SNR Scalability Based on Matching Pursuits

Christophe De Vleeschouwer and Benoit Macq, *Member, IEEE*

***Abstract*—In this paper, SNR scalable representations of video signals are studied. The investigated codecs are well suited for communications applications because they are all based on backward motion-compensated predictive coding, which provides the necessary low-delay property. In a very-low bit rate context (VLBR), the matching pursuits (MP) signal representation algorithm is used to represent the displaced frame difference (DFD) of each layer of a multilevel decomposition of the video signal. A number of conventional prediction schemes that can be generalized to any DFD representation technique are considered. They are compared with an original and MP specific DFD prediction method. Two scenari have been considered. In the first scenario, an enhancement layer is built on a base layer that has been encoded using a classical, i.e., nonscalable scheme. In that case, all methods appear to be comparable. In the second scenario, the fact that the base layer is used as a reference for an enhancement layer is taken into account to build it. In that case, the proposed MP prediction method clearly outperforms all other conventional approaches. Additional lessons can be drawn from this work. The same motion vectors can be used in both SNR layers, and the DFD prediction between layers improves coding efficiency. Moreover, the MP representation of the signal enable us to measure the predictability of the high SNR layer DFD from the low SNR layer DFD, i.e., to quantify the part of the low SNR layer information that also belongs to the high SNR layer.**

***Index Terms*—Hierarchical video coding, matching pursuits, SNR scalability.**

I. Introduction

In addition to the new requirements for real-time transmission, video applications are quickly evolving from one-to-one communications to one-to-many and many-to-many communications. Widespread use of these applications can easily overload existing networks when the same bits of information have to be transmitted to different users at the same time. Multicast, where a single packet is addressed to all intended recipients, and where the network replicates packets only as needed, make it possible to reduce unnecessary duplication of bits. It relieves some of the network load [1], [2]. Nevertheless, in this context, a single, fixed-rate video stream can not satisfy the conflicting requirements of a heterogeneous set of receivers. One approach for delivering multiple resolution, frame rate, or levels of quality across multiple network connections is to encode the video signal with a set of independent encoders, each producing a different output rate. This approach, called *simulcast*, does not exploit statistical correlations across subflows. Its compression performance is thus suboptimal. By contrast, a layered, or scalable, coder exploits correlations across subflows to achieve better overall compression [3]–[5]. The input signal is compressed into a number of discrete layers, arranged in hierarchically to provide progressive refinement. With *scalable* coding algorithms, one original compressed video bitstream is generated; but different subsets of the bitstream can then be selected at the decoder end to support a multitude of display specifications such as the quality level (SNR scalability), the spatial resolution (spatial scalability), and the frame rate (temporal scalability).

Developing (spatially) scalable video compression algorithms has attracted considerable attention in recent year [6]–[11]. In this paper, we focus our attention to the study of SNR scalable schemes based on the extension of the hybrid motion prediction/matching pursuits coding algorithm [12]. While not inherently suited for scalability, hybrid schemes provide the low delay property, a major feature for communication applications.

Section II discusses the problem encountered when coping with SNR scalability. In Section III, a method is proposed to predict the high SNR layer DFD from the low SNR layer reconstructed DFD. This method is specific to the use of the Matching Pursuits (MP) representation technique and has been implemented to minimize the size of either the low SNR layer bitstream, or the complete scalable bitstream. In Section IV, it is explained how the MP framework permits to quantify the relevance of performing prediction between the DFD of each layer. For the sake of comparison, other prediction schemes have been considered in Section V. They include frame-based or macroblock-based DFD prediction modes but also propose others. Results and discussions are provided in Section VI. Section VII concludes.

II. SNR Scalability: Terms of the Problem

As noted before, SNR scalability allows for the decoding of appropriate subsets of a single output bitstream to generate gradual quality approximations of the original sequence. In the following, studies are restricted to a two-layer system. The generalization to multilevel decompositions is straightforward. The *low SNR layer* codes the information required for low quality delivery. The *high SNR layer* provides the additional information required to display a high quality sequence. We refer to the frames generated using both the low and high SNR layers information as to high SNR layer frames.

Given a quality constraint for each layer, the designer of the video scalable coder has to choose between two distinct objectives, depending on the application (s)he deals with.

1) *Low quality delivery with minimal bitrate:* the purpose is to minimize the size of the bitstream subset providing the low quality video display.

Manuscript received April 7, 1999; revised September 28, 2000. The research of C. De Vleeschouwer was supported by the Belgian NFS. The associate editor coordinating the review of this paper and approving it for publications was Dr. Hong-Yuan Mark Liao.

The authors are with the Laboratoire de Telecommunications et Teledetection, Universite Catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium (e-mail: devlees@tele.ucl.ac.be; macq@tele.ucl.ac.be).

Publisher Item Identifier S 1520-9210(00)11072-7.

2) *High quality delivery with minimal bitrate:* the purpose is to minimize the total size of the bitstream, i.e., the size required to achieve high visual quality display.

Obviously, an equivalent choice appears when the constraint on each layer is expressed in terms of rate rather than quality. The first choice consists in providing maximal quality to the receivers having a small bandwidth and accessing only to the low SNR layer. The second choice is to maximize the quality offered to the receivers accessing to the complete bitstream.

The choice depends on the application. For example, for video transmission over a time varying channel such as the Internet, if, on the one hand, the available bandwidth is at full capacity most of the time, it is worthwhile to keep the best possible quality on the high SNR layer. We will see that it can be done at the expense of the low SNR layer. On the other hand, in applications where the channel is expected to spend most of the time at the lower bandwidth, one should take care to the low SNR layer quality. Another example is for the multicast video transmission. Users with different rate or computational capabilities access either the low or high SNR layer. Depending on its goal (wide distribution of advertising, delivering of a high quality service to the users who pay while providing a minimal service to others, ...), the source may desire to favor the high or the low SNR layer at the expense of the other.

In the next sections, different strategies are investigated to enable SNR scalability for matching pursuits (MP) video coding.

III. Atom-Based Prediction Using Matching Pursuits

The Matching Pursuits video coding algorithm is a hybrid motion-compensated algorithm for which the displaced frame difference (DFD) is expanded into waveforms, called *atoms*, chosen among an overcomplete dictionary (see Appendix A and [12], [13]). Stating that a motion-compensation loop is used in each layer, the aim of this section is to study and exploit the statistical dependencies between the prediction errors or displaced frame difference (DFD) of both layers to design an efficient scalable codec. The ability of Matching Pursuits to perform a signal analysis when decomposing it into waveforms is used to predict the main structures of the high SNR layer DFD from the low SNR layer reconstructed DFD.

A. Layers Prediction Errors (DFD)

Fig. 1 presents the components of each layer of the scalable scheme. $I_o(t-1)$ and $I_o(t)$ refers to the original picture at time $t-1$ and t respectively. Given any *single* set of motion vectors $V(t)$ and the linear motion-compensation operator $\Gamma(., V(t))$, the motion-compensation provides $P_o(t) = \Gamma(I_o(t-1), V(t))$. We can observe that the displaced frame difference $\mathrm{DFD}_o(t) = I_o(t) - P_o(t)$ is not strictly null due to the fact that motion compensation is not able to perform perfect prediction (nonlinearity of the motion, occlusions, new objects, ...). Actually, this frame difference is not computed in a predictive scheme because, for convergence purpose, it is necessary to compute the difference between the original frame and the previously reconstructed compensated frame. $I_l(t-1)$ and $I_h(t-1)$ being the low and high SNR reconstructed frames at time $t-1$, the motion-compensation provides a prediction at time t for the low and high SNR layers, denoted $P_l(t)$ and $P_h(t)$ respectively. The prediction error (DFD) is then computed for each layer:

$$\begin{aligned}\mathrm{DFD}_l(t) &= I_o(t) - P_l(t)\\ &= I_o(t) - \Gamma(I_l(t-1), V(t))\\ \mathrm{DFD}_h(t) &= I_o(t) - P_h(t)\\ &= I_o(t) - \Gamma(I_h(t-1), V(t)).\end{aligned} \tag{1}$$

Denoting $E_l(t-1) = I_o(t-1) - I_l(t-1)$ and $E_h(t-1) = I_o(t-1) - I_h(t-1)$ the reconstruction (coding) error of the low and high SNR layers at time $t-1$, we obtain that

$$\begin{aligned}\mathrm{DFD}_l(t) &= I_o(t) - \Gamma(I_l(t-1), V(t))\\ &= I_o(t) - \Gamma(I_o(t-1) - E_l(t-1), V(t))\\ &= \mathrm{DFD}_o(t) + \Gamma(E_l(t-1), V(t))\end{aligned} \tag{2}$$

$$\begin{aligned}\mathrm{DFD}_h(t) &= I_o(t) - \Gamma(I_h(t-1), V(t))\\ &= I_o(t) - \Gamma(I_o(t-1) - E_h(t-1), V(t))\\ &= \mathrm{DFD}_o(t) + \Gamma(E_h(t-1), V(t)).\end{aligned} \tag{3}$$

So, linearity of the motion-compensation operator allows expressing both low and high SNR prediction errors as the sum of two terms. The first one, $\mathrm{DFD}_o(t)$, is common for both layers. It is due to the erroneous prediction obtained when applying motion compensation to the original picture. The second one results from the introduction of the residual coding error within the prediction loop. This term differs in both layers.

The presence of a common term in the prediction errors of both layers allows the reconstructed (encoded/decoded) low SNR layer DFD to predict the high SNR layer prediction error. Matching pursuits, by their ability to analyze locally the signal they represent, allow selecting predictable structures among both errors. Two selection strategies are proposed in Sections III-B and III-C. In each section, a single motion vectors set is used to compensate both layers. In Section III-B, motion vectors $V_l(t)$ are estimated on the low SNR layer, while in Section III-C motion vectors $V_h(t)$ are estimated on the high SNR layer.

B. Low SNR Layer Delivery with Minimal Bitrate

Delivering the low quality with a minimum bitrate means that the nonscalable video coder generates the low SNR layer bitstream. In order to provide scalability, additional information must complete it in order to achieve a better quality when the whole information is accessed.

According to Fig. 2, given the *motion vectors set $V_l(t)$ used to predict the low SNR layer* at time t, one may compute the high SNR layer prediction $P_h(t) = \Gamma(I_h(t-1), V_l(t))$. The resulting prediction error, i.e., $\mathrm{DFD}_h(t)$, has to be encoded using matching pursuits. Nevertheless, as told in Section III-A, the low and high SNR prediction errors have common structures. It is highly probable that the atoms used to represent the low SNR layer prediction error also match the structures of the high SNR layer DFD.

The high SNR layer DFD representation algorithm is modified as follows. At each step, i.e., each time an atom is selected, two possibilities are considered.

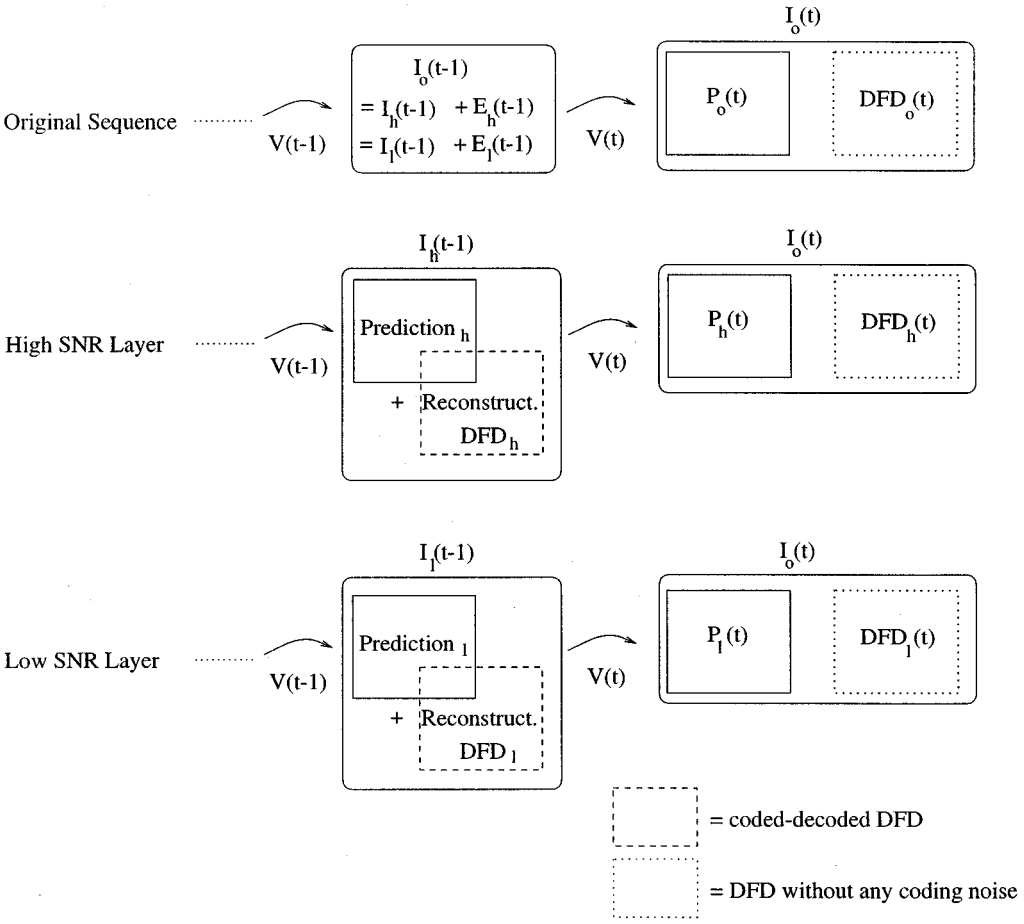

Fig. 1. High and low PSNR layers illustration.

- As in the conventional MP representation technique, the best matching dictionary function is searched for. Note that the search indirectly selects which one of the luminance (Y) or chrominances (U, V) components is represented by the atom. This component is denoted C_{YUV}.
- Among the atoms defining the low SNR layer structures, the best matching function is also searched for. Note that only the atoms describing the C_{YUV} component of the low SNR layer are considered.

In order to decide which one of the two pre-selected atoms is really transmitted, a Rate/Distortion criterion is used. In the first case, the atom is specific to the high SNR layer. Its coding cost can be estimated from the atoms that have been encoded in the previous frame. We denote $COST_{h,\bar{l}}$ the mean coding cost of the high SNR layer atoms that do not belong to the low SNR layer. In the second case, the atom has already been defined in the low SNR layer. Only its amplitude has to be transmitted. In practice, a VLC code transmits it differentially with respect to the amplitude in the low SNR layer. An escape code is used for the atoms of the low SNR layer that are not used in the high SNR layer. The coding cost is estimated by the mean differential coding cost of the atoms of the previous frame. We denote it $COST_{h,l}$. For both atoms, the decrease of distortion achieved by the representation of the atom is the atom energy, i.e., the square of its amplitude, respectively $A_{h,\bar{l}}$ and $A_{h,l}$. In order to maximize the ratio $-\Delta D/\Delta R$, the atom specific to the high SNR layer is selected if:

$$\frac{A_{h,l}^2}{COST_{h,l}} \leq \frac{A_{h,\bar{l}}^2}{COST_{h,\bar{l}}}. \tag{4}$$

The computation overhead of this method is due to the search of the best matching function among the atoms of the low SNR layer. At each step, it requires one additional inner product computations per atom of the low SNR layer. As long as the number of atoms in the low SNR layer is small in comparison with the size of the MP dictionary, the overhead can be neglected. Typically, the dictionary contains 400 functions that can be located in every pixel of a 16 $\times$ 16 search window. That means a size of 400 $\times$ 256 for the MP dictionary. In practice, efficient implementations are possible to compute the inner products with the dictionary functions. So, the overhead becomes significant before the number of atoms in the low SNR layer reaches 400 $\times$ 256. Typically, if the number of atoms is one order of magnitude below, the overhead is significant.

C. High SNR Layer Delivery with Minimal Bitrate

The optimal way to provide the high quality level is to use the nonscalable MP codec for the high SNR layer. This nonscalable codec is used as reference. The aim is to generate a bitstream of minimal size providing the high quality sequence while allowing the extraction of a subset for the reconstruction of a lower quality sequence.

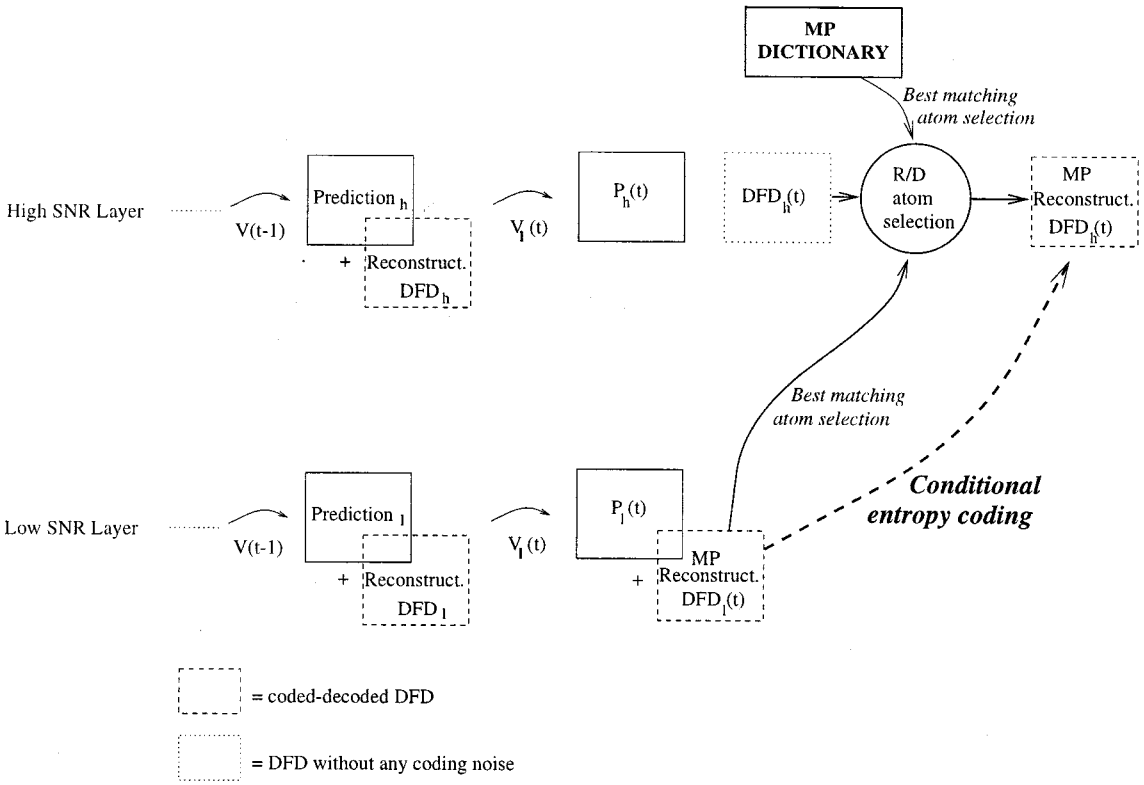

Fig. 2. Atom-based prediction of the high SNR layer DFD from the low SNR layer reconstructed DFD.

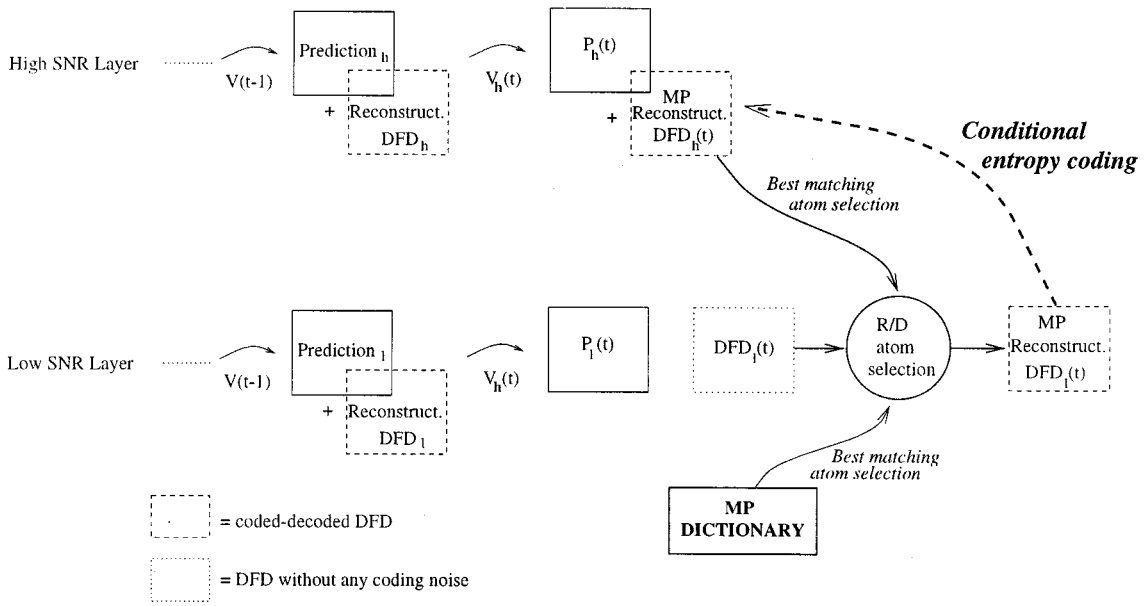

Fig. 3. Construction of the low SNR layer DFD for optimal prediction of the high SNR layer DFD.

Fig. 3 presents the followed strategy. The set of motion vectors is estimated based upon the high SNR layer content. Following the conventional nonscalable scheme, atoms are selected to represent the high SNR layer prediction error, i.e., $\mathrm{DFD}_h(t)$. Once this set has been selected, the cheapest way to provide a lower quality layer is searched for, assuming that prediction is possible between atoms of the low and high SNR layers.

First, *the previous reconstructed low SNR frame is compensated using the motion vectors estimated on the high SNR layer*, i.e., $P_l(t) = \Gamma(I_l(t-1),\, V_h(t))$. The resulting prediction error is represented using either an atom specific to the low SNR layer, or one that has been defined for the high SNR layer. R/D considerations lead to the optimal choice. $COST_{\bar{l}}$ is the mean coding cost of a low SNR layer atom in the previous frame, $COST_{h,\bar{l}}$ is the mean cost of an high SNR layer atom that does not belong to the low SNR layer, and $COST_{h,l}$ is the mean cost of coding the differential amplitude of the atoms that belong to both layers in the previous frame. $A_{l,h}$ is the amplitude, in the

low SNR layer, of the atom pre-selected among the ones of the high SNR layer. $A_{l,\overline{h}}$ is the amplitude of the atom specific to the low SNR layer. In order to maximize the ratio $-\Delta D/\Delta R$, the atom specific to the low SNR layer is selected if

$$\frac{A_{l,h}^2}{COST_l - COST_{h,\overline{l}} + COST_{h,l}} \leq \frac{A_{l,\overline{h}}^2}{COST_l}. \quad (5)$$

Once enough atoms have been selected in the low SNR layer to achieve the required quality, the vectors $V_h(t)$ and the atoms of the low SNR layer are encoded and the low SNR layer bitstream is generated. To provide the high quality bitstream, it is completed by the VLC codes defining the differential amplitude of the atoms present in both layers (conditional entropy coding) and the features of the atoms specific to the high SNR layer.

Note that the same comment than the one in Section III-B can be made about the additional complexity of the method.

IV. DFD Prediction Analysis Using Matching Pursuits

In this section, we show how the approaches proposed in Sections III-B and III-C allow analyzing the correlation between the high SNR and the low SNR DFD layers. More specifically, DFD prediction between layers raises some questions: which part of the low SNR layer reconstructed DFD information can really be exploited to predict the high SNR DFD? If some information appears to be specific to the low SNR layer, what is its importance?

The atoms chosen by the matching pursuits representation of a signal identify the coherent structures of that signal [14]. This feature allows quantifying the answer to the above questions. Indeed, after MP representation of the DFD signals, it is easy to count the number of atoms that are specific to the low SNR layer. It is also easy to measure the cost of the differential coding of the amplitudes of the atoms that match the error of both layers.

A deeper analysis is relevant in the case of high quality delivery with minimum bitrate (Section III-C only). The reconstructed high SNR layer DFD is the one built in the nonscalable scheme. So, for the high SNR layer, information extracted from the scalable bitstream is identical to the information contained in the nonscalable one. The additional bit-budget for the scalable case is due to

- a decrease of entropy coding efficiency due to the partition of the set of atoms. In [13], it has been shown that the efficiency of atom position coding improves as the number of atoms increases. Here, the cloud of atoms representing the high SNR layer DFD is split into a set of atoms encoded in the low SNR layer and a set transmitted in the high SNR layer. It deteriorates coding efficiency;
- the presence of information specific to the low SNR layer. This information is of course not transmitted in the nonscalable scheme.

Formally, let X_h and X_l be the stochastic variables associated to the atom selection process of respectively the high and low SNR layer. For a given frame, their realizations x_h and x_l define the atom set parameters (quantized amplitude, shape and position). From information theory, the following entropy (in) equalities hold:

$$H(X_h) \leq H(X_h, X_l) = H(X_h|X_l) + H(X_l). \quad (6)$$

The inequality becomes equality if X_l is a subset of X_h.

Considering the coding cost C associated to the entropy coding of the realizations x_l and x_h, we observe

$$C(x_h) \leq C(x_h, x_l) \leq C(x_h|x_l) + C(x_l). \quad (7)$$

$C(x_h)$ is the coding cost of the high SNR layer atoms parameters, i.e., the coding cost resulting from the nonscalable scheme. $C(x_l, x_h)$ is the cost that would result from encoding both sets of atoms in a single layer. $C(x_h|x_l) + C(x_l)$ is the scalable scheme coding cost. Due to the entropy coding efficiency deterioration mentioned in, this cost is higher than $C(x_l, x_h)$.

Differentiating four sets of atoms, we identify four stochastic variables associated to the atom parameters definition:

- $X_{l,\overline{h}}$, defining the atoms of the low SNR layer that are not present in the high SNR layer;
- $X_{l,h}$, defining the atoms of the low SNR layer that are present in the high SNR layer;
- $X_{h,l}$, defining the atoms of the high SNR layer that are also present in the low SNR layer;
- $X_{h,\overline{l}}$, defining the atoms of the high SNR layer atoms that are not present in the low SNR layer.

The entropy equality of (6) is decomposed into

$$\begin{aligned} H(X_h, X_l) &= H(X_{h,\overline{l}}, X_{h,l}, X_{l,h}, X_{l,\overline{h}}) \\ &= H(X_{h,\overline{l}}|X_{h,l}, X_{l,h}, X_{l,\overline{h}}) \\ &\quad + H(X_{h,l}|X_{l,h}, X_{l,\overline{h}}) \\ &\quad + H(X_{l,h}|X_{l,\overline{h}}) + H(X_{l,\overline{h}}). \end{aligned} \quad (8)$$

Similarly, we can differentiate four coding costs.

- $C_{l,\overline{h}}$, for the low SNR layer atoms that do not belong to the high SNR layer;
- $C_{l,h}$, for the low SNR layer atoms that are present in the high SNR layer;
- $C_{h,l}$, for the high SNR layer atoms whose shape and position have been defined in the low SNR layer. This cost is equal to the differential amplitude coding cost;
- $C_{h,\overline{l}}$, for the parameters of the high SNR layer atoms that have not been selected to construct the low SNR layer.

Practically, each coding cost is measured by the size of the ad hoc bitstream subset. As parameters of the low SNR layer atoms are encoded as a whole, disregarding whether they are specific to the low SNR layer or not, the coding cost for each category is not available. Each cost is estimated by the corresponding fraction of the total low SNR layer cost. In Section VI-B, we observe and discuss the following inequality:

$$\begin{aligned} C(x_h) &\leq C(x_h, x_l) \leq C(x_h|x_l) + C(x_l) \\ &= C_{h,\overline{l}} + C_{h,l} + C_{l,h} + C_{l,\overline{h}}. \end{aligned} \quad (9)$$

It explains the enlargement of the bitstream in the scalable scheme. The cost of the information specific to the scalable scheme is estimated by $C_{h,l} + C_{l,\overline{h}}$. The remaining excess with regards to $C(x_h)$ is mainly due to the decrease of the source

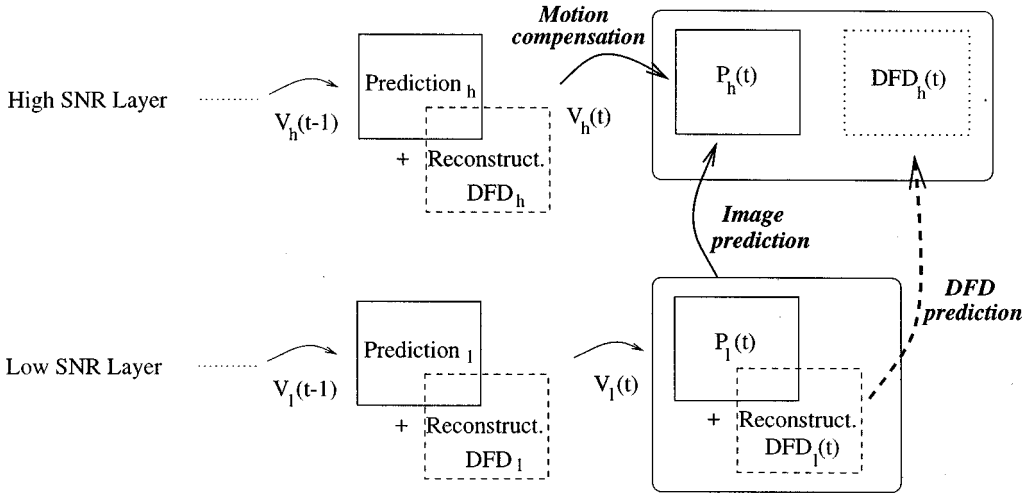

Fig. 4. Multimodal macroblock-based prediction of the high SNR layer. 3 modes are considered: (a) MB is motion-compensated, (b) MB is predicted from the low SNR layer, or (c) MB is motion-compensated and the DFD is predicted from the low SNR layer reconstructed DFD.

coding efficiency. In this reasoning, the assumption is made that, for atoms belonging to both layers, the high SNR layer amplitude coding cost would be close to the amplitude coding cost in the low SNR layer.

V. SNR Scalability: Block- or Frame-Based Prediction

For the sake of comparison and before presenting the results achieved with the atom-based prediction scheme, three conventional scalable schemes are considered. It is worth noting that, on the contrary of the atom-based prediction schemes, these three schemes are not specific to the use of the Matching Pursuits prediction error representation. They can be generalized to any hybrid motion-compensated scheme, whatever the DFD representation technique is. Some of the conclusions drawn from the simulations can thus be generalized to hybrid coders involving other DFD representation techniques.

A. MPEG-4 Scheme for Scalability Using DCT

The first scheme is the versatile MPEG-4 scalable scheme. MPEG-4 is a DCT hybrid motion-compensated scheme. The low SNR layer is encoded as in the nonscalable encoding scheme. For the high SNR layer, bidirectional (spatio-temporal) prediction is used. The two references are the previous high SNR layer frame (with motion-compensation) and the current low SNR layer frame. Fig. 4 presents the outline of the scheme. Solid lines represent the two prediction modes considered. The dashed arrow is not relevant here.

Selection of the prediction mode is macroblock-based [15]. When motion prediction is selected, a high SNR layer motion vector is encoded. *Two sets of motion vectors* (high and low SNR layer) are thus defined in this scheme.

B. MacroBlock-Based Prediction Using Matching Pursuits

The second scheme is also a macroblock-based prediction scheme. Matching Pursuits is used to represent the DFD. Fig. 4 presents the outline of the scheme. Both solid and dashed arrows are relevant in that case. It differentiates itself from the MPEG-4 scheme by the fact that three modes of prediction are considered.

- MB estimation results from the motion compensation of the previous frame of the high SNR layer (MC mode).
- MB prediction is the sum of the motion compensation and of the low SNR layer reconstructed DFD (MC+DFD mode).
- MB is predicted by the current low SNR layer reconstructed image (L-SNR-I mode).

The chosen prediction mode is the one that minimizes the prediction error energy, i.e., the sum of its square values. The prediction mode along with the MB atom flag, i.e., a flag indicating whether the MB contains atom(s) or not, are encoded using a VLC table. The usefulness of a three-modes prediction strategy toward a conventional bimodal strategy (like the one used in the MPEG-4 scalable coder) is demonstrated in Section VI-A-1. Moreover, in Section VI, either a single set of motion vectors, i.e., $V_l(t) = V_h(t)$, estimated on the low or on the high SNR layer, or two sets of motion vectors are used for motion compensation. It allows discussing the usefulness of two sets of motion vectors (MV).

C. Frame-Based Prediction of the DFD

The third scheme is the one proposed by UC Berkeley [13]. In this method, a single set of motion vectors, estimated on the low SNR layer, was used to compensate both layers. According to Fig. 1, both layers are motion compensated and the resulting low and high SNR prediction errors are mentioned as DFD_l and DFD_h. To generate the low SNR layer bitstream, a normalized weighted sum of both prediction errors is encoded, i.e., $(1-\gamma).\mathrm{DFD}_l + \gamma.\mathrm{DFD}_h$, $0 \leq \gamma \leq 1$. The reconstructed weighted prediction error is added to the low SNR layer compensated frame to build the low SNR layer reconstructed image. It is also added to the high SNR predicted frame to generate a new reference for that layer. That reference is then improved by additional atoms, which are specific to the high SNR layer. In [13], γ parameter is used to adjust the quality between both layers while keeping their relative bit rates fixed. Nevertheless,

it appears that increasing the γ value only allow for a small improvement of the high SNR layer (between 0.1 and 0.5 dB) in return for a large quality degradation of the low SNR layer (between 2 and 3 dB). These poor performances are confirmed by the results presented in Section VI.

In Section VI, targeting each one of the two objectives formulated in Section II, the scheme proposed in [13] has been extended and the results are provided for a set of motion vectors estimated either on the low or high SNR layer. With $\gamma = 0$ and the motion vectors estimated on the low SNR layer, low SNR layer quality is delivered with a minimal bit-budget. Estimating motion vectors on the high SNR layer and increasing γ value allows reducing the size of the complete scalable bitstream, but of course makes the access to the minimal quality service more expensive. Experimental results learn us that estimating the set of motion vectors from the high SNR layer rather than from the low SNR one allows for a larger decrease of the complete bitstream size than the one allowed by increasing the γ value.

VI. Results

The results have been generated using a set of four video sequences selected among the ones recommended in the MPEG-4 video coding group for the test of VLBR coding algorithm. For the sake of comparison, common coding conditions (see Table I) have been fixed. For each sequence, PSNR quality levels (in dB) have been settled for both the low and the high SNR layers. Matching Pursuits allow respecting this constraint stringently as, for each frame, atoms can be added to the reconstructed prediction error until the required quality level has been achieved. Comparison between the considered coding schemes is based on the size of the bitstreams.

A. Comparison of the Considered Scalable Schemes

1) Low Quality Delivery with Minimal Bitstream: As the low SNR layer quality has to be provided with a minimal bit-budget, it is encoded using the nonscalable scheme. In Table II, the bit-rate (in bits/s) is presented for five schemes. In the simulcast scheme, both layers are encoded independently. It is the nonscalable scheme proposed as a reference, for comparison purposes. The atom-based prediction scheme refers to the one presented in Section III-B, while the block-based and frame-based schemes refer to Section V-B and V-C. Results obtained when using the MPEG-4 DCT-VM are also presented. In this case, the quantization parameter has been adjusted to provide a mean SNR quality that is just below the quality required in Table I. For each scheme, the bit-rates for transmitting only the low SNR layer, only the high SNR layer and both layers together are given. Of course, for scalable schemes, the bitstream providing the high SNR layer also provides the low SNR layer. On the contrary, for the simulcast scheme, providing both the low and the high SNR layers requires both bitstreams' transmission.

An obvious conclusion that can be drawn from Table II is the superiority of the MP schemes vis-a-vis the DCT MPEG-4 verification model. Another conclusion is that all scalable schemes considered manifest very similar behavior-patterns. For the chosen test conditions, they allow saving about 20% of the total (low + high SNR layers) bit-budget used when providing the high and low quality level separately. It is obtained in exchange for an increase (15–25%) in the necessary bit-budget for providing the high SNR layer only.

TABLE I
Summary of the Coding Conditions for Each Considered Sequence

	Frame resolution	Frame rate	Base layer PSNR quality	Enh. layer PSNR quality
Hall	QCIF	7.5 f/s	31 dB	34 dB
Silent	QCIF	10 f/s	31 dB	34 dB
Foreman	QCIF	10 f/s	31 dB	33 dB
Coastguard	QCIF	10 f/s	30 dB	32 dB

Among the considered MP scalable schemes, the frame-based prediction is the most favorable from a computational point of view. Moreover, its performances often equal or outperform those of the other schemes. This prediction scheme can be recommended when trying to improve the quality of a nonscalable low SNR layer.

Nevertheless, for some sequences, the macroblock(MB)-based prediction scheme performs better than the frame-based one. The frame-based scheme systematically predicts the high SNR layer DFD. On the contrary, the MB-based scheme allows either no prediction at all, or DFD or image MB prediction from the low SNR layer. These modes of prediction are useful for MB's for which prediction errors differ in both layers. Referring to the discussion in Section III-A, these MB's are the ones for which the difference between compensation errors of both layers is significant compared with the original image motion compensation error. It occurs when the translational motion model fits the motion of objects that are encoded with different qualities in both layers. It is the case for the "Coastguard" sequence. The more the motion compensation is efficient, the less the DFD prediction is useful.

The usefulness of three prediction modes is emphasized by Table III. Using three modes always outperforms a bidirectional prediction. Yet, when only two modes are used, prediction of the high SNR layer DFD is preferable to the prediction of the high SNR layer frame. For the "Coastguard" sequence, both modes are nearly equivalent.

Table III also presents the results that are obtained when using two distinct sets of motion vectors for the low and the high SNR layers. Low SNR layer Intra MB's are forced to remain Intra in the high SNR layer. MB's that are Inter in the low SNR layer may either be predicted from the low SNR layer or motion-compensated using a motion vector specific to the high SNR layer. For motion-compensated macroblocks, the low SNR layer reconstructed DFD can either or not predict the high SNR layer DFD. VLC's are used to encode the chosen prediction mode. Due to the use of a set of optimal motion vectors to predict the high SNR layer, the motion prediction error of the high SNR layer has less energy. Nevertheless, it also suffers from a smaller correlation with the low SNR layer DFD. The efficiency of the

TABLE II
LOW SNR LAYER, HIGH SNR LAYER, AND BOTH BIT-RATES LAYERS (BITS/S) FOR FIVE SCHEMES. THE LOW SNR LAYER IS ENCODED AS IN THE NONSCALABLE SCHEME. THE SET OF MOTION VECTORS IS ESTIMATED ON THE LOW SNR LAYER

Sequence	Layer	Simulcast with MP	Atom-based prediction	Block-based prediction	Frame prediction	MPEG-4
Hall	Low SNR	9053	9053	9053	9053	13594
	High SNR	16076	19510	20332	19605	33660
	Low+High	25129	19510	20332	19605	33660
Silent	Low SNR	22178	22178	22178	22178	25218
	High SNR	42419	52311	51397	51673	69667
	Low+High	64597	52311	51397	51673	69667
Foreman	Low SNR	47990	47990	47990	47990	49404
	High SNR	76515	94530	87460	86890	107424
	Low+High	124505	94530	87460	86890	107424
Coastguard	Low SNR	52804	52804	52804	52804	56676
	High SNR	88097	115697	112401	115837	141040
	Low+High	140901	115697	112401	115837	141040

TABLE III
COMPARISON OF THE BIT-RATES (BITS/S) REQUIRED WHEN CONSIDERING THREE OR TWO MODES OF PREDICTION IN A BLOCK-BASED PREDICTION SCHEME. THE LOW SNR LAYER IS ENCODED AS IN THE NONSCALABLE SCHEME

Sequence	Layer SNR	MacroBlock-based prediction schemes				
		Base layer motion vectors set			Two motion vectors sets	
		3 modes	MC, L-SNR-I	MC, MC+DFD	3 modes	MC, L-SNR-I
Hall	Low	9053	9053	9053	9053	9053
	High	20332	22586	20453	21058	21255
	Low + High	20332	22586	20453	21058	21255
Silent	Low	22178	22178	22178	22178	22178
	High	51397	53984	52384	50451	50739
	Low + High	51397	53984	52384	50451	50739
Foreman	Low	47990	47990	47990	47990	47990
	High	87460	92853	89004	88686	92253
	Low + High	87460	92853	89004	88686	92253
Coastguard	Low	52804	52804	52804	52804	52804
	High	112401	118123	117441	112686	118019
	Low + High	112401	118123	117441	112686	118019

prediction between the layers decreases. As expected, we observe that the number of blocks predicted by the MC + DFD mode decreases when two sets of motion vectors are used (see Table IV). From Table III, it appears that the use of two sets of motion vectors is not useful. The additional motion vectors coding cost and the loss of correlation between the DFD structures of both layers outstrip the gain resulting from the improved motion compensation. It is worth noting that this improved motion prediction also requires a significant additional computational load.

2) High Quality Delivery with a Minimal Bitstream: Our aim is to investigate how to reduce the size of the complete scalable bitstream, i.e., how to generate a bitstream of minimal size able to deliver two quality levels. Table V presents the bit-rates (bits/s) obtained for a number of scalable schemes.

For all considered scalable schemes, the set of motion vectors has been estimated on the high SNR layer. It permits a significant reduction of the scalable bitstream size in return for a small increase in the low SNR layer bit-budget. This is observed when comparing the frame-based ($\gamma = 0$) and the macroblock-based prediction schemes of Table V with the ones of Table II.

TABLE IV
COMPARISON OF THE MODES OF PREDICTION IN TWO BLOCK-BASED PREDICTION SCHEMES: (LEFT COLUMN) A SINGLE SET OF MOTION VECTORS HAS BEEN ESTIMATED ON THE LOW SNR LAYER. (RIGHT COLUMN) TWO SETS OF MOTION VECTORS ARE ESTIMATED (ONE FOR EACH LAYER). WE OBSERVE THAT THE NUMBER OF BLOCKS PREDICTED BY THE MC + DFD MODE DECREASES WHEN TWO SETS OF MOTION VECTORS ARE USED

Sequence	Pred. Mode	Number of Macroblocks	
		1 set of MV	2 sets of MV
Hall	L-SNR-I	276	341
	MC	6596	6622
	MC+DFD	454	363
Silent	L-SNR-I	959	909
	MC	7926	8029
	MC+DFD	916	863
Foreman	L-SNR-I	1905	1944
	MC	5647	5997
	MC+DFD	2249	1860
Coastguard	L-SNR-I	2898	2930
	MC	4184	4432
	MC+DFD	2719	2439

TABLE V
LOW SNR LAYER, HIGH SNR LAYER AND BOTH BIT-RATES LAYERS (BITS/S) FOR ATOM-BASED, MACROBLOCK-BASED AND FRAME-BASED PREDICTION SCHEMES. THE SET OF MOTION VECTORS IS ESTIMATED ON THE HIGH SNR LAYER

Sequence	Layer SNR	Simulcast with MP	Atom-based prediction	Frame-based prediction			MB-based prediction
				γ=0	γ=0.2	γ=1	
Hall	Low	9053	12132	9155	9642	14300	9108
	High	16073	17178	19477	18911	18341	19953
	Low + High	25126	17178	19477	18911	18341	19953
Silent	Low	22178	29331	22888	24186	37711	22584
	High	42418	44454	48376	47856	49914	48419
	Low + High	64596	44454	48376	47856	49914	48419
Foreman	Low	47990	57137	50012	53747	87848	49746
	High	76514	81652	82053	83809	104564	84196
	Low + High	124504	81652	82053	83809	104564	84196
Coastguard	Low	52804	65662	54621	56152	82129	54398
	High	88096	92956	108814	104031	92860	108622
	Low + High	140900	92956	108814	104031	92860	108622

When trying to minimize the size of the complete bitstream, the best result is obtained with the atom-based prediction method presented in Section III-C. The additional bit-budget of the scalable scheme is limited to 5–7%. This figure appears from Table V. It comes from the comparison between the total (high + low) cost of the atom-based prediction scheme and the high SNR cost of the simulcast scheme. These 5–7% are a significant reduction with regards to the 20% additional budget obtained when low quality was delivered with a minimal bit-budget (see Table II in Section VI-A1). This is also a significant decrease (3–12%) compared to the overhead of all other schemes considered in Table V. For the frame-based prediction scheme, a set of γ parameters has been considered. Most often, only small decreases in the complete scalable bitstream are obtained in return for large increases in the low SNR layer bit-budget. The best compromises are obtained with $\gamma = 0.2$. To conclude, we also note from Table V that the reduction of the total size of the scalable scheme is always obtained in exchange for a significant increase of the low SNR layer bit-rate. It confirms that the choice of the scalable prediction scheme should be application-driven. Achieving a minimal size for the low SNR layer and for the total stream are conflicting goals.

3) Improvement: Tradeoff for Atom Selection: In Section VI-A1 and VI-A2, a set of atoms is first selected using the conventional MP algorithm. According to the goal, the selection is performed either on the low or on the high SNR layer. But it never takes the other layer into account. Obviously, this initial selection constraints the final results. To improve the proposed method, we should try to get rid of that constraint. A way to proceed could be to compute the inner products of a candidate "initial" atom with the DFD of each layer. We denote the inner products A_l and A_h respectively. To select the atom, an increasing function of both A_l and A_h should be maximize. A candidate function f is:

$$f = \alpha.A_l^2 + \beta.A_h^2. \tag{10}$$

Note that in Section VI-A1, $\beta = 0$, while in Section VI-A1, $\alpha = 0$. When α and β are both non zero, they can be fixed or adaptive. They can, for example, be adapted dynamically as a function of the unbalance between the reconstruction errors in both layers. If the error in one layer remains two high with regards to the other, α and β parameters are modified to favor the choice of an atom that matches that layer. Note also that once the quality constraint is reached for one layer, the R/D optimal selection method proposed in Section VI-A1 or Section VI-A2 is applied.

B. Prediction Errors Analysis

Table VI compares the costs (bits/s) of the atom-based prediction scalable scheme described in Section III-C, and the nonscalable scheme. Motion and intra-information are common to both schemes. In the nonscalable scheme a single bitstream is generated. In the scalable scheme, bits are allocated between the layers. For the sake of analysis, the scalable cost has been partitioned into two parts (see Section IV). The first includes the coding cost of the atoms that are used to reconstruct the high SNR DFD. As told in Section III-C, these atoms are identical to the ones selected in the nonscalable scheme, but they are spread in both layers. A coding efficiency decrease results from the splitting. It is measured by the difference between the third and fourth columns of Table VI. The fifth and last column of the table presents the coding cost of the atoms that are specific to the low SNR layer and the coding cost of the differential amplitude of the atoms belonging to both layers. The sum of these two costs is due to the specific information that has to be transmitted in order to provide scalability. From the table, it appears that both terms of that sum contribute equally (3–4% each) to the scalable encoder additional bit-budget (7–8%). Improving either the prediction strategy or the entropy coding method should thus not permit a significant decrease in the size of the complete scalable bitstream.

Considering now both atom-based prediction schemes (Section III-B and III-C), it can be noticed, from Table VII, that most of the low SNR layer atoms are used in the high SNR layer. The

TABLE VI
COMPARISON OF COSTS (BITS/S) BETWEEN THE NONSCALABLE SCHEME AND THE ATOM-BASED PREDICTION SCHEME, TARGETING A MINIMAL SIZE OF THE TOTAL BITSTREAM. MOTION- AND INTRA-INFORMATION ARE COMMON TO BOTH SCHEMES. IN THE NONSCALABLE SCHEME A SINGLE BITSTREAM IS GENERATED. IN THE SCALABLE SCHEME, BITS ARE ALLOCATED BETWEEN THE LAYERS

Sequence	Motion + Intra	Non-scalable high SNR layer	Scalable cost $C_{h,l}+C_{l,h}$	Scalable cost $C_{l,\bar{h}}+C_{h,l}$
Hall	1559+1842	12672	13366	379
Silent	7821+2296	32301	33260	810
Foreman	16451+1533	58530	59944	3406
Coastguard	8826+1625	77645	79236	3267

TABLE VII
ATOM-BASED SCALABLE PREDICTION SCHEMES: LOW QUALITY DELIVERY (LQ OPTI.) AND HIGH QUALITY DELIVERY (HQ OPTI.) WITH A MINIMAL SIZE BITSTREAM. NUMBER OF ATOMS ALLOCATED TO EACH LAYER

Sequence		Number of atoms		
		Low and High	Low specific	High specific
Hall	LQ opti.	2225	72	4120
	HQ opti.	3555	1	1886
Silent	LQ opti.	4433	143	12817
	HQ opti.	6979	0	6909
Foreman	LQ opti.	11602	166	20770
	HQ opti.	14892	533	10381
Coastguard	LQ opti.	18635	184	28878
	HQ opti.	25318	10	11261

number of atoms that are specific to the low SNR layer is small. Nevertheless, allocation of atoms between layers is quite different in the two cases considered. When trying to minimize the complete scalable bitstream, many more atoms are required in the low SNR layer to achieve the quality constraint. Actually, these atoms are selected among the high SNR layer ones and do not necessarily represent the same low SNR layer structures as the ones that would have been represented using the nonscalable scheme for coding the low SNR layer. Yet, the visual quality of the reconstructed low SNR sequence is very similar in both cases. The number of atoms specific to the low SNR layer is often larger when this layer is encoded using the nonscalable scheme. These atoms are used to represent the structures due to the compensation of the previous frame residual errors, which are different in both layers.

VII. CONCLUSIONS

A number of SNR scalable schemes based on the hybrid motion-compensated Matching Pursuits video coding algorithm have been considered. Given an SNR quality constraint for each layer, two distinct objectives have been highlighted:

- generation of a scalable bitstream delivering the low quality (low SNR layer) with a minimal bit-budget;
- generation a scalable bitstream with a minimal total bit-budget.

It appeared that these are conflicting schemes. The achievement of these objectives is thus open to compromise. The targeted goal should be application-driven. With regards to the performances, for the first objective, all schemes provide similar results. Computational complexity or implementation considerations should guide the choice of the scalable codec designer. On the contrary, when targeting the second objective, the schemes are not equivalent any more. Based on the matching pursuits DFD representation, the proposed atom prediction scheme allows for a much higher compaction of the scalable bitstream than any other conventional scheme. This compaction is obtained in return for a significant increase of the low SNR layer bit-budget. For conventional prediction schemes, that can be generalized to any DFD representation technique, we learned that the estimation of the set of motion vectors on the high SNR layer brings a better compaction of the total bitstream than the modification of the low SNR layer DFD coding strategy proposed in [13].

Exploiting the matching pursuits intrinsic analysis abilities, the DFD predictability between layers has been discussed. The scalable scheme presents an additional bit-budget in comparison with the nonscalable scheme. This additional cost is due to a loss of coding efficiency and to the appearance of information specific to the low SNR layer. The partition of the extra cost between these causes has been discussed. We learned that both causes have a similar impact on the bitstream size.

Another lesson from this work is that the use of two sets of motion vectors is not useful. The additional motion vectors coding cost and the loss of correlation between the DFD structures of both layers outstrip the gain resulting from the improved motion prediction.

Eventually, the efficiency of matching pursuits SNR scalable schemes has been demonstrated in comparison with the MPEG-4 generalized DCT scalable scheme.

APPENDIX
MATCHING PURSUITS VIDEO CODING

A. Basic Principles of Matching Pursuits

In this section, in order to simplify notation, we consider the expansion of a one-dimensional (1-D) signal. The extension of the results to the two-dimensional (2-D) DFD signal is immediate.

Given a large and redundant dictionary $\mathcal{D}=(g_\gamma)_{\gamma\in\Gamma}$ of vectors in $L^2(R)$, such that $||g_\gamma||=1$, matching pursuits [14] perform an adaptive expansion of any vectors f in $L^2(R)$ over a set of waveforms selected from $\mathcal{D}$, in order to best match its structures. This is done by the successive approximations of f through orthogonal projections on elements of $\mathcal{D}$. Let $g_{\gamma_0}\in\mathcal{D}$. The vector f can be decomposed into

$$f \leq f, g_{\gamma_0} > g_{\gamma_0} + Rf \tag{11}$$

where Rf is the residual vector after approximating f in the direction of g_{γ_0}. Clearly, g_{γ_0} is orthogonal to Rf, hence

$$||f||^2 = | < f, g_{\gamma_0} > |^2 + ||Rf||^2. \tag{12}$$

To minimize $||Rf||$, we must choose g_{γ_0} such that $|\langle f, g_{\gamma_0}\rangle|$ is maximum. This is the first step of the approximation proce-

dure. It is repeated iteratively on the obtained residue. So, after n steps, the nth order residue $R^n f$ is decomposed into

$$R^n f \leq R^n f, g_{\gamma_n} > g_{\gamma_n} + R^{n+1} f \tag{13}$$

with g_{γ_n} chosen to match $R^n f$, i.e., to maximize $|\langle R^n f, g_{\gamma_n}\rangle|$. If the decomposition is carried through to order m, f is decomposed into a sum of m *atoms* $\langle R^n f, g_{\gamma_n}\rangle g_{\gamma_n}$ and of the mth order residue $R^m f$, i.e.,

$$f = \sum_{n=0}^{m-1} \langle R^n f, g_{\gamma_n}\rangle g_{\gamma_n} + R^m f. \tag{14}$$

The energy conservation (12) yields

$$\|f\|^2 = \sum_{n=0}^{m-1} |\langle R^n f, g_{\gamma_n}\rangle|^2 + \|R^m f\|^2. \tag{15}$$

Convergence of the process is ensured by energy conservation, i.e., the fact that the residue energy is still decreasing.

So, matching pursuits [14] can be viewed as a way of building signal adapted bases. Starting from an overcomplete dictionary, it defines an adaptive time-frequency transform. The signal is expanded into waveforms called *atoms*, whose time-frequency properties are adapted to the signal's local structures. *It is worth noting that MP, together with a signal projection, performs a signal analysis.* MP explicitly selects the information for transmission among a large and overcomplete set of functions. The most significant coefficients are first extracted.

B. DFD Coding using Matching Pursuits

Matching pursuits expansion techniques have been successfully applied in the framework of DFD coding by Neff and Zakhor [12] and Banham [16]. They have thoroughly proven that this technique is competitive toward the DCT-based standard. In their research, a dictionary composed of a set of 2-D Gabor functions has been chosen. Once the set of 2-D functions with finite extent is fixed, the dictionary is extended to the picture domain by allowing the center of each function to be translated into each pixel position.

A direct extension of the MP algorithm requires examining each 2-D dictionary structure at all possible pixel locations in the DFD. As stated by Neff and Zakhor [12], assuming that the DFD is sparse in pockets of energy where motion prediction was inadequate, we can limit the search around these high-energy pockets. Actually, each luminance (Y) or chrominances (U, V) of the DFD is divided into a set of overlapping blocks (12×12 pixels) located at the center of each block of a grid of 8×8 blocks. For each block, the sum of the squares of all pixel intensities is computed, providing a *block energy* value. The inner product search is then performed in a $S \times S$ search window around the center of the block with the largest energy value. The search window selects thus both the (Y, U, V) component and the spatial area in which the signal representation fidelity is improved by the new atom. Once selected, each atom is characterized by its shape (specified by the indices of the chosen function), its position in the picture, the space it belongs to (Y, U or V) and its amplitude. Entropy coding is required to transmit these parameters efficiently (see [12], [13]).

REFERENCES

[1] D. Kosiur, *IP MULTICASTING: The Complete Guide to Interactive Corporate Networks.* New York: Wiley, 1998.

[2] C. Diot, W. Dabbous, and J. Crowcroft, "Multipoint communication: A survey of protocols, functions, and mechanisms," *IEEE J. Select. Areas Commun.*, vol. 15, pp. 277–290, Apr. 1997.

[3] N. F. Maxemchuk, "Video distribution on multicast networks," *IEEE J. Select. Areas Commun.*, vol. 15, pp. 357–372, Apr. 1997.

[4] S. McCanne, M. Vetterli, and V. Jacobson, "Low-complexity video coding for receiver driven layered multicast," *IEEE J. Select. Areas Commun.*, vol. 15, pp. 983–1001, Aug. 1997.

[5] H. Gharavi and M. H. Partovi, "Multilevel video coding and distribution architectures for emerging broadband digital networks," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, pp. 459–469, Oct. 1996.

[6] K. Illgner and F. Muller, "Spatially scalable video compression employing resolution pyramids," *IEEE J. Select. Areas Commun.*, vol. 15, pp. 1688–1703, Dec. 1997.

[7] M. Ghanbari and V. Seferidis, "Efficient H.261-based two-layer video codecs for ATM networks," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 5, pp. 171–175, Apr 1995.

[8] G. Morrison and I. Parke, "A spatially layered hierarchical approach to video coding," *Signal Process.: Image Commun.*, vol. 5, no. 5-6, pp. 445–462, Dec. 1993.

[9] F. Fan, S. Simon, I. Bruyland, W. Zhu, B. De Canne, and M. Van Bladel, "A method for hierarchical subband HDTV splitting," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 5, pp. 225–230, June 1995.

[10] J. Y. Tham, S. Ranganath, and A. A. Kassim, "Highly scalable wavelet based video codec for very low bit-rate environment," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 12–27, Jan. 1998.

[11] D. Taubman and A. Zakhor, "Multirate 3-D subband coding of video," *IEEE Trans. Image Processing*, vol. 3, pp. 572–588, Sept. 1994.

[12] R. Neff and A. Zakhor, "Very low bit rate video coding based on matching pursuits," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, pp. 158–171, Feb. 1997.

[13] O. Al-Shaykh, E. Miloslavsky, T. Nomura, R. Neff, and A. Zakhor, "Video compression using matching pursuits," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, pp. 123–143, Feb. 1999.

[14] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Processing*, vol. 41, pp. 3397–3415, Dec. 1993.

[15] T. Sikora, "The MPEG-4 video standard verification model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, pp. 19–31, Feb. 1997.

[16] M. Banham and J. Brailean, "A selective update approach to matching pursuits video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, pp. 119–129, Feb. 1997.

Christophe De Vleeschouwer was born in Namur, Belgium, in 1972. He received the Ing. and Ph.D. degrees from the Université Catholique de Louvain (UCL), Louvain-la-Neuve, Belgium, in 1995 and 1999, respectively.

From 1999 to 2000, he was with the Interuniversity Microelectronics Centre (IMEC) R&D, Belgium. He is now a Senior Researcher of the Belgian NFS within the Telecommunication Laboratory of UCL. His main interests concern the algorithmic design of multimedia applications, including video, 3-D, and networking technologies.

Benoit Macq (S'83–M'84) is Professor at the Université Catholique de Louvain (UCL), Louvain-la-Neuve, Belgium. His main interests rely on technologies for visual communications. He is involved in many European projects related to video compression, image watermarking and image analysis for mixed realities. He is the Leader of a research group in the Telecommunication Laboratory of UCL. He has been Researcher with Philips Research Laboratory Belgium. Earlier, he was a Visitor Scientist at the Massachusetts Institute of Technology, Cambridge, and a Visiting Professor at ENST Paris, France, and the University of Nice-Sophia-Antipolis, France.