real-world applications. In the following experiments, dictionary I is used.

E. Impacts of Gradient Operator and Sparse Representation

In the proposed method, a gradient operator is applied before sparse coding, which is mainly adopted for representing image blur better. This is not hard to understand because blur is characterized by the spread of image edges, so gradient domain is more effective for blur representation. Meanwhile, readers may wonder why don't we compute image energy using gradient directly. In this part, we conduct two separate experiments to investigate the contributions of gradient and sparse coding in the proposed method. In the first experiment, we remove the gradient operator and conduct sparse coding directly. In the

TABLE IV
PERFORMANCES OF THE PROPOSED METHOD WHEN ONLY SPARSE CODING OR GRADIENT IS USED

Database	w/o gradient		w/o sparse coding	
	PLCC	SRCC	PLCC	SRCC
LIVE	0.8624	0.8519	0.9344	0.9289
CSIQ	0.8637	0.8211	0.9125	0.8843
TID2008	0.7336	0.7421	0.8492	0.8332
TID2013	0.7357	0.7207	0.8599	0.8355

TABLE V
PERFORMANCE COMPARISON OF SPARISH AND STATE-OF-THE-ART GENERAL-PURPOSE NR IMAGE QUALITY METRICS

Metric	LIVE		CSIQ	
	PLCC	SRCC	PLCC	SRCC
BIQI [12]	Training on LIVE		0.8556	0.7713
BLIINDS-II [13]	Training on LIVE		0.9102	0.8915
BRISQUE [14]	Training on LIVE		0.9273	0.9025
NIQE [15]	0.9433	0.9328	0.9260	0.8943
SPARISH	**0.9595**	**0.9593**	**0.9380**	**0.9141**
Metric	TID2008		TID2013	
	PLCC	SRCC	PLCC	SRCC
BIQI [12]	0.7550	0.7468	0.7819	0.7642
BLIINDS-II [13]	0.8415	0.8388	0.8580	0.8557
BRISQUE [14]	0.8043	0.7990	0.8248	0.8143
NIQE [15]	0.8302	0.8150	0.8146	0.7958
SPARISH	**0.8891**	**0.8869**	**0.9004**	**0.8927**

second experiment, we remove sparse coding and compute block energy using gradient. The experimental results are listed in Table IV.

It is observed from Table IV that the performances of both cases are significantly worse than that of the original model. This further indicates that both gradient and sparse coding are indispensable for the proposed model.

F. Comparison With General-Purpose NR Quality Metrics

Several effective general-purpose NR metrics have been proposed to evaluate image quality without knowing the distortion types. Among them are the well-known Blind Image Quality Index (BIQI) [12], BLind Image Integrity Notator using DCT Statistics (BLIINDS-II) [13], Blind/Referenceless Image Spatial QUality Evaluator (BRISQUE) [14] and Natural Image Quality Evaluator (NIQE) [15]. In this part, we compare our method with these metrics. The simulation results are listed in Table V, where the best result for each database is marked in boldface. It should be noted that BIQI, BLIIND-II and BRISQUE use the images in the LIVE database to train the quality model, so their results in LIVE database do not provide fair comparisons. As a result, they are marked by "Training on LIVE" in the table.

It is observed from the table that the proposed method produces the best results in all the databases. Especially in TID2008 and TID2013, SPARISH significantly outperforms the general-purpose quality metrics. This further confirms the advantage of the proposed method over the general-purpose quality metrics with respect to sharpness assessment.

G. Performance on Multiply Distorted Images

In real-world applications, images are very likely to suffer multiple distortions. Therefore, it is important to evaluate the performance of a sharpness metric on real images. To this end, we further test SPARISH on LIVE MD [47] and CID2013 [49], [50], two datasets with multiply distorted images. LIVE MD consists of images corrupted under two multiple distortion scenarios [48]: 1) image storage where images are first blurred and then compressed by a JPEG encoder; 2) camera image acquisition where images are first blurred due to narrow depth of field or other defocus and then corrupted by white Gaussian noise to simulate sensor noise. For each scenario, there are 225 images with different distortion levels. CID2013 contains more realistic images with multidimensional distortions, which are taken by consumer cameras and mobile phones. In the current version, there are 474 images captured by 79 imaging devices. For each image, the subjective scores of sharpness, graininess, lightness, color saturation are provided, together with an overall quality score. In our experiment, we only use the sharpness scores. It should be noted that although these two databases are employed to test the performance of the proposed method on more realistic distortions, they are quite different. To be specific, the images in LIVE MD are obtained by sequentially adding distortions to high quality pristine images (using MATLAB functions), which aims to simulate realistic distortions. By contrast, CID2013 consists of images taken by consumer cameras and mobile phones, so it is built in real-world environment. Fig. 12 shows some example images in LIVE MD and CID2013 databases. Table VI summarizes the experimental results of SPARISH and the state-of-the-art metrics on LIVE MD and CID2013 databases, where the best results are marked in boldface.

Evaluating the sharpness of multiply distorted images is challenging. This is because most metrics evaluate image sharpness by computing the attenuation of high-frequency components, but distortions may lead to increase of high-frequency information, such as noise and JPEG blocking artifacts. It is observed from Table VI that the proposed method performs the best for "Blur and JPEG" in LIVE MD, and the results are significantly better than the compared metrics. For "Blur and noise," SPARISH achieves the best prediction monotonicity (SRCC), while the prediction accuracy (PLCC) is medium. In CID2013, FISH_{bb} [20] achieves the best performance. The proposed method also produces very promising results, and the prediction accuracy and monotonicity rank the second and fourth, respectively. In fact, the performances of SPARISH, S3, LPC and MLV are comparable, and they significantly outperform the other metrics. Another finding from the results is that while many of the metrics achieve very promising results on "Blur and JPEG" (simulating distortions in image storage), none of them of perform well on "Blur and noise" and CID2013 (simulating distortions in image acquisition). This indicates that quality

Fig. 12. Example images in LIVE MD and CID2013 databases.

TABLE VI
PERFORMANCE COMPARISON OF SPARISH AND STATE-OF-THE-ART SHARPNESS METRICS ON LIVE MD AND CID2013 DATABASES

Database	Criterion	Marz. [16]	JNB [17]	CPBD [18]	S3 [19]	FISH [20]	FISH$_{bb}$ [20]	LPC [21]	SVC [22]	MLV [23]	SPARISH
LIVE MD (Blur + JPEG)	PLCC	0.4707	0.8109	0.5174	0.7479	0.8221	0.8330	0.8130	0.8114	0.8588	**0.9163**
	SRCC	0.4235	0.7677	0.4366	0.5802	0.7407	0.7439	0.6935	0.6506	0.8039	**0.8890**
LIVE MD (Blur + Noise)	PLCC	0.3092	0.1868	0.3168	0.3890	0.2343	0.5224	0.5740	0.2362	**0.5824**	0.3125
	SRCC	0.1689	0.0361	0.2248	0.2823	0.1845	0.3031	0.4323	0.1217	0.4824	**0.5114**
CID2013	PLCC	0.4241	0.4402	0.4177	0.6845	0.6383	**0.7229**	0.6343	0.4251	0.6901	0.6608
	SRCC	0.3248	0.3406	0.2932	0.6480	0.5870	**0.6830**	0.6048	0.2330	0.5943	0.6520

metrics that can effectively measure the distortions in real image acquisition are still lacking.

V. CONCLUSION

High-level features are important in our understanding of visual scenes. Extraction of high-level features is expected to facilitate IQA. Towards this direction, we have proposed a new NR image sharpness metric based on learned dictionary from natural image patches, which has been proved to be able to capture high-level features for image analysis. Our method employs the variance-normalized energy to evaluate image sharpness across images. The energy measures the relative strength of blur in an image, and the variance-based normalization removes the influence of image contents. The proposed method can evaluate the sharpness of different images using a universal dictionary, which can be learned using a small number of natural images. This is beneficial for real-world applications.

We have done extensive experiments on six public image quality databases consisting of both purely blurred images and multiply distorted images. The experimental results demonstrate that the proposed method can produce sharpness scores highly consistent with human perception. We have also compared the proposed method with the state-of-the-art image sharpness metrics as well as several general-purpose NR image quality metrics, and the results demonstrate the advantages of the proposed method.

ACKNOWLEDGMENT

The authors would like to thank both the Editor and the Reviewers for the invaluable comments and suggestions that helped greatly in improving the quality this paper. They would also like to thank Prof. Q. Sang from Jiangnan University, Wuxi, China, for providing their codes in [22] to conduct a comparison in the revision of this paper.

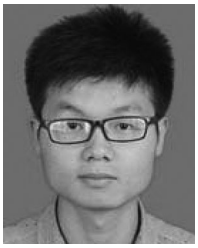
REFERENCES

[1] W. S. Lin and C.-C. Jay Kuo, "Perceptual visual quality metrics: A survey," *J. Vis. Commun. Image Represent.*, vol. 22, no. 4. pp. 297–312, May. 2011.
[2] L. Ma, C. W. Deng, K. N. Ngan, and W. S. Lin, "Recent advances and challenges of visual signal quality assessment," *China Commun.*, vol. 20, no. 5, pp. 62–78, May. 2013.
[3] D. M. Chandler, "Seven challenges in image quality assessment: Past, present, and future research," *ISRN Signal Process.*, vol. 2013, pp. 1–53, 2013, Art. no. 905685.
[4] 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.
[5] L. Zhang, L. Zhang, X. Q. 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.
[6] W. F. Xue, L. Zhang, X. Q. Mou, and A. C. Bovik, "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index," *IEEE Trans. Image Process.*, vol. v, no. 2, pp. 684–695, Feb. 2014.
[7] L. Ma, S. N. Li, F. Zhang, and K. N. Ngan, "Reduced-reference image quality assessment using reorganized DCT-based image representation," *IEEE Trans. Multimedia*, vol. 13, no. 4, pp. 824–829, Aug. 2011.
[8] J. J. Wu, W. S. Lin, G. M. Shi, and A. M. Liu, "Reduced-reference image quality assessment with visual information fidelity," *IEEE Trans. Multimedia*, vol. 15, no. 7, pp. 1700–1705, Nov. 2013

[9] L. D. Li, H. C. Zhu, G. B. Yang, and J. S. Qian, "Referenceless measure of blocking artifacts by Tchebichef kernel analysis," *IEEE Signal Process. Lett.*, vol. 21, no. 1, pp. 122–125, Jan. 2014.
[10] L. D. Li, W. S. Lin, and H. C. Zhu, "Learning structural regularity for evaluating blocking artifacts in JPEG images," *IEEE Signal Process. Lett.*, vol. 21, no. 8, pp. 918–922, Aug. 2014.
[11] H. T. Liu, N. Klomp, and I. Heynderickx, "A no-reference metric for perceived ringing artifacts in images," *IEEE Trans. Circuits Syst. Video*, vol. 20, no. 4, pp. 529–539, Apr. 2010.
[12] A. K. Moorthy and A. C. Bovik, "A two-step framework for constructing blind image quality indices," *IEEE Signal Process. Lett.*, vol. 17, no. 5, pp. 513–516, May 2010.
[13] M. A Saad and A. C. Bovik, "Blind image quality assessment: A natural scene statistics approach in the DCT domain," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3339–3352, Aug. 2012.
[14] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4695–4708, Dec. 2012.
[15] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a completely blind image quality analyzer," *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 209–212, Mar. 2013.
[16] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, "Perceptual blur and ringing metrics: Application to JPEG2000," *Signal Process., Image Commun.*, vol. 19, no. 2, pp. 163–172, Feb. 2004.
[17] R. Ferzli and L. J. Karam, "A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB)," *IEEE Trans. Image Process.*, vol. 18, no. 4, pp. 717–728, Apr. 2009.
[18] N. D. Narvekar and L. J. Karam, "A no-reference image blur metric based on the cumulative probability of blur detection (CPBD)," *IEEE Trans. Image Process.*, vol. 20, no. 9, pp. 2678–2683, Sep. 2011.
[19] C. T. Vu, T. D. Phan, and D. M. Chandler, "S3: A spectral and spatial measure of local perceived sharpness in natural images," *IEEE Trans. Image Process.*, vol. 21, no. 3, pp. 934–945, Mar. 2012.
[20] P. V. Vu and D. M. Chandler, "A fast wavelet-based algorithm for global and local Image sharpness estimation," *IEEE Signal Process. Lett.*, vol. 19, no. 7, pp. 423–426, Jul. 2012.
[21] R. Hassen, Z. Wang, and M. Salama, "Image sharpness assessment based on local phase coherence," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2798–2810, Jul. 2013.
[22] Q. B. Sang, H. X. Qi, X. J. Wu, C. F. Li, and A. C. Bovik, "No-reference image blur index based on singular value curve," *J. Vis. Commun. Image Representation*, vol. 25, no. 7, pp. 1625–1630, Oct. 2014.
[23] K. Bahrami and A. C. Kot, "A fast approach for no-reference image sharpness assessment based on maximum local variation," *IEEE Signal Process. Lett.* vol. 21, no. 6, pp. 751–755, Jun. 2014.
[24] L. J. Li, H. Su, Y. W. Lim, and F. F. Li, "Object bank: An object-level image representation for high-level visual recognition," *Int. J. Comput. Vis.* vol. 107, no. 1, pp. 20–39, Mar. 2014.
[25] Q. V. Le *et al.*, "Building high-level features using large scale unsupervised learning," in *Proc. Int. Conf. Mach. Learn.*, 2012, pp. 81–88.
[26] L. Jiang, A. G. Hauptmann, and G. Xiang, "Leveraging high-level and low-level features for multimedia event detection," in *Proc. 20th ACM Int. Conf. Multimedia*, 2012, pp. 449–458.
[27] H. W. Chang, H. Yang, Y. Gan, and M. H. Wang, "Sparse feature fidelity for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 22, no. 10, pp. 4007–4018, Oct. 2013.
[28] T. Guha, E. Nezhadarya, and R. K. Ward, "Sparse representation-based image quality assessment," *Signal Process., Image Commun.*, vol. 29, no. 10, pp. 1138–1148, Nov. 2014.
[29] L. H. He, D. C. Tao, X. L. Li, and X. B. Gao, "Sparse representation for blind Image quality assessment," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recog.*, Jun. 2012, pp. 1146–1153.
[30] Y. Ding, H. Dai, and S. Z. Wang, "Image quality assessment scheme with topographic independent components analysis for sparse feature extraction," *Electron. Lett.*, vol. 50, no. 7 pp. 509–510, Mar. 2014.
[31] B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, vol. 381, no. 13, pp. 607–609, Jun. 1996.
[32] B. A. Olshausen and D. J. Field, "Sparse coding with an overcomplete basis set: A strategy employed by V1?" *Vis. Res.*, vol. 37, no. 23, pp. 3311–3325, Dec. 1997.
[33] H. Lee, A. Battle, R. Raina, and A. Y. Ng, "Efficient sparse coding algorithms," in *Proc. 19th Annu. Conf. Neural Inform. Process. Syst.*, 2007, pp. 801–808.
[34] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.
[35] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proc. 27th Annu. Asilomar Conf. Signals, Syst. Comput.*, 1993, pp. 40–44.
[36] J. C. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Signal Process.*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010.
[37] S. Q. Wu *et al.*, "Blind blur assessment for vision-based applications," *J. Vis. Commun. Image Represent.*, vol. 20, no. 4, pp. 231–241, May 2009.
[38] 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.
[39] 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. 1–21, Jan. 2010, Art. no. 011006.
[40] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti, "TID2008-a database for evaluation of full-reference visual quality assessment metrics," *Adv. Mod. Radioelectron.*, vol. 10, no. 4, pp. 30–45, 2009.
[41] N. Ponomarenko *et al.*, "Color image database TID2013: Peculiarities and preliminary results," in *Proc. Eur. Workshop Visual Inform. Process.*, 2013, pp. 106–111.
[42] A. V. Murthy and L. J. Karam. (2010). *VBQUEST—Visual Blur Quality Evaluation Software*. [Online]. Available: http://ivulab.asu.edu/Quality/VBQUEST
[43] C. T. Vu, T. D. Phan, P. V. Vu, and D. M. Chandler. (2012). *Source Codes of S3, FISH, and FISH$_{bb}$ Indices*. [Online]. Available: http://vision.okstate.edu
[44] R. Hassen, Z. Wang, and M. Salama. (2013). *Source Codes of LPC Index*. [Online]. Available: https://ece.uwaterloo.ca/ rhassen/LPC-SI/
[45] K. Bahrami and A. C. Kot. (2014). *Source Codes of MLV Index*. [Online]. Available: https://sites.google.com/site/khosrobahrami2010/publications
[46] D. Jayaraman, A. Mittal, A. K. Moorthy, and A. C. Bovik, "Objective quality assessment of multiply distorted images," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, 2012, pp. 1693–1697.
[47] T. Virtanen, M. Nuutinen, M. Vaahteranoksa, P. Oittinen, and J. Häkkinen, "CID2013: A database for evaluating no-reference image quality assessment algorithms," *IEEE Trans. Image Process.*, vol. 24, no. 1, pp. 390–402, Jan. 2015.
[48] T. Virtanen, M. Nuutinen, M. Vaahteranoksa, P. Oittinen, and J. Häkkinen. (2013). *Camera Image Database (CID2013)* [Online]. Available: http://www.helsinki.fi/∼msjnuuti/CID2013/
[49] D. M. Chandler and S. S. Hemami, "VSNR: A wavelet-based visual signal-to-noise ratio for natural images," *IEEE Trans. Image Process.*, vol. 16, no. 9, pp. 2284–2298, Sep. 2007.
[50] M. Evans, N. Hastings, and B. Peacock, *Statistical Distributions*. Hoboken, NJ, USA: Wiley-Interscience, 2000.

Leida Li received the B.S. and Ph.D. degrees from Xidian University, Xi'an, China, in 2004 and 2009, respectively.

From February 2008 to June 2008, he was a Visiting Ph.D. Student with the Department of Electronic Engineering, National Kaohsiung University of Applied Sciences, Kaohsiung, Taiwan. From January 2014 to January 2015, he was a Visiting Research Fellow with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. He is currently a Professor with the School of Information and Electrical Engineering, China University of Mining and Technology, Xuzhou, China. His research interests include multimedia quality assessment, information hiding, and image forensics.



Dong Wu received the B.E. degree from Hubei Engineering University, Hubei, China, in 2013, and is currently working toward the M.S. degree in information and electrical engineering at the China University of Mining and Technology, Xuzhou, China.

His research interests include image quality assessment and sparse representation.

Jinjian Wu received the B.Sc. and Ph.D. degrees from Xidian University, Xi'an, China, in 2008 and 2013, respectively.

From September 2011 to March 2013, he was a Research Assistant with Nanyang Technological University, Singapore. From August 2013 to August 2014, he was a Postdoctoral Research Fellow with the Nanyang Technological University. From July 2013 to June 2015, he was a Lecturer with Xidian University. Since July 2015, he has been an Associate Professor with the School of Electronic Engineering, Xidian University. His research interests include visual perceptual modeling, saliency estimation, quality evaluation, and just noticeable difference estimation.

Prof. Wu has served as a TPC Member for ICME2014, ICME2015, PCM2015, and ICIP2015. He was the recipient of the Best Student Paper Award of ISCAS 2013.

Haoliang Li received the B.S. degree from the University of Electronic Science and Technology of China, Chengdu, China, in 2013, and is currently working toward the Ph.D. degree in electrical and electronic engineering at Nanyang Technological University, Singapore.

His research interests include image processing and multimedia forensics.

Weisi Lin (S'91–M'92–SM'00–F'16) received the B.S. and M.S. degrees from Zhongshan University, Guangzhou, China, in 1982 and 1985, respectively, and the Ph.D. degree from King's College, London University, London, U.K., in 1992.

He was previously the Laboratory Head of Visual Processing and the Acting Department Manager of Media Processing for the Institute for Infocomm Research, Singapore. He is currently an Associate Professor with the School of Computer Engineering, Nanyang Technological University, Singapore. He has authored or coauthored more than 200 refereed papers in international journals and conferences. His research interests include image processing, perceptual modeling, video compression, multimedia communication, and computer vision.

Prof. Lin is a Chartered Engineer (U.K.), a Fellow of the Institution of Engineering Technology, and an Honorary Fellow of the Singapore Institute of Engineering Technologists. He is on the Editorial Board of the IEEE TRANSACTIONS ON IMAGE PROCESSING, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the IEEE SIGNAL PROCESSING LETTERS, and the *Journal of Visual Communication and Image Representation*. He served as the Lead Guest Editor for a Special Issue on perceptual signal processing, the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING (2012). He is a Co-Chair of the IEEE MMTC Special Interest Group on Quality of Experience. He has been elected as a Distinguished Lecturer of APSIPA (2012/2013). He was the Lead Technical Program Chair for the Pacific-Rim Conference on Multimedia 2012, and was a Technical Program Chair of the IEEE International Conference on Multimedia and Expo 2013.

Alex C. Kot (S'85–M'89–SM'98–F'06) has been with the Nanyang Technological University, Singapore, since 1991. He was the Head of the Division of Information Engineering at the School of Electrical and Electronic Engineering for eight years and served as an Associate Chair/Research and Vice Dean Research for the School of Electrical and Electronic Engineering. He is currently a Professor and an Associate Dean for the College of Engineering and the Director of the Rapid-Rich Object Search Lab, Singapore. His research interests include the areas of signal processing for communication, biometrics, data-hiding, image forensics, and information security.

Prof. Kot is an IEEE SPS Distinguished Lecturer and a Fellow of the Academy of Engineering, Singapore. He served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE TRANSACTIONS ON MULTIMEDIA, the IEEE SIGNAL PROCESSING LETTERS, the *IEEE Signal Processing Magazine*, the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS PART I and PART II, and the IEEE TRANSACTIONS ON IMAGE PROCESSING. He also served as the Guest Editor for the Special Issues for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY and the EURASIP *Journal on Advances in Signal Processing*. He is currently an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY. He is also the Editor for the EURASIP *Journal on Advances in Signal Processing*. He has served the IEEE SP Society in various capacities such as the General Co-Chair for the 2004 IEEE International Conference on Image Processing and the Chair of the worldwide SPS Chapter Chairs and the Distinguished Lecturer Program. He was a Member of the IEEE Fellow Evaluation Committee. He is currently the Vice President for the IEEE Signal Processing Society. He was the recipient of the Best Teacher of the Year Award and is a coauthor for several Best Paper Award recipients including ICPR, IEEE WIFS, ICEC, and IWDW.