A Modified Pothole Detection Approach to Capture their Width

Phan Thanh An⋆ · Nguyen Minh Phuong · Dang Le Quang · Le Hong Trang

Received: date / Accepted: date

Abstract Detecting potholes is an ongoing problem that require extensive investigations as potholes significantly deteriorates urban beauty and can cause serious harm to commuters. Multiple techniques have been developed over the years to address that issue. However, there has been no formal way to define the size, particularly the width, of a pothole. This work presents a mathematical definition for the width of an arbitrary 2D pothole at a point inside. The width of each pothole is calculated by taking the supremum of all width values at all points in a given subset of that pothole region. We also extend the detection pipeline by Koch and Brilakis in 2011 to prove the practicality of these definitions. From our experimental results, this modification produces observable improvements in accuracy and recall and segments potholes with a more accurate size.

Keywords Convex sets · Convex hulls · Definition · Orthogonally convex hulls · Orthogonal polygons · Pothole detection · Quickhull algorithm · Segmentation · Width of potholes

⋆ Corresponding author
P. T. An1,4, N. M. Phuong2,4, D. L. Quang5, L. H. Trang3,4
1 Institute of Mathematical and Computational Sciences and Faculty of Applied Science
2 Faculty of Transportation Engineering
Ho Chi Minh City University of Technology (HCMUT), 268 Ly Thuong Kiet Street, District 10, Ho Chi Minh City, Vietnam
3 Faculty of Computer Science and Engineering
Ho Chi Minh City University of Technology (HCMUT), 268 Ly Thuong Kiet Street, District 10, Ho Chi Minh City, Vietnam
4 Vietnam National University Ho Chi Minh City, Linh Trung Ward, Thu Duc City, Ho Chi Minh City, Vietnam
5 Foreign Trade University, Ho Chi Minh City Campus, Ho Chi Minh City, Vietnam
E-mail: thanhan@hcmut.edu.vn, sci.ngmphuong@gmail.com, quang.dangle@hcmut.edu.vn, lhtrang@hcmut.edu.vn
1 Introduction

Potholes have been an expensive and recurring problem not only in underdeveloped countries but all over the world. While small potholes may only cause minor damage to a vehicle, larger ones can actually cause collisions and severe injuries to drivers, passengers, cyclists, and potentially pedestrians. Moreover, the authorities have to spend a massive amount of money to repair them each year. Therefore, appropriate measures need to be taken in order to solve this issue.

There have been a wide variety of computer vision-based techniques available to detect potholes in road images. They are classified into three main types, namely classical 2D image processing, 3D point cloud modeling and segmentation, and machine learning/deep learning [5]. Researchers have also combined multiple methods from different categories to further improve the performance of their algorithm. Nevertheless, these approaches only aim to find potholes in an image and omit their sizes. These, however, are an important factor that directly affects road maintenance plans. Therefore, it is needed to establish a unified way to calculate the width of a pothole after it has been detected.

Fig. 1: Pothole detection algorithms from 2011 to 2021 [5]
In this work, we propose a definition for a pothole’s width in the plane. To show its applicability, we employ the strategy proposed by Koch and Brilakis [3] and make some adjustments to obtain excelling results in both tasks - detection and segmentation. We modify their classical approach by replacing the elliptical approximation with convex hull and orthogonally convex hull approximation. As a result, the potholes are segmented more precisely, leading to an increase in the overall accuracy. We then test this method on the same dataset to further illustrate the remarkable performance.

This paper is structured into two parts as follows. Section 2 introduces the mathematical aspects of the definition of width of 2D potholes. Section 3 describes the techniques used in our approach and the results from our experiments.

2 Width of potholes

In this section, we present two definitions that provide an initial understanding of how a pothole’s width can be measured mathematically. The first one defines a specific width of a compact region in the plane with respect to an arbitrary point residing in it.

**Definition 1.** Given a compact subset $S$ in the plane and $p \in S$,

$$W_S(p) := \inf \{ \| u - v \| : u, v \in \partial S, p \in [u, v], [u, v] \subset \text{int} S \}$$

is said to be the width of $S$ at $p$.

Observe that if $p \in \partial S$, then we can choose $u = v$ and obtain $W_S(p) = 0$.  

**Definition 2.** Given a set $S$ and a nonempty subset $G \subset S$. The number $\sup_{p \in G} W_S(p)$ is said to be the width of $S$ w.r.t. $G$ and denoted by $W_S(G)$.

The definition of width of a compact set in $\mathbb{R}^m$ and some analytic properties can be seen in [1]. Using the two definitions above, we can compute the width a region of any arbitrary shape knowing its boundary. The two examples below illustrate how the width of a simple geometry, i.e. an ellipse, and of a complex pothole region, is specified.

**Example 1.** For an ellipse as in Fig. 2, the set $S$ contains all points inside the ellipse and is illustrated as the shaded region. The major axis (orange line) of the ellipse is used as the subset $G$. For every point $p$ in $G$, $W_S(p)$ is the length of the line segment that passes through $p$ and is perpendicular to the major axis. The width of $S$ with respect to $G$ from Def. 2 is the length of the perpendicular line segment passing through the midpoint of the major axis, which is also the ellipse’s minor axis.

**Example 2.** For a specific pothole as in Fig. 3, the set $S$ is the inside region of the pothole that is darkened. The subset $G$ in this case is taken as the pothole’s skeleton in orange obtained from morphological thinning. For every point $p$ in
Unlike in the case of an ellipse, $W_S(p)$ is the length of a completely different line segment, as a pothole is of arbitrary shape. From this subset $G$, $\sup_{p \in G} W_S(p)$ is approximated by $\max \{W_S(p_i) : i = 1, 2, \ldots, n\}$. In the following we consider an approximation of $W_S(p)$.

**Proposition 1.** Let $P$ be a pothole region, $p \in P$ and

$$F(n) := \{ \|u_i - v_i\| : u_i, v_i \in \partial S, p \in [u_i, v_i], \quad [u_i, v_i] \subset \text{int} S, i = 0, \ldots, n \}$$

Set $W(n) := \min \{F(n)\}$. Then the sequence $\{W(n)\}$ is convergent and $\lim_{n \to \infty} W(n) \geq W_S(p)$. 
Proof. By the definitions of \( F(n) \) and \( W_S(p) \), we have \( W(n) \geq W(n + 1) \geq W_S(p) \) and therefore the sequence \( \{W(n)\} \) is convergent and \( \lim_{n \to \infty} W(n) \geq W_S(p) \).

3 Detecting potholes using their orthogonally convex hulls

The two definitions stated in the previous section are to be used to detect and measure potholes or cracks.

**Proposition 2.** Let \( P \) be a pothole region, \( O \) be an orthogonally convex hull of \( P \) and \( E \) be an ellipse enclosing the region. Then, \( W_P(G) \leq W_O(G) \leq W_E(G) \).

**Proof.** As \( O \) is an orthogonally convex hull of \( P \), we have \( P \subset O \). Let \( p \in O \) and \( W_P(p) > 0 \). As \( p \in \text{int}\text{ }P \) and \( P \subset O \), we have \( p \in \text{int}\text{ }O \). By Proposition 1 in [1], there are \( u, v \in \partial O \) such that \( W_O(p) = \|u - v\| \) and \( p \in [u,v] \subset \text{int}\text{ }O \).

According to Lemma 1 in [1], there are \( u', v' \in \partial P \) with \( p \in [u',v'] \subset [u,v], [u',v'] \subset \text{int}\text{ }P \). As \( W_P(p) \leq \|u' - v'\| \) and \( \|u' - v'\| \leq \|u - v\| = W_O(p) \) we have \( W_P(p) \leq W_O(p) \leq W_O(G) \). Hence \( W_P(G) \leq W_O(G) \leq W_E(G) \).

As \( E \) is a circumscribed ellipse of the pothole, we have \( P \subset O \subset E \). Using the same proof as above, we can obtain \( W_O(G) \leq W_E(G) \) and therefore \( W_P(G) \leq W_O(G) \leq W_E(G) \).

To prove its applicability, we propose a slight modification to the long-established pothole detection method by Koch and Brilakis [3]. We then compare the widths generated by both approaches to demonstrate that the results are true to the property above.

3.1 Experiment setup

The chosen approach is one of the first works addressing the computer vision problem of detecting potholes [5]. They proposed a two-stage 2D image processing model that can be applied directly to a vehicle’s rear camera. First, frames that may have road defects are extracted from the captured video, then they are processed as in Fig. 4 to show whether they really contain potholes. Our work will mostly explore the second stage.

The original method consists of three main parts: image segmentation, shape extraction, and texture extraction and comparison. In the segmentation task, the image is converted grayscale and applied a median filter to reduce noise. Then a histogram shape-based thresholding algorithm is applied to obtain the new enhanced image \( G_{enh} \), followed by a binary conversion using the Eq. 1 with a threshold \( T \). This results in two possible classifications of the segmented regions, pothole shade and entire pothole.

\[
B(i,j) := \begin{cases} 
1 & \text{if } G_{enh}(i,j) \leq T \\
0 & \text{otherwise}
\end{cases}
\] (1)
In the shape extraction task, miscellaneous components (e.g., cracks, artifacts) are removed. The segments are then classified into two categories using the Eq. 2 and approximated accordingly using elliptical regression.

\[
type(R) := \begin{cases} 
  \text{pothole shade if } & P_{\text{cent}} \in R \lor (\varepsilon > \varepsilon_{\text{max}} \land (l_{\text{max}}/s > r_{\text{max}})) \\
  \text{entire pothole candidate otherwise} 
\end{cases}
\]

(2)

where

- \( l_{\text{max}} \) is the length of the major axis
- \( P_{\text{cent}} \) is the position of the centroid
- \( \varepsilon \) is the region’s eccentricity
- \( s \) is the absolute size of the region
- \( r_{\text{max}} \) is the ratio threshold
In final task, the inside region $R_i$ and the outside region $R_o$ are properly extracted. Based on the assumption that the texture inside a pothole is grainier than the outside, two feature vectors $f_i$ and $f_o$ are computed in Eq. 3 and compared in Eq. 4 to conclude if the candidate is indeed a pothole.

$$f_i := (\text{std}(R_i), \text{std}(LM40(R_i)), \text{std}(LM42(R_i)), \text{std}(LM44(R_i)), \text{std}(S4(R_i)))$$

$$f_o := (\text{std}(R_o), \text{std}(LM40(R_o)), \text{std}(LM42(R_o)), \text{std}(LM44(R_o)), \text{std}(S4(R_o)))$$

$$\text{type}(R_i) := \begin{cases} \text{pothole if } |f_o| < |f_i| \\ \text{no pothole otherwise} \end{cases}$$

Although using ellipses for shape approximation is reasonable, there are still some disadvantages:

- A pothole is not always elliptical. Its shape varies vastly.
- A shade region does not always have the same length as the entire pothole, which means approximating an ellipse from its skeleton can sometimes capture only a smaller area.
- An ellipse containing parts of the outside region or not containing the entire pothole will reduce the accuracy of the feature vector, and potentially the final classification.

For these reasons, we propose replacing the ellipses with convex hulls for approximation. Using convex hulls, the approximated areas fits better to the pothole candidate and produces a better result. The new pipeline is illustrated in Fig. 5.

We use orthogonally convex hulls generated from the $O$-QUICKHULL algorithm [4] to segment full pothole candidates more precisely. However, a pothole shade candidate can only be transformed into a full pothole using convex hulls, as orthogonally convex hulls will only capture the shade region. Then, two feature vectors are computed and compared to determine whether the candidate is indeed a pothole.

3.2 Experimental results

We use Prop. 1 to calculate approximately the width of potholes. To show how this modification helps improve the performance, we test on the same dataset that the original method was built upon. The pothole images were taken from different locations with different severities. This dataset can help to determine the feasibility of this approach in a real-world condition. We then forward the images through the two pipelines to compare the results. For evaluation, we compute the width of each correctly detected pothole using new definitions.

A quantitative comparison is conducted to further see how this approach improves the results. In this comparison, we manually count the number of
Fig. 5: Our proposed pipeline

Table 1: Performances on the original dataset

<table>
<thead>
<tr>
<th></th>
<th>Koch and Brilakis’s method</th>
<th>Our approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total TP</td>
<td>25</td>
<td>28</td>
</tr>
<tr>
<td>Total FP</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Total TN</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Total FN</td>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td>Accuracy</td>
<td>69.23%</td>
<td>76.92%</td>
</tr>
<tr>
<td>Precision</td>
<td>96.15%</td>
<td>93.33%</td>
</tr>
<tr>
<td>Recall</td>
<td>69.44%</td>
<td>80%</td>
</tr>
</tbody>
</table>

true positives (TP, potholes that are correctly detected), false positives (FP, wrongly detected regions), true negatives (TN, correctly classified as background) and false negatives (FN, classified as background but are actually potholes).
holes). While precision represents the detection exactness or fidelity \((\frac{TP}{TP + FP})\), recall measures the detection completeness \((\frac{TP}{TP + FN})\). The accuracy metric indicates the average correctness of the classification process \((\frac{TP + TN}{TP + FP + TN + FN})\). Table 1 shows that this new pipeline outperforms the original one in accuracy and recall. The slight drop in precision is due to one small region incorrectly identified as a pothole. This indicates that the modification has helped the image processing approach capture the potholes better than it originally does.

Among 39 regions in the dataset, there are 25 potholes that both approaches detected correctly. We then compute the width of the ground truth and of the segments outputted from the pipelines for each of these potholes. From Fig 6, it can be seen that the widths of the potholes detected by our approach are closer to the ground truth line, meaning that this modification has created a more precise segmentation model. However, it should be noted that width of the potholes and their orthogonally convex hull in the chart are not always smaller than that of their corresponding ellipses as in Proposition 2. The reason is that this algorithm generates ellipses by fitting a multivariate normal distribution to the data points, which does not strictly enclosing them.

![Fig. 7: Pothole widths comparison](image)

The visual differences are shown in Fig. 7. As can clearly be seen, the convex hulls and orthogonally convex hulls segment the potholes more precisely. In some cases, this leads to a better approximation of the inside and outside regions of each candidate, making the feature vectors capture the texture more easily. The two latter examples show that this technique can extract some more correct information that the original pipeline couldn’t.
4 Conclusion

This work has presented two new definitions to be used to compute a pothole’s width. This will create a way for researchers to better measure and compare the size of a pothole. We have also proposed an effective modification to the pothole detection approach from Koch and Brilakis that reduces unnecessary information from the elliptical regression by replacing it with orthogonally convex hulls. This has resulted in a better approximation of the pothole candidates, thus generated some improved textural information to be processed. The comparison on the pothole images shows that this approach produces superior statistics to the original one.
Conflict of Interests

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

Acknowledgment

This research is funded by Vietnam National University Ho Chi Minh City (VNU-HCM) under grant number T2022-20-01.

The authors acknowledge Ho Chi Minh City University of Technology (HC-MUT), VNU-HCM for supporting this study. The authors also want to thank Prof. Dr.-Ing. Christian Koch and his research group for providing the data and the MATLAB code.

References