Decoding of purely compressed-sensed video

Ying Liu, Ming Li, and Dimitris A. Pados

Department of Electrical Engineering, State University of New York at Buffalo,
Buffalo, NY 14260

ABSTRACT

We consider a video acquisition system where motion imagery is captured only by direct compressive sampling (CS) without any other form of intelligent encoding/processing. In this context, the burden of quality video sequence reconstruction falls solely on the decoder/player side. We describe a video CS decoding method that implicitly incorporates motion estimation via sliding-window sparsity-aware recovery from locally estimated Karhunen-Loeve bases. Experiments presented herein illustrate and support these developments.

Keywords: Compressed sensing, compressive sampling, Karhunen-Loeve basis, video, motion estimation, motion imagery, Nyquist theorem, sparse signals.

1. INTRODUCTION

Conventional signal acquisition schemes follow the general Nyquist/Shannon sampling theory: To reconstruct a signal without error, the sampling rate must be at least twice as much as the highest frequency of the signal. Compressive sampling (CS), also referred to as compressed sensing, is an emerging bulk of work that deals with sub-Nyquist sampling of *sparse signals* of interest [1]-[3]. Rather than collecting an entire Nyquist ensemble of signal samples, CS can reconstruct sparse signals from a small number of (random [3] or deterministic [4]) linear measurements via convex optimization [5], linear regression [6],[7], or greedy recovery algorithms [8].

A somewhat extreme example of a CS application that has attracted much interest is the "single-pixel camera" architecture [9] where a still image can be produced from significantly fewer captured measurements than the number of desired/reconstructed image pixels. Arguably, a natural highly desirable next-step development may be compressive video streaming. In this present work, we consider a video transmission system where the transmitter/encoder performs nothing more than compressed sensing acquisition without the benefits of the familiar sophisticated forms of video encoding. Such a set-up may be of particular interest, for example, in problems that involve large wireless multimedia networks of primitive low-complexity, low-cost video sensors. In such a case, the burden of quality video reconstruction falls solely on the receiver/decoder side.

The quality of the reconstructed video is determined by the number of collected measurements, which, based on CS principles, should be proportional to the sparsity level of the signal. Therefore, the challenge of implementing a well-compressed and well-reconstructed CS-based video streaming system rests on developing effective sparse representations and corresponding video recovery algorithms. Several important methods for CS video recovery have already been proposed, each relying on a different sparse representation. An intuitive (JPEG-motivated) approach is to independently recover each frame using the 2-dimensional discrete cosine transform (2D-DCT) [10] or a 2-dimensional discrete wavelet transform (2D-DWT). To enhance sparsity by exploiting correlations among successive frames, several frames can be jointly recovered under a 3D-DWT [11] or 2D-DWT applied on inter-frame difference data [12].

In standard video compression technology around us, effective encoder-based motion estimation (ME) is a defining matter in the feasibility and success of digital video. In the case of CS-only video acquisition that we study in this paper, ME can be exploited at the receiver/decoder side only. In current approaches [13],[14], a video sequence is divided into key frames and CS frames. While each key frame is reconstructed individually

---

Further author information: (Send correspondence to D.A.P)
Y.L.: E-mail: yl72@buffalo.edu, Telephone: 1 716 645 1207
M.L.: E-mail: mingli@buffalo.edu, Telephone: 1 716 645 1207
D.A.P: E-mail: pados@buffalo.edu, Telephone: 1 716 645 1150

using a fixed basis (e.g., 2D-DWT or 2D-DCT), each CS frame is reconstructed conditionally using an adaptively generated basis from adjacent already reconstructed key frames.

In this work, we propose a new sparsity-aware video decoding algorithm for compressive video streaming systems to exploit inter-frame similarities and pursue most efficient and effective utilization of all available measurements. For each video frame, we operate block-by-block and recover each block using a Karhunen-Loève transform (KLT) basis adaptively generated/estimated from previously reconstructed reference frame(s) defined in a fixed-width sliding window manner. The scheme essentially implements motion estimation and compensation at the decoder by sparsity-aware reconstruction using inter-frame KL basis estimation.

The rest of the paper is organized as follows. In Section 2, we briefly review the CS principles that motivate our compressive video streaming system. In Section 3, the proposed sliding-window sparsity-aware video decoding algorithm is described in detail. Some experimental results are presented and analyzed in Section 4 and, finally, a few conclusions are drawn in Section 5.

2. COMPRESSIVE SAMPLING BACKGROUND AND FORMULATION

In this section we briefly review the CS principles for signal acquisition and recovery that are pertinent to our CS video streaming problem. A signal vector $\mathbf{x} \in \mathbb{R}^N$ can be expanded/represented by an orthonormal basis $\mathbf{\Psi} \in \mathbb{R}^{N \times N}$ in the form of $\mathbf{x} = \mathbf{\Psi}\mathbf{s}$. If the coefficients $\mathbf{s} \in \mathbb{R}^N$ have at most k non-zero components, we call $\mathbf{x}$ a k-sparse signal with respect to $\mathbf{\Psi}$. Many natural signals -images most notably- can be represented as a sparse signal in an appropriate basis.

Traditional approaches to sampling signals follow the Nyquist/Shannon theorem by which the sampling rate must be at least twice the maximum frequency present in the signal. CS emerges as an acquisition framework under which sparse signals can be recovered from far fewer samples or measurements than Nyquist. With a linear measurement matrix $\mathbf{\Phi}_{P \times N}, P \ll N$, CS measurements of a k-sparse signal $\mathbf{x}$ are collected in the form of

$$\mathbf{y} = \mathbf{\Phi}\mathbf{x} = \mathbf{\Phi}\mathbf{\Psi}\mathbf{s}. \tag{1}$$

If the product of the measurement matrix $\mathbf{\Phi}$ and the basis matrix $\mathbf{\Psi}$, $\mathbf{A} \triangleq \mathbf{\Phi}\mathbf{\Psi}$, satisfies the Restricted Isometry Property (RIP) [3], then the sparse coefficient vector $\mathbf{s}$ can be accurately recovered via the following linear program

$$\widehat{\mathbf{s}} = \arg \min_{\widetilde{\mathbf{s}}} \|\widetilde{\mathbf{s}}\|_{\ell_1} \qquad \text{subject to} \quad \mathbf{y} = \mathbf{\Phi}\mathbf{\Psi}\widetilde{\mathbf{s}}. \tag{2}$$

Afterwards, the signal of interest $\mathbf{x}$ can be reconstructed by

$$\widehat{\mathbf{x}} = \mathbf{\Psi}\widehat{\mathbf{s}}. \tag{3}$$

In most practical situations, $\mathbf{x}$ is not exactly sparse but approximately sparse and measurements may be corrupted by noise. Then, the CS acquisition/compression procedure can be formulated as

$$\mathbf{y} = \mathbf{\Phi}\mathbf{\Psi}\mathbf{s} + \mathbf{e} \tag{4}$$

where $\mathbf{e}$ is the unknown noise bounded by a known power amount $\|\mathbf{e}\|_{\ell_2} \leq \epsilon$. To recover $\mathbf{x}$, we can use ℓ_1 minimization with relaxed constraint in the form of

$$\widehat{\mathbf{s}} = \arg \min_{\widetilde{\mathbf{s}}} \|\widetilde{\mathbf{s}}\|_{\ell_1} \qquad \text{subject to} \quad \|\mathbf{y} - \mathbf{\Phi}\mathbf{\Psi}\widetilde{\mathbf{s}}\|_{\ell_2} \leq \epsilon. \tag{5}$$

Specifically, if the isometry constant δ_{2k} associated with RIP satisfies $\delta_{2k} < \sqrt{2} - 1$ [3], then recovery by (5) guarantees

$$\|\widehat{\mathbf{s}} - \mathbf{s}\|_{\ell_2} \leq c_0 \|\mathbf{s} - \mathbf{s}_k\|_{\ell_1}/\sqrt{k} + c_1 \epsilon \tag{6}$$

where c_0 and c_1 are positive constants, and $\mathbf{s}_k$ is the k-term approximation of $\mathbf{s}$ by enforcing all but the largest k components of $\mathbf{s}$ to be zero.
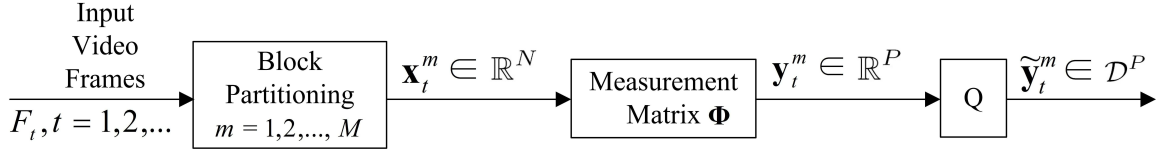
Figure 1. A simple compressed sensing (CS) video encoder system with quantization alphabet $\mathcal{D}$.

Equivalently, the optimization problem in (5) can be reformulated as the following unconstrained problem

$$\widehat{\mathbf{s}} = \arg\min_{\widetilde{\mathbf{s}}} \|\mathbf{y} - \boldsymbol{\Phi}\boldsymbol{\Psi}\widetilde{\mathbf{s}}\|_{\ell_2}^2/2 + \lambda\|\widetilde{\mathbf{s}}\|_{\ell_1}, \tag{7}$$

where λ is a regularization parameter that tunes the sparsity level. The problem in (7) is a convex quadratic minimization program that can be efficiently solved. Again, after we obtain $\widehat{\mathbf{s}}$, $\mathbf{x}$ can be reconstructed by (3). As for selecting a proper measurement matrix $\boldsymbol{\Phi}$, it is known [3] that with overwhelming probability probabilistic construction of $\boldsymbol{\Phi}$ with entries drawn from independent and identical distributed (i.i.d.) Gaussian random variables with mean 0 and variance $1/P$ obeys RIP provided that $P \geq c \cdot k \log(N/k)$. For deterministic measurement matrix constructions, the reader is referred to [4] and references therein.

3. PROPOSED CS VIDEO DECODING SYSTEM

The CS-based signal acquisition technique described in Section 2 can be applied to video acquisition on a frame-by-frame, block-by-block basis. In the simple compressive video encoding block diagram shown in Fig. 1, each frame F_t, $t = 1, 2, ...$, is virtually partitioned into M non-overlapping blocks of pixels with each block viewed as a vectorized column of length N, $\mathbf{x}_t^m \in \mathbb{R}^N$, $m = 1, ..., M$, $t = 1, 2, ...$. Compressive sampling of $\mathbf{x}_t^m$ is performed by random projection in the form of

$$\mathbf{y}_t^m = \boldsymbol{\Phi}\mathbf{x}_t^m \tag{8}$$

with a Gaussian generated measurement matrix $\boldsymbol{\Phi}_{P \times N}$. Then, the resulting measurement vector $\mathbf{y}_t^m \in \mathbb{R}^P$ is processed by a fixed-rate uniform scalar quantizer. The quantized indices $\widetilde{\mathbf{y}}_t^m$ are encoded and transmitted to the decoder.

In the CS video decoder of [10], each frame is individually decoded via sparse signal recovery algorithms with fixed bases such as block-based 2D-DCT (or frame-based 2D-DWT). With a received (dequantized) measurement vector $\widehat{\mathbf{y}}$ and a block-based 2D-DCT basis $\boldsymbol{\Psi}_{\mathrm{DCT}}$, video reconstruction becomes an optimization problem as in (7)

$$\widehat{\mathbf{s}} = \arg\min_{\widetilde{\mathbf{s}}} \|\widehat{\mathbf{y}} - \boldsymbol{\Phi}\boldsymbol{\Psi}_{\mathrm{DCT}}\widetilde{\mathbf{s}}\|_{\ell_2}^2/2 + \lambda\|\widetilde{\mathbf{s}}\|_{\ell_1} \tag{9}$$

where the original video block $\mathbf{x}$ is recovered as

$$\widehat{\mathbf{x}} = \boldsymbol{\Psi}_{\mathrm{DCT}}\widehat{\mathbf{s}}. \tag{10}$$

However, such intra-frame decoding using a fixed basis does not provide sufficient sparsity level for the video block signal. Consequently, higher number of measurements is needed to ensure a required level of reconstruction quality. To enhance sparsity, in [11] the correlation among successive frames was exploited by jointly recovering several frames with a 3D-DWT basis, assuming that the video signal is more sparsely represented in a 3D-DWT domain. In [12], a sparser representation is provided by exploiting small inter-frame differences within a spatial 2D-DWT basis. Nevertheless, in all cases, these decoders cannot pursue/capture local motion effects which can significantly increase sparseness and are well-known to be a critical attribute to the effectiveness of conventional video compression. Below, we propose and describe a new motion-capturing sparse decoding approach.

The founding concept of the proposed CS video decoder is shown in Fig. 2. The decoder consists of an initialization stage that decodes F_t, $t = 1, 2$, and a subsequent operational stage that decodes F_t, $t \geq 3$. At the initialization stage, F_1 is first reconstructed using the block-based fixed DCT basis exactly as described in (9)

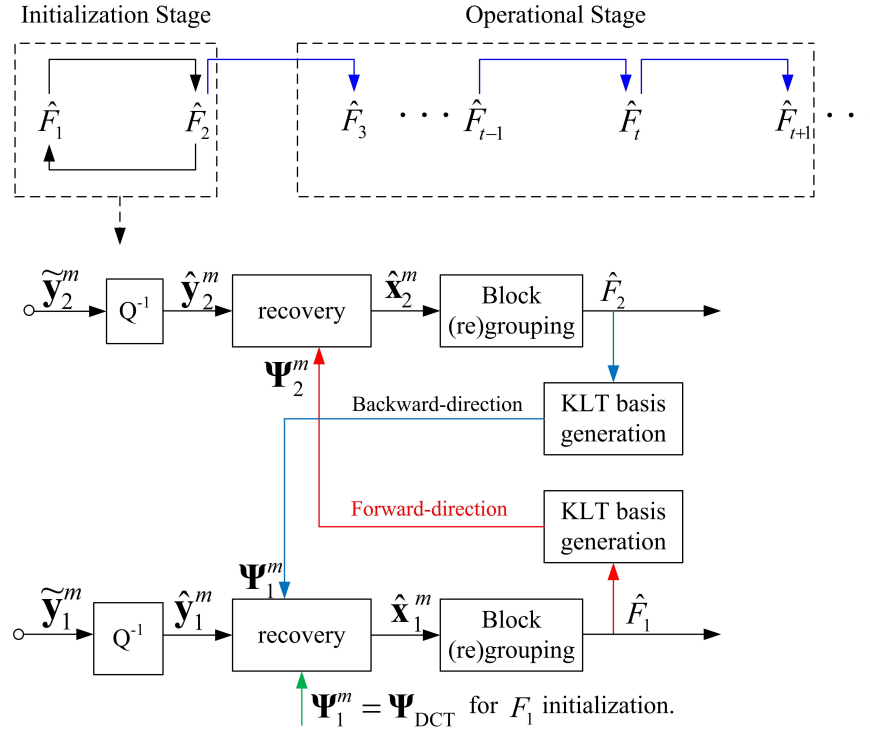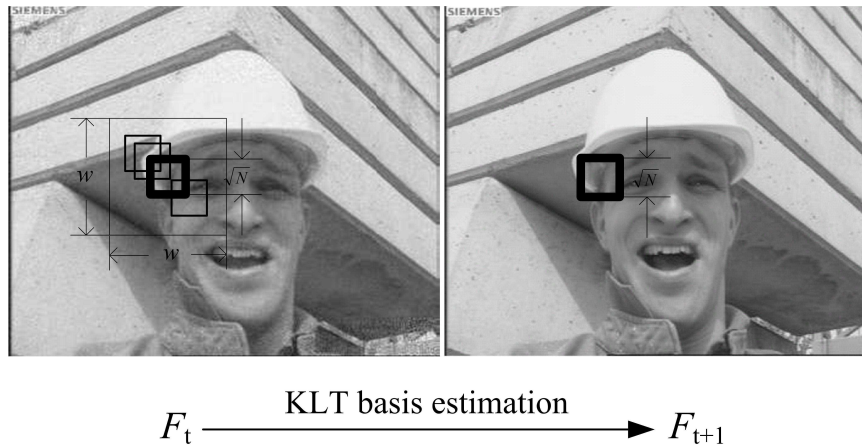Figure 2. Proposed CS decoder system (1st-order decoding algorithm).

Figure 3. KLT basis estimation illustration (1st-order decoding).

and (10). Then, we attempt to reconstruct each block of F_2 based on the reconstructed previous frame $\widehat{F}_1$. Our sparsity-aware ME decoding approach is based on the fact that the pixels of a block in a video frame may be satisfactorily predicted by using a linear combination of a small number of nearby blocks in adjacent (previous or next) frame(s). In particular, for our set-up the blocks in F_2 may be sparsely represented by a few neighboring blocks in $\widehat{F}_1$. We propose to use the KLT basis for this representation. For each block $\mathbf{x}_2^m$ in F_2, $m = 1, ...M$, a group of neighboring blocks that lie in a window of a square $w \times w$ region centered at $\mathbf{x}_2^m$ are extracted from $\widehat{F}_1$. Then, the KLT basis for $\mathbf{x}_2^m$, $\mathbf{\Psi}_{2,\mathrm{KLT}}^m$, is formed *by the eigenvectors of the correlation matrix of the extracted*
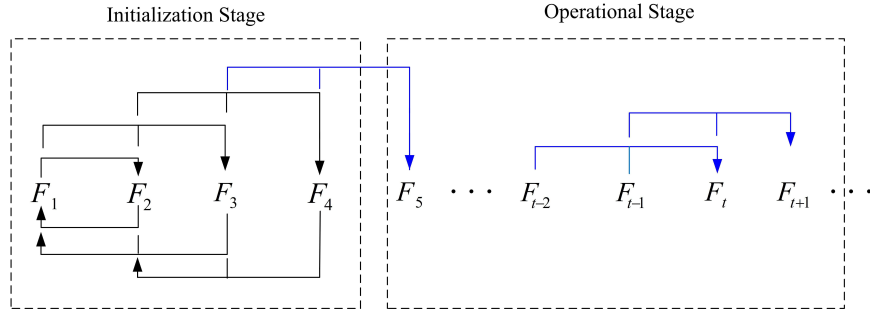
Figure 4. CS decoder of order 2.

blocks from $\widehat{F}_1$. Fig. 3 illustrates the block extraction procedure. Given a block $\mathbf{x}_2^m$ to estimate/reconstruct (block in bold of size $\sqrt{N} \times \sqrt{N}$ in F_2), one can find its co-located block $\widehat{\mathbf{x}}_1^m$ (block in bold of size $\sqrt{N} \times \sqrt{N}$ in $\widehat{F}_1$). Neighboring blocks (other overlapping blocks of size $\sqrt{N} \times \sqrt{N}$ in $\widehat{F}_1$) $\mathbf{d}_i$, $i = 1, ..., B$, can be extracted from a $w \times w$ area carrying out one-pixel shifts in all directions. When, say for example, w equals three times the block width $\sqrt{N}$ and block $\mathbf{x}_2^m$ is well in the interior of F_2, then the total number of available neighboring blocks is $B = (w - \sqrt{N})^2$; for blocks near the edge of F_2, B will be smaller accordingly.

Considering now all the extracted neighboring blocks as different realizations of an underlying vector stochastic process, the correlation matrix can be estimated by the sample average

$$\widehat{\mathbf{R}}_2^m = \frac{1}{B} \sum_{i=1}^{B} \mathbf{d}_i \mathbf{d}_i^T. \tag{11}$$

We form the KLT basis for Frame 2, Block m, $\boldsymbol{\Psi}_{2,\text{KLT}}^m$, by the eigenvectors of $\widehat{\mathbf{R}}_2^m = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^T$,

$$\boldsymbol{\Psi}_{2,\text{KLT}}^m = \mathbf{Q}, \tag{12}$$

where $\mathbf{Q}$ is the matrix with columns the eigenvectors of $\widehat{\mathbf{R}}_2^m$ and Λ is the diagonal matrix with the corresponding eigenvalues. Next, we recover the sparse coefficients $\mathbf{s}_2^m$ by solving

$$\widehat{\mathbf{s}}_2^m = \arg\min_{\widetilde{\mathbf{s}}} \|\widehat{\mathbf{y}}_2^m - \boldsymbol{\Phi}\boldsymbol{\Psi}_{2,\text{KLT}}^m \widetilde{\mathbf{s}}\|_{\ell_2}^2 / 2 + \lambda\|\widetilde{\mathbf{s}}\|_{\ell_1} \tag{13}$$

and we reconstruct the video block $\mathbf{x}_2^m$ by

$$\widehat{\mathbf{x}}_2^m = \boldsymbol{\Psi}_{2,\text{KLT}}^m \widehat{\mathbf{s}}_2^m. \tag{14}$$

After all M blocks are reconstructed, they are grouped again to form the complete decoded frame $\widehat{F}_2$.

So far, during the initialization stage, we have carried out forward only frame F_2 reconstruction accounting for motion from the DCT reconstructed frame F_1. For improved initialization, we may repeat the algorithm backward and reconstruct again F_1 using KLT bases generated from $\widehat{F}_2$. This forward-backward approach iterates for the initial two frames as shown in some detail in Fig. 2 until no significant further reconstruction quality improvement can be achieved.

At the normal operational stage that follows, the decoder recovers the blocks of F_t, $t \geq 3$, based on the KLT bases estimated from $\widehat{F}_{t-1}$. Since only one previous reconstructed frame is used as the reference frame in KLT bases estimation, we refer to this approach as 1st-order sparsity-aware ME decoding.

To exploit the correlation within multiple successive frames and achieve higher ME effectiveness in decoding, we may extend the 1st-order sparsity-aware ME decoding algorithm to an nth-order procedure where at the operational stage each frame is recovered from the past n reconstructed frames. For illustration purposes, Fig. 4 depicts the order $n = 2$ scheme. At the initialization stage, F_1 and F_2 are first reconstructed with forward-backward estimation as in 1st-order decoding. Then, F_3 is decoded with KLT bases estimated from both $\widehat{F}_1$
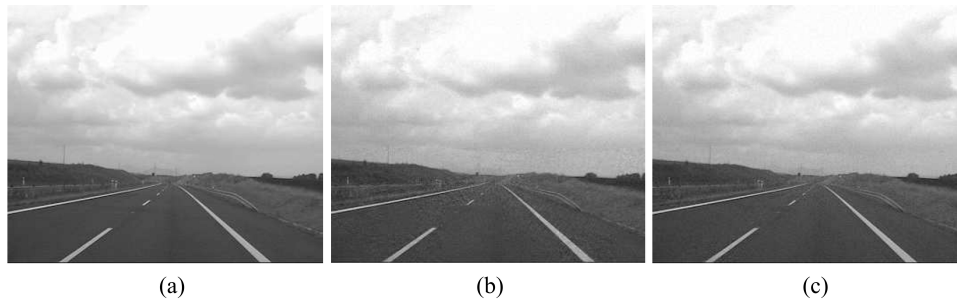
Figure 5. Different decodings of the 11*th* frame of Highway: (a) Original; (b) using the 2D-DCT basis intra-frame decoder ($P = 0.625N$); (c) using the order-5 sparsity-aware ME decoder ($P = 0.625N$).
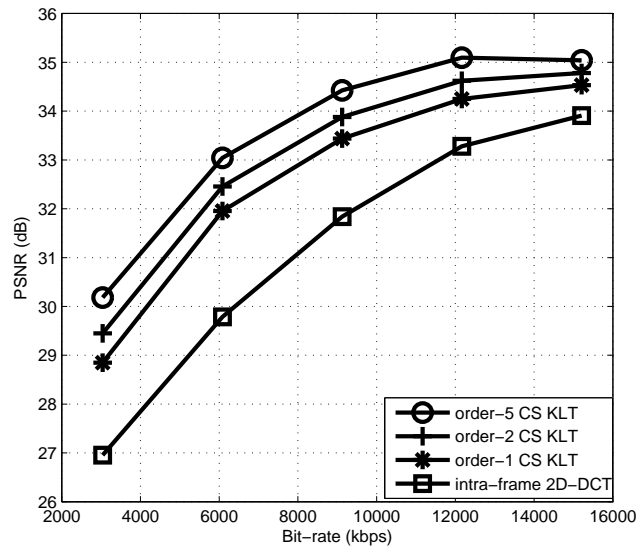

Figure 6. Rate-distortion studies on the Highway sequence.

and $\widehat{F}_2$. After $\widehat{F}_3$ is obtained, F_1 is decoded again in the backward direction with KLT bases estimated from both $\widehat{F}_2$ and $\widehat{F}_3$. The same 2nd-order decoding is performed in the forward direction for F_4 and in the backward direction for F_2, so that each of the initial frames F_t, $1 \le t \le 4$, has been reconstructed with implicit ME from two adjacent frames (Fig. 4). In the subsequent operational stage, each frame F_t ($t \ge 5$) is decoded by the two previous reconstructed frames $\widehat{F}_{t-1}$ and $\widehat{F}_{t-2}$. The concept is immediately generalizable to n^{th}-order decoding with $2n$ initial frames F_1, F_2, ..., F_{2n}.

A defining characteristic of the proposed CS video decoder in comparison with existing CS video literature [10]-[17] is that the order-n sliding-window decoding algorithm utilizes the spatial correlation within a video frame and the temporal correlation between successive video frames, which essentially results to implicit joint spatial-temporal motion-compensated video decoding. The adaptively generated block-based KLT basis provides a much sparser representation basis than fixed block-based basis approaches [10]-[12],[15] as demonstrated experimentally in the following section.

4. EXPERIMENTAL RESULTS

In this section, we study experimentally the performance of the proposed compressive sampling video decoders by evaluating the peak-signal-to-noise ratio (PSNR) (as well as the perceptual quality) of reconstructed video
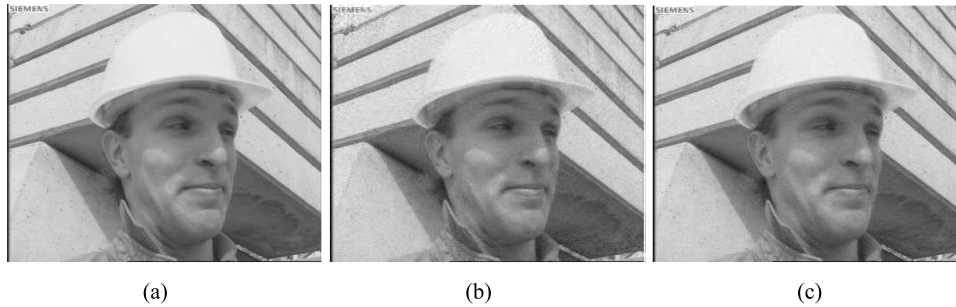
(a)   (b)   (c)

Figure 7. Different decodings of the $5th$ frame of Foreman: (a) Original; (b) using the 2D-DCT basis intra-frame decoder $(P = 0.625N)$; (c) using the order-5 sparsity-aware ME decoder $(P = 0.625N)$.
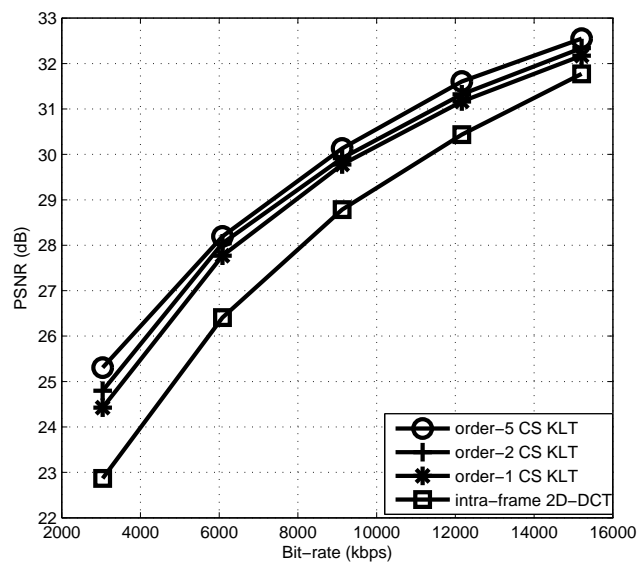

Figure 8. Rate-distortion studies on the Foreman sequence.

sequences. Two test sequences, Highway and Foreman, with CIF resolution 352×288 pixels and frame rate of 30 frames/second are used. Processing is carried out only on the luminance component.

At the encoder side, each frame is partitioned into non-overlapping blocks of 32×32 pixels. Each block is compressively sampled using a $P \times N$ measurement matrix with elements drawn from i.i.d. zero-mean, unit-variance Gaussian random variables. The captured measurements are quantized by an 8-bit uniform scalar quantizer and then sent to the decoder.

At the decoder side, we choose the Least Absolute Shrinkage and Selection Operator (LASSO) algorithm [6],[7] for sparse recovery motivated by its low-complexity and satisfactory recovery performance characteristics. In our experimental studies, four CS video decoders are examined: (i) fixed 2D-DCT basis intra-frame decoder used as a reference benchmark [10]; (ii) order-1; (iii) order-2; and (iv) order-5 sparsity-aware ME decoding.

Fig. 5 shows the decodings of the $11th$ frame of Highway produced by the 2D-DCT basis intra-frame decoder (Fig. 5(b)) and the order-5 CS decoder (Fig. 5(c)). It can be observed that the 2D-DCT basis intra-frame decoder suffers much noticeable performance loss over the whole image, while the proposed order-5 sparsity-aware ME

decoder demonstrates considerable reconstruction quality improvement[*].

Fig. 6 shows the rate-distortion characteristics of the four decoders (fixed 2D-DCT intra-frame, order-1, order-2, and order-5 CS decoding) for the Highway video sequence. The PSNR values (in dB) are averages over 100 frames. Evidently, the proposed order-1 sparsity-aware ME decoder outperforms significantly the fixed basis intra-frame decoder, especially at the low-to-medium bit rate ranges of interest with gains as much as 2dB. The 2nd-order and 5th-order decoders further improve performance by up to one additional dB.

The same rate-distortion performance study is repeated in Figs. 7 and 8 for the Foreman sequence. By Fig. 8, the proposed 1st-order sparsity-aware ME decoder again outperforms significantly the fixed basis intra-frame decoder, with gains approaching 1.5dB at the low-to-medium bit rate range of interest. The performance is enhanced by more than 0.5dB as the decoder order increases to five.

5. CONCLUSIONS

We proposed a sparsity-aware motion-accounting decoder for video streaming systems with plain compressive sampling encoding. The decoder performs sliding-window inter-frame decoding that adaptively generates KLT bases from adjacent previously reconstructed frames to enhance the sparse representation of each video frame block, such that the overall reconstruction quality is improved at any given fixed compressive sampling rate. Experimental results demonstrate that the proposed sparsity-aware decoders outperform significantly the conventional fixed basis intra-frame CS decoder. The performance is improved as the number of reference frames (what we call "decoder order") increases with order values in the range two to five appearing as a good compromise between computational complexity and reconstruction quality.

REFERENCES

[1] E. Candès and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inform. Theory*, vol. 52, pp. 5406-5425, Dec. 2006.

[2] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, pp. 1289-1306, Apr. 2006.

[3] E. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Proc. Magazine*, vol. 25, pp. 21-30, Mar. 2008.

[4] K. Gao, S. N. Batalama, D. A. Pados, and B. W. Suter, "Compressive sampling with generalized polygons," *IEEE Trans. Signal Proc.*, vol. 59, pp. 4759-4766, Oct. 2011.

[5] E. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure and Applied Math.*, vol. 59, pp. 1207-1223, Aug. 2006.

[6] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Stat. Soc. Ser. B*, vol. 58, pp. 267-288, 1996.

[7] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least angle regression," *Ann. Statist.*, vol. 32, pp. 407-451, Apr. 2004.

[8] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inform. Theory*, vol. 53, pp. 4655-4666, Dec. 2007.

[9] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Proc. Magazine*, vol. 25, pp. 83-91, Mar. 2008.

[10] V. Stankovic, L. Stankovic, and S. Cheng, "Compressive video sampling," in *Proc. European Signal Proc. Conf. (EUSIPCO)*, Lausanne, Switzerland, Aug. 2008.

[11] M. B. Wakin, J. N. Laska, M. F. Duarte, D. Baron, S. Sarvotham, D. Takhar, K. F. Kelly, and R. G. Baraniuk, "Compressive imaging for video representation and coding," in *Proc. Picture Coding Symposium (PCS)*, Beijing, China, Apr. 2006.

[12] R. F. Marcia and R. M. Willet, "Compressive coded aperture video reconstruction," in *Proc. European Signal Proc. Conf. (EUSIPCO)*, Lausanne, Switzerland, Aug. 2008.

---

[*]As usual, pdf formatting of the present article tends to dampen perceptual quality differences between Figs. 5 (a), (b), and (c) that are in fact pronounced in video playback. Fig. 6 is the usual attempt to capture average differences quantitatively.

[13] H. W. Chen, L. W. Kang, and C. S. Lu, "Dynamic measurement rate allocation for distributed compressive video sensing," in *Proc. Visual Comm. and Image Proc. (VCIP)*, Huang Shan, China, July 2010.

[14] J. Y. Park and M. B. Wakin, "A multiscale framework for compressive sensing of video," in *Proc. Picture Coding Symposium (PCS)*, Chicago, IL, May 2009.

[15] L. W. Kang and C. S. Lu, "Distributed compressive video sensing," in *Proc. IEEE Intern. Conf. on Acoustics, Speech, and Signal Proc. (ICASSP)*, Taipei, Taiwan, Apr. 2009, pp. 1393-1396.

[16] J. Prades-Nebot, Y. Ma, and T. Huang, "Distributed video coding using compressive sampling," in *Proc. Picture Coding Symposium (PCS)*, Chicago, IL, May 2009.

[17] T. T. Do, Y. Chen, D. T. Nguyen, N. Nguyen, L. Gan, and T. D. Tran, "Distributed compressed video sensing," in *Proc. IEEE Intern. Conf. on Image Proc. (ICIP)*, Cairo, Egypt, Nov. 2009, pp. 1169-1172.