

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

(19) World Intellectual Property Organization
International Bureau



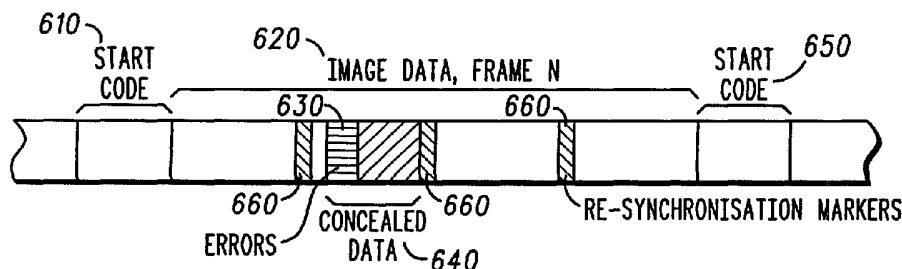
(43) International Publication Date
15 May 2003 (15.05.2003)

PCT

(10) International Publication Number
WO 03/041382 A2

- (51) International Patent Classification⁷: **H04N**
- (21) International Application Number: PCT/EP02/12323
- (22) International Filing Date:
4 November 2002 (04.11.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
0126747.5 7 November 2001 (07.11.2001) GB
- (71) Applicant (for all designated States except US): **MOTOROLA INC** [US/US]; 1303 E.Algonquin Road, Schaumburg, IL 60196 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **HARE, Jonathan** [GB/GB]; 2a Farrington Way, Tadley, Hampshire RG26 3UA (GB). **DOLBEAR, Catherine** [GB/GB]; 6 Theobald Drive, Reading, Berkshire RG31 6YA (GB). **SHAN-ABLEH, Tamer** [GB/GB]; 44 Vitellius Gardens, Basingstoke, Hampshire RG24 9SG (GB).
- (74) Agent: **MCCORMACK, Derek**; Motorola European Intellectual, Property Operations, Midpoint, Alencon Link, Basingstoke, Hampshire RG212 7PL (GB).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:**
— without international search report and to be republished upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: SCALABLE VIDEO TRANSMISSIONS



(57) Abstract: A method for improving a quality of MPEG-4, or similar, scalable video enhancement layers transmitted over an error-prone network. The method includes the steps of inserting one or more re-synchronisation marker (660, 760, 860) into one or more enhancement layers of a scalable video sequence. When one or more errors (630, 730, 830) is detected in said one or more enhancement layers of said scalable video sequence, data is concealed (640, 740, 840) in the video sequence from a video bit position at which said one or more errors is detected, to a video bit position of a subsequent error free re-synchronisation marker. In this manner, a reduction in the amount of error-free data that has to be discarded can be achieved. This, in turn, reduces the amount of visual degradation in the video transmission.

WO 03/041382 A2

Scalable Video TransmissionsField of the Invention

This invention relates to video transmission systems and
5 video encoding/decoding techniques. The invention is
applicable to a video compression system where the video
has been compressed using a scalable compression technique.

Background of the Invention

10 In the field of video technology, it is known that video is
transmitted as a series of still images/pictures. Since
the quality of a video signal can be affected during coding
or compression of the video signal, it is known to include
additional information or 'layers' based on the difference
15 between the video signal and the encoded video bit stream.
The inclusion of additional layers enables the quality of
the received signal, following decoding and/or
decompression, to be enhanced. Hence, a hierarchy of base
pictures and enhancement pictures, partitioned into one or
20 more layers, is used to produce a layered video bit stream.

A scalable video bit-stream refers to the ability to
transmit and receive video signals of more than one
resolution and/or quality simultaneously. A scalable video
25 bit-stream is one that may be decoded at different rates,
according to the bandwidth available at the decoder. This
enables the user with access to a higher bandwidth channel
to decode high quality video, whilst a lower bandwidth user
is still able to view the same video, albeit at a lower
30 quality. The main application for scalable video
transmissions is for systems where multiple decoders with
access to differing bandwidths are receiving images from a
single encoder.

Scalable video transmissions can also be used for bit-rate adaptability where the available bit rate is fluctuating in time. Other applications include video multicasting to a
5 number of end-systems with different network and/or device characteristics. More importantly, scalable video can also be used to provide subscribers of a particular service with different video qualities depending on their tariffs and preferences. Therefore, in these applications it is
10 imperative to protect the enhancement layer from transmission errors. Otherwise, the subscribers may lose confidence in their network operator's ability to provide an acceptable service.

15 In a layered (scalable) video bit stream, enhancements to the video signal may be added to a base layer either by:
(i) Increasing the resolution of the picture (spatial scalability);
(ii) Including error information to improve the Signal
20 to Noise Ratio of the picture (SNR scalability);
(iii) Including extra pictures to increase the frame rate (temporal scalability); or
(iv) Providing a continuous enhancement that may be truncated at any chosen bit rate (Fine Granular
25 Scalability).

Such enhancements may be applied to the whole picture or to an arbitrarily shaped object within the picture, which is termed object-based scalability.

30

In order to preserve the disposable nature of the temporal enhancement layer, the H.263+ ITU H.263 [ITU-T Recommendation, H.263, "Video Coding for Low Bit Rate

Communication"] standard dictates that pictures included in the temporal scalability mode should be bi-directionally predicted (B) pictures. These are as shown in the video stream of FIG. 1.

5

FIG. 1 shows a schematic illustration of a scalable video arrangement 100 illustrating B picture prediction dependencies, as known in the field of video coding techniques. An initial intra-coded frame (I_1) 110 is followed by a bi-directionally predicted frame (B_2) 120. This, in turn, is followed by a (uni-directional) predicted frame (P_3) 130, and again followed by a second bi-directionally predicted frame (B_4) 140. This again, in turn, is followed by a (uni-directional) predicted frame 15 (P_5) 150, and so on.

As an enhancement to the arrangement of FIG. 1, a layered video bit stream may be used. FIG. 2 is a schematic illustration of a layered video arrangement, known in the 20 field of video coding techniques. A layered video bit stream includes a base layer 205 and one or more enhancement layers 235.

The base layer (layer-1) includes one or more intra-coded 25 pictures (I pictures) 210 sampled, coded and/or compressed from the original video signal pictures. Furthermore, the base layer will include a plurality of subsequent predicted inter-coded pictures (P pictures) 220, 230 predicted from the intra-coded picture(s) 210.

30

In the enhancement layers (layer-2 or layer-3 or higher layer(s)) 235, three types of picture may be used:

- (i) Bi-directionally predicted (B) pictures (not shown);
- (ii) Enhanced intra-coded (EI) pictures 240 based on the
5 intra-coded picture(s) 210 of the base layer 205; and
- (iii) Enhanced predicted (EP) pictures 250, 260, based on
the inter-coded predicted pictures 220, 230 of the base
layer 205.

10 The vertical arrows from the lower, base layer illustrate that the picture in the enhancement layer is predicted from a reconstructed approximation of that picture in the reference (lower) layer.

15 If prediction is only formed from the lower layer, then the enhancement layer picture is referred to as an EI picture. It is possible, however, to create a modified bi-directionally predicted picture using both a prior enhancement layer picture and a temporally simultaneous
20 lower layer reference picture. This type of picture is referred to as an EP picture or "Enhancement" P-picture.

The prediction flow for EI and EP pictures is shown in FIG. 2. Although not specifically shown in FIG. 2, an EI picture
25 in an enhancement layer may have a P picture as its lower layer reference picture, and an EP picture may have an I picture as its lower-layer enhancement picture.

For both EI and EP pictures, the prediction from the
30 reference layer uses no motion vectors. However, as with normal P pictures, EP pictures use motion vectors when predicting from their temporally, prior-reference picture in the same layer.

Current standards incorporating the aforementioned scalability techniques include MPEG-4 and H.263. These standards create highly compressed bit-streams, which
5 represent the coded video. However, due to this high compression, the bit-streams are very prone to corruption by network errors as they are transmitted. For example, in the case of streaming video over an error prone network, even with existing network level error protection tools
10 employed, it is inevitable that some bit-level corruption will occur in the bit-stream and be passed on to the decoder.

To counter these bit-level errors, the coding standards
15 have been designed with various tools incorporated that allow the decoder to cope with the errors. These tools enable the decoder to localise and conceal the errors within the bit-stream.

20 However, the current MPEG-4 standard does not allow the use of these error resilience tools within the scalable enhancement layers.

The MPEG-4 standard does define three tools for error
25 resilience of video bit-streams. These are re-synchronisation markers, data partitioning (DP) and reversible variable length codes (RVLCs). These tools are defined for use in the base layer.

30 Data Partitioning (DP) separates motion parameters from the texture information, i.e. discrete cosine transform (DCT) coefficients, within a video transmission. Data partitioning is achieved by separating the motion and

macroblock header information away from the texture information. This approach requires that a second re-synchronisation marker be inserted between motion and texture information. The use of data partitioning, in the same manner as the use of RVLCs, is signalled to the decoder in the Video Object Layer's (VOL's) header. If the texture information is lost, this approach utilises some motion information to conceal these errors. That is, due to the errors the texture information is discarded, while the motion is used to motion compensate the previously decoded Video Object Plane (VOP). This separation is performed on a Video Packet basis where an additional re-synchronisation marker called a Motion Marker separates the motion parameters from the DCT coefficients. If an error occurs in the DCT coefficients portion, then the correctly decoded motion vectors are employed to conceal the remaining portion of the Video Packet.

The MPEG-4 visual standard associates another useful tool with data partitioning called Reversible Variable Length Codes (RVLCs). It is noteworthy that, according to the MPEG-4 video standard, RVLC can only be used in conjunction with data partitioning within the base layer. Here the DCT coefficients are coded with RVLCs rather than the traditional Variable Length Codes (VLCs). As such, parts of a bit stream, which cannot be decoded in the forward direction due to the presence of errors, can often be decoded in the reverse or backward direction. Hence, the number of discarded bits can be reduced.

30

It is worth re-iterating that the MPEG-4 bit stream syntax and the Simple Scalable Profile do NOT currently allow the use of error resiliency tools in the Enhancement layer.

This was because enhancement layers were considered as an error resilience tool. Enhancement layer information contains visual information that enhances the decoding quality of the more important base layer. Hence, as
5 enhancement layer information was not deemed essential, no further resiliency was anticipated.

Thus, the focus for higher levels of protection in a video bit sequence in current communications systems is the base
10 layer. This means that when an error occurs in an enhancement layer bit-stream, the decoder, wishing to keep the enhancement layer, has to conceal much more data, potentially in error, than it would have to if the error resilience tools could be used.

15

In summary, there exists a need in the field of video communications, and in particular in scalable video communications, for an apparatus and a method for improving the quality of scalable video enhancement layers
20 transmitted over an error-prone network, where the abovementioned disadvantages with prior art arrangements may be alleviated.

Known prior art documents include:

- 25 (i) US-A-200221761 (Zhang et al.), published 21 February 2002;
- (ii) 'Error Resilience Methods for FGS Video Enhancement Bitstream', R. Yan et al, The First Pacific Rim Conference on Multimedia (IEEE-PCM 2000), Dec 13-15, 2000;
- 30 (iii) 'Robust Video Coding Algorithms and Systems', J.D. Villasenor, et al., Proceedings of the IEEE, Vol. 87, No. 10, October 1999.

Statement of Invention

The present invention provides a method for improving a quality of scalable video enhancement layers transmitted over an error-prone network, as claimed in claim 1, a
5 method for improving a quality of scalable video enhancement layers transmitted over an error-prone network, as claimed in claim 2, a method for improving a quality of scalable video enhancement layers transmitted over an error-prone network, as claimed in claim 6, a video
10 communication system, as claimed in claim 10, a fifth aspect of the present invention, there is provided a video communication system, as claimed in claim 13, a video communication unit, as claimed in claim 16, a video encoder or a video decoder, as claimed in claim 17, and a
15 mobile radio device, as claimed in claim 19.

Further aspects of the present invention are as claimed in the dependent claims.

20 In summary, an apparatus and methods for improving the quality of scalable video enhancement layers transmitted over an error-prone network by the use of re-synchronisation markers are described. This invention provides a method by which re-synchronisation markers are
25 added to enhancement layer bit-streams in scalable MPEG-4 or similar video compression systems, or H.263+ or similar video compression systems. This is in order to promote better bit-level error resilience capabilities, and reduce the amount of concealment of error-free data that takes
30 place.

In addition, or in the alternative, this invention proposes the use of the data partitioning, and preferably a

Reversible Variable Length Coding, error resilience tool in the video enhancement layers of the MPEG-4 Simple Scalable Profile. One aim is to reduce or to substantially eliminate the macroblock displacements, caused by error concealment
5 utilising estimated motion vectors. A further aim is to minimise the number of discarded bits by performing reverse decoding.

Brief Description of the Drawings

- 10 FIG. 1 is a schematic illustration of a video coding arrangement showing picture prediction dependencies, as known in the field of video coding techniques.
FIG. 2 is a schematic illustration of a known layered video coding arrangement.
- 15 Exemplary embodiments of the present invention will now be described, with reference to the accompanying drawings, in which:
- FIG. 3 is a schematic representation of a scalable video communication system adapted to use re-synchronisation
20 markers in an enhancement layer of a video sequence in accordance with the preferred embodiment of the present invention.
- FIG. 4 illustrates a first enhancement layer bit-stream frame configuration.
- 25 FIG. 5 illustrates a second enhancement layer bit-stream frame configuration.
- FIG. 6 illustrates a first enhancement layer bit-stream frame configuration adapted to incorporate the preferred embodiment of the present invention.
- 30 FIG. 7 illustrates a second enhancement layer bit-stream frame configuration adapted to incorporate the preferred embodiment of the present invention.

FIG. 8 illustrates a third enhancement layer bit-stream frame configuration adapted to incorporate the preferred embodiment of the present invention.

FIG. 9 illustrates an example of syntax code for use in the
5 MPEG-4 standard, for implementing a preferred embodiment of the present invention.

FIG. 10 illustrates a graph highlighting the improvement of decoded SNR scalable video quality, in accordance with a preferred embodiment of the present invention.

10 FIG. 11 illustrates a graph highlighting the improvement of decoded temporal scalable video quality, in accordance with a preferred embodiment of the present invention.

FIG. 12 illustrates a data format of an enhancement layer bit stream that can be adapted to incorporate the use of
15 data partitioning and RVLCs in accordance with the preferred further embodiment of the present invention.

FIG. 13 illustrates at least one benefit from introducing data partitioning, and preferably reverse variable length coding, in an enhancement layer bit stream, for example the
20 data format of FIG. 12.

FIG. 14 illustrates the objective degradations caused by estimating values of lost motion vectors, or using original motion vectors, associated with the error scenarios of FIG. 13.

25 FIG. 15 illustrates proposed syntax amendments to section 6.2.5 "Video Object Plane and Video Plane with Short Header" of the MPEG-4 standard.

Description of Preferred Embodiments

30 This invention applies to SNR scalable encoded video. However, it is within the contemplation of the invention that the inventive concepts described herein can be applied to other types of scalable video sequences such as temporal

scalability, spatial scalability and Fine Granular scalability (FGS).

The inventive concepts herein described find particular
5 application in the current MPEG technology arena, and in future versions of MPEG technology. However, the inventors have also recognised the applicability of the present invention to H.263+ technology.

10 The preferred embodiment of the present invention illustrates a technique for introducing re-synchronisation markers in enhancement layer bit-streams to improve their error resilience.

15 In the example context of MPEG-4 video, the video bit-stream syntax is modified to enable the use of the standard base layer error resilience tools within the enhancement layer bit-stream. Advantageously, the decoder can then make use of the re-synchronisation markers in the
20 enhancement layer bit-stream in the same manner as for a base layer bit-stream.

An analogous technique can be used with H263+ video.

25 Referring first to FIG. 3, a schematic representation of a video communication system 300, including video encoder 315 and video decoder 325, adapted to incorporate the preferred embodiment of the present invention, is shown.

30 In FIG. 3, a video picture F_0 is compressed 310 in a video encoder 315 to produce the base layer bit stream signal to be transmitted at a rate r_1 kilobits per second (kbps).

This signal is decompressed 320 at a video decoder 325 to produce the reconstructed base layer picture F_0' .

The compressed base layer bit stream is also decompressed
5 at 330 in the video encoder 315 and compared with the original picture F_0 at 340 to produce a difference signal 350. This difference signal is compressed at 360 and transmitted as the enhancement layer bit stream at a rate r_2 kbps. This enhancement layer bit stream is decompressed at
10 370 in the video decoder 325 to produce the enhancement layer picture F_0'' which is added to the reconstructed base layer picture F_0' at 380 to produce the final reconstructed picture F_0''' .

15 In accordance with the preferred embodiment of the present invention, the compression function 360 in the video encoder 315 has been adapted to incorporate one or more re-synchronisation markers in the enhancement layer bit-stream. Furthermore, the decompression function 370 in the
20 video decoder 325 has been adapted to determine the location of these one or more re-synchronisation markers in the enhancement layer bit-stream. The video decoder 325 has also been adapted to conceal errors in the received scalable video sequence up to the next available error-free
25 re-synchronisation marker. The use of such re-synchronisation markers is further described with regard to FIG's 6 to 8.

It is within the contemplation of the invention that
30 alternative encoding and decoding configurations could be adapted to use re-synchronisation markers within the

enhancement layer bit-stream. As a result, the inventive concepts hereinafter described should not be viewed as being limited to the example configuration provided in FIG. 3.

5

As discussed above, in prior art arrangements, an MPEG-4 enhancement layer bit-stream does not have the facility to include any re-synchronisation markers within it. Hence, if the video decoder encounters an error, it has no choice but to conceal all of the data from the point at which the error occurred to the start of the next readable start code (in the case of MPEG-4 this would be a VOP_Start_Code) at the beginning of a subsequent frame of video.

10 The problem with this prior art arrangement is that the actual amount of error-free data concealed could be very large, depending on the position of the error. This is illustrated in FIG. 4 and FIG. 5.

15 Referring now to FIG. 4, a first enhancement layer bit-stream frame 400 is illustrated. The first enhancement layer bit-stream frame 400 includes a start code 410, followed by image data of a frame 'N' 420. The image data of frame 'N' 420 is shown to contain errors 430 within the frame 420. FIG. 4 highlights a problem associated with such an enhancement layer bit-stream frame in that a substantial amount of error-free data 440 subsequent to the errors 430 within the frame 420 has to be concealed. Clearly, it is undesirable, as well as inefficient, to
20
25
30 conceal such error-free data.

Referring next to FIG. 5, a second enhancement layer bit-stream frame 500 is illustrated that highlights a further

problem with errors encountered in enhancement layer bit-streams. The second enhancement layer bit-stream frame 500 illustrates errors being encountered within image data of frame 'N' 520 as well as within the subsequent start code 510 of the next 'N+1' frame 525. When errors 530 occur over the start code in this manner, a whole frame 540 may end up being concealed, as shown.

Typically, the effect to the viewer of the aforementioned concealed error-free frames depends on the type of scalability being used. In the case of spatial or SNR scalability, where the enhancement layer increases the spatial resolution or quality of the base image, the viewer may witness frames that are enhanced at the top of the image, but not enhanced at the bottom. Alternatively, if the errors affect a start code, the viewer may experience an annoying flicker, as the enhancement layer disappears entirely over a frame. In the case of temporal scalability, the concealment can lead to some strange effects, with the viewer seeing part of the current frame and parts of previously decoded frames.

In accordance with the preferred embodiment of the present invention, re-synchronisation markers 660, 760, 860 are introduced to the enhancement layer bit-stream transmitted by the video encoder 315, to improve the quality of the decoded video image. The improvement is achieved by reducing the amount of error-free data that has to be concealed by the decoder. Instead of the video decoder 325 having to conceal all of the data in a frame, from the point the error occurred to the end of the frame, the concealment only has to continue up to the first occurrence of either:

- (i) The first valid re-synchronisation marker, bearing in mind that errors could occur in the markers themselves; or
- (ii) The next valid start code.

5

In this manner, the amount of error-free data concealment is reduced substantially.

Referring next to FIG. 6, a first enhancement layer bit-
10 stream frame configuration 600 is shown, adapted by the introduction of re-synchronisation markers 660 in accordance with the preferred embodiment of the present invention. The configuration is similar to that shown in FIG. 4, where the errors 630 occur within the image data
15 frame 620 and are determined to be located between two start codes 610, 650. As shown, by introducing re-synchronisation markers in the manner described above, the video decoder only conceals error-free data up to the re-synchronisation marker 660 that immediately follows the
20 errors 630. Hence, the amount of concealed error-free data 640 is reduced significantly, when compared to the prior art arrangement described with respect to FIG. 4.

Referring next to FIG. 7, a second enhancement layer bit-
25 stream frame configuration 700 is shown, adapted by the introduction of re-synchronisation markers 760 in accordance with the preferred embodiment of the present invention. The configuration is similar to that shown in FIG. 5, where the errors 730 occur across a start code 710,
30 as well as within the image data frame N 720. As shown, by introducing re-synchronisation markers in the manner described above, the video decoder only conceals error-free data up to the first re-synchronisation marker 760, located

in the subsequent image data frame 'N+1' 725. This is because the corrupted VOP header is replicated in the Header Extension Code (HEC) of the Video Packet Header and therefore the Video Object Plane's (VOP's) header can be restored. Hence, the amount of concealed error-free data is reduced significantly, when compared to the prior art arrangement described with respect to FIG. 5.

Referring next to FIG. 8, a third enhancement layer bit-stream frame configuration 800 is shown, adapted by the introduction of re-synchronisation markers 860 in accordance with the preferred embodiment of the present invention.

In FIG. 8, it is shown that the errors may occur around, and affecting, the re-synchronisation markers 860. This may be, for example, around the first re-synchronisation marker 860 within the image data frame N 820. As shown for this situation, the video decoder only conceals error-free data up to the second re-synchronisation marker 860, located within the same image data frame 'N' 820. Hence, the amount of concealed error-free data is still reduced significantly, and only extends into the period between re-synchronisation markers 860.

For completeness, it should also be noted that state of the art techniques for error resilience have been developed that allow decoding of video frames even if the frame header is corrupted by errors, provided re-synchronisation markers exist in the bit-stream. It is envisaged that such techniques can also be used to complement the performance of the present invention.

Referring now to FIG. 9, the existing MPEG-4 syntax code 900 can be readily adapted to incorporate the preferred embodiment of the present invention using re-synchronisation markers in one or more enhancement layers. The required
5 change 910 to the MPEG-4 syntax code, to accommodate the preferred embodiment of the present invention, is illustrated in FIG. 9. Thus, in accordance with the preferred embodiment of the invention, a processor in a video communication unit may be re-programmed, and thereby
10 adapted, to incorporate the inventive concepts hereinbefore described.

More generally, the adaptation may be implemented in the respective video communication unit in any suitable manner.
15 For example, new apparatus may be added to a conventional video communication unit, or alternatively existing parts of a conventional video communication unit may be adapted, for example by reprogramming one or more processors therein. As such, the required adaptation may be
20 implemented in the form of processor-implementable instructions stored on a storage medium, such as a floppy disk, hard disk, PROM, RAM or any combination of these or other storage multimedia.

25 It is also within the contemplation of the invention that such adaptation of a video encoding or video operations decoding may be facilitated by any video communication unit operating in a video communication system. This may be, for example, user equipment such as mobile or portable radios
30 or telephones, or wireless or wired serving communication units such as base transceiver or intermediate (repeater) equipment.

With the revised syntax code in place, re-synchronisation markers can be placed in the enhancement layer bit-stream, and used to reduce the amount of error-free data concealment. Based on Profiles and levels, the MPEG-4 standard states the maximum size of Video Packets, or the number of bits between two consecutive resynchronisation markers, in a given base layer VOP. This aids the decoder, as it roughly knows when to expect the next marker.

10 It is envisaged that the period between re-synchronisation markers in the enhancement layer may be user and/or implementation defined. A smaller distance between re-synchronisation markers leads to a reduction in the amount of data that has to be concealed, should an error occur.

15

However, the distance between the re-synchronisation markers also affects the overall number of bits in the video bit-streams. So, in a fixed bit-rate environment the quality of the enhancement layer has to be reduced, in order for the video to meet the bit budget of the transmission system.

25 The inventors of the present invention have performed extensive experiments that have shown that by placing re-synchronisation markers at approximately every eight hundred bits, the overall increase in bit-stream size is just less than four percent.

30 The current MPEG standard, both for scalable MPEG-4 and FGS MPEG-4, explicitly states that error resilience tools are not to be used in the enhancement layer(s). Hence, the standard, as well as public domain literature, do not allow the use of re-synchronisation markers in an enhancement

layer of scalable MPEG-4. In fact, the MPEG 4 standard, as well as public domain literature, teaches away from the present invention.

5 The current teaching also states that no foreseeable benefit can be gained in using error resilience tools in scalable (including FGS) MPEG-4 video communication systems. Furthermore, it is generally understood in the field of video communication that MPEG-4 enhancement layers
10 would not be used in a mobile communication environment.

Despite the prior art teaching away from the present invention, the results obtained by the inventors of the present invention, shown below in FIG. 10 and FIG. 11, have
15 shown that resynchronisation markers in the enhancement layer(s) are of great benefit to the use of MPEG-4 scalable video in a mobile environment. As such, it is envisaged that emerging 3rd generation 3G mobile communication systems will require scalable video to accommodate different bit
20 rates at different parts of the coverage area. It is further envisaged that future video technology, similar to today's current MPEG-4 standard, will similarly benefit from the inventive concepts described herein.

25 Referring now to FIG. 10, a graph 1000 highlights the improvement of decoded SNR scalable video quality, in implementing the preferred embodiment of the present invention. As clearly shown, the peak signal to noise ratio (PSNR) is substantially improved when re-
30 synchronisation markers are incorporated into the enhancement layer bit-stream, as compared to the PSNR levels without re-synchronisation markers.

Referring now to FIG. 11, a graph 1100 highlights the improvement of decoded temporal scalable video quality, when implementing the preferred embodiment of the present invention. As clearly shown, the peak signal to noise ratio (PSNR) is substantially improved when re-synchronisation markers are incorporated into the enhancement layer bit-stream, as compared to the PSNR levels without re-synchronisation markers.

10 The measurements of FIG. 10 and FIG. 11 were made by adding re-synchronisation markers to the enhancement layers of SNR and temporally scalable video bit-streams transmitted over a GPRS network. In particular, a noteworthy feature is the relative smoothness of the lines. For example, the lines
15 illustrating the video without re-synchronisation markers tend to oscillate much more because of the larger amounts of data concealment taking place. In contrast, the lines illustrating video with re-synchronisation markers are much flatter due to the reduced amount of data concealment that
20 has to take place.

Finally, the applicant notes that future versions of the MPEG communication standard, such as the Joint Video Team (JVT) (from MPEG-4 and H26L) configuration are currently
25 under development. The present invention is not limited to the MPEG-4 standard, and is envisaged by the inventors as applying to future versions of scalable video compression.

It is within the contemplation of the present invention
30 that the aforementioned inventive concepts may be applied to any video communication unit and/or video communication system. In particular, the inventive concepts find particular use in wireless (radio) devices, such as mobile

telephones/mobile radio units and associated wireless communication systems. Such wireless communication units may include a portable or mobile PMR radio, a personal digital assistant, a laptop computer or a wirelessly
5 networked PC.

Although the preferred embodiment of the present invention has been described with reference to the MPEG-4 standard, scalable video system technology may be implemented in the
10 3rd generation (3G) of digital cellular telephones, commonly referred to as the Universal Mobile Telecommunications Standard (UMTS). Scalable video system technology may also find applicability in the packet data variants of both the
15 current 2nd generation of cellular telephones, commonly referred to as the general packet-data radio system (GPRS), and the TERrestrial Trunked Radio (TETRA) standard for digital private and public mobile radio systems.
Furthermore, scalable video system technology may also be utilised in the Internet. The aforementioned inventive
20 concepts will therefore find applicability in, and thereby benefit, all these emerging technologies.

It will be understood that the video transmission arrangement described above provides at least the advantage
25 of achieving a reduction in the amount of error-free data that has to be discarded. This, in turn, reduces the amount of visual degradation in the video transmission.
This is especially important when transmitting video over wireless channels and the Internet, where the errors can be
30 severe. Applications of digital video where the preferred method improves decoder performance are one-way and two-way video communications, surveillance applications, and video streaming applications.

Method of the invention

Summarising the discussion above, a method for improving a quality of MPEG-4 (or similar) scalable video enhancement
5 layers transmitted over an error-prone network has been described. The method includes the steps of inserting one or more re-synchronisation markers into one or more enhancement layers of a scalable video sequence in a video encoder. A video decoder detects one or more errors in the
10 one or more enhancement layers of the scalable video sequence; and, in response to detecting one or more errors, conceals data in the received scalable video sequence from a video bit position at which the one or more errors is detected, to a video bit position of a subsequent error-
15 free re-synchronisation marker.

The inventors have furthermore realised that the present invention may also be applied to video encoding in accordance with the H.263+ standard. This is the method
20 claimed in appended claim 2. H.263+ video is generally used for lower bandwidth transmission systems than MPEG-4 video. The H.263+ standard actually differs from the MPEG-4 standard in several aspects. One of these aspects is that the H.263+ standard does not expressly state that error
25 resilience tools should not be used in the enhancement layers.

Apparatus of the invention

A video communication system has been described that
30 includes a video encoder comprising a processor for encoding a video sequence into an MPEG-4 (or similar) scalable video sequence having a plurality of enhancement layers; and insertion means to insert one or more re-

synchronisation markers into one or more enhancement layers. A transmitter transmits the scalable video sequence containing said one or more re-synchronisation markers. A video decoder includes a receiver for receiving
5 the scalable video sequence and a detector that detects one or more errors in the one or more enhancement layers of the scalable video sequence. A processor is operably coupled to the detector for concealing data in the received
scalable video sequence, when one or more errors are
10 detected. The video data is concealed from a video bit position at which said one or more errors is detected to a video bit position of a subsequent error-free re-synchronisation marker.

15 A video communication unit, an adapted video encoder, an adapted video decoder, and a mobile radio device incorporating any one of these units, have also been described.

20 Generally, the inventive concepts contained herein are equally applicable to any suitable video or image transmission system. Whilst specific, and preferred, implementations of the present invention are described above, it is clear that one skilled in the art could
25 readily apply variations and modifications of such inventive concepts.

Thus, an improved apparatus and methods for improving the quality of scalable video enhancement layers transmitted
30 over an error-prone network have been provided, whereby the aforementioned disadvantages with prior art arrangements have been substantially alleviated.

Enhanced approach

The inventors have also considered a further enhancement of the above invention. The enhanced solution may further improve the subjective quality of video for a viewer.

5

The inventors have considered the problem that error concealment techniques in the temporally scalable enhancement layer conceal data from the point that the error occurred to the next re-synchronisation marker, or the start code at the beginning of the next frame. If a large amount of motion occurred between the current frame and the previous frame, then the area that is concealed can look discontinuous from the rest of the frame.

10 Therefore, the inventors have developed a method to selectively conceal an entire frame of video, in order to reduce the discontinuities in the image, thus improving subjective quality for the viewer.

15 This method improves the image quality by adaptively deciding on how much data should be concealed, to give the best image quality. If the motion between the previous frame and the current frame with errors is large, then the method involves a decision to conceal the whole frame to avoid any discontinuity in the decoded image. If the motion is small, then the decision is to conceal as little data as possible.

20 A detailed arrangement starts from the realisation that discontinuities in the decoded image can be very visible to a viewer, using concealment techniques that conceal errors from the point the error occurred to the next re-

synchronisation point, i.e. re-synchronisation marker or start code.

The visibility of the discontinuity between the unconcealed
5 and concealed areas depends on the amount of motion that is
taking place in the scene. For example, visualise a scene
with little background motion, but people running in the
foreground. The people create lots of foreground motion,
which leads to a large discontinuity, which is visually
10 disturbing. In order to improve the quality of the video
for the viewer (subjective quality will increase, but
objective quality measures such as PSNR will decrease), the
inventors propose a method in which a normalised Sum of
Absolute Difference (SAD) is calculated between the current
15 frame and previous frame over the non-concealed area. If
the normalised SAD exceeds a pre-defined threshold
(indicating a large amount of motion) then the whole frame
is concealed.

20 The enhancement, using the frame suppression method, can in
fact even be used in a video transmission system that does
not have re-synchronisation markers in the enhancement
layers. So it can be seen both as an enhancement to the
claimed invention, and as a potential alternative solution
25 if employed alone.

The inventors of the present invention have further
recognised a potential problem in the concealment of
corrupted visual data by means of estimating the values of
30 lost motion vectors and motion compensating previous VOPs.
As lost motion vectors are estimated from either the
surrounding macroblocks, or the co-located macroblocks from
previously decoded VOPs, it follows that these motion

vectors are not accurate, resulting in non-accurate motion compensation. Hence, estimated motion vectors may often cause the displacement of concealed macroblocks.

5 Thus, the inventors of the present invention propose an enhancement to the preferred embodiment of the present invention described above. In this enhanced further embodiment of the present invention, it is proposed that data partitioning, currently limited to use in a base layer
10 of a video data stream, is also employed in video enhancement layers. In a preferred further embodiment, the data partitioning error resilience tool is supplemented with reversible variable length coding.

15 In the example context of MPEG-4 video, the video bit-stream syntax is modified to enable the use of these standard base layer error resilience tools within the enhancement layer bit-stream. Advantageously, the decoder and decoding operation are further adapted to make use of
20 the data partitioning, into two or more portions of bit streamed data, in the enhancement layer bit-stream in the same manner as for a base layer bit-stream.

An analogous technique can be used with H263+ video.

25

The focus for the enhanced further embodiment is that by employing data partitioning in the enhancement layer, either in isolation or in addition to the embodiments described with reference to FIG's 4 to 6, a (further) reduction in the
30 amount of data discarded when an error is detected can be achieved.

Further, the inventors of the present invention have recognised a potential deficiency of using a specific (re-synchronisation) Motion Marker to separate the motion parameters from DCT coefficients within a video packet header in the data partitioning. If an error occurs in the DCT coefficients part, then the original motion vectors are employed, as opposed to the estimated motion vectors, to conceal the remaining part of the video packet.

Thus, a corruption of one or more variable length coded (VLC) DCT coefficients impairs the rest of a video packet regardless of its error status. In this regard, the further embodiment of the present invention proposes to additionally use reversible variable length codes (RVLCs) to allow parts of a video sequence bit stream that could not be decoded in the forward direction due to the presence of errors, to be decoded in the reverse (backward) direction. A preferred example would be within enhancement layers of the MPEG-4 Simple Scalable Profile.

20

In accordance with the preferred further embodiment of the present invention, the compression function 360 in the video encoder 315 of FIG. 3 has been adapted to incorporate data partitioning in the enhancement layer bit-stream.

Furthermore, the decompression function 370 in the video decoder 325 of FIG. 3 has been adapted to determine that data partitioning in the enhancement layer bit-stream has been performed. The video decoder 325 has also been adapted to conceal errors within one or more data portions of the partitioned data that are detected as having errors. The video decoder 325 has further been adapted to perform RVLC on the enhancement layer, when an error in one or more data portions of a data-partitioned sequence is detected,

in a similar manner to that currently performed on a base layer of a video sequence. The use of such data partitioning and RVLC is further described with regard to FIG's 12 to 14.

5

It is within the contemplation of the invention that alternative encoding and decoding configurations could be adapted to use DP, and preferably RVLC within the enhancement layer bit-stream. As a result, the inventive concepts hereinafter described should not be viewed as
10 being limited to the example configuration provided in FIG. 3.

The introduction of such further error resiliency tools
15 within the enhancement layer reduces, or substantially eliminates, the macroblock displacements caused by error concealment utilising estimated motion vectors. Additionally, the use of DP and RVLCs minimises the number of discarded bits by facilitating the performance of reverse
20 decoding on a video bit stream.

FIG. 12 illustrates a data format 1200 of an enhancement layer bit stream that can be adapted to incorporate the use of data partitioning and RVLCs in accordance with the preferred further embodiment of the present invention. The
25 data format 1200 includes a video packet header 1210. The data format 1200 of the enhancement layer has been adapted to incorporate data partitioning, whereby the motion and header information 1220 has been separated from the texture
30 information 1240. A re-synchronisation motion marker 1230 indicates the location where the motion parameters are separated from the DCT coefficients. A video packet header

1250 of a subsequent video packet, VOP start code or GOV start code, may follow the data-partitioned data format.

In addition, a texture header 1235 has been introduced to
5 indicate texture related information to the decoder. Errors
1244 within the texture information can then be localised by
decoding in the forward direction 1242 and in the reverse
direction 1246, by use of RVLCs.

10 FIG. 13 illustrates the benefits of introducing data
partitioning and RVLC in an enhancement layer bit stream,
for example the data format of FIG. 12. Three enhancement
layer data formats are shown, all of which include initial
re-synchronisation markers 1320, as described above, and
15 video packet headers 1210. The first data format 1310
indicates a data structure that does not employ data
partitioning (and thereby does not employ RVLC). In this
regard, the body information contains both motion parameters
and DCT coefficients 1330.

20

An asterisk indicates an example of an error location 1244.
In the first data format 1310, a remaining part of the Video
Packet 1340, subsequent to the error 1244, is concealed.
The concealed part may include motion parameters and DCT
25 coefficients. Hence, as motion vectors may be corrupted,
estimated motion vector values must be used in the
concealing process.

However, the second data format 1350 has added data
30 partitioning only, as described with regard to FIG. 12. In
this regard, motion parameters 1220 have been separated from
the DCT coefficients 1240 by a motion marker 1230. Again,
an asterisk indicates an example of an error location 1244.

In the second data format 1350, a remaining part of the Video Packet 1360, subsequent to the error 1244, is concealed. However, in this regard, the concealed part only includes DCT coefficients, as the original motion vectors 5 1220 have been decoded prior to the motion marker 1230. Hence, the original motion vectors 1220 may be employed to produce a higher quality and less disturbing concealment.

As a further enhancement to the use of data partitioning in 10 data format 1350, RVLC can be employed. In this case, if an illegal or invalid VLC code is encountered, the decoder starts the reverse decoding from the end of the video packet until it encounters another invalid VLC code. If only one error occurs in the second data format 1350, the use of 15 RVLC, to decode what would otherwise be concealed DCT coefficients, results in a minimal or substantially eliminated concealment process.

Note that the benefits gained from using RVLC in addition to 20 data partitioning are dependent upon the number of errors occurring within a video packet. The third data format 1370 illustrates an example of this limitation, where two errors 1244 and 1245 are encountered in two different macroblocks. In this regard, the forward decoding stops at the first 25 error 1244, after the motion marker 1230. The reverse decoding process 1385 initiated by the decoder using RVLC, decodes the DCT coefficients from the end of the video packet back to the second error 1245. This leaves the concealed parts of the DCT coefficients being limited to the 30 intermediate region 1365. Clearly, this is a substantial improvement over the second data format 1350, data-partitioning only, case, where no reverse decoding is performed.

FIG. 14 illustrates the objective degradations caused by estimating values of lost motion vectors, or using original motion vectors, associated with the error scenarios of the first and second data formats of FIG. 13. Three plots are

- (i) An error free transmission of the enhancement layer;
- (ii) An enhancement layer with video packets where data partitioning is employed with an error 1244, and hence the original motion vectors are used for the concealment;
- (iii) An enhancement layer with video packets where motion parameters and DCT coefficients are not separated with an error 1244 introduced. Hence, estimated motion vectors are used for the concealment process.

15

In FIG. 14, the test sequence Foreman is coded at 18.8 kbit/s and 17.2 Kbit/s for the base and the temporal enhancement layers respectively. The size of the Video Packets was set to 800 bits. Errors in the enhancement layer were generated using a General Packet Radio System (GPRS) physical link layer simulator. The resultant Frame Erasure Rate (FER) is 6.8% and the Residual Bit Error Rate (RBER) is 0.1%.

Although it is evident from the results illustrated in FIG. 14, that data partitioning enhances the quality of concealed macroblocks, the absence of the DCT coefficients (or the prediction error in this case) is still noticeable. Any such degradation is preferably overcome by reducing the number of concealed macroblocks by the use of RVLCs. RVLCs can only be used in conjunction with the data-partitioning tool. Hence, by superimposing the reverse decoding technique of RVLC on the arrangements in FIG. 13, the DCT

30

coefficients concealed due to error 1244 can be minimised or substantially eliminated.

For the use of data partitioning and RVLCs in the
5 enhancement layers, a simple amendment to the current MPEG-4
video bit stream syntax (section 6.2.5 "Video object plane
and video plane with short header", VideoObjectPlane()) is
needed. The syntax code change, which is a further
enhancement to that illustrated in FIG. 9, is illustrated in
10 FIG. 15.

It is noteworthy that the benefits to be gained from
employing data partitioning (and subsequently RVLC) also
varies dependent upon the scalability mode used. For
15 instance in Spatial and SNR scalabilities, predicted VOPs
are always reconstructed without motion compensation, i.e.
with nil motion vectors. Separating motion parameters from
texture data in this case does not make much sense unless
RVLC is employed. This is because corrupted macroblocks are
20 often concealed by copying them from the corresponding co-
located ones of the base layer VOP without motion
compensation. On the other hand, for all other cases where
the macroblocks are motion compensated for prediction, data
partitioning is a useful tool regardless of the use of
25 RVLCs. Hence, it is within the contemplation of the
invention that data partitioning can be used in isolation,
or in tandem with RVLC, within enhancement layer video bit
streams.

30 As described above, the inventive concepts of the further
preferred embodiment enhances the concealment quality of
motion compensated macroblocks of the MPEG-4 enhancement
layers. Additionally, the use of RVLCs reduces the number

of concealed macroblocks by reverse decoding from the end of a Video Packet up to the erroneous VLC code.

Advantageously, the inventive concepts described above
5 enable error resilience tools to be used within scalable video transmissions in a wireless, error prone environment. Consequently, as the inventive concepts described herein enable error resilience tools to be employed, it is now
10 practically feasible to use scalable video in conjunction with network Quality of Service (QoS) information in order to deliver optimal quality to users in situations where network throughput and BER are likely to vary.

It is within the contemplation of the enhanced further
15 embodiment of the present invention, with the use of DP within an enhancement layer of a video sequence, for DP to be used in isolation, or together with RVLC, or in addition to the re-synchronisation marker technique described above. Clearly, wherever practically possible, it is preferred to
20 use all three error resilience tools within an enhancement layer video sequence. However, improved video quality benefits can still be gained by using re-synchronisation markers with or without data partitioning.

Claims

1. A method for improving a quality of MPEG-4, or similar, scalable video enhancement layers transmitted over
5 an error-prone network, the method comprising the steps of:
inserting one or more re-synchronisation markers (660, 760, 860) into one or more enhancement layers of a scalable video sequence;
detecting one or more errors (630, 730, 830) in said one or
10 more enhancement layers of said scalable video sequence;
and
concealing data (640, 740, 840) in said received scalable video sequence when said error is detected, from a video bit position at which said one or more errors is detected,
15 to a video bit position of a subsequent error-free re-synchronisation marker.
2. A method for improving a quality of H.263+, or similar, scalable video enhancement layers transmitted over
20 an error-prone network, the method comprising the steps of:
inserting one or more re-synchronisation markers (660, 760, 860) into one or more enhancement layers of a scalable video sequence;
detecting one or more errors (630, 730, 830) in said one or
25 more enhancement layers of said scalable video sequence;
and
concealing data (640, 740, 840) in said received scalable video sequence when said error is detected, from a video bit position at which said one or more errors is detected,
30 to a video bit position of a subsequent error-free re-synchronisation marker.

3. A method in accordance with claim 1 or claim 2, wherein said one or more errors (630, 730, 830) are detected over a frame start code (710) or across a re-synchronisation marker (660, 760, 860).

5

4. A method in accordance with any previous claim, wherein a number of said one or more re-synchronisation markers to be inserted into said one or more enhancement layers is user definable or pre-determined, for example to provide a trade off between error resiliency and bit efficiency.

10

5. A method in accordance with any previous claim, wherein a location of said one or more re-synchronisation markers is user definable or pre-determined, for example to provide a trade off between error resiliency and bit efficiency.

15

6. A method for improving a quality of MPEG-4 or H.263+ or similar, scalable video sequence transmitted over an error-prone network, in particular a method for improving a quality of MPEG-4, or similar, scalable video enhancement layer transmitted over an error-prone network according to any preceding Claim, the method comprising the steps of:

25

data partitioning (1220, 1240) of a video bit stream into two or more data portions in one or more enhancement layers of said scalable video sequence;
detecting one or more errors (1244, 1245) in one or more of said data portions; and
concealing data (1340, 1360) in one or more data portions where said one or more errors is detected.

30

7. A method according to Claim 6, the method further comprising the step of:
inserting one or more re-synchronisation markers (1230) into one or more enhancement layers of a scalable video sequence to indicate data partitioning between said two or more data portions.
8. A method according to Claim 6 or Claim 7, wherein said two or more data portions in said one or more enhancement layers include a data portion of motion parameters and a data portion of DCT coefficients.
9. A method according to any of preceding Claims 6 to 8, the method further comprising the step of:
performing reversible variable length coding on said one or more of said data portions of said scalable video sequence when said one or more errors (1244, 1245) is detected in said one or more of said data portions.

10. A video communication system (300) comprising:
- a video encoder (315) comprising:
a processor for encoding a video sequence into an MPEG-4,
5 or similar, scalable video sequence having a plurality of enhancement layers;
insertion means to insert one or more re-synchronisation markers into one or more enhancement layers; and
a transmitter for transmitting said scalable video sequence
10 containing said one or more re-synchronisation markers; and
- a video decoder (325) comprising:
a receiver for receiving said scalable video sequence containing said one or more re-synchronisation markers from
15 said video encoder;
a detector detecting one or more errors in said one or more enhancement layers of said received scalable video sequence; and
a processor operably coupled to said detector for
20 concealing data in said received scalable video sequence when said one or more errors is detected, wherein said video data is concealed from a video bit position at which said one or more errors is detected, to a video bit position of a subsequent error free re-synchronisation
25 marker.
11. The video communication system (300) according to claim 10, wherein said detector detects one or more errors over a frame start code or across a re-synchronisation
30 marker contained within said received scalable video sequence.

12. The video communication system (300) according to claim 10 or claim 11, wherein said video encoder further comprises user input means to enable a user to select:
a number of said one or more re-synchronisation markers to
5 be inserted into said one or more enhancement layers;
and/or
a location of said one or more re-synchronisation markers.

13. A video communication system (300), in particular a
10 video communication system according to any of preceding
Claims 10 to 12, the video communication system comprising:

a video encoder (315) comprising:
a processor for encoding a video sequence into an MPEG-4,
15 or similar, scalable video sequence having a plurality of
enhancement layers;
data partitioning means to partition data of a video bit
stream into two or more data portions in one or more
enhancement layers of said scalable video sequence; and
20 a transmitter for transmitting said scalable video sequence
containing said one or more data portions; and

a video decoder (325) comprising:
a receiver for receiving said scalable video sequence
25 containing said one or more data portions from said video
encoder;
a detector detecting one or more errors in said one or more
enhancement layers of said received scalable video
sequence; and
30 a processor operably coupled to said detector for
concealing data in said received scalable video sequence
when said one or more errors is detected, wherein said

video data is concealed in one or more data portions where said one or more errors is detected.

14. The video communication system (300) according to
5 claim 13, wherein said two or more data portions in said one or more enhancement layers include a data portion of motion parameters and a data portion of DCT coefficients.

15. The video communication system (300) according to
10 Claim 13 or Claim 14, wherein said video encoder further comprises means to perform reversible variable length coding on said one or more of said data portions of said scalable video sequence when said one or more errors (1244, 1245) is detected in said one or more of said data
15 portions.

16. A video communication unit (315, 325) adapted for use in the method of any of claims 1 to 9 or adapted for use in the communication system of any of claims 10 to 15.
20

17. A video encoder (315) or a video decoder (325) adapted for use in the method of any of claims 1 to 9 or adapted for use in the communication system of any of claims 10 to 15.
25

18. A mobile radio device comprising a video communication unit in accordance with claim 16 or a video encoder or a video decoder in accordance with claim 17.

30 19. A mobile radio device according to claim 13, wherein the mobile radio device is a mobile phone, a portable or mobile PMR radio, a personal digital assistant, a lap-top computer or a wirelessly networked PC.

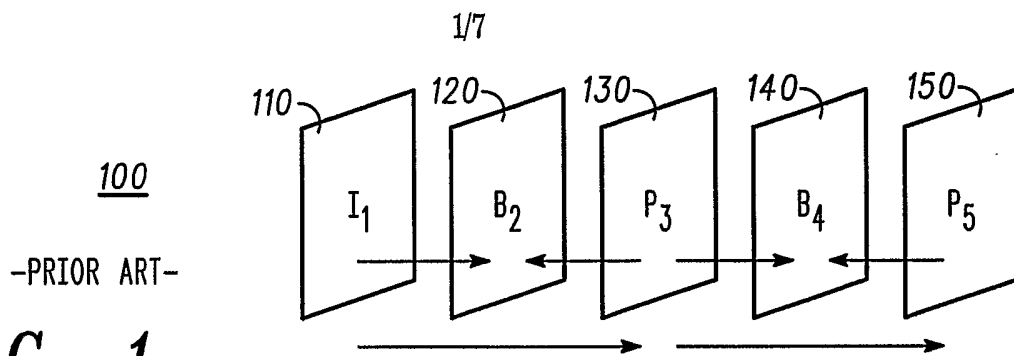


FIG. 1

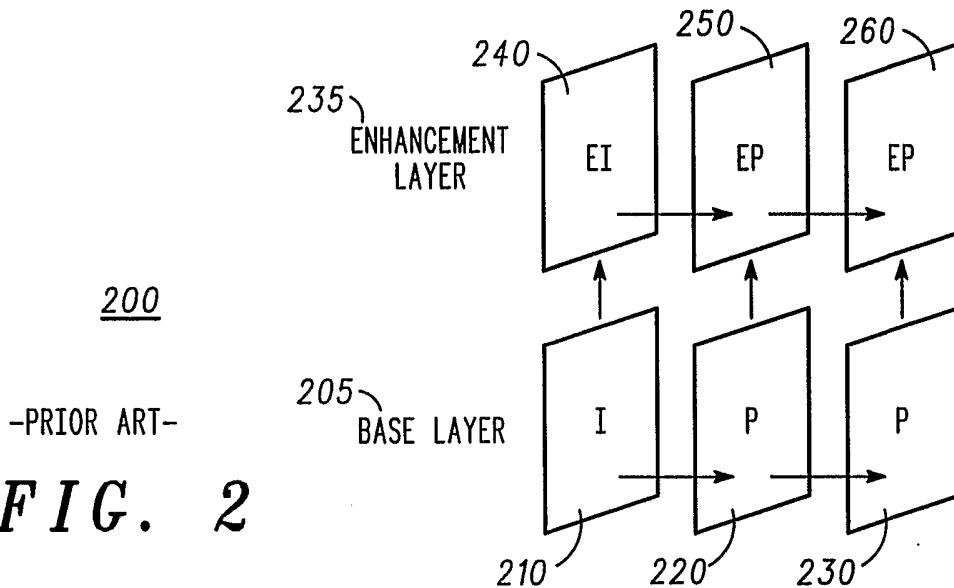
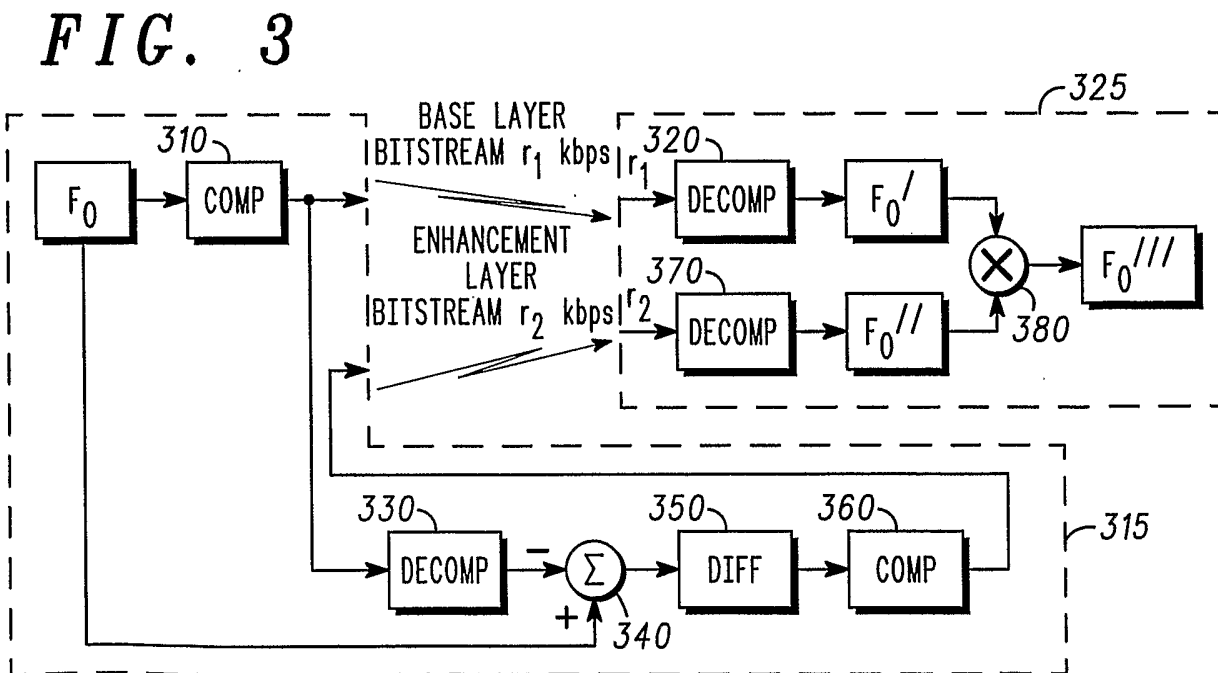


FIG. 2



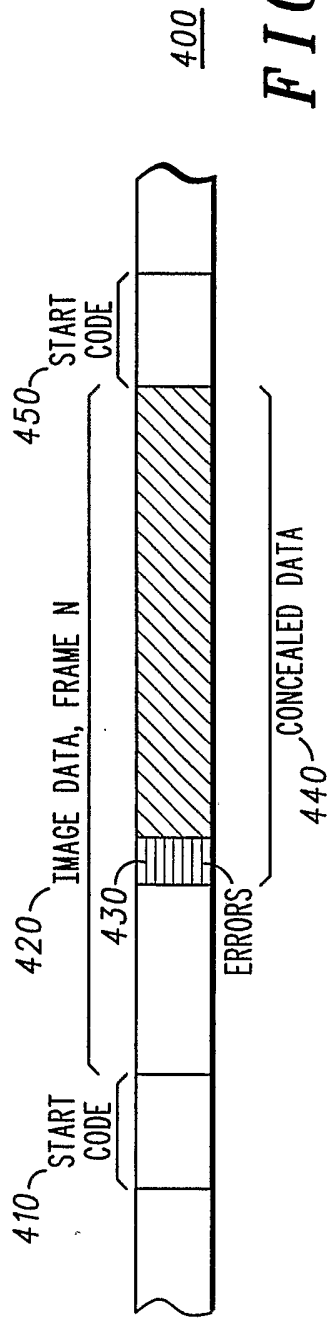


FIG. 4

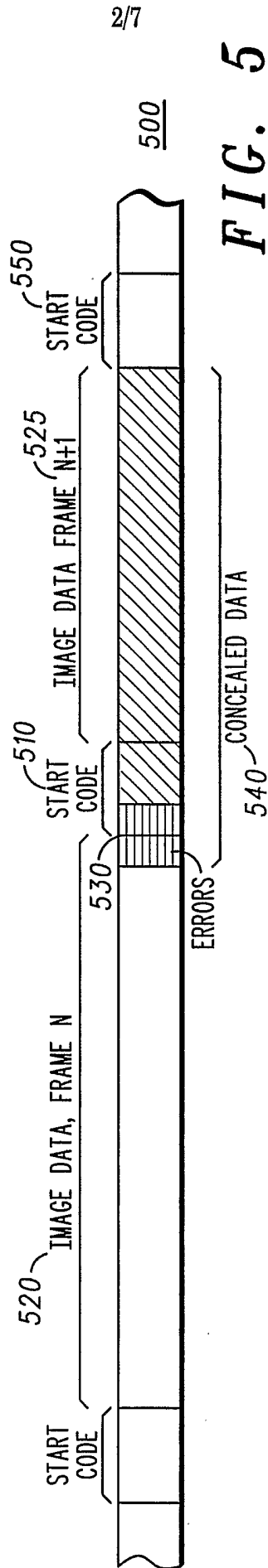


FIG. 5

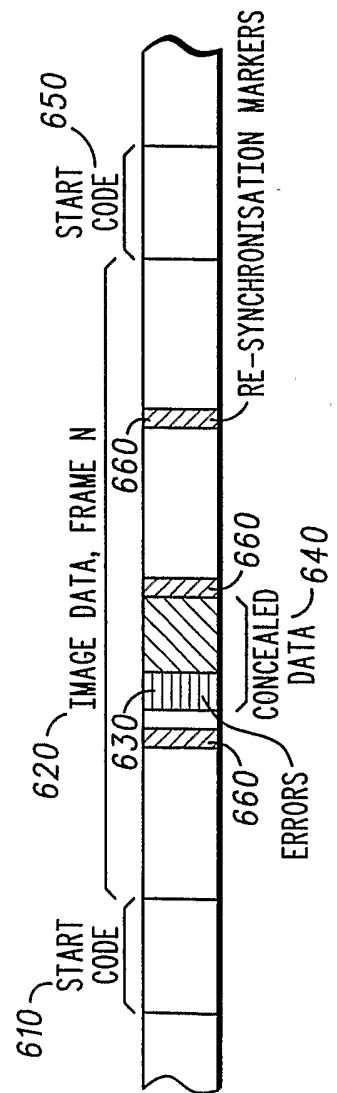


FIG. 6

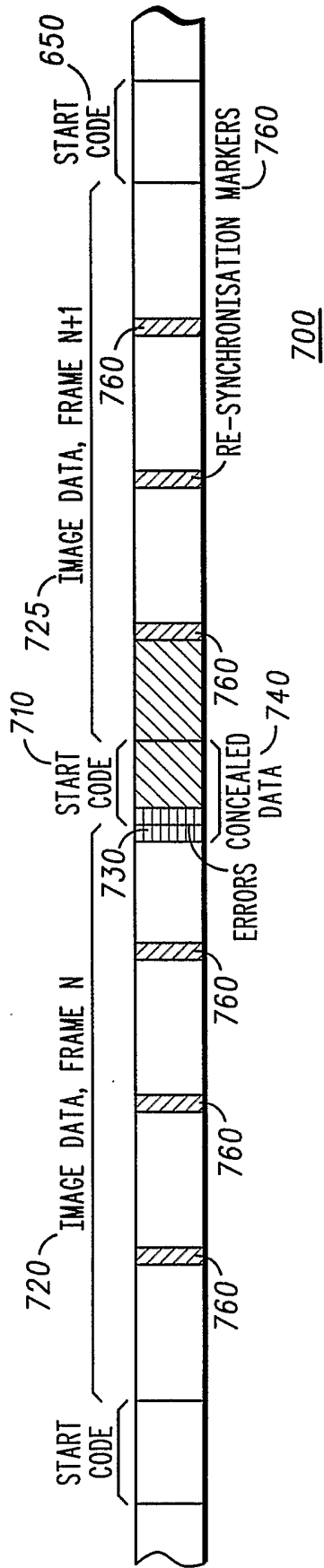


FIG. 7

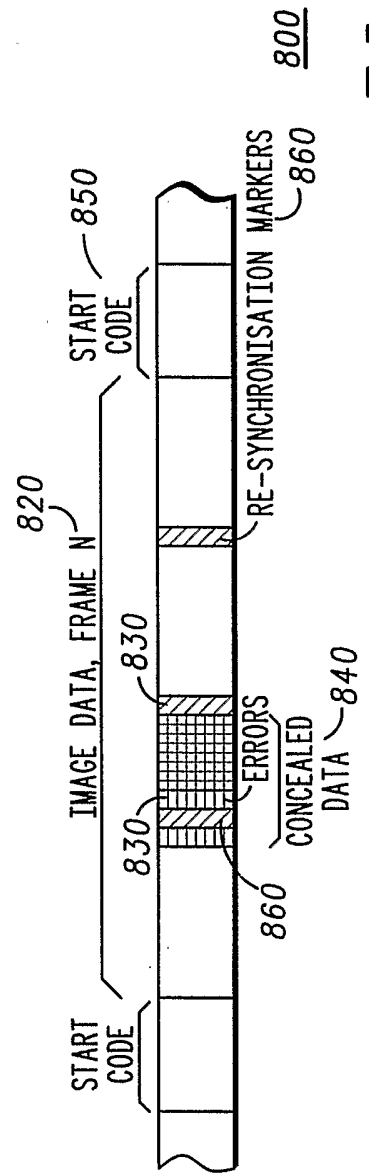


FIG. 8

```

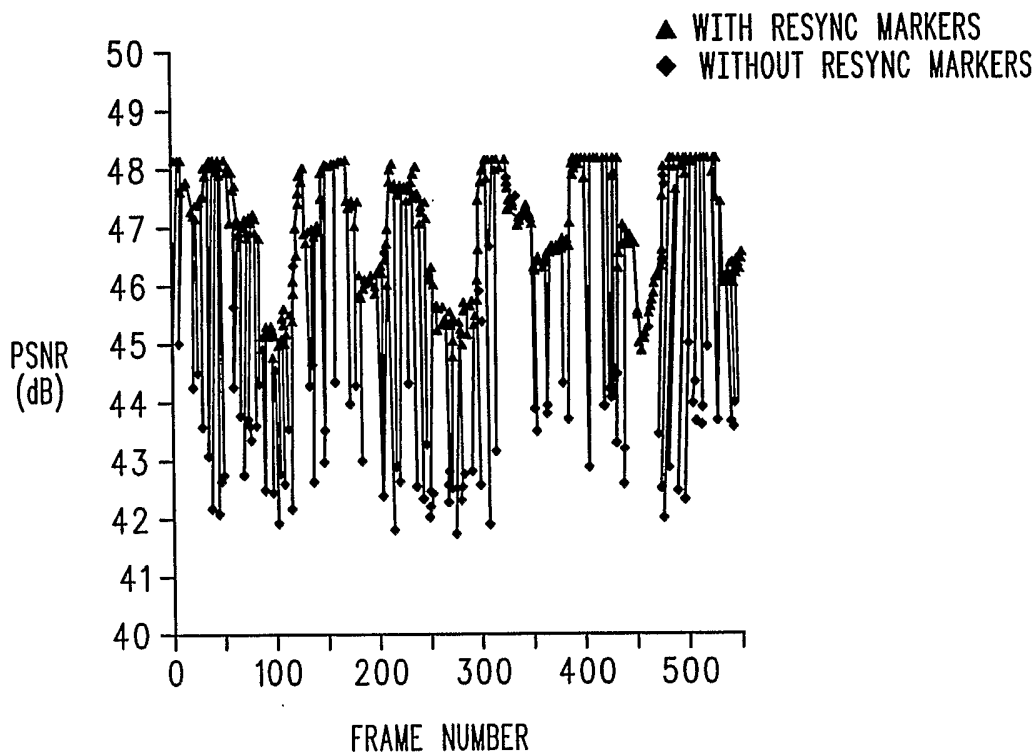
FROM
...
  if (video_object_layer_shape != "binary only") {
    if (!scalability) {
      ...
    }
    else {
      ...
      ref_select_code
      combined_motion_shape_texture()
    }
  }
}
to:
...
  if (video_object_layer_shape != "binary only") {
    if (!scalability) {
      ...
    }
    else {
      ...
      ref_select_code
      while (nextbits_bytealigned()==resync_marker)
      {
        video_packet_header()
        motion_shape_texture()
      }
    }
  }
}
...

```

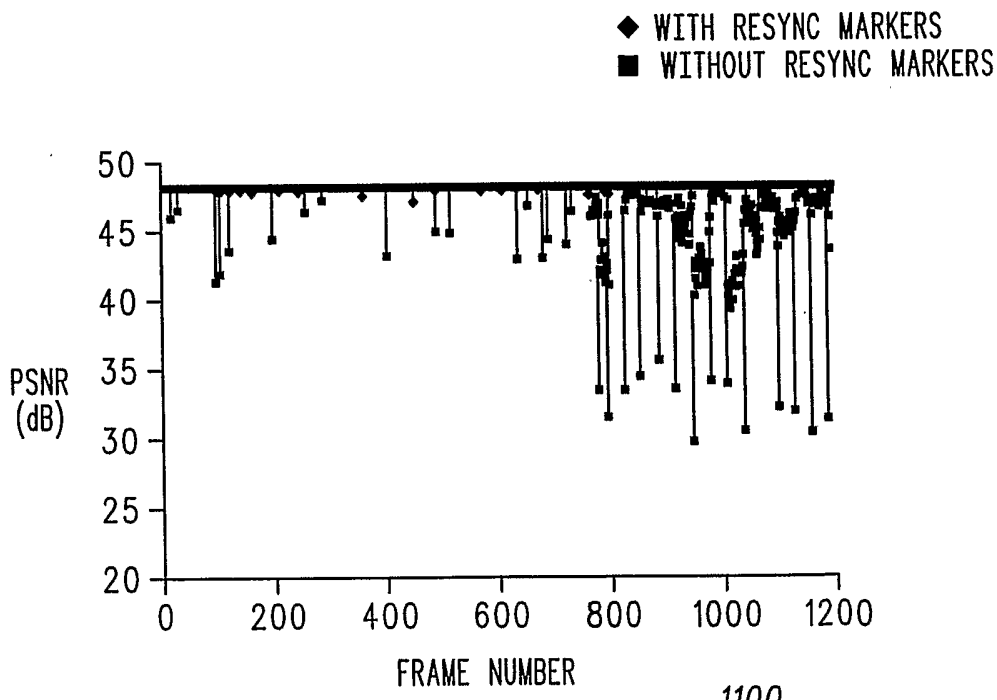
900

910

FIG. 9



1000
FIG. 10



1100
FIG. 11

6/7

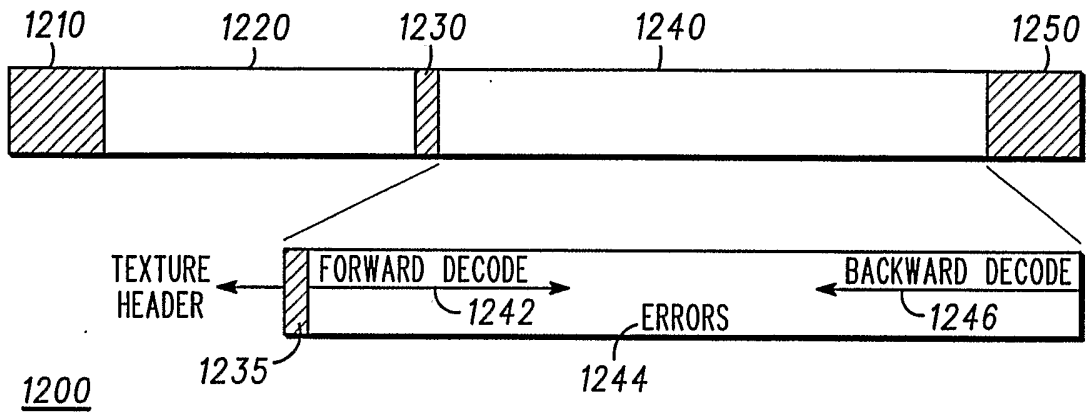


FIG. 12

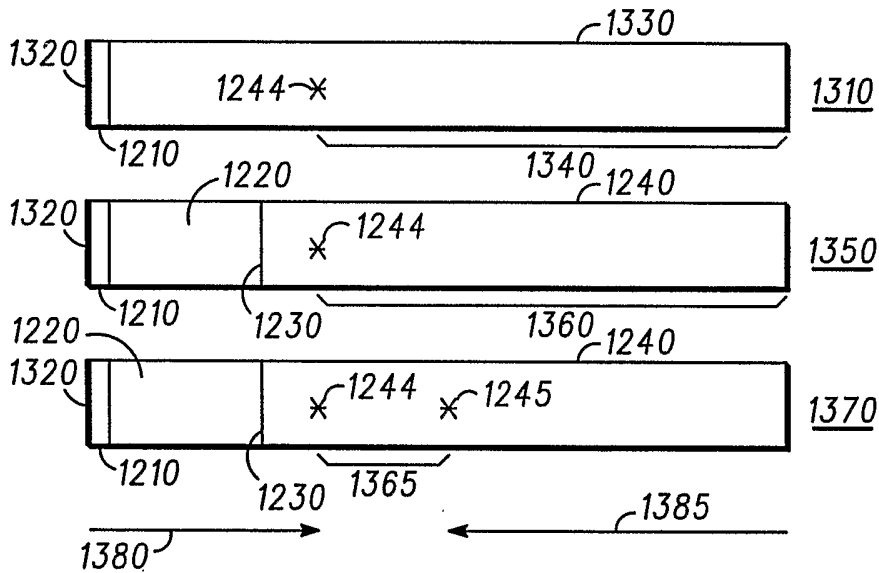


FIG. 13

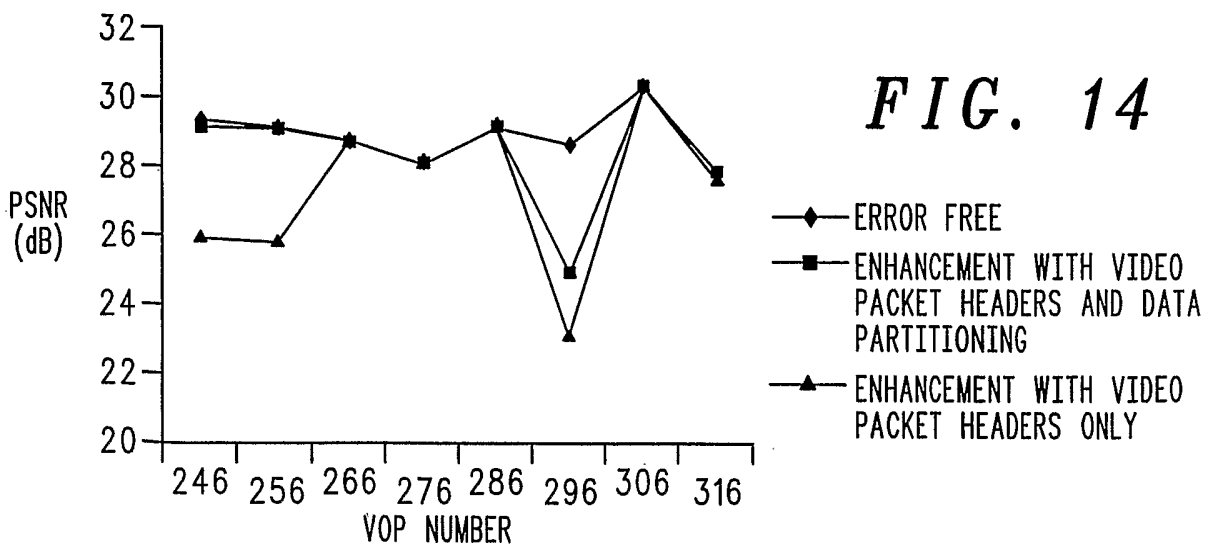


FIG. 14

FIG. 15

From:

(The syntax to be amended is in italic font)

```

...
  if (video_object_layer_shape != "binary only") {
    ...
    if (!scalability) {
      ...
      motion_shape_texture()
      while (nextbits_bytealigned() == resync_marker)
      {
        video_packet_header()
        motion_shape_texture()
      }
    }
    else {
      ...
      ref_select_code
      combined_motion_shape_texture()
    }
  }
...

```

To:

(The modified syntax is in **bold** font)

```

...
  if (video_object_layer_shape != "binary only") {
    ...
    if (!scalability) {
      ...
    }
    else {
      ...
      ref_select_code
    }
  }
  motion_shape_texture()
  while (nextbit_bytealigned() == resync_marker) {
    video_packet_header()
    motion_shape_texture()
  }
...

```