Information fusion of GNSS sensor readings, field notes, and expert's a priori knowledge

Vasenev A., Ionita D., Bijleveld F., Hartmann T., Dorée A.
University of Twente, the Netherlands
a.vasenev@utwente.nl

Abstract. Documenting machinery movements by using positioning technologies, such as global navigation satellite systems (GNSS), is essential to understand and further improve construction processes. However, before measurements can be meaningfully analysed the documented movements should be filtered to exclude outliers. Eliminating outliers manually is a time-demanding process, while automatic filtering can be inaccurate. In particular, path elements may get lost if machine-specific movements are misconceived as noisy data. As a trade-off, we propose an information fusion approach to filter paths of construction machines in a semi-automatic way. The approach allows an expert to relate "hard" sensor and "soft" field records with his or her expectations about how machines can move in real construction projects. Specially developed open-source software illustrates the proposed method for filtering the documented paths of machines involved in road paving projects. The initial testing of the developed software showed its suitability to filter outliers in GNSS data and identified possibilities for further improvements.

1. Introduction

Documenting movements of machines during road construction projects is beneficial both to analyze the progress of the construction process and to identify best and poor practices (Miller et al. 2011). The commonly used technologies to document machine movements on construction sites include GNSS (Global Navigation Satellite System) and laser based positioning sensors. The choice between the two systems can depend on specifics of the construction project. The utilization of laser guided systems provides highly precise data, but could require additional stations to follow all machines during the construction project. For instance, additional hardware might be needed if the direct line of view between machines and a positioning station is obscured by buildings, trees, or other objects located on or next to the construction site. Alternatively, GNSS sensors can be applied to track the construction machinery using a single base station that transmits the correction signal via radio frequencies to multiple GNSS sensors located nearby. This characteristic can justify the choice of using GNSS for example to track road paving projects, where not only machines continuously move, but also the location of the construction site continuously change. However, the correctness of GNSS data is lower than in case of laser based positioning and highly depends on signals from satellites. If the signals are temporarily lost or wrongly interpreted, the documented GNSS readings can include outliers that should be filtered out (Bijleveld et al. 2011).

Removing outliers in the collected data can either rely on automated processing algorithms, such as moving average and local regression smoothing, or demands intensive manual data analysis. Both approaches have their limitations. The automated approach is relatively fast, but some segments of the documented path can be erroneously removed. For example, if the machine's heading is changing, the automated approach can misconceive this change as noisy data. Alternatively, a human expert can perform the careful filtering by removing outliers in GNSS data. In this way, the expert could consider additional information about specific events (such as the process disrupted due to lunch breaks or machinery breakdowns) or changes in the process (for instance, a rush at the end of the working shift). However, the manual filtering is time consuming.

In this paper we propose an information fusion approach to remove outliers in GNSS data. The approach combines both automated and manual methods and extends the previously described automated path filtering methods (Bijleveld et al. 2011). The paper continues with outlining some features of information fusion systems, applicable to refining paths of construction machines. Then, the paper proposes an information fusion approach and describes its implementation to filter GNSS data. Finally, initial testing of the developed open-source software prototype for asphalt paving operations and possible further work directions are depicted.

2. Human-generated Information in Information Fusion Systems

Information fusion (IF) systems are oriented to fuse any kind of data (Raol 2010) including "hard" sensor and "soft" human-generated records (Khaleghi & Karray 2012, Pravia et al. 2009). Both data types have their specific qualities, being precise and additionally descriptive respectively: while the detailed sensor readings can be obtained relatively straightforward, the soft data can expound the process' context and descriptions of particular events. For example, the soft data can include some description based on attributes of interest (Gross et al. 2012) described in a specific (such as illustrated by Pravia et al. 2009) or an ill-structured format to reduce the burden of formalizing data collection. Similarly to analysing the soft data per se, the fusing of soft and hard data be performed by involving an expert because humans have the ability "to gather and organize unstructured, *a priori* information about a problem and then mix that information with measured sensor data, making inferences that could not have been made using the sensor data alone" (Bath et al. 2005, p. 309).

Though human analysis appears to be essential in fusing information (Blasch & Plano 2002, Nilsson et al. 2012) and can naturally support information fusion processes (Blasch & Plano 2003), most fusion systems in the construction domain are not oriented towards active co-operation with a human expert during the data processing process. The commonly used fusion approaches, such as introduced by Haas (2006) and Shahandashti and Razavi (2011), are mainly related to "hard" data fusion. Similarly, specialized data fusion systems to estimate resource locations (Razavi 2010 & Haas 2010) or to support construction management knowledge discovery (Liu, Soibelman & Wu 2004) do not imply continuous interactions with the user during data analysis.

By describing the information fusion approach that utilize both soft data and human reasoning for path filtering, this paper depicts a user-centred interactive solution for a specific task of civil engineering. We presume that utilization of additional descriptive data together with involving a human expert into information fusing can significantly benefit the outcome of the fusing process. For example, in relation to tasks of path filtering, an expert can consider *a priori* knowledge about possible machine movements together with specifics of the construction project, collected as soft data. In particular, the soft data can describe the context of the construction process, such as site geometry and obstacles on site, as well as depict intents of machine operators about how they plan to conduct their work. To support experts in relating such soft data to particular segments of the documented machine paths we propose an information fusion approach, described in the next section.

3. The Proposed Information Fusion Approach to Combine Soft and Hard Data

To filter the documented machine paths based on hard and soft data we propose a user-centred information fusion approach (Figure 1). The approach is based on experts' activities related to the Observe-Orient-Decide-Act (OODA) decision-making model that was introduced in the mid 1950s (Nilsson et al. 2012). In relation to filtering paths of construction machines, the loop can be operationalized as follows: an expert considers (or in other words makes decisions based on) the incoming information combined with understanding of rules about how machines can move. Then, the experts can adjust the sensor readings by eliminating outliers by means of several actions: changing the parameters to find the next outlier, automatically identifying the next outlier, manually adjusting the outlier's limits, and linearizing the outlier.

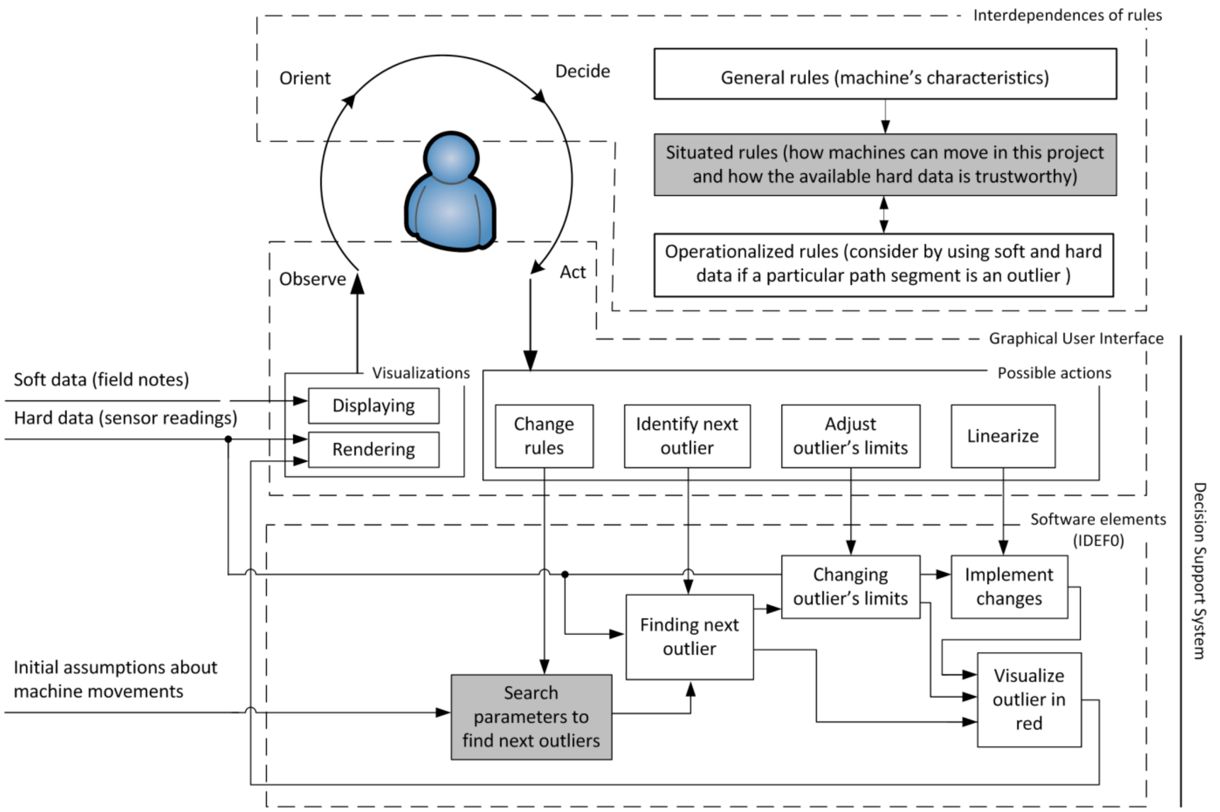

Figure 1. Information Fusion Approach for Retrospective Analysis of the Positional Data by a Single Expert

The experts' activities are driven by their understanding, or in other words "rules", about how machines can move during the analyzed project. In this way, the rules can be seen as part of the experts' knowledge. These sets of rules are considered by an expert during the "orient" phase of the OODA process and can be adjusted during the "decide" phase. Within the proposed approach we outline three interdependent sets of rules: general, situated, and operationalized ones. The rules correspond to how machines can move in general; during the analyzed project; and in particular periods of time during the analyzed project.

The general rules are based on machine characteristics and describe how particular construction machines can move due to their specifications. For instance, machine speed or turning angle can be limited. These machine-specific rules are not expected to differ between different projects and supply an expert with initial understanding about how machines can move during the examined project.

The situated rules describe how construction machines can move during the analysed project. Such rules are related to general rules and to project-specific information, for instance to assumptions about the expected movement patterns and understanding if the collected sensor readings are noisy. In these settings, the machine's maximal speed during the project can be expected to remain significantly less than maximal speed specified in the machine's specifications. For example, the user can consider that: (1) a machine during specific time periods is not expected to make rapid turns and move faster than a certain speed due to the geometry of the construction site; and (2) that outliers can be grouped in a way that if a single outlier was identified, the next *N* machine positions should be carefully analysed as potentially prone to errors.

The situated rules are interrelated with the "operationalized rules". These rules exist in relation to a particular machine path's section within the beginning and the end of the identified outlier. In this way, the expert judges if the particular section is indeed an outlier or if the sensor readings accurately describe machine movements. Such judgments are rooted in situated rules and supported by visual representations of the documented machine's path and additional soft data. The operationalized rules can, in turn, influence the situated rules. For example, the existing expectations about the maximal machine's speed during the analyzed project can be updated if that machine had begun to move faster than expected. Similarly, if the expert had identified too many outliers within the project, he or she can adjust personal assumptions about the noise level in the GNSS data. Such assumptions can be operationalized by changing parameters to find the next outlier, e.g. by increasing the amount of *N* consequent error-prone positions to be critically analyzed.

Within the proposed classification of rules, the situated rules are central in describing how machines can move during the analysed project. These rules, if described as statements, can support automated search for next outliers. For instance, the statements – that can change during the project analysis – could have the following formats: (1) "the maximum angle between the previous and current heading of the machine should not exceed «*20*» degrees"; (2) "the machine's maximal speed should not exceed «*2.7*» m/s"; (3) or "after identifying an outlier the next *N* path points should be checked and all of them should be considered as wrong if the next outlier will be found".

Once the statements are formalized, they become a central element in the proposed semi-automatic IF approach to refine the documented machine paths. The expert can operationalize the situated rules as the program's search parameters to find next outliers (the corresponding blocks in Figure 1 are shown in grey). After the next outlier is identified, the expert can change limits and eliminate the outlier (for example, by linearizing the selected set of locations). The latter action is performed if the selection is considered as an outlier and, therefore, corresponds to the operationalized rules. The sequence of the possible actions is not predefined. For example, a user can find the next outlier and remove it without adjusting the outlier's limits. Similarly, the rules to search for next outliers can be changed at any moment of time.

The interaction and visualization means correspond to software elements of the information system and define the ways the expert can interact with the elements. To illustrate how the proposed IF approach can support users in finding and filtering out outliers in GNSS data in a semi-automatic way, we developed a specialized open-source Java application. The system implements the described approach for the case of asphalt paving operations. The specific of asphalt paving operations and the software implementation details are described in the next section.

4. Illustrative Implementation of the Proposed IF Approach

To illustrate the proposed IF approach we developed an open-source application that imports hard and soft data, processes them, and uploads the processed data to a remote server. The application is oriented for a specific type of construction activities: asphalt paving operations. Specifics of asphalt paving can illuminate the differences between different sets of general, situated, and operationalized rules. During road paving projects, machines normally move relatively slow in comparison to their specifications and often change their directions. These changes of directions can be misinterpreted as noise in GNSS data by filtering algorithms. At the same time, machine movements are relatively easy predictable by a human who expects machines to move in particular patterns (Bijleveld et al. 2011).

To illustrate how specifics of machine movements can influence considerations about possible outliers in GNSS data, this paper further depicts specifics of the machine movements during the asphalt paving process. Then, particular software implementation and initial testing of the software are described.

4.1 Characterizing Machine Movements during Asphalt Paving Projects

The paving operations necessitate coordinated movements of multiple specialized construction machines, such as pavers and rollers. The paver evenly distributes the asphalt mixture and a fleet of rollers continuously compact the deployed asphalt mat until reaching a desired density. The machine movements highly depend on characteristics and functions of construction machines, geometry of the construction site, and specifics of the paving process.

During road construction the paver normally changes its headings based on the site geometry while moving forwards. At the end of the paved lane, it can move backwards to reach the beginning of the next one. Though the paver should, ideally, advance without stops to sustain the continuity of the paving process, in practice the continuity can often be disturbed by external events. For instance, delays of asphalt trucks, breaks, and other events can negatively influence the continuity and result in discrepancies in the paving process. Additionally, the paver's speed varies according to project conditions. Thus, even if the maximum paver's speed is known (e.g. 25 m/min for a particular model (VÖGELE 2012)), the actual speed during road construction can depend on the desired layer thickness, continuity of asphalt delivery, and movements of other machines. Altogether, the additional information about a paving process, such as description of the site geometry, can support in anticipating machine movements, such as the maximum expected angle of changes in the paver's heading of paver's movement direction (forwards or backwards) at any moment of time.

The paver is closely followed by rollers, whose goal is to achieve an optimum density and to provide a smooth surface by compacting the asphalt mat in a specific pattern. In general, rollers move back and forth at slow but uniform speed. The roller's driving direction should not be suddenly changed or rapidly reversed, as these actions can displace the mixture. Additionally, rollers should not stay still over the freshly paved asphalt mat and might avoid overusing stop/start sequences. Roller movements form a rolling pattern that normally progress from the lower to the higher side of the asphalt lane. The maximum speed and the maximum turning angle vary from one roller to another. For example, a three-drum roller can have the maximum steering angle of 40 degrees and can reach the speed of 10.2 km/h (HAMM 2012-1), and a tandem roller can perform 25 degree turn and can speed up till 12.0 km/h (HAMM 2012-2).

In summary, the information on how the construction machines can move in general, the collected GNSS data, and the specifics of the project can support a human in refining the machine paths. To allow the human expert to refine paths of road construction machines we developed a specialized application, described in the next section.

4.2 Implementations Details

An application to import the data, parse it, remove outliers, and upload data to a remote server was developed by the authors using a set of standard and third-party Java components and data structures. In particular, input data files, which include documented GNSS data and a logbook, are parsed on-the-fly as the user selects a data file. If the parsing reveals issues with input data, the program informs the user immediately. The parsed data are stored in memory for fast rendering and quick data modification by using index-based structures, named LinkedHashMaps. These structures link machine's positional data to their respective timestamps and allow employing fast index-based retrieval of elements while maintaining their order. The interface of the developed application employs the adjusted Wizard API framework (http://java.net/projects/wizard) and the Minigeo (http://code.google.com/p/minigeo/) library to visualize machine paths.

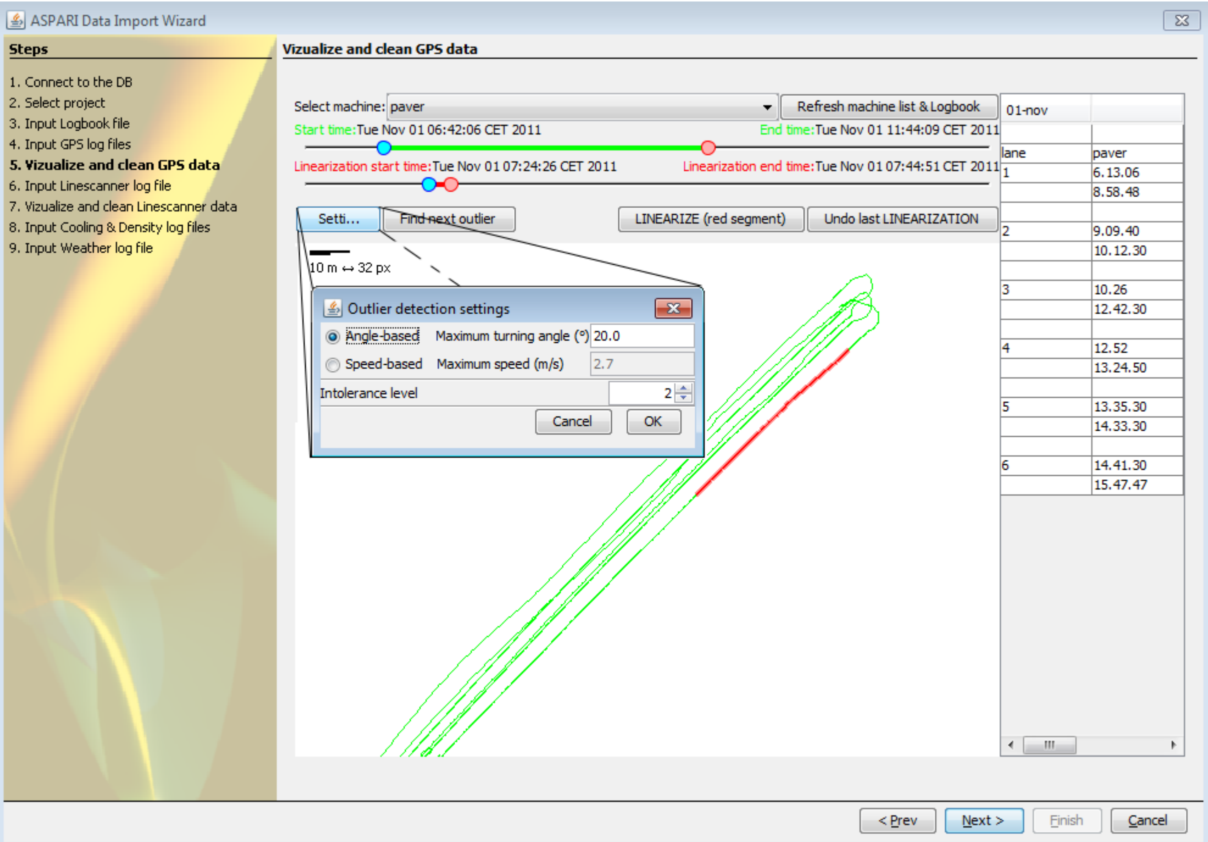

Figure 2. Graphical User Interface of the Developed Application

For users' convenience, the developed graphical user interface renders the machine's path and displays the project's logbook next to it (Figure 2). Additionally, an exemplary path segment is selected and shown in red (the selection made bold in Figure 2 for illustration purposes).

The users communicate their expectations about how machines can move within a particular project by defining the outlier detection settings. In our implementation, the user can choose between two means to detect outliers: *angle-based* search (by setting the maximal expected machine's turning angle) and *speed-based* search (by defining the maximal expected speed of the machine). Correspondingly, the next outlier will be identified either if the machines direction significantly differs from the previous machine's heading, or if the adjunct path points are located too far away from each other. After an outlier was identified, the program considers the next N data points. The user-entered value N – called the "intolerance level" – defines the amount of consequent points that will be considered as being potentially incorrect and corresponds to the program's ability of group outliers. If the next outlier is identified within the next N machine's locations, the limits of the selected outlier interval will be automatically increased to include the next outlier. Additionally, the user can (re-)define the outlier's limits by moving a specific horizontal slider. Then, the user can relate the selected path with the soft data displayed next to the path visualization. If the selected interval is considered as noisy sensor readings, the selection can be linearized. The last linearization can be cancelled, for example if that action was done by mistake.

After identified outliers are eliminated, the developed program can, based on user's decision, generate a triangular mesh that corresponds to the constructed pavement layer. The mesh is determined based on movements and the width of the paver and can later be visualized by other software, such as AsphaltOpen (Miller et al. 2011), to support the analysis of the documented paving project. The mesh, the logbook, and the GNSS data are finally uploaded to a remote server by groups of 1000 statements each.

4.3 Initial Testing of the Developed Software

To test the developed software we used data gathered during a typical asphalt paving project. The logbook and GNSS sensor readings were collected at the road construction site near Alkmaar, a city in the Netherlands. During one 7 hour work shift about 1200 tons of asphalt mixture was paved, which resulted in an 800 m long, 7 m wide, and 7 cm thick asphalt layer.

The collected GNSS readings and the project's logbook were provided to a person familiar with the program to identify and filter out path outliers. The user chose the angle-based detection method for this task. The reconstructed sequence of identifying and correcting a set of outliers in the GNSS data is represented in Figure 3. Some outliers were found automatically based on predefined parameters and linearized right after the detection, while other outliers' limits were changed by the user. For example, the rightmost part of the Figure 3 illustrates a manually selected outlier. The automatically found interval was extended by the user who included the next few path points as he considered them as noisy data. The user assumed that the machine continued to move without changing its direction, as no discrepancies were noted by the logbook and the further machine's path constituted a relatively straight line. Though the user decided to adjust the interval manually, he stated that the next time he could increase the "intolerance level" setting to put a longer set of consequent points under scrutiny. After all the GNSS data were analysed, the corrected machine paths together with the logbook data were uploaded to the remote database server. In summary, this initial test demonstrated the applicability of the developed software to filter out incorrect segments in documented machine paths by involving an expert into data analysis.

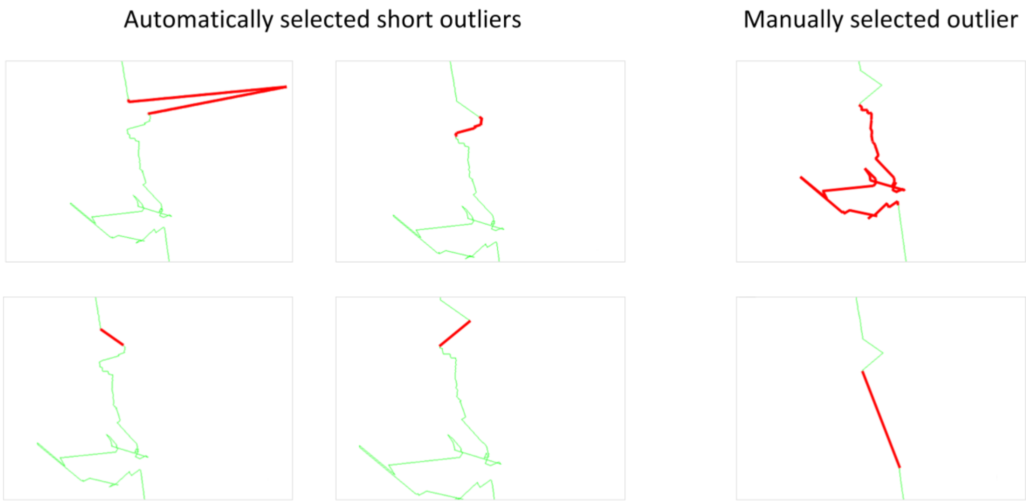

Figure 3. Examples of Identified and Linearized Sections of the Roller's Path

5. Discussion and Future Work

The proposed semi-automated information fusion approach to filter outliers in GNSS data benefits from including soft data and human reasoning about how machines can move within a particular construction project. In particular, the soft data can support user's decisions if some segments of the documented machines' paths do correspond to actual machines' movements at particular moments. Additionally, the soft data can support constructing the situated rules according to user's expectations about how machines can move during the analysed project. For example, if the logbook mentions trees, bridges, or high buildings next to the road, the user can pay additional attention to the data collected next to such obstacles and can accordingly adjust program parameters to search for the next outlier.

Though the testing of the developed software shows the applicability of the proposed information fusion approach to support analysis of GNSS data, several limitations characterize the research scope. In particular, only support of a particular type of the soft data – a logbook of the construction project – was implemented while other potentially useful data, such as on-site photos and construction plans, were not presented to the user. Also, the given example characterizes potential applicability of the approach, rather than calculate direct advantages of use the proposed solution in comparison to fully automated or human analysis. The performance of the approach was not subject of study in this research. Overall, we consider this characteristic as hardly generalizable, as the necessary amount of human effort is highly project specific and depends not only on accuracy of GNSS equipment, but also on other parameters, such as ambient weather conditions and unobscured transmission signals from navigation satellites. Additionally, this research mainly focused on refining GNSS readings and though the machine paths documented by using laser-based positioning systems could also be filtered similarly, additional tests are desired. Thus, further analysis of construction projects under diverse conditions with relatively "noisy" data is needed to study advantages of the approach.

Future work can include studying opportunities of developing machine learning algorithms to automatically correct the outlier detection settings. For instance, if GNSS data are found to be noisy within a specific region, the software could automatically adjust some parameters to effectively discriminate possible outliers in that zone. This study direction could ultimately lead to a computer-pull communication pattern that can reduce the need for the expert's attention due to querying for information “only when the expected value of their observation exceeded the cost of obtaining it” (Kaupp 2008, p.140). Studying possibilities to generalize the learning during the post processing could be essential to advance in this direction.

6. Conclusions

Information fusion of the documented GNSS sensor readings, field notes, and *a priori* knowledge of a human operator can support removing outliers in the documented machine paths. The interactive way of finding and fixing outliers in machine paths, related to noisy sensor readings, ultimately aims at improving the accuracy of the collected data.

In this paper, we proposed a human-centred information fusion approach to combine soft and hard data together with expert’s reasoning. The documented machine paths are corrected in a semi-automatic way. This information fusion approach aspires to improve automated data filtering that can misinterpret some segments related to machine-specific movements during construction operations.

To illustrate the proposed approach, we developed open-source software to refine paths of construction machines involved in asphalt paving operations. The software assists users in data analysis by supporting tasks of easy identifying and eliminating outliers based on expectations how machines could move during a particular paving project. The groups of outliers can be selected according to a specific set of search rules or by adjusting the time limits of the identified outlier. The beta-version of the software can be freely downloaded from http://asphaltopen.svn.sourceforge.net (“DataImportWizard” project) and could potentially be useful to analyse documented paths of construction machines.

References

Bath, W. G., Boswell, C. M., Sommerer, S., Wang, I. J. (2005). Detection systems information fusion. Johns Hopkins Apl Technical Digest, 26(4), pp. 306-314.

Bijleveld, F. R., Vasenev, A., Hartmann, T., Doree, A. G. (2011). Smoothing GPS data of rollers to visualize asphalt paving operations. In: EG-ICE Workshop, 2011, University of Twente.

Blasch, E. P., Plano, S. (2002). JDL level 5 fusion model "user refinement" issues and applications in group tracking. Signal Processing, Sensor Fusion, and Target Recognition Xi, 4729, pp. 270-279.

Blasch, E. P., Plano, S. (2003). Level 5: user refinement to aid the fusion process, pp. 288-297.

Gross, G. A., Nagi, R., Sambhoos, K., Schlegel, D. R., Shapiro, S. C., Tauer, G. (2012). Towards hard+soft data fusion: Processing architecture and implementation for the joint fusion and analysis of hard and soft intelligence data. In: 15th International Conference on Information Fusion, pp. 955-962.

Haas, C. (2006). "A model for data fusion in civil engineering." Intelligent Computing in Engineering and Architecture 4200: pp. 315-319.

HAMM (2012-1) HW 90B/10 Static three-wheel roller 10t. http://www.hamm.eu/en/produkte/statische-walzen/serie-hw/hw-90b-10.227.php#!tab=technische-daten, retrieved March 2012.

HAMM (2012-2) DV 70 VV Pivot steered tandem roller, double vibration roller. http://www.hamm.eu/en/produkte/tandemwalzen/serie-dv/dv-70-vv.176.php#!tab=technische-daten, retrieved March 2012.

Kaupp, T. (2008). Probabilistic Human-Robot Information Fusion, PhD thesis, University of Sydney.

Khaleghi, B., Karray, F. (2012). Human-centered Multi-target Tracking: A Random Set Theoretic Approach, In: Ieee 13th International Conference on Information Reuse and Integration 2012 (Iri), pp. 193-200.

Liu, L., Soibelman, L., Wu, J. (2004). Data Fusion and Modeling for Construction Management Knowledge Discovery.

Miller, S. R., Hartmann, T., Doree, A. G. (2011). Measuring and visualizing hot mix asphalt concrete paving operations. Automation in Construction, 20(4), pp. 474-481.

Nilsson, M., van Laere, J., Susi, T., Ziemke, T. (2012). Information fusion in practice: A distributed cognition perspective on the active role of users. Information Fusion, 13(1), pp. 60-78.

Pravia, M. A., Babko-Malaya, O., Schneider, M. K., White, J. V., Chong, C. Y., Willsky, A. S. (2009). Lessons learned in the creation of a data set for hard/soft information fusion. In: 12th International Conference on Information Fusion, pp. 1037-1046.

Raol, J.R. (2009). Multi-Sensor Data Fusion with MATLAB, CRC Press.

Razavi, S. N. C. T. Haas (2010). "Multisensor data fusion for on-site materials tracking in construction." Automation in Construction 19(8): pp. 1037-1046.

Shahandashti, S. M., S. N. Razavi, et al. (2011). "Data-Fusion Approaches and Applications for Construction Engineering." Journal of Construction Engineering and Management-Asce 137(10): pp. 863-869.

VÖGELE (2012) VÖGELE SUPER 1900-3i http://www.voegele.info/en/produkte/superserie/superhighwayclass/super19003/spezifikation.html, retrieved March 2012.