Reviewer evaluation report

on

Doctoral (Ph.D). thesis entitled

Automotive and mechatronics applications of IMU sensors

prepared by

Roland Nagy

supervisor:

István Szalai, DSc, professor

Reviewer:

Dr. Zsolt János Viharos, Ph.D.

HUN-REN Institute for Computer Science and Control (SZTAKI), senior research fellow, leader of the Intelligent Processes research group;

John von Neumann University, vice-dean for science, department leader, associate professor.

Budapest, 2026.06.14

1. Topic and goal of the Ph.D. thesis

The Ph.D. thesis addresses a very interesting and important field of road management, which is used every day over the year, and it is one of the key infrastructure components of our social and business life. It has relations to health, safety, and financial aspects as well, consequently, keeping our roads in good condition is a fundamental interest of our society. For ensuring this expectation a first step is to measure the road condition, it is a preliminary task before implementing any maintenance or correction actions, consequently, the topic of this Ph.D. thesis is well selected, and it is a significant practical and scientific field.

2. Structure and content of the thesis

The thesis has 146 pages and 6 main paragraphs and various sub-paragraphs extended with the formulation of the theses, publications of the Ph.D. candidate (connected and independent), references and acknowledgement. The structure of the thesis is logical and ordinary for Ph.D. theses, in many cases the acknowledgement is at the beginning of the Ph.D. thesis, but this is not an important consideration, neither the abstracts at the beginning of the document in three languages is not a typical for this kind of document.

After the short introduction the "*Fundamentals and literature research*" paragraph reviews the state-of-the-art of the addressed field, it lists the related scientific publications and evaluates them in a critical way. This led to a clear positioning of the Ph.D. research field, the short conclusions at the end of various state-of-the-art sub-topic reviews are very well formulated and it is one of the main strengths of the document. Road condition evaluation methods were summarized: it is surprising that the road quality is not fully standardized and it depends e.g. on countries or operation companies. Later on, it gives special importance to the clustering results, considering the results on the (~optimal) number of distinctive clusters. A separate sub-paragraph details the vehicle body characteristics, dealing also with the possible positioning of various sensors and sensor types. It also considers the cost and complexity of measurements, which is a good argumentation for utilizing cheap and simple sensors, detailed also at the next paragraph as vibration-based road monitoring systems. Probably in long term the camera-based solution will realize road monitoring,

considering today's strong trend on vision systems using deep neural networks, however it is still an open scientific question.

The "*Concept of the novel system*" paragraph describes the basics of the Ph.D. idea together with its hardware and software realization.

The "*Implementation of the software framework*" paragraph (probably a title similar to e.g. "Calculation algorithm chain" would be a better for this methodological paragraph, software is "only" a realization of the detailed processing functions) is the first core part of the thesis detailing the individual information processing steps, the following preprocessing sequence was described in detail:

1. GPS data upsampling,

2. trajectory resampling and smoothing,

3. bias and trend removal from accelerometer and gyroscope signals,

4. sensor synchronization verification,

5. spatial domain resampling,

6. and infinity norm normalization.

It is very interesting how many considerations and solutions are needed to establish a reliable and robust road condition monitoring solution (e.g. trend removal, frequency harmonizations, etc.).

The effects of aggressive maneuvers was one of the factors that was eliminated from the monitoring system – so, an aggressive maneuver detection solution was introduced using the same measurement signals. A relatively simple solution was applied at the end.

Principal Component Analysis (PCA) helped to reduce the dimensionality of the problem (capturing 92.2% of the total variance), moreover, the introduced change point detection solution enabled the variation of applied signals at different segments.

A sophisticated feature extraction methodology was also described.

The introduced Machine-in-the-Loop annotation method for road qualification realizes a balance between the large-scale labeling of road segments and the relatively low-effort human involvement.

Finally, the fourth paragraph details the quality measures for clustering and the reverse geocoding solution.

The “Results and discussion” paragraph describes and evaluated the concrete results of the introduced system. The “two-car” realization (available at ZalaZONE) mirrors good robustness of the introduced system. This paragraph serves with all the concrete results of the research ensuring the basis for the three theses. 82 kms of training sample was applied and 267 kms of testing and validation were used near Zalaegerszeg. The maneuver detection was well realized, Figure 24. shows good separation of clusters of different maneuver types. The maximal of 10% error in the maneuver recognition is an acceptable solution (Figure 25.). In quality clustering, it was interesting that four clusters were found as the best (vs. IQL proposes “only” three clusters), robust separation for road quality. Finally, the Random Forest method served with the best classification accuracy (and F1 score). Having a classification model, the “lift” effects on the accuracy was determined applying the introduced maneuver detection algorithm and also the additional value of using parallel sensors was also determined. The importance of features and the confusion matrix shows that the results are at the same level as the *best* state-of-the are solutions in the scientific literature.

3. **Minor weaknesses**

The thesis is a well-prepared, robust document; however, some smaller correction advice can be appointed:

- Features were evaluated e.g. by SHAP values, however, the classical Feature Selection (FS) techniques could be applied, that could enable generating hundreds of descriptive features and selecting the best among them. Unfortunately, FS is not mentioned in the thesis, moreover this extended approach will certainly increase the accuracy of the results road qualification for sure, moreover it can serve with additional knowledge. This could be a direction for further research.
- The thesis contains 94 references that is a good number, however, probably the highest weakness is that at many points inside the document the described statements are not sufficiently supported by references. E.g. “While some studies do not specify their labeling methodology for real-world experiments,

and others rely entirely on computer simulations…"; "Furthermore, affordability is critical for widespread deployment, especially across diverse vehicle fleets."; "In this study, "x = 0.2 m was selected, with a trade of between sufficient detail for localized terrain analysis and oversampling."; "Since road vehicles operate within a limited dynamic range, previous studies…"; "…which are the two core indicators of surface uniformity and variability…"; "This methodology has already proven successful across various domains."; "BIC", "AIC"… (page 78) – no reference for them ; "Table 3: Absolute threshold values applied for the determination of maneuver classes along the main axes"; In the second part of the document (e.g. method/algorithm description and also in the results) there are statements and decisions that were not supported by appropriate references.

- "The selection of sensor locations was driven by practical engineering constraints rather than an explicit optimization procedure." – the beginning of the document addressed an optimal sensor selection approach.
- "…infinity norm normalization was employed.". It is very risky to apply this kind of normalization because it is very sensitive to (e.g. only one) outlier(s). So, the 0 mean, 1 standard deviation normalization is more robust.
- At "Table 5: SMOTE algorithm steps" the variable 'x' should be probably 'x(min)'.
- "After comparative analysis mainly based on classification accuracy and robustness" – these comparisons were not described.

4. Some open questions

- At "Mathematically, the correction can be expressed as: corrected(t) = s(t) · w(t), where s(t) is the original signal and w(t) → [0, 1] is the Gaussian-weighted correction mask." – how the values for w(t) were determined?
- "The BIC and AIC criteria were applied to evaluate a series of GMM models trained with varying numbers of clusters. The model corresponding to the lowest BIC value was selected as the final configuration" – what is the argumentation supporting this selection?
- "The preprocessing pipeline started with the removal of sensor errors" – what was the method for this removal?

- Figure 25: What is the reason that acceleration and break have the higher classification error?
- “These observations and experiences are supported by my visual observations.” – what does it mean?

5. Evaluation of the theses

All of the theses (three theses) are accepted (all of them are supported by publications appropriately), however some components of them are not accepted:

1. ***Thesis #1*. I investigated the added information content of vehicle cabin and suspension sensor data for vibration-based road surface quality measurement and classification using a custom-developed multi-sensor measurement system.**
    a. ***This thesis is accepted*.**
        - *Robust scientific results, supporting the acceptance of the thesis*:
            - Among six tested models, the Random Forest classifier achieved the best performance”.
            - “The final classification accuracy was 91.2%, and the area under the ROC curve (ROC–AUC) was 98.6%.” – In the literature review it was identified: “The average classification accuracy from the reviewed studies was on average between 85% and 93%.”, consequently, this Ph.D. results are among the best in the state-of-the-art literature (but not beyond).
            - “The incorporation of suspension acceleration data resulted in a 6.9% increase in classification accuracy compared to models trained exclusively on cabin acceleration and angular velocity signals.”
    b. *The following parts of the thesis are not accepted*:
        - “• I developed and built a data acquisition system for vehicle vibration measurements….” – the complete bullet point.
            - Reason: it is a technical development;

- "I conducted long-term vehicle vibration measurements on public road networks mainly in Zala and Vas counties,…." – the complete bullet point.
    - Reason: it is a test progress description;

2. ***Thesis #2.*** **I introduced and evaluated the applicability of a Machine-In-the-Loop (MIL) annotation method for vibration-based road surface quality classification.**
    a. ***This thesis is accepted*****.**
        - *Robust scientific results, supporting the acceptance of the thesis*:
            - "Due to the absence of ground-truth annotations for the measured road sections, a semi-automated Machine-In-the-Loop annotation approach was developed." & "Machine-In-the-Loop annotation approach is suitable for generating reliable training data for supervised models in vehicle vibration-based road surface quality assessment."
            - "I demonstrated that the developed MIL-based clustering algorithm was capable of grouping road sections of varying lengths into coherent clusters…".
            - "In addition to manual verification, the optimal number of clusters and thus the number of distinguishable road quality classes were determined".
            - "I established that the balance between model complexity and cluster robustness was achieved for 3 <= K <= 5" (small typo in the thesis).
            - "…four identified quality classes, approximately 80% of all data points were assigned to clusters with confidence above 90%."

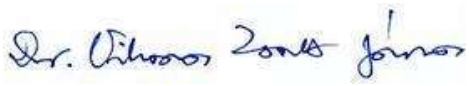
3. ***Thesis #3. I investigated the negative impact of lateral and longitudinal aggressive vehicle maneuvers on vibration-based road surface quality assessment and introduced a detection and correction method to mitigate the signal-distorting elects of such abrupt maneuvers.***
    a. ***This thesis is accepted*****.**
        - *Robust scientific results, supporting the acceptance of the thesis*:

- “A machine learning–based algorithm was developed to detect such abrupt driving events, focusing on four primary maneuver types: left turn, right turn, sudden braking, and sudden acceleration.”
- “The results demonstrated higher accuracy for lateral maneuvers, with the right-turn class exhibiting the lowest error rate. Longitudinal maneuvers, such as braking and acceleration, yielded slightly higher errors, particularly for the hatchback vehicle, which may be attributed to its softer suspension characteristics. The mean classification error was 7.3% for the Nissan Leaf and 6.7% for the BMW 318is, resulting in an overall average classification accuracy of 93%.”
- “…the occurrence and impact of aggressive maneuvers…may distort the vibration signals and negatively a!ect road surface classification accuracy.”
- “…during normal, steady driving on straight roads, the dominant features influencing classification were vertical acceleration and longitudinal angular velocity, both of which are directly related to the pavement condition. However, during sudden turning maneuvers, the dominant signal component shifted to the yaw angular velocity around the vertical axis, which represents vehicle motion rather than road quality. These observations confirmed that maneuver detection and subsequent correction are necessary…”
- “I introduced a correction method based on the SHAP analysis results….For abrupt lateral maneuvers, the transverse acceleration and yaw angular velocity variables were downscaled, while for longitudinal maneuvers, the longitudinal acceleration and pitch angular velocity components were corrected. The correction was applied proportionally to the temporal proximity of the maneuver events, ensuring smooth signal recovery. As a result of this

correction method, the overall classification accuracy of the vibration-based road quality model improved by 2.9%.”

As conclusion, congratulations for the scientific research activities and for the nice scientific results, all of the three theses were accepted by me, consequently, *after the successful defense I suggest awarding Roland Nagy the Ph.D. title*.

Budapest, 2026.06.14



Dr. Zsolt János Viharos