

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
24 December 2008 (24.12.2008)

PCT

(10) International Publication Number
WO 2008/157072 A1

(51) International Patent Classification:
H04N 5/21 (2006.01)

(21) International Application Number:
PCT/US2008/066005

(22) International Filing Date: 6 June 2008 (06.06.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
11/765,029 19 June 2007 (19.06.2007) US

(71) Applicant (for all designated States except US): **MICROSOFT CORPORATION** [US/US]; One Microsoft Way, Redmond, WA 98052-6399 (US).

(72) Inventors: **ZHANG, Cha**; Microsoft Corporation, One Microsoft Way, Redmond, Washington 98052-6399 (US). **ZHANG, Zhengyou**; Microsoft Corporation, One Microsoft Way, Redmond, Washington 98052-6399 (US). **LIU, Zicheng**; Microsoft Corporation, One Microsoft Way, Redmond, Washington 98052-6399 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA,

CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, 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, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))
- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))

Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

(54) Title: VIDEO NOISE REDUCTION

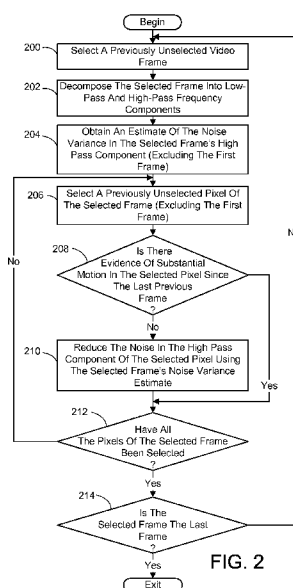


FIG. 2

(57) Abstract: A video noise reduction technique is presented. Generally, the technique involves first decomposing each frame of the video into low-pass and high-pass frequency components. Then, for each frame of the video after the first frame, an estimate of a noise variance in the high pass component is obtained. The noise in the high pass component of each pixel of each frame is reduced using the noise variance estimate obtained for the frame under consideration, whenever there has been no substantial motion exhibited by the pixel since the last previous frame. Evidence of motion is determined by analyzing the high and low pass components.

WO 2008/157072 A1

VIDEO NOISE REDUCTION

BACKGROUND

5

[001] Online video conferencing has attracted more and more interest recently due to the improvements in network bandwidth, computer performance and compression technologies. Webcams are getting popular with home users, allowing them to chat with families and friends over the Internet for free. However, the quality of webcams varies significantly from one to another. Beside common issues caused by poor lighting conditions and sensor color drifting/desaturation, low-end webcams often exhibit significant sensor noises. These noises are visually annoying, and can have negative impact on many post processing algorithms such as compression, object tracking/recognition, among others.

15

SUMMARY

[002] It should also be noted that this Summary is provided to introduce a selection of concepts, in a simplified form, that are further described below in the Detailed Description. This Summary is not intended to identify key features or essential features of the claimed subject matter, nor is it intended to be used as an aid in determining the scope of the claimed subject matter.

[003] The present technique for video noise reduction generally involves first decomposing each frame of the video into low-pass and high-pass frequency components. Then, for each frame of the video after the first frame, an estimate of the noise variance in the high pass component is obtained. This is followed by reducing the noise in the high pass component of each pixel for each frame using its obtained noise variance estimate whenever there has been no substantial motion exhibited by the pixel since the last previous frame. Evidence of motion is determined by analyzing the high and low pass components.

[004] The foregoing general video noise reduction technique can also be implemented in schemes where motion compensation techniques are employed to produce reference and predicted frames (and sometimes bi-directional frames). These techniques involve the use of motion vectors that can be used to assist in the noise reduction process. Still further, in cases where the transmittal of the frames of a video can be delayed, or the video can be processed offline, a backward updating operation based on the present technique can be applied to eliminate possible ghost effects caused by occluded regions.

[005] It is noted that while the foregoing limitations in existing video systems described in the Background section can be resolved by a particular implementation of the present video noise reduction technique according to the present invention, this is in no way limited to implementations that just solve any or all of the noted disadvantages. Rather, the present technique has a much wider application as will become evident from the descriptions to follow.

[006] In addition to the just described benefits, other advantages of the present invention will become apparent from the detailed description which follows hereinafter when taken in conjunction with the drawing figures which accompany it.

DESCRIPTION OF THE DRAWINGS

[007] The specific features, aspects, and advantages of the present invention will become better understood with regard to the following description, appended claims, and accompanying drawings where:

[008] FIG. 1 is a diagram depicting a general purpose computing device constituting an exemplary system for implementing the present invention.

[009] FIG. 2 is a flow diagram generally outlining one embodiment of a process for reducing noise in the frames of a video in accordance with the present video noise reduction technique.

[0010] FIG. 3 is a flow diagram generally outlining one embodiment of a process for reducing noise in the high pass component of a video frame on a pixel by pixel basis when no substantial motion is exhibited by the pixel as part of the present video noise reduction technique.

5

[0011] FIG. 4 is a flow diagram generally outlining one embodiment of a process for computing the noise variance estimate for each frame of a video as part of the present video noise reduction technique.

10 [0012] FIGS. 5A-B are a continuing flow diagram generally outlining one embodiment of a process for reducing noise in the frames of a video encoded using motion compensation in accordance with the present video noise reduction technique.

15 [0013] FIGS. 6A-C are a continuing flow diagram generally outlining one embodiment of a process for reducing noise in the frames of a video, which includes backward updating to alleviate ghosting caused by occlusions in accordance with the present video noise reduction technique.

20 [0014] FIG. 7 is a flow diagram generally outlining one embodiment of a process for backward updating which can be employed as part of the present video noise reduction technique.

25

DETAILED DESCRIPTION

[0015] In the following description of embodiments of the present invention reference is made to the accompanying drawings which form a part hereof, and in which are shown, by way of illustration, specific embodiments in which the
30 invention may be practiced. It is understood that other embodiments may be utilized and structural changes may be made without departing from the scope of the present invention.

1.0 The Computing Environment

[0016] Before providing a description of embodiments of the present video noise reduction technique, a brief, general description of a suitable computing

environment in which portions thereof may be implemented will be described.

The present technique is operational with numerous general purpose or special purpose computing system environments or configurations. Examples of well

known computing systems, environments, and/or configurations that may be suitable include, but are not limited to, personal computers, server computers,

hand-held or laptop devices, multiprocessor systems, microprocessor-based systems, set top boxes, programmable consumer electronics, network PCs, minicomputers, mainframe computers, distributed computing environments that include any of the above systems or devices, and the like.

[0017] Fig. 1 illustrates an example of a suitable computing system environment.

The computing system environment is only one example of a suitable computing environment and is not intended to suggest any limitation as to the scope of use or functionality of the present video noise reduction technique. Neither should the computing environment be interpreted as having any dependency or requirement

relating to any one or combination of components illustrated in the exemplary operating environment. With reference to Fig. 1, an exemplary system for

implementing the present technique includes a computing device, such as computing device 100. In its most basic configuration, computing device 100 typically includes at least one processing unit 102 and memory 104. Depending

on the exact configuration and type of computing device, memory 104 may be volatile (such as RAM), non-volatile (such as ROM, flash memory, etc.) or some combination of the two. This most basic configuration is illustrated in Fig. 1 by dashed line 106. Additionally, device 100 may also have additional

features/functionality. For example, device 100 may also include additional

storage (removable and/or non-removable) including, but not limited to, magnetic or optical disks or tape. Such additional storage is illustrated in Fig. 1 by

removable storage 108 and non-removable storage 110. Computer storage media includes volatile and nonvolatile, removable and non-removable media implemented in any method or technology for storage of information such as

computer readable instructions, data structures, program modules or other data. Memory 104, removable storage 108 and non-removable storage 110 are all examples of computer storage media. Computer storage media includes, but is not limited to, RAM, ROM, EEPROM, flash memory or other memory technology, CD-ROM, digital versatile disks (DVD) or other optical storage, magnetic cassettes, magnetic tape, magnetic disk storage or other magnetic storage devices, or any other medium which can be used to store the desired information and which can be accessed by device 100. Any such computer storage media may be part of device 100.

[0018] Device 100 may also contain communications connection(s) 112 that allow the device to communicate with other devices. Communications connection(s) 112 is an example of communication media. Communication media typically embodies computer readable instructions, data structures, program modules or other data in a modulated data signal such as a carrier wave or other transport mechanism and includes any information delivery media. The term “modulated data signal” means a signal that has one or more of its characteristics set or changed in such a manner as to encode information in the signal. By way of example, and not limitation, communication media includes wired media such as a wired network or direct-wired connection, and wireless media such as acoustic, RF, infrared and other wireless media. The term computer readable media as used herein includes both storage media and communication media.

[0019] Device 100 may also have input device(s) 114 such as keyboard, mouse, pen, voice input device, touch input device, etc. Output device(s) 116 such as a display, speakers, printer, etc. may also be included. All these devices are well known in the art and need not be discussed at length here.

[0020] Of particular note is that device 100 can include a camera 118 (such as a digital/electronic still or video camera, or film/photographic scanner), which is capable of capturing a sequence of images, as an input device. Further, multiple cameras 118 could be included as input devices. The images from the one or more cameras are input into the device 100 via an appropriate interface (not shown). However, it is noted that image data can also be input into the device

100 from any computer-readable media as well, without requiring the use of a camera.

[0021] The present video noise reduction technique may be described in the
5 general context of computer-executable instructions, such as program modules,
being executed by a computing device. Generally, program modules include
routines, programs, objects, components, data structures, etc. that perform
particular tasks or implement particular abstract data types. The present
technique may also be practiced in distributed computing environments where
10 tasks are performed by remote processing devices that are linked through a
communications network. In a distributed computing environment, program
modules may be located in both local and remote computer storage media
including memory storage devices.

[0022] The exemplary operating environment having now been discussed, the
15 remaining parts of this description section will be devoted to a description of the
program modules embodying the present video noise reduction technique.

2.0 The Video Noise Reduction Technique

[0023] The present technique provides an effective and efficient way to reduce
20 the noise in the frames of a video. After noise reduction, the video can be fed into
a video encoder where it has been found there can be saving in bit rate of around
5-75% depending on the original video noise level.

[0024] In general, the present technique decomposes videos into low and high
25 frequency components, and reduces the noise in the high-pass component. In
one embodiment, this is implemented as shown in Fig. 2. The procedure begins
by selecting a frame of the video (200) and then decomposing the selected frame
30 into low-pass and high-pass frequency components (202). In addition, for each
frame after the first, an estimate of the noise variance in the high pass component
of the selected frame is obtained (204). A previously unselected pixel of the
selected frame (excluding the first) is then selected (206), and it is determined if
there is evidence of any substantial motion associated with the selected pixel

since the last previous frame (208). Generally this can be accomplished by analyzing the high pass component of the selected pixel in the current frame, and the low pass component of the selected pixel in the current and previous frame. If substantial motion has not occurred, the noise is reduced in the high pass component of the selected pixel using the noise variance estimate (210). Otherwise the high pass component of the selected pixel is not changed. It is then determined if all the pixels of the selected frame have been selected (212). If not, actions 206 through 212 are repeated. When all the pixels have been considered, it is determined if the selected frame is the last frame (214). If it is not, the procedure is repeated by repeating actions 200 through 214. If it is the last frame, the procedure ends.

[0025] As noted, the foregoing technique can employ the low pass components in the current and preceding frame in order to determine if motion has occurred and so whether a particular pixel under consideration will undergo noise reduction. This unique aspect of the technique, among others, will be described in more detail in the sections to follow.

[0026] In addition, it is noted that the foregoing technique employs a unique and simple model for automatically estimating the noise variance in the high pass component. This will also be described in more detail in the sections to follow.

2.1 Decomposing A Video Frame

[0027] In one embodiment of the present video noise reduction technique, each frame of the video is decomposed into low-pass and high-pass frequency components by first generating the low pass component $I_{low}(t)$ for each pixel using a low pass filter (where t refers to the frame number). For example, a box filter or Gaussian filter could be used. The high pass component $I_{high}(t)$ for each pixel is then generated by subtracting the low pass component of the pixel from the original pixel of the video frame under consideration. Thus,

$$I_{high}(t) = I(t) - I_{low}(t). \quad (1)$$

2.2 Finding Evidence Of Substantial Motion

[0028] As indicated previously, noise reduction is applied to a pixel in a frame only when there has not been substantial motion exhibited by that pixel based in part on the pixel's condition in the immediately preceding frame. This motion determination is generally accomplished on a pixel-by-pixel basis using both the low pass and high pass components by determining if two conditions are met. First, it is determined whether the pixel under consideration satisfies a condition that the absolute value of the difference between the high pass component pixel value and an initial value of I_{est} is less than a first prescribed multiple of the noise variance estimate. I_{est} represents an estimate of the high pass component intensity of the pixel under consideration as it would exist without noise. The aforementioned first prescribed multiple can generally range between about 3 and 5, however, a value of 5 was used with success in tested embodiments of the present technique. Thus, for tested embodiments, the first condition involved whether:

$$|I_{high}(t) - I_{est}(t-1)| < 5\delta, \quad (2)$$

where δ is the aforementioned noise variance estimate.

[0029] The second condition involves, for the pixel under consideration, whether the absolute value of the difference between the low pass component associated with the current frame and the low pass component computed for the frame immediately preceding the current frame is less than a second prescribed multiple of the noise variance estimate. The second prescribed multiple can generally range between about 0.5 and 2.0, however, a value of 1.25 was used with success in tested embodiments of the present technique. Thus, for tested embodiments, the second condition involved whether:

$$|I_{low}(t) - I_{low}(t-1)| < 1.25\delta, \quad (3)$$

where $I_{low}(t-1)$ refers to the low pass component of the correspondingly located pixel in the frame coming just before the current frame being processed.

[0030] The first condition ensures the noise associated with a pixel is not too large as to make the motion determination unfeasible. The second condition ensures there has been no big intensity changes, which would be indicative of motion.

[0031] Each pixel in each frame after the first frame ($t=0$) of the video is tested using the foregoing two conditions. For the second frame ($t=1$) of the video, I_{est} is initialized as the high pass component of the correspondingly-located pixel in the first frame (i.e., $I_{est} = I_{high}(t=0)$). As will be described shortly, it may be refined as part of the noise reduction procedure. For all subsequent frames, the initial I_{est} for a pixel is set to the final I_{est} of the correspondingly-located pixel of the immediately preceding frame, and is used to determine if the foregoing conditions are satisfied. In the preceding equations, as well as those to follow, the initial I_{est} for a frame will be designated as $I_{est}(t-1)$ since it comes from the immediately preceding frame ($t-1$). Likewise, the final I_{est} for a frame will be designated as $I_{est}(t)$.

[0032] As for the noise variance estimate δ , it is initialized for use in the second frame as a prescribed value. This value can range between about 3 and 20, depending on the prior knowledge about the noise level. However, in tested embodiments a value of 10 was employed with success. For all subsequent frames, δ is computed as will be described shortly. It is noted that a single estimate of δ is used for all the pixels of a frame for efficiency.

[0033] If a pixel satisfies both the foregoing conditions it is deemed that no substantial motion has occurred since the immediately preceding frame. Otherwise, it is deemed that substantial motion has occurred.

2.3 Reducing The Noise

[0034] If a pixel exhibits evidence of substantial motion, the noise associated with the high pass component of that pixel cannot be reduced. However, if no substantial motion is evident, this noise can be reduced. Given this, the following

technique can be used to reduce the noise in the high pass component of a pixel under consideration using the noise variance estimate.

[0035] Referring to Fig. 3, for each pixel of a video frame under consideration, in one embodiment it is first determined whether both of the previously-described conditions are satisfied by the pixel (300). If so, the I_{est} is set to a value representing the product of the initial I_{est} and a prescribed weighting factor α , added to the product of one minus the weighting factor and the high pass component pixel value (302). Thus, for each pixel,

$$I_{esti}(t) = \alpha I_{est}(t-1) + (1-\alpha) I_{high}(t). \quad (4)$$

[0036] If either condition is not met, then the final I_{est} for the pixel under consideration is set to the high pass component pixel value $I_{high}(t)$ of the pixel (304). The value of the pixel under consideration is then set to a pixel value representing the sum of the final I_{est} and the low pass component pixel value (306). Thus, for each pixel in the frame under consideration,

$$I'(t) = I_{est}(t) + I_{low}(t), \quad (5)$$

where $I'(t)$ is a noise reduced pixel value.

[0037] In one embodiment of the present video noise reduction technique, the prescribed weighting factor α is set as $N/(N+1)$, where N is number of consecutive frames back from and including the current frame where the pixel satisfies both the aforementioned first and second conditions (i.e., exhibited no substantial motion). In an alternate embodiment a prescribed fixed value of the weighting factor is used for convenience. For example, in tested embodiments, a fixed α value of $7/8$ was employed for the weighting factor.

2.4 Estimating The Noise Variance

[0038] As indicated previously, a noise variance estimate is obtained for each frame considered after the first frame of the video, and this estimate is used in the

calculation associated with the next frame in determining whether a pixel satisfies the aforementioned first and second conditions. As indicated previously, the estimate is obtained from the immediately preceding frame. The noise variance is estimated for the immediately preceding frame as follows.

5

[0039] In one embodiment of the present technique, the noise variance estimate is computed for each frame (after the first) as shown in Fig. 4. First, a previously unselected pixel of the frame under consideration is selected (400). It is then determined if the selected pixel satisfies both the first and second conditions (402). If not, the next pixel is considered by repeating actions 400 and 402. If the pixel does satisfy both conditions, then the square of the difference between the I_{est} and the high pass component value for the selected pixel is computed (404). It is then determined if all the pixels of the frame under consideration have been selected and processed (406). If not, then actions 400 through 406 are repeated. When all the pixels have been considered as described above, a mean square difference is computed from the squares of the differences computed for the pixels satisfying both the first and second conditions (408). The noise variance estimate for the frame under consideration is then set to a value representing the square root of the mean square difference (410). Thus,

20

$$\delta = \text{sqrt}(\text{mean} \{(I_{est} - I_{high}(t))^2\}) \quad (6)$$

3.0 The Video Noise Reduction Technique With Motion Information Available

25

[0040] Modern video codecs make use of similarity across frames to improve coding efficiency, often termed motion compensation. Generally speaking, in motion compensation there are reference frames and predicted frames. Predicted frames are first predicted from the reference frames using motion vectors, and the residue is then coded. If motion vectors are available, it is possible to extend the present technique to provide noise reduction in both the reference and predicted frames of a video. In general, the technique is the same except that the I_{est} from the reference frames is used to construct an initial I_{est} for its associated predicted frames based on the motion vectors.

30

[0041] In one embodiment, this is implemented as shown in Figs. 5A-B. First, a frame of the video is selected (500) and then decomposed into low-pass and high-pass frequency components (502). In addition, for each frame after the first reference frame, an estimate of the noise variance in the high pass component of the selected frame is obtained (504). A previously unselected pixel of the selected frame (excluding the first reference frame) is then selected (506), and it is determined if the selected pixel satisfies a first condition that the absolute value of the difference between the high pass component pixel value and an initial value of I_{est} is less than a first prescribed multiple of the noise variance estimate (508). If not, the I_{est} is set to the high pass component value of the selected pixel (510). If the first condition is satisfied, then it is determined whether the selected pixel satisfies a second condition that the absolute value of the difference between the low pass component associated with the frame currently under consideration and the low pass component computed for the frame immediately preceding the current frame is less than a second prescribed multiple of the noise variance estimate (512). If not, the I_{est} is set to the high pass component value of the selected pixel (510). However, if the second condition is also satisfied, the I_{est} for the selected pixel is set to the product of the initial I_{est} and a prescribed weighting factor, added to the product of one minus the weighting factor and the high pass component pixel value (514). In this case, the initial I_{est} for a pixel of the first prediction frame of the video coming immediately after the first reference frame is set to the high pass component pixel value of the pixel in the first reference frame identified as corresponding to the pixel of the first prediction frame via the set of motion vectors associated with it. In addition, the initial I_{est} for a pixel of each reference frame after the first reference frame is set to the I_{est} of the correspondingly-located pixel of the immediately preceding prediction frame. And finally, the initial I_{est} for a pixel of each prediction frame coming after the first prediction frame is set to the I_{est} of the pixel in its associated reference frame identified as corresponding to the pixel of the prediction frame under consideration via the set of motion vectors associated with it. Once the I_{est} for a pixel has been computed, the current value of the selected pixel is replaced with the sum of the pixel's I_{est} and the pixel's low pass component value (516). It is then determined if all the pixels of the selected frame have been selected (518). If not, actions 506 through 518 are repeated. When all the pixels have been considered, it is

determined if the selected frame is the last frame (520). If it is not, the procedure is repeated by repeating actions 500 through 520. If it is the last frame, the procedure ends.

5 [0042] The foregoing version of the present technique is also applicable to video codecs employing bi-directional frames that are predicted from a reference frame preceding the bidirectional frame and a prediction frame coming after the bidirectional frame. This prediction is accomplished by combining a set of motion vectors from both the associated reference and prediction frames. The only
10 difference in the present procedure is how the initial I_{est} is set for a pixel of a bi-directional frame. More particularly, the initial I_{est} for a pixel of the first bidirectional frame of the video coming immediately after the first reference frame is computed as the combination of the high pass component pixel value of the pixel in the first reference frame identified as corresponding to the pixel of the first bidirectional
15 frame via the set of motion vectors associated with the first reference frame and the I_{est} of the pixel of the associated prediction frame identified as corresponding to the pixel of the first bidirectional frame via the set of motion vectors associated with the prediction frame. Similarly, the initial I_{est} for a pixel of each bidirectional frame coming after the first bidirectional frame is computed as a combination of
20 the I_{est} of the pixel in the associated reference frame identified as corresponding to the pixel of the bidirectional frame under consideration via the set of motion vectors associated with the reference frame and the I_{est} of the pixel in the associated prediction frame identified as corresponding to the pixel of the bidirectional frame under consideration via the set of motion vectors associated
25 with the prediction frame.

4.0 The Video Noise Reduction Technique With Backwards Updating

[0043] If the transmittal of the frames of a video can be delayed a number of
30 frames, or if a completed video can be processed offline, it is possible to enhance the previously described embodiments of the present technique to eliminate possible ghost effects due to occluded regions. For example, assume a background wall is partially occluded by a person. If the person leaves, the previously-described first and second conditions would not be satisfied at the

instance un-occlusion occurs, owing to the motion. Accordingly, no noise reduction would take place in the affected frame(s). In addition, the averaging aspects of computing an I_{est} for a pixel would take a few frames to take full effect. During this period, the un-occluded region would be slightly noisier than the parts that were never occluded, creating a ghosting effect. Performing a backward updating of the I_{est} values for the aforementioned period can alleviate this problem.

[0044] In one embodiment of the present video noise reduction technique, the backward updating is implemented as shown in Figs. 6A-C. The noise in the frames of the video is reduced in the forward direction as described in the previous embodiments. More particularly, the procedure begins by selecting a frame of the video (600) and then decomposing the selected frame into low-pass and high-pass frequency components (602). An estimate of the noise variance in the high pass component of the selected frame is obtained for each frame after the first (604). A previously unselected pixel of the selected frame (excluding the first) is then selected (606), and it is determined if the selected pixel satisfies a first forward updating condition that the absolute value of the difference between the high pass component pixel value and an initial value of I_{est} is less than a first prescribed multiple of the noise variance estimate (608). If not, the I_{est} is set to the high pass component value of the selected pixel (610). If the first forward updating condition is satisfied, then it is determined whether the selected pixel satisfies a second forward updating condition that the absolute value of the difference between the low pass component associated with the frame currently under consideration and the low pass component computed for the corresponding pixel in the frame immediately preceding the current frame is less than a second prescribed multiple of the noise variance estimate (612). If not, the I_{est} is set to the high pass component value of the selected pixel (610). However, if the second forward updating condition is also satisfied, the I_{est} for the selected pixel is set to the product of the initial I_{est} and a prescribed weighting factor, added to the product of one minus the weighting factor and the high pass component pixel value (614). Once the I_{est} has been computed, the current value of the selected pixel is replaced with the sum of the pixel's I_{est} and the pixel's low pass component value (616).

[0045] It is then determined whether the selected pixel, or a correspondingly-located pixel in any of a prescribed number of frames plus one in the sequence of frames preceding the selected frame, failed to satisfy either the first or second forward updating conditions (618). If so, the selected pixel is designated as a candidate for backward updating (620). If not, no action is taken. Either way, it is also determined whether the selected pixel has a correspondingly-located pixel in any of the prescribed number of frames in the sequence of frames preceding the current frame that is designated as a backward updating candidate (622). If it does not, then no backwards updating is necessary and the procedure skips to action 632. However, if it does, it is then determined if the selected pixel has satisfied a first backward updating condition in that it has satisfied both the first and second forward updating conditions and the correspondingly-located pixel in each of the prescribed number of frames minus one in the sequence of frames preceding the selected frame also satisfied both the first and second forward updating conditions (624). If this is not true, it is determined if the selected pixel has satisfied a second backward updating condition in that it fails to satisfy either the first or second forward updating conditions (626). Otherwise, action 626 is skipped. If it is determined the selected pixel satisfied either the first or second backward updating condition, then for the prescribed number of frames starting with the selected frame and moving back in the sequence of frames preceding the selected frame, the first frame where the selected pixel or a correspondingly-located pixel in a preceding frame satisfied both the first and second forward updating conditions, is identified (628). If neither backward updating condition is satisfied, no backwards updating is done and the procedure skips to action 632. Next, for each frame starting with the identified frame and moving back to each frame preceding the identified frame in turn, up to the prescribed number of frames back from the selected frame or up to the first frame encountered where the correspondingly-located pixel did not satisfy both the first and second forward updating conditions, whichever is first, backward updating is applied to the selected pixel or a correspondingly-located pixel in a frame preceding the selected frame, as the case may be (630). It is then determined if all the pixels of the selected frame have been selected (632). If not, actions 606 through 632 are repeated. When all the pixels in the selected frame have been considered, it is

determined if the selected frame is the last frame (634). If it is not, the procedure is repeated by repeating actions 600 through 634. If it is the last frame, the procedure ends.

5 [0046] As for the prescribed number of frames that the foregoing procedure looks back, in one embodiment, the number is fixed. For example, in tested embodiments, the prescribed number of frames was set to 10 frames. It is noted, however, that each frame included in the prescribed number of frames results in a measurable amount of improvement in a pixel's signal-to-noise ratio. Thus, it is possible in an alternate embodiment to set the prescribed number of frames to the number of frames that will produce the improvement in the signal-to-noise ratio needed to achieve a target ratio.

15 [0047] As for the backward updating task, this is generally just the present video noise reduction technique applied in a reverse order of frames. More particularly, referring to Fig. 7, the backward updating task involves first decomposing the pixel under consideration into low-pass and high-pass frequency components (700). The I_{est} of the pixel is then set to a value representing the product of the pixel's initial I_{est} and the aforementioned prescribed weighting factor, added to the product of one minus the weighting factor and the high pass component pixel value (702). Here again, the initial I_{est} for a pixel is set to the I_{est} of the correspondingly-located pixel of the frame considered immediately before the current frame under consideration. Next, the current pixel value of the pixel is replaced with a pixel value representing the sum of the pixel's new I_{est} and its low pass component pixel value (704).

5.0 Other Embodiments

30 [0048] In the foregoing description of the embodiments of the present video noise reduction technique, the pixel intensity was referred to as if it were a single entity. This is true in the case of gray scale images. However, in color video frames, each pixel intensity is actually a combination of individual color intensities. For example, when a Red-Green-Blue (RGB) color model is used to characterize the pixel color, there is a red color channel intensity, green color channel intensity

and blue color channel intensity. In view of this, when the noise in a color video frame is being reduced, the present technique involves repeating the foregoing noise reduction procedures for each color channel. It is further noted that this repetition of the procedures per color channel also applies to other color models and color space characterizations of the pixel intensities.

[0049] It should also be noted that any or all of the aforementioned embodiments throughout the description may be used in any combination desired to form additional hybrid embodiments. Although the subject matter has been described in language specific to structural features and/or methodological acts, it is to be understood that the subject matter defined in the appended claims is not necessarily limited to the specific features or acts described above. Rather, the specific features and acts described above are disclosed as example forms of implementing the claims.

Wherefore, what is claimed is:

1. A computer-implemented process for reducing noise in a video comprising a sequence of video frames, comprising using a computer to perform the following process actions:

decomposing each frame of the video into low-pass and high-pass frequency components (202),

for each frame of the video after the first,

obtaining an estimate of the noise variance in the high pass component (204), and

on a per pixel basis, reducing the noise in the high pass component using the noise variance estimate (210) whenever there has been no substantial motion since the last previous frame as determined by analyzing the high and low pass components.

2. The process of Claim 1, wherein the process action of reducing the noise in the high pass component using the noise variance estimate on a per pixel basis, comprises an action of determining whether the pixel under consideration satisfies a set of conditions, comprising,

whether the pixel under consideration satisfies a first condition that the absolute value of the difference between the high pass component pixel value and an initial value of I_{est} is less than a first prescribed multiple of the noise variance estimate, wherein I_{est} represents an estimate of the high pass component intensity of the pixel under consideration as it would exist without noise, and

whether the pixel under consideration satisfies a second condition that the absolute value of the difference between the low pass component associated with the frame currently under consideration and the low pass component computed for the frame immediately preceding the current frame is less than a second prescribed multiple of the noise variance estimate.

3. The process of Claim 2, wherein the first prescribed multiple of the noise variance estimate ranges between about 3 and about 5, and the second prescribed multiple of the noise variance estimate ranges between about 0.5 and about 2.0.

4. The process of Claim 2, wherein the first prescribed multiple of the noise variance estimate is 5, and the second prescribed multiple of the noise variance estimate is 1.25.

5

5. The process of Claim 2, wherein the initial I_{est} for a pixel of the second frame of the video is set to the high pass component pixel value of the correspondingly-located pixel in the first frame and the initial I_{est} for a pixel of each frame after the second frame of the video is set to the I_{est} of the correspondingly-located pixel of the immediately preceding frame, and wherein the process action of reducing the noise in the high pass component using the noise variance estimate on a per pixel basis, further comprises the actions of:

10

for each pixel of the video frame under consideration,

15

setting the I_{est} to a value representing the product of the initial I_{est} and a prescribed weighting factor, added to the product of one minus the weighting factor and the high pass component pixel value, whenever the first and second conditions are satisfied, and

20

setting the I_{est} to the high pass component pixel value, whenever either the first or second condition is not satisfied, replacing the current pixel value with a pixel value representing the sum of the I_{est} and the low pass component pixel value.

25

6. The process of Claim 5, wherein the process action of obtaining the estimate of the noise variance in the high pass component for each frame of the video after the first, comprises the actions of:

for each video frame after the first frame,

30

computing the square of the difference between the I_{est} and the high pass component pixel value, for each pixel where the first and second conditions are satisfied,

computing a mean square difference from the squares of the differences computed for the pixels where the first and second conditions were satisfied,

setting the noise variance estimate to a prescribed value for the second frame of the video, and

setting the noise variance estimate to a value representing the square root of said mean square difference computed for the immediately preceding video frame, for each frame after the second frame of the video.

5 7. The process of Claim 6, wherein the prescribed initial value of the current noise variance estimate is 10.

 8. The process of Claim 5, wherein the prescribed weighting factor is set to $N/(N+1)$, wherein N is the number of video frames including the current
10 frame back to the last video frame where the pixel under consideration does not satisfy both the first and second conditions.

 9. The process of Claim 5, wherein the prescribed weighting factor is set to a fixed value of 7/8.

15 10. The process of Claim 1, wherein the process action of decomposing the video frame into low-pass and high-pass frequency components, comprises the actions of:

 generating the low pass component using a low pass filter; and
20 generating the high pass component by subtracting the low pass component from the video frame under consideration.

 11. The process of Claim 10, wherein the low pass filter is one of (i) a box filter, and (ii) a Guassian filter.

25 12. The process of Claim 1, wherein each frame of the video is a color frame and the intensity value of each pixel comprises a combination of the individual intensity of each of a group of color channels, and wherein all the process actions are performed separately on each color channel.

30 13. A computer-implemented process for reducing noise in a video comprising a sequence of video frames, comprising using a computer to perform the following process actions:

decomposing each frame of the video into low-pass and high-pass frequency components (602),

for each frame of the video after the first,

obtaining an estimate of the noise variance in the high pass component (604), and

on a per pixel basis, reducing the noise in the high pass component using the noise variance estimate whenever the pixel under consideration satisfies a set of conditions comprising,

a first forward updating condition that the absolute value of the difference between the high pass component pixel value and an initial value of I_{est} is less than a first prescribed multiple of the noise variance estimate (608), wherein I_{est} represents an estimate of the high pass component intensity of the pixel under consideration as it would exist without noise, and

a second forward updating condition that the absolute value of the difference between the low pass component associated with the frame currently under consideration and the low pass component computed for the frame immediately preceding the current frame is less than a second prescribed multiple of the noise variance estimate (612).

14. The process of Claim 13, further comprising a process action of performing a backward updating operation for pixels in a group of frames, wherein the backward updating operation comprises the actions of:

whenever a pixel under consideration in the current frame under consideration fails to satisfy either the first or second forward updating conditions or a correspondingly-located pixel in any of a prescribed number of frames plus one in the sequence of frames preceding the current frame under consideration fails to satisfy either the first or second forward updating conditions, designating the pixel under consideration as a candidate for backward updating; and

whenever a pixel under consideration in the current frame under consideration has a correspondingly-located pixel in any of the prescribed number of frames in the sequence of frames preceding the current frame that is designated as a backward updating candidate,

determining if the pixel under consideration has satisfied a first backward updating condition in that it has satisfied both the first and second

forward updating conditions and the correspondingly-located pixel in each of the prescribed number of frames minus one in the sequence of frames preceding the current frame under consideration also satisfied both the first and second forward updating conditions,

5 whenever the pixel under consideration fails to satisfy the first backward updating condition, determining if the pixel under consideration has satisfied a second backward updating condition in that it fails to satisfy either the first or second forward updating conditions, and

 whenever the pixel under consideration satisfies either the
10 first or second backward updating condition,

 identifying the prescribed number of frames starting with the current frame and moving backward in the sequence of frames preceding the current frame,

 identifying the first frame wherein the pixel under
15 consideration or a correspondingly-located pixel in a preceding frame satisfied both the first and second forward updating conditions, in the identified prescribed number of frames, and

 applying backward updating to the pixel under
consideration or a correspondingly-located pixel in a frame preceding the current
20 frame, as the case may be, for the identified first frame and each frame preceding the identified first frame in the identified prescribed number of frames, in turn, up to the first frame encountered where the correspondingly-located pixel did not satisfy both the first and second forward updating conditions, wherein backward updating of the pixel under consideration or a correspondingly-located pixel in a
25 frame preceding the current frame comprises,

 decomposing the pixel under consideration into low-pass and high-pass frequency components,

 setting the I_{est} of the pixel under consideration to a value representing the product of an initial I_{est} and a prescribed weighting factor,
30 added to the product of one minus the weighting factor and the high pass component pixel value, wherein the initial I_{est} for a pixel is set to the I_{est} of the correspondingly-located pixel of the frame immediately preceding the current frame under consideration, and

replacing the current pixel value of the pixel under consideration with a pixel value representing the sum of the I_{est} and the low pass component pixel value.

5 15. The process of Claim 14, wherein the prescribed weighting factor is set to a fixed value of $7/8$.

 16. The process of Claim 14, wherein the prescribed number of frames is set to 10 frames.

10

 17. The process of Claim 14, wherein the prescribed weighting factor is set to $N/(N+1)$, wherein N is the number of video frames including the current frame back to the last video frame where the pixel under consideration does not satisfy both the first and second conditions.

15

 18. The process of Claim 16, wherein each frame included in the prescribed number of frames results in a known amount of improvement in a pixel's signal-to-noise ratio, and wherein the prescribed number of frames is set to the number of frames that will produce improvement in the signal-to-noise ratio
20 needed to achieve a target ratio.

20

 19. A computer-implemented process for reducing noise in a video comprising a sequence of video frames comprising reference frames and predicted frames that are predicted from an associated reference frame using a set of motion vectors and residue values, comprising using a computer to perform
25 the following process actions:

 decomposing each frame into low-pass and high-pass frequency components (502);

 for each frame after the first reference frame,

30

 obtaining an estimate of the noise variance in the high pass component (504), and

 on a per pixel basis, determining whether the pixel under consideration satisfies a set of conditions comprising,

a first condition that the absolute value of the difference between the high pass component pixel value and an initial value of I_{est} is less than a first prescribed multiple of the noise variance estimate (508),

wherein I_{est} represents an estimate of the high pass component intensity of the pixel under consideration as it would exist without noise, and

a second condition that the absolute value of the difference between the low pass component associated with the frame currently under consideration and the low pass component computed for the frame immediately preceding the current frame is less than a second prescribed multiple of the noise variance estimate (512), and

for each pixel of the video frame under consideration,

setting the I_{est} to a value representing the product of the initial I_{est} and a prescribed weighting factor, added to the product of one minus the weighting factor and the high pass component pixel value, whenever the first and second conditions are satisfied (514), wherein the initial I_{est} for a pixel of a first prediction frame of the video coming immediately after the first reference frame is set to the high pass component pixel value of the pixel in the first reference frame identified as corresponding to the pixel of the first prediction frame via the set of motion vectors associated therewith, and the initial I_{est} for a pixel of each reference frame after the first reference frame is set to the I_{est} of the correspondingly-located pixel of the immediately preceding prediction frame, and the initial I_{est} for a pixel of each prediction frame coming after the first prediction frame is set to the I_{est} of the pixel in its associated reference frame identified as corresponding to the pixel of the prediction frame under consideration via the set of motion vectors associated therewith;

setting the I_{est} to the high pass component pixel value, whenever either the first or second condition is not satisfied (510), and

replacing the current pixel value with a pixel value representing the sum of the I_{est} and the low pass component pixel value (516).

20. The process of Claim 19, wherein the sequence of video frames further comprises bidirectional frames that are predicted from an associated reference frame preceding the bidirectional frame and an associated prediction frame coming after the bidirectional frame, using a set of motion vectors from both

the associated reference and prediction frames, and wherein the process action of setting, for each pixel of the video frame under consideration, the I_{est} to a value representing the product of an initial I_{est} and a prescribed weighting factor, added to the product of one minus the weighting factor and the high pass component pixel value, whenever the first and second conditions are satisfied, further
5 comprises, in the case of a bidirectional frame:

computing the initial I_{est} for a pixel of a first bidirectional frame of the video coming immediately after the first reference frame as a combination of the high pass component pixel value of the pixel in the first reference frame identified
10 as corresponding to the pixel of the first bidirectional frame via the set of motion vectors associated with the first reference frame and the I_{est} of the pixel of the associated prediction frame identified as corresponding to the pixel of the first bidirectional frame via the set of motion vectors associated with the prediction frame; and

15 computing the initial I_{est} for a pixel of each bidirectional frame coming after the first bidirectional frame as a combination of the I_{est} of the pixel in the associated reference frame identified as corresponding to the pixel of the bidirectional frame under consideration via the set of motion vectors associated with the reference frame and the I_{est} of the pixel in the associated prediction frame
20 identified as corresponding to the pixel of the bidirectional frame under consideration via the set of motion vectors associated with the prediction frame.

25

1/9

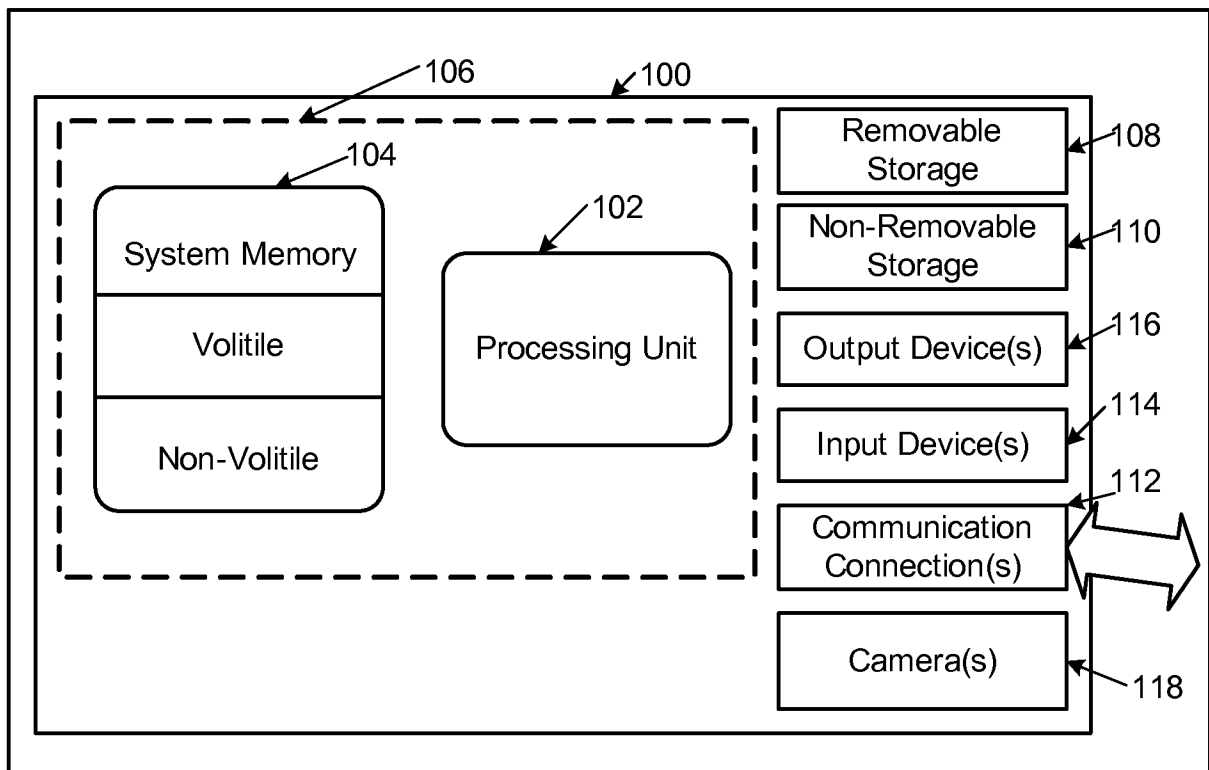


FIG. 1

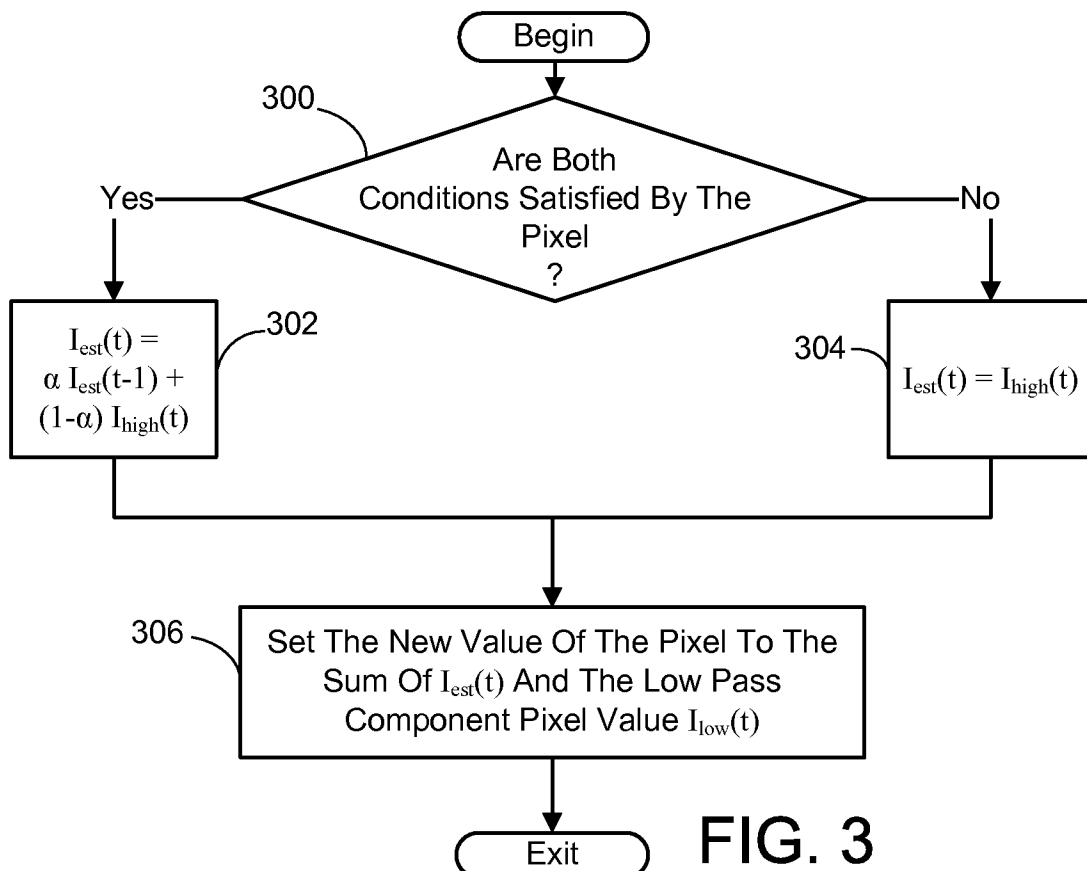


FIG. 3

2/9

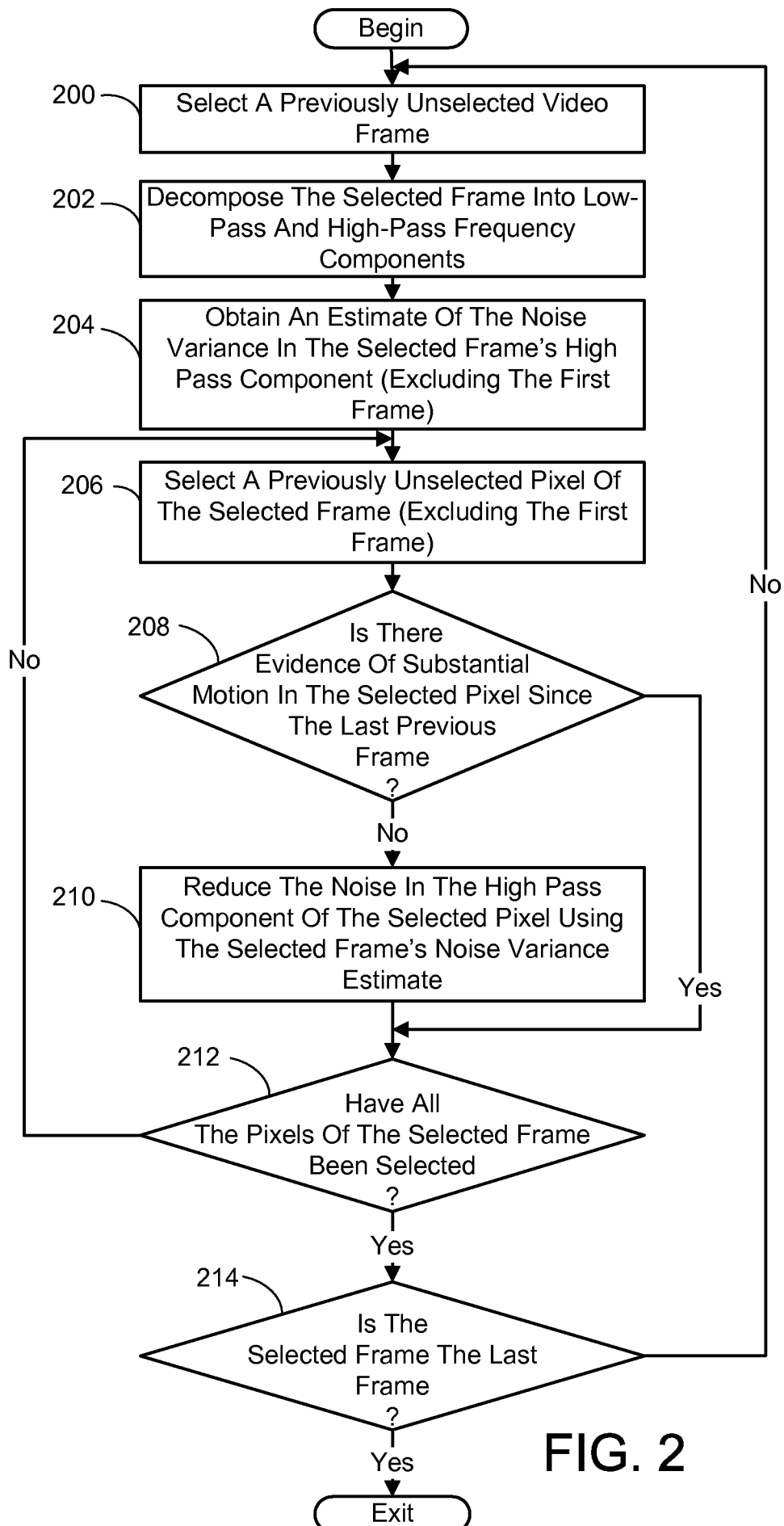


FIG. 2

3/9

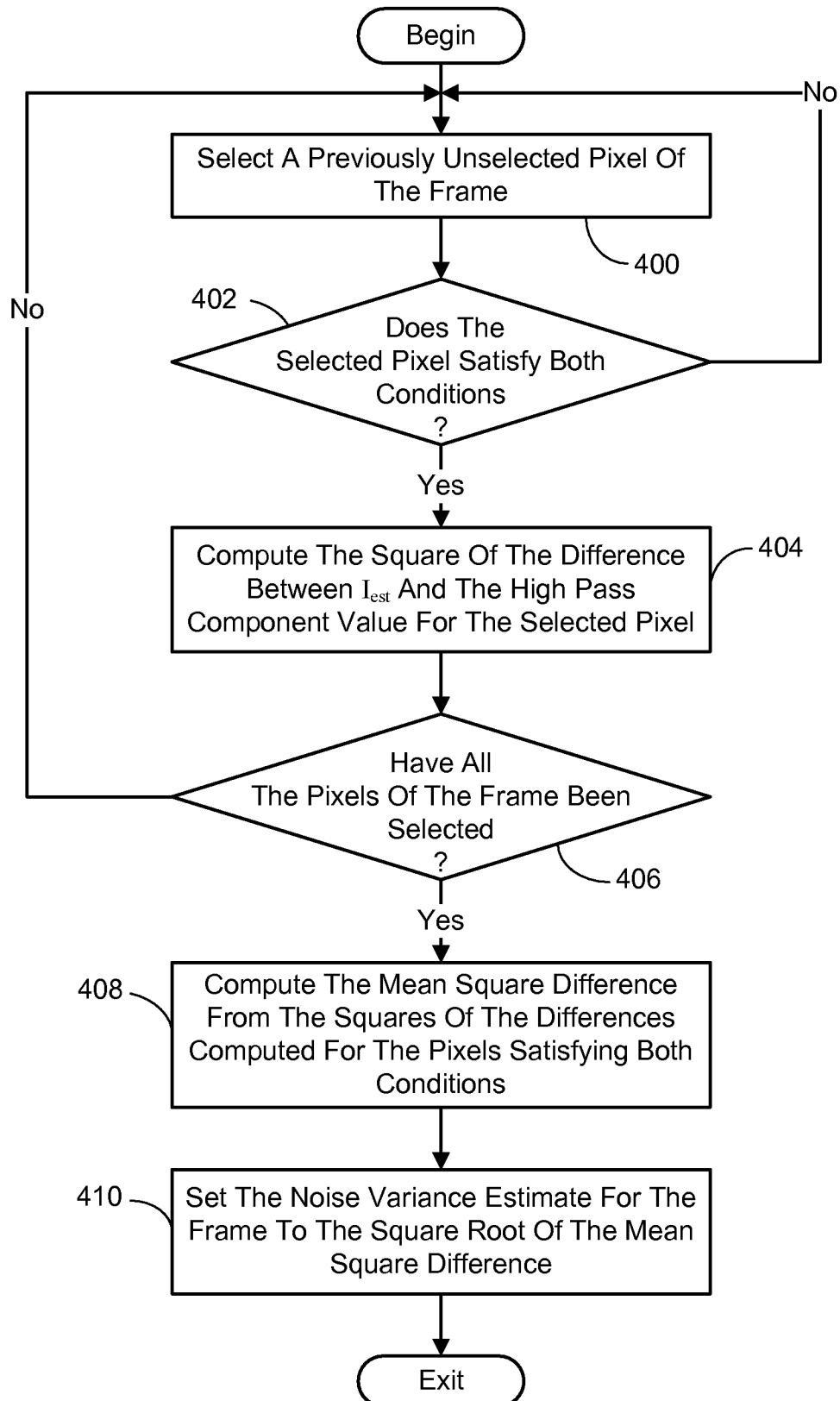


FIG. 4

4/9

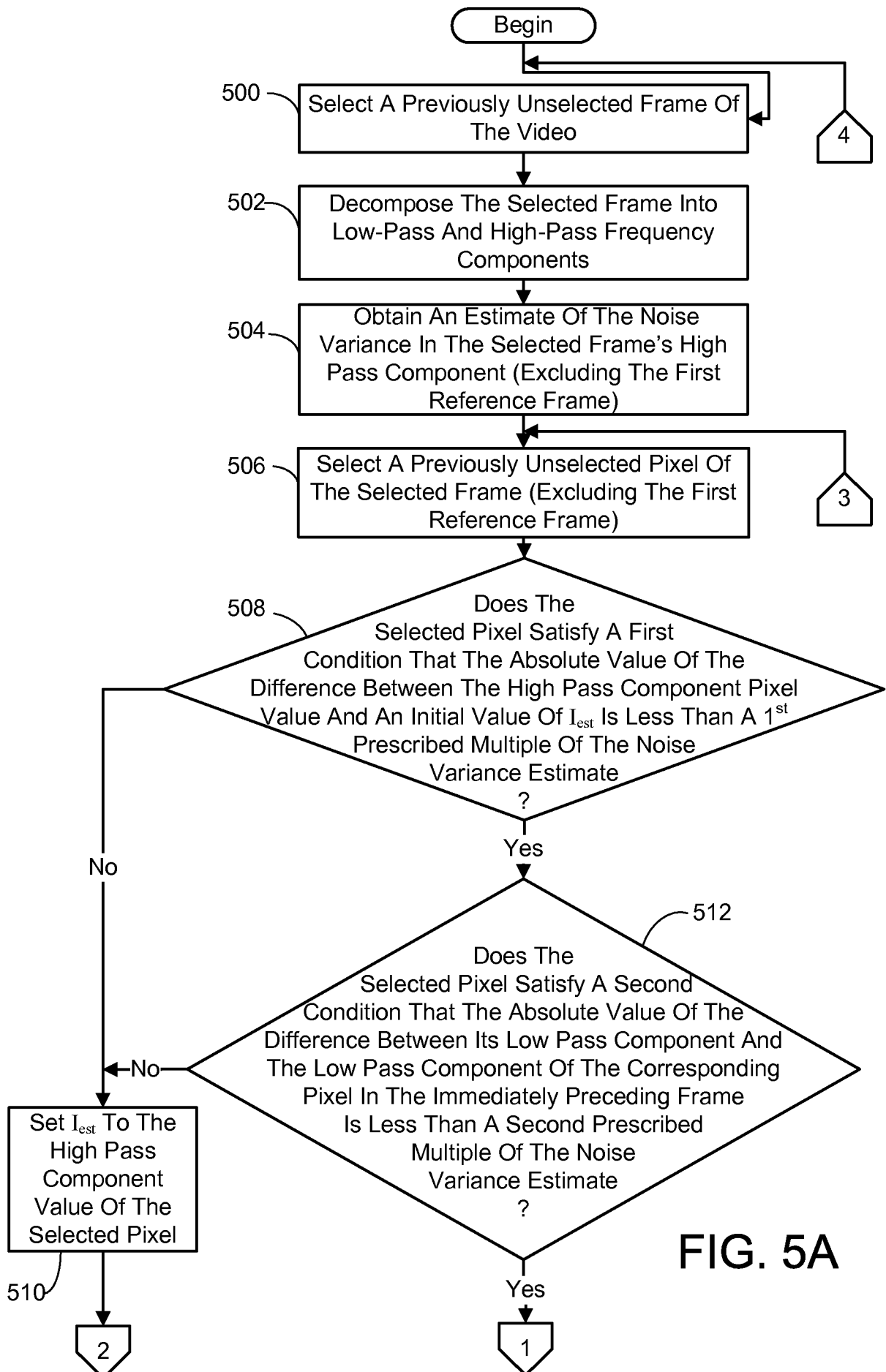


FIG. 5A

5/9

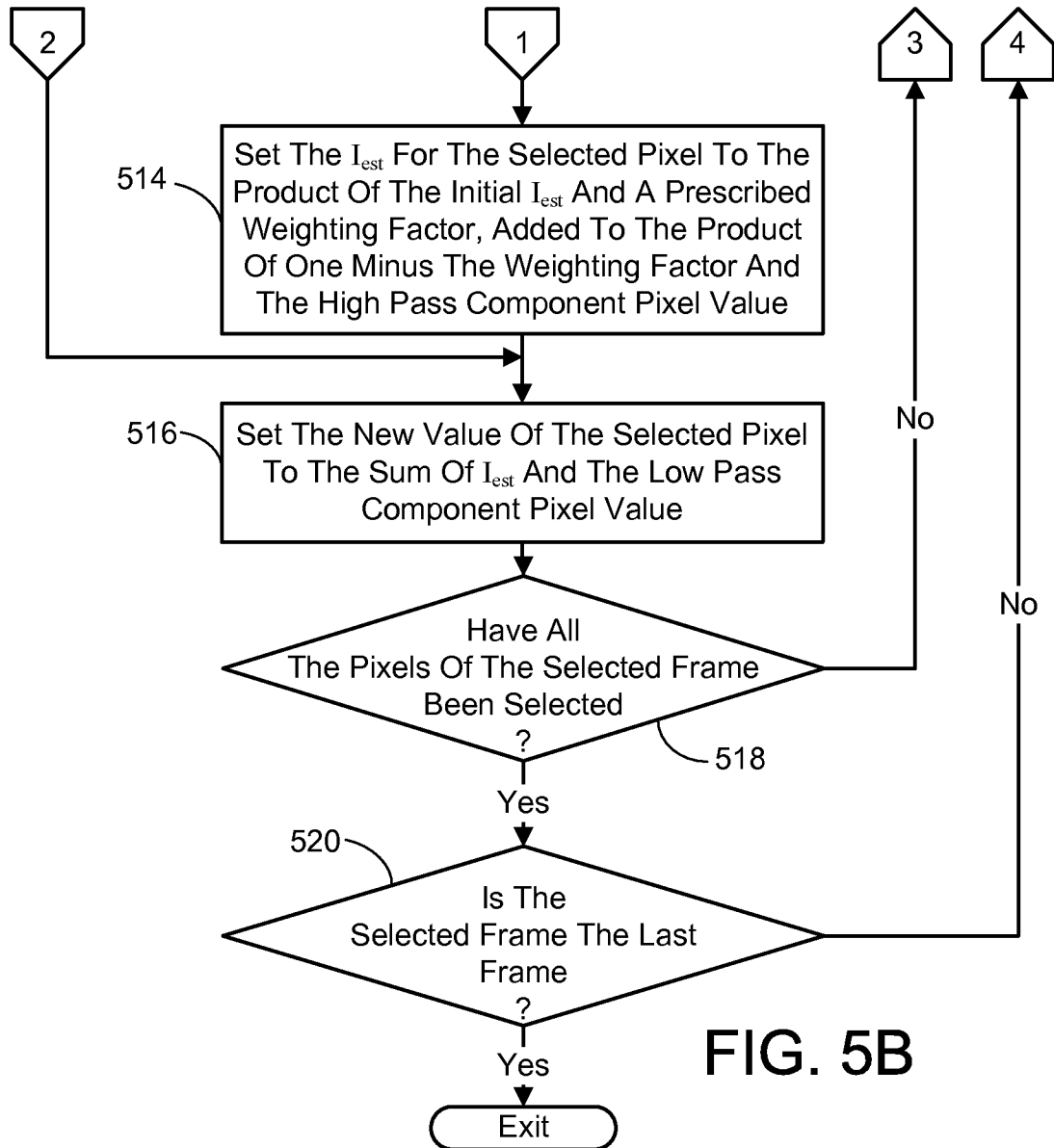


FIG. 5B

6/9

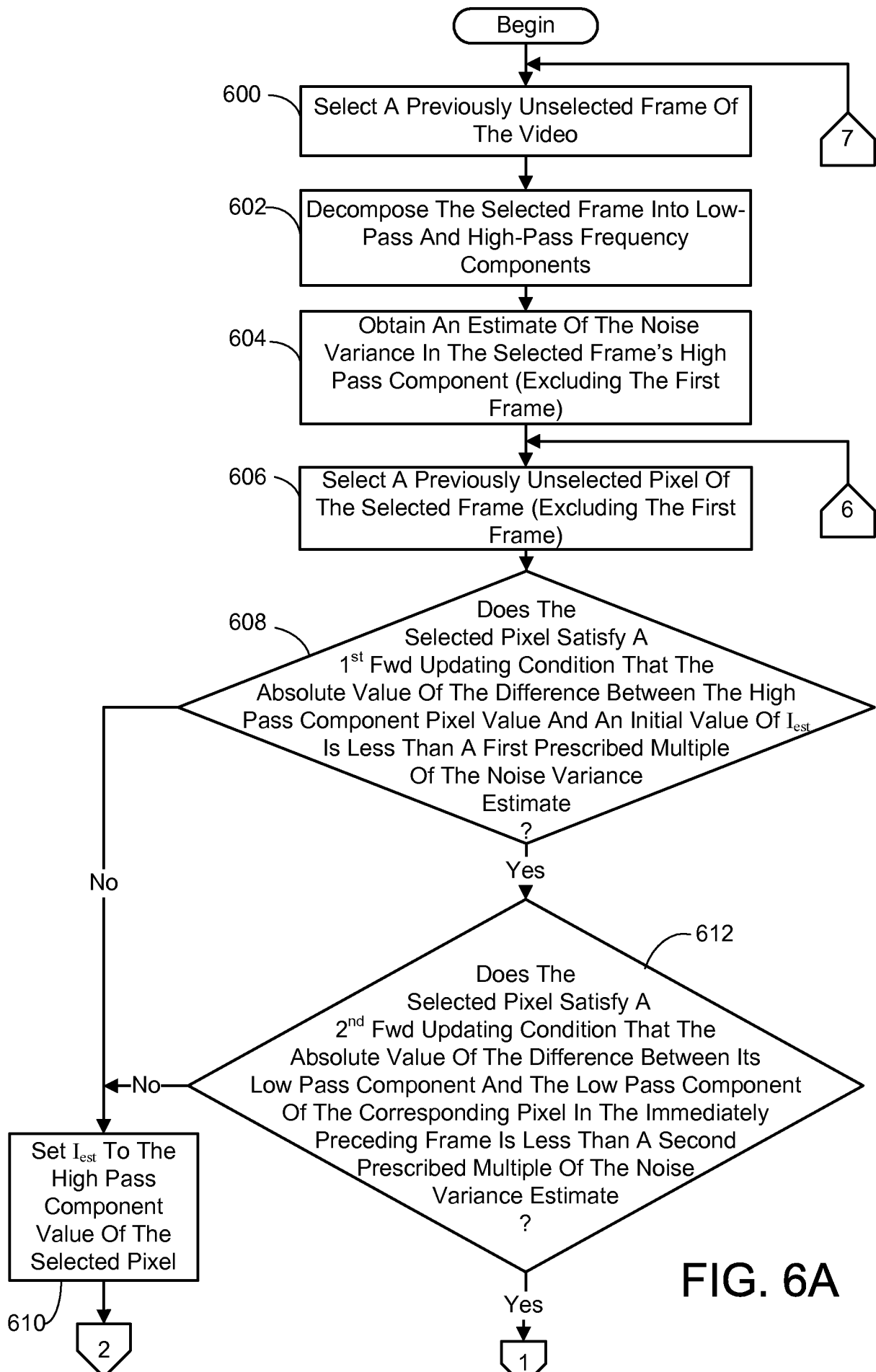
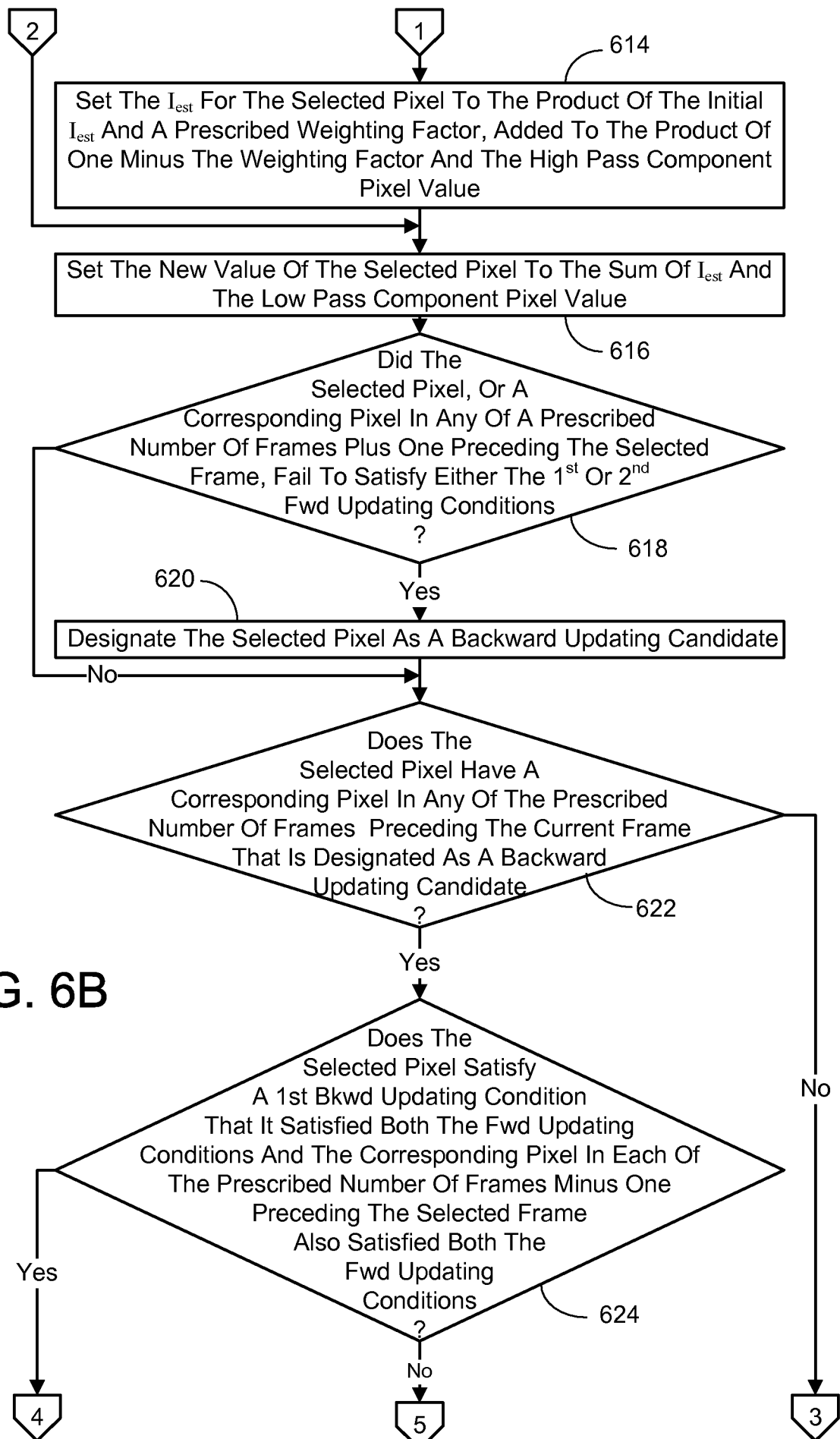


FIG. 6A

7/9



8/9

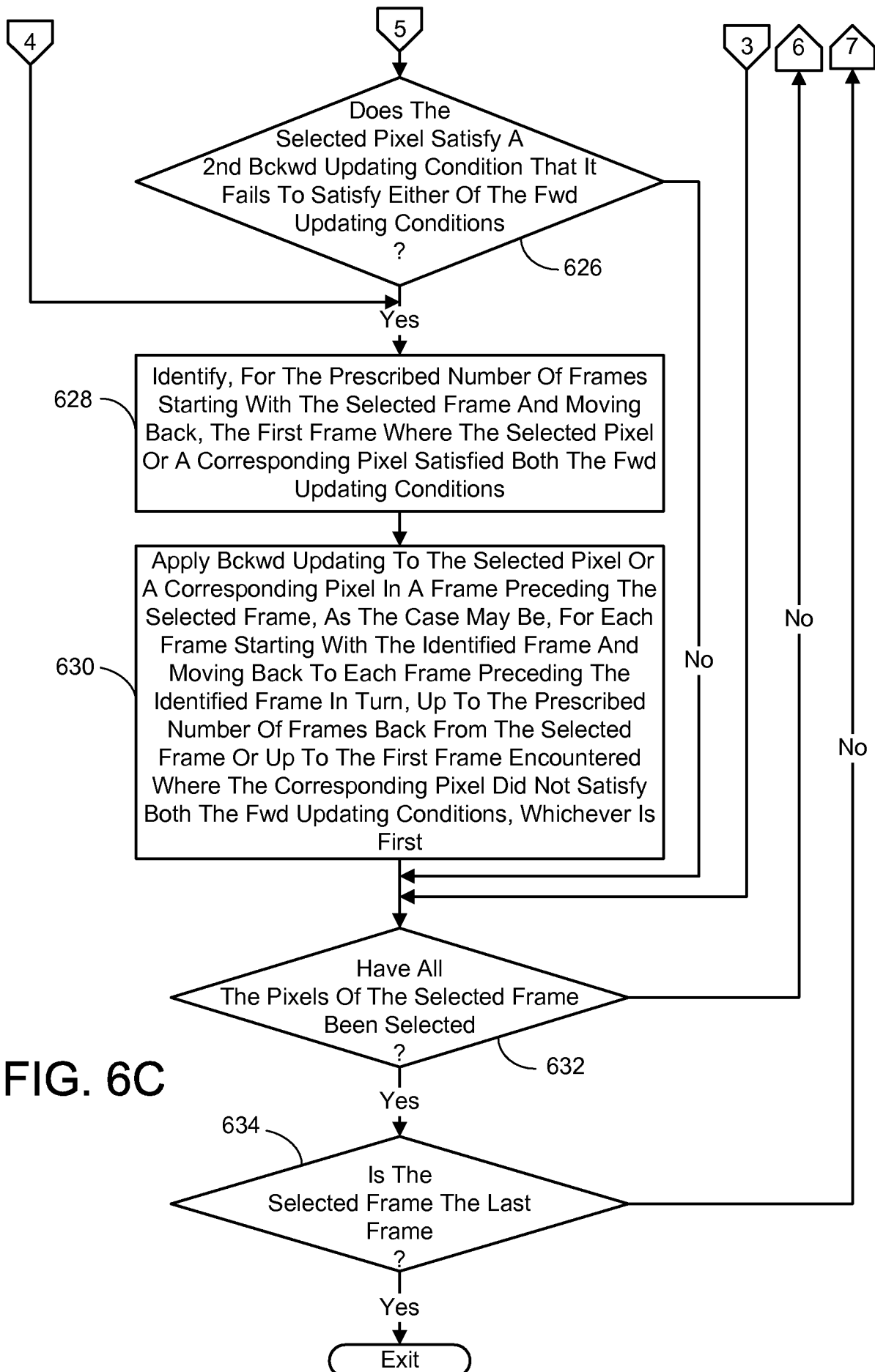


FIG. 6C

9/9

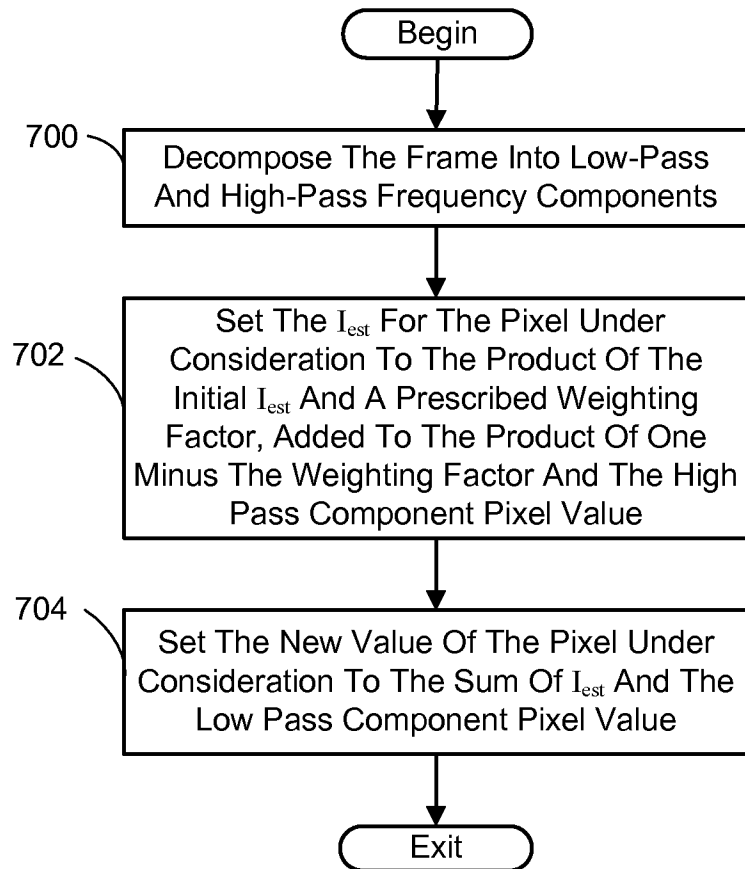


FIG. 7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2008/066005**A. CLASSIFICATION OF SUBJECT MATTER****H04N 5/21(2006.01)i**

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 8 : H04N, G06F

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

Korean Utility Models and applications for Utility Models since 1975

Japanese Utility Models and applications for Utility Models since 1975

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

eKIPASS(KIPO internal) "video conference, frame, noise, reduction, high pass, low pass"

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 7,016,540 B1 (NEC CORPORATION) 21 March 2006 See whole document.	1-20
A	US 7,007,098 B1 (NORTEL NETWORKS LIMITED) 28 February 2006 See abstract, figure 1 and claim 1.	1-20
A	US 2007/0083597 A1 (JOSEPH SALESKY et al.) 12 April 2007 See abstract, figure 2 and claims 1, 4.	1-20
A	US 2004/0015597 A1 (BARRY W. THORNTON) 22 January 2004 See abstract, claim 1 and figure 24.	1-20

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

15 OCTOBER 2008 (15.10.2008)

Date of mailing of the international search report

15 OCTOBER 2008 (15.10.2008)

Name and mailing address of the ISA/KR

Korean Intellectual Property Office
Government Complex-Daejeon, 139 Seonsa-ro, Seo-
gu, Daejeon 302-701, Republic of Korea

Facsimile No. 82-42-472-7140

Authorized officer

MOON, Young Jae

Telephone No. 82-42-481-8592



INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2008/066005

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 7,016,540 B1	21.03.2006	JP P2001-155169 A US 7,151,852 B2 US 7,016,540 B1	08.06.2001 19.12.2006 21.03.2006
US 7,007,098 B1	28.02.2006	NONE	
US 2007/0083597 A1	12.04.2007	US 07,310,675 B1 US 2004/0080504 A1 US 2005/0080850 A1 US 2005/0163062 A1 US 2005/0169197 A1 US 2007/0083597 A1	27.03.2007 29.04.2004 14.04.2005 28.07.2005 04.08.2005 12.04.2007
US 2004/0015597 A1	22.01.2004	AU 2003254018 A1 US 07,325,033 A1	09.02.2004 29.01.2008