

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization International Bureau



(43) International Publication Date  
13 October 2005 (13.10.2005)

PCT

(10) International Publication Number  
**WO 2005/096228 A1**

(51) International Patent Classification<sup>7</sup>:

**G06T 9/00**

(21) International Application Number:

PCT/EP2005/002072

(22) International Filing Date: 28 February 2005 (28.02.2005)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

04290695.8 12 March 2004 (12.03.2004) EP

(71) Applicant (for all designated States except US): **THOM-SON LICENSING S.A.** [FR/FR]; 46 Quai A. le Gallo, F-92100 Boulogne-Billancourt (FR).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **MARQUANT, Gwenaelle** [FR/FR]; 1 rue du Général de Gaulle, F-35340 Liffre (FR). **VIERON, Jérôme** [FR/FR]; 15 rue Frédéric Cournet, F-35000 Rennes (FR). **BOISSON, Guillaume** [FR/FR]; 12 rue Jean Malo-Renault, F-35000 Rennes (FR).

(74) Agent: **RITTNER, Karsten**; Deutsche Thomson-Brandt GmbH, European Patent Operations, Karl-Wiechert-Allee 74, 30625 Hannover (DE).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

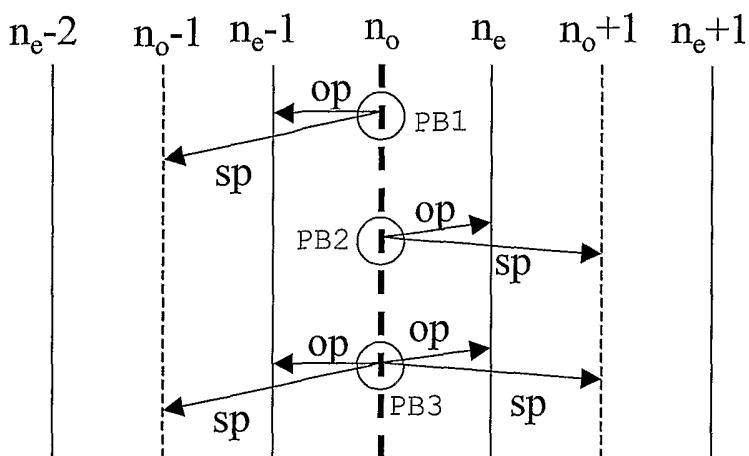
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHOD FOR ENCODING INTERLACED DIGITAL VIDEO DATA



(57) Abstract: Interlaced video can be encoded in two layers, base layer and enhancement layer. A method for optimizing the encoding of motion vectors for enhancement layer is proposed. It comprises defining various different ways to encode the motion vectors, estimate the coding costs for each way, and select the way with the lowest coding cost. The various ways to encode enhancement layer motion vectors can be combined. Either motion vectors from base layer are reused, resulting in a scaling factor and an update vector, or new motion vectors are calculated. Either forward or backward or bi-directional prediction can be used. Either one frame or multiple frames can be taken as reference. The reference frame can be

WO 2005/096228 A1

from the enhancement layer, base layer or both. Either the complete motion vectors are encoded, or only the difference between the motion vector and the corresponding BL motion vector is encoded.

**Method for encoding interlaced digital video data**Field of the invention

5 This invention relates to a method for video compression. More particularly, it relates to a method for encoding data for interlaced video using motion compensated predictive coding.

10 Background

Various video coding algorithms use motion compensated predictive coding for reduction of transmission bandwidth. This is known as Motion Compensated Temporal Filtering (MCTF). In such hybrid schemes, temporal redundancy is reduced using motion compensation, and spatial redundancy is reduced using motion compensation residual transform coding. Motion compensation includes usage of motion estimation (ME) to generate motion vectors (MV), which are 15 used in order to minimize the amount of data to be transmitted when an object moves its position in one picture relative to another picture. The motion vectors are then encoded and transmitted to the decoder, which can then locate the object in a reference picture and move it according to the motion vectors. Motion estimation uses 20 either forward or backward or bi-directional referencing. 25

New display types, e.g. miniature mobile displays, get by with progressive video sequences of low spatial and 30 temporal resolution, while others, e.g. HDTV receivers, require interlaced sequences with high spatial resolution and standard temporal resolution. But while most standard definition (SD) video material is interlaced, recent

research in scalable video coding has concentrated on evaluation of progressive video. A method for transmission of spatially and temporally scalable, interlaced video that provides a wide range of scalability is described in the 5 European Patent Application EP03290507. It is based on separating a video frame sequence into a base layer (BL) and an enhancement layer (EL).

10 Summary of the Invention

Video encoding generally aims at minimizing transmission data rate. For multi-layer approaches for generating spatially and temporally scalable interlaced video, e.g. 15 the described dual layer approach, there may exist further possibilities for redundancy reduction, depending on the actual video sequence. The purpose of the invention is to exploit these possibilities, and thereby improve the efficiency of encoding of interlaced video, and 20 particularly of scalable interlaced video.

A method that is suitable for solving this problem is disclosed in claim 1. An encoder that utilizes the method is disclosed in claim 3.

25

According to the invention, using a flexible scheme for selecting and encoding the EL motion vectors can further reduce the required amount of bits for encoding, the so-called coding costs. This results in an adaptive approach 30 that enables optimization of the data rate with regard to the actual video sequence.

The inventive method for encoding data for interlaced video comprises the steps of separating video frames into a BL and an EL; performing motion estimation for the BL frames, referencing 5 BL frames and resulting in first motion vectors; performing motion estimation for the EL, referencing BL frames, EL frames or combinations thereof and resulting in different sets of second motion vectors; calculating coding costs for the different sets of second 10 motion vectors, the coding costs corresponding or being equal to the number of bits required for encoding; estimating a distortion value for each of the different sets of second motion vectors, the distortion value being based on the difference between the enhancement layer frame 15 and a frame that is predicted based on the current set of second motion vectors and the current reference; selecting the reference and set of second motion vectors that provides the best compromise of low coding costs and low distortion; and 20 transmitting an identifier with the encoded video data stream, the identifier indicating the utilized encoding type. If e.g. for a particular EL frame the optimal encoding is based on forward prediction referencing a single BL frame, an identifier assigned to this combination 25 can be transmitted, so that the decoder is able to recreate the EL frame by applying the motion vectors on the reference frame indicated by the identifier.

The best compromise between low coding costs and low 30 distortion can be found e.g. by assigning weighting factors to the coding cost value and the distortion value, adding the two weighted values and selecting combination with the

minimum sum. The weighting factors can in a trivial case be 1, or can be selected e.g. according to experience.

Further, the inventive method may additionally comprise the 5 step of selecting for the motion vectors of each of the different sets of second motion vectors reference motion vectors belonging to the first motion vectors, and calculating a scaling factor and a difference vector, wherein the scaling factors and difference vectors are used 10 for the calculation of the coding costs and the following steps, and/or for subsequent decoding.

An encoder according to the invention uses motion compensated predictive coding and comprises 15 means for separating video frames into base layer frames and enhancement layer frames; means for performing motion estimation for the base layer frames, referencing base layer frames and resulting in first motion vectors; 20 means for performing motion estimation for the enhancement layer frames, wherein for a single enhancement layer frame different sets of second motion vectors based on different references are generated, the different references being one or more base layer frames, or one or more enhancement 25 layer frames, or combinations thereof; means for calculating coding costs for each of the different sets of second motion vectors, the coding costs corresponding or being equal to the number of bits required for encoding the enhancement layer frame; 30 means for estimating a distortion value for each of the different sets of second motion vectors, the distortion value being based on the difference between the enhancement layer frame and a frame that is predicted based on the

current set of second motion vectors and the current reference;

means for storing the estimated coding costs and distortion values corresponding to the different types of motion

5 estimation, wherein each stored coding cost value and distortion value is associated with an identifier for the corresponding set of second motion vectors;

means for comparing the coding costs for the different sets of second motion vectors, selecting the motion vectors that

10 provide the least coding costs and determining the associated motion estimation; and

means for transmitting an identifier with the encoded video data stream, the identifier indicating the utilized encoding type.

15

A decoder for generating interlaced video data comprises means for receiving an identifier, the identifier being associated to a motion vector or motion field and indicating a reference frame for a region of a prediction

20 frame or a complete prediction frame, and means for decoding a prediction frame or a region within a prediction frame based on said identifier and on said motion vector.

One embodiment of the inventive encoder further comprises

25 means for calculating scaling factors and difference vectors between each of the different sets of second motion vectors and the first motion vectors, wherein the difference vectors are used for the calculation of the coding costs and the following steps, and/or for subsequent decoding.

In one embodiment of the invention, the motion estimation for the enhancement layer is initialized with the first

motion vectors.

In one embodiment of the invention, the motion estimation for the enhancement layer comprises at least two out of the 5 following possibilities: backward prediction, forward prediction and bi-directional prediction.

In one embodiment of the invention the motion estimation for a frame of the enhancement layer references to a single 10 frame, while in another embodiment it references to a plurality of frames.

Further advantageous embodiments of the invention are disclosed in the dependent claims, the following 15 description and the figures.

#### Brief description of the drawings

20 Exemplary embodiments of the invention are described with reference to the accompanying drawings, which show in

Fig.1 backward, forward and bi-directional prediction;

25 Fig.2 multiple reference prediction;

Fig.3 interpolation of EL MV from BL collocated MV;

Fig.4 extrapolation of EL MV from BL collocated MV; and

30

Fig.5 the structure of an encoder for scalable interlaced video.

Detailed description of the invention

Fig.1 shows the principle of a first type of different motion prediction modes that can be combined according to 5 the invention. The motion prediction is based on a number of interlaced frames that usually belong to the same group of frames. The frames are divided into odd and even frames. In this example, odd frames belong to the EL and even frames belong to the BL.

10

A current EL frame  $n_o$  being an odd frame contains several pixel blocks  $PB1, \dots, PB3$ . Motion estimation tries to locate for as many as possible of the pixel blocks a reference in one of the neighbor frames, wherein neighbor frames can 15 have the same parity  $sp$  as the current frame, or opposite parity  $op$ . Therefore there are two neighbor frames in each direction, one in the same layer and one in the other layer. In Fig.1, for the current odd frame  $n_o$  the neighbor frames are: the preceding even frame  $n_{e-1}$  from the BL, the 20 preceding odd frame  $n_{o-1}$  from the EL, the next even frame  $n_e$  from the BL and the next odd frame  $n_{o+1}$  from the EL.

The method according to the invention comprises calculation of different types of motion vectors for the blocks of the 25 current frame, estimating the corresponding coding costs and distortion values and selecting the motion vectors with the best coding cost/distortion compromise for encoding. For the first pixel block  $PB1$  of Fig.1 two possibilities for backward motion estimation are shown, namely either 30 based on the preceding frame of same parity  $sp$  or on the preceding frame of opposite parity  $op$ . Another pixel block  $PB2$  uses forward motion estimation, with the two possibilities of using the next frame of same parity  $sp$  or

the next frame of opposite parity as reference. Yet another pixel block PB3 uses bi-directional motion vectors, with four options for any combination of preceding or succeeding frames of same or opposite parity: op-op, sp-op, op-sp and 5 sp-sp. According to the invention, the coding costs for the described combinations are calculated, the corresponding distortion values estimated, and the combination with the best coding cost/distortion compromise is selected for encoding.

10

Thus, the proposed algorithm makes use of intra/inter layers correlation, i.e. between fields of same or opposite parity. In one embodiment, it is possible to use multiple reference fields to efficiently predict the EL. Fig.2 shows 15 the principle of a second type of different motion prediction modes that can be combined according to the invention. The motion prediction is based on a defined amount of preceding or succeeding frames, wherein it can use more than one frame at a time. This approach is known 20 as multi-referencing. E.g. a fourth region PB4 of the current odd frame  $n_o$  can refer to one or more of the three preceding frames, which are one odd frame  $n_{o-1}$  and two even frames  $n_{e-2}, n_{e-1}$ . Another region PB5 of the current odd frame  $n_o$  may refer to one or more of the three following 25 frames, being also one odd frame  $n_{o+1}$  and two even frames  $n_e, n_{e+1}$ . A further region PB6 of the current odd frame  $n_o$  may refer to one or more of the six neighboring frames, being the three preceding frames and the three following frames mentioned before.

30

Generally, two distinct concepts for encoding of motion vectors can be used, which are called "a priori" and "a posteriori". "A priori" means that the previous base layer

information is exploited: motion vectors are rescaled as initialisation, and a research around this initialisation is performed between surrounding fields. Besides, the coding step benefits from only a residual part of each motion vector has to be transmitted, coupled with the chosen MV, or predictor. "A posteriori" means that the EL motion estimation stage is performed without taking into account the base layer motion information. This latter information may however be used for encoding the EL vectors.

In the "a posteriori" concept, various motion vector prediction modes can be used: backward, forward or bi-directional. The best prediction vector is chosen, either in the same layer field or in the opposite layer field. The choice may be limited to directly contiguous fields, as in Fig.1. This operation mode can also be "multiple reference", i.e. the vector choice is not limited to contiguous neighbours fields, like in Fig.2. In addition to the methods illustrated in Fig.1 and Fig.2, the re-use of motion vector information from the BL makes it easy to encode only the residual part of each motion vector between BL and EL fields. The advantage of such a solution is the reduction of the MV coding cost.

In the "a priori" concept, it is assumed that the BL motion vectors are available. In this case they can be used to efficiently initialise the EL motion estimation process. This is shown in Fig.3. In this example, odd frames  $n_{o-2}$ ,  $n_{o-1}, n_o, n_{o+1}$  are assumed to be used for the base layer. A reference motion vector  $V$  belongs to the BL estimation.  $V$  permits to propose several potential predictors for the EL among all fields, whatever their parity. A particular

predictor  $i$  is given by  $V_i = \beta_i V$ , with  $\beta_i \in Q$ ,  $Q$  being the rational numbers set, i.e.  $V_i$  and  $V$  are collinear. This scaled method can be very similar to the "dual-prime" approach. By coupling this latter method with the "direct mode" it is possible to deduce bi-directional vectors from  $V$ , which can be used to construct motion vectors by adding update vectors. Advantageously, it may be more efficient to encode the factor  $\beta_i$  and the update vector than using another representation of a MV.

10

An advantage of the inventive method is that it enables the usage of potentially long filters in the processing of interlaced video material, which is not known in the art. Filtering more than two frames allows higher efficiency in the re-use of motion vectors. E.g. when an object performs a linear motion over several frames, the same reference vector  $V$  can be used for all these frames. If the motion is not exactly linear, a reference MV can be used for more efficient encoding, as shown in Fig.3. When re-using motion vector information  $V$  from the BL, it is possible to encode only the residual part of each motion vector between BL and EL fields. The advantage is the reduction of the MV coding cost. In Fig.3, predictors are interpolated from the BL collocated MV. E.g. at a specific region of the frame the reference motion vector  $V$  is calculated between two frames  $n_{o-2}, n_{o+1}$  of the BL. For different sub-regions within the specific region, initialisation of a predictor is performed by using the collinear value  $\beta V$  and adding an update vector to adjust the prediction. Thus, the final MV can be defined by the respective  $\beta$  and update vector combination. E.g. for a sub-region PB9 the final motion vector MV3 can be constructed by adding the initial predictor IV and the update vector UV. The performance of the algorithm depends

on the quality of the BL motion information: the better BL motion information is, the more effective is the algorithm.

Fig.4 shows how predictors can be extrapolated from the BL 5 collocated MV. The BL reference vector V2 is in this case calculated between successive BL fields  $n_{o-1}$  and  $n_o$ . It can be used to describe MVs from the EL as a function of a BL MV. Update vectors are used to adjust the prediction.

10 In one embodiment of the invention, it is also possible to consider the complexity of the calculations in order to minimize the required processing power.

Fig.5 shows the structure of an encoder for scalable 15 interlaced video according to the invention. The input signal I is an interlaced video sequence with a temporal resolution of  $2\alpha$  Hz, i.e.  $2\alpha$  fields per second or  $\alpha$  frames per second, and a spatial resolution of  $2M \times 2N$ , i.e. a frame having  $2N$  rows with  $2M$  pixels each, or a field having  $N$  20 rows with  $2M$  pixels each. A splitting block SP distributes the sequence such that the fields are passed alternatingly to a first or a second branch F1, F2. The first branch F1 generates BL information and EL information, but will be referred to herein as BL branch. The other branch generates 25 only EL information and will be referred to herein as EL branch. Each of the branches receives a video sequence with a temporal resolution of  $\alpha$  frames per second and a spatial resolution of  $2M \times N$  pixels. E.g. the BL sequence F1 contains even fields and the EL sequence contains odd fields. The 30 splitting block SP executes a splitting transformation. It may also utilize various filters for more sophisticated processing, e.g. low/high filtering.

The BL branch performs spatial filtering SF1 on the F1 sequence along the x-axis, i.e. within lines, and generates from the lower frequency portions of F1 a video sequence L with a spatial resolution of MxN pixels, and from the 5 higher frequency portions of F1 a spatial resolution enhancement signal H. In a simple example the filtering can be performed by a low pass filter LP<sub>R</sub> and a high pass filter HP<sub>R</sub>, both performing horizontal subsampling of e.g. factor 2, or another factor. The low-resolution sequence L 10 is input to a spatio-temporal filter STF and to a Motion Estimating (ME) block ME1. The spatio-temporal filter STF performs motion compensated temporal transformation, e.g. 2D+t wavelet transform. It may be implemented e.g. as Haar transform filter of length 2, or Daubechies filter of 15 length 9 and 7, or length 5 and 3. The ME block ME1 generates motion vectors MV1, which serve three purposes: first they are used to optimize the spatio-temporal filtering process, second they are transmitted to the receiver to be used e.g. for the inverse filtering process, 20 and third they will be used in the EL branch of the coder.

The output signal P of the spatio-temporal filter STF is transmitted, and can be combined with the motion vectors MV1 by a receiver to reproduce e.g. an MxN pixel,  $\alpha$  Hz 25 progressive mode video sequence. With the described implementation it is also possible to design the spatio-temporal filter such that this progressive sequence is scalable with respect to temporal resolution and to spatial resolution in both x and y direction. The filter may 30 contain several filtering levels, depending on the target frame rate or format, so that e.g. a spatial resolution of M'xN' and/or a temporal resolution of  $\theta\alpha$  Hz can be achieved, with M' and N' ranging in [0,M] [0,N] and  $\theta \in Q_{01}$ ,

Q<sub>01</sub> being the rational numbers ranging in [0,1]. This filtering can deliver different spatio-temporal frequencies whose fittings, e.g. organization within the data stream, make the base layer fully scalable. The output P of this 5 block STF is a progressive sequence with a resolution of M'xN', which is a sub-resolution of the original 2Mx2N format, and it represents the lowest spatial and temporal resolution level the inventive system generates.

10 The spatial resolution enhancement signal H, containing the high frequency fields along the x-axis, belongs to the enhancement layer EL, and can be used by advanced receivers to reconstruct frames with higher spatial resolution in x direction, up to the original resolution.

15 The EL branch performs MCTF on the F2 sequence. It also uses MVs in order to minimize the amount of data to be transmitted when an object moves its position in one picture relative to another picture, but other than the BL 20 it can use different sets of MVs, as described above, namely MVs generated by comparing fields of same or opposite parity. Advantageously, the MV1 vectors generated in the ME1 block of the BL branch can be used, since it is very likely that the motion between e.g. two successive 25 even fields is very similar to the motion between the two corresponding odd fields, or vice versa. Still it is possible that motion changes within a sequence of four fields, so that it may be advantageous to use better motion vectors MV2, generated from the interlaced sequence I in a 30 second motion estimation block ME2. Therefore a decision block CTL is integrated in the EL branch in order to decide which set of MVs is to be used for MCTF. The decision block CTL evaluates the different MV encoding options for the EL,

calculates the respective coding cost, estimates the resulting distortion, weights and adds these values and stores the result for each evaluated option in a memory element MEM. When it has evaluated all allowed options, it 5 selects the option that results in the lowest value, and enables the output of the respective data. Therefore it can control the outputs of the ME blocks ME1,ME2, the MCTF block MCTF and the spatio-temporal filter block STF via a control bus CB. Further, it may transmit an identifier ID 10 for the selected option.

The motion compensated temporal filtering block MCTF makes it possible to generate filtered F2 fields FF2, i.e. F2 fields with lower cost than the original F2 fields. These 15 filtered fields FF2 have the same spatial and temporal resolution as the input F2 fields. Coding cost reduction is achieved by using the described MVs, since the amount of transmitted data is reduced. ME and motion compensation can be performed across fields, independent from parity, in 20 order to capture intra- and inter-fields correlation. Any prediction mode, backward, forward or bi-directional, as well as a multiple reference fields algorithm can be used. In order to reduce MV coding cost and the estimation 25 complexity, motion information from the base layer may be used 'a priori' or 'a posteriori', depending on the targeted requirements.

For transmission, the resulting streams are usually multiplexed into one bitstream. The described encoding 30 method decomposes the interlaced original video sequence in two layers, BL and EL, which allows the resulting bitstream to support a variety of spatial and temporal resolution levels, at the coder side as well as at the decoder side.

The described EL may also be viewed upon as a plurality of layers, since it contains several levels of resolution.

The option of using frames from the other layer is  
5 available to the EL, but not to the BL, since the EL data  
are not available to the BL.

After reception, a decoder can convert the signals to video sequences at different levels of temporal resolution,  
10 either progressive or interlaced, and at various levels of spatial resolution. It can utilize the identifier ID to choose for each EL frame, or region within an EL frame, the correct prediction source and MV.

15 The BL contains a fully scalable description of the input video source in progressive mode, wherein scalability refers to spatial resolution, temporal resolution and quality. The EL, when combined with the BL, allows a low-cost, scalable reconstruction of the original interlaced  
20 video material, including the achievement of original spatial resolution, temporal resolution and SNR quality.

Distortion is in video coding commonly understood as a measure of the quality of an encoded picture compared to  
25 the source picture. It can be estimated using objective criteria, e.g. the sum of square differences, the sum of absolute differences etc. It can be estimated on intermediate pictures, before having effectively computed the encoded and decoded picture.

30

The inventive method can be applied to all video encoding schemes with motion vectors on different quality layers.

Claims

1. Method for encoding data for interlaced video using motion compensated predictive coding, wherein video frames are separated into base layer frames and enhancement layer frames, the method comprising the steps of
  - performing motion estimation for the base layer frames, wherein base layer frames are referenced, resulting in first motion vectors;
  - performing motion estimation for the enhancement layer frames, wherein for a single enhancement layer frame different sets of second motion vectors based on different references are generated, the different references being one or more base layer frames, or one or more enhancement layer frames, or combinations thereof;
  - calculating coding costs for each of the different sets of second motion vectors, the coding costs corresponding or being equal to the number of bits required for encoding the enhancement layer frame;
  - estimating a distortion value for each of the different sets of second motion vectors, the distortion value being based on the difference between the enhancement layer frame and a frame that is predicted based on the current set of second motion vectors and the current reference;
  - selecting the reference and set of second motion vectors that provides the best compromise of low coding costs and low distortion; and
  - transmitting an identifier for the selected reference.

2. Method according to claim 1, further comprising the  
step of selecting for the motion vectors (MV3) of  
each of the different sets of second motion vectors  
5 reference motion vectors (V) belonging to the first  
motion vectors, and calculating a scaling factor and  
a difference vector (UV), wherein the scaling  
factors and difference vectors are used for the  
calculation of the coding costs and the following  
10 steps.

3. Encoder for interlaced video data, the encoder using  
motion compensated predictive coding and comprising  
- means (SP,SF<sub>1</sub>) for separating video frames into  
15 base layer frames and enhancement layer frames;  
- means (ME1) for performing motion estimation for  
the base layer frames, referencing only base layer  
frames and resulting in first motion vectors  
(MV1);  
- means (MCTF,ME2) for performing motion estimation  
20 for the enhancement layer frames, wherein for a  
single enhancement layer frame different sets of  
second motion vectors based on different  
references are generated, the different references  
being one or more base layer frames, or one or  
25 more enhancement layer frames, or combinations  
thereof;  
- means (CTL) for calculating coding costs for each  
of the different sets of second motion vectors,  
30 the coding costs corresponding or being equal to  
the number of bits required for encoding the  
enhancement layer frame;

- means (CTL) for estimating a distortion value for each of the different sets of second motion vectors, the distortion value being based on the difference between the enhancement layer frame and a frame that is predicted based on the current set of second motion vectors and the current reference;
- means (CTL, MEM) for storing the estimated coding costs and distortion values corresponding to the different types of motion estimation, wherein each stored coding cost value and distortion value is associated with an identifier for the corresponding set of second motion vectors;
- means (CTL) for comparing the coding costs for the different sets of second motion vectors, selecting the motion vectors (MV2) that provide the least coding costs and determining the associated motion estimation; and
- means (CTL) for transmitting an identifier (ID) with the encoded video data stream, the identifier indicating the utilized encoding type.

4. Encoder according to claim 3, further comprising means (CTL) for calculating scaling factors and difference vectors between each of the different sets of second motion vectors and the first motion vectors, wherein the difference vectors are used for the calculation of the coding costs and the following steps.

5. Method or encoder according to any of the preceding claims, wherein the motion estimation for the enhancement layer is initialized with the first

motion vectors.

6. Method according to any of the preceding claims,  
wherein the motion estimation for the enhancement  
5 layer comprises at least two out of the following:  
backward prediction, forward prediction and  
bi-directional prediction.

10 7. Method according to any of the preceding claims,  
wherein the motion estimation for a frame of the  
enhancement layer may reference to a single frame or  
to a plurality of frames.

15 8. Decoder for generating interlaced video data  
comprising  
- means for receiving an identifier, the identifier  
being associated to a motion vector or motion  
field and indicating a reference frame for a  
region of a prediction frame or a complete  
20 prediction frame;  
- means for decoding a prediction frame or a region  
within a prediction frame based on said identifier  
and on said motion vector.

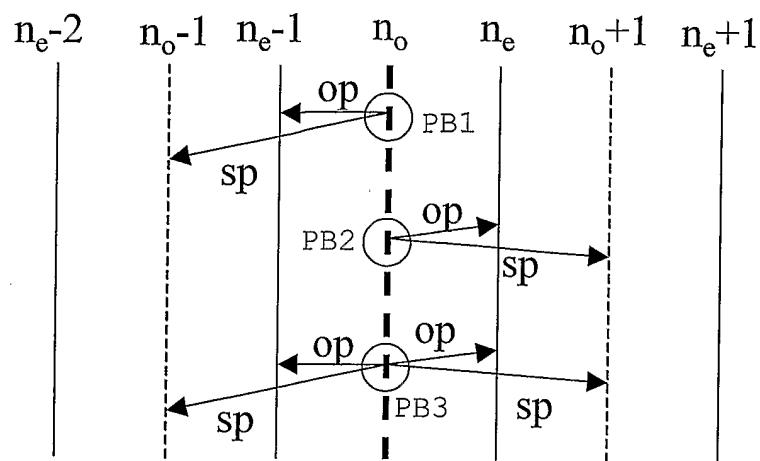


Fig.1

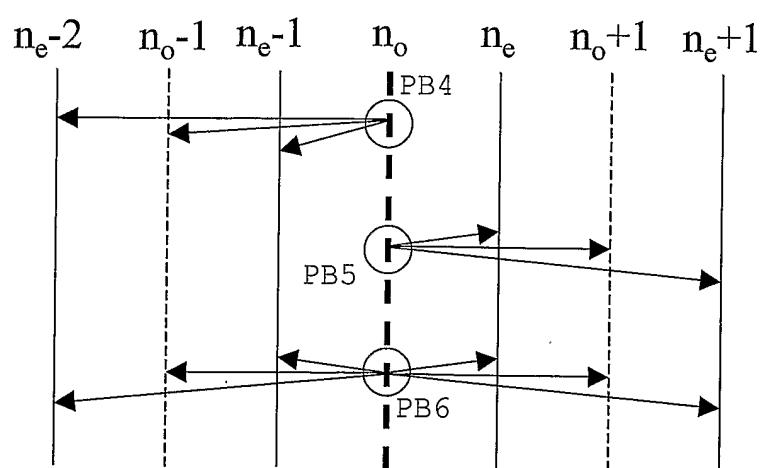
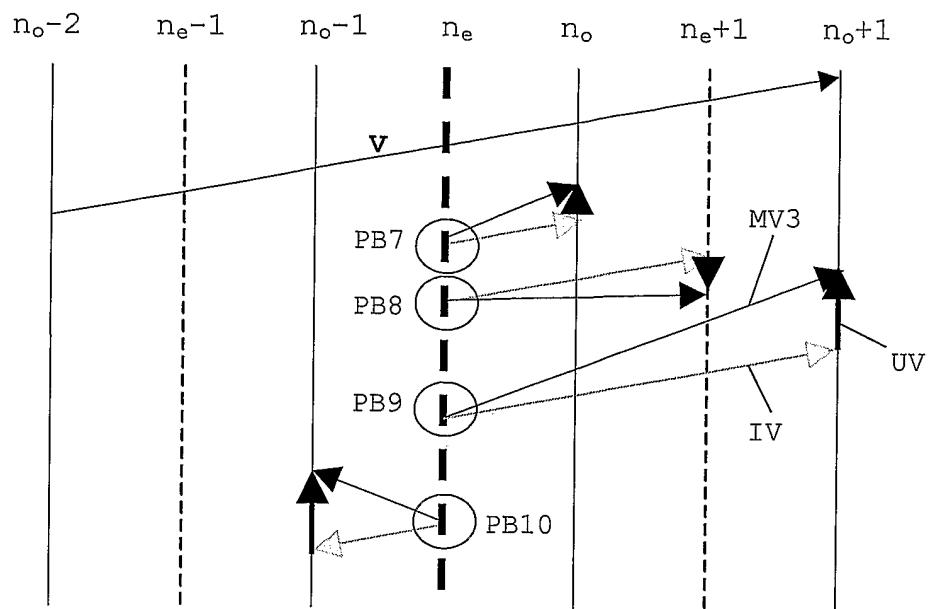
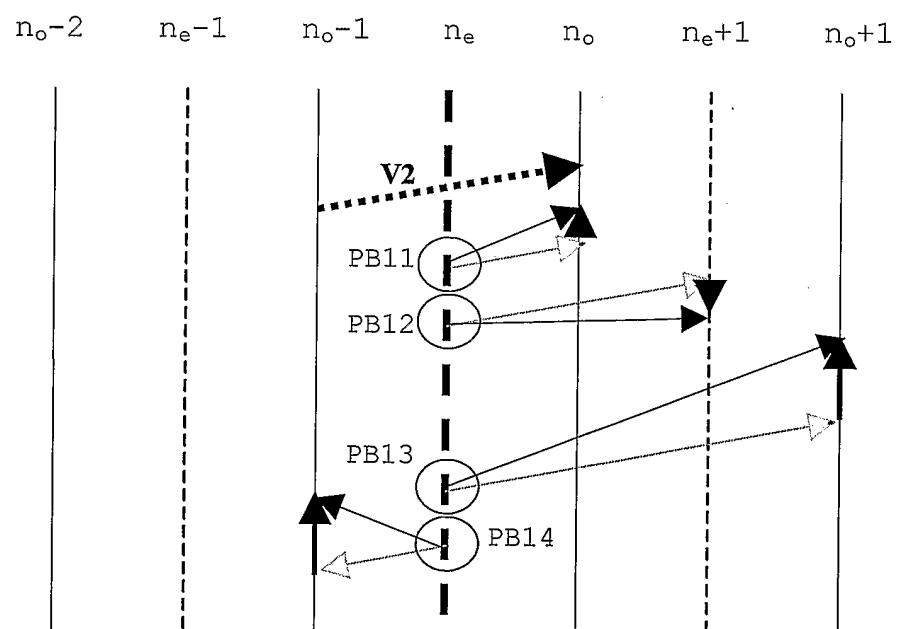


Fig.2

**Fig. 3****Fig. 4**

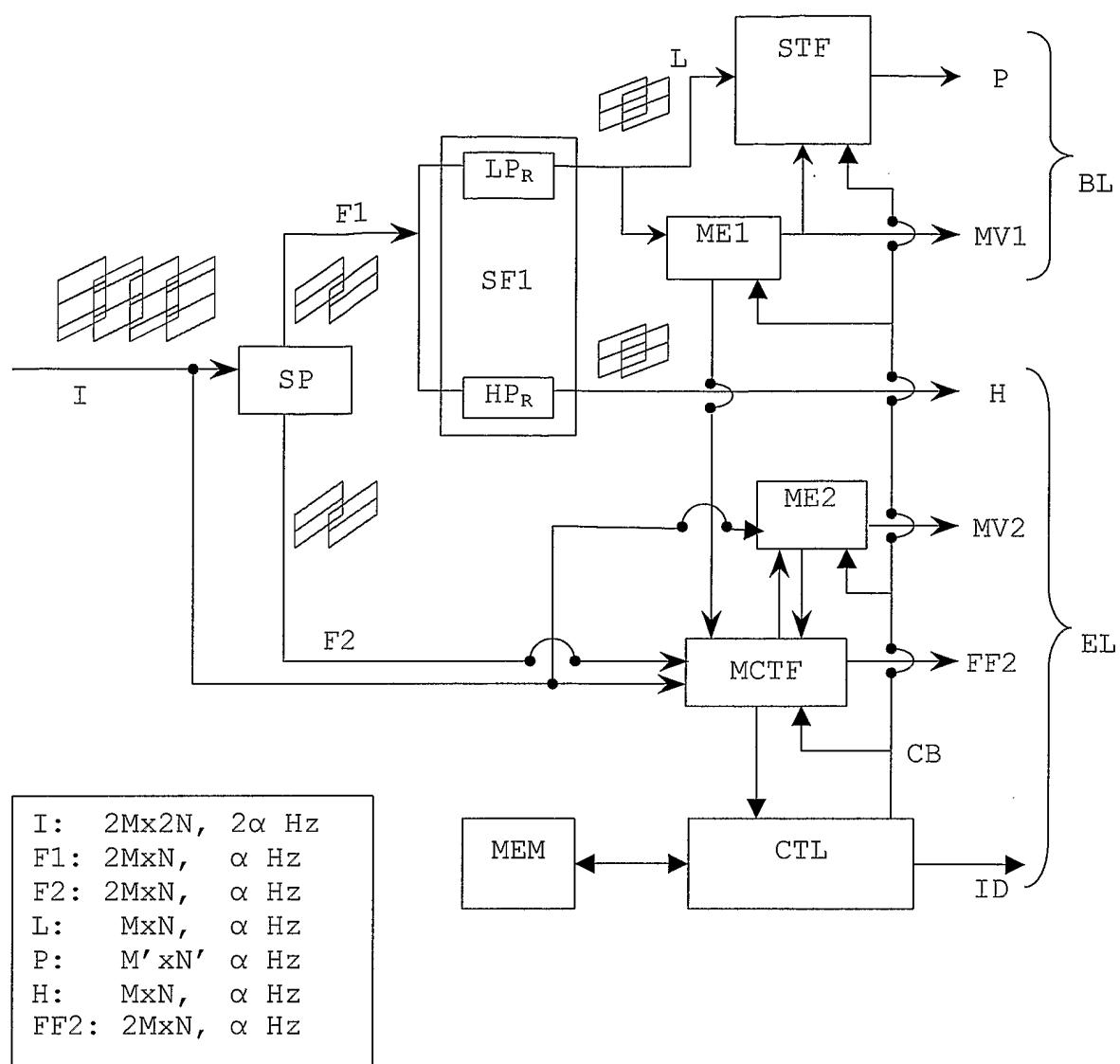


Fig. 5

# INTERNATIONAL SEARCH REPORT

Intel. Application No  
PCT/EP2005/002072

**A. CLASSIFICATION OF SUBJECT MATTER**  
IPC 7 G06T9/00

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)  
IPC 7 G06T H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, INSPEC, COMPENDEX, IBM-TDB

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0 644 695 A (BELL COMMUNICATIONS RES ;AT & T CORP (US)) 22 March 1995 (1995-03-22) abstract column 1, line 3 - line 7 column 2, line 33 - line 42 column 3, line 1 - line 26 column 5, line 8 - column 6, line 36 column 9, line 21 - column 11, line 44 column 13, line 1 - column 19, line 10 figures 1,14 ----- -/-	8
A		1-7



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

\* Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*&\* document member of the same patent family

Date of the actual completion of the international search

13 July 2005

Date of mailing of the international search report

29/07/2005

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL-2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl  
Fax: (+31-70) 340-3016

Authorized officer

Luckett, P

## INTERNATIONAL SEARCH REPORT

International Application No  
PCT/EP2005/002072

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0 883 300 A (GEN INSTRUMENT CORP) 9 December 1998 (1998-12-09)	8
A	abstract page 2, line 5 - line 8 page 2, line 27 - page 3, line 28 page 3, line 50 - page 4, line 10 page 4, line 35 - line 45 page 5, line 16 - line 46 page 10, line 35 - page 12, line 5 page 12, line 33 - page 13, line 6 tables 2,3 figures 2,4,5 -----	1-7
X	US 5 510 840 A (YONEMITSU JUN ET AL) 23 April 1996 (1996-04-23)	8
A	abstract column 1, line 34 - column 2, line 6 column 3, line 63 - line 65 column 13, line 66 - column 14, line 5 column 14, line 43 - line 51 column 16, line 59 - column 18, line 41 figures 23,27 -----	1-7

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

 International Application No  
 PCT/EP2005/002072

Patent document cited in search report		Publication date		Patent family member(s)		Publication date
EP 0644695	A	22-03-1995	CA EP JP	2127151 A1 0644695 A2 7162870 A		22-03-1995 22-03-1995 23-06-1995
EP 0883300	A	09-12-1998	US AU AU BR CA CN CN EP JP NO TW	6057884 A 733055 B2 6993498 A 9801762 A 2238900 A1 1209020 A 1551636 A 0883300 A2 11018085 A 982508 A 406512 B		02-05-2000 03-05-2001 10-12-1998 29-06-1999 05-12-1998 24-02-1999 01-12-2004 09-12-1998 22-01-1999 07-12-1998 21-09-2000
US 5510840	A	23-04-1996	AT DE DE EP WO JP JP JP KR	171831 T 69227185 D1 69227185 T2 0573665 A1 9313626 A1 3331351 B2 3449370 B2 2002252855 A 260475 B1		15-10-1998 05-11-1998 01-04-1999 15-12-1993 08-07-1993 07-10-2002 22-09-2003 06-09-2002 01-07-2000