(12)Indian Patent Application

(21) Application Number: 201741046582

(22) Filing Date: 26/12/2017 (43) Publication Date: 28/06/2019

(71) Applicant(s): L&T TECHNOLOGY SERVICES LIMITED

(72) Inventor(s): GOPINATH, C
RAMESH BABU, D

(51) International Classifications: G06T 7/00

(54) Title: A METHOD FOR GENERATING A 3D REAL MODEL

(57) Abstract: According to an embodiment, a method for generating a real-life customized scenario from a video having a plurality of image frames is disclosed. The image frame may be divided into a plurality of regions and at least one characteristic of the image frame may be determined. The method may include identifying image pixels related to at least one object in the image frame using a convolution network method. The method may further include identifying image pixels related to at least one object by segmenting & parsing the image frame. The image pixels of the objects identified through both methods may be compared pixel by pixel basis to generate an object identification map. After objects identification, the characteristics of the image frame may be corrected and a three-dimensional model may be generated using the corrected image frame and the object identification map.

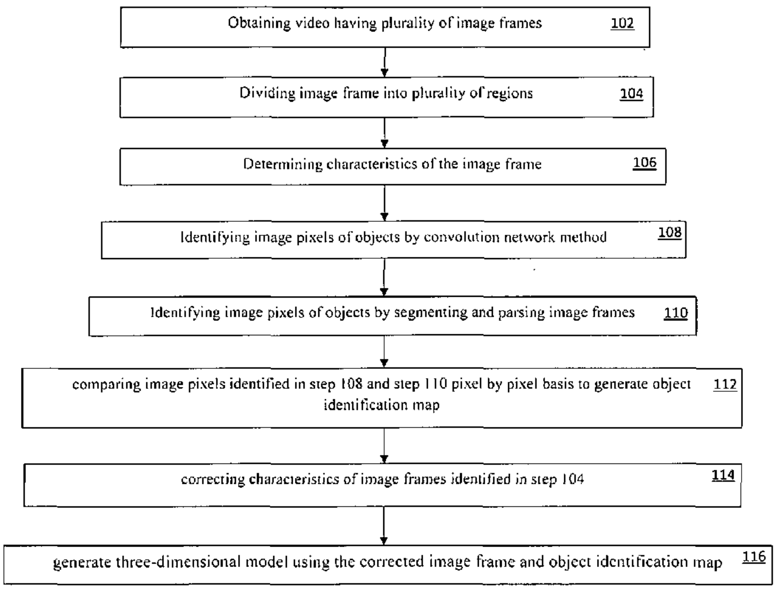
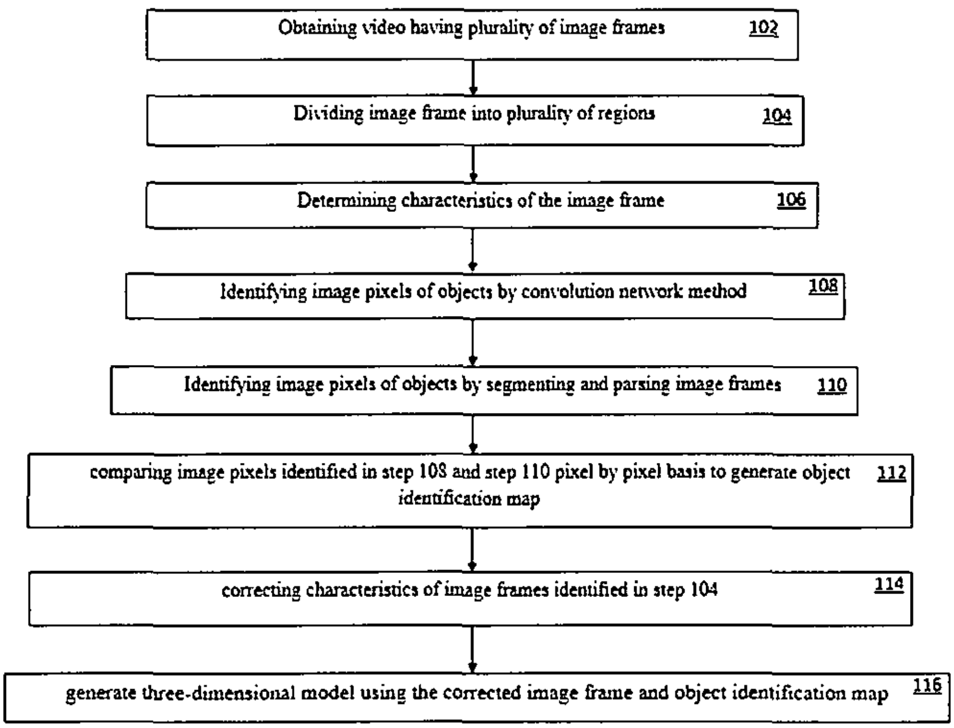


Figure 1

ABSTRACT

A Method for Generating a 3D Real Model

According to an embodiment, a method for generating a real-life customized scenario from a video having a plurality of image frames is disclosed. The image frame may be divided into a plurality of regions and at least one characteristic of the image frame may be determined. The method may include identifying image pixels related to at least one object in the image frame using a convolution network method. The method may further include identifying image pixels related to at least one object by segmenting & parsing the image frame. The image pixels of the objects identified through both methods may be compared pixel by pixel basis to generate an object identification map. After objects identification, the characteristics of the image frame may be corrected and a three-dimensional model may be generated using the corrected image frame and the object identification map.



FIELD OF INVENTION

The invention generally relates to image and video processing technology and more particularly to a method for generating a real-life customized scenario from a video.

BACKGROUND

With the rapid growth in technology, the ability to provide a computer generated virtual environment has become a reality. Such virtual environments have been proven effective for training systems, such as for driver training, pilot training, surgical procedures etc. The system involves combining pre-recorded or computer generated visual information with a real-world environment to provide the perception of a desired environment. For example, a driver's training simulator may include a physical-representation of the driver's seat of an automobile with a video or computer-generated image of a road. The image is made to be reactive to the actions of the driver, by changing speeds and perspectives in response to acceleration, braking and steering by the driver.

However, the testing in virtual environments has showed that the virtual environment does not satisfy all the real-time conditions. Besides, the virtual environment appears cartoonish and doesn't provide the same feel as that of the real-life driving. There is a further possibility that an ADAS (Advanced Driver Assistance Systems) algorithm, that works fine in the virtual environment might not work in real time condition because of continuous changing of environmental conditions.

Testing in real time has also been performed, but it is difficult to test all the scenarios. For recording all possible test cases, where each test case covers all possible scenarios, we need huge database or we need to record a video for more than a million hours. As per Euro NCAP (New Car Assessment Programme) statement, at least 9.5 billion hours of data is required for real time testing of ADAS algorithm.

Hence there is a need for an improved method for generating a real-life customized scenario from a video.

SUMMARY OF THE INVENTION

According to an exemplary embodiment of the invention, a method for generating a customized real-life scenario from a video is disclosed. The video may include a plurality of image frames, where each of the image frame may be divided into a plurality of regions. The method may include determining at least one or more of the characteristics of the image frame. The method may further include a step for identifying one or more image pixels related to at least one object in the image frame using a convolution network. The image pixels related to at least one object in image frame may be again identified by segmenting & parsing the image frame. The image pixels of the objects identified through the convolution network method and the segmenting and parsing may be compared on pixel by pixel basis to generate an object identification map. After the objects are identified, the characteristics of the image frame identified earlier may be corrected as required. A three-dimensional model may be generated using the corrected image frame and the object identification map.

BRIEF DESCRIPTION OF DRAWINGS

Other objects, features, and advantages of the invention will be apparent from the following description when read with reference to the accompanying drawing.

Figure 1 illustrates a flow chart of a method for generating a real-life customized scenario from a video having a plurality of image frames, according to an exemplary embodiment of the invention.

DETAILED DESCRIPTION OF DRAWINGS

The following description with reference to the accompanying drawing is provided to assist in a comprehensive understanding of exemplary embodiments. It includes various specific details to assist in that understanding but these are to be regarded as merely exemplary. Accordingly, those of ordinary skilled in the art will recognize that various changes and modifications of the embodiments described herein can be made without departing from the scope and spirit of the invention. In addition, descriptions of well-known functions and constructions are omitted for clarity and conciseness.

Figure 1 illustrates a flow chart of a method 100 for generating a real-life customized scenario from a video, according to an exemplary embodiment of the invention. The method 100 may be employed for generating various real-life environments for purposes such as, but not limited to, testing, gaming, entertainment etc. The video used for generating the real-life environment may include a plurality of image frames which when viewed in a sequence, plays as the video. The video having a plurality of image frames may be of a real-life environment. According to an embodiment, the video may be of a real-life road traffic condition including multiple objects such as, but not limited to, vehicles, pedestrians, buildings, trees etc. According to another embodiment, the video may be of a real-life flight condition. According to an embodiment, for capturing the road traffic conditions, a video recording device may be mounted on a vehicle and the vehicle may be driven on the road which needs to be recorded. The video recording device may be a camera or a camcorder mounted on the vehicle such as, but not limited to car, truck, bus etc. The distance driven by the vehicle along with the video recording device may depend on the tests to be performed on the real-life customized scenario and the distance may vary from metres to kilometres. By way of an example, the real-life road traffic environments are used in Automated Driver Assistance System (ADAS) for testing road vehicle driving systems. The real-life flight conditions are used in flight simulation for training civilian and military pilots. The real-life environment video recorded by the video recording device may be stored in a video repository.

At step 102, the method 100 collects the recorded video from the video repository. The video repository may be a data storage server/device such as, but not limited to cloud, hard drive, pen drive, compact disk etc. The recorded video may be fed to a real drive simulator system. The real drive simulator system may be used for driver trainings and for conducting research in human factors such as, but not limited to, monitoring driver behaviour, monitoring driver performance, monitoring driver attention etc. The real drive simulator system may be further used in vehicle industry to design and evaluate new vehicles or new advanced driver assistance systems.

At step 104, the method 100 divides each of the image frame in the real drive simulator system into a plurality of regions. According to an embodiment, the image frame may be divided into

3*3 regions. According to another embodiment, the image frame may be divided into 4*4 regions. According to yet another embodiment, the number of regions into which the image frame is divided may be decided by the user depending on the parameters such as, but not limited to, size of the image frame, number of objects in the image frame, clarity of the image frame, etc. As will be appreciated by those skilled in the art that the image frame may be divided for determining the regions of interest and the number of divided regions may vary based on different environments.

At step 106, the method 100 determines the value of at least one or more characteristics of the image frame. The value of at least one or more characteristics may be determined by analysing each of the plurality of regions of the image frame. The characteristics may indicate the day or night time of the environment. In other words, the characteristics of the image frame may indicate whether the image is of morning or afternoon or night. The characteristics of the image frames may be hue, saturation, gamma and white balance information. The values of the characteristics of the image frames may be altered to modify the image frame as per user requirements. By way of an example, the user may change the characteristics to convert an image frame having morning characteristics into an image frame having night characteristics. The change of characteristics may enable the user to create the real-life customised environment as per his requirement.

At step 108, the method 100 identifies at least one or more image pixels of at least one object in the image frame. The identification of the at least one or more image pixels of the at least one object may be detected from the regions of interest. The objects in the image frame may include elements such as, but not limited to, road, sky, person, trees, vehicles, buildings, clouds etc. By way of an example, the objects in the video of the real-life road traffic condition may be road, pedestrians, vehicle etc. and the objects in the video of real-life flight condition may be sky, clouds, birds etc. According to an embodiment, the image pixel of the object may be identified by a convolution network. The Convolution network may analyse the image frame with convolution layers to identify the objects in the image frame.

At step 110, the method 100 repeats the step of identifying at least one or more image pixels related to at least one or more object in the image frame. The objects in the image frame may include elements such as but not limited to road, sky, person, trees, vehicles, buildings, clouds etc. By way of an example, the objects in the video of real-life road traffic condition may be road, pedestrians, vehicle etc. and the objects in the video of real-life flight condition may be sky, clouds, birds etc. In this step 110, the image pixels related to the object are identified by segmenting and parsing the image frame. In segmenting and parsing, the image frame is segmented and parse into different image regions associated with semantic categories, such as road, person, vehicle etc.

At step 112, the image pixels of the objects identified in step 108 by convolution network and the image pixels of the objects identified in step 110 by segmenting and parsing the image frame may be compared on pixel by pixel basis to generate an object identification map. The image pixels of the objects are collected across multiple image frames.

At step 114, the characteristics of the image frames identified in step 106 may be corrected. The characteristics of each of the image frames may be corrected to perform changes in the image frame. These characteristics may enable a user to convert the lighting conditions of the environment to day or night conditions in the image frame. The conversion of the characteristics may be achieved by methods or a combination of methods such as, but not limited to adaptive gamma correction, convolution networks etc. The characteristics may be further corrected to synchronise each of the objects along with the other objects in the image frame.

At step 116, the method 100 maps all the image frames in a three-dimensional space to generate a three-dimensional model. The three-dimensional model may include the objects in a three-dimensional view.

According to an embodiment, the objects in the image frame may be removable from the video captured by the video recording device. To remove the object, the pixels of the object may be removed from the image frame. On removing the pixels of the object, the pixels are replaced by the neighbouring pixels. The neighbouring pixels may replace the space left by the pixels of the object.

According to an embodiment, the objects may be added or moved in the video captured by the video recording device. To add an object, the three-dimensional object may be added at the required position in the three-dimensional or two-dimensional model. The added object may be further synchronized with the other objects in the three-dimensional or two-dimensional model.

It is understood that the above description is intended to be illustrative, and not restrictive. It is intended to cover all alternatives, modifications and equivalents as may be included within the spirit and scope of the invention as defined in the appended claims. Many other embodiments will be apparent to those of skill in the art upon reviewing the above description. The scope of the invention should, therefore, be determined with reference to the appended claims, along with the full scope of equivalents to which such claims are entitled. In the appended claims, the terms "including" and "in which" are used as the plain-English equivalents of the respective terms "comprising" and "wherein," respectively.

We claim:

1. A method for generating a real-life customized scenario from a video, the video having a plurality of image frames, the method comprising:

   step 1 - dividing each image frame into a plurality of regions;

   step 2 - determining at least one of the characteristics of the image frame;

   step 3 - identifying one or more image pixels related to at least one object in the image frame using convolution network method;

   step 4 - repeating the step of identifying the image pixels related to at least one object in image frame by segmenting & parsing the image frame;

   step 5 - comparing the image pixels of the objects identified in step 3 and step 4 on pixel by pixel basis to generate an object identification map;

   step 6 - correcting the characteristics of the image frame identified in step 2; and

   step 7 - generate a three-dimensional model using the corrected image frame and the object identification map.

2. The method for generating a real-life customized scenario from the video as claimed in claim 1, wherein the characteristics of the image frame include hue, saturation, gamma and white balance information.

3. The method for generating a scenario from the video as claimed in claim 1, wherein the image pixels of the objects identified in the object identification map are mapped in a three-dimensional space to generate a three-dimensional view of the objects.

4. The method for generating a scenario from the video further comprising:

   removing the image pixels of the objects identified in the object identification map; and

   replacing the image pixels of the object with the neighboring pixels.

5. The method for generating a scenario from the video further comprising:

   placing the three-dimensional object at a desired position in the three-dimensional or two-dimensional model; and

   synchronizing the three-dimensional object with the three-dimensional or two-dimensional model.

Dated this 26th day of December 2017

Mohammed Faisal (INPA No: 1941)
Head, IPR Dept.
L&T Technology Services Limited
DLF 3rd Block, 2nd Floor,
Manapakkam, Chennai – 600089

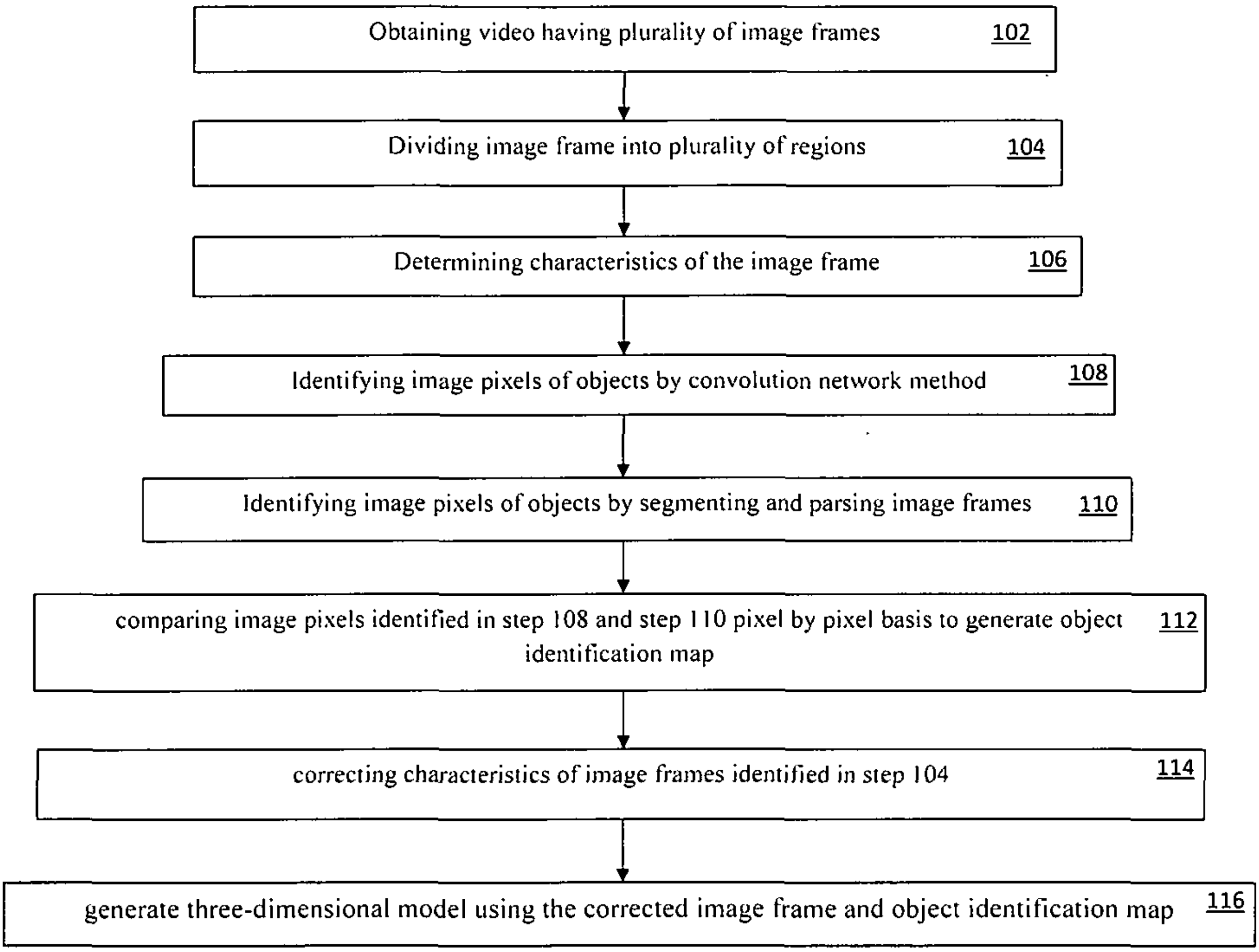


Figure 1

Mohammed Faisal (INPA No: 1941)
Head, IPR Dept.
L&T Technology Services Limited
DLF 3rd Block, 2nd Floor,
Manapakkam, Chennai – 600089