Virtual road concept as a tool for road quality research

Marcin Badurowicz[1,*], and *Jerzy* Montusiewicz[1]

[1]Lublin University of Technology, Faculty of Electrical Engineering and Computer Science, Institute of Computer Science, Nadbystrzycka 36B, 20-618 Lublin, Poland

Abstract. Road quality assessment using crowdsourced data gathered by smartphone users, based on acceleration data, is an interesting subject on using modern technology for improvements of the infrastructure. The algorithms – for both road quality assessment and detection of different elements on the road – need to be tested, especially in the field. To facilitate building sets of different data and sharing them in a standardised way, the authors propose extraction of known road fragments with known types of surface degradation and construction of virtual streams of data, thus "virtual roads". The procedure for data extraction and building a database of segments, combining them into virtual road, as well as testing real-world algorithm using the constructed virtual road are presented in the paper.

1 Introduction

Road quality research, using crowdsourced data acquisition based on the smartphone accelerometer is an intriguing issue, especially in the rural areas, where professional grade assessment equipment cannot be used from the economic point of view. Crowdsourced methods are hugely discussed in the literature, and in previous works of the authors.

The huge and still rising set of data acquired from the crowdsourcing is a challenge in the fact of proper testing different algorithms, especially by different research groups where there is no such a possibility to drive repeatedly over a finite set of real roads to test different possibilities and situations. On the other hand, one research group may drive multiple hours in search for a specific road artefact. Multiple research in the terrain is time-consuming and not efficient. There is also a need for recording and storage of road artefact data in case of creation of new detection algorithms and techniques.

To cope with that, the authors are willing to propose the "virtual road concept". The virtual road is a digital-only set of data, carefully extracted and processed from the real-world data, to be used in building and testing different kinds of detection algorithms for potholes, and other kinds of road quality factors; as well as overall road quality assessment methods, all based on accelerometer data.

The goal of this paper is to present the concept of virtual road, its creation procedure as well as data acquisition which was needed before the actual construction, and performing the virtual road assessment using indicators already published by the authors.

1.1 Similar research

There is a set of established works on the topic of measuring road quality using crowdsourced mobile devices, especially smartphones and similar, as they are equipped with necessary environmental sensors, e.g. accelerometers, starting from the 2008 systems Nericell and TrafficSense [1], through the usage of different sensors, like microphones [2]. There is a research in detection mechanisms for finding different constructional and degenerative changes on the road surface. There are also works covering road quality calculation from accelerometer data [3, 4], especially in the crowdsourcing manner [5, 6] Finally, there are also discussed proposals of creation of road artefacts database for further reuse [7]. The authors were discussing usage of streaming analytics in the cloud computer pattern for such a research, as well as described road artefact detection algorithms [8, 9].

Taking into account results of all the previously performed experiments, which used different cars from different manufacturers and in a different overall state, there is always a direct connection between data recorded by the accelerometer and road profile [3]. The lacking factor now is no possibility of testing new road artefacts detection methods on the set of roads not existing in the real world, the need of maintaining a huge internal database of accelerometer readings and overall problems with sharing these data between scientists.

2 The virtual road concept

When driving on a road, drivers may encounter different types of road surface changes, impacting negatively their overall experience. Potholes, speed bumps, bad road

---

*Corresponding author: m.badurowicz@pollub.pl

segments, manhole covers and similar were dubbed as "road artefacts" by the authors in the previous research [8] and such term will be used in this paper.

The whole concept is based on the idea of the "building blocks", virtual road components, a digital representation of a signal when driving over the specific kind of the road artefact. These blocks may be set up into time-series, allowing to reduce the amount of non-significant data and simulating different kinds of roads; with different types of artefacts of different quantity. The building blocks database will be used to create the virtual road and will be published later for the other researchers to base similar research on.

To create the virtual road, there is a need to perform a series of operations on data acquired during regular experiments, until they are ready for the actual virtual road construction procedure.
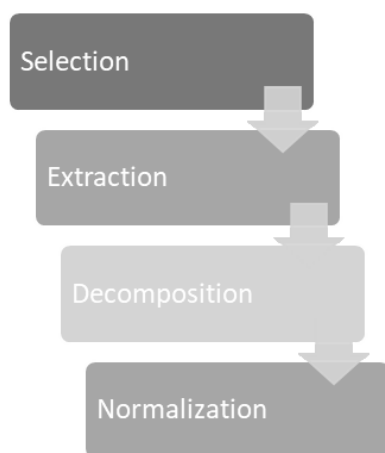

Fig. 1. Steps of building a virtual road components database.

Based on the procedure presented in the Figure 1, the authors are willing to describe it step-by-step, starting with the selection and classification.

2.1 Road artefact classification

First, there is a need of **selection** between types of road artefacts. To create a set of "building blocks", creation of simple classification of the road artefacts, based on their types is indispensable.

The authors have divided road artefacts to the 3 main classes: wide, narrow and point. Examples of the common road artefacts and their respective classes are presented in the Table 1.

The base of the classification is the length of the road artefact in comparison to the length of the car, or the car's wheelbase. The "Wide" class corresponds with the road artefacts longer drastically than the wheelbase – wide road artefacts are commonly relative to the overall surface type – for example paved or rubble road surface.

The "Narrow" class is a set of road artefacts which are about the size of the wheelbase – this includes railroad crossings, tramway crossings, some speed bumps or elevated pedestrian crossings. The "Point" type is smaller than wheelbase, typically the size of the only one wheel – many potholes, manhole covers (raised or lowered), smaller speed bumps.

The generic road artefact types, presented in the Table 1, are the base for building blocks for the virtual road. These 8 types are used and may be used to construct the virtual road with. There is also one more building block type – a baseline asphalt road, being a representation of signal for a regular, good quality road.

Table 1. Road artefact types and classes.

Road artefact	Class
Paved road	Wide
Rubble road	Wide
Railroad crossing	Narrow
Elevated pedestrian crossing	Narrow
Wide speed bump	Narrow
Pothole	Point
Manhole cover	Point
Speed bump	Point

2.2 Data capture and extraction

As mentioned in the introduction, the acquisition device for the whole concept was the smartphone device, mounted in the car in a stable way, so all vibrations from the road profile could be directly represented in the accelerometer data.

The acquisition device, smartphone, reads accelerometer data in 3 dimensions (X, Y and Z) with a frequency of 10 Hz, as well as current time (synchronised with the NTP protocol), speed and location (from the internal Global Positioning System receiver).

As the previous research in this topic suggests, there is a need for calibration to remove tiny accelerometer oscillations [8]. The coordinate-agnostic Global Coordinate System should be also used [10], so the device does not need to be orientated in any specific way. Global Coordinate System is a reorientation procedure, allowing the acquisition device to record acceleration values in axis N, E and Z2, which are the magnetic north, the magnetic east, and vertical (perpendicular to the Earth's surface), respectively.

To gather data about acceleration values from the different types of coordinates, the study was performed, using two cars, driving over the same route, covering multiple roads of different types, in the real-world environment.

One of the cars was an older Peugeot 1007 from 2006, a B-class car, while the second one was a 2014 Toyota Auris, C-class limousine of a higher standard. The differentiation was to achieve different results in overall road readings. The same acquisition device, procedure and methodology were used as in the previous research by the authors: two devices Lumia 820 with the same software were lying in the central tunnel of the car, with a calibration procedure started before entering the traffic [8].

The video camera mounted in cars was used for additional data acquisition step as well as verification. As software the authors were using allowed only to mark geographical position of the road artefact by the user, audio commentary of the artefacts was recorded by the

video camera, along with the overall view of the road. HD action camera (GoClever Action Silver) was mounted with a view on the road in the Peugeot 1007, with a microphone recording inside of the car, an example of the video camera frame is in the Figure 2.

That means that after road artefact encounter, the passenger in each car was tapping the device to mark the position and commented loudly type of the artefact using simple descriptions: "pothole, left", "bump" and similar.

Fig. 2. Example of the video camera data feed frame.

The video/audio tracks and the acceleration data streams from two devices were synchronised using the NTP protocol by time, allowing the authors to perform the data extraction by searching in the audio track names of desired artefacts, searching for the same position in the acceleration track, and final **extraction** of the known road artefact data. This step was performed manually using software developed by the authors and not automated. Software is presented in the Figure 3. The visual recording of the road was also checked with audio commentary to remove false identification from inside the car.
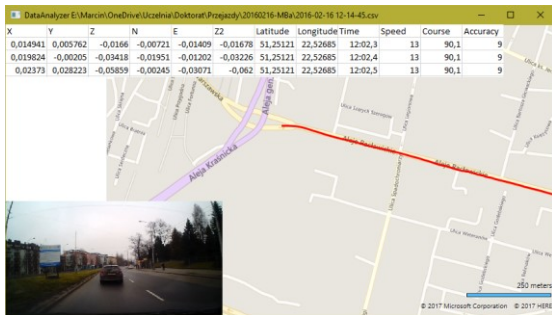

Fig. 3. Application developed for data extraction procedure.

Because driving over the road artefact is not instant, especially in the class of "wide" artefacts, every road artefact segment, and thus building block of the virtual road, must be composed of multiple data points.

Based on the previous research in this topic, the representative time was chosen as 50 data points, (5 seconds).

In case of road artefacts which are recognisable from the vertical acceleration stream (every "narrow" and "point" class), the procedure of extracting the signal was to use 25 data points (2.5 seconds) before the highest peak of vertical acceleration as well as 25 data points after this peak. This 50-data point wide data segment was a base for the proposed digital representation of the road artefact type.

2.3 Decomposition of the individual streams

Signals from driving over the road artefact were the series of tuples – or multi-dimensional vectors in other words, so they needed to be decomposed into multiple parallel streams. Some data which is acquired along with acceleration data needed to be discarded, as was not useable in the virtual road concept. Data streams acquired when driving:
- X, Y, Z axis acceleration;
- N, E, Z2 axis (reoriented) acceleration;
- Current speed;
- Current time;
- Current GPS coordinates.

In this set, both Time and GPS coordinates were not used in the virtual road concept, as they are linked to the real-world fragment.

For every road artefact data extracted in the previous step, these two were discarded. While removal of timestamp did not need additional description, based on the previous research, GPS readings accuracy was only sufficient only to mark the position of the road, but not to mark its beginning or ending position. This data is also needed only in the real-world context, so it was discarded completely.

2.4 Data normalisation procedure

Of course, different artefacts, even in the same class, may be very different one from another in the signal received from driving over them, so the "building blocks" representing the real-world road artefacts are going to be different one from another and there will be multiple recordings for the same road artefact type.

But this problem seems to be more complex as data obtained from two different cars are going to be vary significantly.

To cope with former, the authors are willing to normalise data, based on the known factor, ANF. Accelerometer Normalisation Factor is being calculated when the test car is driven over the known road segment with the same speed as the baseline car. The road segment is carefully selected as one of the A-class road segments, for example newly built highways [11]. First, every sample of dataset (X) to be normalised (x) are scaled into values from the set between minimum and maximum values for the baseline (Y), resulting in normalised x (x_n). This is presented in the equation 1:

$$x_n = x - \min(X) \cdot \frac{\max(Y) - \min(Y)}{\max(X) - \min(X)} + \min(Y) \quad (1)$$

Then, the final ANF should be calculated as an average of every normalised x_n divided by every baseline value (y_i), as presented in the equation 2.

$$ANF = \frac{\sum \frac{x_{n_i}}{y_i}}{n} \quad (2)$$

This procedure is applied to every element of an input vector, thus the final ANF is a vector consisting of

normalisation values for reoriented acceleration values, and is also a transformation of acceleration values from the regular units (g) into dimensionless values. Similarly, the authors are not using real time as the base for the calculations, but data points number, which is also dimensionless.

The normalised values for every known road artefact are grouped into classes of the road artefacts, building the database of different readings for the same road artefact class. The Figure 4, below, presents the example of normalised readings of an example road artefact, the speed bump.
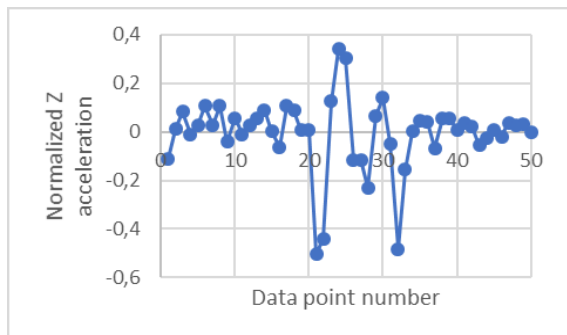

Fig. 4. Normalised readings when driving over a speed bump.

On the other hand, in the Figures 5 and 6 there are signals for road artefacts from different classes: pothole and rubble road, respectively.
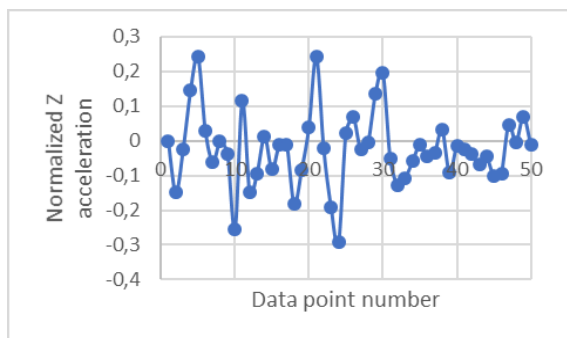

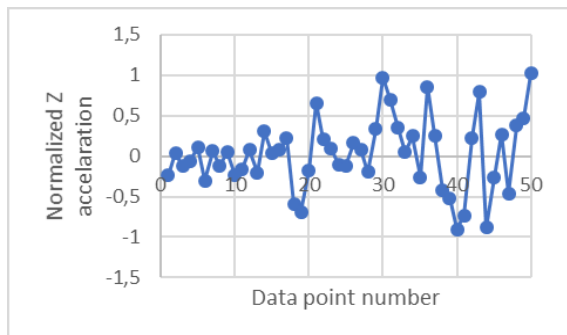**Fig. 5.** Vertical acceleration readings (normalised) when driving over a pothole.

Fig. 6. Vertical acceleration (normalised) on a fragment of a rubble road.

The first problem mentioned – the variety of road artefacts in the same class – is not the problem itself in the case of this research. The authors are willing to include multiple road artefacts of the same class to achieve more diversity in the constructed virtual road, similar to the real world.

3 Virtual road construction

Providing the set of the virtual road artefacts classes, and the "building blocks" of the virtual road gives a possibility to prepare a virtual road, consisting of baseline asphalt fragments and including different types of road artefacts, to check different scenarios and test algorithms. The procedure, presented in the Figure 7, is based on three steps.
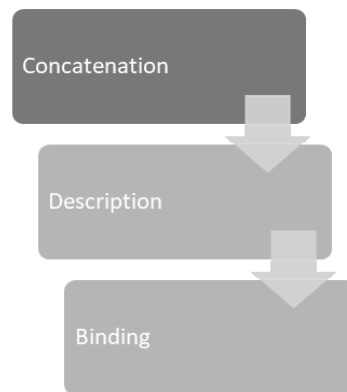

Fig. 7. Procedure for the actual virtual road construction.

The "building blocks" are simply needed to be settled one after another – to **concatenate** their signals. Because data streams for each road artefact type are not tuples anymore, as have been decomposed, there is also a possibility that virtual road concept will be a set of less than all original data streams, allowing testing of different techniques.

3.1 Concatenation of the building blocks

For the testing of the algorithms the authors already have worked with, the building blocks were used of only $Z2$ (global)-axis acceleration. The building blocks for the virtual road were chosen pseudo-randomly in the manner of:
- between 1 and 10 good road segments,
- between 10 and 20 rubble road segments,
- between every 5 road segments there may be a speed bump.

When selection of the given road artefact class (good road, rubble road, speed bump), the actual signal representing the artefact (variant) was selected randomly from the set of normalised signals for this road artefact class.

The built for the purpose of generation of virtual road application was however able to not only generate a virtual road data file for further research, but also had a possibility to save a virtual road and load it from a simple text file, with a strict format.

3.2 Text description of the virtual road

To achieve greater readability and reusability, the authors proposed a method for **description** of the virtual road timeseries in a text format.

The format is based on the number of segments of the given type, then a capital letter describing the road artefact and optionally a lowercase letter describing the variant (the specific real-world equivalent) of the road artefact. The letters are corresponding to the road artefacts classes, as presented in the Table 2. Such format was pseudo-randomly generated, within rules described in the previous section.

Table 2. Road artefact types and corresponding letter code.

Road artefact	Code
"Good" road	A
Paved road	B
Rubble road	C
Railroad crossing	D
Elevated pedestrian crossing	E
Wide speed bump	F
Pothole	G
Manhole cover	H
Speed bump	I

Using the proposed system, for example 3Aa 1Ia 2Ab 3Ad 4Aa will be the equivalent of the virtual road consisting of 3 "good" road segments (variant a), then 1 speed bump (variant a), and then again 9 segments of the "good" road segments – but first 2 variants b, 3 variants d and 4 times variant a.

```
test1.virtualroad  ●

1    Af Ab 2Ag Ac Ia
2    Ag Bc Bd Ba Bc Bb
3    Bc Be 2Bd Ic Be
4    Ba Bd Ia
5
```

Fig. 8. Example description of the virtual road to be constructed.

Such text file, which another example is presented in the Figure 8 above, was read by the application and the actual values were placed in place by selection of the proper road artefacts from the internal database.

3.3 Virtual GPS coordinates binding

Sometimes, there is a need for using data that was discarded in the decomposition stage – in this case being this a GPS data stream, which the authors are willing to use, because requirements of road assessment method used to validate the virtual road concept.

To cope with that, the authors are changing the signal into a stream of tuples consisting a vertical acceleration and corresponding fake GPS coordinates. These coordinates will be used only for the testing procedure, which is based on it, and for easily visualisation.

In general, there may be not only need for GPS coordinates, but also for other values (i.e. speed), which is used in some road artefact detection methods.

4 Experimental results

To confront validity of the constructed virtual road against its real world counterpart, the authors used the same procedure for assessing the virtual road quality as for the real roads – the Road Relative Unevenness Index (RRUI) was used [11].

The virtual road was generated, which text representation has been presented in the previous example in the Figure 8. Such data was than bound with virtual GPS coordinates. In the presented solution, the authors decided to include fake GPS coordinates changing by the 20 meters per second, which was an equivalent of speed in urban areas. To calculate from the meters into GPS degrees, the WGS84 parameters for the Earth were used [12], resulting in usage of a value 111.32 km per decimal degree.

To achieve best results in presenting an output in the paper, only longitude was changed, meaning the virtual road goes straight from the south to the north.
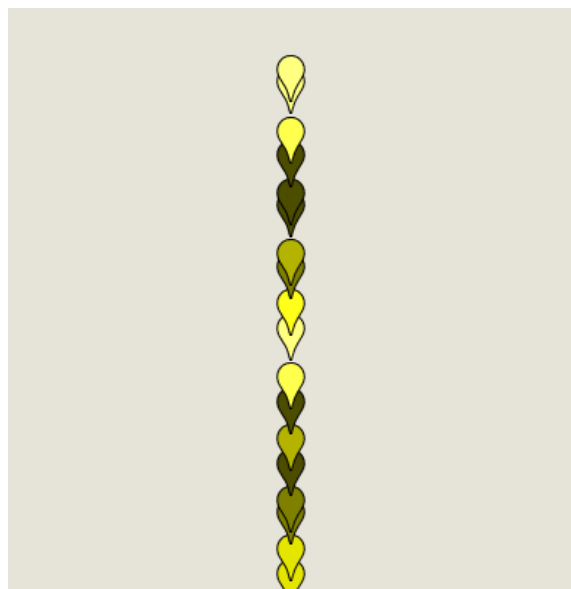

Fig. 9. RRUI pushpins of the virtual road.

As presented in the Figure 9, the RRUI assessment scale has been translated into a 7-point colour scale, marking start points of every new road segments assessed, which were 50 meters long, with a pushpin of different colour. The darker the colour, the higher the respective RRUI value was, marking greater relative unevenness, thus the lower overall road segment quality. Internally, the RRUI scale is a dimensionless value ranging in the case of the presented road from -2 to 51, where -2 is a road better than the baseline, and 51 is worse.

As the virtual road was constructed first from the "good" road segments, the initial RRUI values for the first road segments are low – from about 5 to 22, where high value of 22 is in the position of the road artefact, the speed bump, as is presented in the Table 3. The next virtual road segments, described in the Table 4, built upon rubble road fragments, should be much higher, which is correct again, with values ranging from 23 to even 50.

Table 3. RRUI values for virtual road fragments constructed from good roads.

Segment number	RRUI value
1	14.56
2	13.04
3	18.23
4	14.16
5	10.07
6	22.18
7	10.07
8	5.58

Table 4. RRUI values for virtual road fragments constructed from rubble roads.

Segment number	RRUI value
16	23.76
17	28.26
18	20.91
19	51.49
20	26.42

To check an another algorithm, the road artefact detection method using the modified Z-STDEV method [8] was performed on the same virtual road as before. Using these methods, road artefacts which were included in the virtual road (on position of 6th, 17th and 21st segment, respectively), were detected with one false positive in data point number 106, which is in a second good road segment (class A, variant b).

Similar tests were performed for the set of multiple generated virtual roads and in every case the RRUI values were correctly corresponding to the type of the road.

5 Conclusions and future work

The virtual road concept is an interesting way of constructing purely artificial signals for road artefacts algorithms detection training and testing. The authors also believe that building a normalised database of road artefacts and marking them in a simple text format would be useful for sharing virtual roads between researchers.

The streams, which were decomposed after the extraction phase, may be modified before virtual road construction. The authors are aware of the possibilities of this step, introducing changes to different data streams – for example scaling the Z-acceleration value to simulate deeper potholes. This step is not taken into consideration at this stage, but it will be one of the interesting points in the future for the presented concept.

The second shortcoming in the presented stage of research is no possibility for road artefacts of "point" class to be included into the middle of the "wide" class, being placed only sequentially. This is the next issue for the authors to solve into changing the presented concept to cover more realistic simulation of the real-world roads.

References

1. P. Mohan, V. N. Padmanabhan, R. Ramjee, Nericell: Rich Monitoring of Road and Traffic Conditions using Mobile Smartphones. Proc. 6th ACM Conf. Embed. Netw. Sens. Syst. - SenSys '08 (2008), p. 323

2. A. Mednis, G. Strazdins, M. Liepins, An. Gordjusins, L. Selavo, 136 (2011), doi:10.1007/978-3-642-22185-9

3. V. Astarita, R. Vaiana, T. Iueleand, V. C. Maria, V. P. Giofre, F. De Masi, Procedia - Soc. Behav. Sci. 111, 242–251 (2014)

4. M. Afrin, M. R. Mahmud, M. A. Razzaque, 2015 IEEE Int. WIE Conf. Electr. Comput. Eng. WIECON-ECE 2015, 495–498 (2016)

5. K. Laubis, V. Simko, A. Schuller, C. Weinhardt, Proc. 50th Hawaii Int. Conf. Syst. Sci., 1582–1591 (2017)

6. Lars Forslöf, Hans Jones, J. Civ. Eng. Archit. 9, 485–496 (2015)

7. L. C. Gonzalez, R. Moreno, H. J. Escalante, F. Martinez, M. R. Carlos, IEEE Trans. Intell. Transp. Syst., 1–13 (2017)

8. M. Badurowicz, J. Montusiewicz, in Information and Software Technologies, ICIST 2015 [WOS], R. Damasevicius, G. Dregvaite, Eds. (Springer, Berlin, Germany, 2015; http://link.springer.com/10.1007/978-3-319-24770-0_43), pp. 503–514

9. M. Badurowicz, T. Cieplak, J. T. Montusiewicz, Appl. Comput. Sci. 12, 18–27 (2016)

10. J. Smolka, M. Skublewska-Paszkowska, Hum. Syst. Interact. (HSI), 2016 9th Int. Conf., 126–132 (2016)

11. M. Badurowicz, T. Cieplak, J. Montusiewicz, in Computer Networks: 23rd International Conference, CN 2016, Brunów, Poland, June 14-17, 2016, Proceedings, P. Gaj, A. Kwiecień, P. Stera, Eds. (Springer International Publishing, Cham, 2016; http://dx.doi.org/10.1007/978-3-319-39207-3_31), pp. 360–369

12. World Geodetic System 1984, (available at http://spatialreference.org/ref/epsg/wgs-84/)