

Compressive Sampling-Based Image Coding for Resource-Deficient Visual Communication

Xianming Liu, *Member, IEEE*, Deming Zhai, *Member, IEEE*, Jiantao Zhou, *Member, IEEE*, Xinfeng Zhang, Debin Zhao, *Member, IEEE*, and Wen Gao, *Fellow, IEEE*

Abstract—In this paper, a new compressive sampling-based image coding scheme is developed to achieve competitive coding efficiency at lower encoder computational complexity, while supporting error resilience. This technique is particularly suitable for visual communication with resource-deficient devices. At the encoder, compact image representation is produced, which is a polyphase down-sampled version of the input image; but the conventional low-pass filter prior to down-sampling is replaced by a local random binary convolution kernel. The pixels of the resulting down-sampled pre-filtered image are local random measurements and placed in the original spatial configuration. The advantages of the local random measurements are two folds: 1) preserve high-frequency image features that are otherwise discarded by low-pass filtering and 2) remain a conventional image and can therefore be coded by any standardized codec to remove the statistical redundancy of larger scales. Moreover, measurements generated by different kernels can be considered as the multiple descriptions of the original image and therefore the proposed scheme has the advantage of multiple description coding. At the decoder, a unified sparsity-based soft-decoding technique is developed to recover the original image from received measurements in a framework of compressive sensing. Experimental results demonstrate that the proposed scheme is competitive compared with existing methods, with a unique strength of recovering fine details and sharp edges at low bit-rates.

Index Terms—Low bit-rates image coding, multiple description coding, local random sampling, compressive sensing.

Manuscript received August 5, 2015; revised December 8, 2015, February 3, 2016, and April 1, 2016; accepted April 9, 2016. Date of publication April 14, 2016; date of current version April 29, 2016. This work was supported in part by the Major State Basic Research Development Program of China (973 Program) under Grant 2015CB351804, in part by the National Science Foundation of China under Grant 61300110, Grant 61502122, and Grant 61402547, in part by the Fundamental Research Funds for the Central Universities under Grant HIT.NSRIF.2015067 and Grant HIT.NSRIF.201653, in part by the Macau Science and Technology Development Fund under Grant FDCT/009/2013/A1 and Grant FDCT/046/2014/A1, and in part by the Research Committee at the University of Macau under Grant MRG007/ZJT/2015/FST and Grant MYRG2015-00056-FST. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Aljosa Smolic. (*Corresponding author: Deming Zhai.*)

X. Liu, D. Zhai, and D. Zhao are with the School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China (e-mail: xmliu.hit@gmail.com; zhaideming@gmail.com; dbzhaog@hit.edu.cn).

J. Zhou is with the Department of Computer and Information Science, Faculty of Science and Technology, University of Macau, Macau 999078, China (e-mail: jtzhou@umac.mo).

X. Zhang is with the Rapid-Rich Object Search Laboratory, Nanyang Technological University, Singapore 639798 (e-mail: xfzhang@ntu.edu.sg).

W. Gao is with the National Engineering Laboratory for Video Technology and the Key Laboratory of Machine Perception, Ministry of Education, School of Electrical Engineering and Computer Science, Peking University, Beijing 100871, China (e-mail: wgao@pku.edu.cn).

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.2016.2554320

I. INTRODUCTION

WITH the recent advances in imaging sensor technology, digital cameras have been constantly improved in terms of image quality, spatial resolution and cost. As a result, it becomes popular to integrate imaging ability into wireless sensor networks. Such visual sensors collect and transmit critical visual information that characterizes physical phenomena around them, leading to many emerging applications such as visual surveillance of wild animals, vehicle traffic monitoring, and mobile multimedia.

Although cameras integrated in wireless sensors are becoming increasingly powerful, there still exist many obstacles. Inexpensive wireless visual sensors are usually equipped with batteries of limited capacity, and hence, cannot sustain heavy computations involved in visual data compression and communication. In addition, the energy provisioned for visual sensors is not expected to be renewed throughout their mission, because sensor nodes may be deployed in a hostile or unpractical environment. Storing raw visual data without compression is clearly not an option, due to the limitation of storage capacity. The only feasible solution is an asymmetric visual codec design, in which the encoder is made computationally simple and hence energy efficient, while shifting the heavy computation burdens of retaining high-quality images to the decoder side [1]. The heavy-duty decoding process can be performed by powerful computers or perceivably in the near future by cloud computing, upon receiving the compressed visual data. Meanwhile, it is necessary to provide some error resilience mechanism against instability of wireless channels. In a nutshell, low encoder complexity, satisfactory compression performance in low bit-rates, and error resilience for robust image transmission are the main design goals for wireless visual communications.

In this work, we focus on still image compression for resource-deficient wireless visual sensors. To meet the design goal of light-duty encoder, theoretically, the approach could be distributed source coding (DSC) [2], [3] or compressive sensing (CS) [4], [5], as many researchers recently advocated. DSC allows simple encoding and shifts the heavy complexity to the decoder side through the mechanism of separately encoding but jointly decoding. DSC has an inbuilt robustness against channel losses, thanks to the duality between DSC and channel coding. Song *et al.* [6] proposed a cloud-based distributed image coding scheme. In their method, an input image is reconstructed in the cloud using retrieved correlated images, which serve as the side information (SI). The image is then

compressed through a transform-domain syndrome coding to correct the disparity between the original image and the SI by an iterative refinement process. While this scheme performs comparably with HEVC-intra coding at low bit-rates, it highly relies on whether there are correlated images in the cloud.

The CS theory implies that it is possible to acquire relatively few measurements of the signal in a linear fashion while still permitting exact reconstruction via a complex recovery process at the decoder. By unifying sampling and compression, CS-based image coding techniques have received significant interests recently [11]–[13]. However, up to now, the rate-distortion (RD) performance of CS-based image coding methods is quite disappointing compared with conventional image coding standards. The main cause is the wide use of a uniform scalar quantizer (SQ), which overlooks hidden correlations between different subsets of CS measurements. To alleviate this problem, Mun and Fowler [11] proposed to conjunct SQ with differential pulse code modulation of the block-based CS measurements, in which only the resulting prediction residual is scalar quantized. Wang *et al.* [12] presented a progressive fixed-rate SQ (PSQ) with binning that enables the CS decoder to exploit hidden correlations among measurements. However, even though the performance of these methods improves significantly over the pure SQ-based schemes, they can only match or slightly outperform JPEG in some cases. Furthermore, all these CS-based image coding schemes do not consider the error resilience ability to combat packet loss in wireless transmission. Recently, Deng *et al.* [13] suggested to combine the CS principle with multiple description coding (MDC), which permits a light-duty encoder as well as a robust transmitter. However, this CS-based MDC scheme dramatically changes the standardized code stream syntax.

In this paper, we develop a novel CS-based image coding scheme via local random sampling (LRS), which achieves state-of-the-art coding efficiency with lower encoder computational complexity, while still supporting error resilience. The short version of this work was presented in [10]. The main contributions of this paper are highlighted as follows:

1) Firstly, to obtain a light-duty encoder, we investigate the problem of compact image representation in an approach of compressive sampling in the spatial domain. Different from previous down-sampling based image coding schemes [14], [15], we replace the low-pass down-sampling filter, such as Gaussian and box filters, by a $w \times w$ binary convolution kernel of random coefficients. As the latter is a broadband filter (see Fig. 1 to compare the spectrum responses of random convolution and Gaussian kernels), it retains high-frequency information in the down-sampled image and hence leaves the decoder the possibility of reconstructing sharper and clearer images than the low-pass pre-filters do.

Sparse sampling greatly reduces the encoder complexity, since the down-sampled image is only a small fraction of the original size. This property allows the system to shift the computation burden from encoder to decoder, making our scheme a viable asymmetric compression solution when the encoder is severely resource-deficient. It is worth noting that the low encoder complexity here refers to computational complexity, rather than that involved for hardware implementation.

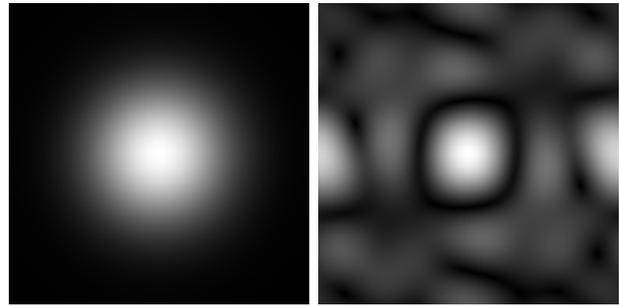


Fig. 1. Fourier spectra of Gaussian kernel (left) and local random convolution kernel (right).

2) Secondly, the proposed scheme naturally possesses the advantages of MDC, and therefore, it permits robust transmission of image data. The down-sampled image generated by LRS can be inherently considered as one description of the original image. We can obtain a number of descriptions (≥ 2) by using different local random sampling matrices. Moreover, the corresponding multiple description (MD) decoding is carried out in the well-known framework of compressive image recovery, which is performed in the same way independently with the number of available descriptions. At this point, when used as a MDC scheme, a unique advantage of our method is noteworthy: the unification of the central decoder and all side decoders.

3) Thirdly, the proposed LRS strategy is flexible to be incorporated with existing image coding standards. Since natural images comprise large smooth regions, there still exist statistical redundancies among locally sampled measurements. Moreover, the sampled measurements remain the conventional form of pixel grid, making the down-sampled image readily compressible by any existing standard image codec.

4) Fourthly, we propose an effective sparsity-based soft-decoding technique for CS image recovery at the decoder side. Specifically, to alleviate the negative influence of compressed and incomplete measurements, we propose the collaborative sparse representation (CSR) method to explore the high-order correlation among local patches in the process of dictionary learning and sparse decomposition simultaneously. This novel image restoration strategy gives the proposed soft decoder its superior RD performance to other low bit-rates and MD image coding methods.

The rest of the paper is organized as follows. Section II provides an overview our compressive sampling based image coding scheme. In Section III, we first briefly introduce the CS theory, then detail the proposed LRS strategy. In Section IV, the proposed sparsity-based soft decoding technique is described. Section V shows experimental results and discussions about the behavior and properties of the new image coding technique in comparison with others. Section VI concludes the paper.

II. SYSTEM OVERVIEW

We depict the architecture of the proposed compressive sampling based image coding system in Fig. 2. As stated previously, our scheme is not only an asymmetric image codec,

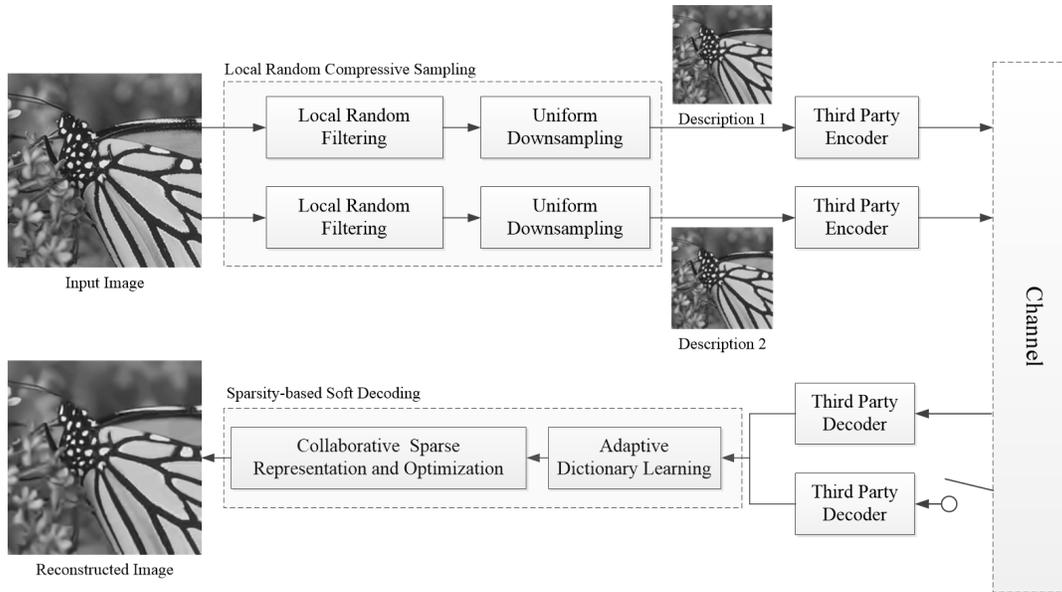


Fig. 2. Block diagram of the proposed LRS-based image coding system. The third party codec can be any image coding standards. In this work, we test JPEG2000 and HEVC-Intra.

but also able to serve as a MDC codec. To avoid unnecessary clutters, the system is illustrated for two descriptions. If both descriptions are received, a high-quality reconstruction can be obtained. While if one of them is lost, a lower-quality, but acceptable, reconstruction can be decoded.

At the encoder side, the input image is pre-filtered by a local binary convolution kernel of random coefficients, and is then uniformly down-sampled to get a more compact representation. Since the down-sampled image has the conventional form of rectangular pixel grid, it can be fed directly to a third party or standard image codec to further remove redundancy. The proposed image compression system has an encoder with lower computational complexity, since only a small fraction of the original image is compressed, while shifting the burden of achieving high coding performance to the decoder side. Such a coding structure is somewhat similar to that of DSC.

The decoder is a cascade of the straightforward hard decoding (which is conventional standard image decoder) of $k \geq 1$ received descriptions and a sparsity-based reconstruction of the transmitted image using all the kWH local random measurements of the image \mathbf{X} , where $W \times H$ is the resolution of the down-sampled images. This hard decoding process prepares a set of kWH local random measurements of \mathbf{X} , and interfaces with a heavy-duty soft decoder that restores \mathbf{X} from these local random measurements. The soft decoding is carried out in the well-known framework of compressive image recovery, which is carried out independently with the number of available descriptions. Therefore, when used as a MDC codec, the proposed system unifies the central decoder and all side decoders. It uses only one MD decoder that suits all possible channel deliveries of the k descriptions.

In the following two sections, we will elaborate the proposed LRS and sparsity-based soft decoding.

III. LOCAL RANDOM COMPRESSIVE SAMPLING

In this section, we first briefly review the CS theory, and then introduce the proposed local random compressive sampling strategy.

A. Compressive Sensing

The emerging CS theory [16], [17] challenges the conventional practice of “oversampling followed by massive dumping”. It has been enthusiastically promoted as a joint sampling and compression approach. The advantages of CS over conventional signal compression techniques are architectural: the CS encoder is made signal independent and computationally inexpensive by shifting the bulk of system complexity to the decoder.

The CS theory reveals the possibility of reconstructing a signal from a small number of random measurements, as long as the signal is sparse in some domain. Denote \mathbf{X} as the image captured by a visual sensor, which can be stacked into a vector $\mathbf{x} \in \mathfrak{R}^N$ according to the lexicographical order. \mathbf{x} is said to be sparse in space Ψ , if the transformation coefficients $\boldsymbol{\alpha} = \Psi^T \mathbf{x}$ are mostly zero or close to zero. The sparsity of \mathbf{x} in Ψ is quantified by the number of significant (nonzero) coefficients K . The CS theory states that \mathbf{x} can be perfectly recovered from $M = O(K \log(N/K))$ random measurements $\mathbf{y} = \Phi \mathbf{x}$, where Φ is a $M \times N$ measurement matrix [17].

The CS recovery of \mathbf{x} from \mathbf{y} can be formulated as the following constrained optimization problem:

$$\min_{\boldsymbol{\alpha}} \|\mathbf{x}\|_1, \quad \text{s.t. } \mathbf{y} = \Phi \mathbf{x}, \quad (1)$$

which can be further written as:

$$\min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_1, \quad \text{s.t. } \mathbf{y} = \Phi \Psi \boldsymbol{\alpha}. \quad (2)$$

After obtaining α by solving (2), the final reconstruction of \mathbf{x} is $\hat{\mathbf{x}} = \Psi\alpha$. Since the matrix $\Phi\Psi$ is rank deficient, the recovery of \mathbf{x} from \mathbf{y} is an ill-posed problem.

In the current task of image coding, the process of quantization will introduce compression noise into CS measurements, that is:

$$\hat{\mathbf{y}} = \Phi\mathbf{x} + \mathbf{n}, \quad (3)$$

where \mathbf{n} is the vector of compression noise. One can relax the equality constraint of (2) to the following formulation:

$$\min_{\alpha} \|\alpha\|_1, \quad \text{s.t.} \quad \|\hat{\mathbf{y}} - \Phi\Psi\alpha\|_2 \leq \mu, \quad (4)$$

where $\mu > 0$ is a tolerance parameter.

One fundamental consideration in employing the above model is the choice of the measurement matrix Φ . The most commonly used family of measurement matrices consists of fully dense random Gaussian/Bernoulli matrices. Their main advantage is that they are universally incoherent with any sparsity bases. What is more remarkable is that they satisfy the Restricted Isometry Property (RIP) with high probability [17]. However, they inherently have one major drawback in practical image compression applications: since the sampled measurements are *globally* linear combination of the original signal, they are completely unstructured. Therefore, it is hard to realize and exploit the statistical redundancy in CS measurements, making it difficult to further compress them.

B. Local Random Filtering

In conventional down-sampling based image coding methods [14], [15], the 2D low-pass Gaussian filtering is exploited to avoid aliasing effect. However, the low-pass filtering discards all high-frequency information beyond the cut-off frequency, leading to limited reconstruction quality (especially the subjective quality) at the decoder side. From the CS theory, we know that perfect reconstruction can still be achieved even though aliasing occurs, provided that the signal is sparse. A natural question arising is: why do we need to throw away the high-frequency information at the pre-filtering stage?

In our method, we pursue more intelligent sparse sampling; a local random filter is employed, keeping not only low-frequency information, but also certain amount of high-frequency information. Certainly, the aliasing problem will occur, as the local random filter is not low-pass, as shown in Fig. 1. Similar to the idea of CS, the aliasing problem can be solved via the proposed collaborative sparsity-based soft-decoding algorithm, which will be elaborated in the next section.

Specifically, LRS is performed with a local binary convolution kernel of size $w \times w$, which generates random measurements \mathbf{Y} of the input image \mathbf{X} . We model the sampling process as a series of inner products against different waveforms $\{\phi_{u,v}\}$

$$\mathbf{Y}_{u,v} = \langle \phi_{u,v}, \mathbf{X}_{u,v} \rangle, \quad (5)$$

where $\phi_{u,v}$ of length w^2 is a random sequence with each element being an *i.i.d* random variable chosen from $\{0, 1\}$. $\mathbf{X}_{u,v}$ of length w^2 is the vector representation of the $w \times w$ window of the original image centered at $(2u, 2v)$. $\mathbf{Y}_{u,v}$ is further normalized to the pixel value interval $[0, 255]$.

C. Uniform Downsampling

The local sampling window goes around the whole image with a preset step, which determines the ratio of down-sampling. Larger sampling step results in a lower-resolution description, which costs lower bit-rate for compression; but at the expense of worse reconstruction quality at the decoder. In this work, in order to achieve good tradeoff between bit-rate and reconstruction quality, the step is set as two pixels so that every other row and every other column of pixels are sampled. By stacking all random measurements in vector \mathbf{y} we can write the above equation in matrix form as:

$$\mathbf{y} = \Phi\mathbf{x}, \quad (6)$$

where Φ is the LRS matrix generated from $\{\phi_{u,v}\}$.

Different from the global random sampling in conventional CS, measurements generated by LRS still preserve local structures of the 2D image signal. Therefore, the measurements constitute a low-resolution image, which can be compressed by any existing standard image codec.

D. Discussion

In this work, we only claim that the proposed LRS matrix is helpful for CS-based image coding applications, in which the compression of measurements is a critical factor affecting the overall RD performance. Measurements generated by LRS matrix preserve local image structures, and thus, can be effectively compressed by a subsequent standard image codec. In contrast, there is little correlation among measurements generated by global dense random matrix, making it difficult to further compress them. However, in general sense, we do not claim that the LRS matrix is better than global ones. In fact, for pure CS sampling and recovery problems, which do not consider measurements compression, local random matrix is suboptimal. This is because LRS requires more measurements to satisfy RIP and lacks the universality compared with global dense random matrices [18].

IV. SPARSITY-BASED SOFT DECODING FOR COMPRESSIVE SENSING RECOVERY

The performance of the proposed LRS-based image coding system is heavily dependent on the capability of designing an effective decoding scheme, which attempts to maximize the quality of the signal decoded. The reconstruction of \mathbf{x} from the set of received hard-decoded image(s) $\hat{\mathbf{y}}$ is an ill-posed inverse problem, which is more difficult than general CS recovery since the measurements are both noisy (due to quantization) and incomplete (due to down-sampling). How to generate robust image representations which can effectively exploit priors in regulating solutions of the inverse problem, is critical for boosting the performance of the proposed soft decoding algorithm.

A. Adaptive Dictionary Learning

One popular technique to incorporate the prior knowledge about images is via a so-called sparsity model, in which an

image is approximated by a sparse linear combination of elements in an appropriately chosen dictionary Ψ :

$$\mathbf{x} = \Psi\boldsymbol{\alpha} + \boldsymbol{\varepsilon}, \quad (7)$$

where $\boldsymbol{\varepsilon}$ is the approximation error. In order to jointly exploit the priors of the random convolution kernels Φ that produces \mathbf{y} and the dictionary Ψ that sparsely represents the input image \mathbf{x} , the soft decoding problem is formulated as

$$\arg \min_{\{\boldsymbol{\alpha}, \Psi\}} \|\widehat{\mathbf{y}} - \Phi\Psi\boldsymbol{\alpha}\|_2^2 + \lambda \|\boldsymbol{\alpha}\|_1, \quad (8)$$

where λ is a Lagrangian multiplier. The above formulation leads to a nonlinear optimization problem. It can be made linear and solved by alternatingly fixing one of $\boldsymbol{\alpha}$ and Ψ and optimizing over the other.

From Fig. 2, it is easy to see that the derived measurements from LRS still keep compact image structure. Therefore, we can learn adaptive patch-based dictionaries from the decoded image(s) directly. Specifically, using $\widehat{\mathbf{y}}$, we build a set of M dictionaries $\{\Psi_m\}_{1 \leq m \leq M}$ in a learning process, in which each Ψ_m is associated with a pattern class of $\sqrt{d} \times \sqrt{d}$ pixel patches. Locally adaptive dictionaries are needed because natural images typically exhibit non-stationary statistics, consisting of many heterogeneous regions of significantly different geometric structures or statistical characteristics. To learn dictionaries $\{\Psi_m\}$, we classify extracted overlapping $\sqrt{d} \times \sqrt{d}$ pixel patches of $\widehat{\mathbf{y}}$ (or multi-scaled version of $\widehat{\mathbf{y}}$) into M groups \mathcal{S}_m , $1 \leq m \leq M$, by k -means clustering [24]–[26]. Considering each group \mathcal{S}_m as a set of samples generated by Ψ_m , we perform the principle component analysis (PCA) on \mathcal{S}_m and let the resulting PCA bases be the words of dictionary Ψ_m .

Within each subset, the informative structures dominate the rare noisy structures introduced by quantization. Thus, by learning the primary atoms of dictionary from the mass of local features within the same subset, we can effectively alleviate the negative influence of noisy features in sparse coding.

B. Collaborative Sparse Representation

After getting the dictionary, the following step is how to derive optimal sparse codes of (8). In conventional standard sparse representation, local patches are considered independently. In our problem, however, this approach is not robust since it neglects the influence of compression noises in sparse decomposition, which are ubiquitous in hard-decoded image(s) due to quantization. Such noises will contaminate the sparse representation and consequently degrade the quality of soft-decoded image. To alleviate the negative influence of compression noises, we propose to exploit the correlation among patches in order to suppress the interference of noises via weakening their response on dictionary atoms.

As stated previously, in the process of dictionary learning, we explicitly divide the extracted patches into different subsets. In each subset, the patches are ensured to be highly correlated, and thus, the produced sparse codes of them should be encouraged to activate on identical dictionary atoms. In this work, we propose a *collaborative sparse representation* strategy, which explicitly introduces a regularization term to

preserve the consistency of sparse codes for similar local patches:

$$\min_{\{\boldsymbol{\alpha}\}} \|\widehat{\mathbf{y}} - \Phi\Psi\boldsymbol{\alpha}\|_2^2 + \lambda \sum_i \|\boldsymbol{\alpha}_i\|_1 + \gamma \sum_{m=1}^M \sum_{i,j \in \mathcal{C}_m} \|\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_j\|_1 \mathbf{W}_{ij}, \quad (9)$$

where λ and γ are regularization parameters; \mathcal{C}_m is the coordinates set of patches in the cluster \mathcal{S}_m ; \mathbf{W}_{ij} measures the similarity between a pair of patches $(\widehat{\mathbf{y}}_i, \widehat{\mathbf{y}}_j)$, which is defined as:

$$\mathbf{W}_{ij} = \begin{cases} \exp\left\{-\frac{\|\widehat{\mathbf{y}}_i - \widehat{\mathbf{y}}_j\|^2}{\sigma^2}\right\}, \sigma > 0 & \text{if } \widehat{\mathbf{y}}_i \text{ and } \widehat{\mathbf{y}}_j \in \mathcal{S}_m \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

where $\widehat{\mathbf{y}}_i$ is a pixel patch in $\widehat{\mathbf{y}}$ centered at pixel location i .

In the new regularization term of (9), ℓ_1 -norm is used to characterize the consistency of sparse codes. Through incorporating the consistency preservation term into the objective function, the sparse codes of similar patches are derived collaboratively. Consequently, the response of noises is dominated by true image features due to their rarity in compressed images.

C. Optimization

Since \mathbf{W}_{ij} becomes zeros for patches not belonging to the same subset, the above optimization process can be performed subset by subset. We use the subset \mathcal{S}_m as an example to show how to get optimal sparse codes. For \mathcal{S}_m , the above objective function can be rewritten as:

$$\arg \min_{\{\boldsymbol{\alpha}_i\}} \left\{ \sum_{i \in \mathcal{C}_m} \|\widehat{\mathbf{y}}_i - \Phi_m \Psi_m \boldsymbol{\alpha}_i\|_2^2 + \lambda \sum_{i \in \mathcal{C}_m} \|\boldsymbol{\alpha}_i\|_1 + \gamma \sum_{i \in \mathcal{C}_m} \sum_{j \in \mathcal{C}_m} \|\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_j\|_1 \mathbf{W}_{ij} \right\} \quad (11)$$

where Φ_m is the reorganized measurement matrix from Φ , which contains the local random convolution kernels of pixels in $\widehat{\mathbf{y}}_i \in \mathcal{S}_m$.

The above objective function involves two ℓ_1 norms, making it difficult to perform optimization. In practical implementation, we relax the objective function into the following form:

$$\arg \min_{\{\boldsymbol{\alpha}_i\}} \left\{ \sum_{i \in \mathcal{C}_m} \|\widehat{\mathbf{y}}_i - \Phi_m \Psi_m \boldsymbol{\alpha}_i\|_2^2 + \lambda \sum_{i \in \mathcal{C}_m} \|\boldsymbol{\alpha}_i\|_1 + \gamma \sum_i \sum_j \|\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_j\|_2^2 \mathbf{R}_{ij} \mathbf{W}_{ij} \right\}. \quad (12)$$

where we replace $\|\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_j\|_1$ with $\|\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_j\|_2^2 \mathbf{R}_{ij}$ and \mathbf{R}_{ij} measures activated atoms response intersection of $\boldsymbol{\alpha}_i$ and $\boldsymbol{\alpha}_j$, which is defined as:

$$\mathbf{R}_{ij} = \frac{\text{sum}(s(\boldsymbol{\alpha}_i) \oplus s(\boldsymbol{\alpha}_j))}{\Gamma}, \quad (13)$$

Here, Γ denotes the cardinality of $\boldsymbol{\alpha}_i$, and $s(\boldsymbol{\alpha}_i)$ is a 0-1 vector, whose element $s_k(\boldsymbol{\alpha}_i) = 1$ if the k -th atom is activated. Also, \oplus represents the operator of exclusive or (XOR).

We further define $\widehat{\mathbf{W}}_{ij} = \mathbf{R}_{ij} \mathbf{W}_{ij}$. Then, the additional regularization term in (12) becomes the well-known graph-Laplacian form [31], [32]. We define the degree

matrix $\mathbf{V} = \text{diag}(v_1, \dots, v_l)$, where $v_i = \sum_j \widehat{\mathbf{W}}_{ij}$, and $\mathbf{L} = \mathbf{V} - \widehat{\mathbf{W}}$ as the Laplacian matrix. The regularization term can then be written as:

$$\sum_i \sum_j \|\alpha_i - \alpha_j\|_2^2 \widehat{\mathbf{W}}_{ij} = \text{Tr}(\mathbf{A}\mathbf{L}\mathbf{A}^T), \quad (14)$$

which can be further expressed as:

$$\text{Tr}(\mathbf{A}\mathbf{L}\mathbf{A}^T) = \text{Tr} \left(\sum_i \sum_j \mathbf{L}_{ij} \alpha_i \alpha_j^T \right) = \sum_i \sum_j \mathbf{L}_{ij} \alpha_i^T \alpha_j, \quad (15)$$

where \mathbf{A} is the sparse codes matrix. With all the above definitions, the optimization problem can be reformulated as:

$$\arg \min_{\{\alpha_i\}} \sum_i \|\widehat{\mathbf{y}}_i - \Phi \Psi \alpha_i\|_2^2 + \lambda \sum_i \|\alpha_i\|_1 + \gamma \sum_i \sum_j \mathbf{L}_{ij} \alpha_i^T \alpha_j. \quad (16)$$

where we drop the subscript of Φ_m and Ψ_m for ease of presentation.

Since computing \mathbf{R}_{ij} involves operations of α_i and α_j , in this work, we adopt an iterative optimization strategy to solve the above objective function. That is, the derived α_i and α_j in the last iteration are used to compute \mathbf{R}_{ij} in the current iteration. Furthermore, in each iteration, instead of directly optimizing all the sparse codes, we optimize each code α_i individually while keeping all the remaining sparse codes $\alpha_j (j \neq i)$ fixed. When optimizing α_i , we can get the following optimization problem:

$$\begin{aligned} & \arg \min_{\alpha_i} f(\alpha_i) \\ & = \arg \min_{\alpha_i} \left\{ \|\widehat{\mathbf{y}}_i - \Phi \Psi \alpha_i\|_2^2 + \gamma L_{ii} \alpha_i^T \alpha_i + \alpha_i^T \mathbf{h}_i \right. \\ & \quad \left. + \lambda \sum_{j=1}^q |\alpha_i^{(j)}| \right\}, \quad (17) \end{aligned}$$

where $\mathbf{h}_i = 2\gamma (\sum_{j \neq i} L_{ij} \alpha_j)$, and $\alpha_i^{(j)}$ is the j -th coefficients of α_i .

The well-known feature-sign search algorithm [29] is employed to solve α_i . Firstly, we define

$$J(\alpha_i) = \|\widehat{\mathbf{y}}_i - \Phi \Psi \alpha_i\|_2^2 + \gamma L_{ii} \alpha_i^T \alpha_i + \alpha_i^T \mathbf{h}_i, \quad (18)$$

which implies

$$f(\alpha_i) = J(\alpha_i) + \lambda \sum_{j=1}^q |\alpha_i^{(j)}|. \quad (19)$$

In non-smooth optimizations, a necessary condition for a parameter vector α_i to be a local minima is that the zero-vector $\mathbf{0}$ is an element of the subdifferential $\partial J(\alpha_i)$, the set containing all subgradients α_i [30]. We define $\nabla_i^{(j)} |\alpha_i|$ as the subdifferentiable value of the j -th coefficient of α_i . If $|\alpha_i^{(j)}| > 0$, the absolute value function $|\alpha_i^{(j)}|$ is differentiable, therefore, $\nabla_i^{(j)} |\alpha_i|$ is given by the $\text{sign}(\alpha_i^{(j)})$. If $\alpha_i^{(j)} = 0$, the subdifferentiable value $\nabla_i^{(j)} |\alpha_i|$ is within the

set $\{-1, 1\}$. Therefore, the optimality conditions for achieving the optimal solution of $f(\alpha_i)$ becomes

$$\begin{cases} \nabla_i^{(j)} J(\alpha_i) + \lambda \text{sign}(\alpha_i^{(j)}) = 0, & \text{if } |\alpha_i^{(j)}| > 0 \\ |\nabla_i^{(j)} J(\alpha_i)| \leq \lambda, & \text{if } \alpha_i^{(j)} = 0 \end{cases} \quad (20)$$

Then, we consider how to select the optimal subgradient $\nabla_i^{(j)} f(\alpha_i)$ when the optimality conditions are violated, i.e., when $|\nabla_i^{(j)} J(\alpha_i)| > \lambda$ if $\alpha_i^{(j)} = 0$. Suppose that $\nabla_i^{(j)} J(\alpha_i) > \lambda$. This means that $\nabla_i^{(j)} f(\alpha_i) > 0$ regardless of the sign of $\alpha_i^{(j)}$. In this case, in order to decrease $f(\alpha_i)$, it is desirable to decrease $\alpha_i^{(j)}$. Since $\alpha_i^{(j)}$ starts at zero, the very first infinitesimal adjustment to $\alpha_i^{(j)}$ will make it negative. Accordingly, we let $\text{sign}(\alpha_i^{(j)}) = -1$. Similarly, if $\nabla_i^{(j)} J(\alpha_i) < -\lambda$, we let $\text{sign}(\alpha_i^{(j)}) = 1$.

To update sparse codes α_i , supposing we know the signs of $\{\alpha_i^{(j)}\}$ at the optimal values, we can remove the ℓ_1 -norm on $\alpha_i^{(j)}$ by replacing each term $|\alpha_i^{(j)}|$ with either $\alpha_i^{(j)}$ if $\alpha_i^{(j)} > 0$; or $-\alpha_i^{(j)}$ if $\alpha_i^{(j)} < 0$; or 0 if $\alpha_i^{(j)} = 0$. Therefore, (17) can be reduced to a standard, unconstrained quadratic programming (QP) optimization problem, which can be solved by the linear system.

In the optimization, we maintain an active set $\mathcal{A} \triangleq \{j | |\alpha_i^{(j)}| = 0, |\nabla_i^{(j)} J(\alpha_i)| > \lambda\}$ for potentially nonzero coefficients and their corresponding signs θ while updating each α_i . Then, it searches for the optimal active set and coefficient signs which minimize the objective function (17). In each activate step, the algorithm uses the zero-value whose violation of the optimality condition $|\nabla_i^{(j)} J(\alpha_i)| > \lambda$ is largest. The algorithm proceeds in a series of feature-sign steps: given a current value for the active set and signs, it computes the analytical solution $\widehat{\alpha}_i^{new}$ of the resulting unconstrained QP optimization; then it updates the solution, the active set and the signs using a discrete line search between the current solution and $\widehat{\alpha}_i^{new}$.

V. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, extensive experimental results are presented to demonstrate the superior performance of the proposed LRS-based image coding technique. As stated previously, the proposed method reduces the data volume by LRS, and hence, permits an encoder with lower computational complexity. Moreover, our method could be used as an effective low bit-rates image coding scheme when the bandwidth of wireless transmission channel is limited. In the case when there is a requirement for robust transmission, our scheme can also serve as a MDC scheme. Therefore, the following comparative study includes three parts: RD performance comparison for low bit-rates, MD image coding, and encoder complexity comparison.

For generality of our comparative study, we select five images of size 256×256 widely used in the literature and one additional image extracted from the first frame of *Foreman* CIF sequence as the test set. These images are shown in Fig. 3. In the paper, we restrict our attention to graylevel images. The proposed scheme can be easily extended to compress color images by independently processing luminance channel Y, and



Fig. 3. Six test images.

chrominance channels U and V . In the following experiments, the CDF 9/7 wavelet transform is used in JPEG2000. There are some parameters that need to be clarified. The size of local random kernel is set to $w = 3$. Note that it is difficult to optimally determine two regularization parameters λ , γ , and the parameter σ^2 in (10). In our work, we empirically set their values. λ is set to 0.01 for all cases. γ is empirically selected from $\{0.001, 0.01, 0.05\}$ for different bit-rates. In addition, σ^2 is fixed to 80. When performing the k -means classification, the cluster number is initialized to 70 for each bit-rate.

A. Performance Comparison of Low Bit-Rates Image Coding

The RD performance of low bit-rates image coding is examined first. In our comparative study, we test eight image compression schemes:

- 1) The JPEG codec.
- 2) The Progressive Scalar Quantization (PSQ) based codec [12], which is a state-of-the-art CS-based image coding scheme.
- 3) The JPEG2000 codec.
- 4) The CADU algorithm coupled with JPEG2000 (CADU-J2K). The CADU algorithm is a state-of-the-art down-sampling-interpolation based image compression technique [15]. JPEG2000 is used to compress the down-sampled images.
- 5) The proposed LRS coupled with JPEG2000 (LRS-J2K).
- 6) The HEVC-Intra codec.
- 7) The CADU algorithm coupled with HEVC-Intra (CADU-HEVC). It is the counterpart of CADU-J2K in which HEVC-Intra is alternatively used to compress down-sampled images.
- 8) The proposed LRS coupled with HEVC-Intra (LRS-HEVC).

The schemes 3), 4) and 5) constitute the JPEG2000-based group, and the last three schemes form the HEVC-based group.

Table I tabulates the PSNR and SSIM results of the compared codecs on six test images, against various bit-rates from 0.1bpp to 0.4bpp. Some table entries for JPEG at very low bit-rates are marked with “-”, since the DCT-based JPEG cannot operate at such low bit-rates. From Table I, we can see that the PSQ algorithm works poorly. Its performance is worse than JPEG2000 and HEVC-Intra by large margin. This observation is consistent with the results in [12]: the PSQ algorithm just matches or slightly outperforms JPEG for some test images.

In the JPEG2000-based compression group, the proposed LRS-J2K method works better than the state-of-the-art

CADU-J2K on all bit-rates of all test images with respect to both PSNR and SSIM criteria. For *Butterfly*, the PSNR gain of LRS-J2K against CADU-J2K is up to 1.02dB (27.38dB vs. 26.36dB), which is achieved when the bit-rate is 0.4bpp. LRS-J2K also outperforms JPEG2000, when the bit-rate is up to 0.4bpp, for test images *Butterfly*, *Leaves*, *Bike*, *Foreman*, and *Monarch*. For *Lena*, while LRS-J2K produces lower PSNR than JPEG2000 when the bit-rate is greater than 0.25bpp, its SSIM performance is better than JPEG2000, implying higher subjective quality. These results are encouraging since JPEG2000 fed with images of the original resolution is widely regarded as an excellent low bit-rates image codec.

In the HEVC-based compression group, the proposed LRS-HEVC also outperforms CADU-HEVC for all cases. For the test images *Butterfly* and *Leaves*, LRS-HEVC achieves better performance than HEVC-Intra when the bit-rate is equal to or greater than 0.25bpp. The LRS-HEVC method achieves up to 0.49dB higher PSNR than HEVC-Intra (22.07dB vs. 21.58dB) when the test image is *Butterfly* and the rate is 0.1bpp. The SSIM performance of LRS-HEVC is more encouraging. For *Leaves*, the SSIM value of LRS-HEVC is higher than that of HEVC-Intra even when the bit-rate is up to 0.4bpp. Note that, HEVC-Intra achieves higher compression efficiency than JPEG2000 at the expense of higher encoder computational complexity, as will be shown in Section V-C. In this paper, we consider resource-deficient wireless visual communications. We recommend to use JPEG2000 as the third-party image codec when the wireless device has limited power, since it achieves good tradeoff between computational complexity and coding efficiency.

Besides the PSNR and SSIM performance, we also report the perceptual quality comparison results. Fig. 4 illustrates the the reconstructed images of compared methods. We omit the results of JPEG and PSQ here, since their perceptual qualities are poor, which can be reflected by their very low SSIM values. In the JPEG2000-based and HEVC-based compression groups, our method achieves the best subjective reconstruction results among all compared methods. The proposed method relies on LRS to preserve parts of high-frequency information, and uses the powerful collaborative sparse representation to obtain high-quality CS image recovery. In each group, the advantage of the proposed method over the other two methods in recovering sharp edges and fine details is significant. Our method preserves edges well and reconstructs the original image with better visual quality, whereas JPEG2000 and HEVC-Intra produce noticeable visual artifacts (e.g., jaggies, ringings and aliasing) in areas of edges and textures.

TABLE I
PSNR (dB) AND SSIM RESULTS FOR DIFFERENT COMPRESSION METHODS

Images	Rate(bpp)	JPEG				PSQ				JPEG2000-based						HEVC-based					
		PSNR		SSIM		PSNR		SSIM		JPEG2000		CADU-J2K		LRS-J2K		HEVC-Intra		CADU-HEVC		LRS-HEVC	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Butterfly	0.10	-	-	17.16	0.4706	19.63	0.6529	19.65	0.6542	20.09	0.7108	21.58	0.7597	21.92	0.7682	22.07	0.7958				
	0.15	-	-	18.79	0.6253	20.96	0.7069	21.27	0.7115	21.69	0.7802	24.13	0.842	23.59	0.8226	24.31	0.8556				
	0.20	-	-	20.11	0.6784	22.45	0.7588	22.69	0.7671	23.16	0.8292	25.27	0.8707	24.62	0.8534	25.41	0.8781				
	0.25	-	-	20.79	0.7022	23.75	0.8031	23.91	0.8088	24.72	0.8669	26.5	0.8918	25.62	0.8778	26.52	0.8966				
	0.30	21.56	0.7203	21.80	0.7341	24.57	0.8250	24.90	0.8347	25.48	0.8788	27.78	0.9101	26.39	0.8959	27.38	0.9108				
	0.35	22.45	0.7465	22.65	0.7571	25.33	0.8461	25.56	0.8570	26.35	0.8901	28.43	0.9187	27.20	0.9173	27.82	0.9183				
	0.40	23.28	0.7737	23.36	0.7787	25.92	0.8621	26.36	0.8763	27.38	0.9069	29.77	0.9318	27.38	0.9211	28.41	0.9261				
Leaves	0.10	-	-	16.41	0.4249	18.99	0.6091	18.93	0.6002	19.10	0.6326	20.57	0.7024	20.65	0.722	20.86	0.7521				
	0.15	-	-	17.58	0.5194	20.41	0.6826	20.43	0.6853	20.93	0.7363	22.58	0.8032	22.37	0.7901	22.68	0.8144				
	0.20	-	-	18.66	0.5612	21.91	0.7507	22.06	0.7566	22.21	0.7709	24.60	0.8647	23.97	0.8475	24.94	0.8830				
	0.25	-	-	19.73	0.6843	23.14	0.8002	23.30	0.8058	23.69	0.8412	26.06	0.9013	25.09	0.8823	26.28	0.9155				
	0.30	20.32	0.7087	20.40	0.7107	23.88	0.8220	24.14	0.829	24.76	0.8728	26.71	0.9102	25.95	0.9019	27.02	0.9228				
	0.35	21.52	0.7509	21.60	0.7529	24.76	0.8474	25.06	0.8559	25.94	0.8990	28.01	0.9312	26.75	0.9195	27.91	0.9341				
	0.40	22.40	0.7765	22.49	0.7775	25.49	0.8638	25.82	0.8765	26.56	0.9109	28.68	0.9405	27.39	0.9341	28.37	0.9412				
Bike	0.10	-	-	17.53	0.3529	19.59	0.4960	19.38	0.4798	19.61	0.4990	21.01	0.5795	20.71	0.5547	20.79	0.5901				
	0.15	-	-	18.70	0.4442	20.83	0.5743	20.86	0.5725	20.88	0.5802	22.03	0.6352	21.77	0.6324	21.89	0.6424				
	0.20	-	-	19.52	0.5131	21.60	0.6290	21.89	0.6361	22.01	0.6433	23.24	0.7037	22.77	0.691	23.08	0.7087				
	0.25	20.01	0.5567	20.07	0.5587	22.28	0.6701	22.61	0.6803	22.74	0.6921	23.85	0.7282	23.60	0.7414	23.48	0.7373				
	0.30	20.94	0.6111	20.98	0.6116	23.23	0.7159	23.45	0.7276	23.62	0.7402	25.24	0.7848	24.30	0.7779	24.15	0.7601				
	0.35	21.68	0.6524	21.72	0.6530	23.77	0.7309	23.92	0.7574	24.11	0.7588	25.87	0.8127	24.81	0.808	24.89	0.8083				
	0.40	22.29	0.6880	22.32	0.6881	24.42	0.7608	24.28	0.7775	24.63	0.7793	26.56	0.8350	25.02	0.8235	25.27	0.8253				
Foreman	0.10	-	-	23.23	0.5787	30.46	0.8456	30.51	0.8449	30.99	0.8627	32.96	0.8874	33.02	0.891	34.24	0.9081				
	0.15	25.08	0.7007	24.60	0.6788	32.17	0.876	32.52	0.8801	33.08	0.8917	34.98	0.9137	34.75	0.9148	35.06	0.9170				
	0.20	28.43	0.7755	28.43	0.7755	33.84	0.8989	34.04	0.9019	34.23	0.9097	36.67	0.9225	35.69	0.9272	36.12	0.9285				
	0.25	30.93	0.8385	30.47	0.8292	34.81	0.9125	35.01	0.9159	35.29	0.9199	37.16	0.9366	36.17	0.935	36.66	0.9373				
	0.30	32.51	0.8734	32.25	0.8682	35.7	0.924	35.65	0.9255	35.84	0.9256	38.35	0.9473	36.54	0.9406	37.30	0.9407				
	0.35	33.40	0.8907	33.21	0.8869	36.56	0.9314	36.12	0.9333	36.92	0.9354	38.98	0.9522	36.88	0.9465	37.51	0.9429				
	0.40	34.21	0.9060	34.21	0.9060	37.29	0.9379	36.49	0.9394	37.41	0.9403	39.61	0.9570	36.98	0.9483	38.15	0.9508				
Lena	0.10	-	-	20.23	0.5434	25.06	0.7151	25.01	0.7132	25.09	0.7224	26.40	0.7599	26.78	0.7817	26.82	0.7809				
	0.15	-	-	22.54	0.5963	26.70	0.7736	26.68	0.7693	26.91	0.7874	28.14	0.814	28.12	0.819	28.25	0.8276				
	0.20	24.16	0.6658	23.19	0.628	27.79	0.8091	27.85	0.8147	27.95	0.8160	29.36	0.8445	29.16	0.853	29.22	0.8487				
	0.25	25.81	0.726	25.10	0.6976	28.91	0.8377	28.74	0.8377	28.94	0.8406	30.63	0.8708	29.71	0.8694	29.96	0.8728				
	0.30	26.88	0.7716	26.37	0.7527	29.81	0.8559	29.35	0.8562	29.76	0.8673	31.97	0.8939	30.10	0.8817	30.60	0.8872				
	0.35	27.70	0.8009	27.33	0.7896	30.61	0.8682	29.95	0.8783	30.34	0.8776	32.57	0.9013	30.50	0.8972	30.99	0.8976				
	0.40	28.61	0.8316	28.33	0.8224	31.34	0.8839	30.28	0.8884	30.68	0.8875	33.33	0.9135	30.69	0.9042	31.32	0.9052				
Monarch	0.10	-	-	18.79	0.5013	21.98	0.6806	22.23	0.6915	22.45	0.7110	24.52	0.7812	24.27	0.7782	24.59	0.7897				
	0.15	-	-	20.31	0.7544	24.04	0.7595	24.04	0.7595	24.38	0.7867	26.41	0.8414	25.79	0.8284	26.17	0.8435				
	0.20	21.13	0.6231	21.12	0.6226	25.40	0.808	25.56	0.8115	26.01	0.8371	28.20	0.8811	27.36	0.8715	27.87	0.8816				
	0.25	22.89	0.6932	21.76	0.6507	26.60	0.8427	26.60	0.8427	27.08	0.8649	29.42	0.9039	28.11	0.8925	28.74	0.8991				
	0.30	24.61	0.7619	23.80	0.7201	27.27	0.8633	27.63	0.8716	28.12	0.8859	30.09	0.9147	28.90	0.9139	29.82	0.9167				
	0.35	25.22	0.784	24.61	0.7619	28.20	0.8873	28.20	0.8873	28.86	0.9004	30.81	0.9250	29.25	0.9241	30.53	0.928				
	0.40	26.25	0.8171	25.75	0.8031	28.81	0.8864	28.62	0.9006	29.49	0.9143	32.25	0.9389	29.41	0.9287	31.10	0.9381				

The CADU method ranks in the middle in visual quality, and it also produces some ghosting and aliasing artifacts along edges. It is worth noting that, in medium bit-rates, even though the PSNR values are lower than HEVC-Intra, our method produces much better subjective results.

B. Performance Comparison of MDC

In this subsection, we evaluate the performance of our method as a MDC scheme. We choose the following two image MDC techniques as the benchmarks: spatial multiplexing MD (SMMD) using low-pass

prefiltering [19], and polyphase down-sampling transform multiplexing (PDTM) [20]. All methods are tested for two balanced descriptions, each of which is coded by JPEG2000 and transmitted over a lossy network independently of the other. The description erasure probability p is the same for the two descriptions.

In the proposed MDC scheme, when considering one description, the RD performance of side decoder is exactly the same as the one reported in the last subsection. Here, the central RD performance (when both descriptions are received) is further provided to demonstrate the superior performance

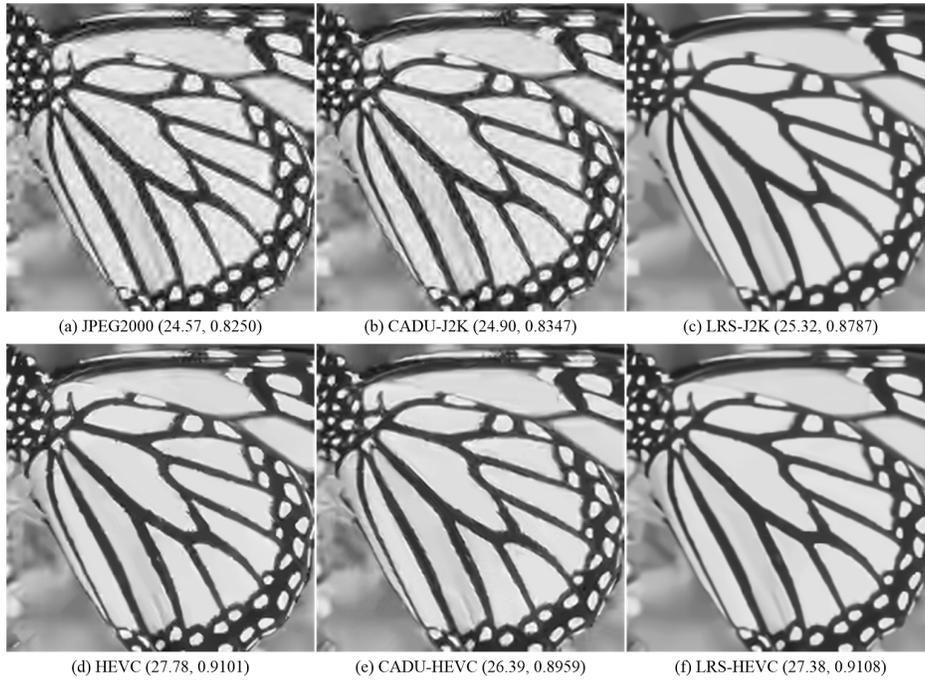


Fig. 4. Subjective performance comparison of JPEG2000-based and HEVC-based groups when the bit rate is 0.3bpp for *Butterfly*. The corresponding PSNR (in dB) and SSIM values are also given. Even though the PSNR is lower than that of HEVC-Intra, the proposed LRF-HEVC achieves better subjective quality than HEVC-Intra.

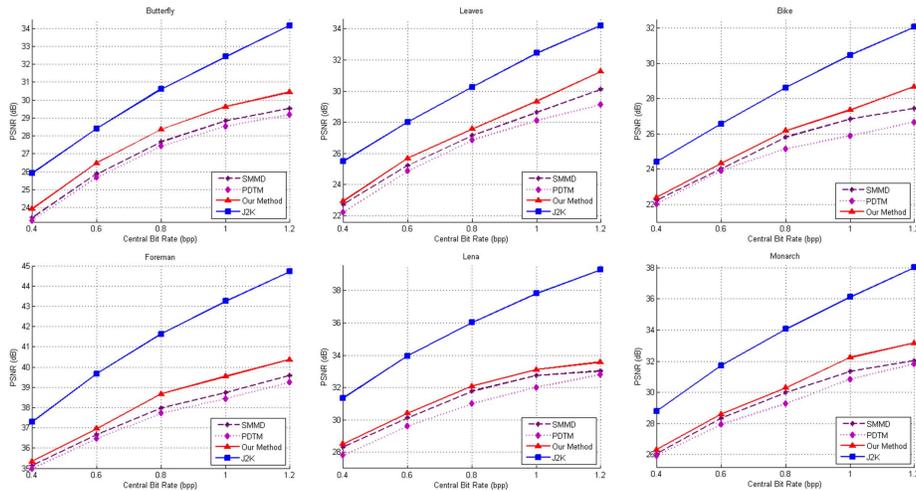


Fig. 5. Comparison of three image MDC methods in PSNR values of central decoders versus the central rate. The central rate is two times of the side rate. The JPEG2000 results are shown as the upper bounds.

of our method. Fig. 5 presents the fidelity-rate (PSNR vs. central rate) curves of three compared methods and JPEG2000. Since two balanced descriptions are considered, the central rate is two times of the side rate. The results on six test images: *Butterfly*, *Leaves*, *Bike*, *Foreman*, *Lena* and *Monarch* are reported. In all test cases, the central decoder of our method outperforms those of SMMD and PDTM. The improvements over SMMD and PDTM are up to 1.2dB and 2dB, respectively.

The MDC algorithms achieve error resilience by introducing redundancy, which is at the price of degraded RD performance. In Fig. 5, the JPEG2000 results are shown as the upper bounds. It should be noted that JPEG2000 does not own the advantages

of MDC in error-robust image transmission. In JPEG2000, the whole image is coded into a single stream, without offering MDC protection against packet losses.

It is also necessary to measure perceptual image quality of MDC. As illustrated in Fig. 6, the proposed method appears to have superior visual quality to its competitors. The image produced by the proposed method is with better preserved image structures and textures. Common compression artifacts observed in SMMD and PDTM, such as jaggies along edges, are substantially suppressed.

Furthermore, we compare the above three MDC methods in term of average distortion versus the description erasure



Fig. 6. Comparison of decoded *Leaves* images when both descriptions are received at total rate 0.6bpp. Left: PDTM; Middle: SMMD; Right: the proposed method.

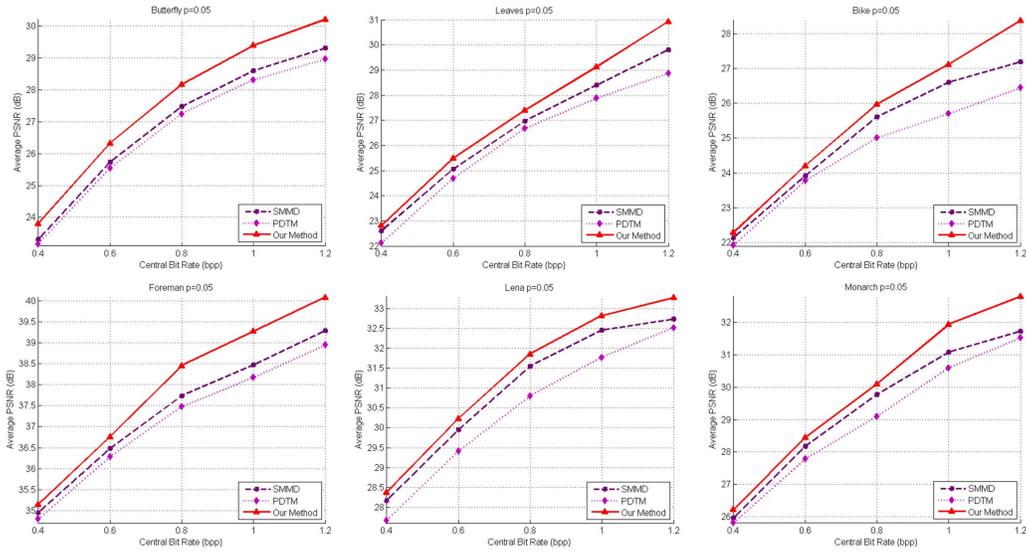


Fig. 7. Comparison of three tested methods in average PSNR (dB) versus the central rate (bpp), when $p = 0.05$.

probability p :

$$\bar{D}(p) = (1 - p)^2 D_c + 2p(1 - p)D_s. \quad (21)$$

Here p denotes the probability of description loss, which is assumed to be equal for all descriptions. D_c and D_s are the central and side distortions, respectively. Two cases $p = 0.05$ and $p = 0.15$ are examined, where we exploit PSNR metric to measure D_c and D_s , and then the average distortion $\bar{D}(p)$ is converted to average PSNR. As shown in Fig. 7 and Fig. 8, the proposed method enjoys fairly large gains in average PSNR against SMMD and PDTM.

As stated previously, at the decoder side, the recovery algorithm is performed in the same way for any number of available descriptions. It would then be interesting to show how the reconstruction performance varies with the number of descriptions. To this end, in Fig. 10, we demonstrate such relationship for the number of descriptions ranging from 1 to 4. As can be seen, the reconstruction performance tends to be linearly improved with respect to the increasing number of descriptions. This is consistent with the expectation of MDC: the more descriptions received, the higher the reconstruction fidelity.

C. Encoder Complexity Comparison

We now show the encoder running time comparison results of our method with JPEG2000 and HEVC-Intra on six test images. In this test, the proposed LRS-based image coding algorithm employs JPEG2000 and HEVC-Intra as the third-party image codec for measurements compression. The results are average running times over seven rate conditions, which are consistent with the test rates shown in Table I. For evaluating the corresponding quality comparisons of compared methods, please refer to Table I.

The compared methods are tested on a typical laptop computer (Intel Core i7 CPU 2.6GHz, 16G Memory, Win10, Matlab R2014a). For each image, we keep the bit-rates of compared methods almost the same. As demonstrated in Fig. 9, when JPEG2000 is used as the third-party image codec, the computational complexity of our LRS-J2K scheme is much lower than that of state-of-the-art image coding standards JPEG2000 and HEVC-Intra. When HEVC-Intra is used as the third-party image codec, the computational complexity of our LRS-HEVC-Intra scheme is also much lower than that of HEVC-Intra, and even lower than that of JPEG2000 on some test images. It should also be noted that the decoder complexity may not be a critical issue in our scheme. This is

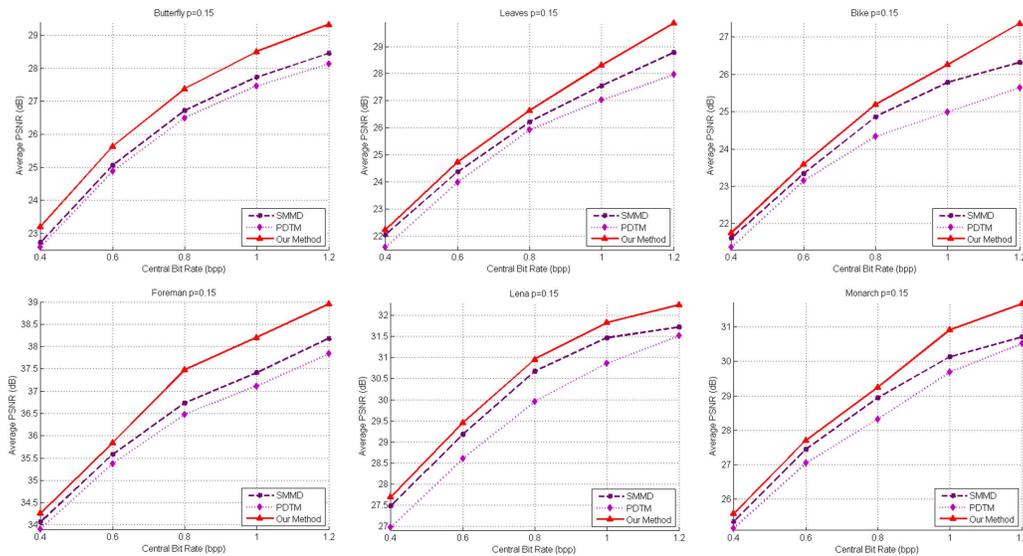


Fig. 8. Comparison of three tested methods in average PSNR (dB) versus the central rate (bpp), when $p = 0.15$.

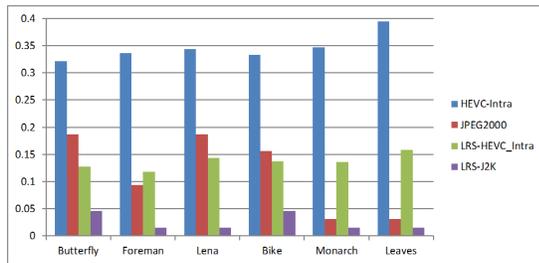


Fig. 9. Encoder complexity comparison.

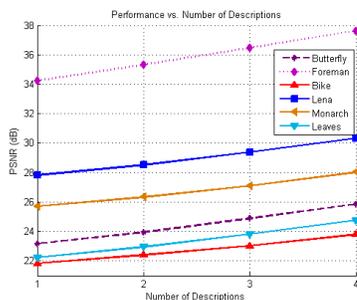


Fig. 10. The relationship between the reconstruction performance and the number of descriptions.

because the heavy duty decoding process could be performed by powerful computers or perceivably in the near future by cloud computing platform.

VI. CONCLUSIONS

We developed an effective CS-based image coding technique of low encoder computational complexity, and a unified decoder when applied as a multiple description coding scheme. In order to recover high-frequency image structures even at low bit-rates, the encoder generates compact image representation(s) by polyphase down-sampling after broad-band

prefiltering of random convolution kernel. The decoder is an image restoration process in a framework of CS recovery, in which the collaborative sparse representation scheme is developed to thoroughly explore high-order correlation among local patches. Experimental results demonstrate that the new image coding method has lower encoder complexity, at the same time, outperforms existing image coding standards at low bit-rates in RD performance, and has unique advantages in perceptual image quality. When applied as a MDC scheme, our method achieves better objective and subjective performance than state-of-the-art MDC algorithms.

REFERENCES

- [1] X. Liu, X. Gao, D. Zhao, J. Zhou, G. Zhai, and W. Gao, "Model-based low bit-rate video coding for resource-deficient wireless visual communication," *Neurocomputing*, vol. 162, pp. 180–190, Aug. 2015.
- [2] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. 22, no. 1, pp. 1–10, Jan. 1976.
- [3] X. Xiong, A. D. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Process. Mag.*, vol. 21, no. 5, pp. 80–94, Sep. 2004.
- [4] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [5] V. K. Goyal, A. K. Fletcher, and S. Rangan, "Compressive sampling and lossy compression," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 48–52, Mar. 2008.
- [6] X. Song, X. Peng, J. Xu, G. Shi, and F. Wu, "Cloud-based distributed image coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 12, pp. 1926–1940, Dec. 2015, doi: 10.1109/TCSVT.2015.2416562.
- [7] M. Alasti, K. Sayrafian-Pour, A. Ephremides, and N. Farvardin, "Multiple description coding in networks with congestion problem," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 891–902, Mar. 2001.
- [8] V. K. Goyal, "Multiple description coding: Compression meets the network," *IEEE Signal Process. Mag.*, vol. 18, no. 5, pp. 74–93, Sep. 2001.
- [9] Y. Wang, A. R. Reibman, and S. Lin, "Multiple description coding for video delivery," *Proc. IEEE*, vol. 93, no. 1, pp. 57–70, Jan. 2005.
- [10] X. Liu, X. Wu, and D. Zhao, "Multiple description image coding with local random measurements," in *Proc. DCC*, pp. 123–132, Mar. 2014.
- [11] S. Mun and J. E. Fowler, "DPCM for quantized block-based compressed sensing of images," in *Proc. 20th Eur. Signal Process. Conf.*, pp. 1424–1428, Aug. 2012.

- [12] L. Wang, X. Wu, and G. Shi, "Binned progressive quantization for compressive sensing," *IEEE Trans. Image Process.*, vol. 21, no. 6, pp. 2980–2990, Jun. 2012.
- [13] C. Deng, W. Lin, B.-S. Lee, and C. T. Lau, "Robust image coding based upon compressive sensing," *IEEE Trans. Multimedia*, vol. 14, no. 2, pp. 278–290, Apr. 2012.
- [14] W. Lin and L. Dong, "Adaptive downsampling to improve image compression at low bit rates," *IEEE Trans. Image Process.*, vol. 15, no. 9, pp. 2513–2521, Sep. 2006.
- [15] X. Wu, X. Zhang, and X. Wang, "Low bit-rate image compression via adaptive down-sampling and constrained least squares upconversion," *IEEE Trans. Image Process.*, vol. 18, no. 3, pp. 552–561, Mar. 2009.
- [16] R. G. Baraniuk, "Compressive sensing," *IEEE Signal Process. Mag.*, vol. 24, no. 4, pp. 118–121, Jul. 2007.
- [17] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.
- [18] A. Eftekhari, H. L. Yap, C. J. Rozell, and M. B. Wakin, "The restricted isometry property for random block diagonal matrices," *Appl. Comput. Harmon. Anal.*, vol. 38, no. 1, pp. 1–31, Jan. 2015.
- [19] X. Zhang and X. Wu, "Standard-compliant multiple description image coding by spatial multiplexing and constrained least-squares restoration," in *Proc. MMSP*, pp. 349–354, Oct. 2008.
- [20] T. Tillo and G. Olmo, "Data-dependent pre- and postprocessing multiple description coding of images," *IEEE Trans. Image Process.*, vol. 16, no. 5, pp. 1269–1280, May 2007.
- [21] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736–3745, Dec. 2006.
- [22] X. Liu, D. Zhai, D. Zhao, and W. Gao, "Image super-resolution via hierarchical and collaborative sparse representation," in *Proc. DCC*, pp. 93–102, Mar. 2013.
- [23] W. Dong, L. Zhang, G. Shi, and X. Wu, "Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization," *IEEE Trans. Image Process.*, vol. 20, no. 7, pp. 1838–1857, Jul. 2011.
- [24] 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.
- [25] I. Tomic and P. Frossard, "Dictionary learning: What is the right representation for my signal?" *IEEE Signal Process. Mag.*, vol. 28, no. 2, pp. 27–38, Mar. 2011.
- [26] A. Coates and A. Y. Ng, "Learning Feature Representations with K-means," in *Neural Networks: Tricks of the Trade* (Lecture Notes in Computer Science), vol. 7700. 2012, pp. 561–580.
- [27] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010.
- [28] S. Gao, I. W.-H. Tsang, and L.-T. Chia, "Laplacian sparse coding, hypergraph Laplacian sparse coding, and applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 92–104, Jan. 2013.
- [29] H. Lee, A. Battle, R. Raina, and A. Y. Ng, "Efficient sparse coding algorithms," in *Proc. NIPS*, pp. 801–808, 2006.
- [30] R. Fletcher, *Practical Methods of Optimization*. Hoboken, NJ, USA: Wiley, 1987.
- [31] X. Liu, D. Zhao, J. Zhou, W. Gao, and H. Sun, "Image interpolation via graph-based Bayesian label propagation," *IEEE Trans. Image Process.*, vol. 23, no. 3, pp. 1084–1096, Mar. 2014.
- [32] X. Liu, D. Zhai, D. Zhao, G. Zhai, and W. Gao, "Progressive image denoising through hybrid graph Laplacian regularization: A unified framework," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1491–1503, Apr. 2014.



Computer Science, HIT.

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, all in computer science. He has been a Project Researcher with the National Institute of Informatics, Tokyo, Japan, since 2014. In 2011, he spent half a year with the Department of Electrical and Computer Engineering, McMaster University, Canada, as a Visiting Student, where he was a Post-Doctoral Fellow from 2012 to 2013. He is currently an Associate Professor with the Department of



Professor with the Department of Computer Science, HIT. Her research interests include machine learning and its application in computer version.

Deming Zhai (M'14) received the B.S., M.S., and Ph.D. (Hons.) degrees from the Harbin Institute of Technology (HIT), Harbin, China, in 2007, 2009, and 2014, respectively, all in computer science. In 2011, she was with the Hong Kong University of Science and Technology, as a Visiting Student. In 2012, she was with the GRASP Laboratory, University of Pennsylvania, PA, USA, as a Visiting Scholar. From 2015 to 2016, she was a Project Researcher with the National Institute of Informatics, Tokyo, Japan. She is currently an Assistant



Professor with the Department of Computer Science, HIT. Her research interests include machine learning and its application in computer version.

Jiantao Zhou (M'11) received the B.Eng. degree from the Department of Electronic Engineering, Dalian University of Technology, Dalian, China, in 2002, the M.Phil. degree from the Department of Radio Engineering, Southeast University, Nanjing, China, in 2005, and the Ph.D. degree from the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong, in 2009. He has held various research positions with the University of Illinois at Urbana-Champaign, the Hong Kong University of Science and Technology, and McMaster University. He is currently an Assistant Professor with the Department of Computer and Information Science, Faculty of Science and Technology, University of Macau. He holds three granted U.S. patents and two granted Chinese patents. His research interests include multimedia security and forensics, and high-fidelity image compression. He was a co-author of a paper that received the best paper award in the IEEE Pacific-Rim Conference on Multimedia in 2007.



Xinfeng Zhang received the B.S. degree in computer science from the 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 Post-Doctoral Fellow with Nanyang Technological University, Singapore. His research interests include image and video processing, image, and video compression.



Debin Zhao received the B.S., M.S., and Ph.D. degrees from the Harbin Institute of Technology, China, in 1985, 1988, and 1998, respectively, all in computer science. He is currently a Professor with the Department of Computer Science, Harbin Institute of Technology. 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.



Wen Gao (M'92–SM'05–F'09) received the Ph.D. degree in electronics engineering from the University of Tokyo, Japan, in 1991. He was a Professor with the Harbin Institute of Technology from 1991 to 1995, and a Professor with the Institute of Computing Technology, Chinese Academy of Sciences, from 1996 to 2006. He is currently a Professor of Computer Science with Peking University. He has published 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, computer vision, multimedia retrieval, multimodal interface, and bioinformatics. He is a member of the Chinese Academy of Engineering, and a fellow of ACM. He chaired a number of prestigious international conferences on multimedia and video signal processing, such as the IEEE ICME 2007, ACM Multimedia 2009, and the IEEE ISCAS 2013, and also served on the advisory and technical committees of numerous professional organizations.

Prof. Gao served on the Editorial Board for several journals, such as the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the IEEE TRANSACTIONS ON MULTIMEDIA, the IEEE TRANSACTIONS ON AUTONOMOUS MENTAL DEVELOPMENT, *EURASIP Journal of Image Communications*, and *Journal of Visual Communication and Image Representation*.