

SPARSE REPRESENTATION OF IMAGES USING OVERLAPPING FRAMES

Karl Skretting, Kjersti Engan, John Håkon Husøy, and Sven Ole Aase

Stavanger University College, Department of Electrical and Computer Engineering
P. O. Box 2557 Ullandhaug, N-4091 Stavanger, Norway
Phone: +47 51 83 20 16, Fax: +47 51 83 17 50, E-mail: karl.skretting@tn.his.no

ABSTRACT

The use of *frames* and *matching pursuits* for signal representation are receiving increased attention due to their potential in various signal processing applications. Good design algorithms for *block oriented* frames have recently been published. Viewing these block oriented frames as generalizations of block oriented transforms, it was natural to seek corresponding generalizations of critically sampled filter banks leading to *overlapping frames*. Here we adapt these design methods to be used on two dimensional signals, that is images, and develop a design scheme for block oriented and overlapping frames suitable for images. Experiments show that the designed frames achieve better results than standard orthogonal decomposition methods when the goal is sparse representation of images.

1. INTRODUCTION

Sparse image representations are useful in different applications. In commonly used block oriented transforms, e.g. the Discrete Cosine Transform (DCT), and the more recent wavelet based decomposition methods, the sparseness is introduced through thresholding of the transform or wavelet coefficients. Thus only a limited number of non-zero coefficients represent the image. This introduces errors in the reconstructed image. The goal is to minimize these errors while fulfilling a sparsity constraint making the number of non-zero coefficients small compared to the number of pixels in the original image. The non-zero coefficients as well as their position information constitute a sparse representation of the image, useful in many applications like compression, feature extraction, modelling, and classification.

This has an analogy in the human visual system, where an image is interpreted as many objects placed relative to each other. This is a sparse representation of the image. We do not perceive the millions of pixels that are received in the retina, but rather a sparse image of some few constituent parts (objects) with different shapes and textures.

Block oriented orthogonal transforms may be used to get a sparse representation of a signal vector, \mathbf{x}_l . The forward transform, which we for notational convenience denote by $\mathbf{T}^{-1} = \mathbf{T}^T$, is used to compute the transform coefficient vector, denoted \mathbf{y}_l , for each block by what is commonly called the analysis equation, $\mathbf{y}_l = \mathbf{T}^{-1}\mathbf{x}_l$. The reconstructed signal vector is then given by the corresponding synthesis equation, where tilde indicates possibly approximated quantities:

$$\tilde{\mathbf{x}}_l = \mathbf{T} \tilde{\mathbf{y}}_l = \sum_{n=1}^N \tilde{y}_l(n) \mathbf{t}_n. \quad (1)$$

The *synthesis vectors*, denoted $\{\mathbf{t}_n\}_{n=1}^N$, are the columns of the matrix \mathbf{T} . The approximated signal vector is built up as a linear

combination of these synthesis vectors. The $\{\tilde{y}_l(n)\}_n$ are thresholded versions of the coefficients, giving a sparse representation.

An overcomplete set of N -dimensional vectors, spanning the space, $\{\mathbf{f}_k\}_{k=1}^K$ where $K \geq N$, is a *frame*. Interpreting these vectors as columns of an $N \times K$ matrix \mathbf{F} we have a more general situation than that of Equation 1. If we use a frame instead of an orthogonal transform, the coefficients can not be computed using the traditional analysis equation. Practical solutions employ *vector selection* algorithms such as Matching Pursuit (MP) [1], Orthogonal Matching Pursuit (OMP) [2], and Fast Orthogonal Matching Pursuit (FOMP) [3].

While the use of frames on images has been reported by several authors, the development of procedures for their design is still in its infancy. In fact most authors use ad-hoc frame designs, [4, 5, 6, 7, 8]. Some attempts at the design of optimal block oriented frames, given a sparseness constraint, are given in [9].

The first attempt at designing frames where the frame vectors are allowed to overlap, referred to as *overlapping frames*, is reported in [10]. While producing good results, this method for designing overlapping frames is both computationally and conceptually more demanding than block oriented frame design. In [11] a method for designing overlapping frames that is as simple as the design procedure for block oriented frames was proposed. Note that these methods for design of overlapping frames have previously only been used on one dimensional signals. In this paper we adapt the method of [11] to be used also on images. We describe how the one dimensional frame design methods may be extended to the two dimensional case, both in the block oriented and the overlapping case. Subsequently, we describe some practical improvements or adjustments that we suggest for the two dimensional case. This is followed by some examples, a comparison to orthogonal decomposition methods, and a brief discussion.

2. FRAME DESIGN

The task is to design a frame, \mathbf{F} , of size $N \times K$, $K \geq N$, adapted to a class of signals, represented by a large set of training vectors, $\{\mathbf{x}_l\}_{l=1}^L$. The expansion corresponding to Equation 1 is

$$\tilde{\mathbf{x}}_l = \mathbf{F} \mathbf{w}_l = \sum_{k=1}^K w_l(k) \mathbf{f}_k, \quad (2)$$

where we have replaced the transform coefficients by weights, denoted \mathbf{w}_l . For frame design it is convenient to collect the vectors, $\{\mathbf{x}_l\}$, $\{\tilde{\mathbf{x}}_l\}$, and $\{\mathbf{w}_l\}$, into matrices, \mathbf{X} , $\tilde{\mathbf{X}}$, and \mathbf{W} . The size of the matrices will be $N \times L$, $N \times L$, and $K \times L$ respectively. The synthesis equation may now be written

$$\tilde{\mathbf{X}} = \mathbf{F} \mathbf{W}. \quad (3)$$

The sparsity of the representation is expressed by the *sparseness factor*

$$S = \frac{\text{number of non-zero coefficients in } \mathbf{W}}{\text{number of samples in } \tilde{\mathbf{X}}}. \quad (4)$$

We point out that this is a global definition; for each column of $\tilde{\mathbf{X}}$ a larger or a smaller number of vectors than the average, SN , may be selected. The use of a global sparseness constraint makes the vector selection procedure more computational demanding, but the benefit is that we get a better global approximation quality compared to the situation where we select the same number of vectors, SN , for each column of $\tilde{\mathbf{X}}$.

2.1. General design method

The optimal frame will depend on the target sparseness factor and the class of signals we want to represent. The problem of finding the optimal frame, \mathbf{F} , for a given class of signals and a given sparseness factor, was treated in [12] and can briefly be summarized as follows: We want to find the frame, \mathbf{F} , and the sparse weights, \mathbf{W} , that minimize the sum of the squared errors. The object function, which we want to minimize, is

$$J = J(\mathbf{F}, \mathbf{W}) = \|\mathbf{X} - \tilde{\mathbf{X}}\|^2 = \|\mathbf{X} - \mathbf{F}\mathbf{W}\|^2. \quad (5)$$

The norm used is $\|\mathbf{A}\|^2 = \text{trace}(\mathbf{A}^T \mathbf{A})$. This norm is used both for matrices and vectors, for vectors it corresponds to the 2-norm. To find the optimal solution to this problem is difficult if not impossible. We split the problem into two parts to make it easier to solve (that is find a reasonably good solution) by using an iterative algorithm.

The algorithm starts with a user supplied initial frame $\mathbf{F}^{(0)}$ and then improves it by iteratively repeating two main steps:

1. $\mathbf{W}^{(i)}$ is found by vector selection using frame $\mathbf{F}^{(i)}$, where the object function is $J(\mathbf{W}) = \|\mathbf{X} - \mathbf{F}^{(i)}\mathbf{W}\|^2$.
2. $\mathbf{F}^{(i+1)}$ is found from \mathbf{X} and $\mathbf{W}^{(i)}$, where the object function is $J(\mathbf{F}) = \|\mathbf{X} - \mathbf{F}\mathbf{W}^{(i)}\|^2$. Then we increment i and go to step 1.

i is the iteration number. The first step is sub-optimal due to the use of practical vector selection algorithms, while the second step finds the \mathbf{F} that minimizes the object function. More details on this algorithm can be found in [12].

2.2. Block oriented frame images

The frame design method in Section 2.1 is independent of what the signal vectors and their sparse approximations, \mathbf{x}_l and $\tilde{\mathbf{x}}_l$, actually represent. The method is easily adapted to images. The training images are divided into non-overlapping blocks; we have used blocks of size 8×8 . Traditional orthogonal transforms, such as the DCT, are separable. The separability property puts restrictions on the 2D transform. Training a frame to be separable is probably hard, and since it also would limit the flexibility of the frame, we do the frame based image approximation in a non-separable manner. This means that the $8 \times 8 = 64$ pixels from each image block are reordered, by lexicographically ordering of the columns of the image block, into a 64×1 vector, \mathbf{x}_l .

Since a reconstructed vector, $\tilde{\mathbf{x}}_l$, is built up as a linear combination of a few of the frame vectors, $\{\mathbf{f}_k\}_{k=1}^K$, a reconstructed image block is built up as a linear combination of a few of the



Fig. 1. The training set consists of 8 images, each have size 512×512 pixels, 256 gray levels.

frame images. A frame image is the image corresponding to a frame vector. Note that the frame images are block oriented, they do not overlap with adjacent blocks.

The operator that splits an image into blocks and order each block as a column vector is denoted R . This operator conserves the (trace) norm, $\|R\mathbf{A}\| = \|\mathbf{A}\|$. We use \mathbf{A} for general matrices and \mathbf{B} for matrices that represent images. For the eighth ($M = 8$) training images in Figure 1, denoted $\{\mathbf{B}_m\}_{m=1}^M$, \mathbf{X} is formed as

$$\begin{aligned} \mathbf{X} &= [\mathbf{X}_1 | \mathbf{X}_2 | \dots | \mathbf{X}_M], \text{ where} \\ \mathbf{X}_m &= R\mathbf{B}_m, \quad m = 1, 2, \dots, M. \end{aligned} \quad (6)$$

The size of \mathbf{X}_m is $64 \times L_m$, where L_m is number of pixels in \mathbf{B}_m divided by 64. An approximation to \mathbf{X} is made as in Equation 3, for each training image individually as $\tilde{\mathbf{X}}_m = \mathbf{F}\mathbf{W}_m$. A reconstructed image is $\tilde{\mathbf{B}}_m = R^{-1}\tilde{\mathbf{X}}_m$. As we want to use the same sparseness factors for each image, we have for each image

$$S = \frac{\text{number of non-zero coefficients in } \mathbf{W}_m}{\text{number of pixels in } \mathbf{B}_m}. \quad (7)$$

Frame design can now be done by the method described in Section 2.1.

2.3. Overlapping frames

The essence of [11], which is currently in the review phase, is as follows: For a one dimensional signal and the block oriented case the synthesis equation for several adjacent blocks can be written as $\tilde{\mathbf{x}} = \mathcal{F}\mathbf{w}$ or

$$\begin{bmatrix} \vdots \\ \tilde{\mathbf{x}}_l \\ \tilde{\mathbf{x}}_{l+1} \\ \tilde{\mathbf{x}}_{l+2} \\ \vdots \end{bmatrix} = \begin{bmatrix} \ddots & & & & \\ & \mathbf{F} & & & \\ & & \mathbf{F} & & \\ & & & \mathbf{F} & \\ & & & & \ddots \end{bmatrix} \begin{bmatrix} \vdots \\ \mathbf{w}_l \\ \mathbf{w}_{l+1} \\ \mathbf{w}_{l+2} \\ \vdots \end{bmatrix}. \quad (8)$$

When we extend this to the general overlapping case, the large frame, \mathcal{F} , can be written as

$$T\mathbf{A} = \mathbf{T}_v \mathbf{A} \mathbf{T}_h^T. \quad (12)$$

\mathbf{T}_v and \mathbf{T}_h are selected to be matrix representations of orthogonal critically sampled filter banks, they have the same structure as \mathcal{G} in Equation 10. T is now a separable and unitary operator and it conserves the (trace) norm, $\|T\mathbf{A}\| = \|\mathbf{A}\|$. The important point is that T , and the reorder operator R defined in Section 2.2, are operators that conserve the (trace) norm.

The extension from the one dimensional case to the two dimensional case can be done by transferring the image into the coefficient domain, where we denote it $\mathbf{Y} = T\mathbf{B}$. The image in the coefficient domain, \mathbf{Y} , has the same size as the image itself, \mathbf{B} . Another important fact is that \mathbf{Y} may be divided into (for the experiments here 8×8) blocks in a sensible way, by grouping the coefficients for the different basis images together.

As for the one dimensional case the approximation is done in the coefficient domain rather than in the signal domain. To achieve this, \mathbf{X} is defined differently than for the block oriented case in Equation 6:

$$\begin{aligned} \mathbf{X} &= [\mathbf{X}_1 | \mathbf{X}_2 | \dots | \mathbf{X}_M], \text{ where} \\ \mathbf{X}_m &= R(\mathbf{Y}_m) = R(T\mathbf{B}_m), \quad m = 1, 2, \dots, M. \end{aligned} \quad (13)$$

The design method of Section 2.1 may now be used as before. The frame \mathbf{F} is now used to represent blocks of the image in the coefficient domain, and not as for the block oriented case, blocks of the image itself. This makes it easy to design a frame where the *frame images* overlap, since each frame image now will be a linear combination of the overlapping basis images of the filter bank T .

An approximation to \mathbf{X} is made as in Equation 3, and the reconstructed image is given as:

$$\tilde{\mathbf{B}}_m = T^{-1} \tilde{\mathbf{Y}}_m = T^{-1} R^{-1} \tilde{\mathbf{X}}_m = T^{-1} R^{-1} \mathbf{F} \mathbf{W}_m, \quad (14)$$

where \mathbf{W}_m , as before, is the weight matrix calculated by a vector selection algorithm. Using the same notation as in Equation 14 the basis images of T may be written as $\{T^{-1} R^{-1} \mathbf{e}_n\}_{n=1}^{64}$ where \mathbf{e}_n is a unit vector (size 64×1) with a 1 in position n and zeros elsewhere. Some basis images are shown in Figure 2. The frame images will now be a linear combination of such basis images. A frame image may be written as $(T^{-1} R^{-1} \mathbf{f}_k)$. Some of these frame images are shown in Figure 3, in the left part for the initial frame in the right part and for the optimized frame.

Investigating the object function from Equation 5, we have

$$\begin{aligned} J &= \|\mathbf{X} - \tilde{\mathbf{X}}\|^2 = \sum_{m=1}^M \|\mathbf{X}_m - \tilde{\mathbf{X}}_m\|^2 \\ &= \sum_{m=1}^M \|T^{-1} R^{-1} (\mathbf{X}_m - \mathbf{F} \mathbf{W}_m)\|^2 \\ &= \sum_{m=1}^M \|\mathbf{B}_m - \tilde{\mathbf{B}}_m\|^2. \end{aligned} \quad (15)$$

J is minimized in the design algorithm, i.e. we minimize the sum of the squared norms of the residual images. Peak Signal to Noise Ratio (PSNR) is defined using this measure

$$\text{PSNR} = 10 \log_{10} \frac{255^2 \cdot (\text{number of pixels})}{\|\mathbf{B} - \tilde{\mathbf{B}}\|^2}. \quad (16)$$

3. PRACTICAL ADAPPTIONS

In the previous section we described the most important modification to the frame design methods presented earlier which were needed when going from the one dimensional case to the two dimensional case and from the block oriented case to the overlapping case. Here we describe some minor modifications giving some improvements in the results. The impatient reader may jump directly to the experimental part without missing the essential ideas.

3.1. Subtracting the low pass image

Several observations motivate subtracting the low pass image from the original image. Vector selection schemes have been shown experimentally to work well when the signal has zero mean, [9]. Due to the large variation in the mean in different parts of an image, it is more advantageous to subtract a locally varying mean than a global mean. Similarly, in DCT based coding schemes (JPEG) the DC components of each image block are handled separately.

We have chosen an approach using a separable biorthogonal low-pass filter, in each dimension downsampled by a factor of 16, to get a low pass version of the image. To get the detailed image we take the original image and subtract the reconstructed low pass image. This is illustrated in Figure 4. For a 512×512 image the low pass image may be stored (without further compression) in 1024 bytes, that is a bit rate of 0.03 bits per pixel. For the test images, Lena and Barbara in Figure 5, PSNR for the low pass images are 23.66 and 19.89 respectively. Since we have already used one coefficient value per 16×16 block, we should use a sparseness factor of $S = S_t - \frac{1}{256}$ in vector selection on the detailed image, where S_t is the actual (total) target sparseness factor.

3.2. Global Matching Pursuit

Images are largely non-stationary and we should select more vectors to represent parts with many details than relative flat parts. Prior to vector selection the image is represented as a matrix, \mathbf{X}_m of size $64 \times L_m$. \mathbf{X}_m is formed as in Equation 6 for the block oriented case or as in Equation 13, in the coefficient domain, for the overlapping case. In both cases the problem is to find the most appropriate number of frame vectors, $\{S'(l)\}_{l=1}^{L_m}$, to use for approximating each column of \mathbf{X}_m , under the constraint, $\sum_{l=1}^{L_m} S'(l) = SNL_m$.

Matching Pursuit (MP), as presented in [1], approximate each signal vector (column of \mathbf{X}_m) independently of other signal vectors, using vectors from a frame. The frame vectors are selected, one at the time, until some stop criterion is met. The stop criterion could be that a predefined number of frame vectors are selected or that an acceptable small value on the residual is achieved (used in [9]). This does not give precise control of the total number of selected vectors in the entire image.

To solve this a variant of MP, Global Matching Pursuit (GMP), was made. The inputs are the matrix \mathbf{X} , the frame \mathbf{F} and the sparseness factor ($0 < S < 1$). For notational convenience we have dropped the index m . The output is the weight matrix, \mathbf{W} , selected so that the sparseness constraint is fulfilled, that is the number of non-zero elements in \mathbf{W} is SNL , and at the same time the error $\|\mathbf{X} - \mathbf{F}\mathbf{W}\|$ is as small as possible. The distribution vector $S'(l)$ is given as the number of non-zero elements in each

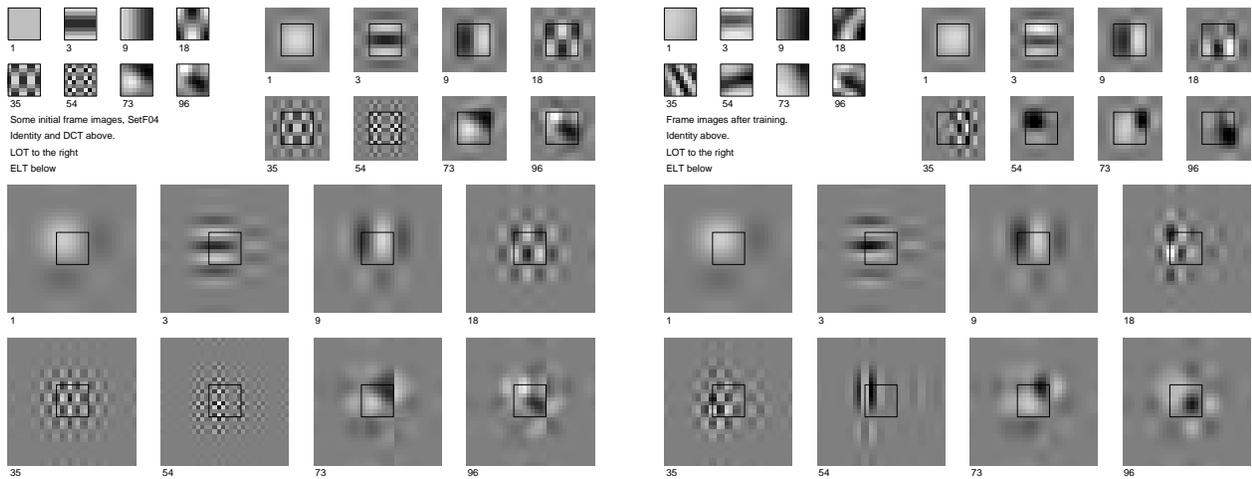


Fig. 3. To the left we have some of the initial frame vectors or frame images set by using basis images and random values. To the right we have some of the frame images after training. We have used the target sparseness factor, $S = \frac{12}{256}$.



Fig. 4. An image (to the left) is splitted into an low-pass image (in the middle) and an image of details (to the right).



Fig. 5. The test images. Lena of size 512×512 and Barbara of size 576×720

column of \mathbf{W} . The algorithm for doing global vector selection is as follows:

1. Set the weight matrix to zero, $\mathbf{W} = \mathbf{0}$. Find all the inner products, $\mathbf{Q} = \mathbf{F}^T \mathbf{X}$. The size of \mathbf{Q} is $K \times L$, $Q(k, l)$ is the inner product of frame vector k and column l in \mathbf{X} .
2. Find the largest (in absolute value) SL inner product of \mathbf{Q} but not more than one from each column. This gives a set of SL indices, simply denoted (k, l) .
3. For each of the SL different pair of indices (k, l) , use the frame vector given by k to approximate column l of \mathbf{X} :
 - (a) Update the weight, $W(k, l) = W(k, l) + Q(k, l)$.
 - (b) Update column l of \mathbf{X} , $X(:, l) = X(:, l) - Q(k, l) \cdot F(:, k)$
 - (c) Update the inner products for vector l of \mathbf{X} , $Q(:, l) = \mathbf{F}^T \cdot X(:, l)$
 - (d) Stop if we have selected enough weights.
4. If more frame vectors need to be selected goto 2.

The modification compared with ordinary MP can be seen in step 2. Ordinary MP would select the L largest inner products, one for each column of \mathbf{Q} . By reducing the number of selected inner products we make sure that the loop in step 3 runs at least N times, making it possible to select many frame vectors for those vectors (columns) of \mathbf{X} that require so.

By using $S'(l)$, found from \mathbf{W} returned by GMP, as the number of frame vectors to use for each vector in matrix \mathbf{X} in the FOMP algorithm [3] we can reduce the error even more. This is because FOMP often finds better coefficients than MP. Thus the best global vector selection algorithm we have found is the combination GMP+FOMP. Note that we use GMP on each image by itself, not globally on all the images.

GMP+FOMP is computationally demanding using approximately 5 times more time than FOMP only, when implemented in Matlab. FOMP need $S'(l)$ as an extra input, so FOMP alone is only an alternative if we already have found the distribution.

Experiments have confirmed that the frame design method performs better when the proposed modifications are used.

4. DESIGN EXAMPLES AND DISCUSSION

The described method is used to design frames for representing images. The overlap factor for the frame images will be decided by the choice of T . Three different T 's are used: the identity transform, corresponding to the pure block oriented case, the 8 channel, 16 taps LOT, and the 8 channel, 32 taps ELT. The size of the frame \mathbf{F} in all the experiments is 64×128 . As initial frame we chose values such that the first 64 frame images are the basis images of the DCT (when T is identity), or the basis images of the LOT or the ELT. Some of these basis images are shown in Figure 2. The rest of the frame images are random low pass images. Some of the initial frame images are shown in the left part of Figure 3.

The images in Figure 1 are used as training data. They are preprocessed by the low pass filtering described in Section 3.1, and the detail images are used for the rest of the process. Each detail image is transformed into a matrix of training vectors, as in Equation 6 for the block oriented case (T is identity) and as in Equation 13 for the overlapping cases (T is LOT and ELT), each case giving one set of training data. Frames are trained for seven different target sparseness factors, belonging to the set

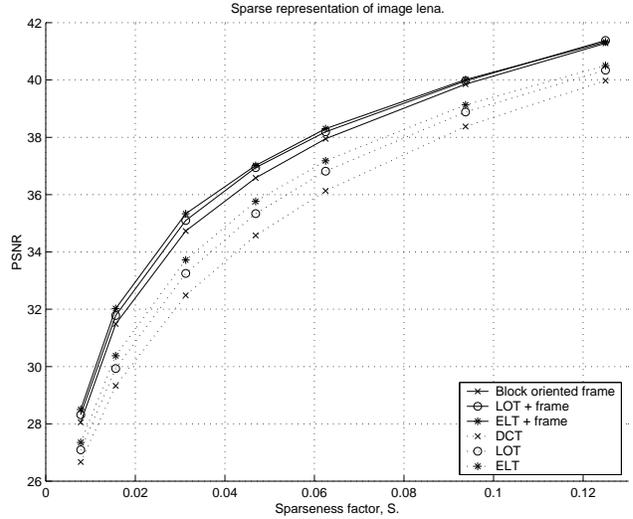


Fig. 6. Sparse representation results for Lena.

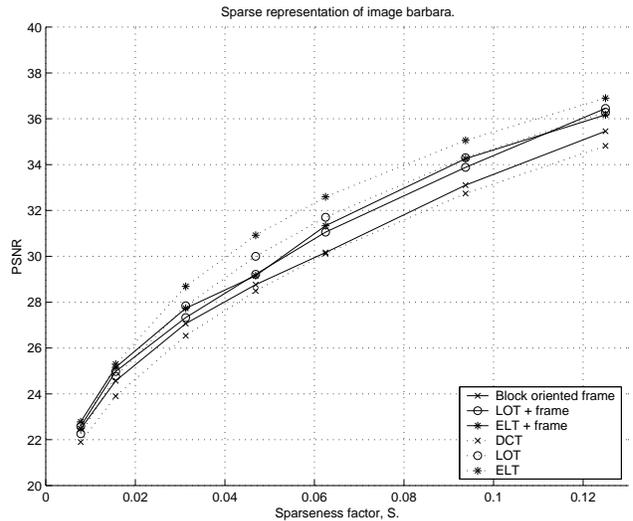


Fig. 7. Sparse representation results for Barbara.

$S_t \in \{ \frac{2}{256}, \frac{4}{256}, \frac{8}{256}, \frac{12}{256}, \frac{16}{256}, \frac{24}{256}, \frac{32}{256} \}$. Since the low pass image has used $\frac{1}{256}$ of the sparseness factor, the sparseness factor used as constraint in Global Matching Pursuit (GMP) will be $S = S_t - \frac{1}{256}$. 21 frames were designed in the experiments.

The frames have been optimized to represent images in a sparse manner, and the frames' capability of sparse representation of an image is examined. One application of sparse representations is compression, but it is not the only one. We focus on the representation capabilities of the frames in these experiments, consequently we do not include quantization or coding of the non-zero coefficients.

We use two test images, shown in Figure 5. An experiment was done on the test image Lena, measuring the quantitative quality in terms of PSNR. For each of the 21 optimized frames, the image was represented using the same sparseness factor as the frame was designed for. Figure 6 shows these calculated PSNR values as solid lines, one line for each of the different types of transform T . We see that the overlapping frames do more than 0.5 dB bet-

ter than block oriented frame, when the sparseness factor is small. For larger sparseness factors the difference is smaller. The frame which use a T based on ELT do marginally better than the frame which use a T based on LOT. For the sake of comparison we also tested sparse representation using the DCT (8×8 blocks) and the (8 channel) filter banks that use LOT and ELT. We keep only the SNL , the sparseness factor multiplied by the number of pixels, largest coefficients, setting the rest to zero. The results are shown in Figure 6 as dotted lines. The filter banks perform considerable better than the DCT, and the frames performs significantly better than the orthogonal filters (DCT, LOT and ELT).

The same sparse representation test was done on the test image Barbara, the results are shown in Figure 7. Also on this image the filter banks and the overlapping frames perform considerably better than the DCT and the block oriented frame. But the optimized frames perform worse than the orthogonal filter banks. This is probably because the Barbara image is quite different from the training images, Barbara include more of high frequency components. The high frequency basis images are to a large degree removed from the frame images during training. This is clearly illustrated when we see what happens with the high frequency frame image numbered 54 in Figure 3, the left part is initial frame images and the right part is the frame images after training. The test image Lena is more similar to the training images than the test image Barbara.

Some qualitative tests are done using both Lena and Barbara as test images. To get noticeable differences and artifacts, a sparseness factor as low as $S = \frac{4}{256}$ was used. We focus on a small detail of the reconstructed images, and show this for both the traditional methods and the designed frames. The result for the detail of Lena is shown in Figure 8, and for Barbara in Figure 9.

On Lena we clearly see the results of higher PSNR for the frames, the quality is overall better. We notice that the blockiness is considerable reduced on the block oriented frame compared to the DCT case, but it is still present. The LOT and ELT cases are not blocky, but we notice some ringing (along the edge of the hat below the eye, and above the eyebrow) and some smearing of details (the spot on the hat above the eye). The best visual quality is provided by the frame based on the ELT, and this is also the one with the best PSNR.

On Barbara the PSNR's are lower than on Lena and the quality of all the reconstructed images are worse. None of them are able to show the shawl at this low sparseness factor. The blockiness of the block oriented frame is here easily seen, but it performs much better than the DCT based reconstruction. The ringing in the LOT and ELT bases reconstructions is most visible in the chin where the stripes of the shawl appear. Note that even though the reconstruction based on the largest ELT coefficients performs best in terms of PSNR, the reconstruction using the ELT based frame is the one that represent the face best (less smearing and the tip of the nose and the eyes are better visualized). This may be because Barbara's face is more within the class of images in the training set than the rest of the Barbara image.

5. CONCLUSION

Overlapping frames for images used with a sparsity constraint performs well, both in quantitative and qualitative measures. Frames do better than transforms, which is very much as expected since the frame is overcomplete. Also overlapping frames do better than block oriented frames, however the improvement is not as large as when we compare filter banks (LOT and ELT) to a transform (DCT).

6. REFERENCES

- [1] S. G. Mallat and Z. Zhang, "Matching pursuit with time-frequency dictionaries," *IEEE Trans. Signal Processing*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.
- [2] G. Davis, *Adaptive Nonlinear Approximations*, Ph.D. thesis, New York University, Sept. 1994.
- [3] M. Gharavi-Alkhansari and T. S. Huang, "A fast orthogonal matching pursuit algorithm," in *Proc. ICASSP '98*, Seattle, USA, May 1998, pp. 1389–1392.
- [4] A.P. Berg and W.B. Mikhael, "An efficient structure and algorithm for image representation using nonorthogonal basis images," *IEEE Trans. Circuits, Syst. II: Analog and Digital Signal Processing*, vol. 44, no. 10, pp. 818–828, Oct. 1997.
- [5] R. Neff and A. Zakhor, "Very low bit-rate video coding based on matching pursuit," *IEEE Trans. Circuits, Syst. for Video Tech.*, vol. 7, no. 1, pp. 158–171, Feb. 1997.
- [6] M. Donahue, D. Geiger, R. Hummel, and T.-L. Liu, "Sparse representations for image decompositions with occlusions," in *Proc. Computer Vision and Pattern Recognition*, San Francisco, USA, 1996, pp. 7–12.
- [7] H. R. Rabiee, R. L. Kashyap, and S. R. Safavian, "Adaptive image representation with segmented orthogonal matching pursuit," in *Int. Conf. on Image Proc.*, Hillsboro, USA, Oct. 1998, vol. 2, pp. 233–236.
- [8] V. DeBrunner, L. Chen, and H. Li, "On the use of (lapped) multiple transforms in still image compression," in *Proc. International Conference on Image Processing*, CA, USA, 1995, vol. 1, pp. 294–297.
- [9] Kjersti Engan, *Frame Based Signal Representation and Compression*, Ph.D. thesis, Norges teknisk-naturvitenskapelige universitet (NTNU)/Høgskolen i Stavanger, Sept. 2000, Available at <http://www.ux.his.no/~kjersti/>.
- [10] S. O. Aase, K. Skretting, J. H. Husøy, and K. Engan, "Design of signal expansions for sparse representation," in *Proc. ICASSP 2000*, Istanbul, Turkey, June 2000, pp. 105–108.
- [11] K. Skretting, J. H. Husøy, and S. O. Aase, "A simple design of sparse signal representations using overlapping frames," Submitted for publication, available at <http://www.ux.his.no/~karlsk/>.
- [12] K. Engan, S. O. Aase, and J. H. Husøy, "Multi-frame compression: Theory and design," *Signal Processing*, vol. 80, pp. 2121–2140, Oct. 2000.
- [13] Tor Audun Ramstad, Sven Ole Aase, and John Håkon Husøy, *Subband Compression of Images – Principles and Examples*, ELSEVIER Science Publishers BV, North Holland, 1995.
- [14] H. S. Malvar and D. H. Staelin, "The LOT: Transform coding without blocking effects," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 4, pp. 553–559, Apr. 1989.
- [15] H. S. Malvar, "Extended lapped transforms: Fast algorithms and applications," in *Proc. ICASSP '91*, Toronto, Canada, May 1991, pp. 1797–1800.

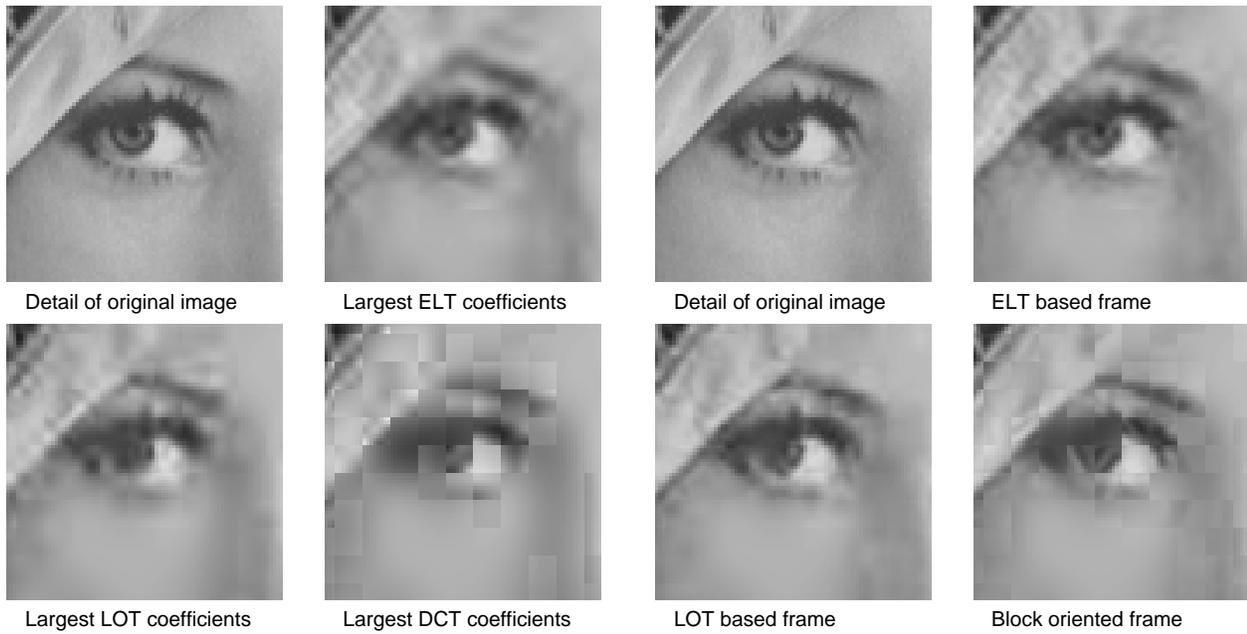


Fig. 8. Qualitative results on Lena. To the left we have the results using the largest coefficient from orthogonal transforms and filter banks, DCT, LOT and ELT. To the right we have the results when we have used frames for sparse representation. The sparseness factor is $S = \frac{1}{64}$ for all images. The PSNR, as plotted in Figure 6, are for the orthogonal cases 29.33 (DCT), 29.93 (LOT), 30.38 (ELT) and for the frames based on these 31.49, 31.78 and 32.02.

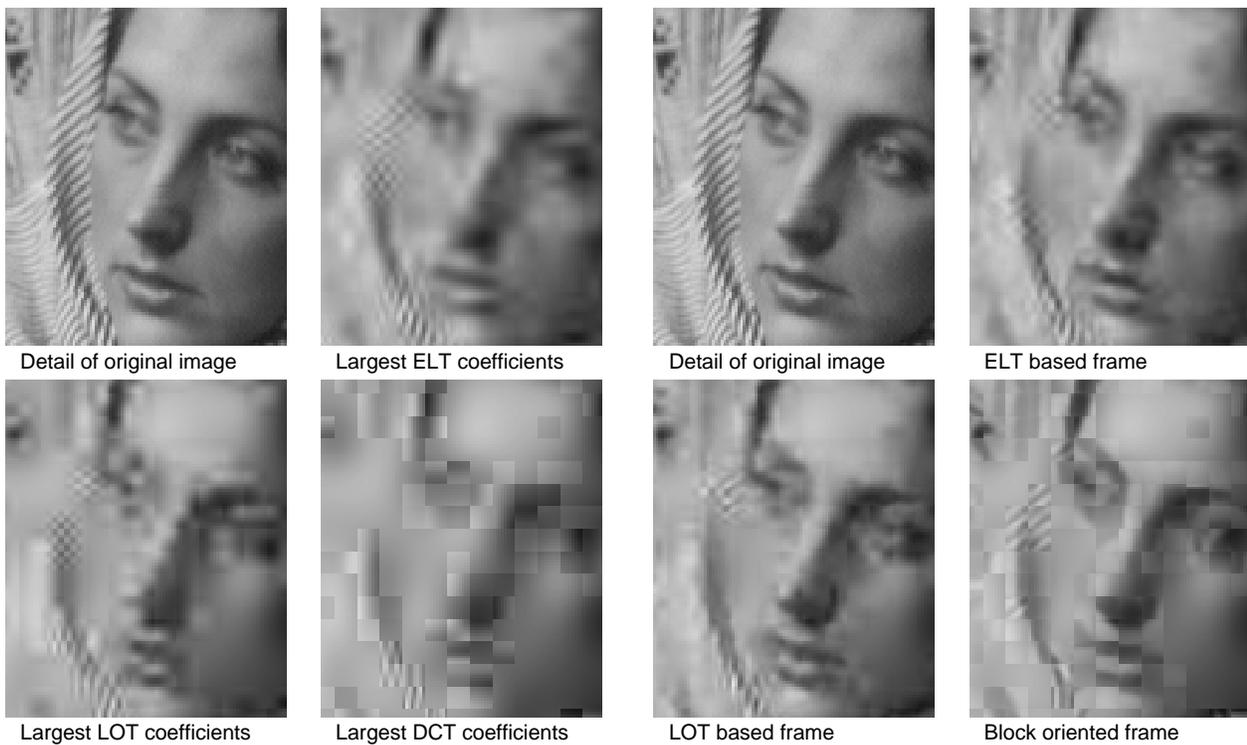


Fig. 9. Like Lena in Figure 8 but here a detail of Barbara. The PSNR, as plotted in Figure 7, are for the orthogonal cases 23.90 (DCT), 24.76 (LOT), 25.30 (ELT) and for the frames based on these 24.57, 24.98 and 25.16.