relaxation of a transient constriction by means of the aforementioned series of packet trains in which the peak packet flow rate exceeds the delivered flow rate.

[00113] The buffer manager 1470 accepts data from the flow meter 1465, which data is retained until it is consumed or passed over the content consumer 1480. The buffer will typically be implemented as a ring buffer in memory or on disk. The write cursor will lag the read cursor by a no less than a specifiable time interval so as to permit rewinds of acceptable scope without requiring selective data re-retrieval from the server.

Flow Optimizer 1430

[00114] The flow optimizer 1430 assumes control over a session once it has been launched, and its channel capacity has been determined. Subsequently, it modulates the flow rate set-point of every session such as to optimize aggregate flow or charges across all active sessions, subject to a variety of min/max constraints on sessions flows, and a maximum constraint on aggregate session flow, as previously discussed.

[00115] The flow optimizer 1430 continues to modulate session flow until all content has been delivered, or one of a number of session anomalies intervene, as follows:

[00116] A downgraded session flow capacity drops below the minimum session flow rate. The session is dropped by the optimizer 1430 as a solution involving this channel does not exist. The user is forced to apply for re-admission, with the option of accepting slower than real-time delivery at the present encoding rate, or accepting a degraded viewing experience at a reduced encoding rate better adapted to the presently available bandwidth.

[00117] A pause VCR command is received from the client. Flow to the client can continue until the client-side buffer is topped off, after which time flow drops to zero. The session is permanently dropped should the pause persists too long. Until that time, the CAC safety margin is temporarily increased by the jit (Just-In-Time) flow rate of the paused session. In this way the session can be resumed without the need for re-admission and the jit bandwidth is made available to currently sessions so long as the pause persists.

[00118] A VCR Jump-Forward command is received from the client. This action is problematic as it causes the just-in-time flow rate (calculated at the new forward offset) to increase, in violation of the assumptions of the CAC algorithm. It must be noted that this violation will occur even if the content that is sought has already been delivered to the client-side buffer. One workable policy is to force the client to apply for re-admission at the forward offset and with a suitably modified buffer level.

[00119] A VCR Jump-Backward (Rewind) command is received. This action may entail the need for re-transmission of content, as previously discussed. The just-in-time flow rate will decrease as per the new position and adjusted buffer contents.

Tunable Minimum Constraint

[00120] A parameterized family of minimum flow constraints can be defined as follows:

(22)

wherein: $f_k(t)$ = Target content flow rate for session k;

$f_k^{JIT}(t)$ = Just-in-time flow rate for session k;

(RHS of Equation (5))

β = Burst tuning factor (positive or zero); and

$f_k^{min}(t)$ = Minimum (reserved) flow rate for session k.

[00121] This family converges on $f_k(t) >= f_k^{JIT}(t)$ when β approaches 0. A β value of .1, for example, ensures a minimum flow rate that exceeds the JIT flow rate by 10%. Thus, non-zero values of β force early delivery of content to a player's media buffer, with beneficial effects on the viewing experience.

Alternative Call Admission Control

[00122] The traditional CAC algorithm compares the flow rate requirement of a prospective client with the unused bandwidth (obtained by subtracting the aggregate flow target from the server flow capacity) and grants service if unused bandwidth is not exceeded. This procedure may break down as the optimizer causes all available bandwidth to be used if it can, thereby reducing unused bandwidth to zero in the best case, even when a new client could be accommodated with a healthy margin to spare.

[00123] Nevertheless, a simple variation on this idea can be used. Available bandwidths can be defined as the difference between server flow capacity and the sum of minimum flow rates for each session computed according to (22) as follows:

(23) $f_{available} = f_{Svr}^{max} - (f_1^{min} + \ldots + f_n^{min})$

wherein: $f_{available}$ = Available bandwidth

f_{Svr}^{max} = Server aggregate flow capacity; and

f_n^{min} = Minimum (reserved) flow rate for session n.

[00124] Service is granted if the computed minimum flow rate of the prospective client is less than available bandwidth even when unused bandwidth is zero.

Toward a Tunable QOS

[00125] In the preceding discussion it is assumed that all clients share the same value of β. One might also consider a scheme whereby a client receives a β value according to the class/cost of service that client has signed up for: the greater the cost (and expected quality of service, or QOS) the greater the value of β.

[00126] Instead of maximizing charges by the methods described above, a variation on the algorithm to maximize flow is considered whereby the reserve flow computation for every session takes into account the individual β value of the associated client. In this fashion a single algorithm can serve either end. The optimizer first ensures that every session receives its required minimum bandwidth (which would now typically depend on β and elapsed session time) before allocating any excess bandwidth more or less evenly to all by interpolation. (According to Figure 15.)

[00127] The scheme is simpler computationally than the iterative interpolation method shown in Figure 16; it is also arguably more fair and stable with respect tot he allocation of available bandwidth.

Deadline-Driven Content Distribution

[00128] In delivering viewable or audible content to a player, the optimization algorithm imposes a minimum flow rate that ensures that all content is delivered by the implied deadline, which is the last moment of a play session. Notwithstanding, the optimizer attempts to deliver content earlier according to bandwidth availability within the server taken as a whole. A similar, though far simpler scenario arises relative to the distribution of content from an optimized server to a target device, whenever the delivery must be accomplished by a designated time, or equivalently, within a given elapsed time, and when overlapped non-blocking consumption of the content at a fixed rate is not required.

[00129] One example concerns the application where a session bundle comprises a session associated with the featured program together with sessions for every advertisement scheduled to interrupt the feature program at a designated time. Each session is endowed with a separate media buffer, which, in the case of ads, would typically be large enough to accommodate the ad in full, thereby permitting a seamless transition to the ad at the appointed time.

[00130] In adapting the optimizer to the needs of content distribution we must revisit the formulation of the flow constraints, as follows:

1. Media buffers are not of a concern, let alone buffer overflow, and it can be assumed that all content received is saved away in a file on disk, or in memory awaiting further use.

2. For these same reasons, the value of β may be zero, at which the minimum flow rate equals the just-in-time rate obtained by dividing the content as yet undelivered by the time till deadline:

(24) $f_k^{JIT}(t_k) = L_k^{TOGO} / (T_k - t_k)$

wherein: L_k^{TOGO} = Content to go for session k, i.e., as yet undelivered by the server at time t, (client.lookup (k).session.payloadToGo).

Given the arbitrary format of the file to be distributed, T no longer signifies the "duration of play" at the average play rate (averagePlayRate); rather, T represents the elapsed time to the deadline measured at the moment the delivery session was launched, and t represents the elapsed time since session launch.

Session Bundles

[00131] A session bundle is a set of sessions that flow along a shared channel, and are thus subject to an aggregate flow constraint equal to the capacity of the shared channel. One important application of session bundles is connected with ad insertion.

[00132] It is not enough to treat each of the constituent sessions as independent with respect to the flow optimizer, for this practice might well result in session flow targets that do not exceed channel capacity when considered singly, yet exceed this same shared capacity when summed together.

[00133] A solution to this problem involves two distinct steps:

1. The sessions bundle is treated as a virtual session by the server optimizer, subject to the minimum flow constraint obtained by summation of the individual session minima, and a maximum constraint that is the least of (a) the sum of the least of constraints (3) and (4) for each session, and (b) the shared

channel capacity client.lookup(k).channel.maxFlowRate (f_k^{cap}). The flow optimizer then generates an aggregate flow target for all channels virtual sessions by interpolation, using:

(25) $Cost(t) = f_1(t)+..+f_n(t)$

[00134] The cost vector in this instance is the n-dimensional unit vector, where n is the number of active clients.

(26) $c=(1, 1,....,1)$

[00135] Now it is possible to have an achievable aggregate flow target for the session bundle k that we must apportion optimally among the constituent sessions. Fortunately for us, our interpolation procedure of Figure 15 can be applied again to the sessions in our bundle, but where the aggregate flow target generated above replaces the aggregate flow constraint f_{svr}^{max}

Specifically:

For every flow f_{ki}, within session bundle k, the optimal flow rate is obtained by interpolation between minimum and maximum session flows:

(27) $f_{ki}^{opt} = f_{ki}^{min} + \alpha * (f_{ki}^{max} - f_{ki}^{min})$

(28) $\alpha = [(f_K^{bundle} - f_{k1}^{min} + ...f_{kn}^{min})]/[(f_{k1}^{max} - f_{k2}^{min}) + + (F_{kn}^{max} - f_{kn}^{min})]$

and the value of f_k^{bundle} is obtained from a higher level optimization in which session bundle k is treated as a virtual session for which:

(29) $f_k^{min} = f_{k1}^{min} + ... + f_{kn}^{min}$

(30) $f_k^{max} = \min(f_{k1}^{max} + .. + f_{kn}^{max}, f_k^{cap})$

It must be noted that session bundles are typically dynamic in nature, outliving many if not all of their constituent sessions. Before a session is added to the bundle two CAC decisions must be reached:

(1) establish bandwidth availability within the bundle; and

(2) bandwidth availability within the server.

Relationship to Adaptive Rate Control

[00136] The flow optimization methods should be viewed as constituting the lowest and perhaps only the first of many supervisory control layers to sit astride all streaming sessions, which comprise their own hierarchical stack of control functions; rate adaptive control sitting above a flow rate regulator which relies on a transport entity.

[00137] Channel capacity is known to vary widely over time. A capacity drop to below the minimum reserve flow rate is problematic as it causes the latter to rise in violation of its monotonicity assumption. In a server with available bandwidth to spare such a rise can be accommodated by the flow optimizer (by bumping the session's minimum flow reservation to the higher value) provided the flow deficit can be made up in the near term by draining pre-delivered content from the media buffer. Should either of these resources approach depletion, stream degradation to a lesser bit-rate or frame rate may be the only viable option short of a pause. Whether up or down, any rate change is a change to *fe*, which must be preceded by change to session's reserve flow rate allocation by the optimizer.

[00138] As can be seen, the optimizer and rate adapter must be integrated with care. One possible approach would be to modify a rate adapter such that it renegotiates its proposed *fe* value with the flow optimizer prior to effecting the rate change it deems appropriate based on information provided by the flow regulator. Rate increases might well be subject to deferral under heavy load conditions, awaiting bandwidth availability. The opposite is true of rate flow decreases: these would be forestalled as long as possible by the optimizer on a system that is not fully loaded, as previously discussed.

Live Events

[00139] Referring to Figure 17, the method of the present application can be extended to the optimization of "live", time-shifted "near-live" events, and cascaded delivery, wherein one server delivers content to another while the latter accepts clients for the same content. The extension assumes the following characteristics of the server and the live performance:

1. The performance is of a fixed duration T known in advance, encoded at an average rate *fe*, with a projected content payload of *fe*T*;

2. The origin server at the point of injection delays the stream by a small dwell time T^{DWELL} on the order of seconds (e.g. 10). Notwithstanding, the delayed content is available for early delivery to mirrored servers and players with the aim of providing these with a modicum of isolation from network jitter affecting the timely delivery of the live video frames.

3. Throughout the live event, captured content is stored in a file, thereby permitting late joiners to view the "live event" starting from any point between the present and the event start time;

4. At the conclusion of the live event, the encoded content, now fully captured to a file, is available for viewing over the balance of the admission window (e.g., 60 minutes in Figure 17).

[00140] The optimisation of live events and video on demand (VOD) differ principally in one respect, namely in the formulation of the server content underflow constraint.

For VOD, this constraint is

(31)

$L_k^{TOGO}(t)$ is initialised to the content size (*fe*T*) at that start of play, and decremented as content is delivered to the client. Δt is the time interval between successive periodic optimizations (e.g. 1 second). In practice this constraint rides

well above other operative maximum flow constraints (e.g. buffer overflow, channel capacity) until the very last calculation of the session flow target by the optimiser.

[00141] For a live stream, this constraint must additionally reflect the fact that content is streaming into the server (where it builds up), even as the server streams it out toward clients, and that the latter activity cannot possibly overtake the former.

[00142] Referring to Figure 17, (which illustrates how optimised live streams behave) line L1 depicts the build-up of freshly captured content to the server. Line L2, which is delayed (time-shifted) relative to L1 by the dwell time T^{DWELL}, represents the live performance as cast and viewable by clients. One such client is admitted at time A and joins the live session at the live moment B. The dashed curved line linking points B and C shows a plausible build-up of content in the client's media buffer in advance of play, as the viewer continues its advance along line L2. Past point C the top of the media buffer tracks the live capture stream and thenceforth advances along line L1. In this instance, the time shift between the live performance as cast (L2) and the live performance as captured (L1) imposes a maximum on the amount of content that could ever accumulate in a client media buffer, irrespective of its actual capacity.

[00143] Another client, admitted at time $t=10$, is interested in viewing the live stream from the beginning without missing anything. As before, the dashed line joining 10 and D depicts a plausible build-up of content in the client's media buffer in advance of play, as the viewer continues its advance along line L3, which is time-shifted by $(10+T^{DWELL})$ minutes relative to L1. The greater time shift (now between L1 and L3) permits more extensive pre-buffering of "near live" content.

[00144] The present constraints do not impose suitable flow limits at and beyond points C and D when the client delivery streams finally links up with the live capture stream, in the live moment. Consequently the optimiser may propose flow targets that cannot be realized.

[00145] Accordingly, the following modification for use in connection with live or cascaded streams is made:

(32) $f_k(t) <= [(L_k^{TOGO}(t) - Live^{TOGO}(t)] / \Delta t$

for a live event delivered at flow rate *fe*, we have

(33) $Live^{TOGO}(t) = \max(fe*(T - T^{DWELL} - t), 0)$

[00146] $Live^{TOGO}(t)$ is the span between projected content size *fe* **T* (level line L0 in Figure 17) and the content capture level line L1, whereas $L_k^{TOGO}(t)$ is the gap between L0 (i.e., *fe* * *T*) and the dashed curved line representing the content level delivered to the media buffer. The latter must always exceed the former. As the two approach one another this constraint takes effect, causing $L_k^{TOGO}(t)$ to always exceed $Live^{TOGO}(t)$ as it decreases at rate *fe*.

[00147] Viewers joining late and willing to see the live performance with some delay relative to the live case enjoy higher potential levels of isolation from network disturbances, and thus QOS, by virtue of their greater effective buffer capacity. By the same token, such viewers afford the optimizer greater latitude to optimize the flow of VOD and "live" content across load fluctuations, thereby enhancing server capacity. Viewers that wish to experience the live event in "real-time" nevertheless benefit by a modestly enhanced QOS conferred by a dwell time's worth of content in their buffers.

Summary

[00148] A method and system for call/connection acceptance and flow modulation for network delivery of video/audio programming is thus provided. Although several embodiments have been illustrated and described, it will be apparent to those skilled in the art that various changes and modifications may be made without departing from the spirit of the invention or the scope of the claims.

CLAIMS:

1. A method of bandwidth allocation for delivery of stored digital content from at least one server device to at least one client device by way of a network, the method comprising the steps of:

a) prescribing a control variable which represents a target flow rate from the server device to each client device;

b) determining time-varying constraints on the flow rate of the content;

c) determining a cost function of the control variables for all clients; and

d) prescribing bandwidth to all clients based upon the value of the control variables that maximize the cost function.

2. The method of Claim 1 wherein step (d) comprises performing periodic computations to update the value of the control variable such that the bandwidth can be continuously allocated to each client.

3. The method of Claim 2 wherein a new client is accepted by:

i) determining an admission capacity of the bandwidth;

ii) admitting a prospective client if the clients minimum allowed value of the control variable is less than the admission capacity; and

iii) wherein a client admitted for service is guaranteed to have sufficient content flow over the entire session.

4. The method of Claim 3 wherein:

The admission capacity equals a server swing capacity reduced by a predetermined safety margin which may be zero, the swing capacity equaling the difference between the server flow capacity and the sum of the minimum allowed flow rate for all clients.

5. The method of Claim 4 wherein the minimum allowed value of the control variable is initialized to the average consumption rate.

6. The method of Claim 4 wherein the minimum allowed value of the control variable is initialized to the average consumption rate multiplied by a factor greater than one.

7. The method of Claim 4 wherein the server flow capacity varies in a step-wise function over time according to a predetermined schedule and the method of accepting a new client further comprises:

i) determining a sequence of future step changes of server flow capacity;

ii) determining at each time in the sequence of future step changes the value for a future worst swing case capacity, the worst case swing capacity being obtained from the server flow capacity at the time by subtracting an extrapolated present minimum allowed flow rate for all active clients that are potentially active in the future;

iii) admitting a prospective client if a minimum allowed value of the control variable is less than the capacity obtained from the sequence and the present capacity;

iv) whereby a new client can be admitted without compromising the service to existing clients in the presence of previously scheduled server flow capacity changes and all clients accumulate content in their buffers to the extent of their respective buffer capacity.

8. The method according to Claim 7 wherein the method of accepting a new client further comprises:

i) partitioning the available bandwidth into partitions representing a prescribed service charge;

ii) determining a maximum-minimum capacity for a service charge to be the maximum allowed sum of all minimum allowed flow rates for all clients incurring an equal or lesser service charge, the maximum-minimum capacity equaling the sum of bandwidth allocations for all service charges less than or equal to the service charge;

iii) determinating a service-charge-specific swing capacity for a given service charge as the difference between the maximum-minimum capacity for the service charge and the sum of all minimum allowed flow rates for all clients incurring the service charge or a lesser service charge;

iv) determining the server swing capacity to be the server flow capacity minus a margin and minus the sum of the minimum allowed flow rate for all clients; and

v) determining the service-charge-specific capacity to be the least of the server swing capacity and the charge specific swing capacity.

9. The method of Claim 8 wherein the method of accepting a new client further comprises:

i) determining the sequence of future times of step changes to server flow capacity over a proposed consumption period;

ii) determining for each time in the sequence of future times, the future worst-case value for the service-charge-specific capacity, the worst case value being obtained by using the extrapolated present minimum allowed flow rate of all presently active clients that are potentially active at the time in the future; and

iii) admitting a client incurring a given service charge if the minimum flow rate of the client is less than the capacity obtained from the sequence and the present service-charge-specific capacity.

10. The method of Claim 3 wherein step (b) comprises determining time-varying constraints wherein:

i) an aggregate flow rate for all clients does not exceed a predetermined server flow capacity;

ii) the flow rate from the server to the client does not exceed a maximum allowed flow rate for the client;

iii) the flow rate to the client will never cause a buffer of the client to overflow;

iv) the flow rate to the client stops when the content is exhausted; and

v) the flow rate from the sever will never be less than the client's minimum allowed flow rate which may vary over time and may be zero.

11. The method of Claim 10 wherein the maximum allowed flow rate to the client is given by the minimum of one selected from the following:

i) a client flow rate ceiling;

ii) the flow rate required to fill the client buffer before the next periodic computation; and

iii) the flow rate required to complete the delivery of the content before the next periodic computation.

12. The method of Claim 11 wherein:

i) the minimum allowed client flow rate cannot increase over time, but may decrease such that the initial value of the control rate is no less than the average consumption rate;

ii) the control variable for the client is always greater than or equal to the current minimum allowed client flow rate; and

iii) the client buffer will never underflow until the content is fully consumed at the average consumption rate.

13. The method of Claim 12 wherein the minimum client flow rate is the product of the constant content flow rate from the client to the server that causes the last piece of content to be delivered at the last possible moment as the client buffer is being drained at the average consumption rate, with a factor greater or equal to one.

14. The method of Claim 13 wherein the minimum client flow rate is initialized to the average consumption rate.

15. The method of Claim 13 wherein the minimum client flow rate is initialized to the average consumption rate multiplied by a factor greater than one.

16. The method of Claim 13 further comprising the step of calculating the minimum client flow rate periodically.

17. The method of Claim 1 wherein the cost function represents an aggregate flow rate and is the sum of the flow rates for all clients.

18. The method of Claim 1 wherein the cost function represents a service charge and is the sum of the client flow rates multiplied by the client's cost of service.

19. The method of Claim 18 wherein:

step (a) comprises:

i) determining a maximum allowed flow rate and a minimum allowed flow rate for each client;

ii) determining a flow rate range for each client as the difference between the maximum allowed flow rate and the minimum allowed flow rate; and

iii) initializing a current flow rate for each client as the minimum allowed flow rate and summing the flow rate into the total server flow rate; and

step (d) comprises:

i) computing remaining server bandwidth as the difference between the maximum server flow capacity and the total server flow rate; and

ii) allocating remaining server bandwidth to remaining clients wherein all clients with a given cost of service receive bandwidth equally to the extent of their respective ranges, and no client with a given cost of service being allocated bandwidth unless all clients with a higher cost of service have received their maximum possible allocation of bandwidth.

20. The method of Claim 17 wherein step (a) prescribing the control variable comprises:

i) forming a hyperplane of control variables that aggregate to the maximum allowed aggregate flow rate;

ii) computing a first vector of minimum allowed client flow rates;

iii) computing a second vector of maximum allowed client flow rates;

iv) determining a multi-dimensional shape of which an interior diagonal spans the first and second vectors; and

v) finding a solution of control variables at the point of intersection between the diagonal and the hyperplane.

21. The method of Claim 20 wherein step (v) comprises finding a solution of control variables at the second vector if there is no point of intersection between the diagonal and the hyperplane.

22. The method of Claim 17 wherein step (a) comprises:

i) computing the sum of the minimum allowed flow rates to all active clients;

ii) computing the sum of the flow rate ranges for all active clients;

iii) computing the difference between the maximum allowed aggregate flow rate and the sum of the minimum;

iv) computing a factor as the ratio of the difference over the sum of the ranges;

v) if the factor exceeds 1, the control variable is set to the maximum allowed client flow rate; and

vi) if the factor is less than 1, computing the control variable by multiplying the client range by the factor and then adding the result to the minimum allowed client flow rate.

23. The method of Claim 17 wherein:

step (a) comprises:

i) determining the maximum allowed flow rate and minimum allowed flow rate for each client;

ii) determining the flow rate range for each client as the difference between the maximum flow rate and the minimum flow rate; and

iii) initializing the control variable for each client as said minimum allowed flow rate and summing the flow rate into a total server flow rate; and

step (d) comprises:

i) computing remaining server bandwidth as the difference between the maximum server flow capacity and the total server flow rate; and

ii) allocating remaining server bandwidth to remaining clients to the extent of their respective ranges.

24. The method of Claim 23 wherein the step of allocating remaining server bandwidth further comprises the clients receiving the bandwidth equally to the extent of their respective ranges.

25. The method of Claim 17 wherein the maximized value of the control variable can be determined by:

i) determining a maximimum and minimum flow rate for each client;

ii) determining a range between the maximimum and mininmum flow rates in order to find a flow rate range;

iii) determining available bandwidth by finding the difference between an aggregate flow capacity and a sum of the minimum flow rate;

iv) determining a flow variable by dividing the unused bandwidth by the flow rate range;

v) prescribing the control variable to be the minimum flow rate added to the flow rate range if the flow varible is less than one or prescribing the control variable to be the minimum flow rate added to the flow rate range corrected by the flow variable if the flow favorable is greater than one such that the control variable is easily calculated for each client.

26. The method of Claim 17 wherein at least one session is a session bundle and:

1) the minimum flow constraint is computed as the sum of the minimum flow rate constraints for all bundled sessions;

2) the maximum flow constraint is the least of 1) the sum of the maximum individual flow rate constraints of each session and 2) the flow capacity of the channel shared among the bundled sessions; and

3) the target flow rate is computed for the session is apportioned among the bundled constituent sessions.

27. The method of Claim 26 wherein the bundled session is an advertisement.

28. A system for allocating bandwidth between a server device and at least one client device, the system comprising:

a call acceptance module operative to receive an incoming request for service;

a flow regulator configured to deliver content at a modulated target flow rate, the content being delivered between the server device and a respective client device when a call is accepted by the call acceptance module; and

a flow optimizer configured to modulate the target flow rate of the flow regulator in order to optimize the aggregate flow of content.

29. The system of Claim 28 wherein the flow optimizer is configured to modulate the flow rate of the flow regulator in order to optimize charges for the content.

30. The system of Claim 29 wherein the flow regulator is configured to deliver the modulated flow rate of content in response to the control variable.

31. The system of Claim 28 wherein the flow optimizer is configured to determine a control variable which corresponds to an optimized flow rate.

32. The system of Claim 31 wherein the flow optimizer is configured to generate a cost function of the control variable that corresponds to a maximized value of the control variable.

33. The system of Claim 28 wherein the flow regulator is configured to deliver the modulated flow rate of content in response to the control variable.

34. The system of Claim 28 wherein the flow optimizer is configured to determine a maximized value of the control variable by:

i) determining a maximum and minimum flow rate for each client;

ii) determining a range between the maximum and minimum flow rates in order to find a flow rate range;

iii) determining unused bandwidth by finding the difference between an aggregate flow capacity and a sum of the minimum flow rate;

iv) determining a flow variable by dividing the unused bandwidth by the flow rate range;

v) prescribing the control variable to be the minimum flow rate added to the flow rate range if the flow variable is less than one or prescribing the control variable to be the minimum flow rate added to the flow rate range corrected by the flow variable if the flow favorable is greater than one such that the control variable is easily calculated for each client.

35. The system of Claim 33 wherein the call acceptance module is configured to accept a call based upon the value of the control variable.

36. A method of bandwidth allocation for delivery of stored digital content from at least one server device to at least one client device by way of a network, the method comprising the steps of:

a) prescribing a control variable which represents a target flow rate from the server to each client device based upon the amount of buffer of the client;

b) determining time-varying constraints on the flow rate of the content;

c) determining a cost function of the control variables for all clients in response to the size of the client's buffer; and

d) prescribing bandwidth to all clients based upon the value of the control variables that maximized the cost function.

37. A method for apportioning a given total flow rate (bandwidth) among a plurality of sessions for clients, the method comprising the steps of:

a) computing a sum of minimum flow rate constraints for each session;

b) computing a sum of flow rate ranges for all active clients;

c) computing a difference between the given total flow rate and a sum of the flow rate minima;

d) computing a factor as the ratio of the difference over the sum of the flow rate ranges;

e) if the factor exceeds 1, setting the flow rate to the maximum allowed client flow rate; and

f) if the factor is less than 1, computing the flow rate by multiplying the client range by the factor and then adding the result to the minimum allowed client flow rate.

38. The method of Claim 37 wherein the clients are distinguished by their cost of service and bandwidth is apportioned to clients iteratively, the amount apportioned to each client in a cost category is the amount remaining from prior allocations to clients in higher cost categories.

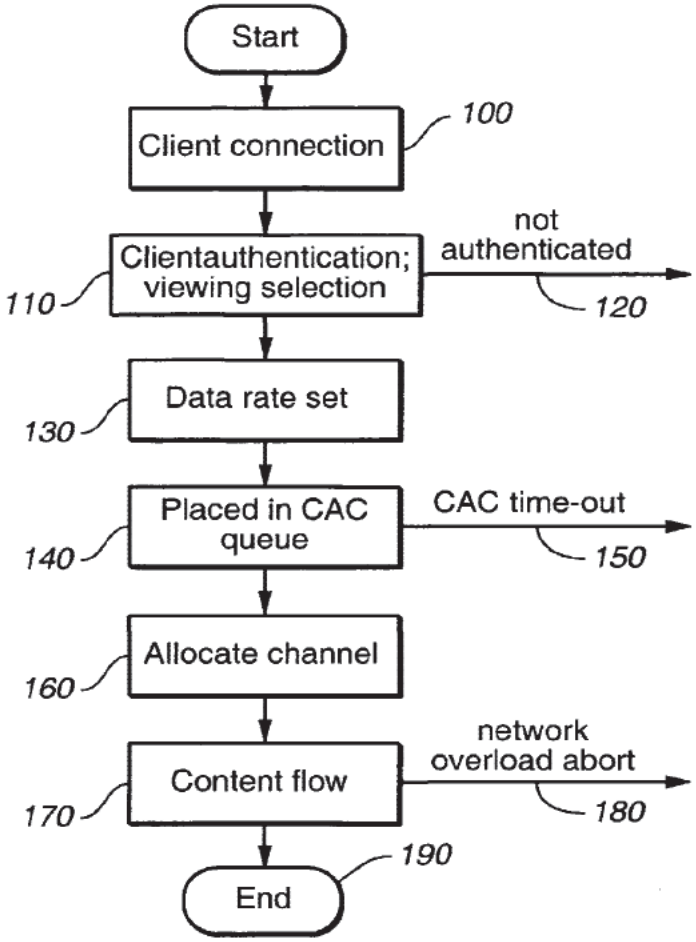

Start
Client connection
100
Clientauthentication; viewing selection
not authenticated
110
120
Data rate set
130
Placed in CAC queue
CAC time-out
140
150
Allocate channel
160
Content flow
network overload abort
170
180
End
190

FIG._1

Ex. 1018-1027

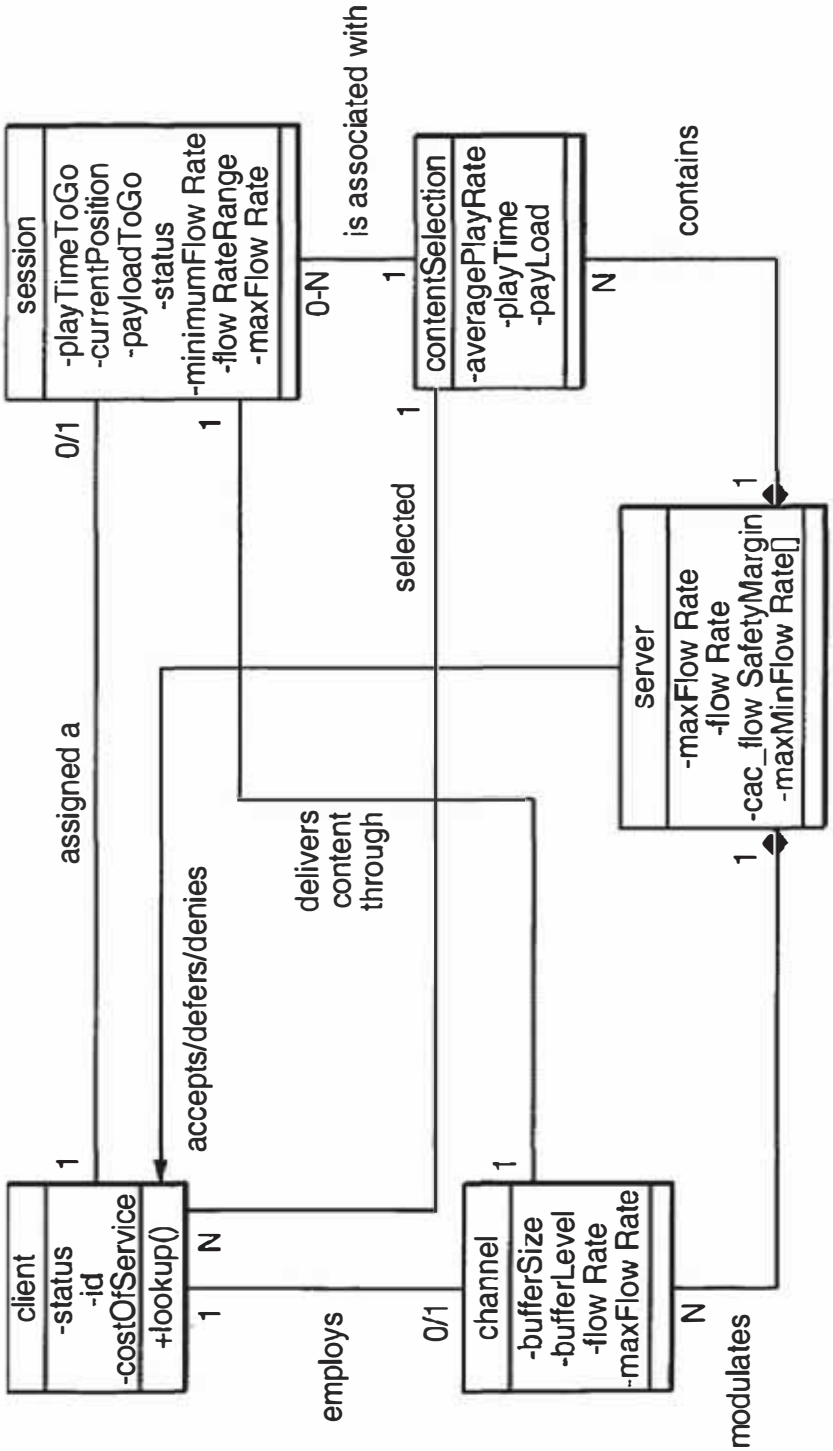


FIG._2

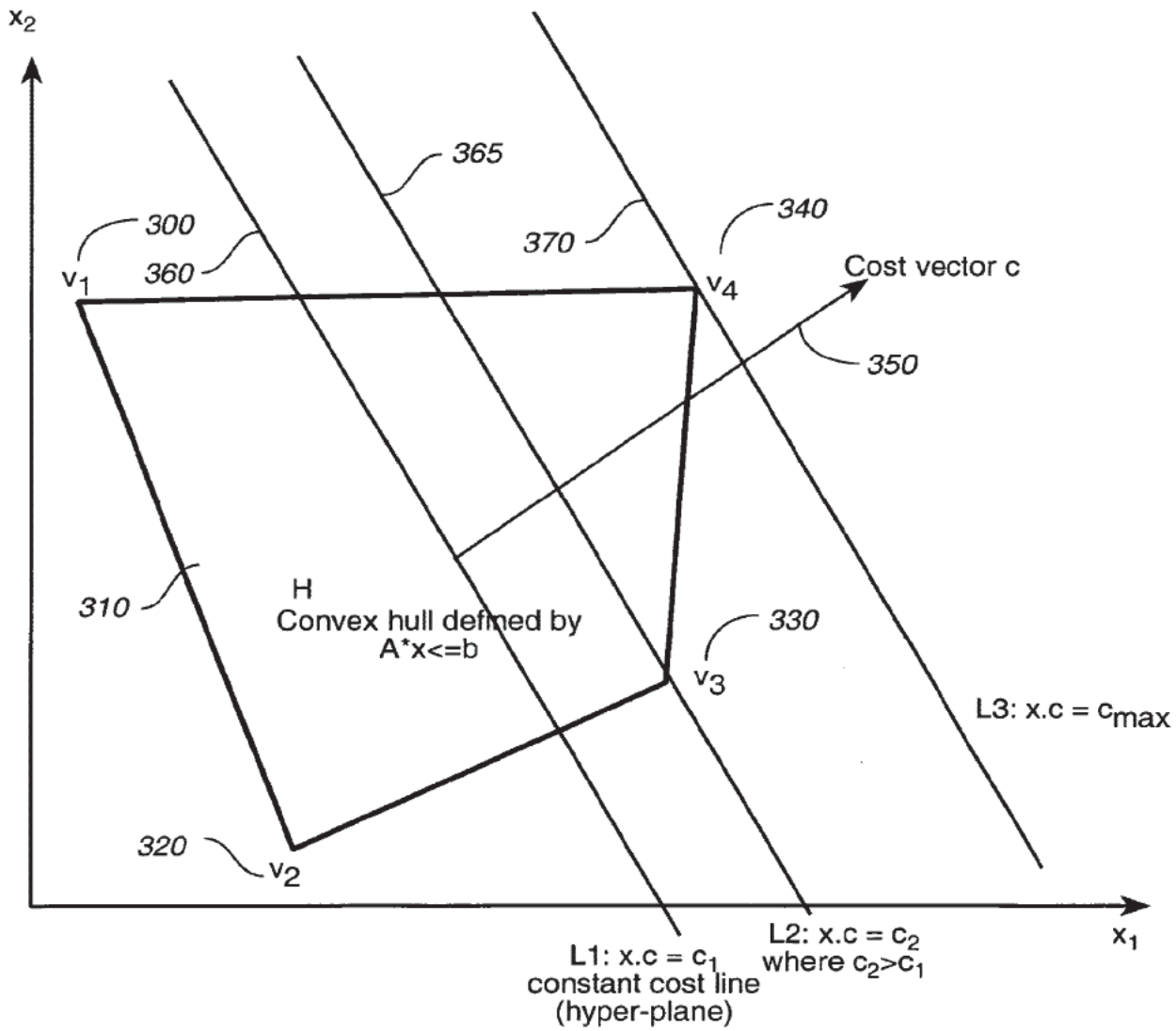

x2
x1
365
300
360
370
340
v1
v4
Cost vector c
350
310
H
Convex hull defined by
A*x<=b
330
v3
L3: x.c = cmax
320
v2
L1: x.c = c1
constant cost line
(hyper-plane)
L2: x.c = c2
where c2>c1

FIG._3

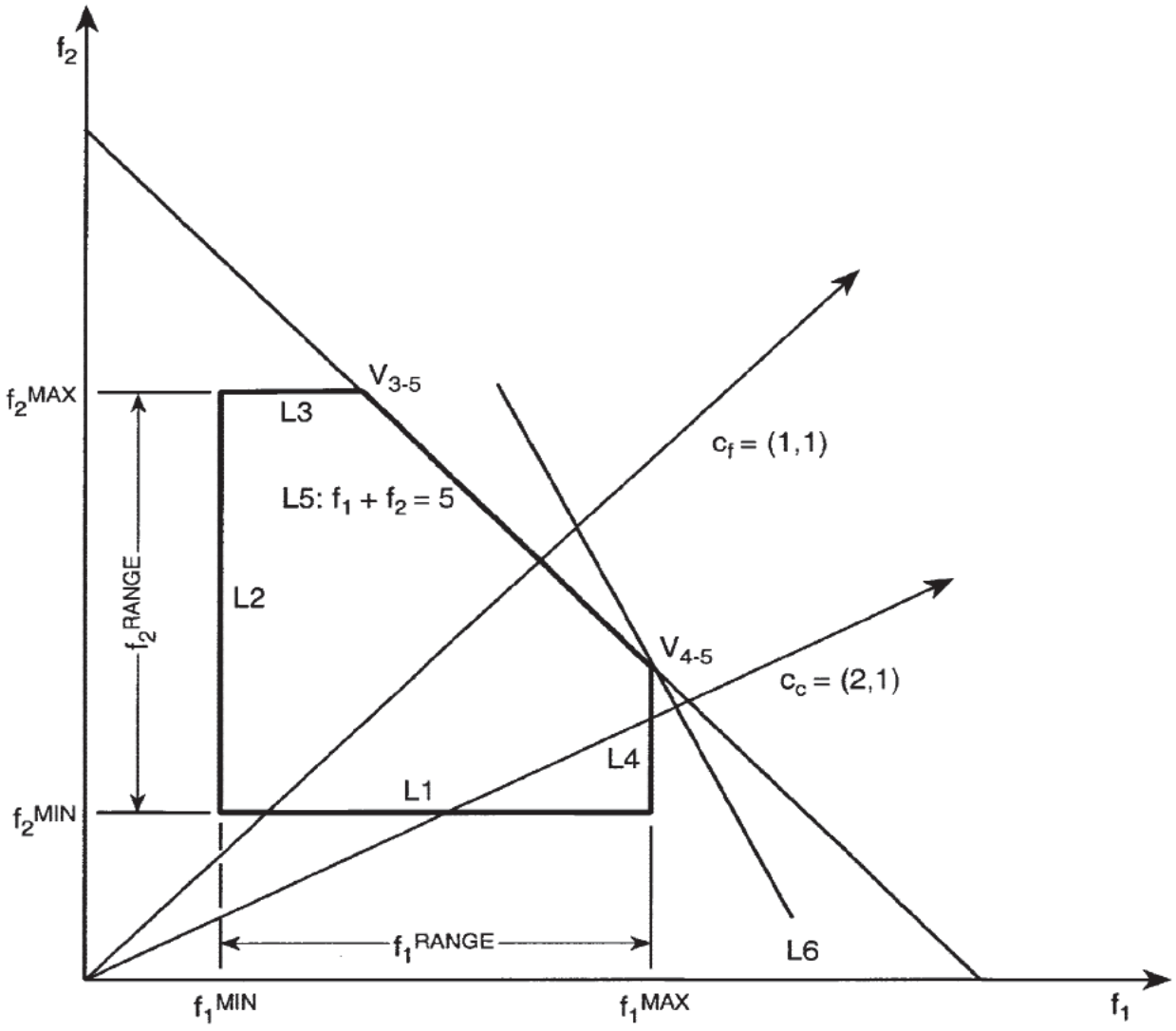

f2
f2MAX
f2MIN
f2RANGE
V3-5
L3
L5: f1 + f2 = 5
L2
V4-5
cf = (1,1)
cc = (2,1)
L4
L1
L6
f1RANGE
f1MIN
f1MAX
f1

FIG._4a

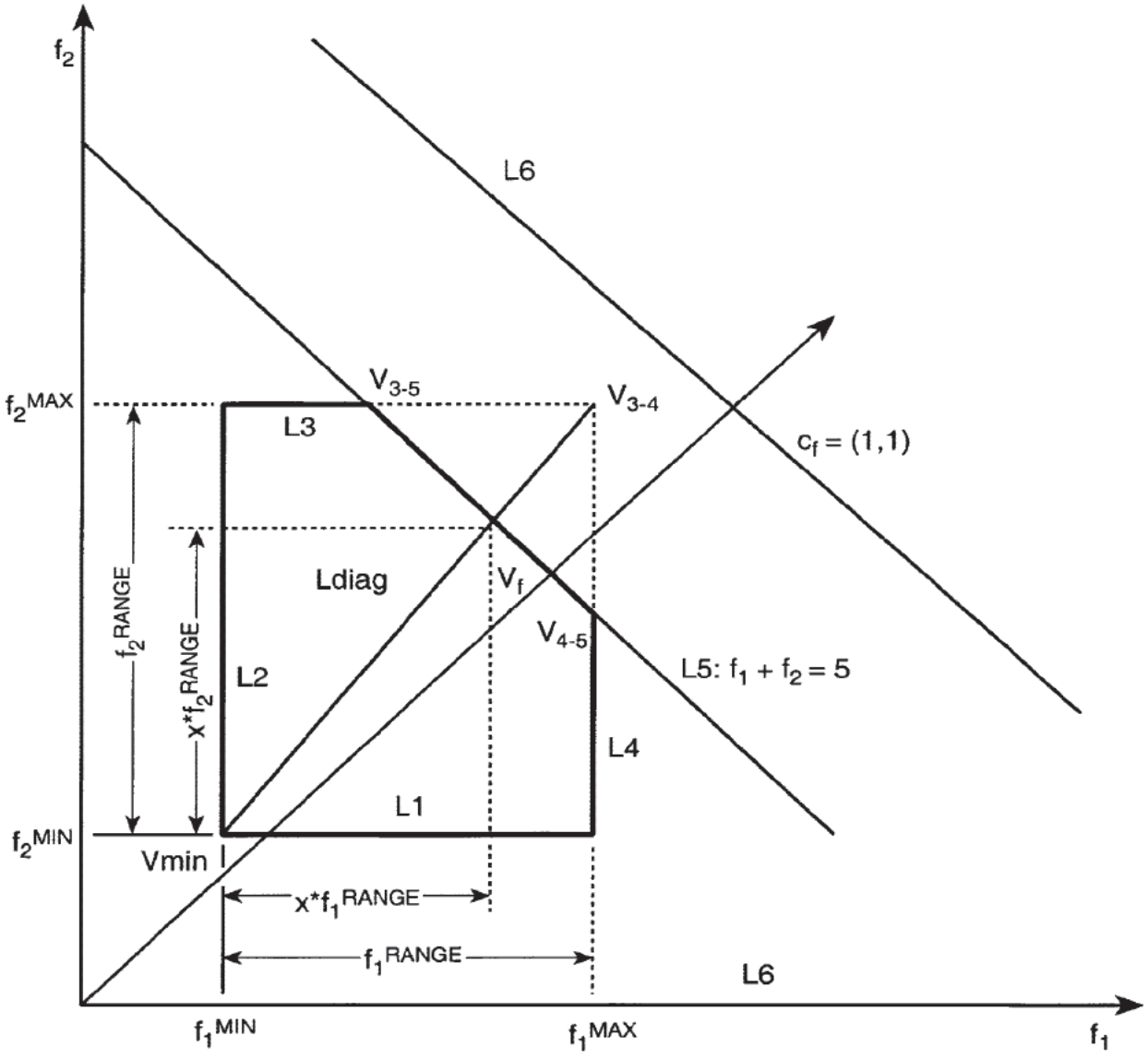

f2
L6
V3-5
V3-4
f2MAX
L3
cf = (1,1)
Ldiag
Vf
f2RANGE
x*f2RANGE
V4-5
L2
L5: f1 + f2 = 5
L4
L1
f2MIN
Vmin
x*f1RANGE
f1RANGE
L6
f1MIN
f1MAX
f1

FIG._4b

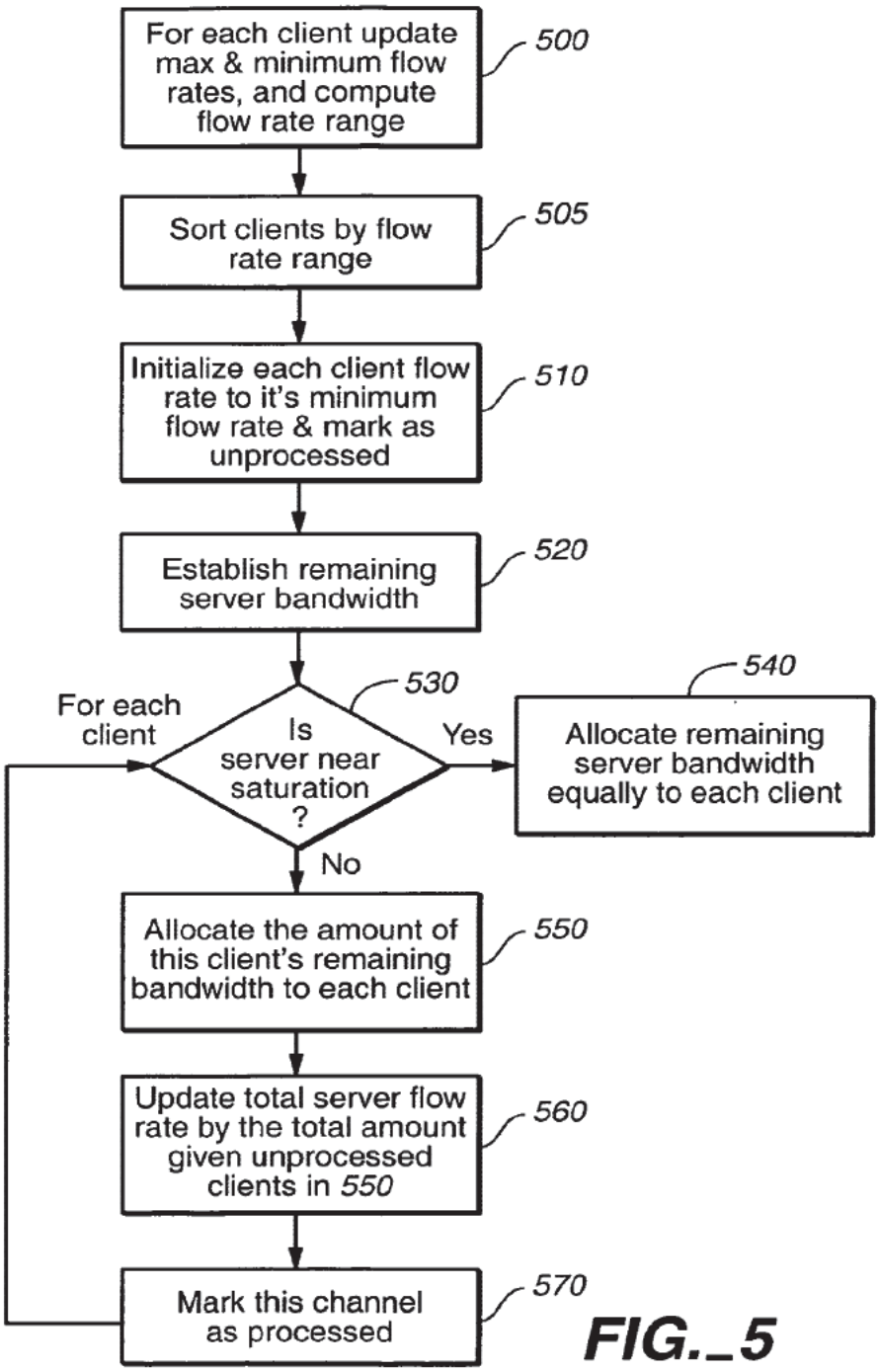

For each client update max & minimum flow rates, and compute flow rate range
500
Sort clients by flow rate range
505
Initialize each client flow rate to it's minimum flow rate & mark as unprocessed
510
Establish remaining server bandwidth
520
530
For each client
Is server near saturation ?
Yes
540
Allocate remaining server bandwidth equally to each client
No
Allocate the amount of this client's remaining bandwidth to each client
550
Update total server flow rate by the total amount given unprocessed clients in 550
560
Mark this channel as processed
570

FIG._5

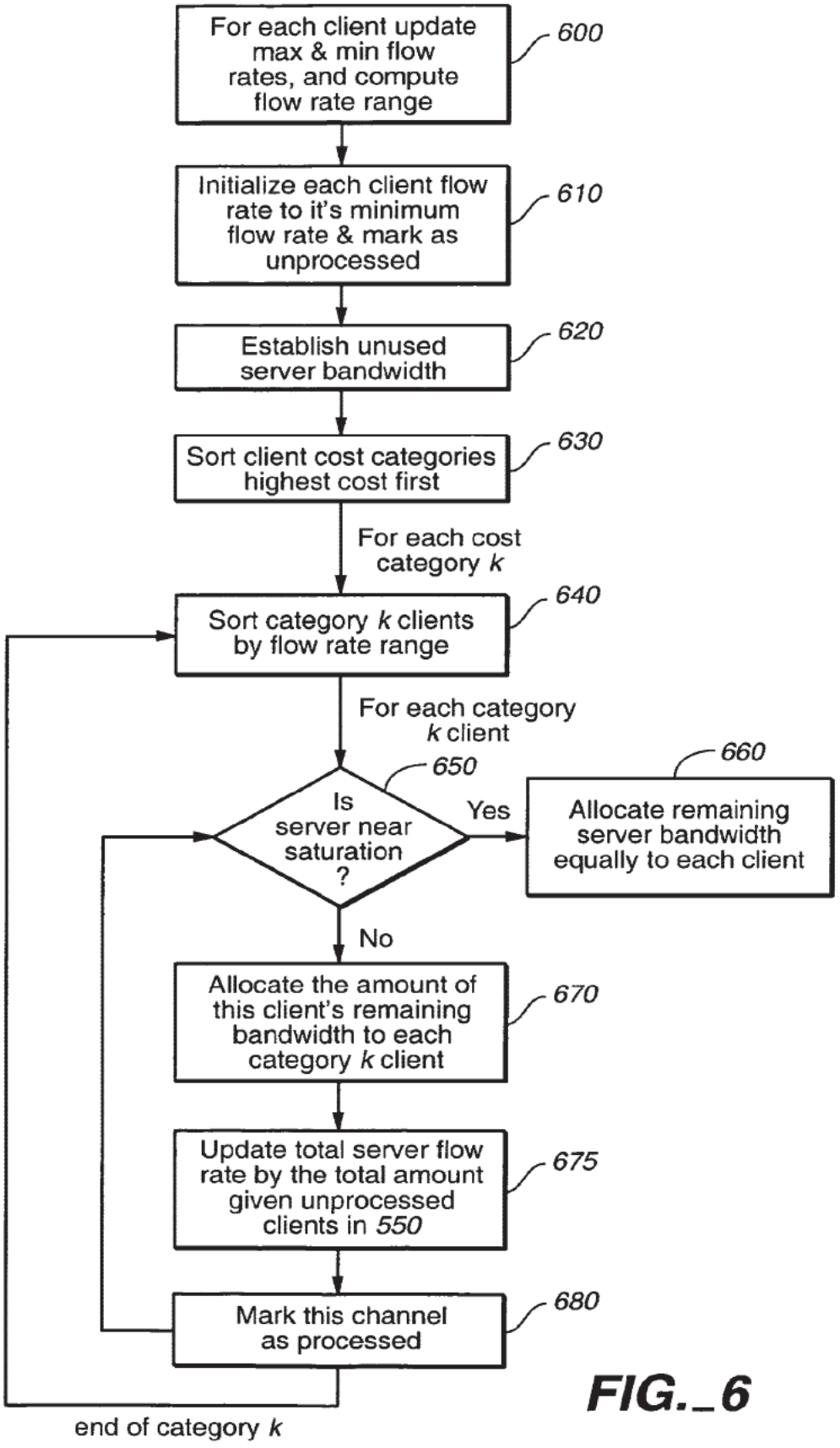

For each client update max & min flow rates, and compute flow rate range
600
Initialize each client flow rate to it's minimum flow rate & mark as unprocessed
610
Establish unused server bandwidth
620
Sort client cost categories highest cost first
630
For each cost category k
Sort category k clients by flow rate range
640
For each category k client
650
Is server near saturation ?
Yes
660
Allocate remaining server bandwidth equally to each client
No
Allocate the amount of this client's remaining bandwidth to each category k client
670
Update total server flow rate by the total amount given unprocessed clients in 550
675
Mark this channel as processed
680
end of category k

FIG._6

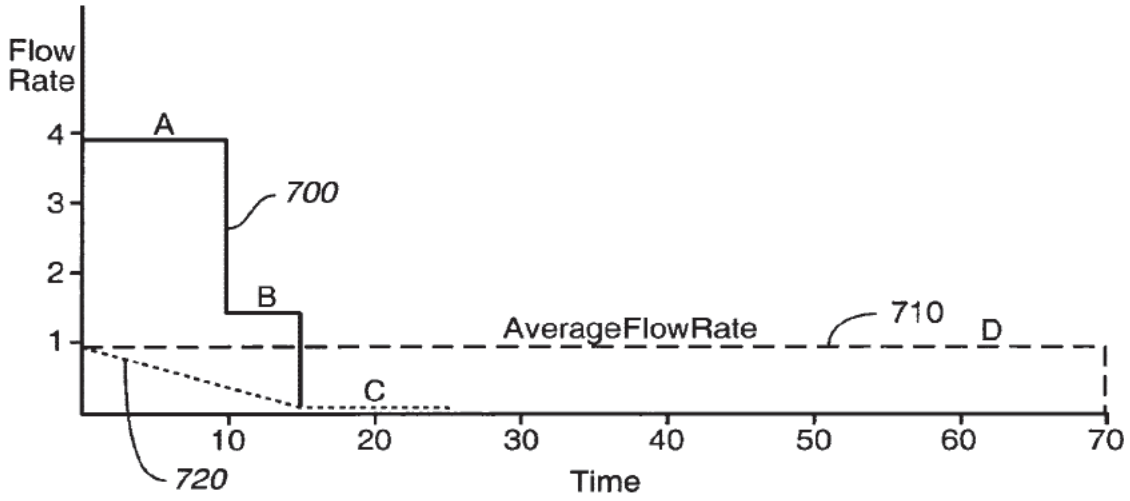

Flow Rate
4
3
2
1
A
700
B
C
AverageFlowRate
710
D
720
10
20
30
40
50
60
70
Time

FIG._7

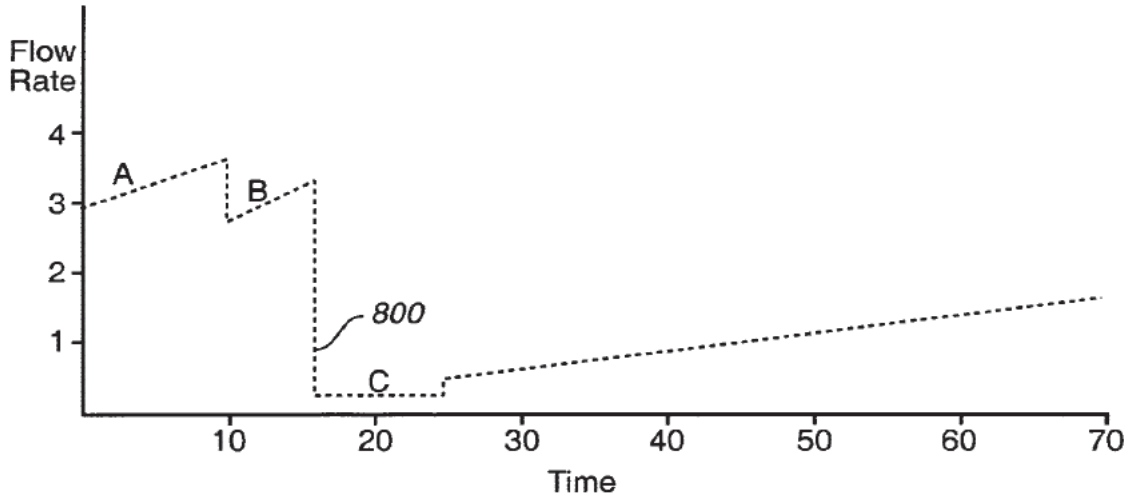

Flow Rate
4
3
2
1
A
B
800
C
10
20
30
40
50
60
70
Time

FIG._8

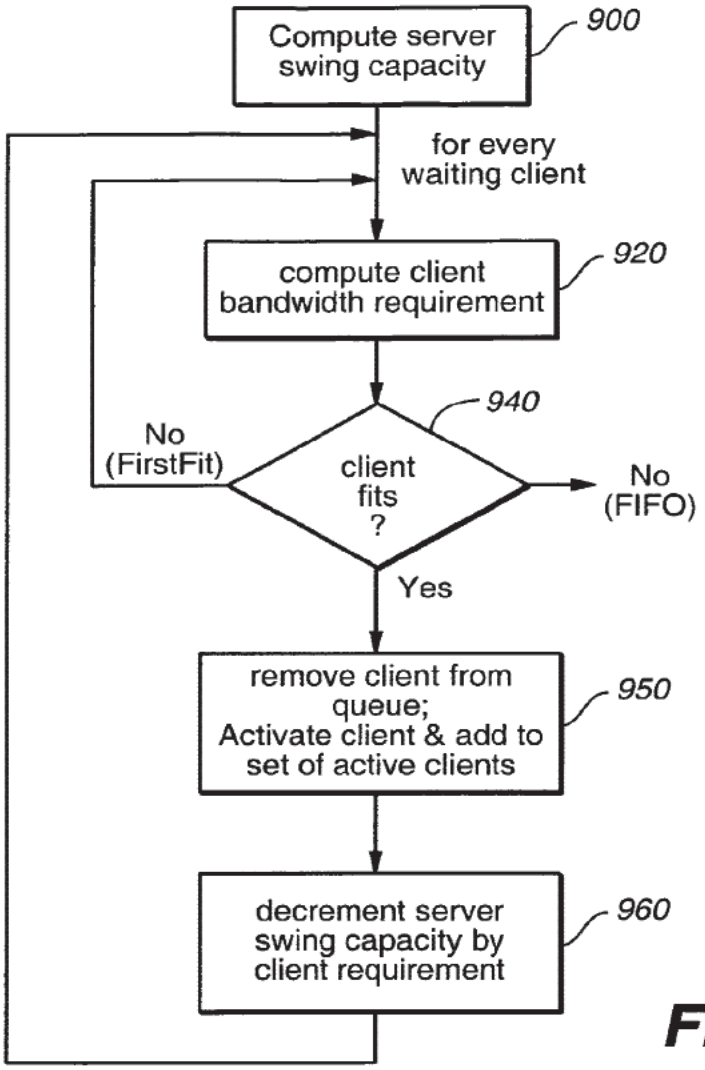

Compute server swing capacity
900
for every waiting client
compute client bandwidth requirement
920
940
client fits ?
No (FirstFit)
No (FIFO)
Yes
remove client from queue; Activate client & add to set of active clients
950
decrement server swing capacity by client requirement
960

FIG._9

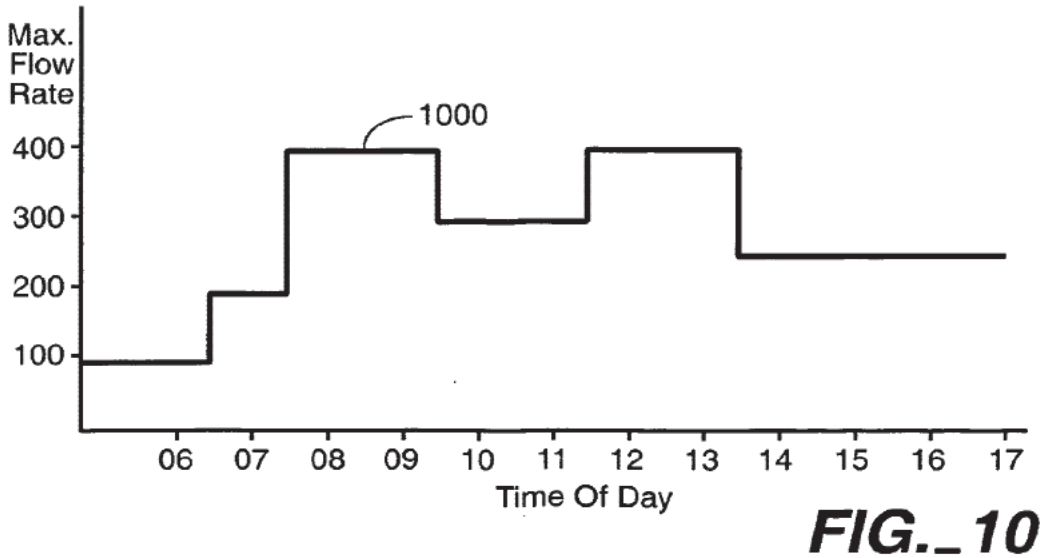

Max. Flow Rate
400
300
200
100
1000
06
07
08
09
10
11
12
13
14
15
16
17
Time Of Day

FIG._10

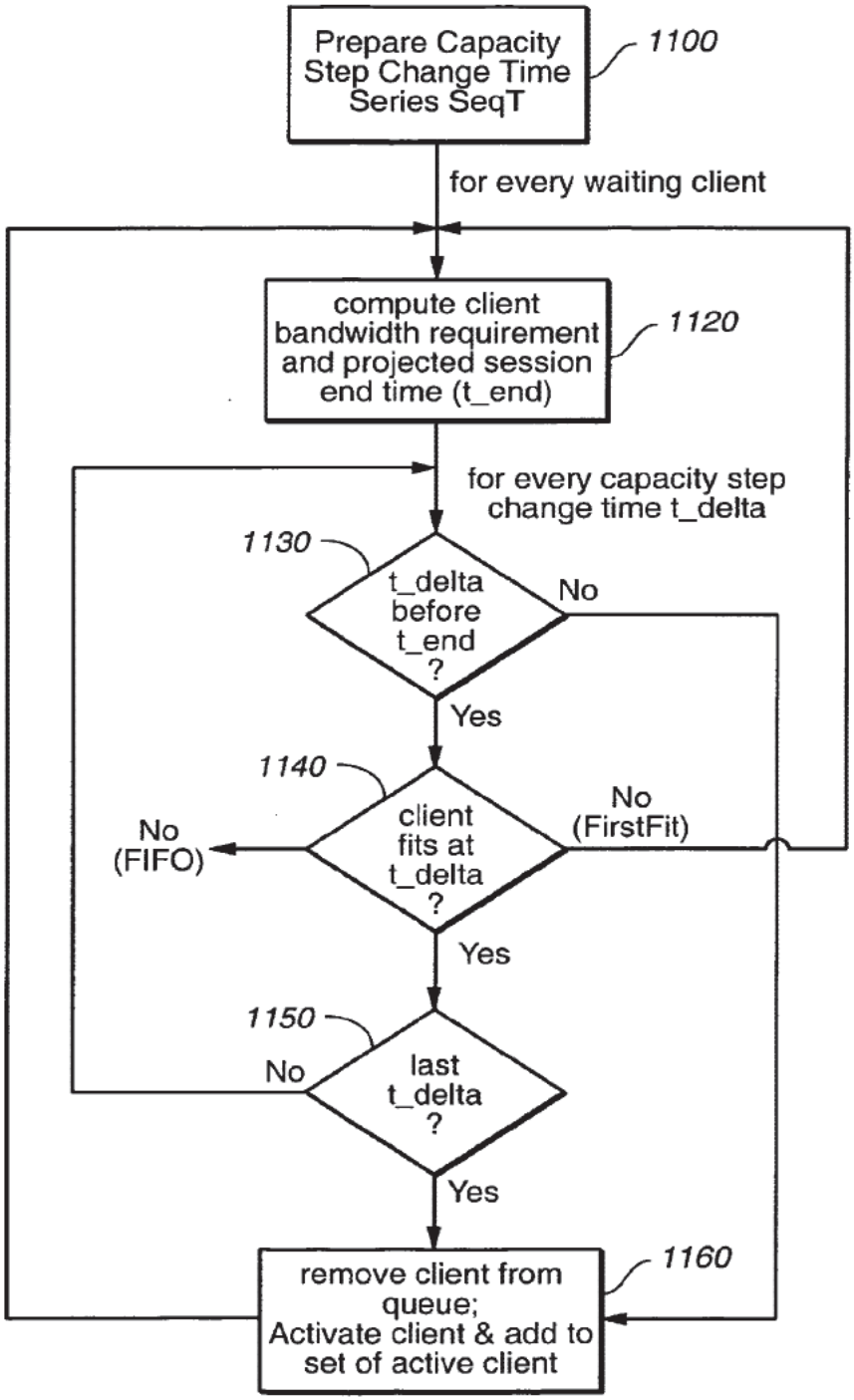

Prepare Capacity Step Change Time Series SeqT
1100
for every waiting client
compute client bandwidth requirement and projected session end time (t_end)
1120
for every capacity step change time t_delta
1130
t_delta before t_end ?
No
Yes
1140
client fits at t_delta ?
No (FIFO)
No (FirstFit)
Yes
1150
last t_delta ?
No
Yes
remove client from queue; Activate client & add to set of active client
1160

FIG._11

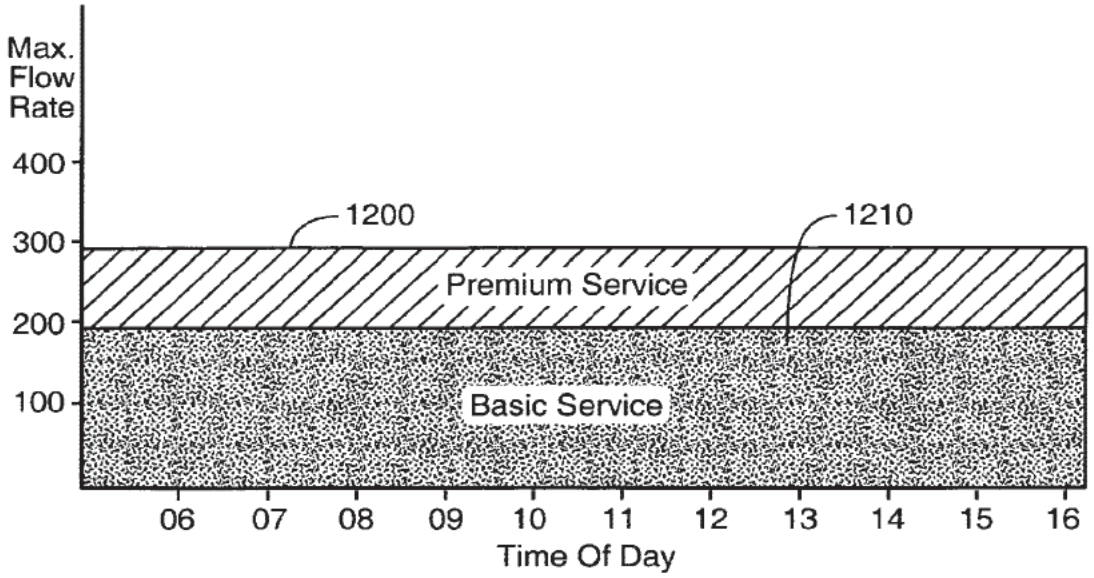

Max.
Flow
Rate
400
300
200
100
1200
1210
Premium Service
Basic Service
06
07
08
09
10
11
12
13
14
15
16
Time Of Day

FIG._12

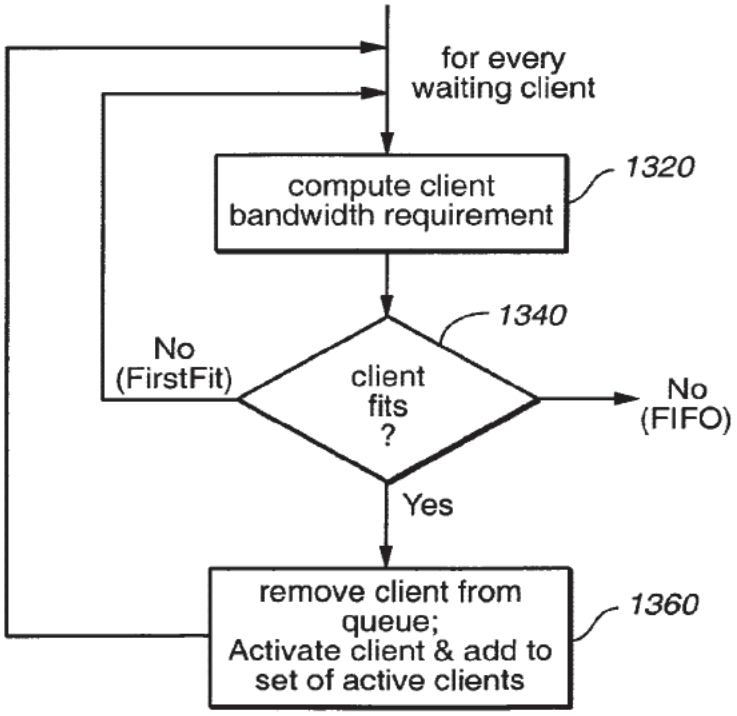

for every
waiting client
compute client
bandwidth requirement
1320
1340
No
(FirstFit)
client
fits
?
No
(FIFO)
Yes
remove client from
queue;
Activate client & add to
set of active clients
1360

FIG._13

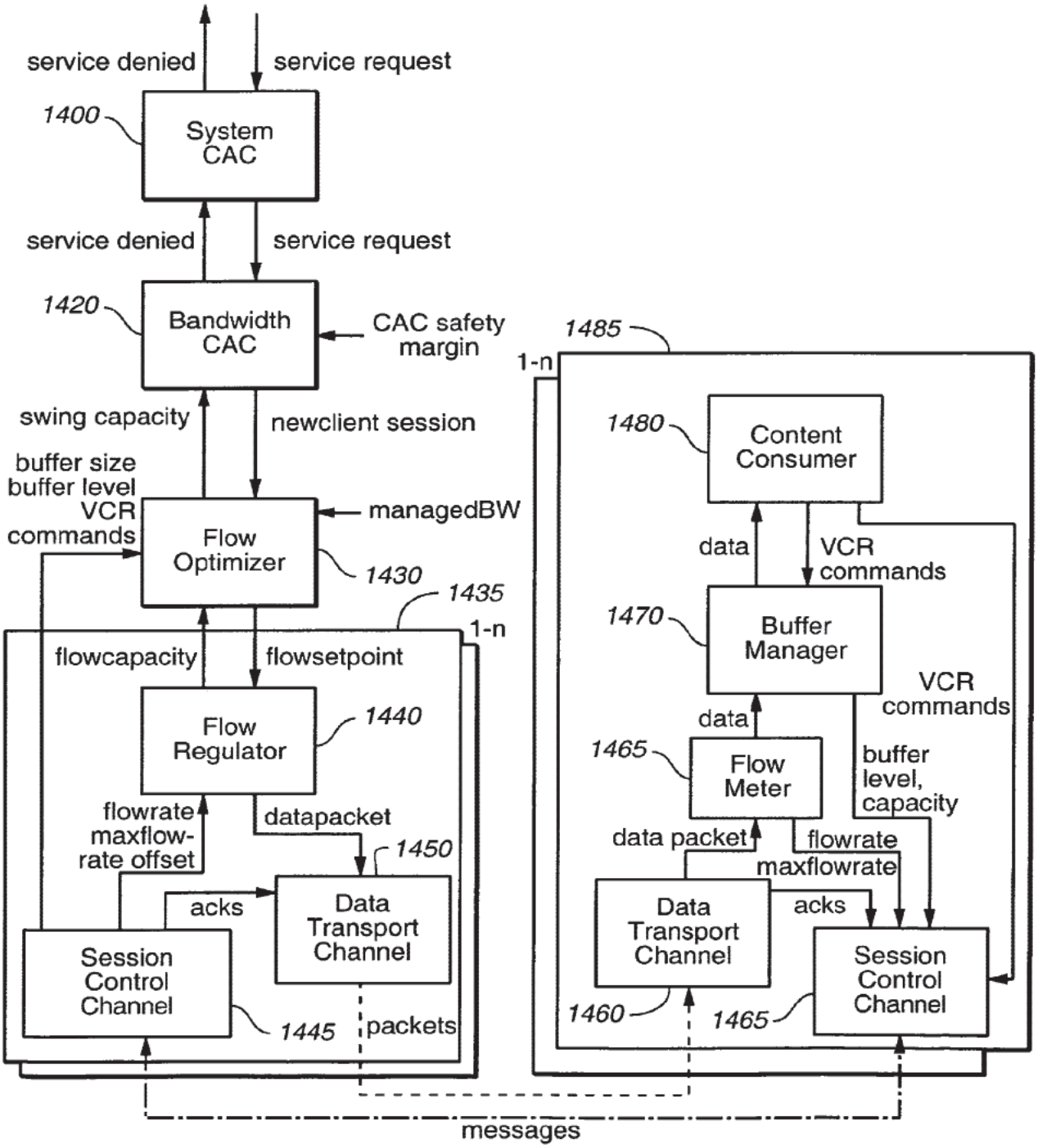

service denied
service request
1400
System CAC
service denied
service request
1420
Bandwidth CAC
CAC safety margin
swing capacity
newclient session
buffer size
buffer level
VCR commands
Flow Optimizer
managedBW
1430
1435
1-n
flowcapacity
flowsetpoint
Flow Regulator
1440
flowrate maxflow-rate offset
datapacket
1450
acks
Data Transport Channel
Session Control Channel
1445
packets
1485
1-n
1480
Content Consumer
data
VCR commands
1470
Buffer Manager
VCR commands
data
1465
Flow Meter
buffer level, capacity
data packet
flowrate maxflowrate
Data Transport Channel
acks
Session Control Channel
1460
1465
messages

FIG._14

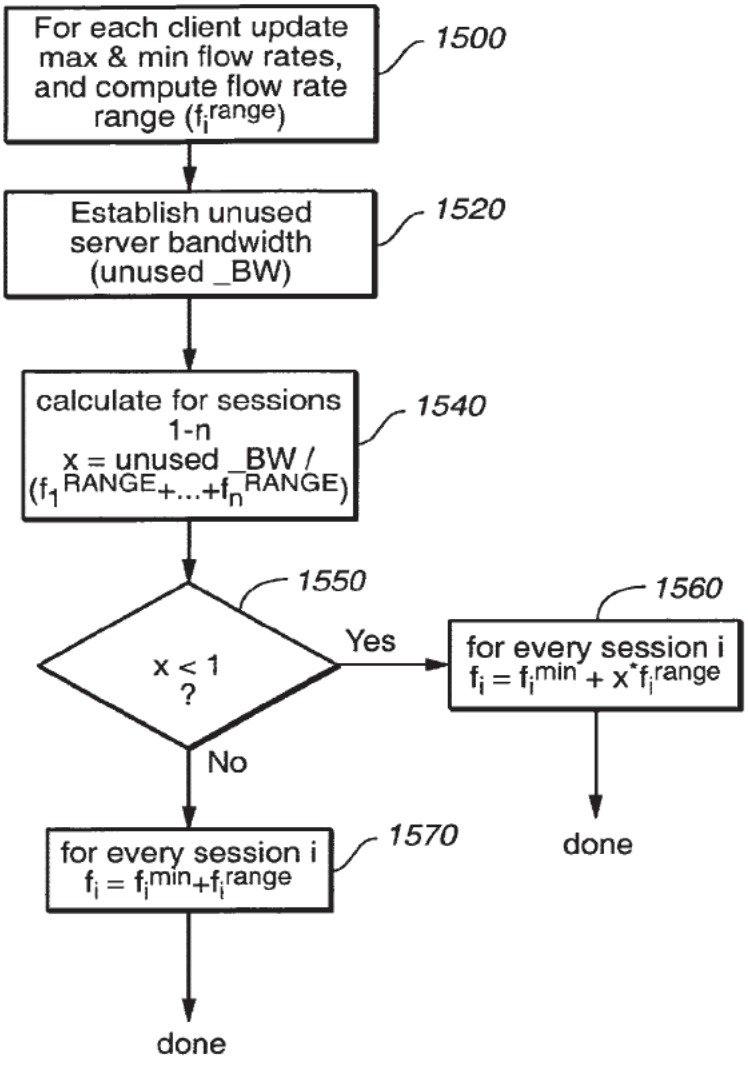

For each client update max & min flow rates, and compute flow rate range (f_i^{range})
1500
Establish unused server bandwidth (unused _BW)
1520
calculate for sessions 1-n
x = unused _BW / (f_1^{RANGE}+...+f_n^{RANGE})
1540
1550
x < 1 ?
Yes
1560
for every session i
$f_i = f_i^{min} + x^* f_i^{range}$
done
No
for every session i
$f_i = f_i^{min} + f_i^{range}$
1570
done

FIG._15

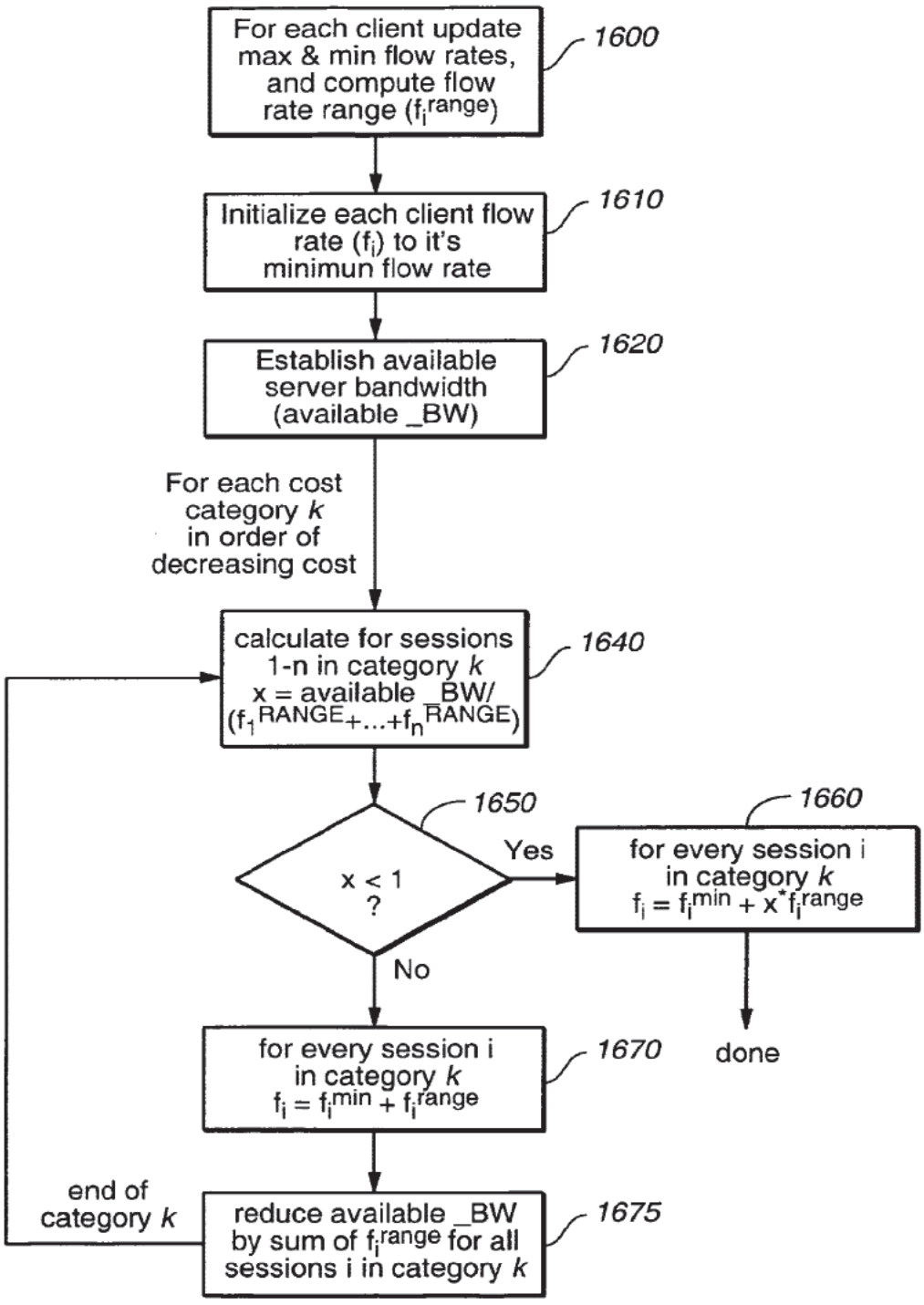

For each client update max & min flow rates, and compute flow rate range (f_i^{range})
1600
Initialize each client flow rate (f_i) to it's minimun flow rate
1610
Establish available server bandwidth (available _BW)
1620
For each cost category *k* in order of decreasing cost
calculate for sessions 1-n in category *k* x = available _BW/ (f_1^{RANGE}+...+f_n^{RANGE})
1640
1650
x < 1 ?
Yes
1660
for every session i in category *k* $f_i = f_i^{min} + x^* f_i^{range}$
done
No
for every session i in category *k* $f_i = f_i^{min} + f_i^{range}$
1670
end of category *k*
reduce available _BW by sum of f_i^{range} for all sessions i in category *k*
1675

FIG._16

Ex. 1018-1040

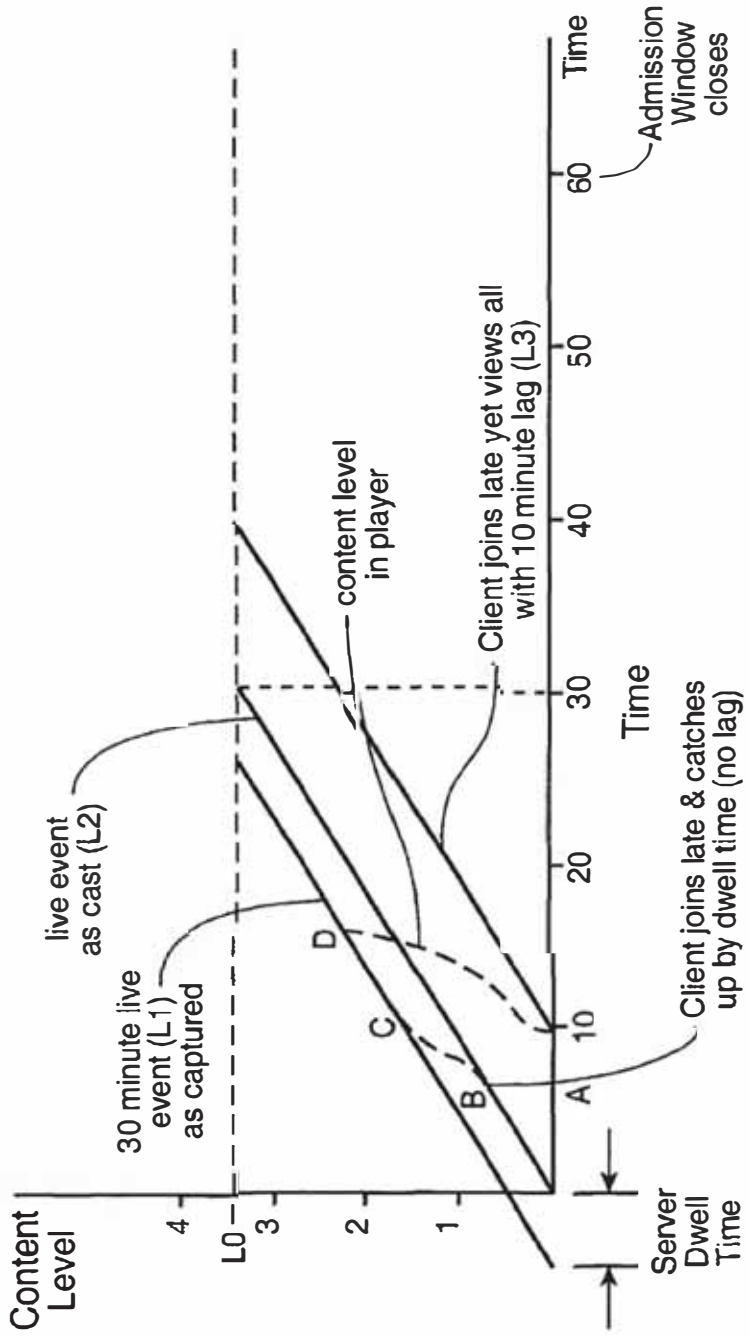


FIG._17

PATENT COOPERATION TREATY

RECEIVED
NOV 10 2008
BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN
SUNNYVALE

From the INTERNATIONAL SEARCHING AUTHORITY

To:
LESTER J. VINCENT
BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN LLP
1279 OAKMEAD PARKWAY
SUNNYVALE, CA 94085-4040

PCT

NOTIFICATION OF TRANSMITTAL OF THE INTERNATIONAL SEARCH REPORT AND THE WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY, OR THE DECLARATION

(PCT Rule 44.1)

MAILED
NOV 17 2008
BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN
LEGAL DEPT

Date of mailing *(day/month/year)* 05 NOV 2008	
Applicant's or agent's file reference 8532P004PCT	**FOR FURTHER ACTION** See paragraphs 1 and 4 below
International application No. PCT/US 08/09281	International filing date *(day/month/year)* 01 August 2008 (01.08.2008)
Applicant MOVE NETWORKS, INC.	

1. ☒ The applicant is hereby notified that the international search report and the written opinion of the International Searching Authority have been established and are transmitted herewith.

   Filing of amendments and statement under Article 19:
   The applicant is entitled, if he so wishes, to amend the claims of the international application (see Rule 46):

   When? The time limit for filing such amendments is normally two months from the date of transmittal of the international search report.

   Where? Directly to the International Bureau of WIPO, 34 chemin des Colombettes 1211 Geneva 20, Switzerland, Facsimile No.: +41 22 740 14 35

   For more detailed instructions, see the notes on the accompanying sheet.

2. ☐ The applicant is hereby notified that no international search report will be established and that the declaration under Article 17(2)(a) to that effect and the written opinion of the International Searching Authority are transmitted herewith.

3. ☐ **With regard to the protest** against payment of (an) additional fee(s) under Rule 40.2, the applicant is notified that:

   ☐ the protest together with the decision thereon has been transmitted to the International Bureau together with the applicant's request to forward the texts of both the protest and the decision thereon to the designated Offices.

   ☐ no decision has been made yet on the protest; the applicant will be notified as soon as a decision is made.

4. **Reminders**

   Shortly after the expiration of **18 months** from the priority date, the international application will be published by the International Bureau. If the applicant wishes to avoid or postpone publication, a notice of withdrawal of the international application, or of the priority claim, must reach the International Bureau as provided in Rules 90*bis*.1 and 90*bis*.3, respectively, before the completion of the technical preparations for international publication.

   The applicant may submit comments on an informal basis on the written opinion of the International Searching Authority to the International Bureau. The International Bureau will send a copy of such comments to all designated Offices unless an international preliminary examination report has been or is to be established. These comments would also be made available to the public but not before the expiration of 30 months from the priority date.

   Within **19 months** from the priority date, but only in respect of some designated Offices, a demand for international preliminary examination must be filed if the applicant wishes to postpone the entry into the national phase **until 30 months** from the priority date (in some Offices even later); otherwise, the applicant must, **within 20 months** from the priority date, perform the prescribed acts for entry into the national phase before those designated Offices.

   In respect of other designated Offices, the time limit of **30 months** (or later) will apply even if no demand is filed within 19 months.

   See the Annex to Form PCT/IB/301 and, for details about the applicable time limits, Office by Office, see the *PCT Applicant's Guide*, Volume II, National Chapters and the WIPO Internet site.

Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/220 (January 2004) *(See notes on accompanying sheet)*

Entered into PIP
by: dlm

DATE IN TO FOREIGN DOCKETING 11/11/08
DOCKETED BY
REVIEWED BY
DATE OUT

PATENT COOPERATION TREATY

From the INTERNATIONAL SEARCHING AUTHORITY

To:
LESTER J. VINCENT
BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN LLP
1279 OAKMEAD PARKWAY
SUNNYVALE, CA 94085-4040

PCT

NOTIFICATION OF TRANSMITTAL OF THE INTERNATIONAL SEARCH REPORT AND THE WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY, OR THE DECLARATION

(PCT Rule 44.1)

	Date of mailing *(day/month/year)* 05 NOV 2008
Applicant's or agent's file reference 8532P004PCT	**FOR FURTHER ACTION** See paragraphs 1 and 4 below
International application No. PCT/US 08/09281	International filing date *(day/month/year)* 01 August 2008 (01.08.2008)
Applicant MOVE NETWORKS, INC.	

1. ☒ The applicant is hereby notified that the international search report and the written opinion of the International Searching Authority have been established and are transmitted herewith.

   Filing of amendments and statement under Article 19:
   The applicant is entitled, if he so wishes, to amend the claims of the international application (see Rule 46):

   When? The time limit for filing such amendments is normally two months from the date of transmittal of the international search report.

   Where? Directly to the International Bureau of WIPO, 34 chemin des Colombettes 1211 Geneva 20, Switzerland, Facsimile No.: +41 22 740 14 35

   For more detailed instructions, see the notes on the accompanying sheet.

2. ☐ The applicant is hereby notified that no international search report will be established and that the declaration under Article 17(2)(a) to that effect and the written opinion of the International Searching Authority are transmitted herewith.

3. ☐ **With regard to the protest** against payment of (an) additional fee(s) under Rule 40.2, the applicant is notified that:

   ☐ the protest together with the decision thereon has been transmitted to the International Bureau together with the applicant's request to forward the texts of both the protest and the decision thereon to the designated Offices.

   ☐ no decision has been made yet on the protest; the applicant will be notified as soon as a decision is made.

4. **Reminders**

   Shortly after the expiration of **18 months** from the priority date, the international application will be published by the International Bureau. If the applicant wishes to avoid or postpone publication, a notice of withdrawal of the international application, or of the priority claim, must reach the International Bureau as provided in Rules 90*bis*.1 and 90*bis*.3, respectively, before the completion of the technical preparations for international publication.

   The applicant may submit comments on an informal basis on the written opinion of the International Searching Authority to the International Bureau. The International Bureau will send a copy of such comments to all designated Offices unless an international preliminary examination report has been or is to be established. These comments would also be made available to the public but not before the expiration of 30 months from the priority date.

   Within **19 months** from the priority date, but only in respect of some designated Offices, a demand for international preliminary examination must be filed if the applicant wishes to postpone the entry into the national phase **until 30 months** from the priority date (in some Offices even later); otherwise, the applicant must, **within 20 months** from the priority date, perform the prescribed acts for entry into the national phase before those designated Offices.

   In respect of other designated Offices, the time limit of **30 months** (or later) will apply even if no demand is filed within 19 months.

   See the Annex to Form PCT/IB/301 and, for details about the applicable time limits, Office by Office, see the *PCT Applicant's Guide*, Volume II, National Chapters and the WIPO Internet site.

Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/220 (January 2004) *(See notes on accompanying sheet)*

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference 8532P004PCT	FOR FURTHER ACTION	see Form PCT/ISA/220 as well as, where applicable, item 5 below.
International application No. PCT/US 08/09281	International filing date *(day/month/year)* 01 August 2008 (01.08.2008)	(Earliest) Priority Date *(day/month/year)* 06 August 2007 (06.08.2007)
Applicant MOVE NETWORKS, INC.		

This international search report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This international search report consists of a total of 9 sheets.

☐ It is also accompanied by a copy of each prior art document cited in this report.

1. **Basis of the report**

a. With regard to the **language**, the international search was carried out on the basis of:

☒ the international application in the language in which it was filed.

☐ a translation of the international application into _____ which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b)).

b. ☐ This international search report has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43.6*bis*(a)).

c. ☐ With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, see Box No. I.

2. ☐ **Certain claims were found unsearchable** (see Box No. II).

3. ☐ **Unity of invention is lacking** (see Box No. III).

4. With regard to the **title**,

☒ the text is approved as submitted by the applicant.

☐ the text has been established by this Authority to read as follows:

5. With regard to the **abstract**,

☒ the text is approved as submitted by the applicant.

☐ the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box No. IV. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. With regard to the **drawings**,

a. the figure of the **drawings** to be published with the abstract is Figure No. 1

☐ as suggested by the applicant.

☒ as selected by this Authority, because the applicant failed to suggest a figure.

☐ as selected by this Authority, because this figure better characterizes the invention.

b. ☐ none of the figures is to be published with the abstract.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 08/09281

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - G06F 15/16 (2008.04)
USPC - 709/231
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)
USPC: 709/231

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC: 709/206, 207,.219, 225, 230-235

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Electronic Databases Searched: PubWEST (PGPB,USPT,USOC,EPAB,JPAB); GoogleScholar
Search Terms Used: streamlets, adaptive stream, multi-rate, multi-bitrate, rate-shifting, skip, ignore, scrub, streaming video/media, filter, block, simulcoding, timeline, live horizon, ratings, porn, mute etc.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2005/0262257 A1 (MAJOR et al.) 24 November 2005 (24.11.2005) (shared author) entire document, especially: Abstract, para [0045], [0047], [0073]	1-26
Y	Clement, B., Page One PR, "Move Networks Closes $11.3 Million In First Round VC Funding," Move Networks, Inc. Press Releases, 07 February 2007 (07.02.2007) [retrieved 12 October 2008 (12.10.2008)] Retrieved from the Internet. <URL: http://www.move.tv/press/press20070207.html> entire document, especially: pg 1, para 4-6	1-26
Y	US 2006/0277564 A1 (JARMAN) 07 December 2006 (07.12.2006) entire document, especially: Abstract, para [0010], [0015], [0068]	4-6, 12-13, 17-19, 22-24
A	Move Networks, "The Next Generation Video Publishing System," 11 April 2007 (11.04.2007) [retrieved 13 October 2008 (13.10.2008)]. Retrieved from the Internet. <URL: http://www.movenetworks.com/wp-content/uploads/move-networks-publishing-system.pdf>	1-26

☐ Further documents are listed in the continuation of Box C. ☐

* 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 another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
14 October 2008 (14.10.2008)	05 NOV 2008
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Lee W. Young PCT helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (April 2007)

PATENT COOPERATION TREATY

From the
INTERNATIONAL SEARCHING AUTHORITY

To:
LESTER J. VINCENT
BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN LLP
1279 OAKMEAD PARKWAY
SUNNYVALE, CA 94085-4040

PCT

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

(PCT Rule 43*bis*.1)

Date of mailing *(day/month/year)*	05 NOV 2008
Applicant's or agent's file reference 8532P004PCT	FOR FURTHER ACTION See paragraph 2 below

International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/US 08/09281	01 August 2008 (01.08.2008)	06 August 2007 (06.08.2007)

International Patent Classification (IPC) or both national classification and IPC
IPC(8) - G06F 15/16 (2008.04)
USPC - 709/231

Applicant MOVE NETWORKS, INC.

1. This opinion contains indications relating to the following items:

- [X] Box No. I Basis of the opinion
- [ ] Box No. II Priority
- [ ] Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- [ ] Box No. IV Lack of unity of invention
- [X] Box No. V Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- [ ] Box No. VI Certain documents cited
- [ ] Box No. VII Certain defects in the international application
- [ ] Box No. VIII Certain observations on the international application

2. **FURTHER ACTION**

If a demand for international preliminary examination is made, this opinion will be considered to be a written opinion of the International Preliminary Examining Authority ("IPEA") except that this does not apply where the applicant chooses an Authority other than this one to be the IPEA and the chosen IPEA has notified the International Bureau under Rule 66.1*bis*(b) that written opinions of this International Searching Authority will not be so considered.

If this opinion is, as provided above, considered to be a written opinion of the IPEA, the applicant is invited to submit to the IPEA a written reply together, where appropriate, with amendments, before the expiration of 3 months from the date of mailing of Form PCT/ISA/220 or before the expiration of 22 months from the priority date, whichever expires later.

For further options, see Form PCT/ISA/220.

3. For further details, see notes to Form PCT/ISA/220.

Name and mailing address of the ISA/US	Date of completion of this opinion	Authorized officer:
Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	14 October 2008 (14.10.2008)	Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/237 (cover sheet) (April 2007)

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US 06/09281

Box No. I Basis of this opinion

1. With regard to the **language**, this opinion has been established on the basis of:

   ☒ the international application in the language in which it was filed.

   ☐ a translation of the international application into _____ which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b)).

2. ☐ This opinion has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43*bis*.1(a))

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, this opinion has been established on the basis of:

   a. type of material

      ☐ a sequence listing

      ☐ table(s) related to the sequence listing

   b. format of material

      ☐ on paper

      ☐ in electronic form

   c. time of filing/furnishing

      ☐ contained in the international application as filed

      ☐ filed together with the international application in electronic form

      ☐ furnished subsequently to this Authority for the purposes of search

4. ☐ In addition, in the case that more than one version or copy of a sequence listing and/or table(s) relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

5. Additional comments:

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US 06/09281

Box No. V **Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**

1. Statement

Novelty (N)	Claims	1-26	YES
	Claims	None	NO
Inventive step (IS)	Claims	None	YES
	Claims	1-26	NO
Industrial applicability (IA)	Claims	1-26	YES
	Claims	None	NO

2. Citations and explanations:

Claims 1-3, 7-11, 14-16, 20-21 and 25-26 lack inventive step under PCT Article 33(3) as being obvious over US 2005/0262257 A1 to Major et al. (hereinafter 'Major') in view of the article entitled "Move Networks Closes $11.3 Million in First Round VC Funding," (Clement).

As to claims 1, 9, 15, 21 and 26, Major teaches a system, method, apparatus and computer program product for multi-bitrate content streaming (adaptive-rate shifting of streaming content, see Abstract), wherein the media content comprises a plurality of streamlets (plurality of streamlets, see Abstract); and at least one data module configured to maintain multi-bitrate streamlet information (staging module is configured to stage the streamlets and arrange the streamlets for playback, see Abstract).
Major does not expressly teach a timeline module configured to maintain a programming lineup of media content available over a network. However, Clement teaches multi-bitrate content streaming (adaptive stream that upshifts and downshifts to match the current constraints, see pg 1, para 4) and a timeline module configured to maintain a programming lineup of media content available over a network (TiVo style navigation, see pg 1, para 6) and a client module configured to acquire content based upon the programming lineup provided by the timeline module (Move Media Player - Viewer Controls, TiVo style, see para 6).
It would have been obvious to one of ordinary skill in the art to combine the timeline module, as taught by Clement, with the Major system for multi-bitrate content streaming in order to enable the user to efficiently select and manage their desired content, thus increasing the ease of use for the end-user and increasing the revenue potential for the content providers.

As to claims 2 and 10, Clement teaches each streamlet comprises a portion of the media content encapsulated as an independent media object (Each streamlet 212 may comprise a portion of the content contained in stream 210, encapsulated as an independent media object, see para [0047]), and wherein each streamlet comprises a predetermined duration of time (In one embodiment, the content contained in each streamlet 212 has a duration of two seconds, see para [0047]).

As to claims 3 and 11 and 16, Major teaches multi-bitrate streamlet information is selected from the group consisting of start time, end time, live video, publisher data, encryption, duration, bitrate values, frame size, channels, codecs, sample rate, and frames parser (content file 200 may also be live or archived content, see para [0045] and Fig. 2a).

As to claims 7 and 25, Clement teaches a content management system configured to dynamically replace portions of the media content (shows can even be pre-empted and replaced by alternative programming, see pg 1, para 5).

As to claim 8, Major teaches an anchor module configured to associate the start of the timeline module with a point in time such that playback of the media content begins at a point corresponding to a live horizon (presentation at the current playback time index, see pg 1, para [0073]).

As to claims 14 and 20, Clement teaches a content management system configured to dynamically replace portions of the media content (shows can even be pre-empted and replaced by alternative programming, see pg 1, para 5).

(See Supplemental Box)

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US 06/09281

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.

Continuation of:
Box No. V --- Reasoned Statement
2. Citations and Explanations:

Claims 4-6, 12-13, 17-19 and 22-24 lack inventive step under PCT Article 33(3) as being obvious over Major in view of Clement as applied above, and further in view of US 2006/0277564 A1 (Jarman).

As to claims 4, 12, 17, 22 and 23, the combination of Major and Clement teach the apparatus of claim 1, the system of claim 9, the method of claim 15 and the computer program product of claim 21, and Major further teaches streamlets providing content (a staging module configured to stage the streamlets and arrange the streamlets for playback on a content player , see para [0015]), but the combination of Major and Clement does not expressly teach a module configured to skip content in response to a rating system level. However, Jarman teaches a programming content distribution system which further comprising a client module configured to skip content in response to a rating system (selectively mutes context specific words as a function of program specific criteria, see Abstract).
It would have been obvious to one of ordinary skill in the art to combine the Jarman content filtering system with the content distribution system taught by Major and Clement because this would enable the user to instruct the more precisely control the desired viewing content; further, Clement suggests content filtering (complete editing and publishing control, see pg 1, para 4) such as that taught by the Jarman content filtering system.

As to claims 5 and 18, Jarman teaches the client module is configured to skip streamlets that do not conform to a selected rating level (allow play of only programs associate with selected program specific information; and blocking display of all programs not including the selected program specific information, see para [0010]).

As to claim 6, Jarman teaches the client module is configured to mute the audio of streamlets that do not conform to a selected rating level (mute words for an audio or audio/video presentation that may be objectionable in one context, see para [0068]).

As to claims 13, 19 and 24, the combination of Major and Clement teach the system of claim 9 and the method of claim 15, but does not expressly teach the client module is configured to mute the audio of streamlets that do not conform to a selected rating level.
However, Jarman teaches multi-bitrate content streaming (adaptive stream that upshifts and downshifts to match the current constraints, see pg 1, para 4) and a timeline module configured to maintain a programming lineup of media content available over a network (TiVo style navigation, see pg 1, para 6) and wherein the client module is configured to mute the audio of streamlets that do not conform to a selected rating level (electively mutes context specific words as a function of program specific criteria, see Abstract. It would have been obvious to combine the Jarman selective content filtering and muting system with the content distribution system taught by Major and Clement because this would enable the user to instruct the more precisely control the desired viewing content; further, Clement suggests content filtering (complete editing and publishing control, see pg 1, para 4).

Claims 1-26 have industrial applicability as defined by PCT Article 33(4) because the subject matter claimed can be made or used in industry.

NOTES TO FORM PCT/ISA/220

These Notes are intended to give the basic instructions concerning the filing of amendments under Article 19. The Notes are based on the requirements of the Patent Cooperation Treaty, the Regulations and the Administrative Instructions under that Treaty. In case of discrepancy between these Notes and those requirements, the latter are applicable. For more detailed information, see also the *PCT Applicant's Guide*, a publication of WIPO.

In these Notes, "Article," "Rule" and "Section" refer to the provisions of the PCT, the PCT Regulations and the PCT Administrative Instructions, respectively.

INSTRUCTIONS CONCERNING AMENDMENTS UNDER ARTICLE 19

The applicant has, after having received the international search report and the written opinion of the International Searching Authority, one opportunity to amend the claims of the international application. It should however be emphasized that, since all parts of the international application (claims, description and drawings) may be amended during the international preliminary examination procedure, there is usually no need to file amendments of the claims under Article 19 except where, e.g. the applicant wants the latter to be published for the purposes of provisional protection or has another reason for amending the claims before international publication. Furthermore, it should be emphasized that provisional protection is available in some States only (see *PCT Applicant's Guide*, Volume I/A, Annexes B1 and B2).

The attention of the applicant is drawn to the fact that amendments to the claims under Article 19 are not allowed where the International Searching Authority has declared, under Article 17(2), that no international search report would be established (see *PCT Applicant's Guide*, Volume I/A, paragraph 296).

What parts of the international application may be amended ?

Under Article 19, only the claims may be amended.

During the international phase, the claims may also be amended (or further amended) under Article 34 before the International Preliminary Examining Authority. The description and drawings may only be amended under Article 34 before the International Preliminary Examining Authority.

Upon entry into the national phase, all parts of the international application may be amended under Article 28 or, where applicable, Article 41.

When ? Within 2 months from the date of transmittal of the international search report or 16 months from the priority date, whichever time limit expires later. It should be noted, however, that the amendments will be considered as having been received on time if they are received by the International Bureau after the expiration of the applicable time limit but before the completion of the technical preparations for international publication (Rule 46.1).

Where not to file the amendments ?

The amendments may only be filed with the International Bureau and not with the receiving Office or the International Searching Authority (Rule 46.2).

Where a demand for international preliminary examination has been/is filed, see below.

How ? Either by cancelling one or more entire claims, by adding one or more new claims or by amending the text of one or more of the claims as filed.

A replacement sheet must be submitted for each sheet of the claims which, on account of an amendment or amendments, differs from the sheet originally filed.

All the claims appearing on a replacement sheet must be numbered in Arabic numerals. Where a claim is cancelled, no renumbering of the other claims is required. In all cases where claims are renumbered, they must be renumbered consecutively (Section 205(b)).

The amendments must be made in the language in which the international application is to be published.

What documents must/may accompany the amendments ?

Letter (Section 205(b)):

The amendments must be submitted with a letter.

The letter will not be published with the international application and the amended claims. It should not be confused with the "Statement under Article 19(1)" (see below, under "Statement under Article 19(1)").

The letter must be in English or French, at the choice of the applicant. However, if the language of the international application is English, the letter must be in English; if the language of the international application is French, the letter must be in French.

PATENT COOPERATION TREATY

From the INTERNATIONAL SEARCHING AUTHORITY

To:
BRIAN C. KUNZLER
8 EAST BROADWAY, SUITE 600
SALT LAKE CITY, UT 84111

RECEIVED
JAN 09 2009

PCT

NOTIFICATION OF TRANSMITTAL OF THE INTERNATIONAL SEARCH REPORT AND THE WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY, OR THE DECLARATION

(PCT Rule 44.1)

Date of mailing *(day month year)*	12 DEC 2008
Applicant's or agent's file reference 2082.2.4PCT	**FOR FURTHER ACTION** See paragraphs 1 and 4 below
International application No. PCT/US 08/61035 8532P003PCT	International filing date *(day month year)* 21 April 2008 (21.04.2008)
Applicant MOVE NETWORKS, INC.	

1. ☒ The applicant is hereby notified that the international search report and the written opinion of the International Searching Authority have been established and are transmitted herewith.

Filing of amendments and statement under Article 19:
The applicant is entitled, if he so wishes, to amend the claims of the international application (see Rule 46):

When? The time limit for filing such amendments is normally two months from the date of transmittal of the international search report.

Where? Directly to the International Bureau of WIPO, 34 chemin des Colombettes 1211 Geneva 20, Switzerland, Facsimile No.: +41 22 740 14 35

For more detailed instructions, see the notes on the accompanying sheet.

2. ☐ The applicant is hereby notified that no international search report will be established and that the declaration under Article 17(2)(a) to that effect and the written opinion of the International Searching Authority are transmitted herewith.

3. ☐ **With regard to the protest** against payment of (an) additional fee(s) under Rule 40.2, the applicant is notified that:

☐ the protest together with the decision thereon has been transmitted to the International Bureau together with the applicant's request to forward the texts of both the protest and the decision thereon to the designated Offices.

☐ no decision has been made yet on the protest; the applicant will be notified as soon as a decision is made.

4. **Reminders**

Shortly after the expiration of **18 months** from the priority date, the international application will be published by the International Bureau. If the applicant wishes to avoid or postpone publication, a notice of withdrawal of the international application, or of the priority claim, must reach the International Bureau as provided in Rules 90*bis*.1 and 90*bis*.3, respectively, before the completion of the technical preparations for international publication.

The applicant may submit comments on an informal basis on the written opinion of the International Searching Authority to the International Bureau. The International Bureau will send a copy of such comments to all designated Offices unless an international preliminary examination report has been or is to be established. These comments would also be made available to the public but not before the expiration of 30 months from the priority date.

Within **19 months** from the priority date, but only in respect of some designated Offices, a demand for international preliminary examination must be filed if the applicant wishes to postpone the entry into the national phase **until 30 months** from the priority date (in some Offices even later); otherwise, the applicant must, **within 20 months** from the priority date, perform the prescribed acts for entry into the national phase before those designated Offices.

In respect of other designated Offices, the time limit of **30 months** (or later) will apply even if no demand is filed within 19 months.

See the Annex to Form PCT/IB/361 and, for details about the applicable time limits, Office by Office, see the *PCT Applicant's Guide*, Volume II, National Chapters and the WIPO Internet site.

Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/220 (January 2004) *(See notes on accompanying sheet)*

Date In to SV FF: 1-12-09
Docketed by: MFa
Reviewed by:
Sent Out by SV FF on:

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference 2082.2.4PCT	FOR FURTHER ACTION	see Form PCT/ISA/220 as well as, where applicable, item 5 below.
International application No. PCT/US 08/61035	International filing date *(day/month/year)* 21 April 2008 (21.04.2008)	(Earliest) Priority Date *(day/month/year)* 19 April 2007 (19.04.2007)
Applicant MOVE NETWORKS, INC.		

This international search report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This international search report consists of a total of 9 sheets.

☐ It is also accompanied by a copy of each prior art document cited in this report.

1. **Basis of the report**

a. With regard to the **language**, the international search was carried out on the basis of:

☒ the international application in the language in which it was filed.

☐ a translation of the international application into _____ which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b)).

b. ☐ This international search report has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43.6*bis*(a)).

c. ☐ With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, see Box No. I.

2. ☐ **Certain claims were found unsearchable** (see Box No. II).

3. ☐ **Unity of invention is lacking** (see Box No. III).

4. With regard to the **title**.

☒ the text is approved as submitted by the applicant.

☐ the text has been established by this Authority to read as follows:

5. With regard to the **abstract**,

☒ the text is approved as submitted by the applicant.

☐ the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box No. IV. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. With regard to the **drawings**,

a. the figure of the **drawings** to be published with the abstract is Figure No. 4

☐ as suggested by the applicant.

☐ as selected by this Authority, because the applicant failed to suggest a figure.

☒ as selected by this Authority, because this figure better characterizes the invention.

b. ☐ none of the figures is to be published with the abstract.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 08/61035

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - G06F 15/16 (2008.04)
USPC - 709/227
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)
USPC: 709/227

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC: 709/201, 203, 223, 227, 228, 238; 715/736, 744, 748, 968

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PubWest(PGPB,USPT,USOC,EPAB,JPAB); Google
Search Terms Used: content acquisition, resilient, networks, tcp, directive handling module, content acquisition module, directive server, network connection, remote host etc.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2005/0188051 A1 (SNEH) 25 August 2005 (25.08.2005), para [0005], para [0055]-[0097], para [0121]-[0136], para [0188]	1-36
Y	US 2003/0014684 A1 (KASHYAP) 16 January 2003 (16.01.2003), para [0101]	1-36
A	US 2006/0206246 A1 (WALKER) 14 September 2006 (14.09.2006)	1-36

☐ Further documents are listed in the continuation of Box C. ☐

* 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 another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25 November 2008 (25.11.2008)	12 DEC 2008

Name and mailing address of the ISA/US	Authorized officer:
Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (April 2007)

PATENT COOPERATION TREATY

From the
INTERNATIONAL SEARCHING AUTHORITY

To:
BRIAN C. KUNZLER
8 EAST BROADWAY, SUITE 600
SALT LAKE CITY, UT 84111

PCT

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

(PCT Rule 43*bis*.1)

Date of mailing *(day/month/year)*	12 DEC 2008
Applicant's or agent's file reference 2082.2.4PCT	FOR FURTHER ACTION See paragraph 2 below

International application No.	International filing date *(day month year)*	Priority date *(day month year)*
PCT/US 08/61035	21 April 2008 (21.04.2008)	19 April 2007 (19.04.2007)

International Patent Classification (IPC) or both national classification and IPC
IPC(8) - G06F 15/16 (2008.04)
USPC - 709/227

Applicant MOVE NETWORKS, INC.

1. This opinion contains indications relating to the following items:

- [X] Box No. I Basis of the opinion
- [ ] Box No. II Priority
- [ ] Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- [ ] Box No. IV Lack of unity of invention
- [X] Box No. V Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- [ ] Box No. VI Certain documents cited
- [ ] Box No. VII Certain defects in the international application
- [ ] Box No. VIII Certain observations on the international application

2. **FURTHER ACTION**

If a demand for international preliminary examination is made, this opinion will be considered to be a written opinion of the International Preliminary Examining Authority ("IPEA") except that this does not apply where the applicant chooses an Authority other than this one to be the IPEA and the chosen IPEA has notified the International Bureau under Rule 66.1*bis*(b) that written opinions of this International Searching Authority will not be so considered.

If this opinion is, as provided above, considered to be a written opinion of the IPEA, the applicant is invited to submit to the IPEA a written reply together, where appropriate, with amendments, before the expiration of 3 months from the date of mailing of Form PCT/ISA/220 or before the expiration of 22 months from the priority date, whichever expires later.

For further options, see Form PCT/ISA/220.

3. For further details, see notes to Form PCT/ISA/220.

Name and mailing address of the ISA/US	Date of completion of this opinion	Authorized officer:
Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	26 November 2008 (26.11.2008)	Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/237 (cover sheet) (April 2007)

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US 08/61035

Box No. I Basis of this opinion

1. With regard to the **language**, this opinion has been established on the basis of:

   ☒ the international application in the language in which it was filed.

   ☐ a translation of the international application into _____ which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b)).

2. ☐ This opinion has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43*bis*.1(a))

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, this opinion has been established on the basis of:

   a. type of material

      ☐ a sequence listing

      ☐ table(s) related to the sequence listing

   b. format of material

      ☐ on paper

      ☐ in electronic form

   c. time of filing/furnishing

      ☐ contained in the international application as filed

      ☐ filed together with the international application in electronic form

      ☐ furnished subsequently to this Authority for the purposes of search

4. ☐ In addition, in the case that more than one version or copy of a sequence listing and/or table(s) relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

5. Additional comments:

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US 08/61035

Box No. V **Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**

1. Statement

Novelty (N)	Claims	1-36	YES
	Claims	None	NO
Inventive step (IS)	Claims	None	YES
	Claims	1-36	NO
Industrial applicability (IA)	Claims	1-36	YES
	Claims	None	NO

2. Citations and explanations:

Claims 1-36 lack an inventive step under PCT Article 33(3) as being obvious over US 2005/0188051 A1 (Sneh) in view of US 2003/0014684 A1 (Kashyap).

Regarding claim 1, Sneh discloses an apparatus for resilient content acquisition (content, para [0005]), the apparatus comprising: a directive handling module configured to implement a directive (directives, para [0055]); wherein the directive defines requirements for establishing, locating, and maintaining a network connection (maintain download state, para [0097]); and a content acquisition module configured to acquire content over a network according to the directive (content acquisition, para [0081]). Sneh does not expressly disclose maintaining a network connection with a remote host. Kashyap teaches maintaining a network connection with a remote host (remote host, para [0101]). It would have been obvious to one of ordinary skill in the art at the time of the applicant's invention to modify the method of Sneh to include maintaining a network connection with a remote host as taught by Kashyap because for the advantage of expanding the network beyond immediate systems for content acquisition.

Regarding claim 2, in the combination of Sneh and Kashyap, Sneh further discloses the content acquisition module is further configured to transmit a directive server query to a directive server (CA Server, para [0136]).

Regarding claim 3, in the combination of Sneh and Kashyap, Sneh further discloses the directive handling module is further configured to receive the directive from the directive server (directives, para [0055]).

Regarding claim 4, in the combination of Sneh and Kashyap, Sneh further discloses the content acquisition module is further configured to locate the content, authorize the content, retrieve the content, display the content, and report statistics in accordance with the directive (content acquisition, para [0081]).

Regarding claim 5, in the combination of Sneh and Kashyap, Sneh further discloses a directive server module configured to maintain a list of directive servers and to select a directive server for each directive request (CA Server, para [0136]).

Regarding claim 6, in the combination of Sneh and Kashyap, Sneh further discloses a server status module configured to maintain a status indicator for each of a plurality of content delivery network servers (status, para [0188]).

Regarding claim 7, in the combination of Sneh and Kashyap, Sneh further discloses a broadcast module configured to broadcast content availability over a network such that a second apparatus acquires the content from the apparatus over the network (content acquisition, para [0081]).

Regarding claim 8, in the combination of Sneh and Kashyap, Kashyap further discloses a discovery module configured to discover available content on the network such that content is acquired from a second apparatus (remote host, para [0101]).

Regarding claim 9, in the combination of Sneh and Kashyap, Kashyap further discloses the directive is further configured to identify a plurality of backup remote hosts (remote host, para [0101]).

(See Supplemental Box)

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.

Continuation of:
Box No. V — Reasoned statement
2. Citations & explanations:

Regarding claim 10, Sneh discloses a system for resilient content acquisition (content, para [0005]), the system comprising: a client module in communication with a network (modules, para [0121]-[0127]), the client module comprising: a directive handling module configured to implement a directive (directives, para [0055]); wherein the directive defines requirements for establishing, locating, and maintaining a network connection (maintain download state, para [0097]); a content acquisition module configured to acquire content over the network according to the directive (content acquisition, para [0081]); wherein the content acquisition module is further configured to communicate a directive server query with a directive server (CA Server, para [0136]); and wherein the directive server is configured to parse the content request and generate a directive in response to the content request (CA Server, para [0136]). Sneh does not expressly disclose maintaining a network connection with a remote host. Kashyap teaches maintaining a network connection with a remote host (remote host, para [0101]). It would have been obvious to one of ordinary skill in the art at the time of the applicant's invention to modify the method of Sneh to include maintaining a network connection with a remote host as taught by Kashyap because for the advantage of expanding the network beyond immediate systems for content acquisition.

Regarding claim 11, in the combination of Sneh and Kashyap, Sneh further discloses the directive handling module is further configured to receive the directive from the directive server (directives, para [0055]).

Regarding claim 12, in the combination of Sneh and Kashyap, Sneh further discloses the content acquisition module is further configured to locate the content, authorize the content, retrieve the content, display the content, and report statistics in accordance with the directive (content acquisition, para [0081]).

Regarding claim 13, in the combination of Sneh and Kashyap, Sneh further discloses a directive server module configured to maintain a list of directive servers and to select a directive server for each directive request (CA Server, para [0136]).

Regarding claim 14, in the combination of Sneh and Kashyap, Sneh further discloses a server status module configured to maintain a status indicator for each of a plurality of content delivery network servers (status, para [0188]).

Regarding claim 15, in the combination of Sneh and Kashyap, Sneh further discloses a broadcast module configured to broadcast content availability over a local network such that a second apparatus acquires the content from the apparatus over the network (content acquisition, para [0081]).

Regarding claim 16, in the combination of Sneh and Kashyap, Kashyap further discloses a discovery module configured to discover available content on a network such that content is acquired from a second apparatus (remote host, para [0101]).

Regarding claims 17, 25 and 33, Sneh discloses a method, computer program product, and apparatus for resilient content acquisition (content, para [0005]), comprising: implementing a directive (directives, para [0055]); defining requirements for locating, establishing, and maintaining a network connection (maintain download state, para [0097]); and acquiring content over a network according to the directive (content acquisition, para [0081]). Sneh does not expressly disclose maintaining a network connection with a remote host. Kashyap teaches maintaining a network connection with a remote host (remote host, para [0101]). It would have been obvious to one of ordinary skill in the art at the time of the applicant's invention to modify the method of Sneh to include maintaining a network connection with a remote host as taught by Kashyap for the advantage of expanding the network beyond immediate systems for content acquisition.

Regarding claims 18 and 26, in the combination of Sneh and Kashyap, Sneh further discloses transmitting a directive server query to a directive server (CA Server, para [0136]).

Regarding claims 19 and 27, in the combination of Sneh and Kashyap, Sneh further discloses receiving the directive from the directive server (CA Server, para [0136]).

Regarding claims 20 and 28, in the combination of Sneh and Kashyap, Sneh further discloses locating the content, authorizing the content, retrieving the content, displaying the content, and reporting statistics in accordance with the directive (content acquisition, para [0081]).

Regarding claims 21 and 29, in the combination of Sneh and Kashyap, Sneh further discloses maintaining a list of directive servers and selecting a directive server for each directive request (CA Server, para [0136]).

Regarding claims 22 and 30, in the combination of Sneh and Kashyap, Sneh further discloses maintaining a status indicator for each of a plurality of content delivery network servers (status, para [0188]).

Regarding claims 23 and 31, in the combination of Sneh and Kashyap, Kashyap further discloses broadcasting content availability over a local network such that a second apparatus acquires the content from the apparatus over the network (remote host, para [0101]).

Regarding claims 24 and 32, in the combination of Sneh and Kashyap, Kashyap further discloses discovering available content on a network such that content is acquired from a second apparatus (remote host, para [0101]).

(See Next Supplemental Box)

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.

Continuation of:
Preceding Supplemental Box
Box No. V --- Reasoned statement
2. Citations & explanations:

Regarding claim 34, Sneh discloses an apparatus for resilient content delivery (content, para [0005]), the system comprising: a parsing module configured to receive a content request from a source and identify content information (modules, para [0121]-[0127]); a response module configured to generate a directive in response to the content information (directives, para [0055]), wherein the directive defines requirements for establishing, locating, and maintaining a network connection (maintain download state, para [0097]); and wherein the response module is further configured to communicate the directive with the source (content acquisition, para [0081]). Sneh does not expressly disclose maintaining a network connection with a remote host. Kashyap teaches maintaining a network connection with a remote host (remote host, para [0101]). It would have been obvious to one of ordinary skill in the art at the time of the applicant's invention to modify the method of Sneh to include maintaining a network connection with a remote host as taught by Kashyap because for the advantage of expanding the network beyond immediate systems for content acquisition.

Regarding claim 35, in the combination of Sneh and Kashyap, Sneh further discloses the content information is selected from a group consisting of host, domain, content request source, path, and content name (content acquisition, para [0081]).

Regarding claim 36, in the combination of Sneh and Kashyap, Sneh further discloses the response module is further configured to update and communicate the directive with the source (CA Server, para [0136]).

Claims 1-36 have industrial applicability as defined by PCT Article 33(4) because the subject matter can be made or used in industry.

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
30 January 2003 (30.01.2003)

PCT

(10) International Publication Number
WO 03/009581 A1

(51) International Patent Classification[7]: H04N 5/00, 7/24

(21) International Application Number: PCT/GB02/03307

(22) International Filing Date: 19 July 2002 (19.07.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
01306206.2 19 July 2001 (19.07.2001) EP

(71) Applicant *(for all designated States except US)*: **BRITISH TELECOMMUNICATIONS PUBLIC LIMITED COMPANY** [GB/GB]; 81 Newgate Street, London EC1A 7AJ (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **JACOBS, Richard, James** [GB/GB]; 15 Through Duncans, Woodbridge, Suffolk IP12 4EA (GB). **WALKER, Matthew, David** [GB/GB]; 9 Thirlmere Court, Felixstowe, Suffolk IP11 9SN (GB). **MIRAS, Dimitrios** [GR/GR]; 12 Proti Parodos Avras, 24100 Kalamata (GR).

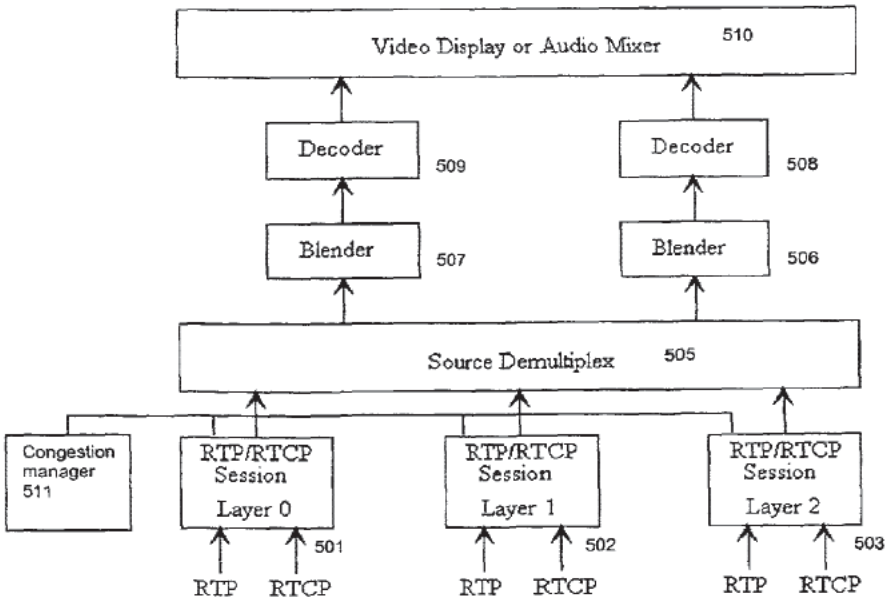
(74) Agent: **WILLIAMSON, Simeon, Paul**; 0T Group Legal, Intellectual Property Department, Holborn Centre 8th Floor, 120 Holborn, London EC1N 2TE (GB).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (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, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: VIDEO STREAM SWITCHING



(57) Abstract: A technique is described for providing a client on a packet based network with a stream of encoded video data. The system is able to maximise the bit-rate of the video stream by adapting to fluctuations in network capacity. The technique is characterised in that adaptation of the bit-rate of the transmitted encoded video data is timed to occur upon a scene change in the video sequence. In this way the interruption to the viewer when the perceived quality of the video sequence increases or decreases is minimised as it is 'hidden' in the scene change. The technique is described as applied to hierarchically encoded video data but equally may be applied to other encoding techniques which adapt to network conditions.

WO 03/009581 A1

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Video Stream Switching

The invention is in the field of video streaming over packet networks and in particular concerns the adaptive transmission of data in response to network congestion.

In recent years, the Internet has experienced a proliferation in transmission of real-time multimedia, mainly in the form of streamed audio-visual content, either delivered live or from pre-recorded sources. Furthermore, traditional forms of multimedia such as streaming and conferencing are being followed by applications with richer content such as Internet multi-channel TV and complex immersive environments. This increase in traffic will put strains on the network and it is therefore desirable that application programs are designed to respond to congestion if stability of the network is to be maintained. It is desirable that network conditions are monitored and output bit-rates adjusted to the available bandwidth.

There is a more immediate advantage to the user in having an adaptive output bit-rate which is that the highest possible transmission bit-rate that the network will allow is used and therefore the user will always receive the best possible image quality. There are a number of known approaches to adaptive quality video streaming, one of which is hierarchical coding. In this technique the original video data is encoded into a number of discreet streams called layers, where the first layer consists of basic data of a relatively poor quality and where successive layers represent more detailed information so that layers can be added to increase the image quality or layers can be taken away, depending on the available bandwidth.

As the bit-rate available to a session is subject to significant variations in the number of layers that are transmitted, quality fluctuations occur in the decoded image as layers are added or dropped. When layers are added or dropped frequently the fluctuations in quality may become disturbing to a viewer.

In a first embodiment of the present invention there is provided a method of operating a multimedia server, said method comprising:

providing a stream of video data representing a video sequence to an output of the multimedia server, wherein the output of the media server is connected to a packet based network,

measuring a property of the video data in order to determine the occurrence of a scene change in the video sequence,

detecting the available bandwidth on the network,

varying the bit-rate of the stream of video data,

wherein the method is characterised in that

variation in the bit-rate of the video data is controlled to occur in response to variations in the capacity of the network (203) and preferentially with a scene change in the video sequence.

The term preferentially used herein is meant to indicate that in certain circumstances (discussed in greater detail below) it may not be practical to wait for a change of scene to occur in the video sequence before varying the bit-rate of the video stream, and in such circumstances the bit-rate will be changed at points in the video stream which do not correspond to a change of scene.

The term scene change (or change of scene) is intended to refer to a sudden change in a video sequence within the space of a one or a very few frames such as typically occurs at a change of scene, whether or not there has been an actual change of scene.

Embodiments of the present invention will now be described, by way of example only, with reference to the following figures, where;

Figure 1 is a schematic diagram of the content-based inter-stream session bandwidth sharing architecture.

Figure 2 shows the order of packets in a layered coding system.

Figure 3 is a diagram of a media server.

Figure 4 is a diagram of the network interface of the media server shown in Figure 3.

Figure 5 is a diagram of a client.

Figure 6 is a diagram of a client according to a first embodiment of the present invention.

The client/ server arrangement for a known hierarchical streaming technology is shown in Figure 1. A media server 202 is provided which has access to compressed audiovisual data 201, which may be fed 'live' from an outside broadcast or may be pre-compressed and stored in a database. The data source 201 may be on the same premises as the media server 202 and linked via an intranet. The media server 202 runs on a suitable server computer and which has access to the Internet 203.

A video viewer, hitherto referred to as the client 204, running on a PC suitably configured to have access to the Internet 203, may connect to the media server 202 via the Internet 203 and thus the client 204 is able to access content. A suitable PC terminal is used.

Layered video compression is achieved with the 1998 version of H.263 but equally may be any other codec, such as MPEG4. Each layer in the hierarchy is coded in such a way as to allow the quality of individual pictures to be enhanced and their resolution to be increased, and additional pictures to be included to increase the overall picture rate, as explained with reference to Figure 2. Figure 2 illustrates a typical dependency between pictures in an H.263 scalable layered coder, with the boxes representing the frames for each layer and arrows showing the dependency between frames. The lowest row shows original, un-coded frames. The next row shows the lowest layer (Layer 0) of the hierarchy which is coded at half the frame rate of Layer 1. Frames in Layer 0 are predicted from the previously encoded frame, as in conventional video compression. Frames in Layer 1 may be predicted from the previously encoded frame in Layer 1 and, if present, the temporally simultaneous Layer 0 encoded frame. Frames in Layer 2 may be predicted from the previously encoded frame in Layer 2 and, if present, the temporally simultaneous Layer 1 or Layer 0 encoded frame. The H.263 specification allows for 15 layers; in the present embodiment server and client software is not limited in the number of layers that can be used, but in this case a database has been generated with video streams containing four layers.

Figure 3 shows the architecture of media server 201. In the present embodiment the audiovisual source is a database 301 of compressed video data and the media server 201 is responsible for reading the compressed data from a database, packetising and distributing it. The data is distributed according to the

Real-time Transport Protocol (RTP) which provides end-to-end network transport functions suitable for applications transmitting real-time data such as audio, video or simulation data. RTP does not address resource reservation and does not guarantee quality-of-service for real-time services. The data transport is augmented by a control protocol (RTCP) to allow monitoring of the data delivery and to provide minimal control and identification functionality. RTP and RTCP are designed to be independent of the underlying transport and network layers. The media server 201 has four major components: database reader 302, video RTP packetiser 303, audio RTP packetiser 305 and a network interface 304. Video and audio data are stored as files of compressed bits in the database 301. The database reader 302 retrieves and synchronises the compressed audio and video from the database 301.

The audio and video data is then sent to a RTP packetiser 303,305. Audio and video information is transmitted over the IP network using the User Datagram Protocol (UDP) and Real-time Transport Protocol (RTP) packetisation. UDP provides a checksum to detect transmission errors, but does not guarantee data delivery: packets may be lost, duplicated or re-ordered. RTP provides end-to-end delivery services, such as payload type identification, sequence numbering, time-stamping and delivery monitoring. RTP packetisers 303,305 attach the RTP Header and, in the case of video packets, the H.263 Payload Header which provides some protection from packet loss. The Payload Header contains information specific to the video stream, such as motion vector predictors, which is obtained by decoding the compressed bit stream.

The rate at which data is read from the database 301 is controlled from the network interface 304, which is illustrated in greater detail in Figure 4. With reference to Figure 4 the arrangement of the network interface 304 for the video data will now be described. A packet numbering module 401 assigns a number to each packet of data. These packet numbers are used by the client to assemble packets into the sequence required for decoding. The number of bits produced by compressing a picture is in general not constant, even when the transmission rate is constant. The first picture of a new scene will usually produce a larger number of bits than a picture where there is little movement or detail. The encoding process will use a control strategy together with an output data buffer to smooth these variations in preparation for transmission at a

constant bit rate. In the case of layered coding, each layer has its own control strategy and buffer as it has its own constant bit rate constraint to meet. The delay across each of these buffers will not be constant and will not be the same for each layer. Delay variation between the layers is therefore introduced at the source of the data. Transmission across a network may cause this delay variation to increase. A packet numbering scheme is therefore needed to ensure that a client can arrange the packets it receives into the order required for decoding. This scheme is required to handle the cases of the client not receiving all of the layers and of packet loss in the layers that it does receive.

Each layer is transmitted as an independent RTP Session on a separate IP address by a session handler 403, 405, 407, 409. The rate at which data is transmitted is controlled by the Transfer Rate Control module 402, which counts Layer 0 bytes to ensure that the correct number are transmitted in a given period of time. The transmission rate of the other layers is smoothed and locked to the rate of Layer 0 using First-In First-Out (FIFO) buffer elements 404, 406, 408.

The client 204 of the known hierarchical streaming technology will now be described with reference to Figure 5. Each RTP/RTCP Session associated with each layer of encoded data has a session handler 501, 502, 503, 504 at the client which is responsible for receiving RTP packets from the network.

The packets are then passed to a source demultiplex module 505. This receives packets from all layers and all sources, demultiplexes them by source, and routes them to a blender module 506, 507 for that source.

The blender module 506, 507 receives packets from all layers from one source in the order that they were received from the network. This may not be the order required for decoding because of packet inter-arrival jitter or packet loss. The blender module 506, 507 uses the packet numbers in the RTP headers to arrange the packets from each layer in order and then combines the packets from all layers together. The output of the blender module 506, 507 is a single stream of packets which are in the correct order for decoding. This is the same order as they come out of the packet numbering module 401 of the media server 202.

Packets are then sent to the decoder 508, 509 where the packets are decoded into 20ms blocks of audio samples or to video pictures. In the case of pictures, these are rendered to a window on the display 508.

Also provided is a congestion manager 511 to which the session handlers 501-504 report packet loss. If packets are being consistently lost, indicating network congestion, the congestion manager 511 will instruct the session handler responsible for the highest layer of compressed data to terminate the RTP/RTCP Session. Periodically the congestion manager 511 will instruct an experimental joining of a layer via the appropriate session handler to test whether there is available bandwidth in the network. If this experiment is successful, i.e. that substantially all of the packets of each layer are getting reaching the client, then the new layer will be adopted. In this way the maximum available bandwidth is employed.

It possible for the congestion manager 509 to instruct layers to be dropped and restored rapidly as the network bandwidth fluctuates. This could be annoying for a viewer and so a way of 'hiding' the shift is employed. The layered encoding method described above is adapted so that if a layer is to be dropped or added, the changeover preferentially occurs during a scene-change in the video data. A video scene is typically perceived as the number of consecutive frames within a video sequence that do not show significant changes in the video content. Within a video scene or shot, the camera action may be fixed or may exhibit a number of relatively uniform changes like panning, zooming, tracking etc. Scene changes may be recognised as abrupt transitions of the camera action or gradual transitions. In order to identify the scenes within a video sequence the assumption is that the levels of motion energy as well as those of luminance and colour do not change much between successive frames within a single scene. Techniques of scene boundary identification include pixel differencing, motion vector and block matching techniques. A very sudden change in the content of temporally adjacent frames will thus appear as a change of scene; clearly it is conceivable that such a change might not actually relate to a scene change at all, but, for example, may instead be the consequence of a large foreground object coming rapidly into view. However, such an abrupt change will nonetheless be a good place to add or remove a

session layer and thus the term scene change is meant to cover such large changes in frame content from one frame to another.

Figure 6 illustrates a system which detects congestion in the network and also imposes the extra condition of adding or dropping a layer upon a scene change. The server 601 comprises the components of the media server 202 of Figure 1 and the unit 603 comprises the components of the client 204 also shown in Figure 1. Video data from the audiovisual source is also passed to a content analyser 604 which segments the video sequence into scenes. The segmented video is then passed to an objective quality assessment module (OQAM) 605 which assesses the contribution to the perceived quality of a scene that each component layer provides.

These results are then passed to an inter layer adaptation module (ILAM) 607. The function of the ILAM 607 is to continuously calculate the number of layers that maximise the perceived quality for the session. The ILAM 607 also receives input from a sender congestion manager 608 which reports on the bandwidth available to the session on the network 203. The sender congestion manager 608 receives feedback from a client congestion manager 609 on the number of packets that have been received. If this matches the number of packets that were sent then the bandwidth is known to be the current transmission rate. If packets are being lost then the bandwidth available to the session is less than the transmission rate, in which case the sender congestion manager 608 informs the ILAM 607 that a layer should be dropped.

In order to select which layer should be dropped the ILAM 607 couples the bandwidth required by a layer with its contribution to the quality of the complete image, as calculated by the OQAM 605. The ILAM 607 performs an exhaustive search on all of the bandwidth/quality values. When the ILAM 607 has selected which layer is to be dropped from a particular scene, the timing of the drop is preferentially set to coincide with the transmission of that particular scene. In this way the drop in quality occurs with the scene change and is thus much less noticeable to a viewer than if the quality change had occurred during a scene. In the case that no packets have been lost for a preset period of time the sender congestion manager 608 will request the ILAM 607 to add in a layer to test

whether all of the available bandwidth is being employed. If no packets are lost during this experiment then the newly added layer is maintained. This process of experimentation with adding layers is continued until a significant proportion of packets are lost, in which instance the system can be confident that all of the available bandwidth is being employed. Again, the timing of the introduction of a layer is set to occur preferentially as the scene changes.

Under a given bit-rate allocation the levels of perceived quality do not change considerably within a scene, but scene cuts cause considerable changes in perceived quality, especially when the content features (spatial and motion energy) change a lot between subsequent scenes. As a consequence there will also be a significant difference in the corresponding quality scores for those successive scenes which may justify a rescheduling of the number of layers in the stream.

The invention is not limited in use to hierarchical encoding schemes. The invention is suitable for use in any encoding technique where the adaptation of transmission bit-rate to accommodate fluctuations in network capacity occurs. For instance the invention may be applied to a transcoding system where encoded data streams are transcoded from a high bit-rate to a low bit-rate or from a low bit-rate to a high bit-rate. The present invention would be suitable to reduce the impact for the viewer as the output bit-rate shifts in response to network conditions by timing the transition to occur upon a scene change in the encoded video sequence. Another example of an adaptive video streaming technique to which the invention may be applied is where multiple independent video streams of different bit-rates are transmitted. In this case the client chooses which stream to accept based on session bandwidth. The client may switch from one stream to another as bandwidth fluctuates; the present invention ensures that the switch is timed to coincide with a scene change in the encoded video stream.

Naturally, there may be times when a change of scene is such a long way away that it is advantageous to switch from one bit-rate to another other than during a change of scene. For example, consider the case where multiple independent video streams at different bit-rates are available for transmission by a media server to a client as, for example, described in co-pending European patent

application No. 00310594.7 the contents of which are hereby incorporated herein by way of reference. In such a case, the server may be capable of transmitting a first stream at a bit-rate of 500kbit/s and a second, higher quality stream at a bit-rate of 1500kbit/s. The client may initially request that the server transmit the first stream at a transmission rate of 1000kbit/s. If the network is not congested and all of the packets transmitted are successfully received by the client, the receive buffer at the client will start to fill with data at a rate of 500kbit/s, since the client will only be removing data from the buffer at a rate of 500kbit/s. After say 10 seconds of filling up the buffer at this rate, the client will have a buffer of ten seconds worth of data at which point it may decide it can attempt to receive the higher bit-rate second stream of video data from the server and thus sends an appropriate request to the server to this effect. If, however, the server is aware that the client has a receive buffer of 5 Mbytes size, it knows that it may continue sending data from the first stream at the rate of 1000kbit/s for at least another 150 seconds before the receive buffer overflows, causing problems for the client. Therefore, the server may attempt to wait for a specified period to see if a scene change occurs during this interval. If so, a switch to the second higher bit-rate signal is not made until the change of scene occurs. Of course, if there is no change of scene within the determined period, the server switches to the higher rate anyway. In this example, a period of only ten seconds is deemed appropriate as the waiting time.

Note that instead of simply waiting to see if a change of scene occurs in the specified period and switching at the end of the period if no such change of scene is detected, an alternative method would be to pre-analyse the video to be sent (clearly this only applies to pre-stored video data and not live video) and to note when changes of scene occur. In such a case, the server, upon receipt of a request from the client to switch streams, could search to see if a suitable change of scene will occur within the predetermined period and if not to switch immediately to the new bit stream.

In the present example, upon switching to the higher rate bit stream, the client may have requested a transmission rate of 1500kbit/s corresponding to the rate at which the data will be drawn from the receive buffer by the client. In such a case, the buffer size of 10 seconds should remain constant so long as all of the

transmitted packets are successfully received by the client. However, in the event of congestion on the network, a proportion of packets may fail to arrive at the client. In such a case, the server will be warned of this via the RTCP. If the congestion is sufficiently severe, the server may deduce that the buffer is in danger of emptying which would cause a break in the video displayed by the client to occur. To prevent this, the server may switch back to the lower bit rate stream. Via the notification of how many packets are being lost, the server can deduce how long it will be before the buffer is emptied. For this time, the server can wait to see if a change of scene occurs, and if so, the new stream will be switched to at that point. Note that it would also be possible for the client to simply request that the new stream be switched to upon detecting that its receive buffer is emptying at an unsustainable rate.

Note that the amount of data in the buffer is actually of less significance than the rate of change of the amount of data in the buffer. Thus it is preferably this quantity which either the server or the client measures in order to determine whether to change from one bit stream to another (or whether to add or drop a layer in the first example).

CLAIMS

1. A method of operating a multimedia server (202), said method comprising;

providing a stream of video data representing a video sequence to an output of the multimedia server (202), wherein the output of the media server (202) is connected to a packet based network (203),

measuring a property of the video data in order to determine the occurrence of a scene change in the video sequence,

detecting the available bandwidth on the network (203),

varying the bit-rate of the stream of video data,

wherein the method is characterised in that

variation in the bit-rate of the video data is controlled to occur in response to variations in the capacity of the network (203) and preferentially with a scene change in the video sequence.

2. A method of operating a multimedia server (202) in accordance with claim 1 wherein upon detecting a decrease in the available bandwidth the bit-rate of the video data is decreased.

3. A method of operating a multimedia server (202) in accordance with claim 1 wherein the bit-rate of the video data is periodically increased in order to ensure that all of the available bandwidth is being used.

4. A method of operating a multimedia server 202 in accordance with claims 1 to 3 wherein said stream of video data is comprised of a plurality of hierarchically encoded layers.

5. A method of operating a multimedia server (202) in accordance with claim 4 wherein the bit-rate of the video data is varied by adding or dropping layers from the stream of data transferred to the output port.

6. A method of operating a multimedia server (202) in accordance with claims 4 or 5 wherein the objective quality contribution of each layer to the

overall quality of the video sequence is measured in order to determine which of the layers is to be added or dropped from the stream of data transferred to the output port.

7. A method of operating a multimedia server 202 in accordance with claims 1 to 3 wherein said stream of video data is comprised of a plurality of independently encoded flows each of which encoded at different bit-rates.

8. A method of operating a multimedia server (202) in accordance with claim 7 wherein the bit-rate of the video data is varied by switching between independently encoded flows.

9. A method of operating a multimedia server 202 in accordance with any of the preceding claims in which the source 201 of video data is stored in encoded form in a database.

10. A method of operating a multimedia server 202 in accordance with any of claims 1 to 8 in which the source 201 of video data is encoded in real-time.

11. A method of operating a multimedia server 202 in accordance with any of the preceding claims wherein, upon detecting a variation in the capacity of the network such as to render it appropriate for a change in bit rate to occur, the server waits for a predetermined time to attempt to detect a change of scene, and if such a change of scene is detected within the predetermined time, the change of bit rate is caused to occur at the time of the change of scene, but if a change of scene is not detected within the predetermined time, then the change of bit rate is caused to occur at the expiry of the predetermined time.

12. A method of operating a multimedia server 202 in accordance with any of the preceding claims wherein, upon detecting a variation in the capacity of the network such as to render it appropriate for a change in bit rate to occur, the server determines if there is going to be a change of scene within a predetermined time, and if it is determined that there will be such a change of scene within the predetermined time, the change of bit rate is caused to occur at the time of the change of scene, but if it is determined that there is not going to

be a change of scene within the predetermined time, then the change of bit rate is caused to occur substantially immediately after having made such a determination.

13. The method of claim 12 further including the step of, prior to providing a stream of video data to the output of the media server, processing the video data to determine the positions within the video data in which changes of scenes occur and making this information available for permitting whether a scene change will occur within a predetermined period at any point along the stream of video data.

14. A multimedia server (202) comprising;

a reader (302) for reading video data representing a video sequence from a source (301),

a scene change detector (604) for detecting changes of scene in the video sequence,

a bit-rate controller (607) for controlling the rate at which the video data is transferred from the reader to an output port, wherein the output port is capable of communicating with a client (204) on a packet network (203);

the media server further comprising means (608) for detecting available bandwidth between the multimedia server (202) and client (204) on the packet network,

characterised in that

the bit-rate controller (607) is arranged to vary the bit-rate of the video data transferred to the output port in dependence on the detected bandwidth preferentially in correlation with a scene change detected by the scene change detector (604).

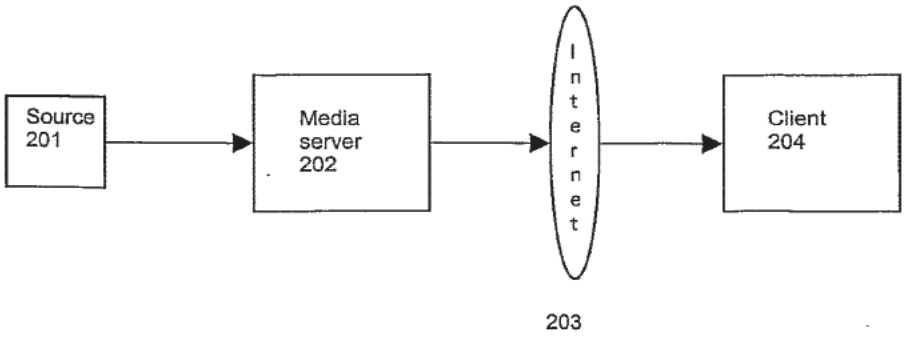


Figure 1

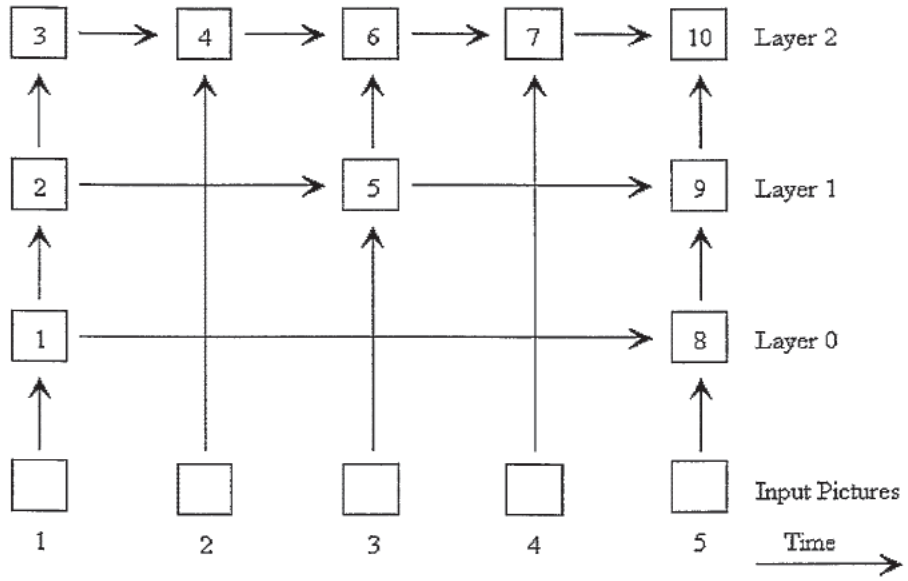


Figure 2

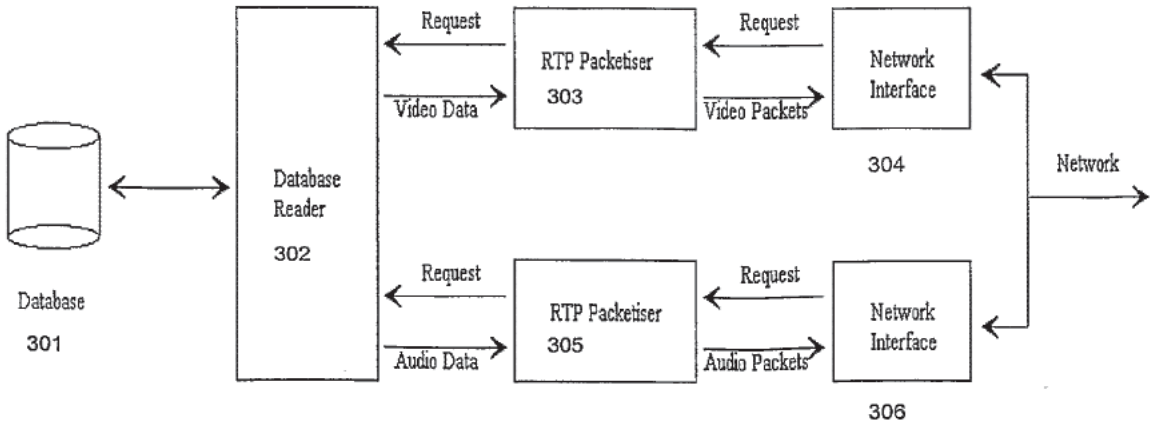


Figure 3

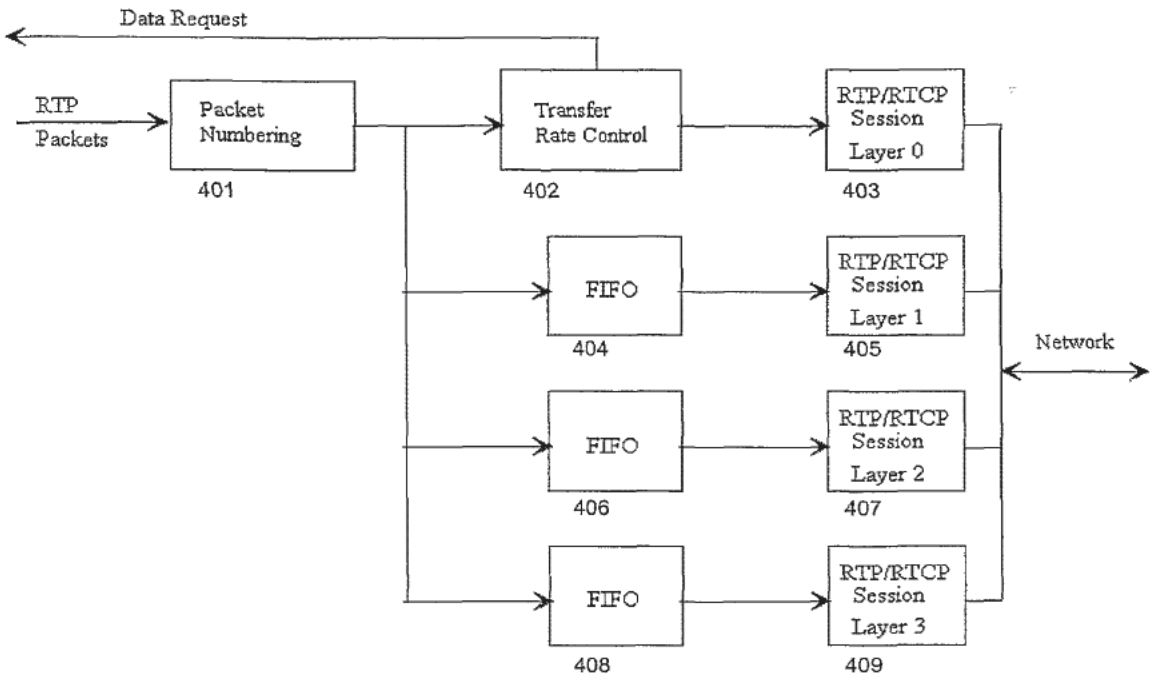


Figure 4

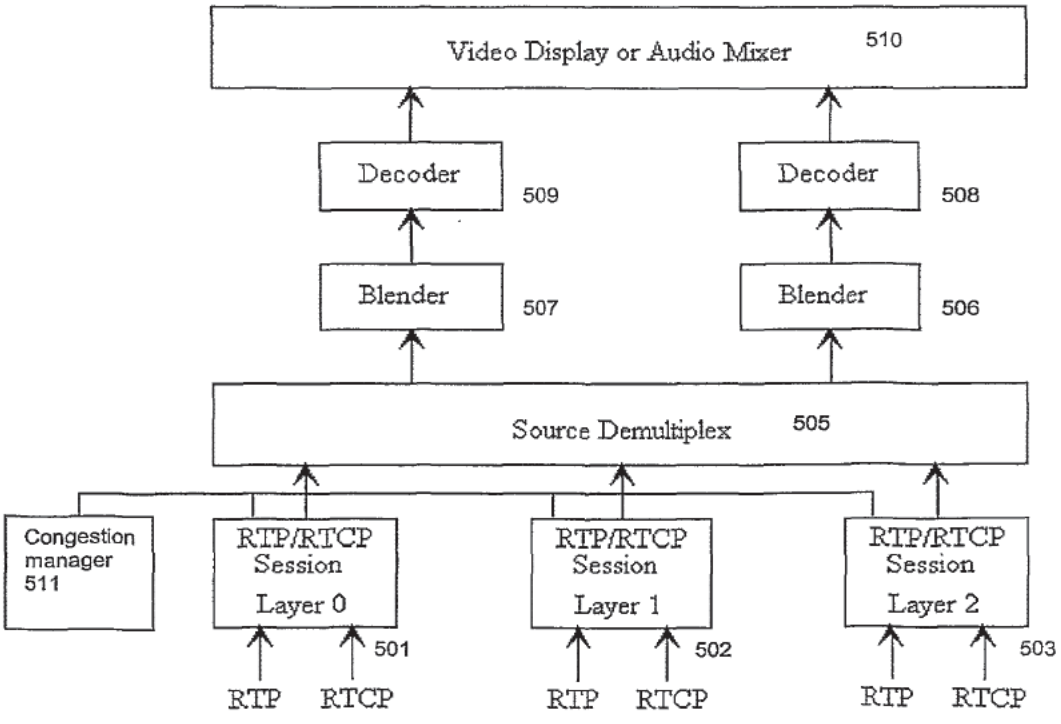


Figure 5

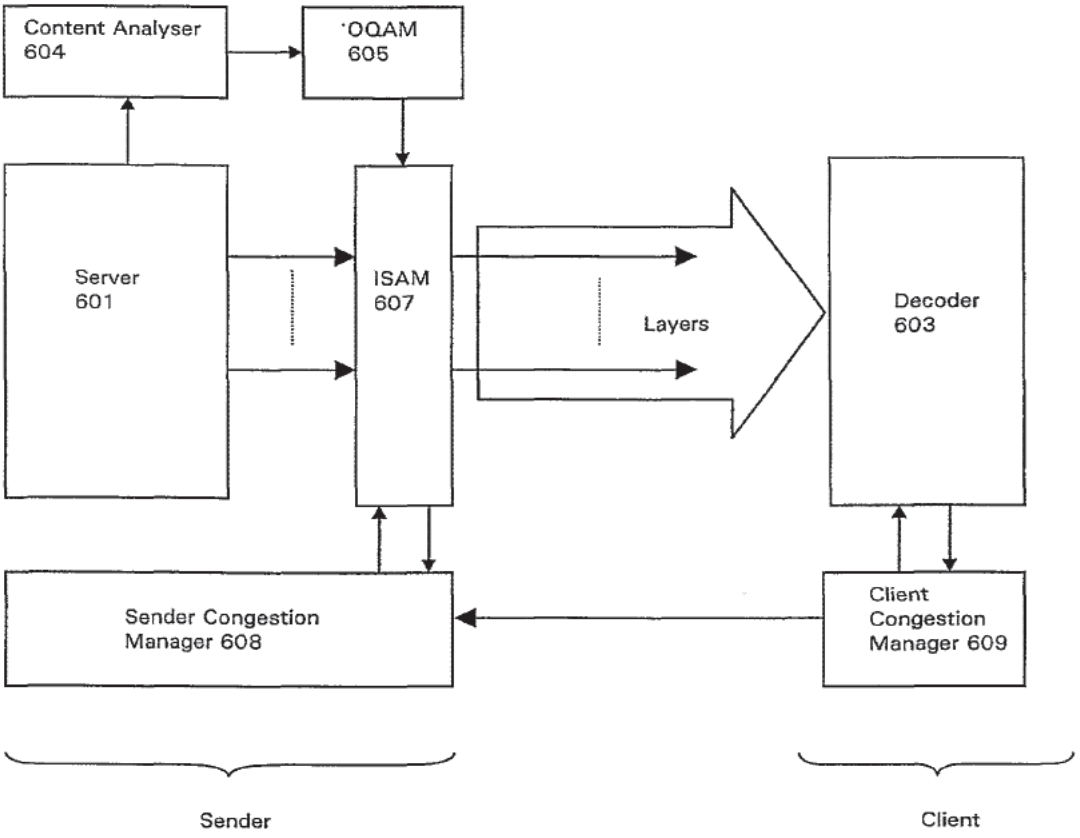

Content Analyser
604
OQAM
605
Server
601
ISAM
607
Layers
Decoder
603
Sender Congestion
Manager 608
Client
Congestion
Manager 609
Sender
Client

Figure 6

INTERNATIONAL SEARCH REPORT

tional Application No
PCT/GB 02/03307

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 H04N5/00 H04N7/24

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)
IPC 7 H04N H04L G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, PAJ, IBM-TDB

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 6 014 694 A (AHARONI AMIR ET AL) 11 January 2000 (2000-01-11)	1-10,14
Y	abstract column 2, line 29 - line 65 column 3, line 62 -column 4, line 34 column 8, line 54 -column 9, line 35 column 10, line 34 - line 49 column 14, line 34 -column 15, line 6	11-13
X	ANONYMOUS: "Method to Deliver Scalable Video across a Distributed Computer System" IBM TECHNICAL DISCLOSURE BULLETIN, IBM CORP. NEW YORK, US, vol. 5, no. 37, 1 May 1994 (1994-05-01), pages 251-256, XP002079392 ISSN: 0018-8689	1-10,14
A	the whole document	11-13

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

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

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

Date of the actual completion of the international search	Date of mailing of the international search report
16 September 2002	23/09/2002

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Fantini, F

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
Y	US 6 038 000 A (HURST JR ROBERT NORMAN) 14 March 2000 (2000-03-14)	11-13
A	column 2, line 52 -column 3, line 22 column 4, line 20 - line 49 column 6, line 9 - line 45 column 7, line 1 - line 24 column 13, line 46 -column 14, line 18 column 20, line 13 - line 30	1,14
A	EP 1 079 631 A (SONY UK LTD) 28 February 2001 (2001-02-28) page 3, line 1 - line 4	1-14
A	US 5 990 945 A (DENTENEER THEODORUS J J ET AL) 23 November 1999 (1999-11-23) abstract column 3, line 13 - line 35 column 4, line 42 - line 44 column 5, line 37 - line 39	11-13

INTERNATIONAL SEARCH REPORT

Information on patent family members

ational Application No
PCT/GB 02/03307

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 6014694 A	11-01-2000	AU 7928998 A	19-01-1999
		EP 0992160 A1	12-04-2000
		WO 9900984 A1	07-01-1999
		JP 2002511216 T	09-04-2002
US 6038000 A	14-03-2000	US 6137834 A	24-10-2000
		EP 0954924 A1	10-11-1999
		JP 2001509354 T	10-07-2001
		WO 9832281 A1	23-07-1998
		EP 0901718 A1	17-03-1999
		JP 2002503402 T	29-01-2002
		KR 2000016034 A	25-03-2000
		WO 9745965 A1	04-12-1997
EP 1079631 A	28-02-2001	GB 2353653 A	28-02-2001
		EP 1079631 A1	28-02-2001
		JP 2001119305 A	27-04-2001
US 5990945 A	23-11-1999	CN 1199532 A	18-11-1998
		EP 0860087 A1	26-08-1998
		WO 9809442 A1	05-03-1998
		JP 11514825 T	14-12-1999

(12) UK Patent Application (19) GB (11) 2 367 219 (13) A

(43) Date of A Publication 27.03.2002

(21) Application No 0023062.3

(22) Date of Filing 20.09.2000

(71) Applicant(s)
Vintage Global
(Incorporated in the British Virgin Islands)
c/o Ernst & Young Trust Corp (BVI) Ltd, P.O Box 3340, Road Town, Tortola, British Virgin Islands

(72) Inventor(s)
Benjamin Neeman
Yossi Rubner
Nehemia Davidson
Uri Salomon

(74) Agent and/or Address for Service
Geoffrey L Melnick
c/o Solo Photographic, 12 Catherine Street, ST ALBANS, Herts, AL3 5BX, United Kingdom

(51) INT CL[7]
H03M 7/30 , H04L 1/00

(52) UK CL (Edition T)
H4P PDCFM

(56) Documents Cited
EP 1049074 A1 **WO 99/00984 A1**
WO 98/09387 A1 **WO 00/72601 A2**
US 5847760 A

(58) Field of Search
UK CL (Edition S) **H4P PDCFM PDCFP PDCFX PF**
INT CL[7] **H03M 7/30 , H04L 1/00 , H04N 7/26 7/32**
Online: WPI, EPODOC, JAPIO

(54) Abstract Title
Streaming of media file data over a dynamically variable bandwidth channel

(57) A method of transmitting data streams over a variable bandwidth channel, comprising modifying the quantity of data to be injected into the channel to accord with a measured capacity or bandwidth of the channel. An input stream is passed through a plurality of parallel compressors 40.1-40.4 which each produce an output stream of a different rate to the other streams. Each stream preferably comprises key or reference frames and delta frames. The reference frames are synchronized between the streams so that a data unit in one stream contains the same information as the corresponding data unit in another stream, but is compressed to a different extent. A data stream for transmission along the channel is produced by selecting data units from more than one compressor output stream so as to produce a stream which makes optimum use of the available channel bandwidth. In one embodiment the duration of the data units are one of a group comprising fixed predetermined durations and duration determined by changes in the channel bandwidth. The method is useful in sending multimedia data streams such as audio and video over a channel such as a cellular telephone network or the internet.

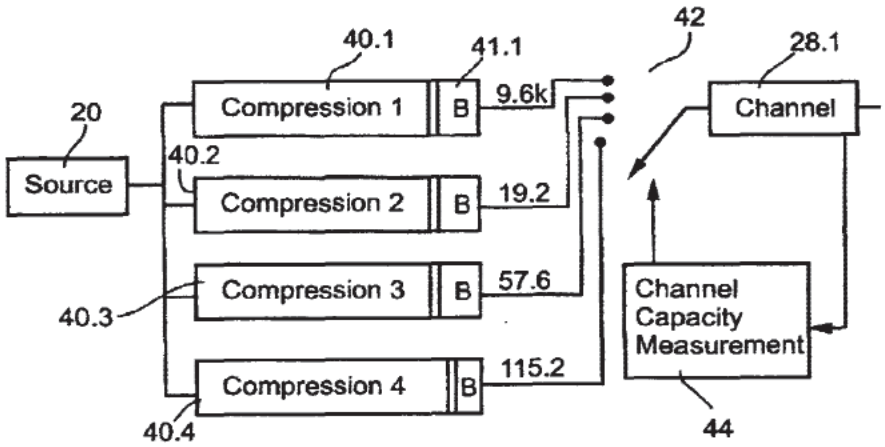


Fig. 4

At least one drawing originally filed was informal and the print reproduced here is taken from a later filed formal copy.

GB 2 367 219 A

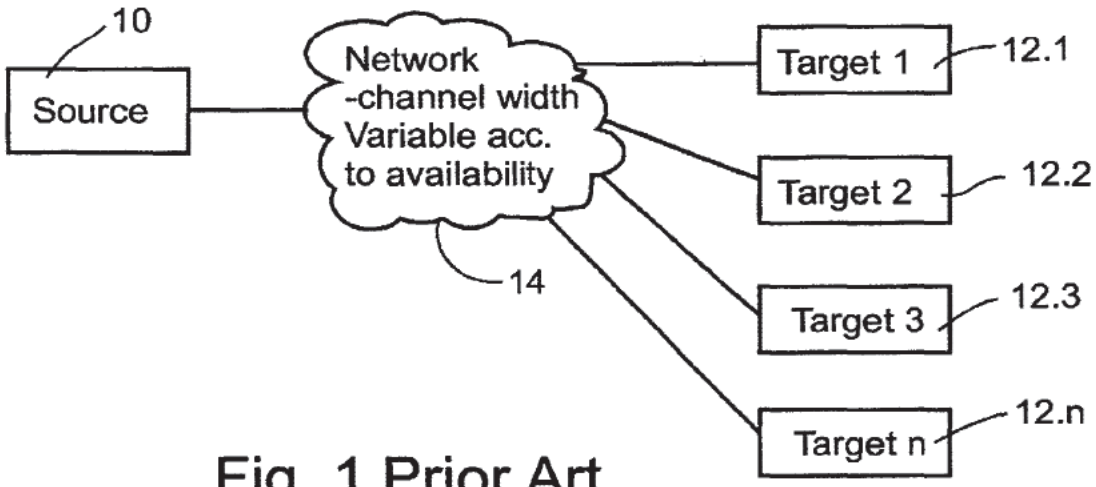


Fig. 1 Prior Art

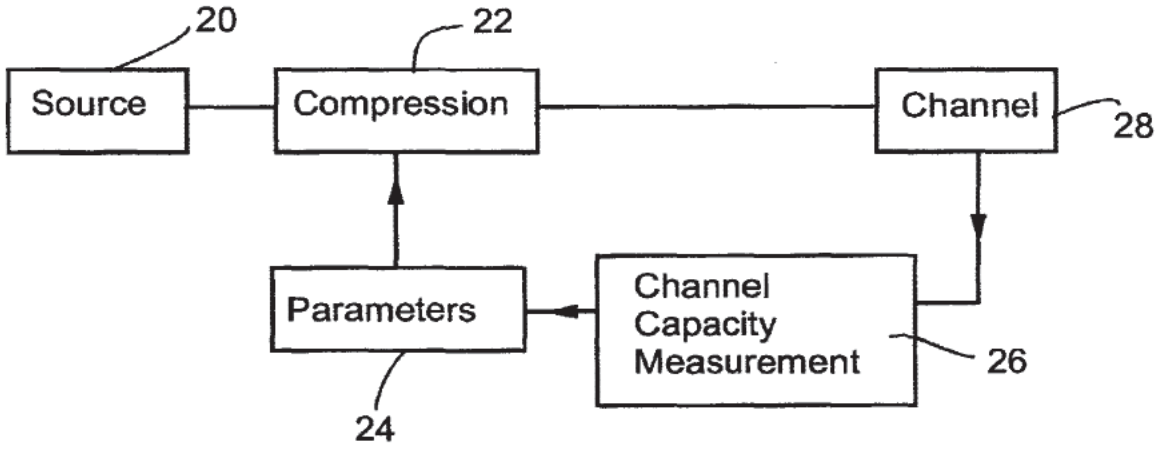


Fig. 2

Ex. 1018-1081

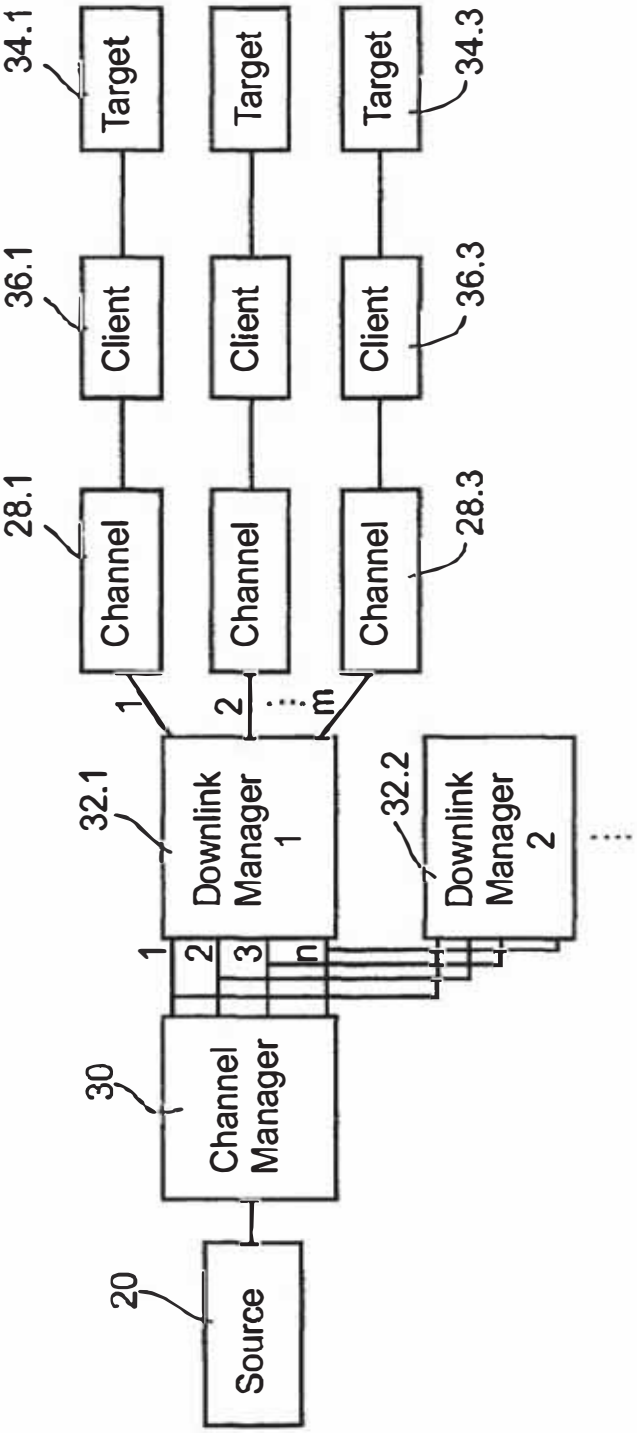


Fig. 3

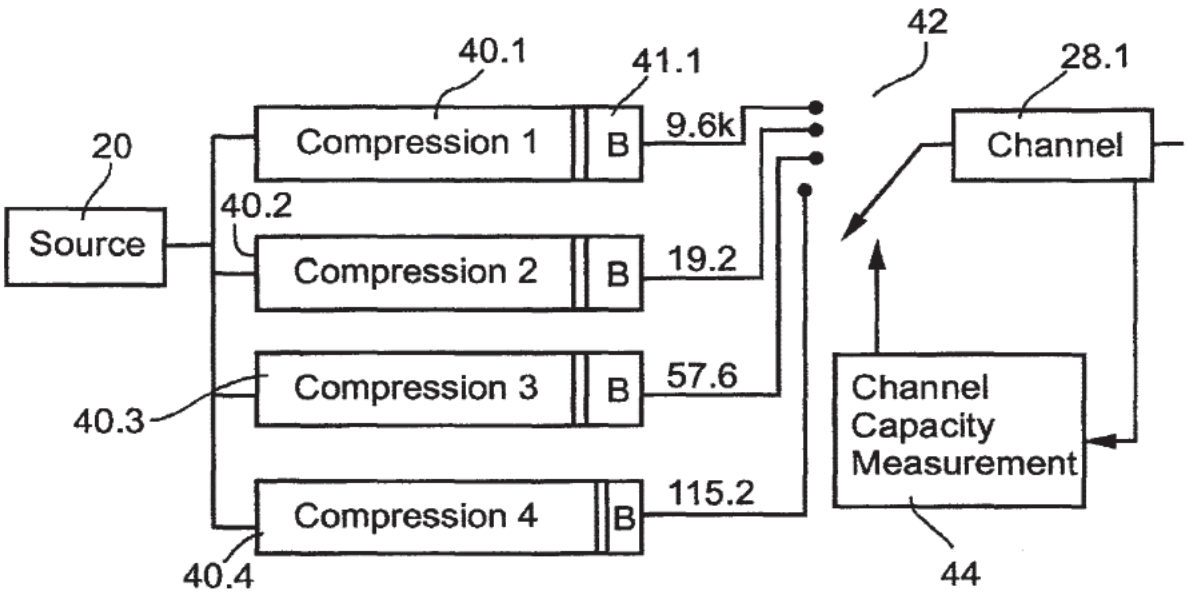

42
40.1
41.1
28.1
Compression 1
B
9.6k
Channel
20
Source
40.2
Compression 2
B
19.2
Compression 3
B
57.6
40.3
Channel
Capacity
Measurement
Compression 4
B
115.2
40.4
44

Fig. 4

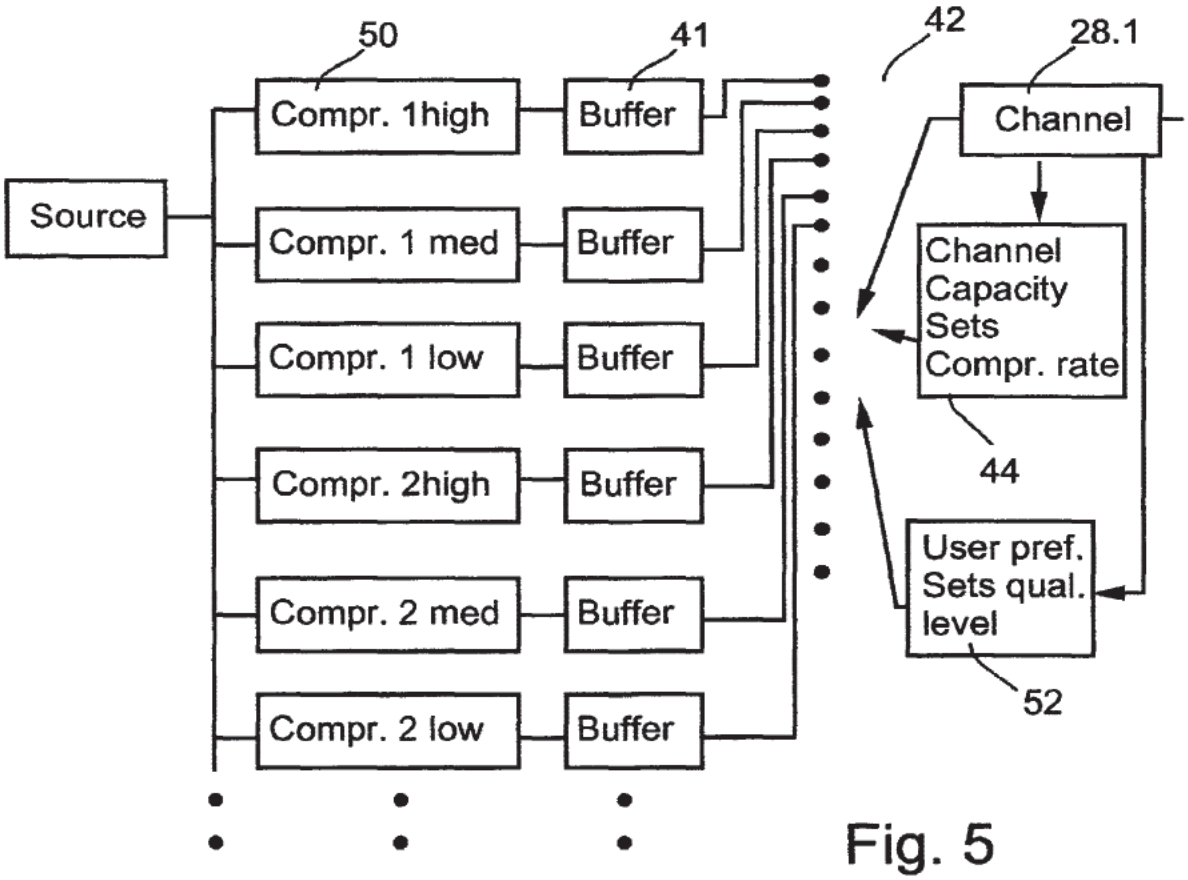

50
41
42
28.1
Compr. 1high
Buffer
Channel
Source
Compr. 1 med
Buffer
Channel
Capacity
Sets
Compr. rate
Compr. 1 low
Buffer
44
Compr. 2high
Buffer
User pref.
Sets qual.
level
Compr. 2 med
Buffer
52
Compr. 2 low
Buffer

Fig. 5

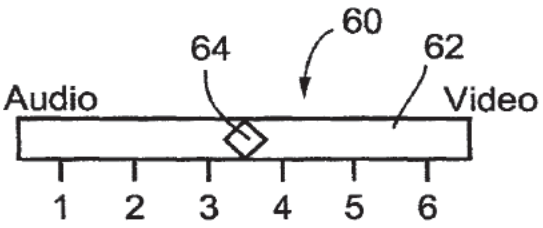


Fig. 6

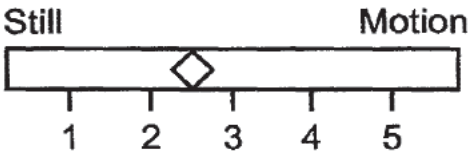


Fig. 7

00/20757 -UK. 9/12/00

Title of the Invention

Streaming of Media File Data Over a Dynamically Variable Bandwidth Channel

Field of the Invention

The present invention relates to streaming of media file data over a channel having a dynamically variable bandwidth, and more particularly but not exclusively to streaming of multimedia data such as video and /or audio data over a channel having a dynamically variable bandwidth.

Background of the Invention

Dynamically varying channel bandwidths are particularly encountered over the Internet and like networks and over radio and cellular telephone links. On the Internet the total bandwidth is fixed and the bandwidth available per channel tends to depend on the amount of usage, the greater the usage the smaller the bandwidth available per user. In radio and cellular links the available bandwidth depends on usage, but the overall capacity may also be dependent on environmental factors.

The Internet relies on packet switched communication, whereas radio and cellular telephone links may be either circuit-switched or packet switched. Either way, the same considerations of channel bandwidth broadly apply.

As the bandwidth changes, a number of approaches are currently in use. One of these is simply to buffer received video frames, to start to show the frames as soon as the buffer is full, and to continue to show the video until the buffer is empty, then to halt the display until the buffer is full again. Another method is to show the video at a real time rate, displaying all frames that arrive in time and dropping all frames that arrive too late.

In both of these methods, data is sent at the same high rate. In the first method, unacceptable delays in showing the data are experienced. In the second method, valuable channel bandwidth is wasted transmitting data that does not get used.

Whether the network is a cellular telephone network or the Internet or like wide area data network, the bandwidth of the data network is variable. In all cases the

Ex. 1018-1084

larger the number of users the lower the bandwidth per user, and radio-based networks may also suffer from reduced bandwidth due to environmental conditions. Furthermore, in the case of cellular telephone networks, events such as handover between two base stations may also affect the available bandwidth.

In general, cellular networks are able to absorb changes in bandwidth without the user noticing a change in the quality of the voice signal. However, it is desirable to go beyond mere voice communication over cellular telephone networks. Already, low level data capability is being built into mobile telephones and it is envisaged that within the next few years, mobile telephones will be able to receive and display multimedia data. Multimedia data uses large amounts of bandwidth and generally requires to be displayed in real time. Thus bandwidth fluctuations will be immediately noticeable. Currently, however, transmission of video data over cellular phone channels is not being attempted on a commercial basis.

At present, mobile telephone systems are on their second generation and a third generation is on the drawing board but will not be available for some years. The problem is particularly associated with the design of advanced second generation equipment (generation 2.5) which, it is envisaged will be able to provide multimedia within the limitations of second generation capacity. It should be noted however that the same considerations apply to further generations of mobile telephones and also to other environments in which it is desired to transmit multimedia data, such as the Internet.

Generally, multimedia data is sent over a network as a datastream in compressed form. There are numerous algorithms for compressing data, and different forms of compression are appropriate for compressing different forms of data. Furthermore, a single form of compression may give rise to different levels of compression according to compression parameters that may be set. The greater the compression the lower will be the resulting data rate and a high level of compression may be used to overcome channel capacity constraints. A solution may involve selection of the compression parameters to meet a likely worst-case scenario. However, the solution is not complete as a high level of compression leads to low quality output and no advantage is taken of the availability for most of the transmission time of a higher bandwidth availability.

US Patent No. 6,014,694, Aharoni et al, discloses a system for adaptively transporting video over networks wherein the bandwidth varies with time. The

Ex. 1018-1085

system comprises a video/audio codec that functions to compress, encode, decode and decompress video streams that are transmitted over networks having available bandwidths that vary with time and location. Depending on the channel bandwidth, the system adjusts the compression ration to accommodate a plurality of bandwidths ranging from 20kbps for POTS to several Mbps for switched LAN and ATM environments. Bandwidth adjustability is provided by offering a trade-off between video resolution, frame rate and individual frame quality. The system generates a video data stream comprising of key, P and B frames from a raw video source. Each frame type is further comprised of multiple levels of data representing varying degrees of quality. In addition, several video server platforms can be utilized in tandem to transmit video/audio information with each video server platform transmitting information for a single compression/resolution level.

P and B frames are delta frames that depend on previous frames to allow a complete picture to be generated. Thus it is not possible simply to mix groups of frames of different quality levels since the deltas between frames are different. The above US Patent solves the problem by taking advantage of what is known as a group of pictures (GOP) which is a self-contained unit such as a scene. Because they are self-contained, GOPs of different compression levels can be successfully mixed. However, GOP boundaries are content related and thus irregular. When GOP boundaries are widely spaced, dynamic following of bandwidth changes are highly inaccurate or generate delays for new users.

Ex. 1018-1086

Summary of the Invention

Embodiments of the present invention attempt to solve the above problems by modifying the quantity of data to be injected into the channel to accord with a measured capacity or bandwidth of the channel. The quantity of data is preferably modified only by altering compression parameters, so that data fully usable in real time can be received from the channel. Thus a reduction in bandwidth of the channel leads to a drop in quality of the received signal and not to dropped frames or a reduction in performance.

According to a first aspect of the present invention there is thus a device for streaming of media file data to be sent over at least one channel having a dynamically variable bandwidth, comprising

an input for receiving media file data at a first rate,

a multiple stream creator for creating a plurality of streams of said media file data, each stream compressed to predetermined other rates, each stream comprising a series of autonomous compressed data units which are mutually compatible, and

a stream selector for dynamically matching a bandwidth of the channel to rates of said streams to send said data file as a combination of said autonomous units from said streams via said channel,

wherein said autonomous units have durations which are one of a group comprising fixed predetermined durations and durations determined by changes in said channel bandwidth.

Preferably, said stream selector is operable to select a combination of said streams in succession said combination being chosen to produce an average stream rate over a period of time which matches said channel bandwidth.

Preferably, said bandwidth of said channel is liable to vary dynamically and wherein said stream selector is operable to repeat said matching throughout said streaming operation.

Preferably, said bandwidth of said channel is liable to vary dynamically and wherein said stream selector is operable to repeat said matching throughout said streaming operation.

Preferably, said autonomous units match each other across said plurality of streams, thereby to enable switching between said streams.

Ex. 1018-1087

Preferably, said data file is arranged in said streams in a compressed media format utilizing reference frames and delta frames and wherein each autonomous unit is self-contained in respect of reference frames.

In an embodiment, said channel is a cellular telephone channel. Alternatively, said channel is an Internet channel.

Preferably, the predetermined other rate is different for each stream. In another embodiment, the predetermined other rate is identical for a set of said streams, each stream in said set being assigned a different balance between quality and speed of said media.

A preferred device comprises a plurality of media types, said predetermined other rate being identical for a set of said streams, each stream in said set being assigned a different balance between quality of said different media types.

Preferably, streams in said set are selectable by a signal received from said channel.

Preferably, streams in said set are selectable by a signal received from said channel.

An embodiment comprises a detector for detecting a change in the bandwidth of said at least one channel and operable to request a corresponding change in the data rate, and wherein the device is operable, upon receipt of said request, to set a delay, and to issue a reference frame within said delay.

The device preferably issues said reference frame at the end of said delay.

The device alternatively issues said reference frame at the occurrence of a largest detected interframe change within the time delay.

According to a second aspect of the invention there is provided a multiple stream creator for receiving a data stream at a first data rate and converting said data stream into a plurality of data streams at a plurality of data rates, said creator comprising an input for receiving said data stream at said first data rate, and a plurality of outputs, each operable to receive said data stream at said first rate and to output said data stream at a predetermined rate individually set for each output as a series of autonomous units having a duration, and wherein said duration is one of a group comprising a fixed predetermined duration and a duration dynamically changeable in accordance with bandwidth conditions in a transmission channel.

Preferably said outputs each comprise data buffers.

Ex. 1018-1088

The multiple stream creator preferably comprises a data stream processor for processing said data stream received at said first rate into said plurality of autonomous units, said autonomous units being synchronized blocks, thereby allowing said output streams to be mixed in said transmission channel to produce continuously usable output.

Preferably, said received data stream is in an uncompressed data format and said data is compressed into a compressed data format independently for each one of said output data streams.

According to a third aspect of the invention there is provided a method of sending a datastream over a variable bandwidth channel comprising modifying a quantity of data to be injected into the channel to accord with a measured capacity of the channel, said method comprising the steps of

compressing the datastream, using an algorithm having compression parameters, to produce a plurality of compressed datastreams each comprising reference frames and delta frames, each compressed datastream produced by using the algorithm together with different parameters, and

synchronizing reference frames between the compressed datastreams so that combinations of said compressed datastreams can be sent via said channel.

Preferably, said step of synchronizing reference frames is carried out at intervals determined by a dynamic status of said channel bandwidth.

Alternatively or additionally, said step of synchronizing reference frames is carried out at predetermined intervals.

Alternatively or additionally, said combination of said compressed datastreams is selected substantially dynamically to match the measured capacity of the channel.

Preferably, a change in bandwidth of a channel causes a request for commencement of a new autonomous unit, and wherein said request causes initiation of measurement of a time delay, and wherein a new autonomous unit is commenced within said time delay.

Preferably, the new autonomous unit is commenced at the end of said time delay.

Preferably, the step is provided of determining the largest delta frame during the time delay and sending a reference frame, to initiate said new autonomous unit, in place of said determined largest delta frame.

Ex. 1018-1089

According to a fourth aspect of the present invention there is provided a tool for controlling receipt of media file data over a network, wherein said media file data is available at a plurality of quality levels, the tool comprising a quality selector to enable a user to select between said different quality levels for receipt.

Preferably, said media file comprises a plurality of data groups and each quality level is a different balance between quality levels of the plurality of data groups.

Preferably, said quality selector comprises a slider operable to be moved along a scale, the position of the slider on the scale indicating a quality level.

Brief Description of the Drawings

For a better understanding of the invention and to show how the same may be carried into effect, reference will now be made, purely by way of example, to the accompanying drawings, in which:

Fig. 1 is a generalized diagram showing a data source sending data to a series of targets over a variable bandwidth network,

Fig. 2 is a generalized block diagram showing a first embodiment of the present invention in which feedback concerning available bandwidth within a channel is used to modify parameters governing compression of the datastream,

Fig. 3 is a generalized block diagram showing a second embodiment of the present invention in which a series of compressed datastreams are continually created and are selected for use in individual channels according to measured parameters of the channel,

Fig. 4 is a generalized block diagram showing the channel manager and part of a downlink manager of Fig. 3 in greater detail,

Fig. 5 is a generalized block diagram showing a third embodiment of the present invention in which the user is able to set a quality level of the data that is being sent,

Fig. 6 is a generalized schematic diagram of a quality level selection tool for use by a user in selecting a quality level in the embodiment of Fig. 5, and

Fig. 7 is a generalized schematic diagram of another quality level selection tool for use with the device of Fig. 5.

Ex. 1018-1090

Description of the Preferred Embodiments

Reference is now made to Fig. 1, which is a generalized diagram showing a data source 10 sending data to a series of targets 12.1..12.n over a variable bandwidth network 14. The network may be a cellular telephone network or the Internet or like wide area data network or a combination. In both cases, the bandwidth of the data network is variable. In all cases the larger the number of users the lower the bandwidth per user, and radio-based networks may also suffer from reduced bandwidth due to environmental conditions. Furthermore, in the case of cellular telephone networks, events such as handover between two base stations may also affect the available bandwidth.

Data is generally compressed before sending over the channel. The level of compression may be preselected to accord with known or likely channel capabilities, as described in the introduction. However such a solution has to concentrate on the lower margins of available bandwidth and thus fails to make full use of higher bandwidth levels when available. Conversely, solutions which send datastreams at higher data rates and do not take the lower bandwidth levels into account lead to unacceptable breaks and delays in performance at the receiving end.

Reference is now made to Fig. 2 which is a generalized block diagram showing a first embodiment of the present invention in which feedback concerning available bandwidth within a channel is used to modify parameters governing compression of a datastream.

A multimedia datastream is received from a source 20 such as a host or server. The datastream may be a video datastream and may be in uncompressed or barely compressed format, for example AVI or MPEG1. The datastream is passed to a compression unit 22 which compresses the data into a very low bit-rate stream using compression parameters. A typical format that may be used for the very low bit-rate stream is MPEG 4 which uses a system based on reference frames and delta frames to compress the image part of the video data. H.263 is another typical format that may be used. A video frame is selected as a reference frame and encoded in its entirety. Then a series of subsequent frames are selected, but for each of these frames, the only data that is encoded is that which has changed with respect to the previous frame. The encoded changes form what is known as the delta frame.

Whichever of the available compression formats are used, compression is governed by a series of parameters 24. This allows the compression level to be raised

Ex. 1018-1091

or lowered depending on the data rate required and the quality of picture desired. A greater level of compression gives a lower quality of data at a lower data rate and a lower level of compression gives a higher quality of data at a higher data rate.

In the embodiment of Fig. 2 a channel capacity measurement 26 is received from a channel 28 over which the transmission occurs. The measurement may be received from a receiving client who simply reports the received data rate. If the received data rate is reported to be lower than the rate at which data is being sent then it is assumed that data is being produced too fast for the channel and the parameters 24 are altered to increase the compression level until the data output rate is the same as the channel rate. Conversely, if the data rate rises at the receiving end, the parameters 24 may be altered to reduce the compression level and increase the output data rate.

Reference is now made to Fig. 3, which is a generalized block diagram showing a second embodiment of the present invention in which a series of compressed datastreams are continually created and are selected for use in individual channels according to measured parameters of the channel. Parts that are the same as those mentioned in previous embodiments are given the same reference numerals and are not referred to again except as necessary for an understanding of the present embodiment.

In the embodiment of Fig. 3, an input datastream is received from a source 20 as before. The stream is passed to a channel manager 30, which compresses the input datastream. However, instead of compressing it once only, it compresses it into a plurality of output data streams, each one using different compression parameters so as to produce a plurality of streams each having a different data rate. The various streams are then passed to one or more downlink managers 32.1...32.n, each of which is able to manage a limited number of channels 28 for sending output data to a limited number of targets 34. Each target 34 is associated with a client 36, typically client software located at the target 34, which is able to send back, to the downlink manager 32, feedback data such as the received channel rate.

The downlink manager 32 simply obtains the received channel rate from the client 36 and mixes the various streams from the channel manager to produce an average stream rate which matches the channel rate as closely as possible. For example if the received channel rate is 39.5 kbits/sec and the downlink manager has available a 35 kbits/sec stream and a 45 kbits/sec stream then the downlink manager

Ex. 1018-1092

will attempt to mix the two streams in suitable proportions to match the 39.5 kbits/sec channel. Thus the target 36 may receive a continuous stream of data, the quality of the channel influencing only the quality of the data and not the continuity.

It will be recalled that the compression systems in use include reference and delta frame style compression. Such compression systems, when used to produce streams at different data rates, do not produce delta frames which are interchangeable between the different data rates. Thus in order to create streams that are interchangeable, it is preferable to arrange all of the streams as a series of self-contained packets of given time duration. Each of the packets is self-contained in that it begins with a reference frame and hence does not depend on previously transmitted data. Thus the target is able to use consecutive packets from different streams without any data discontinuity.

The self-contained units are preferably of relatively short duration. The shorter the duration the better the system is able to follow dynamically changes in the bandwidth of the connection. On the other hand, since each unit must start with a reference or key frame, and such frames are very costly in bandwidth since they must be transmitted in full, the use of short self-contained units itself increases the required bandwidth. A compromise unit length of less than ten seconds, preferably 3, 5 or 7 seconds, is preferred.

In an alternative embodiment the length of the self-contained unit is itself allowed to vary dynamically in accordance with circumstances. When the downlink manager 32 detects a significant change in the channel bandwidth it issues a request for a new data rate from the channel manager 30. Alternatively, a new target may be requesting the data feed. The channel manager 30 initiates a time delay when a first request is received, the delay being less than ten seconds, as discussed above, and at the end of the delay starts new self-contained units at all data rates. Thus dynamic following of changes in bandwidth is provided whilst transmitting the minimum possible number of key or reference frames. If no request is received, then the system continues to use delta frames as long as it is practical.

In a variation of the above, instead of waiting to the end of the time delay, the channel manager first of all looks ahead at the approaching data stream and checks whether there is a reference frame approaching during the time delay period in any case. If so, no further action is needed. If there is no reference frame then the channel manager looks ahead to find the biggest delta frame, representing the biggest

Ex. 1018-1093

instantaneous change in the picture, over the delay time period. A reference frame is then inserted in place of the identified largest delta frame. Thus, the insertion of a full reference frame is optimized to the data.

Reference is now made to Fig. 4 which is a generalized block diagram showing the channel manager 30 and a part of a downlink manager 32 of Fig. 3 in greater detail. Parts that are the same as those mentioned in previous embodiments are given the same reference numerals and are not referred to again except as necessary for an understanding of the present embodiment.

As shown in Fig. 4, the channel manager 30 comprises a series of compression units 40.1, 40.2, 40.3, etc. Each of these compression units is preferably operative to compress an input signal received from source 20, using a different compression parameter or set of compression parameters to produce an output having a different data rate. Each compression unit is followed by a buffer 41. The buffer 41 may be used to arranged the compressed data into packets of fixed time duration.

The downlink manager 32 comprises a series of selector switches 42, one for each channel being managed. For the sake of simplicity, only one such switch 42 is shown in the figure. The switch 42 receives each one of the data streams produced by the source 20 and selects one of them at any given time to be sent to the channel 28. Selection is carried out based on a channel capacity measurement 44, which is received from the client 36.

Reference is now made to Fig. 5, which is a generalized block diagram showing a third embodiment of the present invention in which the user is able to set a quality level of the data that is being sent. Parts that are the same as those mentioned in previous embodiments are given the same reference numerals and are not referred to again except as necessary for an understanding of the present embodiment. As described above, compression parameters are altered in order to change the data rate. Generally speaking, there are numerous parameters that can be set in order to alter the data rate, and there are numerous combinations of these parameters that can give the same data rate. For example, parameters that may be altered include:

the extent of image compression within an individual frame,

the number of frames per second to be transmitted,

the proportion of reference frames to delta frames, and

a balance between compression levels of different parts of the data – for example video data includes both sound and image. It is possible to increase a level

Ex. 1018-1094

of compression to the image and to decrease a level of compression to the sound or vice versa.

In Fig. 5, part of a channel manager 30 is shown as well as a switch 42 of a downlink manager. In the channel manager of the embodiment of Fig. 5, a series of data compressors 50 are provided for each required data rate. The data compressors are labeled high, medium and low, although in fact they are different compromises between compression parameters. For example "high" may indicate high image quality but low frame rate, whereas "low" may indicate low image quality but high frame rate. The selector switch 48 is the same as in the previous embodiment except that it now chooses channels based on two different items of information. One item of information is the received channel rate 44, which selects the data rate as before. The other item of information is a data quality request which is received from the target. This selects, within the given data rate, between the different quality levels.

Reference is now made to Fig. 6, which is a generalized schematic diagram of a quality level selection tool 60 for use by a user in selecting a quality level in the embodiment of Fig. 5. The selection tool 60 comprises a simple longitudinal bar 62 which may be marked with labels and numbers, and a slider 64 for moving along the bar 62. The bar shown in Fig, 6 indicates that the user may select different compromises between sound quality and image quality.

Reference is now made to Fig. 7, which is a generalized schematic diagram of another quality level selection tool 70 for use by a user in selecting a quality level in the embodiment of Fig. 5. The selection tool 70 is likewise in the form of a bar and slider but this time the choice is between high quality images at a low frame rate (marked "still") and low quality images at a high frame rate (marked "motion"). It will be appreciated that a low bandwidth "high" may yield a lower image quality than a high bandwidth "low" and vice versa.

Embodiments of the present invention thus provide a system and method in which the measured received data level of a channel is used to select a transmitted data rate so that full advantage may be taken of fluctuating channel capacity levels to send the highest possible quality multimedia data.

It is appreciated that certain features of the invention, which are, for clarity, described in the context of separate embodiments, may also be provided in combination in a single embodiment. Conversely, various features of the invention

Ex. 1018-1095

which are, for brevity, described in the context of a single embodiment, may also be provided separately or in any suitable subcombination.

It will be appreciated by persons skilled in the art that the present invention is not limited to what has been particularly shown and described hereinabove. Rather the scope of the present invention is defined by the appended claims and includes both combinations and subcombinations of the various features described hereinabove as well as variations and modifications thereof which would occur to persons skilled in the art upon reading the foregoing description.

Ex. 1018-1096

Claims

1. A device for streaming of media file data to be sent over at least one channel having a dynamically variable bandwidth, comprising

an input for receiving media file data at a first rate,

a multiple stream creator for creating a plurality of streams of said media file data, each stream compressed to predetermined other rates, each stream comprising a series of autonomous compressed data units which are mutually compatible, and

a stream selector for dynamically matching a bandwidth of the channel to rates of said streams to send said data file as a combination of said autonomous units from said streams via said channel,

wherein said autonomous units have durations which are one of a group comprising fixed predetermined durations and durations determined by changes in said channel bandwidth.

2. A device according to claim 1, wherein said stream selector is operable to select a combination of said streams in succession said combination being chosen to produce an average stream rate over a period of time which matches said channel bandwidth.

3. A device according to claim 1, wherein said bandwidth of said channel is liable to vary dynamically and wherein said stream selector is operable to repeat said matching throughout said streaming operation.

4. A device according to claim 2, wherein said bandwidth of said channel is liable to vary dynamically and wherein said stream selector is operable to repeat said matching throughout said streaming operation.

5. A device according to claim 1, wherein said autonomous units match each other across said plurality of streams, thereby to enable switching between said streams.

6. A device according to claim 5, wherein said data file is arranged in said streams in a compressed media format utilizing reference frames and

Ex. 1018-1097

delta frames and wherein each autonomous unit is self-contained in respect of reference frames.

7. A device according to claim 1, wherein said channel is a cellular telephone channel.

8. A device according to claim 1, wherein said channel is an Internet channel.

9. A device according to claim 1, said predetermined other rate being different for each stream.

10. A device according to claim 1, said predetermined other rate being identical for a set of said streams, each stream in said set being assigned a different balance between quality and speed of said media.

11. A device according to claim 1, said media file comprising a plurality of media types, said predetermined other rate being identical for a set of said streams, each stream in said set being assigned a different balance between quality of said different media types.

12. A device according to claim 10, said streams in said set being selectable by a signal received from said channel.

13. A device according to claim 11, said streams in said set being selectable by a signal received from said channel.

14. A device according to claim 1, comprising a detector for detecting a change in the bandwidth of said at least one channel and operable to request a corresponding change in the data rate, and wherein the device is operable, upon receipt of said request, to set a delay, and to issue a reference frame within said delay.

Ex. 1018-1098

15. A device according to claim 14, operable to issue said reference frame at the end of said delay.

16. A device according to claim 14, operable to issue said reference frame at the occurrence of a largest detected interframe change within the time delay.

17. A multiple stream creator for receiving a data stream at a first data rate and converting said data stream into a plurality of data streams at a plurality of data rates, said creator comprising an input for receiving said data stream at said first data rate, and a plurality of outputs, each operable to receive said data stream at said first rate and to output said data stream at a predetermined rate individually set for each output as a series of autonomous units having a duration, and wherein said duration is one of a group comprising a fixed predetermined duration and a duration dynamically changeable in accordance with bandwidth conditions in a transmission channel.

18. A multiple stream creator according to claim 17, wherein said outputs each comprise data buffers.

19. A multiple stream creator according to claim 17, comprising a data stream processor for processing said data stream received at said first rate into said plurality of autonomous units, said autonomous units being synchronized blocks, thereby allowing said output streams to be mixed in said transmission channel to produce continuously usable output.

20. A multiple stream creator according to claim 17, wherein said received data stream is in an uncompressed data format and wherein said data is compressed into a compressed data format independently for each one of said output data streams.

21. A method of sending a datastream over a variable bandwidth channel comprising modifying a quantity of data to be injected into the channel to accord with a measured capacity of the channel, said method comprising the steps of

Ex. 1018-1099

compressing the datastream, using an algorithm having compression parameters, to produce a plurality of compressed datastreams each comprising reference frames and delta frames, each compressed datastream produced by using the algorithm together with different parameters, and

synchronizing reference frames between the compressed datastreams so that combinations of said compressed datastreams can be sent via said channel.

22. A method according to claim 21, wherein said step of synchronizing reference frames is carried out at intervals determined by a dynamic status of said channel bandwidth.

23. A method according to claim 21, wherein said step of synchronizing reference frames is carried out at predetermined intervals.

24. A method according to claim 21, wherein said combination of said compressed datastreams is selected substantially dynamically to match the measured capacity of the channel.

25. A method according to claim 22, wherein a change in bandwidth of a channel causes a request for commencement of a new autonomous unit, and wherein said request causes initiation of measurement of a time delay, and wherein a new autonomous unit is commenced within said time delay.

26. A method according to claim 25, wherein said new autonomous unit is commenced at the end of said time delay.

27. A method according to claim 25, comprising the step of determining the largest delta frame during the time delay and sending a reference frame, to initiate said new autonomous unit, in place of said determined largest delta frame.

28. A tool for controlling receipt of media file data over a network, wherein said media file data is available at a plurality of quality levels, the

Ex. 1018-1100

tool comprising a quality selector to enable a user to select between said different quality levels for receipt.

29. A tool according to claim 28, wherein said media file comprises a plurality of data groups and wherein each quality level is a different balance between quality levels of the plurality of data groups.

30. A tool according to claim 28, wherein said quality selector comprises a slider operable to be moved along a scale, the position of the slider on the scale indicating a quality level.

31. A device for streaming of media file data to be sent over at least one channel having a dynamically variable bandwidth, substantially as hereinbefore described with reference to the accompanying drawings.

32. A multiple stream creator for receiving a data stream at a first data rate and converting said data stream into a plurality of data streams at a plurality of data rates, substantially as hereinbefore described with reference to the accompanying drawings.

33. A method of sending a datastream over a variable bandwidth channel comprising modifying a quantity of data to be injected into the channel to accord with a measured capacity of the channel, substantially as hereinbefore described with reference to the accompanying drawings.

34. A tool for controlling receipt of media file data over a network, substantially as hereinbefore described with reference to the accompanying drawings.

Ex. 1018-1101

PATENTS • DESIGNS • COPYRIGHT • TRADE MARKS
The Patent Office
INVESTOR IN PEOPLE

Application No: GB 0023062.3 **Examiner:** John Cullen
Claims searched: 1-27, 31-33 **Date of search:** 6 November 2001

Patents Act 1977
Search Report under Section 17

Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK Cl (Ed.S): H4P (PDCFM, PDCFP, PDCFX, PF)

Int Cl (Ed.7): H03M 7/30; H04L 1/00; H04N 7/26, 7/32

Other: Online: WPI, EPODOC, JAPIO

Documents considered to be relevant:

Category	Identity of document and relevant passage		Relevant to claims
A, E	EP 1049074 A1	(LUCENT) See Abstract, Figs. 2 and 4, and paragraphs 14-18 and 28-35.	---
A, E	WO 00/72601 A2	(BIGBAND NETWORKS) See Figs. 1-4, 6, 7 and 9, line 26 of p15 to line 24 of p17 and line 10 of p21 to line 2 of p22.	---
A	WO 99/00984 A1	(CITRIX) See Fig. 15 and line 13 of p30 to line 14 of p33.	---
A	WO 98/09387 A1	(ERICSSON) See Figs. 3A, 6 and 7, and line 35 of p14 to line 26 of p15.	---
A	US 5847760	(OPTIBASE) See whole document.	---

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 202 487 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
02.05.2002 Bulletin 2002/18

(51) Int Cl.7: **H04L 1/00**

(21) Application number: **01309276.2**

(22) Date of filing: **31.10.2001**

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE TR
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **31.10.2000 JP 2000333916**
29.06.2001 JP 2001200156

(71) Applicant: **Kabushiki Kaisha Toshiba**
Tokyo 105-8001 (JP)

(72) Inventors:
- **Nagai, Takeshi, c/o Kabushiki Kaisha Toshiba**
  Minato-ku, Tokyo 105-8001 (JP)
- **Kikuchi, Yoshihiro,**
  c/o Kabushiki Kaisha Toshiba
  Minato-ku, Tokyo 105-8001 (JP)
- **Masuda, Tadaaki, c/o Kabushiki Kaisha Toshiba**
  Minato-ku, Tokyo 105-8001 (JP)

(74) Representative: **Midgley, Jonathan Lee**
Marks & Clerk
57-60 Lincoln's Inn Fields
GB-London WC2A 3LS (GB)

(54) Data transmission apparatus and method

(57) A data transmission apparatus comprises a network condition determination device for determining a transmission condition of the transmission channel in the transmitter (103) based on at least Jitter information or a packet loss rate obtained from a receiver (104), and a video encoder (102) including means for changing at least one of a bit rate of transmission data and a error resilience level according to the transmission condition.

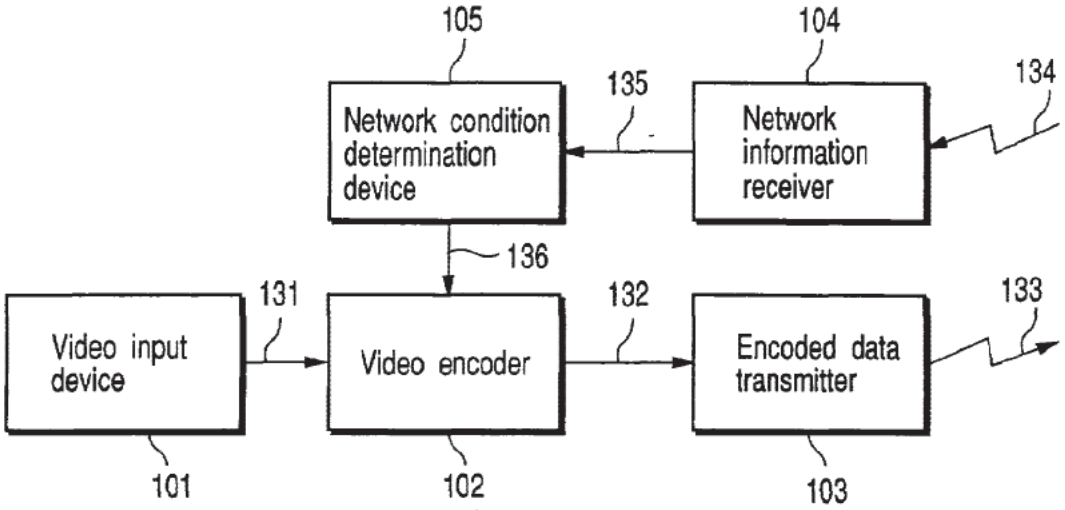


FIG. 1

EP 1 202 487 A2

Printed by Jouve, 75001 PARIS (FR)

Description

[0001] The present invention relates to a data transmission method for transmitting encoded moving/still picture by using a wired communication network such as ISDN (Integrated Services Digital Network) or a wireless communication network such as PHS (Personal handy-phone System) or satellite communication and a data transmission apparatus to which such method is applied.

[0002] In recent years, with progress of digital encoding technology of a variety of information such as image and a broadband network technology, the development of applications using these technologies becomes active, and a system of transmitting compression-encoded image by utilizing a communication network is developed.

[0003] For example, an input video signal is encoded at a video encoder, the obtained encoded data is provided to an encoded data transmitter, and transmission data is transmitted to a transmission channel. In this way, a video is encoded and transmitted. In recent years, with the spread of Internet/Intranet, there has been increased an application or system of transmitting/receiving a packet of data. Providing data as a packet is very effective means for efficiently sharing a bandwidth of a communication channel by a plurality of users.

[0004] In the meantime, as a protocol for transmitting/receiving packet data via Internet/Intranet, there exists TCP/IP (Transmission Control Protocol/Internet Protocol) or UDP/IP (User Datagram Protocol/Internet Protocol) and the like.

[0005] Among them, TCP/IP is an Internet standard protocol. This protocol is applied as the world's best known protocol to Intranet or LAN as well as Internet. TCP works in a transport layer that is a fourth layer of an OSI basic reference model, and IP works in a network layer that is a third layer.

[0006] In addition, UDP/IP is one of the TCP/IP suite transport layer protocols. TCP is a connection type protocol. UDP is a disconnection type protocol. As with TCP, a host application is identified by using a port number. This protocol is often used to carry short data for monitor control of a network or real time oriented data such as audio data or video data.

[0007] TCP/IP incorporates a frame such as retransmission. Thus, this protocol is strong against errors or the like. Even if a large amount of time is required, this protocol is effective to download type applications if an attempt is made to validly receive data. However, this protocol is not powerful for applications requiring real time properties.

[0008] In contrast, although UDP/IP does not have a retransmission frame, a delay associated with retransmission does not occur. Thus, this protocol is very effective to applications requiring real time processing.

[0009] Transmission of moving picture is exemplified as a typical example of application requiring real time processing. However, in the case of general video communication, video data has a very large amount of data, and overflows a network bandwidth in most cases. In this case, a technique for encoding a video data signal, and transmitting a reduced amount of data is used. Techniques such as motion compensation, discrete cosine transform (DCT), sub-band encoding, pyramid encoding, and variable-length encoding or a combination of these techniques are developed as compression-encoding of a video signal.

[0010] As an international standard system for moving picture encoding, there are ISO MPEG-1, MPEG-2, ITU-T H. 261, H. 262, and H. 263. In addition, as an international standard system for multiplying an encoded data stream having a speech/audio signal compressed or other data, there are an ISO MPEG system, ITU-T H. 221, and H. 223.

[0011] Internet or the like is encompassed via an infinite number of networks. Usually, no one knows which network is used or what state is. In addition, an amount of data flowing a network changes from time to time. Thus, a system of judging what amount of data can be communicated in a real time is required.

[0012] Therefore, with further advancement from real time applications utilizing UDP/IP, there have been increased in number applications using a packet format called RTP (Real-time Transport Protocol) for adding and transmitting time information or the like to a packet.

[0013] This RTP is a protocol for transferring audio data signal or video data signal ruled in RFC 1889 in real time. Usually, data is carried on UDP (User Datagram Protocol). This protocol is assumed to be applied to a multimedia system such as video conference, and data can be exchanged in a real time interactive scheme. However, the protocol has no sound quality or picture quality assurance function. Sequential number or time stamp data is assigned into an RTP header, thereby supporting a read time operation.

[0014] RFC is an abbreviation of Request for Comments, and denotes a document such as technical proposal or comments disclosed by TETF (Internet Engineering Task Force). Many de-fact standards such as a variety of protocols of TCP/IP suite are described in RFC.

[0015] By utilizing this RTP, time information and a packet number are added to a packet, so that a receiver can display a voice or image by using correct time information, can determine a packet or the like whose order is replaced in a network, and can detect that a packet is lost by referring to a packet number.

[0016] Moreover, the RTP provides with a system (RTCP) for notifying network state information such as jitter or packet loss rate from a transmitter to a receiver.

[0017] However, how this RTCP information is utilized depends on applications, and is not determined in standards.

[0018] In addition, in the case of an image, a bandwidth corresponding to video transmission cannot be al-

Ex. 1018-1104

located in a network bandwidth, thus making it necessary to compress and deliver a video signal in an encoding system such as MPEG, as described previously. Although this encoding system is effective in a decreased amount of data, the system becomes very fragile to a packet loss or error entry by supplying data to unstable Internet. This is because a moving picture encoding system transmits only a difference between a current frame and a previous frame. Thus, partial data missing causes a very serious problem. In the case of using the UDP or RTP, data retransmission is not basically carried out, thus making it necessary to counteract this problem.

[0019] There generally exist two modes in picture encoding, that is, one is an inter-frame encoding mode for transmitting a difference between the current frame and the previous frame and an intra-frame encoding mode for carrying out encoding in one frame. In general, intra-frame encoded pictures are provided at a proper timing, and inter-frame encoded pictures are provided between them. An interval between the frames encoded in an intra-frame encoding mode is called a GOP (Group of Picture) interval.

[0020] When an encoded picture is missing during transmission to the decoder, the decoded picture corresponding to the missing encoded picture is destroyed. Thereafter, the intra-encoded pictures following the missing encoded picture are decoded based on the destroyed picture. Thus, all the subsequent decoding pictures are affected by the destroyed picture, and decoding cannot be carried out correctly. Therefore, an intra-frame encoded picture is inserted in the way of a stream of frames to cut off a propagation of missing error and recover adequate decoding.

[0021] In a conventional technology, it can determine that errors occur in a network. However, how to utilize the determination result has been unknown. Also, a system for notifying the network information to the transmitter in an unique format is conceived. However, the system is based on the specification of associated applications and lacks a general-purpose use. The system receives transmission data using an information receiver, and controls a video encoder within a range of the information received from the network via a transmission channel.

[0022] A conventional network condition determination is conceived on assumption of network congestion. However, since Internet is popularized in mobile environment, it is required to consider an error or the like in wireless environment. Current technique does not assume countermeasures against such error.

[0023] Further, in the case of dealing with the error, a case in which a network condition changes, and an error rate is not constant is not considered. If a short GOP interval is set in order to counteract a case in which errors frequently occur, the intra-frame encoding increases resulting in impairing the encoding efficiency. When errors occur frequently qualitatively, it may be required to establish such state in which such errors occur frequently. However, in the case of a network in which an error hardly occurs in a normal state, and an error occurs only at one moment, this setting is very wasteful. Conversely, when a long GOP interval is set, it is troublesome that a great influence occurs when an error occurs.

[0024] Thus, a real time transmission using RTP is noted. As described above, by utilizing the RTP that is a protocol for transferring audio data signal or video data signal in real time, time information and a packet number are added to a packet, whereby the receiver can display a voice or image by using correct time information, can determine a packet or the like whose order is replaced in a network, or can detect that a packet is lost by referring to a packet number.

[0025] Moreover, at the RTP, there is provided a system (RTCP) for notifying network state information such as jitter or packet loss rate from the transmitter or receiver.

[0026] It is an object of the present invention to provide a data transmission apparatus and method capable of adjusting a transmission rate by using the RTP characteristics in accordance with a transmission state of a transmission channel based on notification of network state information such as jitter or packet loss rate, obtained from the receiver in the transmitter or capable of carrying out controls such as changing error resilience, thereby making it possible to efficiently transmit data to the maximum, and fully utilize transmission that requires real time properties.

[0027] According to an aspect of the invention, there is provided a data transmission apparatus that transmits transmission data from a transmitter to a receiver via a transmission channel, comprising: an estimation device configured to estimate a transmission condition of the transmission channel in the transmitter based on at least Jitter information or a packet loss rate obtained from the receiver; and a controller configured to change at least one of a bit rate of transmission data and a error resilience level according to the estimated transmission condition.

[0028] According to another aspect of the invention, there is provided a data transmission method comprising: estimating a transmission condition of the transmission channel in the transmitter based on at least Jitter information or a packet loss rate obtained from the receiver; and changing at least one of a bit rate of transmission data and an error resilience level according to the estimated transmission condition.

[0029] According to another aspect of the invention, there is provided a data transmission program stored on a computer readable medium, comprising: instruction means for instructing a computer to estimate a transmission condition of the transmission channel in the transmitter based on at least Jitter information or a packet loss rate obtained from the receiver; and instruction means for instructing the computer to change at least

Ex. 1018-1105

one of a bit rate of transmission data and a error resilience level according to the estimated transmission condition.

[0030] This summary of the invention does not necessarily describe all necessary features so that the invention may also be a sub-combination of these described features.

[0031] The invention can be more fully understood from the following detailed description when taken in conjunction with the accompanying drawings, in which:

FIG. 1 is a view showing a basic configuration of a data transmission apparatus according to a first embodiment of the present invention;
FIG. 2 is a view showing another configuration in the present embodiment;
FIG. 3 is a view showing a configuration of a network condition determination device in the present embodiment;
FIG. 4 is a view showing a configuration of a video encoder in the present embodiment;
FIG. 5 is a view showing a configuration of an encoded parameter determining device in the present embodiment;
FIG. 6 is a view showing a basic configuration of a data transmission apparatus in a second embodiment of the present invention;
FIG. 7 is a view showing a configuration of an encoded data switch device in the present embodiment;
FIG. 8 is a view illustrating a timing of encoded data switching at an encoded data switch device in the present embodiment;
FIG. 9 is a view illustrating a timing of switching encoded data with its different frame interval at the encoded data switch device in the present embodiment;
FIG. 10 is a flow chart showing a basic example of the network condition determination method at the network condition determination device in the first embodiment of the present invention;
FIG. 11 is a flow chart showing a basic example when a delay is forced to occur by using the network condition determination method at the network condition determination device in the first embodiment of the present invention;
FIG. 12 is a flow chart showing a basic example considering a time direction change with the network condition determination method at the network condition determination device in the present embodiment;
FIG. 13 is a view showing an exemplary configuration according to a third embodiment of the present invention;
FIG. 14 is a flow chart showing an example of a GOP interval calculation method at a GOP interval calculator in the present embodiment;
FIG. 15 is a flow chart showing another example of a GOP interval calculation method at a GOP interval calculator in the present embodiment; and
FIG. 16 is a flow chart showing still another example of a GOP interval calculation method at a GOP interval calculator in the present embodiment.

[0032] Hereinafter, embodiments of the present invention will be described with reference to the accompanying drawings.

[0033] The present invention utilizes RTP or RTCP that is a protocol for transferring audio data signal or video data signal in real time. The transmitter performs controls such as adjusting a transmission rate or changing error resilience in accordance with a transmission state of a transmission channel based on notification of network state information acquired from the receiver such as jitter (Jitter information) or packet loss rate. In this manner, data transmission can be carried out efficiently to the maximum. Hereinafter, embodiments of the present invention will be described in detail.

(First Embodiment)

[0034] FIG. 1 shows a basic configuration of an image transmission apparatus according to a first embodiment of the present invention. In FIG. 1, a video signal 131 input from an image input device 101 is processed to be encoded. That is, an encoded data transmitter 103 processes encoded data 132 that has been encoded by a video decoder 102 in the form which conforms to a network, and transmits the encoded data to a receiver. A network information receiver 104 receives network information 134 delivered from the receiver, and outputs the received information to a network condition determination device 105. This network condition determination device 105 has a function that determines a network condition from network information 135 output from the network information receiver 104, and notifies the result as network condition information 136 to the video encoder 102.

[0035] The video encoder 102 encodes the video signal 131 using the network condition information 136. This encoding is carried out as follows. That is, in the case of RTP or RTCP, a packet loss rate or Jitter information is provided to the network condition determination device 105. The network condition determination device 105 determines that, if a packet loss rate is not "0" or if a delay time is longer than a predetermined value, any load is applied to a network, and a presumed amount of data cannot be supplied. Then, the network condition determination device 105 is configured so as to carry out processing so as to issue an instruction (network condition information 136) for lowering a bit rate to the video encoder 102, and to lowly set a target bit rate in encoding.

[0036] According to the thus configured apparatus, when the video signal 131 is inputted from the image input device 101, the inputted video signal 131 is encod-

Ex. 1018-1106

ed by the video encoder 102. Encoded data 132 encoded by the video encoder 102 is inputted to the encoded data transmitter 103.

[0037] The encoded data transmitter 103 processes the encoded data in the form suitable to a network, and receives the processed data to a receiver. The network information receiver 104 receives the network information 134 delivered from the receiver, and outputs the received information to the network condition determination device 105.

[0038] The network condition determination device 105 determines a network condition from the network information 135 outputted from a network information receiver, and notifies the result as the network condition information 136 to the video encoder 102.

[0039] The video encoder 102 encodes the video signal 131 using the thus notified network condition information 136.

[0040] FIG. 10 shows a very simple example of a method for determining a network condition at the network condition determination device 105 by way of showing a flow chart.

[0041] According to this flow chart, at the step S1001, a packet loss rate of network information is checked to see if the rate is "0" or not. As the result of this check, if the packet loss is not "0", an effective bit rate is calculated from the packet loss rate in accordance with the step S1002. This bit rate can be calculated by formula (1) shown below.

$$b' = b \times (1 - r) \qquad (1)$$

where "b" denotes a current bit rate, b' denotes a new bit rate, and "r" denotes a packet loss rate.

[0042] Next, at the step S1003, a parameter (encoding parameter) obtained at the step S1002 is notified to the video encoder 102. Here, the above parameter is used to determine a rate control or error resilience level for encode processing at the image encoder 102. This parameter is also used for adjustments such as adjusting and controlling the encode processing speed of the input image signal 131 to this parameter, thereby changing a bit rate, a frame interval, or an error resilience level.

[0043] On the other hand, at the step S1001, when the packet loss rate is "0", no parameter change occurs at the step S1004. This determination method is provided as a mere example. It is possible to employ any other method for obtaining a bit rate based on a case in which the criterion of the step S1001 is determined by thresholds without being set to "0" or others or based on a formula other than the above formula (1).

[0044] Therefore, according to the first embodiment, even when a transmission network bandwidth is unknown or when the bandwidth changes halfway, it is possible to reset a parameter or the like based on the receiver's network information, and then, encode and transmit an image in accordance with the parameter suitable to a current network. In this manner, even in the case of a network congestion, although a bit rate is automatically lowered, and an image quality is lowered, it is possible to communicate an image without an occurrence of a phenomenon that an image is destroyed due to such packet loss. This method is very effective in image transmission in real time.

[0045] The first embodiment can be carried out by modifying it as follows. For example, as shown in FIG. 11, a method other than that shown in FIG. 10 is employed as a method for determining a network condition at the network condition determination device 105. That is, in the flow chart shown in FIG. 10, a parameter is set at only a packet loss rate. In contrast, in the method shown in FIG. 11, even when no packet loss occurs at the step S1104, a delay quantity is checked. As a result of this check, if a delay occurs in excess of a predetermined value, it is determined that a network is congested. When it is determined that a network is congested, a bit rate is calculated, and a new parameter is set by using formula (2) below, for example, at the step S1105.

$$b' = b \times \{(TN-TS)+(d-dth)\}/(TN-TS) \qquad (2)$$

where "b" denotes a delay time, b' denotes a new bit rate, TN denotes a current time, TS denotes a start time, "d" denotes a delay time, and "dth" denotes delay time thresholds.

[0046] FIG. 2 shows an example when the network condition determination device 105 is provided with a function for generating the network condition information 136 so as to provide the information to the image encoder 102 by extending the configuration shown in FIG. 1.

[0047] According to a configuration shown in FIG. 2, the network condition determination device 105 is provided with a function for acquiring a current parameter at the current image encoder 102 and a function for generating the network condition information 136 on what parameter should be set next from the thus acquired current parameter and network information 135, thereby providing the parameter to the image encoder 102.

[0048] In the case of this construction, the network condition determination device 105 can recognize a state of a video parameter, i.e., a state of the encode processing currently done at the video encoder 102, thus making it possible for the network condition determination device to calculate and grasp a bit rate or the like. This construction is advantageous in the case of determining a network condition by comparing the network information 135 with parameter information 137. Thus, a system is realized which is capable of changing a parameter so as to set an optimal bit rate in accordance with the network condition, and providing the parameter to the video encoder 102, thereby adjusting the encode processing.

[0049] In this way, according to the present invention,

Ex. 1018-1107

transmission is adjusted in accordance with a transmission state of a transmission channel or error resilience is changed. The transmission state is adjusted by changing a bit rate or changing a frame interval. With respect to a change in quality of a transmission channel, the error resilience is changed according to the transmission state of the transmission channel (for example, the error resilience is changed by changing an interval of an MPEG4 synchronous signal). Therefore, data transmission can be carried out efficiently to the maximum, thereby enabling data transmission that makes it possible to fully utilize transmission requiring real time processing.

[0050] FIG. 3 shows a configuration when the configuration shown in FIG. 2 is further developed. This network condition determination device 105 comprises a network information storage device 301 that stores past network information and an encoded parameter information storage device 302 that stores encoded parameter information supplied to the current video encoder 102. A network condition determination device 303 recognizes the past network information and the current encoded parameter from these elements, decides an elapse of time in network condition from these elements to determine an optimal parameter assigned to the video encoder 102, and provides the parameter to the video encoder 102.

[0051] With this construction, it is possible to determine a network condition more accurately by referring to an elapse of time relevant to network information. FIG. 12 shows a flow chart of a network condition determination method in the case of using this system. Processing at the network condition determination device 303 will be described in accordance with this flow chart. First, the network condition determination device 303 determines whether or not packet loss is "0" at the step S1201. As a result of this determination, when the packet loss is present, it is determined by the step S1202 whether or not the past packet loss occurs. As a result, when the past packet loss occurs, the loss rate is compared with the current rate. If the current packet loss rate is higher than the past packet loss rate, it can be determined that the past correction has been invalid.

[0052] It is understood from this determination that the bandwidth overflow of transmission data such as network congestion is not troublesome, but a possibility that a noise or the like occurs over a channel resulting in destroying is high. Therefore, the network condition determination device 303 does not change a bit rate at the step S1204, and carries out parameter setting to ensure reliable error correction, and notifies the fact to the video encoder 102 at the step S1205.

[0053] If the current packet loss rate is lower than the past packet loss rate at the step S1203, the network condition determination device 303 determines that the past change works efficiently, and re-calculates a valid bit rate from the packet loss rate at the step S1206. When the past packet loss does not occur at the step S1202, the last loss is determined as a channel noise, and processing goes to the step S1204.

[0054] On the other hand, when no packet loss occurs at the step S1201, the network condition determination device 303 checks a delay quantity at the step S1207. As a result, when the delay quantity exceeds thresholds "Th", a valid bit rate is calculated from the delay quantity at the step S1208, and the calculation result is notified to the video encoder 102 at the step S1205.

[0055] Lastly, when the delay quantity does not exceeds the thresholds "Th", the network condition determination device 303 determines that successful communication is established, and does not change a parameter at the step S1209.

[0056] By causing the network condition determination device 303 to carry out such processing, the network condition determination device 105 can have a function capable of determining whether a packet loss occurs on a transmission channel due to network congestion or whether a noise is generated, and the loss occurs due to the impairment of a channel state such as radio equipment.

[0057] This function is provided as a mere example. For example, when the past packet loss does not occur at the step S1202, a control for lowering a bit rate may be made by determining network congestion instead of determination of noise. If the current packet loss rate is improved more remarkably than the past packet loss rate at the step S1203, similar parameter change may be made after checking how the past change has been made. In this way, in this system, a variety of determination systems can be set according to a network.

[0058] In the foregoing, although there has been shown an example in which the video encoder 102 carries out encode processing based on encoded parameters determined and assigned at the network condition determination device 105, a configuration for the video encoder 102 to determine a parameter can be provided without being limited to the network condition determination device 105. An example is shown in FIG. 4.

[0059] FIG. 4 is a block diagram depicting an inclusive configuration when the video encoder 102 determines an encoded parameter. In the case of this example, the video encoder 102 comprises an encoded parameter determination device 401 and a signal processor 402.

[0060] In the case of this configuration, the network condition determination device 105 is configured to output the network condition information 136 from the network condition determination device 303, and then, provide the information to the video encoder 102.

[0061] In such a configuration, the network condition determination device 105 provides the network condition information 136 delivered from the network condition determination device 303 to the video encoder 102. Then, the video encoder 102 first inputs the network condition information 136 to an encoded parameter determination device 401. Then, the encoded parameter determination device 401 recognizes a network condi-

Ex. 1018-1108

tion from this network condition information 136, and generates an encoded parameter 431 in the form suitable to the network condition.

[0062] The thus generated encoded parameter 431 is inputted to a signal processor 402. Then, this signal processor 402 encodes the video signal 131 inputted from the video input device 101 by using this encoded parameter 431.

[0063] Encoding information 432 such as the number of encoded bits after encoding is inputted to the encoded parameter determination device 401, and is used for the next encoded parameter determination. Encoded data 132 encoded at the signal processor 402 is output to the encoded data transmitter 103.

[0064] FIG. 5 is a block diagram depicting a configuration including a intra-frame encoding determination device 501 that forcibly sets intra-frame encoding at the encoded parameter determination device 401 shown in FIG. 4. As shown in the figure, the encoded parameter determination device 401 comprises an intra-frame encoding/determination device 501 and an encoded parameter determination device 502.

[0065] In the case of this configuration, the network condition information 136 outputted from the network condition determination device 105 is input to the encoded parameter determination device 502 possessed by the encoded parameter determination device 401. Then, encoded parameter information 531 corresponding to the network condition information 136 is output from the encoded parameter determination device 502.

[0066] The encoded parameter information 531 is input to the intra-frame encoding determination device 501, and it is determined whether inter-frame encoding is carried out as is or intra-frame encoding is forcibly carried out. When it is determined as intra-frame encoding, the encoding parameter information 531 is updated so as to carry out intra-frame encoding, and is output as the encoded,parameter information 431 to the signal processor 402 of the video encoder 102.

[0067] By doing this, even if the past data is not transmitted correctly to a receiver, when the data can be received correctly after parameter change, a correct image can be reproduced.

[0068] It is possible to configure a video encoder capable of determining a network condition, which has functions of the video encoder 102 and network condition determination device 105, thereby making it possible to provide the previously described function in one block.

[0069] The encoded parameter determination device 401 or network condition determination device 105 makes determination concerning the next encoded parameter setting. However, if a picture quality is changed suddenly due to temporary instability of network, an unclear picture may be produced. In order to suppress such a phenomenon, it is possible to incorporate a mechanism in which a change in encoded parameter is suppressed within its predetermined range. In the above described circumstance, this mechanism functions efficiently.

[0070] Although a bit rate or an error resilience parameter may be set as its calculated value, it is possible to employ a method of predetermining some patterns, and then, selecting the closest pattern from among them. Further, instead of the closest pattern, it is possible to make a selection of the closest bit rate which is lower than the calculated bit rate. This makes it possible to prevent the generation of a video image with its unexpected image quality due to a combination of a variety of parameters, and makes it possible to prepare a combination of encoded parameters that have been tested to some extent.

[0071] In the foregoing example, when the packet loss rate is not "0" or when a delay exceeds the thresholds, there has been made a control such that parameters such as bit rate or error resilience are changed, and the changed is reset to that suitable to the current state.

[0072] However, after recovery of a network condition, when no packet loss occurs or when a small delay occurs, it is possible to make a control for increasing a bit rate or decreasing an error resilience parameter.

[0073] The foregoing example will be described below according to a second embodiment.

(Second Embodiment)

[0074] FIG. 6 is a view showing a basic configuration of an image transmission apparatus according to the second embodiment of the present invention. The image transmission apparatus shown in FIG. 6 comprises a plurality of media 601, an encoded data switch device 602, an encoded data selector 603, an encoded data transmitter 103, a network information receiver 104, and a network condition determination device 105.

[0075] The storage media 601 are used to store encoded data 631 on contents that have been encoded in advance. The storage media 601 each store encoded data on the same contents encoded by different encoded parameters, respectively.

[0076] The encoded data switch device 602 selects one of a plurality of storage media 601, and provides encoding data 631 stored in the thus selected storage medium 601 to the encoded data transmitter 103. A selection of the storage media is carried out in accordance with switch information 632 from the encoded data selector 603.

[0077] The network information receiver 104 has a function that receives network information 134 such as packet loss rate or Jitter information delivered from a receiver or a network, and provides the received information as network information 135 to the network condition determination device 105.

[0078] The network condition determination device 105 determines a network condition from the input network information 135, and outputs the result as network condition information 136 to the encoded data selector

Ex. 1018-1109

603. The encoded data selector 603 has a function for estimating, from this network condition information 136, which of the encoded data on encoding parameters is optimal to the current network, and then, outputting to the encoded data switch device 602 the switch information 632 used for switching in order to select encoded data caused by the estimated optimal encoding parameter as an output.

[0079] The encoded data switch device 602 selects one item of encoded data from among the input encoded data 631, and outputs encoded data 132.

[0080] With such a configuration, the encoded data 631 encoded in advance and store in the storage medium 601 or the like is input to an encoding data switch device 602 when the data is reproduced from the storage medium 601. That is, respective encoded data encoded by the differential encoding parameters are maintained in a plurality of storage media 601. As a result of reproducing these items of data, the encoded data encoded by the different encoding parameters, respectively, is inputted. The encoded data switch device 602 selects one from among these items of data, and outputs it to the encoded data transmitter 103.

[0081] This selection is carried out in accordance with switch information 632 from the encoded data selector 603.

[0082] On the other hand, network information 134 such as packet loss rate or Jitter information from a receiver is received at the network information receiver 104, and then, the received information is delivered to the network condition determination device 105. The network condition determination device 105 determines a network condition from the inputted network information 135, and outputs the result as network condition information 136 to an encoded data selector 603.

[0083] The encoded data selector 603 estimates, from the network condition information 136, which of the encoded data of the encoding parameters is optimal to the current network, and outputs the switch information 632 to the encoded data switch device 602.

[0084] The encoded data switch device 602 selects one item of the encoded data from among the input encoded data 631, and outputs encoded data 132.

[0085] The present embodiment will be described by way of a specific example.

[0086] Now, assume that there exist five storage media 601a to 601e each having stored therein five items of encoded data encoded at 384 kbps, 128 kbps, 64 kbps, 32 kbps, and 16 kbps.

[0087] That is, for example, a storage medium having stored therein encoded data on contents encoded at 384 kbps is defined as a storage medium 601a; a storage medium having stored therein encoded data on contents encoded at 128 kbps is defined as a storage medium 601b; a storage medium having stored therein encoded data on contents encoded at 64 kbps is defined as a storage medium 601c; a storage medium having stored therein encoded data on contents encoded at 32 kbps is defined as a storage medium 601d; and a storage medium having stored therein encoded data on contents encoded at 16 kbps is defined as a storage medium 601e.

[0088] In the case of using a transmission channel having its quality capable of being transmitted at 384 kbps, first, reproduced encoded data is selected by the storage medium 601a that is a storage medium having stored therein encoded data on contents encoded at 384 kbps; transmission is started at 384 kbps, and reception of network information from the receiver starts at the same time.

[0089] Although network information is received at a predetermined interval, if a problem occurs with a network based on the received information, a valid bit rate is calculated at a stage at which a packet loss or delay occurs. The encoded data selector 603 selects encoded data on a bit rate close to the calculated bit rate, and the encoded data switch device 602 switches and transmits the selected data.

[0090] For example, the calculated bit rate is 140 kbps, a value close to this bit rate is 128 kbps. In this case, the above switch device selects and switches encoded data reproduced from the storage medium 601b having stored therein encoded data on contents encoded at 128 kbps. In addition, if the calculated bit rate is 100 kbps, a value close to a bit rate lower than the calculated bit rate is 64 kbps. In this case, the switch device selects and switches encoded data reproduced from the storage medium 601c having stored therein encoded data on contents encoded at 64 kbps.

[0091] In this way, in the present embodiment, contents are prepared after encoded in advance for type of transmission bit rate, and a current optimal transmission rate is obtained based on network information such as packet loss rate or Jitter information from a receiver that is a function possessed by RTP. A storage medium having stored therein encoded data obtained at the bit rate corresponding to this transmission rate is selected by the switch 602, and reproduced encoded data output of the storage medium is transmitted. In this manner, there is no need for a server to carry out encode processing in real time, making it possible to reduce a load on the server. This is very effective in distributing a pre-produced video such as broadcast in particular.

[0092] Now, a modified example of the second embodiment will be described here. FIG. 7 is a block diagram depicting a method for carrying out control for a switch timing of encoded data to be inputted at the encoded data switch device 602.

[0093] Encoded data 631 is inputted to the switching position detecting device 701 and a switch device 702. Switch information 632 is inputted to the switching position detecting device 701. If there occurs a need to execute switching from the current encoded data to the other encoded data according to the switch information 632, the switching position detecting device 701 analyzes encoded data, and detects a switching enable posi-

Ex. 1018-1110

tion.

[0094] This means that a search is made for a frame encoded by intra-frame encoding (I-Picture).

[0095] When the switching enable position is set, an instruction for switching encoded data is supplied to the switch device 702 by using the switching instruction information 731. FIG. 8 shows an example of this instruction. Assume that encoded data A is currently selected, and there occurs a need to execute switching to encoded data B according to the switch information 632.

[0096] At this stage, assume that a frame is processed at the time of "switching instruction" shown in FIG. 8. In such a case, if encoded data is switched, an unmatched image occurs. This is because a difference from the previous image is encoded in intra-frame encoding (P-Picture). Therefore, if encoded data is switched at the time of P-Picture, an attempt is made to reproduce the next picture (picture P11 of encoded data (B)) by using immediately preceding picture (decoded picture 111 of encoding data (A)). This provides a problem that a correct decode picture cannot be obtained, and a picture is destroyed. Thus, there occurs a need to execute switching at a timing of a picture of intra-frame encoding (I-Picture) undergoing encoding by that frame only. This timing is defined as a timing at a position marked with "execute switching".

[0097] As in FIG. 9, a case in which the number of frames or time position differs depending on encoded data is considered. In this case as well, a switching enable position of encoded data at a switching destination is detected, and switching of encoded data is executed when the switching enable position is set.

[0098] At the switching position detecting device 701 shown in FIG. 7, when a need to execute switching according to the switch information 632 occurs, it is possible to employ a method for temporarily stopping an output 132 to the encoded data transmitter 103, and then, outputting new encoded data to the encoded data transmitter 103 at a stage at which the switching position is set.

[0099] This method is effective when it is necessary to suppress a buffer overflow or the like by switching encoded data. When an underflow is problematic, a stuffing bit is forcibly inserted, thereby making it possible to counteract such an underflow.

(Third Embodiment)

[0100] Now, a third embodiment according to the present invention will be described here.

[0101] The present embodiment describes an exemplary configuration of an encoding parameter determination device 401 of the video encoder 102 shown in FIG. 4 according to the first embodiment when a time interval (GOP interval) between frames encoded in a intra-frame encoded mode is adjusted based on the network condition information 136 outputted from the network state determination device 105. Here, a description will be given by focusing on differences from the first embodiment.

[0102] FIG. 13 shows an exemplary configuration of the encoding parameter determination device 401. This device comprises a GOP interval calculator (GOP interval calculator) 1501 and an encoding parameter determination device 502.

[0103] The network condition information 136 outputted from the network condition determination device 105 (including at least packet loss rate information) is input to an encoding parameter determination device 502. The parameter determination device 502 inputs packet loss rate information 1531 of the network condition information 136 to the GOP interval calculator 1501. The GOP interval calculator 1501 calculates a GOP interval from the packet loss rate information 1531. The GOP interval information 1532 is notified to the encoding parameter determination device 401. The encoding parameter determination device 401 outputs encoding parameter information 431 that contains the inputted GOP interval information 1532. When internally generated parameter information exists, the encoding parameter determination device 401 outputs GOP interval information 1532 and encoding parameter information 431 that contains the internally generated parameter information.

[0104] FIG. 14 shows an example of a method for determining a GOP interval at the GOP interval calculator.

[0105] At the step S1601, it is determined whether or not a packet loss occurs from the network condition information 136. When such a packet loss occurs, the GOP interval is estimated by formula (3) below, for example, at the step S1602.

$$Gop = (TN - TL)/\{(FN - FL) \times r\} \quad (3)$$

where "gop" denotes a GOP interval, "r" denotes a packet loss rate, FN denotes a total number of current frames, FL denotes a total number of frames during past calculation, TN denotes a time, and TL denotes a time during past calculation.

[0106] The GOP interval obtained at the step S1602 is notified to the signal processor 402 at the step S1603, and encoding is carried out based on the notified value.

[0107] On the other hand, when no packet loss has occurred at the step S1601, the default GOP thresholds are read out at the step S1604. At the step S1603, this value is notified to the signal processor 402.

[0108] According to the present embodiment, it is possible to change dynamically the GOP interval to a value suitable to a network. In this manner, when a small number of errors occur, the GOP interval is increased, thereby making it possible to reduce wasteful intra-frame encoded modes in number. In contrast, when a large number of errors occur, the GOP interval is decreased, and a frame in the intra-frame encoded mode appears quickly, thereby making it possible to acceler-

Ex. 1018-1111

ate recovery. In the present embodiment, it is possible to efficiently select an intra-frame encoded mode.

[0109] FIG. 15 shows another example of a method for determining a GOP interval at the GOP interval calculator in accordance with a flow chart.

[0110] In this example, even when packet loss is fewer than that in the system shown in FIG. 14, the network condition is estimated from the past history of packet loss at the step S1704, and an optimal GOP interval is calculated. This technique makes it possible to estimate whether a network is completely free of an error or an error may occurs by checking the past history instead of resetting the GOP interval to the default value by determining that no packet loss occurs at one moment. In this manner, it is possible to determine the network condition more accurately.

[0111] FIG. 16 shows still another example of a method for determining a GOP interval at the GOP internal calculator in accordance with a flow chart.

[0112] In this example, the GOP interval is determined by a unique calculation system instead of determination based on the presence/absence of a packet loss. This system can be used to determine the GOP interval from the current packet loss rate in accordance with a single formula or to determine the most suitable GOP interval by checking the past history.

[0113] The following formula (4) shows an example of calculation formula for determining the GOP interval.

$$Gop = [(TN - TL)/\{(FN - FL) \times r\}] \times \alpha \qquad (4)$$

where "gop" denotes a GOP interval, "r" denotes a packet loss rate, α denotes a sensitivity coefficient, FB denotes a total number of current frames, FL denotes a total number of frames during past calculation, TN denotes a current time, and TL denotes a time during past calculation.

[0114] In this example, the GOP interval calculated from an actual value is multiplied by a sensitivity coefficient, thereby making it possible to meet a requirement for accelerating recovery by sacrificing efficiency or a requirement when one does not take care even if an effect of error is somewhat extended. For example, if the sensitivity coefficient α is set to be smaller than 1, a GOP interval which is smaller than that calculated from an actually measured value is output. In this manner, although the efficiency is sacrificed, when an error occurs, it is possible to estimate faster recovery.

[0115] The present embodiment is not limited to formulas (3) and (4). For example, in these formulas, a value is calculated by using an increment during past calculation and subsequent, it is possible to calculate the value by taking a margin such as the n preceding increment. In this manner, it is possible to cope with a gradual change hidden in a change in detailed state or to eliminate a temporal change to some extent. Similarly, algorithms are not limited to those shown in the present embodiment.

[0116] The foregoing embodiment has described an exemplary configuration when a time interval (GOP interval) is adjusted, the time interval being between the frames encoded in the intra-frame encoded mode based on the network condition information 136 at the encoding parameter determination device 401 of the video encoder 102 shown in FIG. 4 according to the first embodiment. Alternatively, there can be provided a configuration in which the network condition determination device 105 according to the first embodiment (FIGS. 1, 2, or 3) determines a time interval (GOP interval) encoded in the intra-frame encoded mode based on information from the network information or information from the network information receiver 104 and information from the video encoder 102, and provides the encoded time interval to the image encoder 102.

[0117] As described in the first embodiment, it is possible to configure a video encoder capable of determining a network condition, the video encoder having functions of the video encoder 102 and network condition determination device 105, thereby providing the previously described functions in one block.

[0118] In the foregoing, although a variety of embodiments have been described, in short, the present invention utilizes characteristics of RTP, and comprises a system (RTCP) for notifying network state information such as jitter or packet loss rate from a transmitter or receiver. At the transmitter, a bit rate of transmission data at the transmitter is adjusted according to a transmission state of a transmission channel based on the notification of network state information acquired from the receiver in the transmitter such as jitter or packet loss rate so that control such as change of the error resilience level can be carried out. Therefore, according to the above described present invention, data can be transmitted efficiently to the maximum so as to achieve data transmission that makes it possible to fully utilize transmission requiring real time properties as well.

[0119] The techniques described in the embodiments of the present invention can be distributed to be stored as a computer executable program in a recording medium such as a magnetic disk (such as a flexible disk or a hard disk); an optical disk (such as CD-ROM, CD-R, CD-RW, DVD, or MO), and a recording medium such as a semiconductor memory. Further, the above program can be distributed by means of transmission via a network.

[0120] According to the present invention, it is possible to determine a network condition, and then, optimally set an encoding parameter. In addition, when a packet loss caused by congestion and a loss of a channel error coexist in a network, it is possible to determine the loss, and then, to set a parameter.

[0121] Therefore, according to the present invention, there can be provided a data transmission apparatus and method in which data can be transmitted efficiently to the maximum so as to achieve data transmission that

Ex. 1018-1112

makes it possible to fully utilize transmission requiring real time properties.

Claims

1. A data transmission apparatus for transmitting transmission data from a transmitter to a receiver via a transmission channel, **characterized by** comprising:

   means (105) for estimating a transmission condition of the transmission channel in the transmitter based on at least Jitter information or a packet loss rate obtained from the receiver; and
   means (102) for changing at least one of a bit rate of transmission data and an error resilience level according to the estimated transmission condition.

2. An apparatus according to claim 1, **characterized in that** the transmitter includes means (102) for generating encoded data obtained by encoding a single content at different transmission rates, and means (603) for selecting optimum encoded data having optimum bit rate from the encoded data of the encoded data generating means (102) to adjust a bit rate of the transmission data in the transmitter according to the estimated transmission condition.

3. A data transmission method of transmitting transmission data from a transmitter to a receiver via a transmission channel, **characterized by** comprising:

   generating encoded data obtained by encoding a single content at different transmission rates;
   estimating a transmission condition of the transmission channel in the transmitter based on at least Jitter information or a packet loss rate obtained from the receiver; and
   selecting optimum encoded data having an optimum bit rate from the encoded data to adjust a bit rate of the transmission data in the transmitter according to the estimated transmission condition.

4. A data transmission method **characterized by** comprising:

   estimating a transmission condition of the transmission channel in the transmitter based on at least Jitter information or a packet loss rate obtained from the receiver; and
   changing at least one of a bit rate of transmission data and an error resilience level according to the estimated transmission condition.

5. A data transmitting apparatus that transmits transmission data from a transmitter to a receiver via a transmission channel, **characterized by** comprising:

   means (102) for generating plural encoded data obtained by encoding a single content at different bit rates, in the transmitter;
   means (105) for estimating a condition of the transmission channel based on Jitter information or packet loss rate; and
   means (603) for selecting encoded data having an optimum transmission rate from the plural encoded data according to the estimated transmission condition.

6. A method of transmitting transmission data from a transmitter to a receiver via a transmission channel, the method **characterized by** comprising:

   storing plural encoded data obtained by encoding a content at different bit rates;
   estimating a condition of the transmission channel based on Jitter information or packet loss rate;
   selecting encoded data having optimum transmission rate from the plural encoded data to adjust a bit rate of the transmission data in the transmitter.

7. A data transmission program stored on a computer readable medium, the program **characterized by** comprising:

   instruction means for instructing a computer to estimate a transmission condition of the transmission channel in the transmitter based on at least Jitter information or a packet loss rate obtained from the receiver; and
   instruction means for instructing the computer to change at least one of a bit rate of transmission data and a error resilience level according to the estimated transmission condition.

8. A data transmission apparatus, **characterized by** comprising:

   receiver means (104) for receiving network information regarding a condition of a network;
   determination means (105) for determining a network condition based on the network information from the receiver means;
   encoder means (102) for encoding a video data signal to generate encoded video data and control a bit rate or an error resilience level of the encoded video data according to the network condition information; and
   transmitter means (103) for transmitting the en-

Ex. 1018-1113

coded video data to the network.

9. An apparatus according to claim 6, **characterized in that** the encoder means (102) is supplied with an encoding parameter as the network condition information and generates encoded data whose bit rate is adjusted according to the encoding parameter,

the determination means (105) includes first storage means (302) for storing encoding parameter information obtained by the encoder means, second storage means (301) storing the network information input by the network information receiver, and determination means (303) for determine the encoding parameter by deciding the network condition based on network information input by the network information receiver, past encoding parameter information from the first storage means and past network information from the second storage means.

10. An apparatus according to claim 8, **characterized in that** the encoder means (102) includes parameter determination means (401) for determining the encoding parameter from the network condition information input from the receiver means, the encoder means (102) encoding the video data signal to generate the encoded video data whose bit rate is adjusted in correspondence with the encoding parameter determined by the determination means (401).

11. An apparatus according to claim 10, **characterized by** including encoding determination means (501) determine whether a next frame is to be encoded forcibly in an intra-frame encoding mode, when the parameter determination means (401) determines the encoding parameter using the network condition information.

12. An apparatus according to claim 8 or 9, **characterized in that** the determination means (105) includes computation means for computing a GPO interval from the network information input from the receiver means (104) to generate network condition information including GOP interval information indicating the GOP interval obtained by the computation means, and the encoder means (102) encodes the video data signal according to the GOP interval information included in the network condition information.

13. An apparatus according to claim 8 or 10, **characterized in that** the encoder means (102) includes GOP interval computation means for computing a GOP interval based on the network information input from the receiver means (104) and encodes the video data signal according to the GOP interval information indicating the GOP interval computed by the GOP interval computation means.

14. An apparatus according to claim 12 or 13, **characterized in that** the determination means (105) determines at least a packet loss rate as the network condition, and the GOP interval computation means computes the GOP interval based on at least the packet loss rate.

15. An apparatus according to claim 8, **characterized by** including switching means (602) for switching among plural encoded data encoded at different transmission rates, respectively, to select encoded data having a transmission rate corresponding to switching information, and selector means (603) for selecting a transmission rate of encoded data to be transmitted from the transmitter based on the network condition information outputted from the determination means (105) and output the selected transmission rate as the switching information to the switching means.

16. An apparatus according to claim 15, **characterized in that** the switching means (602) includes a switching position detector for detecting a switching position at which one of the encoded data can be switched to the other, and when an encoded data switching signal is generated by the switching information output from the selector means (603), the switching position detector detects the switching position, to switch the encoded data at the switching position.

Ex. 1018-1114

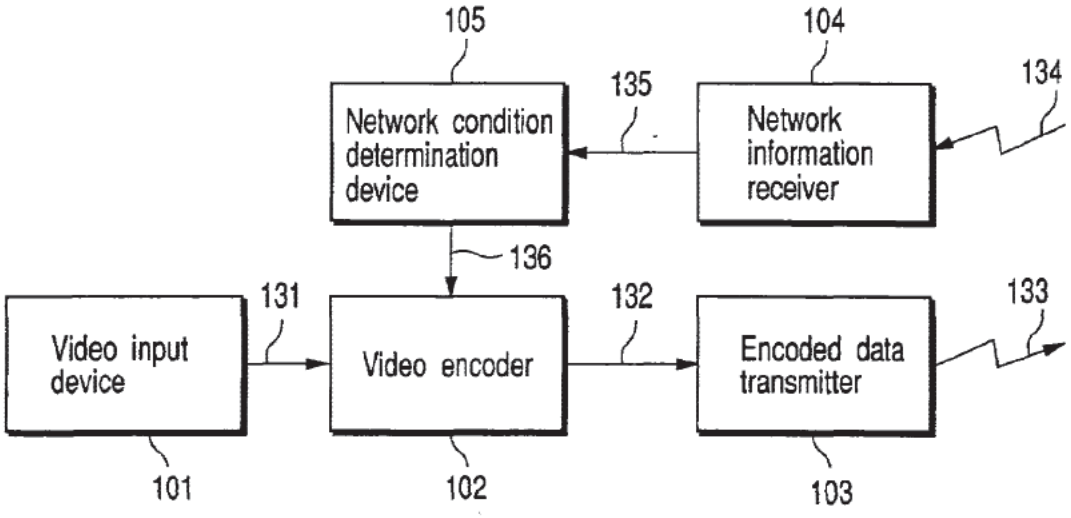


FIG. 1

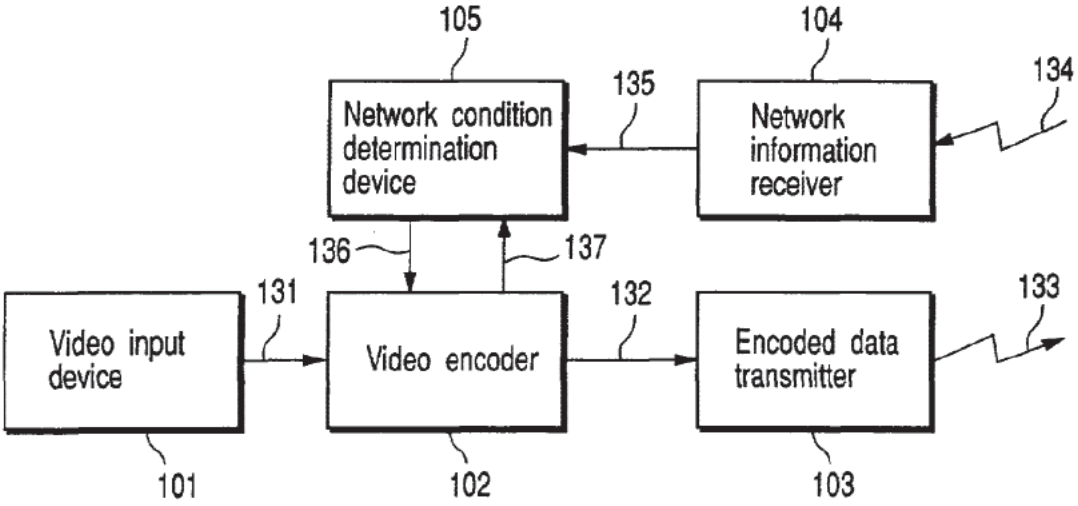


FIG. 2

Ex. 1018-1115

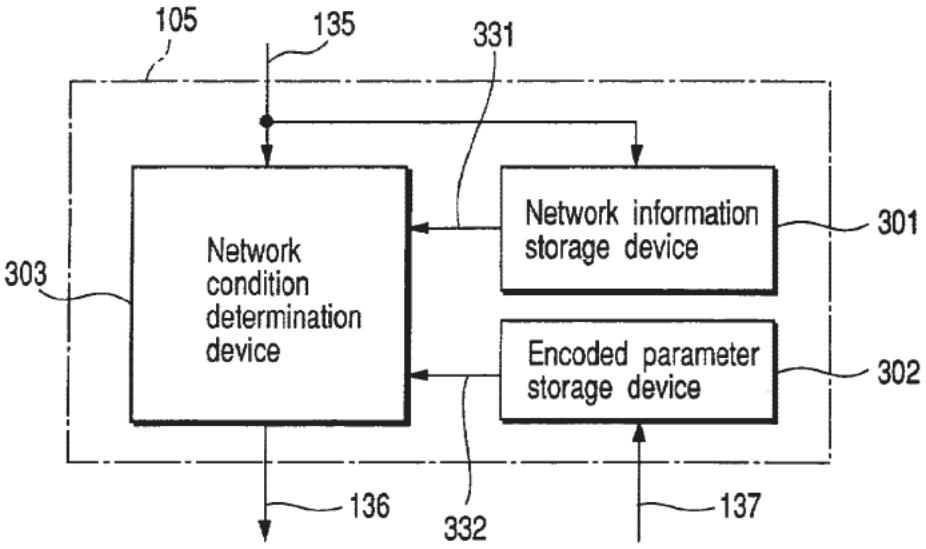


FIG. 3

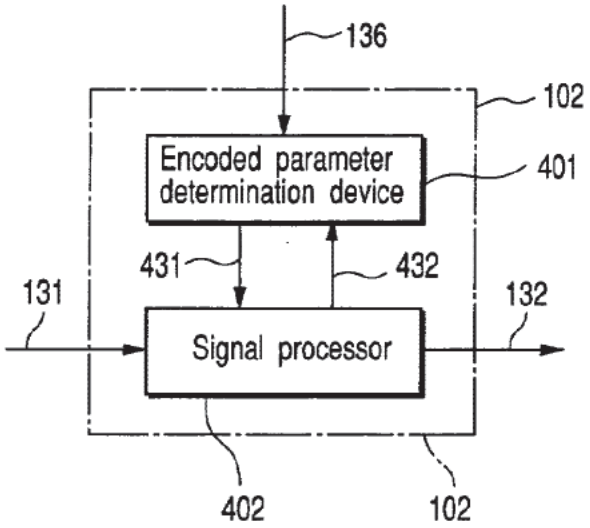


FIG. 4

Ex. 1018-1116

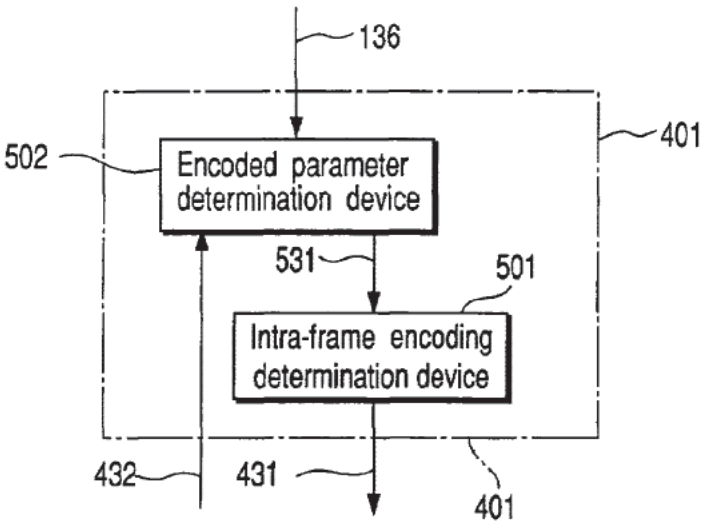


FIG. 5

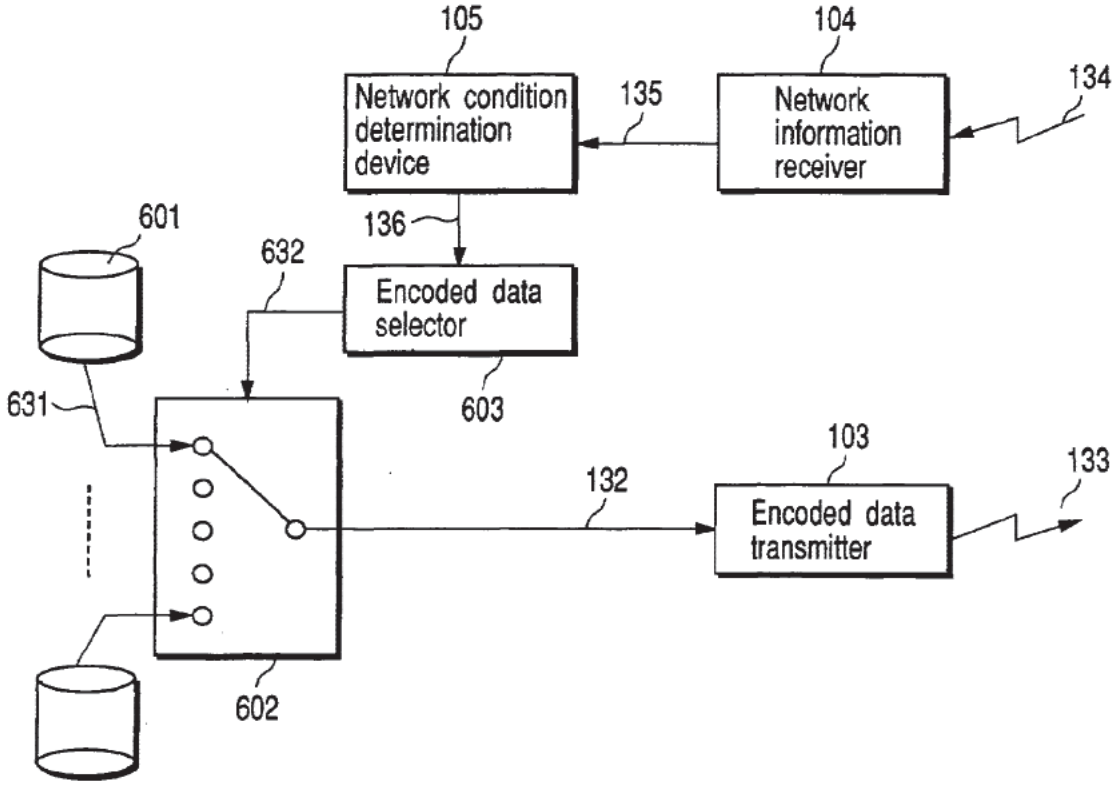


FIG. 6

Ex. 1018-1117

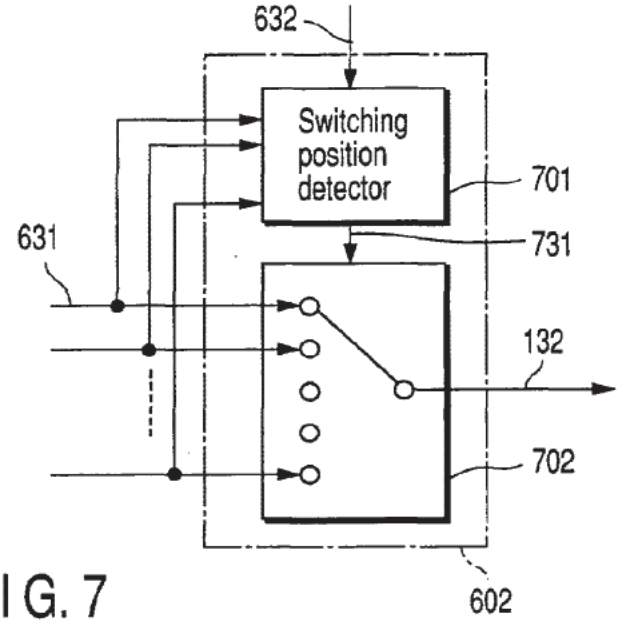


FIG. 7

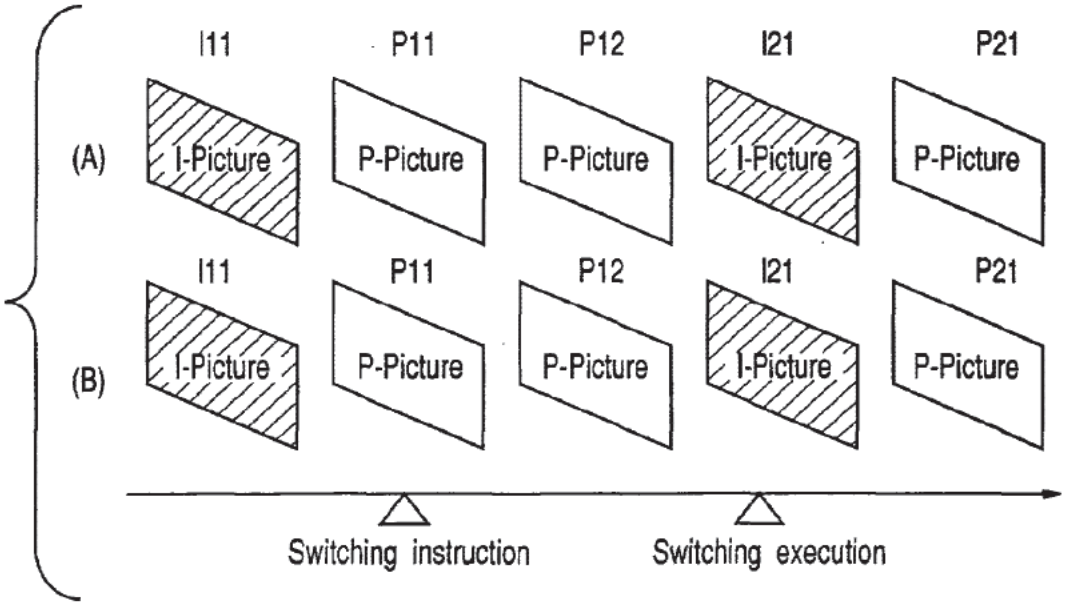


FIG. 8

Ex. 1018-1118

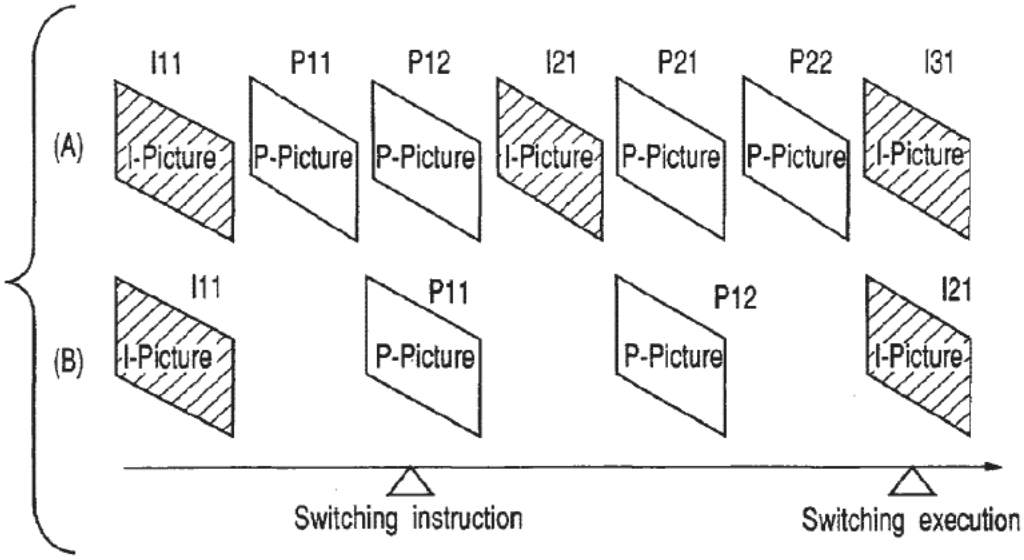


FIG. 9

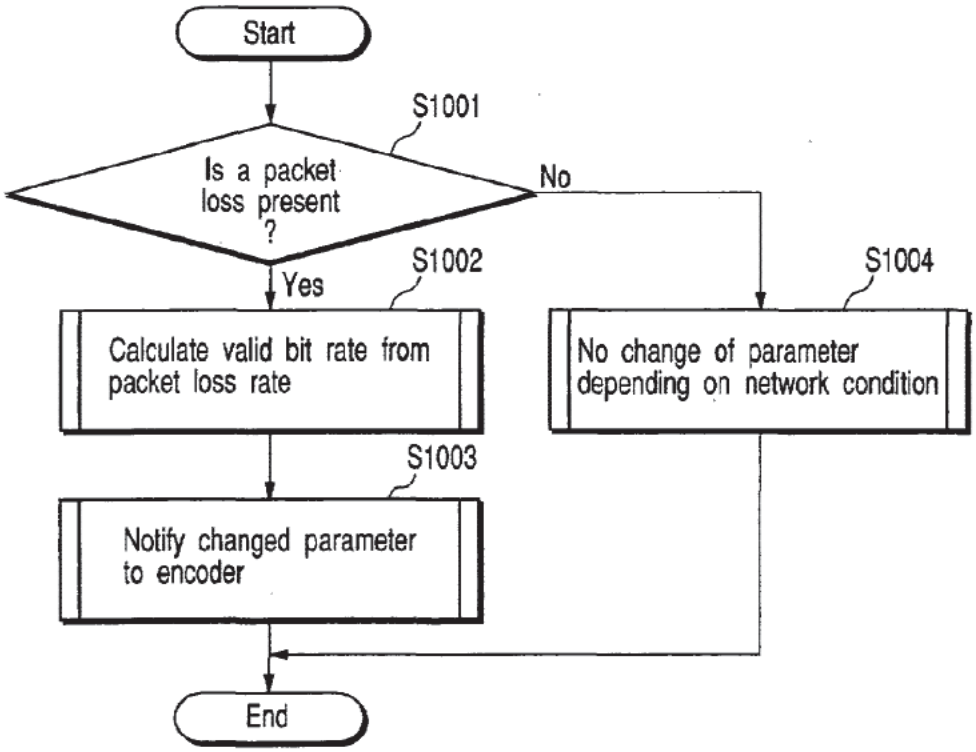


FIG. 10

Ex. 1018-1119

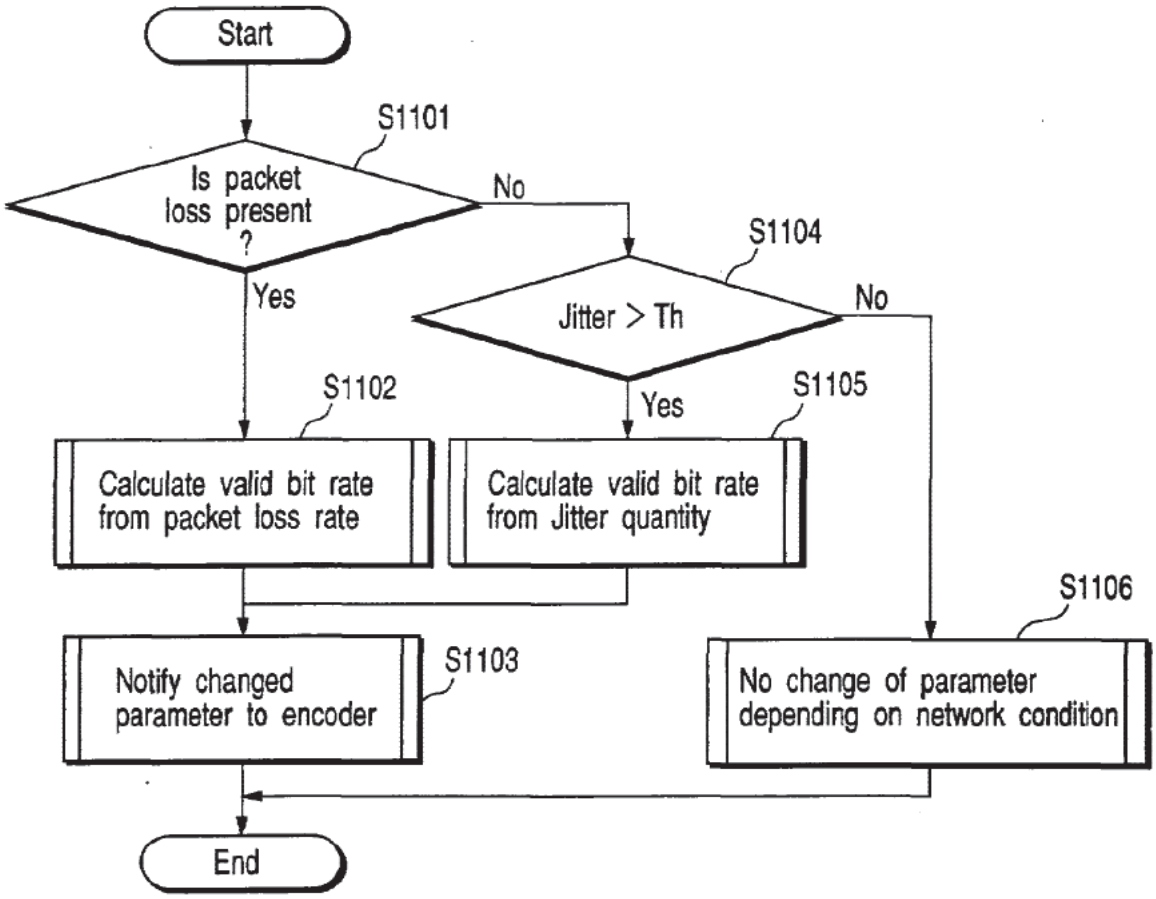


FIG. 11

Ex. 1018-1120

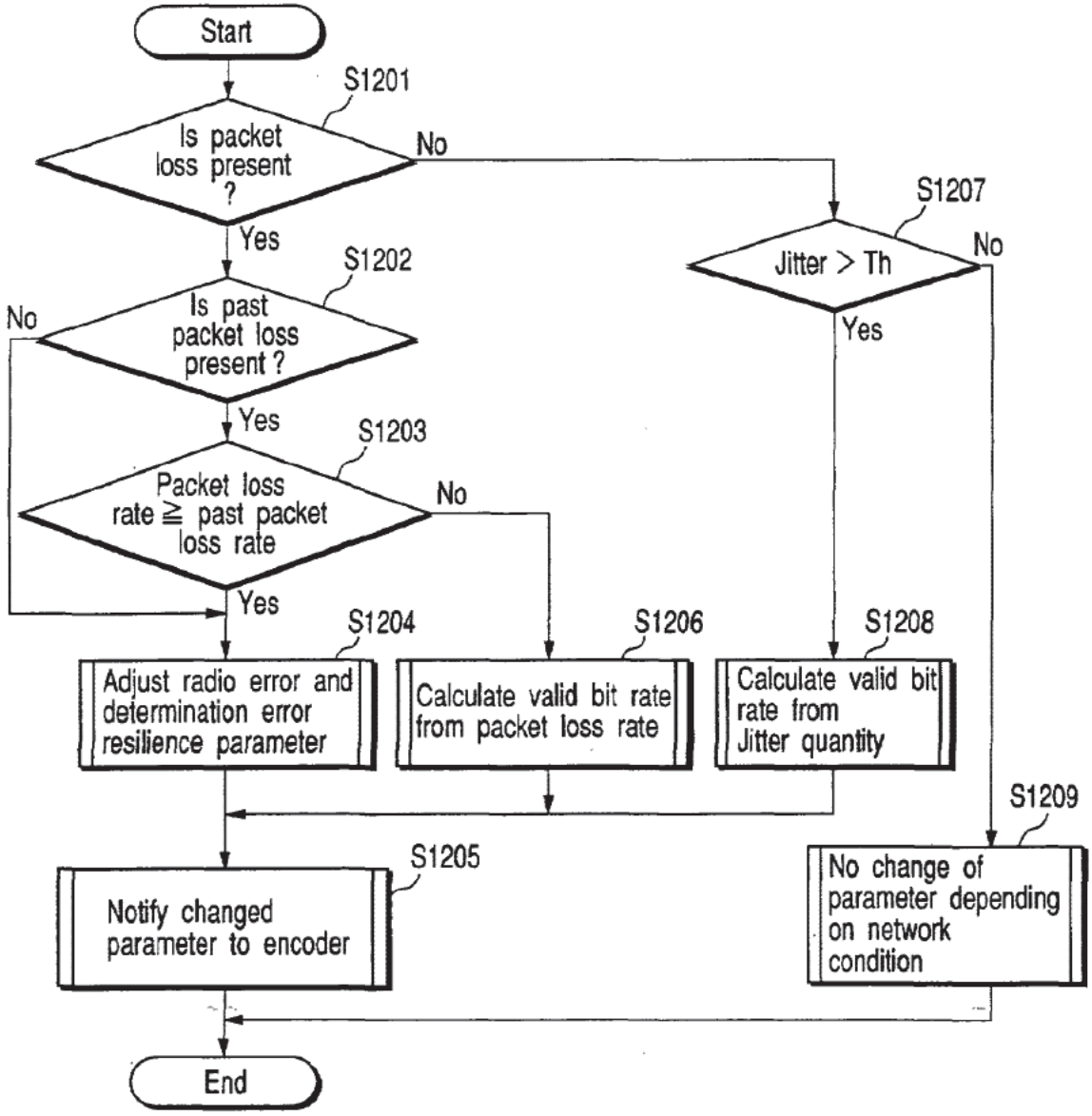


FIG. 12

Ex. 1018-1121

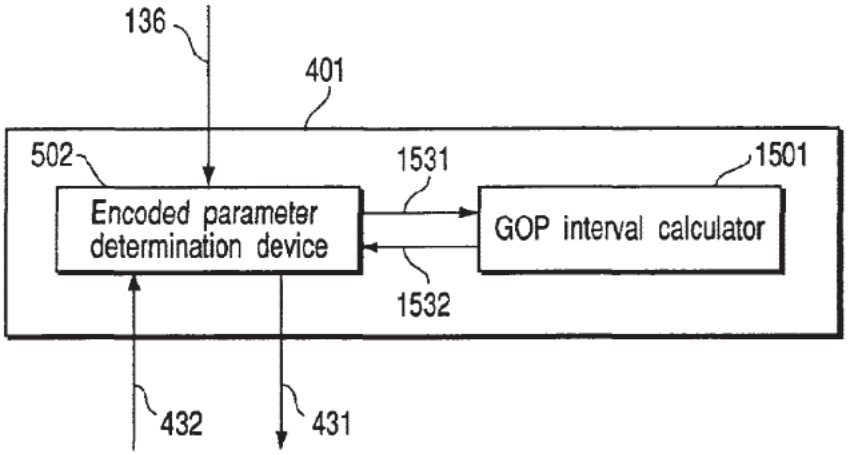


FIG. 13

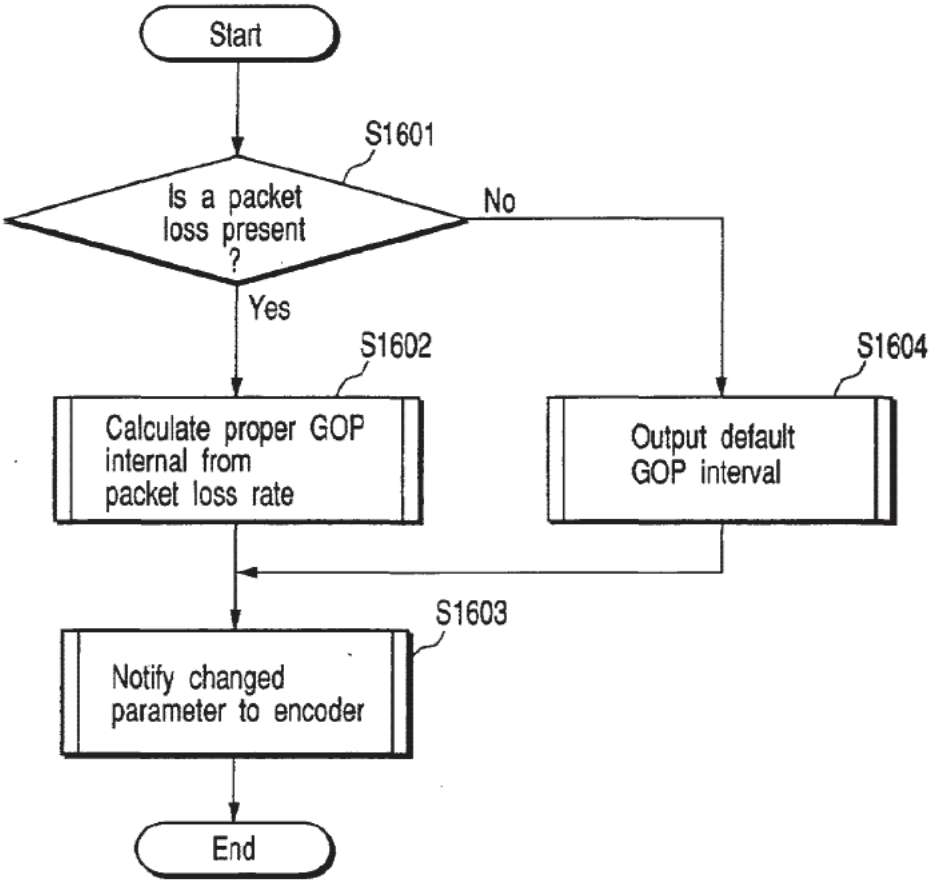


FIG. 14

Ex. 1018-1122

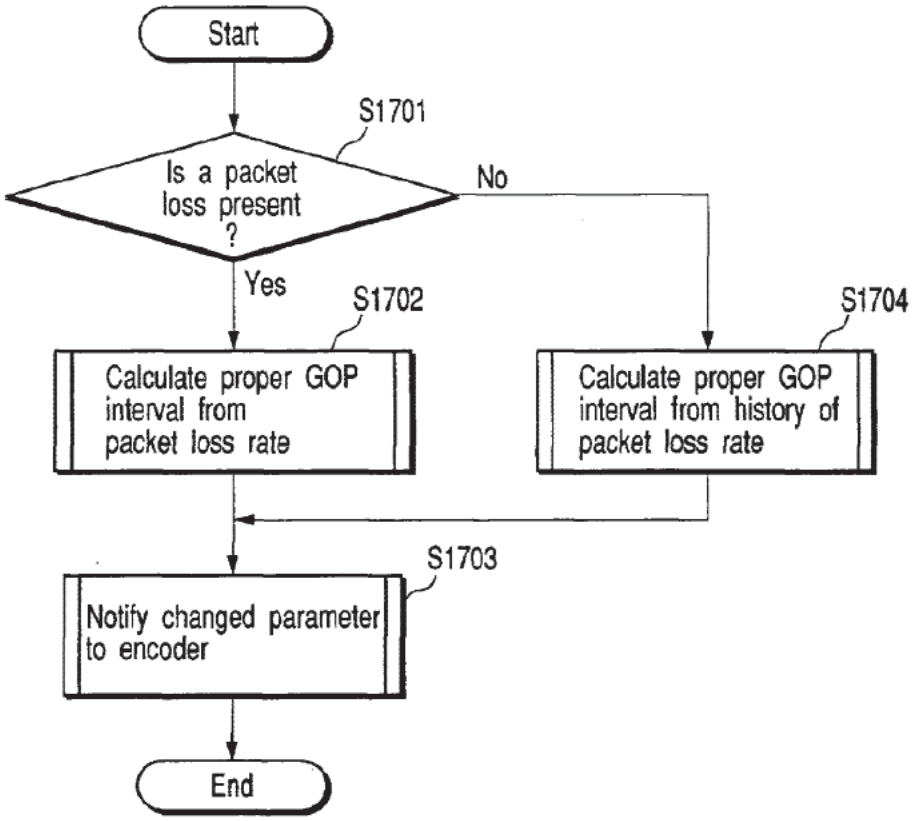


FIG. 15

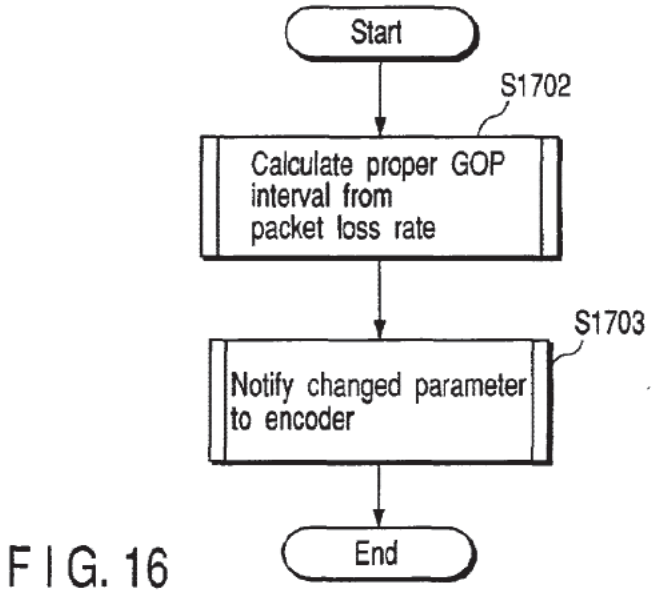


FIG. 16

Ex. 1018-1123

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 298 931 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
02.04.2003 Bulletin 2003/14

(51) Int Cl.7: **H04N 7/24**

(21) Application number: **02396145.1**

(22) Date of filing: **20.09.2002**

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

(30) Priority: **20.09.2001 FI 20011856**

(71) Applicant: **Oplayo Oy**
00180 Helsinki (FI)

(72) Inventor: **Ala-Honkola, Pekka**
00180 Helsinki (FI)

(74) Representative: **Savela, Antti-Jussi**
Patent Agency Compatent Ltd.
Hämeentie 29, 4th. FI.
00500 Helsinki (FI)

(54) Adaptive media stream

(57) This invention relates to media streams transmitted in a network, such as the Internet. The idea is to change the source of the content of a media stream when an available transmission speed rate changes. Each source contains essentially the same information (such as video and voice), but the suitability of each source for transmission has been adjusted to a certain speed rate.

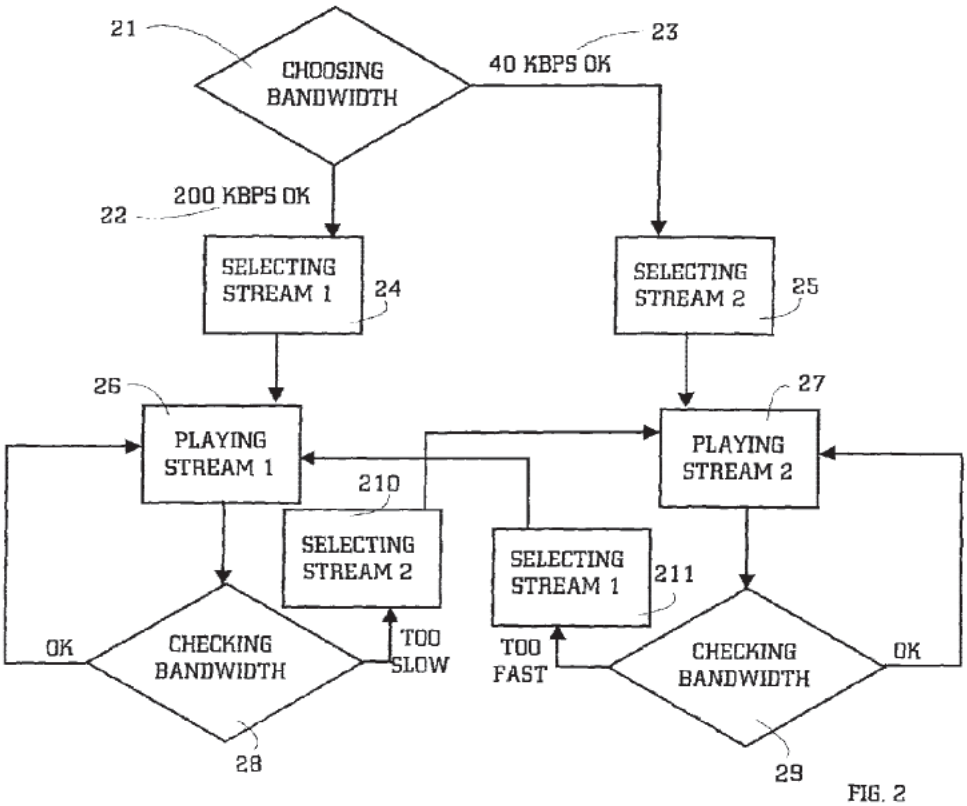


FIG. 2

Printed by Jouve, 75001 PARIS (FR)

EP 1 298 931 A2

Description

FIELD OF THE INVENTION

[0001] This invention relates to media streams transmitted in a network, such as the Internet. Especially, the invention relates to media streams concerning video pictures and voice.

BACKGROUND OF THE INVENTION

[0002] When media streams that contain continuous data, for example video pictures and voice, are transmitted in a network, such as the Internet, the speeds of the media streams have to be adjusted with the speed of the transmission connection available at the time. Often, in the case of video pictures, the adjustment is made by changing the visual quality of the pictures. The speed of a transmission connection in the Internet depends strongly on a user's Internet terminal. Some terminals support technically higher speed rates than others (usually older terminals). Thus, speed adjusted media streams must be offered for a variety of terminals.

[0003] Transmission speed may also vary with time. For example, when congestion happens in the network, the transmission speed may decrease considerably. After the congestion, the transmission speed returns to the normal speed rate.

[0004] Nowadays, most of the content of the Internet is in WWW servers, which use the HTTP protocol for transmitting desired contents to users. Originally, the HTTP protocol was not designed for continuous media streams (video, voice). The problem of the known solutions is that when a media stream has to be adjusted for a changed transmission condition, i.e. to be an adaptive media stream, a special software or protocol must be used, instead of the HTTP protocol. Real Server® (RealNetworks Inc, Seattle, Washington), Quicktime® Server (Apple Computers, Cupertino California), and Windows Media Server® (Microsoft, Redmond, Washington) are commercial server software, which support adaptive media streams. These software packages use special protocols, such as RTP (Real-Time Transport Protocol) and RTSP (Real Time Streaming Protocol), which are designed for the adaptive media streams.

[0005] However, using these software packages usually requires a user to invest and maintain parallel devices and software in addition to the WWW server that is used to serve other Internet content, such as web pages. User communities such as corporations require support staff which must be educated to use these systems. For a service provider this means in practice that a special media stream server must be maintained. The goal of this invention is to eliminate these drawbacks. This is achieved in a way described in the claims.

SUMMARY OF THE INVENTION

[0006] The following definitions should be kept in mind when reading these specifications:

Connection: A transport layer virtual circuit established between two programs for the purpose of communication. A connection can have varying transmission speed rates.

HTTP transport: An application layer transmission of data from a file on a server computer to a user terminal using HTTP 1.1 protocol. One HTTP 1.1 connection can consist of one or several HTTP 1.1 transports.

Stream: A continuous transmission of data over for example HTTP transport

[0007] An important aspect of the invention is to change the source of an adaptive media stream when an available connection speed rate changes. Each source contains essentially the same information (such as video and voice), but the suitability of each source for transmission has been adjusted to a certain connection speed rate. The coding of the source information can be different among the sources, which affects the speed rate. The picture information can be black-and-white instead of a color picture. The picture size can be smaller or larger. The video can be with or without voice. When the source is changed, also a new HTTP transmission for the new source is started for a new stream containing the information of the new source. In other words, the adaptive media stream, i.e. the adaptive stream, comprises at least one media stream for transmitting information.

[0008] In an arrangement according to the invention a source file is formed. The source file contains a plurality of sources intended for different media streams, according to such parameters as bandwidth, computing capability of the receiving terminal, etc. At the beginning of the transmission the optimal transmission speed rate at the time is chosen according to an information transfer condition. The source in the source file, which corresponds to the available speed rate, is selected to be the source of the adaptive media stream. When the available transmission speed rate changes, either to be faster or slower, a command for changing the source of the adaptive media stream, is sent from the user's terminal to the service provider's server, i.e. the source server, wherein the file of the service is. Preferably commands are sent using an HTTP protocol. When the available transmission speed rate changes again, a new command for changing the source is sent from the user's terminal to the source server. The commands are sent when it is possible to use a higher speed rate or when a lower speed rate must be used.

[0009] In one aspect, the invention thus provides a method for providing an adaptive media stream between a sending terminal and a user's terminal wherein

Ex. 1018-1125

the adaptive media stream comprises at least one media stream and the method comprises the steps of choosing a transmission speed rate for the media stream, selecting a suitable data source for the media stream among sources containing essentially the same information content, each source intended for a certain information transfer condition and playing the media stream. During such playing, checking the suitability of the actual transmission speed rate, and continuing playing the media stream if the transmission speed is still suitable or selecting a new data source, which is suitable at the moment as a response for the selection of the new data source, playing a new media stream. Those selection actions continue until the played stream stops.

[0010] Optionally, the step of choosing comprises the phase of reading a header field of a file, which contains at least two data sources, the header indicating the starting points of the sources and to which transmission speed rate each source is indented. Yet another option calls for the selecting step to include the phase of starting to read from the starting point of the selected source.

[0011] Other aspect of the invention provide for the step of monitoring to include the phase of monitoring a filling degree of a buffer for incoming data in the user's terminal, e.g. if the buffer is becoming empty or full, then selecting to continue to play the present media stream or to start playing another media stream. The information of the selection of the new data source may transmitted from the user's terminal to the sending terminal using a GET method of an HTTP protocol with an optional RANGE field. When a new stream is played, an optional phases of terminating the old media stream after starting to play the new media stream and creating the new media stream before playing the new media stream is also contemplated.

[0012] If a new stream is selected, according to another aspect of the invention, the step of reading timing marks in said adaptive media stream for defining the right point to start playing the new media stream is also performed. Further, the step of selecting the initial speed may comprise the phase of reading a header field of files, where each contains one data source, the header indicating the starting points of the source and to which transmission speed rate the source is indented.

[0013] An arrangement for providing an adaptive media stream between a sending terminal and a user's terminal, is also contemplated, wherein the adaptive media stream comprises at least one media stream and the arrangement comprises:

a) Means in the client terminal for choosing a transmission speed rate for the media stream,
b) A data structure in the sending terminal comprising at least two data sources containing essentially the same information, each source intended for a certain transmission speed rate,
c) Means in the client terminal and the sending terminal for selecting a suitable data source from among the sources for the media stream,
d) Means in the user's terminal for checking the suitability of the transmission speed rate when the media stream is played,
e) Means in the user's terminal for reselecting a suitable data source from among the sources for a new media stream.

[0014] An arrangement as described above may further comprise means for playing the data structure in a way that the playing is started from the right point of the data structure. Preferably, means a) comprises means for reading a message from the user's terminal, which defines alternative transmission speed rates. Further preferably, each source comprises timing marks, which indicate the corresponding point in the other sources. More preferably, the arrangement further comprises a GET method of an HTTP protocol with an optional RANGE field for transmitting the information of the reselection of the suitable data source from the user's terminal to the sending terminal. In another preferable embodiment, means d) comprises means for monitoring a filling degree of a buffer for incoming data in the user's terminal. It also preferable that the data structure is a file containing the data sources. Alternatively, the data structure may be a number of files, each of them containing one data source.

[0015] Means c) may further comprise means for reading a header of the data structure, which indicates the starting points of the sources and to which transmission speed rate each source is intended. Means c) may further comprise means for reading headers of the files, which indicate the starting point of the source and to which transmission speed rate the source is intended.

[0016] In yet another aspect, the invention teaches a method for storing adaptive media stream data comprising the steps of storing a plurality of media streams, each representing a substantially similar information content, and each adapted for a different set of information transfer conditions; and providing a plurality of pointers into a first of said media streams, said pointers associated with a source location in one of said media streams and pointing to at least one target location in a second of said media streams; wherein the information content of said target location is substantially a continuation of the information content prior to said source location.

[0017] Preferably, said plurality of media streams is contained in a single file. Said pointers are preferably embedded within said adaptive media stream data, and comprise target locations pointer comprises a byte offset into said single file. Alternatively, said target location pointer comprises an offset into said second source stream, within said single file. Another alternative teaches that said target locations pointers comprise a pointer to another file containing said second media stream or a portion thereof. Optionally in this case, said target location pointer further comprises a byte offset into said

Ex. 1018-1126

other file.

BRIEF DESCRIPTION OF THE DRAWINGS

[0018] In the following the invention is described in more detail by means of Figs. **1 - 3** in the attached drawings where.

FIG. **1** illustrates an example of the structure of a file according to the invention,
FIG. **2** illustrates an example of a flow chart describing the function of the inventive arrangement,
FIG. **3** illustrates an example of an arrangement according to the invention.

DETAILED DESCRIPTION OF THE INVENTION

[0019] FIG. 1 describes an example of a file structure according to the invention. The file is in the service provider's server (source server) from where users can stream it. For example the file contains a video in two different formats. At the beginning of the file there exists a header **1**, which contains starting bytes for two different media sources - the sources for the media stream. Source 1, called stream 1 contains video data in a resolution of 320*240, for a speed rate 200 kbps (kilobit per second) **2.** The starting byte **4** for that video data is byte 2000 in the file. Source 2, called stream 2 contains video data in a resolution of 160*120 optimized for a speed rate 40 kbps **3.** The starting byte of that video data **5** in the file is at byte 800000. If the network and a user's terminal support the transmission speed rate 200 kbps then stream **1** is sent to the user - otherwise stream 2 is sent. In this embodiment, all data feeds are assumed to support communication at a speed of at least 40 kbps. It should be noted that stream 1 and 2 are sources i.e. the parts of the file structures in this example, and a media stream is a continuous transmission of data of a source over a transmission protocol. The transmission protocol used can be any protocol, in this text the HTTP protocol is used.

[0020] Stream 1 is sent, i.e. streamed, to the user. Stream 1 (and the media stream for it as well) contains timing marks **6, 7, 8, 9.** The timing marks are embedded in stream 1 at a various locations, and correlate the information in the two separate resolution streams. Thus the timing marks of stream 1 in the example of Fig. 1 indicate the byte of the file where the stream is running at the time and the corresponding byte **6A, 7A, 8A, 9A,** in stream 2. If for some reason, the network can not transmit stream 1 with 200 kbps, or the user's terminal can not receive at this speed rate, the transmission of the video is changed to stream 2. For example, when the need to change the speed rate to a lower one is noticed and the most recently received (or alternatively the next) timing mark is, let's say, the mark **7** indicating time 1000 ms, it tells the corresponding byte 805000 in stream 2 to where the source of the transmission of the video transfers **11.**

[0021] Similarly, stream 2 (and the media stream for it) contains timing marks **10**, **11**, **12**, **13**, which point to the location of corresponding information in stream 1, as shown by the corresponding byte **10A, 11A**, **12A**, **13A**, in stream 1. When the capability of transmitting 200 kbps is noticed during the transmission at 40 kbps rate, the transmission of the video is changed to stream 1. For example, when the availability of the higher speed rate is noticed and the next timing mark is the mark **13** indicating time 3000 ms, the corresponding byte 77000, located in stream 1, points to the location of corresponding information in stream **1** so the source of transmission is shifted to stream 1, using data stored in location **9** onwards.

[0022] FIG. **2** shows an example of a flow chart, which illustrates a function and method of the invention. FIG. **3** shows an example of the inventive arrangement. When a user visits a home page of a service provider, he/she may find something he wants to order, such as a video show. The video player client running in the user terminal sends **34** a request message from the terminal **33** to the service provider's server **31** through the network **32**, such as the Internet, to get the show using a GET method, which is one of the methods defined in the HTTP1.1 protocol. Let the request be:

GET http://www.serviceprovider.com/video/music/song1.video. **HTTP/1.1**

[0023] Often, the connection speed rate depends on the available (and optional) connection and processing speeds of the terminal. At the beginning, the first stream which is sent works like a test stream. The test stream, i.e. the first stream preferably forms the first part of the file structure, which preferably contains data for the lowest bandwidth. Thus the first stream over the HTTP transport also contains a list of alternative speed rates. The player in the client terminal chooses **21** one of the supported speed rate from the list. (It should be noted that the server can be informed of the capability of a terminal or the available bandwidth using other techniques and the specific method of learning the link capabilities is a matter of technical choice to one skilled in the art.) In this example, the user's terminal supports speed rates of 200 kbps and 40 kbps. Let's assume that the desired file **36** is the file in FIG. **1**. If the speed rate of 200 kbps **22** is available, stream 1 is selected **24**. Otherwise, the speed rate of 40 kbps **23** is chosen and stream 2 is selected **25.** (Notice that the higher speed rate means the higher bandwidth.)

[0024] In the selection of the speed rate of 200 kbps the server starts to play **26** stream 1 beginning at byte (2000) of the file. The data and the timing marks are sent **35** to the user's terminal during the play. The terminal checks **28** frequently that the currently available bandwidth is sufficient for the selected stream. If the bandwidth is sufficient, playing stream 1 is continued. If the bandwidth becomes unavailable for stream 1, stream 2 is selected **210** and stream 2 starts to play **27**, **311** be-

Ex. 1018-1127

ginning at the byte of the file defined by the latest timing mark. The bandwidth of stream 2 is checked **29** similarly as the bandwidth of stream 1, and when the bandwidth is again available for the speed rate of 200 kbps, stream 1 is selected **211**.

[0025] In the selection of the speed rate of 40 kbps the server starts to play **25** stream 2 from the right byte (800000) of the file. The data and the timing marks are sent **35** to the user's terminal during the play. The terminal checks **29** frequently that the bandwidth is okay for the selected stream. If the bandwidth is still restricted to 40 kbps, playing stream 2 is continued. If the bandwidth rise to a level sufficient to support 200 kbps transmission, stream 1 is selected **211** and stream 1 starts to play **26**, and is sent **311** to the terminal **33** beginning at the byte of the file defined by the latest timing mark. The bandwidth of stream 1 is checked **28** similarly as the bandwidth of stream 2, and when the bandwidth is again too slow for the speed rate of 200 kbps, stream 2 is selected **210**.

[0026] The checking of the bandwidth of the streams can be made, for example, measuring a filling degree of a buffer for incoming data connection in the user's terminal. When the buffer is becoming empty, the required bandwidth for current stream is too high compared to what is available through the network connection, and the speed rate should be dropped. When the buffer is becoming full, bandwidth for current stream is lower than what is available through the network connection, and the speed rate can be raised. After the measurement, the suitable stream source is selected **210**, **211** and a GET command with an additional field RANGE is sent **310** from the user's terminal to the server for changing the stream. For example, let the GET command be GET http://www.serviceprovider.com/video/music/song3video RANGE: bytes= 805000- when the transmission speed rate changes from 200 kbps to 40 kbps.

[0027] It is worth noting that preferably, when a stream is changed, a new HTTP transport (and naturally a new media stream) is started and the old media stream is terminated after the new media stream starts playing. From the view of the user the desired file, such as a video show, seems to continue without any break because the new stream starts playing from the byte indicated by the timing mark, thus the information continuity is maintained, despite the media stream switch. So, when the source is changed, also a new HTTP transport is created for the new stream, i.e. media stream, instance and for the new source. Thus the adaptive media stream, i.e. the adaptive stream, is made up of one media stream or several consecutive media streams. Since the consecutive media streams carry essentially the same data, a user gets a visual impression of continuous shows. In other words, the adaptive media stream comprises at least one media stream for transmitting information. The HTTP 1.1 protocol provides a feature called Persistent Connections. Persistent Connection means that several requests (such as GET), can be sent via one TCP connection. Thus one connection may carry several consecutive streams. However, it should be noted that several TCP connections may be needed or used for transmitting consecutive streams.

[0028] Furthermore, it may be said that the adaptive media stream comprises several (or one if the speed rate remains the same during the whole transmission) instances (separate streams) of the same object (the adaptive media stream containing information such as video). Naturally, there are more ways to model the adaptive media stream, but common for these models are that transmission of information of the adaptive stream, (the stream of a continuous data, such as video over one or several connections) happens in one or several consecutive media streams over a transport protocol, each media stream adapted for certain conditions, and certain links or pointers periodically connect or cross connect the media streams in essentially content coherent fashion. Thus varying transmission conditions may dictate how many suitable media streams are sent forming together the adaptive stream.

[0029] For creating the inventive arrangement an initial selection module **39** (i.e. means) for choosing an original transmission speed rate and selecting a suitable source in a service provider's server is needed. In a user's terminal another estimation module **37** (i.e. means for checking transmission rate) is needed, for checking the suitability of the transmission speed rate as reselection module **38** (i.e. means) for reselecting a suitable source are needed.

[0030] Although, the processing and network connection speed of a user's terminal is mostly the decisive element when selecting a suitable transmission speed rate, it may happen that the network itself becomes congested. In congestion, the available speed rate decreases and the transmission speed rate must be lowered. On the other hand, if the network allows a higher speed, the transmission speed rate may be increased.

[0031] Using the inventive arrangement, an adaptive media stream can be served from with a common HTTP server. Therefore, there is no need for the use of a special media stream server, or parallel devices and software. Remarkable cost savings are achieved, since usually service providers and users support the HTTP protocol. Also the reliability of the system increases because of fewer sources of faults.

[0032] It is clear that inventive arrangement can be implemented in many ways. For example, the time period between two timing marks can be different than in the example of this text. It can be any suitable period depending on the features of a solution, such as 2 or 4 seconds. The format of the inventive file can be any suitable format as well, such as MPEG or another.

[0033] The inventive file may comprise more than two different streams (containing the same content), from which the most suitable stream is selected at any given time. Therefore, timing marks must contain a number of references for the corresponding byte in the other data

Ex. 1018-1128

sources.

[0034] All the sources do not have to be in the same file, but each source can be a separate file. However, the use of this kind of structure is not as efficient when attempting to use as little processing power as possible compared to the use of one file structure. In the structure of several files, timing marks must indicate the right file with the corresponding byte.

[0035] Although, a sending terminal have been described to be a server in the examples above, it is clear that the sending terminal can be another terminal capable of sending information.

[0036] Similarly other mechanisms are known and suitable to measure and select the most appropriate source. Such methods may be server based or user terminal based, or even utilize intermediate network nodes to estimate network congestion and bandwidth, and similar measurements that are indicative of total node to node capacity, as well as specific node capabilities, and the selection of sources (and media streams) may be carried out in various ways that will provide the best end-user experience according to varying conditions.

[0037] According to the above-mentioned matters, it is evident that the invention is not restricted to the solutions described in this text, but it can be modified for various solutions, in the scope of the inventive idea.

Claims

1. A method for providing an adaptive media stream between a sending terminal and a user's terminal **characterized in that** the adaptive media stream comprises at least one media stream and the method comprises the steps of:

    a) choosing a transmission speed rate for the media stream,
    b) selecting a suitable data source for the media stream among sources containing essentially the same information, each source intended for a certain transmission speed rate,
    c) playing the media stream,
    d) checking the suitability of the transmission speed rate,
    e) either continuing playing the media stream if the transmission speed is still suitable or,
    f) selecting a new data source, which is suitable at the moment,
    g) as a response for the selection of the new data source, playing a new media stream,
    h) repeating steps d) to g) until the stream, which is played, stops.

2. A method according to claim 1 **characterized in that** step b) comprises the phase of reading a header field of a file, which contains at least two data sources, the header indicating the starting points of the sources and to which transmission speed rate each source is intended.

3. A method according to claim 2 **characterized in that** step c) comprises the phase of starting to read from the starting point of the selected source.

4. A method according to claim 1 or 3 **characterized in that** step d) comprises the phase of monitoring a filling degree of a buffer for incoming data in the user's terminal.

5. A method according to claim 4 **characterized in that** if the buffer is becoming empty or full, then step f) is chosen, else step e) is chosen.

6. A method according to claim 1 or 5 **characterized in that** the information of the selection of the new data source is transmitted from the user's terminal to the sending terminal using a GET method of an HTTP protocol with an optional RANGE field.

7. A method according to claim 1 **characterized in that** step g) comprises phases of terminating the old media stream after starting to play the new media stream and creating the new media stream before playing the new media stream.

8. A method according to claim 1 or 7, **characterized in that** step g) further comprises the phase of reading timing marks in the data source for defining the right point to start playing the new media stream.

9. A method according to claim 1 **characterized in that** step b) comprises the phase of reading a header field of files, where each contains one data source, the header indicating the starting points of the source and to which transmission speed rate the source is intended.

10. An arrangement for providing an adaptive media stream between a sending terminal and a user's terminal, **characterized in that** the adaptive media stream comprises at least one media stream and the arrangement comprises:

    a) means in the sending terminal for choosing a transmission speed rate for the media stream,
    b) a data structure in the sending terminal comprising at least two data sources containing essentially the same information, each source intended for a certain transmission speed rate,
    c) means in the sending terminal for selecting a suitable data source from among the sources for the media stream,
    d) means in the user's terminal for checking the suitability of the transmission speed rate when the media stream is played,

Ex. 1018-1129

e) means in the user's terminal for reselecting a suitable data source from among the sources for a new media stream.

11. An arrangement according to claim 10, **characterized in that** the arrangement further comprises means for playing the data structure in a way that the playing is started from the right point of the data structure.

12. An arrangement according to claim 11, **characterized in that** means a) comprises means for reading a message from the user's terminal, which defines alternative transmission speed rates.

13. An arrangement according to claim 11, **characterized in that** each source comprises timing marks, which indicate the corresponding point in the other sources.

14. An arrangement according to claim 13, **characterized in that** the arrangement further comprises a GET method of an HTTP protocol with an optional RANGE field for transmitting the information of the reselection of the suitable data source from the user's terminal to the sending terminal.

15. An arrangement according to claim 11, **characterized in that** means d) comprises means for monitoring a filling degree of a buffer for incoming data in the user's terminal.

16. An arrangement according to claim 11, **characterized in that** the data structure is a file containing the data sources.

17. An arrangement according to claim 16, **characterized in that** means c) comprises means for reading a header of the data structure, which indicates the starting points of the sources and to which transmission speed rate each source is intended.

18. An arrangement according to claim 11, **characterized in that** the data structure is a number of files, each of them containing one data source.

19. An arrangement according to claim 18, **characterized in that** means c) comprises means for reading headers of the files, which indicate the starting point of the source and to which transmission speed rate the source is intended.

Ex. 1018-1130

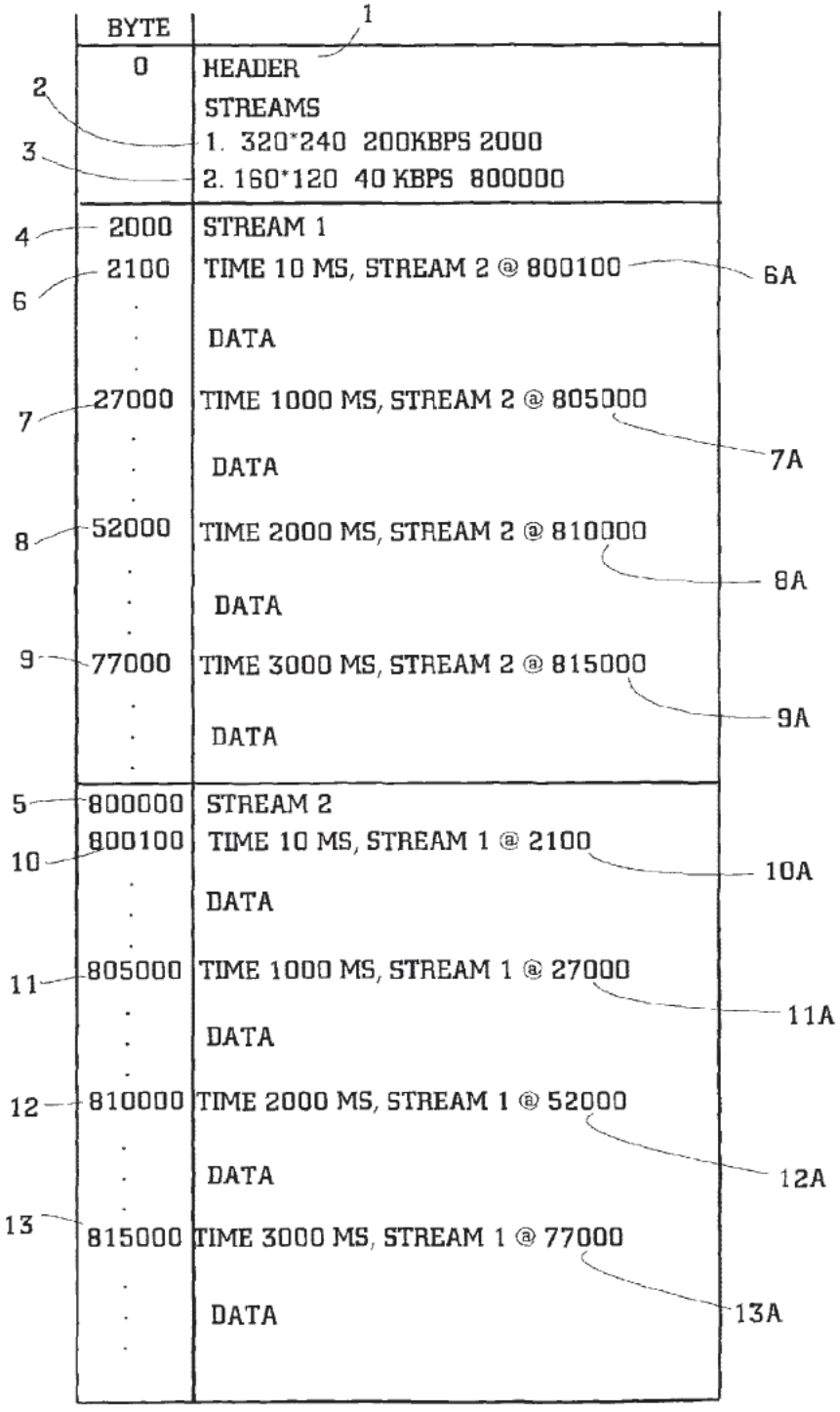


FIG. 1

Ex. 1018-1131

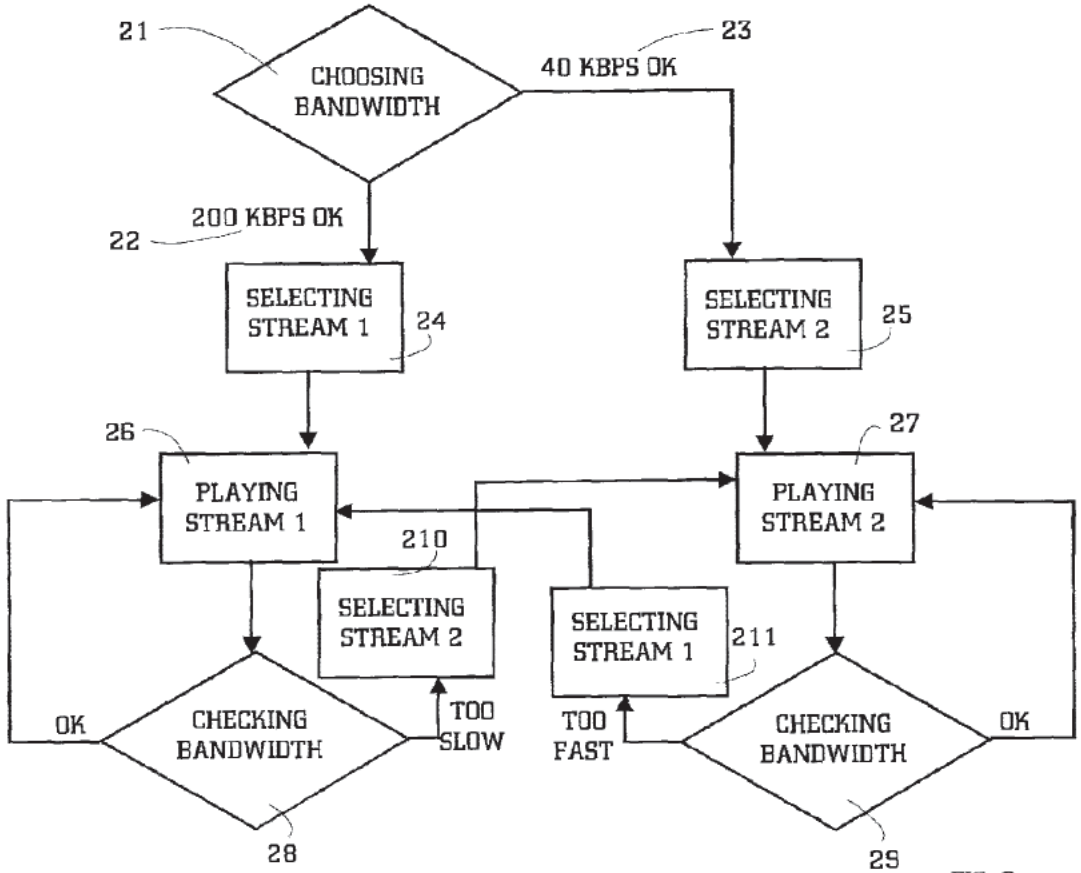

21
CHOOSING BANDWIDTH
23
40 KBPS OK
22
200 KBPS OK
SELECTING STREAM 1
24
SELECTING STREAM 2
25
26
PLAYING STREAM 1
27
PLAYING STREAM 2
210
SELECTING STREAM 2
SELECTING STREAM 1
211
OK
CHECKING BANDWIDTH
TOO SLOW
TOO FAST
CHECKING BANDWIDTH
OK
28
29

FIG. 2

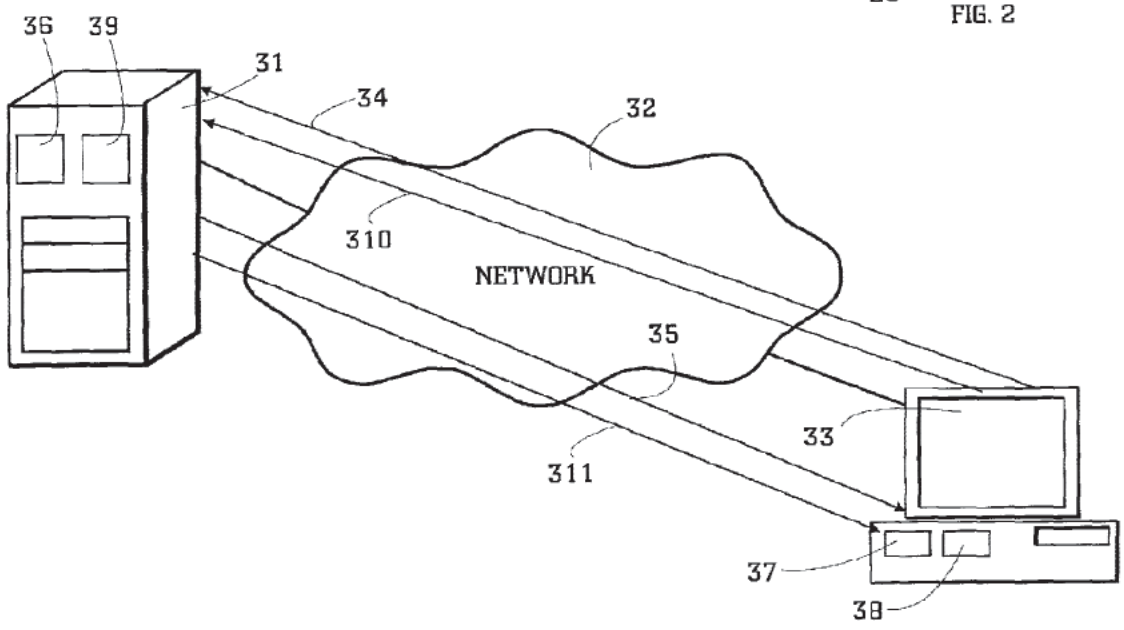

36
39
31
34
32
310
NETWORK
35
33
311
37
38

FIG. 3

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-201343
(43)Date of publication of application : 18.07.2000

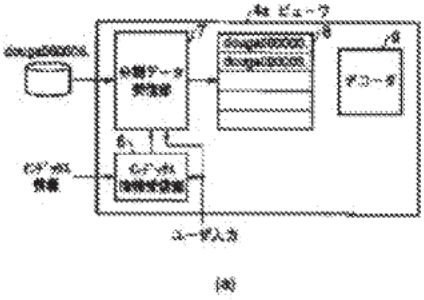
(51)Int.Cl. H04N 7/173

(21)Application number : 11-000497
(22)Date of filing : 05.01.1999

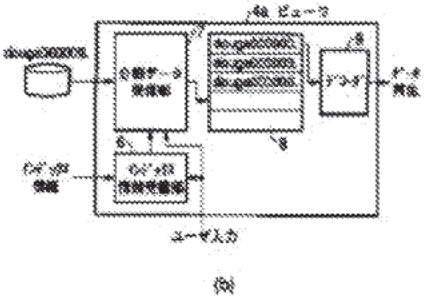
(71)Applicant : TOSHIBA CORP
(72)Inventor : CHIJITANI MASAHIDE

(54) MOVING IMAGE DATA REPRODUCING DEVICE, MOVING IMAGE DATA DIVIDING DEVICE AND RECORDING MEDIUM READABLE BY COMPUTER IN WHICH PROGRAM IS RECORDED

(57)Abstract:
PROBLEM TO BE SOLVED: To enable random access to moving image data without changing a server function.
SOLUTION: A moving image data reproducing device with an index information receiving means 6 to receive index information to be generated by relating plural pieces of divided data to be obtained by deviding the moving image data to its reproduction time information by requesting the information to a server, a divided data receiving means 7 to receive the divided data related to reproduction time information at the reproduction time information to which starting of reproducing is instructed and afterward by successively requesting the information to the server by referring to the index information received by the index information receiving means 6, a storage means 9 to store the divided data received by the divided data receiving means 7 and a reproducing means to successively reproduce the divided data stored in the storage means 9 according to the index information received by the index information receiving means 6 is provided.



(a)



(b)

Ex. 1018-1133

(19)日本国特許庁（ＪＰ）　(12) **公開特許公報（Ａ）**　(11)特許出願公開番号
特開2000－201343
（P2000－201343A）
(43)公開日　平成12年７月18日(2000.7.18)

(51)Int.Cl.[7]	識別記号	ＦＩ		テーマコード（参考）
Ｈ０４Ｎ　7/173	６３０	Ｈ０４Ｎ　7/173	６３０	５Ｃ０６４

審査請求　未請求　請求項の数13　ＯＬ　（全 13 頁）

(21)出願番号　特願平11－497

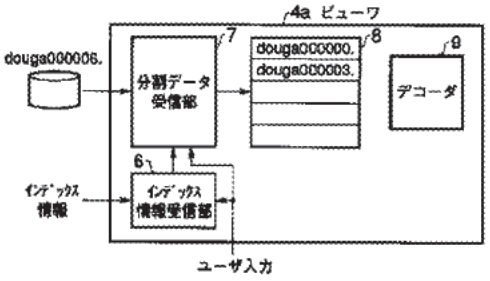
(22)出願日　平成11年１月５日(1999.1.5)

(71)出願人　000003078
株式会社東芝
神奈川県川崎市幸区堀川町72番地
(72)発明者　千々谷　眞英
東京都府中市東芝町１番地　株式会社東芝府中工場内
(74)代理人　100058479
弁理士　鈴江　武彦　（外６名）

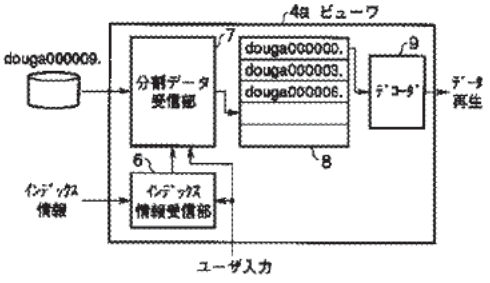
Ｆターム(参考)　5C064 BA07 BB10 BC16 BC20 BC27 BD16

(54)【発明の名称】　動画像データ再生装置及び動画像データ分割装置並びにプログラムを記録したコンピュータ読み取り可能な記録媒体

(57)【要約】
【課題】サーバの機能を変更することなく、動画像データのランダムアクセスを可能にする。
【解決手段】動画像データが分割されて得られる複数の分割データとその再生時刻情報とを関連付けて生成されるインデックス情報を、サーバに要求して受信するインデックス情報受信手段と、当該インデックス情報受信手段に受信されたインデックス情報を参照し、再生開始を指示された再生時刻情報以降の再生時刻情報に関連付けられた分割データをサーバに順次要求して受信する分割データ受信手段と、当該分割データ受信手段に受信された分割データを記憶する記憶手段と、インデックス情報受信手段に受信されたインデックス情報にしたがって、記憶手段に記憶されている分割データを順次再生させる再生手段とを具備した動画像データ再生装置を提供する。



(a)



(b)

1

【特許請求の範囲】

【請求項１】　動画像データが分割されて得られる複数の分割データとそれぞれの再生時刻情報とを関連付けて生成されるインデックス情報を、サーバに要求して受信するインデックス情報受信手段と、

当該インデックス情報受信手段に受信されたインデックス情報にしたがって、再生開始を指示された再生時刻情報及びその後の再生時刻情報に関連付けられた分割データを、サーバに順次要求して受信する分割データ受信手段と、

当該分割データ受信手段によって受信された分割データを記憶する記憶手段と、

前記記憶手段に記憶されている分割データを順次再生させる再生手段とを具備したことを特徴とする動画像データ再生装置。

【請求項２】　請求項１記載の動画像データ再生装置において、

前記分割データの送信をサーバに要求してから、この分割データを受信するまでの時間を推定する時間推定手段を付加し、

前記分割データ受信手段は、当該時間推定手段によって推定された時間とほぼ同じ時間分の分割データが前記記憶手段に記憶されていない場合に、前記インデックス情報にしたがってサーバに分割データを要求し、

前記再生手段は、前記時間推定手段によって推定された時間とほぼ同じ時間分の分割データが前記記憶手段に記憶された場合に、この記憶手段に記憶されている分割データの再生を開始することを特徴とする動画像データ再生装置。

【請求項３】　動画像データが分割されて得られる複数の分割データについて伝送速度が異なる複数の種別が生成されている場合に、この複数種別の各分割データとそれぞれの再生時刻情報とを関連付けて生成されるインデックス情報を、サーバに要求して受信するインデックス情報受信手段と、

前記分割データの送信をサーバに要求してから、この分割データを受信するまでの時間を推定する時間推定手段と、

当該時間推定手段によって推定された時間に基づいて、要求する分割データの種別を選択する分割データ選択手段と、

前記インデックス情報受信手段に受信されたインデックス情報にしたがって、再生開始を指示された再生時刻情報及びその後の再生時刻情報に関連付けられ、かつ前記分割データ選択手段によって選択されている種別の分割データを、サーバに順次要求して受信する分割データ受信手段と、

当該分割データ受信手段によって受信された分割データを記憶する記憶手段と、

前記記憶手段に記憶されている分割データを順次再生させる再生手段とを具備したことを特徴とする動画像データ再生装置。

2

【請求項４】　請求項３記載の動画像データ再生装置において、

前記分割データ受信手段は、前記時間推定手段に推定された時間とほぼ同じ時間分の分割データが前記記憶手段に記憶されていない場合に、前記データ選択手段によって選択されている種別の分割データを、前記インデックス情報にしたがって前記サーバに要求し、

前記再生手段は、前記時間推定手段によって推定された時間とほぼ同じ時間分の分割データが前記記憶手段に記憶された場合に、この記憶手段に記憶されている分割データの再生を開始することを特徴とする動画像データ再生装置。

【請求項５】　動画像データを複数に分割して複数の分割データを生成し、かつ前記複数の分割データとこの分割データの再生時刻情報とを関連付けたインデックス情報を生成する動画像データ分割装置。

【請求項６】　動画像データが分割されて得られる複数の分割データについて伝送速度が異なる複数の種別を生成し、この複数の種別の各分割データをそれぞれの再生時刻情報に関連付けたインデックス情報を生成する動画像データ分割装置。

【請求項７】　コンピュータに、

動画像データが分割されて得られる複数の分割データとそれぞれの再生時刻情報とを関連付けて生成されるインデックス情報を、サーバに要求して受信するインデックス情報受信機能と、

当該インデックス情報受信機能によって受信されたインデックス情報にしたがって、再生開始を指示された再生時刻情報及びその後の再生時刻情報に関連付けられた分割データを、サーバに順次要求して受信する分割データ受信機能と、

当該分割データ受信機能によって受信された分割データを記憶する記憶機能と、

前記記憶機能によって記憶されている分割データを順次再生させる再生機能とを実現させるためのプログラムを記録したコンピュータ読み取り可能な記録媒体。

【請求項８】　請求項７記載のコンピュータ読み取り可能な記録媒体において、

前記分割データの送信をサーバに要求してから、この分割データを受信するまでの時間を推定する時間推定機能を付加し、

前記分割データ受信機能は、当該時間推定機能によって推定された時間とほぼ同じ時間分の分割データが前記記憶機能によって記憶されていない場合に、前記インデックス情報にしたがってサーバに分割データを要求し、

前記再生機能は、前記時間推定機能によって推定された時間とほぼ同じ時間分の分割データが前記記憶機能によって記憶された場合に、この記憶機能によって記憶され

ている分割データの再生を開始することを特徴とするコンピュータ読み取り可能な記録媒体。

【請求項９】　コンピュータに、
動画像データが分割されて得られる複数の分割データについて伝送速度が異なる複数の種別が生成されている場合に、この複数種別の各分割データとそれぞれの再生時刻情報とを関連付けて生成されるインデックス情報をサーバに要求して受信するインデックス情報受信機能と、
前記分割データの送信をサーバに要求してから、この分割データを受信するまでの時間を推定する時間推定機能と、
当該時間推定機能によって推定された時間に基づいて、要求する分割データの種別を選択する分割データ選択機能と、
前記インデックス情報受信機能によって受信されたインデックス情報にしたがって、再生開始を指示された再生時刻情報及びその後の再生時刻情報に関連付けられ、かつ前記分割データ選択機能によって選択されている種別の分割データを、サーバに順次要求して受信する分割データ受信機能と、
当該分割データ受信機能によって受信された分割データを記憶する記憶機能と、
前記記憶機能によって記憶されている分割データを順次再生させる再生機能とを実現させるためのプログラムを記録したコンピュータ読み取り可能な記録媒体。

【請求項１０】　請求項９記載のコンピュータ読み取り可能な記録媒体において、
前記分割データ受信機能は、前記時間推定機能によって推定された時間とほぼ同じ時間分の分割データが前記記憶機能に記憶されていない場合に、前記データ選択機能によって選択されている種別の分割データを、前記インデックス情報にしたがって前記サーバに要求し、
前記再生機能は、前記時間推定機能によって推定された時間とほぼ同じ時間分の分割データが前記記憶機能によって記憶された場合に、この記憶機能によって記憶されている分割データの再生を開始することを特徴とするコンピュータ読み取り可能な記録媒体。

【請求項１１】　コンピュータに、
動画像データを複数に分割して複数の分割データを生成し、かつ前記複数の分割データとこの分割データの再生時刻情報とを関連付けたインデックス情報を生成する機能を実現させるためのプログラムを記録したコンピュータ読み取り可能な記録媒体。

【請求項１２】　コンピュータに、
動画像データが分割されて得られる複数の分割データについて伝送速度が異なる複数の種別を生成し、この複数の種別の各分割データをそれぞれの再生時刻情報に関連付けたインデックス情報を生成する機能を実現させるためのプログラムを記録したコンピュータ読み取り可能な記録媒体。

【請求項１３】　サーバに記憶されている動画像データをクライアントに配信する動画像データ配信システムにおいて、
前記サーバは、
前記動画像データが分割されて得られる複数の分割データと、この各分割データとそれぞれの再生時刻情報とを関連付けたインデックス情報とを保持し、
前記クライアントは、
前記インデックス情報を前記サーバに要求して受信するインデックス情報受信手段と、
当該インデックス情報受信手段に受信されたインデックス情報にしたがって、再生開始を指示された再生時刻情報及びその後の再生時刻情報に関連付けられた分割データを前記サーバに順次要求して受信する分割データ受信手段と、
当該分割データ受信手段によって受信された分割データを記憶する記憶手段と、
前記記憶手段に記憶されている分割データを順次再生させる再生手段とを備えることを特徴とする動画像データ配信システム。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】本発明は、サーバの構成及び機能を変更することなく、動画像データのランダムアクセスを可能にするための動画像データ再生装置及び動画像データ分割装置並びにプログラムを記録したコンピュータ読み取り可能な記録媒体に関する。

【０００２】

【従来の技術】従来の動画像データ配信システムは、主にクライアントとＷＷＷサーバとから構成される。この従来の動画像データ配信システムでは、まずクライアントからＷＷＷサーバに再生を求める動画像データ（動画コンテンツ）の伝送要求が送信される。

【０００３】ＷＷＷサーバは各種データを保持しており、このデータの中から要求された動画像データを選択し送信する。

【０００４】送信された動画像データは、伝送要求を送信したクライアントに受信され、ストリーム再生がなされる。

【０００５】このような従来の動画像データ配信システムにおいて、クライアント及びＷＷＷサーバがインターネットを介して接続されている場合には、一般的にＨＴＴＰプロトコルにより動画像データの伝送がなされる。

【０００６】ここで、例えば動画像データの再生を求めるユーザの中には、動画像データの途中位置からの再生（以下、「ランダムアクセス」という）を求める者もいる。

【０００７】従来の動画像データ配信システムによってランダムアクセスを実現するためには、ランダムアクセスを可能とする専用のＨＴＴＰサーバをこの動画像デー

タ配信システムに適用する必要がある。

【０００８】また、別の手法として、ランダムアクセスを可能とする専用のプログラムであるＣＧＩ（Common Gateway Interface）をWWWサーバに実装させる必要がある。

【０００９】この専用のＨＴＴＰサーバや、ＣＧＩを実装したWWWサーバは、動画像データを再生するクライアントから独自の要求を受信し、動画像データを指定された位置からＨＴＴＰプロトコルによって送信する機能を持つ。

【００１０】

【発明が解決しようとする課題】しかしながら、上記のように、ランダムアクセスを可能とするために専用のＨＴＴＰサーバを用いたり、サーバにＣＧＩを実装する方法においては、従来の動画像データ配信システムのWWWサーバに手を加える必要があり、現状のインフラのままでは実現することができない。

【００１１】したがって、ランダムアクセスの実現、導入が困難であり、作業が複雑である。

【００１２】また、この従来の方法では、ランダムアクセスのための処理が主にサーバでなされるため、サーバの負荷が増加するという問題が生じる。特に、動画像データ配信システムが複数のクライアントを有する場合には、その傾向が強くなる。

【００１３】本発明は、以上のような実状に鑑みてなされたもので、サーバの構成や機能に手を加えることなく動画像データのランダムアクセスを可能とするための動画像データ再生装置及び動画像データ分割装置並びにプログラムを記録したコンピュータ読み取り可能な記録媒体を提供することを目的とする。

【００１４】

【課題を解決するための手段】本発明の骨子は、動画像データを複数の分割して分割データとし、この分割データにそれぞれの再生時間を示す再生時刻情報を関連づけてインデックス情報とし、このインデックス情報にしたがって分割データを順次再生する点にある。

【００１５】以下、本発明を実現するにあたって講じた具体的手段について説明する。

【００１６】第１の発明は、動画像データが分割されて得られる複数の分割データとそれぞれの再生時刻情報とを関連付けて生成されるインデックス情報をサーバに要求して受信するインデックス情報受信手段と、当該インデックス情報受信手段に受信されたインデックス情報にしたがって、再生開始を指示された再生時刻情報及びその後の再生時刻情報に関連付けられた分割データをサーバに順次要求して受信する分割データ受信手段と、当該分割データ受信手段によって受信された分割データを記憶する記憶手段と、インデックス情報受信手段に受信されたインデックス情報の再生時刻情報にしたがって、記憶手段に記憶されている分割データを順次再生させる再生手段とを具備した動画像データ再生装置である。

【００１７】この第１の発明の動画像データ再生装置においては、再生開始を指示された再生時刻情報以降の分割データが順次受信されストリーム再生される。

【００１８】また、この動画像データ再生装置を適用する場合には、サーバは自己が保持するデータのうち要求されたデータを送信する機能を有していればよい。

【００１９】したがって、第１の発明においては、サーバに新しい構成や機能を付加することなく容易にランダムアクセスが実現される。

【００２０】また、ランダムアクセスを実現させるための処理が、サーバではなく動画像データ再生装置（クライアント）側によって実行されるため、サーバの負荷の増加を防止することができる。

【００２１】第２の発明は、第１の発明の動画像データ再生装置において、分割データの送信をサーバに要求してから、この分割データを受信するまでの時間を推定する時間推定手段を付加し、分割データ受信手段は、当該時間推定手段によって推定された時間とほぼ同じ時間分の分割データが記憶手段に記憶されていない場合に、前記インデックス情報にしたがってサーバに分割データを要求し、再生手段は、時間推定手段によって推定された時間とほぼ同じ時間分の分割データが記憶手段に記憶された場合に、この記憶手段に記憶されている分割データの順次再生を開始する動画像データ再生装置である。

【００２２】この第２の発明の動画像データ再生装置においては、分割データを要求してから受信するまでの間に再生されると推測される分割データが、記憶手段に保持される。

【００２３】したがって、この第２の発明においては、第１の発明と同様の作用効果に加えて、必要最低限の量の分割データが格納されしだい、動画像データの再生がなされるため、再生が開始されるまでの時間を最適化することができる。

【００２４】これにより、再生開始までの時間を短縮させることができる。

【００２５】また、次の分割データが要求されてから受信されるまでの間に必要な分割データは確保済みであるため、動画像データの再生が途切れることを防止することができる。

【００２６】さらに、分割データをサーバに要求してから受信するまでの時間が逐次測定され、この時間にしたがって再生可能な状態で記憶しておく分割データの量が決定される。

【００２７】ゆえに、回線の動的な負荷変動にも対応することができ、適切な量の分割データが確保される。

【００２８】したがって、スムーズなストリーム再生が可能となる。

【００２９】第３の発明は、動画像データが分割されて得られる複数の分割データについて伝送速度が異なる複

数の種別が生成されている場合に、この複数種別の各分割データとそれぞれの再生時刻情報とを関連付けて生成されるインデックス情報をサーバに要求して受信するインデックス情報受信手段と、分割データの送信をサーバに要求してから、この分割データを受信するまでの時間を推定する時間推定手段と、当該時間推定手段によって推定された時間に基づいて、要求する分割データの種別を選択する分割データ選択手段と、インデックス情報受信手段に受信されたインデックス情報にしたがって、再生開始を指示された再生時刻情報及びその後の再生時刻情報に関連付けられ、かつ分割データ選択手段によって選択されている種別の分割データをサーバに順次要求して受信する分割データ受信手段と、当該分割データ受信手段によって受信された分割データを記憶する記憶手段と、インデックス情報受信手段に受信されたインデックス情報の再生時刻情報にしたがって、記憶手段に記憶されている分割データを順次再生させる再生手段とを具備した動画像データ再生装置である。

【００３０】この第３の発明の動画像データ再生装置においては、伝送速度の異なる複数種別の分割データのうち、回線の負荷変動を考慮して適した伝送速度の分割データが選択され受信される。

【００３１】したがって、回線の伝送速度が遅い場合には、高速伝送可能な分割データにより伝送がなされるため、再生すべきデータが未伝送となることを防止することができる。

【００３２】また、第１の発明と同様の作用により、サーバを変更することなく、ランダムアクセスが可能である。

【００３３】第４の発明は、第３の発明の動画像データ再生装置において、分割データ受信手段は、時間推定手段に推定された時間とほぼ同じ時間分の分割データが記憶手段に記憶されていない場合に、データ選択手段によって選択されている種別の分割データを、インデックス情報にしたがってサーバに要求し、再生手段は、時間推定手段によって推定された時間とほぼ同じ時間分の分割データが記憶手段に記憶された場合に、この記憶手段に記憶されている分割データの順次再生を開始する動画像データ再生装置である。

【００３４】この第４の発明の動画像データ再生装置においては、第３の発明と同様の作用効果に加えて、分割データをサーバに要求してから受信するまでの間の再生に必要と推測される分の分割データが記憶手段に保持される。

【００３５】したがって、先の第２の発明と同様の効果を得ることができる。

【００３６】第５の発明は、動画像データを複数に分割して複数の分割データを生成し、かつ前記複数の分割データとこの分割データの再生時刻情報とを関連付けたインデックス情報を生成する動画像データ分割装置である。

【００３７】この第５の発明の動画像データ分割装置を用いて複数の分割データ及びインデックス情報を作成してサーバに備え、第１又は第２の発明の動画像データ再生装置を用いることで、サーバの構成、機能に手を加えることなくランダムアクセスが可能となる。

【００３８】また、この動画像データ分割装置を用いることで、複数の分割データ及びインデックス情報を容易に作成できる。

【００３９】第６の発明は、動画像データが分割されて得られる複数の分割データについて伝送速度が異なる複数の種別を生成し、この複数の種別の各分割データをそれぞれの再生時刻情報に関連付けたインデックス情報を生成する動画像データ分割装置である。

【００４０】この第６の発明の動画像データ分割装置を用いて複数種別の各分割データ及びインデックス情報を作成してサーバに備え、第３又は第４の発明の動画像データ再生装置を用いることで、サーバの構成、機能に手を加えることなくランダムアクセスが可能となり、伝送にかかる時間を調節することができる。

【００４１】また、この動画像データ分割装置を用いることで、複数種別の分割データ及びインデックス情報を容易に作成できる。

【００４２】第７乃至第１２の発明は、それぞれ上記第１乃至第６の発明と同様の機能をコンピュータに実現させるためのプログラムを記録したコンピュータ読み取り可能な記録媒体である。

【００４３】この第７乃至第１２の発明のようなプログラムを記録した記録媒体を用いることによって、上述した機能を有していない装置や、計算機においても、簡単に上述した機能を付加することができる。

【００４４】第１３の発明は、サーバに記憶されている動画像データをクライアントに配信する動画像データ配信システムにおいて、サーバは、動画像データが分割されて得られる複数の分割データと、この各分割データとそれぞれの再生時刻情報とを関連付けたインデックス情報とを保持し、クライアントは、インデックス情報をサーバに要求して受信するインデックス情報受信手段と、当該インデックス情報受信手段に受信されたインデックス情報にしたがって、再生開始を指示された再生時刻情報及びその後の再生時刻情報に関連付けられた分割データをサーバに順次要求して受信する分割データ受信手段と、当該分割データ受信手段によって受信された分割データを記憶する記憶手段と、インデックス情報受信手段に受信されたインデックス情報の再生時刻情報にしたがって、記憶手段に記憶されている分割データを順次再生させる再生手段とを備える動画像データ配信システムである。

【００４５】この第１３の発明の動画像データ配信システムにおいては、クライアントが先の第１の発明の動画

像データ再生装置と同様に機能する。

【００４６】これにより、第１の発明と同様の作用効果を得ることができる。

【００４７】なお、第２乃至第１２の発明も、この第１３の発明と同様に、動画像データ配信システムに適用可能である。

【００４８】

【発明の実施の形態】以下、図面を参照しながら本発明の実施の形態について説明する。

【００４９】（第１の実施の形態）本実施の形態においては、動画像データ（動画コンテンツ）を複数に分割した分割データと、この複数の分割データとそれぞれの再生時間情報を関連付けたインデックス情報とを用いることで、サーバを変更することなく動画像データのランダムアクセスを可能とする動画像データ配信システムについて説明する。

【００５０】図１は、本実施の形態に係る動画像データ配信システムの構成を示すブロック図である。

【００５１】この動画像データ配信システム１は、動画像データから分割データとインデックス情報とを生成するエンコーダ２と、内部の記憶装置３ａに各種データをファイル形式で記憶するサーバ３と、動画像データ再生用のビューワ４ａを有するクライアント４とから構成されている。

【００５２】この動画像データ配信システム１では，動画像データがスプライスポイントでファイル分割されており、ビューワ４ａがインデックス情報に基づいて好きな位置の分割データをサーバ３に要求し、受信する。

【００５３】以下、この動画像データ配信システム１の各構成要素について詳しく説明する。

【００５４】図２は、本実施の形態に係る動画像データ配信システム１に備えられるエンコーダ２の機能を示すブロック図である。

【００５５】このエンコーダ２は、動画エンコード時に、ビデオ等の動画素材からなる動画像データ（動画コンテンツ）を所定のスプライスポイント毎に分割して複数の分割データを生成する。また生成した各分割データには異なるファイル名を付けて出力する。この図２では、分割データのファイル名の例として「douga000000」、「douga000003」、「douga000006」等が示されている。

【００５６】なお、この各分割データは、ビューワ４ａ内の受信ページバッファの１ページ内に収まるサイズで生成される。

【００５７】さらに、このエンコーダ２は、各分割データのファイル名と、それぞれの再生時刻の情報とを関係付けてインデックス情報５を生成し、ファイル形式で出力する。

【００５８】図３は、本実施の形態に係る動画像データ配信システム１に備えられるサーバ３の構成を示すブロック図である。

【００５９】サーバ３は、主に記憶装置３ａと、サーバ本体３ｂとから構成される。

【００６０】サーバ本体３ｂは、記憶装置３ａに様々なデータを蓄積する機能と、クライアント４からの要求にしたがって記憶装置３ａの内容を読み出して送信する機能を有する。

【００６１】記憶装置３ａは、各種のデータを記憶し、ここではエンコーダ２によって生成された複数の分割データ及びインデックス情報５も記憶している。

【００６２】クライアント４は、図１で示すように、主にクライアント本体４ｂと、動画像データ再生装置として機能するビューワ４ａとから構成されており、クライアント本体４ｂ上でビューワ４ａが稼働する。

【００６３】図４は、本実施の形態に係る動画像データ配信システム１に備えられるビューワ４ａの機能を示すブロック図である。

【００６４】なお、図４（ａ）は再生開始前のバッファリング中の状態を示し、３ページ分の分割データがページバッファに格納されるまでの状態を示している。

【００６５】一方、図４（ｂ）は再生中の状態を示しており、１ページ分の分割データがデコーダに読み出されると、再びサーバ３から分割データが読み出される状態を示している。

【００６６】このビューワ４ａは、動画像データを再生する装置であり、インデックス情報受信部６、分割データ受信部７、ページバッファ８、デコーダ９から構成される。

【００６７】インデックス情報受信部６は、再生を求める動画像データの指定をユーザから入力する。また、指定された動画像データに関するインデックス情報５をサーバ３に要求し、この要求の応答としてサーバ３からインデックス情報５を受信して保持する。

【００６８】分割データ受信部７は、動画像データの再生開始位置となる再生時刻情報をユーザから入力する。また、インデックス情報受信部６のインデックス情報５を参照し、ユーザから入力された再生時刻情報以降の再生時刻情報に関連付けられているファイルの送信をサーバ３に要求する。この要求は、ページバッファ８に格納された分割データの量が一定量（ここでは３つとする）になるまで順次出力する。

【００６９】そして、この分割データ受信部７は、この要求の応答として分割データを順次受信し、ページバッファ８に出力する。

【００７０】ページバッファ８は、複数のページからなり、各ページ毎に一つの分割データをバッファリングする。また、各ページの内容を再生時刻順に並び替えて記憶する。さらに、デコーダ９によって読み出されたページ内の分割データを破棄し、次に入力される分割データを新規に確保する。

【００７１】デコーダ９は、ページバッファ８内に一定量の分割データが記憶されると、デコード処理を開始し、その後順次デコード処理を続ける。これにより滞りなく動画像データの再生がなされる。

【００７２】次に、上記のような構成を有する動画像データ配信システム１の動作について以下に説明する。

【００７３】まず、エンコーダ２によって動画像データが分割され、分割データ及びインデックス情報が生成され、サーバ３に記憶される。

【００７４】次に、再生を求める動画像データの指定及び再生を開始させる位置が、ユーザによってビューワ４ａに入力される。

【００７５】次に、ユーザに指定された動画像データから生成されたインデックス情報５が、ビューワ４ａのインデックス情報受信部６からサーバ３に要求され、受信される。

【００７６】この受信されたインデックス情報５には、動画像データの再生時刻情報に対して分割データのファイル名が記述されている。

【００７７】分割データ受信部７では、受信されたインデックス情報５に基づいて、再生を開始したい位置以降に対応付けられている分割データがサーバ３に要求され、受信される。

【００７８】受信された分割データは、ページバッファ８の各ページに格納され、再生される時刻順に並び替えられる。

【００７９】また、ページバッファ８において一定量のバッファリングがなされると、デコーダ９によってストリーム再生が開始される。

【００８０】ページバッファ８においては、分割データが１ページ分再生される度にそのデータが破棄されてページが空けられる。

【００８１】そして、分割データ受信部７において、未受信の次の分割データが、インデックス情報５に基づいて検索され、サーバ３に要求、受信され、再生が繰り返される以上説明したように、本実施の形態に係る動画像データ配信システム１においては、動画像データがスプライスポイントにおいてファイル分割されており、その状態がインデックス情報５に記述されている。

【００８２】また、ビューワ４ａが、このインデックス情報５を要求し、続いてこのインデックス情報５に記されている好きな位置から分割データを順次要求・受信し、再生する。

【００８３】したがって、ユーザの望む位置からの分割データが順次再生されるため、動画像データのランダムアクセスが可能となる。

【００８４】また、本実施の形態においては、サーバ３の機能はファイルを保持し、またこの保持されているファイルの中から要求されたファイルを送信するのみである。したがって、この動画像データ配信システム１を導入する場合にはサーバ３に新しい機能や作用を追加する必要がない。

【００８５】ゆえに、クライアント４側のビューワ４ａを備え、サーバ３に分割データとインデックス情報５とを格納するのみで、容易にランダムアクセスを実現することができる。

【００８６】さらに、ランダムアクセスを実現させるための処理が、サーバ３ではなくビューワ４ａによって実行されるため、クライアント４が複数存在してもサーバ３の負荷の増加を防止することができる。

【００８７】なお、本実施の形態に係る動画像データ配信システム１では、エンコーダ２の機能をクライアント４に実装させ、クライアント４がネットワークを介して分割データ及びインデックス情報５をサーバ３に格納するとしてもよい。

【００８８】また、本実施の形態においては、ビューワ４ａの機能をブラウザ等の機能として実現可能であり、サーバ３はWWWサーバとすることができる。

【００８９】これにより、ブラウザで参照する動画コンテンツのランダムアクセスが可能となる。

【００９０】さらに、本実施の形態に係る動画像データ配信システム１では、クライアント４の台数を複数にすることができる。

【００９１】（第２の実施の形態）本実施の形態においては、第１の実施の形態に係る動画像データ配信システム１に備えられるビューワ４ａを変形したビューワについて説明する。

【００９２】先の第１の実施の形態においては、ビューワ４ａがサーバ３に分割データを要求してから受信するまでの時間がネットワークの回線の負荷に依存する。また、この回線の負荷は、様々な要因により時々刻々と変化する。

【００９３】先のビューワ４ａでは、伝送に要する時間や、回線の負荷による伝送速度の変化を吸収するために、動画像データの再生開始時に一定量分（しきい値を設定する）の分割データをバッファリングし、以後順次分割データを読み出している。

【００９４】この手法では、ビューワ４ａにバッファリングされる分割データの量を大きくすればその分だけ回線の負荷による遅延に対応可能となるが、その分再生開始までの待ち時間が長くなる。

【００９５】この再生開始までの待ち時間は、回線の負荷による遅延を吸収できる範囲で最小化されることが望ましい。したがって、バッファリングされる分割データの量も、回線負荷による伝送時間の変化に対応可能な範囲で最小化されることが望ましい。

【００９６】本実施の形態においては、分割データを要求してから受信するまでの時間を推定する。また、この推定された時間に基づいて、ページバッファに記憶しておく分割データの量を変更し、バッファリングされる分

割データの量を動的に最適化するビューワに関して説明する。

【００９７】図５は、本実施の形態に係るビューワの構成を示すブロック図であり、図４と同一の部分には同一の符号を付してその説明を省略し、ここでは異なる部分についてのみ詳しく説明する。

【００９８】このビューワ１０は、第１の実施の形態において説明したビューワ４ａの分割データ受信部７の機能を変更した分割データ受信部１１を備え、新規の構成として時間推定部１２を付加している。また、このビューワ１０は、第１の実施の形態と同様の動画像データ配信システム１に設置可能である。

【００９９】時間推定部１２は、分割データ受信部１１がサーバ３に分割データを要求し、この分割データを受信するまでの時間（以下、「要求受信時間」という）を測定する。

【０１００】なお、この測定された要求受信時間を厳密に用いると、バッファリングの際に余裕がなくなる。したがって、時間推定部１２は多少余分を持たせるために若干多めの値を要求受信時間として推定する。例えば実際の要求受信時間が０．５秒であってとしても、要求受信時間を０．７秒と推定し、設定する。

【０１０１】この時間推定部１２において、要求受信時間は、再生開始以降には分割データが要求されて受信される度に再推定され、更新される。

【０１０２】分割データ受信部１１は、先の第１の実施の形態で説明した分割データ受信部７と同様の処理を行うが、サーバ３から分割データを読み出す際の条件が異なっている。

【０１０３】再生開始前においては、この分割データ受信部１１は、ページバッファ８内の分割データの再生時間の合計が要求受信時間を超えるまで、サーバ３から分割データを受信する。

【０１０４】再生中においては、現在再生中の分割データに関連付けされている再生時刻情報に現状の要求受信時間を加算し、この加算値に該当する再生時刻情報に関連付けされた分割データを要求し、受信する。すなわち、例えば要求受信時間が０．７秒と設定されているとすると、０．７秒後に再生されることになる分割データを要求し、受信する。

【０１０５】上記のような構成を有するビューワ１０の動作について以下に説明する。

【０１０６】図６は、本実施の形態に係るビューワ１０の再生開始前までの動作を示すフローチャートである。

【０１０７】このビューワ１０においては、まずインデックス情報５がサーバ３から受信される（ｓ１）。

【０１０８】次に、このインデックス情報５とユーザの指定内容に基づいて、再生開始位置の分割データが要求される（ｓ２）。

【０１０９】次に、要求された分割データが受信され（ｓ３）、この際に余分を持たせた要求受信時間が推定される（ｓ４）。

【０１１０】次に、インデックス情報５を参照して、次の分割データが要求され（ｓ５）、受信される（ｓ６）。

【０１１１】そして、要求受信時間分の分割データが、ページバッファ８に受信されたか否かが判定され（ｓ７）、受信されている場合には、再生が開始される。

【０１１２】受信されていない場合には、再度次の分割データが要求され、受信される。

【０１１３】図７は、本実施の形態に係るビューワ１０の再生開始後の動作を示すフローチャートである。

【０１１４】再生開始後においては、ページバッファ８に空きがあり（ｔ１）、かつ要求受信時間分の分割データがページバッファ８にない場合に（ｔ２）、次の分割データが要求される（ｔ３）。

【０１１５】なお、この間にもページバッファ８内の分割データはデコードされ再生されている（ｔ４、ｔ９）。

【０１１６】次に、要求された分割データが受信され（ｔ５）、この分割データが最後の分割データか否かが判定される（ｔ６）。

【０１１７】最後の分割データの場合にはデコード処理及び再生が継続され（ｔ７）、処理が終了される。

【０１１８】最後の分割データでない場合には、余分を持たせた要求受信時間が推定され（ｔ８）、デコード処理の実行により動画像データが再生されつつ（ｔ９）、上記の処理が繰り返される。

【０１１９】以上説明したように、本実施の形態に係るビューワ１０においては、サーバに分割データを要求してから受信するまでの時間に多少の余分を加えた時間分の分割データがページバッファ８に蓄積される。

【０１２０】したがって、必要最小限の分割データのみが先行確保されるため、再生開始までの時間を最適化することができる。

【０１２１】（第３の実施の形態）動画像データの再生中に、ネットワークの負荷により回線の伝送速度が分割データの伝送速度未満となると、動画像データがビューワにおいて途切れなく再生されない場合がある。

【０１２２】本実施の形態においては、このような場合に、再生の途中からでも高速伝送可能な分割データに切り換えて伝送を行い、これによりスムーズな再生を維持する動画像データ配信システムについて説明する。

【０１２３】本実施の形態に係る動画像データ配信システムは、図１と同様の構成を有するが、エンコーダ及びビューワに特別の機能が付加されている。

【０１２４】本システムのエンコーダは、元の動画像データが同一であり、スプライスポイントも同一であるが同一の条件における伝送速度が異なる複数種別の分割データ及びインデックス情報を生成する。

【０１２５】また、本システムのビューワは、要求受信時間に基づいて分割データの種別を選択する機能を有する。

【０１２６】以下、図１の場合と異なる機能を有するエンコーダ、及びビューワについて詳しく説明する。

【０１２７】図８は、本実施の形態に係る動画像データ配信システムに備えられるエンコーダの機能を示すブロック図である。

【０１２８】このエンコーダ１３は、エンコード時に、１つの動画像データから伝送速度の異なる複数種別の分割データを生成し、出力する。

【０１２９】なお、以下においては伝送速度の異なる２つの種別の分割データが生成される場合を例として説明するが、３つ以上の種別が生成される場合も同様である。また、この２つの種別のうち、一方を標準分割データとし、この標準分割データよりも伝送速度の速いもう一方を高速分割データとして説明する。

【０１３０】さらに、このエンコーダ１３は、生成された標準分割データのファイル名と高速分割データのファイル名を再生時刻情報に関連付けたインデックス情報１４を生成し、ファイル形式で出力する。

【０１３１】すなわち、このインデックス情報１４には、各再生時刻情報に複数列の分割データのファイル名が関連付けられており、この各列は異なる伝送速度になるようにエンコードされた分割データ列である。

【０１３２】サーバは、標準分割データ、高速分割データ、インデックス情報をファイルとして記憶し、クライアントのビューワから要求されたファイルを送信する。

【０１３３】ビューワは、第２の実施の形態に係る図５のビューワ１０と同様の構成を有するが、分割データ受信部１１に分割データ選択機能が付加されている点が異なる。

【０１３４】この分割データ選択機能は、標準分割データを要求してから受信するまでの時間が、時間推定部１２で推定された要求受信時間よりも大幅に大きい場合に、ネットワークの伝送速度が分割データの伝送速度よりも遅いと判断する。この場合、分割データ選択機能は、インデックス情報１４中の分割データの参照を、標準分割データ列から高速分割データ列に切り換える。

【０１３５】一方、高速分割データを要求してから受信するまでの時間が、時間推定部１２で推定された要求受信時間よりもある程度小さくなった場合には、ネットワークの伝送速度が分割データの伝送速度よりも速くなったと判断し、インデックス情報１４中の分割データの参照を、高速分割データから伝送速度は遅いが画質のよい標準分割データ列に切り換える。

【０１３６】図９は、このビューワにおいて再生開始後に実行される処理を示すフローチャートである。

【０１３７】このビューワでは、動画像データがデコードされ、再生されるが（ｕ１）、それと並列に要求伝送時間分の分割データがページバッファ８に格納されているか否かが判定される（ｕ２）。

【０１３８】ここで、要求伝送時間分の分割データがページバッファ８に格納されていない場合には、この要求伝送時間が大きすぎるか否かが分割データ選択機能によって判定される（ｕ３）。

【０１３９】要求伝送時間が大きすぎる場合には、インデックス情報１４の分割データ列のうち伝送速度の速い方が選択される（ｕ４）。

【０１４０】一方、要求伝送時間が大きすぎない場合には、要求伝送時間が小さすぎないか否かが判定される（ｕ５）。

【０１４１】要求伝送時間が小さすぎる場合には、インデックス情報１４の分割データ列のうち伝送速度の遅い方が選択される（ｕ６）。

【０１４２】そして、この分割データ選択機能によって選択された種別における次の分割データが、サーバに要求される（ｕ７）。

【０１４３】この間にも、動画像データのデコード処理が実行され、再生がなされている（ｕ８）。

【０１４４】要求された分割データが受信されると（ｕ９）、余分を持たせた要求受信時間が推定される（ｕ１０）。

【０１４５】そして、要求する分割データがなくなり（ｕ１１）、かつ再生する分割データもなくなると（ｕ１２）、処理が終了する。

【０１４６】要求する分割データが存在する場合には、上記の処理が繰り返される。

【０１４７】以上説明したように、本実施の形態に係る動画像データ配信システムにおいては、異なる伝送速度の分割データが生成され、分割データを要求してから受信するまでの時間が長い場合には、ビューワは伝送速度の速い分割データを読み出す。

【０１４８】また、分割データの種別の切り替えは各分割データの読み出し毎に行うことができるため、動画像データの再生開始後においても伝送速度を変更することができ、ネットワークの状況変更に柔軟に対応することができる。

【０１４９】さらに、伝送速度に余裕のある場合には画質のよい分割データに変更することができる。

【０１５０】ゆえに、クライアントのビューワで動画像データを途切れなく、良好に再生することができる。

【０１５１】なお、上記の各実施の形態において説明したビューワ、エンコーダの機能は、コンピュータに実行させることのできるプログラムとして、例えば磁気ディスク（フロッピーディスク、ハードディスク等）、光ディスク（ＣＤ－ＲＯＭ、ＤＶＤ等）、半導体メモリなどの記録媒体に書き込んで適用したり、通信媒体により伝送して各種装置に適用することも可能である。このビューワやエンコーダの機能を実現するコンピュータは、記

録媒体に記録されたプログラムを読み込み、プログラムによって動作が制御されることにより、上述した処理を実行する。

【０１５２】

【発明の効果】以上詳記したように本発明では、動画像データを分解した複数の分割データと、この各分割データとそれぞれの再生時刻情報とを関連付けたインデックス情報をサーバに備える。

【０１５３】そして、動画像データ再生装置は、まずインデックス情報を受信し、再生を開始したい位置の時間再生情報以降に関連付けられている分割データを順次読み出す。

【０１５４】これにより、サーバの構成、機能に手を加えることなく、ランダムアクセスが可能になる。

【０１５５】また、このランダムアクセスのための処理がサーバではなく動画像データ再生装置によって実行されるため、ランダムアクセス実行によるサーバの負荷の増加を防止することができる。

【０１５６】さらに、本発明では、動画像データ再生装置に蓄える分割データの量を、分割データを要求してからこの分割データを受信するまでの時間分とすることで、再生開始までの待ち時間を最適化することができる。

【０１５７】さらに、本発明では、伝送速度の異なる複数種別の分割データを用意しておき、ネットワークの状況によって受信する分割データの種別を切り換えることで、ネットワークにおける負荷の影響を受けることなくスムーズな再生が可能となる。

【図面の簡単な説明】

【図１】本発明の第１の実施の形態に係る動画像データ配信システムの構成を示すブロック図。

＊【図２】同実施の形態におけるエンコーダの機能を示すブロック図。

【図３】同実施の形態に係る動画像データ配信システムに備えられるサーバの構成を示すブロック図。

【図４】同実施の形態に係る動画像データ配信システムに備えられるビューワの機能を示ブロック図。

【図５】本発明の第２の実施の形態に係るビューワの構成を示すブロック図。

【図６】同実施の形態に係るビューワの再生開始前までの動作を示すフローチャート。

【図７】同実施の形態に係るビューワの再生開始後の動作を示すフローチャート。

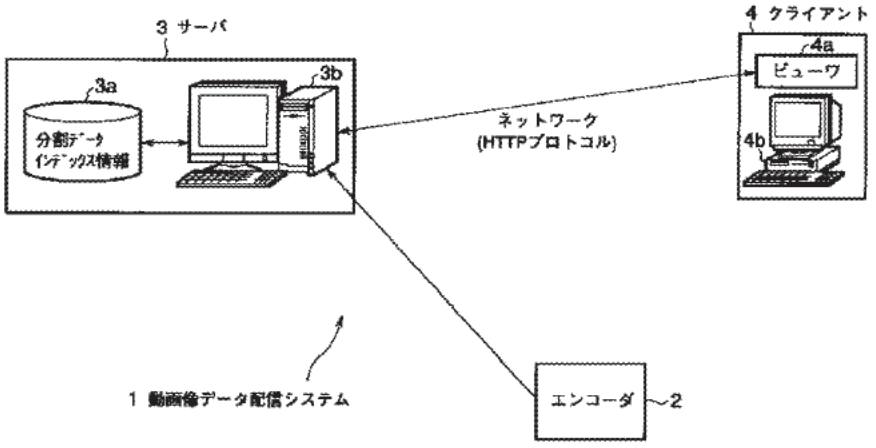
【図８】本発明の第３の実施の形態に係る動画像データ配信システムに備えられるエンコーダの機能を示すブロック図。

【図９】同実施の形態に係る動画像データ配信システムに備えられるビューワの再生開始後の動作を示すフローチャート。

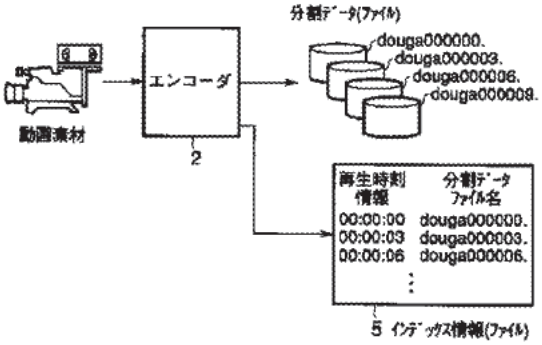
【符号の説明】

１…動画像データ配信システム
２、１３…エンコーダ
３…サーバ
４…クライアント
４ａ、１０…ビューワ
５、１４…インデックス情報
６…インデックス情報受信部
７、１１…分割データ受信部
８…ページバッファ
９…デコーダ
１２…時間推定部

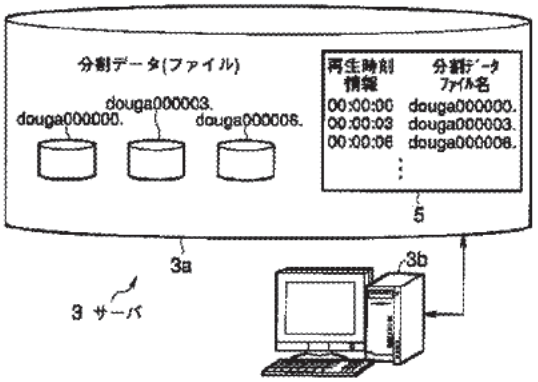
【図１】



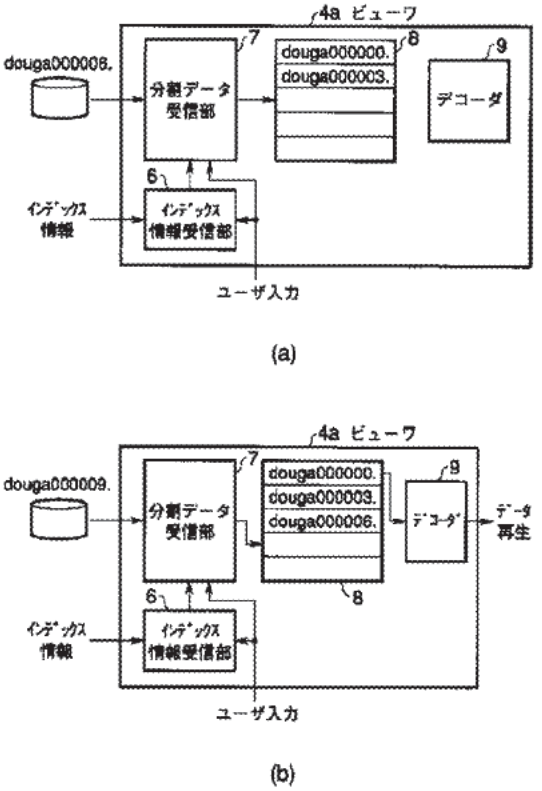
【図２】


分割データ(ファイル)
douga000000.
douga000003.
douga000006.
douga000009.
動画素材
エンコーダ
2
再生時刻情報
分割データファイル名
00:00:00 douga000000.
00:00:03 douga000003.
00:00:06 douga000006.
5 インデックス情報(ファイル)

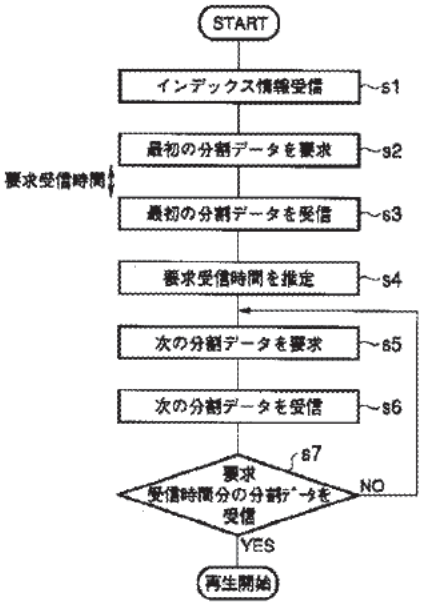
【図３】


分割データ(ファイル)
douga000000.
douga000003.
douga000006.
再生時刻情報
分割データファイル名
00:00:00 douga000000.
00:00:03 douga000003.
00:00:06 douga000006.
5
3a
3b
3 サーバ

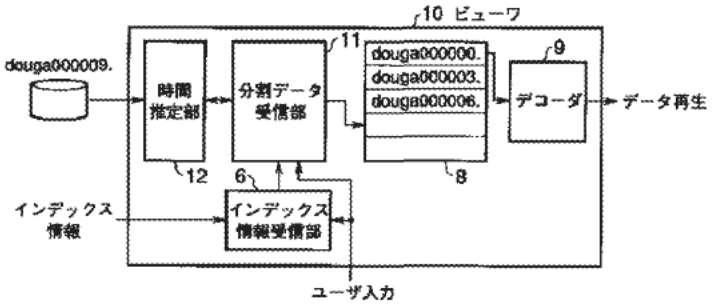
【図４】


4a ビューワ
douga000006.
分割データ受信部
7
douga000000.
douga000003.
8
デコーダ
9
6
インデックス情報
インデックス情報受信部
ユーザ入力
(a)
4a ビューワ
douga000009.
分割データ受信部
7
douga000000.
douga000003.
douga000006.
8
9
デコーダ
データ再生
6
インデックス情報
インデックス情報受信部
ユーザ入力
(b)

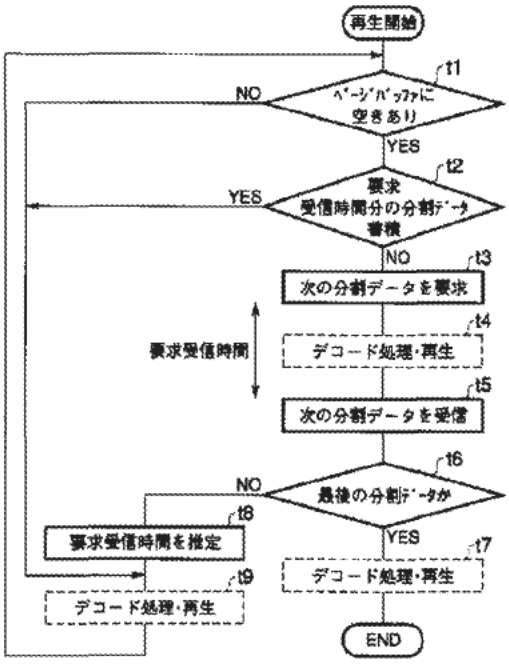
【図６】


START
インデックス情報受信
s1
最初の分割データを要求
s2
要求受信時間
最初の分割データを受信
s3
要求受信時間を推定
s4
次の分割データを要求
s5
次の分割データを受信
s6
s7
要求受信時間分の分割データを受信
NO
YES
再生開始

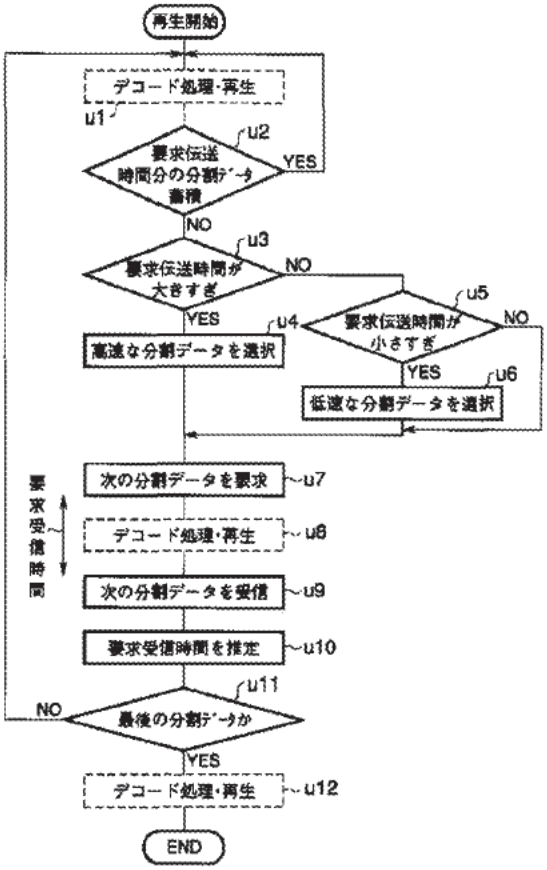
【図５】


10 ビューワ
douga000009.
時間推定部
分割データ受信部
11
douga000000.
douga000003.
douga000006.
9
デコーダ
データ再生
12
6
8
インデックス情報
インデックス情報受信部
ユーザ入力

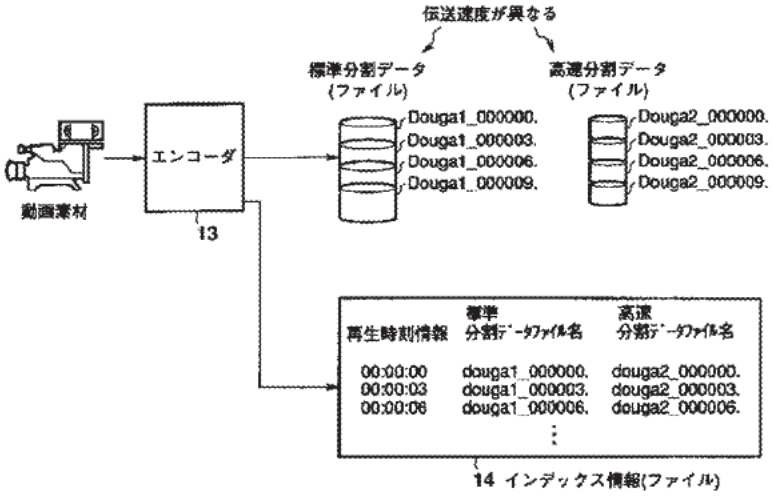
【図７】


再生開始
t1
NO
バージバッファに空きあり
YES
t2
YES
要求受信時間分の分割データ蓄積
NO
t3
次の分割データを要求
要求受信時間
t4
デコード処理・再生
t5
次の分割データを受信
t6
NO
最後の分割データか
t8
要求受信時間を推定
YES
t7
デコード処理・再生
t9
デコード処理・再生
END

【図９】


再生開始
デコード処理・再生
u1
u2
要求伝送時間分の分割データ蓄積
YES
NO
u3
要求伝送時間が大きすぎ
NO
YES
u4
高速な分割データを選択
u5
要求伝送時間が小さすぎ
NO
YES
u6
低速な分割データを選択
要求受信時間
次の分割データを要求
u7
デコード処理・再生
u8
次の分割データを受信
u9
要求受信時間を推定
u10
u11
NO
最後の分割データか
YES
デコード処理・再生
u12
END

【図８】


伝送速度が異なる
標準分割データ
(ファイル)
高速分割データ
(ファイル)
Douga1_000000.
Douga1_000003.
Douga1_000006.
Douga1_000009.
Douga2_000000.
Douga2_000003.
Douga2_000006.
Douga2_000009.
エンコーダ
動画素材
13
再生時刻情報
標準
分割データファイル名
高速
分割データファイル名
00:00:00
00:00:03
00:00:06
douga1_000000.
douga1_000003.
douga1_000006.
douga2_000000.
douga2_000003.
douga2_000006.
14 インデックス情報(ファイル)

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
3 April 2003 (03.04.2003)

PCT

(10) International Publication Number
WO 03/027876 A1

(51) International Patent Classification[7]: G06F 15/16

(21) International Application Number: PCT/US02/30874

(22) International Filing Date:
26 September 2002 (26.09.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/325,483 26 September 2001 (26.09.2001) US

(71) Applicants and
(72) Inventors: **REYNOLDS, Jodie, Lynn** [US/US]; 728 King Way, Folsom, CA 95630 (US). **INGRAHAM, Robert, Walter** [US/US]; 127 Honey Cook Circle, Folsom, CA 95630 (US).

(74) Agents: **GALLENSON, Mavis, S.** et al.; Ladas & Parry, Suite 2100, 5670 Wilshire Boulevard, Los Angeles, CA 90036-5679 (US).

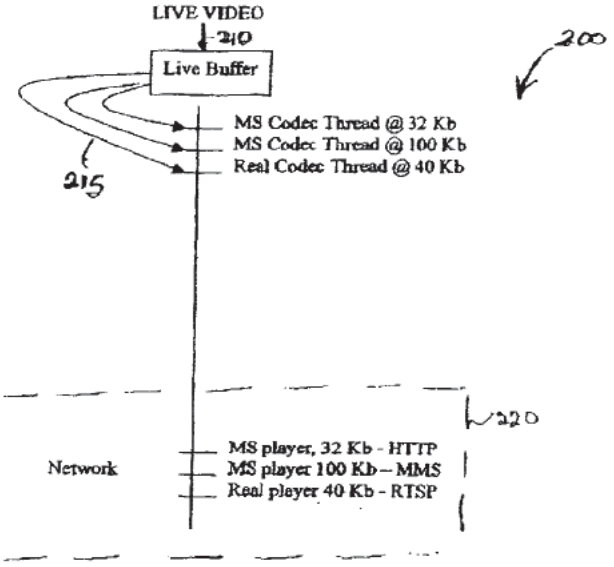
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (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, IE, IT, LU, MC, NL, PT, SE, 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*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: SYSTEM AND METHOD FOR COMMUNICATING MEDIA SIGNALS



(57) Abstract: A media streaming system for streaming media signals is provided. The media streaming system takes a library of separate and distinct CODECs (145) that are provided as a searchable CODEC library and used in determining specific characteristics in the media signal to identify similar sections of the signal. The media streaming system uses a computer implemented intelligence system, such as an intelligence mechanism to learn and capture the unique charateristics of a signal as the signal is being streamed (150). The media streaming system also compresses and decompresses the media signal as the signals are streamed from a source media (100) to a destination device (130, 135, 140).

WO 03/027876 A1

SYSTEM AND METHOD FOR COMMUNICATING MEDIA SIGNALS

RELATED APPLICATION

This application is a non-provisional application of "Reynolds et al. a provisional application entitled "System and Method for Communicating Media Signals" attorney docket number ID-PAT-001PR1, application number 60/325,483, filed on September 26, 2001.

FIELD

This disclosure relates to a system and method for communicating media signals between source and destination devices. More specifically, it relates to a system and method for compressing and decompressing streaming and static media signals for efficiently communicating those signals between source and destination devices using artificial intelligence mechanism.

BACKGROUND

The ability to efficiently communicate streaming and static media between remotely located devices is a significant need that has emerged exponentially with the advent of networked communications such as the Internet. This need has been recently addressed with substantial development resources on a worldwide scale.

The term "media" is herein intended to mean information that may be communicated in the form of a signal from a source device to a destination device for use by the destination device; and, where used herein, media is generally contemplated to comprise either streaming or static media signals. For the purpose of this disclosure, the term "use" as applied to a destination device's operation on media signals is intended to include playing (e.g. sounds, images, video), processing (e.g. telemetry data), or any other use or operation that is the intended purpose of the media signal.

The terms "streaming media" are herein intended to mean media signals that comprise information intended to be communicated to and used by a destination device in a temporal, streaming fashion. The term "streaming" as applied to streaming media signals is herein intended to include signals communicated and processed in a continuous manner over time, or signals that may be

signals communicated and processed in a continuous manner over time, or signals that may be communicated in a series of discrete packets, pieces, or blocks that are interrelated and may be thereafter used by the destination device in a continuous, interrelated fashion. Examples of streaming media signals for the purpose of this disclosure therefore include, without limitation, the following types of media: video, audio, audio combined with video, and data strings such as temporal telemetry. The terms "streaming media" are most typically used by reference to digitized forms of data representing the subject media.

The terms "static media" are herein intended to generally mean media that is not "streaming" as defined above. Static media signals are of the type that generally may be communicated and are intended to be used as a single packet, block, or piece. Static media therefore may include for example, without limitation the following: a discrete image, an individual and relatively temporally short video clip, a sound or sound bite, or a piece or block of information such as telemetry information. It is contemplated, however, that such a "single piece" of static media may be of sufficient magnitude to consist of a plurality of smaller pieces or sub-parts, such as for example regions or pixels of an overall image, individual frames that together form a video clip, digital bits that together comprise a sound, a group of sounds that comprise a sound bite, or bits of information that together comprise a larger block of information.

Streaming media generally includes data files that are significantly larger than static media files, and also often represent many more variables over the temporal communication of such files than experienced with most static media files. Therefore, the ability to efficiently compress streaming media for appropriate communication to destination devices for use is often a much more complex and difficult to achieve goal. Accordingly, much of this disclosure is provided by reference specifically to streaming media communication, and the present invention has been observed to provide significant benefits for such communication. However, where streaming media is specifically referenced herein with respect to this background, and further with respect to the many benefits of the present invention herein disclosed, static media is also further contemplated where appropriate according to one of ordinary skill.

Many different "type-specific" media systems have been in use for quite a long time for transmitting specific types (e.g. video, audio, image, voice, etc.) of streaming and static media signals between sources and remote destinations. Typical examples of such type-specific media systems include television transmission systems, telephone line systems, and radio transmission systems, and every television, telephone, and radio is therefore a receiving device for media. Accordingly, the needs

for efficient communication of streaming and static media touch upon many diverse communications industries, including for example the telephone, television, movie, music, and more recently interactive gaming industries.

Moreover, many medial communications systems, including the various long-standing type-specific systems, are also "format specific", wherein the subject media signals are communicated in a particular format such that the source, transmission channel, and destination device must be specifically compliant to work within that format. Examples of format specific media systems include for example encoded cable television systems that work only for certain types of media and only delivered in particular encoded formats from the cable carrier. Therefore, these systems, in hardware and software, are generally dedicated to only the type and format of media to be provided by the content provider.

Society's needs have outpaced the abilities of these dedicated, content-specific and format-specific systems. In particular, these dedicated systems are not structured to accommodate the ever increasing client demand, real-time, for specified streaming media. Still further, technology developments in the recently interconnected world has tempted the palate of society for the ability to pull, receive, push, and send multiple types of media in multiple formats using one device. Moreover, content providers need to be able to deliver many different media signals to many different types of devices in their clients' offices, living rooms, and hands. Individuals and corporations also desire to communicate with each other using various different formats and using various different respective devices.

Accordingly, a significant industry has emerged for delivering streaming and static media over the centralized network of the Internet. Content delivery companies are currently delivering a wide range of streaming media, from live horse racing and entertainment to medical telemetry and education, over the Internet, and in video and audio formats. According to one published report from DFC Intelligence, video streaming on the Internet grew 215 percent in 2000 to over 900 million total streams accessed. This includes broadband streams, which made up almost 29 percent of total accesses. This same report also estimates that as much as 15 percent of available stream inventory is now being exploited with in-stream advertising. In another report published by Internet researcher Jupiter Media Metrix, business spending alone on streaming video technology will balloon from one-hundred forty million (US$140M) US dollars in 2000 to nearly three billion (US$3B) US dollars by 2005 as companies turn to electronic interaction in communicating with employees, consumers and other businesses.

Still further, the population explosion and increasing number of people transmitting on these systems has severely impacted the available bandwidth for available information. Therefore, the ability to stream media efficiently, using limited bandwidth resources and limited available transmission speeds, is of increased societal importance.

Compression/Decompression Algorithms ("CODECS")

In view of the exponential demand for communicating the different types of media, various compression/decompression systems ("CODEC(s)") have been developed over many years, and have in particular become the recent topic of significant research and development. Specific types of CODECS and systems for managing the operation of CODECS with respect to communicating streaming and static media signals have been developed for specific types of media, including for example still-frame images such as graphics and photographs, and streaming media.

Image CODECS

Various different types of static media CODECS have been developed, and a wide variety of these CODECS are widely known and used. One specific type of static media that has been the topic of particular attention includes images (though a long series of interrelated image frames such as in video context is generally treated as streaming media due to more complex variables, e.g. size and temporal relationship between frames, that significantly impact appropriate compression/decompression needs). Examples of static media CODECing is therefore herein exemplified by reference to certain specific types of conventional image CODEC technologies and methods.

The two most common file formats for graphic images on the world wide web are known as "GIF" and "JPEG" formats, generally considered the respective standards for drawings (e.g. line art) and photographs, and are further described together with other image compression modalities for the purpose of further understanding as follows.

"JPEG" is an acronym for "Joint Photographic Experts Group", and is a graphic image file that complies with ISO standard 10918. Commonly used for photograph compression/decompression, a JPEG file is created by choosing from a range of compression qualities, or, as has also been described, by choosing from one of a suite of compression algorithms. In order to create a JPEG file, or convert

an image from another format to JPEG, the quality of image that is desired must be specified. In general, because the highest quality results in the largest file, a trade-off may then be made, as chosen by the user, between image quality and image size. The JPEG mode of compression generally includes 29 distinct coding processes although a JPEG implementer may not use them all. A JPEG image is typically given a name suffix ".jpg".

"GIF" is an acronym for "Graphics Interchange Format", and is generally considered the de facto standard form of drawing image compression/decompression for Internet communication. GIF formatting uses a compression algorithm known as the LZW algorithm, which was developed by Abraham Lempel, Jacob Ziv, and Terry Welch and made commercially available by Unisys Corporation (though in general such algorithm has been made publicly available without requiring fee-bearing licenses). More specifically, a "LZW" compression algorithm takes each input sequence of bits of a given length (e.g. 12 bits) and creates an entry in a table, sometimes called a "dictionary" or "codebook", for that particular bit pattern. The entry consists of the pattern itself and a shorter code. As input is read, any pattern that has been read before results in the substitution of the shorter code, effectively compressing the total amount of input to something smaller. Earlier approaches, known as LZ77 and LZ78, did not include the look-up table as part of the compressed file. However, the more recent LZW algorithm modality does include the table in the file, and the decoding program that decompresses the file for viewing is able to build the table itself using the algorithm as it processes the encoded input. The GIF format uses the 2D raster data type (associated with display screens using raster lines) and is encoded in binary.

Two versions of GIF formats include GIF 87a, and more recently GIF89a that allows for "animated GIF" file creation, or short sequences of images within a single GIF file that are played in sequence to present movement or change in the image (either in an endless loop or through a progression that reaches an end). GIF89A also allows for, and also for "interlaced GIF", which is a GIF image that arrives and is displayed by the receiver first as a fuzzy outline of an image that is gradually replaced by seven successive waves of bit streams that fill in the missing lines until full resolution is reached. Interlaced GIF allows, for example, a viewer using 14.4 Kbps and 28.8 Kbps modems to observe a briefer wait-time before certain information in a subject image may be processed, such as for example to make decisions (e.g. to click on the image to execute an operation such as a link).

By presenting waves of resolution filling image sequences, interlaced GIF is similar to "Progressive JPEG", which describes an image created using the JPEG suite of compression algorithms

that will "fade in" in successive waves. While the progressive JPEG is often observed to be more appealing way to deliver an image at modem connection speeds, users with faster connections may not likely notice a difference.

"PNG" or "Portable Network Graphics" format has been more recently developed for image compression and that, in time, has been publicized to replace the GIF format for Internet use (though not generally the JPEG format allowing size/quality trade-offs). This format has been developed for public consumption and development. Similar to GIF, PNG is considered a "lossless" compression format, and therefore all image information is restored when a compressed file is decompressed during viewing. However, PNG formatted files are generally intended to be from 10 to 30 percent more compressed than with a GIF format. Further aspects of PNG file formats are provided as follows: (i) color transparency may not be limited to only one color, but the degree of transparency may be controlled ("opacity"); (ii) "interlacing" of images is improved versus standard GIF; (iii) "gamma correction" is enabled, allowing for "tuning" of images in terms of color brightness required by specific display manufacturers; (iv) images can be saved using true color, palette, and gray-scale formats similar to GIF; and (v) "animation" is generally not supported, though PNG is generally considered extensible and therefore software may be layered to provide for such scriptable image animation.

"TIFF" is an acronym for "Tag Image File Format", and is a common format for exchanging raster graphics (or "bitmap") images between application programs, such as for example graphics used for scanner images. A TIFF file is usually given a name suffix of ".tif" or ".tiff", and had generally been developed in the mid-1980's with the support of Adobe Software, Microsoft, and Hewlett-Packard. TIFF files can be in any of several classes, including gray scale, color palette, or RGB full color, the descriptions and differences of which are further developed elsewhere herein this disclosure. TIFF files may also include files with JPEG, LZW, or CCITT Group 4 standard run-length image compression, which are also further described elsewhere herein. As one of the most common graphic image formats, TIFF files are typically used in desktop publishing, faxing, 3-D applications, and medical imaging applications.

Video CODECS

Video compression has been the topic of intense development for various applications, including, for example: pre-recorded video (e.g. "video-on-demand"), teleconferencing, and live video (e.g. broadcasts). "Desk-top" computers, wireless devices, conventional televisions, and high

definition televisions are examples of the different types of receiving devices that an efficient video compression system must serve.

In general, video CODEC algorithms operate on either or both of an individual, frame-by-frame basis, and/or on a "temporal compression" basis wherein each frame is the most common video compression algorithms in conventional use are based on several mathematic principles, including the following: Discrete Cosine Transforms ("DCT"), Wavelet Transforms and Pure Fractals.

"Discrete Cosine Transforms" or "DCT's" are by far the most popular transforms used for image compression applications . In general, DCT is a technique for representing waveform data as a weighted sum of cosines. The DCT is similar to the discrete Fourier transform: it transforms a signal or image from the spatial domain to the frequency domain. The DCT helps separate the image into parts (or spectral sub-bands) of differing importance (with respect to the image's visual quality). Reasons for its popularity include not only good performance in terms of energy compaction for typical images but also the availability of several fast algorithms. DCTs are used in two international image/video compression standards, JPEG and MPEG.

"Wavelet transforms" are generally mathematical algorithms that convert signal data into a set of mathematical expressions that can then be decoded by a destination receiver device, such as for example in a manner similar to Fourier transform. Wavelets have been observed to enhance recovery of weak signals from noise, and therefore images processed in this manner can be enhanced without significant blurring or muddling of details. For this reason, wavelet signal processing has been particularly applied to X-ray and magnetic-resonance images in medical applications. In Internet communications, wavelets have been used to compress images to a greater extent than is generally possible with other conventional methods. In some cases, the wavelet-compressed image can be as small as about 25 percent the size of a similar quality image using the more familiar JPEG format, which is discussed in further detail elsewhere in this disclosure. Thus, for example, a photograph that requires 200Kb and takes a minute to download in JPEG format may require only 50Kb and take only 15 seconds to download in wavelet-compressed format. A wavelet-compressed image file is often given a name suffix ".wif", and either the receiver (e.g. Internet browser on a computer receiver) must support these format specific files, or a plug-in program will be required to read such file.

Fractal image compression is a modern technique of lossy image coding that provides several improvements over existing Fourier series compression schemes. Edge depiction is improved since, when modeled as a step function, edges require a large number of Fourier series terms to properly depict. Other advantages of fractals include fast decoding time and scale independence. Fractal

compression is based on Mandelbrot sets which take advantage of a self similar, scaling dependent, statistical feature of nature (Mandelbrot, 1983). Fractal compression and decompression involves a clustering approach to find regions which show the same characteristics as a sample region independent of rotation and scale. The fractal image compresses images as recursive equations and instructions about how to reproduce them. The equations describe the image in terms of the relationships between its components. The reduction in storage need is due to the fact that fractal compression saves equations and instructions instead of a pixel representation of the image.

"MPEG" is an acronym for Moving Picture Experts Group and has come to be used synonymously with certain evolving video and audio compression standards promulgated therefrom. In general, to use MPEG video files, a personal computer is required with sufficient processor speed, internal memory, and hard disk space to handle and play the typically large MPEG file, usually given the name suffix ".mpg". A specified MPEG viewer or client software that plays MPEG files must be available on the client system, and generally can be downloaded shareware or versions of commercial MPEG players from various sites on the Web. The modes of operation for MPEG formatted media are herein described by reference to these sequentially evolved standards as follows.

More specifically, MPEG-1 standard was designed for coding progressive video generally at a transmission rate of about 1.5Mbps. This was generally designed for the specific application for Video-CD and CD-I media. MPEG-1 audio layer-3 ("MP3") has also evolved from early MPEG work. "MPEG-2" is a standard generally designed for coding interlaced images at transmission rates above 4 Mbps, and was generally intended for use with digital TV broadcast and digital versatile disk. Though it is generally observed that many MPEG-2 players can handle MPEG-1 data as well, the opposite is not generally observed to be true and MPEG-2 encoded video is generally incompatible with MPEG-1 players. Yet another progressive standard, "MPEG-3", has also been proposed for use with high definition television ("HDTV"), though in general MPEG-3 has merged with MPEG-2 which is generally believed to meet the HDTV requirements. Finally, an "MPEG-4" standard has also been most recently developed and is intended to provide a much more ambitious standard to address speech and video synthesis, fractal geometry, and computer visualization, and has further been disclosed to incorporate artificial intelligence in order to reconstruct images.

MPEG-1 and –2 standards define techniques for compressing digital video by factors varying from 25:1 to 50:1. This compression is achieved according to these standards generally using five different compression techniques: (i) discrete cosine transform (DCT), which is a frequency-based transform; (ii) "quantization", which is a technique for losing selective information, e.g. lossy

prediction", wherein some images are predicted from the pictures immediately preceding and following the image.

Further more detailed examples of commercially available video compression technologies include: Microsoft Media Player™ (available from Microsoft Corporation), RealPlayer™ or RealSystem G2™ (commercially available from Real Networks™), Apple's QuickTime™ (commercially available from Sorenson™); and "VDO". The Microsoft Media Player™ is generally believed to apply the MPEG standard of CODEC for compression/decompression, whereas the others have been alleged to use proprietary types of CODECS. Standard compression algorithms, such as MPEG4, have made their way into the hands of developers who are building embedded systems for enterprise streaming, security, and the like.

One example of a more recent effort to provide streaming video solutions over Wireless and IP networks has been publicized by a company named Emblaze Systems (LSE:BLZ). This company has disclosed certain technology that is intended for encoding and playback of live and on-demand video messages and content on any platform: PC's, PDA's, Video cell phones and Interactive TV. Emblaze Systems is believed to be formerly GEO Interactive Media Group. The following Published International Patent Applications disclose certain streaming media compression technologies that is believed to be related to Emblaze Systems to the extent that GEO Interactive Media Group is named as "Assignee": WO9731445 to Carmel et al.; and WO9910836 to Carmel. The disclosures of these references are herein incorporated in their entirety by reference thereto.

Another company that has published CODEC technology that is intended to improve communication of streaming media for wireless applications is Packetvideo™ Corporation, more specifically intending to communicate streaming video to cellular phones. In addition, they are believed to be promoting CODEC technology that is intended to track temporal scalability and signal error resistance in order to protect video and audio streams from the hazards of the wireless environment. U.S. Patent No. 6,167,092 to Lengwehasatit discloses further examples of certain streaming media compression/decompression technology that are believed to be associated with Packetvideo as the named "Assignee" on the face of this Patent reference. The disclosure of this patent reference is herein incorporated in its entirety by reference thereto.

Another prior reference discloses CODEC technology that is intended to provide a cost effective, continuously adaptive digital video system and method for compressing color video data for moving images. The method involves capturing an analog video frame and digitizing the image into a preferred source input format for compression using a combination of unique lossy and lossless digital

Another prior reference discloses CODEC technology that is intended to provide a cost effective, continuously adaptive digital video system and method for compressing color video data for moving images. The method involves capturing an analog video frame and digitizing the image into a preferred source input format for compression using a combination of unique lossy and lossless digital compression techniques including sub-band coding, wavelet transforms, motion detection, run length coding and variable length coding. The system includes encoder and decoder (CODEC) sections, generally disclosed to use a "Huffman" encoder, for compression and decompression of visual images to provide high compression that is intended to provide good to excellent video quality. The compressed video data provides a base video layer and additional layers of video data that are multiplexed with compressed digital audio to provide a data stream that can be packetized for distribution over inter or intranets, including wireless networks over local or wide areas. The CODEC system disclosed is intended to continuously adjust the compression of the digital images frame by frame in response to comparing the available bandwidth on the data channel to the available bandwidth on the channel for the previous frame to provide an output data stream commensurate with the available bandwidth of the network transmission channel and with the receiver resource capabilities of the client users. The compression may be further adjusted by adjustment of the frame rate of the output data stream.

Further more detailed examples of CODEC systems that are intended for use at least in part for streaming video communication are disclosed in the following U.S. Patent Nos.: 6,081,295 to Adolph et al.; 6,091,777 to Guetz et al.; 6,130,911 to Lei; 6,173,069 B1 to Daly et al.; 6,263,020 B1 to Gardos et al.; 6,272,177 to Murakami et al.; and 6,272,180 B1 to Lei. The disclosures of these references are herein incorporated in their entirety by reference thereto.

Most if not all prior streaming video compression methodologies look to the extremely complex mathematical tools within such CODECS, and subtle changes to them, to carry "one size fits all" video over public and private networks of all types, from ultra-low bandwidth networks such as that found in wireless networks, to satellite communications to ultra-high speed fiber optic installations. Among the various conventional methods of compression, there are generally user-definable parameters, including tradeoffs between image size, frame rate, color depth, contrast, brightness, perceived frame quality, buffer length, etc. Further, within the algorithms themselves there are numerous non-user definable qualities and weighted calculations. It is up to the developers to set these one time for one "general" interest, and then package and ship the product.

However, while the video streaming market continues to grow rapidly, the world has not chosen one standard for compression as no one algorithm is ideal for all video sources, destinations, or transmission modalities. While a first CODEC may be best for one type of signal, or for a first portion of a signal (e.g. frame or scene comprising a series of frames), another second CODEC may be best for another type of signal, or even another second portion of the same signal. Still further, one CODEC may be best suited for compression/decompression of a particular streaming signal among send, receive, and transmission devices in a communications network; another second CODEC may be better suited than the first for the same streaming media signal but for another set of communication device parameters. For example, some video streams may deliver color to handheld devices while other video streams can take advantage of the loss of pixels in a black and white transmission to a cellular phone to increase frame rate. Required sound quality, frame rates, clarity, and buffering tolerance all decidedly impact the compression algorithm of choice for optimized video and audio delivery across multiple platforms.

In fact, certain communication device parameters may be sufficiently transient during the streaming media transmission such that an initially appropriate CODEC for an initial set of parameters may be rendered less efficient than another CODEC due to changes in those parameters during the same streamed signal transmission. Examples of such transient parameters include, without limitation: available band width in the data transmission channel, available memory or processing power in either the send or receiving devices, and dedicated display resolution/window in the receiving device (e.g. minimizing windows on a screen). These problems are compounded exponentially by a vast number of iterations of different combinations of such factors that may differentiate one CODEC from another as being most efficient for compression, decompression, and delivery of a specific streaming media signal along a particular communications device system.

As CODEC systems are "format-specific", source and destination devices must be "pre-configured" to communicate media signals between each other according to common, specific compression/decompression modalities, else transcoders must be used. However, even if conventional transcoders are used, constraints in the communication system (e.g. source, transmission channel, destination device) are not generally considered and the communication may be significantly faulty. For the purpose of further illustration, Figures 1A and 1B show two different schematic representations of conventional methods for communicating media between source 110 -120 and destination devices 130 -140. These illustrations specifically exemplify streaming video communication, though other media forms may be represented by similar systems.

Ex. 1018-1158

It has been observed that CODEC algorithms can generally be modified for a specific application and then perform better than a similar unmodified set over that limited example. However, this generally must be done either for a series of frames, or ideally, for each individual frame. Some DCT based algorithms have as many as two billion mathematical operations occurring for each frame at higher resolutions and lower perceived quality. This is entirely too much math for average machines, even commercial servers, to perform thirty to sixty times in a single second. This is the reason for the advent of the dedicated compression board or ASIC.

Audio CODECS

In addition to society's recent interests in improving video compression, audio compression has likewise been the topic of significant efforts also for various live or pre-recorded applications, including audio broadcast, music, transmission synchronized with video, live interactive voice (e.g. telephone). Any and all of these audio compression applications must be compatible with a wide range of client-side receiver/players, such as on a multitude of handheld or desk-top devices having widely varied capabilities and operating parameters.

Conventional audio CODECS generally comprise several different types, a few of which are herein briefly summarized for the purpose of illustration.

"Code Excited Linear Prediction" or "CELP" is a type of speech compression method using waveform CODECS that use "Analysis-by-Synthesis" or "AbS" within the excitation-filter framework for waveform matching of a target signal. CELP-based CODECS have recently evolved as the prevailing technique for high quality speech compression, and has been published to transmit compressed speech of toll-quality at data rates nearing as low as about 6 kbps. However, at least one publication discloses that the quality of CELP coded speech is reduced significantly for bit rates at or below 4 kbps.

"Vocoders" are speech CODECS that are not based on the waveform coding scheme, but rather use a quantized parametric description of the target input speech to synthesize the reconstructed output speech. Vocoders have been disclosed to deliver better speech quality at low bit-rates, such as about 4 kbps, and have been developed for such applications. Low bit rate vocoders use the periodic characteristics of voiced speech and the "noise-like" characteristics of stationary unvoiced speech for speech analysis, coding and synthesis. Some early versions of vocoders (e.g. federal standard 1015 LPC-10, use a time-domain analysis and synthesis method. However, most of the more recent versions,

which at least one publication labels "harmonic coders", utilize a harmonic spectral model for voiced speech segments.

Notwithstanding the previous description of certain specific speech compression techniques, a vast number of speech CODECs and standards have been developed by industry and managed by industry and nonprofit groups. Examples of such groups include without limitation the following, and their standards are often used as reference types of CODECS: the European Telecommunications Standards Institute ("ETSI"); the Institute of Electrical and Electronics Engineers ("IEEE"); and the International Telecommunication Union Telecommunications Standards Sector ("ITU-T"), formerly the "CCITT".

One more recently disclosed method and apparatus for hybrid coding of speech, specified at 4 Kbps, encodes sppech for communication to a decoder for reproduction of the speech where the speech signal is classified into three types: (i) steady state voiced or "harmonic"; (ii) stationary unvoiced; and (iii) "transitory" or "transition" speech. A particular type of coding scheme is used for each class. Harmonic coding is used for steady state voiced speech, "noise-like" coding is used for stationary unvoiced speech, and a special coding mode is used for transition speech, designed to capture the location, the structure, and the strength of the local time events that characterize the transition portions of the speech. The compression schemes are intended to be applied to the speech signal or to the LP residual signal.

Another recently disclosed method and arrangement for adding a new speech encoding method to an existing telecommunication system is also summarized as follows. A CODEC is introduced into a speech transmitting transceiver of a digital telecommunications system in order to use a "new" CODEC and an "old" CODEC in parallel in the system. A CODEC is selected by implementing a handshaking procedure between transceivers where a speech encoding method is implemented in all transceivers and previously used in the telecommunications system concerned. The handshaking is used at the beginning of each connection. At the beginning of a phone call and after handover, the method checks whether both parties can also use the new speech encoding. The handshaking messages have been selected so that their effect on the quality of speech is minimal, and yet so that the probability of identifying the messages is maximal.

Still another relatively recent reference discloses a tunable perceptual weighting filter for tandem CODECS intended for use in speech compression. Specific filter parameters are tuned to provide improved perfromance in tandeming contexts. More specifically, the parameters used are 10^{th}

order LPC predictor coefficients. This system is specified to use "Low-Delay Excited Linear Predictive" CODECS, or "LD-CELP".

Further more detailed examples of streaming audio communications systems using CODECS such as according to the examples just described are provided in the following US Patent references: 6,144,935 to Chen et al.; 6,161,085 to Haavisto et al.; and 6,233,550 to Gersho. The disclosures of these references are herein incorporated in their entirety by reference thereto.

Artificial Intelligence ("AI") and Neural Networks with CODECS

Various systems and methods have been recently disclosed that are intended to integrate artificial intelligence ("AI") or neural networks with the compression and decompression of streaming media signals.

The terms "artificial intelligence" are herein intended to mean the simulation of human intelligence processes by computer systems, including learning (the acquisition of information and rules for using the information), reasoning (using the rules to reach approximate or definite conclusions), and self-correction. Particular applications of AI include "expert systems", which are computer programs that simulate judgment and behavior of a human or organization that has expert knowledge and experience in a particular field. Typically, expert systems contain a knowledge base to each particular situation that is described to the program, and can be enhanced with additions to the knowledge base or to the set of rules.

The terms "neural network" are herein intended to mean a system of programs and data structures that approximates the operation of the human brain, usually involving a large number of processors operating in parallel, each with its own small sphere of knowledge and access to data in its local memory. Typically, a neural network is initially trained or fed large amounts of data and rules about data relationships, after which a program can tell the network how to behave in response to an external stimulus (e.g. input information). In making determinations, neural networks use several principles, including without limitation gradient-based training and fuzzy logic. Neural networks may be further described in terms of knowledge layers, generally with more complex networks having deeper layers. In "feedforward" neural network systems, learned relationships about data can "feed forward" to higher layers of knowledge. Neural networks can also learn temporal concepts and have been widely used in signal processing and time series analysis. Other published applications of neural

networks include oil exploration data analysis, weather prediction, the interpretation of nucleotide sequences in biology labs, and the exploration of models of thinking and consciousness.

The terms "fuzzy logic" are herein intended to mean an approach to computing based upon "degrees of truth" rather than "Boolean logic" which operates within only a true/false (or "binary", as 1 or 0) domain. Fuzzy logic was first advanced by Dr. Lotfi Zadeh of the University of California at Berkeley in the 1960's in relation to work on a problem of computer understanding of natural language, which is not easily translated into absolute Boolean logic terms. Fuzzy logic often does include the cases of 0 and 1 as extreme cases of truth, but also includes the various states of truth in between (e.g. a determination of that the state of being is at some threshold, such as 0.98, may assists in making a decision to assign a 1 with an acceptably low occurrence of error in an operation).

One example of a previously disclosed streaming media compression/decompression system intended to use with artificial intelligence through a neural network uses a Radon transform in order to compress data such as video data. Several previously disclosed AI and/or neural network systems are intended to use AI and/or neural networks for the purpose of error correction during use of certain specified lossless compression CODECS. For example, a learning system is employed to determine a difference between what was received by a receiver after compression and transmission and what is predicted to have been received at the transmission end. That difference is processed as learning to modify the tuning of the CODEC for an additional transmission.

Another example of a disclosed method and device is intended to extrapolate past signal-history data for insertion into missing data segments in order to conceal digital speech frame errors. The extrapolation method uses past-signal history that is stored in a buffer. The method is implemented with a device that is disclosed to utilize a finite-impulse response ("FIR"), multi-layer, feed-forward, artificial neural network that is trained by back-propagation for one-step extrapolation of speech compression algorithm ("SCA") parameters. Once a speech connection has been established, the speech compression algorithm device begins sending encoded speech frames. As the speech frames are received, they are decoded and converted back into speech signal voltages. During the normal decoding process, pre-processing of the required SCA parameters will occur and the results stored in the past-history buffer. If a speech frame is detected to be lost or in error, then extrapolation modules are executed and replacement SCA parameters are generated and sent as the parameters required by the SCA. In this way, the information transfer to the SCA is intended to be transparent, and the SCA processing continues as usual. This disclosure alleges that the listener will not normally notice that a

speech frame has been lost because of the smooth transition between the last-received, lost, and next-received speech frames.

Further more detailed examples of systems that are intended to use artificial intelligence and/or neural networks in systems for media compression and/or decompression, generally relating to media type-specific CODEC methods (e.g. speech, video), are variously disclosed in the following U.S. Patent References: 5,005,206 to Naillon et al.; 5,041,916 to Yoshida et al.; 5,184,218 to Gerdes; 5,369,503 to Burel et al.; 5,598,354 to Fang et al.; 5,692,098 to Kurdziel; 5,812,700 to Fang et al.; 5,872,864 to Imade et al.; 5,907,822 to Prieto, Jr.; and 6,216,267 to Mitchell. Still further examples are provided in the following Published International Patent Applications: WO 01/54285 to Rising; EPO 0372608 A1 to Naillon et al.. The disclosures of all these references cited in this paragraph are herein incorporated in their entirety by reference thereto.

Other disclosures of CODEC systems using feedback or other systems for operating CODECS for use in processing a variety of streaming media signals, but that are not believed to specifically use the labels "AI" or "neural networks", are disclosed in the following U.S. Patent Nos.: 6,072,825 to Betts et al.; 6,182,034 B1 to Malvar; 6,253,165 B1 to Malvar; 6,256,608 B1 to Malvar. The disclosures of these references are herein incorporated in their entirety by reference thereto.

Notwithstanding the significant advancements in CODEC algorithms themselves, and despite prior intended uses of AI and other feedback systems for operating CODECS in order to improve compression efficiencies in communication, there is still a need for significant improvement in the ability to efficiently provide a wide variety of streaming media signals to a wide variety of destination receiver devices over a wide variety of transmission channels with varied bandwidths and communication protocols.

There is still a need to incorporate AI and/or neural networks to apply an appropriate CODEC for communication of a streaming media signal based upon a variety of parameters, including without limitation one or more of the following: (a) the automated choosing of an appropriately optimized CODEC from a library of available CODECS of different types and operation, including in particular based upon an intelligent knowledge of the chosen CODEC's operation compared to the other CODEC's operation and/or against a standard, (b) a pre-trained and/or iteratively learned knowledge of the particular CODEC's operation within a given set of operating parameters representative of the existing situation; and (c) a tuning of the appropriate CODEC based upon an intelligent knowledge of

its operation with respect to either or both of the existing situation or a test situation with reference parameters.

In particular, there is still a need for such an intelligent CODEC system that bases an applied CODEC upon an existing situation that is defined by one or more of the following: parameters of the streaming media signal itself; parameters of the transmission channel capabilities and constraints; and parameters of the receiver device capabilities and constraints.

Still further, there is also still a need for such an intelligent CODEC system that operates based upon an intelligent knowledge with respect to all of these operations and situational parameters in order to optimize the appropriate compression, transmission, decompression, and playing of the subject streaming media signal.

Conventional Transcoders for Streaming Media

Also of recent interest in the field of streaming media communication is providing intercommunication between the wide array of "format-specific" encoding systems in present use. An existing field of various different format-specific systems and pre-encoded content has created a widely fragmented ability to process encoded content, resulting in a significant quagmire of compatibility issues between content providers and client users. If one client desires to see or hear streaming content from a particular source and that content must be put through a CODEC for compression, a compatible CODEC must be used on the client side for decompression to enjoy the signal. Unfortunately, source content is often married to only a few, and often only one, specific CODEC schemes. Therefore, if a client requests such encoded content (or if the source desires to push the encoded content to a particular client), one of two criteria must be met: (1) the client must download or otherwise possess the format-specified CODEC (decoder); or (2) the source media must be put through a "transcoder" in order to decode the source media from the first format into a second format that is compatible with the client's device/system. The term "transcoder" is herein intended to mean a system that converts a media signal from one encoded (i.e. compressed) format to another.

Various techniques for transcoding one media format into another have been previously disclosed. Figure 1C shows one illustrative example of the general process that is characteristic of many known transcoding techniques. More specifically, a request 159 is first received from a particular type of device or player for content that exists in an initial, uncompatible format. According to the specific example shown in Figure 1C, a request 159 from a Microsoft Media™ Player for Real™ Video Content is received. As the content is specifically requested, the content is decoded from the

initial format (e.g. Real-encoded format), and is then "re-encoded" into the appropriate format for the requesting player (e.g. Microsoft Media™ format). This re-encoded media is then served to the requesting client for decoding within the resident system of that player.

This conventional system has significant scalability limitations, in that simultaneous feeds on multiple channels for multiple clients must be supported by an equal number of transcoders. For example, Figure 1D shows a schematic implementation of the conventional transcoding technique just described as it manages four simultaneous stream requests from four Microsoft Media Players, wherein the requested content is initially encoded in Real™ format. The system architecture necessary to support the four encoders 151 - 154 and four decoders 155 - 158 requires significant computing resources. For example, it is believed that each encoder 151 - 154 provided in the example requires a computer having a 600 MHz (e.g. Pentium™ III) having 128 Mbytes of RAM available, or dual 400 MHz processors (e.g. Pentium II) with 256 Mb available RAM. It is further believed that each decoder 155 – 158 needs a 233 MHz machine (e.g. Pentium™ II) having 64 Mb of available RAM. So, four such streams requires the equivalent of a Quad 900 Xeon (available from Compaq, Hewlett Packard, Dell and IBM, estimated to cost at the time of this disclosure about $9K retail). This is for four simultaneous streams – society is presently demanding thousands upon thousands of simultaneous streams.

There is still a need for a transcoder system that efficiently converts multiple format-specifically encoded streaming media signals into multiple other formats using minimal computing resources and in a cost-efficient manner.

Parameters Affecting Media Communication

For the purpose of further illustrating the many variables that may impact the choice of an appropriate CODEC in order to communicate a particular streaming media signal to a desired target, the following is a brief summary of various different types of streaming video formats and processing systems. It is believed that these different systems each generally require different types of compression modalities (e.g. CODECS) in order to optimize communication and playing of streaming media signals in view of available transmission speeds and bandwidth, as well as receiver processing parameters.

Although certain specific types of communications formats and systems are further herein described in detail, the following Table 1 provides a summary of a significant cross-section of the

various different communications systems and transmission carriers currently available or disclosed in view of available speed or bandwidth.

TABLE 1: Data Rates of Various Communications Carrier systems

Technology	Speed	Physical Medium	Application
GSM mobile telephone service	9.6 to 14.4 Kbps	RF in space (wireless)	Mobile telephone for business and personal use
High-Speed Circuit-Switched Data service (HSCSD)	Up to 56 Kbps	RF in space (wireless)	Mobile telephone for business and personal use
Regular telephone service (POTS)	Up to 53 Kbps	Twisted pair	Home and small business access
Dedicated 56Kbps on Frame Relay	56 Kbps	Various	Business e-mail with fairly large file attachments
DS0	64 Kbps	All	The base signal on a channel in the set of Digital Signal levels
General Packet Radio System (GPRS)	56 to 114 Kbps	RF in space (wireless)	Mobile telephone for business and personal use
ISDN	BRI: 64 Kbps to 128 Kbps PRI: 23 (T-1) or 30 (E1) assignable 64-Kbps channels plus control channel; up to 1.544 Mbps (T-1) or 2.048 (E1)	BRI: Twisted-pair PRI: T-1 or E1 line	BRI: Faster home and small business access PRI: Medium and large enterprise access
IDSL	128 Kbps	Twisted-pair	Faster home and small business access
AppleTalk	230.4 Kbps	Twisted pair	Local area network for Apple devices; several networks can be bridged; non-Apple devices can also be connected
Enhanced Data GSM Environment (EDGE)	384 Kbps	RF in space (wireless)	Mobile telephone for business and personal use
Satellite	400 Kbps (DirecPC and others)	RF in space (wireless)	Faster home and small enterprise access
Frame relay	56 Kbps to 1.544 Mbps	Twisted-pair or coaxial cable	Large company backbone for LANs to ISP ISP to Internet infrastructure
DS1/T-1	1.544 Mbps	Twisted-pair, coaxial cable, or optical fiber	Large company to ISP ISP to Internet infrastructure
Universal Mobile Telecommunications Service (UMTS)	Up to 2 Mbps	RF in space (wireless)	Mobile telephone for business and personal use (available in 2002 or later)
E-carrier	2.048 Mbps	Twisted-pair, coaxial cable, or optical fiber	32-channel European equivalent of T-1
T-1C (DS1C)	3.152 Mbps	Twisted-pair, coaxial cable, or optical fiber	Large company to ISP ISP to Internet infrastructure
IBM Token Ring/802.5	4 Mbps (also 16 Mbps)	Twisted-pair, coaxial cable, or optical fiber	Second most commonly-used local area network after Ethernet
DS2/T-2	6.312 Mbps	Twisted-pair, coaxial cable, or optical fiber	Large company to ISP ISP to Internet infrastructure
Digital Subscriber Line (DSL)	512 Kbps to 8 Mbps	Twisted-pair (used as a digital, broadband medium)	Home, small business, and enterprise access using existing copper lines
E-2	8.448 Mbps	Twisted-pair, coaxial cable, or optical fiber	Carries four multiplexed E-1 signals
Cable modem	512 Kbps to 52 Mbps (see "Key and explanation" below)	Coaxial cable (usually uses Ethernet); in some systems, telephone used for upstream requests	Home, business, school access
Ethernet	10 Mbps	10BASE-T (twisted-pair); 10BASE-2 or -5 (coaxial cable); 10BASE-F (optical fiber)	Most popular business local area network (LAN)
IBM Token Ring/802.5	16 Mbps (also 4 Mbps)	Twisted-pair, coaxial cable, or optical fiber	Second most commonly-used local area network after Ethernet
E-3	34.368 Mbps	Twisted-pair or optical fiber	Carries 16 E-1 signals
DS3/T-3	44.736 Mbps	Coaxial cable	ISP to Internet infrastructure Smaller links within Internet infrastructure

OC-1	51.84 Mbps	Optical fiber	ISP to Internet infrastructure Smaller links within Internet infrastructure
High-Speed Serial Interface (HSSI)	Up to 53 Mbps	HSSI cable	Between router hardware and WAN lines Short-range (50 feet) interconnection between slower LAN devices and faster WAN lines
Fast Ethernet	100 Mbps	100BASE-T (twisted pair); 100BASE-T (twisted pair); 100BASE-T(optical fiber)	Workstations with 10 Mbps Ethernet cards can plug into a Fast Ethernet LAN
Fiber Distributed-Data Interface (FDDI)	100 Mbps	Optical fiber	Large, wide-range LAN usually in a large company or a larger ISP
T-3D (DS3D)	135 Mbps	Optical fiber	ISP to Internet infrastructure Smaller links within Internet infrastructure
E-4	139.264 Mbps	Optical fiber	Carries 4 E3 channels Up to 1,920 simultaneous voice conversations
OC-3/SDH	155.52 Mbps	Optical fiber	Large company backbone Internet backbone
E-5	565.148 Mbps	Optical fiber	Carries 4 E4 channels Up to 7,680 simultaneous voice conversations
OC-12/STM-4	622.08 Mbps	Optical fiber	Internet backbone
Gigabit Ethernet	1 Gbps	Optical fiber (and "copper" up to 100 meters)	Workstations/networks with 10/100 Mbps Ethernet plug into Gigabit Ethernet switches
OC-24	1.244 Gbps	Optical fiber	Internet backbone
SciNet	2.325 Gbps (15 OC-3 lines)	Optical fiber	Part of the vBNS backbone
OC-48/STM-16	2.488 Gbps	Optical fiber	Internet backbone
OC-192/STM-64	10 Gbps	Optical fiber	Backbone
OC-256	13.271 Gbps	Optical fiber	Backbone

Comments & Key for Table:

(i)The term "Kbps" as the abbreviation for "thousands of bits per second." In international English outside the U.S., the equivalent usage is "kbits s^{-1}" or "kbits/s".
(ii) Engineers use *data rate* rather than *speed*, but speed (as in "Why isn't my Web page getting here faster?") seems more meaningful for the less technically inclined.
(iii) Relative to data transmission, a related term, bandwidth or "capacity," means how wide the pipe is and how quickly the bits can be sent down the channels in the pipe. These "speeds" are aggregate speeds. That is, the data on the multiple signal channels within the carrier is usually allocated by channel for different uses or among different users.

Key: (i) "T" = T-carrier system in U.S., Canada, and Japan....(ii) "DS"= digital signal (that travels on the T-carrier or E-carrier)...(iii) "E" = Equivalent of "T" that uses all 8 bits per channel; used in countries other than U.S. Canada, and Japan....(iv) "OC" = optical carrier (Synchronous Optical Network) "STM" = Synchronous Transport Modules (see Synchronous Digital Heirarchy). (v) Only the most common technologies are shown. (vi) "Physical medium" is stated generally and doesn't specify the classes or numbers of pairs of twisted pair or whether optical fiber is single-mode or multimode. (vii) The effective distance of a technology is not shown. (viii) There are published standards for many of these technologies.

Cable modem note: The upper limit of 52 Mbps on a cable is to an ISP, not currently to an individual PC. Most of today's PCs are limited to an internal design that can accommodate no more than 10 Mbps (although the PCI bus itself carries data at a faster speed). The 52 Mbps cable channel is subdivided among individual users. Obviously, the faster the channel, the fewer channels an ISP will require and the lower the cost to support an individual user.

Internet Carrier Systems

Communication of streaming video via the Internet may take place over a variety of transmission modalities, including for example digital subscriber lines ("DSL"), "T1" lines, cable modem, plain old telephone service ("POTS") dial-up modem, and wireless carriers. While a description of the many different wireless transmission modalities is treated separately herein, a

summary of various of these other transmission modes is herein provided immediately below for the purpose of further illustration as follows.

The terms "POTS" or "plain old telephone service", or "dial-up", as applied to communications transmission channels, are herein interchangeably used. These terms are intended to mean "narrow-band" communication that generally connects end users in homes or small businesses to a telephone company office over copper wires that are wound around each other, or "twisted pair". Traditional phone service was created to let you exchange voice information with other phone users via an analog signal that represents an acoustic analog signal converted into an electrical equivalent in terms of volume (signal amplitude) and pitch (frequency of wave change). Since the telephone company's signaling is already set up for this analog wave transmission, it's easier for it to use that as the way to get information back and forth between your telephone and the telephone company. Therefore, dial-up modems are used to demodulate the analog signal and turn its values into the string of 0 and 1 values that is called digital information. Because analog transmission only uses a small portion of the available amount of information that could be transmitted over copper wires, the maximum amount of data that you can receive using ordinary modems is about 56 Kbps. The ability of your computer to receive information is constrained by the fact that the telephone company filters information that arrives as digital data, puts it into analog form for your telephone line, and requires your modem to change it back into digital. In other words, the analog transmission between your home or business and the phone company is a bandwidth bottleneck.

With "ISDN", or "Internet subscriber digital network", which some consider to be a limited precursor to DSL, incoming data rates up to about 128 Kbps may be achieved for some end user clients.

A "DSL" or "digital subscriber line" is generally defined as a "broadband" transmission carrier for communicating high-bandwidth communication over ordinary copper telephone lines. Many different types of DSL services have been disclosed, having generally varied data rates and intended applications. Though further discussion is herein provided about certain of these DSL types, the following Table 2 provides a summary of information for certain of these DSL types for the purpose of further developing an overview understanding:

TABLE 2: Types of known DSL services.

DSL Type	Description	Data Rate Downstream; Upstream	Distance Limit	Application
IDSL	ISDN Digital Subscriber Line	128 Kbps	18,000 feet on 24 gauge wire	Similar to the ISDN BRI service but data only (no voice on the same line)
CDSL	Consumer DSL from Rockwell	1 Mbps downstream; less upstream	18,000 feet on 24 gauge wire	Splitterless home and small business service; similar to DSL Lite
DSL Lite (same as G.Lite)	"Splitterless" DSL without the "truck roll"	From 1.544 Mbps to 6 Mbps downstream, depending on the subscribed service	18,000 feet on 24 gauge wire	The standard ADSL; sacrifices speed for not having to install a splitter at the user's home or business
G.Lite (same as DSL Lite)	"Splitterless" DSL without the "truck roll"	From 1.544 Mbps to 6 Mbps , depending on the subscribed service	18,000 feet on 24 gauge wire	The standard ADSL; sacrifices speed for not having to install a splitter at the user's home or business
HDSL	High bit-rate Digital Subscriber Line	1.544 Mbps duplex on two twisted-pair lines; 2.048 Mbps duplex on three twisted-pair lines	12,000 feet on 24 gauge wire	T1/E1 service between server and phone company or within a company; WAN, LAN, server access
SDSL	Symmetric DSL	1.544 Mbps duplex (U.S. and Canada); 2.048 Mbps (Europe) on a single duplex line downstream and upstream	12,000 feet on 24 gauge wire	Same as for HDSL but requiring only one line of twisted-pair
ADSL	Asymmetric Digital Subscriber Line	1.544 to 6.1 Mbps downstream; 16 to 640 Kbps upstream	1.544 Mbps at 18,000 feet; 2.048 Mbps at 16,000 feet; 6.312 Mpbs at 12,000 feet; 8.448 Mbps at 9,000 feet	Used for Internet and Web access, motion video, video on demand, remote LAN access
RADSL	Rate-Adaptive DSL from Westell	Adapted to the line, 640 Kbps to 2.2 Mbps downstream; 272 Kbps to 1.088 Mbps upstream	Not provided	Similar to ADSL
UDSL	Unidirectional DSL proposed by a company in Europe	Not known	Not known	Similar to HDSL
VDSL	Very high Digital Subscriber Line	12.9 to 52.8 Mbps downstream; 1.5 to 2.3 Mbps upstream; 1.6 Mbps to 2.3 Mbps downstream	4,500 feet at 12.96 Mbps; 3,000 feet at 25.82 Mbps; 1,000 feet at 51.84 Mbps	ATM networks; Fiber to the Neighborhood

Typically published data rates for DSL service, which may vary depending upon distance from the central office of the offering service company, includes rates up to 6.1 Mbps (theoretically published at 8.448 Mbps), which is believed to enable continuous transmission of motion video, audio, and 3-D effects. More typical individual connections provide from 512 Kbps to 1.544 Mbps downstream and about 128 Kbps upstream. A DSL line can carry both data and voice signals and the data part of the line is continuously connected. DSL has been anticipated in some publications to replace ISDN in many areas and to compete with cable modem for multimedia communication to homes and businesses. DSL operates purely within the digital domain and does not require change into analog form and back. Digital data is transmitted to destination computers directly as digital data and this allows the phone company to use a much wider bandwidth for forward transmission. Meanwhile, if

a client user chooses, the signal can be separated so that some of the bandwidth is used to transmit an analog signal so that a telephone and computer may be used on the same line and at the same time.

Most DSL technologies require that a signal splitter be installed at a home or business, requiring the expense of a phone company visit and installation. However, it is possible to manage the splitting remotely from the central office. This is known as splitterless DSL, "DSL Lite," G.Lite, or Universal ADSL (further defined below) and has recently been made a standard. Several modulation technologies are used by various kinds of DSL, although these are being standardized by the International Telecommunication Union (ITU). Different DSL modem makers are using either Discrete Multitone Technology (DMT) or Carrierless Amplitude Modulation (CAP). A third technology, known as Multiple Virtual Line (MVL), is another possibility.

A variety of parameters of DSL operation are variable and affect the effective data rates that can be achieved. DSL modems generally follow the data rate multiples established by North American and European standards. In general, the maximum range for DSL without a repeater is 5.5 km (18,000 feet). As distance decreases toward the telephone company office, the data rate increases. Another factor is the gauge of the copper wire. The heavier 24 gauge wire carries the same data rate farther than 26 gauge wire. For destination devices beyond the 5.5 kilometer range, DSL may still be provided, though only generally if the respective phone company provider has extended the local loop with optical fiber cable.

To interconnect multiple DSL users to a high-speed network as a "backbone", the telephone company uses a Digital Subscriber Line Access Multiplexer ("DSLAM"). Typically, the DSLAM connects to an asynchronous transfer mode ("ATM") network that can aggregate data transmission at gigabit data rates. At the other end of each transmission, a DSLAM demultiplexes the signals and forwards them to appropriate individual DSL connections.

"ADSL" or "Asymmetric Digital Subscriber Line" is the form of DSL that will become most familiar to home and small business users. ADSL is called "asymmetric" because most of its two-way or "duplex" bandwidth is devoted to the downstream direction, sending data to the user. Only a small portion of bandwidth is available for upstream or user-interaction messages. However, most Internet and especially graphics- or multi-media intensive Web data need lots of downstream bandwidth, but user requests and responses are small and require little upstream bandwidth. Using ADSL, up to 6.1 megabits per second of data can be sent downstream and up to 640 Kbps upstream. The high downstream bandwidth means that a telephone line may carry motion video, audio, and 3-D images to destination computers or television displays. In addition, a small portion of the downstream bandwidth

can be devoted to voice rather data, and phone conversations may be carried without requiring a separate line. Unlike a similar service over "cable" television lines, ADSL does not compete for bandwidth with neighbors in a given area. In many cases, your existing telephone lines will work with ADSL. In some areas, they may need upgrading.

"CDSL" or "Consumer DSL" is a trademarked version of DSL, to be made available by Rockwell Corporation, that is somewhat slower than ADSL (1 Mbps downstream, generally predicted to be lower upstream) but has the advantage that a "splitter" does not need to be installed at the user's end. Hardware may be required to carry CDSL by local phone companies to homes or businesses. CDSL uses its own carrier technology rather than DMT or CAP ADSL technology.

Various companies have worked with telephone companies in developing a standard and easier installation version of ADSL, called "G.Lite", that is believed to be under deployment at the time of this disclosure. "G.Lite" or "DSL Lite" (also known as "splitterless ADSL", and "Universal ADSL") is believed to be essentially a slower ADSL that doesn't require splitting of the line at the user end but manages to split it for the user remotely at the telephone company, which is believed to lower costs. G.Lite, officially ITU-T standard G-992.2, is published to provide a data rate from 1.544 Mbps to 6 Mpbs downstream and from about 128 Kbps to about 384 Kbps upstream. At least one publication has predicted G.Lite to become the most widely installed form of DSL.

"HDSL" or "High bit-rate DSL" is believed to be the earliest variation of DSL to be widely used for wideband digital transmission within a corporate site and between the telephone company and a customer. The main characteristic of HDSL is that it is symmetrical: an equal amount of bandwidth is available in both directions. For this reason, the maximum data rate is generally lower than for ADSL. HDSL can carry as much on a single wire of twisted-pair as can be carried on a T1 line in North America or an E1 line in Europe (up to about 2.32 Mbps).

"IDSL" or "ISDN DSL" is somewhat of a misnomer since it's really closer to ISDN data rates and service at about 128 Kbps than compared with the much higher rates generally associated with ADSL.

"RADSL" or "Rate-Adaptive DSL" is an ADSL technology to be made available from Westell company in which software is able to determine the rate at which signals can be transmitted on a given customer phone line and adjust the delivery rate accordingly. Westell's "FlexCap2™" version system uses RADSL to deliver from about 640 Kbps to about 2.2 Mbps downstream and from about 272 Kbps to about 1.088 Mbps upstream over an existing line.

"SDSL" or "Symmetric DSL" is similar to HDSL with a single twisted-pair line, carrying about 1.544 Mbps (U.S. and Canada) or about 2.048 Mbps (Europe) each direction on a duplex line. It's symmetric because the data rate is the same in both directions.

"UDSL" or "Unidirectional DSL" is a proposal from a European company, and is generally believed to provide a unidirectional version of HDSL.

"VDSL" or "Very high data rate DSL" is believed to be a technology under development that promises much higher data rates over relatively short distances, for example between about 51 and about 55 Mbps over lines up to about 1,000 feet or about 300 meters in length. At least one publication has predicted that VDSL may emerge somewhat after ADSL is widely deployed and co-exist with it. The transmission technology (CAP, DMT, or other) and its effectiveness in some environments is not yet determined. A number of standards organizations are working on it.

"x2/DSL" is modem from 3Com that supports 56 Kbps modem communication but is upgradeable through new software installation to ADSL when it becomes available in the user's area. At least one publication cites 3Com as describing this technology to be "the last modem you will ever need."

A "T1" transmission line is generally considered a "broadband" carrier and is defined as a type of "T-carrier" system, which is believed to have been first introduced by the Bell System in the U.S. in the 1960's as the first successful system that supported digitized voice transmission. The T-carrier system is entirely digital, using pulse code modulation and time-division multiplexing. Voice signals are typically sampled at about 8,000 times a second and each sample is digitized into an 8-bit word. With 24 channels digitized at the same time, a 192-bit frame, representing 8-bit words on each of 24 channels, is thus transmitted about 8,000 times a second. Each frame is separated from the next by a single bit, resulting in a 193-bit block. The T-1's published data rate of 1.544 Mbps generally represents the 192 bit frame, and the 1-bit signaling bit, multiplied by 8,000.

A T-1 system typically uses 4 wires and provides duplex capability, with two wires dedicated for receiving and two for sending at the same time. The T-1 digital stream includes 24, 64 Kbps channels that are multiplexed, wherein the standard 64 Kbps channel is based on the bandwidth required for a voice conversation. The four wires were originally a pair of twisted pair copper wires, but more recent systems provide coaxial cable, optical fiber, digital microwave, and other carrier technologies. The number and use of the channels may be varied from the standard guidelines.

The original transmission rate (1.544 Mbps) for T-1 lines is in common use today in Internet service provider ("ISP") connections to the Internet. Another level, the T-3 line, is published to

provide 44.736 Mbps, and is also commonly used by Internet service providers. Another commonly used service is "fractional T-1", which is the rental of some portion of the 24 channels in a T-1 line, with the other channels unused.

Display Capabilities/Constraints & Related Standards

Various different types of receiver display capabilities may also significantly impact the appropriate CODEC modality for efficiently communicating particular streaming media signals for display by the receiver. A brief summary of certain examples to illustrate such varied display parameters (e.g. resolution, clarity, color, depth, size, type/format-specific) is provided for a better understanding as follows.

One parameter that is highly variable between different types and makes of streaming media receiver devices, and therefore that may have significant impact on the appropriate CODEC to be used, is the range of colors that may be expressed by a display device, or "palette". A standard "browser-safe" palette, which may be accommodated by most software for Internet-based streaming media display, may include for example about 216 colors, though for web-based streaming media the computer display capability as well as the browser software capability must be understood.

With respect to computer display technology, a color is set for each individual pixel or addressable illumination element on the screen. Each pixel has a red, green, and blue (RGB) component. By specifying the amount of intensity for each of these components, a distinct color is given to that pixel. A "true color" display generally defines the color of a pixel on a display screen using a 24-bit value, allowing the possibility of up to 16,777,216 possible colors. The number of bits used to define a pixel's color shade is called the "bit-depth". True color is sometimes referred to as "24-bit color", though many modern color display systems offer a 32-bit color mode. An extra byte, called the "alpha channel", is typically used for control and special effects information. A "gray scale" (composed of discrete shades of gray) display setting is generally defined as having N bits of depth where N represents the saturation of black within the pixel. If N=1, the image is not called gray scale but instead monochrome, or black and white, as the bit can only be on or off and can contain no shading information.

Common computer resolutions include for example and without limitation the following:

(i) VGA or Video Graphics Array capable of displaying 640x480 pixels in 16 colors or 320x240 pixels in 256 colors in a 4:3 aspect ratio;

(ii) SVGA or Super Video Graphics Array capable of 800x600x6 bits/pixel (16 colors) or 650x480x8 bits/pixel (256 colors). SVGA was created by the Video Electronics Association (VESA); and

(iii) XGA (v1-4) or eXtended Graphics Array capable of 1024x768 pixels at 32,768 colors.

Additional standards have been added such as SXGA, defining pixel sizes above 1960x1440 and color depths of 32 bits/pixel and higher.

In the event that a larger range of colors (or palette) is used by a media signal than a particular display or browser can handle, most browsers are typically adapted to "dither" the colors, which is herein intended to mean that the browser will find colors within its palette that it can substitute for any color that is outside of its palette. To further illustrate the wide range of different system display capabilities, systems using WindowsTM (commercially available from Microsoft Corporation) and MacintoshTM (commercially available from Apple Corporation) based operating systems do not have identical palettes; within the usual 256 color palette, 216 are common to both types of browsers, whereas 40 are different and therefore require dithering by a browser operating within one of the systems if an image signal is communicated to that type of system in a format specified by the other.

Many different technologies also exist with respect to how a visual display is enabled from electronic information. The terms "VDT" or "Video Display Terminals" are generally used within the computer industry and are herein intended to be used interchangeably with simple references to "display". With respect to computer terminal use, VDT's comprise a computer output surface and projecting mechanism that shows text and graphic images to the computer user. VDT's may use a variety of specific display technologies, including for example cathode ray tubes ("CRTs"), liquid crystal displays ("LCDs"), light-emitting diodes ("LEDs"), gas plasma, or other image projection technology. The display is usually considered to include the screen or projection surface and the device that produces the information on the screen. In some computers, the display is packaged in a separate unit or "monitor", or the display may be fully integrated in a single unit with the computer processor.

With respect to LCD's in particular, this technology generally requires minimal volume and physical depth compared to other VDT's, and therefore is typically used in laptop computers and cellphone/PDA's. LCD's consume much less power than LED and gas–display VDT's because they work on the principle generally of blocking light rather than emitting it. An LCD may be either

"passive matrix" or "active matrix", which is also known as "thin film transistor" or "TFT" display. The passive matrix LCD has a grid of conductors with pixels located at each intersection in the grid. A current is sent across two conductors on the grid to control the light for any pixel. An active matrix has a transistor located at each pixel intersection, requiring less current to control the luminance of a pixel. For this reason, the current in an active matrix display can be switched on and off more frequently, improving the screen refresh time and therefore efficacy for higher speeds of streaming media (e.g. action video). Some passive matrix LCD's have dual scanning, in that they scan the grid twice with current in the same time as one scan in earlier versions; however, the active matrix is still generally considered to be the superior technology of the two. Reflective color display technology—the integration of color filters into passive-matrix display construction—is a low-power, low-cost alternative to active-matrix technology. Because they reflect ambient light, reflective LCDs deliver particularly high performance during use outside in daylight. Various different display technologies, and therefore transmission formats, have also been specifically developed for television viewing. Thus several different standards have evolved for television transmission, and their differences may significantly impact the nature and extent of compression desired (and therefore the choice of a particular CODEC) for communicating streaming media signals in television environs. These standards include in particular and without limitation: standard definition television ("SDTV"); and high definition television ("HDTV").

"SDTV" or "standard definition television" and "HDTV" or "high definition television" are the two categories of display formats for digital television ("DTV") transmissions, which are becoming the standard. These formats provide a picture quality similar to digital versatile disk ("DVD"), and are summarized relative to their similarities and differences as follows.

HDTV provides a higher quality display, with a vertical resolution display from about 720p to at least about 1080i and an aspect ratio (the width to height ratio of the screen) of generally 16:9, for a viewing experience similar to watching a movie. In comparison, SDTV has a range of lower resolutions and no defined aspect ratio. New television sets will be either HDTV-capable or SDTV-capable, with receivers that can convert the signal to their native display format. SDTV, in common with HDTV, using the MPEG-2 file compression method in a manner that generally reduces a digital signal from about 166 Mbps to about 3 Mbps. This allows broadcasters to transmit digital signals using existing cable, satellite, and terrestrial systems. MPEG-2 uses the lossy compression method, which means that the digital signal sent to the television is compressed and some data is lost, but this lost data may or may not affect how the human eye views the picture. Both the ATSC and DVB standards selected

MPEG-2 for video compression and transport. The MPEG-2 compression standard is elsewhere herein described in further detail.

Because a compressed SDTV digital signal is smaller than a compressed HDTV signal, broadcasters can transmit up to five SDTV programs simultaneously instead of just one HDTV program, otherwise known as "multicasting". Multicasting is an attractive feature because television stations can receive additional revenue from the additional advertising these extra programs provide. With today's analog television system, only one program at a time can be transmitted. Note that this use of the term "multicasting" is distinct from its use in streaming video where it involves using special addressing techniques.

When the United States decided to make the transition from analog television to DTV, the Federal Communications Commission decided to let broadcasters decide whether to broadcast SDTV or HDTV programs. Most have decided to broadcast SDTV programs in the daytime and to broadcast HDTV programs during prime time broadcasting. Both SDTV and HDTV are supported by the Digital Video Broadcasting (DTV) and Advanced Television Systems Committee (ATSC) set of standards.

HDTV as a television display technology provides picture quality similar to 35 mm. movies with sound quality similar to that of today's compact disc (further with respect to audio quality, HDTV receives, reproduces, and outputs Dolby Digital 5.1). Some television stations have begun transmitting HDTV broadcasts to users on a limited number of channels. HDTV generally uses digital rather than analog signal transmission. However, in Japan, the first analog HDTV program was broadcast on June 3, 1989. The first image to appear was the Statue of Liberty and the New York Harbor. It required a 20 Mhz channel, which is why analog HDTV broadcasting is not feasible in most countries.

HDTV provides a higher quality display than SDTV, with a vertical resolution display from 720p to 1080i. The *p* stands for *progressive scanning*, which means that each scan includes every line for a complete picture, and the *i* stands for *interlaced scanning* which means that each scan includes alternate lines for half a picture. These rates translate into a frame rate of up to 60 frames per second, twice that of conventional television. One of HDTV's most prominent features is its wider aspect ratio (the width to height ratio of the screen) of 16:9, a development based on a research-based belief that the viewer's experience is enhanced by screens that are wider. HDTV pixel numbers range from one to two million, compared to SDTV's range of 300,000 to one million. New television sets will be either HDTV-capable or SDTV-capable, with receivers that can convert the signal to their native display format.

In the United States, the FCC has assigned broadcast channels for DTV transmissions. In SDTV formats, DTV makes it possible to use the designated channels for multiple signals at current quality levels instead of single signals at HDTV levels, which would allow more programming with the same bandwidth usage. Commercial and public broadcast stations are currently deciding exactly how they will implement their use of HDTV.

Simulcast is the simultaneous transmission of the same television program in both an analog and a digital version using two different channels or frequencies. At the end of the DTV transition period, it is believed by that analog transmission will be substantially replaced such that current analog channels will be used solely for DTV. The extra channels that were used for digital broadcasting may for example then be auctioned and used for more television channels or other services such as datacasting. Simulcast is also used for the transmission of simultaneous television and Internet services, the transmission of analog and digital radio broadcasts, and the transmission of television programs in different screen formats such as the traditional format and the wide screen format. Simulcast broadcasting is used worldwide.

The transition to DTV is not an easy or inexpensive transition. For a television station to transmit DTV programming, it must build its DTV facilities, but a station must have revenue to build these facilities. Simulcast allows stations to continue receiving revenues from traditional analog programming and also gain extra revenues from the extra digital programming. Another obstacle in the transition to DTV is lack of interest among consumers. The need for special equipment is prohibiting viewers from seeing the difference between digital and analog programs, which is also slowing down public enthusiasm for DTV.

The equipment needed for operating DTV depends on whether terrestrial, cable, or satellite services are used as the transmission channel/carrier. In any event, and according to known or anticipated systems, it is generally believed that consumers will, at a minimum, have to purchase a converter to view DTV transmissions on their old television sets. In addition, consumers that use terrestrial services or antennas to receive television signals need an antenna equipped for digital signals. A consumer located in mountainous terrain in an ATSC-compliant country may not be able to receive terrestrial-based digital signals because of multipath effects. This is common even with today's analog television system. In DVB compliant countries, terrain does not affect the reception of digital signals. Satellite users are already enjoying DTV broadcasting, but a larger satellite dish might be needed to view HDTV programming.

A "set-top" box is herein defined as a device that enables a television set to become a user interface to the Internet and also enables an analog television set to receive and decode DTV broadcasts. DTV set-top boxes are sometimes called receivers. It is estimated that 35 million homes will use digital set-top boxes by the end of 2006, the estimated year ending the transition to DTV.

A typical digital set-top box contains one or more microprocessors for running the operating system, usually Linux or Windows CE, and for parsing the MPEG transport stream. A set-top box also includes RAM, an MPEG decoder chip, and more chips for audio decoding and processing. The contents of a set-top box depend on the DTV standard used. DVB-compliant set-top boxes contain parts to decode COFDM transmissions while ATSC-compliant set-top boxes contain parts to decode VSB transmissions. More sophisticated set-top boxes contain a hard drive for storing recorded television broadcasts, for storing downloaded software, and for other applications provided by the DTV service provider. Digital set-top boxes can be used for satellite and terrestrial DTV but are used mostly for cable television. A set-top box price ranges from $100 for basic features to over $1,000 for a more sophisticated box.

In the Internet realm, a set-top box often really functions as a specialized computer that can "talk to" the Internet - that is, it contains a Web browser (which is really a Hypertext Transfer Protocol client) and the Internet's main program, TCP/IP. The service to which the set-top box is attached may be through a telephone line as, for example, with Web TV or through a cable TV company like TCI.

To take advantage of Dolby Digital 5.1 channel for satellite broadcasts, a satellite receiver that provides a Dolby Digital output is necessary. For cable users, all digital set-top boxes are equipped with a Dolby Digital two-channel decoder. To use 5.1 channel sound, a 5.1 channel-compliant set-top box is needed or an external 5.1 channel decoder unit.

The most dramatic demonstration of digital television's benefits is through a high-end HDTV, because of the larger screen, wider aspect ratio and better resolution. Like most new technologies, however, HDTV is expensive. Nevertheless, less expensive digital TVs provide a markedly improved viewing experience over regular TV, and for those who choose to retain their old sets, even the addition of a set-top converter will deliver a discernibly improved picture and sound.

The FCC's schedule for transition to DTV proposes that everyone in the U.S. should have access to DTV by 2002 and that the switch to digital transmission must be completed either by 2006 or when 85% of the households in a specific area have purchased digital television sets or set-top converters.

In the early 1990s, European broadcasters, consumer equipment manufacturers, and regulatory bodies formed the European Launching Group (ELG) which launched a "DVB" or "Digital Video Broadcasting" project in order to introduce DTV throughout Europe. DVB is intended to provide an open system as opposed to a closed system. Closed systems are content-provider specific, not expandable, and optimized only for the system they were developed for. An open system, such as DVB, allows the subscriber to choose different content providers and allows integration of PCs and televisions. DVB systems are intended to be optimized for television, but as well as supporting home shopping and banking, private network broadcasting, and interactive viewing. DVB is intended to open the possibilities of providing crystal-clear television programming to television sets in buses, cars, trains, and even hand-held televisions. DVB is also promoted as being beneficial to content providers because they can offer their services anywhere DVB is supported regardless of geographic location. They can also expand their services easily and inexpensively and ensure restricted access to subscribers reducing lost revenues due to unauthorized viewing. Today, the DVB Project consists of over 220 organizations in more than 29 countries worldwide and DVB broadcast services are available in Europe, Africa, Asia, Australia, and parts of North and South America

Format-Specific Media

Various different formats for the streaming media signals themselves are also herein summarized by way of non-limiting example to also provide a further understanding of how CODECS may vary for a particular case.

"DVD" is an acronym for "digital versatile disc" is generally defined as a relatively recent optical disc technology that holds up to about 4.7 Gigabytes of information on one of its two sides, or enough for a movie about 133 minutes long on average. With two layers on each of its two sides, it may hold up to 17 Gigabytes of video, audio, or other information, compared to current CD-ROM discs of approximately the same physical size that hold about 600Mbytes (DVD holds more than about 28 times the information). DVD players are required to play DVD's, though they will also play regular CD-ROM discs. DVDs can be recorded in any of three general formats variously optimized for: (i) video (e.g. continuous movies); (ii) audio (e.g. long playing music); and (iii) or a mixture (e.g. interactive multimedia presentations). The DVD drive has a transfer rate somewhat faster than an 8-speed CD-ROM player. DVD format typically uses the MPEG-2 file and compression standard, which has about 4-times the resolution of MPEG-1 images and can be delivered at about 60 interlaced fields

per second where two fields constitute one image (MPEG-1 delvers about 30 non-interlaced frames per second. MPEG-2 and -1 standards are elsewhere herein defined in more detail. Audio quality on DVD is comparable to that of current audio compact discs.

"DVD-Video" is the name typically given for the DVD format designed for full-length movies and is a box that will work with a television set. "DVD-ROM" is a name given to the player that is believed by some to be the future replacement to CD-ROM drives in computers, as these newer drives are intended to play both regular CD-ROM discs as well as DVD-ROM discs. "DVD-RAM" is the name given to writeable versions of DVDs. "DVD-Audio" is the name typically given to players designed to replace the compact disc player.

"VHS" is an acronym for "Video Home System" and is generally defined as a magnetic videotape cartridge format, typically a half-inch wide, developed for home use with the ability to record and playback analog video and audio signals. VHS has become a popular format and the de facto standard for home movie distribution and reproduction mainly due to its pervasive presence and recordability. VHS stores signals as an analog format on a magnetic tape using technology similar to that of audiocassettes. The tapes are played back and recorded on using VHS video cassette recorders (VHS VCRs). VHS tapes store up to around two hours of video typically, although some VCRs are able to record to them at a slower speed allowing up to six or even eight hours of recording per tape.

The VHS format outputs a little over 200 lines of horizontal resolution. This compares to DVDs that output over 500 lines of horizontal resolution. Technically and perceptually, VHS is a format that has been surpassed by other formats, including for example DVD, S-VHS, Hi-8 and others. However, VHS remains a pervasive means for viewing video, and VHS tapes are still easily found across the country and around the world everywhere from movie rental stores to grocery stores making then easily accessible.

"CD" is an acronym for "compact disc" and is generally defined as small, portable, round medium for electronically recording, storing, and/or playing audio, video, text, and other information in digital form. Initially, CD's were only read-only; however, newer versions allow for recording as well (e.g. "CD-RW").

"Super audio compact disc" or "SACD" is a high resolution audio CD format that, together with DVD-Audio ("DVD-A"), are the two formats competing to replace the standard audio CD (though most of the industry generally is backing DVD-A, with a general exception believed to be Philips and Sony). SACD, like DVD-A, offers 5.1 channel surround sound in addition to 2-channel stereo. Both formats improve the complexity of sound by increasing the bit rate and the sample rate, and can be

played on existing CD players, although generally only at quality levels similar to those of traditional CDs. SACD uses Direct Stream Digital ("DSD") recording, which is published as being proprietary to Sony, that converts an analog waveform to a 1-bit signal for direct recording, instead of the pulse code modulation ("PCM") and filtering used by standard CDs. DSD uses lossless compression and a sampling rate of about 2.8 MHz to improve the complexity and realism of sound. SACD may also include additional information, such as text, graphics, and video clips.

Also for the purpose of further understanding, Internet-based communications also have particular protocols for communication that must be accommodated by a streaming media communications system using the Internet "superhighway". These protocols, in particular with respect to streaming media communication, are briefly summarized immediately below for the purpose of providing a more detailed understanding.

With respect to Internet communication, streaming media signals are generally communicated in digital format via data packets. The terms "packets" are herein intended to mean units of data that are routed between an origin and a destination via the Internet or any other packet-switched network. More specifically, when a file is sent, the protocol layer of the communications system (e.g. TCP layer of TCP/IP based system) divides the file into chunks of efficient size for routing. Each of these packets is separately numbered and includes the Internet address of the destination. The individual packets for a given file may travel different routs through the Internet. When they have all arrived, they are reassembled into the original file, for example by the TCP layer at the receiving end. A packet-switching scheme is an efficient way to handle transmissions on a connectionless network such as the Internet. An alternative scheme, circuit-switched, is used for networks allocated generally for voice connections. In circuit-switching, lines in the network are shared among many users as with packet-switching, but each connection generally requires the dedication of a particular path for the duration of the connection.

Wireless Communications & the WAP Gateway

Of equal importance to the contemporary age of the Internet, the age of wireless communications has significantly extended society's ability to interact outside of the fixed confines of the home and office, allowing our remote communications to break free from the umbilical cords of wires and cables. For example, in 2000, the number of mobile subscribers grew by close to 50%.

However, wireless communications systems, protocols, and enabling technologies have developed in a significantly fragmented, "format-specific" market on a world-wide scale. This is particularly true in comparing systems in wide use in the United States as compared to the rest of the world. Therefore, much effort has been expended in overcoming compatibility issues between format-specific systems and between the related wireless devices operating on different platforms. For the purpose of further understanding wireless communication as it is later related to the present invention, the following is a brief overview of significant technologies, systems, and protocols used in the wireless communication industry.

In general, the progression of wireless communications systems for cellular telephones is colloquially given the terms "1G", "2G", "2.5G", and "3G", representing respectively first generation, second generation, and so-on. Initial systems were purely analog, known as the 1G phones and systems. However, with rapid growth, available bandwidth for cellular phone use quickly eroded, giving way to digital signal processing in the 2G, which significantly widened the available bandwidth and ability for complex signal processing for advanced telecommunications. However, as demand progressed for wireless Internet access, so went the technology development from 2G phones (generally not Internet enabled), to 2.5G and 3G (progressively more enabled). As will be further developed immediately below, the systems, protocols, and enabling technologies thus have developed toward a concentrated focus in bringing the 2.5G and 3G modes to industry and consumers.

In general, there are four major digital wireless networks based upon 2G technology: time division multiple access ("TDMA"), code division multiple access ("CDMA"), Global System for Mobile communication ("GSM") and cellular digital packet data ("CDPD"). These are briefly herein described as follows.

Time division multiple access ("TDMA") is a technology used in digital cellular telephone communication that divides each cellular channel into three time slots in order to increase the amount of data that can be carried. TDMA is used by Digital-American Mobile Phone Service (D-AMPS), Global System for Mobile communication ("GSM"), and Personal Digital Cellular ("PDC). However, each of these systems implements TDMA in a somewhat different and incompatible way. An alternative multiplexing scheme to TDMA and FDMA (frequency division multiple access) is code division multiple access ("CDMA").

Code division multiple access ("CDMA") refers to any of several protocols used in 2G and 3G wireless communications. As the term implies, CDMA is a form of multiplexing that allows numerous signals to occupy a single transmission channel, optimizing the use of available bandwidth. the

technology is used in ultra-high-frequency (UHF) cellular telephone systems in the 800 MHz to 1.9 GHz bands. CDMA uses analog-to-digital conversion (ADC) in combination with spread spectrum technology. Audio input is first digitized into binary elements. The frequency of the transmitted signal is then made to vary according to a defined pattern (code), so it can be intercepted only by a receiver whose frequency response is programmed with the same code, so it follows exactly along with the transmitter frequency. There are trillions of possible frequency-sequencing codes, thus enhancing privacy and making cloning difficult. The CDMA channel is nominally 1.23 MHz wide. CDMA networks use a scheme called "soft handoff", which minimizes signal breakup as a handset passes from one cell to another. The combination of digital and spread spectrum modes supports several times as many signals per unit bandwidth as analog modes. CDMA is compatible with other cellular technologies; this allows for nationwide roaming.

The original CDMA, also known as CDMA One, was standardized in 1993 and is considered a 2G technology that is still common in cellular telephones in the U.S. One version of cdmaOne, IS-95A, is a protocol that employs a 1.25 MHz carrier and operates in RF bands at either 800 MHz or 1.9 GHz; this supports data rates of up to 14.4 Kbps. Another version, IS-95B, is capable of supporting speeds of up to 115 Kbps by bundling up to eight channels.

More recent CDMA varieties, CDMA2000 and wideband CDMA offer data speeds many times faster. CDMA2000, also known as IMT-CDMA Multi-Carrier or IS-136, is a CDMA version of the IMT-2000 standard developed by the International Telecommunications Union (ITU). The CDMA2000 standard is a 3G technology that is intended to support data communications at speeds ranging from 144Kbps to 2 Mbps. Companies that have developed versions of this standard include Ericsson and Qualcomm corporations. Wideband CDMA, or "WCDMA", is an ITU standard derived from CDMA that is also known as IMT-2000 direct spread. WCDMA is a 3G technology intended to support data rates of up to 2Mbps for local area access, or 384 Kbps for wide area access, and supports mobile/portable voice, images, data, and video communications at these speeds. WCDMA digitizes input signals and transmits the digitized output in coded, spread-spectrum mode over a 5 MHz wide carrier – a much broader range than the 200 KHz wide narrowband CDMA.

The Global System for Mobile communication ("GSM") is a digital mobile telephone system that is widely used in Europe and other parts of the world; this system uses a variation of "TDMA" (introduced immediately below) and is the most widely used of the three digital wireless telephone technologies (TDMA, GSM, and CDMA). GSM digitizes, compresses, and then sends data down a channel with two other streams of user data, each in its own time slot. It operates at either the 900 Mhz

or 1800 MHz frequency band. At the time of this disclosure, GSM is generally considered the wireless telephone standard in Europe, and has been published to have over 120 million users worldwide and is available in 120 countries. At least one company in the United States, American Personal Communications (SprintTM subsidiary), is using GSM as the technology for a broadband personal communications services ("PCS"). PCS are telecommunications services that bundle voice communications, numeric and text messaging, voice-mail and various other features into one device, service contract and bill. PCS are most often carried over digital cellular links. This service is planned to have more than 400 base stations for various compact mobile handsets that are being made by manufacturers such as Ericsson, Motorola, and Nokia corporations; these devices generally include a phone, text pager, and answering machine. GSM is part of an evolution of wireless mobile telecommunications that includes High-Speed Circuit-Switched Data (HCSD), General Packet Radio System (GPRS), Enhanced Data GSM Environment (EDGE), and Universal Mobile Telecommunications Service (UMTS).

Cellular Digital Packet Data ("CDPD") is a wireless standard providing two-way, 19.2 kbps packet data transmission over existing cellular telephone channels.

Several different protocols have also been put into use for communicating over the various wireless networks. Various specific such protocols are briefly introduced as follows.

"X.25" is a packet-based protocol, principally used at the time of this disclosure in Europe and adapted as a standard by the Consultative Committee for International telegraph and Telephone (CCITT). X.25 is a commonly used network protocol that allows computers on different public networks (e.g. CompuServe, Tymnet, or TCP/IP network) to communicate through an intermediary computer at the network layer level. X.25's protocols correspond closely to the data-link and physical-layer protocols defined in the Open Systems Interconnection ("OSI").

"OSI" is a model of network architecture and a suite of protocols (a protocol stack) to implement it, developed by ISO in 1978 as a framework for international standards in heterogeneous computer network architecture. The OSI architecture is split between seven layers, from lowest to highest: (1) physical layer; (2) data link layer; (3) network layer; (4) transport layer; (5) session layer; (6) presentation layer; and (7) application layer. Each layer uses the layer immediately below it and provides a service to the layer above. In some implementations, a layer may itself be composed of sub-layers.

General Packet Radio Services ("GPRS") is a packet-based wireless communication service that promises data rates from 56 to 114 Kbps and continuous connection to the Internet for mobile phone

and computer users. The higher data rates will allow users to take part in video conferences and interact with multimedia Web sites and similar applications using mobile handheld devices as well as notebook computers. GPRS is based on Global System for Mobile ("GSM") communication and will complement existing services such as circuit-switched cellular phone connections and the Short Message Service ("SMS"). SMS is a message service offered by the GSM digital cellular telephone system. Using SMS, a short alphanumeric message (160 alphanumeric characters) can be sent to a mobile phone to be displayed there, much like in an alphanumeric pager system. The message is buffered by the GSM network until the phone becomes active.

The packet-based service of GPRS is publicized to cost users less than circuit-switched services since communication channels are being used on a shared-use, as-packets-are-needed basis rather than dedicated only to one user at a time. It is also intended to make applications available to mobile users because the faster data rate means that middleware currently needed to adapt applications to the slower speed of wireless systems will no longer be needed. As GPRS becomes widely available, mobile users of a virtual private network ("VPN") will be able to access the private network continuously rather than through a dial-up connection. GPRS is also intended to complement "Bluetooth", a standard for replacing wired connections between devices with wireless radio connections. In addition to the Internet Protocol ("IP"), GPRS supports X.25 protocol. GPRS is also believed to be an evolutionary step toward Enhanced Data GSM Environment ("EDGE") and Universal Mobile Telephone Service ("UMTS").

Universal Mobile Telecommunications Service ("UMTS") is intended to be a 3G, broadband, packet-based transmission of text, digitized voice, video, and multimedia at data rates up to 2 Mbps. UMTS is also intended to offer a consistent set of services to mobile computer and phone users no matter where they are located in the world. This service is based upon the GSM communication standard, and is endorsed by major standards bodies and manufacturers, and is the planned standard for mobile users around the world by 2002. Once UMTS is fully implemented, computer and phone users can be constantly attached to the Internet as they travel.

Enhanced digital GSM enterprise ("EDGE")service is a faster version of the Global System for Mobile (GSM) wireless service, designed to deliver data at rates up to 384 Kbps and enable the delivery of multimedia and other broadband applications to mobile phone and computer users. The EDGE standard is built on the existing GSM standard, using the same time-division multiple access (TDMA) frame structure and existing cell arrangements. EDGE is expected to be commercially

available in 2001. It is regarded as an evolutionary standard on the way to Universal Mobile Telecommunications Service (UMTS).

Wireless Application Protocol ("WAP") is a specification for a set of communication protocols to standardize the way that wireless devices such as cellular telephones and radio transceivers, can be used for Internet access, including e-mail, the World Wide Web, newsgroups, and Internet Relay Chat ("IRC"). While Internet access has been possible prior to WAP, different manufactures have used "format-specific" technologies. WAP enables devices and service systems to intercooperate.

In most recent times, much effort has been expended to merge the fields of wireless communications and the Internet in order to bridge the gap of cords, wires, and cables that had before separated the "information superhighway" from reaching people on wireless devices. Such technology merger has developed for example within the home and office network setting itself, where wireless infrared and radio frequency communications systems have been developed for interfacing equipment within a "wireless" office or home. Another substantial effort has also been underway to communicate and share information with more remote wireless devices, such as cell phones and personal digital assistants ("PDA's").

PDA's are typically small, mobile devices that may be "hand-held" and usually contain limited processors and display screens for managing, storing, and displaying telephone books, calendars, calculator, and the like. Recently available PDA's have been made "wireless enabled", either by having wireless modems embedded within the PDA itself, or by coupling to wireless modem "plug-ins" such as a cell phone. Wireless enabled PDA's are also generally "Internet enabled" with limited "browser" capability allowing the PDA to communicate with server devices over the Internet. Examples of commercially available wireless "enabled" PDA's include the Palm VII (from Palm, Inc.), and the iPAQ™ (from Compaq, Inc.). These PDA's include a Windows CE™ operating system that provides the limited browser capability and screen display for content. These phones have processing capabilities from about 33 MHz to about 220 MHz and varied screen display capabilities, such as for example 320 x 240 pixel screen displays.

Similarly, cellular phones themselves have also been recently rendered "Internet enabled", also with limited browser capability and screens to display content. Examples of "Internet enabled" cellular phones include, for example: Sanyo SCP-4000™, Motorola *i1000plus*™, among a wide range of others; this wide field represents hundreds of different processing and display capabilities.

In either the case of the PDA or the cellular phone that is "Internet-enabled", compatibility with the Internet protocols of communication must be achieved. In general, wireless communications take

place over a wireless applications protocol ("WAP"), whereas communications over the Internet proceed according to one of several different protocols, the most common being Transmission Control Protocol/Internet Protocol ("TCP/IP"). Therefore, a WAP Gateway, as shown in Figure 1E, forms a bridge between the world of the Internet (or any other IP packet network) and the wireless phone/data network, which are fundamentally different in their underlying technologies. The gateway, in essence, does the interpretation between these two distinct entities, allowing the consumer to use their cell phone or hand held computing device (e.g. PDA) to access the Internet wirelessly.

However, streaming media that is formatted for transmission to higher power computing devices such as desk-top computers having significant display capabilities is not generally compatible for receipt and viewing on these devices that have severely limited processing and display functionality. Particular "format-specific" compression schemes have been developed for use specifically with only these devices, and only specific media content may be transmitted to these devices in those formats.

There is still a need for a streaming media communications system that is adapted to transmit a wide variety of streaming media signals in appropriate formats to be played by wireless devices such as cellular phones and PDA's having unique constraints, such as, for example, limited and variable processing, memory, and display capabilities.

SUMMARY OF THE INVENTION

The present invention addresses and overcomes the various limitations, inefficiencies, resource limitations, and incompatibilities of prior known methods for streaming media communication, and is provided in various beneficial modes, aspects, embodiments, and variations as follows.

The present invention according to one embodiment is a streaming media communications system that uses a computer implemented intelligence system, such as artificial intelligence, in a network system, such as a neural network, to communicate a streaming media signal between a transmission device and at least one destination device.

The present invention according to another embodiment is a system for communicating a streaming media signal between a transmission device and a plurality of destination devices each having different media signal processing capabilities.

The present invention according to another embodiment is a streaming media communications system that is adapted to communicate a streaming media signal from a single transmission device and

at least one destination device over a plurality of different transmission channels, each having a different transmission capability or constraints with respect to communicating the streaming media signal.

The invention according to another embodiment is a neural network incorporating an artificial intelligence implementation that is adapted to be trained in an adaptive learning process with respect to the ability of a streaming media compression system's ability to compress a streaming media signal at a source into a compressed representation of the streaming media signal, transmit the compressed representation over a transmission channel to a destination device, and decompress the compressed representation into a decompressed representation of the streaming media signal which is adapted to be played by the destination device.

The invention according to another embodiment is a system for compressing streaming media signals according to a CODEC that is used at least in part based upon at least one parameter affecting communication of the streaming media signals. According to one aspect of this mode, the CODEC is used according to at least one of the following parameters: a previously learned behavior of the CODEC with respect to another reference signal, a previously learned behavior of the CODEC with respect to a prior attempt at compressing or decompressing the same streaming media signal, a comparison of the CODEC's operation with respect to the streaming media signal against a reference algorithm compression of the streaming media signal, a learned constraint of the transmission channel, and a learned constraint of the destination device. In one beneficial embodiment, the CODEC is used based upon more than one of these parameters, and in still a further beneficial variation is used based upon all of the parameters.

The invention according to another embodiment is a system for compressing streaming media signals using a CODEC library that is adapted to store multiple CODECS of different types and operations, and that is adapted to be searched and accessed by a network system, such as a neural network, in order to provide an appropriate CODEC from the CODEC library for use in compressing the input streaming media signal into a compressed representation for transmission to a destination device.

The invention according to another embodiment is a CODEC operating system that is adapted to interface with a CODEC Library and also with a neural network in order to use the neural network in a process, such as an artificial intelligence process, to choose an appropriate CODEC from the CODEC library and use the chosen CODEC for compressing the streaming media signal into a compressed representation of the streaming media signal for transmission to a destination device.

According to one aspect, the CODEC library is adapted to receive and store a new CODEC such that the new CODEC may be interfaced with the neural network in order to be chosen and applied to compress the streaming media signal as provided.

The invention according to another embodiment is a destination agent that is adapted to be stored by a destination device for use in decompressing a compressed representation of a streaming media signal. The destination agent is adapted to communicate with a remotely located, compressed streaming media transmission system in order receive and play streaming media signals therefrom. In a particularly beneficial aspect, the software agent is adapted to deliver information about the destination device to the compressed streaming media transmission system, and is also adapted to receive and decode certain encoded streaming media signals from the compressed streaming media transmission system.

The invention according to another embodiment is a system for communicating a streaming media signal having a destination agent that is adapted to be stored within a destination device for decompressing a compressed representation of a streaming media signal into a decompressed representation that may be played by the destination device.

According to one aspect of this embodiment, the destination agent has a diagnostic agent and also a decompression agent. The diagnostic agent is adapted to determine a value for at least one parameter of the destination device related to the capability for processing, storage, or display. The decompression agent is adapted to apply a CODEC decompressor to decompress the compressed representation of the streaming media signal into the decompressed representation using a CODEC based at least in part upon the value of the at least one parameter.

According to another aspect, the destination agent comprises a software agent. In one variation, the software agent is embedded within the destination device. In another variation, the software agent is adapted to be loaded onto the destination device at least in part by a remotely located source that is adapted to deliver the compressed representation of the streaming media signal to the destination device.

The invention according to another embodiment is a transcoder for trancoding streaming media signals between at least one initial format and at least one transcoded format. The transcoder includes a single thread for each of several streaming media signals.

The invention according to another embodiment is a video-on-demand streaming media system incorporating the embodiments shown in the Figures and otherwise described herein.

Ex. 1018-1189

The invention according to another embodiment is a mobile telephone communications system incorporating the embodiments shown in the Figures and otherwise described herein.

The invention according to another embodiment is an interactive gaming system incorporating the embodiments shown in the Figures and otherwise described herein.

The invention according to another embodiment incorporates the various modes, embodiments, aspects, features, and variations herein disclosed above and elsewhere to static media, as well as to media that is stored locally after processing (e.g. compressing) and not transmitted.

BRIEF DESCRIPTION OF THE FIGURES

Figures 1A-B show schematic block diagrams representing two respective variations of prior media communications systems using conventional CODEC systems.

Figures 1C-D show schematic block diagrams representing two respective variations of prior media transcoder systems.

Figure 1E shows a schematic block flow diagram of the various interrelated components in a prior WAP gateway communications system.

Figures 2-3 show schematic block diagrams of the transcoder system of one embodiment of the present invention during two respective modes of use.

Figures 4A-5 show block flow diagrams in various detail, respectively, of a media communications system according one embodiment of the invention.

Figure 6 shows a schematic block flow diagram of various interrelated components of a "video-on-demand" streaming video communications system according one embodiment of the invention.

Figure 7 shows a schematic block flow diagram of various interrelated components of a wireless streaming video communications system according to one embodiment of the present invention.

Figure 8 shows a schematic block flow diagram of various interrelated components of a WAP gateway media communications system according one embodiment of the present invention.

Figure 9 shows a schematic block flow diagram of various interrelated components of a wireless communications system during backhauling according to one particular mode of use of the media communications system of an embodiment the present invention.

Figure 10 shows a schematic block flow diagram of various interrelated components of an interactive gaming communications system and set-top TV browsing of one embodiment of the present invention.

Ex. 1018-1190

DETAILED DESCRIPTION OF THE EMBODIMENTS

The present invention as illustrated variously through the embodiments below (and by reference to the Figures) provides a media communications system that includes a compression system, a delivery system, and a decompression system, and in another aspect includes a transcoder system. In general, the combination of these individual sub-systems provides a capability to efficiently transcode media between multiple encoding formats, in addition to customize the compression, delivery, and decompression of randomly selected streaming media signals based upon a large array of system parameters as variables. These variables include for example, without limitation, parameters related to the following: the source video signal, the source transmitting device, the transmission modality, and the destination device. The compression, delivery, and decompression of a media signal is thus customized to be optimally efficient for a given, and changing, environment of use. As a result, a wide range of complex streaming media signals may be communicated with a level of efficiency and range of device compatibility that is significantly improved over other known systems.

Notwithstanding the benefits of the overall streaming media communication system herein described, each sub-system described also independently provides beneficially useful results for streaming media communication. The various subsystems themselves, and the various iterations of combinations of these sub-systems apparent to one of ordinary skill based at least in part upon this disclosure, are also contemplated within the scope of the invention. In addition, various aspects of the overall communication system, as well as of each sub-system described, are also contemplated as useful for other applications other than specifically for streaming media communication in particular. Therefore, where apparent to one of ordinary skill, such additional applications are further contemplated within the scope of the invention, despite the particularly useful modes applied to improved streaming media communication.

Transcoder

A video/audio transcoder 200 is provided according to the invention that enables one incoming video source 210 to be streamed across multiple formats 215 (for example MPEG4, Real Video™, and QuickTime™) from one device without human intervention. The transcoder 200 according to the present embodiment provides substantially greater functionality at a fraction of the price of other

commercially available transcoder systems. Moreover, because the system works "on-the-fly," pre-compressing of the video source 210 is significantly diminished.

More specifically, the transcoder 200 system and method according to the invention is adapted to transcode digitized media originating from any compressed or uncompressed format to be reproduced into any other compressed format -- on demand, real-time. The system 200 and method also enables efficient, simultaneous processing of multiple streams 215 of differing data from a multiplicity of different compressed or uncompressed formats into a multiplicity of different compressed formats.

The transcoder 200 of the present embodiment is herein described in an overall system by way of illustration by reference to Figure 3. As shown, a first player initially makes a connection to a server 300 that houses the transcoder 200. The player format (e.g., Microsoft Media), connection speed (e.g., 32 Kbps) and protocol (HTTP) are identified. The server 300 pulls the live or pre-encoded video into a "live buffer" or "cache" 310 and encodes it as digitized but nearly uncompressed data (e.g., AVI or MPEG2). The server 300 then loads an appropriate CODEC thread (e.g. Microsoft Media™) at the connection speed (e.g. 32 Kbps). Next, the server 300 loads a HTTP/MS player thread that serves the first client Then, a second stream is requested by a client using M/S Player at 100Kbps with MMS. The server loads the appropriate MS CODEC thread at the appropriate 100 Kbps rate. Then, the server 300 loads an MMS/MS player thread to serve the second client. Then, a third stream is requested by a client using Real Player at 40 Kbps with RTSP. The server 300 loads the appropriate Real CODEC thread at the appropriate 40 Kbps rate. Then, the server 300 loads an RTSP/Real player thread to serve the third client. Again, this illustration is exemplary, and other specific CODECS may be suitable substitutes, as well as other bit-rates, etc.

In order to provide still a further understanding of the present transcoder embodiment, Figure 3 shows the transcoder 200 by way of further example as applied to serve multiple different video streams to different clients.

In brief, the present transcoder 200 shown and described uses "thread" communications instead of "IPC" or "Inter Processor Communications" that are used according to many conventional transcoding techniques. For the purpose of this transcoder 200 description, the term "thread" is herein intended to mean an encapsulation of the flow of control in a program. Single-threaded programs are those that only execute one path through their code "at a time". Multithreaded programs may have several threads running through different code paths "simultaneously". In a typical process in which multiple threads exist, zero or more threads may actually be running at any one time. This depends on

the number of CPUs the computer on which the process is running, and also on how the threads system is implemented. While a machine or system with a number of n CPUs may be adapted to run no more than n threads in parallel, the threading operation according to the present transcoder invention may give the appearance of running many more than *n* "simultaneously" by sharing the CPUs among threads.

The transcoder 200 provides abstract APIs, and therefore the CODEC is accessed without the (much larger) native encoder overhead. Buffering 310 is created as a function of client pull for different video streams. Moreover, the transcoder 200 of the invention utilizes a network architecture – a single thread for each different connection, combining clients into same thread if they are at they are within the buffered segment of the same content. The transcoder's 200 use of threads in the manner herein shown and described is considered highly beneficial because a context switch between two threads in a single process is believed to be considerably cheaper (processing/memory/IO) than using a context switch between two processes. In addition, the fact that all data except for stack and registers are shared between threads makes them a natural vehicle for implementing tasks that can be broken down into subtasks that can be run cooperatively.

While various specific architectures may be built around the transcoder 200 embodiments just described in order to achieve particularly desired results on a case-by-case basis. However, for the purpose of further illustration, the following is an example of a more detailed system using the transcoder 200 described. The transcoder 200 is provided adapted to support a large number of simultaneous customer streams, each with differing formats. In particular, such system may support more than 5000 simultaneous streams, and in some circumstances more than 7000 simultaneous customer streams, each with differing video formats. Still further, the transcoder 200 may be implemented to convert any of a wide number of video sources to a format uniquely appropriate or required for many different individual clients each having differing needs. In one particular example, a transcoder 200 as herein described may be implemented to support such high demand simultaneously on any of the following formats: MPEG 1; MPEG 2; MPEG 4; Motion JPEG; AVI; H.261; H.263; H.263 +; RealVideo™; G-8; QuickTime™; Shockwave Flash™; Indeo Cinepak™; ASF.

It is further contemplated that the transcoder 200 may be adapted in an overall communication system to be compliant with all existing and soon anticipated fixed and mobile terminals and devices. Moreover, the transcoder 200 may be implemented to adapt output stream format variables to dynamically accommodate the channel and platform conditions of each client. Still further, the system incorporating the transcoder is adapted to support load balancing servers and routers for multi-

transcoder installations. Accordingly, it is believed that the transcoder 200 of the present invention delivers significantly greater functionality for significantly lower cost than other prior transcoding techniques and systems.

As described above, various different system architectures may incorporate the transcoder 200 of the invention without departing from the scope of the invention. However, more details of a particular architecture that is believed to suitably provide the beneficial level of support just described includes the following aspects: (i) dual P3-933 processor; (ii) any variant of Unix OS; (iii) 512 MB RAM; Redundant Firewire or Gigabit Ethernet; Redundant Power Supplies. Such system may be provided in a rack mounted configuration, or otherwise to suit a particular need.

The following aspects of the transcoder 200 of the invention should be contemplated as broadly beneficial, both independently and in various combinations as is apparent to one of ordinary skill based at least in part from this disclosure.

A system and method is provided for utilizing asynchronous software thread communication in both user and kernel space to perform efficient transcoding on multiprocessor and/or distributed computing platforms (such as clustering). It has been observed that this method is more efficient than utilizing traditional IPC methods to implement the transcoder. A shared library of CODEC algorithms is created and used to access the various CODEC algorithms, thereby incurring a lower processing overhead as well as lower memory utilization than that required by the traditional combined encoder functionality such as that used in the majority of commercial encoders. Of particular benefit, common threads may be used for multiple connections, and in fact even a single thread may be used for every individual connection using the present transcoder.

A system and method is also provided for combining multiple clients to be served by the same thread (for efficiency) whenever the same content is demanded and dynamic buffers (caches) can accommodate all of the data points demanded.

Media Compression and Delivery System

A data compression and delivery system 400 and method is also provided according to the invention for real time dynamic data signal processing for optimal reproduction of approximations of original media data over a given set of constraints. This system 400 and method is illustrated schematically by way of block flow diagrams in Figures 4A and 5. Further description of the various

beneficial features and operation of this system is provided as follows by way of exemplary embodiments generally incorporating by reference the description provided by these Figures 4A-5.

Figure 4 A ia a block diagram illustration of one embodiment of the data compression and delivery system 400 of the present invention. As shown in Figure 4A, the data compression and delivery system 400 comprises media module 405, dynamic player module 407, image processor 410, baseline snapshot module 415, classifier 417, quality of standard (QoS) module 420, network layer input module 425 and network output layer module 430. The system 400 further comprises a neural network processing module 440, timer 435, CODEC library module 445, dynamic client request module 450, ICMP module 455, device and network parameters measurement module 460 and delivery and transmit module 465.

In one embodiment, the system 400, resident at a server node(s), processes incoming uncompressed or previously compressed data. The system 400 employs neural networks 440 with artificial intelligence to monitor the incoming data to determine a plurality of key characteristics of each data segment. The system 400 correlates the incoming data characteristics with libraries 445 of pre-developed self-referencing experientially learned rules of the patterns in a scene in a sequence of frames in the input signal (e.g., a video signal) and with externally imposed constraints to optimally choose a preferred commercially available compression/decompression algorithm (e.g. CODEC) for each segment of the data. The system 400 then sets up an extensive array of usage controls, parameters and variables to optimize the chosen algorithm. Choice of algorithm and set up of parameters and variables will dynamically vary with each segment of incoming data depending upon the characteristics of the data as well as the evolving optimization process itself. The set of possible algorithms is numerous, limited only by availability and other commercial considerations. Each segment of data is encoded and compressed in the above manner and then served to a communications channel.

The compression system 400 just described is particularly useful as a streaming media compression engine, which, based upon information from the available CODEC's and the streaming media delivery system, performs frame-by-frame analysis of the incoming video using another artificially intelligent neural network 440. The system 400 then chooses the most appropriate compression format and configures the compression parameters for optimal video compression based on the best quality as measured by , in one embodiment, a selection of a peak signal to noise ratio from the underlying system environment. The result is the "optimal" video and audio service for the device and conditions present.

A more specific account of the artificial intelligence/neural network 440 aspect of this system as applied to streaming media signals is provided as follows. Initially, a library of separate and distinct CODECs are added to the system as a searchable CODEC library 445. Additional libraries of relevant reference information are also provided, including: a Network Transport Standards (NTS) library 443; and a Quality-of-Service (QoS) library 447. Then, a video (media source) is introduced either in a digitized or non-digitized format (AD conversion is used) via image processor 410. Image processor 410 then decompresses the source (if required) and employs various standard image processing algorithms used for "cleaning-up" the source image(s). The resultant source media is then passed to the baseline snapshot 415 repository where it will be used as a "perfect gold standard" for later comparison. Simultaneously, this resultant source media is also fed to the classifier 417.

The classifier 417 analyzes the source media for temporal, spatial and logical features for the purpose of creating source media sub-segments which exhibit similar combinations of temporal, spatial and logic features. "Similar" is defined to mean a contiguous sub-segment of the source media that contains common temporal, spatial and logic features that would lend themselves to a particular encoding/compression algorithm (as found in the CODEC library 445). This source media sub-segment (or, in one embodiment, a group of contiguous video and audio frames) is referred to as a "scene".

The neural network process 440 then operates upon this scene by employing CODECs from the CODEC library 445 to compress the scene. The internal configuration of each CODEC are manipulated/changed in accordance with inputs obtained from the NTS library 443, QoS library 447, Timer Process 435, Network Input Layer 425, ICMP agent 455 and the Device and Network Parameter measurement agent 460. The compressed scene is then decompressed and a comparison is made against the Baseline Snapshot 415 using a quality measurement made by the quality standard process 420. In one embodiment of the present invention, the Quality Standard Process 420 employs a peak signal noise ration (PSNR) algorithm in order to perform the comparison of the decompressed scene against the baseline snapshot of the source media. The comparison process is repeated with various CODECs from the CODEC library 445 until the Neural Network Process 440 is satisfied with the quality of the resultant compressed scene, within the constraints of the inputs received from the NTS library 443, QoS library 447, Timer process 435, Network Input Layer 425, ICMP Agent 455 and the Device and Network Parameter Measurement Agent 460. Finally, the resultant compressed scene is sent to the Network Layer Output 430 which transports the compressed scene to the Client using an appropriate Network Transport protocol and QoS algorithm.

The above process is repeated until the entire source media has been transmitted to the Client or until the process is aborted due to various possible conditions which may include: a Client request to abort, network transport failure, Client hardware failure, etc.

The NTS library 443 is a repository of network transport services that are selected and used by the Network layer output 430 to transport compressed source media to the Client and by the Network Layer Input 425 to receive information from the Client. The selection is based upon qualitative and quantitative inputs received from the Network Layer Input 425, ICMP agent 445 and the Device and Network Parameter Measurement agent 460.

The QoS library 447 is a repository of quality of service algorithms that are selected and used by the network layer output 430 to transport compressed source media to the Client. The selection is based upon qualitative and quantitative inputs received from the Network Layer Input 425, ICMP agent 455 and the Device and Network Parameter Measurement agent 460.

The ICMP agent 455 generates inputs to the neural network process 440 that dynamically provides it with the quantitative and qualitative characteristics of the transport in use between the processor and the client. In one embodiment of the present invention, the ICMP protocol is used for this purpose.

The Device and Network Parameters Measurement agent 460 generates inputs to the neural network process 440 that dynamically provides it with the qualitative and quantitative characteristics of the client's environment. In one embodiment of the present invention, these client environment characteristics include central processing unit (CPU) capacity, network interface characteristics, storage capacity and media rendering devices capabilities.

Still referring to Figure 4A, the Network Layer Input 425 provides inbound (originating from the client) network transport services. The Network Layer Output 430 provides outbound (originating from the processor) network transport services. The Timer Process 435 provides a way for the user of the invention to limit the maximum amount of time that the Neural Network Process 440 will spend in processing a given source media.

Figure 4B is a block diagram illustration of one embodiment of a CODEC selection scheme of the neural network processing module 440 of one embodiment of the present invention. The neural network processing module 440 shown in Figure 4B comprises a video frame selection module 475, CODEC parameters module 480, input layer module 485 , hidden layers 486 – 487 and output module 488. In one embodiment of the present invention, a CODEC representative signal suitable to be used as

a reference baseline signal for incoming signals to the neural network processing module 440 is generated by the neural network processing module 440. In one embodiment, the classifier 417 determines which scenes in segments of an incoming video signal represents the best scene in light of the available parameters of the underlying CODEC. A list of standards are used by the neural network processing module 440 to determine which scene in the signal represents the best scene. In one embodiment, the Neural Network Process 440 samples a number of pixels in a particular frame of video to determine changes in the number of pixels in that particular frame vis-à-vis the pre-determined parameters of the video signal. In another embodiment, significant motion changes in a particular scene in the video signal may be used as the baseline reference scene ("best scene") for subsequent incoming video.

In one embodiment of the present invention, the neural network processing module 440 takes a segment of video from the classifier 417 as an input and subsequently takes a sample of this input to derive enough information that characterizes the video signal. For example, in the scheme illustrated in Figure 4B, the Neural Network Process 440 takes a window snap-shot (e.g., a 176 X 144 pixel window) to examine. It is advantageous for the Neural Network Process 440 to look at the center of the sample window to generate enough information about the video signal. In one embodiment of the present invention, the Neural Network Process 440 uses a minimum of 8 frames to generate the requisite information about the video signal. Information from the sample window is presented with the particular CODEC parameters from parameter module 480 to the input layer 485.

The input layer 485 is coupled to a plurality of hidden layers 486 – 487 via a plurality of neurons with each connection forming either a strong or weak synoptic link from one neuron to the other. In one embodiment, each CODEC supported by the neural network processing module 440 is provided with its own neural network to process the CODEC specific parameters that come with the particular CODEC. The Neural Network Process 440 generates the "best" video signal through a round-robin like process referred to as a "bake-off" from the plurality of CODECs processed during a video sampling capture period. In processing the best video representation from incoming signals, each of the corresponding neural networks for each of the CODECS generates the best representative sample from the hidden layers 486 – 487 and feed the signal to the output module 488. In one embodiment of the present invention, the output data set of the best CODEC from each class of CODECS being processed by the Neural Network Process 440 has two possibilities. The first being the Neural Network Process 440 submitting the best results for each CODEC to the output module 488 to a "bake-off" neural network of the plurality of "best" samples for each of the plurality of CODECS which in

turn generates the winning best CODEC from the plurality of best CODECS. The bake-off neural network is smaller and faster than the neural networks that handle the processing of the CODECS.

In a second processing scheme, the Neural Network Process 440 may implement a genetic algorithm processing of the best CODECS generated by the plurality of CODECS. The genetic algorithm follows the same statistical selection approach of a marble game. Thus, instead of feeding the winning output CODEC from the various neural networks into a "bake-off" neural network, a genetic algorithm processing may be applied to feed the output module 488 from the various neural networks into a bucket and picking the best CODEC representation from a collection of scenes at the end of the source media, for example, a movie, etc. In one embodiment of the present invention, the Neural Network Process 440 uses a combination of forward and backward propagating algorithm to process the CODECS.

Referring back to Figure 4A, for the purpose of providing a further understanding of this artificial intelligence process, the following example of one particular application is provided. It is to be appreciated that the features and operation of the system provided by this exemplary application are to be considered as broadly descriptive of the neural network 440 aspect for data compression and delivery according to the invention. Other applications may be made and fall within the scope of the invention.

A video content provider installs the system of the present invention on its server. Sample videos are introduced to the system in order to perform an initial AI process as described above. A complex matrix of CODEC characterizations, e.g. for each bit rate, pattern of video, etc., is created to be drawn from later. Next, a client end-user connects to the content provider system in order to view a video M. The communication system of the invention residing on the server delivers a software agent to the client's device, thus enabling the client to connect to the communication system in order to deliver device-specific information and receive the appropriate compressed signal with decompression CODEC for playing. Next, the AI system begins loading the video M as a streaming signal into a buffer for the purpose of choosing the appropriate CODEC for each frame and compressing each frame appropriately for transmission. The time period of the buffer depends upon multiple variables, principally the processing power of the system, and may be generally for example approximately 15 seconds for systems having appropriate capability for pre-recorded but uncompressed video media. Within the buffer, each frame is compared against each CODEC according to the "types" of sequences pre-tested in matrix as depicted in the diagram.

Next, the system 400 looks at end-user parameters, e.g. screen resolution, memory available, via information received from the software agent in the client's device. The most appropriate CODEC is then chosen and configured/tuned for optimal performance by setting certain variables within the CODEC to fixed quantities (e.g. based on comparing source video vs. patterns in past, transmission channel capabilities or constraints, and destination device capabilities or constraints). The process just described is generally done frame-by-frame by the classifier 417, but the CODECS are compared for temporal compression efficiency such that the process for each frame contemplates other leading and lagging frames. Once the appropriate CODEC is chosen and tuned for each frame (or block of frames where appropriately determined automatically by the system), the delivery system reports to the client agent and delivers the tuned CODEC ahead of the corresponding frame(s) to be decompressed and played.

It is to be appreciated that the neural network 440 of this system 400 continuously learns and remembers the performance and operation of the CODECS within the CODEC library 445, and continuously uses its learning to improve the compression efficiency of the input media signal. The process of running a signal frame through the library, modifying CODEC operating parameters, comparing compression performance by the compare logic 525 (Figure 5) against reference standard compression, and running the loop again with further modifications, is an iterative 550 (Figure 5) one that generally continues to improve compression efficiency. In fact, compression with one or more CODECS in the library 445 may reach improved levels better than the reference compression algorithm(s).

Nevertheless, when time constraints 435 (Figure 4A) are present (such as in real-time push or pull demand for the streaming media content), this process must eventually be stopped at some point so that a particular frame or series of frames being processed may be compressed 575 and delivered 580 to the destination without unacceptable delay by timer 435. Then, the next frame or series may be operated upon by the neural network 440 within the CODEC operating system. These endpoints may be defined by reaching a predetermined desired result, such as for example but without limitation: (i) reaching a predetermined percentage (%) compression efficiency, such as for example as compared to the reference standard; or (ii) reaching a predetermined or imposed time limit set on the process, such as for example according to a time related to the buffer time (e.g. 15 seconds); or (iii) the earlier occurrence of either (i) or (ii). In any event, though an endpoint is reached for choosing the appropriate CODEC and performing the compression 575 and delivery 580 operations, this does not mark an endpoint for the neural network 440 training which continues. The information that is gathered through

each loop in the process is stored 550. When subsequent similar frames or system constraint parameters in an incoming frame are encountered 545 in the future, the stored information is remembered and retrieved by the neural network 440 for improving compression 575 and delivery 580 efficiency.

While many different communication protocols are contemplated, one particular embodiment which is believed to be beneficial uses a "full duplex network stack" protocol, which allows for bi-directional communication between the server and the client device. Again, while other protocols may be appropriate for a particular application, the full duplex system is preferred.

The system 400 just described addresses the difficulties encountered with previously known CODEC systems by utilizing the streaming media delivery architecture to overcome latency issues and the embedded neural network 440 to overcome speed concerns. The system 400 is then able to reconfigure the algorithms used for compression in the neural network 440, the goal being to achieve optimum results every time over any network configuration.

A wide variety of CODECS may be used within the CODEC library 445 according to the overall compression systems and methods just described, though beneficial use of any particular CODEC according to the invention contemplates such CODEC taken either alone or in combination with other CODECS. For example, an appropriate CODEC library 445 may include one or more of the following types of CODECS: (i) block CODECS (e.g. MPEG versions, such as Microsoft Media™ or QuickTime™); (ii) fractal CODECS; and (iii) wavelet CODECS (e.g. Real™). According to another aspect, an appropriate CODEC library 445 may include one or more of the following types of CODECS: (i) motion predictive CODECS; and (ii) still CODECS. Still further, the CODEC library 445 may contain one or more of the following: (i) lossy CODECS; and (ii) lossless CODECS.

In one embodiment of the present invention, all of these different types of CODECS may be represented by the CODEC library 445 according to the invention; and, more than one particular CODEC of a given type may be included in the library. Or, various combinations of these various types may be provided in order to achieve the desired ability to optimize compression of a streaming media communication over a wide range of real-time variables in the signal itself, transmission channel constraints, or destination device constraints. Still further, an additional highly beneficial aspect of the invention allows for new CODECS to be loaded into the library 445 and immediately available for use in the neural network 440 compression/delivery system 400. Nevertheless, one particular example of a CODEC library 445 which is believed to be beneficial for use in optimally communicating a wide range of anticipated streaming media signals, and of particular benefit for image signals, includes the

following specific CODECS: MPEG versions 1, 2, and 4 (e.g. Microsoft Media™ and QuickTime™); DUCK TruMotion™; ON2; Real Media™; MJPEG: H.261; H.263; H.263+; GIF; JPEG; JPEG2000; BMP; WBMP; DIVX.

The following are further examples of various aspects of the compression system and method just described that should be considered as broadly beneficial, both independently and in various combinations as is apparent to one of ordinary skill based at least in part on this disclosure. Further examples of such broad aspects are elsewhere provided in the "Summary of the Invention" as well as in the appended claims.

Use of neural networks 440 with artificial intelligence to achieve the various CODEC operations described is broadly and uniquely beneficial. In particular, a system and method is provided for pre-processing 410 of source data determined by application of learned responses to the signal quality, data content and format of the data. A system and method is provided for processing each unit (e.g. frame or block of frames) of source data by selection and application of a suitable CODEC (from a set of all available CODECS in the CODEC library 445) dependent upon observed characteristics of the source data and application of past-learned responses to compressing similar data. A system and method is provided for processing each unit of source data by setting a multiplicity of compression characteristics within a chosen compression algorithm to optimize capture and preservation of the original data integrity. Still further, each or all of the aforementioned signal processing steps is applied to each unique, sequential unit of signal data, e.g., signal clip, video frame, or individual packet as appropriate.

It is further contemplated that a CODEC management system 400 according to the invention provides a system and method for image processing that is adapted to normalize original source data/images as well as to resize and resample original data to fit the specification of the neural network processing module 440. An ability to serve any transmission or recording channel with a single system and with any source data stream is also provided. Moreover, the various systems and methods herein described, individually and beneficially in combination, are provided with compatibility to any connection or connectionless protocol, including but not limited to TCP, UDP, WTP/WDP, HTTP, etc.

The invention as herein shown and described also allows for highly beneficial applications for accelerating the learning rate of neural networks 440 while minimizing the data storage requirements to implement said networks. Different classes of data streams each have unique characteristics that require substantially greater processing by neural networks 440. For example, video data streams differ by prevalence and degree of motion, color contrast, and pattern and visibility of details. Greater

Ex. 1018-1202

processing requires longer times to reach optimal functionality. Greater processing also requires more predictive library storage, often growing to unlimitedly large sizes. For real-time neural network processing, processing time and storage can be minimized to greatly increase functionality by providing pre-developed predictive libraries characteristic of the class of data stream.

Accordingly, the following are examples of aspects of the pre-trained neural network 440 aspects of the invention that should be appreciated as broadly beneficial, both independently and in combination (including in combination with other embodiments elsewhere herein shown and described). A system and method is provided that creates and uses artificial intelligence in a neural network 440 and pre-trains that intelligent network for use in solving a problem, which problem may be for example but not necessarily limited to streaming media compression according to a particular beneficial aspect of the invention. A system and method is also provided for subdividing the universe of problems to be solved into useful classes that may be processed according to a learned history by the intelligent network.

An intelligent streaming media delivery system and method is also provided according to the invention that manages content transmission based on end-user capabilities and transmission channel constraints, such as for example, but without limitation, available transmission speeds or bandwidth, and Internet congestion. The data compression and delivery system 400 utilizes a computer implemented intelligence process, such as an artificial intelligence process based on a neural network to analyze aspects of the connection (including without limitation differing bit rates, latencies, transmission characteristics and device limitations) to make modifications in the compression methodology and to manage Quality of Service ("QoS") 420 issues. Compressed, digital, restorable and/or decompressible data streams may be therefore delivered to a multiplicity of different local and/or remote devices via a multiplicity of transmission mediums characterized by differing capabilities. In addition, a decompression system is provided for reproducing the decompressed data at the terminal device.

In one beneficial embodiment, a terminal device establishes a link with the system resident on a server node(s). Except for software normally required to establish communications, the terminal device might not initially have resident software embedded therein associated with the present system. Upon linking the terminal device to the server node, the system transmits a software agent to the terminal device that cooperates with other software modules on the server-side that together form the overall delivery system. The software agent informs the system of the terminal device configuration and processing capacities for decompressing and displaying the data. The software agent also reports

certain relevant information to the system of the characteristics of the communication channel between the terminal and the server. Such information includes, without limitation: latency, bandwidth, and signal path integrity. Based upon terminal device configuration and real time updates of channel characteristics and capabilities, the system actively manages transmission of the compressed data stream by varying parameters such as buffer length, transmitted bit rate, and error correction. The system also feeds operating conditions to the compression system to dynamically alter encoding and compression settings to optimize delivery of the data. The delivery software agent resident on the terminal device decompresses the data stream that is composed of segment-by-segment variations in compression/decompression algorithm and settings thereof. Dependent upon the terminal device configuration, and especially for very thin clients, instructions may be refreshed on a segment-by-segment basis for each decompression algorithm and encoding setting combination. Instructions for decompressing may also be kept resident if appropriate to the terminal device.

The software agent described for transmission to and operation by the destination device is therefore also considered a highly beneficial aspect of the compression/delivery systems and methods described. By delivering the software agent to the device from the source, a wide range of existing destination devices may be used for communication according to methods that may include variable uses of one or more algorithms or other operations at the transmission source. In other words, the destination devices may not be required to be "format-specific" players as is required by much of the conventional streaming and static media communication systems. Also, by providing the destination agent with a diagnostic capability, diagnostic information may be gathered at the destination device and transmitted back to the source in a format that is compliant for use by the source in its neural network process for achieving the proper CODEC operation for a given set of circumstances.

The use of a client-side agent to supply quality of service information including client-side device data and communication channel status in real time is therefore also believed to be broadly beneficial beyond the specific applications and combinations of other aspects of the invention also herein provided. In addition, the processing of each unit of compressed, transmission-ready data to accommodate client-side device and real-time communication channel conditions is also broadly contemplated as having broad-reaching benefits. Still further, a system and method is described that provides instructions to a client-side agent to enable decompression of each sequential, uniquely compressed unit of data. Therefore, another broad benefit of the invention provides a destination device (such as from the transmission source as herein described for the particular embodiments) with a CODEC that is adapted to decompress a compressed representation of an original media signal into a

Ex. 1018-1204

decompressed representation based upon variable parameters related to at least one of the following: aspects of the original media signal, transmission channel constraints, and destination device constraints. In another broad aspect, the destination device is adapted to use a CODEC that is chosen from a library of CODECS based upon a parameter related to an aspect of the original media signal.

The systems and methods herein described are also considered applicable to the signal processing of each unique, sequential unit of signal data, e.g., signal clip, video frame, or individual packet as appropriate. In addition, the system and its various sub-systems may also be purely software that must be loaded into each appropriate device, or it may be embedded in a host hardware component or chip, e.g. on the server side, or in certain circumstances, on the client side (e.g. various aspects of the destination agent), or for example may be stored such as in flash memory.

The various aspects of the media compression system and method just described are considered beneficial for use according to a wide range of known and soon anticipated media communication needs, including for example according to the various communications devices, communication/transmission channel formats and standards, and media types and formats elsewhere herein described (e.g. in the "Background" section above).

However, for the purpose of further understanding, Figure 6 shows a schematic view of the overall streaming media communications system 600 as specifically applied to "video-on-demand" aspects according to one embodiment of the present invention, wherein many different end users 610 – 620 at many different locations may request and receive, real-time (e.g. without substantial delay), pre-recorded video from a remote source. Further to the information provided in Figure 6, at least one specific implementation of the media communication system 600 delivers the following types of video at the following bit-rates (denotes compressed representations of original signals that are convertible by a destination device to decompressed representations having no or insubstantial loss as observed by the eye of the typical human observer): VHS-format video as low as about 250Kbs; DVD-format video at about 400 Kbps; and HDTV-format video at about 900Kbps. According to these speeds, it is believed that video-on-demand may be provided by telephone carriers over resident transmission line channels, such for example over existing DSL lines 630 - 640.

However, as available bandwidth and mass communication continue to present issues, it is believed that even greater efficiencies may be achieved resulting in delivery of compressed representations of these types of video signals at even lower bit rates. Again, as elsewhere herein described, the compression efficiencies of the invention are closely related to and improve as a function of the processing power made available to the neural network 440, and the neural network's 440

continued learning and training with respect to varied types of media. These resources may even make more remarkable compression efficiencies achievable without modification to the fundamental features of the present invention.

Therefore, the following are further examples of transmission rates for certain compressed video signals that are believed to be desirable and achievable according to one embodiment of the invention: VHS-format video as low as about 200Kbps, more preferably as low as about 150Kbps, and still more preferably as low as about 100Kbps; DVD-format video as low as about 350Kbps, more preferably as low as about 300Kbps, and still more preferably as low as about 250Kbps; and HDTV-format video as low as about 800Kbps, and still more preferably as low as about 700 Kbps.

Moreover, at least one implementation of the media communications system 400 of one embodiment of the invention delivers 20-24 frames/sec color video at a transmission rate of 7 Kbps. This is believed to enable substantial advances in communication of streaming media signals via to wireless destination devices via the WAP Gateway, as is further developed elsewhere hereunder.

It is also to be appreciated that, while video communication has been emphasized in this disclosure, other types of streaming or static media are also contemplated. For example, at least one implementation of the compression and delivery embodiments has been observed to provide substantially CD-quality sound (e.g. via compressed representations of original signals that are convertible by a destination device to decompressed representations having no or insubstantial loss as observed by the ear of the typical human observer) at a bit-rate of about 24 Kbps. At these rates, audiophile-quality sound may be delivered for playing over dial-up modems. However, with further regard to available resource commitment and extent of neural network training, it is further contemplated that the invention is adapted to deliver CD-quality sound at speeds as low as about 20Kbps, and even as low as about 15 Kbps or even 10Kbps.

Wireless Audio Communications System

It is further contemplated that the streaming media communication system of the invention has particularly useful applications within wireless audio communications networks, and in particular cellular telephony networks. Therefore, Figures 7 and 8 schematically show, with respectively increasing amounts of detail, streaming media communications systems 700 and 800 respectively specifically applied to wireless audio communications systems according to certain specific, respective embodiments of the present invention. While particular devices, system parameters, or arrangements

of communicating devices shown are believed to be beneficial in the overall application of the invention, they are not to be considered limiting and may be suitably replaced with other substitutes according to one of ordinary skill based upon this disclosure. The various wireless communications systems 700 and 800, standards, and protocols referenced elsewhere in this disclosure are thus incorporated into this section for the purpose of integration with the various aspects of compression, delivery, decompression, and transcoding according to one embodiment of the invention.

Combination of the communications system 400 of one embodiment of the present invention with the other components of a cellular communications network allows for the enhanced compression, delivery, and decompression according to the invention to manifest in an increased quality of service for wireless audio communications. Improvements in cellular communications according to the invention include, without limitation, the following examples: increasing available bandwidth, extending range of reception, and providing graceful degradation while maintaining connectivity during periods of low signal quality or reception levels.

More specifically, cellular telephony signals are characterized by relatively high degrees of variability, due for example to the roaming positions of clients, and limited cell ranges, atmospheric conditions, and significantly limited and changing available bandwidths over daily use cycles. Therefore, a self-optimizing CODEC management system according to the present invention is particularly well suited to adjust the appropriate communications and compression modalities to the changing environment. At the very least, the increase in compression efficiency and resulting decrease in bandwidth used for given signals is a valuable achievement as wireless channel traffic continues to congest.

In one particular regard, the increased compression efficiency according to the present invention is well applied to improving bandwidth issues during "soft hand-offs" between cells, as illustrated in Figure 9. During cellular phone communications, whenever a transmitter or receiver migrates between cell coverage zones, communications bandwidth requirements and resultant costs are increased by systemic requirement to "pass off" active communications between cells. The act of passing off the communication results in a "backhaul" channel from the previously active cellular transmitter to a central office for forwarding to a newly active cellular transmitter. The backhaul channel represents a significant use of bandwidth. Savings will result from increased compression. As Figure shows, such "backhauling" may include a doubling (media sent back from first cell being left and resent to second cell for transmission) or even a quadrupling (overlapping communication from both first and second cells) in the bandwidth used for communicating a particular signal.

Ex. 1018-1207

The media communications system 400 of the present invention may recognize when backhaul is occurring, such as according to the transmission channel diagnostics provided in the software agent(s), and may respond by adjusting the degree of compression to compensate.

WAP Video Gateway

With a particular view of the rapid growth observed and predicted in the wireless or mobile Internet, embodiments of the present invention contemplate application of the intelligent compression/delivery/ decompression embodiments in combination with WAP Gateway functionality.

A system and method is therefore also provided according to the invention for encoding, compressing and transmitting complex digital media (e.g., video pictures) via bandwidth-constrained wireless communications systems utilizing Wireless Applications Protocol (WAP). In one embodiment, data is processed by the system, resident at a server node(s), employing neural networks with artificial intelligence. Sample segments of data are captured from the input stream and processed to comply with requirements unique to the class of clients. As is described in detail above, the system correlates the continuously varying digital data streams' characteristics with libraries of pre-developed experientially learned rules and with externally imposed constraints to optimally choreograph the coherence, continuity and detail of the data as ultimately received, decoded and presented at the client interface.

A gateway provided with the added functionality of the streaming media communications system herein described is shown schematically in Figure 8. According to the WAP gateway system 830 , a client agent is provided that is adapted to run on a variety of platforms, and requires no specialized hardware to decode the video streams. According to use of the streaming media delivery system of the invention elsewhere herein described, the viewer of the WAP device maintains constant communication with the system server upstream, such that the user-side client 825 may provide the encoding platform with relevant information for streaming media communication, including without limitation: available screen size, processing power, client operating system and browser version, connection speed and latency, thereby allowing the streaming media delivery system to tailor the stream to each individual client it "talks" to. Accordingly, an AI driven server 830 incorporating the AI compression as herein described may be combined with a WAP Gateway 830, combining the necessary WAP to TCP/IP protocol (or other protocol, e.g. dual server stack) translation with a Video and Audio Server 835 employing compression, delivery, and decompression systems and methods

herein described. The WAP Gateway 830 may further include a video transcoder, such as for example incorporating the transcoder systems and methods herein described. An appropriate host architecture according to this system (not shown) generally includes a rack mount system running Linux OS with a modified WAP Gateway 830 or as a software plug-in to existing servers.

This WAP gateway system 830 may be further provided in a Master/Slave relationship as another beneficial aspect of the overall streaming media delivery architecture (applicable to other delivery systems other than specifically wireless). Various content distribution networks, such as available through Akamai and Inktomi, have capitalized on the concept of improving data delivery over the Internet by using "smart caching" on servers which reside on the borders of the Internet. Such a Master/Slave relationship is maintained by the present system wherein a Master Server resides at the source of the content to be delivered and Slave Servers reside on the borders. These servers communicate "intelligently" to optimize the content delivery over the Internet and reduce latency, bandwidth and storage requirements, improving the overall quality of the video/audio stream to the end-user and decreasing the cost of media delivery to the content provider.

The WAP gateway 830 of the present invention supports continued growth in mobile communications, as large telecommunications operators are transitioning to multi-service broadband networks, and as the number of subscribers to the mobile Internet continues to expand rapidly. In particular, mobile communications is a broad class of systems and protocols, each having its own constraints and needs for interacting devices to communicate streaming media. The Gateway 830 in a particularly beneficial aspect may support a variety of "2G" systems with upgradability for upcoming "2.5G" and "3G" network technologies (numerical progression of systems generally represents progression of Internet-enabled capabilities).

The following Table 3 provides examples of known mobile communication standards, and provides certain related information used by the AI system of the present invention for optimizing communication of streaming media amongst the field of mobile destination devices as media players:

Table 3: Existing/Soon Anticipated Mobile Communications Standards

MODE	**BAUD RATE (generally)**
GSM (2G)	9.6 Kbps
CDMA	9.6 Kbps
TDMA	14.4 Kbps
CDPD	14.4 Kbps

iMODE	128 Kbps
GPRS	144 Kbps
WCDMA or CDMA2000	144 Kbps to 2 Mbps
GSM (3G)	2 Mbps

In addition, the present invention is particularly beneficial in its ability to stream a wide variety of media signals to various different types of wireless communications devices. Examples of wireless communications devices that are appropriate for use with the streaming media communications systems and methods of the invention, and which the systems and methods support interchangeably, are provided in the following Table 4:

Table 4: Examples of Internet-enabled PDA's

DEVICE	MAKE	SPEED	MEMORY	SCREEN DEPTH	SCREEN SIZE	MODEM TYPE	CONNECT SPEED
iPAQ	Compaq	206 MHz	16-64 Mb	12 b/pixel color	320x240	External (e.g. CDPD)	9600-14.4 Kbps
PalmVII	Palm	33 MHz	4 -16Mb	4 b/pixel b/w 8 b/pixel color		Internal	14.4 Kbps
Handspring	Palm	33 MHz	4 -16Mb	4 b/pixel b/w; 8 b/pixel color		External	14.4 Kbps
Blackberry	Research-In-Motion	33 MHz	4 Mb	2 b/pixel b/w		Internal	9.6-14.4 Kbps
Jornada	HP	133 MHz	16-32 Mb	18 b/pixel	320x240	External	9.6-14.4 Kbps

Casseopeia ▪▪▪▪▪▪▪▪▪	Casio	150 Mhz	16-31 Mb	12 b/pixel	320x240	External	9.6-14.4 Kbps

Various specific examples are described later below that provide observations of actual wireless Internet applications of the invention as herein described. Such examples include use of a CODEC library according to varied parameters associated with at least the following (without limitation): destination wireless communication device; transmission channel; communications protocol; and the respective streaming media signals themselves. The various particular features of the systems and methods used according to these examples are contemplated as further defining independently beneficial aspects of the invention.

Shared Interactive Environment

A system and method is also provided according to the invention for enabling real-time remote client interaction with a high-definition, multi-dimensional, multi-participant simulated environment without the need for significant client-side processing capacity. More specifically, Figure 10 shows an overall streaming media communication system as applied to shared interactive gaming according to the invention.

This system includes: (i) a proxy server; (ii) graphics rendering capabilities; (iii) a client software agent for feedback of client inputs to the game; (iv) a client software agent for supporting the delivery system of the invention; and (v) streaming from the server to the client. It is contemplated that for multiple clients, which typically represent shared interactive gaming by design, multiple components as just described are provided to support each client.

The interactive gaming embodiments contemplate implementation of data compression and delivery embodiments with devices that are also destination devices for compressed signals from other like, remotely located device systems. This arrangement is broadly beneficial, such as for example in further interactive media implementations such as video conferencing and the like. Accordingly, each remote system is both a source and a destination device, and sends and receives agents between it and other remote systems.

Destination Device

Although the communications systems of the present invention enables communication of streaming media signals to a wide variety of destination devices, a further contemplated feature of the invention provides a remote receiver to be housed as a destination device/player by client users. This set-top player may be adapted to serve at least one, though preferably multiple ones, and perhaps all, of the following: Video on Demand (VOD); Music on Demand (MOD); Interactive Gaming on Demand (IGOD); Voice Over Internet Protocol ("VoIP"), any technology providing voice telephony services over IP connections; Television Web Access; Digital Video Recording to record, pause, and playback live television; e-mail; chat; a DVD player; and other applications apparent to one of ordinary skill. All of this may be delivered to existing televisions in the comfort of users' own homes. Moreover, clients utilizing this box, or other systems interfacing with the communications system of the invention, may receive DVD quality video and surround sound over cable and DSL connections.

EXAMPLES

For the purpose of further illustrating the highly beneficial results that may be achieved according to the invention, the following are examples of specific embodiments that have been used for different types of streaming media communication, including observed results with pertinent discussion. These examples illustrate communication the same pre-recorded video over different transmission channels and to different destination devices, wherein the pre-recorded video has the following originating properties: 720lines of resolution and 32bits of color information, an originating file size of about 1.4Gigabytes.

Example 1

An "iPAQ" model 3650 hand-held PDA (commercially available from Compaq, Inc. for approximately $500 at the time of this disclosure) was provided. The PDA was interfaced with a 14.4 Kbps (max) wireless CDPD modem ("AirCard 300" wireless external modem, commercially available from Sierra Wireless for approximately $200 at the time of this disclosure) using an extension assembly (iPAQ™ PCMCIA expansion sleeve from Compaq, Inc.) with a PCMCIA card slot that couples to the wireless modem. The iPAQ™ used is generally characterized as having the following processing parameters: 206 MHz processor; 32 Mb memory; 12 b/pixel color; 240X320 screen dimensions; PocketPC™ operating system version 3.0 from Microsoft Corp. and stereo sound. The iPAQ™ was

Ex. 1018-1212

connected to the Internet in San Francisco, California via the interfaced CDPD modem over the AT&T cellular wireless carrier system at a connection bandwidth of about 13.3Kbit/sec. A server located in San Jose, California (approximately 50mi away) was contacted by the PDA employing the http and rtsp protocols, and the PDA was used to initiate a request for a pre-recorded video having the following originating properties: 720lines of resolution and 32bits of color information, the originating file size was 1.4Gigabytes. Within about seven seconds, a compressed approximation of the pre-recorded video was received, decompressed, and displayed by the PDA on the PDA's screen. The entire video was seen at 240x320x12bpp resolution in full motion without observable delays or defects.

Example 2

A "Jornada™" model 548 hand-held PDA (commercially available from HP, Inc. for approximately $300 at the time of this disclosure) was provided. The PDA was interfaced with a 9.6 Kbps (max) wireless CDMA phone ("Motorola i85s" wireless external digital cellular phone, commercially available from Motorola authorized vendors for approximately $200 at the time of this writing) using adaptor cables (Motorola and HP RS-232 standard interface cables from Motorola and HP.) that couple the phone and PDA together to form a wireless modem. The Jornada model PDA device used is generally characterized as having the following processing parameters: 133 MHz processor; 32 Mb memory; 12 b/pixel color; 240X320 screen dimensions; PocketPC™ operating system version 3.0 from Microsoft Corp. and stereo sound. The Jornada™ was connected to the Internet in Newark, NJ via the interfaced CDMA phone/modem over the Nextel digital cellular wireless carrier system at a connection bandwidth of 8Kbit/sec. A server -located in San Jose, California (approximately 2900mi away) was contacted by the PDA employing the http and WDP protocols, and the PDA was used to initiate a request for a pre-recorded video having the following originating properties: 720lines of resolution and 32bits of color information, the originating file size was 1.4Gigabytes. Within about seven seconds, a compressed approximation of the pre-recorded video was received, decompressed, and displayed by the PDA on the PDA's screen. The entire video was seen at 176x120x8bpp in full motion without observable delays or defects.

Example 3

A "Set-top Box" model st850 book PC (commercially available from MSI, Inc. for approximately $300 at the time of this writing) was provided. The Set-top Box was interfaced with a 10 Mbps (max) ethernet/802.11 connection using CAT5 ethernet cables (Generic) that couple the Set-top Box to a broadband connection (DS3). The Set-top Box used is generally characterized as having the following processing parameters: 400 MHz processor; 64 Mb memory; 32 b/pixel color; 720 lines of screen resolution; Windows CE operating system version 2.11 from Microsoft Corp. and AC3 digital 6 channel surround-sound. The Set-top Box was connected to the Internet in Newark, NJ via the interfaced shared DS3 connection over the Alter.Net Internet Backbone at a connection bandwidth of 376Kbit/sec. A server located in San Jose, California (approximately 2900mi away) was contacted by the Set-top Box employing the http and rtsp protocols, and the Set-top Box was used to initiate a request for a pre-recorded video having the following originating properties: 720lines of resolution and 32bits of color information, the originating file size was 1.4Gigabytes. Within
About nine seconds, a compressed approximation of the pre-recorded video was received, decompressed, and displayed by the Set-top Box on a commercially available reference monitor's (Sony) screen. The entire video was seen at 720linesx32bpp in full motion without observable delays or defects.

While various particular embodiments have been herein shown and described in great detail for the purpose of describing the invention, it is to be appreciated that further modifications and improvements may be made by one of ordinary skill based upon this disclosure without departing from the intended scope of the invention. For example, various possible combinations of the various embodiments that have not been specifically described may be made and still fall within the intended scope of the invention. According to another example, obvious improvements or modifications may also be made to the various embodiments and still fall within the intended scope of this invention.

What is claimed is:

1. A system for processing a media signal, comprising:

a searchable CODEC library that is adapted to store a plurality of separate and unique CODECS and also to be accessed by and searched by at least one of an artificial intelligence (AI) system and an operating system, such that each CODEC stored in the CODEC library may be individually located and accessed for use by at least one of the AI system and operating system in order to compress the media signal for transmission from a source to a destination device.

2. The system of Claim 1, wherein the CODEC library comprises at least one CODEC that is of at least one of the following three types of CODECS: block CODECS; fractal CODECS; and wavelet CODECS.

3. The system of Claim 2, wherein the CODEC library comprises:
a first CODEC of a first type on the three types, and
a second CODEC of a second type of the three types.

4. The system of Claim 3, wherein the CODEC library comprises:
a first CODEC of a first type of the three types,
a second CODEC of a second type of the three types; and
a third CODEC of a third type of the three types.

5. The system of Claim 4, wherein the CODEC library comprises:
a first CODEC that is a motion predictive CODEC; and
a second CODEC that is a still CODEC.

6. The system of Claim 5, wherein the CODEC library comprises:
a first CODEC that is a lossy CODEC; and
a second CODEC that is a lossless CODEC.

7. The system of Claim 1, further comprising:

a registry associated with the CODEC library and that is searchable in order to search and locate a particular CODEC within the CODEC library.

Ex. 1018-1215

8. A system for communicating a media signal from a source to a destination device over a transmission channel, comprising:

a destination agent that is adapted to be used by the destination device to communicate with the source and to receive a compressed representation of the media signal from the source and to decompress the compressed representation into a decompressed representation of the media signal such that the decompressed representation may be stored or used by the destination device.

9. The system of Claim 8, wherein the destination agent comprises:

a diagnostic agent that is adapted to provide diagnostic information for transmission to a neural network associated with the source, which diagnostic information represents an ability of the destination device to receive, store, process, or display the media signal; and

a decompression agent that is adapted to decompress the compressed representation of the media signal into the decompressed representation of the signal to be stored or played by the destination device.

10. The system of Claim 9, wherein the destination agent further comprises:

a decompression agent that is adapted to receive a compressed representation of the media signal and also to receive a CODEC decompressor associated with the compressed representation of the signal from the delivery system, and that is adapted to use the CODEC decompressor to decompress the compressed representation of the signal into a decompressed representation of the signal to be stored or played by the destination device.

11. The system of Claim 10, wherein the destination agent is adapted to be stored and operated for decompressing and playing a media signal on a plurality of different destination devices each having different capabilities for decompressing, storing, or playing the media signal.

12. The system of Claim 11, wherein the destination agent is adapted to decompress a plurality of different compressed representations of the media signal by using different CODEC decompressors that are each associated with a corresponding different CODEC compressor used to compress the signal.

13. The system of Claim 12, wherein the destination agent is adapted to decompress the compressed representation of the media using a CODEC that is provided based upon at least one of the following parameters: a computer implemented previously learned behavior of the CODEC with

Ex. 1018-1216

respect to a reference media signal, a computer implemented previously learned behavior of the CODEC with respect to a prior attempt at compressing or decompressing the same media signal, a comparison of the CODEC's operation with respect to the media signal against a reference algorithm compression of the media signal, a learned constraint of the transmission channel, and a learned constraint of the destination device.

14. The system of Claim 13, wherein the destination agent is adapted to interchangeably use a block CODEC, a fractal CODEC, or a wavelet CODEC.

15. A system for communicating a media signal from a source to a destination device over a transmission channel, comprising:

a CODEC that is adapted to compress the media signal for optimized transmission along the communications channel and decompression by the destination device, wherein the compression is done by the CODEC based upon at least one of the following parameters: a previously learned behavior of the CODEC with respect to another reference signal, a previously learned behavior of the CODEC with respect to a prior attempt at compressing or decompressing the same media signal, a comparison of the CODEC's operation with respect to the media signal against a reference algorithm compression of the media signal, a learned constraint of the transmission channel, and a learned constraint of the destination device.

16. The system of Claim 15, wherein the CODEC is adapted to compress the media signal into n different compressed representations of the media signal that are adapted to be transmitted to the destination device at n unique bit rates, respectively.

17. The system of Claim 15, further comprising:

a searchable quality of service (QoS) library associated with the CODEC and that is adapted to be used by at least one of an artificial intelligence system or an operating system in order to determine an appropriate use of the CODEC with respect to the media signal based upon at least one QoS standard stored in the QoS standards library.

18. A system for processing a media signal, comprising:

Ex. 1018-1217

a CODEC operating system that is adapted to interface with a CODEC library and also with an artificial intelligence (AI) system in order to use the AI system to identify an appropriate CODEC from the library and to process the media signal with the CODEC.

19. The system of Claim 18, further comprising:

at least one CODEC associated with the CODEC operating system for use in processing the media signal based at least in part upon an operation by the CODEC operating system.

20. The system of Claim 19, further comprising:

a searchable CODEC library that is associated with the CODEC operating system and that is adapted to store a plurality of separate and unique CODECS, wherein the CODEC operating system is adapted to cooperate with the searchable CODEC library and also with the AI system in order to determine an appropriate CODEC stored in the CODEC library for use in processing the media signal.

21. The system of Claim 20, further comprising:

a searchable network transport standards library associated with the CODEC operating system and that is adapted to be used by at least one of an AI system or the CODEC operating system in order to determine an appropriate use of a CODEC with respect to the media signal based upon at least one network transport standard stored in the network transport standards library.

22. The system of Claim 21, further comprising:

a searchable quality of standard (QoS) library associated with the CODEC operating system and that is adapted to be used by at least one of an AI system or the operating system in order to determine an appropriate use of a CODEC with respect to the media signal based upon at least one QoS standard stored in the QoS standards library.

23. The system of Claim 22, further comprising:

a destination agent that is associated with the CODEC operating system and is adapted to be used by a destination device to communicate with a source also associated with the CODEC operating system to receive a compressed representation of the media signal from the source and to decompress the compressed representation into a decompressed representation of the media signal such that the decompressed representation may be stored or used by the destination device.

24. A system for processing a media signal, comprising:

a destination device that is adapted to communicate with a source and to receive a compressed representation of the media signal from the source, and that includes a CODEC that is adapted to decompress the compressed representation into a decompressed representation based upon at least one of the following parameters: a previously learned behavior of the CODEC with respect to another reference signal, a previously learned behavior of the CODEC with respect to a prior attempt at compressing or decompressing the same media signal, a comparison of the CODEC's operation with respect to the media signal against a reference algorithm compression of the media signal, a learned constraint of the transmission channel, and a learned constraint of the destination device.

25. The system of claim 24, wherein the media signal is adapted to be used according to an original format according to an original set of parameters, and the destination device is adapted to process and display the decompressed representation of the media signal in another format according to another set of parameters that are different than the original set of parameters, such that the destination device is incapable of processing and displaying the media signal in the original format according to the original set of parameters.

26. The system of claim 25, wherein the CODEC is adapted to compress the media signal into n different compressed representations of the media signal that are adapted to be transmitted to the destination device at n unique bit rates, respectively.

27. The system of claim 26, wherein the n different compressed representations are adapted to be transmitted at least at n substantially different transmission rates that are distinguished based upon at least one of the following parameters: different transmission system standards; different destination device standards; and different media type standards.

28. A system for processing a media signal, comprising:

a CODEC that is adapted to perform at least one of compression or decompression on the media signal; and

an artificial intelligence (AI) system that is adapted to operate in conjunction with the CODEC in order to tune the operation of the CODEC with respect to the media signal, wherein the AI system is adapted to learn a behavior of the CODEC with respect to at least one parameter affecting the CODEC's operation, and also to modify the operation of at least one of the AI system

itself or the CODEC based upon the learned behavior in order to increase the ability to provide the CODEC in an appropriate form for optimally compressing or decompressing the media signal.

29. The system of claim 28, wherein the CODEC comprises a discrete cosine transform CODEC.

30. The system of claim 29, wherein the CODEC comprises a fractal CODEC.

31. The system of claim 30, wherein the CODEC comprises a wavelet CODEC.

32. The system of claim 31, wherein the CODEC comprises a lossy CODEC.

33. The system of claim 32, wherein the CODEC comprises a substantially lossless CODEC.

34. The system of claim 33, further comprising:

a searchable CODEC library that is adapted to store a plurality of separate and unique CODECS,

wherein the CODEC is adapted to be stored in the searchable CODEC library, and the AI system is adapted to cooperate with the CODEC library in order to locate and access the CODEC so that the AI system may perform at least one operation with respect to the CODEC.

35. A system for processing a media signal, comprising:

a CODEC library that is adapted to store a plurality of separate and unique CODECS;

an artificial intelligence (AI) system associated with the CODEC library and that is adapted to perform an operation with respect to the CODEC library in order to determine an appropriate CODEC from the plurality of CODECS for use in processing the media signal;

a CODEC operating system associated with the CODEC library and also the AI system and that is adapted to control the operation of the AI system with respect to the CODEC library;

a source system that is adapted to compress the media signal based upon a CODEC from the CODEC library and based upon an operation of the AI system;

a destination system that is adapted to communicate with the source system and to receive a compressed representation of the media signal from the source system and decompress the compressed representation into a decompressed representation for use by the destination system; and

a communications system associated with the source system and also with the destination system so that the source system is adapted to transmit the compressed representation of the media signal to the destination system, and the destination system is adapted to receive the compressed representation and also to appropriately decompress the compressed representation into the decompressed representation for storage or use by the destination system.

Ex. 1018-1220

36. The system of claim 35, further comprising:

a searchable network transport standards library associated with the AI system and also with the CODEC operating system and that is adapted to be used by the AI system in order to determine an appropriate CODEC from the CODEC library for processing the media signal based upon at least one network transport standard stored in the network transport standards library.

37. The system of claim 36, further comprising:

a searchable QoS standards library associated with the AI system and also with the CODEC operating system that is adapted to be used by the AI system in order to determine an appropriate CODEC from the CODEC library for processing the media signal based upon at least one QoS standard stored in the QoS standards library.

38. The system of claim 37, wherein the AI system is adapted to determine an appropriate use of a CODEC from the CODEC library for processing the media signal based upon at least one of the following parameters: a previously learned behavior of the CODEC with respect to another reference signal, a previously learned behavior of the CODEC with respect to a prior attempt at compressing or decompressing the same media signal, a comparison of the CODEC's operation with respect to the media signal against a reference algorithm compression of the media signal, a learned constraint of the transmission channel, and a learned constraint of the destination device.

39. A system for communicating a media signal between a source and n unique destination devices each having a unique ability to process the media signal, wherein n is an integer that is equal to or greater than one, comprising:

a compression system associated with the source and which is adapted to compress the media signal into n uniquely compressed representations of the media signal,

wherein each of the n uniquely compressed representations is adapted to be transmitted from the source to a unique one of the n destination devices, and is adapted to be decompressed by the respective destination device into a unique one of n decompressed representations of the media signal to be played by the respective destination device.

40. The system of claim 39, wherein the compression system further comprises:

a searchable CODEC library that is associated with the compression system and that is adapted to store a plurality of separate and unique CODECS, wherein the compression system is

adapted to compress the media signal into each of the n uniquely compressed representations based at least in part upon at least one CODEC stored by the CODEC library.

41. The system of claim 40, wherein the compression system further comprises:

a searchable network transport standards library associated with the compression system and that is adapted to be used by the compression system in order to determine at least one appropriate CODEC for compressing the media signal into the n uniquely compressed representations based upon at least one network transport standard stored in the network transport standards library.

42. The system of claim 41, wherein the compression system further comprises:

a searchable QoS standards library associated with the compression system and that is adapted to be used by the compression system in order to determine at least one appropriate CODEC from the CODEC library for compressing the media signal into the n uniquely compressed representations based upon at least one QoS standard stored in the QoS standards library.

43. The system of claim 42, further comprising:

a destination agent that is associated with the compression system and is adapted to be used by each of the n unique destination devices to communicate with the source, so that each of the n destination devices may receive a unique one of the n compressed representation of the media signal from the source system and to use an appropriate CODEC at least in part based upon a CODEC used by the compression system to create the respective compressed representation so that the respective destination device may appropriately decompress the respective compressed representation into an appropriate one of the decompressed representations of the media signal that may be stored or used by the respective destination device.

44. The system of claim 43, wherein the compression system is adapted to compress the media signal into each of the n different compressed representations based at least in part upon at least one of the following parameters: a previously learned behavior of at least one CODEC with respect to another reference signal, a previously learned behavior of at least one CODEC with respect to a prior attempt at compressing or decompressing the same media signal, a comparison of at least one CODEC's operation with respect to the media signal against a reference algorithm compression of the media signal, a learned constraint of at least one transmission channel, and a learned constraint of at least one destination device.

45. A system for communicating n uniquely compressed representations of a media signal from at least one source to a common receiver, comprising:

a destination device that is adapted to communicate with the at least one source and to receive each of the n uniquely compressed representations from the at least one source, wherein the destination device is further adapted to decompress each of the n uniquely compressed representations into n decompressed representations of the media signal and to play each of the n decompressed representations.

46. A system for communicating a media signal from a source to a destination device over a transmission channel, comprising:

a compression system that is adapted to receive information about a parameter of the transmission channel and to compress the media signal into a compressed representation of the media signal by using a compression technique that is based at least in part upon the parameter information.

47. A system for communicating n unique media signals between a source and a destination, comprising:

a compression system that is adapted to receive information about a parameter of the media signal and to compress the media signal into a compressed representation of the media signal by using a compression technique that is based at least in part upon the parameter information.

48. A system for compressing a media signal from a source for transmission to a destination device, comprising:

a compression system that is adapted to compress the media signal into a compressed representation of the media signal based upon at least one of a transmission channel constraint, a destination channel constraint, or a comparison of a parameter of the media signal against a prior compression operation of the compression system.

49. A system for decompressing a compressed media signal received from a source at a destination device, wherein the destination device is adapted to receive a compressed representation of the media signal from the source and also a decompression instruction from the source such that the destination device is adapted to decompress the compressed representation into a decompressed representation of the media signal based at least in part on the instruction.

50. A method for constructing a neural network for use in compressing media signals, comprising:

providing a plurality of CODECS; and

pre-learning at least one parameter associated with each CODEC's compression performance with respect to a predetermined type of media signal in a manner……; and

remembering each pre-learned parameter in a manner such that at least a particular one of the CODECs may be chosen in the future to compress a particular new media signal based upon the pre-learned parameter of the particular CODEC.

51. The method of claim 50, further comprising:

forming a CODEC library with the plurality of CODECS;

compressing at least one predetermined type of a test media signal with each CODEC in the CODEC library;

automatically learning at least one parameter related to each CODEC's performance in compressing the at least one predetermined type of test media signal;

automatically remembering each such learned parameter in a manner such that a particular media signal to be compressed.

52. A system for communicating a media signal between a source and a destination at least in part over a wireless transmission channel, wherein a compressed representation is

a compression system that is adapted to compress the media signal into a compressed representation of the media signal that is further adapted to be transmitted to a wireless destination device over a wireless carrier; and

a neural network that uses artificial intelligence to optimize the compression of the media signal by the compression system by learning at least one of the following parameters: a constraint of the wireless destination device, a constraint of the wireless transmission carrier, and a comparison of a parameter of the media signal against a prior compression operation of the compression system.

53. A system for compressing a media signal for transmission from a wireless source to a destination device.

54. A system for compressing a media signal into a compressed representation of the media signal for transmission to a wireless destination device.

55. A system for decompressing a compressed representation of a media signal into a decompressed representation of the media signal at a destination device, wherein the compressed representation is received from a source over a wireless carrier at the destination device.

56. A system for communicating a media signal over an IP network between a source and a destination device, wherein a compressed representation of the media signal is transmitted from the source to the destination device over the IP network.

57. A system for compressing a media signal into a compressed representation of the media signal, such that the compressed representation of the media signal may be transmitted from a source to a destination device over an IP network.

58. A system for decompressing a compressed representation of a media signal into a decompressed representation of the media signal, wherein the compressed representation is received at a destination device from a source over an IP network.

59. A method for communicating a media signal over an IP network, comprising:

using artificial intelligence to determine the most appropriate CODEC from a plurality of available CODECS for compressing and decompressing the media signal, wherein the appropriate CODEC is determined based upon at least one of the following parameters: a previously learned behavior of the CODEC with respect to another reference signal, a previously learned behavior of the CODEC with respect to a prior attempt at compressing or decompressing the same media signal, a comparison of the CODEC's operation with respect to the media signal against a reference algorithm compression of the media signal, a learned constraint of the transmission channel, and a learned constraint of the destination device.

60. The method of Claim 59, wherein the system is adapted to communicate a streaming media signal over the IP network.

61. A system for communicating an audio signal between a source and a destination over a transmission channel, wherein the audio signal is compressed to form a compressed representation of the audio signal that may be transmitted over the transmission channel to the destination device

so that the destination device may decompress the compressed representation into a decompressed representation for storage or use by the destination device.

62. The system of Claim 62, wherein the system is adapted to communicate a streaming audio signal.

63. A system for compressing an audio signal into a compressed representation of the audio signal for transmission from a source to a destination device over a transmission channel.

64. A system for decompressing a compressed representation of an audio signal at a destination device for storage or use by the destination device, wherein the compressed representation is received by the destination device from a source over a transmission channel.

65. A method for communicating a streaming audio signal from a source to a destination device over a transmission channel, comprising:

delivering CD-quality music from the source to the destination device over the transmission channel at a transmission rate that is as low as about 25 Kbps.

66. The method of claim 65, wherein the CD-quality music is delivered by compressing a source audio signal into a compressed representation and delivering the compressed representation over the transmission channel to the destination device at the transmission rate, wherein the compressed representation may be decompressed by the destination device into a decompressed representation that is substantially CD-quality representation of the original audio signal that may be stored or played by the destination device.

67. A system for communicating a video signal between a source and a destination over a transmission channel, wherein the streaming video signal is compressed into a compressed representation that may be transmitted from the source to the destination device over the transmission channel and decompressed into a decompressed representation by the destination device for storage or playing by the destination device.

68. A system for communicating a streaming video signal between a source and a destination, comprising:

a media processing module;

a dynamic player module;

an image processor;

a baseline snapshot module;

a signal classifier;

a plurality of CODECS dependent neural networks processors;

a CODECS library;

a network layer input module;

a network layer output module; and

a timer.

69. The system of Claim 68, further comprising a dynamic client request processing module.

70. The system of Claim 69, further comprising a device and network parameters measurement module.

Ex. 1018-1227

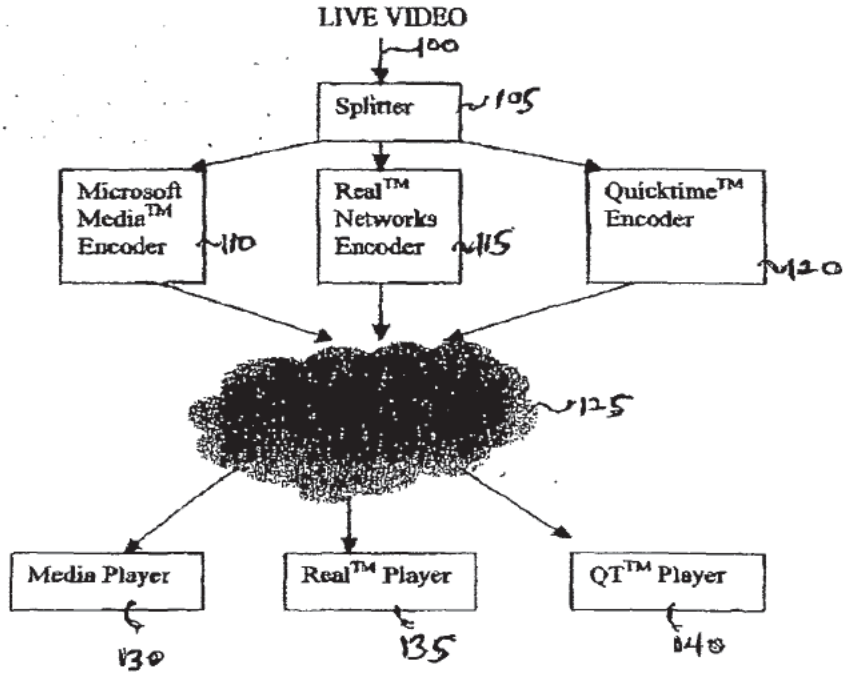


FIGURE 1A
(PRIOR ART)

Ex. 1018-1228

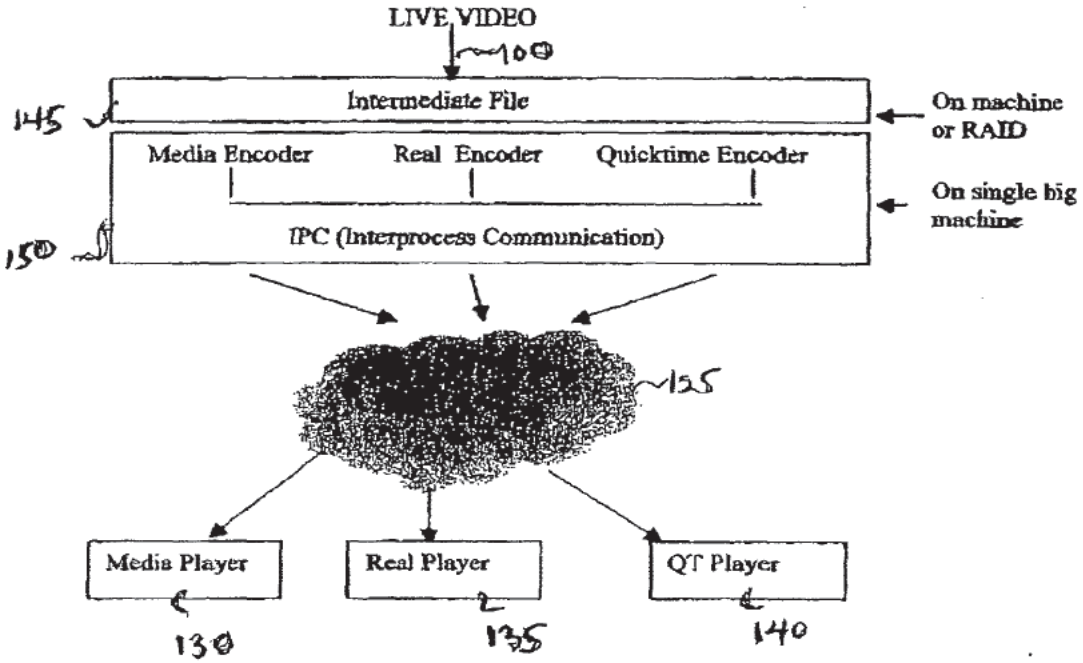


FIGURE 1B
(PRIOR ART)

Ex. 1018-1229

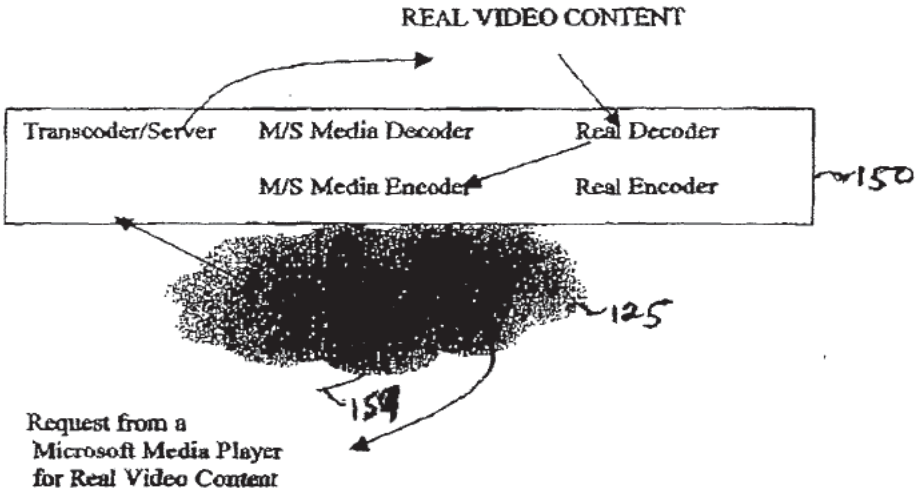


FIG. 1C
(PRIOR ART)

Ex. 1018-1230

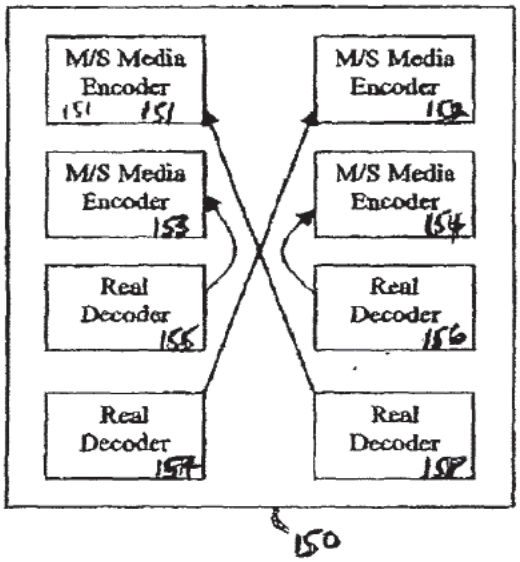


FIGURE 1D
(PRIOR ART)

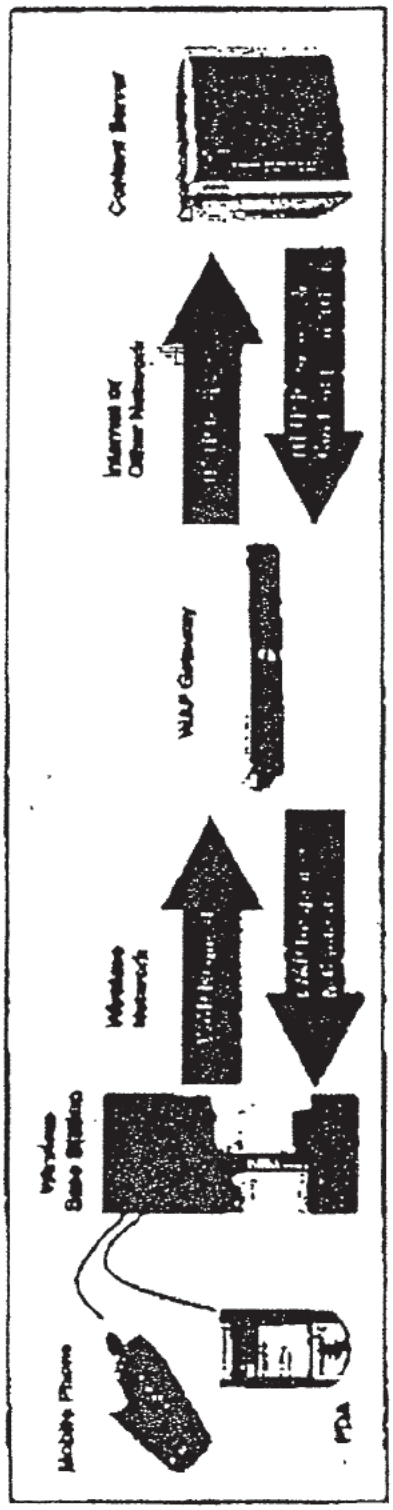


FIGURE 1E
(PRIOR ART)

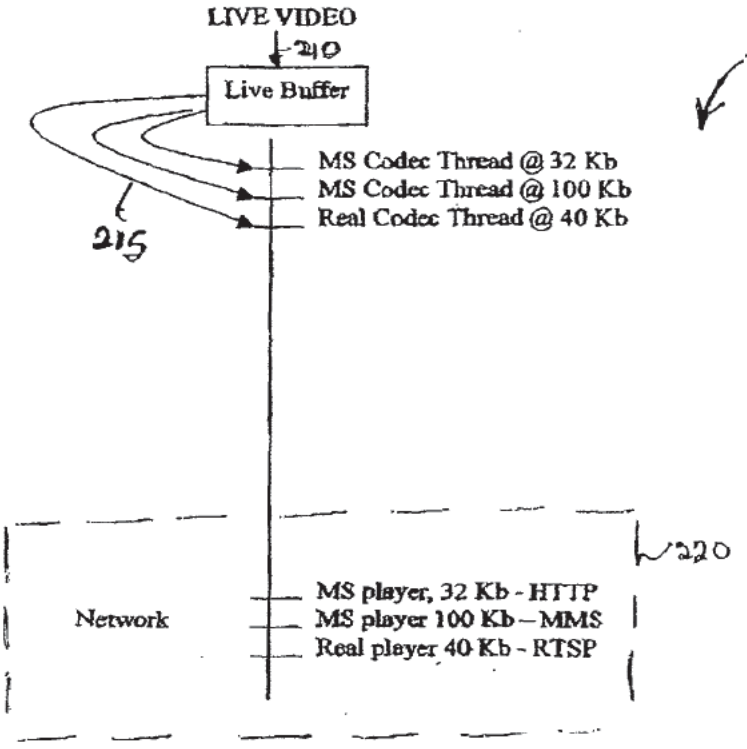


FIGURE 2

Ex. 1018-1233

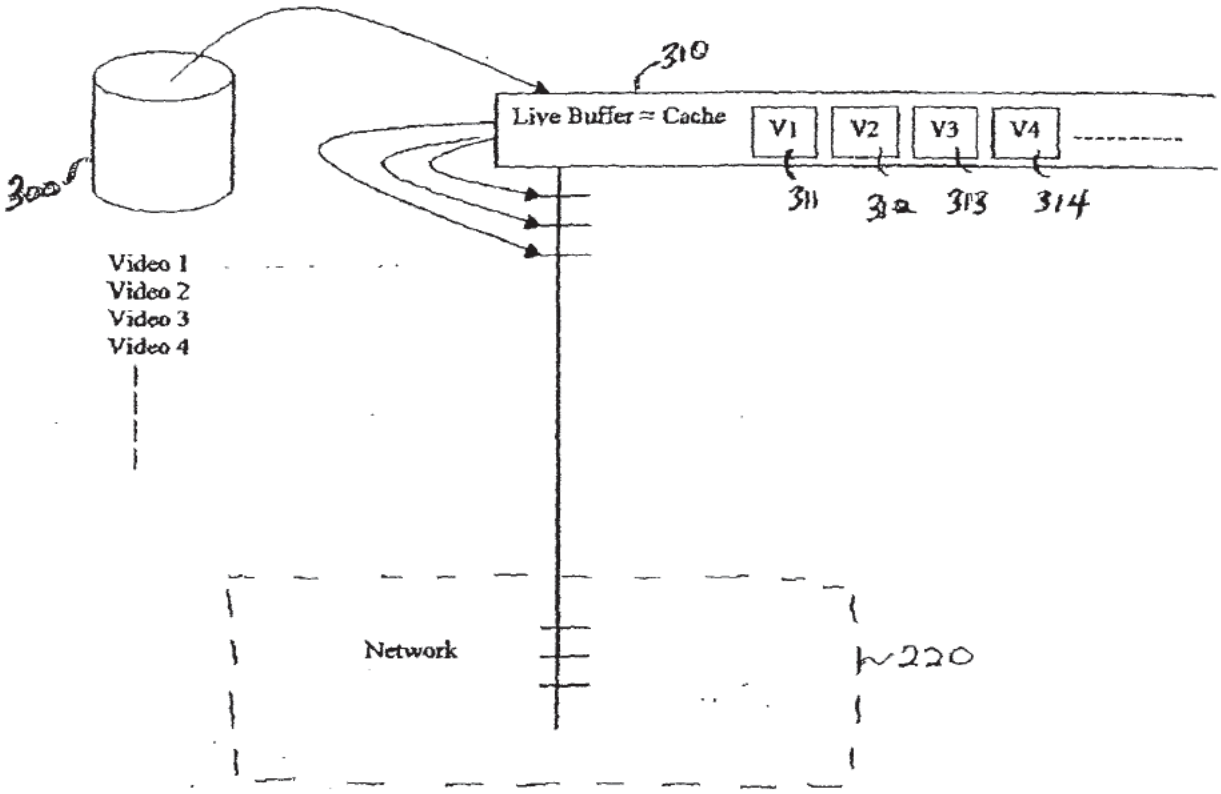
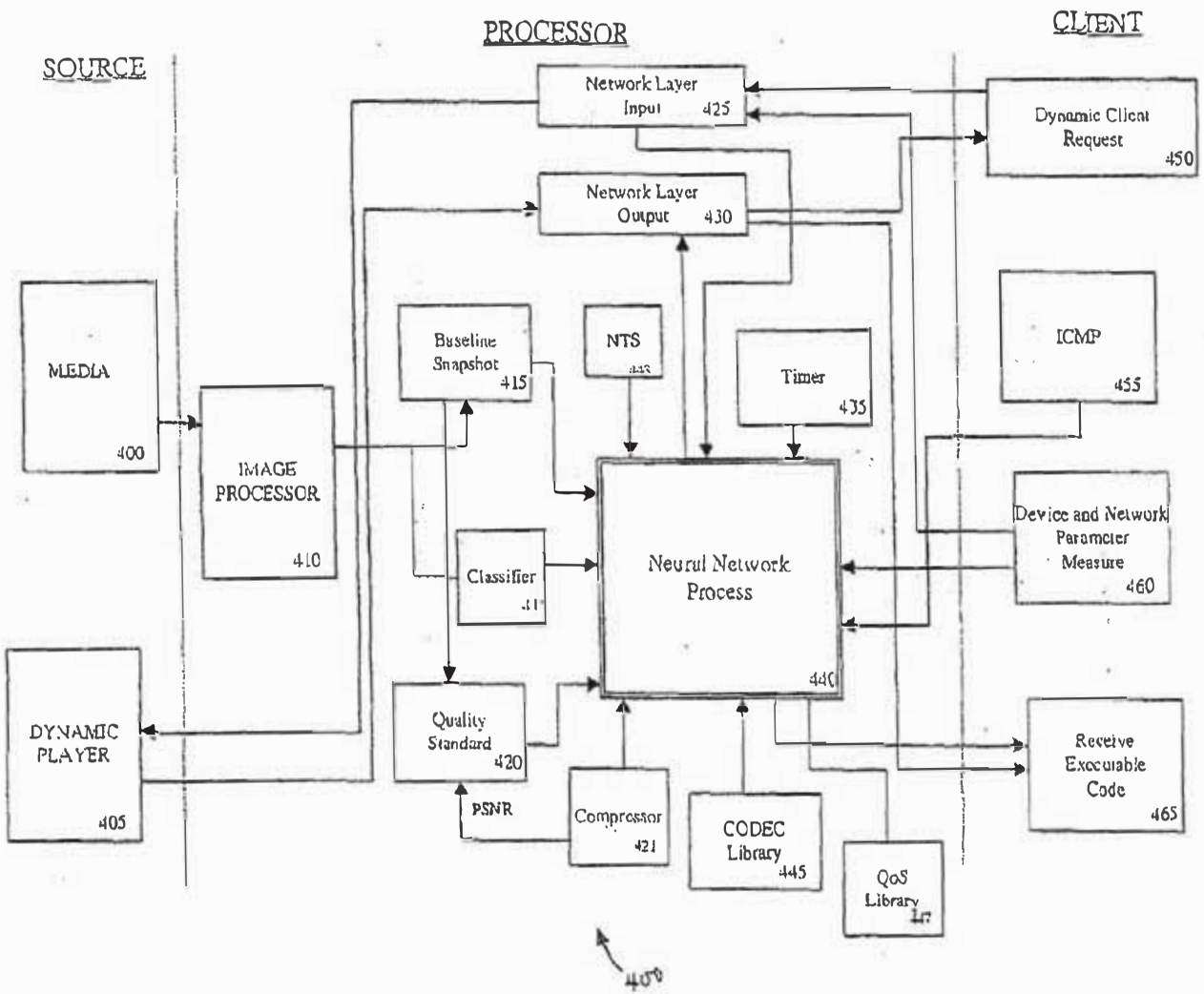

310
Live Buffer = Cache
V1
V2
V3
V4
311
312
313
314
300
Video 1
Video 2
Video 3
Video 4
Network
220

FIGURE 3

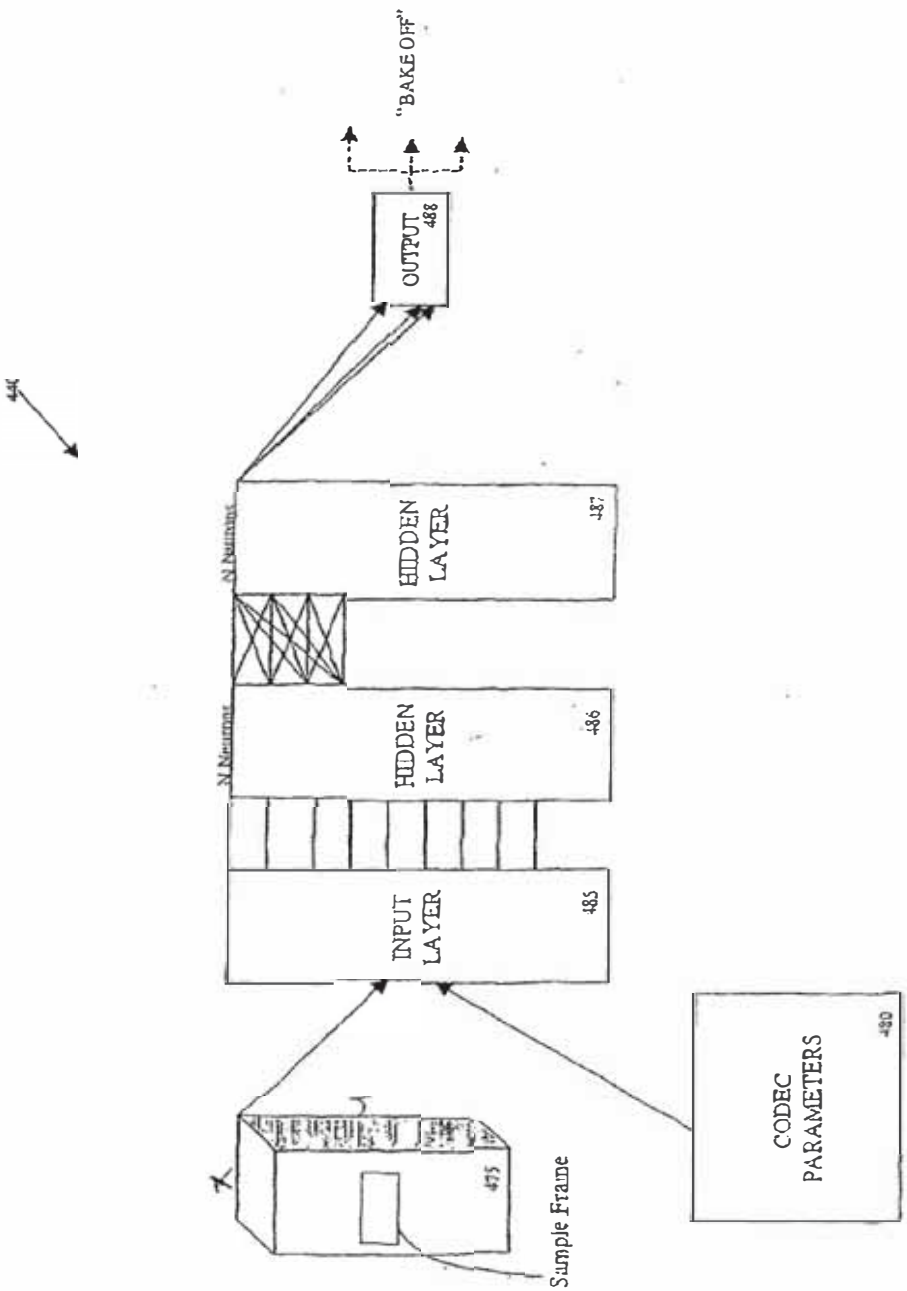
FIGURE 4A



Ex. 1018-1235

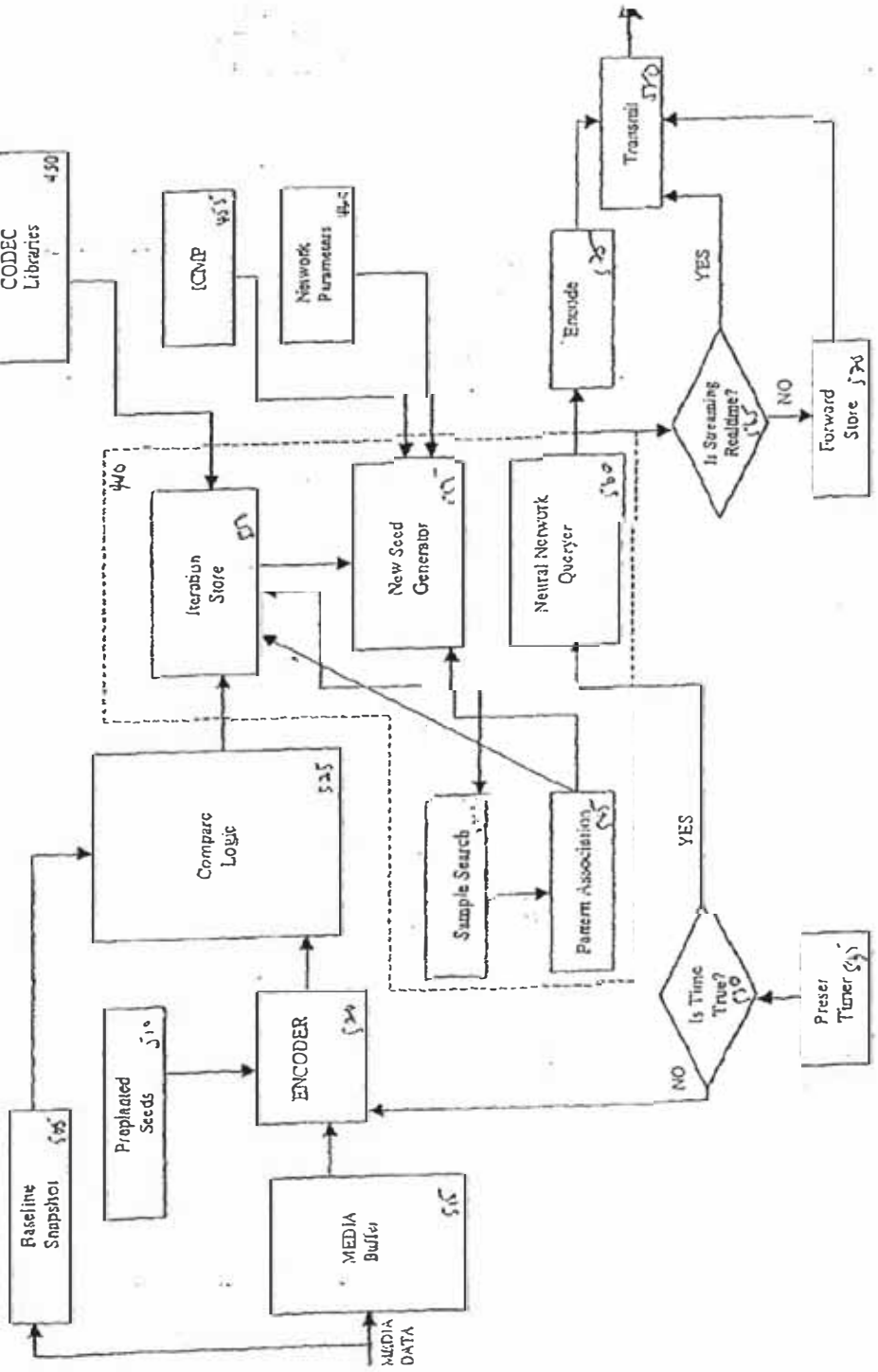
Ex. 1018-1236

FIGURE 4B



Ex. 1018-1237

FIGURE 5



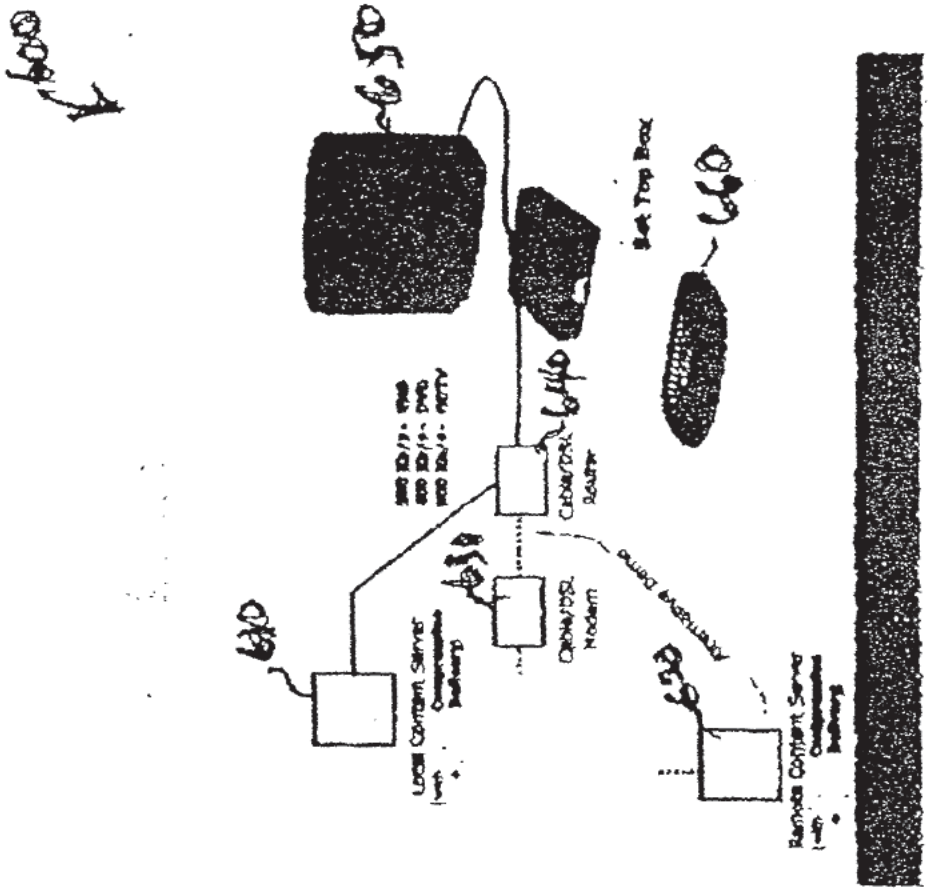

Set Top Box
Cable/DSL Modem
Cable/DSL Router
Local Content Server
Remote Content Server

FIGURE 6

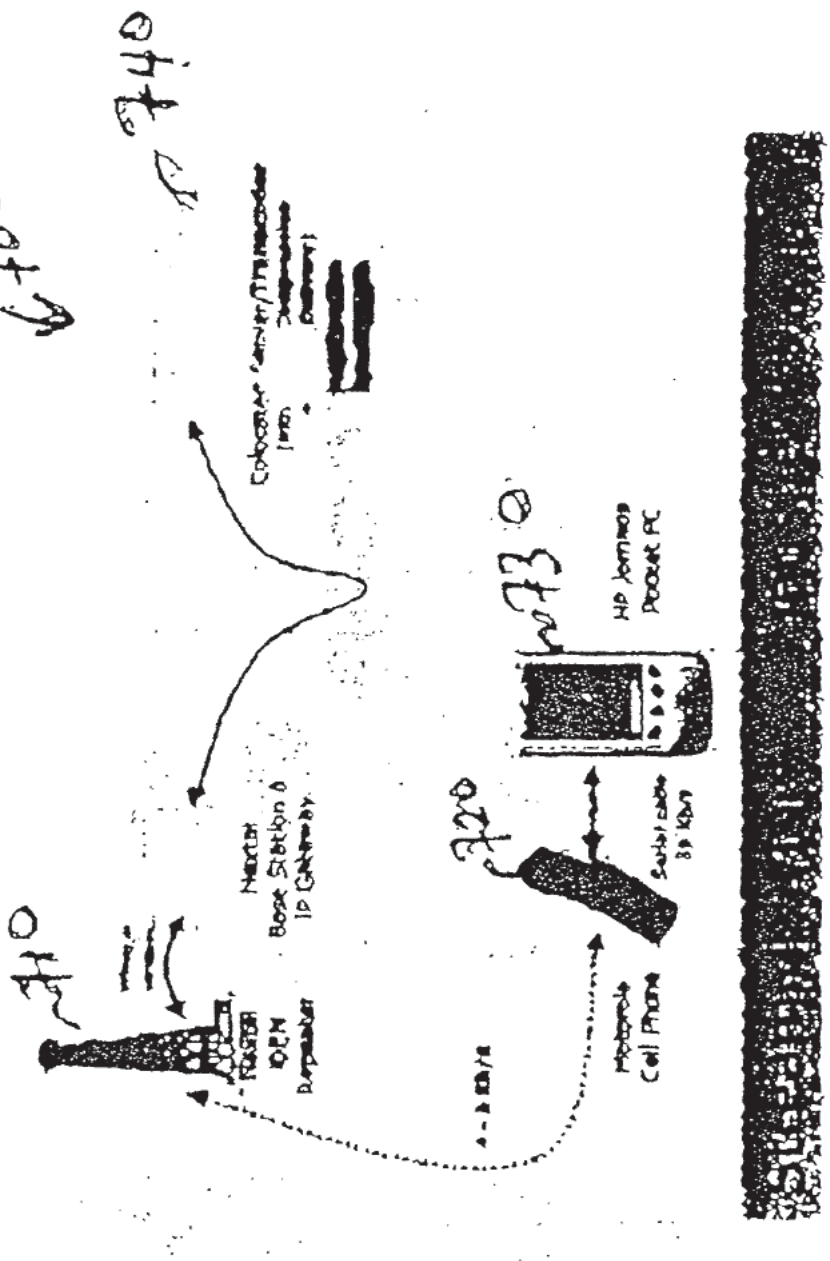


FIGURE 7

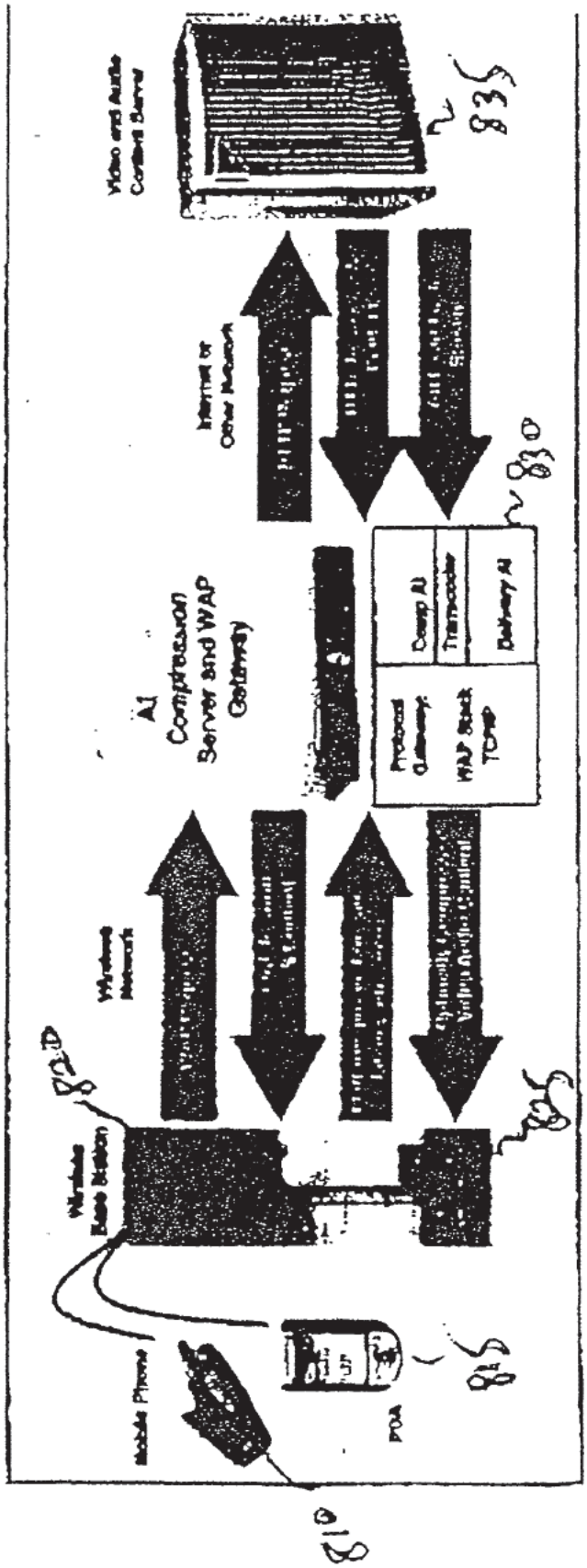
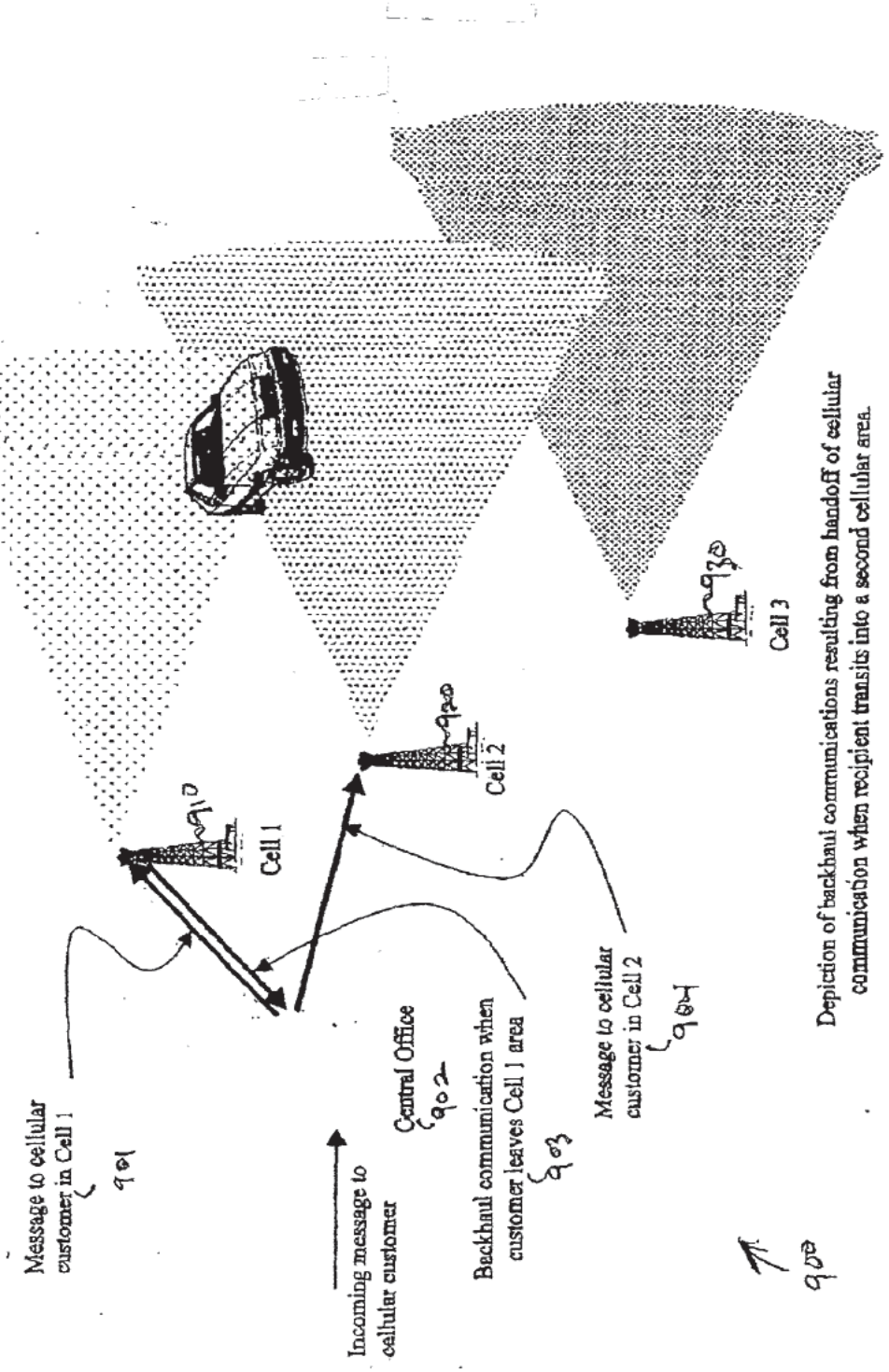
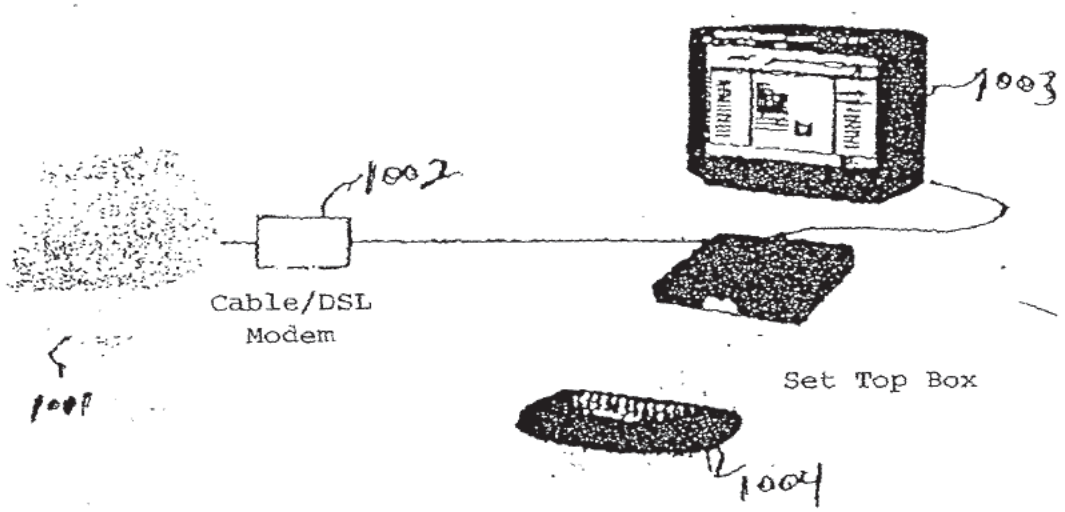

Video and Audio Content Server
835
Internet or Other Network
830
AI Compression Server and WAP Gateway
Wireless Network
820
Wireless Base Station
825
Mobile Phone
PDA
815
810
800

FIGURE 8

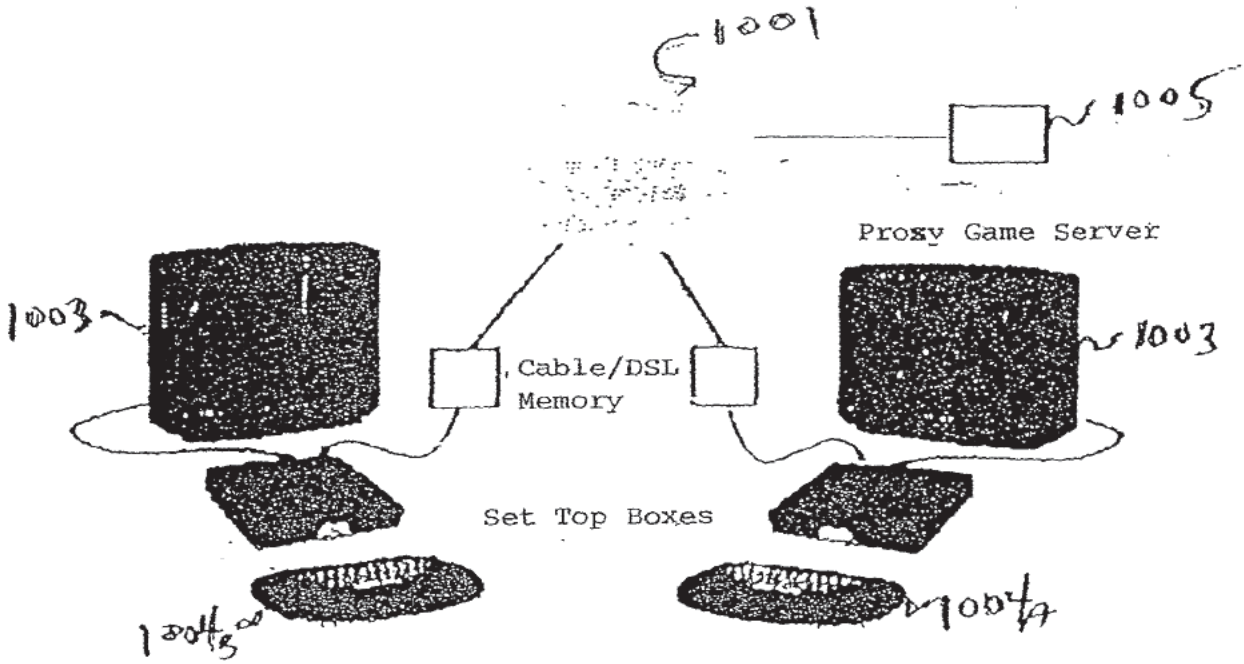
FIG. 9



Ex. 1018-1241



Browsing via the Set Top Box



Interactive Gaming

Fig. 10

Ex. 1018-1242

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US02/30874

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : G06F 15/16
US CL : 709/231,232,217-219,250,220; 704/243

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)
U.S. : 709/231,232,217-219; 345/335,327

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 6,195,692 B1 (HSU) 27 Febuary 2001, column 5 lines 4-19,	1-69
Y	US 6,243,676 B1 (WITTEMAN) 5 June 2001, column 7 line 55-col 8 line 56,	1-69
Y	US 6,115,755 A (KRISHAN) 5 September 2000, column 3 lines 19-45, column 4 lines 54-67	1-69
Y	US 6,085,236 A (LEA) 4 July 2000, column 18 lines 9-20, col 26 line 59-column 27 line 4,	1-69

☐ 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 published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
10 December 2002 (10.12.2002)	15 JAN 2003
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703)305-3230	Authorized officer Mark Powell Telephone No. 703-305-9703

Form PCT/ISA/210 (second sheet) (July 1998)

IDS CERTIFICATION STATEMENTS

I. THIS IDS IS BEING FILED UNDER 37 C.F.R. 1.97(b): (check one box)

☐ Within three months of the filing date of a national application other than a continued prosecution application under 1.53(d).

☐ Within three months of the date of entry of the national stage as set forth in 1.491 in an international application.

☐ Before the mailing date of a first Office Action on the merits. In the event that a first Office Action on the merits has been mailed, please consider this IDS under 37 C.F.R. 1.97(c) and see the statement under 1.97(e) **or** charge deposit account 50-2091 the fee specified by 37 C.F.R. 1.17(p) if such statement is not selected in Section IV.

☐ Before the mailing of a first Office Action after the filing of a request for continued examination under 1.114.

II. THIS IDS IS BEING FILED UNDER 37 C.F.R. 1.97(c):

☒ After the period specified in 37 C.F.R. 1.97(b) and before the mailing date of any of a final action under 1.113, a notice of allowance under 1.311 or an action that otherwise closes prosecution in the application, and accompanied by:

- ☐ a statement specified in 1.97(e) as subsequently provided in this IDS; or
- ■ the fee set forth in 1.17(p).

III. THIS IDS IS BEING FILED UNDER 37 C.F.R. 1.97(d):

☐ After the period specified in 37 C.F.R. 1.97(c) and on or after the mailing date of any of a final action under 1.113, a notice of allowance under 1.311 or an action that otherwise closes prosecution in the application, but before or simultaneous with payment of the issue fee, and accompanied by a statement specified in 1.97(e) as subsequently provided in this IDS **and** the fee set forth in 1.17(p).

IV. Statement under 37 C.F.R. 1.97(e) (check only one box)

The undersigned hereby states that:

☐ Each item of information contained in the IDS was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of this IDS.

☐ No item of information contained in the IDS was cited in a communication from a foreign patent office in a counterpart foreign application, and, to knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the IDS was known to any individual designated in 37 C.F.R. 1.56(c) more than three months prior to the filing of this IDS.

V. Statement under 37 C.F.R. §1.704(d):

☐ The undersigned hereby states that each item of information contained in the information disclosure statement: (i) was first cited in any communication from a patent office in a counterpart foreign or international application or from the Office, and this communication was not received by any individual designated in § 1.56(c) more than thirty days prior to the filing of the information disclosure statement; or (ii) is a communication that was issued by a patent office in a counterpart foreign or international application or by the Office, and this communication was not received by any individual designated in § 1.56(c) more than thirty days prior to the filing of the information disclosure statement..

VI. CONCLUSION

It is respectfully submitted that the claims presently on file patently distinguish the present invention from each of the references cited in this IDS. The references cited in this IDS are being cited only in the interest of candor and without any admission that they constitute statutory prior art or contain matter which anticipates the invention or which would render the same obvious, either singly or in a combination, to a person of ordinary skill in the art. Furthermore the filing of this IDS should not be construed to be an admission that the information cited in this IDS is, or is considered to be, material to patentability as defined in 1.56(b). In addition, the filing of this IDS should not be construed as a representation that a search has been made.

If the Examiner has any questions concerning this IDS, the Examiner is requested to contact the undersigned at the telephone number associated with Customer No. 110240. If it is determined that this IDS has been filed under the wrong rule, the PTO is requested to consider this IDS under the proper rule, with a petition if necessary, and charge the appropriate fee to Deposit Account No. 50-2091.

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
29 April 2004 (29.04.2004)

PCT

(10) International Publication Number
WO 2004/036824 A1

(51) International Patent Classification[7]: H04L 9/06

(21) International Application Number:
PCT/FI2003/000752

(22) International Filing Date: 10 October 2003 (10.10.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
20021820 14 October 2002 (14.10.2002) FI

(71) Applicant *(for all designated States except US)*: **NOKIA CORPORATION** [FI/FI]; Keilalahdentie 4, FIN-02150 Espoo (FI).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **CURCIO, Igor, D.D.** [IT/FI]; Hatanpään valtatie 12 C 57, FIN-33100 Tampere (FI). **LUNDAN, Miikka** [FI/FI]; Männikönkatu 4 A 7, FIN-33820 Tampere (FI).

(74) Agent: **FORSSÉN & SALOMAA OY**; Eerikinkatu 2, FIN-00100 Helsinki (FI).

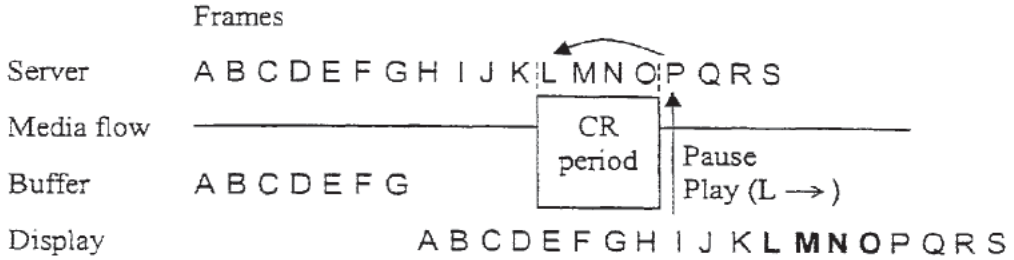
(81) Designated States *(national)*: AE, AG, AL, AM, AT (utility model), AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ (utility model), CZ, DE (utility model), DE, DK (utility model), DK, DM, DZ, EC, EE (utility model), EE, EG, ES, FI (utility model), FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT (utility model), PT, RO, RU, SC, SD, SE, SG, SK (utility model), 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 (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*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: STREAMING MEDIA



(57) Abstract: The invention relates to a method for streaming of media from a streaming server (111) to a mobile client device (101) over an air-interface, wherein, after a cell reselection, the method comprises requesting the streaming server (111) to resend streaming media which the mobile client device (101) was not able to receive due to the cell reselection.

WO 2004/036824 A1

STREAMING MEDIA

FIELD OF THE INVENTION

The invention relates to streaming of media from a streaming server to a mobile client device over an air-interface.

BACKGROUND OF THE INVENTION

A packet switched streaming service PSS is currently being standardized for mobile environment by 3GPP (3rd Generation Partnership Project). An example of a communications system capable of streaming media (streaming video and/or audio) is shown in Figure 1. The system comprises a streaming server 111 which is coupled to an IP-network (Internet Protocol) 104. The IP-network 104 may be, for example, the Internet or a service provider operator's intranet (an intranet network belonging to the operator's domain). The IP-network 104 is coupled to a core network 103 of a mobile communications network via a G_i interface. The mobile communications network also has a radio access network (RAN) 102 coupled to the core network 103. The radio access network 102 provides mobile communications devices 101 with access to the mobile communications network over an air-interface. Said access may be provided either by circuit switched means (circuit switched voice or data call) or packet switched means or both. In the following, GPRS (General Packet Radio Service) is used as an example of a packet switched means to communicate over the air-interface. Concerning GPRS, the generic term radio access network (RAN) is considered to comprise base (transceiver) stations (BS) and base station controllers (BSC).

In streaming media a sequence of 'moving images' (i.e. video) or sound (i.e. audio) or a sequence of 'moving images' together with sound (i.e. multimedia) is sent in a compressed form from the streaming server 111 to a mobile communications device 101 (hereinafter referred to as a client device 101). In contrast to the

technique in which an entire media file has to be arrived at the client before it can be played, the streaming technique enables sending of media (video and/or audio) from the streaming server 111 to the client device 101 in a continuous manner and playing of the media as it arrives at the client.

When media is streamed in cellular mobile communications systems such as the system presented in Figure 1, new problems specific to the mobile environment arise. These problems are mostly due to the different restrictions of mobile systems. One such problem arises in a cell reselection (i.e. handover) situation presented in Figure 2.

In Figure 2, a first base station BS1 of the radio access network 102 serves the area of a first cell 201 whereas a second base station BS2 of the radio access network 102 serves the area of a second cell 202. Before cell reselection, the client device 101 is served by the first base station BS1, i.e. the client device 101 has an active radio link with the first base station BS1. This situation is shown on the left side of the Figure 2. After the cell reselection, the client device 101 is served by the second base station BS2. The radio link with the first base station BS1, shown by a dotted line, is dropped. This situation is shown on the right side of the Figure 2.

The cell reselection (CR) can be divided in three periods in time:

a) Pre-CR period
b) CR period
c) Post-CR period.

During the pre-CR period, received signal quality in the first cell weakens and the client device 101 starts cell reselection signaling. The client device 101 is able to receive streaming media via the first base station BS1 during that period. During the CR period, the actual cell reselection is performed. The client device 101 can

not receive streaming media during that period. During the post-CR period, the client device 101 is able to receive streaming media via the second base station BS2.

The cell reselection may cause a long break in service. For example, a cell reselection (CR period) in a case where GPRS is used as radio bearer may take up to 30-40 seconds. This will affect an ongoing streaming session. For example, part of the packets carrying streaming media may be lost and/or freezing of the played media may occur at the client device 101. If the streaming media comprises a video stream, the freezing means that a still image will appear and stay for a period of time on a client device display before the streaming media is again received and can be played again. If the streaming media comprises an audio stream, the freezing means that no sound is played back (i.e. silence) until the CR period is over.

Accordingly, there is a need to find appropriate means for reducing the effects of cell reselection on an ongoing streaming session.

SUMMARY OF THE INVENTION

According to a first aspect of the invention, there is provided a method for streaming of media from a streaming server to a mobile client device over an air-interface, wherein the method comprises:
requesting the streaming server to send streaming media which the mobile client device is not able to receive due to a cell reselection.

A request requesting the streaming server to send the streaming media may be sent from the mobile client device to the streaming server before or after the mentioned cell reselection.

The mentioned request is not limited to requesting to send only the media which

the mobile client device is (or was) not able to receive due to the cell reselection (during a cell reselection period). In an embodiment, the request may be construed such that what is actually requested is: resending of a portion of a media content not received as well as sending any remaining portion of the media content.

The term media is considered to mean either video or audio or another media, such as still image, or any combination thereof, i.e multimedia.

Preferably, a resending request is sent in response to the cell reselection. The re-sending request is preferably an application layer request sent between a mobile client device application and a streaming server application.

In a preferred embodiment, the streaming server is requested to send at an increased speed for a determined period of time in order to increase at the client device a degree of fullness of a temporary store, such as a buffer, in which the streaming media is temporarily stored before playback. Preferably, this is done by by requesting the streaming server to switch from sending a higher bit rate media stream to sending a lower bit rate media stream at an increased speed. The higher and lower bit rate streams are preferably provided by a multi-rate codec.

According to a second aspect of the invention, there is provided a mobile client device for receiving streaming media from a streaming server over an air-interface, the mobile client device comprising:
means for requesting the streaming server to send streaming media which the mobile client device is not able to receive due to a cell reselection.

Preferably, the air-interface couples said mobile client device to a mobile communications network. The mobile client device preferably comprises a cellular mobile phone.

According to a third aspect of the invention, there is provided a streaming server

for sending streaming media to a mobile client device over an air-interface, the streaming server comprising:
means for receiving a request requesting the streaming server to send streaming media which the mobile client device is not able to receive due to a cell reselection; and
means for acting upon the received request.

According to a fourth aspect of the invention, there is provided a system comprising a streaming server and a mobile client device, for streaming of media from the streaming server to the mobile client device over an air-interface, the system comprising, at the mobile client device:
means for requesting the streaming server to send streaming media which the mobile client device is not able to receive due to a cell reselection, the system further comprising, at the streaming server:
means for receiving the request; and
means for acting upon the received request.

According to a fifth aspect of the invention, there is provided a computer program executable in a mobile client device, the computer program comprising:
program code for causing the mobile client device to request the streaming server to send streaming media which the mobile client device is not able to receive due to a cell reselection.

According to a sixth aspect of the invention, there is provided a computer program executable in a streaming server, the computer program comprising:
program code for causing the streaming server to receive a request requesting the streaming server to send streaming media which the mobile client device is not able to receive due to a cell reselection; and
program code for acting upon the received request.

Dependent claims contain preferable embodiments of the invention. The subject

matter contained in dependent claims relating to a particular aspect of the invention is also applicable to other aspects of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Embodiments of the invention will now be described by way of example with reference to the accompanying drawings in which:

Figure 1 shows a communications system capable of streaming media;

Figure 2 shows a cell reselection situation in the communications system of Figure 1;

Figure 3 shows a client device buffer in a situation in which the client device buffer is shorter than the time that cell reselection takes;

Figure 4 shows a client device buffer in a situation in which the client device buffer is longer than the time that cell reselection takes;

Figure 5 shows another illustration of the situation which is behind Figure 4;

Figure 6 shows preferred action in a preferred embodiment of the invention;

Figure 7 illustrates a preferred embodiment of the invention;

Figure 8 illustrates a preferred embodiment of the invention;

Figure 9 shows a client device according to a preferred embodiment of the invention;

Figure 10 shows a streaming server according to a preferred embodiment of the

invention; and

Figure 11 illustrates three ways of sending messages according to preferred embodiments of the invention.

DETAILED DESCRIPTION

The system shown in Figure 1 can also be used in the preferred embodiments of the invention (the same applies to Figure 2). Accordingly, the system comprises a streaming server 111 which is coupled to an IP-network 104. The IP-network 104 may be, for example, the Internet or a service provider operator's intranet. The IP-network 104 is coupled to a core network 103 of a mobile communications network via a G_i interface. The mobile communications network also has a radio access network (RAN) 102 coupled to the core network 103. The radio access network 102 provides mobile communications devices 101 with access to the mobile communications network over an air-interface. Said access may be provided either by circuit switched means (circuit switched voice or data call) or packet switched means or both. In the following, GPRS (General Packet Radio Service) is used as an example of a packet switched means to communicate over the air-interface.

In order to better understand the preferred embodiments of the invention, let us consider the following example.

In this example, a mobile communications device 101 (again referred to as a client device 101) has established a streaming session with the streaming server 111 when being served by the first base station BS1. In this example and in the following description, mostly a video streaming session is considered. However, a corresponding consideration applies to an audio streaming session as well as to a multimedia (e.g. video with audio) streaming session.

RTSP (Real Time Streaming Protocol) protocol is used in the streaming session

setup. Once the session has been established, the streaming itself may be performed (i.e. media flow may be sent) according to RTP (Real time Transport Protocol) or another protocol. However, if it is desired to make a change in the established session, this will again be done by using RTSP.

At the beginning of the established streaming session, the received streaming media (media flow), transmitted in packets (or frames), is buffered (i.e. temporarily stored) in a buffer at the client device (hereinafter referred to as a client device buffer). When the client device buffer is full, playing (playback) of the streaming media is started at one end of the buffer while the buffer is continuously filled at the other end with received streaming media. Therefore, the buffer should remain nearly full during normal operation of the system.

The client device now undergoes a cell reselection from the first cell 201 to the second cell 202 (Fig. 2). It is to be noted that, although the radio access network (RAN) 102 is illustrated in Figure 2 by a single cloud only, it may comprise different radio access networks. It may comprise a GPRS (GPRS-only) RAN, a 3G (third generation) RAN or a combination thereof. The first cell 201 (served by the base station BS1) and the second cell 202 (served by the base station BS2) may therefore be, for example, cells of the GPRS RAN or 3G RAN. Accordingly, the cell reselection (handover) can occur, for example, between base stations belonging to GPRS RAN, between base stations belonging to 3G RAN or between a base station belonging to GPRS RAN and a base station belonging to 3G RAN.

The client device 101 can not receive streaming media during the CR period. Therefore, when the CR period begins the client device buffer starts emptying because the buffer content is further played while the continuous filling is stopped. Depending on the size of the client device buffer, two different cases can be identified:

i) The client device buffer (in time) is shorter than the time that cell reselection

takes, i.e. the time to empty a full client device buffer is shorter than the CR period,

ii) The client device buffer (in time) is longer than the time that cell reselection takes, i.e. the time to empty a full client device buffer is longer than the CR period.

In the first case (Figure 3), the buffer will become totally empty during the CR period. Since the buffer becomes empty, the client device 101 has no media to play anymore. Therefore, the client device 101 starts rebuffering. Rebuffering can be started during the post-CR period. In order to get the rebuffering started the client device 101 can use an RTSP PAUSE/PLAY method. In this method, the client device 101 first sends an RTSP PAUSE message to the streaming server 111. The PAUSE message will cause the server 111 to pause sending of the media flow. Subsequently, the client device 101 sends an RTSP PLAY message which requests the streaming server 111 to start resending the streaming media from the last received frame onwards. Rebuffering starts when the client again starts receiving the streaming media. When the client device buffer is full, playing of the streaming media is restarted from the last received frame onwards. In the meantime, the last received frame is shown as a still image on a client device display. However, in spite of the drawback of showing a still image, no packets are effectively lost since the media flow which is not received at the client device during the CR period is resent from the server 111 to the client device 101 after the CR period.

In the second case (Figure 4), the whole content of the buffer has not been played during the CR period. Therefore, the buffer only becomes partially empty during the CR period, i.e. there is still streaming media (DATA) in the buffer when the CR period ends. During the post-CR period (after the CR period) playing of the streaming media stored in the buffer before the CR period (i.e. DATA*) is further continued at one end of the buffer while the other end is filled with the streaming

media received after the CR period (i.e. DATA**). If nothing is done, DATA* will eventually end and playing is continued with the first frame of data DATA**. No still image is shown. However, the first frame of DATA** is not the frame that should be played right after the last frame of DATA*. Since the client device 101 could not receive streaming media during the CR period it has not even received the frame that should be showed after the last frame of DATA*. There exists a temporal gap in the buffered media. As a result, the client device 101 will skip showing the sequence of streaming media (video) which could not be received due to the cell reselection Concerning an audio stream, the same effect occurs, i.e. the client device 101 will skip playing sound which could not be received.

The second case is further illustrated with the aid of Figure 5. It is to be noted that Figure 5 only shows a very simplified case. The time is considered to flow from left to right in Figure 5. The streaming media sequence that the server sends comprises video frames A-S. The frames are received at the client device and buffered before playing. The playing of the media starting from frame A starts when the buffer is full (i.e. when the frames A-G have been stored in the buffer), the frames (moving video picture) are subsequently shown on a client device display. The buffering of the frames following frame G is not shown in Figure 5 for clarity reasons.

The media flow from the server to the client device is not possible during the CR period. Therefore, the client device does not receive the frames L-O. Otherwise, the media flow is perfectly received at the client device. Now that the the CR period is shorter than the buffer size, the showing of the frame (frame K) last received before the CR period will be followed by the showing of the frame (frame P) first received after the CR period. Video frames L-O are not shown at all. Concerning an audio stream, the same effect occurs, i.e. the user hears a pause and a temporal gap in the sound.

There exists two type of frames: intra frames and inter frames. The intra frames

contain all necessary information of an image whereas inter-frames only contain changes or predicted changes compared to a previous image. The previous image may be an intra frame or an inter frame. Therefore, if the frame P is an intra frame the shown picture only warps ahead in time from frame K to frame P. In this case the frames L-O are not simply shown to the user. But if the frame P is an inter frame severe distortion in the shown (moving) picture is likely to occur before the next intra frame is received and played.

A preferred embodiment of the invention as shown in Figure 6 concentrates on the above mentioned problem relating to the second case. In this embodiment, the buffer is longer than the duration of the CR period. Therefore, the buffer does not become totally empty during the CR period but only partially empty. It is to be noted that Figure 6 only shows a very simplified case. The time is considered to flow from left to right in Figure 6. The streaming media sequence that the server sends comprises frames A-S. The frames are received at the client device and buffered before playing. The playing of the media starting from frame A starts when the buffer is full, the frames (moving video picture) are subsequently shown on a client device display. The buffering of the frames following frame G is not shown in Figure 6 for clarity reasons. The media flow from the server to the client device is not possible during the CR period. Therefore, the client device does not receive the frames L-O. Otherwise, the media flow is perfectly received at the client device.

When the CR period is over the client device knows exactly what is the last frame that it has received before the CR period (i.e. during the pre-CR period). In this embodiment it is frame K. The client device then requests, right after the CR period has been ended, the streaming server to start resending the streaming media from the last received frame onwards. Upon receiving the request the streaming server starts resending the media flow. The first frame to be sent is in this embodiment frame L. Any frames sent and received after the CR period but before the resending has started (possibly some frames starting from frame P) are

ignored by the client device. When the streaming media (of frames A-K) stored in the buffer before the CR period ends, then the playing is continued with the first frame (frame L) of the resent streaming media received after the CR period. Therefore, no discontinuity in the shown video picture should appear to the user. In Figure 6, the resent video frames L-O which would have otherwise failed to have been received at the client device are shown in bold font. Correspondingly, concerning an audio stream, the streaming server is requested to start resending the audio stream at a point at which reception has been stopped in the client device. No discontinuity in playback should appear.

In practice, the resending request can be performed by the RTSP PAUSE/PLAY method. In this method, the client device first sends, after the CR period has been ended, an RTSP PAUSE message to the streaming server. The PAUSE message will cause the server to pause sending of the media flow. However, the playing of the streaming media received at the client device is not paused unless the buffer gets empty (which should not be the case). Subsequently, the client device sends an RTSP PLAY message to the streaming server 111. The PLAY message contains information on the starting point of the resending. At the end of the CR period the client device knows the time of the last received frame. Based on that, the client device determines the starting point before sending the PLAY message. The PLAY message causes the server to start resending.

An example of the PAUSE message is as follows:

```
PAUSE rtsp://example.com/foo RTSP/1.0
CSeq: 6
Session: 354832
```

The PAUSE message informs the server that a change is coming. An example of the subsequently sent PLAY message is as follows:

PLAY rtsp://example.com/foo RTSP/1.0
CSeq: 7
Session: 354832
Range: npt=28.00-

The message field 'Range' tells the starting point of resending. In this example, the starting point is 28 seconds from the start of the streaming media sequence. In relation to the embodiment shown in Figure 6, this point would be exactly the time of frame L in the streaming media sequence.

In the embodiment just described, the frames (packets) that were lost during the CR period will be resent from the server to the client devide. Also, since the buffer size is longer than the duration of the CR period, no interruption in the shown video picture should appear and the user experience is maximized.

However, as the case is such that during normal streaming the degree of fullness of the buffer does not change, since the buffer is emptied (played) at the same rate as it is filled, this may cause an additional problem. In the embodiment just described, after the cell reselection, the client device buffer stays emptier than before, for the reasons as described in the foregoing. Therefore if, for example, a new cell reselection is to be performed in a near future, the emptier buffer may cause the same drawback as discussed in connection with the first case above (i.e. the case in which the client device buffer was shorter than the time that cell reselection takes).

A preferred embodiment of the invention concentrates on the above identified problem. In this embodiment, to ensure smooth behaviour of the shown video picture (correspondingly: played sound), the client device buffer is filled after the CR period, for a period of time, at a higher rate than it is emptied (played). This period may be called a filling period. When the filling period is over the buffer is full again and normal streaming in which the buffer is filled at the same rate as

emptied is resumed.

It is to be noted that normally raising the degree of fullness of the buffer would require the playing to be paused. In the present embodiment, playing is not paused but the degree of fullness of the buffer can still be raised due to the clever buffer management in which the buffer is filled, during playing, at a rate higher than the playing rate.

To raise the degree of fullness of the buffer without pausing the playback, the client device, in this embodiment, requests the server to switch to sending a lower bit rate streaming media sequence, but to use in the actual sending the same transmission bit rate (hereinafter referred to as an original transmission bit rate) as before. In order to reach the original transmission bit rate the client device requests the server to speed up the transmission of the lower bit rate sequence by a speeding factor. Speeding up the transmission bit rate causes more data to be written in the buffer than read out from the buffer. Therefore, the degree of fullness of the buffer increases as desired.

In other words, the server is asked to switch from sending the original sequence encoded at a first bit rate to sending a new corresponding sequence encoded at a second bit rate lower than the first bit rate and to increase the transmission bit rate of the new sequence in order to reach the original transmission bit rate (bandwidth). It is to be noted that the bit rate at which a media stream is encoded (and will be decoded) is a different concept than the transmission bit rate. The bit rate at which the media stream has been encoded has an effect on the picture quality. If the media stream has been encoded at a higher bit rate, this means that more bits have been used in encoding compared to encoding at a lower bit rate. This typically results in a better picture quality. The transmission bit rate, on the other hand, is the bit rate at which the media stream is actually sent and it depends on the available bandwidth.

The request to switch to sending a lower bit rate sequence and to speed up the transmission bit rate can be delivered by using the RTSP PAUSE/PLAY method. When the client device desires to fill the client device buffer during the post-CR period, it first sends a PAUSE message, corresponding to the one presented already in the preceding description, to the server to indicate that a change is coming. Subsequently it generates a PLAY message and sends it to the server.

An example of such a PLAY message is as follows:

PLAY rtsp://example.com/foo RTSP/1.0
CSeq: 7
Session: 354832
Range: npt=28.00-40.00
Bandwidth: 20000
Speed: 1.5

This message contains two optional message fields, namely 'Bandwidth' and 'Speed' to be understood at the client device and streaming server. These fields are already specified as optional fields by the IETF (Internet Engineering Task Force) in the standard RFC 2326 (Real Time Streaming Protocol).

The message field 'Bandwidth' tells the server to change into sending a lower bit rate sequence (here: a sequence whose bit rate is 20 kbps) and the message field 'Speed' tells the server to speed up the sending by a speeding factor (here: 1.5). The message field 'Range' tells the starting point and stopping point in the streaming media sequence in time units (calculated from the beginning of the streaming media sequence(s)).

The PLAY message shown would suit for an example in which the server initially sends a 30 kbps bit rate sequence at an original transmission bit rate of 30 kbps and in which the server, in order to fill the buffer, is desired to switch into sending

a sequence whose bit rate is 20 kbps and to speed up the sending at a speeding factor 1.5 to reach the original transmission bit rate of 30 kbps during a 12 seconds filling period (i.e. between the time instants of 28 and 40 seconds of the media sequence).

If the client knows the possible bit rate options (they are typically communicated from the server to the client during session setup, for example, with the aid of a 200 OK response to an RTSP DESCRIBE message) it can calculate the speeding factor needed to reach the original bitrate using the following formula:

$$speeding\ factor = \frac{original\ transmission\ bit\ rate}{new\ transmission\ bit\ rate}$$

When the buffer is full, the client device sends another RTSP PAUSE and PLAY message pair to the server. An example of the PLAY message is as follows:

PLAY rtsp://example.com/foo RTSP/1.0
CSeq: 67
Session: 354832
Range: npt=40.00-
Bandwidth: 30000
Speed: 1.0

This exemplary message requests the server to start sending the (original) 30 kbps bit rate media sequence (Bandwidth: 30000) at the original transmission bit rate 30 kbps (Speed: 1.0) starting from the time instant 40 seconds of the media sequence (Range: npt=40.00-).

The client device can calculate the length of the buffer filling period using the following formulas:

$$filling\ period = \frac{LowSeqTime}{speeding\ factor}$$

wherein

$$LowSeqTime = (BufferSize - BufferData) \cdot \frac{original\ transmission\ bit\ rate}{original\ transmission\ bit\ rate - new\ transmission\ bit\ rate}$$

In these formulas, *LowSeqTime* indicates the duration of the lower bit rate sequence playing time in the client device, *BufferSize* indicates the size of the buffer in seconds and *BufferData* indicates the data left in the buffer in seconds.

The filling of the buffer is illustrated in Figure 7. It is to be noted that Figure 7 only shows a very simplified case. All delays, for example, have been ignored. The time is considered to flow from left to right in Figure 7. It can be seen that the buffer starts emptying when the CR period starts. Upon receiving the first PLAY message the server switches to sending a media sequence having a lower bit rate (here: 20 kbps) and speeds up the sending by a speeding factor (here again: 1.5). Since the speeding factor is 1.5, the buffer filling takes twice the time that the emptying period lasted. For example, if the emptying period lasted for 25 seconds, the filling period should be 50 seconds with this speeding factor. When the buffer is full the second PLAY message is sent and the server switches to sending the original media sequence having a higher bit rate (here: 30 kbps) and continues sending at the original transmission bit rate.

In the foregoing, it has been described that, for example, a video stream typically comprises both intra and inter frames, wherein the intra frames are 'independent frames' containing all necessary information of an image whereas inter-frames only contain changes or predicted changes compared to a previous image. In the following preferred embodiment of the invention, the timing of the switch from the lower bit rate sequence to the original bit rate sequence is more closely de-

scribed with this respect.

In this embodiment the purpose is to time, whenever possible, the switch from the lower bit rate sequence to the original bit rate sequence to take place at the point of an intra frame of the original bit rate sequence (or, more generally, at the point of an intra frame of the sequence to which the sending is switched). In this way, prediction errors and/or jumps in media (frame) flow can be avoided. Figure 8 illustrates this embodiment. It is to be noted that Figure 8 only shows a simplified case. The time is considered to flow from left to right in Figure 8. In this embodiment, the first switch from the original bit rate sequence to the lower bit rate sequence is effected at point 28 seconds with the aid of a first PAUSE/PLAY message pair.

As illustrated in Figure 8, the position of intra frames in the original bit rate sequence may differ from the ones of the lower bit rate sequence. However, based on:

- the time required to totally fill the buffer, i.e. '*filling period*',
- the time distance of two adjacent intra frames in the original sequence, i.e. an intra frame rate, '$IFrameTime_{Original}$', and
- the time at which the first switch from the original bit rate sequence to the lower bit rate sequence is effected, i.e. '*SwitchTime*',

it can be calculated a suitable point in time for swiching back to the original bit rate sequence. The parameter '$IFrameTime_{Original}$' can be calculated in the client device from the original bit rate sequence before switching to the lower bit rate sequence. The calculation of the suitable point in time for swiching back to the original bit rate sequence can be performed using the following formula:

$$SeqChangeTime = IFrameTime_{Original} * \left\lfloor \frac{SwitchTime + filling\ period}{IFrameTime_{Original}} \right\rfloor$$

In this formula, '*SeqChangeTime*' indicates the point in time for swiching back to the original bit rate sequence and the brackets indicate a *floor*-function (or *trunc*-function) which truncates the fractional part of the value computed in the brackets. For example, if the time required to totally fill the buffer, i.e. '*filling period*' is 16 seconds, the time at which the first switch is effected, i.e. '*SwitchTime*' is 28 seconds and the time difference between two adjacent intra frames in the original sequence, i.e. '$IFrameTime_{Original}$' is 5 seconds, then the point in time for swiching back to the original bit rate sequence is 40 seconds (*SeqChangeTime*=5*floor((28+16)/5) = 40 s). Accordingly, the second PAUSE/PLAY message pair is sent at point 40 seconds, as indicated in Figure 8. In this embodiment, the starting point and stopping point to be placed in the message field 'Range' of the PLAY message are 28 seconds and 40 seconds, respectively.

It is to be noted that, due to the intra frame rate, it might not in all cases be possible to totally fill the buffer. For example, in the example just described, it would have taken 4 seconds more to totally fill the buffer. However, the formula gives the closest suitable intra frame to the buffer filling period.

In another embodiment, it may well be that, at the point of the last received frame before the CR period, there is not an intra frame in the lower bit rate media sequence. In this embodiment, the starting point of the resending request is adjusted by a necessary amount of frames (or time) backwards such that there will be an intra frame in the lower bit rate sequence at the starting point. In this case, a set of last received frames before the CR period, which belong to a time period subsequent to the starting point, are ignored at the client device in order to guarantee a continuous playback.

In another embodiment of the invention, the switches between two different bit rate sequences are performed without the timing method(s) just described. In this embodiment, the first switching point (from higher to lower bit rate sequence) is directly determined by the last received frame and the second switching point (from lower to higher bit rate sequence) is directly determined by the the time required to totally fill the buffer (*'filling period'*). The switching points can therefore end up being at the point of an intra or inter frame. Should a switching point end up being at the point of an inter frame (for example a P-frame), a small prediction error may occur in the played media, however, in this embodiment, the buffer can be totally filled.

Another way to perform switching is to use so called "switch frames". These are frames which contain "difference" information between corresponding frames in different bit rate sequences. In this embodiment, a bridge and switch between the two sequences is performed with the aid of these frames using the difference information.

Figure 9 shows a client device 101 according to a preferred embodiment of the invention. The client device may be a mobile station of a cellular radio telephone network. The client device 101 comprises a processing unit MCU, a radio frequency part RF, and a user interface UI. The radio frequency part RF and the user interface UI are coupled to the processing unit MCU. The user interface UI typically comprises a display, one or more speaker and a keyboard (not shown) with the aid of which a user can use the device 101.

The processing unit MCU comprises a processor (not shown), a memory 210 and computer software. The software has been stored in the memory 210. The aforementioned client device buffer 240 is also comprised in the memory 210. The processor controls, in accordance with the software, the operation of the client device 101, such as receiving of the streaming media sent from the server 111 and sending of requests to the server 111 via the radio frequency part RF, reading and

writing of received streaming media (video and/or audio) in the buffer 240, and presenting of the received streaming video on the display and audio on the one or more speaker of the user interface UI. A suitable size of the buffer (in time) can be, for example, 1.5 or 2 times the maximum (or average) CR period time.

The software comprises a streaming client software application 220 (hereinafter referred to as client software 220), a protocol stack 230 for implementing the necessary protocol layers such as an RTP layer, an RTSP layer, an SDP (Session Description Protocol) layer, a TCP layer (Transmission Control Protocol), an IP layer and, below the IP-layer, lower protocol layers. In addition, the software comprises as a part of the client software 220 a media player for playing the received media.

The processor generates, based on the client software 220 the aforementioned PAUSE and PLAY messages and sends them to the server 111 via the radio frequency part RF. The processor also performs, based on the client software 220, the necessary calculations relating to the speeding factor, relating to the buffer filling period and relating to the suitable point in time for swiching from a lower bit rate sequence back to the original bit rate sequence.

The generation and sending of the resending request (PLAY message) and other appropriate action is triggered by a cell reselection event occurred in the mobile client device 101. The event can be detected by the client software 220 by reception of an asynchronous message from an API (Application Programming Interface) provided by the lower layers of the protocol stack 230. Alternatively or in addition, the event can be detected by monitoring the buffer level, i.e. the degree of fullness of the client device buffer 240. In this case, if the buffer 240 does not receive data for a certain amount X of time (depending on the implementation, the parameter X can be defined as a constant value, and it is a threshold for the client software 220 to understand that the cell reselection event has occurred), and if the client device 101 afterwards starts to receive data after a certain variable amount

Y of time (wherein Y > X, and Y is the real duration of the CR period), then the client can trigger the action described in this description.

Figure 10 shows a streaming server 111 according to a preferred embodiment of the invention. The streaming server 111 comprises a processing unit CPU, a first memory 310, an IP network interface 350, and a second memory 360. The first memory 310, the IP network interface 350, and the second memory 360 are coupled to the processing unit CPU.

The processing unit CPU controls, in accordance with computer software stored in the first memory 310, the operation of the streaming server 111, such as processing of requests received from the client device 101 and the sending of prerecorded media streams, stored in the second memory (disk) 360, to the client device 101 via the IP network interface 350.

The software comprises a streaming server software application 320 (hereinafter referred to as server software 320), a protocol stack 330 for implementing the necessary protocol layers such as an RTP layer, an RTSP layer, an SDP layer, a TCP layer, an IP layer and lower protocol layers.

The PAUSE and PLAY messages sent from the client device 101 are received via the IP network interface 350. A processor (not shown) of the processing unit CPU processes the messages according to the server software 320 and the protocol stack 330 and takes appropriate action.

The present invention provides means for reducing the effects of cell reselection on an ongoing streaming session. It is to be noted that, according to preferable embodiments of the invention, a request for resending (e.g. a PLAY message) is sent on the application layer, i.e. between the client software application 220 and the server software application 320. Preferably RTSP over TCP (Transmission Control Protocol) or RTSP over another reliable protocol is used in transferring

application layer requests from the client device to the streaming server. The reception of the messages at the streaming server can therefore basically be guaranteed.

Although it has been described that the cell reselection would be performed between two base stations, it should be noted that a cell reselection can be performed also between two sectors of one and the same base station. Also, it should be noted that depending on the implemention, it may well be that separate messages (PAUSE messages) pausing the sending of streaming media are not needed. In an alternative embodiment, both to stop sending streaming media and to start resending streaming media is caused by a single appropriate message.

Further, in relation to the buffer filling embodiment, it has been described in the foregoing that when the server switches to sending a lower bit rate sequence, the original transmission bit rate would be maintained. However, in an alternative embodiment of the invention, a transmission bit rate higher than the original transmission bit rate is used during the filling period in order to fill the buffer more quickly. In this embodiment, it is assumed that a larger bandwidth can be requested by the mobile client device and that a larger bandwidth is actually available in the radio access network.

Further, in relation to the buffer filling embodiment, the bandwidth information of the sequence can be sent to the streaming server 111 also by other means than the RTSP field 'Bandwidth' (for example, by requesting a specific sequence encoded at a bitrate known to the client device 101). In this case, the field 'Bandwidth' is not used, the fields 'Speed' and 'Range' are recalculated according to the actual known sequence bit rate.

It is possible that the network (air-interface) bandwidth changes during the buffer filling period. If the client device 101 supports bandwidth adaptation that involves bit stream switching, the client device 101 should pause the buffer filling (with an

RTSP PAUSE message) before sending a bandwidth adaptation message to the streaming server 111. After the bandwidth adaptation operations are over, the client device can start the buffer filling again and recalculate the 'Bandwidth', 'Speed' and 'Range' values according to the bit rate and timing information of the new media stream.

Also, in relation to the buffer filling embodiment, it should be noted that, in an alternative embodiment, the second PAUSE/PLAY messages are not sent at all but the switch back to sending the original media sequence at the original transmission bit rate will be performed automatically by the server based on the stopping point information contained in the first PLAY message.

The resending request mentioned in the various embodiments of the invention may, in certain cases, actually be a sending request. One such case is considered as an alternative embodiment of the invention. In this embodiment, the client device (101) knowing beforehand that a cell reselection is going to occur in a very near future sends a PAUSE message to the streaming server (111) before the beginning of the cell reselection period (i.e. during the pre-CR period). Sending of the PAUSE message is triggered by a cell reselection initiation event notified to the client software 220 by means of the lower layer API. The PAUSE message causes the streaming server 111 to stop sending the streaming media. When the CR-period is over, the client device 101 then sends a PLAY message causing the streaming server to start sending at the point at which sending was stopped before the CR period. Playing of the streaming media is not stopped at the client device 101 in between, and if the client device buffer 240 has been selected to be longer in time than the time that cell reselection takes, the buffer 240 does not become totally empty during cell reselection and the user does not experience any jumps or interruption in playback. The PLAY message may contain a request to send at an increased speed in order to increase the degree of fullness of the buffer 240.

In yet another embodiment, the mobile client device 101 sends, before the

CR period, to the streaming server 111 an appropriate message requesting the streaming server 111 to stop sending the streaming media and to start sending again at a suitable point in time after the CR period. Sending of the message is triggered, during the pre-CR period, by a cell reselection initiation event notified to the client software 220 by means of the lower layer API. The mobile client device 101, knowing approximately the time that the cell reselection is going to take, estimates a suitable point in time at which it is again able to receive data. It inserts this information to the message so as to prevent the streaming server 111 from starting to send again too early.

An example of the above mentioned message is as follows:

PLAY rtsp://example.com/foo RTSP/1.0
CSeq: 7
Session: 354832
Range: npt=28.00-40.00; **time=19970123T153600Z**
Bandwidth: 20000
Speed: 1.5

This is an RTSP PLAY message having a header field 'Time'. The value of that field schedules the start of the "future" sending of the streaming media.

Figure 11 illustrates (especially relating to the buffer filling embodiments) the following three ways in which messages may be sent:

Case 1: The first PAUSE and PLAY messages are sent after the CR period. The second PAUSE and PLAY messages are sent when the buffer 240 is full.

Case 2: The first PAUSE message is sent before the CR period. The associated PLAY message is sent after the CR period. The second PAUSE and PLAY messages are sent when the buffer 240 is full.

Case 3: The first PAUSE message is sent before the CR period. The associated PLAY message is sent after the PAUSE message but before the CR period. The second PAUSE and PLAY messages are sent when the buffer 240 is full.

The sending of the second PAUSE message is not necessary, if the first PLAY message contains a closed 'Range' field.

Particular implementations and embodiments of the invention have been described. It is clear to a person skilled in the art that the invention is not restricted to details of the embodiments presented above (e.g. message names and message field names), but that it can be implemented in other embodiments using equivalent means without deviating from the characteristics of the invention. The scope of the invention is only restricted by the attached patent claims.

CLAIMS:

1. A method for streaming of media from a streaming server (111) to a mobile client device (101) over an air-interface, wherein the method comprises:

requesting the streaming server (111) to send streaming media which the mobile client device (101) is not able to receive due to a cell reselection.

2. A method according to claim 1, wherein the streaming server is provided with a starting point at which to start sending.

3. A method according to claim 1, wherein streaming server (111) sends the streaming media which the mobile client device (101) is not able to receive due to said cell reselection as well as a remaining portion of streaming media in response to the request.

4. A method according to claim 1, wherein the cell reselection comprises a cell reselection period during which the mobile client device (101) is not able to receive streaming media, the method comprising:

sending from the mobile client device (101) to the streaming server (111), after the cell reselection period, a resending request which requests the streaming server (111) to resend streaming media which the mobile client device (101) was not able to receive during the cell reselection period.

5. A method according to claim 4, wherein the resending request is generated according to RTSP protocol (Real Time Streaming Protocol).

6. A method according to claim 4, wherein the resending request is implemented by an RTSP PAUSE/PLAY message pair.

7. A method according to claim 1, wherein the streaming media is temporarily stored in a temporary store (240), such as a buffer, at the client device (101)

before playing.

8. A method according to claim 7, wherein the temporary store (240) has a size longer in time than a cell reselection period.

9. A method according to claim 7, wherein the streaming server is requested to send streaming media at a rate higher than the playing rate of that media so as to increase a degree of fullness of the temporary store (240).

10. A method according to claim 9, wherein a bandwidth or desired transmission bit rate with speeding factor is communicated to the streaming server (111) in a request.

11. A method according to claim 9, wherein the streaming media is stored at the mobile client device (101) at a rate higher than the playing rate.

12. A method according to claim 9, wherein the streaming server (111) is subsequently requested to resume an original configuration.

13. A method according to claim 7, wherein a degree of fullness of the temporary store (240) decreases during the cell reselection, and the streaming server is requested to send the not received streaming media although the temporary store (240) has not become totally empty, and said requesting is performed without pausing playback at the mobile client device (101).

14. A method according to claim 1, wherein the streaming server has a set of media streams available for transmission in which the same media content has been encoded at different bit rates.

15. A method according to claim 14, wherein information on the available set of media streams is beforehand communicated to the mobile client device (101)

in a streaming session setup.

16. A method according to claim 15, wherein the streaming server (111) is requested to switch from sending a higher bit rate media stream to sending a lower bit rate media stream at an increased speed.

17. A method according to any preceding claim, wherein the streaming media comprise one of the following: a video stream, an audio stream, another stream of single media, a multimedia stream.

18. A method according to any preceding claim, wherein the streaming server (111) sends streaming media to the mobile client device (101) via a mobile communications network.

19. A method according to any preceding claim, wherein the mobile communications network comprises a mobile packet radio network, such as a GPRS (General Packet Radio Service) network.

20. A method according to any preceding claim, wherein said cell reselection is performed between two base stations (BS1, BS2) which are selected from a group comprising: base stations belonging to a GPRS system, base stations belonging to a third generation mobile communications system.

21. A mobile client device (101) for receiving streaming media from a streaming server (111) over an air-interface, the mobile client device (101) comprising:

means (220, 230, MCU) for requesting the streaming server (111) to send streaming media which the mobile client device (101) is not able to receive due to a cell reselection.

22. A streaming server (111) for sending streaming media to a mobile client device (101) over an air-interface, the streaming server (111) comprising:

means (350) for receiving a request requesting the streaming server (111) to send streaming media which the mobile client device (101) is not able to receive due to a cell reselection; and

means (320, 330, CPU) for acting upon the received request.

23. A system comprising a streaming server (111) and a mobile client device (101), for streaming of media from the streaming server (111) to the mobile client device (101) over an air-interface, the system comprising, at the mobile client device (101):

means (220, 230, MCU) for requesting the streaming server (111) to send streaming media which the mobile client device (101) is not able to receive due to a cell reselection, the system further comprising, at the streaming server (111):

means (350) for receiving the request; and

means (320, 330, CPU) for acting upon the received request.

24. A computer program executable in a mobile client device (101), the computer program comprising:

program code for causing the mobile client device (101) to request the streaming server (111) to send streaming media which the mobile client device (101) is not able to receive due to a cell reselection.

25. A computer program executable in a streaming server (111), the computer program comprising:

program code for causing the streaming server (111) to receive a request requesting the streaming server (111) to send streaming media which the mobile client device (101) is not able to receive due to a cell reselection; and

program code for acting upon the received request.

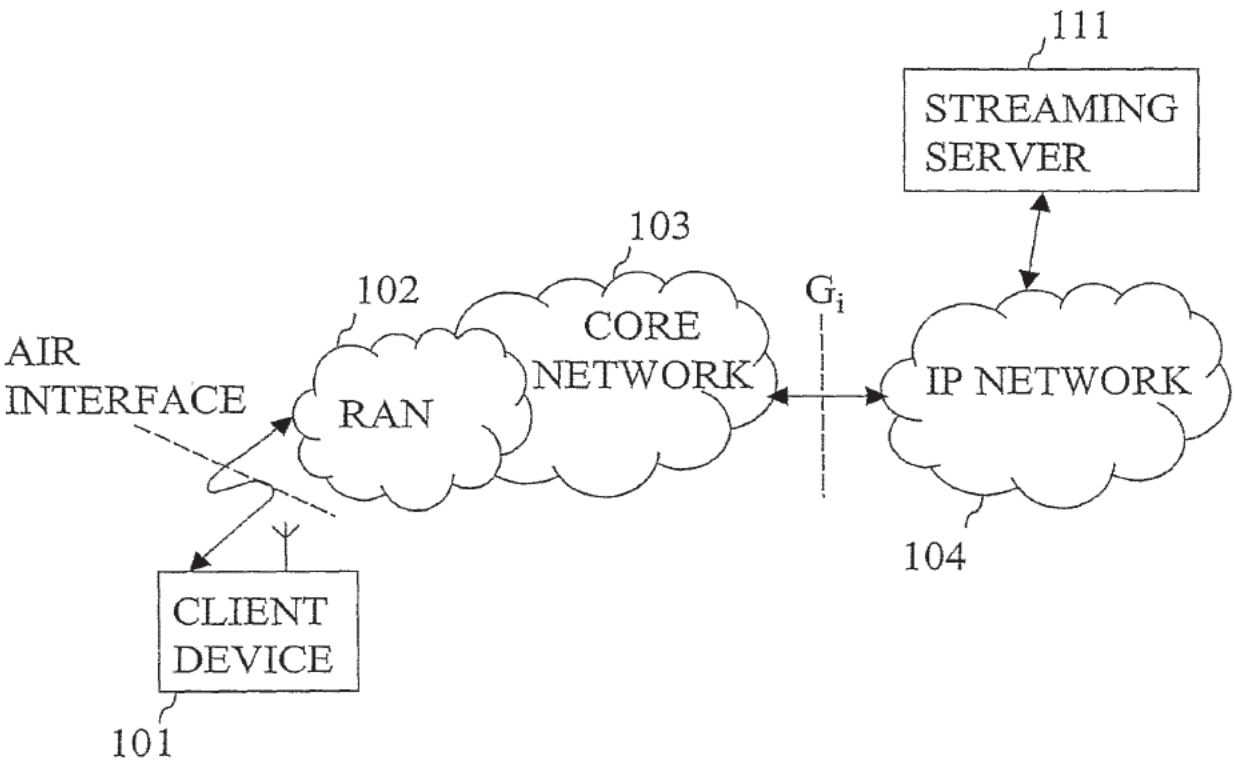

111
STREAMING SERVER
103
102
CORE NETWORK
G_i
AIR INTERFACE
RAN
IP NETWORK
104
CLIENT DEVICE
101

FIG. 1

FIG. 2

Ex. 1018-1277

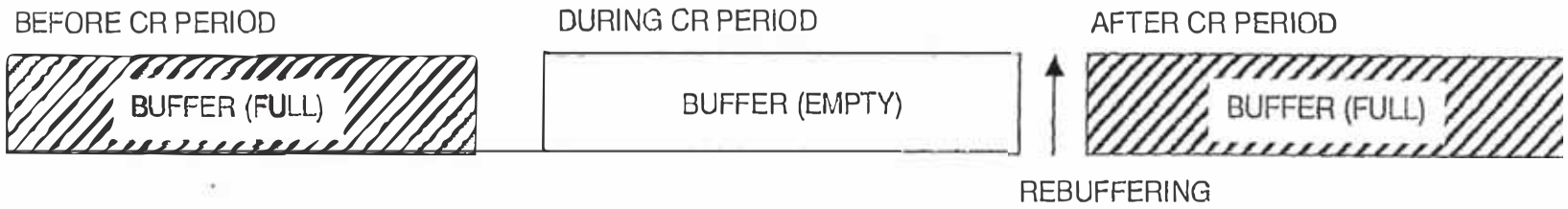


FIG. 3

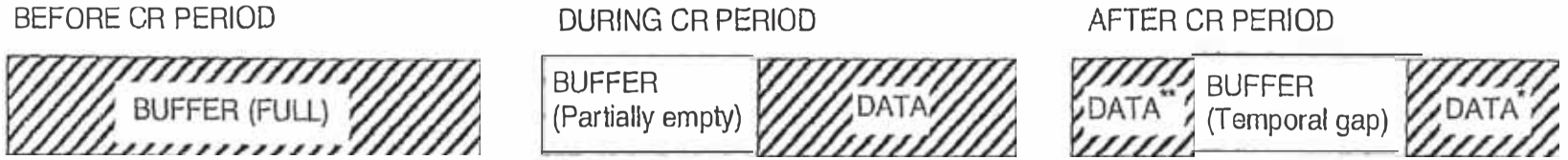


FIG. 4

Ex. 1018-1278

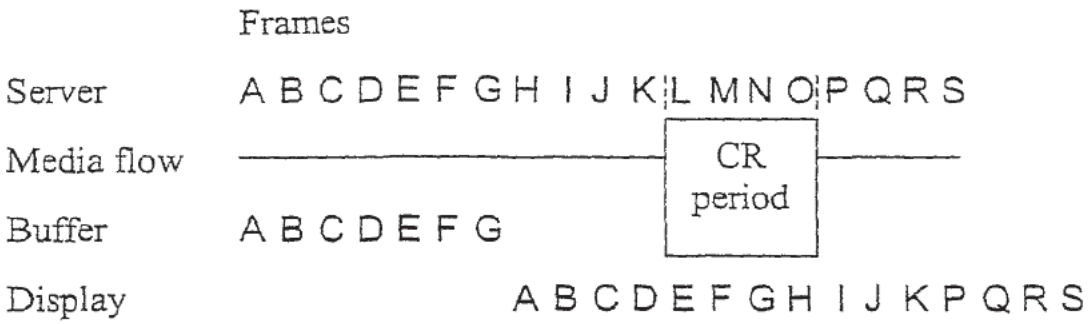

Frames
Server
A B C D E F G H I J K L M N O P Q R S
Media flow
CR period
Buffer
A B C D E F G
Display
A B C D E F G H I J K P Q R S

FIG. 5

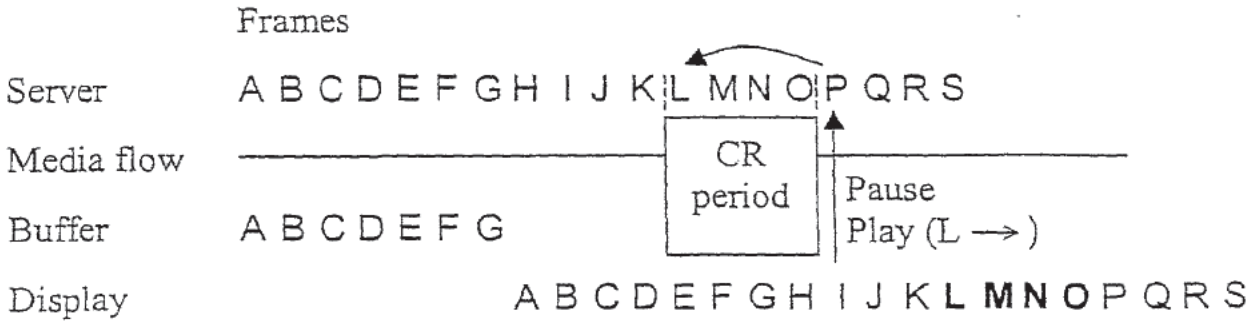

Frames
Server
A B C D E F G H I J K L M N O P Q R S
Media flow
CR period
Pause
Play (L →)
Buffer
A B C D E F G
Display
A B C D E F G H I J K L M N O P Q R S

FIG. 6

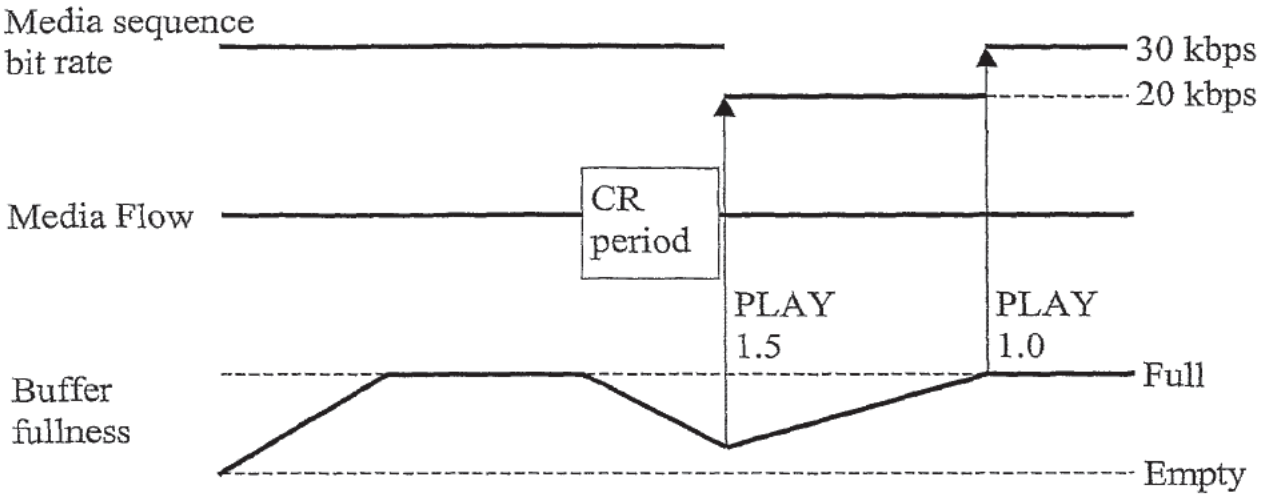

Media sequence bit rate
30 kbps
20 kbps
Media Flow
CR period
PLAY 1.5
PLAY 1.0
Buffer fullness
Full
Empty

FIG. 7

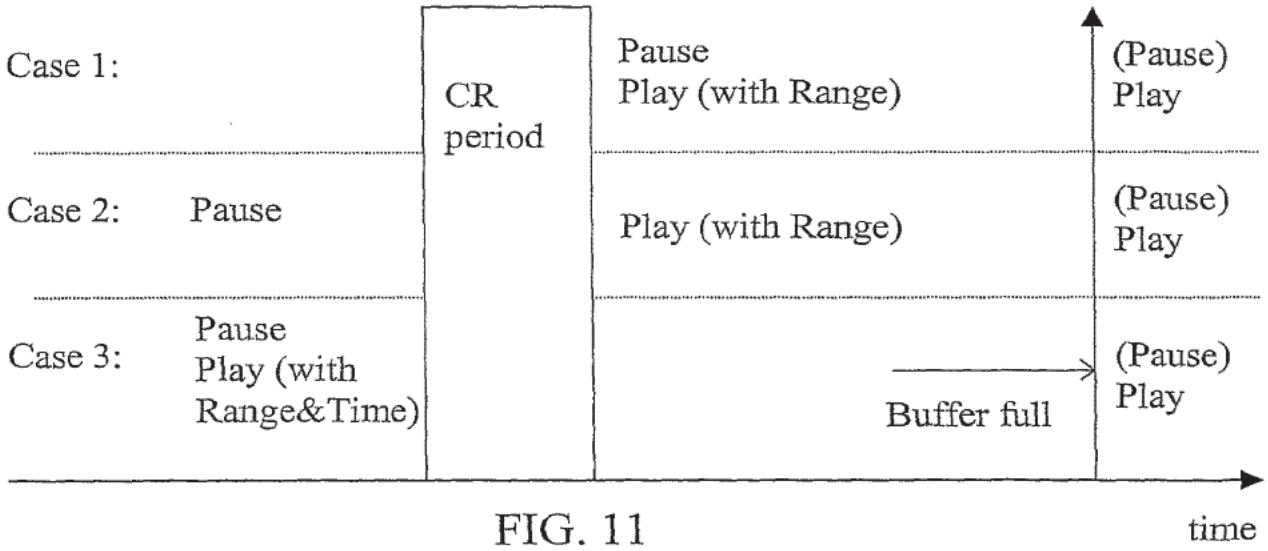

Case 1:
CR period
Pause
Play (with Range)
(Pause) Play
Case 2:
Pause
Play (with Range)
(Pause) Play
Case 3:
Pause
Play (with Range&Time)
Buffer full
(Pause) Play
time

FIG. 11

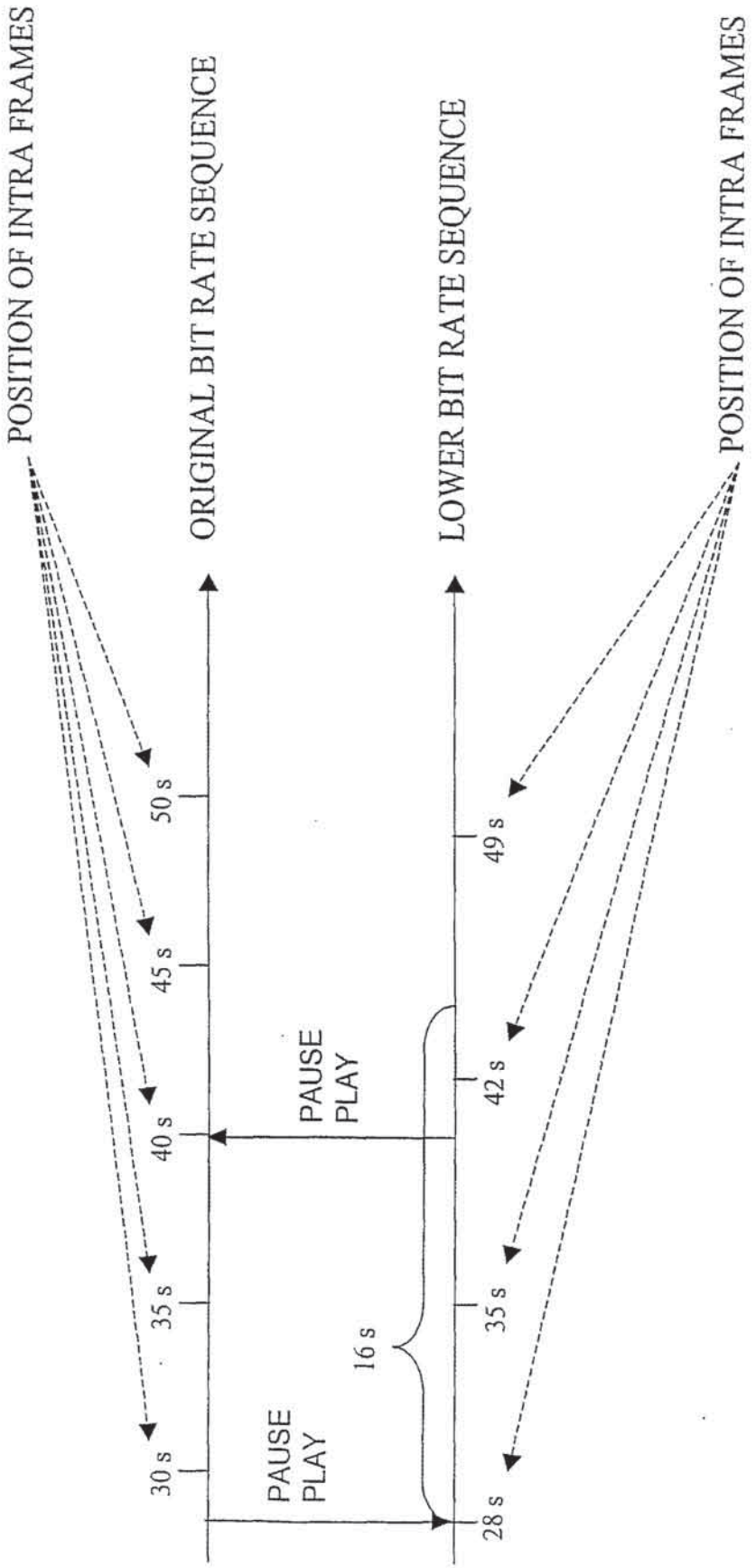


FIG. 8

Ex. 1018-1281

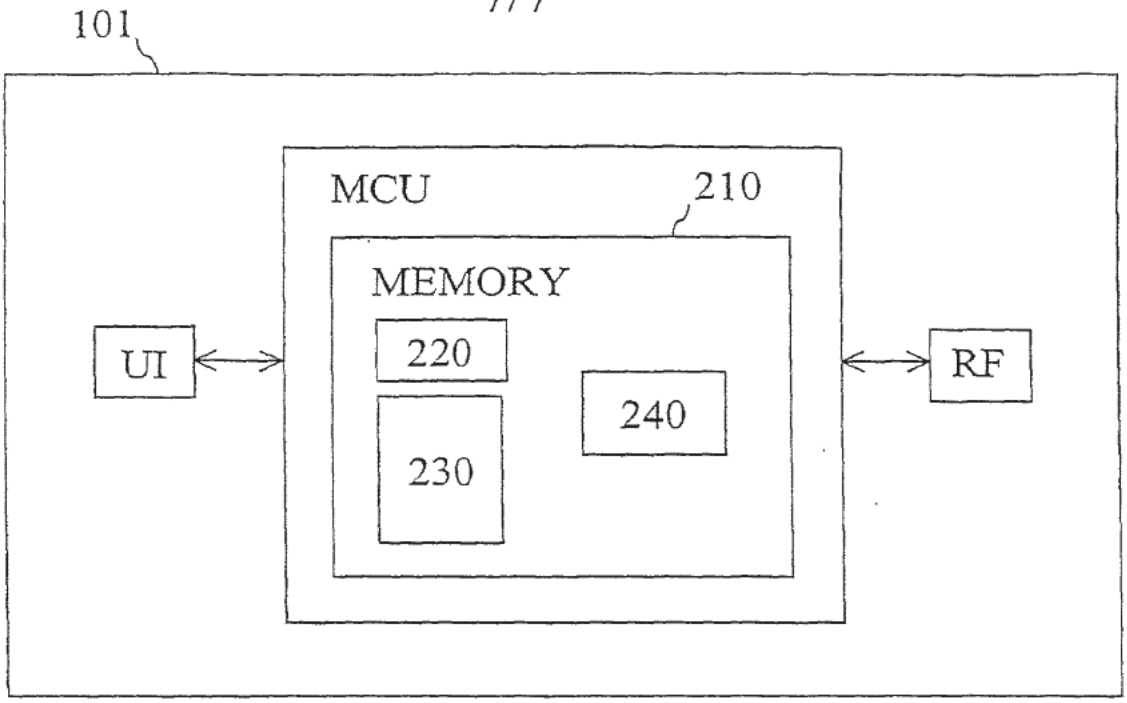

101
MCU
210
MEMORY
220
230
240
UI
RF

FIG. 9

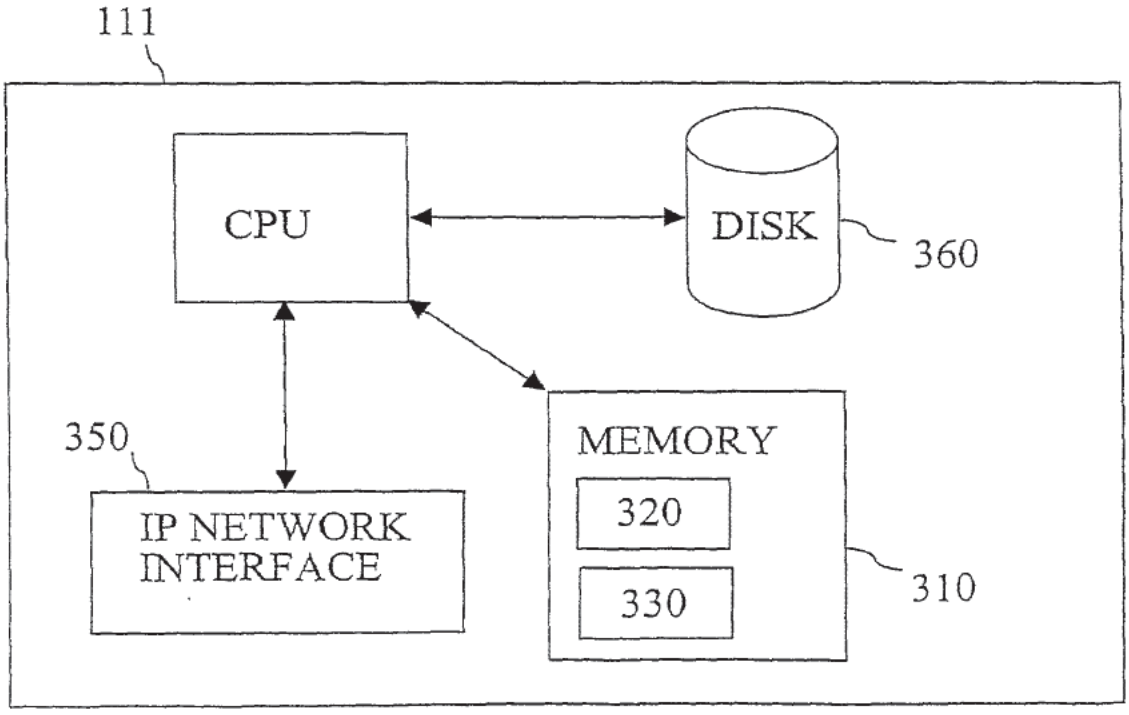

111
CPU
DISK
360
350
IP NETWORK INTERFACE
MEMORY
320
330
310

FIG. 10

INTERNATIONAL SEARCH REPORT

International application No.
PCT/FI 03/00752

A. CLASSIFICATION OF SUBJECT MATTER

IPC7: H04L 9/06
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: H04L, H04Q, H04B

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

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

EPO-INTERNAL, WPI DATA, PAJ, INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,A	US 2003009576 A1 (APOSTOLOPOULOS, J.G. ET AL), 9 January 2003 (09.01.03), paragraphs [0009] - [0011]; figures 8-11	1-25
A	3GPP TS 26.233 V5.0.0: 3rd Generation Partnership project; Technical Specification Group Services and System Aspects; Transparent end-to-end packet switched streaming service (PSS); General description (Release 5), March 2002, pages 1-12, see section 4, "Usage scenarios" and section 5, "General service architecture".	1-25

[X] Further documents are listed in the continuation of Box C. [X] 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 another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
8 December 2003	16 -12- 2003
Name and mailing address of the ISA/ Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86	Authorized officer Fredrik Blomqvist /OGU Telephone No. +46 8 782 25 00

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/FI 03/00752

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 0130090 A2 (TELEFONAKTIEBOLAGET LM ERICSSON (PUBL)), 26 April 2001 (26.04.01), page 2, line 29 - page 4, line 19 -- --------	1-25

Form PCT/ISA/210 (continuation of second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

Information on patent family members

06/09/03

International application No.
PCT/FI 03/00752

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2003009576 A1	09/01/03	WO 03005761 A	16/01/03
WO 0130090 A2	26/04/01	AU 1065601 A	30/04/01
		EP 1224829 A	24/07/02