Distributed Road Surface Condition Monitoring Using Mobile Phones

Mikko Perttunen[1], Oleksiy Mazhelis[2], Fengyu Cong[2], Mikko Kauppila[1], Teemu Leppänen[1], Jouni Kantola[3], Jussi Collin[3], Susanna Pirttikangas[1], Janne Haverinen[1], Tapani Ristaniemi[2], and Jukka Riekki[1]

[1] Computer Science and Engineering Laboratory, University of Oulu
P.O. BOX 4500, Oulu, Finland
[2] Faculty of Information Technology, University of Jyväskylä
P.O. BOX 35, Jyvaskyla, Finland
[3] Department of Computer Systems, Tampere University of Technology
P.O. BOX 553, Tampere, Finland

Abstract. The objective of this research is to improve traffic safety through collecting and distributing up-to-date road surface condition information using mobile phones. Road surface condition information is seen useful for both travellers and for the road network maintenance. The problem we consider is to detect road surface anomalies that, when left unreported, can cause wear of vehicles, lesser driving comfort and vehicle controllability, or an accident. In this work we developed a pattern recognition system for detecting road condition from accelerometer and GPS readings. We present experimental results from real urban driving data that demonstrate the usefulness of the system. Our contributions are: 1) Performing a throughout spectral analysis of tri-axis acceleration signals in order to get reliable road surface anomaly labels. 2) Comprehensive preprocessing of GPS and acceleration signals. 3) Proposing a speed dependence removal approach for feature extraction and demonstrating its positive effect in multiple feature sets for the road surface anomaly detection task. 4) A framework for visually analyzing the classifier predictions over the validation data and labels.

Keywords: accelerometer, signal processing, pattern recognition, support vector machine, classification, road roughness, GPS.

1 Introduction

The need to reduce fuel consumption, traffic accidents, congestion as well as making public transportation more efficient are some of the problems faced worldwide. Aside from developing more efficient motors and more ecological fuels, making traffic infrastructure and vehicles more efficient through the use of advanced information technology is being studied widely.

Friction of the road surface is the most important environmental factor affecting safety. Road surface quality can be characterized using microtexture and

C.-H. Hsu et al. (Eds.): UIC 2011, LNCS 6905, pp. 64–78, 2011.

macrotexture, which contribute to friction, and megatexture and roughness that are formed by the combination of stone particle surfaces and by the gaps between stones used in the surfacing material, respectively [5]. Megatexture refers to potholes, joints, patching, cracks and other small surface defects that cause increased noise levels and rolling resistance. Roughness refers to large surface unevenness that increases vehicle operating costs and decreases driving comfort. In the nordic countries, frost heave causes both seasonal variation in road surface condition as well as permanent cracks and bumps [7].

To manage the quality of the road condition, administrators use special instrumented vehicles to measure the road roughness periodically [4]. To cope with the burden of dedicated condition monitoring, researchers have proposed the use of built-in sensor systems of new passenger vehicles [12] to detect road condition. However, the penetration of vehicles with such integrated sensors will be low at least for the next decade [3] and there is no standard interface to access onboard sensors, rendering it difficult to implement a generic solution using these sensors. Mobile devices with integrated or external sensors provide an alternative traffic sensing and communication system [16]. We envision collecting information about all traffic using mobile phones, for example, pedestrians, bicyclers, and passenger cars.

In the present research, we focus on developing a pattern recognition system for recognizing road surface anomalies that contribute to road roughness, using data from mobile phones installed in vehicles. The problem we consider is to detect road surface anomalies that, when left unreported, can cause wear of vehicles, lesser driving comfort and vehicle controllability, or an accident. It should be noted that this goal is complementary to the usual road roughness evaluation [4], since we consider it important to recognize also individual severe anomalies, not only to categorize the road segments according to average roughness. This work sheds light on the feasibility of the road condition monitoring using sensor-enabled mobile phones and low sampling rates.

The rest of this paper is organized as follows: In the next section we summarize related work. In section 3 we describe data collection setup and initial analysis. Section 4 describes the anomaly recognition system and in section 5 we present the results. In section 6 we discuss and conclude the paper.

2 Related Work

Road roughness is typically measured using special instrumented vehicles [4]. In [12] the authors studied, through simulation, the use of vehicle integrated accelerometers to analyze road roughness. Mobile devices with integrated or external sensors have been proposed as a surrogate traffic sensing and communication system [16]. The suitability of tri-axis accelerometers in an embedded system [11] and integrated in a mobile phone [16] have been experimented together with GPS receivers to recognize road surface anomalies, such as potholes and bumps. In [2] GPS-enabled mobile phones were used to collect traffic information and to estimate the traffic situation.

Traffic conditions monitoring of Google Maps for Mobile [1] is based on using GPS data on mobile phones to estimate vehicle speed. Similarly, in TJam [17], congestions were predicted using estimated vehicle speeds. VTrack [20] utilized location tracks of mobile phones to estimate travel times in real time. In [21] a system was presented to detect car accidents using on-board mobile phones and attemps at filtering out false positives such as high accelerations due to dropping the phone were described.

Using mobile phones makes a large set of sensors readily available, enabling more reliable and exact information, with larger coverage. This kind of setting where citizens contribute to collecting data about the environment has been called participatory sensing [6].

The previous work [16] has identified labeling the road surface condition accurately as a difficult task. We tried to overcome this problem by spectral analysis of the 3D acceleration signals. In the literature, multiple classifiers have been applied as an attempt to solve the problem of speed dependency of accelerometer readings, that is, the fact that driving over the same road surface anomaly with different speeds results in different signal patterns [16]. We propose a robust feature extraction approach which removes speeds dependency of the features. Moreover, while in previous work simple thresholds on various features have been used in anomaly detectors, we use support vector machines [9] to classify road segments. We present also a visualization framework for the classification results to enable visual inspection of, for instance, examples close to the class border.

3 Data Collection

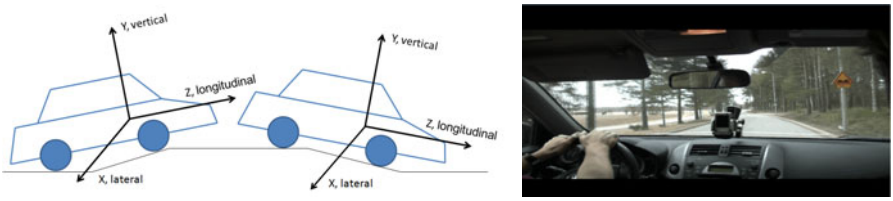
3.1 Data Collection Hardware and Software

The Nokia N95 8GB mobile phone was used to collected data. Accelerometer samples were recorded at 38Hz and GPS readings stored at 1Hz. The GPS readings included only latitude, longitude and timestamp. The data collection tool was written in Java, with a native component for accessing the accelerometer.

3.2 Data Collection Setup

Data was collected using a mobile phone that was attached to a rack on the windshield of a vehicle. The rack was carefully positioned and secured to maintain approximately the same accelerometer coordinates across data collection drives. Camcorder was attached to the head rest of passenger's seat. Fig.1b shows the view of the camcorder, including the phone rack. The accelerometer orientation is shown in Fig.1a.

We have used several vehicles for data collection, however, in this paper we report results on using a single passenger car, for which an accurate - and hence laborous - labeling has been carried out.



(a) (b)

Fig. 1. a) Accelerometer orientation. b) Mobile phone setup.

3.3 Collected Data Sets and Labeling

Several data collection drives were performed. For the work reported in this paper, a drive of about 40 minutes in length (25km) was selected among the drives, as it seemed to contain the most diverse anomalies based on a quick route review.

Two label files were produced by two labelers independently, using the figures (described below) and video. One label file contains the anomalies on the road surface and the other contains the surface types. Two examples of the plots used for labeling and the created labels are shown in Fig.2. However, in the real setup, multiple segments of the full spectrogram were plotted in the same figure (page) to make it easier for the labellers to compare the magnitudes of signals and spectral energies across longer periods. Labeling process was as follows: first, preprocessing for the GPS and acceleration signals was done. Second, a spectrogram of the sum of the band energies of accelerometer signals was plotted and the preprocessed accelerometer signals were plotted on top of the spectrogram. Also speed was superimposed in the same figure. Most of the labelled anomalies were confirmed from the video, but there very a few clear signal artifacts that were not confirmed from video. For each labelled anomaly, also a textual description was added. After the independent labelling, the label sets were merged.

When labeling, the anomalies were categorized into two classes according to their severity. Type 1 represents small potholes, rail road crossings and other road surface roughness. Type 2 represents a) man-made speed bumps and other road surface artifacts directing drivers to slow down and b) severe anomalies that might cause accidents or vehicle breakdown when driven on at a high speed. In this we focus on detecting Type 2 anomalies from asphalt roads. Note that man-made artifacts are included for two reasons: Firstly, their automatic recognition enables adding them to digital maps and thus, warning other drivers. Secondly, they cause signal patterns that are very similar to anomalies caused by worn down road surface, or damage by frost. For these reasons, they represent useful data to build a recognition system. However, type 1 anomalies were also labelled as accurately as possible to enable assessing their contribution to the detection accuracy.

Cobblestone segments from the data set were discarded. This is justified by the fact that the road surface type is available from road databases and having

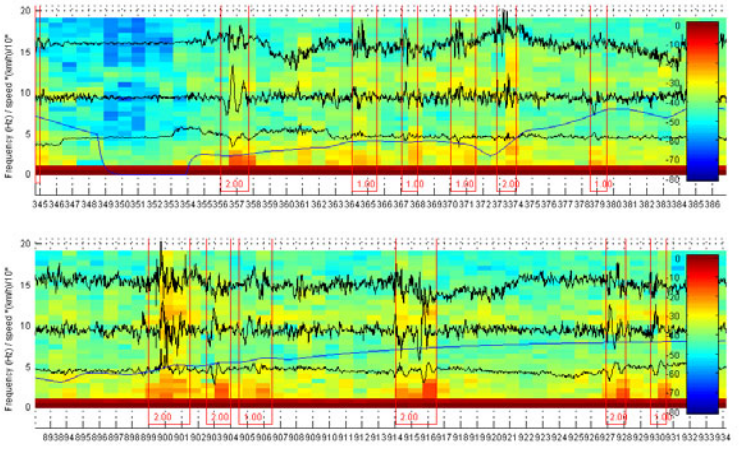


Fig. 2. Two example segments with labels. The sum spectrograms of 3D acceleration signals are plotted with original acceleration signals and speed.

Table 1. Anomaly statistics

Category	Example types	Count	Total length (s)
Type 1	railroad crossing, small pothole, roughness	184	613.9
Type 2	speed bump, bump, pedestrian crossing with cobblestones, large pothole	42	81.6

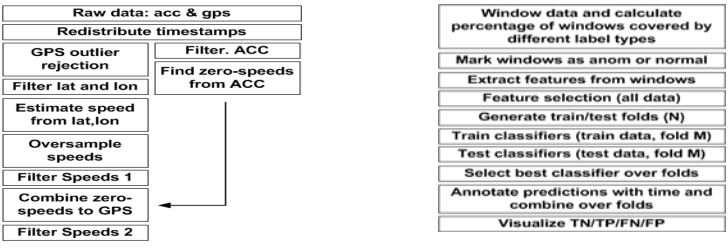
cobblestone in the data can bias the results, since there can be an unusual proportion of "'anomalous"' road surface.

Information about the data sets are shown in Table 1. The mean length of Type 2 anomaly was 1.94s, the maximum and minimum being 5.0s and 1.0s, respectively.

4 Anomaly Recognition

The proposed road surface anomaly recognition system is part of our cooperative traffic sensor network middleware [15] and is based on tri-axis acceleration and GPS data. In brief, GPS was used to estimate speed and several other features were extracted from the acceleration signals. Positioning data was also used to visualize the results on a map. The feature set was used to recognize road surface anomalies and to filter out other similar signal artifacts caused by door slams and jerks at the end of braking, for example.

To put the system into a context, in comparison to [11], we propose an additional preprocessing step: before running the anomaly detection algorithm, our



(a) (b)

Fig. 3. a) Preprocessing flowchart. b) Windowing and classification flowchart.

(a) (b)

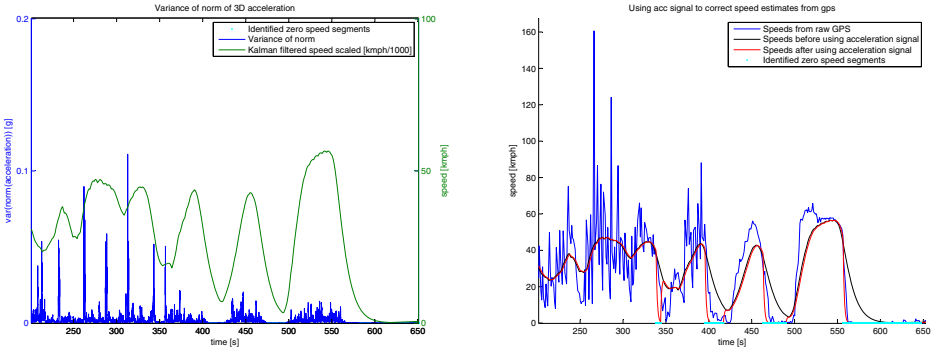
Fig. 4. a) A Segment of raw GPS signal on a map. b) A segment of preprocessed GPS on a map.

system aims to recognize the means of travel of the user [13], [15]. Based on recognizing the means of travel, the system aims to recognize and report road surface anomalies only when the phone is in a car.

4.1 Preprocessing

Using mobile phone sensors was reflected in the signals in two ways: Firstly, the GPS signal of the used phone was very noisy. Secondly, both the GPS measurements and the acceleration measurements were contamined by bursts, which are measurements recorded with the same timestamp. We believe the reason for such bursts is the way that process switches and priorities are handled in the phone.

Fig.3 shows an overview of the anomaly recognition system as a block diagram. As shown in the figure the timestamps of the acceleration and GPS measurements were first redistributed evenly within short time segments. Next, GPS outlier rejection was done and Kalman filter was applied to latitude and longitude to further reduce noise. Example segment of the raw and preprocessed latitude and longitude are shown in Figs. 4a and 4b. In this example, the vehicle was momentarily stopped on a parking lot on the left side of the road.



(a) (b)

Fig. 5. Data collection. a) Fusion of acceleration signals and latitude and longitude based speed estimation through identifying zero speed segments from the acceleration signal. b) Estimated speed signal after applying the identified zero speed segments.

Next, speed was estimated from each consequtive latitude and longitude pairs. The resulting speed signal was oversampled and filtered using reasonable physical limits for acceleration of the vehicle. We used as limits for maximum acceleration 3.5 m/s^2 and for maximum deceleration -11 m/s^2.

While experimenting with the previous steps we noticed that regardless of tuning the filters, we could not reach satisfactory speed estimates. When speed estimates seemed smooth enough, the signal was often contaminated by large latency compared to original (noisy) speed estimate and the speed failed to reach zero on vehicle stops, see Fig.5a. To alleviate the problem, we used two corrections. First, we removed a visually determined latency from the filtered speed estimate. Second, we applied a very simple fusion of the acceleration and GPS signal: the variance of the norm of the acceleration was calculated for the whole acceleration signal and visually examined, see Fig.5b. By simple thresholding, we were able to detect segments of the signal, where speed of the vehicle was zero, or very close to zero. Next, we set the corresponding segments of the speed signal to zero and smoothed the speed signal by applying a Kalman filter. The correctness of the speed estimate was checked to the extent possible from the video. The temporal accuracy of the stops were the easiest to confirm. We recognize a more careful analysis of the preprocessing approach as future work.

4.2 Windowing

As shown in Fig. 3b, the data was framed using a sliding window. We experimented with multiple frame lengths from 0.5s to 2s, a scale we assumed suitable for the anomaly recognition task, as the mean anomaly length was around 2s. For each window, we determined the percentage of the window covered by anomalies (one or more together). In the experiments reported in this paper, we marked

the window as representing anomalous road, when it was covered by more than 50% by anomalies. Temporal order of the windows was kept throughout the processing.

4.3 Feature Extraction and Selection

Feature extraction was done using sliding windows of 2.0s in length, the slide being 0.5s. Several features were extracted from acceleration signals: standard deviation, mean, variance, peak-to-peak, signal magnitude area, 3-order autoregressive coefficients, tilt angles and root mean square for each dimension. Absolute value of correlations of signals between all dimensions were used as well, since it was visually observed that often all the acceleration signals showed similar waveforms in the anomaly segments. This can be seen, for example, in the first anomaly in Fig.2. However, this is not the case when only one side of the car hits a pothole.

Fast Fourier Transformation (FFT) based features were used in order to incorporate information from specific frequencies. This was based on the assumption that bumps and potholes would produce lower frequency components in comparison to vibration originating from the motor and normal road surface. FFT energy was extracted from 17 frequency bands for each acceleration direction (as shown in Fig.2) and mel frequency cepstral coefficients in 4 bands.

We utilized the backwards feature selection algorithm of PRTools [10] to select the optimal feature sets for both the speed scaled and non-speed scaled feature sets.

4.4 Removing Speed Dependence of Features

Most of the features vary as a function of speed. As an example, Fig.6 shows the speed-dependency of the peak of the Y signal. This dependency is considered harmful for classification, because data points at a slow speed may look much different than points at a high speed. However, it is clear that speed cannot by itself be used to classify road surface anomalies, since ideally anomalies can occur equally likely at any speed. Thus, to remove linear dependency on speed, we first fit a line

$$y = ax + b_0 \tag{1}$$

to each feature of the data. Then, we form a new linearly (speed) independent data set

$$\varepsilon_i = y_i - b_0 - ax_i. \tag{2}$$

This method for removing linear dependence is described in the general case in [19].

Fig.6 shows the effect of removing speed dependency for the y-peak feature. This is further illustrated in Fig.7 where both speed dependent (a) and speed independent (b) versions of y-peak and energy of y signal on band 3 (from 2.2Hz

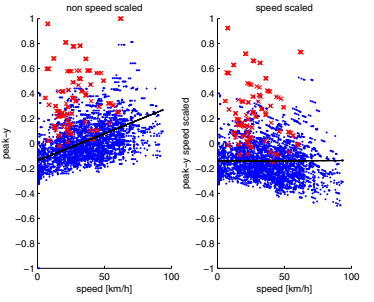
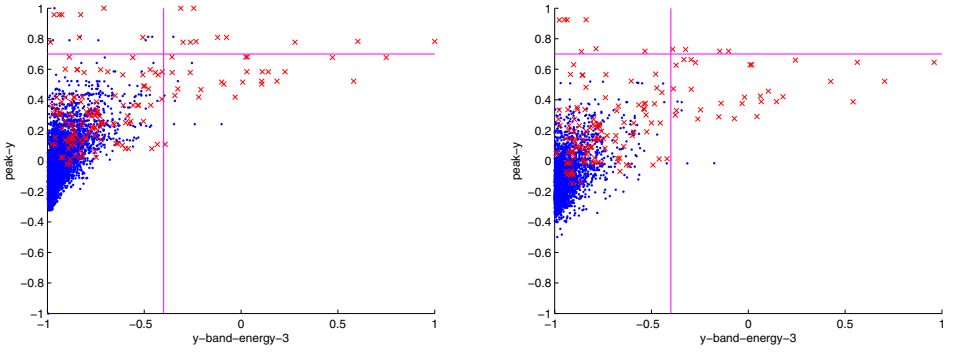


Fig. 6. Peak of y acceleration as a function of time in normal and anomalous windows with a line fitted to the data points. Anomalies are shown as red crosses and normals as blue dots.



(a) (b)

Fig. 7. Peak of y signal plotted against signal energy of y on band 3. Anomalies are shown as red crosses and normals as blue dots. a) Speed dependent. b) Speed independent. The vertical and horizontal lines help illustrating the discrimination improvement due to speed independence removal. Note that there are no normals above the horizontal line in b).

to 3.6Hz) are plotted against each other. The horizontal and vertical lines were added to make it easier to read the figure, that is, to see that more anomalies become linearly separable from the normal road after removing speed dependency.

4.5 Cross-Validation and Evaluation Criteria

All experiments were performed using 5-fold cross-validation, where the training set of each fold contains 4/5 of the total anomaly windows and 4/5 of the normal windows (1/5 for the test set). The folds were created from consequtive windows. We did not use random selection, because then examples from the same anomaly (or same normal) segment could end up in both test and training sets.

Because the datasets are skewed, accuracy is not a reliable measure of model goodness. Instead, we use the geometric mean of accuracies on positive and negative examples, the G-means metric, suggested in [14]. The metric is defined as

$$g = \sqrt{sensitivity \cdot specificity}. \tag{3}$$

Models producing a high G-means have balanced sensitivity and specificity, thus the effect of dataset skew is diminished when using G-means as evaluation criteria. To further illustrate the classification performance, we report also relative sensitivity (RS), a measure proposed by Su and Hsiao [18] to evaluate the relative accuracy on positive and negative examples; the relative sensitivity is defined as

$$RS = \frac{sensitivity}{specificity}. \tag{4}$$

The RS value should be around one for the classification accuracy to be of similar order for the positive and negative classes.

4.6 Classifier Training and Evaluation

To classify the windows representing short road segments we use support vector machines [9], one of the most widely applied classification method. To train a soft margin support vector machine for N input vectors $\mathbf{x}_1 ... \mathbf{x}_n$ with labels $t_n \in \{-1, 1\}$, generally the quadratic optimization problem

$$C \sum_{n=1}^{N} \varepsilon_n + \frac{1}{2} \parallel \mathbf{w} \parallel^2 \tag{5}$$

has to be solved. Here, the slack variables $\varepsilon_n \geq 0$ and $t_n y(\mathbf{x}_n) \geq 1 - \varepsilon_n$ for n = 1,...N. The parameter C controls the cost of misclassification - the larger its value, the more the SVM fits to training data. To enable nonlinear classification boundaries, kernels are used. We utilize the radial basis kernel,

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\gamma \parallel (\mathbf{x}_i - \mathbf{x}_j) \parallel^2\right), \tag{6}$$

where $\gamma > 0$.

As the data sets were dominated by normal road surface with no anomalies, the resulting sets of examples were heavily biased. To improve the recognition results we used a different misclassification weight for the normal and anomaly classes. The misclassification weight for the normal and anomaly classes were set according to their counts in the train dataset, that is,

$$\frac{c_-}{c_+} = \frac{n_+}{n_-}, \tag{7}$$

where c_- is the weight for normal class, c_+ is the weight for the anomaly class, and n_- and n_+ are the corresponding class counts in the training set.

The classifier training and testing procedure is illustrated as a part of the whole process in Fig 3a. Using the test data, we determined the area under ROC curve (AUC) for each classifier. The classifier with the best average G-means over folds was selected. We evaluated 49 SVM parameter combinations using radial basis function kernels (RBF) in LIBSVM [8], corresponding to a grid of values of the parameters γ and C. The parameter γ controls the width of the RBF kernel.

5 Results and Analysis

The classification results for different feature sets are shown in Table 2. Two observations were be done: speed dependence removal improves the classification performance significantly and the feature subset selected by using the backward search procedure worked at least as well as the full feature set.

5.1 Visualization of Recognition Results

Fig. 8 shows the labeled anomalies of two segments of the test set with the predictions superimposed. To avoid clutter, true negatives are not shown. The probability of an anomaly according to the classifier is shown as a continuous red signal.

To further illustrate the results, the predictions are shown on a map in Fig.9. It should be noted from Fig.9c that many of the predictions marked as false positives aren't, strictly speaking, false positives but represent an anomaly of Type 1 or a window partially overlapping with an anomaly of Type 2.

5.2 Result Analysis

The confusion matrix for the best result of Table 2 is shown in Table 3. Note that some of the false positives were a 'natural' consequent of the sliding window approach - some of the windows were overlapped by part of Type 1 anomaly or Type 2 anomaly (see Table 1), but overlapping less than 50% of the window (recall that a window was labelled as anomalous when 50% of it was covered by a labelled anomaly). In this case, there were 127 such false positives, the mean

Table 2. Evaluated feature sets

Name	G-means	AUC	Sens.	Spec.	FPR	FNR	RS
all features	0.68	0.96	0.49	0.99	0.01	0.51	0.49
norm-based features	0.54	0.90	0.32	0.99	0.01	0.68	0.31
95 backward selected	0.67	0.97	0.48	0.99	0.01	0.52	0.49
all features, speed scaled	0.77	0.98	0.61	0.99	0.01	0.39	0.61
norm-based features, speed scaled	0.67	0.96	0.47	0.99	0.01	0.53	0.47
95 backward selected, speed scaled	**0.89**	0.97	**0.82**	0.97	0.03	**0.18**	**0.84**
20 backward selected, speed scaled	0.79	0.98	0.63	0.99	0.01	0.63	0.64

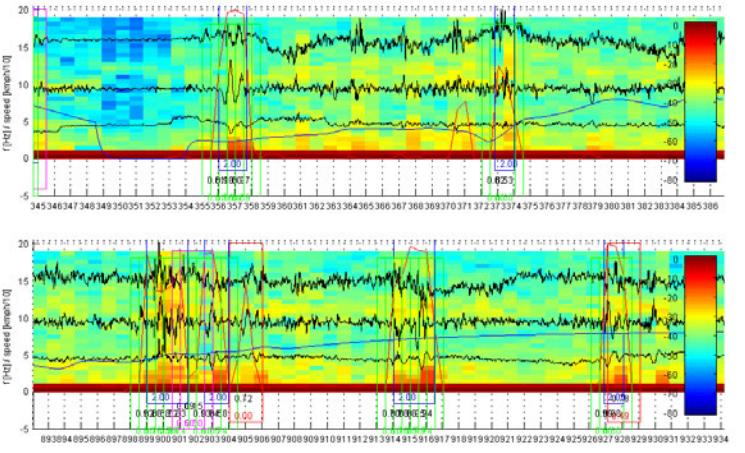
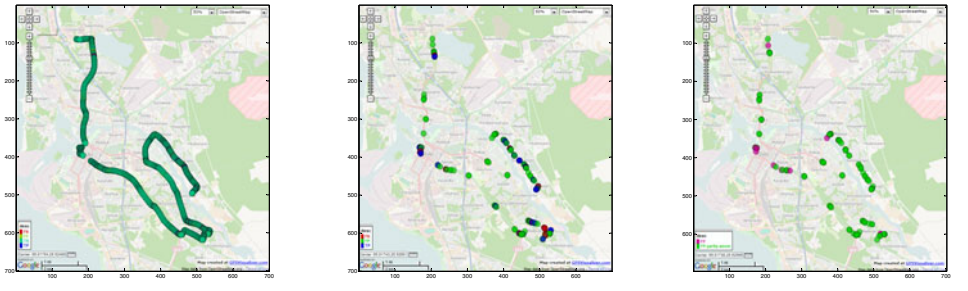


Fig. 8. Two example segments with predictions shown on top of original spectrograms, signals and labels. The original labels are shown as blue rectangles, the true positive windows as green, false positives as red and false negatives as magenta rectangles, correspondingly. The probability of an anomaly according to the classifier is shown as a continuous red signal.



(a) (b) (c)

Fig. 9. Predictions on map. a) All predictions. The breaks represent the removed cobblestone segments. b) False negatives (red), false positives (green) and true positives (blue). c) False positives. The proportion of them overlapping with a Type 1 or Type 2 anomaly are shown in green and the windows not overlapping with an anomaly in magenta.

overlap being 59%. An example can be seen in the bottom of Fig.8, where a false positive prediction occurs after the first labelled anomaly; this is a Type 1 anomaly, as can be seen from Fig.2.

In Table 3 the number in parenthesis after FP signifies the number of false positives after removing the 127 windows that overlapped with an anomaly.

Table 3. Confusion matrix of anomaly detection using 95 features selected by backward search (γ : 1.0 and C : 0.1). The number in parenthesis after FP signifies the number of false positives after removing the (127) windows that overlapped with an anomaly.

	Anomaly	Normal
Anomaly	139	140 (13)
Normal	31	4440

In total, 1623 windows in the normal data overlapped to some extend with a labelled anomaly of Type 1 or Type 2. The mean overlap of the false negatives with the Type 2 anomalies was 71%, while for true positives this was 82%.

The results indicate that Type 2 anomalies could be quite accurately be distinguished from good condition road surface, but the Type 1 and Type 2 anomalies could not be well discriminated. However, this result suggests that locations where a significant amount of both Type 1 and Type 2 anomalies occur will be easy to point out. This is important for road administrators, because maintenance resources can be allocated to where they have the largest effect. It should also be noted that when deployed in practice, the classifier can be followed by a filter that, using a database, removes known speed bumps, cobblestone roads and other segments that are not actually in need of repair but do cause very similar signal patterns to actual road surface anomalies.

5.3 Comparison to Related Work

Similar results were reported in [11] and [16], although exact comparison is not possible due to different data sets and test setups. In [16], dedicated bump detectors were trained for two different speed categories using both known and unknown orientation of the phone. The authors rated as one of their best performing systems a bump detector with FPR=3% and FNR=51%, whereas our best result in terms of those statistics was FPR=3% and FNR=18% – a large reduction in false negatives. In [16], the test setup was not fully described; for example, it was not told whether locations of predictions were compared to locations of labeled bumps, or if time based labels and predictions were used. The data sets were compatible in length, in [16] around 30km and in our data around 25km.

It should also be mentioned that the results on labeled data in [11] were not comparable to ours, because in their data set, unrealistically, the number of examples from normal road surface roughly equalled the number of pothole examples, and they did not report the number of examples from normal road surface predicted as potholes or other anomalies. We used all examples from a continuous drive in the testing.

6 Conclusion

This work deals with monitoring road condition using sensors embedded in mobile phones. We have performed initial data collection and analysis, and

developed road surface condition monitoring system that includes preprocessing, classification and visualization stages. The system was demonstrated to perform favourably when evaluated against a data set which is comparable in size to previously published results in the literature [16].

In this paper we utilized spectral analysis of 3-dimensional acceleration signals in order to get reliable road surface anomaly labels. Preprocessing of GPS and acceleration signals were used to estimate speed and to reduce sampling errors. Moreover, we proposed a speed dependence removal approach to the feature extraction in order to get more robust features. We demonstrated that the speed dependence removal improves the performance of several feature sets for the road surface anomaly detection task. A framework for visually analyzing the classifier predictions over the validation data and labels was presented. Compared to earlier work [11], [16], where only final detections were presented on a map, the framework is clearly advantageous as it allows analyzing what kind of waveforms were falsely classified.

The current recognition performance is not completely satisfactory, but as suggested in [11], collecting data from multiple drivers and clustering the suspected anomalies based on location and requiring an anomaly to be detected a number of times improve the performance. With limited data, we were unable to repeat the suggested clustering experiment, but we plan this as future work.

A practical lesson learned is that labelling driving data from video is time consuming and error prone work. Thus, alternative ways of developing surface anomaly detectors should be studied. For example, unsupervised learning could be experimented with.

Due to the inadequate quality of the data, we plan to experiment with other mobile phones to study the variance of sensor quality between devices and to confirm the results on new datasets. Future work will also address recognizing overall road condition from multiple vehicles of unknown types. This requires robustness against the different dampers and chassis of vehicles. Should this be achievable, the mobile phones based sensors could be carried in any vehicles, producing useful data without pre-configuration.

Acknowledgment. This work was supported by TEKES as part of the Cooperative Traffic ICT program of TIVIT (Finnish Strategic Centre for Science, Technology and Innovation in the field of ICT). The first author would also like to thank the Nokia Foundation.

References

1. Google maps for mobile (2009), `http://www.google.com/mobile/maps`
2. The mobile millenium project (2009), `http://traffic.berkeley.edu/`
3. Final report and integration of results and perspectives for market introduction of ivss, eimpact consortium (August 11, 2008)
4. Comparative performance measurement: Pavement smoothness. NCHRP 20-24(37B), American Association of State Highwayand Transportation Officials (AASHTO) (May 18, 2008)

5. Alauddin, M., Tighe, S.L.: Incorporation of surface texture, skid resistance and noise into pms. In: Proc: 7th International Conference on Managing Pavement Assets. Calgary, Canada (June 24-28, 2008)
6. Burke, J., Estrin, D., Hansen, M., Parker, A., Ramanathan, N., Reddy, S., Srivastava, M.B.: Participatory sensing. In: Workshop on World-Sensor-Web (WSW 2006), pp. 117–134 (2006)
7. Byrne, M., Albrecht, D., Sanjayan, J.G., Kodikara, J.: Recognizing patterns in seasonal variation of pavement roughness using minimum message length inference. Computer-Aided Civil and Infrastructure Engineering 24(2), 120–129 (2009)
8. Chang, C.C., Lin, C.J.: LIBSVM: a library for support vector machines (2001), software `http://www.csie.ntu.edu.tw/~cjlin/libsvm`
9. Cortes, C., Vapnik, V.: Support-vector networks. Machine Learning 20, 273–297 (1995)
10. Duin, R.P.W., Juszczak, P., de Ridder, D., Paclik, P., Pekalska, E., Tax, D.: Prtools, a matlab toolbox for pattern recognition (2004), `http://www.prtools.org`
11. Eriksson, J., Girod, L., Hull, B., Newton, R., Madden, S., Balakrishnan, H.: The pothole patrol: using a mobile sensor network for road surface monitoring. In: MobiSys 2008: Proceeding of the 6th International Conference on Mobile Systems, Applications, and Services, pp. 29–39. ACM, New York (2008)
12. Gonzlez, A., O'brien, E.J., Li, Y.Y., Cashell, K.: The use of vehicle acceleration measurements to estimate road roughness. Vehicle System Dynamics: International Journal of Vehicle Mechanics and Mobility 46, 483–499 (2008)
13. Kantola, J., Perttunen, M., Leppänen, T., Collin, J., Riekki, J.: Context awareness for gps-enabled phones. In: ION 2010 Technical Meeting (January 25-27, 2010)
14. Kubat, M., Matwin, S.: Addressing the curse of imbalanced training sets: One-sided selection. In: Proceedings of the Fourteenth International Conference on Machine Learning, pp. 179–186 (1997)
15. Leppänen, T., Perttunen, M., Riekki, J., Kaipio, P.: Sensor network architecture for cooperative traffic applications. In: 6th International Conference on Wireless and Mobile Communications, September 20-25, pp. 400–403. IEEE, Los Alamitos (2010)
16. Mohan, P., Padmanabhan, V.N., Ramjee, R.: Nericell: rich monitoring of road and traffic conditions using mobile smartphones. In: Proc. ACM SenSys 2008, pp. 323–336. ACM, New York (2008)
17. Riva, O., Nadeem, T., Borcea, C., Iftode, L.: Context-aware migratory services in ad hoc networks. IEEE Trans. Mobile Comput. 6(12), 1313–1328 (2007)
18. Su, C.T., Hsiao, Y.H.: An evaluation of the robustness of mts for imbalanced data. IEEE Trans. Knowl. Data Eng. 19, 1321–1332 (2007)
19. Tanaka, N., Okamoto, H., Naito, M.: Detecting and evaluating intrinsic nonlinearity present in the mutual dependence between two variables. Physica D: Nonlinear Phenomena 147(1-2), 1–11 (2000)
20. Thiagarajan, A., Ravindranath, L.S., LaCurts, K., Toledo, S., Eriksson, J., Madden, S., Balakrishnan, H.: VTrack: Accurate, Energy-Aware Traffic Delay Estimation Using Mobile Phones. In: ACM SenSys 2009, Berkeley, CA (November 2009)
21. Thompson, C., White, J., Dougherty, B., Albright, A., Schmidt, D.C.: Using smartphones to detect car accidents and provide situational awareness to emergency responders. In: Mobile Wireless Middleware, Operating Systems, and Applications. LNICST, vol. 48, pp. 29–42. Springer, Heidelberg (2010)