

Efficient Reconstruction of Frames Based Joint Source-Channel Coded Streams in the Presence of Data Loss

R. Bernardini, A. Tonello, R. Rinaldo

Università di Udine, Dipartimento di Ingegneria Elettrica, Gestionale e Meccanica,
Via delle Scienze 208, Udine, Italy
{bernardini, tonello, rinaldo}@uniud.it

Abstract

Due to their interactive nature, multimedia streams must be sent over UDP with suitable countermeasures for minimizing the effect of data loss. Coding with oversampled filter bank is a promising multiple description technique which allows for signal reconstruction even in presence of packet loss. Although the coded signal can be reconstructed by means of a synthesis filter bank if no coefficient is lost, reconstruction is not trivial in the presence of packet loss. Indeed, although a theoretical solution to such a problem is well-known, its straightforward application would require to wait for all the coefficients to arrive, introducing excessive delay in interactive applications. In this paper we propose a novel alternative approach which puts in front of the synthesis filter bank a “restoring stage” whose purpose is to reconstruct the lost coefficients. It is shown that our approach can be applied in interactive applications, while preserving the optimality properties of the theoretical solution.

1. Introduction

Playing multimedia streams over wireless networks is quite challenging. As a matter of fact, the available bandwidth in a wireless system is limited, thus calling for efficient compression. On the other end, fading channel unreliability require error protection. Note that in the case of multimedia streams, retransmission is useless since the data would arrive too late to meet the constraints on reproduction delay.

Typically the two goals of compression and error protection are achieved by two separate modules, i.e., a source coder followed by a channel coder. Recently, multiple description coding has been suggested as a joint source/channel scheme able to consider both aspect together [1, 2, 3]. In particular, coding with oversampled filter banks has been proposed as a possible solution [4, 5, 6].

A main problem with oversampled filter banks is the reconstruction of the original signal. Such a problem (which can be seen as the infinite-dimensional counterpart of an

over-determined system) can be easily solved by recognizing that processing a signal with an oversampled filter banks can be interpreted as the analysis of a vector (the input signal) by means of a *frame* of vectors [7]. Such an observation allows one to use the known results of frame theory to the case of redundant signal coding. More specifically, it is well-known that the original signal can be reconstructed by means of the *dual* (filter bank) of the analysis part. This allows for an efficient reconstruction of the coded signal.

Coefficient loss can be modeled by supposing that the signal was coded by using a *punctured frame*, i.e., a frame obtained by deleting some vectors from the original one. Even in this case, from a theoretical point of view, reconstruction is not a problem: one computes the dual of the punctured frame and use it to reconstruct the signal (note that the solution of forcing to zero the lost coefficients does not work).

Unfortunately, from a practical point of view, the implementation becomes much more complex: both because the dual frame will not have a filter bank structure anymore and, more important, because the known algorithms (even the fast ones [9]) for frame reconstruction require the knowledge of the whole frame structure. This implies that one should wait for all the future coefficients, in order to know which functions have been deleted. Moreover, this solution, which calls for storing the whole set of coefficients, could prove to be especially unfeasible for mobile applications where the amount of available memory can be limited.

It is worth emphasizing that the reconstruction delay introduced by a naïve application of the dual frame result is *not* due to *computational complexity*, but to the necessity of *knowing the whole loss pattern* before computing the dual frame

In this paper, we present a novel scheme for reconstruction in presence of packet loss. Our scheme puts in front of the standard filter bank synthesis stage a *restoration stage* which recovers, whenever possible, the missing coefficients from the received ones, “hiding” the losses to the synthesis filter bank. It can be shown that this is equivalent to using the dual of the subframe corresponding to

the packet losses. From a mathematical point of view, this corresponds to building the pseudo-inverse of the *lossy* analysis operator. More generally, we can recognize the case when the received coefficients do not correspond to a subframe and use appropriate concealment techniques. By specializing the structure to the case of one-dimensional FIR filter banks, we show that the restoration stage operates “locally,” that is, it reconstructs the missing coefficients as soon as enough information is available. This allows to use the proposed structure in low-delay applications. Note that the known fast algorithms for the construction of the dual frame, such as the one given in [9], are based on a different procedure and require knowledge of the complete loss pattern in order to carry out the reconstruction. This prevents their direct application in low-delay contexts, where signal portions are reconstructed as soon as the corresponding coefficients become available.

2. Oversampled filter banks and frames

A possible way to achieve some resilience against packet losses is to code signal x by means of an oversampled filter bank

$$y_c[n] = \sum_{m \in \mathbb{Z}} x(m) h_c[Mn - m] \quad (1)$$

where h_c , $c = 1, \dots, N$, is the impulse response of the c -th channel and $N > M$. Oversampling implies that there is redundancy in the coefficients $y_c[n]$. It is intuitive that by exploiting such a redundancy one can reconstruct x even if some coefficients are lost.

The main question with this approach is how to reconstruct x from $y_c[n]$. This problem is the infinite-dimensional counterpart of an overdetermined linear system $\mathbf{y} = \mathbf{F}\mathbf{x}$ where \mathbf{F} is a full rank $N \times M$ matrix with $N > M$. If \mathbf{y} belongs to $\text{Im}(\mathbf{F})$, the space generated by the columns of \mathbf{F} , one can find \mathbf{x} by left-multiplying \mathbf{y} by a left-inverse of \mathbf{F} , that is, a matrix \mathbf{F}^{-L} such that $\mathbf{F}^{-L}\mathbf{F} = \mathbf{I}$. Note that, as long as $\mathbf{y} \in \text{Im}(\mathbf{F})$, it does not matter which left-inverse we use¹. However, if $\mathbf{y} \notin \text{Im}(\mathbf{F})$, and no *a priori* information about \mathbf{x} is available, then one could want to solve the system in a least square sense, i.e., to find $\hat{\mathbf{x}}$ such that $\mathbf{y} - \mathbf{F}\hat{\mathbf{x}}$ has minimum length (this corresponds to the maximum-likelihood estimate of \mathbf{x} if $\mathbf{y} = \mathbf{F}\mathbf{x} + \varepsilon$ where ε is independent \mathbf{x} and is a vector of independent Gaussian variables). It is well-known that $\hat{\mathbf{x}}$ can be obtained as $\hat{\mathbf{x}} = \mathbf{F}^\dagger \mathbf{y}$, where \mathbf{F}^\dagger is the *pseudo-inverse* of \mathbf{F} [8]. It is possible to show that, if the columns of \mathbf{F} are linearly independent, \mathbf{F}^\dagger can be expressed as $\mathbf{F}^\dagger = (\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'$. It follows that \mathbf{F}^\dagger is a left-inverse and it is possible to show that among all the left-inverses of \mathbf{F} , \mathbf{F}^\dagger is the only one which maps the orthogonal complement of $\text{Im}(\mathbf{F})$ to zero [8]. The least square property of \mathbf{F}^\dagger descends easily from the fact that $\mathbf{F}\mathbf{F}^\dagger \mathbf{y}$ is the orthogonal projection of \mathbf{y} onto $\text{Im}(\mathbf{F})$ [8].

¹The left inverse is not unique unless \mathbf{F} is square.

In the oversampled filter bank context, the counterparts of \mathbf{y} , \mathbf{x} and \mathbf{F} are, respectively, the sequence of received coefficients $\hat{y}_c[n]$, the input signal x and the linear map F associated with the analysis filter bank. As a matter of fact, one can express the synthesis filter bank operation as an infinite dimension matrix-vector product [10]. Since values $\hat{y}_c[n]$ are obtained by quantizing the filter bank output values $y_c[n]$, the sequence of received coefficients will almost certainly not belong to the analogue of $\text{Im}(\mathbf{F})$, i.e., the linear space of the sequences which can be generated by the analysis filter bank. Similarly to the finite-dimensional case, if no *a priori* information is available about x , the “best” reconstruction of x is the one which best explains the received values, i.e., the signal \hat{x} such that the distance between $F\hat{x}$ and $\hat{\mathbf{y}}$ is minimum. This suggests to reconstruct x by using the “pseudo-inverse” of filter bank (1). As a matter of fact, Eq. (1) can be interpreted as a scalar product between the input sequence and the analysis function $\phi_k = h_c^*[Mn - \cdot]$, with $k = c + nN$. In operator form, we can write $y_k = (Fx)_k \triangleq \langle x, \phi_k \rangle$. In the case of an oversampled filter bank, functions ϕ_k constitute a *frame*, and the general reconstruction formula uses the pseudo-inverse F^\dagger of F , namely [9, 10]

$$\hat{x} = F^\dagger \hat{\mathbf{y}} = (F^*F)^{-1}F^* \hat{\mathbf{y}} = (F^*F)^{-1} \sum_{k \in \mathbb{Z}} \phi_k \hat{y}_k = \sum_{k \in \mathbb{Z}} \tilde{\phi}_k \hat{y}_k \quad (2)$$

The reconstructed signal is obtained by linearly combining, with coefficients \hat{y}_k , functions $\tilde{\phi}_k = (F^*F)^{-1}\phi_k$, which are the infinite-dimensional counterpart of the columns of \mathbf{F}^\dagger . Set $\tilde{\Phi} \triangleq \{\tilde{\phi}_k\}_{k \in \mathbb{Z}}$ is called the *dual frame* of $\Phi \triangleq \{\phi_k\}_{k \in \mathbb{Z}}$ [9, 10].

3. Problem statement and solution

According to the results of Section 2, the best reconstruction of x from the coefficients obtained with an oversampled filter bank (1) is carried out by means of the dual frame. It is possible to show that, in general, the dual frame of a filter bank also has a filter bank structure with filters \tilde{h}_c [7]. Unfortunately, such a nice result will lose validity just for the case when coding via frames would help, i.e., when some coefficients are lost. Moreover, the support of the dual frame functions can in general be larger than that of the filters \tilde{h}_c .

The solution we propose requires to put in front of the standard synthesis filter bank \tilde{h}_c (the dual of the original frame when no coefficients are lost) a *restoration stage* which recovers, whenever possible, the missing coefficients from the received ones. The restoration stage will also detect when the corresponding subset of analysis functions is not a frame anymore and communicate to the synthesis stage which coefficients cannot be recovered in order to allow the application of error concealment techniques. It will turn out that if both analysis and synthesis

banks are made of FIR filters, then the restoration stage operates “locally”, that is, as soon as enough coefficients are received, it reconstructs the lost ones and sends them to the synthesis stage. This allows for the application of the proposed structure to low-delay applications.

Let I be the set of the indexes of the lost coefficients and let F_I be the analysis operator relative to the subset $\Phi_I = \{\phi_k, k \notin I\}$ of the remaining analysis functions. The coefficient loss can be mathematically represented by the operator

$$(\chi_I y)_n \triangleq \begin{cases} y_n & \text{if } n \notin I \\ 0 & \text{if } n \in I \end{cases} \quad (3)$$

By using the just introduced notation we can finally formalize our goal: *we are searching for a linear restoring operator $R : \ell^2(I^c) \rightarrow \ell^2(\mathbb{Z})$ such that $F^\dagger R$ is the pseudo-inverse of $F_I = \chi_I F$.*

In order to find the restoring operator, it is instrumental to express the missing ϕ_k as a linear combination of the known ones by exploiting the fact that ϕ_k and $\tilde{\phi}_k$ are dual frames. To such an end, observe that one can write, for each $k \in I$,

$$\begin{aligned} \phi_k &= \sum_{n \in \mathbb{Z}} \phi_n \langle \phi_k, \tilde{\phi}_n \rangle, \quad k \in I \\ &= \sum_{n \notin I} \phi_n \langle \phi_k, \tilde{\phi}_n \rangle + \sum_{m \in I} \phi_m \langle \phi_k, \tilde{\phi}_m \rangle \end{aligned} \quad (4)$$

By taking the scalar product of both sides of (4) with x one obtains

$$\begin{aligned} y_k = \langle x, \phi_k \rangle &= \langle x, \sum_{n \notin I} \phi_n \langle \phi_k, \tilde{\phi}_n \rangle \rangle + \langle x, \sum_{m \in I} \phi_m \langle \phi_k, \tilde{\phi}_m \rangle \rangle \\ &= \sum_{n \notin I} y_n \langle \tilde{\phi}_n, \phi_k \rangle + \sum_{m \in I} y_m \langle \tilde{\phi}_m, \phi_k \rangle \end{aligned} \quad (5)$$

Equations (5) (there is one equation for each $k \in I$) can be rewritten in matrix form as

$$Y_I = \mathbf{M}' Y_{I^c} + \mathbf{M} Y_I \quad (6)$$

where Y_I (respectively, Y_{I^c}) is a column vector whose elements are coefficients y_k , with $k \in I$ (respectively, with $k \notin I$) and \mathbf{M} and \mathbf{M}' are matrices whose entry in row k and column m is the scalar product $\langle \tilde{\phi}_m, \phi_k \rangle$. Equation (6) can be solved as

$$Y_I = (\mathbf{I} - \mathbf{M})^{-1} \mathbf{M}' Y_{I^c} \quad (7)$$

It is possible to show that matrix $(\mathbf{I} - \mathbf{M})$ is invertible if and only if the subset Φ_I of the analysis functions corresponding to the received coefficients is still a frame and that, in general, the operator $F^\dagger R$ is the pseudo-inverse of F_I .

If Φ_I is not a frame anymore, it is possible to show that although $F^\dagger R$ is such that $\|\hat{y} - F_I F^\dagger R \hat{y}\|$ is minimum for every $\hat{y} \in \ell^2(I^c)$, it is not the true pseudo-inverse

since $F^\dagger R \hat{y}$ is not the minimum-length solution. This implies that $F_I^\dagger = P_I F^\dagger R$, where P_I is the projection over $\text{span } \Phi_I = \text{Ker}(F_I)^\perp$, that is, the true pseudo-inverse can be obtained by projecting the result of $F^\dagger R$ over $\text{span } \Phi_I$.

Note that while the property of making $\|\hat{y} - F_I F^\dagger R \hat{y}\|$ minimum has a perceptive sense ($F^\dagger R \hat{y}$ “best explains” \hat{y}), the fact that $F_I^\dagger \hat{y}$ has minimum length does not grant that it will be “close” to the real solution. This suggests that $F_I F^\dagger R \hat{y}$ could be, from a perceptive point of view, as good as the true pseudo-inverse solution $F_I^\dagger \hat{y}$. Actually, in the experimental section, we show that using $F^\dagger R \hat{y}$ instead of $F_I^\dagger \hat{y}$ as the reconstructed vector gives similar results from a perspective and objective point of view.

Finally, observe that in the case of oversampled FIR filter banks with FIR dual, the scalar product $\langle \tilde{\phi}_m, \phi_k \rangle$ becomes 0 as soon as the distance between indexes m, k is greater than or equal to an appropriate constant D . It is easy to recognize that this implies that infinite matrix \mathbf{M} has a block-diagonal structure, making it possible to compute the inversion as soon as each block is available.

4. Experimental results

In our experiments we used the 512×512 image “monarch” shown in Fig. 1a and an oversampled filter bank obtained by blockwise processing with 8×8 DCTs the image and its translated version by $(4, 4)$ pixels. This is equivalent to oversample the output of the DCT filter bank on the quincunx lattice $\{4 \cdot (i + 2j, i) : (i, j) \in \mathbb{Z}^2\}$ shown in Fig. 2. The resulting frame is a tight one with redundancy 2.

Define the *additional delay* of a given pixel as the difference between the time when a pixel is reconstructed and the time when the same pixel would have been reconstructed if no packet loss had occurred. It is clear that with no errors, the receiver can reconstruct each pixel with an additional delay equal to zero.

In order to compute the delays introduced by the proposed algorithm it is necessary to know the order used for sending the coefficients. In this specific case we used a row-wise ordering by sending the row of blocks starting from $(0, 0)$, the row of blocks starting from $(4, 4)$, the row of blocks starting from $(8, 0)$ and so on... Such an ordering is exemplified for the case of four blocks per row in Fig. 3.

We simulated the coefficient loss within the hypothesis that every coefficient is lost with probability $P_{\text{loss}} = 0.008$ which corresponds, approximately, to losing one coefficient every two blocks. Fig. 4 shows the histogram of the delays introduced in pixel reconstruction. Since each pixel can be reconstructed only when both its 8×8 blocks are completed, it is clear that the delay will always be a multiple the time T necessary to transmit an 8×8 block. Because of this, the delays in Fig. 4 are measured in units of T . From Fig. 4 one easily deduce that the maximum



(a)



(b)

Fig. 1. (a) 512×512 test image “monarch.” (b) The same image reconstructed by forcing to zero the lost coefficients ($P_{\text{loss}} = 0.008$). The procedure proposed in the paper reconstructs (a) exactly with a 16 row average delay.

delay is $217T$ and the mean delay is approximately equal to $88T$. Both delays are much smaller than the delay introduced by the “standard” algorithm which would wait for all the coefficients to arrive introducing a delay equal to the time necessary to transmit all the blocks, i.e., $1024T$.

The strong peak at $17T$ in Fig. 4 is due to the fact that if a lost coefficient is “isolated” (that is, it is the only coefficient in its block), then reconstruction happens after one row of DCT blocks. The delays less than $17T$ are due to the coefficients which were could not be sent to the synthesis stage before the reconstruction of the lost coefficient.

4.1. Proposed solution vs. true pseudo-inverse

In order to assess the differences between $F^\dagger R$ and F_I^\dagger when Φ_I is not a frame, we simulated packet losses with different values of P_{loss} . The results can be seen in Fig 5

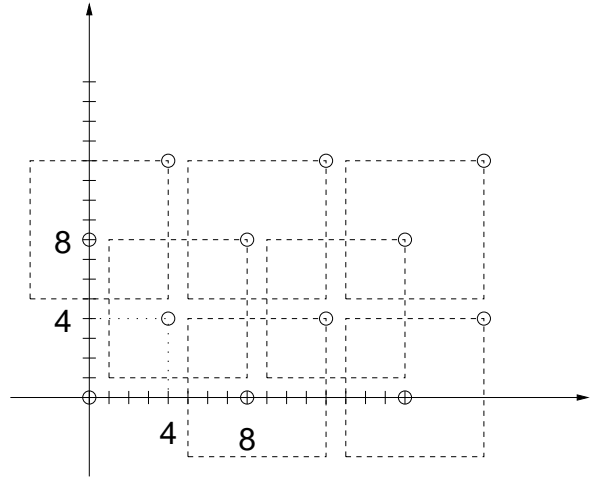


Fig. 2. The dots show the quincunx sampling lattice used with the two-dimensional 8×8 DCT filter bank employed in the experiments. The dashed boxes show the supports of some functions of the corresponding frame.



Fig. 3. Ordering used for sending the coefficients in our experiment. Numbers in the blocks show the order the block coefficients are sent.

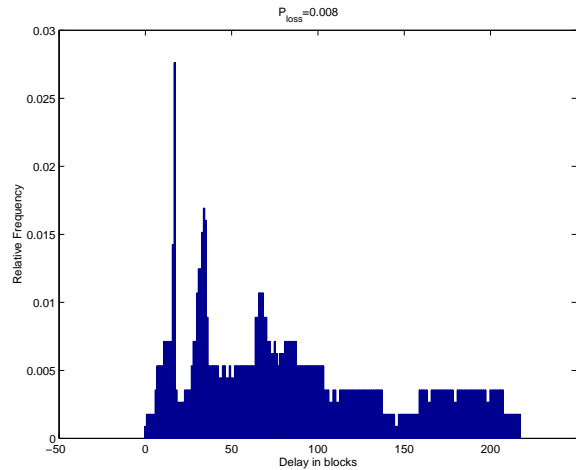


Fig. 4. Histogram of the delays necessary to reconstruct a pixel with the image of Fig. 1a and $P_{\text{loss}} = 0.008$. The unit of the x axis is the time necessary to transmit an 8×8 coefficient block.

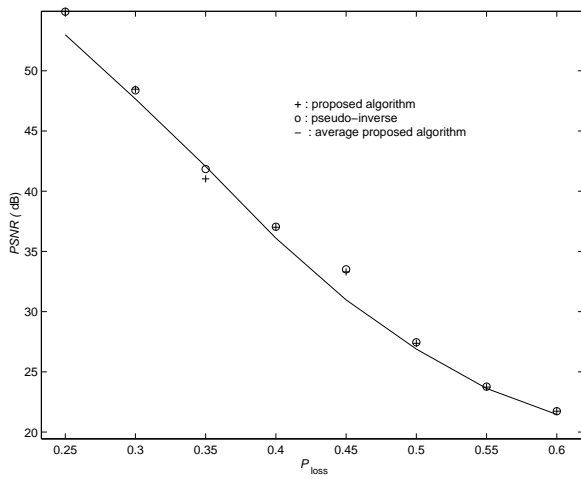


Fig. 5. PSNR vs. P_{loss} for the proposed algorithm with no projection onto $\text{span}\{\Phi_I\}$: one transmission and average. PSNR vs. P_{loss} for the pseudo-inverse solution: one transmission.

which shows the PSNR obtained for different values of P_{loss} . Circles are relative to the results obtained after a single image transmission and using $F^\dagger R \hat{y}$ as the reconstructed vector, while the plus signs are relative to $F_I^\dagger \hat{y}$ for the same loss pattern. One can see from the figure that the overall performance is quite similar in the two cases, and that there is no apparent advantage in performing the projection after the synthesis stage. The plot also reports the average performance when no final projection is computed. The average is computed over 10 independent transmissions for each value of P_{loss} .

5. Conclusions

Using oversampled filter banks in coding applications allows one to reconstruct the coded signal even in presence of coefficient loss. In order to make such a technique effective it is necessary to have a practical algorithm for reconstructing the lost coefficients and for checking which signal parts are definitively lost. In this paper we presented a practical solution to such a problem. The proposed solution puts, before the synthesis filter bank, a restoring stage which reconstructs the missing values and finds which signal parts are definitively lost. If the analysis and synthesis filter banks are made of FIR filters, the

restoring stage operates “locally” reconstructing the lost coefficients as soon as possible. This allows for the application of the proposed structure in low-delay and interactive applications.

6. References

- [1] S. D. Servetto, V. A. Vaishampayan, and N. J. A. Sloane, “Multiple description lattice vector quantization,” in *Proc. IEEE Data Compression Conference*, pp. 13–22, Mar. 1999.
- [2] Y. Wang, M. T. Orchard, and A. R. Reibman, “Multiple description image coding for noisy channels by pairing transform coefficients,” in *Proc. IEEE Workshop on Multimedia Signal Processing*, pp. 419–424, June 1997.
- [3] V. K. Goyal and J. Kovačević, “Generalized multiple description coding with correlating transforms,” *IEEE Transaction on Information Theory*, vol. 47, Sept. 2001.
- [4] V. K. Goyal, M. Vetterli, and N. T. Thao, “Quantized overcomplete expansions in r^n : Analysis, synthesis and algorithms,” *IEEE Transaction on Information Theory*, vol. 44, pp. 16–31, Jan. 1998.
- [5] X. Yang and K. Ramchandran, “Optimal sub-band filter banks for multiple description coding,” *IEEE Transaction on Information Theory*, vol. 46, pp. 2477–2490, Nov. 2000.
- [6] P. Dragotti, J. Kovačević, and V. K. Goyal, “Quantized oversampled filter banks with erasures,” in *Proc. IEEE Data Compression Conference*, pp. 173–182, Mar. 2001.
- [7] Z. Cvetković and M. Vetterli, “Oversampled filter banks,” *IEEE Transaction on Signal Processing*, vol. 46, pp. 1245–1255, May 1998.
- [8] G. Strang, *Linear Algebra and its Applications*. New York: Academic Press, 1980.
- [9] I. Daubechies, *Ten Lecturs on Wavelets*. Philadelphia: SIAM, 1992.
- [10] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1995.