

# A Support Vector method for Automated Road Anomaly Detection using Mobile Device

Md. Saiful Islam  
Department of Computer  
Science and Engineering,  
Shahjalal University of Science  
and Technology, Sylhet.

Shantanu Mandal  
Department of Computer  
Science and Engineering,  
Shahjalal University of Science  
and Technology, Sylhet.

Sazedul Islam  
Department of Computer  
Science and Engineering,  
Shahjalal University of Science  
and Technology, Sylhet.

## ABSTRACT

Surveying highway and road for anomaly detection is a challenging task for the road transport authority in Bangladesh. Lack of sufficient budget and manpower makes this problem really hard. Moreover as a populated and developing country many more roads are being constructed each year. Routine maintenance cannot be conducted due to complex and costly road anomaly survey system. Here we propose a system that can be used to automate the road surveying system for anomaly detection. Road transport authority can easily be benefited from this automated system. Here we use GPS and accelerometer, an inertial sensor that can detect vibration while going on a car as our data source and use signal processing and Support Vector Machine to detect deteriorated road segments.

## Keywords

Road survey, Road anomaly, SVM, Pattern recognition.

## 1. INTRODUCTION

Now-a-days the enhancement of technology is increasing and spreading everywhere, even in the most remote places where people still face basic road transportation difficulties. The roads are becoming congested with the invention of various transports, but people have to ride on them to keep them fast and safe. The road surface can be damaged and deteriorated by many factors related to location, traffic, weather etc. Keeping our roadways in safe condition is a challenging problem.

In developed countries, automated highway road survey system is used to handle the duty of the road maintenance. Often these structures are equipped with sophisticated and expensive equipment installed on vehicle. Developing countries often lack this kind of technology. They sustain their road network through inefficient financial and maintenance planning. Because municipal budgets are generally constrained. So it is difficult to find out which road needs to be repaired first[1]. Moreover routine maintenance could make the road sustainable for more time.

For these purposes in a developing country a system is really necessary which is cost effective, needs less man power and much effective as other system. Here we have proposed an automated highway road survey system which is able to detect the condition of roadways by gathering data from vibration and GPS sensors.

## 2. EXISTING SOLUTION FOR THE PROBLEM

In this section we describe some techniques that are being used for highway road survey. We will also see the advantages and disadvantages of each of these techniques.

### 2.1 Manual Solution

To measure the road condition road engineers make a profile of the road. Profile is the data that shows what the road condition is. They do profiling with following procedure.

- First decide how much road need to be profiling.
- Collect sample data from different segment of the road.
- Do statistical procedure to measure what is the condition of the road.

American Society of Testing and Material (ASTM) is an international standards organization that fixed the standard. If collected data fails to meet the standard then the engineer consider this road as bad.

But this procedure does not show the actual condition of the highway. This survey measure only a small portion of the condition of the road.

### 2.2 Camera Based Systems

Condition of road can be detected with image[2]. A camera is used to capture image and an image processor identify the anomaly of the road. Images are digitized and series of image processing algorithm are applied on them. Information about different anomalies of the road, obstacle, speed breaker, bridge joint and manhole joint can be extracted by using various image processing techniques. Although camera, laser, radar[3] based system is good but it has its own disadvantage. Their performance depends on many factors.

- Camera cannot capture image in night or in bad weather condition.
- As it needs to monitor the highway, one cannot capture image by walking. Highway is hundreds of kilometers long. So it needs vehicle to capture image. So now there are two options. One is, setup camera in public transport and another is setup in private vehicle. As expensive camera is needed, setup camera in public transport is unsecured. But using a private vehicle will increase the cost.
- Let assume the vehicle will run at 40 kilometer per hour. So to survey 100 kilometer of road it needs 2.5 hours. Moreover as the vehicle is running this speed, to capture proper image camera needs to take

many photo per second. So to ensure all these things no ordinary camera can do the task. For maintaining the quality a good expensive camera is needed.

### 3. SYSTEM WE ARE USING

We have already mentioned some of the existing systems that are currently using in various countries. Implementing these systems is somewhat expensive for a developing country like Bangladesh. So we are trying to use a system which is cost effective and accuracy is reasonable.

Traveling through a broken road segment in a vehicle causes a certain level of vibration or shaking according to how much the road is broken in the segment. So if we can somehow record the vibrating rate when traveling in a vehicle then we can determine two things. These are

- There is certainly a road as a vehicle is moving through.
- There may or may not be broken segments in that road segment using our system.

We are using an accelerometer to record the vibration rate when traveling in a vehicle and a GPS to coordinate the place or area of the place of the road. Currently we are using smart phone for our purpose. Because a smart phone device provides both an accelerometer and a GPS sensor in a single package. But in future we are aiming to use a custom made device with a GPS and an accelerometer for our purpose as the custom made device is more accurate and more cost effective. We divided our system in two major parts

- The client side device which will be in the vehicle
- The server side device which will be in the main server computer.

### 4. DATA COLLECTION

Collecting training data is one of the first challenging task for our system to work on testing environment. We have mentioned earlier in the report that we are aiming to use a custom made device to collect the data for the purpose. But as currently we are in short of such a device we are using smart phones.

Smart phones and tablets are consist of a built in 3D accelerometer and GPS sensors. For collecting training data we are currently using smart phones as mentioned earlier. Using these devices has some potential drawbacks. These device do not give the recommended rates of data the system needed to be accurate and putting such an expensive and capable of wide variety of usage devices in the vehicle for such a simple objective is kind of a waste of potential recourses whereas this work can be done by a simple device with a GPS and an accelerometer.

We collected our training data using five smart mobile devices simultaneously. We use these device by placing them in different places in the vehicle. We selected several routes to collect our sample data set. We have used several devices to test weather different position could affect the result. The software we are using to read the accelerator and GPS data is a custom made application.

We collected the sample training data from various road segments around Sylhet city. Observing these route we can however label these route broken, plane or partially broken, etc. We have collected same route data several times. We separately collected data for going and returning. We traveled in a CNG auto-rickshaw. Devices are put in the back seat of the vehicle. We have used following routes.

- SUST Campus to Madina Market: This route has two sides and is considered as broken. The returning route is relatively severe. But in some places one side is plane but the other is broken and vice versa.
- Madina Market to Amberkhana: This route is considered broken in some places. It is approximately 3 kilometers long.
- Madina Market to Medical Road: This route is relatively well and considered as plane. This is newly constructed road. So there is no broken road segments here. We use this road to collect our smooth road training data segments.
- Madina Market to Bandar Bazar: This route is a long route. It is approximately 5 kilometers long. Most of the road is smooth. But in some places there are significant amounts of broken road segments.

### 5. DATA PREPOSSESSING

Data that are being captured with accelerometer has significant amount of noise. This noise need to be filtered. To filter low frequency we use a high pass filter. Moreover, while driving with different speed on same broken roads give different peaks in signal. Different speeds give different peaks. So before segmentation into windows we need to normalize the signal. The accelerometer signal is modulated with different signals like speed, slop degree, engine and tire revolution. Envelop demodulation is often used for empirical mode decomposition (EMD of complex mechanical vibration in Hilbert-Huang transform[4,9]. Demodulated signal  $dS$  is computed from the raw signal  $S$  based on following equation.

$$dS(t) = \frac{H_0 S(t)}{E_o |H_0| X_{acc}(t)}$$

### 6. FEATURE EXTRACTION

Data is prepared for classification phase which are extracted from accelerometer sensor signals. It is computed from time domain and transformation in frequency domain. Data is windowed with 250 sample per window. To differentiate the data we use clustering system. As our system use semi supervised learning we use following steps to label the data.

- Make window with the signal.
- Use clustering to differentiate the window.
- Label different cluster data with loosely labeled data.

### 7. CLASSIFICATION AND RESULT

For the anomaly detection we have used SVM (Support Vector Machine). SVM[5] is a machine learning approach that is used for pattern recognition in supervised learning. We use SVM for our anomaly road patter recognition. We train the SVM with our labeled window data[6]. As we use window system it may occur that same anomaly is presented in different window. This can increase the anomaly portion. We prepare our train and test data separately. To avoid over fitting we use cross validation[7] to find out what model would be better for our system. We use 10-fold cross validation.

When evaluating SVM estimator, we need to set C settings manually that can over fit the data. To avoid these type of scenario total data set is dividing into several parts, where some part is used for training the system and some part is for testing. The following procedure describe how k-folds work

- A model is trained using K-1 of the folds as training data
- The resulting model is validated on the remaining part of the data

We use RBF kernel[8] as our SVM kernel. So here a question may arise that why we are using RBF instead of linear or other kernel. In our problem number of feature is not that much high. We have large number of observation set. RBF kernel outperform linear kernel in the preceding points. So according to our problem definition we have use RBF as our kernel in SVM.

Although our result surplus 90% benchmark, we face difficulties in some places to detect broken segments. Sometimes it seems that there is no broken segments but the road has some other material on it that cause the bumps. Collecting data only once from the segments cause the problem. But if we could collect multiple times with some

time gap (1-2 weeks) then it is highly possible that we would not get the incorrect data. Another problem we face while collecting data from the joint of a road and bridge. Although this joint is not a broken segments our system marks it as a bump. The effect is negligible as the amount of these type of data is so insignificant compared to the vast amount of data.

## 8. CONCLUSION

Here we have proposed a system that can be used to detect road anomaly which is very important to survey highway road. To detect this we have used accelerometer data that captures vibration data while go on a car. After applying some signal processing mechanism we normalize the data and remove noise. Then we have labeled the data with semi supervised style using clustering. SVM which is our main model to detect anomaly pattern is trained with labeled data. Obtained result using this system shows above 90% accuracy regardless of vehicle type and road location.

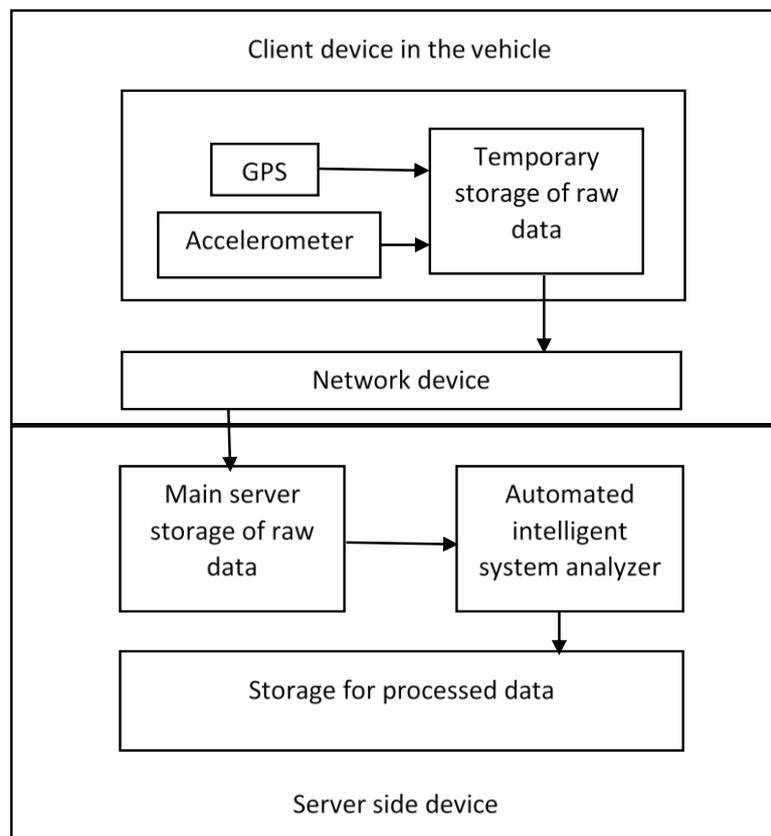


Fig 1: System we are using

Table 1. Sample Data

Timestamp	Latitude	Longitude	X	Y	Z	Tag
1425798449844	24.919950	91.83203	0.765	7.65	4.396	True
1425798450044	24.919950	91.83203	1.53	7.65	4.596	True
1425798450844	24.919906	91.83206	-1.071	10.557	4.496	True
1425798451044	24.919906	91.83206	3.672	8.415	4.396	False
1425798451244	24.919906	91.83206	-1.683	8.262	4.696	False
1425798451445	24.919906	91.83206	0.459	9.027	4.896	False

**Table 2. Camera settings**

Speed	40 kilometer/hour
Travel per second	11.11 meter
Need to take photo per second	12 photos
Need a camera of at least	12 fps
Need a camera of minimum aperture	f/12

**Table 3. Confusion Matrix**

Total Window Frame	180000
True Positive	164214 ( 91.23% )
True Negative	7326 ( 4.07% )
False Positive	6192 ( 3.44% )
False Negative	2268 ( 1.26% )

## 9. REFERENCES

- [1] R. Eaton, R. Joubert, and E. Wright. Pothole primer: A public administrator's guide to understanding and managing the pothole problem.
- [2] Treash, K. and Amaratunga, K. Automatic Road Detection in Grayscale Aerial Images.
- [3] D. R. Huston, N. V. Pelczarski, B. Esser, and K. R. Maser. Damage detection in roadways with ground penetrating radar.
- [4] Norden Eh Huang, Samuel S. Shen. Hilbert-Huang Transform and Its Applications
- [5] Burges, C.J.: A tutorial on support vector machines for pattern recognition. Data mining and knowledge discovery
- [6] Chih-Wei Hsu, Chih-Chung Chang, and Chih-Jen Lin. A Practical Guide to Support Vector Classification
- [7] Ron Kohavi. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection
- [8] S. Sathiya Keerthi, Chih-Jen Lin. Asymptotic Behaviors of Support Vector Machines with Gaussian Kernel
- [9] Huang N., Attah-Okine N.: The Hilbert-Huang Transform in Engineering.