Review Article

Pavement Management Utilizing Mobile Crowd Sensing

Boquan Tian, Yongbo Yuan, Hengyu Zhou, and Zhen Yang

Faculty of Infrastructure Engineering, Dalian University of Technology, Linggong Road, No. 2, 116024 Dalian, China

Correspondence should be addressed to Yongbo Yuan; yongbo@dlut.edu.cn

Received 30 August 2019; Revised 4 May 2020; Accepted 3 June 2020; Published 23 June 2020

Academic Editor: Chongchong Qi

Copyright © 2020 Boquan Tian et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Pavement management, which is vital in road transportation and maintenance, is facing some troubles, such as high costs of labors and machineries, low detecting efficiency, and low update rate of pavement conditions by means of traditional detection ways. Benefiting from the development of mobile communication, mobile computing, and mobile sensing techniques, the intelligence of mobile crowd sensing (MCS), which mainly relies on ubiquitous mobile smart devices in people’s daily lives, has overcome the above drawbacks to a large extent as one new effective and simple measure for pavement management. As a platform for data collection, processing, and visualization, a common smart device can utilize inertial sensor data, photos, videos, subjective reports, and location information to involve the public in pavement anomalies detection. This paper systematically reviewed the studies in this field from 2008 to 2018 to establish an overall knowledge. Through literature collection and screening, a database of studies was set up for analysis. As a result, the year profile of publications and distribution of research areas indicate that there has been a constant attention from researchers in various disciplines. Meanwhile, the distribution of research topic shows that inertial sensors embedded in smartphones have been the most popular data source. Therefore, the process of pavement anomalies detection based on inertial data was reviewed in detail, including preparatory, data collection, and processing phases of the previous experiments. However, some of the key issues in the experimental phases were investigated by previous studies, while some other challenges were not tackled or noticed. Hence, the challenges in both experiment and implementation stages were discussed to improve the studies and practice. Furthermore, several directions for future research are summarized from the main issues and challenges to offer potential opportunities for more relevant research studies and applications in pavement management.

1. Introduction

Pavement management has always been of great importance in road transportation and maintenance, which affects not only the comfort and safety of drivers and passengers but also the service lives, fuel consumption, and emission of vehicles [1]. The coverage of roads and amounts of vehicles continuously increase as road transportation still plays a key role in modern transportation (see Figure 1 [2] and Figure 2 [3]).

Especially in China, roads are frequently being excavated and reconstructed due to the extension and rehabilitation of underground pipelines and cables of pipelines and cables. As a result, potholes and other anomalies are all over the pavements (see Figure 3). Additionally, in the Technical Code of Maintenance for Urban Road [4], detection of urban road includes routine patrol, periodic, and special detections. However, routine patrols consume labor power of experienced road engineers every day. In the meantime, there is at least one-year interval between the implementations of periodic detections, which are incapable of meeting the demand of urban construction and development. Otherwise, periodic detections require expert equipment, such as laser and inertial profilometers, which are usually very expensive. Unfortunately, this situation is not restricted to China [5, 6].

The citizens who drive on the roads every day can record the urban pavement conditions automatically and “subliminally” by the means of smart devices they take along at nearly any time [7]. With the increasing amount of automobile and smartphone possession, mobile crowd sensing (MCS) has been applied to many research fields, such as
The amount of vehicles and private cars (million)

<table>
<thead>
<tr>
<th></th>
<th>2014</th>
<th>2015</th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small passenger service vehicles</td>
<td>117</td>
<td>136</td>
<td>158</td>
<td>180</td>
<td>201</td>
</tr>
<tr>
<td>Private cars</td>
<td>109</td>
<td>127</td>
<td>149</td>
<td>170</td>
<td>189</td>
</tr>
</tbody>
</table>

**Figure 1:** The amount of small passenger service vehicles and private cars in China in the past five years.

Total length of roads

<table>
<thead>
<tr>
<th></th>
<th>2014</th>
<th>2015</th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length (million kilometers)</td>
<td>4.46</td>
<td>4.57</td>
<td>4.69</td>
<td>4.77</td>
<td>4.85</td>
</tr>
<tr>
<td>Density (km/100 km²)</td>
<td>46.5</td>
<td>47.8</td>
<td>48.9</td>
<td>49.7</td>
<td>50.5</td>
</tr>
</tbody>
</table>

**Figure 2:** The total length and density of roads in China in the past five years.

Anomalies on the pavement.

**Figure 3:** Anomalies on the pavement.
environment monitoring, smart transportation, urban management, and public security. This concept came from the idea of crowdsourcing [8] proposed in 2018 and can be divided into two categories, participate sensing and opportunity sensing, according to the initiative of the participants.

The development of MCS provides more efficient approaches to gather, transmit, analyze, and store data of pavement conditions for transportation authorities [9]. Pavement detections (potholes, roughness, etc.) could be conducted through accelerations, images/videos, and even acoustics, by the sensors embedded in vehicle-mounted mobile smart devices, to give guidance for pavement operation and maintenance and to provide timely essential data for work zone coordination on the pavement [10]. Besides, along with positioning technology and wireless transmission technology, MCS basically meets the technical requirements in pavement management [11] and the development of intelligent pavement [12]. Hence, the intelligence of MCS builds a bridge between transportation authorities and the public, which makes pavement management more interactive, economical, efficient, and timely.

In the past decade, utilizing MCS to improve pavement management has been developing rapidly. Nevertheless, researchers barely review the MCS applications for pavement management systematically. Therefore, this paper aims to provide a systematic and comprehensive review of previous studies and to serve as a reference tool for researchers, authorities, and the public of pavement management.

The remainder of this paper is organized as follows: the next section describes the process of literature collection, screening, and analysis to establish an overall understanding of this field. In the light of reviewed literature, Section 3 provides an introduction of detecting pavement anomalies based on MCS in an order of the experimental process. Section 4 discusses the key issues in the experiments, practical implementation, and promotion of the method. In Section 5, several directions for future research are summarized from the challenges and main problems in the previous sections to offer potential opportunities for more relevant researches and applications in pavement management. Finally, the last section states the conclusions.

2. Methodology

2.1. Literature Collection. To obtain a comprehensive dataset of publications, we adopted the Web of Science (WoS) Core Collection as the main data source, EI Village, and Scopus as the supplement for this study. Figure 4 illustrates the process in which we constructed our database through a two-stage literature selection method [13]. Multiple combinations of keywords were utilized to collect an exhaustive list of papers from WoS and Scopus. Although there were plenty of papers related to MCS applications in pavement management, the descriptions of this field were in various ways. For example, when referring to “pavement management,” researchers from different countries use different expressions, such as “road surface management,” “road surface condition management,” and “road technical condition management.” Meanwhile, papers about potholes, bumps, anomalies, defects, and so on, which are obviously related to pavement management, rarely mentioned the relative concepts in the keywords or even in the main body of the text.

After various tests, the search conditions were finalized as "road" AND (smartphone* OR mobile* OR crowd*sens* OR crowd*sourc*) AND (pavement OR surface OR pot*hole* OR bump* OR anomal* OR roughness OR IRI OR maintenance OR deterioration* OR crack* OR defect* OR classif* OR detect* OR survey") in the topic. We used the wildcard notation asterisks to consider plural forms, tenses, and different spelling habits (e.g., “road” and “roads,” “smartphone” and “smart phone” and “smartphones” and “smart phones,” “crowdsensed” and “crowd sensed” and “crowd-sensed” and “crowdsensing” and “crowd sensing” and “crowd-sensing,” etc.).

As a result, there were 307 papers retrieved, including 178 proceedings papers, 127 journal papers, and 5 reviews. Data cleansing was performed after retrieval, because many papers in this dataset did not match our preset topic. First, we excluded 21 obviously irrelevant categories (e.g., Chemistry Analytical) and 23 papers in these categories. Second, the irrelevant papers were manually removed through reading all the abstracts remaining in the dataset (and the articles if necessary). Although some papers mentioned the keywords we chose, these papers do not study our topic about MCS applications for pavement management. For example, papers on Intelligent Transportation mentioned “road,” “smart phone,” and “detection,” but these researches aim to solve urban traffic problems. Another example is that some papers about the unmanned ground vehicle (UGV) utilize smartphones to obtain graphic information of the pavement to avoid obstacles on the road. This reduced our dataset to 109 papers.

In addition, an EI search and a Scopus search were conducted to supplement our dataset, because it was found that there were a few relevant papers highly cited or recently published but not indexed in the WoS database. We used the same search strategy and exclusive rule as the WoS search. This resulted in 7 additional papers. Overall, there were 116 papers in the final database.

2.2. Results

2.2.1. Year Profile of Publications. The first publication about MSC application in pavement management was in 2008 since the wide use of smartphone and the equipment of embedded sensors first started in 2007, the year that the first iPhone appeared on the market. Thus, we collected the papers published from 2008 to 2018. Figure 5 shows the annual publication numbers of the 116 papers in the past decade.

Generally, the annual number of related papers presents a growth trend and reaches an average of 16, although there are fluctuations in annual paper quantities. Before 2012, no more than three publications were related to MCS applications for pavement management annually. In the years of 2012 and 2014, there were obvious increases in this field. In the meantime, the popularization and generalization of
mobile smart devices with embedded sensors created a rich field for MCS based pavement management experiments and applications. The trend of publication number indicates that both researchers and managers in pavement management are paying more and more attention to MCS research studies and applications. The amount slightly decreased in 2018, while there are still plenty of problems to overcome in this field.

2.2.2. Distribution of Research Topic. In general, all the papers used the smartphone as a tool to gather, calculate, display, or share the information of road pavements. At first, we obtained a distribution of research areas from Clarivate Analytics on WoS. Top 10 research areas and the record count in Table 1 indicate that this field has drawn attention from a variety of areas. The efficiency of infrastructure management has been improved by the integration of computer science, telecommunication, and other disciplines.

<table>
<thead>
<tr>
<th>Research areas</th>
<th>Record count</th>
<th>Proportion (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engineering</td>
<td>75</td>
<td>64.655</td>
</tr>
<tr>
<td>Computer science</td>
<td>54</td>
<td>46.552</td>
</tr>
<tr>
<td>Telecommunications</td>
<td>29</td>
<td>25.000</td>
</tr>
<tr>
<td>Transportation</td>
<td>24</td>
<td>20.690</td>
</tr>
<tr>
<td>Instruments instrumentation</td>
<td>7</td>
<td>6.034</td>
</tr>
<tr>
<td>Remote sensing</td>
<td>7</td>
<td>6.034</td>
</tr>
<tr>
<td>Automation control systems</td>
<td>5</td>
<td>4.310</td>
</tr>
<tr>
<td>Optics</td>
<td>4</td>
<td>3.448</td>
</tr>
<tr>
<td>Physics</td>
<td>4</td>
<td>3.448</td>
</tr>
<tr>
<td>Chemistry</td>
<td>3</td>
<td>2.586</td>
</tr>
</tbody>
</table>
Then, a manual reading selection was conducted to establish a more refined classification. Among the 116 papers, 94 papers utilized smartphone embedded inertial sensors to acquire data, which were defined as “inertial” papers, while the remaining 22 papers used smartphones in other ways. Table 2 shows the distribution of research topics, including two main categories, seven subcategories, and the most used methods of each topic. The 22 “noninertial” papers included evaluation of pictures, videos, and subjective reports, image and video recognition, security and privacy, GPR-based and sound pressure-based evaluation, and even task assignment. All the remaining 94 papers utilized inertial sensors to collect data. Most of these papers directly referred to accelerometers, gyroscopes, magnetometers, and other smartphone embedded inertial sensors, while the above papers described the same sensors as vibration sensors. Hence, the “inertial” papers, which held the majority, were further discussed in detail in the next section.

3. Pavement Anomalies Detection Based on Inertial Data

The purpose of this section is to introduce the whole procedure of detecting pavement anomalies based on inertial data and to probe the key issues in each experimental phase, including preparatory phase, data collection phase, and data processing phase (see Figure 6), in order to provide objective supports for the maintenance and reconstruction of road pavements. The scope of the preparatory phase is from the very beginning of experiment design until accomplishing all the considerations in the laboratory. Next, the data collection phase starts when we mount our smartphones in the vehicles and finishes when getting all the data we need. Finally, the data processing phase means the duration when we reduce noise and extract features with specific methods until we draw conclusions from the experiments.

3.1. Preparatory Phase. In the preparatory phase, researchers design the framework of their experiments. First, the targets of pavement anomalies detection utilizing MCS are to find out where the pavement anomaly locates, the kind of destruction form, and how serious it is, when the users are providing the required data unintentionally. Most of the reviewed studies focused on the first target, while few papers classified the types or discussed the severity of pavement anomalies. Jang et al. [36] classified the data into impulse, rough, and smooth using a multilayer feedforward neural network, which was a supervised machine learning technique. Orhan and Eren [37] differentiate the size of potholes as big, medium, small, and no pothole by the means of threshold and Dynamic Time Window (DTW). M. Taha and Nasser [38] provided an example colour code of six hierarchies based on the threshold to classify pavement conditions. Xue et al. [39] inferred the depth and length of the pothole and achieved estimation error rates of 13 and 16 percent. Harikrishnan [40] found that the detection of pothole depth and the hump height were more accurate at a low velocity of 15–20 km/h. Silva et al. [41] identified 382 anomalies and labelled the anomalies with “No anomaly,” “Manhole,” “Short Bump,” “Long Bump,” and “others.”

Second, the parameters that require confirmation usually consist of acceleration, rate of rotation, change of the ambient geomagnetic field, GPS location and velocity [42, 43], and so forth. Z-axis acceleration is the most essential parameter for detecting the vertical vibration caused by pavement anomalies. The rate of rotation and change of the ambient geomagnetic field is mainly used for reorientation since the smartphone could have some orientation and position changes during the vehicle moving on. The rate of rotation can also reflect the pavement anomalies directly according to Alqudah and Sababha [44].

Meanwhile, GPS location information recorded by time can provide a location where pavement anomalies happen. Data from a smartphone GPS can only be recorded one time per second [45] and its precision is 10 meters nowadays. However, smartphone developers are paying more attention to the locate function, such as MI 9 embedded with dual-band and dual-polarized antenna GPS [46, 47]. The precision can also be improved by the assistance of cell-tower triangulation. Peng et al. [48] enhanced the average location accuracy about 25%. Gawad et al. [49] used reinforcement learning to address the challenge of position accuracy.

Additionally, velocity is another important parameter for the success of pavement detection, as velocity can affect the duration in which a vehicle goes through a pavement anomaly. For example, when a vehicle drives at a velocity of 60 km/h (16.7 m/s), it will pass a 20 cm pothole in 0.012 s, which requires the sampling rate of at least 83 Hz to collect a vibration signal. Meanwhile, we can estimate whether the vehicles are on elevated or surface roads by utilizing velocity along with the average number of perceived satellites [50, 51]. Nevertheless, velocity data gained from smartphones at present can only reach a sampling rate of 1 Hz, which comes from GPS location calculation.

It is notable that, nowadays, sampling rates of the sensors are various because the brands and manufacturers are different and although we can set the sampling rate to a certain value to record the accelerations and rotations in the applications, the hardware sometimes could not achieve the purpose. For instance, when the sensors can only provide 100 groups of data per second, we can set the sampling rate up to 300 Hz, which means the application records the value of accelerations and rotations 300 times per second, so that the recorded values would consist of three consecutive identical values for one second 100 times. However, the requirements for sampling rate vary in different reference indices and methods of data processing. Bridgelall [52] recommended that when using the Road Impact Factor (RIF) to measure ride quality, a standard sampling rate of 64 Hz was convenient to estimate the parameters of the suspension system model. Another issue that might need attention is that some of the previous studies used “frequency” to express “sampling rate,” which could possibly result in misunderstanding for the concept in mathematical statistics. For example, in Douangphachan and Onyama [53, 54], a Frequency Domain Analysis (FDA) was conducted, and it was found that the frequency range of 40–50 Hz was more...
appropriate to estimate IRI, in which the “frequency” was relative to the method of FDA, but not the sampling rate.

Third, a suitable application of smartphone should be developed or selected based on the required parameters. The previous studies provided multiple choices of applications from both Android OS and iOS, such as “RoadScan” [55], “UNiquALroad” [56], “RESen” [57], “Streetbump” [58], “PAVVET” [59], and “RoadLab Pro” [60]. The available applications, which can record comprehensive parameters and have an alternative sampling rate, should be considered initially.

Fourth, the precision of embedded sensors is another important issue to consider before conducting the experiments, although some of the previous studies have discussed the usability of embedded sensors. For instance, Mukherjee and Majhi [61] and Zhang et al. [62] compared the accuracy of high precision sensors and sensors of basic android smartphones ($70) in the laboratory and proved there were no obvious differences between them through a vibration table test. Zhao et al. [63] proved that the accelerometers of 6th generation iPod touch and servo-type accelerometer showed good agreement in the sampling rate of 0.2 to 15 Hz. However, in the study of Gurdit Singh et al. [64], the precision of the sensors in Samsung Note 3 is obviously higher than the precision of the other four smartphones. In the study of Aleadelat and Ksaibati [65], the sensors of Samsung Galaxy S III had a lower error than Sony Xperia. Hence, it is an essential work to validate the precision of the embedded sensors if conditions permitted.

### Table 2: Distribution of research topics and main research methods.

<table>
<thead>
<tr>
<th>Category</th>
<th>Research target</th>
<th>Number of studies</th>
<th>Main research methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Inertial”</td>
<td>Data from inertial sensors</td>
<td>94</td>
<td>Threshold, filtering, machine learning, Fourier transforms, frequency domain analysis, and wavelet transform</td>
</tr>
<tr>
<td></td>
<td>Pictures, videos, and subjective reports</td>
<td>7</td>
<td>Subjective evaluation [14–20]</td>
</tr>
<tr>
<td>“Noninertial”</td>
<td>Image and video recognition</td>
<td>7</td>
<td>Machine learning [21–24], Candy edge detection and profile-image [25], multiscale ridge detection filter and boundary contour analysis [26], and Harris corner detection and SIFT descriptor [27]</td>
</tr>
<tr>
<td></td>
<td>Security and privacy</td>
<td>3</td>
<td>Fog computing [28–30]</td>
</tr>
<tr>
<td></td>
<td>GPR</td>
<td>2</td>
<td>Fast Fourier transform [31, 32], Threshold [33, 34]</td>
</tr>
<tr>
<td>“Noninertial”</td>
<td>Sound pressure</td>
<td>2</td>
<td>Formulation and optimal sensing path scheduling [35]</td>
</tr>
<tr>
<td></td>
<td>Task assignment</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Figure 6: Key issues in each phase of the experiment.
3.2. Data Collection Phase. In the data collection phase, four vital issues should be considered, which include choosing appropriate carriers for the smartphone, reasonable placement location for the smartphone, and data storage or transmission.

Smartphones used for data collecting can be placed in vehicles not limited to automobiles but also on a bicycle [66–68], a motorcycle [69], and even in a train [70] or a metro vehicle for different scenarios or purposes. A bicycle’s reaction when going into a pavement anomaly can be more reliable and accurate than that of an automobile. This is because the natural vibration of bicycles is fewer than that of automobiles, which have various kinds of suspension systems and different weights. In the meantime, the velocity of bicycles is much lower than that of automobiles, which means that bicycles can collect more acceleration signals when crossing a pavement anomaly, although it is disadvantageous in long-distance detection missions. Seraj et al. [71] evaluated the IRI (International Roughness Index) of 500 km bike paths, while smartphones are mounted on bicycles.

Previous studies generally placed the smartphones horizontally on the dashboard of the automobiles, in the drivers’ or passengers’ pockets or on smartphone holders and so on, while some researchers compared the effects of different placement locations. González et al. [72] placed the smartphones in the driver door, center console of the vehicle, driver’s shirt pocket, driver’s pants pocket, and lady purse/backpack on the copilot seat, and it showed that the acceleration from five smartphones had the same variation trend, when the smartphone in the pants pocket had the maximum amplitude. Singh et al. [64] placed the smartphones inside the vehicle at different positions (Front Dashboard (Pilot Seat/Copilot Seat), Back Seat) and found that the detection rate of the smartphone placed at the backseat of the vehicle is lower as compared to other smartphones. The different reactions of embedded accelerometers in different placement locations mainly come from the suspension system of automobiles. Hence, the results will be more significant, when the placement location is nearer to the suspension system. Some other researchers placed their smartphones freely, without any special attachment. Koban et al. [73] conducted their experiment using bicycles, putting the smartphone into pants’ side pockets or a bag. Douangphachanh and Oneyama [53] used the data from gyro meters and placed one smartphone inside the driver’s shirt front pocket and the other one in a box near the gearshift, and the smartphone did not need to be reoriented.

With the development of mobile communication technology, 3/4G mobile networks have almost covered the whole world, and even 5G is being available. In this situation, the cost and speed of mobile transmission are adequate to meet the demand of mass data uploading from smartphones in time. Previous studies on this topic either saved the data in smartphones’ internal or external storage and got them out before data processing or uploaded the data in time through mobile networks. Both options are feasible for research on account that the accuracy of road pavement anomalies detection is as important as the in-time data transmission and calculation. Freschi et al. [74] amply explained the issues of data storage and transmission without processing steps. Geem et al. [75] introduced three different data retrieval methodologies based on CAN-bus, OBD, and smartphone in three different stages of their project.

3.3. Data Processing Phase. The data processing phase is a vital step of this topic. The first purpose of data processing was to find out what kinds of patterns of signals indicated the pavement anomalies and where the pavement anomalies located. Most of the time, peaks of reoriented Z-axis acceleration result from the vibrations caused by pavement anomalies, while the natural vibration of the vehicle suspension system and the braking pressed by drivers can also create peaks of Z-axis acceleration [76, 77]. Since then, numerous methods were utilized to reduce the effects of natural vibration and to recognize the feature of sensor signals, so that computers and smartphones could automatically detect the pavement anomalies.

This paper summarized the methods used to process the signal data from smartphone-embedded sensors. The most referred methods and the annual distribution are shown in Figure 7. Among the 94 reviewed inertial papers, the filter was the most frequently mentioned, followed by the threshold, Machine Learning (ML), Fourier Transform (FT), Frequency Domain Analysis (FDA), and Wavelet Transform (WT) technique. Over 40% of the studies referred to filter and threshold each year, while Machine Learning drew attention from the researchers since 2014 and was almost as popular as filter and threshold. Therefore, this section mainly introduces the utilization of filter, threshold, and Machine Learning methods in pavement anomalies detection.

In the reviewed studies, 67 papers referred to methods based on the filter. Multiple filtering methods were implemented to deal with the noise of sensor signals, the contribution of gravity, and the redundancy of collected data. Additionally, some filters were utilized for collecting metadata, categorizing the data access patterns [78], directly measuring impulse patterns [79], and amplifying the acceleration signals [80]. The most popular filters in this topic included high-pass filter [81], low-pass filter [82, 83], and moving average filter [84]. Giacomo Alessandroni et al. [85] improved a fixed first-order low-pass filter by computing a prediction filter with the Levinson–Durbin recursion. Sharma et al. [86] applied different filters for different artifacts such as bump or pothole to remove noise from input data. It is worth mentioning that the Kalman Filter (KF) was utilized over ten times in the past five years. Fouad et al. [87] used KF to reduce noisy samples of GPS readings. Tan et al. [88] regard KF as a general denoising method and utilized KF to remove the noise from the accelerometer and gyroscope. Aly et al. [89] employed a probabilistic framework based on Extended KF to update the map.

The second most mentioned method was the threshold, since the amplitude of the Z-axis acceleration intuitively exhibits the magnitude of the vibration due to pavement anomalies. Mohan et al. [90] considered the spikes as suspect
bumps where the Z-axis acceleration was greater than 1.75 g at a high velocity (>25 km/h) and the spikes as sustained dips where the Z-axis acceleration was lower than 0.8 g at a low velocity. Mednis et al. [91] improved the threshold algorithm and compared the positive rates of Z-THRESH, Z-DIFF, STSEV(Z), and G-ZERO. The results showed that the optimal algorithms for classifying potholes, gaps, and drain pits had a positive rate of 90%. Vittorio et al. [92] not only identified the specific thresholds to compare the estimated impulse values but also registered the location where the events were higher than thresholds which can be a probable localization of road anomalies. Vittorio et al. [92] proved that a vertical acceleration impulse (DVA) threshold of 2.5 m/s² could identify about 98% of stone bumps. Jang et al. [93] improved the reliability of anomalies detection using both threshold-based filtering and a trained classifier. Arroyo et al. [94] provided five adapted thresholds for calibration for five different vehicles. However, fixed thresholds were unable to cope with the variety of vehicles, smartphones, and velocities. In other words, thresholds in the previous studies could be ineffective when changing the experimental parameters and it was a huge work to calculate a suitable threshold for each scenario.

With the development of Computer Science and Artificial Intelligence, Machine Learning method is getting more and more popular in the field of academic research and is widely used in automated reasoning, pattern recognition, computer vision, and so forth. In our dataset, 54 papers mentioned Machine Learning and 25 papers detected the pavement anomalies by the means of this method, including Support Vector Machine (SVM), Artificial Neural Networks (ANN), C4.5 Decision Tree (DT), K-means, Naïve Bayes classifier (NB), Random Forrest (RF), Hidden Markov Model (HMM), and Gradient Boosting Classifier (GB). Machine Learning method was first applied in this topic in the year of 2008. Eriksson et al. [95] combined signal processing method and machine learning method to detect adverse road conditions, and in the proposed system 90% of the reported detections needed to be repaired when evaluated on thousands of kilometers data collected by “uncontrolled” drivers. However, the studies that followed using Machine Learning appeared late, after several years. In the year of 2014, Kattan and Aboalmaaly [96] utilized ANN to differentiate pavement anomalies from normal roads, and the success rate was over 61%. González et al. [97] used ANN and logistic regression, of which the classification accuracy was 86%. Since then, at least five papers utilizing Machine Learning to solve the problem of pavement anomalies based on MCS were published each year until 2017, although the amount of publications utilizing Machine Learning decreased slightly due to the decreasing of total publication amount in 2018.

SVM was the most popular method above all, as a rapid and reliable classification algorithm. From 2015 to 2018, 15 papers utilized SVM to detect pavement anomalies. Five studies took SVM as the main research method (see Table 3). The achievements of these studies have proved that considerable accuracies can be obtained by using SVM.

The other studies compared SVM to other Machine Learning methods (see Table 4), which resulted in multiple and incompatible outcomes, maybe due to various experimental conditions, such as different vehicles, smartphones, pavement conditions, and even regional climate differences. However, there is no research devoted to comparing these contrasts and indicating the most suitable method.

3.4. Brief Summary. This section completely expounded the experimental process of pavement anomalies detection utilizing data from smartphone embedded inertial sensors. Although many attempts have been made, there are still some key issues that need to be noticed in the preparatory...
and data collection phases. In the data processing phase, an accepted suitable method has not been discovered. Since then, the adopted methods in future studies need some indices for comparison with other methods, such as accuracy and false positive rate. Meanwhile, it is necessary to compare the adopted methods and the results of the method and results in previous studies. Although the publication amount decreased in 2018, some challenges in both experiment and implementation stages were not tackled or noticed.

4. Discussion

In this section, we arrange the discussions in two parts, the experiment stage and implementation stage, where the former devotes to the realizing of sensing pavement anomalies and the latter devotes to the popularizing of our proposed approach. The purpose of this arrangement is to provide some references for future studies and some practical issues in real scenarios in pavement management utilizing MCS.

4.1. Experiment Stage. In the experiment stage, there are some improvable issues in the reviewed studies, such as the full utilization of available parameters, the stages of development, and the validation of detection experiments. As mentioned in the preparatory phase, the required parameters should be confirmed before conducting the experiments. Unfortunately, the parameters from previous studies were underutilized. Velocity is an important parameter to be collected. Most of the previous studies realized this but avoided the effects of velocity by setting it at one or two fixed values. Nevertheless, it is difficult to maintain a fixed velocity when we drive in real scenarios, especially in a densely populated city with heavy traffic. Therefore, flexible velocity should be allowed so that the users can provide the required data unintentionally [112, 113]. For instance, the group of different velocity can be divided as 0–20 km/h, 20–40 km/h, 40–60 km/h, and so forth or even more meticulous. Alternatively, Alessandroni et al. [114] figured that vertical acceleration tended to decrease when velocity increased, which depended on a gamma law.

It was generally accepted that the acceleration of the Z-axis of the smart devices would significantly change when vehicles go through pavement anomalies [115–117] (at a fixed velocity in the past). However, the results were not the same as we thought in some researches. Gunawan et al. [118] proved that, in some circumstances, the acceleration of the X-axis is more convicive than that of the Z-axis. Sometimes, the vertical acceleration signals were too random to prove the presence of pavement anomalies [119, 120]. Hence, it is necessary to take advantage of multidimensional data.

It is worth mentioning that, in the age of big data, the types of smartphones and vehicles should be recorded in detail in the smartphone Apps for pavement anomalies detection, so that data from different devices can be classified

<table>
<thead>
<tr>
<th>Authors</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fox et al. [98]</td>
<td>99.6% detection rate with 0.01 false alarms per true pothole in simulated environment</td>
</tr>
<tr>
<td>Nalavde and Surve [99]</td>
<td>78.53% accuracy and 21.10% false positive rate</td>
</tr>
<tr>
<td>Wang et al. [100]</td>
<td>Safety related anomalies: 88%, dangerous anomalies: 82%</td>
</tr>
<tr>
<td>Fox et al. [101]</td>
<td>Detect over 90 percent of potholes with 20 vehicles on a single-lane road and 100 vehicles per lane on a two-lane road with 2 m localization accuracy</td>
</tr>
<tr>
<td>Zhang et al. [102]</td>
<td>91.30% accuracy and 1.70% false positive rate</td>
</tr>
</tbody>
</table>

Table 3: Research studies using SVM as the main method.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Compared methods</th>
<th>Method with best performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mohamed et al. [103]</td>
<td>Three kernel functions of SVM: radial base function (RBF), multilayer perceptron (MLP), and polynomial</td>
<td>Polynomial kernel function</td>
</tr>
<tr>
<td>González et al. [72]</td>
<td>ANN, SVM, DT, RF, NB, K-nearest neighbor (KNN), and Kernel Ridge (KR)</td>
<td>ANN</td>
</tr>
<tr>
<td>Tecimer et al. [104]</td>
<td>KNN, radial basis function network (RBFN), Bayesian network (BN), logistic model trees (LMT), MLP, and SVM</td>
<td>BN</td>
</tr>
<tr>
<td>Kalim et al. [105]</td>
<td>NB, C4.5 DT, SVM, decision tables, and supervised clustering</td>
<td>Two SVM classifiers</td>
</tr>
<tr>
<td>Brisimi et al. [106]</td>
<td>SVM, logistic regression, II-logistic regression, sparse SVM, RF, C-SSVM, and AdaBoost</td>
<td>RF</td>
</tr>
<tr>
<td>Ishikawa and Fujimami [107]</td>
<td>NB, BN, MLP, Sequential Minimal Optimization (SMO, an algorithm of SVM), DT, and RF</td>
<td>RF</td>
</tr>
<tr>
<td>Allouch et al. [108]</td>
<td>C4.5 DT, SVM, and NB</td>
<td>C4.5 DT</td>
</tr>
<tr>
<td>Wang et al. [109]</td>
<td>Mahalanobis–Taguchi system (MTS), classification regression tree, and SVM</td>
<td>MTS</td>
</tr>
<tr>
<td>Carlos et al. [110]</td>
<td>NB, RF, DT, ANN, HMM, and GB</td>
<td>SVM</td>
</tr>
<tr>
<td>Kanarachos et al. [111]</td>
<td>WANEH (combining neural networks, wavelet analysis, and Hilbert transform), a probabilistic method, and a method based on wavelet decomposition and SVM</td>
<td>WANEH</td>
</tr>
</tbody>
</table>

Table 4: Research studies that compared SVM to other methods.
in the future. In this case, we can make a modification for data from different devices, according to the validated accurate data.

With the development of smart devices’ hardware and the data handling capacity, the sources of data keep varying in different stages of development. There were three main periods of development in the reviewed studies. First, most of the previous studies attempted utilizing very few devices to drive hundreds of kilometers on the road and record the signals on smart devices. Then, standard models of these devices were established to detect pavement anomalies in real scenarios. However, it is relatively inefficient when the amount of available devices is too small, which is opposite to our final purpose [79]. The second period is utilizing a certain number of vehicles and smartphones in the same type, such as buses, taxis, trains, and even patrolling cars of traffic and security departments, to minimize differences of data collection devices. In this case, the detection of pavement anomalies can almost cover the main roads of the entire city without any waste of initiative detecting. The ultimate period is to take full advantage of the power of the public, which meets the targets that detecting pavement anomalies locates when the users are providing the required data unintentionally.

Moreover, to prove the feasibility of our approaches, it is more convective to compare the detection results not only with the videos and previous data from road administrations but also with the data of high precision devices, such as road inertial profiler [121, 122], mm-wave Radar [123], and laser scanner. It can be inspired by the simulation method that if we can get an approximate 3D model of the real pavements, we could run our applications on computers, as mentioned in the study of Fox et al. [98], utilizing a simulation software called CarSim.

4.2. Implementation Stage. Many of the recent studies have achieved high detection accuracies with low error rates, but there are still plenty of obstacles to implement our applications in the public, such as the variances of vehicles and smartphones, the willingness of the public to participate in the mobile crowd sensing, and the effectiveness of data from the public.

Various natural vibration from different vehicles and various precision of different smartphone embedded sensors, as well as their combinations, are huge challenges for the application to realize in the public, since it is impossible for the researchers to conduct experiments for every kind of vehicle and smartphone. In the previous studies, acceleration, rotation, GPS location data from various smartphones and vehicles have been collected, but these data were never gathered up. The gathered data can form a database for comparing and correcting the untested data of different devices from the public, if we have complete data from other high-precision equipment and establish a standard model from certain vehicles and smartphones. Thus, the established database would be utilized in other regions in the future and develop fast and strong data from more and more device combinations.

Additionally, the participation cost of the public needs consideration since it would have a significant impact on the willingness of the public and the realizing of our proposed approach. The participation cost mainly comes from four issues: (1) the cost of vehicle damage, (2) the cost of mobile network communication, (3) battery consumption of the smartphones, and (4) the security and privacy of the MCS network. To be specific, firstly, researchers would directly hit the anomalies on a special purpose in the process of experiments, even though it is dangerous for the drivers and passengers and harmful to the vehicles’ lives as well. However, during normal driving, when drivers observe a pavement anomaly, they would slow down to avoid the anomaly by a natural reaction and hit the anomaly only when it is impossible to escape [124]. Secondly, it would cause certain mobile data charges when continuously using the function of mobile positioning and uploading the data from embedded sensors. Mohamed et al. [103] suggested that signal data from the sensors should be processed before being stored or transferred for saving the storage or reducing mobile network flow. The next important hinder to realize our purpose is the consumption of smartphone batteries. Yi et al. [125] tested the power-consuming components and found that the top three components were the screen, the user interface refreshing, and the mobile network interface. At last, users may concern about their privacy and security when they upload their data, including location information in time series, types of smartphones, and vehicles. In these situations, it is better to take an approach of participatory sensing, which adopts some incentive mechanisms, and to establish a privacy policy in order to encourage the public drivers to take part in the pavement anomalies detection. Basuden et al. [28] proposed a certificateless aggregate signcryption (CLASC) scheme to preserve privacy in detecting pavement anomalies utilizing MCS. Some resolutions to the issues of security, privacy, and fairness were provided for various vehicular crowd sensing applications [29, 30].

With other tasks of MCS, a large amount of data would be received from the public, including both useful and useless information. Some of the drivers might miss the pavement anomalies consciously or unconsciously, some of them could provide incomplete information due to their operational errors, and some of them could upload repetitive information or even fake information just to get the reward of the incentive mechanisms [126]. Therefore, there will be a great deal of work to distinguish the useful and useless information, if the detection task faces the public directly without any demands of qualification. In the meanwhile, the pavement anomalies are constantly evolving over time, which means that both the standard model and the data from MCS should be updated continuously and timely to maintain the effectiveness of the database.

5. Directions for Future Research

In this paper, five main directions for future work were identified.

(1) Utilization of multidimensional parameters from smartphone-embedded sensors
At this stage, the application of MCS in pavement management has lower dimensions, including the dimension of time and the dimension of Z-axis acceleration, and there is a lack of utilization of the other parameters that were already collected by the sensors. As mentioned in 4.1, the utilized parameters in the reviewed studies cannot be sufficient to determine the complex motions of moving vehicles. For instance, the moving vehicles have a lower chance to maintain a fixed velocity and have a large possibility to miss or steer by a small pothole on the pavement, while the dimensions of velocity and Y, Z-axis accelerations and even photos, videos, and subjective assessment from the citizens were available but were underappreciated. In the meantime, most of the proposed methods from previous studies could not meet the demand of processing a large amount of multidimensional data. So, it is a vital issue to figure out how to utilize multidimensional parameters in the application of MCS based pavement management.

(2) Validation and modification of data from various smartphones and vehicles
Detecting pavement anomalies with devices in a small quantity cannot satisfy the purpose of utilizing data from the citizens while they are unconscious. The cost of workforce, fuel consumption, and time are not fewer than those of the previous high-precision equipment when there is only one or two validated smartphones and vehicles. In the reviewed studies, it has been proved that various smartphones and vehicles can detect the pavement anomalies resulting in high precisions. Unfortunately, these experiments were not conducted in the same region, so that the results from different devices could not be compared, validated, and modified. Additionally, some studies discarded the devices that performed worse rather than making affords to modify the worse data to a validated data, when they compared the performances of a few smartphones and vehicles, which would cause a limitation of serviceable range in the applications.

(3) Integration pavement data from other equipment
Pavement data acquired by other equipment, such as laser scanner, total station, man-hold GPS, inertial road profiler, and even unmanned aerial vehicle, should be integrated with the data from mobile devices not only to estimate the accuracy of pavement detection utilizing MCS but also to establish a pavement management system (PMS). At an early stage of the proposed approach, all the information from smartphones can be validated and modified by comparison to the information from high-precision equipment, which is expensive and hard to install and maintain. In the meantime, once the data from MCS was processed, the existing PMS can update more rapidly and timely, since MCS would provide a large amount of pavement information every second.

(4) Incentive mechanism, security, and privacy
As mentioned before, there are some obstacles when implementing the proposed approach of this paper, excluding the technique issues. For instance, some of the previous studies mentioned the considerations of the public, such as the cost of vehicle damage and mobile network communication, battery consumption of the smartphones, and the security and privacy of the MCS network. The application of pavement detection utilizing MCS has a particularity, such as the continuous running of vehicles and the embedded sensors, so that the incentive mechanism, security, and privacy are different from those of the other MCS tasks.

(5) Implementation of large-scale patrolling cars in real scenarios
The pavement management utilizing MCS intends to cover the pavement of all the roads in service, within and outside the cities, while a small-scale experiment would cost an unnecessary waste and result in a low coverage on overlapping routes [127]. The large-scale patrolling, such as the researches of Chen et al. [128, 129], can be initiated by departments of road management accompanied by transportation departments and other relevant public departments, which can hire a large number of vehicles in the same type carrying the same smartphones, without any additional devices. In this case, it is more advantageous to conduct large-scale patrolling to solve the issues of vehicle and smartphone diversity and the issues of data validity, as well as the issues of data coverage. Meanwhile, the initiative of participating in data collection can be improved, since it is rewarding for the road management and transportation departments to maintain serviceability of the pavement and for the drivers of public transport to accomplish their routine tasks on comfort roads.

On the other hand, some novel challenges would emerge when implementing large-scale patrolling. For example, the refresh rate of the database should be determined to maintain the timeliness of the pavements and to adapt the capacity of storage on both smartphones and computers. Additionally, the processing capacity of CPU and GPU is considerable when dealing with large amount of uploaded data.

Generally, it is feasible and potential to implement pavement management utilizing MCS. There are block blanks in this field apart from processing the signals from sensors. Therefore, a novel interactive pavement management platform based on Mobile Crowd Sensing is expected to provide new ideas for the application of the pavement management system.

6. Conclusions
This paper reviewed researches of pavement detection utilizing Mobile Crowd Sensing comprehensively and systematically from 2008 to 2018. In this research, a two-stage literature selection method was applied to generalize the
development and the current research status of this field, resulting in a dataset of 116 papers. Then, the dataset was analyzed to provide the year profile of publications and distribution of research topics.

The application of pavement detection utilizing MCS was classified as two main categories, “inertial” and “non-inertial.” Most of the researches utilized the inertial sensors embedded in smartphones, while others estimated the pavement through text description, photos, and videos and approaches of image and video recognition. Therefore, the “inertial” papers were studied as a focus and were reviewed in an order of preparatory phase, data collection phase, and data processing phase. Based on the previous achievements, we discussed the challenges in both the experiment stage and the implementation stage and pointed out future research directions for providing some references in this field.

Generally, Mobile Crowd Sensing provides a good implementation environment and timeliness at a low cost to help with pavement management. The issues in each phase of both experiment and implementation stages can be overcome with the gradual deepening and diverging of researches and the development of hardware, algorithms, and wireless communication. Beyond that, multiple parties, including the citizens, the road management departments, traffic departments, and the construction contractor, are involved in the pavement management, so that the improvement of participation willingness plays a vital role in the promotion of our proposed approach. In conclusion, the development of pavement management utilizing MCS from experiment to implementation still has a long way to go, accompanied with great potential and a bright future.

Data Availability

The literature data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors thank those who assisted them in the process of field detection. This work was supported by the “Fundamental Research Funds for the Central Universities” (DUT18JC44).

References


[73] C. Van Geem, M. Bellen, B. Bogaerts et al., “Sensors on vehicles (SENSOVO)-proof-of-concept for road surface distress detection with wheel accelerations and ToF camera


Advances in Civil Engineering


