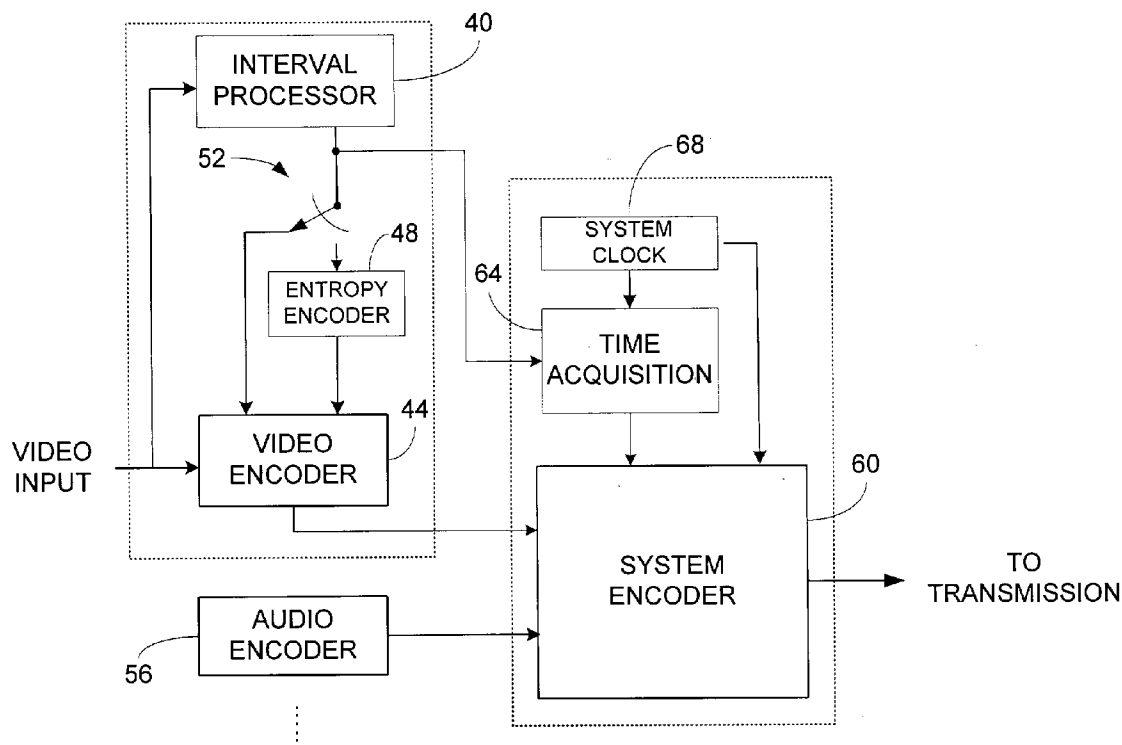


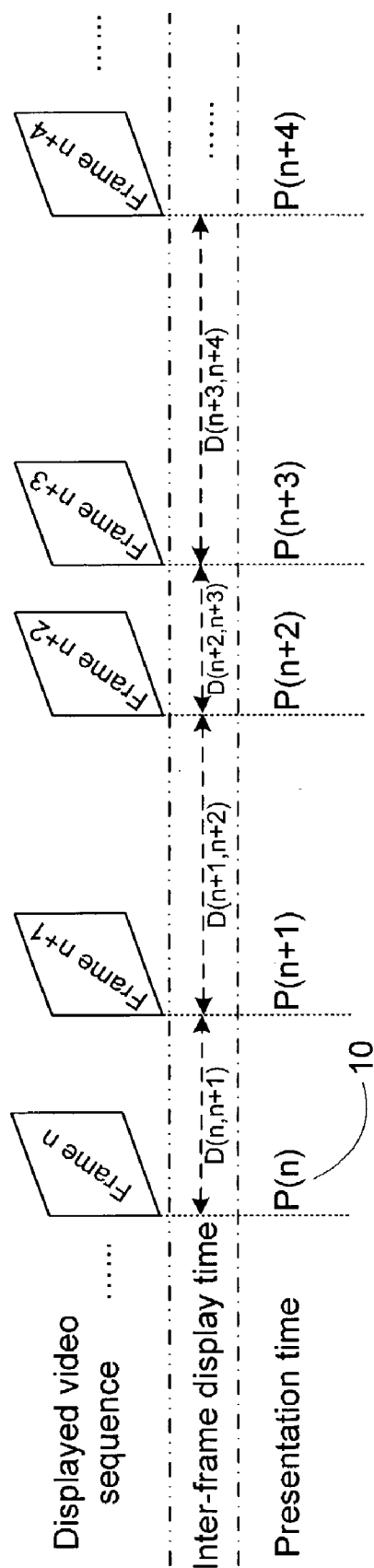


US 20030138051A1

(19) **United States**(12) **Patent Application Publication** (10) **Pub. No.: US 2003/0138051 A1****Chen et al.**(43) **Pub. Date: Jul. 24, 2003**(54) **SYSTEM AND METHOD OF TRANSMISSION
AND RECEPTION OF VIDEO USING
COMPRESSED DIFFERENTIAL TIME
STAMPS**(76) Inventors: **Sherman (Xuemin) Chen**, San Diego,
CA (US); **Alexander G. MacInnis**, Los
Altos, CA (US)Correspondence Address:
MCANDREWS HELD & MALLOY, LTD
500 WEST MADISON STREET
SUITE 3400
CHICAGO, IL 60661(21) Appl. No.: **10/349,705**(22) Filed: **Jan. 22, 2003****Related U.S. Application Data**(60) Provisional application No. 60/351,182, filed on Jan.
22, 2002. Provisional application No. 60/384,234,
filed on May 30, 2002. Provisional application No.
60/441,072, filed on Jan. 17, 2003.**Publication Classification**(51) **Int. Cl.⁷** **H04N 7/12**
(52) **U.S. Cl.** **375/240.26; 375/240.15**(57) **ABSTRACT**

Systems and methods that use differential values to represent decoding and presentation times of pictures are provided. Optionally, some of the systems or the methods may use an entropy coding method such as, for example, the Universal Variable Length Code (UVLC) for the differential time stamps.





Where $P(n+1)=P(n)+D(n, n+1)$, for $k < n+1$;
 $D(k, n+1)=D(k, k+1)+D(k+1, k+2)+\dots+D(n, n+1)$

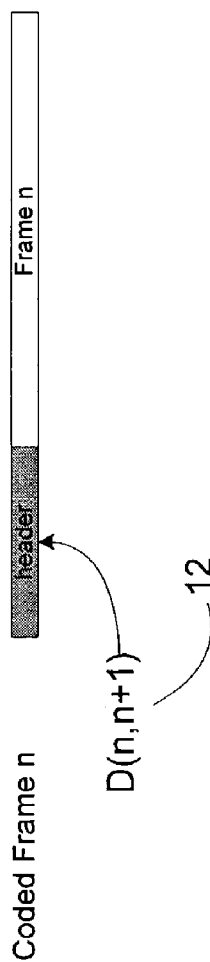


FIG. 1

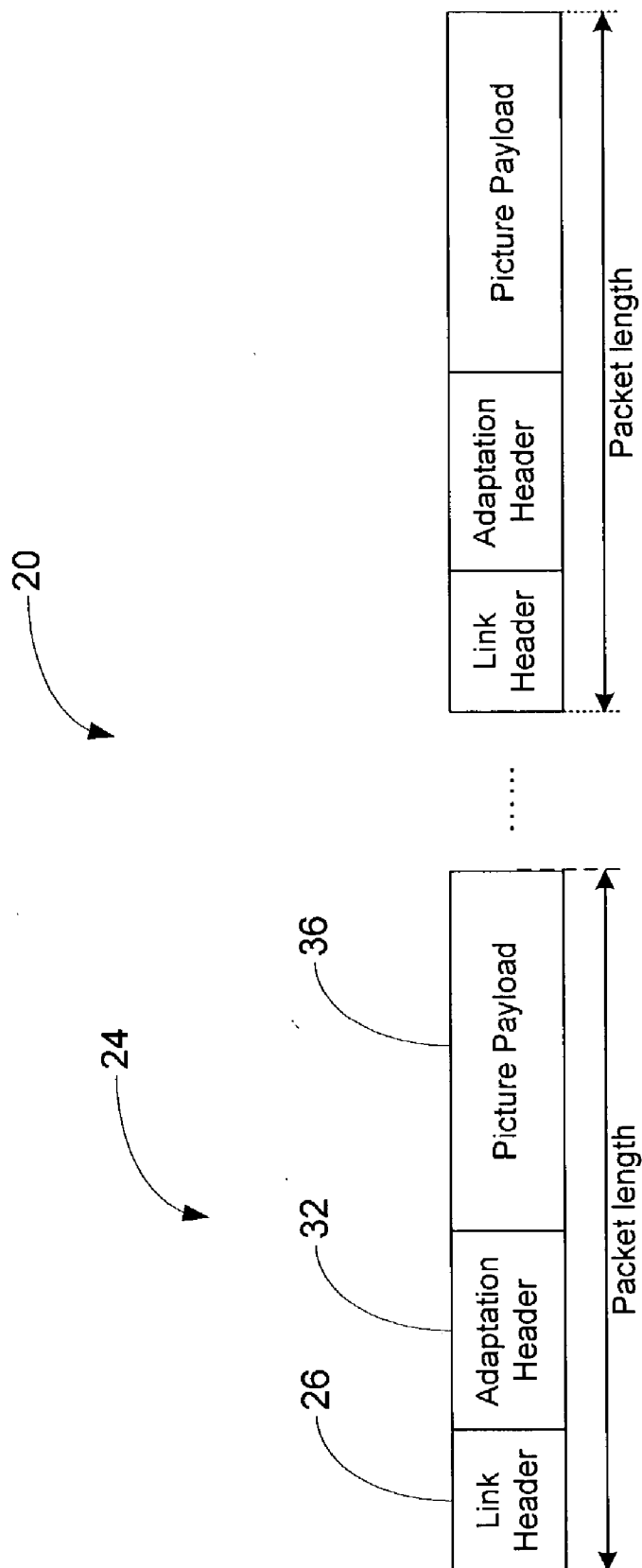


FIG. 2

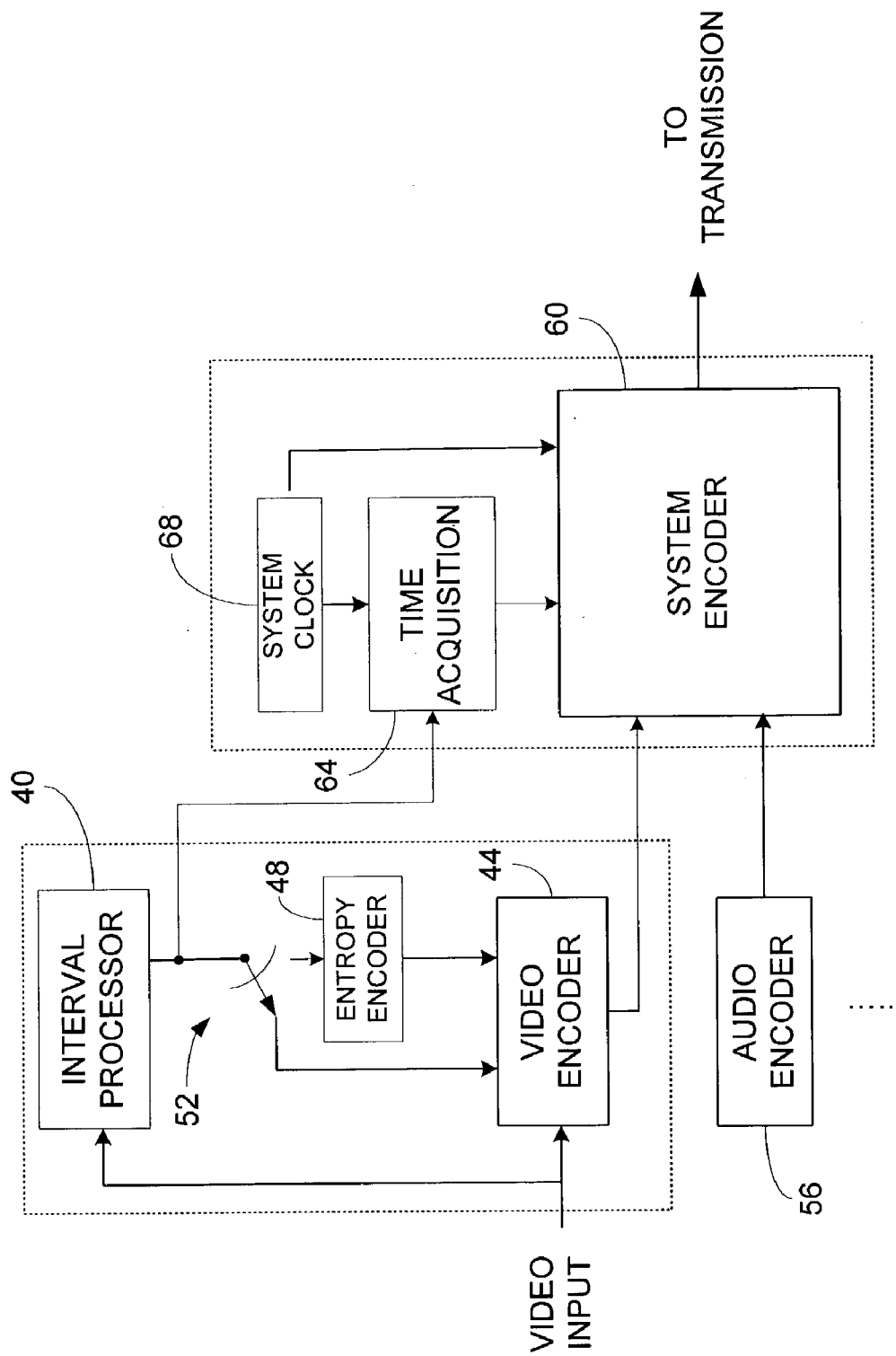


FIG. 3

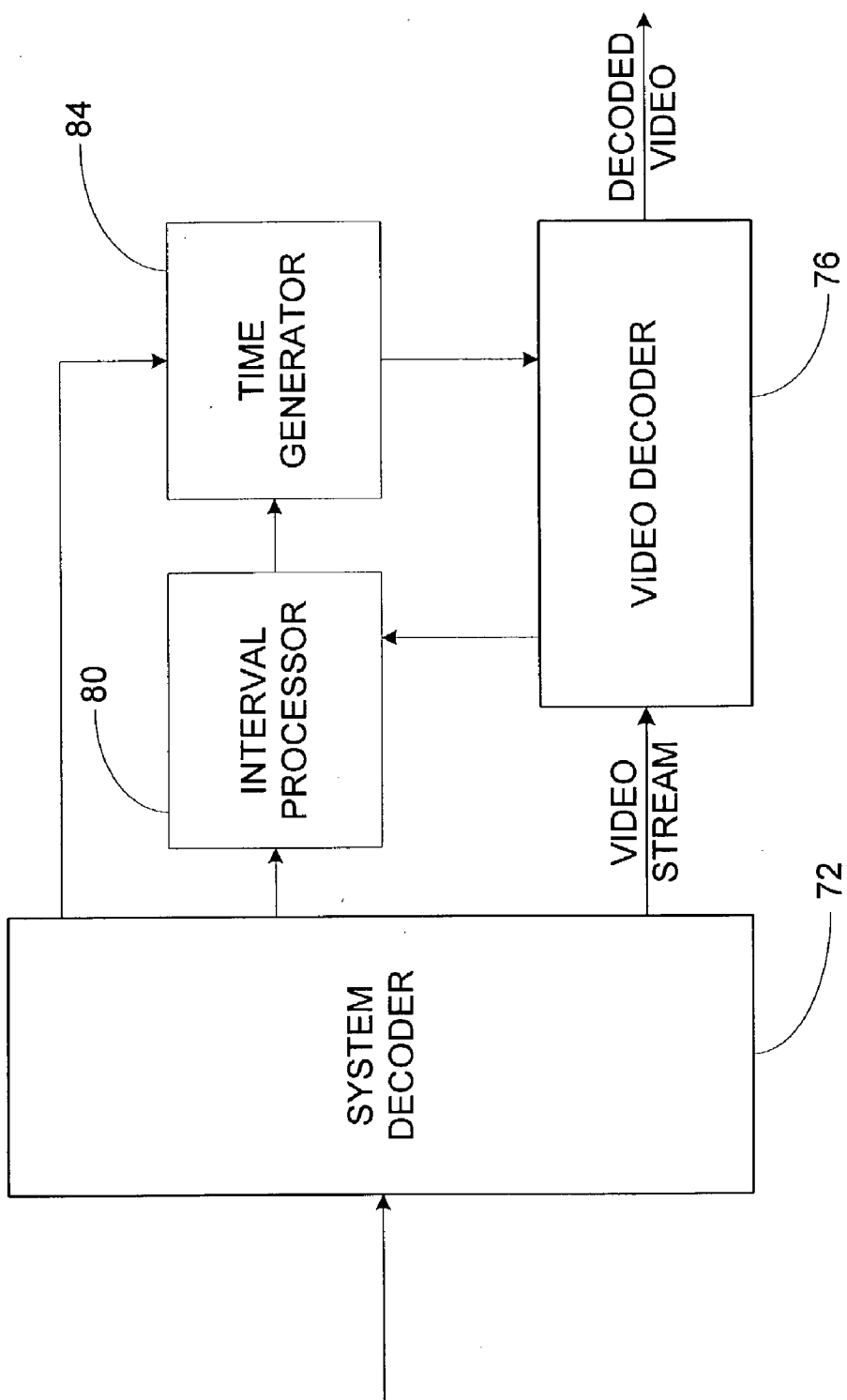


FIG. 4

SYSTEM AND METHOD OF TRANSMISSION AND RECEPTION OF VIDEO USING COMPRESSED DIFFERENTIAL TIME STAMPS

RELATED APPLICATIONS

[0001] This application makes reference to, claims priority to and claims benefit from U.S. Provisional Patent Application Serial No. 60/351,182, entitled "Method for Enabling Transmission and Display of Encoded Compressed Film Material in a Video Bit Stream" and filed on Jan. 22, 2002; U.S. Provisional Patent Application Serial No. 60/384,234, entitled "Method for Enabling Transmission and Display of Encoded Compressed Film Material in a Video Bitstream" and filed on May 30, 2002; and U.S. Provisional Patent Application Serial No. 60_____, entitled "Method for Enabling Transmission and Display of Encoded Compressed Film Material in a Video Bit Stream" and filed on Jan. 17, 2003.

INCORPORATION BY REFERENCE

[0002] The above-identified United States patent applications are hereby incorporated herein by reference in their entirety.

BACKGROUND OF THE INVENTION

[0003] For most applications, digital video is displayed at a known frame rate or at a known field rate. For example, in countries that use 525-line interlaced display systems such as, for example, the United States and Canada, television video signals are sampled and transmitted at approximately 59.94 fields per second (fps). For such countries, digital television video streams are generally encoded and transmitted by using a particular Moving Pictures Experts Group (MPEG) standard (e.g., MPEG-2 Video) at approximately 29.97 frames per second (FPS).

[0004] Hereinafter, an integral value of fps or an integral value of FPS may be an approximation including, within its scope, a range of equivalent values. Thus, for example, the expression 30 FPS may be used to refer to rates such as, for example, approximately 29.97 FPS or approximately 30 FPS. Furthermore, the expression 24 FPS may be used to refer to rates such as, for example, approximately 23.976 FPS or approximately 24 FPS. Similarly, the expression 60 fps may be used to refer to rates such as, for example, approximately 59.94 fps or approximately 60 fps.

[0005] A given frame rate implies a fixed inter-picture display time interval. Therefore, the frame rate can be used to derive the decoding and presentation times for some or all of the pictures in a video sequence.

[0006] In the MPEG-1 Systems and MPEG-2 Systems standards, there are syntax elements such as, for example, a decoding time stamp (DTS) and a presentation time stamp (PTS) which specify the decoding time and presentation time, respectively, of some pictures in terms of a Hypothetical Reference Decoder model called a System Target Decoder (STD). The decoding time refers to the time that a compressed picture is removed from a buffer in the STD model. This is not necessarily the exact time that a practical decoder decodes the picture. The presentation time in the STD model is the time a picture is presented. This may be taken to mean the display time in a real decoder, although

practical decoders may display pictures at slightly different times than those indicated by the video stream. The MPEG-2 Video standard (i.e., ISO/IEC 13818-2) specifies a VBV delay syntax element that indicates similar information to that contained in the DTS in the Hypothetical Video Decoder model and represents the delay from the time certain compressed data enters a buffer in the model to the time a given picture is extracted from that buffer in the model. In the MPEG-1 and MPEG-2 standards, the DTS, PTS and VBV_delay syntax elements specify times in units of a 90 kHz reference clock. The principle of providing decoding and presentation times can apply to any video coding and decoding standard regardless of the syntax specification and the units of time reference. Hereinafter, the terms presentation time and display time are to be used interchangeably, and the terms decoding time and buffer removal time are to be used interchangeably.

[0007] However, for some video applications, the intervals between successive pictures may not be a constant or fixed. It may be desirable for a video coding standard to support dynamically variable frame rates or different types and durations of pictures interspersed to form a video sequence. Thus, variable inter-picture decoding and display times need to be supported. Inter-picture display time can no longer be represented properly by a frame rate. It may be considered necessary to encode the presentation times of all pictures, which is not necessary in MPEG-2 due to the use in MPEG-2 of constant frame rates. The inclusion of a presentation time stamp such as PTS for every picture or the inclusion of a decoding time stamp such as DTS for every picture is undesirable because of the number of bits (e.g., 33 bits per time stamp) and the syntax layers used for encoding in the MPEG-2 standard. Thus, conventional systems and methods may lack an efficient method to represent and to encode presentation times with either fixed or variable inter-picture display intervals.

[0008] The inter-picture decoding interval is, in general, not the same as the inter-picture display interval. For example, in MPEG-2 Video while the frame rate is constant for a given sequence, the interval between the decoding times of successive pictures, represented by DTS or by the VBV_delay field, may vary independently every picture. However, in MPEG-1 and MPEG-2, the decoding time is specified with reference to a time that specific data elements enter a buffer in a Hypothetical model, which requires precise specifications of when each data element enters the Hypothetical buffer. Precise specifications may not be available in all cases. Precise specifications also tend to lead to large syntax elements to specify the decoding time and complexity in specifying and interpreting the DTS and VBV_delay values.

[0009] Further limitations and disadvantages of conventional and traditional approaches will become apparent to one of ordinary skill in the art through comparison of such systems with some aspects of the present invention as set forth in the remainder of the present application with reference to the drawings.

BRIEF SUMMARY OF THE INVENTION

[0010] Some aspects of the present invention may be found in, for example, systems and methods that use differential values to represent decoding and presentation times of

pictures and that, optionally, use an entropy coding method such as, for example, a Universal Variable Length Code (UVLC), for the differential time stamps.

[0011] In one embodiment, the present invention may provide that a display time of a picture may be represented as the difference between the presentation times of the picture and the presentation times of an adjacent picture in display order. The differential time values may be coded by using entropy coding (e.g., UVLC) or by using other methods such as, for example, fixed length codes.

[0012] In another embodiment, the present invention may provide that a decoding time of a picture may be represented by the difference between the decoding times of the picture and the decoding times of an adjacent picture in coded data order. The differential decoding times may be encoded by the use of entropy coding (e.g., UVLC) or by using other methods such as, for example, fixed length codes.

[0013] These and other features and advantages of the present invention may be appreciated from a review of the following detailed description of the present invention, along with the accompanying figures in which like reference numerals refer to like parts throughout.

BRIEF DESCRIPTION OF THE DRAWINGS

[0014] FIG. 1 shows an embodiment of a video sequence according to the present invention.

[0015] FIG. 2 shows an embodiment of a system stream according to the present invention.

[0016] FIG. 3 shows an embodiment of an encoding system according to the present invention.

[0017] FIG. 4 shows an embodiment of a decoder system according to the present invention.

DETAILED DESCRIPTION OF THE INVENTION

[0018] Some embodiments according to the present invention provide an efficient and flexible method to represent and to encode decoding times or presentation times using either fixed intervals or variable intervals between pictures. Some embodiments according to the present invention may be suitable for use in or compatible with new video compression standards (e.g., MPEG-4 AVC standard as described in, for example, document JVT-E146d37ncm.doc, entitled "Joint Video Specification (ITU-T Rec. H.264 [ISO/IEC 14496-10 AVC]", incorporated herein by reference in its entirety) as well as existing video transport standards (e.g., MPEG-2 Systems, IETF RTP, etc.) The MPEG-2 Systems standard may also be known as ISO/IEC 13818-1, and the MPEG-2 Video standard may also be known as ISO/IEC 13818-2; both of which are incorporated by reference herein in their entirety.

[0019] Some embodiments according to the present invention may relate to a system or a method that uses differential values to represent picture display times or presentation times and to represent picture decoding times or buffer removal times. Optionally, some embodiments may provide for the use of entropy coding such as, for example, Universal Variable Length Code (UVLC) for the differential time values. UVLC may also be known as exp-golomb coding. One embodiment of UVLC, referred to as exp-golomb

codes, is described in the MPEG-AVC draft standard previously reference. However, it is understood that the present invention also contemplates using other embodiments of UVLC. Some embodiments according to the present invention may be suitable for use in or compatible with some MPEG standards (e.g., MPEG-4 AVC).

[0020] Some embodiments according to the present invention may relate to a system or a method that uses differential values to represent picture presentation times or picture decoding times. Optionally, some embodiments may include, for example, the use of an entropy coding method such as, for example, UVLC to compress the differential time interval values. Some embodiments may be suitable for use in or compatible with some MPEG standards (e.g., MPEG-4 AVC).

[0021] In one embodiment, the present invention may provide that presentation times of pictures may be represented as differences between the presentation times of two adjacent pictures in presentation order. An embodiment of a video sequence illustrating differentially represented presentation times according to the present invention is shown in FIG. 1. Although inter-picture display time intervals are illustrated as variable, the present invention may be applied to variable time intervals as well as to constant time intervals between frames, pictures or packets.

[0022] In presentation order, the presentation time of the (n+1)-th picture may be generated by using, for example, the presentation time 10 of the n-th picture and the inter-picture presentation interval or time difference 12 between the n-th picture and the (n+1)-th picture. In one embodiment, the presentation time of the (n+1)-th picture may be determined by adding the presentation time 10 of the n-th picture to the inter-picture interval or time difference 12 of the n-th picture. The presentation time of the n-th picture may be encoded by a variety of means such as, for example, via an implied time at the start of a sequence of pictures or a non-differential encoding of the presentation time of the n-th picture. An example of a non-differential coding of presentation time is the PTS specification in the MPEG-2 Systems standard (i.e., ISO/IEC 13818-1). The coded presentation difference time may be included, for example, in the syntax of the picture to which it refers, for example, in the picture header or in other sections of the syntax of the video data such as, for example, in packets that include the compressed video data or in data structures that refer to the picture.

[0023] In another embodiment, the present invention may provide that decoding times of pictures may be represented by differences between the decoding times of two adjacent pictures in coded data order.

[0024] In decoding order, the decoding time of the (m+1)-th picture may be generated by using, for example, the decoding time of the m-th picture and the inter-picture decoding interval or time difference. In one embodiment, the decoding time of the (m+1)-th picture may be determined by adding the decoding time of the m-th picture to the inter-picture decoding interval or time difference between the m-th picture and the (m+1)-th picture. The decoding time of the m-th picture may be encoded by a variety of means such as, for example, via an implied time at the start of a sequence of pictures or a non-differential encoding of the decoding time of the m-th picture. An example of a non-differential coding of decoding time is the DTS specification in the

MPEG-2 Systems standard. Another example is the VBV_delay specification in the MPEG-2 Video standard. The coded decoding difference time may be included, for example, in the syntax of the picture to which it refers, for example, in the picture header or in other sections of the syntax of the video data such as, for example, in packets that include compressed video data or in data structures that refer to the picture.

[0025] FIG. 2 shows an embodiment of a system stream carrying coded video according to the present invention. The system stream 20 may include one or more packets 24. FIG. 2 illustrates two packets which are not necessarily adjacent packets in presentation or decoding order. The packet 24 may include, for example, a link header 26, an adaptation header 32 and a picture payload 36. The adaptation header may include information, for example, relating to one or more of the following: a presentation time, a decoding time, a differential presentation time, a differential decoding time, other types of time stamps or other types of differential time stamps. The adaptation header need not be present in every packet 24. Furthermore, even if the adaptation header 32 is present in a particular packet 24, the adaptation header need not always include information relating to any or all of the above-listed times or differential times.

[0026] In one embodiment, some packets 24, which include picture payloads 36 including information relating to intra-coded pictures, may include adaptation headers 32 that may include, for example, information relating to absolute or non-differential time values. The use of absolute values may be beneficial for purposes such as, for example, random access, fast forward, rewind playback, record and error-resilience. Some packets 24, which include payloads 36 including information relating to inter-coded pictures, may include adaptation headers 32 that may include, for example, information relating to inter-picture or differential time values.

[0027] In another embodiment, the picture payload 36 of the first packet may include, for example, an I-picture payload and the picture payload 36 of a later (although not necessarily adjacent) packet may include, for example, a P-picture payload or a B-picture payload. The adaptation header 32 of the first packet may include, for example, information relating to an absolute decoder time or to an absolute presentation time and the adaptation header 32 of the later packet may include, for example, information relating to a differential decoding time or to a differential presentation time.

[0028] In one embodiment, if a decoder encounters a picture without a non-differential decoding or a non-differential presentation time coded in the stream, then the decoder may determine the respective decoding time or the respective presentation time for the picture based upon, for example, a decoding time or a presentation time of a previous picture (e.g., a previous frame picture or a previous field picture). Further information about the respective decoding time or the respective presentation time may be ascertained, for example, from an encoded differential decoding time or an encoded differential display time of an intervening picture or a current picture and from information about picture structure (e.g., field or frame).

[0029] FIG. 3 shows an embodiment of an encoder system according to the present invention. Video is received by the

encoder system and is processed by the interval processor 40 and the video encoder 44. The interval processor 40 determines, for example, the inter-picture display intervals and decoding intervals from the video. The inter-picture display intervals and the decoding intervals are optionally entropy encoded (e.g., variable length coding or fixed length coding) via the entropy encoder 48. Whether entropy encoding occurs depends upon the position of a switch 52. Depending upon the position of the switch 52, the output of the interval processor 40 or the output of the entropy encoder 48 is received by the video encoder 44 and encoded into the video. Encoded video and encoded audio are sent from the video encoder and an audio encoder 56 to a system encoder 60. The inter-picture display intervals and the decoding intervals are sent from the interval processor 40 to the time acquisition block 64. The time acquisition block 64 acquires the decoding time and the presentation time using the inter-picture display intervals and the decoding intervals and the system clock signal provided by the system clock 68. The system encoder 60 then encodes the video and the audio with information including, for example, one or more of the following: differential decoding times, differential presentation times, non-differential decoding times, non-differential presentation times or other timing or differential timing information. The video stream including the audio stream is then transmitted over a network (not shown) as a video bitstream.

[0030] FIG. 4 shows an embodiment of a decoder system according to the present invention. A system decoder 72 receives the video bitstream via the network. The system decoder 72 demuxes or parses the video bitstream. The video stream is sent to a video decoder 76. A portion of the video bitstream relating to the inter-picture display intervals and the decoding intervals are sent to an interval processor 80. In one embodiment, the portion of the video bitstream parsed out for delivery to the interval processor 80 includes, for example, the adaptation header 32. The interval processor 80 can extract the inter-picture display intervals and the decoding intervals from the parsed video bitstream. Where the differential time values are not stored in system encoding, the interval processor 80 may have to extract the differential time values from the video stream. In one embodiment, differential values can be extracted from the system layer or the video layer depending upon where the differential values are encoded. Accordingly, the video stream may also be provided by the video decoder 76 to the interval processor 80. Once extracted, the inter-picture display intervals and the decoding intervals are sent to the time generator 84. The time generator 84 can generate decoding times and presentation times which are then sent to the video decoder 76. The video decoder 76 decodes the video stream according to the received decoding times and presents the decoded video stream according to the received presentation times.

[0031] The system decoder 72 also parses out portions of the video bitstream that have been encoded with non-differential time values (e.g., absolute presentation times or absolute decoding times). These portions of the video bitstream are sent directly to the time generator 84. The time generator 84 then generates the decoding times and the presentation times from the received parsed video bitstream and sends the information to the video decoder 76. The video decoder 76 decodes the video stream according to the

received decoding times and presents the decoded video stream according to the received presentation times.

[0032] While the present invention has been described with reference to certain embodiments, it will be understood by those skilled in the art that various changes may be made and equivalents may be substituted without departing from the scope of the present invention. In addition, many modifications may be made to adapt a particular situation or material to the teachings of the present invention without departing from its scope. Therefore, it is intended that the present invention not be limited to the particular embodiment disclosed, but that the present invention will include all embodiments falling within the scope of the appended claims.

What is claimed is:

1. A method for facilitating video display, comprising:

embedding timing information into a bitstream, the timing information comprising information related to at least one of a differential decoding time and a differential presentation time; and

determining a decoding time or a presentation time based upon the timing information.

2. The method according to claim 1, wherein embedding the timing information into the bitstream comprises inserting the timing information into a packet of the bitstream.

3. The method according to claim 2, wherein inserting the timing information into the packet comprises inserting the timing information into a header of the packet.

4. The method according to claim 2, wherein the packet comprises a picture payload.

5. The method according to claim 4, wherein the picture payload comprises at least one of an intra-coded picture payload and an inter-coded picture payload.

6. The method according to claim 4, wherein the picture payload comprises at least one of an I-picture payload, a P-picture payload and a B-picture payload.

7. The method according to claim 1,

wherein the differential decoding time comprises a difference between a first decoding time and a second decoding time,

wherein the first decoding time is associated with a first picture of the bitstream, and

wherein the second decoding time is associated with a second picture of the bitstream.

8. The method according to claim 7,

wherein the first picture comprises at least one of a first field picture and a first frame picture, and

wherein the second picture comprises at least one of a second field picture and a second frame picture.

9. The method according to claim 1,

wherein the differential presentation time comprises a difference between a first presentation time and a differential presentation time,

wherein the first presentation time is associated with a first picture of the bitstream, and

wherein the second presentation time is associated with a second picture of the bitstream.

10. The method according to claim 9,

wherein the first picture comprises at least one of a first field picture and a first frame picture, and

wherein the second picture comprises at least one of a second field picture and a second frame picture.

11. The method according to claim 1, further comprising:

applying entropy coding with respect to the differential decoding time or the differential presentation time.

12. The method according to claim 11, wherein applying entropy coding comprises applying a Universal Variable Length Code (UVLC).

13. A method for facilitating video display, comprising:

embedding first timing information into a first packet of a bitstream, the first timing information comprising information related to at least one of a differential decoding time and a differential presentation time.

14. The method according to claim 13, further comprising:

embedding second timing information into a second packet of the bitstream, the second timing information comprising information related to at least one of a non-differential decoding time and a non-differential presentation time.

15. The method according to claim 14, wherein the differential decoding time comprises a difference between a decoding time of the first packet and a decoding time of the second packet.

16. The method according to claim 14, wherein the differential presentation time comprises a difference between a presentation time of the first packet and a presentation time of the second packet.

17. The method according to claim 14, further comprising:

determining a decoding time or a presentation time of the first packet based upon the first timing information and the second timing information.

18. The method according to claim 14, wherein the second packet comprises an intra-coded picture payload.

19. The method according to claim 18, wherein the intra-coded picture payload comprises an I-picture payload.

20. The method according to claim 13, wherein the first packet comprises an inter-coded picture payload.

21. The method according to claim 20, wherein the inter-coded picture payload comprises at least one of a P-picture payload and a B-picture payload.

22. The method according to claim 13, further comprising:

determining a presentation time or a decoding time of the first packet based upon information comprising information relating to picture structure.

23. The method according to claim 13, further comprising:

applying entropy coding with respect to the differential decoding time or the differential presentation time.

24. The method according to claim 23, wherein applying entropy coding comprises applying a UVLC.

25. A system for facilitating video display, comprising:

an encoder system adapted to encode a video bitstream with timing information, the timing information com-

prising information related to at least one of a differential decoding time and a differential presentation time.

26. The system according to claim 25, wherein the encoder system comprises an entropy encoder.

27. The system according to claim 26, wherein the entropy encoder is adapted to apply entropy coding with respect to the differential decoding time or the differential presentation time.

28. The system according to claim 25, further comprising:

a decoder system coupled to the encoder system via a network, the decoder system adapted to determine a decoding time or a presentation time based upon the timing information.

* * * * *