

RoADS: A road pavement monitoring system for anomaly detection using smart phones

Fatjon Seraj¹, Berend Jan van der Zwaag², Arta Dilo, Tamara Luarasi³, and Paul Havinga¹

¹ Pervasive Systems, University of Twente, Enschede, The Netherlands

² Adaptive Systems, Hengelo (O), The Netherlands

³ European University of Tirana, Tirana, Albania

{f.seraj,a.dilo,p.j.m.havinga}@utwente.nl,
berendjan.vanderzwaag.nl@ieee.org, tamara.luarasi@uet.edu.al

Abstract. Monitoring the road pavement is a challenging task. Authorities spend time and finances to monitor the state and quality of the road pavement. This paper investigate road surface monitoring with smart-phones equipped with GPS and inertial sensors: accelerometer and gyroscope. In this study we describe the conducted experiments with data from the time domain, frequency domain and wavelet transformation, and a method to reduce the effects of speed, slopes and drifts from sensor signals. A new audiovisual data labelling technique is proposed. Our system named RoADS, implements wavelet decomposition analysis for signal processing of inertial sensor signals and Support Vector Machine (SVM) for anomaly detection and classification. Using these methods we are able to build a real time multi class road anomaly detector. We obtained a consistent accuracy of $\approx 90\%$ on detecting severe anomalies regardless of vehicle type and road location. Local road authorities and communities can benefit from this system to evaluate the state of their road network pavement in real time.

1 Introduction

The technology is riding fast and is spreading everywhere, even in the most remote places where people still face basic road transportation difficulties. The roads will become obsolete with the invention of teleportation, but until then people have to ride on them fast and safe. Meanwhile, the world road network is estimated at 35,433,439 km¹, and the number of vehicles is estimated at 35 vehicles per 1000 people². Many studies and surveys are made on the topic of roadway deficiencies and their impact on safety and economy [18]. The road surface can wear and deteriorate in time from factors related to location, traffic, weather, engineering solutions and materials used to build. In developed countries, Pavement Management Systems (PMS) are specialized structures that handle the

¹ CIA World Factbook <https://www.cia.gov/library/publications/the-world-factbook/fields/2085.html>

² The World Bank 2011 data: <http://wdi.worldbank.org/table/3.13>

duty of the road maintenance. Often these structures are equipped with sophisticated and expensive equipment installed on specialized Pavement Evaluation Vehicles¹. For example, in the Netherlands since the introduction of PMS in early 1989, 75% of all of the local authorities, municipalities and provinces, are utilising a PMS [23]. Developing countries often lack this kind of technology and the *know – how*. They sustain their road network through inefficient financial and maintenance planning. One way to assess the road pavement is to measure a vehicle’s vibration with inertial sensors found on smartphones. Therefore, we exploit the pervasive and ‘*smart*’ nature of smartphone devices to collect, process and share these data.

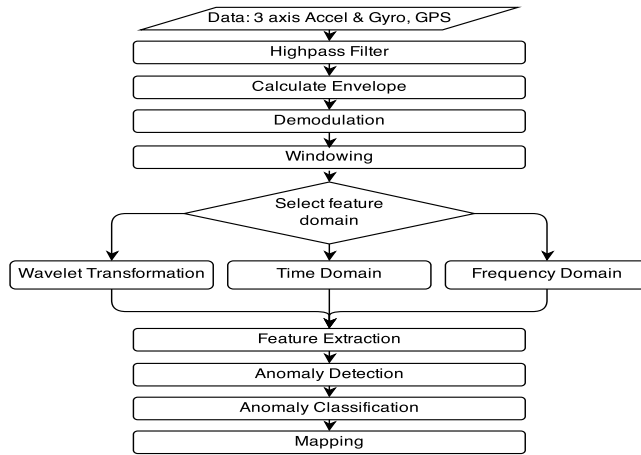


Fig. 1: RoADS flow chart

Other research teams are focused more on pothole detection. Potholes are often results of neglected or bad constructed road segments [17]. We are interested to detect and classify more road surface events. Doing so we can monitor in real time the state and the deterioration of the road segments. Figure 1 shows an overview of our solution, from pre-processing, domain selection and feature extraction to anomaly detection and classification.

In this paper we describe RoADS, a smartphone based **R**oad pavement **A**nomaly **D**etection **S**ystem. Section 2 gives a brief overview of the existing works related to road pavement analysis and anomaly detection. Section 3 outlines our data collection setup, the locations and the methods used to label the collected data. Section 4 explains the preprocessing steps, features extracted from different transformations and the method used to reduce the speed and other dependencies from the sensors signal. Section 5 describes the methods and

¹ Pavement Evaluation Vehicle <https://www.fhwa.dot.gov/research/tfhrc/labs/pavement/index.cfm>

the tools used to classify the data, as well as the obtained results for the labelled and unlabelled roads. Section 6 discusses the conclusion and delineates the future plans.

2 Related work

To determine the road roughness, road engineers measure the profile of the road. A profile is a segment of road pavement, taken along an imaginary line. Usually the longitudinal profiles are subjects of study because they show the design grade, roughness and texture of the profile [22]. Road roughness is defined by American Society of Testing and Materials (ASTM) [1] as: The deviations of a pavement surface from a true planar surface with characteristic dimensions that affect vehicle dynamics, ride quality, dynamic loads, and drainage, for example, longitudinal profile, transverse profile, and cross slope. Equipment and techniques for roughness estimation are usually categorized into [22]:

- Road and Level survey, surveys performed by a survey crew
- Dipstick profiler, a hand-held device commonly used for calibration of complex instruments
- Response type road roughness meters (RTRRMS), transducers that accumulate suspension motions.
- Profiling Devices, sophisticated inertial reference systems with accelerometer and laser sensors to measure the vehicle displacement.

The following works show that inertial sensors alone could be used to detect road surface anomalies. Mainly they were trying to detect potholes, because they are the main concern and also because they are relatively easy to detect based on the energy of the event. Studies and practice suggests that potholes are created as a result of distresses on the road surface [17]. Detecting and classifying more of these distresses we it will be possible to maintain the long term performance of the road pavements. Beside the anomaly detection, all these works faced one major issue, namely the vehicle velocity. The same road anomaly shows different frequencies and amplitudes when driven over with different speeds. Figure 2 shows the signal generated by a manhole when approached with low and high speeds. An important aspect worth mentioning is also the data labelling method used for marking the road anomalies. It is crucial to train a precise detection algorithm with labelling the right anomalous segments.

Pothole Patrol (P^2) [7] uses a high resolution 380Hz accelerometer and a GPS device attached to the car dashboard to collect data and to detect the potholes. Data are transferred to a central server for further processing. Clustering is used to increase detectors precision. Five filters are used, one of them called *z_peak* tries to detect potholes from other high-amplitude road events. Filter *speed vs. z_ratio* is introduced to reject signals with a peak less than a factor t_s times the speed of travel. The labelling technique is based on a trained labeller sitting inside the car and pressing keyboard keys corresponding to predefined anomaly types when they occur.

Nericell [19] uses the Windows Mobile Phone microphone and GPS in con-

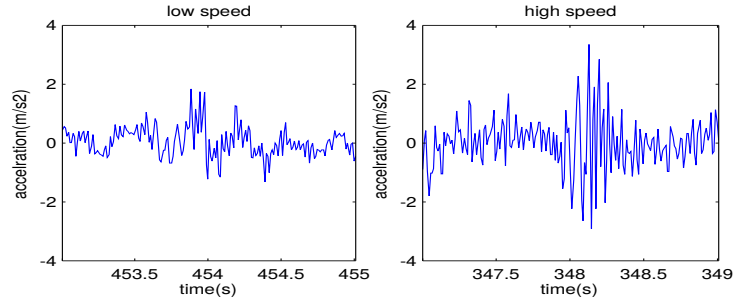


Fig. 2: A 2 second sample of the vertical accelerometer signal when the car drives over a manhole at low speed and at high speed

junction with a high resolution Sparkfun WiTilt accelerometer clocked at 310Hz to monitor traffic and road pavement. The same technique as in P^2 [7] is used to threshold the acceleration signal and to deal with speed. The novelties consist of introducing another filter named z_{sus} for speeds $< 25km/h$, arguing the same anomaly has different shapes for different speeds, and virtual orientation of the phone, using Euler angles to reorient accelerometer data. However, they do not mention the labelling technique they used.

Perttunen et al. [21] use a Nokia N95 mounted on the wind-shield, with accelerometer sampling at 38Hz and GPS to collect the data. Their algorithm classifies the anomalies into two classes: mild and severe. A method of linear regression is introduced to remove the linear dependency of the speed from the feature vector. Labelling is performed with a camcorder attached to the headrest of the front passenger seat, however they realised this method was unreliable to detect the anomalies. A FFT transformation of the signal is performed to extract frequency domain features and to label the data by plotting together the power spectrum and time domain data. Unclear remains the fact how a 38Hz accelerometer sensor can generate 17 frequency bands with 1.4Hz bandwidth.

Tai et al. [24] use a motorcycle riding strictly at two different speeds, 30km/h and 40km/h. An HTC Diamond with accelerometer sampling at 25Hz and a GPS was used to collect the data. Data are preprocessed by the device and sent to a centralised server for classification. Two classification procedures are performed, one to detect the anomalies and the other to rate the road pavement quality from a predefined model of a smooth road. Labelling is performed by the motorcycle rider with a microphone, while riding through an anomaly. An algorithm is used to shift the audio label to correspond with the nearest anomaly event captured by the accelerometer.

3 Data collection

This section describes our data collection setup, the study areas, the type of anomalies, and the technique used to label the data. Throughout this paper we

use the term data referring to the streaming data captured from smartphone inertial sensors.

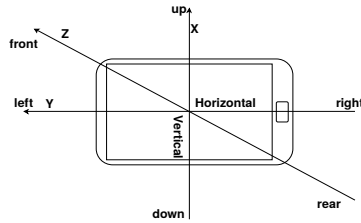


Fig. 3: Smartphone orientation inside the car.

3.1 Data collection setup

Our setup consists of a Samsung Galaxy S2 smartphone running Android 4.0 and an Inertia ProMove 3D¹. The devices were fixed on the windshield of the car with Nokia Universal Holder CR-114. The Android API does not allow to directly access the sensors used in smartphone, however it allows to choose between five predefined delay intervals at which sensor events are sent to the application: UI, Normal, Game, Fastest or user-defined delay [9]. Therefore two modes were used: Game corresponding to $\approx 47\text{Hz}$, and Fastest corresponding to $\approx 93\text{Hz}$ on the Galaxy S2. The Inertia Node was clocked at 200Hz and was used for testing purpose only.

Table 1: Distance between two consecutive accelerometer measurements at different speeds for different systems.

System	Sampling Rate	Speed		
		25km/h	50km/h	75km/h
P^2 [7]	380Hz	1.8cm	3.6cm	5.5cm
Nericell [19]	310Hz	2.2cm	4.5cm	9.0cm
Perttunen [21]	38Hz	18.3cm	36.5cm	54.8cm
Tai [24]	25Hz	27.7cm	55.4cm	83.1cm
Our setup	93Hz	7.0cm	14.0cm	21.0cm

Data collection software registered independently every available sensor on the phone, with corresponding timestamp in nanoseconds since uptime. GPS timestamp is milliseconds since January 1, 1970 [9]. The drive was recorded in a

¹ Inertia ProMove 3D Motion Tracking: <http://inertia-technology.com>

video using the camera and the microphone of the smartphone. All sensor timestamps were synchronised. Table 1 shows how our setup of 47/93Hz compares to other setups in capturing small dimension anomalies.

3.2 Collected data

Data was collected using five different types of cars from different roads in two different cities: in and around the city of Vlora in Albania and in and around the city of Enschede in the Netherlands (see Table 3). The selected roads that were used to collect the data represent different types of anomalies. The Dutch roads are in better shape, are flat without many turns. Common anomalies are manholes, speed humps, patches, cracks, bumps and some small potholes. In contrast, the Albanian roads have a lot of slopes, different types and sizes of potholes, bumps and segments of fully deteriorated or unpaved roads. In total we collected data over a distance of 100.3km on 45.9 unique km of road (see Table 2). Data was collected from accelerometer, gyroscope and GPS sensors

Table 2: Total road coverage in km with different vehicles

Trip	Car	Type	Km	Area	Hz	Location
1	Toyota Corolla	hatchback	22.8	rur/urb	47	Enschede(NL)
2	Peugeot 306	hatchback	22.8	rur/urb	96	Enschede(NL)
3	Skoda Fabia	supermini	16.3	urban	47	Vlora(AL)
4	Toyota Yaris	supermini	33	hway/rur	47	Albania
5	BMW X3	suv	5.46	urban	47	Vlora(AL)
Total			100.3			
Unique			45.9			

of the phone. MEMS gyroscopes are devices that measure the rotation around a specific axis, the *angular velocity*, and are becoming native to smartphones, like accelerometers and compasses. The decision to include the gyroscope in our measurements derives from the fact that rotational velocity describes the behaviour of the car in specific situations. For example, if the car drives through an anomaly that is spanning the road or lane width, such as a speed hump or railroad, the gyroscope measures that rotation around the axis parallel to the anomaly, the pitch, in our case Y_{gyr} . If the car drives over a pothole, the gyroscope measures the anomaly in the axis parallel to the road, the roll, in our case Z_{gyr} .

3.3 Data labelling

We went multiple times over the same road segment, every time with different speed and tried to hit the anomaly in different angles. Labelling was performed by one person using the microphone and the camera of the phone. A detailed

list of all known road surface anomalies was compiled. As the car approached the anomaly the labeller mentioned loudly the type of anomaly. The labeller mentioned everything relevant he saw or felt while sitting beside the driver inside the car. We captured the video footage of the trips in 480x720 pixel resolution. The video footage was used to prove the ground truth. Data collected in Albania was not used for training as they were not voice labelled, and labelling them only by video is inaccurate and time consuming.

Urusoft Subtitle Editor¹ is an application used to transcript the movie subtitles. This application was used to transcript all voice labels with accuracy in the order of milliseconds. The subtitle file is synchronised with the smartphone data timestamp, $t_{acc}^n \leq t_{lab} < t_{acc}^{n+1}$ where t_{acc} is the accelerometer timestamp and t_{lab} is the subtitle time, resulting in an array $[t_{lab_start}, t_{lab_stop}, label]$.

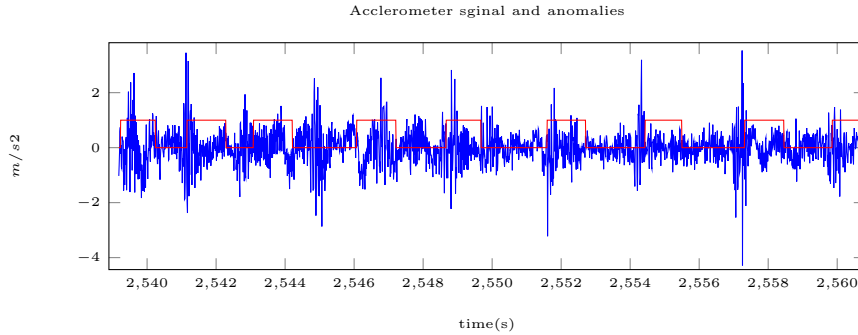


Fig. 4: Signal in blue and binary label (1 for anomaly and 0 for normal road) in red, generated from the voice recordings.

A lag between labelled segments of the data and the actual anomaly was noticed as shown in Figure 4 where Red signal is shifted from or does not correspond with any peak. Also the video footage was shorter and behind, in time about 1s, than accelerometer data length. Because Android is not a real-time operating system (RTOS), data sometimes may drop when the device is busy [16] and some measured data values may have been delayed, resulting in incorrect timestamps. All lags were corrected manually.

The number of individual anomalies was not sufficient for a successful training. We decided to divide the anomalies into event classes. The anomalous data were divided into 3 event classes.

1. Severe, in this category we labelled sunk-in manholes, small potholes and deteriorated and heavily patched road segments.
2. Mild, in this category we placed all those anomalies that happened only in one side of the car such as cracks, one side patches, one side bumps.

¹ <http://www.urusoft.net/products.php?cat=sw>

- Span, in this category were placed road-wide (transversal) bumps, road expansion joins, patches across the road, thick paint, bumps across the road and speed humps.

4 Data preprocessing

The section describes the steps and techniques we used to pre-process the data. A high-pass filter was applied to the signal to remove the low-frequency components, such as turns, acceleration, deceleration, etc. As mentioned in other works, while driving with different speeds over the same anomaly on the road, the signal has different peaks. We call this *speed dependency*. Before segmentation we remove the speed dependency from the signal. It was noticed that when speed signal from GPS and the absolute values of accelerometer x-axis, for a flat road, are plotted as in Figure 7a, the envelope of accelerometer follows the speed signal from GPS. This is not the case if the road is not flat. Figure 5 shows that in the segment between vertical dashed lines the car is driving with high speed through a relatively flat highway, but the peaks are much higher in the segment between vertical black bars, when the car is driving up and down a hill.

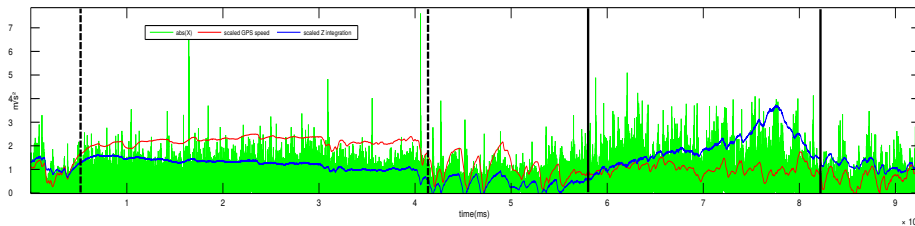


Fig. 5: Readings of linear accelerometer axis $|X|$ (green) with speed from GPS (red) and velocity computed from linear acceleration (blue). Notice the drift accumulated in the segment between two black bars.

4.1 Dealing with speed

To deal with speed dependency we included the speed from GPS into our data logger. Using a device equipped with GPS, one would assume speed and position will always be available, but this is not the case with GPS chips found in smartphones. GPS sensors suffer TTFF (time to first fix) delay. Requirements for a fast TTFF are a clear line of sight and a stable position. To improve the start-up performance of GPS, mobilephones use a technology called A-GPS (Assisted GPS) [14]. A-GPS requires an active data connection with the mobile network operator to receive a preloaded list of available GPS satellites for that location. We did not have an active data connection in our mobile phones at the time we

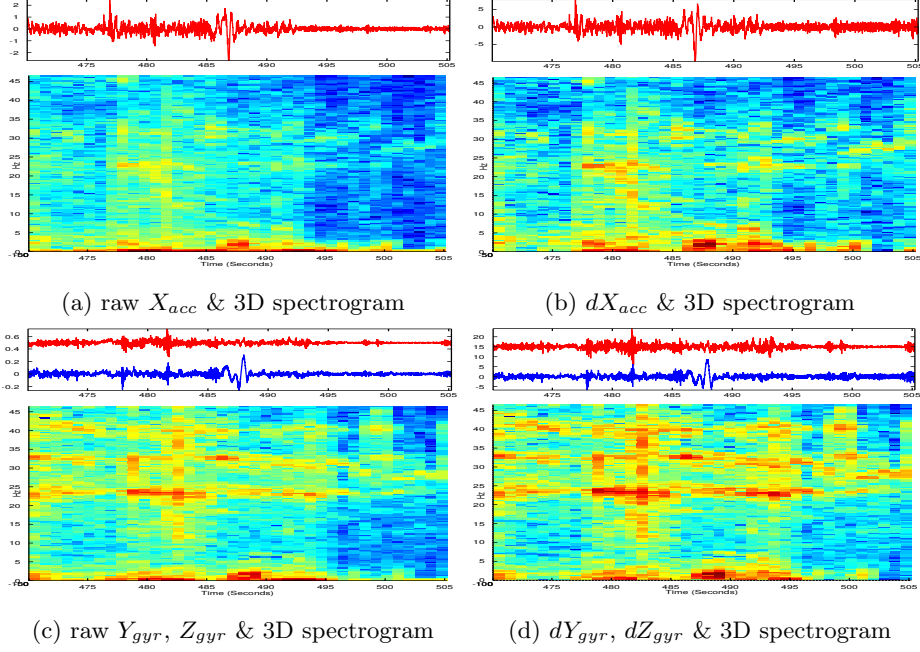


Fig. 6: (a,b) 3D spectrogram of raw and demodulated accelerometer signal with plotted x-axis accelerometer (red) on top; (c,d) 3D spectrogram of raw and demodulated gyroscope signal with plotted y-axis (red) and z-axis gyroscope (blue) on top.

collected the data. On some measurements, data from GPS including the speed were absent. To overcome the absence of speed signal we estimated the velocity by integrating the z-axis of the accelerometer. The results were satisfactory for the flat Dutch roads, but not for the Albanian roads where we were confronted with slopes. The segment between the black bars in Figure 5 shows the car climbing a hill, the integrated velocity is increasing but not the GPS speed. The way the speed, slopes and drift behave on inertial sensors readings has a resemblance with the amplitude modulated signals in radio technology [10]. The accelerometer signal measuring complex mechanical vibrations is modulated with different signals, it acts as a carrier for different signals like speed, slope degree, engine and tire revolutions. Envelope demodulation is often used for empirical mode decomposition (EMD) [8] of complex mechanical vibrations in Hilbert-Huang transform [13]. Demodulated signal dS is computed from the raw signal S based on equation (1),

$$dS(t) = \frac{H \circ S(t)}{E \circ |H| \circ X_{acc}(t)} \quad (1)$$

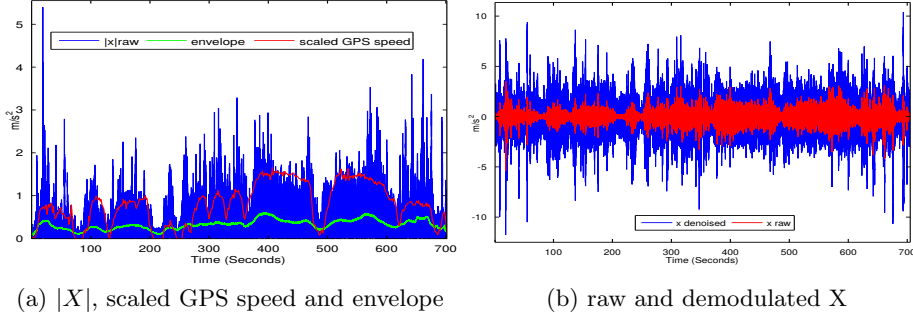


Fig. 7: (a) Absolute value of x-axis accelerometer $|X_{acc}|$ plotted against scaled GPS speed and envelope, (b) Demodulated signal dX_{acc} (blue) plotted against raw signal X_{acc} (red).

where t is time, \circ is the function composition, E is a moving average filter, see Eq. (2), with a large window ($M=2000$ samples rolling window), H is the high-pass filtered applied to the signal and X_{acc} is the x-axis accelerometer.

$$E(t_i) = \frac{1}{M} \sum_{j=0}^{M-1} X(t_{i+j}) \quad (2)$$

The envelope demodulation does not only represent the speed signal but it also compensates for the slope and other low frequency components. From Figure 6 is clear that the demodulated signal shows better the anomalies. Figure 7a shows the signal for the absolute values of x-axis, the speed and the envelope. Figure 7b show how the demodulated signal differs from the raw signal.

4.2 Feature extraction

To prepare data for the classification phase, features representing the data were extracted from accelerometer and gyroscope sensor signals. They are computed from time domain, transformation in frequency domain and wavelet decomposition.

Data was windowed with 256 samples, corresponding to 2.5 seconds and a 170 sample overlap for the data sampled at 93Hz, and 128 samples with 85 samples overlap for the data sampled at 43Hz. First we tried to use the wavelet transformation as a de-noising tool for the signal. Considering the properties of wavelets and their ability to represent the signal in time and frequency, we decided to include wavelet transformation features into our classification algorithm. Discrete Wavelet Transform (DWT) uses multi-resolution filter banks and wavelet filters to analyse and synthesise the original signal [6]. It gives frequency resolution in low frequencies and time resolution in high frequencies [15]. Hesami and McManus [11] showed that DWT analysis outperforms power spectral density (PSD) analysis when used to estimate and analyse road roughness.

Time-domain features extracted: mean, standard deviation, variance, peak to peak, root mean square, zero crossing rate, mean of absolute value, correlation between all axis, tilt angles, wave form length, signal magnitude area.

Frequency-domain features were extracted after FFT transformation with a Hamming window function: mean frequency, median frequency, energy of the frequency bands.

Wavelet decomposition: To decompose the signal we used Stationary Wavelet Transform(SWT) [20] a form of nondecimated DWT. Several experiments were carried out with different wavelet families. Sym5 wavelet from symlet wavelet family, and 4 levels of decomposition gave the best results. Figure 8 shows the plot of the original signal on top and 4 levels of SWT decomposition, the approximations on the left and the details on the right. The high frequencies are visible at the details column d1 and lower frequency signal at a4. Also because we are using SWT we can have a good time resolution for every level of decomposition. The following features were extracted from SWT decomposition: absolute mean, standard deviation, variance, energy for every level of detail and approximation. Four levels of wavelet decomposition were appropriate for our signal.

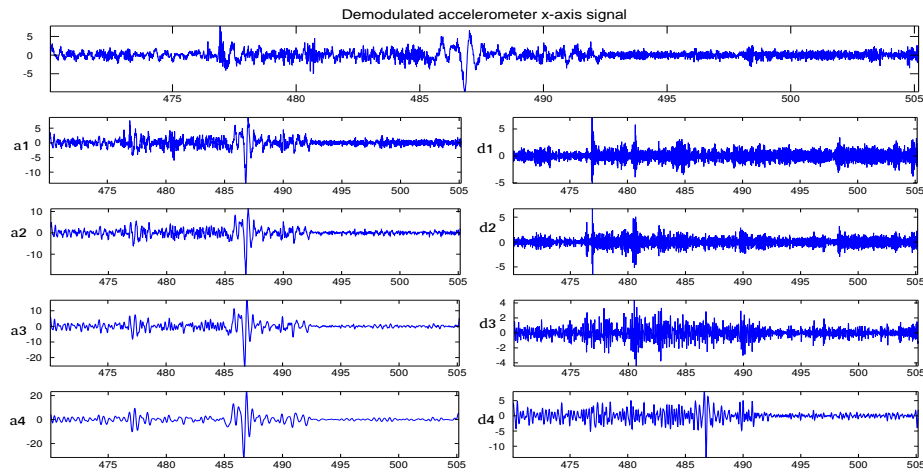


Fig. 8: SWT decomposition at 4 levels with a sym5 wavelet.

5 Classification and Results

For the anomaly detection and classification task we used support vector machines (SVM) [2][5], specifically LIBSVM library [12]. Assuming the number of

anomalous windows is lower than the number of normal windows, we used a 2 step classification. As the first step, all the windows are processed to detect the anomalous windows (those containing road events) from the normal ones. In the second step, the anomalous windows are processed through another classifier to classify the type of anomaly.

5.1 Training SVM

To train the detector we used our labelled data consisting of 3066 windows, 2073 normal windows and 993 anomalous windows. We use a sliding window with an overlap of 66% to frame the data, which means that parts of the anomalous signal would be present in more than one window. Training was performed using 10-fold cross-validation, where 1/10 of both anomalies and normal data were used only for testing purpose. The training data were not stratified. From 993 anomalous windows, we used only those windows with the anomalous signal in the center of the window spreading equally in both sides. The reason behind this decision, is that using a sliding window method, parts of the same anomaly are present in different consecutive overlapping windows and the aim is to train the detector with the best representative of the class. Smartphones are subject of processing power and energy concerns. We try to keep the number of calculated features as small as possible and in the mean time to be flexible with the number of classes. We decided to use a radial basis function (RBF) kernel, as described by Burges [4] and Ben-Hur et al. [3]. A grid search was conducted to find the best values for *hyperparameters* [3]: the kernel parameter $\gamma = 0.002$ and 0.0002 , the cost of misclassification the soft margin constant $C = 320$ and 100 respectively for the detection and classification. Several experiments were performed, training the detector with feature sets from different domains, transformations and combinations thereof, see Table 3. We also experimented with different settings, such as features extracted from the raw signal TD_{raw} , on features from the demodulated signal $TD_{demodulated}$. The best results were achieved by the TD+SWT, the combined feature set from time domain and wavelet transformation. Nevertheless for detection we used only features from SWT.

5.2 Results from training

The results from the training experiments clearly show that the detectors trained with the demodulated features were more accurate on classifying the data than the detectors trained with raw data. Table 3 shows the results of the anomaly detection.

Table 4 shows anomaly classification results of our 10-fold crossvalidation training of the data with features extracted from the wavelet decomposition (SWT) and combined (TD+SWT) features from time domain and wavelet decomposition. The accuracy is the same 91%. However, the severe class is detected better with features from the wavelet transformation only. We also experimented with features from other transformations but the results were lower than 86%.

Table 3: The confusion matrix and accuracy for the classification of anomalous (Positive) and normal (Negative) segments of road, with different feature sets.

Confusion Matrix Anomaly vs Normal											
Method	Accuracy	TP	TN	FP	FN	Spec.	Sens.	G	RS	FPR	FNR
TD _{raw}	80.13%	49	201	39	18	0.82	0.73	0.77	0.89	0.18	0.26
TD _{demodulated}	85.26%	52	214	31	15	0.87	0.82	0.84	0.94	0.13	0.18
FFT _{raw}	77.88%	49	194	51	18	0.79	0.73	0.76	0.92	0.21	0.27
FFT _{demodulated}	79.17%	40	207	38	27	0.84	0.59	0.71	0.70	0.15	0.40
SWT _{raw}	82.69%	50	208	37	17	0.85	0.75	0.80	0.88	0.15	0.25
SWT _{demodulated}	88.14%	57	218	27	10	0.89	0.85	0.87	0.96	0.11	0.15
TD+FFT _{demodulated}	83.01%	48	211	34	19	0.86	0.72	0.79	0.83	0.14	0.28
TD+SWT _{Demodulated}	88.78%	59	218	27	8	0.89	0.88	0.89	0.99	0.11	0.12

*TD= Time Domain, *FD= Frequency Domain, *SWT=Stationary Wavelet Transform
 TP/FP= True/False positive, TN/FN= True/False negative, Spec./Sens.=Specificity/Sensitivity
 G= G-mean, RS= Rlative sensitivity FPR/FNR = False positive/negative rate

Table 4: The confusion matrix and accuracy of the classification for anomalous segments of road with features from SWT and TD+SWT.

Confusion Matrix							
		SWT			TD+SWT		
Class		Sev.	Mild	Span	Sev.	Mild	Span
Severe		17	3	0	16	4	0
Mild		1	15	0	0	16	0
Span		0	0	9	0	0	9
Accuracy		91.1%			91.1%		

5.3 Results on unlabelled data

The evaluation process turned out to be easy using our labelling technique. The system made the predictions and two subtitle files were generated containing the labels of the anomalies detected. The algorithm also generates a KML ¹ file with the location of severe anomalies for Google Maps. We went through the video footage taking notes on reported anomalies from the subtitles. From data collected in Albania some types of severe anomalies were not detected. We believe the reason for that is the fact that our system had not been trained for those anomalies. Table 5 shows the performance of our system with data collected with different cars in different locations. Trip 1 and Trip 2 are made in the Netherlands, in the same road used for training the system. Trip 3, 4 and 5 are trips made in Albania(see Table 2). The trips in the Netherlands have a higher ratio of anomalies per total number of windows than those in Albania. The reason is the driver behaviour in our tests. The video footage showed that the drivers in Albania tried to avoid damage by swerving around or slowing down

¹ Keyhole Markup Language <https://developers.google.com/kml/documentation/kmlreference>

in front of bad road areas, whereas the drivers in the Netherlands did not. The GPS location from mobile phone on the map is not very accurate (see Figure 9). In a wider perspective with more data from different contributors it create the possibility for clustering all measurements by their geo-coordinates.

Table 5: The results of system classifications with models and features from TimeDomain and Wavelet Decomposition

	S/R	#Win	Detection		Classification		Mapping	Evaluation				
			Anom	Sev.	Mild	Span	mapped	Sev.	TD	FD	MA	
Trip 1	TD	47Hz	609	195	107	68	20					
	SWT	47Hz	609	165	119	26	20	113	73	64	15	4
Trip 2	TD	96Hz	764	222	76	105	41					
	SWT	96Hz	764	207	93	72	42	123	61	52	12	3
Trip 3	TD	47Hz	1067	280	119	106	55					
	SWT	47Hz	1067	222	120	51	512	152	82	74	10	8
Trip 4	TD	47Hz	2240	584	349	176	61					
	SWT	47Hz	2240	412	255	98	59	296	173		no video	
Trip 5	TD	47Hz	794	239	113	126	0					
	SWT	47Hz	794	192	99	93	0	134	76	74	6	6

Sev.= Severe, TD= True Detections, FD= False Detections, MA= Missed Anomalies



Fig. 9: Snapshots of map and video anomalies for locations in Vlora, Albania.

To estimate the detector performance for the severe class we counted the false negative windows reported by the detector and also the undetected anomalies from the video footage, see Table 5. The detector detects all windows with anomalous signal, including consecutive windows for the same anomaly. To avoid clutter on the map, if two or more consecutive windows detected the same anomaly, we keep the half of them. This way the number of anomalies shown on the map is fewer than the number of reported windows. As false detections (FD)

were counted all severe anomalies that were detected as mild ones, but that can be subjective based only on video footage. Missed anomalies (MA) is the number of all anomalies, counted by us through the video but not reported from the detector. For Trip 4 associated with Figure 5 the video file was corrupted, but based on GPS map we noticed that the majority of the anomalies were in the segments of deteriorated road, the segment between black lines. For Trip 5, made with the BMW X3 SUV, the results were more correct. Worth mentioning is the fact that the system has a high accuracy on detecting transversal anomalies, spanning the width of the road. On extremely deteriorated road segments, the system detected windows that belong to the transversal anomaly class, see Table 5. In comparison, P^2 [7] and Nericell [19] skipped the windows representing transversal anomalies.

6 Conclusions

In this paper we proposed a system that detects road surface anomalies using mobile phones equipped with inertial sensors: accelerometers and gyroscopes. We applied the stationary wavelet transform analysis and a method to remove effects of speed, slopes and drifts from sensor signals using the envelope technique. Our audiovisual labelling technique was precise and also helpful for the system evaluation. Classifying road anomalies is a rather difficult process and the expectancies to detect all road anomalies on one pass are quite low. Nevertheless, the obtained results showed a consistent accuracy of $\approx 90\%$ on detecting severe anomalies regardless of vehicle type and road location. To increase the accuracy and the number of anomaly classes we will collect more labelled data and improve the training of the SVM through stratification and testing with other feature sets and sensors, such as barometric pressure sensor. For future work, we aim to apply these methods of road anomalies detection in participatory sensing using clustering by the geo-coordinates. We also intend to implement a vehicular network to share that information with other vehicles and to perform road serviceability performance with outputs conform the International Roughness Index and the ISO 2631 standard.

References

1. ASTM Standard E867, Standard Terminology Relating to Vehicle-Pavement Systems (June 2012)
2. Abu-Mostafa, Y.S., Magdon-Ismail, M., Lin, H.T.: Learning From Data. AMLBook (2012)
3. Ben-Hur, A., Weston, J.: A users guide to support vector machines. In: Carugo, O., Eisenhaber, F. (eds.) Data Mining Techniques for the Life Sciences, Methods in Molecular Biology, vol. 609, pp. 223–239. Humana Press (2010)
4. Burges, C.J.: A tutorial on support vector machines for pattern recognition. Data mining and knowledge discovery 2(2), 121–167 (1998)
5. Cortes, C., Vapnik, V.: Support-vector networks. Machine learning 20(3), 273–297 (1995)

6. Daubechies, I.: The wavelet transform, time-frequency localization and signal analysis. *IEEE Transactions on Information Theory* 36(5), 961–1005 (Sep 1990)
7. 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: *Proceedings of the 6th International Conference on Mobile Systems, Applications, and Services*. pp. 29–39. *MobiSys '08*, ACM, New York, NY, USA (2008)
8. Feldman, M.: *Signal Demodulation*. John Wiley & Sons, Ltd (2011)
9. Google: Android developer sensor and location classes, <http://developer.android.com/reference/android/hardware/Sensor.html>
10. Gottlieb, I.: *Understanding amplitude modulation*. Foulsham-Sams techn. books, H. W. Sams (1966)
11. Hesami, R., McManus, K.J.: Signal processing approach to road roughness analysis and measurement. In: *TENCON 2009-2009 IEEE Region 10 Conference*. pp. 1–6. *IEEE* (2009)
12. Hsu, C.W., Chang, C.C., Lin, C.J.: *A Practical Guide to Support Vector Classification*. Department of Computer Science, Taipei 106, Taiwan (2003)
13. Huang, N., Attoh-Okine, N.: *The Hilbert-Huang Transform in Engineering*. Taylor & Francis (2005)
14. LaMance, J., DeSalas, J., Jarvinen, J.: Innovation: Assisted GPS: a low-infrastructure approach. *GPSWorld* (March), 46–51 (2002)
15. Mallat, S.: *A wavelet tour of signal processing: the sparse way*. Academic press (2008)
16. Milette, G., Stroud, A.: *Professional Android Sensor Programming*. Wrox (May 2012)
17. Miller, J.S., Bellinger, W.Y.: *Distress identification manual for the long-term pavement performance program (fourth revised edition)*. Tech. Rep. FHWA-RD-03-031, Federal Highway Administration (June 2003)
18. Miller, T., Zaloshnja, E.: *On a crash course: The dangers and health costs of deficient roadways* (2009)
19. Mohan, P., Padmanabhan, V.N., Ramjee, R.: Nericell: Rich monitoring of road and traffic conditions using mobile smartphones. In: *Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems*. pp. 323–336. *SenSys '08*, ACM, New York, NY, USA (2008)
20. Nason, G.P., Silverman, B.W.: The stationary wavelet transform and some statistical applications. In: *Wavelets and statistics*, pp. 281–299. Springer (1995)
21. Perttunen, M., Mazhelis, O., Cong, F., Kauppila, M., Leppänen, T., Kantola, J., Collin, J., Pirttikangas, S., Haverinen, J., Ristaniemi, T.: Distributed road surface condition monitoring using mobile phones. *Ubiquitous Intelligence and Computing* pp. 64–78 (2011)
22. Sayers, M., Karamihas, S.: *The Little Book of Profiling: Basic Information about Measuring and Interpreting Road Profiles*. University of Michigan. Transportation Research Institute, UMTRI (1996)
23. Schut, P., de Bree, T., Fuchs, G.: Responsible pavement management. In: *First European Pavement Management System: Conference-Procidings and final program* (2000)
24. Tai, Y.c., Chan, C.w., Hsu, J.Y.j.: Automatic road anomaly detection using smart mobile device. In: *Proceedings of the 2010 Conference on Technologies and Applications of Artificial Intelligence (TAAI 2010)*(18-20 Nov. 2010, Hsinchu, Taiwan). pp. 1–8 (2010)