



(51) International Patent Classification:

H04N 19/33 (2014.01) *H04N 19/105* (2014.01)
H04N 19/46 (2014.01)

(21) International Application Number:

PCT/EP2014/050040

(22) International Filing Date:

3 January 2014 (03.01.2014)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

13305007.0 7 January 2013 (07.01.2013) EP

(71) Applicant: THOMSON LICENSING [FR/FR]; 1-5 rue
Jeanne d'Arc, F-92130 Issy-les-Moulineaux (FR).(72) Inventors: LOPEZ, Patrick; c/o Technicolor R&D
France, 975, avenue des Champs Blancs, CS17616, F-
35576 Cesson-Sévigné (FR). BORDES, Philippe; c/o
Technicolor R&D France, 975, avenue des Champs Blancs,
CS17616, F-35576 Cesson-Sévigné (FR). ANDRIVON,
Pierre; c/o Technicolor R&D France, 975, avenue des
Champs Blancs, CS17616, F-35576 Cesson-Sévigné (FR).
SALMON, Philippe; c/o Technicolor R&D France, 975,
avenue des Champs Blancs, CS17616, F-35576 Cesson-

Sévigné (FR). **HIRON, Franck**; c/o Technicolor R&D
France, 975, avenue des Champs Blancs, CS17616, F-
35576 Cesson-Sévigné (FR).

(74) Agents: LORETTE, Anne et al.; Thomson Licensing, 1-5
rue Jeanne d'Arc, F-92130 Issy-Les-Moulineaux (FR).

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

(84) **Designated States** (*unless otherwise indicated, for every
kind of regional protection available*): ARIPO (BW, GH,
GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ,
UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ,
TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,
EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV,
MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,

[Continued on next page]

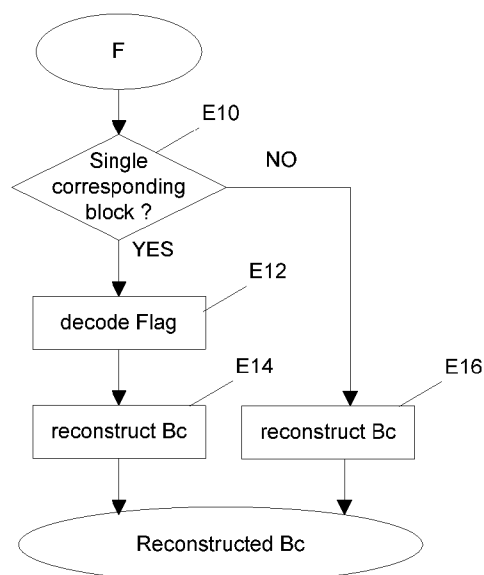
(54) Title: VIDEO CODING AND DECODING METHODS AND CORRESPONDING DEVICES INCLUDING FLAG FOR EN-
HANCEMENT LAYER

FIGURE 3

(57) **Abstract:** A method for decoding a scalable bitstream comprising a base layer and at least one enhancement layer is disclosed. The method comprises for an enhancement layer block, decoding (E10, E12) a flag indicating whether a residue of a spatially corresponding base layer block is used for the reconstruction of the enhancement layer block and reconstructing (E14, E16) the enhancement layer block depending on the flag, wherein the flag is decoded (E10, E12) only when the enhancement block has a single spatially corresponding base layer block.



TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG). **Published:**

— *with international search report (Art. 21(3))*

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

VIDEO CODING AND DECODING METHODS AND CORRESPONDING DEVICES INCLUDING FLAG FOR ENHANCEMENT LAYER

1. FIELD OF THE INVENTION

5 The invention relates to a decoding method for decoding a scalable bitstream comprising one base layer and at least one enhancement layer. The invention further relates to corresponding coding method, coding and decoding devices.

10 2. BACKGROUND OF THE INVENTION

Spatial scalability makes it possible to encode in a single bitstream, i.e. the scalable bitstream, a video content at multiple spatial resolutions. The scalable bitstream usually comprises a base layer and at least one enhancement layer. When decoding only the base layer a low resolution video content is obtained. When decoding the enhancement layer a video content of higher resolution is obtained. In order to increase coding efficiency, it is known to deduce some coding information (e.g. motion vectors) for the enhancement layer from coding information of a lower layer, e.g. the based layer. As an example, it is known for an enhancement layer block to use the motion vector of a spatially corresponding base layer block as a motion vector predictor. The motion vector of the enhancement layer block is either equal to the motion vector predictor or is equal to the motion vector predictor to which a residual vector is added. The residual vector is encoded in the scalable bitstream (respectively decoded from the scalable bitstream).

25 On the other hand, when encoding an enhancement layer block B_c , a predictor P is determined. The predictor is usually a motion compensated reference block, i.e. $P = MC(I_{ref}, MV)$. The reference block is identified in a reference picture I_{ref} by the motion vector MV . The predictor P is then subtracted from the enhancement layer block. The residue thus obtained is encoded in the scalable bitstream. On the decoder side, the residue is decoded from the scalable bitstream and added to the predictor in order to reconstruct the enhancement layer block. In order to further increase coding efficiency, it is known to add the residue $Res(UpscBL)$ of upscaled spatially corresponding base layer blocks to the motion compensated reference block

in order to get the predictor. In this case, $P = MC(I_{ref}, MV) + Res(UpscBL)$. Adding the residue of upscaled spatially corresponding base layer blocks to the motion compensated reference block in order to get the predictor is not systematic and thus requires signalling. However, signalling whether the
5 residue of upscaled spatially corresponding base layer blocks is used for the prediction of the enhancement layer block is bit consuming.

3. BRIEF SUMMARY OF THE INVENTION

The purpose of the invention is to overcome at least one of the
10 disadvantages of the prior art. For this purpose, a method for decoding a scalable bitstream comprising a base layer and at least one enhancement layer is disclosed. The method comprises for an enhancement layer block, decoding a flag indicating whether a residue of a spatially corresponding base layer block is used for the reconstruction of the enhancement layer block and
15 reconstructing the enhancement layer block depending on the flag, wherein the flag is decoded only when the enhancement block has a single spatially corresponding base layer block.

According to a specific embodiment, reconstructing the enhancement layer block depending on the flag comprises:

- 20 - determining a motion vector for the enhancement layer block;
- determining a motion compensated reference block from the motion vector;
- determining a predictor from the motion compensated reference block and depending on the flag from the residue of the spatially corresponding base layer block; and
- 25 - reconstructing the enhancement block from the predictor.

Advantageously, determining the predictor comprises adding a part of a scaled residue of the spatially corresponding base layer block co-located to the enhancement layer block to the motion compensated reference block when the flag indicates that a residue of a spatially corresponding base layer
30 block is used for the reconstruction.

According to a specific characteristic, the motion vector is a motion vector of the spatially corresponding base layer block whose components are scaled according a spatial resolution ratio between the base layer and the enhancement layer.

Advantageously, the flag is decoded only when the enhancement block has a single spatially corresponding base layer block and when the motion vector is the motion vector of the spatially corresponding base layer block whose components are scaled according a spatial resolution ratio between the base layer and the enhancement layer.

A method for encoding picture blocks in a scalable bitstream comprising a base layer and at least one enhancement layer is further disclosed. The method comprises for an enhancement layer block, encoding the enhancement layer block and encoding a flag indicating whether a residue of a spatially corresponding base layer block is used for the encoding of the enhancement layer block, wherein the flag is encoded only when the enhancement block has a single spatially corresponding base layer block.

According to a specific embodiment, encoding the enhancement layer block comprises:

- determining a motion vector for the enhancement layer block;
- determining a motion compensated reference block from the motion vector;
- determining a predictor from at least the motion compensated reference block; and
- encoding the enhancement block from the predictor and encoding the flag whose value depends on whether a residue of a spatially corresponding base layer block is used in determining the predictor.

Advantageously, determining the predictor comprises adding a part of a scaled residue of the spatially corresponding base layer block co-located to the enhancement layer block to the motion compensated reference block when the flag indicates that a residue of a spatially corresponding base layer block is used for the encoding.

According to a specific characteristic, the motion vector is a motion vector of the spatially corresponding base layer block whose components are scaled according a spatial resolution ratio between the base layer and the enhancement layer.

Advantageously, the flag is encoded only when the enhancement block has a single spatially corresponding base layer block and when the motion vector is the motion vector of the spatially corresponding base layer block whose

components are scaled according a spatial resolution ratio between the base layer and the enhancement layer.

A scalable video decoder for decoding a scalable bitstream comprising a base layer and at least one enhancement layer is also disclosed. The decoder

5 comprises means for decoding, for an enhancement layer block, a flag indicating whether a residue of a spatially corresponding base layer block is used for the reconstruction of the enhancement layer block and means for reconstructing the enhancement layer block depending on the flag, wherein the flag is decoded only when the enhancement block has a single spatially
10 corresponding base layer block.

A scalable video encoder for encoding picture blocks in a scalable bitstream comprising a base layer and at least one enhancement layer is further disclosed. The encoder comprises means for encoding an enhancement layer block and means for encoding a flag indicating whether a residue of a
15 spatially corresponding base layer block is used for the encoding of the enhancement layer block, wherein the flag is encoded only when the enhancement block has a single spatially corresponding base layer block.

A scalable video bitstream comprising a base layer and at least one enhancement layer is disclosed. The bitstream encodes for an enhancement
20 layer block a flag indicating whether a residue of a spatially corresponding base layer block is used for the encoding or reconstruction of the enhancement layer block, wherein the flag is encoded only when the enhancement block has a single spatially corresponding base layer block.

A computer program product comprising program code instructions to execute
25 of the steps of the decoding method according to the invention when this program is executed on a computer is also disclosed.

4. BRIEF DESCRIPTION OF THE DRAWINGS

Other features and advantages of the invention will appear with the
30 following description of some of its embodiments, this description being made in connection with the drawings in which:

- Figure 1 represents a quadtree structure as defined in HEVC coding standard;

- Figure 2 represents enhancement layer blocks, e.g. A and B, and corresponding base layer blocks A', B' and C';
- Figure 3 represents the flowchart of a decoding method of a scalable bitstream F with a view to the reconstruction of an enhancement layer block Bc according to the invention;
- Figure 4 depicts in detail a step of the decoding method according to the invention;
- Figure 5 represents the flowchart of a coding method according to the invention;
- Figure 6 depicts in detail a step of the coding method according to the invention;
- Figure 7 depicts a scalable video decoder according to the invention; and
- Figure 8 depicts a scalable video encoder according to the invention.

5. DETAILED DESCRIPTION OF THE INVENTION

The invention relates to a method for reconstructing a block of pixels of a sequence of pictures and a method for coding such a block. A picture sequence is a series of several pictures. Each picture comprises pixels or picture points with each of which at least one item of picture data is associated. An item of picture data is for example an item of luminance data or an item of chrominance data. Hereafter, the coding and decoding methods are described with reference to a picture block. It is clear that these methods can be applied on several picture blocks of a picture and on several pictures of a sequence with a view to the coding respectively the decoding of one or more pictures. A picture block is a set of pixels of any form. It can be a square, a rectangle. But the invention is not limited to such forms. In the following section the word block is used for picture block.

The "predictor" term designates data used to predict other data. A predictor is used to predict a picture block. A predictor or prediction block is obtained from a block or several blocks of the same picture as the picture to which belongs the block that it predicts (spatial prediction or intra-picture prediction) or from one (mono-directional prediction) or several reference blocks (bi-directional

prediction or bi-predicted) of a different picture (temporal prediction or inter-picture prediction) of the picture to which the block that it predicts belongs.

Predictors are also used to predict motion vectors. In this case, the predictor is a motion vector predictor. A motion vector predictor is for example a motion vector associated with a block located in the causal neighborhood of a current block. In case of spatial scalability, the motion vector predictor for an enhancement layer block can be the motion vector associated with a spatially corresponding base layer block after scaling, e.g. the motion vector associated with a base layer block covering the central position of the enhancement layer block. A "causal neighbourhood" of a current block designates a neighbourhood of this block that comprises pixels coded/reconstructed before the current block.

The term "residue" signifies data obtained after extraction of other data. The extraction is generally a subtraction pixel by pixel of prediction data from source data. However, the extraction is more general and notably comprises a weighted subtraction in order for example to account for an illumination variation model. In the case of motion vector, a residual motion vector ΔMV is obtained by subtracting a motion vector predictor MVP from a motion vector MVc.

The term "reconstruction" designates data (e.g. pixels, blocks) obtained after merging residues with a predictor. The merging is generally a sum of prediction data with residues. However, the merging is more general and notably comprises a weighted sum in order for example to account for an illumination variation model. A reconstructed block is a block of reconstructed pixels.

In reference to the decoding of pictures, the terms "reconstruction" and "decoding" are very often used as synonyms. Hence, a "reconstructed block" is also designated under the terminology "decoded block".

The term coding is to be taken in the widest sense. The coding can possibly comprise the transformation and/or the quantization of picture data. It can also designate only the entropy coding.

In the Figures, the represented boxes are purely functional entities, which do not necessarily correspond to physical separated entities. As will be appreciated by one skilled in the art, aspects of the present principles can be

embodied as a system, method or computer readable medium. Accordingly, aspects of the present principles can take the form of an entirely hardware embodiment, an entirely software embodiment (including firmware, resident software, micro-code, and so forth), or an embodiment combining software and hardware aspects that can all generally be referred to herein as a "circuit," "module", or "system." Furthermore, aspects of the present principles can take the form of a computer readable storage medium. Any combination of one or more computer readable storage medium(s) may be utilized.

The flowchart and/or block diagrams in the Figures illustrate the configuration, operation and functionality of possible implementations of systems, methods and computer program products according to various embodiments of the present invention. In this regard, each block in the flowchart or block diagrams may represent a module, segment, or portion of code, which comprises one or more executable instructions for implementing the specified logical function(s). It should also be noted that, in some alternative implementations, the functions noted in the block may occur out of the order noted in the figures. For example, two blocks shown in succession may, in fact, be executed substantially concurrently, or the blocks may sometimes be executed in the reverse order, or blocks may be executed in an alternative order, depending upon the functionality involved. It will also be noted that each block of the block diagrams and/or flowchart illustration, and combinations of the blocks in the block diagrams and/or flowchart illustration, can be implemented by special purpose hardware-based systems that perform the specified functions or acts, or combinations of special purpose hardware and computer instructions. While not explicitly described, the present embodiments may be employed in any combination or sub-combination.

As depicted on **Figure 1**, HEVC video coding standard ISO/IEC ISO/IEC 23008-2 *MPEG-H Part 2* defines a quadtree structure. A picture is divided into coding units CU. Each coding unit is divided into prediction units (PU) and each PU is further divided into transform units (TU). Each TU is encoded by applying an appropriate transform. On Figure 1, the CU is divided into 4 PUs. A PU can then be encoded by applying a single transform. In this case a single TU corresponds to the PU. A PU can be encoded by dividing it

horizontally or vertically into two TUs and applying a transform on each of them. A PU can also be encoded by dividing it in 4 TUs and applying a transform on each of them.

5 In spatial scalability, an enhancement layer block has spatially corresponding base layer blocks. The spatially corresponding base layer blocks are the base layer blocks that cover the enhancement layer blocks when the scaled base layer picture is superposed on the enhancement layer picture. The scaled base layer picture also called upsampled base layer picture is obtained by scaling/upsampling the base layer picture according to the spatial resolution ratio between the two layers. **Figure 2** represents enhancement layer blocks, e.g. A and B, on the top right and base layer blocks A', B', C' on the top left. The scaled base layer blocks are represented with thick lines on the bottom left part of the figure. On the bottom right part of the figure the scaled base layer blocks are superposed on the enhancement layer blocks. To the enhancement layer block A corresponds a single base layer block A' while two base layer blocks B' and C' correspond to the enhancement layer block B. On the bottom right part of the figure, the enhancement layer block A is totally covered by the single scaled base layer block A'. On the contrary, the top part of enhancement layer block B is covered by the scaled base layer block B' and the bottom part is covered by the scaled base layer block C'. To find the spatially corresponding base layer blocks for a given enhancement layer block it is not necessary to effectively do the scaling/upsampling and the superposition.

25

Figure 3 represents the flowchart of a decoding method of a scalable bitstream F with a view to the reconstruction of an enhancement layer block Bc according to the invention. In a step E10, it is checked whether the enhancement layer block Bc has a single spatially corresponding base layer block or a plurality of them. If the enhancement layer block Bc has a single spatially corresponding base layer block, then the method continues to step E12, otherwise it goes to step E16.

30

In a step E12, a flag is decoded for the enhancement layer block Bc. The flag indicates whether a spatially corresponding base layer block residue is used

for the reconstruction of the enhancement layer block Bc, i.e. in the prediction process. This flag is only decoded for the enhancement layer block Bc when the block Bc has a single spatially corresponding base layer block as block A on figure 2. Consequently no such flag is decoded for an enhancement layer block which has more than one spatially corresponding base layer block as block B on figure 2. Advantageously, less bits are used since the flag is not decoded systematically for each block of the enhancement layer.

In a step E14, the enhancement layer block is reconstructed depending on the value of the flag. As an example, when the flag takes the value '0' then the residue of the spatially corresponding base layer block is not used for the reconstruction of the enhancement layer block A and when the flag takes the value '1' then the residue of the spatially corresponding base layer block is used for the reconstruction of the enhancement layer block A. According to a variant, when the flag takes the value '1' then the residue of the spatially corresponding base layer block is not used for the reconstruction of the enhancement layer block A and when the flag takes the value '0' then the residue of the spatially corresponding base layer block is used for the reconstruction of the enhancement layer block A. In the following sections, we suppose that the value '0' indicates that the residue of the spatially corresponding base layer block is not used for the reconstruction of the enhancement layer block while the value '1' indicates that the residue of the spatially corresponding base layer block is used for the reconstruction of the enhancement layer block.

In a step E16, the enhancement layer block is reconstructed without using the residue of spatially corresponding base layer blocks. Classically, the step E16 comprises decoding a residue for the block Bc from the scalable bitstream and adding the residue to a predictor P. The predictor is usually a motion compensated reference block. The motion compensated reference block is determined in a reference picture from a motion vector decoded for the block Bc. Indeed, in this case a plurality of base layer blocks corresponds spatially to the enhancement layer block as for block B on figure 2. In this case, the use of residues of at least two different base layer blocks for the reconstruction of the enhancement layer block is probably not efficient. Indeed, when the residues of the two different base layer blocks are obtained

with different motion vectors using them for the prediction of an enhancement layer block whose motion vector is also probably different is not appropriate.

Figure 4 depicts in detail the step E14 of the decoding method according to the invention. In a step E140, a motion vector MV_c is determined for the enhancement layer block B_c . According to a specific embodiment, the motion vector MV_c is directly decoded from the enhancement layer. According to a first variant, the motion vector is directly derived from the single spatially corresponding base layer block. More precisely, the motion vector MV_c is equal to the motion vector MV_{BL_Upsc} of the single spatially corresponding base layer block scaled according to the spatial resolution ratio between the two layers. With respect to figure 2, the MV_c for the enhancement layer block A is for example the motion vector of the base layer block A' scaled according to the spatial resolution ratio between the two layers. As an example, if the enhancement layer pictures dimensions are twice the dimensions of the base layer pictures then the ratio is 2 and the scaled base layer motion vector components are equal to the components of the base layer motion vector multiplied by 2. In this case, $MV_c = MV_{BL\_Upsc}$, where MV_{BL_Upsc} is the scaled base layer motion vector.

According to another variant, the scaled motion vector is only used as a predictor and thus a residual motion vector ΔMV decoded from the scalable bitstream is added to the scaled base layer motion vector. In this case, $MV_c = MVP + \Delta MV = MV_{BL\_Upsc} + \Delta MV$.

In a step E142, a motion compensated reference block $MC(MV_c, Iref)$ is determined from the motion vector MV_c determined for the enhancement layer block in a step E140. The motion compensated reference block is determined in a reference picture $Iref$.

In a step E144, the value of the flag is checked. If the flag takes the value '0' then the method continues to step E146, otherwise it goes to step E148.

In a step E146, the predictor P of the enhancement layer block B_c is set equal to the motion compensated reference block $MC(MV_c, Iref)$.

In a step E148, the predictor P of the enhancement layer block B_c is set equal to the sum of the motion compensated reference block $MC(MV_c, Iref)$ and the co-located part $ResBL$ of the scaled/upsampled residue of the spatially

corresponding base layer block. With respect to figure 2, the co-located part is the part of the scaled/upsampled residue of the spatially corresponding base layer block A' that covers the enhancement layer block A.

In a step E150, the enhancement layer block Bc is reconstructed by adding
5 the predictor P to the residue, if any, decoded for the enhancement layer block Bc from the scalable bitstream. The decoding of the residue usually comprises entropy decoding, scaling (in the sense of inverse quantizing) and transform. According to a variant no such residue is decoded. In this case, the reconstructed block Bc is equal to its predictor P. When neither a residue nor
10 a residual motion vector is decoded for the enhancement layer block, the enhancement layer block is reconstructed according to a skip mode.

According to another embodiment, the flag is only decoded from the scalable bitstream for the blocks Bc having a single spatially corresponding
15 base layer block and whose motion vector predictor MVP is the motion vector of the single spatially corresponding base layer block scaled according to the spatial resolution ratio between the two layers. This variant encompasses the specific embodiment wherein the motion vector of the enhancement layer block is equal to the motion vector of the single spatially corresponding base
20 layer block scaled according to the spatial resolution ratio between the two layers.

Figure 5 represents the flowchart of a coding method according to the invention. In a step E50, it is checked whether the enhancement layer block
25 Bc has a single spatially corresponding base layer block or a plurality of them. If the enhancement layer block Bc has a single spatially corresponding base layer block, then the method continues to step E52, otherwise it goes to step E56.

In a step E52, the enhancement layer block Bc is encoded. The way the
30 enhancement layer block Bc is encoded depends on a coding mode selected according to a selection criterion. Indeed, on the encoder side several coding modes are checked. The best mode from a rate distortion point of view is usually selected. In the step E52, the enhancement layer block Bc is encoded with or without using the residue of the spatially corresponding base layer

block for the prediction of block Bc depending on the selected coding mode. The invention is not limited by the way the coding mode is selected. One way to select a coding mode is to check all the coding modes and to select the one that reaches the best rate-distortion compromise.

- 5 In a step E54, a flag is encoded for the enhancement layer block Bc. The flag indicates whether a spatially corresponding base layer block residue is used for the encoding (resp. reconstruction on the decoder side) of the enhancement layer block Bc. This flag is only encoded for the enhancement layer block Bc when the block Bc has a single spatially corresponding base layer block as block A depicted on figure 2. Consequently no such flag is
10 encoded for an enhancement layer block which has more than one spatially corresponding base layer block as block B on figure 2. Advantageously, less bits are used since the flag is not encoded systematically for each block of the enhancement layer.
- 15 In a step E56, the enhancement layer block is encoded without using any base layer block residue. Classically, the step E56 comprises encoding a residue for the block Bc in the scalable bitstream F. The residue is obtained by subtracting a predictor P from the block Bc. The predictor is usually a motion compensated reference block. The motion compensated reference
20 block is determined in a reference picture from a motion vector encoded for the block Bc.

Figure 6 depicts detail the step E52 of the coding method according to the invention. In a step E520, a motion vector MVc is determined for the enhancement layer block. According to a specific embodiment, the motion
25 vector is determined by motion estimation. According to a variant, the motion vector MVc equals the motion vector of the single spatially corresponding base layer block scaled according to the spatial resolution ratio between the two layers. With respect to figure 2, the MVc for the enhancement layer block
30 A is for example the motion vector of the base layer block A' scaled according to the spatial resolution ratio between the two layers. As an example, if the enhancement layer pictures dimensions are twice the dimensions of the base layer pictures then the ratio is 2 and the scaled motion vectors components

equal the components of the base layer motion vector multiplied by 2. In this case, $MV_c = MV_{BL\_Upsc}$, where MV_{BL_Upsc} is the scaled base layer motion vector. In a step E522, a motion compensated reference block $MC(MV_c, Iref)$ is determined from the motion vector MV_c determined for the enhancement layer
5 block in the step E520. The motion compensated reference block is determined in a reference picture $Iref$.

In a step E524, it is checked whether the base layer block residue is used for the prediction of block B_c . For a block B_c having a single spatially corresponding base layer block, this choice is encoder dependent. The
10 invention is not at all limited by the way the choice is made. If the encoding method chooses to use the residue of the spatially corresponding base layer block for the prediction of block B_c then the method continues to step E526, otherwise it goes to step E528.

In a step E526, the predictor P of the enhancement layer block B_c is set equal
15 to the sum of the motion compensated reference block $MC(MV_c, Iref)$ and the co-located part $ResBL$ of the scaled/upsampled residue of the spatially corresponding base layer block. With respect to figure 2, the co-located part is the part of the scaled/upsampled residue of the spatially corresponding base layer block A' that covers the enhancement layer block A . The flag for the
20 block B_c is set equal to the value '1' if this value indicates that the base layer block residue is used for the encoding of the enhancement layer block B_c .

In a step E528, the predictor P of the enhancement layer block B_c is set equal to the motion compensated reference block $MC(MV_c, Iref)$. The flag for the block B_c is thus set equal to the value '0' if this value indicates that the base
25 layer block residue is not used for the encoding of the enhancement layer block B_c .

In a step E530, the enhancement layer block B_c is encoded from the predictor P . A residue is possibly encoded for the block B_c by subtracting from the block B_c , the predictor P to get a residue. The residue is usually transformed
30 and quantized before being entropy coded. According to a variant no such residue is decoded. A residual motion vector is also possibly encoded. The residual motion vector is equal to $(MV_c - MVP)$, where MVP is a predictor. The predictor is for example MV_{BL_Upsc} . According to a variant no such residual motion vector is encoded. When neither a residue nor a residual motion vector

is encoded for the enhancement layer block, the enhancement layer block is encoded in a skip mode.

According to another embodiment, the flag is only encoded in the
5 scalable bitstream for the blocks Bc having a single spatially corresponding
base layer block and whose motion vector predictor MVP is the motion vector
of the single spatially corresponding base layer block scaled according to the
spatial resolution ratio between the two layers. This variant encompasses the
specific embodiment wherein the motion vector of the enhancement layer
10 block is equal to the motion vector of the single spatially corresponding base
layer block scaled according to the spatial resolution ratio between the two
layers.

A scalable video bitstream is also disclosed that comprises a base
15 layer and at least one enhancement layer. The video bitstream comprises for
each enhancement layer block Bc having a single spatially corresponding
base layer block a flag specifying whether the residue of said single spatially
corresponding base layer block is used for the encoding/reconstruction of the
block Bc, more precisely for its prediction. According to a variant, the
20 bitstream comprises such a flag only for the enhancement layer block Bc
having a single spatially corresponding base layer block and whose motion
vector predictor MVP is the motion vector of the single spatially corresponding
base layer block scaled according to the spatial resolution ratio between the
two layers. This variant encompasses the specific embodiment wherein the
25 motion vector of the enhancement layer block is equal to the motion vector of
the single spatially corresponding base layer block scaled according to the
spatial resolution ratio between the two layers.

When neither a residue nor a residual motion vector is encoded in the
scalable bitstream for the enhancement layer block, the enhancement layer
30 block is encoded in a skip mode.

An example of a syntax modified according to the invention is provided
for the HEVC video coding standard. All the modifications are provided in
italics. The syntax as defined in document JCT-VC K1003 "*High Efficiency*

Video Coding (HEVC) text specification draft is modified as follows at the prediction unit level:

- idx_mv_baselayer* is defined as the index in the merging candidate list corresponding to the motion vector candidate derived from the base layer collocated PU, i.e. the spatially corresponding base layer block. If such a candidate is not present in the merging candidate list, for instance if the base layer PU was coded in Intra mode, *idx_mv_baselayer* is set to -1.

- A new flag *use_base_residue_flag* is conditionnally read from the bitstream if *merge_idx* corresponds to *idx_mv_baselayer* and if *cu_single_basis[x0][y0]* is true. *cu_single_basis[x0][y0]* indicates that the current enhancement layer PU is included in one single base-layer PU, i.e. has a single spatially corresponding base layer block. Depending on its value, the scaled/upsampled residue will be added to the predictor.

	Descriptor
if(cu_skip_flag[x0][y0]) {	
if(MaxNumMergeCand > 1)	
merge_idx [x0][y0]	ae(v)
if (merge_idx[x0][y0] == idx_mv_baselayer && cu_single_basis[x0][y0])	
use_base_residue_flag	ae(v)
} else { /* MODE_INTER */	
merge_flag [x0][y0]	ae(v)
if(merge_flag[x0][y0]) {	
if(MaxNumMergeCand > 1)	
merge_idx [x0][y0]	ae(v)
} else {	
if(slice_type == B)	
inter_pred_idc [x0][y0]	ae(v)
if(inter_pred_idc[x0][y0] != Pred_L1) {	
if(num_ref_idx_l0_active_minus1 > 0)	
ref_idx_l0 [x0][y0]	ae(v)
mvd_coding(x0, y0, 0)	
mvp_l0_flag [x0][y0]	ae(v)
}	
if(inter_pred_idc[x0][y0] != Pred_L0) {	
if(num_ref_idx_l1_active_minus1 > 0)	
ref_idx_l1 [x0][y0]	ae(v)
if(mvd_l1_zero_flag && inter_pred_idc[x0][y0] == Pred_BI) {	
MvdL1[x0][y0][0] = 0	
MvdL1[x0][y0][1] = 0	
} else	
mvd_coding(x0, y0, 1)	
mvp_l1_flag [x0][y0]	ae(v)
}	
}	
}	
}	

The decoding process (see section 8.5.4 of JCT-VC K1003) is thus modified as follows:

Depending on `rqf_root_cbf`, the following applies:

- 5 – If `rqf_root_cbf` is equal to 0,

- If *use_base_residue_flag* is equal to 0, all samples of the $(nCS_L) \times (nCS_L)$ array *resSamples_L* and all samples of the two $(nCS_C) \times (nCS_C)$ arrays *resSamples_{cb}* and *resSamples_{cr}* are set equal to 0.
- Otherwise (*use_base_residue_flag* is equal to 1) , all samples
5 of the $(nCS_L) \times (nCS_L)$ array *resSamples_L* and all samples of the two $(nCS_C) \times (nCS_C)$ arrays *resSamples_{cb}* and *resSamples_{cr}* are derived from the colocated base layer residue after upsampling.

Figure 7 depicts a scalable video decoder DEC according to the
10 invention. The decoder DEC receives on a first input IN a scalable bitstream. The input is connected to a demultiplexer DEMUX configured to demultiplex the bitstream into a base layer BL and at least one enhancement layer EL. According to a variant not represented on Figure 7, the decoder DEC comprises two inputs and receives on a first input the base layer BL and on
15 the second input the enhancement layer EL. In this case, the decoder DEC does not comprise DEMUX. The DEMUX if present is connected to a base layer decoder 'BL decoder' adapted to decode base layer pictures. The DEMUX if present is further connected to an enhancement layer decoder 'El decoder' adapted to decode enhancement layer pictures. The enhancement
20 layer decoder is adapted to implement the steps E10 to E16 of the decoding according to one of the various embodiments of the decoding method. The decoded base layer pictures are then possibly transmitted on a first output of the decoder OUT1 and the decoded enhancement layer pictures are then transmitted on a second output of the decoder OUT2. As another example,
25 only decoded enhancement layer pictures are transmitted on the second output OUT2 for being displayed on a 2D display such as an UHDTV display. According to a variant not represented on Figure 7, the decoder DEC only comprises the second output of the decoder OUT2.

30 **Figure 8** depicts a scalable video encoder according to the invention. The encoder ENC receives on a first input IN1 base layer pictures and on a second input enhancement layer pictures. The first input IN1 is connected to a base layer encoder adapted to encode and reconstruct the base layer pictures. The second input IN2 is connected to an enhancement layer encoder

adapted to implement the steps E50 to E56 and thus to encode the enhancement layer pictures according to one of the various embodiments of the encoding method.

5 The scalable video coders and decoders according to the invention are for example implemented in various forms of hardware, software, firmware, special purpose processors, or a combination thereof. Preferably, the present principles may be implemented as a combination of hardware and software. Moreover, the software is preferably implemented as an application program
10 tangibly embodied on a program storage device. The application program may be uploaded to, and executed by, a machine comprising any suitable architecture. Preferably, the machine is implemented on a computer platform having hardware such as one or more central processing units (CPU), a random access memory (RAM), and input/output (I/O) interface(s). The
15 computer platform also includes an operating system and microinstruction code. The various processes and functions described herein may either be part of the microinstruction code or part of the application program (or a combination thereof) that is executed via the operating system. In addition, various other peripheral devices may be connected to the computer platform
20 such as an enhancement data storage device and a printing device.

According to variants, the coding and decoding devices according to the invention are implemented according to a purely hardware realisation, for example in the form of a dedicated component (for example in an ASIC (Application Specific Integrated Circuit) or FPGA (Field-Programmable Gate
25 Array) or VLSI (Very Large Scale Integration) or of several electronic components integrated into a device or even in a form of a mix of hardware elements and software elements.

Claims

1. A method for decoding a scalable bitstream comprising a base layer and at least one enhancement layer, the method comprising for an enhancement layer block, decoding (E10, E12) a flag indicating whether a residue of a spatially corresponding base layer block is used for the reconstruction of said enhancement layer block and reconstructing (E14, E16) said enhancement layer block depending on said flag, wherein said flag is decoded (E10, E12) only when said enhancement block has a single spatially corresponding base layer block.
2. The method according to claim 1, wherein reconstructing (E14, E16) said enhancement layer block depending on said flag comprises:
- determining (E140) a motion vector for said enhancement layer block;
 - determining (E142) a motion compensated reference block from said motion vector;
 - determining (E144, E146, E148) a predictor from said motion compensated reference block and depending on said flag from said residue of said spatially corresponding base layer block; and
 - reconstructing (E150) said enhancement block from said predictor.
3. The method according to claim 2, wherein determining (E144, E146, E148) the predictor comprises adding a part of a scaled residue of said spatially corresponding base layer block co-located to said enhancement layer block to said motion compensated reference block when said flag indicates that a residue of a spatially corresponding base layer block is used for the reconstruction.
4. The method according to claim 2 or 3, wherein said motion vector is a motion vector of said spatially corresponding base layer block whose components are scaled according a spatial resolution ratio between the base layer and the enhancement layer.

5. The method according to claim 4, wherein said flag is decoded only when said enhancement block has a single spatially corresponding base layer block and when said motion vector is the motion vector of said spatially corresponding base layer block whose components are scaled according a spatial resolution ratio between the base layer and the enhancement layer.

6. A method for encoding picture blocks in a scalable bitstream comprising a base layer and at least one enhancement layer, the method comprising for an enhancement layer block, encoding said enhancement layer block (E52, E56) and encoding (E50, E54) a flag indicating whether a residue of a spatially corresponding base layer block is used for the encoding of said enhancement layer block, wherein said flag is encoded (E50, E54) only when said enhancement block has a single spatially corresponding base layer block.

7. The method according to claim 6, wherein encoding (E52, E56) said enhancement layer block comprises:

- determining (E140) a motion vector for said enhancement layer block;
- determining (E142) a motion compensated reference block from said motion vector;
- determining (E144, E146, E148) a predictor from at least said motion compensated reference block; and
- encoding (E150) said enhancement block from said predictor and encoding (E54) said flag whose value depends on whether a residue of a spatially corresponding base layer block is used in determining the predictor.

8. The method according to claim 7, wherein determining (E144, E146, E148) the predictor comprises adding a part of a scaled residue of said spatially corresponding base layer block co-located to said enhancement layer block to said motion compensated reference block when said flag indicates that a residue of a spatially corresponding base layer block is used for the encoding.

9. The method according to claim 7 or 8, wherein said motion vector is a motion vector of said spatially corresponding base layer block whose

components are scaled according a spatial resolution ratio between the base layer and the enhancement layer.

10. The method according to claim 7, wherein said flag is encoded only when
5 said enhancement block has a single spatially corresponding base layer block and when said motion vector is the motion vector of said spatially corresponding base layer block whose components are scaled according a spatial resolution ratio between the base layer and the enhancement layer.

10 11. A scalable video decoder for decoding a scalable bitstream comprising a base layer and at least one enhancement layer, the decoder comprising means for decoding, for an enhancement layer block, a flag indicating whether a residue of a spatially corresponding base layer block is used for the reconstruction of said enhancement layer block and means for reconstructing
15 said enhancement layer block depending on said flag, wherein said flag is decoded only when said enhancement block has a single spatially corresponding base layer block.

12. A scalable video decoder according to claim 11, wherein said decoder is
20 adapted to execute the steps of the method according to any of claims 1 to 5.

13. A scalable video encoder for encoding picture blocks in a scalable bitstream comprising a base layer and at least one enhancement layer, the encoder comprising means for encoding an enhancement layer block and
25 means for encoding a flag indicating whether a residue of a spatially corresponding base layer block is used for the encoding of said enhancement layer block, wherein said flag is encoded only when said enhancement block has a single spatially corresponding base layer block.

30 14. A scalable video encoder according to claim 13, wherein said encoder is adapted to execute the steps of the method according to any of claims 6 to 10.

15. A scalable video bitstream comprising a base layer and at least one enhancement layer encoding for an enhancement layer block a flag indicating whether a residue of a spatially corresponding base layer block is used for the encoding or reconstruction of said enhancement layer block, wherein said flag
5 is encoded only when said enhancement block has a single spatially corresponding base layer block.

16. A computer program product comprising program code instructions to execute the steps of the decoding method according to any of claims 1 to 5
10 when this program is executed on a computer.

17. A computer program product comprising program code instructions to execute the steps of the encoding method according to any of claims 6 to 10 when this program is executed on a computer.

1/4

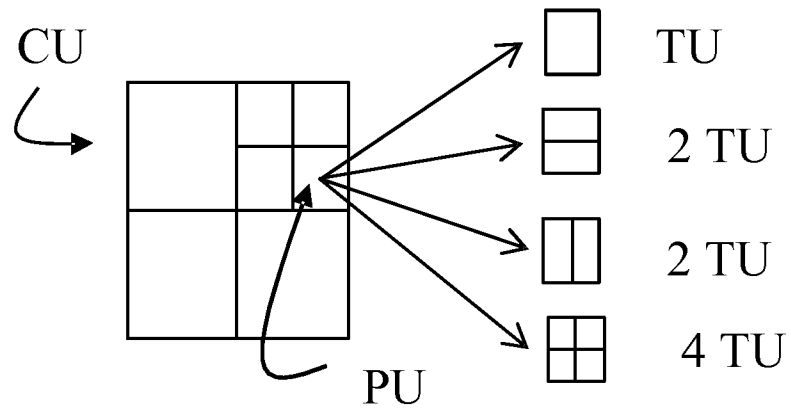


FIGURE 1

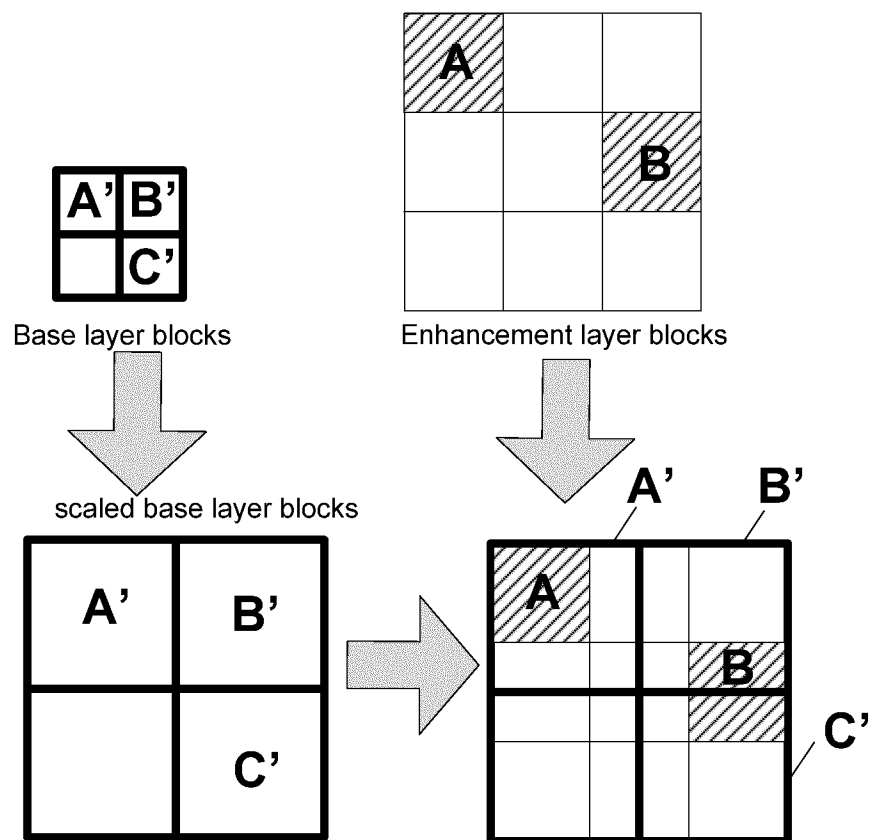


FIGURE 2

2/4

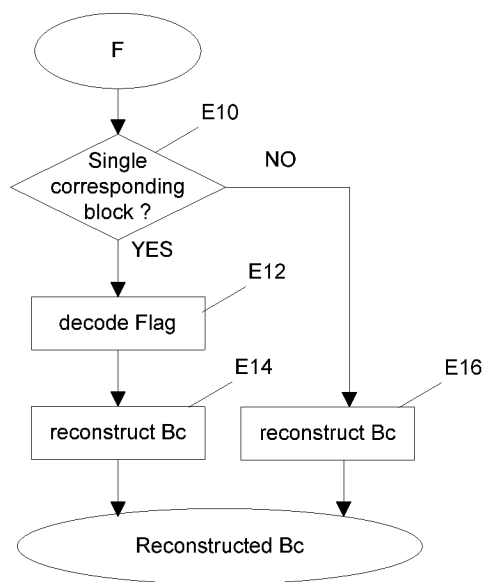


FIGURE 3

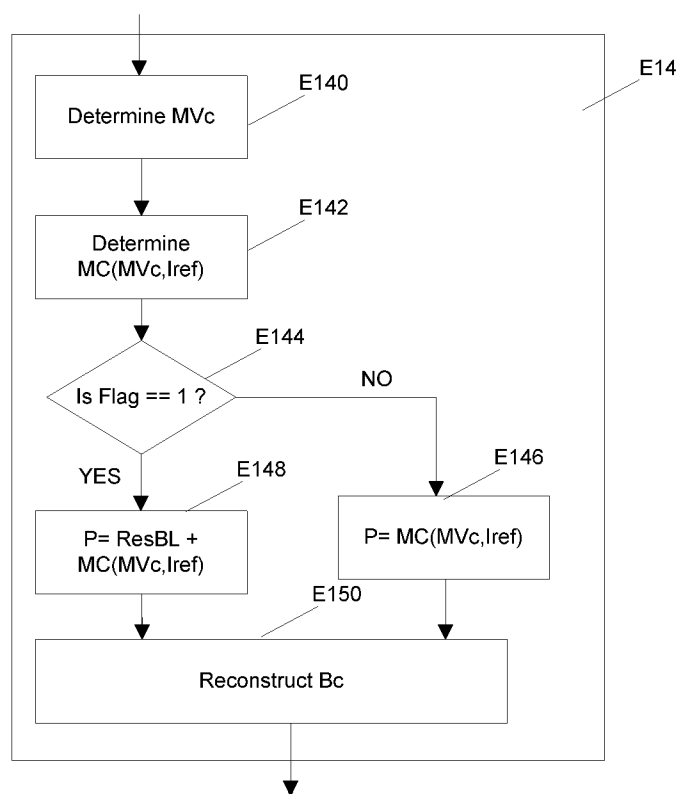


FIGURE 4

3/4

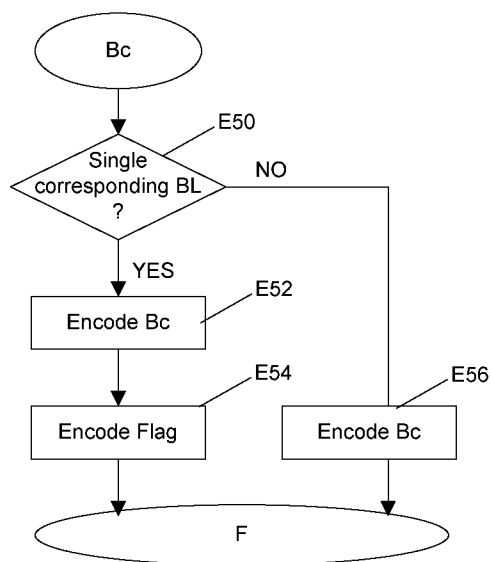


FIGURE 5

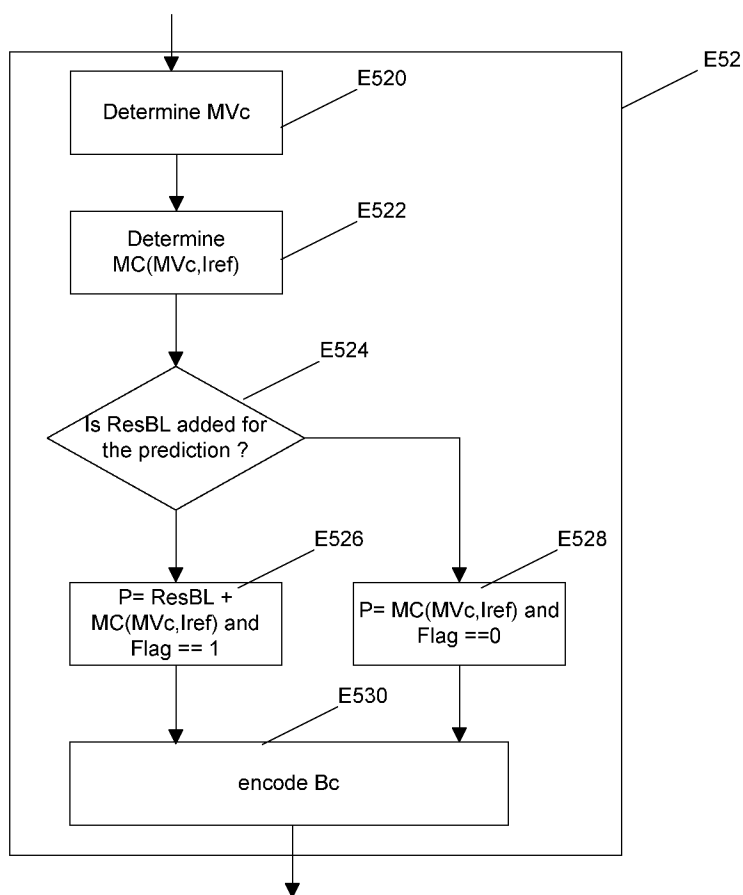


FIGURE 6

4/4

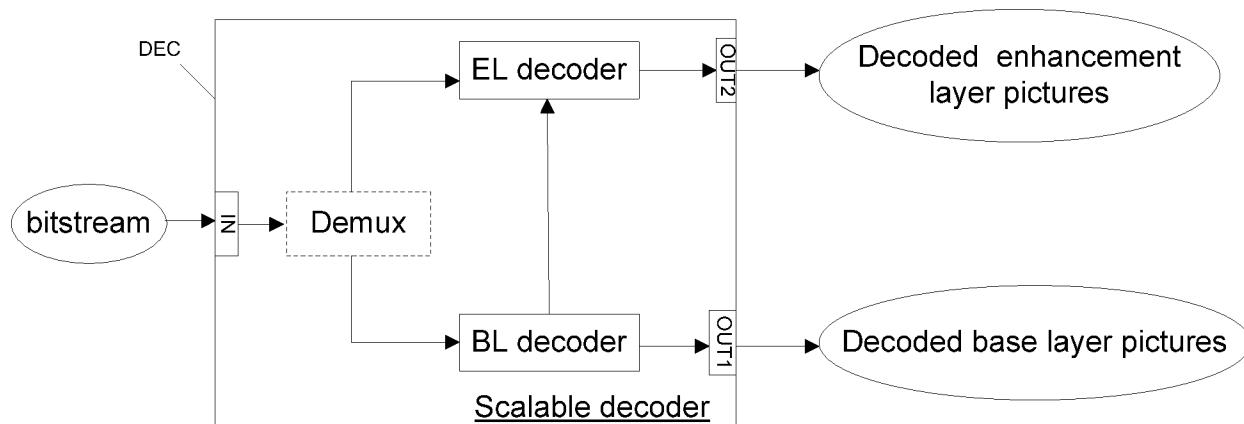


FIGURE 7

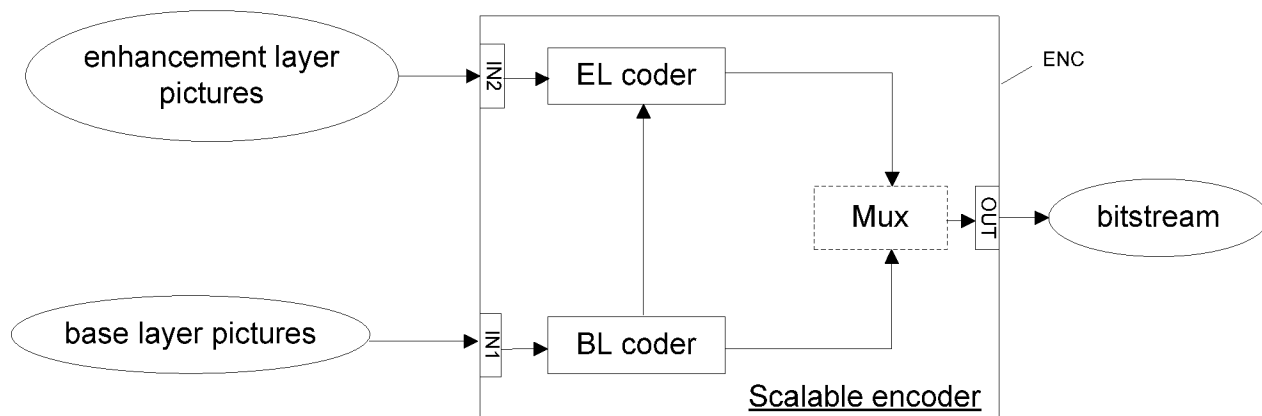


FIGURE 8

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2014/050040

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04N19/33 H04N19/46 H04N19/105
ADD.

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)
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 practicable, search terms used)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2010/158116 A1 (JEON BYEONG MOON [KR] ET AL) 24 June 2010 (2010-06-24)	15
A	paragraphs [0030], [0050] - [0087]; figures 1,4	1-14,16, 17
A	----- US 2008/165850 A1 (SAGETONG PHOOM [US] ET AL) 10 July 2008 (2008-07-10)	1-17
A	paragraphs [0025] - [0067]; figures 2-5 ----- EP 1 694 074 A1 (THOMSON LICENSING [FR]) 23 August 2006 (2006-08-23) the whole document ----- -/-	1-17



Further documents are listed in the continuation of Box C.



See patent family 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 application or patent 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

10 April 2014

Date of mailing of the international search report

22/04/2014

Name and mailing address of the ISA/

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

Authorized officer

McGrath, Simon

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2014/050040

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	SEGALL C A ET AL: "Spatial Scalability Within the H.264/AVC Scalable Video Coding Extension", IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE SERVICE CENTER, PISCATAWAY, NJ, US, vol. 17, no. 9, 1 September 2007 (2007-09-01), pages 1121-1135, XP011193020, ISSN: 1051-8215, DOI: 10.1109/TCSVT.2007.906824 the whole document -----	1-17

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2014/050040

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2010158116	A1	24-06-2010	EP 1980107 A1 15-10-2008
			EP 1985121 A1 29-10-2008
			EP 2082580 A1 29-07-2009
			JP 5063684 B2 31-10-2012
			JP 2009538084 A 29-10-2009
			JP 2009538085 A 29-10-2009
			JP 2009538086 A 29-10-2009
			KR 20080093425 A 21-10-2008
			KR 20080093426 A 21-10-2008
			KR 20080099245 A 12-11-2008
			KR 20090039795 A 22-04-2009
			US 2009010331 A1 08-01-2009
			US 2009010332 A1 08-01-2009
			US 2009060040 A1 05-03-2009
			US 2010158116 A1 24-06-2010
			WO 2008060125 A1 22-05-2008
			WO 2008060126 A1 22-05-2008
			WO 2008060127 A1 22-05-2008
US 2008165850	A1	10-07-2008	CA 2673887 A1 17-07-2008
			CN 101584217 A 18-11-2009
			EP 2116060 A1 11-11-2009
			JP 5108026 B2 26-12-2012
			JP 2010516194 A 13-05-2010
			KR 20090097903 A 16-09-2009
			RU 2009130392 A 20-02-2011
			TW 200838320 A 16-09-2008
			US 2008165850 A1 10-07-2008
			WO 2008086324 A1 17-07-2008
EP 1694074	A1	23-08-2006	CN 101204092 A 18-06-2008
			CN 101213840 A 02-07-2008
			EP 1694074 A1 23-08-2006