LEARNING SCALABLE DICTIONARIES WITH APPLICATION TO SCALABLE COMPRESSIVE SENSING

Bojana Begovic, Lina Stankovic and Vladimir Stankovic

Department of Electrical and Electronic Engineering, University of Strathclyde, Glasgow/UK, G1 1XW

ABSTRACT

Learned sparse signals representations led to state-of-the-art image restoration results for several problems in the field of image processing. In this paper we show that these are achievable for a compressive sensing (CS) scenario based on the sparsity pattern provided by learned dictionary specially designed for the scalable data representation. Experimental results demonstrate and validate the practicality of the proposed scheme making it a promising candidate for many practical applications involving both time scalable image/video display and scalable frame compressive sensing. Provided simulations involve CS scalable sparse recovery of dynamic data changing over time e.g., video. These are important for situations where video streams, tailored to the needs of a diverse user pool operating heterogeneous display equipment, are required. For the aforementioned purpose the proposed method outperforms the conventional K-SVD algorithm.

***Index Terms*—** Sparse encoding, scalable video representation, regularisation, compressive sensing

1. INTRODUCTION

The sparse coding paradigm is a critical factor for recent breakthrough results in the field of image processing. Its main foundation is the assumption that signals (e.g., natural images) are subject to a linear sparse decomposition over a learned dictionary. This so-called sparseland model [1] has led to new generation of state-of-the-art algorithms for several image processing problems [1] stating that the dictionary $\mathbf{D} \in R^{n \times K}$ can be trained for any image signal class $\mathbf{Y}$. Typically the dictionary $\mathbf{D}$ is redundant e.g., the number of its basis vectors (atoms) is greater than the original signal dimension ($K > n$). Given one of the pursuit algorithms e.g., [2][3] and a dictionary $\mathbf{D}$, one can retrieve matrix $\mathbf{X}$ containing sparse approximations $\{\mathbf{x}_i\}_{i=1}^{N} \in R^K$ of each signal $\mathbf{y}_i \in R^n$ i.e. extracted image patch from $\mathbf{Y}$. Thus, we can approximate the image as $\mathbf{Y} \approx \mathbf{DX}$, a collection of weighted linear combinations of a few atoms in $\mathbf{D}$ given each patch $\mathbf{y}_i$. For example, K-SVD algorithm [1][4] is one of the current state-of-the-art algorithms for unsupervised dictionary learning procedure. It efficiently tackles and solves numerous image processing tasks such as denoising, inpainting and demosaicing, both for grayscale [1] and colour [5] images.

Sparse signal decomposition also plays an important role in the reconstruction performance of the compressive sampling (CS) [8][2], commonly in the form of some predefined transform basis e.g., Discrete Cosine Transform (DCT) [6], wavelets [7] etc. This emerging framework has gained an increased interest over the past few years by introducing innovative and revolutionary signal processing mathematics. Specifically, CS carries out joint signal compression and sampling, enabled with specially designed measurement matrix $\mathbf{\Phi} \in R^{S \times n}$ that takes random undercomplete set of samples $\mathbf{y}_{CS} \in R^S$ ($S << n$) as $\mathbf{y}_{CS} = \mathbf{\Phi y} = \mathbf{\Phi D x}$. This simple linear random projections of the source signal $\mathbf{y}$ are aiming to replace the conventional acquisition process i.e. the Nyquist sampling paradigm assuming that treated signal is sparse rather than band-limited. The second important concept upon which CS relies (in order to obtain robust signal reconstruction) is incoherent sampling. That is, vector elements of both representational and sensing basis ($\mathbf{D}$ and $\mathbf{\Phi}$) should exhibit low coherence satisfying the so-called restricted isometry property (RIP) [2]. With one of the non-linear pursuit algorithms e.g., [3], the sparsest representation of $\mathbf{y}$ is commonly restored from sampled measurements $\mathbf{y}_{CS}$. Furthermore, a small number of recent publications challenge the typical CS operational setting by taking into consideration learned dictionaries [9][10] rather than commonly used off-the-shelf sparsifying representational basis e.g., [6][7]. This research direction is backed up by results in [10] where authors put forward a scheme that jointly trains and optimises overcomplete non-parametric dictionary and the CS sensing matrix $\mathbf{\Phi}$.

Additional employment of the sparse signal decomposition is demonstrated in [6] for progressive signal reconstruction in time. By sampling more and more data in incremental steps, one is able to provide a scalable image/video display. In particular, this aims to support diverse applications in terms of user equipment heterogeneity, communication channels and QoS demands. In [6] a scalable signal representation is achieved with the conventional DCT dictionary rather than learning the best dictionary tailored to this task. Hence, a procedure for learning a dictionary capable of adapting to a specific dataset and providing its effective scalable reconstruction is still missing.

This paper addresses for the first time an integration of a trained adaptive dictionary [11] specialised for *scalable* high-motion video sequences reconstruction by extending the K-SVD algorithm and CS *scalable* framework (inspired by [6]). This aims to achieve the frame by frame joint CS *scalable* sampling/compression recovery where progressive restoration in time is achieved by attaining more and more sparse coefficient's entries per patch of each frame. The main contribution of the proposed work is that it provides an unification of the adaptive *scalable* sparse image representation with *scalable* CS sensing. The practicality of the proposed method is illustrated for the general *scalable* recovery and CS *scalable* recovery of high-motion video sequences. All frames are reconstructed via a single trained scalable dictionary. Experimental results confirm that the proposed scalable scheme outperforms conventional K-SVD.

In summary, we emphasize that the simple yet powerful remodelling of the traditional K-SVD algorithm by regularising the dictionary learning process facilitates a sparse signal representation in a *scalable* i.e. progressive manner. This broadens its practical potential especially by integrating it into the traditional CS method aiming to achieve *scalable* CS recovery. Due to space limitations the reader is referred to [1][4] for more details on the the state-of-the-art algorithms for sparse coding K-SVD and to [2][8][10][12] for background information on CS paradigm.

2. SCALABLE DICTIONARY MODEL DESCRIPTION

Adhering closely to the notation used in [4], this section covers the description of the method proposed in [11]. As a basis for our technique we take the regular K-SVD algorithm [4] and build upon it by alternating one of its two main iterative stages i.e. dictionary update. In general, we are given a set of N signals i.e. overlapping image patches size $\sqrt{n} \times \sqrt{n}$ vectorised as $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_N] \in R^{n \times N}$. The classical configuration of the K-SVD algorithm aims to represent these signals as linear combinations of a few dictionary elements: the columns of matrix $\mathbf{D} = [\mathbf{d}_1, \ldots, \mathbf{d}_K] \in R^{n \times K}$. Training of $\mathbf{D}$ is performed simultaneously with an estimation of the matrix containing sparse coefficients $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_N] \in R^{K \times N}$. Once iterative learning is completed each signal in $\mathbf{Y}$ is approximated by $\mathbf{y}_i \approx \mathbf{D}\mathbf{x}_i$.

Note however that this conventional approach is not able to deliver *scalable* and *adaptive* image representation that would be based on the progressive recovery of each image patch $\mathbf{y}_i$. For example, one can form recovery layers denoted as L_{s_a} with $\{s_a \,|1 \le s_a \le \lfloor K/m \rfloor, m < K\}$ over each patch i.e. image. At the beginning of the scalable reconstruction process, the image base layer L_1 consists of the first m sparse coefficients entries per patch. Each subsequent layer L_{s_a} ($s_a > 1$) contains an additional m coefficient values. This problem is addressed by introducing a regularisation scheme over the second K-SVD iterative stage [11] i.e., dictionary atom update step. This provides an effective *scalable* and *adaptive* image reconstruction model for the image data that exhibits sparse representation.

2.1. Sparse coding stage

The first of the two iterative dictionary learning stages (sparse coding) is addressed as a constraint optimisation problem similar to [4], where the optimisation objective:

$$\min_{\mathbf{X}} \left\{ \|\mathbf{Y} - \mathbf{D}_{sc}\mathbf{X}\|_F^2 \right\} \;\; s.t. \; \forall i \;\; \|\mathbf{x}_i\|_0 \le T_0 \qquad (1)$$

is met by OMP [3]. K denotes the number of atoms in $\mathbf{D}_{sc}$ that is the complete scalable dictionary kept fixed during this training phase. Signal $\mathbf{y}_i$ (i = *1, ...,N*), extracted from the original image $\mathbf{Y}$, is mapped into its sparse representation $\mathbf{x}_i$ via [3]. Each of the K entries $\mathbf{x}_i[l]$ corresponds to one of the atoms $\mathbf{d}_l \in \mathbf{D}_{sc}$ (*l=1,...,K*) where $\mathbf{x}_i[l] = 0$ means that particular atom $\mathbf{d}_l$ does not participate in the sparse representation of the signal $\mathbf{y}_i$. The pseudo norm $\|\cdot\|_0$ accounts for the number of non-zero elements in $\mathbf{x}_i$, bounded with an upper constraint T_0. The method in [11] relaxes sparsity constraint T_0 by allowing it to take a greater value than one defined by [4] in order to establish the scalable signal recovery. This change is introduced on an empirical basis, still maintaining sparsity notion of the signal representation.

2.2. Scalable dictionary update stage

The initialisation of the second stage is performed as in [4] where an atom subject to update $\mathbf{d}_j$ is changed while the rest of the $\mathbf{D}_{sc}$ members are kept fixed. Prior to each update, $\mathbf{d}_j$ is set to zero. The integration of the proposed regularisation scheme is executed during the construction of the current representational error matrix $\mathbf{E}_j$ associated with the atom $\mathbf{d}_j$. In [4] $\mathbf{E}_j$ is defined as $\mathbf{E}_j = |\mathbf{Y}_j - \mathbf{D}\mathbf{X}_j|_2^2$. $\mathbf{Y}_j$ is a subset of image patches from $\mathbf{Y}$ which current sparse approximation $\mathbf{X}_j$ includes atom $\mathbf{d}_j$ ($\mathbf{x}_i\,[j] \neq 0 \in X_j$) while $\mathbf{D}$ is a non-scalable dictionary. Unlike [4], we separate the sparse representation $\mathbf{X}_j$ into two components:

- Low frequency (smooth): $\mathbf{Y}_j^{\textbf{low}} = \mathbf{D}_{sc}\mathbf{X}_j^{\textbf{low}}$;
- High frequency (texture): $\mathbf{Y}_j^{\textbf{high}} = \mathbf{D}_{sc}\mathbf{X}_j^{\textbf{high}}$.

These two structural frequencies are denoted with superscripts **low** and **high** respectively. $\mathbf{D}_{sc}$ denotes the dictionary for *scalable* presentation. Subsequently the residual matrix is formed as:

$$\mathbf{E}_j = \left| \mathbf{Y}_j - v_0 \mathbf{D}_{sc}\mathbf{X}_j^{\textbf{low}} - v_1 \mathbf{D}_{sc}\mathbf{X}_j^{\textbf{high}} \right|_2^2. \qquad (2)$$

Weight pair (v_0, v_1) regularises the contribution of the texture and smooth image component to $\mathbf{E}_j$. This provides a mean for controlling the type of the information used for the atom's update. The idea for introduced regularisation scheme is driven by the well known perception characteristic of the Human Visual System (HVS). Specifically, human eyes tend to pay more attention to the edges of an object, thus primarily identifying objects by their shape. This property suggests that HVS is most sensitive to high frequency image content (edges). Motivated by this, we assume the following. In order to achieve effective scalable recovery, the main object shapes appearing within the image need to be identified from the very start of the progressive image restoration process. Hence, higher frequencies are more relevant to scalable dictionary learning and need to be appropriately favored during its training. This reasoning is confirmed by the series of experiments where various weight pairs $[v_0, v_1]$ (with the constraint of $v_0 + v_1 = 1$) show that the introduced regularisation of the texture and smooth image component errors yields the appropriate dictionary for *scalable* data presentation. The overall algorithm is shown as follows:

1. STEP 1 - As in [4] initialise dictionary $\mathbf{D}_{sc}$ with K random extracted overlapping patches;
2. STEP 2 - Unlike [4] split each current sparse approximation in $\mathbf{X}_j$ into:
   - $\mathbf{x}_i^{\textbf{low}} = \mathbf{x}_i \mathbf{T}^{\textbf{low}}$ and $\mathbf{x}_i^{\textbf{high}} = \mathbf{x}_i \mathbf{T}^{\textbf{high}}$;

   where $\mathbf{T}^{\textbf{low}}, \mathbf{T}^{\textbf{high}} \in R^K$ are binary vectors that set to zero any $\mathbf{x}_i\,[l]$ element for $l > \frac{K}{2} + c$ and $l < \frac{K}{2} - c$, respectively. By this the $\mathbf{X}_j$ is decomposed into smooth and texture patch information represented as matrices $\mathbf{X}_j^{\textbf{low}}$ and $\mathbf{X}_j^{\textbf{high}}$ respectively. c is a constant integer term set heuristically.
3. STEP 3 - Unlike [4] after decomposing the sparse representation of $\mathbf{Y}_j$ into $\mathbf{D}_{sc} X_j^{\textbf{low}}$ and $\mathbf{D}_{sc} X_j^{\textbf{high}}$ what accordingly forms their representational residual term as: $\mathbf{E}_j = \left| \mathbf{Y}_j - v_0 \mathbf{D}_{sc} \mathbf{X}_j^{\textbf{low}} - v_1 \mathbf{D}_{sc} \mathbf{X}_j^{\textbf{high}} \right|_2^2$;
4. STEP 4 - As in [4] perform rank-one approximation of $\mathbf{E}_j$ i.e., the Singular Value Decomposition (SVD): for new $\mathbf{d}_j$ atom set the eigenvector corresponding to the largest eigenvalue of the E_j SVD decomposition;
5. STEP 5 - Unlike [4] keep mutually coherent atoms;

By enforcing this type of the frequency content regularisation the learning process generates scalable dictionary $\mathbf{D}_{sc}$ tailored to the characteristics of HVS.

3. APPLICATIONS TO IMAGE PROCESSING

In order to demonstrate the efficacy of the proposed method, we report simulation results both for *scalable* video encoding and *scalable CS*. The performance is evaluated using video test sequence "Stephan" and "Tempete" at CIF resolution 352×288 and a frame rate of 30Hz. Each frame is broken down into total of $N = 96945$ overlapping patches of 8×8 pixels. Hence, the vector signal size $\mathbf{y}_i$ used for the *scalable* dictionary learning algorithm and CS sampling is $n = 64$ pixels. Both dictionaries, $\mathbf{D}_{sc}$ and $\mathbf{D}$, contain $K = n$ entries resulting with redundancy factor $r = 1$. The reason behind employing complete ($K = n$) rather than commonly trained overcomplete dictionary ($K >> n$) stems from conclusions we drawn after performing numerous simulations. They have shown that there is no significant difference between scalable sparse restoration obtained via a complete or overcomplete learnt basis. Secondly, a complete, orthonormal basis is commonly used for the implementation of the CS acquisition procedure. Lastly, the effect of setting the redundancy to $r = 1$ relaxes slightly the algorithm complexity reducing the total training time since less atoms are trained.

The weight pair $[v_0, v_1]$ is set to values 1 and 0, respectively. Detailed simulations shown that this parametrisation yields the best trained scalable dictionary. Furthermore, T_0 is set to 10 (as pointed out in Sec. 2.1) relaxing the upper constraint on the sparsity level allowing for more than four nonzero entries [4] per $\mathbf{x}_i$. The *scalable* training scheme is carried out only once for the first incoming frame in the "Stephan" or "Tempete" sequence i.e., the *training frame* generating the $\mathbf{D}_{sc}$ dictionary. Each subsequent frame for either of the illustrated applications is reconstructed via a single $\mathbf{D}_{sc}$ given *scalable* sparse encoding information or *training frame* CS measurements. Hence, in the reported experimental results only a single dictionary is maintained throughout the entire sequence. Shown approach highly simplifies the computational complexity since training is done only once instead of 300 times. This is significant for the display of the real-time scalable complex video streams. However, when the video scene undergoes significant changes with respect to the *training frame*, a new training frame should be inserted.

3.1. Results for scalable video encoding

For every encoded patch and image, a total of $\lfloor K/m \rfloor = 16$ recovery layers are defined as L_{s_a} (s_a=1,...,16) with scalability step $m = 4$. At the user side, the base layer L_1 is recovered given the trained dictionary $\mathbf{D}_{sc}$ and the first

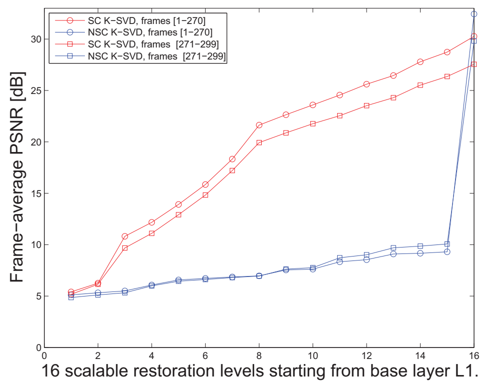
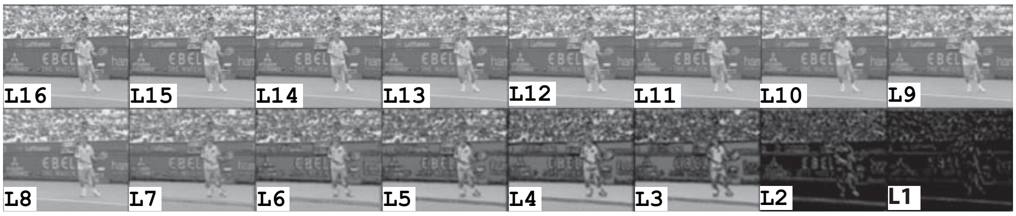
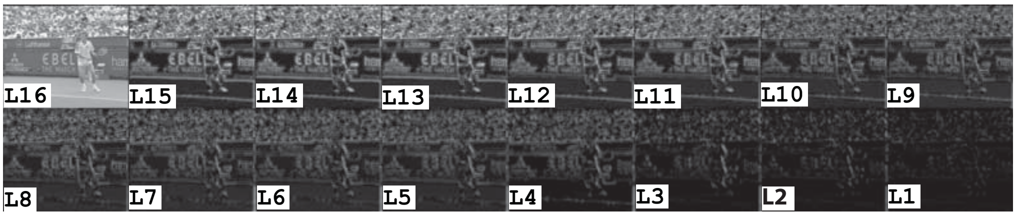


Fig. 1: Frame-average PSNR of the scalable reconstructed video test sequence "Stephan" per each recovery layer L_{s_a} using the scalable and non-scalable K-SVD algorithm



(a) SC K-SVD training frame, “Stephan”



(b) NSC K-SVD training frame, “Stephan”

Fig. 2: Visual assessment of the scalable reconstruction using the scalable/non-scalable K-SVD at every recovery level L_{s_a}.

four incoming coefficient values denoted as $X_1 \in R^{(4)\times N}$ while the rest of unknown entries are treated as zero. The rest of the recovery levels are progressively enhanced by adding four additional entries in each representational vector e.g., $X_2 \in R^{(8)\times N}, \ldots, X_{s_a} \in R^{(s_a * m)\times N}, \ldots, X_{16} \in R^{(64)\times N}$ (per each L_{s_a}, total of $s_a * m$ sparse coefficient values is available). This is done until the final L_{16} restoration level is attained. Constant c (Sec. 2) is heuristically set to value 2.

The objective quality assessment is provided in Fig. 1 in a form of a PSNR only for video sequence “Stephan” due to the space limitation. Given the scalable (“SC”) and conventional (“NSC”) K-SVD algorithm we provide the comparison for the the restoration quality averaged over two groups of frames from the “Stephan” sequence: the first 270 ([1, 270]) and the last 30 ([271, 300]) . By working with these two separate segments, we show to what extent the scalable restoration quality will be affected once the new object is introduced in the upcoming frame (group [271, 300]) while the initial $\mathbf{D}_{sc}$ is maintained for the recovery at the user side. The results clearly demonstrate that the proposed regularised scheme outperforms the standard [4] over all recovery levels L_{s_a} on average by 10.8dB (first 270 frames). This trend is observable as well for the second group of frames that contains a new object i.e. a tennis net with a slight drop in the restoration quality of 1.54dB. Only in the case when all of the information on the sparse coefficients is available ($X_{16} \in R^{(64)\times N}$), does the regular K-SVD algorithm have a slight advantage over the proposed scheme.

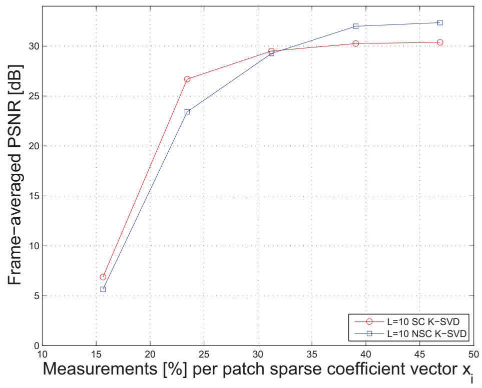
An additional visual assessment of the scalable recovery method is provided in Fig. 2 for reconstruction outcomes at every recovery level L_{s_a}, given the *training frame* from the “Stephan“ sequence. Through carefully observation one can confirm that the proposed scalable scheme is able to recover the frame at a recovery level L_3 ($X_3 \in R^{(12)\times N}$) whereas [4] fails to show any scalable characteristics at any level up to L_{15} ($X_{15} \in R^{(60)\times N}$). Similar visual results were obtained for the second video test sequence “Tempete” which is omitted for brevity.

3.2. Scalable compressive sensing

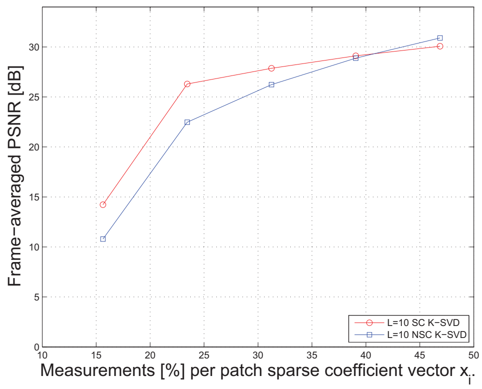
Following closely the experimental layout suggested in [6], here we assess the performance of the proposed *scalable* CS video acquisition scheme unified with the trained *scalable* dictionary D_{sc}. In particular, the proposed framework aims for frame-by-frame progressive CS recovery while analysing the implications of the sub-Nyquist CS paradigm in the *scalable* and *adaptive* representational domain. Unlike in [6], the image is processed block by block, likewise in the previous experimental section. In this way we investigate the effectiveness of the *scalable* CS paradigm when combined with a learned basis rather than a predefined one. Mainly, we take into consideration two cases of CS *scalable* recovery: (*i*) with the *scalable* K-SVD dictionary tailored to this task; (*ii*) with the conventional non-scalable K-SVD dictionary. Sensing is performed sequentially in the sparse representation domain $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_N] \in R^{K\times N}$ for each frame $\mathbf{X}$ in incremental steps. Given the sufficient number of measurements for each image patch represented as a collection of $s_1, s_1+s_2, \ldots, s_1+s_2+\ldots+s_L$ ($\sum_{i=1}^{L} s_i < K$) we are able to recover the frame gradually by collecting entries of sparse coefficients in $\mathbf{X}_i$. Furthermore, each value s_i satisfies the fundamental result of the CS theory [2] that imposes the limit on the necessary number of measurements for exact signal reconstruction. Unlike in the conventional CS setup defined as $\mathbf{y}_{CS} = \mathbf{\Phi}\mathbf{y} = \mathbf{\Phi}\mathbf{D}_{cs}\mathbf{x}$, here we specially structure (size-wise) the sampling matrix $\mathbf{\Phi}$ in order to achieve efficient *scalable* sampling of each image layer. This is carried out using the systematic non-adaptive approach [6] where for each recovery step we have $\mathbf{\Phi}_i \in R^{s_i\times K}$ resulting in $\mathbf{y}^i_{CS}$ samples.

Starting from a base level for $i = 1$ and $s_1 = 10$ with

$\mathbf{y}^1_{CS} = \mathbf{\Phi}_1\mathbf{y} = \mathbf{\Phi}_1\mathbf{D}_{cs}\mathbf{x}$ (around 15 % for each sparse coeff-cient denoted as $\mathbf{x_i}$) we advance through enhancement layers with additional five samples (e.g., $s_1 + s_2 = 15$) in each step until the total number of $S = 30 < K$ samples is reached. Hence, total of five sampled layers are recovered. Fig. 3 shows reconstruction via the proposed *adaptive scalable* CS averaged over frames, given the single trained dictionary $\mathbf{D_{sc}}$ for two video test sequences. The measurement number of interest over each frame is less than 50 %. Reconstruction is perfromed using OMP [3]. The gap between the performance of the two methods is evident for the layers sampled at low subrates i.e. one and two, at around 2.25dB in case of "Stephan" sequence and, in additional third layer for the "Tempete" frames at 2.96dB respectively. We can see that proposed design is successful for the low subsampling factors (e.g., 15%, 23 % and 31 % of sampling information) whereas the conventional K-SVD overtakes the lead as more measurements are added. However, in the context of the CS pardigm, this is not relevant since it imposes a low compression rate during sampling.



(a) Scalable CS reconstruction "Stephan"



(b) Scalable CS reconstruction "Tempete"

Fig. 3: Frame-average PSNR of the scalable CS reconstructed (a) "Stefan" and (b) "Tempete"test sequences as a function of the measurement procent using the scalable (SC) and non-scalable (NSC) K-SVD algorithm

4. CONCLUSION

This work introduces an integrated combination of a learning dictionary framework for scalable image recovery and adaptive scalable CS. Mainly, this is achieved by regularising the K-SVD dictionary update stage together with CS structured sampling. To the best of our knowledge this problem has not been addressed before in the context of learned sparse representations. Overall the experimental results are promising and demonstrate that the proposed method significantly outperforms the classical K-SVD setting and that it can be successfully used for CS *scalable* display of the complex video streams. Future work will analyse achieved coherence levels and the extent to which the RIP term is satisfied, given the *adaptive scalable* learned basis and random sampling matrix.

5. REFERENCES

[1] M. Elad,"Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing," *SPRINGER,* 2010.

[2] D. L. Donoho, "Compressed sensing ," *IEEE Trans. on Information Theory*, vol. 52, pp. 12891306, Apr. 2006.

[3] G. Davis, S. Mallat, and Z. Zhang, "Adaptive time-frequency decompositions," *Opt. Eng*, vol. 33, pp. 2183-2191, Jan. 1994.

[4] M. Aharon, M. Elad and A. Bruckstein, "K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation," *IEEE Trans. Im. Proc.*, vol. 54, pp. 4311-4322, Dec. 2006.

[5] J. Mairal, M. Elad, and G. Sapiro, "Sparse Representation for Color Image Restoration," *IEEE Trans. Im. Proc.*, vol. 17, no.1, pp. 53-69, Jan. 2008.

[6] V. Stankovic, L. Stankovic, and S. Cheng, "Scalable compressive video," *ICIP-2011 IEEE Int. Conf, on Im. Proc.*, Brussels, Belgium , Sep. 2011.

[7] M. Wakin, J. Laska, M. Duarte, D. Baron, S. Sarvotham, D. Takhar, K. Kelly, and R.G. Baraniuk, "Compressive imaging for video representation and coding," *PCS-2006 Picture Coding Symposium*, Beijing, China , Apr. 2006.

[8] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. on Inf. Theory*, vol. 52, Feb. 2006.

[9] M. Elad, "Optimized Projections for Compressed Sensing," *IEEE Tran. on Sig. Proc.*, vol. 55, pp. 5695-5702, Dec. 2007.

[10] J. M. Duarte-Carvajalino and G. Sapiro, "Learning to sense sparse signals: Simultaneous sensing matrix and sparsifying dictionary optimization, " *IEEE Trans. Image Proc.*, vol. 18, no. 7, pp. 1395-1408, July 2009.

[11] B. Begovic, V. Stankovic, L. Stankovic and S. Cheng, "Scalable Sparse Image Representation With Learned Dictionaries," submitted to *ASILOMAR-2012 IEEE Conf.*

[12] E. J. Candès and M. B. Wakin, "An Introduction To Compressive Sampling," *IEEE Tran. on Sig. Proc.*, pp. 2130, March 2008.