Directional Block Compressed Sensing for Image Coding

Lei Liu[1], Anhong Wang[2*], Kongfen Zhu[2], Chunyu Lin[1], and Yao Zhao[1†]

[1]Institute of Information Science
Beijing Jiaotong University
Beijing, China
[†]e-mail: yzhao@bjtu.edu.cn

[2]Institute of Digital Media and Communication
Taiyuan University of Science and Technology
Taiyuan, China
[*]e-mail: wah_ty@yahoo.com.cn

Abstract—**Compared with traditional *Nyquist* sampling, compressed sensing (CS) enables a highly precise reconstruction of the signal from fewer measurements, suggesting great potential for efficient and simplistic data acquisition. In this paper, we propose a directional block-based compressed sensing (DBCS) scheme for image coding, where the directionalities inherently exhibited within image blocks are exploited as the "*a priori*" information. The image block is first directionally scanned following the dominating direction of its edges/textures. Then the vectorized image block is sampled by a block-based compressed sensing (BCS) method. At the decoder, each image block is recovered and then rearranged by the corresponding inverse-scan to obtain the recovered image. Experimental results show that the proposed DBCS scheme outperforms BCS due to the exploitation of the directional information within image blocks.**

I. Introduction

Nearly all the current signal acquisition systems such as digital cameras employ the sample-then-compress framework, where the raw data (suppose with length N) are first obtained by the *Nyquist* sampling and then compressed by encoding the values and locations of the K ($K \ll N$) largest coefficients in terms of a certain transform basis. However, this framework suffers from three inherent inefficiencies [1]. First, the length of the raw data (N samples) is much larger than that of its dense representation (K coefficients). Second, after computation of N transform coefficients, ($N - K$) of them need to be discarded, which makes a vast quantity of computation waste. Third, an overhead is introduced by encoding the locations of the largest coefficients.

To address these inefficiencies, compressed sensing (CS) [1-4] provides a new signal processing paradigm which attracts a widespread interest in the signal processing community. It allows simultaneously signal sampling and compressing at a sub-*Nyquist* rate via a linear random projection under certain conditions, and offers a highly precise reconstruction of the original signal. According to CS principles, fewer sampling data are needed and the digital data acquisition is much simpler, suggesting great efficiency and simplicity in signal processing.

When applied to 2-dimensional (2-D) image, however, several challenges make CS very impractical [5, 6], e.g., it is not suitable for the real-time application as the entire scene needs to be sensed at once, the random projection operator requires a huge memory and the recovery process is very complex. To address these challenges, several fast algorithms [7-9] have been proposed to speed up the CS recovery process. A block-based compressed sensing (BCS) scheme coupled with a smoothed projection-driven Landweber (SPL) recovery is also proposed [5], which is more convenient for real-time applications because of the efficient measurement operation and fast reconstruction. In [6], some directional transforms are applied to the BCS-SPL iterations with a smoothing and a thresholding operation so as to encourage the smoothness and sparseness, thus resulting in a further improved performance.

Meanwhile, several attempts to better preserve the directional information within image blocks are triggered by the fact that many image blocks contain different oriented edges and/or textures [10]. The directionality may be a big thing for the future image/video coding standards with the development of new directional transforms (such as that in [11]). To get better coding performance, the directional information within images needs to be taken more into account. Therefore in this paper, a directional block-based compressed sensing (DBCS) scheme is proposed to incorporate the directionality into CS paradigm. The directionality within an image block is exploited by a selected directional scan mode, which arranges the block into a vectorized version along the dominating direction of its edges/textures. Then the vectorized block is sampled by the BCS method. At the decoder, the image is reconstructed from the measurements with the inverse-scan mode. Two CS reconstruction algorithms, BCS-SPL [6] and a minimum total variation (min-TV) [2, 12] optimization algorithm, are utilized here with some necessary modifications. The proposed DBCS scheme maintains the merits of BCS and leads to a superiority over BCS.

This work was supported in part by NSFC (No. 61073142, 61272262, 61210006), ICP-SX (2011081055), SXNSF (2012011014-3).

II. RELATED WORKS

A. Compressed Sensing

Suppose a given real value discrete signal $x \in \mathrm{R}^N$ with length N can be represented in terms of a basis matrix $\Psi \in \mathrm{R}^{N\times N}$, i.e., $x = \Psi\theta$, where $\theta \in \mathrm{R}^N$ are the coefficients of x in the Ψ domain. The signal x is K-sparse if θ can be well approximated by only $K \ll N$ largest entries. According to the CS principles, the K-sparse signal x can be accurately reconstructed from $M \geq K\log(N/K)$ measurements obtained through a linear and non-adaptive projection [2-4]:

$$y = \Phi x = \Phi(\Psi\theta)\ , \tag{1}$$

where $y \in \mathrm{R}^M$ denotes the measurement vector, and $\Phi \in \mathrm{R}^{M\times N}$ is a measurement matrix which is incoherent with Ψ. The reconstruction can be formulated as an l_1 optimization problem by solving:

$$\min \frac{1}{2}\|\Phi\Psi\theta - y\|_2^2 + \tau\|\theta\|_1\ , \tag{2}$$

where the l_1-norm term promotes the sparsity of the solution in Ψ domain and the l_2-norm term guarantees the fidelity between the solution and the measurements.

For 2-D image sensing, a BCS scheme is proposed in [5, 6], where an image is first divided into $B\times B$ blocks and each block is sampled using a much smaller measurement matrix. Assume that each image block x_i is taken by M_B measurements:

$$y_i = \Phi_B x_i, \tag{3}$$

where Φ_B is an $M_B\times B^2$ orthonormal i.i.d Gaussian matrix with $M_B = \left\lfloor \frac{M}{N}\cdot B^2 \right\rfloor$ (M/N denotes the measurement rate (MR)), and x_i is the vectorized-version of the i-th image block through raster scan. The equivalent measurement matrix for the entire image in (1) is then a block-diagonal one:

$$\Phi = \begin{bmatrix} \Phi_B & & & \\ & \Phi_B & & \\ & & \ddots & \\ & & & \Phi_B \end{bmatrix}. \tag{4}$$

In [6], the BCS reconstruction is deployed in some transform domain (e.g., contourlet transform, dual-tree DWT, DWT and DCT) with the SPL iterations involving a Wiener filter to eliminate blocking artifacts and a thresholding to enforce the sparsity of the solution.

An alternative algorithm to reconstruct an image is to minimize the total variation (min-TV) with quadratic constraints, which finds the solution to the optimization problem [2, 12]:

$$\min \|x\|_{TV} \quad s.t. \quad \|\Phi x - y\|_2 \leq \epsilon\ , \tag{5}$$

where $\|x\|_{TV} = \sum_{i,j}\sqrt{(x_{i+1,j} - x_{i,j})^2 + (x_{i,j+1} - x_{i,j})^2}$ is the total variation of x, i.e., the sum of the magnitudes of the discrete gradient at every point, and ϵ is a parameter associated with the noise level. Compared to (2), the min-TV based reconstruction usually provides a better performance but suffers huge computations.

B. The Exploitation of Directionality in Image/Video Coding

According to the fact that directional edges/textures are ubiquitous within many image blocks, the most recent video coding standard H.264/MPEG4-Part 10 [13] has developed 8 directional prediction modes in the intra prediction coding for 4 × 4 blocks; and the next generation video coding standard – high efficiency video coding (HEVC) [14], which is still under formulation, provides up to 33 directional prediction modes for each prediction unit (PU) (the total number of available prediction modes depends on the size of the corresponding PU). In [11], a directional discrete cosine transform (DDCT) framework is proposed following the directional modes adopted in H.264/AVC, in which the first transform of the conventional 2-D DCT is modified following the dominating direction of an image block. Then the produced coefficients are arranged appropriately to facilitate the second transform which is a horizontal one. By exploiting the directionality within each image block, the DDCT framework achieves remarkable improvement with respect to the rate-distortion performance.

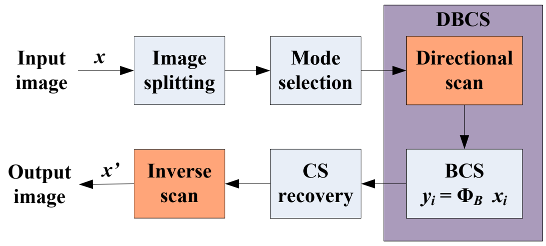

Fig. 1. The block diagram of the proposed scheme.

III. THE PROPOSED DBCS SCHEME

In the BCS scheme [5, 6], the raster scanning for each image block proceeds only along the vertical direction, which however ignores the diversity of directions in image blocks. In this paper, we try to use some directional modes to conduct directional scanning for an image block along its dominating direction. We believe that the directionally-scanned vectorized block may have a more sparse representation under certain transform basis (such as DCT), which will achieve better CS reconstruction.

The scheme of the DBCS-based image coding is shown in Fig. 1. The whole image is first split into some nonoverlapped blocks and each block is scanned according to a selected directional scan mode, and then, a random measurement (sampling) operation like that in BCS method is applied on the vectorized image block. At the decoder, the original image is recovered by using the BCS recovery and an inverse-scan operation, which will be described later.

A. Directional Scan Mode for Image Blocks

The directional scan modes for 8 × 8 blocks are depicted in Fig. 2 (Mode 0 and Mode 1, the vertical and horizontal modes are not included here, and Mode 2, the planar mode, is not considered in this paper). These directional modes are defined as those in the (intra) directional predictions in H.264/AVC [13], and Mode 3 to Mode 8 are named as diagonal down-left, diagonal down-right, vertical-right, horizontal-down, vertical-left and horizontal-up, respectively. As can be seen, the initial points for Modes 4 to Mode 8 are shifted to avoid skips when the block is scanned. For the 16 × 16 and 32 × 32 blocks, 16 and 32 directional scan modes are used following the intra

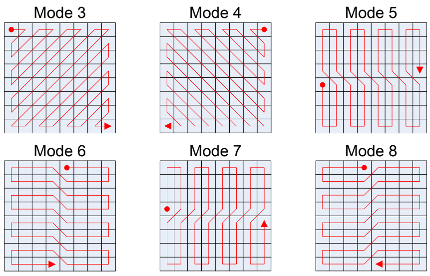

Fig. 2. Directional scan modes for 8×8 blocks.

prediction directions adopted in the HEVC [14], respectively.

After the directional scan modes are determined, the *best* directional mode which follows the dominating direction of an image block is firstly selected, and then the block is arranged into a column-vector by the directional scan. At last, the directionally-scanned and vectorized image block is sampled as that in (3).

It is important to notice that, in a practical image CS scenario, the *a priori* information about the directionality seems to be not available because we cannot get any information about an image before it has been sensed. However, if in a video scenario, it is possible to draw an estimate of a frame before sensing it and then the approximate information is available *a priori*. Here, we ignore this issue and just focus on demonstrating the effectiveness of utilizing directionality in CS image applications.

B. Scan Mode Selection for Each Block

It is not our concern about how to identify the dominating direction of an image block, so we adopt an exhaustive method to test every mode for a block and select the one with highest recovery quality. To this end, all eight modes are used to scan each block, respectively, resulting in eight vectorized versions, then each version is sampled as in (3) with the same sampling operator and recovered using a CS recovery algorithm. The one with highest recovery quality is selected as the best mode for that block.

Obviously, it is very expensive computationally because of the exhaustive selection of the best mode. Moreover, it needs some extra bits to denote the best mode for each block, i.e., 8 × 8 blocks need 3 bits to denote 8 modes, 16 × 16 blocks need 4 bits to denote 16 modes and 32 × 32 blocks need 5 bits to denote 32 modes, respectively. However, the extra bits take a quite tiny portion of the *MR*, i.e., $3/(8\times8)/8 = 5.9\times10^{-3}$, $4/(16\times16)/8 = 2.0\times10^{-3}$ and $5/(32\times32)/8 = 6.1\times10^{-4}$, respectively.

C. The Reconstruction of DBCS

As mentioned previously, two CS recovery algorithms, BCS-SPL [6] and min-TV optimization [2, 12], are adopted in the proposed scheme. Since the measurements are taken from directionally scanned blocks, some necessary modifications have to be done in the recovery algorithms. As to BCS-SPL, a Wiener filter smoothes the outputs of each SPL iteration in the pixel domain, so the inverse scan must be conducted on the current iteration outputs before the smoothing in order to align the outputs with the original-order image; then, after the Wiener filtering, the corresponding directional scan should be applied on the filtered values in order to align with the measurements. On the other hand, for the min-TV based recovery, the smoothness is implicit in the gradient domain which also ensures the sparseness of the solutions; this will naturally benefit greatly from the directionally scanned blocks whose elements are best aligned with each other. In the end, the restored image is obtained by applying the inverse scan on the final solution.

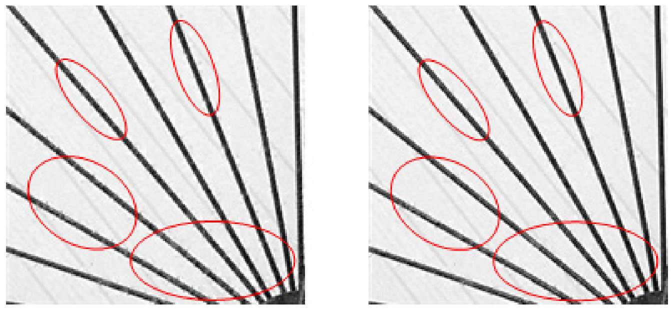

Fig. 3. The subjective comparison between BCS and DBCS for *Spoke* (MR = 0.5, only an enlarged portion is shown here). Left: BCS-TV; Right: DBCS-TV.

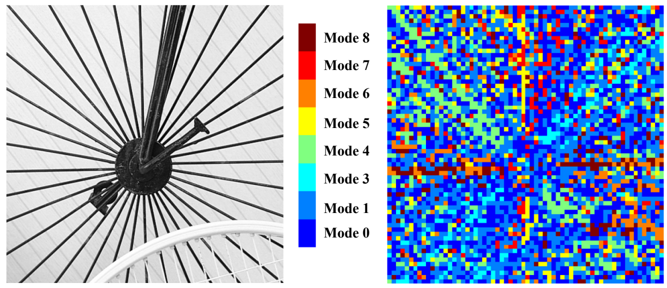

Fig. 4. The distribution of directional modes for *Spoke* with DBCS-TV. Left: original; Right: the directional modes selected in DBCS-TV.

IV. Experimental Results

To validate the performance of the proposed scheme, we compare it with BCS-SPL [6] in the DCT domain; and also with BCS-TV, where BCS sampling is coupled with the min-TV recovery. We refer to the implementations of the proposed scheme as DBCS-SPL and DBCS-TV, respectively. Two 256 × 256 images (*Lena* and *Barbara*) and two 512 × 512 images (*Straw* and *Spoke*) are tested in the experiments.

Table I tabulates the performance (PSNR versus *MR*) comparisons between BCS and DBCS schemes (the overhead of DBCS is not included here). As can be seen, only a very marginal improvement is achieved by using the SPL reconstruction while the TV reconstruction gets a quite remarkable gain. For blocks with size 16 × 16 and 32 × 32, the performance increments are not obvious compared with that of blocks with size 8 × 8. This may because that the selected directional scan modes cannot match the real edges/textures of image blocks with large size. When the block size is 8 × 8, Fig. 3 gives a visual comparison which shows that the proposed scheme achieves significantly better quality than BCS at the regions contain directional edges (see the areas marked by the red ellipse circles); and Fig. 4 shows the directional scan modes distributed over *Spoke*. It can be seen in Fig. 4 that selected modes quite match the regions that contain directional edges, especially at the diagonal directions.

V. Conclusions

This paper proposes a directional block-based compressed sensing (DBCS) scheme which exploits the directionality within image blocks as the *a priori* information. Some directional scan modes are incorporated to directionally scan the image blocks along their dominating directional edges/ textures. The scanned blocks are then sampled using a BCS method. The image is reconstructed by the measurements from directionally-scanned blocks and the corresponding mode information. Experimental results show that, compared with BCS, the proposed scheme achieves remarkable improvement in both PSNR and visual quality, especially in the case of small block size. However, there are still some issues to be considered in the future, e.g., the analysis of the sparsity of the directionally-scanned blocks under some transform basis and the precise definition of directional scan modes to follow the edges/textures within image blocks with a big size.

Acknowledgment

Lei Liu thanks Prof. Bing Zeng for his helpful advices and suggestions.

References

[1] R. G. Baraniuk, "Compressive sensing [lecture notes]," IEEE Signal Process. Magazine, vol. 24, no. 4, pp. 118-121, July 2007.

[2] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," IEEE Trans. Inform. Theory, vol. 52, no. 2, pp. 489-509, Feb. 2006.

[3] D. Donoho, "Compressed sensing," IEEE Trans. Inform. Theory, vol. 52, no. 4, pp. 1289-1306, Apr. 2006.

[4] E. Candès and M. Wakin, "An introduction to compressive sampling," IEEE Signal Processing Magazine, vol. 25, no. 2, pp. 21-30, Mar. 2008.

[5] L. Gan, "Block compressed sensing of natural images," Proc. Int. Conf. Digital Signal Process., Cardiff, UK, pp. 403-406, July 2007.

[6] S. Mun and J. E. Fowler, "Block compressed sensing of images using directional transforms," Proc. Int. Conf. Image Process., Cairo, Egypt, pp. 3021-3024, Nov. 2009.

[7] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," IEEE J. Sel. Topics Signal Process., vol. 1, no. 4, pp. 586-579, Dec. 2007.

[8] T. T. Do, L. Gan, N. Nguyen, and T. D. Tran, "Sparsity adaptive matching pursuit algorithm for practical compressed sensing," Proc. of the 42th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, California, pp. 581–587, October 2008.

[9] J. Haupt and R. Nowak, "Signal reconstruction from noisy random projects," IEEE Trans. Inform. Theory, vol. 52, no. 9, pp. 4036-4048, Sep. 2006.

[10] J. Xu, B. Zeng, and F. Wu, "An overview of directional transforms in image coding," Proc. Int. Symp. Circuits and Systems, Paris, France, pp. 3036-3039, May 2010.

[11] B. Zeng and J.-J. Fu, "Directional discrete cosine transforms—A new framework for image coding," IEEE Trans. Circ. Syst. for Video Technology, vol. 18, no. 3, pp. 305-313, Mar. 2008.

[12] E. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," Comm. Pure and Applied Math., vol. 59, no. 8, pp. 1207–1223, August 2006.

[13] ITU-T Rec. H.264 and ISO/IEC 14496-10 (AVC), "Advanced video coding for generic audiovisual services", Mar. 2005.

[14] JCT-VC, "HM4: High Efficiency Video Coding (HEVC) Test Model 4 Encoder Description," JCTVC-F802, Torino, IT, July 2011.

Table I. PSNR comparative results of DBCS and BCS (in dB)

MR		0.1	0.2	0.3	0.4	0.5
Block Size (8×8)						
Lena (256×256)	BCS-SPL	23.06	24.09	27.78	29.93	31.23
	DBCS-SPL	23.37	25.60	28.39	30.21	31.61
	BCS-TV	21.90	25.03	27.26	29.03	30.77
	DBCS-TV	23.70	26.32	28.36	30.15	31.97
Barbara (256×256)	BCS-SPL	19.77	21.11	22.11	23.16	24.18
	DBCS-SPL	19.91	21.55	22.43	23.54	24.58
	BCS-TV	18.71	20.12	21.09	21.87	22.66
	DBCS-TV	20.17	21.33	22.23	23.15	24.11
Straw (512×512)	BCS-SPL	12.52	14.02	15.25	16.36	17.52
	DBCS-SPL	13.08	15.21	16.31	17.82	19.11
	BCS-TV	10.74	13.87	15.57	16.90	18.27
	DBCS-TV	12.72	14.78	16.36	17.91	19.38
Spoke (512×512)	BCS-SPL	16.26	19.85	22.33	24.48	26.25
	DBCS-SPL	16.30	20.33	22.46	24.52	26.42
	BCS-TV	16.20	19.30	21.68	24.58	26.42
	DBCS-TV	17.85	21.55	23.44	26.04	28.20
Block Size (16×16)						
Lena (256×256)	BCS-SPL	23.08	26.23	28.33	29.83	31.46
	DBCS-SPL	23.22	26.25	28.47	30.07	31.65
	BCS-TV	23.35	26.38	28.49	30.26	31.99
	DBCS-TV	24.09	26.97	28.93	30.70	32.38
Barbara (256×256)	BCS-SPL	19.36	21.01	22.22	23.32	24.49
	DBCS-SPL	19.49	21.19	22.35	23.48	24.76
	BCS-TV	19.32	20.62	21.50	22.25	22.97
	DBCS-TV	19.97	21.16	22.10	22.90	23.91
Straw (512×512)	BCS-SPL	12.16	13.71	15.04	16.31	17.71
	DBCS-SPL	12.68	14.56	16.16	17.54	19.08
	BCS-TV	12.26	14.16	15.75	17.19	18.80
	DBCS-TV	12.75	14.51	16.15	17.55	19.19
Spoke (512×512)	BCS-SPL	16.72	20.43	23.37	25.55	27.62
	DBCS-SPL	16.11	20.49	23.62	25.95	28.09
	BCS-TV	16.99	21.50	24.71	27.08	29.45
	DBCS-TV	18.01	22.49	25.63	27.89	30.38
Block Size (32×32)						
Lena (256×256)	BCS-SPL	23.49	26.26	28.11	29.73	31.28
	DBCS-SPL	23.92	26.42	28.40	30.12	31.71
	BCS-TV	24.61	27.42	29.34	31.18	32.81
	DBCS-TV	24.90	27.61	29.48	31.32	33.14
Barbara (256×256)	BCS-SPL	19.67	21.06	22.35	23.49	24.66
	DBCS-SPL	19.96	21.13	22.40	23.69	24.92
	BCS-TV	19.87	20.88	21.70	22.45	23.30
	DBCS-TV	19.99	21.14	21.91	22.74	23.70
Straw (512×512)	BCS-SPL	12.48	13.89	15.29	16.65	18.02
	DBCS-SPL	13.06	14.68	16.27	17.75	19.22
	BCS-TV	12.73	14.46	15.99	17.49	19.03
	DBCS-TV	12.90	14.63	16.18	17.68	19.26
Spoke (512×512)	BCS-SPL	17.10	20.77	23.31	25.54	27.41
	DBCS-SPL	17.90	21.16	23.95	26.33	27.53
	BCS-TV	19.24	23.79	26.67	28.98	31.48
	DBCS-TV	19.81	24.17	27.02	29.32	31.75