

(19) World Intellectual Property Organization
International Bureau



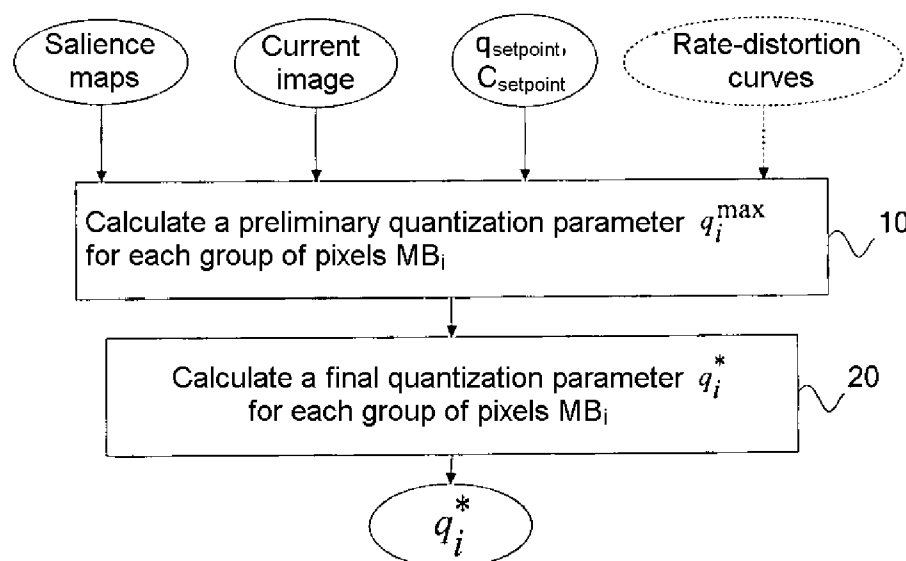
(43) International Publication Date
8 February 2007 (08.02.2007)

PCT

(10) International Publication Number
WO 2007/014850 A2

- (51) International Patent Classification: Not classified
- (21) International Application Number:
PCT/EP2006/064393
- (22) International Filing Date: 19 July 2006 (19.07.2006)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
0552345 28 July 2005 (28.07.2005) FR
- (71) Applicant (for all designated States except US): **THOMSON LICENSING** [FR/FR]; 46 quai Alphonse le Gallo, F-92100 Boulogne-Billancourt (FR).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **LE MEUR, Olivier** [FR/FR]; La Souche, F-35160 Talensac (FR). **THOREAU, Dominique** [FR/FR]; 39 rue du Reage, F-35510 Cesson-Sévigné (FR). **GUILLOT, Philippe** [FR/FR]; 15 rue Auguste Renoir, F-35770 Vern sur Seiche (FR).
- (74) Agents: **LE DANTEC, Claude** et al.; THOMSON, 46, quai Alphonse le Gallo, F-92100 Boulogne-Billancourt (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, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, 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, 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).
- 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: METHOD AND DEVICE FOR DETERMINING QUANTIZATION PARAMETERS IN AN IMAGE



(57) Abstract: The invention relates to a method for determining a quantization parameter for each group of pixels in an image. It comprises the following steps: - Calculating (10) a preliminary quantization parameter ($q_i^{\max}$) for each of the groups of pixels so as to minimize the variation in reconstruction quality between the groups when the preliminary quantization parameters are used for coding the image with a second number of bits ($C_{\min}$), which is less than the first number of bits (C_{setpoint}); and - Calculating (20) a final quantization parameter (q_i^*), which is less than or equal to the preliminary quantization parameter ($q_i^{\max}$), for each of the groups of pixels by reallocating the difference in bits between the first and second numbers of bits to the groups of pixels as a function of their content and their perceptual interest.



WO 2007/014850 A2

METHOD AND DEVICE FOR DETERMINING QUANTIZATION PARAMETERS IN AN IMAGE

1. Field of the Invention

5 The invention relates to a device and a method for determining quantization parameters for each group of pixels in an image, by using information about the perceptual interest of each of the groups of pixels. These quantization parameters are subsequently used for coding the image with a number of bits C_{setpoint} .

2. Prior Art

10 Selective compression makes it possible to locally vary the rate in an image by distributing a number of bits C_{setpoint} non-homogeneously in the image in order to improve the quality of the reconstructed image or reconstruction quality, i.e. of the coded then decoded image. In the case of 15 the MPEG2 and MPEG4 video coding standards, for example, the quantization parameter used for coding each image may vary in a given image from one block of pixels to another, for example from one macroblock (block of 16 pixels by 16 pixels) to another. In this way, the regions of interest 20 in the image can be coded with a reconstruction quality greater than that of the regions of non-interest by allocating them more bits, i.e. by associating a lower quantization parameter with them. In the case of a video conference application, the background is considered to be a region of non-interest whereas the speaker's head and shoulders are considered to be regions of 25 interest. Non-homogeneous allocation of the number of bits thus makes it possible to improve overall the perceived reconstruction quality by allocating more bits to the blocks of pixels belonging to the speaker's face or shoulders than to those belonging to the background.

30 It is therefore necessary to identify the regions of interest in the image and allocate them a greater number of bits for coding them. Identification of the regions of interest may be based for example on modelling the pre-attentive visual attention ("computational modelling of bottom-up visual selective attention"). Such modelling is described in an article by O. Le Meur et al. entitled "Performance assessment of a visual attention system entirely based

on a human vision modeling" published in the ICIP conference proceedings of October 2004, and in European Patent Application EP 1 544 792 published in June 2005.

Once the regions of interest are identified, as indicated previously, a greater
5 number of bits is allocated to them for coding of them in order to improve their reconstruction quality. The conventional solutions make it possible to allocate bits by locally adapting the quantization parameter as a function of the perceptual interest of the regions. However, they introduce numerous spatio-temporal visual defects particularly in the regions of non-interest, for example
10 in the background. These visual defects are problematic because they attract the observer's eye and therefore create fixation points in the image, which reduce the perceived reconstruction quality.

3. Summary of the Invention

15 It is an object of the invention to overcome some or all of these drawbacks. To this end the invention provides a method for determining a quantization parameter for each group of pixels in an image, making it possible to guarantee a minimum reconstruction quality over all the regions of the image, and particularly over the regions of non-interest.

20 The invention relates to a method for determining a quantization parameter for each group of pixels in an image, the quantization parameters being used for coding the image with a first number of bits (C_{setpoint}) corresponding to the number of bits necessary for coding the image with a setpoint quantization parameter (q_{setpoint}). The method comprises the following
25 steps:

- Calculating a preliminary quantization parameter (q_i^{max}) for each of the groups of pixels (MB_i) so as to minimize the variation in reconstruction quality between the groups when the preliminary quantization parameters are used for coding the image with a second number of
30 bits (C_{min}) which is less than the first number of bits (C_{setpoint}); and
- Calculating a final quantization parameter (q_i^*), which is less than or equal to the preliminary quantization parameter (q_i^{max}), for each of the groups of pixels by reallocating the difference in bits ($C_{\text{setpoint}} - C_{\text{min}}$)

between the first and second numbers of bits to the groups of pixels as a function of their content and their perceptual interest.

According to one particular characteristic, the difference in bits
5 $(C_{\text{setpoint}} - C_{\text{min}})$ between the first and second numbers of bits is reallocated to the groups of pixels proportionally to their perceptual interest, i.e. the number of bits reallocated and the value of the perceptual interest to vary in the same sense.

Preferably, the perceptual interest of a group of pixels is characterized
10 by a salience value calculated for this group of pixels.

According to one variant, the step of calculating the preliminary quantization parameter (q_i^{max}) is preceded by a step of associating a set of points with each of the groups of pixels, each point comprising a quantization parameter value, a number of bits necessary for coding the group of pixels
15 with the quantization parameter and an associated distortion value.

Advantageously, the step of calculating the preliminary quantization parameters (q_i^{max}) comprises the following steps:

- a. For each of the groups of pixels (MB_i) , calculating a distortion value $(d_v(i, q_{\text{init}}))$ corresponding to the coding of the group of pixels with an
20 initial quantization parameter (q_{init}) , which is greater than the setpoint quantization parameter (q_{setpoint}) ;
- b. For the image, calculating a current variance value (σ_v^2) of the distortion corresponding to the coding of the groups of pixels in the image with the initial quantization parameter (q_{init}) ;
- 25 c. Identifying a first set of groups of pixels (ES_1) corresponding to the N groups of pixels having the smallest distortion values and a second set of groups of pixels (ES_2) corresponding to the N groups of pixels having the largest distortion values, N being a predetermined integer;
- d. Decreasing the quantization parameters associated with the groups of
30 pixels of the first set (ES_1) by a value n and increasing the quantization parameters associated with the groups of pixels of the second set (ES_2) by a value n, the quantization parameters associated with each

of the groups of pixels other than those belonging to the first and second sets remaining unchanged, n being a predetermined integer;

- e. For the image, recalculating a new variance value (σ_v^2) of the distortion corresponding to the coding of the groups of pixels in the image with the quantization parameters resulting from step d, the current variance value becoming a preceding variance value and the new variance value becoming the current variance value; and
- f. Returning to step c if the absolute value of the difference between the current variance value and the preceding variance value is greater than a threshold (ϵ), otherwise, for each of the groups of pixels, assigning the value of the quantization parameter resulting from step d to the preliminary quantization parameter ($q_i^{\max}$) of this group.

Preferably, the integer N is the integer part of the product M times K, where K is the number of groups of pixels in the image and M is a number lying between 0 and 1.

Advantageously, the step of calculating the final quantization parameters (q_i^*) comprises the following steps:

- a. Calculating a parameter $\lambda(i,0)$, referred to as the initial rate-distortion parameter, for each of the groups of pixels (MB_i) according to the following formula:

$$\lambda(i,0) = \frac{D(i, q_i^{\max}) - D(i, q_i^{\max} + 1)}{R(i, q_i^{\max} + 1) - R(i, q_i^{\max})}$$

where: - $q_i^{\max}$ is the preliminary quantization parameter associated with the pixel group (MB_i) of index i;

- $D(i, A)$ is a perceptual distortion value corresponding to the coding of the group of pixels MB_i with the quantization parameter A; and

- $R(i, A)$ is the number of bits necessary for coding the group of pixels of index i with the quantization parameter A.

- b. Determining the maximum value of the rate-distortion parameters associated with each of the groups of pixels;

- 5 c. Decreasing the quantization parameter associated with the group of pixels of index i_0 having the maximum rate-distortion parameter, referred to as the identified group, by a value m , the quantization parameters associated with each of the groups of pixels other than the identified group remaining unchanged, m being a predetermined integer;
- 10 d. Calculating the difference between the number of bits necessary for coding the identified group with the quantization parameter of the identified group as calculated in step c and the number of bits necessary for coding the identified group with the quantization parameter of the group identified before step c, this difference being referred to as the number of supplementary bits;
- e. Subtracting the number of supplementary bits from the difference in bits ($C_{\text{setpoint}} - C_{\text{min}}$);
- 15 f. For each identified group, recalculating the rate-distortion value according to the following formula:

$$\lambda(i_0, k+1) = \frac{D(i_0, QP(i_0, k)) - D(i_0, QP(i_0, k+1))}{R(i_0, QP(i_0, k+1)) - R(i_0, QP(i_0, k))}$$

where: - $D(i_0, A)$ is the perceptual distortion value corresponding to the coding of the identified group with the quantization parameter A ;

20 - $R(i_0, A)$ is the number of bits necessary for coding the identified group with the quantization parameter A ; and

- $QP(i_0, k)$ is the parameter associated with the identified group at the preceding iteration k and $QP(i_0, k+1)$ is the quantization parameter as calculated at the iteration $k+1$.

- 25 g. Returning to step b if the difference in bits ($C_{\text{setpoint}} - C_{\text{min}}$) is positive, otherwise, for each of the groups of pixels, assigning the value of the quantization parameter resulting from step c to the final quantization parameter (q_i^*) of this group.

30 Advantageously, the perceptual distortion $D(i, q_i)$ associated with a group of pixels of index i , coded with the quantization parameter q_i , is derived

from a conventional distortion value $d_v(i, q_i)$ according to one of the following formulae:

- $D(i, q_i) = d_v(i, q_i) * s(i)$; or
- $D(i, q_i) = d_v(i, q_i) * s^p(i)$.

5 where - $s(i)$ represents a value characterizing the perceptual interest of the group of pixels of index i ;

- p is a positive integer; and
- $*$ is the multiplication operator.

10 The invention also relates to a device for determining a quantization parameter for each group of pixels in an image, the quantization parameters being used for coding the image with a first number of bits (C_{setpoint}) corresponding to the number of bits necessary for coding the image with a setpoint quantization parameter (q_{setpoint}). The device comprises the following
15 means:

- Means for calculating a preliminary quantization parameter (q_i^{max}) for each of the groups of pixels (MB_i) so as to minimize the variation in reconstruction quality between the groups when the preliminary quantization parameters are used for coding the image with a second
20 number of bits (C_{min}), which is less than the first number of bits (C_{setpoint}); and
- Means for calculating a final quantization parameter (q_i^*), which is less than or equal to the preliminary quantization parameter (q_i^{max}), for each of the groups of pixels by reallocating the difference in bits
25 ($C_{\text{setpoint}} - C_{\text{min}}$) between the first and second numbers of bits to the groups of pixels as a function of their content and their perceptual interest.

30 Lastly, the invention relates to a computer program product which comprises program code instructions for carrying out the steps of the method when the program is run on a computer.

4. Lists of the Figures

The invention will be more clearly understood and illustrated by means of entirely non-limiting examples of advantageous embodiments and implementations with reference to the appended figures, in which:

- 5 – Figure 1 illustrates a method for quantization parameter determination according to the invention;
- Figure 2 illustrates a rate-distortion curve associated with a group of pixels of index i ;
- Figure 3 represents two histograms of reconstruction quality
- 10 at two different iterations of the method according to the invention; and
- Figure 4 illustrates a device according to the invention.

5. Detailed Description of the Invention

15 The invention relates to a method for determining quantization parameters in an image, which uses information about the content of this image, more precisely information such as for example salience values characterizing the perceptual interest of the regions or groups of pixels (for example a block or macroblock) in the image. An image comprises pixels, with

20 each of which at least one luminance value is associated. The method may be applied to a single image or a sequence of a plurality of images. Each image is divided into K groups of pixels MB_i , $i \in [1, K]$. Each group of pixels may be a macroblock or, more generally, a block of I pixels by J pixels. Groups of pixels with any shape may equally be envisaged.

25 The invention requires knowledge of the reconstruction quality of an image or an MB_i . To this end, a plurality of reconstruction quality metrics may be used in order to estimate the distortion between a source image (or respectively a source MB_i) and the corresponding reconstructed image (or respectively the corresponding reconstructed MB_i), i.e. the source image

30 coded with a given quantization parameter then decoded (or respectively the coded then decoded MB_i). Among the techniques used for calculating a conventional distortion, the "sum of square errors" (SSE) is defined for an image (or respectively for an MB_i) as the sum over this image (or respectively

over this MB_i) of the squared differences between the luminance value associated with the pixel in the original image and the luminance value associated with the pixel having the same coordinates in the reconstructed image. Another calculation technique (MSE – "mean square error") is defined as being equal to the SSE divided by the number of samples used (i.e. the number of pixels). The invention also uses perceptual distortions, i.e. ones which take into account information about the content of the image and more precisely the perceptual interest of the MB_is in the image. These various perceptual distortions can be derived from the conventional distortions according to various formulae. For instance, a perceptual distortion for an MB_i referenced D(i,q_i) may be derived by the following formulae from a conventional distortion referenced d_v(i,q_i) :

$$D(i,q_i) = d_v(i,q_i) * s(i) \text{ or } D(i,q_i) = d_v(i,q_i) * s^p(i),$$

where: - s(i) represents a value lying between 0 and 1 characterizing the perceptual interest of the MB_i;

- p is a positive integer (for example p=2);
- q_i is the quantization parameter used for coding the image; and
- * is the multiplication operator.

These functions are examples, and other functions may be used. For an MB_i with a value s(i) equal to 1, the conventional distortion d_v(i) over this group of pixels is conserved. For an MB_i having a low value s(i), the conventional distortion is greatly reduced. Advantageously, this value s(i) is a salience value. In this case, a salience map is calculated for an image. A salience map is a two-dimensional topographical representation of the degree of salience of each pixel in the image. This map is normalized for example between 0 and 1, although it may also be normalized between 0 and 255. The salience map therefore provides a salience value S(x,y) per pixel (where (x,y) are the coordinates of a pixel in the image) which characterizes the perceptual interest of this image. The higher the value of S(x,y) is, the more pertinent the pixel of coordinates (x,y) is from a perceptual point of view. In order to obtain a salience value s(i) per MB_i, for example, the mean value of the salience values S(x,y) associated with each of the pixels of an MB_i is calculated. The

median value may also be used instead of the mean value in order to represent the MB_i . A salience map associated with a given image may be obtained by the method comprising the following steps:

- 5 - projection of the image in a psycho-visual colour space according to the luminance component in the case of a monochromatic image, and according to the luminance component and according to each of the chrominance components in the case of a colour image; it will be assumed below that the image being processed is a colour image;
- 10 - perceptual decomposition of the projected components (one luminance component and two chrominance components) into sub-bands in the frequency domain according to a human eye visibility threshold; the sub-bands are obtained by dividing up the frequency domain according to the radial spatial frequency and the orientation (angular selectivity); each sub-band may be considered as the neuronal image corresponding to a population
15 of visual cells attuned to an interval of spatial frequencies and a particular orientation;
- extraction of salient elements of the sub-bands relating to the luminance component and relating to each of the chrominance components, i.e. the most important information in the sub-bands;
- 20 - improvement of the contours of the salient elements in each sub-band relating to the luminance component and relating to each of the chrominance components;
- calculation of a salience map for the luminance on the basis of the improved contours of the salient elements of each sub-band relating to the
25 luminance component;
- calculation of a salience map for each of the chrominance components on the basis of the improved contours of the salient elements of each sub-band relating to the chrominance components; and
- generation of a final salience map on the basis of the luminance and
30 chrominance salience maps.

This method is described in European Patent Application EP 1 544 792 published in June 2005. The article by O. Le Meur et al. entitled "Performance assessment of a visual attention system entirely based on a human vision modeling" and published in the ICIP conference proceedings of October 2004
5 also gives details of the salience model. Other methods may be used for characterizing the perceptual interest of an MB_i .

According to a preferred embodiment, which is illustrated by Figure 1, the method is divided into the 2 steps referenced 10 and 20. The modules
10 represented in Figure 1 are functional units, which may or may not correspond to physically distinguishable units. For example, these modules or some of them may be grouped into a single component or constitute functionalities of the same software. *Conversely*, some modules may optionally be composed of separate physical entities. The object of the method is to achieve a
15 satisfactory compromise between the perceived reconstruction quality of the regions of interest in relation to the regions of non-interest in the image so as to improve the overall perceived reconstruction quality, without introducing other defects such as, for example, spatio-temporal defects in the case of a sequence of images. To this end, steps 10 and 20 of the method consist in
20 distributing a setpoint number of bits C_{setpoint} between the groups of pixels MB_i in an image, referred to as the current image, as a function of their perceptual interest, for example characterized by a salience value $s(i)$, and optionally by using rate-distortion curves associated with each MB_i . More precisely, they consist in associating a final quantization parameter q_i^* with each group of
25 pixels MB_i in the current image. In the case of an image sequence, steps 10 and 20 may be applied successively to all the images in the sequence.

The number C_{setpoint} is an input parameter of the method, and corresponds to the number of bits allocated to the current image for coding it. This number may, for example, be provided by the user of the method as a function of the
30 application. In the case of a sequence of images, this number of bits C_{setpoint} may also be determined by a conventional rate control method such as that defined in document ISO/IEC JTC1/SC29/WG11, *Test model 5*, 1993. This number may vary in particular as the function of the type of current image (for

example intra image, predicted image). Specifically, a larger number of bits is necessary for coding an intra type image (i.e. an image in a sequence of images which is coded without reference to the other images in the sequence) than for a coded image of the predicted type (i.e. an image in a sequence of images which is coded with reference to another image in the sequence). The setpoint number of bits C_{setpoint} corresponds to the number of bits necessary for coding the current image with a unique quantization parameter q_{setpoint} or with a different parameter for each MB_i . For the sake of clarity, a single parameter q_{setpoint} will be used here for describing the invention. The value of q_{setpoint} referenced in Figure 2 may be provided directly by the rate control method indicated above, or it may be determined on the basis of the value of C_{setpoint} and a rate-distortion curve associated with the current image, such as of the one represented by Figure 2 for an MB_i .

The generation of a rate-distortion curve consists in associating a distortion value and a coding cost (i.e. a number of bits) with each quantization parameter in a given interval (for example [0-31] for MPEG-2 and [0-51] for MPEG-4 AVC) specified, for example, by the coding standard. Such a curve may be associated with an image or with an MB_i . A rate-distortion curve may be provided by external means or alternatively generated as follows for an MB_i . The same technique can be used for generating the rate-distortion curve associated with an image. One technique for calculating the points of the rate-distortion curve consists in coding each MB_i with a plurality of quantization parameters (for example 1, 2, ..., q_i , q_i+1 , q_i+2 , ...) and in decoding it in order to generate a set of points. A quantization parameter q_i , a coding cost $R(i, q_i)$ and a conventional distortion value $d_v(i, q_i)$ correspond to each point. The coding cost $R(i, q_i)$ represents the number of bits necessary for coding an MB_i by using the quantization parameter q_i . As indicated previously, the value $d_v(i, q_i)$ is obtained by coding the MB_i with the quantization parameter q_i , decoding it and calculating the conventional distortion between this reconstructed MB_i and the source MB_i . In order to avoid too many coding operations, it is possible to code each MB_i with a reduced number of quantization parameters, for example one out of every two (i.e. 2, 4, ..., q_i , q_i+2 , q_i+4 , ...). The total curve as illustrated in Figure 2 is then

interpolated between the calculated points 30, for example by using a cubic interpolation or by spline curves. It is also feasible to construct only a portion of the curve around the quantization parameter q_{setpoint} . Such a method for constructing this curve consists in using the statistical properties of the images. Specifically, the images are generally modelled by a Gaussian model whose various parameters (i.e. mean, variance) are estimated directly on the basis of the current image or the images in the sequence.

Whatever the way in which the data have been obtained, they may be stored in correspondence tables ("look-up tables"), one per group of pixels MB_i , which associate a conventional distortion value $d_v(i, q_i)$ and a number of bits $R(i, q_i)$ with each quantization parameter q_i .

The input data may be provided to the method of the invention in the form of data files.

Referring again to Figure 1, step 10 consists in calculating a preliminary quantization parameter q_i^{max} for each MB_i so as to minimize the variation in reconstruction quality around an average reconstruction quality. It is carried out in four sub-steps: one initialization sub-step and three sub-steps applied iteratively until a first termination criterion.

To this end, an initial quantization parameter q_{init} is determined, which is uniform over all of the image and greater than the setpoint quantization parameter q_{setpoint} . For example, q_{init} is equal to $q_{\text{setpoint}} + T$ (with, for example, $T=3$).

As a variant, a starting setpoint C_{init} is determined on the basis of C_{setpoint} and other parameters such as, for example, the resolution of the images in the sequence and/or meta-data and/or the spatio-temporal activity of the images. A quantization parameter q_{init} is derived from the value of C_{init} and from the rate-distortion curve associated with the current image. The value C_{init} corresponds to the number of bits used for coding the current image with the quantization parameter q_{init} . The value of C_{init} which is less than the value of C_{setpoint} may also be set empirically to half the value of C_{setpoint} .

For each MB_i in the current image, the initialization sub-step consists in calculating the conventional i.e. non-perceptual distortion $d_v(i, q_{\text{init}})$ associated

with this MB_i coded with the quantization parameter q_{init} . The mean value $\bar{d}_v$ of the conventional distortion, as well as its variance σ_v^2 , are calculated over the current image in question according to the following formulae:

$$\bar{d}_v = \frac{1}{N} \sum_i d_v(i, q_{init}) \text{ and } \sigma_v^2 = \frac{1}{N} \sum_i (d_v(i, q_{init}) - \bar{d}_v)^2$$

- 5 The values $\bar{d}_v$ and σ_v^2 can thus be calculated directly on the basis of the current source image and the current reconstructed image.

The second sub-step consists in identifying a first set of groups of pixels corresponding to the N groups of pixels MB_i having the smallest conventional distortion values, said first set being referenced ES_1 in Figure 3, and a second

10 set of groups of pixels corresponding to the N groups of pixels MB_i having the greatest conventional distortion values, said second set being referenced ES_2 in Figure 3. N is defined for example by the formula $N=E[M*K]$, where $E[.]$ is the integer part function, * is the multiplication operator and M is a number lying between 0 and 1. A value of $M=0.1$ seems well suited.

- 15 The third sub-step consists in decreasing the quantization parameters associated with the groups of pixels MB_i of the first set ES_1 by a value n in order to increase their reconstruction quality, and in increasing the quantization parameters associated with the groups of pixels MB_i of the second set ES_2 by a value n in order to decrease their reconstruction quality,
- 20 n being a predetermined integer. A value of n equal to 1 seems well suited. The other MB_i s keep the same quantization parameter.

The last sub-step consists in recalculating the mean value of the conventional distortion $\bar{d}_v$ of the current image, as well as its variance σ_v^2 . If the absolute value of the difference between the variance value calculated at the preceding

25 iteration and the current value is less than a threshold ε (for example, $\varepsilon = 10^{-6}$), the distribution of bits is terminated. Otherwise, the method returns to the first sub-step in order to continue the distribution.

- For the current image, this step 10 makes it possible to have a reconstruction quality which is less than the setpoint reconstruction quality but is more
- 30 homogeneous. Figure 3 represents two histograms of reconstruction quality at two different iterations of step 10. At the second iteration, the reconstruction

quality of the macroblocks belonging to the first set ES_1 has increased and the reconstruction quality of the macroblocks belonging to the second set ES_2 has decreased to approach the average reconstruction quality. The setpoint reconstruction quality is the reconstruction quality calculated between the current source image and the current reconstructed image, i.e. the current source image coded with the quantization parameter q_{setpoint} then decoded. Specifically, the overall quality over an image is maximal when the local quality is identical and the overall quality drops greatly when the quality drops locally. This step makes it possible for a preliminary quantization parameter

10 q_i^{max} , which corresponds to the last quantization parameter calculated, to be associated with each MB_i in the current image. A new rate C_{min} is calculated, which takes into account the preliminary quantization parameters associated with each of the MB_i : $C_{\text{min}} = \sum_{i=1}^K R(i, q_i^{\text{max}})$.

Step 20 consists in calculating a final quantization parameter q_i^* for each MB_i by reallocating the remaining bits ΔC , i.e. the difference in bits between C_{setpoint} and C_{min} , as a function in particular of the perceptual interest of the MB_i s, a greater number of bits being reallocated to the MB_i s whose perceptual interest is highest. The reallocation of bits is carried out according to three sub-steps: one initialization sub-step and two sub-steps applied iteratively until a second termination criterion.

The first sub-step, referred to as initialization, consists in calculating an initial rate-distortion parameter $\lambda(i,0)$ for each MB_i in the following way on the basis of the rate-distortion curves calculated previously and the salience maps:

$$\lambda(i,0) = \frac{d_{vi}(i, q_i^{\text{max}}) - d_{vi}(i, q_i^{\text{max}} + 1)}{R(i, q_i^{\text{max}} + 1) - R(i, q_i^{\text{max}})}$$

25 where $\lambda(i,k)$ represents the slope of the rate-perceptual distortion curve at a given point on this curve as calculated at iteration k. The rate-perceptual distortion curve is derived directly from the rate-distortion curve provided to the method as input and one of the formulae adopted for calculating a perceptual distortion (for example $D(i) = d_v(i) * s(i)$). The higher the parameter $\lambda(i,k)$ is, the more the distortion decreases strongly with a small extra cost of bits.

Let $QP(i, k)$ be the quantization parameter associated with MB_i at iteration k . During an iteration k , the second sub-step consists in determining the maximum value $\lambda_{\max}(k)$ among all the parameters $\lambda(i, k)$ calculated:

$\lambda_{\max}(k) = \max_i \lambda(i, k)$. A quantization parameter reduced by an integer value m

5 in relation to the preceding iteration will be associated with the group of pixels MB_{i_0} of index i_0 corresponding to $\lambda_{\max}(k)$, i.e. $QP(i_0, k+1) = QP(i_0, k) - m$.

Preferably, m is equal to 1. The other $MB_{i, i \neq i_0}$ keep their quantization parameter, i.e. $QP(i, k+1) = QP(i, k)$. Furthermore, the number of bits to be reallocated is updated in the following way:

10 $\Delta C = \Delta C - (R(i_0, QP(i_0, k+1)) - R(i_0, QP(i_0, k)))$.

During iteration k , the third sub-step consists in recalculating the rate-distortion parameter associated with the MB_{i_0} whose quantization parameter has just been modified, in the following way:

$$\lambda(i_0, k+1) = \frac{D(i_0, QP(i_0, k)) - D(i_0, QP(i_0, k+1))}{R(i_0, QP(i_0, k+1)) - R(i_0, QP(i_0, k))}$$

15 The rate-distortion parameters associated with the other $MB_{i, i \neq i_0}$ remain unchanged, i.e. $\lambda(i, k+1) = \lambda(i, k)$. So long as ΔC is positive, the method returns to the second sub-step. This step 20 makes it possible for a final quantization parameter q_i^* , which corresponds to the last quantization parameter calculated, to be associated with each MB_i in the current image.

20

The present invention also relates to a device, referenced 40 in Figure 4, which implements the method previously described. Only the essential elements of the device are represented in Figure 4. The device 40 comprises in particular a random-access memory 42 (RAM or similar component), a
25 read-only memory 43 (hard disk or similar component), a processing unit 44 such as a microprocessor or a similar component, an input/output interface 45 and a man-machine interface 46. These elements are connected together by an address and data bus 41. The read-only memory 43 contains in particular the algorithms carrying out steps 10 and 20 of the method according to the
30 invention. It may also contain the algorithms for obtaining the input

parameters of the method such as, for example, a rate control algorithm, an algorithm for generating the salience maps as well as an algorithm for coding/decoding the images. Upon start-up, the processing unit 44 loads and executes the instructions of these algorithms. The random-access memory 42
5 comprises in particular the operating programs of the processing unit 44, which are loaded upon start-up of the apparatus, as well as the images to be processed. The purpose of the input/output interface 45 is to receive the input signal (i.e. the sequence of source images and optionally the input parameters such as the setpoint number of bits C_{setpoint} , the associated quantization
10 parameter q_{setpoint} , the salience maps, the rate-distortion curves) and to deliver the quantization parameters determined according to steps 10 and 20 of the method of the invention. The man-machine interface 46 of the device allows the user to interrupt the processing. The results of the determination of the quantization parameters in each image are stored in random-access memory
15 then transferred to read-only memory in order to be archived with a view to subsequent processing operations, for example coding the images with these quantization parameters. The man-machine interface 46 comprises in particular a control panel and a display screen.

20 Of course, the invention is not limited to the exemplary embodiments mentioned above. In particular, the person skilled in the art may apply any variant in the embodiments explained and combine them in order to benefit from their various advantages. For example, perceptual distortion metrics other than those described previously may be used. Likewise, other methods
25 may be used in order to determine the rate-distortion curves associated with each of the groups of pixels MB_i . Instead of determining values C_{setpoint} and C_{init} directly, for example by using a rate control method, it is moreover possible to directly use quantization parameters q_{setpoint} and q_{init} proposed for example by a user as a function of the application. The values C_{setpoint} and C_{init}
30 then correspond to the number of bits used for coding the current image, respectively with q_{setpoint} and q_{init} . According to the invention, furthermore, it is not necessary to construct rate-distortion maps. In fact, a group of pixels MB_i may be coded with a given quantization parameter each time it is essential to know the number of bits necessary for coding this MB_i with the given

quantization parameter and the associated distortion. The input data of the method according to the invention, i.e. the setpoint rate C_{setpoint} , optionally q_{setpoint} , the salience maps and optionally the rate-distortion curves, may be provided by methods other than those described previously.

Claims

- 5 1. Method for determining a quantization parameter for each group of pixels in an image, the said quantization parameters being used for coding the said image with a first number of bits (C_{setpoint}) corresponding to the number of bits necessary for coding the said image with a setpoint quantization parameter (q_{setpoint}), characterized in that it comprises the following steps:
- 10 – Calculating (10) a preliminary quantization parameter (q_i^{max}) for each of the said groups of pixels (MB_i) so as to minimize the variation in reconstruction quality between the said groups when the said preliminary quantization parameters are used for coding the said image with a second number of bits (C_{min}), which is less than the first
- 15 number of bits (C_{setpoint}); and
- Calculating (20) a final quantization parameter (q_i^*), which is less than or equal to the preliminary quantization parameter (q_i^{max}), for each of the said groups of pixels by reallocating the difference in bits ($C_{\text{setpoint}} - C_{\text{min}}$) between the first and second numbers of bits to the said groups
- 20 of pixels as a function of their content and their perceptual interest.
2. Method according to Claim 1, characterized in that the said difference in bits ($C_{\text{setpoint}} - C_{\text{min}}$) between the first and second numbers of bits is reallocated to the said groups of pixels proportionally to their perceptual interest.
- 25 3. Method according to one of Claims 1 and 2, characterized in that the perceptual interest of a group of pixels is characterized by a salience value calculated for this group of pixels.
- 30 4. Method according to one of Claims 1 to 3, characterized in that the step of calculating the said preliminary quantization parameter (q_i^{max}) is preceded by a step of associating a set of points (30) with each of the said groups of pixels,

each point comprising a quantization parameter value, a number of bits necessary for coding the said group of pixels with the said quantization parameter and an associated distortion value.

5 5. Method according to one of Claims 1 to 4, characterized in that the step of calculating the said preliminary quantization parameters ($q_i^{\max}$) comprises the following steps:

- 10 a. For each of the said groups of pixels (MB_i), calculating a distortion value ($d_v(i, q_{init})$) corresponding to the coding of the said group of pixels with an initial quantization parameter (q_{init}), which is greater than the setpoint quantization parameter ($q_{setpoint}$);
- b. For the said image, calculating a current variance value (σ_v^2) of the distortion corresponding to the coding of the said groups of pixels in the said image with the said initial quantization parameter (q_{init});
- 15 c. Identifying a first set of groups of pixels (ES_1) corresponding to the N groups of pixels having the smallest distortion values and a second set of groups of pixels (ES_2) corresponding to the N groups of pixels having the largest distortion values, N being a predetermined integer;
- 20 d. decreasing the quantization parameters associated with the said groups of pixels of the said first set (ES_1) by a value n and increasing the quantization parameters associated with the said groups of pixels of the said second set (ES_2) by a value n, the quantization parameters associated with each of the groups of pixels other than those belonging to the said first and second sets remaining unchanged, n being a
- 25 predetermined integer;
- e. For the said image, recalculating a new variance value (σ_v^2) of the distortion corresponding to the coding of the said groups of pixels in the said image with the quantization parameters resulting from step d, the current variance value becoming a preceding variance value and the
- 30 new variance value becoming the current variance value; and
- f. Returning to step c if the absolute value of the difference between the current variance value and the preceding variance value is greater than a threshold (ϵ), otherwise, for each of the said groups of pixels,

assigning the value of the quantization parameter resulting from step d to the preliminary quantization parameter ($q_i^{\max}$) of this group.

6. Method according to Claim 5, characterized in that the integer N is the integer part of the product M times K, where K is the number of groups of pixels in the said image and M is a number lying between 0 and 1.

7. Method according to Claim 6, characterized in that $M=0.1$, $n=1$ and $\varepsilon=10^{-6}$.

8. Method according to one of Claims 1 to 7, characterized in that the step of calculating the said final quantization parameters (q_i^*) comprises the following steps:

a. Calculating a parameter $\lambda(i,0)$, referred to as the initial rate-distortion parameter, for each of the said groups of pixels (MB_i) according to the following formula:

$$\lambda(i,0) = \frac{D(i, q_i^{\max}) - D(i, q_i^{\max} + 1)}{R(i, q_i^{\max} + 1) - R(i, q_i^{\max})}$$

where: - $q_i^{\max}$ is the preliminary quantization parameter associated with the pixel group (MB_i) of index i;

- $D(i, A)$ is a perceptual distortion value corresponding to the coding of the said group of pixels MB_i with the quantization parameter A; and

- $R(i, A)$ is the number of bits necessary for coding the said group of pixels of index i with the quantization parameter A.

b. Determining the maximum value of the said rate-distortion parameters associated with each of the said groups of pixels;

c. Decreasing the quantization parameter associated with the group of pixels of index i_0 having the said maximum rate-distortion parameter, referred to as the identified group, by a value m, the quantization parameters associated with each of the groups of pixels other than the identified group remaining unchanged, m being a predetermined integer;

- d. Calculating the difference between the number of bits necessary for coding the said identified group with the quantization parameter of the identified group as calculated in step c and the number of bits necessary for coding the identified group with the quantization parameter of the group identified before step c, this difference being referred to as the number of supplementary bits;
- e. Subtracting the said number of supplementary bits from the said difference in bits ($C_{\text{setpoint}} - C_{\text{min}}$);
- f. For each identified group, recalculating the said rate-distortion value according to the following formula:

$$\lambda(i_0, k+1) = \frac{D(i_0, QP(i_0, k)) - D(i_0, QP(i_0, k+1))}{R(i_0, QP(i_0, k+1)) - R(i_0, QP(i_0, k))}$$

where: - $D(i_0, A)$ is the perceptual distortion value corresponding to the coding of the said identified group with the quantization parameter A;

- $R(i_0, A)$ is the number of bits necessary for coding the said identified group with the quantization parameter A; and

- $QP(i_0, k)$ is the parameter associated with the said identified group at the preceding iteration k and $QP(i_0, k+1)$ is the quantization parameter as calculated at the iteration k+1.

- g. Returning to step b if the said difference in bits ($C_{\text{setpoint}} - C_{\text{min}}$) is positive, otherwise, for each of the said groups of pixels, assigning the value of the quantization parameter resulting from step c to the final quantization parameter (q_i^*) of this group.

9. Method according to Claim 8, characterized in that the perceptual distortion $D(i, q_i)$ associated with a group of pixels of index i, coded with the quantization parameter q_i , is derived from a conventional distortion value $d_v(i, q_i)$ according to one of the following formulae:

- $D(i, q_i) = d_v(i, q_i) * s(i)$; or
- $D(i, q_i) = d_v(i, q_i) * s^p(i)$.

where - $s(i)$ represents a value characterizing the perceptual interest of the said group of pixels of index i;

- p is a positive integer; and
- * is the multiplication operator.

10. Method according to Claim 9, characterized in that $m=1$ and $p=2$.

5

11. Device for determining a quantization parameter for each group of pixels in an image, the said quantization parameters being used for coding the said image with a first number of bits (C_{setpoint}) corresponding to the number of bits necessary for coding the said image with a setpoint quantization parameter (q_{setpoint}), characterized in that it comprises the following means:

10

- Means for calculating (40) a preliminary quantization parameter (q_i^{max}) for each of the said groups of pixels (MB_i) so as to minimize the variation in reconstruction quality between the said groups when the said preliminary quantization parameters are used for coding the said image with a second number of bits (C_{min}), which is less than the first number of bits (C_{setpoint}); and
- Means for calculating (40) a final quantization parameter (q_i^*), which is less than or equal to the preliminary quantization parameter (q_i^{max}), for each of the said groups of pixels by reallocating the difference in bits ($C_{\text{setpoint}} - C_{\text{min}}$) between the first and second numbers of bits to the said groups of pixels as a function of their content and their perceptual interest.

15

20

12. Computer program product, characterized in that it comprises program code instructions for carrying out the steps of the method according to one of Claims 1 to 10, when the said program is run on a computer.

25

1/3

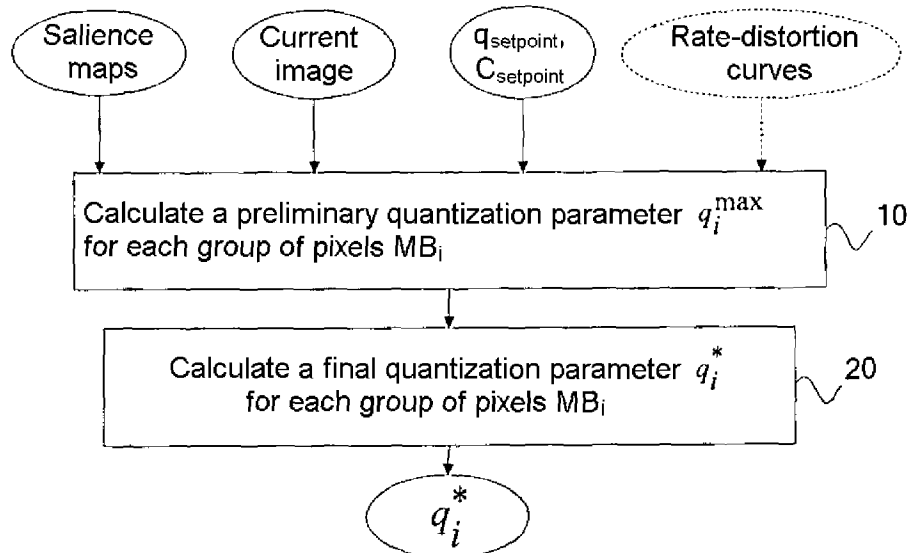


FIG. 1

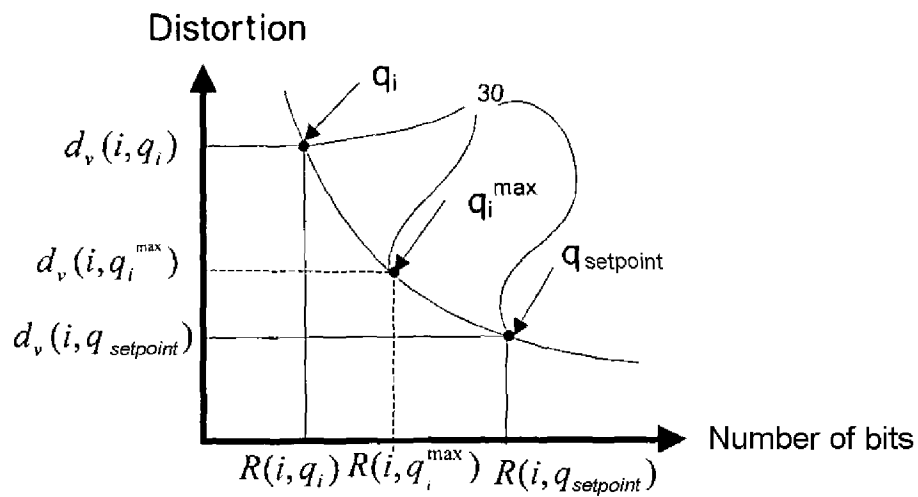


FIG. 2

2/3

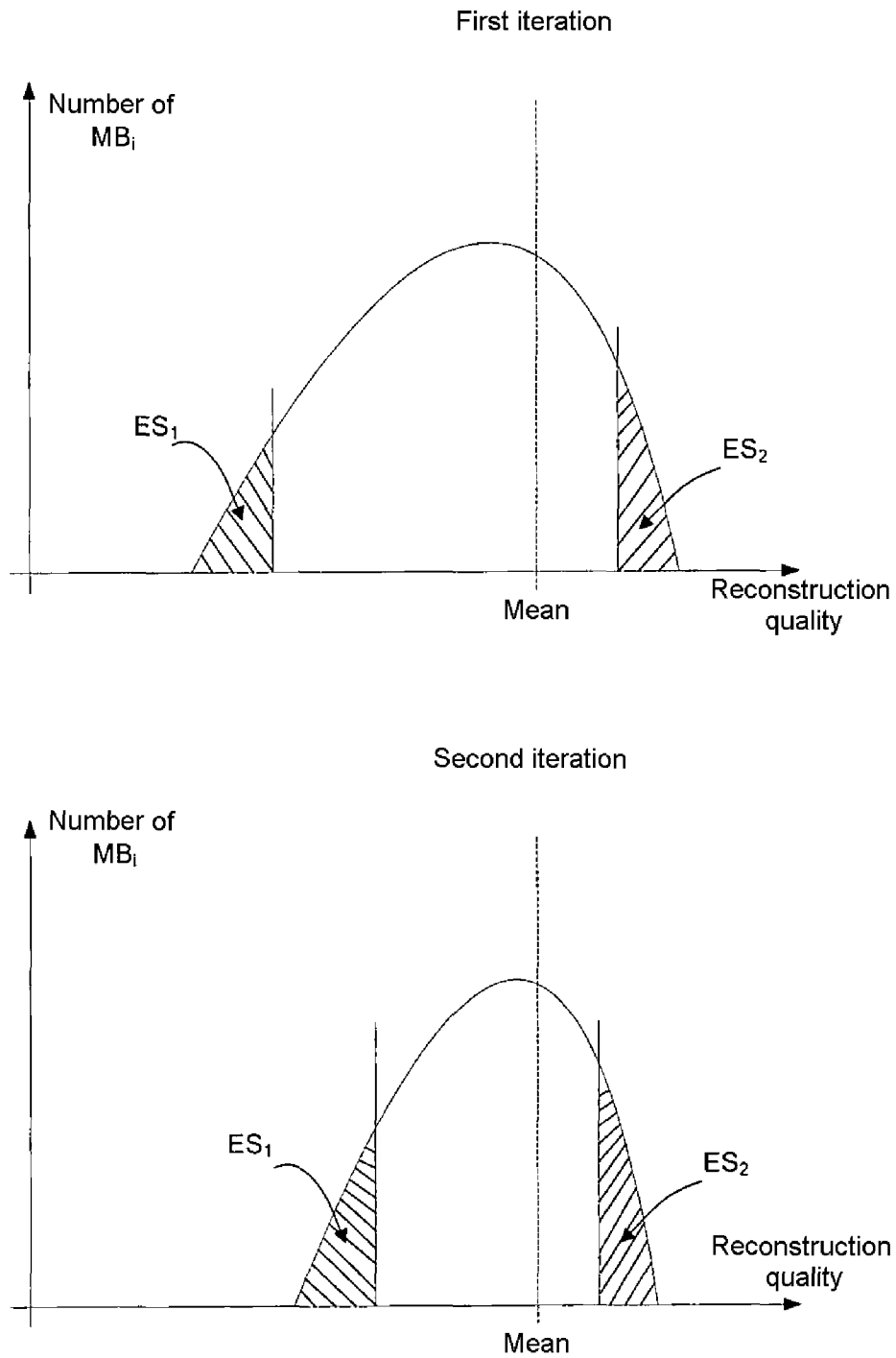


FIG. 3

3/3

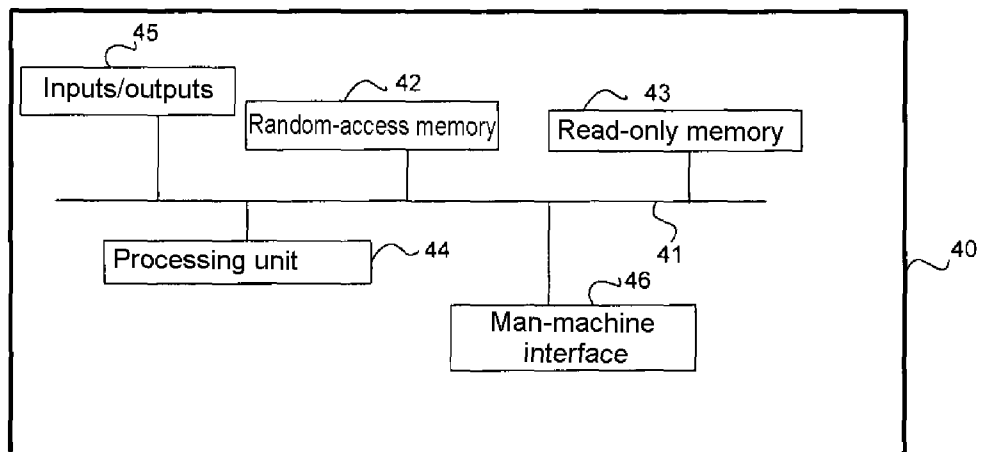


FIG. 4