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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 July 2007 (26.07.2007)

PCT

(10) International Publication Number
WO 2007/084475 A2

(51) International Patent Classification:
H04N 7/26 (2006.01) *H04N 7/50* (2006.01)
H04N 7/36 (2006.01) *H04N 7/68* (2006.01)

(21) International Application Number: PCT/US2007/001067

(22) International Filing Date: 16 January 2007 (16.01.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/759,411 17 January 2006 (17.01.2006) US

(71) Applicant *(for all designated States except US)*: **THOMSON LICENSING** [FR/FR]; 46, Quai A. Le Gallo, F-92100 Boulogne Billancourt (FR).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **YANG, Hua** [CN/US]; 3711 Quail Ridge Drive, Plainsboro, New Jersey 08536 (US). **BOYCE, Jill, MacDonald** [US/US]; 3 Brandywine Court, Manalapan, New Jersey 07726 (US).

(74) Agents: **LAKS, Joseph, J.** et al.; Thomson Licensing LLC, 2 Independence Way, Suite 200, Princeton, New Jersey 08540 (US).

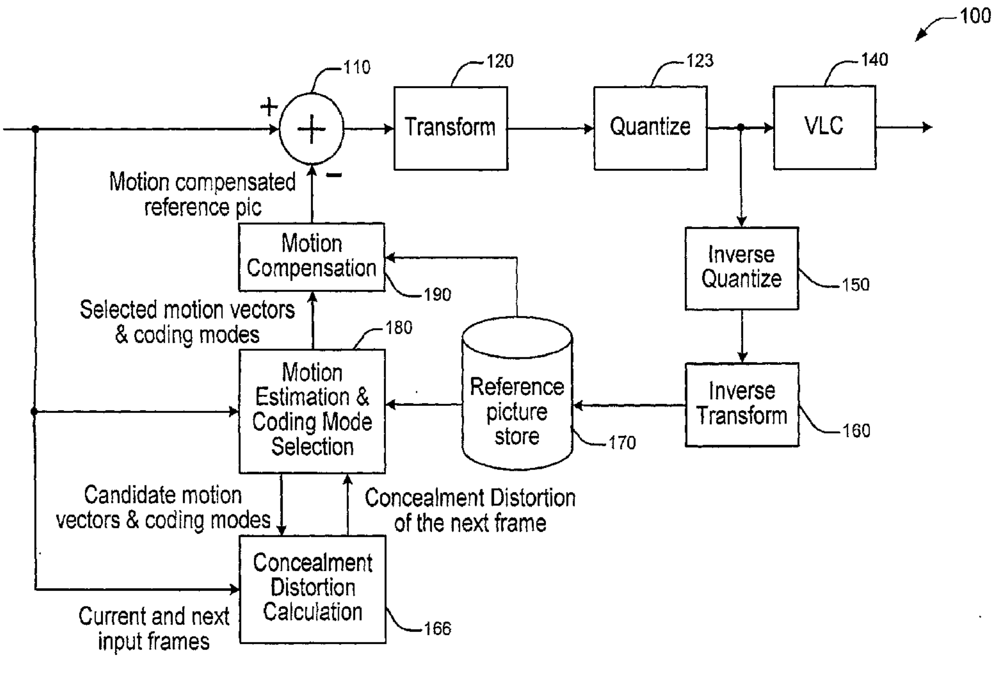
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, 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, 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: METHODS AND APPARATUS FOR LOW COMPLEXITY ERROR RESILIENT MOTION ESTIMATION AND CODING MODE SELECTION



(57) Abstract: There are methods and apparatus for low complexity error resilient motion estimation and coding mode selection. A video encoder includes an encoder (100) for encoding a video sequence using weighted error concealment distortion to account for packet loss impact on a video quality of the video sequence at a corresponding decoder.

WO 2007/084475 A2

METHODS AND APPARATUS FOR LOW COMPLEXITY ERROR RESILIENT MOTION ESTIMATION AND CODING MODE SELECTION

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of U.S. Provisional Application Serial No. 60/759,411, filed 17 January, 2006, which is incorporated by reference herein in its entirety.

FIELD OF THE INVENTION

The present invention relates generally to video encoding and decoding and, more particularly, to methods and apparatus for low complexity error resilient motion estimation and coding mode selection.

BACKGROUND OF THE INVENTION

In video streaming, a known practice for error resilient video coding is to apply end-to-end distortion, rather than source coding distortion, in rate distortion (RD) optimized motion estimation and coding mode selection. However, the end-to-end distortion metric is not necessarily the best metric to accurately measure the subjective or perceptual reconstructed video quality at the decoder. Furthermore, end-to-end estimation itself is a highly challenging and complicated task, which entails intensive computation and storage complexity.

In a typical video encoder such as those conforming to the H.261, H263, H.264, MPEG-1, MPEG-2, or MPEG-4 video coding standards, a video frame is divided into macroblocks, and each macroblock (MB) may be coded in one of several coding modes. In an INTER (inter-coding) mode, a motion vector is first found that points to the best matching block in a previously coded frame, then the difference between this MB and its best matching block is coded. In an INTRA (intra-coding) mode, the MB is either coded directly or predicted from some previously coded pixels in the same frame (called intra prediction). Among all possible modes, the encoder finally chooses the optimal one, based on a mode decision criterion.

In rate-distortion (RD) optimized motion estimation and mode selection, both the motion estimation and mode decision criteria are a weighted sum of the distortion of the decoded MB and the number of bits used. When the underlying transmission network is not reliable, part of the transmitted video bit stream may be lost. Therefore, besides coding efficiency, error resilience is another critical concern in video streaming applications. To improve the error resilience of video coding, a commonly used approach is to select the motion vector and/or MB coding mode of a certain block/MB via end-to-end distortion based RD optimization. However, the end-to-end distortion based approach has several drawbacks.

First, the statistically defined end-to-end mean squared error (MSE) distortion is not always a good metric to truthfully represent the subjective/perceptual decoder reconstructed video quality. In practice, the subjective quality of a reconstructed video sequence is more heavily dependent on its maximum distortion than on the average distortion over all the frames. This insight has already been widely applied in rate control of source video coding, where the criterion of minimum maximum distortion or minimum quality variation is used for optimal bit allocation within a video frame or sequence.

Second, accurately estimating the expected end-to-end distortion itself is a challenging and complicated task. Existing distortion estimation solutions primarily suggest either block-based or pixel-based approaches. A pixel-based ROPE estimation method was proposed, and then subsequently improved upon to extend for sub-pel prediction (hereinafter, the initial method and improved variation thereof collectively referred to as "ROPE"). ROPE achieves accurate distortion estimation at the price of a significant amount of increased computation and storage complexity, since it requires tracking channel distortion for each single pixel. Block-based estimation greatly reduces the incurred complexity, as channel distortion is tracked for each coding block. However, its estimation accuracy is also inevitably compromised. Subsequent to the initial proposal for block-based estimation, a new solution was proposed that renders several important extensions to account for sub-pel prediction, deblocking filtering, and more complicated error concealment schemes. However, even for the block-based approach, it still has to recursively keep track of and calculate various channel distortion quantities from frame to frame and, thus, its total amount of implementation cost is still significant.

Third, in spite of the large variety of effective error concealment schemes previously proposed, for convenience and simplicity, almost all the existing end-to-end distortion based error resilient video coding schemes assume the same simplest "copy from previous frame" error concealment scheme, which is far from satisfactory in practical video streaming applications. Schemes have been proposed in which a more effective "motion copy" error concealment scheme was assumed in the optimized motion estimation and mode selection schemes. However, the schemes still require complicated block-based end-to-end distortion estimation.

SUMMARY OF THE INVENTION

These and other drawbacks and disadvantages of the prior art are addressed by the present invention, which is directed to methods and apparatus for low complexity error resilient motion estimation and coding mode selection.

According to an aspect of the present invention, there is provided a video encoder. The video encoder includes an encoder for encoding a video sequence using weighted error concealment distortion to account for packet loss impact on a video quality of the video sequence at a corresponding decoder.

According to another aspect of the present invention, there is provided a video encoder. The video encoder includes an encoder for encoding a video sequence by modeling error concealment distortion using a current picture to be encoded in the video sequence and a next un-coded picture in the video sequence to optimize a received video quality of the video sequence at a corresponding decoder when the video sequence is received subject to network induced errors.

According to yet another aspect of the present invention, there is provided a video encoding method. The method includes encoding a video sequence using weighted error concealment distortion to account for packet loss impact on a video quality of the video sequence at a corresponding decoder.

According to still another aspect of the present invention, there is provided a video encoding method. The method includes encoding a video sequence by modeling error concealment distortion using a current picture to be encoded in the video sequence and a next un-coded picture in the video sequence to optimize a received video quality of the video sequence at a corresponding decoder when the video sequence is received subject to network induced errors.

According to a further aspect of the present invention, there is provided a video encoder. The video encoder includes a concealment distortion calculator for calculating a weighted error concealment distortion for a block of a particular picture in a video sequence. The video encoder also includes a motion estimator and coding mode selector, in signal communication with the concealment distortion calculator, for selecting a coding mode for the block of the particular picture based on the weighted error concealment distortion, to account for packet loss impact on a video quality of the video sequence at a corresponding decoder.

According to a yet further aspect of the present invention, there is provided a video encoder. The video encoder includes a concealment distortion calculator for modeling error concealment distortion for a video sequence using a current picture to be encoded in the video sequence and a next un-coded picture in the video sequence to optimize a received video quality of the video sequence at a corresponding decoder when the video sequence is received subject to network induced errors.

According to an additional aspect of the present invention, there is provided a video encoding method. The method includes calculating a weighted error concealment distortion for a block of a particular picture in a video sequence. The method also includes selecting a coding mode for the block of the particular picture based on the weighted error concealment distortion, to account for packet loss impact on a video quality of the video sequence at a corresponding decoder.

According to a yet additional aspect of the present invention, there is provided a video encoding method. The method includes modeling error concealment distortion for a video sequence using a current picture to be encoded in the video sequence and a next un-coded picture in the video sequence to optimize a received video quality of the video sequence at a corresponding decoder when the video sequence is received subject to network induced errors.

These and other aspects, features and advantages of the present invention will become apparent from the following detailed description of exemplary embodiments, which is to be read in connection with the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention may be better understood in accordance with the following exemplary figures, in which:

FIG. 1 is a block diagram for an exemplary video encoder to which the present principles may be applied, in accordance with an embodiment of the present principles;

FIG. 2 is a block diagram for an exemplary method for low complexity error resilient motion estimation and coding mode selection in according with an embodiment of the present principles;

FIG. 3 is a block diagram for an exemplary method for SKIP mode selection in accordance with an embodiment of the present principles;

FIG. 4 is a block diagram for an exemplary method for inter-prediction mode selection in accordance with an embodiment of the present principles; and

FIG. 5 is a block diagram for an exemplary method for intra prediction mode selection in accordance with an embodiment of the present principles.

DETAILED DESCRIPTION

The present invention is directed to methods and apparatus for low complexity error resilient motion estimation and coding mode selection.

The present description illustrates the principles of the present invention. It will thus be appreciated that those skilled in the art will be able to devise various arrangements that, although not explicitly described or shown herein, embody the principles of the invention and are included within its spirit and scope.

All examples and conditional language recited herein are intended for pedagogical purposes to aid the reader in understanding the principles of the invention and the concepts contributed by the inventor to furthering the art, and are to be construed as being without limitation to such specifically recited examples and conditions.

Moreover, all statements herein reciting principles, aspects, and embodiments of the invention, as well as specific examples thereof, are intended to encompass both structural and functional equivalents thereof. Additionally, it is intended that such equivalents include both currently known equivalents as well as equivalents developed in the future, i.e., any elements developed that perform the same function, regardless of structure.

Thus, for example, it will be appreciated by those skilled in the art that the block diagrams presented herein represent conceptual views of illustrative circuitry embodying the principles of the invention. Similarly, it will be appreciated that any flow charts, flow diagrams, state transition diagrams, pseudocode, and the like represent various processes which may be substantially represented in computer readable media and so executed by a computer or processor, whether or not such computer or processor is explicitly shown.

The functions of the various elements shown in the figures may be provided through the use of dedicated hardware as well as hardware capable of executing software in association with appropriate software. When provided by a processor, the functions may be provided by a single dedicated processor, by a single shared processor, or by a plurality of individual processors, some of which may be shared. Moreover, explicit use of the term "processor" or "controller" should not be construed to refer exclusively to hardware capable of executing software, and may implicitly include, without limitation, digital signal processor ("DSP") hardware, read-only memory ("ROM") for storing software, random access memory ("RAM"), and non-volatile storage.

Other hardware, conventional and/or custom, may also be included. Similarly, any switches shown in the figures are conceptual only. Their function may be carried out through the operation of program logic, through dedicated logic, through the interaction of program control and dedicated logic, or even manually, the particular technique being selectable by the implementer as more specifically understood from the context.

In the claims hereof, any element expressed as a means for performing a specified function is intended to encompass any way of performing that function including, for example, a) a combination of circuit elements that performs that function or b) software in any form, including, therefore, firmware, microcode or the like, combined with appropriate circuitry for executing that software to perform the function. The invention as defined by such claims resides in the fact that the functionalities provided by the various recited means are combined and brought together in the manner which the claims call for. It is thus regarded that any means that can provide those functionalities are equivalent to those shown herein.

Reference in the specification to "one embodiment" or "an embodiment" of the present principles means that a particular feature, structure, characteristic, and so forth described in connection with the embodiment is included in at least one embodiment of the present principles. Thus, the appearances of the phrase "in one embodiment" or "in an embodiment" appearing in various places throughout the specification are not necessarily all referring to the same embodiment.

Turning to FIG. 1, an exemplary video encoder is indicated generally by the reference numeral 100. An input to the video encoder 100 is connected in signal communication with a non-inverting input of a summing junction 110. The output of the summing junction 110 is connected in signal communication with an input of a transformer 120. An output of the transformer 120 is connected in signal communication with an input of a quantizer 123. An output of the quantizer 123 is connected in signal communication with an input of a variable length coder 140. An output of the variable length coder 140 is available as an output of the encoder 100.

The output of the quantizer 123 is further connected in signal communication with an input of an inverse quantizer 150. An output of the inverse quantizer 150 is connected in signal communication with an input of an inverse transformer 160. An output of the inverse transformer 160 is connected in signal communication with an input of a reference picture store 170. A first output of the reference picture store 170 is connected in signal communication with a first input of a motion estimator and coding mode selector 180. The input to the encoder 100 is further connected in signal communication with a second input of the motion estimator and coding mode selector 180 and a first input of a concealment distortion calculator 166. An output of the concealment distortion calculator 166 is connected in signal communication with a third input of the motion estimator and coding mode selector 180. A first output of the motion estimator and coding mode selector 180 is connected in signal communication with a second input of the concealment distortion calculator 166. A second output of the motion estimator and coding mode selector 180 is connected in signal communication with a first input of a motion compensator 190. A second output of the reference picture store 170 is connected in signal communication with a second input of the motion compensator 190. The output of the motion compensator 190 is connected in signal communication with an inverting input of the summing junction 110.

In accordance with an embodiment of the present principles, a method and apparatus for error resilient motion estimation and mode selection is provided. A new distortion metric is used to more reasonably account for the impact of the "worst-case" distortion on the overall subjective video quality at the decoder. End-to-end distortion estimation is not required, thus achieving low complexity.

Existing error resilient motion estimation and mode selection schemes assume statistically defined end-to-end distortion as their ultimate video quality metric. We use f_n^i, $\hat{f}_n^i$, and $\tilde{f}_n^i$ to denote respectively the original, encoder reconstructed, and decoder reconstructed pixel *i* in frame *n*. Let *Blk* denote the current coding block. Let *p* denote the packet loss rate, and presume one frame per packet. The squared-error end-to-end distortion is defined as follows:

$$E\{D_n\} = \sum_{i \in Blk} E\{(f_n^i - \tilde{f}_n^i)^2\}. \quad (1)$$

For an Inter coded block, one can rewrite $E\{D_n\}$ as follows:

$$E\{D_n\} \approx (1-p)\cdot(D_{Enc,n} + D_{EP,n}) + p \cdot D_{EC,n} \quad (2)$$

$D_{Enc,n}$, $D_{EC,n}$, and $D_{EP,n}$ denote encoder source coding distortion, decoder error concealment distortion, and error propagation distortion, respectively, which are defined as follows. We use $\tilde{f}_{n,EC}^i$ to denote the decoder error concealment pixel for the lost f_n^i. The motion vector for that block is defined as *mv*. Also, presume that motion compensation is only allowed from the previous frame.

$$D_{Enc,n} = \sum_{i \in Blk} (f_n^i - \hat{f}_n^i)^2, \quad (3)$$

$$D_{EP,n} = \sum_{i \in Blk} E\{(f_{n-1}^{i+mv} - \tilde{f}_{n-1}^{i+mv})^2\}, \quad (4)$$

$$D_{EC,n} = \sum_{i \in Blk} E\{(f_n^i - \tilde{f}_{n,EC}^i)^2\}. \quad (5)$$

In the scenario of video streaming, packet loss and consequent error propagation generally incur more serious damage on video quality than quantization distortion (or source coding distortion). Hence, in practice, even if a video presentation yields a lower average PSNR than that of another video presentation due to its compromised "no-loss-case" PSNR (or source coding quality), its overall perceptual quality may still be much better, if it renders a much better "loss-case" quality. However, from Equation (2), one can see that existing $E\{D_n\}$ accounts for the impact of $D_{EC,n}$ simply by the loss probability p of the current frame and, thus, its actual impact on the overall subjective quality is underrated.

To more accurately capture this impact, and in accordance an embodiment of the present principles, we propose a new distortion metric as follows:

$$D_n = D_{Enc,n} + w_n \cdot D_{EC,n}. \quad (6)$$

By properly setting the weighting factor w, one can flexibly control the proportion of error concealment distortion $D_{EC,n}$ in the overall distortion, which directly affects the trade-off between the "no-loss-case" distortion $D_{Enc,n}$ and the "loss-case" distortion $D_{EC,n}$, such that better overall perceptual quality can be achieved.

How to set the value of w is a relevant issue for the application of the new distortion metric, for which several relevant factors may be involved. Presume a video sequence is coded into groups of pictures (GOPs) and each GOP starts with an I-frame (an intra-coded frame), followed by P-frames (inter-coded frames). Ideally, w should be related with the position of the current frame in a GOP. It is generally true that the loss of a frame in the beginning of a GOP causes more serious error propagation effect than the loss of a frame in the middle or end of the GOP. Therefore, denoting the remaining number of frames in the current GOP by k, a larger k should yield a larger w as well. To more accurately capture the error propagation effect, one may further take into account the intra MB percentage (denoted by γ) of a frame. With a higher γ, less error propagation will be incurred in the following frames. Another factor to be considered is the expected packet loss rate *p*. With a smaller *p* more concern will be focused on the "no-loss-case" distortion, rather than on the "loss-case" distortion. In the extreme case of *p=0*, *w* ideally should also be zero. When the actual expected *p* is unknown, one may use a predetermined value for *p*, representing the worst packet loss case to be addressed by the distortion metric. In one embodiment of the present principles, presuming p is unavailable, a heuristically derived scheme is used, which considers *k* only.

$$w_n = \max[\min[a \cdot k(n), w_{\max}], w_{\min}]. \tag{7}$$

The values of the remaining parameters in (7) may be heuristically determined by experiment. (For example, $a = 0.05$, $w_{\max} = 0.75$, $w_{\min} = 0.25$.)

Of course, given the teachings of the present principles provided herein, other calculations of *w* may be performed possible by considering the aforementioned factors, while maintaining the scope of the present principles.

The new distortion metric in Equation (6) can be generally applied with any existing RD optimization based error resilient video coding and streaming technique, by replacing the previously used end-to-end distortion metric. In accordance with an embodiment of the present principles, by applying the new distortion metric and further assuming "motion copy" error concealment, an error resilient motion estimation and mode selection scheme is provided as follows. It is to be noted that in an embodiment, a new weighted error concealment distortion term is added to account for the impact of packet loss.

For motion estimation, the following is used:

$$mv^* = \arg\min_{mv}[D_{DFD,n}(mv) + w_{n+1} \cdot D_{EC,n+1}(mv) + \lambda_{MV} \cdot R_{mv}]. \tag{8}$$

For mode selection, the following is used:

$$mode^* = \arg\min_{mode}[D_{Enc,n}(mode) + w_{n+1} \cdot D_{EC,n+1}(mode) + \lambda_{MODE} \cdot R(mode)]. \tag{9}$$

Herein, *mv** and *mode** represent, respectively, the selected best motion vector and coding mode for a certain block/MB in frame n. R_{mv} and $R(mode)$ denote the corresponding coding rates. λ_{MV} and λ_{MODE} are the related Lagrangian multipliers. $D_{DFD,n}$ denotes displaced frame difference (DFD), which is defined as

$$D_{DFD,n}(mv) = \sum_{i \in Blk} | f_n^i - \hat{f}_{n-1}^{i+mv} | . \qquad (10)$$

The assumed "motion copy" error concealment is a more effective scheme than the simplest "copy from previous frame" scheme which, thus, is more practically applicable. In this scheme, if a frame is lost, then motion information of co-located blocks in the previous reconstructed frame is used to predict the current frame. Accordingly, *mv* and *mode* selection of a current frame block/MB will affect D_{EC} of the collocated block/MB in the next frame. Therefore, in Equations (8) and (9), both *w* and D_{EC} are for the frame (n+1). Note that existing end-to-end distortion based schemes assume the simplest "copy from previous frame" error concealment. As such, *mv* or *mode* selection of the current coding block/MB in frame *n* does not directly affect $D_{EC,n+1}$.

To calculate $D_{EC,n+1}$, one can see from Equation (5) that it still requires complicated end-to-end estimation due to the presence of $\tilde{f}^i_{n,EC}$. Additionally, some *mv* candidates may refer $\tilde{f}^i_{n,EC}$ to an area not yet coded in the current frame, which incurs a causality problem.

Therefore, to avoid such complication and reduce complexity, we simply approximate $D_{EC,n+1}$ using the previous original frame in its calculation. Note that in practice, the ignored source coding distortion generally represents a fairly small portion of the overall concealment distortion. For squared error distortion,

$$D_{EC,n+1} \approx \sum_{i \in Blk} (f_{n+1}^i - f_n^{i+mv})^2 . \qquad (11)$$

Comparing with the conventional non-error-resilient scheme, the new scheme requires the availability of the next original frame, while coding the current frame. This introduces a one-frame coding delay. The new scheme also incurs additional computation of $w_{n+1}D_{EC,n+1}$, and additional storage of the next original frame. However, both the additional one-frame coding delay and coding complexity pose no serious quality degradation or implementation difficulty in practice. Compared with the prior end-to-end distortion based error resilient coding schemes, this new scheme yields much less computation and storage complexity as it requires no end-to-end estimation. Furthermore, in practice, to reduce complexity, $D_{DFD,n}$ is commonly calculated as a sum-of-absolute-error. Either sum-of-absolute error or sum-of-squared-error, as seen in Equation (10), can be used. Via the aforementioned approximation, $D_{EC,n+1}$ in Equation (8) can be also calculated using the low complexity approach of sum-of-absolute-error. This feature, however, is not readily achievable in prior end-to-end distortion based schemes, as estimating the absolute-error distortion is a highly difficult problem.

Another issue in the new mode selection scheme is concerned with intra/inter mode selection. In "motion copy" concealment, if a co-located MB in the previous frame is coded in intra-mode, then it will be treated as a Skip-mode MB, which means its motion information will be derived using neighboring MBs. As the derived motion vector is not particularly selected for good error concealment performance, coding the current MB in intra-mode generally renders poor concealment performance for the co-located MB in the next frame. On the other hand, it is well known that intra coding will stop existing error propagation from the previous frame, which may greatly reduce the error propagation effect in the following frames.

In accordance with an embodiment of the present principles, and to account for this effect in intra/inter mode selection in the embodiment, we switch between the new scheme defined by Equation (9) and a conventional scheme (i.e., with *w=0* in Equation (9)) according to a heuristically derived criterion, as described below. The criterion include: select the best inter mode via Equations (8) and (9); select the best intra mode via the conventional scheme; if $PSNR(D_{EC,n+1}(Inter)) - PSNR(D_{EC,n+1}(Intra)) \geq \Delta_{th}$, then continue using the new scheme for intra/inter selection, otherwise switch to the conventional scheme for intra/inter selection; select the best coding mode for the current MB.

The basic idea is that when the $D_{EC,n+1}$ reduction from the best inter mode is not significantly large, we conservatively switch to use the conventional scheme for intra/inter mode selection. Note that this switching incurs no additional computation complexity, as $D_{Enc,n}$ and $D_{EC,n+1}$ can be simultaneously calculated in practice. Δ_{th} is a pre-determined threshold that balances the trade-off between good error concealment of the next frame, and good error propagation reduction in the remaining frames of the current GOP. For example, an appropriate threshold for the instant embodiment of the present invention is $\Delta_{th} = 7dB$.

Further, it is noted that performance of the new motion estimation and mode selection scheme may be improved using a correspondingly developed rate control algorithm, which properly controls the quantization scale, λ_{MV} and λ_{MODE} of each MB (or frame) and, thus, helps the encoder to achieve a certain given total bit rate. In an embodiment, we use the same well-known Lagrangian multiplier calculation scheme as that adopted in the reference JM H.264 encoder. For better rate control performance, one may develop other ad hoc algorithms particularly addressing the new scheme.

Simulation results have shown that in spite of compromised "no-loss-case" average PSNR, the new scheme effectively improves the performance of "motion copy" error concealment at the decoder, which consequently leads to greatly reduced error propagation effect. Thus, with the presence of packet loss, the new scheme yields much less quality variation than that of the conventional scheme, and an overall improved subjective video quality is achieved at the decoder. It is also shown that higher performance enhancement is achieved for medium motion sequences than that for low motion or high motion sequences, which render either under-challenging or over-challenging error concealment situations.

A low complexity error resilient method for motion estimation and coding mode selection in accordance with an embodiment of the present principles will now be described with respect to FIG. 2. Subsequent FIGs. 3, 4, and 5 further illustrate blocks 225, 230, and 240, respectively, of the method of FIG. 2. As used herein: "CAME" denotes concealment-aware motion estimation, defined in Equation (8);

"CAMS" denotes concealment-aware mode selection, defined in Equation (9); and "conventional MS" denotes conventional non-error-resilient mode selection, i.e., with w=0 in Equation (9).

Turning to FIG. 2, an exemplary method for low complexity error resilient motion estimation and coding mode selection is indicated generally by the reference numeral 200. The method 200 includes a start block 202 that passes control to a loop limit block 205. The loop limit block 205 performs a loop for each frame n, and passes control to a function block 210. The function block 210 obtains k from the GOP position of frame n, and passes control to a function block 215. The function block 215 sets $w_n = \max[\min[a \cdot k(n), w_{\max}], w_{\min}]$, and passes control to a loop limit block 220. The loop limit block 220 performs a loop for each macroblock of a current frame n, and passes control to a function block 225. The function block 225, for SKIP mode, calculates $D_{EC,n+1}$ (SKIP) and the CAMS cost (see FIG. 3), and passes control to a function block 230. The function block 230, for each inter-pred (inter-prediction) mode, selects the motion vector(s) with CAME, calculates the CAMS cost (see FIG. 3), and passes control to a function block 235. The function block 235 selects the best inter mode, and passes control to a function block 240. The function block 240, for each intra mode, selects the intra prediction via conventional MS, calculates the conventional MS cost (see FIG. 4), and passes control to a function block 245. The function block 245 selects the best intra mode, and passes control to a function block 250. The function block 250 calculates $D_{EC,n+1}$ (INTRA) = $D_{EC,n+1}$ (SKIP), and passes control to a decision block 255. The decision block 255 determines whether or not X=PNSR($D_{EC,n+1}$ (Inter))< Y=PSNR($D_{EC,n+1}$ (Intra)), and if (X-Y)>Δ_{th}. If so, then control is passed to a function block 260. Otherwise, control is passed to a function block 265.

The function block 260 uses the CAMS cost to select the best mode, and passes control to a loop limit block 270. The loop limit block 270 ends the loop over each macroblock, and passes control to a loop limit block 275. The loop limit block 275 passes control to an end block 280.

The function block 265 uses the conventional MS cost to select the best mode, and passes control to the loop limit block 270.

Turning to FIG. 3, an exemplary method for SKIP mode selection is indicated generally by the reference numeral 300. The method 300 includes a start block 302 that passes control to a function block 305. The function block 305 finds the motion vector of SKIP mode from the neighboring coded macroblocks, and passes control to a function block 310. The function block 310 calculates the squared-error $D_{EC,n}$, $D_{EC,n+1}$ (SKIP), and R(SKIP), and passes control to a function block 315. The function block 315 calculates the CAMS cost of SKIP mode, and passes control to an end block 320.

Turning to FIG. 4, an exemplary method for inter-prediction mode selection is indicated generally by the reference numeral 400. The method includes a start block 402 that passes control to a loop limit block 405. The loop limit block 405 performs a loop for each inter prediction mode, and passes control to a loop limit block 410. The loop limit block 410 performs a loop for each block, and passes control to a loop limit block 415. The loop limit block 415 performs a loop for each motion vector, and passes control to a function block 420. The function block 420 calculates the absolute-error $D_{DFD,n}$, $D_{EC,n+1}$, R_{mv}, and passes control to a function block 425. The function block 425 calculates the CAM cost, and passes control to a loop limit block 430. The loop limit block 430 ends the loop for each of the motion vectors, and passes control to a function block 435. The function block 435 selects the best motion vector of the block, and passes control to a loop limit block 440. The loop limit block 440 ends the loop for each of the blocks, and passes control to a function block 445. The function block 445 calculates the squared-error $D_{EC,n}$, $D_{EC,n+1}$, and R(inter-pred) of the whole macroblock, and passes control to a function block 450. The function block 450 calculates the CAMS cost of the inter prediction mode, and passes control to a loop limit block 455. The loop limit block ends the loop over each inter prediction mode, and passes control to an end block 460.

Turning to FIG. 5, an exemplary method for intra prediction mode selection is indicated generally by the reference numeral 500. The method 500 includes a start block 505 that passes control to a loop limit block 505. The loop limit block 505 performs a loop for each intra 16x16 or intra 4x4 mode, and passes control to a loop limit block 510. The loop limit block 510 performs a loop for each block, and passes control to a loop limit block 515. The loop limit block 515 performs a loop for each intra-pred (intra prediction) mode, and passes control to a function block 520. The function block 520 calculates the squared error $D_{EC,n}$ and R(intra-pred) of the block, and passes control to a function block 525. The function block 525 calculates the conventional MS cost, and passes control to a loop limit block 530. The loop limit block 530 ends the loop over each intra-pred mode, and passes control to a function block 535. The function block 535 selects the best intra-pred mode of the block, and passes control to a loop limit block 540. The loop limit block 540 ends the loop over each block, and passes control to a function block 545. The function block 545 calculates the squared error $D_{EC,n}$, and R(intra) of the whole macroblock, and passes control to a function block 550. The function block 550 calculates the conventional MS cost of the intra mode, and passes control to a loop limit block 555. The loop limit block 555 ends the loop over each intra mode, and passes control to an end bock 560.

A description will now be given of some of the many attendant advantages/features of the present invention, some of which have been mentioned above. For example, one advantage/feature is a video encoder that includes an encoder for encoding a video sequence using weighted error concealment distortion to account for packet loss impact on a video quality of the video sequence at a corresponding decoder.

Another advantage/feature is the video encoder as described above, wherein the encoder calculates the weighted error concealment distortion on a block-basis.

Yet another advantage/feature is the video encoder as described above, wherein the video sequence includes a Group of Pictures, and the encoder encodes a particular picture in the Group of Pictures by calculating the weighted error concealment distortion for a block in the particular picture using a concealment distortion weighting factor that is based upon a position of the particular picture in the Group of Pictures.

Moreover, another advantage/feature is the video encoder that calculates the weighted error concealment distortion using the concealment distortion weighting factor as described above, wherein the concealment distortion weighting factor is set to a predetermined constant value.

Further, another advantage/feature is the video encoder as described above, wherein the video sequence includes a Group of Pictures, and said encoder encodes a particular picture in the Group of Pictures by calculating the weighted error concealment distortion for a block of the particular picture using a concealment distortion weighting factor that depends upon an intra block percentage of remaining pictures in the Group of Pictures.

Also, another advantage/feature is the video encoder as described above, wherein the encoder encodes a particular picture in the video sequence by calculating the weighted error concealment distortion for the particular picture using a concealment distortion weighting factor that depends upon an expected packet loss rate relating to the particular picture.

Additionally, another advantage/feature is the video encoder as described above, wherein said encoder encodes the video stream using a concealment-aware distortion metric that is based upon the weighted error concealment distortion.

Moreover, another advantage/feature is the video encoder that encodes the video stream using the concealment-aware distortion metric as described above, wherein the concealment-aware distortion metric is further based upon an encoding source coding distortion.

Further, another advantage/feature is the video encoder that encodes the video stream using the concealment-aware distortion metric as described above, wherein the concealment-aware distortion metric is without basis on end-to-end distortion.

Also, another advantage/feature is the video encoder as described above, wherein the encoder performs an error resilient motion estimation to optimize, based upon the weighted error concealment distortion, a received video quality of the video sequence at a corresponding decoder when the video sequence is received subject to network induced errors.

Additionally, another advantage/feature is the video encoder that performs the error resilient motion estimation as described above, wherein the encoder calculates the weighted error concealment distortion using a current picture to be encoded in the video sequence and a next un-coded picture in the video sequence.

Additionally, another advantage/feature is the video encoder that performs the error resilient motion estimation as described above, wherein the video sequence is encoded to optimize a subsequent application of a motion copy process to one or more pictures of the video sequence at the corresponding decoder.

Moreover, another advantage/feature is the video encoder as described above, wherein the encoder performs an error resilient encoding mode selection to select an encoding mode for a block of a current picture to be encoded in the video sequence using the weighted error concealment distortion.

Further, another advantage/feature is the video encoder that performs an error resilient encoding mode selection as described above, wherein the encoder calculates the weighted error concealment distortion using a current picture to be encoded in the video sequence and a next un-coded picture in the video sequence.

Also, another advantage/feature is the video encoder that performs an error resilient encoding mode selection as described above, wherein the video sequence is encoded to optimize a subsequent application of a motion copy process to one or more pictures of the video sequence at the corresponding decoder.

Additionally, another advantage/feature is the video encoder as described above, wherein said encoder performs a selection between intra mode and inter mode to encode a current picture in the video sequence by comparing a concealment distortion Peak Signal to Noise Ratio gain of the inter mode over the intra mode with respect to a threshold.

Moreover, another advantage/feature is the video encoder that includes an encoder for encoding a video sequence by modeling error concealment distortion using a current picture to be encoded in the video sequence and a next un-coded picture in the video sequence to optimize a received video quality of the video sequence at a corresponding decoder when the video sequence is received subject to network induced errors.

These and other features and advantages of the present invention may be readily ascertained by one of ordinary skill in the pertinent art based on the teachings herein. It is to be understood that the teachings of the present invention may be implemented in various forms of hardware, software, firmware, special purpose processors, or combinations thereof.

Most preferably, the teachings of the present invention are implemented as a combination of hardware and software. Moreover, the software may be implemented as an application program tangibly embodied on a program storage unit. The application program may be uploaded to, and executed by, a machine comprising any suitable architecture. Preferably, the machine is implemented on a computer platform having hardware such as one or more central processing units ("CPU"), a random access memory ("RAM"), and input/output ("I/O") interfaces. The computer platform may also include an operating system and microinstruction code. The various processes and functions described herein may be either part of the microinstruction code or part of the application program, or any combination thereof, which may be executed by a CPU. In addition, various other peripheral units may be connected to the computer platform such as an additional data storage unit and a printing unit.

It is to be further understood that, because some of the constituent system components and methods depicted in the accompanying drawings are preferably implemented in software, the actual connections between the system components or the process function blocks may differ depending upon the manner in which the present invention is programmed. Given the teachings herein, one of ordinary skill in the pertinent art will be able to contemplate these and similar implementations or configurations of the present invention.

Although the illustrative embodiments have been described herein with reference to the accompanying drawings, it is to be understood that the present invention is not limited to those precise embodiments, and that various changes and modifications may be effected therein by one of ordinary skill in the pertinent art without departing from the scope or spirit of the present invention. All such changes and modifications are intended to be included within the scope of the present invention as set forth in the appended claims.

CLAIMS:

1. An apparatus, comprising:

an encoder (100) for encoding a video sequence using weighted error concealment distortion to compensate for packet loss impact on a video quality of the video sequence at a corresponding decoder.

2. The apparatus of claim 1, wherein said encoder (100) calculates the weighted error concealment distortion on a block-basis.

3. The apparatus of claim 1, wherein the video sequence includes a Group of Pictures, and said encoder (100) encodes a particular picture in the Group of Pictures by calculating the weighted error concealment distortion for a block in the particular picture using a concealment distortion weighting factor that is based upon a position of the particular picture in the Group of Pictures.

4. The apparatus of claim 3, wherein the concealment distortion weighting factor is set to a predetermined constant value.

5. The apparatus of claim 1, wherein the video sequence includes a Group of Pictures, and said encoder (100) encodes a particular picture in the Group of Pictures by calculating the weighted error concealment distortion for a block of the particular picture using a concealment distortion weighting factor that depends upon an intra block percentage of remaining pictures in the Group of Pictures.

6. The apparatus of claim 1, wherein said encoder (100) encodes a particular picture in the video sequence by calculating the weighted error concealment distortion for the particular picture using a concealment distortion weighting factor that depends upon an expected packet loss rate relating to the particular picture.

7. The apparatus of claim 1, wherein said encoder (100) encodes the video stream using a concealment-aware distortion metric that is based upon the weighted error concealment distortion.

8. The apparatus of claim 7, wherein the concealment-aware distortion metric is further based upon an encoding source coding distortion.

9. The apparatus of claim 7, wherein the concealment-aware distortion metric is without basis on end-to-end distortion.

10. The apparatus of claim 1, wherein said encoder (100) performs an error resilient motion estimation to optimize, based upon the weighted error concealment distortion, a received video quality of the video sequence at a corresponding decoder when the video sequence is received subject to network induced errors.

11. The apparatus of claim 10, wherein said encoder (100) calculates the weighted error concealment distortion using a current picture to be encoded in the video sequence and a next un-coded picture in the video sequence.

12. The apparatus of claim 10, wherein the video sequence is encoded to optimize a subsequent application of a motion copy process to one or more pictures of the video sequence at the corresponding decoder.

13. The apparatus of claim 1, wherein said encoder (100) performs an error resilient encoding mode selection to select an encoding mode for a block of a current picture to be encoded in the video sequence using the weighted error concealment distortion.

14. The apparatus of claim 13, wherein said encoder (100) calculates the weighted error concealment distortion using a current picture to be encoded in the video sequence and a next un-coded picture in the video sequence.

15. The apparatus of claim 13, wherein the video sequence is encoded to optimize a subsequent application of a motion copy process to one or more pictures of the video sequence at the corresponding decoder.

16. The apparatus of claim 1, wherein said encoder (100) performs a selection between intra mode and inter mode to encode a current picture in the video sequence by comparing a concealment distortion Peak Signal to Noise Ratio gain of the inter mode over the intra mode with respect to a threshold.

17. An apparatus, comprising:

an encoder (100) for encoding a video sequence by modeling error concealment distortion using a current picture to be encoded in the video sequence and a next un-coded picture in the video sequence to optimize a received video quality of the video sequence at a corresponding decoder when the video sequence is received subject to network induced errors.

18. A video encoding method, comprising:

encoding (200) a video sequence using weighted error concealment distortion to compensate for packet loss impact on a video quality of the video sequence at a corresponding decoder.

19. The method of claim 18, wherein said encoding step calculates the weighted error concealment distortion on a block-basis (410, 510).

20. The method of claim 18, wherein the video sequence includes a Group of Pictures, and said encoding step encodes a particular picture in the Group of Pictures by calculating the weighted error concealment distortion for a block in the particular picture using a concealment distortion weighting factor that is based upon a position of the particular picture in the Group of Pictures (210).

21. The method of claim 20, wherein the concealment distortion weighting factor is set to a predetermined constant value.

22. The method of claim 18, wherein the video sequence includes a Group of Pictures, and said encoding step encodes a particular picture in the Group of Pictures by calculating the weighted error concealment distortion for a block of the particular picture using a concealment distortion weighting factor that depends upon an intra block percentage of remaining pictures in the Group of Pictures.

23. The method of claim 18, wherein said encoding step encodes a particular picture in the video sequence by calculating the weighted error concealment distortion for the particular picture using a concealment distortion weighting factor that depends upon an expected packet loss rate relating to the particular picture.

24. The method of claim 18, wherein said encoding step encodes the video stream using a concealment-aware distortion metric that is based upon the weighted error concealment distortion.

25. The method of claim 24, wherein the concealment-aware distortion metric is further based upon an encoding source coding distortion.

26. The method of claim 24, wherein the concealment-aware distortion metric is without basis on end-to-end distortion.

27. The method of claim 18, wherein said encoding step performs an error resilient motion estimation to optimize, based upon the weighted error concealment distortion, a received video quality of the video sequence at a corresponding decoder when the video sequence is received subject to network induced errors (230).

28. The method of claim 27, wherein said encoding step calculates the weighted error concealment distortion using a current picture to be encoded in the video sequence and a next un-coded picture in the video sequence.

29. The method of claim 27, wherein the video sequence is encoded to optimize a subsequent application of a motion copy process to one or more pictures of the video sequence at the corresponding decoder.

30. The method of claim 18, wherein said encoding step performs an error resilient encoding mode selection to select an encoding mode for a block of a current picture to be encoded in the video sequence using the weighted error concealment distortion (260).

31. The method of claim 30, wherein said encoding step calculates the weighted error concealment distortion using a current picture to be encoded in the video sequence and a next un-coded picture in the video sequence.

32. The method of claim 30, wherein the video sequence is encoded to optimize a subsequent application of a motion copy process to one or more pictures of the video sequence at the corresponding decoder.

33. The method of claim 18, wherein said encoding step performs a selection between intra mode and inter mode to encode a current picture in the video sequence by comparing a concealment distortion Peak Signal to Noise Ratio gain of the inter mode over the intra mode with respect to a threshold (255).

34. A video encoding method, comprising:

encoding a video sequence by modeling error concealment distortion using a current picture to be encoded in the video sequence and a next un-coded picture in the video sequence to optimize a received video quality of the video sequence at a corresponding decoder when the video sequence is received subject to network induced errors.

35. A video encoder, comprising:

a concealment distortion calculator (166) for calculating a weighted error concealment distortion for a block of a particular picture in a video sequence; and

a motion estimator and coding mode selector (180), in signal communication with said concealment distortion calculator, for selecting a coding mode for the block of the particular picture based on the weighted error concealment distortion, to compensate for packet loss impact on a video quality of the video sequence at a corresponding decoder.

36. The video encoder of claim 35, wherein said concealment distortion calculator (166) calculates the weighted error concealment distortion on a block-basis.

37. The video encoder of claim 35, wherein the video sequence includes a Group of Pictures, the particular picture is included in the Group of Pictures, and said concealment distortion calculator (166) calculates the weighted error concealment distortion for the block of the particular picture in the Group of Pictures using a concealment distortion weighting factor that is based upon a position of the particular picture in the Group of Pictures.

38. The video encoder of claim 37, wherein the concealment distortion weighting factor is set to a predetermined constant value.

39. The video encoder of claim 35, wherein the video sequence includes a Group of Pictures, the particular picture is included in the Group of Pictures, and said concealment distortion calculator (166) calculates the weighted error concealment distortion for the block of the particular picture in the Group of Pictures using a concealment distortion weighting factor that depends upon an intra block percentage of remaining pictures in the Group of Pictures.

40. The video encoder of claim 35, wherein said concealment distortion calculator (166) calculates the weighted error concealment distortion for the particular picture using a concealment distortion weighting factor that depends upon an expected packet loss rate relating to the particular picture.

41. The video encoder of claim 35, wherein said concealment distortion calculator (166) calculates a concealment-aware distortion metric that is based upon the weighted error concealment distortion, the concealment-aware distortion metric for use in encoding the video sequence.

42. The video encoder of claim 41, wherein the concealment-aware distortion metric is further based upon an encoding source coding distortion.

43. The video encoder of claim 41, wherein the concealment-aware distortion metric is without basis on end-to-end distortion.

44. The video encoder of claim 35, wherein said motion estimator and coding mode selector (180) performs an error resilient motion estimation to optimize, based upon the weighted error concealment distortion, a received video quality of the video sequence at a corresponding decoder when the video sequence is received subject to network induced errors.

45. The video encoder of claim 44, wherein said concealment distortion calculator (166) calculates the weighted error concealment distortion using the particular picture and a next un-coded picture in the video sequence.

46. The video encoder of claim 44, wherein the video sequence is encoded to optimize a subsequent application of a motion copy process to one or more pictures of the video sequence at the corresponding decoder.

47. The video encoder of claim 35, wherein said concealment distortion calculator (166) calculates the weighted error concealment distortion using the particular picture and a next un-coded picture in the video sequence.

48. The video encoder of claim 35, wherein the video sequence is encoded to optimize a subsequent application of a motion copy process to one or more pictures of the video sequence at the corresponding decoder.

49. The video encoder of claim 35, wherein said motion estimator and coding mode selector (180) performs a selection between intra mode and inter mode to encode the particualar in the video sequence by comparing a concealment distortion Peak Signal to Noise Ratio gain of the inter mode over the intra mode with respect to a threshold.

50. A video encoder, comprising:

a concealment distortion calculator (166) for modeling error concealment distortion for a video sequence using a current picture to be encoded in the video sequence and a next un-coded picture in the video sequence to optimize a received video quality of the video sequence at a corresponding decoder when the video sequence is received subject to network induced errors.

51. A video encoding method, comprising:

calculating a weighted error concealment distortion for a block of a particular picture in a video sequence; and

selecting (260) a coding mode for the block of the particular picture based on the weighted error concealment distortion, to compensate for packet loss impact on a video quality of the video sequence at a corresponding decoder.

52. The method of claim 51, wherein said calculating step calculates the weighted error concealment distortion on a block-basis (410, 510).

53. The method of claim 51, wherein the video sequence includes a Group of Pictures, the particular picture is included in the Group of Pictures, and said calculating step calculates the weighted error concealment distortion for the block of the particular picture in the Group of Pictures using a concealment distortion weighting factor that is based upon a position of the particular picture in the Group of Pictures (210).

54. The method of claim 53, wherein the concealment distortion weighting factor is set to a predetermined constant value.

55. The method of claim 51, wherein the video sequence includes a Group of Pictures, the particular picture is included in the Group of Pictures, and said calculating step calculates the weighted error concealment distortion for the block of the particular picture in the Group of Pictures using a concealment distortion weighting factor that depends upon an intra block percentage of remaining pictures in the Group of Pictures.

56. The method of claim 51, wherein said calculating step calculates the weighted error concealment distortion for the particular picture using a concealment distortion weighting factor that depends upon an expected packet loss rate relating to the particular picture.

57. The method of claim 51, wherein said calculating step calculates a concealment-aware distortion metric that is based upon the weighted error concealment distortion, the concealment-aware distortion metric for use in encoding the video sequence.

58. The method of claim 57, wherein the concealment-aware distortion metric is further based upon an encoding source coding distortion.

59. The method of claim 57, wherein the concealment-aware distortion metric is without basis on end-to-end distortion.

60. The method of claim 51, further comprising performing an error resilient motion estimation to optimize, based upon the weighted error concealment distortion, a received video quality of the video sequence at a corresponding decoder when the video sequence is received subject to network induced errors. (230)

61. The method of claim 60, wherein said calculating step calculates the weighted error concealment distortion using the particular picture and a next un-coded picture in the video sequence.

62. The method of claim 60, wherein the video sequence is encoded to optimize a subsequent application of a motion copy process to one or more pictures of the video sequence at the corresponding decoder.

63. The method of claim 51, wherein said calculating step calculates the weighted error concealment distortion using the particular picture and a next un-coded picture in the video sequence.

64. The method of claim 51, wherein the video sequence is encoded to optimize a subsequent application of a motion copy process to one or more pictures of the video sequence at the corresponding decoder.

65. The method of claim 51, wherein said selecting step performs a selection between intra mode and inter mode to encode the particular in the video sequence by comparing a concealment distortion Peak Signal to Noise Ratio gain of the inter mode over the intra mode with respect to a threshold (255).

66. A video encoding method, comprising:

modeling error concealment distortion for a video sequence using a current picture to be encoded in the video sequence and a next un-coded picture in the video sequence to optimize a received video quality of the video sequence at a corresponding decoder when the video sequence is received subject to network induced errors.

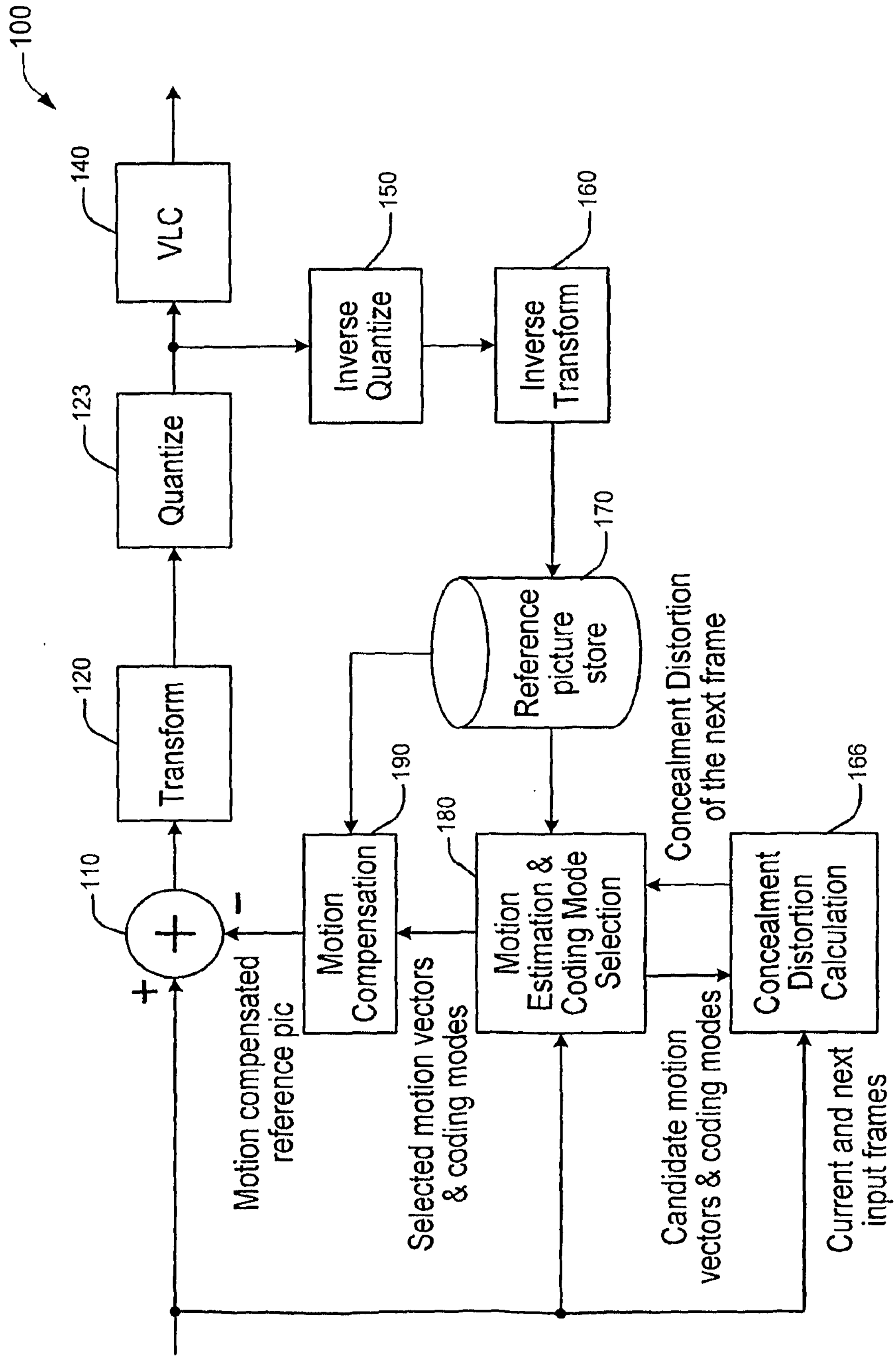

100
110
+
-
120
Transform
123
Quantize
140
VLC
150
Inverse Quantize
160
Inverse Transform
170
Reference picture store
190
Motion Compensation
180
Motion Estimation & Coding Mode Selection
166
Concealment Distortion Calculation
Motion compensated reference pic
Selected motion vectors & coding modes
Candidate motion vectors & coding modes
Concealment Distortion of the next frame
Current and next input frames

FIG. 1

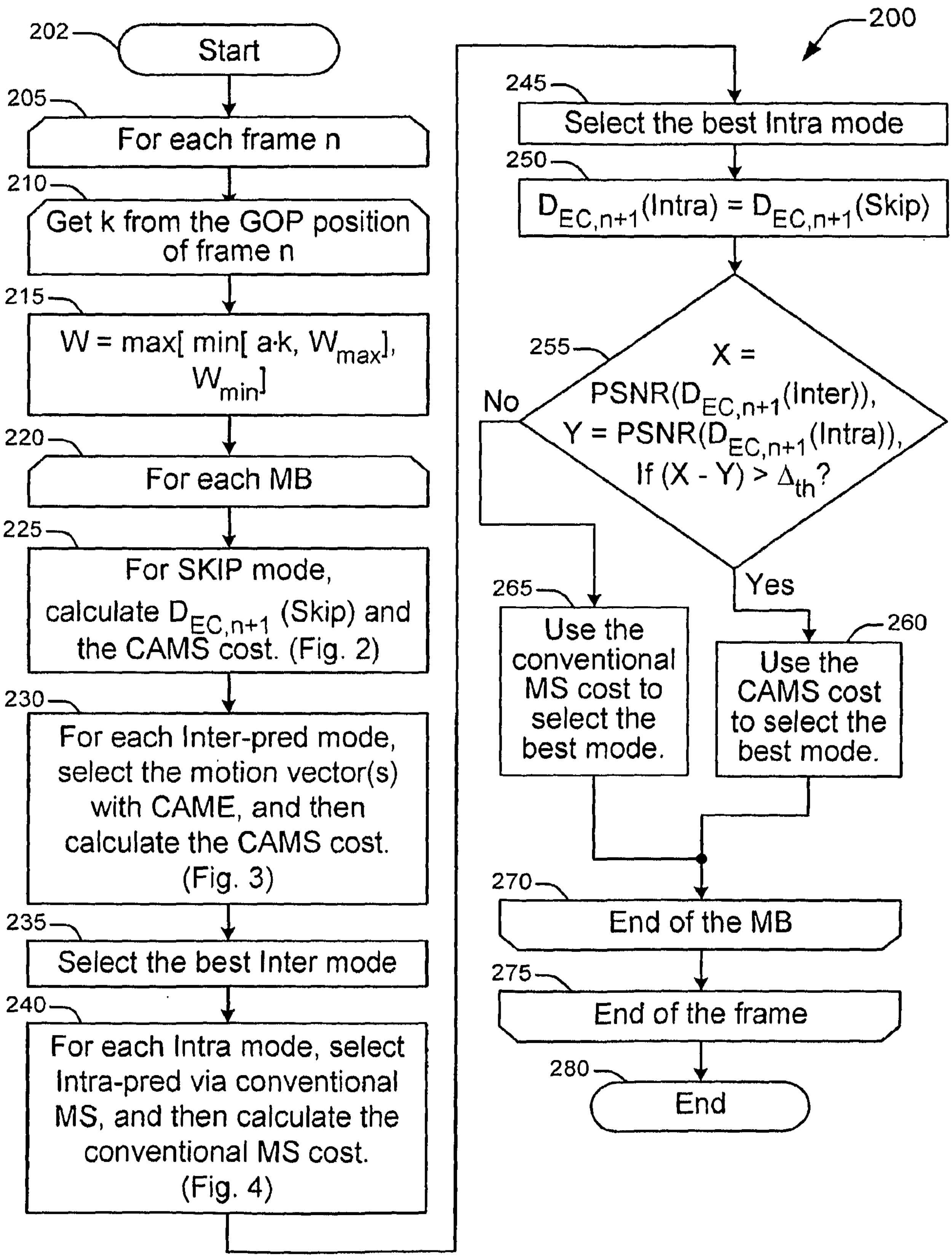

200
202
Start
205
For each frame n
210
Get k from the GOP position of frame n
215
W = max[min[a·k, W_max], W_min]
220
For each MB
225
For SKIP mode, calculate D_EC,n+1 (Skip) and the CAMS cost. (Fig. 2)
230
For each Inter-pred mode, select the motion vector(s) with CAME, and then calculate the CAMS cost. (Fig. 3)
235
Select the best Inter mode
240
For each Intra mode, select Intra-pred via conventional MS, and then calculate the conventional MS cost. (Fig. 4)
245
Select the best Intra mode
250
D_EC,n+1(Intra) = D_EC,n+1(Skip)
255
X = PSNR(D_EC,n+1(Inter)), Y = PSNR(D_EC,n+1(Intra)), If (X - Y) > Δ_th?
No
Yes
265
Use the conventional MS cost to select the best mode.
260
Use the CAMS cost to select the best mode.
270
End of the MB
275
End of the frame
280
End

FIG. 2

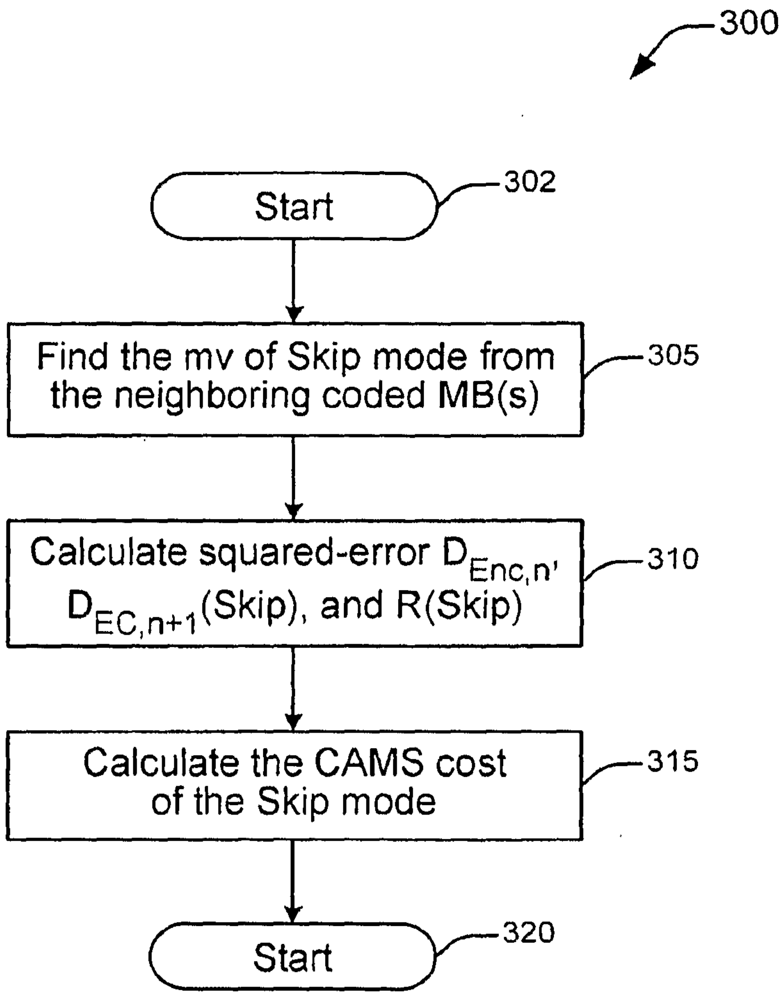

300
Start
302
Find the mv of Skip mode from the neighboring coded MB(s)
305
Calculate squared-error $D_{Enc,n}$, $D_{EC,n+1}$(Skip), and R(Skip)
310
Calculate the CAMS cost of the Skip mode
315
Start
320

FIG. 3

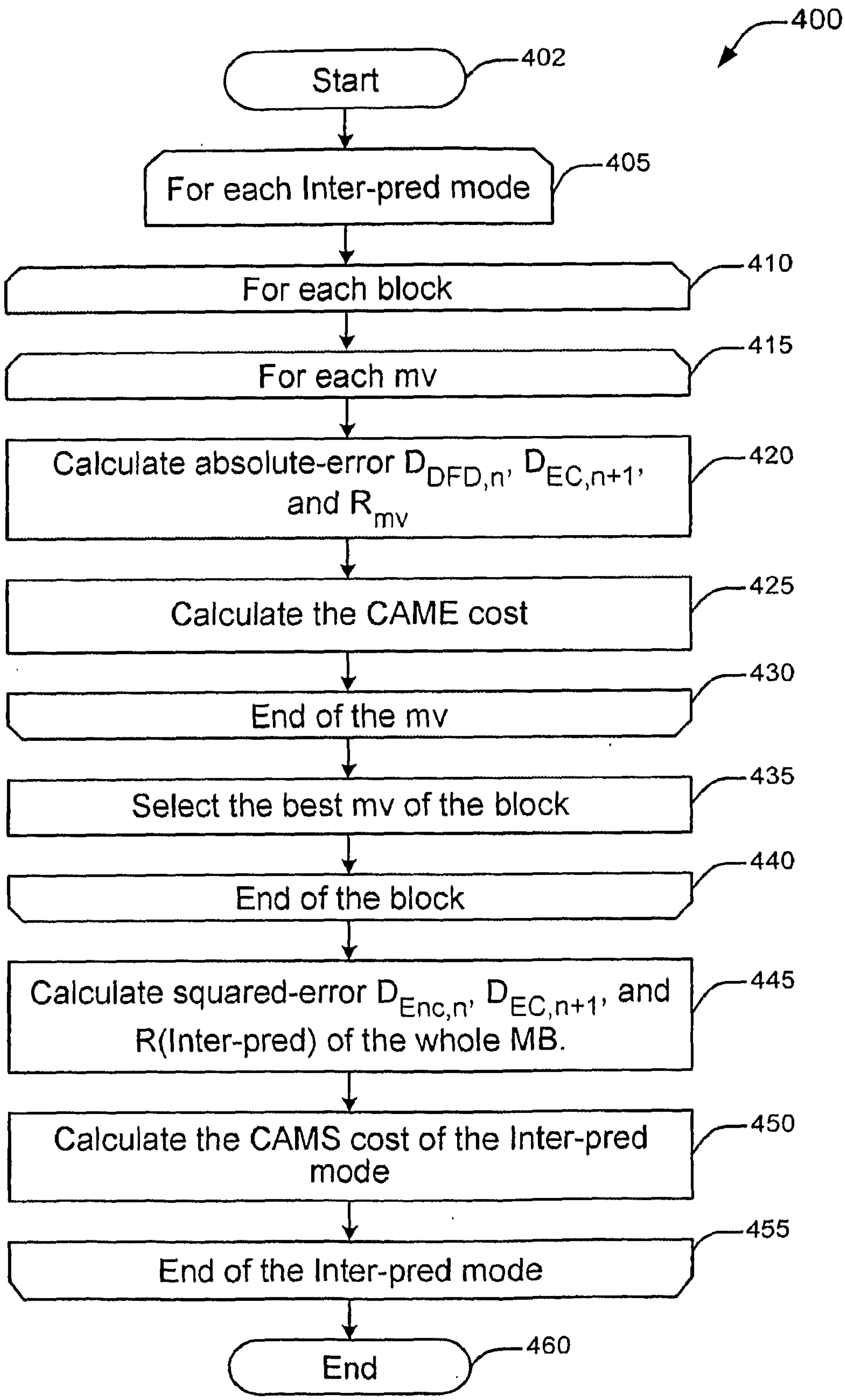

400
Start
402
For each Inter-pred mode
405
For each block
410
For each mv
415
Calculate absolute-error $D_{DFD,n}$, $D_{EC,n+1}$, and R_{mv}
420
Calculate the CAME cost
425
End of the mv
430
Select the best mv of the block
435
End of the block
440
Calculate squared-error $D_{Enc,n}$, $D_{EC,n+1}$, and R(Inter-pred) of the whole MB.
445
Calculate the CAMS cost of the Inter-pred mode
450
End of the Inter-pred mode
455
End
460

FIG. 4

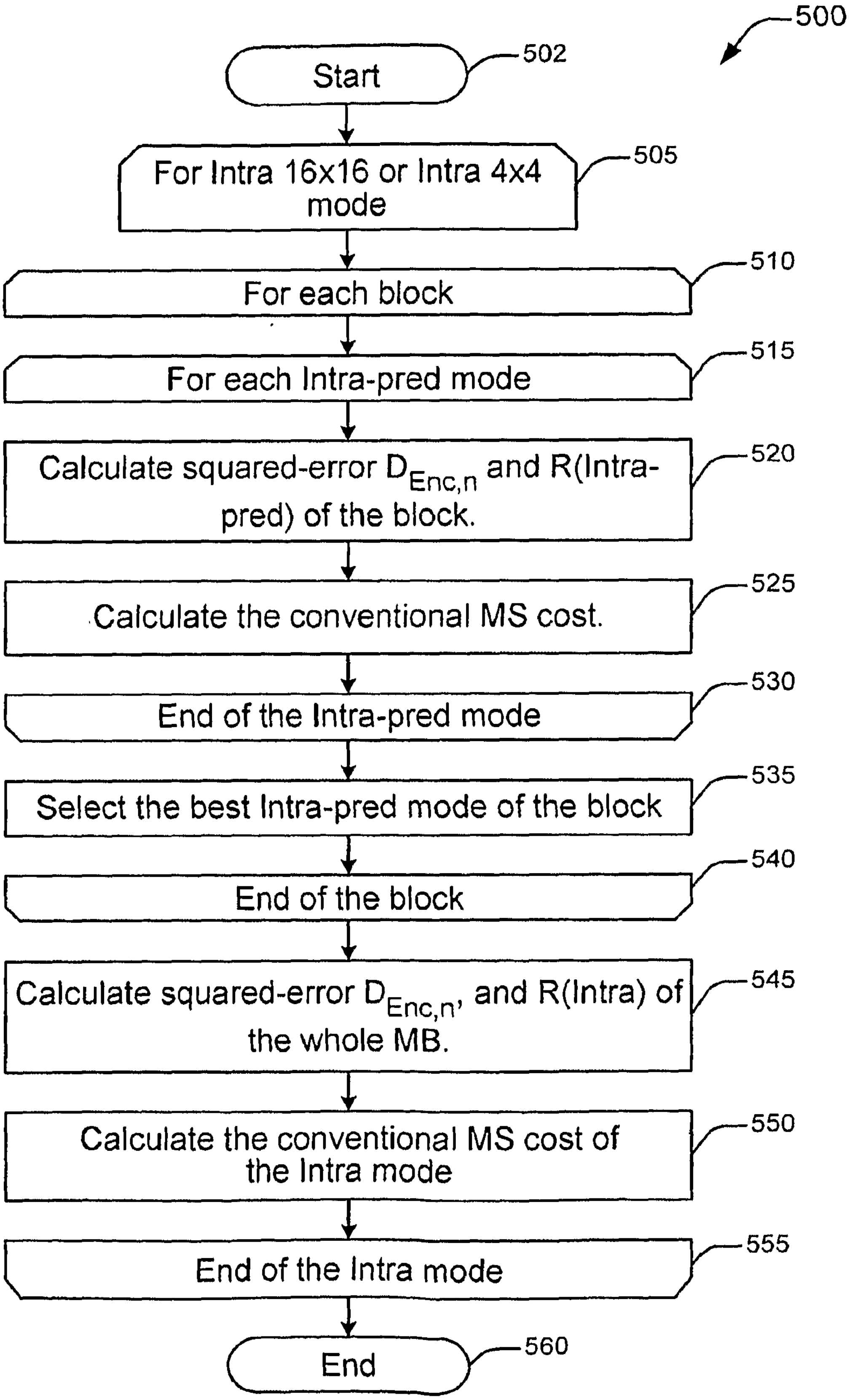

500
Start
502
For Intra 16x16 or Intra 4x4 mode
505
For each block
510
For each Intra-pred mode
515
Calculate squared-error $D_{Enc,n}$ and R(Intra-pred) of the block.
520
Calculate the conventional MS cost.
525
End of the Intra-pred mode
530
Select the best Intra-pred mode of the block
535
End of the block
540
Calculate squared-error $D_{Enc,n}$, and R(Intra) of the whole MB.
545
Calculate the conventional MS cost of the Intra mode
550
End of the Intra mode
555
End
560

FIG. 5