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

(19) World Intellectual Property
Organization
International Bureau

(43) International Publication Date
9 December 2004 (09.12.2004)

PCT

(10) International Publication Number
WO 2004/107755 A1

(51) International Patent Classification⁷: H04N 7/12

(21) International Application Number:
PCT/KR2003/002761

(22) International Filing Date:
17 December 2003 (17.12.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10-2003-0035467   3 June 2003 (03.06.2003)   KR

(71) Applicant and
(72) Inventor: KANG, Sang Hyuk [KR/KR]; #106-1007
Hyundai Apt., 88 Dangsan 4ga, Yeongdeungpo-ku, Seoul
150-044 (KR).

(74) Agent: LEE, Won-Hee; 8th Fl., Sung-ji Heights II, 642-16
Yoksam-dong, Kangnam-ku, Seoul 135-080 (KR).

(81) Designated States (national): 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, HR, HU, ID, IL, IN, IS, JP, KE, KG,
KP, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK,
MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT,
RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR,
TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (BW, GH,
GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE,
ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE,
SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA,
GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— with international search report

[Continued on next page]

(54) Title: PACKET SCHEDULING METHOD FOR STREAMING MULTIMEDIA DATA
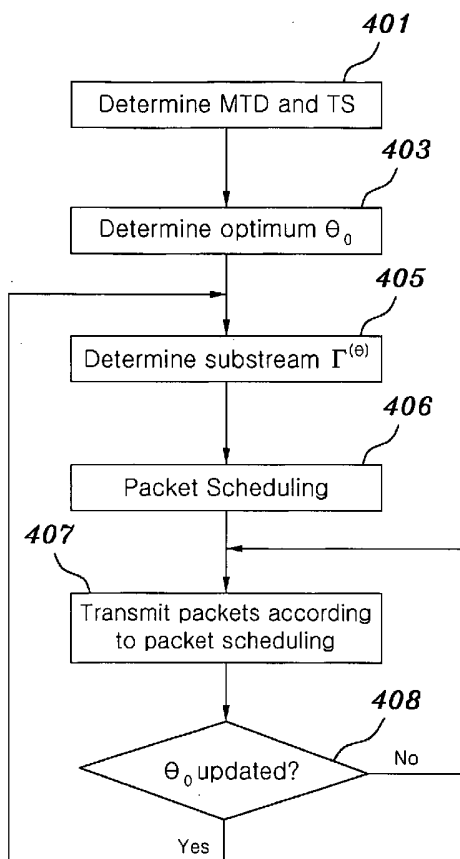
(57) Abstract: The present invention relates, in general, to streaming technology and, more particularly, to a packet scheduling method for streaming multimedia data. The packet scheduling method of the present invention streams multimedia data by a server in a network. The network includes the server for providing multimedia data divided into picture groups each having a sequence of N pictures, and a terminal for displaying the multimedia data received from the server in a streaming manner. In the packet scheduling method, the sequences of the pictures are divided into motion part packets and texture part packets, and priorities are assigned to the packets according to temporal scaling. A threshold for a predetermined priority is determined in consideration of conditions of a channel and a buffer status of the terminal and a substream is constructed using packets with priorities below the threshold within the respective picture groups. The packets in the constructed substream are sequentially transmitted to the terminal.

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.

PACKET SCHEDULING METHOD FOR STREAMING MULTIMEDIA DATA

Field of the Invention

The present invention relates, in general, to streaming technology and, more particularly, to a packet scheduling method for streaming multimedia data.

BACKGROUND ART OF THE INVENTION

Streaming is a technology of processing transmitted data like a continuous flow of water without interruption. With the development of the Internet, streaming technology has become more and more important. The reason for this is that most users do not have high-speed interface lines sufficient to immediately download large capacity multimedia files. By utilizing streaming technology, a client browser or plug-in can begin the display of data even before all the files are transmitted.

Especially, demands for technology of streaming moving picture data have explosively increased on the wired Internet. Therefore, many service providers, such as Internet movie theaters or Internet broadcasting stations, have appeared. Differently from real-time conversation type communication,

moving picture data streaming technology is characterized in that picture data to be transmitted is encoded in advance and stored in a server, and the playing of the moving picture data starts after an initial buffering time of approximately 5
5 through 20 seconds has elapsed when a request for the transmission of the moving picture data is received from a user.

Such streaming data consists of a plurality of packets which have several classes according to influences on service
10 quality when the streaming data are displayed, or preset priorities of data. An operation of determining the classes and transmitting streaming data to clients according to the determined classes by a server is designated as packet scheduling.

15   There are two reference documents for the above-described streaming data scheduling: 1) 'Markov decision process' disclosed in "Rate-distortion optimized streaming of packetized media" by P. A. Chou and Z. Miao, submitted to IEEE Trans. Multimedia, Feb. 2001 and 2) method disclosed in
20 "Expected run-time distortion based scheduling for delivery of scalable media" by Z. Miao and A. Ortega, submitted to Int'l Packetvideo Workshop 2002, Apr. 2002.

However, since the 'Markov decision process' disclosed in the reference document [1] uses a complicated algorithm, it is
25 difficult to apply the 'Markov decision process' to streaming

2

technology certainly requiring real-time implementation. The
method disclosed in the reference document [2] is an algorithm
of approximately and experientially calculating an expected
value of video quality distortion in real-time, and then
5   transmitting optimum video packets. For this algorithm, it is
required to measure packet loss probability in real time.
However, there is a problem in that it is difficult to measure
the packet loss probability, so that performance of the
algorithm is influenced by the precision of the measured
10   packet loss probability.

SUMMARY OF THE INVENTION

Accordingly, the present invention has been made keeping
15   in mind the above problems occurring in the prior art, and an
object of the present invention is to provide a packet
scheduling method for streaming multimedia data, which can
solve the above problems in which it is difficult to apply a
conventional 'Markov decision process' to streaming technology
20   certainly requiring real-time implementation, and the
performance of a conventional scheduling algorithm is
influenced by the precision of a measured packet loss
probability.
Another object of the present invention is to provide a
25   packet scheduling method, which can improve the service

quality of a streaming data service.

A further object of the present invention is to provide a packet scheduling method, which can obtain improved performance in consideration of a channel environment and the

5   size of a receiver buffer.

In order to accomplish the above object, the present invention provides a packet scheduling method for streaming multimedia data by a server in a network, the network including the server for providing multimedia data divided

10   into picture groups each having a sequence of N pictures and a terminal for displaying the multimedia data received from the server in a streaming manner, the method comprising the steps of dividing N-1 P-frames into motion part packets and texture part packets, and assigning priorities to the packets

15   according to temporal scaling, determining a threshold for a predetermined priority in consideration of conditions of a channel and a buffer status of the terminal and constructing a substream using packets with priorities below the threshold within the respective picture groups, and sequentially

20   transmitting the packets in the constructed substream to the terminal.

BRIEF DESCRIPTION OF THE DRAWINGS

25   The above and other objects, features and other

advantages of the present invention will be more clearly understood from the following detailed description taken in conjunction with the accompanying drawings, in which:

FIG. 1 is a view showing the construction of a network to which a packet scheduling method according to an embodiment of the present invention is applied;

FIGS. 2a to 2d are views showing temporal scaling and a layer structure of priorities according to the present invention;

FIG. 3 is a view showing a queuing model according to the present invention;

FIG. 4 is a flowchart of a packet scheduling method according to an embodiment of the present invention;

FIG. 5 is a view showing an example of the construction of a substream according to the present invention;

FIG. 6 is a view showing an example of the packet scheduling of the present invention;

FIG. 7 is a view showing result values obtained by applying the packet scheduling method of the present invention to a wireline network;

FIGS. 8a to 8e are views showing another example of result values obtained by applying the packet scheduling method of the present invention to the wireline network; and

FIGS. 9a and 9b are views showing result values obtained by applying the packet scheduling method of the present

invention to a wireless network.

DETAILED DESCRIPTION OF THE INVENTION

5   Hereinafter, embodiments of the present invention will be

described in detail with reference to the attached drawings.

Reference now should be made to the drawings, in which

the same reference numerals are used throughout the different

drawings to designate the same or similar components. In this

10   specification, a detailed description of related prior art or

constructions will be omitted if the detailed description

makes the gist of the present invention unclear.

FIG. 1 is a view showing the construction of a network to

which a packet scheduling method according to an embodiment of

15   the present invention is applied. Referring to FIG. 1, a

streaming data providing server 10 transmits encoded data to a

pre-roll buffer 30 connected to a client 40 through a network

20, such as the Internet. The buffer 30 provides the received

data to the client 40 by a preset method. The client 40

20   decodes and displays the data.

Related terms required for a later description of the

present invention are summarized in the following Table 1.

Table 1

Picture group	Group Of Pictures (GOP) ·
Sequence of	Packets or frames constituting GOP, in which one GOP
pictures	includes one independently decoded I-frame and N-1 P-

	frames decoded with reference to previous data
Motion Texture Discrimination(MTD)	Method of dividing respective frames of a sequence of pictures into a motion vector part and a texture part and assigning priority thereto. Hereinafter, a motion vector is referred to as a motion.
Temporal scaling	Method of decreasing temporal resolution of a video bit stream by reducing the number of video frames transmitted within a given time. That is, packets constituting one GOP can be classified into important packets and less important packets at the time of encoding, and a reception terminal can decode only important packets to display multimedia data.
Substream	Concept of substream proposed in the present invention means a sequence of packets reconstructed using packets having priorities determined by a predetermined algorithm, which will be described later, in each GOP.
Earliest Deadline First (EDF)	EDF means that an earlier packet is transmitted first. Hereinafter, EDF is referred to as "earliest packet first" transmission.
Receiver	Client terminal provided with a multimedia streaming service from a certain server
Video rate	Value obtained by dividing the size of entire video data by a playing time. For example, if there are moving pictures corresponding to 60 seconds and the size of entire moving picture data is 12 megabits, a video rate is 12Mbits/60seconds = 200,000 bps.

Hereinafter, the temporal scaling and the priority layer structure according to the present invention will be described in detail. Next, a scheduling method of constructing a substream will be described. Then, experimental results 5 obtained by the application of the present invention will be described.

A. Scaling and priority layer structure

FIGS. 2a to 2d are views showing the scaling and the priority layer structure according to the present invention. Hereinafter, the scaling and the layer structure are described with reference to FIGS. 2a to 2d.

5   If it is assumed that there is a video frame sequence Fn (n=0, 1, 2, ...) with a frame rate f, a data receiver displays received data at time t=0 sec, and a total display time is t=n/f sec.   Such a video frame sequence consists of a plurality of picture groups (hereinafter referred to as GOPs).

10   A j-th P frame in a GOP is designated as P_j (j= 1, 2,..., N-1).

FIG. 2a shows a case where frames constituting a GOP are not scaled.

FIG. 2b shows a MTD method, in which respective frames are divided into a motion part and a texture part, so that I-

15   frames and the motion part of P-frames form a layer 0, and the texture part of the P-frames forms a layer 1.   In this case, the layer 0 is a basic layer having a priority higher than the layer 1, which is an enhancement layer.

FIG. 2c shows a two-layer temporal scaling scheme.

20   Referring to FIG. 2c, a layer 0 is composed of I-frames and even-numbered P-frames of a picture sequence, and a layer 1 is composed of odd-numbered P-frames thereof.

FIG. 2d shows a layer structure of a combination of the MTD and the temporal scaling.   Referring to FIG. 2d, a layer 0

25   is composed of I-frames and a motion part of even-numbered P-

frames of a picture sequence. A layer 1 is composed of a motion part of odd-numbered P-frames of the picture sequence. A layer 2 is composed of a texture part of the even-numbered P-frames thereof. Finally, a layer 3 is composed of a texture

5   part of the odd-numbered P-frames thereof.

In the above layer structure, if the priority or importance level is h = 0, 1,..., H-1, h=0 represents highest priority. Therefore, as the layer becomes lower, the priority becomes higher. In the same layer, as an index number of a

10   frame becomes lower, the priority becomes higher.

B. Scheduling method

A principle of the scheduling method of the present invention determines a variable θ according to a predetermined

15   algorithm with conditions of a channel and buffer status of a terminal taken into consideration, and constructs a substream using packets with higher priorities within GOPs using the determined variable θ.

Data in each importance level within a GOP is segmented

20   into fixed-size packets. A random process $\{X_i^{(h)}; i = 0, 1, ...\}$ is defined to be the sequence of the size. In this case, the random process represents the sequence of the size of data with importance level h in an i-th GOP of consecutive GOPs. Further, the characteristics of the random process are assumed

25   to be Wide Sense Stationary (WSS). Further, the cumulative

sum $S_i^{(\theta)}$ (θ = 0, 1, ..., H-1) of the random process is defined as the following Equation [1].

$$S_i^{(h)} = \sum_{j=0}^{h} X_i^{(j)} \tag{1}$$

Further, autocovariance of $S_i^{(\theta)}$ is defined as the following Equation [2],

$$V[k] \equiv E[(S_i^{(h)} - \overline{S}^{(h)})(S_{i+k}^{(h)} - \overline{S}^{(h)})] \tag{2}$$

where k= 0, 1, ...

FIG. 3 is a view showing a queuing model according to the present invention.

An upper part of FIG. 3 shows the status of a buffer of a receiver, and a lower part of FIG. 3 shows a queuing model indicated in consideration of the status of the receiver buffer. In FIG. 3, Ch represents a channel, which can be, for example, the wired/wireless Internet.

If a substream proposed in the present invention is assumed to be $\Gamma^{(\theta)}$, each GOP is composed of data with importance levels h=0 through h=θ, for the entire video sequence. Further, if the length of a GOP is 1 sec, the size of data of an i-th GOP in the substream is represented by $S_i^{(\theta)}$, and an average data rate of the substream $\Gamma^{(\theta)}$ is represented by $E[S_i^{(\theta)}]$ in packets/sec. Next, the queuing model shown in FIG. 3 is introduced for the substream $\Gamma^{(\theta)}$. In this case, the queuing model has the stochastic substream $\Gamma^{(\theta)}$ as an input process, an output service rate equal to a channel bandwidth

C, and a fixed-size buffer with a size $B^{(\theta)}$. The buffer size $B^{(\theta)}$ must be carefully determined in consideration of the buffer status of the receiver. Especially, the $B^{(\theta)}$ is represented by the number of packets according to the average data rate of the substream $\Gamma^{(\theta)}$, and is related to a pre-roll duration for the substream $\Gamma^{(\theta)}$ in the receiver buffer.

The $B^{(\theta)}$ is calculated by Equation [3],

$$B^{(\theta)} = (m_\theta - k)E[S^{(\theta)}] \text{ [packets]} \qquad [3]$$

where both m_θ and k represent GOP numbers. If a GOP is composed of 16 frames I+15P, GOP numbers may be assigned in such a way that a first GOP is designated as GOP 0 and a GOP next to the GOP 0 is designated as GOP 1. A variable k represents a situation in which a video player of a receiver plays a certain frame at the present time and this frame is a frame of a GOP k. Further, m_θ represents a GOP number, indicating the amount of data to later be played, which is prestored in the receiver buffer. Therefore, m_θ is a value greater than k. If there are no data in the receiver buffer at the present time, m_θ has a value equal to k. If there are data corresponding to a GOP in the receiver buffer at the present time, m_θ has a value equal to k+1. More specifically, m_θ is determined to be a largest GOP number of GOPs for which all packets with importance level θ have been satisfactorily transmitted to the receiver without error. Further, if the playing of the video signals starts once by the receiver, the

frames are played at a constant speed according to a given frame per second value, which also means that GOPs are played at a constant speed. Therefore, GOP numbers m_θ and k are analyzed with respect to time.

5   In Equation [3], $B^{(\theta)}$ is determined in consideration of only a case where an importance level is equal to or less than θ. That is, if a GOP duration is 1 second, the receiver has a queuing time of a length $m_\theta - k$.

In the queuing model shown in the lower part of FIG. 3, 10 the average data rate $E[S_i^{(\theta)}]$ of the substream $\Gamma^{(\theta)}$ is less than a channel bandwidth C, and the input process suffers from packet losses due to video rate fluctuations. In order to apply the concept of effective bandwidth to the problem of the packet losses, a packet loss probability $\varepsilon_p^{(\theta)}$ of the substream 15 $\Gamma^{(\theta)}$ is defined as Equation [4],

$$\varepsilon_p^{(\theta)} = e^\beta \qquad [4]$$

where β is defined as Equation [5].

$$\beta = \frac{-2B^{(\theta)}(C - E[S^{(\theta)}])}{\sum_k v^\theta[k]} \qquad [5]$$

Relationships between the above-described probability 20 functions are described below. First, if $B^{(\theta)}$ decreases or the video rate fluctuations become large, the packet loss probability $\varepsilon_p^{(\theta)}$ increases. Further, the average data rate $E[S_i^{(\theta)}]$ of the substream $\Gamma^{(\theta)}$ increases, the packet loss

probability $\varepsilon_p^{(\theta)}$ increases. However, the packet loss probability $\varepsilon_p^{(\theta)}$ must be within a range of $[0, 1]$ and the average data rate $E[S_i^{(\theta)}]$ must be smaller than the channel bandwidth C.

5   In the substream $\Gamma^{(\theta)}$, a decoding failure due to the packet loss of any one motion part within a GOP is defined to be $\varepsilon_G^{(\theta)}$ as probability. That is, in the queuing model, the decoding failure probability $\varepsilon_G^{(\theta)}$ for the substream $\Gamma^{(\theta)}$ is defined as Equation [6].

10      $$\varepsilon_G^{(\theta)} = \begin{cases} \varepsilon_p^{(\theta)} E[S^{(\theta)}], \theta \le N-1 \\ \varepsilon_p^{(\theta)} E[S^{(N-1)}], \theta > N-1 \end{cases} \qquad\qquad [6]$$

It can be seen from the queuing theory that the decoding failure probability $\varepsilon_G^{(\theta)}$ is a monotonically increasing function of θ. In the scheduling method, the decoding failure probability $\varepsilon_G^{(\theta)}$ of the GOP is compared with a requirement

15 value (or threshold value) specified by a user so as to obtain an optimum θ.

If the decoding failure probability $\varepsilon_G^{(\theta)}$ must be less than an arbitrary value γ, the server determines an optimum importance level θ_0 by the following Equation [7] in the given

20 channel bandwidth C.

$$\theta_0 = \text{argmax}_\theta \{\varepsilon_G^{(\theta)} < \gamma, E[S^{(\theta)}]\} < C \qquad\qquad [7]$$

As shown in Equation [7], the optimum importance level is the largest value of the θ values such that the decoding

failure probability $\varepsilon_G^{(\theta)}$ is less than the arbitrary value γ, while keeping an average bit rate below the channel bandwidth C. Due to the limitation of the channel bandwidth, the packet loss probability $\varepsilon_G^{(\theta)}$ must be within the range of [0, 1], as

5 described above. In an ideal case, θ_o is an integer, and, thus, the server sets the optimum importance level to be changed between θ_o and θ_o+1 so as to regulate the receiver buffer and maximize the utilization of the channel bandwidth. Especially, a time spent at θ_o versus θ_o+1 relates to the

10 marginal capacity of the receiver buffer. If there is no marginal capacity in the receiver buffer, θ_o+1 is selected, while if there is any marginal capacity, θ_o is selected. That is, if there are many packets stored in the receiver buffer, it is determined that the channel status is superior, while if

15 there are few packets stored in the receiver buffer, it is determined that the channel status is inferior.

If θ_o is determined, importance levels are divided into two priorities. A high priority part represents a part with importance levels h = 0 through h = θ_o. A low priority part

20 represents a part with importance levels h = θ_o + 1 through h= H-1. In this case, packets with a high priority are scheduled to be transmitted first, and the packets with a relatively low priority are scheduled to be transmitted next. Within each priority, packets are transmitted according to an Earliest to

25 Deadline First (EDF) manner.

14

In the scheduling method according to the present invention, the receiver periodically measures throughput and reports the measured throughput to the server so as to take time-varying channel fluctuations and queue status in the receiver buffer into consideration. The server, having received the measured throughput from the receiver, calculates mean throughput by Exponentially Weighted Moving Average (EWMA) as expressed in Equation [8],

$$C \leftarrow aC + (1-a)Y_m \qquad\qquad [8]$$

where $0 < a < 1$ is satisfied, and Y_m is newly reported throughput by the receiver. For a packet switching network, a sending rate is calculated to be suitable for Transmission Control Protocol (TCP). At this time, the measurement of the packet loss probability and Round Trip Time (RTT) is required.

In order to calculate the variability of channel throughput (or channel capacity), the server calculates the variance σ^2_Y of the Y_m. Therefore, if channel fluctuation, that is, the variance σ^2_Y increases, the display quality at the receiver will be further degraded. Further, since the channel fluctuations and video rate fluctuations are statistically independent, a value β required to calculate the packet loss probability may be defined as Equation [9].

$$\beta = \frac{-2B^{(\theta)}(C - E[S^{(\theta)}])}{\sum_k v^{(\theta)}[k] + \sigma^2_Y} \qquad\qquad [9]$$

The server calculates the packet loss probability by

applying Equation [9] to the above Equation [4].

FIG. 4 is a flowchart of a packet scheduling method according to an embodiment of the present invention. The method, which will be described below, is performed by the server 10. Referring to FIG. 4, the above temporal scaling (TS) and the layer structure for dividing respective frames into motion and texture parts are determined at step 401. An optimum importance level θ_0 satisfying Equations [1] through [9] is determined at step 403. A substream $\Gamma^{(\theta)}$ is determined according to the determined optimum importance level θ_0 at step 405. The substream is constructed (packet scheduling) at step 406, and the transmission of the packets is started depending on the packet scheduling at step 407. The construction of the substream and the transmission of the packets will be described later. As described above, the receiver periodically measures throughput and reports the measured throughput to the server so that time-varying channel fluctuations and the queue status of the receiver buffer are taken into consideration. The server obtains the optimum importance level θ_0 through Equations [7] through [9] in consideration of the measured throughput, and determines whether to update the optimum importance level θ_0 at step 408. If it is determined that the optimum importance level θ_0 is updated, the server returns to step 405 to repeatedly perform steps 405, 406 and 407 using the updated optimum importance

level θ_0.

FIG. 5 is a view showing an example of the construction of a substream according to the present invention.

Referring to FIG. 5, a video sequence is composed of
5   sequential GOPs. First, if θ_0 is 0, a substream is composed of packets with importance level $h=0$ within respective GOPs, in which the packets are scheduled in order in such a way that a packet of an earlier GOP is arranged first. Next, if θ_0 is 1, the substream is composed of packets with importance levels
10   $h=0$ and $h=1$ within the respective GOPs, in which the packets are scheduled in such a way that packets of an earlier GOP are arranged first with a packet with higher importance level $h=0$ being arranged first. Next, if θ_0 is 2, the substream is composed of packets with importance levels $h=0$, $h=1$ and $h=2$
15   within the respective GOPs, in which the packets are scheduled in such a way that packets of an earlier GOP are arranged first with a packet with higher importance level $h=0$ being arranged first.

FIG. 6 is a view showing an example of the packet
20   scheduling of the present invention.

As described above, θ_0 is a periodically updated value, so that, if θ_0 is updated to 2 from 1 after packets with importance levels $h=0$ and $h=1$ are transmitted, the substream is scheduled as in the case of $\theta_0 = 2$ shown in FIG. 6, and a
25   packet with importance level $h=2$ is first transmitted in a

next round.

C: experimental data

Experimental conditions of the scheduling method according to the present invention are described below.

- Transmission of real video data (a talk show with 1670 frames and a movie with 1700 frames) through wireline and wireless channels

    - Quarter Common Intermediate Format (QCIF), 15 frames/sec

    - Using User Datagram Protocol (UDP) as protocol

    - Receiver buffer having pre-roll buffering time of 10 sec

    - Moving Picture Experts Group (MPEG)-4 codec

    - GOP = 16 frames (1 I-frame and 15 P-frames)

    - Wireline network: between Korea Advanced Institute of Science and Technology (KAIST) of Korea and University of California, Berkeley, with artificial error added

    - Wireless network: IEEE 802. 11

FIG. 7 is a view showing result values obtained by applying the packet scheduling method of the present invention to a wireline network. Referring to FIG. 7, SS represents typical sequential sending, EBS represents effective bandwidth scheduling, EBS (MTD) represents a case where differential priorities are assigned to motion and texture parts of the

present invention, EBS (TS) represents a case where the temporal scaling of the present invention is applied to the wireline network, and EBS (MTD + TS) represents a case where the motion and texture parts of the present invention are
5 discriminated and the temporal scaling is applied to the wireline network. It can be seen that the EBS (MTD + TS) outperforms all other schemes at channel throughputs below 95 Kbps.

FIGS. 8a to 8e are views showing another example of
10 result values obtained by applying the packet scheduling method of the present invention to the wireline network. FIGS. 8a to 8e illustrate the results obtained by comparing the SS with the EBS (MTS + TS).

Referring to FIGS. 8a to 8e, FIG. 8a shows a video bit
15 rate and FIG. 8b shows a channel bandwidth (BW). Referring to FIG. 8c, in the case of the SS scheme, an importance level θ_0 is fixed to 30, but, in the case of the EBS (MTD + TS) scheme, the importance level θ_0 is adaptively varied according to the fluctuations of a channel capacity and the status of the
20 receiver buffer. FIG. 8d is a view comparing the performance of both the SS and EBS schemes, in which, in the case of the SS scheme, performance is greatly deteriorated if the video bit rate is rapidly increased or a channel error becomes heavy. The performance of the two schemes, compared with each
25 other in terms of Peak Signal to Noise Ratio (PSNR) [dB],

shows that an average PSNR of the EBS (MTD + TS) scheme is higher than that of the SS scheme by 4.2. FIG. 8e shows the status of the receiver buffer in the case where the EBS (MTD + TS) scheme is applied to the wireline network.

5   FIGS. 9a and 9b are views showing result values obtained by applying the packet scheduling method of the present invention to a wireless network. FIG. 9a shows the transmission of a talk show, in which the average PSNR of the EBS (MTD + TS) scheme is higher than that of the SS scheme by 10 8.8.

Although the preferred embodiments of the present invention have been disclosed for illustrative purposes, those skilled in the art will appreciate that various modifications, additions and substitutions are possible, without departing 15 from the scope and spirit of the invention as disclosed in the accompanying claims. For example, the number of frames within each GOP and the determination of an important frame for temporal scaling can be modified. Therefore, the scope of the present invention cannot be limited to the above embodiments, 20 and must be defined by the equivalents of the following claims as well as the following claims.

INDUSTRIAL APPLICABILITY

As described above, the present invention provides a packet scheduling method for streaming multimedia data, which is based on an effective bandwidth and can be applied to a streaming service with an improved quality.

WHAT IS CLAIMED IS

1. A packet scheduling method for streaming multimedia
data by a server in a network, the network including the
server for providing multimedia data divided into picture
groups each having a sequence of N pictures and a terminal for
displaying the multimedia data received from the server in a
streaming manner, the method comprising the steps of:

dividing the sequences of the pictures into motion part
packets and texture part packets, and assigning priorities to
the packets according to temporal scaling;

determining a threshold for a predetermined priority in
consideration of conditions of a channel and a buffer status
of the terminal and constructing a substream using packets
with priorities below the threshold within the respective
picture groups; and

sequentially transmitting the packets in the constructed
substream to the terminal.

2. The packet scheduling method according to claim 1,
wherein the motion part packets are assigned priorities higher
than those of the texture part packets.

3. The packet scheduling method according to claim 1,
wherein the priorities of the packets are assigned so that an

I-frame and a motion part of even-numbered P-frames of each picture sequence are assigned a highest priority and a texture part of odd-numbered P-frames of each picture sequence is assigned a lowest priority.

4. The packet scheduling method according to claim 1, wherein the threshold θ_0 for the predetermined priority is determined by the following equation:

$$\theta_0 = \text{argmax}_\theta \{\varepsilon_G^{(\theta)} < \gamma, \ E[S^{(\theta)}]\} < C$$

where θ is 0, 1, ... and is equal to a number of packets to which priorities are assigned, γ is a threshold of a preset decoding failure probability, $\varepsilon_G^{(\theta)}$ is the decoding failure probability, $E[S^{(\theta)}]$ is an average data rate of the substream, and C is a channel bandwidth.

5. The packet scheduling method according to claim 1, wherein the threshold for the predetermined priority is updated according to variations of data throughput periodically reported from the terminal.

6. The packet scheduling method according to claim 1, wherein the packets received by the terminal have a loss rate of $\varepsilon_p^{(\theta)}$ calculated by the following Equations:
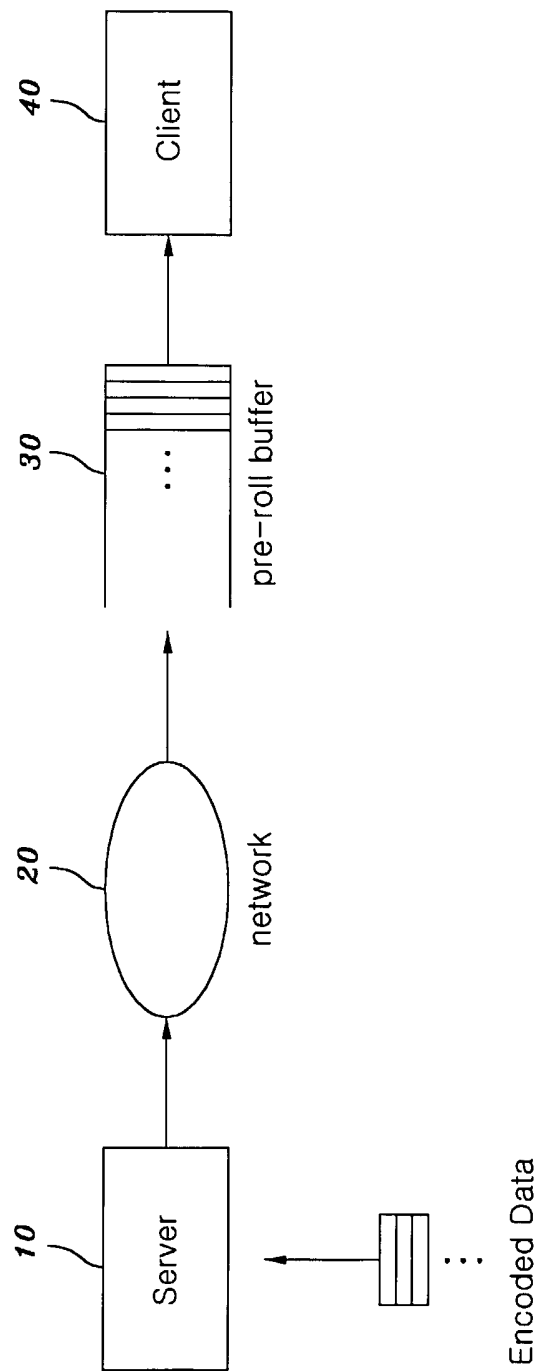
$$\varepsilon_p^{(\theta)} = \varepsilon^\beta;$$

$$\beta = \frac{-2B^{(\theta)}(C - E[S^{(\theta)}])}{\sum_k v^{(\theta)}[k]} \ ; \quad \text{and}$$

$$\varepsilon_G^{(\theta)} = \begin{cases} \varepsilon_p^{(\theta)} E[S^{(\theta)}], \theta \leq N-1 \\ \varepsilon_p^{(\theta)} E[S^{(N-1)}], \theta > N-1 \end{cases} .$$

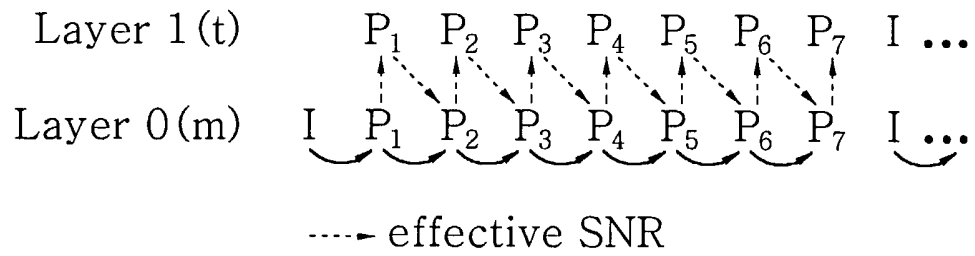7. The packet scheduling method according to claim 1,
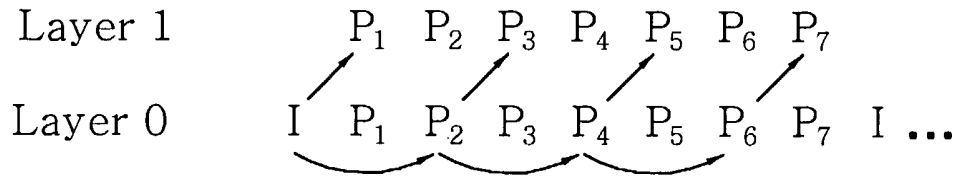5 wherein the packets received by the terminal have a loss rate

of $\varepsilon_p^{(\theta)}$ calculated in consideration of a variance of a channel

bandwidth by the following equations:

$$\varepsilon_p^{(\theta)} = \varepsilon^\beta;$$

$$\beta = \frac{-2B^{(\theta)}(C - E[S^{(\theta)}])}{\sum_k v^{(\theta)}[k] + \sigma_Y^2} \ ; \quad \text{and}$$

10 $$\varepsilon_G^{(\theta)} = \begin{cases} \varepsilon_p^{(\theta)} E[S^{(\theta)}], \theta \leq N-1 \\ \varepsilon_p^{(\theta)} E[S^{(N-1)}], \theta > N-1 \end{cases} .$$

1/17

FIGURES

FIG. 1

2/17

FIG. 2a

$$I \quad P_1 \quad P_2 \quad P_3 \quad P_4 \quad P_5 \quad P_6 \quad P_7 \quad I \, \ldots$$

FIG. 2b

Layer 1(t) P_1 P_2 P_3 P_4 P_5 P_6 P_7 I ...

Layer 0(m) I P_1 P_2 P_3 P_4 P_5 P_6 P_7 I ...

----► effective SNR

FIG. 2c

Layer 1 P_1 P_2 P_3 P_4 P_5 P_6 P_7

Layer 0 I P_1 P_2 P_3 P_4 P_5 P_6 P_7 I ...

FIG. 2d

Layer 3 (t)　　　P_1　P_2　P_3　P_4　P_5　P_6　P_7

Layer 2 (t)　　　P_1　P_2　P_3　P_4　P_5　P_6　P_7

Layer 1 (m)　　　P_1　P_2　P_3　P_4　P_5　P_6　P_7

Layer 0 (m)　　I　P_1　P_2　P_3　P_4　P_5　P_6　P_7　I ...

FIG. 3
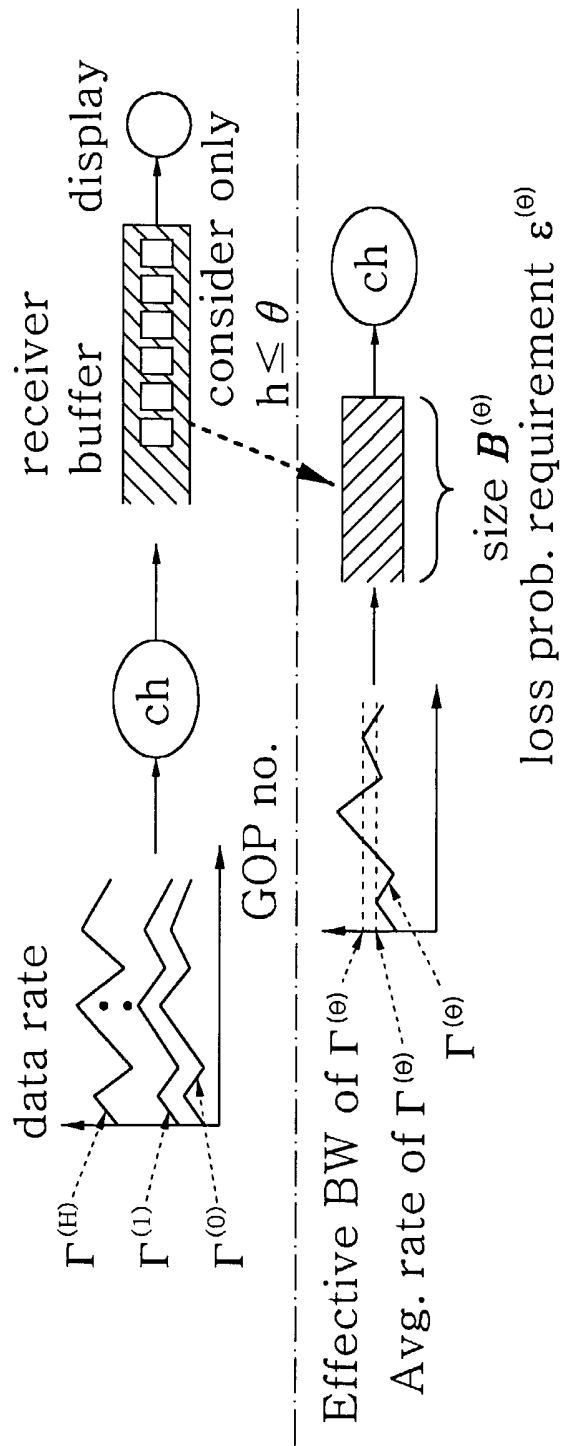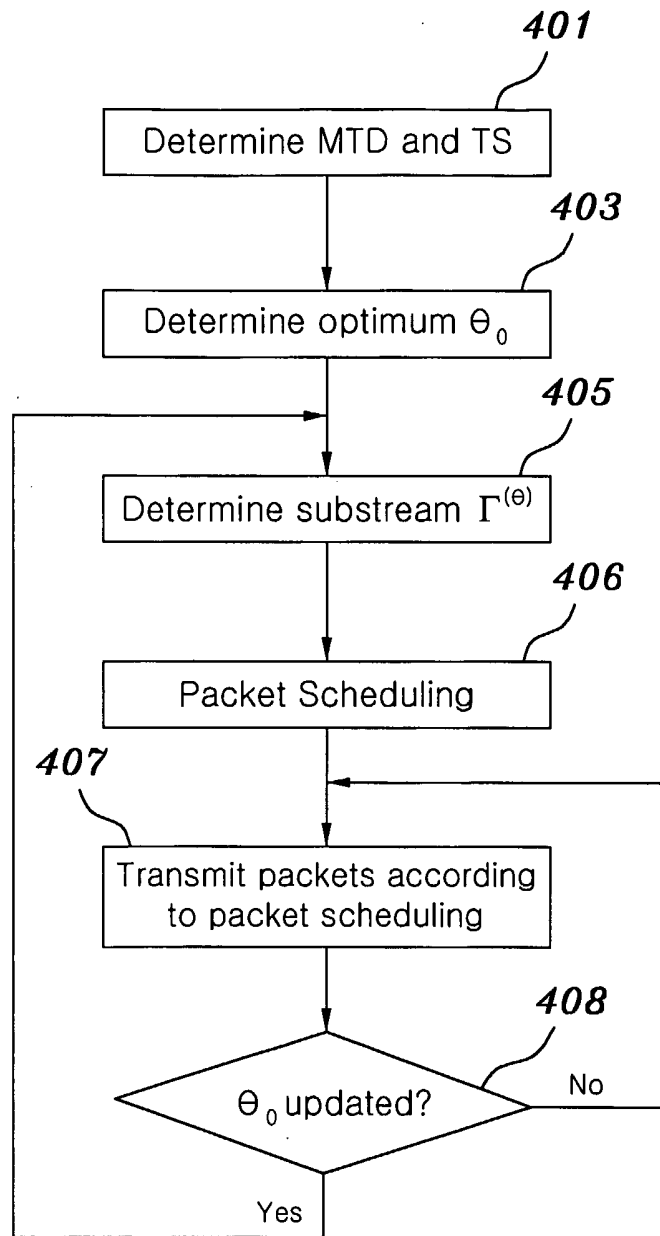
FIG. 4

401

Determine MTD and TS

403

Determine optimum Θ_0

405

Determine substream $\Gamma^{(\theta)}$

406

Packet Scheduling

407

Transmit packets according
to packet scheduling

408

Θ_0 updated?   No

Yes

FIG. 5

FIG. 6

FIG. 7

FIG. 8a

(a) Video bit rate

12/17

FIG. 8b

(b) Channel BW

FIG. 8c

(c) Secope of high priority

14/17

FIG. 8d

(d) Performace

15/17

FIG. 8e

(e) Receiver burrer status(EBS)
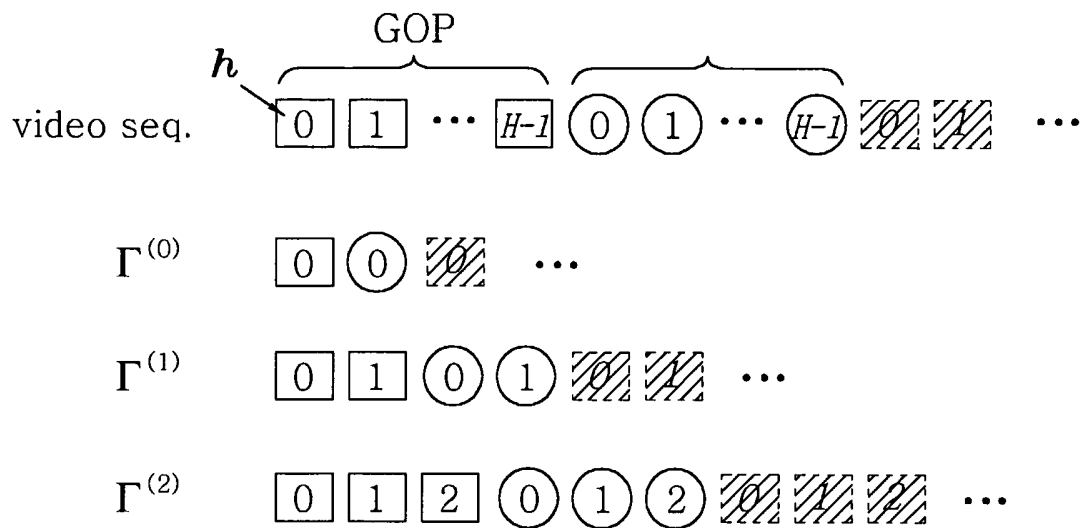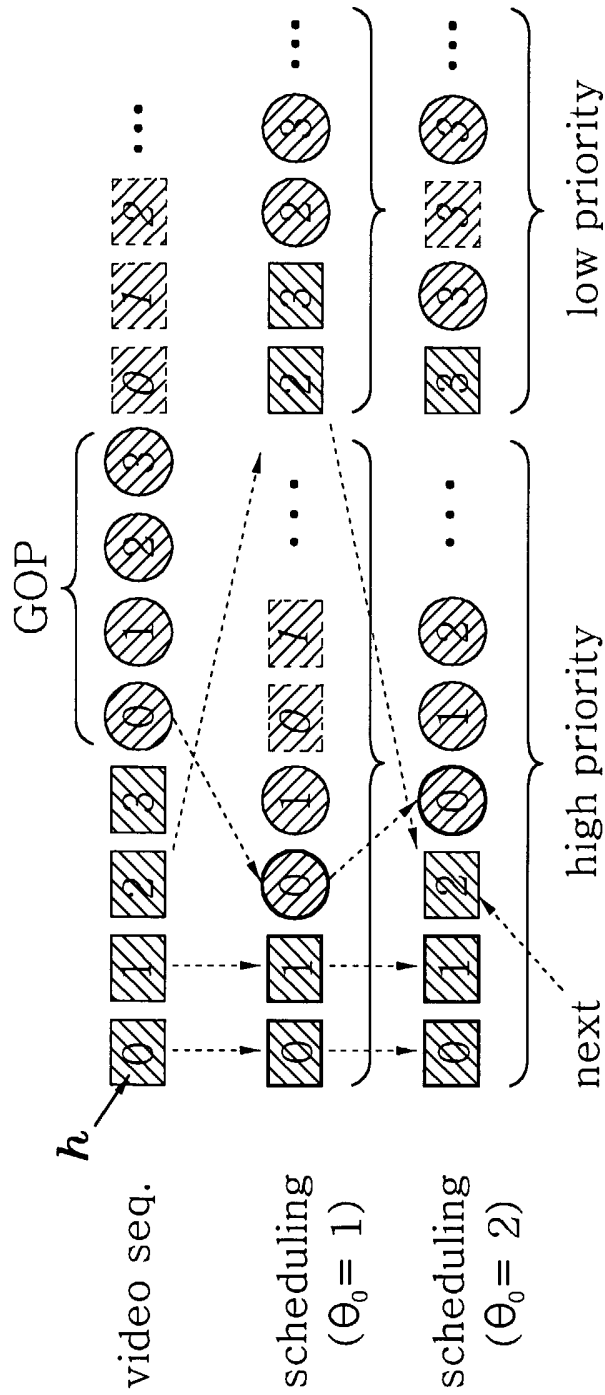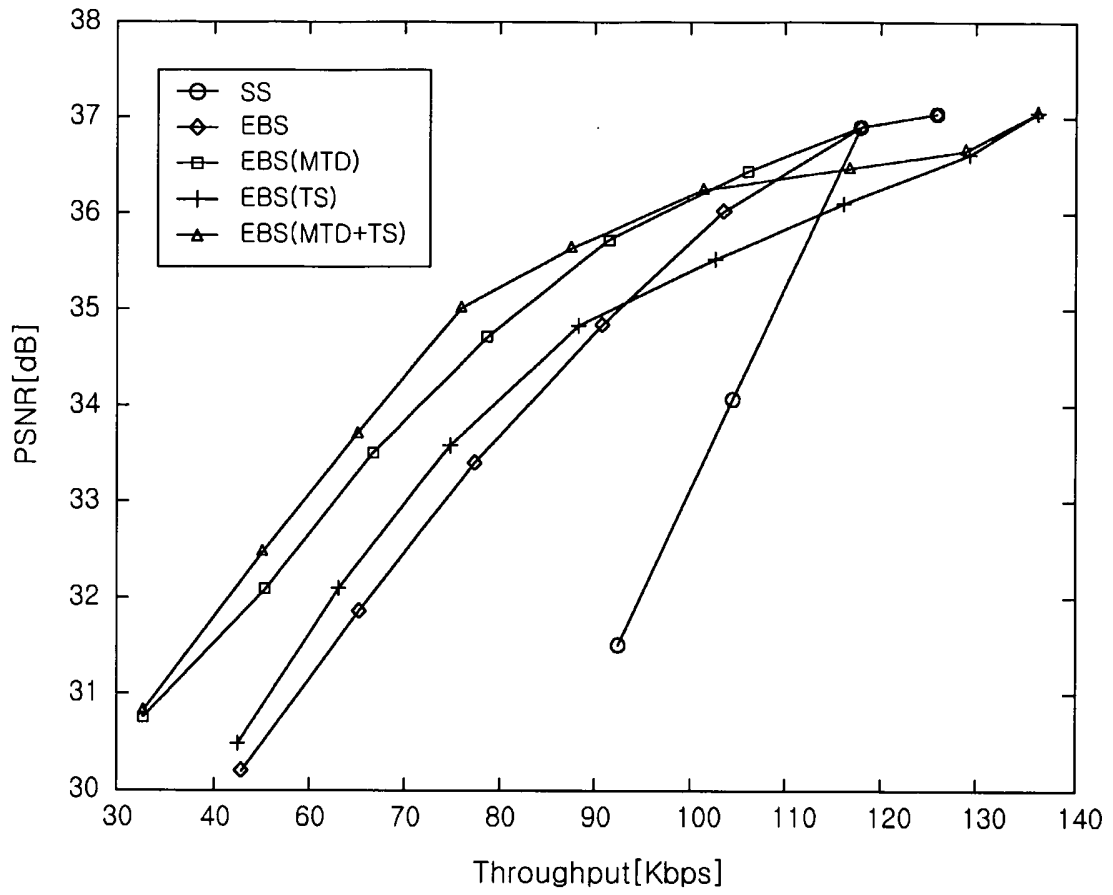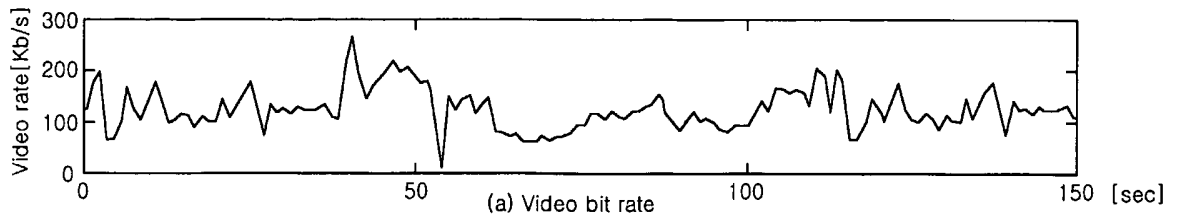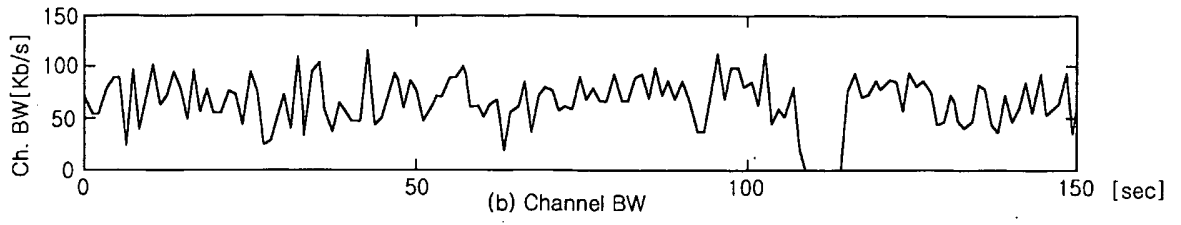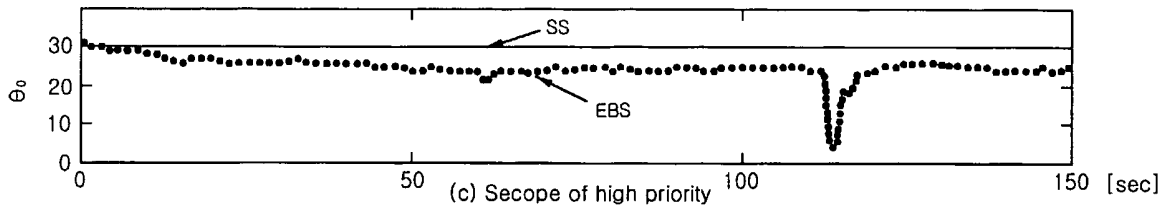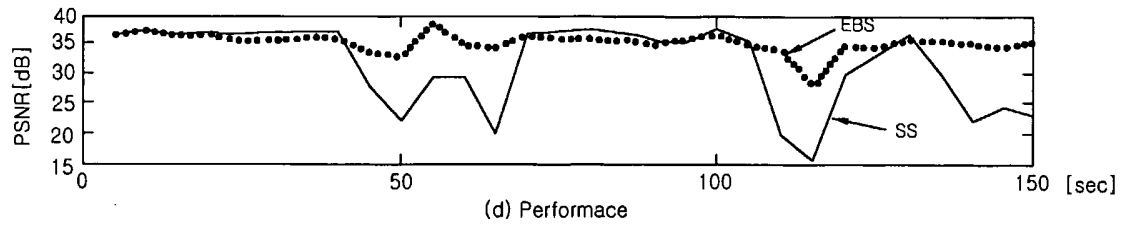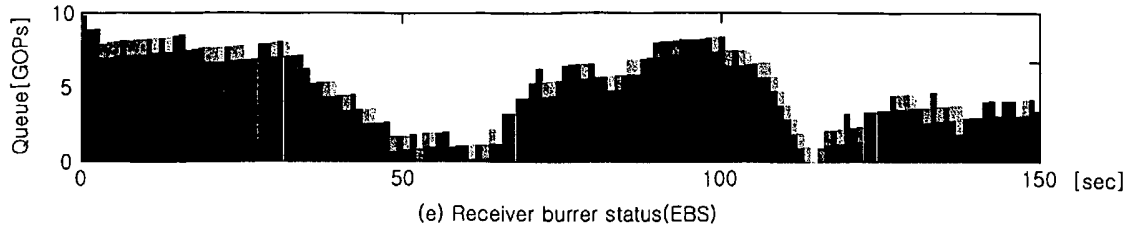
16/17

FIG. 9a

(a) Talk show

FIG. 9b

(b) Movie trailer
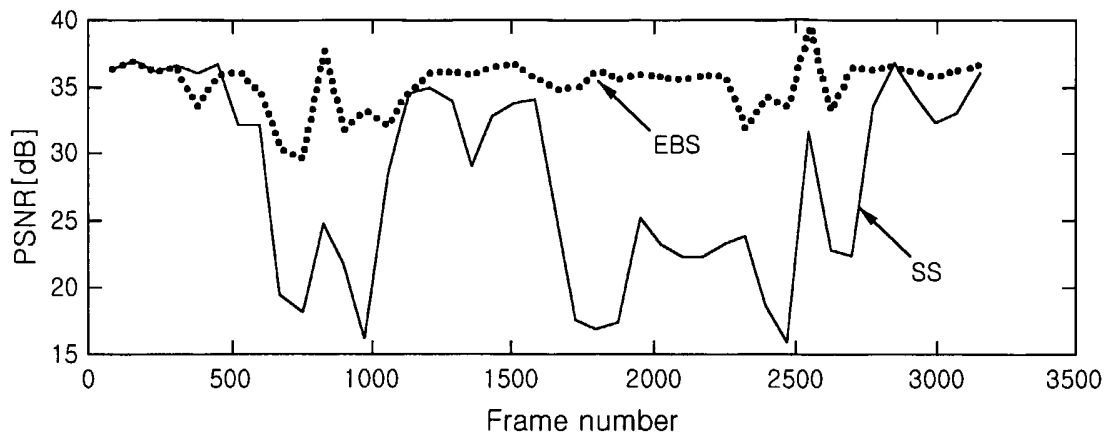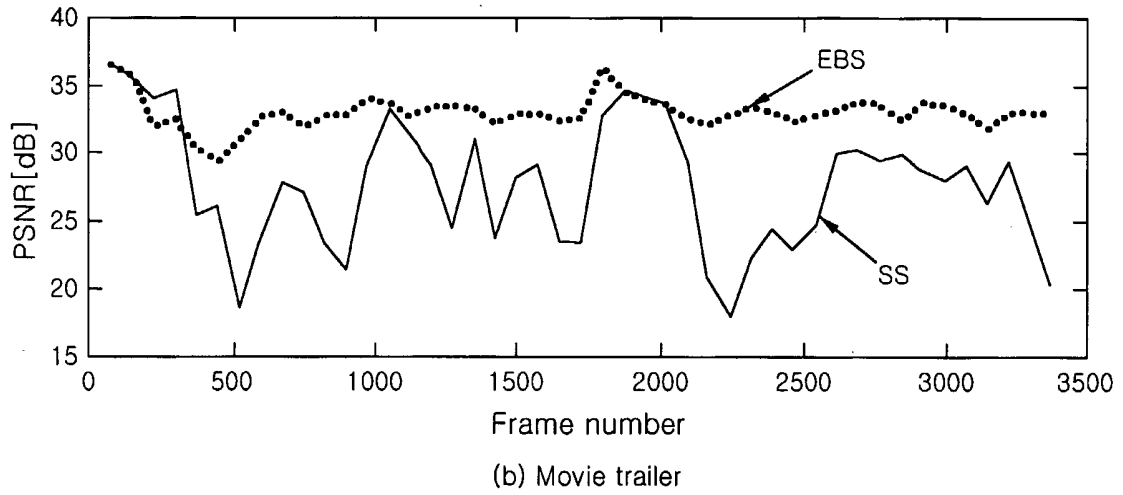
INTERNATIONAL SEARCH REPORT

	International application No.
	PCT/KR2003/002761

A.	CLASSIFICATION OF SUBJECT MATTER

IPC7 H04N 7/12

According to International Patent Classification (IPC) or to both national classification and IPC

B.	FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC7 G06F13/00, H04L12/56, H04N 7/12

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Korean Patents and applications for inventions since 1975
Korean Utility models and applications for Utility models since 1975

Electronic data base consulted during the intertnational search (name of data base and, where practicable, search terms used)
eKIPASS

C.	DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US6047000A (THE HONG KONG UNIVERSITY OF SCIENCE & TECHNOLOGY) 04.APRIL.2000 -See the whole document	1-7
A	US6134217A (THE REGENTS OF THE UNIVERSITY OF CALIFORNIA)17.OCTOBER.2000 -See the whole document	1-7
A	US6262986A (KABUSHIKI KAISHA TOSHIBA) 17.JULY.2001 -See the whole document	1-7
A	US5761417A (INTERNATIONAL BUSINESS MACHINES CORPORATION) 02.JUNE.1998 -See the whole document	1-7

☐ Further documents are listed in the continuation of Box C.	☒ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents,such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
06 FEBRUARY 2004 (06.02.2004)	06 FEBRUARY 2004 (06.02.2004)

Name and mailing address of the ISA/KR	Authorized officer
Korean Intellectual Property Office 920 Dunsan-dong, Seo-gu, Daejeon 302-701, Republic of Korea	KWON, KI WON
Facsimile No. 82-42-472-7140	Telephone No. 82-42-481-8123

Form PCT/ISA/210 (second sheet) (January 2004)

INTERNATIONAL SEARCH REPORT

Information on patent family members

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US6047000A	04.04.2000	US6047000A	04.04.2000
US6134217A	17.10.2000	US6134217A	17.10.2000
US6262986A	17.07.2001	JP9083547A2	28.03.1997
		US6262986A	17.07.2001
US5761417A	02.06.1998	CA2154038A1	09.03.1996
		JP8154236A	11.06.1996
		KR192723B1	15.06.1999
		SG54076A1	16.11.1998
		US5761417A	02.06.1998