Measurement data treatment in multi-sensor applications for railway vehicle inspection

Thomas Maly and Herbert Schweinzer
Institute of Electrical Measurements and Circuit Design, Vienna University of Technology, Gusshausstrasse 25/354, 1040 Vienna, Austria
E-mail: 1thomas.maly@tuwien.ac.at, 2herbert.schweinzer@tuwien.ac.at

Abstract. Achievement of high safety for railways comprises different aspects. Beside train control technology, also the examination of fault states of driving trains makes a contribution to safety assurance. The past inspection through human train station inspectors is going to be replaced by technical equivalents. Available detection systems focus only on particular fault states. Reliability and accuracy of such systems often suffer from heavy measurement conditions (e.g. high train speeds, disturbing environmental influences). Thus, a new approach consists in the usage of data of different sensor systems for evaluating train fault states from a more global point of view (“Checkpoint”). It implies the necessity of data conjunction for gaining more reliable results. In this work a concept for uniform data treatment is presented. It is based on sensor system independent data representation and on application related data storage for easy and flexible data conjunction. The uniform data representation also includes error descriptions of the values. Furthermore, problems to estimate fault state recognition errors are pointed out and a feasible approach is outlined.

1. Introduction
Formerly, most of the railway stations were staffed with station inspectors. Beside the train control technology, the tasks of the inspectors consisted also in the examination of the trains for achieving high driving safety. Thereby they are checking visible, audible, smellable and sometimes tactile criteria of trains which pass or stop in the train stations. However, due to economical reasons infrastructure operators had to reduce the number of employees. Also railway stations are operated unmanned facilitated by a more centralized handling based on modern control technology.

The market is about to react on these trends. At present, diverse detection systems are under development, which are able to identify the occurrence of particular train fault states (e.g. hot bearings or brakes, flat or overloaded wheels, railway loading gauge violations). Some of them are ready for the market or were already used in the field respectively. In order to inspect most of the trains, but to restrict sensors systems to a maintainable number, wayside installation locations have to be orientated on the main lines. Thus, the detection systems have to be designed for measuring under full travelling speed, which mostly means small measured amplitudes, short acquisition time and measurement variations due to dynamic effects. Additionally, outdoor measurements usual are subject to environmental influences. Meeting these requirements, the measurement data allow evaluating the train condition, which may have impacts on the control technology. Hence, there are also extremely high requirements of the reliability of such systems.
2. Checkpoints – data fusion with uniform data treatment

2.1. Demand for data fusion
Detectors focus only on the occurrence of one or few fault states. As a stand-alone system primarily delivering alarm or warning notifications, in general the systems are able to provide the measured properties of the train (e.g. bearing temperature, wheel weight). An enormous advantage in train defect estimation can be achieved through a holistic consideration of the train state. This means that the information of different sensor systems together (e.g. dynamic scale, hot box detector, flat wheel detector, loading gauge detector, derailment detector) build the data basis for a global train inspection. By sophisticated model-based conjunction of data of different sensor systems, physical dependencies can be taken into account for an improvement of the resulting state prediction quality. Therefore an overall system, which is able to combine all sensor systems with minimum integration effort, is under development. It is called “Checkpoint”[1].

2.2. Sensor system independent data representation
In a first step, the concrete output of each sensor system has to be transformed into an abstract data format. This standardisation of the provided data enables a uniform data treatment for data conjunction and evaluation.

In general, sensor systems typically supply following data types: binary data (e.g. warning or alarm notifications, loading gauge violations), enumerations (e.g. big, middle or weak flat wheels), integer or floating point numbers (e.g. bearing temperatures, wheel weight). The first two categories are mostly results of an internal evaluation process of the sensor systems itself. Numeric data may either be pre-evaluated or raw measurement data. Nevertheless, this information is completely transformed by the data abstraction stage (Fig. 1) and delivered as specific characteristics of train properties. Several specific characteristics are defined for each sensor system type. This enables to replace the sensor system in use together with its data abstraction unit by another system of the same type without further modifications in other layers of processing. After calculating characteristic values, they are linked to a database, called “train model”. As an example, the temperature of a wheel bearing is measured and provided as a sensor system independent characteristic “bearing temperature”. Different bearing temperature measurement systems may directly provide temperature values or also IR-radiation intensity values. In both cases the data abstraction stage will calculate the characteristic “bearing temperature”, e.g. by transforming radiation intensity to surface temperature.

For basic error considerations at data evaluation stage, specific characteristics also have to carry information about measurement uncertainties. An often suitable description uses the distribution type with characteristic values (e.g. Gaussian distribution and variance). To fulfil more complex requirements, the error probability density distribution may be specified.
2.3. Train model - an application related data-base and reference method

As mentioned before, sensors systems are mostly designed as stand-alone systems with direct signalling of warnings and alarms. To notify a reference position of detected problems on the train, they are using individual reference methods which often necessitate local measurements of additional reference information, e.g. axle-counting.

Data fusion demands a common reference system. A common time reference cannot be realised because of the insularity of the different sensor systems. However, measurement data are mostly related to properties of train objects. These objects have fixed geometric positions on the train and can be used as references. Using train objects as references allows relating measured data for further conjunction.

Three types of information of the providing systems referenced by the appropriate object types can be identified: train-based information (e.g. train speed, number of wagons, train identification number), wagon-based information (e.g. wagon weight, wagon type and identification number, number of axles) and axle-based information (e.g. axle weight, bearing temperatures, brake disc temperatures). Further geometric details of the objects have to be defined relatively within the object, e.g. axes positions of a wagon, left or right wheel of an axle.

For global data fusion of all systems, the three referencing systems have to be merged to one common referencing system - the train model [2]. The train model is primarily a data-base, but further allows linking of measured data in two orders: a logical order of data of the same characteristic related to a specific object type, and a geometric order according to the hierarchy of objects. Prerequisite for the common base is the use of at least one sensor system, which provides wagon-axle relation data.

2.4. Data analysis

A rule based conjunction of characteristics enables a consideration of physical dependencies. For instance, the bearing temperature depends on the load of the wheel. Higher load causes higher bearing stress and leads to increased bearing temperatures. To avoid wrong decisions, a new characteristic “wheel load compensated bearing temperature” will be created which gets results of processed compensation rules and will be used for evaluating the bearing condition.

Available rules comprise a set of mathematical functions on the defined data types, like linear operations or case decisions. Together with a provided conjunction of characteristics, also the error description of the resulting characteristic has to be calculated according to the algorithm. Using resulting characteristics as inputs for further conjunctions, a multi-stage data fusion can be realised.

After processing, the resulting characteristic is evaluated by simply comparing the value with a threshold. In case of exceeding the threshold, previously defined and approved actions (e.g. alarm message generation, notification of the interlocking machine) are initiated.

3. Method to ensure optimal classification

3.1. Aspects of threshold adjustment

Gaining high operational safety is a big demand of railway operators. Accordingly a high detection rate of fault states is demanded. Thus, the desired recognition probability (RP) of the Checkpoint should be between 95 and 100%. This aim can be achieved by sensor data fusion and by adjusting the alarm threshold towards more sensitive triggering.

On the other hand, main railway lines carry heavy train traffic. In the majority of cases stopped or decelerated trains significantly influence the time schedule of subsequent trains. Economical aspects require a very low proportion of false alarms with respect to the total number of checked objects: a very low recognition error probability (REP). To reach this, the alarm threshold should be adjusted in a more insensitive manner which is contrary to a high recognition probability. Thus, a trade-off has to be found, which satisfies both requirements.
3.2. Problems of fault state evaluation
Considerations concerning the RP and the REP have to include potential sources for incorrect results. The following problems can be identified:

- Significance of measured characteristics: some fault states can be determined directly. But for most of them only secondary characteristics can be used for estimation of the real condition. For example, for evaluating the bearing condition the bearing surface temperature has high significance. But just at the frontier between good and bad bearing condition, the surface temperature is rather an indicator than a reliable value.

- Imperfection of measurements: each measurement is influenced by unknown or unpredictable environmental conditions. This leads to random deviations which can be specified by error probability density functions. Mostly these functions have a mean value zero and are narrow in relation to the measurement range. Thus, mentionable influences on recognition probability or error probability are mainly expected if the measurement value is near the threshold and the uncertainty can cause faulty decisions.

- Influences on measured characteristics: influences of general train properties on specific characteristics are sometimes well known (e.g. train speed). However, if not known influences must be considered as unpredictable. The resulting uncertainty has to be specified again by probability density functions.

3.3. Estimation of fault recognition properties
Aforesaid imperfections may cause faulty decisions during data evaluation: the probability of faulty decisions rises if typical values of characteristics are near the evaluation threshold. Thus an estimation of both RP and REP of an evaluated characteristic has to be based on a frequency distribution of values of the characteristic. The concrete frequency distribution of a characteristic depends on the Checkpoint location. Especially type and arrangement of railway lines influence the frequency distribution: typical wagon weights cause a specific mean wheel load, typical train operation results in different stress and cooling of wheel bearings.

Up to now, no information exists about the frequency distribution of values of most characteristics. Thus, a value logging functionality for characteristics is implemented in the uniform data treatment (Fig. 1). It enables extensive data gathering for subsequent statistical analysis. Based on random samples, the frequency distribution can be built which will be used for estimation of the systems recognition properties. Furthermore, by extensive analysis of collected data physical dependencies between characteristics can be verified and optimised conjunctions may be derived.

4. Conclusion
In this work a system for recognizing train fault states was presented. Based on various sensor system types, the fault recognition properties may be improved by conjunction of measured characteristics. A prototype Checkpoint is already built up and is under test since several months. Gathered sensor data will be used to optimise the recognition performance.

References
[1] Maly T, Schweinzer H and Rumpler M Sept. 2004 Advances in train monitoring by networked checkpoints Proc. 5th IEEE Int. Workshop on Factory Communication Systems (Sept. 2004, Vienna) pp 339-342
[2] Maly T, Rumpler M and Schweinzer H Oct. 2004 Joining sensor systems for railway vehicle inspection Proc. IEEE Sensors (Oct. 2004, Vienna) pp 12-15

1 “Checkpoint” and “set up of a test system” are projects granted by the Austrian government, partners: Alcatel Austria, Austrian Federal Railways and Vienna University of Technology