

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
22 March 2007 (22.03.2007)

PCT

(10) International Publication Number
WO 2007/032986 A2

(51) International Patent Classification:
H04N 7/01 (2006.01)

(21) International Application Number:
PCT/US2006/034785

(22) International Filing Date:
8 September 2006 (08.09.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/715,711 8 September 2005 (08.09.2005) US
11/512,754 29 August 2006 (29.08.2006) US

(71) Applicant (*for all designated States except US*): **ANCHOR BAY TECHNOLOGIES, INC.** [US/US]; 300 Orchard City Drive, S131, Campbell, CA 95008 (US).

(72) Inventor; and

(75) Inventor/Applicant (*for US only*): **ADAMS, Dale** [US/US]; P.O. Box 448, Gualala, CA 95445 (US).

(74) Agent: **COLEMAN, Brian, R.**; Perkins Coie LLP, 101 Jefferson Drive, Menlo Park, CA 94025 (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, 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: SOURCE-ADAPTIVE VIDEO DEINTERLACER

(57) Abstract: A method is described for deinterlacing an interlaced video stream. The method includes detecting an occurrence of groups of adjacent fields that are derived from a common original image frame source, and merging field pairs of the interlaced video stream to create a non-interlaced video stream output.



WO 2007/032986 A2

SOURCE-ADAPTIVE VIDEO DEINTERLACER

FIELD

Embodiments of the present invention pertain to deinterlacing, and, more particularly, to source adaptive video deinterlacing.

BACKGROUND

Interlaced scanning and the progressive scanning are typical scanning methods employed in a video display device. Interlaced scanning has been employed in current National Television Systems Committee ("NTSC") television ("TV") systems. For the video display shown in Figures 1A-1C, even and odd fields would be interlaced sequentially and alternately on the screen when displaying a video image. For Figures 1A-1C, for interlaced scanning, the solid lines would represent lines being scanned at present, and the dotted lines would represent lines scanned for a preceding display.

An interlaced field of a video signal includes half of the number of total lines of an image. A sequence of interlaced fields arranged in display order alternate, field-by-field, between an even field and an odd field. Even fields are those fields that include only the even-numbered lines in the video image. Odd fields are those fields that include only the odd-numbered lines in the video image. For example, the interlaced fields 101 and 103 are even fields because they include only lines 0, 2, 4 and so forth. The interlaced field 102 is an odd field because it includes only lines 1, 3, 5 and so forth.

The progressive or non-interlaced scanning methods has been employed for computer monitors, digital TVs, Liquid Crystal Displays ("LCDs"), Digital Light Processing ("DLP") displays, plasma displays, etc. The non-interlaced or progressive scanning method is a method that sets one image frame as a frame unit and displays an entire frame at a time just like projecting films on the screen. Thus, if the video

display shown in Figures 1A-C instead employed progressive scanning, then the dotted lines would be sequentially scanned together with the solid lines.

A deinterlacing apparatus refers to an apparatus that converts a video signal employing interlaced scanning into a video signal employing progressive scanning. Accordingly, for a video display device that processes a progressively scanned video signal, in order for an interlaced video signal to be processed normally, a deinterlacing apparatus is required for the video display device to convert an interlaced video signal into a progressively scanned video signal.

Deinterlacing also provides other advantages, such as improving video quality. Specifically, deinterlacing can remove interlace motion artifacts, increase apparent vertical resolution, and reduce flicker. Furthermore, deinterlacing can also be used to restore source data that was originally progressive, such as data from film or computer graphics. Deinterlacing is also useful because progressive video compresses better than interlaced video. Additionally, interlaced videos can be scaled, but the quality can often be poor. Deinterlacing is also useful because progressive video scales better. Efficient and reliable deinterlacing techniques are further advantageous because video standards dictate that interlaced video will be around a long time. Indeed, deinterlacing is advantageous not only for so-called "legacy" standard definition video sources, such as NTSC, Phase Alternating Line ("PAL"), and Sequential Color a Memoire ("SECAM"), but also for High Definition TV ("HDTV"), given that the HDTV 1080i format employs interlacing.

A deinterlacer to convert a progressively scanned video signal into an interlaced video signal can be implemented in various ways. The deinterlacing techniques mentioned below are rudimentary techniques and may be used as basis for other deinterlacing techniques. A first deinterlacing technique involves combining two adjacent fields into one frame. Two fields (odd scan lines and even scan lines) are

combined together to create a deinterlaced³ or progressive video signal. This technique maintains vertical resolution, but has the problem of creating interlace motion artifacts, such as combing affects. In particular, combining two fields with moving objects results in comb-like edges.

A second deinterlacing technique involves using a single field to create a frame. A line replication method, in which line field information of the present field itself is used repeatedly, can be used. Alternatively, an intra-field interpolation method is employed in which a linear line interpolation of the present field itself can be used. Line replication typically results in a degradation of overall picture quality, however. The intra-field interpolation typically results in degradation of picture quality after interpolation for portions without motion.

Another deinterlacing technique is motion adaptive deinterlacing. For this technique, adjacent fields are merged for still areas of the picture and scan line interpolation is used for areas of movement. To accomplish this, motion, on a sample-by-sample basis, is detected over the entire picture in real time, requiring processing of several fields of a video signal.

Another deinterlacing technique involves detecting a repeating pattern, also referred to herein as a "cadence," in the video signal data. Repeating patterns are often formed when progressive video sources are transferred to interlaced video signals. The most common cadences are known as 3:2 pulldown and 2:2 pulldown. The former is used to transfer 24 frame per second film sources to 60 field per second interlaced video signals, while the latter is used to transfer either 30 frame per second sources (e.g., computer graphics) to 60 field per second interlaced video signals or to transfer 25 frame per second film (which is really 24 frame per second film sped up by 4%) to 50 field per second interlaced video (e.g., PAL). Once the cadence is detected, the original frame can be reconstructed.

Most cadence detection algorithms⁴ look for certain specific repeating patterns (such as those for 3:2 pulldown and 2:2 pulldown) to lock onto. This is disadvantageous in that many other cadences can exist. Examples include 2:3:3:2, 2:2:2:4, 3:2:3:2:2, 3:3, 5:5, 6:4 and 8:7. In general, when a cadence other than 3:2 and 2:2 is detected, the technique falls back on another deinterlacing technique, such as motion adaptive deinterlacing, instead of doing any further cadence detection.

Also, problem sources can make accurate detection of the cadence difficult. Such problem sources include out-of-sequence source type breaks (such as edits and splices), frame-rate converted material (such as when PAL video data is converted into NTSC data), multiple source types at once (such as fades, wipes, subtitles, special effects, stock tickers, multiple video windows, and the like), and noisy or otherwise corrupted sources.

SUMMARY

5

A method is described for source adaptive deinterlacing of an interlaced video stream. The method includes detecting the occurrence of groups of adjacent fields that are derived from a common frame in the original video signal and merging field pairs from the detected groups in the interlaced video stream to create a non-interlaced video stream output.

Other features and advantages of embodiments of the invention will be apparent from the accompanying figures and from the detailed description that follows below.

BRIEF DESCRIPTION OF THE DRAWINGS

One or more embodiments of the present invention are illustrated by way of example and not limitation in the figures of the accompanying drawings, in which like references indicate similar elements, and in which:

Figures 1A-1C show a video display with screen lines illustrating prior art screen scanning;

Figure 2 schematically illustrates the conversion of two consecutive interlaced fields into a single progressive frame using a deinterlacing module in accordance with an embodiment of the present invention;

Figure 3 illustrates a video system employing an embodiment of the invention;

Figure 4 is a flowchart of a method for converting a sequence of consecutive interlaced fields into a sequence of progressive frames in accordance with an embodiment of the present invention;

Figure 5 is a flowchart of a method for converting a sequence of consecutive interlaced fields into a sequence of progressive frames in accordance with an embodiment of the present invention;

Figure 6 illustrates the utility of using a 1-field difference for a 2:2 pulldown in accordance with an embodiment of the present invention;

Figure 7 illustrates the utility of using a 2-field difference for a 2:2 pulldown in accordance with an embodiment of the present invention;

Figure 8 illustrates the utility of using a 1-field difference for a 3:2 pulldown in accordance with an embodiment of the present invention; and

Figure 9 illustrates the utility of using a 2-field difference for a 3:2 pulldown in accordance with an embodiment of the present invention;

Figure 10 is an illustration of correlation values of temporally consecutive fields over time;

Figure 11 is an illustration of correlation values of temporally consecutive fields over time indicating an increasing maximum correlation value trend;

Figure 12 is an illustration of correlation values of temporally consecutive fields over time indicating a decreasing maximum correlation value trend;

Figure 13 is an illustration of correlation values of temporally consecutive fields over time indicating a change in correlation values around a scene transition;

A method is described for source adaptive deinterlacing of an interlaced video signal to create a progressive video signal. As described further below, the method includes determining if a sequence of two or more fields from the interlaced source is derived from the same original progressive source frame. One method of determining if a sequence of two fields have a common source is to determine how similar the two fields are to each other. The more similar the two fields are to each other, the more likely it is that they have been taken from the same original source frame. Once it has been determined that two fields come from the same original source frame, they can be merged to reconstruct the original source frame.

According to certain embodiments of the present invention, the method of deinterlacing by determining if two fields have a common source can recognize and adapt to any cadence or cadences in the video signal. Further, once it has been determined that two fields come from the same original source frame, they can be merged without need of any further processing, such as calculating pixels.

Figure 2 illustrates a deinterlacing module 203 that converts two temporal interlaced fields 201 and 202 into a progressive frame 204. The interlaced field 201 is an even field because it only includes lines 0, 2, 4, and 6. The interlaced field 202 is an odd field because it only includes lines 1, 3, 5, and 7.

For the example of Figure 2 and subsequent figures, interlaced fields are shown as being relatively small for clarity. For example, the interlaced fields are illustrated as being composed of 4 lines of 8 pixels, while the progressive frames are illustrated as being composed of 8 lines of 8 pixels. Nevertheless, typical interlaced fields and progressive frames will be many times larger than the example interlaced fields and progressive frames used in the figures and in this description. The smaller example fields and frames are used herein for clarity in describing the key principles of

embodiments of the present invention. Embodiments of the present invention are not limited to any particular size of interlaced field or progressive frame.

The deinterlacing module 203 of Figure 2 may be part of a video system, such as the video system 110 of Figure 3. Video system 110 includes a video processing system 60, a video display device 70 and a signal source 80. Video processing system 60 may include one or more programmable general-purpose or special-purpose microprocessors, digital signal processors ("DSPs"), programmable controllers, field programmable gate array ("FPGAs"), application-specific integrated circuits ("ASICs"), programmable logic device ("PLDs"), or the like, or a combination of such devices. The deinterlacing module 203 may be, for example, implemented in a PLD, a FPGA, an ASIC or other type of circuitry or device. Alternatively, or in addition, the deinterlacing module 203 may be one of the application programs and/or other program modules which may be loaded into system memory of processing system 60 in preparation for execution by processors of the processing system 60. For one embodiment, deinterlacing module 203 is part of video processor 60.

Processing system 60 includes a signal input that receives a video signal from a signal source 80. Signal source 80, may be either a single channel signal source or a multiple channel signal source. A single channel signal source provides programming from a recorded medium, such as a videocassette, compact disc, DVD, etc. Examples of a single channel signal source include a videocassette recorder, a CD player, and a DVD player. A multiple channel signal source includes any system or device that is capable of sending a signal that may be received by a satellite receiver, a cable or optic connection, a terrestrial antenna, or the like. Examples of a multiple channel signal source include Digital Satellite System ("DSS"), Digital Video Broadcasting ("DVB"), a cable box, locally broadcast programming (i.e. programming broadcast using Ultra High Frequency ("UHF") or Very High Frequency ("VHF"), and so forth.

The output of video processor 60 goes to a video display device 70, such as an HDTV, a standard definition TV, a computer monitor, etc. The display 70 can employ various display techniques, such a plasma display, a cathode ray tube ("CRT"), a LCD display, a DLP display, and a projector, for example.

A video processor 60 employing a deinterlacing module 203 receives an interlaced video signal from video signal source 80 and provides a deinterlaced signal to display device 70.

For alternative embodiments, the deinterlacing module 203 can be part of the video signal source apparatus 80 or the video display device 70. For alternative embodiments, the video processor 60 can be part of the video signal source 80 or the video display device 70.

Figure 4 is a flowchart of an overall method 300 of converting a sequence of temporally consecutive fields of interlaced video signal (e.g., fields 201 and 202 of Figure 2) into a sequence of progressive frames of video data. The method involves detecting in an interlaced video stream the occurrence of groups of adjacent fields that are derived from the same original video image frame so that the original video image frame may be reconstructed by merging the fields. The detection is done by comparing a pair of fields at block 301, determining a measure of similarity or difference between the fields at block 311, and determining if the fields are derived from a common source frame at block 321. If there is a high level of confidence that the fields are derived from a common source frame, then pairs of fields can be merged to reconstruct the original source frame at block 331. Otherwise, the two fields are deinterlaced using other techniques at block 341.

Method 300 utilizes the observation that the more similar two or more fields are to each other, the more likely they are to have been taken from the same original video image. The method 300 also utilizes the fact that conversion of a slower frame rate

progressive source to a faster frame rate ¹⁰interlaced format requires that multiple sequential fields be taken from the same original source frame. In other words, there will be a sequence of adjacent fields that are all taken from the same source frame.

Figure 5 is a flowchart of a method 400 of converting a sequence of temporally consecutive fields of an interlaced video signal (e.g., fields 201 and 202 of Figure 2) into a sequence of a progressive video signal. At block 401, in order to determine how similar two fields are to each other, a correlation operation is performed on a field pair. Accordingly, a correlation operation may be performed on two fields that are spaced one field apart. This operation is herein known as "1-field operation." Alternatively, a correlation operation may be performed on two fields that are spaced two fields apart. This operation is herein known as "2-field operation." According to certain embodiments of the invention, both the 1-field operation and the 2-field operation may be calculated.

According to embodiments of the present invention, a correlation operation may be performed on two fields that are spaced "x" field is apart, where $x > 0$. However, performing a correlation operation on fields that are in the vicinity of each other, for instance, are spaced 1 or 2 fields apart, is more useful and reliable, because it provides a more frequent indication of any changes in field data.

The 2-field operation compares fields that are spatially coincident. Thus, for an even-odd sequence of fields, the fields being compared are both composed of either even-numbered lines or odd-numbered lines. Accordingly, a 2-field operation can be performed by calculating a pixel-by-pixel difference to obtain a 2-field difference. The magnitude of the pixel difference values can be summed over the entire field, with the resultant sum indicating how different or similar the two fields are overall. A low resultant sum value indicates very similar fields while a high resultant sum value

indicates very different fields. The terms "low" and "high", as well as "similar" and "different" as used herein, are, of course, relative.

The 1-field operation involves comparing two fields that are not spatially coincident -- i.e., for an even-odd sequence of fields, one field is composed of even-numbered lines while the other is composed of odd-numbered lines. Thus, because the two fields are not spatially coincident, they cannot be directly compared. In addition, there may be aliasing present in a single field due to the fact that taking only every other line of a source video image may not generate a high enough vertical sampling rate to represent all the vertical high frequencies in the video signal. The aliases are different between the two fields being compared, causing yet another difference between them.

From each field a comparison field is created by phase-shifting one field up and the other down. For one embodiment, one field is phase-shifted 1/4 line up and a field to be compared 1/4 line down. The phase shift can be performed for instance, using a Finite Impulse Response ("FIR") filter approach, although other techniques such as simple linear interpolation can also be used. For this embodiment, a comparison field is created from each field by computing pixels from the pixels of each field and comparing the calculated pixels. Additionally, a comparison field can be created from one of the two fields by computing pixels from the pixels of that field and the resultant calculated pixels can be compared with the pixels of the other field. The resulting field can be compared with an original field in the same manner as used for the 2-field difference described above.

In general, however, the 1-field difference has a higher baseline or noise level than the 2-field difference due to the fact that the original two fields being compared are not spatially coincident. Therefore, one or more operations may also be performed on each of the two original fields before the phase shift to reduce noise and aliasing artifacts. One such operation is a vertical low-pass filtering operation.

According to certain embodiments ¹² of the invention, both the 1-field difference and the 2-field difference may be calculated. The 2-field difference provides useful measurements when used on certain cadences, like 3:2 pulldown, but is not very useful when used on a cadence that has only two fields taken from the same frame, such as 2:2 pulldown. Figure 6 illustrates a pattern 631 detected by using a 1-field difference for a 2:2 pulldown. Figure 7 illustrates that by using a 2-field difference for a 2:2 pulldown, a detected pattern 641 actually shows no repeating pattern. For both Figures 6 and 7, progressive scan source frames 601, 611, and 621 are converted respectively into interlaced video field pairs 603 and 605, 613 and 615, and 623 and 625. A 1-field difference of the fields displays a detection pattern of 0-1-0-1-0, where a "0" represents that the two fields being compared are similar to each other, and a "1" represents that the two fields being compared are different to each other. A 2-field difference of the fields displays a detection pattern of 1-1-1-1-1. Thus, a 2-field difference for a 2:2 pulldown always conveys that the fields are different to each other, and therefore, in contrast to the 1-field difference does not provide much valuable information about the fields or the cadence.

On the other hand, as illustrated in Figures 8 and 9, both 1-field difference and 2-field difference operations are suitable for use on a 3:2 pulldown video sequence. Figure 8 illustrates a pattern 651 detected by using a 1-field difference for a 3:2 pulldown. Figure 9 illustrates a pattern 661 detected by using a 2-field difference for a 3:2 pulldown. For both Figures 8 and 9, progressive scan source frames 601, 611, and 621 are converted respectively into interlaced video field groups 602, 604, and 606; 612 and 614; and 622, 624 and 626. A 1-field difference of the fields displays a repeating detection pattern of 0-0-1-0-1. A 2-field difference of the fields displays a repeating detection pattern of 0-1-1-1-1. Thus, both the 1-field and 2-field difference for a 3:2 pulldown provide useful information.

In the description of Figure 5, while both 1-field and 2-field correlations are calculated and used, it is possible to perform the source-type detection based on only one of these measurements. Using both measurements together may result in a much more reliable detection of similar fields.

At block 411 of Figure 5, a history of the correlation measurements or values is saved. For one embodiment, the history of the correlation measurements is saved in a history buffer. When a new pair of frame correlation values is available at the end of each field period, the oldest values in the history record are discarded and the new values are added. According to certain embodiments of the present invention, correlation history measurements are saved over time for "n" frame periods. Thus, for both the 1-field and 2-field difference frame sums, a history is kept of the most recent "n" frame periods. For one embodiment, the value of n is chosen based on the longest cadence period that is desirable to detect. Thus, for 3:2 pulldown, a value of n can be 5, which represents the maximum cadence period length.

Based on the saved history, at block 421 of Figure 5 the following calculations can be made: local correlation maximum and minimum values, the dynamic range of the correlation values, and whether scene transitions occur within the vicinity of the current field or in the history range. As illustrated in Figure 10 a local correlation maximum value 402 is the maximum value of the difference frame sums, while a local correlation minimum value 403 is the minimum value of the difference frame sums. The dynamic range 404 of the correlations values is a difference between the maximum and minimum correlation values in the history range.

For one embodiment, new correlation values are adapted before being saved in the history. For instance, if a new correlation value is much larger than the maximum value currently in the history buffer, then the new value will be reduced in amplitude to avoid abnormally skewing the history data. A correlation value can be much higher than

a maximum value due to noise, momentary¹⁴ changes in the image data, or a scene transition. If a long term increase in the correlation value is actually present, the history values will adjust over time to match the incoming values.

Once new correlation values have been added, the updated history record is examined to determine the minimum and maximum correlation values. The sequence of history values is also examined to determine if the higher values (which represent fields which are from different source frames) have a decreasing trend, or if the lower values (which represent fields which are from the same source frame) have an increasing trend. Higher maximums and lower minimums are tracked by their very nature, but a downward maximum or upward minimum trend must be consciously tracked. For instance, as shown in Figure 11, an upward maximum trend is tracked because a new local maximum 405 is greater than the previous local maximum 402. However, as shown in Figure 12, a downward maximum trend is not tracked by its very nature, because the local maximum 402 does not change. This tracking can be done for maximums by looking at the sequence of history values for which the threshold comparison (as described below) yield a "1" (or "different") value, and is done for minimums by looking at only the history values for which the threshold comparison yields a "0" (or "similar") value. When such a trend is identified, the minimum and/or maximum values provided as outputs for subsequent processing steps can be set to the most recent value in the trend as opposed to the absolute minimum and/or maximum in the history record.

As discussed above, the 1-field difference measurement has a higher baseline value than the 2-field difference due to the fact that a pair of fields will almost never perfectly correlate. According to certain embodiments of the present invention, a method to compensate for this discrepancy is provided. A multiple of the minimum value in the history record is subtracted from both the minimum and the maximum

values before they are provided as outputs.¹⁵ A larger multiple is used when the dynamic range value is high and a multiple of 1 is used when the dynamic range is low. One net result of this is that the output minimum is always zero.

In addition, scene transition occurrences in a video sequence can also be detected. A scene transition normally results in a momentary spike, as illustrated in Figure 13, in both the 1-field difference and the 2-field difference values. Detecting the scene transition is important so as to prevent the history values from being artificially skewed upwards, and also to identify that a change in correlation level may be occurring. A scene transition can result in significant changes in the overall video image content, and this can manifest as very different minimum, maximum, and dynamic range values. Furthermore, the scene transition detection information is passed on to subsequent processing stages so that any such changes can be taken into account.

At block 431 of figure 5, based on the correlation history measurements, a threshold value is calculated. If the correlation value is above the threshold, then the fields are different. If the correlation value is below the threshold the fields are considered to be similar. The threshold value calculation is based on a number of factors, such as the minimum and maximum correlation values from the correlation history, the presence or absence of a scene transition, and the presence of a downward trend in the history maximum values. The threshold value is a dynamic value that changes from field to field.

Threshold values for 1-field operations and 2-field operations are calculated slightly differently. For one embodiment, for both operations, the threshold value is a fraction of the dynamic range and is calculated by using the following equation:

$$[\text{Dynamic Range} / \text{Scaling Factor}] + \text{Minimum}.$$
 The scaling factor varies depending on various factors.

The scaling factor for calculating the¹⁶ threshold for the 1-field operation depends largely on the size of the maximum. Because the 1-field minimum is always zero, the maximum is effectively the same as the dynamic range. Larger maximum values cause a larger scaling factor to be used, with a nominal range of 6 to 64. The threshold is biased towards the minimum rather than the middle of the range for two primary reasons. First, the larger correlation values tend to vary more than the minimums, and secondly, because it is better to incorrectly decide that two fields are different than to incorrectly decide that they are the same. In the event of a scene transition, the threshold is decreased to prevent false detections when the transition is from a higher average motion level to a lower average motion level. In such a case, a false detection would indicate that fields are similar when they are really not, resulting in interlace motion artifacts being present in the deinterlaced video signal. If a downward maximum trend is present, then a transition has no effect on the threshold calculation because the threshold level will already have been depressed by the decreasing maximum values.

The 2-field threshold scaling factor is also calculated as a fraction of the 2-field difference dynamic range. The calculation rules are a bit different, however, as the minimum is often not zero and more truly represents the difference between the 2 fields. When the minimum is very small, then the dynamic range scaling factor is chosen based on the minimum amplitude. For larger minimums, the threshold scaling factor is a fixed value. Like the 1-field difference threshold, the 2-field threshold is depressed in the event of a scene transition.

At block 441, the 1-field difference and 2-field difference values are each compared to their respective thresholds. If the difference value is above the threshold, the result is a "1" (or "different") and if it is below the threshold, the result is a "0" (or "similar"). A history of the comparison results for both 1-field and 2-field differences is maintained for a set number "m" of previous field periods. For one embodiment, m is

equal to at least two times n , because to recognize a repeating pattern, at least two instances of the pattern need to be present. The comparison history may be saved in a history buffer.

At block 451 of Figure 5, the comparison history is examined to determine if repeating patterns exist. Many such patterns are indicative of a transfer from a progressive source format to an interlaced one. The most common are known as 3:2 pulldown and 2:2 pulldown. Many other cadences exist, although they are in less common use than the two already cited. Examples include 2:3:3:2, 2:2:2:4, 3:2:3:2:2, 3:3, 5:5, 6:4 and 8:7. The 3:2 pulldown comparison pattern repeats every five field periods for both the 1-field difference and the 2-field difference. The 2:2 pulldown comparison pattern repeats every two field periods, but only for the 1-field difference, because the 2-field difference fails in this case. Thus, for instance, a comparison pattern of 0-1-0-1-0-... using a 1-field difference indicates a 2:2 pulldown. According to certain embodiments of the present invention, all cadences are tested from repeating every two field periods to repeating every fifteen field periods. The detected cadence or cadences are presented as output for subsequent processing steps.

At block 461 of Figure 5, based on the history of the comparison values, as well as the dynamic range value of the correlation values, a state value is assigned to the current field. This is done for both the 1-field and 2-field correlation histories. For one embodiment, a state value includes the values of "Start", "Middle", "End" or "None." The "Start" state indicates that a repeated sequence of fields from the same frame is just beginning. The "End" state indicates that the repeated sequence is ending. The "Middle" state indicates that the current field is between the beginning and end of the repeated sequence. The "None" state indicates that no repeated sequence is deemed to be present. As shown in Figure 8, the repeating state assignments 652 for 1-field difference for a 3:2 pulldown is Start / Middle / End / Start / End.

The basis of the state value assignment is the observation that conversion of a slower frame rate progressive source to a faster frame rate interlaced format requires that multiple sequential fields be taken from the same original source frame. In other words, there will be a sequence of adjacent fields that are all taken from the same source frame. Such a sequence must have a first field and a last field, and may have one or more middle fields.

The possible state values are therefore "Start", "Middle", and "End" when a repeating sequence exists. When no repeating sequence exists a state value of "None" is used. Because the 3:2 pulldown technique is very common, and because a 2-field difference is a very reliable indicator of such a pattern, an additional state assignment can be made to cover the two field periods not handled by the other states for the 2-field difference pattern. This state assignment is termed "InBetween." Thus, as shown in Figure 9, the repeating state assignments 662 for a 2-field difference pattern with a 3:2 pulldown source would be Start / Middle / End / InBetween / InBetween.

The state assignments are based on the comparison history as well as the previous state values. The following is a non-exhaustive list of examples of state assignment:

Table 1

<i>Field Operation</i>	<i>Current Comparison</i>	<i>Previous Comparison</i>	<i>Previous State</i>	<i>State</i>
1-field	0	1	End or None	Start
1-field	1	0	Start or Middle	End
2-field	0	1	None, InBetween, or End	Start

The state assignments are used to determine which fields to use to create the progressive video sequence, as described below.

19
According to certain embodiments of the present invention, other factors may be also used in determining the presence or absence of a repeating field sequence. One such factor is a "pair quality metric." This metric can be assigned when two sequential 2-field difference comparisons are made of fields that come from different original source frames. In such a case, one 2-field difference value is a comparison between the even field of original source frame "N" and the even field of original source frame "N+1", while the other is a comparison of the odd fields of those two source frames. For instance, referring to Figure 9, the 2-field difference value 628 is a comparison between the odd field 604 of original source frame 601 and the odd field 612 of original source frame 611, while the 2-field difference value 630 is a comparison between the even field 606 of original source frame 601 and the even field 614 of original source frame 611. Because both 2-field differences 628 and 630 are comparing the same two original source frames (601 and 611), their values will be very similar. If they are not similar, then either the fields being compared are not from just two source frames, or there is a problem of some type with the video signal. Based on the similarity of the two 2-field difference values, a quality metric can be assigned to the measurement. Thus, the method 400 of Figure 5 determines, when performing two sequential 2-field correlations of fields which come from different original source frames, if the 2-field correlation values are consistent with even/even and odd/odd field comparisons from 2 original source frames.

The pair quality metric is only assigned when the 1-field and/or 2-field state values indicate that the current and previous 2-field difference values were each comparing at least one field from the same source frames. This second of the pair of 2-field difference values occurs when the state is determined to be "End", and the first of the pair occurs in the previous field period. The magnitude of the difference between the pair of 2-field difference values is compared to the larger of the two values. Based

on the comparison result, a quality value (e.g., very good, good, medium, bad, very bad)²⁰ is assigned to that pair.

At block 471, based on the correlation history, state, any detected cadence, dynamic range quality, pair quality and scene transition values, the method determines if a lock can be acquired on the signal. A lock is considered to have occurred on the signal when the indicators as a whole provide a high confidence that a field sequence has come from a progressive source by extracting multiple fields from each original source frame. In other words, the signal has been recognized as having a valid repeating-field pattern and field pairs can be combined to deinterlace the video signal. A lock can be acquired on a sequence as long as a repeating pattern is detected, regardless of what the pattern actually is.

In order for lock to occur, a repeating pattern must exist. Ideally, the 1-field and 2-field difference values both agree on the pattern type. In certain cases, only one of these is required for lock as long as other quality metrics are sufficiently high. Once lock has initially occurred, the system stays in lock until some other event causes lock to be cleared. In general, it is harder to acquire lock than to lose it. The basic idea is that there needs to be a high confidence in the signal before lock is set and field pairs are combined to form the output, but that once lock has occurred many types of variations in the sequence pattern can be tolerated. Once the system is in lock the cadence does not need to remain constant and could, for example, change back and forth between common patterns such as 3:2 or 2:2 pulldown without losing lock.

Some factors which could cause loss of lock include scene transitions which cause the 1-field state to become "None," or both state values being "None," or very poor quality metrics, or conflicting 1-field and 2-field state values. Recognized cadences that can cause lock to be acquired are 2:2, 3:2, 3:3, 4:4, 5:5, 4:2, 2:2:2:4, 2:3:3:2, 3:2:3:2:2, 6:4 and 8:7.

At block 481 of Figure 5, method 400²¹ determines how to construct the deinterlaced output frame from the sequence of fields based on the state value assigned to the field. Accordingly, the current and previous fields are combined if the current field is at the start or in the middle of a repeated sequence. The previous and second previous fields are combined if the current field is at the end of a repeated sequence. Furthermore, motion-adaptive deinterlacing procedure can be performed on the current and previous field pair if there appears to be no valid combination of fields to reconstruct an original source frame.

Once lock has been acquired, field pairs are combined to form a deinterlaced output stream. The lock state alone does not solely enable this combination however. Rather, multiple factors must be present for field pairs to be combined to form the output. These factors include lock, current and previous states, dynamic range and field pair quality, and cadence.

There are "strong" and "weak" detections for the various field combination possibilities. A strong detection generally occurs when both the 1-field and 2-field measurements are in agreement and the quality metrics are not too low. Weak detections occur when only one of the 1-field or 2-field measurements is valid and requires that the quality metrics are high. There can be a prioritized sequence of decisions regarding which fields to combine, with strong detections having precedence and the weak detections being valid only if the strong detections are not. This prevents a weak detection decision criteria from being used when a higher-confidence, strong detection is present.

As described above, when the current state is "Start" or "Middle," then the current and previous fields are combined. When the state is "End," the previous and second previous fields are combined. When the 2-field state is "InBetween" (and there is no conflicting 1-field state), the first of the two "InBetween" states is treated as a

"Start" and the second is treated as an "End," When no valid overall system state is determined to exist, then the lock signal is de-asserted, combination of field pairs stops, and the system drops back to motion-adaptive deinterlacing. At block 491, a deinterlaced frame is output.

All of the various measurements and quality metrics, particularly as described in reference to Figure 5, are not absolutely required. Indeed, a subset of factors and measurements may be used. The more factors that are considered, however, the more confidence that can be achieved in the decision.

While Figure 3 and the corresponding discussion above provide a general description of a suitable environment in which embodiments of the invention may be implemented, features of embodiments of the present invention disclosed herein may be practiced in association with a variety of different system configurations. Embodiments of the invention may be implemented in hardware, software, or any combination thereof, whether now existing or to be developed in the future. Where desired, the operations of the various embodiments of the present invention may be incorporated into computer-executable instructions that may be stored on a computer-readable medium.

Embodiments of the invention can take the form of instructions, such as program modules, being executed by processors. Generally, program modules include routines, programs, objects, components, data structures, etc., that perform particular tasks or implement particular abstract data types. The sequences of instructions implemented in a particular data structure or program module represent examples of corresponding acts for implementing the functions or steps described herein.

In the forgoing specification, the invention has been described with reference to specific exemplary embodiments. It will, however, be evident that various modifications and changes may be made thereto without departing from the broader spirit and scope

23
of the invention. The specification and drawings are accordingly to be regarded in an illustrative sense rather than a restrictive sense.

CLAIMS

24

What is claimed is:

1. A method of deinterlacing an interlaced video stream, comprising:
detecting an occurrence of groups of adjacent fields that are derived from a common original image frame source; and
merging field pairs of the interlaced video stream.
2. The method recited in claim 1, wherein detecting an occurrence of groups of adjacent fields derived from a same original image frame in the interlaced video stream comprises:
determining if field pairs of the interlaced video stream are similar.
3. The method recited in claim 2, further comprising merging field pairs of the interlaced video stream if the field pairs of the interlaced video stream are determined to be similar.
4. The method recited in claim 2, further comprising deinterlacing the interlaced video stream using motion adaptive deinterlacing if the field pairs of the interlaced video stream are determined to be different.
5. The method recited in claim 2, wherein determining if field pairs of the interlaced video stream are similar comprises:
performing a correlation operation between field pairs.
6. The method recited in claim 5, wherein performing a correlation operation between field pairs further comprises:
performing a difference operation between field pairs that are spaced one field apart.

7. The method recited in claim 5, wherein performing a correlation operation²⁵ between field pairs comprises:

performing a difference operation between field pairs that are spaced two fields apart.

8. The method recited in claim 2, wherein determining if field pairs of the interlaced video stream are similar further comprises:

determining whether a scene transition has occurred.

9. The method recited in claim 5, wherein determining if field pairs of the interlaced video stream are similar further comprises:

calculating a threshold value based on one or more factors, wherein the factors comprise a history of correlation operation values between field pairs, a minimum and maximum value in the history of correlation operation values, a minimum and maximum value for a range of correlation values, and detected scene transitions.

10. The method recited in claim 9, wherein determining if field pairs of the interlaced video stream are similar further comprises:

comparing a correlation value for a field pair with the threshold value, wherein a correlation value smaller than the threshold value indicates a field pair that is from a common source.

11. The method recited in claim 10, wherein determining if field pairs of the interlaced video stream are similar further comprises:

examining a history of comparison values obtained from comparing correlation values with the threshold value to determine if a repeating pattern exists in the interlaced video stream.

12. The method recited in claim 11, wherein determining if field pairs of the interlaced video stream are similar further comprises:

assigning a state value to a current field, wherein the state value is used to determine whether two fields can be properly merged.

13. The method recited in claim 11, wherein determining if field pairs of the interlaced video stream are similar further comprises:

acquiring a lock on the repeating pattern.

14. The method recited in claim 3, wherein merging field pairs of the interlaced video stream further comprises:

determining how to construct the deinterlaced output frame from the sequence of fields.

15. The method recited in claim 14, wherein motion adaptive deinterlacing is used if no repeating pattern exists.

16. The method recited in claim 6, further comprising:
phase shifting field pairs to result in two spatially coincident calculated fields.

17. The method recited in claim 6, further comprising:
performing a vertical low-pass filtering operation on each field of the field pairs to remove noise and aliasing artifacts.

18. The method recited in claim 11, wherein determining if a repeating pattern exists in the interlaced video stream comprises:

assigning a quality value to a field pair representing a level of similarity between correlation values of field pairs.

19. A method of deinterlacing an interlaced video stream, comprising:

detecting an occurrence of groups ²⁷ of consecutive temporal fields of interlaced video fields that are derived from a common original image frame source;
detecting and locking onto a repeating pattern in the groups; and
combining field pairs based on the repeating pattern.

20. A computer-readable medium having stored thereon computer-executable instructions to deinterlace image signals, the computer-executable instructions comprising:

detecting an occurrence of groups of adjacent fields that are derived from a common original image frame source; and

merging field pairs of the interlaced video stream.

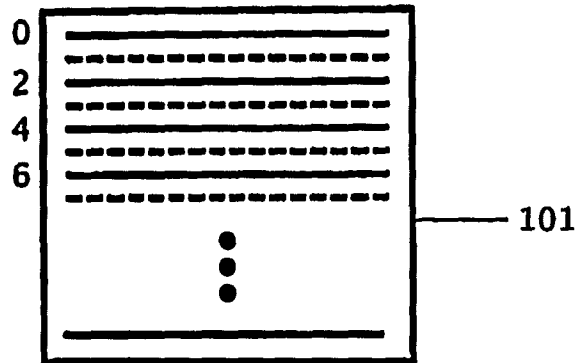


FIG. 1A
PRIOR ART

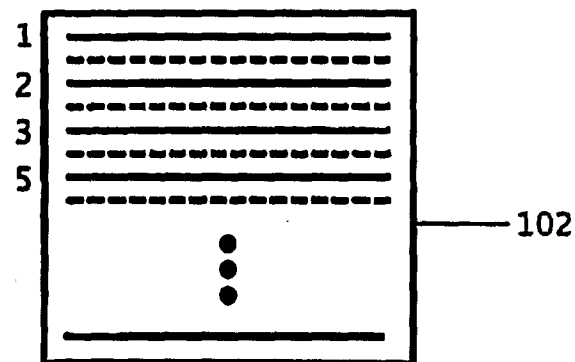


FIG. 1B
PRIOR ART

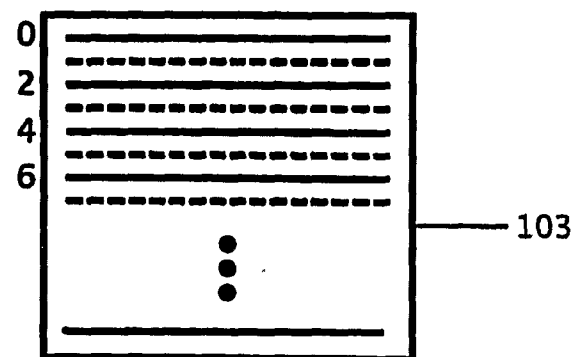


FIG. 1C
PRIOR ART

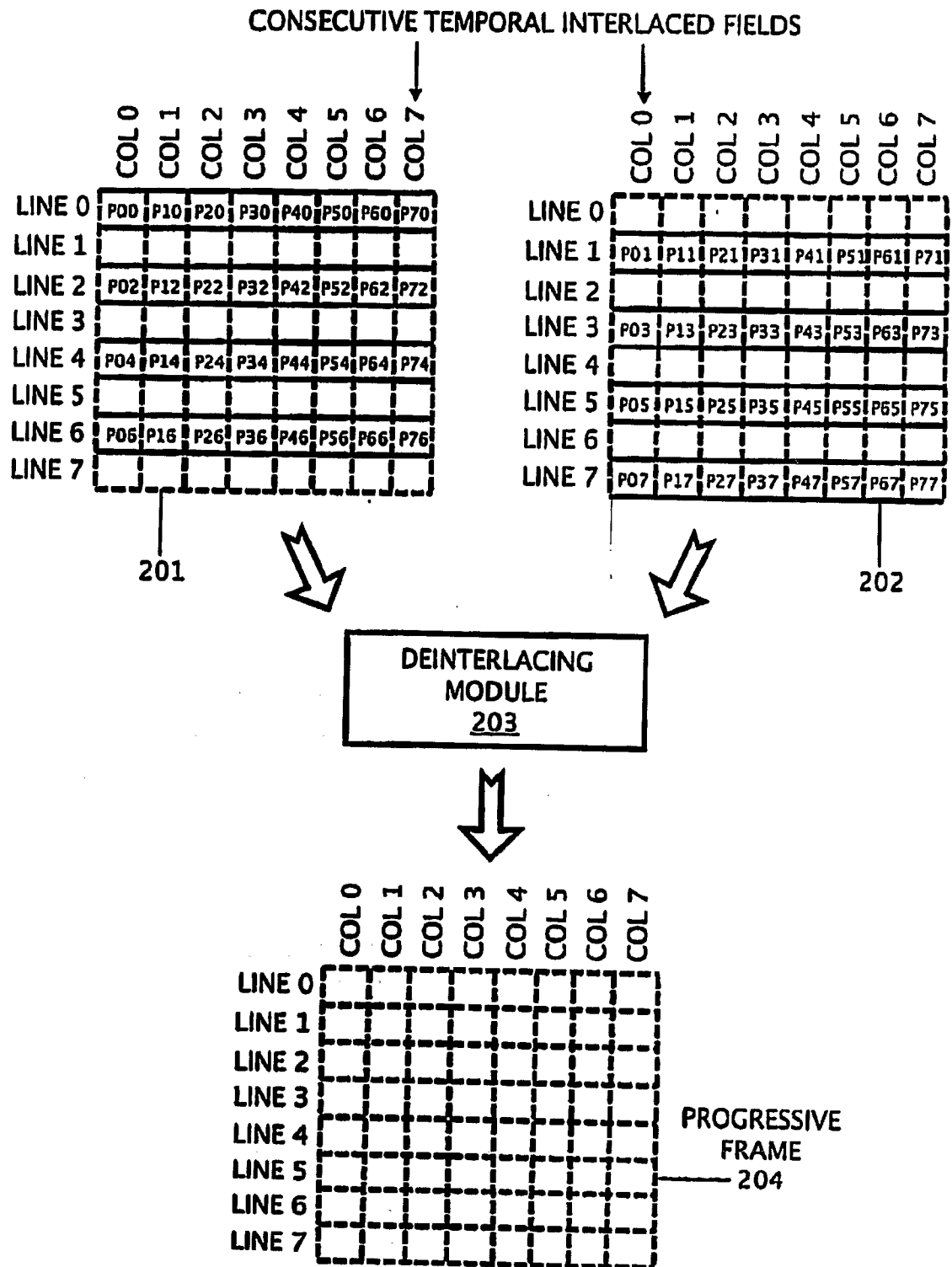


FIG. 2

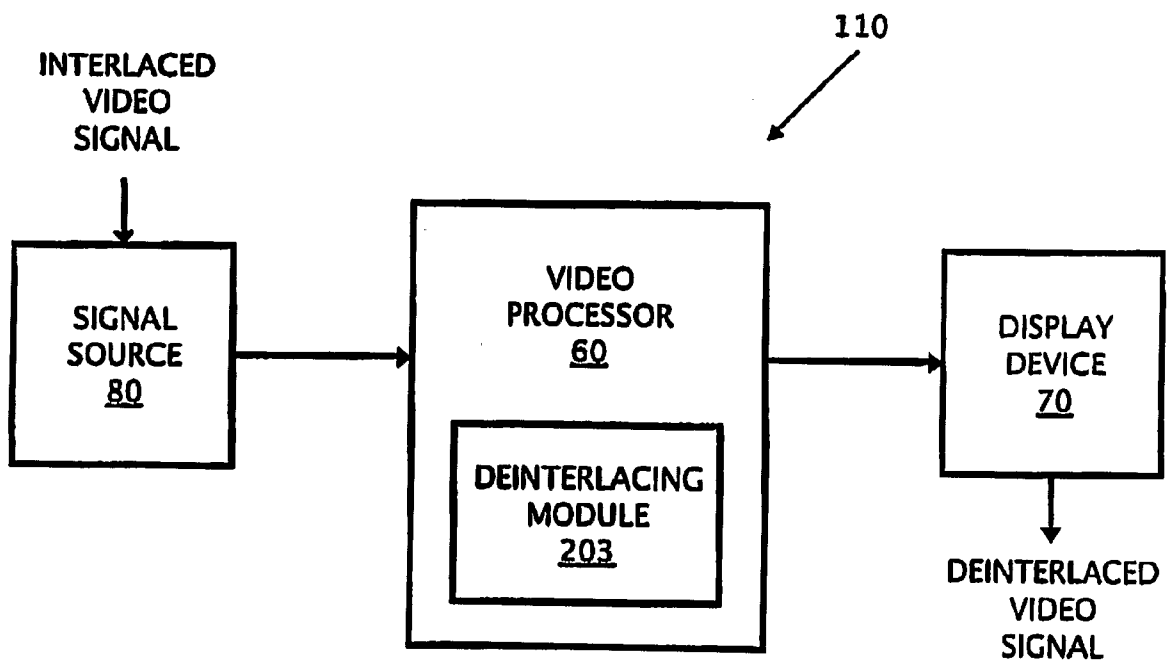


FIG. 3

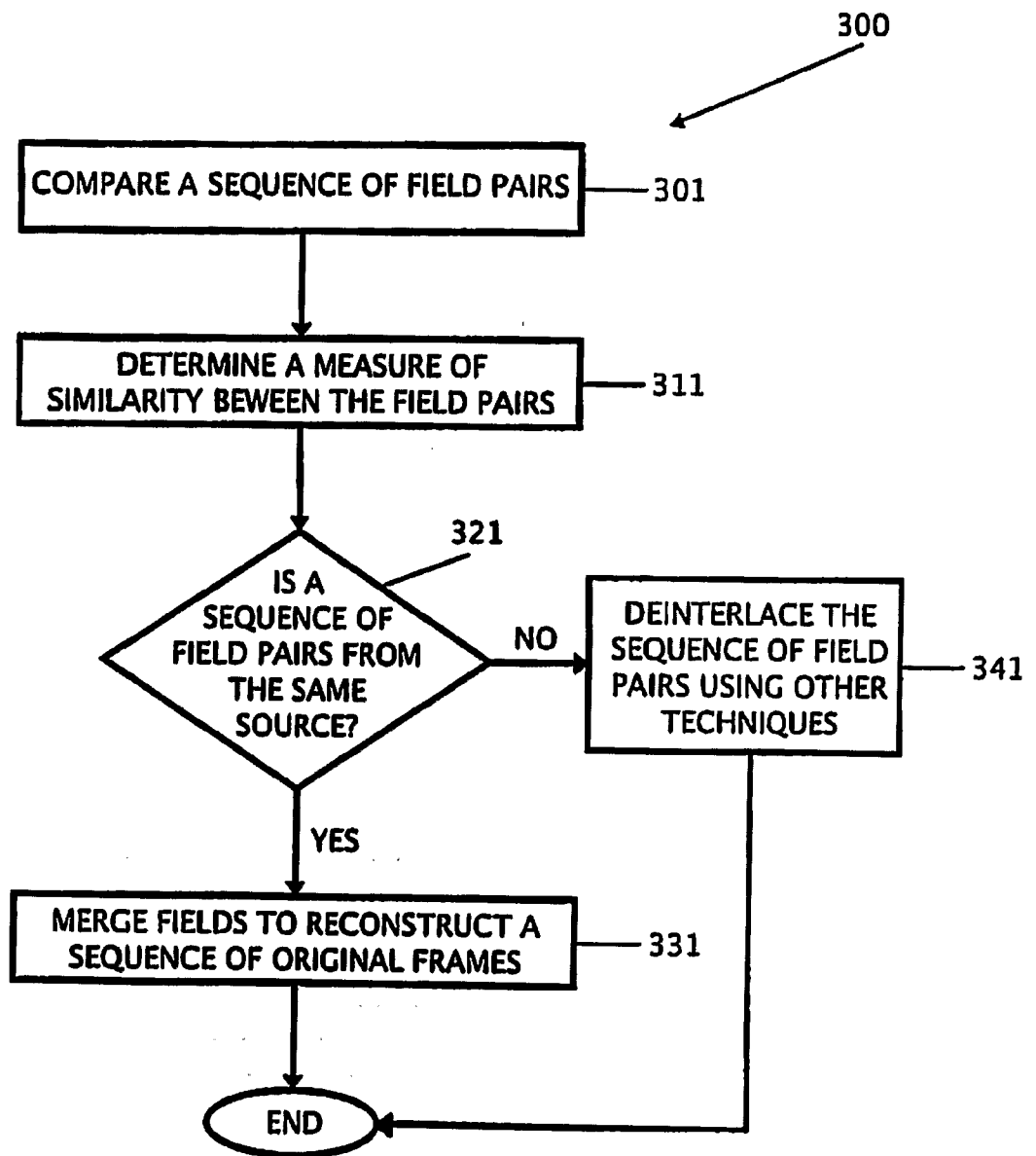


FIG. 4

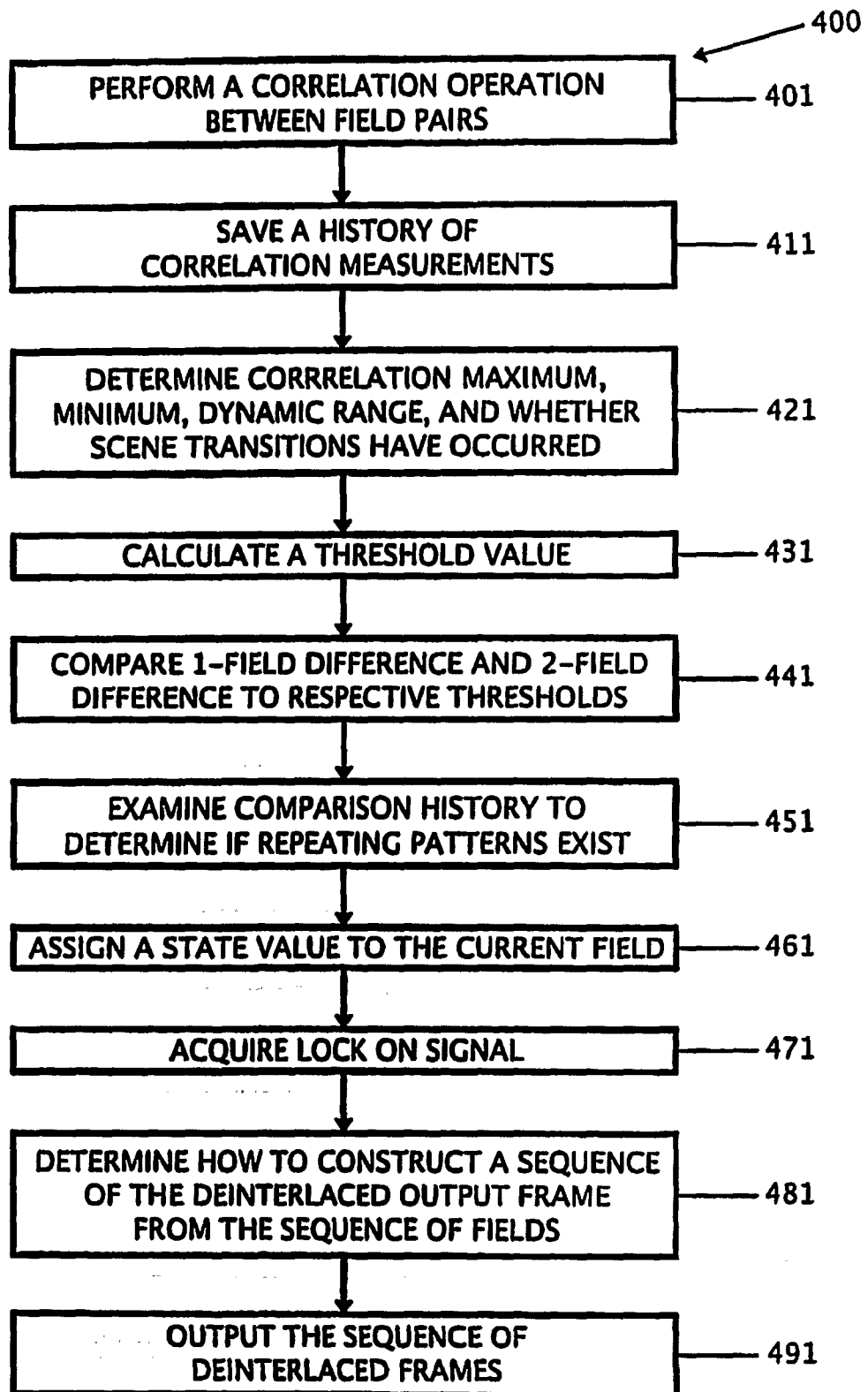


FIG. 5

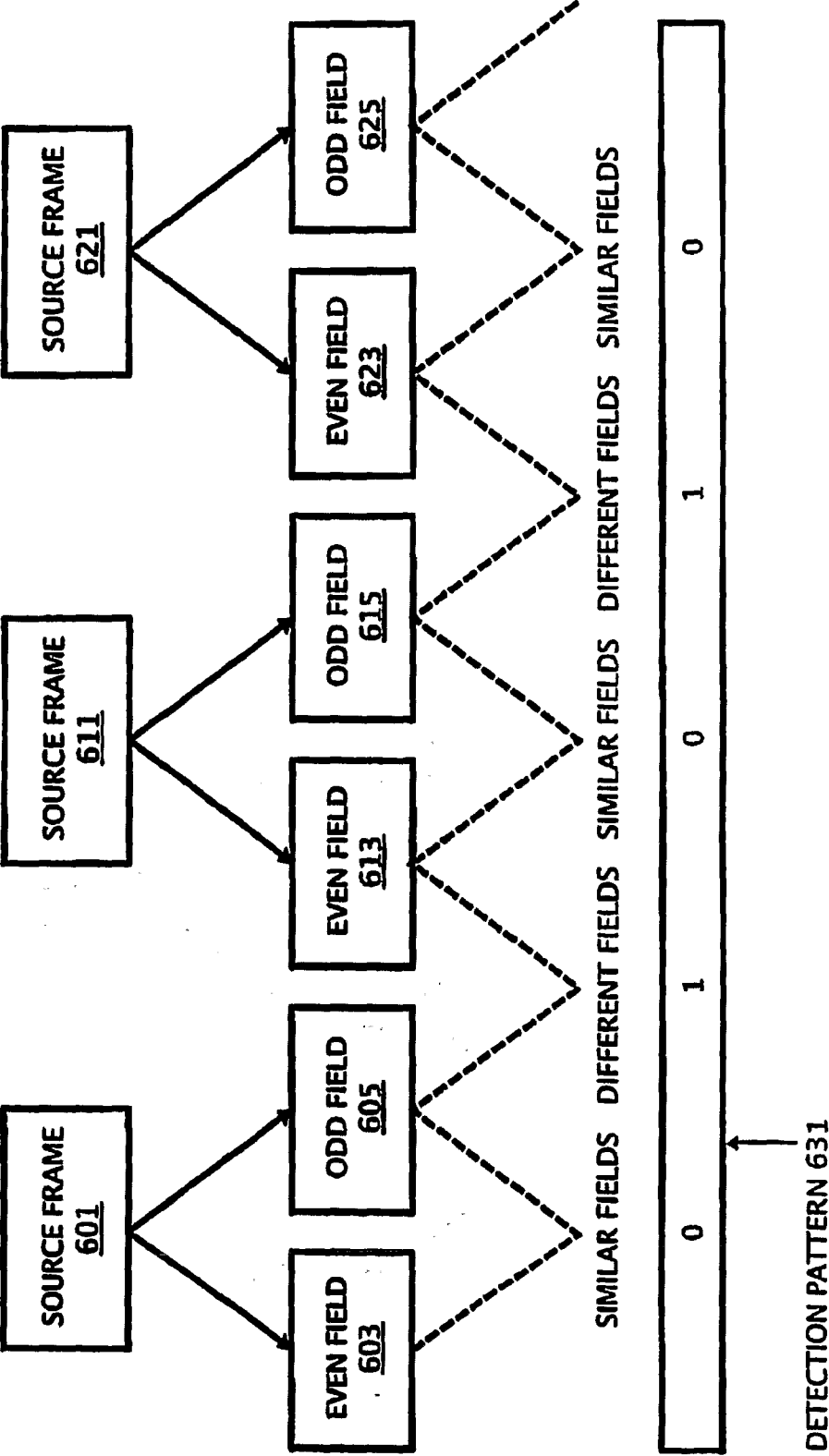


FIG. 6

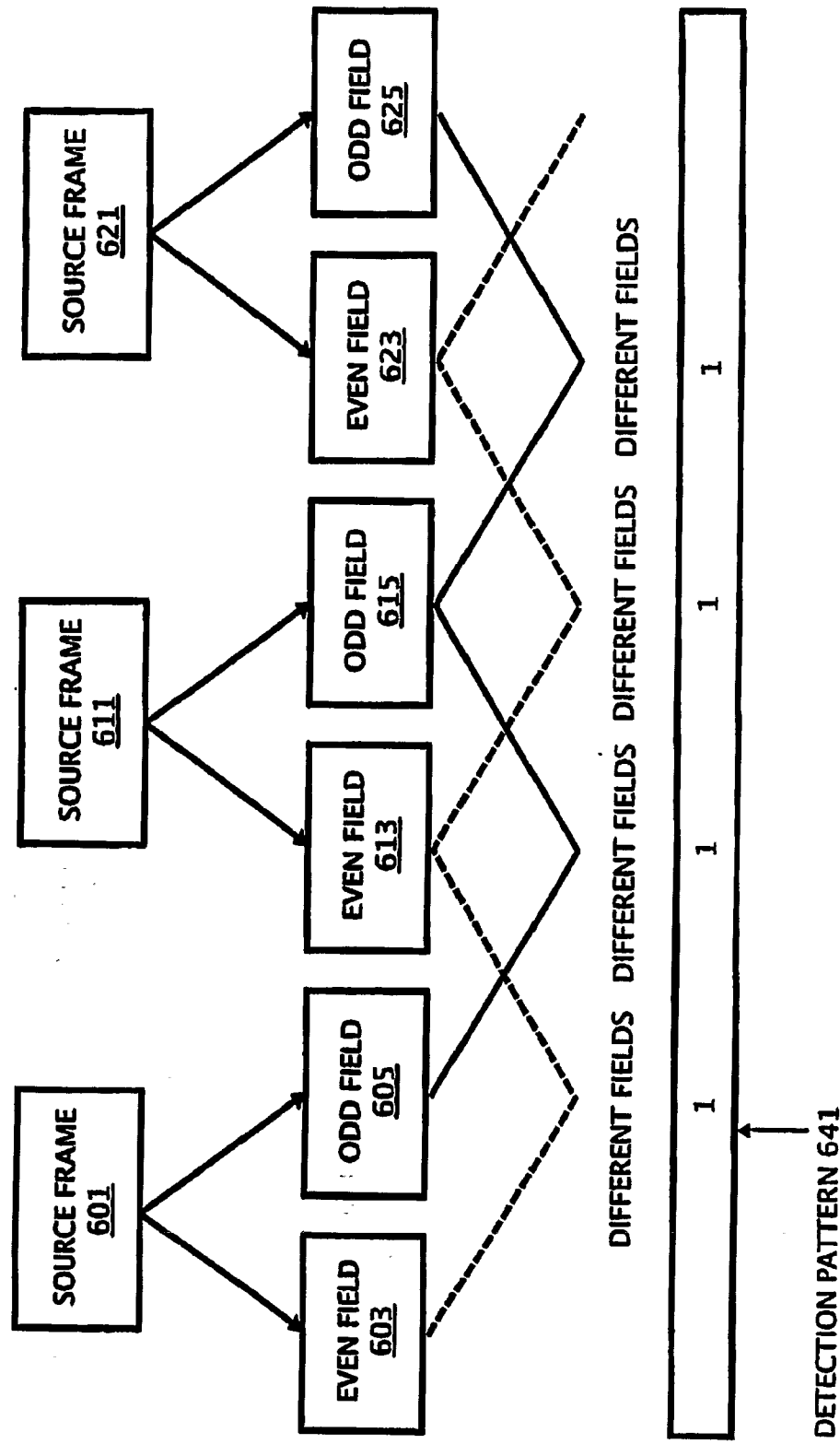


FIG. 7

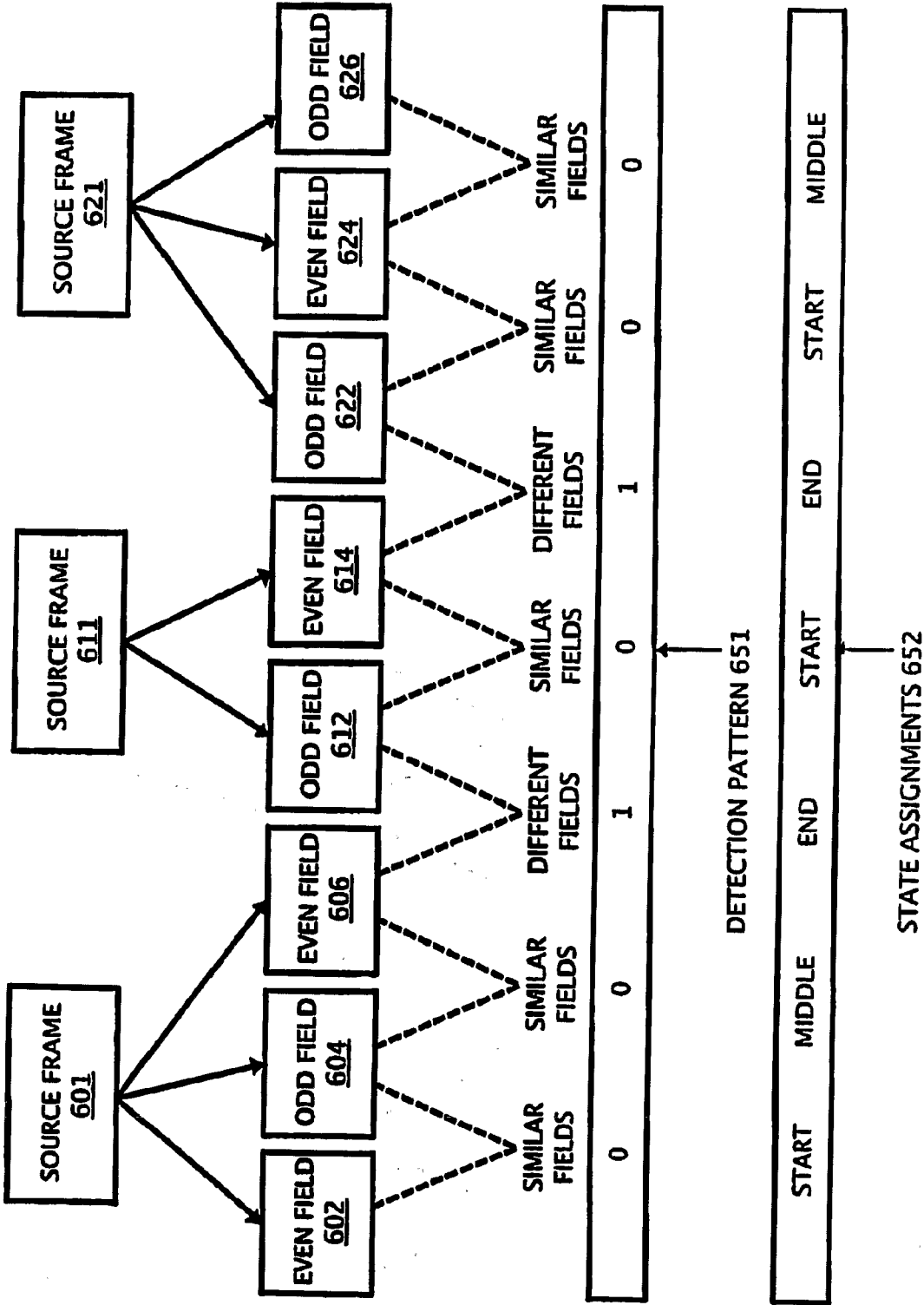
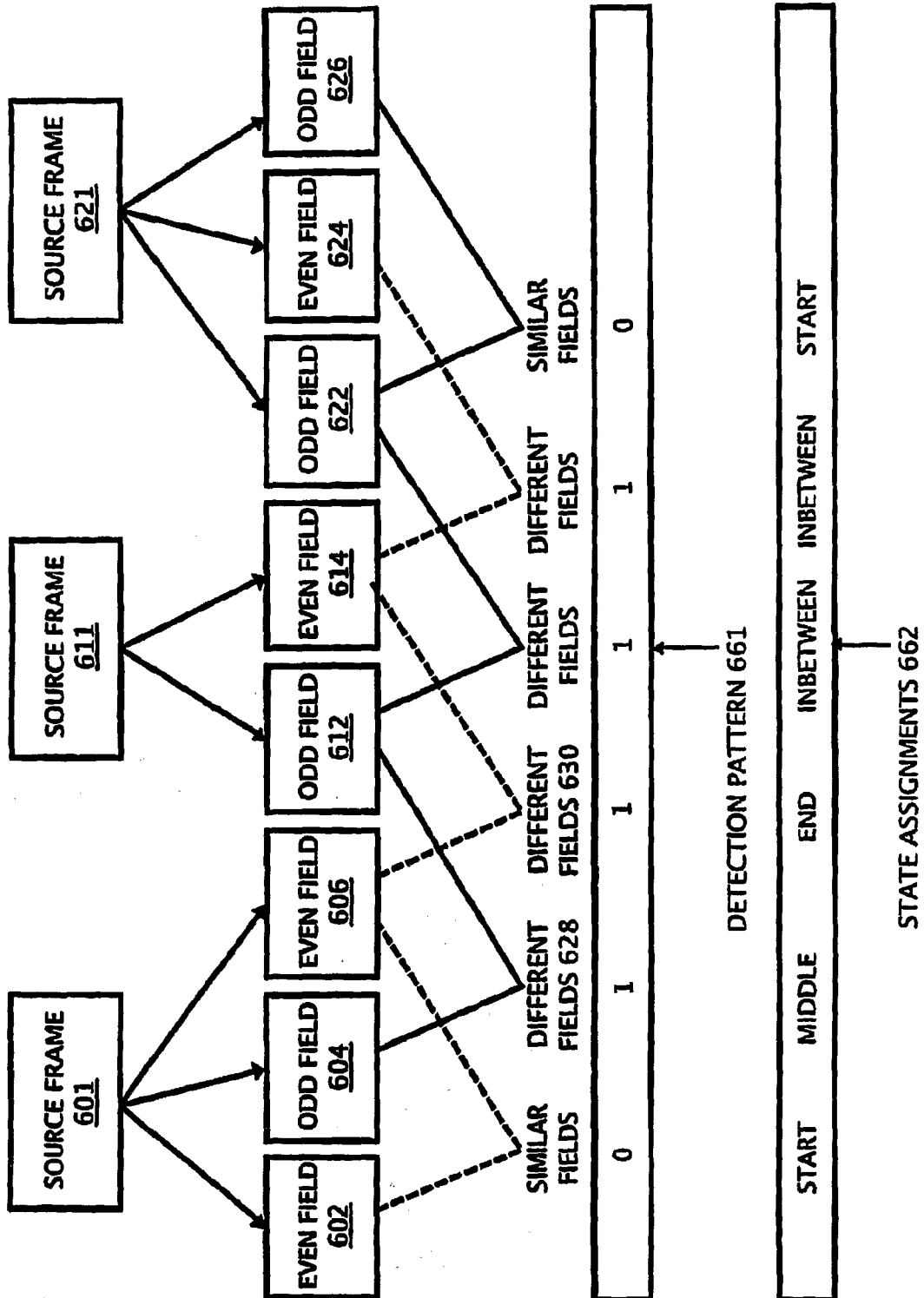


FIG. 8



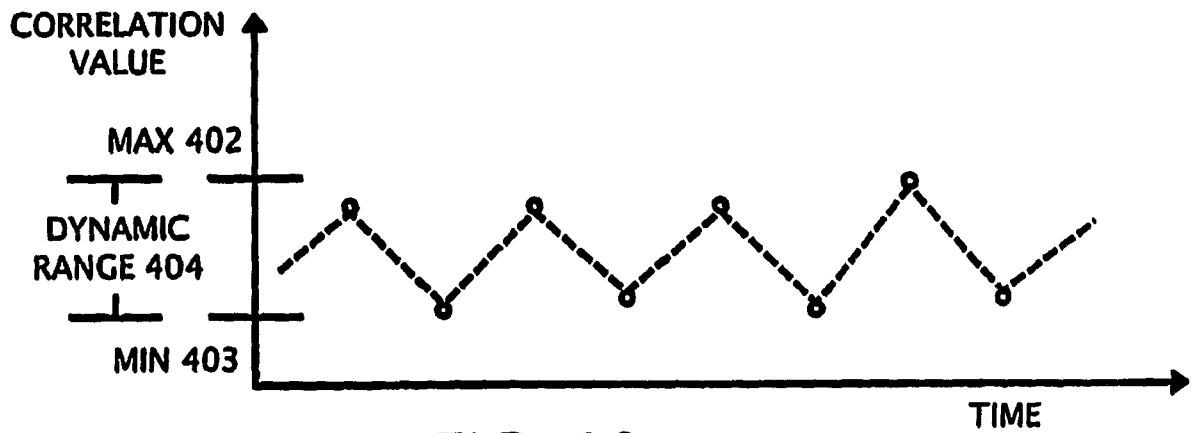


FIG. 10

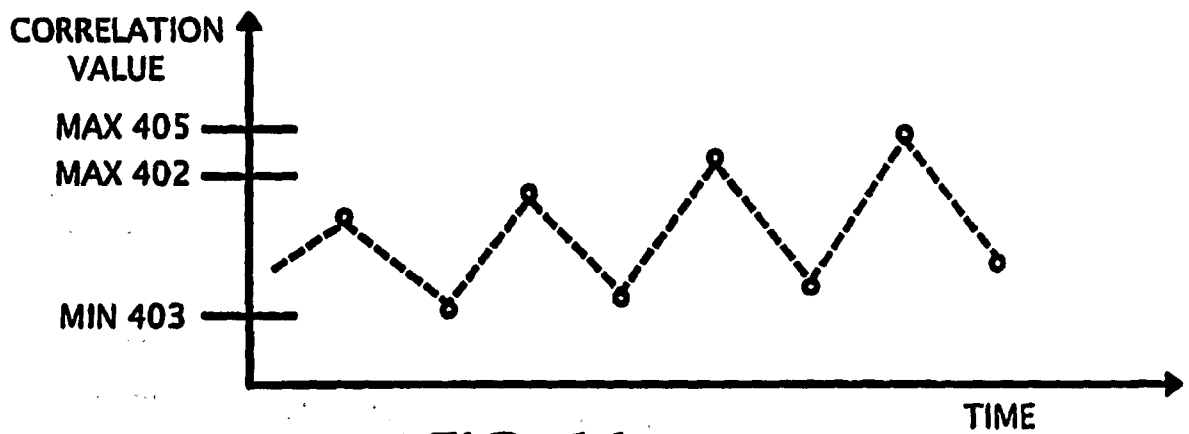


FIG. 11

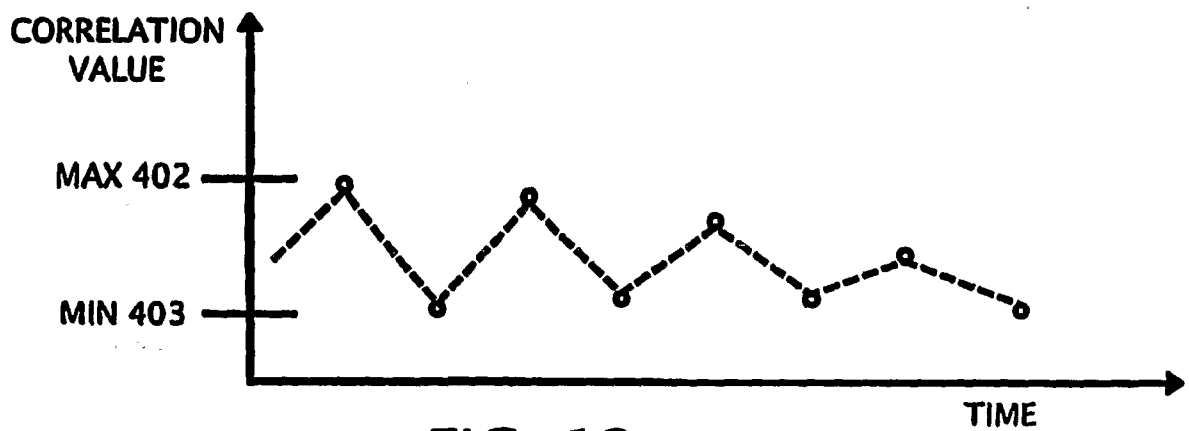


FIG. 12

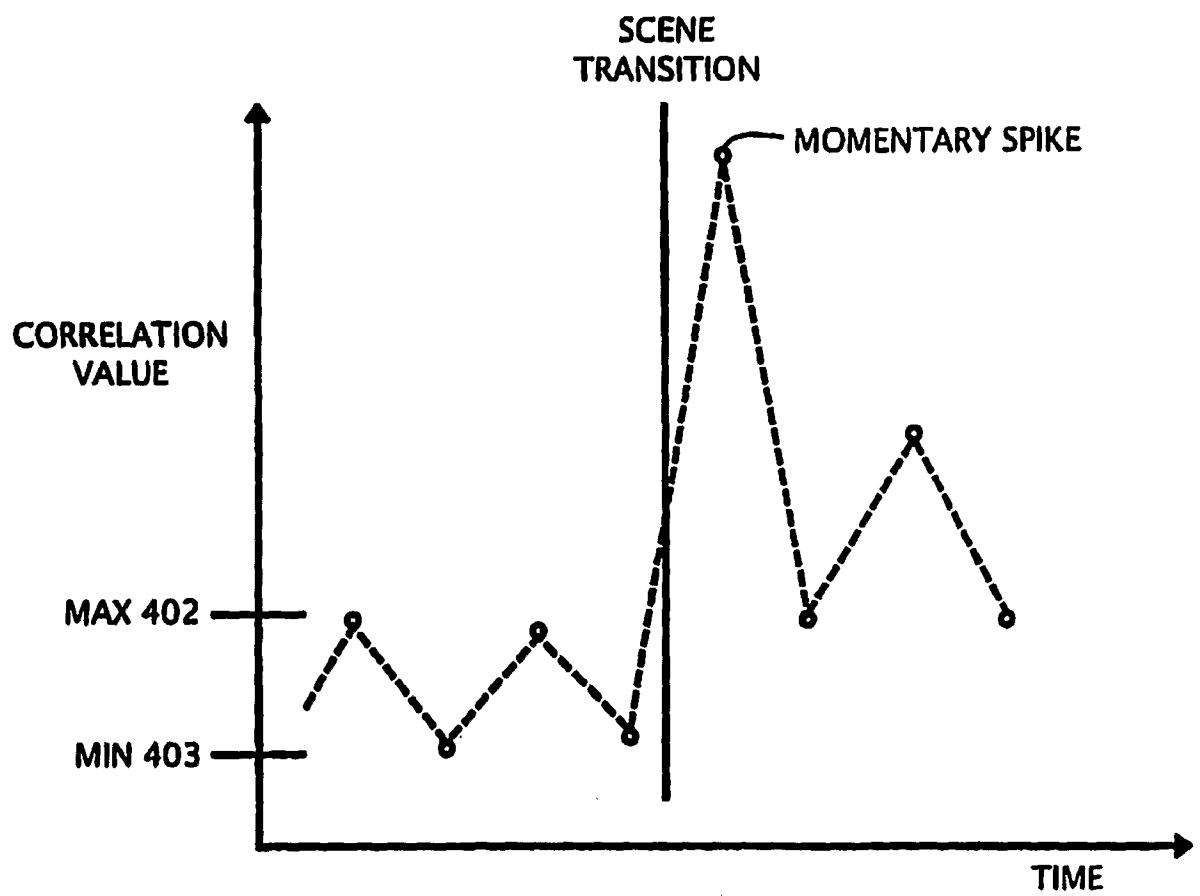


FIG. 13