

(19)



(11)

EP 3 750 776 B1

(12)

EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention of the grant of the patent:
24.08.2022 Bulletin 2022/34

(51) International Patent Classification (IPC):
B61L 23/04 ^(2006.01) **B61L 3/00** ^(2006.01)
G06T 3/00 ^(2006.01)

(21) Application number: **19179792.7**

(52) Cooperative Patent Classification (CPC):
B61L 23/041; B61L 3/006

(22) Date of filing: **12.06.2019**

(54) METHOD AND SYSTEM FOR DETECTING A RAILROAD SIGNAL

VERFAHREN UND SYSTEM ZUM ERFASSEN EINES EISENBAHNSIGNALS

PROCÉDÉ ET SYSTÈME DE DÉTECTION D'UN SIGNAL DE CHEMIN DE FER

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

(74) Representative: **Wildhack & Jellinek**
Patentanwälte OG
Landstraßer Hauptstraße 50
1030 Wien (AT)

(43) Date of publication of application:
16.12.2020 Bulletin 2020/51

(56) References cited:
EP-A1- 3 048 559 US-A1- 2018 057 030

(73) Proprietor: **Mission Embedded GmbH**
1100 Wien (AT)

• **SINA AMINMANSOUR: "Video Analytics for the Detection of Near-miss Incidents at Railway Level Crossings and Signal Passed at Danger Events", QUEENSLAND UNIVERSITY OF TECHNOLOGY, [Online] 10 October 2017 (2017-10-10), XP055640309, https://eprints.qut.edu.au/112765/1/Sina_Aminmansour_Thesis.pdf**

(72) Inventors:
• **Reisner, Clemens**
1050 Wien (AT)
• **Sarwar, Omair**
1130 Wien (AT)
• **Kreilmeier, Michael**
1050 Wien (AT)

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

EP 3 750 776 B1

Description

[0001] The invention relates to a method and a system for the detection of railway signals relevant for the rail track that is currently used by a train.

[0002] More particularly, the invention relates to a method and system for automatically detecting, whether a signal is present and relevant for the rail track that is actually used by a train. If needed, determination of the actual state of the signal, e.g. a stop (typically red), a go (typically green), or another state, is possible after the relevant signal has been detected.

[0003] For the purposes of this application the term "signal" is defined and shall be understood as railway signal.

[0004] The overall method can be used to assist a train driver to find the relevant signal for the respective train and to avoid human failure by misinterpretation of signals. Such misinterpretation can occur in particular on railway networks having many signals for different rail tracks that are visible at the same time. Alternatively, the method can also be used to provide derived and relevant signal information for automatic train operation.

[0005] When driving a train through a region of a railway network with multiple parallel rail tracks there is an increased probability that the driver of the train wrongly identifies a signal, which is relevant for the current rail track of the train. Such misinterpretation may lead to unnecessary stops of the train, or emergency brake application or even accidents when e.g. a stop signal is ignored.

[0006] The identification of the relative position of the signal with respect to the rail track as well as the determination of the state of the signal is of high importance also for automatic train operation.

[0007] For this reason, one objective of the invention is to automatically detect the signals associated with the actual rail track currently being used by the train, in particular in order to assist the driver or for automatic train operation.

[0008] Currently, in the field of the invention there are many different types of methods available to provide information about rail tracks using three-dimensional images of the scene surrounding the train. Such methods have the drawback that they require, for instance, stereo imaging sensors; using such sensors further implies a complex and thus error-prone installation and calibration process to be carried out during assembly. In addition, three-dimensional sensor systems lose their benefits as the distance of the depicted objects from the sensors increases beyond a specific limit; however, being beyond such limits might be the case for practical application of the invention.

[0009] For the use of the invention, however, it is sufficient to have a standard two-dimensional camera, e.g. an area scan camera as image sensor, which produces two-dimensional images. Such image sensors are widely available, easy to assemble and in general simpler and more reliable than other systems for three-dimensional environmental perception, e.g. LIDAR sensors.

[0010] Sina Aminmansour: "Video Analytics for the Detection of Nearmiss Incidents at Railway Level Crossings and Signal Passed at Danger Events", Queensland University of Technology, 10 October 2017, available at:

https://eprints.qut.edu.au/112765/1/Sina_Aminmansour_Thesis.pdf (last accessed 18 March 2021) discloses a method and a system for detecting railway signals associated with rail tracks currently used by a train based on two dimensional images taken by a camera installed on the train. The rail track currently used by the train and signals are identified in the image and the placement of the signals with respect to the rail tracks is identified. Based on the information, if a signal is placed left or right of the rail tracks it is decided, if the signal belongs to the relevant track of the train.

[0011] It is therefore the objective of the invention to provide a simple and robust method for the identification of signals that are associated with the rail track currently used by a train, in particular only based on an image generated by a monocular image sensor.

[0012] The invention solves this problem using a method of claim 1 for detecting a signal, which is associated with the rail track that is used by a train; this method comprises the following steps:

- a) taking at least one two-dimensional image from the train with an image sensor being located in or on the train generating an image of the on-coming rail tracks and signals, wherein the image sensor is mounted in or on the train in a manner that its y-axis is aligned substantially vertically, and its x-axis is aligned substantially horizontally,
- b) using a classification method for identifying image regions corresponding to rail tracks and image regions corresponding to signals within the image,
- c) among said identified image regions corresponding to rail tracks, selecting the image region, which corresponds to the rail track that is used by the train, in particular by identifying the rail track
 - going through a predefined region of the image, and/or
 - having the maximum area or pixel size among the rail tracks identified in step b),
- d) for each of the image regions corresponding to a signal:

- estimating projected points of the signals and projected path of the selected rail track into a common projection plane, *in particular a ground plane*,

wherein the projected point of the signal is estimated as the *position of a ground point of the respective signal within the image*,

- wherein the ground point is the point within the image that depicts a surface point, which is directly underneath the respective signal, and
- wherein the height and the width of the panels of the respective signal (S) are known or estimated by the shape and/or content of the image region of the signal, wherein determination of the ground point of the signal (S) comprises the following steps:
- determining a pixel height h_t and a pixel width w_t of the signal's image region as follows:

- counting a number of pixels A corresponding to the signal's image region
- determining the pixel-width w_t of the signal's image region by using the relation:

$$A = w_t * h_t = w_t * w_t * r$$

- determining the pixel height h_t of said signal's image region using a known ratio r between the height and width of the panel of the signal,
- calculating a y-distance d_t from lower edge of the signal's image region to the respective ground point from the pixel height h_t which corresponds to the known height of the signal by using the relation:

$$d_t = \text{mounting height} * h_t / \text{signal height}$$

- defining the coordinates of the ground point based on the lower edge of the signal's image region as well as a horizontal center of the signal's image region as:

$$G_t = [x_{m,t}; y_{l,t} - d_t]$$

- wherein the lower edge position $y_{l,t}$ of the signal's image region is defined as the highest or least y-position, depending on the orientation of the y-axis of the image sensor, and
- wherein the horizontal center $x_{m,t}$ of the signal's image region (S_t) is defined as the x-coordinate of the centroid of the signal's image region,

e) estimating a perspective transformation to transform the image into the common projection plane, *preferably into a bird's eye view image*, based on visual properties of the image and/or mounting properties of the image sensor,
f) applying the perspective transformation

- to the selected image region or parts thereof, in particular a central path within the rail track, and
- to the projected points, and

g) determining one or more projected points that have minimum distance and/or a distance lower than a predefined threshold distance and/or relative placement to the selected image region under said perspective transformation and identifying the signals and/or image regions corresponding to signals associated with these projected points within the image. The criteria for determining said association may be governed by regulatory constraints and/or constraints defined by the railway network operator.

[0013] The invention further solves this problem using a system of claim 3.

[0014] Such a system is designed to detect a railway signal being associated with the rail track that is used by a train. The system comprises:

- an image sensor mounted or mountable in or on the train that is arranged to take at least one two-dimensional image

of the on-coming rail tracks and signals, wherein the image sensor is mounted in or on the train in a manner that its y-axis is aligned substantially vertically, and its x-axis is aligned substantially horizontally, and

- a processing unit that is connected to the image sensor that is programmed to execute the following steps:

- 5 b) using a classification method for identifying image regions corresponding to rail tracks and image regions corresponding to signals within the image,
- c) among said identified image regions corresponding to rail tracks, selecting the image region, which corresponds to the rail track that is used by the train, in particular by identifying the rail track

- 10 - going through a predefined region of the image, and/or
- having the maximum area or pixel size among the rail tracks identified in step b),

d) for each of the image regions corresponding to a signal:

- 15 - estimating projected points of the signals and projected path of the selected rail track into a common projection plane, *in particular a ground plane, wherein the projected point of the signal is estimated as the position of a ground point of the respective signal within the image,*
- wherein the ground point is the point within the image that depicts a surface point, which is directly underneath the respective signal, and
- 20 - wherein the height and the width of the panels of the respective signal are known or estimated by the shape and/or content of the image region of the signal,

wherein determination of the ground point of the signal (S) comprises the following steps:

- 25 - determining a pixel height h_t and pixel width w_t of the signal's image region as follows:
 - counting a number of pixels A corresponding to the signal's image region
 - determining the pixel-width w_t of the signal's image region (S_t) by using the relation:

- 30
$$A = w_t * h_t = w_t * w_t * r$$
- determining the pixel height h_t of said signal's image region using a known ratio r between the height and width of the panel of the signal,
- 35 - calculating a y-distance d_t from lower edge of the signal's image region to the respective ground point from the pixel height h_t which corresponds to the known height of the signal (S) by using the relation:

$$40 \quad d_t = \text{mounting height} * h_t / \text{signal height}$$

- defining the coordinates of the ground point (G_t) based on the lower edge of the signal's image region as well as a horizontal center of the signal's image region as:

$$45 \quad G_t = [x_{m,t}; y_{1,t} - d_t]$$

- 50 - wherein the lower edge position $y_{1,t}$ of the signal's image region is defined as the highest or least y-position, depending on the orientation of the y-axis of the image sensor, and
- wherein the horizontal center $x_{m,t}$ of the signal's image region is defined as the x-coordinate of the centroid of the signal's image region,

- 55 e) estimating a perspective transformation to transform the image into the common projection plane, *preferably into a bird's eye view image*, based on visual properties of the image and/or mounting properties of the image sensor,

f) applying the perspective transformation

- to the selected image region or parts thereof, *in particular a central path within the rail track*, and
- to the projected points, and

g) determining one or more projected points that have minimum distance and/or a distance lower than a pre-defined threshold distance and/or relative placement to the selected image region under said perspective transformation and identifying the signals and/or image regions corresponding to signals associated with these projected points within the image.

[0015] The criteria for determining said association may be governed by regulatory constraints and/or constraints defined by the railway network operator.

[0016] The invention can be used when a train is on a railtrack. Signals which are assigned to or associated with the respective rail track are typically placed at predefined positions with respect to the rail track. These positions are usually determined by the rules of a signal regulation and/or the operator of the railway network. For example, these predefined positions of the signals may be on the right-hand side of said rail track. Therefore, the respective states of these signals being located on this predefined position with respect to the rail track are valid for the respective train.

[0017] Additionally, signals, which are not placed at predefined positions and are not matching other association criteria as defined in the governing regulations for a given rail track, are not relevant for trains on this rail track. By using the invention human failure is avoided or minimized. In particular, situations in which the driver ignores a signal relevant to the train, or in which the driver incorrectly considers the signal of another rail track relevant, can be avoided. In addition, the derived and relevant signal information from the image region containing the signal is of particular use for the automatic train operation.

[0018] The following steps can be used to further avoid false positive detections of signals that are not relevant for the actual rail track. These steps comprise:

- providing a map of the railway network including the positions of rail tracks and signals, wherein each signal is associated with a rail track, and
- localizing the train's position on the railway map and in particular the rail track it is using,
- using the map for determining the position of the on-coming signal associated with the rail track it is using, and
- only if the on-coming signal is within a predetermined region in front of the train taking an image and carrying out steps a) to g) of claim 1.

[0019] For the same reason, the system according to the invention can be improved by a geolocation unit connected to the processing unit, wherein

- the processing unit further comprises a memory for a map of the railway network including the positions of rail tracks and signals, wherein each signal is associated with a rail track, and
- the processing unit is further programmed to execute the following steps:
 - finding the position of the train provided by the geolocation unit within the railway map *and in particular the rail track it is using*,
 - using the map for determining the position of the on-coming signal associated with the rail track it is using, and
- only if the on-coming signal is within a predetermined region in front of the train taking an image and executing steps b) to g) of claim 3.

[0020] The following detailed description of preferred embodiments of the invention is merely exemplary in nature and is not intended to limit the invention or the application and uses of the invention. The invention will be described with reference to the following figures which are provided by way of explanation only. The scope of protection is defined by the claims.

Fig. 1 depicts a typical image taken from a train.

Fig. 2 schematically shows the detection of rail tracks and signals.

Fig. 3 schematically shows an example for the calculation of the position of the ground points of the signals within the image.

Fig. 4 schematically depicts a perspectival image transformation of the image of Fig. 1 to a bird-eye's view image.

Fig. 5 shows a map as used in an advantageous second embodiment of the invention using geolocation information of the train and the signals.

[0021] In a first step (a) an image sensor, which is located in or on the train and which is at least partly directed in the driving direction of the train, i.e. the direction of travel or movement of the train, takes two-dimensional images I_t of the surrounding area, in particular of the rail track and of the signals. Such image sensors can be simply mounted in the front part of the train.

[0022] One example of such an image I_t is shown in **Fig. 1**. The image contains rail track-regions $T_1, \dots, T_4$ each of which depicts one rail track. Moreover, the image I_t contains signal-regions $S_1, \dots, S_4$, each of which shows one signal.

[0023] The image of the image sensor is transmitted to a processing unit, which is used to further execute the following steps (b) to (g).

[0024] In a second step (b) a classification method is used in order to automatically detect the rail track-regions $T_1, \dots, T_4$ and the signal-regions $S_1, \dots, S_4$. Railtracks can be detected either using classical approaches, e.g. shown in Rui Fan and Naim Dahmoun "Real-Time Stereo Vision-Based Lane Detection System", 2018 or Han Ma et al. "Multiple Lane Detection Algorithm Based on Optimised Dense Disparity Map Estimation", 2018. Classical approaches usually first extract the edges or lines from the rail track-regions and then apply curve fitting to refine the railtracks.

[0025] Alternatively, it is also possible to use machine learning algorithms, which e.g. train an end-to-end neural network to segment the railtracks from the background; such methods are known from the state of the art, in particular from Davy Neven, Bert De Brabandere, Stamatios Georgoulis, Marc Proesmans and Luc Van Gool "Towards End-to-End Lane Detection: an Instance Segmentation Approach", 2018 or Min Bai et al. "Deep Multi-Sensor Lane Detection", 2018.

[0026] In both cases a classification method is used in order to identify the regions T_t corresponding to rail tracks and the regions S_t corresponding to signals in the image I_t taken by the image sensor. However, from this result it is still not possible to determine, which signals are relevant for the driver or for the rail track that is actually used.

[0027] In order to determine the relevant signals, the position of the actual rail track within the image I_t is determined. Therefore, the regions of the current rail track, i.e. the rail track T_C that is actually used by the train, is identified.

[0028] The rail track T_C used by the train and the respective region of the rail track T_C within the image I_t can be found by searching one of the rail track regions T_t containing rail tracks already identified in step (b) that goes through or overlaps with a predefined image area A_p as shown in **Fig. 2**. This image area A_p usually depends on the orientation of the image sensor within the train. If an image sensor is heading straight in the direction of movement, the predefined image area A_p is in the center of the lower part of the image I_t . The image area A_p can be easily defined after mounting the image sensor in or on the train by searching for a region of the image, which is close to the train and which shows a part of the rail track, independently of the curvature of the rail track.

[0029] Alternatively, it is also possible to search for the regions T_t having the maximum area or containing maximum pixels among the rail tracks identified in step (b). As the rail track currently used is close to the image sensor, it typically covers a bigger region within the image I_t than the other rail track regions.

[0030] In this preferred embodiment of the invention the rail track identified in step (c) is represented by a center path P_C .

[0031] In order to determine the signals relevant for the identified rail track T_C in step (d) the positions of the ground points $G_1, \dots, G_4$ of the signals S_T are estimated with respect to the image I_t . As many signals are mounted on poles the ground point $G_1, \dots, G_4$ is the point within the image that shows the base point of the pole. More generally, the ground point is the point within the image I_t that depicts a surface point, which is directly underneath the respective signal S , even if the signal is not mounted on a pole, but e.g. mounted on overhead infrastructure.

[0032] In order to determine the ground point $G_1, \dots, G_4$, the following prerequisites are typically needed. Firstly, the size, i.e. height and width, of the panels of the signals are known. If there are different types of signals available, said type has to be firstly estimated by the shape and/or content of the image region of the signal.

[0033] For instance, a typical kind of signal could have a rectangular panel of 1.4-meters height and 0.7-meters width. Moreover, the signal could typically be mounted in a height of 4.85-meters to 6.95-meters (see **Fig. 3**).

[0034] Moreover, it is assumed that the image sensor is mounted on the train in a manner that its y-axis is - at least approximately - aligned vertically, and its x-axis is - at least approximately - aligned horizontally.

[0035] The height h_t and width w_t of the image region of the signal can be obtained by different methods, such as detection of an axis-parallel bounding box, etc.

[0036] The pixel height h_t and width w_t of the signal's image region S_t is determined as follows: As already mentioned, the ratio r between the height and width of the panel of the signal are known initially, for instance in the above case 2:1. In the following, the number of pixels A corresponding to the signal's image region S_t is counted. By using the relation $A = w_t \cdot h_t = w_t \cdot w_t \cdot r$ the pixel-width w_t of the image region S_t of the signal can be determined easily. Accordingly, the pixel height h_t of said signal region S_t can be determined.

[0037] As the pixel height h_t corresponds to the known height h_t of the signal, one can easily calculate the y-distance d_t from the lower edge of an image region S_t to the respective ground point G_t . Given that the image region S_t has a pixel-height of 56 px as well as the above specifications for signals the y-distance d_t between the lower edge of the image region S_t and the ground point G_t is calculated as:

$$d_t = [4.85, 6.95] \text{ m} * 56 \text{ px} / 1.4 \text{ m} = [194, 278] \text{ px}.$$

[0038] This means, that the ground point is at a position between 194 px and 278 px under the lower edge of the signal's image region S_t . To simplify calculations, an average pole height, for example of 6-meters can be used, which would lead to a y-distance d_t of 240 px.

[0039] In order to define the ground point's coordinates, one has to find the lower edge of the signal region as well as the horizontal center of the signal region. The lower edge position $y_{l,t}$ is defined as the highest or least y-position, depending on the orientation of the y-axis of the image sensor. The horizontal center $x_{m,t}$ of the signal region is defined as the x-coordinate of the centroid of the signal region. The position of the ground point G_t is defined as:

$$G_t = [x_{m,t}; y_{l,t} - d_t].$$

[0040] In general, the ground point G_t can be determined by calculating an offset from an image position within the region of the signal S_t , wherein the offset is defined by the size, in particular by a length on the region of the signal S_t . For example, the ground point G_t can be determined by calculating width and height of a rectangular bounding box that is parallel to the image axes.

[0041] Even if the ground points G_t were determined within the image I_t , it is still not possible to determine the signal which is closest to the actual rail track directly from the image, because the image as taken by the image sensor is distorted, i.e. distances within the image are not directly related to the distances of the objects depicted in the image.

[0042] As shown, the position of a ground point G_t underneath the signal was estimated from the image by projection of the signal within the image I_t to the ground plane.

[0043] Instead of a projection to a ground plane, it is also possible to use any other plane as projection plane. In this case the path P_c of the selected rail track and the position of the signals are projected into this plane.

[0044] For this reason, in a subsequent step (e) a perspective transformation R is determined, by which it is possible to transform the sensor image I_t into the projection plane; as this projection plane is - in this very example - parallel to the ground plane, the transformation returns a bird's eye view image I_B . The transformation R is a perspective transformation, that is based on visual properties of the image and/or mounting properties of the image sensor. One particular method to determine the transformation R is to find a trapezoidal region R_T within the region T_c of the actual rail track. Based on the mounting properties of the image sensor and the assumption of a sufficiently flat and smooth track bed it is clear, that this trapezoidal region R_T typically covers a rectangular region R_R within a bird's eye view image I_B . Even if the rail track is curved, there is a region close to the train, which is sufficiently straight, to obtain an approximate trapezoid structure. The perspective transformation is defined in such way, that it transforms at least a trapezoidal part of the region T_c of the actual rail track to a rectangular region within the bird's eye view image I_B .

[0045] Even if it is not required to transform the overall image I_t into the bird's eye view image I_B , said image I_B is shown for the sake of clarity in **Fig. 4**. In order to determine the signals that are assigned to the actual rail track that is used by the train, the perspective transformation R is in a further step (f) applied to the selected rail track, in particular to a central path $P_1, \dots, P_4$ within the rail track, and to the ground points that were determined in step (d). From these data it is possible to determine real-world distances, or it is at least possible to compare distances of ground points to the actual rail track.

[0046] In order to find the relevant signals, which are associated with the used rail track, one can detect the ground point G_t that has the least distance compared to the other ground points G_t under said perspective transformation R . Alternatively, or in addition, a ground point G_t and its respective signal S_t should be associated with the actual rail track only if the distance from the ground point G_t to the region, in particular the path P_c , of the selected rail track S_t under said perspective transformation R is below a predefined threshold distance.

[0047] In general, it is also possible, that a regulation sets out more complicated rules relating to the association of signals with rail tracks. These regulations can rely on the determined distances and/or relative placement, e.g. right or left from the rail track, only. However, it is also possible that there are more complex rules relating to the association of rail tracks and signals that are based on a sequence of detections or the position of the train and the signal in the railway network. These rules may rely on distances, relative placement and sequences of signals on the rail track.

[0048] An advanced **second embodiment** of the invention uses the method as described above in order to determine images of signals that are associated with the actual rail track. In addition, this embodiment also relies on a geolocation-map, containing the location of the rail tracks $t_A, \dots, t_G$ as well as the position $S_A, \dots, S_L$ of the signals (**Fig. 5**).

[0049] For this second embodiment a geolocation-unit is used, the geolocation-unit being connected or integrated to the processing unit. The geolocation-unit provides the actual geolocation of the train which is further processed by the processing unit as follows.

[0050] When the train is on the rail track t_c , in particular moves on the rail track, its geographical positions (depicted

by dots in **Fig. 5**) is determined using geolocation methods, such as Global Navigation Satellite System (GNSS) or Simultaneous Localization And Mapping (SLAM). As a result, the actual position of the train, which is specified by coordinates, is known at each time. Using said coordinates as well as the map containing the location information of the railway network, it is possible to determine the actual rail track of the train independently from the visual method shown in the first embodiment of the invention. The map of the railway network is preferably stored on a memory connected to the processing system. However, in order to precisely localize the signal as well as the state of the signal the invention of the first embodiment is used. Once the train starts its journey towards its destination, its route or driving direction DD, depicted by an arrow in **Fig. 5**, is already defined and tracked in real-time using the geolocation unit.

[0051] If the map of rail tracks, signals and the driving direction DD are known, it is possible to determine the positions $S_A, \dots, S_L$ of signals associated with the respective rail track. Therefore, it is also possible to determine, if - within a predefined region $E_1, \dots, E_3$ calculated from the actual position of the train $e_1, \dots, e_3$ - a signal is expected to be seen in the image. Such a signal is also referred to as on-coming signal.

[0052] The method for detecting image regions of signals, which are associated with the actual rail track as shown in the first embodiment of the invention only needs to be carried out, if such a signal is expected based on the geolocation method as explained above. If, such as for the position e_3 , there is no on-coming signal within the associated region E_3 , the visual search for signals is turned off in order to reduce the likelihood for false positive matches.

[0053] For the position e_1 the associated region E_1 contains many ground points S_F, S_G, S_I, S_J . However, none of these ground points are associated with the actual rail track t_C so that there is no on-coming signal; the method for detecting image regions of signals is turned off.

[0054] As the position e_2 and its associated region E_2 contain a signal S_D that is associated with the actual rail track t_C , it is to be expected, that there is an on-coming signal S_D . Therefore, the method for detecting image regions of signals is carried out in order to find a visual representation of said signal in the image I_t .

[0055] Alternatively, or in addition, the information relating to the map positions of signals can be used in order to further avoid false positive detections of signals with respect to both further described embodiments of the invention. A matching algorithm is used in order to match the rail tracks $t_A, \dots, t_G$ of the map to the rail tracks as contained in bird's eye view image I_B . If a selected ground point G_i matches the geolocation of a signal of the predefined map, which is associated with the actual path as determined by the geolocation, the overall reliability of the result will be increased.

[0056] For all preferred embodiments of the invention it is possible to display the image region containing an image of the identified signal to the train driver. Alternatively, it is also possible to highlight said image region when showing the overall image to the driver. Typically, the procedure as described above is repeated regularly so as to provide a video in order to inform the train driver about relevant on-coming signals.

[0057] Alternatively, or in addition, any other event can be triggered by the information automatically retrieved from the image such as simple audible, visual, or haptic warning.

[0058] Of course, it is also possible to feed the image data of the identified image region of the relevant signal to an image processing algorithm in order to extract the state of the signal. This state can be used to assist the train driver, e.g. to brake the train accordingly, when a stop signal is identified or even for automatic train operation.

Claims

1. Method for detecting a railway signal, which is associated with the rail track that is used by a train,

a) taking at least one two-dimensional image (I_t) from the train with an image sensor being located in or on the train generating an image of the on-coming rail tracks and signals, wherein the image sensor is mounted in or on the train in a manner that its y-axis is aligned substantially vertically, and its x-axis is aligned substantially horizontally,

b) using a classification method for identifying image regions (T_i) corresponding to rail tracks and image regions (S_i) corresponding to signals within the image (I_t),

c) among said identified image regions (T_i) corresponding to rail tracks, selecting the image region (T_C), which corresponds to the rail track that is used by the train, in particular by identifying the rail track

- going through a predefined region of the image, and/or
- having the maximum area or pixel size among the rail tracks identified in step b),

d) for each of the image regions (S_i) corresponding to a signal (S):

- estimating projected points of the signals and projected path of the selected rail track (T_C) into a common projection plane, *in particular a ground plane*, wherein the projected point of the signal is estimated as the

position of a ground point (G_t) of the respective signal within the image (I_t),

- wherein the ground point (G_t) is the point within the image (I_t) that depicts a surface point, which is directly underneath the respective signal (S), and

- wherein the height and the width of the panels of the respective signal (S) are known or estimated by the shape and/or content of the image region (S_t) of the signal (S), wherein determination of the ground point (G_t) of the signal (S) comprises the following steps:

- determining a pixel height h_t and pixel width w_t of the signal's image region (S_t) as follows:

- counting a number of pixels A corresponding to the signal's image region (S_t)

- determining the pixel-width w_t of the signal's image region (S_t) by using the relation:

$$A = w_t * h_t = w_t * w_t * r$$

- determining the pixel height h_t of said signal's image region (S_t) using a known ratio r between the height and width of the panel of the signal (S),

- calculating a y-distance d_t from a lower edge of the signal's image region (S_t) to the respective ground point (G_t) from the pixel height h_t which corresponds to the known height of the signal (S) by using the relation:

$$d_t = \text{mounting height} * h_t / \text{signal height}$$

- defining the coordinates of the ground point (G_t) based on the lower edge of the signal's image region (S_t) as well as a horizontal center of the signal's image region (S_t) as:

$$G_t = [x_{m,t}; y_{1,t} - d_t]$$

- wherein the lower edge position $y_{1,t}$ of the signal's image region (S_t) is defined as the highest or least y-position, depending on the orientation of the y-axis of the image sensor, and

- wherein the horizontal center $x_{m,t}$ of the signal's image region (S_t) is defined as the x-coordinate of the centroid of the signal's image region (S_t),

e) estimating a perspective transformation (R) to transform the image (I_t) into the common projection plane, preferably into a bird's eye view image (I_B), based on visual properties of the image and/or mounting properties of the image sensor,

f) applying the perspective transformation

- to the selected image region (T_C) or parts thereof, in particular a central path within the rail track, and
- to the projected points, and

g) determining one or more projected points (G_t) that have minimum distance and/or a distance lower than a predefined threshold distance and/or relative placement to the selected image region (T_C) under said perspective transformation (R) and identifying the signals and/or image regions (S_t) corresponding to signals associated with these projected points within the image (I_t).

2. Method according to claim 1, characterized by the following steps:

- providing a map of the railway network including the positions of rail tracks and signals, wherein each signal is associated with a rail track,

- localizing the train's position on the railway map and the rail track it is using,

- using the map for determining the position of the on-coming signal associated with the rail track it is using,

- only if the on-coming signal is within a predetermined region in front of the train taking an image and carrying out the method of claim 1.

3. System for detecting a railway signal, which is associated with the rail track that is used by a train, comprising

- an image sensor mounted or mountable in or on the train that is arranged to take at least one two-dimensional image (I_t) of the on-coming rail tracks and signals, wherein the image sensor is mounted in or on the train in a manner that its y-axis is aligned substantially vertically, and its x-axis is aligned substantially horizontally, and
- a processing unit that is connected to the image sensor that is programmed to execute the following steps:

b) using a classification method for identifying image regions (T_t) corresponding to rail tracks and image regions (S_t) corresponding to signals within the image (I_t),

c) among said identified image regions (T_t) corresponding to rail tracks, selecting the image region (T_C), which corresponds to the rail track that is used by the train, in particular by identifying the rail track

- going through a predefined region of the image, and/or
- having the maximum area or pixel size among the rail tracks identified in step b),

d) for each of the image regions (S_t) corresponding to a signal (S):

- estimating projected points of the signals and projected path of the selected rail track (T_C) into a common projection plane, *in particular a ground plane, wherein the projected point of the signal is estimated as the position of a ground point (G_t) of the respective signal within the image (I_t),*

- wherein the ground point (G_t) is the point within the image (I_t) that depicts a surface point, which is directly underneath the respective signal (S), and

- wherein the height and the width of the panels of the respective signal (S) are known or estimated by the shape and/or content of the image region (S_t) of the signal (S), wherein determination of the ground point (G_t) of the signal (S) comprises the following steps:

- determining a pixel height h_t and a pixel width w_t of the signal's image region (S_t) as follows:

- counting a number of pixels A corresponding to the signal's image region (S_t)
- determining the pixel-width w_t of the signal's image region (S_t) by using the relation:

$$A = w_t * h_t = w_t * w_t * r$$

- determining the pixel height h_t of said signal's image region (S_t) using a known ratio r between the height and width of the panel of the signal (S),

- calculating a y-distance d_t from a lower edge of the signal's image region (S_t) to the respective ground point (G_t) from the pixel height h_t which corresponds to the known height of the signal (S) by using the relation:

$$d_t = \text{mounting height} * h_t / \text{signal height}$$

- defining the coordinates of the ground point (G_t) based on the lower edge of the signal's image region (S_t) as well as a horizontal center of the signal's image region (S_t) as:

$$G_t = [x_{m,t}; y_{1,t} - d_t]$$

- wherein the lower edge position $y_{1,t}$ of the signal's image region (S_t) is defined as the highest or least y-position, depending on the orientation of the y-axis of the image sensor, and

- wherein the horizontal center $x_{m,t}$ of the signal's image region (S_t) is defined as the x-coordinate of the centroid of the signal's image region (S_t),

e) estimating a perspective transformation (R) to transform the image (I_t) into the common projection plane,

preferably into a bird's eye view image (I_B), based on visual properties of the image and/or mounting properties of the image sensor,
f) applying the perspective transformation

- to the selected image region (T_C) or parts thereof, *in particular a central path within the rail track*, and
- to the projected points, and

g) determining one or more projected points (G_i) that have minimum distance and/or a distance lower than a predefined threshold distance and/or relative placement to the selected image region (T_C) under said perspective transformation (R) and identifying the signals and/or image regions (S_i) corresponding to signals associated with these projected points within the image (I_i).

4. System according to claim 3, further **characterized by** a geolocation unit connected to the processing unit, wherein

- the processing unit further comprises a memory for a map of the railway network including the positions of rail tracks and signals, wherein each signal is associated with a rail track, and
- the processing unit is further programmed to execute the following steps:

- finding the position of the train provided by the geolocation unit within the railway map *and the rail track it is using*,
- using the map for determining the position of the on-coming signal associated with the rail track it is using, and
- only if the on-coming signal is within a predetermined region in front of the train taking an image and executing steps b) to g) of claim 3.

Patentansprüche

1. Verfahren zum Erfassen eines Eisenbahnsignals, das mit der Fahrbahn verknüpft ist, die von einem Zug verwendet wird,

a) Aufnehmen mindestens eines zweidimensionalen Bildes (I_i) des Zugs mit einem Bildsensor, der sich im oder am Zug befindet, das ein Bild der entgegenkommenden Fahrbahnen und Signale erzeugt, wobei der Bildsensor derart im oder am Zug montiert ist, dass seine y-Achse im Wesentlichen vertikal ausgerichtet ist und seine x-Achse im Wesentlichen horizontal ausgerichtet ist,

b) Verwenden eines Klassifizierungsverfahrens zum Identifizieren, innerhalb des Bildes (I_i), von Bildbereichen (T_i), die Fahrbahnen entsprechen, und Bildbereichen (S_i), die Signalen entsprechen,

c) aus den identifizierten Bildbereichen (T_i), die Fahrbahnen entsprechen, Auswählen des Bildbereichs (T_C), der der Fahrbahn entspricht, die von dem Zug verwendet wird, insbesondere durch Identifizieren der Fahrbahn,

- die durch einen vorab definierten Bereich des Bildes führt, und/oder
- die maximale Fläche oder Pixelgröße aus den in Schritt b) identifizierten Fahrbahnen aufweist,

d) für jeden der Bildbereiche (S_i), die einem Signal (S) entsprechen:

- Schätzen von projizierten Punkten der Signale und projiziertem Weg der ausgewählten Fahrbahn (T_C) in eine gemeinsame Projektionsebene, insbesondere eine Bodenebene, wobei der projizierte Punkt des Signals als die Position eines Bodenpunkts (G_i) des entsprechenden Signals innerhalb des Bildes (I_i) geschätzt wird,

- wobei der Bodenpunkt (G_i) der Punkt innerhalb des Bildes (I_i) ist, der einen Oberflächenpunkt abbildet, der direkt unterhalb des entsprechenden Signals (S) ist, und
- wobei die Höhe und die Breite der Tafeln des entsprechenden Signals (S) bekannt sind oder anhand der Form und/oder des Inhalts des Bildbereichs (S_i) des Signals (S) geschätzt werden, wobei die Bestimmung des Bodenpunkts (G_i) des Signals (S) die folgenden Schritte umfasst:

- Bestimmen einer Pixelhöhe h_i und einer Pixelbreite w_i des Bildbereichs (S_i) des Signals wie folgt:

EP 3 750 776 B1

- Zählen einer Anzahl von Pixeln A , die dem Bildbereich (S_t) des Signals entspricht
- Bestimmen der Pixelbreite w_t des Bildbereichs (S_t) des Signals unter Verwendung der Beziehung:

5

$$A = w_t * h_t = w_t * w_t * r$$

10

- Bestimmen der Pixelhöhe h_t des Bildbereichs (S_t) des Signals unter Verwendung eines bekannten Verhältnisses r zwischen der Höhe und Breite der Tafel des Signals (S),

- Berechnen eines y -Abstandes d_t von einem unteren Rand des Bildbereichs (S_t) des Signals zu dem entsprechenden Bodenpunkt (G_t) von der Pixelhöhe h_t , die der bekannten Höhe des Signals (S) entspricht, unter Verwendung der Beziehung:

15

$$d_t = \text{Anstiegshöhe} * h_t / \text{Signalhöhe}$$

- Definieren der Koordinaten des Bodenpunkts (G_t) basierend auf dem unteren Rand des Bildbereichs (S_t) des Signals sowie einem horizontalen Mittelpunkt des Bildbereichs (S_t) des Signals als:

20

$$G_t = [x_{m,t}; y_{l,t} - d_t]$$

25

- wobei die Position des unteren Rands $y_{l,t}$ des Bildbereichs (S_t) des Signals als die höchste oder niedrigste y -Position definiert ist, in Abhängigkeit von der Ausrichtung der y -Achse des Bildsensors, und
- wobei der horizontale Mittelpunkt $x_{m,t}$ des Bildbereichs (S_t) des Signals als die x -Koordinate des Schwerpunkts des Bildbereichs (S_t) des Signals definiert ist,

30

- e) Schätzen einer Perspektiventransformation (R), um das Bild (I_t) in die gemeinsame Projektionsebene zu transformieren, *vorzugsweise in eine Vogelperspektivenansicht (I_B)*, basierend auf visuellen Merkmalen des Bildes und/oder Montagemerkmalen des Bildsensors,
- f) Anwenden der Perspektiventransformation

35

- am ausgewählten Bildbereich (T_C) oder Teilen davon, insbesondere einem mittleren Weg innerhalb der Fahrbahn, und
- an den projizierten Punkten, und

40

- g) Bestimmen eines oder mehrerer projizierter Punkte (G_t), die einen minimalen Abstand oder einen Abstand, der geringer als ein vorab definierter Schwellenabstand ist, und/der eine Positionierung relativ zu dem ausgewählten Bildbereich (T_C) unter der Perspektiventransformation (R) aufweisen, und Identifizieren der Signale und/oder Bildbereiche (S_t), die Signalen entsprechen, die mit diesen projizierten Punkten innerhalb des Bildes (I_t) verknüpft sind.

45

2. Verfahren nach Anspruch 1, **gekennzeichnet durch** die folgenden Schritte:

- Bereitstellen einer Karte eines Eisenbahnnetzes, die die Positionen von Fahrbahnen und Signalen beinhaltet, wobei jedes Signal mit einer Fahrbahn verknüpft ist,
- Lokalisieren der Position eines Zuges auf der Eisenbahnkarte und der Fahrbahn, die er verwendet,
- Verwenden der Karte zum Bestimmen der Position des entgegenkommenden Signals, das mit der Fahrbahn verknüpft ist, die er verwendet,
- nur falls das entgegenkommende Signal innerhalb eines vorab bestimmten Bereichs vor dem Zug ist, Aufnehmen eines Bildes und Ausführen des Verfahrens von Anspruch 1.

55

3. System zum Erfassen eines Eisenbahnsignals, das mit der Fahrbahn verknüpft ist, die von einem Zug verwendet wird, umfassend

- einen Bildsensor, der im oder am Zug montiert ist oder montierbar ist, der eingerichtet ist, um mindestens ein

zweidimensionales Bild (I_t) der entgegenkommenden Fahrbahnen und Signale aufzunehmen, wobei der Bildsensor derart im oder am Zug montiert ist, dass seine y-Achse im Wesentlichen vertikal ausgerichtet ist und seine x-Achse im Wesentlichen horizontal ausgerichtet ist, und
 - eine Verarbeitungseinheit, die mit dem Bildsensor verbunden ist, die programmiert ist, um die folgenden Schritte auszuführen:

b) Verwenden eines Klassifizierungsverfahrens zum Identifizieren, innerhalb des Bildes (I_t), von Bildbereichen (T_t), die Fahrbahnen entsprechen, und Bildbereichen (S_t), die Signalen entsprechen,

c) aus den identifizierten Bildbereichen (T_t), die Fahrbahnen entsprechen, Auswählen des Bildbereichs (T_c), der der Fahrbahn entspricht, die von dem Zug verwendet wird, insbesondere durch Identifizieren der Fahrbahn,

- die durch einen vorab definierten Bereich des Bildes führt, und/oder
- die maximale Fläche oder Pixelgröße aus den in Schritt b) identifizierten Fahrbahnen aufweist,

d) für jeden der Bildbereiche (S_t), die einem Signal (S) entsprechen:

- Schätzen von projizierten Punkten der Signale und projiziertem Weg der ausgewählten Fahrbahn (T_c) in eine gemeinsame Projektionsebene, insbesondere eine Bodenebene, wobei der projizierte Punkt des Signals als die Position eines Bodenpunkts (G_t) des entsprechenden Signals innerhalb des Bildes (I_t) geschätzt wird,

- wobei der Bodenpunkt (G_t) der Punkt innerhalb des Bildes (I_t) ist, der einen Oberflächenpunkt abbildet, der direkt unterhalb des entsprechenden Signals (S) ist, und

- wobei die Höhe und die Breite der Tafeln des entsprechenden Signals (S) bekannt sind oder anhand der Form und/oder des Inhalts des Bildbereichs (S_t) des Signals (S) geschätzt werden, wobei die Bestimmung des Bodenpunkts (G_t) des Signals (S) die folgenden Schritte umfasst:

- Bestimmen einer Pixelhöhe h_t und einer Pixelbreite w_t des Bildbereichs (S_t) des Signals wie folgt:

- Zählen einer Anzahl von Pixeln A , die dem Bildbereich (S_t) des Signals entspricht

- Bestimmen der Pixelbreite w_t des Bildbereichs (S_t) des Signals unter Verwendung der Beziehung:

$$A = w_t \cdot h_t = w_t \cdot w_t \cdot r$$

- Bestimmen der Pixelhöhe h_t des Bildbereichs (S_t) des Signals unter Verwendung eines bekannten Verhältnisses r zwischen der Höhe und Breite der Tafel des Signals (S),

- Berechnen eines y-Abstandes d_t von einem unteren Rand des Bildbereichs (S_t) des Signals zu dem entsprechenden Bodenpunkt (G_t) von der Pixelhöhe h_t , die der bekannten Höhe des Signals (S) entspricht, unter Verwendung der Beziehung:

$$d_t = \text{Anstiegshöhe} \cdot h_t / \text{Signalhöhe}$$

- Definieren der Koordinaten des Bodenpunkts (G_t) basierend auf dem unteren Rand des Bildbereichs (S_t) des Signals sowie einem horizontalen Mittelpunkt des Bildbereichs (S_t) des Signals als:

$$G_t = [x_{m,t}; y_{l,t} - d_t]$$

- wobei die Position des unteren Rands $y_{l,t}$ des Bildbereichs (S_t) des Signals als die höchste oder niedrigste y-Position definiert ist, in Abhängigkeit von der Ausrichtung der y-Achse des Bildsensors, und

- wobei der horizontale Mittelpunkt $x_{m,t}$ des Bildbereichs (S_t) des Signals als die x-Koordinate des Schwerpunkts des Bildbereichs (S_t) des Signals definiert ist,

- e) Schätzen einer Perspektiventransformation (R), um das Bild (I_t) in die gemeinsame Projektionsebene zu transformieren, *vorzugsweise in eine Vogelperspektivenansicht (I_B)*, basierend auf visuellen Merkmalen des Bildes und/oder Montagemerkmalen des Bildsensors,
f) Anwenden der Perspektiventransformation

- am ausgewählten Bildbereich (T_C) oder Teilen davon, insbesondere einem mittleren Weg innerhalb der Fahrbahn, und
- an den projizierten Punkten, und

- g) Bestimmen eines oder mehrerer projizierter Punkte (G_t), die einen minimalen Abstand oder einen Abstand, der geringer als ein vorab definierter Schwellenabstand ist, und/der eine Positionierung relativ zu dem ausgewählten Bildbereich (T_C) unter der Perspektiventransformation (R) aufweisen, und Identifizieren der Signale und/oder Bildbereiche (S_t), die Signalen entsprechen, die mit diesen projizierten Punkten innerhalb des Bildes (I_t) verknüpft sind.

4. System nach Anspruch 3, weiter **gekennzeichnet durch** eine Geolokalisierungseinheit, die mit der Verarbeitungseinheit verbunden ist, wobei

- die Verarbeitungseinheit weiter einen Speicher für eine Karte des Eisenbahnnetzes umfasst, die die Positionen von Fahrbahnen und Signalen beinhaltet, wobei jedes Signal mit einer Fahrbahn verknüpft ist, und
- die Verarbeitungseinheit weiter programmiert ist, um die folgenden Schritte auszuführen:
- Auffinden der Position des Zugs, die von der Geolokalisierungseinheit bereitgestellt wird, innerhalb der Eisenbahnkarte und der Fahrbahn, die er verwendet,
- Verwenden der Karte zum Bestimmen der Position des entgegenkommenden Signals, das mit der Fahrbahn verknüpft ist, die er verwendet, und
- nur falls das entgegenkommende Signal innerhalb eines vorab bestimmten Bereichs vor dem Zug ist, Aufnehmen eines Bildes und Ausführen der Schritte b) bis g) von Anspruch 3.

Revendications

1. Procédé de détection d'un signal de chemin de fer associé à la voie ferrée utilisée par un train, comprenant les étapes consistant à

- a) prendre au moins une image bidimensionnelle (I_t) à partir du train avec un capteur d'image situé dans ou sur le train générant une image des voies ferrées venant en sens inverse et de signaux, dans lequel le capteur d'image est monté dans ou sur le train de manière à ce que son axe y soit aligné sensiblement verticalement, et son axe x soit aligné sensiblement horizontalement,
- b) utiliser un procédé de classification pour identifier des régions d'image (T_t) correspondant à des voies ferrées et des régions d'image (S_t) correspondant à des signaux dans l'image (I_t),
- c) parmi lesdites régions d'image identifiées (T_t) correspondant à des voies ferrées, sélectionner la région d'image (T_C) qui correspond à la voie ferrée qui est utilisée par le train, en particulier en identifiant la voie ferrée

- parcourir une région prédéfinie de l'image, et/ou
- présenter la surface ou la taille de pixel maximum parmi les voies ferrées identifiées à l'étape b),

- d) pour chacune des régions d'image (S_t) correspondant à un signal (S) :

- estimer des points projetés des signaux et la trajectoire projetée de la voie ferrée sélectionnée (T_C) dans un plan de projection commun, *en particulier un plan de masse*, dans lequel le point projeté du signal est estimé comme la position d'un point de masse (G_t) du signal respectif à l'intérieur de l'image (I_t),

- dans lequel le point de masse (G_t) est le point à l'intérieur de l'image (I_t) qui représente un point de surface, qui se trouve directement sous le signal respectif (S), et
- dans lequel la hauteur et la largeur des panneaux du signal respectif (S) sont connues ou estimées par la forme et/ou le contenu de la région d'image (S_t) du signal (S), dans lequel la détermination du point de masse (G_t) du signal (S) comprend les étapes suivantes consistant à :

- déterminer une hauteur de pixel h_t et une largeur de pixel w_t de la région d'image (S_t) du signal de la manière suivante :

- compter un nombre de pixels A correspondant à la région d'image (S_t) du signal
- déterminer la largeur de pixel w_t de la région d'image (S_t) du signal en utilisant la relation :

$$A = w_t * h_t = w_t * w_t * r$$

- déterminer la hauteur de pixel h_t de ladite région d'image (S_t) du signal en utilisant un rapport connu r entre la hauteur et la largeur du panneau du signal (S),

- calculer une distance y_{d_t} entre un bord inférieur de la région d'image (S_t) du signal et le point de masse respectif (G_t) à partir de la hauteur de pixel h_t qui correspond à la hauteur connue du signal (S) en utilisant la relation :

$$d_t = \text{hauteur de montage} * h_t / \text{hauteur de signal}$$

- définir les coordonnées du point de masse (G_t) sur la base du bord inférieur de la région d'image (S_t) du signal ainsi que d'un centre horizontal de la région d'image (S_t) du signal sous la forme :

$$G_t = [x_{m,t} ; y_{l,t} - d_t]$$

- dans lequel la position de bord inférieur $y_{l,t}$ de la région d'image (S_t) du signal est définie comme la position y la plus élevée ou la plus faible, en fonction de l'orientation de l'axe y du capteur d'image, et
- dans lequel le centre horizontal $x_{m,t}$ de la région d'image (S_t) du signal est défini comme la coordonnée x du centroïde de la région d'image (S_t) du signal,

e) estimer une transformation de perspective (R) pour transformer l'image (I_t) dans le plan de projection commun, *de préférence dans une image vue du ciel (I_B)* sur la base de propriétés visuelles de l'image et/ou de propriétés de montage du capteur d'image,

f) appliquer la transformation de perspective

- à la région d'image sélectionnée (T_c) ou à des parties de celle-ci, en particulier un trajet central à l'intérieur de la voie ferrée, et
- aux points projetés, et

g) déterminer un ou plusieurs points projetés (G_t) qui présentent une distance minimum et/ou une distance inférieure à une distance de seuil prédéfinie et/ou un placement relatif par rapport à la région d'image sélectionnée (T_c) sous ladite transformation de perspective (R) et identifier les signaux et/ ou les régions d'image (S_t) correspondant à des signaux associés à ces points projetés dans l'image (I_t)

2. Procédé selon la revendication 1, caractérisé par les étapes suivantes consistant à :

- fournir une carte du réseau de chemin de fer incluant les positions des voies ferrées et des signaux, dans lequel chaque signal étant associé à une voie ferrée,
- localiser la position du train sur la carte de chemin de fer et la voie ferrée qu'il utilise,
- utiliser la carte pour déterminer la position du signal d'approche associé à la voie ferrée qu'il utilise,
- uniquement si le signal d'approche se trouve dans une région prédéterminée devant le train, prendre une image et exécuter le procédé de la revendication 1.

3. Système de détection d'un signal ferroviaire, qui est associé à la voie ferrée utilisée par un train, comprenant

- un capteur d'image monté ou pouvant être monté dans ou sur le train qui est agencé pour prendre au moins une image bidimensionnelle (I_t) des voies ferrées et des signaux en approche, dans lequel le capteur d'image est monté dans ou sur le train d'une manière telle que son axe y est aligné sensiblement verticalement, et son

axe x est aligné sensiblement horizontalement, et

- une unité de traitement qui est connectée au capteur d'image qui est programmée pour exécuter les étapes suivantes consistant à :

- 5 b) utiliser un procédé de classification pour identifier des régions d'image (T_t) correspondant à des voies ferrées et des régions d'image (S_t) correspondant à des signaux dans l'image (I_t),
c) parmi lesdites régions d'image identifiées (T_t) correspondant à des voies ferrées, sélectionner la région d'image (T_c) qui correspond à la voie ferrée qui est utilisée par le train, en particulier en identifiant la voie ferrée

- 10 - parcourir une région prédéfinie de l'image, et/ou
- présenter la surface ou la taille de pixel maximum parmi les voies ferrées identifiées à l'étape b),

d) pour chacune des régions d'image (S_t) correspondant à un signal (S) :

- 15 - estimer des points projetés des signaux et la trajectoire projetée de la voie ferrée sélectionnée (T_c) dans un plan de projection commun, en particulier un plan de masse, dans lequel le point projeté du signal est estimé comme la position d'un point de masse (G_t) du signal respectif dans l'image (I_t),

- 20 - dans lequel le point de masse (G_t) est le point dans l'image (I_t) qui représente un point de surface, qui est directement sous le signal respectif (S), et
- dans lequel la hauteur et la largeur des panneaux du signal respectif (S) sont connues ou estimées par la forme et/ou le contenu de la région d'image (S_t) du signal (S),

25 dans lequel la détermination du point de masse (G_t) du signal (S) comprend les étapes suivantes consistant à :

- déterminer une hauteur de pixel h_t et une largeur de pixel w_t de la région d'image (S_t) du signal de la manière suivante :

- 30 - compter un nombre de pixels A correspondant à la région d'image (S_t) du signal
- déterminer la largeur de pixel w_t de la région d'image (S_t) du signal en utilisant la relation :

$$A = w_t * h_t = w_t * w_t * r$$

- 35 - déterminer la hauteur de pixel h_t de ladite région d'image (S_t) du signal en utilisant un rapport connu r entre la hauteur et la largeur du panneau du signal (S),

- 40 - calculer une distance y d_t entre un bord inférieur de la région d'image (S_t) du signal et le point de masse respectif (G_t) à partir de la hauteur de pixel h_t qui correspond à la hauteur connue du signal (S) en utilisant la relation :

$$d_t = \text{hauteur de montage} * h_t / \text{hauteur de signal}$$

- 45 - définir les coordonnées du point de masse (G_t) sur la base du bord inférieur de la région d'image (S_t) du signal ainsi que d'un centre horizontal de la région d'image (S_t) du signal sous la forme :

50
$$G_t = [x_{m,t} ; y_{l,t} - d_t]$$

- dans lequel la position de bord inférieur $y_{l,t}$ de la région d'image (S_t) du signal est définie comme la position y la plus élevée ou la plus faible, en fonction de l'orientation de l'axe y du capteur d'image, et

- 55 - dans lequel le centre horizontal $x_{m,t}$ de la région d'image (S_t) du signal est défini comme la coordonnée x du centroïde de la région d'image (S_t) du signal,

e) estimer une transformation de perspective (R) pour transformer l'image (I_t) dans le plan de projection

commun, de préférence dans une image vue du ciel (I_B), sur la base de propriétés visuelles de l'image et/ou de propriétés de montage du capteur d'image,
f) appliquer la transformation de perspective

- 5 - à la région d'image sélectionnée (T_C) ou à des parties de celle-ci, *en particulier un trajet central à l'intérieur de la voie ferrée*, et
 - aux points projetés, et

- 10 g) déterminer un ou plusieurs points projetés (G_i) qui présentent une distance minimum et/ou une distance inférieure à une distance de seuil prédéfinie et/ou un placement relatif par rapport à la région d'image sélectionnée (T_C) sous ladite transformation de perspective (R) et identifier les signaux et/ou les régions d'image (S_i) correspondant à des signaux associés à ces points projetés dans l'image (I_t).

- 15 4. Système selon la revendication 3, **caractérisé en outre par** une unité de géolocalisation connectée à l'unité de traitement, dans lequel

- l'unité de traitement comprend en outre une mémoire pour une carte du réseau de chemin de fer incluant les positions des voies ferrées et des signaux, dans lequel chaque signal est associé à une voie ferrée, et
20 - l'unité de traitement est en outre programmée pour exécuter les étapes suivantes consistant à :

- trouver la position du train fournie par l'unité de géolocalisation au sein de la carte de chemin de fer et *la voie ferrée qu'il utilise*,
 - utiliser la carte pour déterminer la position du signal d'approche associé à la voie ferrée qu'il utilise, et
25 - uniquement si le signal d'approche se trouve dans une région prédéterminée devant le train, prendre une image et exécuter les étapes b) à g) de la revendication 3.

30

35

40

45

50

55

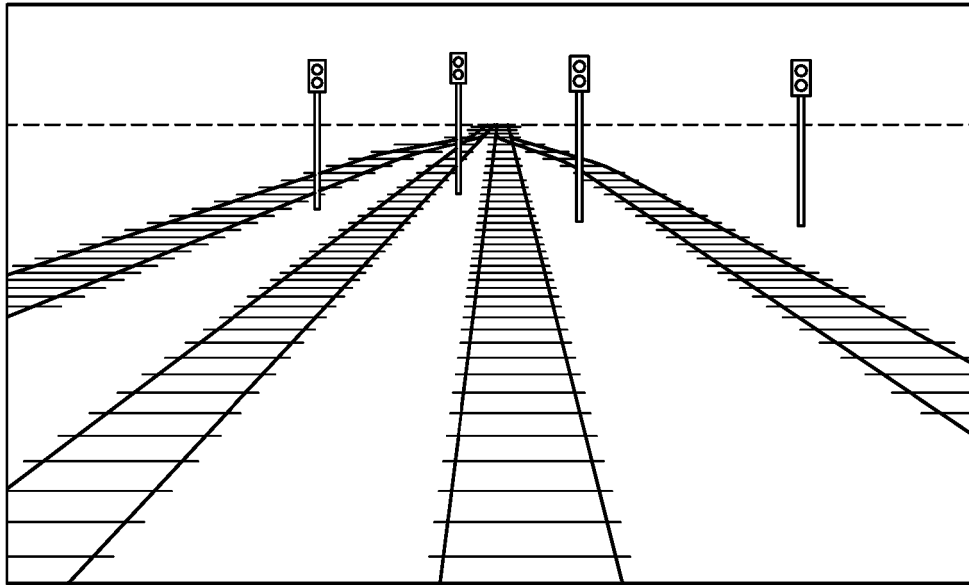


Fig. 1

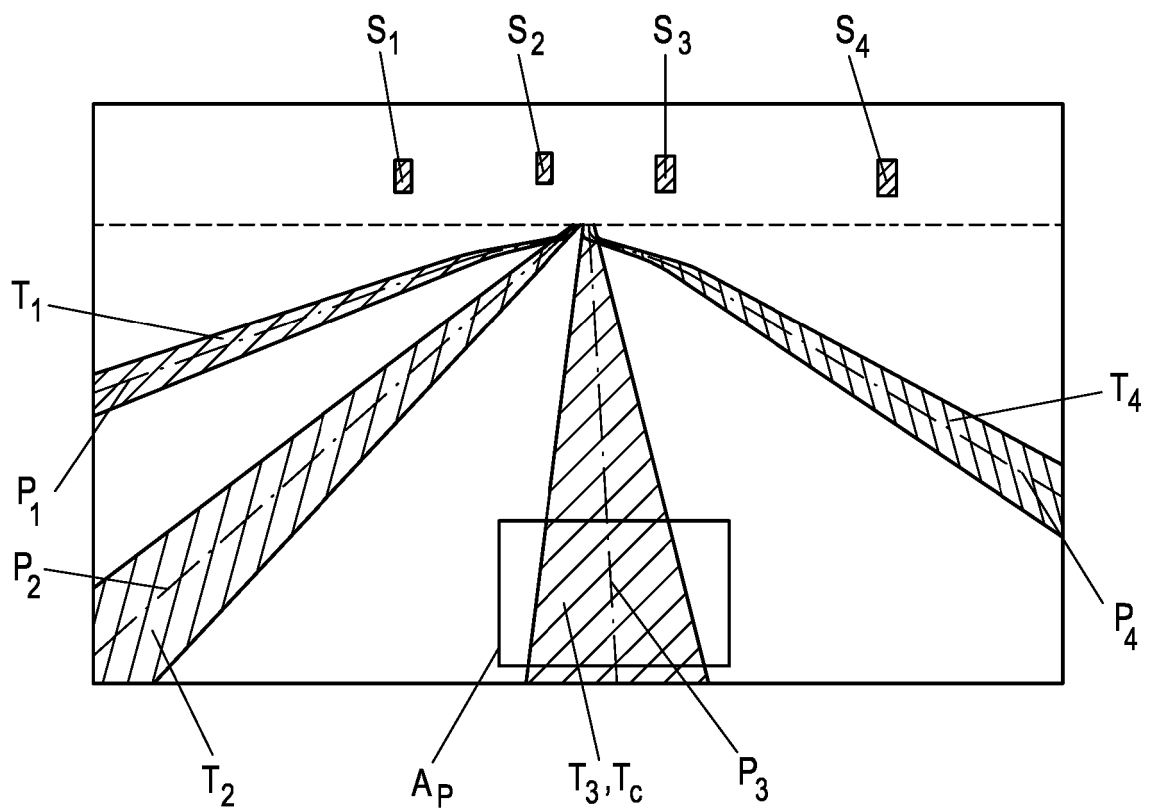


Fig. 2

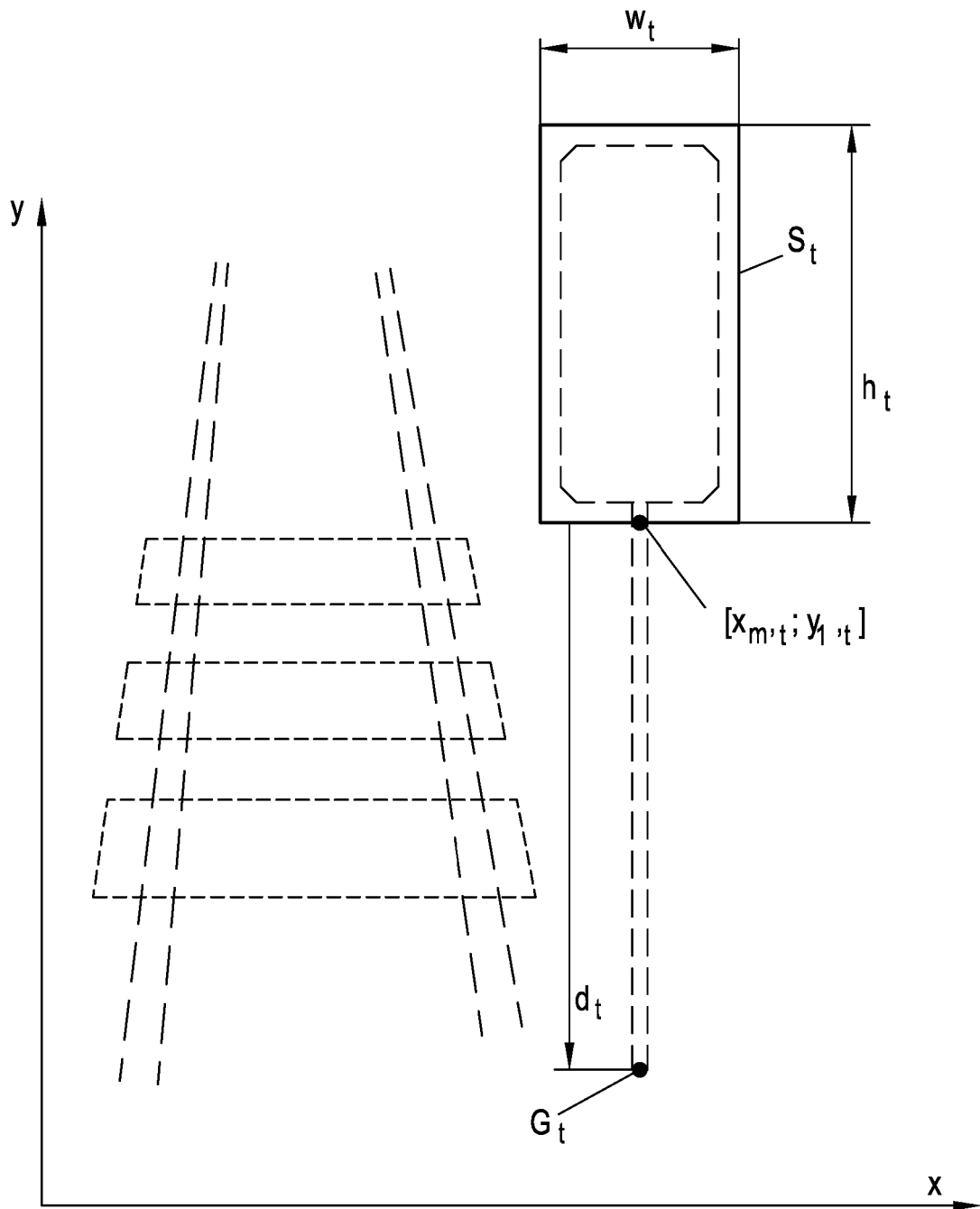


Fig. 3

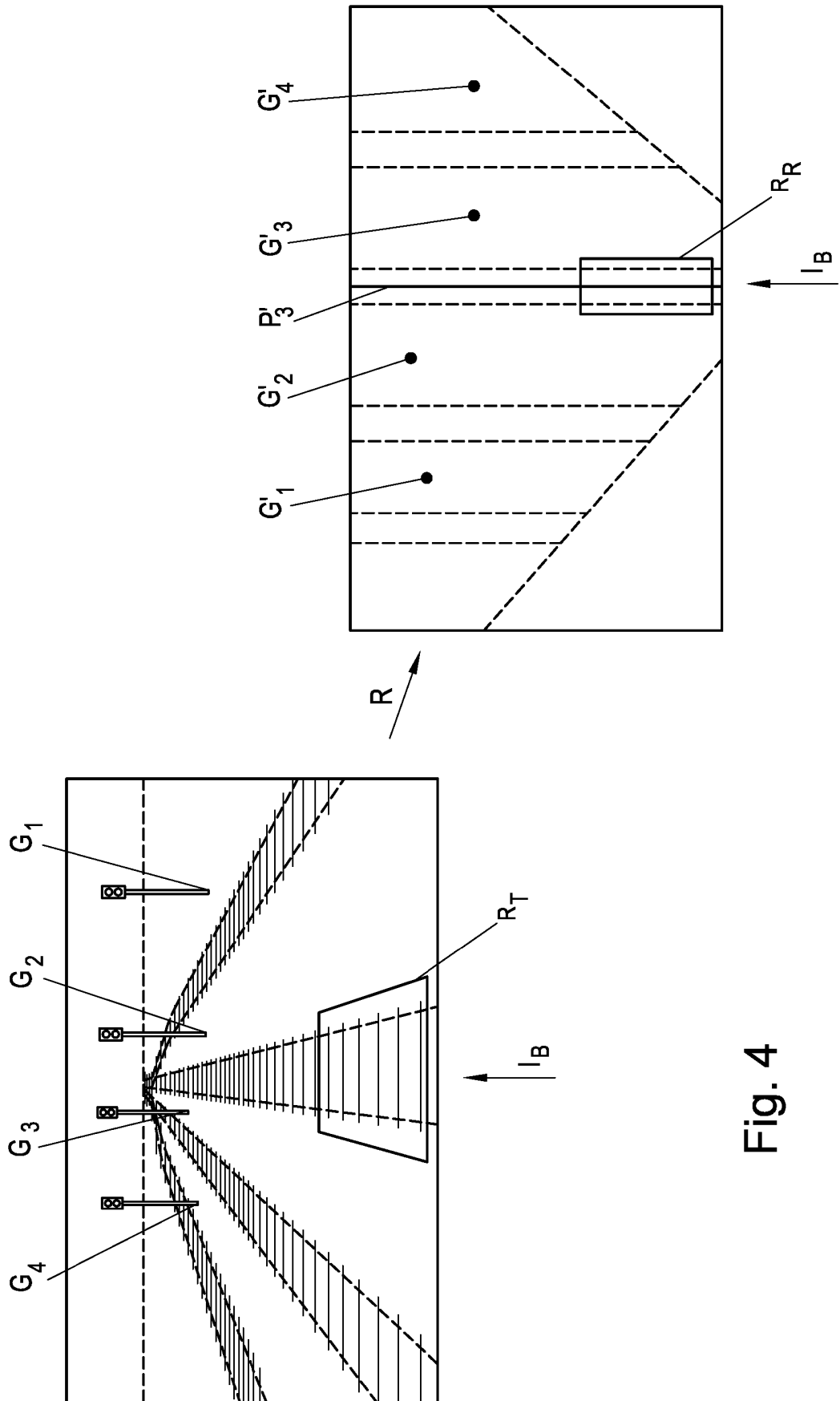


Fig. 4

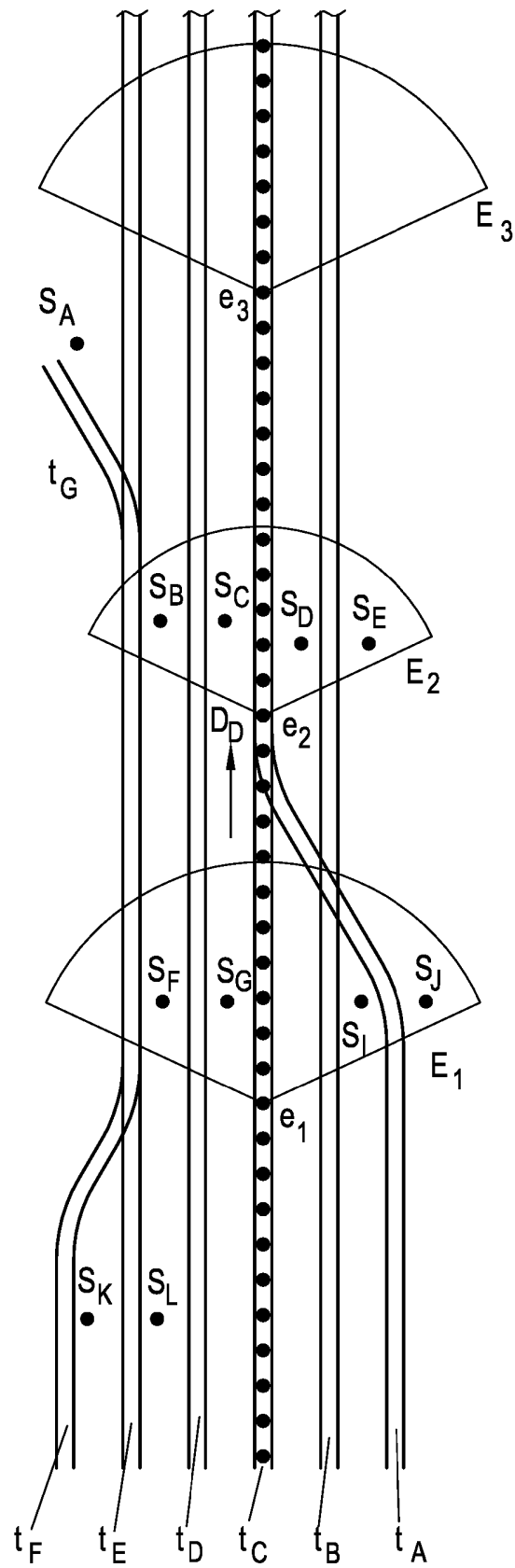


Fig. 5

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Non-patent literature cited in the description

- **SINA AMINMANSOUR.** Video Analytics for the Detection of Nearmiss Incidents at Railway Level Crossings and Signal Passed at Danger Events. Queensland University of Technology, 10 October 2017 [0010]
- **RUI FAN ; NAIM DAHNOUN.** Real-Time Stereo Vision-Based Lane Detection System, 2018 [0024]
- **HAN MA.** Multiple Lane Detection Algorithm Based on Optimised Dense Disparity Map Estimation, 2018 [0024]
- **DAVY NEVEN ; BERT DE BRABANDERE ; STAMATIOS GEORGOULIS ; MARC PROESMANS ; LUC VAN GOOL.** Towards End-to-End Lane Detection: an Instance Segmentation Approach, 2018 [0025]
- **MIN BAI.** Deep Multi-Sensor Lane Detection, 2018 [0025]