

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
1 February 2007 (01.02.2007)

PCT

(10) International Publication Number
WO 2007/014384 A2

(51) International Patent Classification: **Not classified**

(21) International Application Number:
PCT/US2006/029739

(22) International Filing Date: 27 July 2006 (27.07.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/703,156 28 July 2005 (28.07.2005) US

(71) Applicant (for all designated States except US): **THOMSON LICENSING** [FR/FR]; 46, Quai A. Le Gallo, F-92100 Boulogne-billancourt (FR).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **COMER, Mary, Lafuze** [US/US]; 2571 W 975 S, Fairmount, Indiana 46928 (US). **COOK, Gregory, William** [US/US]; 210 Saw Dust Drive, Lafayette, Indiana 47905 (US).

(74) Agents: **LAKS, Joseph, J.** et al.; Thomson Licensing Inc., Two Independence Way, Suite 200, Princeton, New Jersey 08540 (US).

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

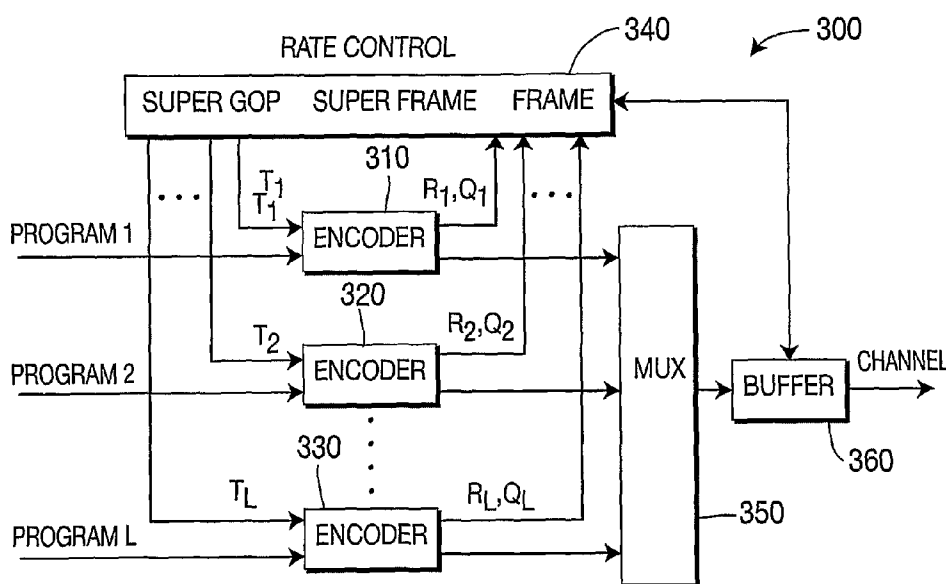
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (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: METHOD AND APPARATUS FOR TRANSMITTING MULTIPLE VIDEO STREAMS OVER A VIDEO CHANNEL



(57) Abstract: There are provided methods and apparatus for transmitting multiple video streams over a video channel. A video encoder is provided for encoding a plurality of streams corresponding to a plurality of programs. The video encoder includes an encoder (300) for encoding each of the plurality of streams for transmission over a communication channel. The encoder imposes a relative positioning on intra-coded frames among the plurality of programs to meet individual stream delay requirements corresponding to the plurality of streams.

WO 2007/014384 A2

METHOD AND APPARATUS FOR TRANSMITTING MULTIPLE VIDEO STREAMS OVER A VIDEO CHANNEL

CROSS-REFERENCE TO RELATED APPLICATIONS

5 This application claims the benefit of U.S. Provisional Application Serial No. 60/703,156, filed July 28, 2005 and entitled "METHOD AND APPARATUS FOR MULTIPLE VIDEO STREAMS OVER A VIDEO CHANNEL," which is incorporated by reference herein in its entirety.

10 FIELD OF THE INVENTION

 The present invention relates generally to video encoding and, more particularly, to methods and apparatus for encoding multiple video streams for transmission over a video channel.

15 BACKGROUND OF THE INVENTION

 When multiple video programs are transmitted over a single channel, the channel bandwidth can be used more efficiently by implementing a joint rate controller to allocate bits among the programs, instead of performing rate control independently for each program. With joint rate control, bits are allocated so that the total bitrate over
20 all the programs meets the channel bandwidth limitations. Normally, there is very little restriction on how bits are distributed within a group of pictures (GOP), but this is not the case if there is a low delay constraint. To achieve low delay, both the total number of bits for a GOP and the distribution of those bits across the GOP must be constrained.

25 SUMMARY OF THE INVENTION

 These and other drawbacks and disadvantages of the prior art are addressed by the present invention, which is directed to methods and apparatus for encoding multiple video streams for transmission over a video channel.

 In accordance with the principles of the present invention, there is described a
30 method and apparatus for video encoding a plurality of streams corresponding to a

plurality of programs. The apparatus includes an encoder for encoding a plurality of streams for transmission over a communication channel. The encoder imposes a relative positioning on intra-coded frames among the plurality of programs to meet individual stream delay requirements corresponding to the plurality of streams.

5 These and other aspects, features and advantages of the present invention will become apparent from the following detailed description of exemplary embodiments, which is to be read in connection with the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

10 The present invention may be better understood in accordance with the following exemplary figures, in which:

FIG. 1 is a block diagram for an exemplary video encoder to which the present principles may be applied, in accordance with an embodiment of the present invention;

15 FIG. 2 is a diagram for exemplary video streams, in accordance with an embodiment of the present invention;

FIG. 3 is a diagram for an exemplary joint coding system to which the present principles may be applied, in accordance with an embodiment of the present invention; and

20 FIG. 4 is a flow diagram for an exemplary method for encoding a plurality of video streams for transmission over a single communication channel, in accordance with an embodiment of the present invention.

DETAILED DESCRIPTION

25 The present invention is directed to methods and apparatus for encoding multiple video streams for transmission over a video channel. Advantageously, the present invention addresses the problem of simultaneously controlling the bitrate and the delay in a multiprogram application using joint rate control. In an embodiment, the present invention utilizes low-delay joint rate control to efficiently allocate bits during the encoding of non-scalable video streams.

The present description illustrates the principles of the present invention. It will thus be appreciated that those skilled in the art will be able to devise various arrangements that, although not explicitly described or shown herein, embody the principles of the invention and are included within its spirit and scope.

5 All examples and conditional language recited herein are intended for pedagogical purposes to aid the reader in understanding the principles of the invention and the concepts contributed by the inventor to furthering the art, and are to be construed as being without limitation to such specifically recited examples and conditions.

10 Moreover, all statements herein reciting principles, aspects, and embodiments of the invention, as well as specific examples thereof, are intended to encompass both structural and functional equivalents thereof. Additionally, it is intended that such equivalents include both currently known equivalents as well as equivalents developed in the future, i.e., any elements developed that perform the same function, regardless of
15 structure.

Thus, for example, it will be appreciated by those skilled in the art that the block diagrams presented herein represent conceptual views of illustrative circuitry embodying the principles of the invention. Similarly, it will be appreciated that any flow charts, flow diagrams, state transition diagrams, pseudocode, and the like represent
20 various processes which may be substantially represented in computer readable media and so executed by a computer or processor, whether or not such computer or processor is explicitly shown.

The functions of the various elements shown in the figures may be provided through the use of dedicated hardware as well as hardware capable of executing
25 software in association with appropriate software. When provided by a processor, the functions may be provided by a single dedicated processor, by a single shared processor, or by a plurality of individual processors, some of which may be shared. Moreover, explicit use of the term "processor" or "controller" should not be construed to refer exclusively to hardware capable of executing software, and may implicitly include,

without limitation, digital signal processor ("DSP") hardware, read-only memory ("ROM") for storing software, random access memory ("RAM"), and non-volatile storage.

Other hardware, conventional and/or custom, may also be included. Similarly, any switches shown in the figures are conceptual only. Their function may be carried out through the operation of program logic, through dedicated logic, through the interaction of program control and dedicated logic, or even manually, the particular technique being selectable by the implementer as more specifically understood from the context.

In the claims hereof, any element expressed as a means for performing a specified function is intended to encompass any way of performing that function including, for example, a) a combination of circuit elements that performs that function or b) software in any form, including, therefore, firmware, microcode or the like, combined with appropriate circuitry for executing that software to perform the function. The invention as defined by such claims resides in the fact that the functionalities provided by the various recited means are combined and brought together in the manner which the claims call for. It is thus regarded that any means that can provide those functionalities are equivalent to those shown herein.

Turning to FIG. 1, an exemplary video encoder to which the present principles may be applied is indicated generally by the reference numeral 100. An input to the encoder 100 is connected in signal communication with a non-inverting input of a summing junction 110. The output of the summing junction 110 is connected in signal communication with a block transformer 120. The transformer 120 is connected in signal communication with a first input of a quantizer 130. The output of the quantizer 130 is connected in signal communication with a variable length coder ("VLC") 140, where the output of the VLC 140 is an externally available output of the encoder 100. A first input of a rate controller 177 is provided from the output of the summing junction 110, a second input of the rate controller 177 is provided from the output of the VLC 140, and an output of the rate controller 177 is connected in signal communication with a second input of the quantizer 130.

The output of the quantizer 130 is further connected in signal communication with an inverse quantizer 150. The inverse quantizer 150 is connected in signal communication with an inverse block transformer 160, which, in turn, is connected in signal communication with a reference picture store 170. A first output of the reference picture store 170 is connected in signal communication with a first input of a motion estimator 180. The input to the encoder 100 is further connected in signal communication with a second input of the motion estimator 180. The output of the motion estimator 180 is connected in signal communication with a first input of a motion compensator 190. A second output of the reference picture store 170 is connected in signal communication with a second input of the motion compensator 190. The output of the motion compensator 190 is connected in signal communication with an inverting input of the summing junction 110.

As noted above, the present principles are advantageously directed to encoding and transmitting multiple video streams over a single channel with a low delay constraint. The joint rate controller allocates bits to each frame of each program in a way that meets constraints on the total channel bandwidth and individual program delays, while not allowing any individual coding buffer to overflow or underflow.

In an embodiment, intra-coded (I) pictures/frames, which tend to be larger than inter-coded (P, B) pictures/frames, are staggered from one program to the next to help meet the low delay requirement. Compared to the case with no delay constraint, there is a much tighter bound on the peak size of each individual coded picture when low delay is required. For this reason, it is advantageous to have the I frames, which are generally significantly larger than motion compensated P and B frames, occur at different times for different programs. A conventional rate control method has been proposed that allows I frames to be at different times for different programs. However, unlike the principles associated with embodiments of the present invention, the conventional rate control method does not impose a relative positioning of I frames. This is because, unlike the present invention, the conventional rate control method does not attempt to constrain delay.

Turning to FIG. 2, exemplary video streams in accordance with the principles of the present invention are indicated generally by the reference numeral 200. In particular, FIG. 2 illustrates the group of pictures (GOP) structures of the L programs for the case where each program uses the GOP pattern IBPBP, and I frames are spread
5 evenly such that a minimum number of I frames appear at any specific moment in time. That is, the I frames are positioned so as to not exceed a pre-specified threshold number of I frames at a given time across the L programs.

Turning to FIG. 3, an exemplary joint video coding system to which the principles of the present invention may be applied is indicated generally by the reference numeral
10 300. The joint video coding system 300 may include one or more encoders such as that shown and described with respect to FIG. 1. It is presumed that there are L programs which can be jointly coded for a particular channel bandwidth. In FIG. 3, T represents a target number of bits from a rate controller 340, R represents a resulting number of compressed bits, and Q represents an average quantization parameter.

15 The joint video coding system 300 includes an encoder 310, an encoder 320, an encoder 330, the rate controller 340, a multiplexer (MUX) 350, and a buffer 360.

First inputs of the encoders 310, 320, and 330 are available as inputs to the joint coding system 300. First, second, and third outputs of the rate controller 340 are connected in signal communication with second inputs of the encoders 310, 320, and
20 330, respectively. First outputs of the encoders 310, 320, and 330 are respectively connected in signal communication with first, second, and third inputs of the rate controller 340. Second outputs of the encoders 310, 320, and 330 are respectively connected in signal communication with first, second, and third inputs of the multiplexer 350.

25 An output of the multiplexer 350 is connected in signal communication with a first input of the buffer 360. A bidirectional input/output of the rate controller 340 is connected in signal communication with a bidirectional input/output of the buffer 360. An output of the buffer 360 is available as an output of the joint video coding system 300.

A description will now be given regarding an exemplary embodiment illustrating the principles of the present invention, followed by a further description thereof with respect to FIG. 4.

As noted above, it is presumed that there are L programs which can be jointly coded for a particular channel bandwidth. To determine the target bit allocation for each frame, the bit allocation algorithm operates in a hierarchical fashion. First, a super GOP is created, which includes a GOP from each program. For illustrative purposes, it is presumed that the GOP size for each stream is N. Next, a target bit allocation, T(n), for each super frame (where each super frame is a set of frames that includes a frame from each of the L programs taken at the same time instant (i.e., having the same time index), each of the frames included in a super frame being also included in a corresponding super group of pictures) is calculated as follows:

$$T_n = \frac{\sum_{l=1}^L \frac{1}{K_{l,n,t}} C_{l,n,t}}{\sum_{l=1}^L \left[N_{l,I} \frac{C_{l,I}}{K_I} + N_{l,P} \frac{C_{l,P}}{K_P} + N_{l,B} \frac{C_{l,B}}{K_B} \right]} T_r \quad (1)$$

where the following designations apply: l is the number of the program (1 to L); n is the frame position in the GOP (0 to N-1); t is the picture type, {I, P, B}, and is a function of n and l, i.e., t=t(n,l); C(l,n,t) is the complexity measure for frame n; K(l,n,t) is the picture type compensation factor, {K(I), K(P), K(B)}; N(l,I), N(l,P), N(l,B) are the remaining number of I, P, B pictures, respectively; and T(r) is the number of remaining bits.

One can compute the target bit allocation for a program, T(l,n), via the following formula:

$$T_{l,n} = \frac{\frac{1}{K_{l,n,t}} C_{l,n,t}}{\sum_{l=1}^L \frac{1}{K_{l,n,t}} C_{l,n,t}} T_n \quad (2)$$

In this case, the rate controller also selects the encoding mode $t(l,n)$ for the frames. I-frames in particular are spread temporally so that multiple I-frames are not encountered in a particular super frame, and a high quality low-delay result is more easily achieved. One possible embodiment for $t(l,n)$ is the following exemplary

5 algorithm for a GOP having a sequence of IBPBPB:

$$\begin{aligned} t(l,n) &= I, && \text{if } (n-m(l)+N) \bmod N = 0 \\ &= B, && \text{if } (n-m(l)+N) \bmod N \bmod 2 = 1 \\ &= P && \text{otherwise} \end{aligned}$$

10

where $m(l)$ = startup offset for each program. For example if $N=2L$, then the following applies:

$$\begin{aligned} m(1) &= 0 \\ m(2) &= 2 \\ &\dots \\ m(x) &= 2(x-1) \\ &\dots \\ m(L) &= N-2 \end{aligned}$$

20

Of course, it is to be appreciated that other distributions can be selected, when N is or is not precisely a multiple of L , while maintaining the scope of the present principles.

Turning to FIG. 4, an exemplary method for encoding a plurality of video streams
25 for transmission over a single communication channel is indicated generally by the reference numeral 400. Each of the streams corresponds to a respective video program embodied therein. As noted above, it is presumed that there are L programs which can be jointly coded for a particular channel bandwidth. The method 400 provides an advantageous bit allocation over prior art encoding methods in determining
30 a target bit allocation for each frame.

The method 400 includes a start block 405 that passes control to a function block 410. The function block 410 creates a super GOP that includes a GOP from each program, and passes control to a function block 415. For illustrative purposes, it is presumed that the GOP size for each stream is N. The function block 415 selects the encoding mode $t(l,n)$ for the frames, and passes control to a function block 420.

Regarding the selection process implemented by the function block 415, I-frames in particular are spread temporally so that multiple I-frames are not encountered in a particular super frame (e.g., the number of I-frames in a particular super frame is within a pre-specified threshold), and a high quality low-delay result is more easily achieved.

The function block 420 calculates a target bit allocation, $T(n)$, for each super frame (where each super frame is a set of frames that includes a frame from each of the L programs taken at the same time instant, i.e., having the same time index, each of the frames included in a super frame being also included in a corresponding super group of pictures), e.g., using Equation (1), and passes control to a function block 425. The function block 425 computes the target bit allocation, $T(l,n)$, for each of the programs using the target bit allocation for the super frames, and passes control to a function block 430. Function block 425 may compute the target bit allocation for each of the programs using, e.g., Equation (2). Function block 430 encodes the frames based on the encoding mode selection (per function block 415) and target bit allocation for each program (per function block 425), and passes control to an end block 435.

A description will now be given of some of the many attendant advantages/features of the present invention, some of which have been previously described. For example, one advantage/feature is a video encoder for encoding a plurality of streams corresponding to a plurality of programs, wherein the video encoder includes an encoder for encoding each of the plurality of streams for transmission over a single communication channel. The encoder imposes a relative positioning on intra-coded frames among the plurality of programs to meet individual stream delay requirements corresponding to the plurality of streams. Another advantage/feature is the video encoder as described above, wherein the relative positioning on the intra-coded frames is imposed on a group of pictures (GOP) basis. Moreover, another

advantage/feature is the video encoder as described above, wherein the relative positioning on the intra-coded frames is such that the intra-coded frames occur at different times in the plurality of programs. Further, another advantage/feature is the video encoder as described above, wherein the encoder creates a super group of pictures from the plurality of programs, and spreads the intra-coded frames temporally within super frames included in the super group of pictures while ensuring that an amount of the intra-coded frames within each of the super frames in the super group of pictures is within a pre-specified threshold amount. The super group of pictures include at least one group of pictures from each of the plurality of programs. Each of the super frames include a set of frames that have a same time index from across the plurality of programs.

These and other features and advantages of the present invention may be readily ascertained by one of ordinary skill in the pertinent art based on the teachings herein. It is to be understood that the teachings of the present invention may be implemented in various forms of hardware, software, firmware, special purpose processors, or combinations thereof.

Most preferably, the teachings of the present invention are implemented as a combination of hardware and software. Moreover, the software may be implemented as an application program tangibly embodied on a program storage unit. The application program may be uploaded to, and executed by, a machine comprising any suitable architecture. Preferably, the machine is implemented on a computer platform having hardware such as one or more central processing units ("CPU"), a random access memory ("RAM"), and input/output ("I/O") interfaces. The computer platform may also include an operating system and microinstruction code. The various processes and functions described herein may be either part of the microinstruction code or part of the application program, or any combination thereof, which may be executed by a CPU. In addition, various other peripheral units may be connected to the computer platform such as an additional data storage unit and a printing unit.

It is to be further understood that, because some of the constituent system components and methods depicted in the accompanying drawings are preferably

implemented in software, the actual connections between the system components or the process function blocks may differ depending upon the manner in which the present invention is programmed. Given the teachings herein, one of ordinary skill in the pertinent art will be able to contemplate these and similar implementations or configurations of the present invention.

Although the illustrative embodiments have been described herein with reference to the accompanying drawings, it is to be understood that the present invention is not limited to those precise embodiments, and that various changes and modifications may be effected therein by one of ordinary skill in the pertinent art without departing from the scope or spirit of the present invention. All such changes and modifications are intended to be included within the scope of the present invention as set forth in the appended claims.

CLAIMS:

1. An apparatus for encoding a plurality of streams corresponding to a plurality of programs comprising an encoder (300) for encoding the plurality of streams for transmission over a communication channel, wherein said encoder imposes a relative positioning on intra-coded frames among the plurality of programs to meet individual stream delay requirements corresponding to the plurality of streams.
2. The apparatus of claim 1, wherein the relative positioning on the intra-coded frames is imposed on a group of pictures (GOP) basis.
3. The apparatus of claim 1, wherein the relative positioning on the intra-coded frames is such that the intra-coded frames occur at different times in the plurality of programs.
4. The apparatus of claim 1, wherein said encoder (300) creates a super group of pictures from the plurality of programs, and spreads the intra-coded frames temporally within super frames included in the super group of pictures while ensuring that an amount of the intra-coded frames within each of the super frames in the super group of pictures is within a pre-specified threshold amount, the super group of pictures including at least one group of pictures from each of the plurality of programs, each of the super frames including a set of frames that have a same time index from across the plurality of programs.
5. A method for encoding a plurality of streams corresponding to a plurality of programs, the method comprising encoding (425) ones of the plurality of streams for transmission over a communication channel, wherein said encoding step comprises imposing a relative positioning on intra-coded frames among the plurality of programs to meet individual stream delay requirements corresponding to the plurality of streams.

6. The method of claim 5, wherein the relative positioning on the intra-coded frames is imposed on a group of pictures (GOP) basis.

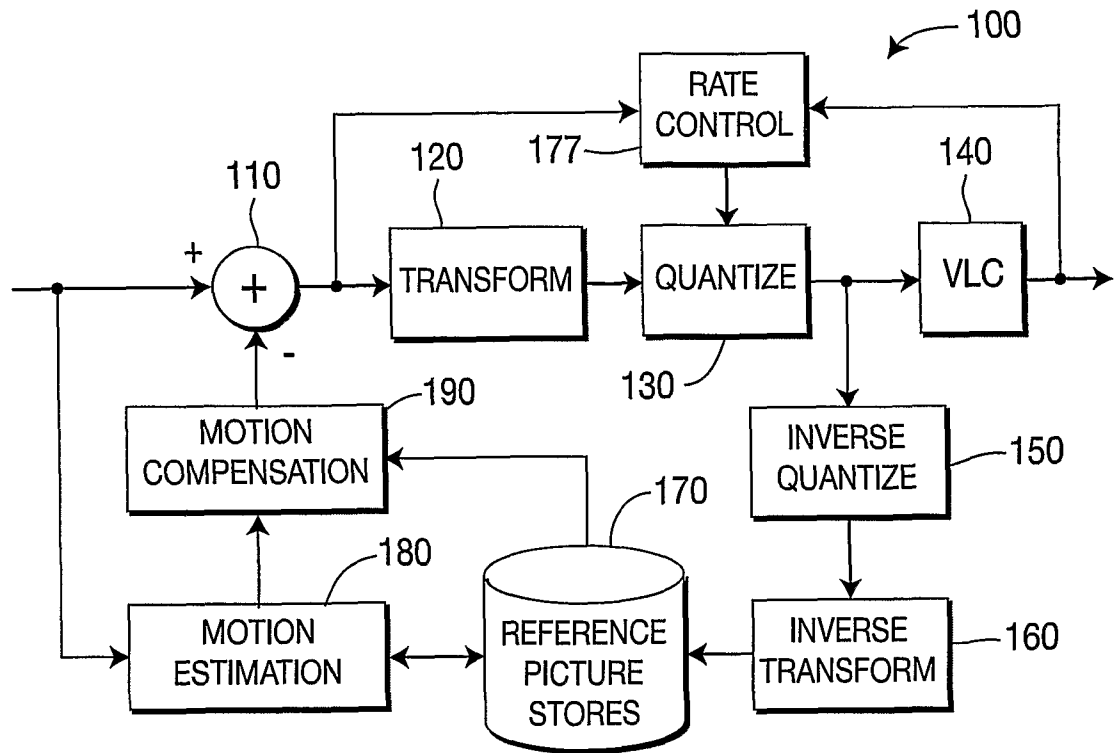
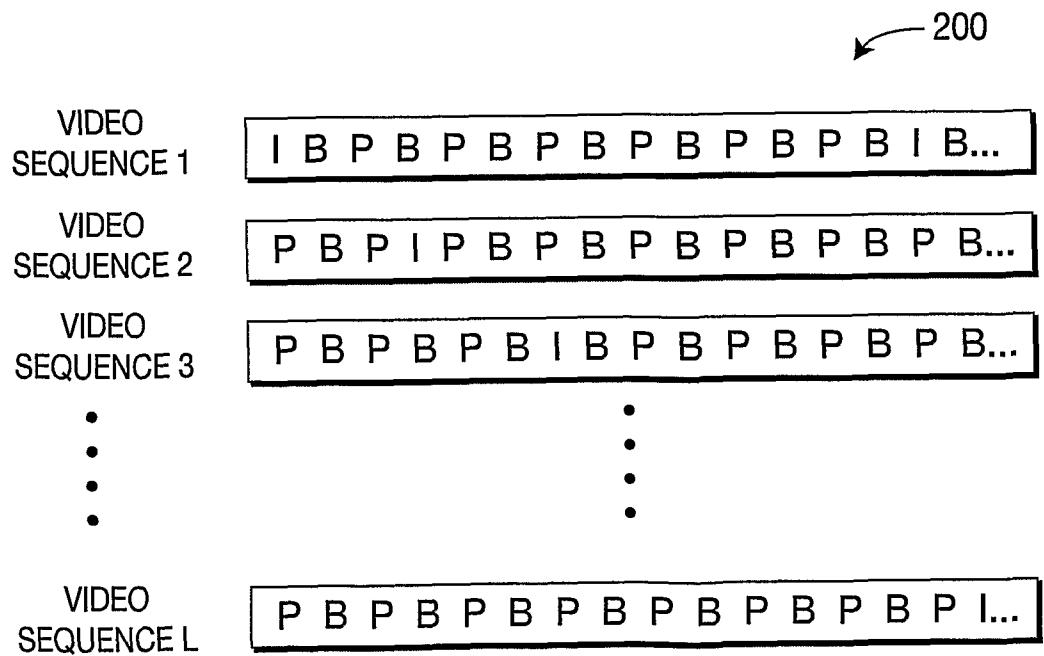
5 7. The method of claim 5, wherein the relative positioning on the intra-coded frames is such that the intra-coded frames occur at different times in the plurality of programs.

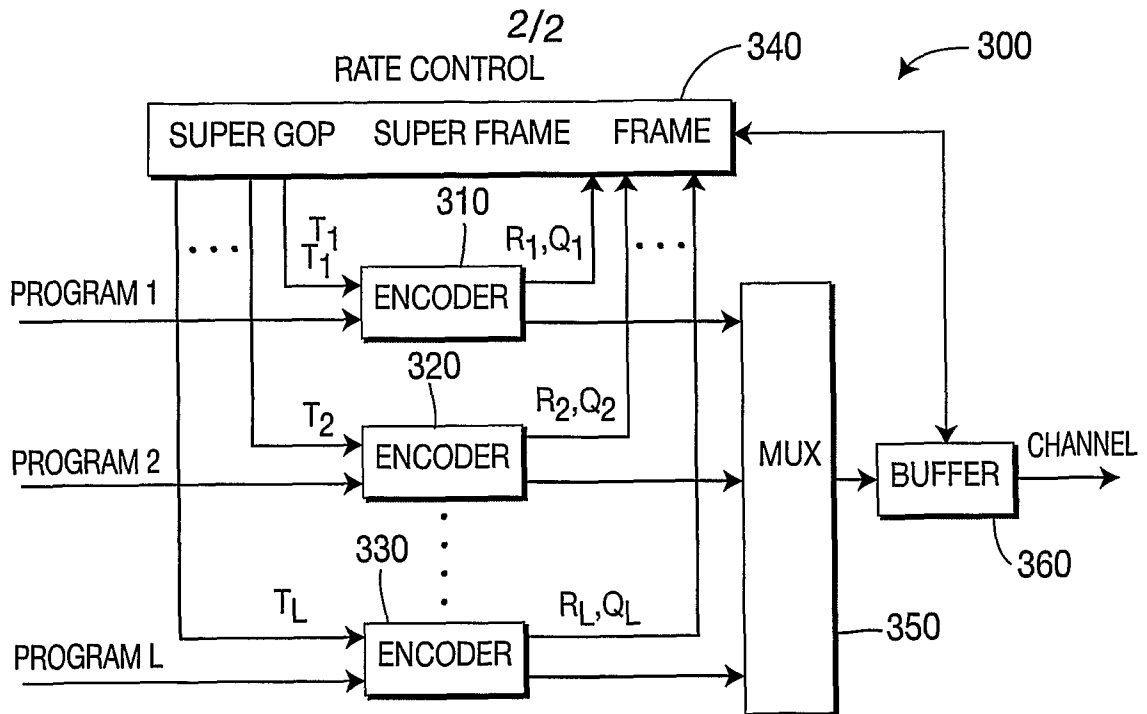
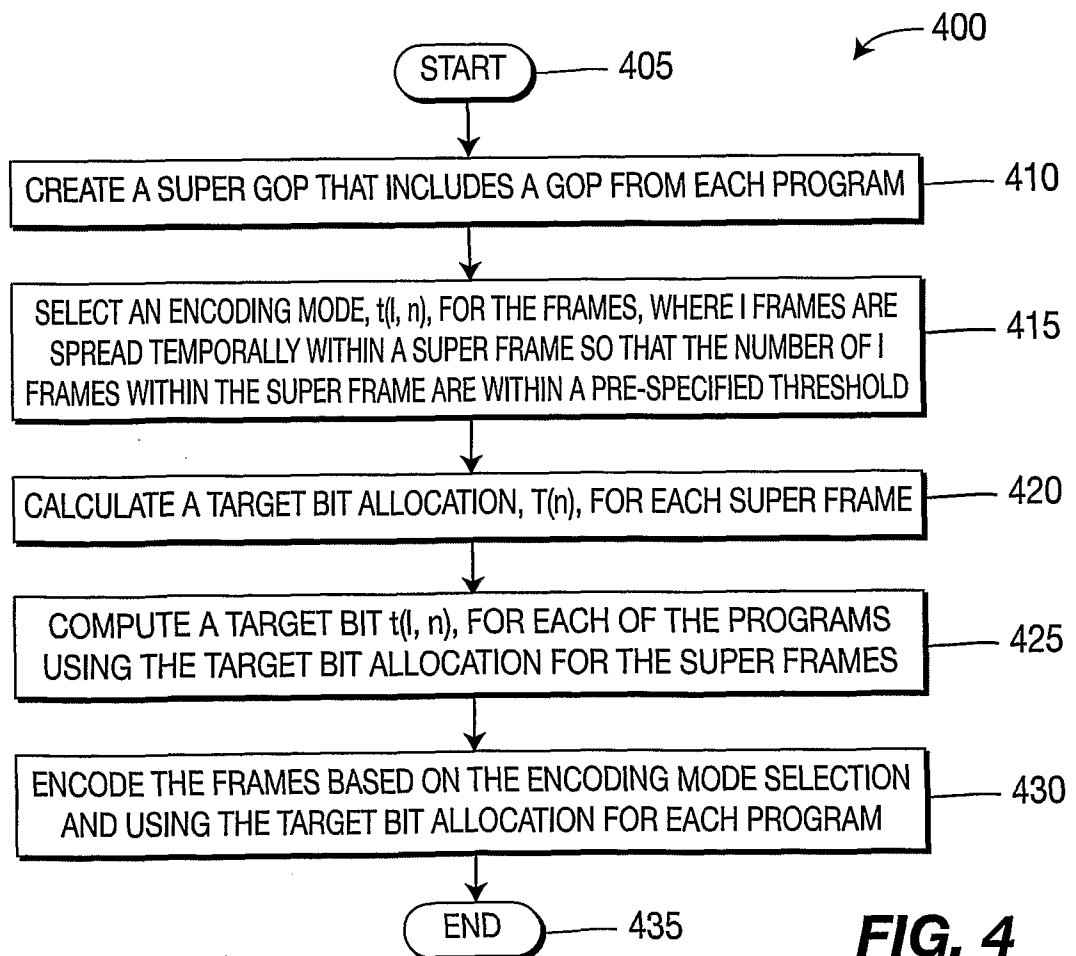
10 8. The method of claim 5, wherein said encoding step comprises:
creating (410) a super group of pictures from the plurality of programs; and
spreading (415) the intra-coded frames temporally within super frames included in the super group of pictures while ensuring that an amount of the intra-coded frames within each of the super frames in the super group of pictures is within a pre-specified threshold amount, the super group of pictures including at least one group of pictures
15 from each of the plurality of programs, each of the super frames including a set of frames that have a same time index from across the plurality of programs.

9. A video processing method comprising:
encoding each one of N video programs to provide N encoded video programs,
20 where $N > 1$; and
transmitting said N encoded video programs over a channel;
wherein said transmitting step temporally staggers intra-coded (I) frames among said N encoded video programs to minimize bitrate peaks over said channel.

25 10. An video processing apparatus comprising:
means for encoding each one of N video programs to provide N encoded video programs, where $N > 1$; and
means for transmitting said N encoded video programs over a channel;
wherein intra-coded (I) frames are temporally staggered among said N encoded
30 video programs to minimize bitrate peaks over said channel.

1/2

**FIG. 1****FIG. 2**

**FIG. 3****FIG. 4**