International Journal of Geo-Information

MDPI

Article

Embracing Crowdsensing: An Enhanced Mobile Sensing Solution for Road Anomaly Detection

Xiao Li [1,*], Da Huo [1], Daniel W. Goldberg [1,2], Tianxing Chu [3,4], Zhengcong Yin [1] and Tracy Hammond [2]

1 Department of Geography, Texas A&M University, 3147 TAMU, College Station, TX 77843-3147, USA
2 Department of Computer Science & Engineering, Texas A&M University, 3112 TAMU, College Station, TX 77843-3112, USA
3 Department of Computing Sciences, Texas A&M University-Corpus Christi, 6300 Ocean Drive, Corpus Christi, TX 78412-5799, USA
4 Conrad Blucher Institute for Surveying and Science, Texas A&M University-Corpus Christi, 6300 Ocean Drive, Corpus Christi, TX 74812-5799, USA
* Correspondence: xiao.li@tamu.edu

Received: 18 August 2019; Accepted: 8 September 2019; Published: 13 September 2019

Abstract: Road anomaly detection is essential in road maintenance and management; however, continuously monitoring road anomalies (such as bumps and potholes) with a low-cost and high-efficiency solution remains a challenging research question. In this study, we put forward an enhanced mobile sensing solution to detect road anomalies using mobile sensed data. We first create a smartphone app to detect irregular vehicle vibrations that usually imply road anomalies. Then, the mobile sensed signals are analyzed through continuous wavelet transform to identify road anomalies and estimate their sizes. Next, we innovatively utilize a spatial clustering method to group multiple driving tests' results into clusters based on their spatial density patterns. Finally, the optimized detection results are obtained by synthesizing each cluster's member points. Results demonstrate that our proposed solution can accurately detect road surface anomalies (94.44%) with a high positioning accuracy (within 3.29 meters in average) and an acceptable size estimation error (with a mean error of 14 cm). This study suggests that implementing a crowdsensing solution could substantially improve the effectiveness of traditional road monitoring systems.

Keywords: Mobile Crowdsensing; Road Anomaly Detection; Continuous Wavelet Transform; Spatial Clustering; Smartphone Sensors

1. Introduction

"No one knows how many potholes are out there, but we all agree there are a ton of them." The U.S. Federal Highway Administration (FHWA) estimates that about 52% of the U.S. highways are in a miserable condition [1]. A newly released report—*Repair Priorities 2019* shows that the percentage of "poor condition" roads in the U.S. has rapidly increased from 14% to 20% between 2009 and 2017 [2]. The category of "poor condition" road is defined by FHWA, which contains excessive road anomalies, such as potholes, bumps, and ruts. Road anomalies can not only negatively impact driving experience, but they also damage vehicle components, cause economic loss, even lead to car crashes. The American Automobile Association estimates that pothole damage costs three billion U.S. dollars in vehicle repairs nationwide annually [3]. Meanwhile, approximately one-third of traffic fatalities occur on poor-condition roads each year [4]. Therefore, effectively detecting road anomalies has become a fundamental social need, which requires immediate attention.

Traditional road anomaly detections were conducted through three main types of approaches, including 3D laser scanning, vision-based image processing, and vehicular vibration-based

analysis [5]. However, implementing these approaches requires costly and sophisticated instruments (e.g., profilometer, 3D laser scanner), which also consumes extensive time, workforce, and expertise. It is worth noting that road surface conditions can vary day by day. Existing potholes can grow larger when driven over by heavy vehicles. How to continuously monitor road anomalies with a low-cost and high-efficiency solution remains a challenging question.

Mobile sensing technology has advanced rapidly over the past a few years [6–8]. Equipped with miniaturized sensors (e.g., accelerometer, gyroscope, GPS, camera), smartphones have become promising data acquisition and computing platforms, which could achieve a high-sampling rate with little or zero economic cost. Smartphones are ubiquitous today, and they empower the citizens to sense their surroundings, generate data, and contribute their observations to achieve a continuous monitoring system in an unprecedented manner [9–11]. By leveraging the power of citizens and the rich sensing resources, mobile crowd sensing has become a popular researching paradigm for large-scale sensing and monitoring in recent years. Mobile crowd sensed data sources are transforming our life. They have been proven to be extremely efficient and have been successfully deployed to solve real-world issues, such as noise monitoring, traffic density estimation, route planning, among others [12,13].

Recent studies attempt to investigate the ability of mobile sensors in road anomaly detection. Studies have proven that smartphone accelerometers can effectively capture the vehicle vibrations caused by the unevenness of the road surface [14–16]. Through analyzing these mobile sensors' signals, we can potentially identify road anomalies.

1.1. Related Studies

Different studies have been conducted to identify road anomalies (e.g., potholes and bumps) using smartphone sensors. Among the available mobile sensors, accelerometers are most sensitive for capturing vehicle jerks when hitting bumps and potholes. The existing methods have been implemented to analyze acceleration signals, which can be broadly classified into two categories: 1) threshold-based methods and 2) machine learning methods. In recent studies, signal processing techniques, such as wavelet transforms, start being adopted to analyze mobile sensed signals. Meanwhile, implementing crowdsensing solutions has become a promising research direction, which shows a significant potential to obtain more reliable detection results by synthesizing data provided by the public.

Threshold-based methods detect road anomalies through extracting extreme values from acceleration signals. Astarita et al. [14] explored the effectiveness of built-in smartphone accelerometers for detecting speed bumps and potholes using threshold-based method. In their study, the extreme peak values along the curve of z-axis acceleration were treated as direct indicators for identifying bumps and potholes. Three filters were utilized to eliminate data noise and enhance the peak signals. The result demonstrated that speed bumps could be successfully identified by the extreme peak values of filtered z-axis acceleration with an accuracy of 90%. However, this method was less useful for locating potholes with a detection rate of around 65%. Mednis et al. [17] compared different threshold-based methods for identifying road anomalies from acceleration signals. A dedicated accelerometer was installed on a vehicle to sense its vibration. The authors found a specific data pattern while hitting potholes—acceleration readings near to be 0 m/s^2 for all three axes. Therefore, they created a G-ZERO algorithm and compared with the other three methods, including Z-THRESH, Z-DIFF, and STDEW(Z). The results demonstrated this new method can achieve 90% accuracy for detecting road anomalies. Rishiwal and Khan [18] proposed a simple threshold-based solution to measure the severity of bumps and potholes. Continuous series of z-axis acceleration were collected to represent vehicle vibrations when driving along a road. A set of thresholds were generated through empirical tests to examine z-axis acceleration, which could extract road anomalies and label their severity levels (1 to 3) with an accuracy of 93.75%. Zang et al. [19] attempted to use bicycle-mounted smartphones to measure the conditions of pedestrian and bicycle lanes. Their study also implemented a threshold-based method to extract significant spikes from the curve of vertical acceleration. These spikes were recognized as

road anomalies. The authors validated their result with 10 ground truth samples and achieved 100% detection accuracy.

Machine learning methods have also been intensively utilized in road anomaly detections. Kalim et al. [20] created a new mobile app called CRATER to identify potholes and speed bumps through machine learning methods. In their study, the authors also used the built-in accelerometer to capture the vehicle shocks and vibrations while driving. A set of features (e.g., mean, maximum, minimum speed, etc.) were generated from the collected signals. Five classifiers were compared, including naïve Bayes, support vector machine (SVM), decision tables, decision tree, and supervised clustering. The results demonstrated that SVM did the best among the five methods, which could successfully identify potholes and speed bumps with accuracy rates of 90% and 95%. Meanwhile, this paper also attempted to obtain more reliable results by leveraging crowdsourced data. The potholes had to be reported by more than five different users before publishing on the web map. Celaya-Padilla et al. [21] utilized a different machine learning approach to check the existence of speed bumps. The authors first installed some hardware sensors (e.g., three-axis accelerometer and gyroscope) on a vehicle to measure vehicle vibration. The collected data series were split into two-second subsets. Each subset was manually labeled as with or without a speed bump. Then, seven statistical features (e.g., mean, variation, skewness, etc.) were generated from each axis of the two sensors' measurements for each subset. These features were selected through a multivariate feature section strategy supported by genetic algorithms. Finally, the selected features were fed to logistic regression models to identify whether a speed bump exists in each subset. This study achieved a detection accuracy of 97.14%. A similar study was conducted by Silva et al. [22]. The authors used random forest classifier to detect road anomalies from mobile sensed data. Fifty statistical features were generated from each subset of the collected data series. Each subset contained 125 continuous three-axis accelerometer measurements. Through applying feature selection procedure, 25 features were selected and used in the classification model. This method achieved a 77.23% - 93.91% accuracy for distinguishing road with and without anomalies in different experimental settings.

Wavelet analysis has a superior ability for analyzing continuous changing signals, which shows a great potential to aid in interpreting mobile sensed data. Wei et al. [23] calculated wavelet statistics using an official roughness dataset to characterize road surface roughness. Results demonstrated that the obtained wavelet statistics showed a high correlation with officially measured roughness indexes. Recent studies attempted to use wavelet transforms to recognize bumps and potholes from mobile sensed data series. For example, Bello-Salau et al. [24] were the first to integrate wavelet transform (WT) into road anomaly detection. In their study, the authors combined a discrete WT model with the scale-space filtering algorithm to denoise the vehicle vibration signals collected from a dedicated accelerometer—NI myRIO-1950. Then, a fixed threshold was used to extract abnormal values from the denoised signals to identify the road anomalies (e.g., bumps and potholes). This study achieved relatively high accuracy for detecting bumps (96%) and potholes (94%). Rodrigues et al. [25] conducted a similar study to evaluate the effectiveness of a different discrete WT—Haar wavelet transform (HWT) for detecting potholes. The authors first created an Android-based mobile app to collect data from the built-in smartphone accelerometer. Then, HWT was applied to the z-axis accelerations in different decomposition levels to generate wavelet coefficients, which could highlight the abnormal variations when hitting potholes. Thresholds were generated based on the mean value and the standard deviation of the calculated wavelet coefficients. These thresholds were used to label the collected signals as potholes, intermediate irregularities, and acceptable perturbations. However, the authors only used two manually collected potholes to validate their result, which was not statistically sufficient.

Implementing crowdsensing solutions would be exceptionally beneficial in road anomaly detection, as it allows continuous monitoring of road surface conditions by leveraging public contributed data with little or even zero economic cost. Li et al. [15] proposed a crowdsensing solution to assess road surface conditions. The authors first used an improved threshold-based method to detect potholes. Then, the crowd sensed potholes within a 10-meter radius were aggregated into one pothole through a

simple averaging procedure. Sabir et al. [26] conducted a similar study to enhance the accuracy of the detected road anomalies. In their study, the public reported potholes within a 5-meter radius were clustered to eliminate duplicated reports. Meanwhile, road anomalies had to be reported by different users before final confirmed. This study could successfully detect 90% of speed breakers and 85% of potholes.

1.2. Knowledge Gaps

Although existing studies have proven efficient to identify road anomalies using mobile sensed data, they also expose some knowledge gaps which need to be addressed, including:

1. **Existing detection methods have apparent limitations.** Threshold-based methods need extensive empirical studies to obtain high-reliable thresholds. However, these thresholds mostly need to be adjusted and even re-tested when applied in different locations, which, in turn, significantly limits the repeatability of threshold-based methods. Machine learning methods usually require an extensive model training process based on a vast amount of labeled data, which is laborious and time-consuming. Utilizing wavelet transform (WT) can be more efficient to analyze mobile sensed data; however, integrating WT into road anomaly detection is still at a preliminary stage. To date, only a few studies reported on the utilization of discrete WT. The implementation of continuous wavelet transform (CWT) is still underexplored.
2. **Pothole size estimation is lacking.** Most existing studies focus only on identifying and locating potholes; however, few studies investigate how to estimate potholes' size using mobile sensed data. The damages caused by potholes vary by their sizes. Patching a pothole can cost about $35 to $50 U.S. dollars. Therefore, accurate and timely pothole size estimation is of great importance, which can help local governments allocate budget to fix hazardous potholes wisely.
3. **Prior crowdsourcing solutions are too simple to synthesize public contributed results efficiently.** How to leverage crowd sensed data to achieve a better road anomaly detection is still an underexplored question. Currently, only a few studies have attempted to address this question with some simple crowdsensing strategies (e.g., average the crowd sensed data). However, these studies cannot effectively integrate public contributions to optimize the detection result.

1.3. Solution and New Contributions

To fill the above-referenced knowledge gaps, we propose an enhanced mobile sensing approach to detect road anomalies. In this study, we first acquire mobile sensors' data, including three-axis accelerometer and GPS, through a customized mobile app: *PotholeAnalyzor*. We then use wavelet analysis to identify road surface anomalies (such as bumps and potholes) and measure their sizes based on the mobile sensed data. Finally, we innovatively synthesize different driving tests' results through a spatial clustering method, Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN), to optimize the detection results.

Compared with prior studies, this study makes three new contributions for road anomaly detection, including:

1. **Implement a new method.** To the best of our knowledge, this study marks the first attempt to test the performance of CWT in road anomaly detection.
2. **Provide a solution for pothole size estimation.** Pothole size estimation plays an important role in road surface management; however, it has not been considered in prior studies. This study uses an innovative wavelet-based approach to extract size information for road surface anomalies, which is a new solution to an existing problem.
3. **Put forward an enhanced mobile sensing approach.** There are some drawbacks associated with the crowd sensed data, such as data inaccuracy and redundancy. This study is among the first to investigate how to optimize road anomaly detection results by spatially clustering different

driving tests' detection results. Implement a new method. To the best of our knowledge, this study marks the first attempt to test the performance of CWT in road anomaly detection.

2. Methods

In this study, we propose an enhanced crowdsensing approach to detect road anomalies by taking advantage of CWT and spatial clustering methods. The detection process goes through three main stages as shown in Figure 1, including (1) mobile sensors' data acquisition and preprocessing, (2) road anomaly detection and size estimation, and (3) result optimization by clustering crowd sensed data.

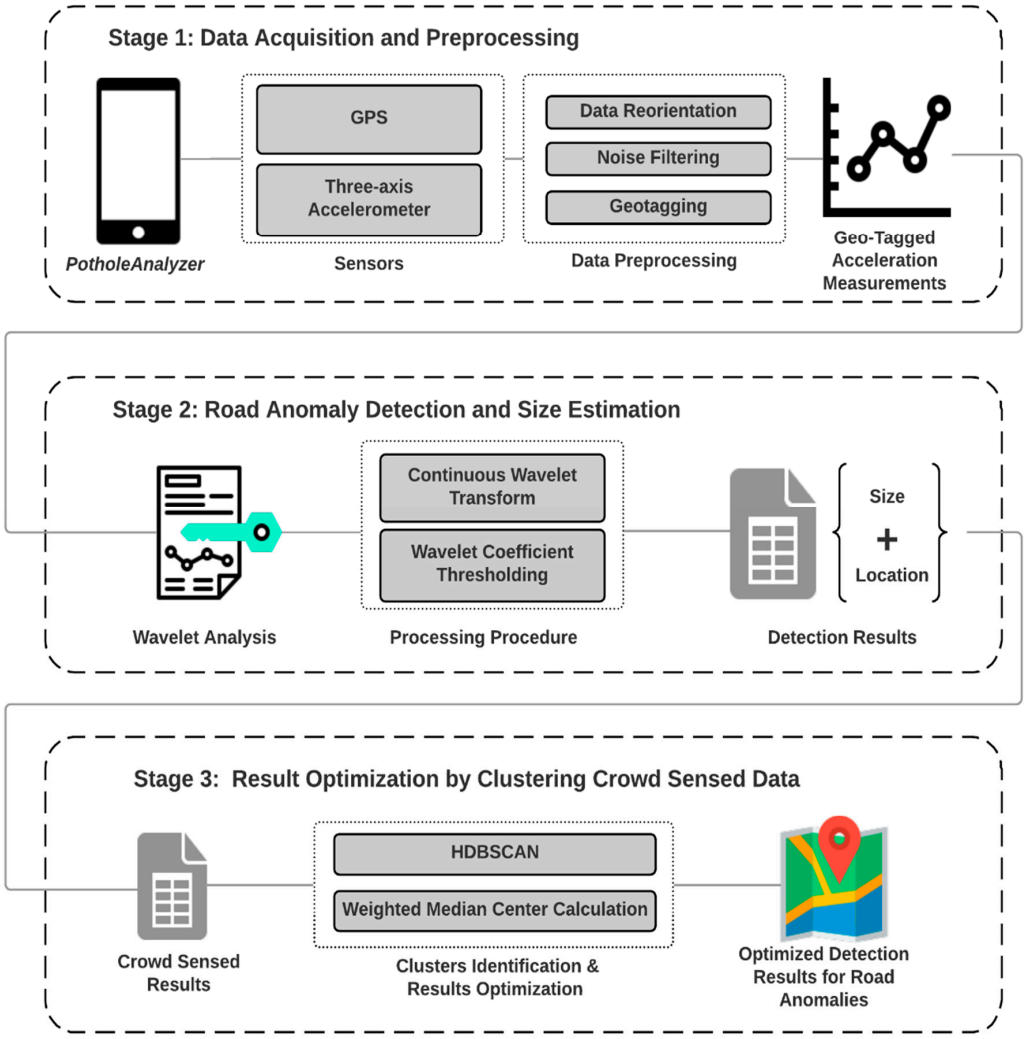

Figure 1. Research workflow.

This section details the data and methods used in each processing stage, respectively. We first create an Android-based mobile app—*PotholeAnalyzor* to acquire research data from two smartphone sensors (e.g., GPS and accelerometer). Next, the mobile collected raw data is preprocessed to clean, transform, and organize datasets before conducting analysis. Then, we make the first attempt to use CWT to analyze mobile sensed signals for identifying road anomalies and estimating their sizes. Finally, the detected bumps and potholes are confirmed and optimized by clustering multiple driving tests' results.

2.1. Data Acquisition and Preprocessing

Former studies have proven that a smartphone accelerometer works well for capturing irregular vehicle vibrations when hitting potholes or bumps [15–17]. By integrating with GPS data, these abnormal acceleration signals can be geotagged, which can aid in identifying and locating road anomalies. Although some studies suggest that the gyroscope can measure smartphone orientation and generate additional features to characterize vehicle motion, this study only utilizes one smartphone motion sensor—accelerometer for two reasons: (1) the accelerometer is the most direct motion sensor measuring vehicle vibrations and proven powerful enough for capturing abnormal signals, (2) utilizing two motion sensors at a high sampling rate (e.g., 100 Hz) can drain the smartphone battery much more faster, which will significantly limits the implementation of the proposed solution.

In this study, we collect data from a smartphone accelerometer and GPS through a customized mobile app. The collected raw accelerometer's data is preprocessed through three steps: data reorientation, data smoothing, and geotagging accelerometer's measurements using GPS data.

2.1.1. Mobile Sensor Data Collection

To obtain the mobile sensors' data, we create a mobile app— *PotholeAnalyzor* using Android application program interfaces (APIs). *PotholeAnalyzor* can record real-time sensed accelerometer measurements, timestamps, and GPS coordinates. Please note that smartphones must be fixed on the vehicle using smartphone holders during data collection, which can avoid some noises caused by devices sliding.

Accelerometer measures both the real acceleration force and earth gravity. To eliminate the influence of earth gravity, Android provides a linear acceleration sensor, which isolates and removes the force of gravity from accelerometer measurements using a low-pass filter and a high-pass filter. Refer to [15,27] for a detailed explanation.

This study analyzes linear accelerometer measurements to detect road anomalies. The sampling rate of the accelerometer is set to 100 Hz while GPS is set to 1 Hz. Figure 2 shows the app's user interface, which contains a dynamic chart showing the z-axis acceleration and a Google Maps visualizer tracking the driving path using GPS.

Figure 2. The user interface of *PotholeAnalyzor*.

2.1.2. Data Reorientation

To ensure the effectiveness of mobile sensed acceleration for capturing vehicle jerks while hitting potholes, data reorientation needs to be implemented to align the accelerometer's axes with the vehicle's axes—x-axis and y-axis of the accelerometer should be used to measure the horizontal movement of the vehicle; z-axis should be perpendicular to the vehicle and senses its vertical vibration, which are directly caused by road anomalies [5]. Euler Angles have been widely proven to be effective for reorienting accelerometers. In this study, we reorient the accelerometer measurements through Euler Angles as follows [14,28]:

$$\alpha = \tan^{-1}\left(\frac{a_y}{a_z}\right), \beta = \tan^{-1}\left(\frac{-a_x}{\sqrt{{a_y}^2 + {a_z}^2}}\right), \tag{1}$$

$$a'_x = \cos\beta \times a_x + \sin\beta \times \sin\alpha \times a_y + \cos\alpha \times \sin\beta \times a_z, \tag{2}$$

$$a'_y = \cos\alpha \times a_y - \sin\alpha \times a_z, \tag{3}$$

$$a'_z = -\sin\beta a_x + \cos\beta \times \sin\alpha \times a_y + \cos\beta \times \cos\alpha \times a_z, \tag{4}$$

where α and β are two Euler Angles, roll and pitch, a_x, a_y, a_z are the raw accelerometer measurements along three axis, and a'_x, a'_y, a'_z are the reoriented three-axis accelerations.

2.1.3. Data Smoothing

Removing data noise is an essential step in signal analysis. Mobile sensed measurements inevitably contain noises. In this study, we implement a high-pass filter to wipe off noises and enhance signal patterns, which is conducted as:

$$y_i = \theta \times y_{i-1} + \theta \times (x_i - x_{i-1}),\ i \in [1, n-1], \tag{5}$$

$$\theta = \frac{t}{t + dT}, \tag{6}$$

where x_i is the ith raw sample data, y_i is the ith smoothed data, t is the current time tag, dT is the event delivery rate, n is the number of samples, which refers to the number of z-axis accelerometer measurements in this study.

Figure 3 shows the comparison between raw data and processed data, which indicates noises can be efficiently eliminated with an enhanced data pattern after filtering.

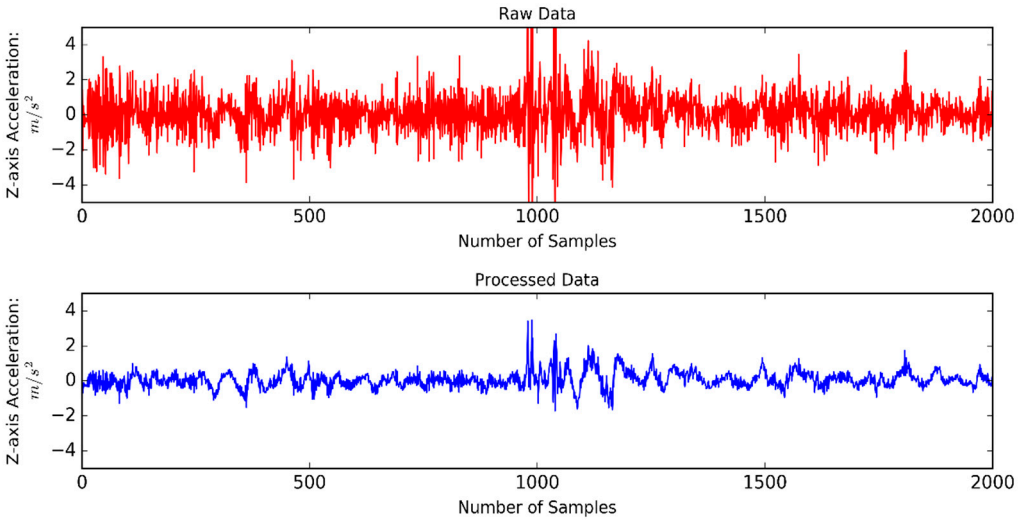

Figure 3. Comparison between raw data and processed data.

2.1.4. Geotagging

The sampling rates of GPS (1 Hz) and accelerometer (100 Hz) are different. To identify the locations of road anomalies, we need to geotag each accelerometer measurement by leveraging GPS readings. In this study, we adopt a scheme proposed in [6] to integrate these two sensors' data. First, the original GPS readings (latitude, longitude, height) are transformed into earth-centered earth-fixed (ECEF) coordinates (x, y, z). Then, we find two temporal-nearest GPS readings for each accelerometer measurement by matching their timestamps. Last, the accelerometer measurement can be geotagged through a linear interpolation scheme based on its temporal distance to its two nearest GPS points.

$$x = x_0 + \frac{(t-t_0)(x_1-x_0)}{(t_1-t_0)},\ y = y_0 + \frac{(t-t_0)(y_1-y_0)}{(t_1-t_0)},\ z = z_0 + \frac{(t-t_0)(z_1-z_0)}{(t_1-t_0)} \tag{7}$$

where (x, y, z) is the calculated ECEF coordinates for the accelerometer measurement with a timestamp t, $(x_0,\ y_0,\ z_0)$ and $(x_1,\ y_1,\ z_1)$ are two consecutive GPS readings with timestamps t_0 and t_1, which are temporally nearest GPS points to the acceleration measurement.

2.2. Road Anomaly Detection and Size Estimation

From a digital signal perspective, each piece of accelerometer recording is a sum of multiple signals with varying frequencies and amplitudes. The amplitude signature of road anomaly is very sensitive to acquisition platform and conditions such as driving speed and the type of vehicle; therefore, amplitude-based detection approaches are often site-specific and unreliable. Frequency-based methods are much more stable because they focus on identifying unique frequency components that are indicative of surface roughness and road anomalies. Fourier analysis and wavelet analysis are the two most popular frequency-based approaches. The use of Fourier analysis in road surface roughness characterization [29,30], however, suffers from a major limitation which is the lack of association between the spatial domain and the frequency domain, such that locating a certain spectral anomaly on the distance profile is difficult with Fourier analysis. Wavelet analysis, on the other hand, is a superior option because it does not only reveal the frequency components of the road profile but also identify where a certain spectral anomaly exists in the spatial domain. Previous applications of wavelet analysis in this field have yielded satisfactory results in road roughness assessment and the detection of surface irregularities, e.g., [23]. In this study, we extend this application and discuss the use of wavelet analysis in pothole detection and pothole size estimation.

2.2.1. Continuous Wavelet Transform

We detect potholes and estimate their sizes by performing the continuous wavelet transform on the preprocessed data. We chose CWT over the discrete wavelet transform (DWT) because CWT results are easier to interpret given that CWT operates at every scale (frequency) and the shifting of the wavelet function is continuous. The one-dimensional CWT is defined as [31]:

$$C(a,\tau) = \int_{-\infty}^{\infty} f(x)\frac{1}{a}\psi^*\left(\frac{x-\tau}{a}\right)dx, \tag{8}$$

where C is the output wavelet coefficient, $f(x)$ is the preprocessed input signal as a function of location x, a is the scale parameter (inversely related to spatial frequency), τ is position parameter and ψ^* is the complex conjugate of the mother-wavelet function that is chosen based on the feature of interest.

In this study, we use order 3 Daubechies wavelet (DB3) as the mother-wavelet (Figure 4) which is recommended by [23]. There is a correspondence between wavelet scales and frequency, such that a smaller scale corresponds to a compressed wavelet, which is high in frequency, while larger scales correspond to a stretched wavelet, representing lower frequency. As defined in Equation (8), a wavelet coefficient is a function of both wavelet scale and position. Scale controls the compression or stretching

of the wavelet and position controls the shifting of the wavelet function. For each scale (corresponding to a certain degree of wavelet compression or stretching), the wavelet examines every location on the input signal by continuously moving along the distance axis. Therefore, the final output is a two-dimensional matrix in scale (frequency)-location space, which is then converted to a matrix of percentage of energy (the sum of all elements in the matrix equals 1).

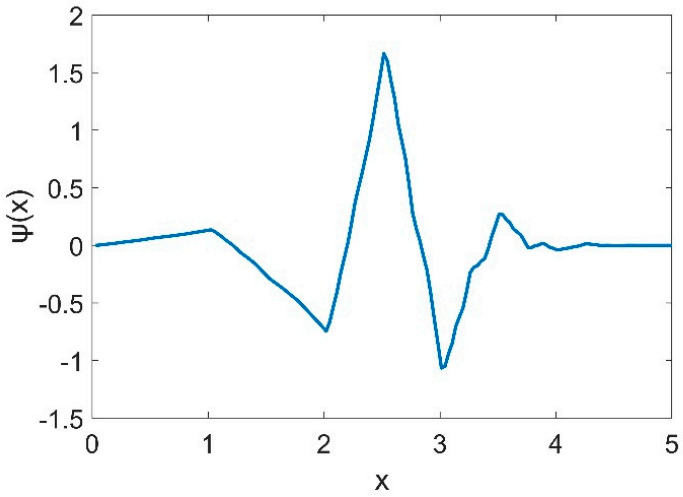

Figure 4. The order 3 Daubechies wavelet (DB3).

CWT produces high wavelet coefficient values at scales where the oscillation in the wavelet correlates best with the signal feature. With a proper choice of mother-wavelet that approximates the target signal (in this case, our target signal is the accelerometer recording when hitting a pothole), the wavelet coefficient image will highlight the target location at the right scale.

2.2.2. Pothole Size Estimation

CWT generates a high value response when the wavelet shifts to a pothole location.

The raw wavelet coefficient images, however, do not come with a meaningful scale that corresponds to pothole size and usually capture irrelevant information such as random road noise and the vibration of the engine. Therefore, we further process the wavelet coefficient images with the following steps:

1. Convert the unitless wavelet scales to physical scales in meters using the algorithm provided by MATLAB Wavelet Toolbox [32].
2. Multiply the scale axis by a scaling factor, which relates the converted wavelet scales to the sizes of target. This scaling factor is determined by field experiments at a test site and is kept as a constant unless the data acquisition platform is changed (in this study, we get a value of 0.3 for generic vehicles including sedan and SUV).
3. Clean the wavelet coefficient images by thresholding (only keep values that are greater than N times of overall average, and in this case, we use N = 18).
4. Apply 2-D Gaussian filter to remove noise and combine detections that correspond to the same pothole. Then the center of each highlighted zone is considered as the center of a detected pothole.
5. Get the size estimation for each detected pothole (highlighted zones on the wavelet coefficient image).

The final result contains two pieces of information: pothole location (step 4) and pothole size (step 5). It is necessary to state that the choice of scaling factor and threshold value may subject to change in other data acquisition settings, because the signals can be influenced by the coupling between road and vehicle. For example, the data acquired by a pickup truck with a large tire and harder suspension may require a different set of processing parameters. Also note that since the mobile

device mainly measures vehicle vibrations along a driving path, we only estimate the maximum driving-dimensional length of road anomalies in this study. Here, the driving-dimension of anomalies is parallel to the road driving direction, as illustrated in Figure 5.

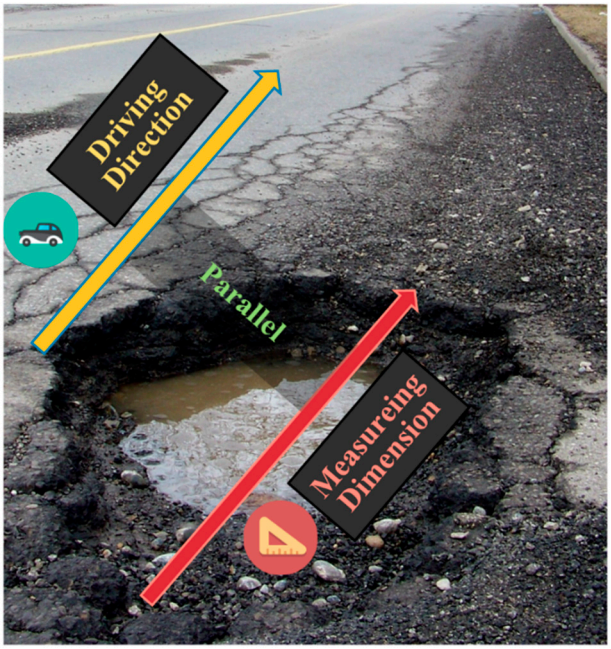

Figure 5. Illustration of the measuring dimension for road anomalies.

2.3. Result Optimiztion by Clustering Crowd Sensed Data

Using smartphone sensors to detect vehicle jerks is a highly efficient solution to identify road anomalies; however, it also has some significant drawbacks. For example, the detection result purely depends on whether the vehicle kicks up road anomalies. However, vehicle wheels only run over a small portion of pavement surface, which significantly limits the detection coverage. Meanwhile, a single user's detection result can be influenced by various factors, such as vehicle models, phone models, driving skills, etc. Therefore, in this study, we implement a crowdsensing solution to optimize the detection results by mining public contributed data. We hypothesize that the significant similarities among crowd sensed data could be used to obtain more reliable detection results than single user's results.

In this study, we innovatively implement spatial clustering methods to group crowd sensed results into clusters based on their similarities. Then, each cluster's member points are further synthesized to form a unique point using weighting schemes, which represents a confirmed road anomaly.

2.3.1. Density-Based Clustering

Density-Based Spatial Clustering of Applications with Noise (DBSCAN) has been extensively utilized to analyze spatial patterns, which can effectively identify concentrated points (clusters) and discrete points (noises) [33,34]. Implementing DBSCAN requires two parameters, including 1) minimum points to form a cluster (C_{min}) and 2) search distance (d) to define neighbors. The clustering procedure can classify data points into three classes, including [34]:

- **Core point**—a point which has at least C_{min} neighbors—points within the d distance to the tested point are counted as its neighbors.
- **Border point**—a point which is counted as a neighbor to core points but does not have its own neighbors (the distance is insufficient, less than C_{min}).
- **Noise point**—a point which is neither a core point nor a border point.

The clustering procedure of DBSCAN contains the following main steps:

1. Choose a random sample point from the dataset as a starting point (p).
2. Identify the neighbors of p using a customized search distance.
3. If p was a core point, it would be marked as visited, a cluster would be formed with the core point and all its connected points. Connected points include p's neighbors and all reachable points (within a d radius) of its neighbors.
4. If p was not a core point, DBSCAN would retrieve an unvisited point from the dataset as a new starting point and repeat the process.
5. The process will end when all points are marked as visited or all points are assigned to a cluster.

Hierarchical DBSCAN (HDBSCAN) is an enhanced density-based clustering method proposed by Campello et al. in 2013 [35]. This method integrates DBSCAN with hierarchical clustering algorithm, which significantly extends the ability of DBSCAN to identify clusters of varying densities. As one of the most data-driven clustering methods, HDBSCAN only has one required parameter C_{min}. One prominent advantage of HDBSCAN is that it can generate probability scores for the sample points. The probability score indicates the likelihood of a point to be involved in a cluster. Refer to [36] for a detailed explanation of HDBSCAN.

In this study, we implement HDBSCAN to group the crowd sensed road anomalies. Each identified cluster is recognized as a unique road anomaly. Meanwhile, this process can also aid in filtering out low-quality public detected results though a simple procedure: points labeled as noises or with low probability scores are eliminated from the clustering result.

2.3.2. Weighting Schemes

After removing the low-quality crowd sensed data, we utilize two weighting schemes to synthesize each cluster's members into one data point. First, we calculate the weighted median center for each cluster to represent the locations of final determined anomalies. The median center is the location which minimizes the distance to all features in a group. The median center is less influenced by outliers than the mean center, which is a more reliable measure of central tendency [37]. Mathematically, the median center needs to satisfy the following objective function [37]:

$$\text{Min} \sum_{i=1}^{n} w_i \sqrt{(x_i - u)^2 + (y_i - v)^2}, \tag{9}$$

where x_i and y_i are coordinates of the ith point, u and v are coordinates of weighted median center, w_i is the weight of the ith point, which refers to the probability score in this study, and n is number of points.

Meanwhile, a weighted average scheme is used to optimize the size estimation result for each cluster.

$$S_{opt} = \frac{\sum_{i=1}^{n} s_i w_i}{\sum_{i=1}^{n} w_i}, \tag{10}$$

where n is number of points in a cluster, s_i the estimated size of the ith point, w_i is the weight of the ith point, which refers to the probability score in this study, and S_{opt} is the recalculated size for each cluster.

Through these two weighting schemes, we can effectively leverage crowd sensed data to obtain an optimized detection result.

3. Experiments and Results

3.1. Experiment Settings

To verify the effectiveness of our method, we manually collected 24 road anomalies as ground truth points from two parking lots at Texas A&M University. These anomalies were positioned through a hand-held GPS—GARMIN GPSMAP 78 with a high positioning accuracy (~3 meters). The 3-meter

positioning accuracy is accurate enough in this study to evaluate the performance of mobile sensed data (with 5 to 10 meters positioning error) and to help road maintainers locate road anomalies. Meanwhile, we carefully measured each pothole's driving-dimensional length using a ruler to form a ground-truth dataset. Figure 6 illustrates the spatial distribution of the obtained ground truth data.

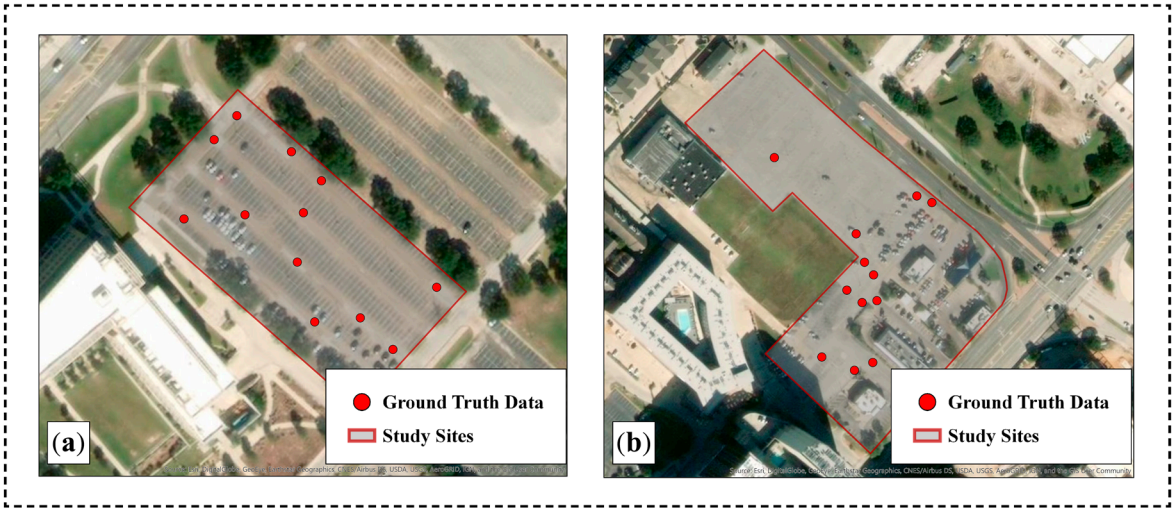

Figure 6. Study sites: (**a**) Parking lot 1; (**b**) Parking lot 2.

Table 1 shows our experiment settings. In this experiment, we tested each parking lot five times by two different drivers, with approximately 30 miles per hour (mph) driving speed. One driver drove a 2009 Toyota Corolla with a Moto X Pure phone running our *PotholeAnalyzor* to detect each parking lot three times. Another driver drove a 2009 Toyota RVA4 with an iPhone 8 running a similar iOS app *CrowdSensor* to detect each parking lot twice. Drivers' explicit permission was required before collecting sensors' data. The sampling rates of accelerometers for both phones were set to 100 Hz. GPS was set to 1 Hz. Through increasing the variability of the experiment (such as drivers, phones, vehicles, etc.), we were able to effectively assess the performance of our method for processing crowd sensed data.

Table 1. Experiment settings.

Experiment Settings	Lot 1	Lot2
Data acquisition time	02/24/2019 2:10 pm	06/01/2019 11:10 am
Road anomalies	12 potholes	8 potholes and 4 bumps
Vehicles models	2009 Toyota Corolla and 2009 Toyota RVA4	2009 Toyota Corolla and 2009 Toyota RVA4
Phone models and apps	Moto X Pure: *PotholeAnalyzor* iPhone 8: *CrowdSense*	Moto X Pure: *PotholeAnalyzor* iPhone 8: *CrowdSense*
Sensors sampling rates	Accelerometer: 100Hz GPS: 1Hz	Accelerometer: 100Hz GPS: 1Hz
Driving tests	2 drivers. Driver 1: test 3 times using Moto X Pure. Driver 2: test 2 times using iPhone 8.	2 drivers. Driver 1: test 3 times using Moto X Pure. Driver 2: test 2 times using iPhone 8.
Ground Truth Acquisition	Manually collected with GARMIN GPSMAP 78 and ruler.	Manually collected with GARMIN GPSMAP 78 and ruler.

3.2. Wavelet Analysis Results

After data collection, we first eliminated the noise of Z-axis acceleration data and geotagged each data point using GPS readings. Then, we analyzed the processed Z-axis acceleration series to identify road anomalies and measure their sizes.

As illustrated in Figure 7, the upper subplot shows the input signals—preprocessed Z-axis acceleration. Then, we performed CWT on the signals to calculate its similarity with mother wavelet at continuous scales, as shown in the middle subplot. The lower subplot shows the filtered high wavelet coefficients, which indicates the high possibility that an anomaly exists with a specific size. The red circles indicate the location and size of ground truth points. The results demonstrated that wavelet analysis can efficiently identify, locate, and measure abnormal signals caused by hitting road anomalies.

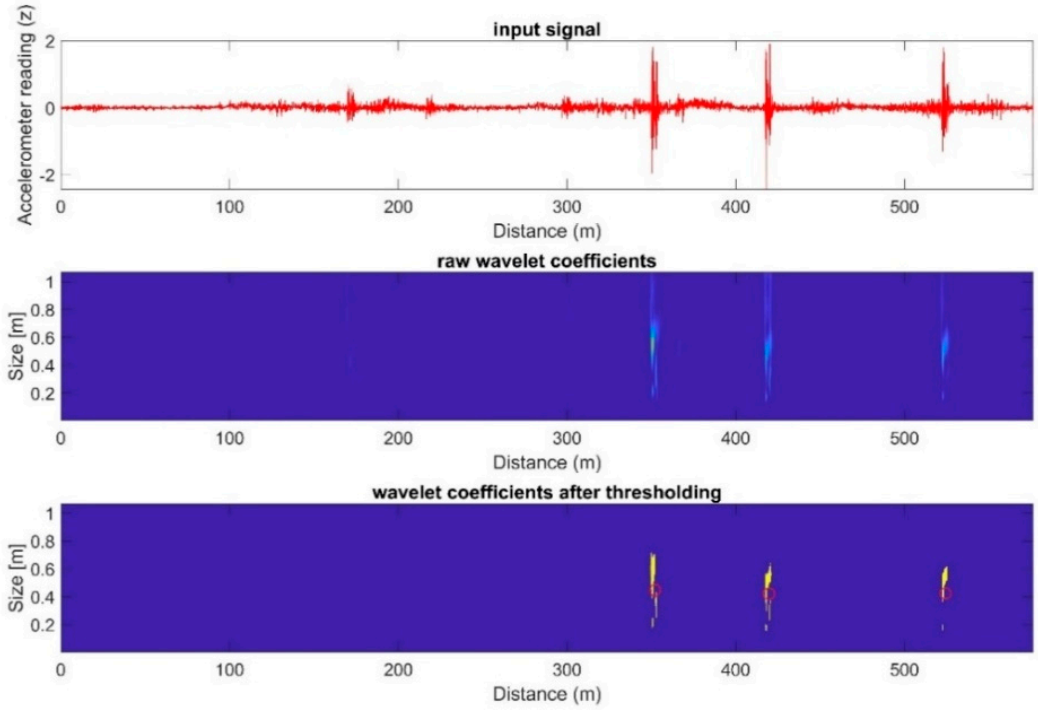

Figure 7. Wavelet analysis results.

Meanwhile, we also further explored the influence of driving speed on the detection result. In this experiment, we tested a road segment from Parking Lot2 three times at different driving speeds (namely, 20 mph, 30 mph, and 40 mph). This road segment contains four bumps with the same size of 0.4 meters. Figure 8 shows the detection results generated from three driving tests. This figure shows that all four bumps can be successfully identified (yellow lines in right-side subplots) from the three driving tests with acceptable size estimation results (~0.25 to 0.5 meters). This indicates that our proposed method achieved a stable performance for detecting road anomalies with different driving speeds. It is also worth noting the detection results (yellow lines) show a positioning difference with the ground truth points (red circles) when driving at 40 mph (bottom-right subplot in Figure 8). This is because the GSP sampling rate is 1 Hz, which is more easily to be influenced by high driving speed. Therefore, we suggest implementing this approach with driving speeds under 40 mph for achieving higher road anomalies positioning accuracy.

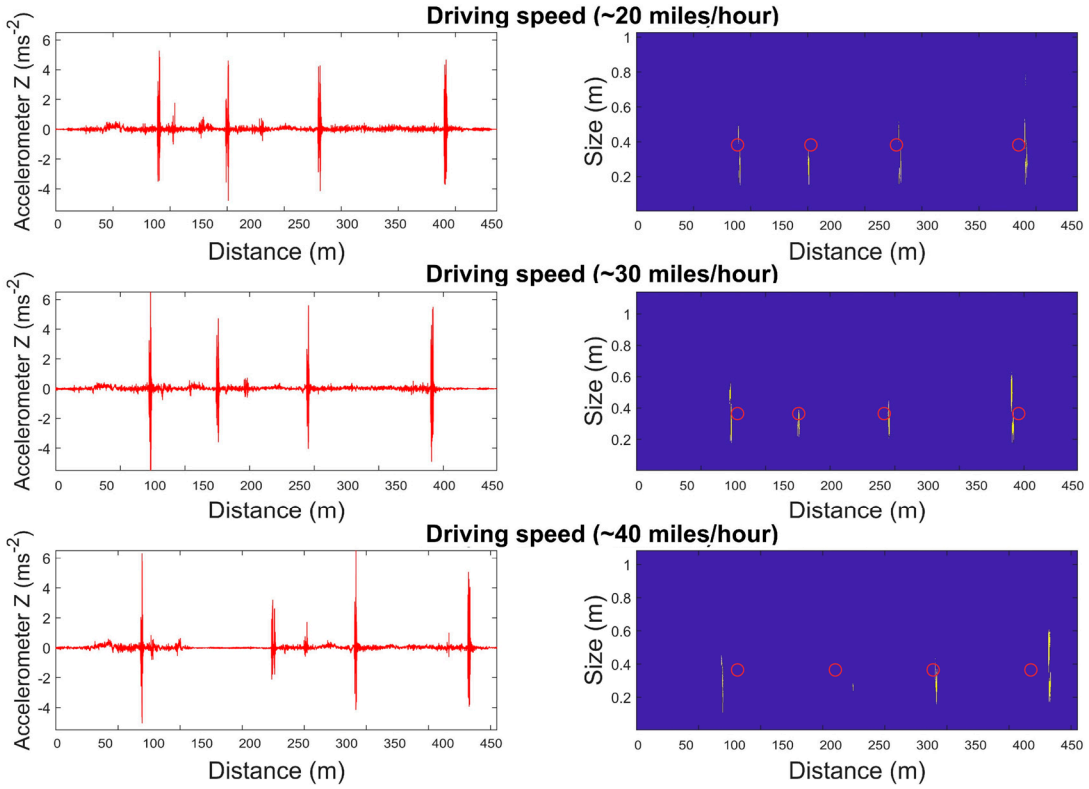

Figure 8. Results generated from three driving tests with different driving speeds.

3.3. Optimized Detection Results by Mining Crowd Sensed Data

After obtaining detection results from each driving test, we implemented HDBSCAN to group the 10 detection results (five for each study site) based on their similarities, which can aid in eliminating low-quality public contributed data and enhancing detection accuracy.

Figure 9a,b illustrate detection results obtained from five driving tests for both study sites. These two subplots show that most of the detected anomalies are concentrated around ground truth points; however, there is still a certain number of (~24% in this study) detected points with a relatively far distance (greater than 10 meters) to ground truth points. It implies that the detection results obtained from one single driving test are not reliable. To optimize our results, we first implemented HDBSCAN on the five times detection results to form clusters. HDBSCAN can automatically group sample points into clusters or noises based on their spatial density patterns. Meanwhile, it also generates a probability score for each point, indicating its likelihood of being involved in a cluster. In this study, clustering noises and cluster member points with low probability scores (less than 0.5) were regarded as low-quality contributed points and eliminated from the detection results. Figure 9c,d show the clustering results for both study sites after eliminating low-quality contributed points. Through this procedure, the points with a large distance to the cluster centers can be successfully removed. Finally, we calculated the weighted median center for each cluster to synthesize multiple contributed points into one point, which represents the optimized location of a detected road anomaly. Figure 9e,f shows that the optimized detection results (yellow dots) can perfectly match with ground truth points (red dots). Meanwhile, we also used a weighted average scheme based on cluster probability scores to recalculate the driving-dimensional size for each final confirmed road anomaly.

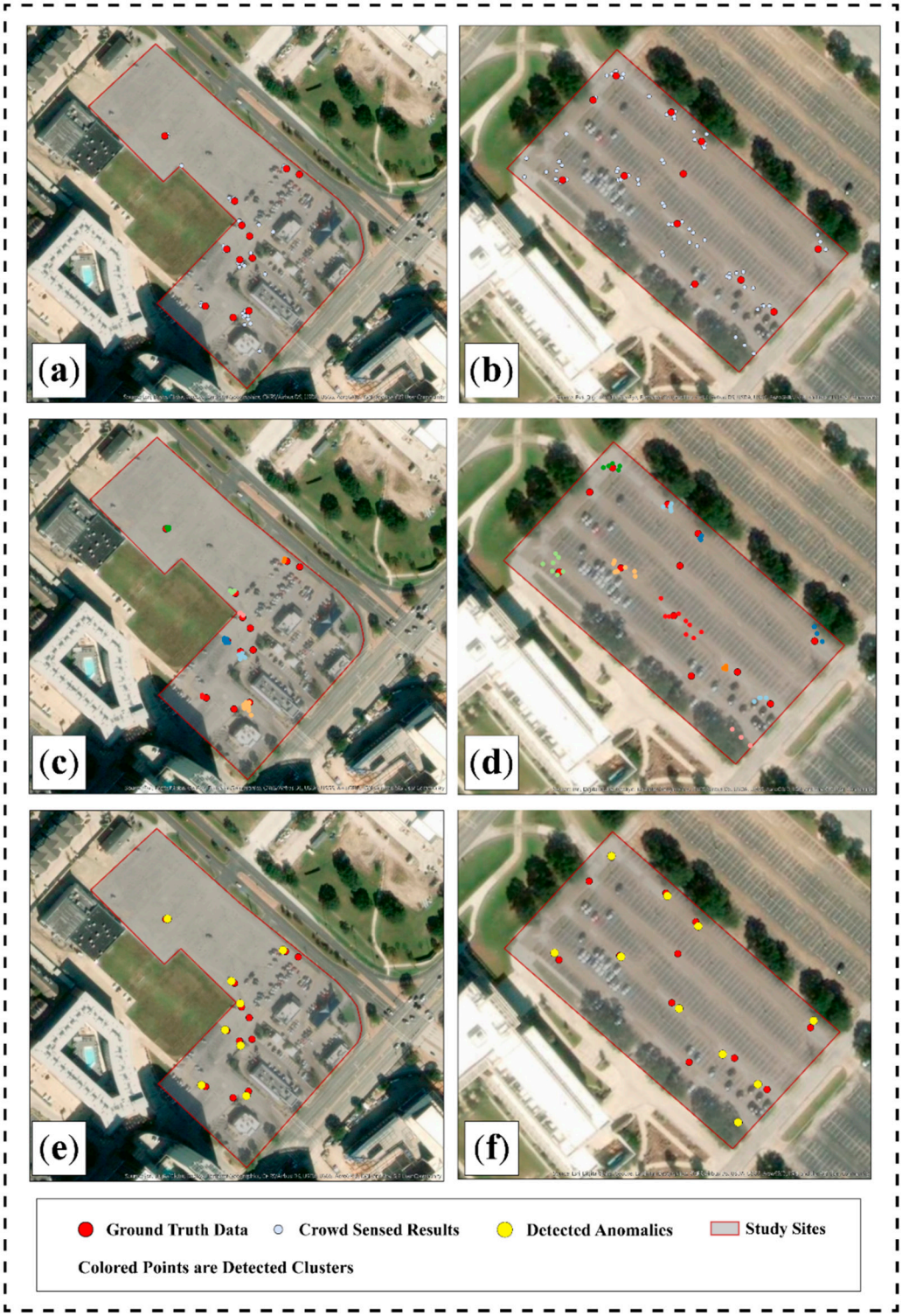

Figure 9. Crowd sensed data integration results: (**a**) and (**b**) are detection results of five driving tests for two study sites; (**c**) and (**d**) show the clustering results after eliminating low-quality contributed points; (**e**) and (**f**) are the optimized detection results by synthesizing each cluster's member points.

3.4. Result Evaluation

To better evaluate the performance of this enhanced crowdsensing solution in road anomaly detection, we compared our method with a widely utilized threshold-based method—Z-THRESH

(Method 1) (Mednis et al. 2011) and a preliminary crowdsensing approach proposed by Li et al. (Method 2).

- Method 1: Z-axis accelerometer measurements exceeding $0.4g$ m/s^2 are counted as road anomalies.
- Method 2: An improved threshold-based detection method integrated with a simple crowdsensing strategy—anomalies need to be reported by more than three users before finally confirmed. The location for the confirmed anomaly is calculated by averaging all the contributed points.

Since Method 1 does not mention how the crowd sensed data was synthesized, we integrated the same crowdsensing strategy used in Method 2 to Method 1 for fusing five driving tests' results. In this study, we compared these two methods with our enhanced solution in terms of detection efficiency and position accuracy.

The detection efficiency is evaluated from three perspectives:

1. Accuracy: Correctly detected anomalies (NCDA)/Total detected anomalies.
2. Coverage Rate: Detected ground truth points (NDGT)/Total ground truth points.
3. Detection Redundancy: (NCDA - NDGT)/(NCDA)

In this experiment, the detected anomalies within a 10-meter radius to any ground truth points are counted as correctly detected anomalies. For each ground truth point, if it can match with any detected anomalies within a 10-meter radius, it would be counted as detected ground truth points. Please note each ground truth point may be matched with more than one detected anomaly; therefore, we also checked detection redundancy for each method.

Meanwhile, we calculated the distance between detected anomalies to their corresponding ground truth points to compare the positioning accuracy while performing different methods.

Table 2 represents the comparison results among these three methods. The results demonstrate that the proposed enhanced crowdsensing solution achieved the highest detected accuracy (94.44%), which is far superior to the other two methods (43.90% and 64.71%). Our approach also achieved the same coverage rate compared to Method 2. Moreover, by applying spatial clustering methods, we can dramatically synthesize crowd sensed points into high-reliable detection results with no redundant detected anomalies and higher positioning accuracy.

Table 2. Performance comparison among three methods.

Criteria	Evaluation Indices	Method 1	Method 2	Our Method
Detection Efficiency	Accuracy	43.90%	64.71%	94.44%
	Coverage Rate	66.67%	70.83%	70.83%
	Detection Redundancy	11.11%	22.22%	00.00%
Positioning Accuracy (meter)	Min	0.60	0.73	0.58
	Mean	3.47	4.07	3.29
	Max	9.88	7.27	6.21
	STEDV	2.58	2.41	1.56

More importantly, this study added a new dimension to road anomaly detections to estimate the driving-dimensional size for each road anomaly. In this study, we used two methods to synthesize the size estimation results of cluster member points into one final result. One is to average all member points' estimation values. Another is to calculate the weighted mean based on the cluster probability scores of each member point. Figure 10 shows the size estimation results by implementing these two methods. The centerline of the box represents the mean value of estimation errors. The box upper and lower bounds represent the mean plus and minus standard deviation, respectively. This figure indicates that our method can effectively estimate the driving-dimensional size for road anomalies with an acceptable detection error. Meanwhile, the weighted mean shows a lower mean error and a

smaller standard deviation in Figure 10. It indicates that using the weighted average scheme can better synthesize crowd-sensed data than calculating the average.

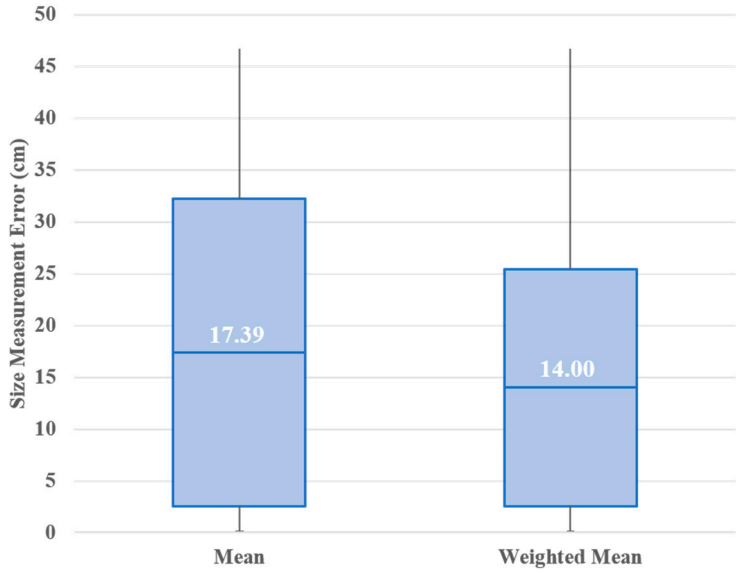

Figure 10. Result evaluation for anomaly size estimation.

4. Discussion and Conclusions

Road anomaly detection is of great importance in road maintenance and management. Continuously monitoring road anomalies with a low-cost and high-efficiency solution is a fundamental social need; however, it remains a complicated and unsolved research task. In this study, we proposed an enhanced mobile sensing approach to detect road anomalies and measure their sizes using smartphone sensors. To the best of our knowledge, this study marks the first attempt to utilize CWT in road anomaly detection. We are also among the first to explore the implementation of spatial clustering methods (HDBSCAN) for synthesizing crowd sensed results.

In this study, a built-in smartphone accelerometer and GPS were first utilized to capture and geotag vehicle vibrations. Next, CWT was adopted to extract and analyze abnormal mobile sensed signals when vehicles are hitting road anomalies. Then, we utilized a spatial clustering method, HDBSCAN, to group different driving tests' detection results into clusters based on their spatial density patterns. Each cluster's member points were finally synthesized into a unique road anomaly.

To verify the effectiveness of the proposed method, we validated it with 24 manually collected road anomalies and compared its performance with a widely utilized threshold-based method, Z-THRESH, and a preliminary crowdsensing approach proposed by Li et al. [15]. Our experiments demonstrated that wavelet analysis outperforms conventional threshold-based methods, which can more effectively identify abnormal vehicle vibrations when hitting road anomalies through analyzing mobile sensed data. Through spatially mining the crowd sensed results, our enhanced mobile sensing solution achieved the highest road anomalies detection accuracy (94.44%) among the three tested methods with a higher positioning accuracy (within 3.29 meters in average). More importantly, our approach could successfully estimate the driving-dimensional size of bumps and potholes based on the calculated wavelet coefficients with an acceptable size estimation error (with a mean error of 14 cm). This could be enormously beneficial for helping local government allocate a road maintenance budget to fix hazardous potholes wisely.

This study demonstrated that the mobile sensing approach is efficient for detecting road anomalies. It also proved the potential and effectiveness of mobile crowdsensing solutions for conducting large-scale sensing and monitoring tasks. Leveraging crowd sensed data could continuously monitor road surface

condition with few additional economic costs, which substantially improves the effectiveness of traditional road monitoring systems.

However, some technical barriers exist, which limit the implementation of crowdsensing solutions at the current stage. For example, mobile crowdsensing is significantly constrained by smartphone hardware. Low-quality mobile sensors' data may lead to unreliable detection results. Collecting mobile sensors' data at a high sampling rate can drain phone battery in several hours, or even faster. To overcome these limitations, a comprehensive crowdsensing-quality-control strategy should be proposed and formalized in future work, which could further eliminate the low-quality crowd sensed data (e.g., data collected using low-quality sensors or devices, data collected while driving at high speed). Meanwhile, we could further optimize the mobile-based analyzing algorithm, reduce computing load, and choose a more appropriate sensor sampling rate instead of using 100 Hz, which may potentially extend the smartphone battery life. Meanwhile, in future work, we will improve the proposed solution from the following ways:

1. **Propose a new anomaly size estimation solution.** In this study, we only estimate the driving-dimensional size of road anomalies. In fact, the depth of potholes is also a critical factor for assessing pothole damages. In future work, we will attempt to measure the depth of road anomalies through analyzing the amplitude of mobile sensed abnormal vibration signals.
2. **Improve the performance of crowdsensing solution.** Using spatial clustering methods can efficiently eliminate low-quality contributed data points and optimize detection results. However, the density-based clustering method may mis-cluster two neighboring potholes into the same group, which could influence the detection accuracy. In future work, we will test different spatial clustering methods, compare their performances, and further form a formalized crowdsensing strategy to synthesize crowd sensed data with further improved accuracy.
3. **Put forward a real-time road anomaly detection system.** Drivers can sense road surface using smartphones at real-time. With a certain number of reliable data contributors, we can potentially update road detection results on a daily, or even hourly basis. In future work, we will attempt to recruit vehicles from local governments (e.g., garbage truck, police vehicles) to put forward a real-time road anomaly monitoring system, which could continuously monitor road surface conditions with high accuracy.

It is worth noting that, to make autonomous vehicles a reality, vehicular sensing techniques are undergoing an unprecedented revolution, which also shows great potential for facilitating the implementation of crowdsensing solutions for assessing road qualities. Nowadays, each commercial vehicle is equipped with approximately 4,000 sensors [38,39]. These sensors empower vehicles to collect thousands of signals through the controller area network (CAN) bus technology, which could monitor the vehicle and its surrounding environment in real-time. These vehicular sensors have a higher sampling rate and a better data quality than that of a smartphone, which facilitates achieving a more precise detection result than smartphone sensors. Meanwhile, light detection and ranging (LiDAR) provides a compelling sensing ability to autonomous vehicles [40,41]. The vehicular LiDAR can simultaneously scan and generate high-resolution 3-D representations of immediate vicinity, which could help us identify road anomalies and bumpy road segments more effectively. Therefore, we believe that the vehicular crowdsensing system could be the next-generation approach for large-scale sensing and monitoring with higher data quality, faster data transmission, and better precision. This proposed solution remains promising and efficient in the foreseeable future.

Author Contributions: Conceptualization, Xiao Li, Da Huo, and Daniel W. Goldberg; Methodology, Xiao Li, Da Huo, and Tracy Hammond; Software, Xiao Li, Tianxing Chu, and Zhengcong Yin; Validation, Xiao Li and Da Huo; Formal Analysis, Xiao Li and Da Huo; Investigation, Xiao Li and Da Huo; Writing-Original Draft Preparation, Xiao Li; Writing-Review & Editing, Xiao Li, Tianxing Chu, Daniel W. Goldberg, and Tracy Hammond; Visualization, Xiao Li and Da Huo; Supervision, Daniel W. Goldberg.

Funding: The open access publishing fees for this article have been covered by the Texas A&M University Open Access to Knowledge Fund (OAKFund), supported by the University Libraries and the Office of the Vice President for Research.

Acknowledgments: We thank Ruizhi Chen for his comments on an earlier version of the manuscript, and Zhe Zhang for suggestions that greatly improved the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bruce, V. Why America Has So Many Potholes. Available online: http://content.time.com/time/magazine/article/0,9171,159579,00.html (accessed on 5 August 2019).
2. Bellis, R.; Osborne, B.; Davis, S.L. *Repair Priorities 2019*; Transportation for America: Washington, DC, USA, 2019.
3. AAA POTHOLE DAMAGE COSTS U.S. DRIVERS $3 BILLION ANNUALLY. Available online: https://www.oregon.aaa.com/2016/02/pothole-damage-costs-u-s-drivers-3-billion-annually/ (accessed on 5 August 2019).
4. The Pothole Facts. Available online: https://www.pothole.info/the-facts/ (accessed on 5 August 2019).
5. Sattar, S.; Li, S.; Chapman, M. Road surface monitoring using smartphone sensors: A review. *Sensors* **2018**, *18*, 3845. [CrossRef] [PubMed]
6. Li, X.; Goldberg, D.W.; Chu, T.; Ma, A. Enhancing driving safety: Discovering individualized hazardous driving scenes using GIS and mobile sensing. *Trans. GIS* **2019**, *23*, 538–557. [CrossRef]
7. Xu, W.; Chen, R.; Chu, T.; Kuang, L.; Yang, Y.; Li, X.; Liu, J.; Chen, Y. A context detection approach using GPS module and emerging sensors in smartphone platform. In Proceedings of the 2014 Ubiquitous Positioning Indoor Navigation and Location Based Service, Corpus Christ, TX, USA, 20–21 November 2014.
8. Chen, R.; Chu, T.; Liu, J.; Chen, Y.; Chen, L.; Xu, W.; Li, X.; Hyyppa, J.; Tang, J. Development of a contextual thinking engine in mobile devices. In Proceedings of the 2014 IEEE Ubiquitous Positioning Indoor Navigation and Location Based Service (UPINLBS), Corpus Christ, TX, USA, 20–21 November 2014; pp. 90–96.
9. Guo, B.; Chen, C.; Zhang, D.; Yu, Z.; Chin, A. Mobile crowd sensing and computing: When participatory sensing meets participatory social media. *IEEE Commun. Mag.* **2016**, *54*, 131–137. [CrossRef]
10. Guo, B.; Yu, Z.; Zhou, X.; Zhang, D. From participatory sensing to Mobile Crowd Sensing. In Proceedings of the 2014 IEEE International Conference on Pervasive Computing and Communication Workshops, PERCOM WORKSHOPS 2014, Budapest, Hungary, 24–28 March 2014.
11. Panichpapiboon, S.; Leakkaw, P. Traffic Density Estimation: A Mobile Sensing Approach. *IEEE Commun. Mag.* **2017**, *55*, 126–131. [CrossRef]
12. Guo, B.; Chen, C.; Yu, Z.; Zhang, D.; Zhou, X. Building human-machine intelligence in mobile crowd sensing. *IT Prof.* **2015**, *17*, 46–52. [CrossRef]
13. Zappatore, M.; Longo, A.; Bochicchio, M.A. Using mobile crowd sensing for noise monitoring in smart cities. In Proceedings of the 2016 International Multidisciplinary Conference on Computer and Energy Science, Split, Croatia, 13–15 July 2016.
14. Astarita, V.; Caruso, M.V.; Danieli, G.; Festa, D.C.; Giofrè, V.P.; Iuele, T.; Vaiana, R. A Mobile Application for Road Surface Quality Control: UNIquALroad. *Procedia-Soc. Behav. Sci.* **2012**, *54*, 1135–1144. [CrossRef]
15. Li, X.; Goldberg, D.W. Toward a mobile crowdsensing system for road surface assessment. *Comput. Environ. Urban Syst.* **2018**, *69*, 51–62. [CrossRef]
16. Li, X.; Chen, R.; Chu, T. A crowdsourcing solution for road surface roughness detection using smartphones. In Proceedings of the 27th International Technical Meeting of the Satellite Division of the Institute of Navigation, ION GNSS 2014, Tampa, FL, USA, 8–12 September 2014.
17. Mednis, A.; Strazdins, G.; Zviedris, R.; Kanonirs, G.; Selavo, L. Real time pothole detection using Android smartphones with accelerometers. In Proceedings of the 2011 International Conference on Distributed Computing in Sensor Systems and Workshops, DCOSS'11, Barcelona, Spain, 27–29 June 2011.
18. Rishiwal, V.; Khan, H. Automatic pothole and speed breaker detection using android system. In Proceedings of the 2016 39th International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2016—Proceedings, Opatija, Croatia, 30 May–3 June 2016.

19. Zang, K.; Shen, J.; Huang, H.; Wan, M.; Shi, J. Assessing and mapping of road surface roughness based on GPS and accelerometer sensors on bicycle-mounted smartphones. *Sensors* **2018**, *18*, 914. [CrossRef] [PubMed]
20. Kalim, F.; Jeong, J.; Ilyas, M.U. CRATER: A Crowd Sensing Application to Estimate Road Conditions. *IEEE Access* **2016**, *4*, 8317–8326. [CrossRef]
21. Celaya-Padilla, J.M.; Galván-Tejada, C.E.; López-Monteagudo, F.E.; Alonso-González, O.; Moreno-Báez, A.; Martínez-Torteya, A.; Galván-Tejada, J.I.; Arceo-Olague, J.G.; Luna-García, H.; Gamboa-Rosales, H. Speed bump detection using accelerometric features: A genetic algorithm approach. *Sensors* **2018**, *18*, 443. [CrossRef] [PubMed]
22. Silva, N.; Shah, V.; Soares, J.; Rodrigues, H. Road anomalies detection system evaluation. *Sensors* **2018**, *18*, 1984. [CrossRef] [PubMed]
23. Wei, L.; Fwa, T.F.; Zhe, Z. Wavelet Analysis and Interpretation of Road Roughness. *J. Transp. Eng.* **2005**, *131*, 120–130. [CrossRef]
24. Bello-Salau, H.; Aibinu, A.M.; Onumanyi, A.J.; Onwuka, E.N.; Dukiya, J.J.; Ohize, H. New road anomaly detection and characterization algorithm for autonomous vehicles. *Appl. Comput. Inform.* **2018**. [CrossRef]
25. Silveira Rodrigues, R.; Pasin, M.; Kozakevicius, A.; Monego, V. Pothole Detection in Asphalt: An Automated Approach to Threshold Computation Based on the Haar Wavelet Transform. In Proceedings of the 2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC), Milwaukee, WI, USA, 15–19 July 2019; pp. 306–315.
26. Sabir, N.; Memon, A.A.; Shaikh, F.K. Threshold Based Efficient Road Monitoring System Using Crowdsourcing Approach. *Wirel. Pers. Commun.* **2019**, *106*, 2407–2425. [CrossRef]
27. Motion Sensors. Available online: https://developer.android.com/guide/topics/sensors/sensors_motion (accessed on 5 August 2019).
28. Johnson, D.A.; Trivedi, M.M. Driving style recognition using a smartphone as a sensor platform. In Proceedings of the IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC, Washington, DC, USA, 5–7 October 2011.
29. Sayers, M.W.; Gillespie, T.D.; Queiroz, C.A.V. *The International Road Roughness Experiment—Establishing Correlation and a Calibration Standard for Measurements*; The World Bank: Washington, DC, USA, 1986; ISBN 0821305891.
30. Hayhoe, G. Spectral Characteristics of Longitudinal Highway Profiles as Related to Ride Quality. In *Vehicle, Tire, Pavement Interface*; ASTM International: West Conshohocken, PA, USA, 2009; pp. 32–53.
31. Daubechies, I. *Ten Lectures on Wavelets*; Siam: Philadelphia, PA, USA, 1992.
32. Misiti, M.; Misiti, Y.; Oppenheim, G.; Poggi, J.-M. *MathWorks Wavelet Toolbox User's Guide*; MathWorks Inc.: Natick, MA, USA, 2015.
33. Esri How Density-Based Clustering Works. Available online: https://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/how-density-based-clustering-works.htm (accessed on 31 July 2019).
34. Ester, M.; Kriegel, H.-P.; Sander, J.; Xu, X. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining, Portland, OR, USA, 2–4 August 1996.
35. Campello, R.J.G.B.; Moulavi, D.; Sander, J. Density-based clustering based on hierarchical density estimates. In Proceedings of the 17th Pacific-Asia Conference on Knowledge Discovery and Data Mining, Gold Coast, Australia, 14–17 April 2013.
36. How HDBSCAN Works. Available online: https://hdbscan.readthedocs.io/en/latest/how_hdbscan_works.html (accessed on 5 August 2019).
37. Lee, J.; Wong, D.W.S. *Point Descriptors. In Statistical Analysis with ArcView GIS*; John Wiley & Sons Inc.: Hoboken, NJ, USA, 2001; pp. 33–58. ISBN 0471348740.
38. Massaro, E.; Ahn, C.; Ratti, C.; Santi, P.; Stahlmann, R.; Lamprecht, A.; Roehder, M.; Huber, M. The Car as an Ambient Sensing Platform. *Proc. IEEE* **2017**, *105*, 3–7. [CrossRef]
39. Fugiglando, U.; Massaro, E.; Santi, P.; Milardo, S.; Abida, K.; Stahlmann, R.; Netter, F.; Ratti, C. Driving Behavior Analysis through CAN Bus Data in an Uncontrolled Environment. *IEEE Trans. Intell. Transp. Syst.* **2019**, *20*, 737–748. [CrossRef]
40. Schwarz, B. Lidar: Mapping the world in 3D. *Nat. Photonics* **2010**, *4*, 429. [CrossRef]

41. Wang, H.; Wang, B.; Liu, B.; Meng, X.; Yang, G. Pedestrian recognition and tracking using 3D LiDAR for autonomous vehicle. *Rob. Auton. Syst.* **2017**, *88*, 71–78. [CrossRef]