PAN POLSKA AKADEMIA NAUK

METROLOGY AND MEASUREMENT SYSTEMS

Index 330930, ISSN 0860-8229

www.metrology.pg.gda.pl

M&MS

ECONOMICAL METHODS FOR MEASURING ROAD SURFACE ROUGHNESS

Dariusz Grabowski, Maciej Szczodrak, Andrzej Czyżewski

Gdańsk University of Technology, Faculty of Electronics, Telecommunications and Informatics,
G. Narutowicza 11/12, 80-233 Gdańsk, Poland (✉ *ksm@sound.eti.pg.edu.pl*, 48 58 347 1114)

Abstract

Two low-cost methods of estimating the road surface condition are presented in the paper, the first one based on the use of accelerometers and the other on the analysis of images acquired from cameras installed in a vehicle. In the first method, miniature positioning and accelerometer sensors are used for evaluation of the road surface roughness. The device designed for installation in vehicles is composed of a GPS receiver and a multi-axis accelerometer. The measurement data were collected from recorded ride sessions taken place on diversified road surface roughness conditions and at varied vehicle speeds on each of examined road sections. The data were gathered for various vehicle body types and afterwards successful attempts were made in constructing the road surface classification employing the created algorithm. In turn, in the video method, a set of algorithms processing images from a depth camera and RGB cameras were created. A representative sample of the material to be analysed was obtained and a neural network model for classification of road defects was trained. The research has shown high effectiveness of applying the digital image processing to rejection of images of undamaged surface, exceeding 80%. Average effectiveness of identification of road defects amounted to 70%. The paper presents the methods of collecting and processing the data related to surface damage as well as the results of analyses and conclusions.

Keywords: image processing, neural networks, pothole detection, Kinect, Raspberry Pi.

1. Introduction

Road damage is caused by many factors: from atmospheric conditions, through heavy traffic, to erosion. In many cases, there is a little time between damage and significant deterioration of the road surface, which in turn can result in damage to vehicles, increase fuel consumption and cause accidents. High costs of carrying out road condition tests are the reason of their reduced frequency, thus extending the time needed to locate damages requiring repair. Very high prices of the equipment used by road managers for this purpose have a significant impact on the costs and on the scale of such research. Meanwhile, the development of ICT enables to invent cost-effective solutions that can be applied for this purpose. Although the precision of low-cost surface scanners may be lower than expensive professional equipment, they can be more widely used and thus may potentially increase effectiveness of the road damage detection.

Any road surface impediment – such as potholes, ruts, railroad crossings or traffic calming devices – becomes an issue, both for the technical conditions of vehicle suspension operation and

Article history: received on Nov. 26, 2017; accepted on May 26, 2018; available online on Sep. 30, 2018, DOI: 10.24425/123905.

the safety of road users. Therefore, the road surface quality assessment plays an important role in the infrastructure management and in effective planning of corrective actions. For drivers and vehicle owners, there is also important increased energy consumption caused by the road surface impediments requiring drivers to slow down and then accelerate [1].

The research presented in the paper was performed by the authors in Poland. In the Ordinance of the Polish Minister of Transport and Maritime Economy of 2 March 1999 on technical conditions to be met by new surfaces of public roads (Journal of Laws, No. 43, item 430), the requirements for the smoothness parameters that should characterise a new pavement of a public road were defined and the requirements for the measurement equipment were also included. Profile measurements can be made using a laser or other devices that meet the requirements of the above regulation.

In the Polish standard, there are distinguished four classes of technical conditions of the pavement [2]. Class A (very good) and class B (satisfactory) conditions constitute the desired level. It applies to new or renewed pavements, where only occasional damages are allowed. Class C is a warning level applied to pavements with a significant damage requiring repair. The critical level or class D requires immediate repair. The road pavements of class D usually mean numerous and extensive damages.

Sensors play an important role in creating a system that analyses the surface condition. They provide the information necessary for analysis of a tested object/environment. The type, quantity and quality of these devices have a major influence on the results of data processing. The variety of used devices enables to analyse many physical quantities, whereas the quantity and quality of the devices influence the reliability of the results obtained from analysing the collected data.

Professional measurement instruments usually use laser distance sensors. The sensors employing a laser to determine the distance generate a beam that is directed towards the object being tested. As a result of the diffuse reflection, part of the laser beam returns to the transmitter. The sensor measures the time between sending and receiving the beam, and on the basis of the speed of light it is capable to determine the distance between the object being tested and the transmitter. These sensors can be used in detection of asphalt discontinuities and in testing of depths of individual defects. Using many of these sensors in a single line enables to effectively represent the surface condition.

The following chapters of the paper present two proposed methods for estimating the surface condition and the results of their operation. The first one is based on the use of acceleration sensors and analysing signals obtained from inside a vehicle. In turn, the second method is based on analysing images obtained from external cameras attached to the vehicle, and it assumes the use of neural networks to detect unevenness in the image representation. The presentation of these methods will be preceded by a review of known solutions in this field, prepared on the basis of a literature review.

2. Review of methods for road surface estimation

2.1. Accelerometer-based solutions

A device equipped with an accelerometer (one or more) and a GPS is located in a vehicle moving along a road. The basic assumption for detecting unevenness with a sensor is that the sensor detects a significant shock due to spotting irregularities of surfaces, such as potholes, manholes, sewers and similar ones. The accelerometer readings are compared with reference shocks when driving on a level surface. Based on the difference in sensor signals, the level

of unevenness on the road is determined. The data from the GPS receiver enable to place the estimated data in the geographical space. This solution was used, among others, in the Pothole Patrol system [3]. In taxi network cars there were installed Linux-based computers, equipped with a set of sensors and Wi-Fi modules. The detection of unevenness was performed locally, while the results and positions were sent to the central server. In this way, the data from a specific vehicle could be verified (or rejected) by other vehicles within the system.

A similar solution was used in CRSM [4]. The crowdsourcing technique was used – to delegate execution of the complex task to a large community. In this case, 100 taxis with dedicated sensors were used. The vehicles collected information on the surface in Shenzhen (China) and by using an appropriate algorithm the information about irregularities was sent to the central server. As a result, 90% accuracy was achieved in the detection of continuity interruptions on roads with negligible horizontal false declarations [4].

Among the available solutions, BusNet should also be mentioned [5]. It was decided to use the public transport to monitor the surface condition by equipping buses with acceleration sensors. The system consisted of a sensory network based on the MICA platform. The role of mobile sensors was played by sensors installed on buses, whereas intermediate stations were located near bus stops. The collected data were eventually transmitted to the central server for analysis.

A different solution was proposed in 2011 in Latvia [6]. A session of 10 car rides on a 4.4 km long road segment was carried out, during which the condition of the surface was analysed using 4 Android phones. Each of them, apart from dedicated software, contained an integrated accelerometer and a GPS receiver. The sensor auto-calibration was an important part of this solution, enabling the device to work in different vehicles.

The work by Tomiyam *et al.* [7] presents the results obtained from two accelerometers mounted on the outside of a vehicle. Although the results are unaffected by suspension damping (sensor in the control arm area), the experiments have been carried out with special sensors, which, in turn, have required additional professional and reliable installation in the vehicle suspension components.

Strazdins *et al.* [8] compared the results of measurements using GPS sensors and accelerometers from three different cell phones. Similar measurements were presented by Tai *et al.* [9], although the vehicle was a motorbike in that case.

2.2. Solutions based on 3D image reconstruction

The first of these systems is based on the structural light technology. The University of Carnegie developed a surface damage detection platform. For this purpose, a laser sensor emitting the pattern, a 15 Hz video recording camera and a GPS were used to determine the location. The sensors scanning a selected road section were mounted on a vehicle, moving at a speed of 14.3 m/s. From the collected data they managed to generate a 3D map of the scanned road, which was later used for the damage detection. Recently, the team of authors of that solution was working on developing an algorithm for distinguishing road cavities from *e.g.* sewer manholes [10].

Another solution used, among others, by the General Directorate for National Roads and Motorways in Poland, is a laser profilograph [11]. One such device is a Dynatest's product, model 5051 Mk-II Test System. The device is installed on a beam mounted at the front of the vehicle, as shown in Fig. 1, equipped with densely distributed laser sensors. They scan the road surface area approximately every 5 mm and then the obtained data are stored in the fixed memory. The test is based on recording the maximum value of ground clearance between the damaged surface and the horizontal beam [12]. Usually the measurements are taken at 50 km/h, but it is also possible

to perform tests in a range from 20 km/h to 110 km/h without any significant deterioration of results [13]. This makes it possible to carry out research also on high-speed roads and motorways without causing traffic difficulties.

Fig. 1. A professional laser profilograph (A) mounted on a vehicle [11].

In addition to laser sensors, the device is equipped with accelerometers, gyros, a GPS receiver and a distance meter. This enables to study the longitudinal and transverse roadway profiles, the curvature of curves, *etc*. [14].

Another solution worth mentioning is a stereovision system constructed at the Silesian University of Technology [15]. The platform performs surface condition assessment based on stereo images – sequentially taken images by two CCD cameras while the vehicle is in motion. Additionally, it is equipped with sensors to identify the movement and to change the rotation angle of the car wheel. During analysis, the sensor assembly produces an image with information on depth. This solution was collated and compared with the measurement of the same surface with a laser rangefinder for 160 different cross-sections. It has been shown that differences between both methods are only incidental, which proves effectiveness of the system. It is worth mentioning that the developed method can be applied regardless of the weather conditions – unlike the laser research, where the road surface must be dry [15].

An interesting solution to scanning bicycle paths, country lanes, unpaved roads, etc. was developed by the *Schniering* company [16]. The *Argus Agil* device is equipped with HD (High Definition) video recording instruments, laser sensors and a GPS receiver. The sensor set enables testing surface features at speeds from 0 km/h to 60 km/h. After driving over 9000 km, *Argus Agil* proved to be a valuable commercial solution for determining the conditions of cycle paths in Germany.

2.3. Solutions based on vision systems

The first system described below for detection of unevenness in the road surface is based on video data [17]. The solution uses the optical data recorded by a dedicated electronic device installed in a car, and then the generated frames analysed to detect the road continuity interruptions. The device is designed with a 1280×720, 60 fps camera, a GPS receiver and a CMOS camera with a signal processor. The image processing is performed by removing noise, increasing contrast, binarization, and finally extracting the area containing damage.

Another solution based on analysing videos was proposed in the paper [18]. In contrast to not very accurate solutions based on shock sensors and accelerometers, as well as expensive laser solutions, the system has been created which enables detecting breakdowns of asphalt continuity with the use of an in-car recorder. The analysis uses a lane detection algorithm. The method consists of digitizing the image to detect lines defining the limits of a lane/roadway. Then, after the image denoising, a fragment of the image is determined in which unevenness is to be detected. The testing platform includes a 1920×1080 pixel camera and a Cortex-A8 processor for recording and processing data. The tests were carried out on the basis of 20 videos recorded in a sunny weather on a motorway. The achieved sensitivity was 71%, while the precision was 88%. A weakness of this solution is false detection with changing light intensity or lack of detection of the presence of flat unevenness.

The last example discussed in this brief overview is detection of unevenness based on static 2D images. In the study there is proposed to divide the images into the areas containing defects and those without them [19]. A potential segment with a defect is subject to geometric approximation. For a chosen area, its texture is taken from the original image and then it is compared with the textures of the nearest areas without defects. If the comparison shows a significant difference between the images, the area is defined as a cavity in the roadway. The algorithm has been implemented in the MATLAB environment with a support of Image Processing Toolbox, tested with 120 images of damaged and undamaged roadway segments. Some of them have been downloaded from the Google Images search engine. The remaining ones were acquired from a remotely controlled robot equipped with an HP Ellite Autofocus webcam with a resolution of 640×480 pixels. Finally, 50 images were used for training of the algorithm, whereas the remaining 70 were used for testing it. The research results have shown 82% precision of this solution.

3. Accelerometer method

The modern car suspension is a sophisticated mechanical system capable to absorb various road irregularities protecting the car users from uncomfortable shakes. The transmission of vibration is described by various suspension models [20, 21].

The aim of suspension design is to minimize the transfer function of vibration caused by a rough road to the vehicle's body, over a range of frequencies between 0.1 and 100 Hz [22]. The transfer function of vibration $|H|$ caused by a rough road to the vehicle's body is defined as the ratio of amplitude X of equivalent mass m and amplitude X_0 of the displacement excitation produced by a rough pavement: $|H| = X/X_0$ [22].

The authors' approach is to use the Gabor transform for the detection of road surface anomalies. The Gabor transform is widely applied in the analysis of time series, particularly those originating from vibrations [23, 24]. The experiments focus on the comparison of results of the method applied to various devices, vehicle speeds and types. The main motivation to use smartphones is the common availability of rich sensor-equipped devices. Such mobile data gathering units do not require any electronic hardware preparation, thus interested car users can easily participate in the measurements. The experimental data were collected from multiple passes of 3 types of cars on roads of various quality. The focus was put on the analysis of the characteristic of a poor-quality road section by means of measuring vibrations by commonly available smartphone sensors.

The first step of data processing is the sensor alignment made to achieve proper values of the acceleration vector. The ideal vector $[a_x, a_y, a_z]$ takes the form: $[0, 0, g]$. In real conditions, the

sensor placement causes that z axis is not perpendicular to the gravity force, thus the readout is: $[b_x, b_y, b_z]$. The rotation matrix that rotates vector **a** onto vector **b** is given by the formula:

$$\mathbf{R} = \mathbf{I} + [\mathbf{v}]_\times + [\mathbf{v}]_\times^2 \frac{1-\mathrm{c}}{\mathrm{s}^2}, \tag{1}$$

where:

$$[\mathbf{v}]_\times = \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix}; \qquad \mathbf{v} = [v_1 \;\; v_2 \;\; v_3] = \mathbf{a} \times \mathbf{b}; \qquad \mathrm{s} = \|\mathbf{v}\|; \qquad \mathrm{c} = \mathrm{a} \cdot \mathrm{b},$$

I – identity matrix.

The Gabor transform is applied to the accelerometer data representation (a_z) aligned according to (1) and then coefficients $G_{n,k}$ are obtained:

$$G_{n,k} = \sum_{i=0}^{N-1} x(i)\gamma^*(i-n)e^{-j2\pi ki/N}, \tag{2}$$

where: $\gamma^* = \overline{g_\infty}$; $g_\infty(t) = \dfrac{1}{2\pi\alpha}e^{-t^2/4\alpha}$; N – the number of frequency bins (carrier frequencies of the Gabor elementary functions), $\propto$ – a scale of window, n,k – location coordinates of $G_{n,k}$ in the time-frequency plane, $x(i)$ – the input signal.

The transform was applied to 100 m long road sections. The experimentally obtained number of frequencies N equalled to 40, and α was equal to 8 [25].

An individual threshold was applied to the calculated magnitude of Gabor coefficients, dependent on the sensor sensitivity and on the vehicle type. The maximum of the magnitude of the coefficients over the length of the examined road section was determined for each pair (vehicle, sensor). The calculated values correspond to the largest variations of the vibration amplitude. The threshold was determined based on the maximal magnitudes of Gabor coefficients, a good quality road surface serving as a reference.

The output value for each road section represents the maximum degree of surface degradation in that road section (the lower the value the better the surface quality).

4. Video-based method

In order to develop and to implement a method of road surface roughness vision detection, a system based on visual classification (RGB image and depth sensor image) with the use of a neural network was developed.

The system was designed to operate in a centralised environment. The idea was to enable simultaneous work of many client units mounted on vehicles. They should be able to continuously transfer the collected data to the central server for further processing. On the other hand, it was also necessary to provide an interface that would enable road users/managers to monitor the current condition of the road without delay. The system consists of 2 elements:

- Client – the data generator – a module located directly on a vehicle, responsible for the collection and preliminary processing of graphical data and GPS positions;
- Server – responsible for the collection and classification of information sent from the client.

Bringing the design assumptions into work, the system was divided into three interrelated applications, as shown in Fig. 2:

- DataCollector – an application running on client devices, whose task is to collect graphical data and GPS positions and to pre-process this information;
- PhotosReceiver – an application running on the server whose task is to analyse the data sent by DataCollector and to save related information in the database;
- WebCgi – a server-based application that provides a web user application that shows the results on the map.

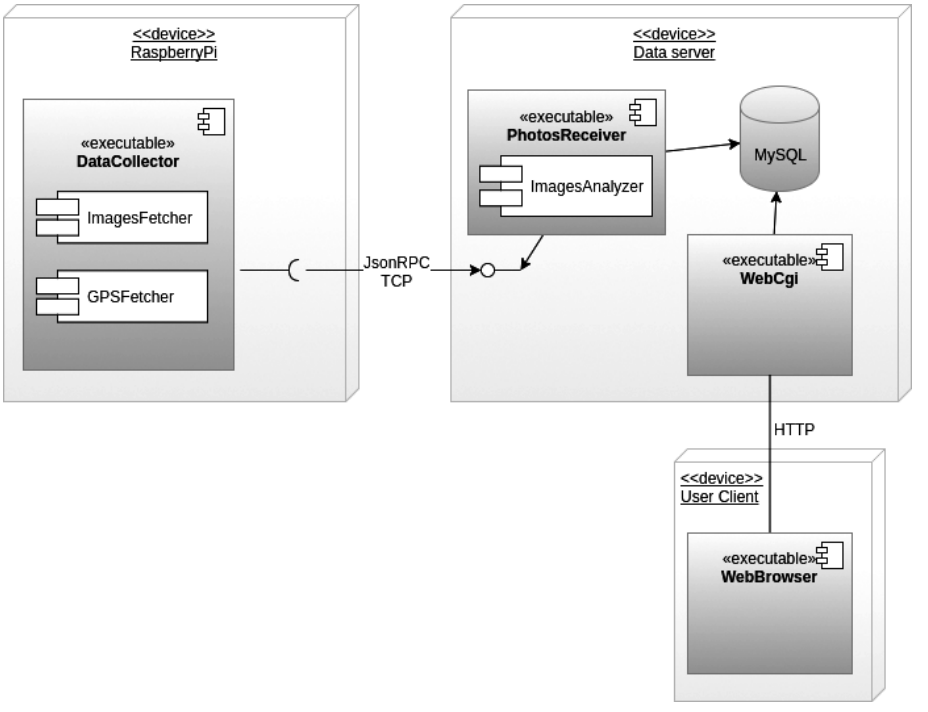

Fig. 2. A diagram of arrangement of software units on processors.

DataCollector and PhotosReceiver applications communicate with each other via the Json-RPC protocol on TCP sockets. It is an RPC (*remote procedure call*) based on the Json format. An open-source libjson-rpc-cpp library is used for this purpose. Based on the interface created in Json, the generator creates C++ header files for client and server applications. This makes it possible to call up the method remotely without the need for additional parsing of transmitted data – the transport layer remaining invisible to the user.

Figures 3 and 4 show the flow of depth and RGB colour images through the system. The functions represented by the blocks are discussed later in the description.

The main design assumption for the hardware platform was the mobility of the solution. Among the client subassemblies, the main component is a Raspberry Pi 2 Model B V1.2 prototype board; an RGB-D Kinect v1 sensor and a GPS receiver are used to collect the video data.

The application together with the libraries was installed on the Raspberry Pi processor board. The aforementioned Kinect cameras and GPS receiver were connected to this computing unit. The complete assembly was powered by the car battery through a mounted inverter. The Kinect

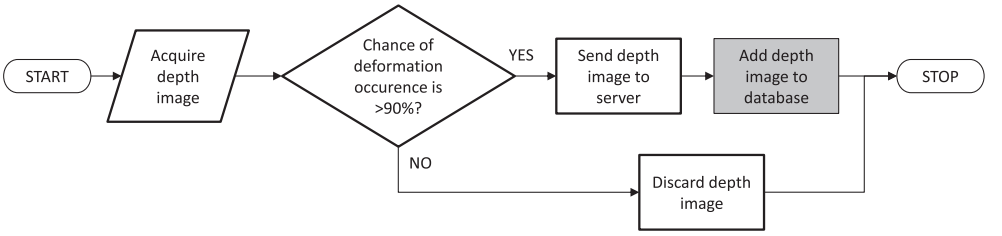

Fig. 3. The flow of depth images through the system. Bolded contours indicate the elements in the client application. The elements of the server application are marked with filling.

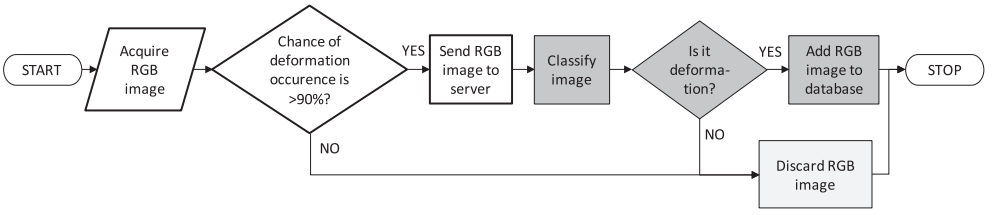

Fig. 4. The flow of RGB images through the system. Unfilled blocks indicate the elements in the client application. The elements of the server application are marked with grey filling. Light grey filling indicates part of both sides.

sensor was mounted on the car engine flap using a mounting tape. The sensor's framerate enabled to drive at a speed of about 30 km/h. A photo of the vehicle setup for the test drive is shown in Fig. 5.

Fig. 5. A car equipped with a Kinect sensor (A).

In the client part, the first step is to analyse depth images before uploading them to the server. The aim is to identify the images that contain a distortion element of the road surface. This approach reduces the amount of data transferred to the server. A set of algorithms has been

developed based on 500 depth image samples taken, which enabled to extract valuable data from a noisy image.

The consecutive steps of the process presented in Fig. 6 are:

a) capturing the original depth picture;
b) equalizing the histogram;
c) applying a box-type linear low-pass filter with a window size of 19×19 pixels;
d) using the Sobel operator with a west-eastern mask of size $k = 5$;
e) using the Sobel operator with an eastern-western mask of size $k = 5$;
f) combining the images obtained in (d) and (e);
g) image binarization using a constant $t = 30$;
h) happlying the dilation with a $k = 3$ mask and double erosion with the same mask.

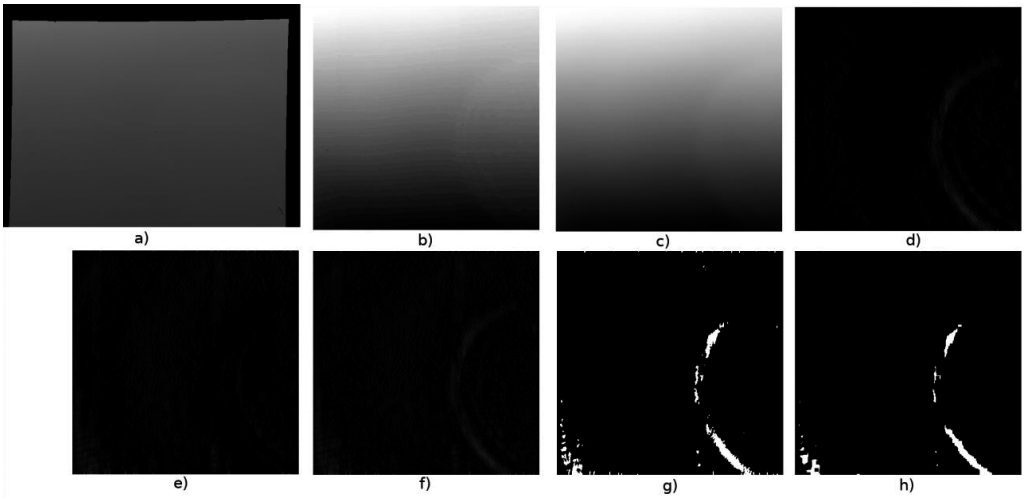

Fig. 6. Steps of processing the depth images.

The biggest challenge in the above analysis was to find a suitable algorithm of edge detection. Due to the angle of attack of the camera, it is possible to observe tonal transition from dark (bottom) to light (up). Most of the methods that work well for RGB images, failed here, in the context of detecting the edges between different stripes of lightness. Only the application of aforementioned Sobel's directional operators enabled to solve the problem.

A preliminary analysis of RGB images is then performed. Due to a limited processing power of the client device, it was decided to make only a preliminary checking if the photo can show a damage of the road. It was assumed that the analysis and classification demanding a higher computing power would take place on the server side.

One of the main assumptions for this analysis was the shortest possible time of its execution, therefore complex image processing algorithms were not used. The consecutive steps in this process are presented in Fig. 7:

a) conversion of the original image to greyscale;
b) the use of a box-type linear low-pass filter with a window size of 11×11;
c) binarization of the image with the Adaptive Gaussian method, a block size $b = 25$;
d) applying the dilation with a $k = 3$ mask and double erosion with the same mask.

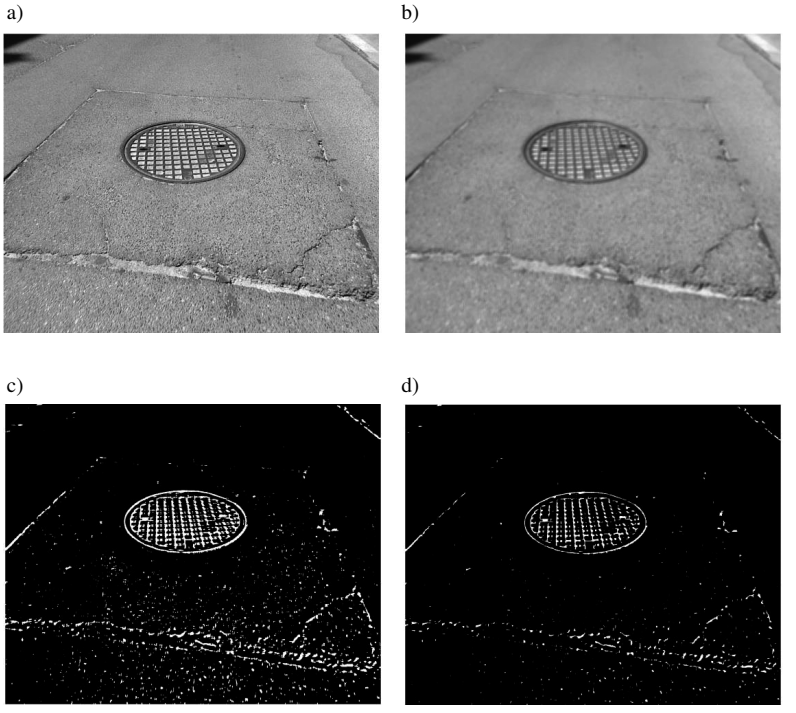

Fig. 7. Steps of initial RGB image processing.

In the second step, the processed image was subject to an assessment of the probability with which the relevant information on it is likely to be present. This was done by counting the white pixels and comparing them with the accepted value determined experimentally by analysing 10,000 photos divided into categories: photographs containing damage and pictures just depicting the roadway. The selected factor enables to reject insignificant images with simultaneous preservation of the proper ones.

The methods of image classification were applied in the server. An important piece of information for the user/road manager is what kind of damage he or she is dealing with. An advantage of classifying these objects is that they can be filtered when searching for specific elements, *e.g.* a sewer manhole. An effective way of doing that is to use a neural network. Based on a large collection of categorized data being its input, the process of network training is carried out. As a result, the network identifies the image characteristics and assigns each of images to one of categories. The network trained in this way can be used to classify images from outside the input set.

In image recognition applications, the CNN (*Convolutional Neural Network*) networks are leading the forefront [26]. For example, in the ILSVRC (Large Scale Visual Recognition Challenge) [27] held since 2010, these are the only solutions used due to their high efficiency. The popularity of these networks has grown along with the increase in the CPU performance and the development of the GPU technology, especially by NVIDIA. The system described above uses an 18-layer CNN network called Resnet-18, developed by Microsoft [28]. The basic parameter for choosing this solution was the ratio of training time and prediction time to the average error rate.

To start the neural network training process, one needs to prepare a collection of images divided into categories. The more numerous and diverse this collection is, the better learning results can be achieved. Therefore, 10.000 photos of a roadway were collected, selected, and then – based on the guidelines [11] – divided manually into the categories:

- potholes – localised losses in bituminous layers;
- sewer manholes;
- mesh cracks and single cracks;
- filled-in and laid-over patches – local areas repaired with the help of an extra layer of bitumen;
- road without any damage.

In the case of a smaller collection (potholes), their photos were downloaded from the Google Images repository. For even better results, it was decided to make a few simple transformations on the collected images, in this way increasing the material input and increasing the learning process efficiency. Examples are shown in Fig. 8, but in practice different combinations of techniques were used. The photos were then binarized by applying:

- noise reduction using a non-local medium algorithm with a search window size of 21 and intensity $h = 19$;
- adaptive Gauss binarization with a 13 pixel window size.

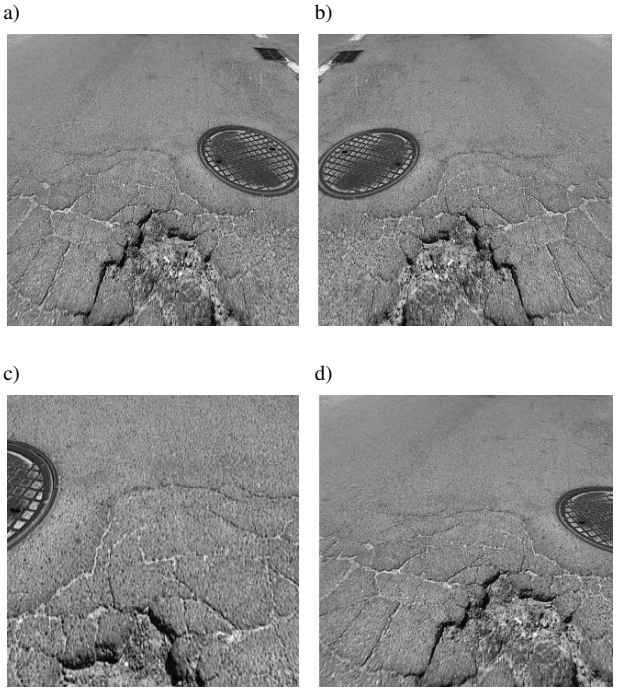

Fig. 8. Selected transformations of the source image original image (a); mirror (b); approximation (c); offset to the right (d).

The next step in the process was transformation of images to the RecordIO format, accepted by the MXNet library. This was done in order to scale the images to square dimensions (in this case 320×320 px), to change the quality of graphics to 90% and to describe an appropriate record label, so that one could identify the group to which it belongs. As a result, a single file was

created containing all the entered images. For the purpose of network training, it was necessary to generate two such files – the first one contained a set of 90% of all photos (training data). The remaining 10% was the content of the second file (verification data).

The network learning was carried out using the Amazon Web Services' cloud computing. The machine selected for the task was a p2.xlarge equipped with an NVIDIA K80 GPU with 12 GB of memory, supporting the CUDA technology. The network learning during the system development was carried out on different machines. Differences in the processing efficiency between them are presented in Fig. 9.

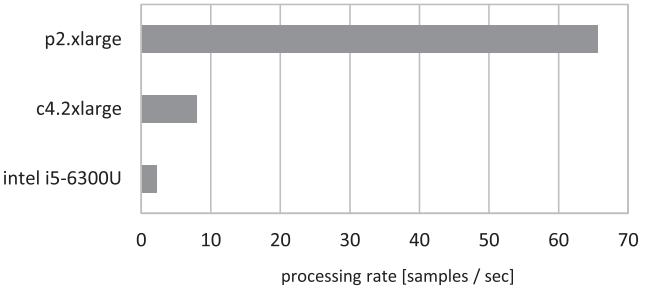

Fig. 9. The amount of processed samples during a 1 s period for different machines: **p2.xlarge** – a machine with 1 Core GPU NVIDIA K80, **c4.2xlarge** – a machine with 8 processor cores Xeon E5-2666 v3, **Intel i5-6300U** – laptop with 4 cores 2.40 GHz.

The network parameters were set as:
– the size of a single sample lot – 128;
– the number of learning iterations – 50;
– the learning rate – 0.1.

4500 previously collected photos were used for the learning process. The progress of the process is shown in Fig. 10. The final learning outcome for the test data amounted to 94%.

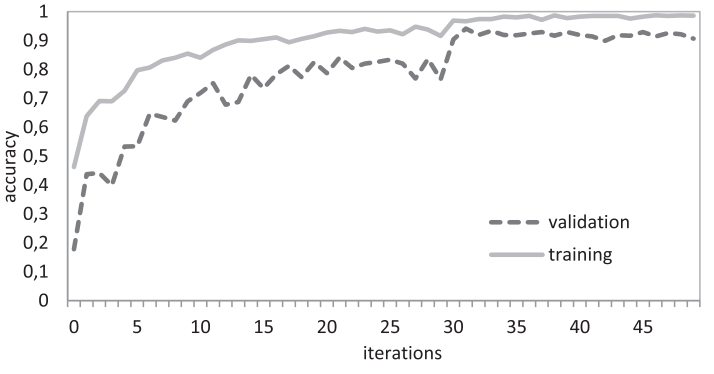

Fig. 10. Learning results in subsequent iterations for the test and learning data.

5. Experiments

5.1. Method using accelerometers

The experimental data were collected from multiple passes of 3 types of cars on roads of various quality. The focus was put on the analysis of the characteristic of a poor-quality road

section by means of vibrations measured by commonly available smartphone sensors. Three smartphones with Android system were used in the test. One Sony Xperia S with a BMA250 accelerometer chip at an 80 S/sec sampling rate was placed in the console at the car centre, and two Xiaomi Redmi with a BMI160 chip at a 200 S/s sampling rate: the Note 3 model was located on the front cockpit and 4 PRO model was placed above the rear axle.

Figure 11 presents two plots of magnitude of the Gabor coefficients calculated for the vertical acceleration signal, obtained for about 1 km long fragments of good- and poor-quality roads. Significantly different values of magnitude can be observed.

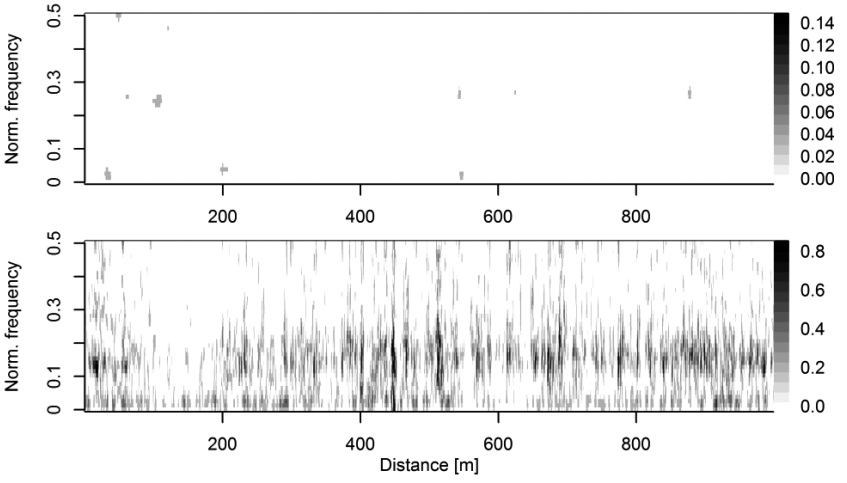

Fig. 11. Comparison of the evaluation method outputs for a flat road (upper plot) and poor-quality road (lower plot).

Table 1 presents the results of assessment of quality of each of 100 metre segments of a poor-quality road. The numbers represent normalized (for a pair "vehicle, device") magnitude maxima of the Gabor coefficients. The results are provided for the following three measurement configu-

Table 1. The values of road quality assessment coefficients (higher value means worse) for various combinations of sensors and vehicles on a poor-quality road surface.

Section [m]	V1 D3 40 km/h	V2 D3 40 km/h	V2 D1 40 km/h	Mean	Std. dev.
0–100	0.59	0.62	0.54	0.58	0.04
100–200	0.97	0.78	0.72	0.82	0.13
200–300	0.89	1.00	0.96	0.95	0.06
300–400	0.51	0.57	0.62	0.57	0.06
400–500	0.52	0.46	0.44	0.47	0.04
500–600	0.56	0.78	0.55	0.63	0.13
600–700	0.68	0.75	0.81	0.75	0.07
700–800	1.00	0.97	0.93	0.97	0.04
800–900	0.85	0.79	0.71	0.78	0.07
900–1000	0.72	0.84	0.68	0.75	0.08

rations: (vehicle 1, device 3, velocity 40 km/h), (vehicle 2, device 3, velocity 40 km/h), (vehicle 2, device 1, velocity 40 km/h).

A total of 600 road segments of various surface conditions were rated, belonging to category A and D. The results obtained with the prepared method showed the correct classification in 550 cases.

5.2. Video-based method

5.2.1. Analyses of images

The first test was the initial processing of the RGB images on a client device. The factor assumed in the application determined the probability with which an obstacle is present in the image. The test consisted of launching the analysis for a representative group of 1000 images with and without damage visible in them. As shown in Table 2, the software effectively eliminates the photos of a damage-less road, thereby reducing the amount of data necessary to be sent to the server.

Table 2. Results of the pre-selection image tests.

Images rightly rejected	Images wrongly rejected
83.80%	21.67%

Another test was performed on the classification effectiveness of the RGB images sent to the server. The idea was to check what percentage of photos was correctly classified into a correct category and what is the error rate. The defective classes were checked. The category "road without damage" included correctly non-qualified photos. To carry out the tests, 2400 photos that did not participate in the neural network training process were used. The studies have shown that – on average – in 70% of cases, images were assigned to appropriate categories. Detailed results are presented in Table 3. A relatively low score for the "Potholes" category needs clarification. The main reason for this is a small set of learning data for this category. An attempt to subside the set with photos from Google Images did not bring a desired result. The solution to it would be to extend the set with data collected from the same sensor. Excluding this category, the global average result of effectiveness reached 83.5%.

Table 3. Results of classification of damage images.

Category of damage	precision %	recall %	specificity %	accuracy %	F1 score %
Netting and single mesh cracks	93.02	57.97	99.69	96.93	71.43
Sewer drain	83.78	95.88	98.22	98.02	89.42
Built-in and laid-over patches	73.79	93.33	83.93	87.00	82.42
Potholes	29.66	98.80	71.91	74.79	45.62

5.2.2. Performance testing

In the case of the client application, not only its functionality but also performance was tested. The research concerned the data transferability through the application. The Kinect sensor, with which the images are taken, generates an average of 10 frames per second. For the colour image processing, an average duration of this operation is 4.21 ms, entailing 84 ms for the depth

image, whereas an average time of uploading images to the server through the Json-RPC protocol running on TCP is 398 ms for the colour images and 62 ms for the depth images. It is worth mentioning that the transmission time depends on the JPEG compression factor and it may differ from the presented results, *e.g.* in the case of a significant difference in illumination.

In the case of the server application, the highest CPU load was generated by the image classification. Time durations of saving a file on the SSD or adding a record to the database had a little impact on performance. The use of the trained neural network model turned out to be a bottleneck. The average processing time per photo was 1.156 s. This made it necessary to queue the submitted images.

6. Conclusions

The obtained results show that the implemented road roughness estimation method based on the accelerometer signal brought satisfactory results in the performed test series. During the experiments the distance of about 60 km was travelled.

The results obtained by means of the Gabor transform-based analysis show that the magnitude of coefficients is highly dependent on road surface impediments, returning larger values for any roughness and considerably lower ones for smooth road surfaces. On the examined roads, the classification effectiveness achieved 91%. Although there is a noticeable correlation between the acceleration bias and imperfection of pathways during the tracking, the dependence of values on damaged or quality-class roads remains a case-sensitive matter. In the non-controlled conditions, an additional factor is the driving style and the driver's reaction to a cavity spotted in the road surface. Nevertheless, the proposed method seems to be applicable to monitoring of road surface conditions employing typical cars instead of special test vehicles.

In the visual modality, a set of depth and RGB image processing algorithms were created. Then, based on the collected input, a neural network model was trained to classify the defects. Finally, the software was developed and the system was tested for effectiveness and reliability. Research has shown 83% effectiveness in the use of digital image processing to reject damage-free images. In the case of damage detection, the average effectiveness was 70%. The created system certainly opens a wide area for further development. In terms of hardware, more cameras/sensors could be installed, which would improve quality of the results. Another idea would be to replace the client device with a minicomputer containing a GPU, such as NVIDIA Jetson. This would make possible to train neural networks using models with more layers and more learning sets.

Acknowledgments

The research was subsidized by the Polish National Centre for Research and Development and the General Directorate of Public Roads and Motorways within the grant No. OT4- 4B/AGH-PG-WSTKT. The authors would like to acknowledge Ms. Karolina Marciniuk for her contribution to collecting accelerometer measurement results.

References

[1] Zaabar, I., Chatti, K. (2011). A field investigation of the effect of pavement type on fuel consumption, *Proc. of Transportation and Development Institute Congress 2011: Integrated Transportation and Development for a Better Tomorrow*. Chicago, IL, USA, 772–781.

[2] Múčka, P. (2016). International Roughness Index specifications around the world. *Road Materials and Pavement Design*., 18(4), 929–965.

[3] Eriksson, J., *et al.* (2008). The pothole patrol: using a mobile sensor network for road surface monitoring. *Proc. of the 6th international conference on Mobile systems, applications, and services*, Breckenridge, CO, USA, 29–39.

[4] Chen, K., Tan, G., Lu, M., Wu, J. (2016). CRSM: a practical crowdsourcing-based road surface monitoring system. *Wireless Networks*, 22(3), 765–779.

[5] De Zoysa, K., Keppitiyagama, C., Seneviratne, G., Shihan, W. (2007). A public transport system based sensor network for road surface condition monitoring. *Proc. of the 2007 workshop on Networked systems for developing regions*, Kyoto, Japan.

[6] Mednis, A., Strazdins, G., Zviedris, R., Kanonirs, G., Selavo, L. (2011). Real time pothole detection using android smartphones with accelerometers. *Proc. of 2011 International Conference on Distributed Computing in Sensor Systems and Workshops (DCOSS)*, Barcelona, Spain.

[7] Tomiyama, K., Kawamura, A., Nakajima, S., Ishida, T., Jomoto, M. (2012). A Mobile Profilometer For Road Surface Monitoring By Use Of Accelerometers. *Proc. of 7th Symposium on Pavement Surface Characteristics: SURF 2012*, Norfolk, VA.

[8] Strazdins, G., Mednis, A., Kanonirs, G., Zviedris, R., Selavo, L. (2011). Towards vehicular sensor networks with android smartphones for road surface monitoring. *Proc. of 2nd international workshop on networks of cooperating objects*, Chicago, IL, USA.

[9] Tai, Y., Chan, C., Hsu, J.Y. (2010). Automatic road anomaly detection using smart mobile device. *Conference on technologies and applications of artificial intelligence*, Hsinchu, Taiwan.

[10] Mertz, C. (2011). Continuous road damage detection using regular service vehicles. *Proc. of the ITS World Congress*, Orlando, FL, USA.

[11] General Directorate for National Roads and Motorways in Poland, Measurement subsystems. *https://www.gddkia.gov.pl/pl/a/6791/podsystemy-pomiarowe*, (2012).

[12] Braga, A., Puodžiukas, V., Čygas, D., Laurinavičius, A. (2003). Investigation of automobile roads pavement deterioration trends in Lithuania. *Journal of Civil Engineering and Management*, 9(1), 3–9.

[13] Cigada, A., Mancosu, F., Manzoni, S., Zappa, E. (2010). Laser-triangulation device for in-line measurement of road texture at medium and high speed. *Mechanical Systems and Signal Processing*, 24(7), 2225–2234.

[14] Talvik, O., Aavik, A. (2009). Use of Fwd Deflection Basin Parameters (Sci, Bdi, Bci) for Pavement Condition Assessment. *The Baltic Journal of Road and Bridge Engineering*, IV(4), 196–202.

[15] Staniek, M. (2017). Detection of cracks in asphalt pavement during road inspection processes. *Scientific Journal of Silesian University of Technology. Series Transport*, 96, 175–184.

[16] Schniering ARGUS AGIL, *http://www.schniering.com*, (Oct. 2017).

[17] Kim, T., Ryu, S. (2014). System and Method for Detecting Potholes based on Video Data. *Journal of Emerging Trends in Computing and Information Sciences*, 5(9), 703–709.

[18] Jo, Y., Ryu, S. (2015). Pothole detection system using a black-box camera. *Sensors*, 15(11), 29316–29331.

[19] Koch, C., Brilakis, I. (2011). Pothole detection in asphalt pavement images. *Advanced Engineering Informatics*, 25(3), 507–515.

[20] Alvarez-Sánchez, E. (2013). A quarter-car suspension system: Car body mass estimator and sliding mode control (The 3rd Iberoamerican Conference on Electronics Engineering and Computer Science). *Procedia Technology*, 7, 208–214.

[21] Thoresson, M., Uys, P., Els, P., Snyman, J. (2009). Efficient optimisation of a vehicle suspension system, using a gradient-based approximation method, part 1: Mathematical modelling. *Mathematical and Computer Modelling*, 50(9), 1421–1436.

[22] Suciu, C.V., Tobiishi, T., Mouri, R. (2012). Modeling and simulation of a vehicle suspension with variable damping versus the excitation frequency. *Journal of Telecommunications and Information Technology*, 1, 83–89.

[23] Russell, P., Cosgrave, J., Tomtsis, D., Vourdas, A., Stergioulas, L., Jones, G. (1998). Extraction of information from acoustic vibration signals using Gabor transform type devices. *Measurement Science and Technology*, 9(8), 1282–1290.

[24] Liang, T.C., Lin, Y.L. (2012). Ground vibrations detection with fiber optic sensor. *Optics Communications*, 285(9), 2363–2367.

[25] Carmona, R., Hwang, W., Torresani, B. (1998). *Practical Time-Frequency Analysis: Gabor and Wavelet Transforms, with an Implementation in S*. San Diego, CA: Academic Press.

[26] Goodfellow, I., Bengio, Y., Courville, A. (2016). *Deep learning*. MIT Press. *http://www.deeplearningbook.org*.

[27] Russakovsky, O., Deng, J., Su, H., Krause, J., *et al*. (2015). Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3), 211–252.

[28] He, K., Zhang, X., Ren, S., Sun, J. (2016). Deep residual learning for image recognition. *Proc. of the IEEE conference on Computer Vision and Pattern Recognition*, 770–778.