

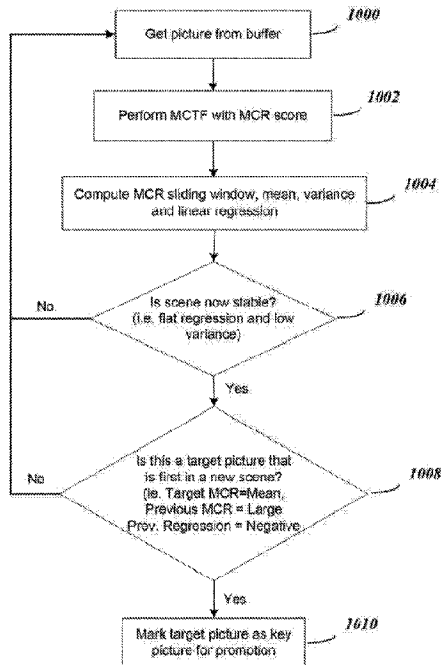


(86) Date de dépôt PCT/PCT Filing Date: 2015/12/09
 (87) Date publication PCT/PCT Publication Date: 2016/07/14
 (45) Date de délivrance/Issue Date: 2021/06/01
 (85) Entrée phase nationale/National Entry: 2017/07/04
 (86) N° demande PCT/PCT Application No.: US 2015/064714
 (87) N° publication PCT/PCT Publication No.: 2016/111800
 (30) Priorités/Priorities: 2015/01/05 (US62/099,674);
 2015/12/09 (US14/963,727)

(51) Cl.Int./Int.Cl. *H04N 19/142* (2014.01),
H04N 19/146 (2014.01), *H04N 19/194* (2014.01),
H04N 19/87 (2014.01)
 (72) Inventeurs/Inventors:
 MICHELSEN, WAYNE D., US;
 CHANDRA, SWAROOP, US;
 ZHOU, WEI, US;
 PERERA, JANE, US
 (73) Propriétaire/Owner:
 ARRIS ENTERPRISES LLC, US
 (74) Agent: GOWLING WLG (CANADA) LLP

(54) Titre : UTILISATION DE STATISTIQUES DE FILTRE TEMPOREL COMPENSE DE MOUVEMENT (MCTF) POUR LA DETECTION DE CHANGEMENT DE SCENE LORSQU'UN FONDU, FONDU ENCHAINE OU COUPURE SURVIENNENT

(54) Title: USING MOTION COMPENSATED TEMPORAL FILTER (MCTF) STATISTICS FOR SCENE CHANGE DETECTION WHEN A FADE, DISSOLVE OR CUT OCCURS



(57) **Abrégé/Abstract:**

A method is provided to better detect a scene change to provide a prediction to an encoder to enable more efficient encoding. The method uses a Motion Compensated Temporal Filter (MCTF) that provides motion estimation and is located prior to an encoder.

(57) **Abrégé(suite)/Abstract(continued):**

The MCTF provides a Motion Compensated Residual (MCR) used to detect the scene change transition. When a scene is relatively stable, the MCR score is also relatively stable. However, when a scene transition is in process, the MCR score behavior changes. Algorithmically, the MCR score is used by comparing the sliding mean of the MCR score to the sliding median. This comparison highlights the transition points. In the case of a scene cut, the MCR score exhibits a distinct spike. In the case of a fade or dissolve, the MCR score exhibits a transitional period of degradation followed by recovery. By implementing the above detection using the MCR, the location of the I-pictures in the downstream encoding process can be accurately determined for the encoder.

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
14 July 2016 (14.07.2016)



(10) International Publication Number
WO 2016/111800 A1

(51) International Patent Classification:

H04N 19/87 (2014.01) H04N 19/615 (2014.01)
H04N 19/172 (2014.01) H04N 19/146 (2014.01)
H04N 19/142 (2014.01) H04N 19/194 (2014.01)
H04N 19/107 (2014.01)

(21) International Application Number:

PCT/US2015/064714

(22) International Filing Date:

9 December 2015 (09.12.2015)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/099,674 5 January 2015 (05.01.2015) US
14/963,727 9 December 2015 (09.12.2015) US

(71) Applicant: ARRIS ENTERPRISES, INC. [US/US];
3871 Lakefield Drive, Suwanee, Georgia 30024 (US).

(72) Inventors: MICHELSEN, Wayne D.; 2450 Walsh Avenue,
Santa Clara, California 95051 (US). CHANDRA, Swaroop;
4461 Lafayette Street, Santa Clara, California

95054 (US). ZHOU, Wei; 684 Costigan Circle, Milpitas,
California 95035 (US). PERERA, Jane; 20218 Pierce
Road, Saratoga, California 95070 (US).

(74) Agents: WARD, Thomas A. et al.; 3871 Lakefield Drive,
Suwanee, Georgia 30024 (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, BN, BR, BW, BY,
BZ, CA, CH, CL, 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, IR, IS, JP, KE, KG, KN, KP, KR,
KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG,
MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM,
PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC,
SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, 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, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ,
TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU,
TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE,

[Continued on next page]

(54) Title: USING MOTION COMPENSATED TEMPORAL FILTER (MCTF) STATISTICS FOR SCENE CHANGE DETECTION WHEN A FADE, DISSOLVE OR CUT OCCURS

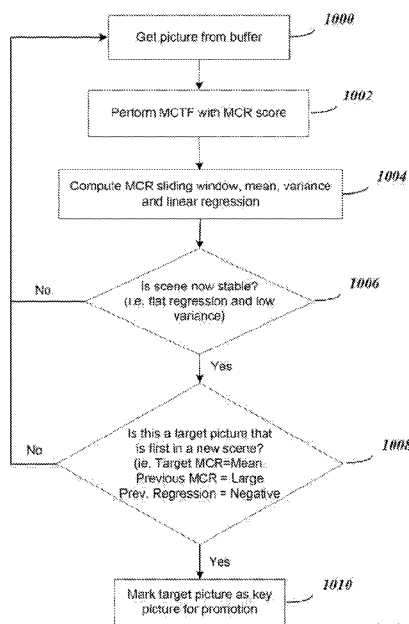


FIG. 10

(57) Abstract: A method is provided to better detect a scene change to provide a prediction to an encoder to enable more efficient encoding. The method uses a Motion Compensated Temporal Filter (MCTF) that provides motion estimation and is located prior to an encoder. The MCTF provides a Motion Compensated Residual (MCR) used to detect the scene change transition. When a scene is relatively stable, the MCR score is also relatively stable. However, when a scene transition is in process, the MCR score behavior changes. Algorithmically, the MCR score is used by comparing the sliding mean of the MCR score to the sliding median. This comparison highlights the transition points. In the case of a scene cut, the MCR score exhibits a distinct spike. In the case of a fade or dissolve, the MCR score exhibits a transitional period of degradation followed by recovery. By implementing the above detection using the MCR, the location of the I-pictures in the downstream encoding process can be accurately determined for the encoder.

WO 2016/111800 A1

WO 2016/111800 A1 

DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, **Published:**
LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, — *with international search report (Art. 21(3))*
SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA,
GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

USING MOTION COMPENSATED TEMPORAL FILTER (MCTF)
STATISTICS FOR SCENE CHANGE DETECTION
WHEN A FADE, DISSOLVE OR CUT OCCURS

[0001]

BACKGROUND

TECHNICAL FIELD

[0002] The present invention relates to detection of a scene change such as when a fade, dissolve or scene cut occurs. More particularly, the invention relates to detection of a scene change and providing the scene change termination points as an indicator to an encoder to enable more efficient video encoding.

RELATED ART

[0003] Scene transitions such as cuts, fades and dissolves adversely affect encoding efficiency because of increased temporal entropy. In other words, as video frames are processed in an encoder, there is a lack of order of predictability when scene changes occur that do not enable the encoder to efficiently allocate bits for encoding to particular frames. Accurate detection of scene transitions is, thus, desirable for encoding. Along with allocation of bits for encoding, detection of a scene change accurately has at least the three following potential uses:

[0004] (1) Enabling specifying a fresh reference picture (I-picture) to be provided as a first clean picture after the scene change transition is complete.

[0005] (2) Enabling a fade termination point to be determined for computed weighted predictions; and

[0006] (3) Enabling beginning and ending termination points of a dissolve to be determined in weighted bi-prediction.

[0007] Detection of scene transition points is usually accomplished by examining reduced picture luma statistics, such as picture-average and block-histograms. However, this method is fragile since entirely different scenes can still have very similar luma variations. It is, thus, desirable to provide improved methods of detecting scene transitions.

SUMMARY

[0008] Embodiments of the present invention provide a method to better detect a scene change to provide a prediction to an encoder to enable more efficient encoding. The method uses a Motion Compensated Temporal Filter (MCTF) that provides a block of picture pre-processing prior to an encoder where the MCTF performs motion-estimation, and produces a motion-compensation prediction on temporally sequential pictures. The MCTF block provides a measure of the Motion Compensated Residual (MCR) that is used in embodiments of the present invention to detect the scene change. The MCR measure is not subject to the limitation of scene luma similarity, and provides a more accurate scene change detection method.

[0009] When a scene is relatively stable, the MCR score is also relatively stable. Both the mean MCR score, and the MCR variance, varies slowly within a scene. However, when a scene transition is in process, MCR score behavior changes because whatever temporal predictability there was within the scene is compromised. Algorithmically, the MCR score is used to predict scene change points by comparing the sliding mean of the MCR score to the sliding median. This comparison highlights the scene transition terminals. In the case of a scene cut, the MCR score exhibits a distinct spike. In the case of a fade or dissolve, the MCR score exhibits a transitional period of degradation followed by recovery.

[0010] By implementing the above detection using the MCR, the location of the I-pictures in the downstream encoding process can be accurately determined for the encoder. Further bits can be better allocated during the encoding process with the scene change terminal points more accurately predicted.

In accordance with an aspect of the present invention there is provided a method for encoding video using scene change detection comprising: obtaining video frames provided to an encoder; obtaining a motion compensated residual (MCR) for the video frames; determining a sliding MCR score for individual ones of the video frames; determining a sliding mean of the MCR score for the video frames; comparing the MCR score with the MCR mean score; and providing a prediction of when a scene change occurs to the encoder based on the comparison of the MCR score with the MCR mean, wherein the MCR measurement is provided from a Motion Compensated Temporal Filter (MCTF) that provides pre-processing prior to encoding to perform motion estimation as well as motion compensation prediction on temporally sequential pictures of the video frames.

In accordance with another aspect of the present invention there is provided an apparatus to encode video frames, the apparatus comprising: an encoder having a first input for receiving video frames to be processed and a second input for receiving parameter data to enable the encoder to allocate bits for frames for encoding; a pre-filter with Motion Compensated Temporal Filtering (MCTF) frame buffer having an input receiving the video frames and an output providing the first input to the encoder; a MCTF statistical analysis module processor that provides a Motion Compensated Residual (MCR) for receiving the video frames from the pre-filter with MCTF and having an output providing the second input to the encoder; a preprocessor memory connected to the MCTF statistical analysis processor for storing code that is executable by the preprocessor to determine the parameter data to enable the encoder to allocate bits, the code causing the preprocessor to perform the following steps: obtaining a MCR for the video frames; determining a sliding MCR score for individual ones of the video frames; determining a sliding mean of the MCR score for the video frames; comparing the MCR score with the MCR mean score; and providing a prediction of when a scene change occurs to the encoder based on the comparison of the MCR score with the MCR mean, wherein the MCR is provided from the MCTF frame buffer that provides pre-processing prior to encoding to perform motion estimation as well as motion compensation prediction on temporally sequential pictures of the video frames.

BRIEF DESCRIPTION OF THE DRAWINGS

[0011] Further details of the present invention are explained with the help of the attached drawings in which:

[0012] Fig. 1 shows a graph of frames vs. bit allocation for a video clip with a cross-fade scene change;

[0013] Fig. 2 shows frame 212 from Fig. 1, where the cross-fade starts

[0014] Fig. 3 shows frame 228 from Fig. 1, where the cross-fade mid-point occurs;

[0015] Fig. 4 shows frame 240 from Fig. 1, where the cross-fade termination was detected using prior art methods;

[0016] Fig. 5 shows frame 244 from Fig. 1, where the cross-fade termination actually occurs;

[0017] Fig. 6 shows a graph of frames vs. bit allocation for a video clip where a scene cut occurs at frame 1304;

[0018] Fig. 7 shows the scene cut start that is graphically illustrated in Fig. 6;

[0019] Fig. 8 shows the scene cut end that is graphically illustrated in Fig. 6.

[0020] Fig. 9 shows the MCR measurement values plotted for the two cases shown in Figs. 1-5 and Figs. 6-8, respectively;

[0021] Fig. 10 is a flow chart showing steps for identification of a frame for a scene change according to embodiments of the present invention; and

[0022] Figs. 11 illustrates components needed for encoding to implement embodiments of the present invention.

DETAILED DESCRIPTION

[0023] For embodiments of the present invention, if a scene clip can be identified as a scene change, good video quality can be achieved in particular by more accurately determining the location for I-pictures. In H262, H264, and H265 (also known as MPEG2, AVC and HEVC respectively) video encoding, Intra-coded pictures (I-pictures) are decodable in isolation, and are not dependent upon information from other pictures. They are used to terminate error propagation, and to provide a clean reference for future Inter-predicted pictures (P and B-pictures). As such, it is important to place I-pictures, within the encoded video stream, at locations that provide a tactical advantage for efficient encoding. When considering

scene changes, the optimum place for an I-picture is at the immediate beginning of a new scene; that being the first picture after the completion of a fade-in, cross-fade, or scene-cut.

[0024] To help understand how a determination of when a frame is a scene change according to the present invention, several different scene clips of data along with per frame bit allocation and MCR score are analyzed.

[0025] Fig. 1 plots frames of a video clip with frames numbered along the along an X-axis with allocated bits shown on the Y-axis. The video clip values plotted in Fig. 1 shows a typical $\frac{1}{2}$ second cross-fade, in an MPEG-2 encoded file. Specific frames 212, 228, 240 and 244 form the beginning, middle and termination of the cross-fade scene change that are shown in subsequent figures. The red lines indicate where prior art scene change points are identified by previous scene change detection methods.

[0026] Figs. 2-5 show the start through termination frames for the cross-fade graphed in Fig. 1. Fig. 2 shows frame 212 of the video clip graphically illustrated in Fig. 1 where the cross fade scene change begins. In Fig 2, the initial scene is a close-up of the nose of the space shuttle, and the new scene being added with cross-fade is a scene with the complete space shuttle and launch pad. Fig. 3 shows frame 228 of the video clip where the cross fade mid-point alpha blended frame occurs. Fig. 4 shows frame 240 of the video clip where the cross fade termination was detected using previous detection methods. In the encoded original content, the end of the fade was not handled optimally. Frame 240 was encoded as an anchor I-frame, but as can be seen in Fig. 4 in area 400 the original space shuttle nose cone scene can still be seen as the cross-fade is still in progress. Finally, Fig. 5 shows frame 244 of the video clip where the cross fade termination actually occurs, the point that termination is detected using embodiments of the present invention.

[0027] Fig. 6 shows frames of a video clip with a scene cut that occurs at frame 1304, near the middle of the graph. In Fig. 6 the frames are numbered along the X-axis and allocated bits are shown on the Y-axis. Figs. 7-8 show the specific scenes before and after scenes for the scene cut of Fig. 6. Fig. 7 shows the scene cut start with the storage bay of the space shuttle. Finally, Fig. 8 shows the scene cut end scene which is the solar panels of the space station.

[0028] Fig. 9 shows MCTF motion compensated residual (MCR) measurement values plotted on the Y-axis respectively versus the frame numbers plotted on the X-axis for the two cases shown in Figs. 1-5 and Figs. 6-8, respectively. In the two cases shown in Figs. 1-5 and Figs. 6-8, the reduced MCR in Fig. 9 shows the upcoming scene transition location prior to encoding. For the cross-fade of Figs. 1-5, the MCR score exhibits a transitional period of degradation followed by recovery which terminates at the termination of frame 244. For the scene cut in Figs. 6-8, when the scene cut frame 1304 is encountered, the MCR score exhibits a distinct spike.

[0029] Fig. 10 is a flow chart showing steps for identification of a frame for a scene change according to embodiments of the present invention. In step 1000 pictures are first taken from a frame buffer that feeds an encoder, one frame at a time for evaluation. Next, in step 1002 the frame is put through the MCTF with a MCR score determined. Next, in step 1004 for the MCR score, a sliding window, mean, variance and linear regression is determined for a group of frames previously received in the frame buffer preceding the encoder, including the present frame being evaluated, to enable later evaluation of the present frame to be performed.

[0030] The next steps of Fig. 10 use the MCR score to evaluate the present frame. In step 1006 a determination if the present scene is stable is made. The scene stability is determined based on whether the MCR linear regression is flat and the MCR variance is low. If not,

operation returns to step 1000 and a next frame is taken from the buffer. If so, the present frame may be the target frame for a scene change and operation proceeds to step 1008. In step 1008, a further evaluation of the present scene is made to determine if the present frame is the target frame that ends the scene change. The evaluation is performed by looking at the target MCR relative to the mean of previous frames, whether the previous frame MCR was significantly larger and whether the regression from the previous frame was negative. If in step 1008 the evaluation does not determine this is the target frame, control returns to step 1000 where the next frame is obtained from the buffer. If in step 1008 the frame is determined to be the target frame that terminates the scene change, control proceeds to step 1010 where the frame is marked as a key picture for promotion as the target frame to the encoder.

[0031] Fig. 11 illustrates components needed to implement embodiments of the present invention. The system shown in Fig. 11 is an integrated pre-filter and two pass encoder. Frames of a video for encoding are provided to buffer registers of a pre-filter 1100 that provides MCTF evaluation of the video frames. A new MCTF statistics evaluation element 1120, that provides MCR values and analysis, receives the output from the pre-filter 1100. The MCTF statistics evaluation element 1120 provides information such as setting of the MCR sliding window for evaluation of frames, the MCR mean of a group of frames, and the variance and linear regression of the MCR. The MCTF statistics evaluation element 1120 further provides information such as determination of a target MCR mean value variance that a termination frame may have, as well as providing a comparison of the present frame being evaluated with the values determined from frames in the sliding evaluation window. The statistics output from the MCTF statistics evaluation element 1120 are then provided to subsequent components for encoding.

[0032] The encoding system of Fig. 11 further includes elements that provide two-pass encoding with a first path including the elements: Encode preparation with reordering element 1102, Encode with reference picture and Motion Estimation (ME) element 1106 and Post encode analysis element 1108. The first path provides statistical data to enable more efficient encoding in the second path. The MCTF statistics evaluation element 1120 output is provided to elements 1102 and 1106 in the first path. Also, the video data output from the pre-filter 1100 is provided through the first path elements 1102, 1106 and 1108.

[0033] The second path in Fig. 11 includes the elements: Delay buffer 1104, Encode preparation with Reordering element 1110 and Encode with reference picture and ME element 1112. The output of the statistics evaluation element 1120 is provided to the elements 1110 and 1112. The video data output from the pre-filter 1100 is also provided through the second path elements 1104, 1110 and 1112. The statistical data from the first path is further provided to elements 1110 and 1112 of the second path to enable more efficient encoding.

[0034] The system shown in Fig. 11 has the Pre-Filter with MCTF element 1100 integrated with the remainder of the encoder components. The MCTF based MCR data may not otherwise be available to the encoding process, because the pre-filter is either in an isolated hardware component, or because the low-level MCR results are not exposed outside the native MCTF hardware. The MCR scores are also more smoothly behaved than a similar process that could be operated using a single pass encoding process.

[0035] For components shown in Fig. 11, the elements can be controlled by one or more processors linked to one or more memories to enable operation. The memory stores code that is executable by a processor to enable the processor to perform the processes described herein.

Further the memory can be used to provide data storage with the data accessible by the processor to store or retrieve when performing operations.

[0036] Although the present invention has been described above with particularity, this was merely to teach one of ordinary skill in the art how to make and use the invention. Many additional modifications will fall within the scope of the invention as that scope is defined by the following claims.

CLAIMS

What is claimed:

1. A method for encoding video using scene change detection comprising:
 - obtaining video frames provided to an encoder;
 - obtaining a motion compensated residual (MCR) for individual ones of the video frames;
 - determining a MCR score for the individual ones of the video frames;
 - determining a mean of the MCR score for the video frames;
 - comparing the MCR score with the MCR mean score; and
 - providing a prediction of when a scene change occurs to the encoder based on the comparison of the MCR score with the MCR mean,
 - wherein the MCR is provided from a Motion Compensated Temporal Filter (MCTF) that provides pre-processing prior to encoding to perform motion estimation as well as motion compensation prediction on temporally sequential pictures of the video frames.
2. The method of claim 1, wherein the scene change comprises at least one of a fade, a dissolve or a scene cut.
3. The method of claim 1, wherein when the scene change comprises a scene cut, the MCR score exhibits a spike.
4. The method of claim 1, wherein when the scene change comprises a fade or a dissolve, the MCR score exhibits a transitional period of degradation followed by a recovery.

5. The method of claim 1, wherein the prediction of when a scene change occurs identifies the transition terminals at the beginning and the end of the scene change.
6. The method of claim 1 further comprising placing an I-frame during the encoding process at a point where the scene change occurs.
7. An apparatus to encode video frames, the apparatus comprising:
 - an encoder having a first input for receiving video frames to be processed and a second input for receiving parameter data to enable the encoder to allocate bits for frames for encoding;
 - a pre-filter with Motion Compensated Temporal Filtering (MCTF) frame buffer having an input receiving the video frames and an output providing the first input to the encoder;
 - a MCTF statistical analysis module processor that provides a Motion Compensated Residual (MCR) for receiving the video frames from the pre-filter with MCTF and having an output providing the second input to the encoder;
 - a preprocessor memory connected to the MCTF statistical analysis processor for storing code that is executable by the preprocessor to determine the parameter data to enable the encoder to allocate bits, the code causing the preprocessor to perform the following steps:
 - obtaining a MCR for individual ones of the video frames;
 - determining a MCR score for the individual ones of the video frames;
 - determining a mean of the MCR score for the video frames;
 - comparing the MCR score with the MCR mean score; and

providing a prediction of when a scene change occurs to the encoder based on the comparison of the MCR score with the MCR mean,

wherein the MCR is provided from the MCTF frame buffer that provides pre-processing prior to encoding to perform motion estimation as well as motion compensation prediction on temporally sequential pictures of the video frames.

8. The apparatus of claim 7, wherein the scene change comprises at least one of a fade, a dissolve or a scene cut.

9. The apparatus of claim 7, wherein when the scene change comprises a scene cut, the MCR score exhibits a spike.

10. The apparatus of claim 7, wherein when the scene change comprises a fade or a dissolve, the MCR score exhibits a transitional period of degradation followed by a recovery.

11. The apparatus of claim 7, wherein the prediction of when a scene change occurs identifies the transition terminals at the beginning and the end of the scene change.

12. The apparatus of claim 7 further comprising placing an I-frame during the encoding process at a point where the scene change occurs.

13. The apparatus of claim 7, wherein the encoder is a two path device with the MCTF statistical analysis module processor provided in a first path and the encoder provided in the second path.

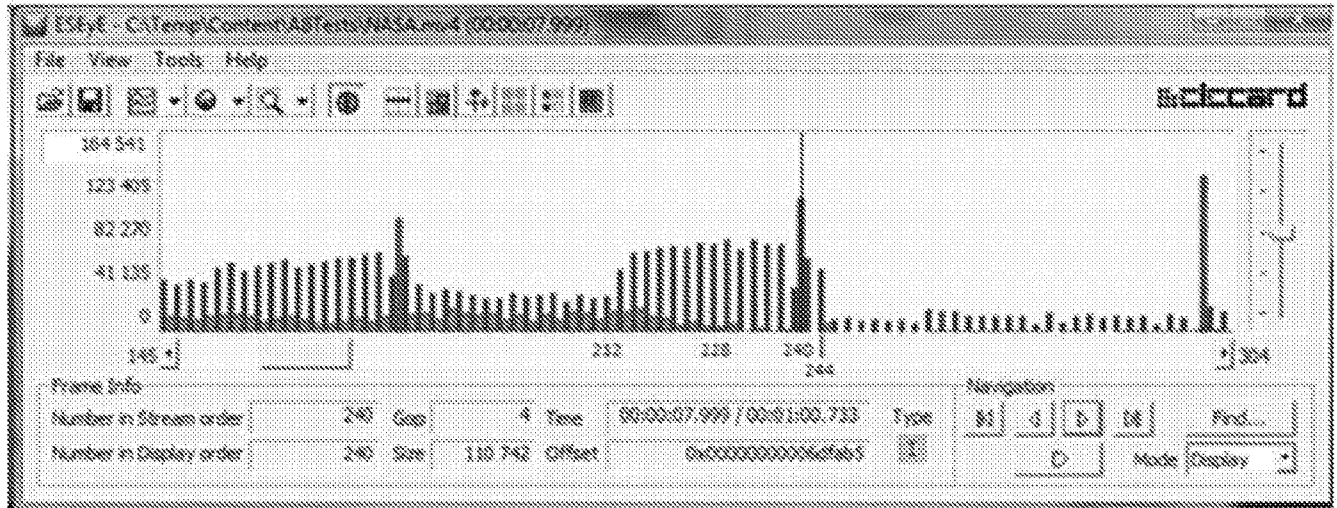
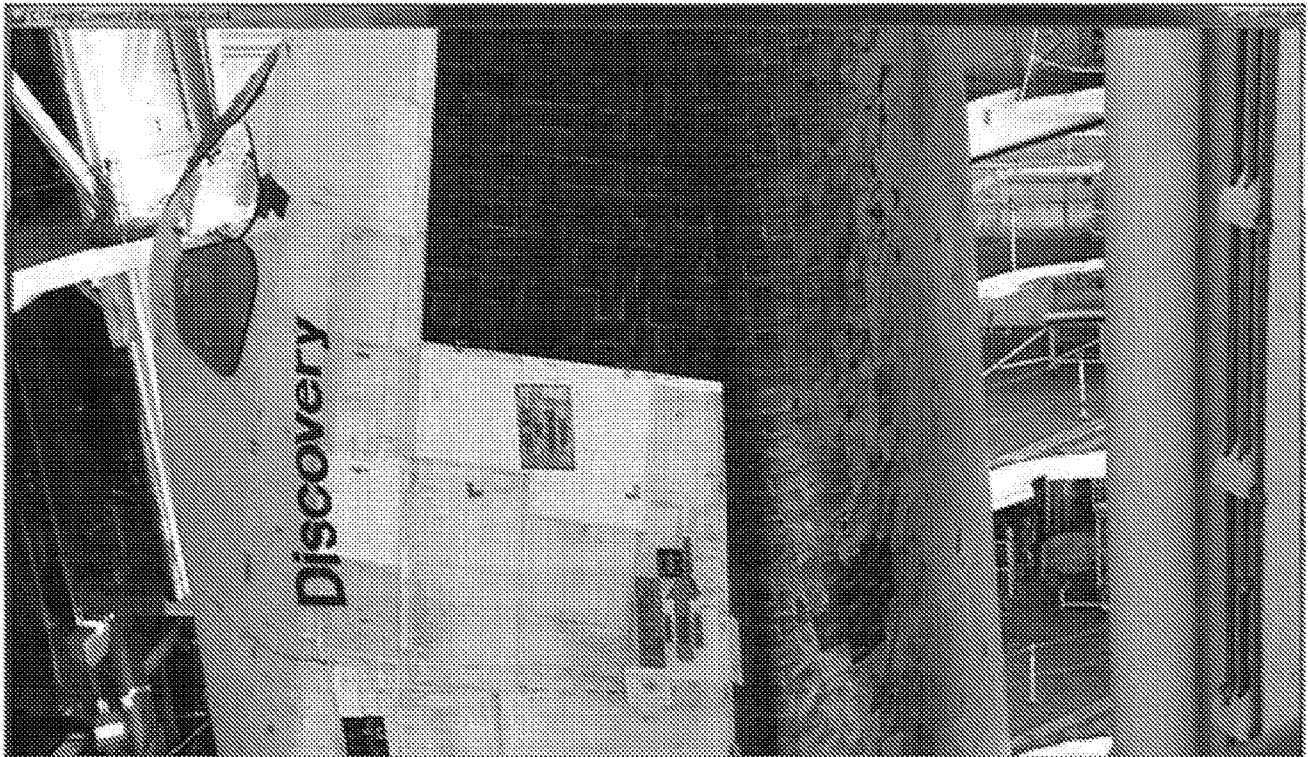
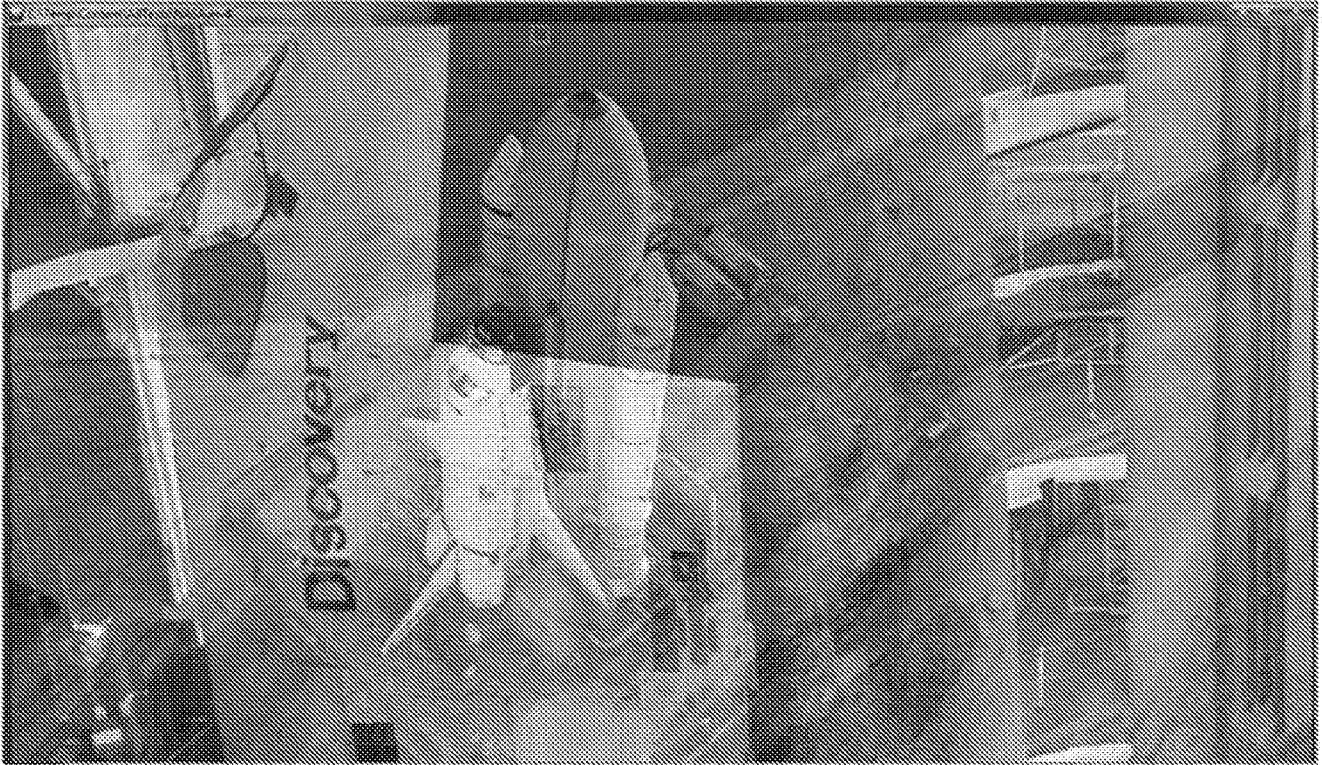


FIG. 1



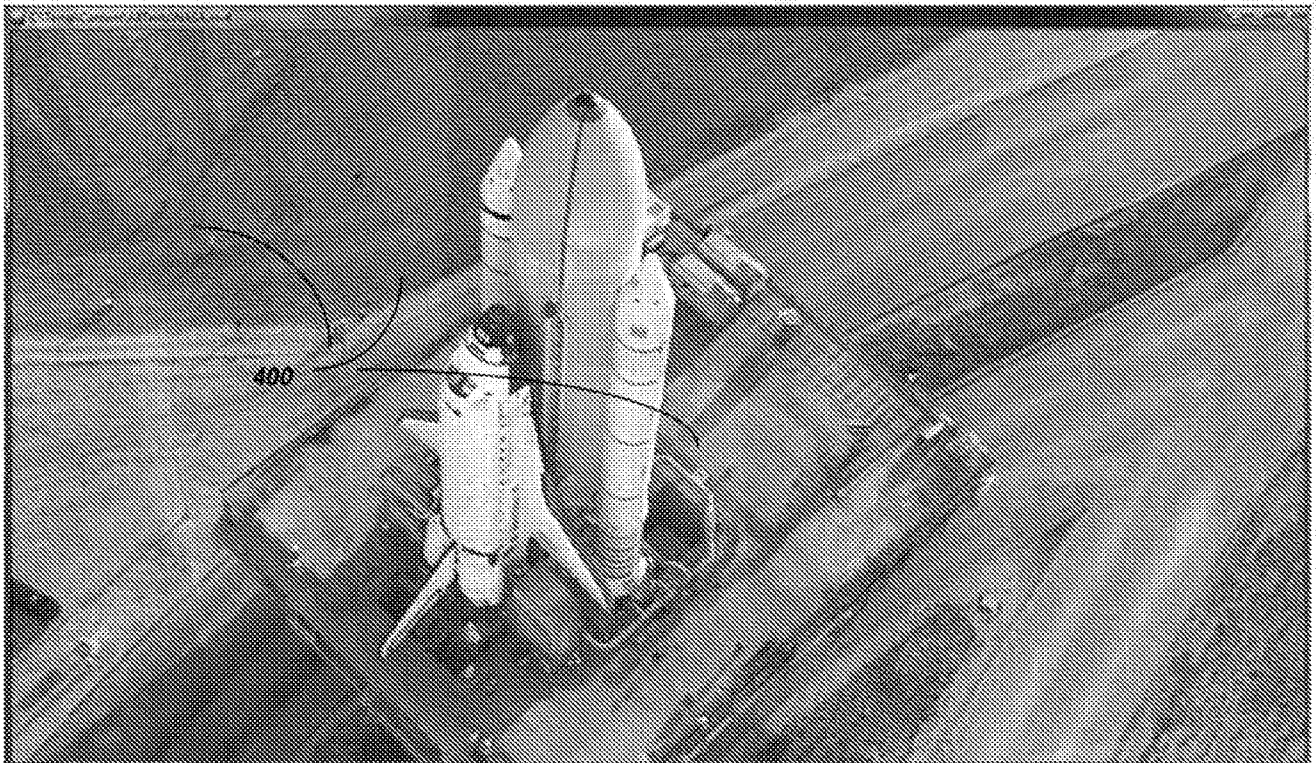
Frame 212: Cross Fade Start Scene

FIG. 2



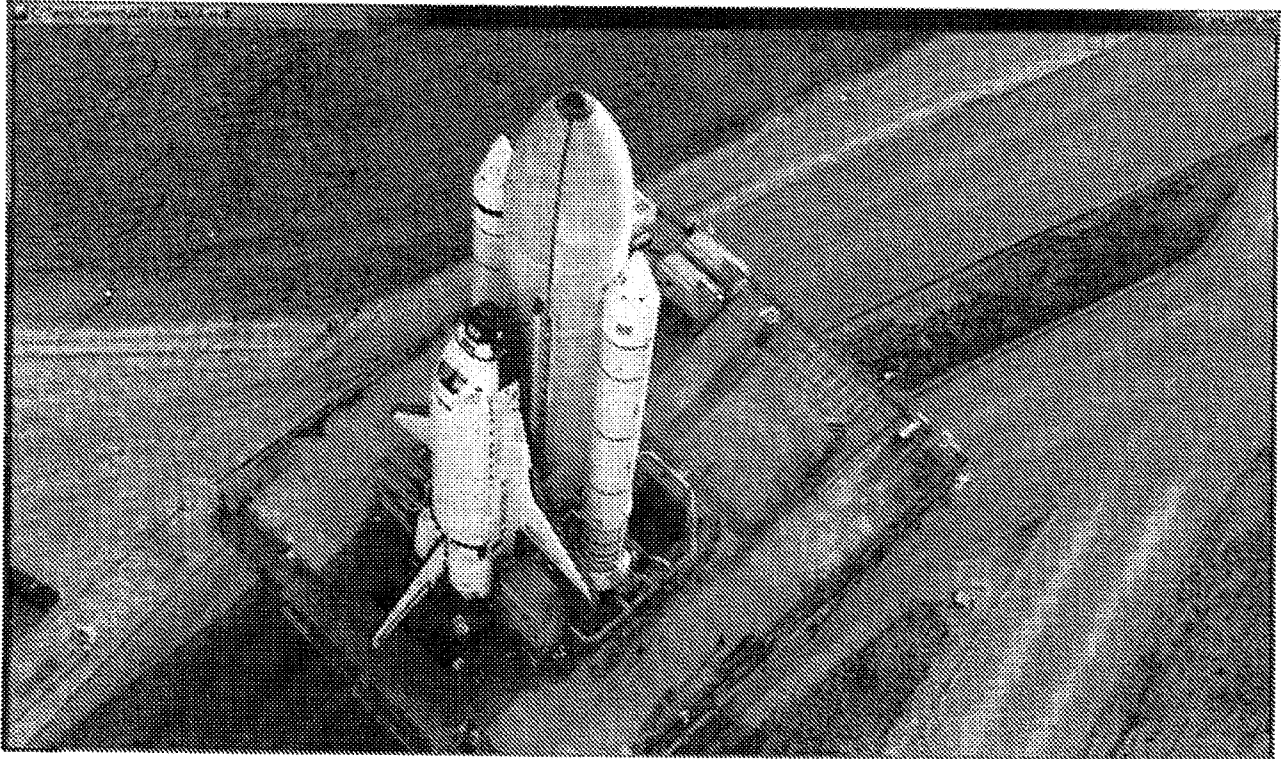
Frame 228: Cross Fade Mid-Scene

FIG. 3



Frame 240: Cross Fade Termination Scene 1

FIG. 4



Frame 244: Cross Termination Scene 2

FIG. 5

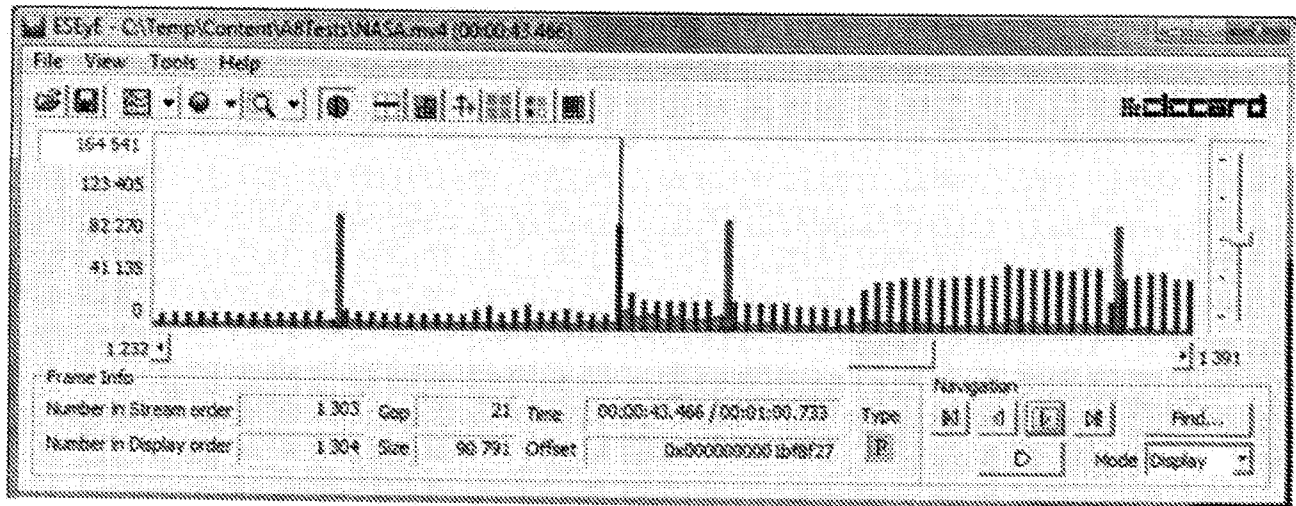


FIG. 6

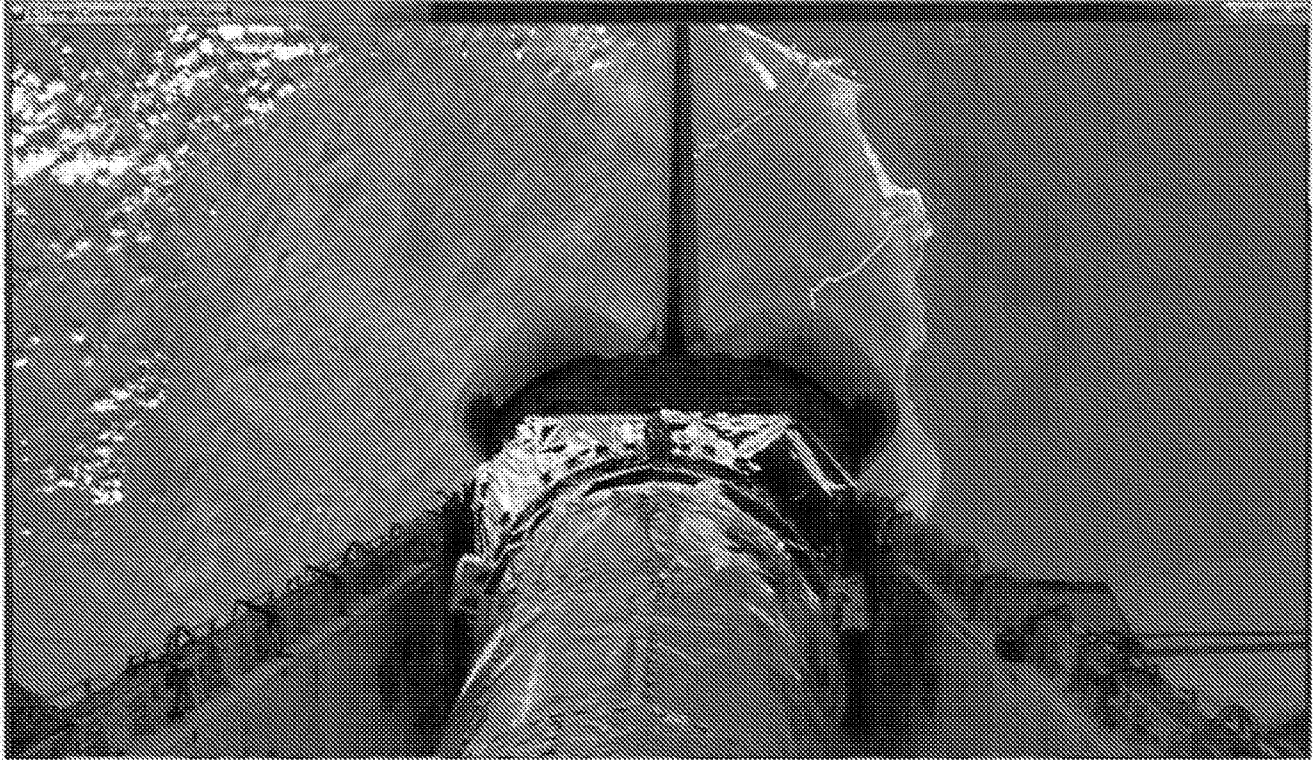


FIG. 7

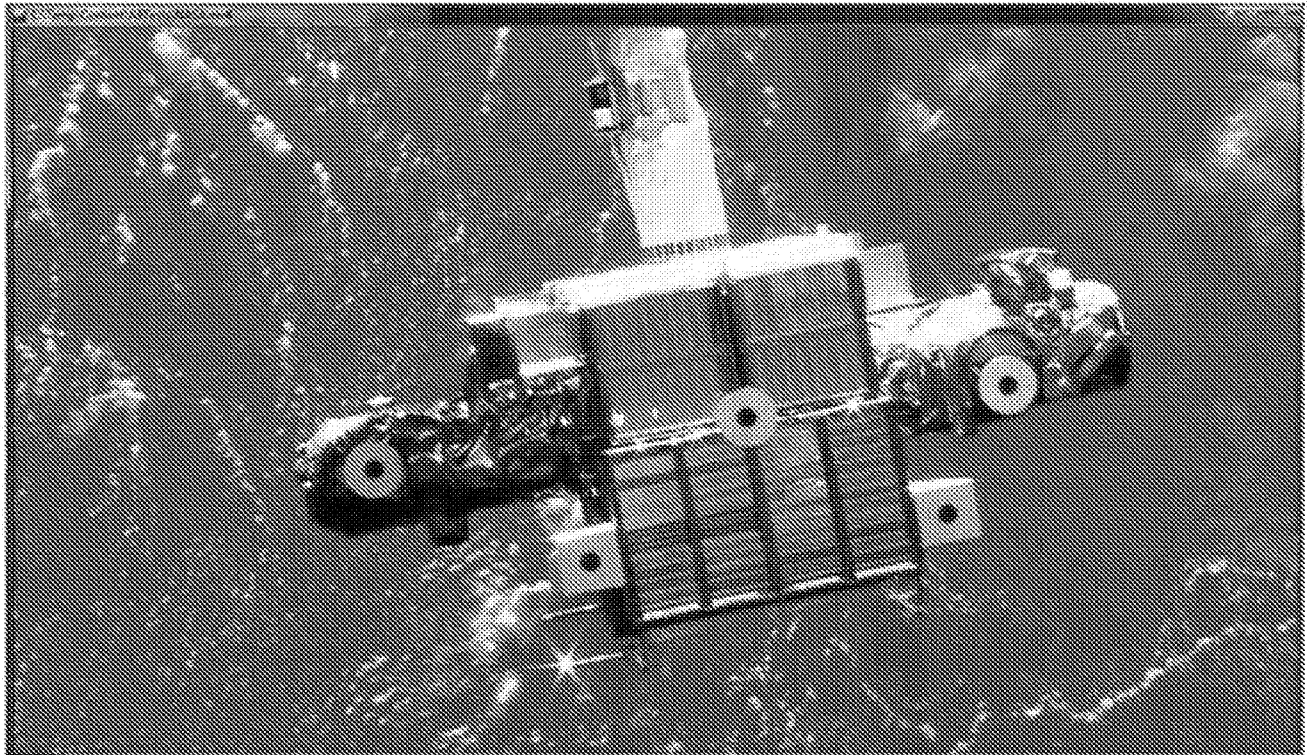


FIG. 8

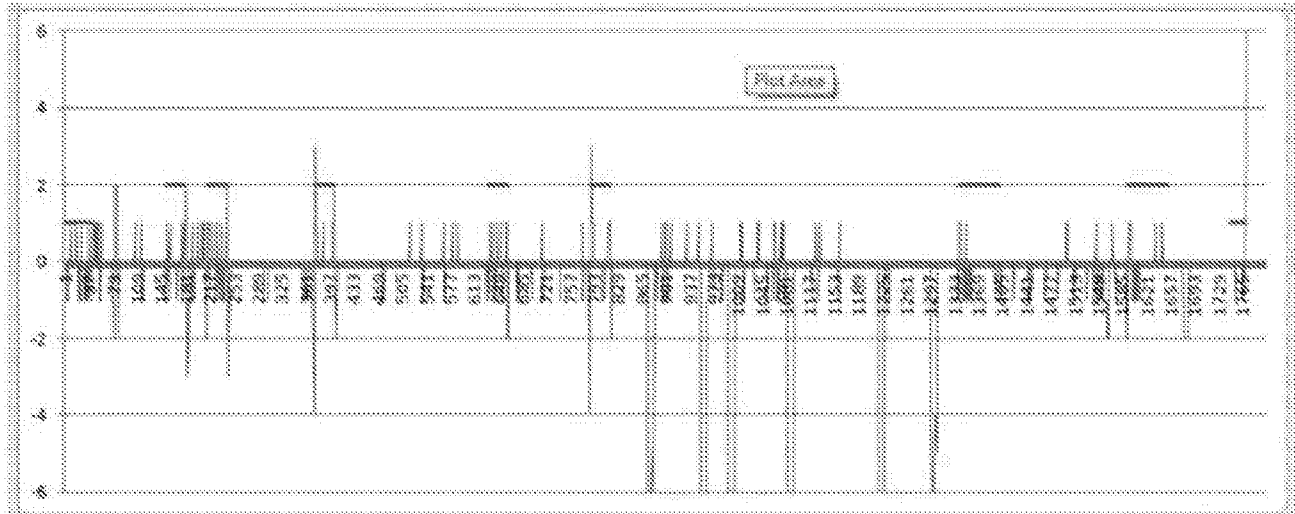


FIG. 9

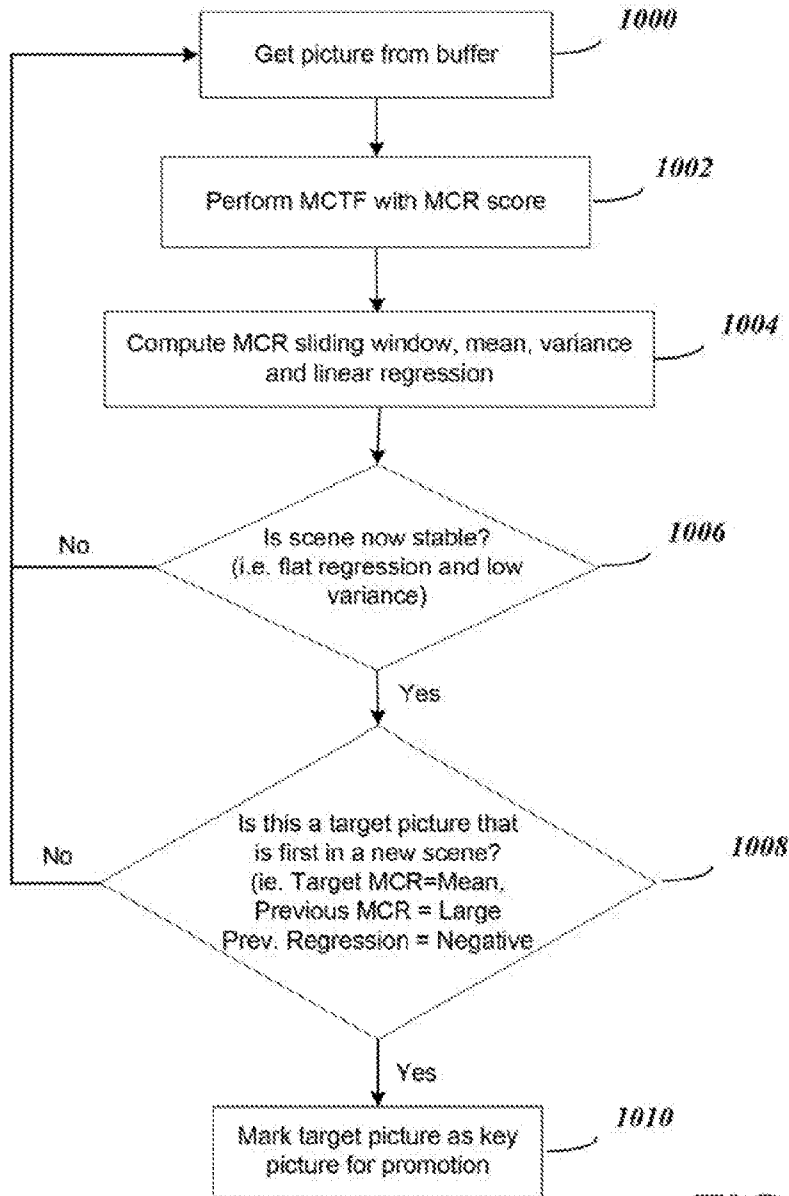


FIG. 10

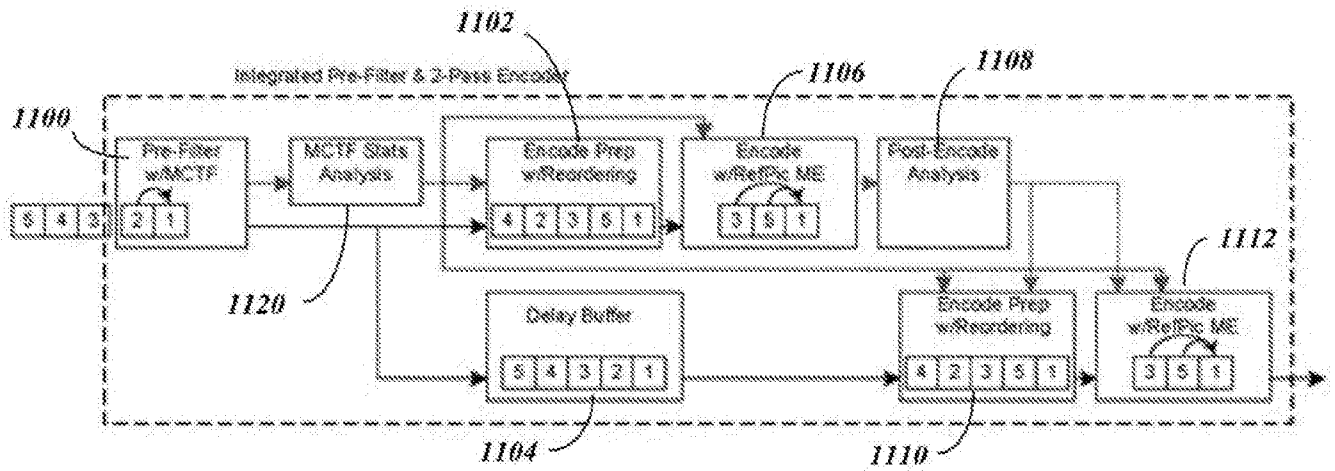


FIG. 11

