PCT

(51) International Patent Classification 6 : H04N 7/50	A1	(11) International Publication Number: WO 97/32436
		(43) International Publication Date: 4 September 1997 (04.09.97)

(21) International Application Number: PCT/EP97/00863

(22) International Filing Date: 22 February 1997 (22.02.97)

(30) Priority Data:
96103269.5   2 March 1996 (02.03.96)   EP
(34) Countries for which the regional or
international application was filed: DE et al.

(71) Applicant *(for all designated States except US)*: DEUTSCHE THOMSON-BRANDT GMBH [DE/DE]; Hermann-Schwer-Strasse 3, D-78048 Villingen-Schwenningen (DE).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: BLAWAT, Meinolf [DE/DE]; Isernhagener Strasse 18, D-30161 Hannover (DE). HERPEL, Carsten [DE/DE]; Große Barlinge 61, D-30171 Hannover (DE).

(74) Agent: HARTNACK, Wolfgang; Deutsche Thomson-Brandt GmbH, Licensing and Intellectual Property, Göttinger Chaussee 76, D-30453 Hannover (DE).

(81) Designated States: AU, BR, CN, JP, KR, SG, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: METHOD AND APPARATUS FOR ENCODING OR DECODING DIGITAL VIDEO DATA

(57) Abstract

In a constant bit rate coder the kernel MPEG encoder circuit is followed by a buffer of a certain physical size. This buffer is used to smooth the variable bit rate output of the kernel MPEG encoder. The video buffering verifier (VBV) is a concept used in encoder control that gives an upper bound for the usable encoder buffer size as well as the minimum buffer size needed by a decoder to decode all bitstreams. The quantiser control parameters can vary more - thereby improving the resulting picture quality - if a virtual buffer size is introduced which is greater than the size of the physical buffer. The bit rate control is intended for applications using storage devices with the capability to process a variable bit rate data stream (DVD). Therefore the encoder control strategy is adapted to this situation. For the first e.g. 80 % of the total playing time an average bit rate is produced which is e.g. 95 % only of the normal average bit rate. For the remaining 20 % of the playing time the value for the desired average bit rate ascends to e.g. 150 % of the normal value. During the first 80 % the virtual buffer size is increased by a factor of e.g. 100 whereas during the last 20 % the virtual buffer size descends to the normal value.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgystan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic	SD	Sudan
CF	Central African Republic		of Korea	SE	Sweden
CG	Congo	KR	Republic of Korea	SG	Singapore
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LR	Liberia	SZ	Swaziland
CS	Czechoslovakia	LT	Lithuania	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	LV	Latvia	TJ	Tajikistan
DK	Denmark	MC	Monaco	TT	Trinidad and Tobago
EE	Estonia	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	UG	Uganda
FI	Finland	ML	Mali	US	United States of America
FR	France	MN	Mongolia	UZ	Uzbekistan
GA	Gabon	MR	Mauritania	VN	Viet Nam

**Method and Apparatus for encoding or decoding digital video
data**

The present invention relates to a method and to an appara-
tus for encoding and decoding digital video data for and
from, respectively, storage media having a given data stor-
age capacity.

Background

ISO/IEC 13818 "Information Technology - Generic coding of
moving pictures and associated audio information: Video"
(MPEG2) discloses an International standard concerning coded
digital video and audio data streams. Due to varying picture
content the generated amount of code varies over time. How-
ever, in case of fixed channel capacity normally a constant
bit rate (CBR) MPEG encoder control is desired which can be
achieved by using an encoder buffer at transmission side and
a decoder buffer at receiver side which is explained in more
detail in annexes C and D.4 of said International standard.
The other mode is variable bit rate (VBR) coding.

Invention

If there is any digital storage medium like DVD (digital
video disc) which is capable of variable bit rate operation
it might be possible to improve the coding quality by using
variable bit rate coding instead of constant bit rate cod-
ing.

It is one object of the invention to disclose a method of
variable bit rate encoding for a fixed-capacity storage me-
dium in which the coding quality is basically improved com-
pared to that of constant bit rate encoding. This object is
achieved by the method disclosed in claim 1.

It is a further object of the invention to disclose a method
of variable bit rate decoding from a fixed-capacity storage
medium in which the decoding quality is basically improved
compared to that of constant bit rate decoding. This object
is achieved by the method disclosed in claim 2.

It is a further object of the invention to disclose an en-
coding and a decoding apparatus which utilises the inventive
encoding and decoding method, respectively. This object is
achieved by the apparatuses disclosed in claims 3 and 4.

It is a further object of the invention to disclose a digi-
tal video signal which is encoded or decoded according to
the inventive method. This object is achieved by the signal
disclosed in claim 5.

"Constant bit rate" shall mean a constant number of bits
output to the channel (or e.g. the Systems multiplexer) in
each field period (20ms in a 50Hz TV signal), while
"variable bit rate" shall cover all other bit delivery
schedules.
In a CBR environment the kernel MPEG encoder circuit is fol-
lowed by a buffer of a certain size. This buffer is used to
smooth the inherently variable bit rate output of the kernel
MPEG encoder. The video buffering verifier VBV is a concept
used in encoder control that gives an upper bound for the
usable encoder buffer size as well as the minimum buffer
size needed by a decoder to decode all bit streams compliant
to this profiles @ levels (P@L). For all defined profiles @
levels maximum VBV_buffer_size values are specified in
clause 8 of ISO/IEC 13818-2. Smaller values are also allowed
and can be signalled in each bitstream by the
VBV_buffer_size element.
In a variable bit rate application there is no urgent need
to use the encoder buffer to smooth the bit rate, however,
it might still be useful to retain the buffer, as will be-
come evident below. So the change is that the channel bit
rate is variable now.

In both, CBR and VBR applications, the MPEG encoder and the
MPEG decoder operate by using the well-known hybrid DPCM/DCT
structures.

The bit rate control goals for applications using storage
devices with the capability to process a variable bit rate
data stream differ from the CBR case. It can now be desir-
able to:

- maintain uniform quality over time;
- maximise storage time on a medium of a given capacity.

Therefore the encoder control strategy needs to be adapted
to this situation.

To ease understanding the encoder control for the CBR case
shall be reviewed first.

The conventional encoder control strategy consists of three
steps or control levels:

1) *bit allocation* (global control):

The allocation of bits to the different coded frame types
(I, P, B frames) is determined according to their relative
complexity X_i, X_p, X_b. The goal is to equalise the subjec-
tive quality of the frame types. In the following feedback
implementation, which does not carry out any pre-analysis,
it is assumed that video scenes are sufficiently stationary
for some time, so that information from the past frames can
be used for the bit allocation to the coming frame or field,
depending on the coding structure. In the remainder of the
specification, "frame" may generally denote a frame or a
field or any other section of a GOP which is greater than a
unit (to be defined below).

The complexity can be estimated by any criterion reflecting
the relative coding cost for I, P, and B frames at a given
image quality. However, it should be considered whether the
necessary measurements can be obtained without too much bur-
den on the encoder complexity.

First the "image quality" should be measured. However, it is
well known that no commonly agreed, sufficiently simple, ob-

jective measure exists. Therefore, usually the signal-to-
noise ratio (SNR) of the reconstructed frame is taken as an
approximation to image quality. In the context of an MPEG
encoder control even this measure is somewhat costly as it
involves computation of SNR for each encoded frame. There-
fore the average quantiser step size (Q_i, Q_p, Q_b) used to
code a frame is taken as a quality indicator instead.
Now, having defined a measure for "quality", the
"complexity" of a frame can be expressed as the product of
bits generated and the quantiser step size used. In other
words: a frame is complex to code if it produces many bits
even if the quantiser step size is coarse. This product is
formed for each unit m (e.g. macroblock or block or slice or
any section which is smaller than a frame) and then summed
over the frame:

$$x_m = b_m * q_m \tag{1a}$$
$$X_c = \Sigma x_m, \tag{1b}$$

wherein b_m is the number of bits for coding unit m, q_m is
the average quantiser step size used in unit m, and c = i, p
or b is the coding type.
At the beginning of each Group Of Pictures GOP the bit
budget $B_G(g)$ available for the current GOP g, considering
also a carry $B_G(g-1)$ from the previous GOP, is

$$B_F(n=0) = B_G(g) = (\text{bit rate}) *$$
$$(\text{frames per GOP}) / (\text{frame rate}) - B_G(g-1) \tag{2}$$

For each frame n in the current GOP, a target number of bits
T_c, depending on its coding type c, is determined according
to the computed complexity and the remaining number of
frames of each coding type (N_i, N_p, N_b) within the current
GOP. Considering the remaining bits $B_F(n)$ after coding frame
n-1 of the GOP, the following formula can be used:

$$T_c = B_F(n) * (X_c * K_c) /$$
$$(N_i*X_i*K_i + N_p*X_p*K_p + N_b*X_b*K_b) \tag{3}$$

The factors K_c are used to reflect different visual importance of frame coding types, typically $K_i = K_p = 1$ and $K_b < 1$.

The complexity computation according to equations (1a) and (1b) is normally updated after each frame, using data from the most recent I, P and B frames, to ensure a reasonable response to scene changes. The bit budget

$$B_F(n+1) = B_F(n) - \Sigma b_m \tag{4}$$

for the remaining frames of the GOP, after subtracting the number of bits Σb_m used for the current frame, is re-distributed using formula (3) with N_i, N_p, N_b set to the remaining number of frames of each type.

After coding the last frame N of the GOP, the remaining bits $B_G(g) = B_F(N)$ are carried to the next GOP g+1.

It is also possible to explicitly allow a non-constant amount of bits per GOP, keeping only the long-time average bit rate constant. Practically this can be implemented by re-filling the bit budget $B_F(n)$ after each frame using

$$B_F(n) = B_F(n-1) - \Sigma b_m + (\text{bit rate})/(\text{frame rate}) \tag{5}$$

instead of re-filling the budget only at the start of a GOP using equation (2). This budget will then always be distributed according to the number of I, P and B frames in a GOP, rather than only on the remaining frames in the current GOP. The effect of this change in case of a scene change in the middle of a GOP is that a scene change P frame might obtain more bits which improves its quality. However, due to VBV fullness and/or the reduced bit budget, most probably the next regular I frame will not get as many bits as it should, hereby decreasing the quality of the whole next GOP. Therefore, preferably, the "constant bit budget per GOP" notion is used for CBR applications.

A minimum bit budget is needed for each frame to be able to encode at least the overhead information. This is most important for B frames that tend to be least complex, therefore getting allocated very few bits. CBR rate control may use fixed minimum bit budgets, however, an estimation of the

necessary overhead bit budget could also be made based on the past frames.

The maximum number of bits that can be allocated to a frame (usually an I frame) is bound by the status of the VBV. No more than VBV_buffer_size bits can be consumed by a single frame, given an initially empty encoder buffer.

2) *buffer control* (local control):

For each individual frame the bit production is controlled to be within some range of the allocated target number of bits, using a virtual buffer concept, that corresponds to a P-controller in control theory. This is depicted in Fig. 5. The virtual buffer VBU_c is filled from the difference e_m between the bits b_m generated by each unit (macroblock) in the MPEG2 encoder MEN, and from the portion of the target for this unit, $t_m = T_c / M$, where M is the number of units per frame.

Independent virtual buffers VBU_c with a possibly different size V_{co} are used for each frame coding type c = i, p, b, since difficulties in matching the target for one frame type should not propagate to other frame types. The status d_m of the virtual buffer VBU_c at the end of a frame is propagated as d_o to the next frame of the same coding type:

$$d_m = d_o + \Sigma (b_m - t_m) \tag{6}$$

For each unit m a quantiser step size q_m is computed in calculating means QC from the fullness of this virtual buffer d_m relative to its size V_{co}.

The so-called "non-linear quantiser" table, see table 7-6 in ISO/IEC 13818-2, (which does *not* refer to non-linear quantisation) is used to obtain a sufficiently progressive non-linear mapping of the available 31 codewords for quantiser_scale_code to actual quantiser step sizes q_m. quantiser_scale_code can be simply set to 32 * d_m / V_{co}. The very coarse quantisation available with this table can be used to deal with "panic" situations when the buffer is close to full.

The size of the virtual buffer, V_{CO}, which in control theory
is called "reaction" parameter of the P-controller, deter-
mines the tightness of control. It can be made very large
(even larger than the VBV) if it is acceptable that the num-
5   ber of bits generated deviates much from the target.
A large virtual buffer can be used, e.g., for P frames,
since it is not known a priori whether a scene change will
occur in the current GOP. In case of scene change the bit
production can be unpredictably high. However, this will re-
10   sult in impaired quality only in such cases where the vir-
tual buffer is filled to such extend that the quantiser step
size is drastically increased.
Advantageously, the reaction to an unpredicted high bit pro-
duction is delayed as much as possible, while an unpredicted
15   low bit production should result in immediate reduction of
the quantiser step size. This can be realised by different
reaction parameters V_{CO}, Vd_{CO}, (i.e. virtual buffer size),
which e.g. depend on the relation between the virtual buffer
fullness d_m when coding unit m and the initial fullness d_0
20   at the start of a frame, using:
if $d_m > d_0$

$$\text{quantiser\_scale\_code} = 32 * d_m/V_{CO}, \qquad (7)$$

else

$$\text{quantiser\_scale\_code} = 32 * (d_0/V_{CO} - (d_0 - d_n)/Vd_{CO}) \, .$$

25

Vd_{CO}, which is the reaction when d_m is going down, should be
small compared to V_{CO}. q_m is derived again by a look-up in
table 7-6 of ISO/IEC 13818-2.
The status of the VBV should be monitored in addition to the
30   control based on the virtual buffers, especially if control
is "loose", i.e. the virtual buffers are big. While the tar-
get for a frame as well as the target for a whole GOP may be
exceeded, the VBV, representing the decoder physical buffer,
may never under- or overflow.
35   Furthermore MPEG disallows frequent occurrence of units
(macroblocks) that exceed a certain number of bits (ISO/IEC
13818-2, table 8-6). This has to be monitored and controlled
as well.

3) *content adaptive quantisation*:

To reflect the varying local complexity of different parts
of an image, a local complexity measure for each unit can be
computed. Currently this is derived from the AC energy of
the luminance blocks of the unit.

The complexity factor obtained is used to modify the previ-
ously computed quantiser step size to some extend, e.g. by a
factor in the range 0.5 to 2.0. Only this final quantiser
step size information (coded as quantiser_scale_code) is
transmitted in the bitstream each time it is different from
that for the previous unit (macroblock). The aim is to dis-
tribute bits in such a way that the subjective quality
within the frame is homogenous. Alternative schemes exploit-
ing the properties of the human visual system to that end
can be applied here.

Preferably, control level 3) is disabled in case of coarse
quantisation since otherwise no strict control of bit gen-
eration is possible.

Advantageously, if the bit allocation in level 1) is suc-
cessful, level 2) should maintain a mostly constant virtual
buffer status, i.e. a constant quantiser step size, through-
out a frame.

However, the bit allocation can fail and/or the complexity
can be distributed very inhomogeneously within the frame, so
that in practice this will never be completely true.

Another reason for sub-optimum performance of the control
scheme is the fact that the complexity values computed in
step 1) are not constant for a given image, but they depend
on the quantisation step size in a way that cannot easily be
modelled. This is depicted in Fig. 6, showing the complexity
X over quantiser step size Q for I frames of 12 different
sequences. When coding natural picture sequences one can ob-
serve some degree of oscillation in bit allocation for sub-
sequent P frames, although the scene was completely tempo-
rally homogeneous. Advantageously, this behaviour can be

avoided if a complexity measure is used that is independent
of the applied quantiser step size.

In VBR applications the following additional considerations
must be made, beginning with the encoder and decoder buff-
ers.

As mentioned above, physical buffers are advantageous even
in the VBR case. The reason for this is the huge amount of
bits produced by I frames compared to other frame types.
When encoding standard TV pictures (CCIR 601 resolution) at
a reasonable quality with CBR, e.g. at 5Mbit/s, a typical
instantaneous data rate for an I frame can be more than
15Mbit/s, with "instantaneous data rate" defined as (bits
generated by a frame) / (one frame period).

As long as the peak data rate of the storage device is below
this value, a buffer is still necessary to be able to
achieve higher quality with VBR than with CBR coding. The
VBV_buffer_size of 1.75Mbit for Main Profile @ Main Level
will limit the peak instantaneous data rate to e.g. a more
than sufficient 45Mbit/s rate assuming an initially empty
buffer.

Advantageously, the VBR output of the encoder buffer and the
VBR input of the decoder buffer can be modelled as leaky
buckets. Data always leaks at the peak rate of the channel.
It leaks out of the encoder buffer as long as data is pre-
sent in it, while it leaks into the decoder buffer as long
as it is not full.

For control purposes the physical buffer status can be moni-
tored once per frame (field) period only. However, this
means that a buffer portion of the size [(peak data rate) /
(frame period)] is needed to cope with a possible unequal
bit production during the frame period. This further limits
the peak instantaneous data rate. Advantageously, buffer
control on a unit by unit (macroblock by macroblock) basis
can be used to avoid this situation.

A VBR control should reach the following goals:
As postulated at the beginning a VBR encoder control should
be more targeted to equalising the quality over the whole
duration of a recording on the storage medium. Especially,

bits should not be spent unless an additional (visible) gain in quality can be expected. From the rate control point of view, two scenarios can be distinguished.

Scenario 1) Fit a maximum duration of video on a storage medium of given capacity:
A sufficient quality can be selected by the application designer by setting the quantiser step size from control step 2) to an appropriate value. The content adaptive quantisation of step 3) should still be done, since it addresses the subjective quality harmonisation within each frame. In fact, improvements in this area are possible if vision models are used that better reflect the artefact visibility by a human observer.
Rate control is now necessary only to avoid overflow of the encoder buffer. Underflow will not occur in a VBR application, since the variable channel data rate is allowed to be zero temporarily.

Scenario 1a) Guaranteed minimum duration of video:
Using a rate control as described in *scenario* 2) below, it would furthermore be possible to guarantee a minimum recording duration on the medium. Depending on the complexity of the video, the recording duration can also be higher than this minimum.

Scenario 2) Fit a video of given duration on a medium of given capacity:
Again, it is assumed that there is no a priori analysis of the complexity distribution throughout the video material. As long as there is no complexity measure available that is independent from the quantiser step size *and* that can be derived from variables that are readily available from a reasonably low cost encoder, control solutions to equalise the video quality throughout its duration have to use some heuristics and will not be optimum.
The basic idea for a VBR control approach is to have a very loose control with varying bit rate at the beginning, while

control becomes tighter when the end of the capacity of the
storage medium is approached or when the remaining duration
of the picture material to be coded is such that a pre-se-
lected minimum bit rate can not be guaranteed throughout
this remaining duration.

It should be noted, however, that this approach privileges
the beginning of a recording, that can now exhibit constant
quality, while a constant rate with possibly lower quality
is enforced toward the end.

Advantageously, this disadvantage can be remedied to some
extent if the constant bit rate R_e that is used toward the
end of the recording is selected to be higher than the aver-
age bit rate R_{avg} for the whole recording. That average bit
rate can be computed as

$$R_{avg} = B \ / \ T \tag{8}$$

with B being the total storage medium capacity and T the
video material duration.

Advantageously, this bit rate R_e is chosen such that most
video content at the end exhibits acceptable quality while
the quality for the first part of the recording is still
better than in case of CBR coding of the whole recording.

So, from the above, for each kind of recording the unknown
variables are the quantiser step size at the start of the
encoding and the constant bit rate R_e toward the end. Unfor-
tunately, it is not easily possible to find a suitable quan-
tiser step size to start with, since without knowledge about
the complexity of a video it is difficult to estimate its
bit production at a given quantiser step size.

Furthermore, it is somewhat messy to build a control scheme
that monitors the physical buffer fullness at all times and
allows the transition from constant quantiser to constant
rate operation. This transition depends on the amount of
bits produced until a point in time t as well as on the du-
ration that remains to be encoded. Early production of an
unproportionately high amount of bits as well as a short re-
maining duration are indicators for the need of tight rate
control.

12

Therefore, the following modified rate control scheme is
proposed:

It is preferable to keep the bit rate controlled during the
5   whole recording. By having a very loose control (= very big
virtual buffers) and by removing the constant bit budget
$B_G(g)$ per GOP at the beginning it is possible to allow
rather large variations in instantaneous bit rate (i.e. a
quasi-constant quantiser step size) so that at least scene
10   changes and difficult scenes of a duration of a couple of
seconds will benefit from this loose control.

For this purpose it is necessary to define the behaviour
change of the rate control scheme over time.

First, reasonable values for the average bit rate at the be-
15   ginning and at the end of the recording need to be defined.
Taking R_{avg} as the average bit rate for this recording, com-
puted as in equation (8), it is advantageous to have for in-
stance:

$$R_{start} = 0.95 * R_{avg} \tag{9a}$$

20   $$R_{end} = 1.5 * R_{avg}, \tag{9b}$$

considering that 0.95 times the average bit rate would in
the worst case lead to a negligible loss if the whole pic-
ture sequence has a constant complexity over time, while 1.5
times the average bit rate at the end is sufficient to prop-
25   erly code most worst case picture content if it happens to
occur just at the end.

Over the whole time (with a linear transition from R_{start} to
R_{end}) R_{avg} can be regarded as composed of

$$R_{avg} = R_{start}*\tau_1 + 0.5*(R_{start}+R_{end})*\tau_2 + R_{end}*\tau_3 \tag{10a}$$

30   with the time periods τ_1, τ_2, and τ_3 normalised to be
$$\tau_1+\tau_2+\tau_3 = 1 \text{ (total time of storage medium)} \tag{10b}$$

With the above assumptions it is necessary to select one
more parameter. Setting τ_3 to 0.01 reflects the fact, that
precise control of the overall bit production can be re-es-
35   tablished in a very short time. Given e.g. one hour video
duration T this corresponds to $\tau_3*T = 36$ seconds. τ_3 can be
regarded as an overall bit number adaptation period.

Solving the equations (9a) to (10b) yields

$\tau_1 = 0.79$ and $\tau_2 = 0.20$.

With these figures loose control will occur during about 79% of the recording time while the transition phase takes about 20% of the total time. Other transition characteristics can be implemented that increase the percentage of time with loose control even more.

The time variant parameters of the rate control can now be described using

$$\alpha = \begin{cases} 0 & , \ \tau < \tau_1 \\ (\tau - \tau_1) \ / \ (\tau_2 - \tau_1) & , \ \tau_1 \le \tau \le \tau_2 \\ 1 & , \ \tau > \tau_2 \end{cases} \qquad (11)$$

Other transition characteristics for α can be implemented between time instants τ_1 and τ_2.

The bit rate used for step 1) of the rate control now is

$$R(\tau) = (1-\alpha) * R_{start} + \alpha * R_{end} \qquad (12)$$

The virtual buffer sizes V_{c0}, $c = i,p,b$, are advantageously modified to be

$$V_c(\tau) = ((1-\alpha) * V_{huge} + \alpha) * V_{c0} \qquad (13)$$

For example, an increase in size of $V_{huge} = 100$ can be used. These huge virtual buffers corresponding to a very slow re-action are intentional for short time intervals of complex picture material. However, it is not desirable that a non-complex part of the video following a long complex part could now be coded with a very coarse quantiser step size. It is therefore necessary to use the above mentioned idea of a quicker reaction parameter in case the bit production (= the virtual buffer status) decreases more than predicted, i.e. falls below the target. This parameter, Vd_c, might also be time variant according to

$$Vd_c(\tau) = ((1-\alpha) * Vd_{huge} + \alpha) * Vd_{c0} , \qquad (14)$$

but as initial guess $Vd_{huge} = 1$ is proposed, since during the whole duration a fast reaction is desirable in this case.

14

Advantageously, the re-filling of the bit budget $B_F(n)$ is
now carried out not only once per GOP.
Refill per frame n:
$B_F(n) = B_F(n-1) + (1-\alpha)*R(\tau) / (frame\ rate)$

5

Refill per GOP g:
$B_F(n=0) = B_G(g) =$
$\alpha*R(\tau)*(frames\ per\ GOP) / (frame\ rate) + B_G(g-1)$

10   Apart from the emptying of the encoder buffer and, hence,
the VBV as a leaky bucket, as described before, these can be
the only changes that need to be made to the CBR rate con-
trol as described in the first part of the specification.

15   In principle, the inventive method of encoding digital video
data is suited for storage on memory or recording means hav-
ing a given data storage capacity and allowing a variable
instantaneous storage bit rate, whereby
     - for a first part of said storage capacity the data is en-
20   coded using a slightly less than average bit rate and a sig-
nificantly more loose than average bit rate control;
     - for a second part of said storage capacity the data is
encoded using a higher than average bit rate and a tight bit
rate control.

25

In principle, the inventive method is suited for decoding
digital video data stored on memory or recording means hav-
ing a given data storage capacity and allowing a variable
instantaneous reading bit rate into decoding buffer means,
30   whereby
     - during a first part of said storage capacity the data is
decoded using a slightly less than average decoder buffer
means input bit rate and significantly higher than average
allowable buffer filling level variations;
35   - during a second part of said storage capacity the data is
decoded using a higher than average decoder buffer means in-
put bit rate and less than average allowable buffer filling
level variations.

15

Advantageous additional embodiments of the inventive methods
are resulting from the respective dependent claims.

In principle the inventive apparatus is suited for encoding
5   digital video data for storage on memory or recording means
having a given data storage capacity and allowing a variable
instantaneous storage bit rate, using an encoding loop in-
cluding discrete cosine transform means, quantiser means,
inverse quantiser means, inverse discrete cosine transform
10   means, picture store means, and entropy encoding means de-
livering the encoded output signal and being connected to
the output of said quantiser means , wherein said quantiser
means, said inverse quantiser means, and said entropy encod-
ing means are controlled in the following way:
15   -   for a first part of said storage capacity the data is en-
coded using a slightly less than average bit rate and a sig-
nificantly more loose than average bit rate control;
-   for a second part of said storage capacity the data is
encoded using a higher than average bit rate and a tight bit
20   rate control.

In principle the inventive apparatus is suited for decoding
digital video data which are stored on memory or recording
means having a given data storage capacity and allowing a
25   variable instantaneous reading bit rate into decoding buffer
means, using a chain of entropy decoding means, inverse
quantiser means, inverse discrete cosine transform means,
and adding means delivering the decoded output signal which
is also fed to picture store means the output signal of
30   which is provided to a second input of said adding means,
wherein said decoding buffer means are controlled in the
following way:
-   during a first part of said storage capacity the data is
decoded using a slightly less than average decoder buffer
35   means input bit rate and significantly higher than average
allowable buffer filling level variations;
-   during a second part of said storage capacity the data is
decoded using a higher than average decoder buffer means in-

put bit rate and less than average allowable buffer filling
level variations.

Advantageous additional embodiments of the inventive appara-
5   tuses are resulting from the respective dependent claims.

In principle, the inventive digital video signal is suited
for storage on memory or recording means having a given data
storage capacity and allowing a variable instantaneous stor-
10   age bit rate, whereby
     - during a first part of said storage capacity the signal
     is encoded using a slightly less than average bit rate and a
     significantly more loose than average bit rate control;
     - during a second part of said storage capacity the signal
15   is encoded using a higher than average bit rate and a tight
     bit rate control.
     Preferably, the inventive digital video signal is an MPEG2
     or MPEG1 or H261 encoded video signal.

20   The inventive storage medium, in particular an optical disc,
     contains or has recorded on it an inventive digital video
     signal.

25   <u>Drawings</u>

     Preferred embodiments of the invention are described with
     reference to the accompanying drawings, which show in:
     Fig. 1   constant bit rate system;
30   Fig. 2   variable bit rate system;
     Fig. 3   basic block diagram of MPEG encoder;
     Fig. 4   basic block diagram of MPEG decoder;
     Fig. 5   structure of a control loop;
     Fig. 6   example functions for complexity over quantiser
35   step size.

<u>Preferred embodiments</u>

In a CBR and in a VBR application the MPEG2 encoder MENC
produces variable bit rate data VBRD which is intermediately
5 stored in the encoder buffer ENCB. The filling level of this
buffer can be used to control the coding process in MENC.
At receiver side variable bit rate data VBRD which is inter-
mediately stored in the decoder buffer DECB are fed to an
MPEG2 decoder MDEC. The reading of that data from DECB is
10 controlled by MDEC.
In Fig. 1 the encoder buffer ENCB delivers a constant bit
rate data stream CBR to the decoder buffer DECB. In Fig. 2,
however, the encoder buffer ENCB delivers a variable bit
rate data stream VBR to the decoder buffer DECB.

15

In Fig. 3 the encoder video data input signal IE includes
macroblock data to be encoded. In case of intraframe data a
subtractor SUB simply passes the data via discrete cosine
transform computing means DCT and quantising means Q to en-
20 tropy encoding means ECOD which may include ENCB and which
deliver the encoder video data output signal OE.
In case of interframe data subtractor SUB subtracts pre-
dicted macroblock data PMD from the input signal and passes
the difference data via discrete cosine transform computing
25 means DCT and quantising means Q to entropy encoding means
ECOD. The output signal of Q is also fed to inverse quantis-
ing means Q_E^{-1} the output signal of which passes through in-
verse discrete cosine transform computing means DCT_E^{-1} to
adder ADDE in the form of reconstructed macroblock differ-
30 ence data RMDD. The output signal of ADDE is intermediately
stored in frame store and motion compensation means FS_MC_E
which also perform motion compensation on reconstructed mac-
roblock data and which deliver such predicted macroblock
data PMD to SUB and to the other input of adder ADDE.
35 Quantising means Q, inverse quantising means Q_E^{-1}, and ECOD
are controlled by the filling level of encoder buffer ENCB
in the inventive way described above. It is also possible
that FS_MC_E does not perform motion compensation.

The variable bit rate output to the storage medium can be
done using the maximum possible bit rate while there is data
available from the encoder buffer, and can be suspended
else.

In Fig. 4 the decoder video data input signal ID passes
through entropy decoding means EDEC which may include DECB,
inverse quantising means Q_D^{-1}, and inverse discrete cosine
transform computing means DCT_D^{-1} to an adder ADDD which de-
livers the decoder video data output signal OD. The output
signal of ADDD is fed to frame store and motion compensation
means FS_MC_D which also perform motion compensation on re-
constructed macroblock data and which in case of interframe
macroblock data deliver such predicted macroblock data PMD
to the other input of adder ADDE. In case of intraframe mac-
roblock data the adder ADDD simply passes the output signal
of DCT_D^{-1}.
Inverse quantising means Q_D^{-1} and EDEC are controlled ac-
cording to the kind of control for Q_E^{-1} and ECOD. Normally,
in motion compensation means FS_MC_D the same motion vector
information is used as in motion compensation means FS_MC_E.
It is also possible that FS_MC_D does not perform motion
compensation.
In order to handle situations of high bit rate demand, e.g.
in case of scene changes and complex scene content, the de-
coder buffer DECB is filled from the storage medium with
preferably the maximum possible bit rate. This can be done
generally in case DECB is not full whereby no data from the
storage medium is stored into DECB if DECB is full. As an
alternative, DECB can be filled with a bit rate which corre-
sponds to the respective bit rate when storing data on the
memory or recording means. In either case the reading bit
rate out of DECB corresponds to that when encoding, i.e. the
allowed and sometimes really occurring average buffer fill-
ing level variations are greater in the first part ($\approx 80\%$) of
the total capacity and are less in the last part ($\approx 20\%$) of
the total capacity. "Average allowable buffer filling level
variations" means the variation over a length of some GOPs.

The invention can be used in any means like DVD, other opti-
cal or magneto-optical or ferro-optical discs, electronic or
optical RAM or ROM, and magnetic tape systems which use a
given capacity for storing video data.
The invention is not limited to MPEG2 encoders and decoders,
but can also be used with MPEG1 (ISO/IEC 11172-2), H261
(ISO/IEC), and other coding and decoding methods in e.g.
standard TV, HDTV, and videophone applications.

Claims

1. Method of encoding digital video data for storage on
   memory or recording means having a given data storage
   capacity and allowing a variable instantaneous storage
   bit rate (VBR), **characterised** in that
   - for a first part of said storage capacity the data is
     encoded (MENC; Q, Q_E^{-1}, ECOD) using a slightly less than
     average bit rate and a significantly more loose than av-
     erage bit rate control;
   - for a second part of said storage capacity the data is
     encoded (MENC; Q, Q_E^{-1}, ECOD) using a higher than aver-
     age bit rate and a tight bit rate control.

2. Method of decoding digital video data stored on memory
   or recording means having a given data storage capacity
   and allowing a variable instantaneous reading bit rate
   (VBR) into decoding buffer means, **characterised** in that
   - during a first part of said storage capacity the data is
     decoded (MDEC) using a slightly less than average de-
     coder buffer means (DECB) input bit rate and signifi-
     cantly higher than average allowable buffer filling
     level variations;
   - during a second part of said storage capacity the data
     is decoded (MDEC) using a higher than average decoder
     buffer means (DECB) input bit rate and less than average
     allowable buffer filling level variations.

3. Apparatus for encoding digital video data for storage on
   memory or recording means having a given data storage
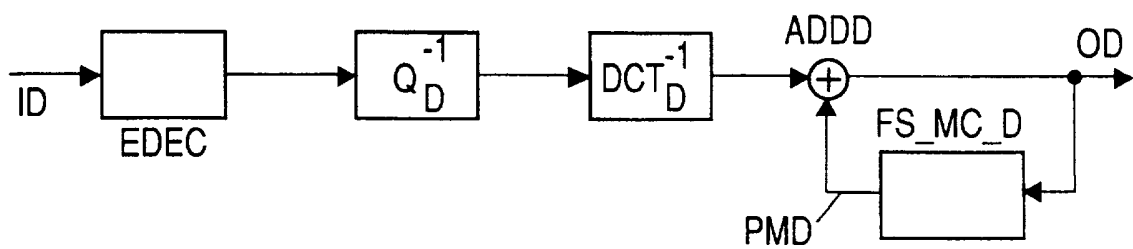   capacity and allowing a variable instantaneous storage
   bit rate (VBR), using an encoding loop including dis-
   crete cosine transform means (DCT), quantiser means (Q),
   inverse quantiser means (Q_E^{-1}), inverse discrete cosine
   transform means (DCT_E^{-1}), picture store means (FS_MC_E),
   and entropy encoding means (ECOD) delivering the encoded
   output signal (OE) and being connected to the output of
   said quantiser means, **characterised** in that said quan-

tiser means (Q), said inverse quantiser means (Q_E^{-1}),
and said entropy encoding means (ECOD) are controlled in
the following way:
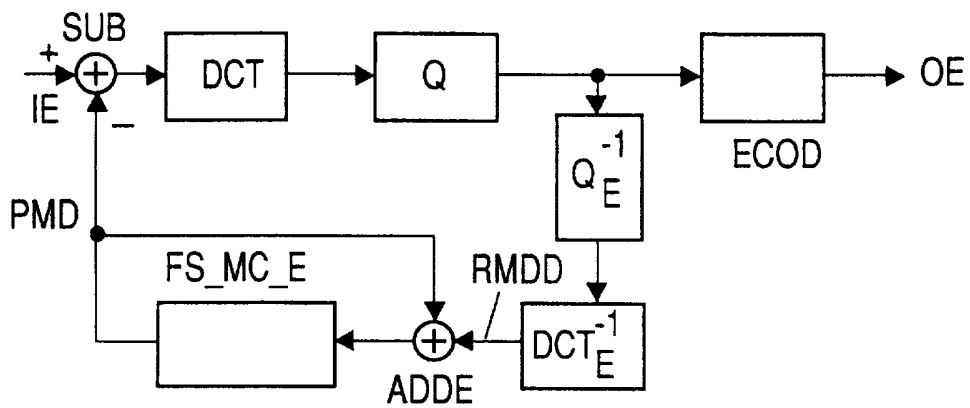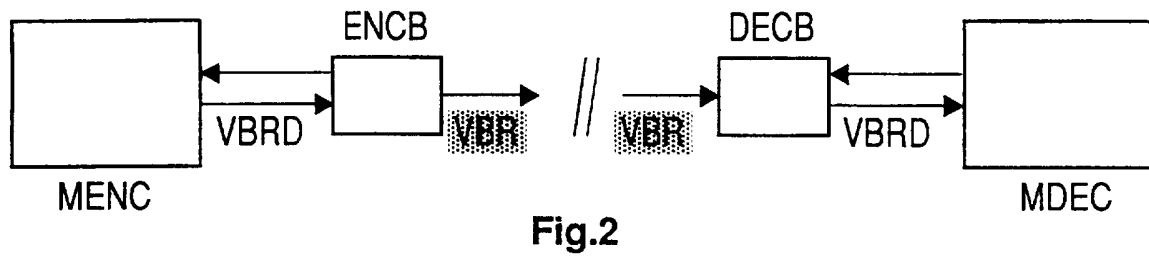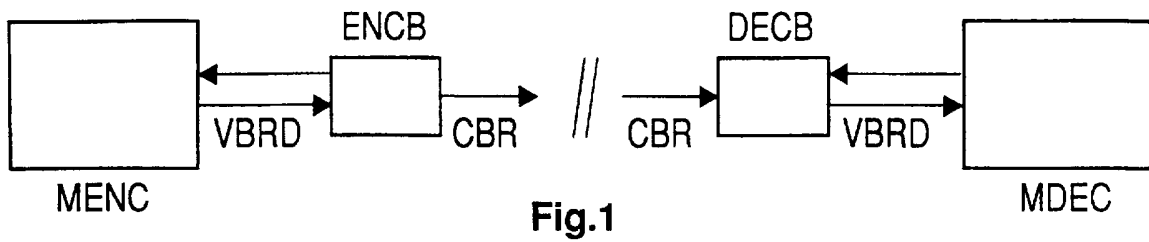
-    for a first part of said storage capacity the data is
5   encoded using a slightly less than average bit rate and
          a significantly more loose than average bit rate con-
          trol;
-    for a second part of said storage capacity the data is
          encoded using a higher than average bit rate and a tight
10   bit rate control.

4.   Apparatus for decoding digital video data which are
          stored on memory or recording means having a given data
          storage capacity and allowing a variable instantaneous
15   reading bit rate (VBR) into decoding buffer means
          (DECB), using a chain of entropy decoding means (EDEC),
          inverse quantiser means (Q_D^{-1}), inverse discrete cosine
          transform means (DCT_D^{-1}), and adding means (ADDD) deliv-
          ering the decoded output signal (OD) which is also fed
20   to picture store means (FS_MC_D) the output signal of
          which is provided to a second input of said adding
          means, **characterised** in that said decoding buffer means
          (DECB) are controlled in the following way:
-    during a first part of said storage capacity the data is
25   decoded (MDEC) using a slightly less than average de-
          coder buffer means (DECB) input bit rate and signifi-
          cantly higher than average allowable buffer filling
          level variations;
-    during a second part of said storage capacity the data
30   is decoded (MDEC) using a higher than average decoder
          buffer means (DECB) input bit rate and less than average
          allowable buffer filling level variations.

5.   Digital video signal which is suited for storage on mem-
35   ory or recording means having a given data storage ca-
          pacity and allowing a variable instantaneous storage bit
          rate, whereby

- during a first part of said storage capacity the signal is encoded using a slightly less than average bit rate and a significantly more loose than average bit rate control;

5   - during a second part of said storage capacity the signal is encoded using a higher than average bit rate and a tight bit rate control.

6. Signal according to claim 5, which is an MPEG2 or MPEG1
10   or H261 encoded video signal.

7. Storage medium, in particular an optical disc, contain-
        ing or having recorded on it a signal according to
        claims 5 or 6.

15

8. Method, apparatus, signal, or storage medium according
        to any of claims 1 to 7, wherein in said first part of
        said storage capacity said encoding and decoding, re-
        spectively, is carried out using an average data rate
20   which is about 95% of that basic average data rate which
        would be allowed for said memory or storage means if no
        separation into said first and second part would take
        place, and wherein in said second part of said storage
        capacity said encoding and decoding, respectively, is
25   carried out using an average data rate which is about
        150% of that basic average data rate.

9. Method, apparatus, signal, or storage medium according
        to any of claims 1 to 8, wherein said first part of said
30   storage capacity occupies about 80% of the total avail-
        able storage capacity for said digital video data and
        wherein said second part of said storage capacity occu-
        pies about 20% of said total available storage capacity.

35 10. Method, apparatus, signal, or storage medium according
        to any of claims 1 to 9, wherein between said first part
        and said second part of said storage capacity there is a
        further part in which

for encoding the average bit rate and bit rate control
and
for decoding the average decoder buffer means (DECB) in-
put bit rate and the average allowable buffer filling

5   level variations
are changed to form a smooth transition.

Fig.1

Fig.2

Fig.3

Fig.4
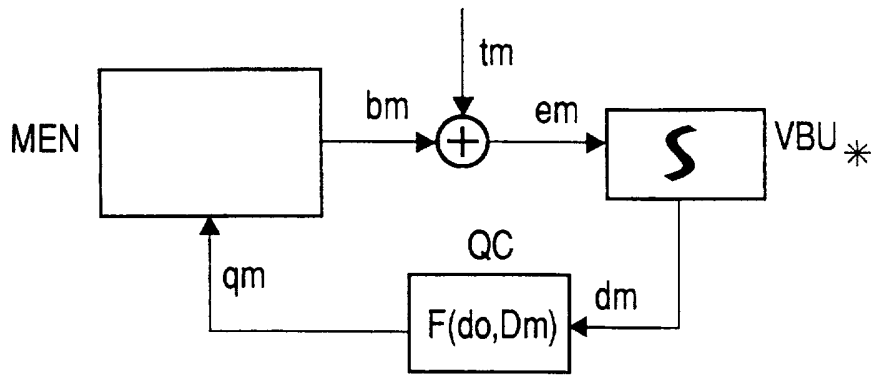
Fig.5

Fig.6

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6   H04N7/50

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 6   H04N

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)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP 0 595 268 A (VICTOR COMPANY OF JAPAN) 4 May 1994 see column 1, line 29 - column 2, line 17 see column 7, line 53 - column 8, line 17 see column 16, line 26 - column 17, line 12 see column 19, line 25 - line 47 see claim 3 see figures 4,7 ---	1-6,8,9
A	EP 0 664 651 A (HITACHI LTD) 26 July 1995 see column 4, line 8 - line 41 see column 6, line 18 - column 7, line 10 see column 10, line 40 - line 51 see column 11, line 14 - line 24 see claims 9-11,17-19 --- -/--	1-7

[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
4 June 1997	17.06.97

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	Dippel, U

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

1

INTERNATIONAL SEARCH REPORT

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5 134 476 A (ARAVIND RANGARAJAN ET AL) 28 July 1992 see column 2, line 6 - line 24 see column 4, line 60 - column 5, line 17 see column 6, line 36 - line 64 see figure 3 ---	1-6
A	EP 0 637 891 A (SONY UK LTD) 8 February 1995 see column 1, line 38 - column 3, line 11 see column 4, line 53 - column 5, line 9 see column 6, line 31 - line 41 see column 7, line 36 - column 8, line 1 ---	1-5
A	EP 0 670 663 A (PHILIPS ELECTRONIQUE LAB ;PHILIPS ELECTRONICS NV (NL)) 6 September 1995 see page 2, line 36 - line 54 -----	5,6,8-10

1

INTERNATIONAL SEARCH REPORT

information on patent family members

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
EP 0595268 A	04-05-94	JP	6141298 A	20-05-94
		JP	6165125 A	10-06-94
EP 0664651 A	26-07-95	JP	7212710 A	11-08-95
US 5134476 A	28-07-92	NONE		
EP 0637891 A	08-02-95	GB	2280566 A	01-02-95
		JP	7093891 A	07-04-95
		US	5612788 A	18-03-97
EP 0670663 A	06-09-95	FR	2717029 A	08-09-95
		JP	7284109 A	27-10-95