



- (51) International Patent Classification:  
*G06K 9/00* (2006.01)
- (21) International Application Number:  
PCT/EP2016/056064
- (22) International Filing Date:  
18 March 2016 (18.03.2016)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
S2015/0085 19 March 2015 (19.03.2015) IE
- (71) Applicant: **IDASO LIMITED** [IE/IE]; Curraghmore, Mullingar, County Westmeath (IE).
- (72) Inventors: **KERR, Ian**; 22 Pearse Street, Mullingar, County Westmeath (IE). **GRAHAM, Mervyn**; Drumna-cooha, Ballinallee, County Longford (IE). **KRISTÓF, Csorba**; Csanak u. 13, 4032 Debrecen (HU). **MÁRTON BÁLINT, Varga**; Farkasréti út 102, 2040 Budaörs (HU).

**LEVENTE, Lipták**; Fehérhegyi utca 15, 3881 Abaújszántó (HU). **LÁSZLÓ, Kundra**; Wéber Ede park 56, 2145 Kerepes (HU). **FODOR, Szabolcs**; Bornemissza u. 18 fszt. 1, 1119 Budapest (HU). **ALBERT, Istvan**; Dobsinai utca 6/b, 1124 Budapest (HU).

(74) Agents: **LITTON, Rory** et al.; Cruickshank & Co., 8A Sandycroft Business Centre, Sandycroft, Dublin 18 (IE).

(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,

[Continued on next page]

(54) Title: A METHOD OF ANALYSING A MULTIMEDIA FILE

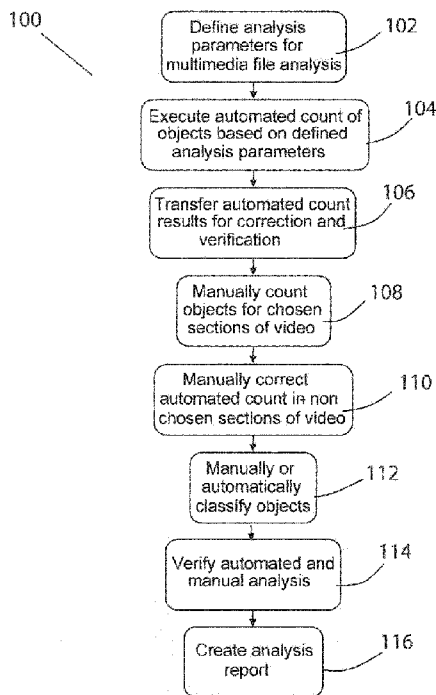


Figure 1

(57) Abstract: The present invention is directed to a method of analysing a multimedia file for the detection, counting, tracking and/or classification of objects displayed in images contained within the multimedia file. The method comprises the steps of executing an automated analysis of the multimedia file and then performing a manual analysis of selected portions of the multimedia file. The analysis comprises observing objects seen in images of the multimedia file by detecting, counting, tracking and/or classifying the objects. An accuracy rating for each observed object is produced. When the accuracy rating is below a predefined threshold, that observed object is selected for manual analysis. The advantage of providing such a semi-automated method is that the object analysis of a multimedia file can be carried out in a relative short amount of time, whilst the accuracy of the detections is maintained at a relatively high standard due to the manual intervention in chosen sections of the multimedia file.

WO 2016/146847 A1



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

**Declarations under Rule 4.17:**

- *as to the identity of the inventor (Rule 4.17(i))*
- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

**Published:**

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

**“A method of analysing a multimedia file”**

**Introduction**

5

This invention relates to a method of analysing a multimedia file. In particular, the present invention is directed towards a method for analysing a multimedia file for purposes of observing objects appearing in the multimedia file by detecting, counting, tracking and/or classifying such objects.

10

Throughout this specification, the term “multimedia file” shall be understood to encompass any type of multimedia file which may be stored and downloaded for analysis or streamed from a remote storage for analysis or streamed live from a content source location for real-time analysis. The multimedia file may be a video file and such a video file can be interpreted as containing a plurality of images which together make up the video file. It may not, and is likely not to, have any audio component.

15

The multimedia file which is to be analysed is envisaged to be a video file showing a road junction, and the objects to be observed would be vehicles passing through the road junction. In such cases, the multimedia file will be created using content recorded from a traffic camera, closed-circuit television camera or similar type of camera, which records images of the junction to be analysed.

20

A further aspect of the invention envisages observing people as the objects. This could be used in analysing public spaces to determine the flow of people through a pedestrian junction, or monitor the flow of people through public spaces such as stadiums, shopping centres and the like.

25

The primary implementation of the present invention is envisaged to assist with object (e.g. vehicles) movement observation, specifically, object detection, object counting, object tracking and object classification from video footage.

30

The detection, counting and tracking of vehicles as they pass through a junction can be of great assistance for traffic management purposes. The present invention

35

- 2 -

is able to count where the vehicles travel from, and travel on to as they pass through a specific junction, during a time period, for example 12 hours. A further aspect of the invention is to classify the vehicles as this is useful from a statistical point of view for planning road extensions and road capacity requirements for heavy loads.

It will be readily understood that the term "junction" refers to any roadway, or route which has a plurality of arms, along each of which arms a vehicle may travel and endpoints of the arms meet one another to form the junction. A 3-arm junction is typically known as a T-junction; a 4-arm junction may be a crossroads or a 4-arm roundabout.

A vehicle passing through the junction will enter along one arm of the junction and will exit along an arm. It is a main object of the present invention to detect and count the number of vehicles passing through a selected junction and to also track the movement of the vehicles through the junction, during a time period, by analysing the entry and exit arms for each of the vehicles passing through the junction. The classification of the vehicle is also determined. The class of a vehicle may be one of: a car, a bus, a taxi, a truck, a van, a lorry and so on.

Currently in most cases the detecting, counting and classifying of vehicles is carried out entirely manually. In this manner, a traffic survey company will install a camera at a specific junction and a person will be employed to watch the video recorded from that camera. The person will count, track and classify the vehicles. The results are normally written by the person on paper sheets. It is common that the person will have to view the recorded video several times in order to generate accurate results, particularly where the junction is a complicated junction that may have a large number of vehicles passing through it. The person usually has to rewind and watch forward a section of video several times if a number of cars are passing through the junction concurrently. This manual process is extremely time intensive and labour intensive. The efficiency of the process is relatively low and can be prone to human error also.

Automated detection, counting and tracking of vehicles has been considered in the prior art. The objects (e.g. vehicles) shown in the video file can be automatically

- 3 -

5 detected using a tracking algorithm. The tracking algorithm requires some configuration to allow the algorithm to analyse the video and produce the results. Such automated systems are described in U.S. Patent Number US 8,204,955 (MIOVISION TECHNOLOGIES INCORPORATED). US 8,204,955 discloses a method and system which are provided for remotely analysing multimedia content, in particular video.

10 U.S. Patent Number US 7,460,691 (CET TECHNOLOGIES PTE LIMITED) is a further example of the prior art and discloses image processing techniques which are applied video images for the purposes of acquiring traffic data. The disclosure discusses a traffic monitoring system, the basic function of which is for traffic data acquisition and incident detection and in particular the application of image processing techniques for the detection of a vehicle, from sequence of video images, as well as the acquisition of traffic data and detection of any traffic incidents.

15 However, there are a number of issues with fully automated systems such as those described in U.S. Patent Number US 8,204,955 and U.S. Patent Number US 7,460,691. Oftentimes, count errors occur due to weather conditions, poor camera views, high density traffic and so on. In such instances, the accuracy of the automated systems drops dramatically and the system is ineffectual as the count errors which are unavoidable due to the weather conditions, poor camera views, high density traffic, etc. are incorporated into the automated analysis. It is also known for automated systems to miss some shapes/sizes of objects (such as small mopeds for example), or, produce alternative errors such as false positives, duplicated detections and so on.

20 In short, using the fully automated tracking produces several errors which results in a lower accuracy than the manual counting, although the fully automated systems are time efficient. The manual tracking is better from an accuracy point of view but is both labour intensive and time intensive and therefore relatively costly in comparison to the automated process described in the prior art.

25 It is a goal of the present invention to provide a system and method that overcomes at least one of the above mentioned problems.

### **Summary of the Invention**

The present invention is directed to a method of analysing a multimedia file for the detection, counting, tracking and/or classification of objects displayed in images  
5 contained within the multimedia file, wherein the method comprises the steps of executing an automated analysis of the multimedia file, wherein, the automated analysis comprises observing the objects by detecting, counting, tracking and/or classifying the objects in the multimedia file, and producing an accuracy rating for at least each observation of an object; marking chosen sections of the multimedia file,  
10 where observations of the objects in those chosen sections have an accuracy rating below a predefined threshold; and, manually observing objects in the chosen sections of the multimedia file.

The advantage of providing such a semi-automated method is that the object analysis  
15 of a multimedia file can be carried out in a relative short amount of time, whilst the accuracy of the detections is maintained at a relatively high standard due to the manual intervention in chosen sections of the multimedia file where the automated detection has been deemed to have an accuracy rating below an acceptable predefined threshold. The method also allows for correction of the automated  
20 detection on the fly in non-chosen sections of the multimedia file.

In a further embodiment, the accuracy rating is produced separately for the detection, counting, tracking and/or classification of the objects.

25 In a further embodiment, the accuracy rating is an accumulated accuracy rating which is produced based on any combination of the detection, counting, tracking and/or classification of the objects.

In a further embodiment, the accuracy rating is calculated on the probability of the  
30 detection, counting, tracking, and/or classifying being correct.

In a further embodiment, the probability of the detection, counting, tracking, and/or classifying of the object being correct is based on one or more of: to what extent the motion of the object complies with expected motions of the object, such as  
35 kinematic and dynamic capabilities of the object; whether the object has properly

- 5 -

approached, entered and exited an internal area, which internal area is specified by a user and is a portion of an image contained in the multimedia file; whether a direction of motion of the object complies with a preset movement path on the image of the multimedia file, which movement path has been defined by the user;

5 whether the object is distinct and separated from other objects by a predefined distance in the image of the multimedia file; and/or, whether a tracked path of the object is distinct from other tracked paths of other objects.

In a further embodiment, the method comprises a further step of producing an

10 object analysis report.

In a further embodiment, the method comprises a further step of manually correcting erroneous detections in non-chosen sections of the multimedia file.

15 In a further embodiment, the method comprises a further step of manually classifying the automatically detected objects and/or the manually detected objects.

In a further embodiment, the step of manually classifying the automatically detected objects and/or the manually detected objects includes displaying a thumbnail image

20 of the object to be classified to a user, and allowing the user to call up a short video of the object to be classified if the thumbnail image is deemed to be insufficient to classify the object.

In a further embodiment, the method comprises a further step of validating portions

25 of the automatically detected objects, the manually detected objects and the manually classified objects, by way of making a comparison with manually detected objects and manually classified objects respectively.

In a further embodiment, the multimedia file is a video file. In a further embodiment,

30 the multimedia file is a video file of a junction.

In a further embodiment, the objects are moving objects. In a further embodiment, the objects are vehicles.

35 In a further embodiment, a user manually detects, counts, tracks and/or classifies

- 6 -

objects using a game controller suitable for a games console.

In a further embodiment, background areas on images contained in the multimedia file may be selected by a user in advance of executing the automated analysis of the multimedia file such that those selected background areas in the images of the  
5 multimedia file can be excluded from the automated analysis.

In a further embodiment, internal areas on images contained in the multimedia file may be selected by a user in advance of executing the automated analysis of the multimedia file so that the selected internal areas in the images of the multimedia  
10 file can be used during the automated analysis to assist in calculating the accuracy rating and/or confirming a valid detection of an object.

In a further embodiment, prior to executing the automated analysis of the multimedia file, analysis parameters for analysing the multimedia file are input by a  
15 user.

In a further embodiment, the analysis parameters input by the user comprise one or more of: movement paths, camera settings, time settings, background and internal  
20 areas, and, tracker configuration.

The present invention is further directed towards a tracker system for analysing a multimedia file so as to detect, count, track and/or classify objects displayed in images contained within the multimedia file; the tracker system comprising means  
25 for executing an automated analysis of the multimedia file, wherein, the automated analysis comprising observing objects in the multimedia file by detecting, counting, tracking and/or classifying the objects in the multimedia file, and means for producing an accuracy rating for each observation of an object; and, means for marking chosen sections of the multimedia file, where observations of the objects in  
30 those chosen sections have an accuracy rating below a predefined threshold; and, means to manually detect objects in the chosen sections of the multimedia file.

In a further embodiment, the means for producing an accuracy rating for each observation of the object produces an accuracy rating separately for each of the  
35 detection, counting, tracking and/or classification of the objects.



- 7 -

In a further embodiment, the means for producing an accuracy rating for each observation of the object produces an accumulated accuracy rating based on a combination of the accuracy ratings for the detection, counting, tracking and/or  
5 classification of the objects.

In a further embodiment, the accuracy rating is calculated on the probability of the detection, counting, tracking, and/or classifying being correct.

10 In a further embodiment, the probability of the detection, counting, tracking, and/or classifying of the object being correct is based on one or more of: to what extent the motion of the object complies with expected motions of the object, such as kinematic and dynamic capabilities of the object; whether the object has properly approached, entered and exited an internal area, which internal area is specified by  
15 a user and is a portion of an image contained in the multimedia file; whether a direction of motion of the object complies with a preset movement path on the image of the multimedia file, which movement path has been defined by the user; whether the object is distinct and separated from other objects by a predefined distance in the image of the multimedia file; and/or, whether a tracked path of the  
20 object is distinct from other tracked paths of other objects.

In a further embodiment, the tracker system further comprises means to manually correct erroneous detections in non-chosen sections of the multimedia file.

25 In a further embodiment, the tracker system further comprises means to manually classify the automatically detected objects and/or the manually detected objects.

In a further embodiment, the means to manually classify the automatically detected objects and/or the manually detected objects comprises means to display a  
30 thumbnail image of the object to be classified by a user, and means to allow the user to call up a short video of the object to be classified if the thumbnail image is deemed to be insufficient to classify the object.

In a further embodiment, the tracker system further comprises means to validate  
35 portions of the automatically detected objects, the manually detected objects and

- 8 -

the manually classified objects being validated by way of a comparison with manually detected and classified objects.

In a further embodiment, the tracker system further comprises means to produce an  
5 object analysis report.

In a further embodiment, the multimedia file is a video file. In a further embodiment, the multimedia file is a video file of a junction.

10 In a further embodiment, the objects are moving objects. In a further embodiment, the objects are vehicles.

In a further embodiment, the user manually detects, counts, tracks and/or classifies objects using a game controller suitable for a games console.

15

In a further embodiment, the tracker system further comprises means for selecting background areas on images contained in the multimedia file in advance of executing the automated analysis of the multimedia file such that those selected background areas in the images of the multimedia file can be excluded from the  
20 automated analysis.

20

In a further embodiment, the tracker system further comprises means to select internal areas on images contained in the multimedia file in advance of executing the automated analysis of the multimedia file so that the selected internal areas in  
25 the images of the multimedia file can be used during the automated analysis to assist in calculating an accuracy rating and/or confirming a valid detection of an object.

25

In a further embodiment, the tracker system further comprises input means to allow  
30 a user input analysis parameters for analysing the multimedia file to prior to executing the automated analysis of the multimedia file.

30

In a further embodiment, the analysis parameters input by the user comprise one or more of: movement paths, camera settings, time settings, background and internal  
35 areas, and, tracker configuration.

35

The present invention is further directed to a method of analysing a multimedia file for the detection, counting, tracking and/or classification of objects displayed in images contained within the multimedia file, wherein the method comprises the steps of executing an automated analysis of the multimedia file, wherein, the automated analysis detects, counts, tracks and/or classifies objects in the multimedia file by at least detecting an object, and producing an accuracy rating for at least each detection of an object; marking chosen sections of the multimedia file, where detections of the objects in those chosen sections have an accuracy rating below a predefined threshold; manually detecting objects in the chosen sections of the multimedia file; manually correcting erroneous detections in non-chosen sections of the multimedia file; and producing an object analysis report.

The present invention is further directed to a method of analysing a multimedia file for the detection, counting, tracking and/or classification of objects displayed in images contained within the multimedia file, wherein the method comprises the steps of executing an automated analysis of the multimedia file, wherein, the automated analysis detects, counts, tracks and/or classifies objects in the multimedia file by at least detecting an object, and producing an accuracy rating for at least each detection of an object; marking chosen sections of the multimedia file, where detections of the objects in those chosen sections have an accuracy rating below a predefined threshold; manually detecting objects in the chosen sections of the multimedia file; manually correcting erroneous detections in non-chosen sections of the multimedia file; manually classifying the automatically detected objects and the manually detected objects; and producing an object analysis report.

The present invention is further directed to a method of analysing a multimedia file for the detection, counting, tracking and/or classification of objects displayed in images contained within the multimedia file, wherein the method comprises the steps of executing an automated analysis of the multimedia file, wherein, the automated analysis detects, counts, tracks and/or classifies objects in the multimedia by at least detecting an object, and producing an accuracy rating for at least each detection of an object; marking chosen sections of the multimedia file, where detections of the objects in those chosen sections have an accuracy rating below a predefined threshold; manually detecting objects in the chosen sections of

- 10 -

the multimedia file; manually correcting erroneous detections in non-chosen sections of the multimedia file; manually classifying the automatically detected objects and the manually detected objects; validating portions of the automatically detected objects, the manually detected objects and the manually classified objects by way of a comparison with manually detected and classified objects; and producing an object analysis report.

In a further embodiment, the multimedia file is a video file. In a further embodiment, the objects are moving objects. In a further embodiment, the objects are vehicles. In a further embodiment, the multimedia file is a video file of a junction.

In a further embodiment, the accuracy rating is calculated on the probability of the detection, counting, tracking, and/or classifying being correct.

In a further embodiment, a user manually detects, counts, tracks and/or classifies objects using a game controller. In a further embodiment, the game controller can be programmed.

In a further embodiment, the manual classification of the automatically detected objects and the manually detected objects includes displaying a thumbnail image of the object to be classified to a user, and further allows a user to call up a short video of the object to be classified if the thumbnail image is deemed to be insufficient to classify the object.

In a further embodiment, background areas on images from the multimedia file may be selected by a user in advance of executing the automated analysis of the multimedia file so that background areas in the images of the multimedia file can be excluded from the automated analysis.

In a further embodiment, internal areas on images from the multimedia file may be selected by a user in advance of executing the automated analysis of the multimedia file so that the internal areas in the images of the multimedia file can be used during the automated analysis to assist in calculating the accuracy rating and/or confirming a valid detection of an object.

- 11 -

In a further embodiment, prior to executing the automated analysis of the multimedia file, analysis parameters for analysing the multimedia file are input by a user.

- 5 In a further embodiment, the analysis parameters input by the user comprise one or more of: movement paths, camera settings, time settings, background and internal areas, and, tracker configuration.

10 The present invention is further directed towards a tracker system for analysing a multimedia file so as to detect, count, track and/or classify objects displayed in images contained within the multimedia file, characterised in that the tracker system comprises means for executing an automated analysis of the multimedia file, wherein, the automated analysis detects, counts, tracks and/or classifies objects in the multimedia file in the form of a detection of an object, and produces an accuracy  
15 rating for each detection of an object; and, means for marking chosen sections of the multimedia file, where detections of the objects in those chosen sections have an accuracy rating below a predefined threshold; means to manually detect objects in the chosen sections of the multimedia file.

20 The present invention is further directed towards a tracker system for analysing a multimedia file so as to detect, count, track and/or classify objects displayed in images contained within the multimedia file, characterised in that the tracker system comprises means for executing an automated analysis of the multimedia file, wherein, the automated analysis detects, counts, tracks and/or classifies objects in  
25 the multimedia file in the form of a detection of an object, and produces an accuracy rating for each detection of an object; means for marking chosen sections of the multimedia file, where detections of the objects in those chosen sections have an accuracy rating below a predefined threshold; means to manually detect objects in the chosen sections of the multimedia file; means to manually correct erroneous  
30 detections in non-chosen sections of the multimedia file; means to manually classify the automatically detected objects and the manually detected objects; means to validate portions of the automatically detected objects, the manually detected objects and the manually classified objects being validated by way of a comparison with manually detected and classified objects; and means to produce an object  
35 analysis report.

### Detailed Description of Embodiments

The invention will be more clearly understood from the following description of some  
5 embodiments thereof, given by way of example only, with reference to the  
accompanying drawings, in which:

10 Figure 1 is a flow diagram detailing some of the steps involved in the present  
invention;

Figure 2 is an image taken from a multimedia file, whereby, objects  
15 appearing in the image are to be analysed in accordance with the present  
invention;

Figure 3 is a screenshot of a user interface for manually counting and  
20 tracking objects using a tracker system in accordance with the present  
invention; and,

Figure 4 is a screenshot of a user interface for manually classifying objects  
25 using a tracker system in accordance with the present invention.

Referring to Figure 1, there is provided a multimedia file analysis method indicated  
generally by reference numeral 100. The multimedia file analysis method 100  
30 comprises a plurality of steps, 102 to 116, which enable the analysis of objects which  
appear in the multimedia file.

In step 102, a user will define analysis parameters for analysing a multimedia file.

In step 104, an automated observation of objects displayed in the multimedia file is  
35 carried out, based on defined analysis parameters. Figure 1 refers to an automated  
count as opposed to an automated observation; however, it will be understood that  
throughout this specification, the term 'observation of objects', and its grammatical  
equivalents, shall be interpreted as detecting objects displayed in one or more images  
which make up a part of the multimedia file, counting the objects displayed in the one  
or more images which make up a part of the multimedia file, tracking the objects

- 13 -

displayed in the one or more images which make up a part of the multimedia file, and/or, classifying the objects displayed in the one or more images which make up a part of the multimedia file. The automated observation of objects may be carried out over the full length of the multimedia file, which will typically be a video file; or, the  
5 automated observation may be carried out during a predefined selected portion of the multimedia file.

In step 106, automated count results from step 104 are transferred for correction and verification by a user.

10

In step 108 the user will manually count objects, which shall be understood to be manually observing objects, for chosen sections of the multimedia file. The chosen sections will be sections of the multimedia file where the automated observation has a relatively low accuracy rating as the probability of the observations regarding the  
15 observed object being correct have been adjudged to be low. The method of the present invention then specifies for these sections to be manually counted.

In step 110, the user will manually correct the automated count in non-chosen sections of the multimedia file, if the user notices that there are any errors in these  
20 non-chosen sections when having a brief look through the multimedia file. This step is of course not mandatory and may be omitted in some embodiments of the invention.

In step 112, the user manually classifies objects in the multimedia file according to preset criteria, or, objects are automatically classified according to the preset criteria. It  
25 will be appreciated that it is foreseen to automatically classify at least some of the objects detected in the multimedia file based on assessing the size and/or shape of the object detected in the multimedia file. As will be discussed in greater detail hereinbelow, the angle of view which the camera has relative to the objects passing through the field of vision of the camera will be taken into account when determining  
30 the criteria for assessing the shape and size of objects appearing in the images of the multimedia file.

In step 114 of the method, the automated and manual analysis results are verified by have a user check randomly chosen portions of the multimedia file.

35

- 14 -

In the final step 116, an analysis report is created.

It will be understood that steps such as the creation of a report, or the verification of data, or the classification of objects may be omitted.

5

For the purposes of this description, the implementation of the method will focus on the observation of objects, thus incorporating the detection, counting, tracking and classification of objects appearing in a multimedia file, such as detecting, counting, tracking and classifying vehicles which appear on a video file of a junction with the vehicles passing through the junction. Alternative implementations for this methodology are envisaged for observing the movement of people, animals and other types of transportation at junctions or open public spaces.

It will be further understood that one or more users may be used to complete the steps given hereinbefore; the invention is thus not limited to the same user completing all of the steps. Indeed, in one embodiment of the present invention, it is envisaged that the system will allow manually implemented steps to be carried out by users derived from a crowd sourced community of workers. In this embodiment, an individual or group of individuals would log into the system, and possibly complete a training exercise, prior to then carrying out the manual detection, counting, tracking, classifying and/or verifying of objects in an assigned multimedia file. The system would assign work to the logged in users and would track the work carried out by each user to ensure it is of a sufficient quality rating before authorising a payment for the user.

25 Now turning to consider each of the steps in more detail.

With reference to Figure 2, there is shown an image indicated generally by reference numeral 200, from a multimedia file (not shown) in which a number of vehicles 202, 204 are seen to be passing through a three-armed junction indicated generally by reference numeral 206. In this case the three-armed junction 206 is a round-about. A background area (discussed hereinbelow) 208 has been defined by a user. An internal area (discussed further below) 210 has also been defined by the user. And, a movement path (discussed hereinunder) 212, which is indicated by a plurality of arrows, has been defined by the user.

35



- 15 -

As per step 102 of Figure 1, a user will define the analysis parameters for a tracker system which will be used for the multimedia file analysis of the image of Figure 2. The analysis parameters are required by the tracker system, in addition to the multimedia file, to be able to detect the objects, in this case vehicles, in the multimedia file.

5

The analysis parameters comprise, but are not limited to, one or more of the following:

- 10
- 15
- 20
- 25
- 30
- 35
- a) Movement routes; each possible movement through the junction is represented as a matrix of movement directions. The matrix of movement directions is shown in a top left corner of Figure 2, where each direction of each arm is accorded a separate reference letter. A vehicle entering the junction from the bottom of the Figure 2 and turning left, from a driver's perspective, will travel through cell C and then pass through cell A, and exit the junction through cell D. If the vehicle entering the junction from the bottom of the Figure 2 turns right off of the round-about, again from a driver's perspective, the vehicle will travel through cell C and then pass through cell A, and exit the junction through cell F. Cell A is the junction/round-about itself.
  - b) Camera settings; a plurality of settings which indicate the position and/or orientation of the camera are used to allow the method to determine the view contained in the multimedia file so that a plane may be established. For example, one or more of: a height of installation of the camera from ground level (e.g. 5m), a field of view of the camera (e.g. 65°), a horizon height which refers to the location of the horizon relative to a chosen border of the camera view (e.g. 23% below the top border), a cell size for the movement matrix (e.g. 1m), and a camera angle which is the angle formed between the direction in which camera faces and the horizontal (e.g. 30° downward) are provided by a user in order to allow the view of the junction to be contextualised by the method, for example, to understand how far into the distance the camera view can see. This can assist with detecting and classifying objects as the size and shape of the objects will vary according to the angle/plane of view of the camera, and the distance towards the horizon which the camera view can see, and hence provides in the

multimedia file.

- 5 c) Time settings; frames per second in real time are made known to the method to allow the method to be able to calculate the real velocity (m/s) of the objects appearing in the multimedia file, which will be the velocity of the vehicles passing through the junction.
- 10 d) Background and internal areas; the user can select areas of the image from the multimedia file which can be ignored from the point of view of the analysis. For example, flags or trees fluttering due to wind, or, a pedestrian walking on the footpath may be excluded, if desired by using a background area. The flags/trees could be confused for objects that are to be tracked, and, it may not always be desired to detect and analyse pedestrians in some circumstances. Although it will be understood, that in other embodiments, the roadways and carriageways may be excluded so that the method can focus on analysing pedestrians on the footpaths, if desired. The internal areas can be used to help trigger an observation of an object by setting the parameters such that an object will only be counted, tracked and or
- 15 or classified if it at least enters the internal area, or enters and subsequently leaves the internal area.
- 20 e) Tracker configuration; the method employed by the tracker system can be adjusted to take account of several pre-defined conditions. For example, different video formats will require different configurations due to the difference image contrasts that exist between the different configurations. The tracker system may also be used during different lighting conditions, such as low light at dusk or dawn, or at night time.
- 25
- 30 The tracker system provides the user with a user interface to allow the user to easily input and set these parameters. The background areas, the internal areas and the movement paths can be defined on a drawing area which can be overlaid on an image from the multimedia file. The movement paths can be drawn using a stroke of a brush-type software tool which can define a possible travel path. In this manner, each of the
- 35 possible travel paths can be input quickly. The software tool may also take account of

- 17 -

the angle/plane of the camera when accepting the path drawing stroke so that the depth perspective of the camera view is taken into account by automatically narrowing the path appropriately when the drawing stroke is made towards the horizon. In further embodiments, the background areas and the internal areas may be automatically defined by the tracker system.

As can be seen from Figure 2, the internal area 210 is needed for the tracker system to be able to distinguish between valid vehicles and invalid vehicles. A valid vehicle is defined as an object/vehicle which is initially detected outside of the internal area 210, is detected inside the internal area 210 and hence is considered to pass through the internal area 210 through the internal area 210, and subsequently is detected outside of the internal area 210 again.

With reference now to step 104, the automated count of objects in the multimedia file is executed based on defined analysis parameters from step 102. Firstly, the tracker system exports the analysis parameters set by the user defined in the previous section for the tracker and then executes the tracker system to begin scrolling through the multimedia file and counting the objects shown in the multimedia file. During the tracking the tracker system provides the user with progress information and estimates the total time of execution. This is important as the tracking execution time for a longer video (12 hr) can last many hours. The user can also cancel/restart the tracking at any time.

The tracker runs through the entire multimedia file and runs the object/vehicle detection algorithm to count the objects/vehicle. An object or vehicle is detected by comparing pixels of adjacent frames to detect the presence of an object, and its motion relative to the background pixels which will not substantively change. After the detection is complete the tracker system exports the tracking results. The results includes one or more of: the frame number(s) in the multimedia file (e.g. video file) where the object entered and exited the junction; the detected object movement (e.g. the number of the movement based on the cells of the movement matrix); an estimated accuracy; the detected object class, if the classification is being done automatically based on the size and/or shape of a cluster of pixels which are determined to represent an object; and/or, a screenshot of each of the objects for manual classification, if manual classification is being done.

- 18 -

The estimated accuracy is the probability that the detected object is actually an object.

The probability is determined by assessing:

- 5       - to what extent the motion of the object complies with expected movements of the object. The expected movements could be based on the kinematic and dynamic capabilities of the object. Taking the example of a vehicle, the vehicle will have a maximal acceleration level, a constraint against motion sideways without turning, and so on;
- 10       - whether the object has properly approached and entered the internal area 210, exited the internal area 210, and did not re-enter the internal area 210 again;
- whether a direction of motion of the object complies with the movement path 212 defined by the user;
- whether the object is distinct from other objects (so as to avoid, for example, counting a front cab and a rear trailer of a truck as two separate vehicles);
- 15       - whether a tracked path of the object is distinct from other tracked paths of other objects, so as to determine whether there are multiple objects which are occluding each other for a period of time;
- whether the object has a size and/or shape which is within an expected range of sizes and/or set of expected shape types. It could be that a saloon car, an estate car, a SUV, a truck, a van, a motorbike, a pushbike, and so on all have different expected sizes and shapes and the detected size and shape would be compared to the expected sizes and shapes to determine if one of them substantially matches. As the plane of view of the camera will be known, the position of the object on the image can be used to determine the distance of the object from the camera as this will clearly have a bearing on the expected size and shape of the object;
- 20       - whether the object is observed to move in a manner consistent with the known plane of view of the camera. Such that, the object size should reduce if the object moves towards a horizon and vice versa; and/or,
- 25       - whether the object enters and leaves the cells of the movement matrix in an expected direction(s).
- 30       - whether the object enters and leaves the cells of the movement matrix in an expected direction(s).

The estimated accuracy is an important aspect of the invention and as many information sources (such as information provided by the user, information retrieved 35 by the tracking system, and, information extracted from the images) are used to

- 19 -

determine the estimated accuracy of the observation of the object in successive images from the camera.

As mentioned above, the tracker system estimates the accuracy of the observation of an object by estimating the probability that the object has been detected, counted, tracked and/or classified correctly. The tracker will mark chosen sections of the multimedia file, in the form of time ranges during the multimedia file, where the tracker was not able to observe the object with a high accuracy rating, as compared with an acceptable preset threshold. These are the sections of the multimedia file where there are either objects with low accuracy ratings (doubtful cases), or objects which do not seem to be vehicles but nonetheless have a higher accuracy than conventional noise (possible noise cases). These possible noise cases are where the tracker system definitely sees motion but it is not clearly an expected object, such as a vehicle. For example, traffic jams cause issues as there will be many objects moving slowly and stopping behind each other; the tracker system thus detects a single, large object moving in complex ways and the tracker system will be unable to identify the motion of individual objects. However, the tracker system will understand that there are vehicle-like objects moving and the system needs to ask for the user's manual assistance. These chosen sections of the multimedia file will then be processed manually in accordance with step 108.

In a preferred embodiment, the chosen section of the multimedia file will be grouped in accordance with the arms, or cells, of the junction under review. In this manner, it is easier for a manual count of the chosen sections of the multimedia file to be done for several objects all emanating from the same arm on the junction.

The tracker also determines a number in the form of a percentage value, which is referred to as a certainty ratio, for each arm of the junction. This certainty ratio indicates how much of the multimedia file requires manual processing, which is an indicator of the quality of the automated analysis. This ratio is calculated based on the number of chosen sections of the multimedia file, the duration of the chosen sections and the distribution of the chosen sections. If the chosen sections are relatively long and are distributed relatively far from each other, then the user of the tracker system will not have any issues to operate the system as the user can simply use the time ratio of the chosen intervals, as chosen, to manually process the chosen sections.

- 20 -

However, if the chosen intervals are relative short and are relatively close to one another, then the user cannot easily manually process two seconds of video, then skip one second of video, then manually process three seconds of video, and so on. So before reporting the final chosen sections and calculating the certainty ratio, the tracker system merges chosen sections which are close to each other, even though there will be a short certain section between the two chosen sections. The user can set a parameter, such as five seconds, which determines if chosen sections are too close to one another and need to be merged to make the manual processing easier for the user. In this manner, if the distribution of the chosen intervals is many relative short sections that are relative close to each other, the tracker system will merge these chosen sections (and any short certain sections falling therebetween) into a single longer, continuous section for manual processing. It should be noted that if an object or vehicle is correctly observed by the tracker system in a short certain section which is subsequently merged with chosen sections for manual processing, then the count, classification, detection and so on for that vehicle is removed so as to avoid double counting.

All of the results of the automated analysis are transfer, following step 106, for correction and verification by a user.

In step 108, a user will manually count objects for the chosen sections of video, which had a relatively low accuracy estimated value. In this manner, any objects which the tracker system was uncertain about are manual counted by a user. Moreover, the user can correct non-chosen sections of the multimedia file and correct any errors that may be present, even though the tracker system would have evaluated these non-chosen sections to have a sufficiently high accuracy rating. As the multimedia is played back to the user, during these non-chosen sections a user can still override the automated result and correct errors in this way.

After the automated analysis has been executed, the manual analysis may be carried out. The manual analysis consists of the following tasks:

- i) Manually observe, including one or more of: detect, count, track and/or classify objects in the chosen section(s) of the multimedia file which has been analysed automatically already;
- ii) Correct any errors that may be present in the non-chosen sections of the

- 21 -

multimedia file; and,

- iii) Manually classify both automatically detected vehicles and manually detected vehicles.

5 The first two tasks can be completed on the same screen, which is shown in Figure 3. The third task is completed on a separate screen as shown in Figure 4.

There are different of errors which can occur during the automated analysis of a multimedia file. The types of known errors comprise:

- 10 a) Duplication: this is where a single object is detected as two objects. This is typical where the objects may be a vehicle with a trailer. In this situation the user needs to delete one of the detections.
- b) Incorrectly detected origin arm or incorrectly detected destination arm: in such cases the user needs to delete the detection and add the vehicle manually with the correct origin arm and the correct destination arm.
- 15 c) The detected and tracked paths of two different objects become either joined or switched. For example, if two vehicles are moving in the opposite direction to one another, and they meet at the center of the junction, the tracker system may interpret this as being two U-turns. In such cases, the user needs to manually delete both detections and add the two vehicles correctly.
- 20 d) Detect noise as a valid vehicle. For example, pedestrians, cyclists, trees, flags and other items may cause false detections which must then be manually deleted by the user.

25

Referring now to Figure 3, the correction user interface indicated generally by reference numeral 300, which allows a user to manually analyse objects in the chosen section(s) of the multimedia file which has been analysed automatically already, and, correct any errors that may be present in the non-chosen sections of the multimedia file is shown. The correction interface 300 comprises a main panel 302 which contains the video of the junction with an overlay of detections and the thumbnail of the junction in the upper left corner of the main panel 302, which indicates the cells of the matrix which represent the arms of the junction, which in the given example are represented by reference letters A to G. A progress bar indicated generally by reference numeral 304 is provided below the main panel 302. The progress bar 304 indicated the overall

30

35

- 22 -

length of playback of the multimedia file, and where the user currently is within that playback period. Chosen and non-chosen sections of the multimedia file can be differentiated along the progress bar 304 using different colours or the like. A classification table 306 is provided to one side of the main panel 302. The classification table 306 lists the different type of classifications which may be accorded to the detected objects and may also provide a running count of each classification type. An origin arm and destination arm indicator 308 is provided on the correction user interface 300 also. As described hereinbefore, playback will be carried out for one origin arm at a time, so for the most part the origin arm will be locked. If the user changes the origin arm on the origin arm and destination arm indicator 308, then the progress bar 304 and the main panel 302 will be updated to reflect the relevant information, such as the chosen and non-chosen sections for the newly selected origin arm. Information on the frame number of the multimedia file, and a recording time, may be provided as indicated by reference numeral 310.

Within the main panel 302, each detected object 312, which in this embodiment will be a vehicle, is displayed on the main panel 302 with an overlay of a bounding box 314, a detected and tracked path 316 for that object and additional information such as the origin arm/destination arm pair reference letters 318 and the object ID number for that object 320 which is useful for identification purposes. The bounding box 314 and the detected and tracked path 316 of the object 312 is visible and may be colour coded to differentiate numerous objects from one another when there are several objects being shown together on the same frame. In this manner, the user is able to quickly understand and recognise which objects have been detected and what is the detected and tracked path for that object is. Any errors are easily decipherable and the user can take manual correction action as required. Where the object has not been detected at all, such as during the chosen sections of the multimedia file, the user can add the object and classify the object, at the same time, using just one input device. During the chosen sections of the multimedia file, the user can be alerted by way of a red border, flashing border or some other audio, visual or tactile alarm so as to encourage the user to pay particular attention during those periods.

Therefore, as can be ascertained from the above, the user will carry out the following steps of playing the multimedia file for a selected origin arm. During playback for the selected origin arm, the user will manually count and classify the vehicles not have



- 23 -

been detected, and, correct the tracker system errors. The user repeats this process for all of the origin arms at the junction.

5 The manual analysis of an object requires the user to select the destination arm of the vehicle, bearing in mind that the origin arm will be fixed, and, select the class of the object. These selections can be made using any known input devices such as a mouse, by clicking on the appropriate destination arm button and the number next to the selected class for the object; using a keyboard in a similar fashion by way of shortcuts with predefined keys, or with a game controller which is used with a games  
10 console. The game controller is seen as a preferred input device as the inputs can be made in a very quick and ergonomic manner. For example, the destination arm can be selected using a D-pad or control stick, by pushing the D-pad or control stick in one of the four available directions with can be associated with the different arms. Up to eight different arms could be accommodated using such a D-pad or control stick using the  
15 horizontal, vertical and diagonal movements of the D-pad or control stick.

The class of the object can also be specified with the game controller by using the buttons on the game controller, whereby the different buttons are each associated with one of the different classes of object. As the correction of errors, may require a  
20 deletion of a detected object, the deletion can be executed with mouse by right clicking for example, or by using a keyboard with a key such as the 'delete' key, or by using a game controller by selecting the object to be deleted with a dedicated controller button. The selected object is then highlighted and additional information appears on the vehicle overlay. If there are multiple detections on the current frame, the detection  
25 which has the highest probability that it needs to be deleted (by having the lowest accuracy rating, for example) is selected first although a user can scroll to other detected objects within that frame. The user can accomplish this by scrolling to the next detected object by clicking the same dedicated controller button again. The dedicated controller button could be a shoulder button or rear touch pad as is found on  
30 many game controllers now.

With reference now to step 112 in Figure 1 and to Figure 4, there is provided in Figure 4 an object classification user interface indicated generally by reference numeral 400, which allows a user to manually classify objects which have been either manually  
35 detected by a user or automatically detected by the tracker system in the multimedia

- 24 -

file. The classification interface 400 displays a plurality of objects to be classified 402. Each object to be classified 402 comprises a completed field box 404 and an automatic classification indicator 406. If the automatic classification is accurate enough, it may not be necessary for manual classification, although at present it is currently envisaged to manually classify each detected object. The plurality of objects to be classified 402 are exported by the tracker system as thumbnail images. These thumbnail images are presented in a grid on the classification user interface 400. A classification table 410 is provided on the classification user interface 400. The classification table 410 lists the different type of classifications which may be accorded to the detected objects and may also provide a running count of each classification type. Progress data 408 may also be provided on the classification user interface 400.

In use, the user will classify each detected object separately. The selected object can be classified using an input device such as a mouse, a keyboard or a game controller. With a mouse, the user would click on the required class, and the selected object would become the next object to be classified and the user would move through the entire grid of objects to be classified in this manner. Alternatively a keyboard may be used, with predesignated keys relating to each of the classification types. With these shortcuts the classification can be very effective. There can be some situations where it is not easy to recognise the class of the object from the thumbnail image alone. For these situations the classification user interface 400 provides a keyboard shortcut for the user to open a video playback around the frame of the thumbnail so the object to be classified can be seen in a short moving video as this may assist in the classification. The object to be classified will be marked on the video playback using a bounding box overlay in a preferred embodiment.

Once the multimedia file has been thoroughly analysed and all of the objects have been detected, counted, tracked and/or classified, a final verification step is carried out in accordance with step 114 of Figure 1. This is to ensure that the analysis results comply with predefined accuracy levels. The validation algorithm marks a plurality of relatively short time periods randomly selected from the multimedia file and preferably distributed over the length of the multimedia file. A user then manually counts, tracks and classifies the objects in these short time periods and sends the results to the tracker system. The validation algorithm compares the original analysis results with the results of the validation analysis. Using this comparison, the validation algorithm

- 25 -

can estimate the accuracy of the original analysis results. The acceptable levels of accuracy of the original analysis results vary from market to market, but may be approximately 5% in the United Kingdom and Ireland for the enumeration of vehicles, and, approximately 15% for the classification of vehicles. There is no published standard and different levels of accuracy of the original analysis will be accepted in different jurisdictions. An analysis report is then created as detailed in step 116 of Figure 1 and this report can be used to ascertain traffic flow, or object movement in a videoed space/area.

10 A skilled person in the art will understand that the automated detection, counting, tracking and classification of objects in the multimedia file may be accomplished in a number of well-known ways. For example, an algorithm may detect an object by comparing the pixels of one frame to those in an adjacent frame and thus detect movement of objects through a sequence of frames. The shape and size of the objects, along with their trajectory and approximate velocity can be determined when taking into account the parameters input by a user with regard to the camera location and angle and so on.

The terms "comprise" and "include", and any variations thereof required for grammatical reasons, are to be considered as interchangeable and accorded the widest possible interpretation.

Throughout the preceding specification, the term "analysis" shall be understood to refer to the detection and/or tracking and/or counting and/or classification of objects in a multimedia file, and has been used interchangeably as such with these latter terms. The objects shall be understood to be moving objects, such as vehicles as being the primary example and embodiment described hereinbefore.

It will be understood that throughout the preceding specification, the term "vehicle" shall be afforded the broadest possible scope and shall refer to motorised vehicles such as cars, trucks, vans, lorries, motorbikes, mopeds, scooters, self-balancing wheeled hoverboards and so on, and, also to non-motorised vehicles such as pushbikes, wheelchairs, scooters, skateboards, and so on. As discussed hereinbefore, the invention can also be applied to people and can be adapted to accommodate people pushing trolleys, buggies, pushchairs and so on.

- 26 -

It will be understood that the components shown in any of the drawings are not necessarily drawn to scale, and, like parts shown in several drawings are designated the same reference numerals.

5

The invention is not limited to the embodiments hereinbefore described which may be varied in both construction and detail.

**CLAIMS**

1. A method of analysing a multimedia file for the detection, counting, tracking and/or classification of objects displayed in images contained within the multimedia file, wherein the method comprises the steps of executing an automated analysis of the multimedia file, wherein, the automated analysis comprises observing the objects by detecting, counting, tracking and/or classifying the objects in the multimedia file, and producing an accuracy rating for at least each observation of an object; marking chosen sections of the multimedia file, where observations of the objects in those chosen sections have an accuracy rating below a predefined threshold; and, manually observing objects in the chosen sections of the multimedia file.
2. A method of analysing a multimedia file as claimed in claim 1, wherein, the accuracy rating is produced separately for the detection, counting, tracking and/or classification of the objects.
3. A method of analysing a multimedia file as claimed in claim 1, wherein, the accuracy rating is an accumulated accuracy rating which is produced based on any combination of the detection, counting, tracking and/or classification of the objects.
4. A method of analysing a multimedia file as claimed in any preceding claims, wherein, the accuracy rating is calculated on the probability of the detection, counting, tracking, and/or classifying being correct.
5. A method of analysing a multimedia file as claimed in claim 4, wherein, the probability of the detection, counting, tracking, and/or classifying of the object being correct is based on one or more of:
- to what extent the motion of the object complies with expected motions of the object, such as kinematic and dynamic capabilities of the object;
  - whether the object has properly approached, entered and exited an internal area, which internal area is specified by a user and is a portion of an image contained in the multimedia

- 28 -

file;

- whether a direction of motion of the object complies with a preset movement path on the image of the multimedia file, which movement path has been defined by the user;
- 5       - whether the object is distinct and separated from other objects by a predefined distance in the image of the multimedia file; and/or,
- whether a tracked path of the object is distinct from other tracked paths of other objects.

10

6.       A method of analysing a multimedia file as claimed in any of the preceding claims, wherein, the method comprises a further step of producing an object analysis report.

15

7.       A method of analysing a multimedia file as claimed in any of the preceding claims, wherein, the method comprises a further step of manually correcting erroneous detections in non-chosen sections of the multimedia file.

20

8.       A method of analysing a multimedia file as claimed in any of the preceding claims, wherein, the method comprises a further step of manually classifying the automatically detected objects and/or the manually detected objects.

25

9.       A method of analysing a multimedia file as claimed in any of claim 8, wherein, the step of manually classifying the automatically detected objects and/or the manually detected objects includes displaying a thumbnail image of the object to be classified to a user, and allowing the user to call up a short video of the object to be classified if the thumbnail image is deemed to be insufficient to classify the object.

30

10.      A method of analysing a multimedia file as claimed in any of the preceding claims, wherein, the method comprises a further step of validating portions of the automatically detected objects, the manually detected objects and the manually classified objects, by way of making a comparison with manually detected objects and manually classified objects respectively.

35

- 29 -

11. A method of analysing a multimedia file as claimed in any preceding claims, wherein, the multimedia file is a video file.
- 5 12. A method of analysing a multimedia file as claimed in any preceding claims, wherein, the multimedia file is a video file of a junction.
13. A method of analysing a multimedia file as claimed in any preceding claims, wherein, the objects are moving objects.
- 10 14. A method of analysing a multimedia file as claimed in any preceding claims, wherein, the objects are vehicles.
- 15 15. A method of analysing a multimedia file as claimed in any preceding claims, wherein, a user manually detects, counts, tracks and/or classifies objects using a game controller suitable for a games console.
- 20 16. A method of analysing a multimedia file as claimed in any preceding claims, wherein, background areas on images contained in the multimedia file may be selected by a user in advance of executing the automated analysis of the multimedia file such that those selected background areas in the images of the multimedia file can be excluded from the automated analysis.
- 25 17. A method of analysing a multimedia file as claimed in any preceding claims, wherein, internal areas on images contained in the multimedia file may be selected by a user in advance of executing the automated analysis of the multimedia file so that the selected internal areas in the images of the multimedia file can be used during the automated analysis to assist in calculating the accuracy rating and/or confirming a valid detection of an object.
- 30 18. A method of analysing a multimedia file as claimed in any preceding claims, wherein, prior to executing the automated analysis of the multimedia file, analysis parameters for analysing the multimedia file are input by a user.
- 35 19. A method of analysing a multimedia file as claimed in claim 18, wherein, the

- 30 -

analysis parameters input by the user comprise one or more of: movement paths, camera settings, time settings, background and internal areas, and, tracker configuration.

- 5     20.     A tracker system for analysing a multimedia file so as to detect, count, track and/or classify objects displayed in images contained within the multimedia file; the tracker system comprising means for executing an automated analysis of the multimedia file, wherein, the automated analysis comprising observing objects in the multimedia file by detecting, counting, tracking and/or classifying the objects in the multimedia file, and means for producing an accuracy rating for each observation of an object; and, means for marking chosen sections of the multimedia file, where observations of the objects in those chosen sections have an accuracy rating below a predefined threshold; and, means to manually detect objects in the chosen sections of the multimedia file.
- 10
- 15
21.     A tracker system for analysing a multimedia file as claimed in claim 20, wherein, the means for producing an accuracy rating for each observation of the object produces an accuracy rating separately for each of the detection, counting, tracking and/or classification of the objects.
- 20
22.     A tracker system for analysing a multimedia file as claimed in claim 20, wherein, the means for producing an accuracy rating for each observation of the object produces an accumulated accuracy rating based on a combination of the accuracy ratings for the detection, counting, tracking and/or classification of the objects.
- 25
23.     A tracker system for analysing a multimedia file as claimed in any of claims 20 to 22, wherein, the accuracy rating is calculated on the probability of the detection, counting, tracking, and/or classifying being correct.
- 30
24.     A tracker system for analysing a multimedia file as claimed in claim 23, wherein, the probability of the detection, counting, tracking, and/or classifying of the object being correct is based on one or more of:
- 35             - to what extent the motion of the object complies with expected



- 31 -

- 5 motions of the object, such as kinematic and dynamic capabilities of the object;
- whether the object has properly approached, entered and exited an internal area, which internal area is specified by a user and is a portion of an image contained in the multimedia file;
  - whether a direction of motion of the object complies with a preset movement path on the image of the multimedia file, which movement path has been defined by the user;
  - 10 - whether the object is distinct and separated from other objects by a predefined distance in the image of the multimedia file; and/or,
  - whether a tracked path of the object is distinct from other tracked paths of other objects.
- 15
25. A tracker system for analysing a multimedia file as claimed in any of claims 20 to 24, wherein, the tracker system further comprises means to manually correct erroneous detections in non-chosen sections of the multimedia file.
- 20 26. A tracker system for analysing a multimedia file as claimed in any of claims 20 to 25, wherein, the tracker system further comprises means to manually classify the automatically detected objects and/or the manually detected objects.
- 25 27. A tracker system for analysing a multimedia file as claimed in claim 26, wherein, the means to manually classify the automatically detected objects and/or the manually detected objects comprises means to display a thumbnail image of the object to be classified by a user, and means to allow the user to call up a short video of the object to be classified if the thumbnail
- 30 image is deemed to be insufficient to classify the object.
28. A tracker system for analysing a multimedia file as claimed in any of claims 20 to 27, wherein, the tracker system further comprises means to validate portions of the automatically detected objects, the manually detected objects
- 35 and the manually classified objects being validated by way of a comparison

- 32 -

with manually detected and classified objects.

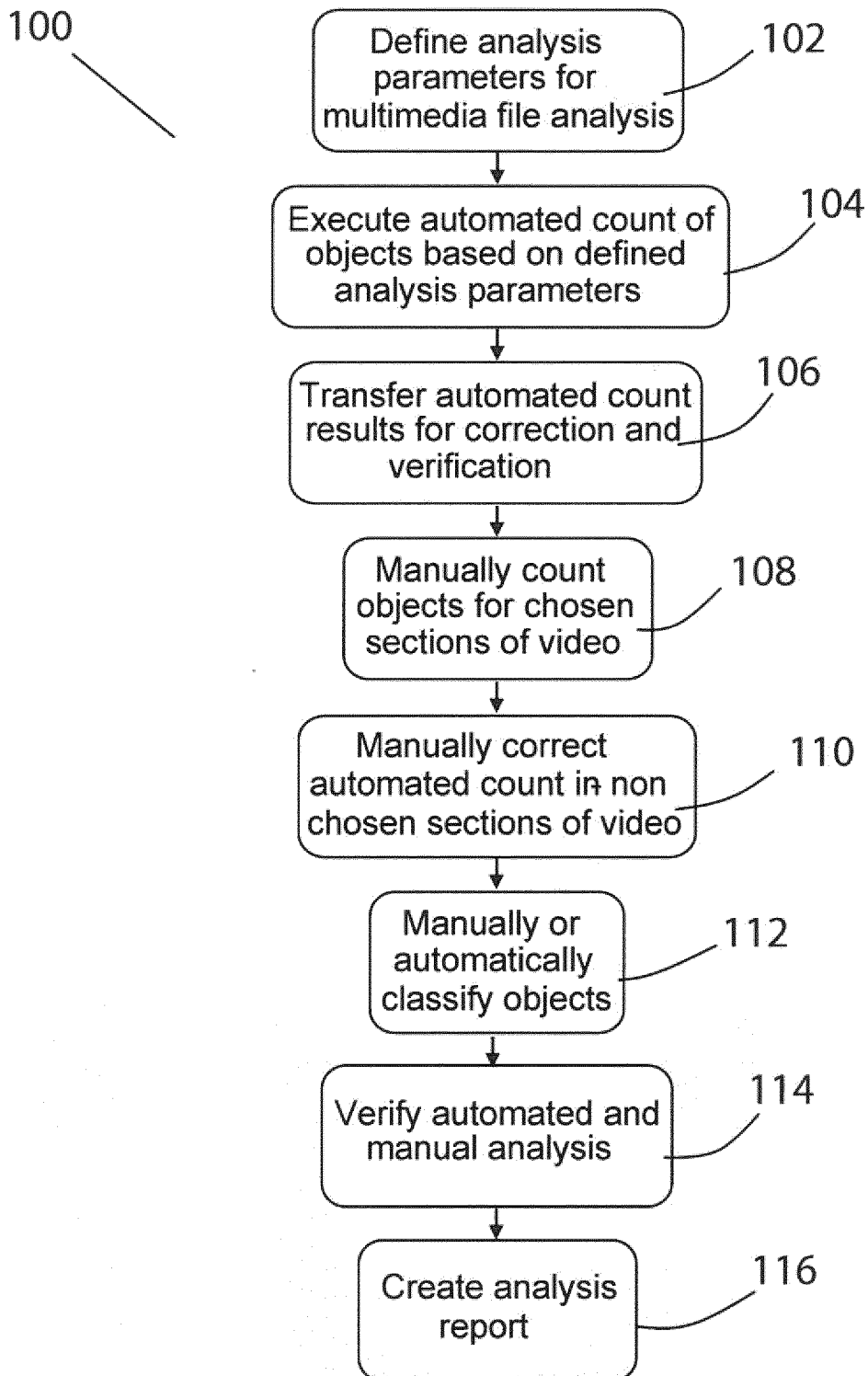
- 5
29. A tracker system for analysing a multimedia file as claimed in any of claims 20 to 23, wherein, the tracker system further comprises means to produce an object analysis report.
30. A tracker system for analysing a multimedia file as claimed in any of claims 20 to 29, wherein, the multimedia file is a video file.
- 10 31. A tracker system for analysing a multimedia file as claimed in any of claims 20 to 30, wherein, the multimedia file is a video file of a junction.
32. A tracker system for analysing a multimedia file as claimed in any of claims 20 to 31, wherein, the objects are moving objects.
- 15 33. A tracker system for analysing a multimedia file as claimed in any of claims 20 to 32, wherein, the objects are vehicles.
34. A tracker system for analysing a multimedia file as claimed in any of claims 20 to 33, wherein, the user manually detects, counts, tracks and/or classifies objects using a game controller suitable for a games console.
- 20 35. A tracker system for analysing a multimedia file as claimed in any of claims 20 to 34, wherein, the tracker system further comprises means for selecting background areas on images contained in the multimedia file in advance of executing the automated analysis of the multimedia file such that those selected background areas in the images of the multimedia file can be excluded from the automated analysis.
- 25 36. A tracker system for analysing a multimedia file as claimed in any of claims 20 to 35, wherein, the tracker system further comprises means to select internal areas on images contained in the multimedia file in advance of executing the automated analysis of the multimedia file so that the selected internal areas in the images of the multimedia file can be used during the automated analysis to assist in calculating an accuracy rating and/or
- 35

- 33 -

confirming a valid detection of an object.

37. A tracker system for analysing a multimedia file as claimed in any of claims  
20 to 36, wherein, the tracker system further comprises input means to allow  
5 a user input analysis parameters for analysing the multimedia file to prior to  
executing the automated analysis of the multimedia file.
38. A tracker system for analysing a multimedia file as claimed in claim 37,  
wherein, the analysis parameters input by the user comprise one or more of:  
10 movement paths, camera settings, time settings, background and internal  
areas, and, tracker configuration.

1 / 4

Figure 1

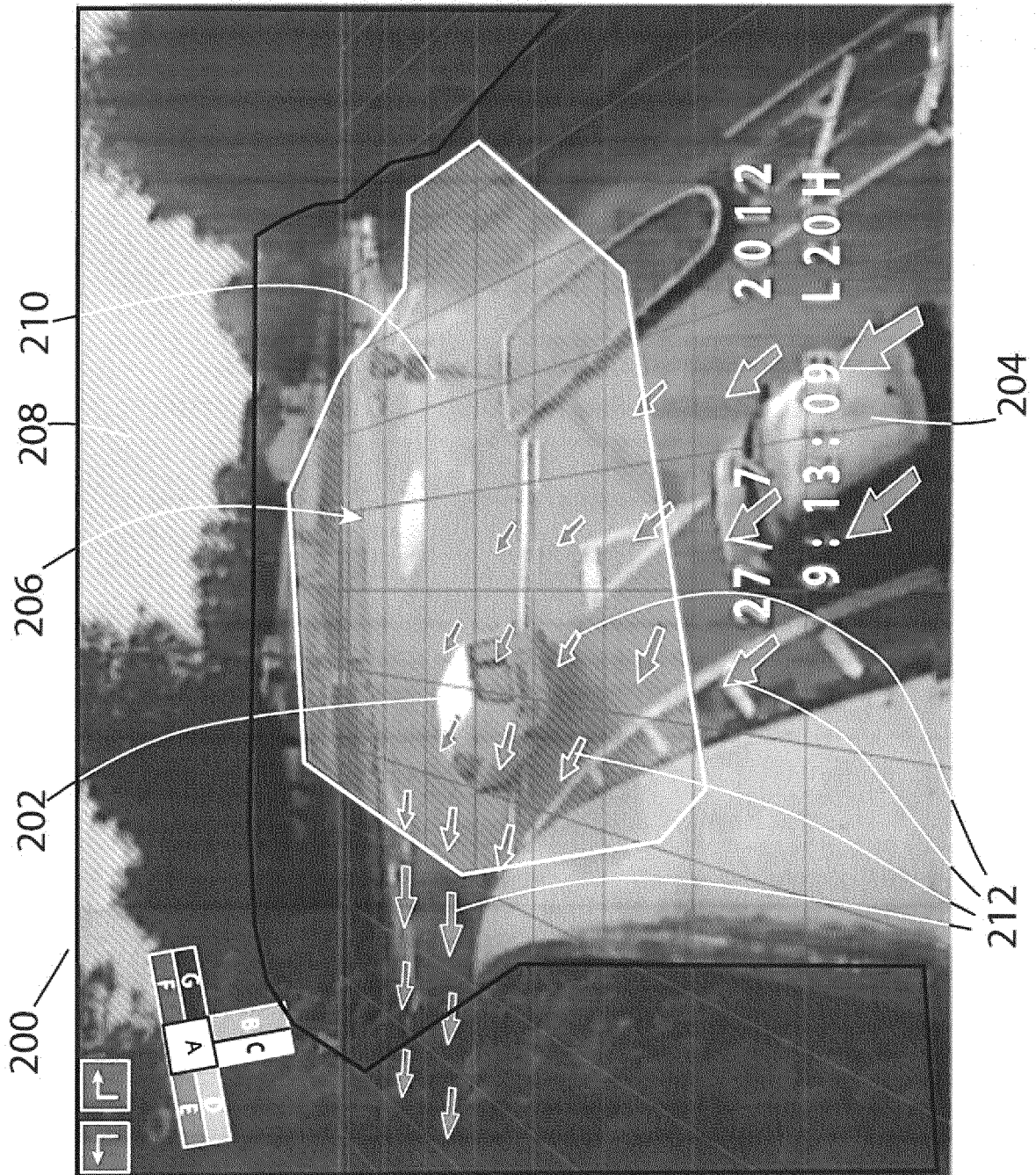


Figure 2

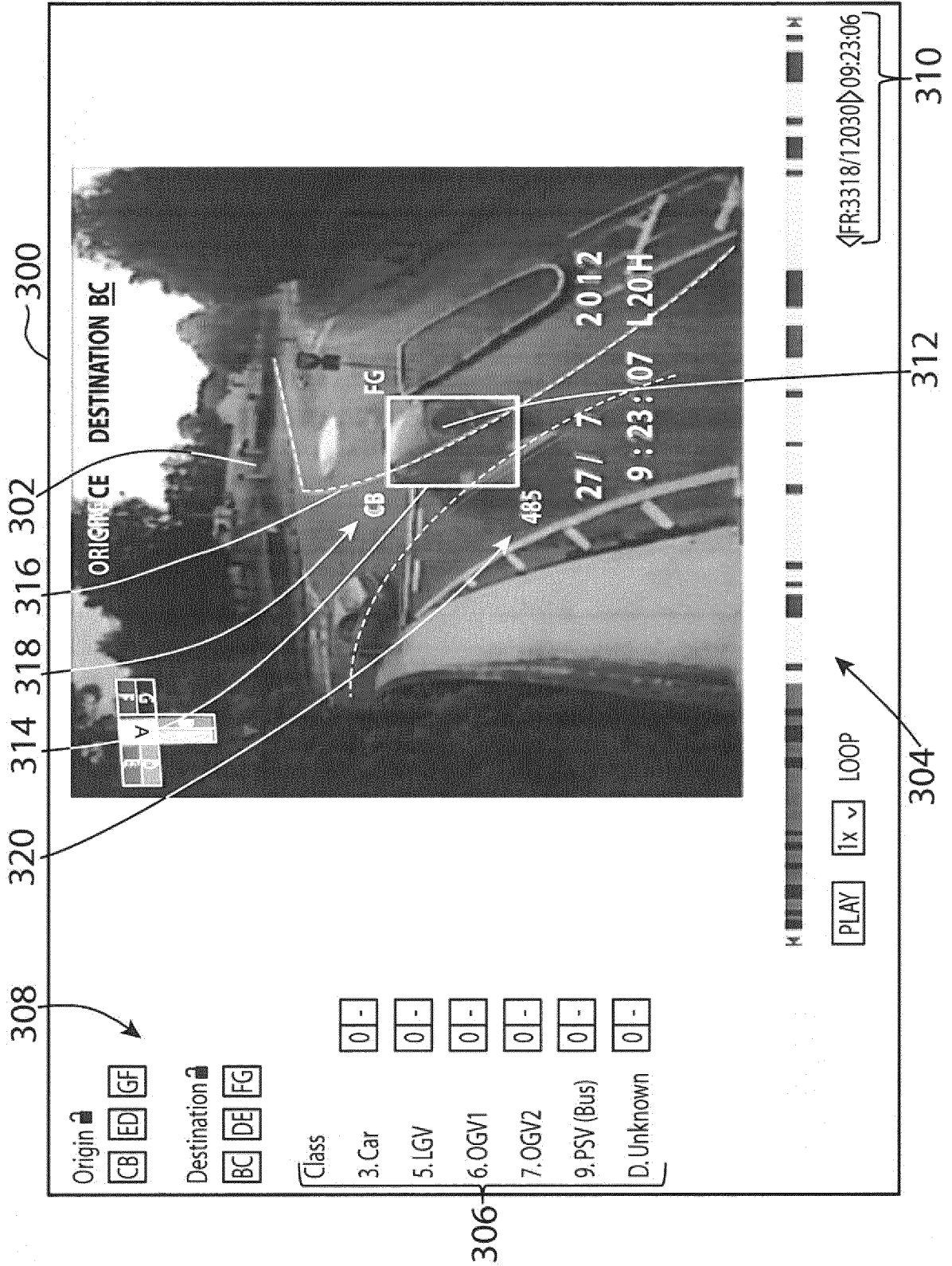


Figure 3

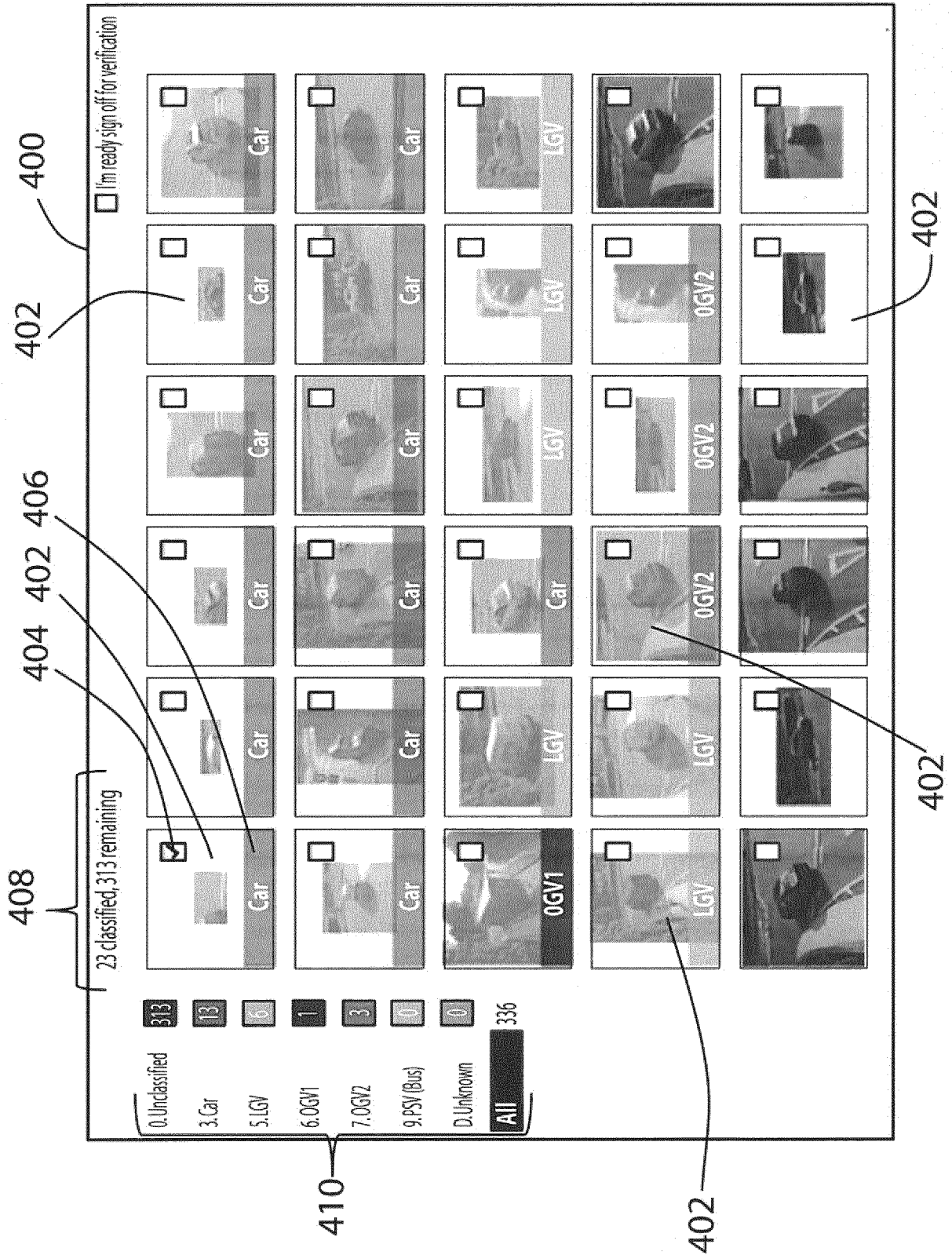


Figure 4

INTERNATIONAL SEARCH REPORT

International application No  
PCT/EP2016/056064

A. CLASSIFICATION OF SUBJECT MATTER  
INV. 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)  
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 practicable, search terms used)  
EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2009/244291 A1 (SAPTHARISHI MAHESH [US] ET AL) 1 October 2009 (2009-10-01)	1-4, 6-18, 20-23, 25-37
Y	figures 2,3,7,10 paragraph [0039] - paragraph [0062] paragraph [0087] - paragraph [0095] -----	5,19,24, 38
Y	EP 1 921 581 A1 (SMARTEREE GMBH [CH]) 14 May 2008 (2008-05-14)	5,19,24, 38
A	paragraph [0037] - paragraph [0042]  -----  -/--	1,11-14, 16-18, 20, 30-33, 35-37

Further documents are listed in the continuation of Box C.

See patent family annex.

\* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier application or patent but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of 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  13 May 2016	Date of mailing of the international search report  24/05/2016
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  de Bont, Emma



## INTERNATIONAL SEARCH REPORT

International application No  
PCT/EP2016/056064

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>COIFMAN B ET AL: "A real-time computer vision system for vehicle tracking and traffic surveillance", TRANSPORTATION RESEARCH. PART C, EMERGING TECHNOLOGIES, PERGAMON, NEW YORK, NY, GB, vol. 6C, no. 4, 1 August 1998 (1998-08-01), pages 271-288, XP008088782, ISSN: 0968-090X, DOI: 10.1016/S0968-090X(98)00019-9 the whole document -----</p>	1-38

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2016/056064

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2009244291	A1	01-10-2009	
		CN 101965576 A	02-02-2011
		CN 101965729 A	02-02-2011
		DE 112009000480 T5	07-04-2011
		DE 112009000485 T5	17-03-2011
		GB 2470520 A	24-11-2010
		GB 2471036 A	15-12-2010
		GB 2491987 A	19-12-2012
		GB 2492246 A	26-12-2012
		GB 2492247 A	26-12-2012
		GB 2492248 A	26-12-2012
		KR 20100124309 A	26-11-2010
		KR 20100126468 A	01-12-2010
		KR 20150123968 A	04-11-2015
		TW 200945218 A	01-11-2009
		TW 201005656 A	01-02-2010
		TW 201535276 A	16-09-2015
		US 2009244291 A1	01-10-2009
		US 2009245573 A1	01-10-2009
		US 2012274777 A1	01-11-2012
		US 2014176727 A1	26-06-2014
		US 2015092054 A1	02-04-2015
		US 2015093035 A1	02-04-2015
		US 2015278607 A1	01-10-2015
		WO 2009111498 A2	11-09-2009
		WO 2009111499 A2	11-09-2009
EP 1921581	A1	14-05-2008	NONE