

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
20 September 2007 (20.09.2007)

PCT

(10) International Publication Number
WO 2007/105118 A2

(51) International Patent Classification: Not classified

(21) International Application Number:
PCT/IB2007/001768

(22) International Filing Date: 9 March 2007 (09.03.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0650866 14 March 2006 (14.03.2006) FR

(71) Applicant (for all designated States except US): CANON
KABUSHIKI KAISHA [JP/JP]; 3-30-2, Shimomaruko,
3-chome, Ohta-ku, Tokyo (JP).

(72) Inventors; and

(75) Inventors/Applicants (for US only): LE FLOCH, Hervé
[FR/FR]; 10, rue Monseigneur Duchesne, F-35000 Rennes
(FR). GISQUET, Christophe [FR/FR]; 61, bd de Sévigné,
F-35700 Rennes (FR).

(74) Agent: SANTARELLI; 14, Avenue de la Grande-Armée,
F-75017 Paris (FR).

(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,

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

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

Declaration under Rule 4.17:

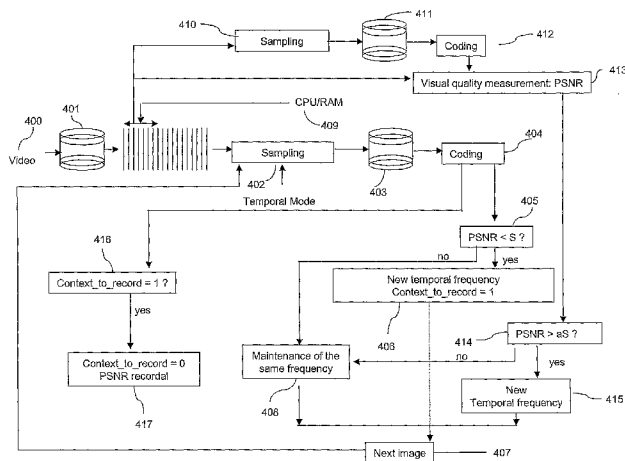
— of inventorship (Rule 4.17(iv))

Published:

— without international search report and to be republished
upon receipt of that report

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

(54) Title: A METHOD AND DEVICE FOR ADAPTING A TEMPORAL FREQUENCY OF A SEQUENCE OF VIDEO IMAGES



(57) Abstract: The invention concerns a method of adapting a temporal frequency of a sequence of video images for the purpose of its transmission over a communication network, characterized in that images of the sequence having been sampled at a temporal frequency f_1 . The method and device for adapting a temporal frequency of a sequence of video images, the method comprises a step of deciding as to the carrying out of a step of simulating a coding of images of the video sequence sampled at a temporal frequency $f_2 > f_1$, for the purpose of determining whether the sampling temporal frequency f_1 of the sequence can be increased, the decision being taken on the basis of at least one criterion (409; 513) relative to the resources of a communication apparatus able to perform the simulation step (412; 516) and/or on the basis of the evolution over time of the characteristics of the video sequence and/or of the network (512, 515).

WO 2007/105118 A2

"A method and device for adapting a temporal frequency of a sequence of video images"

5 The invention concerns a method and device for adapting a temporal frequency of a sequence of video images for the purpose of its transmission over a communication network.

 When a video sequence is compressed, for example using a coding algorithm in accordance with the MPEG-4 standard, the quality of the images of
10 the video sequence after decompression may prove to be bad.

 This case is encountered generally when the images of the video sequence are highly textured and have strong motion and/or when the rates are low.

 In such conditions, it is known to temporally downsample the video
15 sequence, which amounts to deleting certain of the images of the sequence.

 Thus, the quality of the images resulting from the downsampling is better in that the compression rate is lower.

 If it is desired for example to compress into one Megabit 50 images in one second, it can be understood that with a temporal downsampling by a
20 factor of two, only 25 images per second are compressed into one Megabit.

 On account of this, the compression rate applied to each of the images of the video sequence is reduced and their quality is thereby improved.

 The impression of fluidity resulting from such a downsampling is often less good but can be accepted in certain cases.

25 This may occur when it is considered that the clarity of the images resulting from the downsampling takes priority, or else that it is compensated for by a temporal interpolation after the decompression of the video sequence.

 It is known from the document US 6 633 609 to improve the conventional methods of video coding.

Conventionally, each image of a video sequence is coded by a coder if the calculating resources of that coder are available. It may prove that the coder is already busy and cannot therefore process the current image of the video sequence. In this case, the current image is deleted and the same applies
5 for the other images of the sequence when the coder is already active.

Since these image deletions are not regular, jerks arise in the sequence.

In this document it is proposed to delete images at regular intervals in order to avoid that phenomenon.

10 To that end, the method proposed aims to evaluate the mean time for compression of an image by the coder and to generate a sampling temporal frequency of the video which conforms to that mean time.

However, this method is not sufficiently effective since the frequency is so generated once and for all, even if the activity of the coder varies over
15 time.

A coding method using rate-distortion models is also known from the document entitled "Rate-Distortion Models for Video Transcoding", SPIE Conference on Image and Video Communications and Processing, January 2003.

20 According to this method, a first rate-distortion model is used in the case of a simple quantization of the images.

In this case, the sampling temporal frequency of the video is assumed to be maximum and the value of the distortion depending on the intended rate is supplied by that first model.

25 The first rate-distortion model involves an equation linking the rate and the distortion which is simple and well-known to the person skilled in the art.

This method also involves a second rate-distortion model which is used in the case in which images are regularly deleted and it is then considered
30 that the sampling temporal frequency is reduced in comparison with the preceding case.

As regards the second model, this presupposes that the images of the same scene are stationary and that the distortion of the missing images (the missing images are replaced by the closest decoded image from a temporal point of view) may be deduced using the property of stationarity.

5 To do this, an analytical temporal distortion model is generated from a parameter learning phase and a phase of segmenting the video into homogeneous scenes.

Thus, according to the teaching of this document, the two rate-distortion models provide, over a time interval, two measurements of the distortion, i.e. a mean distortion provided by the first model at maximum
10 temporal resolution and a mean distortion provided by the second model and which takes into account the temporal downsampling of the images.

The decision to downsample the video sequence will then be taken on the basis of the distortion values calculated by those models.

15 It will be noted that this method is particularly complicated to implement, in particular as regards the learning and segmenting phases, and involves numerous calculations. Furthermore, it is based on an interpolation model which can prove to be of low reliability.

The present invention aims to mitigate at least one of the drawbacks
20 mentioned above by providing to adapt the sampling temporal frequency of a video sequence in a simple manner.

To that end, the invention concerns a method of adapting a temporal frequency of a sequence of video images for the purpose of its transmission over a communication network, characterized in that images of the sequence
25 having been sampled at a temporal frequency f_1 , the method comprises a step of deciding as to the carrying out of a step of simulating a coding of images of the video sequence sampled at a temporal frequency $f_2 > f_1$, for the purpose of determining whether the sampling temporal frequency f_1 of the sequence can be increased, the decision being taken on the basis of at least one criterion
30 relative to the resources of a communication apparatus able to perform the simulation step and/or on the basis of the evolution over time of the characteristics of the video sequence and/or of the network.

Thus, before deciding on an increase in the sampling temporal frequency, it will be decided whether it is opportune to carry out a simulation of coding at that frequency, this being done according to different conditions.

5 The invention is thus particularly flexible since it makes it possible to adapt the sampling frequency dynamically, according to different conditions being met which may evolve over time.

Moreover, the invention is particularly simple to implement and proves to be more precise than the technique used in the prior art which involves rate-distortion models.

10 It will be noted that the video sequence may be characterized by an energy or a video activity which may be stronger or weaker and by a visual quality which may be higher or lower.

As regards the network, this may be characterized by its transmission capacity (e.g. bandwidth available, data transmission time, etc.)
15 which may be better or worse (e.g. higher or lower bandwidth, greater or lesser data transmission time, etc.).

According to a feature, the evolution over time of the characteristics of the video sequence and/or of the network is noted with respect to the initial characteristics presented by the video sequence and/or by the network when it
20 has been decided to use the temporal frequency f_1 for the sampling.

Observation will now be made of the temporal evolution of the context in which the sampling temporal frequency has been modified to pass to the frequency f_1 . The context is defined by the state of the video sequence and/or of the network at a given moment.

25 According to a feature, the method comprises, prior to the step of deciding as to the carrying out of a simulating step, a step of storing in memory characteristics of the video sequence and/or characteristics of the network at a given time in relation with a sampling temporal frequency of the images of the video sequence.

30 It is thus provided to store in memory the context defined by the video sequence and/or the network in order to be able to follow its evolution over time.

It will be noted that the storage in memory of this context may take place before or after having modified the sampling temporal frequency to the value f_1 .

5 This storage in memory may be useful in particular for reasons of following the evolution of the video context and/or of the network context, for example, for statistical purposes.

According to a feature, the step of storing in memory is carried out after it has been decided to reduce the sampling temporal frequency of the images of the video sequence from a frequency f_0 to the frequency f_1 .

10 It will be noted that that the storage in memory may take place when the sampling temporal frequency of the sequence is reduced and/or, as mentioned above, at other times, for example, to obtain a history of the evolution of the context over time.

The recording of the video context and/or of the network context 15 when the sampling frequency is reduced will make it possible later, on examining the evolution of that context, to decide whether or not to increase the sampling frequency of the sequence.

According to a feature, the method comprises, prior to the step of deciding as to the carrying out of a simulating step, the following steps:

- 20
- sampling images of the video sequence at a temporal frequency $f_0 > f_1$,
 - coding the sampled images,
 - determining the quality of the coded images,
 - comparing the determined quality with respect to a predetermined
- 25 threshold,
- according to the result of the comparison, deciding as to a reduction of the sampling temporal frequency of the images of the video sequence from the frequency f_0 to the frequency f_1 .

30 Thus, the decision to reduce the sampling temporal frequency of the video sequence to the value f_1 has been taken after estimation of the quality of the coded sampled images.

It will be noted that when a step of memory storage is provided, this may be carried out at any time with respect to any one of the aforementioned steps of sampling, coding, quality determination, comparison and decision.

5 The recording may also take place in parallel with any one of those steps.

By way of example, when it is decided to reduce the sampling frequency from f_0 to f_1 , the recording may take place before that decision is taken, after it, or after modification of the frequency, or in parallel with the frequency modification.

10 According to a feature, the method comprises a step of comparing between the current characteristics presented by the video sequence and/or the network and the initial characteristics presented by the video sequence and/or the network when it was decided to use the temporal frequency f_1 for the sampling.

15 Current characteristics means characteristics of the video sequence and/or of the network after passage of a certain time, consecutively to the first taking of a decision to modify the temporal frequency.

These current characteristics are, for example, those existing at the time of the decision taking as to the coding simulation.

20 This comparison makes it possible to determine the evolution over time of the features of the video sequence and/or of the network.

According to a feature, the step of comparing the characteristics is in particular performed in the form of a step of comparing the qualities of the video sequence obtained respectively with the current characteristics (current context)
25 and initial characteristics (initial context) of that sequence.

It will be noted that this comparing step assumes that the quality of the video sequence in the initial context has been stored in memory and that it is determined in the current state.

30 According to a feature, the method comprises, according to the result of the comparing step, a step of deciding as to an increase in the sampling temporal frequency from f_1 to f_2 .

It is thus possible to decide on the increase of the temporal frequency directly according to the result of the comparison of the current and initial characteristics and, more generally, according to the evolution over time of those characteristics, and thus to dispense with the coding simulation step.

5 This evolution provides an approximate indication which makes it possible to take a rapid decision. However, if according to the circumstances (e.g. type of video data to transmit) it is preferred to obtain more detail on the evolution of the context before deciding on an increase in frequency, then the prior step of coding simulation is preferable.

10 According to a feature, the method comprises a step of increasing the sampling temporal frequency, when the current characteristics of the video sequence and/or of the network have improved over time.

Thus, when the video and/or network context has favorably evolved, it is possible to envisage increasing the sampling temporal frequency directly, without having recourse to the coding simulation step.

15 This enables time to be saved and to reduce the calculation cost of the method.

According to a feature, when the current characteristics of the video sequence and/or of the network have improved over time, the carrying out of the step of simulating coding of sampled images at the temporal frequency $f_2 > f_1$ depends on the state of the resources of the communication apparatus with respect to a predetermined threshold.

When the video and/or network context has favorably evolved, the resources of the communication apparatus are taken into account before deciding on carrying out the coding simulation step.

It can however be envisaged in certain circumstances not to take into account those resources and nevertheless to carry out the coding simulation. This may be envisaged when there is no need to rapidly take a decision to increase the frequency or when the video data may possibly be coded more slowly.

30 According to a feature, the state of the resources of the communication apparatus being below the predetermined threshold, the method

comprises a step of increasing the sampling temporal frequency without having recourse to the coding simulation step.

Thus, when the state of the resources (calculation capacity, memory space) of the communication apparatus is insufficient, it can be provided, in certain circumstances, to dispense with the coding simulation and to directly increase the sampling frequency.

It will however be noted that when the state of the resources permits, the coding simulation step can also be envisaged in order to evaluate the quality of the video sequence so coded in a simulated manner, before deciding on an increase in sampling frequency.

According to a feature, when the current characteristics of the video sequence and/or of the network have degraded over time, the coding simulation step is not carried out.

Thus, depending on the following of the video and/or network context, in particular when the context has degraded, it can be deduced thereby that the coding simulation step is of no use since the quality of the video sequence at an increased sampling frequency will very probably be insufficient.

According to another feature, the method comprises a step of simulating coding of images of the video sequence sampled at the temporal frequency $f_2 > f_1$ when the state of the resources of the communication apparatus is greater than a predetermined threshold.

Thus, when the state of the resources of the apparatus permits, a coding simulation is carried out.

More particularly, the simulation step is subdivided into several sub-steps:

- sampling images of the video sequence at the temporal frequency f_2 ,
- simulating coding of the sampled images,
- determining the quality of the coded images,
- comparing the determined quality with respect to a predetermined threshold,

- in case the threshold is exceeded, increasing the sampling temporal frequency of the images of the video sequence.

Thus, if the quality of the coded images arising from the simulation
5 proves insufficient, the same sampling temporal frequency of the images of the video sequence is kept.

According to a feature, the characteristics of the video sequence are the video activity of the sequence, for example, the variance of the prediction errors, the variance of the motion vectors, and/or the quality of the video
10 sequence.

This quality of the video sequence or of an image can be expressed with respect to the signal to noise ratio of the video sequence or of the image or of several images after coding.

Moreover, the characteristics of the network are for example defined
15 by the bandwidth of the network.

The invention also concerns a device for adapting a temporal frequency of a sequence of video images for the purpose of its transmission over a communication network, characterized in that images of the sequence having been sampled at a temporal frequency f_1 , the device comprises means
20 for deciding as to the carrying out of a simulation of coding of images of the video sequence sampled at a temporal frequency $f_2 > f_1$, for the purpose of determining whether the sampling temporal frequency f_1 of the sequence can be increased, the decision being taken on the basis of at least one criterion relative to the resources of a communication apparatus able to perform the
25 simulation step and/or on the basis of the evolution over time of the characteristics of the video sequence and/or of the network.

This device for implementing the method described above has the same advantages as it does.

The invention also relates to:

30 - an information carrier readable by a computer system, possibly wholly or partly removable, in particular a CD-ROM or magnetic medium, such as a hard disk or a diskette, or a transmissible medium such as an electrical or

optical signal, said information carrier comprising instructions of a computer program, characterized in that it enables the implementation of the method briefly described above, when that program is loaded and executed by the computer system.

5 - a computer program loadable into a computer system, said program containing instructions enabling the implementation of the method briefly described above, when that program is loaded and executed by the computer system.

Other features and advantages will appear in the following
10 description, which is given solely by way of non-limiting example and made with reference to the accompanying drawings, in which:

- Figure 1 is a diagram of a communication apparatus in which the invention may be implemented;

15 - Figure 2 is a schematic representation of the environment of the invention;

- Figure 3 is a schematic view of an algorithm for determining a temporal frequency of a sequence of video images according to the invention;

20 - Figure 4 is a schematic view of an algorithm for determining a temporal frequency of a sequence of video images according to a first embodiment of the invention;

- Figure 5 is a schematic view of an algorithm for determining a temporal frequency of a sequence of video images according to a second embodiment.

25 As represented in **Figure 1**, a device 110 for implementing the invention is for example implemented in the form of a micro-computer connected to different peripherals.

Among the peripherals are a digital video camera 1100 connected to a graphics card not shown and which provides data to be processed to the device 110.

30 It will be noted that the video camera may be replaced by any means for image acquisition or storage, or even by a scanner able to communicate data to the device 110.

The device 110 comprises a communication interface 1102 connected to a communication network 1103 over which digital data are transmitted.

5 The device 110 may receive those data to be processed from the network 1103 or may transmit them over the network after having processed them.

The device 110 also comprises a data storage means 1104 such as a hard disk.

10 A drive 1105 for a disc 1106 is also to be found in the device 110, it being possible for the disc to be a diskette, a CD-ROM or a DVD-ROM.

The disc 1106 just as for the hard disk 1104 may contain data processed according to the invention as well as a computer program or programs implementing the invention.

15 This program or programs may for example be contained in the storage medium 1106 and transferred into the device 110 to be stored there, for example, on the hard disk 1104.

According to a variant, the program or programs enabling device 110 to implement the invention may be stored in read only memory 1107 (ROM).

20 According to another variant, the program or programs may be received by the device 110 from the communication network 1103 to be stored there in identical manner to that which has already been described.

The device 110 is also connected to a microphone 1108 to process audio data.

25 A screen 1109 makes it possible to view the data to be processed or the processed data, or to serve as an interface with the user who may thus parameterize certain processing modes, using a keyboard 1110 or any other means such as a mouse or another pointing device.

The device also comprises a central processing unit 1111 (CPU) which executes the instructions relative to the implementation of the invention.

30 These instructions or lines of code are stored in the read only memory 1107 or in the other aforementioned storage means.

On powering up the device, the processing program or programs according to the invention which are stored in a non-volatile memory, for example the memory 1107 (ROM), are transferred into the random access memory 1112 (RAM) which will then contain the executable code of the program or programs, as well as registers for storing the variables necessary for the implementation of the invention.

More generally, a data storage means, readable by a computer or a micro-computer, stores the program or programs implementing the method according to the invention and, more particularly, a method of coding, transmission and decoding of data.

It will be noted that the data storage means may be integrated or not into the device 110 and may possibly be removable.

The device 110 also comprises a communication bus 1113 enabling the different aforesaid components to be linked together, whether they are integrated into the device 110 or connected thereto, and so makes it possible to establish the communication between those different elements.

The representation of the bus 1113 is not limiting and in particular the central processing unit 1111 is capable of communicating instructions to any component of the device 110 or component connected thereto, whether directly or via another component of the device.

It will be noted that the data processed by the device 110 are data from a sequence of video images.

As represented in **Figure 2**, the invention applies in particular in the context of a transmission of a sequence of video images over a communication network from a communication apparatus which is, for example, identical to the device 110 of Figure 1.

Upstream of the transmission, a module 200 for acquiring a video sequence is provided, for example, in the form of a camera delivering images in a non-compressed format.

In the example illustrated it is assumed that the frequency of video acquisition is 30 images per second.

The images acquired by the module 200 are next transferred to the video coding module 201 which is, for example, a video coder in accordance with the MPEG-4 standard.

5 Each image compressed by the module 201 is next cut up into data packets by the module 203 and the packets so formed are transmitted over the network by the transmission module 204.

It should be noted that the transmission of the packets over the network is carried out in conformity with the constraint of bandwidth $B(t)$ of the network, under the supervision of the control module 205.

10 The variable t is a time index and the bandwidth of the network which is determined at a given time may thus evolve over time.

It should furthermore be noted that the value of the bandwidth $B(t)$ is known to the video coding module 201 which thus adapts the compression rate of the images, and thus their quality, so as to be able to transmit all the packets
15 over the network. When the bandwidth value $B(t)$ is too low, the compression rate is too high and the quality of the video strongly decreases.

In such a case, it is provided to adapt the sampling temporal frequency of the video sequence by deleting some of the images provided by the video acquisition module 200.

20 The module 202 has the role of determining the appropriate sampling temporal frequency of the images of the video sequence.

The video sequence of which the temporal frequency has been adapted one or more times by the module 202 is transmitted over the network.

The network is for example a wireless network.

25 It will be noted that the modules 200 to 205 form part of the communication apparatus referred to as sender.

The transmitted packets are successively received by a main data reception module 206 and by a packet reception module 207 in which they are assembled together to constitute a binary file.

30 The data constituting this file are then processed by the video data decoding module 208.

When the decoding of the images of the video sequence has been carried out, those images, or the video in its entirety, may undergo post-processing in order to improve the visual quality.

Such processing is carried out by the post-processing module 209
5 and may, for example, recover the initial temporal frequency of the video sequence via a temporal interpolation method.

The module 209 may furthermore implement methods of suppressing block effects and numerous other methods known to the person skilled in the art.

10 The display module 210 next carries out the display of the video sequence.

The modules 206 to 210 form part of a communication apparatus referred to as receiver and which is, for example, identical to the device 110 of Figure 1.

15 It will be noted that, in the context described above, the video acquisition and the coding thereof are performed in real time.

However, the adaptation of the temporal frequency of the video sequence according to the invention may also be carried out on a video that has already been compressed, for example, in MPEG-4 or other format.

20 In that case, transcoding of the compressed video is then necessary in order to adapt the size of the compressed video to the bandwidth constraints of the network.

This transcoding may consist in requantizing and/or modifying the temporal frequency.

25 The algorithm represented in **Figure 3** illustrates in more detail a part of the different functionalities implemented by the module 202 of Figure 2.

It will be noted that in general, the module 202 of Figure 2 must take a decision as to the temporal frequency to adopt for the sampling of the video sequence on the basis of criteria which will be defined below. This decision thus
30 leads either to downsampling of the images of the video sequence, or to increasing the sampling temporal frequency.

The algorithm of Figure 3 comprises a first step 300 of acquiring a video sequence, for example with a camera.

On acquiring the video sequence, the images thereof are sampled at a temporal frequency f_0 .

5 During the following step 301 this video sequence is coded and a step 303 enables control of the rate allocated to each image of the video sequence.

More particularly, during step 303, the rate control makes it possible to adapt the coding parameters taking into account the bandwidth $B(t)$ available over the communication network.

10 During the following step 304, the visual quality of the sampled and coded images is determined.

It is thus possible, for example, to use the PSNR (Peak Signal to Noise Ratio) as a measure of the visual quality of an image the video sequence.

15 The peak signal to noise ratio is determined by the following formula:

$$\text{PSNR} = 20\text{Log}_{10} (255/\text{RMSE}),$$

where RMSE designates the square root of the MSE and MSE designates the Mean Square Error on a color component of an image (such as the luminance or the chrominance), and is determined by the following formula, where L represents the width of the image and H its height:

$$\text{MSE} = \frac{1}{L \times H} \sum_{i=0}^{L-1} \sum_{j=0}^{H-1} \left(X(i, j) - \tilde{X}(i, j) \right)^2$$

It will be noted that the Mean Square Error may be calculated directly during the quantization phase which is implemented at the video coding step 301.

25 After determining the visual quality of a coded image, comparison is carried out during the following step 305 of that quality to a predetermined threshold S.

When the visual quality of the image sampled at the frequency f_0 and coded is less than the predetermined threshold, this means that the spatial quality of the images must be improved.

30

It will be noted that this threshold is determined empirically and depends on the type of the video data and/or on the envisaged application. Thus, for example, it may be equal to 29dB for an application related to video conferencing and may be lower for video surveillance applications.

5 To that end the sampling temporal frequency of the images of the video sequence should thus be reduced.

This decision is taken at step 306.

It will be noted that the case in which the visual quality of the coded images is greater than the threshold S has not been envisaged in Figure 3 in
10 the interest of clarity.

However, in such a case the sampling frequency f_0 of the video sequence is not modified.

As soon as the decision to reduce the frequency f_0 to the frequency f_1 has been taken, it is provided during the step 307 to record the conditions
15 which have given rise to that decision being taken.

More particularly, storage in memory is for example made of the characteristics or properties of the video sequence and/or of the characteristics or properties of the network at a given time.

The context which thus led to reducing the sampling temporal
20 frequency of the video sequence is stored in memory and a variable denoted "context_to_record" which initially has the value 0, is set to 1.

Thus, on coding the following image, if that variable has the value 1, the quality determined at step 304 (e.g. PSNR) is recorded as the initial context value and the variable is immediately reset to 0 after that recordal.

25 It will be noted that the context which is stored in memory at step 307 is for example the bandwidth $B(t)$ available during the change in frequency and the video activity (the variance of the prediction errors, the variance of the motion vectors).

Further to the decision taking of step 306, step 308 provides for
30 determining a new reduced temporal frequency f_1 .

For example, the temporal frequency f_0 is divided by two.

Thus, the algorithm of Figure 3 makes it possible to decide on a temporal downsampling on the basis of a given criterion and stores in memory a context in which that decision was taken and which will serve later to return to a higher sampling temporal frequency.

5 The algorithm of **Figure 4** which will now be described defines conditions in which the sampling frequency may be increased according to a first embodiment of the invention.

 According to this algorithm, the decision to carry out a step of simulating coding of images of the video sequence sampled at a temporal
10 frequency f_2 higher than f_1 is taken on the basis of at least one criterion relative to the resources of the communication apparatus which may perform the simulation step.

 This simulating step is there to precisely determine, according to real conditions, whether a new frequency may be adopted for the sampling.

15 The algorithm of Figure 4 comprises a first step 400 of acquiring a video sequence with a camera and a step 401 is for example provided for temporarily storing the video data so acquired.

 During the following step 402 downsampling of the video sequence is carried out at the reduced frequency f_1 .

20 This is because, at step 306 of Figure 3 the decision to reduce the sampling frequency from f_0 to f_1 has already been taken.

 The images of the sequence so sampled may next, for example, be stored temporarily at step 403, then be coded at step 404.

 The following step 405 provides for determining, for example, the
25 quality of an image so coded and to compare it to the threshold S as the two steps 304 and 305 of Figure 3 provide.

 It will be noted that the processing envisaged here is made image by image.

 When the quality obtained is less than the predetermined threshold,
30 a new reduced temporal frequency is selected (step 406), and the conditions (context) in which that decision to reduce the frequency was taken are stored in

memory (characteristics of the video sequence and/or characteristics of the network).

More particularly, the variable "context_to_record" is set to 1.

It will be noted that for more detail reference may be made to the description made above in relation to Figure 3 which defines the passage to a lower frequency during steps 307 and 308.

The following step 407 makes it possible to select the following image of the sequence sampled at the new frequency and the following operations are then carried out in identical manner:

- 10 - coding of that new image,
- determining the quality of that coded image and
- comparing with the threshold, as has just been described for the preceding image.

Returning to step 405, when the visual quality of the coded image is greater than the threshold, the following step 408 provides for keeping the same sampling frequency and the following image of the sequence sampled at the frequency f_1 is passed on to (step 407) as described above.

In parallel with these operations, it is provided at step 409 to analyze the state of the resources of the communication apparatus and, in particular, to determine whether calculating resources and memory space are available in that apparatus to perform a coding simulation.

For example, that availability may be determined with respect to a threshold defining a maximum level of occupancy of the calculating unit and of the memory space.

When the state of the resources so permits, a certain number of consecutive images of the video sequence are selected on leaving step 401. It will be noted that, according to the state of these resources, it may be decided to perform the coding simulation by adapting the number of images selected and, for example, merely using a few images (1, 2 or 3) if the resources are close to the threshold. The time for calculating the coding of these images may also be spread out over a lapse of time greater than that imposed by real time. This will induce a slight temporal offset in a possible decision to increase the

temporal frequency, but this will enable a few more images to be taken into account during the decision taking (5 or 6 images).

The selected images are next downsampled at step 410 at a temporal frequency f_2 greater than the frequency f_1 used for the sampling of step 402.

The level of downsampling applied at step 410 is, for example, two times less than the level of downsampling applied at step 402.

The images so downsampled are for example stored temporarily at step 411, then coded at step 412.

It will be noted that to the extent that the coding steps 404 and 412 use the same images, some calculations which are carried out at step 412, during the simulation of the second coding, can be reused subsequently at the coding step 404.

During the following step 413, determination is made of the quality of each of the images of which the coding has been simulated, for example by determining their visual quality as described above at step 304 of Figure 3.

During the following step 414, comparison is made of the quality of the coded images to a threshold aS , with $a > 1$, in order to determine whether the quality is considerably greater than the quality threshold S .

In practice, if there are several coded images this step is only carried out on the last selected image in order to ensure that the quality of the image used for this test is as stable as possible (and thus the most representative possible) of the qualities of all the selected images.

In the affirmative, step 414 is followed by step 415 which authorizes the increase in the temporal sampling frequency from f_1 to f_2 .

For example, the rate of downsampling is then divided by two during this step.

Returning to step 414, when the quality of the images of which the coding has been simulated proves to be insufficient, it is decided not to modify the sampling frequency as already explained with reference to step 408 described above.

Thus, as has just been described at steps 409 to 415, the availability of the calculating resources is determined and, possibly, the memory space, in order to decide whether it is possible to simulate, in parallel with a first coding (step 404) made at a given temporal frequency, a second coding with a higher temporal frequency.

By way of example, if it is found that the coding of one second of video with the current temporal frequency uses 50% of the machine resources, it will be possible to select 0.5 seconds of video to simulate a second coding at steps 410 and 412.

It is known that in a video coder the estimation of motion between two images proves to be very costly in terms of calculations.

Taking this into account, it is thus possible to estimate the motion between two images during the coding step 404 on the basis of the motion which was estimated at the coding step 412.

Thus, if at step 412 the motion is calculated between the images $I(0)$ and $I(1)$, and between the images $I(1)$ and $I(2)$, then by simple addition, the motion between $I(0)$ and $I(2)$ is estimated and may serve as a first approximation during the coding step 404. The contrary, i.e. the re-use at step 412 of calculations carried out at step 404, is also possible.

This make it possible to greatly reduce the search space and calculation time.

It should be noted that when the decision is taken to sample the images at the reduced frequency f_1 at step 402, the steps 307 and 308 of Figure 3 have been carried out.

Thus, the initial context has been recorded (for example: video activity and bandwidth) and the variable "context_to_record" has been set to 1.

The following image sampled at the frequency f_1 is then coded at step 404 before its visual quality is determined at step 405.

Nevertheless, in parallel with the coding or thereafter, step 416 is carried out which verifies whether the aforementioned variable is the value 1.

In the affirmative, the following step 417 provides for setting that variable to 0 and for recording the value of the last visual quality (PSNR) determined at step 304 of Figure 3.

It will however be noted that this recording (apart from that of the PSNR value) could alternatively take place at step 307 of Figure 3.

Generally, the visual quality of an image depends on the video activity and on the transmission capacity of the network. More particularly, for a given bandwidth, the higher the video activity, the lower its visual quality.

If, during the test of step 416, the variable is at 0 then no context (PSNR) is to be recorded since this signifies that the sampling frequency has not been reduced.

Figure 5 illustrates an algorithm for adapting the sampling frequency of a video sequence according to the evolution over time of characteristics of the video sequence and/or of the network (context), in accordance with a second embodiment of the invention.

As will be seen below this algorithm makes it possible in particular, under certain circumstances, to reduce the use of the machine resources (calculation unit and/or memory space) while taking into account the temporal evolution of the aforementioned context.

The algorithm of Figure 5 comprises a first step 500 of acquiring a video sequence which is then stored at step 501, downsampled at the frequency f_1 at step 503 according to a decision to reduce the temporal frequency concerned at step 502 (corresponding to step 306 of Figure 3), stored at step 504, then coded at step 505.

All these steps are identical to the respective steps 400, 401, 402, 306, 403 and 404 of Figure 4.

Similarly, step 506 of comparing with respect to the threshold S of the quality of the coded images and the decision to reduce the temporal frequency of sampling and recording of the new context concerned at step 507, as well as the decision to keep the same sampling frequency concerned at step 509 are identical to the respective steps 405, 406 and 408 of Figure 4.

After these steps, processing of the following image is proceeded with at step 508 which is analogous to step 407 of Figure 4.

After coding the following image at step 505, step 510 is provided in order to determine the manner in which the context has evolved over time.

5 Depending on the result of this test step, it will be decided whether the simulation of a second coding at a higher sampling temporal frequency proves to be useful or not.

It will be noted that step 510 may occur at other locations in the algorithm, and not necessarily after the coding, when the context is characterized by the video activity and/or the transmission capacity of the network.

10 More particularly, during step 510 comparison is made between the context representing the video sequence and/or the network when it has been decided to reduce the sampling temporal frequency (context 515 recorded beforehand), and the so-called current context (video activity and current bandwidth 512) which represents the state of the video sequence and/or of the network at the present time or shortly before.

It will be noted that the initial context 515 was recorded at step 307 of Figure 3, the recording 515 also including the recording of the visual quality of the coded image (PSNR) just after (step 417 of Figure 4) reduction of the frequency to f_1 .

25 The context is for example defined by the characteristics presented by the video sequence and/or by the network at a given time and it may, for example, be the visual quality (e.g. PSNR) of the video or of an image, the video activity (for example the variance of the prediction errors and/or the variance of the motion vectors), as well as the available bandwidth at the time of reference.

30 However, it is possible, during this step, to take into account only the visual quality of the images and thus to compare the visual quality of an image recorded initially at step 417 of Figure 4, that is to say when it has been decided to reduce the sampling temporal frequency of the video sequence, with the visual quality, referred to as current, of the image coded at step 505.

This coded image results from a sampling at the reduced frequency.

Thus, when the temporal evolution of the context shows that the current characteristics of the video sequence and/or of the network have improved over time, that is to say for example if the visual quality of the current video sequence is greater than that of the initial video sequence or if the bandwidth is greater than the prior bandwidth, then it is probable that the visual quality of the video sequence at a higher sampling temporal frequency is good.

In this case, the simulation can be envisaged of a second coding of images of the video sequence which are sampled at a temporal frequency greater than the frequency f_1 of step 503.

It will however be noted that this coding simulation may be subordinated to the state of the machine resources (calculation unit and/or memory space) available at a given time.

The step of verifying the state of the machine resources with respect to a predetermined threshold (level of occupancy of the calculation unit or of the memory space) is carried out at step 513 which is identical to step 409 of Figure 4.

It will however be noted that it is also possible to perform the step of coding simulation without taking into account the state of the machine resources.

Thus, on adopting a sampling frequency greater than the sampling frequency used at step 503, the steps are executed of downsampling 514, storage 515, coding 516, determining a visual quality 517 and comparing 518 that quality with the threshold aS .

These steps are identical to the respective steps 410, 411, 412, 413 and 414 of Figure 4.

If the visual quality of the video sequence at the greater frequency proves to be sufficiently good ($PSNR > aS$), then the following step 519 provides for adapting the frequency by selecting a greater temporal frequency which is, for example, that at which the coding simulation was carried out.

On the other hand, if the visual quality proves insufficient ($PSNR \leq aS$), the sampling frequency used at step 503 is kept (step 509).

It should be noted that when the state of the machine resources used at step 513 proves to be less than the predetermined threshold (level of occupancy), it is possible to envisage directly increasing the sampling temporal frequency without having recourse to the coding simulation provided at steps 5 514 and following.

Thus calculation time is saved and decision speed is increased.

Returning to the comparing step 510, when the context has degraded over time, which results, for example, in a current visual quality (PSNR) of the video sequence which is less than or equal to the visual quality of the video 10 sequence at the time the sampling temporal frequency was reduced, it is then very probable that the visual quality of the video sequence at a higher temporal frequency will be insufficient.

More particularly, this is explained by the fact that the current context proves to be less good than that which had already led to a reduction in the 15 temporal frequency.

In this case, the coding simulation provided at steps 514 and following proves to be of no use and step 510 is followed by step 520 which resets the value of the determined visual quality (PSNR) to 0. Step 520 is then followed by step 518 already described earlier and which, given the value of the 20 PSNR, directly leads to step 509 of keeping the same sampling frequency.

It will also be noted that the recording of the context via the variable "context_to_record" is carried out in an identical manner to that which was described with reference to Figures 416 and 417 of Figure 4.

It should be noted that when a higher sampling temporal frequency is 25 selected, for example when it is doubled, the current context corresponds to the context which was recorded during the previous reduction in sampling temporal frequency.

CLAIMS

1. A method of adapting a temporal frequency of a sequence of video
5 images for the purpose of its transmission over a communication network,
characterized in that images of the sequence having been sampled at a
temporal frequency f_1 , the method comprises a step of deciding as to the
carrying out of a step of simulating a coding of images of the video sequence
sampled at a temporal frequency $f_2 > f_1$, for the purpose of determining whether
10 the sampling temporal frequency f_1 of the sequence can be increased, the
decision being taken on the basis of at least one criterion (409 ; 513) relative to
the resources of a communication apparatus able to perform the simulation step
(412 ; 516) and/or on the basis of the evolution over time of the characteristics
of the video sequence and/or of the network (512, 515).

15 2. A method according to claim 1, characterized in that the evolution
over time of the characteristics of the video sequence and/or of the network is
noted with respect to the initial characteristics presented by the video sequence
and/or by the network when it has been decided to use the temporal frequency
 f_1 for the sampling.

20 3. A method according to claim 1 or 2, characterized in that it
comprises, prior to the step of deciding as to the carrying out of a simulating
step, a step of storing (307) in memory characteristics of the video sequence
and/or characteristics of the network at a given time in relation with a sampling
temporal frequency of the images of the video sequence.

25 4. A method according to claim 3, characterized in that the step of
storing (307) in memory is carried out after it has been decided (306) to reduce
the sampling temporal frequency of the images of the video sequence from a
frequency f_0 to the frequency f_1 .

30 5. A method according to one of claims 1 to 4, characterized in that it
comprises, prior to the step of deciding as to the carrying out of a simulating
step, the following steps:

- sampling (300) images of the video sequence at a temporal frequency $f_0 > f_1$,
- coding (301) the sampled images,
- determining the quality (304) of the coded images,
- 5 - comparing (305) the determined quality with respect to a predetermined threshold,
- according to the result of the comparison, deciding (306) as to a reduction of the sampling temporal frequency of the images of the video sequence from the frequency f_0 to the frequency f_1 .

10 6. A method according to one of claims 1 to 5, characterized in that it comprises a step of comparing (510) between the current characteristics (512) presented by the video sequence and/or the network and the initial characteristics (515) presented by the video sequence and/or the network when it was decided to use the temporal frequency f_1 for the sampling.

15 7. A method according to claim 6, characterized in that the step of comparing (510) the characteristics is in particular performed in the form of a step of comparing the qualities of the video sequence obtained respectively with the current and initial characteristics.

20 8. A method according to claim 6 or 7, characterized in that it comprises, according to the result of the comparing step, a step of deciding (518) as to an increase in the sampling temporal frequency from f_1 to f_2 .

25 9. A method according to claim 8, characterized in that the method comprises a step of increasing (519) the sampling temporal frequency, when the current characteristics of the video sequence and/or of the network have improved over time.

30 10. A method according to claim 6 or 7, characterized in that, when the current characteristics of the video sequence and/or of the network have improved over time, the decision to carry out the step of simulating coding of sampled images at the temporal frequency $f_2 > f_1$ depends on the state of the resources (513) of the communication apparatus with respect to a predetermined threshold.

11. A method according to claim 10, characterized in that, the state of the resources of the communication apparatus being below the predetermined threshold, the method comprises a step of increasing (507) the sampling temporal frequency without having recourse to the coding simulation
5 step.

12. A method according to claim 6 or 7, characterized in that, when the current characteristics of the video sequence and/or of the network have degraded over time, the coding simulation step is not carried out.

13. A method according to one of claims 1 to 10, characterized in
10 that it comprises a step of simulating coding (412 ; 516) of images of the video sequence sampled at the temporal frequency $f_2 > f_1$ when the state of the resources (409 ; 513) of the communication apparatus is greater than a predetermined threshold.

14. A method according to claim 13, characterized in that the
15 simulation step is subdivided into several sub-steps:

- sampling (410 ; 514) images of the video sequence at the temporal frequency f_2 ,
- simulating coding (412 ; 516) of the sampled images,
- determining the quality (413 ; 517) of the coded images,
- 20 - comparing (414 ; 518) the determined quality with respect to a predetermined threshold,
- in case the threshold is exceeded, increasing (415 ; 519) the sampling temporal frequency of the images of the video sequence.

15. A method according to one of claims 1 to 14, characterized in
25 that the characteristics of the video sequence are the video activity and/or the quality of the video sequence.

16. A method according to one of claims 1 to 15, characterized in that the quality of the video sequence or of an image is expressed with respect to the signal to noise ratio of the coded image or video sequence.

30 17. A method according to one of claims 1 to 16, characterized in that the characteristics of the network are the bandwidth of the network.

18. A device for adapting a temporal frequency of a sequence of video images for the purpose of its transmission over a communication network, characterized in that images of the sequence having been sampled at a temporal frequency f_1 , the device comprises means for deciding as to the carrying out of a simulation of coding of images of the video sequence sampled at a temporal frequency $f_2 > f_1$, for the purpose of determining whether the sampling temporal frequency f_1 of the sequence can be increased, the decision being taken on the basis of at least one criterion relative to the resources of a communication apparatus able to perform the simulation step and/or on the basis of the evolution over time of the characteristics of the video sequence and/or of the network.

19. A device according to claim 18, characterized in that it comprises, means for storing in memory characteristics of the video sequence and/or characteristics of the network at a given time in relation with a sampling temporal frequency of the images of the video sequence.

20. A device according to one of claims 18 to 19, characterized in that it comprises:

- means for sampling images of the video sequence at a temporal frequency $f_0 > f_1$,
- means for coding of the sampled images,
- means for determining the quality of the coded images,
- means for comparing the determined quality with respect to a predetermined threshold,
- means for deciding as to a reduction of the sampling temporal frequency of the images of the video sequence from the frequency f_0 to the frequency f_1 , said means for deciding being adapted to take a decision depending on the result of the comparison.

21. A device according to one of claims 18 to 20, characterized in that it comprises means for comparing between the current characteristics presented by the video sequence and/or the network and the initial characteristics presented by the video sequence and/or the network when it was decided to use the temporal frequency f_1 for the sampling.

22. A device according to claim 21, characterized in that it comprises, means for deciding as to an increase of the sampling temporal frequency from f_1 to f_2 , said means for deciding being adapted to take a decision depending on the result of the comparison.

5 23. A device according to claim 22, characterized in that it comprises means for increasing the sampling temporal frequency which are adapted to increase the frequency, when the current characteristics of the video sequence and/or of the network have improved over time.

10 24. A device according to one of claims 18 to 23, characterized in that it comprises means for simulating coding of images of the video sequence sampled at the temporal frequency $f_2 > f_1$, said means are adapted to simulate the coding when the state of the resources of the communication apparatus is greater than a predetermined threshold.

15 25. A device according to claim 24, characterized in that said simulating means comprise:

- means for sampling images of the video sequence at the temporal frequency f_2 ,
- means for simulating coding of the sampled images,
- means for determining the quality of the coded images,
- 20 - means for comparing the determined quality with respect to a predetermined threshold,
- means for increasing the sampling temporal frequency of the images of the video sequence which are adapted to increase the temporal frequency in case of exceeding the threshold.

25 26. An information carrier readable by a computer system, possibly wholly or partly removable, in particular a CD-ROM or magnetic medium, such as a hard disk or a diskette, or a transmissible medium, such as an electrical or optical signal, this information carrier comprising instructions of a computer program characterized in that it enables the implementation of the method
30 according to any one of claims 1 to 17, when that program is loaded and executed by the computer system.

27. A computer program that can be loaded into a computer system, said program containing instructions enabling the implementation of the method according to any one of claims 1 to 17, when that program is loaded and executed by the computer system.

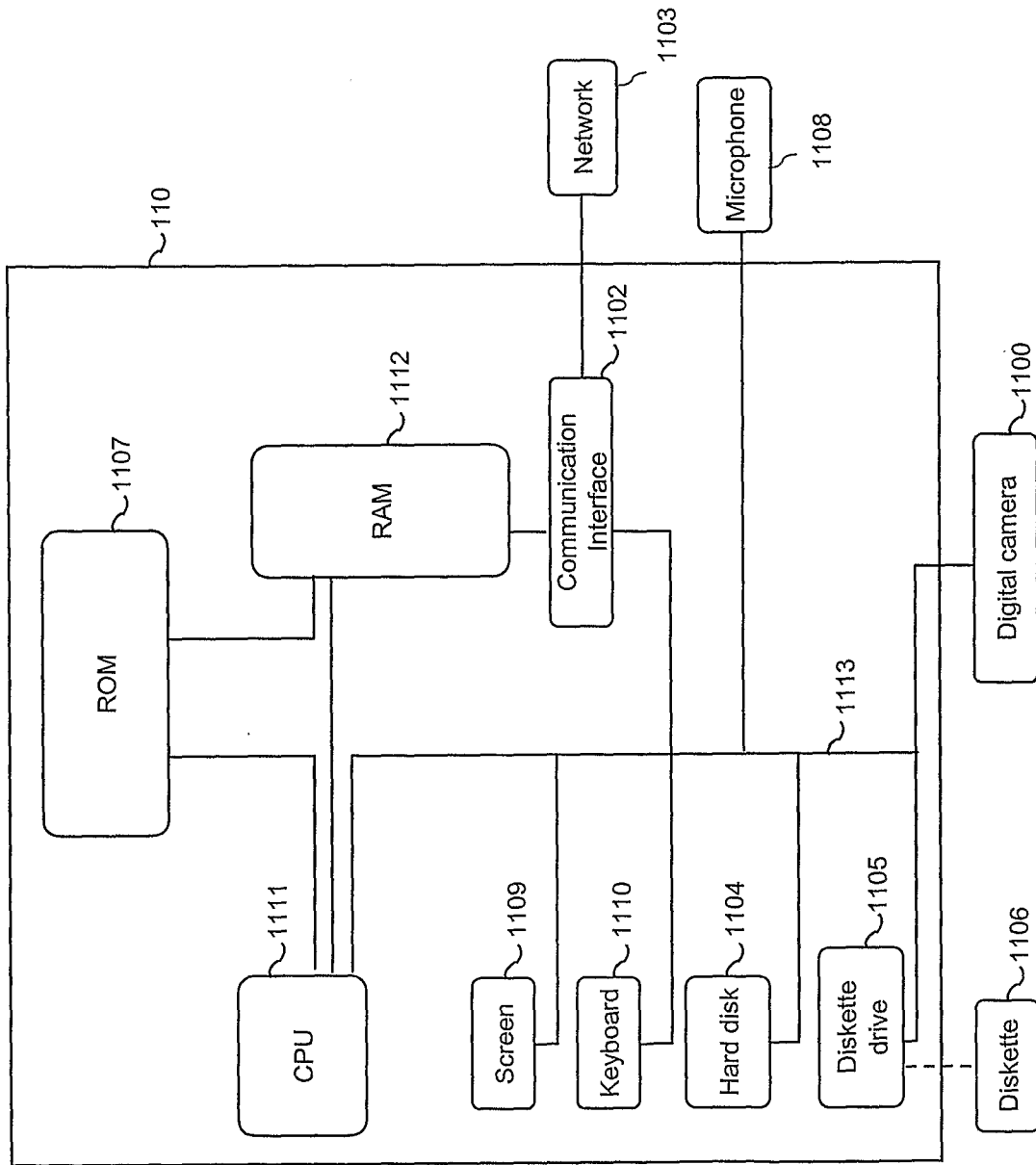


Fig. 1

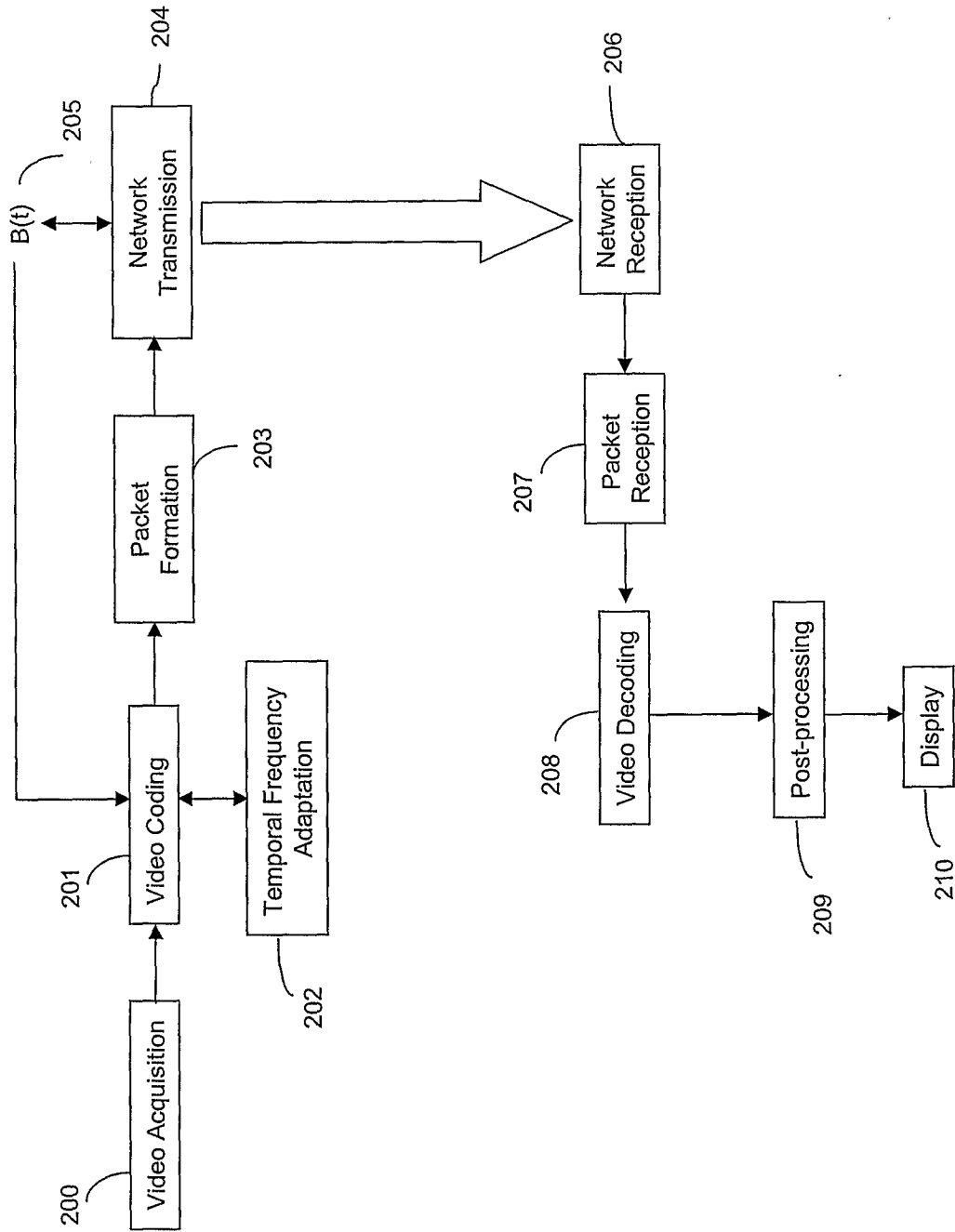


Fig. 2

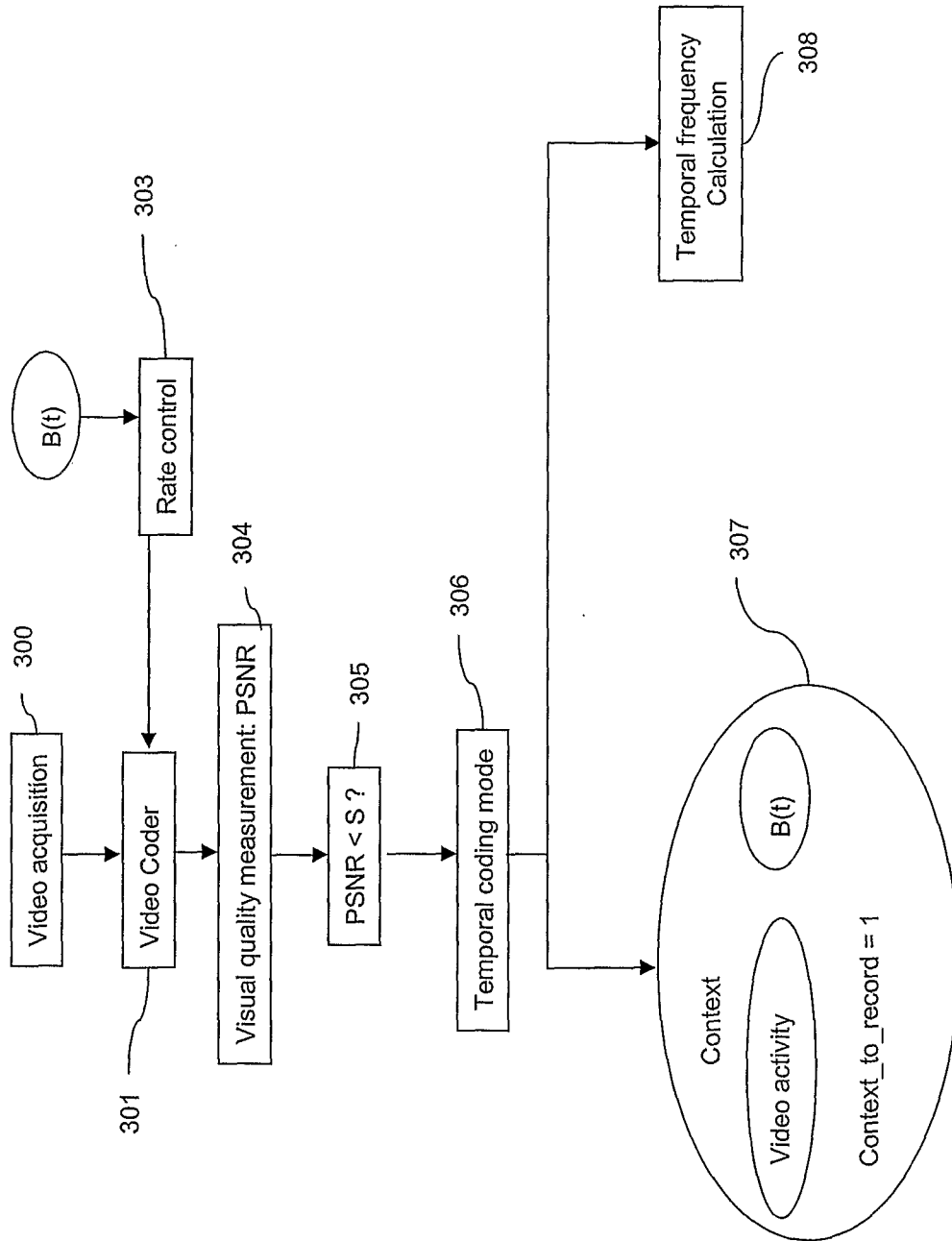


Fig. 3

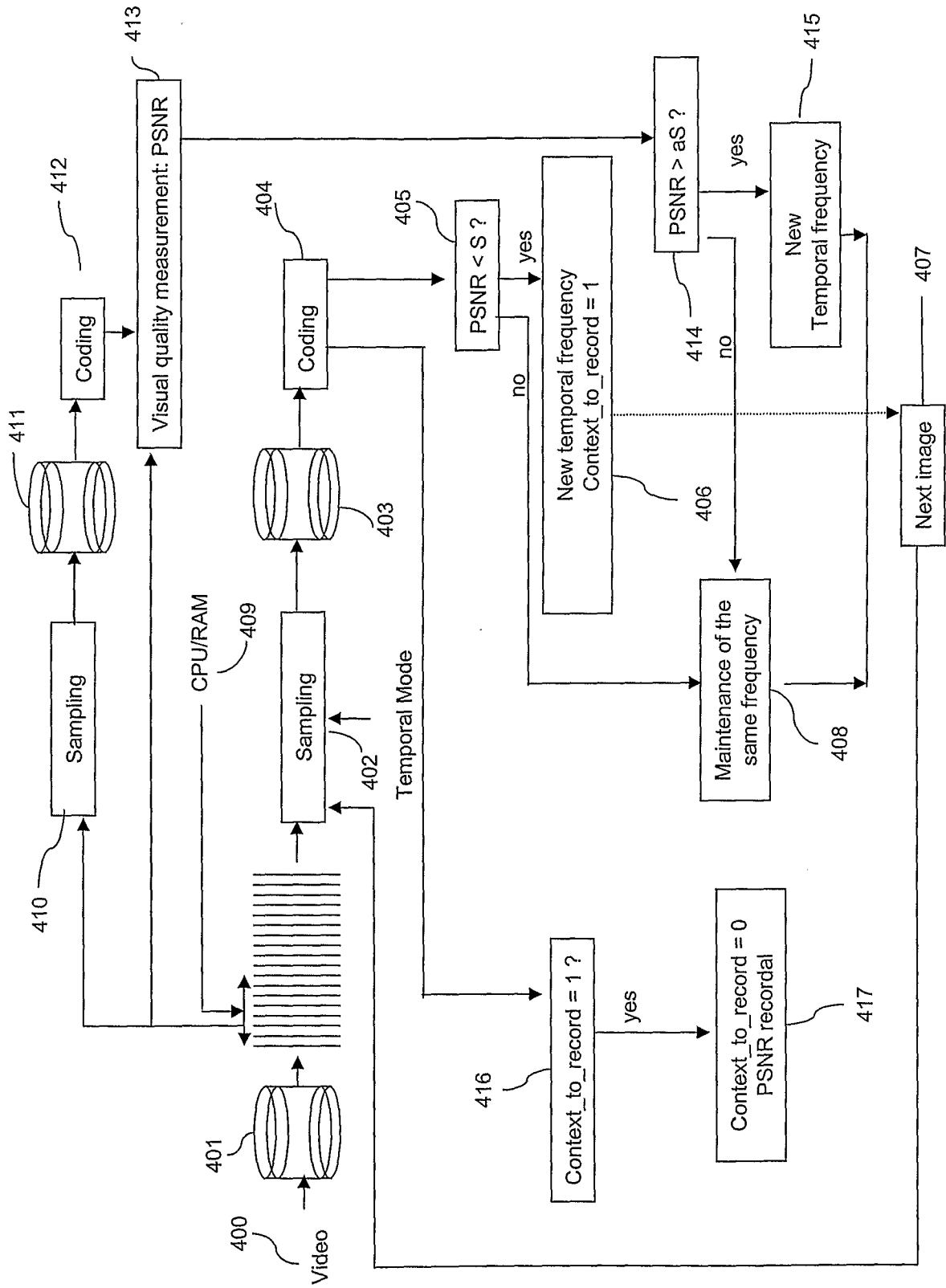


Fig. 4

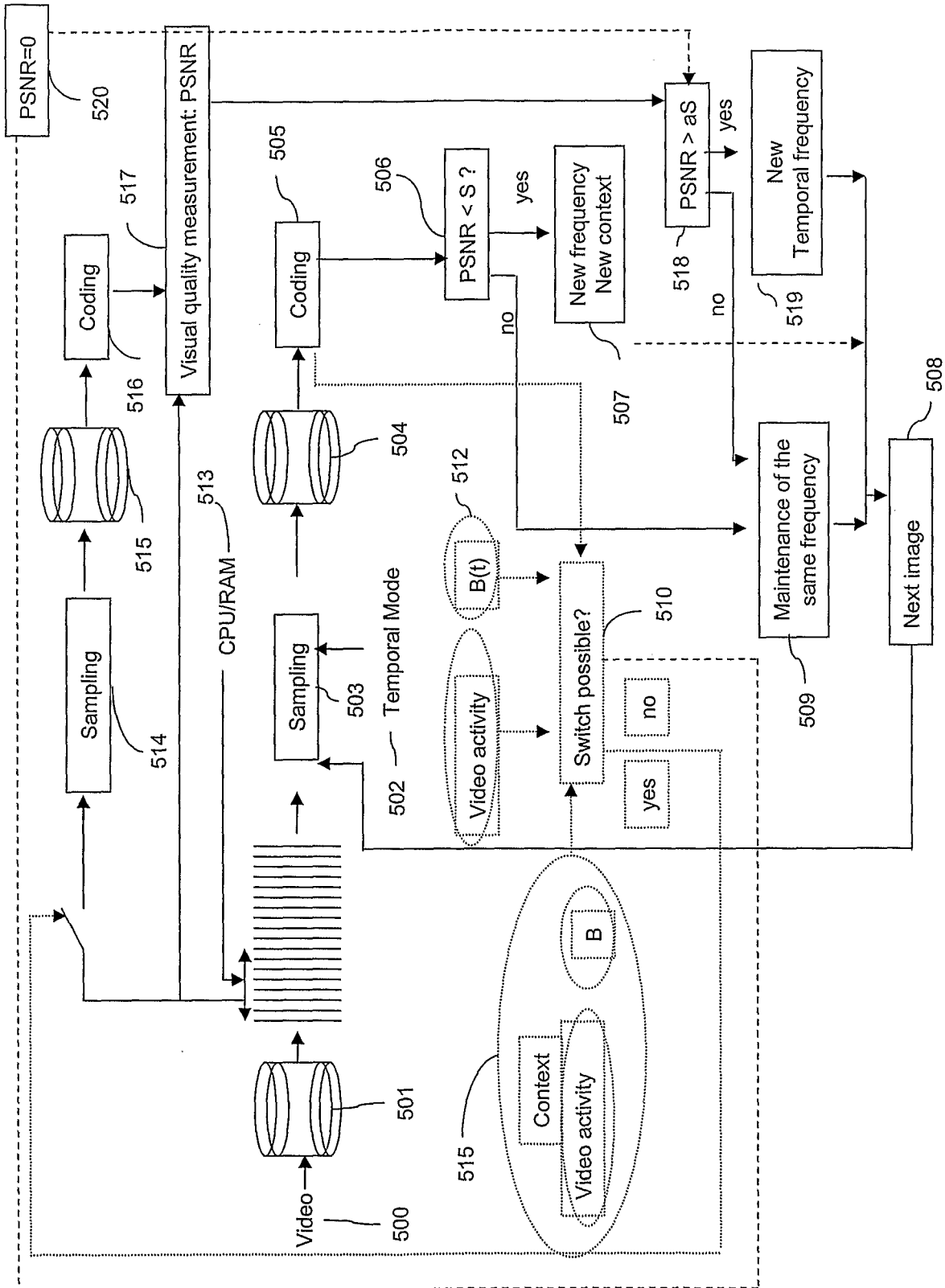


Fig. 5