(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 December 2020 (17.12.2020)

WIPO | PCT

(10) International Publication Number
WO 2020/249558 A1

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

(21) International Application Number:
PCT/EP2020/065968

(22) International Filing Date:
09 June 2020 (09.06.2020)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
19179792.7 12 June 2019 (12.06.2019) EP

(71) Applicant: **MISSION EMBEDDED GMBH** [AT/AT]; Gutheil-Schoder-Gasse 8-12, 1100 Wien (AT).

(72) Inventors: **REISNER, Clemens**; Diehlgasse 33/21, 1050 Wien (AT). **SAWAR, Omair**; Schrotzbergstrasse 1/6, 1020 Wien (AT). **KREILMEIER, Michael**; Hartmanngasse 12/20, 1050 Wien (AT).

(74) Agent: **WILDHACK & JELLINEK PATENTANWÄLTE**; Landstraßer Hauptstraße 50, 1030 Wien (AT).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA,

(54) **Title:** METHOD FOR DETECTING A RAILROAD SIGNAL

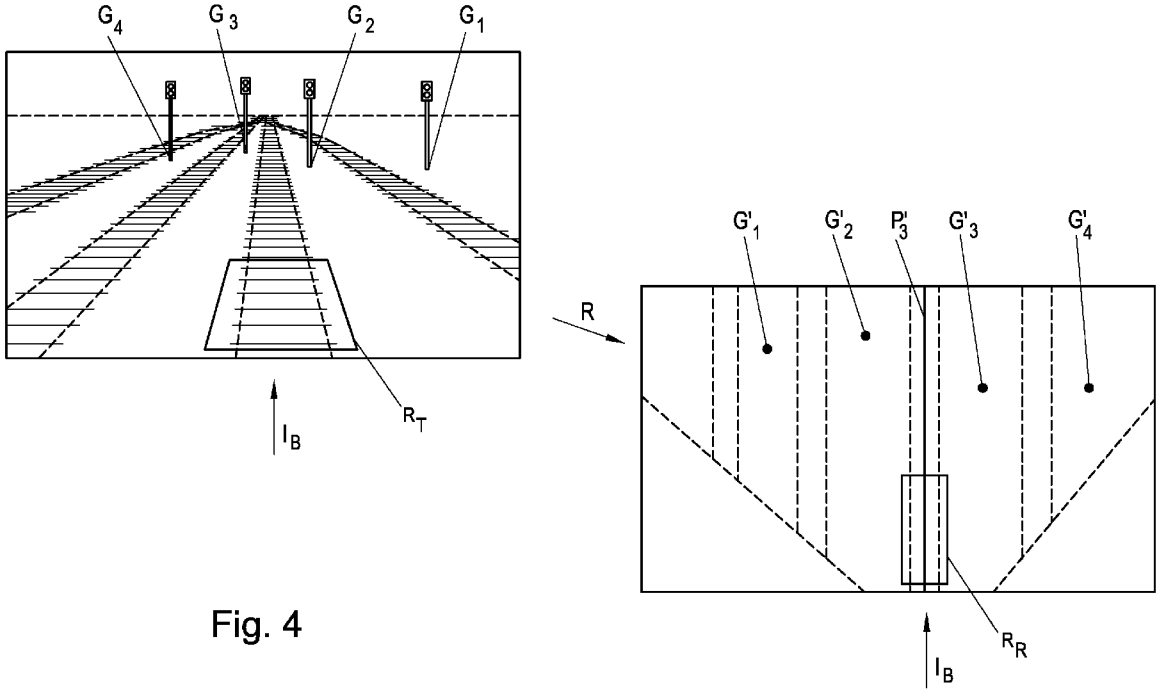


Fig. 4

(57) **Abstract:** 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 a train with an image sensor being located in or on the train generating an image of the on-coming rail tracks and signals, 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: -

WO 2020/249558 A1

[Continued on next page]

SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

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

estimating projected points of the signals and/or the projected path of the selected rail track (T_c) into a common projection plane, *in particular a ground plane, preferably estimating the position of a ground point (G_t) of the respective signal within the image (I_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).

Method for detecting a railroad signal

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.

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.

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

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.

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.

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.

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.

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.

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.

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.

The invention solves this problem using a method of claim 1for 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 a train with an image sensor being located in or on the train generating an image of the on-coming rail tracks and signals,

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/or the projected path of the selected rail track into a common projection plane, *in particular a ground plane, preferably estimating the position of a ground point of the respective signal within the image,*

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.

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

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, 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 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/or the projected path of the selected rail track into a common projection plane, *in particular a ground plane, preferably estimating the position of a ground point of the respective signal within the image,*

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.

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.

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.

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 its 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.

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.

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.

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 one 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.

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.

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

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

In a second step (b) a classification method is used in order to automatically detect the rail track-regions T_1, ..., T_4 and the signal-regions S_1, ..., S_4. Railtracks can be detected either using classical approaches, e.g. shown in *Rui Fan and Naim Dahnoun "Real-Time Stereo Vision-Based Lane Detection System", 2018* or *Han Ma et el. "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.

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 el. "Deep Multi-Sensor Lane Detection", 2018.*

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.
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.

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.

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.

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

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, ..., 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, ..., 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.

In order to determine the ground point G_1, ..., 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.

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**).

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.

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.

Alternatively, it is also possible to determine the pixel height h_t and width w_t of the signal's image region S_t 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 $A = w_t*h_t = w_t*w_t*r$ 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.

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 can be calculated as:

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

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.

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 $y_{1,t}$ can be 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 can be defined as the x-coordinate of the centroid of the signal region. The position of the ground point G_t can be defined as:

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

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.

There could be other possible ways to determine the ground point of a signal. For example, using the known gauge-width of the rail track and the signal panel's dimensions (i.e. real height and width), one could estimate the ground point of the signal in the two-dimensional image by comparing the signal panel's pixel-width to the gauge's pixel-width of the gauge of the rail track, both scaled in an appropriate manner. In order to determine correctly scaled pixel-heights and/or pixel-widths it is advantageous to consider the image's perspective distortion and the orientation of the rail tracks.

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.

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.
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.

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. As an example, the transformation R can be 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 Tc 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.

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, ..., 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.

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.

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.

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, ..., t_G as well as the position S_A, ..., S_L of the signals (**Fig. 5**).

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.

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.

If the map of rail tracks, signals and the driving direction DD are known, it is possible to determine the positions S_A, ..., S_L of signals associated with the respective rail track. Therefore, it is also possible to determine, if - within a predefined region E_1, ... E_3 calculated from the actual

position of the train e_1, ..., e_3 – a signal is expected to be seen in the image. Such a signal is also referred to as on-coming signal.

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.

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.

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.

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, ..., 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_t 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.

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.

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.

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 a train with an image sensor being located in or on the train generating an image of the on-coming rail tracks and signals,
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:
- estimating projected points of the signals and/or the projected path of the selected rail track (T_C) into a common projection plane, *in particular a ground plane, preferably estimating the position of a ground point (G_t) of the respective signal within the image (I_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 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,
- 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. Method, in particular according to claim 1 or 2, for determination, in particular in step d) of claim 1 or 2, of the ground point (G_t) of a signal (S), in particular of a signal (S) mounted on a pole or on overhead infrastructure,
- wherein the ground point (G_t) is the point within an 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),

comprising the following steps:
- taking at least one two-dimensional image (I_t) from a train with an image sensor being located in or on the train generating an image of at least the on-coming signals,
- wherein the image sensor is mounted on the train in a manner that its y-axis is aligned substantially

vertically, and its x-axis is aligned substantially horizontally,

- using a classification method for identifying image regions (S_t) corresponding to signals (S) within the image (I_t),
- determining the pixel height h_t and pixel width w_t of the signal's image region (S_t) as follows:
  - counting the 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 the known ratio r between the height and width of the panel of the signal (S),
- calculating the y-distance d_t from the 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 = mounting height * h_t / 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 the 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).

4. 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, 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:

- estimating projected points of the signals and/or the projected path of the selected rail track (T_C) into a common projection plane, *in particular a ground plane, preferably estimating the position of a ground point (G_t) of the respective signal within the image (I_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).

5. System according to claim 4, 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 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 4.

6. System, in particular according to claim 4 or 5, for determination of the ground point (G_t) of a signal (S), in particular of a signal (S) mounted on a pole or on overhead infrastructure,
- wherein the ground point (G_t) is the point within an 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), 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 at least the on-coming signals,

  wherein the image sensor is mounted 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 configured to execute the following steps:

  - taking at least one two-dimensional image (I_t) with the image sensor being located in or on the train generating an image of at least the on-coming signals,

  - using a classification method for identifying image regions image regions (S_t) corresponding to signals within the image (I_t),

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

    - counting the 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 the known ratio r between the height and width of the panel of the signal (S),

  - calculating the y-distance d_t from the 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 the horizontal center of the signal's image region (S_t) 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 (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).

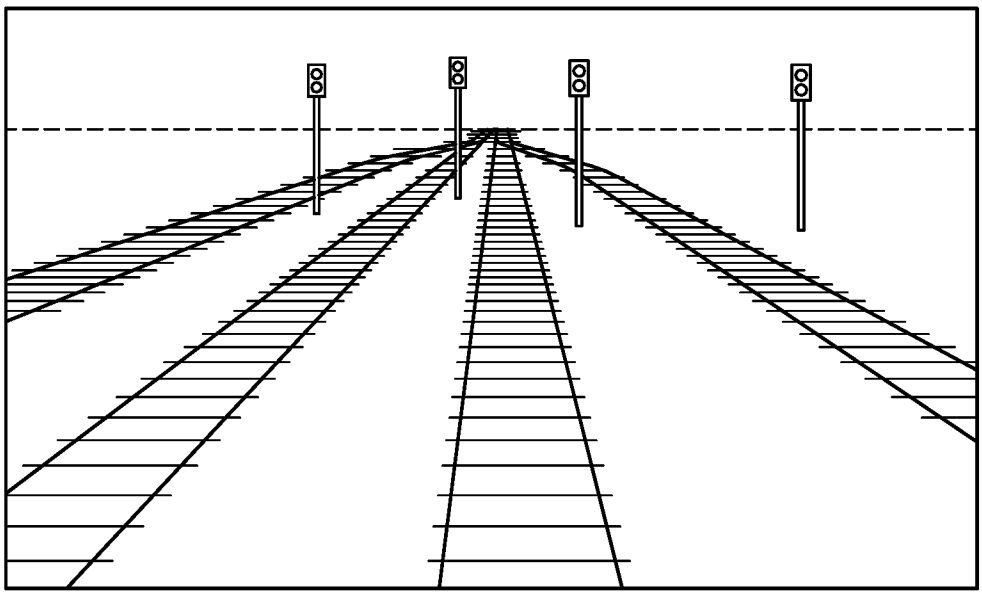


Fig. 1

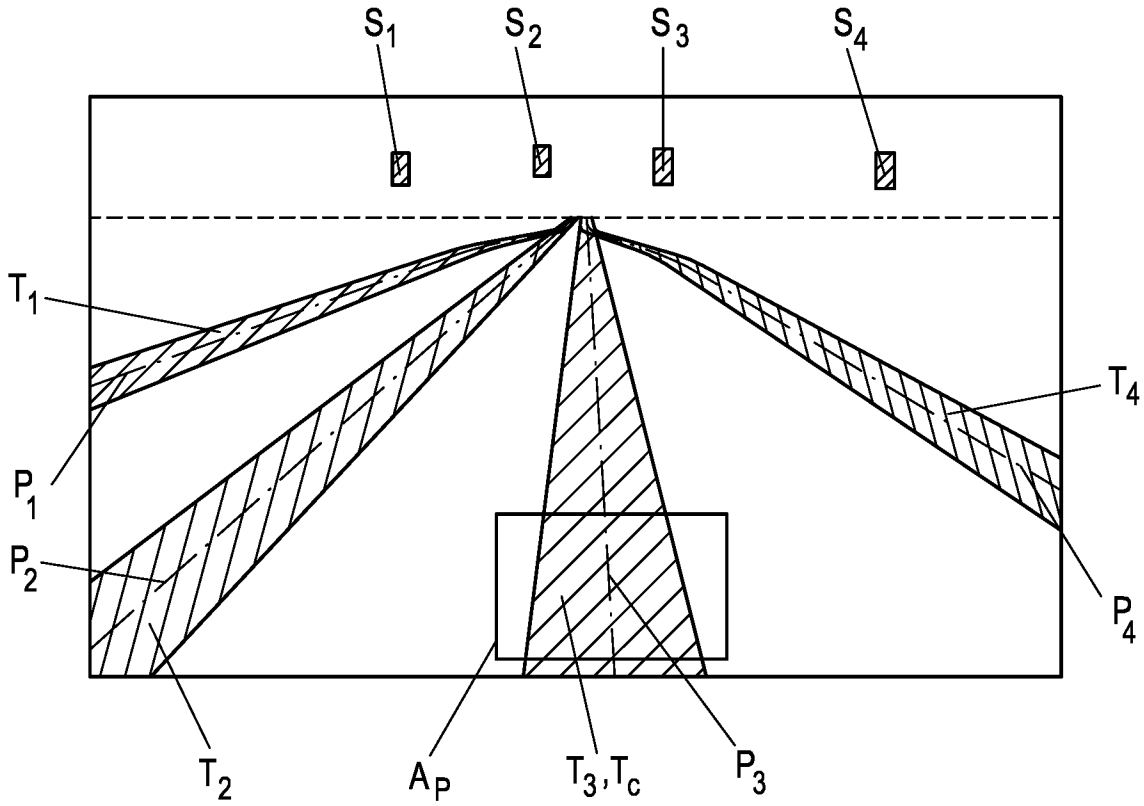


Fig. 2

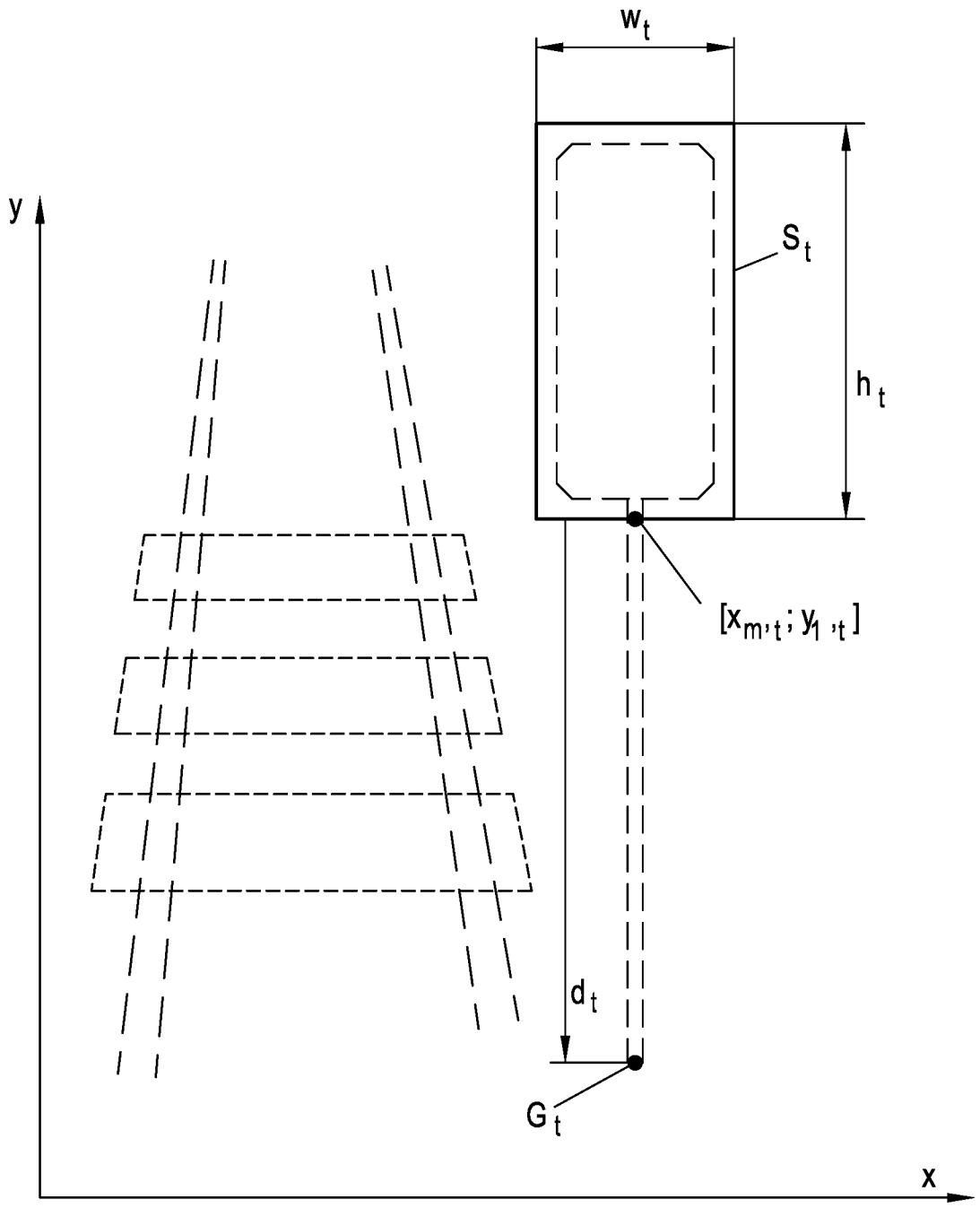


Fig. 3

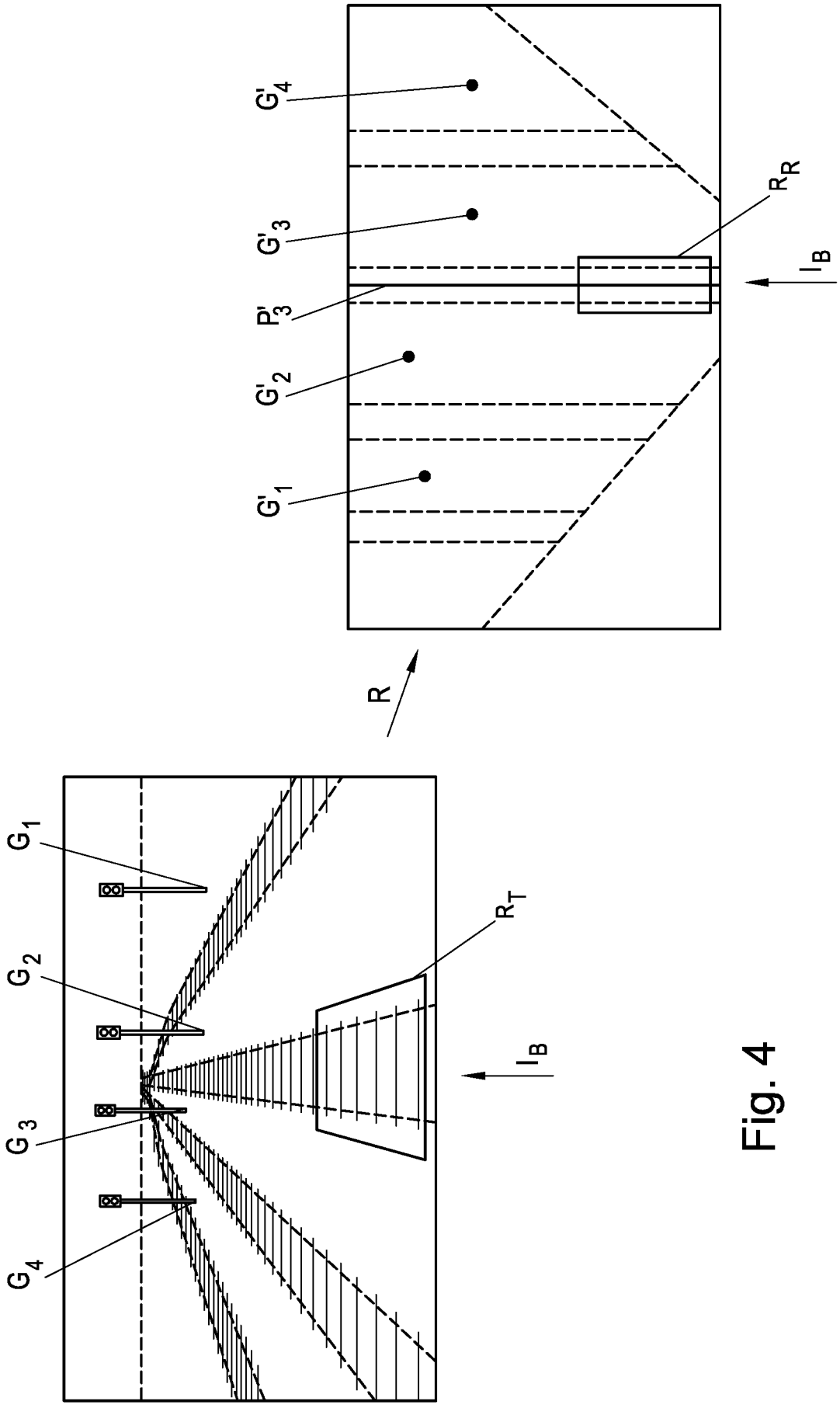


Fig. 4

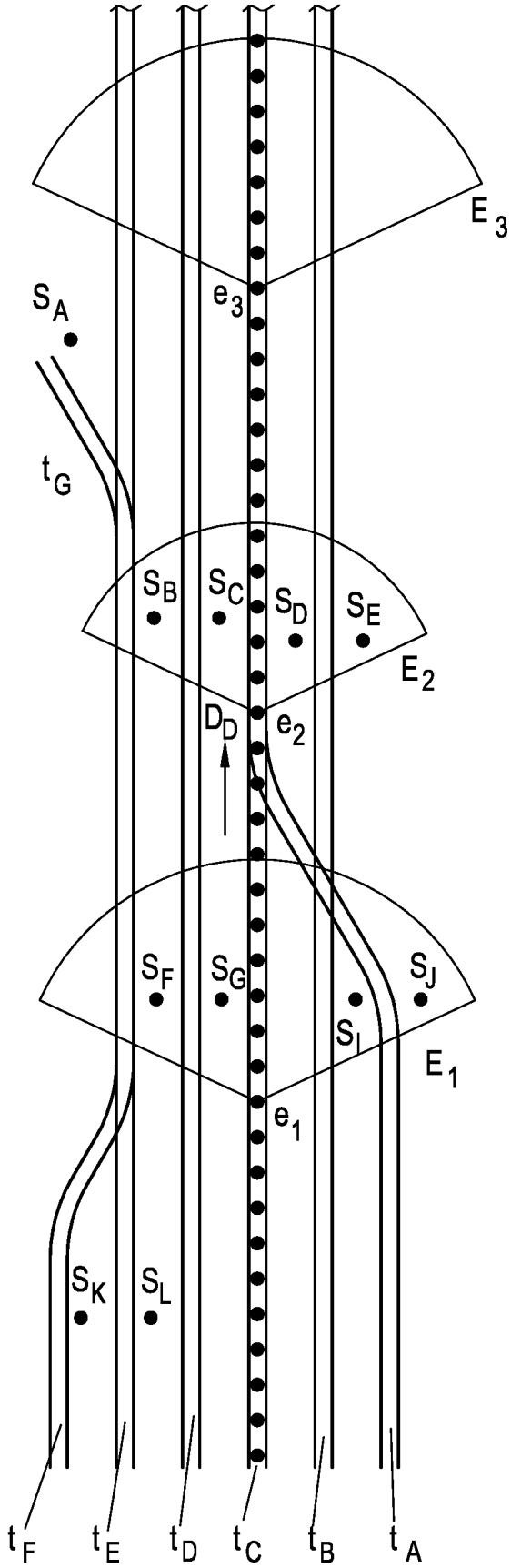

E3
e3
SA
tG
SB
SC
SD
SE
E2
DD
e2
SF
SG
SJ
SI
E1
e1
SK
SL
tF
tE
tD
tC
tB
tA

Fig. 5

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2020/065968

A. CLASSIFICATION OF SUBJECT MATTER
INV. B61L23/04
ADD. B61L3/00 G06T3/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
B61L G06T

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	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, 10 October 2017 (2017-10-10), XP055640309, https://eprints.qut.edu.au/112765/1/Sina_Aminmansour_Thesis.pdf	1,2,4,5
A	https://eprints.qut.edu.au/112765/1/Sina_Aminmansour_Thesis.pdf page 15 page 35; figure 2.15 page 45 page 52 - page 53 page 57 page 67 page 70 - page 72; figure 4.4 page 88 page 130	3,6

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
29 July 2020	06/08/2020

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Pita Priegue, Miguel

3

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	page 132 - page 133; figure 7.5 page 136 -----	
A	US 2018/057030 A1 (PUTTAGUNTA SHANMUKHA SRAVAN [US] ET AL) 1 March 2018 (2018-03-01) the whole document -----	1-6
A	EP 3 048 559 A1 (RINDINVEST AB [SE]) 27 July 2016 (2016-07-27) the whole document -----	1-6

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2020/065968

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2018057030 A1	01-03-2018	US 2016221592 A1	04-08-2016
		US 2018057030 A1	01-03-2018
EP 3048559 A1	27-07-2016	NONE	