Low-Rank Decomposition-Based Restoration of Compressed Images via Adaptive Noise Estimation

Xinfeng Zhang, *Member, IEEE*, Weisi Lin, *Fellow, IEEE*, Ruiqin Xiong, *Member, IEEE*, Xianming Liu, *Member, IEEE*, Siwei Ma, *Member, IEEE*, and Wen Gao, *Fellow, IEEE*

***Abstract*— Images coded at low bit rates in real-world applications usually suffer from significant compression noise, which significantly degrades the visual quality. Traditional denoising methods are not suitable for the content-dependent compression noise, which usually assume that noise is independent and with identical distribution. In this paper, we propose a unified framework of content-adaptive estimation and reduction for compression noise via low-rank decomposition of similar image patches. We first formulate the framework of compression noise reduction based upon low-rank decomposition. Compression noises are removed by soft thresholding the singular values in singular value decomposition of every group of similar image patches. For each group of similar patches, the thresholds are adaptively determined according to compression noise levels and singular values. We analyze the relationship of image statistical characteristics in spatial and transform domains, and estimate compression noise level for every group of similar patches from statistics in both domains jointly with quantization steps. Finally, quantization constraint is applied to estimated images to avoid over-smoothing. Extensive experimental results show that the proposed method not only improves the quality of compressed images obviously for post-processing, but are also helpful for computer vision tasks as a pre-processing method.**

***Index Terms*— Block transform coding, compression noise, patch clustering, denoising, low-rank, SVD.**

I. Introduction

ALONG with the fast development of portable digital devices, *e.g.*, digital cameras and smart phones, more and more images and videos are captured and shared through Internet or mobile networks. Due to the limitation of bandwidth, the shared images and videos are usually compressed at low bit rates, and their qualities are severely degenerated by compression noise. These low quality images are not only with poor user experience but also deteriorate the performance of many computer vision algorithms, which are mainly designed for uncompressed images and videos.

In order to improve the quality of compressed images, there are lots of image denoising methods proposed in recent years. Most of the existing denoising algorithms assume additive white Gaussian noises, which are independent and with identical distribution in a whole image. In general denoising procedures, the standard deviation of noise is usually assumed known and utilized to control filtering strength to avoid smoothing image structures excessively. However, in practical scenarios, compression noise is content correlated and is also difficult to estimate with existing noise level estimation methods.

In this paper, we investigate the compression noise estimation and reduction for block-discrete cosine transform (BDCT) based compression. Since compression noises are mainly generated by quantizing transform coefficients, they are dependent on the distribution of coefficients. To remove compression noise as much as possible, we propose a content-aware method to reduce compression noise by dividing images into different groups of similar image patches. For each group of similar image patches, we formulate the compression noise reduction as a low-rank optimization problem, and solve it via soft-thresholding the singular values in singular value decomposition (SVD) of group of similar image patches. Since the thresholds for singular values are directly related with compression noise levels, we also propose a content-dependent compression noise estimation algorithm. First, we derive the distribution parameters of coefficients from image correlation model. Then, we derive the standard deviation of compression noise for each group according to coefficient distribution and quantization steps. Finally, we take the weighted average of all these estimated patches to restore original images. To avoid over-smoothing, narrow quantization constraint (NQC) [1] is applied to the restored image. Therefore, our method with content-dependent noise estimation is different from previous works, which only utilize a global noise level by assuming i.i.d for compression noise. Extensive experimental results show that our method can significantly improve the quality of compressed images, and it is also helpful for computer vision tasks as a pre-processing method.

Manuscript received September 17, 2015; revised March 21, 2016 and June 27, 2016; accepted June 28, 2016. Date of publication July 7, 2016; date of current version July 21, 2016. This work in part by the Rapid-Rich Object Search (ROSE) Lab at the Nanyang Technological University, Singapore. The ROSE Lab is supported by the National Research Foundation, Singapore, under its Interactive Digital Media (IDM) Strategic Research Programme. This work was also cooperated with Peking University, and partially supported by the National Basic Research Program of China (973 Program) under Grant 2015CB351800. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Jean-Francois Aujol. *(Corresponding author: Weisi Lin.)*

X. Zhang and W. Lin are with the Rapid-Rich Object Search (ROSE) Lab, Nanyang Technological University, Singapore 637553 (e-mail: xfzhang@ntu.edu.sg; wsling@ntu.edu.sg).

R. Xiong, 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: rqxiong@pku.edu.cn; swma@pku.edu.cn; wgao@pku.edu.cn).

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/TIP.2016.2588326

TABLE I

MAIN NOTATIONS AND THEIR DESCRIPTIONS

Notations	Descriptions
$\boldsymbol{x}_{\mathcal{B}}$, $\boldsymbol{x}_{G_i}^{(k)}$	Image patches in original image with size of $p \times p$
$\mathbf{X}_{\mathcal{B}}$	Transformed image patches in original image with size of $p \times p$
$\boldsymbol{y}_{\mathcal{B}}$, $\boldsymbol{y}_{G_i}^{(k)}$	Image patches in compressed image with size of $p \times p$ in spatial domain
$\mathbf{Y}_{\mathcal{B}}$	Transformed image patches in compressed image with size of $p \times p$
$\mathbf{X}_{G_i}$	Image patch matrix extracted from original image, $p^2 \times K$
$\mathbf{Y}_{G_i}$	Image patch matrix extracted from compressed image, $p^2 \times K$
$\mathbf{N}_{G_i}$	Compression noise in image patch matrix, $p^2 \times K$
Q	Quantization matrix
$g(n\|y = tq)$	Noise distribution function when dequantized value is equal to tq
$\sigma_{\boldsymbol{x}}$, $\mathbf{\Sigma}_{\mathbf{X}}$	Standard deviation of image patches in spatial and transform domain, respectively
σ_{n,G_i}, $\mathbf{\Sigma}_{n,G_i}$	Standard deviation of compression noise for image patch group G_i in spatial and transform domain, respectively

The remainder of this paper is organized as follows. Section II reviews related works and introduces useful notations. Section III introduces the framework of the proposed compression noise reduction method based on low-rank decomposition. Section IV introduces the proposed content-dependent compression noise estimation method for each group of similar image patches. The iterative implementation of the proposed method is introduced in Section V. Section VI report and analyses the experimental results, and conclusion of this paper is in Section VII.

II. RELATED WORKS

In this section, a few concepts and notations related with block transform image coding are briefly reviewed. Then, we review the image compression noise reduction methods and the compression noise level estimation methods, respectively. The main notations in this paper are listed in Table I.

A. Block Transform Image Coding

Block transform coding is the most widely used image coding framework, in which the block discrete cosine transform (BDCT) is adopted by most of popular image/video coding standards, *e.g.*, JPEG [2]. In a typical BDCT coding framework, an input image, $\mathcal{I}$, is divided into non-overlapped $N \times N$ blocks. Each block is transformed into frequency domain using DCT, and then the transform coefficients are quantized independently, and compressed by entropy coding. At the decoder side, the inverse procedure is carried out to reconstruct images. The whole process can be described as,

$$\mathbf{X}_{\mathcal{B}} = \mathcal{T}(\boldsymbol{x}_{\mathcal{B}}), \tag{1}$$

$$\mathbf{Y}_{\mathcal{B}}(u, v) = round\left(\frac{\mathbf{X}_{\mathcal{B}}(u, v)}{Q(u, v)}\right)Q(u, v), \tag{2}$$

where $\mathcal{T}$ is a transform operator, and $\boldsymbol{x}_{\mathcal{B}}$ and $\mathbf{X}_{\mathcal{B}}$ are the original image data of block $\mathcal{B}$ in spatial domain and transform domain respectively. $Q(u, v)$ is the quantization step for frequency band (u, v) and $round(\cdot)$ means rounding real values into their nearest integers. $\mathbf{Y}_{\mathcal{B}}(u, v)$ is the reconstructed coefficient.

During the above process, the main source producing compression noise is quantization in Eqn.(2). Since coefficients in high frequency bands are very small for most image blocks, the quantization operator may directly make them zeros, which leads to insufficient coefficients to represent image local structures and may generate ringing artifacts around edges. In addition, due to the loss of inter-block correlation in quantization process, similar coefficients in neighboring blocks may be quantized into different ranges, which leads to discontinuities at block boundaries, referred as to be blocking artifacts. We will consider these distortions due to lossy compression as a specific type of content correlated noise in the following sections of this paper.

B. Compression Noising Reduction Methods

During the past three decades, numerous image denoising methods have been proposed in the image processing field. These methods can be roughly classified into two categories, general denoising methods and specific denoising methods. General denoising methods usually assume noise following independent identical distribution, and utilize some image prior models, *e.g.*, local smoothness priors and sparse priors, to depress the noise magnitudes, which do not conform with image prior models. Besides the image prior models, specific denoising methods further take advantage of some side information on noise to improve their performance, *e.g.*, quantization steps in compression noise reduction. This paper will focus on compression noise reduction problem.

Although general denoising methods also can alleviate the compression artifacts, e.g., BM3D [3] and CSR [4], their performances are not as good as that in removing noise with independent identical distribution (i.i.d). Therefore, a lot of denoising methods specially designed for compression noise are proposed in literatures, e.g., [5]–[20]. Reeve and Lim [5] smooth out the blocking artifacts by directly applied a 3×3 Gaussian filter to the boundary pixels. In [6], a nonlinear space-variant filters based on edge-oriented classifiers is utilized to reduce blocking artifacts while preserving image edges. To avoid oversmoothing image textures, Minami and Zakhor [7] utilized the quantization intervals of transform coefficients as a convex set to constrain the filtered coefficients. In the latest video coding standards, *e.g.*, HEVC [8], the strength of deblocking filter is adaptive according to coding modes, and the adaptive loop filter (ALF) [9], [10] in HEVC is utilized to reduce the coding artifacts by deriving Wiener filters in encoder side. Maggioni *et al.* [21] extended BM3D by utilizing temporal information to reduce video compression noise and estimated a global noise level for each frame based on quantization steps.

Besides the filtering approaches, there are also many compression noise reduction schemes with probability estimation based on image prior models. Sun and Cham [11] used a high order Markov random field (MRF) to mode original images based on the field of experts (FoE) framework to reduce compression noise. Zhang *et al.* [12], [13] proposed a multi-prediction adaptive fusion framework to remove compression noise with multiple image prior models. Ren *et al.* [14] applied low rank constraint to groups of similar image patches to reduce compression artifacts, and Zhang *et al.* [20] further improved this method by estimating noise level adaptively. Chang *et al.* [15] first learned a dictionary from compressed image, and then restored latent original image by jointly minimizing image total variations and difference between estimated image and its sparse representation with learned dictionaries. Liu *et al.* [16], [22] introduced quantization in DCT domain dictionary learning, and proposed to remove compression noise by jointly sparse coding for compressed images in spatial and transform domains. There are also numerous methods specially designed for compression noise reduction in literatures, *e.g.*, [17] and [18].

C. Noise Level Estimation

Noise level, *e.g.*, standard deviation of noise, is generally utilized to adjust the filtering strength of denoising. However, in practice, the noise level is unknown and need to be estimated from noisy images, which is also a difficult problem, especially for content correlated noise. An intuitive method to estimate noise level is to calculate the standard deviation of difference between noisy image and filtered image by low-pass filter [23]. Obviously, this method is difficult to get accurate noise level estimation, because low pass filter is inefficient to remove content-dependent noises. Based on i.i.d assumption, Lee [24] proposed to calculate standard deviation of noise from homogeneous regions, where pixel variance is mainly dominated by noise. However, it is also a difficult problem to identify the homogeneous regions in noisy images. Sutour *et al.* [25] proposed a non-parametric approach to detect homogeneous regions based on Kendall's τ coefficient and estimated noise level function of stationary noise assuming that the noise is spatially uncorrelated.

Considering image sparsity in transform domain, Donoho proposed to estimate noise variance from the coefficients in diagonal bands of image wavelet decomposition [26]. Liu *et al.* [27] estimated content uncorrelated noise level from single image by applying PCA to weak texture patches. Ponomarenko *et al.* [28] utilized the K most similar blocks in an image to estimate noise variance by applying median filtering operation to these blocks in DCT domain. Colom *et al.* [29] further extended the method in [28] to estimate the variance of noise according to the intensity and the frequency. Ponderated MSE is proposed to measure the similarity among image patches to reduce the negative effects of noise, by giving higher importance to the low frequencies of the blocks. Colom *et al.* [30] further proposed a multi-scale approach to improve the estimation accuracy for noise low frequencies, which may not be captured by a small image patch in original resolution. This method efficiently improves the performance of the intensity-frequency dependent noise reduction [31]. However, the compression noise is more complex, and is difficult to describe by a simple model, e.g., polynomial, based on image intensity and frequency.

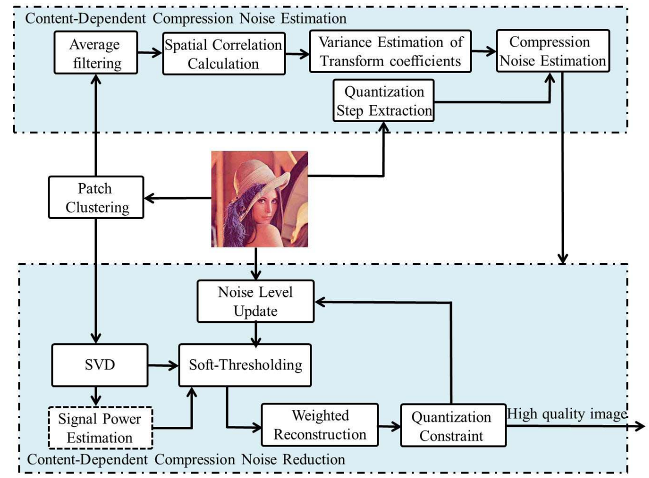


Fig. 1. Framework of content-dependent compression noise level estimation and reduction.

Furthermore, for images compressed at very low bit rates, these methods can be inefficient to estimate noise level since noise is highly correlated. This is also explained in the conclusion section of [30].

III. Compression Noise Reduction via Low-Rank Decomposition

In this section, we introduce the framework of the proposed compression noise reduction method via low-rank decomposition, which is illustrated in Fig.1. The top part shows the framework of the compression noise level estimation method, and the bottom part shows the framework of the proposed compression noise reduction method. Herein, the adaptive thresholds with content-dependent noise level in noise reduction method and the corresponding content-dependent estimation method for compression noise are the main contributions of this paper.

In the proposed compression noise reduction method, a compressed image, $\mathcal{I}_{\boldsymbol{y}}$, is divided into a set of overlapped $p \times p$ image patches firstly, denoted as *target patches*, which are extracted every s pixels (denoted as *overlapped step*) along raster scanning order, and they are overlapped when $s < p$. For each *target patch*, the K most-similar patches (including itself) are found in a $R \times R$ neighborhood around it ($R = 31$ in this paper), and these similar patches are denoted as *reference patches* and selected according to their similarity measured with the following equation,

$$d_j = \|\boldsymbol{y}_t - \boldsymbol{y}_j\|_F^2, \tag{3}$$

where $\boldsymbol{y}_t$ and $\boldsymbol{y}_j$ are *target patch* and *reference patch* respectively, which are rearranged into vectors according to row-major order. $\|\cdot\|_F^2$ is the Frobenius norm of patch vector. After that, the compressed image is organized into different image patch groups, and each group of image patches are arranged into a matrix,

$$\mathbf{Y}_{G_i} = \left[\boldsymbol{y}_{G_i}^{(1)}, \boldsymbol{y}_{G_i}^{(2)}, \cdots, \boldsymbol{y}_{G_i}^{(K)}\right]. \tag{4}$$

We call this procedure *Patch clustering* as shown in Fig.1. Since compression noise levels for image patches with similar structure are usually approximate, a potential advantage of similar patch clustering is that it is convenient to adjust the control parameters of denoising algorithms according to image content group by group.

In order to reduce compression noise, we assume the compression noise as additive one, and the noisy image patch matrix can be rewritten as,

$$\mathbf{Y}_{G_i} = \mathbf{X}_{G_i} + \mathbf{N}_{G_i} \tag{5}$$

where $\mathbf{X}_{G_i}$ and $\mathbf{N}_{G_i}$ are the corresponding original image patch matrix and compression noise matrix respectively. Since column elements in $\mathbf{Y}_{G_i}$ are similar image patches, the matrix is able to be approximated via low-rank decomposition [14]. Then, compression noise reduction can be formulated as the following low-rank optimization problem,

$$\min_{\mathbf{X}_{G_i}} \|\mathbf{Y}_{G_i} - \mathbf{X}_{G_i}\|_F^2, \quad s.t. \ rank\left(\mathbf{X}_{G_i}\right) \leq \tau. \tag{6}$$

Due to the rank of $\mathbf{X}_{G_i}$ equal to the number of its non-zero singular values, it is difficult to solve efficiently. In order to solve the optimization problem in Eqn.(6), we take Nuclear norm [32] of $\mathbf{X}_{G_i}$ to approach its rank as done in many existing works (such as [33]),

$$\|\mathbf{X}_{G_i}\|_* = trace(\sqrt{\mathbf{X}_{G_i}^* \mathbf{X}_{G_i}}) = \sum_{k=1}^{min(p^2,K)} \lambda'_{k,G_i}, \tag{7}$$

where p^2 and K are the dimensions of matrix $\mathbf{X}_{G_i}$ in row and column directions respectively. λ'_{k,G_i} is the k^{th} singular value of $\mathbf{X}_{G_i}$. In further, based on the Lagrange multiplier method, the problem in Eqn.(6) can be rewritten as,

$$\widehat{\mathbf{X}}_{G_i} = \arg\min_{\mathbf{X}_{G_i}} \|\mathbf{Y}_{G_i} - \mathbf{X}_{G_i}\|_F^2 + \tau \|\mathbf{X}_{G_i}\|_*, \tag{8}$$

There are lots of methods ([33]–[35]) proposed to solve the low-rank constraint optimization problem in Eqn.(8). In this paper, we apply the soft-thresholding method to singular values to solve the problem considering its efficiency and less parameters to adjust. The solution to the problem in Eqn.(8) is as follows,

$$\begin{cases} \mathbf{Y}_{G_i} = \mathbf{U}_{G_i} \Lambda_{G_i} \mathbf{V}_{G_i}^* \\ \Lambda_{\boldsymbol{\tau},G_i} = D_{\boldsymbol{\tau}}(\Lambda_{G_i}) \\ \widehat{\mathbf{X}}_{G_i} = \mathbf{U}_{G_i} \Lambda_{\boldsymbol{\tau},G_i} \mathbf{V}_{G_i}^*, \end{cases} \tag{9}$$

where $D_{\boldsymbol{\tau}}(\cdot)$ is a nonlinear function which applies a soft-thresholding rule at level $\boldsymbol{\tau}$ to the singular values of the input matrix. Traditional methods, e.g., [4] and [14], utilize a global noise level to control the threshold for different image patch groups,

$$\tau_{k,G_i} = \frac{\gamma \sigma_n^2}{\sqrt{\lambda'_{k,G_i}}}, \tag{10}$$

where σ_n is the standard deviation of compression noise for the whole image, and γ is a constant. λ'_{k,G_i} is estimated by,

$$\lambda'_{k,G_i} = \sqrt{\lambda_{k,G_i} - \sigma_n^2}, \tag{11}$$

where λ_{k,G_i} is the k^{th} singular value of $\mathbf{Y}_{G_i}$. This procedure is denoted as *Signal Power Estimation* of Fig.1, which is the same as that in [4] and [14]. And the soft-thresholding $D_{\boldsymbol{\tau}}(\cdot)$ for the k^{th} singular value is defined as,

$$D_{\boldsymbol{\tau}}(\lambda_{k,G_i}) = \begin{cases} \lambda_{k,G_i} - \boldsymbol{\tau}_{k,G_i} sign(\lambda_{k,G_i}), & if \ |\lambda_{k,G_i}| > \boldsymbol{\tau}_{k,G_i} \\ 0, & if \ |\lambda_{k,G_i}| \leq \boldsymbol{\tau}_{k,G_i}, \end{cases} \tag{12}$$

Therefore, $\Lambda_{\boldsymbol{\tau},G_i}$ is the singular value matrix with the shrunken singular values.

In this paper, we proposed an adaptive soft-thresholding method to improve the performance of compression noise reduction via content-dependent noise level estimation. We estimate compression noise level for every image patch group, which is shown in the upper part of Fig.1. In our method, the threshold for the k^{th} singular value of image patch matrix $\mathbf{Y}_{G_i}$ is,

$$\tau_{k,G_i} = \frac{\gamma \sigma_{n,G_i}^2}{\sqrt{\lambda'_{k,G_i}}}, \tag{13}$$

$$\lambda'_{k,G_i} = \sqrt{\lambda_{k,G_i} - \sigma_{n,G_i}^2}. \tag{14}$$

The standard deviation of compression noise, σ_{n,G_i}, is estimated from the corresponding groups of similar image patches, which is introduced in Section IV-B.

Considering the overlapping of image patches, there may be multiple estimations for one pixel generated from different groups. In order to reconstruct image while avoiding over-smoothing, after all the image patch groups are processed with soft-thresholding, we reconstruct image by taking the weighted average of overlapped image pixels from different image patches.

$$\hat{\boldsymbol{x}}(i,j) = \sum_{\mathcal{B} \in \Omega_{i,j}} w_{\mathcal{B}} \hat{\boldsymbol{x}}_{\mathcal{B}}(i',j'). \tag{15}$$

Here $\boldsymbol{\Omega}_{i,j}$ is the set of image patches including the pixel $\hat{\boldsymbol{x}}(i,j)$. (i',j') and (i,j) indicate the same pixel under the image patch coordinate and the whole image systems, respectively. $w_{\mathcal{B}}$ is the weight for pixels in patch $\hat{\boldsymbol{x}}_{\mathcal{B}}$, which is determined according to the rank of image patch matrix,

$$\begin{aligned} w_{\mathcal{B}} &= \frac{1}{Z} max\left((1 - \frac{r}{M}), \frac{1}{M}\right), \quad \boldsymbol{x}_{\mathcal{B}} \in G_i, \\ M &= min(p^2, K), \end{aligned} \tag{16}$$

where r is the rank of the matrix $\mathbf{Y}_{G_i}$, and Z is a normalized constant. The weight design is based on our denoising assumption that similar image patch matrix is low-rank. If a group of image patches are decomposed with less non-zero singular values, they more conform to low-rank constraint and the corresponding estimation may be more reliable. This procedure corresponds to the *Weighted Reconstruction* of Fig.1.

In order to avoid oversmoothing, a narrow quantization constraint operator is applied to DCT coefficients based on the theory of narrow quantization constraint set (NQCS) [1], which is also utilized in [17]. In this procedure,

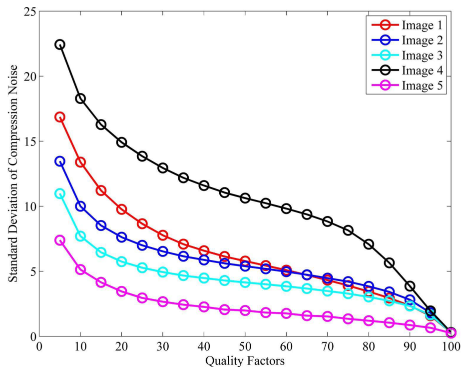


Fig. 2. Compression noise level variations for different images.

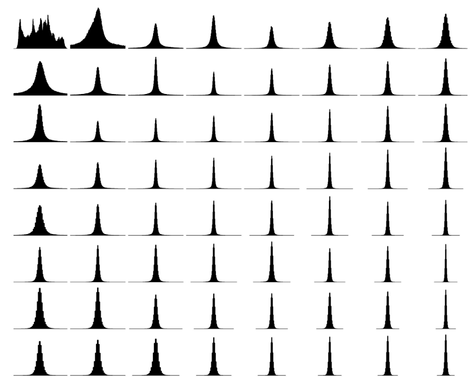


Fig. 3. The histogram of DCT coefficients in 8×8 blocks.

the reconstructed image from Eqn.(15) is further divided into non-overlapped blocks as that in image coding process. For each block, $\widehat{\boldsymbol{x}}_{\mathcal{B}}$, we first transform it into DCT domain,

$$\widehat{\mathbf{X}}_{\mathcal{B}} = \mathcal{T}(\widehat{\boldsymbol{x}}_{\mathcal{B}}) \tag{17}$$

where $\mathcal{T}$ is the transform operation, e.g., 8×8 DCT in JPEG. Then, each coefficient is processed with the narrow quantization constraint operator as follows,

$$\widehat{\mathbf{X}}_{\mathcal{B}}(u,v) = \begin{cases} \mathbf{Y}_{\mathcal{B}}(u,v) + \dfrac{cQ(u,v)}{2}, & \\ \qquad if\ \widehat{\mathbf{X}}_{\mathcal{B}}(u,v) > \mathbf{Y}_{\mathcal{B}}(u,v) + \dfrac{cQ(u,v)}{2} \\ \mathbf{Y}_{\mathcal{B}}(u,v) - \dfrac{cQ(u,v)}{2}, & \\ \qquad if\ \widehat{\mathbf{X}}_{\mathcal{B}}(u,v) < \mathbf{Y}_{\mathcal{B}}(u,v) - \dfrac{cQ(u,v)}{2} \\ \widehat{\mathbf{X}}_{\mathcal{B}}(u,v), \quad others. \end{cases} \tag{18}$$

Here c $(0 < c < 1)$ is a constant, and $c = 0.7$ in our proposed method. This operation corresponds to the *Quantization Constraint* of Fig.1.

IV. Content-Dependent Compression Noise Level Estimation via Patch Grouping

In this section, we first analyze the characteristics of compression noise, and then introduce the proposed content-dependent estimation method for compression noise via patch grouping.

A. Compression Noise Analysis

Since compression noise is directly generated by quantizing transform coefficients, its distribution is not only related with quantization steps but also correlated with distribution of transform coefficients. A simple but intuitive experiment is illustrated in Fig.2, where images are compressed at different quality factors (QF) by JPEG. The range of QF is from 1 to 100, which corresponds to the image quality from low to high with the increase of quantization steps. In Fig.2, the horizontal axis indexes quality factor, and the vertical axis is the standard deviation of compression noise in images (denoted as global noise level). The compression noise level not only changes along with quantization steps, but also varies significantly in different images, which verifies that compression noise is content-dependent.

The distribution of DCT coefficients has been widely investigated in literatures, which is able to be well modeled by Gaussian, Laplacian distribution [36] or Generalized Gaussian [13]. Fig.3 shows histogram of DCT coefficients in 8×8 blocks. We can see that Gaussian distribution can well model the coefficient in low frequency, while Laplace distribution can well model high frequency coefficients. Therefore, the Generalized Gaussian Distribution (GGD) with zero mean is suitable for coefficient modelling with the following distribution,

$$GG(x;\beta,\rho) = \frac{\beta^{1/2}}{2\Gamma(1+1/\rho)} e^{-\beta^{\rho/2}|x|^{\rho}}, \tag{19}$$

where ρ is the shape parameter and β is the inverse scale parameter. When ρ is equal to 1, GGD is equivalent to Laplace distribution. When ρ is equal to 2, GGD is equivalent to Gaussian distribution. However, the shape parameter is difficult to determine for different frequency bands, which is also beyond the scope of this paper. Therefore, we take Gaussian distribution (i.e., $\rho = 2$) to model the DCT coefficients in this paper, which is also an good approximation for DCT coefficients, especially for low frequency bands where compression noise is dominant. The proposed compression noise estimation method also can be forwardly applied to other distribution models, and we also show the results with Laplace distribution (i.e., $\rho = 1$). If we take x as a Gaussian signal, and t as the index of quantization interval, or the quantized value, the reconstructed value of x in the t^{th} quantization interval can be written as,

$$y = tq, \quad x \in \left[\frac{(2t-1)q}{2}, \frac{(2t+1)q}{2}\right], \tag{20}$$

where $q = Q(u,v)$ is a scalar quantization step. When the reconstruction value y is known, the distribution of quantization noise (in this paper, we regard the two terms, compression noise and quantization noise, to be equivalent) can be written as,

$$n = x - y \tag{21}$$

$$g(n|y=tq) = \begin{cases} \dfrac{1}{C_t\sqrt{2\pi}\sigma} e^{-\frac{(n+tq)^2}{2\sigma^2}}, & n \in \left[-\dfrac{q}{2}, \dfrac{q}{2}\right] \\ 0, & others, \end{cases} \tag{22}$$

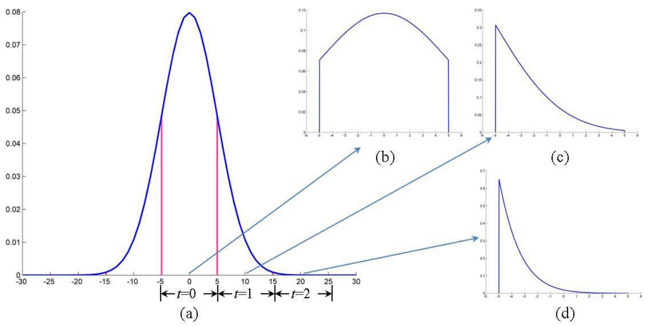


Fig. 4. The distribution of quantization noise in different quantization intervals for signals following Gaussian distribution, (b) the noise distribution for signals in $t = 0$ interval, (c) the noise distribution for signals in $t = 1$ interval, (d) the noise distribution for signals in $t = 2$ interval.

where C_t is the integration of the Gaussian signal,

$$C_t = \int_{\frac{(2t-1)q}{2}}^{\frac{(2t+1)q}{2}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{n^2}{2\sigma^2}} dn. \quad (23)$$

Here, quantization noise n is limited in the range defined by quantization step, $[-\frac{q}{2}, \frac{q}{2}]$. In this range, the quantization noise follows the same distribution with that of DCT coefficients x due to the additive noise assumption with Eqn.(21), while beyond the range its probability is zero. C_t is the normalized parameter to make the integration of the probability density function in the whole range is 1.

Fig.4 illustrates the distribution of quantization noise in different quantization intervals for signals following Gaussian distribution with zero mean. Fig.4(a) shows Gaussian distribution of signal, x, and Fig.4(b)-(d) illustrate the distribution of quantization noise when $t = 0, 1, 2$, respectively. We can see that the quantization noise follows Gaussian distribution in the dead-zone quantization interval (*i.e.*, $t = 0$), while in other intervals its distribution more resembles χ^2 distribution. The standard deviation of quantization noise distribution is dependent on that of original signal distribution and the quantization steps, which will be estimated in the following subsection.

B. Content-Dependent Compression Noise Level Estimation

Based on the above analysis, there are three necessary factors for image compression noise level estimation, *i.e.*, quantized values or reconstructed values (t or y), quantization steps and standard deviations of image coefficients. The first two factors are easily retrieved from the compressed bitstream. However, the third factor is difficult to be derived from compressed images. There are two reasons, 1) image structure varies with different contents, which makes it difficult to formulate image coefficient distribution with a single fixed global standard deviation; 2) quantization with large quantization steps directly removes most coefficients in high frequency bands, which makes it difficult to derive coefficient distribution parameters for every band by directly calculating statistics from reconstructed values.

In order to estimate compression noise level, we propose a content-dependent estimation method for compression noise via patch grouping. First, we also divide image into patches and classify them into different groups according their similarity as that in section III. These similar image patches in the same group can be regarded to share the same distribution parameter, *i.e.*, the standard deviation of coefficient distribution. In this paper, we take the most widely used image spatial correlation model [37], as illustrated in Eqn.(24), to derive the parameters of Gaussian distribution for DCT coefficients (The derivation process is also applicable to other distributions, *e.g.*, Laplace distribution). For a 2D image $\mathcal{I}_{\boldsymbol{y}}$ with size of $H \times W$, the correlation between two pixels x_1 and x_2 at locations (m_1, n_1) and (m_2, n_2) is modeled as,

$$\begin{aligned} r(x_1, x_2) = E(x_1 x_2) = \sigma_{\boldsymbol{x}}^2 \rho_H^{|m_1-m_2|} \rho_V^{|n_1-n_2|}, \\ 0 \le m_1, m_2 \le W, \quad 0 \le n_1, n_2 \le H, \end{aligned} \quad (24)$$

where ρ_H and ρ_V are the correlation coefficients of neighboring pixels in the horizontal direction and vertical direction, respectively. $\sigma_{\boldsymbol{x}}$ is the standard deviation of image signals in spatial domain.

We further derive the coefficient distribution from the image spatial correlation model. For a 2D image patch, $\boldsymbol{x}_{\mathcal{B}}$ with size of $p \times p$, we reorganize it into vector, $\boldsymbol{v}_{\mathcal{B}}$, in row-major order, and the corresponding transform coefficient vector denoted as $\boldsymbol{f}_{\mathcal{B}}$,

$$\begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_{p^2} \end{pmatrix} = \begin{pmatrix} c_{1,1} & \cdots & c_{1,p^2} \\ \vdots & \ddots & \vdots \\ c_{p^2,1} & \cdots & c_{p^2,p^2} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_{p^2} \end{pmatrix}, \quad (25)$$

where $\boldsymbol{f}_{\mathcal{B}} = \{f_i\}$, $\boldsymbol{v}_{\mathcal{B}} = \{v_i\}$, and transform matrix $\mathcal{T} = (c_{i,j})_{p^2 \times p^2}$. The variance of the transform coefficients is,

$$\begin{aligned} \Sigma_{\mathbf{X}}^2(u, v) &= E(f_i^2) = E\{(\sum_{j=1}^{p^2} c_{i,j} v_j)^2\} \\ &= \sum_{(j_1,j_2)\in S} c_{i,j_1} c_{i,j_2} E(v_{j_1} v_{j_2}) \\ &= \sigma_{\boldsymbol{x}}^2 \sum_{(j_1,j_2)\in S} \{c_{i,j_1} c_{i,j_2} \rho_H^{|j_1-j_2|_H} \rho_V^{|j_1-j_2|_V}\}. \end{aligned} \quad (26)$$

Here $|j_1 - j_2|_H$ and $|j_1 - j_2|_V$ represent the absolute difference of pixel coordinates in horizontal direction and vertical direction, respectively. S represents the set of all combinations of the element indices of $\boldsymbol{v}_{\mathcal{B}}$. Herein, the index i of f corresponds to transform band (u, v) of 2D image patch.

Furthermore, in order to estimate the variance of DCT coefficients in different bands, we need to estimate spatial correlation of adjacent pixels. For a group of similar image patches, we estimate the adjacent pixel correlations from the average of these similar patches $\bar{\boldsymbol{y}}_{G_i}$ directly, which is regarded as an initial estimation of the original image patch and corresponds to the *Average filtering* in Fig.1.

$$\rho_H = \frac{\sum_{i,j=1}^{p} (\bar{\boldsymbol{y}}_{G_i}(i+1, j) - \mu)(\bar{\boldsymbol{y}}_{G_i}(i, j) - \mu)}{\sum_{i,j=1}^{p} (\bar{\boldsymbol{y}}_{G_i}(i, j) - \mu)^2}, \quad (27)$$

$$\rho_V = \frac{\sum_{i,j=1}^{p} (\bar{\boldsymbol{y}}_{G_i}(i, j+1) - \mu)(\bar{\boldsymbol{y}}_{G_i}(i, j) - \mu)}{\sum_{i,j=1}^{p} (\bar{\boldsymbol{y}}_{G_i}(i, j) - \mu)^2}, \quad (28)$$

$$\bar{\boldsymbol{y}}_{G_i}(i, j) = \frac{1}{K} \sum_{k=1}^{K} \boldsymbol{y}_{G_i}^{(k)}(i, j), \quad \mu = \frac{1}{p^2} \sum_{i,j=1}^{p} \bar{\boldsymbol{y}}_{G_i}(i, j), \quad (29)$$

TABLE II
STANDARD DEVIATION OF DCT COEFFICIENTS WHEN SPATIAL CORRELATION $\rho_H = 0.4$, $\rho_V = 0.9$ AND $\sigma_{\boldsymbol{x}} = 1$

8.087	6.312	4.876	3.601	2.744	2.213	1.901	1.737
2.068	1.614	1.247	0.921	0.702	0.566	0.486	0.444
0.712	0.555	0.429	0.317	0.241	0.195	0.167	0.153
0.341	0.266	0.206	0.152	0.116	0.093	0.080	0.073
0.215	0.168	0.130	0.096	0.073	0.059	0.051	0.046
0.156	0.122	0.094	0.069	0.053	0.043	0.037	0.033
0.127	0.099	0.076	0.056	0.043	0.035	0.030	0.027
0.112	0.088	0.068	0.050	0.038	0.031	0.026	0.024

where $\boldsymbol{y}_{G_i}^{(k)}$ is the k^{th} image patches in group G_i. In addition, the standard deviation $\sigma_{\boldsymbol{x},G_i}$ of the corresponding original image patches in spatial domain also can be calculated from the initial estimation, $\bar{\boldsymbol{y}}_{G_i}$,

$$\sigma_{\boldsymbol{x},G_i} = \alpha \sqrt{\sum_{i,j=1}^{p} (\bar{\boldsymbol{y}}_{G_i}(i,j) - \mu)^2}, \tag{30}$$

where α is a scale factor.

Based on Eqn. (26), we can derive standard deviation $\boldsymbol{\Sigma}_{\mathbf{X}}(u,v)$ of coefficients in band (u,v). An example of $\boldsymbol{\Sigma}_{\mathbf{X}}(u,v)$ for 8×8 DCT block is illustrated in TABLE II when $\rho_H = 0.4$, $\rho_V = 0.9$ and $\sigma_{\boldsymbol{x}} = 1$. In the example, the derived standard deviations along horizontal direction are larger than that along vertical direction, which implicitly shows that there is relative higher probability for large coefficients along horizontal direction than that along vertical direction.

Based on the derived standard deviation of coefficients and coefficient prior model, we can estimate the compression noise level in each band for every group as follows.

$$\boldsymbol{\Sigma}_{n,G_i}(u,v) = \sqrt{\int_{-\frac{q}{2}}^{\frac{q}{2}} (n - \mu_n(u,v))^2 g_{u,v,G_i}(n|y = tq)dn}, \tag{31}$$

$$\mu_n(u,v) = \int_{-\frac{q}{2}}^{\frac{q}{2}} n g_{u,v,G_i}(n|y = tq)dn, \tag{32}$$

where $q = Q(u,v)$ is the quantization step for band (u,v). g_{u,v,G_i} is the distribution function for DCT coefficients in band (u,v) of group G_i in Eqn.(22) with the standard deviation, $\boldsymbol{\Sigma}_{n,G_i}(u,v)$. Therefore, the compression noise level for group G_i in Eqn.(13) is calculated by averaging noise standard deviations in all the bands, which is shown as follows,

$$\sigma_{n,G_i} = \sum_{u,v=1}^{p} w_{G_i}(u,v)\boldsymbol{\Sigma}_{n,G_i}(u,v), \tag{33}$$

$$w_{G_i}(u,v) = \boldsymbol{\Sigma}_{n,G_i}(u,v)/S, \;\; S = \sum_{u,v=1}^{p} \boldsymbol{\Sigma}_{n,G_i}(u,v). \tag{34}$$

V. ITERATIVE IMPLEMENTATION FOR CONTENT-DEPENDENT COMPRESSION NOISE REDUCTION

In order to further improve the performance of the compression noise reduction, we will introduce the iterative implement for the proposed compression noise reduction and noise level update procedure.

Algorithm 1 Compression Noise Estimation

Input: Compressed image, $\mathcal{I}_{\boldsymbol{y}}$.
Patch Clustering: Classify image patches into N_g groups according to their similarity measure in Eqn. (3);
for $i = 1$ *to* N_g **do**
  Initialization: average of image patches in group G_i, generate $\bar{\boldsymbol{y}}_{G_i}$;
  Correlation estimation:signal spatial correlation calculation from $\bar{\boldsymbol{y}}_{G_i}$ with Eqn.(27) and (28);
  Coefficient variance: calculate the variance of transform coefficients based on Eqn.(26);
  Noise estimation: calculate the standard deviation of compression noise via Eqns.(31)$\sim$ (33);
end
Output: Standard deviation of compression noise for all image patch groups, $\{\sigma_{n,G_i}\}$.

Algorithm 2 Compression Noise Reduction

Input: Compressed image, $\mathcal{I}_{\boldsymbol{y}}$.
Patch Clustering: Classify image patches into N_g groups according to their similarity measure in Eqn. (3);
Estimate compression noise level with **Algorithm 1**;
while *stop condition* **do**
  for $i = 1$ *to* N_g **do**
    Singular value decomposition on image patch group $\mathbf{Y}_{G_i}$;
    Soft-Thresholding the singular values of $\mathbf{Y}_{G_i}$ via Eqn.(12) and (13);
    Image patch reconstruction with shrunken singular values with last equation in Eqn.(9);
  end
  Image reconstruction with weighted average in Eqn.(15);
  Apply narrow quantization constraint to estimated coefficients with Eqn.(18);
  Update image noise with the Eqn.(35);
end
Output: Restored high quality image, $\mathcal{I}_{\hat{\boldsymbol{x}}}$

Since the standard deviation of image signals is estimated from an initial estimation, it is not so accurate. Therefore, we take an iterative implementation for removing compression noise, and update the compression noise level after each iteration. When the whole image is reconstructed, the latest reconstructed image is utilized as input to the next iteration with the updated compression noise levels as follows,

$$\sigma_{n,G_i}^{(k)} = max\left(0, \sigma_{n,G_i}^{(0)} - std\left(\widehat{\mathbf{X}}_{G_i}^{(k-1)} - \mathbf{Y}_{G_i}\right)\right), \tag{35}$$

where $\sigma_{n,G_i}^{(0)}$ is estimated from the decoded image via Eqn.(33) and $\widehat{\mathbf{X}}_{G_i}^{(k)}$ is the image patch matrix extracted from the reconstructed image after the k^{th} iteration. $std(\cdot)$ is the function to calculate standard deviation. $\sigma_{n,G_i}^{(k)}$ is the standard deviation of compression noise of image patches in group G_i

TABLE III
PSNR RESULTS OF RESTORED JPEG IMAGES AT QF = 10 (PSNR: dB)

Images	JPEG	BM3D	CSR	FoE	PSW	ReJPEG	Ren's	Proposed-G	Proposed-L
Barbara	26.34	27.34	27.50	27.09	26.90	27.13	27.67	29.05	29.01
Lena	30.41	31.40	31.65	31.66	30.98	31.48	31.79	32.24	32.25
Pepper	30.14	31.11	31.36	31.57	30.83	31.21	31.46	31.92	31.92
House	30.32	31.29	31.58	31.89	31.25	31.44	31.65	32.50	32.49
Bike	26.63	27.82	28.02	27.48	27.14	27.28	28.06	28.83	28.81
Flower	31.78	32.93	33.33	33.68	32.43	33.33	33.46	34.05	34.07
Gantrycrane	27.22	28.21	28.50	28.19	27.55	27.97	28.59	29.39	29.38
Girl	30.69	31.90	32.17	31.94	31.01	31.97	32.31	32.79	32.79
Hat	30.64	31.41	31.61	31.71	31.24	31.36	31.70	32.13	32.13
Window	29.73	30.76	31.05	30.95	30.02	30.70	31.17	31.47	31.47
Sailboats2	29.67	30.53	30.75	30.68	30.00	30.41	30.83	31.31	31.31
Parrot	31.73	32.45	32.70	33.09	32.40	32.72	32.81	33.56	33.56
Monarch	29.47	30.49	30.82	30.90	30.12	30.57	30.96	31.44	31.44
Car	29.70	30.70	30.97	30.82	30.23	30.66	31.12	31.82	31.83
Building	30.11	31.54	31.93	31.77	30.77	31.76	32.04	32.88	32.89
Pens	30.41	31.83	32.14	31.93	31.23	31.87	32.25	32.59	32.61
Average	29.69	30.73	31.00	30.96	30.26	30.74	31.12	31.75	31.75

TABLE IV
SSIM RESULTS OF RESTORED JPEG IMAGES AT QF = 10

Images	JPEG	BM3D	CSR	FoE	PSW	ReJPEG	Ren's	Proposed-G	Proposed-L
Barbara	0.781	0.817	0.817	0.810	0.798	0.802	0.821	0.853	0.852
Lena	0.818	0.849	0.853	0.857	0.840	0.853	0.859	0.867	0.868
Pepper	0.784	0.818	0.824	0.832	0.813	0.823	0.828	0.837	0.836
House	0.808	0.832	0.837	0.849	0.835	0.841	0.839	0.850	0.850
Bike	0.764	0.807	0.808	0.799	0.783	0.794	0.808	0.822	0.822
Flower	0.856	0.891	0.900	0.913	0.891	0.902	0.904	0.918	0.918
Gantrycrane	0.862	0.899	0.904	0.906	0.875	0.891	0.910	0.923	0.922
Girl	0.809	0.850	0.855	0.843	0.819	0.849	0.859	0.869	0.869
Hat	0.821	0.847	0.851	0.859	0.848	0.849	0.856	0.867	0.867
Window	0.854	0.888	0.893	0.895	0.867	0.889	0.898	0.905	0.905
Sailboats2	0.806	0.838	0.842	0.841	0.819	0.837	0.846	0.854	0.854
Parrot	0.850	0.877	0.883	0.898	0.885	0.887	0.888	0.901	0.901
Monarch	0.870	0.905	0.911	0.922	0.903	0.910	0.916	0.926	0.926
Car	0.837	0.865	0.869	0.871	0.855	0.866	0.874	0.885	0.885
Building	0.871	0.912	0.919	0.918	0.897	0.914	0.922	0.934	0.934
Pens	0.826	0.870	0.874	0.868	0.849	0.869	0.877	0.881	0.882
Average	0.826	0.860	0.865	0.868	0.849	0.861	0.869	0.881	0.881

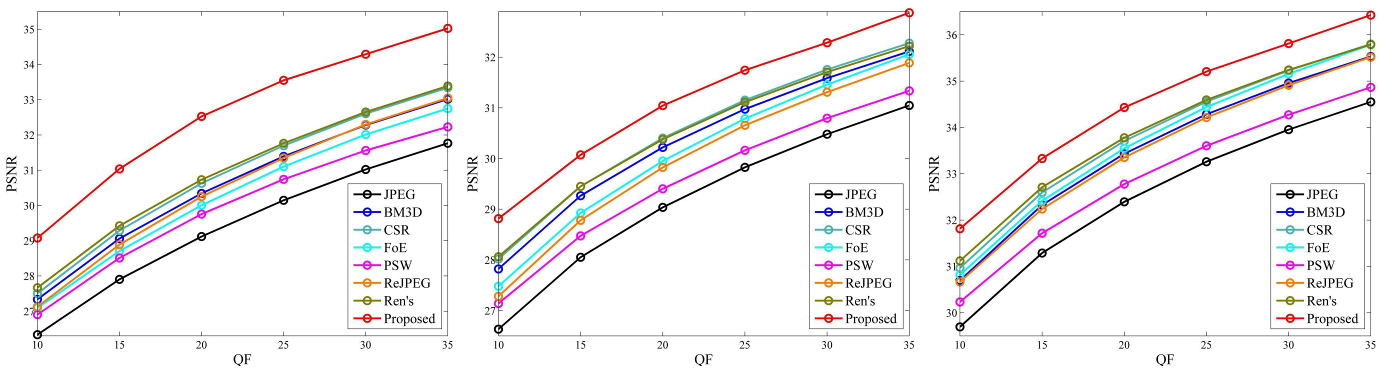
at the k^{th} iteration. Finally, the proposed content-dependent compression noise estimation and reduction algorithms are described in **Algorithm 1** and **Algorithm 2**, respectively. In our method, we take the mean absolute difference of the restored images of two consecutive iterations to decide the termination of **Algorithm 2**, which is denoted as $MAD^{(k)}$ after the k^{th} iteration. When $MAD^{(k)}$ is smaller than 0.08, **Algorithm 2** will stop.

VI. EXPERIMENTS AND ANALYSIS

In this section, we test the proposed method on JPEG images, which is the most widely used image compression format. The test images used in our experiments include popular images, *e.g.*, *Barbara*, *Lena*, *Peppers* and some standard images in Kodak. These test images are high quality with lossless compression, which have been updated to Google Drive[1] to be connivent for users viewing them. These high quality images are compressed by JPEG codec[2] at different quality factors (QF), and then restored with different denoising methods. We first verify the performance of the proposed compression reduction method by comparing it with other state-of-the-art denoising methods according to the quality of restored images. Furthermore, we also verify the usage of the proposed method in computer vision field based on the performance of some basic operations in computer vision applications, *i.e.*, edge detection and image segmentation. Finally, we analyze the complexity of the proposed method. The main parameters are predefined in our method as, patch

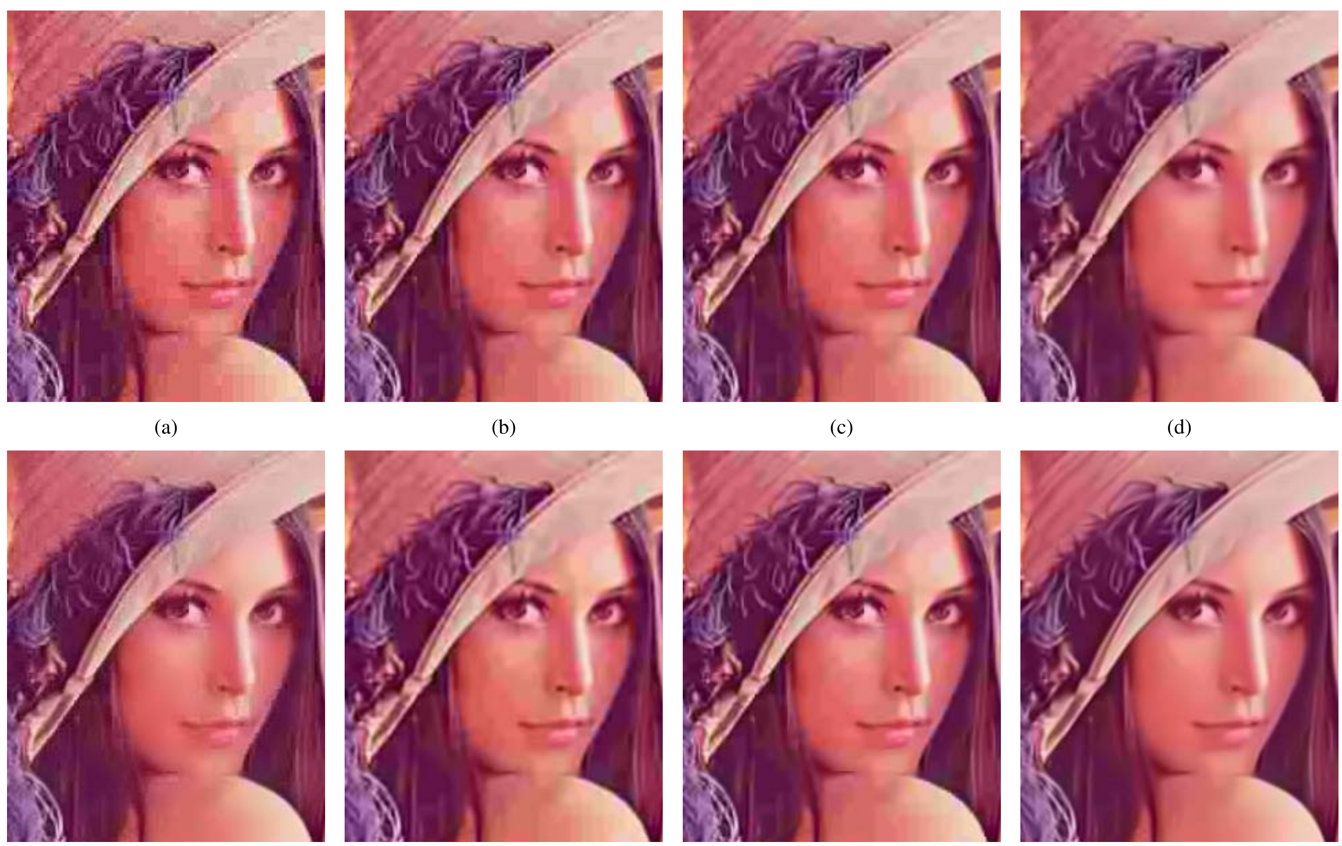
[1]https://drive.google.com/folderview?id=0B09lxMhXreF1UDJEZy1rLUlHQU0&usp=sharing

[2]http://www.ijg.org/



(a) (b) (c)

Fig. 5. The PSNR result comparison among different denoising methods for images compressed at different QFs. (a) *Barbara*. (b) *Bike*. (c) *Car*.



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

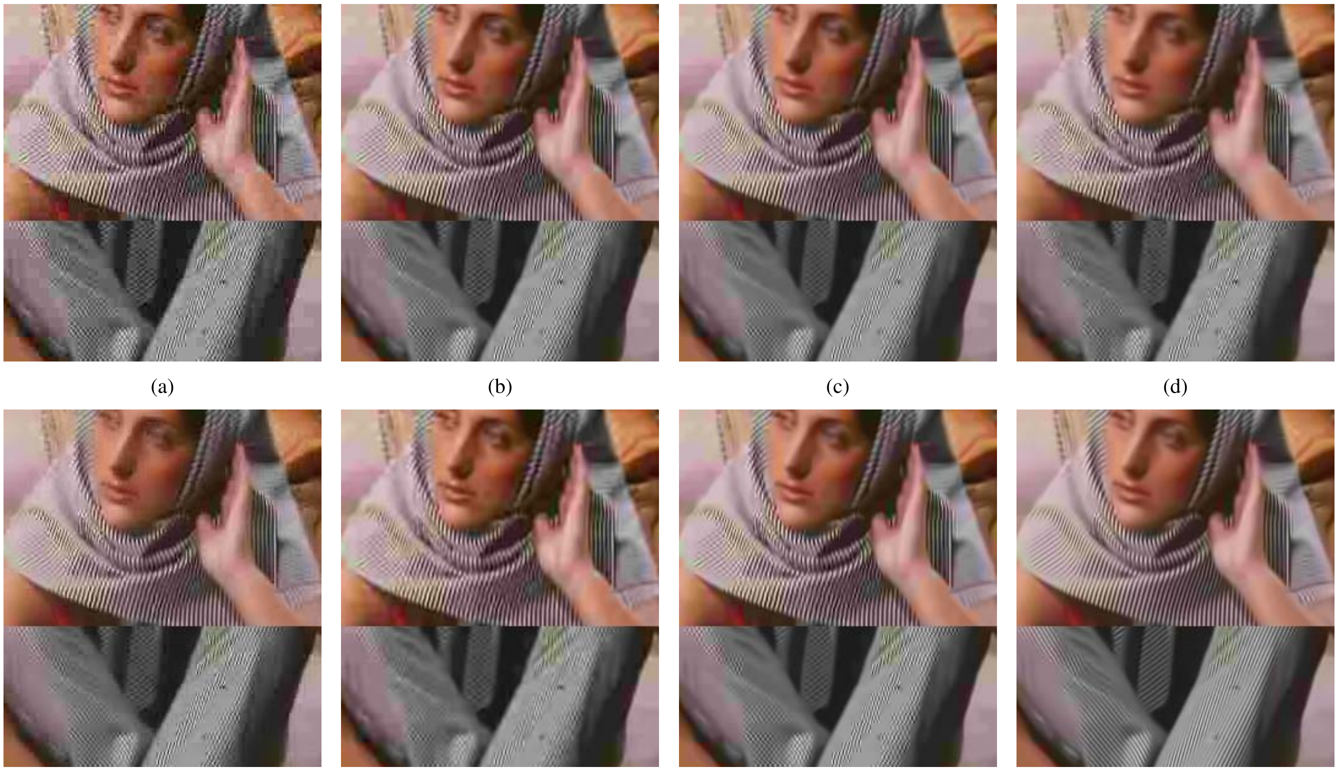
(e) (f) (g) (h)

Fig. 6. The restored JPEG image, *Lena*, at QF = 10. (a) JPEG. (b) BM3D. (c) CSR. (d) FoE. (e) PSW. (f) ReJPEG. (g) Ren's. (h) Proposed.

size $p = 8$, overlapped step $s = 5$, $K = 40$ in Eqn.(4), $\gamma = 2\sqrt{2}$ in Eqn.(13), $c = 0.7$ in Eqn.(18), and $\alpha = 8$ in Eqn.(30).

A. Performance Comparison on Compression Noise Reduction

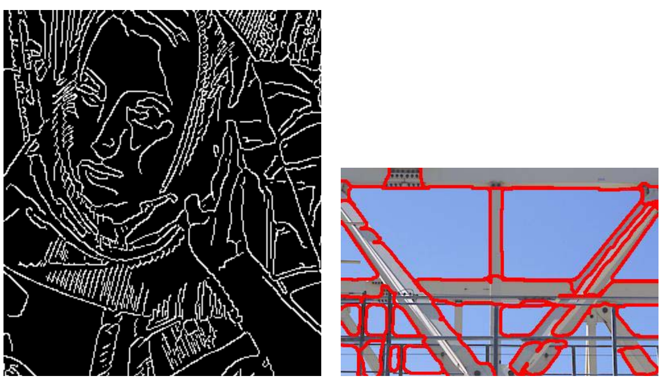
In this subsection, we compare the proposed compression noise reduction method with state-of-the-art denoising methods, including BM3D [3], CSR [4], FoE [11], PSW [17], ReJPEG [38], and Ren's method [14]. Although BM3D and CSR are general denoising methods, they also use similar patch grouping and thresholding, which are similar with ours. The others are denoising methods specially for removing compression noise. For the compared methods, BM3D, CSR and Ren's methods need a global parameter, *i.e.*, standard deviation of compression noise. In order to get the best performance of the compared methods, we utilize the original image to calculate standard deviation of compression noise for them. In practice, the performance of the compared methods may be lower than that with actual standard deviation of compression noise.



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

(e) (f) (g) (h)

Fig. 7. The restored JPEG image, *Barbara*, at QF = 10. (a) JPEG. (b) BM3D. (c) CSR. (d) FoE. (e) PSW. (f) ReJPEG. (g) Ren's. (h) Proposed.



(a) (b)

Fig. 8. (a) Edge detection result on original image, *Barbara*. (b) Segmentation result on original image *Gantrycrane*.

In the proposed algorithm, we take the Gaussian distribution to model the DCT coefficients. In order to verify its efficiency, two widely used distributions for DCT coefficients are compared, i.e., Gaussian distribution and Laplace distribution. From the results in Table III, the proposed method with the two distributions (denoted as Proposed-G and Proposed-L for using Gaussian distribution and Laplace distribution) achieves very approximate results, which shows that the two distributions can well model the DCT coefficients. Although Gaussian distribution may be not the perfect model for DCT coefficients in different bands, it also can well approximate the distributions of DCT coefficients, especially for that in low frequency bands. In addition, considering that the compression noise mainly exists in low frequency bands when using the same quantization steps, the Gaussian distribution is reasonable in modelling DCT coefficients in the proposed method. Therefore, in the following results, we only show the results of the proposed method with Gaussian distribution model.

Based on the PSNR results of the restored images with different denoising methods in Table III, the proposed method significantly improves the quality of decoded JPEG images, and achieves up to 2.06 dB gain over JPEG decoder on average. Especially, the proposed method achieves up to 2.7 dB gain for images, *Barabra*, which have more patterns with similar structure and are more suitable for low-rank approximation than other images. The proposed method also outperforms other denoising methods, and achieves about 0.43$\sim$1.49 dB gain on average. To better show the perceptual quality of the restored images, we also show the results with another widely used quality metric, Structural Similarity Index Metric (SSIM) [39], to further verify the superiority of the proposed method. Table IV shows the corresponding SSIM results for JPEG images compressed at QF=10, and our method achieves about 0.012$\sim$0.055 compared with JPEG and other denoising methods. We also compares the performance of NQCS with normal quantization constraint, and NQCS achieves about 0.32 dB gain for these images in Table III.

Fig.5 illustrates the performance of different methods on a large bitrate range, *i.e.*, for images compressed at different QFs. The proposed method works well over a wide

TABLE V

F-Measure Results for Edge Detection on Restored JPEG Images at QF=10

Images	JPEG	BM3D	CSR	FoE	PSW	ReJPEG	Ren's	Proposed
Barbara	0.666	0.710	0.714	0.679	0.669	0.683	0.719	0.731
Lena	0.549	0.591	0.597	0.586	0.567	0.584	0.607	0.629
Pepper	0.491	0.545	0.563	0.554	0.523	0.554	0.549	0.579
House	0.589	0.646	0.665	0.662	0.632	0.659	0.652	0.700
Bike	0.690	0.730	0.728	0.704	0.701	0.706	0.731	0.742
Flower	0.489	0.517	0.535	0.559	0.527	0.540	0.550	0.578
Gantrycrane	0.783	0.808	0.803	0.790	0.763	0.785	0.810	0.811
Girl	0.550	0.585	0.589	0.559	0.543	0.578	0.598	0.603
Hat	0.422	0.447	0.453	0.483	0.470	0.447	0.474	0.502
Window	0.659	0.685	0.694	0.680	0.634	0.680	0.697	0.707
Sailboats2	0.532	0.582	0.595	0.571	0.546	0.563	0.603	0.610
Parrot	0.367	0.401	0.405	0.444	0.424	0.398	0.418	0.472
Monarch	0.578	0.645	0.628	0.644	0.615	0.644	0.649	0.655
Car	0.631	0.667	0.679	0.658	0.648	0.660	0.687	0.709
Building	0.777	0.796	0.800	0.781	0.769	0.786	0.804	0.818
Pens	0.538	0.553	0.553	0.513	0.517	0.533	0.554	0.555
Average	0.582	0.619	0.625	0.617	0.597	0.613	0.631	0.650

(a) JPEG (b) BM3D (c) CSR (d) FoE

(e) PSW (f) ReJPEG (g) Ren's (h) Proposed

Fig. 9. Edge detection results using Canny operator on image, *Barbara*, (a) edge detection result on JPEG image compressed at QF=10, (b)~(h) edge detection result on restored image with methods, BM3D, CSR, FoE, PSW, ReJPEG, Ren's method and the proposed method orderly.

quality or bitrate range. Fig.6 and Fig.7 show the subjective quality of restored images, *Lena* and *Barabra*, with different methods respectively. We can see that the proposed method reconstructs more visual pleased results by removing most of the compression artifacts obviously. Especially, at the face area in *Lena* and scarf area in *Barabra*, the blocking and ringing artifacts are removed much cleaner than that in compared methods.

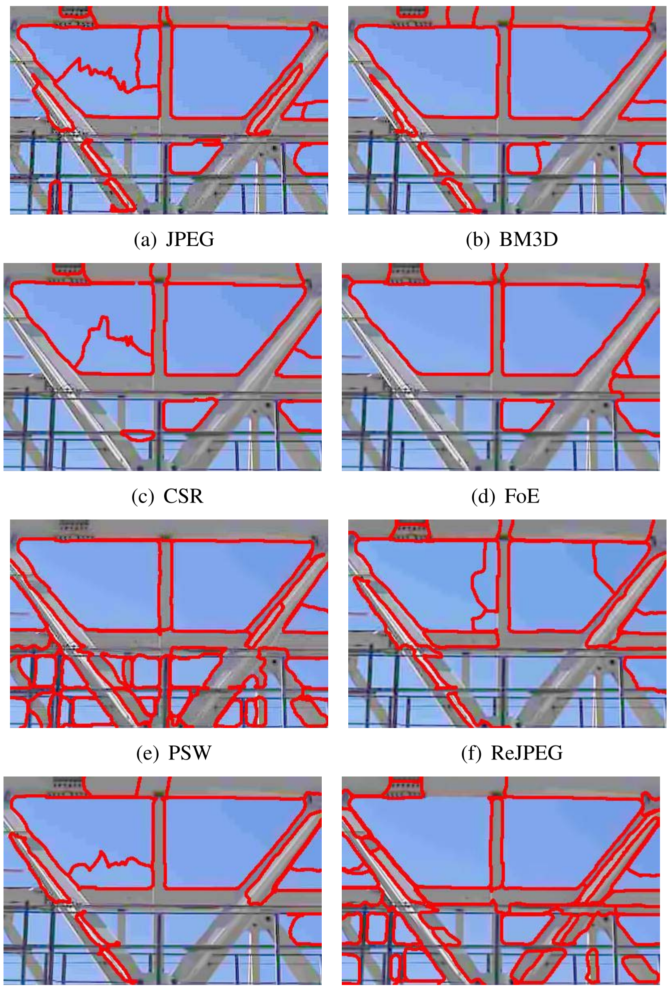
B. Application in Basic Computer Vision Tasks

Besides improving image quality, our method may help improve the performance of basic computer vision tasks when it is utilized as a pre-processing method. As an initial exploration, we carry out two basic computer vision operations, *i.e.*, edge detection and image segmentation, on images reconstructed by JPEG decoder and denoising methods. Fig.9 shows the results of edge detector, Canny operator [40],

TABLE VI
RESULTS FOR SEGMENTATION ON RESTORED JPEG IMAGES AT QF=10

	GCE	PRI	VoI
JPEG	0.215	0.890	1.865
BM3D	0.204	0.860	1.961
CSR	0.206	0.874	1.892
FoE	0.195	0.861	1.910
PSW	0.190	0.863	1.935
ReJPEG	0.211	0.848	1.946
Ren's	0.208	0.857	1.980
Proposed	0.198	0.906	1.643

*GCE ranges between [0,1], the lower value is better, PRI ranges between [0,1], the higher vale is better and VoI ranges between $[0, \infty)$, the lower value is better.



(a) JPEG (b) BM3D (c) CSR (d) FoE (e) PSW (f) ReJPEG (g) Ren's (h) Proposed

Fig. 10. Performance of image segmentation operation on image, *Gantrycrane*, (a) segmentation result on JPEG image compressed at QF=10, (b)~(h) segmentation result on restored image with methods, BM3D, CSR, FoE, PSW, ReJPEG, Ren's method and the proposed method orderly.

on JPEG image at QF=10 and the restored images by denoising methods. We can see that the edges detected from the restored image by our method more approach with that detected from original image in Fig.8(a), while lots of false edges are detected from JPEG image due to severe

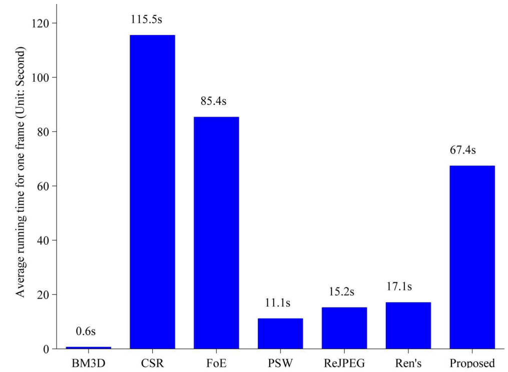


Fig. 11. The average running time of different methods on images with size of 256 × 256.

compression noise. Table V shows the F-measure results, which take the detection results on original image as anchors. The F-measure is calculated with the following equation,

$$F_1 = 2\frac{precision \cdot recall}{precision + recall}. \quad (36)$$

Based on the results, more true edges are detected from the restored image by our proposed method and the best F-measure results are achieved compared with other methods.

Fig.10 shows image segmentation results of method in [41] for JPEG image at QF=10 and the corresponding restored images by denoising methods, respectively. We can see that the image segmentation method can efficiently separate different contents from the image restored by our method as that from original image in Fig.8(b), while its performance is degraded fast on the JPEG image due to the negative effects of compression noise. Table VI shows the results with three objective metrics for image segmentation, *i.e.*, Global Consistency Error (GCE) [42], Variation of Information (VoI) [43], and Probabilistic Rand Index (PRI) [44], which take the segmentation results on original images as anchors. The segmentation results on the images restored by our method achieves the best results according to all the metrics, which shows that our proposed compression noise reduction method can effectively improve the performance of the computer vision algorithm.

C. Complexity Analysis

To evaluate the computational cost of the proposed method, we compare the running time of different methods on three images with size of 256 × 256 compressed at QF = 10. We calculate the average running time of different methods for these test images with MATLAB 2012, Intel (R) Core (TM) i5-4570@3.20GHz, and 64bit Windows 7 operating system. Except for BM3D being optimized with C++, our method and other comparison methods are all implemented only with MATLAB. Fig.11 shows the average running time for different methods. Compared with other methods implemented with MATLAB, our method needs about 67.4s to process one image on average, which is significantly faster than CSR (115.5s) and FoE (85.4s). In addition, our method can be further speeded up by processing every image patch group in parallel and implementing with C++. Therefore, it is possible to apply the proposed method for practical applications even with the real-time requirement.

VII. Conclusion

In this paper, we have proposed a content-dependent compression noise level estimation and reduction framework via similar patch clustering and low-rank constraint. The compression noise is estimated based on quantization steps, and image prior models, *i.e.*, a transform coefficient prior model and an image spatial correlation model. The compression noise is removed by soft-thresholding the signular values of similar image patch matrices adaptively according to their noise levels instead of a global noise level. Extensive experimental results have verified that the proposed method not only significantly improves the quality of compressed images against the relevant existing works, but also benefits computer vision tasks by removing compression noise.

Acknowledgment

The authors would like to thank Prof. Jiaying Liu of Peking University for her help in providing source code about their related work. They would like to also thank the Associate Editor and the three anonymous reviewers for their helpful comments and suggestions.

References

[1] S. H. Park and D. S. Kim, "Theory of projection onto the narrow quantization constraint set and its application," *IEEE Trans. Image Process.*, vol. 8, no. 10, pp. 1361–1373, Oct. 1999.
[2] G. K. Wallace, "The JPEG still picture compression standard," *IEEE Trans. Consum. Electron.*, vol. 38, no. 1, pp. 18–34, Feb. 1992.
[3] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.
[4] W. Dong, X. Li, D. Zhang, and G. Shi, "Sparsity-based image denoising via dictionary learning and structural clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2011, pp. 457–464.
[5] I. H. Reeve, III, and J. S. Lim, "Reduction of blocking effect in image coding," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 8. Apr. 1983, pp. 1212–1215.
[6] B. Ramamurthi and A. Gersho, "Nonlinear space-variant postprocessing of block coded images," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, no. 5, pp. 1258–1268, Oct. 1986.
[7] S. Minami and A. Zakhor, "An optimization approach for removing blocking effects in transform coding," in *Proc. Data Compress. Conf. (DCC)*, Apr. 1991, p. 442.
[8] A. Norkin *et al.*, "HEVC deblocking filter," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1746–1754, Dec. 2012.
[9] X. Zhang, R. Xiong, S. Ma, and W. Gao, "Adaptive loop filter with temporal prediction," in *Proc. Picture Coding Symp. (PCS)*, May 2012, pp. 437–440.
[10] C.-Y. Tsai *et al.*, "Adaptive loop filtering for video coding," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 6, pp. 934–945, Dec. 2013.
[11] D. Sun and W.-K. Cham, "Postprocessing of low bit-rate block DCT coded images based on a fields of experts prior," *IEEE Trans. Image Process.*, vol. 16, no. 11, pp. 2743–2751, Nov. 2007.
[12] X. Zhang, R. Xiong, S. Ma, and W. Gao, "Reducing blocking artifacts in compressed images via transform-domain non-local coefficients estimation," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2012, pp. 836–841.
[13] X. Zhang, R. Xiong, X. Fan, S. Ma, and W. Gao, "Compression artifact reduction by overlapped-block transform coefficient estimation with block similarity," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 4613–4626, Dec. 2013.
[14] J. Ren, J. Liu, M. Li, W. Bai, and Z. Guo, "Image blocking artifacts reduction via patch clustering and low-rank minimization," in *Proc. Data Compress. Conf. (DCC)*, Mar. 2013, p. 516.
[15] H. Chang, M. K. Ng, and T. Zeng, "Reducing artifacts in JPEG decompression via a learned dictionary," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 718–728, Feb. 2014.
[16] X. Liu, X. Wu, J. Zhou, and D. Zhao, "Data-driven sparsity-based restoration of JPEG-compressed images in dual transform-pixel domain," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 5171–5178.
[17] G. Zhai, W. Zhang, X. Yang, W. Lin, and Y. Xu, "Efficient image deblocking based on postfiltering in shifted windows," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 1, pp. 122–126, Jan. 2008.
[18] X. Zhang, R. Xiong, S. Ma, and W. Gao, "Artifact reduction of compressed video via three-dimensional adaptive estimation of transform coefficients," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 4567–4571.
[19] X. Zhang, R. Xiong, W. Lin, S. Ma, J. Liu, and W. Gao, "Video compression artifact reduction via spatio-temporal multi-hypothesis prediction," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 6048–6061, Dec. 2015.
[20] X. Zhang, W. Lin, J. Liu, and S. Ma, "Compression noise estimation and reduction via patch clustering," in *Proc. Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA)*, Dec. 2015, pp. 715–718.
[21] M. Maggioni, G. Boracchi, A. Foi, and K. Egiazarian, "Video denoising, deblocking, and enhancement through separable 4-D nonlocal spatiotemporal transforms," *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 3952–3966, Sep. 2012.
[22] X. Liu, X. Wu, J. Zhou, and D. Zhao, "Data-driven soft decoding of compressed images in dual transform-pixel domain," *IEEE Trans. Image Process.*, vol. 25, no. 4, pp. 1649–1659, Apr. 2016.
[23] S. I. Olsen, "Estimation of noise in images: An evaluation," *CVGIP: Graph. Models Image Process.*, vol. 55, no. 4, pp. 319–323, Jul. 1993.
[24] J.-S. Lee, "Refined filtering of image noise using local statistics," *Comput. Graph. Image Process.*, vol. 15, no. 4, pp. 380–389, Apr. 1981.
[25] C. Sutour, C.-A. Deledalle, and J.-F. Aujol, "Estimation of the noise level function based on a nonparametric detection of homogeneous image regions," *SIAM J. Imag. Sci.*, vol. 8, no. 4, pp. 2622–2661, 2015.
[26] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 613–627, May 1995.
[27] X. Liu, M. Tanaka, and M. Okutomi, "Single-image noise level estimation for blind denoising," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 5226–5237, Dec. 2013.
[28] N. N. Ponomarenko, V. V. Lukin, K. O. Egiazarian, and J. T. Astola, "A method for blind estimation of spatially correlated noise characteristics," *Proc. SPIE*, pp. 753208-1–753208-12, Feb. 2010.
[29] M. Colom, M. Lebrun, A. Buades, and J. M. Morel, "A non-parametric approach for the estimation of intensity-frequency dependent noise," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 4261–4265.
[30] M. Colom, M. Lebrun, A. Buades, and J.-M. Morel, "Nonparametric multiscale blind estimation of intensity-frequency-dependent noise," *IEEE Trans. Image Process.*, vol. 24, no. 10, pp. 3162–3175, Oct. 2015.
[31] M. Lebrun, M. Colom, and J.-M. Morel, "Multiscale image blind denoising," *IEEE Trans. Image Process.*, vol. 24, no. 10, pp. 3149–3161, Oct. 2015.
[32] C.-J. Hsieh and P. Olsen, "Nuclear norm minimization via active subspace selection," in *Proc. 31st Int. Conf. Mach. Learn. (ICML)*, 2014, pp. 575–583.
[33] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Found. Comput. Math.*, vol. 9, no. 6, pp. 717–772, Apr. 2009.
[34] S. G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Trans. Image Process.*, vol. 9, no. 9, pp. 1532–1546, Sep. 2000.
[35] K. Mohan and M. Fazel, "Iterative reweighted algorithms for matrix rank minimization," *J. Mach. Learn. Res.*, vol. 13, no. 1, pp. 3441–3473, Nov. 2012.
[36] E. Y. Lam, "A mathematical analysis of the DCT coefficient distributions for images," *IEEE Trans. Image Process.*, vol. 9, no. 10, pp. 1661–1666, Oct. 2000.
[37] M. A. Robertson and R. L. Stevenson, "DCT quantization noise in compressed images," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 1, pp. 27–38, Jan. 2005.
[38] A. Nosratinia, "Enhancement of JPEG-compressed images by re-application of JPEG," *J. VLSI Signal Process. Syst. Signal, Image Video Technol.*, vol. 27, no. 1, pp. 69–79, 2001.
[39] 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.

[40] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986.

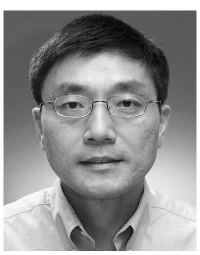
[41] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *Int. J. Comput. Vis.*, vol. 59, no. 2, pp. 167–181, Sep. 2004.

[42] 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. 8th IEEE Int. Conf. Comput. Vis. (ICCV)*, vol. 2. Jul. 2001, pp. 416–423.

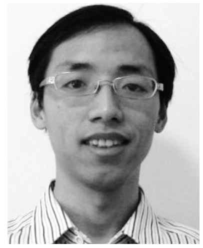
[43] M. Meilă, "Comparing clusterings: An axiomatic view," in *Proc. 22nd Int. Conf. Mach. Learn.*, 2005, pp. 577–584.

[44] C. Pantofaru and M. Hebert, "A comparison of image segmentation algorithms," Robot. Inst., Carnegie Mellon Univ., Pittsburgh, PA, USA, Tech. Rep. CMU-RI-TR-05-40, 2005.

Xinfeng Zhang (M'16) received the B.S. degree in computer science from Hebei University of Technology, Tianjin, China, in 2007, and the Ph.D. degree in computer science from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2014. He is currently a Research Fellow with Nanyang Technological University, Singapore. His research interests include image and video processing, and image and video compression.



Weisi Lin (M'92–SM'98) received the B.Sc. degree in electronics and the M.Sc. degree in digital signal processing from Zhongshan University, Guangzhou, China, in 1982 and 1985, respectively, and the Ph.D. degree in computer vision from Kings College, London University, London, U.K., in 1992. He was involved in teaching and research with Zhongshan University, Shantou University, Shantou, China, Bath University, Bath, U.K., the National University of Singapore, the Institute of Microelectronics, Singapore, and the Institute for Infocomm Research, Singapore. He was the Laboratory Head of Visual Processing and the Acting Department Manager of Media Processing, Institute for Infocomm Research. He is currently an Associate Professor with the School of Computer Engineering, Nanyang Technological University, Singapore. His current research interests include image processing, perceptual modeling, video compression, multimedia communication, and computer vision.



Ruiqin Xiong (M'08) received the B.S. degree from the University of Science and Technology of China, Hefei, China, in 2001, and the Ph.D. degree from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2007. He was a Research Intern with Microsoft Research Asia from 2002 to 2007 and a Senior Research Associate with the University of New South Wales, Australia, from 2007 to 2009. He joined Peking University, Beijing, in 2010. His research interests include statistical image modeling and sparse representation, image and video processing, video compression, and multimedia communication.

Xianming Liu (M'12) received the B.S., M.S., and Ph.D. degrees from the Harbin Institute of Technology (HIT), Harbin, China, in 2006, 2008, and 2012, respectively. In 2011, he spent half a year with the Department of Electrical and Computer Engineering, McMaster University, Canada, as a Visiting Student, where he then worked as a Post-Doctoral Fellow from 2012 to 2013.

He has been a Project Researcher with National Institute of Informatics, Tokyo, Japan, since 2014. He is currently an Associate Professor with the Department of Computer Science, HIT.

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 held the post-doctoral position with the University of Southern California. He joined the Institute of Digital Media, EECS, Peking University, where he is currently an Associate Professor. He has authored over 100 technical articles in refereed journals and proceedings in the areas of image and video coding, video processing, video streaming, and transmission.

Wen Gao (M'92–SM'05–F'09) received the Ph.D. degree in electronics engineering from the University of Tokyo, Tokyo, Japan, in 1991. He was a Professor in computer science with the Harbin Institute of Technology, Harbin, China, from 1991 to 1995, and a Professor with the Institute of Computing Technology, Chinese Academy of Sciences, Beijing. He is currently a Professor in computer science with the Institute of Digital Media, School of Electronic Engineering and Computer Science, Peking University, Beijing, China. He has authored extensively including 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 interfaces, and bioinformatics.