



US 20150012200A1

(19) **United States**

(12) **Patent Application Publication**
JOSHI et al.

(10) **Pub. No.: US 2015/0012200 A1**

(43) **Pub. Date: Jan. 8, 2015**

(54) **SYSTEM AND METHOD FOR TRAFFIC SIGNAL RECOGNITION**

Publication Classification

(71) Applicant: **Toyota Motor Engineering & Manufacturing North America, Inc.**, Erlanger, KY (US)

(51) **Int. Cl.**
B60R 11/04 (2006.01)
G05D 1/02 (2006.01)

(72) Inventors: **Aydhut S. JOSHI**, Ann Arbor, MI (US);
Danil V. PROKHOROV, Canton, MI (US)

(52) **U.S. Cl.**
CPC **B60R 11/04** (2013.01); **G05D 1/0276** (2013.01); **B60R 2300/205** (2013.01)
USPC **701/70**; 348/148

(73) Assignee: **Toyota Motor Engineering & Manufacturing North America, Inc.**, Erlanger, KY (US)

(57) **ABSTRACT**

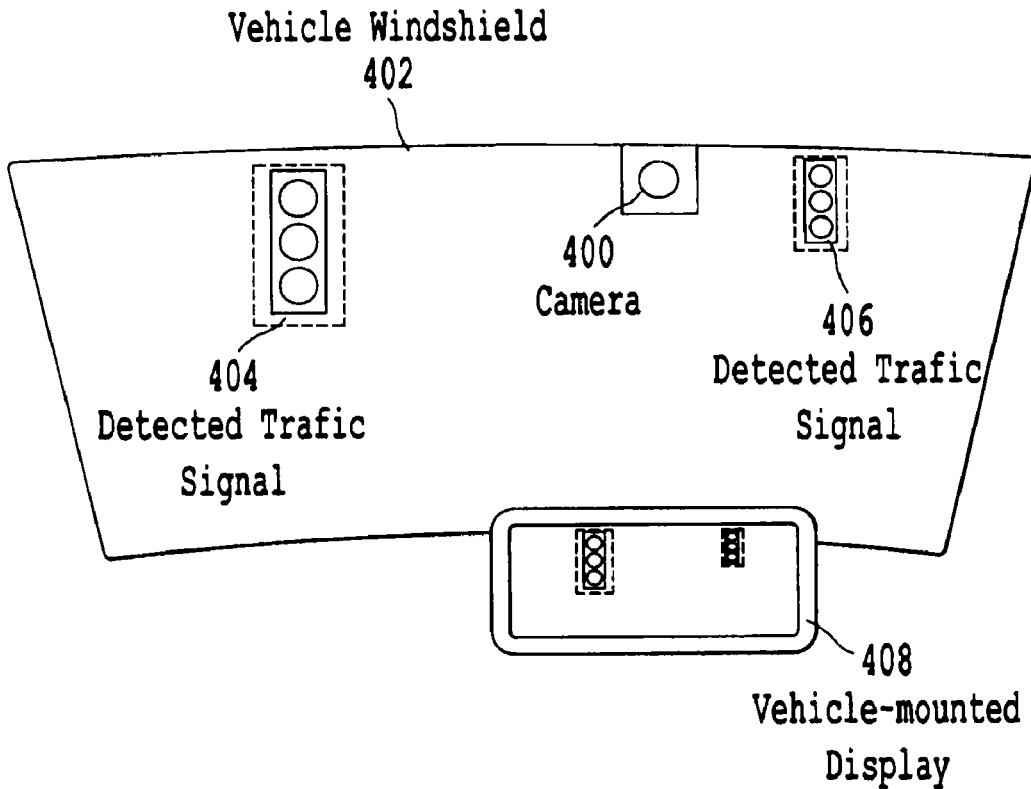
Detecting traffic signaling of a mobile environment includes receiving image frames captured by an imaging device to detect a traffic signal candidate therefrom based on at least a vehicle location. A score having a predefined value is established for the traffic signal candidate. The traffic signal candidate is tracked by incrementing the score when the traffic signal candidate is determined to appear, based on a probabilistic estimator that indicates an expected value for a location of the traffic signal candidate, in a tracked image frame of the captured image frames. The score is decremented when the traffic signal candidate is determined not to appear, based on the probabilistic estimator, in the tracked image frame. The traffic signal candidate is recognized as a traffic signal based on whether the score is above a specified threshold.

(21) Appl. No.: **14/455,381**

(22) Filed: **Aug. 8, 2014**

Related U.S. Application Data

(63) Continuation of application No. 13/371,722, filed on Feb. 13, 2012, now Pat. No. 8,831,849.



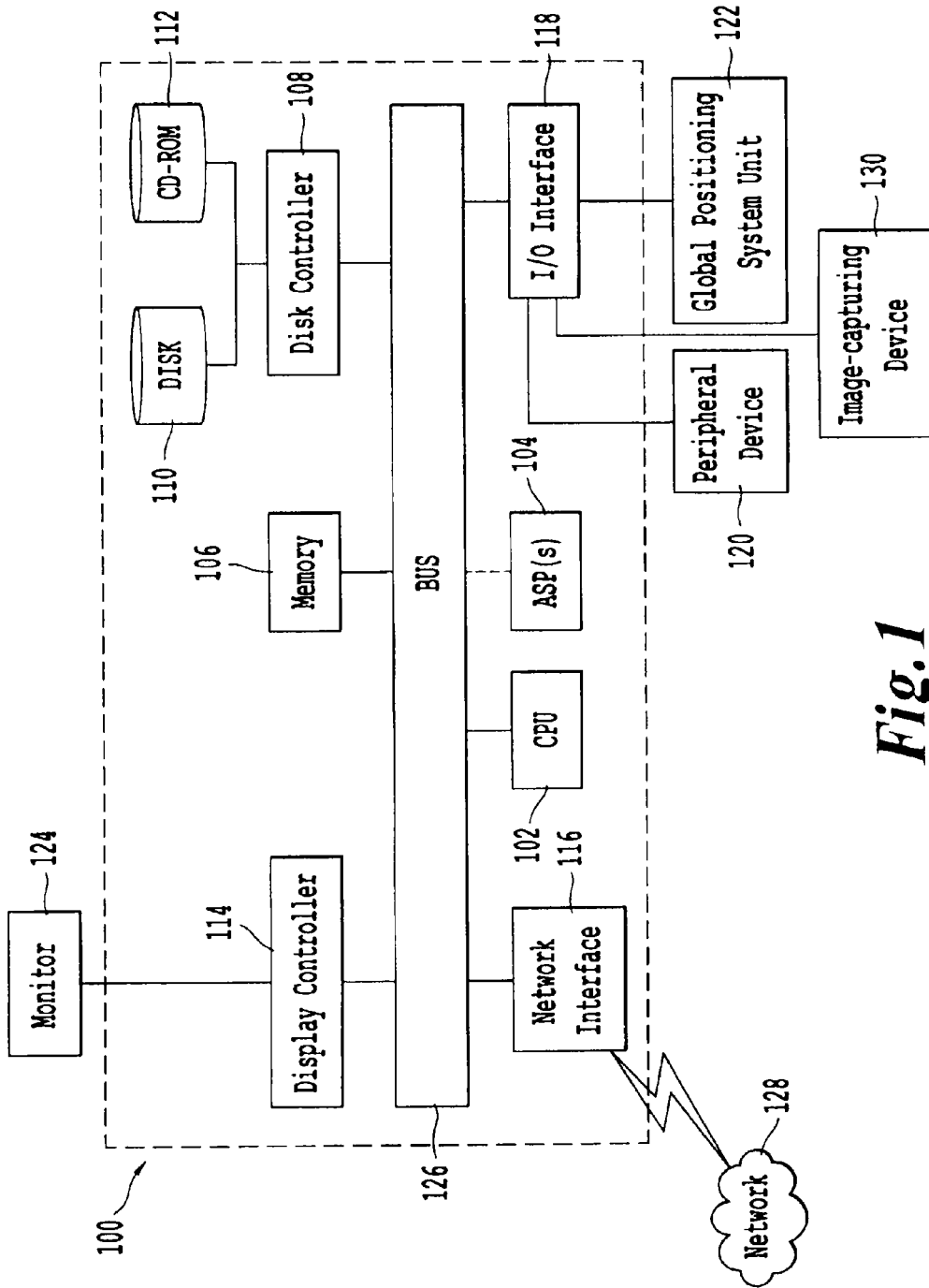


Fig. 1

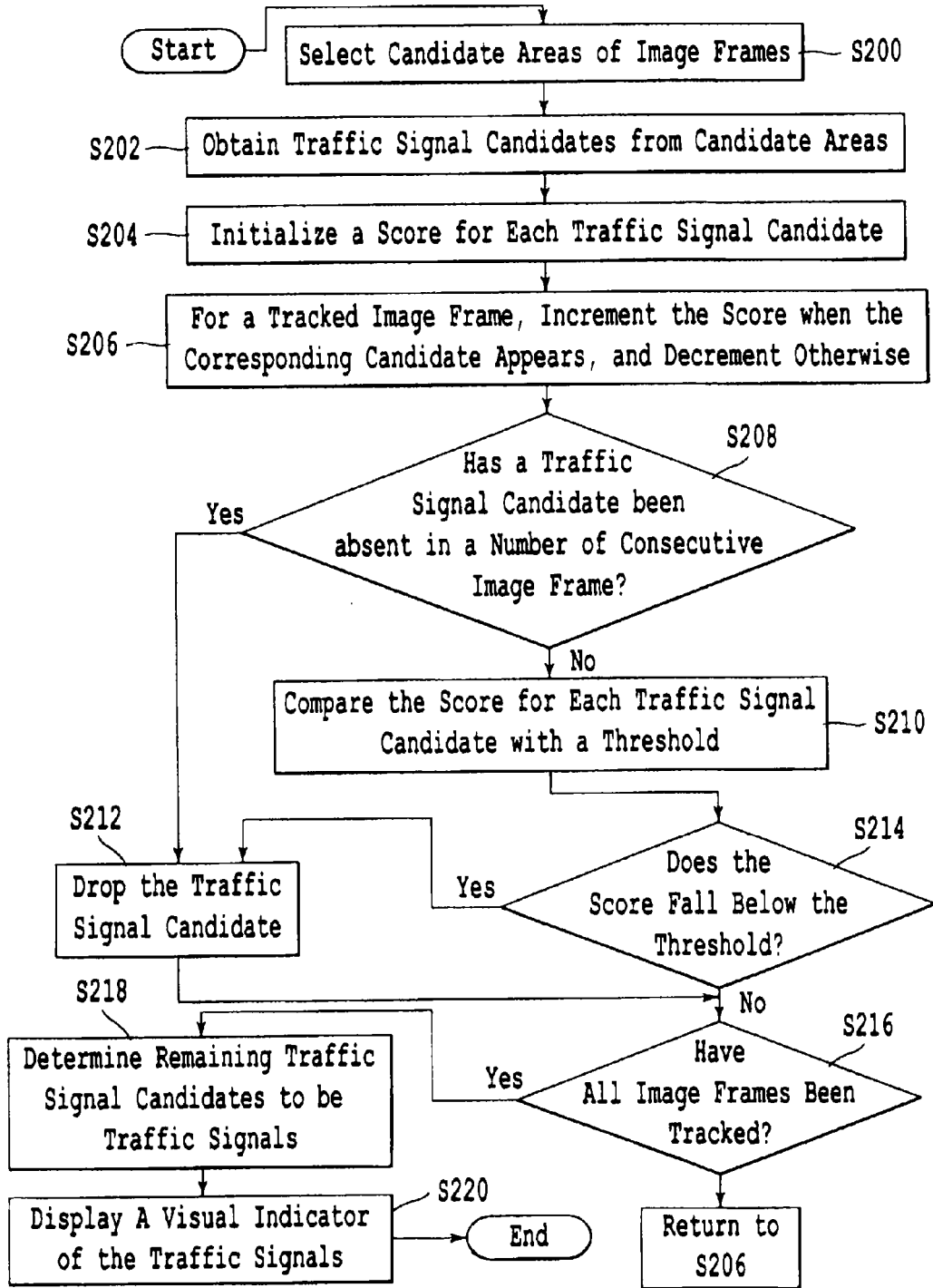


Fig. 2

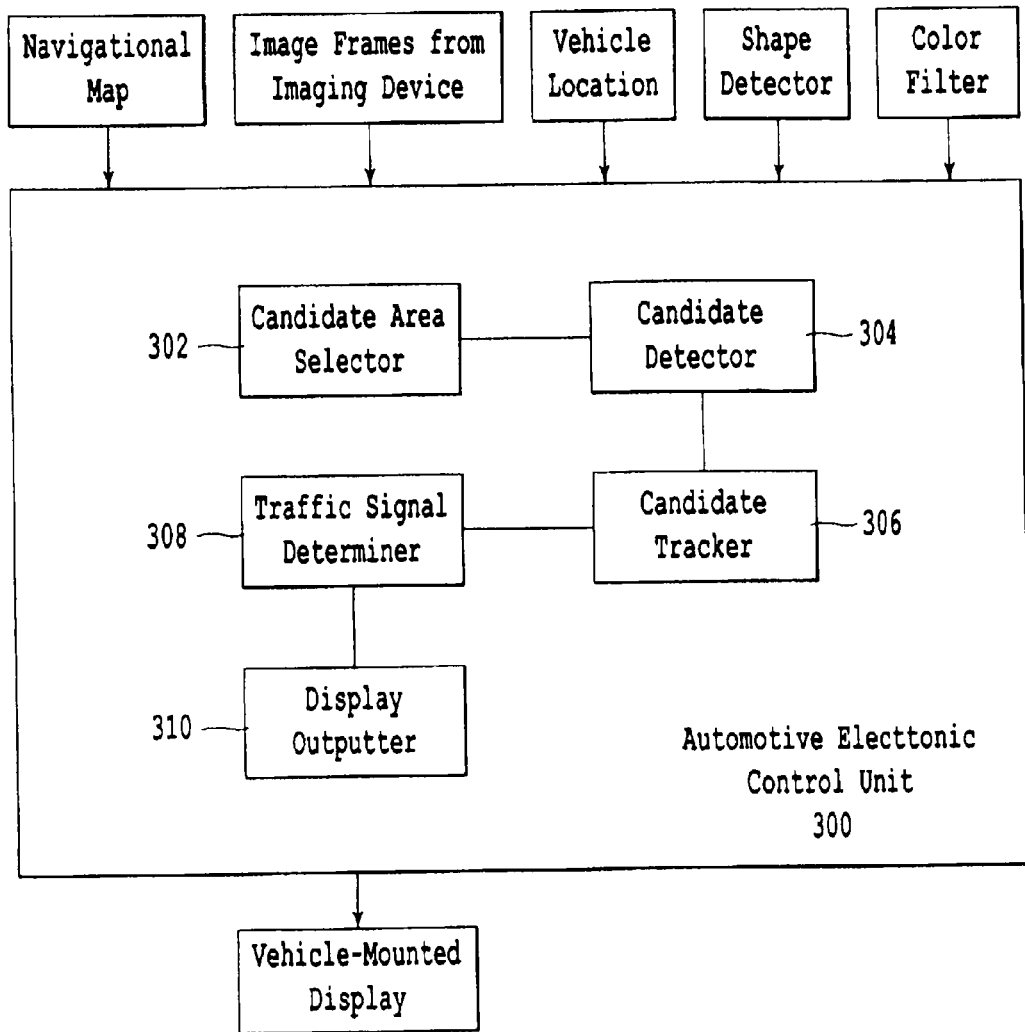


Fig. 3

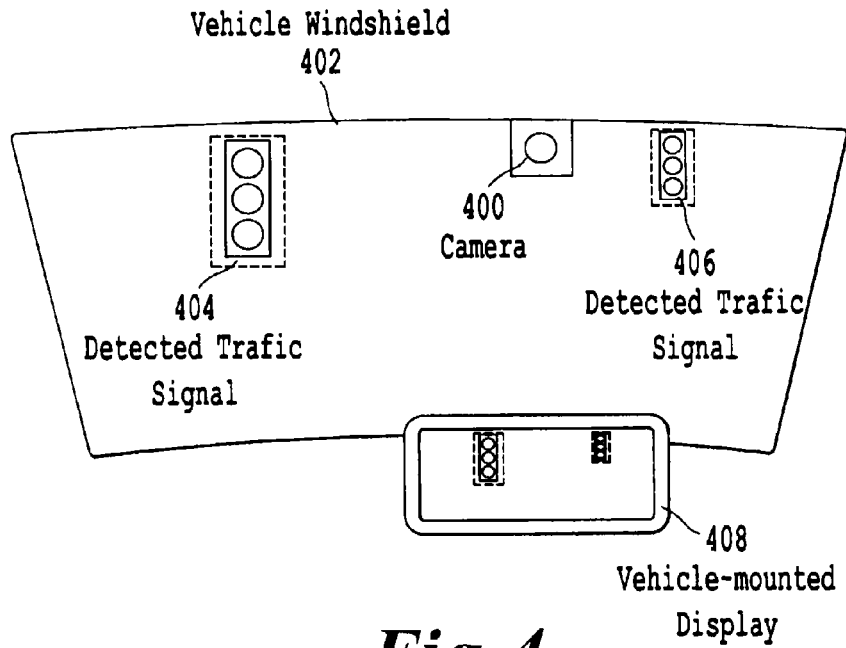


Fig. 4

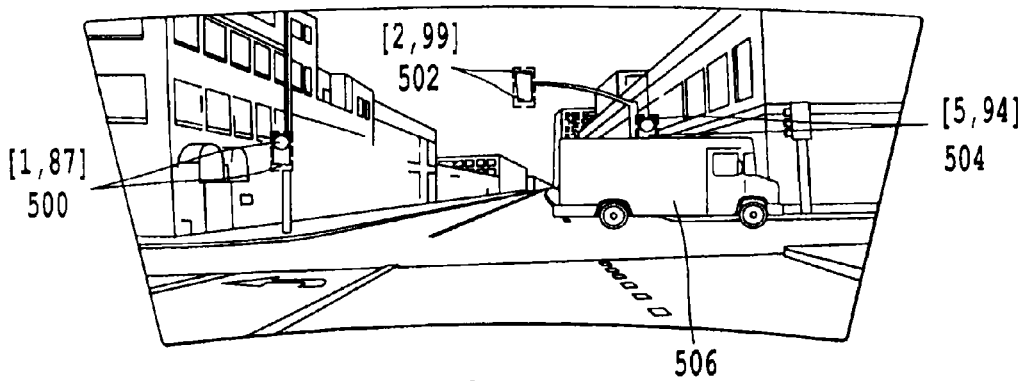


Fig. 5

SYSTEM AND METHOD FOR TRAFFIC SIGNAL RECOGNITION

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application is a continuation of U.S. application Ser. No. 13/371,722 filed Feb. 13, 2012, the entire content of which is incorporated herein by reference.

FIELD OF THE DISCLOSURE

[0002] This application relates to recognizing the presence of traffic signals from traffic signal candidates detected from image frames captured by a vehicle-mounted camera, by using historical observation data obtained using a probabilistic estimator.

BACKGROUND

[0003] The “background” description provided herein is for the purpose of generally presenting the context of the disclosure. Work described herein, to the extent it is described in this background section, as well as aspects of the description which may not otherwise qualify as prior art at the time of filing, are neither expressly or impliedly admitted as prior art.

[0004] Traffic signal recognition refers to the process of identifying and marking areas of an image frame that include a traffic signal, for example, a traffic light. A camera obtains the image frames, which are processed individually. Traffic light candidates and areas in the immediate vicinity are determined based on matching an actual traffic template with a template from a template library. The accuracy of traffic signal recognition may be increased by using adaptive templates. Adaptive templates are adjusted according to an expected size of the traffic light candidate and/or the actual traffic light template, which is based on a distance of a vehicle from the traffic light candidate. A best matched template is validated based on a number of possible templates that are matched in size to the actual image frame area corresponding to the traffic light candidate. For example, one of the possible templates may be used when a green circle in the traffic light candidate is illuminated and another of the possible templates may be used when a red circle in the traffic light candidate is illuminated. This is because the red and green circles are at different locations within the traffic light housing.

SUMMARY

[0005] This disclosure addresses various problems associated with known traffic signal recognition processes which the inventors have discovered. In particular, in the traffic signal recognition process involving template matching, a template description is required for each template in a template library, which is difficult or impossible to implement for previously unseen traffic lights. Further, adjacent image frames in a captured image frame sequence or video are processed independently, without taking into account temporal and spatial relationships. For example, in order to identify traffic light candidates within an image frame, all areas of the frame must be analyzed, which is a computationally intensive process. In addition, when the image frames include many objects, the process can require even more computer resources. Moreover, objects within the image frame may be incorrectly identified based on a traffic light recognition algo-

rithm used to perform the traffic light detection. This results in an increase in the amount of time a user must spend to correct these misidentifications.

[0006] This disclosure is directed to a traffic signal recognition method that, in one non-limiting embodiment, detects candidates for traffic signals based on at least one of a navigational map or a global positioning system (GPS) location of a vehicle. The method uses a color filter to filter image frames obtained from a vehicle-mounted camera. The filtered images are used to detect traffic signal candidates. In a non-limiting embodiment, the traffic signal recognition method is implemented in any of the following, but not limited to: a personal communication device (e.g., cellular phone, data-enabled mobile phone, smartphone) and a portable electronic device (e.g., handheld GPS receiver, portable gaming console, laptop, tablet computer). In another non-limiting embodiment, the device in which the traffic signal recognition method is implemented includes any of the following, but not limited to: an image capturing function; a network connection (e.g., cloud network, the Internet, intranet); and a GPS function.

[0007] In another non-limiting embodiment, a distance of the vehicle from the traffic signal candidate is used to determine an expected size of the traffic signal candidate. The expected size of the traffic signal is used to search for traffic signal candidate areas and to examine neighborhoods of possible traffic signal candidate areas. A shape detector is used to reject false traffic signal candidates that are generated due to noise in image frames captured by the camera.

[0008] A probabilistic estimator is subsequently employed in a frame-by-frame tracking of the detected traffic light candidates. The tracking process allows evidence to be accumulated in favor of each traffic light candidate when the traffic light candidate is observed in a frame. A traffic light candidate is dropped from consideration when the accumulated evidence falls below a specified threshold, for example, when the traffic light candidate is not observed for a specified amount of time.

[0009] In another non-limiting embodiment, the method also includes monitoring traffic signals for an instance in which a particular color light turns off and another color of light is subsequently illuminated. This enables a traffic light candidate to be detected in the immediate vicinity of the light that turned off (e.g., in the up or down directions). Therefore, traffic signal candidates are tracked so as to reliably recognize traffic signals, for example traffic lights, irrespective of the shape and size of a housing for the traffic light. False positive identifications of traffic signals are also rejected based on consistent spatial and/or temporal behavior.

[0010] The foregoing paragraphs have been provided by way of general introduction, and are not intended to limit the scope of the following claims. The described embodiments, together with further advantages, will be best understood by reference to the following detailed description taken in conjunction with the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

[0011] A more complete appreciation of this disclosure and many of the attendant advantages thereof will be readily obtained as the same becomes better understood by reference to the following detailed description when considered in connection with the accompanying drawings, wherein:

[0012] FIG. 1 is a block diagram showing hardware components of a traffic signal recognition system;

[0013] FIG. 2 shows an exemplary algorithm for traffic signal recognition;

[0014] FIG. 3 shows an exemplary automotive electronic control unit implementing traffic signal recognition;

[0015] FIG. 4 illustrates a view from a vehicle including the traffic signal recognition system and displaying recognized traffic lights; and

[0016] FIG. 5 is an exemplary image frame that indicates recognized traffic signals.

DETAILED DESCRIPTION

[0017] In the drawings, reference numerals designate identical or corresponding parts throughout the several views. FIG. 1 shows an exemplary structure for a recognition system 100. The recognition system 100 includes computer hardware components that are either individually programmed or execute program code stored on various recording media. In an exemplary embodiment, the recognition system 100 includes a central processing unit (CPU) 102, one or more application-specific processors (ASPs) 104, memory 106, a disk controller 108 controlling operations of a hard disk drive 110 and/or an optical disk drive 112 (e.g., a CD-ROM drive), a display controller 114, a network interface 116, and an I/O interface 118. The I/O interface 118 is connected to a peripheral device 120, a Global Positioning System (GPS) Unit 122, and an image-capturing device 130. The display controller 114 is connected to a monitor 124. In some implementations, the monitor 124 is a component of the peripheral device 120, which may share a communication/data channel with the monitor 124. In some implementations, the I/O interface 118 is further connected to user-input peripherals such as touch-screen monitor controller, a keyboard and a mouse. Further, the components shown in FIG. 1 share a common BUS 126 for transferring data, and the controllers may include application specific integrated controllers and other programmable components.

[0018] The network interface 116 allows the recognition system 100 to communicate via a network 128, such as the Internet. In some aspects, any of the peripheral device 120, the GPS unit 122, and the image-capturing device 130 are coupled to the recognition system 100 via the network 128.

[0019] Referring now to FIG. 2, in step S200, candidate areas having a relatively high prior probability of including a traffic signal are selected from image frames obtained by a vehicle-mounted camera. In one non-limiting embodiment, a navigational map is used to determine the most likely positions for traffic signals to reduce or eliminate the need to initially search in large areas of the image frames. In another non-limiting embodiment, a GPS location and orientation of the vehicle is used to narrow down a list of candidate areas. In yet another non-limiting embodiment, a saliency map is created from the camera projection.

[0020] In step S202, traffic signal candidates are obtained from the candidate areas and are each instantiated as an object. In one non-limiting embodiment, a filter, for example, a top hat filter, is used on specific color channels, of a captured image frame, that correspond to certain frequencies. In Fourier space, the top hat filter selects a desired frequency band by specifying a lower and upper bounding frequency. In another non-limiting embodiment, a shape detector is used on the filtered image frame to reject false traffic signal candidates generated due to noise in the captured image frame. When the traffic signal is known to be a certain distance from the vehicle-mounted camera, for example, based on informa-

tion obtained from the navigational map, objects of the size corresponding to the given distance having, for example, a round shape are searched. Alternatively, neighborhoods of possible traffic signal candidates on the image frame corresponding to the previous time step are examined. In yet another non-limiting embodiment, a subset of the image frames is searched to reduce usage of processing resources. In this manner, traffic signals of many shapes and forms are accurately recognized. Step S202 also includes monitoring for traffic signals for an instance in which a particular traffic light color turns off and another traffic light color turns on. A traffic signal candidate is therefore detected in the immediate vicinity of the light that turned off (e.g., in the up or down directions).

[0021] In step S204, each traffic signal candidate is associated with an initialized score to track and accumulate evidence in favor of the traffic signal candidates.

[0022] In step S206, frame-by-frame tracking occurs for each traffic light candidate using a probabilistic estimator. In one non-limiting embodiment, the probabilistic estimator is any of the following, but not limited to: a Kalman filter, an iterated Kalman filter, an extended Kalman filter, a sigma-point filter. The present disclosure describes the probabilistic estimator as implemented by a Kalman filter. However, as understood by those of ordinary skill in the art, the probabilistic estimator is implementable by other filters, such as those set forth above. The probabilistic estimator (e.g., a Kalman filter) is used to track traffic signals. That is, the Kalman filter generates a predicted location for each traffic signal candidate based on previous estimations and observations of traffic signal candidate in previous image frames. The prediction is modeled by the state dynamics of the Kalman filter and is based on information about the ego-vehicle motion (i.e., the motion of the camera in the vehicle), the relatively stationary nature of traffic signals and in one non-limiting embodiment, the size of the traffic signal candidate, to improve the outcome of the tracking algorithm (described in further detail below).

[0023] The tracking algorithm accumulates evidence, i.e., by incrementing the score each time a traffic signal candidate appears in a tracked image frame, in favor of each traffic signal candidate. The score associated with the traffic signal candidate is decremented when the traffic signal candidate does not appear in the tracked image frame. More particularly, the Kalman filter is updated based on x-, y-, and in one non-limiting embodiment, z-coordinate information for each traffic signal candidate as well as a probability, P, and a prediction is generated for the most likely location of the traffic signal, based on the previous detection or classification. The prediction is associated with a traffic signal candidate in the stored list of previously observed traffic signal candidates. If a new traffic signal candidate is observed, a new Kalman filter and a new object are initialized. A Kalman filter posterior update is made using a new observation, and the score is incremented for observed traffic signals. In one non-limiting embodiment, when the prediction is not available for the previously observed objects (e.g., due to their temporary occlusions), the score is decremented.

[0024] A Kalman filter is based on a linear dynamic system discretized in the time domain. It is modeled based on a Markov chain built on linear operators perturbed by Gaussian noise. The state of the system is represented as a vector of real numbers. At each discrete time increment, a linear operator is applied to the state to generate a new state, mixed with some noise, and information from controls in the system, if they are

known. Then, another linear operator mixed with more noise generates the observed outputs from the true (“hidden”) state:

$$x_k = F_k x_{k-1} + B_k u_k + w_k \quad (\text{Eq. 1})$$

where

[0025] F_k is the state transition model which is applied to the previous state x_{k-1} ;

[0026] B_k is the control-input model which is applied to the control vector u_k ; and

[0027] w_k is the process noise which is assumed to be drawn from a zero mean multivariate normal distribution with covariance Q_k .

$$w_k \sim N(0, Q_k) \quad (\text{Eq. 2})$$

[0028] At time k an observation (or measurement) z_k of the true state x_k is made according to

$$z_k = H_k x_k + v_k \quad (\text{Eq. 3})$$

where

[0029] H_k is an observation model which maps a true state space into an observed space; and

[0030] v_k is the observation noise which is assumed to be zero mean Gaussian white noise with covariance R_k .

$$v_k \sim N(0, R_k) \quad (\text{Eq. 4})$$

[0031] Thus, the process is robust with respect to instantaneous classification errors (e.g., false positives due to noise, which typically do not propagate temporally) and temporary occlusions of the traffic signal candidate.

[0032] Also in step S206, the traffic signal candidates are monitored for changes, for example, light changes. The monitoring is initiated when the light turns off and detects any traffic signal candidates in the immediate vicinity of the light that turned off (e.g., in the up or down directions) having a different light illuminated. An object is instantiated and a filter is initialized for the new traffic light candidate. Note that shape of the new traffic signal candidate can change, at least slightly, based on different angles of viewing and due to different orientation (e.g., yaw and pitch angles) of the on-board camera in the vehicle.

[0033] In step S208, a determination is made, based on the Kalman filter and the tracking algorithm, as to whether the traffic signal candidate has been absent in a specified number of consecutive image frames. The traffic signal candidate is dropped from consideration, in step S212, if traffic signal candidate has been absent in the specified number of consecutive image frames. If the traffic signal candidate has not been absent in the specified number of consecutive image frames, the process proceeds to step S210 in which the score is compared with a predetermined threshold. A determination as to whether the score is less than the predetermined threshold in step S214 and if so, the traffic signal candidate is dropped as a candidate in step S212.

[0034] Otherwise, the process continues to step S216 where it is determined whether each image frame to be tracked has been tracked and if so, the remaining traffic signal candidates are recognized and classified as traffic signals in step S218. If not, the process returns to step S206. In step S220, a visual indicator of the recognized traffic signals determined in step S218 is displayed on a screen, dashboard, windshield, or other vehicle-mounted display, and the process ends.

[0035] Referring now to FIG. 3, a system diagram of an automotive electronic control unit 300 is shown. FIG. 3 illustrates a functional relationship of hardware as modified by

software or computer-executable instructions. That is, the automotive electronic control unit 300 shown in FIG. 3 corresponds to hardware, or a combination of hardware and software. More particularly, the automotive electronic control unit is implemented by at least a portion of the tangible hardware shown in FIG. 1 and optionally, software or computer-executable instructions corresponding to the algorithm shown in FIG. 2.

[0036] The automotive control unit 300 receives image frames from a vehicle-mounted imaging device, for example, a video camera. The automotive control unit 300 also receives any one or more of the following, but not limited to: a navigational map including for example, streets and intersection locations, a vehicle location and orientation (e.g., pose or GPS coordinates), a shape detector algorithm, and a color filter algorithm. Using the image frames, a candidate area selector 302 uses either or both of the navigational map and the vehicle location and orientation to determine candidate areas for traffic signal candidates in each of the image frames. The image frames, with the identified candidate areas, are passed to a candidate detector 304.

[0037] The candidate detector 304 detects traffic signal candidates from the candidate area and each is associated with an object or data structure. Alternatively or additionally, the traffic signal candidates are detected based on an analysis of the entire image frame. As described above, the traffic signal candidates are obtained by employing either or both a color filter and a shape detector. A list of traffic signal candidates is passed to a candidate tracker 306.

[0038] As described above, each traffic signal candidate is associated with an initialized score. A probabilistic estimator, for example, a Kalman filter, is used to track the traffic signal candidates. As described above, the Kalman filter generates a predicted location for each traffic signal candidate based on previous estimations and observations of traffic signal candidate in previous image frames. Evidence in favor of the traffic signal candidates is accumulated by incrementing a score associated with a traffic signal candidate each time the traffic signal candidate appears in a tracked image frame. The candidate tracker 306 decrements the score associated with the traffic signal candidate when the traffic signal candidate does not appear in the tracked image frame. When the score associated with the traffic signal candidate falls below a predetermined threshold, or when the traffic signal candidate has not appeared in a specified consecutive number of image frames, the traffic signal candidate is dropped and its associated object is removed from the list. The list of traffic signal candidates is then passed to a traffic signal determiner 308.

[0039] The traffic signal determiner 308 determines whether each image frame to be tracked has been tracked and if so, determines that the remaining traffic signal candidates in the list are classified and recognized as traffic signals. Position information corresponding to the traffic signal candidates is passed to a display outputter 310. The display outputter 310 outputs a signal to draw a visual indication to be displayed on at least one of, but not limited to: a screen, dashboard, windshield, vehicle-mounted display, personal communication device, and a portable electronic device. For example, the visual indication is a colored polygon or free-form shape enclosing or otherwise indicating a location corresponding to the recognized traffic signal on the display. As another example, the visual indication is patterned to assist colorblind or color-impaired drivers. As yet another example,

the visual indication is text, for example, words such as “GO”, “STOP”, “READY”, “RED”, “GREEN”, or “YELLOW”.

[0040] In another non-limiting embodiment, either in addition or in alternatively from outputting a visual indication, position information corresponding to the traffic signal candidates determined by the traffic signal determiner **308** is used as input to a vehicle control system. For example, the position information can be input to a vehicle controller that directs an active safety module that controls the vehicle to automatically stop, go, or accelerate.

[0041] In FIG. 4, a view from inside the vehicle is illustrated. A camera **400** is mounted near a windshield **402** of the vehicle. In one non-limiting embodiment, the camera **400** mounted outside of or on top of the vehicle. As will be understood by one of ordinary skill in the art, the camera **400** is mountable anywhere on or around the vehicle as long as image frames showing approaching terrain (with a sufficiently wide angle of view) are captured. In one non-limiting embodiment, detected traffic signals **404** and **406** are enclosed with dashed line boxes that are illuminated, or otherwise indicated, on the vehicle windshield **402**. In another non-limiting embodiment, the detected traffic signals are enclosed with dashed line boxes that are illuminated, or otherwise indicated, on a vehicle-mounted display **408**.

[0042] In FIG. 5, a sample image frame including visual indicia around three traffic signals is shown. Two traffic signals **500** and **502** are recognized with a specified degree of confidence, and therefore, are marked with boxes of a particular color, shape, or pattern. The third traffic signal **504**, which is temporarily occluded by a passing vehicle **500**, but the presence of the traffic signal **504** is still recognized based on evidence accumulated during detection results in the previous image frames and the Kalman filtering for the candidate corresponding to the specified traffic signal. In FIG. 5, each traffic signal **500**, **502**, and **504** is also shown with its index in a list of traffic signal candidates and an accumulated evidence counter that is based on, for example, a number of image frames that the traffic signal candidate appears in. In the exemplary illustration of FIG. 5, traffic signal **500** has an index of 1 and an accumulated evidence counter of 87; traffic signal **502** has an index of 2 and an accumulated evidence counter of 99; and traffic signal **504** has an index of 5 and an accumulated evidence counter of 94. The index of 5 for the traffic signal **504** indicates that traffic signal candidates at indices 3 and 4 were dropped as candidates during the filtering.

[0043] Obviously, numerous modifications and variations of the present disclosure are possible in light of the above teachings. It is therefore to be understood that within the scope of the applied claims, the invention may be practiced otherwise than as specifically described herein.

1-20. (canceled)

21. A controller for detecting traffic signaling of a mobile environment, comprising:

processing circuitry configured to:

receive image frames captured by an imaging device to detect a traffic signal candidate therefrom based on at least a vehicle location, independently of a size and a shape of a traffic light housing associated with the traffic signal candidate;

establish a score having a predefined value for the traffic signal candidate and track the traffic signal candidate by incrementing the score when the traffic signal candidate is determined to appear, based on a probabilistic estima-

tor that indicates an expected value for a location of the traffic signal candidate, in a tracked image frame of the captured image frames, and decrementing the score when the traffic signal candidate is determined not to appear, based on the probabilistic estimator, in the tracked image frame; and

determine whether the traffic signal candidate is a traffic signal based on whether the score is above a specified threshold.

22. The controller according to claim **21**, wherein the circuitry is further configured to determine that the traffic signal candidate is not a traffic signal when the traffic signal candidate appears in a specified number of consecutive image frames of the captured image frames or less, the specified number corresponding to the specified threshold.

23. The controller according to claim **21**, wherein the circuitry is further configured to project a visual indicator signal that specifies a location of the traffic signal on a vehicle windshield.

24. The controller according to claim **21**, wherein the probabilistic estimator is implemented by one of a Kalman filter, an iterated Kalman filter, an extended Kalman filter, and a sigma-point filter.

25. The controller according to claim **21**, wherein the probabilistic estimator takes as an input an ego-vehicle motion to determine the expected value for the location of the traffic signal candidate in the tracked image frame.

26. The controller according to claim **21**, wherein the traffic signal candidate is determined to be the traffic signal independently of whether the traffic signal changes color during tracking with respect to the captured image frames.

27. The controller according to claim **21**, wherein the circuitry is further configured to instantiate a new probabilistic estimator for each time a new traffic signal candidate is detected.

28. A method of detecting automotive traffic signaling in a mobile environment, comprising:

receiving image frames, captured by an imaging device, to detect a traffic signal candidate therefrom based on at least a location of a vehicle independently of a size and a shape of a traffic light housing associated with the traffic signal candidate;

establishing a score having a predefined value for the traffic signal candidate and tracking, by processing circuitry, the traffic signal candidate by incrementing the score when the traffic signal candidate is determined to appear, based on a probabilistic estimator that indicates an expected value for a location of the traffic signal candidate, in a tracked image frame of the captured image frames, and decrementing the score when the traffic signal candidate is determined not to appear, based on the probabilistic estimator, in the tracked image frame; and

determining, by the processing circuitry, whether the traffic signal candidate is a traffic signal based on whether the score is above a specified threshold.

29. The method according to claim **28**, wherein the traffic signal candidate is determined to be the traffic signal when the traffic signal candidate is occluded in at least one tracked image frame of the captured image frames.

30. The method according to claim **28**, further comprising: controlling, by the processing circuitry, vehicular movement including accelerating, decelerating, and braking, when it is determined that the traffic signal candidate is the traffic signal.

31. A non-transitory computer-readable storage medium having computer readable program codes embodied in the computer readable storage medium for detecting automotive traffic signaling in a mobile environment, the computer readable program codes that, when executed, cause a computer to execute:

receiving image frames, captured by an imaging device, to detect a traffic signal candidate therefrom based on at least a location of a vehicle independently of a size and a shape of a traffic light housing associated with the traffic signal candidate;

establishing a score having a predefined value for the traffic signal candidate and tracking the traffic signal candidate by incrementing the score when the traffic signal candidate is determined to appear, based on a probabilistic estimator that indicates an expected value for a location

of the traffic signal candidate, in a tracked image frame of the captured image frames, and decrementing the score when the traffic signal candidate is determined not to appear, based on the probabilistic estimator, in the tracked image frame; and

determining whether the traffic signal candidate is a traffic signal based on whether the score is above a specified threshold.

32. The non-transitory computer-readable storage medium according to claim **31**, wherein the traffic signal candidate changes shape in at least one of the captured image frames based on an orientation of the imaging device.

33. The non-transitory computer-readable storage medium according to claim **31**, wherein the traffic signal candidate changes shape in at least one of the captured image frames based on a viewing angle of the imaging device.

34. The non-transitory computer-readable storage medium according to claim **31**, further comprising: outputting at least one of a patterned and textual visual indicator specifying the traffic signal to a vehicle-mounted display.

* * * * *