



US 20050002458A1

(19) **United States**(12) **Patent Application Publication** (10) **Pub. No.: US 2005/0002458 A1****Bruls et al.**(43) **Pub. Date:****Jan. 6, 2005**(54) **SPATIAL SCALABLE COMPRESSION**(30) **Foreign Application Priority Data**(76) Inventors: **Wilhelmus Hendrikus Alfonsus Bruls**,
Eindhoven (NL); **Reinier Bernardus**
Maria Klein Gunnewiek, Eindhoven
(NL)

Oct. 26, 2001 (EP) 01204066.3

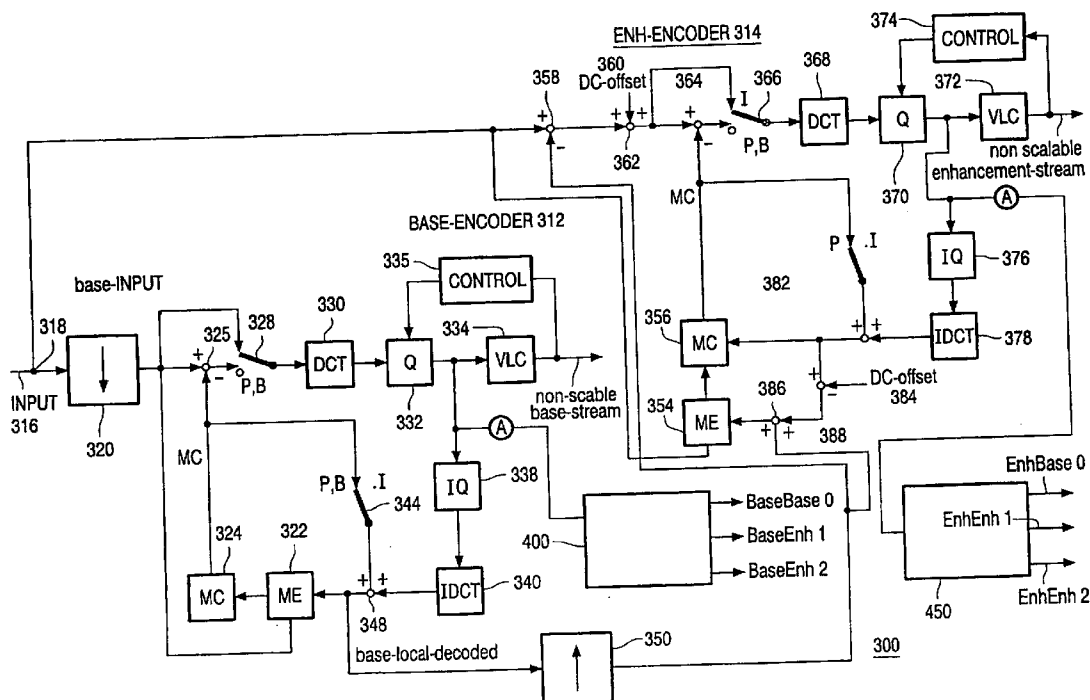
Mar. 8, 2002 (EP) 02075916.3

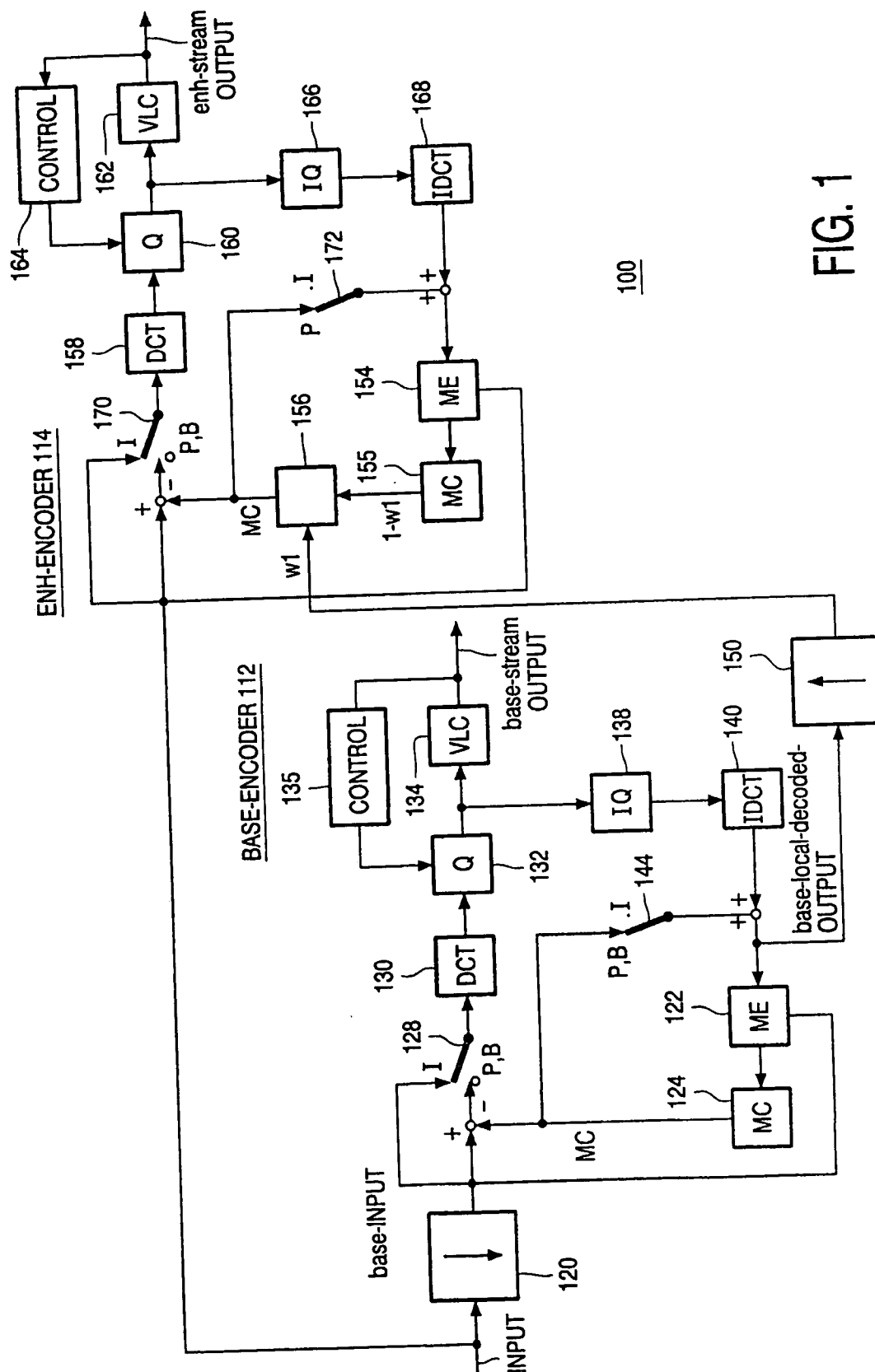
Publication Classification(51) **Int. Cl.⁷** **H04N 7/12**(52) **U.S. Cl.** **375/240.21; 375/240.03; 375/240.2;**
375/240.12

Correspondence Address:

U S Philips Corporation
Intellectual Property Department
P O Box 3001
Briarcliff Manor, NY 10510 (US)(21) Appl. No.: **10/493,267**(22) PCT Filed: **Oct. 21, 2002**(86) PCT No.: **PCT/IB02/04370**(57) **ABSTRACT**

An apparatus for efficiently performing spatial scalable compression of an input video stream is disclosed. A base encoder encodes a base encoder stream. Modifying means modifies content of the base encoder stream to create a plurality of base streams. An enhancement encoder encodes an enhancement encoder stream. Modifying means modifies content of the enhancement encoder stream to create a plurality of enhancement streams.





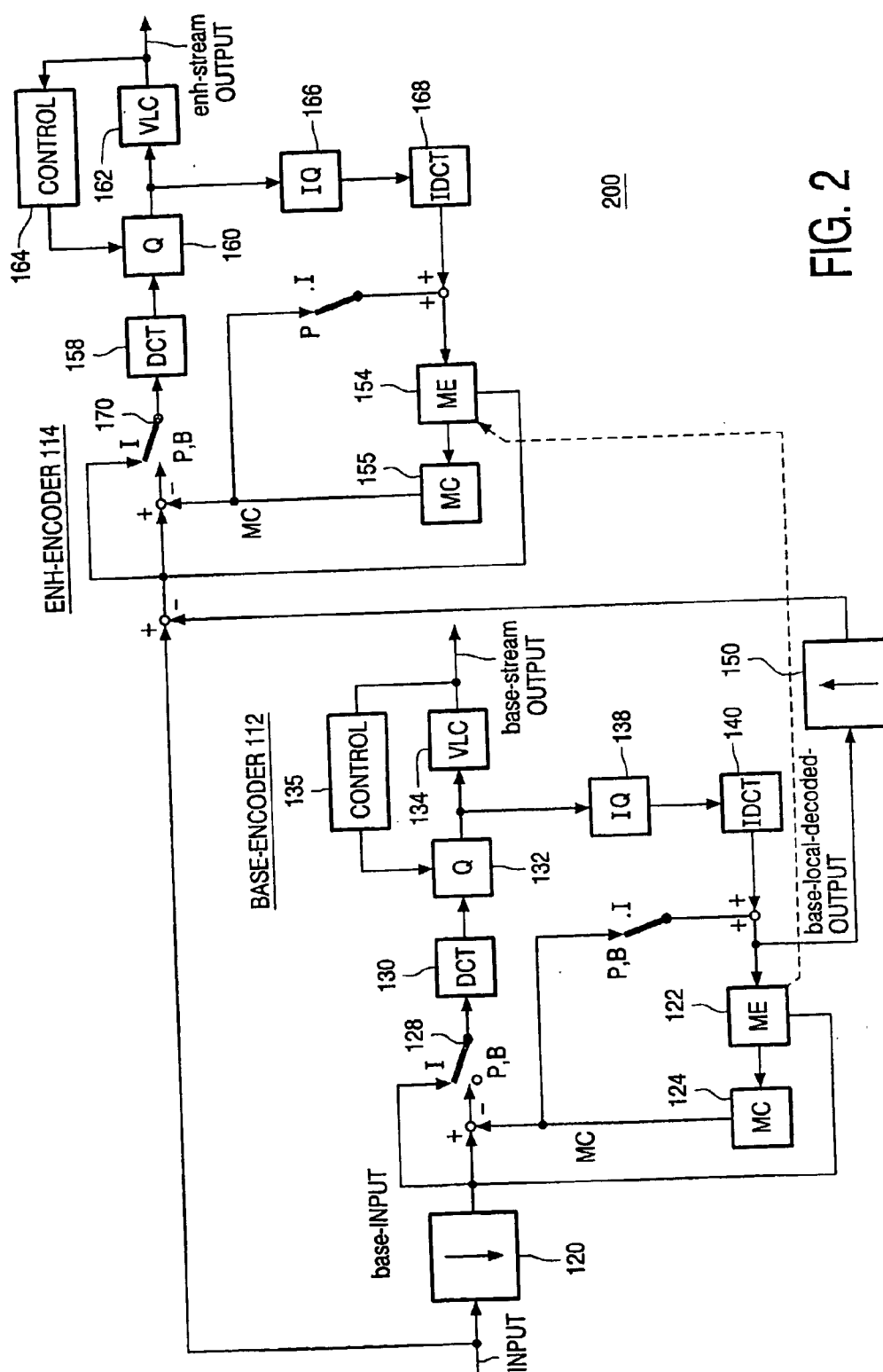


FIG. 2

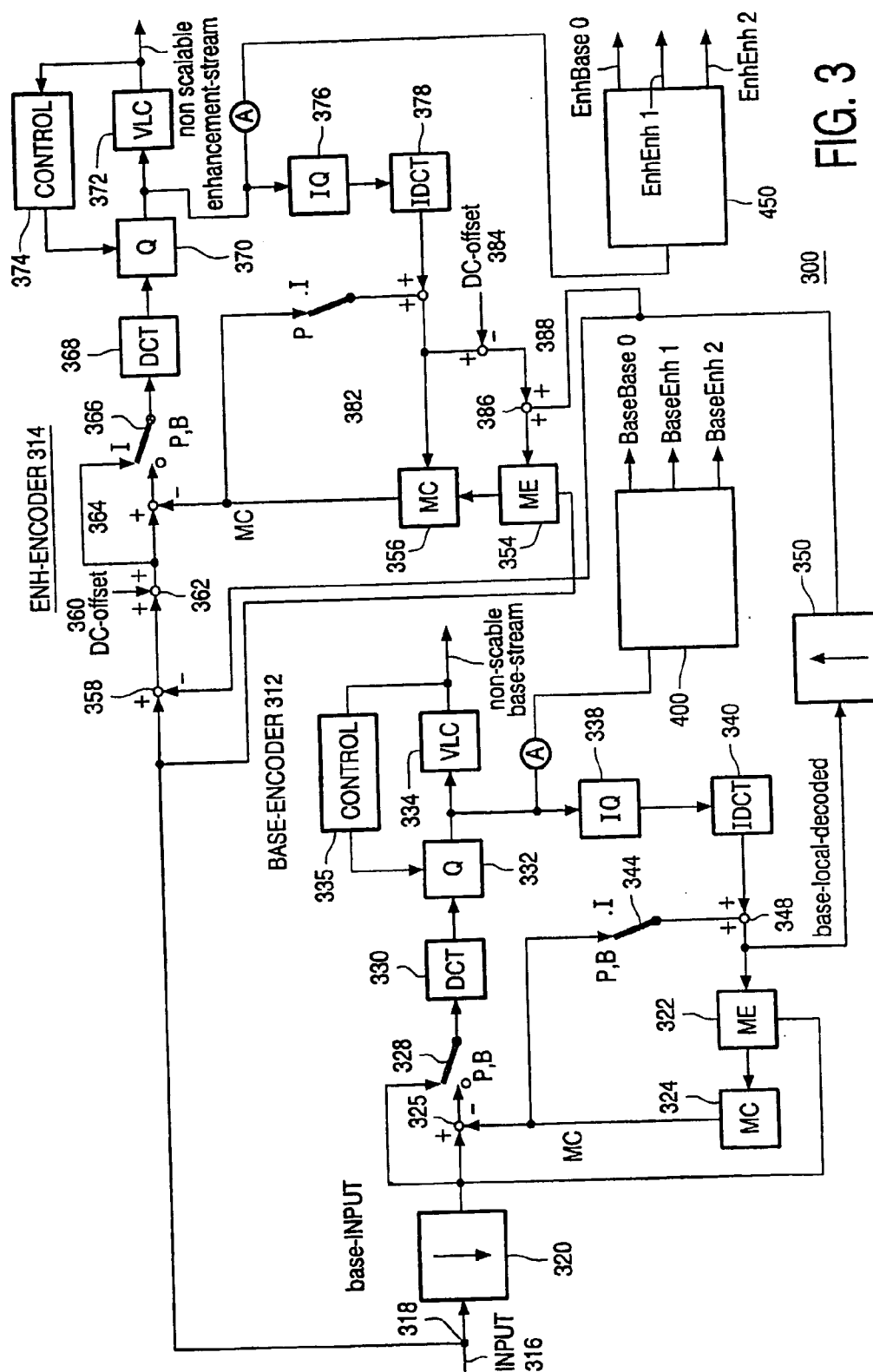


FIG. 3

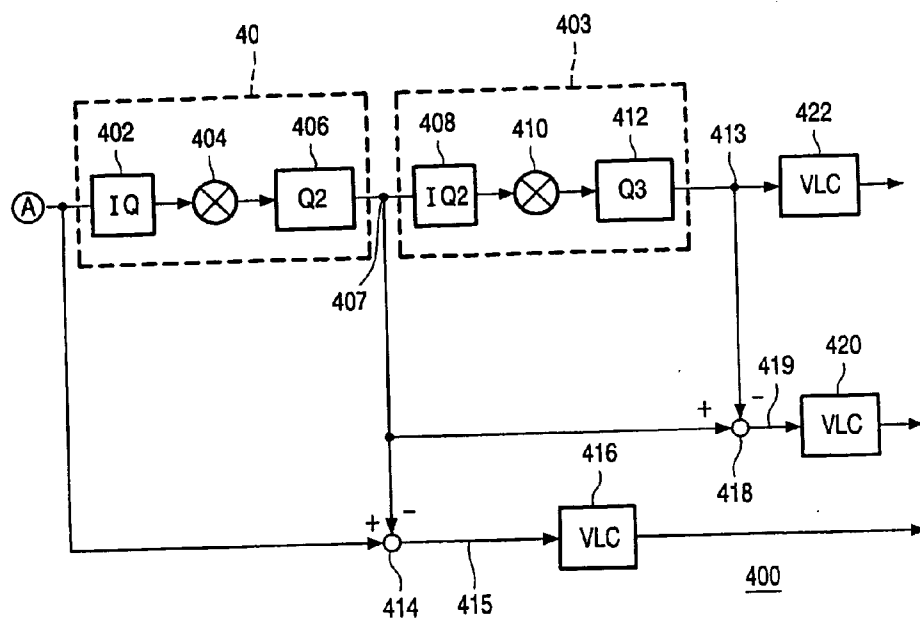


FIG. 4

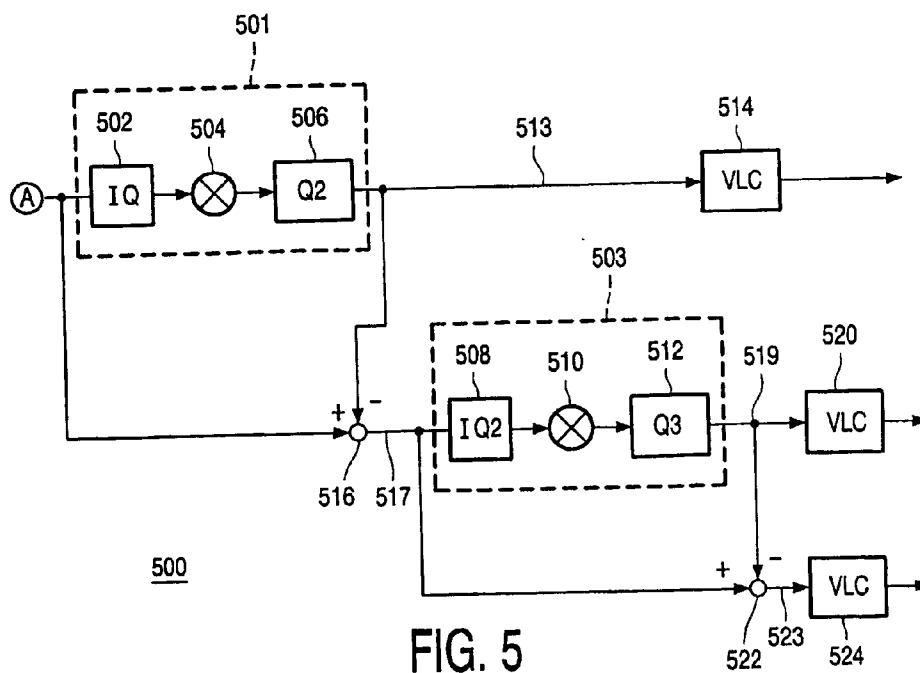


FIG. 5

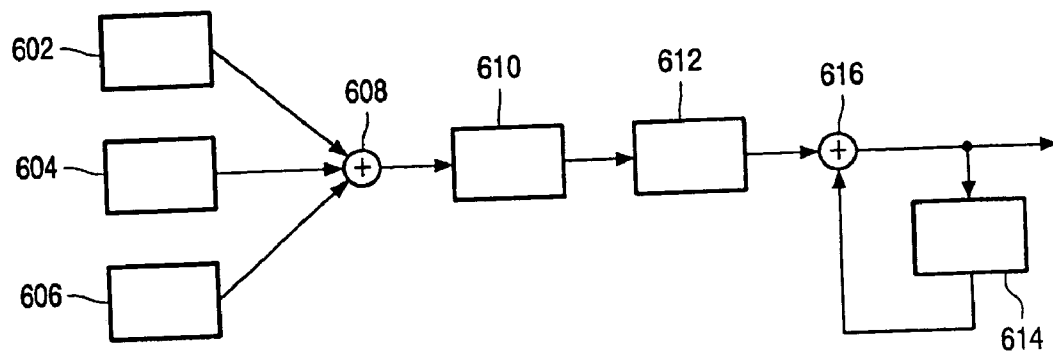


FIG. 6

SPATIAL SCALABLE COMPRESSION

FIELD OF THE INVENTION

[0001] The invention relates to a video encoder, and more particularly to a video encoder which uses spatial scalable compression schemes to produce a plurality of base streams and a plurality enhancement streams.

BACKGROUND OF THE INVENTION

[0002] Because of the massive amounts of data inherent in digital video, the transmission of full-motion, high-definition digital video signals is a significant problem in the development of high-definition television. More particularly, each digital image frame is a still image formed from an array of pixels according to the display resolution of a particular system. As a result, the amounts of raw digital information included in high resolution video sequences are massive. In order to reduce the amount of data that must be sent, compression schemes are used to compress the data. Various video compression standards or processes have been established, including, MPEG-2, MPEG-4, H.263 and H.264.

[0003] Many applications are enabled where video is available at various resolutions and/or qualities in one stream. Methods to accomplish this are loosely referred to as scalability techniques. There are three axes on which one can deploy scalability. The first is scalability on the time axis, often referred to as temporal scalability. Secondly, there is scalability on the quality axis, often referred to as signal-to-noise scalability or fine-grain scalability. The third axis is the resolution axis (number of pixels in image) often referred to as spatial scalability or layered coding. In layered coding, the bitstream is divided into two or more bitstreams, or layers. Each layer can be combined to form a single high quality signal. For example, the base layer may provide a lower quality video signal, while the enhancement layer provides additional information that can enhance the base layer image.

[0004] In particular, spatial scalability can provide compatibility between different video standards or decoder capabilities. With spatial scalability, the base layer video may have a lower resolution than the input video sequence, in which case the enhancement layer carries information which can restore the resolution of the base layer to the input sequence level.

[0005] Most video compression standards support spatial scalability. FIG. 1 illustrates a block diagram of an encoder 100 which supports MPEG-2/MPEG-4 spatial scalability. The encoder 100 comprises a base encoder 112 and an enhancement encoder 114. The base encoder is comprised of a low pass filter and downsampler 120, a motion estimator 122, a motion compensator 124, an orthogonal transform (e.g., Discrete Cosine Transform (DCT)) circuit 130, a quantizer 132, a variable length coder 134, a bitrate control circuit 135, an inverse quantizer 138, an inverse transform circuit 140, switches 128, 144, and an interpolate and upsample circuit 150. The enhancement encoder 114 comprises a motion estimator 154, a motion compensator 155, a selector 156, an orthogonal transform (e.g., Discrete Cosine Transform (DCT)) circuit 158, a quantizer 160, a variable length coder 162, a bitrate control circuit 164, an inverse quantizer 166, an inverse transform circuit 168, switches

170 and 172. The operations of the individual components are well known in the art and will not be described in detail.

[0006] Unfortunately, the coding efficiency of this layered coding scheme is not very good. Indeed, for a given picture quality, the bitrate of the base layer and the enhancement layer together for a sequence is greater than the bitrate of the same sequence coded at once.

[0007] FIG. 2 illustrates another known encoder 200 proposed by DemoGrafx. The encoder is comprised of substantially the same components as the encoder 100 and the operation of each is substantially the same so the individual components will not be described. In this configuration, the residue difference between the input block and the upsampled output from the upsampler 150 is inputted into a motion estimator 154. To guide/help the motion estimation of the enhancement encoder, the scaled motion vectors from the base layer are used in the motion estimator 154 as indicated by the dashed line in FIG. 2. However, this arrangement does not significantly overcome the problems of the arrangement illustrated in FIG. 1.

SUMMARY OF THE INVENTION

[0008] It is an object of the invention to overcome at least part of the above-described deficiencies of the known spatial scalability schemes by providing a spatial scalable compression scheme which produces a plurality of base streams with differing quality levels and a plurality of enhancement streams with differing quality levels.

[0009] According to one embodiment of the invention, an apparatus for efficiently performing spatial scalable compression of an input video stream is disclosed. A base encoder encodes a base encoder stream. Modifying means modifies content of the base encoder stream to create a plurality of base streams. An enhancement encoder encodes an enhancement encoder stream. Modifying means modifies content of the enhancement encoder stream to create a plurality of enhancement streams.

[0010] According to another embodiment of the invention, a method and apparatus for providing spatial scalable compression of an input video stream is disclosed. The input video stream is downsampled to reduce the resolution of the video stream. The downsampled video stream is encoded to produce a base encoder stream. A plurality of base streams are created from the base encoder stream. The base encoder stream is decoded and upconverted to produce a reconstructed video stream. The expected motion between frames from the input video stream and the reconstructed video stream is estimated and motion vectors for each frame of the received streams is calculated based upon an upscaled base layer plus enhancement layer. The reconstructed video stream is subtracted from the video stream to produce a residual stream. A predicted stream is calculated using the motion vectors in a motion compensation unit. The predicted stream is subtracted from the residual stream. The resulting residual stream is encoded and an enhancement encoder stream is outputted. A plurality of enhancement streams are created from the enhancement encoder stream.

[0011] According to another embodiment of the invention, a method and apparatus for decoding a plurality of coded video signals is disclosed. Each of the video streams is decoded and then the video streams are combined. An

inverse quantization operation is performed on quantization coefficients in the decoded video streams to produce DCT coefficients. An inverse DCT operation is performed on the DCT coefficients to produce a first signal. Predicted pictures are produced in a motion compensator and the first signal and the predicted pictures are combined to produce an output signal.

[0012] These and other aspects of the invention will be apparent from and elucidated with reference to the embodiments described hereafter.

BRIEF DESCRIPTION OF THE DRAWINGS

[0013] The invention will now be described, by way of example, with reference to the accompanying drawings, wherein:

[0014] **FIG. 1** is a block schematic representation of a known encoder with spatial scalability;

[0015] **FIG. 2** is a block schematic representation of a known encoder with spatial scalability;

[0016] **FIG. 3** is a block schematic representation of an encoder with spatial scalability according to one embodiment of the invention;

[0017] **FIG. 4** illustrates a modifying device with attenuators in series according to one embodiment of the invention;

[0018] **FIG. 5** illustrates a modifying device with attenuators in cascade according to one embodiment of the invention; and

[0019] **FIG. 6** illustrates a decoder according to one embodiment of the invention.

DETAILED DESCRIPTION OF THE INVENTION

[0020] **FIG. 3** is a schematic diagram of an encoder according to one embodiment of the invention. The depicted encoding system **300** accomplishes layered compression, whereby a portion of the channel is used for providing a plurality of lower resolution base layers and the remaining portion is used for transmitting a plurality of enhancement layers, whereby various base layers and base and enhancement layers can be combined to create video streams of differing quality levels. It will be understood by those skilled in the art that other encoding arrangements can also be used to create multilayered base and enhancement video streams and the invention is not limited thereto.

[0021] The encoder **300** comprises a base encoder **312** and an enhancement encoder **314**. The base encoder is comprised of a low pass filter and downsampler **320**, a motion estimator **322**, a motion compensator **324**, an orthogonal transform (e.g., Discrete Cosine Transform (DCT)) circuit **330**, a quantizer **332**, a variable length coder (VLC) **334**, a bitrate control circuit **335**, an inverse quantizer **338**, an inverse transform circuit **340**, switches **328**, **344**, and an interpolate and upsample circuit **350**.

[0022] An input video block **316** is split by a splitter **318** and sent to both the base encoder **312** and the enhancement encoder **314**. In the base encoder **312**, the input block is inputted into a low pass filter and downsampler **320**. The low pass filter reduces the resolution of the video block which is

then fed to the motion estimator **322**. The motion estimator **322** processes picture data of each frame as an I-picture, a P-picture, or as a B-picture. Each of the pictures of the sequentially entered frames is processed as one of the I-, P-, or B-pictures in a pre-set manner, such as in the sequence of I, B, P, B, P, . . . , B, P. That is, the motion estimator **322** refers to a pre-set reference frame in a series of pictures stored in a frame memory not illustrated and detects the motion vector of a macro-block, that is, a small block of 16 pixels by 16 lines of the frame being encoded by pattern matching (block Matching) between the macro-block and the reference frame for detecting the motion vector of the macro-block.

[0023] In MPEG, there are four picture prediction modes, that is an intra-coding (intra-frame coding), a forward predictive coding, a backward predictive coding, and a bi-directional predictive-coding. An I-picture is an intra-coded picture, a P-picture is an intra-coded or forward predictive coded or backward predictive coded picture, and a B-picture is an intra-coded, a forward predictive coded, or a bi-directional predictive-coded picture.

[0024] The motion estimator **322** performs forward prediction on a P-picture to detect its motion vector. Additionally, the motion estimator **322** performs forward prediction, backward prediction, and bi-directional prediction for a B-picture to detect the respective motion vectors. In a known manner, the motion estimator **322** searches, in the frame memory, for a block of pixels which most resembles the current input block of pixels. Various search algorithms are known in the art. They are generally based on evaluating the mean absolute difference (MAD) or the mean square error (MSE) between the pixels of the current input block and those of the candidate block. The candidate block having the least MAD or MSE is then selected to be the motion-compensated prediction block. Its relative location with respect to the location of the current input block is the motion vector.

[0025] Upon receiving the prediction mode and the motion vector from the motion estimator **322**, the motion compensator **324** may read out encoded and already locally decoded picture data stored in the frame memory in accordance with the prediction mode and the motion vector and may supply the read-out data as a prediction picture to arithmetic unit **325** and switch **344**. The arithmetic unit **325** also receives the input block and calculates the difference between the input block and the prediction picture from the motion compensator **324**. The difference value is then supplied to the DCT circuit **330**.

[0026] If only the prediction mode is received from the motion estimator **322**, that is, if the prediction mode is the intra-coding mode, the motion compensator **324** may not output a prediction picture. In such a situation, the arithmetic unit **325** may not perform the above-described processing, but instead may directly output the input block to the DCT circuit **330**.

[0027] The DCT circuit **330** performs DCT processing on the output signal from the arithmetic unit **33** so as to obtain DCT coefficients which are supplied to a quantizer **332**. The quantizer **332** sets a quantization step (quantization scale) in accordance with the data storage quantity in a buffer (not illustrated) received as a feedback and quantizes the DCT coefficients from the DCT circuit **330** using the quantization

step. The quantized DCT coefficients are supplied to the VLC unit **334** along with the set quantization step.

[0028] The VLC unit **334** converts the quantization coefficients supplied from the quantizer **332** into a variable length code, such as a Huffman code, in accordance with the quantization step supplied from the quantizer **332**. The resulting converted quantization coefficients are outputted to a buffer not illustrated. The quantization coefficients and the quantization step are also supplied to an inverse quantizer **338** which dequantizes the quantization coefficients in accordance with the quantization step so as to convert the same to DCT coefficients. The DCT coefficients are supplied to the inverse DCT unit **340** which performs inverse DCT on the DCT coefficients. The obtained inverse DCT coefficients are then supplied to the arithmetic unit **348**.

[0029] The arithmetic unit **348** receives the inverse DCT coefficients from the inverse DCT unit **340** and the data from the motion compensator **324** depending on the location of switch **344**. The arithmetic unit **348** sums the signal (prediction residuals) from the inverse DCT unit **340** to the predicted picture from the motion compensator **324** to locally decode the original picture. However, if the prediction mode indicates intra-coding, the output of the inverse DCT unit **340** may be directly fed to the frame memory. The decoded picture obtained by the arithmetic unit **340** is sent to and stored in the frame memory so as to be used later as a reference picture for an inter-coded picture, forward predictive coded picture, backward predictive coded picture, or a bi-directional predictive coded picture.

[0030] The quantization coefficients from the quantizer **332** are also applied to a modifying means **400**. The modifying device **400** comprises a plurality of attenuation steps which can be arranged in series as illustrated in FIG. 4 or in cascade or parallel as illustrated in FIG. 5. As illustrated in FIG. 4, the quantization coefficients from the quantizer **332** are applied to an attenuator **401**. The signal is then attenuated by the attenuator **401** which results in attenuated DCT coefficients carried by a signal **407**. In series with the attenuator **401**, a second attenuator **403** attenuates the amplitude of the DCT coefficients carried by the signal **407** and delivers new attenuated coefficients carried by signal **413**, that are variable length coded by a variable length coder **422** for generating a first base video stream BaseBase0.

[0031] The attenuators **401** and **403** are composed of an inverse quantizer **402** and **408**, respectively, a weighting device **404** and **410**, respectively, followed in series by a quantizer **406** and **412**, respectively. The quantization coefficients from the quantizer **332** are inverse quantized by the inverse quantizer **402**. The weighting is performed by a 8*8 weighting matrix multiplied to DCT blocks, each DCT coefficient being thus multiplied by a weighting factor contained in the matrix, the results of each multiplication being rounded to the nearest integer, weighting matrix being filled by values which amplitude are between 0 and 1, set for example to non-uniform values close to 1 for low frequential values and close to 0 for high frequential values, or to uniform values so that all coefficients in the 8*8 DCT block are equally attenuated. The quantization step consists of dividing weighted DCT coefficients by a new quantization factor for delivering quantized DCT coefficients, said quantization factor being the same for all coefficients of all 8*8 blocks composing a macroblock.

[0032] The coding error **415** relative to the attenuator **401** is generated by subtracting signal **407** from a signal from the quantizer **332** by means of a subtraction unit **414**. The coding error **415** is then variable length coded by a variable length coder **416** for generating a base enhancement video stream BaseEnh2. The coding error **419** relative to the attenuator **403** is generated by subtracting a signal **413** from signal **407** by means of a subtraction unit **418**. The coding error **419** is then variable length coded by a variable length encoder **420** for generating a second base enhancement video stream BaseEnh1.

[0033] In this example, the minimum quality base resolution would be provided by the video stream BaseBase0. A medium quality base resolution would be provided by combining the video stream BaseBase0 with the video stream BaseEnh0. A high quality base resolution would be provided by combining the video stream BaseBase0, BaseEnh0 and BaseEnh1.

[0034] The enhancement encoder **314** comprises a motion estimator **354**, a motion compensator **356**, a DCT circuit **368**, a quantizer **370**, a VLC unit **372**, a bitrate controller **374**, an inverse quantizer **376**, an inverse DCT circuit **378**, switches **366** and **382**, subtractors **358** and **364**, and adders **380** and **388**. In addition, the enhancement encoder **314** may also include DC-offsets **360** and **384**, adder **362** and subtractor **386**. The operation of many of these components is similar to the operation of similar components in the base encoder **312** and will not be described in detail.

[0035] The output of the arithmetic unit **340** is also supplied to the upsampler **350** which generally reconstructs the filtered out resolution from the decoded video stream and provides a video data stream having substantially the same resolution as the high-resolution input. However, because of the filtering and losses resulting from the compression and decompression, certain errors are present in the reconstructed stream. The errors are determined in the subtraction unit **358** by subtracting the reconstructed high-resolution stream from the original, unmodified high resolution stream.

[0036] According to one embodiment of the invention illustrated in FIG. 3, the original unmodified high-resolution stream is also provided to the motion estimator **354**. The reconstructed high-resolution stream is also provided to an adder **388** which adds the output from the inverse DCT **378** (possibly modified by the output of the motion compensator **356** depending on the position of the switch **382**). The output of the adder **388** is supplied to the motion estimator **354**. As a result, the motion estimation is performed on the upscaled base layer plus the enhancement layer instead of the residual difference between the original high-resolution stream and the reconstructed high-resolution stream. This motion estimation produces motion vectors that track the actual motion better than the vectors produced by the known systems of FIGS. 1 and 2. This leads to a perceptually better picture quality especially for consumer applications which have lower bit rates than professional applications.

[0037] Furthermore, a DC-offset operation followed by a clipping operation can be introduced into the enhancement encoder **314**, wherein the DC-offset value **360** is added by adder **362** to the residual signal output from the subtraction unit **358**. This optional DC-offset and clipping operation allows the use of existing standards, e.g., MPEG, for the enhancement encoder where the pixel values are in a pre-

determined range, e.g., 0 . . . 255. The residual signal is normally concentrated around zero. By adding a DC-offset value **360**, the concentration of samples can be shifted to the middle of the range, e.g., 128 for 8 bit video samples. The advantage of this addition is that the standard components of the encoder for the enhancement layer can be used and result in a cost efficient (re-use of IP blocks) solution.

[0038] The various enhancement layer video streams are created in a similar manner as the creation of the multiple base video streams described above. The quantization coefficients from the quantizer **370** are also applied to the modifying device **450**. The modifying device **450** may have the same elements as the modifying device **400** illustrated in **FIG. 4**, and in the following description the same reference numerals will be used for like elements. The quantization coefficients from the quantizer **370** are applied to the attenuator **401**. The signal is then attenuated by the attenuator **401** which results in attenuated DCT coefficients carried by a signal **407**. In series with the attenuator **401**, a second attenuator **403** attenuates the amplitude of the DCT coefficients carried by the signal **407** and delivers new attenuated coefficients carried by signal **413**, that are variable length coded by a variable length coder **422** for generating a first enhancement video stream EnhBase0.

[0039] The attenuators **401** and **403** are composed of an inverse quantizer **402** and **408**, respectively, a weighting device **404** and **4410**, respectively, followed in series by a quantizer **406** and **412**, respectively. The weighting is performed by a 8*8 weighting matrix multiplied to DCT blocks, each DCT coefficient being thus multiplied by a weighting factor contained in the matrix, the results of each multiplication being rounded to the nearest integer, weighting matrix being filled by values which amplitude are between 0 and 1, set for example to non-uniform values close to 1 for low frequential values and close to 0 for high frequential values, or to uniform values so that all coefficients in the 8*8 DCT block are equally attenuated. The quantization step consists of dividing weighted DCT coefficients by a new quantization factor for delivering quantized DCT coefficients, said quantization factor being the same for all coefficients of all 8*8 blocks composing a macroblock.

[0040] The coding error **415** relative to the attenuator **401** is generated by subtracting signal **407** from a signal from the quantizer **370** by means of a subtraction unit **414**. The coding error **415** is then variable length coded by a variable length coder **416** for generating a second enhancement video stream EnhEnh2. The coding error **419** relative to the attenuator **403** is generated by subtracting a signal **413** from signal **407** by means of a subtraction unit **418**. The coding error **419** is then variable length coded by a variable length encoder **420** for generating a third base enhancement video stream EnhEnh1.

[0041] In this example, the minimum quality full resolution would be provided by adding the video stream EnhBase0 to the high quality base resolution video stream. A medium quality full resolution would be provided by combining the video streams EnhBase0 and EnhEnh1 with the high quality base resolution. A high quality full resolution would be provided by combining the video streams EnhBase0, EnhEnh1 and EnhEnh2 with the high quality base resolution.

[0042] **FIG. 5** illustrates a modifying device wherein the attenuators are connected in cascade or parallel. It will be

understood that the modifying device **500** can be used in both the base layer and the enhancement layer as a substitute for modifying devices **400** and **450**. The quantization coefficients from the quantizer **332** (or quantizer **370**) are supplied to the first attenuator **501**. The attenuator **501** comprises an inverse quantizer **502**, a weighting device **504** and a quantizer **506**. The quantization coefficients are inverse quantized in the inverse quantizer **502**, then weighted and requantized, as described above with respect to **FIG. 4**, in the weighting device **504** and the quantizer **506**. The attenuated DCT coefficients carried by a signal **513** are then coded in a variable length coder **514** to produce a first base (enhancement) stream.

[0043] The coding error **517** of the attenuator **501** is generated by subtracting the signal **517** from the signal from the quantizer **332** (quantizer **370**) by means of a subtraction unit **516**. The coding error is applied to the second attenuator **503** which is comprised of an inverse quantizer **508**, a weighting device **510** and a quantizer **512**. The attenuated signal **519** is encoded by a variable length coder **520** which produces a second base (or enhancement) stream. The coding error **523** of the attenuator **503** is generated by subtracting the signal **519** from the signal **517** by means of a subtraction unit **522**. The coding error **523** is encoded by a variable length coder **524** which produces a third base (enhancement) stream.

[0044] **FIG. 6** illustrates a decoder according to one embodiment of the invention for decoding the multiple base or enhancement streams produced by the modifying devices. The multiple base (enhancement) streams are decoded by a plurality of variable length decoders **602**, **604** and **606**. The decoded streams are then added together in an arithmetic unit **608**. The decoded quantization coefficients in the combined stream are supplied to an inverse quantizer **610** which dequantizes the quantization coefficient in accordance with the quantization step so as to convert the quantization coefficients into DCT coefficients. The DCT coefficients are supplied to the inverse DCT unit **612** which performs inverse DCT on the DCT coefficients. The obtained inverse DCT coefficients are then supplied to the arithmetic unit **614**. The arithmetic unit **614** receives the inverse DCT coefficients from the inverse DCT unit **612** and data (produced in a known manner) from a motion compensator **616**. The arithmetic unit **614** sums the stream from the inverse DCT unit **612** to the predicted picture from the motion compensator **616** to produce the decoded base (or enhancement) stream. The decoded base and enhancement streams can be combined in a known manner to create the decoded video output.

[0045] It will be understood that the different embodiments of the invention are not limited to the exact order of the above-described steps as the timing of some steps can be interchanged without affecting the overall operation of the invention. Furthermore, the term "comprising" does not exclude other elements or steps, the terms "a" and "an" do not exclude a plurality and a single processor or other unit may fulfill the functions of several of the units or circuits recited in the claims.

1. An apparatus for efficiently performing spatial scalable compression of an input video stream, comprising:

a base encoder (312) for encoding a base encoder stream; means for modifying (400) content of the base encoder stream to create a plurality of base streams;

an enhancement encoder (314) for encoding an enhancement encoder stream; and

means for modifying (450) content of the enhancement encoder stream to create a plurality of enhancement streams.

2. The apparatus according to claim 1, wherein said modifying is performed by a set of attenuation steps (401, 403) applied to coefficients composing said base encoder stream being assembled in series and a re-encoding step associated to each of said attenuation steps for delivering one of said plurality of base streams from a coding error by each attenuation step.

3. The apparatus according to claim 1, wherein said modifying is performed by a set of attenuation steps (401, 403) applied to coefficients composing said enhancement encoder stream being assembled in series and a re-encoding step associated to each of said attenuation steps for delivering one of said plurality of enhancement streams from a coding error by each attenuation step.

4. The apparatus according to claim 1, wherein said modifying is performed by a set of attenuation steps (501, 503) applied to coefficients composing said base encoder stream being assembled in cascade and a re-encoding step associated to each of said attenuation steps for delivering one of said plurality of base streams from a coding error by each attenuation step.

5. The apparatus according to claim 1, wherein said modifying is performed by a set of attenuation steps (501, 503) applied to coefficients composing said enhancement encoder stream being assembled in cascade and a re-encoding step associated to each of said attenuation steps for delivering one of said plurality of enhancement streams from a coding error by each attenuation step.

6. A layered encoder for encoding an input video stream, comprising:

a downsampling unit (320) for reducing the resolution of the video stream;

a base encoder (312) for encoding a base encoder stream; means for creating (400) a plurality of base streams by modifying content of the base encoder stream;

an upconverting unit (350) for decoding and increasing the resolution of the base encoder stream to produce a reconstructed video stream;

a motion estimation unit (354) which receives the input video stream and the reconstructed video stream and calculates motion vectors for each frame of the received streams based upon an upscaled base layer plus enhancement layer;

a first subtraction unit (362) for subtracting the reconstructed video stream from the input video stream to produce a residual stream;

a motion compensation unit (356) which receives the motion vectors from the motion estimation unit and produces a predicted stream;

a second subtraction unit (364) for subtracting the predicted stream from the residual stream;

an enhancement encoder (314) for encoding the resulting stream from the subtraction unit and outputting an enhancement encoder stream;

means for creating (450) a plurality of enhancement streams by modifying content of the enhancement encoder stream.

7. The layered encoder according to claim 6, wherein said means for creating a plurality of base streams comprises:

a set of attenuation means (401, 403) applied to coefficients composing the base encoder stream, said attenuation means being assembled in series for delivering one of said plurality of base streams;

re-encoding means (416, 420, 422) associated with each attenuation means for delivering one of said plurality of base streams, from a coding error generated by each attenuation means.

8. The layered encoder according to claim 6, wherein said means for creating a plurality of base streams comprises:

a set of attenuation means (501, 503) applied to coefficients composing the base encoder stream, said attenuation means being assembled in cascade for delivering one of said plurality of base streams;

re-encoding means (514, 520, 524) associated with each attenuation means for delivering one of said plurality of base streams, from a coding error generated by each attenuation means.

9. The layered encoder according to claim 7, wherein said means for creating a plurality of enhancement streams comprises:

a set of attenuation means (401, 403) applied to coefficients composing the enhancement encoder stream, said attenuation means being assembled in series for delivering one of said plurality of enhancement streams;

re-encoding means (416, 420, 422) associated with each attenuation means for delivering one of said plurality of enhancement streams, from a coding error generated by each attenuation means.

10. The layered encoder according to claim 8, wherein said means for creating a plurality of enhancement streams comprises:

a set of attenuation means (501, 503) applied to coefficients composing the enhancement encoder stream, said attenuation means being assembled in cascade for delivering one of said plurality of enhancement streams;

re-encoding means (514, 520, 524) associated with each attenuation means for delivering one of said plurality of enhancement streams, from a coding error generated by each attenuation means.

11. The layered encoder according to claim 7, wherein the attenuation means comprises frequential weighting means (404, 410) followed in series by quantization means (406, 412) for quantizing the coefficients, performed at the block level.

12. The layered encoder according to claim 7, wherein each re-encoding means comprises subtracting means (414,

418) for subtracting an output signal from an input signal of the associated attenuation means for delivering the coding error, and variable length coding means for creating one of said base streams from the coding error.

13. The layered encoder according to claim 8, wherein the attenuation means comprises frequential weighting means (**504, 510**) followed in series by quantization means (**506, 512**) for quantizing the coefficients, performed at the block level.

14. The layered encoder according to claim 13, wherein each re-encoding means comprises subtracting means (**516, 522**) for subtracting an output signal from an input signal of the associated attenuation means for delivering the coding error, and variable length coding means for creating one of said base streams from the coding error.

15. The layered encoder according to claim 9, wherein the attenuation means comprises frequential weighting means (**404, 410**) followed in series by quantization means (**406, 412**) for quantizing the coefficients, performed at the block level.

16. The layered encoder according to claim 7, wherein each re-encoding means comprises subtracting means (**414, 418**) for subtracting an output signal from an input signal of the associated attenuation means for delivering the coding error, and variable length coding means for creating one of said enhancement streams from the coding error.

17. A method for providing spatial scalable compression of an input video stream, comprising the steps of:

downsampling the input video stream to reduce the resolution of the video stream;

encoding the downsampled video stream to produce a base encoder stream;

creating a plurality of base streams by modifying content of the base encoder stream;

decoding and upconverting the base stream to produce a reconstructed video stream;

estimating the expected motion between frames from the input video stream and the reconstructed video stream and calculating motion vectors for each frame of the received streams based upon an upscaled base layer plus enhancement layer;

subtracting the reconstructed video stream from the video stream to produce a residual stream;

calculating a predicted stream using the motion vectors in a motion compensation unit;

subtracting the predicted stream from the residual stream; encoding the resulting residual stream and outputting an enhancement encoder stream; and

creating a plurality of enhancement streams by modifying content of the enhancement encoder stream.

18. A decoder for decoding a plurality of coded video signals, comprising:

a plurality of decoders (**602, 604, 606**), one for each video stream, for decoding said video streams;

arithmetic unit (**608**) for combining said decoded video streams;

inverse quantization means (**610**) for performing an inverse quantization operation on quantization coefficients in said decoded video streams to produce DCT coefficients;

inverse DCT means (**612**) for performing an inverse DCT operation on the DCT coefficients to produce a first signal;

a motion compensation unit (**616**) for producing predicted pictures;

arithmetic unit (**614**) for combining the first signal and the predicted pictures to produce an output signal.

19. The decoder according to claim 18, wherein the plurality of coded video streams are base streams.

20. The decoder according to claim 18, wherein the plurality of video streams are enhancement streams.

21. A method for decoding a plurality of coded video signals, comprising:

decoding each of said video streams;

combining said decoded video streams;

performing an inverse quantization operation on quantization coefficients in said decoded video streams to produce DCT coefficients;

performing an inverse DCT operation on the DCT coefficients to produce a first signal;

producing predicted pictures in a motion compensator;

combining the first signal and the predicted pictures to produce an output signal.

* * * * *