EVENT DETECTION IN THE TERRAIN SURFACE

Weixiao Dong

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science
In
Mechanical Engineering

John B. Ferris
Robert L. West
Scott L. Hendricks

April 29, 2016
Blacksburg, VA

Keywords: Event Detection, Edge Detection, Edge Stitching, Comparative Nodal Uncertainty Method, Curved Regular Grid

Event Detection in the Terrain Surface

Weixiao Dong

ABSTRACT

Event Detection is a process of identifying terrain flatness from which localized events such as potholes in the terrain surface can be found and is an important tool in pavement health monitoring and vehicle performance inspection. Repeated detection of terrain surfaces over an extended period of time can be used by highway engineers for long term road health monitoring. An accurate terrain map can allow maintenance personnel for identifying deterioration in road surface for immediate correction. Additionally, knowledge of the events in terrain surface can be used to predict the performance the vehicles would experience while traveling over it.

Event detection is composed of two processes: event edging and stitching edges to events. Edge detection is a process of identifying significant localized changes in the terrain surface. Many edge detection methods have been designed capable of capturing edges in terrain surfaces. Gradient searches are frequently used in image processing to recover useful information from images. The issue with using a gradient search method is that it returns deterministic values resulting in edges which are less precise. In order to predict the precision of the terrain surface, the individual nodal probability densities must be quantified and finally combined for the precision of terrain surface. A Comparative Nodal Uncertainty Method is developed in this work to detect edges based on the probability distribution of the nodal heights within some local neighborhood. Edge stitching is developed to group edges to events in a correct sequence from which an event can be determined finally.

GENERAL AUDIENCE ABSTRACT

The main topic in this work is a process of identifying terrain flatness from which localized events such as potholes in the terrain surface can be found. This process is called Event Detection and it is an important tool in pavement health monitoring and vehicle performance inspection. The knowledge of terrain shape and characterization has uses in many different areas of road surface research. Determining road surface deterioration has values to vehicle engineers, pavement engineers and agricultural engineers since terrain serves as the primary input to any land transportation tool. The accurate surface texture provides the ability to keep track of the road condition and provide maintenance crews information to make an immediate repairs.

In this work, Event Detection is composed of event edging and stitching edges to events. Event edging is a process of identifying the significant changes in the terrain surface. Edge stitching is a process of connecting those changes detected in the process of event edging in a correct sequence in order to determine the final event in the terrain surface. Both processes use mathematical methods and statistical methods to deal with the data obtained from the terrain measurement system. Comparing the height of each point on the road surface can find which point is on the high place and which point is on the low place. Then the edge points which are located between the points which are on the high place and the points which are on the low place can be detected out. The advantage of this novel Event Detection method compared to traditional Event Detection methods is that this method compares the height distribution of each point on the road surface instead of comparing two deterministic height values used in traditional Event Detection method. This novel Event Detection method gives a more accurate result in the end than traditional Event Detection methods due to the use of nearby information of each point on the road surface.

ACKNOWLEDGEMENTS

First and foremost, I would like to sincerely thank my advisor, Dr. John Ferris for the source of professional guidance and support throughout this academic journey. Thank you also for your consistent encouragement which increases me much confidence on my research. I would also like to thank my committee members, Dr. Hendricks and Dr. West for always being considerable and taking time out of their busy schedules to deal with the issues of completing degree.

Thank you to Virginia tech for providing the academic environment and unparalleled facilities in which I was able to perform this research. I would also like to thank the research sponsors Chrysler who made this research possible. Special thanks goes to Dr. Tjhung for gladly giving advices in the bi-weekly Chrysler meeting.

I would also like to thank all my colleges in Vehicle Terrain Performance Laboratory I have worked with over the past couple of years. It is a pleasure to work with all of you. More thanks to my friends, whose support and encouragement always cheer me up.

Lastly but surely not least, I would like to express my sincerest gratefulness to my parents for their steadfast support and understanding throughout my graduate life.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

This thesis is focused on the development of event detection in the terrain surfaces. Event Detection is a process of identifying terrain flatness from which localized events such as potholes in the terrain surface can be found. A Comparative Nodal Uncertainty Edge Detection Method and Edge Stitching that have been developed over the course of this work have been capable of providing edges which are relatively clearer than results found in those existing gradients methods.

Comparative Nodal Uncertainty Method is developed to detect edges based on the probability distribution of the nodal heights within some local neighborhood. Expected Distance Method and Number Difference Method are two methods which have been developed in this work. The output of these two methods is respectively an Expected Distance matrix and a Number Difference Matrix. The entries of Expected Distance matrix are expected to be small for the points on the edge or the entries of Number Difference matrix are expected to be large for the points on the edge. All edge points are obtained from Expected Distance matrix or Number Difference matrix through a threshold.

Edge Stitching is a necessary step after finding all edge points which groups edge points to events. Edge Stitching uses the result of Expected Distance matrix or Number Difference matrix from Comparative Nodal Uncertainty Edge Detection Method and develops a search algorithm in order to reach the goal of connecting edge nodes to events sequentially.

In the formation of this event detection method, many fundamental principles of mathematics and statistics are brought together with the intention of detecting road edges quickly and accurately. The accuracy of this method depends on the accuracy of the input data measured from terrain surfaces, the proper selection of threshold in both Expected Distance matrix and Number Difference matrix from Comparative Nodal Uncertainty Method, and the efficiency of the Edge Stitching.

1.1 Motivation

The knowledge of terrain shape and characterization has uses in many different areas of road surface research. Determining road surface deterioration has values to vehicle engineers, pavement engineers and agricultural engineers since terrain serves as the primary input to any land transportation tool. The accurate surface texture provides the ability to keep track of the road condition and provide maintenance crews information to make an immediate repairs. Additionally, the surface condition gives land vehicle drivers signals of path ahead to take an immediate action of maintenance for driving stability. Therefore, it is beneficial for multiple parties to have some method in place to record these terrain characteristics. For example, this is helpful in vehicle deign since it can determine the characteristic of vehicle through the simulation in different road flatness conditions in order to improve vehicle durability.

Many measurement systems are available today to measure terrain surface with high accuracy. Ability to extract useful information from this vast amount of data is important towards analyzing those characteristics on road surfaces. Some techniques such as Curved Regular Grid (CRG) have been developed to convert rough data to regular data providing convenience to analyze the data from terrain measurement system [1]. Edge detection has been developed based on this technology. With the availability of existing measurement systems and techniques which prepares data to be usable, event detection method is capable of being verified.

Edge detection is primarily focused on detecting edges in the terrain surfaces and Edge Stitching aims to complete event detection by connecting edge points to events. In order to obtain clear edges reflecting the real road condition, all possible edge points need to be found and connected in a correct sequence for the sake of event independency.

Taking all of these considerations into account, an event detection method must include an edge detection method which accurately detects out all possible edges and an edge stitching method which is capable of connecting all edges points to events sequentially.

1.2 Problem Statement

Historically, different edge detection methods have been used to locate edges in images. Traditional edge detection methods use image processing methods (Prewitt, Sobel etc.) to emphasize regions that correspond to edges. Typically, they are used to find the approximate absolute gradient magnitude at each point in an input grayscale image. These methods are capable of processing data and locating the edge fast but they only use the deterministic gradient values for each node obtained from comparing nodal heights. The will result in edges which are less precise.

In order to reach a higher accuracy of data processing, a novel edge detection method Comparative Nodal Uncertainty Edge Detection Method is developed which is based on the probabilistic distribution of a groups of points. The output of Comparative Nodal Uncertainty Edge Detection Method is an Expected Distance matrix whose entries are obtained from the comparison of two groups of neighbor points, instead of only focusing on one pixel in traditional gradient methods. This explains the higher accuracy of data processing.

Edge Stitching which groups edges to an event needs to be taken into consideration when edges in terrain surfaces has been detected out. Although many edge detection methods exist, there are no edge stitching methods that exists to locate the event by connecting the edges to events sequentially.

A difficulty of edge stitching is the uncertainty of searching direction of connecting edge points. Therefore, an Edge Stitching Method is needed to successfully search the correct direction of the path for connecting edge points. In this work, an Edge Stitching method is developed based on the results of Comparative Nodal Uncertainty Method, the Expected Distance matrix to group edge points to events sequentially.

1.3 Thesis Statement and Scope of Work

Thesis Statement: The localized events in the terrain surfaces can be determined by Comparative Nodal Uncertainty Edge Detection Method which detects out all edge points and Edge Stitching Method which is capable of connecting edges to events.

The objective of this work is to describe the principles and algorithm used in the design and development of Comparative Nodal Uncertainty Edge Detection Method and Edge Stitching Method and how it compares with traditional gradient methods. A proof-of-concept of this thesis is presented primarily from the Vehicle Terrain Performance Lab (VTPL). Examples will be developed in order to illustrate those concepts in this thesis. The scope of this work will be restricted to verify the feasibility of Comparative Nodal Uncertainty Edge Detection Method by comparing the results of examples with other gradient methods, and the feasibility of Edge Stitching Method. The final result of this work will be an illustration of how the Comparative Nodal Uncertainty Edge Detection Method and Edge Stitching Method works in detecting events in road surfaces.

1.4 Main Contributions

The main focus of this research is to explore a novel event detection method with a higher accuracy of detecting localized events by using an existing data set measured from a terrain measurement system and converted to regular spaced data by a Curved Regular Grid technique. The following are the original contributions to the state-of-art that are developed in this work:

1.4.1 A novel Edge Detection Method

Comparative Nodal Uncertainty Edge Detection Method is a novel edge detection method developed to detect edges based on the probability distribution of the nodal heights within some local neighborhood. There are two branches of this edge detection method which are Expected Distance method and Number Difference method and each of them have their own advantages and applicability. All edge points are obtained from Expected Distance matrix or Number Difference matrix through a threshold. Through simulations provided the Comparative Nodal Uncertainty Edge Detection Method is at least as well as gradients methods which basically use the deterministic value of gradients of nodal heights to detect edges. Two matrices Expected Distance matrix and Number Difference matrix are obtained can both be used in Edge Stitching.

1.4.2 Edge Stitching

Edge Stitching is a necessary step after finding all edge points. This method is used to group edge points to events. Edge Stitching uses the result of Expected Distance matrix or Number Difference matrix from Comparative Nodal Uncertainty Edge Detection Method. Several algorithms for searching right path of edges are developed in order to reach the goal of connecting edge nodes to events sequentially and properly. Two filtering methods with applicability and limitations are developed to filter self-defined outliers in the result of Comparative Nodal Uncertainty Edge Detection Method. Two methods connecting filtered edge points to events are following developed in the following to connect all possible edge points to events by filling gaps between edge points. Through simulations provided the developed Edge Stitching Method can connect edge points to events properly.

1.5 Thesis Outline

This work will be organized in the following manner. The first chapter contains an overview of the purpose of event detection as well as how this novel event detection method contributes into the techniques and works already completed on event detection. The aim of this thesis is to show the algorithm of Comparative Nodal Uncertainty Method as well as the Edge Stitching Method. The second chapter contains a review of the background literature on the topics of Curved Regular Grid (CRG). Chapter three provides a background of existing edge detection methods gradient methods, the Nodal Uncertainty Method and Rank-Sum Test that will be used in the next section, and the development of Comparative Nodal Uncertainty Edge Detection Method. Example section and Discussion section are provided in the end of this chapter for comparing the results of using different edge detection methods. Chapter four provides a background of existing image stitching methods, the Fourier series and Least Square Curve Fitting methods which will be used in the next section, the development of Edge Stitching Method which complete the process of event detection by grouping edge points to events. Furthermore, examples will be carried out for explaining each filtering edges method and connecting edge points to events method. A Discussion section is provided to discuss the advantages and disadvantages of each filtering method and connecting edge-to-event method. Finally, in chapter five, a summary of the research and final conclusions will be given. This thesis is focused on the development of techniques for detecting localized events in terrain surfaces.

2. Background

This chapter provides a background of a technique that converts the irregularly spaced data into regular spacing. With the availability of existing measurement systems and this technique which prepares data to be usable, event detection method is capable of being verified by this data.

2.1 Curved Regular Grid

Many systems exist that accurately measure road surfaces [1]. From the laser scanner of a terrain measurement system a "point cloud" of terrain height data at irregular locations in the horizontal plane, shown as blue dots in Figure 2-1, is formed. The global coordinate system is defined by black axes X and Y, Easting and Northing respectively. Due to the large size of terrain profile data collection and computational power limitations, these irregularly spaced data sets are not typically used directly in the simulation process. Therefore, a compact and organized terrain surface and terrain profile is needed for simulation.

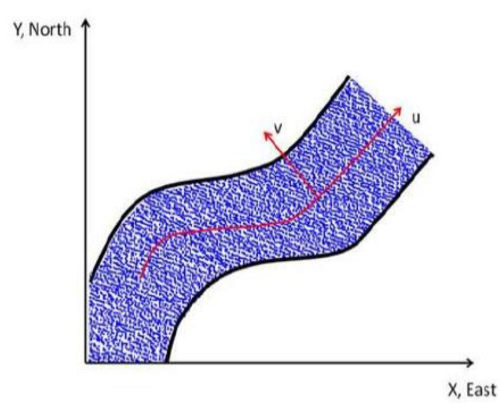


Figure 2-1 Irregular data obtained by terrain measurement system

In Figure 2-1 a path coordinate system is defined by red axes u and v. Where u is located in the direction along the road surface and v lies perpendicularly to u. A stochastic gridding method known as Curved Regular Grid (CRG) [1] is developed for the conversion from the cloud data (rough, non-regular data) into the regularly spaced grid for the purpose of storage and computational simulation. The terrain heights are estimated at each regularly spaced grid point in the (u, v) coordinate system. In Figure 2-2 the intersection between the horizontal and vertical lines indicates the grid node.

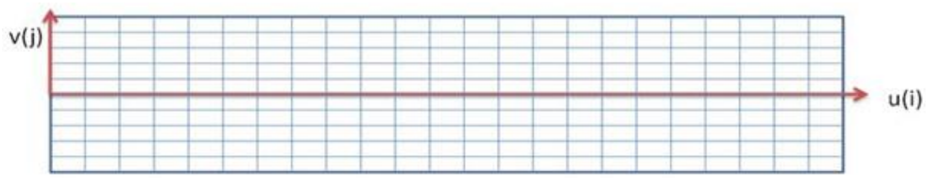


Figure 2-2 Small curve regular gird example

The single nodal height at each intersection is obtained by applying an interpolation algorithm to the data points enclosed inside a search area where all points heights are obtained, as shown in **Error! Reference source not found.**. The crosses indicate data points and the red circle indicates the search radius within which the data points are evaluated. Therefore, this method enables the conversion from non-regularly rough data, from the terrain measurement system, to regularly gridded data. Furthermore, if a single point estimate is required as a measure of location for the height, then an average measure (e.g., a trimmed mean) is calculated so that a deterministic surface (u, v, Z) is defined [1].

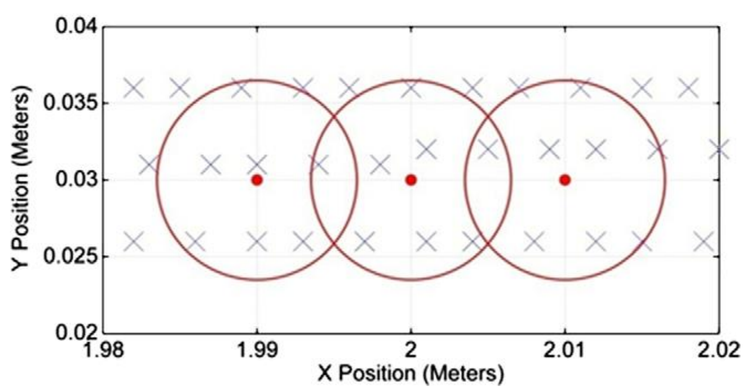


Figure 2-3 Regularly spaced grid example

Each grid node forms a corresponding cumulative probability distribution, F(Z), where Z is a random variable in the estimate of each grid node, shown in Figure 2-4.

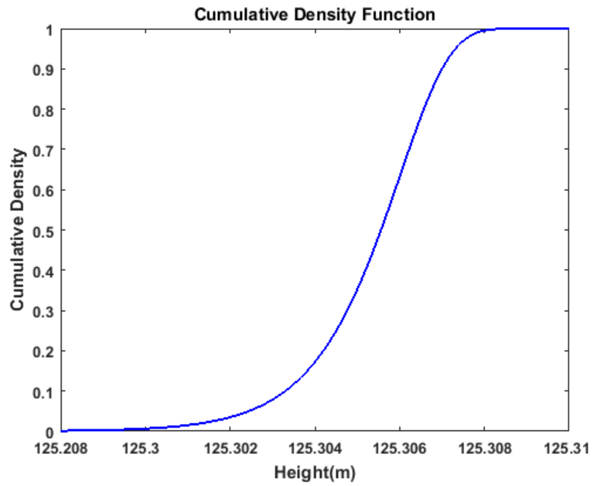


Figure 2-4 Cumulative probability function for typical grid node

A compact and organized terrain surface and profile are obtained for simulation. Specifically, each grid node in Figure 2-2 is represented by nine corresponding heights at cumulative percentages from 10 to 90. Thus, each nodal height is not a deterministic value but nine values extracted from its own cumulative density function at specific percentages. These nine nodal heights will be used in Chapter 3.

3. Edge Detection

3.1 Introduction

Comparative Nodal Uncertainty Edge Detection Method is developed to detect edges based on the probability distribution of the nodal heights within some local neighborhood. The output of one method the Expected Distance Method is an expected distance matrix where entries are calculated by a Two-Sample Wilcoxon Sum-Rank Test are expected to be small for the points on the edge. The output of another method the Number Difference Method is a Number Difference matrix where entries with a large number indicates they are on the edge. All edge points are obtained from Expected Distance matrix or Number Difference matrix through a threshold. This method is novel because it is comparing the similarity of two probability distributions instead of comparing two deterministic values.

In this Chapter, Section 3.2 provides a background of existing edge detection methods. Section 3.3 develops how to obtain the output: Expected Distance matrix or the Number Difference matrix. Section 3.4 provides some examples verifying the Comparative Nodal Uncertainty Edge Detection Method. In the end, the results from the Comparative Nodal Uncertainty Edge Detection Method examples and previous methods will be displayed in Section 3.4. Discussion of the advantages and disadvantages, the scope of applicability, the justification and implication of choosing thresholds will be provided in Section 3.5. Finally, in Section 3.6, a summary of contribution of Comparative Nodal Uncertainty Edge Detection Method will be given.

3.2 Background

Any image stores a collection of pixels or points forming the picture elements [3] and images contain a great deal of information [4]. Image processing is any form of information processing where the input is an image, and the output is a set of characteristics of the image. Edge Detection, as a fundamental tool in image processing, machine vision, and computer vision, provides a way to extract useful information from an image including a vast amount of pixels or points [14]. This section gives a review of some existing Edge Detection methods. In Section 3.4

provides a few examples using these image processing methods whose data input are based upon the technique Curved Regular Grid (CRG).

In this work, a data file including all grid nodes is created as an image file. The data points contained in the grid are used to determine the edges. A major variation in picture quality can be created through the detected areas with strong intensity contrasts from one pixel to the next. However, instead of using changes in contrasts in the picture, the edge detection functions use a change in the gradients of the heights. Gradients practically refer to the magnitude of the gradient or its approximation, whose values can be related to the degree of intensity change in areas of variable intensity **Error! Reference source not found.**]. The first-order derivative in both horizontal and vertical directions at each gridded node has been used as a 2D function $f(\mathrm{x}, \mathrm{y})$, whose vector and magnitude are provided in Equation 3-1 and Equation 3-2.

$$\nabla f = \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \qquad \text{Equation 3-1}$$

$$\begin{aligned} \nabla f &= mag(\nabla f) = \left[{g_x}^2 + {g_y}^2\right]^{\frac{1}{2}} \\ &= \left[\left({}^{\partial f}/_{\partial x}\right)^2 + \left({}^{\partial f}/_{\partial y}\right)^2\right]^{\frac{1}{2}} \end{aligned} \qquad \text{Equation 3-2}$$

The quantity can be approximated sometimes by removing the square root operation, shown as Equation 3-3, or by applying absolute values shown as Equation 3-4

$$\nabla f \approx {g_x}^2 + {g_y}^2 \qquad \text{Equation 3-3}$$

$$\nabla f \approx |g_x^2| + |g_y^2| \qquad \text{Equation 3-4}$$

Traditional edge detection methods (Prewitt, Sobel etc.) are used to identify edges and extract information from the image by using gradients. An example of a 3×3 Matrix with different z value corresponding to each different node height is provided in Figure 3-1. Many gradients edge detection methods introduced in this chapter use the 3×3 Matrix by assigning values through its own mask.

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

Figure 3-1 The 3×3 Matrix of data points

3.2.1 Sobel

The Sobel method detects the edges by searching for the maximum of first derivative which is located on the edge. The Sobel operator is a discrete differentiation operator by computing an approximation of the gradient of the image intensity function [3, 17]. The operator consists of a pair of 3×3 convolution masks as shown in Figure 3-2. G_x and G_y are used by the operator as the gradients matrices. One mask is simply the other rotated by 90° **Error! Reference source not found.**].

-1	0	+1
-2	0	+2
-1	0	+1

Gx

+1	+2	+1
0	0	0
-1	-2	-1

Gy

Figure 3-2 The 3 × 3 Sobel mask

3.2.2 Prewitt

The Prewitt edge detector is similar to Sobel method. The difference bettwen Sobel edge detector and Prewitt edge detector is that Prewitt edge detector does not emphasize pixels that are closer to the center of the masks [5, 16]. Figure 3-3**Error! Reference source not found.** shows the mask applied by the Prewitt edge operator.

-1	0	+1
-1	0	+1
-1	0	+1

Gx

+1	+1	+1
0	0	0
-1	-1	-1

Gy

Figure 3-3 The 3 × 3 Prewitt mask

3.2.3 Roberts Cross Method

The Roberts Cross method [6, 18] provides a similar way with the Prewitt and the Sobel methods where use the derivative to determine the edges. Instead of using 3 × 3 Matrix mask, Roberts Cross method applied a basic 2 × 2 mask to calculate the gradient in the direction of x and y. Figure 3-4 indicates the simple 2 × 2 mask used in edge detection. Unfortunately, one disadvantages of the Robert Cross method is that it suffers from the sensitivity due to the noise [19].

+1	0
0	-1

Gx

0	+1
-1	0

Gy

Figure 3-4 The 2 × 2 Roberts Cross mask

3.2.4 Canny

The Canny edge detector smoothes the image with a Gaussian filter by computing the gradient magnitude and direction using finite-difference approximations for the first derivatives [7, 15]. The Canny method, without using a mask over the grid nodes, locates the edges which are marked as the maxima in gradient magnitude. Thresholding values identifed by the Gaussian filter value, σ, aim to detect and link edges, and remove signal noise as well.

3.2.5 Laplacian of Guassian (LoG)

LoG is combined with a Laplace filter providing a discrete approximation to a mathematical Laplace operator **Error! Reference source not found.Error! Reference source not found.Error! Reference source not found.**] shown in Equation 3-5.

$$\nabla^2 f(x,y) = \frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2} \qquad \text{Equation 3-5}$$

LoG detector detects edges based on the variable intensity of the pixels by testing a wider area around a pixel. LoG masks can have window with different size based on different pixels to which different weights are assigned. An example of 5 × 5 mask is provided in Figure 3-5**Error! Reference source not found.**.

0	0	-1	0	0
0	-1	-2	-1	0
-1	-2	16	-2	-1
0	-1	-2	-1	0
0	0	-1	0	0

Figure 3-5 The 5 × 5 Laplacian of Gaussian mask

3.2.6 Zero-Cross Method

The zero-cross edge operator **Error! Reference source not found.**] is on the basis of a similar mask of LoG method by searching for zero crossings of the second derivative. Equation 3-6 shows the second derivative function $f_\alpha''(\mathrm{r}, \mathrm{c})$ for a pixel (r, c) in the direction of α where r and c represent the row and the column of the selected window. The identified location of zero-crossing can determine the pixels related to the edges.

$$f_\alpha'' = \frac{\partial^2 f}{\partial r^2}\sin^2\alpha + \frac{2\partial^2 f}{\partial rc}\sin\alpha\cos\alpha + \frac{\partial^2 f}{\partial c^2}\cos^2\alpha \qquad \text{Equation 3-6}$$

All edge detection methods listed from 3.2.1 to 3.2.6 work very well in terms of a single value associated with each location (each pixel takes on a single value). Also, all edge detection methods require threshold values in order to separate objects from the background [20]. Therefore, the only means available to detect an edge is a comparison of these deterministic values. However, in the case of measured terrain surfaces, the height at each location is probabilistic, described by a cumulative probability function. A novel edge detection technique is developed in this work based on the comparison between two cumulative probability distributions of two neighborly nodes. This additional information is leveraged in the work of Comparative Nodal Uncertainty Edge Detection which is described in Section 3.3. The Wilcoxon Rank-Sum Test is used to compare two distribtuion. A description of Wilcoxon Rank-Sum Test is given in Section 3.2.7.

3.2.7 Two-Sample Wilcoxon Rank-Sum Test

The Two-Sample Wilcoxon Rank-Sum Test is one of nonparametric statistic methods which does not rely on assumptions that the data are drawn from a given probability distribution. It is based solely on the rank in which the observations from the two samples fall.

Specifically, suppose that you have samples of observations from each of two populations, X and Y, containing m and n observations, respectively. Two-sample Wilcoxon rank-sum test is to test the hypothesis that the distribution of m-measurements in population X is the same as the distribution of n-measurements in population Y, which can be written symbolically as in Equation 3-7, where $\mathrm{H_0}$ is null hypothesis and $\mathrm{H_a}$ is alternative hypothesis in this test.

$$\begin{aligned} &\mathrm{H_0}: X = Y \\ &\mathrm{H_a}: X \neq Y \end{aligned} \qquad \text{Equation 3-7}$$

Wilcoxon Rank-Sum Test **Error! Reference source not found.**] is selected to provide a quantification of how much intermingling there exist in the two independent groups of population X and Y. This quantification can be used as a measurement of equality of two nodal heights by generating two nodal height cumulative probability distributions. Specifically, if the obtained probability is very high which means the null hypothesis is true, it is more confident to say that the two research nodes are not on the edge due to the little variation between them. Otherwise these two research nodes are on the edge since the null hypothesis is not true which indicates there is much variation among them. The null hypothesis and alternative hypothesis in the Wilcoxon Rank-Sum Test could also be written as,

$\mathrm{H_0}$: All observations comes from the same population.

H_a: All observations comes from different population.

The test procedures start from ranking the m + n observations of the combined sample from smallest to largest, where the smallest has rank 1 and the largest has rank of the total count of two samples. The data set applied in this work has nine elements from the cumulative probability distribution of each sample node. The nine elements include nodal heights at 10%, 20% ... 90%.

The Wilcoxon rank-sum test statistic W is the sum of the observed rank for observations from X, listed in Equation 3-8.

$$W = \begin{gathered} the\ sum\ of\ the\ ranks\ in\ the\ combined \\ sample\ associated\ with\ X\ observations \end{gathered} \qquad \text{Equation 3-8}$$

Analytically, the smallest possible value of W is $W = 1 + 2 + \cdots + 9 = 45$ (if all nine samples from X are smaller than all nine samples from Y) and the largest possible value is $W = 10 + 11 + \cdots + 18 = 126$ (if all nine samples from X are larger than all nine samples from Y). The test procedure based on the statistic W is to reject H_0 if the computed value W is "too extreme" which means it is very close to 45 and 126. In other words, if the computed value W is very close to the average which is 85.5 it indicates not to reject H_0 which means all observations comes from the same population. Additionally, when both data sets have more than 8 elements, the distribution of W can be approximated by an appropriate normal curve, and its population mean μ_w and standard deviation σ_w can be calculated by Equation 3-9 and Equation 3-10, respectively, where m and n are sample sizes of two groups which are both nine in this work.

$$\mu_w = \frac{m(m+n+1)}{2} \qquad \text{Equation 3-9}$$

$$\sigma_w = \sqrt{\frac{mn(m+n+1)}{12}} \qquad \text{Equation 3-10}$$

A Central Limit Theorem can then be applied to conclude that when H_0 is true, the test statistic can be shown by Equation 3-11.

$$Z = \frac{W - \mu_w}{\sigma_w} = \frac{W - \frac{m(m+n+1)}{2}}{\sqrt{\frac{mn(m+n+1)}{12}}} \qquad \text{Equation 3-11}$$

This statistic (Z) is used for a two-tailed test, in this case. P-value, which indicates the probability that two groups of samples are from the same population is obtained with the z value by checking in the z-table. Thus, if the obtained P-value is very high, close to 1, the two samples of nodal heights probability distributions are very smilar which means there is little variation between the two samples. Otherwise, there exists a significant variation among them.

3.3 Development of Comparative Nodal Uncertainty Method

A novel edge detection technique is developed in this section based on the comparison between two cumulative probability distributions of two neighborly nodes. The Wilcoxon Rank-Sum Test is used to obtain the probability of two nodes having the same height. Expected Distance Method is developed to obtain the Expected Distance to node with same height matrix where entries of edges are expected to be small. Number Difference Method is developed based on Wilcoxon Rank-Sum Test and the selected significance level to obtain the Number Difference Matrix where entires of edges are expected to be large. Expected Distance Method and Number Difference Method are devloped in Section 3.3.1 and Section 3.3.2, respectively.

3.3.1 Expected Distance Method

After obtaining the probability (p-value from Wilcoxson Rank-Sum test) that two samples are from the same population, the Expected Distance to node with the same height expression is developed in Equation 3-12, where, d_i is the distance between the research node and the i-th node nearby, $P(d_i)$ represents the probability of the research node having the same height with the i-th node nearby. The Expected Distance to node with same height of each point is the sum of Expected Distance of nodes nearby. The amount of nodes nearby included in the calculation depends on how big the mask it applies which is also explained in this section.

$$E[D] = d_1 P(h_0 = h_1) + d_2 P(h_0 = h_2) + d_3 P(h_0 = h_3) + \cdots = \sum_{i=1} d_i P(h_0 = h_i) \qquad \text{Equation 3-12}$$

Distance between nodes and calculated probability from Wilcoxon Rank-Sum Test are two factors affecting Expected Distance to node with the same height. The value of Expected Distance is expected to be small for the node on the edge when the calculated probability and distance are both small. Small calculated probability from Wilcoxon Rank-Sum Test indicates there exists much variation on nodal heights among these two nodal height distributions. Small distance indicates that two nodes are close to each other which means they affect each other more significantly. Four major combinations of calculated probability and distance are listed in Table 3-1. A threshold is needed in order to determine all the possible edges from the Expected Distance Matrix.

Table 3-1 Combination of calculated probability and distance between nodes

	Probability	
	Large	Small

Distance	Large	Large Expected Distance to node with same height, not on the edge.	On the edge if Expected Distance to node with same height below a threshold
	Small	On the edge if Expected Distance to node with same height below a threshold	Small Expected Distance to node with same height, on the edge.

In order to simplify the calculation, two parameters are defined. Mask is defined as the range of area needed to look at or how many nodes counted in calculating the expected distance to node with same height of a research node. Level is required to capture all possible nodes within the radius, where the level number is the round-up approximation of the input radius. Thus, the size of biggest mask can be enlarged to 3×3, 5×5, 7×7, etc, using Equation 3-13.

$$\text{largest mask size} = 2 \times \text{level} + 1 \qquad \text{Equation 3-13}$$

For example, if input radius is "2", then level is "2" and according to Equation 3-13 the size of the largest mask should be 5×5 which is shown in Figure 3-6. Any node in a grid will become the green center node and the nodes in purple represent all nodes counted by the biggest mask.

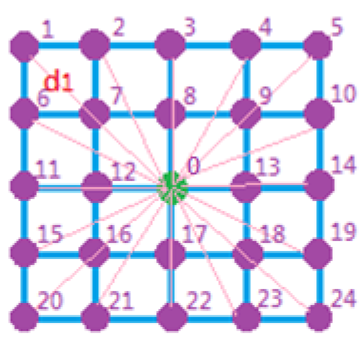
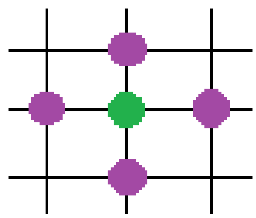


Figure 3-6 The 5×5 mask with the research node in the center

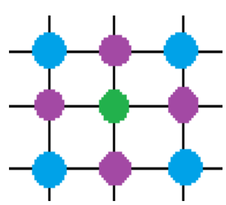
Radius, the range determining the area the user wants to look in calculating the Expected Distance to node with same height is needed as input which is possibly a decimal number. Then the calculation will start from the smallest mask 3-by-3 to the largest mask whose mask number is related to the round-up approximation of the input radius. In this example, the input radius is “2”, thus the level number is “2” and the size of the largest mask is 5×5. The smallest mask and its distance matrix is shown in Figure 3-7.



$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Figure 3-7: Smallest mask of radius of 2 and its distance matrix

The first mask considers any position with distance “1” to the research node located in the center. Any node with distance to center node larger than “1” is not considered and set it as “0”. The Expected Distance of this first mask is the sum of p-values from Wilcoxon Rank-Sum Test between green node and each purple node multiplied by the distance matrix. The second mask of radius of 3 is shown in Figure 3-8.

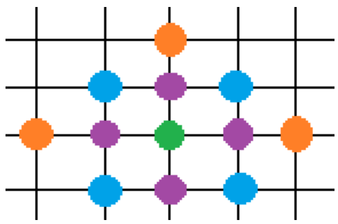


$$\begin{bmatrix} \sqrt{2} & 1 & \sqrt{2} \\ 1 & 0 & 1 \\ \sqrt{2} & 1 & \sqrt{2} \end{bmatrix}$$

Figure 3-8: The second mask of radius of 3 and its distance matrix

The second mask considers any position with distance “1” and “$\sqrt{2}$” to the research node located in the center. Any node with distance to center node larger than “$\sqrt{2}$” is not considered and set it as “0”. The Expected Distance of this second mask is the sum of p-values from Wilcoxon

Rank-Sum Test between green node and each node in purple or blue multiplied by the distance matrix. The third mask of radius of 3 is shown in Figure 3-9.



$$\begin{bmatrix} 0 & 0 & 2 & 0 & 0 \\ 0 & \sqrt{2} & 1 & \sqrt{2} & 0 \\ 2 & 1 & 0 & 1 & 2 \\ 0 & \sqrt{2} & 1 & \sqrt{2} & 0 \end{bmatrix}$$

Figure 3-9: The third mask of radius of 3 and its distance matrix

The third mask considers any position with distance "1", "$\sqrt{2}$" and "2" to the research node located in the center. Any node with distance to center node larger than "2" is not considered and set it as "0". The Expected Distance of this third mask is the sum of p-values from Wilcoxon Rank-Sum Test between green node and each node in purple, blue or orange multiplied by the distance matrix. Since the entries of the mask have reached the level which is the round-up approximation of the input radius of 2, no more further mask is needed to be considered. Thus, the Expected Distance Matrix of each mask of the grid in Figure 3-6 is calculated by Equation 3-14.

$$E[D] = d_1 P(h_0 = h_1) + d_2 P(h_0 = h_2) + \cdots + d_{24} P(h_0 = h_{24}) = \sum_{i=1}^{24} d_i P(h_0 = h_i) \qquad \text{Equation 3-14}$$

The mask will be moved to each node in the grid except the perimeter which can be approximated by the same value of the node most nearby. In the end, Expected Distance to node with same height matrix for all nodes will be obtained. The final step of edge detection is to filter out the noise background with Expected Distance value higher than the set threshold which can be defined on a scale from 0 to 1 and the left nodes are the edge nodes. The filter equation is shown in Equation 3-15. There are two methods to remove noise background which will be discussed in Section 3.3.1.1 and Section 3.3.1.2.

$$\text{Edge Nodes} = E[D] < \text{threshold} \qquad \text{Equation 3-15}$$

3.3.1.1 Cumulative density function of expected distance

After obtaining Expected Distance matrix where each entry locates between 0 and 1, a cumulative density function of expected distance can be drawn out. As stated in Section 3.3.1, points with lowest expected distance values are expected to be on the edge. Therefore, the amount of edge points is less than points on surfaces. Thus, from the cumulative density function of expected distance, if select small percentages, such as, 10%, 20%, and find the corresponding expected distance value which can be set as expected distance threshold, then the noise background can be removed by discarding points whose expected distance value are above this selected expected distance threshold.

Once obtaining the cumulative density function of expected distance, a percentage can be chosen by user and the threshold will be calculated automatically. Finally, noise background will be removed from the results. Section 3.3.1.2 provides another way of removing noise background which user can select threshold manually, and more conditions are included in the filtering process.

3.3.1.2 Search for nodes with the lowest expected distance values

Expected Distance matrix has all entries located between 0 and 1. As stated in Section 3.3.1, the entry, the expected distance is expected to be small if this node is on the edge. In other words, the smaller the entry is, the more possible the node is on the edge. Therefore, the node with the smaller value of expected distance has the larger possibility to be the points on the edge.

The process of removing noise background can be considered as finding all possible edge points. The first step is to set a threshold that entries are only lower than this threshold can be counted as edge nodes. The searching process starts by locating the smallest expected distance value. The threshold is adjustable to include more or less possible edges nodes. The founded expected distance value gradually increases until reaching the threshold. If the expected distance value of a founded node is larger than this threshold, then the searching process will be terminated. All founded points by this threshold are considered as possible edge points.

Once found, the node with lowest Expected distance value is set to "1" or greater than 1. To distinguish with other nodes who also have value of "1", the Expected Distance matrix is rescaled to 0 to 0.99. This founded node with the lowest value of expected distance can be the first node of the event to be found.

Then, look at the eight surrounded nearby nodes (in the direction of up, down, left and right or up-left, up-right, down-left, and down-right added) and search for the node with lowest expected distance value and set its expected distance value as "1", too. This means this node is connected to the first node which yields the edge in a sequence. Same searching process will be gone through to find other nodes in the event. There are three conditions determining whether to stop the searching process.

- Find the edge node with same index which means the event is closed
- Reach the perimeter of grid
- Reach the threshold of Expected Distance value

In the third condition, when reaching the expected distance threshold, while having not found the edge node with the same index yet, this indicates this event may be a crack or non-closed event. All possible edge points will be found under these three conditions, and the other points will be considered as noise background.

This method of removing noise background is not as simple as drawing the cumulative density function in Section 3.3.1.1. However, this method does have advantages. Firstly, this method doesn't rely on the cumulative density function and the percentages the user needs to select. Attempt of choosing a threshold is more direct to see the results. Secondly, in the three conditions listed in this section, this method is also capable of closing an event, as well as eliminating non-closed events. In this way, edges points will be found more accurately. Finally, the process of searching for nodes with the lowest expected distance values is as fast as the method of drawing a cumulative density function.

3.3.2 Number Difference Method

Number Difference Method is also based on Wilcoxon Rank-Sum Test. The only difference of output matrix with the output matrix from the Expected Distance Method is that the

Number Difference Method uses significance level of Wilcoxon Rank-Sum Test to reject the null hypothesis which supports two group of data lie on the same height distribution. The significance level can also be understood as the probability of rejecting the null hypothesis given that it is true and is most often set at 0.05 (5%), shown in Equation 3-16. For independent tests the procedure has type I error probability equal to α [35].

$$\text{significance level } \alpha = 0.05 \qquad \text{Equation 3-16}$$

All probability obtained from Wilcoxon Rank-Sum Test will be compared to this significance level and for probabilities which are lower than significance level will be counted as edge points. Repeat the process for all nodes in the 3-by-3 mask of a research node defined in Section 3.3.1. Count the number of probability which is lower than the significance level and this total amount is defined as the entry of Number Difference Matrix.

Figure 3-10 shows several Number Difference Matrix examples which may be obtained from the Number Difference Method. True edges are indicated by the circles or squares in red. Marked number indicates the amount of points which are located at the opposite side in a mask of a research node. For example, if a research node is located inside of the edge, then any point located outside of the edge will contribute “+1” to the Number Difference of this research node. All marked numbers consist of the Number Difference Matrix of this grid. As seen in Figure 3-10, points with larger Number Difference are closer to the edge.

A threshold is needed to filter points with number lower than this threshold. If the threshold is set as 3, then the points with numbers lower than 3 which are also indicated in black will be removed off from the Number Difference; any points with numbers equal to or higher than 3 which are indicated in blue will remain as edge points. Points with larger Number Difference are considered as edge points except point with “extreme” numbers like “7” and “8”. In those cases the event is too small to be counted as an event. More discussion about the selection of threshold will be provided in Section 3.5.

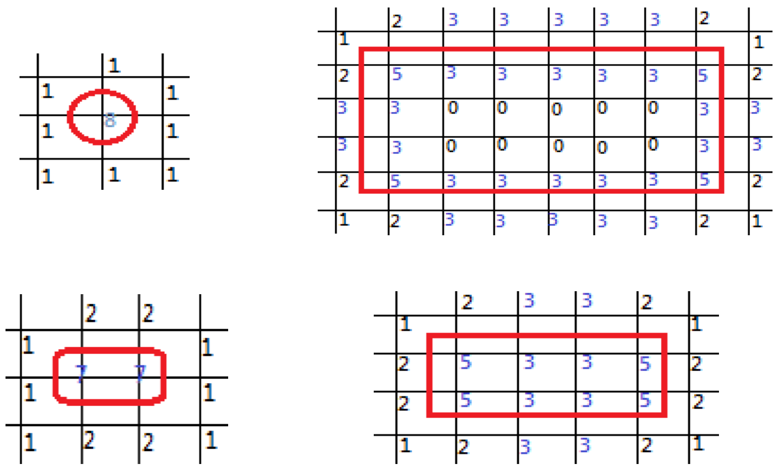


Figure 3-10: Number Difference Matrix example of a grid

3.4 Examples and Results

This section provides a few examples of manmade pothole from a road data set to test the Comparative Nodal Uncertainty Method. Section 3.4.1 provides the results of examples using gradients method. Section 3.4.2 provides the results of examples using Expected Distance Method. Section 3.4.3 provides the results of examples using Number Difference Method. In each example, the contour plot on original data will be firstly given to give a look of a single pothole. Then Expected Distance Method and Number Difference Method will be applied.

3.4.1 Gradients Method

The first example is a pothole of 20×20 locating at the center of a 50×50 grid. The contour plot of the real pothole is shown in Figure 3-11. The area in yellow represents the high place and the area in blue represents the low place. The color bar on the right gives the height of each node on the grid with the unit of millimeter.

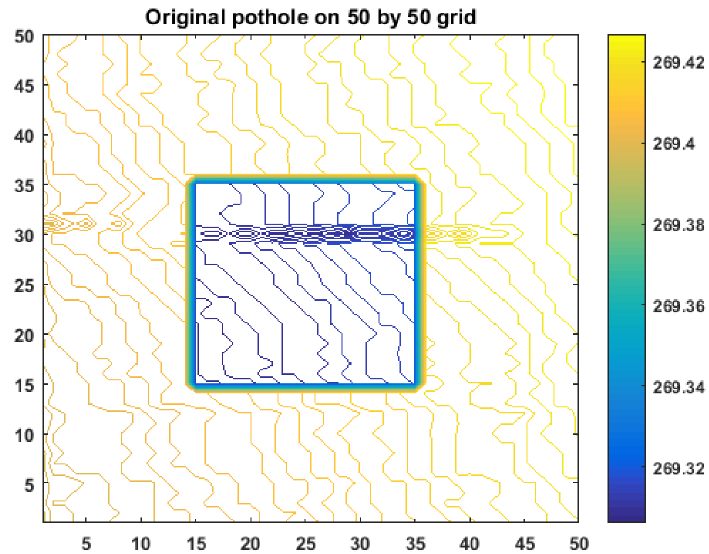


Figure 3-11: Contour plot of a 20×20 pothole with steep edges in a grid of 50×50

As shown in Figure 3-12, this 20×20 square pothole has a very steep edge. The area in yellow represents the high place and the area in blue represents the low place. The color bar on the right gives the height of each node on the grid with the unit of millimeter.

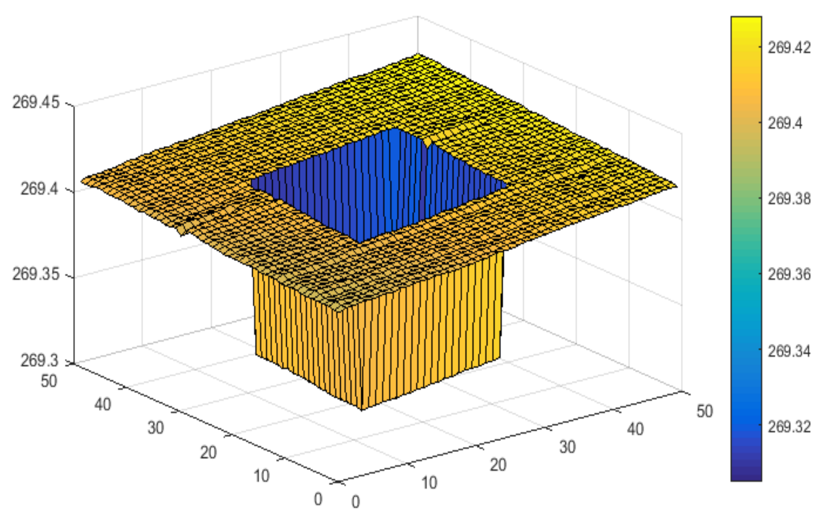


Figure 3-12: Surf plot of pothole with steep edges

Gradient methods results are shown in Figure 3-13. There are six gradient methods and they are Sobel, Prewitt, Roberts, Laplacian of Gaussian, Canny, Zero-cross, respectively. The horizontal coordinate u has an equivalent interval of 0.1 mm and the vertical coordinate v has an equivalent interval of 0.05 mm. From the color bar located in the right of each figure, all obtained values lie between 0 and 1. It is apparent that the yellow area represents the edges.

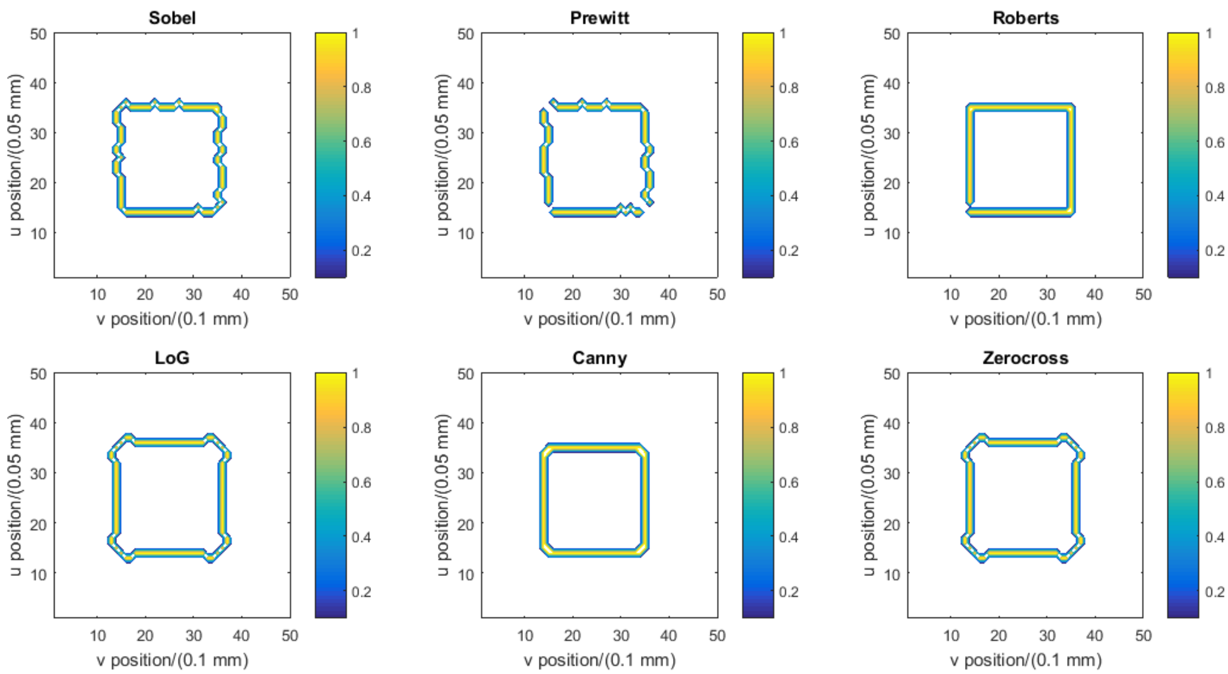


Figure 3-13: Gradients plot of pothole with steep edges

The second example is a pothole of 20×20 locating at the center of a 50×50 grid. The only difference with the first example is the edge of this manmade pothole gradually changes. The contour plot of the real pothole is shown in Figure 3-14.

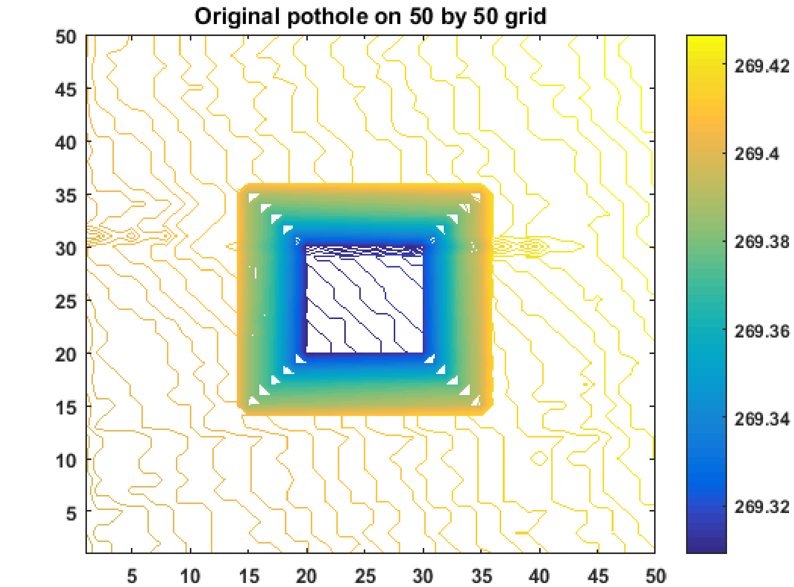


Figure 3-14: Contour plot of a 20 × 20 pothole with gradual edges in a grid of 50 × 50

As shown in Figure 3-15, this 20 × 20 square pothole has a gradual edge. The area in yellow represents the high place and the area in blue represents the low place. The maximum and minimum value in the color bar are exactly same with the values shown in Figure 3-12 color bar, which means the only difference between this pothole and the first one is the edges are gradual in this case.

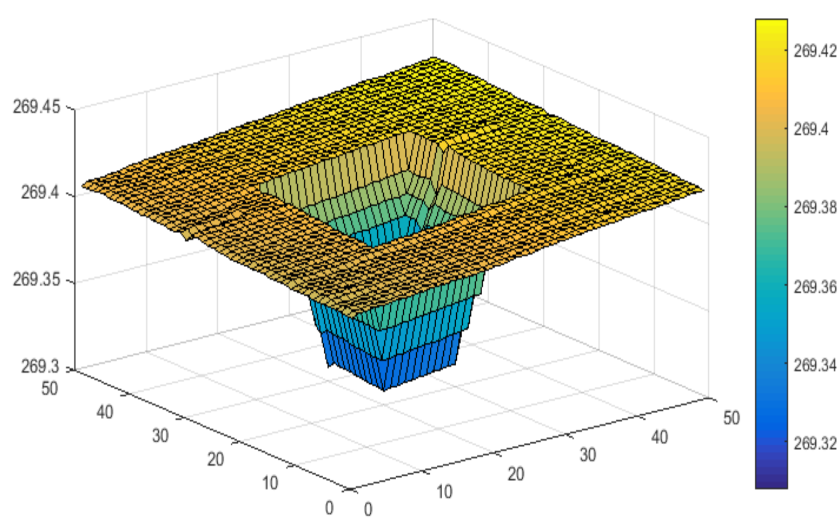


Figure 3-15: Surf plot of pothole with gradual edges

Gradient methods results are shown in Figure 3-16. All yellow area represents the detected edges. The comparison between this pothole with gradual edges and the first example the pothole with steep edges will be discussed in Section 3.5.

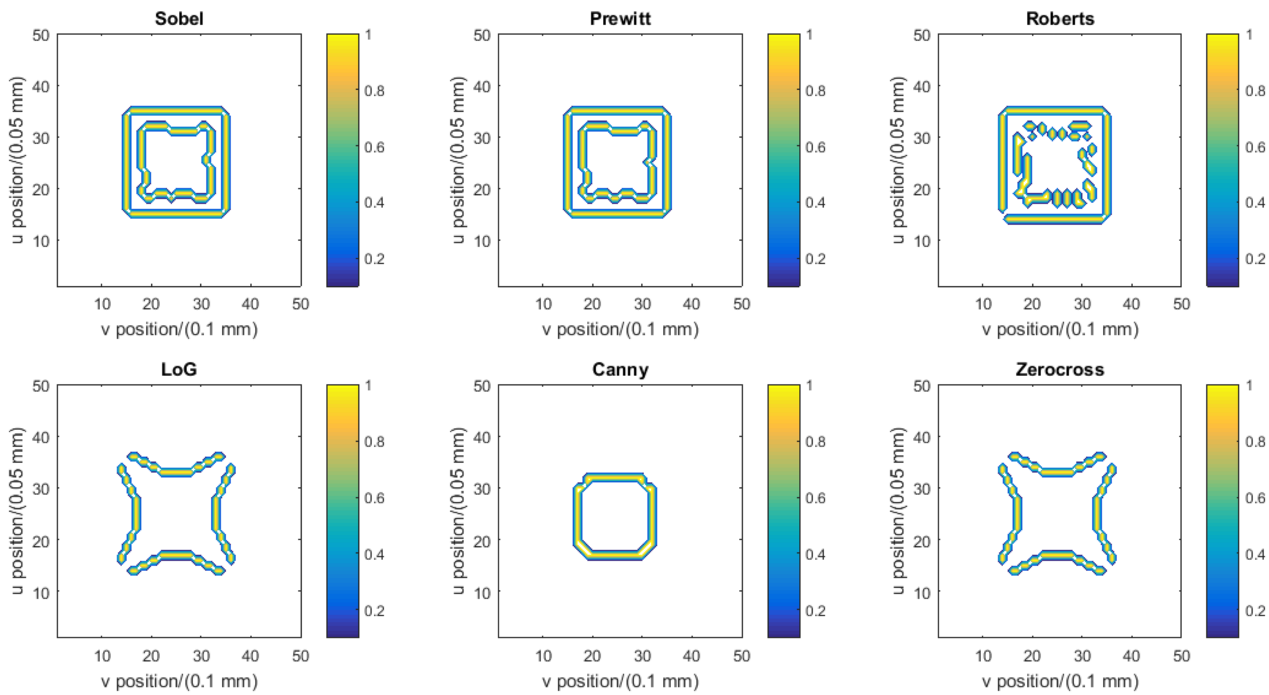


Figure 3-16: Gradients plot of pothole with gradual edges

3.4.2 Expected Distance Method

Expected Distance Method is applied on manmade pothole with steep edges and gradual edges. Figure 3-17 displays the output of Expected Distance Method on the pothole with steep edges in the case of radius is 3. There are six masks obtained from Expected Distance Method on this pothole.

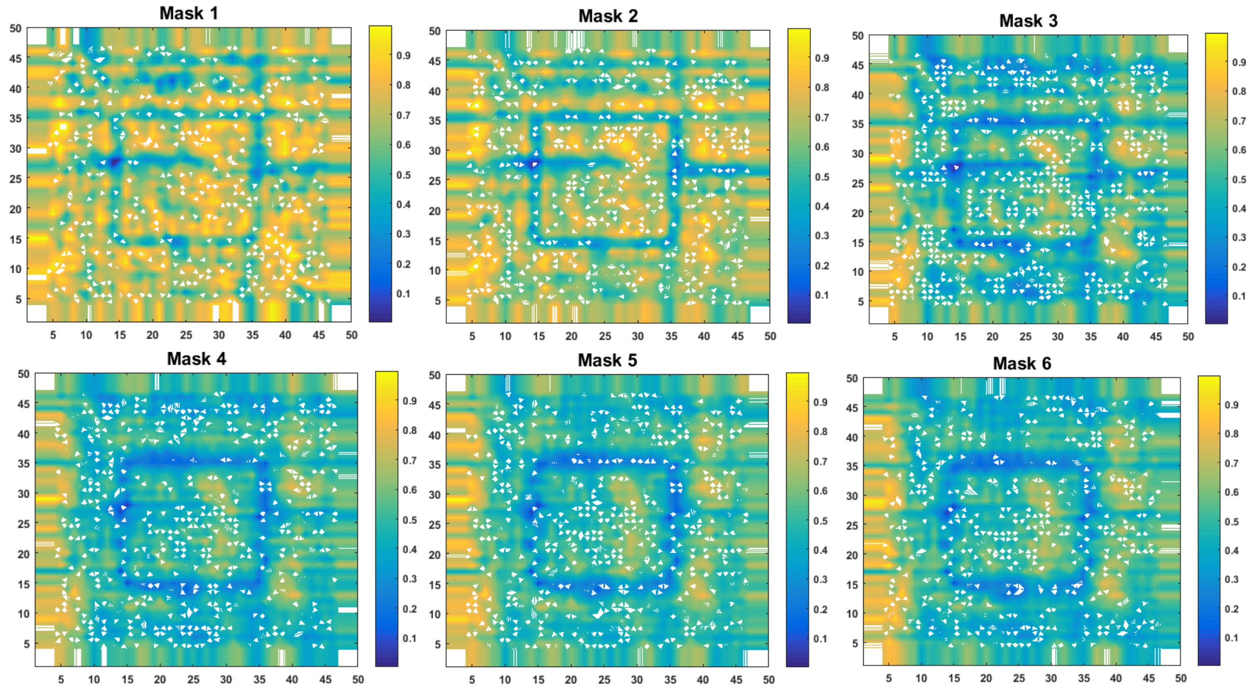


Figure 3-17: Expected Distance Edge Detection Method on pothole with steep edges

After obtaining Expected Distance matrix from Expected Distance Edge Detection Method, a cumulative density function shown in Figure 3-18 can be drawn out and used to select threshold to remove off noise background.

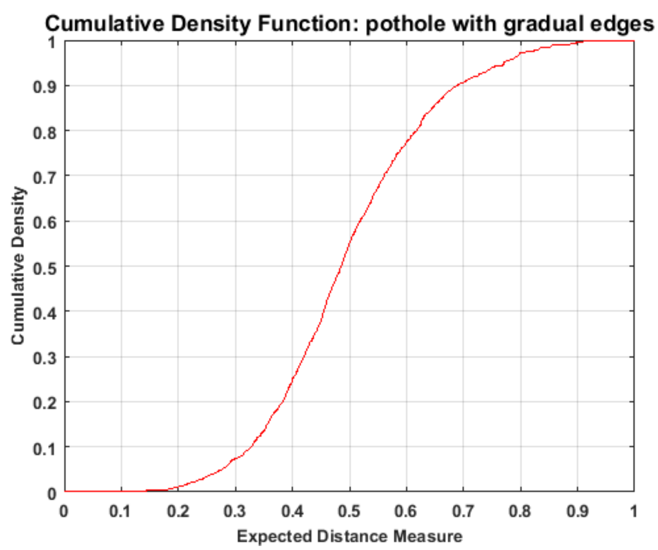


Figure 3-18: Cumulative Density Function of expected distance on pothole with steep edges

Since all edge points should have a relatively low expected distance value compared to other points, it is reasonable to select a percentage to filter out all points whose expected distance measure is higher than this corresponding threshold according to the selected percentage. **Error! Reference source not found.** displays filtered edge points selecting three different percentages. These results will be discussed in Section 3.5 and used in Section 4.4. Since this manmade pothole is single and simple, noise background removal method of searching for nodes with the lowest expected distance values is tested to have the same result as this method of drawing a cumulative density function of expected distance values.

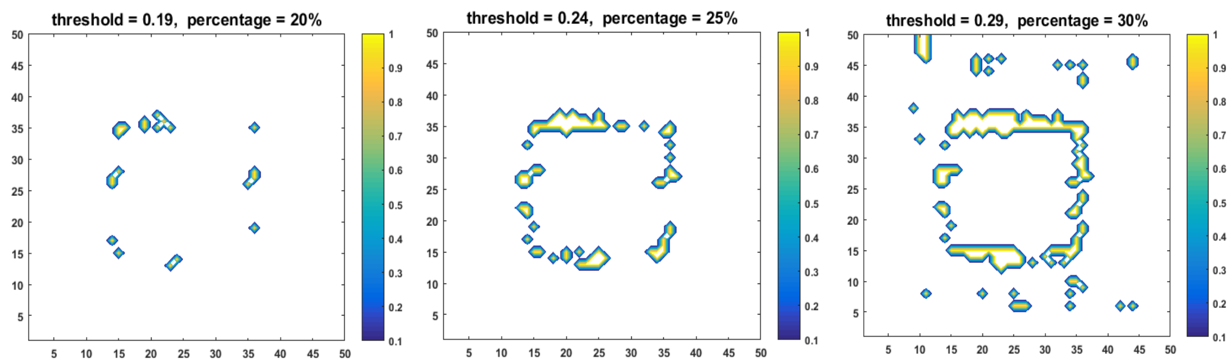


Figure 3-19: Removed noise background using different threshold on pothole with steep edges

In order to simulate the real pothole in the terrain surface, a manmade pothole with gradual edges is used to test the applicability of Expected Distance Edge Detection Method. Figure 3-20 displays the output of Expected Distance Method on the pothole with gradual edges in the case of radius is 3. There are six masks obtained from Expected Distance Method on this pothole.

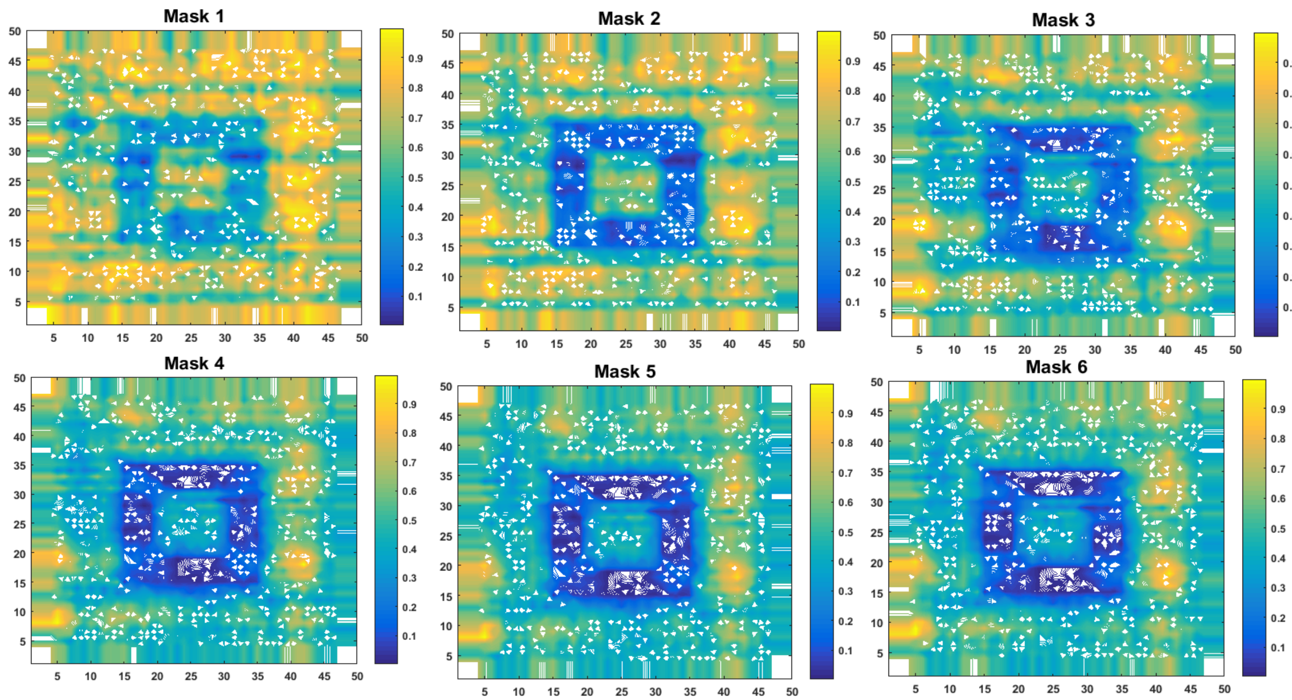


Figure 3-20: Expected Distance Edge Detection Method on pothole with gradual edges

Similarly, a cumulative density function shown in Figure 3-21 can be drawn out and used to select threshold to remove off noise background.

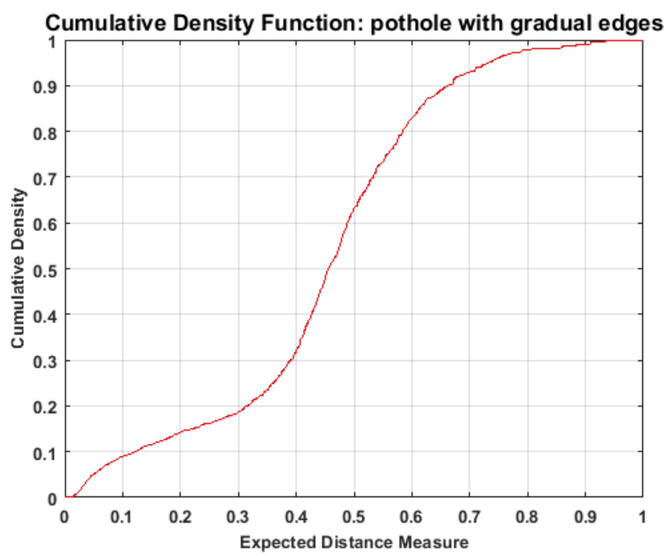


Figure 3-21: Cumulative Density Function of expected distance on pothole with gradual edges

Error! Reference source not found. displays filtered edge points selecting three different percentages. These results will be discussed in Section 3.5 and used in Section 4.4.

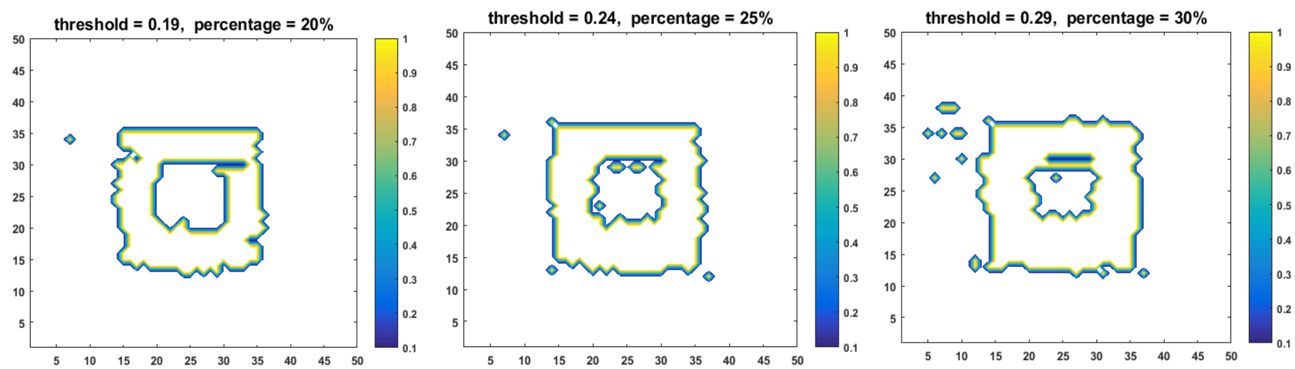


Figure 3-22: Removed noise background using different threshold on pothole with gradual edges

3.4.3 Number Difference Method

Number Difference Method is applied on an artificial pothole with steep edges and gradual edges. **Error! Reference source not found.** displays the output of Number Difference Edge Detection Method on the pothole with steep edges in the case of different significance level.

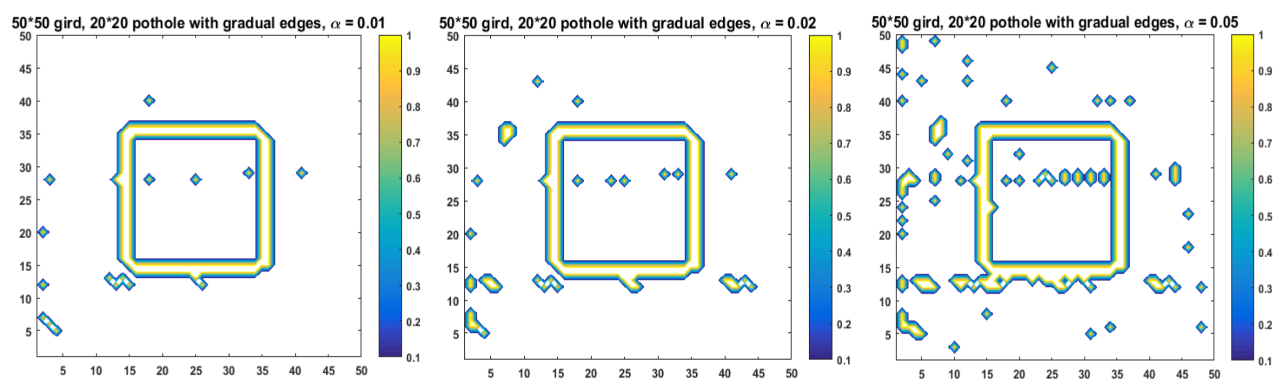


Figure 3-23: Number Difference Method on pothole with steep edges in the case of $\alpha = 0.01, 0.02,$ and 0.05

Error! Reference source not found. displays the output of Number Difference Edge Detection Method on the pothole with gradual edges in the case of different significance level.

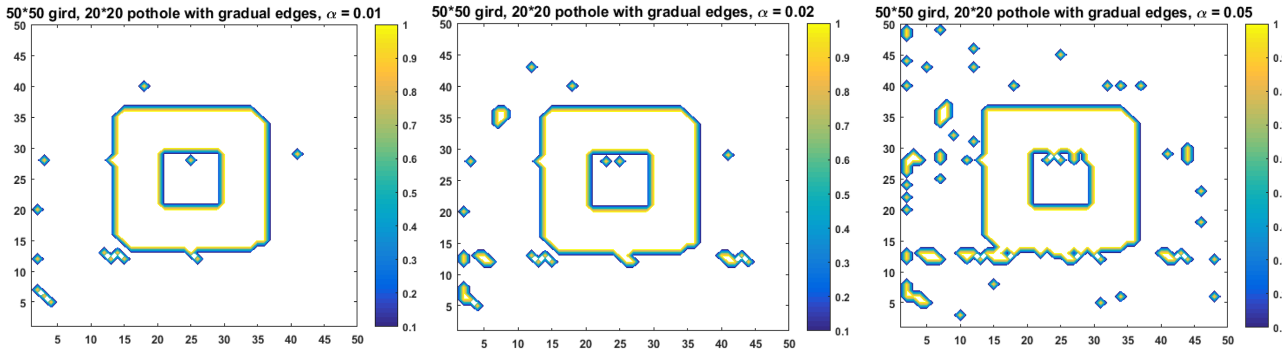


Figure 3-24: Number Difference Method on pothole with gradual edges in the case of $\alpha = 0.01, 0.02, \text{and } 0.05$

3.5 Discussion

As shown in Figure 3-13, Figure 3-16, **Error! Reference source not found.**, **Error! Reference source not found.**, **Error! Reference source not found.** and **Error! Reference source not found.**, all edge detection methods can detect out edges in the case of a pothole with steep edges and a pothole with gradual edges. However, the accuracy of detected edges varies among edge detection methods. This section will provide discussion in three combinations. Section 3.5.1 will discuss the comparison between gradients methods and Expected Distance Method. Section 3.5.2 will discuss the comparison between gradients methods and Number Difference Method. Finally, the comparison between Expected Distance Method and Number Difference Method will be discussed in Section 3.5.3. Two circumstances of pothole with steep edges and pothole with gradual edges will be discussed in each comparison.

3.5.1 Comparison between gradients methods and Expected Distance Method

Traditional gradient methods detect edges from a field of scalar values (pixels in grayscale). Therefore, the only means available to detect an edge is a comparison of these deterministic values.

As demonstrated in Figure 3-13, all gradient methods can approximately detect out steep edges of a single pothole, but Roberts and Canny can accuately detect edges more precisely than other gradients methods. In real terrain surfaces, it is relatively rare to find steep edges so the simulation of a pothole with gradual edges is necessary.

In the case of a pothole with gradual edges, none of the gradient methods except Canny can detect the correct edges. Sobel and Prewitt detect two complete edges since edges, in gradient methods, are defined as the abrupt change in surface, which should be two edges connected between flat surface and steep surface, as shown in Figure 3-16. Roberts method can detect the outer edges very well but the inner edges are not continuous. Laplacian of Gaussion and Zerocross methods don't produce the correc shape of the edges. As stated in Section 3.2.6, in the case of a real measured terrain surfaces, the height at each location is probabilistic, described by a cumulative probability function. The Expected Distance Edge Detection Method and the Number Difference Method are based on the comparison between two cumulative probability distributions of two neighboring nodes, both using the Wilcoxon Rank-Sum Test.

From Figure 3-17, mask 4, 5, and 6 can detect the steep edges relatively clearly without noise in the background. In the case of a pothole with gradual edges, mask 4, 5 and 6 in Figure 3-20 give clear edges. Masks become bigger with the increase of the mask numberand includes more points in the calculation. The detected edges in Figure 3-20 are wider than edges in Figure 3-17 because the edges in Figure 3-20 change gradually. The entry of the expected distance matrix are the expected distance measure which lies between 0 and 1. Edge points with low expected distance are indicated as blue in Figure 3-17. Figure 3-18 is used to determine expected distance threshold which removes off the noise background. Expected distance threshold is most possibly located below 30% of expected distance measure of all points in the grid. Percentage is preferable than directly choosing an expected distance threshold. The corresponding expected distance threshold will be calculated after setting a percentage. From the three plots in **Error! Reference source not found.** and **Error! Reference source not found.**, the percentage of 25% can detect out the edges completely with minimizing the effect of outliers. Although the other two expected distance thresholds can not detect out the edges clearly, the edges can still be connected completely using Edge Stitching Method which will be introduced in Section 4.

The advantage of Expected Distance Method over gradients method can be found when comparing Figure 3-16 and **Error! Reference source not found.**. Compared to most of gradients method applied on pothole with gradual edges, Expected Distance Method detects edges more robust on the inner edges. The reason is that Expected Distance Method detects edges using probabilistic distribution of a groups of points instead of only using the deterministic gradient values for each node obtained from comparing nodal heights. Therefore, Expected Distance Method is more robust than gradients method on detecting edges.

3.5.2 Comparison between gradients methods and Number Difference Method

As stated in Section 3.3.2, Number Difference Method is also based on Wilcoxon Rank-Sum Test. The only difference of output matrix with one from Expected Distance Method is that the Number Difference Method uses significance level of Wilcoxon Rank-Sum Test to reject the null hypothesis which supports two group of data lie on the same height distribution. In the example part in Section 3.4.3, three circumstances of significance level are considered to compare which one is more applicable and appropriate on this method.

From **Error! Reference source not found.** and **Error! Reference source not found.**, in terms of the part of outliers, significance level of 0.01 gives clearer edges than 0.02 and 0.05. Significance level is used to judge whether the test results are statistically significant, as well to determine the probability of error that is inherent in the test. Statistical significance relates to how likely the observed effect is due to random error instead of a "true" difference between groups [significance level]. Technically, significance level α is the maximum acceptable level of risk for rejecting a true null hypothesis (Type I error) and is expressed as a probability ranging between 0 and 1. For example, a significance level of 0.05 indicates a 5% risk of concluding that a difference exists when there is no actual difference. Choosing a smaller significance level, 0.01, is to be more certain that it will only detect a difference that really does exist, which also means that the Type I error is small.

The advantage of Number Difference Method over gradients method can be found when comparing Figure 3-16 and **Error! Reference source not found.**. Similarly, compared to most of gradients method applied on pothole with gradual edges, Number Difference Method detects edges more robust on the inner edges. The reason is that Number Difference Method detects edges using probabilistic distribution of a groups of points, as well as the significance level, instead of only using the deterministic gradient values for each node obtained from comparing nodal heights. Therefore, Number Difference Method is more robust than gradients method on detecting edges.

3.5.3 Comparison between Expected Distance Method and Number Difference Method

As stated in Section 3.3.1 and Section 3.3.2, both Expected Distance Method and Number Difference Method are based on probabilistic distribution of a groups of points using Wilcoxon Rank-Sum Test. The only difference of output matrix with one from Expected Distance Method is that Number Difference Method uses significance level of Wilcoxon Rank-Sum Test to reject the null hypothesis which supports two group of data lie on the same height distribution.

Compared **Error! Reference source not found.** with **Error! Reference source not found.**, after removing off noise background, Number Difference Method gives more complete and clear edges than Expected Distance Method on the pothole with steep edges. Although with the increase of the significance level in Number Difference Method, increasing outliers appear in the result of detected edges, smaller significance level is good enough to obtain complete edges. However, if fewer outliers are preferable, smaller expected distance threshold in Expected Distance Method cannot detect out clear and complete edges.

Compared **Error! Reference source not found.** with **Error! Reference source not found.**, in the case of pothole with gradual edges, the results of Expected Distance Method and Number Difference Method are similar except the detected edges from Number Difference Method produces more straight boundaries than ones from Expected Distance Method. Furthermore, the significance level of 0.01 is preferable in Number Difference Method to detect out relatively complete edges with fewer outliers.

3.6 Contribution

This chapter reviews traditional gradients edge detection methods and introduces novel edge detection method Expected Distance Method and Number Difference Method. The principal contributions of this chapter are:

- Developed Expected Distance Method and Number Difference Method using probabilistic distribution of grid nodes instead of only using the deterministic gradient values for each nodes obtained from comparing nodal heights.

- Explored the justification of choosing threshold values in Expected Distance Method and given an appropriate choosing range, as well as choosing significance level in Number Difference Method.
- Compared Expected Distance Method, Number Difference Method with gradients methods and found that Expected Distance Method and Number Difference Method are able to reach a higher accuracy of actual edges in the simulated potholes.
- Compared Expected Distance Method with Number Difference Method and found that Number Difference Method is more robust than Expected Distance.

4. Edge Stitching

4.1 Introduction

The Comparative Nodal Uncertainty Edge Detection Method is developed to detect edges based on the probability distribution of the nodal heights within some local neighborhood. However, as shown in **Error! Reference source not found.**, edge points may not be continuous, forming a complete event. Furthermore, false-positive identification of an edge point often occur, as shown in **Error! Reference source not found.**. In order to solve this ambiguity, a novel Edge Stitching method is developed to group edge points together in a sequential chain to form the final events. Some existing image mosaicking, as a key concept in image processing field, works by combining two or more images by stitching them together so they appear to be one image [21].

In this work, the novel Edge Stitching method uses the results of Expected Distance matrix or Number Difference matrix from Comparative Nodal Uncertainty Edge Detection Method to develop a search algorithm in order to identify a sequential chain of edge nodes to define complete and continuous events.

The reminder of this chapter proceeds as follows. Section 4.2 provides an overview of a series of techniques which will be used in Section 4.3. The Edge Stitching Method is developed in section 4.3 by separating into three parts: obtaining all possible edge points, filtering edge points, and connecting filtered edge points to form events. Section 4.4 provides some examples verifying the newly developed Edge Stitching Method and the results of examples are discussed. The chapter concludes with a discussion of the advantages and disadvantages of the methods developed in Section 4.3 and a summary of contribution of Comparative Nodal Uncertainty Edge Detection Method is provided.

4.2 Background

This section gives a review of Comparative Nodal Uncertainty Method from which Edge Stitching will be developed, as well as a series of techniques that will be used in Section 4.3. Section 4.2.1 provides an overview of Comparative Nodal Uncertainty Edge Detection Method.

An explanation of Fourier series is provided in Section 4.3.2. This Section concludes with an explanation of Least Squares in regression analysis.

4.2.1 Comparative Nodal Uncertainty Edge Detection Method

Comparative Nodal Uncertainty Edge Detection Method is a novel edge detection method based upon the Two-sample Wilcoxon Sum-Rank test, which is primarily used to obtain an Expected Distance to node with same height matrix to determine the edges. Expected Distance Method and Number Difference Method are two edge detection methods developed in Chapter 3.

The output of Comparative Nodal Uncertainty Edge Detection Method is an Expected Distance Matrix or a Number Difference Matrix in that entries in both matrices lie between 0 and 1. These two matrices are used to determine edges points in a grid through a threshold. The threshold in the former method is the expected distance measure chosen from any value in a cumulative density function of expected distance values. The threshold in the latter method is the significance level that is also the Type I error in Wilcoxon Rank-Sum Test. After the removal of noise background it will be the input data set of Edge Stitching.

4.2.2 Normal Distribution

The normal distribution is the most popular distribution in statistics [22]. Some of its mathematical properties are very useful in data analysis. Data are said to be normally distributed if their frequency histogram approximates to a bell shaped curve. Practically, the data can be determined as normally distributed by looking their bell-shaped curve histogram. One of the fundamental theorems of probability is the Central limit theorem [23]. If the samples are obtained from a large number of observations and each observation does not depend on other observations, then the computed average of observations will be distributed according to the normal distribution [24], no matter what the population histogram looks like. In other words, the distribution of the sample mean tends toward the normal distribution as the sample size increases. The mean of the sampling distribution of the sample mean is equal to the mean of the population from which samples are drawn.

The Normal Probability Plot [25, 26] aims at graphically assessing normality of data. If the data are normal the plot will be linear. For normal data the points plotted in the probability plot should fall approximately on a straight line, indicating high positive correlation. These plots are intuitive to interpret and also have the benefit of identifying outliers easily. Figure 4-1 displays an example of Normal Probability Plot. The blue data points marked with "+" symbol follow well along the red line to indicate a normal distribution. Other tests are also available for the proof of normal distribution [27, 28].

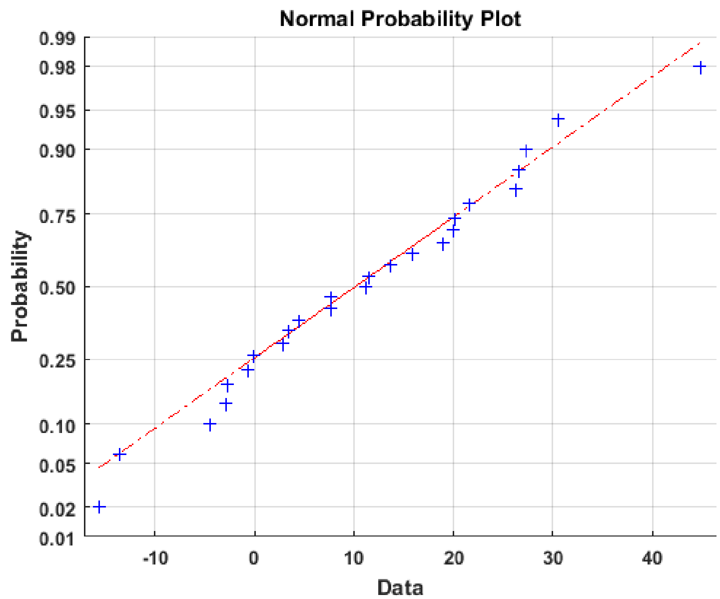


Figure 4-1: An example of Normal Probability Plot

Once the distribution has been confirmed as normal distribution, the Empirical Rule, also known as Three-Sigma Rule can mathematically express that "nearly all" values are taken to lie within three standard deviations of the mean [29]. Specifically, approximately 68% of the data values fall within 1 standard deviation of the mean; approximately 95% of the data values fall within 2 standard deviation of the mean; and approximately 99.75% of the data values fall within 3 standard deviation of the mean.

4.2.3 Fourier series

The basic assumption behind Fourier series is that any given function can be expressed in terms of a series of sine and cosine functions. Sine and cosine functions are basis for the linear space of functions to which the given function belongs [30]. It is the properties of the inner product space (as reviewed in Appendix A.1), coupled with the analytically familiar properties of the sine and cosine functions that make Fourier series useful and powerful.

The sequence of functions forms an infinite orthonormal sequence in the space of all piecewise continuous functions [31] within the interval $[-\pi, \pi]$ where the inner product $\langle f, g\rangle$ is defined in Equation 4-1, with the overbar denoting complex conjugate [30].

$$\langle f, g\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f\overline{g}\,dx \qquad \text{Equation 4-1}$$

It is not difficult to prove the set $\left\{\frac{1}{\sqrt{2}}, \sin(x), \cos(x), \sin(2x), \cos(2x), \ldots\right\}$ is orthonormal since the inner product of any two different terms in this set is zero and the inner product of any term and itself in this set is one in terms of the definition of orthonormality. Having proved that, it holds true that this sequence forms an orthonormal basis for the space of piecewise continuous functions in the interval $[-\pi, \pi]$. This enables an arbitrary element of this linear space of piecewise continuous functions to be expressed as a linear combination of the elements of this sequence, as shown in Equation 4-2 and Equation 4-3.

$$\begin{gathered} f(x) \sim \frac{a_0}{\sqrt{2}} + a_1\cos(x) + a_2\cos(2x) + \cdots \\ + a_n\cos(nx) + \cdots \\ + b_1\sin(x) + b_2\sin(2x) + \cdots + b_n\sin(nx) + \cdots \end{gathered} \qquad \text{Equation 4-2}$$

$$\begin{gathered} f(x) \sim \frac{a_0}{\sqrt{2}} + \sum_{n=1}^{\infty}(a_n\cos(nx) + b_n\sin(nx)) \\ -\pi < x < \pi \end{gathered} \qquad \text{Equation 4-3}$$

This is the standard Fourier series expansion for $f(x)$. The tilde was interpreted in terms of the discontinuity of points. The left hand-side converges to the value of the mean of the two one-sided limits as dictated in Equation 4-4 in the case of having both left and right derivatives.

The tilde can mean equal which to indicate that at points where the function is continuous, the right-hand side converges to $f(x)$.

$$f(x) = \frac{f(x_-) + f(x_+)}{2} \qquad \text{Equation 4-4}$$

After multiplying by $\sin(nx)$ and integrating term by term on the interval $[-\pi, \pi]$, all but the b_n on the right disappears. Similarly, a_n can be given in Equation 4-5 and Equation 4-6, where $n = 0,1,2, \ldots$

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)dx \qquad \text{Equation 4-5}$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)dx \qquad \text{Equation 4-6}$$

It is not convenient and practical to use $\frac{1}{\sqrt{2}}a_0$ as the constant term. A practical Fourier series are shown in Equation 4-7 with $\frac{1}{2}a_0$ as the constant term.

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos(nx) + b_n \sin(nx)) \qquad \text{Equation 4-7}$$
$$-\pi < x < \pi$$

Here is a simple example of Fourier series on a ramp function. In Figure 4-2, the original ramp function is shown by the red line. Fourier series expansion is applied on this function with different amounts of coefficients (constraints) that are the unknown a_n and b_n in Equation 4-7. It is apparent that by increasing the amount of coefficients in Fourier series expansion, the ramp function can be described more and more accurately. The line with “129 constraints” is nearly overlaps with the original function. However, no matter how close the Fourier series expansion is to the original function, any line of the results is still wave-like since the estimated function is composed of sine and cosine waves. This concept will be used in Section 4.3.2.2.

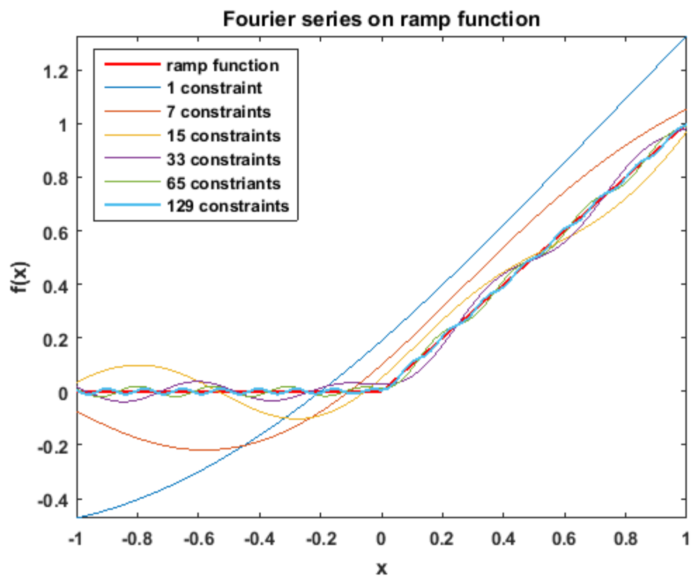


Figure 4-2: Fourier series on ramp function with different amounts of constraints

4.2.4 Least square curve fitting

An important area in approximation is the problem of fitting a curve to experimental data [32]. It is essential to assume that the data set is polluted with some degree of measurement error (“noise”) since the data appears experimental. However, it is not necessarily to construct a curve that goes through every experimental data point. It is desired to construct a function that represents the “sense of the data” or a close approximation to the data. In other words, curve fitting only needs to capture the trend in the data by assigning a single function across the entire range.

The method of curve fitting used in this paper is a most common approach - least square curve fitting. The least square approach defines the “correct” curve that minimizes the sum of squares of the distances between the data points and the curve. For example, Figure 4-3 displays a simple example where a linear straight line can better represent the data with noise.

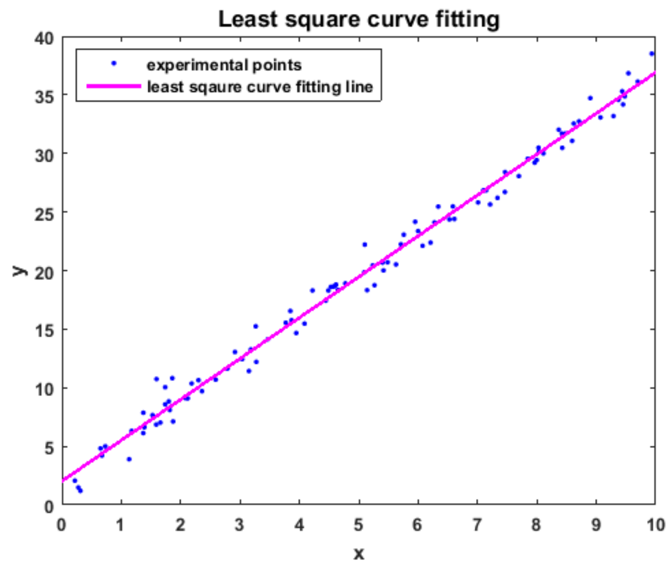


Figure 4-3: Example of least square curve fitting

The principle is shown as follows. Let the noisy data be defined as pairs $(x_k, y_k), 1 \leq k \leq n$ for some n. Thus, in this case, it is desired to find the coefficients p and q in the equation $\mathrm{y} = \mathrm{p}x + q$ such that Equation 4-8 is minimized.

$$\mathrm{S(p,q)} = \sum_{k=1}^{n} \big(y_k - (px_k + q)\big)^2 \qquad \text{Equation 4-8}$$

After solving the straightforward problem of multivariable calculus, the solution is shown in Equation 4-9 and Equation 4-10.

$$\mathrm{p} = \frac{n\sum_{i=1}^{n} x_k y_k - (\sum_{i=1}^{n} x_k)(\sum_{i=1}^{n} y_k)}{n\sum_{k=1}^{n} x_k^2 - (\sum_{i=1}^{n} x_k)^2} \qquad \text{Equation 4-9}$$

$$q = \frac{(\sum_{i=1}^{n} x_k^2)(\sum_{i=1}^{n} y_k) - (\sum_{i=1}^{n} x_k)(\sum_{i=1}^{n} x_k y_k)}{n\sum_{k=1}^{n} x_k^2 - (\sum_{i=1}^{n} x_k)^2} \qquad \text{Equation 4-10}$$

In this example, the notion of a least square data fit is simply a straight line to data. However, in most cases, it can be generalized beyond simply fitting a straight graph to data. Instead, a higher degree polynomials are needed especially for those with higher dimensional data sets. In Section 4.3.2.2, Fourier series will be combined with least square fitting to complete the process of connecting edge points to event.

4.3 Development of Edge Stitching

The goal of Edge Stitching is to develop an algorithm so as to group edge nodes sequentially to its corresponding event. From Section 3, Comparative Nodal Uncertainty Edge Detection Method yields an Expected Distance Matrix or a Number Difference Matrix where entries are the measures to determine edges. Research on Edge Stitching will use the results of Expected Distance Matrix or Number Difference Matrix to find true edge points and connect those points to events.

In Section 3.4.2, Expected Distance Method provides a way of removing noise background with a given threshold. However, as shown in **Error! Reference source not found.**, many outliers do exist that do not belong to the expected event. Thus, the first step of Edge Stitching is to remove the outliers and two methods are developed in Section 4.3.1 to do so. The second step of Edge Stitching is to connect the filtered edge points to events and two methods are developed in Section 4.3.2 as well.

4.3.1 Filtering edge points

An edge point can be defined as an outlier if it is far away from the cluster of edge points. To connect edge points to events and make the obtained event more robust and accurate, it is necessary to filter edge points by removing the edge outliers after all possible edge points are obtained. This section provides two ways of removing edge outliers from all possible edge points obtained from Expected Distance Edge Detection Method and Number Difference Edge Detection Method.

4.3.1.1 Filter by standard deviation of radius distribution

Filtering edges points by standard deviation of radius distribution is based upon the property of normal distribution. As stated in Section 4.2.2, following the Three-Sigma Rule of a distribution, it is necessary to check whether all possible edge points form a normal distribution using the Normal Probability Plot.

First, with all indices of possible edge points stored in a matrix, an approximate center of edge points can be found by averaging indices of horizontal and vertical coordinates using Equation 4-11, where x is the index of horizontal coordinate and y is the index of vertical coordinate and c represents center. The edges of event to be found are described by a cluster of points with an averaged center in Cartesian coordinates.

$$(x_c, y_c) = \left(\frac{\sum_{i=1}^{n} x_i}{n}, \frac{\sum_{i=1}^{n} y_i}{n}\right) \qquad \text{Equation 4-11}$$

Secondly, the distance between each edge point to center can be obtained in Equation 4-12, where i is the number of edge points.

$$\mathrm{d_i} = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2}\ , \quad \mathrm{i} = 1,2,3, \dots, \mathrm{n} \qquad \text{Equation 4-12}$$

Distances of all found edge points to the averaged center consist of a distance distribution where a mean and standard deviation can be obtained from this distribution. Standard deviations can tell us about the shape of the distribution and how close the individual data values are from the mean value. The mean or standard deviation is often estimated from a sample of elements and presented with the standard errors [33, 34]. As stated in Section 4.2.2, the Three-Sigma Rule can be used on a normal distribution to filter outliers. Outliers of edge points can be filtered out by picking edge nodes located only within one standard deviation. The left edge points after outlier filtering are the true edge points among all possible edge points. Examples will be provided in Section 4.4.1.

4.3.1.2 Filter by looking at local points

Outliers are defined as the points which are far away from the cluster of edge points. In other words, outliers are not surrounded by possible edge points as many as the true edge points. The edge points filtering method, instead of using the radius to the center in the filtering method as stated in Section 4.3.1.1, looks at each possible edge point itself while; searching whether more than a fixed number of "N" edge points are around this possible edge point. If there are more than or same number of N points around this point, it will not be defined as an outlier and will stay in the group of edge points; otherwise, it will be considered as an outlier. In doing so, all potential outliers can be picked out no matter where the group of edge points is located. The fixed number "N" can be chosen according to the size of the mask where the edge points are counted. Examples will be provided in Section 4.4.3.

4.3.2 Connecting filtered edge points to events

A cluster of true edge points with a center are obtained after filtering edge points by removing outliers. However, as shown in **Error! Reference source not found.**, the edges, even after filtering outliers, are not complete and may give an unclear view of localized events. Thus, to further characterize events, it would be better to display a complete edge that separates the event from the terrain surface. Connecting filtered edge points to events is therefore necessary to achieve this goal.

Two methods of connecting filtered edge points to events are developed in Section 4.3.2.1 and Section 4.3.2.2. Before connecting filtered edge points to events, it is better to describe each edge point as an angle and radius related to the center. This can be done by converting the

Cartesian coordinates to Polar coordinates. However, edge points have to be translated prior to this conversion so that the center found is fixed in the center of the polar plot. This translation can be achieved quickly and simply with Equation 4-13. This translation fixes the center to origin (0, 0) so that the cluster of edge points will appear within a circle in the polar plot.

$$\begin{pmatrix} x^* \\ y^* \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} x_c \\ y_c \end{pmatrix} \qquad \text{Equation 4-13}$$

Then, the conversion between Cartesian coordinates and polar coordinates can be described by two simple equations shown in Equation 4-14 and Equation 4-15, where r is the radius for each point to the fixed center of polar plot and α is the angle about the origin (0, 0).

$$r = \sqrt{x^2 + y^2} \qquad \text{Equation 4-14}$$

$$\alpha \approx \tan^{-1}\left(\frac{y}{x}\right) \qquad \text{Equation 4-15}$$

Now, all filtered edge points are converted from Cartesian coordinates to polar coordinates and each edge point is described by an angle α and its radius to origin (0, 0). In Section 4.3.2.1 and Section 4.3.2.2, two approaches to connect edge points to event are described based on the converted data in polar coordinates.

4.3.2.1 Connecting edge points to event by discretizing the angle by $2\pi/n$

A further definition of the edge helps to better describe the position of each edge point by looking at discretized angles. Discretizing a circle by number "n" will get "n" segments over $(-\pi, \pi)$ and the increment would be $2\pi/n$. In each discrete angle, an edge representative can be calculated by averaging the radius of all edge points located around this angle within one tolerance. The tolerance can be defined as a half of increment so the range being looked around an angle is one increment. Thus, the cluster of edge points can be sorted out to edge representatives located at discretized angles.

However, it is very possible that some discretized angles might not have an edge representative when there are no edge points located within the tolerance of this discretized angle. Actually, this always happens in the low density area. Hence, in order to have a continuous edge after Edge Stitching, it is necessary to fill the gap angle where no edge representative exists. This can be done by averaging the edge representatives on the two angles most nearby and available. Then the edge representative at each discretized angle can be obtained finally. Generally speaking, the process of discretizing and approximation yields "n" edge representatives at fixed discrete angles characterized by a radius.

Finally, edge points can be connected to an event by converting polar coordinates back to Cartesian coordinates. Examples will be provided in Section 4.4.3.

4.3.2.2 Connecting edge points to event by least square curve fitting

This method of connecting edge points to event uses Fourier series (as reviewed in Section 4.2.3) and least square curve fitting (as reviewed in Section 4.2.4). The combination of Fourier series and least square curve fitting yields a function by which edge points can be connected to event. Now, all filtered edge points are converted from Cartesian coordinates to polar coordinates and each edge point is described by an angle α and its radius to origin (0, 0). Equation 4-16 displays a Fourier series expansion where α is the angle of edge representatives in polar system, ω is the wave frequency that can be estimated as "1", $A_0, A_1, A_2, \ldots$ are the coefficients (constraints) of the estimated function and $f(\alpha)$ is the estimated radius to the center.

$$f(\alpha) = A_0 + A_1 \sin(\omega\alpha + A_2) + A_3 \sin(2\omega\alpha + A_4) + A_5 \sin(3\omega\alpha + A_6) + \cdots \qquad \text{Equation 4-16}$$

The goal is to find an estimated function that can best represent the distribution of edge points. Once the radius and angle of each edge point to center are found in polar coordinates, the plot of angle α versus estimated radius will provide a clear view of the distribution of all edge points with both parameters extracted from Fourier series. The curve on this plot reflects all information included in the Fourier series equation.

The curve obtained in this way is able to represent the distribution of all filtered edge points once all coefficients in the Equation 4-16 have been determined. Instead of determining the coefficients with Equation 4-5 and Equation 4-6, another way to determine all coefficients is Least Square Fitting which is a mathematical procedure to find the best-fitting curve to a given set of points. This process can be done by minimizing the sum of the squares of the offsets (“the residuals”) of the points from the curve, Equation 4-17 is used in this case, where r is the real radius from every edge point to center and $f(\alpha_i)$ represents the estimated distance from every each point to center. Using the sum of the squares of the offsets instead of the sum of the offsets allows the error to be treated as a positive measure. One of the concerns of using squares of the offsets is that it will make the outlying points have a disproportionate effect on the fit. However, we can minimize such effect by removing the outliers in the first step of Edge Stitching. Thus, Least Curve Fitting is a good way to find the best fit coefficients to achieve the goal of connecting edge points to event.

$$\text{Error}^2 = \sum [r_i - f(\alpha_i)]^2 \qquad \text{Equation 4-17}$$

4.4 Example and Result

This section provides a few examples of manmade pothole using the developed Edge Stitching method to identify the event. Since there are two ways of filtering edge points and two ways of connecting edge points to event, totally four Edge Stitching method can be developed. This section is organized as follows. Section 4.4.1 provides an example of filtering edge points by standard deviation of radius distribution followed by discretizing angles from $-\pi$ to π to connect edges. Section 4.4.2 provides an example of filtering edge points by standard deviation of radius distribution followed by least square curve fitting to connect edges. Section 4.4.3 provides an example of filtering edge points by looking local points followed by discretizing angles from $-\pi$ to π to connect edges. Finally, Section 4.4.4 provides an example of filtering edge points by looking local points followed by least square curve fitting to connect edges. The results of two typical potholes, one with steep edges and the other one with gradual edges as mentioned in Section 3 will be listed in this section. Table 4-1 displays a clear organization of this section. Four Edge Stitching methods use both the result of Expected Distance Edge Detection Method and the result of Number Difference Edge Detection Method for pothole with steep edges and pothole with gradual edges.

Table 4-1: Summary of four Edge Stitching methods

Edge Stitching	Discretizing angles from $-\pi$ to π	Least square curve fitting
Standard deviation of radius distribution	Section 4.4.1	Section 4.4.2
Looking local points	Section 4.4.3	Section 4.4.4

4.4.1 Filter by standard deviation of radius distribution, connect edges by discretizing angles

The original pothole with steep edges can be seen in Figure 3-11 and Figure 3-12. The original pothole with gradual edges can be seen in Figure 3-14 and Figure 3-15. Both pothole are approximately 20×20 square located in a 50×50 grid.

<u>Results from Expected Distance Edge Detection Method on pothole with steep edges</u>

Three cases of different threshold are tested in Section 3.4.2 and their results are used in this Edge Stitching method. The data store shown in **Error! Reference source not found.** is used for the analysis of generating edges with a center, fixed angle, and respective radius in both polar coordinates and Cartesian coordinates. The image has a suspected edge indicated by the high intensity blue shapes.

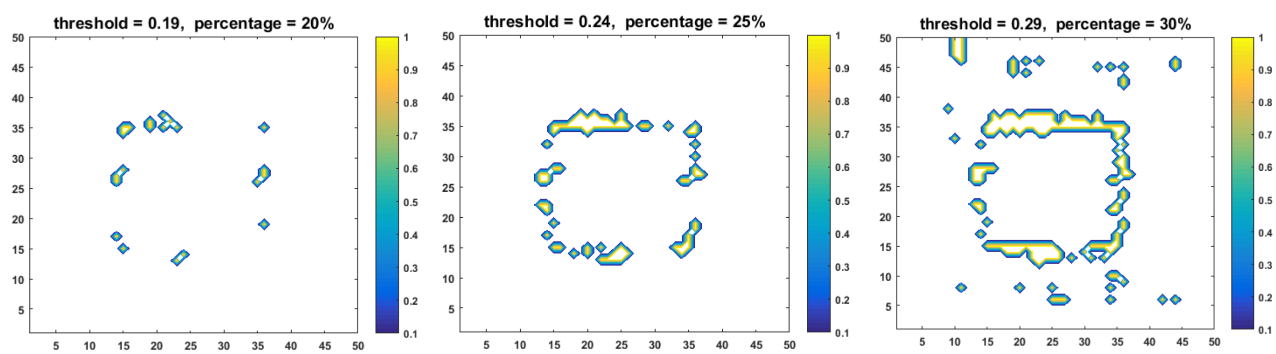


Figure 4-4: Results of Expected Distance Method on pothole with steep edges in cases of different threshold

- **Step 1:** Verify normal distribution

In Figure 4-5, the blue data points indicated as "+" fall approximately on a straight line which indicates the expected distance, and the radius in polar system fall in a normal distribution. Therefore, this Edge Stitching method can be used on this data set.

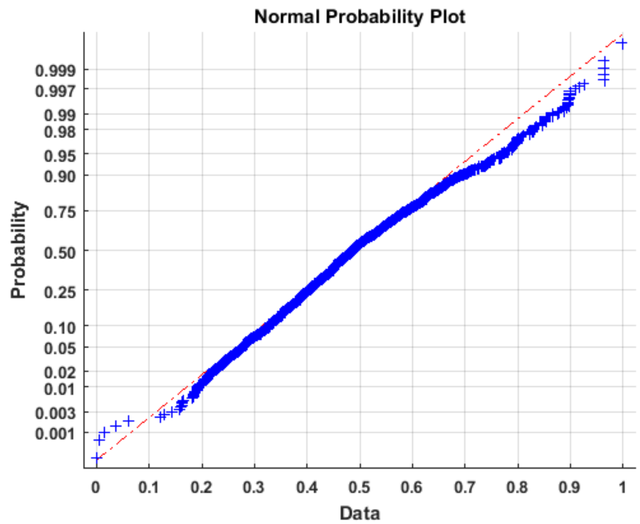


Figure 4-5: Normal probability plot of expected distance of possible edges on pothole with steep edges

- **Step 2:** Find a center in Cartesian coordinates

The edge is then described by a cluster of points with a center in Cartesian coordinates, as shown in **Error! Reference source not found.**. Apparently, outliers exist in all figures. To better characterize the edge points and obtain an accurate approximation of event, outliers need to be removed beforehand. Outliers of edge points are filtered out by picking only edge nodes located within one standard deviation. The edge points remained after outliers are filtered are the true edge points among all possible edge points.

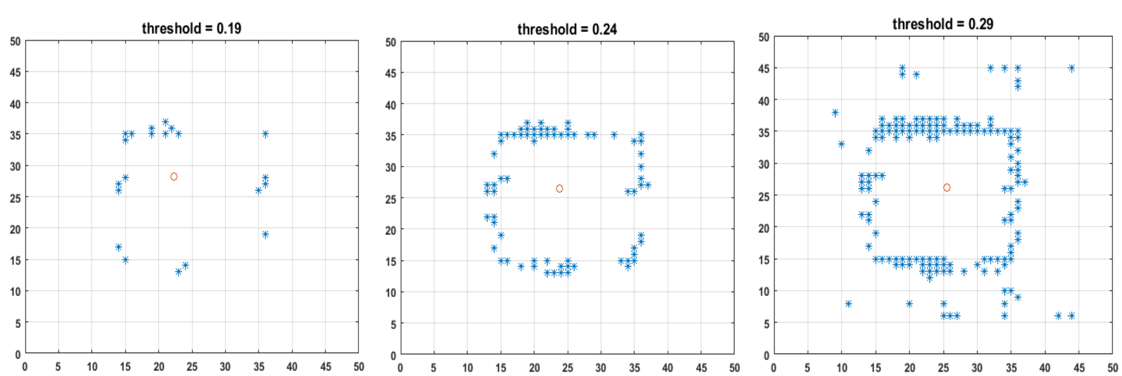


Figure 4-6: Edge points with center in Cartesian coordinates in cases of threshold 0.19, 0.24, and 0.29

- **Step 3:** Filter outliers in Cartesian coordinates

Error! Reference source not found. shows edge points after outliers are removed. Obviously, a more accurate center is found to represent the edge points with the effect of outliers being diminished. The rightmost figure still keeps some points that are a little far away from the cluster of edge points. These points, however, are not considered outliers since they are not quite far away from the center when compared with those edge points removed.

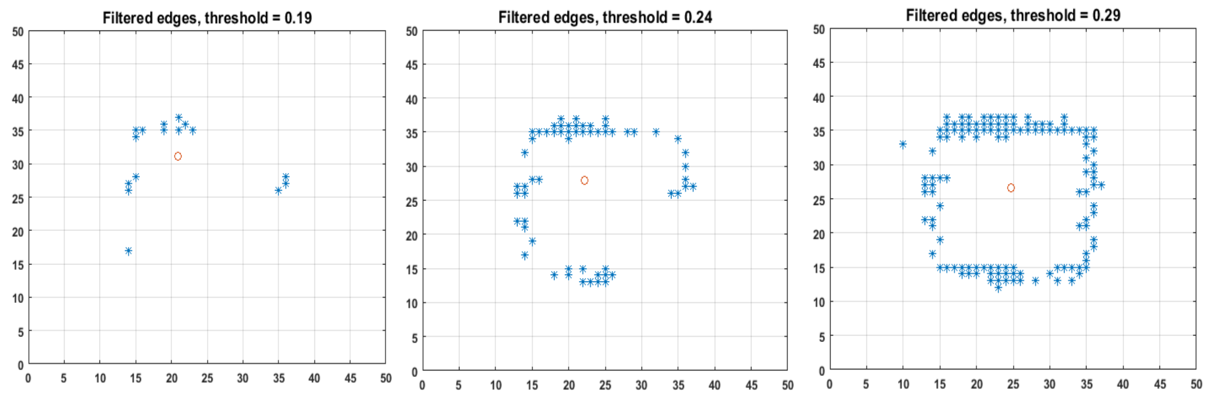


Figure 4-7: Edge points with outliers removed in cases of threshold 0.19, 0.24, and 0.29

- **Step 4:** Convert from Cartesian coordinates to polar coordinates

To further characterize the edge, a conversion is made from Cartesian coordinates to polar coordinates to describe every edge point by an angle and radius relative to the center. A translation is made to fix the center to origin (0, 0) so that the cluster of points will appear within a circle in the polar plot. **Error! Reference source not found.** displays the results of edge points converted into polar coordinates.

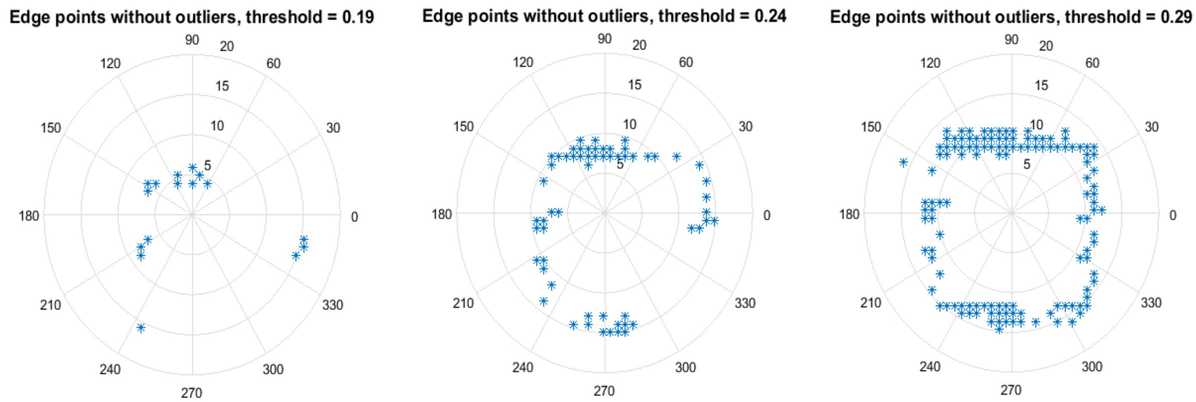


Figure 4-8: Edge points converted into polar coordinates in cases of threshold 0.19, 0.24, and 0.29

- **Step 5:** Discretize edge points and find edge representatives on every angle

The angle of 2π is divided by 30 to obtain 30 increments to approximate an event. **Error! Reference source not found.** shows the approximation of discretized edge points before the gap increments are filled, **Error! Reference source not found.** shows the approximation of discretized edge points after the gap increments are filled.

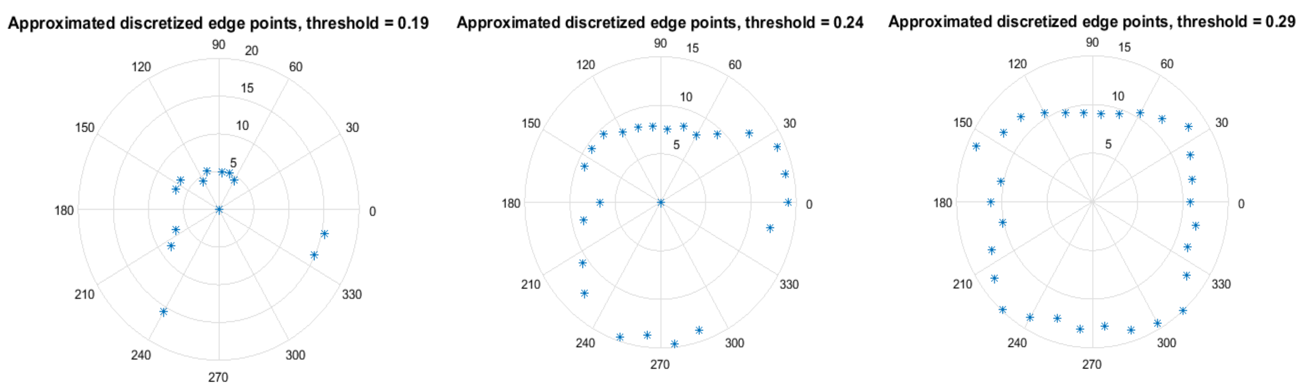


Figure 4-9: Approximation of discretized edge points in cases of threshold 0.19, 0.24, and 0.29

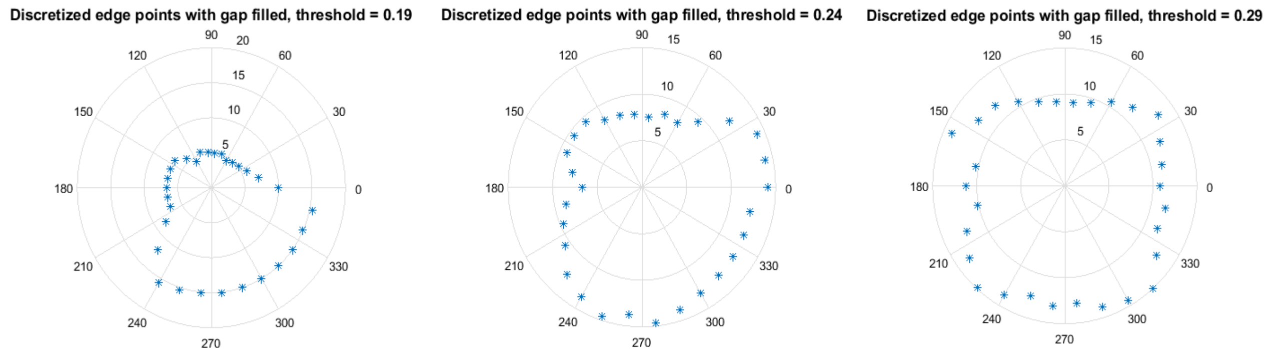


Figure 4-10: Approximation of discretized edge points with gap filled in cases of threshold of 0.19, 0.24, and 0.29

- **Step 6:** Connect edge points to event and convert back to Cartesian coordinates

Finally, all edge points can be connected to an event and displayed in Cartesian coordinates according to the converted center, as shown in **Error! Reference source not found.**.

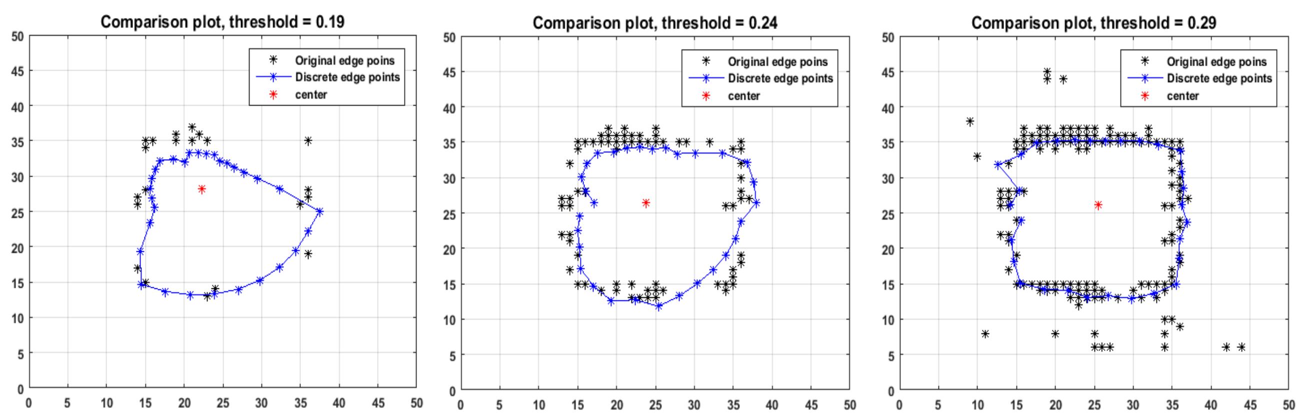


Figure 4-11: Comparison between original edge points and estimated edge points on pothole with steep edges developed from Expected Distance Edge Detection Method

Results from Expected Distance Edge Detection Method on pothole with gradual edges

Three cases of threshold are tested on pothole with gradual edges (as described in Section 3.4.2) and their results are used in this Edge Stitching method. The data store in **Error! Reference source not found.** is used for the analysis of generating edges with a center, fixed angle, and respective radius in both polar coordinates and Cartesian coordinates. The image has a suspected edge indicated by the high intensity blue shapes.

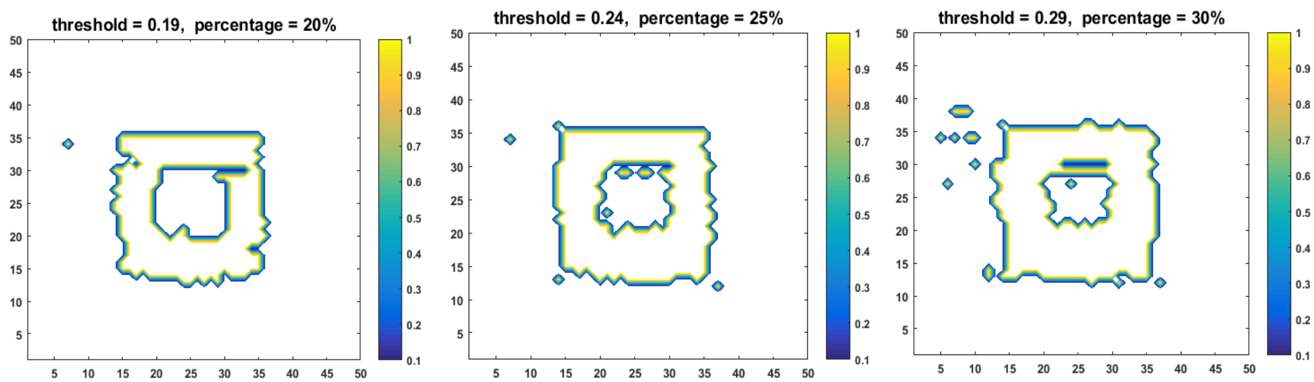


Figure 4-12: Results of Expected Distance Method on pothole with gradual edges in cases of different threshold

In Figure 4-13, the blue data points indicated as "+" fall approximately, but not as closely as in Figure 4-5, on a straight line which indicates the expected distance, and the radius in polar system fall in a normal distribution. However, the Edge Stitching method is still used on this data set to make a comparison of same method but applied on a different data set.

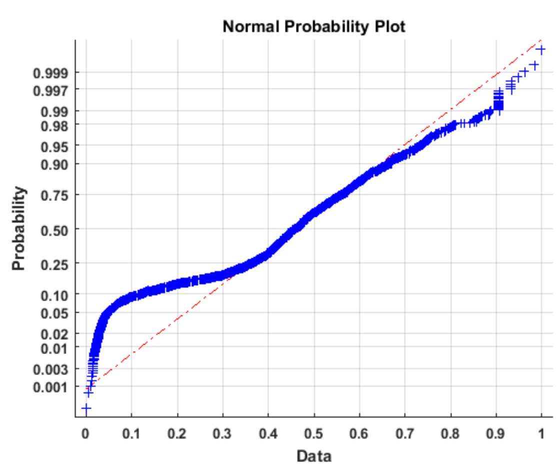


Figure 4-13: Normal probability plot of expected distance of possible edges on pothole with gradual edges

Since this example of pothole with gradual edges goes through exactly same procedure as the example of pothole with steep edges, only the results of final step of the comparison plots are shown here. All edge points can be connected to an event and displayed in Cartesian coordinates according to the converted center, as shown in **Error! Reference source not found.**.

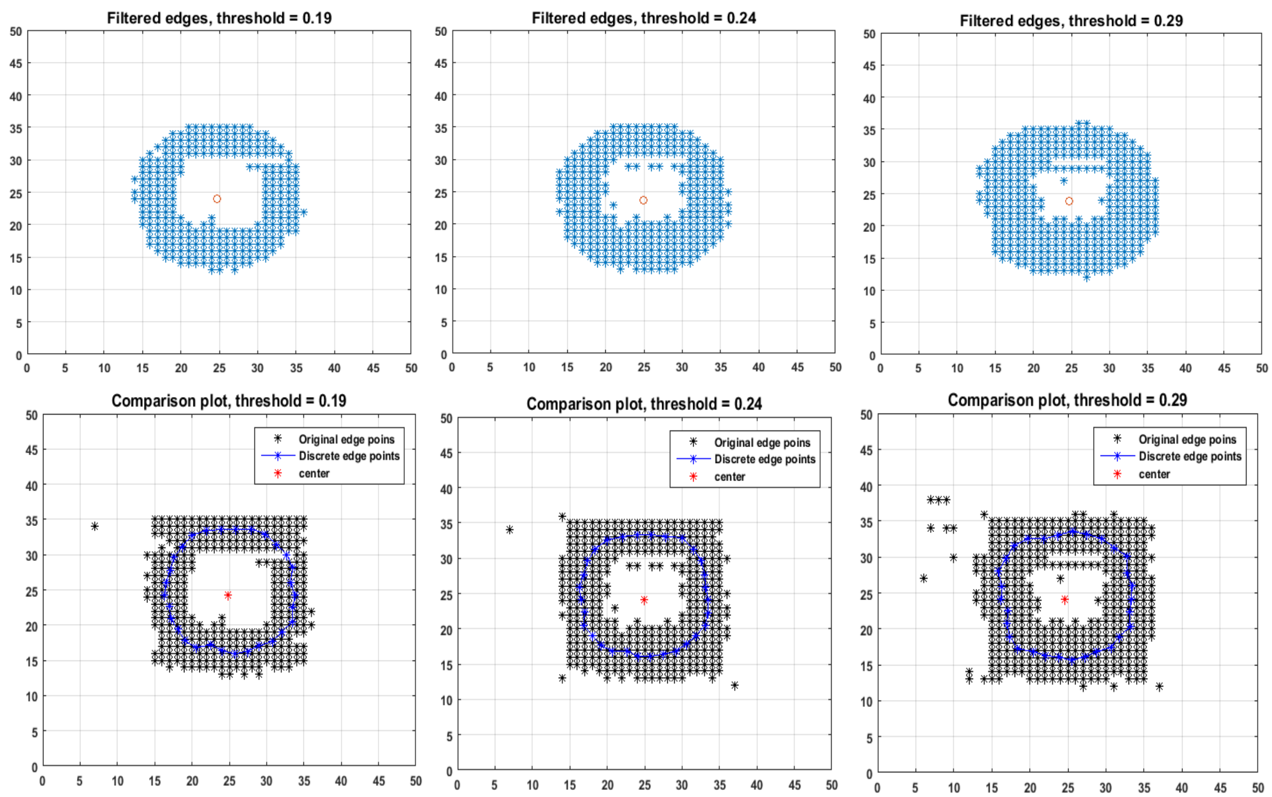


Figure 4-14: Comparison between original edge points and estimated edge points on pothole with gradual edges developed from Expected Distance Edge Detection Method

Results from Number Difference Edge Detection Method on pothole with steep edges

Three cases of threshold are tested on pothole with steep edges (as described in Section 3.4.3) and the test results are used in this Edge Stitching method. The data store in Figure 4-15 is used for the analysis of generating edges with a center, fixed angle, and respective radius in both polar coordinates and Cartesian coordinates. The image has a suspected edge indicated by the high intensity blue shapes.

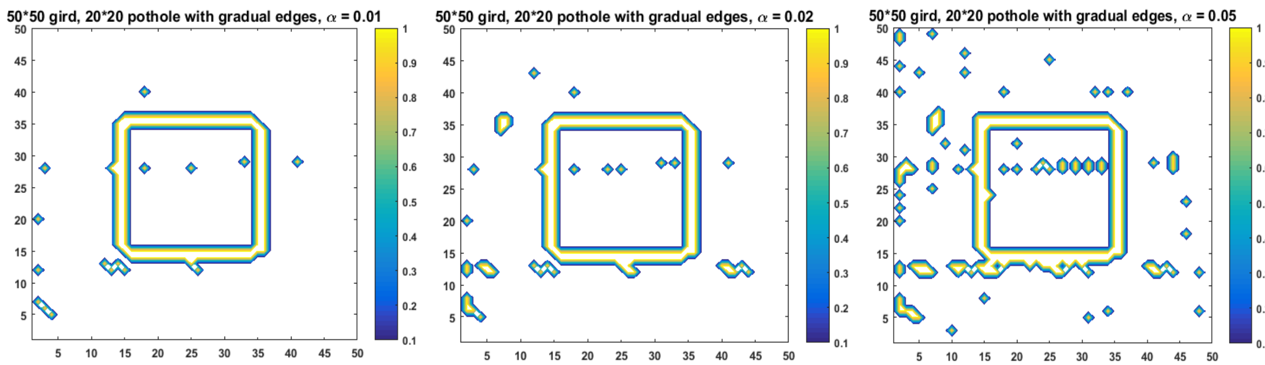


Figure 4-15: Results of Number Difference Method on pothole with steep edges in cases of different significance level

In **Error! Reference source not found.**, Normal Probability Plot is firstly used to investigate whether the data of radius obtained from Number Difference matrix exhibit the standard normal curve. To make a clear comparison, the result of detected edges on pothole with steep edges are shown in Figure 4-16.

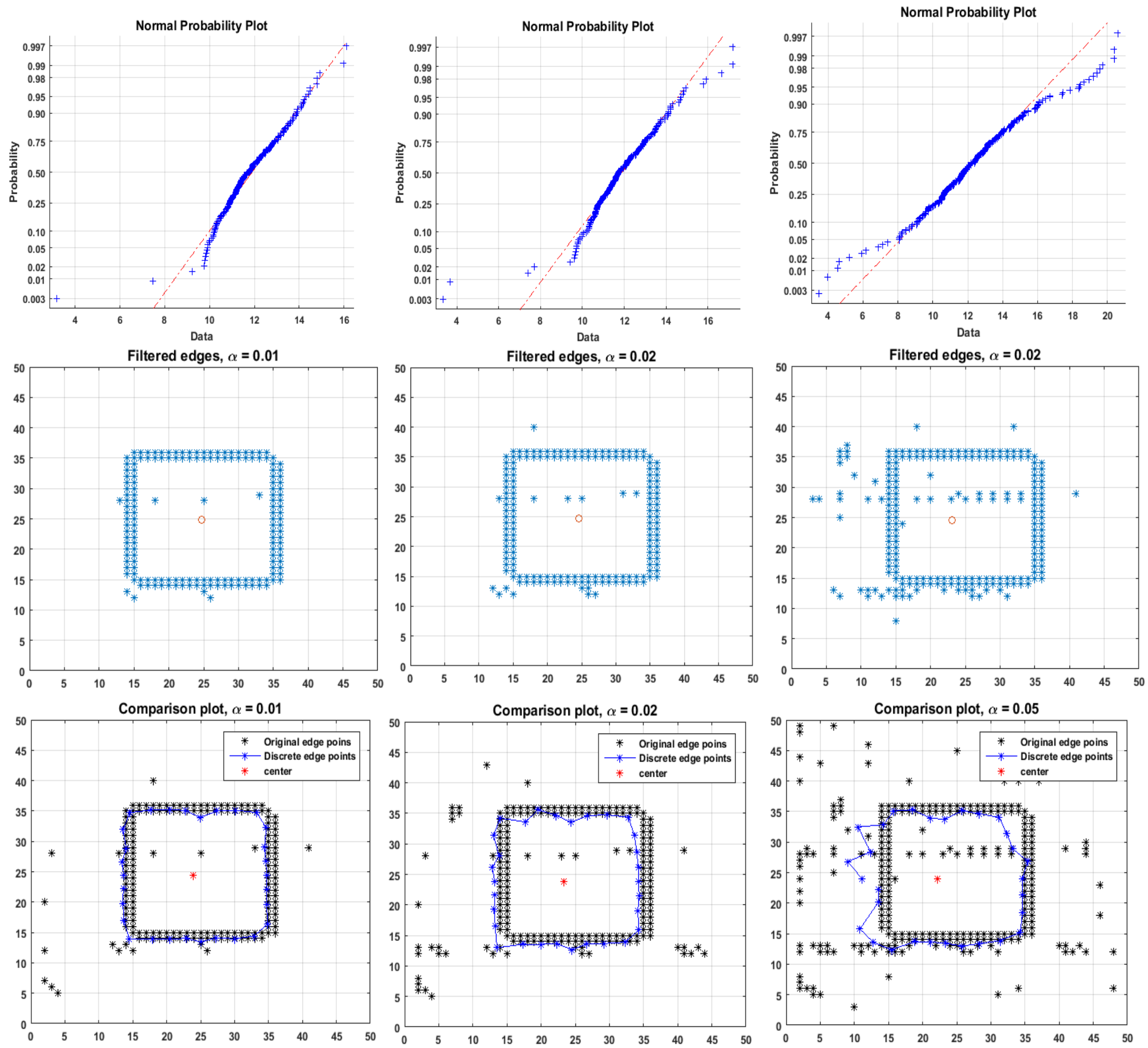


Figure 4-16: Comparison between original edge points and estimated edge points on pothole with steep edges developed from Number Difference Edge Detection Method

Results from Number Difference Edge Detection Method on pothole with gradual edges

Three cases of different significance level are tested in Section 3.4.3 and their results are used in this Edge Stitching method. The data store in Figure 4-17 is used for the analysis of generating edges with a center, fixed angle, and respective radius in both polar coordinates and Cartesian coordinates. The image has a suspected edge indicated by the high intensity blue shapes.

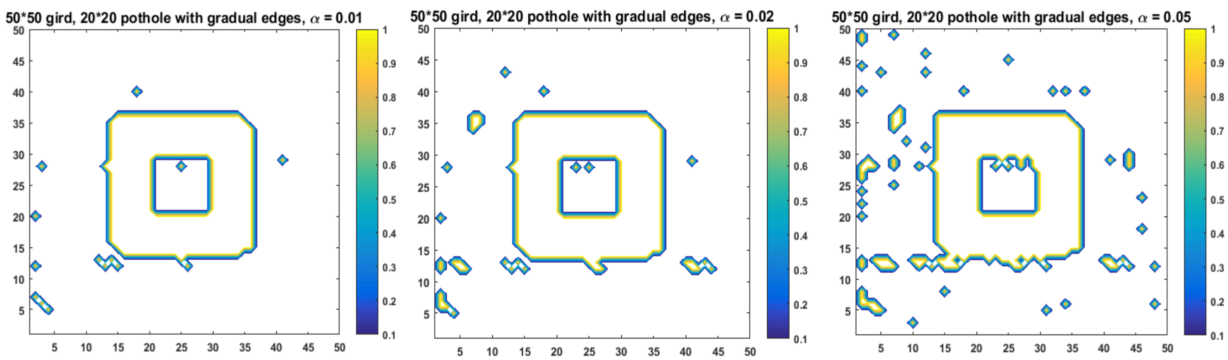


Figure 4-17: Results of Number Difference Method on pothole with gradual edges in cases of different significance level

In Figure 4-18, Normal Probability Plot is firstly used to investigate whether the data of radius obtained from Number Difference matrix exhibit the standard normal curve. To make a clear comparison, the result of detected edges on pothole with gradual edges are shown in Figure 4-18.

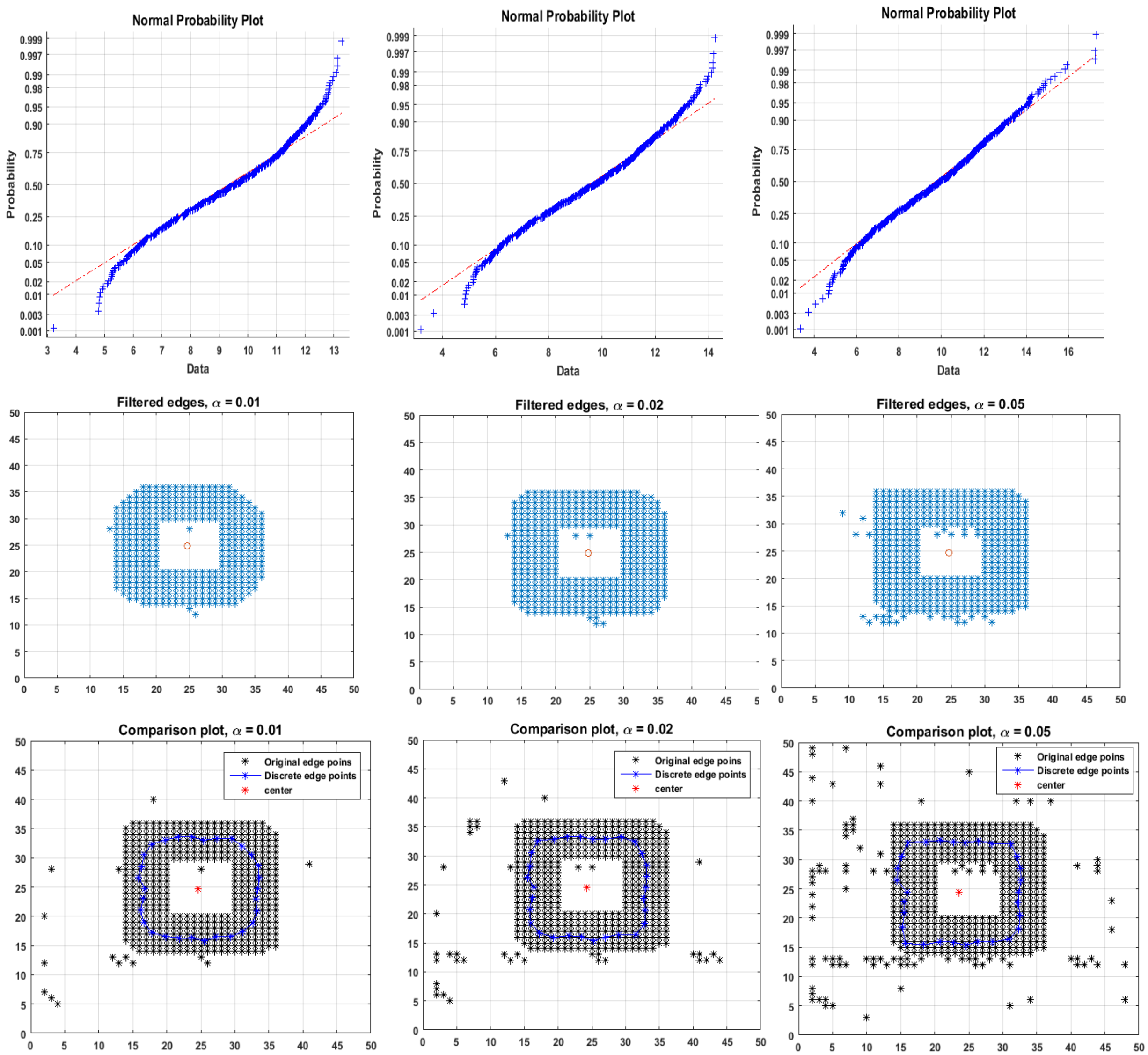


Figure 4-18: Comparison between original edge points and estimated edge points on pothole with steep edges developed from Number Difference Edge Detection Method

4.4.2 Filter by standard deviation of radius distribution, connect edges by least square curve fitting

The original pothole with steep edges depicted in Figure 3-11 and Figure 3-12. The original pothole with gradual edges depicted in Figure 3-14 and Figure 3-15. Both pothole are approximately squares of 20×20 located in a 50×50 grid.

<u>Results from Expected Distance Edge Detection Method on pothole with steep edges</u>

Three cases of different threshold are tested in Section 3.4.2 and their results are used in this Edge Stitching method. The procedures of filtering edges by standard deviation of radius distribution have been described in Section 4.4.3. The results of filtering edges are listed in Step 1.

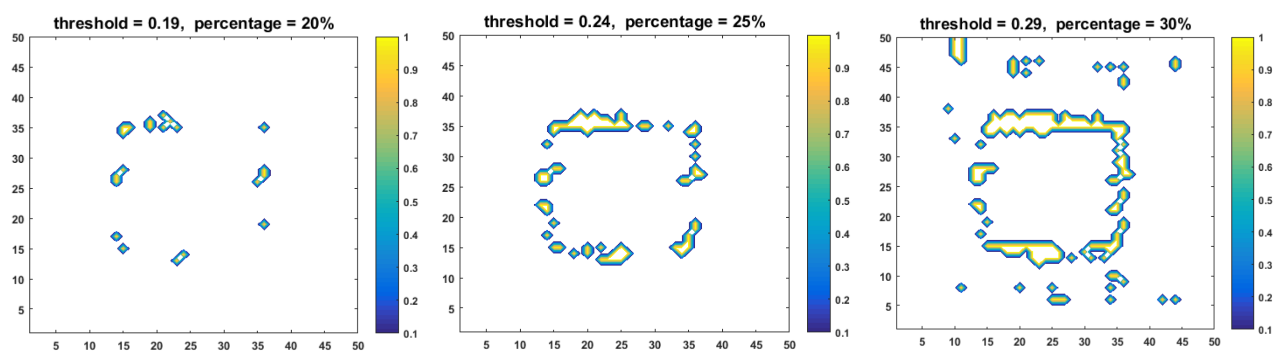


Figure 4-19: Results of Expected Distance Method on pothole with steep edges in cases of different threshold

The data store in Figure 4-19 is used for the analysis of generating edges with a center, fixed angle, and respective radius in both polar coordinates and Cartesian coordinates. The image has a suspected edge indicated by the high intensity blue shapes.

- **Step 1:** Obtain filtered edges in polar coordinates

Figure 4-20 displays the results of edge points converted into polar coordinates, with every edge point described by an angle and radius relative to the origin (0, 0).

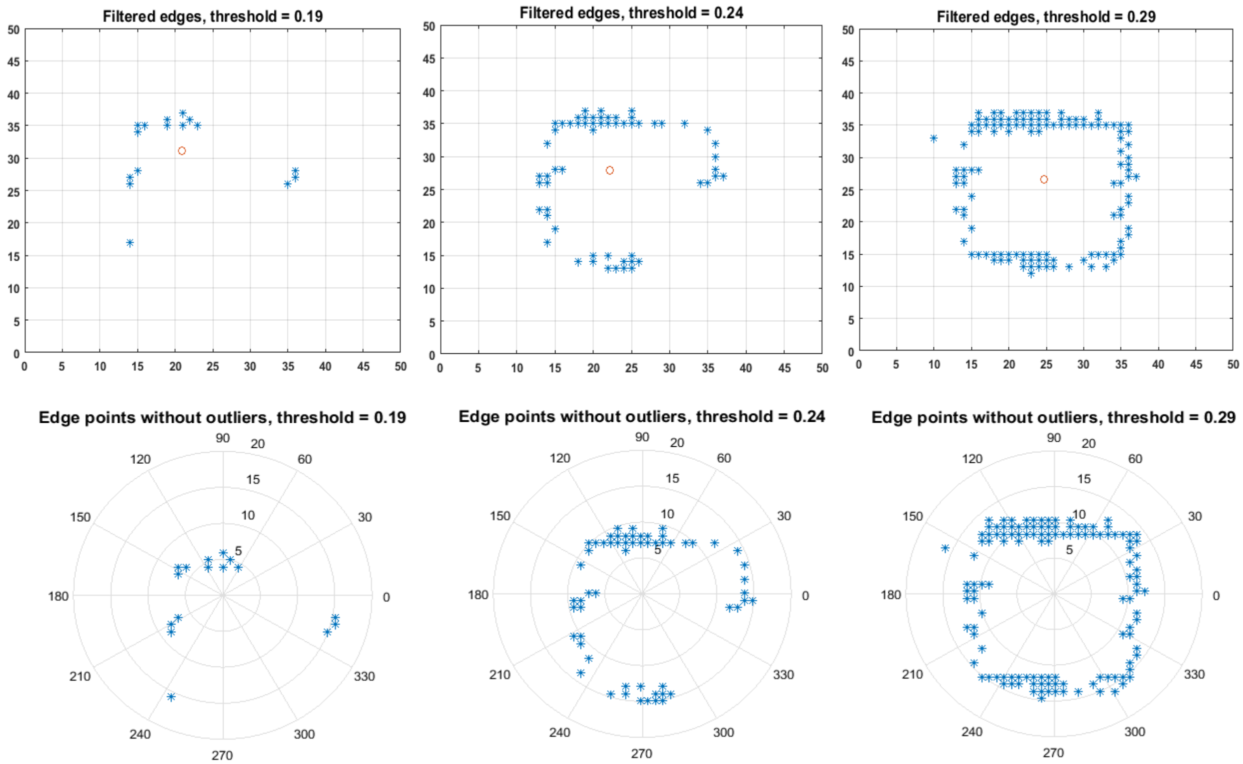


Figure 4-20: Filtered edge points in polar coordinates in cases of threshold 0.19, 0.24, and 0.29

- **Step 2:** Least square curve fitting of alpha vs. radius plot in Cartesian coordinates

Alpha and radius are obtained from polar coordinates and plotted in Cartesian coordinates. In Figure 4-21, red circles represent the original filtered edge points. Colorful lines represent least square fitting curve with the increased constraints determined in Fourier series expansion (as explained in Section 4.3.2.2.

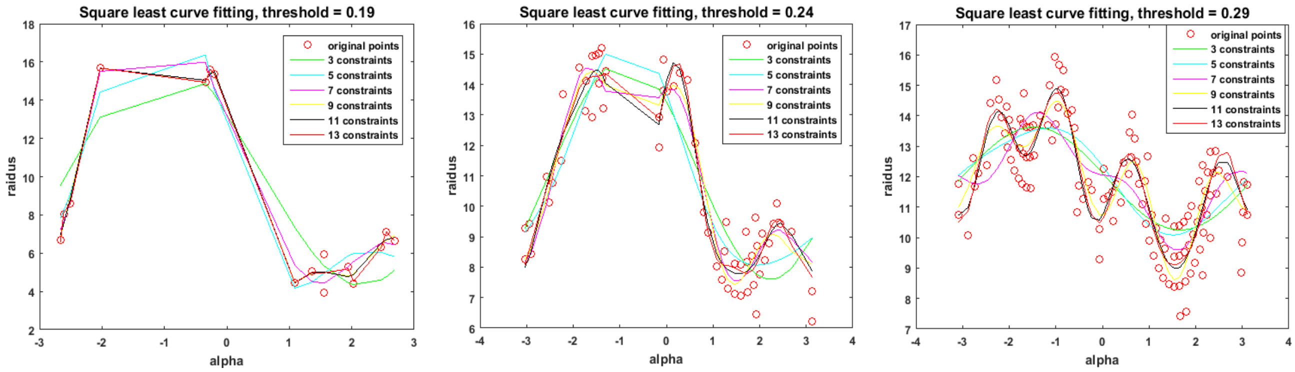


Figure 4-21: Radius of filtered edge points of pothole with steep edges vs. alpha and least square curve fitting in Cartesian coordinates (Expected Distance Method)

- **Step 3:** Obtain final event in Cartesian coordinates

Finally, edge points on event can be obtained from the fitting curve found in Step 2 with different amount of constraints, as shown in Figure 4-22.

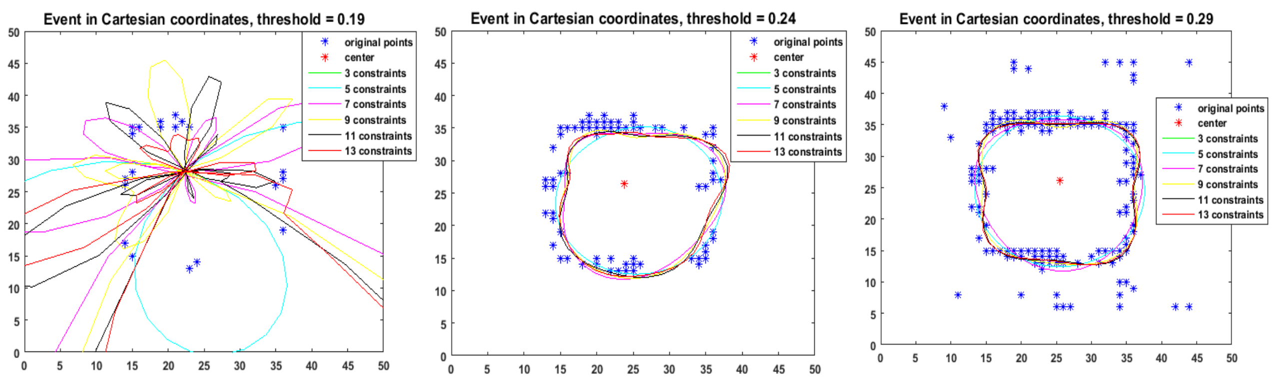


Figure 4-22: Comparison between original edge points and estimated edge points on pothole with steep edges developed from Expected Distance Edge Detection Method

Results from Expected Distance Edge Detection Method on pothole with gradual edges

Three cases of difference expected distance threshold are tested in Section 3.4.2 and their results are used in this Edge Stitching method. Since this example of pothole with gradual edges goes through exactly same procedure as the example of pothole with steep edges, only the results of the final step for the comparison plots are provided here.

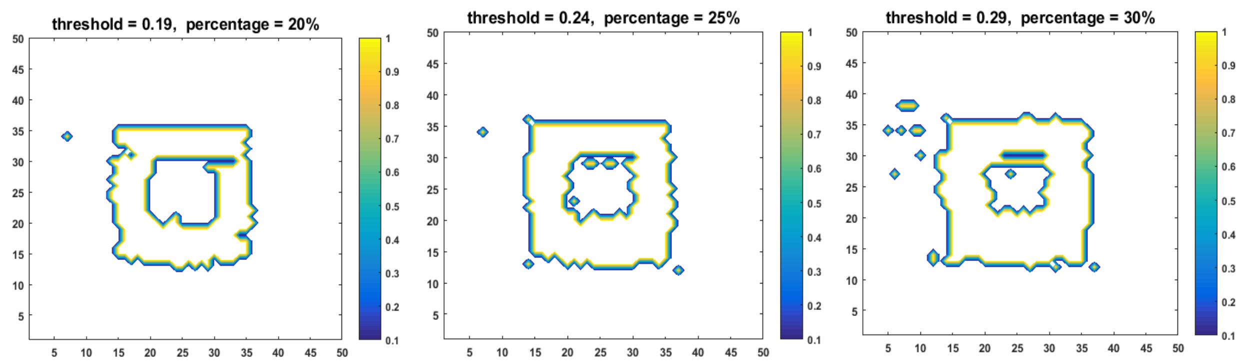


Figure 4-23: Results of Expected Distance Method on pothole with gradual edges in cases of different threshold

The data store in Figure 4-23 is used for the analysis of generating edges with a center, fixed angle, and respective radius in both polar coordinates and Cartesian coordinates. The image has a suspected edge indicated by the high intensity blue shapes.

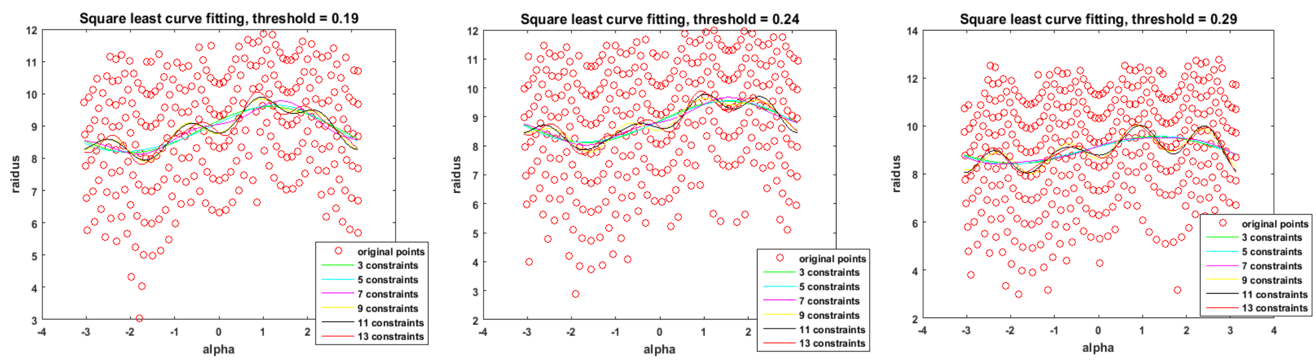


Figure 4-24: Radius of filtered edge points of pothole with steep edges vs. alpha and least square curve fitting in Cartesian coordinates (Expected Distance Method)

Alpha and radius are obtained from polar coordinates and plotted in Cartesian coordinates. In Figure 4-24, red circle represent the original filtered edge points while colorful lines represent least square fitting curve with the increased constraints determined in Fourier series expansion (as explained in Section 4.3.2.2).

Finally, edge points on event can be obtained from the fitting curve found in Figure 4-24 with different amount of constraints, as shown in Figure 4-25.

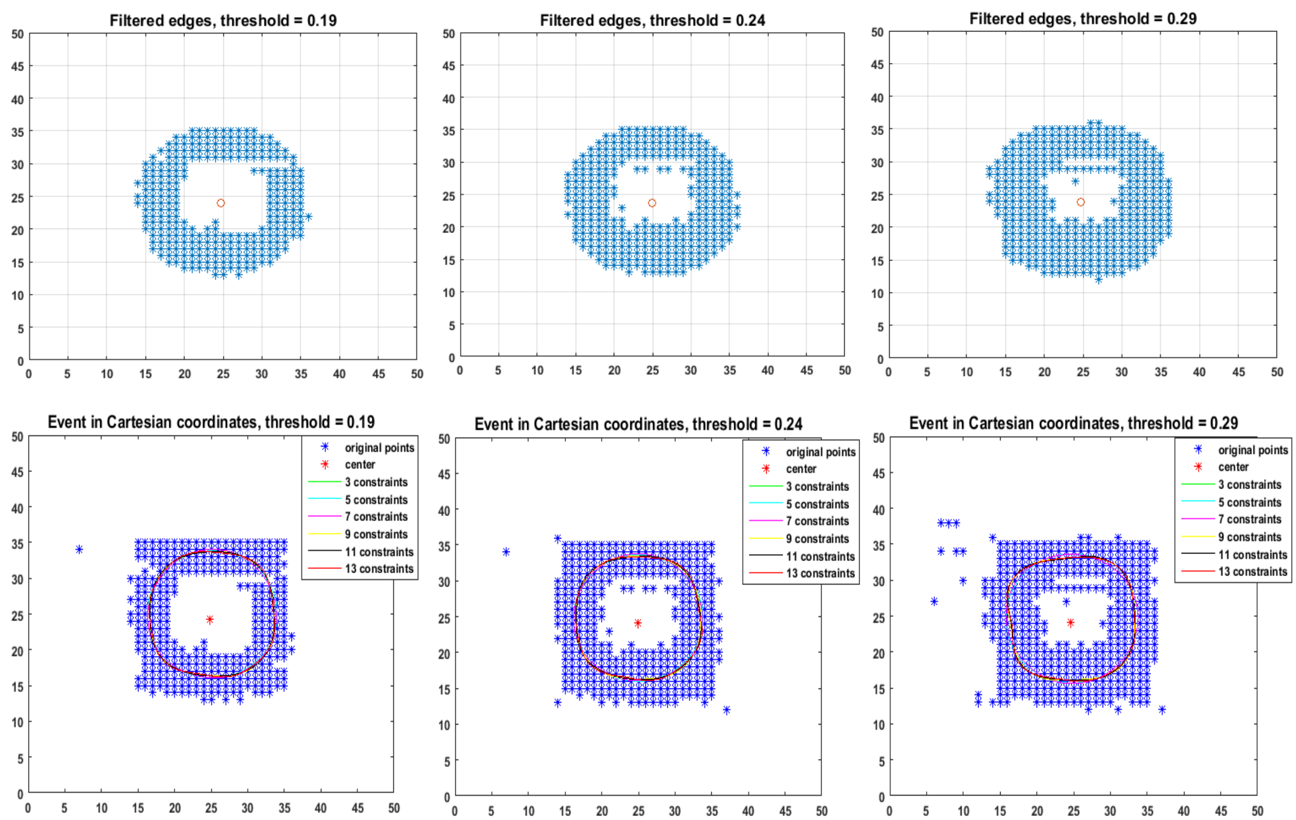


Figure 4-25: Comparison between original edge points and estimated edge points on pothole with gradual edges developed from Number Difference Edge Detection Method

Results from Number Difference Edge Detection Method on pothole with steep edges

Three cases of different significance level are tested in Section 3.4.3 and their results are used in this Edge Stitching method.

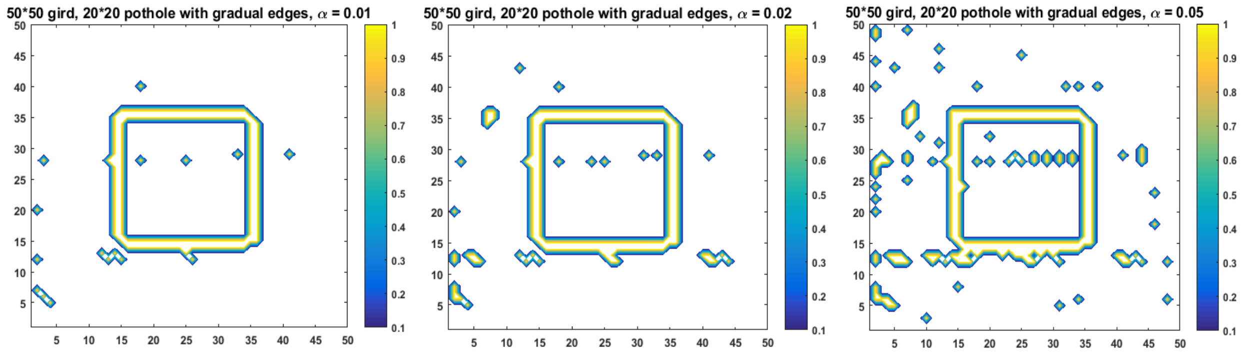


Figure 4-26: Results of Number Difference Method on pothole with steep edges in cases of difference significance level

The data store in Figure 4-26 is used for the analysis of generating edges with a center, fixed angle, and respective radius in both polar coordinates and Cartesian coordinates. The image has a suspected edge indicated by the high intensity blue shapes.

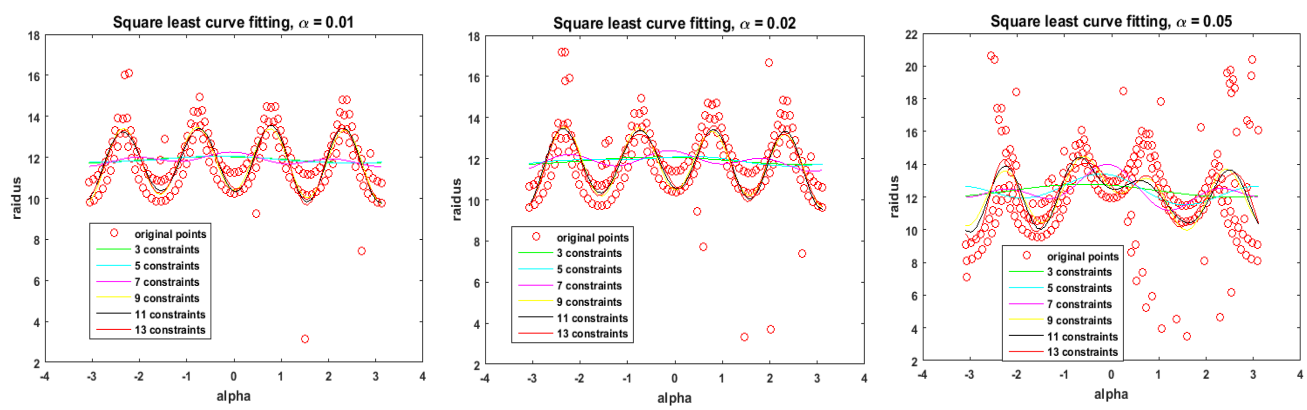


Figure 4-27: Radius of filtered edge points of pothole with steep edges vs. alpha and least square curve fitting in Cartesian coordinates (Number Difference Method)

Alpha and radius are obtained from polar coordinates and plotted in Cartesian coordinates. In Figure 4-27, red circles represent the original filtered edge points while colorful lines represent least square fitting curve with the increased constraints determined in Fourier series expansion (as explained in Section 4.3.2.2).

Finally, edge points on event can be obtained from the fitting curve found in Figure 4-27 with different amount of constraints, as shown in Figure 4-28.

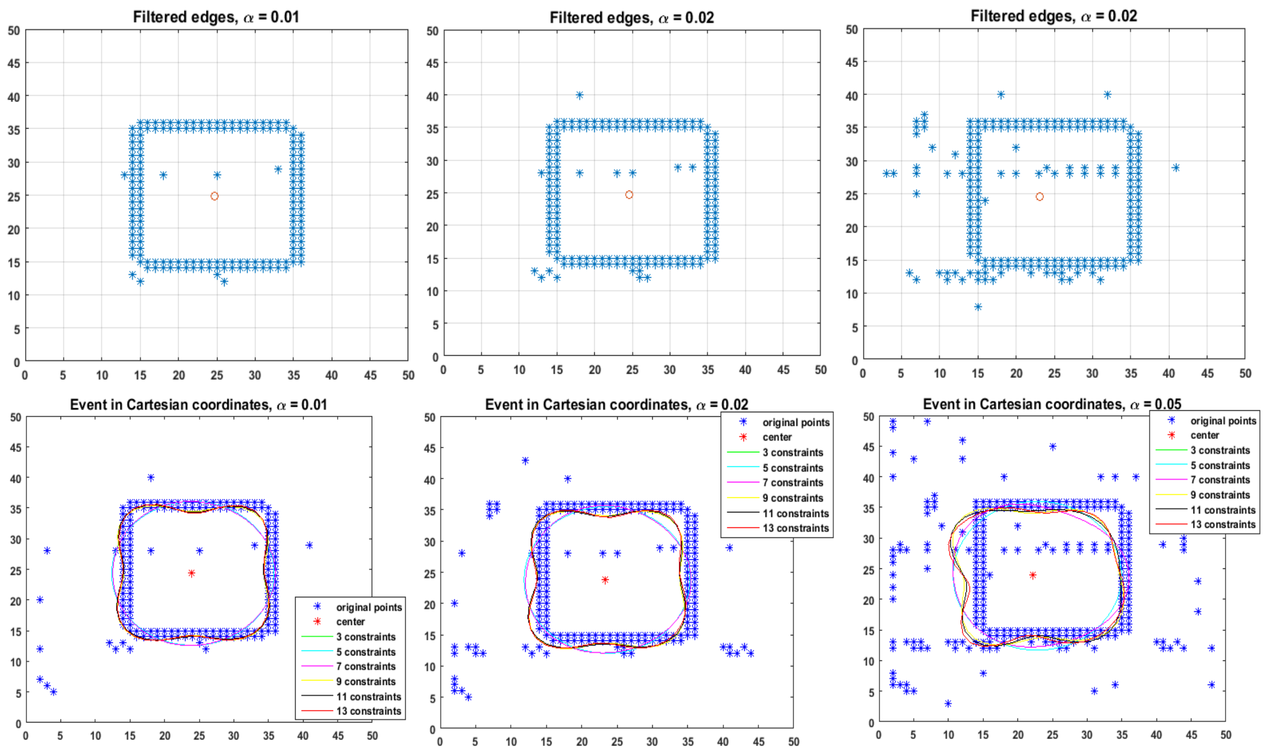


Figure 4-28: Comparison between original edge points and estimated edge points on pothole with steep edges developed from Number Difference Edge Detection Method

Results from Number Difference Edge Detection Method on pothole with gradual edges

Three cases of different significance level are tested in Section 3.4.3 and their results are used in this Edge Stitching method.

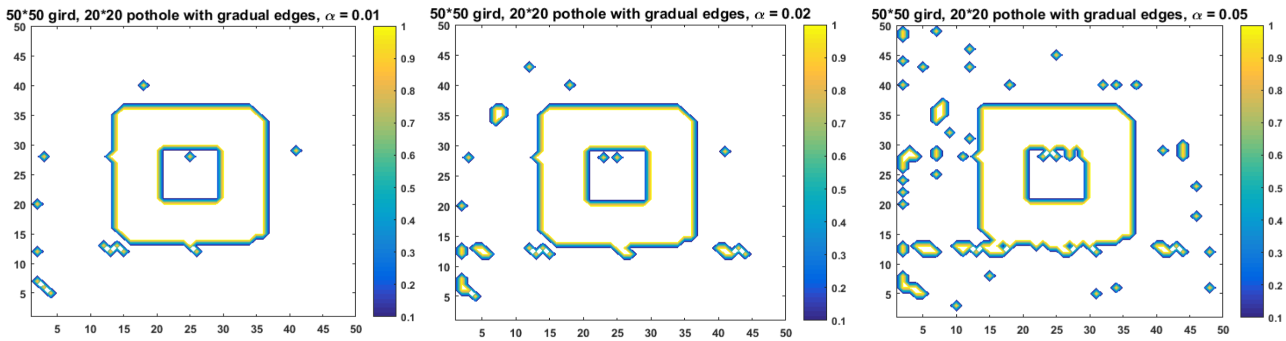


Figure 4-29: Results of Number Difference Method on pothole with gradual edges in cases of different significance level

The data store in Figure 4-29 is used for the analysis of generating edges with a center, fixed angle, and respective radius in both polar coordinates and Cartesian coordinates. The image has a suspected edge indicated by the high intensity blue shapes.

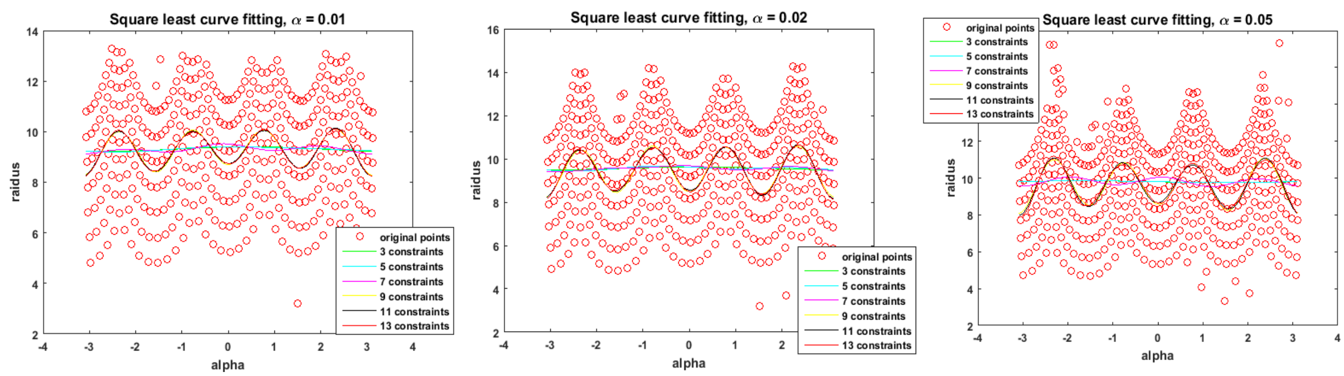


Figure 4-30: Radius of filtered edge points of pothole with gradual edges vs. alpha and least square curve fitting in Cartesian coordinates (Number Difference Method)

Alpha and radius are obtained from polar coordinates and plotted in Cartesian coordinates. In Figure 4-30, red circles represent the original filtered edge points while colorful lines represent least square fitting curve with the increased constraints determined in Fourier series expansion (as explained in Section 4.3.2.2).

Finally, edge points on event can be obtained from the fitting curve found in Figure 4-30 with different amount of constraints, as shown in Figure 4-31.

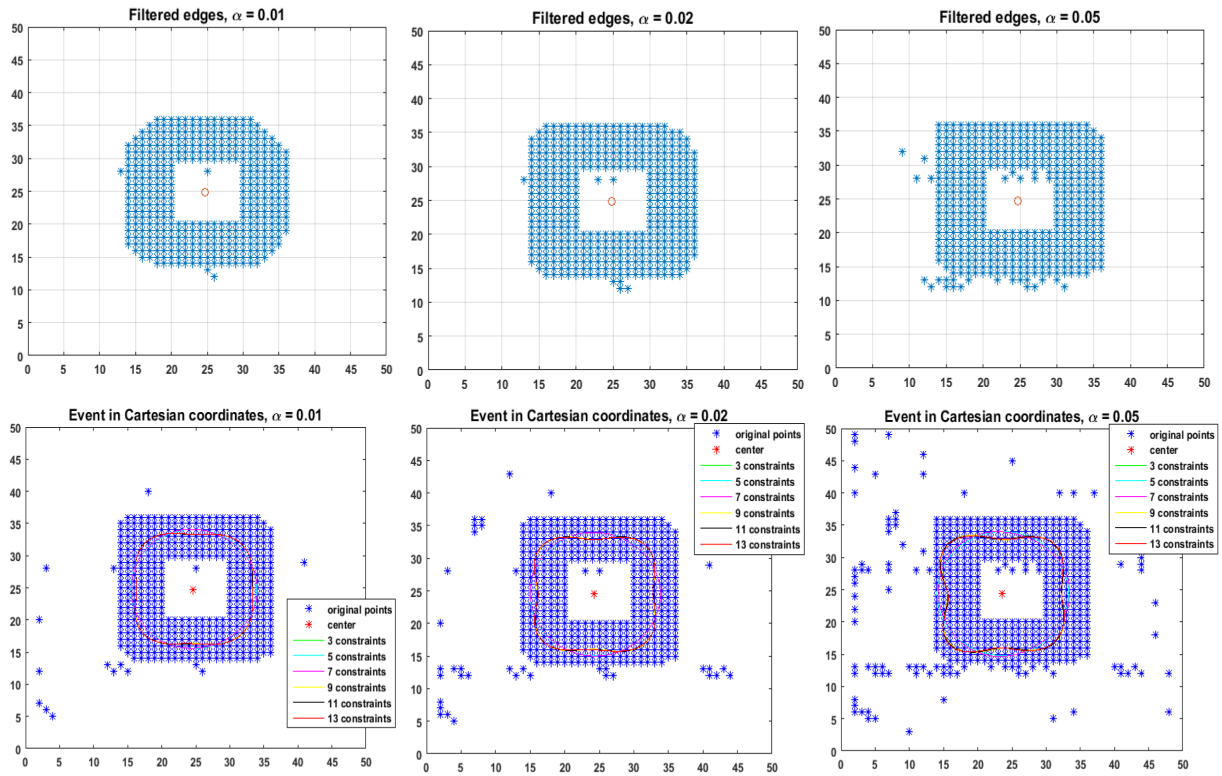


Figure 4-31: Comparison between original edge points and estimated edge points on pothole with gradual edges developed from Number Difference Edge Detection Method

4.4.3 Filter by looking local points, connect edges by discretizing angels

The original pothole with steep edges depicted in Figure 3-11 and Figure 3-12. The original pothole with gradual edges depicted in Figure 3-14 and Figure 3-15. Both potholes are approximately squares of 20×20 located in a 50×50 grid.

Results from Expected Distance Edge Detection Method on pothole with steep edges

Three cases of different threshold are tested in Section 3.4.2 and their results are used in this Edge Stitching method. The data store in Figure 4-32 is used for the analysis of generating edges with a center, fixed angle, and respective radius in both polar coordinates and Cartesian coordinates. The image had a suspected edge indicated by the high intensity blue shapes.

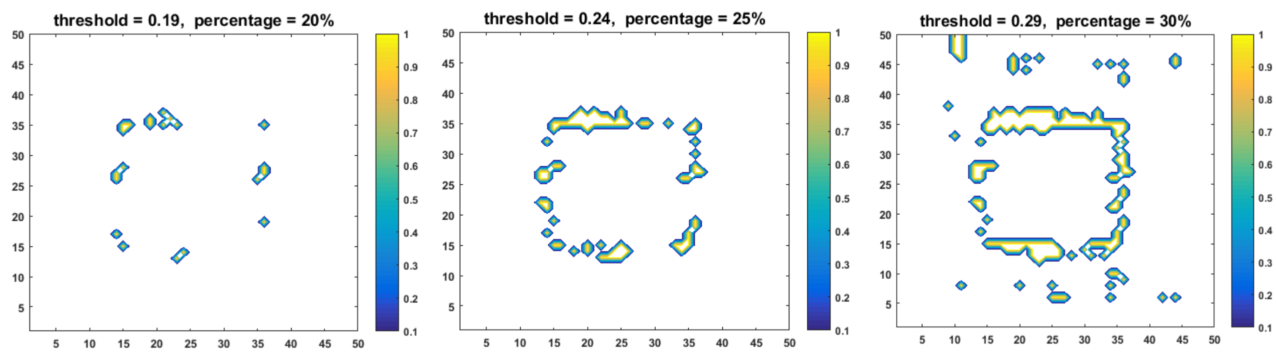


Figure 4-32: Results of Expected Distance Method on pothole with steep edges in cases of different threshold

As stated in Section 4.3.1.2, filtering edge by looking at local points is carried out by looking at each possible edge point itself. The results are displayed in Figure 4-33.

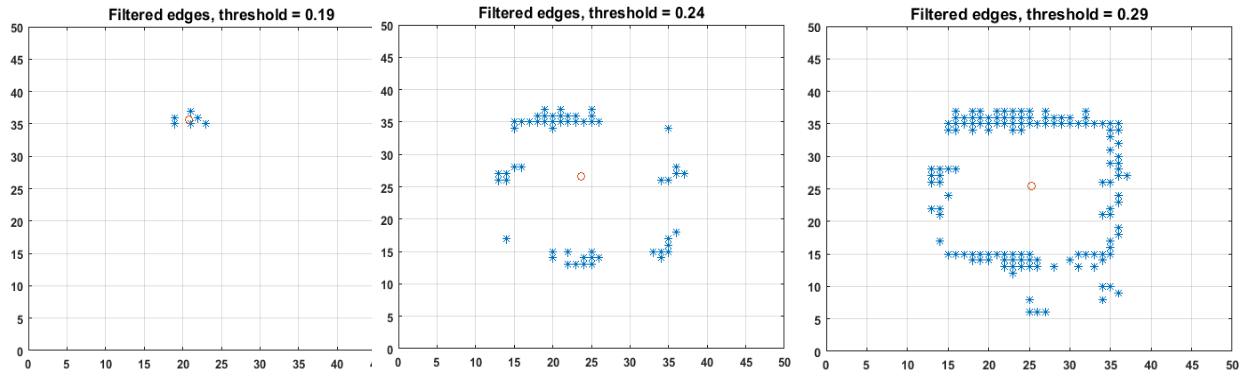


Figure 4-33: Filtered edge points using method of looking local edges points of pothole with steep edges in cases of threshold 0.19, 0.24, and 0.29

Given the method of connecting edge points to event has been stated in Section 4.4.1, only the results of connected event represented in Cartesian coordinates according to the converted center are shown in Figure 4-34.

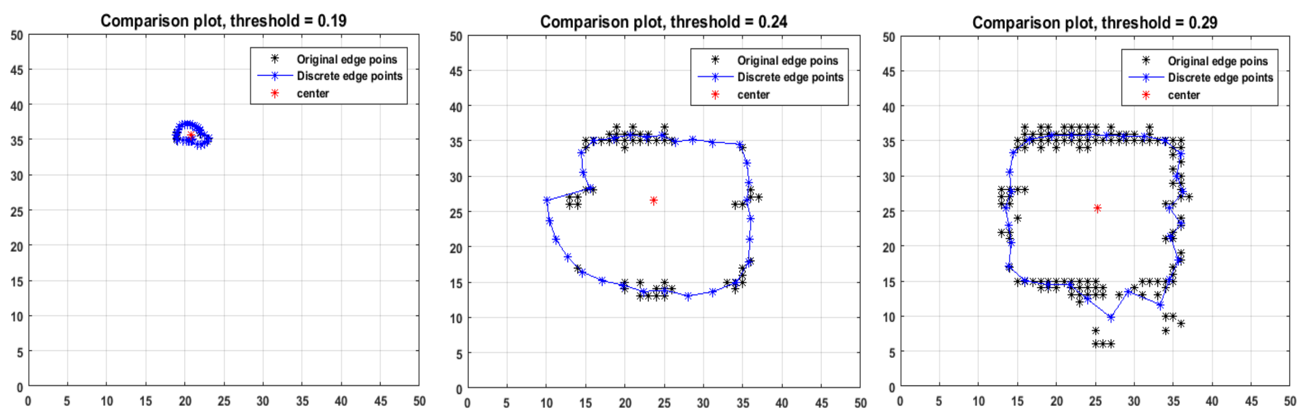


Figure 4-34: Comparison between original edge points and estimated edge points on pothole with steep edges developed from Expected Distance Edge Detection Method

Results from Expected Distance Edge Detection Method on pothole with gradual edges

Three cases of different threshold are tested in Section 3.4.2 and their results are used in this Edge Stitching method. The data store in Figure 4-32 is used for the analysis of generating edges with a center, fixed angle, and respective radius in both polar coordinates and Cartesian coordinates. The image has a suspected edge indicated by the high intensity blue shapes.

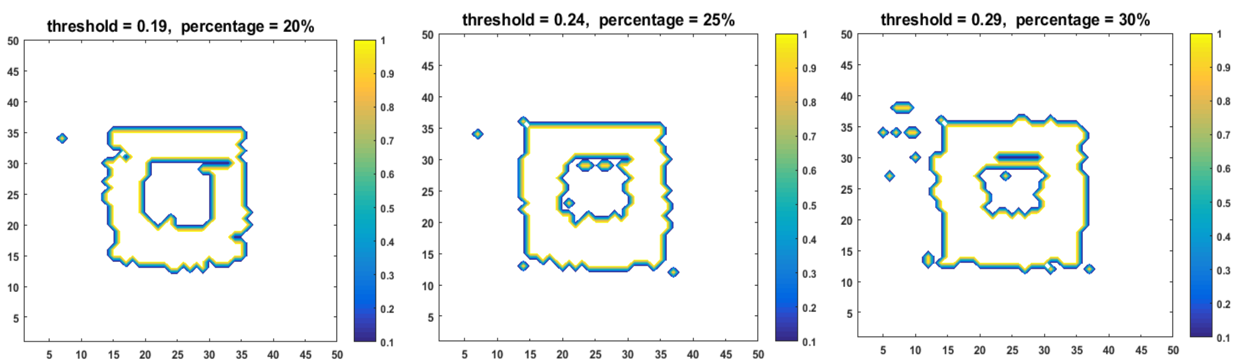


Figure 4-35: Results of Expected Distance Method on pothole with gradual edges of pothole in cases of different threshold

As stated in Section 4.3.1.2, filtering edge by looking local points is looking at each possible edge point itself. The result display as Figure 4-36.

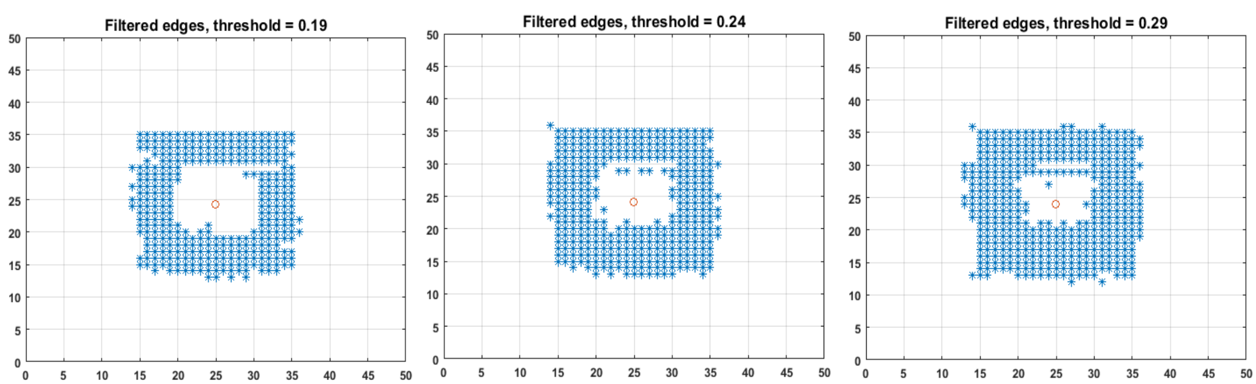


Figure 4-36: Filtered edge points using method of looking at local edges points of pothole with gradual edges in cases of threshold 0.19, 0.24, and 0.29

Given the method of connecting edge points to events has been stated in Section 4.4.1, only the results of connected event represented in Cartesian coordinates according to the converted center are shown in Figure 4-37.

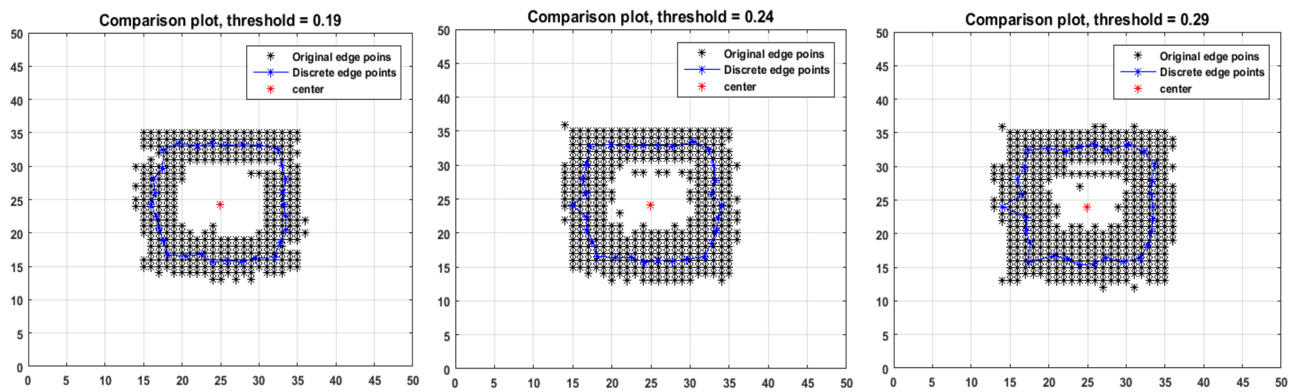


Figure 4-37: Comparison between original edge points and estimated edge points on pothole with gradual edges developed from Expected Distance Edge Detection Method

Results from Number Difference Edge Detection Method on pothole with steep edges

Three cases of different significance level are tested in Section 3.4.3 and their results are used in this Edge Stitching method. The data store in Figure 4-38 is used for the analysis of generating edges with a center, fixed angle, and respective radius in both polar coordinates and Cartesian coordinates. The image has a suspected edge indicated by the high intensity blue shapes.

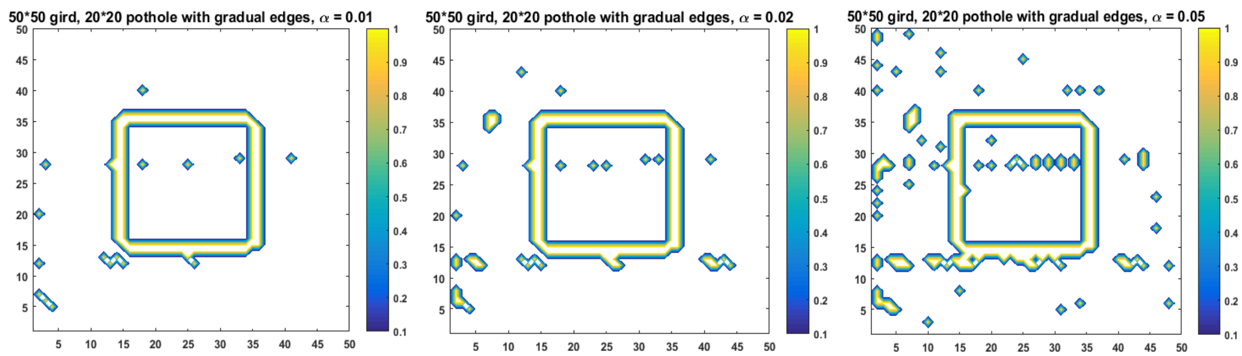


Figure 4-38: Results of Number Difference Method on pothole with steep edges of pothole in cases of different significance level

As stated in Section 4.3.1.2, filtering edge points by looking at local points is carried out by looking at each possible edge point itself. The results are displayed in Figure 4-39.

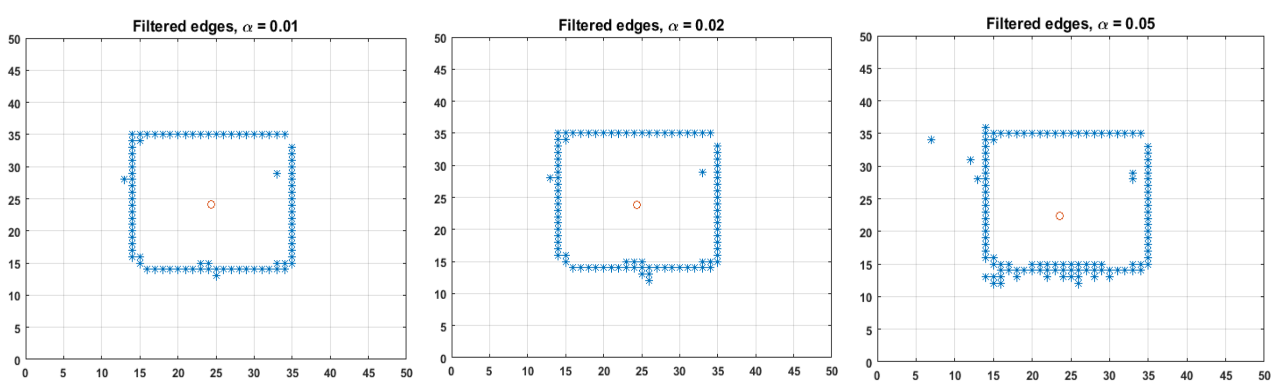


Figure 4-39: Filtered edge points using method of looking at local edges points of pothole with steep edges in cases of significance level 0.01, 0.02, and 0.05

Given the method of connecting edge points to event has been stated in Section 4.4.1, here only the results of connected event represented in Cartesian coordinates according to the converted center are shown in Figure 4-40.

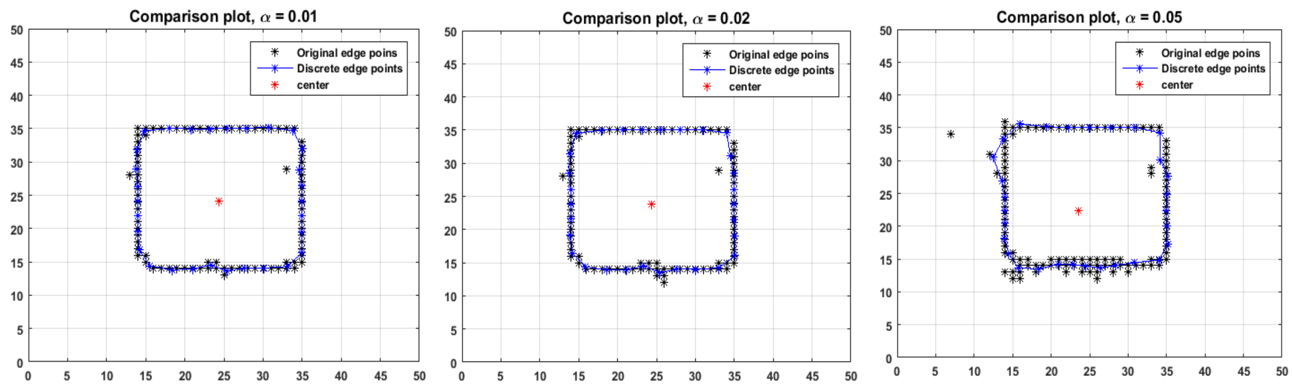


Figure 4-40: Comparison between original edge points and estimated edge points on pothole with steep edges developed from Number Difference Edge Detection Method

Results from Number Difference Edge Detection Method on pothole with gradual edges

Three cases of different significance level are tested in Section 3.4.3 and their results are used in this Edge Stitching method. The data store in Figure 4-41 is used for the analysis of generating edges with a center, fixed angle, and respective radius in both polar coordinates and Cartesian coordinates. The image has a suspected edge indicated by the high intensity blue shapes.

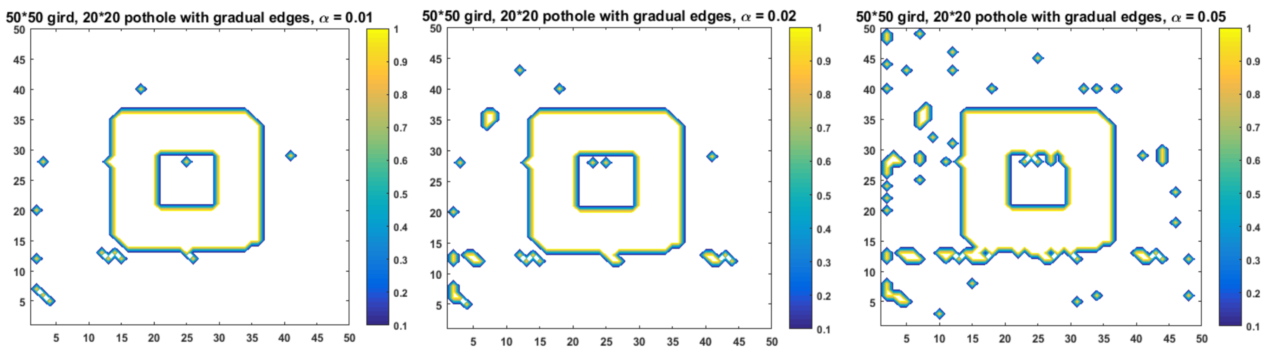


Figure 4-41: Results of Number Difference Method on pothole with gradual edges of pothole in cases of different significance level

As stated in Section 4.3.1.2, filtering edge points by looking at local points is cairred out by looking at each edge point itself. The results are displayed in Figure 4-42.

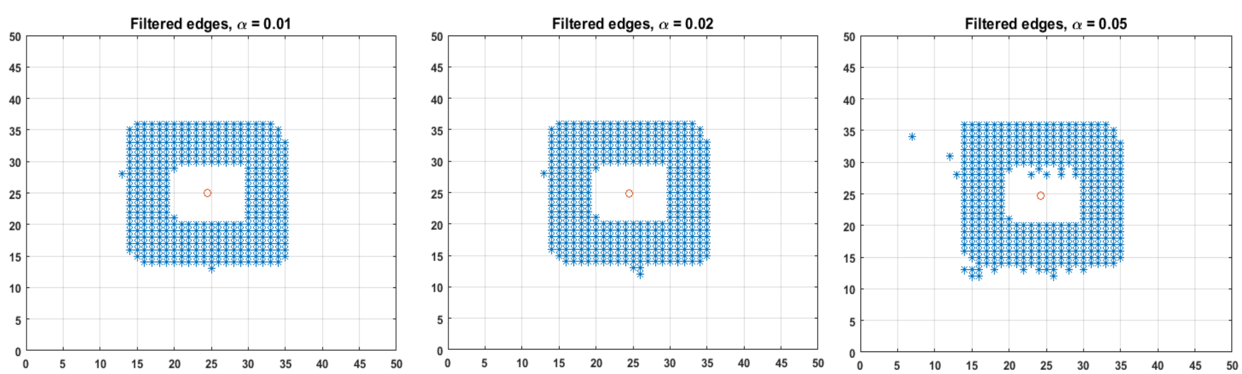


Figure 4-42: Filtered edge points using method of looking at local edges points of pothole with gradual edges in cases of significance level 0.01, 0.02, and 0.05

Given the method of connecting edge points to event has been stated in Section 4.4.1, only the results of connected event represented in Cartesian coordinates according to the converted center are shown in Figure 4-43.

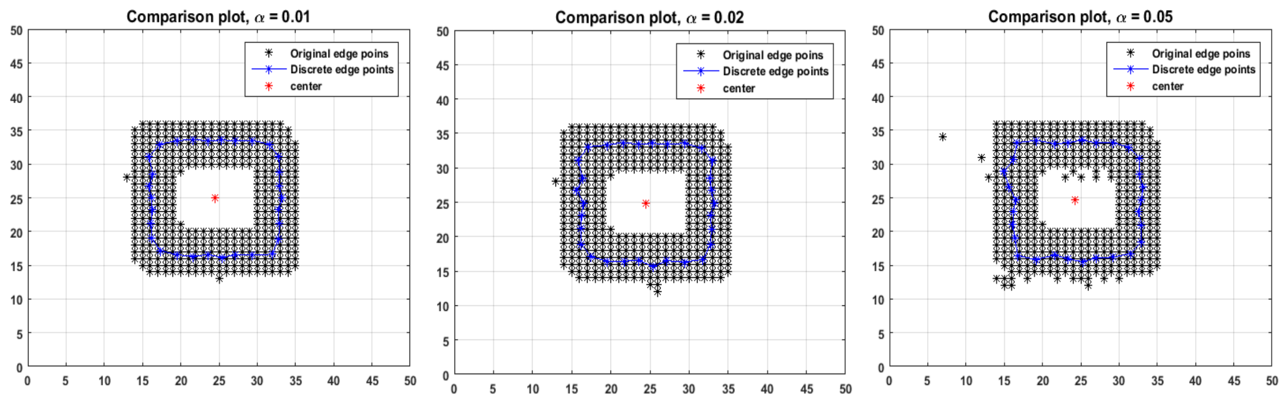


Figure 4-43: Comparison between original edge points and estimated edge points on pothole with gradual edges developed from Number Difference Edge Detection Method

4.4.4 Filter by looking at local points, connect edges by least square curve fitting

The original pothole with steep edges depicted in Figure 3-11 and Figure 3-12. The original pothole with gradual edges depicted in Figure 3-14 and Figure 3-15. Both potholes are approximately square of 20×20 located in a 50×50 grid.

<u>Results from Expected Distance Edge Detection Method on pothole with steep edges</u>

Three cases of different threshold are tested in Section 3.4.2 and their results are used in this Edge Stitching method. The data store in Figure 4-44 is used for the analysis of generating edges with a center, fixed angle, and respective radius in both polar coordinates and Cartesian coordinates. The image has a suspected edge indicated by the high intensity blue shapes.

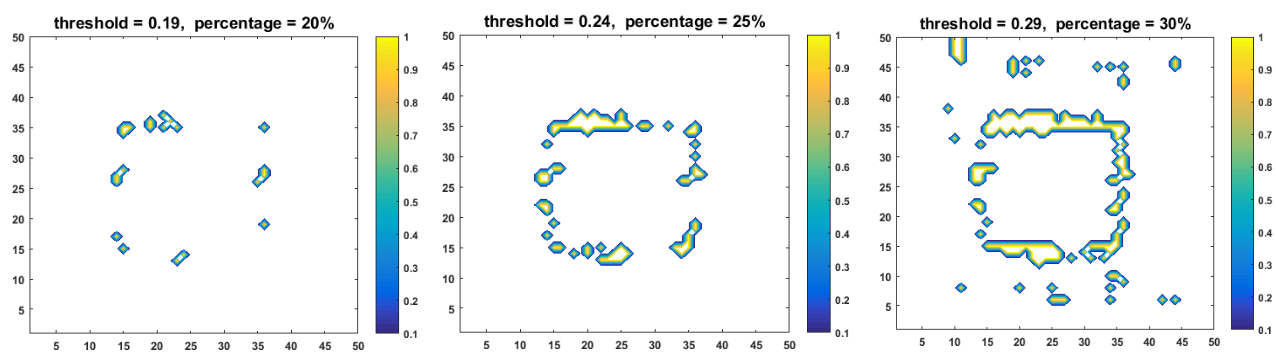


Figure 4-44: Results of Expected Distance Method on pothole with steep edges in the case of different threshold

As stated in Section 4.3.1.2 and Section 4.3.2.2, the methods of filtering edge points by looking at local points and connecting edges by least square curve fitting will be used in this example. Figure 4-45 displays the filtered edges and Figure 4-46 displays the connected edges.

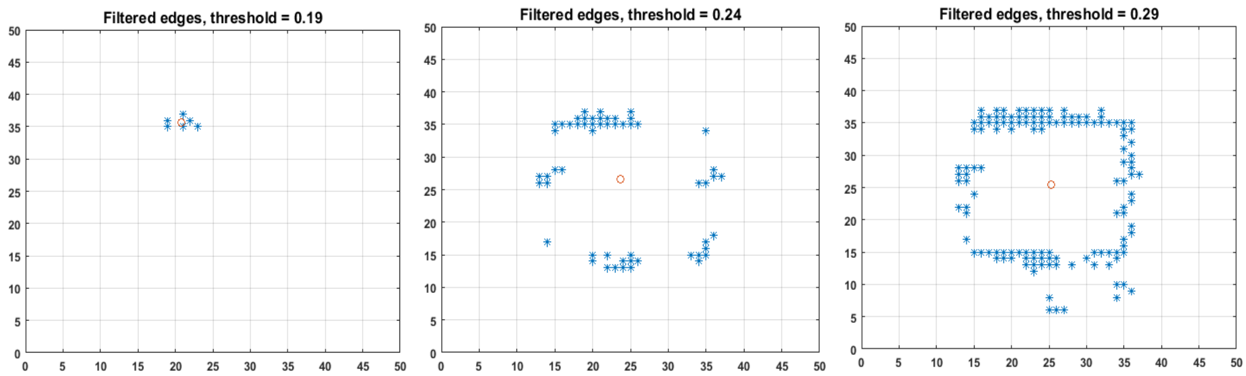


Figure 4-45: Filtered edge points using method of looking at local edges points of pothole with steep edges in cases of threshold 0.19, 0.24, and 0.29

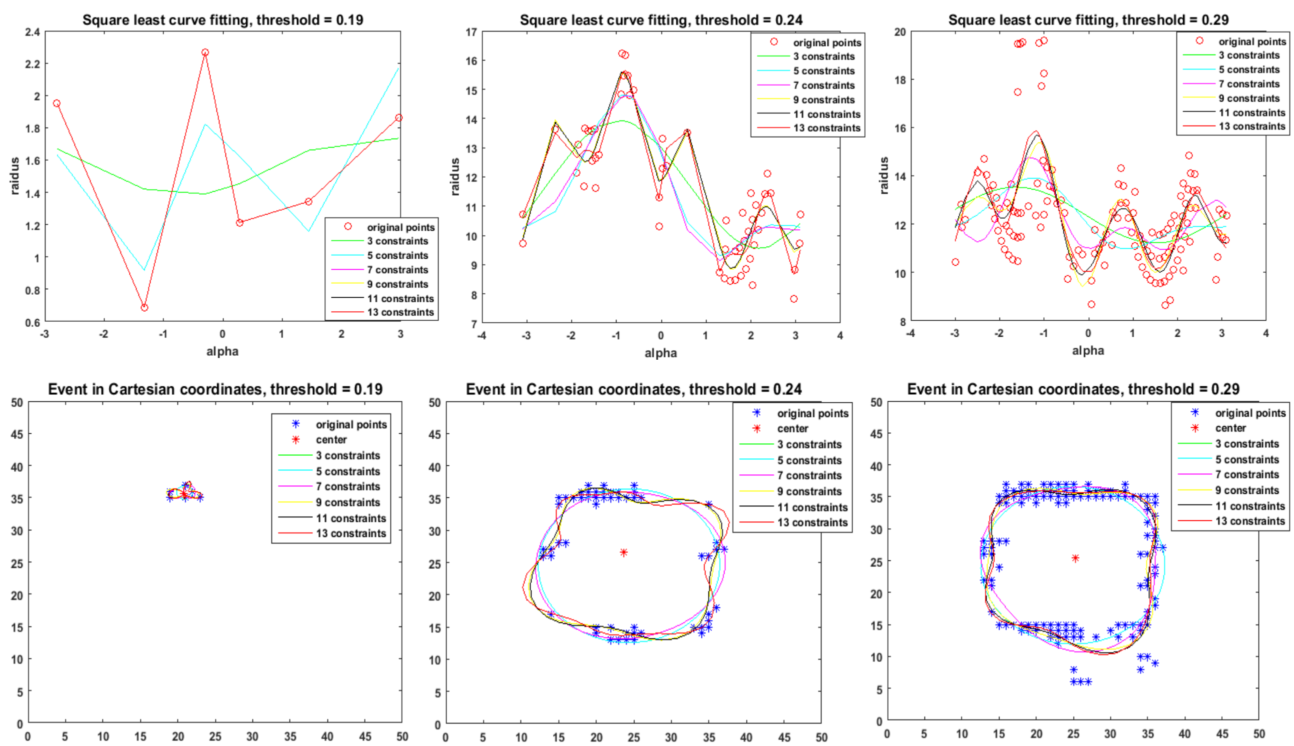


Figure 4-46: Comparison between original edge points and estimated edge points on pothole with steep edges developed from Expected Distance Edge Detection Method

Results from Expected Distance Edge Detection Method on pothole with gradual edges

Three cases of different threshold are tested in Section 3.4.2 and their results are used in this Edge Stitching method. The data store in Figure 4-47 is used for the analysis of generating edges with a center, fixed angle, and respective radius in both polar coordinates and Cartesian coordinates. The image has a suspected edge indicated by the high intensity blue shapes.

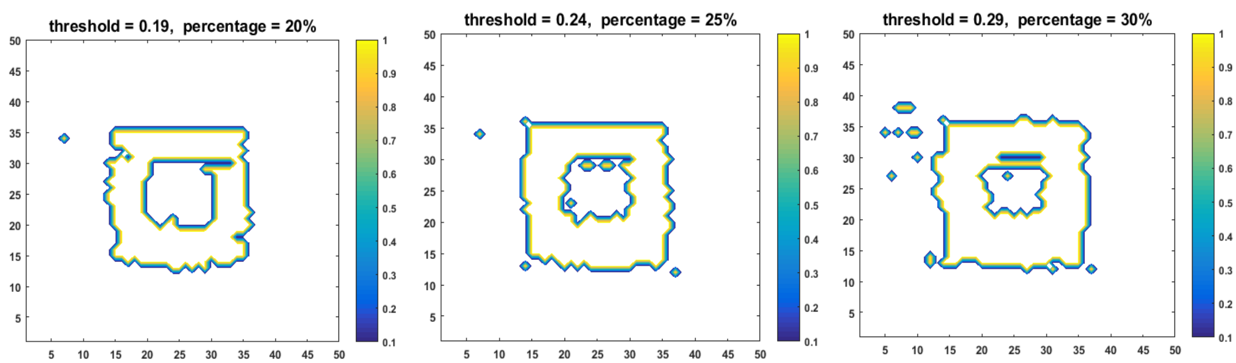


Figure 4-47: Results of Expected Distance Method on pothole with gradual edges in cases of different threshold

As stated in Section 4.3.1.2 and Section 4.3.2.2, the methods of filtering edge points by looking at local points and connecting edges by least square curve fitting will be used in this example. Figure 4-48 displays the filtered edges and Figure 4-49 displays the connected edges.

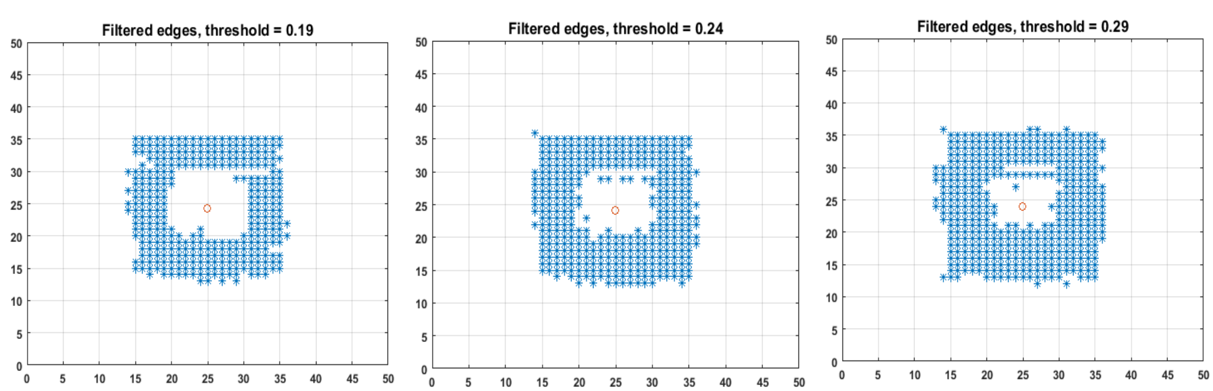


Figure 4-48: Filtered edge points by looking at local edges points of pothole with gradual edges in cases of threshold 0.19, 0.24, and 0.29

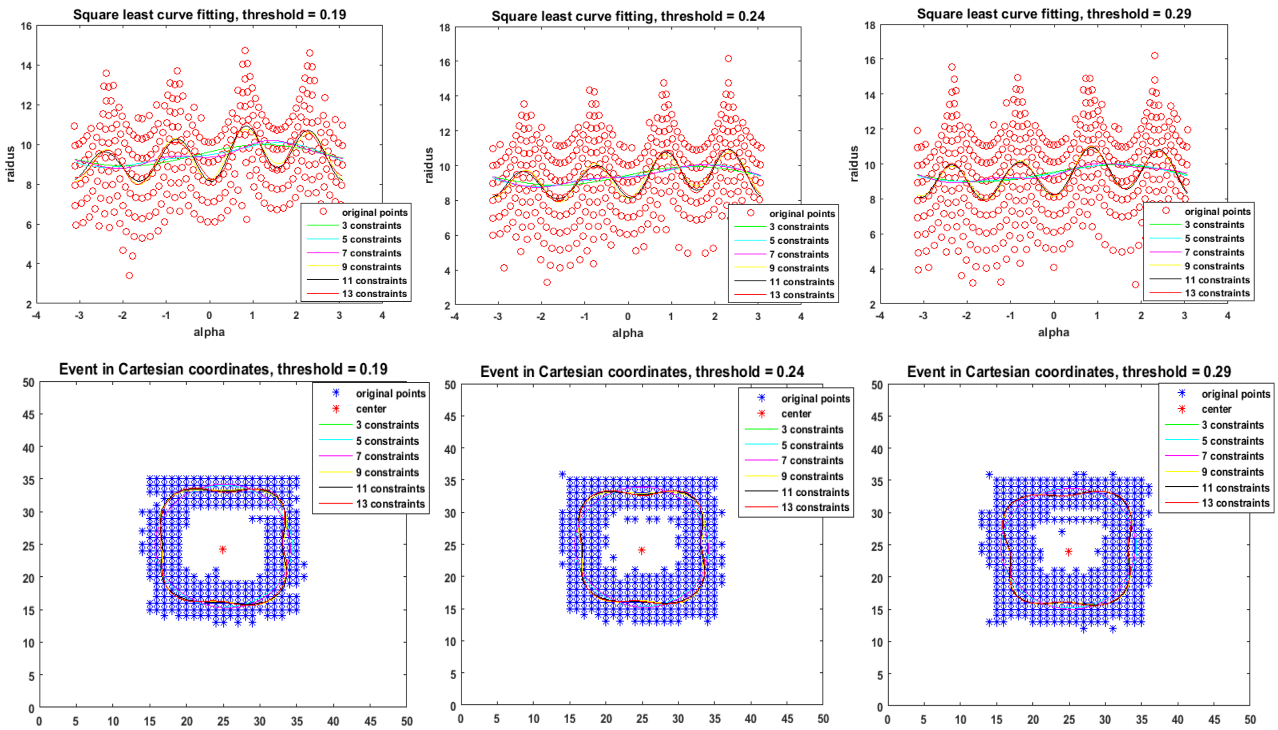


Figure 4-49: Comparison between original edge points and estimated edge points on pothole with gradual edges developed from Expected Distance Edge Detection Method

Results from Number Difference Edge Detection Method on pothole with steep edges

Three cases of different threshold are tested in Section 3.4.2 and their results are used in this Edge Stitching method. The data store in Figure 4-50 is used for the analysis of generating edges with a center, fixed angle, and respective radius in both polar coordinates and Cartesian coordinates. The image has a suspected edge indicated by the high intensity blue shapes.

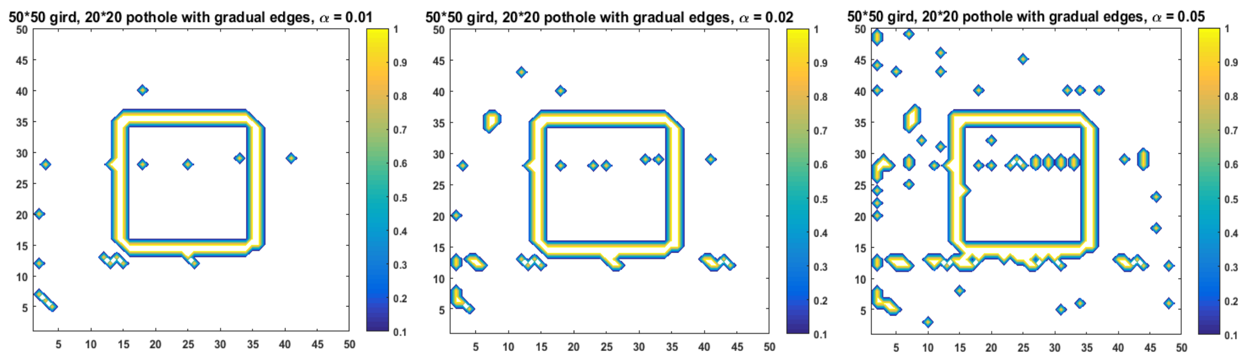


Figure 4-50: Results of Number Difference Method on pothole with steep edges in cases of different significance level

As stated in Section 4.3.1.2 and Section 4.3.2.2, methods of filtering edge points by looking at local points and connecting edges by least square curve fitting will be used in this example. Figure 4-51 displays the filtered edges and Figure 4-52 displays the connected edges.

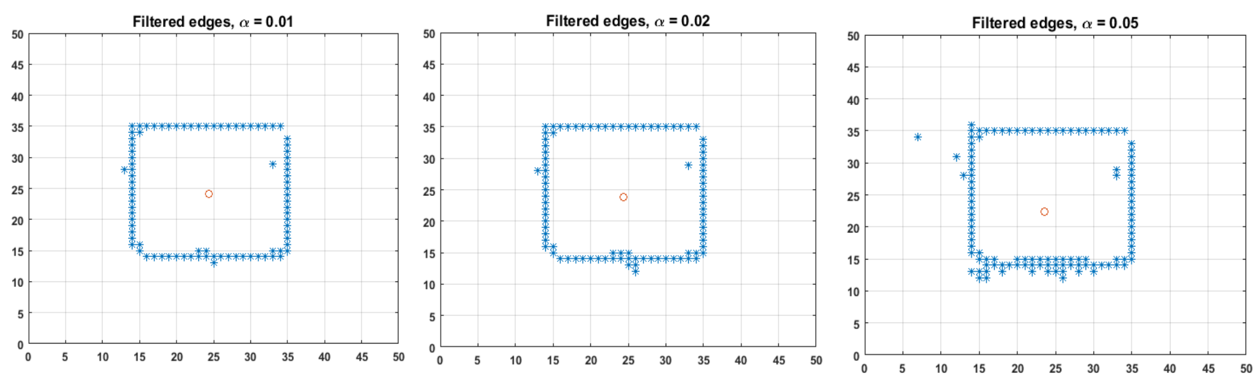


Figure 4-51: Filtered edge points by looking at local edges points of pothole with steep edges in cases of significance level 0.01, 0.02, and 0.05

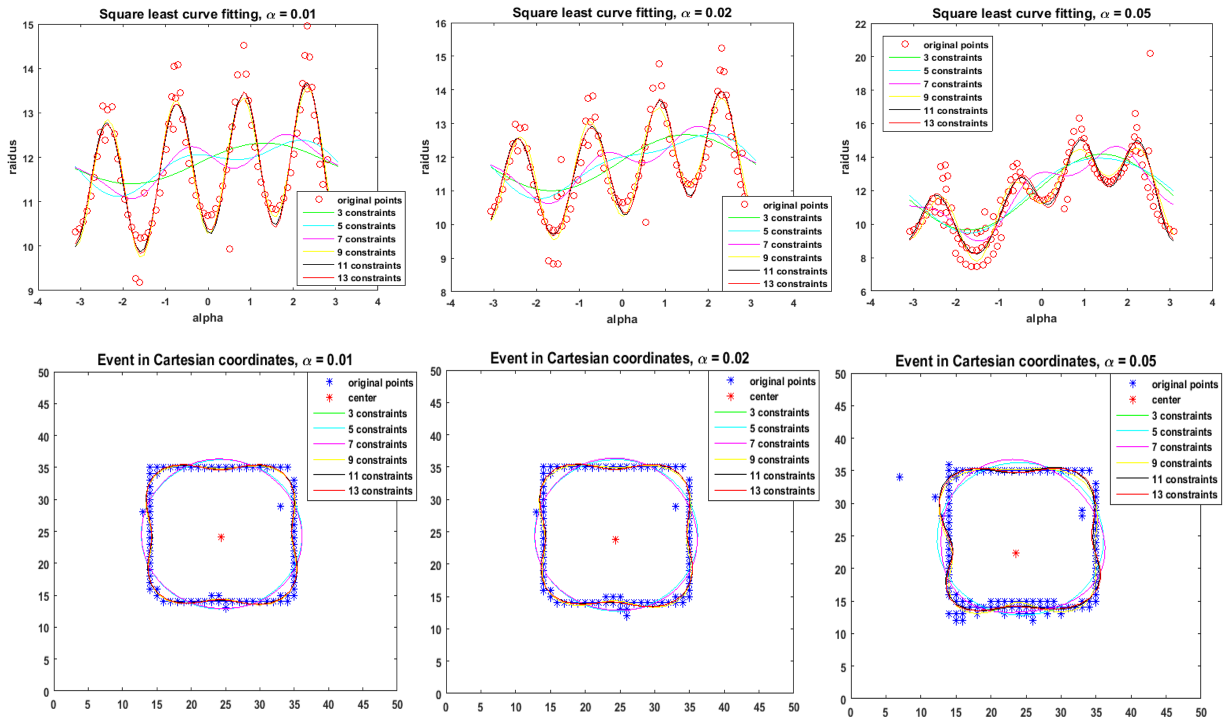


Figure 4-52: Comparison between original edge points and estimated edge points on pothole with steep edges developed from Number Difference Edge Detection Method

Results from Number Difference Edge Detection Method on pothole with gradual edges

Three cases of different threshold are tested in Section 3.4.2 and their results are used in this Edge Stitching method. The data store in Figure 4-53 is used for the analysis of generating edges with a center, fixed angle, and respective radius in both polar coordinates and Cartesian coordinates. The image has a suspected edge indicated by the high intensity blue shapes.

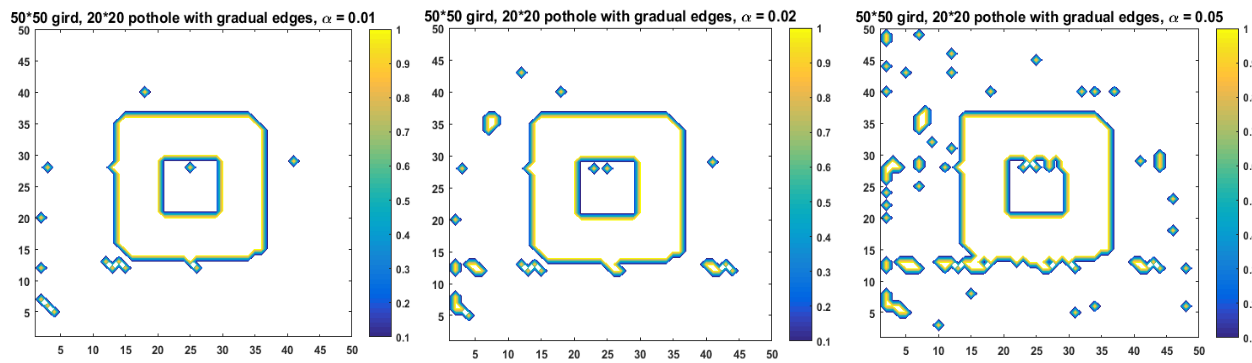


Figure 4-53: Results of Number Difference Method on pothole with steep edges in cases of different significance level

As stated in Section 4.3.1.2 and Section 4.3.2.2, methods of filtering edge points by looking at local points and connecting edges by least square curve fitting are used in this example. Figure 4-54 displays the filtered edges and Figure 4-55 displays the connected edges.

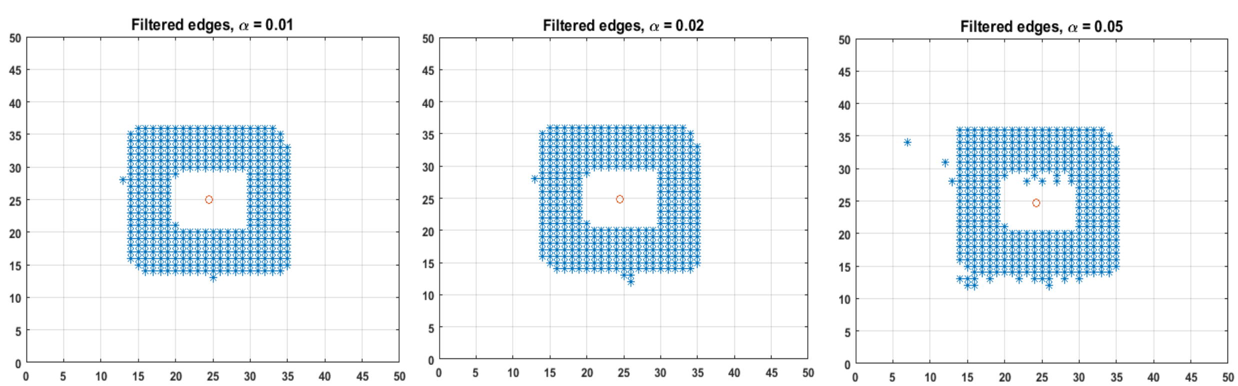


Figure 4-54: Filtered edge points by looking at local edges points of pothole with gradual edges in cases of significance level 0.01, 0.02, and 0.05

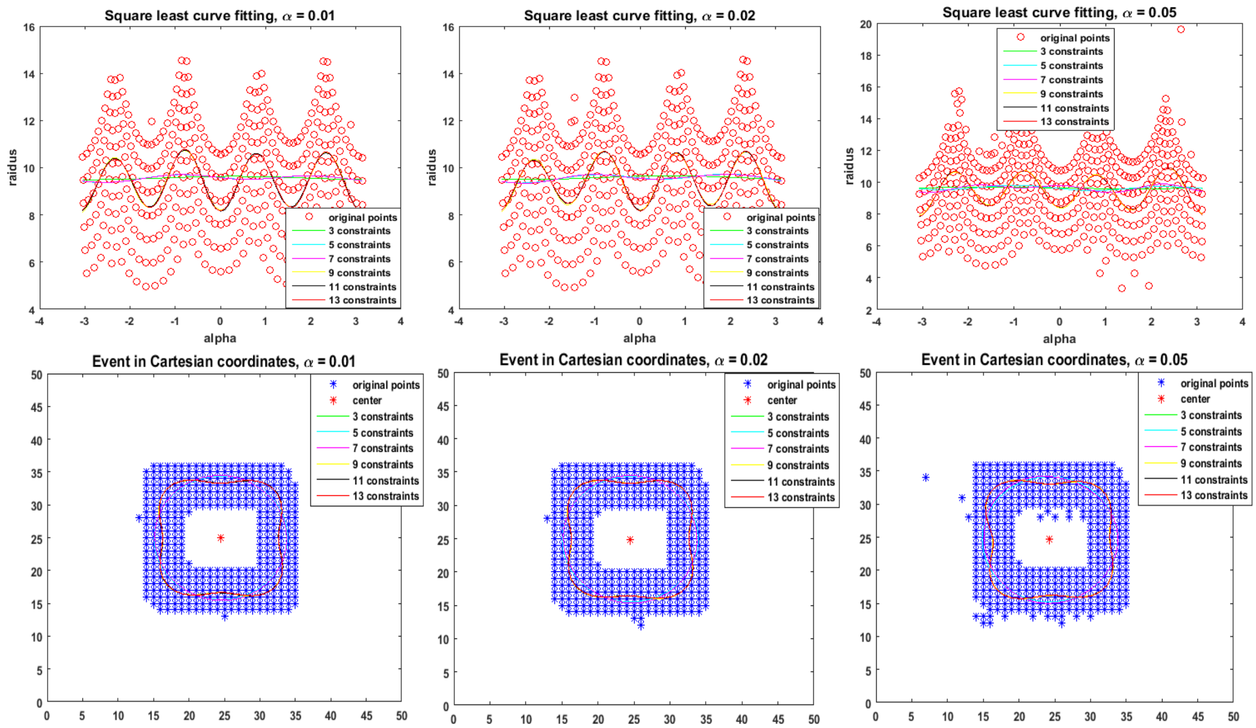


Figure 4-55: Comparison between original edge points and estimated edge points on pothole with gradual edges developed from Number Difference Edge Detection Method

4.5 Discussion

This section provides comparisons among different methods of filtering edge points and connecting edge points to event. Section 4.5.1 compares two methods of filtering edge points. Section 4.5.2 compares two methods of connecting edge points to event. Finally, Section 4.5.3 will compares four methods of Edge Stitching. Two instances of pothole with steep edges and gradual edges will be discussed in each comparison.

4.5.1 Comparison between two filtering edge points methods

Filtering edges points by standard deviation of radius distribution is based upon the properties of normal distribution. This method, therefore, only applies on the data whose distribution is normal. As shown in Figure 4-7, this filtering method yields better shape of event when chosen threshold is 0.29 other than 0.19 or 0.24. The main reason is that smaller threshold produces fewer possible edge points and thus results in the uneven distribution of edge clusters. Filtering these distributed edge clusters by standard deviation of radius distribution tends to rely on a biased averaged center and may cause the removal of real edge points. Another reason of yielding incomplete edges after filtering is that the pothole tested in this case has with steep edges. This will radically give fewer points than simulating a pothole with gradual edges. However, all results (as shown in Figure 4-16 and Figure 4-18) of test with Number Difference Edge Detection Method show a clear shape of filtered edges in every case of significance level. The pothole with gradual edges obtained from Expected Distance Edge Detection Method also gives a clear shape of filtered edges in every case of significance level (as shown in Figure 4-14).

Furthermore, for long-narrow events, this method may filter out edge points on the short perimeter since those points are unlikely located within one standard deviation. Although this filtering method is easy to execute, it may result in decrease of accuracy for event detection on long-narrow events. Additionally, this method is based upon the radius distribution of all possible edge points, it is hence very likely to cause incorrect match in cases of multiple events. Therefore, the clusters of edge points have to be separated before this filtering method is adopted.

Filtering edge points by looking at each possible edge point itself is another filtering method that is based upon a fixed number “N”. This number “N” will determine whether this potential edge point will stay in the group of edge points or be thrown away as an outlier. If there are more than or equal number of N points around this point, it will not be considered an outlier and will stay in the group of edge points; otherwise, it will be considered an outlier.

Figure 4-33 displays the filtered edge points obtained from the filtering method of looking at local points. This filtering method yields better shape of event when the threshold is 0.29 other than 0.19 or 0.24. The main reason is that smaller threshold produces fewer possible edge points and thus results in uneven distribution of edges clusters. Such clusters are more likely to be removed when very few points are within them. This will lead to the incompletion of detected events. Another reason of yielding incomplete edges after filtering by this method is that the pothole tested in this case has steep edges. This will undoubtedly produce fewer edge points than that when a pothole with gradual edges is applied. As shown in Figure 4-36, in all cases of different threshold, this filtering method yields relatively clear edges of pothole with gradual edges. Similarly, combined with filtering method by standard deviation of radius distribution, this method yields accurate results of filtered edges as long as the results of Number Difference Edge Detection Method are used regardless of a pothole with steep edges or gradual edges, as shown in Figure 4-39 and Figure 4-42.

This filtering method has an apparent advantage over the method of filtering edges points by standard deviation of radius distribution. That is, in such way, all outliers will be selected out no matter where the group of edge points is located. Since this method determines outliers by looking at local edge points that are not related to clusters of edge points belonging to other events, this method can be applied on multiple events and will not cause inaccurate match.

4.5.2 Comparison of two connecting edge points to events methods

Connecting edge points by discretizing angles is to describe the position of each edge point at discretized angles from 0 to 2π. At each discrete angle, an edge representative is calculated by averaging the radius of all edge points located around this angle within one tolerance. The cluster of edge points, therefore, can be sorted to edge representatives located at discretized angles. In Figure 4-14, Figure 4-18, Figure 4-37, Figure 4-40 and Figure 4-43, this method gives a clear shape of event when the filtered edge points consist of a nearly complete shape of event. However, if the filtered edge points are not totally complete (as shown in Figure 4-11 and Figure 4-34) or if outliers occupy many edge points (as shown in Figure 4-16 when significance level is 0.05), the connected edge points tend not to have a right shape of event. The reason is that some discretized angles possibly do not have an edge representative when no edge points are located within the tolerance of this discretized angle. This always happens in the low density area. Even though edge representatives at each discretized angle can be obtained finally, there is a high probability that the edge points at certain discretized angles are far away from their supposed positions.

Another method of connecting edge points to event is based on the method of least square curve fitting, which yields a smoother contour than the method of discretized angles. Such examples are shown in Figure 4-22 and Figure 4-25. It is noteworthy that for threshold 0.19, the final event cannot be detected out completely. However, as shown in other two pictures in Figure 4-22 and all pictures in Figure 4-25, the edge points can be connected to events correctly and clearly with different thresholds. This method has an important advantage in that the best fit in the least square curve minimizes the sum of squared residuals - the difference between observed values and their fitted values provided by a model (the model is developed with Fourier series equations which as reviewed in Section 4.2.3).

However, this method only applies well on cases when the filtered edges points are as enough as input of building equations. In Figure 4-22, when the threshold is 0.19, the result comes out is not in a right shape, but a chaos. However, when the filtered edge points number more than points for threshold 0.19, this method can give a relatively right shape of event and the estimated event contour becomes more accurate with the increase of coefficients in Equation 4-16. For pothole with gradual edges, this method nonetheless can connect edge points to events accurately (as shown in Figure 4-24 and Figure 4-31). This does not mean that edge points in these cases are not affected by outliers. The truth is that the filtered edge points in this case are much more than those on pothole with steep edges, so the outliers will not affect the detected events that much.

4.5.3 Comparison of four Edge Stitching methods

The results of four edge stitching methods in cases of threshold 0.19, 0.24, and 0.29 are compared and displayed in **Error! Reference source not found.**. The first method filters edge points by standard deviation of radius distribution and connects edges by discretizing angles. The second method filters edge points by standard deviation of radius distribution and connects edges by least square curve fitting. The third method filters edge points by looking local points and connecting edges by discretizing angles. The last method is using filtering edge points by looking at local points and connects edges by least square curve fitting. The edges are taken from Expected Distance Edge Detection Method.

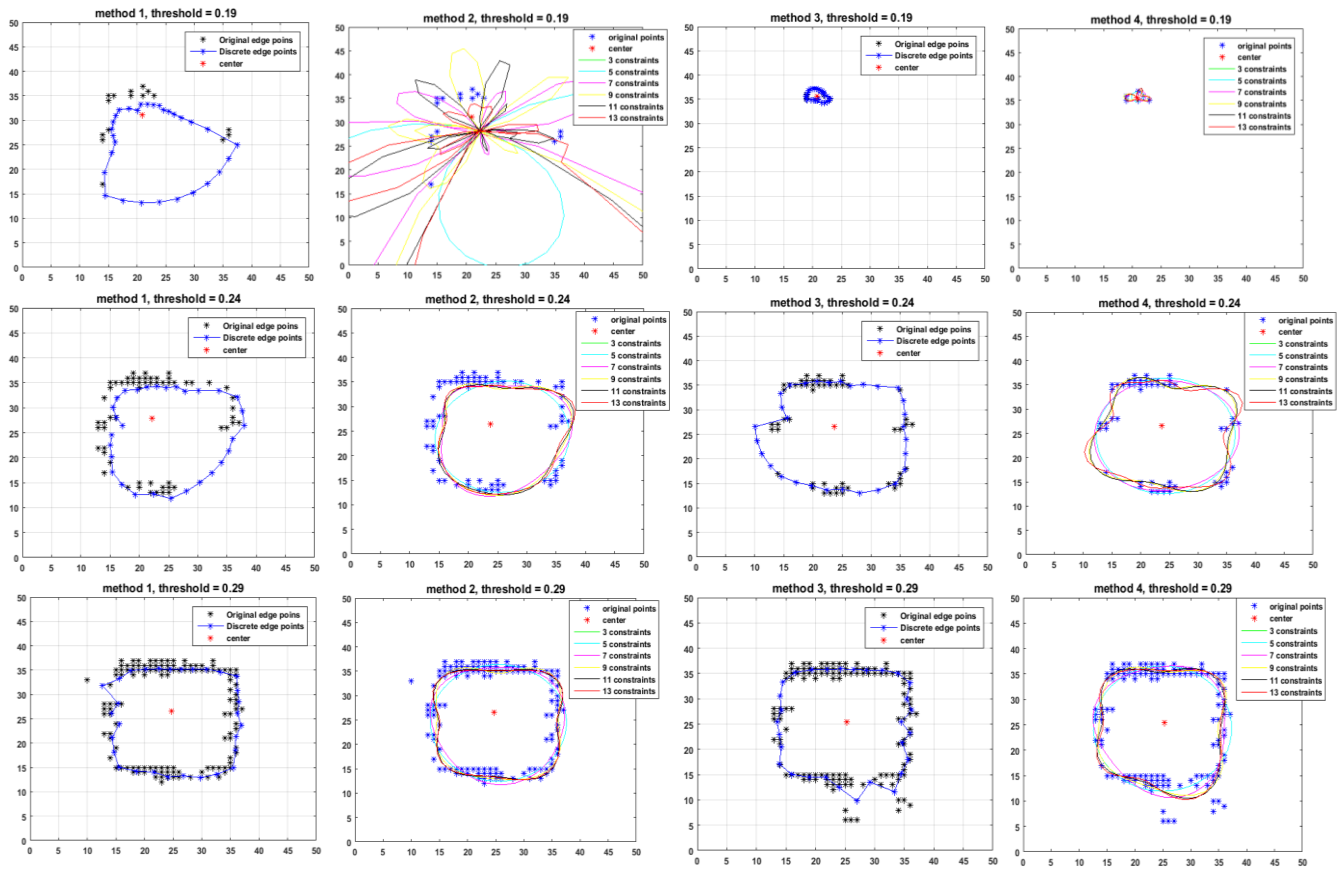


Figure 4-56: Comparison of four Edge Stitching Method on pothole with steep edges obtained from Expected Distance Edge Detection Method

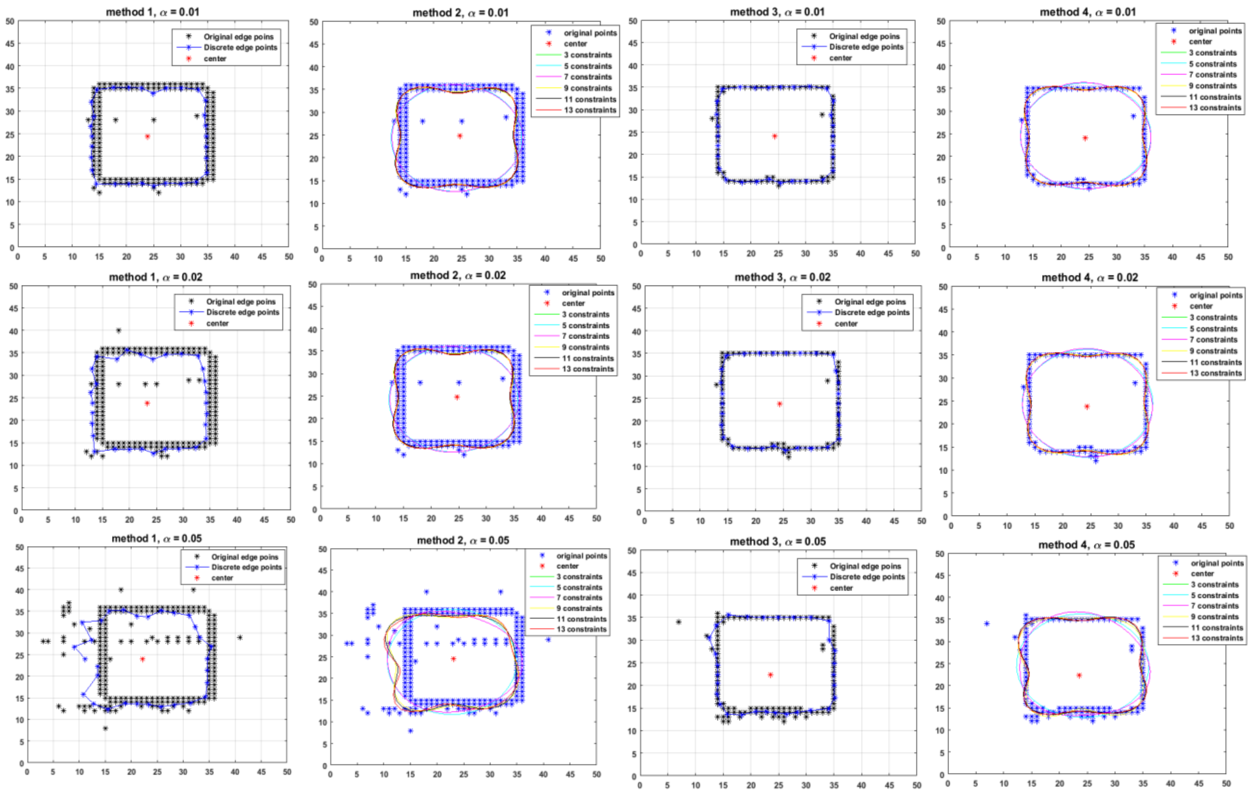


Figure 4-57: Comparison of four Edge Stitching Method on pothole with steep edges obtained from Number Difference Edge Detection Method

As shown in **Error! Reference source not found.**, method 1 apparently capable of connecting edges in all three thresholds. Seemingly, the method of filtering edge points by standard deviation of radius distribution behaves better than the other filtering method of looking at local points. However, as discussed in Section 4.5.1, this method of filtering edge points by standard deviation of radius distribution is not suitable for long-narrow events, in that it is based on radius distribution of all possible edges points in such a way that edge points on short perimeter are very likely to be filtered out. Furthermore, this method does not fit multiple events well either. When comparing all four test results for threshold 0.29, it is not difficult to find that the curve fitting method yields smoother contour than discretizing angles. In other words, curve fitting method minimizes the residual. This can be seen in method 3 and method 4 for threshold 0.24 as well. With respect to outliers filtering, the comparison of method 2 and method 3 for threshold 0.24 manifests that filtering method of looking at local points leaves fewer points than filtering method of using radius distribution. The reason why this happens is that when the pothole tested has steep edges, if the threshold is too small, there will be no enough edge points yielded from Edge Detection Method. In this case, it is very possible that edges points are considered “outliers” and removed with filtering method by looking at local points. However, this problem does not exist in the case of pothole with gradual edges.

In Figure 4-57, all pictures give clear edges no matter which one of the four Edge Stitching methods is applied on the results obtained from Number Difference Edge Detection Method. Selection of threshold may affect accuracy of detected edges. For example, with method 1 and method 2, remaining outliers will cause incorrect shape of events when significance level is set to 0.05. For method 3 and method 4, however, all three cases of significance level give correct shape of events.

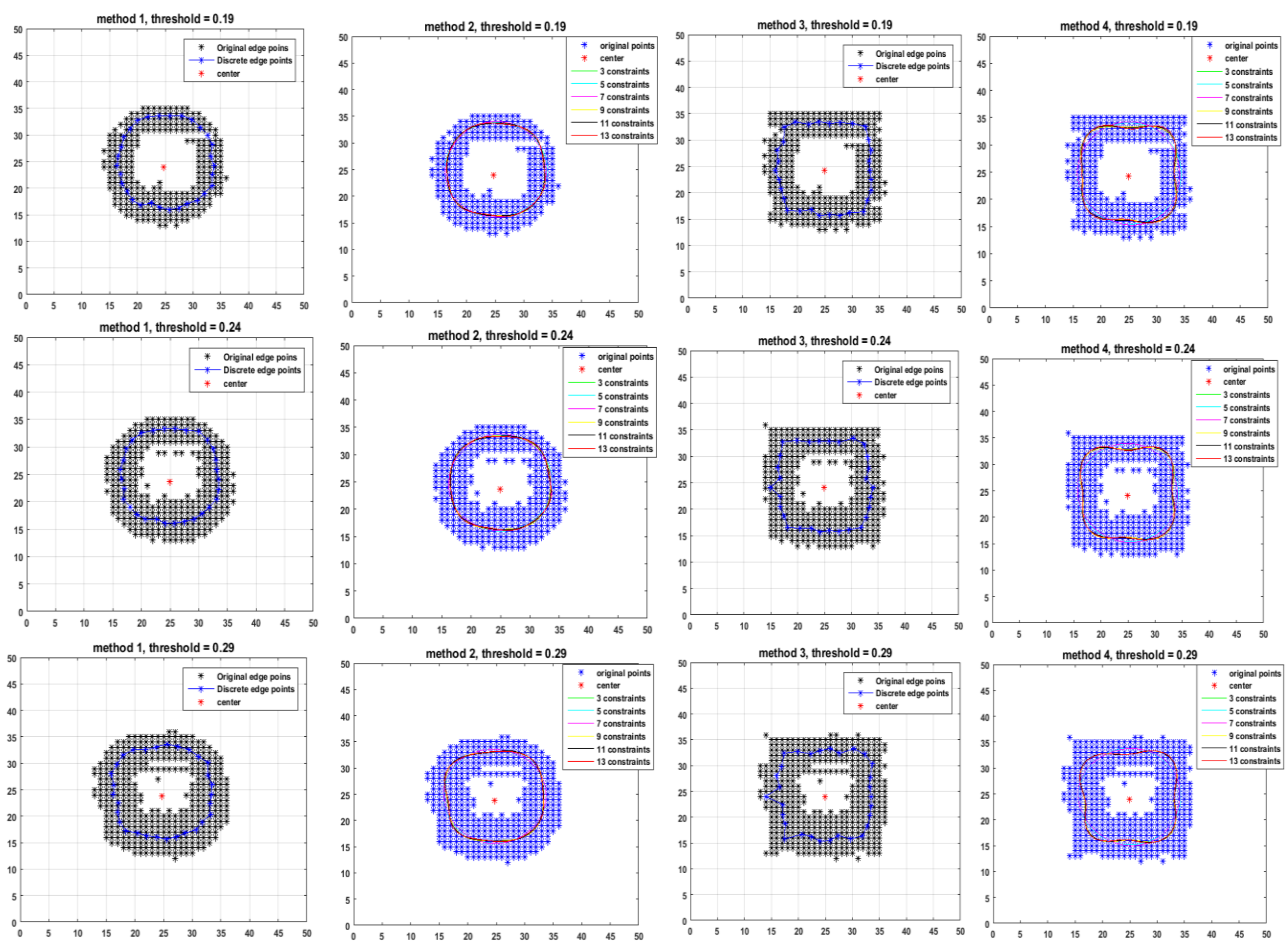


Figure 4-58: Comparison of four Edge Stitching Method on pothole with gradual edges obtained from Expected Distance Edge Detection Method

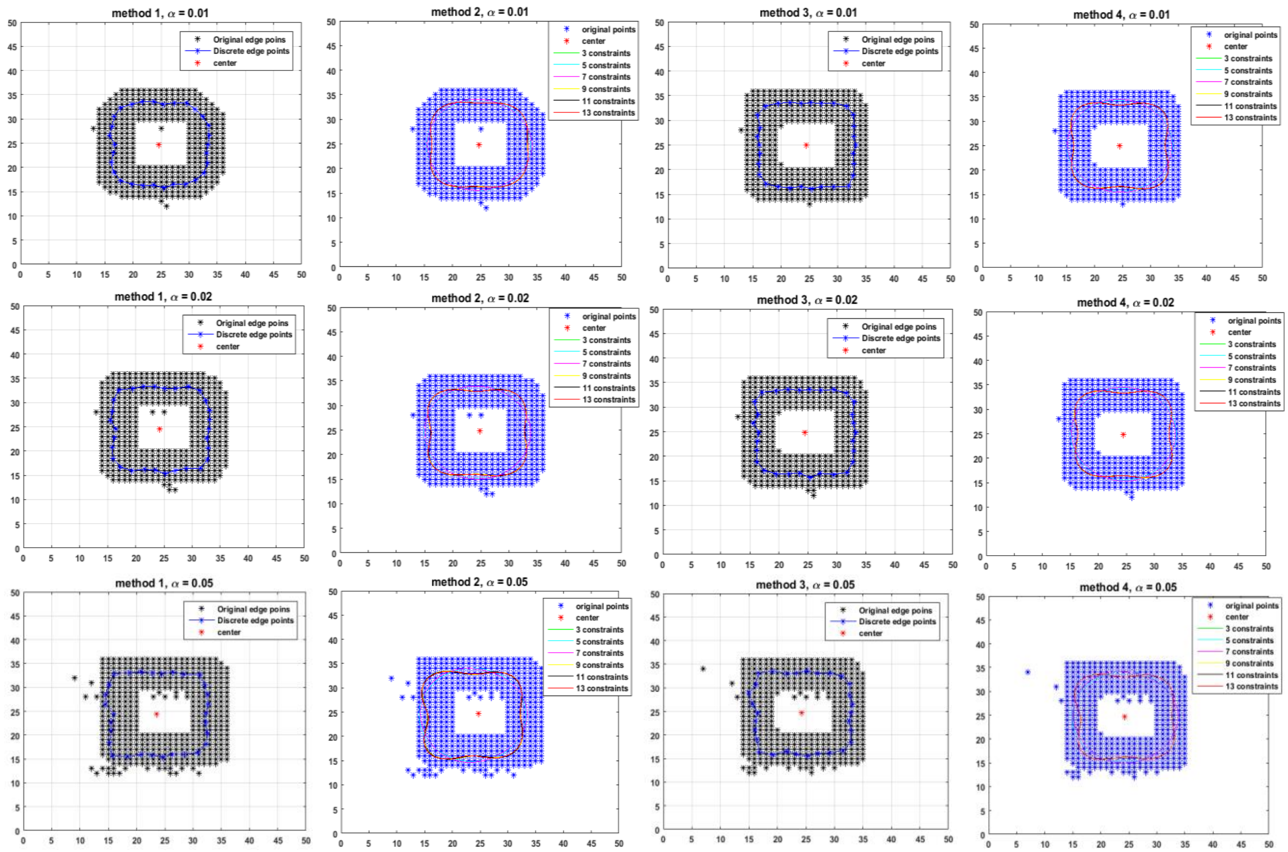


Figure 4-59: Comparison of four Edge Stitching Method on pothole with gradual edges obtained from Number Difference Edge Detection Method

Figure 4-58 displays the results of Edge Stitching Method on pothole with gradual edges based on the data from Expected Distance Edge Detection Method. Obviously, method 1 and method 2, which filter edge points by standard deviation of radius distribution maintain fewer original points than that from method 3 and method 4, which filter edge points by looking at local points. When looking at the standard deviation of radius distribution, edges points tend to be filtered if they are located out of the tolerance. Consequently, even though those edge points on the corner are not supposed to be removed from entire set of edge points, it still occurs when this filtering method is used. This problem, however, does not happen when edge points are filtered by looking at local points in that this filtering method only looks at edge point itself, instead of relating to radius distribution.

Noticeably, unlike method 1 and method 2, method 3 and method 4 yield nearly rectangular event, which is very close to the shape of original pothole. The difference between these two groups of methods lies on the filtering method involved. Filtering edge points by standard deviation of radius distribution tends to remove corner points of the rectangular pothole since they are located out of one standard deviation. Consequently, the filtering method by standard deviation of radius distribution usually yields an event with round corners instead of sharp corners when the pothole is in rectangular shape. This undesired problem will not happen if the method of filtering edge points by looking at local points is employed, as indicated by the results from method 3 and method 4 in Figure 4-58.

Figure 4-59 displays the results of Edge Stitching Method on pothole with gradual edges based on the data processed from Number Difference Edge Detection Method. All four Edge Stitching methods yields clear events. Please note that for method 1 and method 2, with edge points in the corner filtered out when significance level is 0.01, clear events can still be obtained though.

To sum up, Edge Stitching is a critical step after all edge points are found. Using the result of Expected Distance matrix or Number Difference matrix from Comparative Nodal Uncertainty Edge Detection Method, Edge Stitching procedure groups edges to events. Several algorithms for searching the right path of edges can be executed successfully to reach the goal of connecting edge nodes to events sequentially and properly. Two filtering methods with different level of applicability and limitations are successfully developed and employed to filter self-defined outliers for the result of Comparative Nodal Uncertainty Edge Detection Method. Two methods of connecting filtered edge points to events are successfully developed and employed to connect all possible edge points to events by filling edge gaps afterwards as well.

4.6 Contribution

This study mainly illustrates two filtering methods for edge points and two stitching methods for connecting edges to events. Four combinations of Edge Stitching method are tested and compared in this effort. Principal contributions of this study are:

- Different methods are developed and tested to filter edge points by standard deviation of radius distribution by looking at local edge points based upon the results of Expected Distance Edge Detection Method and Number Difference Edge Detection Method.
- Examples are explored for four combinations of two methods of filtering edge points and two methods of connecting edges to events with different thresholds based upon the results of Expected Distance Edge Detection Method or different significance levels based upon the results of Number Difference Edge Detection Method.
- Four Edge Stitching methods are compared and analyzed for their applicability, advantages, and disadvantages, respectively.
- At least one of these four Edge Stitching methods is proved to be effective in filtering outliers from all possible edge points and connecting filtered edges to events. The process of event detection can be finalized to achieve a positive result with such method.

5. Conclusion

This study is aimed at the development of event detection in terrain surfaces, a process of identifying localized events such as potholes in the terrain surface based on terrain flatness. Comparative Nodal Uncertainty Method and Edge Stitching Methods are developed and applied in the course of this study. Some combinations of these methods are capable of providing a relatively clear shape of edges for the pothole on the road surface.

Comparative Nodal Uncertainty Method is developed to detect edges based on the probability distribution of the nodal heights within some local neighborhood using Wilcoxon Rank-Sum Test. The output of this method is an Expected Distance matrix or a Number Difference matrix. Obtained through a threshold, the entries of Expected Distance matrix are expected to be small for the nodes on the edge. On the contrary, obtained through a significance level, the entries of Number Difference matrix are expected to be large for the nodes on the edge.

Edge Stitching is a step further in terms of finding out edge nodes in order to group nodes to events. Edge Stitching methods use the results of Expected Distance matrix or Number Difference matrix from Comparative Nodal Uncertainty Method. Several search algorithms are developed in order to group edge nodes to events sequentially. Totally four Edge Stitching methods are employed in this study to finalize the process of Event Detection.

Based on the results as illustrated in Section 3.4 and Section 4.4, conclusion can be made that successful event detection process is possible to detect events in different cases when the right combination of edge point filtering and edge point stitching methods is adopted.

References

1. Ma, R., Ferris, J.B., 2011, "Terrain Gridding Using a Stochastic Weighting Function," *Proceedings of the ASME Dynamic Systems and Control Conference*. October 31 - November 2, Arlington, Virginia.

2. Acbarjya, Pinaki, Das Ritaban, Ghoshal, Dibyendu. "A Study on Image Edge Detection Using the Gradients." *International Journal of Scientific and Research Publications*. Volume 2, Issue 12. December 2012. ISSN 2250-3153

3. Vincent, O.R., Folorunso, O., 2009, "A Descriptive Algorithm for Sobel Image Edge Detection," *Proceedings of Informing Science & IT Education Conference.*

4. Gupta S., Mazumdar S. G., "Sobel Edge Detection Algorithm" *International Journal of Computer Science and Management Research.* Volume 2 Issue 2. February 2013. ISSN 2278-733SX

5. Seif, A.,et.al. ;"A hardware architecture of Prewitt edge detection", *Sustainable Utilization and Development in Engineering and Technology (STUDENT), 2010 IEEE Conference,* Malaysia, pp. 99 – 101, 20-21 Nov. 2010.

6. L. Roberts. "Machine Perception of 3-D Solids" *Optical and Electro-optical Information Processing,* MIT Press 1965.

7. Shrivakshan, G.T., Chandrasekar, Dr.C., "A Comparison of various Edge Detection Techniques used in Image Processing," *IJCSI International Journal of Computer Science Issues*, Vol. 9, Issue 5, No 1, September 2012. ISSN(Online) 1694-0814

8. Vliet, L.J, Young, I.T., Beckers, G.L., 1989, "A Nonlinear Laplace Operator as Edge Detector in Noisy Images," *Computer Vision, Graphics, and Image Processing 45, 167-195*

9. Liqin, Shen, Dinggang Shen, Feihu, Qi, 1994 "Edge Detection on Real Time Using LOG Filter". *International Symposium on Speech, Image Processing and Neural Networks.* April 13 – 16, Hong Kong.

10. D. Marr and H. Hildreth, "Theory of edge detection" Proc. Roy. Soc. London, vol. B207, pp. 187-217, 1980.

11. Haralick, Robert M. "Digital Step Edges from Zero Crossing of Second Directional Derivatives" *IEEE Transactions on Pattern Analysis and Machine Intelligence.* Volume PAMI-6, Number 1, Pp 58-68. January 1984.

12. Devore, J.L., (1987). *Probability and Statistics for Engineering and the Science.* Belmont, California: Brooks/Cole Publishing Company

13. Frinkle, M., Weir, R., and Ferris, J. B., 2004, "A Graphical Representation of Road Profile Characteristics," *Proceedings of the SAE World Congress.* Paper number 2004-01-0769, SAE, Detroit, MI.

14. Umbaugh, Scott E. "*Digital image processing and analysis: human and computer vision applications with CVIPtools."* (2nd ed. ed.). Boca Raton, FL: CRC Press. 2010. ISBN 9-7814-3980-2052

15. Canny, J., "A Computational Approach to Edge Detection." *IEEE Trans. Pattern Analysis and Machine Intelligence*. 8(6):679–698, 1986.

16. J. M. S. Prewitt, "Object Enhancement and Extraction," *Picture Processing and Psychopictorics*, B. Lipkin and A. Rosenfeld, Eds. New York: Academic, 1970, pp. 75-149.

17. Sobel, I., "*An Isotropic 3×3 Gradient Operator, Machine Vision for Three – Dimensional Scenes*", Freeman, H., Academic Press, NY, 376-379, 1990.

18. L. Roberts. *"*Machine Perception of 3-D Solids*" Optical and Electro-optical Information Processing*, MIT Press 1965.

19. LS. Davis, "A survey of edge detection techniques", *Computer Graphics and Image Processing*, Volume 4, Number 3, pp 248-260, 197.

20. Sezgin, Mehmet, Sankur, Bulent, "Survey over Image Thresholding Techniques and Quantitative Performance Evaluation." *Journal of Electronic Imaging.* 13(1), pp 146-165. January 2004.

21. Inad Aljarrah, "Image Mosaicing Using Binary Edge Detection." *Proceedings of the International conference on Computing Technology and Information Management*. Dubai, UAE, 2014.

22. Saralees Nadarajah, "A generalized normal distribution." *Journal of Applied Statistics*, 32:7, 685-694, DOI: 10.1080/02664760500079464, 2005

23. Deli Li, "On the central limit theorem along subsequences of sums of i.i.d. random variables." *Published online,* August 2014

24. Rice, John, "Mathematical Statistics and Data Analysis (Second ed)." Duxbury Press, ISBN 0-534-20934-3

25. Shibdas Bandyopadhyay, "Normal Probability plot", *AMS*, subject classification: 62P, 1991

26. W.C. Hamilton, S.C. Abrahams, "Normal probability plot analysis of small samples." *Acta Cryst.* A28, 215-218, DOI: 10.1107/S0567739472000476

27. Carlos M. Jarque, Anil K. Bera, “A Test for Normality of Observations and Regression Residuals.” *International Statistical Review,* Vol. 55, No. 2, pp. 163-172, August, 1987.

28. Smith, S.P., “A test to determine the multivariate normality of a data set.” Pattern Analysis and Machine Intelligence, IEEE Transactions on, Volume: 10, Issue: 5, pp. 757-761, 1988.

29. Friedrich Pukelsheim, “The Three Sigma Rule”, The American Statistician, Vol. 48, No. 2, pp. 88-91, May, 1994.

30. Phil Dyke, “An Introduction to Laplace Transforms and Fourier Series (Second Edition)”, pp. 83-87, ISSN: 1615-2085, 1999.

31. Gordon Haigh, “Piecewise Continuous Functions”, Mathematics in School, Vol. 8, No. 5, pp. 9-11, Nov., 1979.

32. Epperson, James F., “Introduction to Numerical Methods and Analysis (2nd Edition)”, pp. 234, ISBN: 1-118-39513-1, 978-1-118-39513-4, Oct., 2013.

33. Sangtae Ahn, “Standard Errors of Mean, Variance, and Standard Deviation Estimators.” July, 2003.

34. E.L.Lehmann, G.Casella, “Theory of point estimation”, Springer-Verlag, New York, 1998.

35. Andrea C. Skelly, “Probability, proof, and clinical significance”, Volume: 2, Issue: 4, 2011.

Appendix

A.1 Linear Space

This appendix, provides some fundamental concepts of linear space. The proof processes are intentionally omitted but can be found by following the references. Basic knowledge to mathematics and relevant mathematical notions and notations are essential to understand the processing methods of this study. This appendix is intended to enlighten readers with the underpinning mathematical methodology, other than preventing them from exploring the findings of this study if put in the main body of this paper. Several concepts are described here for readers to review and understand the Fourier series deeper.

A.1.1 Sequence

The elements of a set are usually written as a sequence of lower case letter as shown in Equation A-1.

$$\mathrm{A} = \{\mathrm{a}_1, \mathrm{a}_2, \mathrm{a}_3, \dots, \mathrm{a}_\mathrm{n}\} \qquad \text{Equation A-1}$$

With regard to Fourier series, sets have infinitely number of elements. So a row of dots should appear after a_n. To demonstrate an element that is included in a set, one usual notation is shown in Equation A-2.

$$\mathrm{S} = \{x|\ f(x)\} \qquad \text{Equation A-2}$$

The vertical line is read as "such that" so that $f(x)$ describes some property that x holds in order that $\mathrm{s} \in \mathrm{S}$, which means that s is an element of the set S. A sequence of function shown in

Equation A-3 is often used in Fourier series. The concepts of sequence is also needed to define the particular normed spaces within which Fourier series operate.

$$\left\{\frac{1}{\sqrt{2}}, \sin(x), \cos(x), \sin(2x), \cos(2x), \dots\right\} \qquad \text{Equation A-3}$$

A.1.2 Vector space

A vector space V is an algebraic structure that consists of elements (also called vectors) and two operations as addition "+" and multiplication "×". A vector space over F is the set V whose elements obey a list of properties by vectors a, b, c and scalar ξ, ϕ ∈ F where F is basically a field of real number or complex numbers as shown in Equation A-4 and Equation A-5, where x is the real part of z, and y is the imaginary part of z.

$$\mathbb{R} = \{x \mid x\ is\ a\ real\ number\} = \{-\infty, \infty\} \qquad \text{Equation A-4}$$

$$\mathbb{C} = \left\{ z = x + iy \mid x, y \in \mathbb{R}, \mathrm{i} = \sqrt{-1}\right\} = \{-\infty, \infty\} \qquad \text{Equation A-5}$$

The list of properties satisfying a vector space is:

1. $a + \mathrm{b}$ is also a vector (closure under addition).
2. $(a + \mathrm{b}) + \mathrm{c} = a + (\mathrm{b} + \mathrm{c})$ (associativity under addition).
3. There exists a zero vector denoted by 0 such that $0 + a = a + 0 = a, \forall a \in \mathrm{V}$ (additive identity).
4. For every vector a ∈ V, there is a vector $-a$ such that $a + (-a) = 0$.
5. $a + \mathrm{b} = \mathrm{b} + a$ for every a, b ∈ V (additive commutativity).
6. $\xi a \in \mathrm{V}$ for every $\xi \in \mathrm{F}, a \in \mathrm{V}$ (scalar multiplicity).
7. $\xi(a + \mathrm{b}) = \xi a + \xi \mathrm{b}$ for every $\xi \in \mathrm{F}, a, \mathrm{b} \in \mathrm{V}$ (first distributive law).
8. $(\xi + \phi)a = \xi a + \phi a$ for every $\xi, \phi \in \mathrm{F}, a \in \mathrm{V}$ (second distributive law).
9. For the unit scalar 1 of the field F, and every $a \in \mathrm{V}, 1 \cdot a = a$ (multiplicative identity).

In the study of Fourier series, vectors are basically functions. The name linear space emphasizes the linearity property which is confirmed by the following definition and properties. These definition and properties are useful for the understanding of Fourier series.

Definition A.1 (linear independence) If V is a vector (linear) space, the vectors $a_1, a_2, \dots, a_n \in V$ are said to be linearly independent if Equation A-6 exists.

$$\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_n a_n = 0 \qquad \text{Equation A-6}$$

This indicates that all of the scalars are zero, as shown in Equation A-7.

$$\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0, (\alpha_1, \alpha_2, \dots, \alpha_n \in F). \qquad \text{Equation A-7}$$

Thus, $\alpha_1, \alpha_2, \dots, \alpha_n$ are said to be linearly independent, otherwise, they are said to be linearly dependent.

Definition A.2 A finite set of vectors $a_1, a_2, \dots, a_n$ is said to be a basis for the linear space V if the set of vectors $a_1, a_2, \dots, a_n$ is linearly independent and $\mathrm{V} = \mathrm{span}\{a_1, a_2, \dots, a_n\}$. The natural number n is called the dimension of V and is written as $\mathrm{n} = \dim(\mathrm{V})$.

Definition A.3 Let V be a real or complex linear space. An inner product is an operation between two elements of V which results in a scalar, denoted by $\langle a_1, a_2 \rangle$ and has the following properties.

1. For each $a_1 \in V, \langle a_1, a_2 \rangle$ is a non-negative real number.
2. For each $a_1 \in V, \langle a_1, a_2 \rangle = 0$ if and only if $a_1 = 0$.
3. For each $a_1, a_2, a_3 \in V, and\ \alpha_1, \alpha_2 \in F$
   $\langle \alpha_1 a_1 + \alpha_2 a_2, a_3 \rangle = \alpha_1 \langle a_1, a_3 \rangle + \alpha_2 \langle a_2, a_3 \rangle$.
4. For each $a_1, a_2 \in V, \langle a_1, a_2 \rangle = \overline{\langle a_2, a_1 \rangle}$, where the overbar denotes the complex conjugate.