

Video surveillance system installed on board of a rail vehicle

The present invention relates to a video surveillance system on board a rail vehicle, and a related video surveillance method.

The invention is in the general field of automated video surveillance, based on video streams from video surveillance cameras, and in particular for the application of video surveillance in the railroad field.

In this field there are specific constraints, linked in particular to railroad standards, vibrations caused by the rolling of a vehicle (train or metro for example), available space, and limited power consumption.

The algorithms for processing and analyzing video streams are increasingly complex and consume computing power and energy to enable analysis and detection of situations on board a vehicle, such as the detection of abnormal or aggressive behavior, falling passengers, or overloading in the occupation of a rail vehicle. Such known algorithms are currently implemented on computers located in a ground processing center, which have sufficient computing capacity, on video streams transmitted from video surveillance cameras on board a rail vehicle. Processing in a ground-based computer center is limited by the bandwidth capacities of a communication network enabling video streams to be routed to the computers, which leads to a time delay and consequently makes it difficult to obtain real-time situation analysis results, which can be critical, especially if a detected situation presents a risk to passenger safety.

There are also video surveillance cameras, or CCTV (closed-circuit television) cameras, which incorporate calculation processors and are suitable for implementing image analysis processing. However, each camera is only capable of processing the images of the video streams it acquires, and therefore it is not possible in this case to carry out a situation analysis by correlating information from several cameras. Moreover, cameras compatible with railroad standards have limited computing power and therefore can only implement simple processing, such as algorithms for detecting motion in successive images of a video stream. Such processing is insufficient for a situation analysis of persons and objects present in a part of a rail vehicle, and therefore for detecting possible risk situations.

The object of the invention is to remedy the drawbacks of the state of the art.

To this end, the invention proposes a video surveillance system on board a rail vehicle, comprising a plurality of video surveillance cameras on board the rail vehicle, each camera having a spatial position and an associated field of view, each camera being capable of capturing a video stream made up of a succession of digital images. This

system further comprises an on-board electronic calculation device comprising at least one calculation acceleration device. Each camera is connected to the on-board electronic calculation device and is adapted to transmit video streams captured to the said electronic calculation device, the electronic calculation device being capable of obtaining a
 5 breakdown of the plurality of video surveillance cameras into a sequence of groups of cameras to be processed successively over time, on a circular repetition basis, and for each group of cameras processed successively:

- obtaining video streams from said group of cameras, each video stream containing at least one digital image,
- 10 - applying a digital image analysis to calculate values of descriptor values of objects present in the or each image of each video stream from a camera in said group of cameras,
- applying at least one situation analysis algorithm using the descriptor values calculated for at least one image of each video stream of said group of
 15 cameras processed to obtain a situation analysis result in the field of view of the cameras of said group of cameras.

Advantageously, the system of the invention makes it possible to carry out an automatic analysis of video streams captured by groups of cameras, the groups of cameras being processed successively, with circular repetition, by the same on-board
 20 electronic calculation device comprising a calculation accelerator. In this way, railroad constraints are respected, and the repetition of the processing of video streams by groups of cameras enables automated video surveillance in almost real time for the entire rail vehicle.

According to the invention, the video surveillance system may also have one or
 25 more of the features below, taken independently or in any technically conceivable combination.

In each digital image of a video stream, the digital image analysis implements a detection of persons and/or an extraction of a skeleton related to a person present in the vehicle.

30 The characteristic descriptor values comprise spatial positions, in a reference frame related to the digital image, of joints associated with a person present in the scene.

The system is installed on board a rail vehicle comprising a plurality of carriages, each group of cameras consisting of cameras installed on board one of the vehicle's carriages.

The system is capable of implementing multiple different situation analysis algorithms depending on the state of the rail vehicle.

The situation analysis algorithm implements abnormal event detection, and, in the event of a positive detection for a current camera group, the system is capable of continuing to process video stream images from the current camera group.

In addition, the system is suitable for raising an alert after the situation analysis application.

The system comprises a timing module capable of selecting the next group of CCTV cameras to be processed based on the results of the situation analysis.

According to another aspect, the invention relates to a video surveillance method implemented by a video surveillance system as briefly described above. This method comprises steps consisting of:

- obtaining a breakdown of the plurality of video surveillance cameras into a series of groups of cameras to be processed successively over time, according to a circular repetition, and, for each group of cameras processed successively:

- a) obtaining video streams from said group of cameras, each video stream containing at least one digital image,

- b) applying digital image analysis to calculate values of characteristic descriptors of objects present in the or each image of each video stream from a camera of said group of cameras,

- c) applying at least one situation analysis algorithm using the descriptor values calculated for at least one image of each video stream of said group of cameras processed to obtain a situation analysis result in the field of view of the cameras of said group of cameras.

The method has the same advantages as the video surveillance system briefly described above.

In one embodiment, the method includes a step of selecting the next group of cameras to be processed.

In one embodiment, the situation analysis algorithm uses abnormal event detection, and, in the event of positive detection of an abnormal event for a current group of cameras, the method repeats steps a), b) and c) for the current group of cameras.

In one embodiment, an alert is also raised if an abnormal event is positively detected.

According to another aspect, the invention also relates to a computer program containing software instructions which, when executed by a programmable electronic device, implement a video surveillance method as briefly described above.

Other features and advantages of the invention will be apparent from the description given below, which is indicative and in no way limitative, with reference to the appended figures, among which:

- Figure 1 is a schematic representation of an on-board video surveillance system according to an embodiment of the invention;

- Figure 2 is a schematic illustration of a first method of carrying out a circular repetition of processing by groups of cameras;

- Figure 3 is a schematic illustration of a second method of carrying out circular repetition of processing by groups of cameras;

- Figure 4 is a synopsis of the main steps of an automatic video surveillance method according to an embodiment.

The invention will be described below in its application in a railroad video surveillance system, on board a rail vehicle, such as a train or a metro.

Figure 1 schematically illustrates the functional blocks of a video surveillance system 1 in one embodiment, the video surveillance system being embedded in a rail vehicle 2, such as a train.

The video surveillance system 1 comprises a plurality of on-board cameras $4_1, \dots, 4_i, \dots, 4_N$, the number N being an integer chosen according to various constraints, such as the number N being the number of carriages forming part of rail vehicle 2.

For example, in one embodiment, vehicle 2 comprises 3 carriages, and 2 video surveillance cameras are spatially positioned in each carriage, for example integrated into the ceiling or side walls of the car and adapted to capture images with a related field of view, the field of view being directed towards the interior of the car. Preferably, the system consists of 2 to 8 cameras per rail vehicle carriage, their spatial position being chosen so that their combined fields of view cover the interior space of the rail vehicle.

For example, in one embodiment, each camera is capable of acquiring video with a resolution of 1080x720 pixels.

In the video surveillance system 1, the cameras are grouped into camera groups rated G_1 to G_L , forming a sequence of camera groups. For example, $L=3$ camera groups are distinguished in the example in Figure 1. The order of the camera groups in this

sequence of camera groups is chosen according to an order of railroad carriages, for example.

The video surveillance system 1 also includes a programmable electronic device 10 capable of making calculations, on board the vehicle 2, referred to below simply as a calculation device 10.

Each video surveillance camera 4_i includes a communication module capable of implementing a given communication protocol, for example the Internet Protocol (IP), connected by a communication link 8_i to the calculation device 10, for example a wired link.

The calculation device 10 also includes communication modules $9_1, \dots, 9_i, \dots, 9_N$ enabling making a communication connection to be made with each on-board camera 4_i .

Thus, the device 10 is capable of acquiring video streams captured by each of the on-board video surveillance cameras, in real time.

In one embodiment, each video stream made up of a sequence of digital images is coded in a compressed format so as to reduce the amount of binary data to be transmitted to the computing device 10. For example, the H264 compression format is used. The calculation device 10 is capable of decoding video streams in H264 format.

The calculation device 10 includes one or more central processing units (CPUs) with an operating frequency of at least 1.4 GHz. It also includes at least one calculation acceleration device 14, physically adapted to carry out image processing calculations, for example convolutions, in an optimized manner. The calculation acceleration device 14 is, for example, a GPU ("Graphics Processing Unit"), TPU ("tensor processing unit"), DSP ("digital signal processor") or FPGA ("Field Programmable Gate Array") processor.

Preferably, the calculation acceleration device 14 is capable of performing at least one trillion floating point operations per second while consuming less than 15W.

The calculation device 10 also includes a memory 16 capable of storing digital data, including computer program software instructions, configuration files and parameter values used by software implementing a video surveillance method as described when executed by the processors 12 and the calculation acceleration device(s) 14.

The video-surveillance method of the invention is implemented, in a production mode, by software instruction modules 20, 22, 24, 26.

The module 20 for obtaining video streams makes it possible to obtain video streams, over a time interval, preferably time-synchronized, coming from all the video surveillance cameras in the same group of video surveillance cameras. Each video stream includes at least one digital image acquired at a given time.

The image analysis module 22 implements one or more image processing algorithms to calculate characteristic values of objects present in each image. The term "objects" is to be understood here in the broadest sense, and includes inanimate objects, people and animals. For example, it connotes a person detection algorithm, and a skeleton detection algorithm, as described in more detail below.

The situation analysis module 24 implements one or more situation analysis algorithms according to the characteristic values of objects calculated on several images of all the video streams from the video surveillance cameras of the group being processed.

The timing module 26 enables the selection of a next group of video surveillance cameras, allowing all the camera groups to be selected in succession, and to perform circular repetition, that is, to select the first camera group in the sequence of camera groups again after processing the last camera group in this sequence. In other words, the video streams provided by the video surveillance cameras are processed in groups of video streams using carousel processing.

The rotation timing between successive groups of cameras providing video streams to be processed is chosen, and varies, for example, according to the situation analysis to be carried out by the module 24.

For example, for an estimate of passenger density, the analysis of a single image of a video stream per camera is sufficient, whereas for the detection of an act of aggression, for example, it is necessary to process a number of images, for example equal to 48, of each video stream of a group of cameras before moving on to the next group of cameras.

Figure 2 schematically illustrates a first method of processing two groups of cameras, G_1 and G_2 , each comprising two distinct cameras and providing two video streams, using the time diagrams 32, 34, 36, 38 shown in parallel. Figure 2 schematically illustrates the temporal sequencing of the processing of the respective video streams by the processor 12 and the calculation acceleration device 14 of the calculation device 10.

Thus, at the time noted t_0 , a first image $I_{1,1}$ of the first video stream $V_{1,1}$ of the first group of cameras G_1 is acquired, as well as a first image $I_{1,2}$ of the second video stream $V_{2,1}$ of the first group of cameras G_1 . Then, at a point in time t_1 , an analysis of the $I_{1,1}$ image is performed, and at the point in time t_2 , an analysis of the $I_{1,2}$ image is performed. For example, the analysis processing is the detection of a skeleton.

At a later point in time t_3 , the results of the previous analyses are merged to obtain characteristic descriptor values for the objects in the scene captured by the two cameras in group G_1 , and these characteristic descriptor values are stored.

The method then processes the video streams from the second camera group G_2 .

Thus, at the time t_4 , a first image $J_{1.1}$ of the first video stream V_{12} of the second camera group G_2 is acquired, as well as a first image $J_{1.2}$ of the second video stream V_{22} of the second camera group G_2 . Then, at a point in time t_5 , an analysis of the $J_{1.1}$ image is performed, and at a point in time t_6 , an analysis of the $J_{1.2}$ image is performed. For example, the analysis processing is skeleton detection.

At a later point in time t_7 , the results of the previous analyses are merged to obtain characteristic descriptor values for the objects in the scene captured by the two cameras in the second camera group G_2 , and these characteristic descriptor values are stored.

The method then re-processes the first group of cameras G_1 , starting at a time t_8 when second images from the first and second video streams of the first group of cameras G_1 are obtained, and so on.

In this example, only one image is acquired and analyzed per camera group before moving on to the next group.

After performing M processing sequences for the same camera group, where M is 4 in this example, a situation analysis is performed at a time t_a . For example, this involves an integration of the previously stored results in order to carry out a count of people present in the joint field of vision of a group of cameras. In a practical case, this involves counting a number of people per rail carriage.

Figure 3 schematically illustrates a second way of processing two camera groups G_1 and G_2 , each comprising two separate cameras and providing two video streams, using time diagrams 42, 44, 46, 48 shown in parallel. Figure 2 schematically illustrates the temporal sequencing of the processing of the respective video streams by the processor 12 and the calculation acceleration device 14 of the calculation device 10.

In contrast to the embodiment illustrated in Figure 2, in this second embodiment, a series of 5 images from the first and second video streams of the first group of cameras G_1 are acquired and processed before processing the first and second video streams of the second group of cameras G_2 . For example, in this second production mode, the processing of each image extracted from a video stream consists of detecting a skeleton, and the joint processing of the descriptor values consists of detecting an abnormal event relating to people, for example a fall or an assault.

Figure 3 illustrates the processing of the first group of cameras up to the implementation of the situation analysis at the instant t'_a , enabling the detection of abnormal behavior.

In this second embodiment, for the implementation of the situation analysis, the acquisition of images from the first and second video streams of the first group of cameras is continued in parallel with the implementation of the situation analysis. If an abnormal event is detected, it is planned to continue the image and situation analysis, and to raise alarms following the positive detection of the occurrence of an abnormal event. In this case, the processing timing of the video streams from the camera groups in the sequence of camera groups is dynamically modified to take into account an abnormal situation detected.

For example, in an embodiment, the processing of the video streams from the first group is continued until a negative detection of an abnormal event, then the next camera group is processed.

In the example shown in Figure 3, there is no detection of an abnormal event, and therefore the acquisition and processing of the images of the first and second video streams of the second camera group G_2 is carried out from the time t'_a of the end of processing the first camera group G_1 .

It should be noted that in one embodiment, the image acquisition is carried out in parallel with the situation analysis, as the situation analysis does not occupy all the computing resources of the calculation device 10.

Figure 4 is a synopsis of the main steps in a video surveillance method according to an embodiment, implemented by a calculation device 10.

The method includes a first step of obtaining a breakdown into a series of groups of cameras $\{G_1...G_L\}$ to be processed, from among all the N cameras connected to the calculation device 10. For example, the breakdown is obtained from a configuration file. Alternatively, the breakdown is performed by calculation, depending on the spatial position of each camera in relation to a reference frame related to the rail vehicle.

Preferably, the cameras in a given camera group have joint fields of view, that is, they have an intersection.

Each camera group provides a synchronized or quasi-synchronized video stream, acquired over a given time interval, to be processed before moving on to processing the video stream provided by the next camera group.

In the embodiment described above, the video streams acquired by a group of cameras are considered synchronized, that is, the images of the respective video streams

from each camera in the group of cameras are acquired at predetermined moments in time.

The method consists of selecting a group of cameras to be processed, known as the current camera group, and then acquiring the synchronized video streams from the current camera group.

A digital image analysis 56 is performed on the acquired images.

In one embodiment, the acquisition 54 and analysis 56 steps follow one another on a plurality of synchronized images extracted from the video streams from the cameras in the current camera group, as shown in Figure 3.

In one embodiment, the analysis step 56 uses a skeleton extraction algorithm. A detection of image salient points corresponding to joints is performed, each point corresponding to a joint having a related spatial position, expressed by coordinates (x,y) in the digital image frame of reference. The detected joints each have a related type of joint, for example: eyes, ears, neck, shoulders, elbows, wrists, hips, knees and ankles. Given the type of joint detected, the joints are linked together to form a skeleton, showing a person.

For example, the algorithm proposed in the paper "Realtime multi-person 2d pose estimation using part affinity fields" by Cao et al, Conference on Computer Vision and Pattern Recognition 2017, is used.

Of course, any other method that can be used to form a skeleton showing a person's skeleton from points related to the joints known to the person in the trade can be used.

The spatial positions of the articulations and their links form characteristic descriptors of objects present in the digital image analyzed.

The descriptor values are stored for each image analyzed.

Then a situation analysis step 58 is performed using a merging of the calculated and stored descriptor values.

For example, in step 58, an event detection algorithm is implemented (that is, a fall or assault) based on the trajectories of the joints of different people in a plurality of images.

According to a variant, step 58 implements an algorithm for estimating the occupancy rate (also called passenger density) of the rail vehicle car observed by the current camera group, which can be performed on a single digital image of a video stream.

In another variant, an intrusion detection or presence detection algorithm is implemented in step 58.

Preferably, the selection of the algorithm implemented in step 58 is made on the basis of a train status, such as in service, at the end of service, at the depot.

5 For example, a configuration file is related to each state of the train, and stored in memory 16 of the calculation device 10.

Indeed, it is useful to implement an occupancy rate estimation or event detection when the train is in service, a detection of presence of persons at the end of service when the train goes to the depot and a presence or intrusion detection when the train is at the
10 depot.

For certain situation analyses, for example when an event is detected in step 58, steps 54 to 58 are repeated for the current camera group.

If the event detection is negative, or if the situation analysis step carries out another situation analysis that does not require special vigilance, step 58 is followed by
15 step 60 to select the next camera group to be processed in the sequence of camera groups.

The result of the situation analysis is passed on in step 62 to a client application.

In addition, in the event of a positive verification, an alert is raised in step 62, for example an alert is sent to a control center, to the train driver. Thus, advantageously, if an
20 incident is detected, it is possible to deal with it quickly.

Advantageously, the video surveillance system and the video surveillance method described make it possible to carry out one or more situation analyses on board a rail vehicle, taking into account images captured by several cameras. Thanks to processing by groups of cameras, it is possible to carry out situation analyses for the entire rail
25 vehicle, in a manner compatible with railroad constraints.

CLAIMS

1.- A video surveillance system on board a rail vehicle, comprising a plurality of video surveillance cameras ($4_1, 4_i, 4_N$) on board the rail vehicle (2), each camera ($4_1, 4_i, 4_N$) having a spatial position and a related field of view, each camera being adapted to capture a video stream composed of a succession of digital images, the system (1) further comprising an on-board electronic calculation device (10), characterized in that:

- said electronic on-board calculation device (10) comprises at least one calculation acceleration device (14),

and in that each camera ($4_1, 4_i, 4_N$) is connected to the on-board electronic calculation device (10) and is adapted to transmit captured video streams to said electronic calculation device (10),

the electronic calculation device (10) being adapted for:

-obtaining a breakdown of the plurality of video surveillance cameras into a series of groups of cameras to be processed successively in time, according to a circular repetition, and for each group of cameras processed successively:

-a) obtaining (20) video streams from said group of cameras, each video stream containing at least one digital image,

-b) applying (22) digital image analysis to calculate values of characteristic descriptors of objects present in the or each image of each video stream from a camera of said group of cameras,

-c) applying (24) at least one situation analysis algorithm using the descriptor values calculated for at least one image of each video stream of said group of cameras processed to obtain a situation analysis result in the field of view of the cameras of said group of cameras.

2. The video surveillance system according to claim 1, wherein said digital image analysis implements, in each digital image of a video stream, a detection of persons and/or a skeleton extraction related to a person present in the vehicle.

3. The video surveillance system according to claim 2, wherein the values of characteristic descriptors comprise spatial positions, in a reference frame related to the digital image, of joints associated with a person present in the scene.

4. The video surveillance system according to any of claims 1 to 3, on board a rail vehicle comprising a plurality of carriages, each group of cameras being made up of cameras installed on board a carriage of the vehicle.

5 5. A system according to any one of claims 1 to 4, adapted for implementing several different situation analysis algorithms depending on the state of the rail vehicle.

6. The system according to any one of claims 1 to 5, wherein the situation analysis algorithm implements abnormal event detection, and in the event of a positive detection
10 for a current camera group, the system is adapted for continuing to process video stream images from the current camera group.

7. The system according to any of claims 1 to 6, further adapted for raising a warning alert after the situation analysis application.
15

8. The system according to any of claims 1 to 7, comprising a timing module (26) adapted for selecting a next video surveillance camera group to be processed, based on a situation analysis result.

20 9. A video surveillance method implemented by a video surveillance system in accordance with any one of claims 1 to 8,

comprising steps consisting of:

-obtaining (50) a breakdown of the plurality of video surveillance cameras into a series of groups of cameras to be processed successively in time, according to a circular repetition, and for each group of cameras processed successively:
25

-a) obtaining (54) video streams from said group of cameras, each video stream containing at least one digital image,

-b) applying (56) digital image analysis to calculate values of characteristic descriptors of objects present in the or each image of each video stream from a camera of
30 said group of cameras,

-c) applying (58) at least one situation analysis algorithm using the descriptor values calculated for at least one image of each video stream of said group of cameras processed to obtain a situation analysis result in the field of view of the cameras of said group of cameras.
35

10. A method according to claim 9, further comprising a step of selecting the next camera group to be processed.

11. A computer program containing software instructions which, when executed
5 by a programmable electronic device, implement a video surveillance method according to claim 9.

10

15

20