



(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
15.10.2014 Bulletin 2014/42

(51) Int Cl.:
G08G 1/07 (2006.01) **G08G 1/015 (2006.01)**
G08G 1/042 (2006.01)

(21) Application number: **13162865.3**

(22) Date of filing: **09.04.2013**

(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
Designated Extension States:
BA ME

- **Fasih, Alireza**
8020 Graz (AT)
- **Al Machot, Fadi**
9020 Klagenfurt (AT)
- **Haj Mosa, Ahmad**
9020 Klagenfurt (AT)
- **Ali, Mouhannad**
9020 Klagenfurt (AT)

(71) Applicant: **SWARCO Traffic Systems GmbH**
72669 Unterensingen (DE)

(74) Representative: **TBK**
Bavariaring 4-6
80336 München (DE)

- (72) Inventors:
- **Junghans, Ralf**
73272 Neidlingen (DE)
 - **Grüner, Karl**
81241 München (DE)
 - **Kyamakya, Kyandoghre**
9020 Klagenfurt (AT)

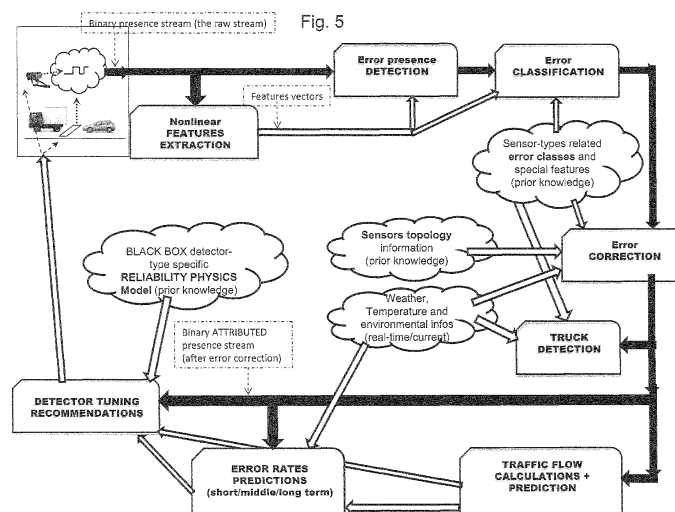
Remarks:

Amended claims in accordance with Rule 137(2) EPC.

(54) **Quality determination in data acquisition**

(57) The present invention relates to a method, apparatus and a computer program product for quality determination in data acquisition. The present invention includes a signal processing apparatus and method for processing at least one detection signal, the apparatus comprising a receiver configured to receive the at least one detection signal from at least one sensor, a model generator configured to generate a model of signal sequences based on previously received detection signals

while taking into consideration a topology of the sensors and/or environmental condition in the vicinity of the at least one sensor, comparing means configured to compare the at least one detection signal with the model of signal sequences generated by the model generator, and a prediction means configured to predict quality of the at least one detection signal based on a comparison of the at least one detection signal with the model generated by the model generator.



Description

Field of the invention

[0001] The present invention relates to quality determination in data acquisition. In particular, the present invention relates to a method, an apparatus and a computer program product for quality determination in data acquisition.

Background

[0002] Inductive loops and other well-known data suppliers like video cameras with virtual loops, infrared sensors, radar sensors and the like are called traffic detectors or sensors, respectively, and are especially used in traffic control systems for delivering a good picture of the traffic situation based on the cross section and lane-related measurement principle of traffic parameters.

[0003] Ideally, each vehicle driving over/through these physical/virtual sections on given lanes is recognized and causes the electronic detector component to generate a pulse for each vehicle presence in form of digital information.

[0004] For example, inductive loop detectors generate digital information out of the analogue signal response. The analogue and digital circuits of the detector component together with the connected loop-wire located in the road concrete/bitume may also be capable of using special parameters as measurement mechanism.

[0005] The obtained digital information is provided to the traffic control software within the intersection traffic controller and to any traffic monitoring system.

[0006] The operability and data quality of the measurement system has been so far (in the state-of-the-art) solely roughly determined through a simple evaluation considering aggregated or interval based vehicle counts on a higher system level. Some types of traffic detectors provide also a binary good/bad statement about their overall functionality.

[0007] An essential prerequisite for being able to cope with the increasing traffic volume and the resulting challenging traffic situations in terms of traffic control quality is a high data quality at the data acquisition level and it is further much more important to be informed constantly about the trend of the data quality short-, middle- and long-term. The evaluation and constant monitoring of both the reliability and the data quality of the data acquisition infrastructure in operation are playing a prominent role in modern traffic management concepts.

[0008] Normally, an eventually decreasing quality of data acquisition systems takes place slowly and gradually. This cannot be detected appropriately by existing traffic control systems and leads continuously to a deterioration of the traffic actuated control algorithm both in traffic controllers and traffic management systems at the strategic decision-making level. These decisive factors (related to sensor quality deterioration) are hardly ever

taken into account in present traffic management solutions.

Summary of the invention

[0009] In view of the above, it is an object of the present invention to provide a method, an apparatus and a computer program product in order to overcome such problems.

[0010] According to the present invention, this problem is solved by picking up the digital pulse information directly at the sensor output. A sensor reliability data model is trained offline to gather all significant characteristics of the sensor component, the physical object in the road surface/concrete, environmental influences, all this for enabling a reliable assessment of the data quality of the single sensor. Considering the topological information including lane direction, signaling and the relative position of the sensors to each other at a complete intersection, mutual interferences are also trained in the sensor reliability data model.

[0011] The resulting model or rather the set of models offers for the first time the possibility of monitoring (in real- or quasi real-time and proactively) both the current state and the evolution of sensor data quality especially for sensors delivering a digital pulse stream caused induced by the presence of vehicles in traffic.

[0012] The present invention presents a global and comprehensive concept for a robust "traffic detectors' quality management". The overall concept relates to all traffic sensors types, independently of whether they are point-sensors or area-sensors. Examples of sensors types are: loop detectors systems, video-based systems, radar-based systems (all forms of radars are meant: radio-wave based; infrared-light-based, light-based, sound-based), and (earth) magnet fields based systems.

[0013] According to an aspect of the present invention there is provided a signal processing apparatus for processing at least one detection signal, the apparatus comprising:

a receiver configured to receive the at least one detection signal from at least one sensor,
a model generator configured to generate a model of signal sequences based on previously received detection signals while taking into consideration a topology of the sensors and/or environmental condition in the vicinity of the at least one sensor,
comparing means configured to compare the at least one detection signal with the model of signal sequences generated by the model generator; and
a prediction means configured to predict quality of the at least one detection signal based on a comparison of the at least one detection signal with the model generated by the model generator.

[0014] According to further refinements of the present invention as set out under the above aspects,

- the apparatus further comprises
- determining means configured to determine whether or not the at least one detection signal comprises an error, and

a classification means configured to classify the error into a plurality of predetermined error classes, if it is determined that the detection signal comprises an error,

- correcting means configured to correct the error, if it is determined that the detection signal comprises an error,
- storing means configured to store a history of the detection signal, and prediction means configured to generate a detection signal error rate prediction based on the stored history;
- the prediction is at least one of a short term prediction, a middle term prediction, and a long term prediction;
- the apparatus further comprises determining means configured to determine settings of the sensor based on the detection signal error rate prediction such that the detection signal error rate becomes minimum; and
- the quality of the detection signal includes at least one of an overall error rate of the detection signal, and an error rate of each of the error classes of the detection signal.

[0015] According to another aspect of the present invention there is provided a method for processing at least one detection signal, the method comprising:

receiving the at least one detection signal from at least one sensor,
generating a model of signal sequences based on previously received detection signals while taking into consideration a topology of the sensors and/or environmental condition in the vicinity of the at least one sensor,
comparing the at least one detection signal with the model of signal sequences generated by the model generator; and
predicting quality of the at least one detection signal based on a comparison of the at least one detection signal with the model generated by the model generator.

[0016] According to further refinements of the present invention as set out under the above aspects, the method further comprises

- determining whether or not the at least one detection signal comprises an error, and classifying the error into a plurality of predetermined error classes, if it is determined that the detection

signal comprises an error;

- correcting the error, if it is determined that the detection signal comprises an error;
- storing a history of the detection signal, and generating a detection signal error rate prediction based on the stored history;
- the prediction is at least one of a short term prediction, a middle term prediction, and a long term prediction;
- the method further comprises determining settings of the sensor based on the detection signal error rate prediction such that the detection signal error rate becomes minimum;
- the quality of the detection signal includes at least one of an overall error rate of the detection signal, and an error rate of each of the error classes of the detection signal;

[0017] According to an exemplary aspect of the present invention, there is provided a computer program product comprising computer-executable computer program code which, when the program is run on a computer (e.g. a computer of an apparatus according to any one of the aforementioned apparatus-related exemplary aspects of the present invention), is arranged to cause the computer to carry out the method according to any one of the aforementioned method-related exemplary aspects of the present invention.

[0018] Such computer program product may comprise or be embodied as a (tangible) computer-readable (storage) medium or the like on which the computer-executable computer program code is stored, and/or the program may be directly loadable into an internal memory of the computer or a processor thereof.

[0019] Advantageous further developments or modifications of the aforementioned exemplary aspects of the present invention are derivable from the following detailed description.

Brief Description of the Drawings

[0020] For a more complete understanding of exemplary embodiments of the present invention, reference is now made to the following description taken in connection with the accompanying drawings in which:

Fig. 1 is an overview of an exemplary application of the assessment and prediction of the data quality of inductive loop systems according to certain embodiments of the present invention;

Fig. 2 is an overview illustrating the traffic control module shown in Fig. 1;

Fig. 3 is a diagram illustrating the reliability physics of a given traffic sensor system type according to certain embodiments of the present invention;

Fig. 4 is a diagram illustrating an internal parameter tuning model concept according to certain embodiments of the present invention;

Fig. 5 is a diagram illustrating the schematic overall concept of the invention according to certain embodiments.

Description of exemplary embodiments of the invention

[0021] Exemplary aspects of the present invention will be described herein below. More specifically, exemplary aspects of the present are described hereinafter with reference to particular non-limiting examples and to what are presently considered to be conceivable embodiments of the present invention. A person skilled in the art will appreciate that the invention is by no means limited to these examples, and may be more broadly applied. Further, it is noted that each of the below described aspects or embodiments can be arbitrarily combined.

[0022] Fig. 1 is an overview of a scenario to which the present invention is applicable. In particular, Fig. 1 shows an example of a crossroad that is equipped with various traffic lights and sensors. In Fig. 1, reference signs TS1 to TS4 indicate traffic signal lights and reference signs D1 to D10 indicate inductive loop detectors. Further, reference sign TC denotes a traffic controller for traffic light signaling including a traffic control module 10. As shown in Fig. 1, analogue signal from the various inductive loop detectors are input into the traffic control module 10.

[0023] Fig. 2 is a block diagram illustrating the traffic control module 10 shown in Fig. 1. The traffic control module 10 comprises a detection module 11, a RoSiT (Robust Sensors in Traffic) module 20 and traffic programming module 12. As shown in Fig. 2, analogue signals from the inductive loop detectors D1 to D10 are input to the detection module 11 and the detection module 11 outputs digitized signals to the RoSiT module 20. Further, the RoSiT module 20 outputs an indication regarding loop data quality to the traffic programming module 12.

[0024] The RoSiT module 20 comprises a pattern recognition module 21, an environmental influences module 22, an error classification module 23, a quality assessment and forecast module 24, an error correction module 25, a detector parameter recommendation module 26, and a truck detection module 27. The pattern recognition module 21, the environmental influences module 22, the error classification module 23, the error correction module 25, the detector parameter recommendation module 26, and the truck detection module 27 each input data to the quality assessment and forecast module 24, which outputs loop data quality information to the traffic programmers module 12.

[0025] In short, the present invention comprises the following aspects. Each of the below aspects will be described in more details later on.

[Aspect 1]

[0026] The first aspect concerns the from and granularity of the detector's signal raw data observation level. That is, according to a first aspect, solely microscopic detector data is used for quality management.

[Aspect 2]

[0027] According to a second aspect, out of the microscopic data, and while considering weather data and sensor topology information we do extract pattern for detector quality assessment and management.

[0028] Then, it is determined whether the sensor data contain errors or not. That is, based on extracted pattern, it is assessed whether the data contain errors or not.

[Aspect 3]

[0029] A third aspect deals with the issue of how to differentiate and classify the error types observed. That is, according to the third aspect, in case of errors/faults presence, these faults are classified. In this regard, it is proposed to robustly classify and identify OCCURRING error/faults types, the error types depending on the underlying sensor type.

[Aspect 4]

[0030] According to a fourth aspect, the type of vehicle observed is classified. That is, from the pattern collected out of non-erroneous data, the presence of either a car or a truck is classified (this classification is independent of the sensor type).

[Aspect 5]

[0031] A fifth aspect concerns the issue of how to perform erroneous detector data correction, whenever possible. According to the fifth aspect, while involving sensor type related error types knowledge, sensor topology, environmental information, and observed error types, the errors are robustly corrected and a cleaned data stream is generated.

[Aspect 6]

[0032] The sixth aspect relates to how to build a black box model of the error rate related detector's behavior in dependence of both environmental information (weather and temperature) and traffic volume.

[Aspect 7]

[0033] According to the seventh aspect, it is described how to perform short, middle and long term detector error rate performance prediction while involving the black-box model. That is, based on archived pattern history and

environmental conditions information we do predict detector faults rate for short, middle and long term. This does also give indication about the remaining detector life time.

[Aspect 8]

[0034] An eighth aspect deals with how to either automatically tune or recommend contextually appropriate detector 'sensitivity' values and/or 'measurement time' values and/or other equivalent or relevant settings in order to minimize detector's error rate performance.

[0035] Based on detector faults rate behavior, a detector tuning recommendation (that takes environmental information and a detector reliability physics model into consideration) is calculated and provided. The recommendation concerns measurement time and/or sensitivity level, and/or other equivalent detector settings.

[0036] The processing related to all of the above described 8 aspects is performed in real-time and does reach a detection rate above 96% according to very extensive tests that did involve real field data.

[0037] In the following, the aspects of the present invention will be described in more detail.

[Aspect 1]

[0038] Concerning the granularity of the observation of the raw data originating from the detector, solely the microscopic level is considered. With "microscopic level", it is referred to the level at which the detector raw data have been transformed into an information about presence or non-presence of a vehicle on the road. Thus, the atomic information (or atomic data stream) considered in the present invention and on which all further processing steps are based is that about

- a) vehicle presence within detector's range or region of interest, and
- b) relation to the respective timings, i.e. "start" and "end" of each vehicle presence.

[0039] No further form of aggregation is considered, contrarily to approaches in the related state-of-the-art, which mainly rely on aggregated data.

[0040] Therefore, the first step in the global detectors quality management is to transform the detector raw data into an "atomic data stream". For (rare) systems that do not explicitly generate an "atomic data stream" a pre-processing module for generating/guessing/synthesizing the related 'atomic data stream' can be designed. Such a module to be calibrated once by real sample observations may eventually contain some form of stochastic intelligence.

[Aspect 2]

[0041] The atomic data stream is to be observed and

analyzed in order to determine whether it contains errors or not. In this point, the overall methodology of this determination is presented.

[0042] There is involved a pattern recognition concept that involves the following considerations: sensor type, sensor network topology, weather information and traffic volume information. Hidden markov models, bayesian networks and support vector machines are involved in the overall process.

[0043] The output of this processing is information-related to the presence or not of errors in the observed atomic data stream. Either a single cross-section or a set of consecutive cross-sections is/are considered simultaneously depending on their logical interdependency as determined/fixed by the underlying sensor network topology.

[0044] In the following, a detailed specific example will be given. However, the present invention is not limited to the specific example.

[0045] Offline process Phase 1

1. Errorless reference data Stream collection or generation
2. Logical-topological correlation between neighbouring loops and sections + weather & environment information

[0046] A reference clean or errorless data stream of appropriate significant length is collected or generated. This corresponds to an observation from typical portions from several typical hours up to days.

[0047] Besides, all logical and topological correlation between neighboring loops and sections are recorded for a set of reference topologies or sensor network configurations.

[0048] This is done once for a given sensor type for different weather conditions.

Offline process Phase 2

[0049]

1. Transformation of reference streams into a set of images/pictures of different sizes while considering the logical-topological information
2. Transformation of the differently sized various images of (1) using Radon transformation (*)
3. Features extraction using a cellular neural network (CNN) processor (**)
4. Use a discrete-time CNN processor (***) or PCA (4*) to create reference classes for the various images sizes

[0050] The collected reference data stream is transformed into various images of different sizes (using sliding windows of different sizes) while involving the logical-topological information.

[0051] The different images are transformed though

nonlinear processing (radon transformation and CNN processing) out of which a set of reference classes is determined and saved.

[0052] This is done once for any given sensor type.

(*): Reference to radon transformation: <http://www.tmrfindia.org/ijcsa/v7i34.pdf>, last access Sept 14th, 2012

(**): Example of CNN based feature extraction, see: http://www.itk.ppke.hu/Szlavik_diss_ENG.pdf, last access Sept 14th, 2012

(***): Example of a discrete time CNN classifier: <http://www.ieice.org/proceedings/NOLTA2005/HTMLS/paper/7141.pdf>

(4*): Example of PCA based classifier: <http://www.tmrfindia.org/ijcsa/v7i34.pdf>, last access Sept 14th, 2012

Online Phase 3

[0053]

1. Online data stream collection
2. Remember the current logical-topological correlation between neighbouring loops and sections (from phase 1)
3. Remember the results of all processes of Phase 2

Online Phase 4

[0054]

1. Transformation of ONLINE streams into a set of images/pictures of different sizes while considering the logical-topological information
2. Transformation of the differently sized various images of (1/Phase 4) using Radon transformation
3. Features extraction using a cellular neural network (CNN) processor system
4. Use a discrete-time CNN processor system or PCA to classify and compare with the REFERENCE CLASSES of Phase 2

[0055] All processes of Phase 2 are done, but now on real-time/online collected data streams.

[0056] The discrete time CNN classification allows a comparison with the references classes.

```

IF
    all online classes fit to the reference
classes,
    THEN
        the data stream is error free;
    OTHERWISE
        There is one or more errors in certain
identified
    portions of the data stream.
    The portions containing errors are kept in

```

```

view
    for further processing.

```

[Aspect 3]

5

[0057] Portions of the "atomic data stream" containing errors are identified in and marked as such. In this point, marked atomic data stream portions are processed and errors pattern are differentiated and then clustered. A semi-supervised labeling involving for example rough-set theory is then performed offline by a human expert. This labeling depends on the underlying sensor type and is done once. For loop detectors, for example and just for illustrative purposes, the labeling will differentiate 6 error classes: weak-signal, pulse break-up, chattering, splash-over, over-counting and under-counting. The error classes are of course not limited to the above mentioned 6 error classes.

10

[0058] For other sensor types, the labeling will depend on their inherent physics and on the sensor network topological information.

[0059] The following concepts are involved in the error type classification: signal energy, statistical analysis, support vector machine and rough-set theory.

15

[0060] In the following, a detailed specific example will be given. However, the present invention is not limited to the specific example.

Offline Phase 1

20

[0061]

1. Record and/or generate a set of atomic data streams of different lengths and containing all possible known types of errors. Using "Aspect 2" to observe real traffic can help recordings from the field
2. Consider thereby different logical-topological correlations between neighbouring loops and sections + weather and environment information
3. Use, once, a semi-supervised labeling of the different error types by a human expert. The possible error types depend on the current sensor type (or sensor types combinations)
4. Extract the signatures of all identified error types by using SVM or better a CNN processing. An alternative is a statistical measure over sliding windows.
5. Save/record the extracted signatures for use online.

25

30

35

40

45

50

55

[0062] The offline phase 1 is used to record reference "error containing data streams or data stream pairs" of different lengths while taking the following into consideration: reference logical-topological contexts, weather and environmental conditions and sensor type or sensor types combinations.

[0063] A clustering of the different error types is performed by involving for example rough-set theory.

[0064] A human expert is used once in order to label

the different types/classes of errors observed/identified/known. The latter is done once for every sensor type or sensors type combinations.

[0065] The reference data streams being known for each labeled error types, corresponding signatures are fixed/extracted by using either a nonlinear feature extractor like CNN or by a statistical measure performed over sliding windows.

[0066] The signatures related to each error type are saved for further use in the online phase.

[0067] This OFFLINE phase is performed once and the output is the set of signatures for all relevant error types for a given sensor type or sensor types combination. These signatures will then be used in phase 2 (ONLINE) in order to detect/ classify errors (types).

Online Phase 2

[0068]

1. Identify the error containing sets of atomic data streams detected in (Aspect 2/Phase 4)
2. If not yet done, extract relevant features by involving preferably CNN or some other appropriate feature extractor
3. Use a robust classifier (e.g. CNN based or SVM baser or other) to detect/identify the underlying error type
4. Alternative to (3):

After the error signatures are known, this error type classifier can be run directly on the online data stream obtained after Aspect 2/Phase 2/(1)

[0069] The phase 2 is an online process. It assumes that the signatures of all possible errors are known (while considering related logical-topological and weather + environment information).

[0070] Thus, a robust classifier (CNN based, SVM based, or other) is used to detect the underlying error within a given data-stream portion.

[0071] If the data stream portion has been provided by Aspect 2/Phase 4, the related features have been already extracted by Aspect 2 and can be provided to Aspect 3/Phase 2.

[0072] An alternative is, provided Phase 1 of Aspect 3 has been done once and thus has already provided the signatures of all possible errors, to skip all steps of Aspect 2 coming after Aspect 2/Phase 2/(1). The online data streams of different lengths would then be collected from here and processed for error classification. In this case Aspect 2 is merely de facto skipped and considered or integrated as an offline part of Aspect 3.

[0073] This Online phase 2 can be directly performed on the online data-stream while skipping most of the online phase (Phase 2) of Aspect 2 as explained in the note "Alternative to (3)".

[Aspect 4]

[0074] Clean portions of the "atomic data stream" (i.e. not containing errors), originated from a single detector system, are processed in order to classify the type of vehicle observed. This classification relates to vehicle length and respective headways as sole differentiating features. Thus, it is mainly differentiated between trucks and cars. Further differentiations are possible so far they are related to vehicle length.

[0075] The main classification is based on the features sets constituted by the presence data pulse lengths and the related distribution of the respective headways. A shock wave pattern is then extracted and involved in a support vector machine based classification.

[0076] The classification output is whether the vehicle is a car or a truck, and whether it has been in movement or stopping, the latter for example due to red traffic light phase.

[0077] In the following, a detailed specific example will be given, however, the present invention is not limited to the specific example.

Offline Phase 1

[0078]

1. Select explicitly or generate, similar to Aspect 2/Phase 1, error free reference data streams of different lengths, where, through human supervision, one knows there are trucks inside those streams.
2. Different contexts should be considered hereby. Additionally to those of Aspect 2, the consideration of the presence or not of stops at red traffic lights and that of different traffic conditions is mandatory
3. Then apply the processes described in Phase 2 of Aspect 2 and obtain the various signatures of the truck's presence in data streams of various lengths
4. Save the obtained signatures for use in the online truck detection of Phase 2

[0079] This phase 1 is an offline process. It is done once and intends to produce the signatures of a truck presence within the atomic data streams while taking related context information into consideration. The reference data streams involved here must be error free.

[0080] The obtained various signatures of the truck's presence in a traffic stream are then saved for use in the online truck detection process.

[0081] Since sliding windows are used, the exact position of the truck within the stream will also be determined. That is, the presence pulse corresponding to a truck will be explicitly identified. The signatures are formed accordingly.

[0082] This offline phase results in a set of signatures indicating the presence and the position of a truck within a traffic stream from different perspectives.

Online Phase 2

[0083]

1. Extract current online data streams of different lengths. These atomic data streams should be error free.
2. Extract relevant features by involving an appropriate nonlinear feature extractor, preferably CNN
3. Then use a classify involving the different truck signatures obtained in Aspect 4/Phase 1. Any classifier can be used, but preferably a CNN based one
4. The result will be:

- a. That the data stream contains a truck or not
- b. In the positive case, the position of the truck

[0084] This phase 2 is an online process.

[0085] The online data streams are assumed to be error free. If they are not error free, the stream should be taken after error correction or the portions containing errors should be skipped for the truck detection.

After

[0086]

- a) the selection of appropriate data streams,
- b) the related features extractions (preferably through CNN), and
- c) a classification involving the pre-stored signatures obtained from Phase 1 (of Aspect 4),

[0087] it will be assessed whether a given data stream contain a truck or not. In the positive case, even the position will be determined. Thus, the exact presence pulse corresponding to a truck will be identified. A CNN based classifier is preferred due to the high performance. Other classifiers can however also been used.

[0088] This online phase detects the presence of a truck within a traffic stream and gives also the exact position. The atomic 'presence pulse' corresponding to a truck will be identified.

[0089] This can be done in real-time within less than a second and can thus be considered, amongst others, in highly precise adaptive future traffic control systems

[Aspect 5]

[0090] Error correction is based on a reasoning based on some prior knowledge formulated logically by a human expert and that depends on error classes related to a given detector type and on fixed sensor network topological configurations. The topological information is related to how detectors lay as a group in a common cross-section or are situated with a given spatial separation within a direct longitudinal neighborhood.

[0091] The prior knowledge that expresses the es-

sence of the rules for the real-time, logic-programming-based reasoning correction is based on a plausibility matrix fixing the possible either simultaneous or consecutive error occurrences depending on the underlying topology and while considering the detector-type related variety of error classes.

[0092] The output of the correction process is, for each given error (detected by the concept of Aspect 3), that:

- a) whether a correction is possible or not;
- b) in cases a correction is possible, a corrected/clean "atomic data stream" is generated

[0093] In the following, a detailed specific example will be given; however, the present invention is not limited to the specific example.

Offline Phase 1

[0094]

1. The error correction is based on:

- a) The logical-topological information (see Aspect 2)
- b) The different error types identified in Aspect 3 and especially how they correlate with the logical-topological information

2. Based on (1) of Aspect 5 and on information gathered from Aspect 2 and Aspect 3, a human expert will define logic-programming rules that will determine:

- a) If a detected error can be SURELY corrected
- b) In the positive case (i.e. a correction is possible), how it should be surely corrected
- c) In the negative case:

- i) Formulate a default recommendation (some examples for illustration: e.g.: delete the pulse(s); or set erroneous pulse to normal; or replace erroneous long pulse by 2 normal pulses, etc.)
- ii) Or state that even a default recommendation is not possible due to lack of information

3. The defined logic-programming rules will be integrated in a logic-programming REASONER to be used online for performing error corrections

[0095] This phase 1 is an offline process.

[0096] In essence, knowledge from Aspect 2 and Aspect 3 is used by a human expert to formulate logic-programming rules that are then integrated in a logic-programming reasoner to be used in real-time for error cor-

rection.

[0097] The reasoner will involve an appropriate ontology that includes the context information indicated in Aspect 2: logical-topology, weather and environment, data streams lengths, error types, etc. The ontology and the reasoner will be both designed during this offline phase.

[0098] Due to real-time constraints a solver based on Answer Set Programming (ASP) is recommended, although any other solver can be used.

[0099] How probable each default recommendation is, is determined by extensive offline observation tests.

[0100] This OFFLINE phase will result in a reasoner to be used for online corrections and/or default recommendations for cases a SURE correction is not possible

Online Phase 2

[0101]

1. Get the error containing atomic data stream portions (see Aspect 3)

2. Use the REASONER designed in Phase 1 of Aspect 5 to determine:

a) If a the detected error can be SURELY corrected

b) In the positive case (i.e. a correction is possible), the reasoner will corrected it

c) In the negative case:

i) Depending on current context information, the reasoner may formulate a default recommendation and assign it a probability value, which was obtained in Phase (1) of Aspect 5

ii) Or the reasoner will just state that even a default recommendation is not possible due to lack of information

3. A new corrected data stream may be generated or the old one may be maintained but just augmented with the reasoner's output obtained in (2) of Aspect 5/Phase2. This means that related labels will be added to the erroneous portions. This may be of relevance for those functions using the sensor data upwards.

[0102] This phase 2 is an online process.

[0103] In essence the reasoner built in Phase 1 of Aspect 5 will be used online on error containing atomic data stream portions.

[0104] The reasoner will determine whether the error can be corrected. If yes, it will correct it. If not it will either state a default recommendation or state that neither default recommendation nor correction is possible due to lack of information.

[0105] Each default recommendation should be ended with a probability value. This latter value is obtained

offline as explained in Phase (1) of Aspect 5.

[0106] The reasoner will preferably involve a solver based on ASP (Answer Set Programming) to ensure a fast speed under real-time constraints.

[0107] This Online phase involves the use of an ultra-fast logic-programming reasoner for online error corrections and/or for default recommendations for cases a sure correction is not possible.

[Aspect 6]

[0108] A black-box model of the detector's performance is constructed with the objective of probabilistically describing the error rate related detector's behavior. In the essence, the intention is to determine the detector's reliability physics in dependence of both internal and external parameters.

[0109] Internal detector parameters are different settings determining and/or related to its core functioning principle. External parameters are traffic volume, detector's network topology, weather and environmental related ones.

[0110] To construct the detector's black-box model a combination of the following probabilistic instruments is involved: structural equation modeling, bayesian (belief) networks, and hidden markov models.

[0111] In the following, a detailed specific example will be given; however, the present invention is not limited to the specific example.

1. Collect a significant amount of real historic field data (containing also classified faults positions) according to parameters of Fig. 3, i.e. related inputs and outputs

2. Use an appropriate statistical method to assess which of the inputs do have a significant impact on the error/fault rates of different types. Example of method: SEM (structural equation modeling), etc.

3. Suppress from the model of Fig. 3 those input parameters with meaningless impact (according to results of (2)) on the rates of the outputs

4. Use now an appropriate stochastic modeling instrument to build the probability reliability physics model involving the remaining input parameters of Fig. 3. This black box model building (training) involves both historic data (see (1) and the findings from (2). Examples of appropriate methods: bayesian networks, fuzzy logic, neural networks, (hidden) markov models, etc.

5. The obtained black box model is saved for use latter in either occasional or online dynamic/automatic optimal setting of the sensor systems in dependence of current external and internal conditions (i.e. the remaining significant inputs of Fig. 3.)

6. This process must be done once for every sensor system type, e.g.: inductive loops, video, radar, etc.

[0112] In essence, this is an offline process aiming at

assessing the sensor's reliability physics in dependence of both internal and external parameters.

[0113] Potential/candidate internal parameters are for example: sensor sensitivity level, measurement time level, etc. These parameters may differ depending on the specific sensor system type.

[0114] Potential/candidates for external parameters are: weather, temperature, traffic level, proportion of trucks in the traffic, etc.

[0115] This offline process results into a sensor system type specific reliability physics black box model. This model will be used later for online detector tuning in dependence of both current internal and external parameter conditions

[Aspect 7]

[0116] A robust detectors' quality management requires the capability to predict the detector's reliability in three time horizons: short term (range of some minutes up to one hour); middle term (range of some hours, up to some days, and up to a week); long term (range of several weeks, up to several months, and up to some years).

[0117] The short term prediction is of relevance for managing the impact of detector quality on traffic management. The middle term prediction is mainly of relevance for short term traffic infrastructure maintenance activities. And the long term prediction is mainly of relevance to predict the operationally acceptable lifetime of a given detector or detector group.

[0118] Overall, the prediction does involve the prior knowledge that is compressed in the black-box model obtained in Aspect 6. Further, for both short-term and middle term detector's-performance prediction, a concept involving mainly hidden markov models and the black-box model is used. For both short and middle term, all external parameters are either known or can be easily predicted (e.g. traffic volume) or obtained from external sources (e.g. weather information).

[0119] For a long-term prediction, however, external parameters are hardly known. Thus, here another prediction approach is used. It does rather predict an envelope that fixes with a high reliability, as proven by extensive tests involving real data, the upper-bound of the error-rates' detector performance. Thus, the real performance lies below the envelope constituted by the upper-bound. The available detector's data history to be involved in the long-term prediction should be at least two times longer than the future time-frame ahead for prediction. For example, to predict for the next three weeks (or months), the available observed history should be longer than 6 weeks (or months). Thus, the appropriate level of data aggregation should be selected in a way that the available history should satisfy this length requirement.

[0120] The envelope's prediction involves two major steps: a) time series filtering through a quadratic trend estimation and a consecutive smoothing; and b) estimate

a cubic trend of the later-resulting time series and use it for future prediction.

[0121] In the following, a detailed specific example will be given, however, the present invention is not limited to the specific example.

Phase 1 Short Term Prediction

Step 1 Offline Short Term Prediction

[0122]

1. Collect a significant amount of real historic field data according to parameters of Fig. 3, i.e. related inputs and outputs
2. Observe (see Aspect 6/(2) & (4)) and note how far historical data on traffic volume is associated with error rates
3. Observe (see Aspect 6/(2) & (4)) and note how far historical data on weather and temperature are associated with error rates

Step 2 Online Short Term Prediction

[0123]

1. Use a stochastic state space model (for instance Kalman or particle filter, etc.) to predict the traffic volume. Use the stored historical data of traffic volume to train the model
2. Train a probabilistic model to model the dynamic error rate in dependence of traffic volume, weather and temperature.
3. Use the trained probabilistic model (e.g. HMM) to predict the error rate using the predicted traffic volume (Online step 1) and the forecasted environmental conditions (temperature, weather)
4. Future time horizon: a couple of hours: 1-4 hours in the future

[0124] In essence, this process is aiming at predicting the error rate in order to evaluate the SHORT TERM performance of the sensor system. The aggregation time unit is hours; one predict for the next couple of hours (1 to 4 hours). The black box reliability physics model is taken into consideration. Thus, the short term prediction considers the history of error rates associated with the conditions (weather, temperature and traffic volume). State space models (stochastic and probabilistic) are proposed for this prediction. For instance, hidden Markov model (HMM), Bayesian network and Kalman filter.

[0125] The sensor system internal parameter tuning control system (see Aspect 8) will need the prediction result as one of the inputs if a tuning assessment for the corresponding time horizon is needed.

[0126] This online short term prediction model covers the next couple of hours. It takes the sensor system reliability black box model into consideration. State space

predictions models are appropriate to implement this prediction. We do favor a HMM combined with Kalman Filter.

Phase 2 Middle Term Prediction

Step 1 Offline Middle Term Prediction

[0127] 1. Collect a significant amount of real historic field data according to parameters of Fig. 3, i.e. related inputs and outputs

Step 2 Online Middle Term Prediction

[0128]

1. Estimate a logarithmic local trend for error rate historical data
2. Estimate the trend error deviation
3. Build a model that combines the logarithmic trend with the error deviation
4. Use the estimated model to predict the upcoming error rate
5. Time horizon: a couple of days or a couple of weeks; aggregation level: days or respectively weeks

[0129] In essence, this process is aiming at predicting the error rate in order to evaluate the MIDDLE TERM performance of the sensor. The aggregation time unit is days and weeks. According to these aggregations, the error rate time series is smoother than in short term, and the environmental conditions have due to the aggregation a relatively smaller impact on error rates.

[0130] The proposed middle term prediction uses a stochastic state space model (e.g. Auto Regression Moving Average Model, Local Trend Model, etc.).

[0131] The sensor system internal parameter tuning control system (see Aspect 8) will need the prediction result as one of the inputs if a tuning assessment for the corresponding time horizon is needed.

Phase 3 Long Term Prediction

Step 1 Offline Long Term Prediction

[0132]

1. Collect a sufficient amount of real historic error rate data for a given sensor system

Step 2 Online Long Term Prediction

[0133]

1. Extract from this the worst cases of error rate (i.e. all data points that are higher than the quadratic trend)
2. Estimate the logarithmic trend of the worst cases

from historical data constituted of the extracted points from (1)

3. Predict the future of the worst case based on the logarithmic trend

4. Future time horizon: 3 months and more; aggregation of data points: in days or in weeks

[0134] In essence, this process is aiming at LONG TERM predicting the error rate in order to estimate amongst others the lifetime of a given sensor.

[0135] There is a specific level of the predicted error rate for which the sensor is considered unreliable. This threshold should be sensor system type specific. If for example an average of 10% error rate (@ 500 vehicles) is no more acceptable, then it will be fixed as a threshold on which basis the potential remaining life time of a sensor can be estimated based on the long term prediction functionality.

[0136] A future 'time horizon' of 3 months and more is taken, for example.

[0137] The Long term prediction model estimates the trend of the worst cases of error rates. It involves similar performance worst cases observed in the recent related history of the historical data.

[0138] This online process enables a future prediction of the sensor system performance. It allows to know in advance whether the sensor system may reach performance regions that are unacceptable.

[Aspect 8]

[0139] The extensive study of the detectors' reliability physics has shown that their errors rate performance depends on both traffic and environmental conditions on one hand, and on internal detector settings on the other hand.

[0140] This close interdependency has been expressed within the black box model (Aspect 6).

[0141] The core question here is, however, the following: while knowing both environmental and traffic conditions, how to determine the appropriate internal detector settings in order to minimize the detector's error rate for a given short-term future observation window?

[0142] This question is equivalent to the inverse problem to the black box modeling.

[0143] This inverse problem is solved by involving mainly two complementary methods: bayesian networks and fuzzy logic programming.

[0144] Through this concept it is therefore possible to:
a) formulate/determine in an offline mode appropriate recommendations for maintenance technicians about the optimal settings of internal detector's parameters depending on underlying conditions (traffic and environmental); and b) automatically, and in an online mode, set and control the detector settings through an embedded platform/scheme. The automatic adaptivity of internal detector's settings will significantly improve the detector's performance quality despite volatile/varying external un-

derlying hard conditions.

[0145] In the following, a detailed specific example will be given; however, the present invention is not limited to the specific example.

Phase 1 Offline / Model Training

[0146]

1. Involve data from Aspect 6/(1)
2. Take the black box model obtained in Aspect 6/(5); also involve findings of Aspect 6/(2)
3. Combine data from (1) and the models and findings in (2) to perform extensive stochastic evaluations and/or observations for various scenarios.
4. The observations of (3) will be used to train the probabilistic black box model of Fig. 4. Possible appropriate instruments to realize/implement this black box model are: bayesian networks, (hidden) markov models, fuzzy logic controller, etc. Extensive experiences have shown that all instruments perform quite well; we do however prefer/favor the fuzzy logic controller

[0147] In essence, this process is a partial inverse problem of the black-box model reliability physics model of Aspect 6/(5). See Fig. 4.

[0148] The reliability physics model is an input used to extensively train the new sensor system internal parameter tuning control model of Fig. 4.

[0149] The tuning can be calculated for the current time or for a future time interval. Respective related inputs will be then needed.

[0150] This online sensor system tuning model controller is in essence a REASONER that builds on knowledge from the sensor system reliability physics model. It ensures an adaptivity of the sensor system to keep the error rates performance the lowest possible at any moment.

Explanations of key words

[0151] In the following, some key words and expressions used in the above description are explained for better understanding of the principles of the invention.

Point sensors:

[0152] A point sensor is typically stationed/placed at a fixed location along/on/under a roadway and watches (i.e. counts, records, classify, etc.) vehicles passing at this particular location in the time domain. The fixed location may be:

- (a) a complete cross-section of the roadway at a given location, or
- (b) just a single lane of the cross-section.

[0153] Examples of point sensors are: loop detectors, radars, ultrasonic radars, Lidars (light detection and ranging) and video cameras.

[0154] Area sensors or Space sensors:

An area sensor typically covers/watches a wide surface/area of a roadway or road junction. It is generally placed high in the sky/air over the ground surface and takes snapshots of traffic at an instant of time. It does therefore simultaneously cover/watches many cross-sections and many lanes.

[0155] Examples of area sensors are: aerial photography or satellite imagery, and mobile sensors such as automatic vehicle location and global positioning system.

Loop detectors systems:

[0156] The inductive-loop detector is the most utilized sensor in a traffic management system. The principal components of an inductive-loop detector are:

- One or more turns of insulated loop wire wound in a shallow slot sawed in the pavement
- Lead-in cable from the curbside pull box to the intersection controller cabinet
- Detector unit: Electronics unit housed in a nearby controller cabinet

[0157] Cars passing over or that stop within the detection area of an inductive-loop detector do decrease the inductance of the loop. The electronics unit senses this event as a decrease in frequency and sends a pulse to the controller signifying the passage or presence of a vehicle.

Video-based traffic sensor systems:

[0158] Video-based detectors use advanced image processing schemes running on an appropriate micro-processor to analyze a video image input. Different approaches are used by video detection sensors. Some analyze the video image of a target area on the pavement. A change in the image of the target area as a vehicle passes through the target area is analyzed. Another approach does rather identify when a target vehicle enters the camera field of view and tracks the target vehicle through this field of view. Finally, other video sensors use a combination of these two approaches.

Radar-based detector system:

[0159] A radar operates according to the following principle: "The radar dish or antenna transmits pulses of radio waves or microwaves which bounce off any object in their path. The object returns a tiny part of the wave's energy to a dish or antenna which is usually located at the same site as the transmitter" (Source: [http://en.wikepe-](http://en.wikipedia-)

dia.org/wiki/Radar; latest access: Sept. 10th, 2012). Through an appropriate processing of the reflected energy the radar can be used to determine presence, range, altitude, speed and direction of targeted objects. In traffic detection, the radar is used as a point sensor which can detect both presence and speed of a vehicle in a fixed sport (or cross-section) on the road.

Detector's signal raw data:

[0160] By "detector's signal raw data" is meant the time series generated by a point detector. It is constituted of the succession of vehicle presence detection pulses. The time series contains beginnings and ends of the successive pulses, which are separated by so-called time headways. An area sensor will generate simultaneously many of such time series for fixed areas or locations.

Black-box model:

[0161] Commonly a black box model can be defined as follows. A black box is a module with known inputs, known outputs, a known function but with an unknown internal mechanism. A black box module: a) acts predictably, b) can be used without knowledge of its internal details, c) hides information from the rest of the system. Thus, a black box is defined by "what" it does and not by "how" it does it. For building a black box model one has to start from measurements of the behavior of the system and the related external influences (inputs to the system) and try to determine a mathematical relation between them without going into the details of what is actually happening inside the system.

Atomic information:

[0162] Atomic information refers to a minimal amount of information which makes it possible to distinguish two particular items. Atomic information is monovaluate and does refer to only one thing. Failure to capture atomic information means inaccurate, unreliable or missing data.

Vehicle presence:

[0163] In traffic management, for the planning of a junction one needs to know about the presence or not of a vehicle at a set of fixed locations on the road. A location is a generally rectangular area on the lane surface of approx. 1m x 2m. The selection of these positions serve different needs, for example a) provide a trigger or parameter on which a vehicle actuated traffic management system is based for extending the current green time or for giving green to a particular approach; b) measure queue length; etc.

Hidden markov models:

[0164] According to "https://cwiki.apache.org", HMM is defined as follows:

"A Hidden Markov Model (HMM) is a statistical model of a process consisting of two (in our case discrete) random variables O and Y , which change their state sequentially. The variable Y with states $\{y_1, \dots, y_n\}$ is called the "hidden variable", since its state is not directly observable. The state of Y changes sequentially with a so called Markov Property. This means, that the state change probability of Y only depends on its current state and does not change in time. "

[0165] Further, the Hidden Markov Model (HMM) is a variant of a *finite state machine* that does have a set of: a) hidden states, Q ; b) an output *alphabet*(observations), O ; c) transition probabilities, A ; d) output (emission) probabilities, B ; e) and initial state probabilities, M . The current state is not observable. Instead, each state produces an output with a certain probability (B). Usually the states, Q , and the outputs, O , are understood; thus an HMM is said to be a triple, (A, B, M) .

Sources for detailed information:

[0166]

- Lawrence R. Rabiner (February 1989). "A tutorial on Hidden Markov Models and selected applications in speech recognition". *Proceedings of the IEEE* 77 (2): 257-286. doi:10.1109/5.18626.
- Hidden Markov Models, Theory and Applications. Edited by Przemyslaw Dymarski, ISBN 978-953-307-208-1, Hard cover, 314 pages, Publisher: InTech.

Bayesian networks:

[0167] A Bayesian network (or - other naming of the same: Bayes network, belief network, Bayes(ian) model, probabilistic directed acyclic graphical model) is a probabilistic graphical model (a type of statistical model) that represents a set of random variables and their conditional dependencies via a directed acyclic graph. For example, a Bayesian network could represent the probabilistic relationships between diseases and symptoms. Given symptoms, the network can be used to compute the probabilities of the presence of various diseases (Source: Wikipedia).

[0168] In other words, Bayesian nets are a network-based framework for representing and analyzing models involving uncertainty. They are different from other knowledge-based systems concepts because uncertainty is handled in mathematically rigorous yet efficient and simple way.

[0169] The general probabilistic inference problem here is in fact to find the probability of an event given a set of evidences. This can be done in Bayesian nets with sequential applications of Bayes Theorem.

Support vector machines:

[0170] Support vector machines (SVM) are concepts used for classification. They belong to the class of supervised learning models with associated learning algorithms that analyze data and recognize patterns. The fundamental SVM takes a set of input data and predicts, for each given input, which of two possible classes forms the output. Thus, SVM is a non-probabilistic binary linear classifier.

[0171] Given a set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that assigns new examples into one category or the other.

[0172] An SVM model is a representation of provided examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on.

[0173] Further, SVMs can efficiently perform non-linear classification using what is called a kernel function, and thereby implicitly mapping their inputs into high-dimensional feature spaces.

Fuzzy logic programming:

[0174] Fuzzy logic is an approach to computing based on "degrees of truth" rather than the usual "true or false" (1 or 0) Boolean logic on which the modern computer is based. Fuzzy logic includes 0 and 1 as extreme cases of truth (or "the state of matters" or "fact") but also includes the various states of truth in between (i.e. very unlikely' (gray areas of probability)). Software based on application of fuzzy-logic (as compared with that based on Formal Logic) allows computers to mimic human reasoning more closely, so that decisions can be made with incomplete or uncertain data.

[0175] A fuzzy inference system is a system that uses fuzzy set theory to map inputs (features in the case of fuzzy classification) to outputs (classes in the case of fuzzy classification). In other words, a fuzzy inference is an extended form of formal inferences that enables qualitative and/or quantitative evaluation of the degree of confidence level for a given causality on the basis of fuzzy expressions.

Weak-signal:

[0176] It is a particular type of faulty detection signal for vehicle presence detection. The presence detection of a vehicle at a fixed location is generally expressed at the detector/sensor output in form of a quadratic pulse

in time. The width of this quadratic pulse does generally correlate with the speed of the vehicle which presence is detected. Thus, a fast vehicle will generate a pulse with a smaller width; a long vehicle like a truck will result in a longer pulse width.

[0177] The generated 'presence pulse' (from a given sensor/detector) is named "weak signal" if its width is too small/short and thereby does significantly deviates from what should be expected normally (i.e. according to the current context related to maximum speed, current speed and average vehicle length).

Pulse break-up:

[0178] It is a particular type of faulty detection signal for vehicle presence detection. The generated 'presence pulse' or rather 'pulse-series' (from a given sensor/detector) is named "pulse breakup" if instead of one pulse, two consecutive pulses are generated whereby one or both of the following situations are additionally true:

- The time separation between the two pulses, we call it headway, does significantly deviate from what should be expected normally (i.e. according to the current context)
- One of the two pulses is a 'weak signal'

Chattering:

[0179] It is a particular type of faulty detection signal for vehicle presence detection.

[0180] The generated 'presence pulse' or rather 'pulse-series' (from a given sensor/detector) is named "chattering" if instead of one pulse, multiple (more than three) consecutive pulses are generated whereby both of the following situations are additionally true:

- The time separation between two consecutive components of the pulse-series, the headway, does significantly deviate from what should be expected normally (i.e. according to the current context)
- All or at least more than two of the pulse-series are 'weak signals'

Splash-over:

[0181] It is a particular type of faulty detection signal for vehicle presence detection.

[0182] The generated 'presence pulse' (from a given sensor/detector) is named "splash-over" if both of the following are true:

- It is generated rather on (or for) a directly/closely neighboring lane (in cross-section) and not on the lane on which the vehicle is effectively present or passing
- It is either a "weak signal" or a "pulse break-up"

Over-counting:

[0183] It is a particular type of faulty detection signal for vehicle presence detection.

[0184] It is a particular form of splash-over or in other words a "normal pulse" splash-over. The generated 'presence pulse' (from a given sensor/detector) is named "over-counting" if both of the following are true:

- It is generated rather on (or for) a directly/closely neighboring lane (in cross-section) and not on the lane on which the vehicle is effectively present or passing
- It is a normal signal (i.e. its width is normal according to the current relevant context) and neither "weak signal" nor "pulse breakup"

Under-counting:

[0185] It is a particular type of faulty detection signal for vehicle presence detection.

[0186] It is a particular anomaly of the detection signal. Instead of generating a separate pulse for each individual vehicle passing a sensor, the sensor/detector does generate a single presence pulse for 2 or more consecutive vehicles passing by. The resulting pulse length will be higher than normal and an undercounting of the effective number of vehicles will be a further consequence. This generally happens when the headways between slowly moving vehicles is relatively small, as it is for example the case during traffic jams.

Rough-set theory:

[0187] Rough set theory is a relatively new mathematical tool for imperfect data analysis. The foundation of the rough set philosophy is the assumption that to every object of the universe of discourse some information (data, knowledge) is associated. Objects characterized by the same information are indiscernible (similar) in view of the available information about them. The indiscernibility relation generated in this way is the mathematical basis of rough set theory.

[0188] Rough set theory can be approached as an extension of the classical set theory, for use when representing incomplete knowledge. Rough sets can be considered as sets with fuzzy boundaries; that is, sets that cannot be precisely characterized using the available set of attributes. The basic concept of the rough set theory is the notion of approximation space.

Detailed information, see:

[0189]

<http://www.nit.eu/czasopisma/JTIT/2002/3/7.pdf> (last access: Sep. 12th, 2012)

Signal energy:

[0190] The signal considered is the one formed by the succession of the different consecutive presence pulses over time over a given location at a cross-section. We call the energy of a signal (the above mentioned signal is meant) the area under the squared signal.

Shock wave pattern:

[0191] These pattern do reflect within the 'vehicle presence pulses' series over time the well-known traffic wave phenomenon. Traffic waves, also called stop waves or traffic shocks, are travelling disturbances in the distribution of cars on a highway. Traffic waves usually travel backwards in relation to the motion of the cars themselves, or "upstream".

[0192] The shock wave pattern is a particular distribution of both consecutive pulse-lengths and headways that do reflect the existence of a traffic wave in the underlying traffic stream.

Time series filtering:

[0193] Each time series data that reflects the output of an experiment or coming out as a signal from a dynamical system will expectedly contain some amount of noise that is embedded in it. The analysis of such data in presence of noise may often fail to give accurate information.

[0194] A method for filtering a time series data is a tool to clean as much noise from it as possible as the data should be made compatible for further analysis.

Quadratic trend estimation; cubic trend estimation:

[0195] Trends are patterns of a variable over time. In mathematics, the term quadratic describes something that pertains to squares, to the operation of squaring, to terms of the second degree, or equations or formulas that involve such terms.

[0196] If one looks at a time series, the first thing that stands out is the obvious tendency of the series to grow or to fall over time.

[0197] That is, it is immediately apparent from a time series plot that the average change in the series is either positive or negative. This tendency is the series's trend.

[0198] A general approximation of the trend of a time series is expressed by the following polynomial trend model:

$$T_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \dots + \beta_p t^p$$

where p is a positive integer. The coefficients β_i are obtained from a regression process over the underlying time series.

[0199] Here, of p=1, the trend is called "linear". How-

ever, if $p=2$, the trend is called "quadratic". And if $p=3$, the trend is called "cubic".

Detector sensitivity:

[0200] A sensor's sensitivity indicates how much the sensor's output changes when the measured quantity changes. For many sensor systems, the possibility exists of varying or setting the sensitivity, either manually or automatically.

Detector measurement time:

[0201] This is the effective time that the measurement process within a sensor or a detector is exposed to the physical environment from which a given parameter is measured. A high measurement time will result in a lower measurement update frequency.

Microscopic data:

[0202] In the context of traffic sensing microscopic data refer to single-vehicle related data. In this case here, these will be the time series of individual vehicle presence pulses generated by the traffic detectors.

Microscopic traffic analysis means analyzing the behavior of the individual driver. The characteristics of individual driving (i.e. microscopic data) can be described with the following variables: speed, acceleration, time headway, distance headway, relative speed between two consecutive vehicles.

Macroscopic data:

[0203] In the context of traffic sensing, macroscopic data refer to an aggregation of single-vehicle related data. In this case here, these will be the time series of aggregated numbers of individual vehicle presences per time unit, for example per hour.

[0204] The macroscopic quantities are the flow rate (the number of cars passing a cross section per unit of time, q), the density (the number of cars per unit of distance, k), and the space mean speed.

Mesoscopic data:

[0205] Instead of time-aggregated data, characteristics of individual vehicles can be obtained for cross sections. Then the characteristics of the distributions of the measured data can be examined (mesoscopic traffic flow characteristics) which also differ by flow condition. For example, the distributions of time headways are well studied because they can, for instance, be used to estimate capacity.

Pattern recognition (in the context here):

[0206] A pattern is a particular signature that is expressed by 'elementary features' of an object. In the context here, the object is the time series of consecutive detector pulse signals. In this context the 'elementary features' are for example a part or all microscopic data attributes or a combination of them.

[0207] Informally, a pattern is defined by the common denominator among the multiple instances of an entity. The pattern of relevance here are for example the different faulty signal types.

Detector Tuning (in the context here):

[0208] Most detector systems do offer the possibility of setting, manually or dynamically/automatically two key parameters of the sensor system or their related equivalent: sensitivity, measurement time, and eventually some more, depending on the specific sensor system.

[0209] Tuning refers to setting (manually or dynamically or a recommendation) these key parameters in ranges that put the sensor systems in the best sensing conditions resulting in the best sensor performance and consequently in the lowest error/faults rate.

Classification error:

[0210] The classification error is a measure of the misclassification rate or proportions.

[0211] It is known a measurement error is the difference between the true value of a measurement and the value obtained during the measurement process. A classification error is a type of measurement error by which the respondent does not provide a true response to a survey item. This can occur in one of two ways: a false negative response or a false positive response. A 'false negative response' corresponds to when the measurement system indicates that an event did not occur although it did.

[0212] In summary, the quintessence of the overall concept is presented in Fig. 5.

[0213] Thus, in view of the above, there is presented a sensor reliability data model that is trained offline to gather all significant characteristics of the sensor component, the physical object in the road surface/concrete, environmental influences, all this for enabling a reliable assessment of the data quality of the single sensor. Considering the topological information including lane direction, signaling and the relative position of the sensors to each other at a complete intersection, mutual interferences are also trained in the sensor reliability data model.

[0214] The resulting model or rather the set of models offers for the first time the possibility of monitoring (in real- or quasi real-time and proactively) both the current state and the evolution of sensor data quality especially for sensors delivering a digital pulse stream caused induced by the presence of vehicles in traffic.

[0215] Thus, in view of the above, a global and comprehensive concept for a robust "traffic detectors' quality management" is presented.

Claims

1. A signal processing apparatus for processing at least one detection signal, the apparatus comprising:

a receiver configured to receive the at least one detection signal from at least one sensor,
a model generator configured to generate a model of signal sequences based on previously received detection signals while taking into consideration a topology of the sensors and/or environmental condition in the vicinity of the at least one sensor,
comparing means configured to compare the at least one detection signal with the model of signal sequences generated by the model generator; and
a prediction means configured to predict quality of the at least one detection signal based on a comparison of the at least one detection signal with the model generated by the model generator.

2. The apparatus according to claim 1, further comprising
determining means configured to determine whether or not the at least one detection signal comprises an error, and
a classification means configured to classify the error into a plurality of predetermined error classes, if it is determined that the detection signal comprises an error.

3. The apparatus according to claim 2, further comprising
correcting means configured to correct the error, if it is determined that the detection signal comprises an error.

4. The apparatus according to claim 2 or 3, further comprising:

storing means configured to store a history of the detection signal, and
prediction means configured to generate a detection signal error rate prediction based on the stored history.

5. The apparatus according to claim 4, wherein the prediction is at least one of a short term prediction, a middle term prediction, and a long term prediction.

6. The apparatus according to claim 4 or 5, further com-

prising:

determining means configured to determine settings of the sensor based on the detection signal error rate prediction such that the detection signal error rate becomes minimum.

7. The apparatus according to any one of claims 1 to 6, wherein

the quality of the detection signal includes at least one of an overall error rate of the detection signal, and an error rate of each of the error classes of the detection signal.

8. A method for processing at least one detection signal, the method comprising:

receiving the at least one detection signal from at least one sensor,
generating a model of signal sequences based on previously received detection signals while taking into consideration a topology of the sensors and/or environmental condition in the vicinity of the at least one sensor,
comparing the at least one detection signal with the model of signal sequences generated by the model generator; and
predicting quality of the at least one detection signal based on a comparison of the at least one detection signal with the model generated by the model generator.

9. The method according to claim 8, further comprising
determining whether or not the at least one detection signal comprises an error, and
classifying the error into a plurality of predetermined error classes, if it is determined that the detection signal comprises an error.

10. The method according to claim 9, further comprising
correcting the error, if it is determined that the detection signal comprises an error.

11. The method according to claim 10, further comprising:

storing a history of the detection signal, and
generating a detection signal error rate prediction based on the stored history.

12. The method according to claim 11, wherein the prediction is at least one of a short term prediction, a middle term prediction, and a long term prediction.

13. The method according to claim 11 or 12, further comprising:

determining settings of the sensor based on the

detection signal error rate prediction such that the detection signal error rate becomes minimum.

14. The method according to any one of claims 8 to 13, wherein the quality of the detection signal includes at least one of an overall error rate of the detection signal, and an error rate of each of the error classes of the detection signal.
15. A computer program product comprising computer-executable computer program code which, when the program is run on a computer, is arranged to cause the computer to carry out the method according to any one of claims 8 to 14.

Amended claims in accordance with Rule 137(2) EPC.

1. A signal processing apparatus for processing at least one detection signal, the apparatus comprising:

a receiver configured to receive the at least one detection signal from at least one sensor,
a model generator configured to generate a model of signal sequences based on previously received detection signals while taking into consideration a topology of the sensors and/or environmental condition in the vicinity of the at least one sensor,
comparing means configured to compare the at least one detection signal with the model of signal sequences generated by the model generator;
a prediction means configured to predict, for a predetermined time period in future, quality of the at least one detection signal based on a comparison of the at least one detection signal with the model generated by the model generator, and
adjusting means for adjusting settings of the at least one sensor based on the predicted quality, wherein the quality of the detection signal includes at least an overall error rate of the detection signal.

2. The apparatus according to claim 1, further comprising determining means configured to determine whether or not the at least one detection signal comprises an error, and
a classification means configured to classify the error into a plurality of predetermined error classes, if it is determined that the detection signal comprises an error.

3. The apparatus according to claim 2, further comprising

prising

correcting means configured to correct the error, if it is determined that the detection signal comprises an error.

4. The apparatus according to claim 2 or 3, further comprising:

storing means configured to store a history of the detection signal,
and
prediction means configured to generate a detection signal error rate prediction based on the stored history.

5. The apparatus according to claim 4, wherein the prediction is at least one of a short term prediction, a middle term prediction, and a long term prediction.

6. The apparatus according to claim 4 or 5, further comprising:

determining means configured to determine settings of the sensor based on the detection signal error rate prediction such that the detection signal error rate becomes minimum.

7. The apparatus according to any one of claims 2 to 6, wherein the quality of the detection signal further includes an error rate of each of the error classes of the detection signal.

8. A method for processing at least one detection signal, the method comprising:

receiving the at least one detection signal from at least one sensor,
generating a model of signal sequences based on previously received detection signals while taking into consideration a topology of the sensors and/or environmental condition in the vicinity of the at least one sensor,
comparing the at least one detection signal with the model of signal sequences generated by the model generator;
predicting, for a predetermined time period in future, quality of the at least one detection signal based on a comparison of the at least one detection signal with the model generated by the model generator, and
adjusting settings of the at least one sensor based on the predicted quality, wherein the quality of the detection signal includes at least an overall error rate of the detection signal.

9. The method according to claim 8, further comprising

ing

determining whether or not the at least one detection signal comprises an error, and
classifying the error into a plurality of predetermined error classes, if it is determined that the detection signal comprises an error. 5

10. The method according to claim 9, further comprising
correcting the error, if it is determined that the detection signal comprises an error. 10

11. The method according to claim 10, further comprising: 15

storing a history of the detection signal, and
generating a detection signal error rate prediction based on the stored history.

12. The method according to claim 11, wherein the prediction is at least one of a short term prediction, a middle term prediction, and a long term prediction. 20

13. The method according to claim 11 or 12, further comprising: 25

determining settings of the sensor based on the detection signal error rate prediction such that the detection signal error rate becomes minimum. 30

14. The method according to any one of claims 9 to 13, wherein
the quality of the detection signal further includes an error rate of each of the error classes of the detection signal. 35

15. A computer program product comprising computer-executable computer program code which, when the program is run on a computer, is arranged to cause the computer to carry out the method according to any one of claims 8 to 14. 40

45

50

55

Fig. 1

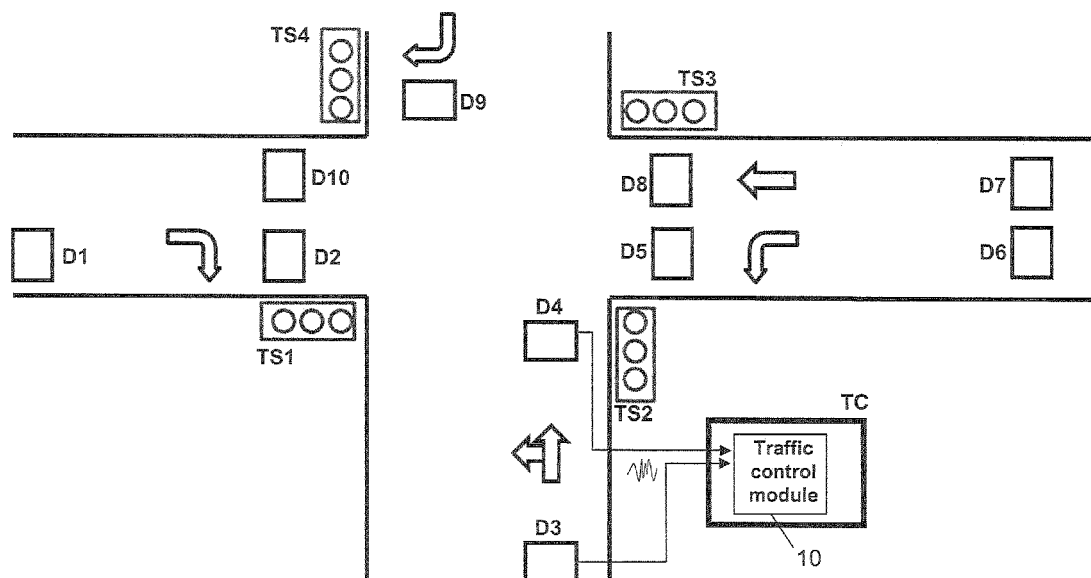


Fig. 2

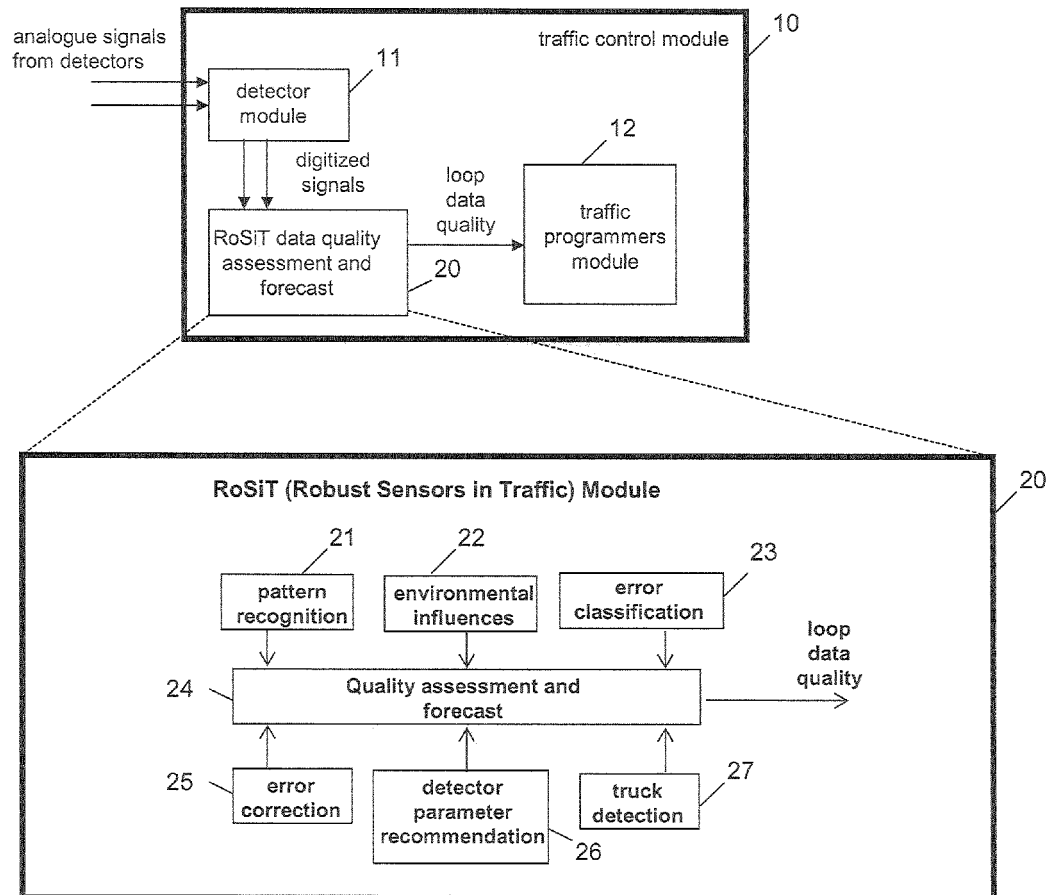


Fig. 3

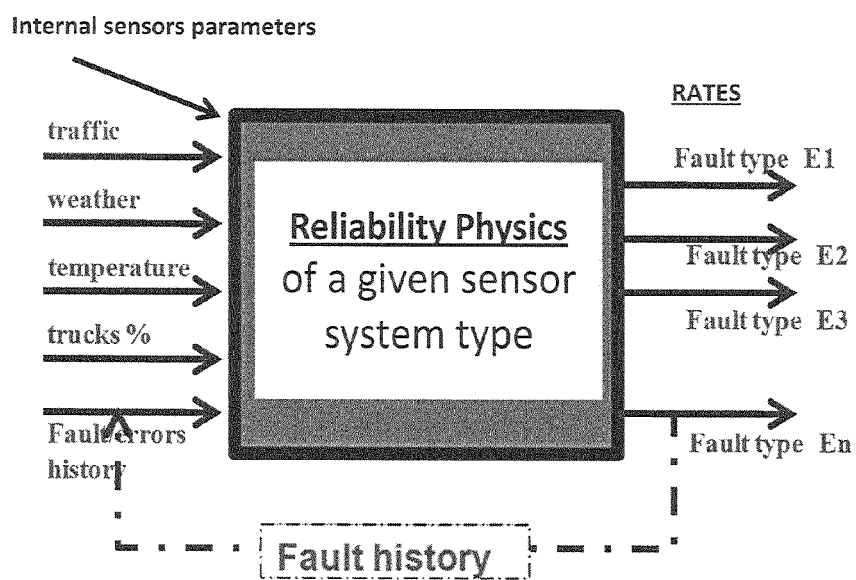


Fig. 4

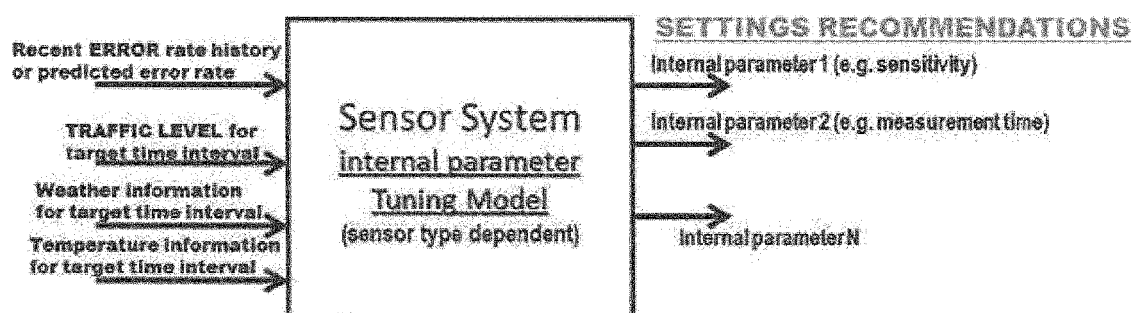
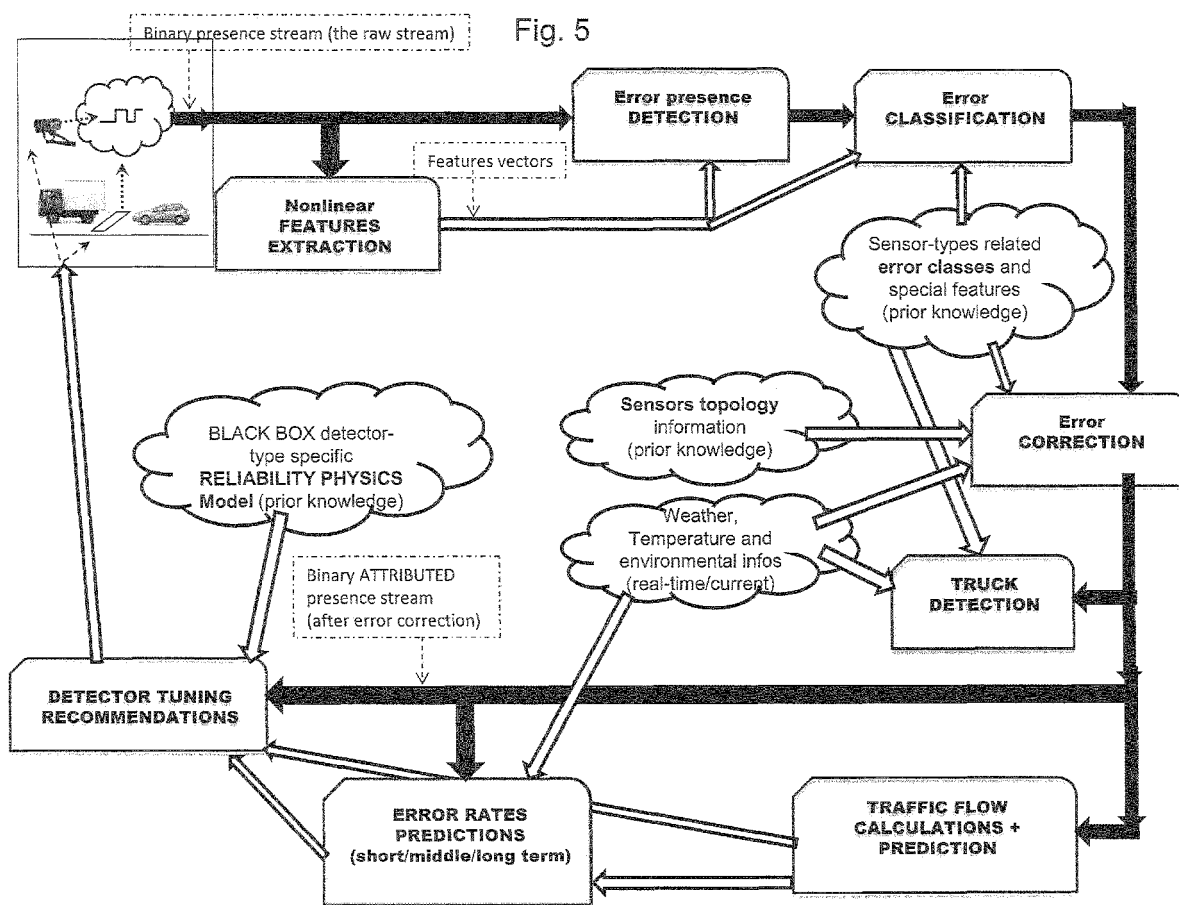


Fig. 5





EUROPEAN SEARCH REPORT

Application Number
EP 13 16 2865

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
X	US 2012/161987 A1 (YOO JAE-JUN [KR] ET AL) 28 June 2012 (2012-06-28) * paragraphs [0113] - [0120]; figure 14 *	1-15	INV. G08G1/07 G08G1/015 G08G1/042
A	AU 2009 236 035 A1 (INTELEMATICS AUSTRALIA PTY LTD) 10 June 2010 (2010-06-10) * page 6, line 4 - line 16; figure 1 *	1-15	
A	US 2002/154032 A1 (HILLIARD STEVEN R [US] ET AL) 24 October 2002 (2002-10-24) * paragraph [0078] *	1-15	
A	US 2004/252193 A1 (HIGGINS BRUCE E [US]) 16 December 2004 (2004-12-16) * paragraph [0117] *	1-15	
			TECHNICAL FIELDS SEARCHED (IPC)
			G08G
The present search report has been drawn up for all claims			
Place of search		Date of completion of the search	Examiner
The Hague		14 February 2014	Créchet, Patrick
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document			

 1
EPO FORM 1503 03.82 (P04C01)

**ANNEX TO THE EUROPEAN SEARCH REPORT
ON EUROPEAN PATENT APPLICATION NO.**

EP 13 16 2865

5

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

14-02-2014

10

15

20

25

30

35

40

45

50

55

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2012161987 A1	28-06-2012	KR 20120072139 A US 2012161987 A1	03-07-2012 28-06-2012
-----	-----	-----	-----
AU 2009236035 A1	10-06-2010	NONE	
-----	-----	-----	-----
US 2002154032 A1	24-10-2002	NONE	
-----	-----	-----	-----
US 2004252193 A1	16-12-2004	AT 504906 T AU 2004202617 A1 CA 2470744 A1 EP 1486928 A2 ES 2364056 T3 PT 1486928 E US 2004252193 A1 WO 2004111971 A2 ZA 200509921 A	15-04-2011 06-01-2005 12-12-2004 15-12-2004 23-08-2011 04-07-2011 16-12-2004 23-12-2004 25-10-2006
-----	-----	-----	-----

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

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

- *The radar dish or antenna transmits pulses of radio waves or microwaves which bounce off any object in their path. The object returns a tiny part of the wave's energy to a dish or antenna which is usually located at the same site as the transmitter*, 10 September 2012, <http://en.wikipedia.org/wiki/Radar> **[0159]**
- **LAWRENCE R. RABINER**. A tutorial on Hidden Markov Models and selected applications in speech recognition. *Proceedings of the IEEE*, February 1989, vol. 77 (2), 257-286 **[0166]**
- Hidden Markov Models. Theory and Applications. In-Tech, 314 **[0166]**