



(51) International Patent Classification:
G06T 5/00 (2006.01) **G06K 9/00** (2006.01)
G06T 7/20 (2006.01)

(21) International Application Number:
PCT/US2010/002028

(22) International Filing Date:
19 July 2010 (19.07.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/271,381 20 July 2009 (20.07.2009) US

(71) Applicant (for all designated States except US): **THOMSON LICENSING** [FR/FR]; 1 rue Jeanne d'Arc, F-92130 Issy-les moulineaux (FR).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **BHAGAVATHY, Sitaram** [IN/US]; 5910 Hunters Glen Drive, Plainsboro, New Jersey 08536 (US). **ZHANG, Dong-Qing** [CN/US];

4812 Quail Ridge Drive, Plainsboro, New Jersey 08536 (US).

(74) Agents: **SHEDD, Robert, D.** et al.; c/o Thomson Licensing LLC, 2 Independence Way, Suite 200, Princeton, New Jersey 08540 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, 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, 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, PE, PG, PH, PL, PT, RO, RS, RU, 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, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ,

[Continued on next page]

(54) Title: A METHOD FOR DETECTING AND ADAPTING VIDEO PROCESSING FOR FAR-VIEW SCENES IN SPORTS VIDEO

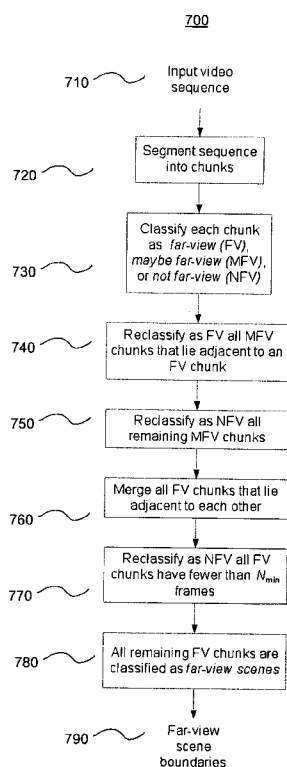


FIG. 7

(57) Abstract: A method is disclosed for analyzing video to detect far-view scenes in sports video to determine when certain image processing algorithms should be applied. The method comprises analyzing and classifying the fields of view of images from a video signal, creating and classifying the fields of view of sets of sequential images, and selectively applying image processing algorithms to sets of sequential images representing a particular type of field of view.



TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, **Published:**

EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU,

LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK,

SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,

GW, ML, MR, NE, SN, TD, TG).

— *with international search report (Art. 21(3))*

A METHOD FOR DETECTING AND ADAPTING VIDEO PROCESSING FOR FAR-VIEW SCENES IN SPORTS VIDEO

CROSS REFERENCE TO RELATED APPLICATION

This application claims priority to and all benefits accruing from provisional
5 application filed in the United States Patent and Trademark Office on July 20, 2009
and assigned serial number 61/271,381.

BACKGROUND OF THE INVENTION

The present invention generally relates to a method and associated apparatus
for analyzing video to detect far-view scenes in sports video to determine when
10 certain image processing algorithms should be applied. The method comprises
analyzing and classifying the fields of view of images from a video signal, creating
and classifying the fields of view of sets of sequential images, and selectively
applying image processing algorithms to sets of sequential images representing a
particular type of field of view.

15 This section is intended to introduce the reader to various aspects of art,
which may be related to various aspects of the present invention that are described
below. This discussion is believed to be helpful in providing the reader with
background information to facilitate a better understanding of the various aspects of
the present invention. Accordingly, it should be understood that these statements
20 are to be read in this light, and not as admissions of prior art.

As mobile devices have become more capable and mobile digital television
standards have evolved, it has become increasingly practical to view video
programming on such devices. The small screens of these devices, however,
present some limitations, particularly for the viewing of sporting events. Small
25 objects, such as the ball in a sporting event, can be difficult to see. The use of high

video compression ratios can exacerbate the situation by significantly degrading the appearance of small objects like a ball in a far-view scene.

5 It is possible to enhance the appearance of these objects, but such algorithms may be computationally costly or degrade the overall image quality if applied when not needed. It would be desirable to be able to detect particular types of scenes that could benefit from object enhancement algorithms such that the algorithms may be selectively applied. The invention described herein addresses this and/or other problems.

SUMMARY OF THE INVENTION

In order to solve the problems described above, the present invention concerns analyzing video to detect far-view scenes in sports video to determine when certain image processing algorithms should be applied. The method
5 comprises analyzing and classifying the fields of view of images from a video signal, creating and classifying the fields of view of sets of sequential images, and selectively applying image processing algorithms to sets of sequential images representing a particular type of field of view. This and other aspects of the invention will be described in detail with reference to the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

The above-mentioned and other features and advantages of this invention, and the manner of attaining them, will become more apparent, and the invention will be better understood, by reference to the following description of embodiments of the invention taken in conjunction with the accompanying drawings, wherein:

FIG. 1 is a diagram of a prior art video processing system with object localization and enhancement;

FIG. 2 is a diagram of a video processing system with object localization and enhancement utilizing far-view scene detection;

FIG. 3 is a flowchart of a method for classifying the field of view represented by a frame;

FIG. 4 is an illustration of the results of applying portions of a method for classifying the field of view represented by a frame;

FIG. 5 is a flowchart of a method of segmenting a video sequence into chunks;

FIG. 6 is a flowchart of a method of classifying the field of view of a chunk of a video sequence; and

FIG. 7 is a flowchart of a method of far-view scene detection utilizing the chunk-level classifications.

The exemplifications set out herein illustrate preferred embodiments of the invention. Such exemplifications are not to be construed as limiting the scope of the invention in any manner.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

As described herein, the present invention provides a method and associated apparatus for analyzing video to detect far-view scenes in sports video to determine when certain image processing algorithms should be applied. The method
5 comprises analyzing and classifying the fields of view of images from a video signal, creating and classifying the fields of view of sets of sequential images, and selectively applying image processing algorithms to sets of sequential images representing a particular type of field of view.

While this invention has been described as having a preferred design, the
10 present invention can be further modified within the spirit and scope of this disclosure. This application is therefore intended to cover any variations, uses, or adaptations of the invention using its general principles. Further, this application is intended to cover such departures from the present disclosure as come within known or customary practice in the art to which this invention pertains and which fall within
15 the limits of the appended claims.

In a preferred embodiment, the present invention may be implemented in signal processing hardware within a television production or transmission environment. The method is used to detect far-view scenes in sports video, with an exemplary application in soccer video. Far-view scenes are those corresponding to
20 wide-angle camera views of the play field, wherein the objects of interest, for instance, the players and ball, are small enough to be easily degraded by video compression or not be clearly visible.

Figure 1 illustrates a conventional image processing system 100 with object
25 localization and enhancement applied to all portions of a video, regardless of the field of view of a particular scene. Input video 110 is first processed through object localization algorithms 120 to produce information 130 about detected objects, such as balls or pucks. Various techniques are known in the art for detecting such

objects, such as the use of filters corresponding to characteristics of the object of interest.

Information 130 about the identified objects is passed to the object enhancement stage 150 and to an object aware encoder 170. This information may include, for instance, location, size, trajectory, or a mask of a ball. Video frames 140 are processed at object enhancement stage 150 using the information 130 about the identified objects. For instance, a highlight color may be placed over the location of a ball or puck to allow the viewer to more easily identify its location.

The resulting video frames 160 with enhancement applied to the detected objects are then encoded by the object-aware encoder 170, resulting in an output bitstream 180. The use of object information 130 by the object-aware encoder 170 may allow the encoding to be adjusted to preserve the visibility and appearance of identified objects in the far-view scenes, such as players or the ball. For instance, a lower compression ratio may be used for scenes in which the ball appears as a small object, or for particular areas of frames where the ball appears.

In a system like system 100, object localization and enhancement are performed without regard to the type of view represented in the frames being processed. Thus, unnecessary processing is performed on some types of scenes, potentially resulting in wasted time, wasted processing resources, or image quality degradation.

Figure 2 illustrates the data flow in a video processing system in accordance with the present invention. The input video 210 is first processed by a far-view scene detection algorithm 220, which will be described in detail below. With the addition of far-view scene detection 220, the scenes 230 classified as far-view may then be processed separately from those scenes 240 classified as not being far-view. This may provide savings in time, savings of processing resources, or improvements in image quality.

Far-view scenes 230 are sent to object localization and enhancement processing 250. This processing may include highlighting of a detected ball, illustration of a ball trajectory, or other enhancements. Non-far-view scenes 240 bypass the object localization and enhancement stage and are sent directly to the object-aware encoder 280. Object information 260 produced by the object localization and enhancement stage 250 and the enhanced far-view scenes 270 are also sent to the object-aware encoder 280, which produces an encoded output bitstream 290.

As described earlier, the use of object information 260 by the object-aware encoder 280 allows the encoding to be adjusted, for instance, to preserve the visibility and appearance of identified objects, such as soccer balls, in the far-view scenes. For instance, a lower compression ratio may be used for scenes in which the ball appears, or for particular areas of frames where the ball appears.

The detection of far-view scenes of step 220 comprises the following stages:

1. Classification of each frame in the sequence as far-view (FV), maybe far-view (MFV), or not far-view (NFV), described below with respect to Figure 3.
2. Segmentation of the sequence into a number of chunks, each chunk being a contiguous set of "similar-looking" frames, described below with respect to Figure 5.
3. Classification of each chunk as FV, MFV, or NFV, described below with respect to Figure 6.
4. Detection of far-view scenes by merging and reclassifying chunks, described below with respect to Figure 7.

FIG. 3 is a flowchart of a method 300 of far-view frame classification, described with respect to an exemplary application of soccer video. An input frame 310 from a soccer video is processed into a binary field mask at a soccer field

detection step 320. In a simple case, the soccer field pixels are identified using the knowledge that the field is generally covered in grass or grass-colored material. The result is a binary mask classifying all field pixels with a value of 1 and all non-field pixels, including objects in the field, with a value of 0. Various image processing techniques may then be used to then identify the boundaries of the playing field, ignore foreground objects, and create a solid field mask. For instance, all pixels within a simple bounding box encompassing all of the contiguous regions of field pixels above a certain area threshold may be included in the field mask. Other techniques, including the use of filters, may be used to identify the field and eliminate foreground objects from the field mask.

At step 330, "player-like" objects are identified through analysis of the foreground objects, which are connected sets of non-field pixels within the field boundary identified in step 320. For a foreground object o with constituent pixels $\{(x_i, y_i)\}$, the following object parameters are computed:

area: $a_o = \sum_i 1$, (i.e., the number of pixels in the object),

height: $h_o = \max_i y_i - \min_i y_i + 1$,

width: $w_o = \max_i x_i - \min_i x_i + 1$,

compactness: $c_o = a_o / (h_o \times w_o)$, and

aspect ratio: $r_o = h_o / w_o$.

Objects are considered "player-like" when the area, compactness, and aspect ratio each exceed a threshold related to known characteristics of players in a soccer video. Stated otherwise, if $(a_o > t_a)$, $(c_o > t_c)$, and $(r_o > t_r)$, the object o is considered "player-like." In a preferred embodiment directed toward 320×240-pixel soccer videos, the threshold values of $t_a = 10$, $t_c = 0.1$, and $t_r = 0.5$ are used.

At step 340, the maximum area $A_{\max}$ and median area A_{med} of all player-like objects are computed. As described above, the area of a particular object may be computed simply as the number of pixels comprising the object. Likewise, the area of the field A_{field} may be computed as the number of pixels comprising the field mask.

5 At step 350, the area of the field, the median area of player objects, and the maximum area of player objects are compared to thresholds related to the expected areas in a far-view scene. If ($A_{\text{field}} > T_{\text{field}}$), ($A_{\text{med}} < T_{\text{med}}$), and ($t_{\max} < A_{\max} < T_{\max}$), the frame is labeled as FV at step 360. That is, if the field area in the frame is large, the median player area is small, and the maximum player area is within an expected
10 range, the field-of-view of the scene is wide, or far.

If the frame is not labeled as FV, further determinations are made as to whether the frame may be far view, or is most likely not far view. The frame is classified as MFV at step 380 if the criteria ($A_{\text{field}} > t_{\text{field}}$) and ($A_{\max} \leq t_{\max}$) are met at step 370. Stated otherwise, if the field area is above a lower threshold, but the
15 maximum area of a player-like object is not above a minimum threshold, a reliable determination cannot be made based upon the single frame. If the frame is not labeled as FV or MFV, it is classified as NFV at step 390.

In a preferred embodiment directed toward 320×240-pixel soccer videos, the threshold values of $T_{\text{field}} = 0.4 \times H \times W$ (40% of the frame area), $t_{\text{field}} = 0.3 \times H \times W$ (30%
20 of the frame area), $T_{\text{med}} = 600$, $t_{\max} = 100$, and $T_{\max} = 2400$, were used, where H and W are the height and width of the frame in pixels.

The steps of method 300 are repeated for each frame in the sequence. Thus, a sequence of classifications of field of view, one per frame, is obtained.

Figure 4 illustrates the field detection and player-like object detection
25 processes of steps 320 and 330. Given a frame 410, the soccer field detection first results in a binary mask 420, classifying all field-colored pixels with a value of 1 and all non-field pixels, including objects in the field, with a value of 0. Objects on the field, such as players, lines, and the ball, appear as holes in the mask since they are

not the expected color of the field. The result of the determination of the boundaries of the soccer field is shown in 430. The holes in the mask from players, lines, and the ball are removed, creating a large contiguous field mask. Image 440 shows the detected foreground objects, the connected sets of non-field pixels within the field boundaries shown in 430. These include, for instance, binary masks of players, field lines, and the ball. Image 450 represents the result of the detection of player-like objects. Objects not meeting the thresholds for area, compactness, or aspect ratio, such as the field lines, have been eliminated.

The frame-level classification of method 300 will generally produce some erroneous frame classifications. In order to mitigate the effects of such errors, the video is segmented into sets of contiguous "similar-looking" frames called "chunks." The process of identifying chunks is described below with respect to Figure 5. The field of view of each chunk is then classified as a whole based on the statistics of the classifications of its constituent frames, as described below with respect to Figure 6.

Fig. 5 outlines a method 500 used for segmenting a video sequence 510 into a number of chunks based upon similarity of color histograms of contiguous sequences of frames. A color histogram consists of counts of the number of pixels in the frame within various ranges of color values. Frames from the same scene are likely to have similar histograms, as the camera will be pointed at the same objects, which generally have constant color. A change of scene will place different objects of different colors within the field of view, thereby generally changing the color histogram.

At step 515, the list of chunks C is empty. The starting frame number j , the first frame in the chunk under construction, is initialized to a value of 1, the first frame in the video sequence.

At step 520, the color histogram of the j^{th} frame is computed. In a preferred embodiment, H_j is a 256-bin histogram of grayscale values of pixels in frame j . A smaller number of larger histogram bins may be utilized to reduce the computational

intensity of histogram comparison. The color histogram of the j^{th} frame serves as a basis for assembly of the chunk.

At step 525, a loop begins for frames $i = j+1, \dots, N$, where N is the number of the last frame in the sequence. The frames following frame j will be analyzed one at a time for similarity to frame j to determine if they should be included in the chunk that begins with frame j .

At step 530, the color histogram H_i of the i^{th} frame is computed using the same technique used at step 520. Then at step 535, the histogram difference between the i^{th} frame and the j^{th} frame, d_{ij} is computed. The difference may be computed as $d_{ij} = \|H_i - H_j\|_1$, where $\|\cdot\|_1$ refers to the 1-norm or sum of absolute differences (SAD) between two vectors.

At step 540, a determination is made based on the difference d_{ij} from step 535 as to whether the color histogram of frame i is similar enough to that of frame j for frame i to be included in the chunk beginning with frame j . The frames are considered similar enough if the distance d_{ij} between their color histograms H_i and H_j is below a threshold T_{chunk} . In a preferred embodiment, $T_{\text{chunk}} = 0.4$.

If $d_{ij} > T_{\text{chunk}}$, that is, if the color histogram difference of frame i from frame j is too large for frame i to be included in the chunk, the interval $[j, i-1]$ is added to the list of chunks at step 545. Thus, the current chunk is terminated at the previous frame, frame $i-1$, the last frame meeting the similarity threshold. The starting frame number j for the new chunk is set to the current value of i at step 565, and the process returns to step 520 for building of the next chunk.

However, if d_{ij} is less than or equal to T_{chunk} , the frame i is considered similar enough to the initial frame of the chunk, j , to be added to the current chunk. A determination is then made at step 555 as to whether $i = N$ (i.e., the last frame in the video has been reached). If not, the process returns to the beginning of the loop at step 525 and the next frame is considered for inclusion in the current chunk.

The ending frame number of the chunk under construction is thereby increased until either a sufficiently dissimilar frame is located or until the last frame is reached. If the last frame has been reached, that is, $i = N$, the interval $[j \ N]$ consisting of the final frames is added to the list of chunks 575 at step 570 and the process terminates.

At the end of this process, a list of chunks C has been produced. Each chunk is represented by a pair $[b \ e]$, where b is the beginning frame of the chunk and e is the ending frame.

Fig. 6 shows the method 600 used to classify the field of view of each chunk as FV, MFV, or NFV. In a preferred embodiment, chunk classification is based on a strict form of majority voting among the labels of the frames in the chunk, mitigating labeling errors that occurred in the frame-level classification.

At step 620, each frame of the input video chunk 610 is classified as FV, MFV, or NFV. This frame-level classification may be performed using method 300 described above.

At step 630, the percentage of FV frames is computed for the chunk. If more than 50% of the frames in the chunk are determined to be FV at step 640, the whole chunk is classified as FV at step 650. That is, if the majority of constituent frames are far view, the chunk is considered far view.

If the percentage of FV frames is not above 50%, the percentage of MFV frames in the chunk is computed at step 660. If more than 50% of frames are determined to be MFV at step 670, the chunk is classified as MFV at step 680. If neither criterion is satisfied, the chunk is classified as NFV at step 690. In an alternative embodiment, chunks may be classified as NFV if the frame count is below a certain threshold.

Fig. 7 is a flowchart of a method 700 used in the overall far-view scene detection process. An input video sequence 710 is segmented into chunks at step

720, as described above with respect to Figure 5. At step 730, each chunk is classified as FV, MFV, or NFV, as described with respect to Figure 6.

At step 740, if an MFV chunk lies adjacent to an FV chunk, it is reclassified as FV. That is, if a determination regarding the field of view could not be made at step 730, the chunk will be considered far-view if it is adjacent to a far-view chunk. A chunk $[b_1 \ e_1]$ is said to be adjacent to a chunk $[b_2 \ e_2]$ if $b_1 = e_2 + 1$ or $e_1 = b_2 - 1$. In a preferred embodiment, only MFV chunks adjacent to original FV chunks are reclassified and reclassification based on adjacency to other reclassified FV chunks is not allowed.

At step 750, all remaining MFV chunks are reclassified as NFV. That is, if a determination regarding the field of view of the chunk could not be made at step 620 and the chunk is not adjacent to a chunk identified as far-view, the chunk will be assumed to not be far-view.

At step 760, the process merges all FV chunks that lie adjacent to each other into larger chunks. By merging two adjacent FV chunks $[b_1 \ e_1]$ and $[b_2 \ e_2]$, where $e_1 = b_2 - 1$, C is modified by removing these two chunks and adding $[b_1 \ e_2]$ in their place. The new chunk inherits the FV label of its constituent parts. This merging process is repetitive and is performed until there are no adjacent FV chunks left.

If an FV chunk has fewer than $N_{\min}$ frames, it may be reclassified as NFV. In one particular embodiment, $N_{\min}$ is chosen to be 30 frames. Thus, processing of short scenes may be avoided.

Finally, at step 780, all of the remaining FV chunks in C are classified as far-view scenes. The beginning and ending frames of each FV chunk indicate the boundaries 790 of the corresponding far-view scene. Similarly, all remaining NFV chunks are classified as non-far-view scenes (after merging adjacent ones as described earlier).

At the end of this process, a list of far-view scenes S_{FV} and a list of non-far-view scenes S_{NFV} are obtained. Each scene is represented by a pair $[b\ e]$, where b is the beginning frame of the scene and e is its end frame. These scene boundaries may then be used by the object highlighting system shown in Fig. 2 to control the application of image processing techniques applicable only to certain views. The far-view scenes are sent to the object localization and enhancement modules that enhance objects of interest and generate object metadata for the object-aware encoder. The non-far-view scenes may be sent directly to the encoder to be encoded without object highlighting.

While the present invention has been described in terms of a specific embodiment, it will be appreciated that modifications may be made which will fall within the scope of the invention. For example, various processing steps may be implemented separately or combined, and may be implemented in general purpose or dedicated data processing hardware or in software. The overall complexity of the method may be reduced by relaxing the criteria for objects in the field to be considered in decision-making. For example, instead of detecting player-like objects, all objects larger than a threshold area may be considered. Also, instead of using grayscale pixel values for computing histograms during chunk segmentation, it is possible to use full-color values (e.g. RGB, YUV). Furthermore, distance measures other than the SAD may be used for comparing histograms. Instead of using threshold-based criteria for frame classification, a classifier (e.g., support vector machine) may be learned from a labeled training set of FV, MFV, and NFV frames. Furthermore, the proposed method may be applied to other sports or events with moving objects of interest. Finally, the method may be used to detect other types of scenes than far view for specialized processing.

CLAIMS

1. A method of processing video in a video processing system comprising the steps of:

analyzing a plurality of images from a video signal;

5 classifying the field of view of each of said plurality of images from said video signal in response to at least one size of a feature of the image;

creating sets of sequential images from said plurality of images from said video signal based upon comparisons of metrics of adjacent images;

classifying the field of view of each of said sets of sequential images; and

10 selectively applying an image processing algorithm to at least those sets of sequential images representing a particular type of field of view.

2. The method of claim 1 wherein each of said plurality of images from said video signal is a frame.

3. The method of claim 1 wherein said size is a length, width, or area.

15 4. The method of claim 1 wherein said metrics comprise color histogram data.

5. The method of claim 1 wherein classifying the field of view of an image comprises associating the image with data indicating whether the image represents a far field of view shot, may represent a far field of view shot, or does not represent a far field of view shot.

20 6. The method of claim 1 wherein classifying the field of view of an image comprises creating a mask representing areas of the image that may represent a playing field.

7. The method of claim 6 wherein classifying the field of view of an image comprises comparing a computed area of said mask with at least one threshold.

8. The method of claim 6 wherein classifying the field of view of an image comprises identifying and computing areas of player-like objects within the
5 boundaries of said mask.

9. The method of claim 8 wherein classifying the field of view of an image comprises comparing the maximum computed area or median computed area of said at least one - objects with a threshold.

10. The method of claim 1 wherein classifying the field of view of a set of
10 sequential images comprises a statistical analysis of the classifications of the images in the set.

11. The method of claim 10 wherein said statistical analysis comprises identifying a classification pertaining to more than half of the images from the video signal in the set.

15

12. An apparatus for processing video in a video processing system comprising:

means for analyzing a plurality of images from a video signal;

5 means for classifying the field of view of each of said plurality of images from said video signal in response to at least one size of a feature of the image;

means for creating sets of sequential images from said plurality of images from said video signal based upon comparisons of metrics of adjacent images;

means for classifying the field of view of each of said sets of sequential images; and

10 means for selectively applying an image processing algorithm to at least those sets of sequential images representing a particular type of field of view.

13. The apparatus of claim 12 wherein each of said plurality of images from said video signal is a frame.

14. The apparatus of claim 12 wherein classifying the field of view of an image
15 comprises associating the image with data indicating whether the image represents a far field of view shot, may represent a far field of view shot, or does not represent a far field of view shot.

15. The apparatus of claim 12 wherein classifying the field of view of an image
20 comprises creating a mask representing areas of the image that may represent a playing field.

16. The apparatus of claim 15 wherein classifying the field of view of an image comprises comparing a computed area of said mask with at least one threshold.

17. The apparatus of claim 15 wherein classifying the field of view of an image comprises identifying and computing areas of player-like objects within the boundaries of said mask.

5 18. The apparatus of claim 12 wherein classifying the field of view of a set of sequential images comprises a statistical analysis of the classifications of the images in the set.

19. An apparatus for processing video in a video processing system comprising:

a processor for:

10 analyzing a plurality of images from a video signal;

classifying the field of view of each of said plurality of images from said video signal in response to at least one size of a feature of the image;

creating sets of sequential images from said plurality of images from said video signal based upon comparisons of metrics of adjacent images;

15 classifying the field of view of each of said sets of sequential images;
and

selectively applying an image processing algorithm to at least those sets of sequential images representing a particular type of field of view.

20 20. The apparatus of claim wherein classifying the field of view of a set of sequential images comprises a statistical analysis of the classifications of the images in the set.

100

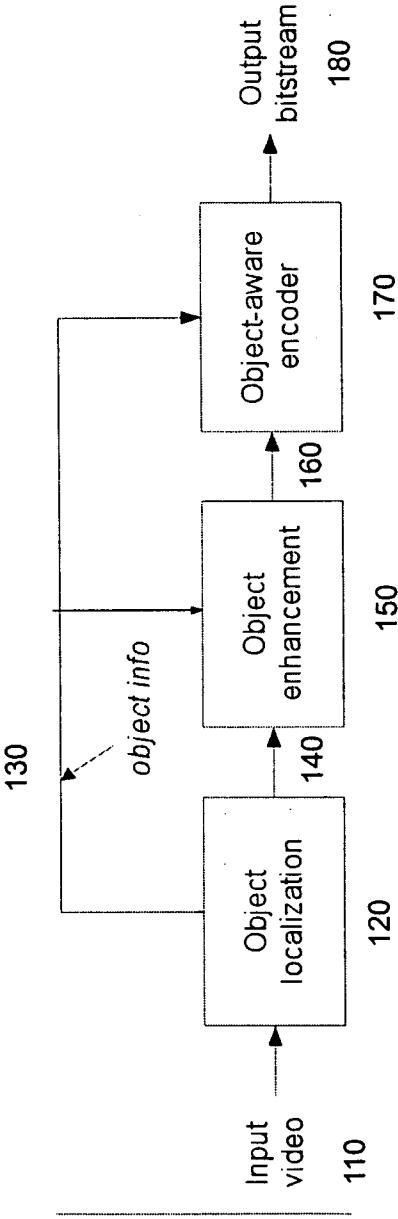


FIG. 1 PRIOR ART

200

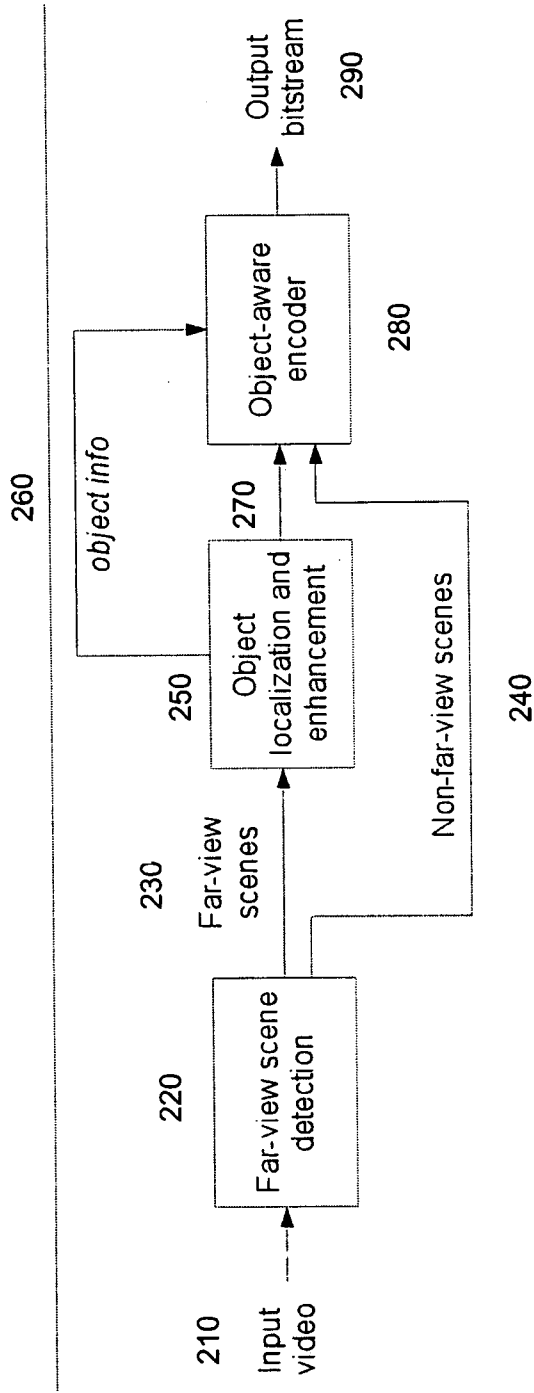


FIG. 2

3/7

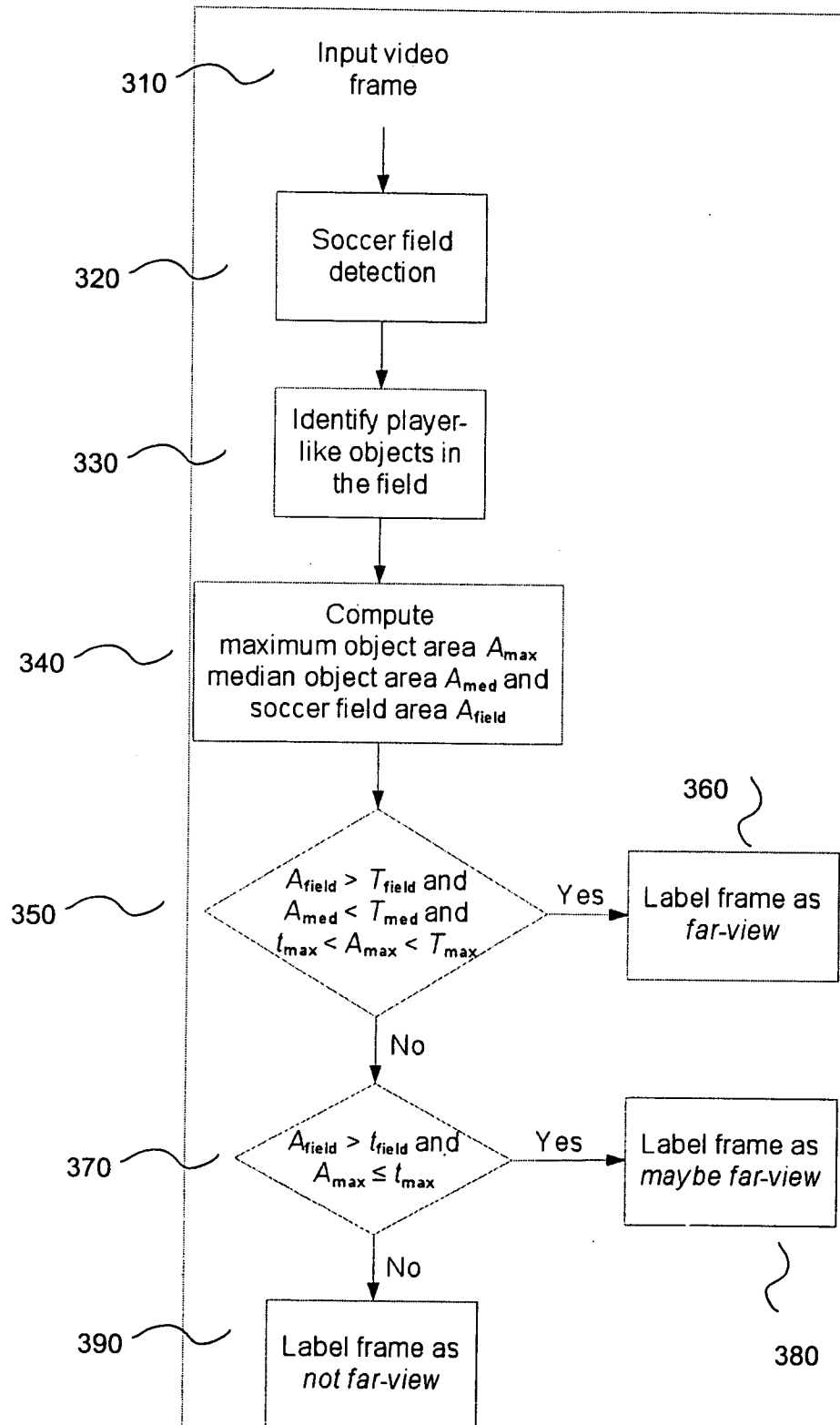
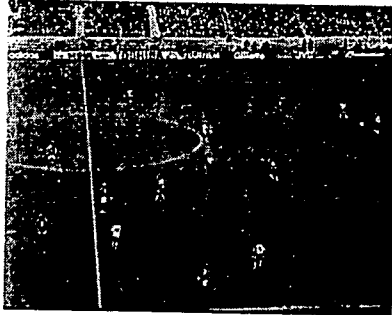
300

FIG. 3

4/7

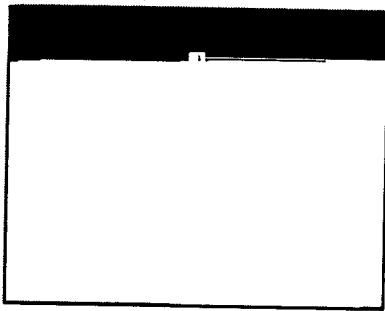
400



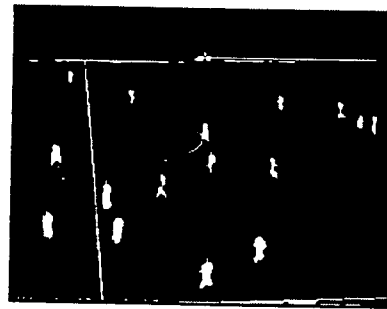
410



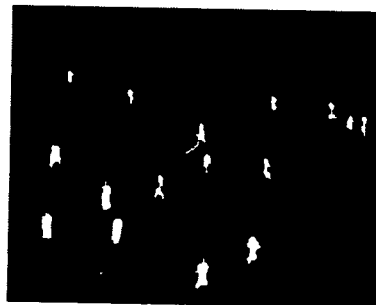
420



430



440



450

FIG. 4

5/7

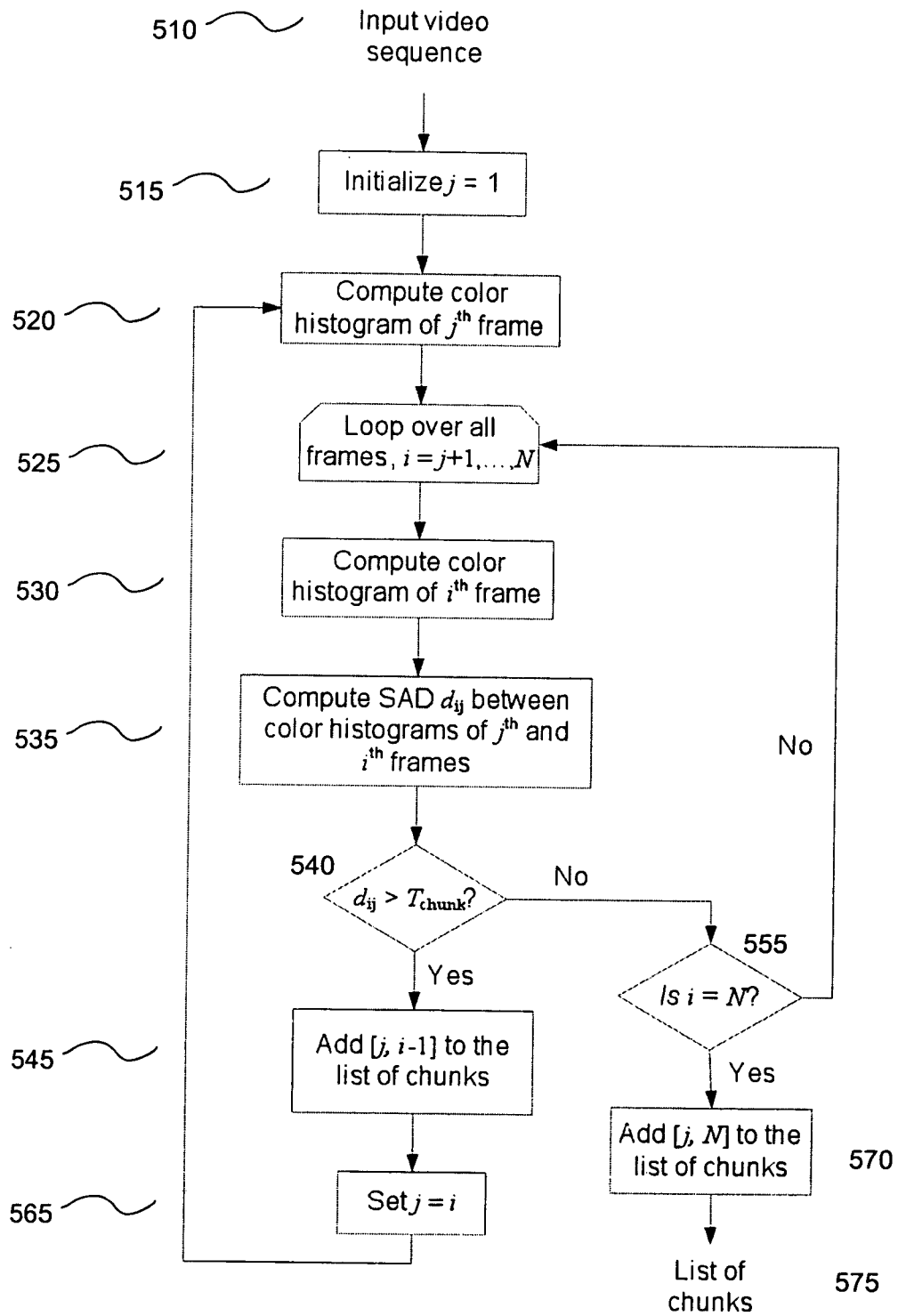
500

FIG. 5

6/7

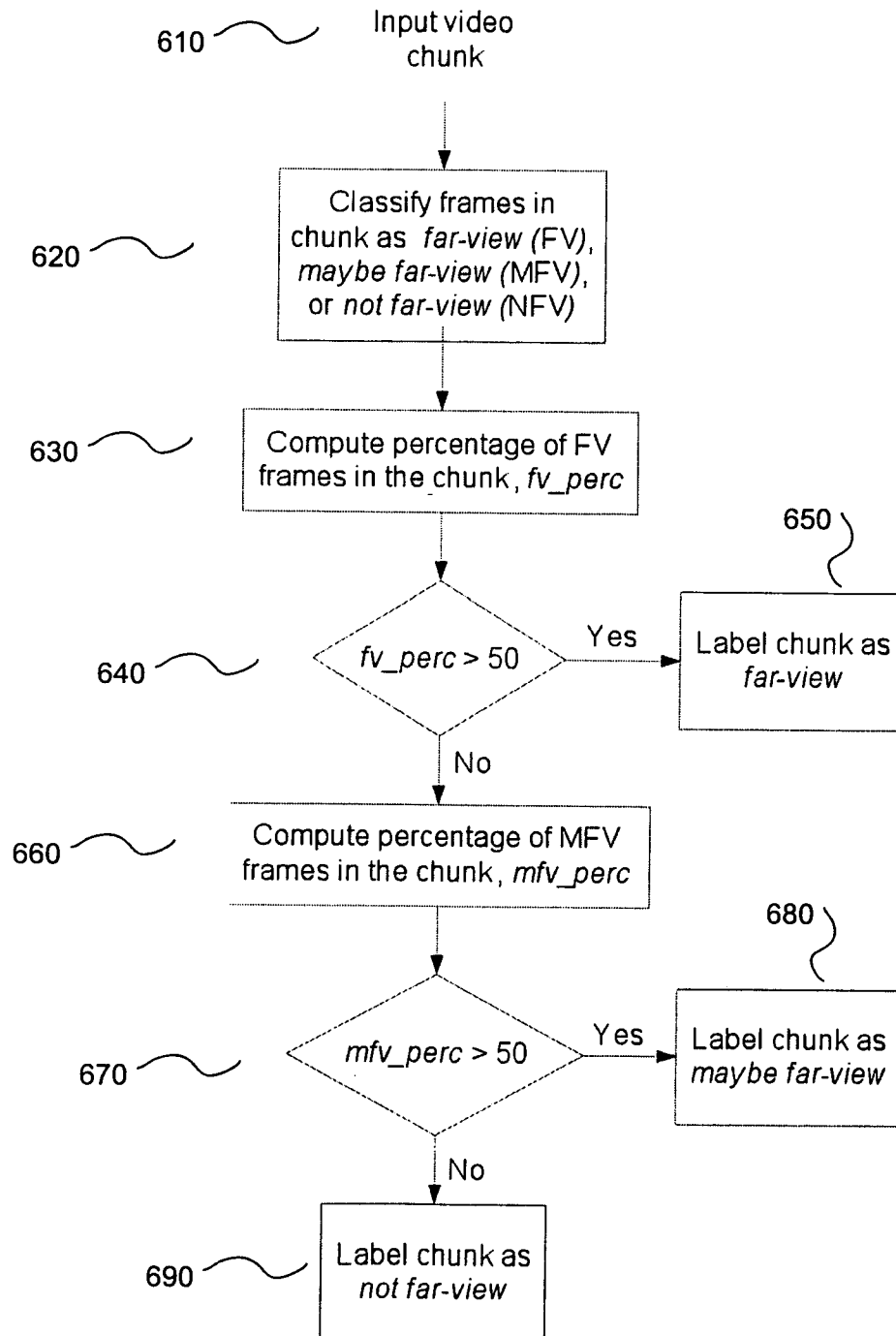
600

FIG. 6

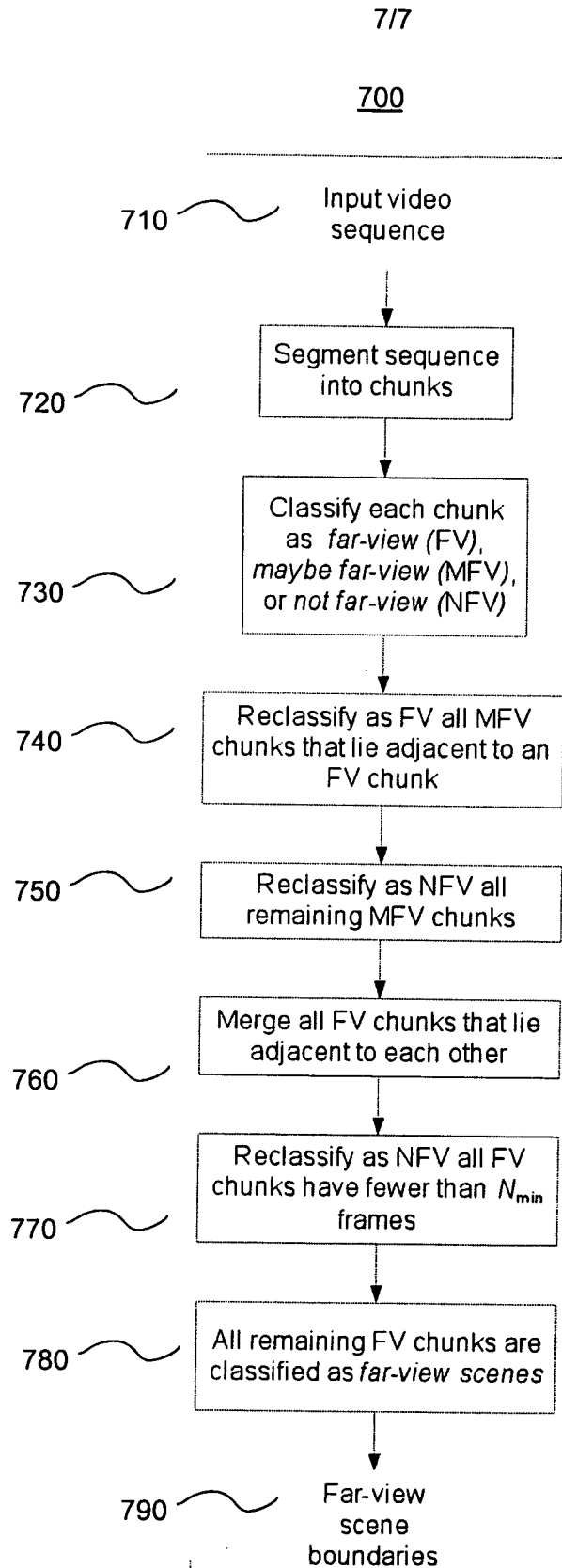


FIG. 7

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2010/002028

A. CLASSIFICATION OF SUBJECT MATTER

INV. G06T5/00 G06T7/20 G06K9/00
ADD.

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)

G06T G06K

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

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

EP0-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2007/242088 A1 (KIM CHANGICK [KR] ET AL) 18 October 2007 (2007-10-18) * abstract paragraphs [0025] - [0028], [0049] - [0063]	1-9, 12-17,19
A	----- EKIN A ET AL: "Framework for tracking and analysis of soccer video" PROCEEDINGS OF THE INTERNATIONAL SOCIETY FOR OPTICAL ENGINEERING (SPIE), SPIE, USA LNKD- DOI:10.1117/12.453120, vol. 4671, 1 January 2002 (2002-01-01), pages 763-774, XP002393277 ISSN: 0277-786X * abstract Section 2, figure 3 ----- -/--	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 document 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 another 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

14 October 2010

Date of mailing of the international search report

27/10/2010

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Kessler, Cordula

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2010/002028

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>SADHER D A ET AL: "A combined audio-visual contribution to event detection in field sports broadcast video. case study: Gaelic football" SIGNAL PROCESSING AND INFORMATION TECHNOLOGY, 2003. ISSPIT 2003. PROCEEDINGS OF THE 3RD IEEE INTERNATIONAL SYMPOSIUM ON DARMSTADT, GERMANY 14-17 DEC. 2003, PISCATAWAY, NJ, USA, IEEE, 14 December 2003 (2003-12-14), pages 552-555, XP010729214 ISBN: 978-0-7803-8292-3 * abstract paragraphs [05.1], [06.3] -----</p>	1-20

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2010/002028

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2007242088 A1	18-10-2007	KR 20060060630 A	05-06-2006