



(12) **EUROPEAN PATENT APPLICATION**  
 published in accordance with Art. 158(3) EPC

(43) Date of publication: **29.11.2006 Bulletin 2006/48**

(51) Int Cl.: **G06T 9/00 (2006.01) G06T 15/00 (2006.01)**  
**H04N 13/00 (2006.01)**

(21) Application number: **04715594.0**

(86) International application number: **PCT/MX2004/000012**

(22) Date of filing: **27.02.2004**

(87) International publication number: **WO 2005/083637 (09.09.2005 Gazette 2005/36)**

(84) Designated Contracting States:  
**AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LI LU MC NL PT RO SE SI SK TR**

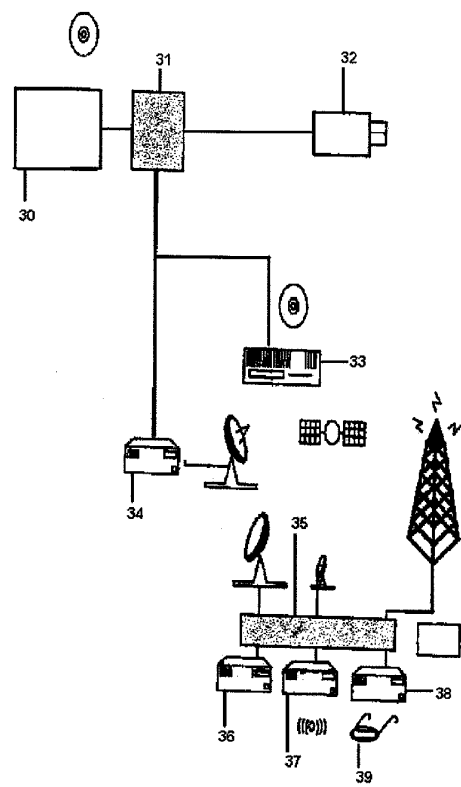
(71) Applicant: **Tdvision Corporation S.A. DE C.V. México, D.F. 02800 (MX)**

(72) Inventor: **GUTIÉRREZ NOVELO, Manuel Rafael México, D.F. 02800 (MX)**

(74) Representative: **Grey, Ian Michael Venner, Shipley LLP 20 Little Britain GB-London EC1A 7DH (GB)**

(54) **METHOD AND SYSTEM FOR DIGITAL DECODING 3D STEREOSCOPIC VIDEO IMAGES**

(57) MPEG-2 compatible stereoscopic 3D-video image digital decoding method and system, using its own coding algorithm. In order to obtain 3D-images from a digital video stream, modifications have been made to the current MPEG2 decoders, by means of software and hardware changes in different parts of the decoding process. Namely, the video\_sequence structures of the video data stream are modified via software to include the necessary flags at the bit level of the image type in the TD-Vision® technology; the modifications are made in the decoding processes as well as in decoding the information via software and hardware, where a double output buffer is activated, a parallel and differences decoding selector is activated, the decompression process is executed, the corresponding output buffer is displayed; the decoder must be programmed via software to simultaneously receive and decode two independent program streams, each with an TDVision® stereoscopic identifier.



**Fig. 1**

**EP 1 727 090 A1**

**Description****FIELD OF THE INVENTION**

5 [0001] The present invention is related to stereoscopic video image display in the 3DVisor® device and, particularly, to a video image decoding method by means of a digital data compression system, which allows the storage of three-dimensional information by using standardized compression techniques.

**BACKGROUND OF THE INVENTION**

10

[0002] Presently, data compression techniques are used in order to decrease the bits consumption in the representation of an image or a series of images. The standardization works were carried out by a group of experts of the International Standardization Organization. Presently, the methods are usually known as JPEG (Joint Photographic Expert Group), and MPEG (Moving Pictures Expert Group).

15

[0003] A common characteristic of these techniques is that the image blocks are processed by means of the application of a transform adequate for the block, usually known as Discrete Cosine Transform (DCT). The formed blocks are submitted to a quantization process, and then coded with a variable-length code.

[0004] The variable-length code is a reversible process, which allows the exact reconstruction of that which has been coded with the variable-length code.

20

[0005] The display of digital video signals includes a certain number of image frames (30 to 96 fps) displayed or represented successively at a 30 to 75 Hz frequency. Each image frame is still an image formed by a pixels array, according to the display resolution of a particular system. By example, the VHS system has a display resolution of 320 columns and 480 rows, the NTSC system has a display resolution of 720 columns and 486 rows, and the high definition television system (HDTV) has a display resolution of 1360 columns and 1020 rows. In reference to a digitized form of low resolution, 320 columns by 480 rows VHS format, a two-hour long movie could be equivalent to 100 gigabytes of digital video information. In comparison, a conventional compact optical disk has an approximate capacity of 0.6 gigabytes, a magnetic hard disk has a 1-2 gigabyte capacity, and the present compact optical disks have a capacity of 8 or more gigabytes.

25

[0006] All images we watch at the cinema and TV screens are based on the principle of presenting complete images (static images, like photographs) at a great speed. When they are presented in a fast and sequential manner at a 30 frames per second speed (30 fps) we perceive them as an animated image due to the retention of the human eye.

30

[0007] In order to codify the images to be presented in a sequential manner and form video signals, each image needs to be divided in rows, where each line is in turn divided in picture elements or pixels, each pixel has two associated values, namely, luma and chroma. Luma represents the light intensity at each point, while luma represents the color as a function of a defined color space (RGB), which can be represented by three bytes.

35

[0008] The images are displayed on a screen in a horizontal-vertical raster, top to bottom and left to right and so on, cyclically. The number of lines and frequency of the display can change as a function of the format, such as NTSC, PAL, or SECAM.

[0009] The video signals can be digitized for storage in digital format, after being transmitted, received, and decoded to be displayed in a display device, such as a regular television set or the 3DVisor®, this process is known as analog-to-digital video signal coding-decoding.

40

[0010] By definition, MPEG has two different methods for interlacing video and audio in the system streams.

[0011] The transport stream is used in systems with a greater error possibility, such as satellite systems, which are susceptible to interference. Each package is 188 bytes long, starting with an identification header, which makes recognizing gaps and repairing errors possible. Various audio and video programs can be transmitted over the transport stream simultaneously on a single transport stream; due to the header, they can be independently and individually decoded and integrated into many programs.

45

[0012] The program stream is used in systems with a lesser error possibility, as in DVD playing. In this case, the packages have a variable-length and a size substantially greater than the packages used in the transport stream. As a main characteristic, the program stream allows only a single program content.

50

[0013] Even when the transport and program streams handle different packages, the video and audio formats are decoded in an identical form.

[0014] In turn, there are three compression types, which are applied to the packages above, e.g. time prediction, compression, and space compression.

55

[0015] Decoding is associated to a lengthy mathematical process, which purpose is to decrease the information volume. The complete image of a full frame is divided by a unit called macroblock, each macroblock is made up of a 16 pixels x 16 pixels matrix, and is ordered and named top to bottom and left to right. Even with a matrix array on screen, the information sent over the information stream follows a special sequential sequence, i.e. the macroblocks are ordered

in ascending order, this is, macroblock0, macroblock1, etc.

**[0016]** A set of consecutive macroblocks represents a slice; there can be any number of macroblocks in a slice given that the macroblocks pertain to a single row. As with the macroblocks, the slices are numbered from left to right and bottom to top. The slices must cover the whole image, as this is a form in which MPEG2 compresses the video, a coded image not necessarily needs samples for each pixel. Some MPEG profiles require handling a rigid slice structure, by which the whole image should be covered.

**[0017]** USP No. 5,963,257 granted on October 5th, 1999 to Katata et al., protects a flat video image decoding device with means to separate the coded data by position areas and image form, bottom layer code, predictive coding top layer code, thus obtaining a hierarchical structure of the coded data; the decoder has means to separate the data coded in the hierarchical structure in order to obtain a high quality image.

**[0018]** USP No. 6,292,588 granted on September 18th, 2001 to Shen et al., protects a device and method for coding predictive flat images reconstructed and decoded from a small region, in such way that the data of the reconstructed flat image is generated from the sum of the small region image data and the optimal prediction data for said image. Said predictive decoding device for an image data stream includes a variable-length code for unidimensional DCT coefficients. USP No. 6,370,276 granted on April 9th, 2002 to Boon, uses a decoding method similar to the above.

**[0019]** USP No. 6,456,432 granted on September 24th, 2002 to Lazzaro et al., protects a stereoscopic 3D-image display system, which takes images from two perspectives, displays them on a CRT, and multiplexes the images in a field-sequential manner with no flickering for both eyes of the observer.

**[0020]** USP No. 6,658,056 granted on December 2, 2003 to Duruoz et al., protects a digital video decoder comprising a logical display section responding to a "proximal field" command to get a digital video field of designated locations in an output memory. The digital video display system is equipped with a MPEG2 video decoder. Images are decoded as a memory buffer, the memory buffer is optimized maintaining compensation variable tables and accessing fixed memory pointer tables displayed as data fields.

**[0021]** USP No. 6,665,445 granted on December 16th, 2003 to Boon, protects a data structure for image transmission, a flat images coding method and a flat images decoding method. The decoding method is comprised of two parts, the first part to codify the image-form information data stream, the second part is a decoding process for the pixel values of the image data stream, both parts can be switched according to the flat image signal coding.

**[0022]** USP No. 6,678,331 granted on January 13th, 2004 to Moutin et al., protects a MPEG decoder, which uses a shared memory. Actually, the circuit includes a microprocessor, a MPEG decoder, which decodes a flat image sequence, and a common memory for the microprocessor, and the decoder. It also includes a circuit for evaluating the decoder delay, and a control circuit for determining the memory priority for the microprocessor or the decoder.

**[0023]** USP No. 6,678,424 granted on January 13th, 2004 to Ferguson, protects a behavior model for a real-time human vision system; actually, it processes two image signals in two dimensions, one derived from the other, in different channels.

### **BRIEF DESCRIPTION OF THE INVENTION**

**[0024]** It is an object of the present invention to provide a stereoscopic 3D-video image digital decoding system and method, comprised of changes in software and changes in hardware.

**[0025]** It is an additional object of the present invention to provide a decoding method where the normal video\_sequence process is applied to the coded image data, i.e. variable\_length\_decoding (VLD), inverse\_scan; inverse\_quantization, inverse\_discrete\_cosine\_transform (IDCT), and motion\_compensation.

**[0026]** It is also an object of the present invention to make changes in the software information for decoding the identification of the video format, 2D-images MPEG2 backward compatibility, discriminating a TDVision® type image, storing the last image buffer, applying information decoding, applying error correction and storing the results in the respective channel buffer.

**[0027]** It is still another object of the present invention to provide a decoding method with the video\_sequence process normal form, in such a way that when a TDVision® type image is found, the buffer of the last complete image is stored in the left or right channel buffers.

**[0028]** It is also another object of the present invention to provide a decoding process in which two interdependent (difference) video signals can be sent within the same video\_sequence, in which information decoding is applied and is stored as a B type frame.

**[0029]** It is still another object of the present invention to provide a decoding process in which error correction is applied to the last obtained image when the movement and color correction vectors are applied.

**[0030]** It is also an object of the present invention to program the decoder by software, to simultaneously receive and codify two independent program streams.

**[0031]** It is still another object of the present invention to provide a decoding system, which decodes the 3D-image information via hardware, in which a double output buffer is activated.

[0032] It is another object of the present invention to provide a decoding system of 3D-image information, which activates an image-decoding selector in parallel and by differences.

[0033] It is also another object of the present invention to provide a 3D-image information decoding system, which executes the decompression process and displays the corresponding output buffer.

5

**DETAILED DESCRIPTION OF THE INVENTION.**

[0034] The combination of hardware and software algorithms makes possible the stereoscopic 3D-image information compression, which are received as two independent video signals but with the same time\_code, corresponding to the left and right signals coming from a 3Dvision® camera, by sending two simultaneous programs with stereoscopic pair identifiers, thus promoting the coding-decoding process. Also, two interdependent video signals can be handled by obtaining their difference, which is stored as a "B" type frame with the image type identifier. As the coding process was left open in order to promote the technological development, it is only necessary to follow this decoding process, namely: apply variable-length decoding to the coded data where a substantial reduction is obtained, but a look-up table must be used to carry out decoding; apply an inverse scan process; apply an inverse quantization process in which each data is multiplied by a scalar; apply the inverse cosine transform function; apply error correction or motion compensation stage and eventually obtain the decoded image.

10

15

[0035] The novel characteristics of this invention in connection with its structure and operation method will be better understood from the description of the accompanying figures, together with the attached specification, where similar numerals refer to similar parts and steps.

20

[0036] Figure 1 represents the technology map to which the subject object of the present invention pertains. It shows a stereoscopic 3D-image coding and decoding system and corresponding method. The images come from a stereoscopic camera (32), the information compiled in (31) and are displayed in any adequate system (30) or (33). The information is coded in (34) and then it can be transmitted to a system having an adequate previous decoding stage such as (35), which may be a cable system (36), a satellite system (37), a high definition television system (38) or a stereoscopic vision system such as TDVision®'s 3DVisors® (39).

25

[0037] Figure 2 shows a flowchart in which the steps of the process are outlined. The objective is to obtain three-dimensional images from a digital video stream by making modifications to the current MPEG2 decoders, and changes to software (3) and hardware (4) in the decoding process (2): the decoder (1) must be compatible with MPEG2-4.

30

[0038] Figure 3 outlines the structures that must be modified and the video\_sequence of the data stream in order to identify the TDVision® technology image type at the bit level.

[0039] Each of the stages of the decoding process is detailed below (20):

The coded data (10) are bytes with block information, macroblocks, fields, frames, and MPEG2 format video images.

35

[0040] Variable\_length\_decoding (11) (VLC, Variable-length Decoder) is a compression algorithm in which the most frequent patterns are replaced by shorter codes and those occurring less frequently are replaced by longer codes. The compressed version of this information occupies less space and can be transmitted faster by networks. However, it is not an easily editable format and requires decompression using a look-up table.

40

[0041] For example, the word BEETLE

Letter	ASCII Code	VLC
B	01000010	0000 0010 10
E	01100101	11
L	0110 1100	0001 01
T	0111 0100	0100

45

[0042] Therefore, the ASCII code for the word is: 0100 0010 0110 0101 0110 0101 0111 01000 0110 1100 0110 0101 in VLC: 0000 0010 10 11 11 0100 00010 01 11.

50

[0043] A substantial decrease is noted, however, in order to go back from VLC to the word 'Beetle' a search in the look-up table is needed to decode the bit stream, this is made by exact comparison of the read bits.

[0044] Inverse scan (12): The information must be grouped by blocks, and by coding the information with the VLC a linear stream is obtained. The blocks are 8x8 data matrixes, so it is necessary to convert the linear information in a square 8x8 matrix. This is made in a descending zigzag manner, top to bottom and left to right in both sequence types, depending on whether it is a progressive image or an interlaced image.

55

[0045] Inverse Quantization (13): It consists simply in multiplying each data value by a factor. When codified, most of the data in the blocks are quantized to remove information that the human eye is not able to perceive, the quantization

## EP 1 727 090 A1

allows to obtain a greater MPEG2 stream conversion, and it is also required to perform the inverse process (Inverse quantization) in the decoding process.

**[0046]** Inverse cosine transform (14) (IDCT, inverse\_discrete\_cosine\_transform): The data handled within each block pertain to the frequency domain, this inverse cosine transform allows to return to the samples of the space domain.

Once the data in the IDCT have been transformed, pixels, colors and color corrections can be obtained.

**[0047]** Motion compensation (15) allows to correct some errors generated before the decoding stage of MPEG format, motion compensation takes as a reference a previous frame and calculates a motion vector relative to the pixels (it can calculate up to four vectors), and uses them to create a new image. This motion compensation is applied to the P and B type images, where the image position is located over a "t" time from the reference images. Additionally to the motion compensation, the error correction is also applied, as it is not enough to predict the position of a particular pixel, but a change in its color can also exist. Thus, the decoded image is obtained (16).

**[0048]** To decode a P or B type image, the reference image is taken, the motion vectors are algebraically added to calculate the next image, and finally the error correction data is applied, thus generating the decoded image successfully. Actually, in the video\_sequence, two interdependent video signals exist, "R-L= delta, the delta difference is that stored as a B type stereoscopic pair frame with TDVision® identifier and which is constructed at the moment of decoding by differences from the image. This is, R-delta= L and L-delta= R, the left image is constructed from the difference with the right image, which in turn is constructed from the difference with the left image.

**[0049]** The previous process is outlined in such a way that the left or right signal is taken, both are stored in a temporary buffer, then the difference between the left and right signals is calculated, and then it is coded as a B type image stored in the video\_sequence to be later decoded by differences from said image.

**[0050]** In the decoding process it can be deduced that the data inputted by the VLC stage are much smaller than the data outputted by the same stage.

**[0051]** MPEG video sequence structure: This is the maximum structure used in the MPEG2 format and has the following format:

Video sequence (Video\_Sequence)  
Sequence header (Sequence\_Header)  
Sequence extension (Sequence\_Extension)  
User Data (0) and Extension (Extension\_and\_User\_Data (0))  
Image group header (Group\_of\_Picture\_Header)  
User Data (1) and Extension (Extension\_and\_User\_Data (1))  
Image header (Picture\_Header)  
Coded image extension (Picture\_Coding\_Extension)  
User Data (2) and Extensions (Extension\_and\_User\_Data (2))  
Image Data (Picture\_Data)  
Slice(Slice)  
Macroblock (Macroblock)  
Motion vectors (Motion\_Vectors)  
Coded Block Pattern (Coded\_Block\_Pattern)  
Block (Block)  
Final Sequence Code (Sequence\_end\_Code)

**[0052]** These structures make up the video sequence. A video sequence is applied for MPEG format, in order to differentiate each version there must be a validation that immediately after the sequence header, the sequence extension is present; should the sequence extension not follow the header, then the stream is in MPEG1 format.

**[0053]** At the beginning of a video sequence, the sequence\_header and sequence\_extension appear in the video\_sequence. The sequence\_extension repetitions must be identical on the first try and the "s" repetitions of the sequence\_header vary little compared to the first occurrence, only the portion defining the quantization matrixes should change. Having sequences repetition allows a random access to the video stream, i.e., if the decoder wants to start playing at the middle of the video stream this may be done, as it only needs to find the sequence\_header and sequence\_extension prior to that moment in order to decode the following images. This also happens for video streams that could not start from the beginning, such as a satellite decoder turned on after the transmission time.

**[0054]** The full video signal coding-decoding process is comprised of the following steps:

Digitizing the video signals, which can be done in NTSC, PAL or SECAM format.  
Storing the video signal in digital form  
Transmitting the signals  
Recording the digital video stream in a physical media (DVD, VCD, MiniDV)

## EP 1 727 090 A1

Receiving the signals  
Playing the video stream  
Decoding the signal  
Displaying the signal

5

**[0055]** It is essential to double the memory to be handled by the adequate DSP and have the possibility of disposing of up to 8 output buffers, which allow the previous and simultaneous representation of a stereoscopic image on a device such as TDVision®'s 3DVisor®

10

**[0056]** Actually, two channels must be initialized when calling the programming API of the DSP as, by example, the illustrative case of the Texas Instruments TMS320C62X DSP.

**[0057]** MPEG2VDEC\_create (const IMPEG2VDEC\_fxns\*fxns, const MPEG2VDEC\_Params\* params).

**[0058]** Where IMPEG2VDEC\_fxns y MPEG2VDEC\_Params are pointer structures defining the operation parameters for each video channel, e.g.:

15

```
3DLhandle=MPEG2VDEC_create (fxns3DLEFT,Params3DLEFT).  
3DRhandle=MPEG2VDEC_create(fxns3DRIGHT,Params3DRIGHT).
```

**[0059]** Thereby enabling two video channels to be decoded and obtaining two video handlers, one for the left-right stereoscopic channel.

20

**[0060]** A double display output buffer is needed and by means of software, it will be defined which of the two buffers must display the output by calling the AP function:

```
Namely, MPEG2VDEC_APPLY(3DRhandle, inputR1, inputR2, inputR3, 3dourright_pb, 3dourright_fb).  
MPEG2VDEC_APPLY(3DLhandle, inputL1, inputL2, inputL3, 3doutleft_pb, 3doutleft_fb).
```

25

**[0061]** This same procedure can be implemented for any DSP, microprocessor or electronic device with similar functions.

30

**[0062]** Where 3DLhandle is the pointer to the handle returned by the DSP's create function, the input1 parameter is the FUNC\_DECODE\_FRAME or FUNC\_START\_PARA address, input2 is the pointer to the external input buffer address, and input3 is the size of the external input buffer size.

**[0063]** 3doutleft\_pb is the address of the parameter buffer and 3doutleft\_fb is the beginning of the output buffer where the decoded image will be stored.

**[0064]** The timecode and timestamp will be used for output to the final device in a sequential, synchronized manner.

35

**[0065]** It is essential to double the memory to be handled by the DSP and have the possibility of disposing of up to 8 output buffers which allow the previous and simultaneous display of a stereoscopic image on a device such as TDVision® Corporation's 3DVisor®.

**[0066]** The integration of software and hardware processes is carried out by devices known as DSP, which execute most of the hardware process. These DSP are programmed by a C and Assembly language hybrid provided by the manufacturer. Each DSP has its own API, consisting of a functions list or procedure calls located in the DSP and called by software.

40

**[0067]** With this reference information, the present application for MPEG2 format-compatible 3D-images decoding is made.

45

**[0068]** Actually, at the beginning of a video sequence the sequence header (sequence\_header) and the sequence extension always appear. The repetitions of the sequence extension must be identical to the first. On the contrary, the sequence header repetitions vary a little as compared to the first occurrence, only the portion defining the quantization matrixes should change.

50

**[0069]** Figure 4 shows the compilation software format for the TDVision® decoding method (40), where the video\_sequence (41) of the digital stereoscopic image video stream is identified, which may be dependent or independent (parallel images), in the sequence\_header (42). If the image is TDVision® then the double buffer is activated and the changes in the aspect\_ratio\_information are identified. The information corresponding to the image that can be found here is read in the user\_data (43). The sequence\_scalable\_extension (44) identifies the information contained in it and the base and enhancement layers, the video\_sequence can be located here, defines the scalable\_mode and the layer identifier. extra\_bit\_picture (45) identifies the picture\_estructure, picture\_header and the picture\_coding\_extension (46) reads the "B" type images and if it is a TDVision® type image, then it decodes the second buffer. picture\_temporal\_scalable\_extension () (47), in case of having temporal scalability, is used to decode B type images.

55

**[0070]** Namely, the sequence header (sequence\_header) provides a higher information level on the video stream, for clarity purposes the number of bits corresponding to each is also indicated, the most significative bits are located within the sequence extension (Sequence\_Extension) structure, it is formed by the following structures:

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55

**Sequence\_Header**

Field	bits	Description
Sequence_Header_Code	32	Sequence_Header_Start 0x00001 B3
Horizontal_Size_Value	12	less significant bits for width *
Vertical_Size_Value	12	12 less significant bits for length
<b>Aspect Ratio Information</b>	4	image aspect
		0000 forbidden
		0001 n/a TDVision®
		0010 4:3 TDVision®
		0011 16:9 TDVision®
		0100 2.21:1 TDVision®
		0111 will execute a logical "and" in order to obtain backward compatibility with 2D systems.
		0101...1111 reserved
<b>Frame rate code</b>	4	0000 forbidden
		0001 24,000/1001 (23.976) in TDVision® format
		0010 24 in TDVision® format
		0011 25 in TDVision® format
		0100 30,000/1001 (29.97)"
		0101 30 in TDVision® format
		0110 50 in TDVision® format
		0111 60,000/1001 (59.94) " (will execute a logical "and" in order to obtain backward compatibility with 2D systems.)
		1000 60
		1111 reserved

Bit_rate_value	18	The 18 less significant bits of the video_stream bit rate (bit_rate= 400 x bit_rate_value + bit_rate_extension<< 18) the most significant bits are located within the sequence_extension structure.
Marker_bit	1	Always 1 (prevents start_code failure).
Vbv_buffer_size_value	10	The 10 less significant bits of vbv_buffer_size, which determines the size of the video buffering verifier (Vbv), a structure used to ensure that a data stream can be used decoding a limited size buffer without exceeding or leaving too much free space in the buffer.
Constrained_parameters_flag	1	Always 0, not used in MPEG2.
Load_intra_quantizer_matrix	1	Indicates if an intra-coded quantization matrix is available.
If (load_intra_quantizer_matrix)	8x64	If a quantization matrix is indicated, then it must be specified here, it is a 8x64 matrix.

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55

	(continued)	
Field	Description	
Load_non_intra_quantizer_matrix	If load_non_intra_quantizer_matrix	bits 1
If		
load_non_intra_quantizer_matrix		
Non_intra_quantizer_matrix (64)	If the previous flag is activated, the 8x64 data forming the quantized matrix are stored here.	8x64

\* The most significant bits are located within the sequence\_extension structure.



5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55

**Picture\_coding\_extension**

Field	bits	Description
Extension_start_code	3	Always 0x000001B5
Extension_start_code_identifier	4	Always 1000
F_code(0)(0)	4	Used to decode motion vectors; when it is a type I image, this data is filled with 1111.
F_code(0)(1)	4	
F_code(1)(0)	4	Decoding information backwards in motion vectors (B), when it is a (P) type image it must be set to 1111, because there is no backward movement.
F_code(1)(1)	4	Decoding information backwards in motion vectors, when it is a P type image it must be set to 1111, because there is no backward movement.
Intra_dc_precision	2	precision used in the inverse quantizing of the coefficients of the DC discrete cosine transform. 00 8 bits precision 01 9 bits precision 10 10 bits precision 11 11 bits precision
Picture_structure	2	Specifies if the image is divided in fields or in a full frame. 00 reserved (image in TDV <sup>ision</sup> ® format) 01 top field 10 bottom field 11 by-frame image
Top_field_first	1	0= decode bottom field first 1= decode top field first
Frame_pred_frame_dct	1	
Concealment_motion_vectors	1	
Q_scale_type	1	
Intra_vic_format	1	
Alternate_scan	1	
Repeat_first_field	1	0= display a progressive frame 1= display two identical progressive frames

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55

(continued)

Field	bits	Description
Chroma_420_type	1	If the chroma format is 4:2:0, then it must be equal to progressive frame, otherwise it must be equal to zero.
Progressive_frame	1	0= interlaced 1= progressive
Composite_display_flag	1	warns about the originally coded information
V_axis	1	
Field_sequence	3	
Sub_carrier	1	
Burst_amplitude	7	
Sub_carrier_phase	8	
Next_start_code()		

## EP 1 727 090 A1

### Picture\_temporal\_scalable\_extension()

**[0071]** Two spatial resolution streams exist in case of having temporal scalability, the bottom layer provides a lesser index version of the video frames, while the top layer can be used to derive a greater index version of frames of the same video. The temporal scalability can be used by low quality, low cost or free decoders, while the greater frames per second would be used for a fee.

### Picture\_temporal\_scalable\_extension()

Field	bits #	Definition
Extension_start_code_identifier	4	Always 1010
Reference_select_code	2	It is used to indicate that the reference image will be used to decode intra_coded images FOR O TYPE IMAGES 00 enhances the most recent images 01 the lower and most recent frame layer in display order 10 the next lower frame layer in order of forbidden display. 11 forbidden FOR B TYPE IMAGES 00 forbidden 01 most recently decoded images in enhanced mode 10 most recently decoded images in enhanced mode 11 most recent image in the bottom layer in display order
Forward_temporal_reference	10	Temporal reference
Marker_bit	1	
Backward_temporal_reference	10	Temporal reference
Next_star_code()		

### Picture\_spatial\_scalable\_extension()

**[0072]** In the case of image spatial scalability, the enhancement layer contains data, which allow a better resolution of the base layer so it can be reconstructed. When an enhancement layer is used as a function of a base layer as a reference for the motion compensation, then the bottom layer must be escalated and offset in order to obtain greater resolution of the enhancement layer.

### Picture\_spatial\_scalable\_extension()

### [0073]

Field	bits #	Definition
Extension_start_code_identifier	4	Always 1001
Lower_layer_temporal_reference	10	Reference to the lower layer's temporal image
Marker_bit	1	1
Lower_layer_horizontal_offset	15	Horizontal compensation (Offset)
Marker_bit	1	1
Lower_layer_vertical_offset	15	Vertical compensation (Offset)
Spatial_temporal_weight_code_table_index	2	Prediction details
Lower_layer_progressive_frame	1	1= progressive 0= interlaced
Lower_layer_desinterlaced_field_select	1	0= the top field is used 1= the bottom field is used
Next_start_code()		
Copyright_extension()		
Extension_start_code_identifier	4	Always 010
Copyright_flag	1	if it is equal to 1 then it uses copyright If it is zero (0), no additional copyright information is needed

**EP 1 727 090 A1**

(continued)

	Field	bits #	Definition
5	Copyright_identifier	8	1= original 0= copy
	Original_or copy	1	
	Reserved	7	
	Marker_bit	1	
10	Copyright_number_1	20	Number granted by copyright instance
	Marker_bit	1	
	Copyright_number_2	22	Number granted by copyright instance
	Marker_bit	1	
	Copyright_number_3	22	Number granted by copyright instance
15	Next_start_code()		
	Picture_data()		

**[0074]** This is a simple structure, it does not have field in itself.

20 Slice()

**[0075]** Contains information on one or more macroblocks in the same vertical position.

25	Slice_start_code	32
	Slice_vertical_position_extension	3
	Priority_breakpoint	7
	Quantizer_scale_code	5
	Intra_slice_flag	1
30	Intra_slice	1
	Reserved_bits	7
	Extra_bit_slice	1
	Extra_information_slice	8
	Extra_bit_slice	1
35	Macroblock()	
	Macroblock_modes()	
	Motion_vectors()	
	Motion_vector()	
40	Coded_block_pattern()	
	Block()	

EXTENSION AND\_USER\_DATA(2)

45 **[0076]** The image can be displayed in:

- DVD (Digital Versatile Disks)
- DTV (Digital Television)
- HDTV (High Definition Television)
- 50 CABLE (DVB Digital Video Broadcast) SATELLITE (DSS Digital Satellite Systems); and it is the software and hardware process integration.

**[0077]** The decoding compilation format in the hardware (50) section of Figure 5, is duplicated in the DSP input memory, at the same time, the simultaneous input of two independent or dependent video signals is allowed, corresponding to the left-right stereoscopic existing signal taken by the stereoscopic TDVision® camera. In the procedure the video\_ sequence (51) is detected to alternate the left and right frames or sending them in parallel, sequence\_header (52) identification, the image type (53) is identified, it passes to the normal video stream (54), then it is submitted to an error

## EP 1 727 090 A1

correction process (55), the video image information is sent to the output buffer (56) which in turn shares and distributes the information to the left channel (57) and the right channel (58) in said channels the video stream information is displayed in 3D or 2D.

**[0078]** Consists in storing both L (left) and R (right) video streams in simultaneous form as two independent video streams, but synchronized with the same time\_code, so they can later be decoded and played back in parallel in large output buffers. They can also be dependent and decoded by differences.

**[0079]** Regarding hardware, most of the process is executed by devices known as DSP (Digital Signal Processors). As an example, namely, the Motorola and the Texas Instruments (TMS320C62X) models can be used.

**[0080]** These DSP are programmed by a hybrid language from C and Assembly languages, provided by the manufacturer in question. Each DSP has its own API, consisting of a functions list or procedure calls located in the DSP to be called by software. From this reference information, the 3D-images are coded, which are compatible with the MPEG2 format and with their own coding algorithm. When the information is coded, the DSP is in charge of running the prediction, comparison, quantization, and DCT function application processes in order to form the MPEG2 compressed video stream.

**[0081]** In order to obtain three-dimensional images from a digital video stream, certain modifications have been made to the current MPEG2 decoders, by software and hardware changes in different parts of the decoding process. The structures and the video\_sequence of the video data stream must be modified to include the necessary flags to identify at the bit level the TDVision® technology image type.

**[0082]** The modifications are made in the next decoding steps. **Software:**

- Video format identification.
- Application of a logical "and" for MPEG2 backward compatibility in case of not being a TDVision® video.
- Image decoding in normal manner (previous technique) scanning the video\_sequence.
- In case of a TDVision® type image:
  - \_ Discriminating if they are dependent or independent video signals
  - \_ Store the last complete image buffer in the left or right channel buffer.
  - \_ Apply the B type frame information decoding.
  - \_ Apply error correction to the last obtained image by applying the motion and color correction vectors.
  - \_ Store the results in their respective channel buffer.
  - \_ Continue the video sequence reading.

### Hardware:

**[0083]** When the information is decoded via hardware;

- \_discriminate if the image is 2D or 3D
- \_Activate a double output buffer (memory is increased).
- \_The difference decoding selector is activated.
- \_The parallel decoding selector is activated.
- \_ The decompression process is executed.
- \_ The image is displayed in its corresponding output buffer.

**[0084]** The following structures, sub-structures and sequenced will be used in specific ways; they belong to the video\_sequence structure for the hardware implementation of the MPEG2 backward compatible TDVision® technology.

**[0085]** Actually:

Sequence\_header  
Aspect\_ratio\_information  
1001 n/a in TDVision®  
1010 4:3 in TDVision®  
1011 16:9 in TDVision®  
1100 2.21:1 in TDVision®

**[0086]** A logical "and" will be executed with 0111 to obtain the backward compatibility with 2D systems, when this occurs, the instruction is sent to the DSP that the buffer of the stereoscopic pair (left or right) must be equal to the source, so all the images decoded will be sent to both output buffers to allow the image display in any device.

Frame\_rate\_code  
1001 24,000/101 (23.976) in TDVision® format

## EP 1 727 090 A1

1010 24 in TDVision® format.  
 1011 25 in TDVision® format.  
 1100 30,000/1001 (29.97) in TDVision® format.  
 1101 30 in TDVision® format.  
 5 1110 50 in TDVision® format.  
 1111 60,000/1001 (59.94) in TDVision® format.

**[0087]** A logical "and" with 0111 will be executed in order to obtain backward compatibility with 2D systems.

```

User_data()
Sequence_scalable_extension
Picture_header
      Extra_bit_picture
                                0 = TDVision®
                                1 = normal
Picture_coding_extension
      Picture-structure
      00= image in TDVision® format
Picture_temporal_scalable_extension()
```

**[0088]** At the moment of coding the information a DSP is used which is in charge of executing the prediction, comparison, and quantization processes, applies the DCT to form the MPEG2 compressed video stream, and discriminates between 2D or 3D-images.

**[0089]** Two video signals are coded in an independent form but with the same time\_code, signals corresponding to the left signal and the right signal coming from a 3DVision® camera, sending both programs simultaneously with TDVision® stereoscopic pair identifiers. This type of decoding is known as "by parallel images", consisting in storing both left and right (L and R) video streams simultaneously as two independent video streams, but time\_code-synchronized. Later, they will be decoded and played back in parallel. Only the decoding software must be decoded, the coding and the compression algorithm of the transport stream will be identical to the current one.

**[0090]** Software modifications in the decoder.

**[0091]** In the decoder, two program streams must be programmed simultaneously, or two interdependent video signals, i.e., constructed from the difference between both stored as a B type frame with an identifier, following the programming API as in the example case, in the use of the TMS320C62X family Texas Instruments DSP.

**[0092]** DSP's programming algorithm and method.

- Create two process channels when starting the DSP (primary and secondary buffers or left and right when calling API).
- Get the RAM memory pointers for each channel (RAM addresses in the memory map)
- When a TDVision® type video sequence is obtained it is taken as B type
  - the image is decoded in real-time
  - the change or difference is applied to the complementary buffer
  - the results are stored in the secondary buffer.

**[0093]** In that related to the software in the video\_sequence data stream, two options are implemented:

- 1.- One modifies only the software and uses the user\_data() section to store the error correction that allows to regenerate the stereoscopic signal.
- 2.- The other enables by hardware the PICTURE\_DATA3D() function which is transparent to MPEG2-compatible readers, and which it can be decoded by a TDVision®-compatible DSP.

**[0094]** At the moment that the MPEG2 decoder detects a user\_data() code, it will search the 3DVISION\_START\_IDENTIFIER = OXOOOABCD 32-bit identifier, which is an extremely high and difficult to reproduce code, or which does not represent data. Then, the 3D block length to be read will be taken into account, which is a 32-bit "n" data. When this information is detected within the USER\_DATA(), a call to the special decoding function will be made which is then compared to the output buffer and applied from the current read offset of the video\_sequence, the n bytes as a typical correction for B type frames. The output of this correction is sent to other output address, which is directly associated to a video output additional to that existing in the electronic display device.

**[0095]** If the PICTURE\_DATA3D() structure is recognized, then it proceeds to read the information directly by the

decoder; but it writes the information in a second output buffer, which is also connected to a video output additional to that existing in the electronic display device.

[0096] In case of the program stream, two signals (left and right) are synchronized by the time\_code, which will be decoded in parallel by a MPEG decoder with enough simultaneous multiple video channels decoding capability, or which can send two interdependent video signals within the same video\_sequence, e.g., "R-L = delta", where delta is the difference stores as a "B" type frame with stereoscopic pair TDVision® identifier and which can be reconstructed at the moment of the decoding by differences from the image, i.e., "R-delta = L" or "L-delta = R", as in the case of the aforementioned Texas Instruments DSP, which is considered as an illustrative but not limiting example.

[0097] A video containing a single video sequence is also implemented; but alternating the left and right frames at 60 frames per second (30 frames each) and when decoded place the video buffer image in the corresponding left or right channel.

[0098] It will also have the capacity of detecting via hardware if the signal is of TDVision® type, if this is the case, it will be identified if it is a transport stream, program stream or left-right multiplexion at 60 frames per second.

[0099] In the case of the transport stream the backward compatibility system is available in the current decoders, having the ability to display the same video without 3d characteristics but only in 2D, in which case the DSP is disabled to display the image in any TDVision® or previous technique device.

[0100] In the case of the program stream unmodified coders are used, such as those currently used in satellite transmission systems; but the receptor and decoder have a TDVision® flag identification system, thus enabling the second video buffer to form a left-right pair.

[0101] Finally, in the case of multiplexed video, the MPEG decoder with two video buffers (left-right) is enabled, identifying the adequate frame and separating each signal at 30 frames per second, thus providing a flickerless image, as the video stream is constant and due to the characteristic retention wave of the human eye the multiplexion effect is not appreciated.

[0102] Particular embodiments of the invention have been illustrated and described, it will be obvious for those skilled in the art that several modifications or changes can be made without departing from the scope of the present invention. All such modifications and changes are intended to be covered by the following claims, so that all changes and modifications fall within the scope of the present invention.

## Claims

1. A stereoscopic 3D-video image digital decoding system and method, in which the structures of the video\_sequence of the video data stream are modified via software, to include the necessary flags at the bit level for the image type, **characterized by** only modifying the software and by using the user\_data() section to store the error correction which allows to regenerate the stereoscopic video signal, thereby actually identifying the video format; applying a logical "and" for MPEG2 backward compatibility in case it is not a TDVision® video; typically decoding by scanning the video\_sequence; when the image is a TDVision® image:

- a) storing the last complete image buffer in the left or right channel buffer.
- b) applying the differences or parallel decoding for B type frame information,
- c) applying error correction to the last image obtained by applying the motion and color correction vectors,
- d) storing the results in their respective channel buffer,
- e) continuing with the video\_sequence reading.

2. Stereoscopic 3D-video image digital decoding method and system, in which the video\_sequence structures of the video data stream are modified via software to include the necessary flags at the bit level of the image type according to Claim 1, further **characterized by** the decoder compilation format being as follows:

- a) reading video\_sequence ,
- b) discriminating the sequence\_header, if a TDVision® image is identified, then activating the double buffer,
- c) reading in the user\_data the image as if it was contained in said structure,
- d) adding in the sequence\_scalable\_extension information to the video\_sequence MPEG, said information could be contained within said structure,
- e) finding in the picture\_header the TDVision® image identifier in the extra\_bit\_picture,
- f) reading the "B" type image in the picture\_coding\_extension, and if it is a TDVision® type image, decoding then the second buffer,
- g) if the image is temporarily scalable, applying "B" to the decoder.

3. Stereoscopic 3D-video images digital decoding method and system, in which the structures and the video\_sequence of the video data stream are modified to include the necessary flags at the bit level of the image type according to Claim 1, further **characterized in that** when the decoder detects a user\_data() code, it searches the 32-bit 3DVision®\_start\_identifier = 0x000ABCD identifier, upon detecting this information a call is made to the special decoding function which compares the output buffer and applies it from the current reading offset of the video\_sequence.
4. Stereoscopic 3D-video images digital decoding method and system, in which the video\_sequence structures of the video data stream are modified via software to include the necessary flags at the bit level of the image type according to Claim 1, further **characterized in that** the decoder must be programmed via software to simultaneously receive and decode two program streams.
5. Stereoscopic 3D-video images digital decoding method and system, in which the video\_sequence structures of the video data stream are modified via software to include the necessary flags at the bit level of the image type according to Claim 1, further **characterized in that** two interdependent video signals can be sent within the same video\_sequence; said signals depending one from the other, and coming from a 3DVision® camera; in terms of their algebraic addition (R-L=delta), each signal is stored as a B type frame, which decoding is by differences from one of them.
6. Stereoscopic 3D-video images digital decoding method and system, in which the video\_sequence structures of the video data stream are modified via software to include the necessary flags at the bit level of the image type according to Claim 1, further **characterized in that** two independent video streams L and R are stored in simultaneous form, but being synchronized with the same time\_code, and decoded and displayed in parallel.
7. Stereoscopic 3D-video images digital decoding method and system, in which the video\_sequence structures of the video data stream are modified via hardware, **characterized by** the specific use of the structures, substructures and sequences belonging to the video\_sequence to implement the MPEG2 backward-compatible TDVision® technology via hardware, in effect, discriminating whether it is a 2S or 3D signal; activating a double output buffer (additional memory); activating a parallel decoding selector, activating a difference- decoding selector; executing the image decompression process, displaying the image in the corresponding output buffer; enabling the PICTURE\_DATA3D() function, which is transparent for the compatible MPEG2 readers.
8. Stereoscopic 3D-video images digital decoding method and system, in which the video\_sequence structures of the video data stream are modified via hardware according with Claim 7, **characterized by** the specific use of the structures, substructures and sequences belonging to the video\_sequence in order to implement the MPEG2 backward-compatible TDVision® technology via hardware:
- a) sequence\_header
    - aspect\_ratio\_information
    - 1001 n/a in TDVision®
    - 1010 4:3 in TDVision®
    - 1011 16:9 in TDVision®
    - 1100 2.21:1 in TDVision®
    - a logical "and" with 0111 will be executed to obtain backward compatibility with 2D systems, where an instruction is sent to the DSP stating that the stereoscopic pair buffer (left or right) must be equal to the source;
  - b) frame\_rate\_code
    - 1001 24,000/1001 (23.976) in TDVision® format
    - 1010 24 in TDVision® format
    - 1011 25 in TDVision® format
    - 1100 30,000/1001 (29.97) in TDVision® format
    - 1101 30 in TDVision® format
    - 1110 50 in TDVision® format
    - 1111 60,000/1001 (59.94) in TDVision® format
    - a logical "and" with 0111 will be executed to obtain backward compatibility with 2D systems, where an instruction is sent to the DSP stating that the stereoscopic pair buffer (left or right) must be equal to the source;
  - c) user\_data()
  - sequence\_scalable\_extension
  - d) picture\_header



extra\_bit\_picture  
0 = TDVision®  
1 = normal  
e) picture\_coding\_extension  
5 picture\_structure  
00 = image in TDVision® format  
f) picture\_temporal\_scalable\_extension().

10 9. Stereoscopic 3D-video images digital decoding method and system, in which the video\_sequence structures of the video data stream are modified via hardware according with Claim 7, **characterized in that**, when the PICTURE\_DATA3D()structure is recognized, it proceeds to read the information directly by the decoder, but it writes the information in a second output buffer also connected to a video output additional to that existing in the electronic display device.

15 10. Stereoscopic 3D-video images digital decoding method and system, in which the video\_sequence structures of the video data stream are modified via hardware according with Claim 7, **characterized in that**, if the signal is of TDVision® type, it is identified if it is a transport stream, program stream or left or right multiplexion at 60 frames per second; when it is a transport stream it has backward compatibility in the current 2D coders; where an instruction is sent to the DSP stating that the stereoscopic pair buffer (left or right) must be equal to the source, having the  
20 ability to display the video without 3D characteristics of TDVision®.

25

30

35

40

45

50

55

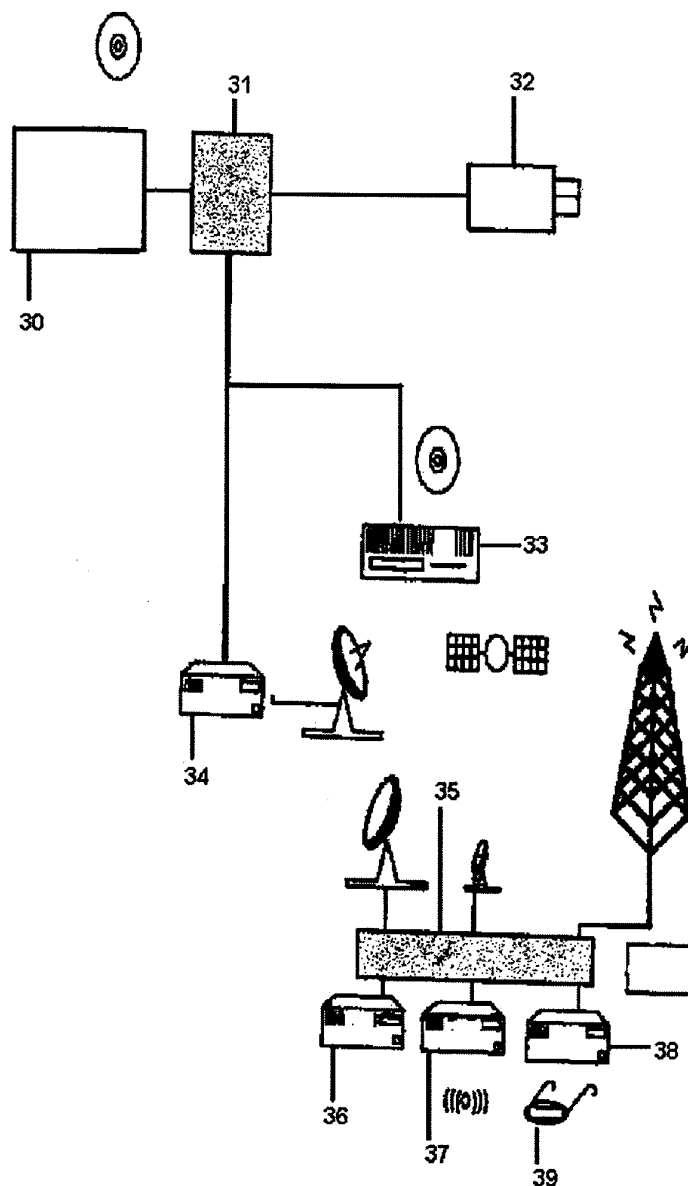
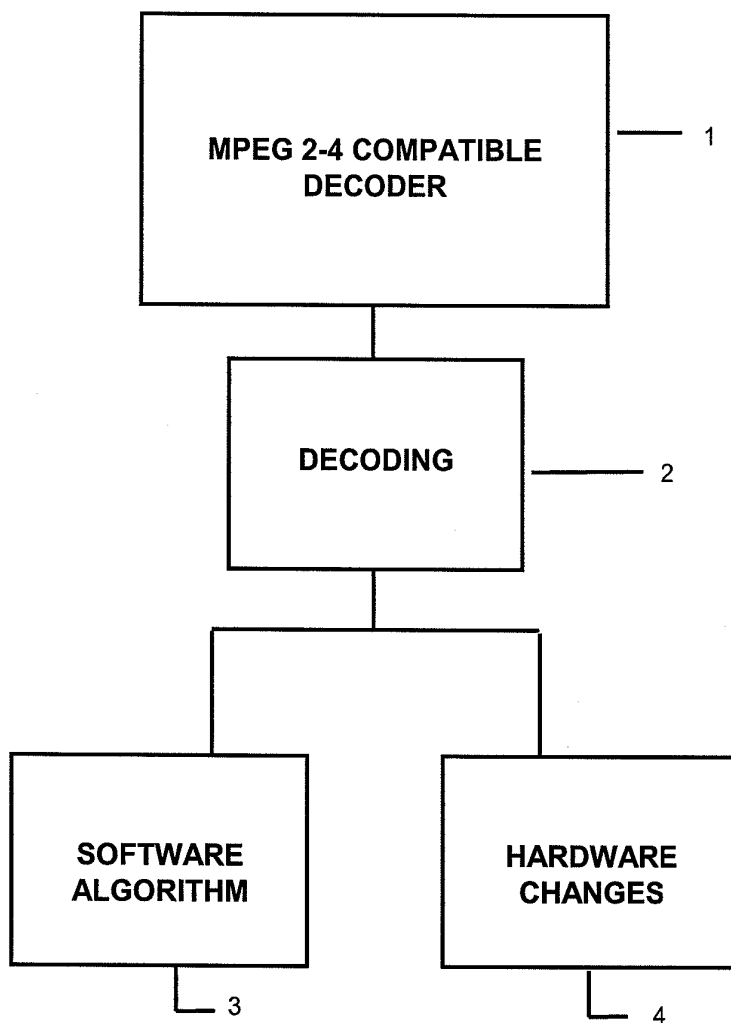
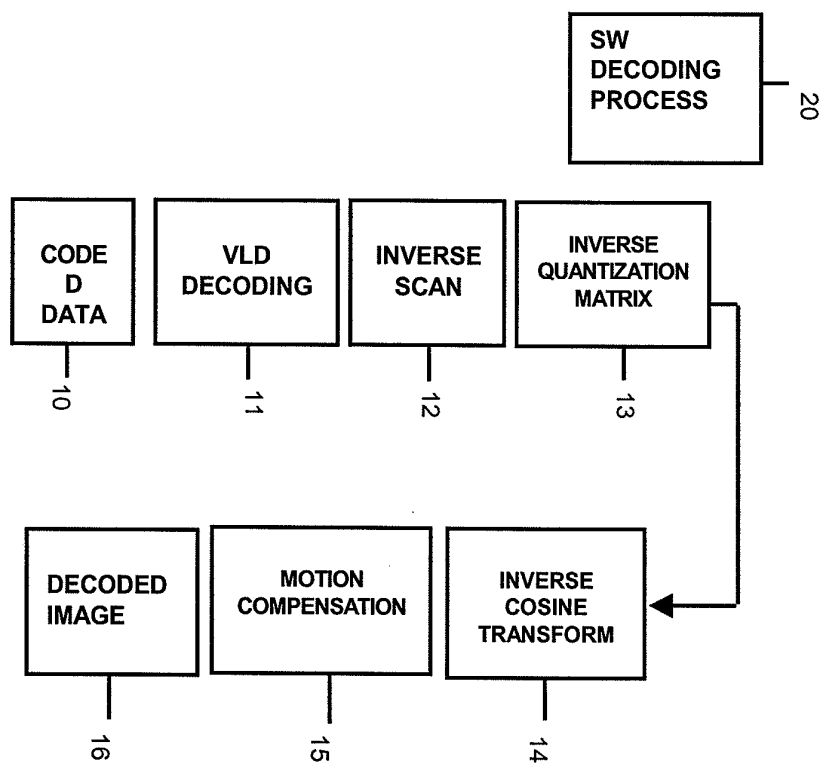


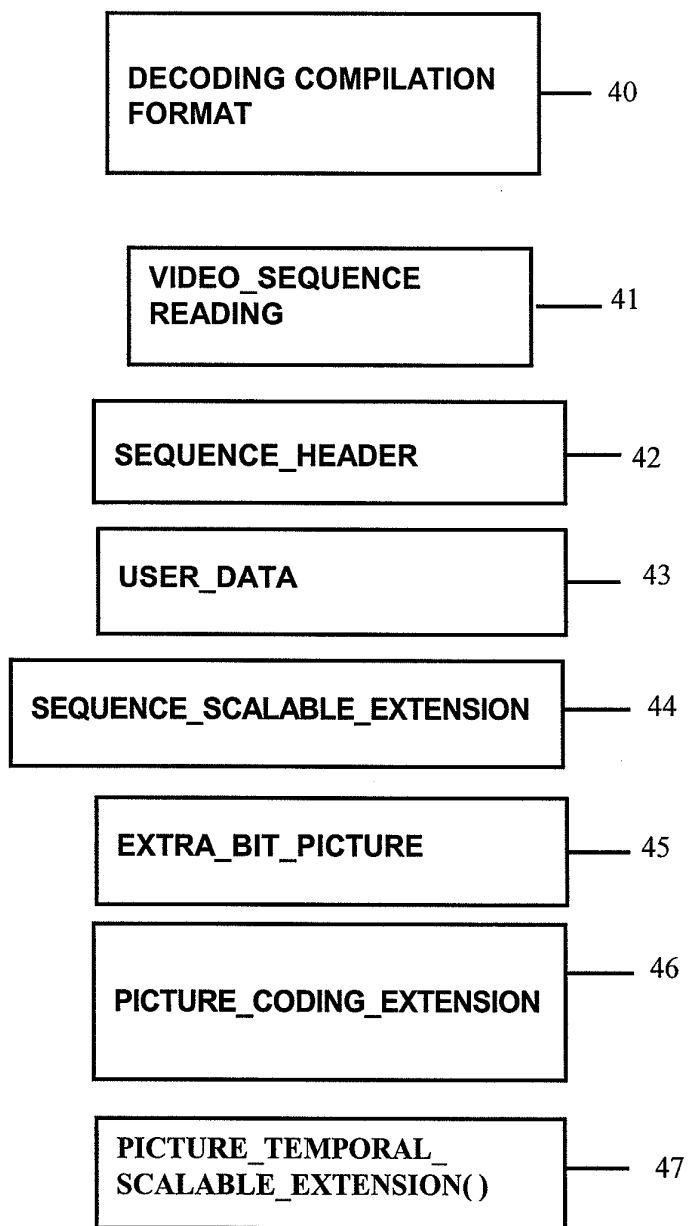
Fig. 1



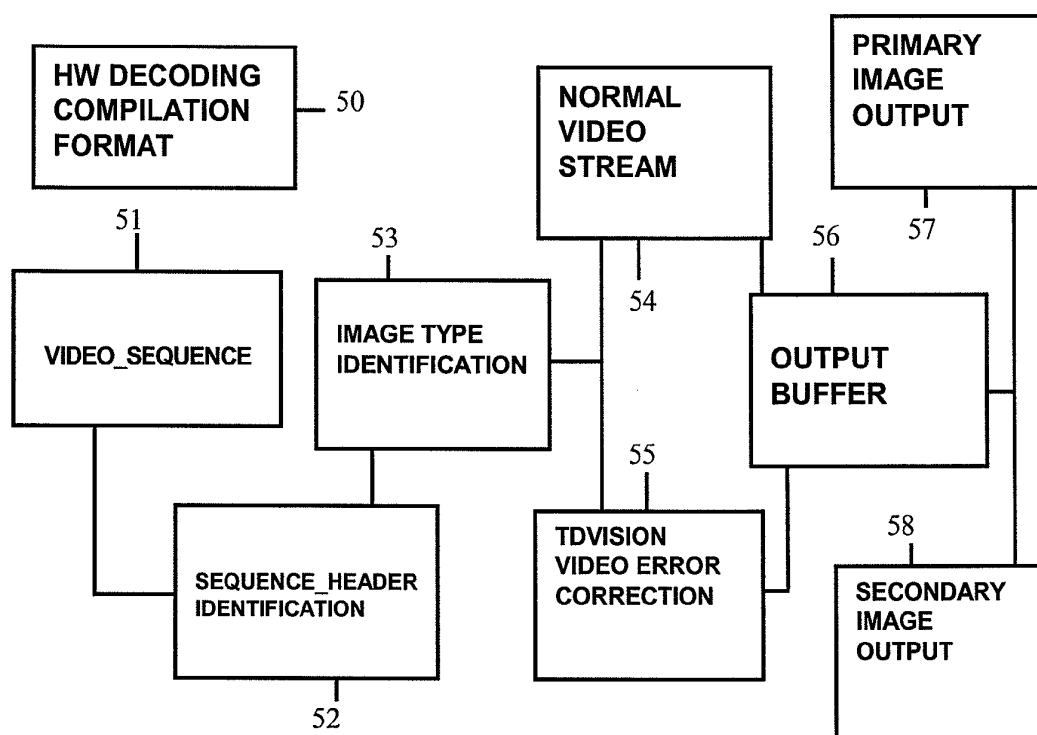
**Fig. 2**



**Fig. 3**



**Fig. 4**



**Fig. 5**

## INTERNATIONAL SEARCH REPORT

International application No. PCT/ MX 2004/000012
--

A. CLASSIFICATION OF SUBJECT MATTER		
<b>IPC<sup>7</sup> G06T 9/00, G06T 15/00, H04N 13/00</b> According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols)		
<b>IPC<sup>7</sup> G06T+, H04N+</b>		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
<b>CIBEPAT, EPODOC, PAJ, WPI, INSPEC</b>		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2004008893 A (SATOSHI ITOI et al.) 15.01.2004, <b>paragraphs 13-16, 69-72, 79-88, 149-170; figures 1,3,4.</b>	1
A	US 2003190079 A (GOUTELLE GILLES et al.) 09.10.2003, <b>paragraphs 18-32; figures.</b>	1
A	WO 9817068 A (SARNOFF CORPORATION) 23.04.1998, <b>page 4, line 21-page 6, line 7; figure 2.</b>	1
A	US 6370193 B (LEE YONG-WOON) 09.04.2002, <b>column 4, line 21-column 8, line 19; figures 2-8.</b>	1
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
<b>24 August 2004 (24.08.2004)</b>		<b>15 September 2004 (15.09.2004)</b>
Name and mailing address of the ISA/		Authorized officer
<b>S.P.T.O.</b>		
Facsimile No.		Telephone No.

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International Application No  
PCT/ MX 2004/000012

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2004008893 A	15.01.2004	JP 2004048293 A	12.02.2004
US 2003190079 A	09.10.2003	WO 0176257 A CN 1365577 T EP 1269761 A JP 2003530038 T	11.10.2001 21.08.2002 02.01.2003 07.10.2003
WO 9817068 A	23.04.1998	EP 0931420 A US 6144701 A JP 2001502504 T	28.07.1999 07.11.2000 20.02.2001
US 6370193 B	09.04.2002	JP 10327407 A KR 273110 B	08.12.1998 01.12.2000



**REFERENCES CITED IN THE DESCRIPTION**

*This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.*

**Patent documents cited in the description**

- US 5963257 A, Katata [0017]
- US 6292588 B, Shen [0018]
- US 6370276 B, Boon [0018]
- US 6456432 B, Lazzaro [0019]
- US 6658056 B, Duruoz [0020]
- US 6665445 B, Boon [0021]
- US 6678331 B, Moutin [0022]
- US 6678424 B, Ferguson [0023]