Spring 2014

A GPR-GPS-GIS-integrated, information-rich and error-aware system for detecting, locating and characterizing underground utilities

Shaui Li
Purdue University

Graduate School ETD Form 9
(Revised 01/14)

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Shuai Li

Entitled
A GPR-GPS-GIS-INTEGRATED, INFORMATION-RICH AND ERROR-AWARE SYSTEM FOR DETECTING, LOCATING AND CHARACTERIZING UNDERGROUND UTILITIES

For the degree of Master of Science in Civil Engineering

Is approved by the final examining committee:

Hubo Cai

Dulcy M. Abraham

Phillip S. Dunston

To the best of my knowledge and as understood by the student in the *Thesis/Dissertation Agreement. Publication Delay, and Certification/Disclaimer (Graduate School Form 32)*, this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

Hubo Cai

Approved by Major Professor(s): _____

_____

Approved by: Mike Kreger 04/11/2014

Head of the Department Graduate Program Date

A GPR-GPS-GIS-INTEGRATED, INFORMATION-RICH AND ERROR-AWARE SYSTEM FOR DETECTING, LOCATING AND CHARACTERIZING UNDERGROUND UTILITIES

A Thesis

Submitted to the Faculty

of

Purdue University

by

Shuai Li

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science in Civil Engineering

May 2014

Purdue University

West Lafayette, Indiana

望前驱之英华卓荦，应后起之勇力追踪。

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere thanks to my advisor, Professor Hubo Cai, for his patience, guidance, and financial support. Without his great advice, encouragement and active involvement, this thesis would not have been done. Professor Cai is always willing to help and strategically educate his student. I am indebted to Professor Cai for helping me hone my skills in research and study.

I would like to sincerely thank Professor Dulcy M. Abraham and Professor Phillip S. Dunston for being on my committee. I am grateful to Professor Abraham for providing valuable feedback on how to improve this thesis and helping me hone my professional skills.

I also want to thank my lab mate Xing Su and Chenxi Yuan. I sincerely thank Chenxi for teaching me wonderful stuffs, helping me doing field experiments and encouraging me when I am facing difficulties. I am thankful to all my friends back home in China and all the new friends that I made in America for their direct and/or indirect help, their understanding and their trust. Special thanks to my girlfriend, Shiyun Zhang, for her enthusiasm and support.

My highest gratitude goes to my parents for their continuous support and unconditional love throughout my life!

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

Figure Page

ABSTRACT

Li, Shuai. M.S.C.E., Purdue University, May 2014. A GPR-GPS-GIS-Integrated, Information-Rich and Error-Aware System for Detecting, Locating and Characterizing Underground Utilities. Major Professor: Hubo Cai.

Underground utilities have proliferated throughout the years. The location and dimension of many underground utilities have not always been properly collected and documented, leading to utility conflicts and utility strikes, and thus resulting in property damages, project delays, cost overruns, environment pollutions, injuries and deaths. The underlying reasons are twofold. First, the reliable data regarding the location and dimension of underground utility are missing or incomplete. Existing methods to collect data are not efficient and effective. Second, positional uncertainties are inherent in the measured utility locations. An effective means is not yet available to visualize and communicate the inherent positional uncertainties associated with utility location data to end-users (e.g., excavator operator). To address the aforementioned problems, this research integrate ground penetrating radar (GPR), global positioning system (GPS) and geographic information system (GIS) to form a total 3G system to collect, inventory and visualize underground utility data. Furthermore, a 3D probabilistic error band is created to model and visualize the inherent positional uncertainties in utility data.

Three main challenges are addressed in this research. The first challenge is the interpretation of GPR and GPS raw data. A novel method is created in this research to simultaneously estimate the radius and buried depth of underground utilities using GPR scans and auxiliary GPS data. The proposed method was validated using GPR field scans obtained under various settings. It was found that this newly created method increases the accuracy of estimating the buried depth and radius of the buried utility under a general scanning condition. The second challenge is the geo-registration of detected utility locations. This challenge is addressed by integration of GPR, GPS and GIS. The newly created system takes advantages of GPR and GPS to detect and locate underground utilities in 3D and uses GIS for storing, updating, modeling, and visualizing collected utility data in a real world coordinate system. The third challenge is positional error/uncertainty assessment and modeling. The locational errors of GPR system are evaluated in different depth and soil conditions. Quantitative linkages between error magnitudes and its influencing factors (i.e., buried depths and soil conditions) are established. In order to handle the positional error of underground utilities, a prototype of 3D probabilistic error band is created and implemented in GIS environment. This makes the system error-aware and also paves the way to a more intelligent error-aware GIS.

To sum up, the newly created system is able to detect, locate and characterize underground utilities in an information-rich and error-aware manner.

CHAPTER 1. INTRODUCTION

This chapter describes the backgrounds and research motivations, states the goal and objectives, and outlines the scope and organization of this thesis.

1.1 Background and Problem Statement

Detecting, locating and characterizing underground utilities are critical tasks throughout all the life cycle stages of building and civil infrastructure systems. Four motivating cases in the phases of design and planning, construction, operation and maintenance, and data management are presented below to highlight the importance of a system for utility detection, localization and characterization. In addition, the management of data uncertainty is also underscored.

Case 1: Subsurface mapping - a critical assistant to design and planning

The location of underground utilities is crucial for construction design and planning, which is highlighted by Federal Highway Administration (FHWA 2011) in a number of case studies. On a major highway project in Richmond, Virginia, the utility location data was collected by Virginia Department of Transportation (DOT)'s consultant to assist VDOT's roadway and hydraulics designers to determine possible highway utility conflicts. As a result, design changes were made and 61 of the 75 potential conflicts were

eliminated, saving $ 731, 425 worth of utility adjustment. In addition, the Virginia DOT credits subsurface utility engineering (SUE) for reducing 20% time needed for highway design. On a utility project in Columbus, Ohio, accurate utility information presented at the pre-bid meeting increased the bidder's confidence in the construction plans, resulting in a bid that was $400,000 less than anticipated. Additionally, there were no change orders, contractor claims, utility relocation and utility damages on the project owing to the correct utility information.

Great benefits can be gained in the design and planning phase with quality underground utility information. Various techniques are applied to collect the underground utility location data. However, most of them are destructive, time-consuming, and laborious. They are not productive in large-scale survey work and only provide limited information. For example, vacuum excavation methods are based on measurements at specific locations and unable to locate the utilities beneath pavement. Therefore, to assist construction design and planning, a system is needed for detecting and locating subsurface utilities in a way of timely-fashion, non-destructive and high coverage.

Case2: Damages to existing utilities during construction - a consequence of missing location data

Data regarding the true locations of subsurface utilities is often missing, incomplete, or in most cases, unreliable due to fragmented utility ownerships, poor data management, and utility relocation (Kumar and Sommerville 2012, Sterling et al. 2009, Lanka et al. 2002). The lack of reliable and accurate utility location data is the main reason for excavations

being the single largest cause of utility strikes, causing damages to existing utilities and thus high risks of cost overrun, project delay, environmental pollution, deaths and injuries (NTSB 2007, Felt 2007, Costello et al. 2007, Lester and Leonard 2007, Metje et al. 2007, Rogers et al. 2012). The cost of pipe damages amounts to billions of dollars each year (McMahon et al. 2005, NTSB 2007).

A utility is inadvertently struck or scratched by excavator every minute in the United States (CGA 2008). Industry statistics of underground utility damages indicated that of all accidents reported in 2011, 22% were due to poor locating practice (MUUDS 2011). Vermont Department of Public Service published a statistics of underground utility damages (2005-2009) showing 20.8% were due to facility location error and 21.2 % were excavation related (VDOPS 2010). The emerging trenchless technologies also impose high risks upon existing utilities. For example, Damage Information Reporting Tool (DIRT) reported 258 utility hits related with horizontal directional drilling (HDD) in 2005 across the United States (DIRT 2006). These unsatisfactory situations necessitate the need for a system that can accurately locate underground utilities to avoid utility strikes during construction activities.

Case 3: Restore and improve subsurface infrastructure - a grand challenge

The nation's civil infrastructures are aging and failing. The newly released 2013 Report Card from American Society of Civil Engineers (ASCE) graded various categories of U.S. infrastructures. The average grade was D+, which illustrates the growing need for restoring and improving urban infrastructures. Between 700,000 and 800,000 miles of

public sewer mains exist in the United States, many of which were installed after World War II and now are approaching the end of their useful lives. Because of the aging sewer mains and insufficient capacity, 900 billion gallons of untreated sewage were discharged into creeks and rivers each year. In addition, many catastrophic tragedies occurred due to ineffective infrastructure maintenance. For instance, the tragic 2010 San Bruno pipeline explosion was caused by a break which was not properly detected and repaired.

A main problem resides in maintaining and upgrading subsurface infrastructures is how to detect, locate and characterize underground utilities effectively. The existing infrastructures are buried in congested underground with unknown locations and dimension (radius), posing great difficulty for rehabilitation or renewal. The lack of a system to accurately retrieve location and dimension information of subsurface utility often renders the rehabilitation work ineffective. Among the millions of holes being dug every year in the United Kingdom to find buried utilities for maintenance, a significant proportion are dry holes i.e., excavations that fail to find the pipeline or cable being sought (Farrimond and Parker 2008). Such situations result in unnecessary traffic congestion, material wastage (e.g., material to landfill and new materials for reinstatement), visual intrusion, and wasted time and energy (Rogers and Cohn 2013). Therefore, devising methods for mapping and labeling buried infrastructure is critical for tackling the challenges in maintaining and upgrading subsurface infrastructures.

Case 4: Waste of collected data - a result of poor data management

Utility data in paper format is out of date and is of limited use in construction and asset management since it is not spatially context-aware and difficult for storage, real-time access and update. The emerging geographical information system (GIS) has been steadily replacing paper format drawings to create, inventory and manage utility data (Su et al. 2013). But it is still suffering from the absence of depth values and attributes. For instance, GIS typically models underground utilities as 2D polyline in GIS, missing the vertical depth information (Su et al. 2013). However, the depth is critical for the construction of accurate 3D plans (TTN 2002).

The depth of some buried utilizes can be derived through the reference surface and buried depth (Su et al. 2013). However, such derived depth, usually stored as attribute in GIS, is unreliable since the reference surface and/or the depth of cover are very susceptible to changes. Hence, lack of reliable vertical information renders the utility data worthless, which wastes data collection efforts and collected data. Utility attributes such as pipe radius, are of importance to decision-makers to reach better decisions. For instance, more accurate excavation area can be marked out for utility damage prevention, if utility radius is known.

Positional uncertainties are inherent in the collected data. Identification and modeling of utility positional uncertainties are the bedrocks of proximity queries and monitoring during excavation activities to improve safety and prevent utility damages. Visualization of positional errors in an easy-to-understand manner communicates the uncertainties to

end-users effectively and efficiently, which ensures their operating in an informed state. Therefore, an urgent need exists for modeling and visualizing the uncertainties associated with measured utility locations in 3D space. However, the current practice of GIS is deterministic and falls short of modeling and visualizing positional error/uncertainty for underground utility.

Both the utility and the construction industries call for an intelligent data management approach to extend the current 2D utility data to 3D and to integrate positional error/uncertainty and additional utility attributes into the decision making process.

A Substantial number of similar cases exist in every phase of infrastructure life cycle management. We conclude that:

1) In civil infrastructure management, location and dimension data of subsurface utility is missing or in poor quality. However, such data is really a necessity.
2) Traditional methods for data acquisition are destructive, laborious, time-consuming, and fail to provide adequate information.
3) The current poor data management approach wastes data collection efforts and collected data. It fails to handle the inherent positional uncertainty of utility data and thus is not adequate to support decision-making throughout the life cycle of civil infrastructures.

We argue that we need to create an effective detecting, locating and characterizing system and improve GIS data management approach. We envision that such system can collect data and retrieve useful information in a non-destructive, reliable, responsible,

responsive and respected way. The GIS data management approach with its capability to house uncertainty-aware data is more intelligent. Figure 1.1 summarizes the research motivations.

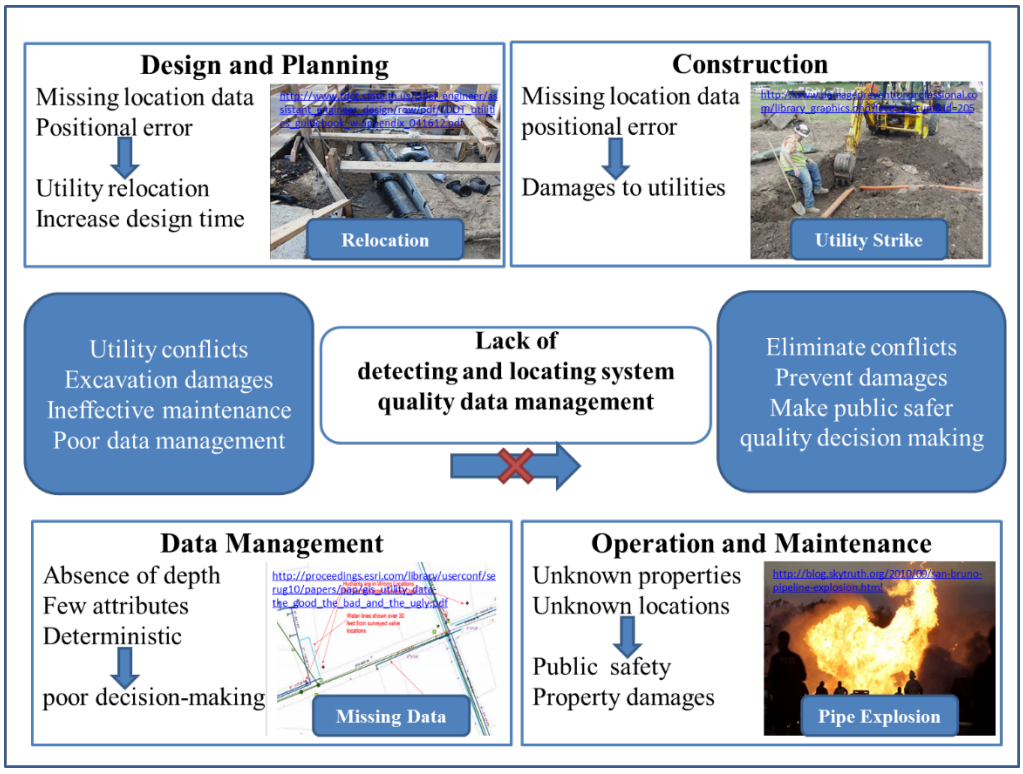


Figure 1.1 Research motivations

1.2 Goal and Objectives

Before stating our research goal and objectives, it is necessary to explain the implication and extension of the term "detecting, locating and characterizing". Detecting implies discover or discern the subsurface utilities from its surrounding environment. The term "locating" refers to the determination of the object's position. There are different implications for different entities. For a contractor or owner, locating means the process

of marking the objects out for damage prevention or further assessment. For an engineer, this term means the process of precisely and accurately measuring and documenting its three-dimensional location (Sterling 2009). This engineers' implication is adopted in our research. The term "characterizing" means the evaluation of object's attribute such as the radius of underground utilities.

The overarching goal is to create a quantitative, information-rich and uncertainty-aware system for detecting, locating, and characterizing underground utilities using ground penetrating radar (GPR), global positioning system (GPS) and geographical information system (GIS).

To achieve this goal, the specific research objectives are to:

1) Address GPR data interpretation challenge through signal and image processing.
2) Address data registration challenge through integration of GPR, GPS and GIS.
3) Address data quality issue through error measurement and uncertainty modeling.
4) Validate the newly created system in field experiments.

Among the line of sensing technologies and methods such as radio frequency (RF) detection, electromagnetic methods (EM), acoustic emission method, and magnetic methods, GPR was chosen as the sensor for detecting, locating and characterizing subsurface utilities for the reasons listed below.

1) GPR is capable of detecting nonmetallic objects, which may be the biggest advantage of GPR over other magnetic sensor technology (Sterling 2009).

2) GPR has a sound detectability (Sterling 2009). GPR does not require a physical connection to the utility. It does not rely on the detectable EM fields radiated by the utility itself.
3) GPR has the highest resolution among the geophysical methods, and it is possible to reach centimeter scale resolution (Jeong 2003).
4) GPR raw data can be processed to retrieve size, geometric characteristics, and spatial orientation in addition to location information (Herman 1997).
5) GPR also has the merits of fast data acquisition, cost effectiveness for mapping large areas, and better results compared to other non-destructive technologies (Jaw et al 2013).

Despite these merits, we have identified three challenges of GPR to be addressed in this research. The first challenge is data interpretation - the retrieval of useful information from GPR raw data. The second challenge is geo-registration - the relay of GPR detected location to the real world geospatial coordinates. The third challenge is data quality issue - error measurement and modeling.

1.3 Scope and Organization

The remainders of this thesis are organized as follows. Chapter 2 reviews related studies, and identifies the knowledge gaps. Chapter 3 describes the main components of the novel system, i.e. GPR, GPS and GIS, and how they are integrated as a total system. Chapter 4 mainly elucidates the techniques used to interpret GPR and GPS raw data. GPR raw data are processed and extracted from original GPR scan; a generic hyperbola equation is

proposed to model GPR raw data, and a novel very important point (VIP) algorithm is created to estimate location and radius of buried utilities with auxiliary GPS data. The proposed method is validated by field experiments. Chapter 5 measures and assesses the locational errors of the newly created system, and put forward a 3D probabilistic error band model to handle the positional uncertainties of underground utilities. Chapter 6 summarizes the finding, conclusions, and point out the direction for future research.

CHAPTER 2. LITERATURE REVIEW

This chapter first reviews a number of techniques for mapping underground utilities and justifies the utilization of GPR, GPS and GIS to create the novel system. Following that are literature reviews with an emphasis on GPR raw data interpretation, geo-registration with GPS and GIS data management. Thereafter, three knowledge gaps in the current practice are identified and summarized towards the end of this chapter.

2.1 Techniques Used for Mapping Underground Utilities

There are a variety of methods capable of detecting, locating and characterizing underground utilities. They can be categorized into destructive and non-destructive techniques. Each technology has its own merits and demerits. A brief review of these techniques is as follows.

2.1.1 Destructive Methods: Excavation

The most accurate approach to measure the horizontal and vertical locations of underground utilities is to expose it by excavation (Sterling et al. 2009). However, the varied excavation methods all bear some risks of damaging existing utilities. This is even true for the emerging vacuum excavation. The vacuum excavations, including air vacuum and water vacuum excavation, use a powerful vacuum to suck soil from the test hole. It

requires pavement or concrete to be removed first by jackhammers, rock drills or concrete saws. Hence, high risks of damages arise to shallow utilities or those embedded in the roadway. Furthermore, vacuum excavations are labor-intensive, time-consuming, and fall short in large scale survey.

Sterling et al. (2009) identified three limitations for the vacuum excavation.

1) The vacuum excavations may damage the wrappings and coatings on cathodically protected gas lines since a great force is applied to break up the surrounding soils.
2) Another concern with water vacuum is the subsequent soil compaction and paving integrity. The saturated soil is not suitable for backfill. The soil surrounding the test hole may be disturbed by water saturation, which may result in ground settlement. In addition, the introduced moisture around a pipe may lead to potential corrosion.
3) Operational difficulty is certain to arise during water vacuum excavation when the air or ground temperature is below freezing.

The most challenging point is that if the horizontal locations are unknown or inaccurate, how can surveyor know where to excavate? Recently, the priority of non-destructive methods over destructive methods (excavation methods) for locating underground utilities is underscored by US Environmental Protection Agency (USEPA) (Thomson et al. 2009). Therefore, a host of non-destructive sensing techniques for mapping underground utilities are introduced and discussed below.

2.1.2 Non-Destructive Methods: Sensing Technology

Sterling et al. (2009), Costello et al. (2007), Metje et al. (2007) and Jeong et al. (2003) investigated a number of sensing technologies for detecting and locating underground utilities. These include pipe and cable locator, terrain conductivity method, infrared method, resistivity measurement, magnetic method, elastic wave method, radio frequency identification and GPR. A brief description, advantages and disadvantages for each of these eight technologies are provided as follows.

1) Pipe and Cable Locator

Pipe and cable locator is the most common instrument for detecting and tracing underground utilities. It is based on electromagnetic theory. A transmitter emits an electromagnetic wave and a receiver is tuned to detect changes in the wave. Once the wave encounters a metallic object, an electromagnetic current is generated on that object, creating a magnetic field around that conductor. This magnetic field can be detected by the receiver and thus the operator is able to detect the subsurface utility. The advantages of pipe and cable locator are its low cost and effectiveness in tracing metallic utilities. A major concern of pipe and cable locator is the limited range of detectable materials. Non-metallic utilities without the assistance of wires or installed metallic tape cannot be detected using this method. Additionally, it cannot measure the depths of buried metallic pipes.

2) Terrain Conductivity Method

Terrain conductivity (TC) method is also based on the electromagnetic theory. TC detects utilities by measuring the average conductivity of a cone-spaced volume beneath the transmitter and receiver. The effective penetrating depth is typically 15 to 20 feet. This technology is built on the premise that there is a substantial difference of conductivity between utilities and their surrounding medium. It is an effective detection method to find metallic utilities in a non-congested and dry environment. The disadvantages of this technology are the limited detectability, the massive and complicated data interpretation, and incapability of depth estimation. In areas with high moisture or high water table, it may be impossible to detect any kind of utility unless the utility is watertight, empty, large, and relatively shallow. Massive amount of data with different antenna orientations are needed for detection and tracing. Furthermore, estimation of utility depth with TC method is not realistic.

3) Infrared Method

Infrared method is mainly used to detect utilities that have operating temperature distinguishable from the temperature of the surrounding soils. The temperature difference can be detected at the ground surface using an infrared camera. The infrared method is rarely used or just for very specific scenarios because of its inherited limitations. The deeper a utility is buried the less chance it can be detected at the surface due to the decreasing temperature difference. Infrared methods are difficult to use and interpret in a congested urban environment with lots of cement paving. This technology is also largely

influenced by climate, site, geology and utility conditions. More importantly, Infrared method provides no depth-information.

4) Resistivity Measurement

In Resistivity measurement, a direct current is injected into the ground using two or more electrodes, the resultant voltages are then measured and the average resistivity is calculated. The measurement depth is determined by the spacing of electrodes. A utility with a resistivity different from that of the surrounding medium will be detected given enough data. However, resistivity measurement is more of a searching technique than a tracing technique. The data processing procedures is cumbersome and the rate of data acquisition is slow. This method is rarely employed. In most cases, it is used for other tasks while the detection of utility is just a by-product.

5) Magnetic Method

Iron is a material commonly used in pipe fabrications, which opens the opportunity for using magnetic properties to detect and sometimes trace an iron or steel pipe. There are two types of magnetic surveys: total field and gradient. The total field magnetic survey is usually used for environmental surveys, and rarely employed as a utility localization method. With the gradient method, as the detector moves close to a magnetic object, the shape and intensity of the magnetic field creates an interpretable reading in the equipment. Magnetic method can detect only ferrous metallic utilities. It is difficult to detect pipes buried several feet below the surface due to low initial magnetic strength. Several

uncontrollable factors, e.g. the object shape, internal structure and purity of materials will affect the performance. Depth estimation is impossible with magnetic method.

6) Elastic Waves

When pipes are nonmetallic and a metallic conductor cannot be inserted into it, elastic waves methods may be employed to detect and trace the utility. The elastic wave must be introduced into the medium first and detected after its reflections or refractions occurred due to buried structures. Basically, there are three techniques for imaging utilities, i.e. seismic reflection, seismic refraction and acoustic emission. Seismic reflection and refraction are only useful in specialized conditions and must follow rigorous procedures. Most utilities are too small to be detected by the large wavelength of seismic waves. Acoustic emission is fairly for tracing nonmetallic water lines; however it is less useful as a searching technique. During measurements, it can be influenced by background noise such as traffic noise. It also needs access to pipeline in active mode. The most non-negligible limitation is that acoustic emission has no depth-estimation capability.

7) Radio Frequency Identification method

Radio frequency identification (RFID) technology is rapidly developed as a method to detect, locate or even characterize utilities at a specific point. Dziadak et al (2009) proposed a model for locating buried assets based on RFID technology. Kumar and Sommerville (2012) created an algorithm to obtain 3D location of buried utilities using GPS and RFID. The RFID device that consists of a reader and several tags works in such a way. The reader use radio frequency (RF) waves to transmit a signal, which activates

the tag attached on a buried utility. Once the tag within the reading zone has been activated, the tag transmits data back to the reader. After appropriate processing, the utility location can be estimated. However, the attachments of tags to utilities remain a practical problem. This is especially true for small utilities. The presence of high moisture and metallic materials will negatively affect RFID performance. The various size and shape of detection zones also render this technique relative inaccurate for searching and locating utilities.

8) Ground Penetrating Radar

Ground Penetrating Radar (GPR) is a well-accepted geophysical technology for detecting and locating underground utilities. As a chosen technology in this research, a detailed description of GPR with regards to its components, principles and data formats is given in section 2.3 and Chapter 3.

2.1.3 Selection of Subsurface Sensor

Despite an impressive number of sensing technologies, there is not a total solution for detecting and locating underground utilities since each sensor has its limitations. Table 2.1 lists the main concerns for all eight technologies considered in this study. GPR is chosen to create our detecting and locating system because it overcomes the most concerned limitations of other sensors in mapping underground utilities.

1) Many of these technologies are only used for detection. Several of them can provide 2D (horizontal position) data. Few are capable of estimating the buried depth.

However, the third dimension, depth, is crucial for developing accurate 3D plans for designing and planning, construction and asset management. With GPR, it is possible to derive vertical depth of a buried utility.

2) Most of them cannot achieve high localization accuracy (Costello et al. 2007, Cullen 2005, Dziadak et al. 2009). Authorities such as American Society of Civil Engineering (ASCE) (2002) and National Underground Asset Group in UK (2006) recommended an accuracy of ± 100 mm in three dimensions for locating buried utilities. Among the existing technologies, GPR has great potential to meet the requirement due to its high accuracy and resolution.

3) They often have a limited range of detectable materials. Some can only locate metallic or ferrous pipes while some are more suited for nonmetallic pipes. While theoretically, GPR can detect all the pipes fabricated with different materials.

4) Currently, few of them can retrieve geometric information, spatial orientation of the utility from the collected data. However, through appropriate processing, this is possible with GPR.

Table 2.1 Limitations of sensing technologies for locating underground utilities

Technologies	Limitations
Magnetometer	• Can only detect ferrous utilities buried shallowly • Depth estimation is not considered realistic
Pipe and Cable Locator	• Can only detect metallic utilities • Possible interference by nearby metallic utilities
Terrain Conductivity Method	• Unable to detect utilities in high moisture environment • Massive amount of data are needed for processing • Depth estimation is not considered realistic
Infrared Method	• Unable to detect utilities buried deeply • Not suitable for congested urban environment • Depth estimation is not considered realistic
Resistivity Measurement	• May not be useful as a trace technique • Data setup and collection is cumbersome • May not be applicable for mapping paved areas
Elastic Waves	• Most utilities are too small to be detected • May be influenced by background noise • Depth estimation is not possible with acoustic emission
RFID	• Attachment of tag to existing utilities is difficult • Negatively affected by moisture and metallic materials • Detectability and accuracy are relatively low
GPR	• Large signal attenuation in conductive soils • Highly skilled crews are needed for data interpretation

2.2 GPS and GIS for Mapping Underground Utilities

GPS and GIS, though not detecting techniques, are crucial for the documentation of 3D locations of detected utilities and associated data management tasks.

2.2.1 GPS for Locating Underground Utilities

GPS is an increasingly common way to obtain the horizontal position of buried utilities, which is requested by many project owners to acquire permanent record of utilities locations (Jeong et al. 2003). Several benefits of GPS are listed below to explain why it is necessary for locating utilities (USNCO 2013).

1) GPS can significantly improve survey productivity in terms of time, equipment and labor required.
2) GPS has few operational limitations compared to conventional techniques, e.g. GPS surveying is not constrained by line-of-sight visibility between survey stations.
3) GPS supports real-time accurate positioning, mapping and modeling of the physical world.
4) Information collected with GPS can be delivered to GIS that can inventory, manipulate and display geographically referenced data.
5) The location of utilities can be registered to geospatial referencing system by GPS as real-world coordinates for future analysis.

With respect to locating underground utilities using GPS, there are two main scenarios. When the underground utilities are exposed by excavation or before burying during installation, the GPS can work alone to obtain 3D locations of the utilities as a permanent records. However in most cases, the utilities are buried with unknown positions. In such a scenario, GPS often works with other sensing techniques such as GPR to first detect the utilities and then locate them. The horizontal coordinates are obtained by GPS and the buried depths are derived by GPR.

The integration of GPS and GPR to produce 3D displays of underground utilities are highly recommended (DiBenedetto et al. 2010, Sterling et al. 2009a, Sterling et al 2009b, Manacorda 2007). One distinct advantage of such integration is that the location of detected utilities can be register to real world spatial referencing system. As indicated by

Su et al. (2013) and Lew (2000), the utility depths are rarely referenced to a recognized elevation datum. This missing step results in inaccurate or unreliable location data if the cover of the utility is changed due to excavation, road construction or erosion. With the help of GPS, the derived depth can be converted to a vertical position that references a vertical datum, and thus making the collected data more reliable. Furthermore, GPS also serves as a bridge transferring GPR data to GIS, which automates the inventory and update of utility location in GIS, and guide future field localization.

2.2.2 GIS for Mapping Underground Utilities

GIS is an emerging effective tool to inventory and manage utility data. Environmental System Research Institute (ESRI) (1995) defined GIS as "an organized collection of computer hardware, application software, geographic data, and personnel designed to efficiently capture, store, update, manipulate, analyze, and display all forms of geographic referenced Information." It is also an information system that embraces advanced analysis, modeling and prediction capabilities to work with data referenced by spatial or geographical coordinates (Huxhold 1999). Due to its distinguished merits such as integration of spatial and non-spatial information from different sources and in different forms, registration of locations to real world coordinates, advanced spatial analysis capabilities, and etc., GIS becomes an efficient and effective data management tool in civil engineering (Su et al. 2013, Poku and Arditi 2006).

Currently, computer aided design and drafting (CADD) system is the prevailing utility data management system, but GIS appears to be the next generation of utility data

management system. Utility industry is a geospatial information-intensive sector. For instance, more than 80% of all the information associated with water and waste water utilities is geographically referenced (Shamsi 2002). The utility data proliferates and accumulates throughout all the life cycle stages. On that regard, GIS is a more suitable approach for storing, tracking and updating the utility information (Yelakanti et al. 2003).

Many utility owners have undergone the transition from paper maps or CADD to GIS for creating, organizing and managing geospatial utility information (Su et al. 2013, Crawford 2012, Jeong et al. 2004). This is due to the distinguish advantages of GIS over traditional CADD systems. The easy data transformation with GPS makes GIS a more expedient way to create and inventory geospatial utility data. In addition, in traditional approach the spatial data (e.g. the alignment of utilities) are maintained in CADD while the attribute data (e.g. sizes, materials and conditions of utilities) are stored in database software such as MS Excel or MS Access. Because of the separation, it requires much more time and effort to store and update the data (Yelakanti et al. 2003). However, GIS provide the platform for integrating the geospatial and attribute data, hence cutting down the time consumed and potential errors. Moreover, the integrated data can be used to produce a new set of data in tabular or visual formats to assist utility design, construction, inspection and rehabilitation (Jeong et al. 2004, Shamsi 2002).

The merits of GPS, GIS and GPR are supplementary to each other. The integration of GPR, GPS and GIS composes a total system for mapping and managing underground utilities. In such a system, GPR detects underground utilities and estimate their buried

depth and radius. GPS is used to register the location of detected utilities to real world coordinates. The GIS is responsible for managing the collected utility data.

2.3 GPR for Detecting, Locating and Characterizing Underground Utilities

Utilization of GPR for underground utility localization and characterization began in the 1960s with the advent of plastic gas pipes (Sterling et al. 2009). Since then, much research effort has been poured to facilitate and improve GPR performance in locating and characterizing underground utilities. This section introduces the related work and aims to identify the challenges needed to be addressed in existing methods. The challenges faced by using GPR for mapping underground utilities are discussed in the aspects of raw data interpretation and error measurement. While the challenges of hardware development, selection or optimization of GPR equipment are not considered in this research.

2.3.1 Locating and Characterizing Underground Utilities

In addition to the detection of buried utilities, some detailed information such as depth, radius and materials of the utilities are needed in many scenarios because they are crucial for effective planning, excavation damage prevention and asset management. Olhoeft (2000) proposed to maximize the information returned from GPR raw data through appropriate data analysis techniques. The most significant information may be the utility radius (R) and GPR wave propagation velocity (v). The depth of the underground utility can be calculated as the product of wave propagation velocity v and its travel time t. The

time t is directly available from GPR scans and thus, the estimation of depth is transferred to the estimation of EM wave velocity v.

The first attempt in estimating R was proposed by Stolte and Nick (1994) who derived the functional dependence between radius and hyperbolic eccentricity (the reflection of cylindrical utility in GPR scan is hyperbolic). Shihab et al. (2004) found that hyperbolic eccentricity is only a function of v. The direct relationship between velocity (v) and radius (R) was the main reason for the false functional dependence between R and eccentricity in the study by Stolte and Nick in 1994. A number of studies have investigated the estimation of utility radius and GPR wave velocity from GPR raw data. Some of the most noteworthy works are referenced and commented on.

1) Hough Transform

Many studies have employed the Hough Transform to estimate the velocity and radius of the buried cylindrical utility based on hyperbolas present in GPR raw scans. The Hough Transform was patented by Hough (1962) and extended by Duda and Hart (1972) to model simple geometrical shapes in binary images through a voting procedure (Maas and Schmalzl 2013). In the context of processing GPR scans, the Hough Transform is used to characterize the parameters of the hyperbola. Specifically, the hyperbola is described by a set of parameters, e.g., the radius and EM wave velocity, via explicit mathematical expressions. All the GPR raw data points on the hyperbolic arc are transformed from the GPR scan into the parameter of Hough space (Maas and Schmalzl 2013). Thereafter, the

parameters, i.e., radius and velocity that fit the hyperbola best are determined by selecting the local maximum in the parameter space.

Windsor et al. (2005 a, b) and Li et al. (2012) discussed two scenarios when applying Hough Transform, depending on whether the EM wave velocity is a known priori. Windsor et al. (2005b) used a known velocity and Li et al. (2012) measured the range of velocity based on the electromagnetic characteristics of surrounding soils. The classical Hough Transform was successful in deriving the radius under this known EM wave velocity scenario. However, under the unknown EM wave velocity scenario, the solution of Hough Transform was unsatisfactory, as demonstrated by Windsor et al. (2005 a, b).

Several studies have also attempted to refine the Hough Transform to increase its accuracy and efficiency under the unknown velocity scenario. For instance, Brogioli et al. (2008) extended the Hough Transform by giving higher weights to optimally placed sets of data pairs than the "ill-conditioned" sets. Maas and Schmalzl (2013) used the Viola-Jones algorithm to narrow down the hyperbola location to certain areas and thus reduced the number of inputs for the computationally expensive Hough Transform.

Despite a number of advantages of the Hough Transform, such as its reliability in noisy data sets (Maas and Schmalzl 2013), three main limitations constrain its use in GPR applications. Firstly, because of the strong correlation between parameters (e.g., radius and velocity), the robustness and accuracy of estimation cannot be guaranteed (Windsor et al. 2005b). Secondly, the Hough Transform has a high computational requirement,

especially for a massive amount of raw data (Maas and Schmalzl 2013). Thirdly, a large input dataset is required to calculate radius even if wave velocity is known a priori (Ristic et al. 2009).

2) Mathematical Fitting

The underground utilities often present hyperbolic reflections in GPR scans. Hence, the second approach is based on a mathematical fitting procedure of hyperbola. The cylinder radius and wave propagation velocity are estimated from the geometry interpretation of the fitted hyperbola (Naganuma et al. 2011, Chen and Cohn 2010 a b, Ristic et al. 2009, Dolgiy et al 2006, Shihab and AI-Nuaimy 2005).

Shihab and AI-Nuaimy (2005) proposed a direct conic least-square fitting technique to retrieve geometric information from single radargram, which can be considered as an extension of the work done by Fitzgibbon et al. (1999), O'Leary and Zsombor-Murray (2004). Dolgiy et al. (2006) used several techniques (i.e. weighted least square method, the recursive Kalman filter, the maximum likelihood method, the direct least-square fitting method and the Nelder-Mead direct search method of optimization) to fit the hyperbola and estimate the radius. They recommended the recursive Kalman filter technique because it is the most expedient for practical application. However, the main limitation is that this technique is based on a priori known v that is acquired using the common midpoint (CMP) or point reflector analysis methods.

By using a nonlinear least-square method, Naganuma et al. (2011) proposed an estimation method of the geometric condition of buried pipes. Chen and Cohn (2010 a) created an algorithm based on algebraic distance fitting and applied a probabilistic conic mixture model to mining GPR data. Later on, the author noticed this algorithm is not applicable for GPR data with relative large amount of noise. Then they extended the previous algorithm by using a more robust orthogonal distance fitting algorithm in the probabilistic mixture model which handles the noise nicely (Chen and Cohn, 2010 b).

Ristic et al. (2009) presents a new method to simultaneously estimate the radius and velocity based on a nonlinear least squares fitting procedures. The author used a modified Levenberg-Marquardt method to estimate the hyperbola apex in order to reduce the number of correlation between parameters. Then the boundary of the velocity is determined according to zero and a predefined radius. Finally, an optimal velocity is selected in the solution space to calculate radius R. The author claimed that his algorithm is more accurate and robust with regard to noise and large amount of raw data. However, one potential concern is that this algorithm needs to predefine a maximum R and a searching step, which may in turn affect the accuracy as well as the computational effort.

3) Interactive Technique

The third approach is based on an interactive interpretation of hyperbolic reflection from radar scans. The interactive technique described by Olhoeft (2000) involves visual overlap of a predefined hyperbola with the hyperbola in the GPR scan. The velocity and the depth are determined by the angle between hyperbola asymptotes. R is estimated from

the curvature of the hyperbola apex with human intervention. However, this method do not thoroughly characterize the hyperbola in terms of semi-major and semi-minor axes, thus fall short of providing necessary information for geometric characterization. According to Olhoeft (2000), the processing results are not unique, and a high-quality scan is crucial. Moreover, the determination of the exact hyperbola center position is another problem. Yufryakov and Linnikov (2006) also use the interactive procedure to estimate the radius, velocity as well as spatial orientation parameters from a 3D scan based on a detailed geometrical description.

4) Artificial Intelligence

One example cited for artificial intelligence methods is the work of Shaw, et al. (2003), which developed a neural network approach to automatically estimate the radius of rebar in concrete. However, this approach may not be applicable for estimating pipe radius, although both pipes and rebar are cylinders. The main reason is that it is extremely difficult or even impractical to acquire sufficient training data because of the heterogeneity of the subsurface soils and different utility radius.

2.3.2 Error Measurement and Modeling for GPR Data

Measurement and modeling of GPR locating errors has received very little attention. The locating error is defined as the difference in distance between the estimated position and the actual position of the utility. Most of the researches (including those reviewed before) only conducted error measurement while missed the error modeling part. For instance, Jaw and Hashim (2013) measured the locating error using different scanning methods for

data acquisition, namely: perpendicular-to-pipe scanning, along-pipe scanning and variation-angles (30°, 45° and 60°) scanning. Their experiments showed the along-pipe scanning method yields the highest accuracy. The locating errors of shallow utility are 0.098 m for horizontal position and 0.095 m for vertical position. However, few of the researches have attempted to model the locating error, which is considered as a knowledge gap.

2.4 GIS for Utility Data Management

Although the proliferation of GIS in utility data management has been seen in recent years, there are still several limitations need to be addressed.

1) Utility are modeled as 2D polylines and the depth value is missing (Arnott and Keddie 1992, Halfawy et al. 2005, Chasey and Cowan 2008).
2) Attributes regarding the utility such as radius are seldom recorded.
3) It lacks a method for managing and rendering positional inaccuracy/uncertainty of utility data (Su et al. 2013).

The first two drawbacks of current GIS practice are mainly due to the absence of data. With an effective detecting, locating and characterizing system for data acquisition, these two limitations can be tackled. Therefore the limitation lies in how to handle positional uncertainty of linear utility lines, and model and render it in GIS to support decision-making in infrastructure management. The problem is further narrowed down to uncertainty modeling of linear objects in GIS. A brief review is as follows.

A promising approach to address the limitation stated before is an "error-aware" GIS (Duckham and McCreadie 2002, 1999), which can be aware of the uncertainty in utility locations and capable of supporting visualization and analysis of that positional uncertainty. To develop such an "error-aware" GIS for utility data management, the main task is uncertainty modeling of the linear, geospatial utility in GIS.

The uncertainty is defined as the positional discrepancy between the records-indicated locations and their real world locations. The term "uncertainty" is interchangeable with positional error/inaccuracy (Su et al. 2013, Goodchild 1998). Several error models for line in GIS were proposed. In 2D, the uncertainty of a straight line is initially modeled as an error epsilon band. This concept is first proposed by Perkal (1956), who defines a 2D error region enclosed by two parallel lines. However, this epsilon band is deterministic because it assumed that the true line is definitely within the band that has a uniform width (Tong et al. 2013). Goodchild and Hunter (1997) proposed a simple buffering approach that can estimate the percentage of lines within the buffer and thus is capable of evaluating the positional accuracy.

Generally, there are two different normality assumptions with regard to errors of the endpoints of a line segment. Different assumptions lead to different error models. One normality assumption is that the errors of two endpoints of a line segment are independent and obey a two dimensional normal distribution. Based on that, Caspary and Scheuring (1993) applied the error propagation law to the points on the line segment and developed an error band model. Shi (1994) derived the error distribution of the

intermediate points on the line segment and proposed a confidence region model. The confidence region of the line segment is the union of the confidence regions of all the points on the line segment. Furthermore by assuming the confidence region of each point on the line is a rectangle, Shi (1994, 1998) modeled the confidence region of a line segment as a rectangle-based shape.

The error ellipse model is another kind of error model. On the basis of another assumption that the error of two endpoints are correlated and follow a four dimensional normal distribution, several error model were proposed. Shi and Liu (2000) introduced a more generic error band model, known as "G-band", based on stochastic process theory. The nature of error at each point is indicated by an error ellipse. In this model, a number of error ellipses of the points on the line segment are drawn around the line segment. The G-band is defined as the envelope of these ellipses. Probabilities can be determined for G-bands with various sizes to model the uncertainty in lines (Heuvelink et al. 2007, Wu and Liu 2008). Chapman et al. (2003) derived the error ellipses of arbitrary points on a line segment and constructed the error band from the ellipses of points on the line segment. However, such ellipse-based error band cannot provide the probability of the line segment falling within the error band (Tong et al. 2013).

In addition to analytical methods, some studies also utilize simulation method for error modeling in GIS (Dutton 1992, Zhang and Goodchild, 2002, Zhang et al. 2006, Su et al. 2013, Tong et al. 2013). Dutton (1992) observed the error distribution of line segments based on Monte Carlo simulation. He assumed the error of the points on the line segment

follow a circular normal distribution where the errors are not correlated with respect to x and y coordinates and the error values for both coordinates are the same.

Zhang and Goodchild (2002) assumed the error of endpoints follow a two-dimensional normal distribution and conducted thousand simulations to generate the density surface of the line segment. Zhang et al. (2006) applied Monte Carlo approach to simulate the probability density function of a line segment with the assumption that the error of endpoints follows a two-dimensional normal distribution.

Su et al. (2013) proposed a 3D Probability G-band assuming error is normally distributed in all dimensions and performed simulations to observe the error band. The results confirmed that 3D ellipsoids can be used to model the error of points on line. However, the authors didn't mathematically describe the 3D Probability G-band. They suggested such 3D Probability G-band can be simplified to a probability band of uniform size along the line for practical application. Tong et al. (2013) incorporated analytical and simulation methods and proposed a statistical simulation error model by taking the line segments, instead of points, as variables.

However, with regard to modeling positional uncertainty of linear utility lines in GIS, the current researches involve certain limitations. Most of the error models for lines in GIS are 2D. To accommodate the positional uncertainty for 3D utility location data, the 2D error model need to be extended to 3D. Another limitation is that the current error model focused on the shape and size of the error bands while lack a method to estimate the

probability of the lines falling within the error band. The capability of probability estimation is crucial since this information can be further communicated to end user such as excavator operator to support decision making.

2.5 Knowledge Gaps and Research Tasks

Knowledge gap and research task 1: interpretation of GPR raw data

The existing methods (i.e. Hough transform, curve fitting, interactive techniques and artificial intelligence) all have certain limitations for effectively estimating the depth and radius of underground utilities.

1) The current methods for estimating radius and depth is not practical because almost all the existing methods require GPR antenna to move perpendicularly to the buried pipe or need two parallel slice of GPR scan. If the horizontal position of utility is absent, it is not possible to move GPR perpendicularly to the buried pipe. It is also time consuming and operational difficult to acquire two parallel scan.
2) It lacks an effective algorithm to simultaneously estimate wave velocity and utility radius. Current methods are either too computational expensive for on-site application, or need a priori known velocity.
3) Few of the existing methods provide the information regarding utility spatial orientation.

Therefore, the research task 1 is to close this technical gap by creating a novel algorithm to effectively retrieve the wave velocity, radius, spatial orientations of the buried pipe from GPR raw data in a more practical and expedient way.

Knowledge gap and research task 2: error measurement and modeling

The second gap in current knowledge is positional error modeling. Most current researches have not investigated the patterns of GPR locating errors. Research questions such as "is there any functional dependency between errors and buried depth of utilities?" has not been resolved. Therefore, the corresponding research task is to:

1) Measure the locational error of GPR in different scenario, i.e. in different buried depths and different soils.
2) Examine the error patterns and model the locational errors.

Knowledge gap and research task 3: Integrated GPR-GPS-GIS (3G) system

The third knowledge gap is the lack of a GPR-GPS-GIS (3G) integrated system with all the components functioning effectively and efficiently for detecting, locating and characterizing underground utilities, and managing the collected data in an uncertainty-aware manner. The specific research tasks for closing this knowledge gap are listed as follows.

1) Integrate GPR, GPS and GIS to form a novel system for underground utility mapping.
2) Create an error/uncertainty model in GIS to manage the positional uncertainty of underground utilities and visualize it in GIS environment.

CHAPTER 3. DETECTING, LOCATING AND CHARACTERIZING SYSTEM

This chapter describes the principles and roles of GPR, GPS and GIS. Section 3.1 mainly introduces GPR fundamentals in aspects of system components, working principles, data formats and reflection patterns of underground utilities in GPR scans. Section 3.2 illustrates locating principle of GPS. Section 3.3 describes functions of GIS and multi-patch technique for 3D object visualization.

Figure 3.1 clearly illustrates the system configuration conceptually. The system hardware consists of GPR and GPS. GPR is utilized to detect buried utilities, estimate buried depth and radius of the detected utilities through signal and image processing. GPS is employed to register the location of detected utilities to real world coordinate system, and transfer data to GIS for further analysis. The GPS is also used to assist GPR raw data interpretation, which will be discussed in detail in Chapter 4.

The measurements with GPR will involve inevitable errors, which need to be considered and modeled. Hence, the positional errors are evaluated in terms of utility buried depth and soil conditions. An error model for the linear, geospatial underground utilities is designed to make the system error-ware. The system software consists of the error model, the algorithm used to retrieve information from GPR and GPS raw data, and the GIS

platform for inventory, visualize and update utility data. The location and radius of buried utilities estimated from GPR and GPS raw data and the associated error information are relayed to GIS for management and visualization.

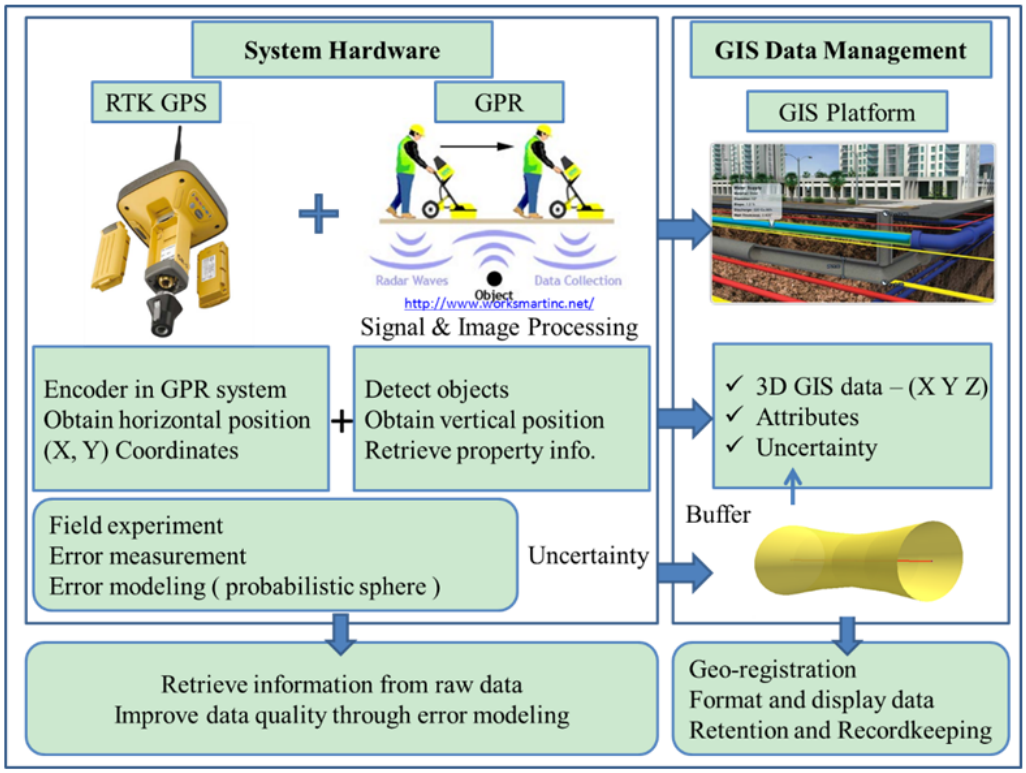


Figure 3.1 GPR-GPS-GIS system overview

3.1 GPR Fundamentals

A brief description of GPR fundamentals is presented as follows.

3.1.1 Components of GPR System

This section describes the main components of GPR with an emphasis on their functions and interactions with each other. It builds up an overall picture of how this system works.

Figure 3.2 shows a typical GPR system with its five main parts interfacing and communicating with each other. These five components are encoder, electronic unit, monitor (PC), control unit and antenna. The functions of these components are described following a measurement cycle.

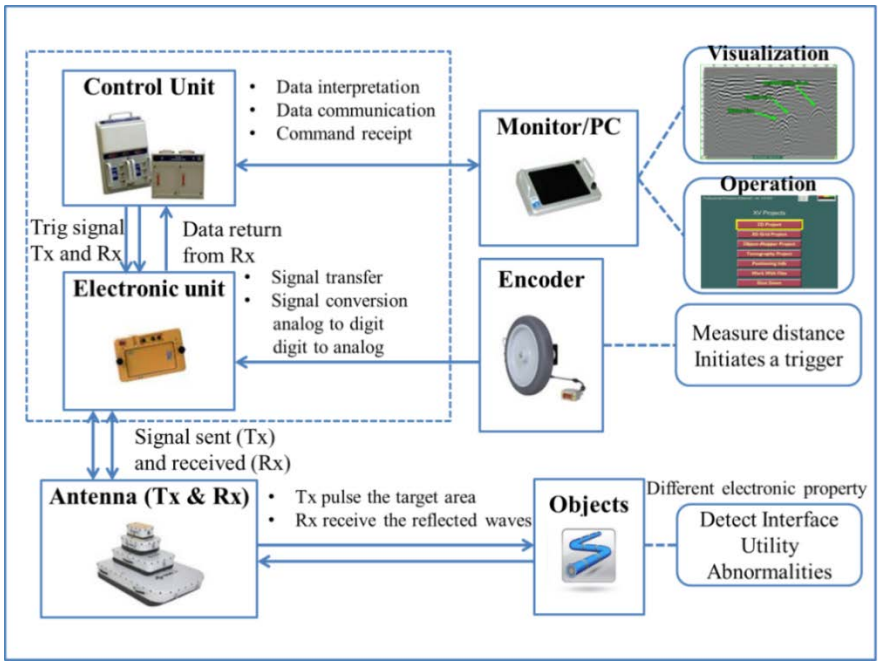


Figure 3.2 Components of GPR system

The starting point is encoder, generally a mechanical device such as a wheel that measures the distance of the survey track and initiates a trigger request for pulsing a radar signal at predetermined distance. Electronic unit, also referred to as AD converter, connects to encoder and receives trigger requests from it. Electronic unit serves as the interface between control unit and antenna by converting signals from analog to digit and visa-versa depending on the direction. The received trigger request is passed by electronic unit to control unit. On the receipt of the trigger request, a digital radar signal

is generated by control unit and then passed back to electronic unit for analog conversion. Resulting analog radar signal is passed to antenna.

Transmitter (Tx) antenna receives the analog radar signal to pulse the target area. Meanwhile, a signal is sent to receiver (Rx) antenna informing that a pulse has been sent and to expect reflected waves. Every radar pulse is digitally encoded for matching of transmitted and reflected signals. The reflected waves are captured by the receiver antenna and relayed to electronic unit for digitalizing. This digitized information is then passed to control unit where it is interpreted, buffered and relayed to monitor and/or data storage device (usually a PC) for visualization and analysis. The process described above is repeated as the GPR is moved across the surface of the target area for every measurement cycle.

With regard to objects detection, the process can be described with “5Rs”. The first R is radiation. The antenna radiates electromagnetic (EM) wave from the transmitter, which will propagates through the medium underneath the antenna. The second R is recognition. As the propagating wave encounters an object or interface with different electronic property, the wave recognize this change. The third R is reflection and refraction. When the wave hit the object or interface, a portion of the wave energy is reflected back to the receiver antenna and other potion of the wave is refracted through the medium. The fourth R is receipt of the wave and the fifth R is record of the interpreted signal. This “5Rs” briefly summarizes the procedure of detecting subsurface objects using GPR system.

3.1.2 GPR Principles

The principle behind the scene is the electromagnetic (EM) theory. It outlines the building blocks required to work quantitatively with GPR to detect and locate underground utilities. In this section, the electromagnetic properties of materials are introduced. Based on that, two EM mechanisms involved in the interaction between GPR and external objects, i.e. electromagnetic propagation through homogeneous medium and electromagnetic scattering through heterogeneous medium are described. At last, GPR radiation pattern is discussed.

1) Electromagnetic properties

The material properties that govern the behavior of electromagnetic wave in a medium are dielectric permittivity (ε), electrical conductivity (σ) and magnetic permeability (μ). Dielectric permittivity is measured in units of electrical capacitance (farads, F) per meter and is a representation of the material's ability to store electrical charge (Neal 2004). Dielectric permittivity is the main factor that limits the size of GPR footprints and determines the wave velocity in low-loss medium such as clean sand or gravel. Electrical conductivity characterizes the ability to transport charge on application of a static electric field. Magnetic permeability, measured in inductance (henrys, H) per meter, is essentially the magnetic equivalent of dielectric permittivity. It is a measure of magnetic field energy stored and lost through induced magnetization (Neal 2004).

The electrical conductivity and magnetic permeability are two factors affecting EM waves. Dielectric materials such as clean sands, gravels will allow a great amount of

radar energy pass through without dissipating it while electrical conductive material will attenuate the energy at a much shallower depth. Such electric conductive media include wet clay and those contain salt water, dissolvable minerals, and etc. Under very unfavorable condition, e.g. in a wet calcareous soil that contain certain clay-rich minerals, the maximum penetrating depth of GPR in the ground can be much less than a meter, no matter what frequency of the antenna is used (Conyers). With respect to magnetic permeability, the higher the magnetic permeability, the more electromagnetic energy will be attenuated during its transmission. Media that contain magnetite minerals, iron oxide cement or iron-rich soils can all have a high magnetic permeability and therefore transmit radar energy poorly (Conyers). Most often, the permittivity and permeability of materials are expressed as relative permittivity or dielectric constant ε_r and relative permeability μ_r, which are defined as Equation 3.1 and 3.2.

$$\varepsilon_r = \frac{\varepsilon}{\varepsilon_0} \tag{3.1}$$

$$\mu_r = \frac{\mu}{\mu_0} \tag{3.2}$$

$\varepsilon_0 = 8.854 \times 10^{-12} F/m$, $\mu_0 = 4\pi \times 10^{-7} H/m$ is the permittivity and permeability of free space (i.e. a region where there is no matter and no electromagnetic or gravitational fields) respectively. In most GPR applications, variations in ε and σ are most important while variations in μ are seldom of concern (Annan 2009, Loulizi 2001).

2) Electromagnetic scattering

Electromagnetic scattering occurs due to that EM waves encounter a discontinuity in the electromagnetic properties in the medium. In the case of GPR application for mapping underground utilities, the discontinuity could be either the interface between two layers with different dielectric constant, e.g. interface between pavement and soil or soil and utilities, or abnormalities within a layer, e.g. compacted soil surrounding the utilities. At the interface, the EM wave experience reflection, refraction or diffraction depending on the geometry of the discontinuity, the properties of the materials, the incoming angle and wavelength of the signal, and etc. The scattering will yield a reflected wave and a transmitted wave. If the reflected waves are captured by the receiver antenna, the discontinuity is detected.

The reflection strength is proportional to the magnitude of changes of electromagnetic properties (Van Dam 2001). Meanwhile, the transmitted wave will continue propagating in the medium. The amounts of reflected and transmitted energy, with respect to signal amplitude, are determined by the reflection coefficient and transmission coefficient. Considering a normal incidence of incoming wave, the reflection coefficient γ and transmission coefficient τ are determined as Equation 3.3 and 3.4 respectively, given that σ and μ_r contrasts are negligible (Lahouar 2003, Neal 2004).

$$\gamma = \frac{\sqrt{\varepsilon_{r,2}} - \sqrt{\varepsilon_{r,1}}}{\sqrt{\varepsilon_{r,1}} + \sqrt{\varepsilon_{r,2}}} \tag{3.3}$$

$$\tau = \frac{2\sqrt{\varepsilon_{r,1}}}{\sqrt{\varepsilon_{r,1}} + \sqrt{\varepsilon_{r,2}}} \tag{3.4}$$

Where $\varepsilon_{r,1}$ $\varepsilon_{r,2}$ are the relative dielectric permittivity of adjacent layers 1 and 2.

As for utility detection, layer 1 is the surrounding medium while layer 2 is the surface of the utility. Different utility materials result in different amounts of reflected energy, and thus different GPR profile. The larger the dielectric contrast between the utility and surrounding medium, the larger the reflection coefficient and subsequently, the delineation of utility are more evident. For instance, a Polyvinyl Chloride (PVC) pipe and a metallic pipe are buried in dry sands. The dielectric permittivity is 3 for dry sands, 3.4 for PVC and an infinite number for metal. With Equation 3.3, the calculated reflection coefficient for PVC pipe and metallic pipe are 0.03128 and 1 respectively. Hence, it is difficult to discern PVC pipe from GPR images because of the tiny amount of reflected energy. On the contrary, metallic pipe is much easier to be detected because almost full amount of energy are reflected by the interface. The electromagnetic scattering mechanism stated herein delineates how underground utilities are detected by GPR system.

3) Electromagnetic propagation

The behavior of an electromagnetic wave propagating through a homogeneous medium is governed by Maxwell's equations and constitutive relations. Maxwell's equations mathematically describe the physics of EM fields and constitutive relations relate the EM field to the material properties. The wave propagation properties, e.g. wave propagation velocity, energy attenuation constant could be derived from the Maxwell's equation and constitutive relations (Lahouar 2003). The velocity (v) of an electromagnetic wave is a

function of its frequency (f), the speed of light in free space, and the host medium's relative dielectric permittivity (ε_r), relative magnetic permeability (μ_r) and electric conductivity (σ). Mathematically it is defined as Equation 3.5 (Harris 2006).

$$v = \frac{c_0}{\sqrt{\varepsilon_r \mu_r \frac{1+\sqrt{1+(\sigma/\omega\varepsilon)^2}}{2}}} \tag{3.5}$$

Where c_0 is the electromagnetic wave velocity in free space ($3\times10^{-8} m/s$), ε is the permittivity of the material ($\varepsilon = \varepsilon_r \times \varepsilon_0$), and $\sigma/\omega\varepsilon$ is a loss factor where $\omega = 2\pi f$ is angular frequency (rad/s).

This relationship is the basis for locating subsurface objects using GPR (Harris 2006). In low-loss material where GPR is effective and most GPR surveys are conducted, the influence of conductivity σ upon the GPR frequency range is minimal and $\sigma/\omega\varepsilon$ is assumed to be 0 (Neal 2004). Furthermore, the influence of relative magnetic permeability μ_r is also assumed to be negligible, and a value corresponding to nonmagnetic material ($\mu_r = 1$) is given (Neal 2004). As a result, Equation 3.5 can be simplified to Equation 3.6

$$v = \frac{c_0}{\sqrt{\varepsilon_r}} \tag{3.6}$$

Therefore, the distance of a buried point detected by GPR in homogeneous medium can be calculated as Equation 3.7.

$$d = \frac{vt}{2} \tag{3.7}$$

Where d is the distance of the buried point to the GPR antenna; t is the two-way travel time (i.e. the time it takes for the wave travel from the transmitter to the reflector and back to the receiver) of the GPR wave propagating in the medium; which can be obtained from GPR output data; v is the wave propagation velocity in the homogeneous medium. Equation 3.7 is the basic equation for locating objects in the subsurface.

The amplitude (A) of the radar waves shows an exponential decline from its initial value (A_0) as the electromagnetic wave propagates through the medium, as Equation 3.8 shows (Neal 2004).

$$A = A_0 e^{-\alpha z} \tag{3.8}$$

Where α is the attenuation constant; and z is the traveled distance. According to Lahouar (2003), α is calculated as Equation 3.9. For a low-loss material, Equation 3.9 is simplified to Equation 3.10. The notation of Equation 3.9 and 3.10 is the same as Equation 3.5.

$$\alpha = \omega\sqrt{\mu\varepsilon}\sqrt{\frac{\sqrt{1+(\sigma/\omega\varepsilon)^2}-1}{2}} \tag{3.9}$$

$$\alpha = \frac{\sigma}{2}\sqrt{\frac{\mu}{\varepsilon}} \tag{3.10}$$

As indicated in Equation 3.10, conductivity exerts the greatest impact on the attenuation constant, which proves that there is a high conduction-based energy loss in conductive medium. In order to ensure or improve the quality of GPR images, an adequate attenuation correction is needed (Neto and Medeiros 2006).

4) Radiation Pattern

It is worth to examine the radiation pattern of GPR and estimate the effective detection range. It is also referred to as Fresnel zone or antenna footprint. The footprint is essential for determining transect spacing within a survey grid to detect all the subsurface features of interests. Instead of a pencil-like beam, the radar energy radiated from the standard commercial GPR is in an elliptical cone with the apex at the center of the transmitter antenna (Yalciner 2009). The approximate shape of the footprint is shown in Figure 3.3 and the approximate size is computed as Equation 3.11 (Yalciner, 2009). The receiving pattern of the antenna is exactly the same as the transmitting pattern with the same degree of directionality for given ground conditions (Roberts and Daniels, 1996).

Generally, the angle of the cone is a function of the relative dielectric permittivity of the medium through which the wave propagates, and the frequency of the radar energy emitted from the antenna. More specifically as Equation 3.11 shows, the lower the antenna frequency, the longer the major semi-axis and thus the broader the transmission cone. In addition, higher relative dielectric permittivity leads to lower velocity of the radar wave and more focused conical transmission pattern.

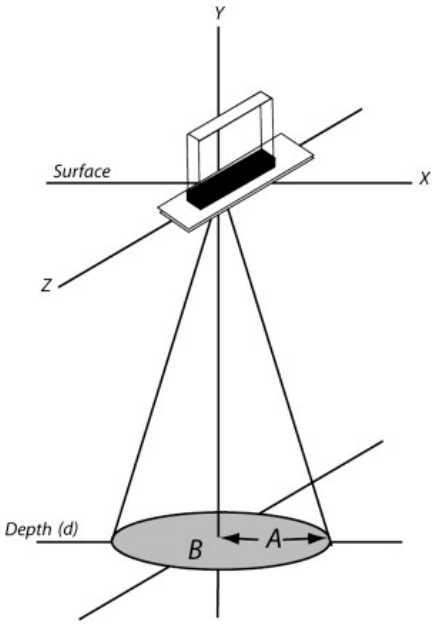


Figure 3.3 Approximate size of the footprint (Yalciner 2009)

$$A = \frac{\lambda}{4} + \frac{d}{\sqrt{(\varepsilon + 1)}}$$
$$B = \frac{A}{2} \tag{3.11}$$

Where A is the major semi-axis and B is the minor semi-axis, λ is the center frequency wavelength of radar energy, d is the depth from ground surface to reflection surface, ε is the average relative dielectric permittivity of material from ground surface to the depth.

In order to get the deepest penetrating depth, the antenna must be pointing normal to the ground plane and located very close to the ground. The maximum penetrating depth is controlled by the antenna frequency (Herman 1997). Generally, the high frequency antenna has a high resolution with a shallow penetrating depth while the low frequency antenna has a low resolution and a deep penetrating depth. For instance, the approximate

maximum penetrating depth is 6 meters for 500 MHz antenna and 2.5 meters for that of 800 MHz antenna. There is always a trade-off between the resolution and penetrating depth such that one needs to select the correct operating frequency for the depth necessary and the resolution desired in a specific task.

3.1.3 GPR Data Format

Figure 3.4 illustrates GPR data formats and gives each an example. The four GPR data formats, i.e. samples, A-scan, B-scan and C-scan are introduced in an evolutionary sequence.

1) Sample

The basic GPR data is samples, i.e. the instant digital values (electric field intensity (amplitude)) of the recorded radar signal at specific times. Three parameters in sampling the signal are amplitude, sampling interval and sample number. The amplitude is a function of transmitted signal strength, coupling between GPR antenna and ground, the signal travel path, and electromagnetic attenuation of subsurface materials and interfaces encountered by the signal (Plati and Loizos, 2013). Samples are recoded every sampling interval, i.e. the time between points for each recorded waveform (Harris 2006).

The maximum sampling interval that is suited for survey can be calculated using Equation 3.12 (Dojack 2012).

$$t = \frac{1000}{6f} \tag{3.12}$$

Where t is the maximum sampling interval (in ns) and f is the center frequency of the antenna (in MHz). Sample number is the order in which the samples were measured, recorded and stored. Sample numbers serves as the reference to the sequential positions of samples in A-scan, which will be introduced next.

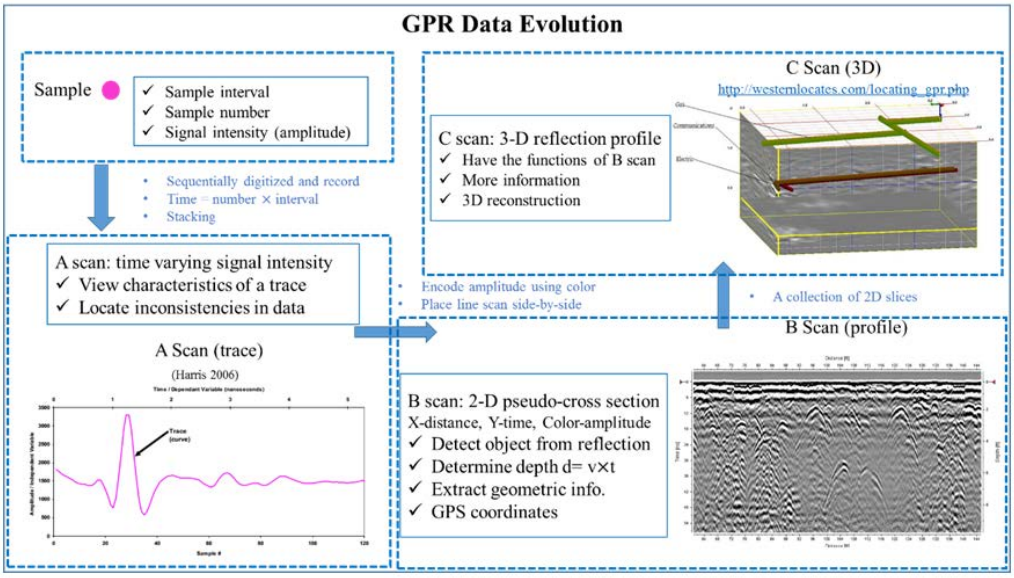


Figure 3.4 GPR data formats

2) A-scan

The basic output of GPR is A-scan, i.e. a 1-D time varying signal intensity (amplitude). A-scan is built up by placing a certain number of samples sequentially. Figure 3.5 is an example of A-scan. The vertical axis represents the normalized amplitude and the horizontal axis represents the two-way time of reflected signal. The two-way time is the multiplication of sample number and sampling interval. At the beginning of the signal, there is a transmitted pulse resulted from coupling between transmitter antenna and receiver antenna. A series of reflected signal is following the transmitted pulse.

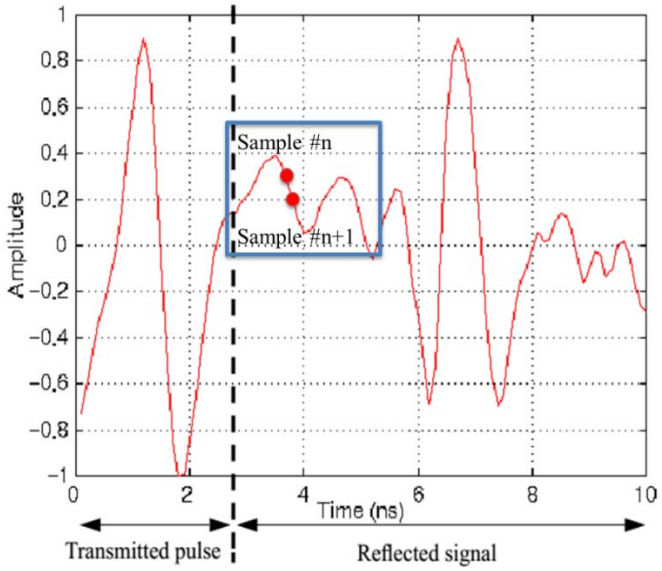


Figure 3.5 Example of A-scan (Herman 1997)

A-scan illustrates how the signal is reflected back to the antenna with arrival times. The arrival times vary according to the location of the discontinuity within the medium. Figure 3.6 shows a series of GPR A-scan of a metallic plate at various distances from the antenna. As the distance between the antenna and the plate increases, the propagation time increase and thus the reflected signals are recorded at a later time (Herman 1997). A-scan is often used to view the characteristics of an individual trace, or to locate inconsistencies in the data (Harris 2006).

3) B-scan (profile)

B-scan is obtained as GPR antenna moves along a line. The generation of B-scan is a two-step process. A series of A-scans are triggered at a predefined time or distance interval when the antenna is pulled along the survey path. Firstly, the amplitudes of the A

scans are encoded using intensity or color to convert it to line scans. Then these color encoded strips are stacked sequentially side by side to form a B-scan (see Figure 3.7).

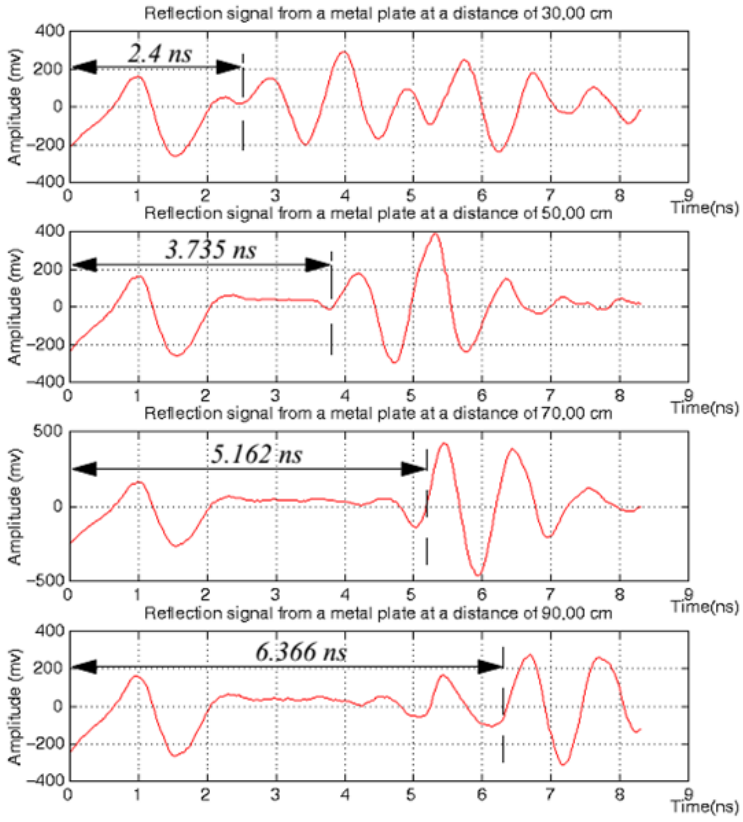


Figure 3.6 A-scans of a metallic plate at various distances (Herman 1997)

Figure 3.8 is an example of B-scan. The X axis of B-scan represents the moved distance of GPR antenna along the survey path, Y axis represents the two-way travel time of the radar wave, and the color is added as a third dimension to depict the amplitudes. B-scan is the main GPR data format used for interpretation in most GPR application scenarios. Not only the buried utilities can be detected, but also the depth and even geometric

information such as pipe radius can be obtained through appropriate image analysis techniques.

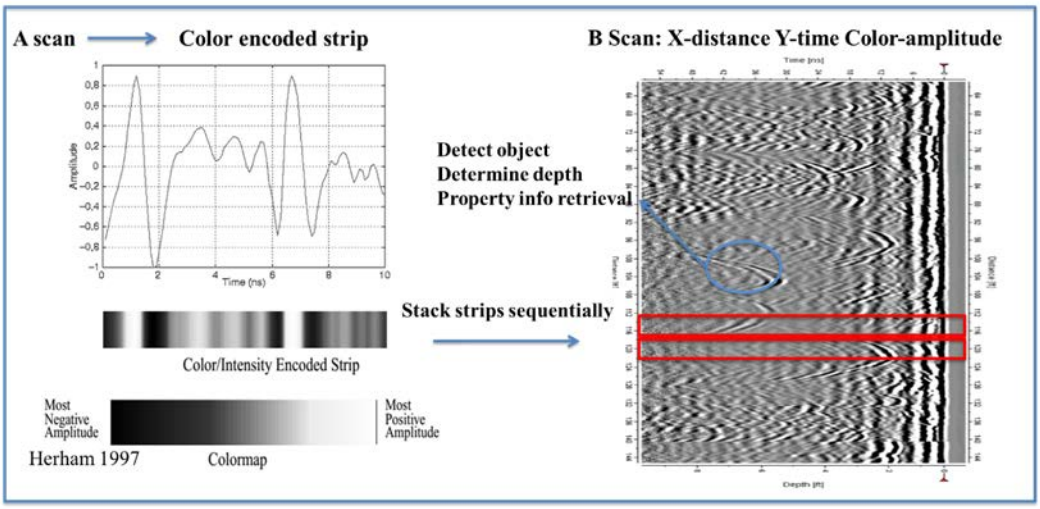


Figure 3.7 Formation of B-scan by a collection of A-scans

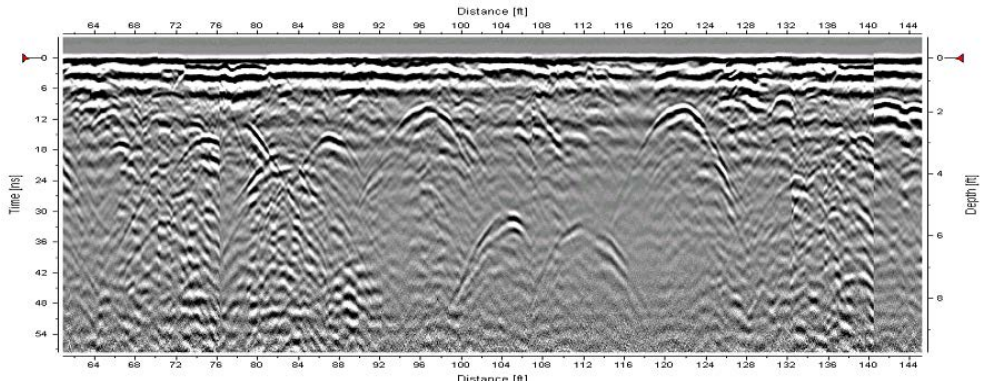


Figure 3.8 Example of B-scan

4) C-scan (3D data)

The 3D GPR data, often known as C-scan, is obtained by surveying along a series of parallel lines. The B-scans acquired in each line are stacked slice by slice to form the C-scan. Figure 3.9 shows how a collection of B-scans can form a C-scan. C-scan inherits all

the features and capabilities of B-scan, and furthermore it contains more information such as orientation of the objects, the relative position of several objects in the subsurface, and etc. However it also requires much more efforts to collect data and is computational expensive to process the collected data.

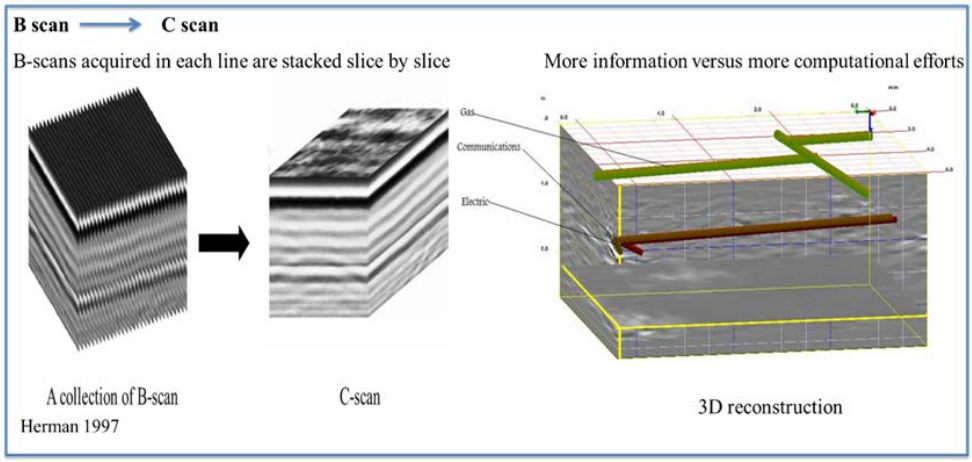


Figure 3.9 Formation of C-scan by a collection of B-scans

3.1.4 Reflection Patterns of Buried Utility in B-scan

The EM wave reflected by subsurface utilities result in a wide range of signal patterns that depend on the sizes, relative orientations of the antenna and utilities, materials of the utilities, and characteristics of subsurface features. This section first illustrates the reflection patterns of buried point and line in a B-scan. Following that, the reflection patterns of cylindrical objects are formulated.

1) Reflection of Buried point

The reflection pattern of a buried point in a B-scan is a hyperbola. The hyperbola generated because of two reasons. The first is that the conical radar energy projected into the ground allows GPR to detect the buried point in an oblique direction. Additionally in the B-scan, the two-way time is plotted directly below where the antenna measured the signal. Hence as the antenna moves toward and then away from the buried point, the reflections form a hyperbola (see Figure 3.10). It is worthy to note that only the apex of the hyperbola denotes the actual location of the buried point.

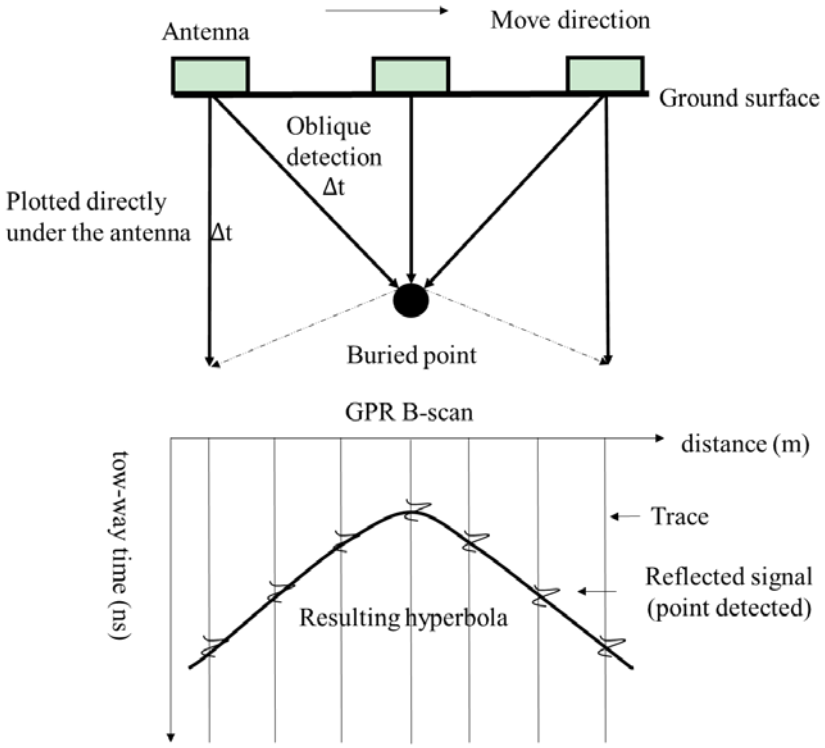


Figure 3.10 Hyperbola generated by point reflector

2) Reflection of buried line

In order to derive the reflection patterns of buried lines in GPR B-scan, the line is modeled as a succession of points. Generally, buried lines will generate different reflections as the GPR surveys in different directions. Two directions, i.e. parallel and perpendicular to the line are considered with regard to the reflection profile.

Figure 3.11 shows the reflection pattern of a buried line when GPR antenna moves perpendicularly to it. As shown in Figure 3.11, a plane passing through the center of GPR and perpendicular to the buried line intersects that line at a point. Obviously, this point is always the first one to be detected and will generate a hyperbola in the B-scan. Other points distributing along the two sides of that point will be detected at a later time and form a series of hyperbolae enveloped beneath the hyperbola formed by the point shown in Figure 3.11. As the points are placed closely to form a line, the reflection pattern in B-scan shows a hyperbolic reflection.

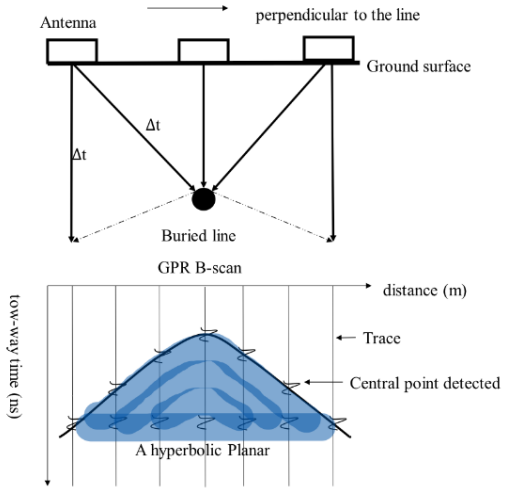


Figure 3.11 Hyperbolic reflection generated by buried line

The reflection pattern of a buried line in B-scan is linear strip when the GPR antenna moves parallel to the line (see figure 3.12). The point (i.e. the intersection of the buried line and the vertical wave emitted from transmitter) will be detected first, and a reflected pulse will be generated in B-scan. While the points detected by the oblique radar waves will cause the amplitudes in the scan vary at a later time. Therefore, the reflection pattern of a buried line is a collection of closely stacked lines or a strip, since the line is modeled as a succession of points.

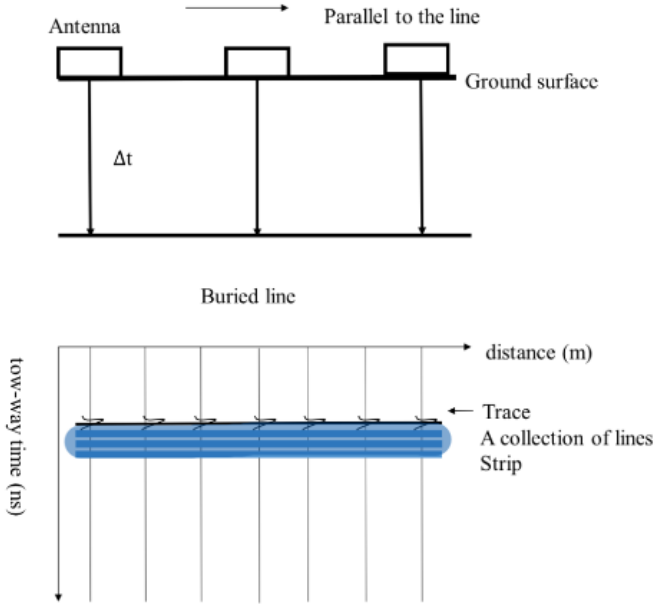


Figure 3.12 Linear strip generated by buried line

An interesting problem arises when GPR antenna moves along the buried line but not parallel to it (see figure 3.13). Figure 3.13 shows a tilted buried line and its reflection profile. It is important to note that the apparent angle β in the B-scan is different from

the actual angle α . The relationship between α and β is shown in Equation 3.13 (Herman 1997).

$$\sin\alpha = \tan\beta \tag{3.13}$$

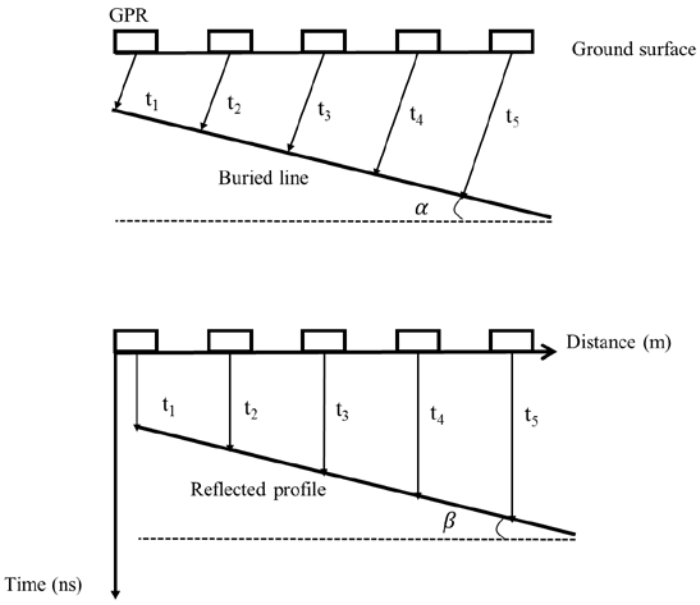


Figure 3.13 A tilted buried line and its reflection profile

3) Reflection of cylindrical utilities

To illustrate the reflection patterns of a cylindrical object in B-scan, an example in an ideal situation is cited. Figure 3.14 shows a nonmetallic pipe is buried in a homogeneous dielectric medium, and GPR antenna moves perpendicularly to it. The GPR sampling interval is assumed to be infinitesimal meaning all the waves reflected by the cylinder and received by the antenna can be continuously measured and stored. The outside upper half of the cylinder is interface 1 and the inside lower half of the cylinder is interface 2.

Tow definition, i.e. direct wave and key point are given for illustrating the cylinder reflection pattern. The direct wave is defined as the part of energy that travels the shortest distance between transmitter and receiver. In Figure 3.14, the slice is perpendicular to the pipe and passes through the center of GPR antenna. Within this plane, the footprint of GPR is a fan and the cross section of the pipe is a circle. The key point is defined as the intersections of the circle and the line that connects the centers of GPR antenna and the circle. Therefore apparently, the path of direct wave reflected by the interface is the line that connects the key point on that interface and the center of GPR antenna.

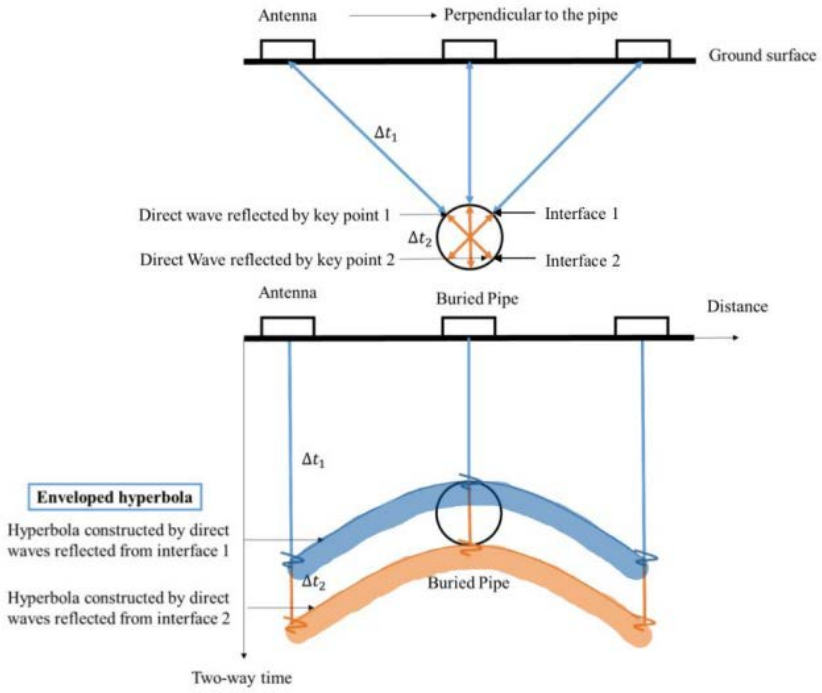


Figure 3.14 Hyperbolic strips generated by buried cylindrical utility

The reflected direct waves will generate two enveloped hyperbolae in the reflection profile (see Figure 3.14). The two hyperbolae are “constructed” by a number of key

points on the interfaces as GPR antenna move towards and away from the pipe. Therefore, for a nonmetallic pipe, the reflection pattern in B-scan is two hyperbolic strips when GPR antenna moves perpendicularly to it (see Figure 3.14). The enveloped hyperbolae are essential to determine the depth and geometric information of buried utilities, which will be discussed in Chapter 4. Similarly, it is straightforward that reflection pattern will be two linear strips when the GPR antenna moves parallel to the pipe (see Figure 3.15).

There are several inconsistences when the data is acquired in real world conditions. The first inconsistence is that the GPR misses recording the reflected direct waves because they arrive right in the sampling interval. In other words, the enveloped hyperbola is not really constructed by the key points. However, this may not be a big concern since the sampling interval is very short and results in very small spacing. For example, a 0.02 ns sampling interval (which is typical in utility locating) translates into a spacing of 0.3 cm assuming a free space (15 cm/ns round trip propagation velocity). In sand, the same sampling interval translates into a 0.15 cm spacing (7.5 cm/ns round trip propagation velocity). Therefore, the small spacing away from the key points can be ignored and the enveloped hyperbolae can be considered to be constructed by the key points. The second inconsistency is the noise generated by the heterogeneity of subsurface, refraction and diffraction of cylindrical interface. The negative impacts of the noise are inevitable and will generate locating and characterizing errors during the measurements.

In addition, the reflection pattern varies according to different pipe materials, products conveyed in the pipe, surrounding medium and pipe radius. For instance, if the buried

pipe in Figure 3.14 and Figure 3.15 is metallic, the hyperbolic and linear strip generated by interface 2 will not appear in the radargram. This is because all the radar energy is reflected by the metallic interface 1, which "shadows" interface 2 and makes it undetected. Pipes that contain water may show duplicate hyperbolae as the radar waves echo from the top of the pipe, the water in the pipe and the bottom of the pipe. If the pipe is buried in a trench with compacted walls, the radar wave reflected from the trench walls may form an X shape above the hyperbola. More importantly, the radius of the pipe also affects the shape of the hyperbola, which will be discussed in detail in Chapter 4.

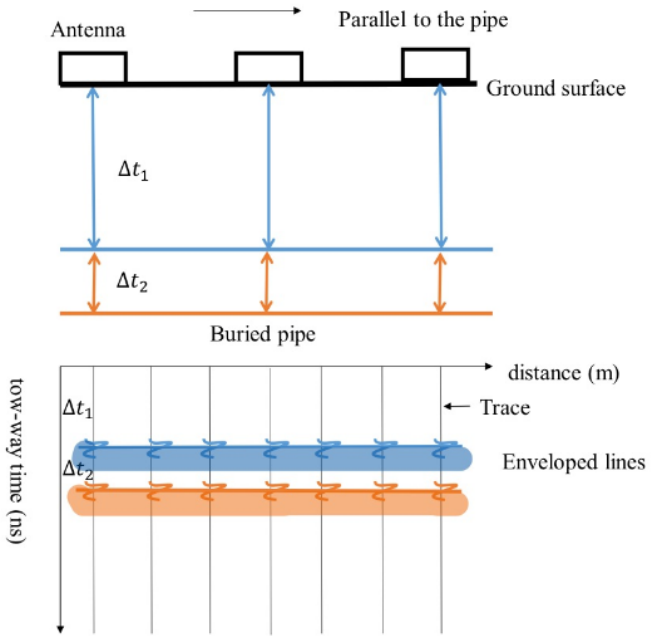


Figure 3.15 Linear strips generated by buried cylindrical utility

3.2 GPS Fundamentals

The global positioning system (GPS) is a U.S.-owned system that provides users with positioning, navigation and timing (PNT) services. It can locate a point by tracking signal from satellites. The GPS fundamentals are introduced as follows.

3.2.1 GPS Elements

GPS consists of three parts: space segment, control segment and user segment (see Figure 3.16). U.S. National Coordination Office gives an overview of GPS system (USNCO 2013).

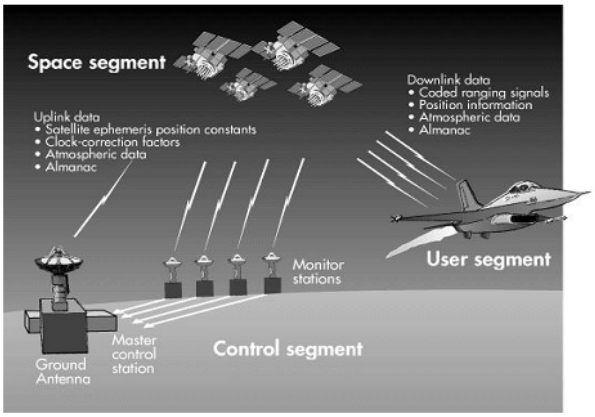


Figure 3.16 Three major segments of GPS (Cai 2003)

1) Space Segment

GPS space segment consists of a constellation of satellites transmitting radio signals to control segments and users. The constellation is managed by Air Force to ensure at least 24 GPS satellites are available for 95% of the time. The GPS satellites fly in medium Earth orbit at an altitude of approximately 20,200 km and circle the Earth twice a day.

The satellites are arrayed into six equally-spaced orbital planes surrounding the Earth so that each orbit contains four “slots” occupied by baseline satellites (see Figure 3.17). This 24-slot configuration ensures at least four satellites are available in view from virtually any point on the Earth. In June 2011, a GPS constellation expansion, known as “Expandable 24” configuration, is completed. As a result, GPS now operates as a 27-slot constellation with improved coverage in most parts of the world.

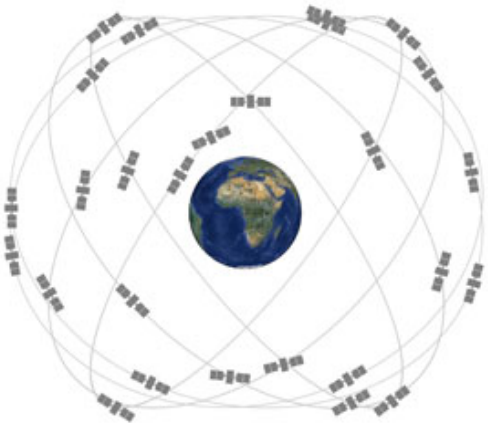


Figure 3.17 Expandable 24-slot satellites constellation
(http://www.gps.gov/systems/gps/space/)

2) Control Segment

GPS control segment is a global network of ground facilities that track and monitor GPS satellites, perform analysis, send commands and upload data to the constellation. The current operational control segment consists of a master control station, an alternate master control station, 12 command and control antennas, and 16 monitoring sites, which are distributed around the world (see Figure 3.18).

The master control station (MCS) in Colorado performs the primary function of control segment. Based on the collected navigation information from the monitor stations, the MCS computes the precise locations of the GPS satellites in space and upload this data to the satellites. The MCS also undertakes the maintenance of constellation health and its accuracy. In the case of a satellite failure, the MCS can reposition satellites to maintain an optimal GPS constellation. The monitor stations track the GPS satellites and collect atmospheric data, range/carrier measurements, and navigation signals. The information collected by monitor stations is communicated to the MCS. Ground antennas are used to communicate with the GPS satellites for command transmission and control purposes. In addition, the control segment is connected to the 8 remote tracking stations in Air Force Satellite Control Network (AFSCN) worldwide, which increases visibility, flexibility and robustness for telemetry, tracking, and command.

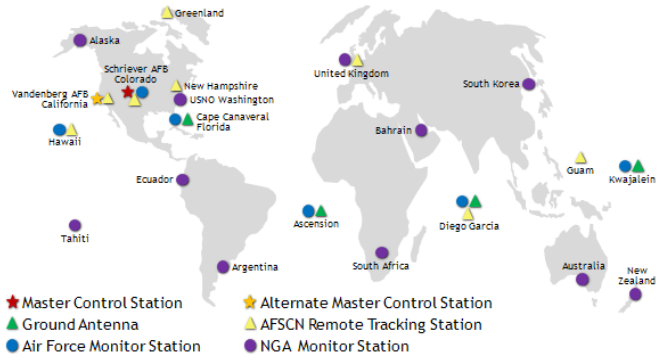


Figure 3.18 Control segment

(http://www.gps.gov/systems/gps/control/)

3) User Segment

GPS user segment consists of the GPS receiver equipment, which receives the signal from the GPS satellites and uses the transmitted information to calculate the user's three dimensional position and time. The principle of how the position is determined will be described below.

3.2.2 GPS Principles

Although GPS technologies have experienced rapid developments, the basic principle remains the same. The distance from a position on Earth to a satellite can be determined via satellite ranging. The satellites send radio signals to reach a specific position on the Earth surface with light speed (186,000 miles per second). The traveling time is measured and times the traveling speed to derive the distance. The resulting distance is referred to as "pseudorange" due to some inherent errors in the time measurement $\Delta t_{measured}$ (Cai 2003). A GPS receivers need to receive signals from four different satellites to enable it to calculate signal transit time Δt_1, Δt_2, Δt_3, Δt_4 and thus the range of the user to the four satellites R_1, R_2, R_3, R4 (see Figure 3.19). As the spatial positions of the four satellites are known, the range can be determined in a Cartesian coordinate system as Equation 3.14.

$$\begin{aligned} R_i &= c \times \Delta t_i \\ R_i &= \sqrt{(\mathrm{X} - \mathrm{X}_{sat,i})^2 + (\mathrm{Y} - Y_{sat,i})^2 + (\mathrm{Z} - Z_{sat,i})^2} \end{aligned} \tag{3.14}$$

Where R_i is the range of the user to the i^{th} satellites, c is the light speed, Δt_i is the transit time of signal from i^{th} satellite, (X, Y, Z) is the coordinates of the user and ($X_{sat,i}$, $Y_{sat,i}$, $Z_{sat,i}$) is the coordinates of i^{th} satellites.

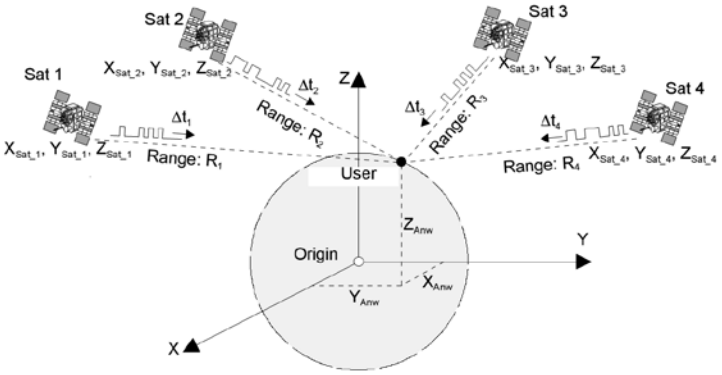


Figure 3.19 Position determination form four satellites (Zogg 2002)

However, the time measured $\Delta t_{measured}$ by the GPS receiver is different from the signal transit time Δt_i. The time when the satellite signal is transmitted is known very precisely because of the atomic clocks on the satellites. In addition, all satellite clocks are adjusted or synchronized with each other and Coordinated Universal Time (UTC). But the receiver clock is not synchronized to UTC and thus slow or fast by Δt_0. Δt_0 is positive when the user clock is fast. Therefore the measured time $\Delta t_{measured}$ will generate an inaccurate/incorrect "pseudorange" PSR (see Equation 3.15).

$$\begin{aligned} \Delta t_{measured,i} &= \Delta t_i + \Delta t_0 \\ PSR_i &= c \times \Delta t_{measured,i} = c \times \Delta t_i + c \times \Delta t_0 \end{aligned} \tag{3.15}$$

Where $\Delta t_{measured,i}$ is the measured signal transit time from i^{th} satellite, PSR_i is the pseudorange of user to i^{th} satellite, Δt_0 is the time error of GPS receiver. With Equation 3.14 and 3.15, the position of user is determined as Equation 3.16.

$$\sqrt{(\mathrm{X}-\mathrm{X}_{sat,i})^2 + (\mathrm{Y}-Y_{sat,i})^2 + (\mathrm{Z}-Z_{sat,i})^2} + c \times \Delta t_0 = c \times \Delta t_{measured,i} \tag{3.16}$$

If four satellites are tracked simultaneously by the GPS receiver, the system of four equations can be solved and thus the position (X, Y, Z) and Δt_0 can be derived.

There are different types of GPS with different accuracy. The static GPS only uses satellites to determine the user's position. It is not suitable for utility survey due to large localization error (possible horizontal error is 10.2 m and vertical error is 12.8m (Zogg 2002)) and long measuring time. The differential GPS (DGPS) is applied to greatly improve the locating accuracy. In principle, a reference receiver located at an accurately measured reference point (i.e. the exact coordinates are known) is used in addition to the user receiver. Many errors can be eliminated through continually comparing the user receiver with the reference receiver (Zogg 2002). Generally, the compensation of error can be described as a three-step process. First is to

1. Determine the correction values at the reference station.
2. Relay the correction values from the reference station to the GPS user.
3. Correct the position measured by the GPS user.

There are two different types of Differential GPS (see Figure 3.20 and Figure 3.21):

1. Differential GPS (DGPS) is based on the measurement of signal transit time and the achievable accuracy is approximately 1m.
2. Real Time Kinematic GPS (RTK GPS) is based on phase measurement of the carrier signal and the achievable accuracy is approximately 2cm.

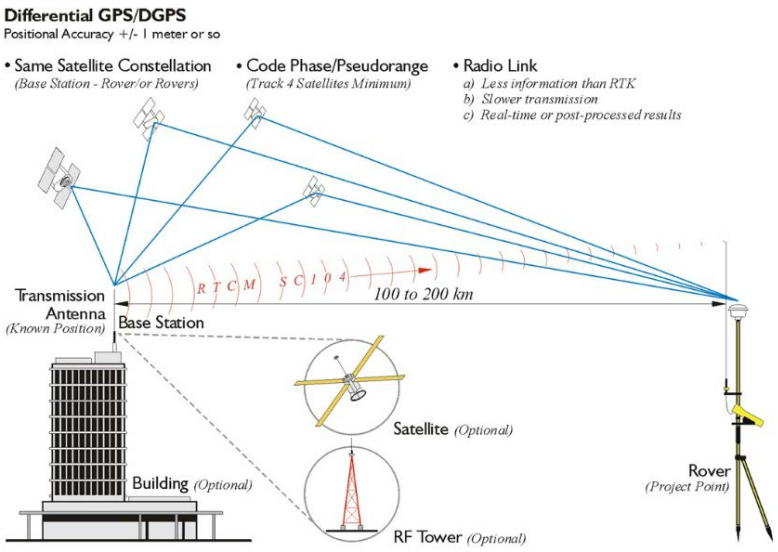


Figure 3.20 Differential GPS (Sickle 2009)

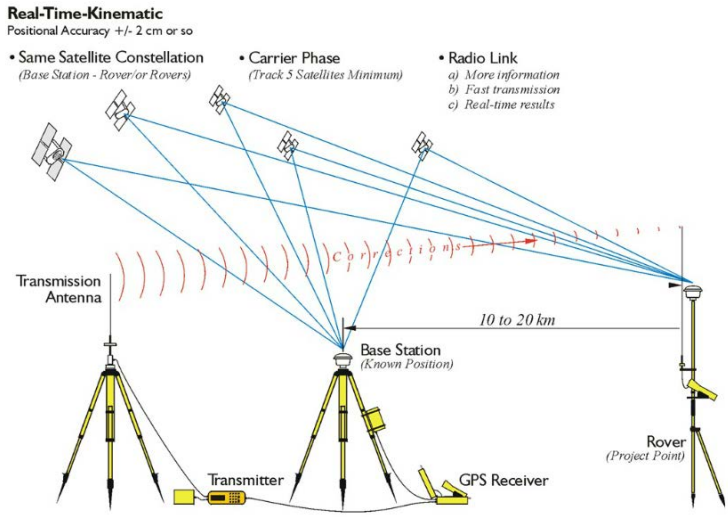


Figure 3.21 RTK GPS (Sickle 2009)

Instead of measuring the transit time of satellite signal, RTK GPS evaluates the satellite signal carrier phase to calculate the distance from user to satellite. The carrier wavelength

λ is approximately 19 cm. The range to a satellite can be determined using Equation (3.17) (see Figure 3.22).

$$D = (\mathrm{N} \times \lambda) + (\varphi \times \lambda) \tag{3.17}$$

Where D is the distance between satellite and user, N is the number of complete cycles, λ is the wave length and φ is the phase.

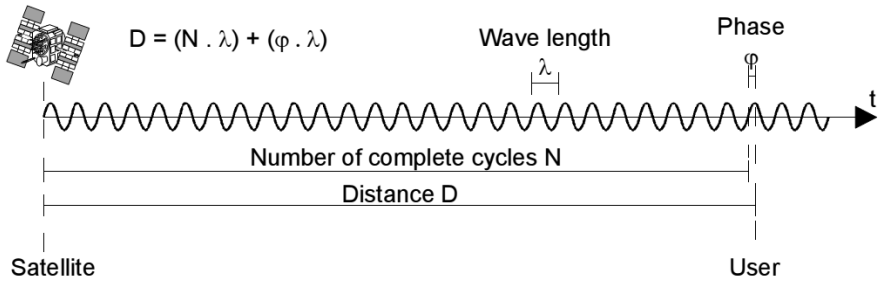


Figure 3.22 Principle of phase measurement (Zogg 2002)

By tracking several satellites at different times and through comparison between the user receiver and reference receiver, the user's position can be determined after solving a number of equations. The accuracy is within a few millimeters. As authorities such as ASCE recommended, the accuracy of utility mapping need to be within ± 100 mm. Hence, in order to achieve that accuracy, RTK GPS is adopted in this research to work with GPR for utility locating.

3.3 GIS Fundamentals

A geographic information system (GIS) integrates hardware, software and data for capturing, manaing, analyzing and displaying all forms of geographically referenced

information. GIS enable us to view, understand, question, interpret, and visualize data in many ways that reveal relationship, patterns, and trends in the form of maps, globes, reports and charts (ESRI). A brief introduction of GIS with regard to utility data management is presented in section 3.3.1. Section 3.3.2 introduces the mechanism of multi-patch for 3D object modeling in GIS.

3.3.1 GIS for Utility Data Management

The collected utility data, i.e. the utilities location and their attributes including radius, materials, ownship, is transferred to GIS for data management. GIS is a database system designed to work with the spatially referenced data. GIS makes use of two types of databases. One is a table including geographic data that is contained in a shape file. The other type of database is linked with the shape file and contains attribute information, which is imported from data soures and/or generated in the GIS application (Poku and Arditi 2006). The concepts are described in detail with an example of sewer utility data management in GIS (the data is from city of Bellingham, washiongton, U.S., http://www.cob.org/services/maps/gis/sewer.aspx).

The shape file constructs maps in a view with points, lines and polygons, which are known as features. A point feature is a GIS object that stores its geographica representation (i.e. x and y coordiante pair) in the database. Some point features such as utility apexes, need to also include a z-value or height, to correctly locate itself in 3D space. These objects, modeled as 3D point features, embed their z coordiantes inside the geometry or shape file to enable it to represent any 3D position in space. Point features

can be utilized to model manholes, discharge points and valves in utility management. Figure 3.24 shows the manhole locations of Bellingham in Arc GIS.

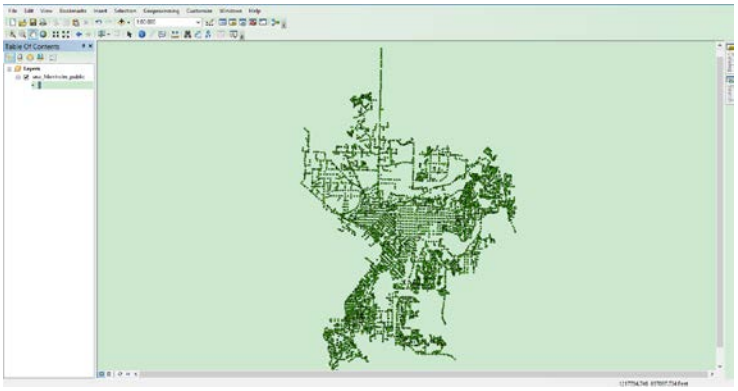


Figure 3.24 Point features for manhole modeling in Arc GIS

A Line feature have several locations (a series of x and y coordiante pairs) strung out along the line in sequence. Some lines need to also include z-values, or heights to correctly locate themsevlves in 3D space. The z-aware polyline allows it to connect any two 3D points together. Figure 3.25 shows a 2D polyline employed to model the sewer mains in Bellingham in Arc GIS.

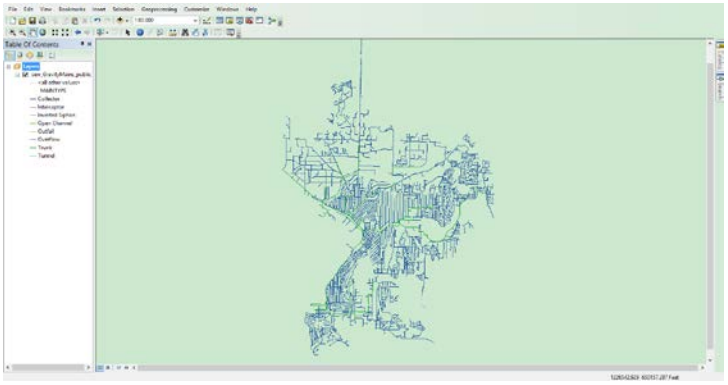


Figure 3.25 Line features for sewer modeling in Arc GIS

Polygon features consist of one or more lines that form a loop, i.e. a series of x and y coordiante pairs that enclose an aera, an example being the sewer basin (see figure 3.26).

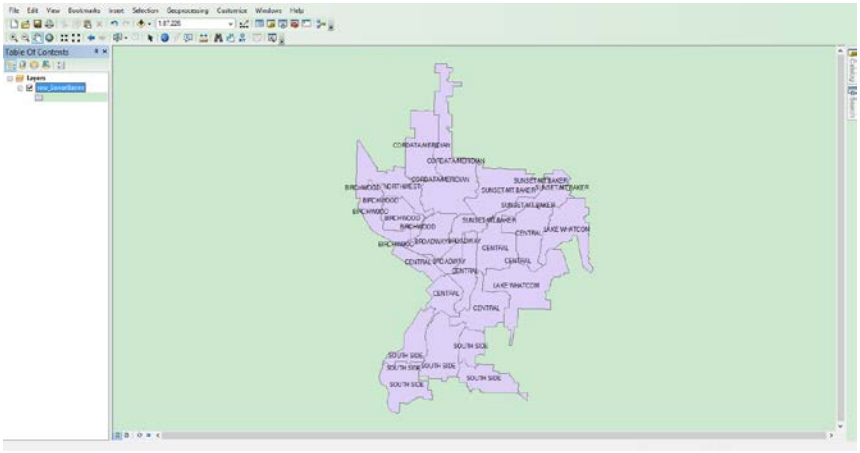


Figure 3.26 Polygon Features for sewer basin modeling in Arc GIS

The other type of database file is an attribute table storing non-geospatial information. Figure 3.27 shows an example of the attribute table as well as the linkage with the shape file. The attributes of the highlighted features are also been highlighted in the attribute table. The attributes table includes information such as the diameter, material and ownership of the utilities, and data source and accuracy of manhole locations. Such attribute information is critical to support decision-making throughout utility life cycle management.

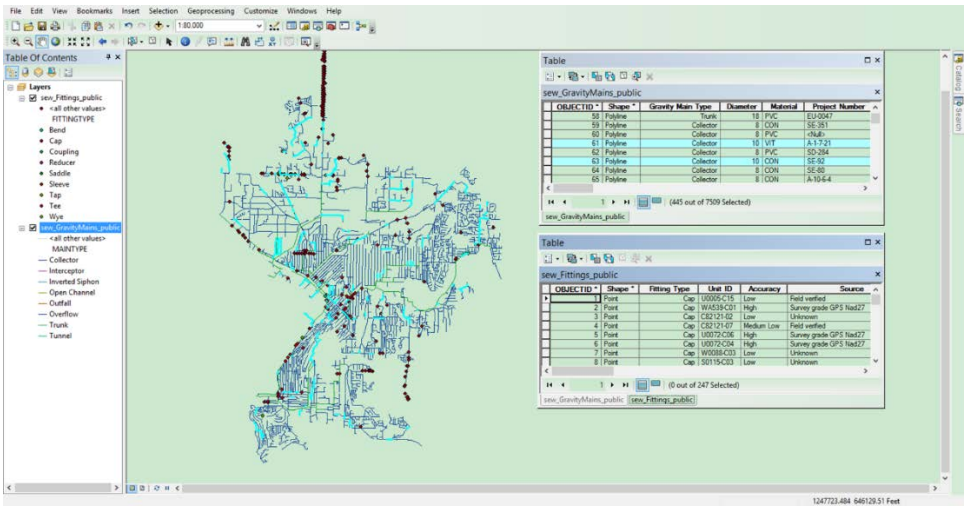


Figure 3.27 Attribute table of features in Arc GIS.

The current practices showed herein are predominantly being 2D and deterministic. To modeling the underground utilities and the associated positional uncertainty in 3D, multi-patch method for 3D object modeling in GIS is introduced next.

3.3.2 A Multi-patch Method for 3D Modeling in GIS

Multi-patch data format, a geographic information system (GIS) industry standard developed by ESRI in 1997, is a geometry used to represent the boundary of 3D objects (ESRI 2008). A multi-patch feature is a GIS object that can be made up of triangle strips, triangle fans, triangles, or rings. The texture, color, transparency and geometric information can be stored in the multi-patch, which results in an ideal data type for 3D feature representation. This data type is mainly used in Arc GIS and supported by many other commercial or open-source 3D software packages (e.g. SketchUp).

A multi-patch can be viewed as a container for a collection of geometries that represent 3D surfaces. As illustrated in Figure 3.28, these sub-geometries can be triangle strips, triangle fans, triangles, or groups of rings. A single multi-patch may consists one or a combination of these geometries.

Triangle strip is a continuous linked strip of 3D triangles where every vertex after the first two completes a new triangle. As shown in Figure 3.29 (a), a new triangle is formed by connecting the new vertex with its two immediate predecessors: (0, 1, 2), (2, 1, 3), (2, 3, 4), (4, 3, 5). The triangle fan geometries are formed in a similar way. It is a continuous fan of 3D triangles where the first point defines the apex or origin that all triangles share as a common pivot point and is included in all triangle surfaces. Every vertex after the first two completes a new triangle, and a new triangle is always formed by connecting the new vertex to its immediate predecessor and the first vertex of the part. For a triangle fan with six points, the triangle surfaces are defined by points: (0, 1, 2), (0, 2, 3), (0, 3, 4), (0, 4, 5) (see Figure 3.29 (b)). A collection of 3D triangles where each consecutive triplet of vertices define a new triangle. The size of a triangles part must be a multiple of three. For a triangles part with six points, the triangle surfaces are defined by points: (0, 1, 2), (3, 4, 5) (see Figure 3.29 (c)). A ring is a geometric element from which polygons are constructed, defined by an area bounded by one closed sequence of connected segments. Figure 3.29 (d) illustrates several geometries formed by a combination of rings. With multi-patch, the underground utilities and associated positional error can be modeled and rendered in GIS. The technical details will be discussed in chapter 5.

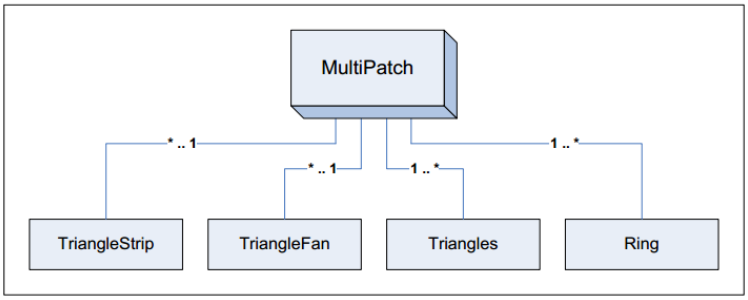


Figure 3.28 multi-patch sub-geometry type (ESRI 2008)

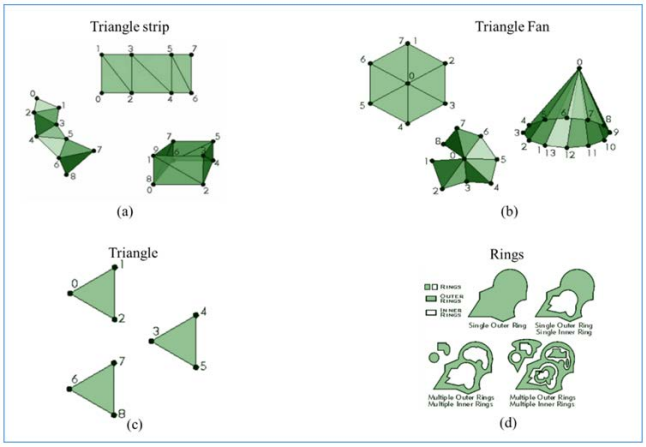


Figure 3.29 Examples of Multi-patch components (ESRI 2008)

CHAPTER 4. INTERPRETATION OF GPR AND GPS RAW DATA

The proposed method consists of three major phases. Phase 1 is GPR and GPS data interpretation, which includes GPR raw data preprocessing, extraction and modeling. A novel VIP algorithm is created to retrieve location, radius and spatial orientation of underground utilities from GPR and GPS raw data. A number of field experiments were carried out to validate VIP algorithm. The technical details are described in this chapter. Phase 2 is locational error assessment, which evaluates the error magnitudes of GPR in different scenarios. Quantitative linkages between error magnitudes and its influencing factors, i.e. buried depth and soil conditions are established. Phase 3 takes the error magnitudes as inputs to model the positional uncertainty/error of underground utilities in 3D space. Phase 3 aims to extend the current 2D utility data to 3D and paves the way to an error-aware GIS practice. The implementation of phase 2 and phase 3 will be introduced in chapter 5.

4.1 Preprocessing and Extraction of GPR Raw Data

GPR raw data contains information including antenna position x (m), two-way time t (ns), corresponding GPS coordinates and signal amplitudes. The original GPR B scans are required to be processed for visual quality enhancement and raw data extraction. In this section, REFLEXW 7.0, a successful program package for GPR data analysis, is utilized

to preprocess and extract GPR raw data. The general processing sequence including subtract-mean (dewow), time-zero correction, gain and background removal is described below.

1) Dewow

Dewow, referred to correction of low-frequency and DC bias in data, reduces the data to a mean zero level (Cassidy 2009). It is a critical step as it allows positive-negative color filling to be used in the recorded trace to reveal important information. Figure 4.1 illustrates the concept of dewow filter correction. Figure 4.2 gives an example of applying dewow to a GPR B scan (left is original scan and right is processed scan).

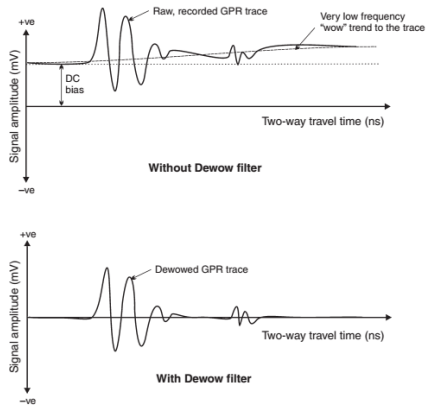


Figure 4.1 Dewow filter correction on a raw GPR trace (Cassidy 2009)

2) Time-zero Correction

Time-zero point is defined as the first arrival time of EM wave reflected by the ground. This point may deviate from its real position in GPR scan due to thermal drift, electric instability and etc. (Cassidy 2009). Time-zero point affects the ground interface

identification and the time sequence of later events. Therefore, GPR scans need to be adjusted to a common time-zero position before further processing. It is usually achieved by moving the time-zero point to the first negative peak of the trace. Figure 4.3 gives an example of time-zero correction in GPR scan (left is original scan and right is processed scan).

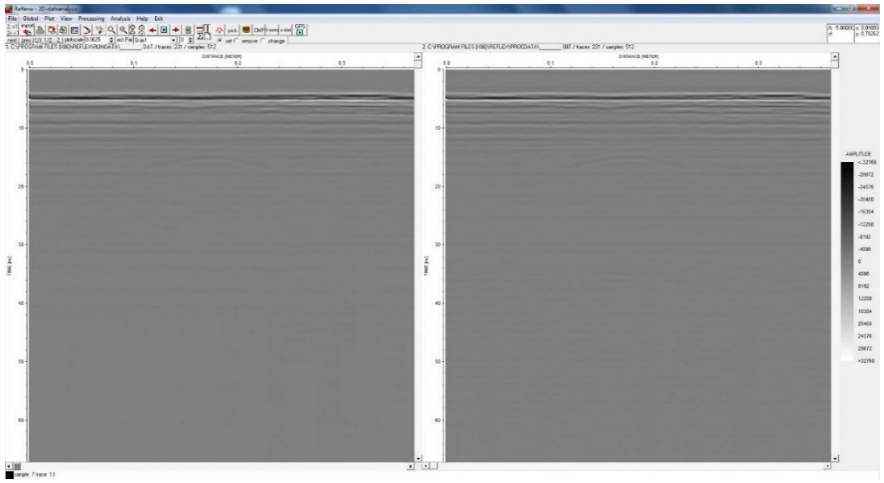


Figure 4.2 Example of dewow

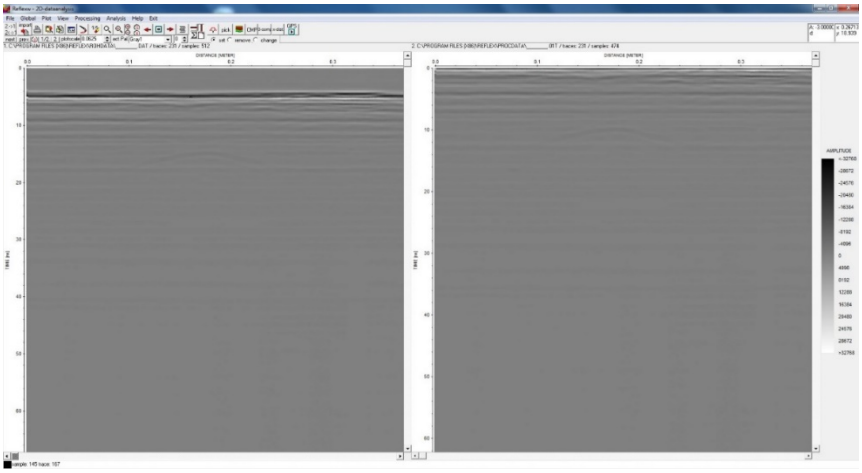


Figure 4.3 Example of time-zero correction

3) Gain

Gain aims to improve the visual form of GPR scan and/or change the data structure by altering the relative amplitudes (Cassidy 2009). The appearance of later arrivals in GPR scan may be blurry due to signal attenuation and geometrical scattering losses. In such situations, gains are needed to enhance the visibility by increasing the amplitudes of traces. In Figure 4.4, gain is applied to a B scan of a metallic pipe buried in clay. The hyperbola in the original scan is blurry due to large signal attenuations in clay soil (left in Figure 4.4). After gain is applied, the outline of the resulting hyperbola becomes clear, which assists the detection of buried utilities and further processing of GPR raw data.

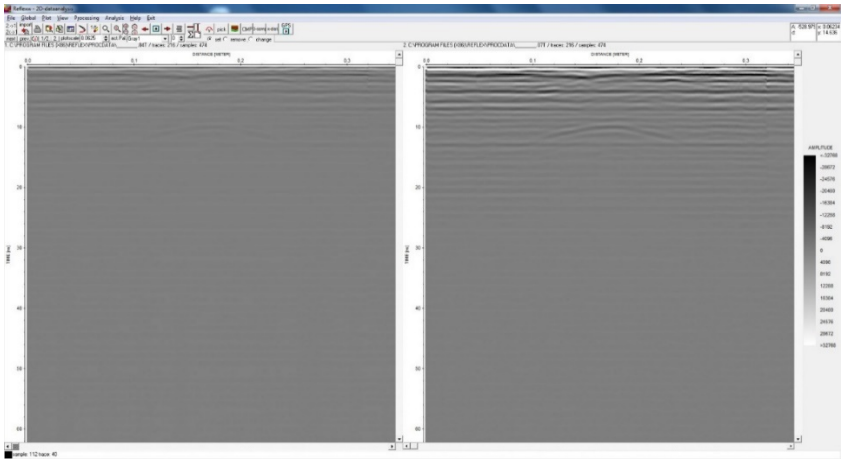


Figure 4.4 Example of gain in GPR scan

4) Background Removal

Background removal aims to remove human-induced and/or system noise. It is usually achieved by removing high-frequency ‘speckle’ from radio transmissions or the striping effect from antenna ringing (Cassidy 2009). The visual quality of GPR B scan will be enhanced as it shows in Figure 4.5 (left is original scan and right is processed scan).

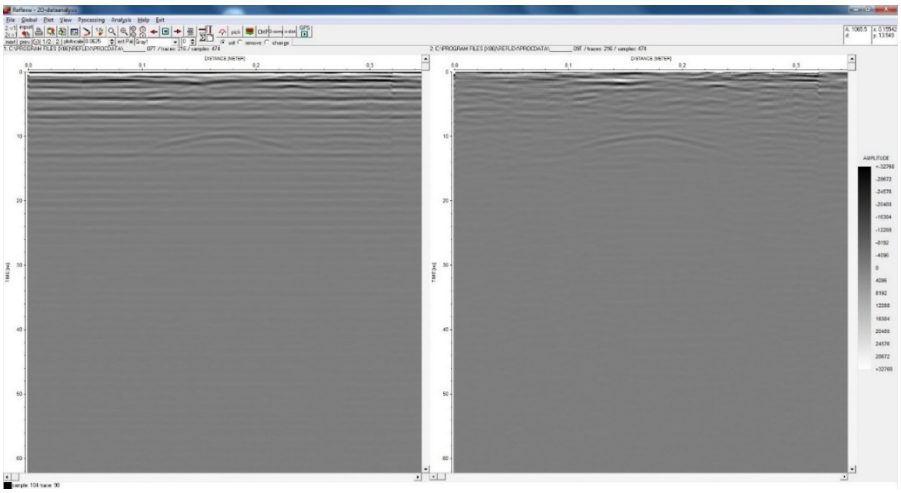


Figure 4.5 Example of background removal in GPR scan

After preprocessing, GPR raw data are extracted from the radagram. It is performed in the scan by picking the points with highest reflected amplitude from the zone that features hyperbolic patterns. Figure 4.6 shows a greyscale GPR scan of a steel pipe buried in clay soil. The white bands indicate positive and black bands represent negative peaks. The red dots are raw data points extracted from the hyperbola reflection. The extracted information is stored in an ASCII format, which includes antenna position x (m) on the scan trajectory, two-way time t (ns), amplitudes and GPS coordinates of that point. The pick with highest amplitude and shortest two-way time is considered as the hyperbola apex. The hyperbola apex indicates the position of utility apex. Table 4.1 gives an example of the extracted raw data.

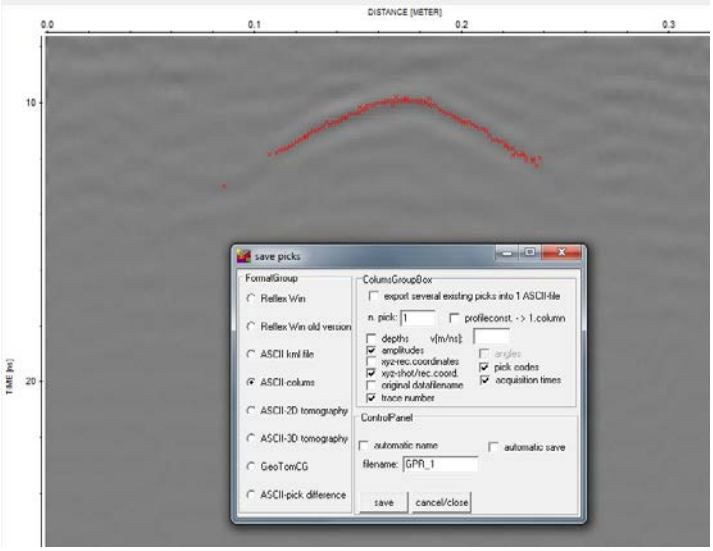


Figure 4.6 Raw data extraction from GPR scan

Table 4.1 Example of extracted raw data

Data point	Distance along the trajectory (m)	Tow-way time (ns)	Signal amplitude	GPS Coordinates of antenna	
				Northing	Easting
Point 1	0.11935	10.1482	-3828.3	4474337.734	505802.984
Point 2	0.14550	10.0485	-3886.2	4474337.712	505802.998
Point 3	0.17265	9.9873	-3947.6	4474337.689	505803.012
Point 4	0.19741	10.0545	-3905.5	4474337.668	505803.025
Point 5	0.22030	10.1515	-3833.9	4474337.648	505803.037

4.2 GPR Raw Data Modeling

Hyperbola equation is employed to model GPR raw data. The reason is that most reflection patterns are hyperbolic. Linear reflections are rarely seen because GPR scan directions are seldom along the buried utilities. The degree of accuracy for locating and characterizing highly depends on how a number of parameters are considered in the model. The most notable parameters are wave propagation velocity, radius and spatial orientation between GPR scan direction and buried utility. There is an evolution of

hyperbola equations used for GPR raw data modeling. They are point reflector model, cylindrical reflector model in perpendicular-to-utility scan, and cylindrical reflector model in generic scan. Thereafter, a detailed analysis of these three models is provided. It concludes that the cylindrical reflector model in generic scan is most suitable for GPR raw data modeling because it is more practical and can eliminate a large portion of estimation errors.

4.2.1 Point Reflector Model

Initially, underground utilities are modeled as long lines buried in homogeneous medium. GPR moves perpendicularly across the buried line (see Figure 4.7). Hence, the section of buried utility is regarded as a point with zero radius. Based on this assumption, AI-Nuaimy et al. (2000) put forward a model that relates two-way travel time with antenna position. EM wave propagation velocity is the primary parameter considered in that model.

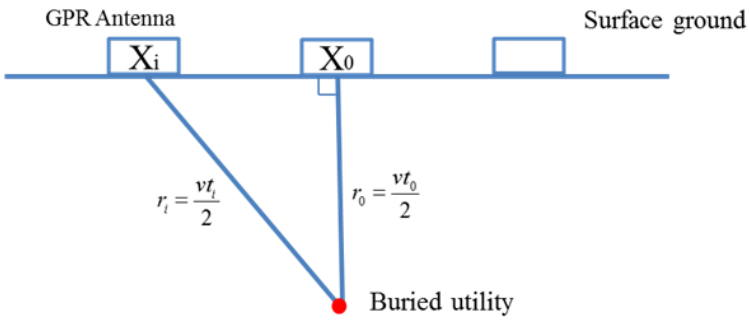


Figure 4.7 Hyperbola equation resulted from point reflector

In Figure 4.7, x_i denotes an arbitrary position along GPR survey trajectory, and x_0 is the position where GPR antenna is right above the buried utility. r_i is the distance from GPR antenna to the buried utility at position x_i. t_i is the two-way travel time, and v is the wave propagation velocity. A right triangle is formed, which formulates Equation 4.1.

$$r_i^2 = r_0^2 + (x_i - x_0)^2 \tag{4.1}$$

The substitution of r_i and r_0 with Equation 4.2 leads to Equation 4.3.

$$\begin{aligned} r_i &= \frac{vt_i}{2} \\ r_0 &= \frac{vt_0}{2} \end{aligned} \tag{4.2}$$

$$(\frac{t_i}{t_0})^2 - (\frac{2(x_i - x_0)}{vt_0})^2 = 1 \tag{4.3}$$

However, it has the least benefits to use Equation 4.3 for GPR raw data modeling. Equation 4.3 applies only if the utility is buried very deep and with a very small radius. Otherwise, large estimation errors will be incurred. This is demonstrated by Shihab and AI-Nuaimy (2005) by a quantitative analysis of the incurred errors.

4.2.2 Cylindrical Reflector Model in Perpendicular Scan

The point reflector model is improved by Shihab and AI-Nuaimy (2005) to take radius into consideration. While GPR scan direction is still perpendicular to the buried cylindrical utility. According to Chapter 3, the raw data (x_n, t_n) can be considered as the reflection of n^{th} key point on the cylindrical utility. In Figure 4.8, r_i denotes the distance from the antenna center at a specific position x_i ($i = -N, ...0..., N$) to the corresponding

key point on the cylindrical object. R is radius and v represents wave propagation velocity. As discussed before, a right triangle is formed and thus Equation 4.4 exists.

$$(r_i + R)^2 = (r_0 + R)^2 + (x_i - x_0)^2 \tag{4.4}$$

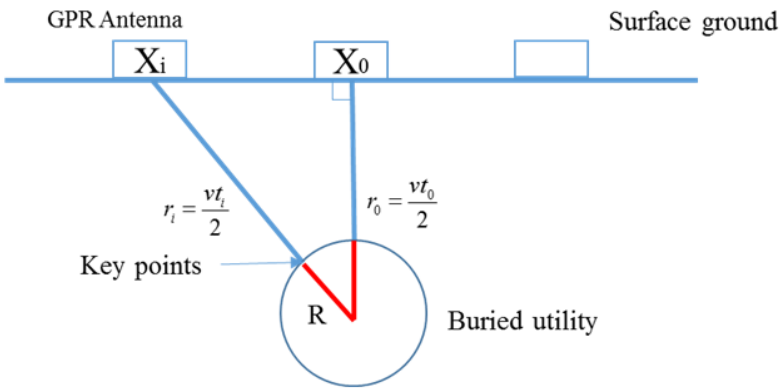


Figure 4.8 Cylindrical reflector model in perpendicular-to-utility scan

With Equation 4.2, it is possible to write the hyperbola equation as a function of R, v, x_0 and t_0 (see Equation 4.5).

$$t_i = \frac{2}{v}\left(\sqrt{\left(\frac{vt_0}{2} + R\right)^2 + (x_i - x_0)^2} - R\right) \tag{4.5}$$

Each pair of raw data (x_i , t_i) extracted from a hyperbolic reflection satisfies Equation 4.5 depending on the noise level. Equation 4.5 can be rewritten as the canonical hyperbola equation (see Equation 4.6).

$$\frac{(t_i + (2R/v))^2}{(t_0 + (2R/v))^2} - \frac{(x_i - x_0)^2}{((v/2)t_0 + R)^2} = 1 \tag{4.6}$$

The hyperbola is centered at (x_0 , $\frac{-2R}{v}$), and the hyperbola semi axes are defined as Equation 4.7 (see Figure 4.9).

$$\begin{aligned} a &= t_0 + \frac{2R}{v} \\ b &= \frac{v}{2} t_0 + R = \frac{v}{2} a \end{aligned} \tag{4.7}$$

With Equation 4.7, the angle between the hyperbola asymptotes and axis is given by Equation 4.8.

$$\varphi = \arctan(\frac{b}{a}) = \arctan(\frac{v}{2}) \tag{4.8}$$

Equation 4.8 indicates that the angle φ is directly proportional to v. Angle φ increases with increasing v, and vice versa (see Figure 4.9).

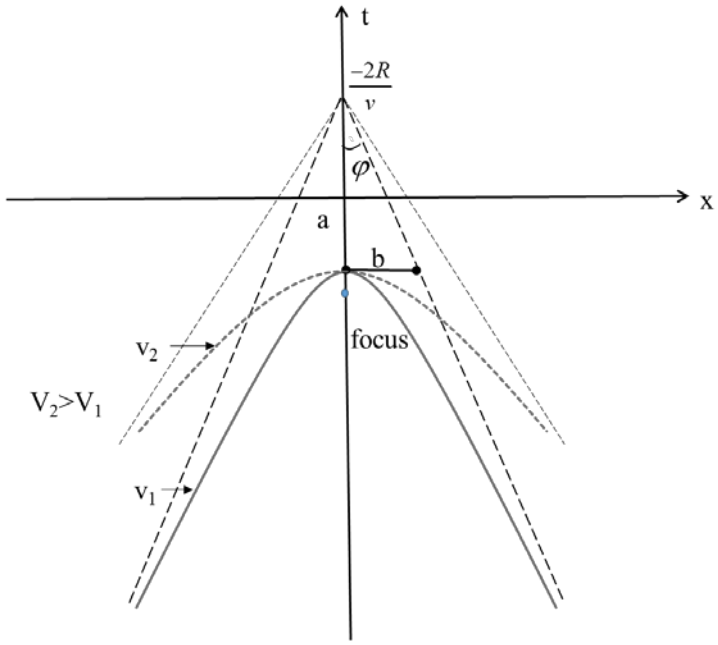


Figure 4.9 A hyperbola with its asymptotes and related parameters indicated

The consideration of radius R in the model eliminates a large portion of errors and greatly increases the accuracy, as demonstrated next. Figure 4.10 shows the comparison between hyperbolae resulting from line and cylinder. The apexes of the hyperbolae are located at the same position since they are buried in the same depth in the subsurface. The hyperbola resulting from line is centered at (x_0, 0). While the center of hyperbola resulting from cylinder moves away from the origin by $\frac{-2R}{v}$. Since the wave velocity remains the same and so does the angle between hyperbola asymptotes, the hyperbola resulting from cylinder is wider.

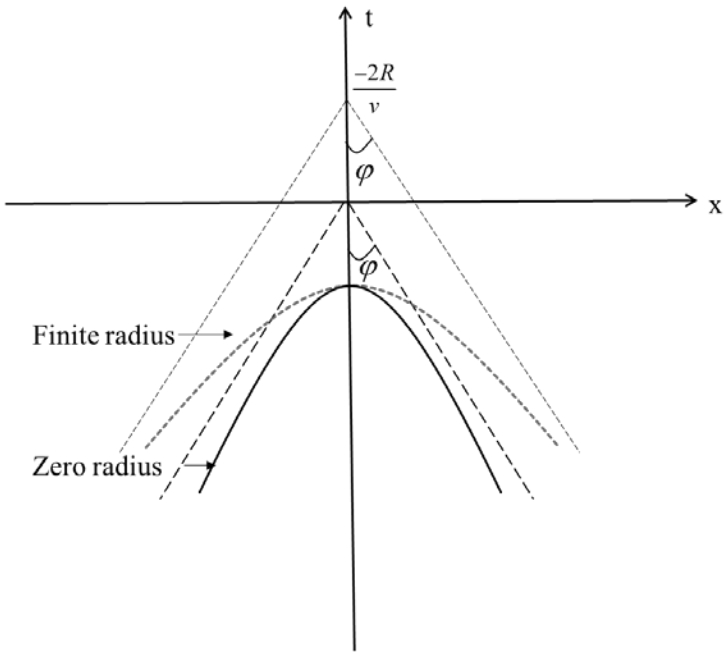


Figure 4.10 Comparison between zero radius hyperbola and finite radius hyperbola

If the radius of a cylinder is assumed to be zero and Equation 4.3 is used to model GPR raw data, then the estimated wave propagation velocity v will be artificially enlarged. As

illustrated in Figure 4.11, the true asymptotes of hyperbola resulting from cylinder are the red lines with an angle of Ω. When Equation 4.3 is applied, the asymptotes will be assumed to be the yellow dash lines with angle of ω. Apparently, $\omega > \Omega$ and thus the estimated velocity from Equation 4.3 is exaggerated. As a result, the buried utility will be estimated at a location deeper than where they are in reality, causing disastrous consequences during excavation.

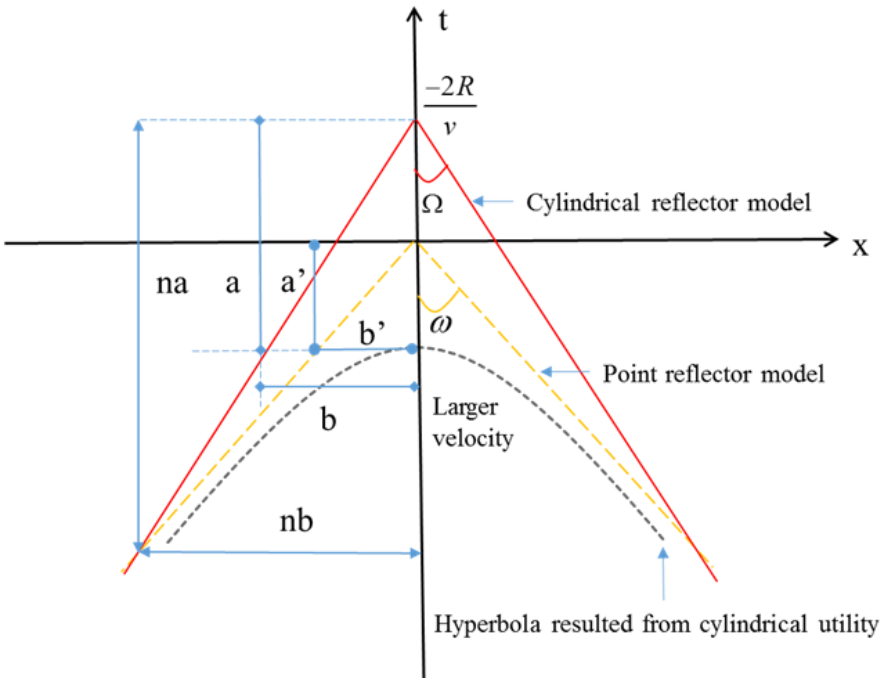


Figure 4.11 Impacts of applying point reflector model to cylindrical utility (modified from Shihab and AI-Nuaimy 2005)

Mathematically, Shihab and AI-Nuaimy (2005) quantifies the velocity estimation error of applying point reflector model to cylindrical utility. In figure 4.11, the angles between

axis and asymptote of these two models can be derived in Equation 4.9 (Shihab and AI-Nuaimy 2005).

$$\tan\omega = \frac{b'}{a'} = \frac{nb}{na - \frac{2R}{v}} = \frac{b}{a - \frac{2R}{nv}}$$
$$\tan\Omega = \frac{b}{a} \tag{4.9}$$

Where n is a constant representing the ratio between the horizontal offset part where the asymptotes of the two models intersect and the depth of the target. The enlarged velocity v' from point reflector model can be calculated as:

$$v' = \frac{2b'}{a'} \tag{4.10}$$

The percentage error in the velocity estimation is given by Shihab and AI-Nuaimy (2005) in Equation 4.11.

$$e_v = \frac{v' - v}{v} \times 100\% = \frac{R}{nb - R} \times 100\% \tag{4.11}$$

Since depth is related directly to speed, therefore this is also the percentage error for depth estimation. The value of n can be obtained from Figure 4.11 and is given by Equation 4.12 (Shihab and AI-Nuaimy 2005).

$$n = \frac{R/b}{(\tan\varphi \tan\omega - 1)} \tag{4.12}$$

The value of b increases with increasing depth. R/b is called radius-to-depth-to-center radio. A numerical analysis conducted by Shihab and AI-Nuaimy (2005) elucidated that the error percentage increases sharply as the utility is buried shallower and with larger

radius. Even for modest R/b ratios, the estimation error can be significant (see Figure 4.12).

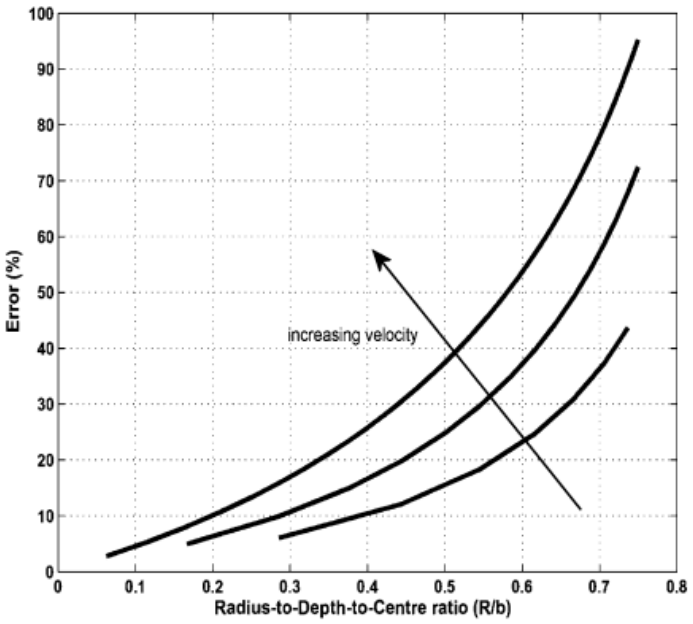


Figure 4.12 Percentage error with respect to R/b ratio (Shihab and AI-Nuaimy 2005)

Therefore, it confirms that the point reflector model is not appropriate for modeling GPR raw data and the cylindrical reflector model is more realistic and practical.

4.2.3 Cylindrical Reflector Model: Generic Scan

The previous models are all based on the assumption that GPR moves perpendicularly to scan buried utilities. However, this is not realistic and practical since most often the horizontal positions of existing underground utilities are unavailable. In this thesis, we take the spatial orientation between GPR scan trajectory and buried utility into

consideration to refine previous hyperbola model. The generic cylinder model extends the application scenarios and significantly improves locating and characterizing accuracy.

Figure 4.13 presents the geometric aspects of the generic cylinder model. In Figure 4.13, x-y plane represents the ground surface. A utility is buried in the subsurface. The vertical inclination of the buried utility is β, which is the angle between the utility and ground surface (x-y plane). The horizontal orientation is α, which is defined as the angle between the projection of utility on x-y plane and GPR scan direction. The GPR scan direction in Figure 4.13 is parallel to the x axis. The angle between the utility and the GPR scan trajectory is θ.

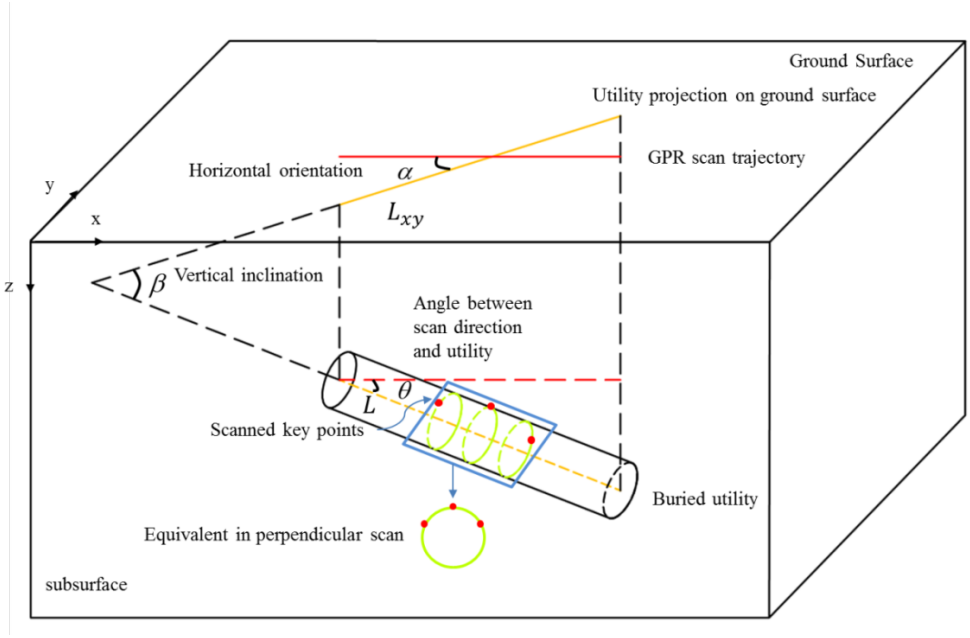


Figure 4.13 Geometric illustration for generic cylindrical reflector model

The generic hyperbola model to represent the GPR raw data points is derived as follows.

1) Difference between the perpendicular-to-utility scan and the generic scan

As GPR moves across the buried utility, the corresponding points on cylinder are scanned and the associated two-way times are recorded. The scanned points in a generic scan are equivalent to those in a perpendicular-to-utility scan (see Figure 4.13). Given the high velocity of EM wave (i.e. the speed of light) and the relative short traveling distance, the variations of two-way times of scanned points between a generic scan and a perpendicular-to-utility scan are negligible. In other words, the difference between a generic scan and a perpendicular-to-utility scan lies in the distance of GPR scan trajectory. Figure 4.14 illustrates this difference that is also captured in Equation (4.13).

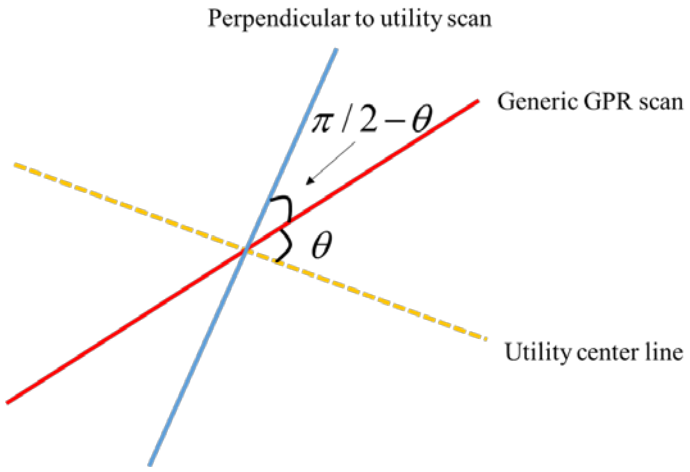


Figure 4.14 Difference in trajectory distance between generic GPR scan and perpendicular-to-utility scan

$$L_{GP} = L_G \times \sin\theta \tag{4.13}$$

Where L_{GP} is the length of GPR trajectory in the perpendicular-to-utility scan, and L_G is that of a generic scan.

The substitution of Equation (4.13) into Equation (4.6) results in Equation (4.14), a more generic hyperbola equation.

$$\frac{(t_i + (2R/v))^2}{(t_0 + (2R/v))^2} - \frac{(x_i - x_0)^2 \sin\theta^2}{((v/2)t_0 + R)^2} = 1 \tag{4.14}$$

2) Consideration of the Relative Angle

The relative angle θ is difficult to accurately acquire. To address this issue, the angle θ is analytically expressed in terms of horizontal orientation α and vertical inclination β. In the third phase of the methodology, the determination of α and β will be explained in detail. Here, it only describes how θ can be expressed by α and β.

The relation among θ, α and β is derived by three projections. First, the utility centerline is projected to the ground surface using Equation (4.15).

$$L_{xy} = L \times \cos\beta \tag{4.15}$$

Where L_{xy} is the length of projected utility centerline, and L is the length of utility centerline. Second, L_{xy} is projected to the GPR scan trajectory using Equation (4.16).

$$L_G = L_{xy} \times \cos\alpha \tag{4.16}$$

Where L_G is the length of GPR scan trajectory. Equation (4.15) and (4.16) lead to Equation (4.17).

$$L_G = L \times \cos\alpha \times \cos\beta \tag{4.17}$$

Third, the utility centerline can be projected to the GPR scan trajectory directly using Equation (4.18).

$$L_G = L \times \cos\theta \tag{4.18}$$

Combining Equations (4.17) and (4.18) leads to Equation (4.19) that mathematically expresses the relationship between θ and α and β.

$$\cos\theta = \cos\alpha \times \cos\beta \tag{4.19}$$

3) Impacts of Relative Angle on Estimation Accuracy

It is essential to consider the relative angle in estimating wave velocity and radius, otherwise large error will occur as demonstrated by the quantitative analysis below.

1. Error in Estimating Velocity

The semi-axes of the hyperbola in a generic scan are given in Equation (4.20).

$$\begin{aligned} a &= t_0 + \frac{2R}{v} \\ b &= \frac{\frac{v}{2}t_0 + R}{\sin\theta} \end{aligned} \tag{4.20}$$

The EM wave velocity is then calculated by Equation (4.21).

$$v = \frac{2(b\sin\theta - R)}{t_0} \tag{4.21}$$

However in previous studies and current practices, if the relative angle is unknown, it is assumed that GPR moves across the buried utility perpendicularly. Therefore, the resulting velocity v' is calculated by Equation (4.22).

$$v' = \frac{2(b - R)}{t_0} \tag{4.22}$$

Therefore, error is incurred and the error percentage e_v is calculated by Equation (4.23).

$$e_v = \frac{v' - v}{v} = \frac{(1 - \sin\theta)}{\sin\theta - \frac{R}{b}} \tag{4.23}$$

It is clear from Equation (4.23) that the error percentage is a function of the relative angle, and the R/b ratio. The error increases with decreasing θ and increasing R/b ratio. The value of b increases with increasing burying depth, and vice versa. In other words, the error in estimating velocity is significant if ignoring the relative angle given that:

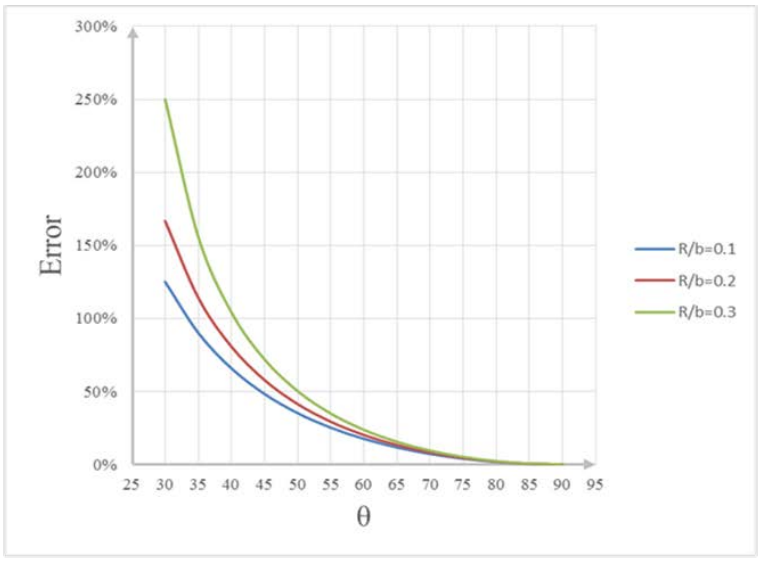
1) the utility is buried in a shallow depth, i.e., R/b is large;
2) the relative angle θ is small.

The error percentage e_v can also be calculated by Equation (4.24) given the relationship among the relative angle θ, horizontal orientation α and vertical inclination β.

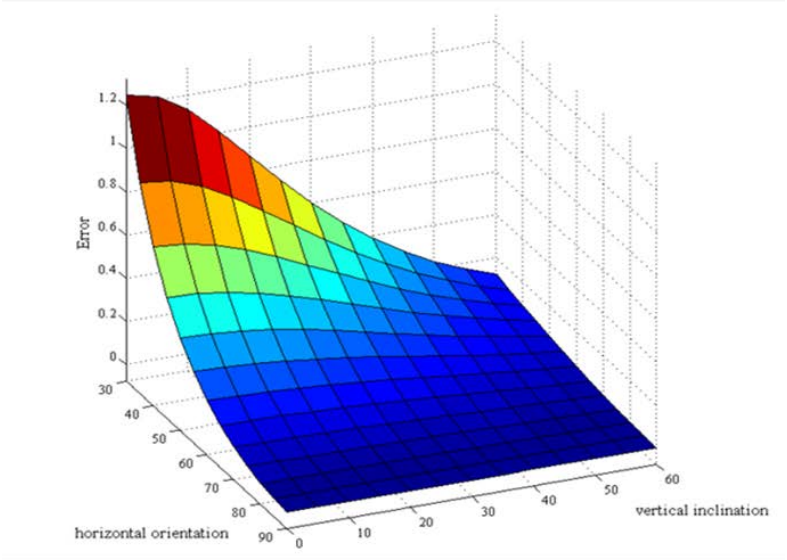
$$e_v = \frac{(1 - \sqrt{1 - \cos^2\alpha\cos^2\beta})}{\sqrt{1 - \cos^2\alpha\cos^2\beta} - \frac{R}{b}} \tag{4.24}$$

Figure 4.15(a) presents the error percentage varying with the relative angle and different R/b ratios. Figure 4.16 (b) plots the error percentage versus the horizontal orientation and vertical inclination of the buried utility given that R/b equals 0.1.

The error induced in velocity estimation can be costly and even catastrophic. In addition, the wave velocity derived without considering the relative angle is always artificially magnified. Therefore, the buried utility is always assumed at a depth much greater than they are in real world. This may lead to falsely instilled confidence and potentially misleads excavation operators into unintentional strikes.



(a)



(b)

Figure 4.15 Error magnitudes of velocity when ignoring the relative orientation

2. Error in Estimating Radius

Similarly, the error percentage e_R in estimating radius is given by Equation (4.25).

$$e_R = \left| \frac{1-\sin\theta}{\sin\theta - \frac{vt_0/2}{b}} \right| \tag{4.25}$$

When the utility is buried very deep and the radius is relatively small (i.e., the ratio $\frac{vt_0/2}{b}$ is proximate to 1), the error percentage in estimating radius can be as high as 100%. When the utility is buried shallow and the radius is relative large (take the ratio $\frac{vt_0/2}{b}$ to be 1/3 as an example), the error percentage e_R varies with the relative angle, as presented in Figure 4.16.

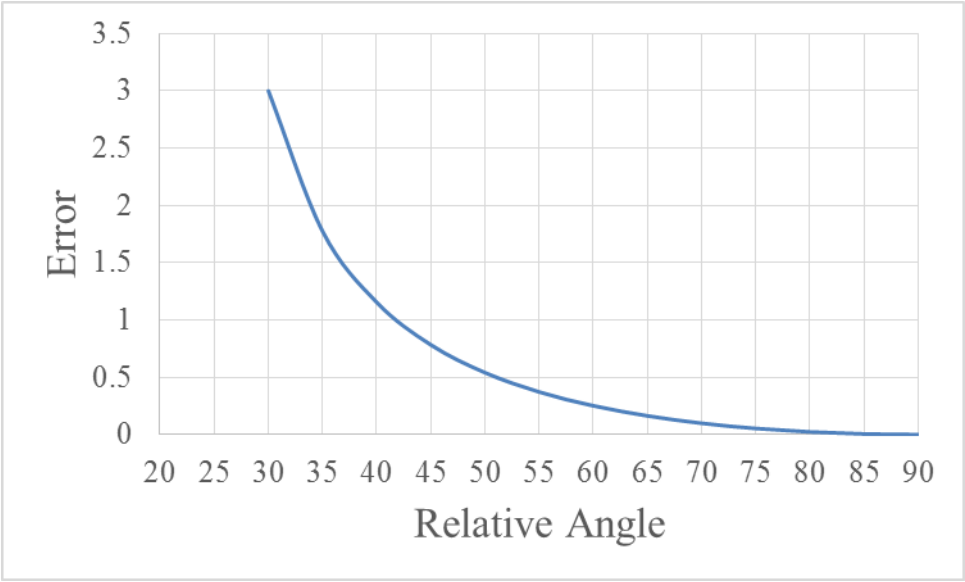


Figure 4.16. Error percentage in estimating radius when ignoring the relative angle

4.3 Very Important Point Algorithm

This section first elucidates the VIP model given the perpendicular-to-utility scenario, and then formulates the two nonlinear equations of EM wave velocity and utility radius.

Thereafter, it describes a mechanism to search the VIP among massive GPR raw data points to derive the velocity and radius. Towards the end of this section, the VIP algorithm is extended to a more generic scenario (i.e., GPR does not scan the buried utility perpendicularly) with assistance of GPS data.

4.3.1 VIP Model

The very important point is defined as the key point on cylinder that is detected by the boundary of GPR footprint. Figure 4.17 clearly presents the concept. The estimation of utility radius and wave propagation velocity is through solving two nonlinear equations derived from the VIP model.

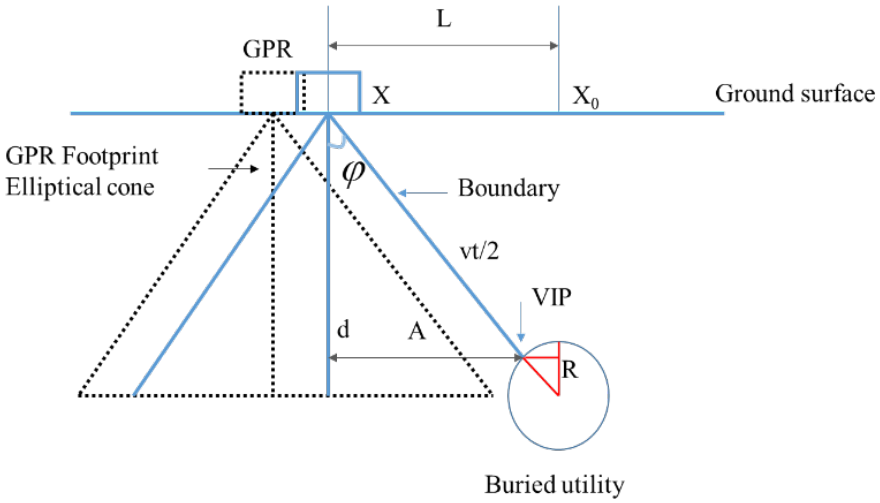


Figure 4.17 Very Important Point Model

In Figure 4.17, position x is where the very important point is detected by GPR, and position x_0 is where GPR antenna is directly above the utility apex. The distance between

these two points is L that can be derived from GPR scan. The distance between the center of GPR at location x and the very important point is given by vt/2 (v is the EM wave velocity and t is the two way time obtained from GPR raw data). A is the major semi-axis of GPR footprint. It is easy to derive Equation (4.26) and (4.27).

$$L - A = R \times \sin\varphi \tag{4.26}$$

$$\sin\varphi = \frac{2A}{vt} \tag{4.27}$$

Substituting Equation (4.27) to Equation (4.26) leads to Equation (4.28).

$$R = \frac{(L-A)vt}{2A} \tag{4.28}$$

A can be calculated using Equation (4.29) (Yalciner 2009).

$$A = \frac{\lambda}{4} + \frac{d}{\sqrt{(\varepsilon+1)}} \tag{4.29}$$

Where d is the depth from the center of GPR to the VIP point. It is calculated by Equation (4.30).

$$d = \sqrt{\left(\frac{vt}{2}\right)^2 - A^2} \tag{4.30}$$

Substitution of d to Equation (4.29) results in Equation (4.31).

$$A = \frac{\frac{\lambda(\varepsilon+1)}{2} + \sqrt{\frac{\lambda^2(\varepsilon+1)^2}{4} - 4(\varepsilon+2)[\frac{\lambda^2(\varepsilon+1)}{16} - \frac{(vt)^2}{4}]}}{2(\varepsilon+2)} \tag{4.31}$$

The EM wave length λ and average dielectric constant ε of the subsurface medium are given by Equation (4.32) and (4.33).

$$\lambda = \frac{v}{f} \tag{4.32}$$

$$\varepsilon = \left(\frac{c}{v}\right)^2 \tag{4.33}$$

Where f is the frequency of GPR antenna and c is the speed of light, both of which are constant values. Therefore apparently, A is a univariate function of EM wave velocity. In other words, Equation (4.28) represents one relationship between the radius R and EM wave velocity v.

Furthermore, the VIP is on the hyperbola and thus satisfies the hyperbola equation that is rewritten as Equation (4.34).

$$R = \frac{L^2}{v(t - t_0)} - \frac{v(t + t_0)}{4} \tag{4.34}$$

Where t_0 is the two way time where GPR is directly above the utility apex. It can be read from GPR scan. Equation (4.34) presents another relation between R and v. The radius and EM wave propagation velocity are estimated by solving these two nonlinear equations (Equation (4.28) and (4.34)) derived from VIP model. Once the wave velocity v is estimated, the burying depth is calculated using Equation (4.35).

$$d = \frac{vt_0}{2} \tag{4.35}$$

Previous deduction is based on the premise that the VIP is known. The following part illustrates how the VIP can be found among the massive GPR raw data points.

4.3.2 Search for VIP

The VIP is searched following an assume-and-check strategy. Every point in GPR raw data is assumed to be VIP to estimate the radius and velocity. Based on the derived radius and velocity, these points are checked with two statistical criterions to find out the VIP.

To save computational efforts, the searching area is narrowed down to certain area, see Figure 4.18. It was found that the resulting R becomes "abnormal" as the points do not fall into the searching area. Specifically, if the point approaches the apex of hyperbola, the resulting R will become less than zero. While if the point approaches to the endpoints of hyperbola, the resulting R will be extremely large and thus can be discarded. In this sense, the search for VIP starts at the two endpoints of hyperbola assuming every point is VIP to estimate R and v, and stops until the resulting R is equal to or less than 0.

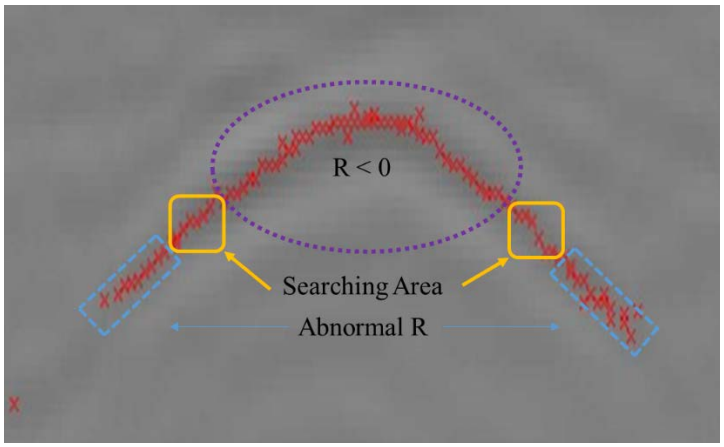


Figure 4.18 Searching area for VIP

In order to enhance the algorithm robustness, a searching buffer is utilized to handle the noise produced by heterogeneity of subsurface soil, GPR survey process and etc. The existence of noise will cause VIP "jump" away from its real position. The searching buffer takes such deviation into consideration. Figure 4.19 illustrates the mechanism of searching buffer. A rectangular buffer with $2N\Delta t$ and $2N\Delta d$ is built around a point (x, t) in GPR raw data. All the points in this buffer with coordinates $(x+i\Delta d, t+i\Delta t)$ (i=-N, …, 0, …, N) will be searched and checked.

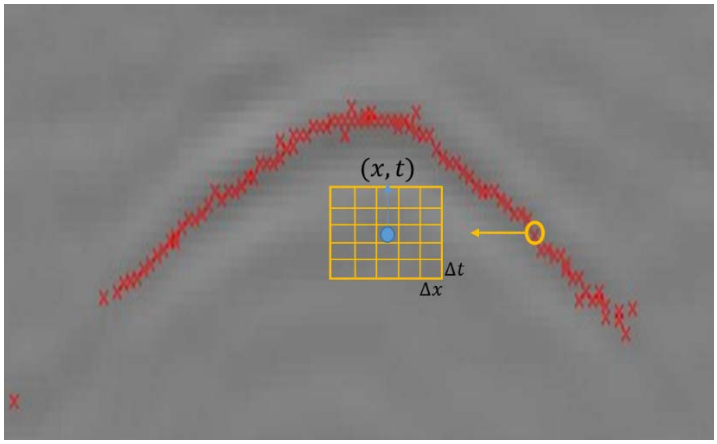


Figure 4.19 Searching buffer for VIP

Once all the potential VIPs are searched and corresponding R and v are derived, two statistical criterions are used to check the VIP out. The first criterion is based on the root mean square error (RMSE) between GPR raw data and a fitted hyperbola. The estimated radius R and wave velocity v are substituted to the hyperbola equation to calculate the RMSE of the GPR raw data. A perfect hyperbola is synthesized and the points on that hyperbola are assumed to be VIP for estimating R and v. It reveals that theoretically the R

and v derived at the VIP leads to the least RMSE, see Figure 4.20. As such, the first criterion is called least RMSE (LRMSE) rule. However, the GPR profile of cylindrical buried utility is imperfect hyperbola that has certain deviations due to heterogeneity in subsurface medium, imperfect survey and processing methods. Therefore the LRMSE rule itself cannot guarantee in real world conditions that the point with Least RMSE is definitely the VIP. But it has been noted that the VIP is certainly with a very small RMSE that is proximate to the least value. Hence, several points with least RMSE are selected for further check based on the second criterion.

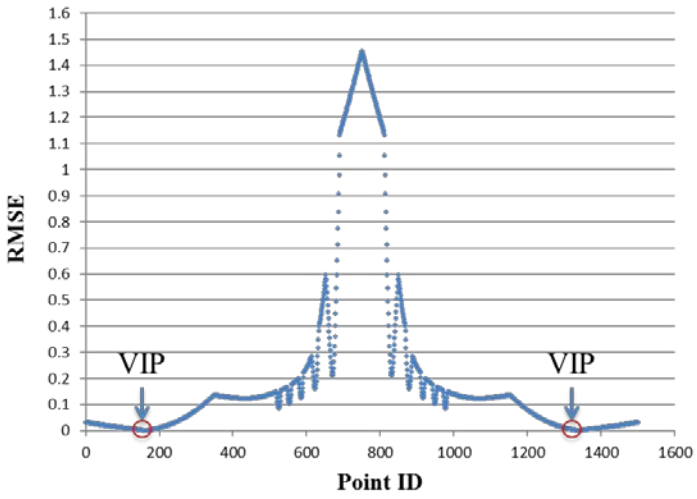


Figure 4.20 LRMSE rule

The second criterion is called highest votes rule. It works in such a way to filter out the unwanted points.

1) For a derived velocity v, the coordinates (x, t) of all the raw data points are used to calculate the radius R since R is the only unknown variable in the hyperbola equation.

This results in a histogram generated by a number of R values. The counts in each column are referred to votes.

2) The second step is to check if the value of R derived from VIP model falls into the column with highest votes. If it doesn't, the point will be discarded.
3) If there are more than one point satisfy (2), then the one with higher votes is considered as the VIP.

Once the VIP is found, the depth and radius of buried utility can be estimated as discussed before. Next, we will explain how to extend the VIP algorithm to a generic scenario where GPR does not scan the buried utility perpendicularly.

4.3.3 Extension to Generic Scenario

It is essential to estimate the horizontal orientation and vertical inclination of the buried utility to achieve the extension of VIP algorithm to generic scenario. In this study, the horizontal orientation is measured using auxiliary GPS data and the vertical inclination is estimated based on a trial-and-error process. The steps are detailed as follows.

Step 1 is to estimate the horizontal orientation α, illustrated in Figure 4.21. GS 1 and GS 2 are two successive generic scans; P1 and P2 are the corresponding hyperbola apex in GS1 and GS2. P1 and P2 indicate the apexes of the buried utility, and thus determine the direction of the buried utility section. The scanning direction can be easily determined using the endpoints of the scanning trajectory (e.g., P3 and P4 of GS1). Therefore, the horizontal orientation α is calculated using Equation (4.36).

$$\alpha = \arctan\left|\frac{k_1 - k_2}{1 - k_1 k_2}\right| \tag{4.36}$$

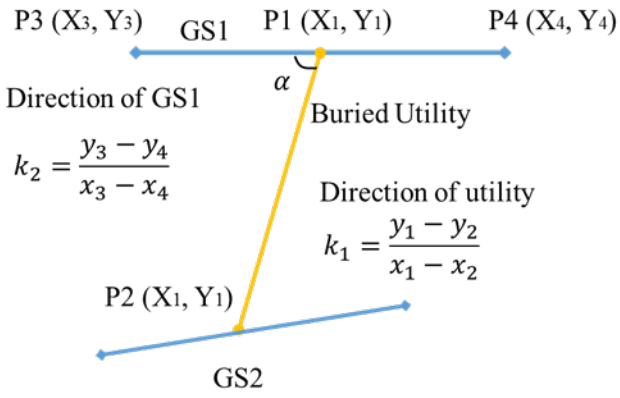


Figure 4.21 Horizontal orientation estimation

Step 2 is to estimate the vertical inclination β and eliminate the error term in wave velocity through a trial-and-error method. Based on previous analysis, the real velocity v is calculated as Equation (4.37).

$$v = \frac{v'}{(1 + e_v)} \tag{4.37}$$

Where v' is derived by VIP algorithm assuming a perpendicular-to-utility scan. e_v is the error percentage that is computed using Equation (4.24). Substitution of Equation (4.24) to Equation (4.37) leads to Equation (4.38).

$$v = \frac{v'(\sqrt{1 - \cos^2\alpha\cos^2\beta} - R'/b')}{1 - R'/b'} \tag{4.38}$$

Where R' and b' are derived in VIP model assuming a perpendicular-to-utility scan.

The trial-and-error process starts by substituting the estimated α to Equation (4.38) and

simultaneously initializing the unknown β to zero to calculate v. The estimated velocity is then used to calculate the angle β using Equation (4.39) (see Figure 4.22).

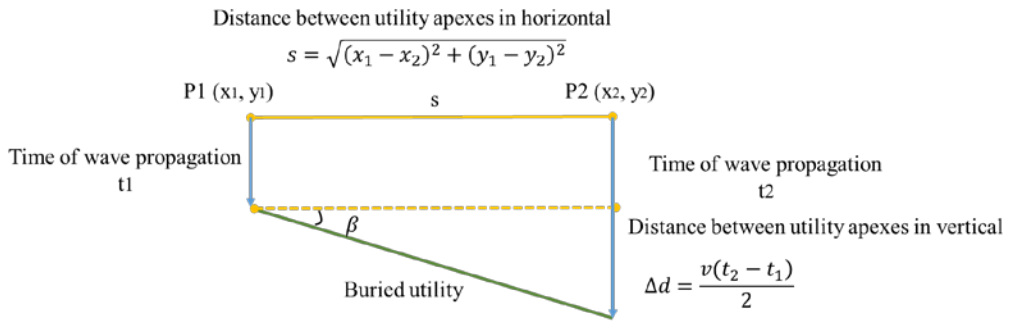


Figure 4.22 Vertical inclination estimation

$$\beta = \arctan(\frac{\Delta d}{s}) \tag{4.39}$$

Subsequently, the new value of β is again substituted to Equation (4.38) and the process is repeated until the difference between the two values of vertical inclination is less than a threshold (e.g., 0.5 degree). Consequently, the relative angle and real velocity are derived. Step 3 is to use the relative angle and velocity derived in Step 2 to estimate the radius as discussed before. In sum, the whole process of the VIP algorithm is summarized in the flowchart shown in Figure 4.23.

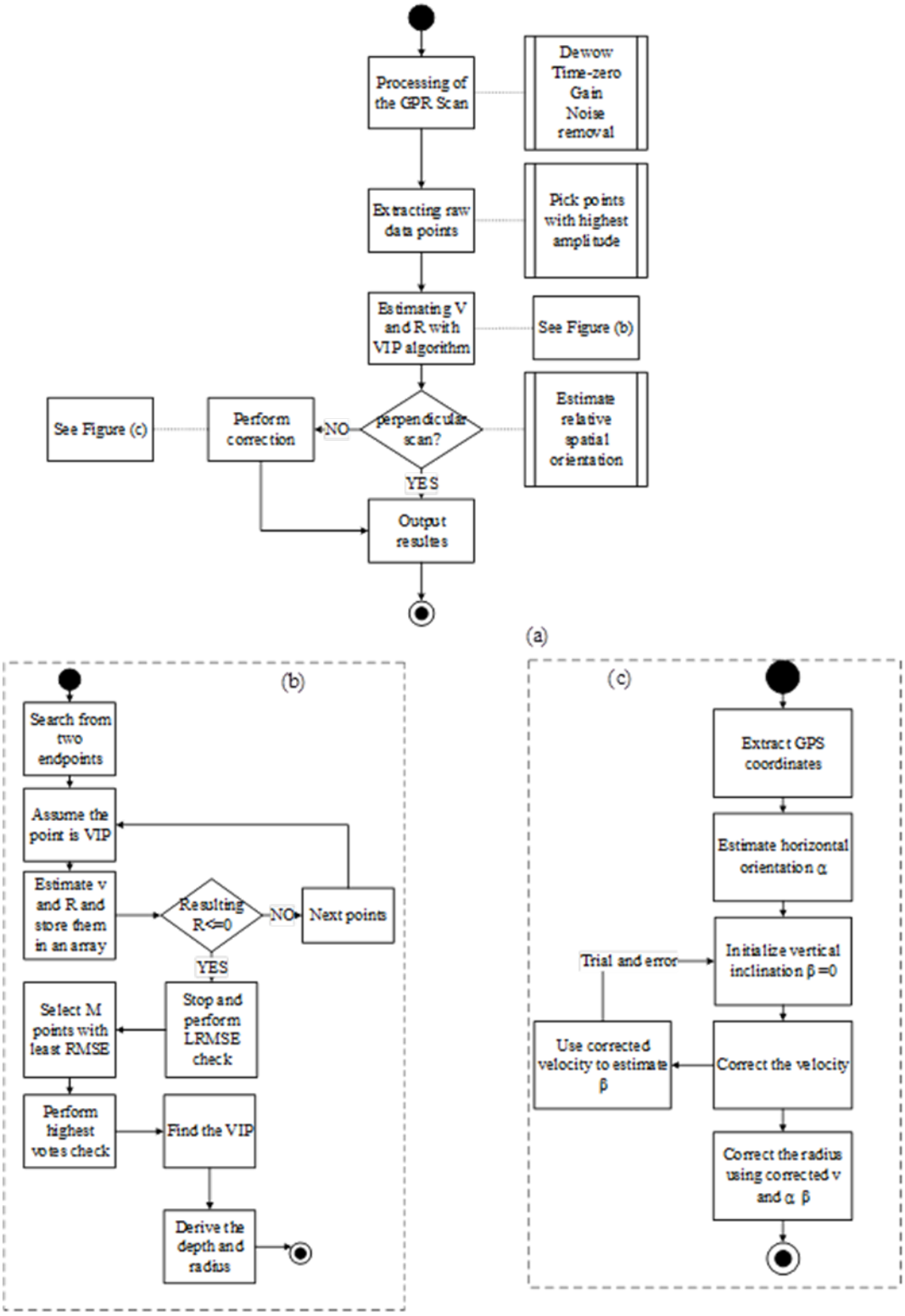


Figure 4.23 Overview of VIP algorithm

4.4 Validation of VIP Algorithm

This section validates the VIP algorithm by six examples to illustrate that it can estimate wave velocity and radius under various settings. Examples 1-4 are taken from the study of Ristic et al. (2009). Examples 5-6 were obtained from field experiments. A description of these six examples is provided in Table 4.2.

Table 4.2 Common characteristics of the GPR raw data

Example	Location	Pipe type	Radius (cm)	Depth (cm)	GPR scan characteristics
1	Urban street area Novi Sad, Serbia	Gas line metal pipe	17.78	Not provided	Perpendicular-to-pipe scan, 400MHz Antenna
2	Urban street area Novi Sad, Serbia	Gas line metal pipe	17.78	Not provided	Perpendicular-to-pipe scan, 400MHz Antenna, different soil from example 1
3	Test site Latvia/Estonia	Metal pipe	26.5	Not provided	Perpendicular-to-pipe scan, 900 MHz Antenna
4	Urban street area Kikinda, Serbia	Gas line metal pipe	10.995	Not provided	Perpendicular-to-pipe scan, 400 MHz Antenna
5	Purdue University, U.S.	Metal pipe	10.5	25.3	Perpendicular-to-pipe scan, 800 MHz Antenna
6	Purdue University, U.S.	Metal Pipe	10.5	21.6	Generic scan, 800 MHz Antenna

Figure 4.24 illustrates the GPR raw data of the first four examples. Table 4.3 shows the results of the VIP algorithm together with results obtained by Ristic et. al. (2009). The results demonstrate that the VIP algorithm can simultaneously estimate radius R and wave velocity v in perpendicular-to-utility scans. The results are close to those obtained by Ristic et al. (2009).

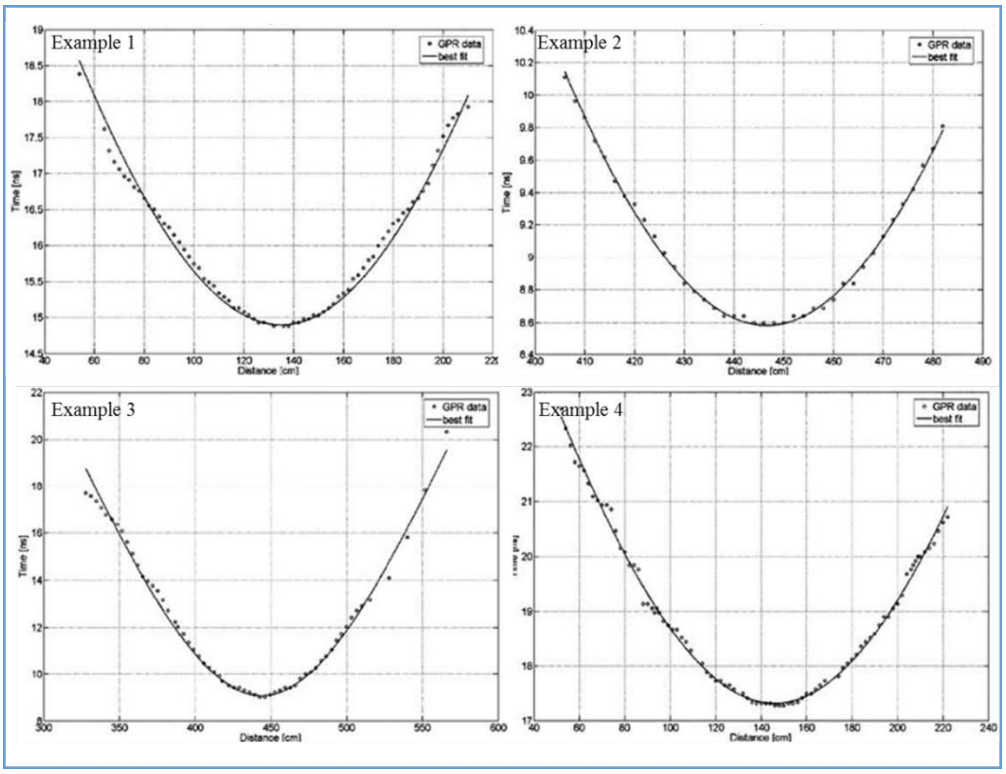


Figure 4.24 GPR raw data of examples 1-4 (Ristic et al. 2009)

Table 4.3 Results of Example 1-4

Example	True Radius (cm)	VIP algorithm			Algorithm of Ristic et al. (2009)		
		Estimated Radius (cm)	Radius Estimation Error	Estimated Velocity (cm/ns)	Estimated Radius (cm)	Radius Estimation Error	Estimated Velocity (cm/ns)
1	17.78	17.140	3.6%	13.605	18.17	2.2%	13.57
2	17.78	18.093	1.8%	13.161	18.2	2.4%	13.21
3	10.995	10.659	3.1%	12.228	10.26	6.7%	12.51
4	26.5	26.992	1.6%	11.998	25.83	2.5%	12.41

Examples 5-6 were conducted in West Lafayette, Indiana to evaluate the VIP algorithm in estimating burying depth of the utility under generic scanning scenarios. Figure 4.25 and Figure 4.26 illustrates the usage of a metallic cylindrical object in the experiment and the usage of RTK GPS and GPR in data collection.

Figure 4.25 Metallic cylindrical object buried in wet clay soil

RTK GPS is used before and after burial to measure the actual buried depth of the metallic object. Sequentially, GPR is used to scan the buried object (see Figure 4.26).

Figure 4.26 Buried depth measurement using RTK GPS and GPR

Figure 4.27 illustrates the GPR raw data of example 5. Table 4.4 presents the results of applying the VIP algorithm to estimate the velocity, depth, and radius. The error in velocity and radius are 16.87% and 7.42%, respectively. The relatively large error is

attributed to two possible causes, the backfill not being compacted, which caused the EM wave to scatter between layers and a large deviation of the VIP from its true position; and the poor performance of GPR in wet clay.

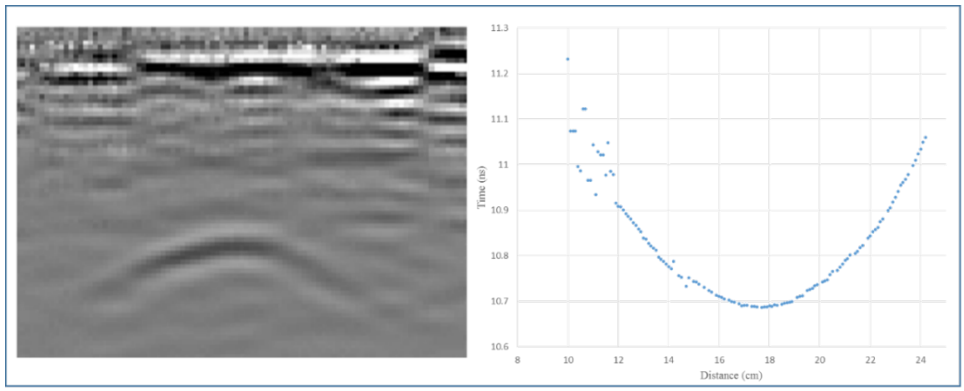


Figure 4.27 GPR raw data of example 5

Table 4.4 Results of example 5

	Radius (cm)	Velocity (cm/ns)	Depth (cm)
True value	10.50	4.74	25.30
VIP derived results	11.28	3.94	21.10
Estimation error	0.78	0.80	4.20
Error percentage	7.42%	16.87%	16.6%

In example 6, the metal cylinder was buried with a vertical inclination of five degrees. The raw data of example 6 is shown in Figure 4.28. Two generic scans are obtained to retrieve the coordinates of the two apexes on the cylindrical object (see Table 4.5). Since the buried cylindrical object is short and GPR scanned it with an angle, the reflection is not a completed hyperbola.

Table 4.5 Coordinates retrieved from GPR scan

Points	Utility apex P1	Utility apex P2	Track point P3	Track point P4
Coordinate (x)	4474337.639N	505803.282E	4474337.942N	505802.952E
Coordinate (y)	4474337.648N	505802.971E	4474336.798N	505803.207E

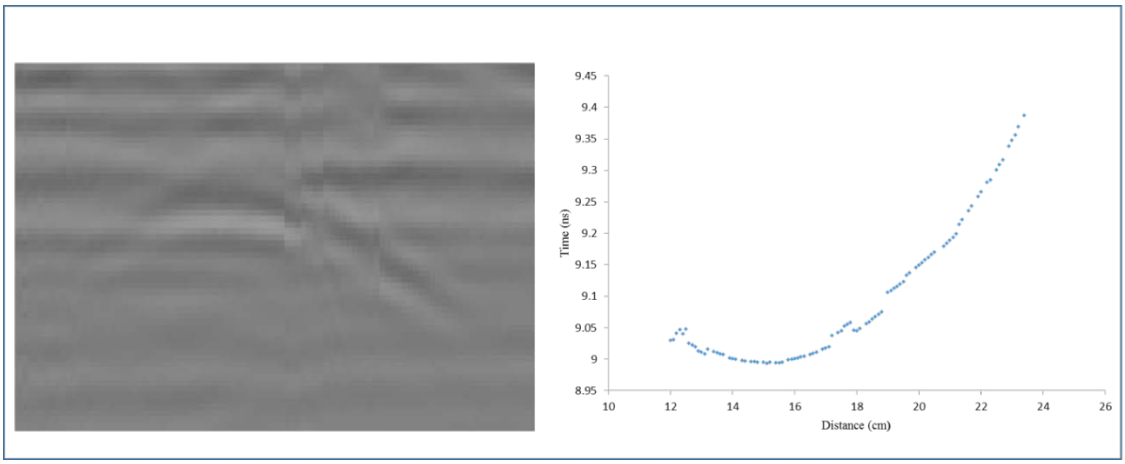


Figure 4.28 GPR raw data of example 6

Table 4.6 presents the results for example 6 for both cases of considering and not considering the relative angle. Larger errors associate the results that did not consider the relative angle. The error percentage in velocity is larger than that of radius when ignoring the relative angle. This is mainly because the cylinder has a relative large radius (compared to its shallow burying depth), which means the R/b ratio is large, leading to a larger error percentage in velocity estimation. Comparing example 6 with example 5, it is found that the errors incurred by the unawareness of relative angle is eliminated. The remaining error is close to that of example 5. This example validates the applicability of VIP algorithm in estimating burying depth and radius of underground utilities from generic GPR scans. It demonstrates the correction process can greatly improve the accuracy in estimating both radius and depth of the underground utilities.

Table 4.6 Results of Example 6

	Radius (cm)	Velocity (cm/ns)	Depth (cm/ns)	Radius Error	Velocity Error
True value	10.5	4.805	21.60	/	/
VIP ignoring relative angle	13.063	6.437	28.93	24.41%	33.96%
VIP considering relative angle	12.41	5.409	24.31	18.19%	12.57%

These filed experiments validate that the VIP algorithm can estimate the depth and radius of buried utilities in an acceptable accuracy in both perpendicular-to-utility scan and generic scan. It also proves that ignoring the relative spatial orientation between utility and GPR scan direction will result in a large estimation error.

4.5 Chapter Summary

This section describes the methods utilized in this research for preprocessing and extraction of GPR raw data. The visual quality of GPR scan is improved by appropriate preprocessing. The raw data are extracted for further analysis. A refined model that considers wave velocity, radius and relative orientation of buried utility is proposed in this thesis. It mathematically quantifies the error percentage in estimation for ignoring the relative orientation between buried utility and GPR scan direction. The newly created VIP algorithm is able to retrieve depth and radius information from GPR raw data. A number of field experiments validate its efficiency and effectiveness.

CHAPTER 5. ERROR ASSESSMENT AND MODELING

This chapter describes the assessment and modeling of the positional errors inherent in GPR system. The locational error magnitudes are evaluated via a number of simulated GPR scans in terms of mean value and standard deviation. The output is a quantitative linkage between the error magnitude and its influencing factors, i.e. buried depths and soil conditions. The error magnitudes can be queried using the established quantitative linkage. A 3D probabilistic error band is created in this research to account for the positional errors of underground utilities in GIS. This 3D error band takes the error magnitudes as input to automatically create an enclosed buffer in GIS, which represents the zone containing the true utility line with a certain probability. This newly created 3D probabilistic error band extends the current 2D utility data to 3D and paves the way to an error-aware GIS.

5.1 Error Measurement and Assessment

This section aims to measure the locating errors and examine the error pattern. The positional error consists of systematic errors and random errors. The systematic errors can be remedied via a translation in a 3D coordinate system and thus are not considered herein. It is well accepted that the random errors are normally distributed in x, y and z dimensions. Hence, the error magnitude is characterized by a multivariate normal

distribution. The main task is to estimate the mean value and standard deviation of the multivariate normal distribution upon the assumption that the errors in all three dimensions are independent.

5.1.1 Laboratory Experiments for Error Measurement

A reliable way to measure the locating error magnitude is through a number of field experiments. An experiment process for error measurements is proposed in Figure 5.1. After site preparation, RTKGPS is utilized to obtain the accurate 3D location of the installed pipe. Since RTK GPS can achieve a positional accuracy of 6mm in horizontal and 12 to 18 mm in vertical, hence the locations indicated by RTK GPS are considered as the true location of the buried pipe. Following that, GPR is employed to locate the buried pipe. The localization is conducted for a number of times to get adequate data. The error magnitudes, assumed to be a multivariate normal distribution, can be characterized through comparing the GPR-indicated locations with the ground truth. In addition, the buried depth and soil conditions are changed, and the processes are repeated to obtain the error magnitudes. The expected output of the laboratory experiments is a quantitative linkage between positional error and buried depth and soil conditions.

In practice, the positional error can be queried from the quantitative linkage as it shown in Figure 5.2. The outcome of such error query is the mean values and standard deviations of the error magnitudes in x, y and z dimensions. Such information is then input in phase 3 to model and visualize the positional uncertainty of geospatial, linear underground utilities.

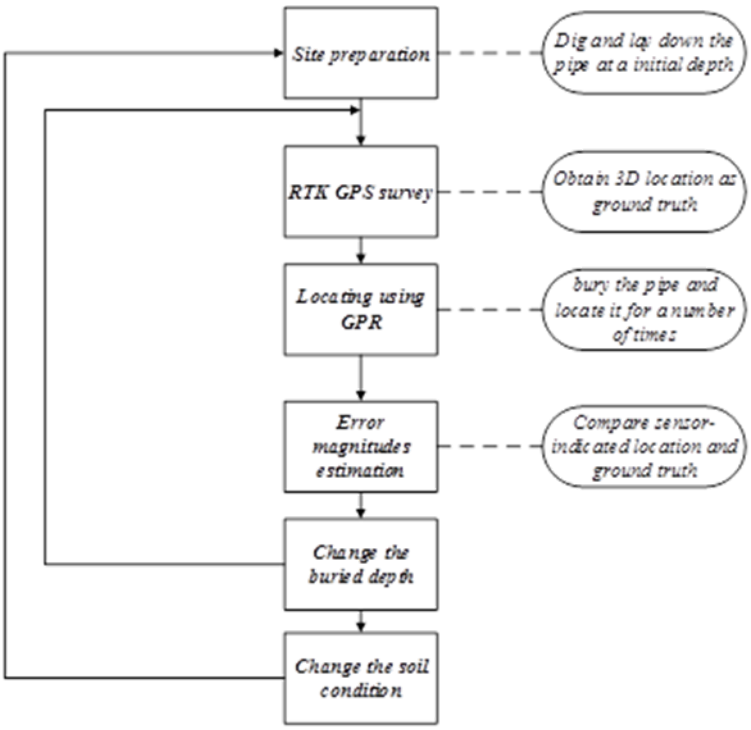


Figure 5.1 General process of laboratory experiments

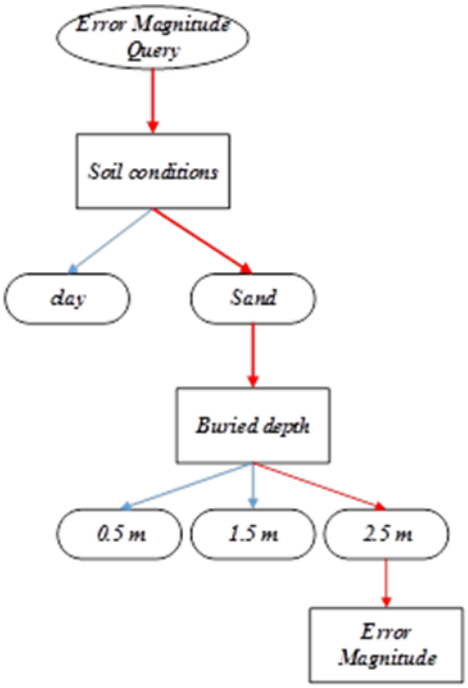


Figure 5.2 Query process for positional error magnitude

5.1.2 Simulated GPR Scans for Error Measurement

Real data are rare and a thorough laboratory experiments are expensive and extensive to be conducted. Hence, in this research the VIP algorithm is applied to a number of simulated GPR scans for error magnitude measurements. GPR scans are produced by GprMax 2.0, a successful simulator for modeling GPR data based on the Finite Difference Time Domain (FDTD) numerical method. The input file of the simulator specifies the parameters for generating GPR scans. The simulator output can be hyperbolic image in binary file (file*b. out) or ASCII file (file *a. out). Figure 5.3 gives an example and provides explanations of the input and output file for GprMax simulation.

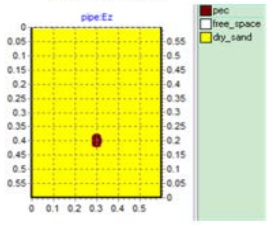
A number of GPR scans are generated by GprMax 2.0 to simulate a 10 cm radius metallic pipe is buried in different depths and soils. The buried depth increases from 0.3 m to 0.8 m with a step of 0.1 m. Tow soil conditions, dry sand and dry clay, are considered in the simulation. The antenna frequency is set to be 800 MHz and the scan direction is perpendicular to the cylindrical pipe. VIP algorithm is applied to estimate the buried depth of the pipe. The mean value and standard deviation of the error magnitude with respect to buried depths and soil conditions are measured and evaluated.

Figure 5.4 shows the error magnitude of locating a metallic pipe in dry sand with various buried depth. Table 5.1 provides the mean values and standard deviations of the error magnitude in various depths.

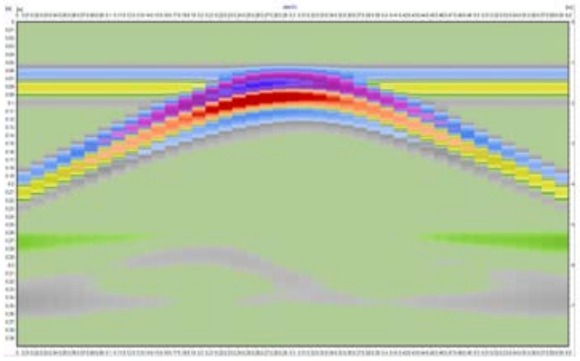
Input file

```
#medium: 3.0 0.0 0.0 0.01 1.0 0.0 dry_sand   Electromagnetic properties of medium
---------------------------------------------------------
#domain: 0.6 0.6          Size in meter of the model
#dx_dy: 0.0025 0.0025     Discretization steps
#time_window: 8.0e-9      Total required simulated time
---------------------------------------------------------
#box: 0.0 0.0 0.6 0.6 dry_sand    Introduce a rectangle of specifc properties
#cylinder: 0.3 0.2 0.025 pec      Introduce a cylinder in the 2D model
---------------------------------------------------------
#line_source: 1.0 900e6 ricker MyLineSource   Excitation parameters
#analysis: 41 pipe.out b
#tx: 0.075 0.2525 MyLineSource 0.0 8.0e-9    Creating scans
#rx: 0.125 0.2525
#tx_steps: 0.01 0.0
#rx_steps: 0.01 0.0
#end_analysis:
---------------------------------------------------------
#title: Model of pipe in sand
#messages: y
#geometry_file: pipe.geo
```

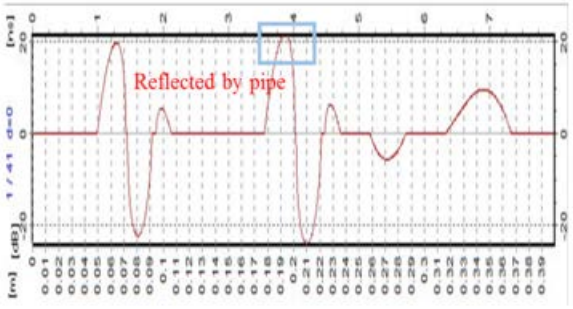
Geometry Profile



Output file



GPR scan plotted by GprProfile 1.2



Time varying amplitude of trace 1 in B scan

Figure 5.3 An example of the GprMax input and output file

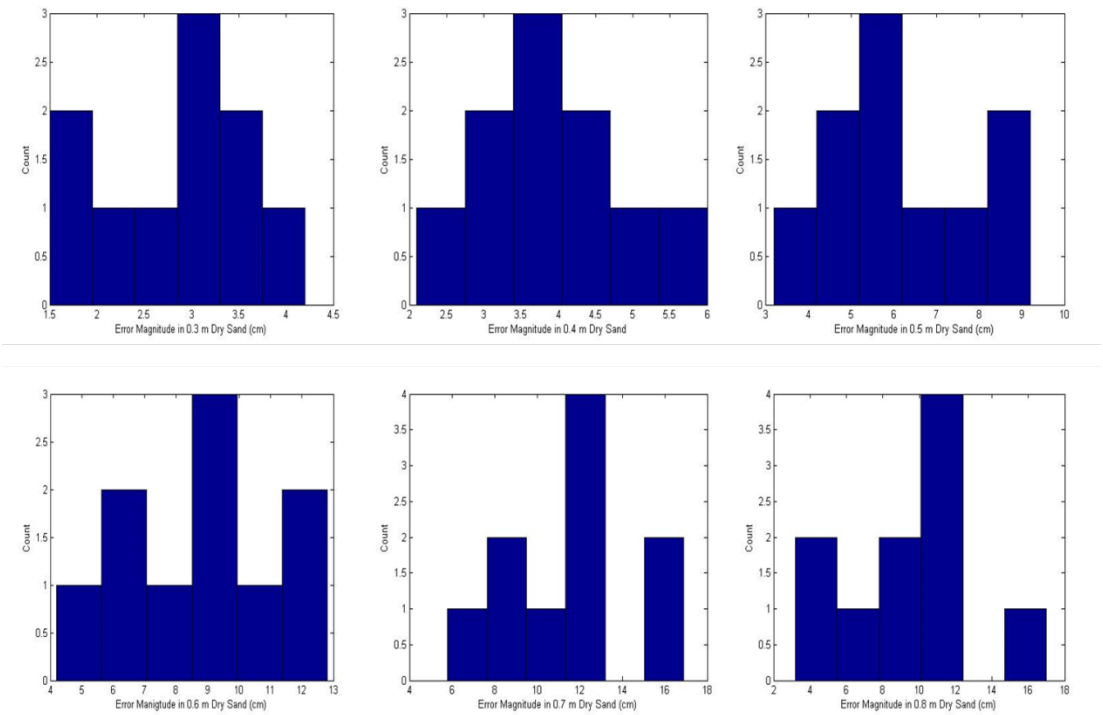


Figure 5.4 Error Magnitude with respect to buried depth in Sand

Table 5.1 Mean and standard deviation of error magnitude in dry sand

Buried depth (m)	0.3	0.4	0.5	0.6	0.7	0.8
Mean of Error magnitude (cm)	2.844	3.970	5.98	8.8	11.49	9.49
Standard deviation of Error magnitude (cm)	0.8928	1.1314	1.8701	2.8636	3.2889	3.945

Figure 5.5 shows the regressions of the mean value and standard deviation versus buried depth. It shows that the mean value and standard deviation of error magnitudes increases linearly with increasing buried depth. Figure 5.6 shows the error magnitude of locating a metallic pipe in dry clay with various buried depths.

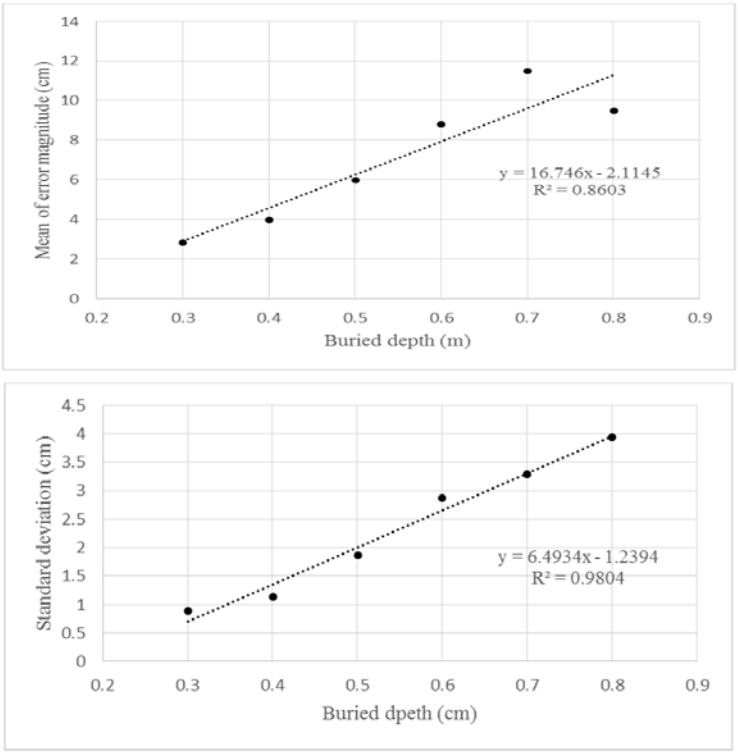


Figure 5.5 Regression of error magnitudes with respect to buried depth in sand

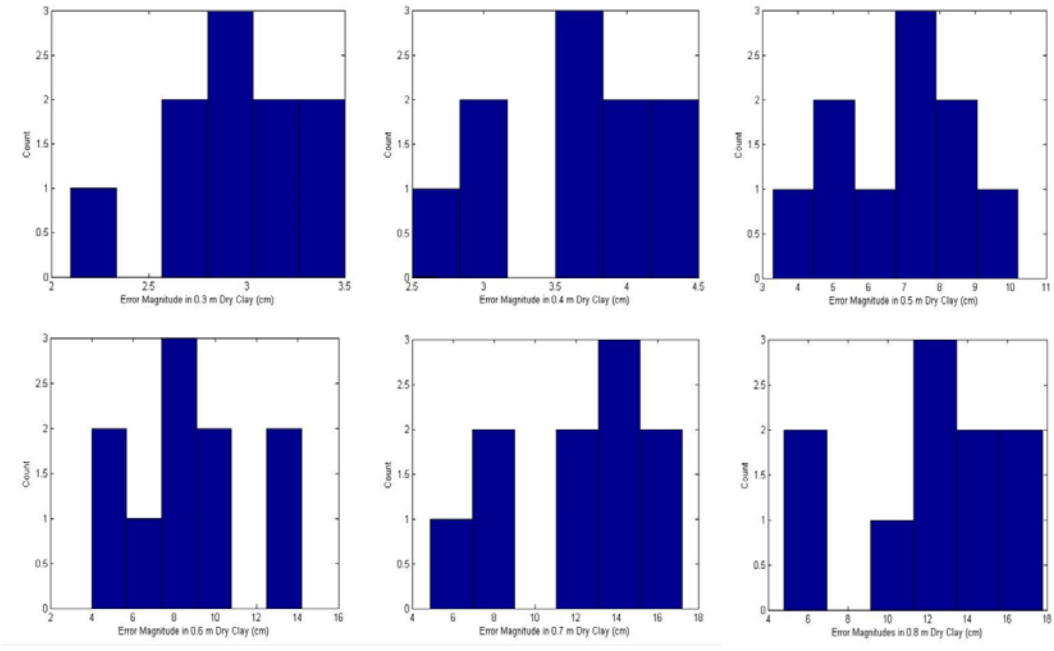


Figure 5.6 Error Magnitude with respect to buried depth in clay

Table 5.2 provides the mean values and standard deviations of the error magnitude in various depths.

Table 5.2 Mean and standard deviation of error magnitude in dry clay

Buried depth (m)	0.3	0.4	0.5	0.6	0.7	0.8
Mean of Error magnitude (cm)	2.965	3.645	6.85	8.77	12.3	12.2
Standard deviation of Error magnitude (cm)	0.3958	0.6265	2.0266	3.2565	4.0233	4.1721

Figure 5.7 shows the regressions of the mean value and standard deviation versus buried depth.

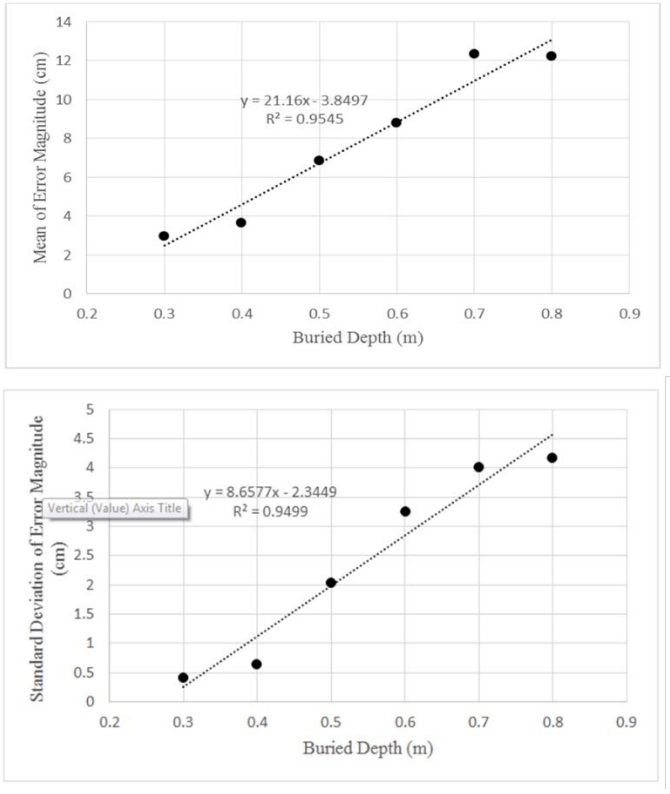


Figure 5.7 Regression of error magnitudes with respect to buried depth in clay

Similar pattern of GPR locating errors is found in both dry clay and sand. The regression models serve as the quantitative linkages for error magnitude query. From the data nature, i.e. buried depth and soil conditions (the soil conditions can be roughly estimated in the site), the mean value and error magnitude can be queried from the regression models.

The horizontal error magnitudes, i.e. error magnitudes in x and y dimension, is assumed to solely depend on GPS accuracy. For the vertical error magnitude, i.e. error magnitude in z dimension, a worst scenario is considered in error propagation between GPR and GPS. For instance, the vertical error magnitude for RTK GPS is 0.018 m and the buried depth error produced by GPR is 0.1 m. The error magnitude for the integrated system is assumed to be 0.118 m. Therefore, the error magnitude with its mean value and standard deviation can be estimated from the data collected by GPR and GPS. The error magnitude will be input into phase 3 for error modeling and rendering.

5.2 Positional Error Modeling for Underground Utilities

The underground utility lines were modeled as 3D straight-line segments connected at surveying points. The positional uncertainties of the underground utilities, expressed as 3D probabilistic uncertainty bands, were derived from the utility locations and their positional errors at certain surveying points. This strategy is reasonable since utility mapping is often conducted in discrete points, and these points totally determine the line segment. This section first illustrates the 3D error model designed for the geospatial utility points. Thereafter, it formulates the 3D probabilistic error model for a utility line segment and describes its visualization mechanism in GIS.

5.2.1 Positional Error Model for Points

Positional errors consist of systematic errors and random errors. Systematic errors were not considered in this research; and the random errors e_x, e_y, and e_z in the x, y, and z dimensions were assumed to be normally distributed and independent of each other. Hence, the random error e of a specific point can be described as a non-singular multivariate normal distribution (i.e., $e \sim \mathcal{N}(\mu, \Sigma)$), where $\mu = [E[e_x], E[e_y], E[e_z]]$; $\Sigma = [Cov[e_i, e_j]]$, i=x, y, z; j=x, y, z. The geometry or equidensity contours for a non-singular multivariate normal distribution are ellipsoids, which are shown in Figure 5.8 (Hansen 2011, Greenwalt and Shultz 1968). The ellipsoid can be formulated as Equation (5.1) and follows a χ^2 distribution with three degrees of freedom.

$$(\frac{x}{\sigma_x})^2 + (\frac{y}{\sigma_y})^2 + (\frac{z}{\sigma_z})^2 = c^2 \tag{5.1}$$

Where σ_x, σ_y and σ_z are the standard deviation of the error magnitudes in the x, y, and z dimensions. Its size is determined by the error magnitude and the probability of containing the true point

For an ellipsoid that specifies a 3D geospatial extent (e.g., given three semi-principal axes of $c\sigma_x$, $c\sigma_y$ and $c\sigma_z$), the probability of containing the actual point can be calculated via direct integration of the probability density functions or by referring to χ^2 distribution tables. This process can be reversed to derive the size of the ellipsoid that corresponds to

a specific probability. For instance, a 90% probability (e.g., $P=[\mu \le c^2]=0.9$) corresponds to 6.251 in χ^2 table, meaning $c^2=6.25$ and c=2.5. Therefore, the size of the ellipsoid was determined by three semi-principal axes at the size of 2.5 $(\sigma_x, \sigma_y, \sigma_z)$

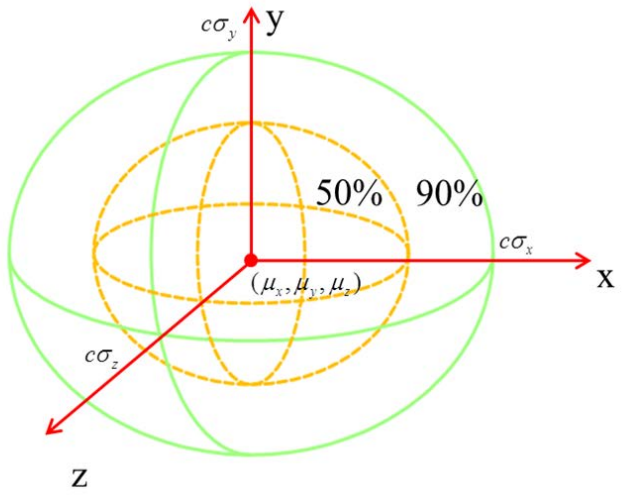


Figure 5.8 Positional error model for point

5.2.2 Positional Error Model for Line Segment

5.2.2.1 General Concept

In this thesis, utility lines are modeled as straight line segments that connect two end points. We model the positional error of such line segment as a 3D probabilistic error band, which encloses the true location of the utility line at a particular probability. The 3D error band corresponding to a specific probability will enclose and be tangent to all the error ellipsoids of the points on the line with the same probability. Figure 5.9 clearly illustrate this concept with 90% and 50% error bands. The probability 90% means the probability of containing the true line in that error band is 0.9.

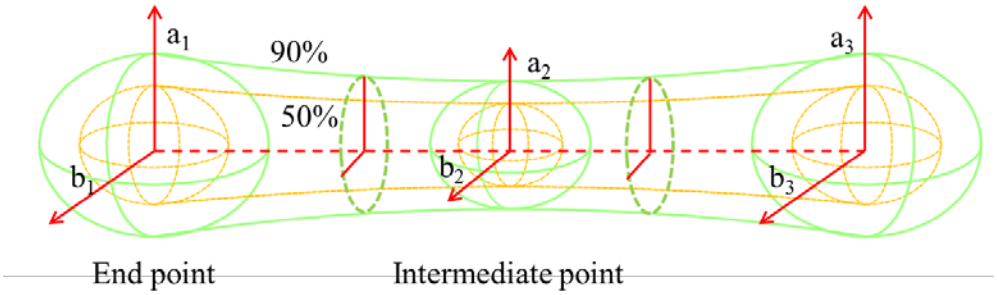


Figure 5.9 Formation of 3D probabilistic error band

It has been noted that every section of the 3D error band is an ellipse with various semi-axes (a, b). Therefore, the shape and size of the band can be described via a mathematical function, see Equation 5.2.

$$(a,b)_{x,p\%} = f(x, p\%, e) \tag{5.2}$$

Where $(a,b)_{x,p\%}$ represents the semi-axes of the ellipse that corresponds to p% probability at location x along the centerline of the utility, and e is the error magnitudes that can be queried as described in section 5.1. Hence, the remaining problem is to derive the semi-axes of the enveloping ellipses that correspond to a specific probability. There are two characteristics of these enveloping ellipses: 1) the ellipses are orthogonal to the line segments and 2) the ellipses are the largest ones along the line segments. The following parts illustrate the derivation of the semi-axes of the enveloping ellipses.

5.2.2.2 Description of line segments

A line segment P_0P_1 is determined by the two endpoints P_0 and P_1. P_t is an arbitrary point on the line segment P_0P_1. The geometry and size of the error ellipsoid of P_t is determined

by its position and the error magnitudes of the endpoints. An assumption adopted in this model is that the two endpoints P_0 and P_1 are independent. This assumption is reasonable since the measurements in utility survey are independent processes. The two end points are represented by the vector $P_0 = (X_{0,} Y_{0,} Z_0)$ and $P_1 = (X_{1,} Y_{1,} Z_1)$ (see Equation 5.3).

$$\begin{aligned} X_i &= X_i^* + \mu_{i,x} \\ Y_i &= Y_i^* + \mu_{i,y} \\ Z_i &= Z_i^* + \mu_{i,z} \end{aligned} \tag{5.3}$$

Where (X_i^*, Y_i^*, Z_i^*) is the record-indicated location of i[th] points, and $(\mu_{i,x}, \mu_{i,y}, \mu_{i,z})$ is the mean value of the error magnitudes of i[th] points, i=0, 1 (See Figure 5.10). Therefore the positions of endpoints follow a multivariate normal distribution with mean value (X_i, Y_i, Z_i) and standard deviation $(\sigma_x, \sigma_y, \sigma_z)$. $(\sigma_x, \sigma_y, \sigma_z)$ is the standard deviation of error magnitudes in x, y and z dimensions. The coordinates of an arbitrary point on the line segment can be represented as Equation 5.4 (Shi and Liu, 2000).

$$\begin{aligned} X_t &= (1-t)X_0 + tX_1 \\ Y_t &= (1-t)Y_0 + tY_1 \\ Z_t &= (1-t)Z_0 + tZ_1 \end{aligned} \tag{5.4}$$

X_t, Y_t, and Z_t are linear transformations of normally distributed variables, hence, they also follow a normal distribution (Mikhail and Ackermann 1976). The mean values are calculated by Equation 5.4. The standard deviations $\sigma_{t,i}$ are computed by Equation 5.5.

$$\begin{aligned} &\sigma_{t,i} = \sqrt{(1-t)^2 \times {\sigma_{0,i}}^2 + t^2 \times {\sigma_{1,i}}^2} \\ &i = X, Y, Z \end{aligned} \tag{5.5}$$

$\sigma_{0,i}$ and $\sigma_{1,i}$ are the standard deviation of error magnitudes of two endpoints. Hence, the position and error ellipsoid of intermediate points on the line are determined based on the position and error magnitudes of two endpoints.

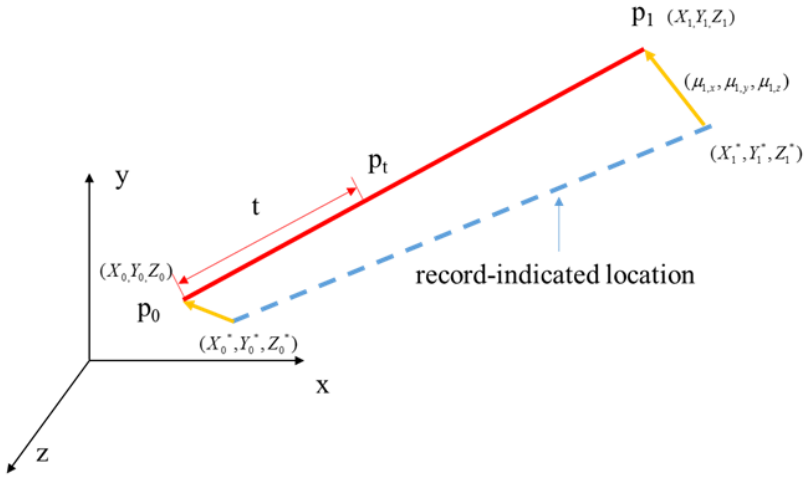


Figure 5.10 Description of line segments

5.2.2.3 Ellipses orthogonal to the line

Derivation of the ellipses that are orthogonal to a line segment is a vital step in building the 3D error band. In Figure 5.11, the ellipse in red passes through the center of the ellipsoid and is orthogonal to the line segment. The shaded volume of the ellipsoid will be enclosed by the enveloping ellipses along the line. Two coordinate transformations are performed to calculate the semi-axes.

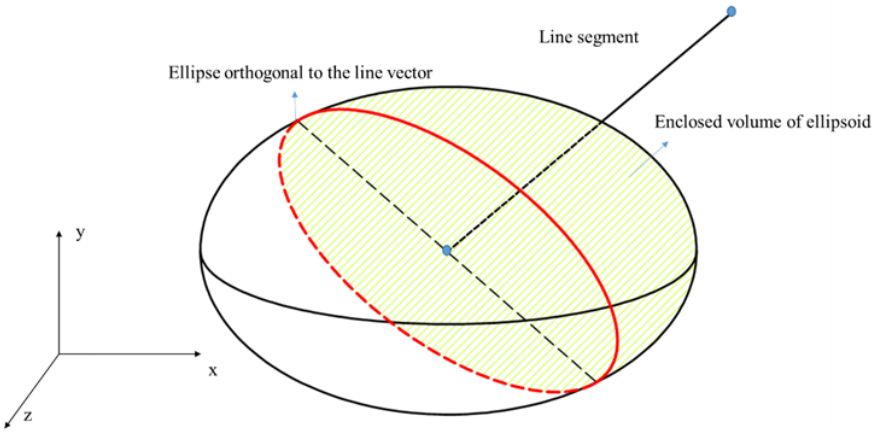


Figure 5.11 Ellipse orthogonal to the line vectors

The first transformation is to rotate the original coordinate system. The vector of the line segment is (i, j, k). Let (i, j, k) be the direction numbers of z' axis in the new x'y'z' coordinate system. Let $(-k, 0, i)$ be the direction number of the x' axis in the x'y'z' coordinate system, which makes the x' axis perpendicular to the z' axis. The y' axis is perpendicular to the x'z' plane and the direction number of the y' axis thus is $(ij, -(i^2 + k^2), jk)$. The direction numbers are normalized to direction cosines, see Equation 5.6-5.8.

$$\begin{aligned} \lambda_1 &= \frac{-k}{\sqrt{i^2 + k^2}} \\ \mu_1 &= 0 \\ \eta_1 &= \frac{i}{\sqrt{i^2 + k^2}} \end{aligned} \tag{5.6}$$

$$\lambda_2 = \frac{ij}{\sqrt{i^2+j^2+k^2}\times\sqrt{i^2+k^2}}$$
$$\mu_2 = \frac{-(i^2+k^2)}{\sqrt{i^2+j^2+k^2}\times\sqrt{i^2+k^2}} \tag{5.7}$$
$$\eta_2 = \frac{jk}{\sqrt{i^2+j^2+k^2}\times\sqrt{i^2+k^2}}$$

$$\lambda_3 = \frac{i}{\sqrt{i^2+j^2+k^2}}$$
$$\mu_3 = \frac{j}{\sqrt{i^2+j^2+k^2}} \tag{5.8}$$
$$\eta_3 = \frac{k}{\sqrt{i^2+j^2+k^2}}$$

Taking a matrix format, the origin coordinate system are rotated, see Equation 5.9.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \lambda_1 & \lambda_2 & \lambda_3 \\ \mu_1 & \mu_2 & \mu_3 \\ \eta_1 & \eta_2 & \eta_3 \end{bmatrix} \times \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} \tag{5.9}$$

The equation of the ellipsoid is transformed to the new x'y'z' system. The intersection of plane Z'=0, and the ellipsoid results in the ellipse that is orthogonal to the line vector. The substitution of Z'=0 in the ellipsoid equation yields the equation of the ellipse in a general form (see Equation (5.10)).

$$AX'^2 + 2BX'Y' + CY'^2 + 2DX' + 2EY' + F = 0 \tag{5.10}$$

where A, B, C, D, E and F are constant values. In order to derive the canonical ellipse equation, another coordinate transform is needed to remove the X'Y' term, see Equation 5.11.

$$\begin{bmatrix} X' \\ Y' \end{bmatrix} = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix} \times \begin{bmatrix} X'' \\ Y'' \end{bmatrix} \tag{5.11}$$

Where t is computed as Equation 5.12.

$$t = \frac{\arctan(2B/(A-C))}{2} \tag{5.12}$$

This coordinate transformation results in the canonical ellipse equation:

$$\frac{(x''-h)}{a^2} + \frac{(y''-k)}{b^2} = 1 \tag{5.13}$$

Where a, b, h and k are constant values. The semi-axes (a, b) can be directly extracted from the Equation 5.13.

5.2.2.4 The largest ellipses along the line

This section illustrates the determination of the enveloping ellipses along the line. Firstly, an illusion will be clarified as shown in Figure 5.12. The enveloping ellipses at a specific point (i.e., P_2 in Figure 5.12) may not always be the one that passes through the center of an error ellipsoid of that point (i.e., green ellipse in dash line in Figure 5.12). Unawareness of this illusion will result in the exclusion of some positional errors.

To derive the enveloping ellipses along a line segment, for easy illustration, the line segment is assumed to be the x axis of the coordinate system. In Figure 5.13, P_0 is one endpoint of the line segment and $P_{t1}$and P_{t2} are two intermediate points on the line.

Ellipse 1 (y-z plane) and Ellipse 2 (x-y) plane are the central ellipses of the error ellipsoids at P_{t1} and P_{t2}, respectively.

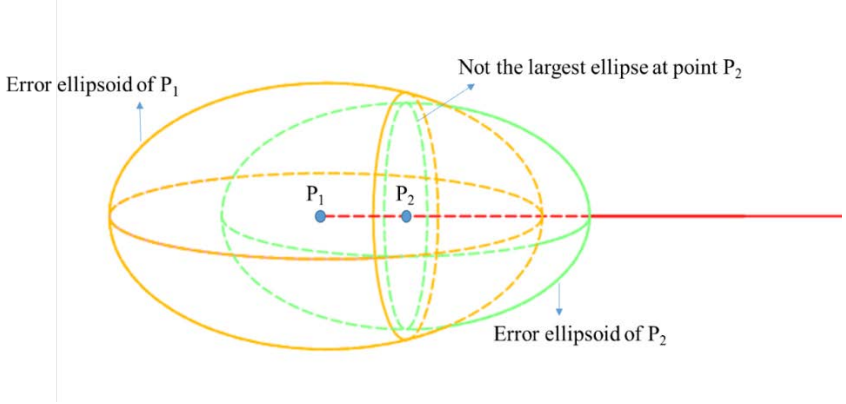


Figure 5.12 An illusion in determination of the enveloping ellipses

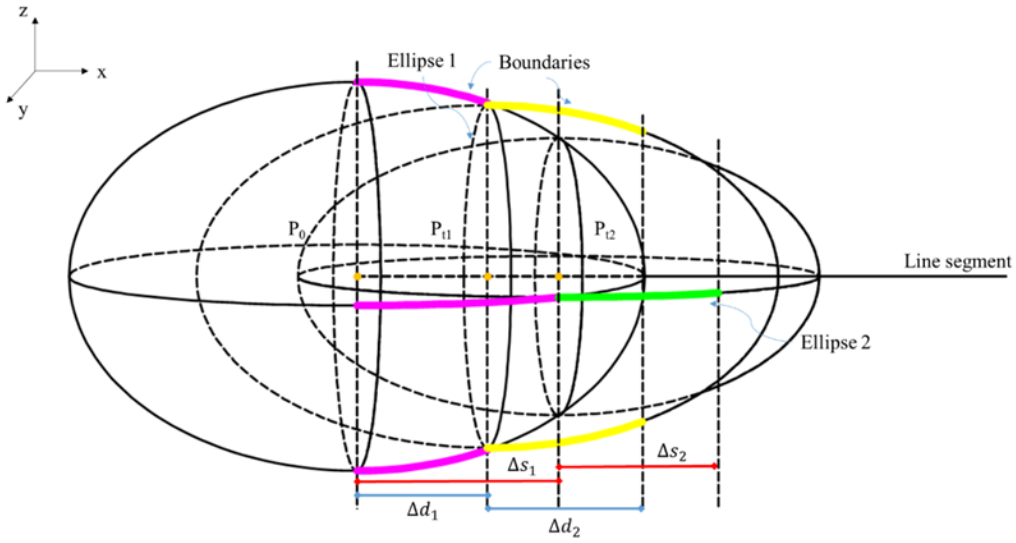


Figure 5.13 Derivation of enveloping ellipse along the line segment

On the basis of previous analysis, the equation of ellipse 1 can be written as Equation (5.14).

$$\frac{y^2}{\sigma_{t_1,y}^2}+\frac{z^2}{\sigma_{t_1,z}^2}=1 \tag{5.14}$$

Its semi-axes can be computed by Equation (5.15).

$$\begin{aligned}\sigma_{t_1,y}^2&=(1-\frac{\Delta d_1}{d})^2\sigma_{0,y}^2+(\frac{\Delta d_1}{d})^2\sigma_{1,y}^2\\ \sigma_{t_1,z}^2&=(1-\frac{\Delta d_1}{d})^2\sigma_{0,z}^2+(\frac{\Delta d_1}{d})^2\sigma_{1,z}^2\end{aligned} \tag{5.15}$$

Where d is the distance between two endpoints P_0 and P_1, Δd_1 is the distance between P_0 and P_{t1} (see Figure 5.13). $\sigma_{i,j}$ (i=0, 1, t_1, t_2 j=x, y, z) is the standard deviation of the error magnitude of point P_i in j dimension. Substitution of $x=\Delta d_1$ in the ellipsoid equation of P_0 results in Equation (5.16).

$$\frac{y^2}{\sigma_{0,y}^2(1-\frac{\Delta d_1^2}{\sigma_{0,x}^2})}+\frac{z^2}{\sigma_{0,z}^2(1-\frac{\Delta d_1^2}{\sigma_{0,x}^2})}=1 \tag{5.16}$$

Letting z semi-axis be equal to that of ellipse 1 leads to Equation 5.17.

$$\sigma_{0,z}^2(1-\frac{\Delta d_1^2}{\sigma_{0,x}^2})=(1-\frac{\Delta d_1}{d})^2\sigma_{0,z}^2+(\frac{\Delta d_1}{d})^2\sigma_{1,z}^2 \tag{5.17}$$

Δd_1 is the only unknown variable in Equation (5.17) and thus can be calculated easily. Likewise, Δs_1, the distance between P_0 and P_{t2}, can be computed by solving Equation (5.18)

$$\sigma_{0,y}^2(1-\frac{\Delta s_1^2}{\sigma_{0,x}^2})=(1-\frac{\Delta s_1}{d})^2\sigma_{0,y}^2+(\frac{\Delta s_1}{d})^2\sigma_{1,y}^2 \tag{5.18}$$

Let the distance away from P_0 be Δt. The semi-axes of the enveloping ellipse are derived based on the relationship between Δd_1, Δs_1, and Δt.

1) If $\Delta t < \Delta d_1$, the semi-axes in the y and z dimensions are calculated by substituting $x = \Delta t$ into the ellipsoid equation of P_0. The red curves in Figure 14 indicate the boundaries in this scenario.
2) If $\Delta d_1 < \Delta t < \Delta s_1$, the semi-axis in z dimension is calculated by substituting $x = \Delta t - \Delta d_1$ and y=0 into the ellipsoid equation of P_{t1}. The yellow curve in Figure 14 indicates the boundary in z dimension. The semi-axis in the y dimension is still calculated by substituting $x = \Delta t$ and $z = 0$ into the ellipsoid equation of P_0.
3) If $\Delta t > \Delta s_1$, the semi-axis in the z dimension is calculated by substituting $x = \Delta t - \Delta d_1$ and $y = 0$ into the ellipsoid equation of P_{t1}. The semi-axis in the y dimension is calculated by substituting $x = \Delta t - \Delta s_1$ and z=0 into the ellipsoid equation of P_{t2}. The green curve in Figure 14 indicates the boundary in the y dimension.

The following critical distances Δd_i and Δs_i ($i = 1, 2, ..., n$) can be computed by repeating the process described above. Within each critical distance, the boundaries can be fully determined as was previously proven. As a number of enveloping ellipses along the line segment can be determined, the 3D error band is formulated by connecting these ellipses to form an enclosed 3D surface. Figure 5.14 graphically illustrate this concept.

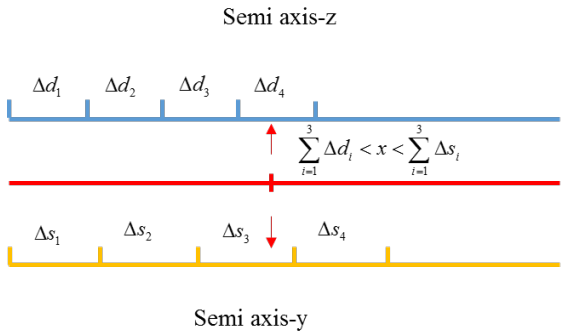


Figure 5.14 Determination of ellipse boundaries

As a number of enveloping ellipses along the line segment can be determined, the 3D error band is formulated by connecting these ellipses to form an enclosed 3D surface. The following part will illustrate how to render this error band in GIS using multipatch.

5.2.3 Error Model Rendering in GIS

The positional error model for linear, geospatial underground utilities is rendered in GIS using multipatch. The enveloping ellipses are drawn around the points along the line segment, and then are connected by triangular strips (see Figure 5.15).

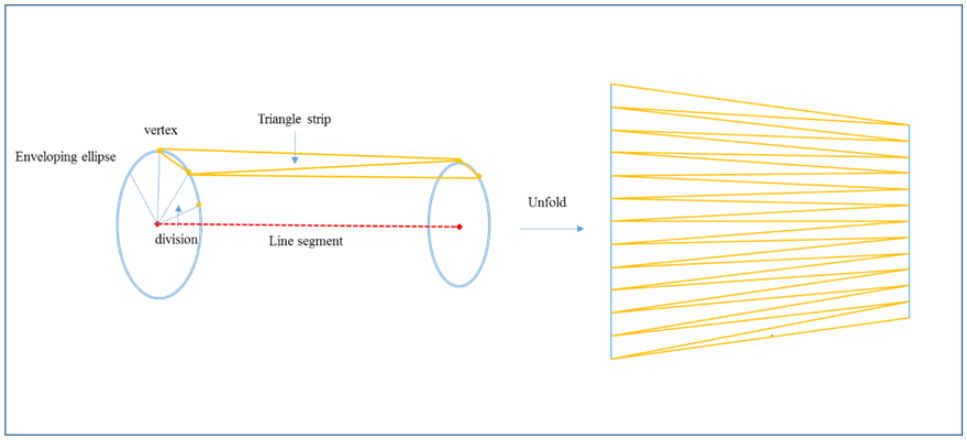


Figure 5.15 Error band formed by triangle strips

An add-in was created in Arc Scene using Arc Object to render the 3D error band, which takes the location and the associated positional error magnitudes of the two endpoints as input to construct the 3D error band. The steps were as follows.

1) Determine the center and semi-axes of the enveloping error ellipses along the line as described in previous sections.
2) Define the layer, feature class (point class, polyline class, and multipatch class) in the Arc Scene environment and set the division numbers. The ellipse will be more accurately constructed if more divisions are applied.
3) Take two points and construct the vertexes of their enveloping ellipses (see Figure 5.16). First, the normal vector of the line segment is obtained. Second, the magnitude of the normal vector is calculated based on the semi-axes and a specific rotation angle α using Equation (5.19)

$$r = \frac{ab}{\sqrt{a^2 \sin^2 \alpha + b^2 \sin^2 \alpha}} \tag{5.19}$$

Where r is the length of the normal vector, a and b are the semi-axes of the enveloping ellipse. α is defined as the angle between the normal vector and the y axis. Third, a vertex is constructed whose y and z coordinates are the sum of the y and z coordinates of the center and the y and z components of the normal vector. The next vertex on the ellipse is constructed following the same procedure as the angle α increases by a division degree.

4) Construct the triangular strips with the vertexes and add them to the multipatch geometry. Hence, the 3D error band between the first two points is constructed.
5) Take the next two points and repeat the previous process until the endpoint of the line

segment is reached. Figure 5.17 summarizes the algorithm for creating 3D error band in GIS.

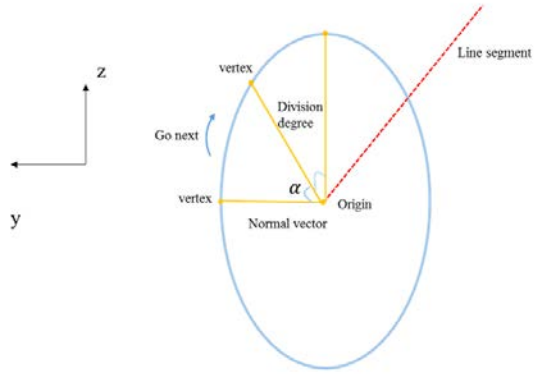


Figure 5.16 Vertex construction for enveloping ellipse

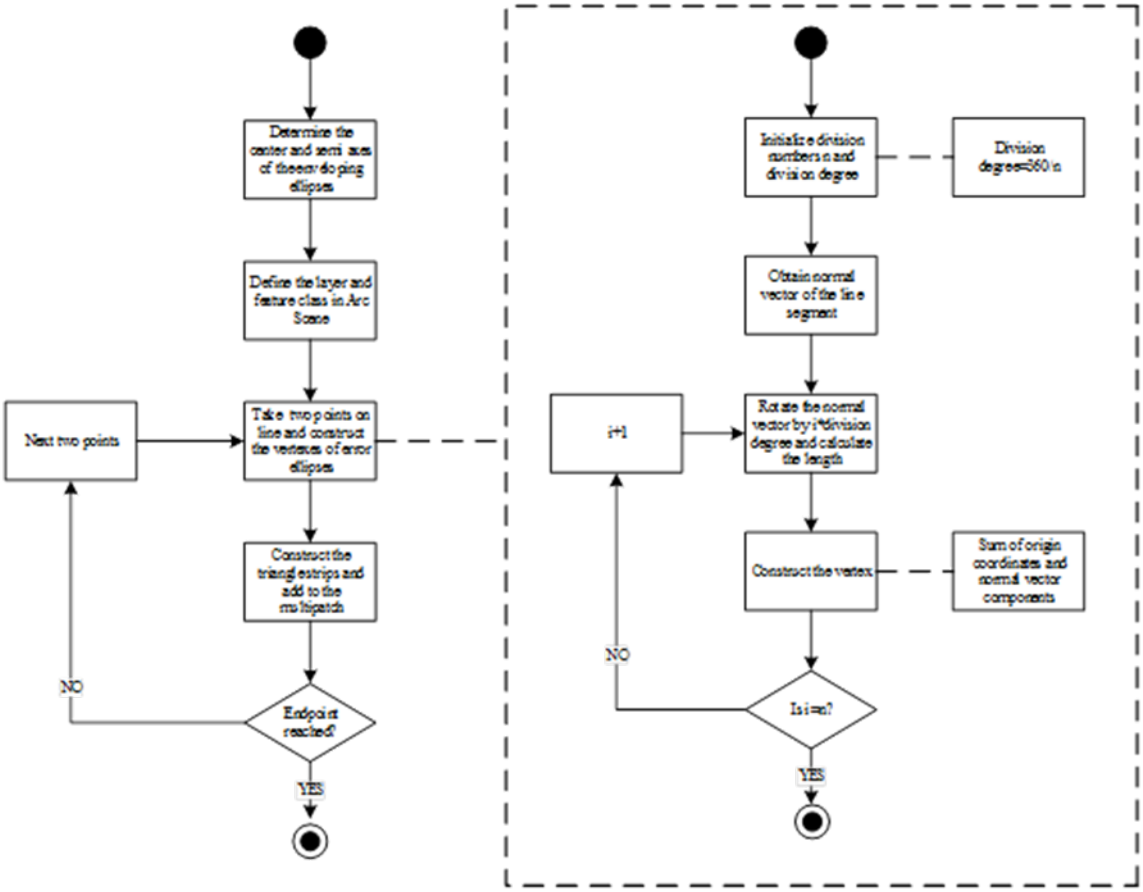


Figure 5.17 Algorithm for creating 3D error band in Arc Scene

Following the procedure stated before, a 3D error band that corresponds to a specific probability can be visualized in GIS. An example of the created 3D error band in GIS is shown in Figure 5.18, which validate the approach stated before.

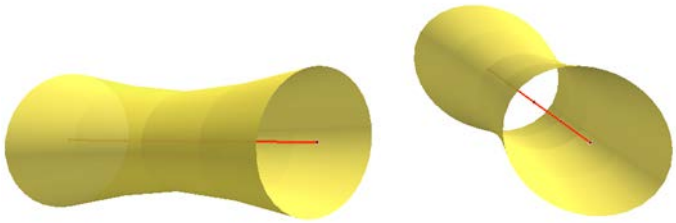


Figure 5.18 An example of created 3D error band in GIS

5.3 Chapter Summary

This section evaluates the locating error magnitude of the GPR system via a number of simulated GPR scans. The error magnitudes are assessed in different buried depth and soil types. A quantitative linkage between the error magnitude and buried depth and soil type is established. The error magnitudes serve as the input for error modeling in GIS. A 3D probabilistic error model in GIS is created for linear, geospatial underground utilities. This 3D error model can be implemented in GIS environment, which extends the 2D utility data to 3D and makes the current GIS more intelligent.

CHAPTER 6. CONCLUSIONS AND RECOMMENDATIONS

This chapter summarizes the research findings and conclusions. The research limitations are discussed and future research directions are pointed up.

6.1 Summary and Conclusions

The information regarding the location and dimension of many underground utilities have not always been collected and properly documented, leading to detrimental consequences such as utility conflicts and utility strikes throughout all the life cycle stages of building and civil infrastructures. This is attributed to two root causes. First, there is a lack of an effective and efficient system to collect the location and dimension information in a reliable and responsive manner. Second, the positional uncertainty or error that are inherent in the collected data has not been considered and modeled. This research aims to address these problems by creating an error-aware and geospatial system to map and characterize the buried underground utilities.

In this research, ground penetrating radar (GPR), global positioning system (GPS) and geographical information system (GIS) are integrated to form a total 3G system. It was found that the GPR, GPS and GIS can be integrated seamlessly to detect, locate and characterize the buried utilities, geo-register the detected utility to real world coordinate

system, and visualize the underground utilities with the associated positional uncertainties.

This research also creates a novel very important point (VIP) algorithm to measure the buried depth and radius of cylindrical underground utilities from GPR and GPS raw data. This algorithm is expected to eliminate two impractical constraints of GPR applications in underground utility mapping in current practice, i.e., 1) the requirement of moving GPR perpendicularly across the buried utility to scan it, and 2) the requirement of knowing either the EM wave velocity or the range of the radius of the underground utility.

The visual quality of original GPR scan is enhanced by a series of image processing stages to facilitate utility detection and raw data extraction. To represent the raw data points, a generic hyperbola equation is proposed that considers the relative angle between the buried utility and GPR scan direction. Based on the generic hyperbola equation, quantitative analyses of errors in estimating velocity and radius are carried out. The quantitative analyses reveal that, if ignoring the relative angle, the error in estimate for velocity increases with decreasing relative angle. The error will be further exaggerated if the utility is buried shallowly and has a large radius. The velocity derived assuming a perpendicular-to-utility scan is always artificially magnified, and thus the utility is assumed deeper than its real world position. This may cause utility strikes in in utility expansion and/or maintenance. The error in estimates for radius also increases with decreasing relative angles. The error analyses confirm the necessity of considering the relative angle in estimations.

To simultaneously estimate the radius and burying depth of underground utilities from GPR and GPS raw data, a novel very important point (VIP) algorithm is created. Two nonlinear equations regarding the two unknown variables, i.e., EM wave velocity and utility radius, are formulated and solved at the VIP that is detected by the boundary of GPR footprint. The filed experiments validate that this newly created algorithm can effectively estimate burying depth and radius of underground utilities from GPR raw data in various settings. It is also worth emphasizing that the estimation accuracy is greatly improved (e.g., the error percentages of depth and radius are reduced from 33.96% and 24.41% to 12.57% and 18.19% respectively in a generic scan) when taking the relative angle into consideration and applying the correction process. The RTK GPS is found to be effective to assist the interpretation of GPR raw data and extend the VIP algorithm to a more generic scenario.

To evaluate and model the positional uncertainties/errors of this system, a prototype of 3D probabilistic error band is proposed in this study. The positional errors of the GPR system is first assessed through a number of simulated GPR scans in different settings. The error magnitudes are assessed in terms of mean value and standard deviation in different depth and soil conditions. It was found that in dry sand and dry clay, the mean value and standard deviation of the error magnitudes increase linearly with the buried depth. Quantitative linkages between the error magnitude and buried depth in both sand and clay are established. Therefore, the error magnitude can be queried based on the quantitative linkages and the buried depth derived from GPR raw data.

A 3D probabilistic error band model is created in this study to handle the positional uncertainties inherent in underground utility location data. The shape and size of the uncertainty band were found to depend on the positional errors of the utility points, the probability of containing the true utility locations, and the distance/interval between the surveying points. With this novel 3D probabilistic error band model, the positional accuracy of underground utility data can be described by continuous probability distributions and quantified as 3D geospatial extents with associated probabilities of containing the true location of utilities. Furthermore, a multipatch surface model was found to be effective in rendering the 3D error bands in an easy-to-understand manner in GIS. Its visualization in GIS clearly conveys the positional uncertainty synergistically to downstream applications. On-going research is integrating this newly created 3D probabilistic uncertainty band model with augmented reality and machine control to protect the underground infrastructure by preventing unintended excavation damage, which potentially could transform excavator operations from a primarily skill-based activity to one that is knowledge-based.

To sum up, with all the system hardware, i.e., GPR, GPS and GIS, and software, i.e., the VIP algorithm, 3D probabilistic error band in GIS, the newly created system is effective and efficient in detecting, locating and characterizing underground utilities in an information-rich and error-aware manner.

6.2 Limitations and Future Researches

The limitations and future research directions are discussed in the following aspects.

(1) The interpretation of GPR data is not completely automated. Given the large amount of GPR raw data points, it is very time-consuming, tedious and error-prone to manually extract every raw data points from the GPR scans. Therefore, future research will be devoted to create effective and efficient algorithms to automatically recognize hyperbolae in GPR scans and extract the GPR raw data points.

(2) The limitations of the VIP algorithm are twofold. First, it requires a relative high quality of GPR scan that contains a relative large amount of raw data points. Sparse raw data may lead to failure of finding VIP. Second, heterogeneous or layered soils will lead to scattering of EM waves and thus deviations of VIP. Hence, the performance of the VIP algorithm will be reduced. Future researches will focus on improving the robustness and performance of VIP algorithm.

(3) To accurately evaluate the relationship between positional errors of the system and its influencing factors, other parameters such as the pipe radius and materials may need to be considered. Furthermore, instead of using simulated data, filed experiments should be performed to assess the positional uncertainties. Thus, a number of field experiments will be performed to assess the positional errors.

(4) The prototype of the 3D probabilistic error band is only a visual representation of the location and positional uncertainties of the buried utilities. It can be used, for example, to determine the digging zone of excavations activities. However, more analytical capabilities, for example, spatial proximity analysis, are needed. Therefore, future

research will integrate other technologies (e.g., augmented reality) with this 3D buffer to monitor excavation activities in the field.

LIST OF REFERENCES

AI-Nuaimy, W., Huang, Y., Nakhkash, M., Fang, M., Nguyen, V. and Eriksen, A. (2000). "Automatic Detection of Buried Utilities and Solid Objects Using Neural Networks and Pattern Recognition." Journal of Applied Geophysics, 43(2), 157-165.

American Society of Civil Engineering (ASCE). (2002). "Standard Guideline for the Collection and Depiction of Existing Subsurface Utility Data." ASCE Code and Standards Activity Committee (CSAC), New York, US, pp. 4-6.

Annan, A. P (2009). "Electromagnetic Principles of Ground Penetrating Radar." In Ground Penetrating Raear: Theory and Applications, H. M. Jol, Ed., Chapter 1, Elsevier, Amsterdam, The Netherlands, 2009.

Arnott, D. and Keddie, A. (1992). "Data Capture – The Standards and Procedures Utilized within Northumbrian Water Group." In Proceedings of the Association for Geographical Information Conference, pp. 2.12.1-2.12.6.

Borgioli, G., Capineri, L., Falorni, P., Matucci, S. and Windsor, C. (2008). "The Detection of Buried Pipes from Time-of-flight Radar Data." IEEE Transactions on Geoscience and Remote Sensing, 46(2008), 2254-2266.

Cai, H. (2004). "Accuracy Evaluation of a 3-D Spatial Modeling Approach to Model Linear Objects and Predict Their Lengths." PhD Dissertation, North Carolina State University, North Carolina.

Caspary, W. and Scheuring, R. (1993). "Positional Accuracy in Spatial Databases." Computers, Environment and Urban System, 17(2), 103-110.

Cassidy, N. J. (2009). "Ground Penetrating Radar Data Processing, Modeling and Analysis." In Ground Penetrating Radar: Theory and Applications, H. M. Jol, Ed., Chapter 1, Elsevier, Amsterdam, The Netherlands, 2009.

Chapman, M. A., Alesheikh, A., Karimi, H. (2003). "Error Modeling and Management for Data in Geospatial Information Systems." In: Green, D. R., King, S. D. (Eds.), Coastal and Marine Geo-Information Systems: Applying the Technology to the

Environment. Kluwer Academic Publishers, Dordrecht, the Netherlands, pp. 355-369.

Chasey, A. D. and Cowan, R. (2008). "Integration of GIS and GPS to Improve Utility Mapping and Damage Prevention." FIATECH Annual Technology Conference and Showcase, New Orleans, LA, March 31-April 2.

Chen, H. and Cohn, A. G. (2010a). "Probabilistic Conic Mixture Model and Its Application to Mining Spatial Ground Penetrating Radar Data." In Workshops of SIAM Conference on Data Mining (WSDM 10).

Chen, H. and Cohn, A. G. (2010b). "Probabilistic Robust Hyperbola Mixture Model for Interpreting Ground Penetrating Radar Data." In Neural Networks (IJCNN), The 2010 International Joint Conference, 1(8), 18-23.

Common Ground Alliance (CGA). (2008). "Call Before You Dig." Retrieved May 28, 2013, from http://www.call811.com/for-the-media/once-per-minute.aspx

Conyers, L. B. "Variables Affecting GPR Resolution and Energy Penetration." Retrieved July 4, 2013, from http://mysite.du.edu/~lconyers/SERDP/variablesaffecting2.htm

Costello, S. B., Chapman, D. N., Rogers, C. D. F. and Metje, N. (2007). "Underground Asset Location and Condition Assessment Technologies." Journal of Tunnelling and Underground Space Technology, 22(2007), 524-542.

Crawford, D. (2012). "Implementing a Utility Geographic Information System for Water, Sewer, and Electric." Master Thesis, Northwest Missouri State University, 2012.

Cullen, M. (2005). "Use of A common Framework for Positional Referencing of Buried Assets." ICE.

Damage Information Reporting Tool (DIRT). (2006). "CGA DIRT Analysis and Recommendations for Calendar Year 2005". Retrieved May 28, 2013, from https://www.damagereporting.org/dr/download/2005_DIRT_Report.pdf

DiBenedetto, S. P., Mochny, H. R. and Hanson, D. R. (2010). "Pipeline Case Histories with 2D and 3D Underground Imaging." Pipelines 2010: Climbing New Peaks to Infrastructure Reliability – Renew, Rehab, and Reinvest, 1038-1047.

Dojack, L. (2012). "Ground Penetrating Radar Theory, Data Collection, Processing and Interpretation: A Guide for Archaeologists.

Dolgiy, A., Dolgiy, A., Zolotarev, V. (2006). "Optimal Radius Estimation for Subsurface Pipes Detected by Ground Penetrating Radar." In Proceedings of 11th

International Conference on Ground Penetrating Radar, Columbus, Ohio, USA, 4.1 Utility Detection, UTL.6 on CD.

Duckham, M. and McCreadie, J. E. (2002). "Error-Aware GIS Development." In W Shi, PF Fisher, & MF Goodchild (Eds.), Spatial Data Quality, pp. 63-75, Taylor & Francis: London.

Duckham, M. and McCreadie, J. E. (1999). "An Intelligent, Distributed, Error-Aware Object-Oriented GIS." In Proceedings 1st International Symposium on Spatial Data Quality, Hong Kong, pp. 496-506.

Duda, R. and Hart, P. (1972). "Use of the Hough Transformation to Detect Lines and Curves in Pictures." Communication of the ACM, 15(1), 11-15.

Dutton, G. (1992). "Handling Positional Uncertainty in Spatial Databases." In: Proceedings of the 5th International Symposium on Spatial Data Handling, August 1992, Charleston, South Carolina, IGU Commission of GIS, Charleston.

Dziadak, K., Kumar, B., and Sommerville, J. (2009). "Model for the 3D Location of Buried Assets Based on RFID Technology." Journal of Computing in Civil Engineering, 23(3), 2009, 148-159.

Environmental System Research Institute (ESRI) (2013). "What is GIS?" Retrieved August 20, 2013 from http://www.esri.com/what-is-gis/overview#overview_panel

Environmental System Research Institute (ESRI). (2008). "The Multipatch Geometry Type." An ESRI White Paper, Redlands, California.

Environmental System Research Institute (ESRI). (1995). "Introduction to ArcView." Environmental System Research Institute, Redlands, California, USA.

Federal Highway Administration (FHWA) of U.S. Department of Transportation. (2011). "Case Studies of Subsurface Utility Engineering." Retrieved May 28, 2013, from http://www.fhwa.dot.gov/programadmin/case.cfm

Farrimond, M. S. and Parker, J. M. (2008). "The Importance of Seeing Through the Ground – A Utility Perspective." Proceedings of the 12th International Conference on Gournd Penetrating Radar, 16-19 June, 2008, Birmingham, UK.

Felt, T. (2007). "Pipeline Safety: Momentum on Damage Prevention." In the Pipe, Association of Oil Pipe Lines, American Petroleum Institute, 1(2), 2007.

Fitzgibbon, A., Pilu, M. and Fisher, R. (1999). "Direct Least-squares Fitting of Ellipses." Pattern Analysis and Machine Intelligence, IEEE Transactions on, 21(5), 476-480.

Goodchild, F. M. (1998). “Uncertainty: The Archilles Heel of GIS?” Geo Info System, 8(11), 50-52.

Goodchild, M. F. and Hunter, G. J. (1997). “A Simple Positional Accuracy Measure for Linear Features.” International Journal of Geographical Information Science, 11(3), 299-306.

Greenwalt, C. R. and Shultz, M. E. (1968). “Principles of Error Theory and Cartographic Applications.” ACIC Technical Report No. 96, United States Air Force, Aeronautical Chart and Information Center, St. Louis, Missouri.

Halfawy, M. R., Newton, L. A. and Vanier, D. J. (2005). “Municipal Infrastructure Asset Management Systems: State-of-the-art Review.” In CIB W78 Conference on Information Technology in Construction, pp. 1-8.

Hansen, N. (2011). “The CMA Evolution Strategy: A Tutorial.” Retrieved September 16, 2012 from http://www.lri.fr/~hansen/cmatutorial.pdf

Harris, D. A. (2006). “Pavement Thickness Evaluation Using Ground Penetrating Radar.” PhD Dissertation, Purdue University, West Lafayette, Indiana.

Herman, H. (1997). “Robotic Subsurface Mapping Using Ground Penetrating Radar.” Doctoral dissertation, tech. report CMU-RI-TR-97-19, Robotics Institute, Carnegie Mellon University.

Heuvelink, G. B. and Brown, J. D. and Van Loon, E. E. (2007). “A Probabilistic Framework for Representing and Simulating Uncertain Environmental Variables.” International Journal of Geographical Information Science, 21(5), 497-513.

Hough, P. (1962). “Method and Means for Recognizing Complex Patterns.” US Patent N0. 3,069,654, 1962.

Huxhold, W. E. (1991). “An Introduction to Urban Geographical Information System.” Oxford University Press, New York.

Jaw, S. W., Hashim, M. (2013). “Locational Accuracy of Underground Utility Mapping Using Ground Penetrating Radar.” Journal of Tunnlling and Underground Space Technology, 35(2013), 20-29.

Jeong, H. S., Abraham, D. M. and Lew, J. J. (2004). “Evaluation of an Emerging Market in Subsurface Utility Engineering.” Journal of Construction Engineering and Management, 130(2), 2004, 225-234.

Jeong, H. S., Arboleda, C. A., Abraham, D. M., Halpin, D. W. and Bernold, L. E. (2003). “Imaging and Locating Buried Utilities.” Publication FHWA/IN/JTRP-2013/12.

Joint Transportation Research Program, Indiana Department of Transportation and Purdue University, West Lafayette, Indiana, 2003, doi:10.5703/1288284313237.

Kumar, B. and Sommerville, J. (2012). "A Model for RFID-based 3D Location of Buried Assets." Automation in Construction, 21(2012), 121-131.

Lahouar, S. (2003). "Development of Data Analysis Algorithms for Interpretation of Ground Penetrating Radar Data." PhD Dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Lanka, M., Butler, A. and Sterling, R. (2002). "Use of Approximate Reasoning Techniques for Locating Underground Utilities." Journal of Tunnelling and Underground Space Technology, 16(2002), 13-31.

Lester, J. and Leonard, E. B. (2007). "Innovative Process to Characterize Buried Utilities Using Ground Penetrating Radar." Automation in Construction, 16(2007), 546-555.

Lew, J. J. (2000). "Cost Savings on Highway Projects Utilizing Subsurface Utility Engineering." Federal Highway Administration, Washington, DC.

Li, W., Zhou, H. and Wan, X. (2012). "Generalized Hough Transform and ANN for Subsurface Cylindrical Object Location and Parameters Inversion from GPR Data." In Proceedings of the 14th International Conference on Ground Penetrating Radar, Shanghai, China, pp. 281-285.

Loulizi, A. (2001). "Development of Ground Penetrating Radar Signal Modeling and Implementation for Transportation Infrastructure Assessment." Ph.D. Dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Maas, C., Schmalzl, J. (2013). "Using Pattern Recognition to Automatically Localize Reflection Hyperbolas in Data from Ground Penetrating Radar." Computer & Geosciences, 58(2013), 116-125.

Manacorda, G., Scott, G. H., Rameil, M., Courseille, R., Farrimond, M., and Pinchbeck, D. (2007). "The ORFEUS Project: A Step Change in Ground Penetrating Radar Technology to Locate Buried Utilities." Proc., Presented at ISTT NoDig 2007, Rome.

Marking Underground Utility Depth System (MUUDS). (2011). "Industry Statistics for Underground Utility Damages." Retrieved May 28, 2013, from http://www.muuds.com/images/TSCFactsheet.pdf

McMahon, W., Burtwell, M. H. and Evans, M. (2005). "Minimising Street Works Disruption: The Real Cost of Street Works to the Utility Industry and Society." UK Water Industry Research, London.

Metje, N., Atkins, P. R., Brennan, M. J., Chapman, D. N., Lim, H. M., Machell, J., Muggleton, J. M., Pennock, S., Ratcliffe, J., Redfern, M., Rogers, C. D. F., Saul, A. J., Shan, Q., Swingler, S. and Thomas, A. M. (2007). "Mapping the Underworld – state-of-the-art review." Journal of Tunnelling and Underground Space Technology, 22(2007), 568-586.

Naganuma, Y., Kotaki, H., Nomura, Y., Kato, N. and Sudo, Y. (2011). "Method of Estimating the Radius, Depth and Direction of Buried Pipes With Ground Penetrating Radar." Journal of The Japanese Society for Non-Destructive Inspection, 60(9), 548-553.

National Transportation Safety Board (NTSB). (2007). "National Transportation Safety Board (NTSB) on Damage Prevention." In the Pipe, Association of Oil Pipe Lines, American Petroleum Institute, 1(2), 2007.

National Underground Asset Group (NUAG). (2006). "Capturing, Recording, Storing and Sharing Underground Asset Information – A Review of Current Practice and Future Requirements." UK Water Industry Research Ltd., London.

Neal, A. (2004). "Ground Penetrating Radar and Its Use in Sedimentology: Principles, Problems and Progress." Earth Science Reviews, 66(3), 261-330.

Neto, P. X. and Medeiros, W. E. (2006). "A Practical Approach to Correct Attenuation Effects in GPR Data." Journal of Applied Geophysics, 59(2), 140-151.

O'Leary, P. and Zsombor-Murray, P. (2004). "Direct and Specific Least-square Fitting of Hyperbolae and Ellipses." Journal of Electronic Imaging, 13(3), 492-503.

Olhoeft, G. R. (2000). "Maximizing the Information Return from Ground Penetrating Radar." Journal of Applied Geophysics, 43(2000), 175-187.

Perkal, J. (1956). "On Epsilon Length." Bulletin De I'Academie Polonaise Des Sciences, 4, 399-403.

Plati, C. and Loizos, A. (2013). "Estimation of In-situ Density and Moisture Content in HMA Pavements Based on GPR Trace Reflection Amplitude Using Different Frequency." Journal of Applied Geophysics, 97(2013), 3-10.

Poku, S. E. and Arditi, D. (2006). "Construction Scheduling and Progress Control Using Geographical Information Systems." Journal of Computing in Civil Engineering, 20(5), 2006, 351-360.

Ristic, A. V., Petrovacki, D., Govedarica, M. (2009). "A New Method to Simultaneously Estimate the Radius of a Cylindrical Object and the Wave Propagation Velocity from GPR Data." Computer & Geosciences, 35(8), 2009, 1620-1630.

Roberts, R. L. and Daniels, J. J. (1996). "Analysis of GPR Polarization Phenomena." Journal of Environmental and Engineering Geophysics, 1(2), 139-157.

Rogers, C. D. F. and Cohn, A. G. (2013). "Mapping the Underworld Brochure-How good are we at finding our own?" Retrieved May 1, 2013, from http://www.mappingtheunderworld.ac.uk/MTU%20Brochure%20Final%20Version.pdf

Shamsi, U. M. (2002). "GIS Tools for Water, Wastewater, and Stormwater system." ASCE Press.

Shaw, M. R., Molyneaux, T. C. K., Millard, S. G., Taylor, M. J. and Bungey, J. H. (2003). "Assessing Bar Size of Steel Reinforcement in Concrete Using Ground Penetrating Radar and Neural Networks." Insight: Non-Destructive Testing and Condition Monitoring, 45(12), 813-816.

Shi, W. Z. and Liu, W. B. (2000). "A Stochastic Process-Based Model for Positional Error of Line Segments in GIS." International Journal of Geographical Information Science, 14(1), 51-66.

Shi, W. Z. (1998). "A Generic Statistical Approach for Modeling Errors of Geometric Features in GIS." International Journal of Geographical Information Science, 12(2), 131-143.

Shi, W. Z. (1994). "Modeling Positional and Thematic Uncertainty in Integration of GIS and Remote Sensing." ITC Publication, 22, Enschede, The Netherlands.

Shihab, S. and AI-Nuaimy, W. (2005). "Radius Estimation for Cylindrical Objects Detected by Ground Penetrating Radar." Subsurface Sensing Technologies and Applications, 6(2), 151-166.

Shihab, S., AI-Nuaimy, W. and Eriksen, A. (2004). "Radius Estimation for Subsurface Cylindrical Objects Detected by Ground Penetrating Radar." In Proceedings of 10th International Conference on Ground Penetrating Radar, Delft, the Netherlands, pp. 319-322.

Sickle, J. V. (2009). "Static GPS, DGPS and RTK GPS." Retrieved July 31, 2013 from https://courseware.e-education.psu.edu/downloads/geog862/Lesson7/index.htm

Sterling, R. L., Anspach, J., Allouche, E., Simicevic, J., Rogers, C. D. F., Weston, K., et al. (2009 a). "Encouraging Innovation in Locating and Characterizing

Underground Utilities." SHRP 2 Report S2-R01-RW. Washington, D.C.: Transportation Research Board.

Sterling, R., Anspach, J., Allouche, E., Simicevic, J. and Rogers, C. D. F. (2009b). "Innovation in Utility Locating and Mapping Technologies." In ICPTT 2009 Advances and Experiences with Pipelines and Trenchless Technology for Water, Sewer, Gas, and Oil Applications, pp. 793-802.

Stolte, C. and Nick, K. (1994). "Eccentricity-migration: A method to Improve the Imaging of Pipes in Radar Reflection Data." In Proceedings of 5th International Conference on Ground Penetrating Radar, Ontario, Canada, pp. 723-733.

Su, X., Talmaki, S., Cai, H. and Kamat, V. (2013). "Uncertainty-aware Visualization and Proximity Monitoring in Urban Excavation: A Geospatial Augmented Reality Approach." Visualization in Engineering, 1(1), 2013, 1-13.

Thomson, J., Wang, L. and Royer, M. (2009). "State of Technology Review Report on Condition Assessment of Ferrous Water Transmission and Distribution Systems." National Risk Management Research Laboratory, Office of Research and Development, US Environmental Protection Agency.

Tong, X., Sun, T., Fan, J., Goodchild, M. F. and Shi, W. (2013). "A Statistical Simulation Model for Positional Error of Line Features in Geographic Information System (GIS)." International Journal of Applied Earth Observation and Geoinformation, 21(2013), 136-148.

Trenchless Technology Network (TTN). (2002). "Underground Mapping, Pipeline Location Technology and Condition Assessment." University of Birmingham, Infrastructure Engineering and Management Research Center.

US National Coordination Office (USNCO). (2013). "GPS Overview." Retrieved June 10, 2013, from http://www.gps.gov/

Van Dam, R. L. (2001). "Cause of Ground Penetrating Radar Reflections in Sediment." PhD Dissertation, Vrije University, Amsterdam, Netherland.

Vermont Department of Public Service (VDOPS). (2010). "Vermont Underground Utility Damage Prevention – Damage Incidents Statistics and Trends." Retrieved May 28, 2013, from http://damageprevention.vermont.gov/trends

Windsor, C., Falorni, P. and Capineri, L. (2005a). "The Estimation of Buried Pipe Diameters by Generalized Hough Transform of Radar Data." In Proceedings of Progress in Electromagnetic Research Symposium, Hangzhou, China, pp. 345-349.

Windsor, C., Capineri, L., Falorini, P., Matucci, S. and Borgioli, G. (2005b). "The Estimation of Buried Pipe Diameters Using Ground Penetrating Radar." Insight-Non-Destructive Testing and Condition Monitoring, 47(2005), 394-399.

Wu, H. S. and Liu, Z. L. (2008). "Simulation and Model Validation of Positional Uncertainty of Line Feature on Manual Digitizing a Map." The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVII, Part B2, Beijing, China.

Yalciner, C. C., Bano, M., Kadioglu, M., Karabacak, V., Meghraoui, M. and Altunel, E. (2009). "New Temple Discovery at the Archaeological Site of Nysa (Western Turkey) Using GPR Method." Journal of Archaeological Science, 36(8), 1680-1689.

Yelakanti, V., Najafi, M. and Habibian, A. (2003). "Role of GIS in Pipeline Industry." In Proceedings of New Pipeline Technologies, Security, and Safety, pp. 267-275.

Yufryakov, B. A. and Linnikov, O. N. (2006). "Buried Cylinders Geometric Parameters Measurement by Means of GPR." In Proceedings of Progress in Electromagnetics Research Symposium, Cambridge, USA, pp. 187-191.

Zhang, L., Deng, M. and Chen, X. (2006). "A New Approach to Simulate Positional Error of Line Segment in GIS." Geo-spatial Information Science, 9(2), 142-146.

Zhang, J. X. and Goodchild, M. F. (2002). "Uncertainty in Geographical Information." Taylor and Francis, New York, USA.

Zogg, J. M. (2002). "GPS Basics: Introduction to the System Application Overview." Retrieved July 31, 2013 from http://geology.isu.edu/geostac/Field_Exercise/GPS/GPS_basics_u_blox_en.pdf