IEEE Access
The journal for rapid open access publishing

Received August 14, 2014, accepted August 17, 2014, date of publication August 28, 2014, date of current version October 1, 2014.

Digital Object Identifier 10.1109/ACCESS.2014.2353056

Compressed Vision Information Restoration Based on Cloud Prior and Local Prior

FENG JIANG[1], (Member, IEEE), XIAODONG JI[2], (Member, IEEE), CHUNJING HU[2], (Member, IEEE), SHAOHUI LIU[1], (Member, IEEE), AND DEBIN ZHAO[1], (Member, IEEE)

[1]School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China
[2]Key Laboratory of Universal Wireless Communications, Ministry of Education, Beijing University of Posts and Telecommunications, Beijing 100876, China

Corresponding author: F. Jiang (fjiang@hit.edu.cn)

This work was supported in part by the Beijing Natural Science Foundation under Grant 4110001, in part by the National High Technology Research and Development Program of China under Grant 2014AA01A701, in part by the State Major Science and Technology Special Projects under Grant 2013ZX03001001, and in part by the National Natural Science Foundation of China under Grant 61272386 and Grant 61100096.

ABSTRACT In wireless communication, compressed vision information may suffer from kinds of degradation, which dramatically influences the final visual quality. In this paper, a compressed vision information restoration method is proposed based on two explored vision priors: 1) the cloud prior and 2) the local prior. The cloud prior can be obtained from the nature images set in the cloud, and fields of experts is used to formulate the statistical character of the nature image contents as a high order Markov random field. The local prior is achieved from the degraded image itself, and K-SVD is adopted to model the sparse and redundant representation characters of nature images. These priors are effectively comprised in the proposed vision information restoration method. The relation between the quantization parameter and the optimal configuration of the prior models is further analyzed. In addition, an enhanced quantization constrained projection algorithm is proposed to refine the high frequency components. We extend this paper to compressed video restoration for H.264/AVC and the experiment results demonstrate that the proposed scheme can reproduce higher quality images compared with conventional H.264/AVC.

INDEX TERMS Nature image prior, high-quality image restoration, fields of experts, K-SVD.

I. INTRODUCTION

With the rapid development of wireless technologies, the concept of the Fifth Generation (5G) wireless communication system started to emerge and the expectations towards 5G are set much higher in terms of capacity and maximum throughput when compared with Fourth Generation (4G). There are also some new technical challenges the system will need to face, like Machine to Machine (M2M) communication, energy efficiency and complete ubiquity. Developing the new technologies provides uninterrupted access to services such as voice and data, Multimedia Message Service (MMS), video chat, Mobile TV, HDTV content and Digital Video Broadcasting (DVB), which opens the gate that creates a new dimension to our lives and changes our lifestyle significantly.

As the most important form of media, the vision information, images and videos, suffering degradations during the acquisition, storage and transmission, which has profound influence on the performance of Fifth Generation (5G) wireless communication system. Block-based discrete cosine transform coding (BDCT), reducing the redundancy between the pixels of images, is an effective way for image and video compression and has been prevalent in the main image and video compression standards. However, in BDCT, motion compensation, and the coarse quantization results in high frequency (HF) coefficients truncation. Intensity discontinuity and in particular the block truncation of high frequency (HF) cause the blocking effect and ringing artifacts around the contour. Many methods have been proposed to suppress blocking artifacts and these methods can be divided in to three categories. Some of them try to suppress the artifacts based on the transform domain methods, such as DCT [1], sparse 3D transform-domain collaborative filtering [4], over complete wavelet representation (OWR) [2], [3] and spatial domain [5]. Some research devote to the post processing or image iteratively recovery methods based on the theory of projections onto convex sets (POCS) [6]. Recently, learning-based image restoration [7]–[9] have emerged to deal with the artifices in the compressed images and videos.

For the deblocking methods in DCT domain, the direct operation of the DCT coefficients try to reduce artifacts before compressed image being fully decoded. In particular, in order to reduce block artifact, an adaptive filtering method was proposed [1], in which the HVS weighted mechanism and masking quantified constraints is adopted to eliminate the block artifacts. In [2], cross-correlation between the scale and the protection of the edge information are explored to reduce artifacts in compressed images. Liew et al. [3] proposed a method to suppress blocking and ringing effects by analyzing the statistical properties of discrete blocks and cross-scale wavelet coefficients. K. Dabov [4] proposed a novel image denoising strategy, based on the sparse representation in the transform domain. In the spatial domain, spatially adaptive filtering is proposed to the deblocking applications. Based on non-local characters of image, the post-filtering transfer window is proposed [5], which moves the window between adjacent blocks to suppress block artifacts. Some other studies regard the image compression as a kind of distortion, and kinds of algorithms have been proposed for the image restoration. The projection onto convex sets (POCS) algorithm denotes the original prior information as convex sets, and obtain the convergence result through iterative projections. Yang et al. [6] proposed image restoration by integrating local HVS statistical characteristics and directional smoothness constraints.

The above conventional post processing strategies cannot recover the high frequency components of compressed vision information, which has been discarded in the quantization step of compression. Recently, learning-based image restoration [7] has been proposed to rebuild a high-quality image from a codebook which contains the HF components for the low quality image. Because a patch of degraded images can be mapped to more than one high-frequency patches, the learning-based image restoration is an ill-posed problem in essence.

Recently, redundant representation and sparsity based signal denoising has attracted a lot of attention. With the over redundant dictionary, the image denoising can be addressed with the introduction of matching pursuit technique. K-SVD is proposed to solve this problem [14], [15], which achieves good performance. Utilize the sparse prior, K-SVD provides an effective solution for the compromise of the reconstruction and the degraded image under a regulation optimization framework. To extend this work to image sequence and exploring the space and time cues in neighboring patches, 3-D atoms are further proposed [16] in which the propagation of the dictionary over time are explored.

Different from simple restoration tasks, the compressed images restoration should take account of the compatibility with the compression standard, and the possibility of the cooperation with other deblocking techniques, e.g. the loop filtering in H.264/AVC, etc. Although 3-D atoms lead to better denoising performance in some cases, it is not suitable to apply these methods to H.264/AVC directly. Firstly, the predictive coding, motion compensation and the loop filtering used in hybrid coding make it difficult to introduce the 3-D atoms into H.264/AVC directly. Secondly, the performance of 3-D atoms-based method is under the assumption that the Gaussian noise added to the original image sequences is pre-known, and this is quite different from the video compression scenario.

Inspired by the recent progress in sparse prior models of nature image [14], a novel compressed video restoration framework is proposed in this paper and our contributions are summarized as follows:

(1) We model the nature image as a special case of high order Markov random field (MRF), and obtain the cloud prior for the compressed images restoration from a large quantity of images with similar contents and categories in the cloud.

(2) K-SVD is adopted to model the local sparse prior and an adaptive dictionary is obtained for the image effective description. The maximum a posteriori (MAP) estimation is used to alleviate the blocking artifacts.

(3) An enhanced quantization constrained projection algorithm is further proposed to correct the high frequency components in the reconstructed images. The reconstructed images are obtained by blending the loop filtered images and the reconstruction images. Some other considerations are also included in this paper: the compatibility of restoration method with the conventional hybrid coding technique and the possible extension to a novel video coding framework based on images sparse and redundant representation.

II. RELATED WORKS

In last ten years, the natural image prior modeling has become a research focus in the field of computer vision, and with energy minimization criterion, significant progress in low-level vision has been obtained. Generally, the criterion function consists of two terms: one is the likelihood term, ensuring the likeness of the degraded image and the restoration one and the other is the constraint term, constraining the restoration signals satisfying some prior. This approach has been explored in many computer vision tasks, such as estimating optimal flow [17], [18], stereo vision and 3D reconstruction [19], [20], image content segmentation and so on. Some of these tasks can be categorized according to their output, the ''natural image.'' The transparency estimation [21], camera blur restoration [22] image denoising and image inpainting [23] fall in to this category. In the prior modeling process, some models capture the image prior with a large quantity of images while some model need only few image. According to the quality of nature images needed in the prior modeling, the prior models can be divided into cloud prior and local prior categories.

A. CLOUD PRIOR: FILTER-BASED MRFS

Markov random field based image prior has been widely explored and utilized, e.g., noise elimination [25]–[31], inpainting [32], segmentation [34] and 3D vision [33]. The exploration of MRF models contains diversity considerations, such as the scale of the cliques, the design of the filters,

and the form of the energy functions, as well as the parameters estimation approaches.

Most of the early Markov random field approaches adopted the quadratic image smoothness prior [35] and the potential function is the squared summation of local derivation operators. It is noted this would be the most appropriate if derivation filter response of the natural image distribution is indeed Gaussian. In fact, when derivation filter is applied to the images, the distribution of the response of the filter is not Gaussian, which is peaked at zero and contains a heavy tail [10], [36], [37]. Non Gaussian marginal response is also observed for optical-flow and 3D vision [9], [38]. Therefore, Gaussian prior is not a good choice. The recent approaches commonly assume a non-quadratic potential function for local derivation response. Generally, one would prefer learning the potential function from training data. The FRAME model proposed by Zhu [26] uses high-order MRFs to model the filter responses. In the FRAME model, discretized energy functions are obtained from training data, and the filters are chosen from pre-designed candidates. Then a potential function which is built on the response of linear filter ω_k to the image x to estimate the probability:

$$\begin{aligned}\Pr(x; \{w_k, \Psi_k\}) &= \frac{1}{Z(\{w_k, \Psi_k\})} e^{-\sum_{i,k} E_k\left(\omega_{ik}^T x\right)} \\ &= \frac{1}{Z(\{w_k, \Psi_k\})} \prod_{i,k} \Psi_k\left(\omega_{i,k}^T x\right),\end{aligned}$$

where i is denoted as the index of pixels and k is the linear filter index. $\omega_{i,k}^T x$ denotes the response of the linear filter ω_k to x at location i. The partition function $Z(\{\omega_k, \Psi_k\})$ is a normalization constant.

Note the above equation provides a general instance of some popular prior explored in low-level vision. If the filter is horizontal and vertical derivation and the potential function has the quadratic form, it provides the smoothness prior of images Zhu and Mumford [26], [33] utilized Gibbs sampling to estimate the statistics. They assumed that the filters are selected from a pre-defined set of derivation like filters and the potential function has no shape constrain. The obtained results are very non-intuitive. For derivation at the finest scale, the learned potential function are similar to the log histograms and peaked at zero. But at the coarse scale the function form are inverted.

Field of Experts (FoE) [9], [40] are also the instances of high-order MRF models. FoE adopts continuous energy function and the filter is learned from training data for better performance. FoE assumes the energy function has the form of student T distribution and the filters are allowed arbitrary. The exploration of the learned filters achieved superior performance compared with the simple derivation filters on a range of restoration tasks [41]. In [42], auxiliary-variable Gibbs sampling is adopted and minimum Bayesian mean squared error estimation is introduced to explore more general and effective MRF model with the Gaussian scale mixture (GSM) model [43]. Although the sampling based approaches have statistical interpretation, they mostly have high computation complexity, which is challenging in applications. Additionally, there are no closed form solutions for the model expectation and the exact inferences are intractable [26]. Recently, maximum likelihood estimation is approximated with contrastive divergence (CD) [44], which do not need costly equilibrium samples.

B. LOCAL PRIOR: SPARSE AND REDUNDANT REPRESENTATION

The sparse prior of natural images [10] can be categorized into transform-domain techniques, such as discrete cosine transform (DCT) and discrete wavelet transform (DWT). When performing transform to natural image, the original image can be represented by some coefficients of the principal components. In contrast, high-frequency components over the whole coefficient fields generally has many zero or small transform coefficients. Based on these observation in transform domain the shrinkage technique is introduced for image denoising [45], [46], in which a threshold operator is adopted for the coefficients in the transform domain to remove smaller ones and keep larger ones. Image sparsity and redundant representation is a general form of transform-domain sparsity. It demonstrates that a nature image can be approximately reconstructed with a dictionary, such that

$$\mathrm{x} \approx \Phi\alpha, \quad \text{s.t.} \ \|\alpha\|_0 \leq \mathrm{T},$$

where T is a predenoted threshold value and l_0-norm $\|\alpha\|_0$ calculate the number of nonzero elements in α. Utilizing the sparse prior, the restoration of image x can be regarded as an optimal estimation by the following MAP:

$$\hat{\alpha} = \arg\max_{\alpha} \Pr(y/a)\Pr(\alpha),$$

where $\Pr(\mathrm{a}) \propto \exp(-\lambda \|\alpha\|_0)$ is the prior function. Considering the computation complexity of the l_0 regularization optimization, approximation approaches are generally adopted, such as greedy pursuit such as the Orthogonal Matching Pursuit (OMP) [47], [48] and convex relaxation algorithms, such as the Basis Pursuit (BP) [49].

Another vital aspect of sparse prior and redundant representation is the design of the dictionary [50]–[52]. Dictionary learning methods are introduced to learning a dictionary which fits a given set of images adaptively. One of the most effective approaches is K-SVD [14] which utilizes either OMP or BP in the iteration procedure for dictionary estimation. K-SVD can be regarded as a general form of K-means clustering, and alternates between a process of sparse coding of image based on the learned dictionary and a process of the dictionary updating. Denote signals Y and threshold T_0, the K-SVD can be expressed as:

$$\min_{D,X} \left\{ \|Y - DX\|_F^2 \right\} \quad \text{s.t.} \quad \forall i, \|x_i\|_0 \leq T_0,$$

where X is the sparse representation coefficient matrix and D is the learned redundant dictionary. Similar with FoE, K-SVD learns dictionary atoms with the overlapping

image patches. The entire image can be sparsely represented by the concatenation of all of sparse codes. The image construction with the sparse codes becomes an over-determined problem, which has a straightforward least-square solution. In practice, dictionary learned with K-SVD has achieved excellent performance in image restoration.

III. CLOUD PRIOR BASED COMPRESSED IMAGES RESTORATION

For an image to be compressed, the DCT coefficients of each patch are further quantized and the quantization noise arises. This process is undertaken for each block independently and the introduction of discontinuity brings blocking artifacts. Additionally, because of the blurred edges and the truncation of the high frequency DCT coefficients during the quantization and the deblocking steps, the ringing artifacts arise.

A. MAP ESTIMATION AND OPTIMIZATION

One of the most crucial parts of the MAP estimation is to adopt appreciate models for the original image X, i.e. the prior model. Assume

$$Y = X + N, \tag{1}$$

where X is denoted the original image, Y is the compressed image, and N is the block artifacts noise, which is caused by quantization process. Then the MAP estimation can be formulated as

$$\begin{aligned}\hat{X}_{MAP} &= \arg\max_X p(X|Y) \\ &= \arg\max_X \log \frac{p(Y,X)}{p(Y)} \\ &= \arg\max_X \{\log p(Y|X) + \log p(X)\},\end{aligned} \tag{2}$$

Where $P(X)$ represents the image prior model for the original and $P(Y|X)$ is the degradation model.

B. ORIGINAL IMAGE PRIOR MODEL

The image probability density function, i.e. the prior p(X), can be described as a Markov random field. To demonstrate this in detail, let x represent a pixel in the image and suppose a neighboring system ∂x contains all the neighbor pixels of X. A clique c is composed of a set of pixels, neighboring to each other:

$$\forall x, y \in c \Rightarrow y \in \partial x. \tag{3}$$

Figure 1 gives an instance of an 8-point neighborhoods.

MRF mode can be represented with Gibbs function:

$$p(X) = \frac{1}{Z} \exp\left\{-\sum_{c\in C} V_c(X)\right\}, \tag{4}$$

where c denotes a clique, $V_c(X)$ is energy function defined on the clique, and Z is the normalization term.

Denote J_k as a $n \times n$ linear filter, the clique c_t contains $n \times n$ pixels with pixel t in the center, N is the total number of filters applied, and α_k is a positive parameter which is used to adjust the ϕ proper distributions. Field of Experts (FoE) was proposed [8] in the following form:

$$\begin{aligned}p(X) &= \frac{1}{Z} \exp\left\{\sum_{t\in C'}\sum_{k=1}^{N} \log \phi(J_k^T X_{c_t}; \alpha_k)\right\} \\ &= \frac{1}{Z}\prod_{t\in C'}\prod_{k=1}^{N} \phi(J_k^T X_{c_t}; \alpha_k),\end{aligned} \tag{5}$$

where

$$\phi(J_k^T X_{c_t}; \alpha_k) = \left(1 + \frac{1}{2}(J_k^T X_{c_t})^2\right)^{-\alpha_k}. \tag{6}$$

One advantage of FOE is that both the filter J_k and the parameter α_k can be estimated from a training image data set.

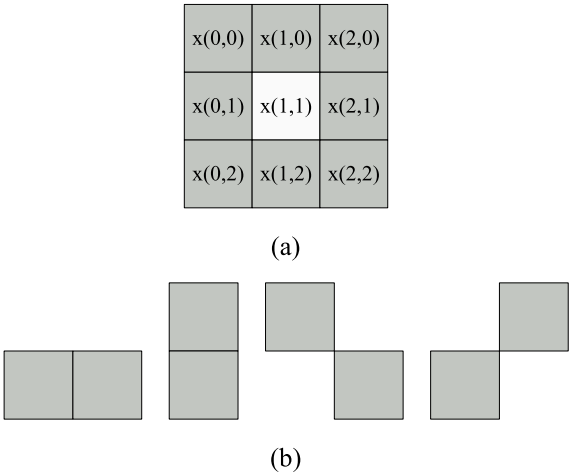


FIGURE 1. An 8-point neighboring system and its 2-point clique. (a) 8-point neighborhood: the gray blocks are the neighbors of x(1,1) (b) 2-point clique.

C. DEGRADATION MODEL

Haven obtained the description of $P(X)$, an appropriate assumption is provided that the quantization noise N and the original image X has the independent relationship,

$$p(Y|X) = p(X + N|X) = p_N(N). \tag{7}$$

We assume that quantization errors satisfies Gaussian distribution:

$$p_N(n(i,j)) = \frac{1}{\sqrt{2\pi}\sigma_N(i,j)} \exp(-\frac{n^2(i,j)}{2\sigma_N^2(i,j)}), \tag{8}$$

where n is an 8×8 patch of the degrade image N, (i,j) denotes the location of the 8×8 block, $\sigma_N^2(i,j)$ represents the variance of quantization error.

There have been many quantization error models been proposed, such as assuming that the quantization error has the same variance at different position of images [12] or exploring the uniform distribution to model the quantization [13]. However, for the pixels at the block boundary, the quantization error tends to be much stronger than those in the center

of block, which is not appreciate to be described with the former variances. We proposes a function to approximately estimate the variance of quantization error:

$$\sigma_N^2(i,j) = a \cdot ((i-3.5)^2 + (j-3.5)^2 + d), \tag{9}$$

where a is denoted as a factor relating with the quantization parameter and d is a constant.

D. OPTIMIZATION METHOD

Once we have obtained the image prior model, the degradation model and by comprising these two to the whole image, an objective function for maximization can be obtained as

$$\begin{aligned}\hat{X}_{MAP} &= \arg\max_X \prod_m p(X(m)|Y(m)) \\ &= \arg\max_X \sum_m (\log p(Y(m)|X(m)) + \log p(X(m)) \\ &= \arg\max_X (\sum_m (-\lambda \cdot \frac{N^2(m)}{2\sigma_N^2}) + \sum_{c_t \in C'} \sum_{k=1}^{N} \log\phi(J_k^T X_{c_t}; \alpha_k))\end{aligned} \tag{10}$$

where m is the patch index, $X(m)$, $Y(m)$ and $N(m)$ are the mth block of the original image, the compressed image and the quantization noise at the same location, respectively. λ is a constant utilized as the weight constant relating to the noise model. Optimizing (10) can be transformed to be a minimization problem with the energy function $E_o(X)$:

$$E_O(X) = \lambda \sum_m \frac{N^2(m)}{2\sigma_N^2} - \sum_{c_t \in C'} \sum_{k=1}^{N} \log\phi(J_k^T X_{c_t}; \alpha_k), \tag{11}$$

which can be expressed as

$$E_O(X) = \lambda \cdot E_N(X) + E(X), \tag{12}$$

where

$$E_N(X) = \sum_m \frac{N^2(m)}{2\sigma_N^2}, \tag{13}$$

$$E(X) = -\sum_{c_t \in C'} \sum_{k=1}^{N} \log\phi(J_k^T X_{c_t}; \alpha_k). \tag{14}$$

Here the conjugate gradient descent is adopted to solve this problem. The gradient of $E(X)$ and $E_N(X)$ can be expressed as:

$$\nabla E_N(X) = \frac{1}{\sigma_N^2}(Y - X), \tag{15}$$

$$\nabla E(X) = -\sum_{k=1}^{N} J_k \cdot \psi_k(J_k X). \tag{16}$$

A further description of the optimization process is provided in [8] and [9].

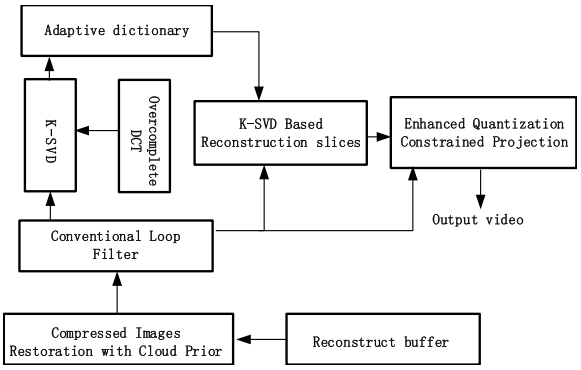


FIGURE 2. Frame work of Images restoration with Cloud prior and Local Prior.

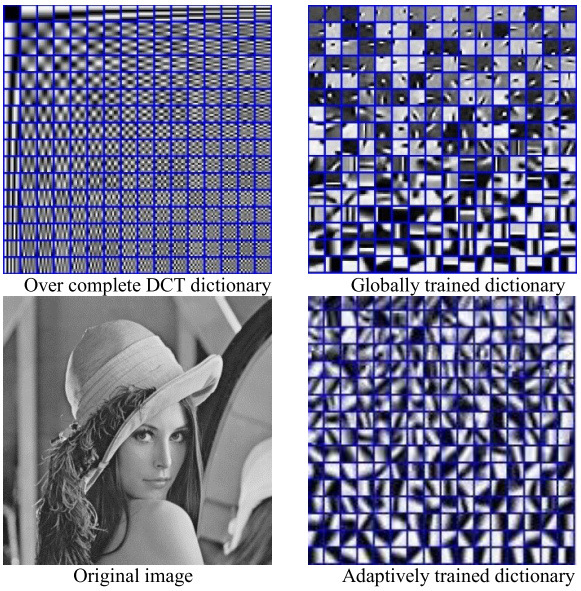


FIGURE 3. Dictionaries for spare representation. Top: Overcomplete DCT dictionary and globally trained dictionary. Bottom: Original image and corresponding adaptively trained dictionary with K-SVD.

IV. LOCAL PRIOR MODELING AND ENHANCEMENT

Having obtained the restoration image with cloud prior, a local prior model is utilized to get refined images. The proposed local prior model and enhancement procedure can be represented as follows: providing a degraded image Y and the quantization parameter Q, a preliminarily restored image $\hat{X}$ is expected to be estimated with MAP strategy as shown in Figure 2.

A. SPARSE REPRESENTATION AND K-SVD ALGORITHM

Most of the traditional approaches are global representation based [11]–[13], in which the features are learned from amount of natural images. The global representation based methods have drawbacks to a certain extend: firstly, it is not precise enough; secondly, it can't obtain full scale features for the original images. Self-adaptive method can be adopted to overcome this problem and obtain effective sparse features and is a perfect solution to this problem. We thus utilize the K-SVD algorithm [14], [15] to estimate the atoms of

the dictionary, which leads to more precise representation of the target image.

B. IMAGE RECONSTRUCTION USING SPARSITY AND REDUNDANCY

Based on sparse and redundant representation a denoising approach is proposed in [14], which has the ability to learn the effective sparse representation for each image block in the training set. A basic assumption is each image block can be reconstructed as a linear combination of a small subset of the basics. Accordingly, the task of image self-adaptive sparse representation can be described as a potential function minimization. The following optimization term describes a combination of three penalties:

$$\left\{\hat{\alpha}_{ij}, \hat{X}\right\} = \arg\min_{\alpha_{ij},X} \lambda \|X - Y\|_2^2 + \sum_{ij} \mu_{ij} \left\|\alpha_{ij}\right\|_0 + \sum_{ij} \left\|D\alpha_{ij} - R_{ij}X\right\|_2^2. \quad (17)$$

The first term is used to constrain the proximity between the degraded image Y and its reconstruction version X. The second term is utilized to satisfy that each block from the reconstruction could be represented up to a bounded error with the dictionary D. The third term demands that the number of coefficients should satisfy representation any patch is small. The values μ_{ij} are patch-specific weights. Minimizing this function yields the sparse representation algorithm and the choice of D is crucial to the performance, as shown in Figure 3. The methods in [8] and [9] are of an iteration block-coordinate relaxation approach, which fixes all the other unknowns except one to be updated. The iteration process can be described as:

1) Updating the sparse representation coefficients $\{\alpha_{ij}\}$: assuming that D and X are fixed, solve the problem as

$$\hat{a}_{ij} = \arg\min_a m_{ij} \|a\|_0 + \left\|Da - x_{ij}\right\|_2^2. \quad (18)$$

In this step, the sparsest vector for each patch in the image is pursuit to describe it with atoms of D. The orthogonal matching pursuit (OMP) algorithm is utilized for this purpose.

2) Updating the atoms in the dictionary D: assuming that X is fixed and update each atom at a time in D, as well as updating the coefficients α_{ij}. This is done via a rank-one approximation of a residual matrix.

3) Updating the reconstructed image X: after rounds of updating $\{\alpha_{ij}\}$ and D, the coefficients of all blocks obtain a sparse representation for the original image

$$\hat{X} = \arg\min_x \lambda \|X - Y\|_2^2 + \sum_{ij} \left\|D\hat{\alpha}_{ij} - R_{ij}X\right\|_2^2. \quad (19)$$

This optimization is solved with a simple weighting of the represented blocks with overlaps.

C. ENHANCED QUANTIZATION CONSTRAINED PROJECTION

After applying K-SVD for single frame, the frame is divided into blocks which will be transformed and quantized to get new quantization coefficients (NQC). If NQC are different from the corresponding original quantization coefficients that will be transmitted to the decoder, the pixel values of the block will be modified by the following methods.

The quantization constraint and the range constraint are respectively imposed on the DCT coefficients and the pixel values during the iteration. It is our priori knowledge that the original DCT coefficients must lie within the quantization intervals and the pixel values between 0 and 255. If either of them is violated, the intermediate result should be set to the nearest value satisfying the corresponding constraint. In the BDCT process, the BDCT quantization coefficients compose a constrained set:

$$C^T = \left\{x : \hat{x}_{(i,j)}^{\min} < BDCT(x) < \hat{x}_{(i,j)}^{\max}, \quad \forall i = 1, 2, \ldots, M; \quad \forall j = 1, 2, \ldots N \right. \quad (20)$$

where C^T is the constrained set, (i, j) is the corresponding pixel location; $\hat{x}_{(i,j)}^{\min}$ and $\hat{x}_{(i,j)}^{\max}$ are the boundaries of quantization interval. The projection of image block $\boldsymbol{x}$ to the convex set C^T can be represented as follows:

$$P^T(x) = BDCT^{-1} \cdot \hat{x}, \quad (21)$$

where P^T is the mapping process, and the quantization constrained pixel value at location (i, j) can be represented as:

$$KSVD\_pixel\ (i,j) = \hat{x}(i,j) = \begin{cases} \hat{x}_{(i,j)}^{\min}, & if\ BDCT(x)_{(i,j)} < \hat{x}_{(i,j)}^{\min} \\ \hat{x}_{(i,j)}^{\max}, & if\ BDCT(x)_{(i,j)} > \hat{x}_{(i,j)}^{\max} \\ BDCT(x)_{(i,j)}, & otherwise \end{cases} \quad (22)$$

The final restored image is achieved by blending the loop filtered results LoopPixel(i, j) with the quantization constrained results KsvdPixel(i, j).

$$X_E = \omega_1 \cdot LoopPixel\ (i,j) + \omega_2 \cdot KsvdPixel\ (i,j) \quad (23)$$

where $w_1 + w_2 = 1$ and w_1 is weighted coefficient for loop filtered results and w_2 for quantization constrained results.

D. THE PROPOSED EDGE ENHANCEMENT

After the aforementioned steps, a restoration image could be obtained. However, because the FoE model based on a set of learned filters is a no-bias model, which processes the smooth area and the contour area of image equally. So the edge may be blurred, even may the ringing artifacts are arisen. Since the edge area of an image is important for visual attention, the loss of edge information could greatly degrade the subjective quality. Therefore, in order to restore the edge information, an edge enhancement algorithm is proposed, which is illustrated in Figure 4.

A Laplacian Image $\hat{X}_{Laps}$ is obtained by constructing a Gaussian/Laplacian image pyramid as shown in Figure 5,

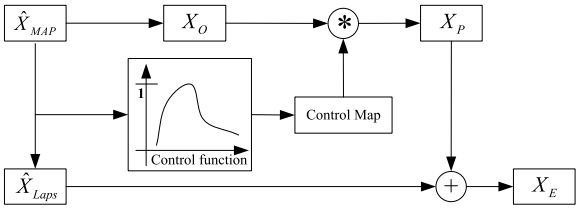


FIGURE 4. The proposed enhancement scheme.

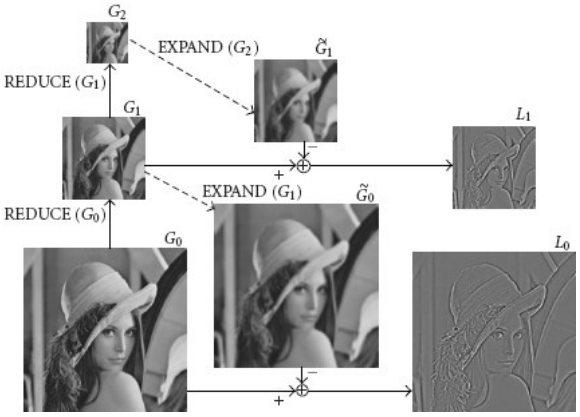


FIGURE 5. Gaussian/Laplacian image pyramid (N = 3).

from which we can further obtain the HF content for edge restoration.

This process can be demonstrated by the following expressions: Denote G_n as the nth Gaussian image of the input G_0, then G_n can be calculated with a $REDUCE(\cdot)$ operation:

$$G_{n+1} = REDUCE(G_n), \quad 0 \le n \le N-1, \tag{24}$$

where N is the levels number in the pyramid structure and G_0 represents the highest resolution. Denote $\tilde{G}_n$ as the image expended from G_{n+1} with an $EXPEND(\cdot)$ operation:

$$\tilde{G}_n = EXPAND(G_{n+1}), \quad 0 \le n \le N-1. \tag{25}$$

The Laplacian pyramid can be got by

$$L_n = \tilde{G}_n - EXPAND(G_{n+1}), \quad 0 \le n \le N-1, \tag{26}$$

where L_n is the *nth* Laplacian pyramid image of G_0.

Here only the 0th Laplacian image of $\hat{X}_{MAP}$ is utilized, denoted as $\hat{X}_{Laps}$. Although $\hat{X}_{Laps}$ contains the HF information of $\hat{X}_{MAP}$, it is inevitable to discard some true HF information of the original image. For the next step we will enhance this HF image as shown in Figure 4, with the following expression:

$$X_E = \hat{X}_{Laps} + X_P = \hat{X}_{Laps} + \gamma \cdot C_M \cdot X_O. \tag{27}$$

By performing the control function with $\hat{X}_{MAP}$, the control map C_M can be obtained and X_O is the HF component of $\hat{X}_{MAP}$. Figure 6 provides an instance of the control function, which as a function of local activity LA, LA_0 and C_0 representing the parameters to adjust the control function.

The control function should satisfy:

$$\lim_{LA(i,j)\to 0} C_f(i,j) = C_0, \quad \lim_{LA(i,j)\to\infty} C_f(i,j) = 0. \tag{28}$$

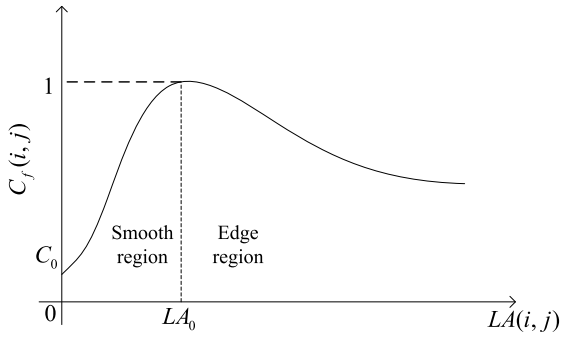


FIGURE 6. Control function $C_f(i,j)$ as a function of the local activity $LA(i,j)$.

With the adoption of the control function, the overshoots near sharp edges can be limited and the noise in the smooth area can be attenuated at the same time. The control function can be defined as:

$$C_f(i,j) = \frac{LA(i,j) + c}{k \cdot (LA(i,j))^2 + l}$$
$$a = \frac{C_0 \cdot LA_0}{2(1-C_0)}, \; l = \frac{LA_0}{2(1-C_0)}, \; k = \frac{1}{2 \cdot LA_0}, \tag{29}$$

where (i,j) is the pixel location in the image. With precise setting these two parameters, it can be ensured that only the selected detail information in the contour regions is refined. By comprising X_P with X_D, enhanced HF image X_E can be obtained with local compatibilities between the primitives enforced. As to the local activity $LA(i,j)$, the response of Sobel operator is utilized to measure the intensity variance of local regions in this paper. By combining the enhancement HF image with the deblocking image, a higher-quality image could be reconstructed. Finally, quantization constrained projection is used to improve the image quality by eliminating possible over blurring influence.

V. EXPERIMENT RESULTS

In this section, extensive experiments are performed to evaluate the performance of the proposed compressed vision information restoration based on cloud priors and local priors, which is compared with the other state-of-the-art methods. In our experiments, the parameters used in the prior modeling are set as follows. The filter's scale of FoE is 5×5, which have the ability to extract the structural information in natural images. The tested images are shown in Fig. 7.

A. COMPRESSED IMAGE RESTORATION

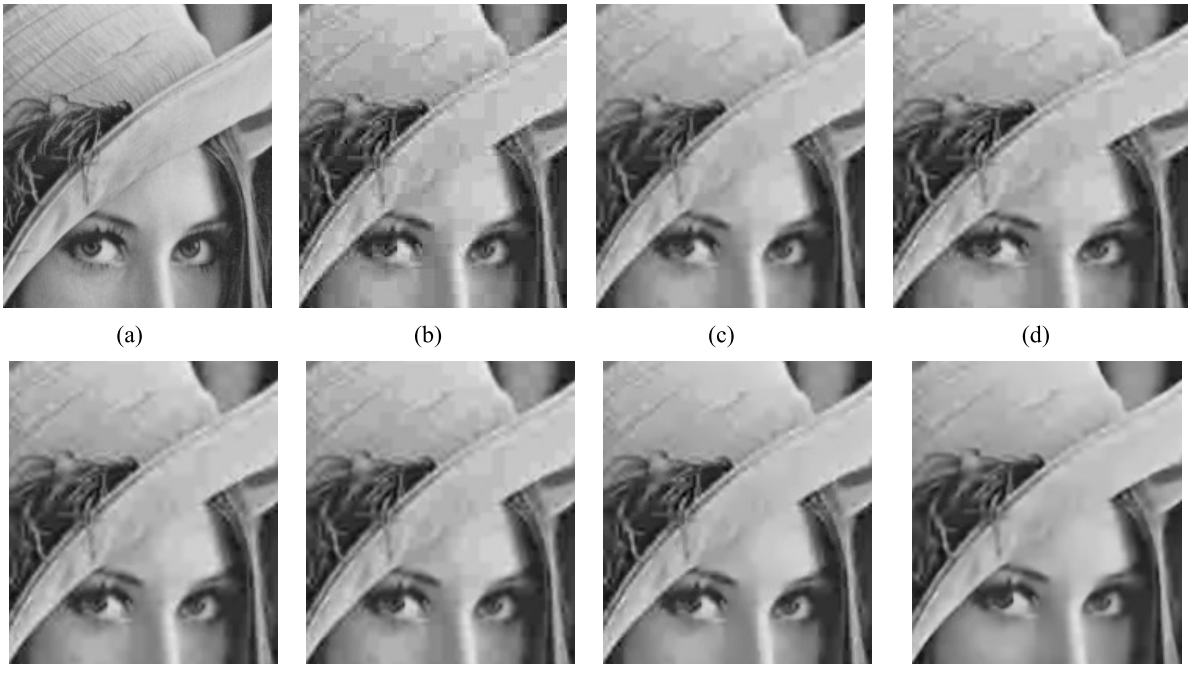
To evaluate the proposed cloud prior and local prior based images restoration, we compare our method with some other popular strategies and adopt the peak signal-to-noise ratio as the objective quality assessment and the comparison results are reported in Table 1. The best two performances for each column are emphasized by the shades. With these comparisons, it can be observed that the proposed method achieves significantly better performance emphasized in bold than

TABLE 1. PSNR comparison of different image restoration schemes.

Image	LENA				PEPPERS				BARBARA				BABOON			
JPEG	28.24	30.41	31.95	32.96	30.18	31.58	32.47	33.61	23.86	25.70	27.05	28.25	21.52	22.47	23.43	24.51
Xiong[2]	29.24	31.05	32.30	33.03	30.45	31.45	32.05	32.76	23.99	25.15	25.80	26.25	21.86	22.59	23.25	23.96
Yang[6]	28.99	31.03	32.42	33.35	30.65	31.93	32.75	33.80	24.25	26.10	27.42	28.60	21.86	22.76	23.69	24.71
Chen[1]	29.14	31.16	32.40	33.28	30.88	31.91	32.70	33.87	24.52	26.11	27.12	28.07	21.97	22.90	23.67	24.49
Zhai[5]	29.23	31.25	32.68	33.57	30.76	32.05	32.85	33.86	24.43	26.15	27.51	28.69	21.98	22.88	23.76	24.79
Liew[3]	29.55	31.44	32.73	33.55	31.16	32.29	33.03	33.99	24.50	26.05	27.38	28.56	21.99	22.80	23.65	24.68
proposed	**30.16**	**32.06**	**33.42**	**34.31**	**31.81**	**32.93**	**33.65**	**34.56**	**24.75**	**26.53**	**27.85**	**29.01**	**22.03**	**22.91**	**23.80**	**24.80**

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

FIGURE 7. Typical test images (a) LENA, (b) PEPPERS, (c) BARBARA, and (d) BABOON.



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

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

FIGURE 8. Subjective quality comparison on LENA. (a) Original image. (b) JPEG 30.41dB. (c) Xiong 31.05dB. (d) Yang 31.03dB. (e) Chen 31.16dB. (f) Zhai 31.25dB. (g) Liew 31.44dB. (h) Proposed 32.05dB.

JPEG and the other deblocking methods, even at higher bit rates. Significant improvements in image quality are obtained both in smooth and texture images. As shown in Fig. 8, the subjective result of the proposed method outperforms the other restoration methods. Fig. 9 illustrates the performance of the proposed edge enhancement. With the proposed

(a) 30.403dB (b) 31.634dB (c) 32.057dB

(d)27.155dB (e)27.846dB (f)27.899dB

FIGURE 9. Restored images of the proposed scheme: a) JPEG-coded "LENA" b) Restored "LENA" without the proposed edge enhancement c) Restored "LENA" with the proposed edge enhancement d) JPEG-coded "BARBARA" e) Restored "BARBARA" without the proposed edge enhancement f) Restored "BARBARA" with the proposed edge enhancement.

TABLE 2. Coding conditions.

Parameter	settings
Reference software	JM10.0
Profile	High(100)
GOP structure	IPPIPP…
Number of frames to be encoded	30
Number of reference frames	5
Rate-distortion optimization	On
Quantization parameter	44,40,36
Adaptive rounding	Off
CABAC	On

cloud prior and local prior based images restoration, greater improvements can be achieved in the images which contain large smooth area, such as LENA and PEPPERS. However, in the images containing most texture information, the improvement is relatively smaller.

B. COMPRESSED VIDEO RESTORATION

The simulation conditions are given in Table II. In our simulations, all intra/inter encoding modes are enabled to evaluate the performance of the proposal methods. In order to highlight the results of our proposal, the loop filtered results rather than the results of our proposal are used as reference for inter prediction.

Enhanced Quantization Constrained Projection uses formula (23) to correct the pixel values. EQCP corrects all the pixels of all modes with $w_1 = (QP/4 - 3)/10$, $w_2 = 1 - w_1$ for P images, and corrects pixels of mode Intra 16×16 with $w_1 = 0.6, w_2 = 0.4$ for I images. The performance comparison of PSNR is depicted in Table III.

The experiment results reveal that for the proposed method, the PSNR of restored I images is better than that of the restored P image with the same QP, which is mostly because the blocking artifacts of I images are more obvious. Another reason for less PSNR gain of P image is that if the quantization coefficients of a block are small enough, they will be discarded for the sake of saving bitrates. As to the choose of the parameter λ, with a large-scale test, we find empirically that the optimized value of $\lambda\lambda$ depends on the QP value and the type of the images. As seen in Table III, the optimized λ of I images is smaller than that of P images; as QP goes down, better results are achieved with larger values of λ. As for the EQCP, the weighted fusion of the loop filter results and those of K-SVD turns out to be better than using either of them. The pixel values of all modes will be processed with the weights related to QP for P images, while those of the mode Intra 16×16 with constant weights, 0.4 and 0.6 will be appropriate for I images. The proposed method achieves better performance for P images than I images, and the confidence of the restoration of P images is higher for the sequences with less movement.

TABLE 3. Comparison of PSNR between JM10.0 and the proposal.

QP	image type	30 / λ	PSNR(dB)				Gain
			jm10.0		*proposal*		
			predict	***loop***	***KSVD***	***EQCP***	
ICE_4cif							
44	I	12	31.32	31.91	32.33	32.34	**0.43**
	P	9	31.79	31.91	32.15	32.16	**0.25**
40	I	9	33.84	34.37	34.78	34.79	**0.42**
	P	7	34.12	34.22	34.43	34.44	**0.22**
36	I	6	36.16	36.58	36.91	36.91	**0.33**
	P	5	36.23	36.32	36.45	36.46	**0.14**
foreman_cif							
44	I	11	27.58	28.09	28.47	28.49	**0.40**
	P	6	28.16	28.21	28.33	28.33	**0.12**
40	I	8	30.02	30.52	30.90	30.91	**0.39**
	P	4	30.60	30.65	30.72	30.74	**0.08**
36	I	6	32.40	32.81	33.16	33.18	**0.37**
	P	3	32.86	32.91	32.97	32.98	**0.07**
football_cif							
44	I	10	25.47	25.78	26.05	26.05	**0.28**
	P	8	25.17	25.26	25.36	25.37	**0.12**
40	I	7	27.92	28.15	28.45	28.47	**0.32**
	P	6	27.37	27.46	27.58	27.60	**0.13**
36	I	4	30.36	30.52	30.88	30.88	**0.36**
	P	4	29.74	29.82	29.94	29.95	**0.13**
foreman_qcif							
44	I	10	25.59	25.98	26.37	26.37	**0.39**
	P	4	26.40	26.41	26.47	26.47	**0.06**
40	I	8	28.15	28.49	28.95	28.94	**0.45**
	P	3	29.14	29.15	29.22	29.21	**0.06**
36	I	6	30.77	31.04	31.52	31.52	**0.49**
	P	3	31.46	31.48	31.55	31.56	**0.08**

VI. CONCLUSION

In this paper, a compressed video restoration scheme based on cloud and local priors for vision information is proposed. A cloud prior model is obtained by training on a huge nature image dataset in the cloud which models the nature image in the form of high order Markov random field. The local prior model explores the sparse characters of nature images. These two models are comprised to eliminate the block artifacts in the compressed vision information. The enhanced quantization constrained projection method is further provided to refine the high frequency components of the vision information. Simulation evaluations demonstrate that the proposed methods outperform the current deblocking algorithms.

REFERENCES

[1] T. Chen, H. R. Wu, and B. Qiu, "Adaptive postfiltering of transform coefficients for the reduction of blocking artifacts," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 5, pp. 594–602, May 2001.

[2] Z. Xiong, M. T. Orchard, and Y.-Q. Zhang, "A deblocking algorithm for JPEG compressed images using overcomplete wavelet representations," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 2, pp. 433–437, Apr. 1997.

[3] A. W.-C. Liew and H. Yan, "Blocking artifacts suppression in block-coded images using overcomplete wavelet representation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 4, pp. 450–461, Apr. 2004.

[4] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[5] G. Zhai, J. Cai, W. Lin, X. Yang, and W. Zhang, "Shifted window based filtering for alleviating blocking artifacts," in *Proc. IEEE Workshop Signal Process. Syst.*, Oct. 2007, pp. 289–294.

[6] Y. Yang, N. P. Galatsanos, and A. K. Katsaggelos, "Projection-based spatially adaptive reconstruction of block-transform compressed images," *IEEE Trans. Image Process.*, vol. 4, no. 7, pp. 896–908, Jul. 1995.

[7] L. Ma, F. Wu, D. Zhao, W. Gao, and S. Ma, "Learning-based image restoration for compressed image through neighboring embedding," in *Advances in Multimedia Information Processing*. Berlin, Germany: Springer-Verlag, 2008, pp. 279–286.

[8] S. Roth and M. J. Black, "Field of experts: A framework for learning image priors," in *Proc. IEEE Comput. Vis. Pattern Recognit.*, vol. 2. Jun. 2005, pp. 860–867.

[9] 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.

[10] 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. 6583, pp. 607–609, 1996.

[11] X. Hou and L. Zhang, "Dynamic visual attention: Searching for coding length increments," in *Advances in Neural Information Processing Systems*. Red Hook, NY, USA: Curran Associates, Inc., 2008, pp. 681–688.

[12] L. Zhang, M. H. Tong, T. K. Marks, H. S. Shan, and G. W. Cottrell, "SUN: A Bayesian framework for saliency using natural statistics," *J. Vis.*, vol. 8, no. 7, 2008, Art. ID 32.

[13] N. Bruce and J. Tsotsos, "Saliency based on information maximization," in *Advances in Neural Information Processing Systems*. Cambridge, MA, USA: MIT Press, 2006, pp. 155–162.

[14] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Image Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.

[15] M. Elad and M. Aharon, "Image denoising via learned dictionaries and sparse representation," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, New York, NY, USA, Jun. 2006, pp. 895–900.

[16] M. Protter and M. Elad, "Image sequence denoising via sparse and redundant representations," *IEEE Trans. Image Process.*, vol. 18, no. 1, pp. 27–35, Jan. 2009.

[17] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, Jul. 2003.

[18] X. Lan, S. Roth, D. Huttenlocher, and M. J. Black, "Efficient belief propagation with learned higher-order Markov random fields," in *Proc. 9th ECCV*, 2006, pp. 269–282.

[19] O. J. Woodford, C. Rother, and V. Kolmogorov, "A global perspective on MAP inference for low-level vision," in *Proc. IEEE 12th ICCV*, Sep./Oct. 2009, pp. 2319–2326.

[20] P. Kohli and M. P. Kumar, "Energy minimization for linear envelope MRFs," in *Proc. IEEE CVPR*, Jun. 2010, pp. 1863–1870.

[21] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 11, pp. 1222–1239, Nov. 2001.

[22] G. Peyré, S. Bougleux, and L. Cohen, "Non-local regularization of inverse problems," in *Proc. Eur. Conf. Comput. Vis.*, Marseille, France, Oct. 2008, pp. 57–68.

[23] X. Zhang, M. Burger, X. Bresson, and S. Osher, "Bregmanized nonlocal regularization for deconvolution and sparse reconstruction," *SIAM J. Imag. Sci.*, vol. 3, no. 3, pp. 253–276, 2010.

[24] M. Jung, X. Bresson, T. F. Chan, and L. A. Vese, "Nonlocal Mumford–Shah regularizers for color image restoration," *IEEE Trans. Image Process.*, vol. 20, no. 6, pp. 1583–1598, Jun. 2011.

[25] F. R. K. Chung, *Spectral Graph Theory*. Providence, RI, USA: AMS, 1997, p. 212.

[26] S. C. Zhu and D. Mumford, "Prior learning and Gibbs reaction-diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 11, pp. 1236–1250, Nov. 1997.

[27] Y. Li and D. P. Huttenlocher, "Sparse long-range random field and its application to image denoising," in *Proc. 10th ECCV*, 2008, pp. 344–357.

[28] M. F. Tappen, "Utilizing variational optimization to learn Markov random fields," in *Proc. IEEE CVPR*, Jun. 2007, pp. 1–8.

[29] K. G. G. Samuel and M. F. Tappen, "Learning optimized MAP estimates in continuously-valued MRF models," in *Proc. IEEE CVPR*, Jun. 2009, pp. 477–484.
[30] M. F. Tappen, C. Liu, E. H. Adelson, and W. T. Freeman, "Learning Gaussian conditional random fields for low-level vision," in *Proc. IEEE CVPR*, Jun. 2007, pp. 1–8.
[31] J. J. McAuley, T. S. Caetano, A. J. Smola, and M. O. Franz, "Learning high-order MRF priors of color images," in *Proc. 23rd ICML*, 2006, pp. 617–624.
[32] S. Roth and M. J. Black, "Steerable random field," in *Proc. IEEE 11th ICCV*, Oct. 2007, pp. 1–8.
[33] J. Sun, H.-Y. Shum, and N.-N. Zheng, "Stereo matching using belief propagation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 7, pp. 787–800, Jul. 2003.
[34] A. Barbu and S. C. Zhu, "Generalizing Swendsen–Wang to sampling arbitrary posterior probabilities," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1239–1253, Aug. 2005.
[35] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artif. Intell.*, vol. 17, nos. 1–3, pp. 185–203, Aug. 1981.
[36] E. Simoncelli, "Statistical models for images: Compression, restoration and synthesis," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Nov. 1997, pp. 673–678.
[37] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain," *IEEE Trans. Image Process.*, vol. 12, no. 11, pp. 1338–1351, Nov. 2003.
[38] J. Huang, A. B. Lee, and D. Mumford, "Statistics of range images," in *Proc. IEEE CVPR*, pp. 1324–1331, Jun. 2000.
[39] G. E. Hinton, "Products of experts," in *Proc. Int. Conf. Artif. Neural Netw. (ICANN)*, vol. 1. 1999, pp. 1–6.
[40] S. Roth and M. J. Black, "Fields of experts," *Int. J. Comput. Vis.*, vol. 82, no. 2, pp. 205–229, 2009.
[41] M. J. Black and P. Anandan, "The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields," *Comput. Vis. Image Understand.*, vol. 63, no. 1, pp. 75–104, 1996.
[42] U. Schmidt, Q. Gao, and S. Roth, "A generative perspective on MRFs in low-level vision," in *Proc. IEEE CVPR*, Jun. 2010, pp. 1751–1758.
[43] S. Lyu and E. P. Simoncelli, "Modeling multiscale subbands of photographic images with fields of Gaussian scale mixtures," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 4, pp. 693–706, Apr. 2009.
[44] D. Krishnan and R. Fergus, "Fast image deconvolution using hyper-Laplacian priors," in *Proc. NIPS*, vol. 22. 2009, pp. 1–9.
[45] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.
[46] E. Simoncelli and E. Adelson, "Noise removal via Bayesian wavelet coring," in *Proc. Int. Conf. Image Process.*, Sep. 1996, pp. 379–382.
[47] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proc. 27th Asilomar Conf. Signals, Syst. Comput.*, vol. 1. Nov. 1993, pp. 40–44.
[48] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.
[49] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Rev.*, vol. 43, no. 1, pp. 129–159, Jan. 2001.
[50] E. J. Candes and D. L. Donoho, "Recovering edges in ill-posed inverse problems: Optimality of curvelet frames," *Ann. Statist.*, vol. 30, no. 3, pp. 784–842, 2002.
[51] M. N. Do and M. Vetterli, "Contourlets: A directional multiresolution image representation," in *Proc. Int. Conf. Image Process.*, 2002, pp. 357–360.
[52] G. Easley, D. Labate, and W.-Q. Lim, "Optimally sparse image representations using shearlets," in *Proc. 14th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2006, pp. 974–978.

FENG JIANG (M'08) received the B.S., M.S., and Ph.D. degrees in computer science from the Harbin Institute of Technology (HIT), Harbin, China, in 2001, 2003, and 2008, respectively. He is currently an Associated Professor with the Department of Computer Science, HIT, and a Visiting Scholar with the School of Electrical Engineering, Princeton University, Princeton, NJ, USA. His research interests include computer vision, pattern recognition, and image and video processing.

XIAODONG JI (M'07) received the B.S. and M.S. degrees in radio communications, and the Ph.D. degree from the Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 1985, 1988, and 1994, respectively. He is currently an Associate Professor with the School of Information and Communication Engineering, BUPT. His research interests include wireless communications and mobile IP.

CHUNJING HU (M'07) is currently an Associate Professor and a Master Tutor with the Wireless Signal Processing and Network Laboratory (WSPN), Beijing University of Posts and Telecommunications (BUPT), Beijing, China, directed by Prof. W. Wenbo. She received the B.S., M.S., and Ph.D. degrees from BUPT in 1991, 1994, and 2007, respectively. Since 1994, she has been with BUPT, where she is involved in research on speech recognition, machine translation, and wireless applications. She has authored four books, and over 10 papers at domestic core journals and international academic conferences.

SHAOHUI LIU (M'08) received the B.S., M.S., and Ph.D. degrees in computer science from the Harbin Institute of Technology (HIT), Harbin, China, in 2000, 2002, and 2007, respectively. He is currently an Associate Professor with the Department of Computer Science, HIT. His research interests include computer vision, pattern recognition, and image and video processing.

DEBIN ZHAO (M'11) received the B.S., M.S., and Ph.D. degrees in computer science from the Harbin Institute of Technology (HIT), Harbin, China, in 1985, 1988, and 1998, respectively.

He is currently a Professor with the Department of Computer Science, HIT. He has authored over 200 technical articles in refereed journals and conference proceedings in the areas of image and video coding, video processing, video streaming and transmission, and pattern recognition.