Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2789998 A1 2004/12/13

(21) 2 789 998

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(22) Date de dépôt/Filing Date: 2004/05/05

(41) Mise à la disp. pub./Open to Public Insp.: 2004/12/13

(62) Demande originale/Original Application: 2 466 458

(30) Priorité/Priority: 2003/06/13 (US10/460,949)

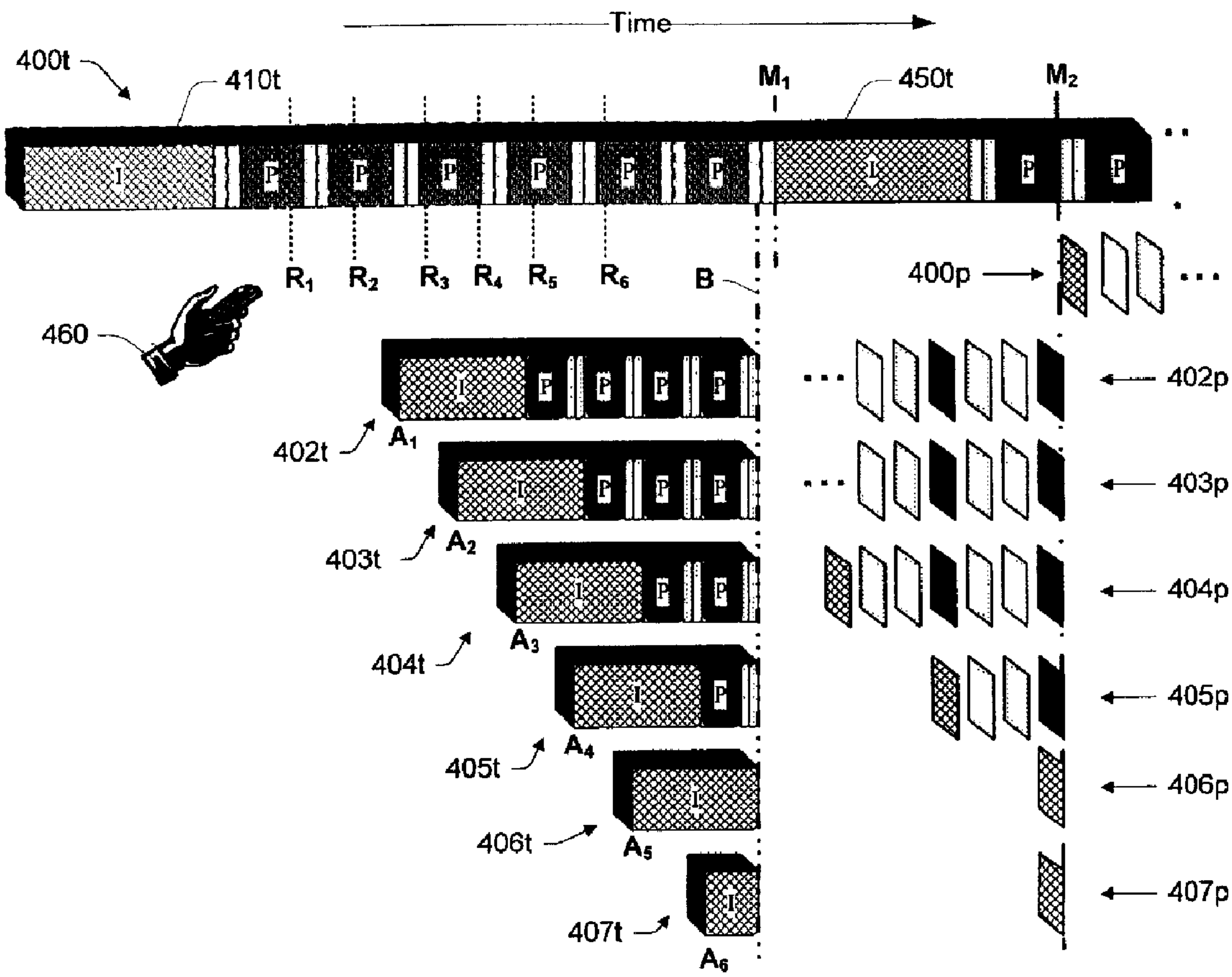
(51) Cl.Int./Int.Cl. *H04N 21/234* (2011.01), *H04N 21/236* (2011.01), *H04N 21/647* (2011.01), *H04N 7/50* (2006.01)

(71) Demandeur/Applicant:
MICROSOFT CORPORATION, US

(72) Inventeurs/Inventors:
BALDWIN, JAMES ARMAND, US;
GREEN, DUSTIN L., US;
GROSSMAN, JOHN H., IV, US

(74) Agent: SMART & BIGGAR

(54) Titre : SYSTEME DE DEMARRAGE RAPIDE POUR TRAINS DE SEQUENCES VIDEO NUMERIQUE
(54) Title: FAST START-UP FOR DIGITAL VIDEO STREAMS



(57) Abrégé/Abstract:
Described herein is a technology facilitating the presentation of digital video streams. An implementation, described herein, reduces the effective start-up delay in the presentation of the first frames of the video content that occurs when a system tunes into a video

(57) Abrégé(suite)/Abstract(continued):
stream by streaming at least an additional stream parallel with the main multicast stream. The alternative parallel stream having offset start points from the main multicast stream to more quickly provide a picture in response to a tuning command.

Abstract

Described herein is a technology facilitating the presentation of digital video streams. An implementation, described herein, reduces the effective start-up delay in the presentation of the first frames of the video content that occurs when a system tunes into a video stream by streaming at least an additional stream parallel with the main multicast stream. The alternative parallel stream having offset start points from the main multicast stream to more quickly provide a picture in response to a tuning command.

FAST START-UP FOR DIGITAL VIDEO STREAMS

This is a divisional of Canadian Patent Application No. 2,466,458 filed May 5, 2004.

TECHNICAL FIELD

This invention generally relates to a technology for digital video streaming.

BACKGROUND

With advent of digital video streaming technology (such as video-on-demand (VOD) systems), users are able to see and hear digital videos, more or less, as the data is being received from a video server.

When video is streamed, the incoming video stream is typically buffered on the user's receiving device (e.g., computer or set-top box) while data is downloaded into it. At some defined point (generally, when the buffer is full), the video contents are presented to the user. As the video content plays, the receiving device empties the data stored in the buffer. However, while the receiving device is playing the stored video, more data is being downloaded to re-fill the buffer. As long as the data can be downloaded at least as fast as it is being played back, the file will play smoothly.

MPEG

The predominant digital video compression and transmission formats are from a family called MPEG (Moving Picture Experts Group). It is the name of family of standards used for coding audio-visual information (e.g., movies, video, music, and such) in a digital compressed format.

For the convenience of explanation of video streaming, the MPEG-family video stream is generally discussed and described herein. However, those who are skilled in the art understand and appreciate that other such digital video compression and transmission formats exist and may be used.

Of course, there are other digital video compression and transmission formats, such as the H.264 codec. Those of ordinary skill in the art will understand how the concepts discussed herein with relationship to MPEG apply to other formats.

GOP and Frames

A MPEG video stream is typically defined by a series of segments called Groups of Pictures (GOP). Typically, a GOP consists of a set of pictures intended to be displayed in sequence over a short duration (e.g., ½ second) when displayed at their intended speed.

A GOP typically includes three types of frames:

- an intra frame (**I**-frame);
- predictive frames (**P**-frames); and
- bi-directionally predictive frames (**B**-frames).

There is no specific limit to the number of frames which may be in a GOP, nor is there a requirement for an equal number of pictures in all GOPs in a video sequence.

The **I**-frame is an encoded still image. It is not dependent upon any other frame that the decoder has already received. Each GOP typically has only one **I**-frame. It is sometimes called a random access point (or "RAP") since it is an entry point for accessing its associated GOP.

From the point of view of a video-stream decoder, the **P**-frames are predicted from the most recently reconstructed **I**- or **P**-frame. A **P**-frame (such as frame 120p) requires data from a previously decompressed anchor frames (e.g., I-frames or P-frames) to enable its decompression.

Switching to the point of view of video stream encoder and transmitter, the B-frames are predicted from the closest two I- or P-frames—one frame in the past and one frame in the future. A B-frame (such as frame 132p) requires data from both preceding and succeeding anchor frames (e.g., I-frames or P-frames) to decode its image. It is bi-directionally dependent.

Of course, other digital video compression and transmission formats (such as H.264 codec) may employ other labels, some different types, and different relationships between frames. For example, in H.264, the frame types, frame dependence relationships, and frame ordering are much more decoupled than they are in MPEG. In H.264, the I-frames are independently decodable and are random access points. Also, frames have defined presentation order (like MPEG does). However, the other frames relate differently than do the MPEG P-frames and B-frames.

So, those of ordinary skill in the art will understand how the concepts discussed herein with relationship to MPEG apply to other formats.

Transmission and Presentation Timelines

Fig. 1 illustrates two manifestations of the same MPEG video stream. The first is the transmission timeline 100t and the other is the presentation timeline 100p.

The transmission timeline 100t illustrates a video stream from the perspective of its transmission by a video-stream encoder and transmitter.

Alternatively, it may be viewed from the perspective of the receiver of the transmission of the video stream.

As shown in Fig. 1, the **I**-frames (e.g., 110t and 150t) are typically temporally longer than the other frames in the transmission timeline. Since it doesn't utilize data from any other frame, it contains all of the data necessary to produce one complete image for presentation. Consequently, an **I**-frame includes more data than any of the other frames. Since the **I**-frame has more data than others, it follows that it typically requires greater time for transmission (and, of course, reception) than the other frame types.

Fig. 1 also shows **P**-frames (such as 120t) and **B**-frames (such as 130t and 132t) of the transmission timeline 100t. Relative to the **B**-frames, the **P**-frames are temporally longer in the transmission timeline because they typically include more data than the **B**-frames. However, **P**-frames are temporally shorter than **I**-frames because they include less data than **I**-frames. Since the **B**-frames rely on data from at least two other frames, they typically do not need as much data of their own to decode their image as do **P**-frames (which rely on one other frame).

Fig. 1 also illustrates the presentation timeline 100p of the video stream from the perspective of its presentation by the video decoder and presenter. In contrast to their transmission duration, the presentation duration of each frame—regardless of type—is exactly the same. In other words, it displays at a fixed frequency.

The incoming frames of the video stream are decoded, buffered, and then presented at a fixed frequency (e.g., 24 frames per second (fps)) to produce a relatively smooth motion picture presentation to the user. In MPEG 2 used to convey NTSC video, the field rate is fixed, and each MPEG 2 picture may produce

1, 2, or 3 fields. Field pictures are required to produce 1 field, and frame pictures may produce 2 or 3 fields. Thus, the frame picture presentation rate may not be fixed, but it is not dictated by the transmission rate of the frame pictures.

Fig. 1 also illustrates a typical decoded GOP 105 of MPEG in its presentation timeline. This GOP example includes an **I**-frame 110p; six **P**-frames (e.g., 120p); and 14 **B**-frames (e.g., 130p and 132p). Typically, each GOP includes a series of consecutively presented decoded frames that begin with an **I**-frame (such as frame 110p).

<u>Order of Transmission and Presentation</u>

As shown in Fig. 1, the order in which the frames are presented typically does not directly match the order in which the frames are transmitted. The arrows shown in Fig. 1 between the frames of the transmission timeline 100t and the presentation timeline 100p illustrate a typical way that frames are reordered between reception and presentation. The tail of each arrow has a bullet (i.e., circle) anchor at the end of a transmitted frame. The head of each arrow has an arrowhead pointing to its corresponding presentation frame.

For example, the transmission **I**-frame 110t corresponds to the presentation **I**-frame 110p. In reality these are the same frames, but their timeline representations indicate their different manifestations.

Returning to the explanation of this example, the transmission **P**-frame 120t corresponds to the presentation **P**-frame 120p. The transmission **B**-frames 130t and 132t corresponds to the presentation **B**-frames 130p and 132p. As shown in Fig. 1, these **B**-frames 130t and 132t are encoded, transmitted, received, and decoded *after* their **P**-frame 120t in the transmission timeline 100t, but their corresponding presentation **B**-frames 130p and 132p are presented *before* their **P**-frame 120p in the presentation timeline 100t. Note that the encoder typically receives the frames in non-compressed form in the same order that the frames are eventually displayed, and the encoder typically performs the frame re-ordering before compressing the frames.

Furthermore, the next GOP to be transmitted starts with **I**-frame 150t, but two **B**-frames 134t and 136t typically come along after this new GOP has begun. As illustrated in Fig. 1, the straggling **B**-frames 134p and 136p are presented in sequence and before the presentation of the **I**-frame 150p of the new GOP.

<u>GOP Presentation Delay</u>

Fig. 1 shows that the **I**-frame 110t of an example GOP is first received beginning at point $\mathbf{T}_1$ in time; however, it is not first presented until point $\mathbf{T}_2$. The time gap between the two points is called herein the "GOP presentation delay" and is labeled 170 in Fig. 1. It represents the delay from when the receiver first begins receiving the first frame of a GOP (which is typically the **I**-frame) until the device first presents the first frame of the GOP.

There are many reasons for this delay. Some are a natural consequence of the video streaming technology and others are imposed into the process to address known technical issues. Some of reasons for the GOP presentation delay include:

- contrast between the time required to receive a frame transmission and the time required to display a frame;
- the time required to decode a frame (especially considering inter-frame dependencies for decoding); and

- built-in delay to facilitate smooth presentation of frames without needed to wait for frame transmission or decoding.

The details of these reasons and the knowledge of other reasons are known to those of ordinary skill in the art.

Video-Stream Presentation Start-up Delay

To tune channels in a video-streaming environment (such as digital cable), a receiver receives a video stream and waits for an access point into the stream. A channel change cannot occur until an access point is received. From the perspective of the user, this can lead to lengthy channel change times.

Fig. 2 illustrates an example of a video-stream presentation start-up delay at 280. The start-up delay is the effective delay experienced by a user. It includes a delay between when a particular video stream is requested and the actual presentation of the first frame of a GOP from the particular video stream. As shown in Fig. 2, the start-up delay 280 includes the GOP presentation delay 270 (discussed above).

Referring to Fig. 2, this example is explained. A GOP, starting with **I**-frame 210t, is being transmitted. This is shown in the transmission timeline 200t. The receiver tunes into this video stream at request point **R**. This selection is illustrated as a user selecting a video-stream channel using a remote control 260.

Again, this is an example illustration for explanatory purpose. This point **R** could be at any moment in time after the beginning (i.e., after the beginning of its **I**-frame 210t) of a GOP.

The receiver must wait for a random access point (or RAP) in order to access the video stream. In this example, each GOP has one RAP. An **I**-frame is an example of a typical RAP. Therefore, each GOP has one **I**-frame. So, the receiver must wait for the next **I**-frame (at the beginning of the next GOP) before it can access the video-stream transmission as shown by transmission timeline 200t.

Once the receiver has an **I**-frame in its buffer, it may refer back to it for dependency decoding of **P**- and **B**-frames. Consequently, a conventional system must wait for a RAP before it can start buffering frames (that are useful).

In Fig. 2, the receiver starts buffering the next GOP at point $\mathbf{M}_1$ with **I**-frame 250t. Thus, the first frame that may be eventually presented to the user is **I**-frame 250t, because it is the first RAP in the stream after the point at which the receiver joined the stream. Because of the GOP presentation delay (discussed above), it actually starts presenting the GOP (with **I**-frame 250p of presentation timeline 200p) at point $\mathbf{M}_2$—which is also the presentation start-up point **S** of the start-up delay 280.

As demonstrated by the screens 262-266, the start-up delay is the effective delay experienced by a user. The user selects a video-stream channel at request point **R** (using, for example, a remote 260) and sees a blank screen, as shown by screen 262. Of course, there may be information presented here (such as electronic programming information), but since it is not yet the desired video-stream content it is effectively blank.

Screen 264 shows that screen remains blank even after the next GOP is currently being received. Screen 266 shows that the first image of frame 250p is finally presented to the user.

The average length of this start-up delay is directly proportional to the average GOP length. Some video-stream providers employ relatively long average

GOP lengths. For these instances, this delay is even more acute because the user is waiting longer for the next GOP to come round after she has changed channels.

In short, this start-up delay is very annoying to the typical users and tries their patience.

SUMMARY

Described herein is a technology facilitating the presentation of digital video streams. An implementation, described herein, reduces the effective start-up delay in the presentation of the first frames of the video content that occurs when a system tunes into a video stream.

In one aspect, there is provided a system comprising: a selector configured to select a lead-in video stream, where the lead-in stream corresponds to a particular main video-stream transmission; a receiver configured to receive the selected lead-in video-stream transmission; a presentation device configured to present content of the selected lead-in video stream transmission; a switcher configured to switch reception from the lead-in to the main video-stream transmission and to switch presentation from the content of the lead in to the content of the main video-stream transmission; an inquisitor configured to query a video-stream provider asking the provider to identify which one of multiple lead-in video-stream transmissions is the first transmission with an available access point, wherein the identified lead-in transmission is the one selected, wherein a first one of the multiple lead-in video stream transmissions includes data for a first number of frames, a second one of the multiple lead-in video-stream transmissions includes data for a second number of frames, the first number of frames being different from the second number of frames, the first number of frames being determined at least in part based on a first time and a next main video-stream access point transmission time, the first time being associated with a first beginning transmission time for the first one of the multiple lead-in video-stream transmissions, the second number of frames being determined at least in part based on a second time and the next main video-stream access point transmission time, the second time being associated with a second beginning transmission time for the second one of the multiple lead-in video-stream transmissions.

In another aspect, there is provided a system comprising: a selector configured to choose a particular main video-stream transmission for reception and presentation; a receiver configured to temporarily receive, before the transmission of the next access point transmitted in the main video-stream transmission, a lead in video stream that corresponds to the particular main video-stream transmission; a switcher configured to switch reception from the lead-in to the main video stream transmission; wherein the lead-in video-stream transmission is scheduled for transmission during transmission of one group-of-pictures (GOP) of the main video-stream transmission, wherein a GOP has only one access point, and wherein the lead-in video-stream transmission includes data for a number of frames, the number of frames being determined at least in part based on the scheduled beginning time for transmission of the lead-in video-stream transmission and a next main video-stream access point transmission time.

This summary itself is not intended to limit the scope of this patent. Moreover, the title of this patent is not intended to limit the scope of this patent. For a better understanding of the present invention, please see the following detailed description and appending claims, taken in conjunction with the accompanying drawings. The scope of the present invention is pointed out in the appending claims.

BRIEF DESCRIPTION OF THE DRAWINGS

The same numbers are used throughout the drawings to reference like elements and features.

Fig. 1 is diagram illustrating a typical video stream transmission timeline and its corresponding presentation timeline.

Fig. 2 is diagram illustrating the presentation start-up delay using a typical video stream transmission timeline and its corresponding presentation timeline.

Fig. 3 is diagram illustrating the new presentation start-up delay in accordance with an implementation, described herein, that employs a single alternative video-stream.

Fig. 4 is diagram illustrating the new presentation start-up delays in accordance with another implementation, described herein, that employs multiple alternative video-streams.

Fig. 5 is a flow diagram showing a methodological implementation described herein.

Fig. 6 is a flow diagram showing a methodological implementation described herein.

Fig. 7 illustrates exemplary environment in which an implementation described herein may be employed.

Fig. 8 illustrates of an example presentation device, a television, and various input devices that interact with the presentation device.

Fig. 9 is a block diagram that illustrates components of the example presentation device(s) shown in Figs. 7 and 8.

DETAILED DESCRIPTION

In the following description, for purposes of explanation, specific numbers, materials and configurations are set forth in order to provide a thorough understanding of the present invention. However, it will be apparent to one skilled in the art that the present invention may be practiced without the specific exemplary details. In other instances, well-known features are omitted or simplified to clarify the description of the exemplary implementations of the present invention, and thereby, to better explain the present invention.

Furthermore, for ease of understanding, certain method steps are delineated as separate steps; however, these separately delineated steps should not be construed as necessarily order dependent in their performance.

The following description sets forth one or more exemplary implementations of a Fast Start-up for Digital Video Streams that incorporate elements recited in the appended claims. These implementations are described with specificity in order to meet statutory written description, enablement, and best-mode requirements. However, the description itself is not intended to limit the scope of this patent.

The inventors intend these exemplary implementations to be examples. The inventors do not intend these exemplary implementations to limit the scope of the claimed present invention; rather, the inventors have contemplated that the claimed present invention might also be embodied and implemented in other ways, in conjunction with other present or future technologies.

An example of an embodiment of a Fast Start-up for Digital Video Streams may be referred to as an "exemplary fast start-up system."

For the convenience of explanation, digital video streams are discussed and described herein in terms of the MPEG-family standard format. However, those who are skilled in the art understand and appreciate that other such digital video compression and transmission formats exist.

Introduction

The one or more exemplary implementations, described herein, of the present claimed invention may be implemented (in whole or in part) by a presentation device 708 (of Figs. 7-9) and/or as part of a computing environment like that shown in Fig. 7.

To tune channels in a digital video multicast (e.g., IP multicast) environment, a receiver receives a multicast video data stream and waits for an access point into the stream. This is sometimes called a random access point (RAP). A channel change cannot occur until a RAP is received. Thus may lead to lengthy channel change times.

A multicast environment is an example of one type of environment that employs video streaming. Such an environment may utilize Internet Protocol multicasting (i.e., IP multicasting). Those of ordinary skill in the art are familiar with multicast and its use in a multicast environment.

In many IP multicast scenarios, there is a limited bit-rate available to the client device (i.e., a receiver). The bit-rate used to the client device typically depends on which IP multicasts the client is currently listening to among all available IP multicasts.

Because the total number of ongoing IP multicasts is not limited by the available bit-rate to any particular client, the exemplary fast start-up system is particularly applicable to such a scenario. One may multicast multiple join-in streams (e.g., alternative transmissions 402t-407t) and a main stream at all times. Furthermore, individual client device may determine for themselves which stream to be listening to at any given point. Also, the bit-rate of the main stream may be as high as the connection to the client device can support—so, a relatively high apparent steady-state video quality may be achieved.

As shown in Fig. 3, the exemplary fast start-up system employs a main multicast video stream transmission 300t and one or more alternative multicast lead-in video stream transmissions (such as transmission 302t) to achieve a faster tuning time. These alternative streams include the same original content as the main stream, but they may have a lower bit-rate than the main stream.

The alternative streams may be low bitrate so that they may be transmitted in less time than is required to transmit the main stream. This is graphically illustrated by alternative video-stream transmission 302t being shorter than the corresponding frames in the main video-stream transmission 300t. Herein, the concept of "low bitrate" is in terms of number of bits per picture (i.e., frame). These may be a normal bitrate in terms of the number of bits per second.

With the exemplary fast start-up system, the video stream content is presented more quickly (than conventional approaches) in response to a tuning command—even if the initial picture is of lesser quality.

While much of the discussion of the exemplary fast start-up system is framed within terms of the MPEG family of digital video compression and transmission formats, those of ordinary skill in the art will understand how the concepts discussed herein with relationship to MPEG apply to other formats, such as H.264.

Exemplary Fast Start-up System with One Alternative Stream

Fig. 3 illustrates the exemplary operation of the exemplary fast start-up system with only one alternative video-stream. Fig. 3 shows the main multicast video stream transmission 300t and the alternative multicast lead-in video stream transmission 302t. Although this discussion refers to the use of only one alternative stream, the bulk of the concepts illustrated in this example apply the use of multiple alternative video-streams.

Fig. 3 also illustrates how this operation reduces the apparent video-stream presentation start-up delay—which is the effective delay experienced by the user. This start-up delay is the delay between the time a particular video stream is requested and the actual presentation of the first frame of a GOP.

Referring to Fig. 3, this example is explained. A GOP starting with **I**-frame 310t is transmitted in the main stream 300t. The receiver tunes into this video stream at request point **R**. This selection is illustrated as a user selecting a video-stream channel using a remote control 360. (However, in a multicast environment this stream 300t is requested, but since the receiver is not yet tuned in, the stream may not actually be transmitted to the receiver until some point after **R**.)

Again, this is an example illustration for explanatory purpose. This point **R** could be at any moment in time within a GOP after its beginning (i.e., after the beginning of its **I**-frame 310t).

The receiver typically waits for a random access point (or RAP) in order to access a video stream. (In this example, each GOP is assumed to have one RAP.) An **I**-frame is an example of a typical RAP. So, the receiver must wait for the next **I**-frame (at the beginning of the next GOP transmission) before it can access a video-stream transmission and present the video.

With conventional approaches, the user would start seeing the video presentation (as shown in presentation timeline 300p) upon the presentation of the first **I**-frame 350p at point $\mathbf{M_2}$. Therefore, the conventional presentation start-up delay (**D**) would be the difference between the request point (**R**) and point $\mathbf{M_2}$. Written as an equation that is

$$\mathbf{D} = \mathbf{M_2} - \mathbf{R} \qquad [1]$$

Conventional Presentation Start-up Delay

However, instead of waiting for the next RAP in the main stream transmission 300t, the exemplary fast start-up system tunes into the alternative video-stream transmission 302t. It starts receiving this GOP transmission at point **A**. The RAP of alternative transmission 302t is **I**-frame 312t; therefore, it can begin presenting the alternative presentation timeline 302p with presentation **I**-frame 312p at point **S'**.

While alternative video-stream is presented (as shown in presentation timeline 302p), the exemplary fast start-up system requests that the multicast router switch the receiver to the main stream multicast transmission 300t at the next RAP (e.g., **I**-frame 350t) of the main stream.

While alternative video-stream is presented (as shown in presentation timeline 302p), the exemplary fast start-up system starts receiving main video-stream transmission 300t starting with the first frame (e.g., **I**-frame 350t) of the next GOP of the main stream. In Fig. 3, the first frame (e.g., **I**-frame 350t) of the next GOP starts at point $\mathbf{M_1}$. Point **B** on the timeline represents the end of the reception of the last frame of the alternative stream transmission 302t.

The switch-over from the alternative stream transmission 302t back to the main stream 300t occurs during or around the gap between points **B** and $\mathbf{M_1}$ in Fig. 3. An exaggerated and visible gap between these points is shown in Fig. 3 to illustrate that there is a clear opportunity for the exemplary fast start-up system to join the main stream. In reality, points **B** and $\mathbf{M_1}$ may occur very nearly concurrently and may indeed be concurrent. It is possible in some embodiments to have an small overlap so that point **B** occurs shortly after point $\mathbf{M}_1$.

At point $\mathbf{M}_2$, the exemplary fast start-up system starts presenting main video-stream presentation 300p starting with the first frame (e.g., **I**-frame 350p) of the next GOP of the main stream. This starts immediately after (or very nearly so) the presentation of the last frame (e.g., frame 322p) of the alternative stream presentation 302p. This presentation switch-over point is designated in Fig. 3 at point $\mathbf{M}_2$.

To accomplish switch-over, it may be desirable for the main stream be tagged for splicing. Alternatively, the RAP locations within the main stream may be conveyed to the multicast router via some other mechanism. Furthermore, it may be desirable for the router have an extension to normal IP multicast so that it starts delivering the main stream multicast transmission when the next RAP in the main stream multicast arrives.

Although not necessarily required, it is desirable for the presentation of the alternative stream 302p to be timed so that the last frame presented (e.g., frame 322p) is presented immediately before the presentation of the first frame (e.g., frame 350p) of the main stream presentation 300p. Doing so enhances the smooth presentation of the video content—in particular, it smoothes the presentation of the switch-over from the alternative to the main stream presentations at point $\mathbf{M}_2$.

As demonstrated by the screens 362-366, the start-up delay is the effective delay experienced by a user. The user selects a video-stream channel at request point **R** (using, for example, a remote 360) and sees a blank screen, as shown by screen 362. Of course, there may be information presented here (such as electronic programming information), but since it is not yet the desired video-stream content it is effectively blank.

Screen 364 shows that screen remains blank even as the RAP of the alternative stream 302t is being transmitted and received. However, screen 366 shows that the first image of frame 312p is presented to the user.

With the exemplary fast start-up system (as shown in Fig. 3), the user first experiences the presentation of the video content (as shown in alternative presentation timeline 302p) upon the presentation of the first **I**-frame 312p at point **S'**. Therefore, the new presentation start-up delay (**D'**) would be the difference between the request point (**R**) and presentation of the alternative stream at point **S'**. Expressed as an equation, that is

$$\mathbf{D'} = \mathbf{S'} - \mathbf{R} \qquad [2]$$

New Presentation Start-up Delay

The time-savings effected by the exemplary fast start-up system is the difference between the **D'** and **D**. Using equations 1 and 2, that time-savings may be expressed as this equation:

$$\mathbf{D} - \mathbf{D'} = \mathbf{M}_2 - \mathbf{S'} \qquad [3]$$

Time-savings

With the exemplary fast start-up system, the user experience is improved because the new start-up delay (**D'**) is less than the conventional start-up delay (**D**) (i.e., **D'** < **D**). The exemplary fast start-up system improves the user's experience by decreasing the effective start-up delay experienced by the user when compared to the delay experienced using conventional approaches.

Exemplary Fast Start-up System with Multiple Alternative Streams

To further minimize the start-up delay, the exemplary fast start-up system may account for the randomness at which a user tunes into a video-stream channel by employing multiple alternative, RAP phase-staggered video-stream transmissions.

Fig. 4 shows the main multicast video stream transmission 400t and phase-staggered multiple alternative lead-in video-stream transmissions 402t-407t. Each alternative transmission may be sent using all of the available bandwidth—thereby minimizing transmission time for each stream. Also, the transmission schedule of the streams are phase-staggered so that the completed reception of each stream's RAP is staggered. Since the RAP of each of the streams illustrated in Fig. 4 is at the beginning of the transmissions, the beginning of each of the transmissions 402t-407t is staggered. This staggered start of the alternative transmissions is illustrated in Fig. 4.

Furthermore, the time range of phase-staggering of the alternative streams 402t-407t is between RAPs of the main stream transmission 400t. In Fig. 4, this is illustrated by the phase-staggering between I-frames 410t and 450t.

By sending multiple different streams, tuning time is improved because the receiver may select one of the lead-in streams to play. The one selected will typically be the one which will be ready to be presented the quickest after the time at which the user tunes.

These alternative transmissions need not be sent concurrently to a particular receiver within a multicast environment. Rather, each one is prepared for transmission, but a particular multicast stream is sent to a particular receiver only when requested by that receiver. Fig. 4 shows six alternative request points (**R_1** through **R_6**) where each one corresponds to a particular alternative video-stream transmission (streams 402t-407t, respectively).

For example, when the exemplary fast start-up system wishes to tune to a channel (examples are indicated by points **R_1** through **R_6**), it queries the multicast server (such as content server 712 of Fig. 7) in order to determine which lead-in alternative stream is the first lead-in that has not started yet, and the receiver joins that alternative multicast transmission. Then, the exemplary fast start-up system requests that the router switch the receiver back to the main stream multicast transmission 400t just before the next RAP (e.g., frame 450t) of the main stream.

Since the alternative stream transmission(s) serves as a "bridge" until the receiver can start receiving the next RAP of the main stream 400t, all of these alternative streams (402t-407t) are shown in Fig. 4 ending at point **B**.

In Fig. 4, the first frame (e.g., **I**-frame 450t) of the next GOP starts at point **M_1**. Point **B** on the timeline represents the end of the reception of the last frame of each of the alternative streams (402t-407t)

The switch-over from each of the alternative streams (402t-407t) back to the main stream 400t occurs during or around the gap between points **B** and **M_1** in Fig. 4. An exaggerated and visible gap between these points is shown in Fig. 4 to illustrate that there is a clear opportunity for the exemplary fast start-up system to join the main stream. In reality, points **B** and **M_1** may occur very nearly concurrently and may indeed be concurrent. It is possible in some embodiments to have a small overlap so that point **B** occurs shortly after point $\mathbf{M_1}$.

Fig. 4 shows the alternative video-stream presentations 402p-407p that correspond to the alterative video-stream transmissions 402t-407t, respectively. Although not necessarily required, it is desirable for the alternative video-stream presentations 402p-407p to be timed so that the last frame presented is presented immediately before the presentation of the first frame of the main stream presentation 400p. Doing so enhances the smooth presentation of the video content—in particular, it smoothes the presentation of the switch-over from the alternative to the main stream presentations at point $\mathbf{M_2}$.

Operation of Exemplary Fast Start-up System

Fig. 5 shows a methodological implementation of the exemplary fast start-up system. This methodological implementation may be performed in software, hardware, or a combination thereof.

At 510 of Fig. 5, the user tunes into a specific main video-stream multicast transmission (e.g., 300t or 400t). This example point is designated at points **R** in Fig. 3 and $\mathbf{R_1}$ through $\mathbf{R_6}$ in Fig. 4.

At 512, the exemplary fast start-up system queries a multicast server (such as content server 712 of Fig. 7) in order to determine which of the alternative RAP-phase-staggered lead-in alternative video streams is the first lead-in that has not started yet. Of course, if there is only one alternative transmission, this query may be viewed simply as a request.

At 514, the exemplary fast start-up system joins the alternative multicast transmission identified by the query. It receives and buffers this alternative transmission. This occurs, for example, at designated points **A** in Fig. 3 and $\mathbf{A_1}$ through $\mathbf{A_6}$ in Fig. 4.

At 516, it presents the video stream of the identified and buffered alternative transmission. This occurs, for example, at designated point **S** in Fig. 3.

At 518, the exemplary fast start-up system switches back to receiving and buffering the main stream multicast transmission (e.g., 300t and 400t) exactly at the next RAP (e.g., frame 350t and 450t) of the main stream. It may do this by requesting that the router switch the receiver back to the main stream multicast transmission exactly at the next RAP of the main stream. This occurs, for example, between or around designated points **B** and $\mathbf{M_1}$ in Figs. 3 and 4. This may also be described as occurring "on or about" such designated points.

At 520, it presents the video stream of the main stream multicast. This occurs, for example, at designated point $\mathbf{M_2}$ in Figs. 3 and 4.

Operation of Exemplary Fast Start-up Multicast System

Fig. 6 shows a methodological implementation of the exemplary fast start-up multicast system, which may be embodied by a content provider 702 and/or a content distribution system 706 of Fig. 7). This methodological implementation may be performed in software, hardware, or a combination thereof.

At 610 of Fig. 6, the exemplary fast start-up multicast system concurrently encodes one or more alternative video streams for transmission. The server system encodes each alternative stream so that the RAP of each is phase-staggered relative to the other streams.

Furthermore, each stream is encoded so that each ends at the same point, which is at or near when the next RAP is available in the main video stream. This designated point, for example, is point **M_1** in Figs. 3 and 4.

At 612, the multicast system receives a query to which of the RAP-phase-staggered lead-in alternative video streams is the first lead-in that has not started yet. Of course, if there is only one alternative transmission, this query may be viewed simply as a request.

At 614, it transmits the alternative multicast transmission identified by the query to the receiver that requested it.

At 616, the multicast system receives a request for the router switch the receiver back to the main stream multicast transmission just before the next RAP of the main stream. At 618, it does so in response to such a request. This occurs, for example, between or around designated points **B** and **M_1** in Figs. 3 and 4.. This may also be described as occurring "on or about" such designated points.

Exemplary Environment

Fig. 7 illustrates an exemplary environment 700 in which the techniques, systems, and other aspects described herein may be implemented (partially or wholly). Exemplary environment 700 is a television entertainment system that facilitates distribution of multi-media.

The environment 700 includes one or more multimedia content providers 702, a content distribution system 706, and one or more presentation devices 708(1), 708(2), ..., 708(N) coupled to the content distribution system 706 via a multicast-capable network 710.

Multimedia content provider 702 includes a content server 712 and stored content 714, such as movies, television programs, commercials, music, and similar audio and/or video content. Content server 712 controls distribution of the stored content 714 from content provider 702 to the content distribution system 706. Additionally, content server 702 controls distribution of live content (e.g., content that was not previously stored, such as live feeds) and/or content stored at other locations to the content distribution system 706.

Content distribution system 706 may be coupled to a network 720, such as an intranet or the Internet. The content distribution system 706 includes a multicast transmitter 728, and one or more content processors 730. Multicast transmitter 728 multicasts signals across multicast-capable network 710.

Content distribution system 706 is representative of a headend service that provides multimedia content to multiple subscribers.

Multicast-capable network 710 can include a cable television network, RF, microwave, satellite, and/or data network, such as the Internet, and may also include wired or wireless media using any multicast format or multicast protocol. Additionally, multicast-capable network 710 may be any type of network, using any type of network topology and any network communication protocol, and may be represented or otherwise implemented as a combination of two or more networks.

Content processor 730 processes the content received from content provider 702 prior to transmitting the content across multicast-capable network 708. A particular content processor 730 may encode, or otherwise process, the received content into a format that is understood by the multiple presentation devices 708(1), 708(2), ..., 708(N) coupled to multicast-capable network 710.

Presentation devices 708 may be implemented in a number of ways. For example, a presentation device 708(1) receives content multicasted from a satellite-based transmitter via a satellite dish 734. Presentation device 708(1) is also referred to as a set-top box or a satellite receiving device. Presentation device 708(1) is coupled to a television 736(1) for presenting the content received by the presentation device (e.g., audio data and video data), as well as a graphical user interface. A particular presentation device 708 may be coupled to any number of televisions 736 and/or similar devices that may be implemented to display or otherwise render content. Similarly, any number of presentation devices 708 may be coupled to a single television 736.

Presentation device 708(2) is also coupled to receive content from multicast-capable network 710 and provide the received content to associated television 736(2). Presentation device 708(N) is an example of a combination television 738 and integrated set-top box 740. In this example, the various components and functionality of the set-top box are incorporated into the television, rather than using two separate devices. The set-top box incorporated into the television may receive multicast signals via a satellite dish or wireless antenna (such as dish 734) and/or via multicast-capable network 710. In alternate implementations, presentation devices 706 may receive content via the Internet or any other multicast medium.

The exemplary environment 700 also includes live or stored pay-per-view (PPV) content 742, such as PPV movie content. The stored or live content is typically multicast on a schedule. When a device joins a PPV multicast channel, the PPV content may be viewed with a presentation device 708.

Exemplary Presentation Device

Fig. 8 illustrates an exemplary implementation 800 of a presentation device 708 shown as a standalone unit that connects to a television 736. Presentation device 708 may be implemented in any number of embodiments, including as a set-top box, a satellite receiver, a TV recorder with a hard disk, a game console, an information appliance, a DVD player, personal video recorder, a personal computer, a home media center, a modem, and so forth.

Presentation device 708 includes a wireless receiving port 802, such as an infrared (IR) or Bluetooth wireless port, for receiving wireless communications from a remote control device 804, a handheld input device 806, or any other wireless device, such as a wireless keyboard. Handheld input device 806 may be a personal digital assistant (PDA), handheld computer, wireless phone, or the like. Additionally, a wired keyboard 808 is coupled to communicate with the presentation device 708. In alternate embodiments, remote control device 804, handheld device 806, and/or keyboard 808 may use an RF communication link or other mode of transmission to communicate with presentation device 708.

Presentation device 708 may have a storage medium reader 809 for reading content storage media, such as DVD disks. A standalone or non-standalone presentation device 708 may include the storage medium reader 809.

Presentation device 708 may receive one or more multicast signals 810 from one or more multicast sources, such from a multicast network.

Presentation device 708 also includes hardware and/or software for providing the user with a graphical user interface by which the user can, for example, access various network services, configure the presentation device 708, and perform other functions.

Presentation device 708 may be capable of communicating with other devices via one or more connections including a conventional telephone link 812, an ISDN link 814, a cable link 816, an Ethernet link 818, a DSL link 820, and the like. Presentation device 708 may use any one or more of the various communication links 812-820 at a particular instant to communicate with any number of other devices. The multicast signals may also be received via the various communication links 812-820.

Presentation device 708 generates video signal(s) 820 and audio signal(s) 822, both of which are communicated to television 736. Alternatively, video and audio signal may be communicated to other audio/visual equipment, such as speakers, a video monitor, a home theater system, an audio system, and the like.

Although not shown in Fig. 8, presentation device 708 may include one or more lights or other indicators identifying the current status of the device. Additionally, the presentation device may include one or more control buttons, switches, or other selectable controls for controlling operation of the device.

Fig. 9 illustrates selected components of presentation device 708 shown in Figs. 7 and 8. Presentation device 708 includes a first tuner 900 and an optional second tuner 902. The tuners 900 and 902 are representative of one or more in-band tuners that tune to various frequencies or channels to receive television signals, as well as an out-of-band tuner or receiver or network interface card that tunes to or receives the multicast communications channel over which other content may be multicast to presentation device 708.

The tuners 900 and 902 may be digital tuners, analog tuners, or any combination of analog and digital components used to get digital data into the client device 708.

Presentation device 708 also includes one or more processors 304 and one or more memory components. Examples of possible memory components include a random access memory (RAM) 906, a disk drive 908, a mass storage component 910, and a non-volatile memory 912 (e.g., ROM, Flash, EPROM, EEPROM, etc.).

Alternative implementations of presentation device 708 can include a range of processing and memory capabilities, and may include more or fewer types of memory components than those illustrated in Fig. 9.

Processor(s) 904 process various instructions to control the operation of presentation device 708 and to communicate with other electronic and computing devices. The memory components (e.g., RAM 906, disk drive 908, storage media 910, and non-volatile memory 912) store various information and/or data such as multimedia content, electronic program data, web content data, configuration information for presentation device 708, and/or graphical user interface information. The device may cache data into any one of these many memory components.

An operating system 914 and one or more application programs 916 may be stored in non-volatile memory 912 and executed on processor 904 to provide a runtime environment. A runtime environment facilitates extensibility of presentation device 708 by allowing various interfaces to be defined that, in turn, allow application programs 916 to interact with presentation device 708.

The application programs 916 that may be implemented on the presentation device 708 may include an electronic program guide (EPG), an email program to facilitate electronic mail, and so on.

Presentation device 708 can also include other components pertaining to a television entertainment system which are not illustrated in this example for simplicity purposes. For instance, presentation device 708 can include a user interface application and user interface lights, buttons, controls, etc. to facilitate viewer interaction with the device.

Network interface 924 and serial and/or parallel interface 926 allows presentation device 708 to interact and communicate with other electronic and computing devices via various communication links. Although not shown, presentation device 708 may also include other types of data communication interfaces to communicate with other devices.

It may include a modem 928 or other communications device that facilitates communication with other electronic and computing devices via a conventional telephone line or other communications mediums.

The presentation device 708 has the ability to receive multicast digital data and it may receive it using the tuners 900 or 902, the network interface 924, the modem 928, or other communications device.

Presentation device 708 also includes an audio/video output 930 that provides signals to a television or other device that processes and/or presents or otherwise renders the audio and video data. This output may be called the display.

Presentation device 708 also includes a fast start-up module 940 that partially or wholly implements the exemplary fast start-up system. It may be application program or a hardware component.

Although shown separately, some of the components of presentation device 708 may be implemented in an application specific integrated circuit (ASIC). Additionally, a system bus (not shown) typically connects the various components within presentation device 708.

A system bus may be implemented as one or more of any of several types of bus structures, including a memory bus or memory controller, a peripheral bus, an accelerated graphics port, or a local bus using any of a variety of bus architectures. By way of example, such architectures can include a CardBus, Personal Computer Memory Card International Association (PCMCIA), Accelerated Graphics Port (AGP), Small Computer System Interface (SCSI), Universal Serial Bus (USB), IEEE 1394, a Video Electronics Standards Association (VESA) local bus, and a Peripheral Component Interconnects (PCI) bus also known as a Mezzanine bus.

Computer-Executable Instructions

An implementation of an exemplary fast start-up system may be described in the general context of computer-executable instructions, such as program modules, executed by one or more computers or other devices. Generally, program modules include routines, programs, objects, components, data structures, etc. that perform particular tasks or implement particular abstract data types. Typically, the functionality of the program modules may be combined or distributed as desired in various embodiments.

Computer Readable Media

An implementation of an exemplary fast start-up system may be stored on or transmitted across some form of computer readable media. Computer readable media may be any available media that may be accessed by a computer. By way of example, and not limitation, computer readable media may comprise "computer storage media" and "communications media."

"Computer storage media" include volatile and non-volatile, removable and non-removable media implemented in any method or technology for storage of information such as computer readable instructions, data structures, program modules, or other data. Computer storage media includes, but is not limited to, RAM, ROM, EEPROM, flash memory or other memory technology, CD-ROM, digital versatile disks (DVD) or other optical storage, magnetic cassettes, magnetic tape, magnetic disk storage or other magnetic storage devices, or any other medium which may be used to store the desired information and which may be accessed by a computer.

"Communication media" typically embodies computer readable instructions, data structures, program modules, or other data in a modulated data signal, such as carrier wave or other transport mechanism. Communication media also includes any information delivery media.

The term "modulated data signal" means a signal that has one or more of its characteristics set or changed in such a manner as to encode information in the signal. By way of example, and not limitation, communication media includes wired media such as a wired network or direct-wired connection, and wireless media such as acoustic, RF, infrared, and other wireless media. Combinations of any of the above are also included within the scope of computer readable media.

Conclusion

Although the invention has been described in language specific to structural features and/or methodological steps, it is to be understood that the invention defined in the appended claims is not necessarily limited to the specific features or steps described. Rather, the specific features and steps are disclosed as preferred forms of implementing the claimed invention.

CLAIMS:

1. A system comprising:

a selector configured to select a lead-in video stream, where the lead-in stream corresponds to a particular main video-stream transmission;

a receiver configured to receive the selected lead-in video-stream transmission;

a presentation device configured to present content of the selected lead-in video stream transmission;

a switcher configured to switch reception from the lead-in to the main video-stream transmission and to switch presentation from the content of the lead in to the content of the main video-stream transmission;

an inquisitor configured to query a video-stream provider asking the provider to identify which one of multiple lead-in video-stream transmissions is the first transmission with an available access point, wherein the identified lead-in transmission is the one selected, wherein a first one of the multiple lead-in video stream transmissions includes data for a first number of frames, a second one of the multiple lead-in video-stream transmissions includes data for a second number of frames, the first number of frames being different from the second number of frames, the first number of frames being determined at least in part based on a first time and a next main video-stream access point transmission time, the first time being associated with a first beginning transmission time for the first one of the multiple lead-in video-stream transmissions, the second number of frames being determined at least in part based on a second time and the next main video-stream access point transmission time, the second time being associated with a second beginning transmission time for the second one of the multiple lead-in video-stream transmissions.

2. The system as recited in claim 1, wherein the selector is further configured to select the selected lead-in video stream from one of multiple lead-in video streams.

3. The system as recited in claim 1, wherein the switcher is configured to switch on or about the occurrence of an access point transmitted in the main video stream transmission.

4. The system as recited in claim 1, wherein the switcher is configured to switch just before the occurrence of an access point transmitted in the main video-stream transmission.

5. The system as recited in claim 1, wherein the available access point is an anchor frame.

6. The system as recited in claim 1, wherein the access point of each of the multiple lead-in video-stream transmissions are scheduled for phase staggered transmission relative to each other.

7. The system as recited in claim 1, wherein each of the multiple lead-in video-stream transmissions are encoded using a lower bitrate than that used by the main video-stream transmission.

8. A system comprising:

a selector configured to choose a particular main video-stream transmission for reception and presentation;

a receiver configured to temporarily receive, before the transmission of the next access point transmitted in the main video-stream transmission, a lead in video stream that corresponds to the particular main video-stream transmission;

a switcher configured to switch reception from the lead-in to the main video stream transmission;

wherein the lead-in video-stream transmission is scheduled for transmission during transmission of one group-of-pictures (GOP) of the main video-stream transmission, wherein a GOP has only one access point, and wherein the lead-in video-stream transmission includes data for a number of frames, the number of frames being determined at least in part based on the scheduled beginning time for transmission of the lead-in video-stream transmission and a next main video-stream access point transmission time.

9. The system as recited in claim 8 further comprising a presenter configured to present content of the lead-in video stream transmission and configured to, after switching reception to the main video-stream, present the content of the main video stream transmission.

10. The system as recited in claim 8, wherein the switcher is further configured to switch on or about the transmission of an access point in the main video-stream transmission.

11. The system as recited in claim 10 wherein the video-stream transmissions are multicast.

12. The system as recited in claim 8, wherein the switcher is further configured to switch on or about the transmission of the next access point to occur during the main video-stream transmission.

13. The system as recited in claim 8, wherein each of the multiple lead-in video-stream transmissions are encoded using a lower bitrate than that used by the main video-stream transmission.

14. The system as recited in claim 8, wherein the video-stream transmissions are multicast.

15. The system as recited in claim 8, wherein the access points are anchor frames.

16. The system as recited in claim 8 further comprising:

the selector is further configured to select one of multiple lead-in video-stream transmissions for reception, where each lead-in stream corresponds to a particular main video-stream transmission, wherein the one lead-in transmission selected is the lead-in transmission that is temporarily received;

a presenter configured to present content of the selected lead-in video stream transmission;

a switcher further configured to switch reception from the lead-in to the main video-stream transmission and doing so on or about the occurrence of an access point transmitted in the main video-stream transmission and switch presentation from the content of the lead-in to the content of the main video stream transmission.

17. The system as recited in claim 16, wherein the access point of each of the multiple lead-in video-stream transmissions are scheduled for phase staggered transmission relative to each other.

18. The system as recited in claim 16, wherein each of the multiple lead-in video-stream transmissions are encoded using a lower bitrate than that used by the main video-stream transmission.

19. The system as recited in claim 16, wherein the access points are anchor frames.

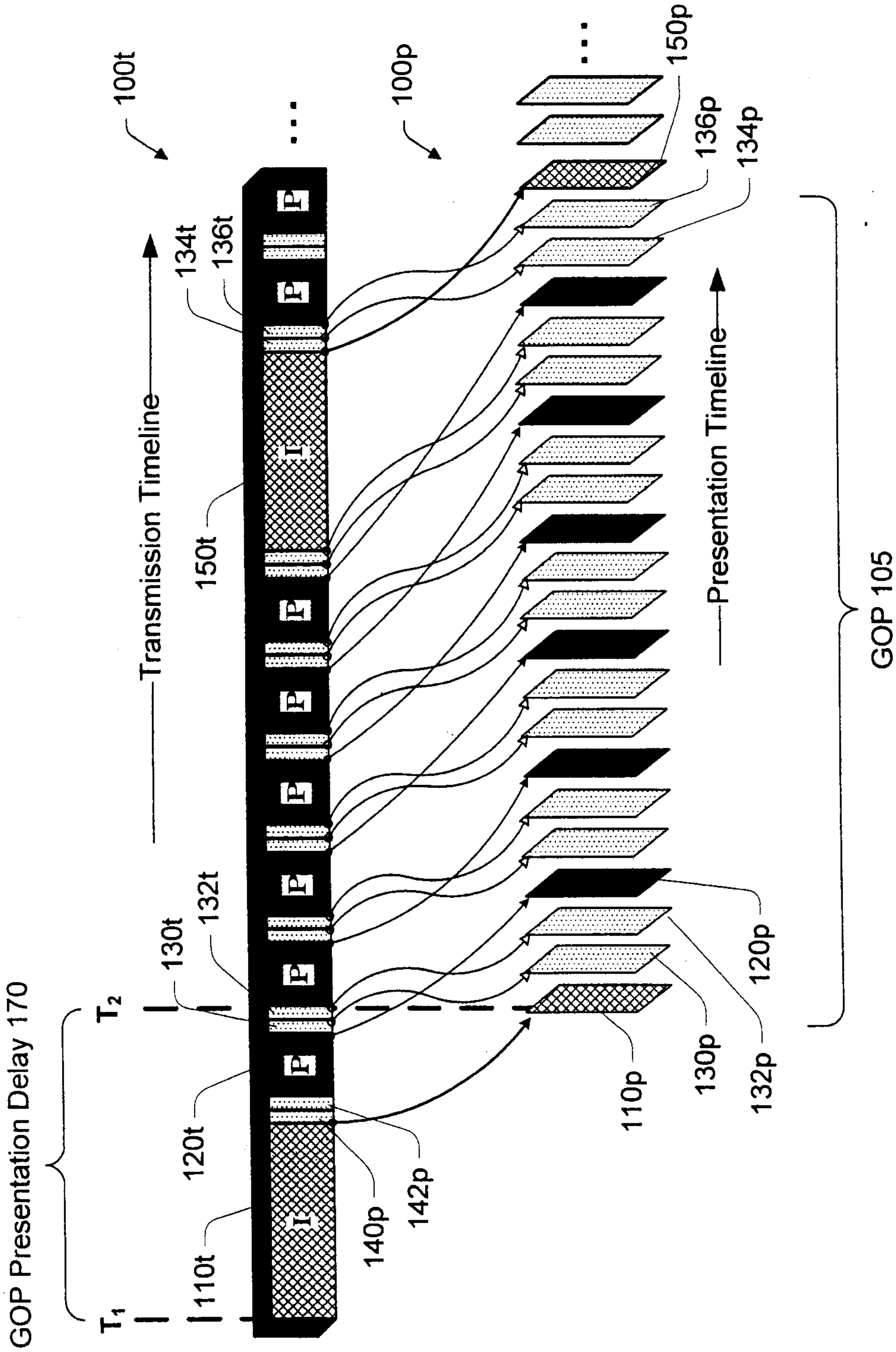


Fig. 1 (background)

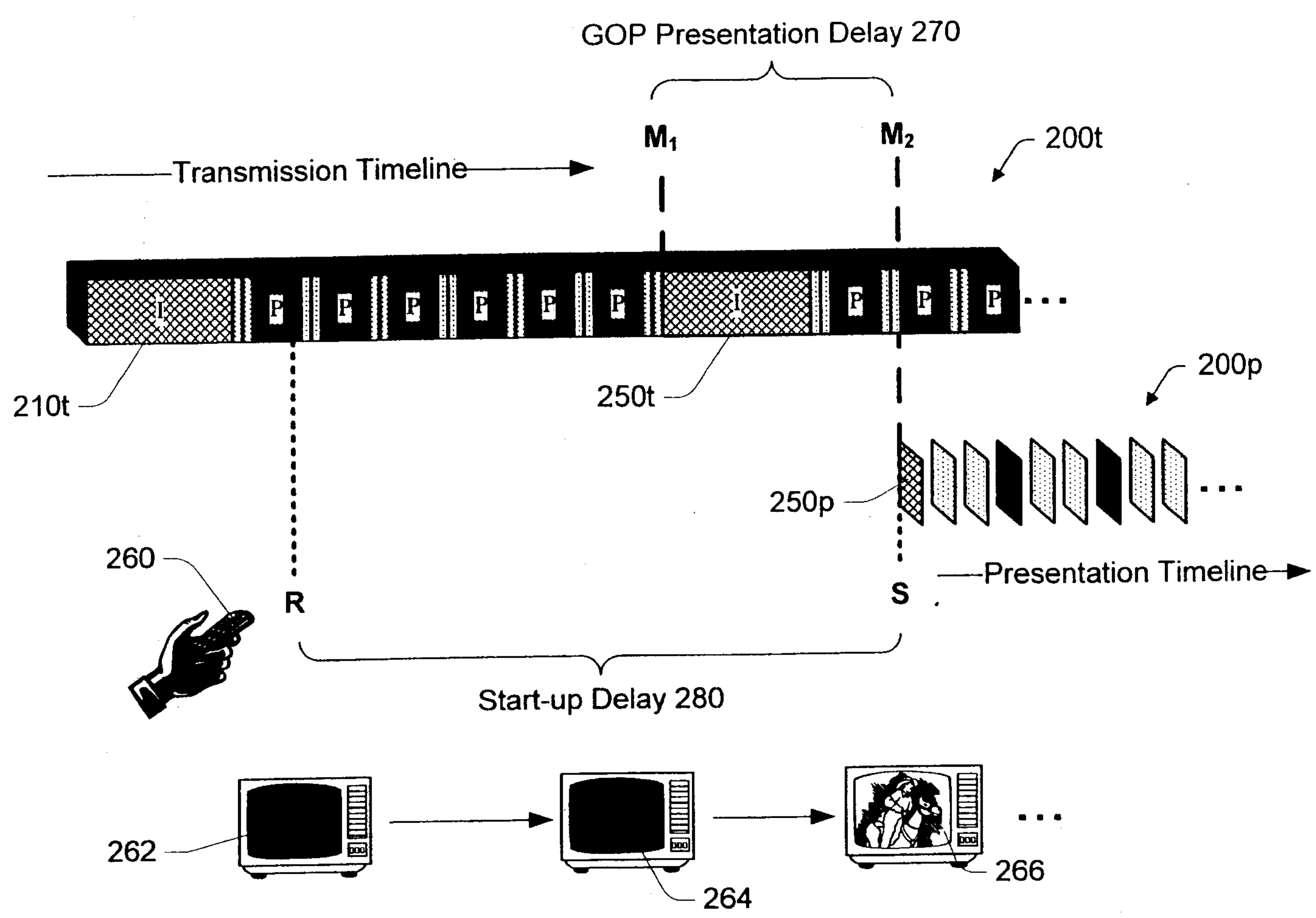


Fig. 2 (background)

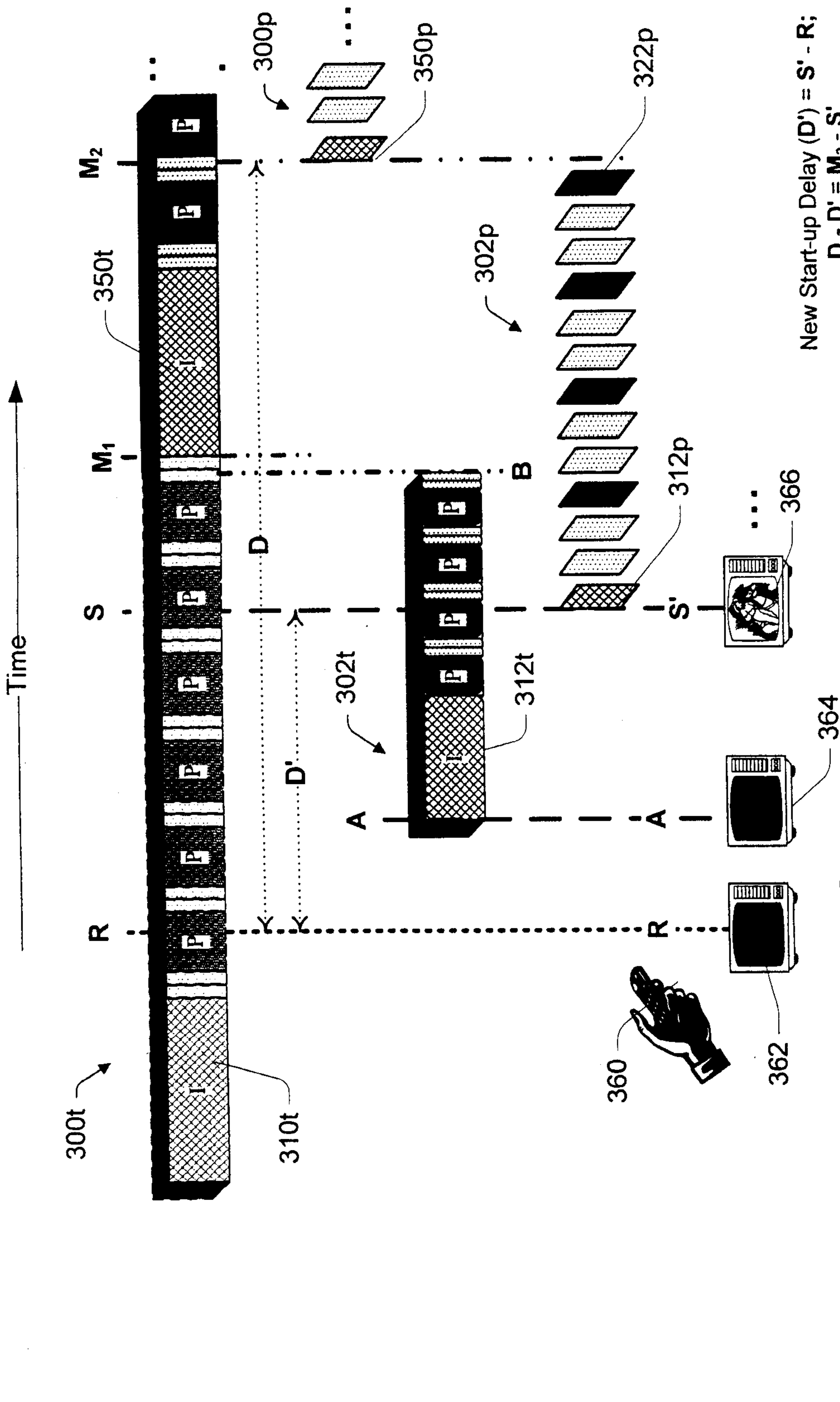
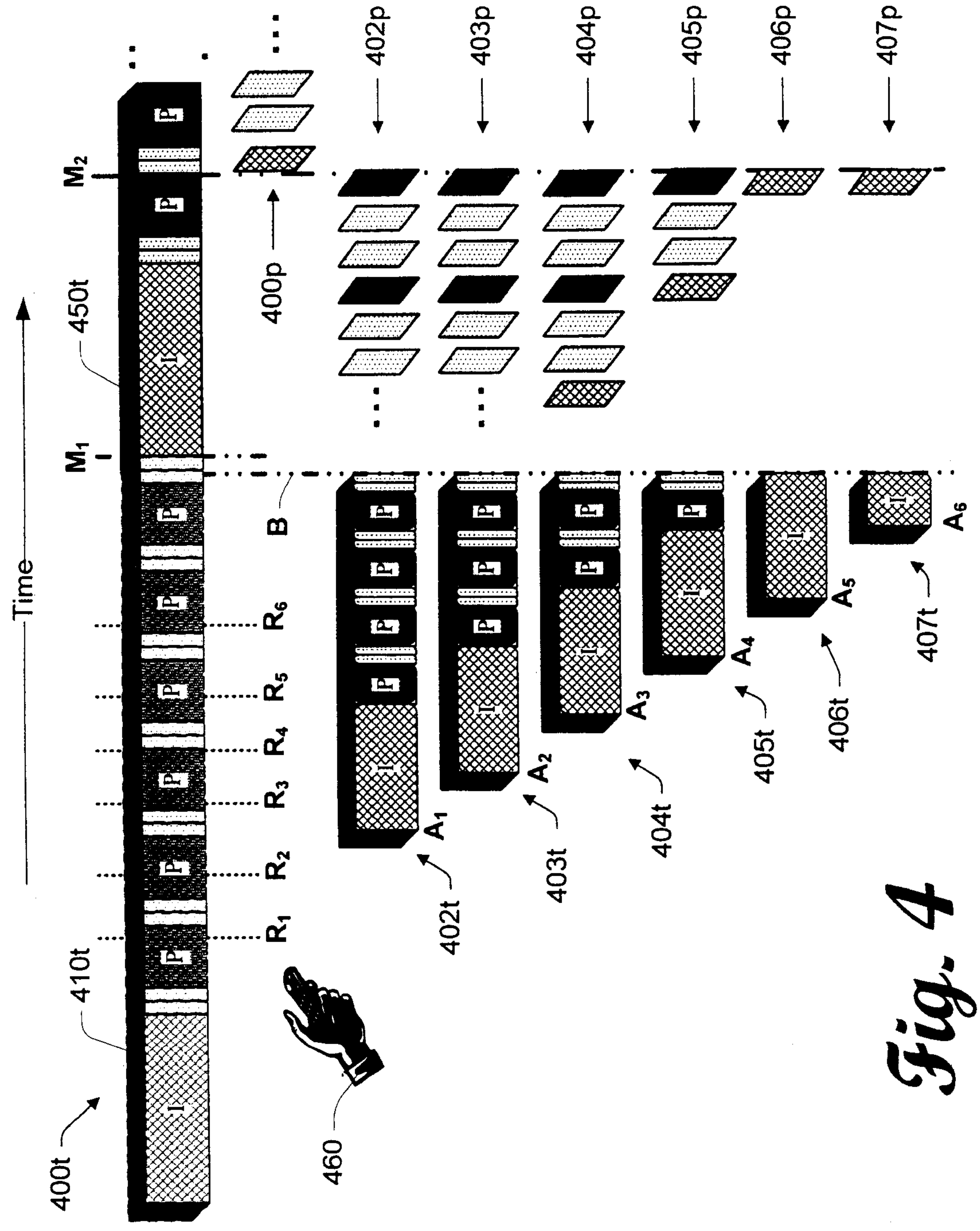

Time
300t
310t
R
S
M1
M2
350t
D
D'
302t
312t
A
B
300p
350p
322p
302p
312p
S'
360
362
364
366
New Start-up Delay (D') = S' - R;
D - D' = M2 - S'

Fig. 3



Time
400t
410t
450t
M1
M2
B
R1
R2
R3
R4
R5
R6
400p
402p
403p
404p
405p
406p
407p
402t
403t
404t
405t
406t
407t
A1
A2
A3
A4
A5
A6
460
I
P

Fig. 4

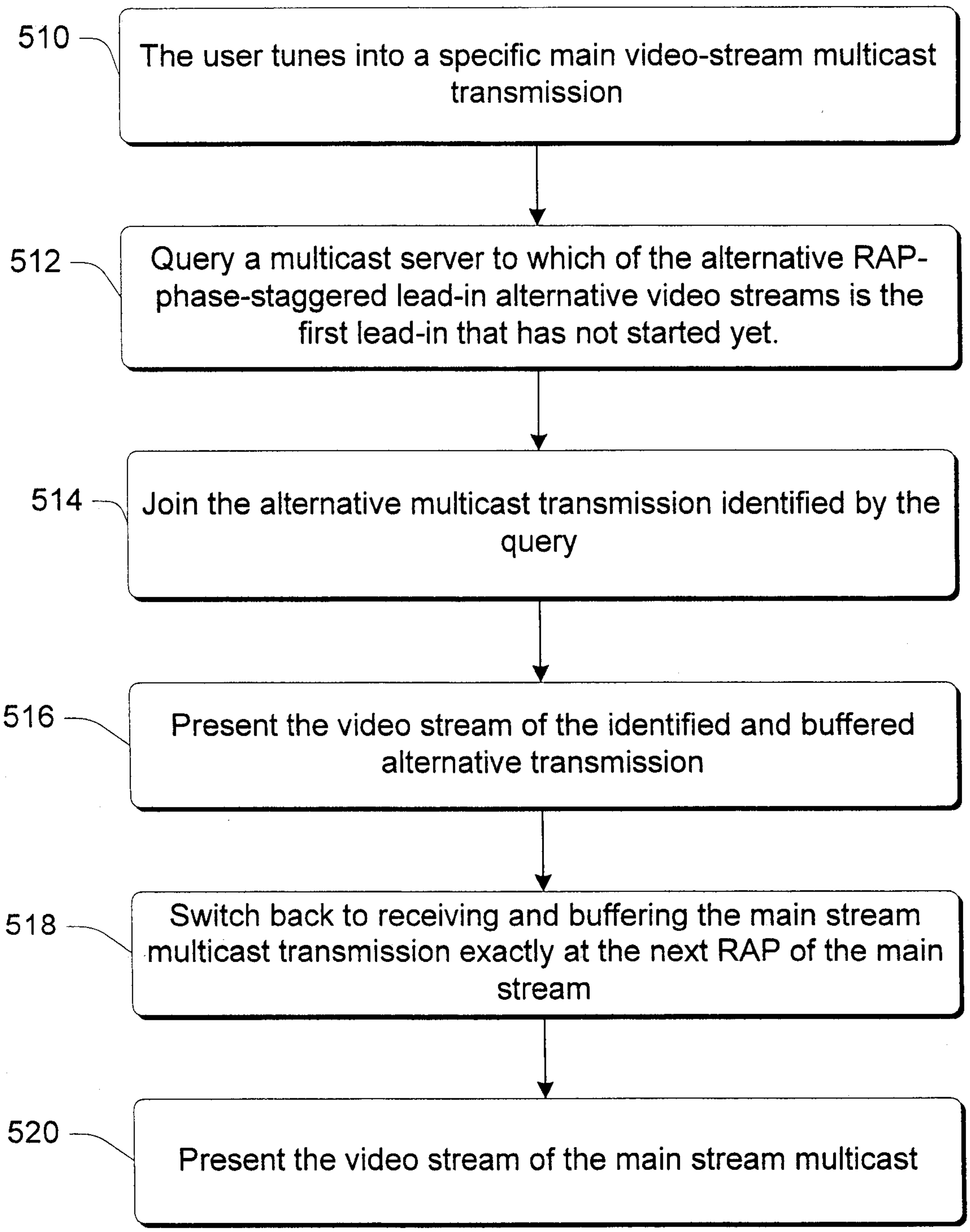

510
The user tunes into a specific main video-stream multicast transmission
512
Query a multicast server to which of the alternative RAP-phase-staggered lead-in alternative video streams is the first lead-in that has not started yet.
514
Join the alternative multicast transmission identified by the query
516
Present the video stream of the identified and buffered alternative transmission
518
Switch back to receiving and buffering the main stream multicast transmission exactly at the next RAP of the main stream
520
Present the video stream of the main stream multicast

Fig. 5

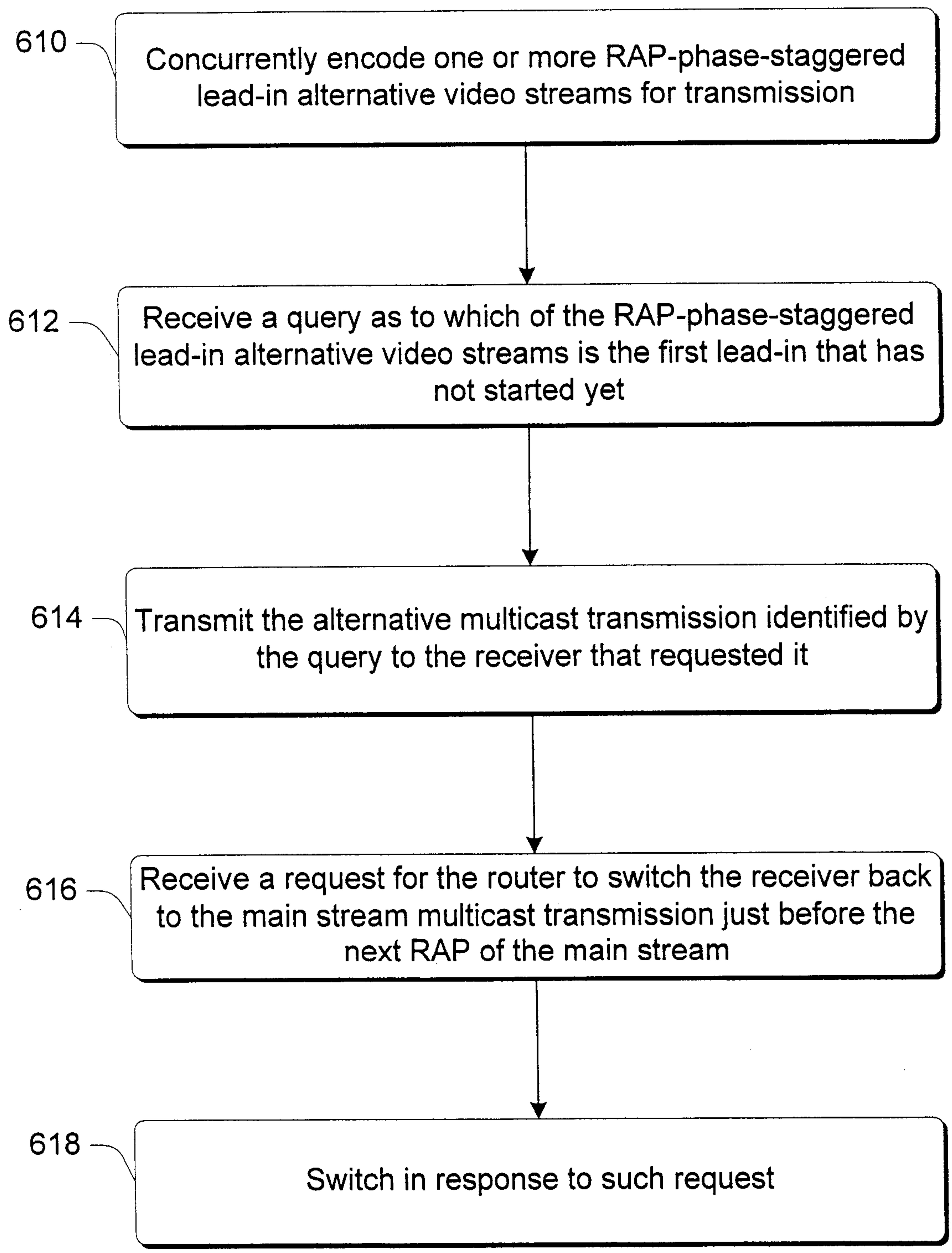

610
Concurrently encode one or more RAP-phase-staggered lead-in alternative video streams for transmission
612
Receive a query as to which of the RAP-phase-staggered lead-in alternative video streams is the first lead-in that has not started yet
614
Transmit the alternative multicast transmission identified by the query to the receiver that requested it
616
Receive a request for the router to switch the receiver back to the main stream multicast transmission just before the next RAP of the main stream
618
Switch in response to such request

Fig. 6

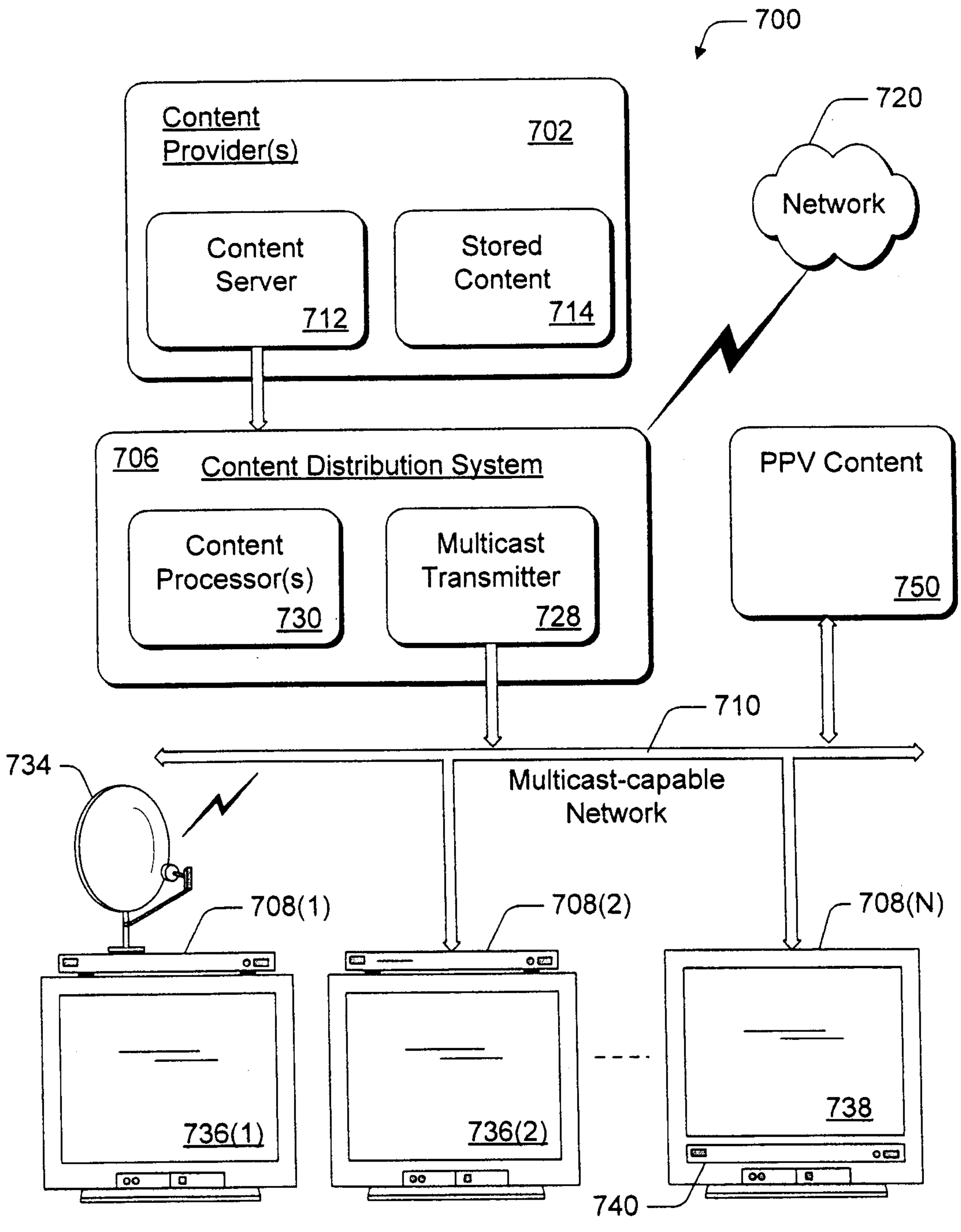

700
Content Provider(s)
702
Content Server
712
Stored Content
714
720
Network
706
Content Distribution System
Content Processor(s)
730
Multicast Transmitter
728
PPV Content
750
710
734
Multicast-capable Network
708(1)
708(2)
708(N)
736(1)
736(2)
738
740

Fig. 7

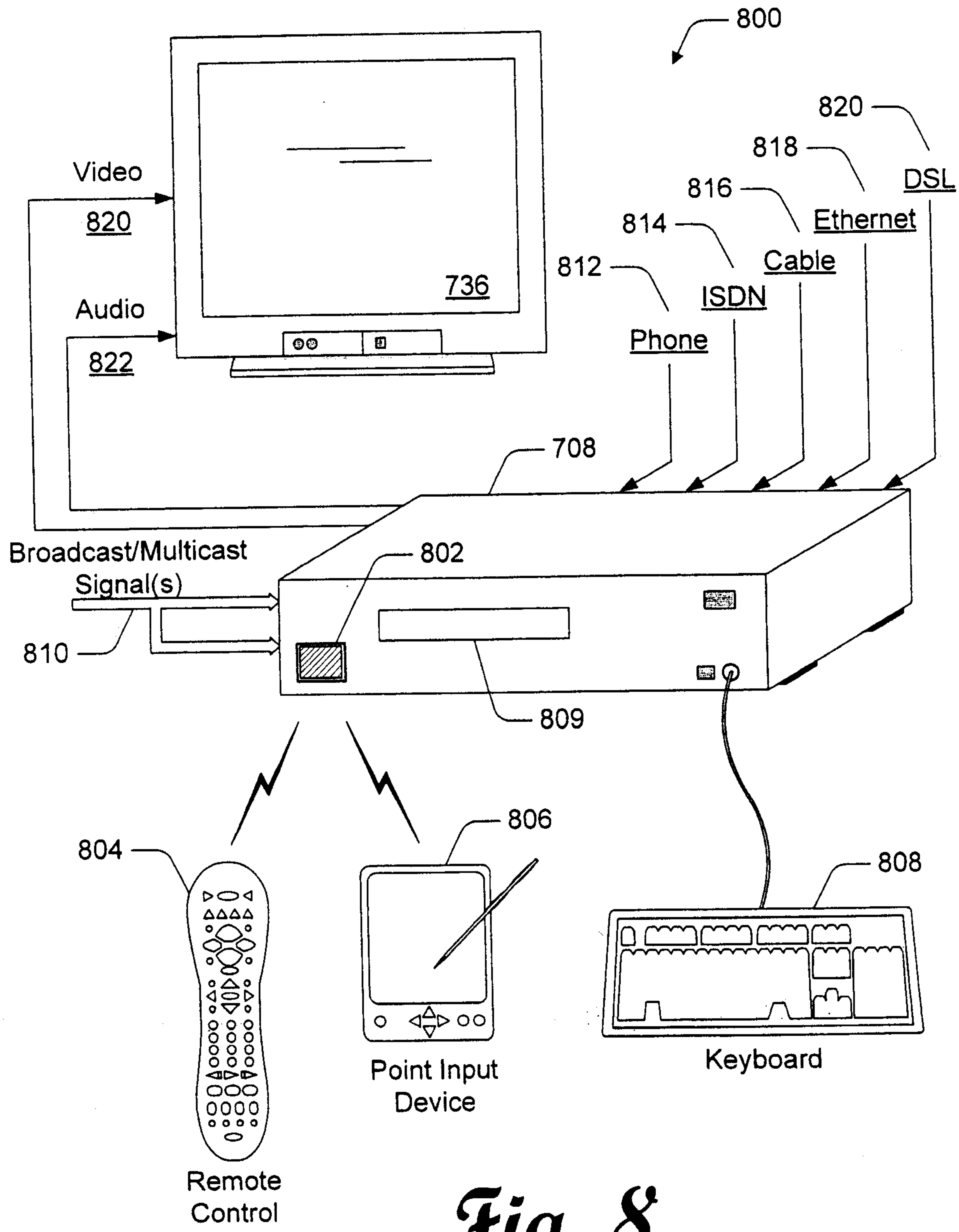

800
820
818
816
814
812
DSL
Ethernet
Cable
ISDN
Phone
Video
820
Audio
822
736
708
Broadcast/Multicast
Signal(s)
802
810
809
804
806
808
Remote
Control
Point Input
Device
Keyboard

Fig. 8

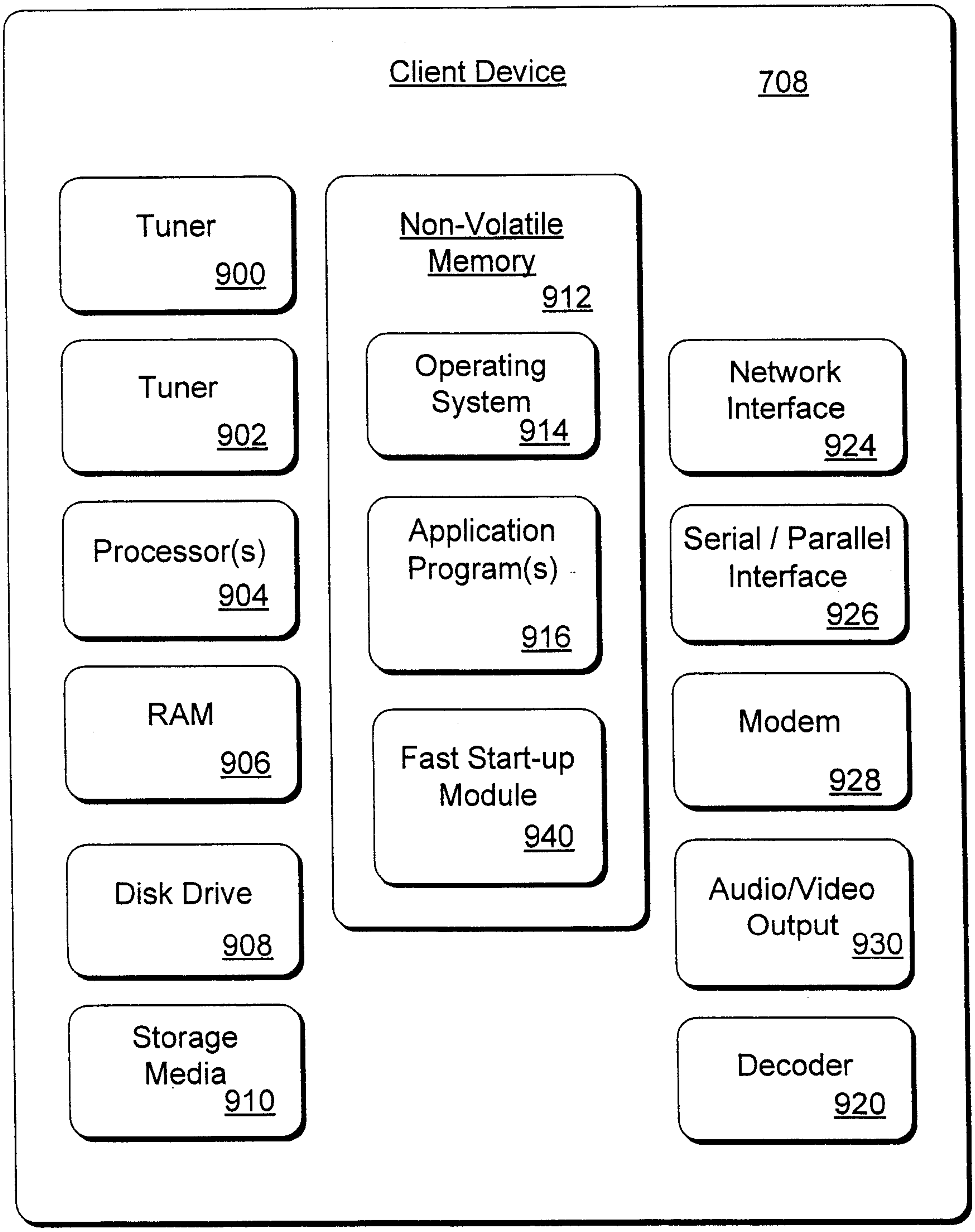
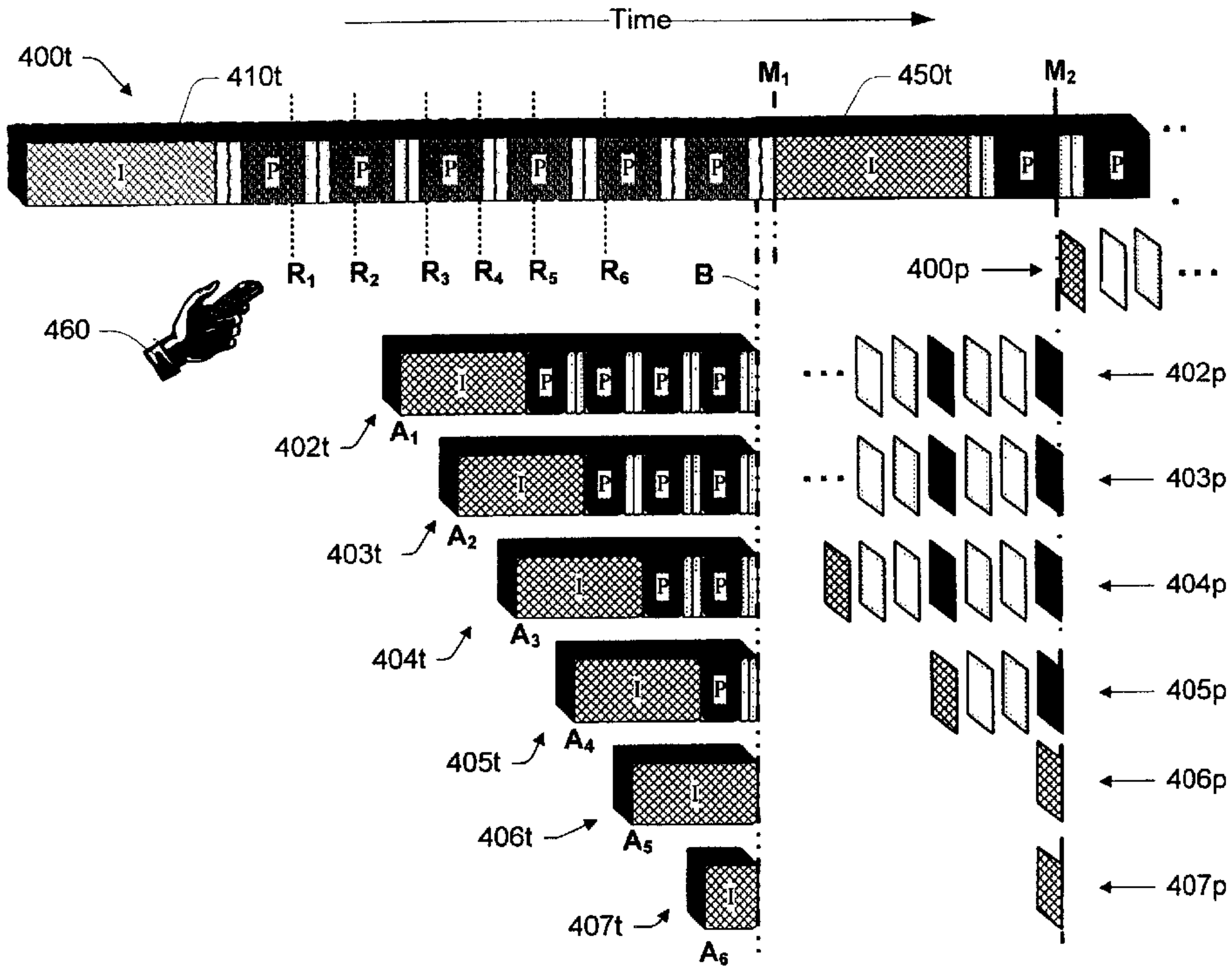

Client Device
708
Tuner
900
Tuner
902
Processor(s)
904
RAM
906
Disk Drive
908
Storage Media
910
Non-Volatile Memory
912
Operating System
914
Application Program(s)
916
Fast Start-up Module
940
Network Interface
924
Serial / Parallel Interface
926
Modem
928
Audio/Video Output
930
Decoder
920

Fig. 9


Time
400t
410t
M1
450t
M2
I
P
P
P
P
P
P
I
P
P
R1
R2
R3
R4
R5
R6
B
400p
460
I
P
P
P
P
402p
402t
A1
I
P
P
P
403p
403t
A2
I
P
P
404p
404t
A3
I
P
405p
405t
A4
I
406p
406t
A5
I
407p
407t
A6