

Characterization of components of water supply systems from GPR images and tools of intelligent data analysis



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

David Ayala Cabrera

FluIng. Instituto U. de Matemática Multidisciplinar
Departamento de Ingeniería Hidráulica y Medio Ambiente

A thesis submitted for the degree of Doctor of Philosophy
in Water and Environmental Engineering

Supervised by:
Prof. Dr. Rafael Pérez García
Prof. Dr. Joaquín Izquierdo Sebastián
Dr. Antonio Manuel Herrera Fernández

Valencia, December 2015

Dedico este trabajo doctoral:

En primera instancia a *Dios*.

A mis padres, *David Ayala* y *Nirza Cabrera*, porque es por ustedes que soy lo que soy. Gracias por brindarme el apoyo y las bases para sacar mis proyectos adelante. Gracias por todo el cariño y por todos los esfuerzos que han y siguen realizando por mí.

A mi adorada esposa, *Anastasia Tsitsou*, gracias por el apoyo incondicional brindado, por todos los momentos sacrificados y especialmente por los bellos momentos a tu lado, por ser mi motivo de inspiración y mis deseos de superación. Gracias por todo el amor incondicional ofrecido. Mil y mil gracias Agapi mou.

Finalmente, a mis hermanos *Ramier*, *Pilar*, *Patricia* y *Katerine*.

Agradecimientos

Este trabajo ha sido posible gracias al Programa de Ayudas de Investigación y Desarrollo (PAID) de la Universitat Politècnica de València, que han concedido una beca para el autor, en el marco del Programa para la Formación de Personal Investigador (FPI) para el año 2011.

Quiero agradecer a mis directores Rafael Pérez García, Joaquín Izquierdo y Manuel Herrera por creer en este proyecto, por su apoyo, por sus buenos consejos tanto en lo emocional como en lo profesional. Adicionalmente, gracias por ser esa chispa que inició mis caminos en la investigación. A Rafael gracias por su apoyo y constancia y por llevarme por ese tan inexplorado camino de los métodos no destructivos. A Joaquín mis agradecimientos por su entrega y su dedicación, su apoyo y comprensión, gracias por esos inesperados artículos a publicar que medían mi capacidad de reacción. Gracias por toda la dedicación a los incomprensibles momentos de mi creatividad. A Manuel, gracias por todo su apoyo y amistad, por incentivar me a publicar y así compartir mis ideas, gracias por ser ese amigo y colega cómplice en el desarrollo de mi trabajo y por sus invaluable consejos. De igual manera quiero hacer extensivo este agradecimiento a todos los integrantes de FluIng, especialmente a Joanna, Eva, Enrique e Idel.

Quiero agradecer a la National Technical University of Athens de Grecia, específicamente al grupo de investigación Microwave and Fiber Optics Laboratory, por acogerme en mi periodo de estancia. Particularmente, quiero agradecer a Manthos Bimpas por su apoyo, amistad y buenos consejos en el ámbito profesional, además de hacer posible el desarrollo de esta estancia. Manthos, su ejemplo de trabajo bien realizado es motivador, Efcharisto Poli.

Mis agradecimientos para el Consorcio AQUARUM por su ayuda en la captura de datos de campo. En especial a los ingenieros Mario Tavera y Halmark Mejía.

Quiero manifestar mi enorme gratitud a los ingenieros Beatriz Wills y Guillermo Chica de la Universidad de Antioquia, con quienes siempre he podido contar con sus valiosos consejos y amistad. Así mismo, hacer extensivo mi gratitud a mis excompañeros de Ingedelta, Wilmar, Mauricio, Andrés y Jorge, quienes me apoyaron desde los inicios del doctorado.

A los Ingenieros Camilo Cruz y Patricia Torres de la Universidad del Valle, gracias por brindarme esa mano amiga.

Gracias a mis amigos Paola Nates, Alejandro Perdomo, Guillermo Sánchez, Julio Rojas, Lina Perea, Nancy Sánchez, Lorena Sánchez, Alexandra Martínez. Finalmente, a todos los que de una u otra manera han pasado por mi vida aportando. Para todos el mayor de los agradecimientos.

Abstract

Over time, due to multiple operational and maintenance activities, the networks of water supply systems (WSSs) undergo interventions, modifications or even are closed. In many cases, these activities are not properly registered. Knowledge of the paths and characteristics (status and age, etc.) of the WSS pipes is obviously necessary for efficient and dynamic management of such systems. But often such knowledge is very poor or simply non-existent. To search for and / or update information on paths and characteristics of WSS pipes, the most common activities are: excavation, processing of surveys among senior operators, and location of visible objects (e.g. valves). However, such activities are often inaccurate and provide only partial information on the WSS networks, thereby leaving many gaps in the information system. This problem is greatly augmented by considering the detection and control of leaks. This is one of the hardest problems that WSSs have to face with the management of the networks constituting the system. However, access to reliable leakage information is a complex task. In many cases, leaks are detected when the damage is already considerable, which brings high social and economic costs. In this sense, non-destructive methods (e.g., ground penetrating radar - GPR) may be a constructive response to these problems, since they allow, as evidenced in this thesis, to ascertain paths of pipes, identify component characteristics, and detect primordial water leaks. Selection of GPR in this work is justified by its characteristics as non-destructive technique that allows studying both metallic and non-metallic objects. Although the capture of information with GPR is usually successful, such aspects as the capture settings, the large volume of generated information, and the use

and interpretation of such information require high level of skill and experience.

This dissertation may be seen as a step forward towards the development of tools able to tackle the problem of lack of knowledge on the WSS buried assets. To this end we have employed the GPR, which is an electromagnetic non-invasive technique, which allows recognition of the materials in the underground without disturbing the environment. The main objective of this doctoral work is thus to generate tools and assess their feasibility of application to the characterization of components of WSSs from GPR images using various computational and data treatment methodologies. These tools have a vocation for facilitating decision-making in technical management of WSSs.

In this work we have carried out laboratory tests specifically designed to propose, develop and evaluate methods for the characterization of the WSS buried components we seek to identify and study. Additionally, we have conducted field tests, which have enabled us to determine the feasibility of implementing such methodologies under uncontrolled conditions. The methodologies developed are based on techniques of intelligent data analysis, such as segmentation and clustering, multi-agent based systems, and other specific processes devised by the author. The basic principle of this work has involved the processing of data obtained through the GPR to look for a configuration, which is stable, repeatable and allows to obtain, with the least amount of data, useful information about WSS components, with special emphasis on the pipes.

After performing numerous activities, one can conclude that, using GPR images, it is feasible to obtain more information than the typical identification of hyperbolae currently performed. In addition, this information can be observed directly, e.g. more simply, using the methodologies proposed in this doctoral work. These methodologies also prove that it is feasible to identify patterns (especially with the preprocessing algorithm termed *Agent race*) that provide fairly good

approximation of the location of leaks in WSSs. Also, in the case of pipes, one can obtain such other characteristics as diameter and material.

The main outcomes of this thesis consist in a series of tools we have developed to locate, identify and visualize WSS components from GPR images. Most interestingly, the data are synthesized and reduced so that the characteristics of the different components of the images recorded in GPR are preserved. These results are promising in the sense of allowing more complex classifications and interpretations that envisage interesting possibilities for the development of automatic techniques of identification and classification. The ultimate goal is that the developed tools facilitate decision-making in the technical management of WSSs, and that such tools can even be operated by personnel with limited experience in handling non-destructive methodologies, specifically GPR.

Resumen

Con el paso del tiempo, y debido a múltiples actividades operacionales y de mantenimiento, las redes de los sistemas de abastecimiento de agua (SAAs) sufren intervenciones, modificaciones o incluso, son clausuradas, sin que, en muchos casos, estas actividades sean correctamente registradas. El conocimiento de los trazados y características (estado y edad, entre otros) de las tuberías en los SAAs es obviamente necesario para una gestión eficiente y dinámica de tales sistemas. Sin embargo, en muchas ocasiones, tal conocimiento es muy deficiente o simplemente inexistente. Para buscar y/o actualizar la información referente a los trazados y características de las tuberías de los SAAs, las actividades más comunes son: excavación, elaboración de encuestas a operarios de mayor antigüedad, y localización de objetos visibles (por ejemplo, válvulas). Sin embargo, tales actividades suelen ser inexactas y brindan solo información parcial de las redes de los SAAs, quedando con ello múltiples vacíos de información en el sistema. A esta problemática se suma la detección y el control de las fugas de agua. Este es uno de los mayores problemas al que los SAAs tienen que hacer frente en la gestión de las redes que componen el sistema. Sin embargo, el acceso a información fiable sobre las fugas es una tarea compleja. En muchos casos, las fugas son detectadas cuando los daños en la red son ya considerables, lo que trae consigo altos costes sociales y económicos. En este sentido, los métodos no destructivos (por ejemplo, ground penetrating radar - GPR), pueden ser una respuesta a estas problemáticas, ya que permiten, como se pone de manifiesto en esta tesis, localizar los trazados de las tuberías, identificar características de los componentes y detectar las fugas de

agua cuando aún no son significativas. La selección del GPR, en este trabajo se justifica por sus características como técnica no destructiva, que permite estudiar tanto objetos metálicos como no metálicos. Aunque la captura de información con GPR suele ser exitosa, la configuración de la captura, el gran volumen de información, y el uso y la interpretación de la información requieren de alto nivel de habilidad y experiencia por parte del personal.

Esta tesis doctoral se plantea como un avance hacia el desarrollo de herramientas que permitan responder a la problemática del desconocimiento de los activos enterrados de los SAAs. Para ello hemos empleado el GPR, que es una técnica electromagnética y no invasiva, que permite el reconocimiento de los materiales enterrados en el suelo sin perturbar su entorno. El objetivo principal de este trabajo doctoral es, pues, generar herramientas y evaluar la viabilidad de su aplicación en la caracterización de componentes de un SAA, a partir de imágenes GPR, empleando diversas metodologías computacionales y de tratamiento de datos. Estas herramientas tienen vocación de facilitar la toma de decisiones en la gestión técnica de los SAAs.

En este trabajo hemos realizado ensayos de laboratorio específicamente diseñados para plantear, elaborar y evaluar metodologías para la caracterización de los componentes enterrados de los SAAs que tratamos de identificar y estudiar. Adicionalmente, hemos realizado ensayos de campo, que han permitido determinar la viabilidad de aplicación de tales metodologías bajo condiciones no controladas. Las metodologías elaboradas están basadas en técnicas de análisis inteligentes de datos, tales como segmentación y clustering, sistemas multi-agente y otros procesos específicos planteados por el autor. El principio básico de este trabajo ha consistido en el tratamiento adecuado de los datos obtenidos mediante el GPR, a fin de buscar una configuración estable, repetible y que permita obtener, con la menor cantidad posible de datos, información de utilidad para los SAAs respecto a sus componentes, con especial énfasis en las tuberías.

Tras la realización de múltiples actividades, se puede concluir que es viable obtener más información de las imágenes de GPR que la que actualmente se obtiene con la típica identificación de hipérbolas. Esta información, además, puede ser observada directamente, de manera más sencilla, mediante las metodologías planteadas en este trabajo doctoral. Con estas metodologías se ha probado que también es viable la identificación de patrones (especialmente el pre-procesado con el algoritmo *Agent race*) que proporcionan aproximación bastante acertada de la localización de las fugas de agua en los SAAs. También, en el caso de las tuberías, se puede obtener otro tipo de características tales como el diámetro y el material.

Como resultado de esta tesis se han desarrollado una serie de herramientas que permiten visualizar, identificar y localizar componentes de los SAAs a partir de imágenes de GPR. El resultado más interesante es que los resultados obtenidos son sintetizados y reducidos de manera que preservan las características de los diferentes componentes registrados en las imágenes de GPR. Estos resultados son prometedores en el sentido de que permiten clasificaciones e interpretaciones más complejas que auguran posibilidades interesantes para el desarrollo de técnicas automáticas de identificación y clasificación. El objetivo último es que las herramientas desarrolladas faciliten la toma de decisiones en la gestión técnica de los SAAs y que tales herramientas puedan ser operadas incluso por personal con una experiencia limitada en el manejo de metodologías no destructivas, específicamente el GPR.

Resum

Amb el temps, a causa de les múltiples activitats d'operació i manteniment, les xarxes de sistemes d'abastament d'aigua (SAAs) se sotmeten a intervencions, modificacions o fins i tot estan tancades. En molts casos, aquestes activitats no estan degudament registrats. El coneixement dels camins i característiques (estat i edat, etc.) de les canonades d'aigua i sanejament fa evident la necessitat d'una gestió eficient i dinàmica d'aquests sistemes. Però sovint aquest coneixement és molt pobre o simplement inexistent. Per cercar i / o actualitzar la informació en els camins i les característiques de les canonades d'aigua i sanejament, les activitats més comunes són: excavació, processament d'enquestes entre els operadors d'alt nivell, i la ubicació dels objectes visibles (per exemple, vàlvules). No obstant això, aquestes activitats són sovint inexactes i proporcionen informació parcial sobre les xarxes d'aigua i sanejament, deixant moltes llacunes en el sistema d'informació. Aquest problema es veu augmentat en gran mesura tenint en compte la detecció i control de fuites. Aquest és un dels majors problemes que els SAAs han d'afrontar en la gestió de les xarxes que componen el sistema. No obstant això, l'accés a informació fiable sobre les fuites és una tasca complexa. En molts casos, les fugues es detecten quan el dany ja és considerable, el que porta costos socials i econòmics. En aquest sentit, els mètodes no destructius (per exemple, ground penetrating radar - GPR) poden ser una resposta constructiva a aquests problemes, ja que permeten, com s'evidencia en aquesta tesi, per determinar rutes de canonades, identificar les característiques dels components, i detectar les fuites d'aigua quan encara no són significatives. La selecció del GPR en aquest treball es justifica

per les seves característiques com a tècnica no destructiva que permet estudiar tant objectes metàl·lics i no metàl·lics. Tot i que la captura d'informació amb GPR sol ser reeixida, aspectes com ara la configuració de captura, el gran volum d'informació que es genera, i l'ús i la interpretació d'aquesta informació requereix alt nivell d'habilitat i experiència.

Aquesta tesi pot ser vista com un pas endavant cap al desenvolupament d'eines capaces d'abordar el problema de la manca de coneixement sobre els actius d'aigua i sanejament enterrat. Amb aquesta finalitat hem emprat el GPR, que és una tècnica no invasiva electromagnètica, que permet el reconeixement dels materials enterrats sense pertorbar el seu entorn. L'objectiu principal d'aquest treball doctoral és, doncs, generar eines i avaluar la seva factibilitat d'aplicació a la caracterització dels components de los SAAs, a partir d'imatges GPR, utilitzant diverses metodologies de càlcul i tractament de dades. Aquestes eines tenen una vocació de facilitar la presa de decisions en la gestió tècnica de SAAs.

En aquest treball s'han dut a terme proves de laboratori específicament dissenyats per proposar, desenvolupar i avaluar mètodes per a la caracterització dels components d'aigua i sanejament soterrat que busquem identificar i estudiar. A més, hem dut a terme proves de camp, que ens han permès determinar la viabilitat de la implementació d'aquestes metodologies en condicions no controlades. Les metodologies desenvolupades es basen en tècniques d'anàlisi intel·ligent de dades, com ara la segmentació i l'agrupació, els sistemes basats en multi-agent, i altres processos específics ideats per l'autor. El principi bàsic d'aquest treball ha consistit en el tractament de dades obtingudes a través del GPR per buscar una configuració, que és estable, repetible i permet obtenir, amb la menor quantitat de dades, informació útil sobre els components d'SAA, amb especial èmfasi en la canonades.

Després de realitzar nombroses activitats, es pot concloure que, amb l'ús d'imatges de GPR, és factible obtenir més informació que la identificació típica d'hipèrboles realitzat actualment. A més, aquesta informació pot ser observada directament, per exemple, més simplement, utilitzant les metodologies proposades en aquest treball doctoral. Aquestes metodologies també demostren que és factible per identificar patrons (especialment el pre-processat amb l'algoritme *Agent race*) que proporcionen bastant bona aproximació de la localització de fuites en SAAs. També, en el cas de tubs, es pot obtenir altres característiques com ara el diàmetre i el material.

Els principals resultats d'aquesta tesi consisteixen en una sèrie d'eines que hem desenvolupat per localitzar, identificar i visualitzar els components dels SAAs a partir d'imatges GPR. El resultat més interessant és que els resultats obtinguts són sintetitzats i reduïts de manera que preserven les característiques dels diferents components registrats en les imatges de GPR. Aquests resultats són prometedors en el sentit de permetre que les classificacions i interpretacions que preveuen interessants possibilitats per al desenvolupament de tècniques automàtiques d'identificació i classificació més complexes. L'objectiu final és que les eines desenvolupades faciliten la presa de decisions en la gestió tècnica de SAA, i que tals eines poden fins i tot ser operades per personal amb poca experiència en el maneig de metodologies no destructives, específicament GPR.

Contents

List of Figures	xix
List of Tables	xxiii
Notations	xxv
1 Introduction	1
1.1 Introduction	3
1.2 Objectives	7
1.2.1 Main objective	7
1.2.2 Specific objectives	7
1.3 Justification	8
1.4 Thesis outline	11
1.5 Contributions of the thesis	15
2 Background	19
2.1 Non-destructive surveying methods	21
2.2 Non-destructive methods in the location of WSS components	23
2.3 GPR overview	25
2.4 Theoretical foundations of GPR	30
2.5 Electromagnetic parameters	31
2.5.1 Conductivity	31
2.5.2 Dielectric permittivity	32
2.5.3 Magnetic permeability	32
2.5.4 Wave propagation velocity	33
2.5.5 Reflection and refraction	34

CONTENTS

2.5.6	GPR equation	37
2.6	GPR interpretations	38
2.7	GPR data processing	42
2.7.1	Filters	44
2.7.2	Predictive deconvolution	45
2.7.3	Migration	46
2.7.4	Fast Fourier Transform – FFT	46
2.7.5	Hilbert transform	46
2.7.6	Signal correction procedures	47
2.7.6.1	Background removal	47
2.7.6.2	Gain	47
3	GPR images as tools for visualizing WSS components	49
3.1	Introduction to methodologies of visualization of objects in GPR images	53
3.1.1	GPR role in the analysis of WSS components	54
3.1.2	Factors influencing the visualization of objects in GPR images	55
3.2	Principles of the proposed visualization process – T_{14} and T_{15} matrices	56
3.2.1	WAV classification	60
3.2.2	WAV extraction	62
3.2.2.1	Generation of matrices D	63
3.2.2.2	Generation of matrices U	65
3.2.3	WAV accumulation	66
3.2.4	WAV rebuild	67
3.3	Experimental study – matrices T_{14} and T_{15}	69
3.3.1	Case 1. Simple case	70
3.3.2	Case 2. Complex case	72
3.4	Implementation issues	75
3.5	Summary and comments	75

4	Location of buried plastic pipes using multi-agent support based on GPR images	77
4.1	Automatic pipe location – multi-agent process	81
4.2	A brief introduction to multi-agent systems	82
4.3	Proposed multi-agent process	84
4.4	Experimental study – multi-agent process	87
4.4.1	Customizing the multi-agent location process	88
4.4.2	Case 1. Simple case	89
4.4.3	Case 2. Complex case	90
4.5	Implementation issues	92
4.6	Summary and comments	93
5	Segmentation of GPR images of WSS components	95
5.1	Principles of the segmentation and cleaning proposed	100
5.2	Segmentation and cleaning methodology	107
5.2.1	Binarization	108
5.2.1.1	Standardizing traces	108
5.2.1.2	Selecting trends	109
5.2.2	Segmentation and cleaning	111
5.2.2.1	Edge detection	112
5.2.2.2	Horizontal straight lines detection – Hough transform	113
5.2.2.3	Removing lines	115
5.2.3	Refining	116
5.3	Density analysis	116
5.4	Experimental studies – segmentation and cleaning	118
5.4.1	Segmentation and cleaning - pre-processing	119
5.4.2	Density analysis	123
5.5	Summary and comments	126
6	Automatic generation of WSS pipeline layout maps based on GPR images	129
6.1	Introduction to the use of GPR for WSS component mapping	134
6.2	Proposal for detecting piping layout	136

CONTENTS

6.2.1	Approach 1	136
6.2.2	Approach 2	137
6.3	Approaches 1 and 2 – general aspects	139
6.4	Sampling of interest area	141
6.4.1	Latin square sampling (LSS)	141
6.4.2	Training for pipe existence plausibility	143
6.4.3	Adaptive sampling route	145
6.4.4	Generation of probability maps	148
6.4.4.1	Obtaining points of view	148
6.4.4.2	Integration of points of view	150
6.5	Image pre-processing	151
6.6	Analysis of the proposed system for detecting the piping layout in a case of an urban WSS	153
6.6.1	Case study - Approach 1	154
6.6.2	Case study - Approach 2	156
6.7	Summary and comments	162
7	Water supply system component evaluation from GPR radar- grams using a multi-agent approach	165
7.1	Principles of the proposed pre-processing – Agents race	170
7.2	Proposed pre-processing algorithm – Agents race	174
7.2.1	Interpretation	175
7.2.2	Decision to move	177
7.2.3	Movement time	177
7.2.4	The race phases	178
7.3	Implementation issues and recommendations	179
7.4	Experimental study – <i>Agent race</i>	181
7.4.1	Case study	181
7.4.2	Data reduction percentages in GPR matrix data	183
7.5	Summary and comments	185

8	Hierarchical agglomerative clustering analysis of GPR data	187
8.1	Motivation and principles of the proposed clustering method – Hierarchical agglomerative clustering	191
8.2	Proposed system architecture	193
8.3	Hierarchical Agglomerative Clustering	194
8.3.1	Distance metrics	195
8.3.2	Linkage methods	196
8.3.3	Cophenetic correlation	197
8.3.4	Cluster analysis	198
8.4	Experimental Study – HAC	200
8.5	Summary and comments	207
9	GPR-based water leak models in water supply systems	209
9.1	Short introduction to the role of GPR to tackle leaks in WSSs	214
9.2	Data capturing: design and layout of the laboratory tests	216
9.3	Analysis of the location and identification of hyperbolas from the raw images	218
9.3.1	Initial state: raw images	218
9.3.2	Final state: raw images	220
9.4	Analysis: contrast between raw images for the initial and final states	223
9.5	Analysis of location and identification of anomalies in pre-processed images	226
9.5.1	Data pre-processing and ordering algorithm	226
9.5.2	Initial state: pre-processed images	229
9.5.3	Final state: pre-processed images	231
9.6	Analysis: contrast between pre-processed images	233
9.7	3D comparison of the analysis of contrasts between raw and pre-processed images	235
9.8	Analysis of field images: a case study	238
9.9	Summary and comments	242

CONTENTS

10 Conclusions and future developments	245
10.1 Summary	247
10.2 Conclusions	251
10.3 Future developments	262
Appendices	265
A File structure: *.dzt	267
B Equipment configuration for the performed tests	271
C Contributions	279
References	289

List of Figures

1.1	Chapter general scheme	14
2.1	Image capture through GPR prospecting	26
2.2	Reflection and refraction of incident energy on a horizontal contact	34
3.1	Chapter 3 general approach	51
3.2	Scheme of the proposed visualization process	57
3.3	Matrix A , formed by the vector traces, X_j	58
3.4	Proposed visualization method	59
3.5	The transformation methodology on a trace	60
3.6	WAV classification, selection of norm h	61
3.7	Generation of matrices B with $h=1000$ for one X_j in A	62
3.8	Generation of matrices D , starting from minimum values of the wave amplitude	63
3.9	Example of generation of matrices D	64
3.10	Generation of matrices U , starting from maximum values of the wave amplitude	65
3.11	Matrices $T3$, $T5$, $T9$ and $T12$ from matrix A	67
3.12	Example of matrices $T14$ y $T15$	68
3.13	Schematic configuration for Case 1, experimental studies	70
3.14	Raw image and $T14$ and $T15$ images obtained with GPR	71
3.15	Case 2, experimental studies	72
3.16	Case 2, application of the proposed method	73
4.1	Chapter 4 scope	79
4.2	Automatic pipe location by multi-agent methods	81

LIST OF FIGURES

4.3	Phases of the proposed multi-agent process	84
4.4	Proposed multi-agent interface	86
4.5	Evolution of the proposed multi-agent process	87
4.6	Simple case	90
4.7	Complex case	91
5.1	Conceptual approach of Chapter 5	98
5.2	Schematic configuration, static tests	101
5.3	Raw images, static tests	102
5.4	Contour lines for raw data with $h = 1500$, static tests	103
5.5	Raw data traces, static tests	106
5.6	Segmentation and cleaning methodology proposed	107
5.7	Trace representations	109
5.8	Selection of wave path trends	110
5.9	Image region, 3×3 mask	112
5.10	Hough transform, Hough parametric space to detect straight lines	115
5.11	Density analysis of the different processes that make up the pro- posed pre-processing	117
5.12	Schematic configuration, experimental studies	119
5.13	Application of the proposed pre-processing	120
5.14	Expansion of pipe localization zone in the resulting images from pre-processing	122
5.15	Density maps	125
6.1	Chapter 6 scope	132
6.2	Schematic representation of the proposed system. Approach 1 . .	137
6.3	Schematic representation of the proposed approach, Approach 2 .	138
6.4	Demarcation of the interest area, PCI capture	140
6.5	Latin square-based initial profile acquisition	143
6.6	Examples of possible decision for the first step of the adaptive route according to end point position of a PL	145
6.7	Interest area with PAd information	148
6.8	Evolution of a set of agents starting from a set of PAds	150
6.9	Vertical and horizontal point of view	150

LIST OF FIGURES

6.10	Example of probability map	151
6.11	Location of the sampling area – Case study	153
6.12	Pipe path location – Real case, Approach 1	155
6.13	Schematic configuration, case study – Approach 2	156
6.14	Comparison of vertical probability maps	157
6.15	Comparison of horizontal probability maps	158
6.16	Combination of pre-processed profiles and probability area	159
6.17	Merging Figure 6.16 with the interpretation of likely location of the pipe	160
6.18	Plan view of Figure 6.17	161
7.1	Chapter 7 scope	167
7.2	Principles of the GPR image pre-processing	171
7.3	Schema of the proposed multi-agent algorithm	175
7.4	Some competition times	181
7.5	Race last time step vs schematic configuration test	182
7.6	Pre-processing results	183
7.7	Data reduction percentages through the multi-agent pre-process developed	184
8.1	Chapter 8 scope	189
8.2	Architecture of the proposed system	193
8.3	Layouts of the prepared configurations	200
8.4	GPR survey files for the subsoil configurations in Figure 8.3	202
8.5	Obtained <i>cophenet</i> correlation coefficients	203
8.6	Clusters obtained for the configurations considered in Figure 8.3 placed back in the original space	204
8.7	Images resulting from the removal of pipe-free layers	206
9.1	Chapter 9 scope	212
9.2	Laboratory layout	217
9.3	Identifiable forms in raw images in Sections 9.3.1 and 9.3.2	218
9.4	Horizontal profiles – Raw images for the initial state	219
9.5	Vertical profiles – Raw images for the initial state	220

LIST OF FIGURES

9.6	Horizontal profiles – Raw images for final state	221
9.7	Vertical profiles – Raw images for final state	222
9.8	Horizontal profiles – Contrasting raw images for initial and final states	224
9.9	Vertical profiles – Contrasting raw images for initial and final states	225
9.10	Schema for the agent-race algorithm – Generation of Output1, and Output2	227
9.11	Output 2 building	228
9.12	Identifiable target forms in pre-processed images in Section 9.5.2 and 9.5.3	228
9.13	Horizontal profiles – Pre-processed images for the initial state	229
9.14	Vertical profiles – Pre-processed images for the initial state	230
9.15	Horizontal profiles – Pre-processed images for the final state	231
9.16	Vertical profiles – Pre-processed images for the final state	232
9.17	Horizontal profiles – Contrast between pre-processed images	234
9.18	Vertical profiles – Contrast between pre-processed images	235
9.19	Example of 3D model construction	236
9.20	Comparison between 3D models generated from the interpretation of the horizontal and the vertical profiles contrasting the initial and the final states	237
9.21	Layout for the fieldwork	238
9.22	Horizontal profiles – Raw images	239
9.23	Vertical profiles – Raw images	239
9.24	Horizontal profiles – Pre-processed images	240
9.25	Vertical profiles – Pre-processed images	241
9.26	Interpretation from pre-processed images, 3D model	242

List of Tables

1.1	Contributions to Congresses	17
1.2	Contributions in high impact Journals	18
2.1	Geophysical methods and physical properties of whom depend . .	22
2.2	Limitations of underground mapping technologies	24
3.1	Characteristics of the buried pipes used for testing	69
4.1	Characteristics of the buried pipes used for testing	88
4.2	Looking-for-pipes excerpt within stage 2 of the working algorithm	92
6.1	Auto-regressive model of order 4, initial capture profiles	154
6.2	Auto-regressive model of order 4, profiles of the guided trajectory	154
7.1	Interpretation rule exceptions pseudo-code	176
7.2	The payoff function for each agent	180
8.1	Characteristics of the buried pipes used for testing	201
8.2	Combinations of metrics and linkage methods used within HAC .	203
A.1	*.dzt date structure	269
A.2	*.dzt header structure	270
B.1	Data capturing parameters for the performed tests – Equipment configuration	273

LIST OF TABLES

Notations

ACo Combination of Average linkage and Cosine distance metric

AE Acoustic emission

AEu Combination of Average linkage and Euclidean distance metric

ASEu Combination of Average linkage and Seuclidean distance metric

CB radio Citizens band radio

CCo Combination of Complete linkage and Cosine distance metric

CCTV Close-circuit television

CEu Combination of Complete linkage and Euclidean distance metric

CMP Common midpoint

CSEu Combination of Complete linkage and Seuclidean distance metric

EC Eddy currents

EM Electromagnetic methods

EW East-West

FDTD Finite-difference time-domain

FFT Fast Fourier transform

Fib Asbestos cement

NOTATIONS

FIR Finite-response filter

Fund Cast iron

GIS Geographic information systems

GPR Ground penetrating radar

GPS Global positioning system

GSSI Geophysical survey systems Inc.

HAC Hierarchical agglomerative clustering

HS Hammer soundings

HT Hydroscope technology

IE Impact-echo

IIR Infinite-response filter

LFEM Low frequency electromagnetic field

LSS Latin square sampling

MAS Multi-agent system

MF Magnetic flux

MFL Magnetic flux leakage

ML Magnetic locators

MSCL Mild steel cement-lined

NS North-South

PAd Profile of adaptive capture

PCCP pre-stressed concrete cylindrical pipe

PCI Profile capture input

- PE** Polyethylene
- PIGs** Pipeline inspection gauges
- PIM** Pulsed induction methods
- PIRAT** Pipe inspection real-time assessment technique
- PL** Profile of Latin squares capture
- PMFs** Passive magnetic fields
- PVC** Polyvinyl chloride
- RGB** Red-Green-Blue color code scale
- RM** Resistivity methods
- RMPS** Rapid magnetic permeability scan
- RT** Radio graphic testing
- RTK** Real-time kinematic
- SCo** Combination of Single linkage and Cosine distance metric
- SEu** Combination of Single linkage and Euclidean distance metric
- SM** Sonar methods
- SN** South-North
- SSET** Sewer scanner and evaluation technology
- SSEu** Combination of Single linkage and Seclidean distance metric
- SVM** Support vector machines
- TLM** Transmission-line-matrix
- US - EPA** United States Environmental Protection Agency
- UT** Ultrasonic testing

NOTATIONS

UWB Time domain ultra wide band

WAV Wave amplitude values

WE West-East

WI Water input

WIP Wave impedance probe

WO Water output

WSS Water supply system

Chapter 1
Introduction

1.1 Introduction

Water supply systems (WSSs) are integrated by a wide variety of elements that are required for proper supply operation. These elements are mainly pipelines, as the components of greatest length, valves, reservoirs and other auxiliary elements. Information about these components, including the undergone changes and abnormalities that may have occurred over time, is critical to the productive control and technical management of WSSs [Bimpas *et al.* (2011)]. Adequate knowledge of the layout and characteristics (such as diameter, material, aging, e.g.) of WSS components is essential to manage these systems efficiently and dynamically. This knowledge is crucial to achieve objectives of the technical management of WSSs such as: identification of illegal connections, leak detection and water control, simulation and operation of networks, study of the evolution of pollutants in networks, maintenance planning, rehabilitation and renovation of WSS components, and applications of Geographic Information Systems (GIS), among others.

There is a large number of WSSs that do not have adequate information about their elements. In other systems information is implemented in an incipient way. It is common to find urban systems with many years of service (even more than 50 years), at which time of implementation the current computational standards were not available. In many systems, the only reference is a set of drawings which, over time, gradually becomes obsolete, deteriorates or is lost so that only the living memory of the system built in the knowledge of the operators remains. If we also consider the constant reforms, justified by operational issues, and the addition of new infrastructure due to population growth, the result is a set of buried system assets that form a real underground city, which is partially or completely unknown. It is a fact that in many cases the layout and such characteristics of pipelines such as diameter, material, time of use, condition of current status, among others, are unknown.

1.1. Introduction

Knowledge of the components of water supply systems, especially pipes, is very important. To find and / or update the layout and characteristics of these components various approaches are commonly used that allow for plans of the network configuration: inspections through excavations on specific pathways, surveys of personnel with seniority and experience, identification of visible components such as valves, inspection wells, and others. In this sense, there are works such as Brás (2012) for the implementation of a digital cadastre of the water supply and sewage networks of the city of Coimbra (Portugal). In Tavera (2013), a plan of the supply network is elaborated as a starting point for developing better planning and management of the drinking water system in the city of Tegucigalpa (Honduras). The plan, in both studies, consisted in locating and linking logically (based on experience) all the installed elements and all the points where the pipes are visible (works, bridge underpasses, river overpasses, reservoir and tank valve chambers, etc.). Additionally, the memory of the workers in relation to the network was used, which allowed to define the layout of the network of aqueducts in the area of study. For each pipe, design, material, diameter and thickness were inventoried. Due to the difficulty of visual accessibility to some pipes, scans through specific excavations were performed (in both studies) to resolve doubts raised during the field work. The information collected was combined with pre-existing system information such as drawings on paper, the repair history of the system, the known location of valves, and others.

Works as the two quoted in the previous paragraph, despite being of great importance and relevance to WSSs, are very complex and costly in time, as far as information retrieval and consolidation concerns. Additionally, the use of excavation as a method of verification, or to complete missing areas, implies high economic and social impact [Royal *et al.* (2011)]. Among the most obvious social impacts, we can mention road closures (pedestrian or civil), traffic congestion, and damage to other nearby facilities. In addition, these intrusive activities are often inaccurate and provide only partial information about WSS networks, thereby leaving many gaps in the information system. As a result, currently it is observed a growing interest in non-destructive examination methods of WSS components, instead of other methods conventionally employed. Studies such as the one of the

United States Environmental Protection Agency (US - EPA), encourage the use of non-destructive tools as methodologies that enable the technical management in WSSs, compared to other more traditional methods (destructive methods) [Thomson *et al.* (2009)].

The development of non-invasive techniques in the exploration and recovery of subsurface information has shown a growing interest from infrastructure facilities both from public and private sectors in recent years. There are applications in different fields such as oil and gas exploration, geology, location of ducts and pipes, and archeology, among others [Pasolli *et al.* (2009a)]. In this regard, non-destructive methods, for example, ground penetrating radar (GPR) may give a suitable response to problems of WSSs to locate the paths of pipes and to detect incipient water leaks, when they are not significant yet. GPR selection in this dissertation is based on its characteristics as a non-destructive technique, which can identify paths and characteristics of both metallic and non-metallic pipes, compared to other non-destructive methods [Cist & Schutz (2001); Takahashi (2006)]. Similarly, as said, GPR has a potential application in the search for other WSS features such as water leaks [Crocco *et al.* (2010)].

Information retrieval by non-destructive methods is thus a good alternative. However, factors such as the complexity of the spatial arrangement of the water and sanitation networks, along with the constant growth of cities, make the huge volume of information generated by these methods to difficult the interpretation and analysis of obtained data. To draw relevant conclusions from the obtained data, system operators need a high level of skill and experience in the use and interpretation of data [Doolittle & Bellantoni (2010); Thomson *et al.* (2009)].

This dissertation may be seen as a breakthrough for the development of tools which can respond to the problem of lack of knowledge on the layout and characteristics of urban WSS components in use. To this purpose, the GPR, which is a non-invasive technique that allows to get knowledge of subsurface materials without altering the medium, is used. GPR operation is based on the emission and posterior reception of electromagnetic signals to the ground [Schrott *et al.* (2013)].

1.1. Introduction

The main objective of this doctoral work is to propose methodologies for the characterization of WSS components, and evaluate their application feasibility. Firstly, pipes are considered and the GPR response to the variation of some kinds of materials is observed. To this purpose, laboratory tests with buried pipes of different material properties in a controlled environment were performed in the facilities of the research group Fluing - Multidisciplinary Institute for Applied Mathematics of the Polytechnic University of Valencia. These tests enable us to analyze the effect of various materials and environment conditions on the signal propagation. Additionally, examples of field tests are also presented in this thesis. Once determined that it is feasible to observe pipes in WSSs through GPR, the work was extended to the application of the obtained results to leakage aspects. The aim of extending the research to water leaks is an additional contribution to evaluate the response of the methodological proposals made in this thesis to other objects with amorphous characteristics which vary over time. In this case, the presence of water pockets or soil moisture caused by leakage is associated with characteristics and / or abnormalities within the GPR images. Additionally, the feasibility of using these results in useful applications for WSSs is assessed.

This research is not intended to cover all aspects and the full spectrum of possibilities offered by non-destructive techniques for WSSs, in our case the GPR. On the contrary, it is intended as an advance to serve as a basis in determining the characteristics of components of WSSs in use through intelligent decision-making systems able to generate effective and efficient actions on the service.

1.2 Objectives

1.2.1 Main objective

- The main objective of this doctoral work is to design tools and assess the feasibility of its application to the characterization of components of WSSs in service by analyzing images obtained by GPR using intelligent data analysis methodologies. We seek that these tools facilitate decision-making in the technical management of WSSs.

1.2.2 Specific objectives

To achieve the main objective of this thesis the following specific objectives are proposed:

- Generation, planning and optimization of methodologies to visualize objects (pipes) in images obtained through GPR.
- Building of visual models of automatic detection of pipeline materials that are not easily identifiable (plastics) and that, however, are used in WSSs, using GPR images.
- Search and identification of pipe features from GPR images eventually aiming at automatic detection.
- Planning and optimization of models for automatic detection of pipe layouts in WSSs using GPR images.
- Exploration of the amount of data necessary for pipe feature extraction from GPR images.
- Development of a pre-processing system favoring posterior data classification so as to improve visualization of pipelines, and reduction of data for

1.3. Justification

analysis, while preserving the characteristics of the objects contained in the images.

- Application of unsupervised clustering techniques for the analysis of GPR images allowing fast and non-subjective classification of the objects contained in the GPR images.
- Generation of interpretations of GPR images allowing the location of pipes and leaks.
- Pipe and leak pattern extraction from GPR images.

1.3 Justification

As mentioned above, a large number of water supply systems are relatively old and in many cases, the exact position of some components (pipes, valves, etc.) is unknown. Usually, in these systems there is no comprehensive control of the installed elements and such relevant pipe characteristics as diameter are unknown. In other cases, the network has to supply very large and / or scattered areas and pipes can be considerably long, which makes the network susceptible of installing illegal connections made by companies and individuals to use water fraudulently. The absence of adequate referencing and control systems clearly hampers system management.

Main cities have, in general, a wide and complete coverage of water supply service. Currently, achieving good service has transcended the implementation issues and has turned to be a problem of adequate management of the system. In this sense, today we find actions that implement management of water supply systems techniques such as sectorization [Herrera *et al.* (2015); Campbell *et al.* (2014)], network optimization [Izquierdo *et al.* (2015); Montalvo *et al.* (2014b)], vulnerability studies [Gutiérrez-Pérez *et al.* (2013); Ramos-Martínez *et al.* (2014)], management techniques for intermittent systems [Ilaya-Ayza *et al.* (2014)] or suitable decision-making actions [Montalvo *et al.* (2014a); Benítez *et al.*

(2015)], among others. However, any action, in addition to obviously consider the system hydraulic variables, should be based on a comprehensive knowledge of the morphological network conditions. Currently there is a lot of work with concepts such as the so-called smart cities, which involve dynamic, active, and real-time management of the system, which should rely on a correct identification and assessment of the state of the assets of the system [Herrera *et al.* (2012)]. In indoor supply networks, we find similar problems, albeit on a smaller scale, as often the site plans of the buildings go astray, while there is always the need to make repairs, especially when the facility is old. Consequently, characterizing such systems is also necessary.

Pipes constitute one of the main assets of a WSS, given its volume compared with other components. Pipes can be divided into mains or pipes of the main network, and secondary pipes performing distribution and services. A variety of materials and technologies are used in the production of water supply pipes. The material of the installed pipes is closely related mainly to the year of installation, the installation conditions and the diameter. Thus, in pipelines of large diameters (over 300 mm), materials of installed pipes commonly found are mild steel cement-lined (MSCL), and pre-stressed concrete cylindrical pipes (PCCPs). For very old distribution networks usually piping materials are cast iron, cement and foremost ductile iron. Polyethylene (PE) and polyvinyl chloride (PVC) are usually found in new networks [Rizo (2010)]. Plastic pipes (PE and PVC pipes) are often used in WSSs, mainly in secondary networks (service) and connections.

Characterizing a pipeline is identifying its conformation and constitution, that is to say, determining material composition, layout, internal and external diameter, operating time (age of the pipe), presence and characteristics of obstructions, usage status (used or unused), and deformation, among other features. The availability of a characterization system allows, among others, the determination of actions against possible failures, the detection of illegal connections, the design of suitable operation, the actions against possible intrusion of contaminants, and the determination of vulnerability issues. Overall, suitable characterizations of supply systems allow decision-making for efficient management and planning.

1.3. Justification

It is at this point that methods that can respond to the problem of lack of knowledge of the layout and characteristics of WSSs components become important. Contrasts among the different physical properties of the materials involved may be used to characterize a supply system. This characterization can be carried out using various geophysical methods. The advantage of the application of non-destructive methods for this purpose is based on the possibility of getting knowledge of the materials in the medium without affecting the environment. In addition, lower social and economic impacts as compared with using conventional methods (e.g. excavation) will be produced. GPR, by itself or in conjunction with other geophysical prospecting methods, is a powerful tool for the investigation of these components. The advantage of GPR application lies in the good properties of working and obtaining results exhibited by this methodology. Among other properties let us quote here its non-destructive nature, its ability to provide an overview of the study area, and its faster and cheaper applicability in comparison with other available technologies [Busquet *et al.* (1996)].

The interpretation of most surveys conducted with GPR are based on the form or shape observed in the radargram, which is determined by the contrasts produced by the characteristics of the materials, after the signal is emitted and subsequently received by the equipment. In other cases, the application of multiple processing steps provides successful results. However, the interpretation of the results, either directly or by the application of data processing methods (pre and post-processing) should be made with particular care, since inadequate application can lead to interpretation errors, yielding misleading results. Identifying buried objects (in our case WSS pipes) in raw GPR images is a complex, nontrivial task [Singh & Nene (2013)]. Therefore, efficient processing of such images requires a methodology that favors the interpretation (either manually or automatically) and resulting in obtaining accurate conclusions.

The generation of tools that enable easy and non-subjective processing of GPR surveys is paramount. These tools should provide the ability to parametrize the information and allow quality classifications, as the ones provided by intelligent classification systems, with applicability in WSS component classification.

1.4 Thesis outline

- The current chapter (**Chapter 1**) provides a general introduction to the content of the work developed in this thesis. We have tried to present some preliminary ideas clearly, and have specified the objectives pursued in the development of the thesis. Also, we have presented very briefly the state of development of the subject, seeking to create a global view of the problem and highlighting its importance for water supply systems. Finally, we have contextualized the research lines followed.
- In **Chapter 2** an introduction to non-destructive prospecting methods and their application to the detection and identification of characteristics of WSS pipes is presented. Next, we give an introduction to the GPR and a review of the literature regarding the generalities of the method and the principles and theory in which the operation with GPR is based. A review of the interpretation techniques for understanding and drawing conclusions from GPR images is performed. The difficulties associated with interpretation are then discussed. Finally, a review of some methods of processing of data associated to the information obtained from GPR is performed.
- In **Chapter 3**, a tool for displaying WSS components from GPR images is developed. This chapter presents the mathematical notation for the formulation and shows the application of each proposed procedure in a visual way, by tracking a given example. The principles of the tool presented in this chapter is based on the wave amplitude obtained from surveys conducted by GPR. With this tool two important matrices called $T14$ and $T15$ are obtained. This methodology is applied to two case studies (identified as simple and complex cases) under controlled laboratory conditions. In both cases different materials commonly employed in WSSs are tested. The findings, conclusions and comments after the application of this methodology close this chapter.

1.4. Thesis outline

- **Chapter 4** addresses the problem of locating and identifying plastic pipes using GPR images. This problem is developed using the transformation presented in the previous chapter. The development and methodological proposal presented in this chapter are carried out from a multi-agent perspective and with the aim of removing all the images of objects and features different than pipes. The design and the behavioral approach of the agents are presented in this chapter. This chapter also presents two case studies, a simple case and a complex case, under controlled laboratory conditions.
- An approach to the interpretation of GPR images is presented in **Chapter 5**. Using this approach a technique for image segmentation is presented. Segmentation is intended to facilitate image interpretation with the aim of facilitating the search for pipes in these images. The segmentation process proposed in this chapter is based on some transformations performed on matrices T_{14} and T_{15} . This chapter highlights the approach of reducing the space to be analyzed (data reduction), while preserving the characteristics of the objects embedded in the GPR images. This chapter also analyzes the volume of data and describes how to get views of the pipes easy to understand by non highly qualified staff in GPR image analysis. The planning and development of the processes in this chapter are presented by using tests conducted under controlled laboratory conditions. Processing and analysis of experimental studies are presented later on.
- The problem of detecting pipes in WSSs is addressed in **Chapter 6**. In this chapter, a system that includes three typical stages used in mapping buried objects is presented: (a) image retrieval, (b) image analysis and interpretation, and (c) mapping (of pipes, in our case). The proposed system is shown in this chapter using two approaches. The first one is an iterative process that guides GPR data retrieval to determine the most plausible location of the pipeline. Once images have been captured, the second process seeks generating probability maps of the paths of the pipes. Both approaches are studied using a multi-agent methodology in an automatic way. The segmentation presented in the previous chapter is used here as the basis for the analysis and interpretation of GPR images. The implementation of the

various processes in this chapter is developed by studying an existing pipe in an urban system (a real case).

- **Chapter 7** is devoted to improve the results obtained with the transformation of GPR images into the matrices T_{14} and T_{15} , using a multi-agent approach. To do this, an algorithm based on a competition of agents (agent race algorithm) is proposed. This algorithm maintains the premise of seeking data reduction, while preserving the characteristics of the elements contained in the images shown in previous chapters. The application process is exemplified in a case study conducted by controlled laboratory conditions.
- In **Chapter 8** exploration of features of data contained in GPR images in search for pipes in water systems is addressed. To this purpose, the use of a well-known technique for unsupervised clustering (Hierarchical agglomerative clustering), in combination with the pre-processing of data mentioned in the previous chapter (agent race algorithm) is proposed. This chapter presents an experimental design that can be adapted to a series of configurations obtained by using different objects and materials. Additionally, a sensitivity analysis of different metric distances and linking methods for cluster analysis is presented aiming at identifying the most suitable combination of both metric and linking methods.
- **Chapter 9**, of exploratory character, raises the problem of water leakage both in terms of location and identification through GPR images. This chapter considers first controlled laboratory conditions, using contrast analysis and the multi-agent pre-processing method proposed earlier (agent race algorithm). Subsequently, the feasibility of application to data from a real case is evaluated. It should be mentioned that in this chapter, a variant of the mentioned pre-processing method that improves visualization of features and encourages interpretations of GPR images is considered.
- Finally, in **Chapter 10** the findings of this work are presented, and lines of work for future research are suggested.

1.4. Thesis outline

Chapter organization is as follows.

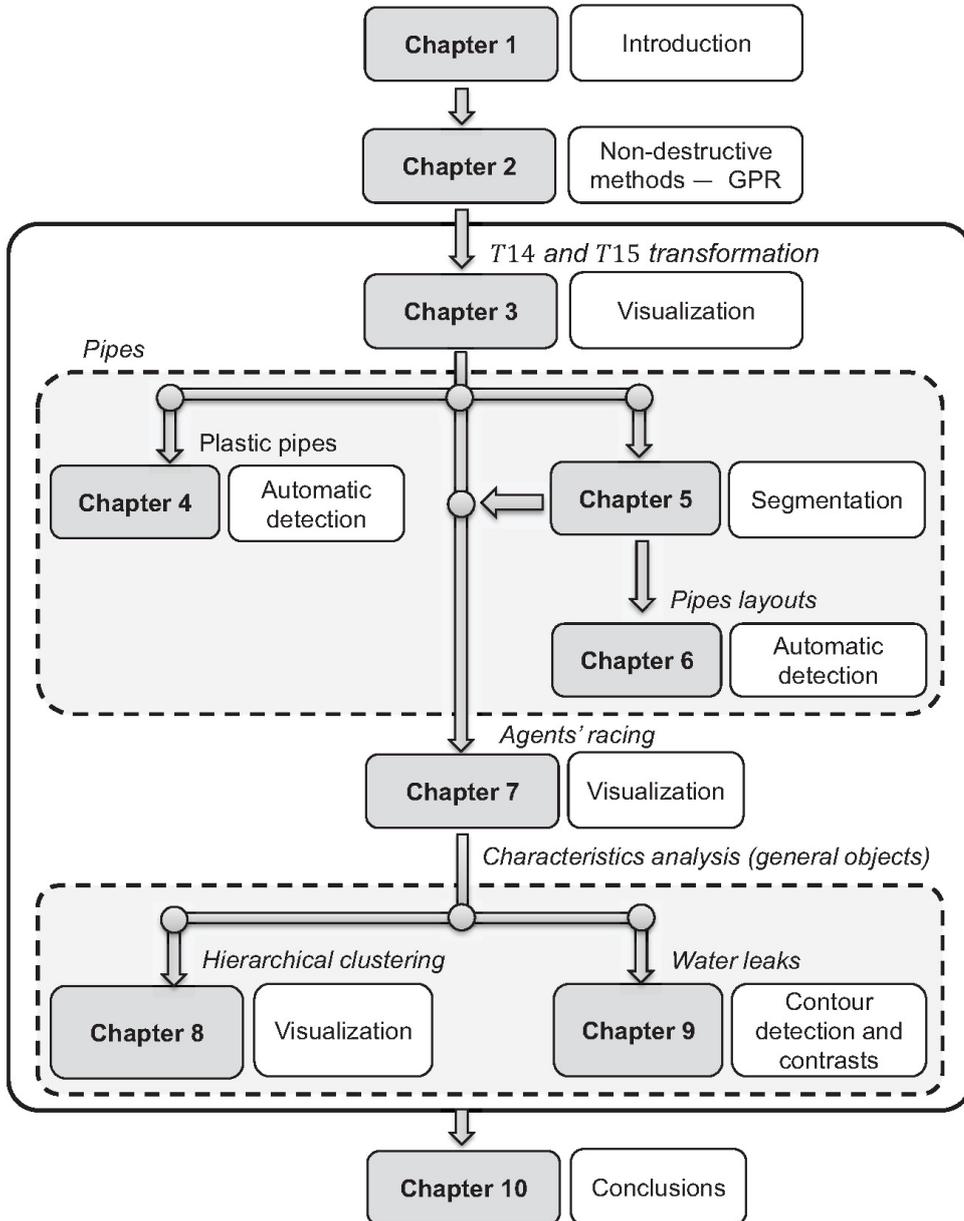


Figure 1.1: Chapter general scheme

In addition to the main body of this thesis, three appendices are of interest for future reference. **Appendix A** presents the formatting structure that

contains the data from surveys conducted with GPR, *.dzt – files GSSI format. **Appendix B** presents the instrumental settings used in the different tests performed. A third appendix presents the contributions in terms of conference papers and Journal papers written in the context of this thesis (**Appendix C**).

1.5 Contributions of the thesis

The main contribution of this thesis is to provide WSSs with methodological tools for analysis of their assets (primarily pipelines in the case of this thesis) using GPR as a non-destructive technique. In addition, we have sought that the presented developments promote image interpretation of GPR by non highly qualified personnel. The specific contributions produced by the development of this doctoral work are listed below.

- Planning, development, building and application of a tool that allows visualization of pipes and other buried objects from GPR images. Transformation into matrices T_{14} and T_{15} .
- Automatic identification of plastic pipes in GPR images through the implementation of multi-agent systems.
- Pipe feature extraction from GPR images and feature interpretation planning.
- The sampling space of GPR images is bounded by a proposal for segmentation of images obtained from the transformation given by T_{14} and T_{15} matrices that uses the Hough transform.
- Two approaches addressing model assessment and optimization for automatic detection of WSS pipes from GPR images have been proposed.
- A methodology for data pre-processing based on multi-agent systems based on the transformation matrices T_{14} and T_{15} and segmentation concepts

1.5. Contributions of the thesis

obtained by the Hough transform is proposed. This methodology reduces the amount of data to analyze, while preserving the characteristics of the buried objects embedded in GPR images.

- Clustering methodologies, based on unsupervised clustering, favoring the analysis and understanding of various GPR images of buried objects have been used.
- The proposed pre-processing approaches based on multi-agent systems have been extrapolated to water leakage, including an additional modification to obtain more understandable information.

Various contributions originated from this doctoral work have been disseminated in a number of national and international Conferences. This activity has helped consolidate the approaches to the underlying problems, and check the utility and acceptability of the proposals made. The main Congresses in which we have participated are given in (Table 1.1).

Table 1.1: Contributions to Congresses

Congress scope	Edition; Congress	
Water and networks	<ul style="list-style-type: none"> • XXVI, XXIV; Congreso Latinoamericano de Hidráulica • 3rd, 2nd; Annual International Forum on Water • 16th; Conference on Water Distribution System Analysis – WDSA • XIV, XIII, XII, XI, X, IX; Seminario Iberoamericano de Planificación, Proyecto y Operación de Abastecimiento de agua – SEREA • I; Seminario Iberoamericano de la Red IT-AGUA • III, II; Jornadas de Ingeniería del Agua – JIA • 8th; International Conference of European Water Resources Association – EWRA 	
	<ul style="list-style-type: none"> • X, IX, VIII, VII; Seminario Euro Latinoamericano de Sistemas de Ingeniería – SELASI 	
	Data analysis and artificial intelligence	<ul style="list-style-type: none"> • 7th; International Congress on Environmental Modelling and Software – iEMSs • 16th; Congrés Internacional de l’Associació Catalana d’Inteligència Artificial – CCIA
		<ul style="list-style-type: none"> • 2012, 2011, 2010; Mathematicall Modelling in Engineering & Human Behaviour
	Mathematical applications	<ul style="list-style-type: none"> • 2012, 2011, 2010; Mathematicall Modelling in Engineering & Human Behaviour

Also, a number of Journal papers have been produced within the thesis development. The main high impact Journal articles (quartile 1 and 2) are presented in Table 1.2.

1.5. Contributions of the thesis

Table 1.2: Contributions in high impact Journals

	Impact index	Category: Journal ranking/total in category
Ayala-Cabrera <i>et al.</i> (2014a)	(JCR 2013) 1.495	<ul style="list-style-type: none"> • Engineering, Electrical & Electronic: 97/244
Ayala-Cabrera <i>et al.</i> (2013a)	(JCR 2013) 2.048	<ul style="list-style-type: none"> • Chemistry, Analytical: 35/74 • Electrochemistry: 15/27 • Instruments & Instrumentation: 12/62
Ayala-Cabrera <i>et al.</i> (2013b)	(JCR 2013) 2.020	<ul style="list-style-type: none"> • Computer Science, Interdisciplinary Applications: 26/104 • Computer Science, Software Engineering: 11/110 • Mathematics, Applied: 16/252
Ayala-Cabrera <i>et al.</i> (2011a)	(JCR 2011) 1.444	<ul style="list-style-type: none"> • Geosciences, Multidisciplinary: 77/169 • Mining & Mineral Processing: 3/23
Ayala-Cabrera <i>et al.</i> (2011b)	(JCR 2011) 2.020	<ul style="list-style-type: none"> • Computer Science, Interdisciplinary Applications: 50/101 • Computer Science, Software Engineering: 26/109 • Mathematics, Applied: 40/247

An exhaustive list including all contributions is presented in Appendix C.

Chapter 2

Background

2.1. Non-destructive surveying methods

This chapter addresses the topic of non-destructive inspection methods and their application in searching and detecting buried objects, with emphasis on components (mainly pipes) of water supply systems (WSS). In addition, this chapter attempts to address the ground penetrating radar (GPR) from a global perspective in order to meet its general characteristics, composition, basic operating principles, and applications. The chapter deals with aspects of GPR image interpretation, introducing their associated main problems and some strategies that are usually approached to deal with them. Moreover, a number of conventional methods in processing GPR data are also presented.

2.1 Non-destructive surveying methods

The non-destructive surveying methods are framed in the field of geophysics. Geophysics studies the Earth in its composition, based on physical measures that are usually done on its surface or from sounding. When this study is related to relatively small areas and depths not exceeding a few of kilometres it is about Applied Geophysics, and the set of methods for this purpose they are the geophysical survey. However, when it deals with areas from meters to ten of meters with application to civil engineering, it is Geotechnics. This study is about *applied geophysics* to water engineering, specifically to water supply systems.

From applying geophysical studies, we can obtain information concerning the composition of the subsoil through any physical parameter measured at the surface, which ranges from mechanical waves velocity to changes in gravitational fields. These measurements try to detect differences in density through the intensity of the current associated to the ease of electric current propagation. This ability to quickly characterize subsurface conditions without disturbing the environment offers the benefit of lower costs, less risk and lower environmental impact, giving a better overall understanding of complex environmental conditions inspected.

Geophysical methods are based on measurements of physical properties of the

2.1. Non-destructive surveying methods

elements of the subsoil (rocks, sediments, water, cavities, objects, etc.) and can be classified into two types: passive and active. On one hand, passive methods are those based on the detection of variations of the natural properties of the ground, without introducing artefacts. On the other hand, active methods are those that study the behaviour of materials contained in the subsoil by creating artificial signals. These methods study the interaction of the signals with the materials in the subsoil. Each geophysical technique is based on the interpretation of the contrasts in specific physical properties of the subsoil (i.e., dielectric constant, electrical conductivity, and density).

The type of physical property which depends on a particular geophysical method determines the response and range of applications on it [Schrott *et al.* (2013)]. For instance, Table 2.1 shows a number of the geophysical methods employed more often and the physical properties of whom depend.

Table 2.1: Geophysical methods and physical properties of whom depend

Geophysical method	Physical property associated
- Gravimetry	- Density
- Magnetometry	- Magnetic susceptibility
- Seismic refraction	- Elastic model, density
- Seismic reflection	- Elastic model, density
- Electrical techniques (i.e. geoelectrical tomography)	- Electrical resistivity
- Spontaneous potential	- Electric potential difference
- Induced polarization	- Capacity and electrical resistivity
- Electromagnetic methods	- Conductivity
- GPR	- Conductivity and Permittivity

2.2. Non-destructive methods in the location of WSS components

The infrastructure of public services (utilities) and especially its failure, is currently an issue of great interest since its interruption causes project delays, cost overruns, damage property, and injuries which in some cases may lead to death, among others. The activities that interfere with the effective management of utilities can be attributed mainly to two causes; firstly, to incomplete, outdated or inaccurate data on the location of the system's assets. These failures may happen due to changes in managers, inadequate data efforts, relocation of public services and operational maintenance activities and operation. The second cause of failures is related to the lack of registers of the damage and their location [Li *et al.* (2015)]. This neglects the failure prediction on these systems. Table 2.2 shows a list of the main non-destructive methods applied by utilities and their associated limitations.

2.2 Non-destructive methods in the location of WSS components

There are several methods and techniques for inspecting and/or mapping pipes in WSS, and these may be classified as destructive and non-destructive processes. Studies, such as several performed by the US Environmental Protection Agency (US - EPA), underline the use of non-destructive tools as methodologies favouring technical management of WSS instead of other destructive testing methods [Thomson *et al.* (2009)]. The most commonly used non-destructive methods include: acoustic emission (AE); eddy currents (EC); electromagnetic methods (EM); impact-echo (IE); hammer sounding (HS); magnetic flux (MF); sonar methods (SM); ultrasonic testing (UT); pulsed induction methods (PIM); magnetic locators (ML); resistivity methods (RM); radio graphic testing (RT); pipeline inspection gauges (PIGs); and GPR.

2.2. Non-destructive methods in the location of WSS components

Table 2.2: Limitations of underground mapping technologies – Source: Li *et al.* (2015)

Technologies	Limitations
Magnetometer	<ul style="list-style-type: none"> - Can only detect ferrous utilities buried shallowly - Depth estimation is not considered realistic
Pipe and cable locator	<ul style="list-style-type: none"> - Can only detect metallic utilities - Possible interference by nearby metallic utilities
Terrain conductivity method	<ul style="list-style-type: none"> - Unable to detect utilities in high moisture environment - Massive amount of data are needed for processing - Depth estimation is not considered realistic
Infrared method	<ul style="list-style-type: none"> - Unable to detect utilities buried deeply - Not suitable for congested urban environment - Depth estimation is not considered realistic
Resistivity measurement	<ul style="list-style-type: none"> - May not be useful as a trace technique - Data setup and collection is cumbersome - May not be applicable for mapping paved areas
Elastic wave method	<ul style="list-style-type: none"> - Most utilities are too small to be detected - May be influenced by background noise - Depth estimation is not possible with acoustic emission
Radio frequency identification	<ul style="list-style-type: none"> - Attachment of tag to existing utilities is difficult - Negatively affected by moisture and metallic materials - Detectability and accuracy are relatively low
Ground penetrating radar	<ul style="list-style-type: none"> - Large signal attenuation in conductive soils - Highly skilled crews are needed for data interpretation

The diversity of techniques shows that each of them has advantages and disadvantages. For example, EC and EM are used for inspecting metallic pipes, while AE and IE are used for concrete pipes [Rizo (2010)]. PIM and RM are useful for locating metallic objects. GPR has been extensively used as a non-destructive methodology to analyse components and anomalies in WSS. The components more frequently analysed are pipes and, especially, metallic pipes. Only a few

incipient attempts have been conducted regarding leaks. GPR has many advantages that support the inspection of WSS components in relation to other non-destructive techniques. We check these advantages throughout this chapter and the subsequent development of the thesis. The benefits in using GPR have led us to select this non-destructive prospective method for the development of this doctoral work. In the next section we present generalities and particularities of GPR.

2.3 GPR overview

GPR is an electromagnetic geophysical technique for non-invasive exploration of the subsoil in its characterization and monitoring. GPR prospecting is a technology that can efficiently and accurately solve problems inherent to the study of subsoil, without resorting to invasive methods on the environment subject to inspection. This technique allows the detection of buried objects and the highly accurate identification of soil layers, in addition to gathering detailed maps of the subsoil. GPR can provide information about the exact location (position and depth) of any detected object [Busquet *et al.* (1996)]. Prospecting with GPR method usually requires a relatively short inspection time and can be done at reasonable cost. The main strength of this technique is the non-destructive character of its tests. One of the main features of GPR is the delineation of the contours of the different materials that make up the soil, provided there is sufficient contrast between the dielectric properties of the structures involved [Pueyo *et al.* (2005)].

GPR survey is based on the emission of electromagnetic pulses from the surface, which are propagated and reflected by the discontinuities encountered underground. This technique can reveal interfaces between the different materials, provided that there is sufficient contrast between the dielectric properties of the materials through environmental signals. The changes (specially the sharper changes) in the dielectric constant of the subsoil materials result in that one part of the energy is reflected [Robinson *et al.* (2013)]. The electromagnetic energy of the reflected pulse is measured on the surface by receiving devices.

2.3. GPR overview

Reflection and refraction wide angle or common midpoint (CMP) often are used to compute the relationships depth / resistance / velocity [Schrott *et al.* (2013)]. The time needed for the pulse to travel from the surface and return provides a measurement of the depth of a reflector, given the positions of the source and receiver and the propagation velocity of the pulses in the soil type [Pueyo *et al.* (2005); Bonomo & De la Vega (2010)]. Thus, a comparison between amplitudes of emitted and received pulses provides additional information about the composition of the subsoil [Götttshe (1997)]. These signals, called traces, are captured at fixed times and contain electromagnetic characteristics of the traversed medium. The subsequent accumulation of traces leads to what is known as radargrams, which are generally displayed using color or grey scales to generate images (Figure 2.1). This aspect is of interest because the analysis conducted in this thesis is based on images of GPR.

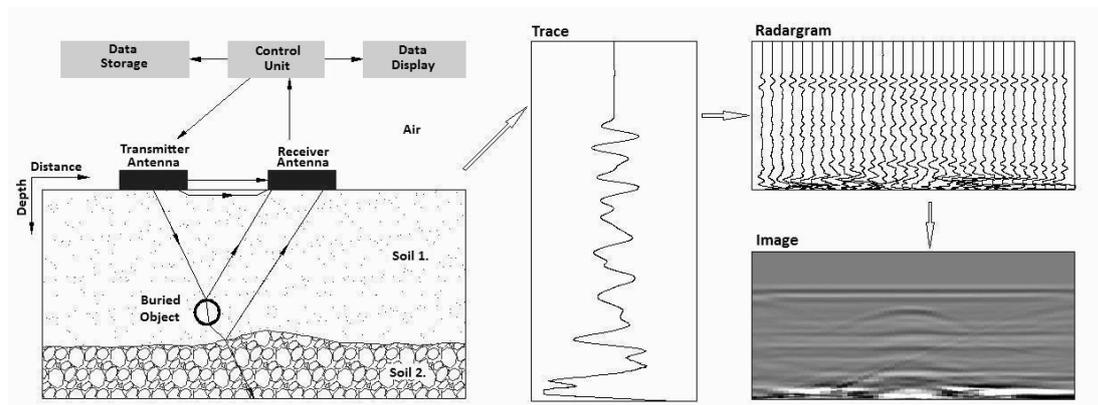


Figure 2.1: Image capture through GPR prospecting – Source: Ayala-Cabrera *et al.* (2011b)

As shown in Figure 2.1, the transmitting antenna sends a signal to the subsoil, whom is subject to reflection, refraction or diffraction. When the radiated wave find out electrical heterogeneities in the material properties of the subsoil, part of the energy is reflected towards the surface and another portion is transmitted deeper. The reflected signal is amplified, converted to its frequency spectrum and registered directly into a computer, where it can be displayed in near real time during the prospecting work. This type of equipment allows some preliminary

processing of data during the survey, allowing to obtain results directly in the field even though further data processing is needed [Pueyo *et al.* (2005)].

The data obtained by the equipment indicate variations of electromagnetic properties of subsurface materials. The characteristics of the waves received depend both variations of the electric permittivity, conductivity and magnetic permeability of the subsoil, and the characteristics of the emitted wave and the objects in their way. These data comprise the intensity of the received signal and the sending/arrival times of waves (usually double time - *TWT*) [Pueyo *et al.* (2005)]. GPR equipments radiate electromagnetic energy pulses to the subsoil by a transmitting antenna and measure twice the round-trip time of the reflected waves at the boundary between materials with different dielectric permittivity.

The GPR has a wide range of applications in various fields such as engineering, environment, geology, archaeological research, military, forensic investigations, the search for avalanche victims, and detecting landmines, among many others. These applications are the identification of a variety of common structures in geotechnical studies being of main interest (as was presented in the previous section) the buried infrastructure of public services (utilities), such as tunnels, canals, drinking water and wastewater pipes.

Moreover, there are many GPR applications in engineering. Studies of different soil layers and road pavement thickness [Rodriguez *et al.* (2015)]; GPR signals simulations by finite-difference time-domain (FDTD) to promote inspection of pavements and generate diagnoses of pavement failures [Tosti & Umiliaco (2014)]; inspection of concrete structures conditions by feature extraction in laboratory under controlled conditions [Queiroz *et al.* (2012)]; clay content in soils evaluation to prevent risks of damage to pavements [Tosti *et al.* (2013)]; inspection of stone arch bridges using numerical models to analyse and interpret the structures in GPR images captured [Solla *et al.* (2012)]. On pollution and environmental protection, there are works as Catapano *et al.* (2014) in oil spill detection and monitoring by GPR and microwave tomography.

2.3. GPR overview

This versatility on GPR applications is based on offering a high-resolution register, it is a quick method in terms of data capture and uses small shooting intervals that allows to consider it as an almost continuous method capturing information [Pueyo *et al.* (2005)]. The sending and receiving antennas are in the same unit or enclosure, usually armoured; even both transmission and reception functions can be performed by a single antenna for the most usual GPR equipments in urban hydraulics survey. This configuration of transmitter and receiver and the associated methodology are referred to as monostatic. Alternatively, the transmitting and receiving antennas can be placed in different boxes, in an assembly called bistatic [Bonomo & De la Vega (2010)].

The application of the GPR method in their monostatic or bistatic versions, with single fold, generally allows to obtain vertical surfaces and even 3D planes with high resolution of the reflectors buried in the subsoil. Alternatively, other methodologies can be used. In the case of multiple folding, a transmitting antenna and a receiving mounting are arranged in a bistatic system. These are usually located symmetrically on both sides of a central fixed point, and the distance between the two antennas is called offset. In some cases it is appropriate to use such borehole measurements. In this case, the recipient, the sender or both move vertically in holes in the ground, thus generating transmission images [Bonomo & De la Vega (2010)]. GPR is a technique that has the advantage of high resolution and the ability to detect non-magnetic objects through environmental signals [Bonomo & De la Vega (2010); Busquet *et al.* (1996)]. In comparison to other techniques, GPR has advantages such as portability, relatively low cost and weight, and ease of application at different land areas.

The main merits of GPR in detecting utilities include the following::

- Ability to detect and locate both metallic and non-metallic objects, which is a great advantage in comparison to other magnetic-based techniques [Li *et al.* (2015)].
- Under the premise that it is possible to obtain a good register of the utilities with GPR, we can estimate the depth properly, through techniques of data

processing. This is relevant, since the depth information obtained by the GPR allows the upgrade of existing 2D GIS systems to more accurately build 3D maps of utilities. 3D geospatial data are crucial to prevent failures during operation and maintenance activities [Li *et al.* (2015)].

- The GPR has a higher level of resolution than other technologies, with a possible resolution of the order of centimetres, which it allows its integration in systems such as real-time kinematic (RTK) GPS, which are able to offer high accuracy (and in three dimensions) of locating buried utilities [Li *et al.* (2015)].

However, despite the above advantages, the obtain field data images can be difficult due to problems such as limited coverage or images of noise in the absence of direct relations between the observed and the physical parameters to be reconstructed [Queiroz *et al.* (2012)].

As with all GPR investigations there are limitations depending on the uncontrolled and controlled groups of factors, and other underground structures in urban areas. For example, soil penetration depends on soil type and antenna type. Soil moisture, as well as highly clay soils, will quickly attenuate the radar signal and decrease performance (i.e. dry sandy soils are best). High frequency antennas, in the order of 1 to 2 GHz, produce the best resolution (i.e., can find small objects), but can only achieve penetration of one to few meters. Low frequency antennas, in the order or 10 to 200 MHz, can achieve penetration up to tens of meters, depending on soil conditions, but may not be able to locate small objects or small diameter pipes [Lane *et al.* (2000); Malvar & Cline (2004)], issues to consider when testing directly from the raw images.

2.4 Theoretical fundations of GPR

Maxwell's equations are the theoretical basis on which the phenomenon of the GPR technique is based together with the so-called constitutive equations, which relate the intensity of the electric and magnetic field to the electric displacement and magnetic induction, respectively, allowing correlate interactions between electromagnetic waves and media in which they propagate. This equations system contains three parameters that characterize an electromagnetic environment: σ , ε , μ . Conductivity (σ) is the measure of a material's ability to conduct electrical current. The dielectric permittivity (ε) provides a measure of the polarizability of the material in the presence of an electric field. Finally, the measure of the attractiveness of the magnetic fields is described by the magnetic permeability (μ) [Jol (2009)]. Below the Maxwell equations are presented.

$$\nabla \vec{D} = \rho_f \quad , \quad (2.1)$$

$$\nabla \vec{B} = 0 \quad , \quad (2.2)$$

$$\nabla \vec{E} = \frac{\partial \vec{B}}{\partial t} + \vec{M} \quad , \quad (2.3)$$

$$\nabla \vec{H} = \frac{\partial \vec{D}}{\partial t} + \vec{J} \quad , \quad (2.4)$$

where E and H represent respectively the electric and magnetic field; D and B the vector electric displacement and magnetic induction vector; ρ free charge density and J the current density. This system of equations in conjunction with continuity equations (Equations 2.5, 2.6, 2.7) allows the understanding of the interactions between electromagnetic waves and the media in which they propagate.

$$\nabla \vec{D} = \varepsilon \vec{E} \quad , \quad (2.5)$$

$$\nabla \vec{J} = \sigma \vec{E} \quad , \quad (2.6)$$

$$\nabla \vec{H} = \frac{\vec{B}}{\mu} \quad . \quad (2.7)$$

2.5 Electromagnetic parameters

The electromagnetic parameters (conductivity, dielectric permittivity and permeability) are those that define a medium to pass an electromagnetic wave. In most of the geological and geotechnical applications of GPR, electrical properties: electrical conductivity, σ and dielectric permittivity, ε tend to be the prevailing factors controlling GPR responses [Ahmadi *et al.* (2015); Schrott *et al.* (2013)]. In the magnetic permeability (μ), variations are usually trivial; it is how in most cases and simulations the specially referred parameters for GPR are σ and ε [Ahmadi *et al.* (2015)]. Following is a brief description of these parameters, given their importance in GPR image interpretation.

2.5.1 Conductivity

The conductivity of a medium provides the measure of the response of the free charges in the presence of an external electric field, being the proportionality factor between the free field applied and the current density due to the movement of these free charges. That is, it provides a measure of the ability of a material to conduct electrical current and is described according to Ohm's law, whose formulation is presented in Equation 2.6.

The conductivity units of measure in the IS are ohms/meter or the equivalent of Siemens/meter. The conductivity of a medium is the inverse of resistivity.

2.5. Electromagnetic parameters

In general we can distinguish between conductive materials, semiconductors and insulators.

2.5.2 Dielectric permittivity

The permittivity is a constant measures the polarizability of a material in the presence of an electric field. It provides a value of the static response of the material when it is in the presence of an external electric field. The dielectric permittivity of vacuum is $\varepsilon_0 = 8,854 \times 10^{-12}$ F/m (farad/meter). The relative permittivity of a material ε_r is the ratio of the dielectric permittivity and the vacuum, $\varepsilon_r = \varepsilon/\varepsilon_0$, and is also called dielectric constant. For most of the materials that can be found in the subsoil when performing electromagnetic surveying, relative permittivity values are found between 1 (value for air) and 81, the latter value is the dielectric constant of water at 20° C temperature [de la Vega (2001)]. Thus, we can obtain the dielectric permittivity of a material multiplying its dielectric constant (or relative permittivity) by ε_0 .

2.5.3 Magnetic permeability

Magnetic permeability (μ) is the variable that relates the magnetic induction with the electric field. It is measured in Henry/m and is defined as the product of the relative magnetic permeability of the material and the magnetic permeability in vacuum.

$$\mu = \mu_r \cdot \mu_0 \quad , \quad (2.8)$$

where: μ_r is the magnetic permeability related to the material and μ_0 is the magnetic permeability in vacuum ($\mu_0 = 4\pi \times 10^{-7}$ H/m) [Martínez & Rodríguez (2011)].

2.5.4 Wave propagation velocity

Electromagnetic waves travel in a medium at a rate depending on the relative dielectric constant of the ground, the velocity of light in vacuum, and the relative magnetic permeability. The velocity of the electromagnetic wave (v_m) in a certain material is given by:

$$v_m = \frac{c}{\sqrt{(\varepsilon_r \mu_r / 2)((1 + P^2) + 1)}} \quad , \quad (2.9)$$

where P is the loss factor, such that $P = \sigma / \omega \varepsilon$ y σ is the conductivity, $\omega = 2\pi f$ (f is the frequency), and ε is the absolute dielectric constant. In a material with a low loss factor, we have $P \rightarrow 0$; then the electromagnetic wave velocity is given by:

$$v_m = \frac{c}{\sqrt{\varepsilon_r}} = \frac{0.3}{\sqrt{\varepsilon_r}} \text{ m/ns} \quad , \quad (2.10)$$

where c is the velocity of light (0.3 m/ns) [Tavera *et al.* (2008)].

The depth of objects can be estimated using Equation 2.11 and considering the properties of the surrounding soil as a homogeneous and relatively consistent with respect to the wave velocity [Li *et al.* (2015)].

$$d = \frac{TWT \cdot v_m}{2} \quad , \quad (2.11)$$

where d is the depth of the object, TWT is the two-way time measured when the GPR antenna is directly over the object, and v_m is the propagation velocity of the electromagnetic wave in the medium. The propagation velocity can be calibrated by a few test holes [Li *et al.* (2015)] or can be obtained by analysing the obtained GPR images [Schrott *et al.* (2013)].

2.5. Electromagnetic parameters

2.5.5 Reflection and refraction

GPR operation is based on the detection of reflections and dispersions of the emitted signals. When the emitted energy reaches a discontinuity in the medium electromagnetic parameters, the phenomena of reflection and refraction occurs, which is related to the change in the direction of displacement of the signal (i.e., the wavefront refracted) in accordance with Snell's law [Jol (2009)]. Figure 2.2 shows schematically the mentioned phenomena of energy reflection and refraction.

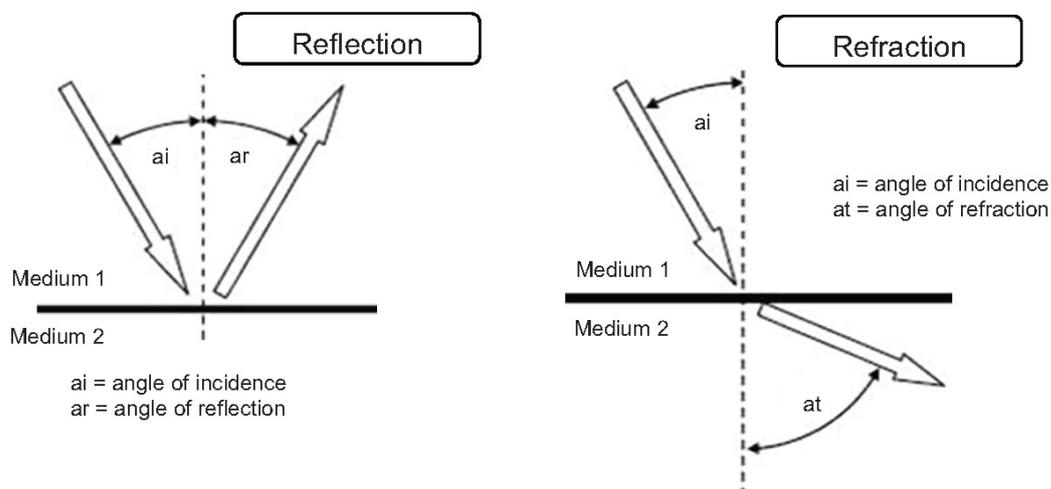


Figure 2.2: Reflection and refraction of incident energy on a horizontal contact

A wave is “reflected” when meets a change in the material and its properties vary. The reflection angle is the same as the incidence angle, generally with reference to the perpendicular to the surface, or the normal line. Furthermore, the change of direction of the signal is given when switching from one medium to another with different velocity (different index of refraction), with a different incidence angle to zero degrees with respect to the normal of the interface.

2.5. Electromagnetic parameters

The impedance of an electromagnetic field is the ratio of the electric field and the magnetic field. We can define an impedance to the incident electromagnetic field, η_1 , coinciding with the reflected magnetic field, and another for the refracted electromagnetic field, η_2 , i.e., transmitted:

$$\eta_1 = \left\{ \frac{\vec{E}_i}{\vec{H}_i} = \sqrt{\frac{\mu_0\mu_r}{\varepsilon_0\varepsilon_r}} \right\} = \left\{ \frac{\vec{E}_r}{\vec{H}_r} = \sqrt{\frac{\mu_0\mu_r}{\varepsilon_0\varepsilon_r}} \right\} , \quad (2.12)$$

$$\eta_2 = \left\{ \frac{\vec{E}_t}{\vec{H}_t} = \sqrt{\frac{\mu_0\mu_{r2}}{\varepsilon_0\varepsilon_{r2}}} \right\} . \quad (2.13)$$

The coefficients of transmission and reflection Fresnel energy can be computed from the expressions of the impedances as the ratio of the incident field and the refracted field in the first case, and the ratio of the incident field and the reflected field in the second case. If ai is the incidence angle of the wave on the flat surface, and at is the refraction angle, these coefficients take the following expression:

$$R_{1 \rightarrow 2} = \frac{\vec{E}_r}{\vec{E}_i} = \frac{\eta_2 \cos(at) - \eta_1 \cos(ai)}{\eta_1 \cos(ai) + \eta_2 \cos(at)} , \quad (2.14)$$

$$T_{1 \rightarrow 2} = \frac{\vec{E}_t}{\vec{E}_i} = \frac{2\eta_2 \cos(at)}{\eta_1 \cos(ai) + \eta_2 \cos(at)} , \quad (2.15)$$

where $T_{1 \rightarrow 2}$ is the Fresnel refraction coefficient between the upper (1) and lower (2) medium, $R_{1 \rightarrow 2}$ is the Fresnel reflection coefficient for contact of the two media, E_i is the incident field, E_r is the reflected field and E_t is the refracted or transmitted field.

These expressions are simplified in GPR studies as the system operates with a very small angle reflection. It simplification may consider incidence and reflection (in the case that the reflective surface is flat) angles of zero degrees which

2.5. Electromagnetic parameters

allows to work with normal incidence. Under these conditions, the coefficients of reflection and transmission only depend on the relationship between complex impedances. Working on non-magnetic media, an adequate approximation to most of the materials in which GPR approaches are applied is the following expression:

$$R = \frac{\eta_1 - \eta_2}{\eta_1 + \eta_2} = \frac{\sqrt{\varepsilon_{r1}} - \sqrt{\varepsilon_{r2}}}{\sqrt{\varepsilon_{r1}} + \sqrt{\varepsilon_{r2}}} , \quad (2.16)$$

$$T = \frac{2\eta_2}{\eta_1 + \eta_2} = \frac{2\sqrt{\varepsilon_{r2}}}{\sqrt{\varepsilon_{r1}} + \sqrt{\varepsilon_{r2}}} . \quad (2.17)$$

The sum of the reflection coefficient (which represents the reflected energy percentage of the incident energy) and the refraction coefficient (which represents the percentage of power transmitted of the incident energy to the second medium), is the unity. From the above expressions, it follows that the greater the difference between the electromagnetic parameters of the media, the higher the reflection coefficient; that is, the greater the contrast between the two media in contact, the higher percentage of the incident energy will be reflected at the discontinuity and, therefore, we have a lower percentage of power transmitted to the lower medium. High values of R in a GPR study implies the possibility to observe easily the reflected wave in the registers. In addition, the penetration decreases and the reflections produced in subsequent contrasts are of smaller amplitude as the percentage of refracted energy is smaller, which contributes to a lower wave amplitude in subsequent reflections [Tavera *et al.* (2008)].

2.5.6 GPR equation

The most commonly used GPR equipment in urban hydraulics consists of a transmitter and a receiving antenna mounted in the same box. This configuration of transmitter-receiver and its associated methodology is referred to as monostatic. In this PhD thesis, monostatic antennas are used to obtain GPR data. The main equation used for monostatic GPR equipment is shown below. It gives the ratio (Q) between the received power, P_R , and the emitted power, P_E (see Equation 2.18)

$$Q = \frac{P_R}{P_E} = \frac{\lambda^2}{4\pi} \cdot G^2 \cdot \frac{S}{(4\pi r^2)^2} \cdot e^{-4\alpha r}, \quad (2.18)$$

where

λ : wavelength in the air.

G : gain of the antenna.

S : effective surface of the reflector.

α : attenuation constant.

r : distance to the reflector.

From Equation 2.18, it is possible to obtain the relationship between wave amplitudes (Q_{Amp}) received, A_R , and emitted, A_E , as follow [Göttshe (1997)]:

$$Q_{Amp} = \frac{A_R}{A_E} \propto \sqrt{\frac{1}{r^4} \cdot e^{-4\alpha r}} = \frac{e^{-2\alpha r}}{r^2} . \quad (2.19)$$

2.6 GPR interpretations

GPR images are complex and difficult to understand, so the GPR image interpretation is inherently subjective and depends on the knowledge, skill and experience of the operator [Lu *et al.* (2014); Robinson *et al.* (2013); Al-Nuaimy (1999)]. Once GPR data is acquired, there are usually tackled by both manual and automatic processes. The analysis and interpretation of large volumes of data generated by GPR surveys are extremely difficult and often become a bottleneck that influences into a suitable application of the technique. The GPR is one of the most reliable tools for obtaining information of buried objects. Since its commercial release, there have been multiple efforts to overcome one of the major limitations of this technology such as the “human factor” in the interpretation of results, given that the operators have to analyse and interpret a large number of GPR images. This requires a huge effort by the operator and it is not always productive in obtaining results, besides being an inefficient process in terms of time-consuming [Simi *et al.* (2008)].

Much of the success of operator skill and interpretations of GPR comes from the combination of observations extracted from the radar images, with the knowledge of the structure and composition of the materials and the medium, as well as the characteristics of the buried objects [Al-Nuaimy (1999)]. Nevertheless, many interpretations based on alternative models and other different data may miss key components that are visible in the GPR profiles [Robinson *et al.* (2013)]. The task of identifying buried objects using GPR images can be performed in two ways. Firstly, by processing raw images captured by GPR. Secondly, by improving and customizing the hardware used in the GPR system, adapting it to the user’s application needs. The first option is possible using various techniques of image processing, while the second option is economically unfeasible. The important task is always to improve the image visualization in a convenient manner, such that allows to identify objects in the images [Singh & Nene (2013)]. The automatically detection and identification of objects becomes a genuine challenge. The automatic GPR image analysis is an interesting subject because it

involves the use of pre-processing, detection and classification tools with the objective of near-real time in approaching data interpretations. The actual use of the GPR system requires tools to use “on the ground”, with a very short time difference between image capture and the subsequent analysis and interpretation of soil composition (objects, pipes, cables, infrastructure, etc.) to update rapidly (possibly in a few seconds), for example, layers of geographic information systems [Belotti *et al.* (2002)].

GPR images interpretation are often based on shapes obtained after reflections; such as lines and hyperbolae, among others. The shape of the overall response of GPR for buried objects (especially in specific objectives and horizontal cylindrical objects) is a hyperbola [Ahmadi *et al.* (2015)]. One of the main strategies used in the interpretation of GPR images is given by the identification and recognition of hyperbolae. We can usually estimate both the spatial position and the depth of the buried object registered in the GPR images [Janning *et al.* (2014)]. At first, no knowledge about how many hyperbolae exist and where they are in a GPR image. In addition, there is the added problem of additional reflections that often appear under (or near) the initial reflector. This happens since the GPR signal has a complex waveform and the single reflector may appear as a pair or more reflections. Sometimes this is overlooked, so there is more than one interpretation for the same GPR reflector profiles [Robinson *et al.* (2013); Janning *et al.* (2014)].

GPR images contain valuable information that can be useful in searching for buried objects. However, the source of all the signals registered on the GPR images should be questioned always. Keep in mind that not all the registered responses in GPR images are originated in the subsoil, as part of the energy is propagated through the air and is also reflected (and reach the GPR receiving antenna). This energy can be generated from the surface due to the possible antenna proximity to objects such as fences, telephone poles and overhead wires, to name a few examples. Such reflections can appear as additional hyperbolae on a GPR image that can be associated with buried non-existent objects. Therefore, the areas must be previously inspected to prevent these events (known as

2.6. GPR interpretations

“airwave”). If a potential source is found, it must firstly studied the its influence on the images, in order to minimize its effect and do not lead to erroneous interpretations of GPR images. There are typically used armoured antennas to prevent these events; however, it should be noted that this type of shielding is not effective in all cases studied [Robinson *et al.* (2013)].

Furthermore, GPR images of GPR may contain other various elements given by unwanted effects such as: coupling effect of antenna, ground echo, scattering problem, noise, and clutter data [Singh & Nene (2013)], among others. Noise and clutters have become a major concern for researchers.

The scattering problem is given by singularities or non-uniformity data of the electromagnetic impedance medium. The effect occurs when the dimensions of these non-uniformities are much smaller than the electromagnetic wavelength [Tosti *et al.* (2013)].

Noise is generally due to multiple contaminant sources of the signal during the formation of the initial wave. Noise can be caused by external sources or other radio-wave-emitting devices such as cell phones, citizens band radio - CB radio; or poor high-frequency electronic design, among others [Harrison (2005); Singh & Nene (2013)]. Noise can intersect with the different hyperbolae belonging to objects or be tangent to them [Janning *et al.* (2014)]. This makes difficult the interpretation of objects in GPR images. An additional type of noise typically found in GPR images is the speckle noise [Belotti *et al.* (2002)]. This noise is caused by random fluctuations in the signal and forms granules added to the image. Speckle noise is critical in GPR image analysis, as it significantly degrades the image quality, which may lead to have wrong or misguided interpretations [El-Mahallawy & Hashim (2013)].

A common form of noise in GPR images is the coupling effect of antenna. This happens when the GPR transmitter emits electromagnetic waves, and some of them are directly received by the GPR receiving antenna. This introduces noise in the image, shown as horizontal lines in the image [Nagashree *et al.* (2014);

Robinson *et al.* (2013)]. These lines cover any reflection with an equivalent depth to the half of the antenna separation [Robinson *et al.* (2013)].

Clutter data is caused by events due to actual reflections from the initial wave, since the GPR is designed such that the receiver is continuously picking up signals even when no longer the transmitter sends pulses. As a result, the receiver picks up the transmitted wave and the reflected signals overlaps [Singh & Nene (2013)], reducing the ability to capture the target signal reflections [Harrison (2005)] and minimizing the potential visualization objects in GPR images [Singh & Nene (2013)].

The above mentioned issues mask the objects contained in the images, thereby hampering their identification. Discerning between noise, clutter data, medium's effects, and other effects of the system and subsoil requires great skill by the operator side. Operator experience becomes important in time to data cleaning and its subsequent interpretation [Al-Nuaimy (1999)]. Therefore, these issues lead to an increased demand for developing accurate and automated detection and identification object techniques in noise contaminated images [Lu *et al.* (2014)]. Data pre-treatment and processing prior to any interpretation becomes a valuable tool which promotes the identification objects captured by GPR images of objects, and their characteristics. Pre-processing GPR images mainly relates to the elimination or reduction of the effect of certain characteristics (as mentioned previously) that suppress or mask the target response [Nagashree *et al.* (2014)].

In the GPR image interpretation, given the types of targets to identify, highlight interpretation strategies based on hyperbola recognition for studies of soil stratification and object recognition, such that:

- In soil stratification studies, a strategy generally applied involves making a first correction on the profiles regarding topographical conditions. Subsequently, data are initially visualized in images, these images identify the main reflections. The ends of the reflections are usually identified and characterized. The ability to identify representative reflection patterns in GPR

2.7. GPR data processing

images allows to locate features such as bedding schemes, bedsets and cross-stratification on soil stratification studies. Another common interpretation strategy on this type of study is to identify reflection patterns with similar geometry; this strategy has been called *radar facies* [Robinson *et al.* (2013)].

- *Object recognition* in GPR images. The main interest of this thesis is *Pipe recognition*, where there are included two subjects: (a) pipe localization, and (b) pipe type estimation [Janning *et al.* (2014)].
 1. The first requires the identification of multiple reflections belonging to the same buried pipe [Janning *et al.* (2014)], and is the first in-depth reflection, which is what is usually sought. This allows to obtain the spatial and depth location of the pipe or any buried object [Singh & Nene (2013)].
 2. The second includes aspects such as the estimation of size, shape, and material, among others [Janning *et al.* (2014); Singh & Nene (2013)].

By applying this last strategy, we want to mention to Unterberger and Ulriksen who were among the first to present works of greatest importance on buried pipes detection. Ulriksen, from the University of Lund (Sweden), described in his doctoral thesis [Ulriksen (1982)] the number of ways in which it is possible to visualize buried pipes depending on their content, material and thickness.

2.7 GPR data processing

Registers obtained from surveys conducted with GPR can be directly interpreted in some cases. However, given the immense variability of media and their constituents, it is often necessary to process the data with the aim of improving the signal/noise ratio or highlight some event (such is the case of buried pipes). Processing data contributes to improve the registering process, but its improper or excessive use may lead to loss of important information or interpretation, subsequently reaching to incorrect conclusions. When the searched structures have an

important relative dielectric constant, the obtained profiles can be subjected to an intense signal filtering (background removal, DC stacking, tracking) thereby obtaining typical anomalies as the usual hyperbolae (hyperbolae of reflection). However, in cases where there are vertical development anomalies (anomalies caused by multiple reflections) the relative dielectric constant is much higher and can not be discriminated directly to each other anomalies. Significant part of information on the flat reflective elements could go undetected with the application of an intense filtering and the ability to discriminate between hyperbolic anomalies and vertical development may be lost. We shall handle with care applying filters to profiles given that potential information lost and the diminishing of the discrimination ability when there are elements with high relative permittivity contrasts.

Anomalies caused by metallic elements (multiple reflections) are easily identifiable with a shallow filter in the study of profiles, while the discrimination capacity decreases with the intensity of filtering [Pueyo *et al.* (2005)]. The success of the technique depends on the natural characteristics of the soil, the analysed structures features, the equipment used, and the correct interpretation of the data. A cause-effect correlation can be enhanced with the generation of synthetic radargrams analogous to the synthetic seismogram of common use in seismic reflection [Busquet *et al.* (1996)]. There is also possible to support the data visualization by the application of qualitative assessments, various degrees of structure identification. A number of the most important procedures used for this purpose are presented below.

2.7. GPR data processing

2.7.1 Filters

The term filtering refers to a wide variety of techniques aimed to selectively modifying the signal received amplitude or frequency. Filtering can be applied to remove noise and to improve the quality of the image visualization. The filtering process can be applied to the raw image for background removal [Singh & Nene (2013)]. This type of procedure could be classified into two groups: horizontal and vertical filters.

- **Horizontal filters – spatial.** It is a treatment horizontally applied to the register. The aim with the application of these filters is to approach a background noise removal from the radargram. This noise may be of high frequency (registered with few traces), generated by sliding the antenna by an uneven ground. This methodology is one of the most applied in the treatment of GPR data registers. For this type of processing, frequent contents are referred to the horizontal length of the anomalies. An anomaly of long wavelength (and therefore low frequency) is registered in a high number of traces, while a small wavelength anomaly (high frequency) is registered in a few traces. Such treatment may be performed with two types of filters: the so-called finite-response filter (FIR) and the so-called infinite-response filter (IIR).
- **Vertical filters – temp.** There are treatments applied in the vertical or temporal direction on each of the traces of the registers. They can be used to make corrections to baseline or to eliminate electronic noise that is inherent to each trace. These filters consist on multiplying each signal by an operator who is the unity in the frequency band which we are interested to conserve. Vertical filters tend to zero in the frequency band that does not matter. When these filters are applied it is appropriate to conduct studies of the frequencies before and after the application thereof to verify that it was successful and to avoid losing information that could be relevant in the study. They can be also high pass or low pass; high pass filters prevent low frequencies are registered, usually those that are below 10 MHz, which are

considered as noise of the system. Therefore, they aim to eliminate low-frequency noise, which largely depends on the used part of the antenna, the wiring and its length. Moreover, lowpass filters are used to remove higher than expected frequencies. With the application of these two filters is demarcated the study frequency range.

- **Gaussian filters.** Gaussian filters are a class of low-pass filters, based on the Gaussian distribution function of probability. The 2D Gaussian function for a space is given by:

$$g(x, y) = \frac{1}{2\pi\sigma^2} \exp^{-\frac{x^2+y^2}{2\sigma^2}} , \quad (2.20)$$

where x is the distance from the origin in the horizontal axis, y is the distance from the origin in the vertical axis, and σ is the standard deviation. This filter is usually applied as image pre-processing. Then, these images allow to discriminate between objects and background [Singh & Nene (2013)].

2.7.2 Predictive deconvolution

It is a procedure that aims to improve the vertical resolution in radargrams. It is a technique that has been applied to GPR surveying and is firstly used in seismic studies. It applies because it is believed that the medium by which the signal propagates acts as a its filter. Thus, the registered trace in the radargram is actually a convolution of the initial emitted pulse. The objective is to make a deconvolution in order to remove the medium effects on the initial pulse. Thus, the aim is to transform the radargram field in a register where signals only represent the primary reflection.

2.7. GPR data processing

2.7.3 Migration

This procedure is performed to move to its actual position the effects of the registered inclinations of reflections and to collapse the produced and registered diffractions. To use this treatment is necessary to estimate the average velocity of propagation through the medium. It can be known from the generated hyperbola diffractions. It basically is computing the equation of each identified hyperbolae, moving the reflector to the hyperbola's vertex, on the basis of returning signals to their "true" position. In this case, the reflection point is retained and the electromagnetic energy is moved towards this point. This technique favours a more accurate localization (in images) [Singh & Nene (2013)]. However, as the hyperbola equation depends on the medium propagation velocity, an error in the determination of this parameter can cause that such hyperbola does not completely collapse (the estimated velocity is lower than the real). On the other way around, it is possible to have the known effect as over-migration (the estimated velocity is higher than the real).

2.7.4 Fast Fourier Transform – FFT

This widely known data processing is based on the realization of a frequency analysis of the registers. This allows to study the evolution of the signal parameters as a function of the frequency. It is often used to obtain the trace frequency spectra, in order to better selection of frequency filters to be applied, or to check the application of any of the above filters has on the frequencies.

2.7.5 Hilbert transform

This mathematical procedure expresses the relationship between the real and the imaginary part of the signal, with which can be obtained in the registers amplitudes envelopes, and the instantaneous phase and frequency. The latter can be used as type of signal filtering indicators during its propagation through

the medium, providing information on the medium behaviour in front of the electromagnetic signal applied. Obtaining a trace envelope provides information about the instantaneous amplitude of the signal and therefore about the energy involved in each event.

2.7.6 Signal correction procedures

2.7.6.1 Background removal

In raw data radargrams usually are horizontal banding typically of high intensity that partially or completely hide signs of interest and should be eliminated. These background signals can be caused by the direct reception of signals from the transmitter, which especially happens when the unit uses GPR shielded antennas. Another instances of cause of this issue are: due to the reflection of signals at the interface between air and soil, because of reflections produced by any object foreign to the subsoil, by reverberations of diverse origin, etc.

A simple technique to remove these horizontal bands is to achieve an average trace computed from all traces that make up the radargram. This average trace is subtracted from each later on. As an alternative to calculate the average trace, sometimes it is considered only a selection of traces. Importantly, background removal from the subtraction of an average trace tends to eliminate not only horizontal unwanted signals but also reflections that may occur in horizontal or near-horizontal layers, which could incur in relevant information losses. It is partly for this reason that the method often applies in only an area or window radargram.

2.7.6.2 Gain

This is a modification (usually amplification) that is introduced to the signal. This modification does not have to be constant throughout the length of the

2.7. GPR data processing

trace, which can vary with the propagation time. This is how data as signal wave amplitude is multiplied by the gain function. The choice of the applied gain function during the data acquisition depends on how the data is treated. Gains may be designed to enhance the data visibility for the interpretation in-situ. It can also be possible to set a constant value throughout the capture for further processing, but the in-situ modification of the gain is the more used alternative among GPR operators [Al-Nuaimy (1999)]. The function aims to amplify events associated with deep or little energetic reflections, increasing amplitudes which by their small size, would not be taken into account in the interpretation of the registers. It is used to correct the effects of declining energy produced during the signal propagation in the medium. The gain function of a GPR equipment is defined by the following expression:

$$G = E_{ant}D = \frac{4\pi}{\lambda^2}A_{eff} \quad , \quad (2.21)$$

where

D : directivity.

A_{eff} : effective antenna aperture.

E_{ant} : antenna efficiency.

λ : wavelength.

Chapter 3

GPR images as tools for visualizing WSS components

In this chapter, we propose a methodological tool to examine layouts and to reveal concealed characteristics of WSSs. For this purpose, we use underground images obtained with GPR as a method that does not alter the system conditions and environment characteristics (non-destructive methods) to improve the visualization of objects (if any) in GPR images. The main objective is to assess the viability of finding buried objects registered by GPR images, and, specifically, WSS components. The main idea is to improve the identification of shapes that can help extract patterns representing and thus detecting those objects. This chapter focuses on pre-processing data from GPR traces and its subsequent transformation into images. The underlying reason for performing this task is to discriminate noise sources that generate spurious images, so that trained intelligent systems efficiently locate buried pipes. The general approach proposed in this chapter, within the thesis framework, is presented in Figure 3.1.

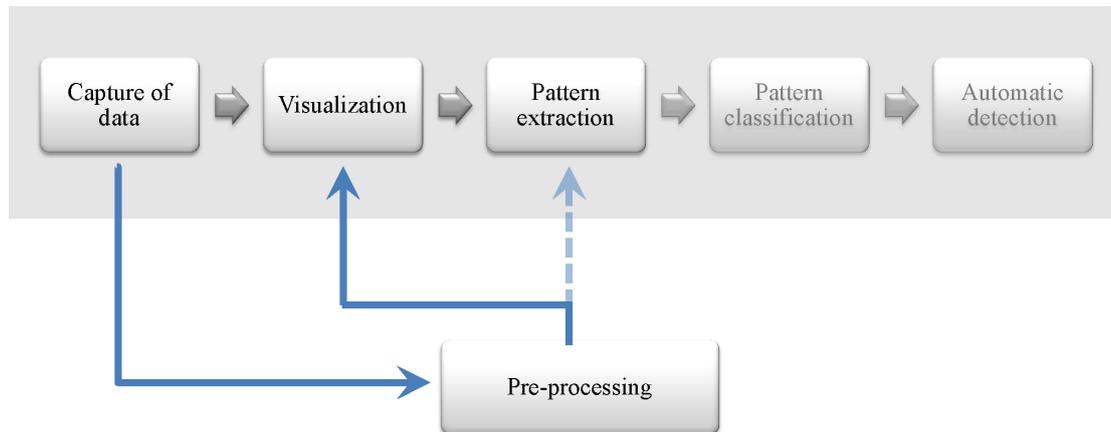


Figure 3.1: Chapter 3 general approach

This chapter contains the most relevant aspects of the *Master thesis* presented by the author [Ayala-Cabrera (2009)]. We have to mention that a preliminary version of this approach was presented in the Congress *IX SEREA* [Ayala-Cabrera *et al.* (2009)]. Then, a substantive extension of the approach was published in a high impact Journal, *Mathematical and computer Modelling* [Ayala-Cabrera *et al.* (2011b)]. The importance of the approach presented in this chapter hinges on the fact that it serves as the starting point for a number of approaches, processes and

3. GPR IMAGES AS TOOLS FOR VISUALIZING WSS COMPONENTS

data analysis techniques that are presented in subsequent chapters. It is worth to note here that the processes presented in this chapter has been fully developed using MatLab. The study is performed based on the selection of the wave amplitude as the starting point. It is well-known that this parameter provides a differential response to the passage of the signals emitted by the GPR through the different inspected materials. We present the principles and procedures used to obtain the tool and its application in laboratory tests. These tests have been performed in a laboratory layout under controlled conditions, using pipes of four different materials buried in dry soil. The objective is to generate the basis and identify the parameters for training intelligent systems that are able to characterize different components of urban WSSs. Promising results are obtained with this methodology, which is non-subjective and repeatable, and visualizes buried pipes efficiently. This tool will help WSS managers to obtain a more accurate picture of their WSS and so, offer better service to users.

This chapter is organized as follows: in this section we have presented a summary of the work developed in the chapter. Section 3.1 presents an introduction to the visualization of objects in GPR images. The next section (Section 3.2) presents the principles and development of the visualization process we have devised through the so-called $T14$ and $T15$ matrices. The experimental part of this chapter, containing a simple case and a complex case developed in laboratory controlled conditions, is presented in Section 3.3. Comments on the implementation of the processes are provided in Section 3.4. Finally, the summary and pertinent comments on the work developed in this chapter appear in Section 3.5.

3.1 Introduction to methodologies of visualization of objects in GPR images

This chapter proposes a first step towards developing a tool for identifying characteristics of unknown component in urban WSSs. It is based on the use of GPR as a non-invasive technique enabling buried objects to be detected without altering the environment. GPR technology development started in 1904, and there have been many studies regarding GPR exploration [He *et al.* (2009)]. Many efforts have sought to improve GPR interpretation and obtain faster image identification and interpretation.

Most surveys conducted with GPR base their interpretation on the forms or shapes observed in the radargram, which are determined by the contrasts produced by the signal first emitted and the received by the equipment. Contrasts are based on the characteristics of the materials. In other cases, the application of multiple processing steps provide successful results. However, result interpretation, either directly or by applying methods of data processing, must be performed with particular care, since inadequate application can lead to misinterpretation, yielding erroneous conclusions.

The methodologies applied in object recognition using GPR images are associated with the following tasks: (1) object detection and location, (2) object material recognition, (3) object size estimation, and (4) object shape estimation [Pasolli *et al.* (2009b)]. There have been interesting works on GPR detection in landmine searching [Gader *et al.* (2000); Chan *et al.* (2007)], archaeological exploration [Cukavac *et al.* (2008)], and steel for the reinforcement of concrete [Shaw *et al.* (2005)], among others.

Even though information retrieval by GPR is worthwhile, the huge volume of generated information and the interpretation of data usually require high levels of skill and experience [Thomson *et al.* (2009)]. The principle behind GPR in detecting buried pipes exploits the differences between the dielectric properties of the pipes and the surrounding soil. As a result, steel or cast iron pipes can

3.1. Introduction to methodologies of visualization of objects in GPR images

be easily detected. However, if the pipe is made out of a material with dielectric properties close to those of the soil, then detection may be difficult [Loulizi (2001)], as is the case of plastic pipes.

3.1.1 GPR role in the analysis of WSS components

The GPR has already been used as a non-destructive method for analyzing WSS components and anomalies. Pipes are the most analyzed components, especially metallic pipes. In the case of water leaks identification is in a very budding stage; details regarding the use of GPR for water leaks are presented in Chapter 9 (Section 9.1).

Various procedures have been adapted to WSSs from other prospection image interpretation methods. Several authors have addressed GPR image pre-processing, trying to improve visualization of buried objects. This process is performed to clean the noise in GPR images. This noise can be originated by the process itself, or by external signals embedded in the underground material converging on the signal path. The aim is better visualization of the contrasting images and observing features of the objects under the ground, thus obtaining clean images for locating objects. This is the way in which image pre-processing in WSSs is performed by the application of procedures adapted from other methodologies. This is the case of the application of Background Removal, Gradient Filter, and Migration, used by Olhoeft (2000) for cleaning GPR images to find metallic pipes. Works such as the one by Simi *et al.* (2008) use the Hough transform in pattern recognition of the hyperbolas formed in GPR images. Thus, an image was clarified to determine the location of a pipe, and to define its depth and size. Liu & Shen (1991) also carried out a numerical simulation using methodologies such as transmission-line-matrix (TLM) and employing Maxwell's equations to identify metallic pipes.

As computer technology has improved, there have been increased efforts to develop tools based on intelligent systems. So, we find proposals for automatic pipe location in GPR images using neural networks [Al-Nuaimy *et al.* (2000);

3.1. Introduction to methodologies of visualization of objects in GPR images

Gamba & Lossani (2000)] and support vector machines (SVM) [Pasolli *et al.* (2008); Pasolli *et al.* (2009a)]; also heuristic methods based on genetic algorithms were proposed in Birkenfeld (2010), where hyperbolas generated by public service components are detected automatically. In the case of [Ciu *et al.* (2010); Delbo *et al.* (2000)] fuzzy set theory is used for GPR image processing. All these works attempt to obtain clean images for locating objects through the implementation of automatic location methods and using an image cleaning process prior to the classification task [Liu & Shen (1991); Pasolli *et al.* (2009b)]. The success of the application of these methodologies hinges mainly on the cleanliness of the images obtained with some classification pre-processing. In most cases, the objective is the identification of the typical hyperbolas identifying the objects of interest in the image under study. In addition, there is an increasing use of plastic elements in WSSs and interpretation methods need to be adapted to this current situation.

3.1.2 Factors influencing the visualization of objects in GPR images

There are multiple factors influencing GPR images. Each factor may substantially alter the signal response passing through different materials. A comprehensive assessment of these factors produces a division into two groups: (a) uncontrolled and (b) controlled factors. The first group includes environmental factors such as: topography, soil humidity, electromagnetic properties of the prospected medium (electrical conductivity, dielectric permittivity, and magnetic permeability; see Chapter 2), medium within which the electromagnetic wave propagates, depth of the pipes, etc. In this chapter this group is not considered for the surrounding medium of the buried pipes, since the tests have been performed under constant, uniform, and controlled laboratory conditions. The second group includes such factors as sampling conditions (GPR design, receiver performance, employed antennas, etc.), data collection, pre-visualization, storage and data processing (pre and post-processing). This chapter is devoted to data pre-processing, while other factors in this group remain constant.

3.2. Principles of the proposed visualization process – T_{14} and T_{15} matrices

The pre-processing considered in this chapter focuses on data from traces and its subsequent transformation into images. The underlying reason for performing this task is to avoid noise sources generating unsuitable images for intelligent systems (the multi-agent systems discussed in Chapter 4) to efficiently locate buried pipes.

3.2 Principles of the proposed visualization process – T_{14} and T_{15} matrices

The traces obtained with GPR prospecting present a quasi-periodic character as a result of the reflections generated by signals passing through the different materials under the prospected soil, following the emission downwards and posterior reception of the reflected signal. The anomalies contained in the soil generate alterations in the wave amplitude as a result of the associated electromagnetic changes they produce. The most important characteristics are the wave amplitude and the frequency. Exploiting the variability of the wave amplitude, which depends on the environment in which the signal propagates, and its quasi-periodic nature, we have developed various procedures that are presented in Section 3.2. The main objective is to generate tools for parametrization of wave amplitude. This seeks to make visible characteristics of buried objects that will generate the basis for training intelligent systems able to provide information on WSS components (Figure 3.2).

3.2. Principles of the proposed visualization process – T_{14} and T_{15} matrices

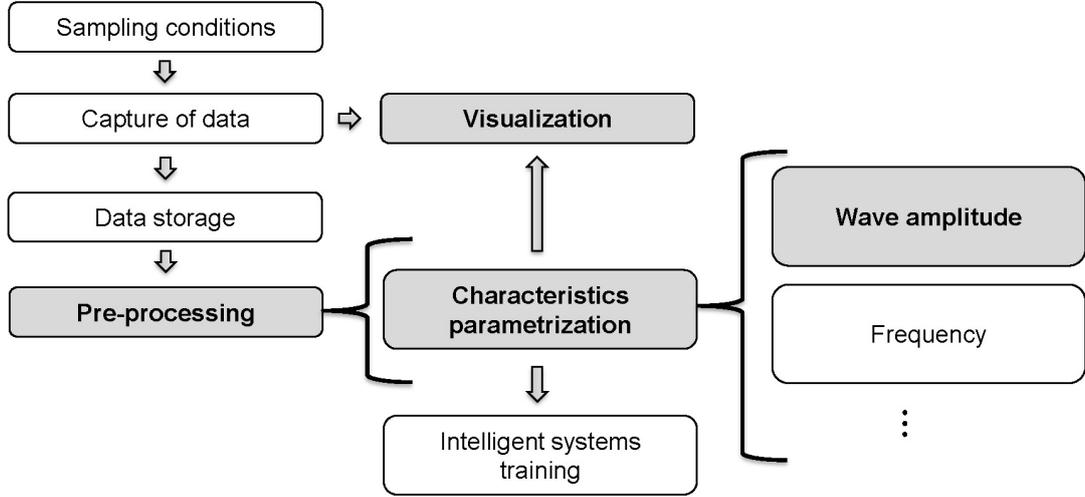


Figure 3.2: Scheme of the proposed visualization process – Parametrization of characteristics

The signals received during GPR prospecting are stored in a matrix, A (called radargram, containing the prospecting raw data), that is made up of m -vectors, X_j , $j = 1, \dots, n$, (called traces), that represent the variation of the soil electromagnetic properties in terms of depth. Let us represent this matrix by columns $A = [X_1, X_2, \dots, X_{n-1}, X_n]$. The length, m , of vectors X_j is the volume of signal data recorded for each vector (Figure 3.3), which depends on the characteristics of the equipment used. The sample size is an equipment parameter; generally sets of 512, 1024, or 2048 samples/trace are used in commercial equipments.

In this chapter, the matrix A will be subjected to different procedures based on the wave amplitude values obtained. This seeks to identify its main characteristics and determine if it is feasible to obtain information from the generated wave amplitudes for the classification of objects in GPR images. In this chapter, these objects are components of WSSs, specifically pipelines.

3.2. Principles of the proposed visualization process – $T14$ and $T15$ matrices

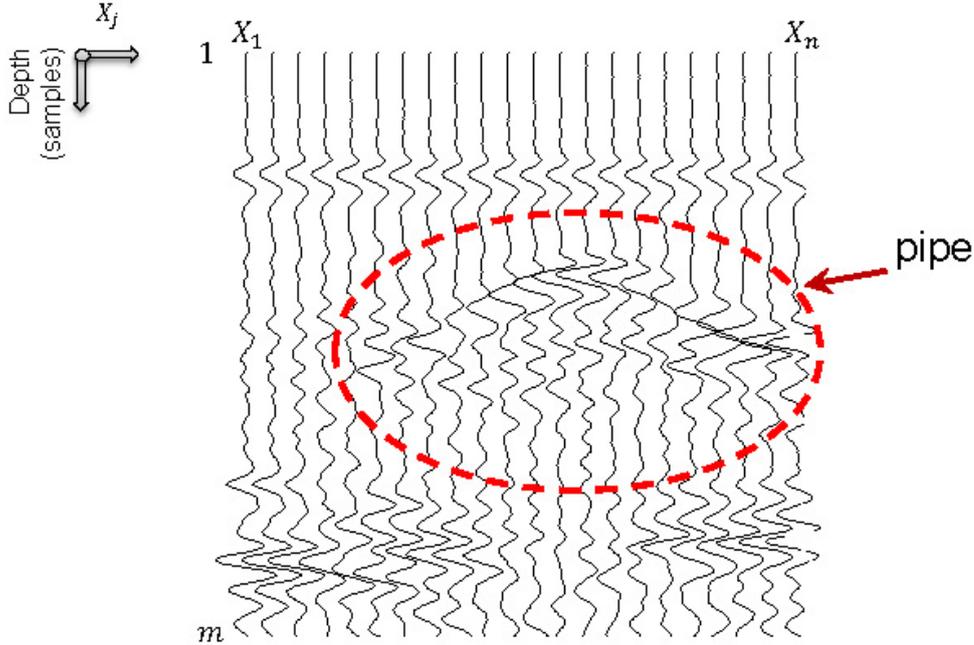


Figure 3.3: Matrix A , formed by the vector traces, X_j

The transformation method used in the initial radargram (raw data) is based on the wave amplitude parameter. The raw data (matrix) is taken as input in the method and divided into ranges of values and then rebuilt into new matrices, as explained below.

The proposed method, after taking the raw data as input, follows two procedures: (a) a matrix-visualization process; and (b) the construction of a classifier. The first procedure builds the so-called $T14$ and $T15$ matrices and comprises four phases: (1) wave amplitude values (WAV) classification, (2) WAV extraction, (3) WAV accumulation, and (4) WAV rebuilding. The second process will be presented in Chapter 4. The visualization method proposed in this chapter is schematically presented in Figure 3.4.

The proposed method (Figure 3.4) gathers various characteristics of the sought anomalies and produces clean images. Usually, the first classification made in GPR images is associated to the different shapes demarcated by the wave amplitudes, after visualizing the images in some color scale (usually gray-scale). From

3.2. Principles of the proposed visualization process – T_{14} and T_{15} matrices

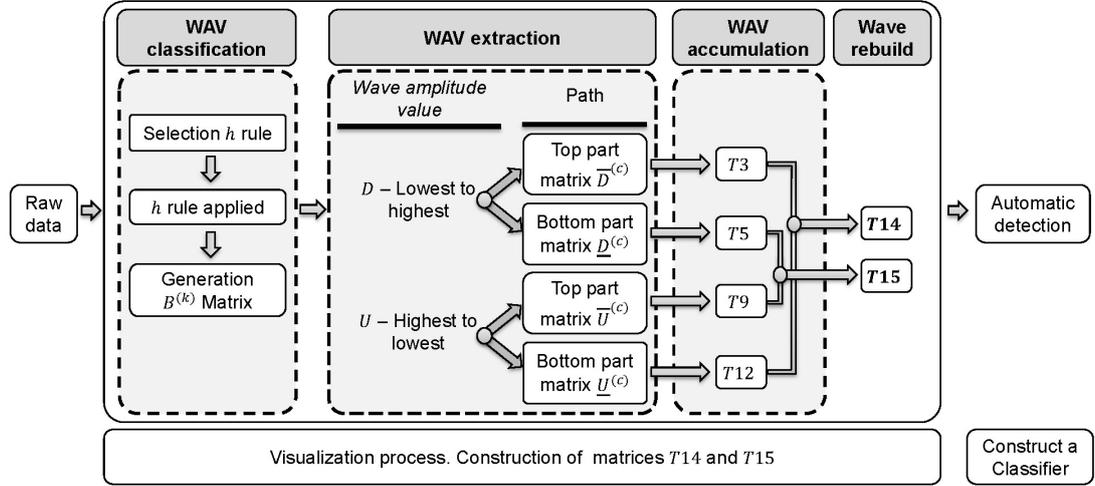


Figure 3.4: Proposed visualization method – Matrices T_{14} and T_{15} . Source: Ayala-Cabrera *et al.* (2011b)

the demarcated forms, operators perform interpretations (based on their experience) to identify different objects that may have been captured in the survey. From this first identification, the following methodology is applied to the matrix A considered: performing a division of wave amplitudes by ranges, extraction of specific parts of the wave, and progressive accumulation of new generated matrices.

Matrices T_{14} and T_{15} constitute the output of the transformation, which may be rendered as images for visualization. The four transformation phases are applied to each trace (column) of the radargram. An instance of the transformation methodology for one trace is shown in Figure 3.5. This figure shows the k columns consecutively ascending, that compose the B matrix and the c columns for the \overline{D} , \underline{D} , \overline{U} and \underline{U} matrices defined below.

3.2. Principles of the proposed visualization process – $T14$ and $T15$ matrices

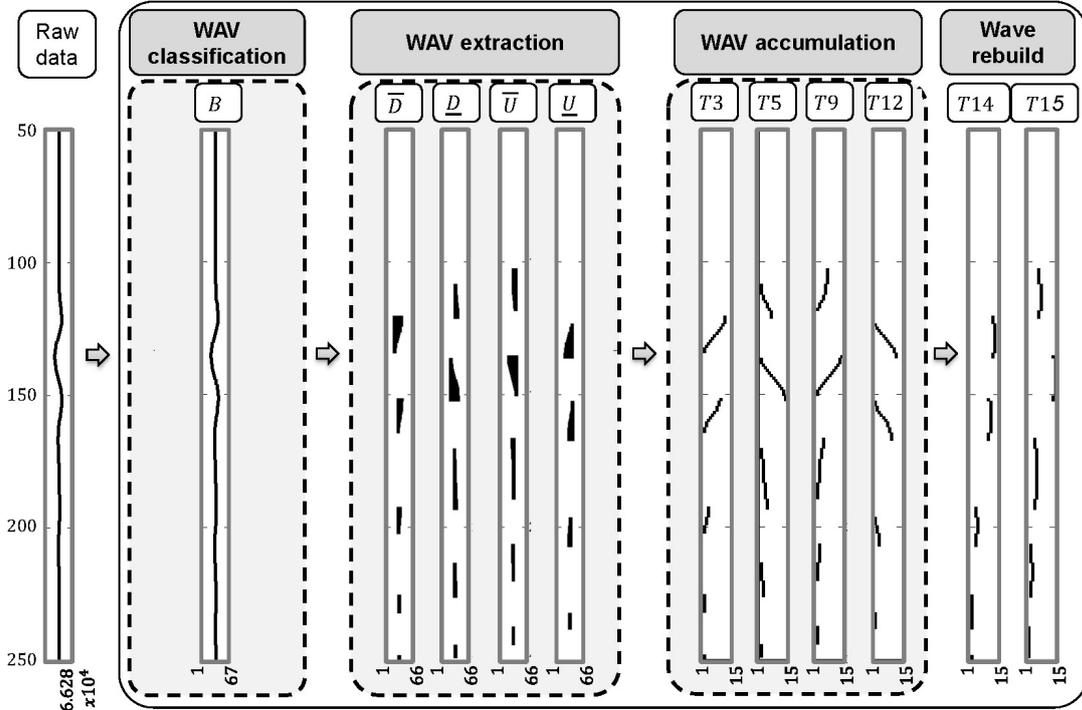


Figure 3.5: The transformation methodology on a trace – Matrices $T14$ and $T15$. Source: Ayala-Cabrera *et al.* (2011a)

The four transformation phases are applied to each trace (column) of the radargram and are described next.

3.2.1 WAV classification

This first phase is based on the conventional gray-scale classification applied to the wave amplitude values, generally used to analyze radargrams. The wave amplitude values in matrix A vary in the interval $[R_1, R_2]$, R_1 and R_2 being the minimum and maximum wave amplitude values adjusted so that the interval length $R = R_2 - R_1$ is a multiple of 1000. Then, we define a partition with norm $h = 1000$, given by numbers a_0, a_1, \dots, a_k . Additional details on the selection of h may be found in Ayala-Cabrera (2009). Taking an arbitrary X_j from matrix A , as shown in Figure 3.3, the trace is subdivided, according to the the equations

3.2. Principles of the proposed visualization process – $T14$ and $T15$ matrices

given below, by amplitude ranges with norm h de 20000, 10000, 5000 and 1000. When representing the effect of color scale classification, the representation in Figure 3.6 is obtained.

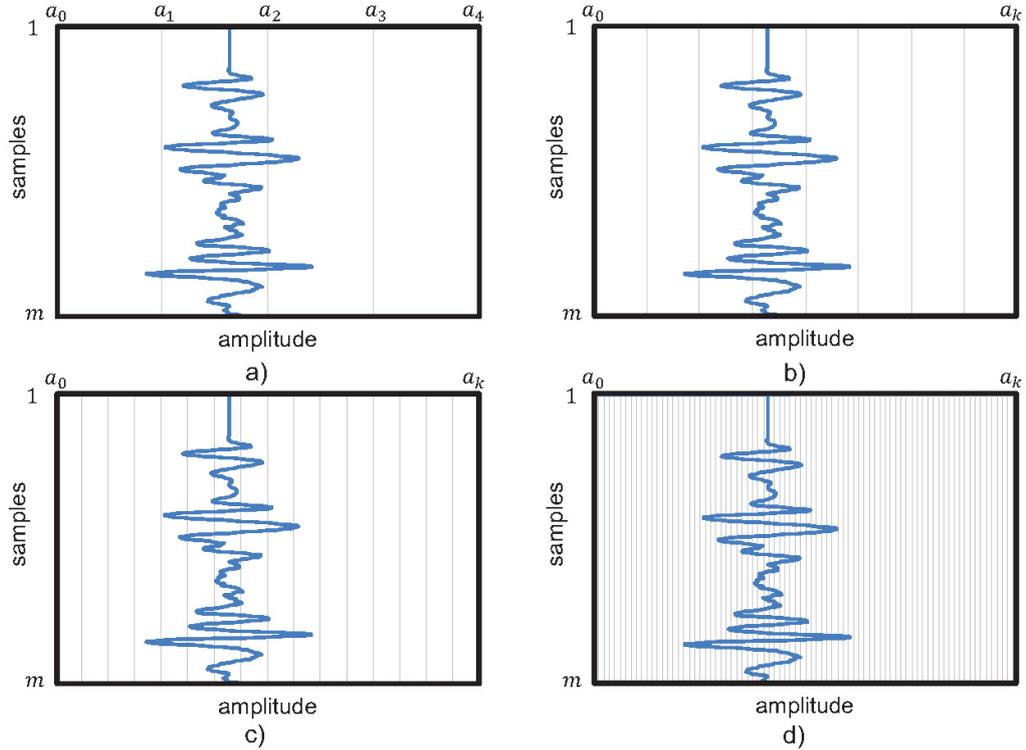


Figure 3.6: WAV classification, selection of norm h . (a) $h = 20000$, (b) $h = 10000$, (c) $h = 5000$, and (d) $h = 1000$

As a result, we have $K + 1$ equidistant elements, where $K = R/h$. Then $B^{(k)}$ matrices are calculated with elements determined by Equation 3.1.

$$b_{ij}^{(k)} = \begin{cases} x_{ij}, & \text{if } a_{k-1} \leq x_{ij} < a_k \\ 0, & \text{otherwise} \end{cases} \quad k = 1, \dots, K. \quad (3.1)$$

The generated matrices $B^{(k)}$ contain the WAV obtained from a GPR prospection, so enabling easy manipulation during the next rebuilding processes. Considering the matrix corresponding to the coarser discretization, Figure 3.6 (part d), we develop an example of the process of generation of matrices B , as represented in Figure 3.7.

3.2. Principles of the proposed visualization process – $T14$ and $T15$ matrices

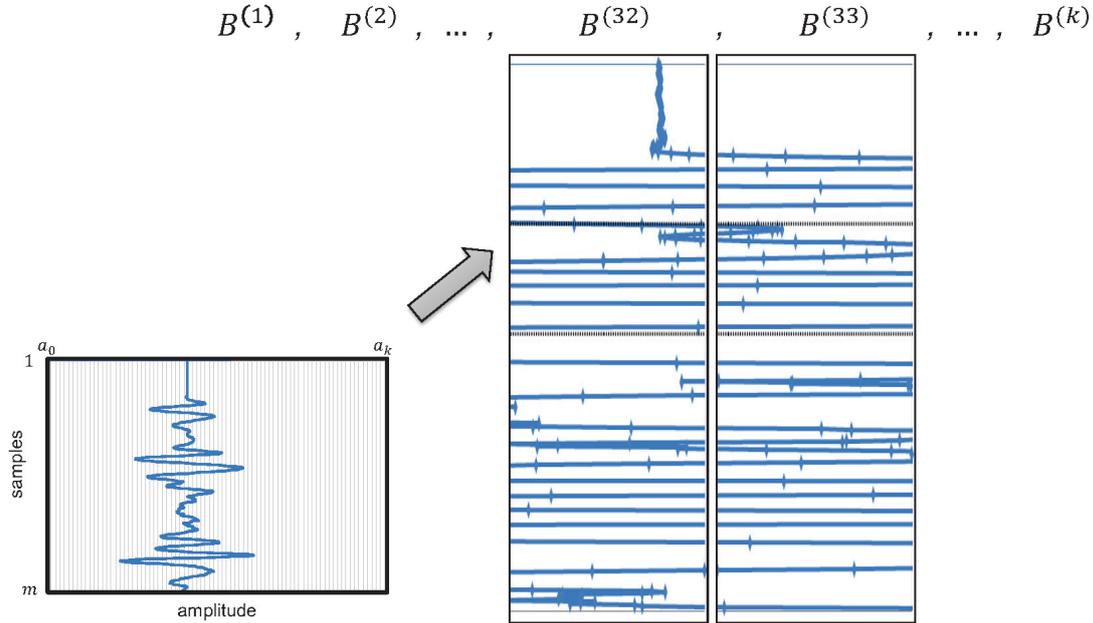


Figure 3.7: Generation of matrices B with $h = 1000$ for one X_j in A

This first classification of the raw data contained in A is the basis of the proposed procedure. We can say that the lower the h , the greater the amount of data to generate more starting points to assess the paths of the wave. However, it should be mentioned that the lower the h , the longer the calculation time employed by the procedure. In contrast, the larger the selected h , the smaller the detail obtained, although the calculation time decreases. We will take $h = 1000$ for the procedure described below.

3.2.2 WAV extraction

The phase after WAV classification is WAV extraction. This phase seeks to contain the values of these wave amplitude variations in separate matrices and it is based on this property. It is used to rebuild separate traces from the data contained in matrices $B^{(k)}$. In the WAV extraction, $B^{(k)}$ matrices are the inputs, and the WAVs are analyzed taking into account the traces' paths. Thus, if the paths are towards higher WAVs, they are extracted and classified into so-called D

3.2. Principles of the proposed visualization process – T_{14} and T_{15} matrices

matrices. Otherwise, if the paths are toward lower WAVs, they are classified into so-called U matrices. Traces under study (X_j) are determined by the successive addition of minimum and maximum WAV (peaks and valleys). Thus, we can take two reference points to rebuild the wave, being either the minimum or the maximum WAV for each wave composing the trace obtained from matrix A .

3.2.2.1 Generation of matrices D

To create the matrices D , the traces are scanned starting from the lowest to highest WAVs. One can follow the paths starting from the top or the bottom part of the wave to produce matrices indicated as \overline{D} and \underline{D} , respectively. The process of generation of matrices D is schematically presented in Figure 3.8.

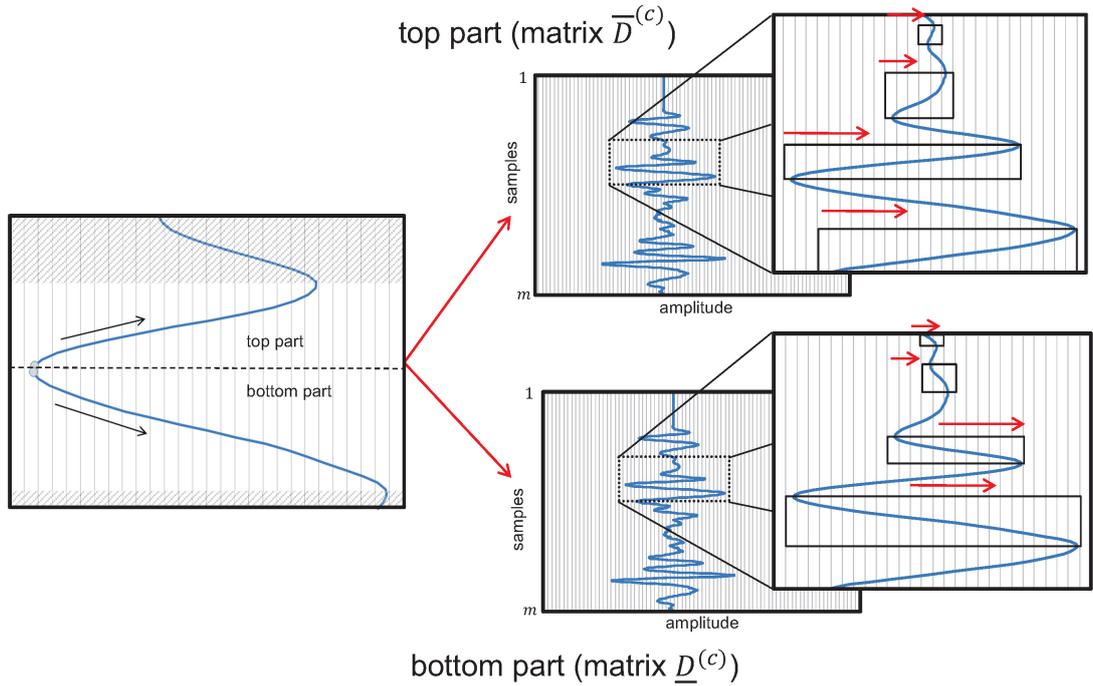


Figure 3.8: Generation of matrices D , starting from minimum values of the wave amplitude – Classification in upper part, \overline{D} , and lower part, \underline{D}

The D matrices (both \overline{D} and \underline{D}) are calculated by using Equation 3.2 for c varying from 1 to $K - 1$, where $B^{*(c)}$ is defined as $b_{ij}^{*(c)} = 1$ if $b_{ij}^{(c)} \neq 0$ and

3.2. Principles of the proposed visualization process – $T14$ and $T15$ matrices

$b_{ij}^{*(c)} = 0$, otherwise (\odot is the Hadammard product of matrices).

$$D^{(c)} = B^{(c)} \odot B^{*(c)} + \sum_{k \in C_1} Br^{(k)}, \quad (3.2)$$

where $C_1 = \{k : c + 1 \leq k < K\}$ is the set of WAV columns to create the D matrix. This equation holds for matrices \overline{D} and \overline{Br} , and \underline{D} and \underline{Br} , where \overline{Br} and \underline{Br} are defined by

$$\overline{br}_{ij}^{(k)} = \begin{cases} 1, & \text{if } \{b_{i-1,j}^{(k-1)}, b_{i-1,j}^{(k)} \leq b_{ij}^{(k)}\} \wedge \{b_{i-1,j}^{(c)}, b_{i-1,j}^{(k-1)}, b_{i-1,j}^{(k)} \neq 0\} \\ 0, & \text{otherwise} \end{cases},$$

$$\underline{br}_{ij}^{(k)} = \begin{cases} 1, & \text{if } \{b_{i+1,j}^{(k-1)}, b_{i+1,j}^{(k)} \leq b_{ij}^{(k)}\} \wedge \{b_{i+1,j}^{(c)}, b_{i+1,j}^{(k-1)}, b_{i+1,j}^{(k)} \neq 0\} \\ 0, & \text{otherwise} \end{cases}.$$

Figure 3.9 present an example of generation of matrices D .

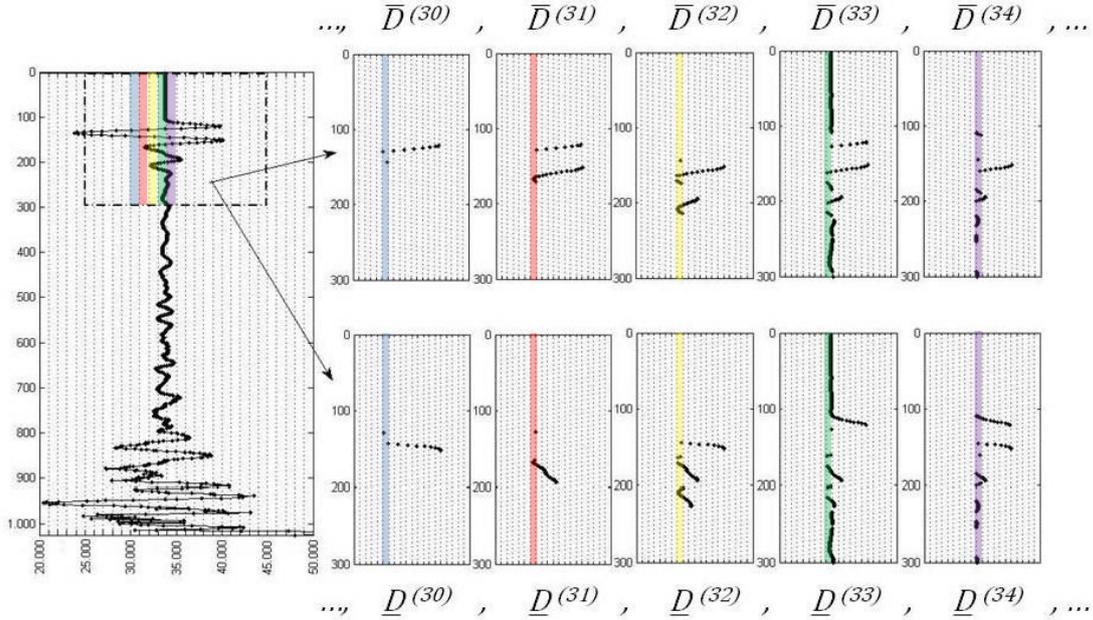


Figure 3.9: Example of generation of matrices D

3.2. Principles of the proposed visualization process – T_{14} and T_{15} matrices

3.2.2.2 Generation of matrices U

The same procedure is performed starting from the maximum WAVs and accumulating from the highest to the lowest values, defining matrices \bar{U} and \underline{U} , top and bottom, respectively. The process of generation of matrices U is schematically presented in Figure 3.10.

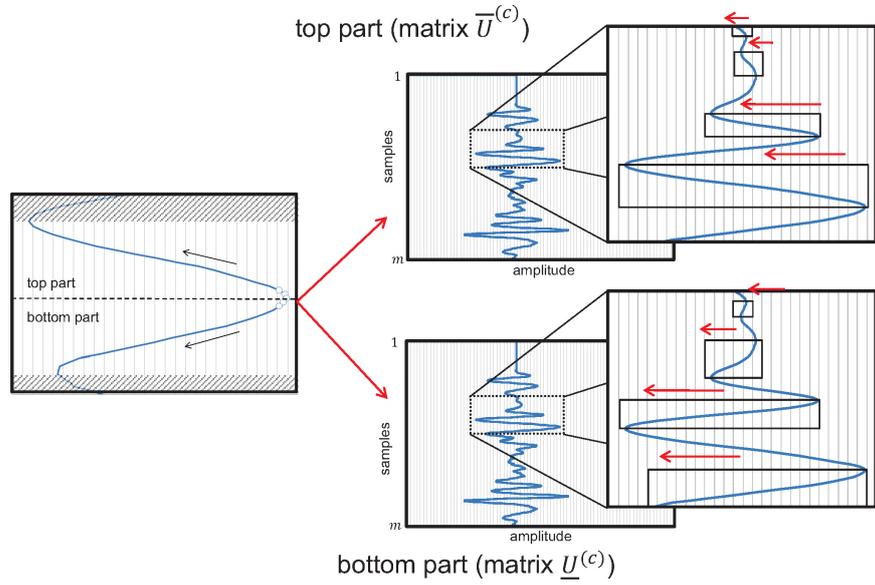


Figure 3.10: Generation of matrices U , starting from maximum values of the wave amplitude – Classification in upper part, \bar{U} , and lower part, \underline{U}

The U matrices (both \bar{U} and \underline{U}) are calculated with Equation 3.3, for c varying from 2 to K .

$$U^{(c)} = B^{(c)} \odot B^{*(c)} + \sum_{k \in C_2} Bl^{(k)}, \quad (3.3)$$

where $C_2 = \{k : c + 1 < k \leq K\}$ is the set of WAV columns to create the U matrix. This equation holds for matrices \bar{U} and \bar{Bl} , and \underline{U} and \underline{Bl} , where \bar{Bl} and \underline{Bl} are defined by

3.2. Principles of the proposed visualization process – $T14$ and $T15$ matrices

$$\bar{b}_{ij}^{(k)} = \begin{cases} 1, & \text{if } \{b_{i-1,j}^{(k-1)}, b_{i-1,j}^{(k)} \geq b_{ij}^{(k)}\} \wedge \{b_{ij}^{(c)}, b_{ij}^{(k)} \neq 0\} \\ 0, & \text{otherwise} \end{cases} ,$$

$$\underline{b}_{ij}^{(k)} = \begin{cases} 1, & \text{if } \{b_{i+1,j}^{(k-1)}, b_{i+1,j}^{(k)} \geq b_{ij}^{(k)}\} \wedge \{b_{ij}^{(c)}, b_{ij}^{(k)} \neq 0\} \\ 0, & \text{otherwise} \end{cases} .$$

3.2.3 WAV accumulation

After obtaining the matrices for the different parts of the wave, we calculate D and U (Equations 3.2 and 3.3, respectively), by using their four side sparse matrices: \bar{D} , \underline{D} , \bar{U} and \underline{U} . We then obtain an accumulated data matrix (that we term $T3$) by adding the generated matrices $\bar{D}^{(c)}$. With $T3$ we highlight the top part of a wave that makes the shape and goes from the lowest to the highest WAVs. By similarly using the $\underline{D}^{(c)}$, $\bar{U}^{(c)}$, and $\underline{U}^{(c)}$ matrices we build matrices $T5$, $T9$, and $T12$, respectively (see Equation 3.4).

$$M = \sum_{c=1}^C N^{(c)} , \quad (3.4)$$

where C is C_1 when the matrices come from D , and C_2 if these matrices come from U . This Equation 3.4 holds for pairs of matrices $(M; N)$ of the form: $(T3; \bar{D})$, $(T5; \underline{D})$, $(T9; \bar{U})$, $(T12; \underline{U})$. Visually, these matrices generated from matrix A , can be seen in Figure 3.11.

As we can observe in Figure 3.11, the matrices obtained from this procedure highlight the anomalies registered by the the signal on its way through the underground. Matrices $T3$ and $T12$ correspond to the top part of the wave, and matrices $T9$ and $T5$ correspond to the extracted wave starting from the bottom part. This correspondence can be visualized in particular characteristics exhibited by each pair of matrices. For example, in Figure 3.11, a red box is used to point to a characteristic that is observed in a similar way in the pair $(T3; T12)$,

3.2. Principles of the proposed visualization process – T_{14} and T_{15} matrices

and in the pair $(T_5; T_9)$, although it is different for both pairs. Based on these similarities we propose a fusion process in the next phase.

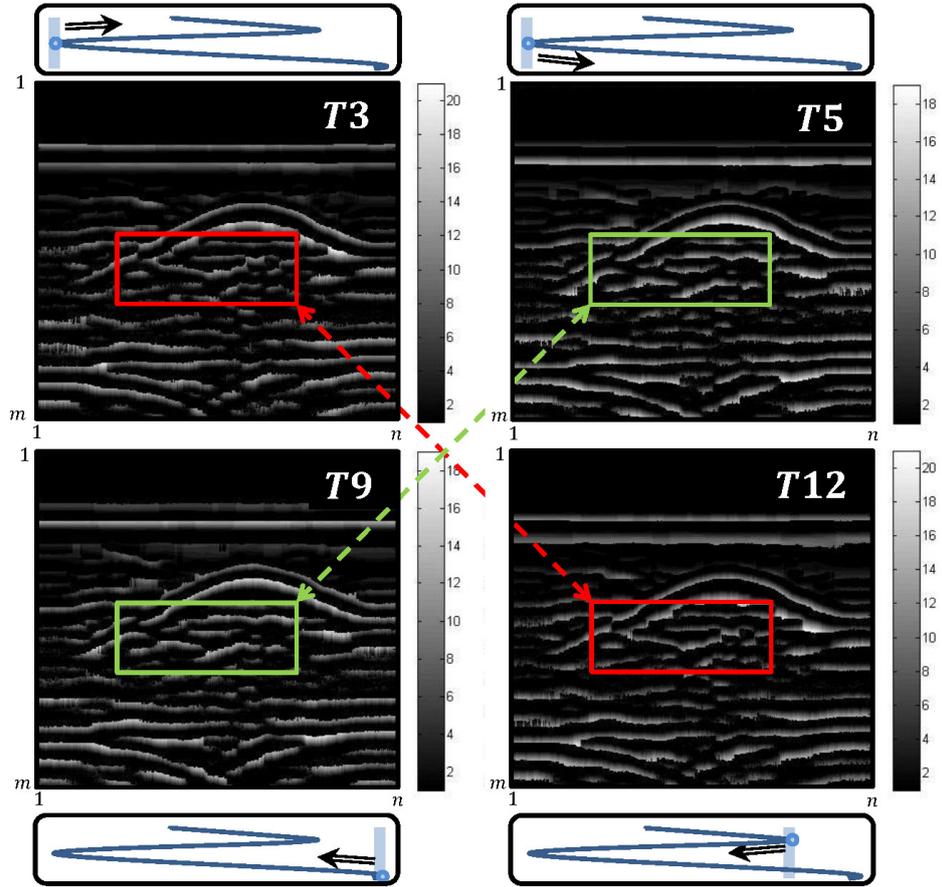


Figure 3.11: Matrices T_3 , T_5 , T_9 and T_{12} from matrix A

3.2.4 WAV rebuild

Exploiting the similarity of characteristics exhibited by the pairs $(T_3; T_{12})$ and $(T_5; T_9)$, which we have presented visually through an example in Figure 3.11, we now propose a fusion process of the matrices of each pair. If we add these matrices following Equations 3.5 and 3.6, we obtain matrices T_{14} and T_{15} .

3.2. Principles of the proposed visualization process – $T14$ and $T15$ matrices

$$T14 = T3 + T12 \quad (3.5)$$

$$T15 = T5 + T9 \quad (3.6)$$

The values in the matrices $T14$ and $T15$ are complementary. These matrices only match in the valleys and peaks. The values in these matrices represent the measurement of every WAV amplitude. An example of these matrices is given in Figure 3.12.

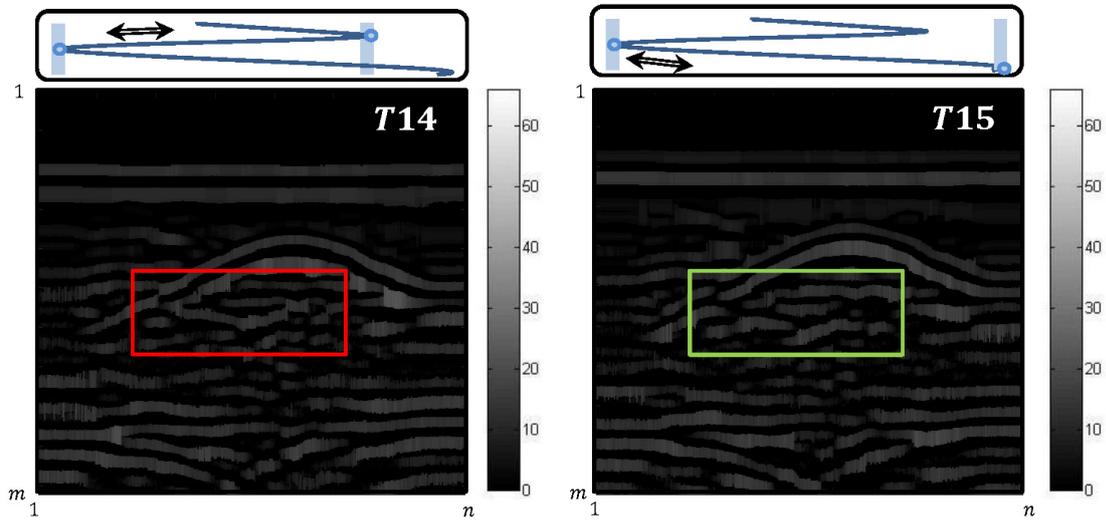


Figure 3.12: Example of matrices $T14$ y $T15$

These matrices are the result of the wave accumulations, represent the ranges of amplitude values of the reflections obtained with the signal passage through the different materials under exploration, and classify the ranges into layers. These matrices minimize the impact of data loss by different times and the accumulation of the data obtained for each survey in an initial single layer.

3.3 Experimental study – matrices T_{14} and T_{15}

This section provides the implementation of the proposed methods of visualization, as described in Section 3.2. To this purpose we address two case studies: (a) a simple case (Case 1); and (b) a complex case (Case 2). The case-studies correspond to GPR images taken from four different pipes commonly used in WSSs. The pipe materials tested were asbestos cement (Fib), cast iron (Fund), polyvinyl chloride (PVC), and polyethylene (PE). The characteristics of the pipes used in these case studies are presented in Table 3.1.

Table 3.1: Characteristics of the buried pipes used for testing

Material	Case	Inner diam. (mm)	Outer diam. (mm)
PVC	1, 2	100	110
Polyethylene (PE1)	2	35	50
Polyethylene (PE2)	1, 2	76	90
Asbestos cement (Fib)	1, 2	80	96
Cast iron (Fund)	1, 2	86	98

The equipment used for these prospections corresponds to a monostatic antenna with a central frequency of 1.5 GHz, that can be considered a good choice for urban underground exploration. Indeed, in this environment, the relevant elements (usually metallic or plastic pipes of many different sections) are relatively close to the ground surface [Gamba & Lossani (2000)]. GPR images were obtained by burying the pipes in dry soil (clayey and silty soil) in the test tank. The following task consist on pre-processing the captured GPR images with the methods proposed in this chapter. Both cases are detailed below.

3.3. Experimental study – matrices T_{14} and T_{15}

3.3.1 Case 1. Simple case

This case corresponds to a GPR survey using four different pipe materials used in WSSs (see Table 3.1). Five assays were performed; one of them corresponds to a no-pipe test – a reference test (Figure 3.13,a) – and the other four correspond, by turn, to the four pipes buried in the tank and surveyed using the equipment (Figure 3.13,b). The equipment parameters for these tests correspond to the configuration of tests number 70 (see Appendix B). We also have to mention that, even though the captured samples amount to 1024 samples/trace, the images in this case only show till sample number 700.

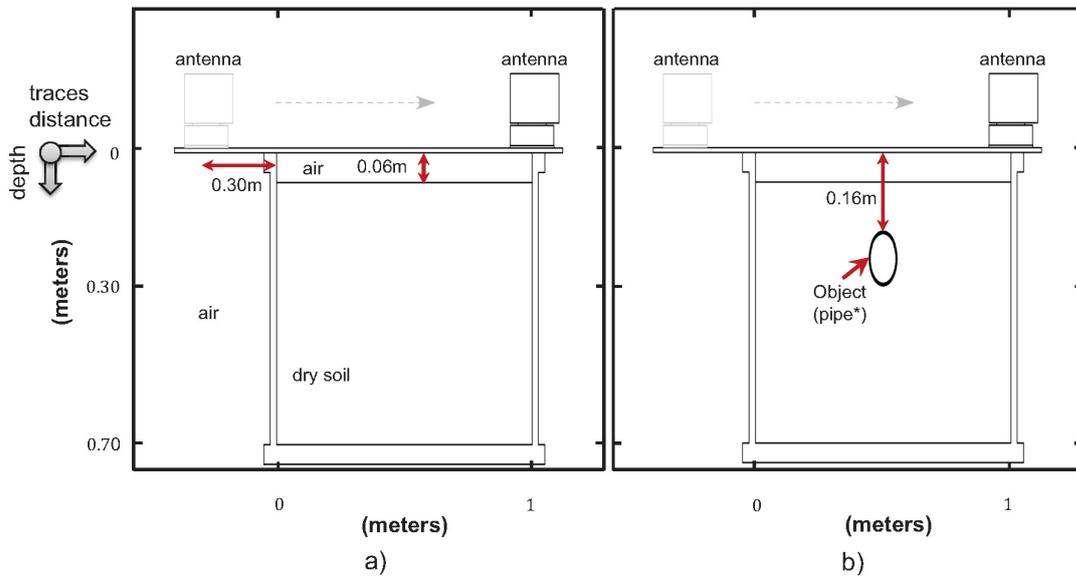


Figure 3.13: Schematic configuration for Case 1, experimental studies – Matrices T_{14} and T_{15} . (a) Capture for no-pipe – reference configuration –, and (b) capture for pipes' configurations

In this first case study, prospection tests were performed for each buried pipe separately. The images without pre-processing (raw matrices) are compared with the T_{14} and T_{15} matrices resulting from the applied procedure described in Section 3.2 on raw data (Figure 3.14). To generate the possibility of making a visual contrast between the pipe images and a reference image we performed a reference test without a pipe, under the same environmental conditions.

3.3. Experimental study – matrices $T14$ and $T15$

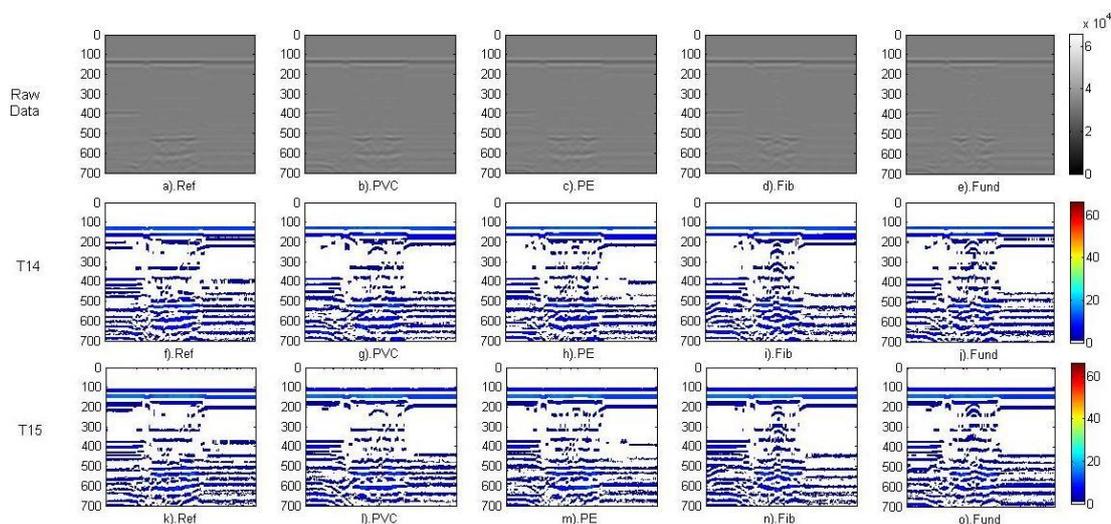


Figure 3.14: Raw image and $T14$ and $T15$ images obtained with GPR – Source: Ayala-Cabrera *et al.* (2011b)

We can see how the color intensity due to material contrast (resulting from the signal passage through soil–pipe–soil) reveals the non-plastic pipes when compared with the raw results (Figure 3.14). Note the intensity of the color images representing asbestos cement and cast iron pipes, the intensity being weaker for the plastic tested pipes (PVC, PE2). However, while PVC and PE2 pipes are visible in Figure 3.14, the contrast of the boundaries is not strong enough for them to be seen at a glance. This is because the low wave amplitude characteristic of these materials also causes a low intensity of color which does not show clear demarcation for PVC and PE2 pipes – and so they become almost invisible. Moreover, lower noise in the image produced by asbestos cement and cast iron pipes is readily appreciated.

The most discernible pipes are those of asbestos cement and cast iron, while plastic materials are displayed with more difficulty. Nevertheless, we can see with clarity and less noise these PVC and PE2 pipes in the images of $T14$ and $T15$ matrices (Figure 3.14).

Asbestos cement and cast iron pipes are already discernible in raw images. These images reveal a solid hyperbola and more color intensity (with the wave

3.3. Experimental study – matrices T_{14} and T_{15}

amplitude) than plastic pipes, which are visible through T_{14} and T_{15} transformations.

3.3.2 Case 2. Complex case

In this case, the prospectings were performed using again the above mentioned pipes (see Table 3.1). However, this time all the pipes were buried simultaneously, at various elevations and positions. With this test we try to approximate the real conditions that may be expected in the field, where various pipes may be buried very close ones to the others. The parameters of the equipment used for this prospecting correspond to the configuration number 12 (see Appendix B). We also have to mention that the captured samples amount to 2048 samples/trace. In Figure 3.15, we present a scheme of the configuration of the pipes in the tank used in this test, and the raw respective image obtained.

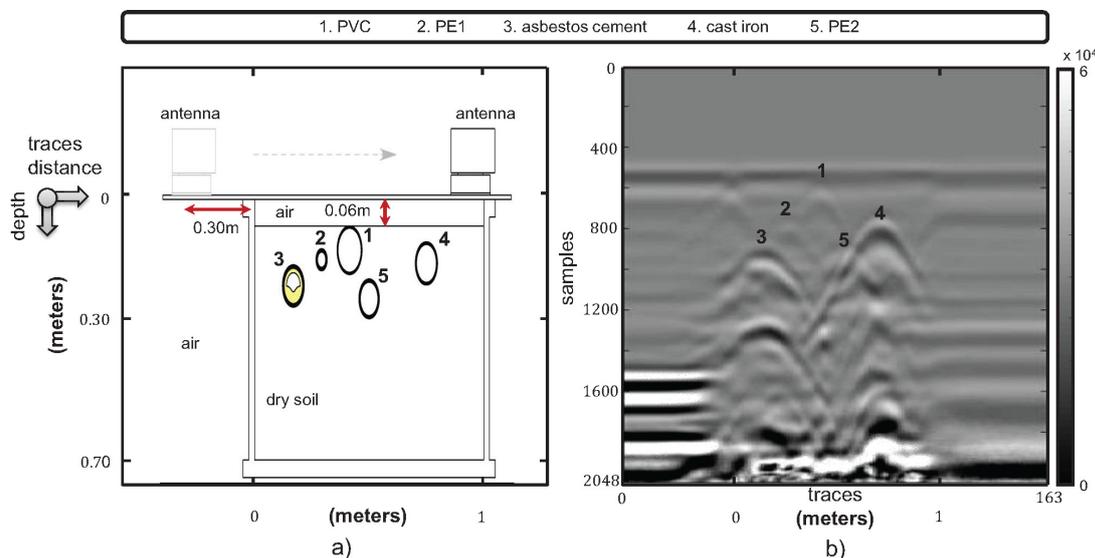


Figure 3.15: Case 2, experimental studies. (a) Schematic configuration, and (b) raw image obtained

In Figure 3.15,b various marks, produced by the passage of the GPR over the location of the pipes in response to the emitted signal, may be seen. We observed

3.3. Experimental study – matrices $T14$ and $T15$

the same behavior in Case 1 (see Section 3.3.1), with significant difference between plastic and non-plastic pipes, according to the color intensity presented. Clearly, asbestos cement and cast iron pipes are better observed, while plastic pipes are not easily identifiable, which is also consistent with the observations in Case 1. It is also shown how the proximity between the pipes overlaps the characteristics of the various pipes, thus hindering identification.

After obtaining the results from this prospection, using the proposed methodology, matrices $T14$ and $T15$ were obtained, thus improving the various pipe characteristics visualization. Figure 3.16 presents the obtained results for the case study 2.

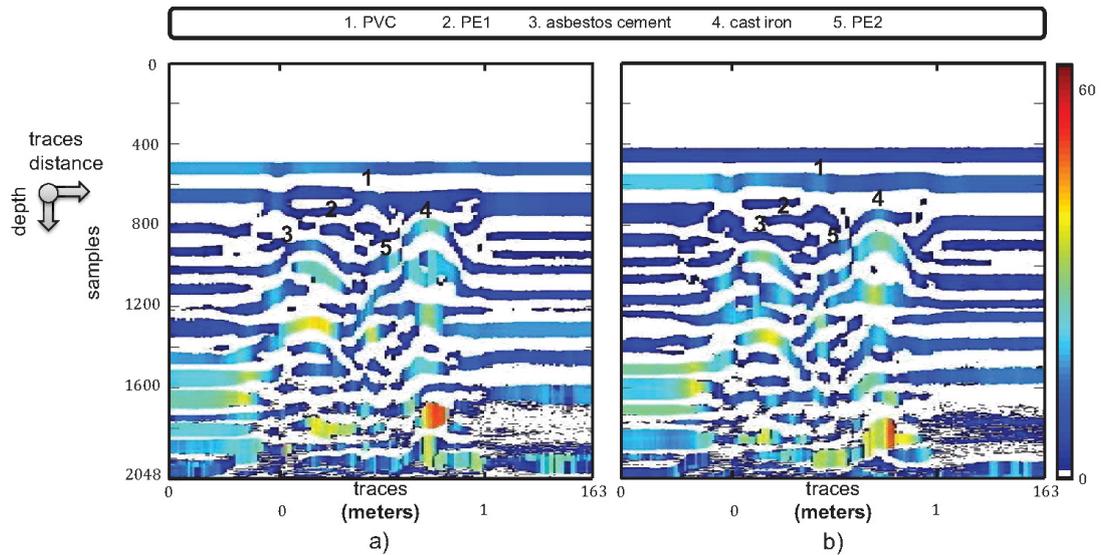


Figure 3.16: Case 2, application of the proposed method. (a) $T14$, and (b) $T15$ matrix images

Figure 3.16,a present the traces left by the various test pipes placed in the tank, which are visualized by layers. Next we describe the observed material characteristics according to the layer discretization of matrix $T14$.

3.3. Experimental study – matrices T_{14} and T_{15}

PVC. For this material, a formation of hyperbolic characteristics is observed in the central part of this figure. Specifically this happens in layer 13, then developing towards layer 9, and presenting asymptotic behavior between layers 8 and 5, which overlap with the medium thus impeding a complete visualization.

PE1, PE2. For material PE1, it can be observed the development in the central area of hyperbolic characteristics, visible starting on layer 9 and further developing the hyperbola till layer 6, without any asymptotes developed. In the case of PE2, the development of characteristics is observed between layers 6 and 2; this material shows low intensity of the wave amplitude due to its scarce conductivity; finally, hyperbolas corresponding to the characteristics of this material are observed that soon mix with other medium reflexion, thus impeding a clear visualization.

Fib y Fund. For these non-plastic materials we can check how hyperbolas develop on the upper layers; in the case of the asbestos cement pipe, hyperbolas appear by the center between layers 18 and 11, and asymptotes develop between layers 10 and 6. In the case of the cast iron pipe, we find hyperbolic development between layers 30 and 18, with asymptote formation observed between layers 17 and 7. Let us pinpoint here how these materials exhibit clear different behavior from plastic pipes, presenting start point and layer formation in higher accumulation levels, and a clear demarcation in the consolidated matrix, thus clearly differing from reflexions produced by the medium.

In Figure 3.16,b we can observe that the central part of the hyperbola corresponding to the characteristics of the various pipes are similar to those observed in Figure 3.16,a, except for the asbestos cement pipe. This fact can be attributed to the fact that the used asbestos cement pipe had an incrustation in its interior, thus provoking the separation of the corresponding hyperbola from those of the other materials, thus separating plastic and non-plastic materials.

Moreover, we can observe that the transformation of the GPR prospection results obtained from the buried pipes, producing T_{14} y T_{15} , considerably improves

the visualization of hyperbolic characteristics for each pipe. This improvement clearly favors the visualization of plastic pipes (PVC, PE1, PE2).

3.4 Implementation issues

The captured GPR data is stored in *.dzt files (GPR survey format). The conversion of these files to Tab Delimited ASCII text format is performed using the RTOAW freeware. With the information in a text file, calculation, manipulation, and organization are performed in MatLab. The mentioned freeware was used in Chapter 3 and in Chapter 4. For the rest of the chapters the code was adapted to dump data directly in the MatLab environment; this code is presented in Appendix A.

3.5 Summary and comments

In this chapter, we propose methods for visualization of pipes in WSSs from GPR images. In the raw captured images (with no pre-processing), we can see how the weakly reflective pipe materials (PVC, PE) are difficult to identify. The transformation of raw data based on the T_{14} and T_{15} matrices improves the visualization of plastic pipe images by producing a better representation of the signal characteristics. PE pipes are more clearly demarcated than when using raw data and the hyperbolic asymptotes of asbestos cement and cast iron pipes are easily visualized. In general, it was much easier to see the pipe features in the transformed images.

The characteristics of the buried pipes, according to their composition, are observed on different layers. So, the possibility of classification opens when groups of pipes are closely spaced. The methodology proposed in this chapter has shown the feasibility of finding in GPR images some typical characteristics of buried objects, such as hyperbolas formed after the passage of the signal by different media.

3.5. Summary and comments

This chapter has further shown that it is feasible to observe these characteristics for not easily identifiable object materials, as is the case of some plastics. This enables us to support the use of GPR in detecting WSS components, given the versatility of the presented method in examining materials of different characteristics. Although the work in this chapter has been performed under controlled laboratory conditions, it enable us to show clear potential to extrapolate the results to uncontrolled conditions of an urban environment. The use of images of GPR in uncontrolled conditions is explored in Chapters 6 y 9.

The two case studies considered in this study enabled us to show that it is possible to obtain the first hyperbola, for objects not only separated in space but also very close spatially (this case is very common in WSSs and in other networks).

Finally, it should be noted that this is a simple process that does not depend on specialist skills (thus being a non-subjective process) and is repeatable. Also, with appropriate adjustments, the radargrams enable the parametrization of supply pipe systems and this could serve as a basis for intelligent training systems. This will help provide WSS managers with a more accurate vision of the systems they operate to offer better service to the users.

Chapter 4

Location of buried plastic pipes using multi-agent support based on GPR images

This chapter focuses on the generation of tools to aid inspection and identify buried plastic pipes in WSSs. A viability study for extracting features from GPR images, and an approach to the above-mentioned application based on multi-agent systems are addressed in this chapter. This chapter proposes a tool for identifying and locating plastic pipes of WSSs in GPR images. This methodology may be adapted for various pipe materials (metal, concrete, etc.) and is based on multi-agent systems working on a transformed GPR image. The conceptual location of this chapter, in the framework of the thesis, is presented in Figure 4.1.

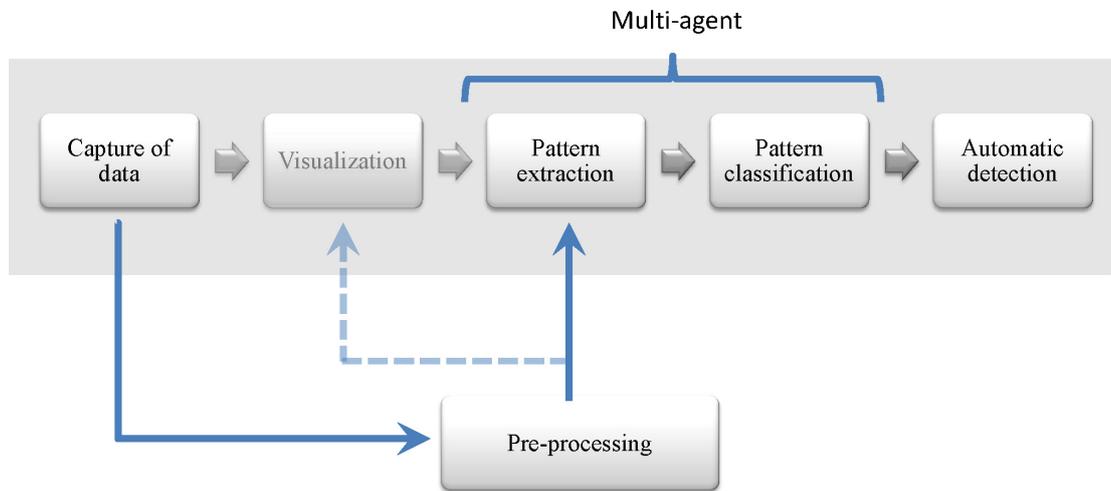


Figure 4.1: Chapter 4 scope

A first approximation of the approach described in this chapter was presented in the *XXIV Congreso Latinoamericano de Hidráulica* [Ayala-Cabrera *et al.* (2010)]. Subsequently, the proposed system was improved and published in a high impact journal, *Journal of Applied Geophysics* [Ayala-Cabrera *et al.* (2011a)]. The main contributions of the procedures proposed in this chapter are: firstly, that operators in charge of GPR exploration will not need great experience or sophisticated skills for locating plastic pipe using GPR images; and, secondly, we have opened a new way to further classification that may be used by other methodologies, as explained later. It should be mentioned that the multi-agent location algorithm proposed in this chapter was implemented in the NetLogo platform [Wilensky (1999)].

4. LOCATION OF BURIED PLASTIC PIPES USING MULTI-AGENT SUPPORT BASED ON GPR IMAGES

The methodology was designed to produce a reliable tool that does not require considerable experience and knowledge from users. The process is divided into three stages: a) transformation of images, b) multi-agent process, and c) location. The first stage transforms data obtained from GPR exploration into two types of images termed $T14$ and $T15$ matrices (see Chapter 3). Firstly, we use intensive matrix manipulation of the GPR output for pre-processing the images. As a result, two matrices are produced that classify initial data based on the original radargram of the wave amplitude parameter. Secondly, we *train* the agents to extract plastic pipe characteristics, hence eventually to automatically locate these pipes in the initial image. Then the plastic pipe characteristics that offer an enhanced likelihood of location are defined. This procedure is evaluated through two case-studies. One study corresponds to a simple case (one pipe) and the other corresponds to various pipes (made of different materials). Both cases were developed under controlled laboratory conditions. The obtained results are promising, and we show that automatic plastic pipe location has been achieved.

This chapter is organized as follows. In this section we have presented a general view of the work developed to build the chapter. Section 4.1 presents a multi-agent system for the automatic detection of plastic pipes in GPR images. The next section provides a succinct introduction to multi-agent systems. Then, (Section 4.3) provides the details of the algorithm for automatic location of plastic pipes from GPR images. In Section 4.4 specific details of the agent configuration are given, and the two application examples, a simple case and a complex one, are developed. Sections 4.5 and 4.6 close the chapter by describing some implementation issues and the main conclusions and comments to this chapter, respectively.

4.1 Automatic pipe location – multi-agent process

In the process described in this chapter a multi-agent methodology is applied to matrices $T14$ (and/or $T15$) for selecting candidate zones for locating plastic pipes in images resulting from GPR surveys. The particular interest of this chapter focuses on the piping material polyvinyl chloride (PVC), due to the increasing use of such materials in WSS lines together with the visualization difficulties posed by these materials.

The development of matrices $T14$ and $T15$ (a method proposed in Chapter 3) provides a visualization tool that enables the identification of objects not easily detectable in GPR images (this is the case of plastic materials, for example, PVC pipes). Exploiting the advantage provided by this visualization system, we propose the methodology described in this chapter. In summary, we start with the $T14$ (or $T15$) images, over which we apply a suitable combination of multi-agent methodologies. This allows automatic detection of plausible areas containing PVC pipes. The proposed multi-agent process is composed of a training phase and an automatic location procedure (see Figure 4.2).

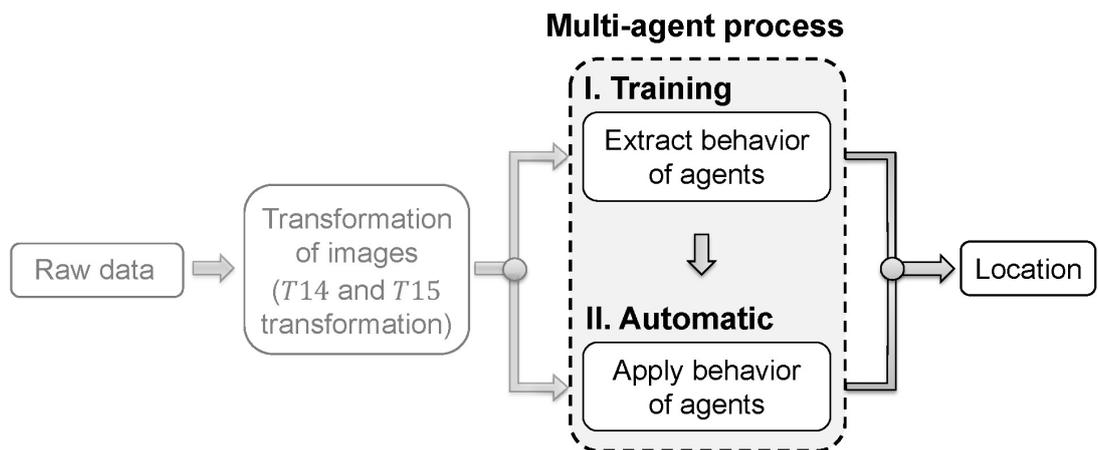


Figure 4.2: Automatic pipe location by multi-agent methods. Source: Ayala-Cabrera *et al.* (2011a)

4.2. A brief introduction to multi-agent systems

Firstly, we propose the type of agents to apply in this process. We then iteratively assess its behavior enabling them to extract the most plausible location of buried plastic pipes. Finally, after applying this second stage, the process finishes with the location of the pipe in the image. The number of agents in each phase and the characteristics of each phase were determined in iterative trials until reaching a stable configuration making it possible to ascertain the plausible area of location of the pipe in the image. These processes are detailed in Section 4.3.

4.2 A brief introduction to multi-agent systems

A multi-agent system consists of a population of autonomous entities (agents) [Sycara (1998)] situated in a shared structured framework (environment) [Weyns & Holvoet (2005)]. An agent is any entity in a system that can generate events affecting itself, other agents and the environment. The main principles guiding an agent behavior are: autonomy, social abilities through some specific communication language, reactivity to the environment, and proactivity enabling the agent to take initiatives.

The system is based on such tools as game theory (e.g., setting the agent preferences by a utility function), economics, and biology, as well as artificial intelligence algorithms [Shoham & Leyton-Brown (2009)]. Other significant contributions on application of multi-agent systems to WSS problems may be found in the works by Gianetti *et al.* (2005) who used agents to control the physical equipment of a water supply, or Izquierdo *et al.* (2009, 2010a,b), working on different aspects of WSS management, such as water-hammer control, or the establishment of criteria to divide the network into district metered areas.

In a system representing some reality (an image in our case) agents may be either external or internal elements of the system. In the first case, agents are disseminated within the system to assess their immediate environments and make decisions about themselves, or their neighboring agents, or their environment. In

4.2. A brief introduction to multi-agent systems

the second case, we understand that it is a system composed of subsystems at arbitrary nesting depths and achieving different levels of abstraction. Given a fixed level, the individual components will be the agents that decompose the whole system into different parts, and these are examined in a decentralized manner. Often, this is more efficient than working directly in some global approach.

Agents operate independently but they are also able to interact with their environment and coordinate with other agents. This coordination may imply cooperation if the agent society works towards common objectives. Thus, in a cooperative community, agents usually have individual capabilities which, when combined, will lead to solving the entire problem. However, cooperation is not always possible and there are instances where agents are competitive and have divergent goals. In this later case, the agent should also take into account the actions of others. However, even if agents are able to act and achieve their goals by themselves, it may be beneficial to partially cooperate for better performance, thereby forming coalitions. When coordinating activities, either in a cooperative or a competitive environment, negotiation may prove a suitable method to solve conflicts among agents. Negotiation may be seen as the process of identifying interactions based on communication and reasoning regarding the state and intentions of other agents [Herrera *et al.* (2010)].

Once agents have been defined and their relationships established, a schedule of combined actions defines the processes occurring over time. Instructions are given to hundreds or thousands of agents that operate independently by interacting with their environment and coordinating among themselves. This coordination may imply cooperation (as in case of this chapter) if the agent society works towards common goals. Thus, in a cooperative community, agents usually have individual capabilities which, when combined, are able to efficiently solve a problem.

4.3 Proposed multi-agent process

The PVC pipe location algorithm is described in Figure 4.3. This figure shows that the process is based on the images associated with the $T14$ (or $T15$) matrices and is composed of three phases: (a) *roll-on*; (b) *termite-eating*; and (c) *final detection*.

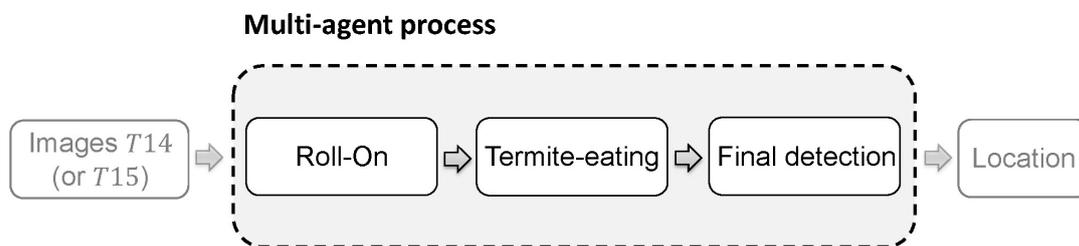


Figure 4.3: Phases of the proposed multi-agent process. Source: Ayala-Cabrera *et al.* (2011a)

In the problem considered here, the agents work on images $T14$ (or $T15$). The underlying rationale behind this process consists in cleaning zones where the presence of a pipe is less likely, taking into account that the whole image is composed of blue and white pixels. Thus, we will discard areas where the values of the color distribution may show that the presence of a pipe is unlikely. To establish this, let us consider the random variable X , which contains the value of a pixel color and we will look for a probability such as:

$$P\left(|\mu - \bar{X}_{Obs}| > \bar{X}_{Color}\right) \leq tol_{\alpha}, \quad (4.1)$$

where μ is the average color in the presence of a pipe, \bar{X}_{obs} is the average color observed in the area, and tol_{α} is the error tolerance. If we fix tol_{α} we can obtain the critical values of \bar{X}_{Color} by proposing the maximum and minimum values for deciding whether to erase, or not, each area under study. Unlikely zones are discarded by agents of different breeds that can cooperate with each other and their behavior is based on Equation 4.1. The stages that compose this process are described below (and the implementation is detailed in Section 4.4):

4.3. Proposed multi-agent process

1. *Roll-on.* A vertical array of agents covering the full height of the image is first defined. Each covers a squared area and measures the conditions to find a pipe. These agents move horizontally by scanning across the image width. After scanning a square, an agent deletes the corresponding sub-image if it corresponds to a zone where virtually there is no pipe. The area is deleted if there are many blue (or white) pixels inside each inspected square.
2. *Termite-eating.* The first phase cannot distinguish where the pipes are located because the agents are constrained to fixed start positions and square dimensions. The creation of a new breed of agents will add flexibility to this method. We call these agents termites. They start randomly in a number of image pixels and their movements depend on the pixel color where they are placed. Taking into account that pipes are blue colored, the termites walk on blue pixels with care and go faster on white pixels. When the termite position is blue, an evaluation is carried out horizontally near the neighboring pixels. This horizontal area will be deleted (eaten by the termite) if either too many or too few blue pixels are detected. This allows us to discriminate sets of blue pixels organized in shapes other than horizontal lines (such as an hyperbola shape, which suggest the presence of a pipe).
3. The *final detection* phase is divided into various parts: firstly, we compute the borders of the undeleted figures. These borders are colored with a color of greater value than the corresponding blue in the RGB scale. Agents similar to those mentioned in phase 1 are then brought into action. However, these agents now have more flexibility in their movements, and cross along the image surface following random walks. This final process stops after an adequate number of iterations (which depends on the number of agents).

For the training stage, we propose a NetLogo interface [Wilensky (1999)] that implements the three multi-agent processes above described. This application is customizable using a simple menu (Figure 4.4).

4.3. Proposed multi-agent process

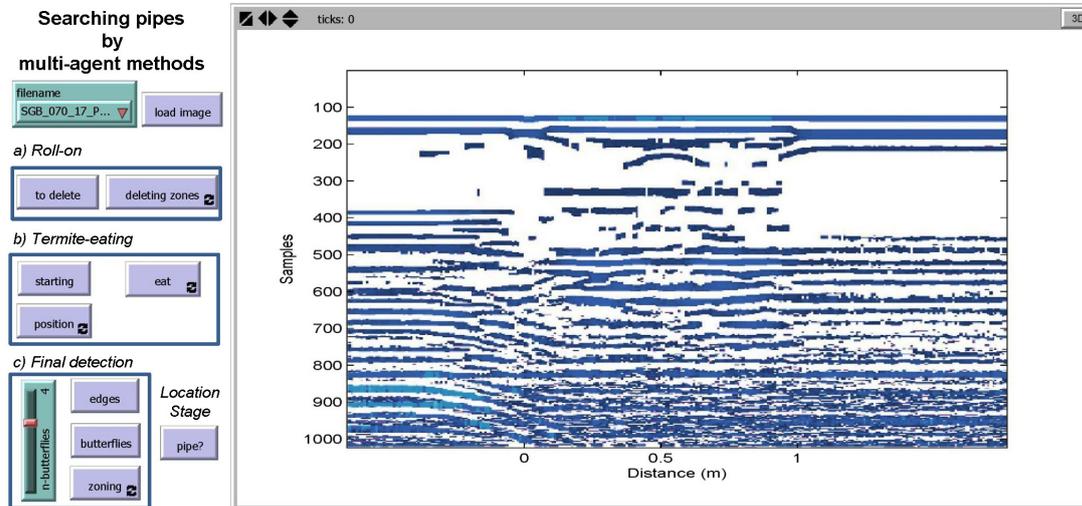


Figure 4.4: Proposed multi-agent interface. Source: Ayala-Cabrera *et al.* (2011a).

Users can interact with the process by changing the default parameters (e.g., number of iterations on each phase and the number of agents) and loading the file with their target image. In this chapter, the interface was employed iteratively, determining suitable agent characteristics to detect and locate plastic pipes (PVC) in GPR images. Nevertheless, this interface may also be used to locate pipes made from other materials (including metal and concrete, among others).

Once a suitable agent-configuration is identified, we extract its behavior to automatically detect pipes in other GPR images. Now, the remaining blue pixels (after the *final detection* phase) are colored in red, and a pipe is easily visualized in the plausible zone of its location when this image is overlapped with the T_{14} (or T_{15}) initial image. The complete automatic procedure is shown in Figure 4.5.

4.4. Experimental study – multi-agent process

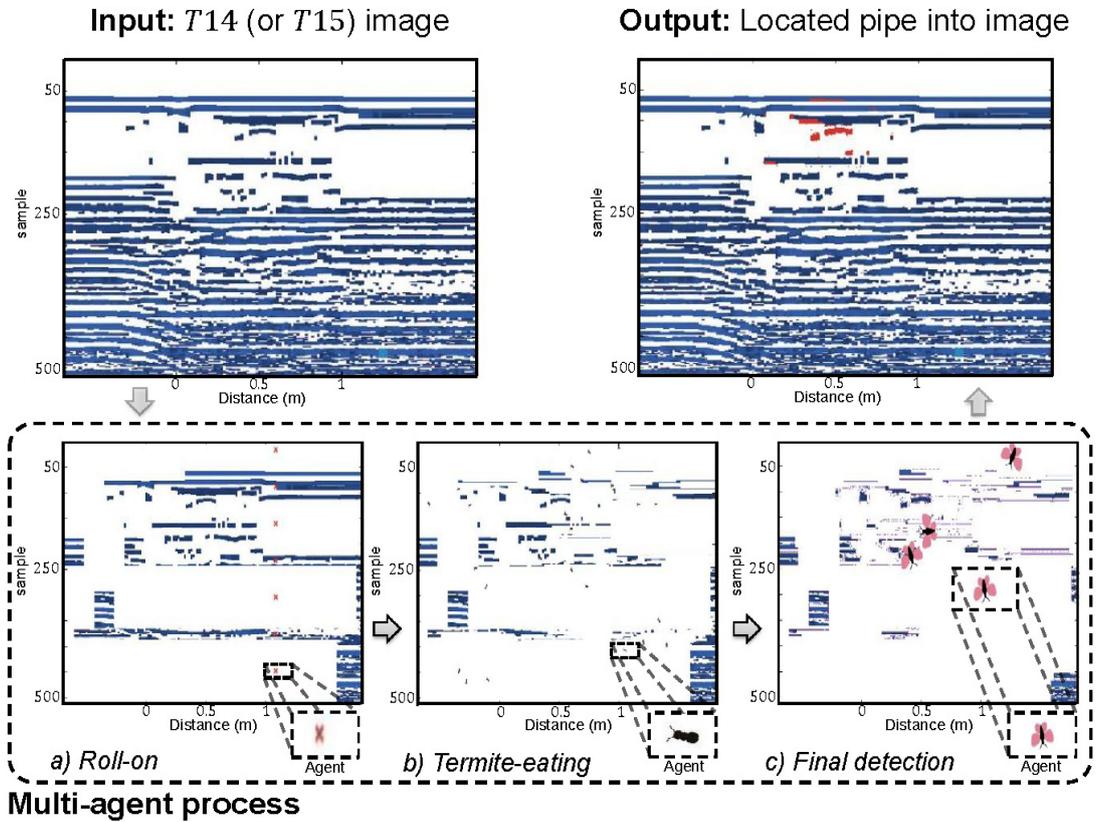


Figure 4.5: Evolution of the proposed multi-agent process. Modified: Ayala-Cabrera *et al.* (2011a)

4.4 Experimental study – multi-agent process

The implementation viability of the multi-agent method proposed in Section 4.3 is evaluated in this section. To do this, we test two case-studies: (a) a simple case (Case 1); and (b) a complex case (Case 2). In the tests we used pipes commonly used in WSSs. The pipes were buried under controlled conditions in dry soil. The pipe materials tested were PVC, asbestos cement (Fib), and cast iron (Fund). The characteristics of the used pipes are shown in Table 4.1.

4.4. Experimental study – multi-agent process

Table 4.1: Characteristics of the buried pipes used for testing

Material	Case	Inner diam. (mm)	Outer diam. (mm)
PVC	1, 2	100	110
Asbestos cement (Fib)	2	80	96
Cast iron (Fund)	2	86	98

The equipment used for these explorations corresponds to a monostatic antenna with a central frequency of 1.5 GHz, which can be considered a good choice for urban underground exploration. Indeed, in this environment, the relevant elements (usually metallic or plastic pipes of many different sections) are relatively close to the ground surface [Gamba & Lossani (2000)]. The PVC pipe was used in both Case 1 and Case 2. The other two pipes were only used in Case 2. Both cases are detailed below.

4.4.1 Customizing the multi-agent location process

To customize the automatic location process, the training stage proposes the following configuration for the multi-agent process:

1. In the first phase, *roll-on*, the $T14$ image is covered with eight square agents that cross the image horizontally. While they are moving, the agents test the possibility of finding a pipe in their current position. These agents will delete the areas with low possibilities.
2. In the *termite-eating* phase, we generate 125 termites to check image positions in their random walks. These new agents will discard (eat) the horizontal blue stripes that are too large (more than 400 blue pixels) or too small (less than 50 blue pixels) to be part of a pipe image. To gain efficiency in their walks of 400 steps, the termites will move pixel to pixel if their current position is a blue pixel (pipes are formed by blue pixels), and jump ten pixels ahead if their position is white.

4.4. Experimental study – multi-agent process

3. In the last *final detection* phase, we color in purple (above blue in the RGB scale) the borders of the remaining images. We then start the process with four squared agents, as above, but moving randomly (200 iterations).

The customization proposed is a consequence of the training stage where we evaluate twenty different agent configurations on forty $T14$ and another forty $T15$ images of PVC pipes (twenty for the simple case, see Section 4.4.2; and another twenty for the more complex case, see Section 4.4.3). We select the best configuration (eight roll-on agents that cross the image, 125 termites – 400 iterations –, and four square agents – 200 iterations). In more than 95% of the instances, the process successfully locates the PVC pipes for the simple case (this number diminishes in the complex case to 75%). The false positive percentage is approximately 2% in the simple case and 10% in the more complex case. The final image obtained with this multi-agent process should be just the pipe (with maybe some noise around it). These pixels have been colored in red and placed over the original image.

4.4.2 Case 1. Simple case

This first case corresponds to a PVC pipe exploration (see Table 4.1). This first case corresponds with the one presented in Chapter 3, (Case 1); the corresponding results, after application of the transformation into the matrices $T14$ and $T15$, have already been presented (Figure 3.14, parts b, g, l) and analyzed in Section 3.3.1. However, the equipment parameters used in this survey correspond to the test configuration number 68 (see Appendix B). The main difference between this assay and the one presented in the previous chapter resides in the amount of samples, respectively, 512 and 1024 samples/trace. The proposed configuration for this case (Figure 4.6,a) was tested in a direction that was transverse to the pipe. The obtained raw data matrix is shown in gray scale (Figure 4.6,b). The results obtained, after applying the automatic location method proposed in this chapter, are shown in Figure 4.6,c.

4.4. Experimental study – multi-agent process

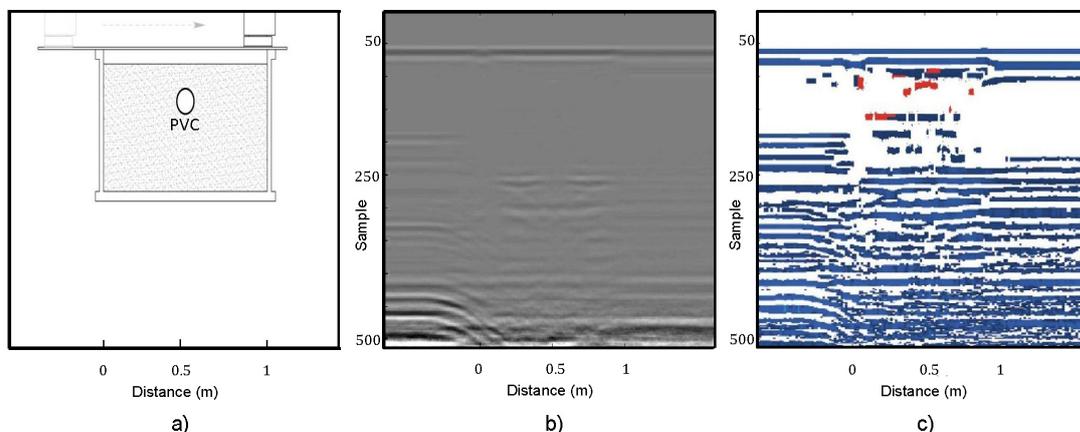


Figure 4.6: Simple case – This shows the correspondence of the result obtained with the proposed schema. (a) Schematic configuration, (b) GPR raw image, and (c) T_{14} image with the identified pipe. Source: Ayala-Cabrera *et al.* (2011a)

In the raw data image (Figure 4.6,b), we notice that the contrast of the PVC pipe borders are not strong enough to be seen immediately. This effect is a result of the low contrast between the permittivity of the soil and the buried pipe. This produces low color intensity that prevents clear PVC pipe demarcation in the image (being almost invisible). Figure 4.6,c shows the result after the application of the transformation methodology and the automatic location process. In the image we can also see the formation of a solid hyperbola (corresponding to the PVC pipe) that is only visible after the T_{14} (or T_{15}) transformation. This hyperbola is successfully selected automatically.

4.4.3 Case 2. Complex case

We prospected for three pipes in this more complex case (see Table 4.1). These pipes were buried simultaneously (at different elevations and positions). This second test attempts to obtain a better recreation of real-world street conditions, where pipes are often buried closely together. This case corresponds with Case 2 presented in Chapter 3. However in this test two out of the five pipes considered in the previous chapter were not used. The equipment parameters used in this test correspond to assay configuration number 79 (see Appendix B). In this case,

4.4. Experimental study – multi-agent process

compared with Case 1 of the previous chapter, in addition to the number of samples per trace (2048 instead of 512 samples/trace), the number of traces per second was also increased from 32 to 120 traces/s. Figure 4.7,a shows the pipe distribution in the experimental tank and the raw image obtained after the exploration (Figure 4.7,b). The methodology proposed in this chapter was tested in this setting. The results are shown in Figure 4.7,c.

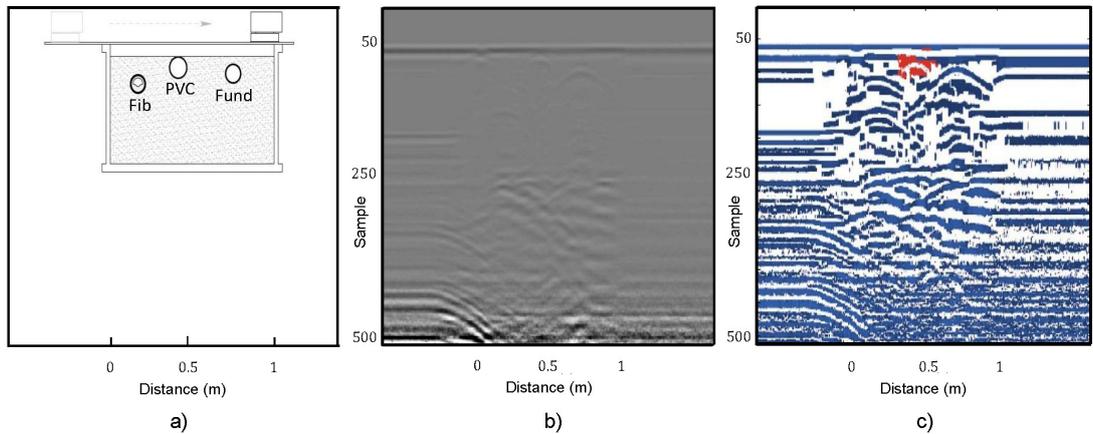


Figure 4.7: Complex case – This shows the correspondence of the obtained result with the proposed schema. (a) Schematic configuration, (b) GPR raw image, and (c) $T14$ image with the identified pipe. Source: Ayala-Cabrera *et al.* (2011a).

Figure 4.7,b suggests traces of different pipes in response to the GPR signal. We can observe that the PVC pipe representation is similar to Figure 4.6,b with little or no demarcation boundaries. The cast iron pipe shows the best demarcation, while there are almost no traces of the asbestos cement pipe.

After obtaining this data, we made a transformation using the proposed methodology. Then, we applied automatic plastic pipe location in the $T14$ image as shown in Figure 4.7,c (showing the hyperbolas of the three pipes). The closeness between the pipes means there is an overlap for the majority of the characteristics. This means it is more difficult to locate the pipes. Nevertheless, by applying the multi-agent process proposed above, we can discriminate the plastic pipe characteristics from the others.

4.5 Implementation issues

The algorithm proposed in this chapter has been implemented on both Windows and Linux environments, using MatLab to carry out the above treatment of matrices, and the NetLogo platform as the multi-agent programming language. NetLogo offers a flexible and powerful tool as it runs on the Java virtual machine and is similar to natural language in some ways (as we can see in Table 4.2, with an excerpt of the algorithm code for phase 2).

Table 4.2: Looking-for-pipes excerpt within stage 2 of the working algorithm

```
... to looking-for  
ask termites with [ [pcolor] of patch-here > 90] [  
set heading 270  
let nearby [list pxcor pycor] of patches with [abs pxcor <= 80  
and abs pycor <= 2]  
set nblue count patches at-points nearby with [pcolor >= 50]  
if nblue > 400 [ask patches at-points nearby [set pcolor white] ] fd 1  
if nblue < 50 [ask patches at-points nearby [set pcolor white] fd 1  
] ] ...
```

The results of the visualization process (Chapter 3) are printed using the RGB color scale, and processed with NetLogo, which, in turn, provides numerical and graphical output for exportation to other programs for further analysis.

4.6 Summary and comments

It is well-known that weakly reflective materials (PVC, PE) are difficult to identify without pre-processing the raw captured GPR images. Plastic pipes show weak values of wave amplitude after reflection, and reveal undefined contours along the boundaries. This chapter introduces an automatic pipe location method based on multi-agent systems. This takes advantage of the quality offered by the $T14$ and $T15$ image representations. The developed methodology is efficient with regard to the computational resources and accurate in its results (even in the challenging case of plastic pipes). The application is carried out by a customizable and simple software environment. It offers graphical and numerical results that could favor future exchanges between different software sources and help WSS managers gain a more accurate vision of their systems and so offer better service to users.

This chapter further shows that pre-processing of the appropriate data is indispensable. It is worth to keep in mind that the cleaner the data, the higher the chances of success for detecting unwanted objects (in this case, plastic pipes) in the GPR images.

The two referred case studies have enabled us to show that it is possible to discriminate and then extract the first hyperbola corresponding to plastic pipes, not only when the object is isolated in space, but also when other objects are very close spatially (a very common situation in WSSs and other networks).

Finally, the proposed methodology to locate pipes in plastic pipe images was effective in both case studies. The obtained results show the viability of using multi-agent methods for locating (plastic) pipes in simple and complex cases. We should highlight that the proposed location process is a management tool that does not require user experience and can be used in both mapping and inspecting WSS pipe systems.

4.6. Summary and comments

Chapter 5

Segmentation of GPR images of WSS components

This chapter focuses on the development of easy-to-apply procedures that allow the visualization and quick identification of WSS pipes in GPR images by non highly qualified staff. In this line, in Chapter 3 the feasibility of displaying objects of different materials registered in GPR images has been shown. We showed that it is feasible to visualize elements in GPR images, including those elements generally difficult to identify, such as plastic materials (e.g. polyethylene, polyvinyl chloride – PVC). By observing the broad spectrum of visual possibilities for objects of different materials provided by the GPR registrations, in Chapter 4 an automatic procedure that evaluates the images and determines the presence of plastic pipes was proposed. This way, it was found that it is possible not only to display but also to automatically detect objects of plastic materials in GPR images.

In this chapter, the feasibility of developing a general tool that allows to expand the range of materials (objects) and their characteristics susceptible to identification, thus favoring GPR image analysis, is evaluated. The aim is to bring to light WSS components contained in GPR images. We want to assess the feasibility of implementing a tool for extracting general characteristics of different pipe materials that enable automatic identification. The study focuses on the wave amplitude, and the Hough transform has been used as a method for segmentation and cleaning of areas on no presence of pipes. We obtain promising results as the methodology, which is non-subjective and repeatable, visualizes buried pipes with efficiency. The objective is to generate the basis and identify the parameters for training intelligent systems that are able to characterize different components of urban WSSs. The approach of this chapter, in the framework of this thesis, can be seen in Figure 5.1.

5. SEGMENTATION OF GPR IMAGES OF WSS COMPONENTS

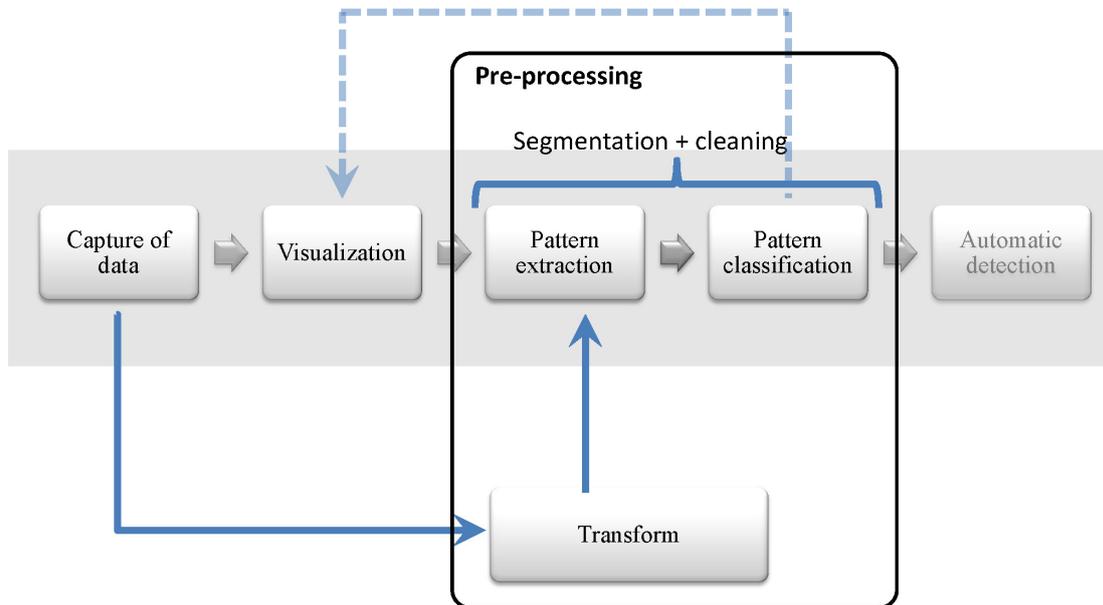


Figure 5.1: Conceptual approach of Chapter 5

A first approximation of the approach described in this chapter was presented in the *X SEREA* [Ayala-Cabrera *et al.* (2011d)]. The main contribution of the GPR image segmentation and cleaning approach presented in this chapter boils down to reducing the amount of data to evaluate, while still preserving the characteristics of the objects embedded in the GPR images. The reduction of the amount of data to analyze constrains the area of interest. We work on two crucial aspects in this thesis: (a) visualization improvement of objects in GPR images and (b) identification of the relevant data necessary for automatic identification of objects. With the first one we seek favoring visual identification and classification. The second one aims at providing data classification systems that are able to identify the most relevant information to achieve automatic detection at reasonable computational cost and in an efficient way for the user. The algorithm presented in this chapter has been implemented in MatLab.

This chapter is organized as follows. In the first section we have shortly introduced the work developed and have related this work with the one in previous chapters. The next section introduces the principles of GPR image segmentation and cleaning focusing on buried pipelines. Then Section 5.2 details the segmentation and cleaning methodology proposed in this chapter. Next, Section 5.3 proposes a method for analyzing data based on the amount of data obtained with the application of the various processes integrating the methodology proposed in the chapter. Then, experimental studies, together with its process and analysis, are presented in (Section 5.4). Finally, a section of conclusions closes the chapter.

5.1 Principles of the segmentation and cleaning proposed

The idea underlying the proposal in this chapter is to look for general characteristics of the objects registered in GPR images. To this aim, first the characteristics of the medium surrounding the sought objects are identified, so as they can be removed from posterior images. This cleaning activity aims at getting images containing only characteristics associated to the objects of interest for the exploration. We thus manage to bound the search area, thus favoring the application of posterior analyses.

The characteristics we sought are anomalies that break either sharply or smoothly the observed trend or continuity of the measured parameters in the recorded images. To this aim, we first identify in the image the typical (or recurrent) characteristics of the medium and, then, determine the nature of these images and how they are associated with the medium properties. To achieve this, we propose a specific test with the objective of detecting a first approximation to the problem. This test is made of three static data captures. The schematic configuration of these captures are presented in Figure 5.2.

The first capture (Figure 5.2,a) is intended to show how the signals constituting the traces (columns) behave in homogeneous conditions, that is to say, for the no-object setting. As captures in media of homogeneous characteristics is extremely complex, this test was performed by placing the antenna in the center of the tank and registering signals during an arbitrary time lapse tc .

With the other captures, recorded in the same homogeneous medium, we registered the object passage under the antenna at various depths. To this purpose we placed the antenna in the same position as in the first (no-object) capture and then we displaced the pipe through the antenna measuring space. The capture time, tc , has a domain $[0, t]$, where t is the time necessary for the pipe to complete the entire planned route (Figure 5.2,b) in captures 2 and 3. By displacing the

5.1. Principles of the segmentation and cleaning proposed

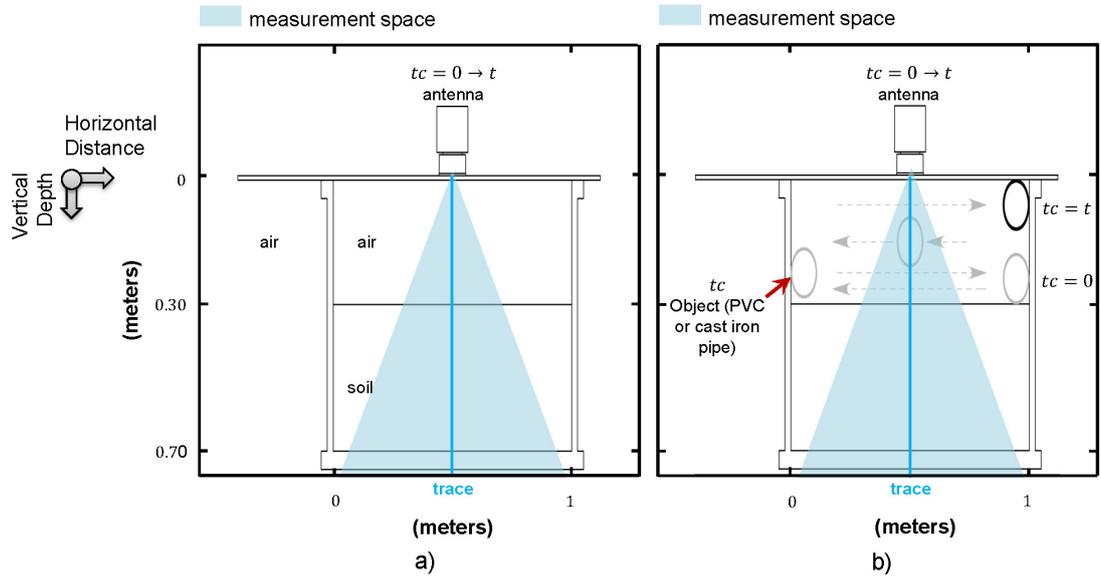


Figure 5.2: Schematic configuration, static tests – Pipe movement along the space. (a) Capture of traces without any object, (b) capture of traces with an object

pipe along the antenna measuring span we seek to simulate the displacement of the antenna over the objects of interest under spatial homogeneous conditions.

Taking into account the big differences between the registers for plastic and metallic objects (already addressed in Chapter 3), captures 2 and 3 were taken by using the same configuration (Figure 5.2,b), just varying the pipe, that is to say the material. PVC was used for plastic materials (capture 2) and cast iron represented the metallic materials (capture 3).

The equipment parameters used in these tests correspond to the assay configuration number 12 (see Appendix B). The GPR images resulting from these tests are presented in Figure 5.3.

5.1. Principles of the segmentation and cleaning proposed

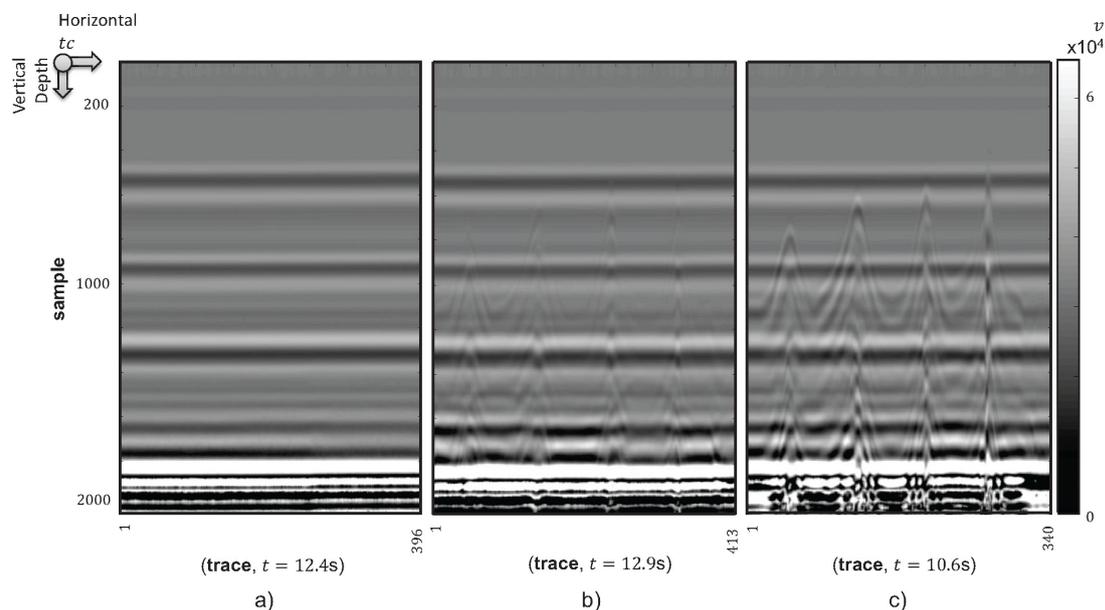


Figure 5.3: Raw images, static tests – Pipe spatial movement. (a) Capture 1; traces for the no-object test, (b) capture 2; traces for plastic object, and (c) capture 3; traces for metallic object

Figure 5.3,a shows that, under relatively invariable spatial conditions, GPR images appear as straight horizontal lines whose color intensity depends on the wave amplitude of the explored medium. In insets (b) and (c) of Figure 5.3 it is easily observed how the passage of the object through the measuring space and the variation of the depth of the object (Figure 5.2,b), generates a wave with increasing intensity as the depth diminishes. Also, it can be observed how the response observed in the image, associated to the passage of the object through the medium, couples with the response (straight horizontal lines) associated to the medium. The most evident variation observed in insets (b) and (c) is given by the color intensity, which in (c) is more clearly demarcated than in (b) and, as a result, better visually perceived.

By simplifying the problem in order to make visible the characteristics obtained in Figure 5.3, we proceeded to visualize the images by using contour lines. We have termed v the wave amplitude values obtained from the GPR images. Firstly, we identify equal values for v , such that those data altogether determine

5.1. Principles of the segmentation and cleaning proposed

different lines for different values of v at distance, h , among them, and such that the resulting lines do not intersect with each other (isolines). Observing Figure 5.3 it can be seen that the most problematic visualization corresponds to the passage of the plastic object (Figure 5.3,b). To cope with this problem the selection of the spacing between the contours lines was performed iteratively looking for a value h that allowed observing the main characteristics when performing capture 2. Eventually a value of $h = 1500$ was selected. We then proceeded to detect the contour lines for captures 1 and 3 (Figure 5.3 parts a and c) using this value of h .

The resulting contour lines are presented in Figure 5.4.

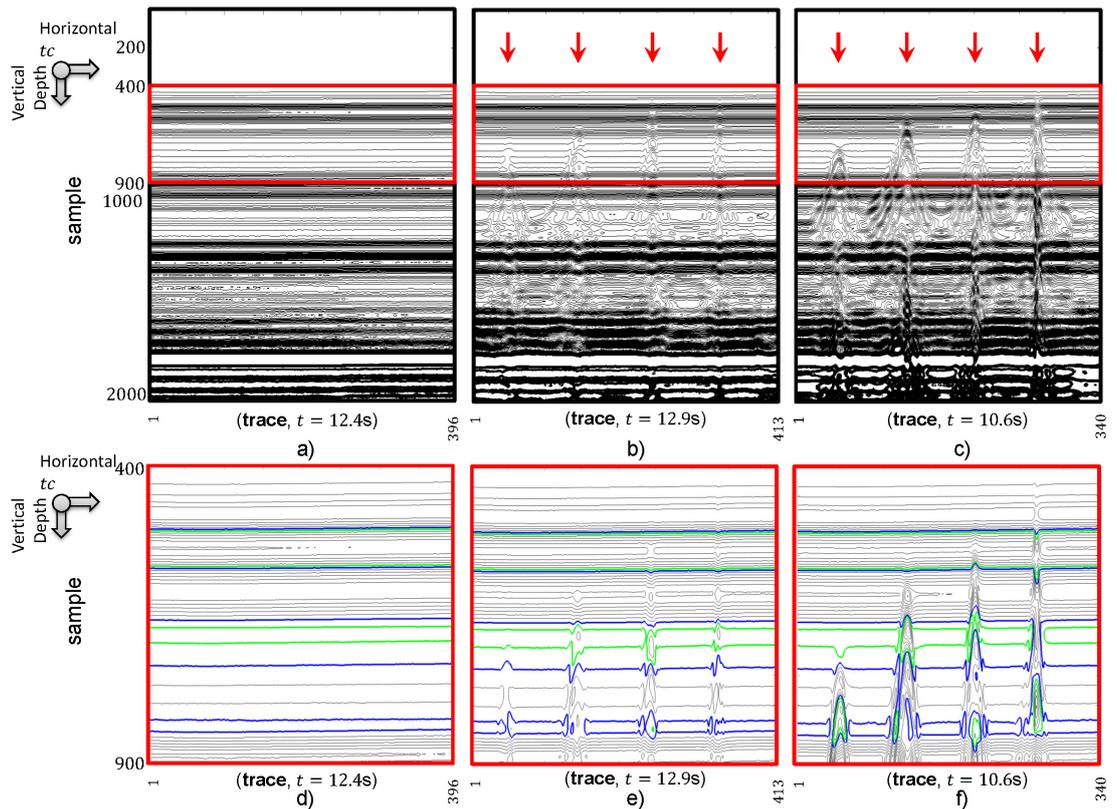


Figure 5.4: Contour lines for raw data with $h = 1500$, static tests – Spatial pipe movement. (a) Isolines without object, (b) isolines for plastic object, and (c) isolines for metallic object. (d–f) Zoom of red rectangle (a–c), respectively

5.1. Principles of the segmentation and cleaning proposed

The problem simplification performed (from Figure 5.3 to Figure 5.4) improves the visualization of the passing objects through the measuring space (Figure 5.4,b and c) and allows better comparison with the medium without anomalies (Figure 5.4,a). The traces produced by the passage of the centre of the object through the measuring spaces have been marked using red arrows in Figure 5.4,b and c.

The signals received in GPR surveys are stored in a matrix, A (radargram, raw data), that is made up of m -vectors, X_j , $j = 1, \dots, n$, (trace, vector), that represent the variation of the soil's electromagnetic properties in terms of depth. Let us represent this matrix by columns $A = [X_1, X_2, \dots, X_{n-1}, X_n]$. The length, m , of vectors X_j , is the volume of signal data recorded for each vector, which depends on the characteristics of the equipment used. The matrix corresponding to the preprocessed image (isolines in this case) will be termed A' , which retains the original size of A ($m \times n$). By ordering the data as they appear as a function of depth, these data may be represented by groups (families) of functions. This can be observed in Figure 5.4. This way, using contour lines enables us to discriminate these data using families of functions. In Figure 5.4,a, corresponding to the medium without anomalies, the family of functions is made out of straight lines which develop along the rows of the image matrix (straight horizontal lines). We denote this family fl . The number of functions integrating this family is noted ii and its domain is $[1, m1]$. The length, $m1$, will depend on the applied process to show the data that best represent both the medium and the embedded objects (relevant data) contained in the analyzed GPR image. In addition, $m1$ will be always lower or equal than m .

In Figure 5.4,a some functions breaking the smoothness of members of fl can be observed. The new individuals only affect some member of the family, without having vertical effect in all the family members. We note the new individuals by r . They represent the noise generated or just the smoothness break on the straight horizontal lines produced by effects different than the object measurement.

In parts (b) and (c) of Figure 5.4, obtained after the passage of the object (four times for the presented captures) through the measurement zone, there is a

5.1. Principles of the segmentation and cleaning proposed

tendency to break the linearity of each member of fl , which is distorted by the new family of functions $\{g_{ii}\}_{obj}$ that correspond to the anomaly which disrupts the space registered in the image. obj represents the number of registered objects in the image. Thus, the matrix A behavior is a priori defined as the array of functions in a matrix shape $A'f$ that is composed by the fusion of function families such as:

$$A'f = \begin{pmatrix} fl_1(i, j, v) + \{g_1\}_{obj}(i, j, v) + r_1(i, j, v) \\ fl_2(i, j, v) + \{g_2\}_{obj}(i, j, v) + r_2(i, j, v) \\ \vdots \\ fl_{m1-1}(i, j, v) + \{g_{m1-1}\}_{obj}(i, j, v) + r_{m1-1}(i, j, v) \\ fl_{m1}(i, j, v) + \{g_{m1}\}_{obj}(i, j, v) + r_m(i, j, v) \end{pmatrix}, \quad (5.1)$$

being $i = 1, \dots, m$ and $ii = 1, \dots, m1$.

In parts (d) and (f) of Figure 5.4, there are shown member representations $A'f$ for any v (highlighted blue lines) and for a $v - h$ (highlighted green lines). By considering that two objects cannot occupy the same space simultaneously, we can claim that every member of the family generated by each object $\{g_{ii}\}_{obj}$ affects to the function fl_{ii} in a range $[a_{ii}, b_{ii}]$ such that $\{a > b \wedge a \geq 1 \wedge b \leq\}$. We can claim that the object is partially in the image if $\{a = 1 \wedge b = n\}$. Otherwise, we claim that the object is totally developed in the image. The importance of clarifying this issue is because along the current thesis we focus only on objects whose range is totally or partially within the image, but not covering the whole image. The object O_{obj} source is approximately the lower of i in g_{obj} for all $\{g_{ii}\}_{obj} \neq 0$.

We define each fl_{ii} as a horizontal straight line, stating that $fl_{ii} = c_i$. Figure 5.5 is proposed in order to observe this constant. This figure is a representation of the wave amplitude values vs. samples, for each of the three assays carried out in the raw A matrix.

5.1. Principles of the segmentation and cleaning proposed

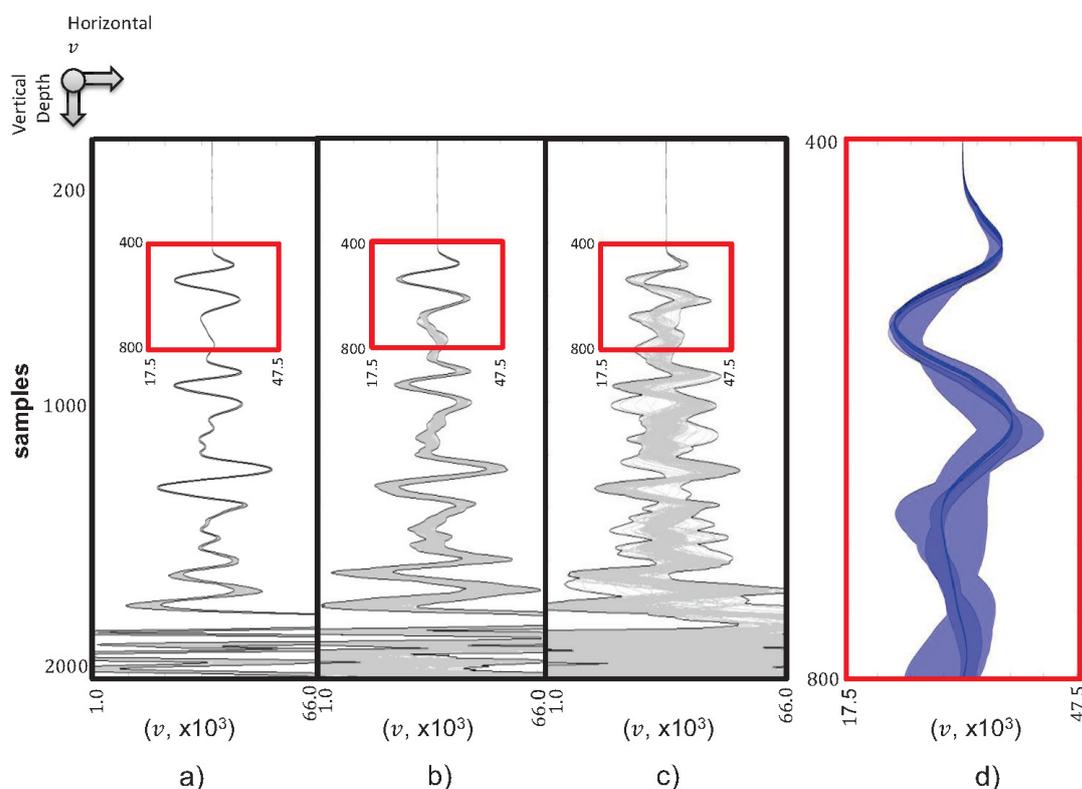


Figure 5.5: Raw data traces, static tests – Pipe movement. (a) Traces without object, (b) traces with plastic object, (c) traces with metallic object, and (d) maximum and minimum trace envelop for the captures

Figure 5.5 shows the variation in wave amplitude for the image with no object under the same starting conditions. These variations are not representative enough when there are compared objects in the same row (Figure 5.5,a). However, the variation between rows becomes significant by introducing plastic objects (Figure 5.5,b) and still more with metallic objects (Figure 5.5,c). The variation of these differences between the three figures is shown in Figure 5.5,d.

We can take into account the multi-agent systems performance to find out plastic pipes in the $T14$ or $T15$ matrices (Section 4.3 – Chapter 4). Three different agents were employed in the following three phases: (1) *roll-on*, (2) *termite-eating* and, (3) *final detection*. The first one was in charge to eliminate squared areas representing low intensity of a specific color. To this purpose, this

5.2. Segmentation and cleaning methodology

agent makes horizontal displacements throughout the image. In some way this agent seeks a similar data behavior than the presented by the fl family for the $A'f$ matrix (Equation 5.1). We know that we do not search for horizontal straight lines in the GPR images; that is, fl members. Because of the good performance already observed in applying multi-agent systems to visualize objects made up of different materials, in the following section we introduce a search and classification method for detecting those not desired elements corresponding to the more likely data of representing the absence of objects (fl family).

5.2 Segmentation and cleaning methodology

The removal of non-likely location zones of buried pipes in GPR images is carried out starting from the raw images and by applying a suitable combination of methodologies. Our proposal takes the radargram obtained from the surveying with GPR (raw matrix, A) as input for a three stage process: (a) binarization, (b) segmentation and cleaning and (c) refinement. These stages are detailed in the following sections. The corresponding schematic approach is shown in Figure 5.6.

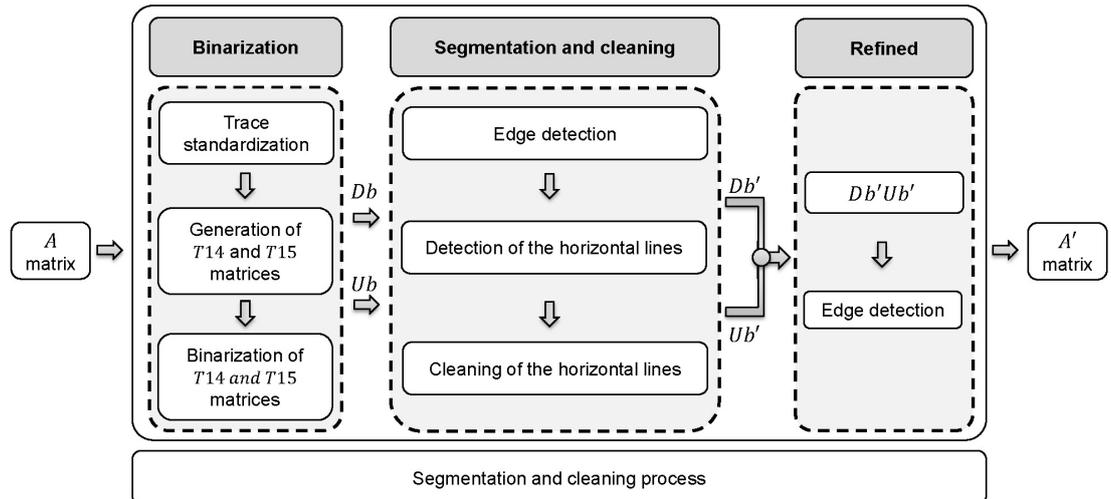


Figure 5.6: Segmentation and cleaning methodology proposed

5.2. Segmentation and cleaning methodology

5.2.1 Binarization

This first stage of the process is proposed through three steps: (a) standardizing traces, (b) selecting trends, and (c) binarization. The first step consists in leading every trace to a common (standard) form to enable its comparison and better management. The second step takes into account the already raised idea for the T_{14} and T_{15} transformations (see Chapter 3).

5.2.1.1 Standardizing traces

Data capture rates for each trace in the used equipment are 512, 1024, 2048 and 4096 samples/trace; all of them are common values for commercial applications. In this step, traces are led to the maximum rate. The step is approached by interpolating data and the subsequent equivalence between the different traces applied. This enables a correspondence between traces plus a greater consolidation of the forms corresponding to the image. The potential precision loss after interpolating values for the wave amplitude is not significant, since the subsequently used procedure applies trends between data rather than between values. In addition to this rate equivalence, interpolating traces allows to find out the true trace peaks. This is essential in this chapter since clipped traces generate additional data, augmenting the noise associated with the images. Among the interpolations most commonly used in GPR to correct and find the true peaks, linear, polynomial, and cubic spline [Jol (2009); Petrinovic (2009)] must be quoted. This work uses cubic spline interpolation. An example of a trace is shown in Figure 5.7.

With the interpolation, the clipped wave parts have been corrected (Figure 5.7). We also use interpolation to get a finer data discretization. Thus, we take the traces values from the original amount to a constant value (4096 samples/trace), which allows the comparison between radargrams with different rates of capture (samples per trace).

Once a standard working rate is obtained, we proceed to work with absolute values of wave amplitude in radargrams. Anomalies are alterations of the wave

5.2. Segmentation and cleaning methodology

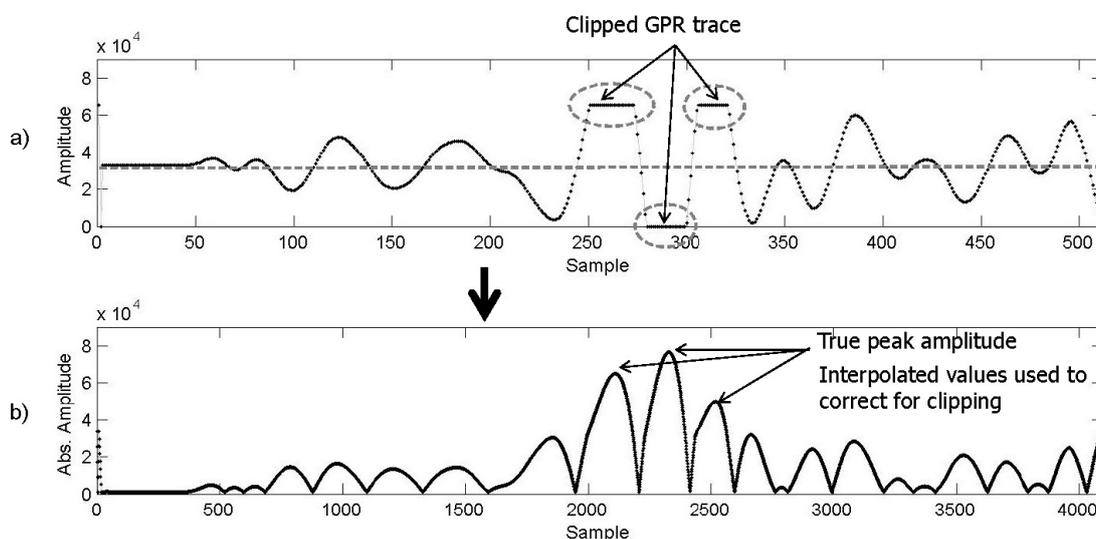


Figure 5.7: Trace representations (a) Trace, and (b) trace in absolute values; example of how the interpolation function can be used to correct the clipped form of the traces

in the image that are marked with great intensity at peaks and valleys obtained, which are maximum and minimum wave amplitude values, respectively. Working in absolute values allows that anomalies are only sought in peaks (Figure 5.7,b), which significantly reduces run-times and demarcates better the image forms contours.

Every value is summed up with the wave amplitude absolute value (this is 1,000 in this case) in order to distinguish such values from the zeros obtained after the following procedures. These traces are accumulated making up a new image to be processed.

5.2.1.2 Selecting trends

In Section 5.1 we have observed that the value h chosen among the wave amplitude values allows to obtain a synthesis of values (isolines), which improve the embedded objects in image visualization. However, choosing a value for h is

5.2. Segmentation and cleaning methodology

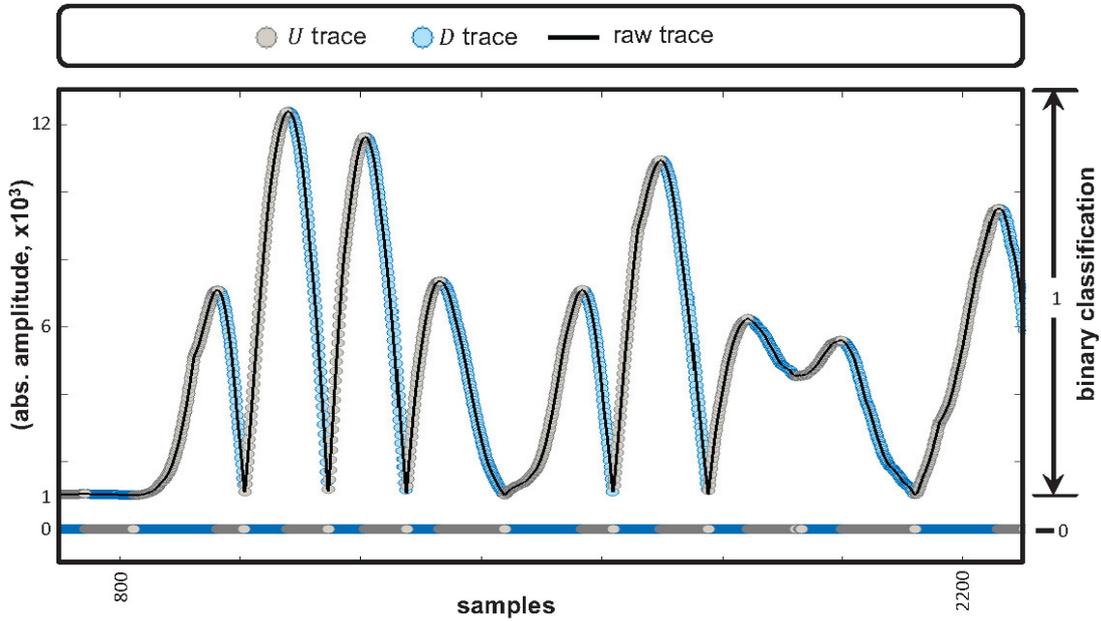


Figure 5.8: Selection of wave path trends

something subjective, so we decided to rely on the maximum and minimum wave amplitude values as a first bounded measure for the $T14$ and $T15$ matrices.

After trace standardizing, the following step is framed in monitoring the wave amplitude value paths. Thus, by following the wave it can be determined if the path is increasing or decreasing (top or bottom part, respectively). In the case of data corresponding to a bottom part, it is classified as $T14$ by the algorithm and in case it correspond to the top part, it is classified as $T15$. In the case of intermediate values, it is computed their trend to carry out the classification as lower (D) or upper (U) (Figure 5.8).

For the proposed methodology, the changes are handled as binary matrices, that is $Db = \{1, \text{ for } T14 > 0 \vee 0, \text{ otherwise}\}$, and $Ub = \{1, \text{ for } T15 > 0 \vee 0, \text{ otherwise}\}$, when the selected trend is increasing or decreasing (top or bottom part, respectively, see Figure 5.8). This is how the new algorithm proposed in this chapter avoids the obtained layers and traces of transformations to be handled in a binary way. By accumulating all traces comprising the radargram,

5.2. Segmentation and cleaning methodology

the output matrices for the first stage of the process, called *Db* and *Ub* matrices, are computed. These two matrices are binary but fully represent the initial matrix.

5.2.2 Segmentation and cleaning

Segmentation is the process to splitting images in regions or objects of similar attributes. Each segmented region usually has a meaning within the image. The objective is to simplify or to change the image representation by another more significant and easier to analyze. Segmentation is used for both to locate objects and to find out their limits within an image. This allows extracting objects from an image for their posterior description and recognition.

Segmentation tools can belong to three blocks: tools based on border detection, thresholding tools, and tools based on pixel aggregation. In the tools based on border detection, segmentation can be tackled by detecting every region limits, that is, detecting images edges. The other two techniques approach segmentation as a pixel or group of pixels classification problem.

The Hough transform, belonging to border detection tools, is used in pre-processing to detect horizontal straight lines and to remove them from the image (cleaning). The amount of data that make up each line and their features for taking the removal decision are tuned by iterative tests until reaching a stable configuration that allows to identify and to eliminate unlikely zones for pipe location in the image.

The segmentation and cleaning stage is managed in three steps: (a) edge detection, (b) straight lines detection, and (c) removal of straight lines. The first step involves edge detection from the *Ub* and *Db* matrices. To do it, it is applied the order 4 convolution Laplacian of the 3×3 mask matrix. Once the edges are detected, the data composing the image matrices are mapped to the Hough parametric space, in order to only identify horizontal straight lines. In the third step, the horizontal straight lines detected are removed. After applying the three

5.2. Segmentation and cleaning methodology

steps above exposed to Db and Ub matrices, there are obtained two new matrices called Ub' and Db' . Each of the steps in this phase is in more detailed presented below.

5.2.2.1 Edge detection

The first step of this phase is to detect the edges in the Db and Ub matrices. The Laplacian is applied to this purpose since it is an operator that works independently of any orientation and edges can be defined as limits between two regions with different values. Thereby, the transition between two regions can be regarded as a sharp change. Edge detection is based on the computation of the derivative, taking into account that for any constant the value of such derivative is zero and for a change this value is different to zero. The Laplacian is a scalar operator of the second derivative and for two-dimensional functions is defined by Equation 5.2.

$$\nabla^2 f(x, z) = \frac{\partial^2}{\partial x^2} f(x, z) + \frac{\partial^2}{\partial z^2} f(x, z) . \quad (5.2)$$

The above equation is implemented approaching the matrix convolution by a 3×3 mask region (Figure 5.9).

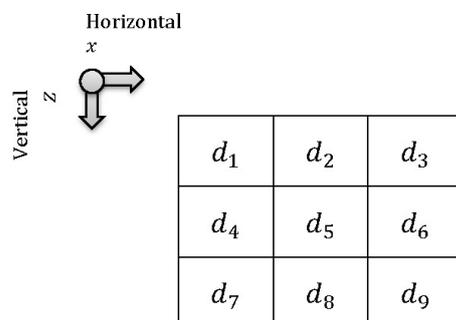


Figure 5.9: Image region, 3×3 mask

5.2. Segmentation and cleaning methodology

A discrete function is computed by taking the second derivative in columns (x) and rows (z); that is, the gradient differences are approached as indicated:

$$f_{xx} = f(x + 1, z) - 2f(x, z) + f(x - 1, z) , \quad (5.3)$$

$$f_{zz} = f(x, z + 1) - 2f(x, z) + f(x, z - 1) , \quad (5.4)$$

adding the components of the two equations (Equation 5.3 and Equation 5.4), and implementing it in a 3×3 mask, it is obtained the Laplacian applied in this chapter:

$$Laplacian = \begin{vmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{vmatrix} . \quad (5.5)$$

The principle for the definition of this operator is that the coefficient associated with the central data (d_5 , Figure 5.9) is positive and the coefficients associated with the exterior data are negative, such as the sum of all the coefficients is zero. By approaching the mask convolution with this matrix, the result is zero when the central data has the same value than its neighbors. For the particular case of a binary image, the Laplacian operator is chosen with central member $d_5 = 4$, defined in Equation 5.6, and stating that if the value of the central point is different from any of its 4 neighbors, it detects an edge [Urrea & Ospina (2004)].

$$\nabla^2 f = 4d_5 - (d_2 + d_4 + d_6 + d_8) . \quad (5.6)$$

5.2.2.2 Horizontal straight lines detection – Hough transform

Once the edges are detected, the Hough transform is applied to detect lines in the Db and Ub matrices. The Hough transform is a tool that allows to isolate shape features (curves) in an image. It is a robust technique in front of noise and the existence of gaps in the border of the object [Al-Nuaimy (1999)]. The

5.2. Segmentation and cleaning methodology

technique is based on finding curves that can be parametrized as straight lines, polynomials and circles. It has been usually applied in GPR images for detecting the first reflection hyperbola given by the objects. Some examples can be found at Simi *et al.* (2008) and Al-Nuaimy (1999). In addition, there are works as Rahkar-Farshi *et al.* (2014) where modifications in Hough transform are done through the Artificial Bee Colony algorithm to optimize the time to detect multiple hyperbolae in GPR images.

A line segment is analytically described in various ways, one of them is in polar coordinates the parametric or normal notation (Equation 5.7). This is the kind of notation we use in this chapter,

$$\rho = x \cos \theta + z \sin \theta \quad , \quad (5.7)$$

where ρ is the length of a normal from the origin to the line and θ is the angle of ρ regarding the x axis. For instance, in Figure 5.10 it is presented the application of the Hough transform, after edge detection, in the Db and Ub matrices for the Section 5.1 tests (see Figure 5.2 configuration).

Figure 5.10 shows great similarity between D and U matrices per each capture and between captures in the Hough space. We also observe that the greatest accumulation of possible straight lines is presented with a normal of $\theta = 90^\circ$ for low ρ , which represents that the majority of curves at the images correspond to horizontal straight lines. Moreover, it increases the accumulation for ρ and θ different from horizontal straight lines by adding an object into the image. This confirms the proposal made in Equation 5.1 regarding the fl family, which is distorted with the addition of the g and r families. In this step only horizontal straight lines are sought, on the basis of the above observations and taking into account that pipes can be represented as g families in the images. Thus, Equation 5.7 reduces to Equation 5.8, with the subsequent decreasing of computational time, thus improving these processes.

$$\rho = z \quad . \quad (5.8)$$

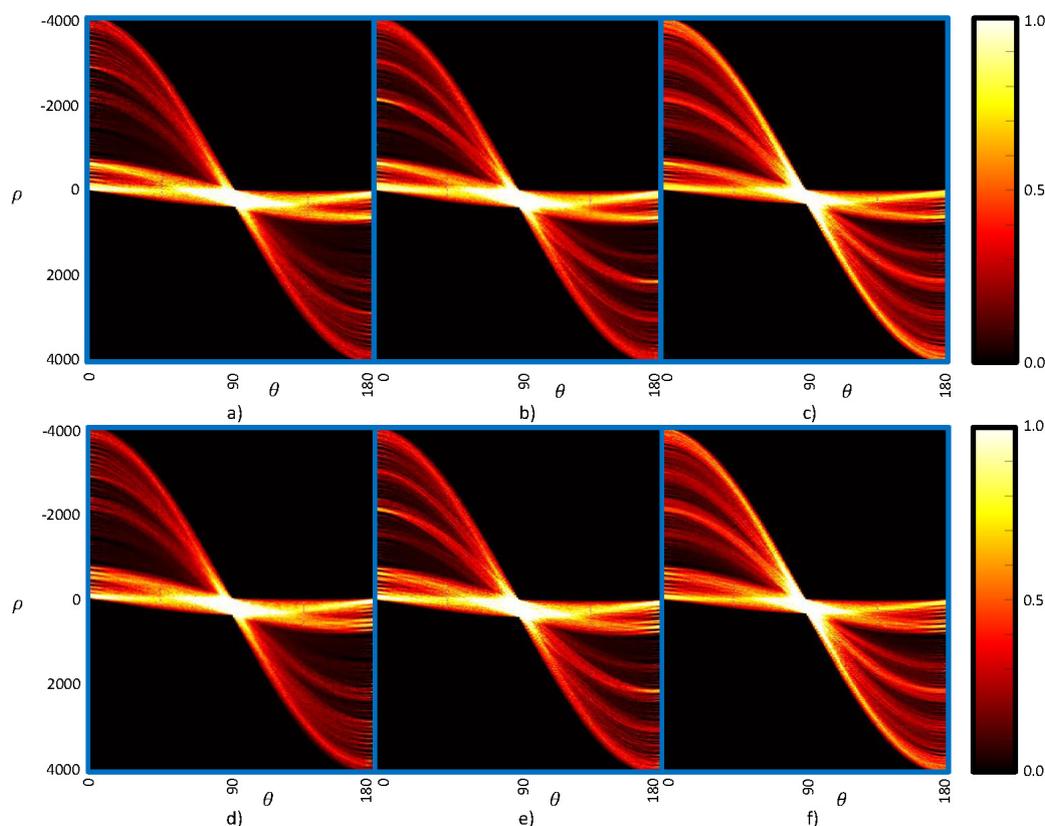


Figure 5.10: Hough transform, Hough parametric space to detect straight lines. (a–c) Matrices D_b , and (d–f) matrices U_b . (a) and (d) Capture 1 (no object); (b) and (e) capture 2 (PVC); and (c) and (f) capture 3 (Fib)

5.2.2.3 Removing lines

After the data parametrization by applying the Hough transform, a new representation of the parametric space by a cell structure whose elements is obtained by accumulating cells $Acum(\rho_i, \theta_i)$ with $\theta_i = 90^\circ$. The accumulated cells with magnitude within an established threshold can be considered as potential horizontal straight lines. The determination of the threshold is performed by varying three characteristics: number of traces to evaluate, and maximum and minimum amount of data that make up the line. The variation of the characteristics is performed in radargrams of pipes with different configurations of equipment and

5.3. Density analysis

pipes of different materials. The threshold for determining the data to be preserved is established in the evaluation of the radargram every 10 traces, with a data accumulation given by $3 \leq Acum(\rho_i, \theta_i) \leq 6$. Data with accumulations outside of this threshold are eliminated. By applying the three previous steps of this stage to the Db and Ub matrices, we obtain two new matrices called Db' and Ub' .

5.2.3 Refining

In this process stage the Db' and Ub' are added up. The edge detection process (Section 5.2.2.1) is applied to this new created matrix. The objective is to eliminate redundant data and minimizing the noise. The output matrix of this process is called A' and corresponds to the segmented and cleaned initial image.

5.3 Density analysis

Once the pre-processing of segmentation and cleaning has been applied, we go back to the configuration shown in Figure 5.2, where we observe the variation of the data behavior with the application of the different processes that make up the image. Pipes of two additional materials are included in this test in addition to the two pipes previously used. In total, these materials are PVC, polyethylene (PE), asbestos cement (Fib) and cast iron (Fund), all of them with a nominal diameter of 0.1 m.

In this case, again a reference capture (capture 1; Figure 5.2,a) is prepared. Each capture is made three times for a total of 15 tests. We call density to the amount of data obtained in each process among the total volume of data. The analyzed process results are two binary matrices, Db and Ub , obtained from $T14$ and $T15$ matrices, respectively. Since the pre-processing output is a combination of these matrices, we also analyze the sum of the amount of data in Db and Ub . To observe the number of selected points to be removed due to the fact that they

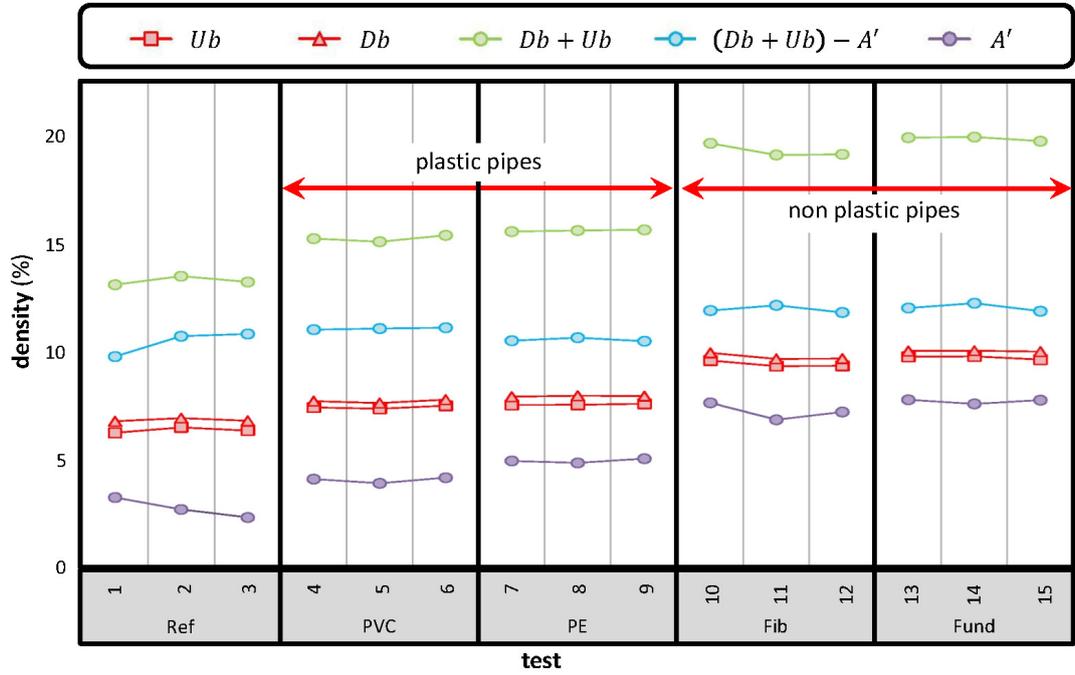


Figure 5.11: Density analysis of the different processes that make up the proposed pre-processing – Under distance homogeneous conditions

are horizontal straight lines, we subtract the amount of data retained in A' to the amount of data in $D_b + U_b$ and we also work with A' . The number of data of these five matrices are divided into the initial amount of data ($m \times n$) of the raw matrix, with which we get what we call density per process. The results of these tests are shown in Figure 5.11.

For the five types of tests, Figure 5.11 shows a degree of correspondence between the obtained densities in each process through the respective iterations (3 iterations per test). The figure also shows that the process clearly decreases the amount of data at the output (A' matrix). The necessary data to analyze in homogeneous conditions regarding the distance, is lower than 8% of the total for each image. We observe that the amount of non-eliminated data (data not representing horizontal straight lines at the reference condition) is lower than 4%. This value indicates that the r family (noise) has density values lower than 4%

5.4. Experimental studies – segmentation and cleaning

in A' , in homogeneous conditions, that is to say, unless there is some additional source of noise. In other words, the noise in the A' matrices (residual noise) is of around 4% of the data obtained, plus the obtained noise from additional sources. It can be seen that the passage of the objects (pipes in this case) is recorded at every process except in the space of $((Db + Ub) - A')$, which reaffirms the hypothesis we made in this chapter. In addition, we can see that the density of data in the remaining processes shows, under these conditions, a tendency to separate plastic and non-plastic objects in a natural way. This relation is marked in the $Db + Ub$ and A' matrices containing all the data.

5.4 Experimental studies – segmentation and cleaning

This section presents the application of the proposed segmentation and cleaning method described in Section 5.2. The case study corresponds to the GPR images of four pipes of different materials usually employed in WSSs. The pipe materials used in the tests are PVC, polyethylene (PE), asbestos cement (Fib) and cast iron (Fund) with a nominal diameter of 0.1 m. The GPR images are obtained by burying pipes in dry ground in a laboratory tank. GPR surveying is carried out with a reference capture (Figure 5.12,a) and burying each pipe separately. The equipment parameters for these tests correspond to the configuration of test number 70 (see Appendix B).

5.4. Experimental studies – segmentation and cleaning

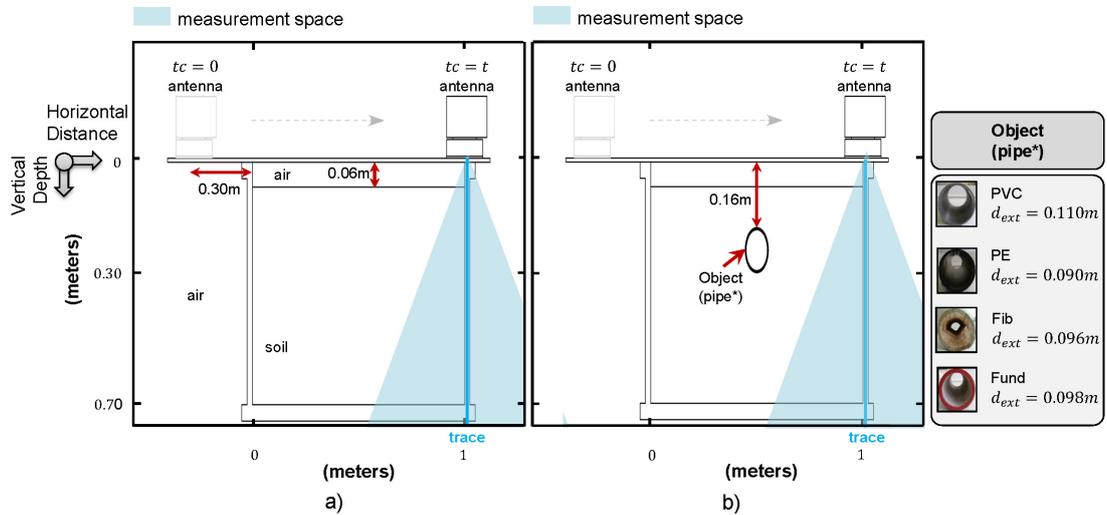


Figure 5.12: Schematic configuration, experimental studies. (a) Capture reference configuration, and (b) capture configuration with pipes

A pre-processing is done to the images obtained under this configuration. The analysis is divided into two sections. Firstly, we show the results obtained in the proposed pre-processing. Secondly, we present a complementary exploratory analysis of the amount of obtained data (density analysis).

5.4.1 Segmentation and cleaning - pre-processing

The raw results of the survey conducted are represented in Figure 5.13 (a – e) in gray-scale. The images obtained after application of the segmentation and cleaning process are presented in Figure 5.13 (f – j).

5.4. Experimental studies – segmentation and cleaning

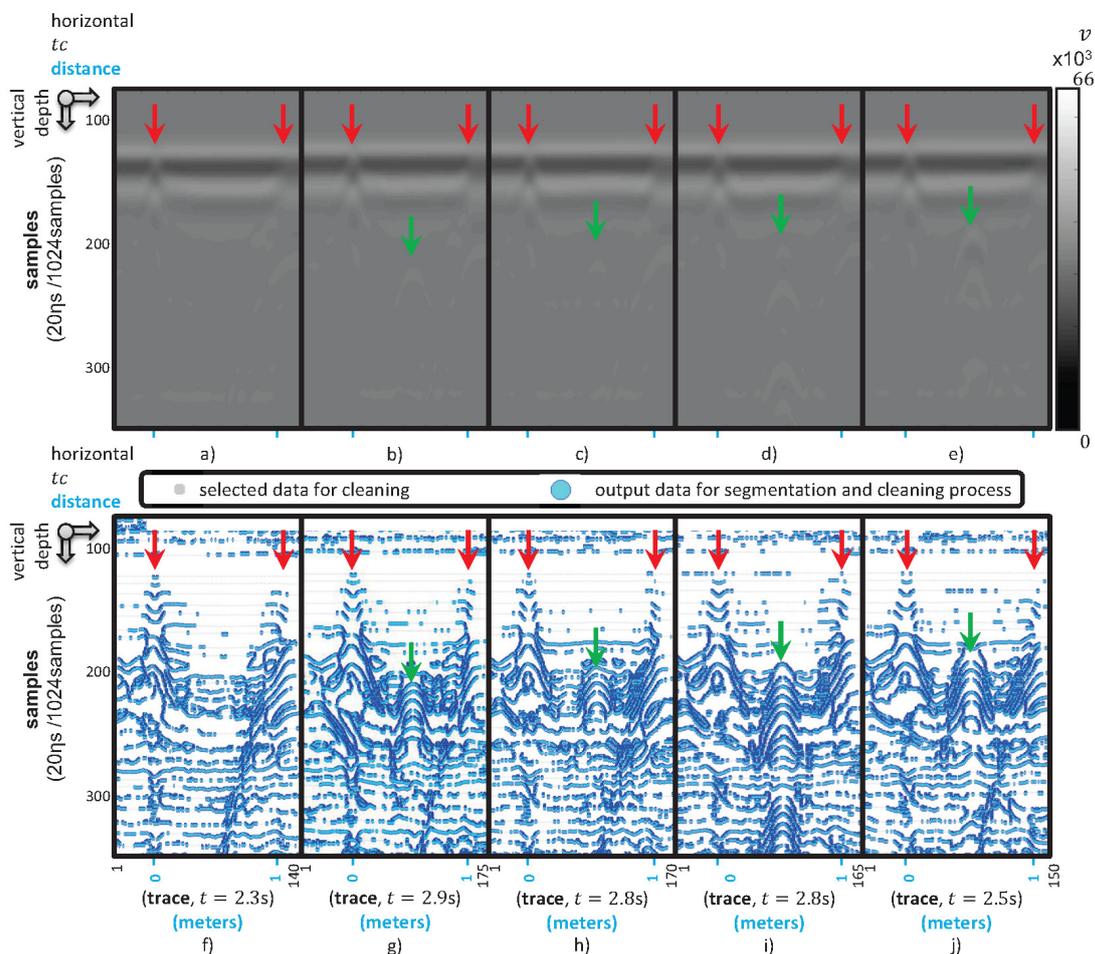


Figure 5.13: Application of the proposed pre-processing. (a–e) GPR raw images, and (f–j) pre-processing result. (a) and (f), Reference; (b) and (g), PVC; (c) and (h), PE; (d) and (i), Fib; and (e) and (j), Fund

In the images of raw data shown in Figure 5.13 (a – e), we can notice that the contrast of borders is not strong enough to be easily identified by direct visualization in PVC and PE pipes. This is a result of low wave amplitude characteristic of these materials and causes a low color intensity that does not allow a strong demarcation of PVC and PE lines (almost invisible). Moreover, Fund and Fib pipes are also difficult to visualize, though the contrast is more pronounced.

5.4. Experimental studies – segmentation and cleaning

Figure 5.13 (f – j) show a sensible improvement in the image shapes and features after the application of the segmentation and cleaning process when the result is compared with their respective raw images. In addition, we can observe that the images show the behavior arisen at the matrix analysis (Equation 5.1). In the same way, we observe a fl family (data selected to be cleaned), such as the family members that are distorted by members of families generated by both objects (g family) and noise (r family). Thus, Figure 5.13,f demarcates the tank walls containing dry soil, which are also demarcated in the same way in parts (g) to (j) of Figure 5.13 (red arrows). Between these two walls a space variation is observed that corresponds to buried pipes as in parts (g) to (j) of the figure (green arrow). The green arrow is located on the first observed deflection (depth) that is attributed to the pipes. By transferring such arrow to its corresponding position in the raw image, we see that a first hyperbola is demarcated at shallow depth. This effect has been usually applied to identify objects in GPR images. The same happens for the highlighted effect by the red arrows (tank walls). In both cases, we see that applying segmentation and cleaning process demarcates other characteristics related to the objects, beyond those corresponding to the first hyperbola. This enable us to make further interpretations of the images. The observed objects in the case study images (tank walls and pipes) correspond to those raised in the schematic configuration of the test (Figure 5.12).

Figure 5.14 shows the expansion of the area where the pipe is observed (Figure 5.13,g – j) and its corresponding space for the reference image (Figure 5.13,f).

In Figure 5.14 we can observe and contrast, using the reference image (a), the marking of the different buried pipes (b – e) after the segmentation and cleaning process of the raw images. In the images obtained through the pre-processing it is possible to observe consolidated and well defined hyperbolae formations, aggregated in a similar way to the hyperbola families ordered in depth (g family). The application of the segmentation and cleaning process to the raw matrices also increases the visualization of these pipes in the images, by enhancing their features, such as the hyperbola asymptotes of the pipe in the image.

5.4. Experimental studies – segmentation and cleaning

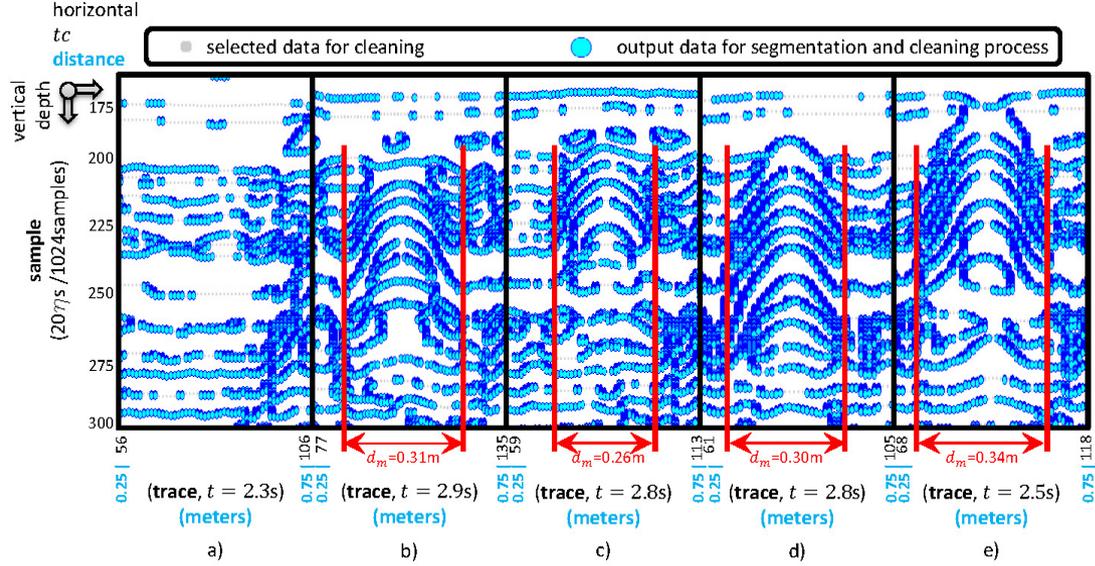


Figure 5.14: Expansion of pipe localization zone in the resulting images from pre-processing. (a) Reference, (b) PVC, (c) PE, (d) Fib, and (e) Fund

From the schematic configuration (for pipes) (Figure 5.12,b) we know that the pipe axis is located at 0.5 m from the internal tank wall. In the images we find out that the hyperbola center is approximately located to 0.5 m from the walls (Figure 5.14,b – e). Thus, as we already defined (Section 5.1), the g families have a range $[a_{ii}, b_{ii}]$ containing to each family member. The ranges with a larger distance between a and b (measured diameter d) for each family at every material analyzed are: PVC, $d_m = 0.31$ m; PE, $d_m = 0.26$ m; Fib, $d_m = 0.30$ m; and Fund, $d_m = 0.34$ m (Figure 5.14,b – e). In we relate d_m to every material with their respective external diameters (Figure 5.12), d_{ext} , we have for each material at given metering conditions, a certain distortion (def) by material: PVC, $def = 2.8$; PE, $def = 2.9$; Fib, $def = 3.2$; and Fund, $def = 3.4$. This shows that for the used depth (0.16 m), pipes present diameters between 2.8 and 3.4 times larger depending on the material.

After raw data, obtained from GPR surveys in laboratory on different pipe materials, has underwent the processing methodology herein proposed, regarding the effect on the various used materials, we can say the following for the proposed methodology.

5.4. Experimental studies – segmentation and cleaning

- **PVC.** It improves image visualization and increases the amount of features related to the signal passing through, when it is compared to raw data analysis.
- **PE.** It allows visualization of this type of material, having a better object demarcation when it is compared to raw data analysis.
- **Fib. and Fund.** It improves visualization of image features, such as asymptotes, and presents further consolidation in the pipe image.

The segmentation and cleaning process was tested in a total of 1,140 pipe images of various materials (PVC, PE, Fib, Fund) and GPR prospected references. By the application of this process we have a significant improvement in more than 85% of the cases, where the image properties were preserved while the amount of data to be analyzed was reduced. The final result of the process should be the reduction of unlikely zones for pipe locating entailed by the segmentation and cleaning of the raw image.

5.4.2 Density analysis

After applying the described pre-processing to the tests, this section introduces the density analysis of the results (process) in a similar way to the analysis already presented in Section 5.3. The density results are obtained by performing tests under homogeneous distance conditions (see configuration Figure 5.2) of the same various pipe materials commonly used in WSS (see Figure 5.11). We obtain a number of density values, which naturally classify data of different pipe materials. This study is performed for the different processes applied in relation to the total data comprising the image analyzed. This first density relationship has a value for $\{(D + U) / \text{total data}\}$ lower than 25%, while pipes are grouped at densities above 14%. However, given that the conditions imposed in this test are far from being the conditions found in the field (under uncontrolled conditions), we claim that this homogeneity condition (in distance) can be obtained by splitting the image into smaller visualization areas, where the obtained densities for the

5.4. Experimental studies – segmentation and cleaning

$\{(D + U)/ \text{total data}\}$ ratio are around 30%. This value is chosen by considering the above-mentioned 25% and assuming a 5% more due to effects of objects not yet considered. The densities for the rest of the matrices are obtained using the number of images needed to meet this condition.

Let us represent by C any of the matrices obtained from the different processes outlined in this chapter, namely, $\{Db, Ub, Db + Ub, (Db + Ub) - A', A'\}$, where all of them are $m \times n$ binary matrices. This matrix may be divided into blocks of a size corresponding to the dimensions of matrices envisaged for non homogeneous cases (field cases). Let $(lx \times lz)$ be the dimension of all these blocks. The new matrix structure will be represented by blocks $C'_{ir,jr}$ where $ir = \{1, \dots, lz\}$, $jr = \{1, \dots, lx\}$, and $(lx \wedge lz) \in \mathbb{N} = \{1, 2, \dots\}$. In the blocks, indices iz and jx run, respectively: $iz = \{1 : (m - 1)/lz : m\}$, and $jx = \{1 : (n - 1)/lx : n\}$. The proposal is that

$$\left(\left(\sum_{iz_{ir}}^{iz_{ir+1}} \sum_{jx_{jr}}^{jx_{jr+1}} c_{(iz_{ir}, iz_{ir+1}], (jx_{jr}, jx_{jr+1}]} \right) \cdot \frac{lz \cdot lx}{(n - 1)(m - 1)} \right) \cdot 100 < \approx 30, \quad (5.9)$$

that is to say, that each block exhibits a density ratio lower than 30%.

The results of this proposal for the test presented in Figure 5.12, are shown in Figure 5.15.

5.4. Experimental studies – segmentation and cleaning

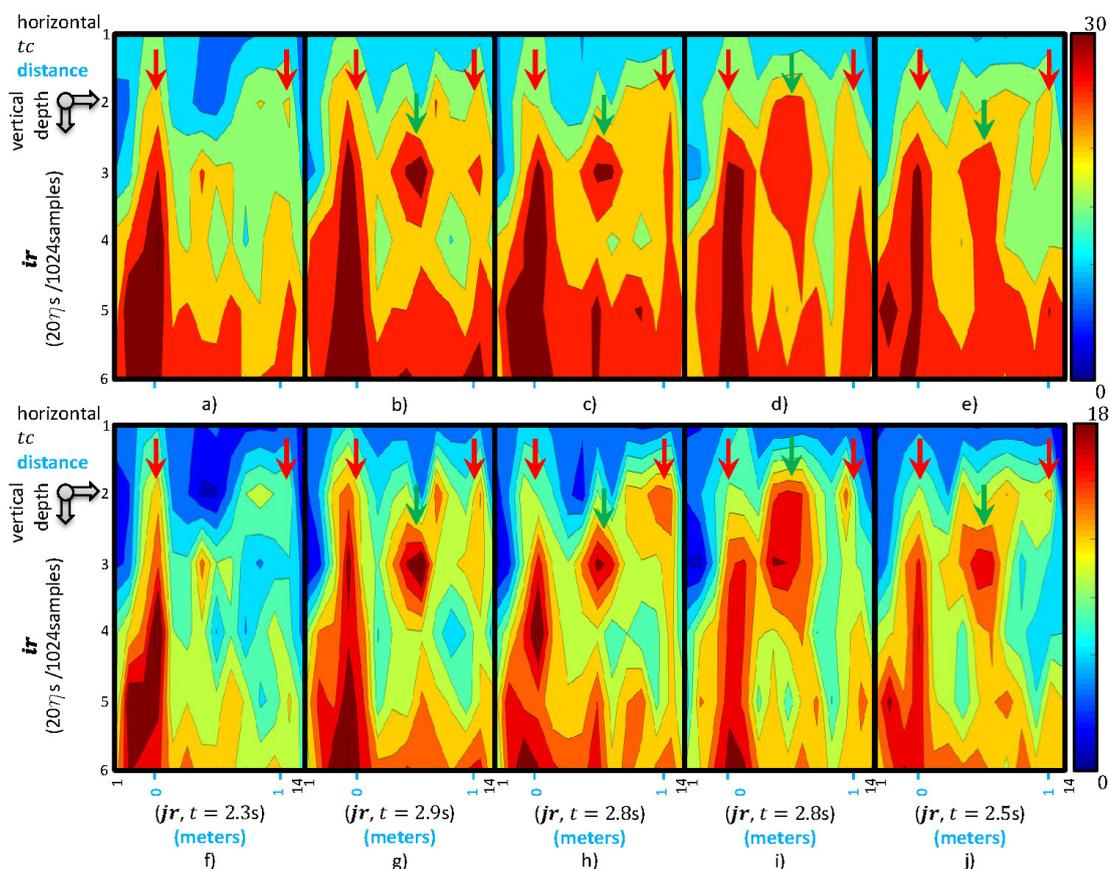


Figure 5.15: Density maps. (a–e) Density maps for $C = D + U$, and (f–j) density maps from $C = A'$. (a) and (f), Reference; (b) and (g), PVC; (c) and (h), PE; (d) and (i), Fib; and (e) and (j), Fund

In Figure 5.15 (a – e) the pair (lx, lz) , when $C = D + U$, for the images is (6,14), with a maximum density per test of: Ref 36%, PVC 35%, PE 36%, Fib 33%, and Fund 34%. Above 30% the density values were considered as 30%. From this series of blocks, it is obtained a density map that is presented in Figure 5.15 (a – e). Using the same pair (lx, lz) the densities for $C = A'$ have been computed, whereby the density maps observed in Figure 5.15 (f – j) are obtained. The red and green arrows shown in the different images retain the same position in distance than the position shown in Figure 5.14. Similarly, the distance in meters on the horizontal axis is given. The position of the arrows also marks the starting point of the object in depth.

5.5. Summary and comments

Figure 5.15 shows that the density maps for conditions $C = D + U$ and $C = A'$ improve GPR data visualization. In insets (a – f) of the figure, the objects and their interactions are demarcated by high density values. In insets (f to j) of the figure the areas containing the objects that caused the deformation are demarcated in more detail.

We can observe in Figure 5.14 that the green arrows are still indicating the hyperbola family associated with the buried pipe. Neither in part a, nor in part f, are the deformation demarcated with such intensity. In insets g to j, the object densities (green arrow) are increasing toward the center with a maximum concentration observed in block 3 (in depth) for the 4 cases. This indicates that through a density analysis it is possible to perform a posterior classification of the object characteristics.

5.5 Summary and comments

In this chapter we assessed the feasibility of obtaining a general tool that allows us to identify objects (specifically pipes) in GPR images. In the exposition of the fundamentals of this chapter, we have found primarily a matrix $A'f$ representing the behavior of the data as a collection of functions (families), consisting of three different spaces: medium (fl), objects (g) and noise (r). The relationship of these spaces builds the initial image, and each space has its own characteristics that differentiate them from each other.

An initial approach led us to find a way to separate these spaces, for which we proposed a segmentation methodology based on the Hough transform to detect and eliminate horizontal straight lines (family fl) and analyze the two remaining spaces (families $g + r$). The analysis was based on the wave amplitude values, using the transformation of the raw image (raw matrix, A), which was obtained with the methodology proposed in Chapter 3 (matrices $T14$ and $T15$).

By burying four pipes of different materials in dry soil in a laboratory tank and after information capture through GPR surveying, the results showed a considerable difficulty in object visualization for weakly reflective materials (PVC, PE). The weak marking of plastic material pipes in raw GPR images is consequence of their particular wave amplitude values. These values are lower than the corresponding to other materials and do not contrast enough with those of the surrounding environment. The non-plastic materials (Fib, Fund) showed certain demarcation, although their identification presented high degree of difficulty in raw images. Therefore we can say that in raw images, visual identification of pipes, regardless of the material, is especially difficult, subjective and without guarantee of repeatability, being necessary the participation of workers with considerable experience in GPR data analysis to identify pipes of different materials.

The separation of spaces by the pre-processing proposed in this chapter improves the image visualization pipes buried in dry soil. In general, after the application of this pre-processing, it is possible to work with a considerable improvement in the visualization and consolidation of the pipe characteristics in the GPR images, observing the pipes clearly delineated in contrast with what happens in the raw data. This is the case of the space distortion at a given depth, which can be related to the diameter of the buried pipe. The approach showed that it is possible to get, beyond the first hyperbola, a family of hyperbolae (g) which gives a more direct approach to the captured object in the image, allowing to progress towards more detailed analysis.

The density analysis of the data previously classified as families of functions (fl , g and r) shows consistency among the various parts of the process. We found out the existence of a relationship between the different process and the added objects, making it possible to limit the influence space of the families (g) by dividing the original space into smaller images. This relationship provides an approximation to more comprehensible images (schematic configuration).

Finally, it should be noted that the pre-process proposed in this chapter is easy to apply and does not depend on the consideration and judgment of a specialist (it is not a subjective process). It is repeatable and also, with appropriate

5.5. Summary and comments

adjustments, allows to parametrize captured radargrams of pipes in WSS, which may serve as a basis for training intelligent classification systems.

Chapter 6

Automatic generation of WSS pipeline layout maps based on GPR images

This chapter is intended as a step forward towards the development of a tool that addresses the problem of lack of knowledge of piping layout in WSSs. We try to generate (automatic) pipeline routing maps from images obtained through the use of GPR as a non-destructive method. In addition, we develop appropriate tools for easy viewing of the results. This is finally intended to improve water supply efficiency and turn management of WSSs more dynamic, in such areas as data collection, inspection, decision making, prevention, and maintenance and implementation of plans. Chapter 5 has shown the importance of suitably bounding the search space, so that once the interest space has been delimited, using the demarcation obtained by the characteristics of the objects in the images, more comprehensive information and more detailed analyzes can be obtained. In this chapter, we seek to detect the spatial continuity of the objects recorded in the GPR images. This will allow us to discriminate pipes from other objects, since pipes exhibit significant spatial continuity. This chapter uses the space segregation obtained in Chapter 5 which made use of the therein proposed segmentation and cleaning methodologies (pre-processing) to evaluate the spatial continuity of the objects recorded in the GPR images. The output provided by this methodology will be the GPR image matrix after having taken off all horizontal straight lines. The scope of this chapter, in the framework of this thesis, is presented in Figure 6.1.

A preliminary version of this approach (approach 1) was presented in the Congress *XI SEREA* [Ayala-Cabrera *et al.* (2012b)]. A short version of the same approach was presented in the Congress *Mathematical Modelling in Engineering & Human Behaviour 2012* [Ayala-Cabrera *et al.* (2012a)]. Afterward, the complete approach was reformulated (approach 2) and published in the journal *International Journal of Complex Systems in Science* [Ayala-Cabrera *et al.* (2014b)].

6. AUTOMATIC GENERATION OF WSS PIPELINE LAYOUT MAPS BASED ON GPR IMAGES

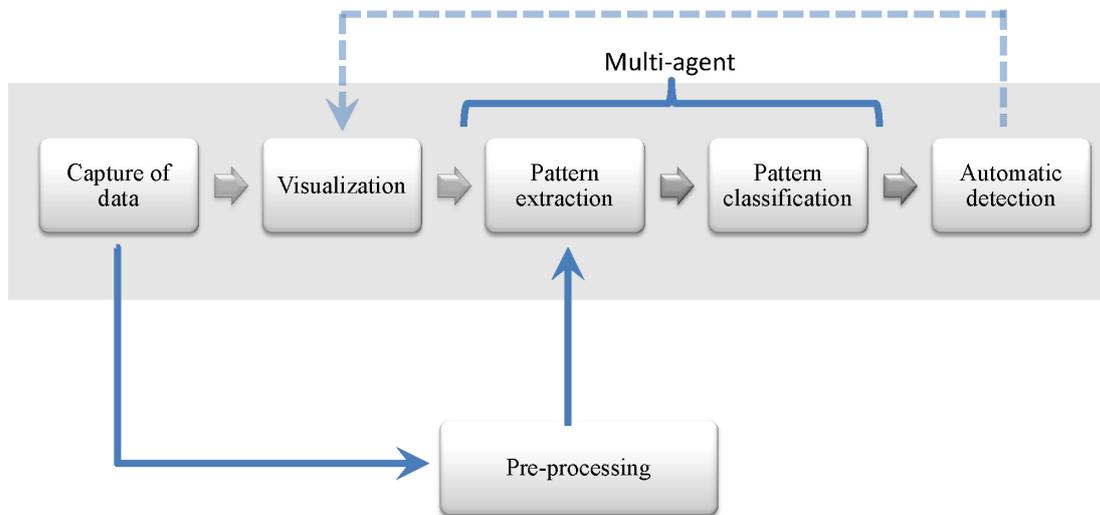


Figure 6.1: Chapter 6 scope

In this chapter, two approaches to detect pipe paths in GPR images are explored: (1) capture routes, and (2) analysis routes. The first approach is aimed at enhancing GPR data capture that can expeditiously identify the route of a pipeline. The second is devised to analyze data quickly and automatically, once it has been captured.

The purpose of the first approach improves the usual exhaustive scanning systems with GPR, thereby the survey time is optimized and the amount of data needed to conduct records of WSS components is minimized (approach 1). Approach 2 presents a probability-based system and a multi-agent treatment of the captured GPR images to serve as a guide to highlight relevant information for mapping WSS pipelines, once the data has been captured, trying to make fast identifications of the objects found without the operator needing considerable skill and experience, thus seeking to minimize the time taken to obtain reliable results of the data analysis.

The main contribution of the approaches outlined in this chapter lies in the automatic detection of courses of WSS pipelines using non-destructive methods (in this thesis, GPR) and by personnel not highly qualified in GPR image analysis. The approaches proposed in this chapter attempt to unify the main three aspects

of pipeline mapping, namely, capture, analysis and interpretation and mapping, in an adaptive way that serves as a guide for recording and / or analyzing the information relevant to WSS pipeline mapping. The algorithms developing these approaches, including the activities of the agents, have been written entirely in MatLab.

This chapter is organized as follows. The first section has briefly introduced the work developed in this chapter. Section 6.1 provides an introduction to the use of GPR images in WSS component mapping. Section 6.2 raises two approaches and their respective architectures for detecting piping layouts automatically. In Sections 6.3, 6.4 and 6.5 a development to generate inspection routes (approach 1) and probability maps (approach 2) for mapping WSS pipelines using GPR images is presented. Subsequently, the implementation of the proposed systems in experimental studies, together with processing and analysis details, is presented. Finally, a concluding section closes the chapter.

6.1 Introduction to the use of GPR for WSS component mapping

Close knowledge of piping layouts and features in WSSs is necessary. However, many WSSs have no information about installed ducts or changes thereof performed; there are WSSs that do not have adequate reference systems, or have very incipient systems. In some cases, paths and piping characteristics, such as diameter, material condition, time of use, among others, are unknown. However, these aspects are necessary to make appropriate decisions and take actions, regarding technical management of WSSs. This knowledge is critical in areas such as detection of illegal connections, estimated volume of leaked water, approach of operating system settings, study of the evolution of pollutants in the network, maintenance, rehabilitation and renovation of supply networks.

The search for WSS assets within conventional management activities is usually carried out through excavation [Royal *et al.* (2011)]. This process involves increased economic and social cost. Thus, at present, non-destructive methods have been fostered in the pursuit of these components. However, although information retrieval using non-destructive methods is successful, the specific way data is captured, the large volume of generated information, and the need of useful interpretation require high level of skill and experience.

The GPR has already been used as a non-destructive method for analyzing WSS components and anomalies. The most analyzed components are pipes, metallic materials being those most frequently evaluated (see Chapter 3, Section 3.1). Currently, detection of pipes using GPR and subsequent mapping is conditioned by the experience of the operators. Sampling (inspection and capture) campaigns are performed following the operator interpretation of GPR images obtained after various passes of the equipment, and the eventual marking, on the surface of the ground, of the area where the pipe is located, according to the operator belief. After multiple passes of the equipment and subsequent marks, the layout of the assets underground is eventually determined. The pipeline mapping

6.1. Introduction to the use of GPR for WSS component mapping

process can be summarized in three stages: (a) GPR image capture, (b) image analysis and interpretation, and (c) pipeline mapping. The second and third stages have been addressed by multiple authors. As a result, procedures for analysis and interpretation of GPR images adapted from other survey methodologies may be applicable to WSSs. (Chapter 3, Section 3.1.1).

Regarding the mapping process of such urban infrastructures as WSSs and waste water collection systems, various works, using GPR, have been developed. In Chen & Cohn (2011) data observed by operators, location plans, and data obtained by GPR are combined using a Bayesian approach to generate maps of the networks assets. In Dell'Acqua *et al.* (2004), a work of general character dealing with mapping of linear objects, a semi-automatic approach for the detection of dispersed linear objects in the set of GPR data, based in the Radon 3D transformation, is proposed. Gurbuz *et al.* (2010) uses an iterative detection technique and the strategy of eliminating linear and planar structures in 3D images obtained from GPR surveys. Al-Shayea & Bahia (2010) focus on GPR profile information to detect property changes (object detection), which is iteratively processed until the eventual location of the object in the underground. In Dutta *et al.* (2013), a framework for fusion of images coming from various sensors was proposed. Specifically, GPR and vibro-acoustic signals were used. In this work dynamic Bayesian networks to detect buried infrastructure, mainly focusing on WSS buried pipes of plastic materials, were used. The outcomes of this study revealed high accuracy in detecting pipes, after comparison of locations with the existing reference plans. Additionally, the results allowed pipe visualization through suitable 3D representations.

6.2 Proposal for detecting piping layout

The approaches proposed in this chapter are based on iterative analysis processes that aims to generate capture routes and, from the analysis of these captures, to generate automatic pipe routes (Approach 1). The second approach seeks, using profiles captured under an established configuration, to iteratively generate route analyzes and eventually produce, in an automatic way, probability maps of pipe paths, using the GPR as a sensor for both cases. Both approaches are described below.

6.2.1 Approach 1

This approach describes an iterative process that will guide data collection using GPR to generate paths of pipes. The system will consist of three stages: (1) capture GPR images, (2) image analysis and interpretation, and (3) route mapping. This system, through the interaction with the GPR equipment, will allow the proposed algorithm and the capture system (either performed by operators or, preferably, mechanically) to meet, in a very short time, the layout of the pipes, with the additional advantage of classifying the various sections of the route for further analysis of other features. The purpose of this work is to improve the conventional GPR scanning methods. As a result, the survey time is optimized and the amount of data needed to conduct records of the components of WSSs is minimized. The system architecture for Approach 1 is presented in Figure 6.2.

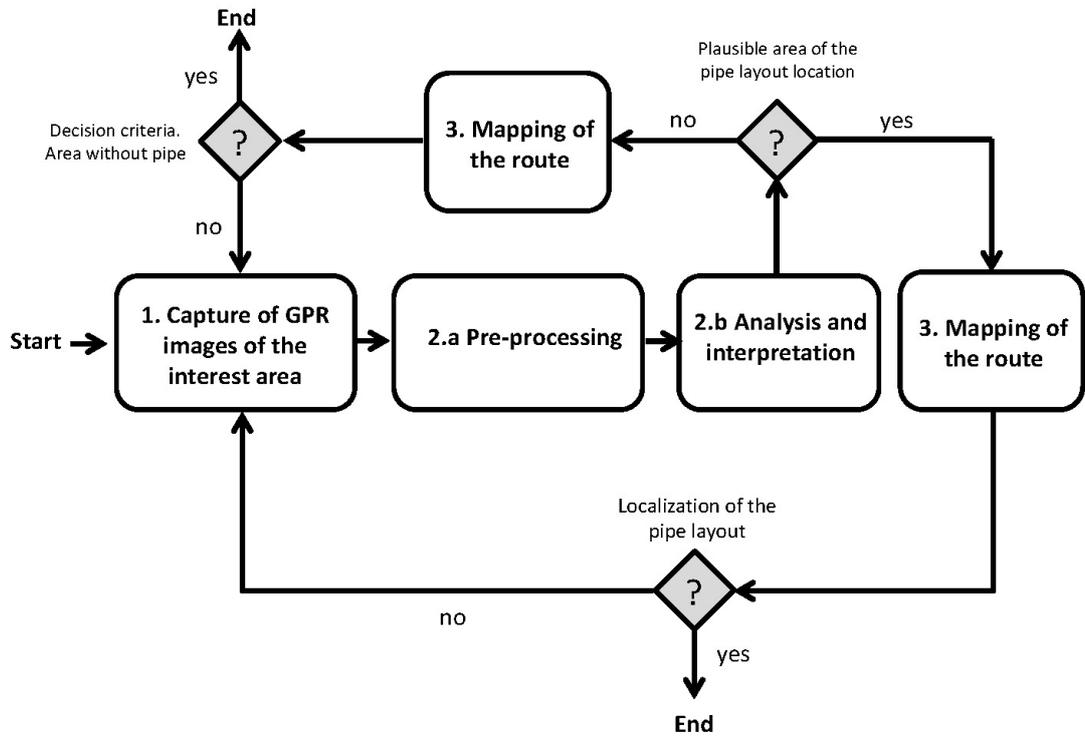


Figure 6.2: Schematic representation of the proposed system. Approach 1 – Automatic detection of piping layout

6.2.2 Approach 2

This approach is based on an iterative analysis process that aims to automatically generate probabilistic maps of WSS pipes using GPR, from profiles previously captured. The objective of this approach is to evaluate the collected data, and to generate efficient analyzes of data routes. The proposed architecture for probabilistic maps is based on an iterative analysis process based on a multi-agent system. This procedure is performed by agents endowed with a knowledge engine that simulates a stochastic process that propose different pseudo-random walks aimed at studying the features provided by the GPR data. In this way, the areas with lower probability of existence of pipes are disregarded by the agents. This procedure is applied iteratively to obtain a process that produces aggregate

6.2. Proposal for detecting piping layout

interpretation of the agents' points of view about the area of interest, and that eventually shape the map locating the buried pipes. The proposed methodology in this approach consists of four steps: (1) GPR image capturing in the area of interest, (2) sampling the area of interest, (3) analysis and interpretation, and (4) generation of probability maps. The first step will be performed by the operator. The remaining three stages will be carried out by agents and supported by an appropriate combination of statistical procedures. The algorithms of these last three stages have been developed in MatLab and their activities are carried out by multi-agent systems. The agents proposed in this system simulate the overall behavior of expert staff to find pipelines in WSSs via GPR. The system architecture for Approach 2 is shown in Figure 6.3.

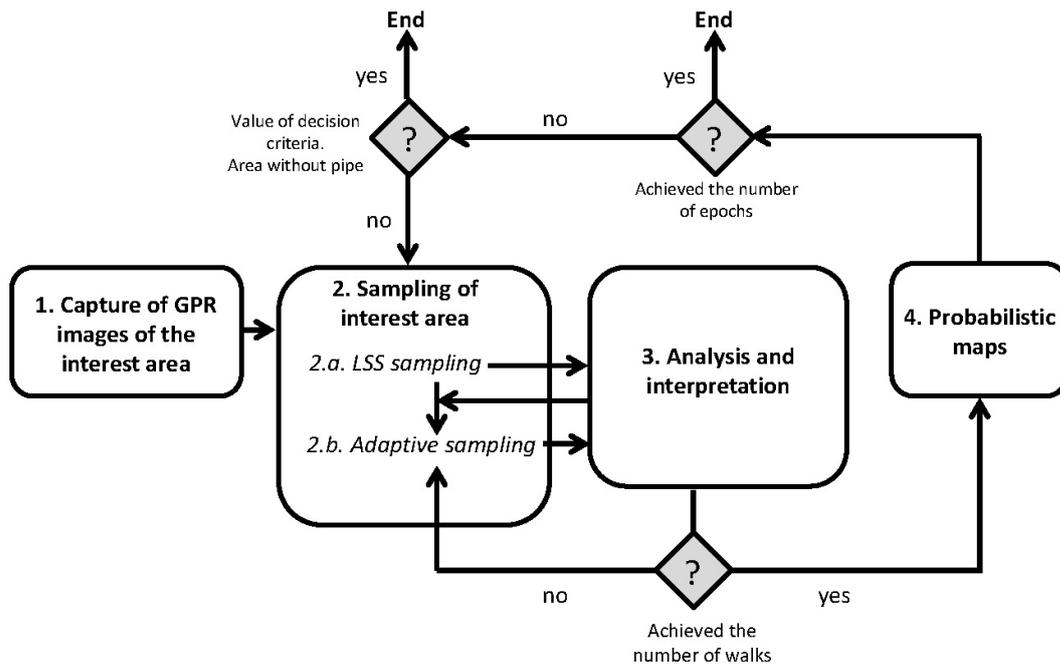


Figure 6.3: Schematic representation of the proposed approach, Approach 2 – Automatic detection of pipe paths

6.3 Approaches 1 and 2 – general aspects

To start with, a square area ($L \times L$ in meters) of interest is selected for inspection; features such as name, origin, axes and coordinate systems must also be established. In this area, images will be taken horizontally and vertically with respect to the axes; we call these images PCIs (profile capture inputs). In Approach 1, these images will correspond to each step taken. In Approach 2 the number of these profiles will be at least two in each direction, and will cover the entire area of interest. These profiles will be interpolated in Approach 2, thus generating a volume of raw or unprocessed data. The values for each point will be interpolated values of the wave amplitude. The number of profiles selected in this approach will influence the detail of the objects found, and facilitate the search. This will reduce the interpolation thereof and give greater reliability to the results. This type of sampling is proposed to minimize data capturing time, and to take into account that capturing is performed by personnel not expert in GPR image interpretation. Thus, the interpretation work and the search for possible pipe traces is assigned to the artificial agents.

As noted in the previous section, the two approaches have similar activities, the most important being data capturing. In the first approach, each profile is captured according to the analysis of the previous capture. Approach 2, however, starts with a predefined number of captures, which are performed in a single attempt by the capture system (operator or mechanical system). Both approaches will work under the premise of conditions of complete lack of knowledge with respect to the path of the pipeline, that is to say, the absence of any guidance (such as recognition of valves, plans, etc.) that would allow to intuitively assume the pipeline path. An example of the initial setting for each approach is presented in Figure 6.4.

6.3. Approaches 1 and 2 – general aspects

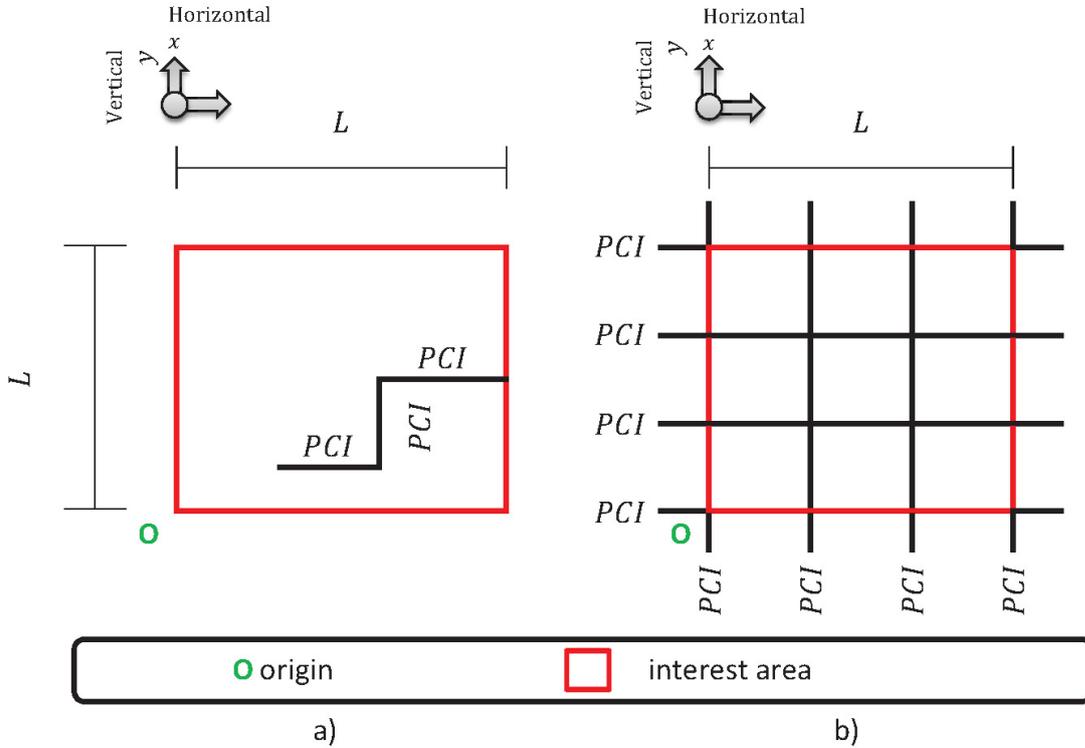


Figure 6.4: Demarcation of the interest area, PCI capture. (a) Approach 1, and (b) Approach 2

In this chapter we will call profiles to the images captured in the various activities. The profiles obtained in this chapter are of three types, depending on the activity to be executed. Thus, we consider initial capture profiles (PCI), profiles of Latin squares (PL) and adaptive profiles (PAd). The axes of the profiles captured are denoted by xm and z , and correspond to the total distance (in meters) captured and the number of captured data in depth, respectively. The coordinates of each profile for xm will be given by the initial pair (x_1, y_1) and its corresponding final pair (x_2, y_2) . The domain for these profiles is $\{NS, SN, WE, EW\}$, and is determined by $\{NS: (x_1 = x_2) \wedge (y_1 < y_2); SN: (x_1 = x_2) \wedge (y_1 > y_2); WE: (y_1 = y_2) \wedge (x_1 < x_2); EW: (y_1 = y_2) \wedge (x_1 > x_2)\}$.

6.4 Sampling of interest area

Once defined the space of interest, in this section we show how the profiles for study are selected. Sampling the area of interest will be the first activity to develop. In Approach 1 this activity is carried out by the capture system (manual or mechanical) and in Approach 2 by the agents. In this activity, two types of samples for both approaches are distinguished: (1) initial step based on Latin square sampling (LSS), and (2) adaptive steps. The first type of sampling makes a global sweep of the surface area of interest. Although for both approaches it is possible to define a random starting point, we have considered a sampling rate based on Latin squares, in order to increase the probability of success in finding the route of the pipeline. The second is based on the result of analysis of the data of the selected profile using the first type of sampling, so as to start from a point (within the area of interest) more likely to contain the sought pipe (if any). This second type of sampling serves to continue sampling the area with a higher probability of finding the route of the pipeline sought. The types of sampling performed are described below.

6.4.1 Latin square sampling (LSS)

We use the LSS technique to explore ground variation sources by means of randomly assigned tasks; this begins the process of finding the would-be buried pipe. The LSS provides the proposed system with the initial exploration criterion for the capture system (Approach 1), or for the agents (Approach 2) to start their searching walks. The LSS technique divides an interest area, usually rectangular, into a set of $m \times m$ plots organized matrix-wise into rows and columns, Figure 6.5, step 1. Then, a symbol from a set of m different elements, which can be identified with the elements of the set $S = \{1, 2, \dots, m\}$, is assigned to every plot (step 2), each symbol occurring exactly once in each row and exactly once in each column [Bailey (2008); Fisher (1926)]. In this case, we have used the symbols a, b, c and d. If each entry of an $m \times m$ Latin square is written as a triple (r, c, s) , where r

6.4. Sampling of interest area

is the row, c is the column, and $s \in S$, we obtain a set of m^2 triples called the orthogonal array representation of the square. A Latin square is the set of all triples (r, c, s) , where $1 \leq r, c, s \leq m$, such that all ordered pairs (r, c) are distinct, all ordered pairs (r, s) are distinct, and all ordered pairs (c, s) are distinct.

To build the PLs, one symbol is randomly selected and thereby m plots for the test are chosen (in this chapter $m = 4$ and the plots are square; see Figure 6.5, step 3). Once the m plots have been chosen, one route is randomly selected in every plot to start with the sampling process. The candidate routes have directions selected from the set $\{NS, SN, WE, EW\}$, these capital letters representing the four cardinal points, and proceed through the plot going past the center of the plot (Figure 6.5, step 4). Thus, any of these routes is an oriented straight segment with starting point (x_1, y_1) and end point (x_2, y_2) given by

$$\begin{cases} x_1 = x_2 = \hat{x}; y_1 = \hat{y} - (ss/2); y_2 = \hat{y} + (ss/2), & \text{if } SN \\ x_1 = x_2 = \hat{x}; y_1 = \hat{y} + (ss/2); y_2 = \hat{y} - (ss/2), & \text{if } NS \\ y_1 = y_2 = \hat{y}; x_1 = \hat{x} - (ss/2); x_2 = \hat{x} + (ss/2), & \text{if } WE \\ y_1 = y_2 = \hat{y}; x_1 = \hat{x} + (ss/2); x_2 = \hat{x} - (ss/2), & \text{if } EW \end{cases}, \quad (6.1)$$

where (\hat{x}, \hat{y}) is the center of the selected plot and ss is the length of its side. Here $ss = L/m$.

The obtained profiles will be numbered according to the column number of the plot within the Latin square (Figure 6.5, step 5).

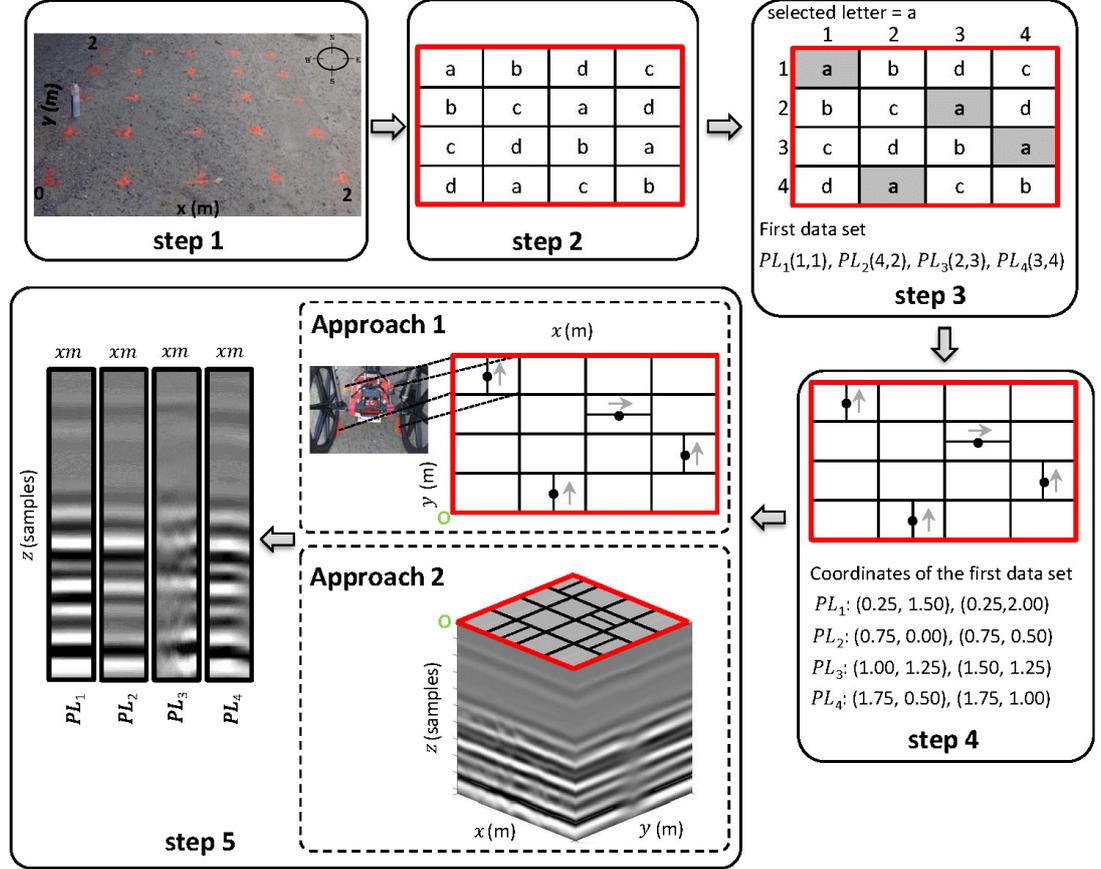


Figure 6.5: Latin square-based initial profile acquisition

In addition, a color representing some plausibility for the route to (partially) follow a pipe underneath will be assigned to any route. The process assigning such a color is described next.

6.4.2 Training for pipe existence plausibility

For the analysis, we flatten the matrix associated to the image at time t onto a vector, crv_t , by row concatenation. In addition, only matrix cells with a value of 1 will contribute to crv_t . Then, an auto-regressive model of order 4,

$$crv_t = \phi_1 \cdot crv_{t-1} + \phi_2 \cdot crv_{t-2} + \dots + \phi_{m_0} \cdot crv_{t-m_0} + a_t, \quad (6.2)$$

6.4. Sampling of interest area

is applied. $\phi_s = 1, \dots, mo$ are the parameters to be estimated, mo is the order of the auto-regressive model, and a_t is the white-noise disturbance.

It can be observed that the pipe existence plausibility in the image substantially increases the value of the model coefficients.

The plausibility indicator, which we call area color indicator (*aci*), belongs to the domain {green, yellow, red}: “green” represents high plausibility of pipe existence, “red” indicates low plausibility, and “yellow” is assigned if a conclusion is not clearly drawn. This indicator is obtained as follows.

Once obtained the parameters of the model for a specific route, the vector $\phi = (\phi_s)_{s=1}^{mo}$ is built. Then, a new vector, \vec{vt}_s , is obtained:

$$\vec{vt}_s = \{1, \text{ for } \text{abs}(v\phi_s) \geq R \vee 0, \text{ otherwise } \},$$

where R is a threshold. R has been obtained after analyzing a set of 200 classified images obtained both in lab and field tests.

A scalar vt is afterward obtained by accumulating the values of \vec{vt}_s :

$$vt = \sum_{s=1}^{mo} \vec{vt}_s .$$

Finally, the indicator *aci*, to establish criteria about pipe existence plausibility when considering a given profile, is built:

$$aci = \begin{cases} green, & \text{if } vt > 3 \\ yellow, & \text{if } vt = 3 \\ red, & \text{otherwise} \end{cases} . \quad (6.3)$$

6.4.3 Adaptive sampling route

After initializing the process with the LSS, including the *aci* plausibility index of the PLs, new sections are added to the routes for the agents to walk. This eventually will build a PAd. The starting situation can be any of the examples shown in Figure 6.6.

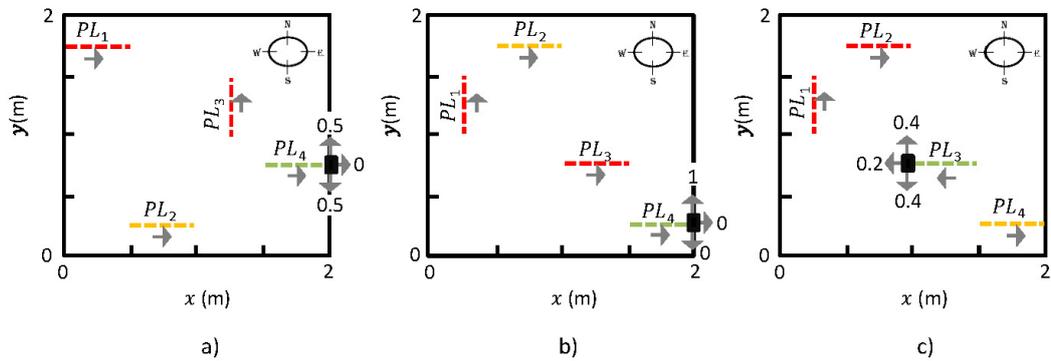


Figure 6.6: Examples of possible decision for the first step of the adaptive route according to end point position of a PL. (a) Side (no corner), (b) corner, and (c) interior of the interest area

For an agent to build up a PAd, firstly, one of the PLs and a starting point within it are selected, according to the following rules.

1. If $aci = \text{“green”}$ for just one PL, then the corresponding route for the PAd will start at the end coordinates of the route corresponding to the “green” PL.
2. If $aci = \text{‘green’}$ for more than one PL, then one of the “green” PLs is randomly selected and, again, the route for the PAd will start at the end coordinates of the route corresponding to the selected PL.
3. If $aci \neq \text{“green”}$ for all PLs, and there is just one PL with $aci = \text{“yellow”}$ the starting point for the PAd route will start at the end coordinates of the route corresponding to the “yellow” PL.

6.4. Sampling of interest area

4. If $aci \neq$ “green” for all PLs, and there is more than one PL with $aci =$ “yellow”, then one of the “yellow” PLs is randomly selected and, again, the route for the PAd will start at the end coordinates of the route corresponding to the selected PL.
5. If $aci \neq$ “red” for all PLs, one of them is randomly selected and the corresponding route for the PAd will start at the end coordinates of the route corresponding to the selected PL.

Then, an agent will take further movements to build up a PAd. The general rules for agent movements are: (a) no reverse direction in the walks, (b) only horizontal or vertical movements are allowed, and (c) perpendicular movements to the previous movement are favored.

To make a decision about the direction to move for the first segment of the PAd, the following specific rules are applied. These rules add constraints to the general rules, according to the position of the starting point within the interest area. The starting point may be located: (1) on one side (not a corner), (2) at one corner, or (3) in the interior of the interest area (see Figure 6.6). The decision is made in terms of a probability calculated for the various possible directions, {NS, SN, WE, EW}, as follows:

1. Starting point on one side (not a corner): both perpendicular movements are assigned 50% of probability, as seen in Figure 6.6, inset a).
2. Starting point at one corner: the whole probability is assigned to the only perpendicular direction allowed, as in Figure 6.6, inset b).
3. Starting point in an interior point: perpendicular directions get double probability than horizontal directions, as in Figure 6.6, inset c).

For the next movements an agent has an additional set of rules related to the fact that the area is limited. These rules will favor the inspection of areas already not walked. The rules, complementing respectively the previous three rules, are the following:

1. If one of the perpendicular ways has already been walked, the whole probability will be assigned to the other perpendicular way. If both perpendicular ways have already been walked, the agent will randomly select one of the ways, but just to skip it by jumping to its other end point.
2. If the remaining perpendicular way has already been walked, the agent will just skip it by jumping to its other end point.
3. If the subsequent horizontal direction has already been walked, the whole probability will be equally assigned to the perpendicular directions. If only one of the perpendicular directions has already been walked, the probability will be assigned 80% to the other perpendicular way and 20% to the horizontal one. If just two ways have already been walked, the whole probability will be assigned to the unexplored direction. If, finally, the three directions have already been explored, the agent will randomly select one of the ways as in rule 3, skip it and jump to its other end point.

Simultaneous with the various parts of the walks, *aci* indexes will be added to each section. After a number of agents have performed their walks, the interest area will be endowed with additional information in the terms of Figure 6.7. This information will be used later by a new breed of agents to help produce the probability map, as explained later.

6.4. Sampling of interest area

1. An agent will be created for each segment with *aci* “green” or “yellow”; the agent will initially be placed surrounding the segment. See Figure 6.8, inset (1,2).
2. An agent will explore perpendicularly the segment it was created from and will never go out of the interest area. See subsequent insets to the right, then to the second row of Figure 6.8.
3. Three of the six sensors will explore in one direction and the other three in the other.
4. Sensor points will move away to expand the agent as much as possible with the only constraint of stopping when encountering a segment with *aci* = “red” parallel to the direction of exploration. Only the sensor points encountering a “red” segment are affected by this rule. See insets (1,3), (1,4), etc.
5. If at the end of the process, two or more agents occupy the same area, it will be considered as a single area (last insets of Figure 6.8).

The patches of the areas identified by the agents are then assigned a value of 1, while the outer patches will be assigned 0. Considering both horizontal and vertical explorations, the developed point of view can be represented as two matrices, which can be represented as in Figure 6.9.

6.4. Sampling of interest area

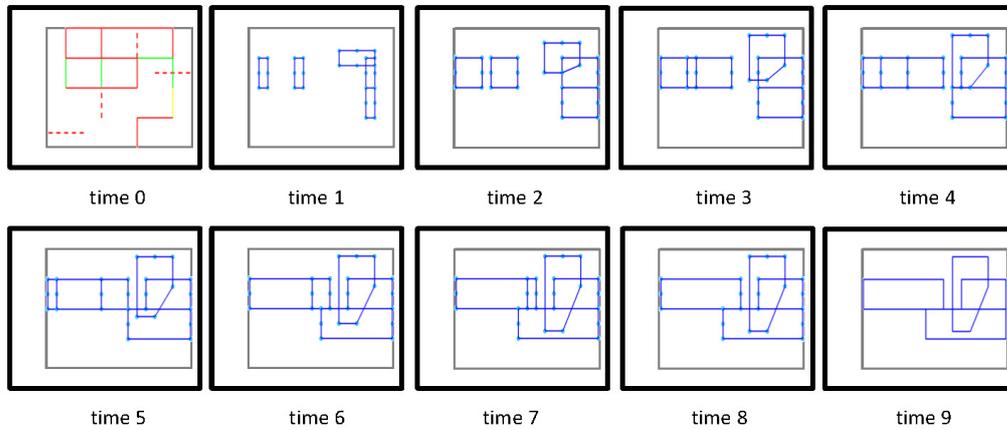


Figure 6.8: Evolution of a set of agents starting from a set of PADs

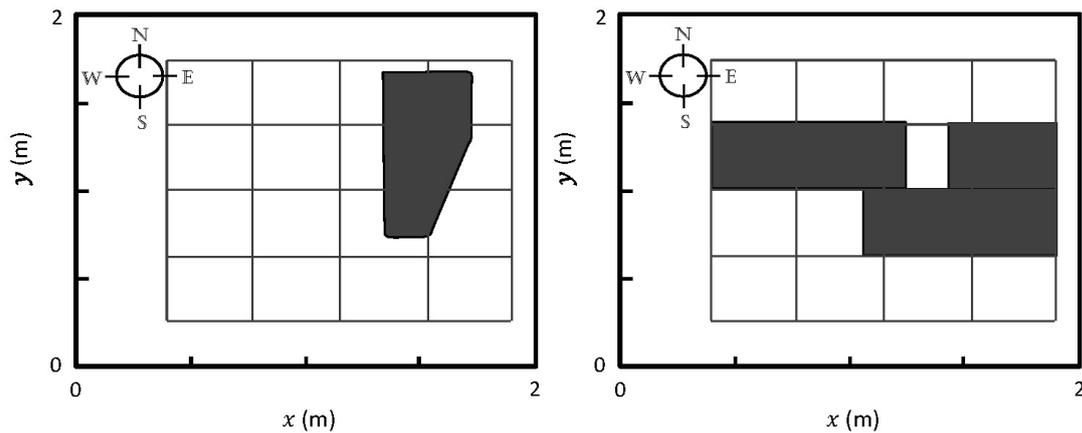


Figure 6.9: Vertical and horizontal point of view

6.4.4.2 Integration of points of view

The aggregation of all these points of views developed by a number of agents after the exploration previously defined generates what we call location probability maps. This aggregation process is performed in two phases: (1) accumulation, and (2) continuity resolution. As presented in Figure 6.3, this process is repeated for a number of epochs decided by the user. Thus, each of these points of view, in both directions, is accumulated, thereby generating a global view of the problem

and a map demarcating the likely location of the pipeline, with high reliability. An example of probability map is presented in Figure 6.10.

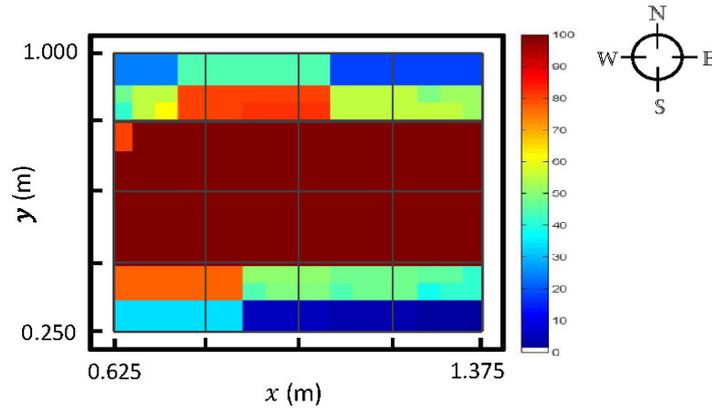


Figure 6.10: Example of probability map

This process determines potential hot spot locations of the pipe in the study area.

Let us finally add that if the study area has no segment with $aci =$ “yellow” or “green”, a number of “red” segments are selected to start the process. If after 20% of the total number of epochs selected by the user, the sum of viewpoints remains equal to zero in the accumulated point of view, both for vertical and horizontal directions, the algorithm stops and a decision is made about no existence of pipes in the interest area.

6.5 Image pre-processing

In the analysis and interpretation activity developed by the agents, the pipe existence plausibility in the analyzed image is evaluated. This activity is composed of two stages: (a) pre-processing and, (b) identification and training. The first stage is used to improve the image visualization and to reduce the amount of data to handle, thus facilitating identification and training. The second stage for this

6.5. Image pre-processing

activity is used to interpret the final image in the last stage, and establish the plausibility of this image belonging to the pipe layout location sought. However, although this activity of analysis and interpretation was selected for its generality for any type of pipe, it can be replaced totally or partially, and even reconsidered, if desired, to find pipes with more specific characteristics, such as the case of plastic pipes using, for example, the approach presented in Chapter 4.

Image pre-processing is paramount for both proposed approaches, since it allows to simplify the image while maintaining the reliability of the data in the resulting images. The pre-processing method used in this chapter was previously addressed in Chapter 5. The underlying rationale behind this stage consists in cleaning zones where the presence of a pipe is less likely in the obtained image. The cleaning of the zones in this stage considers that non-horizontal variations of the wave amplitude value correspond to ‘no-pipe existence’. To this purpose, first, the obtained profiles are transformed into $T14$ and $T15$ images (see, Chapter 3). Then, the Hough transform, a well-known segmentation technique, is used to detect and remove the horizontal lines from both images (see, Chapter 5). These images are then merged; thereby a new image is generated. For this new image the detection and cleaning of horizontal lines is applied with the aim to minimize the noise in the final image. The pre-processing used in this chapter starts with each profile image (PL or PAd) that must be analyzed (raw matrix, A) and consists of three phases: (a) binarization, (b) segmentation and cleaning, and (c) refining (see process details in Chapter 5, Section 5.2). Then the pre-processed matrix A' is obtained. This matrix is the input for the proposed process to obtain the indicator of plausibility pipe existence, aci (Section 6.4.2).

6.6. Analysis of the proposed system for detecting the piping layout in a case of an urban WSS

6.6 Analysis of the proposed system for detecting the piping layout in a case of an urban WSS

In this section we discuss the use of the proposed system in this chapter for automatic detection of paths of pipes. We use the two proposed approaches in a real case, which corresponds to an urban area for which no guidance on the pipe location was available (see Figure 6.11).



Figure 6.11: Location of the sampling area – Case study

The field work was carried out in an urban water supply system and tried to identify a pipeline, whose location was unknown and was subsequently confirmed by excavation. The road is asphalt concrete and the pipe has a nominal diameter of 200 mm cast iron at a depth of 1.0 m. The GPR equipment used and the capture parameters correspond with the same characteristics of the equipment used so far in previous chapters on laboratory tests. However, in this case, the antenna used had a center frequency of 400 MHz, a frequency lower than the previously used (1500 MHz). The equipment configuration used for the surveys in this case study corresponds to configuration 115 (see Appendix B).

6.6. Analysis of the proposed system for detecting the piping layout in a case of an urban WSS

6.6.1 Case study - Approach 1

In this section, the application of the proposed system for automatic mapping of the layout of pipes buried in a WSS is presented by the adaptive setting inspection routes (Approach 1) described in Section 6.2.1. In Figure 6.11 we use the red square to present the area bounded for this approach with $L = 2$ m (size 2×2 m). In Table 6.1 the results provided by the auto-regressive model for the initial profiles are presented.

Table 6.1: Auto-regressive model of order 4, initial capture profiles – Latin sampling

	ϕ_1	ϕ_2	ϕ_3	ϕ_3	a_t	aci
PL ₁	0.7732	-0.0730	0.0289	0.0673	91.82	red
PL ₂	0.6003	-0.2529	0.2731	0.2745	46.52	green
PL ₃	0.7444	-0.0967	0.0343	0.0329	90.78	red
PL ₄	0.6620	-0.2466	0.2967	0.0872	43.50	red

The paths constituting the sampled route have the auto-regressive model coefficients given in Table 6.2.

Table 6.2: Auto-regressive model of order 4, profiles of the guided trajectory

	ϕ_1	ϕ_2	ϕ_3	ϕ_3	a_t	aci
PAd ₁	0.6780	-0.2362	0.2766	0.1489	48.04	green
PAd ₂	0.6787	-0.1920	0.0747	-0.0324	117.70	yellow
PAd ₃	0.5507	-0.2000	0.2638	0.2474	78.84	yellow
PAd ₄	0.6985	-0.0936	0.0087	-0.0835	114.60	red
PAd ₅	0.7085	-0.2060	0.1060	-0.0340	144.40	green
PAd ₆	0.6600	-0.2083	0.2600	0.1588	43.90	green
PAd ₇	0.6419	-0.3246	0.1517	0.0262	155.30	green

6.6. Analysis of the proposed system for detecting the piping layout in a case of an urban WSS

The results obtained after applying the proposed approach are summarized and presented graphically in Figure 6.12.

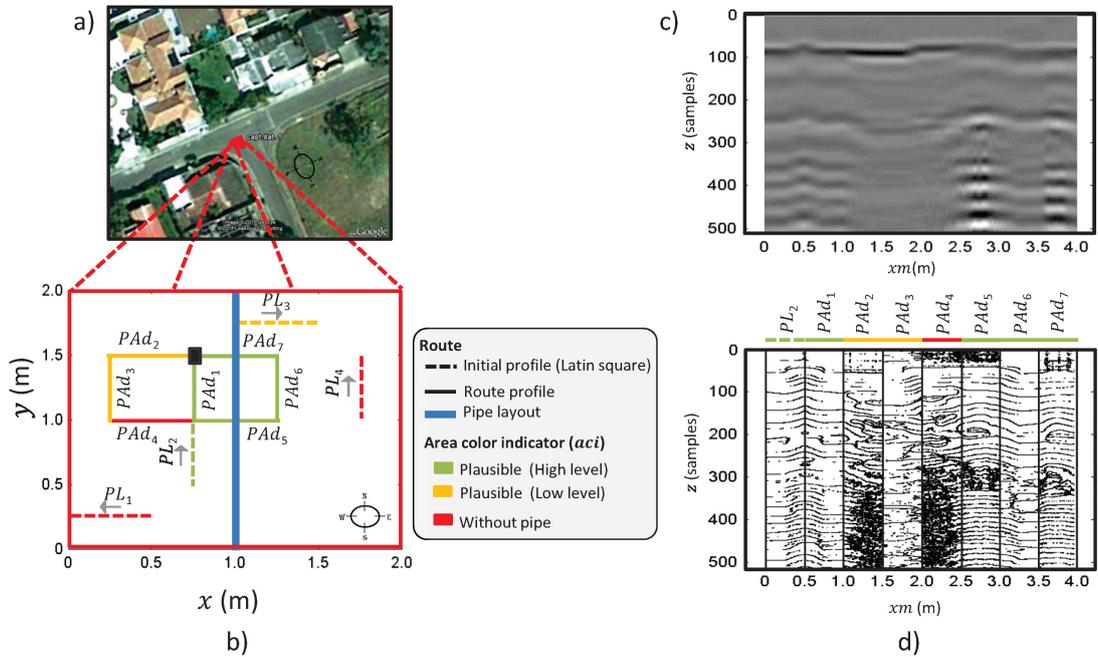


Figure 6.12: Pipe path location – Real case, Approach 1. (a) Interest area, (b) result of application, (c) raw profiles, and (d) pre-processed profiles

The process of generating the search path of the pipe (Figure 6.12) starts by choosing the path of the previous exploration by Latin squares, where more plausible evidence of the existence of the pipe can be found. For our case study (see Table 6.1), the Latin square that generates the path (PL₂), the segment with coordinates $Xm\{(0.75,0.50), (0.75,1.00)\}$, is selected. The other paths are marked with *aci* according to the conditions described in Section 6.4.3. Since the direction of Xm_0 is SN, Xm_1 will take the possible directions: {SN, WE, EW}. WE and EW are equiprobable, while continuing to SN would have less probability, taking into account that in the previous step, Xm_0 , pipe evidence has been found. This way, directions WE and EW (probabilities 0.4 and 0.4) have double probability than SN (probability 0.2). A random selection is performed and a new direction Xm_1 , again SN. The process is repeated taking into consideration the

6.6. Analysis of the proposed system for detecting the piping layout in a case of an urban WSS

recorded information from the previous paths regarding direction and likelihood of pipe existence. After running seven path sections, a trajectory, as in Figure 6.12 and detailed in Table 6.2, eventually identifies the buried pipe in the area under study. Information about the depth of each pipe section that is highly susceptible to contain a pipe is used for two purposes: just as a mere informative information, and also to confirm pipe slopes for different pipe section, thus helping confirm the existence of the pipe.

6.6.2 Case study - Approach 2

In this section, the application of the proposed system for mapping paths of WSS buried pipes by automatically generating probability maps (Approach 2) is presented. The case study correspond to an urban area in which the route of the pipeline (verified in the field) was known. The automatic system was raised in order to compare the results with the field observations. The scheme of the assay is presented in Figure 6.13,a.

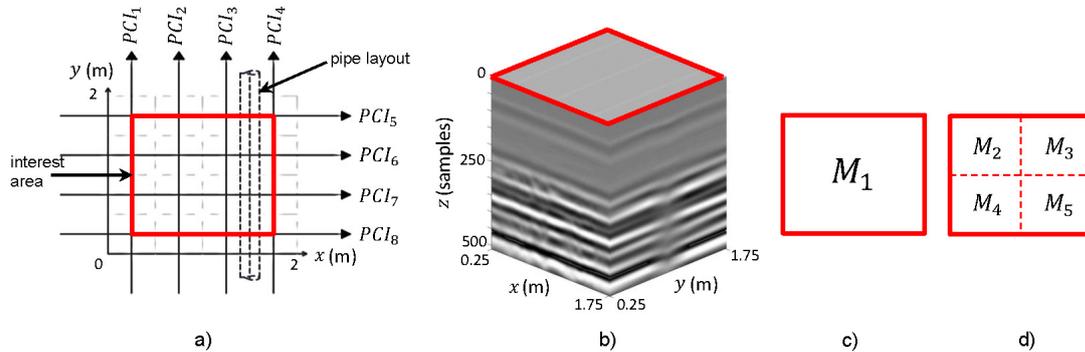


Figure 6.13: Schematic configuration, case study – Approach 2. (a) Scheme of the case study, (b) raw data from the interest area, (c) analysis set A, and (d) analysis set B.

Starting with eight profiles (PCI_1, \dots, PCI_8) in the study area, raw data integrates the bulk of data corresponding to the interest area (Figure 6.13,b). In this view, interpretation is deemed really difficult. So, the analysis strategy for

6.6. Analysis of the proposed system for detecting the piping layout in a case of an urban WSS

this section consists on the application of the methodology just described, through two configurations (set A and set B; Figure 6.13,c y d). Set A is obtained by the action of just one agent, while four agents produce set B. For set A, $L_{setA} = 1.5$ m for the interest area, while for set B, the study area is divided into four interest areas with $L_{setB} = 0.75$ m each. The value of m is 4 for both sets. This way the step size for the agents is $ss_{setA} = 0.375$ m and $ss_{setB} = 0.188$ m, for a total of 100 epochs. At this point we evaluate some possibilities or variants of the model, trying to improve the results and evaluating the repetitiveness of the model and the consistency of the results. So, we evaluate both sets three times each. The results of the proposed system obtained for the two sets is presented in Figures 6.14 and 6.15, showing vertical and horizontal maps of probabilities, respectively.

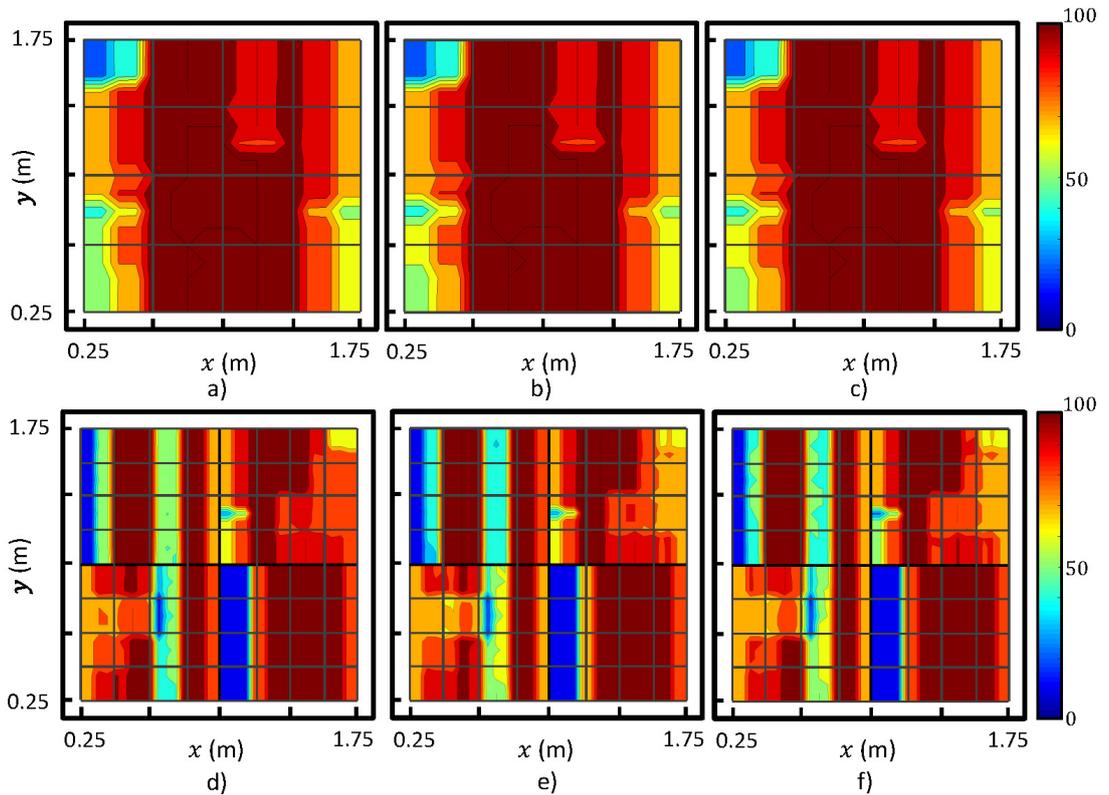


Figure 6.14: Comparison of vertical probability maps. (a – b) set A, and (d – f) set B.

6.6. Analysis of the proposed system for detecting the piping layout in a case of an urban WSS

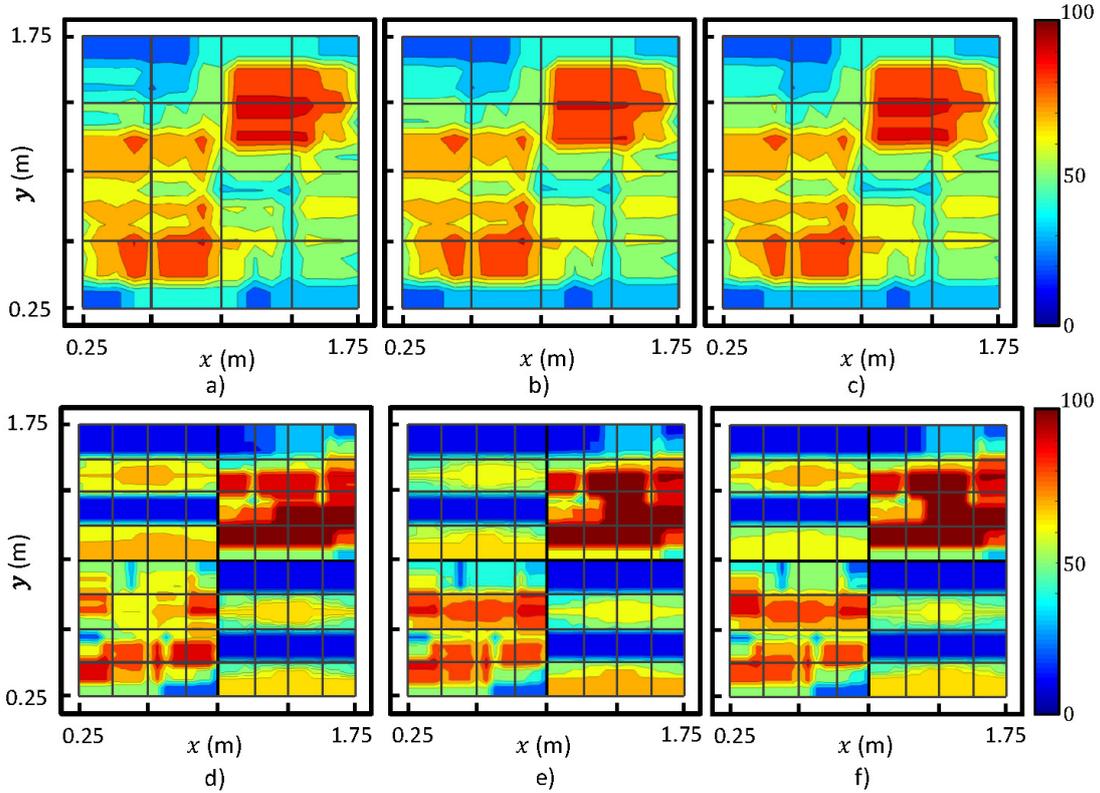


Figure 6.15: Comparison of horizontal probability maps. (a – b) set A, and (d – f) set B.

Figures 6.14 and 6.15 show the consistency of the results for the repetitions of each set. Also, the similarity of high values for both maps is observed in parts (a, b, c) and (d, e, f). It can be observed that a continuous area of bigger size can be seen in the vertical maps, that is to say, the longitudinal direction is clearly marked by the y -axis. Taking into account that the likelihood of pipe existence is very low for horizontal maps and that for repetitions of the same set the maps are very similar, we now evaluate the pipe existence for a probability value of 90% for vertical maps and select one of the three repetitions by set, by contrasting with the pre-processed profiles in Figure 6.16.

6.6. Analysis of the proposed system for detecting the piping layout in a case of an urban WSS

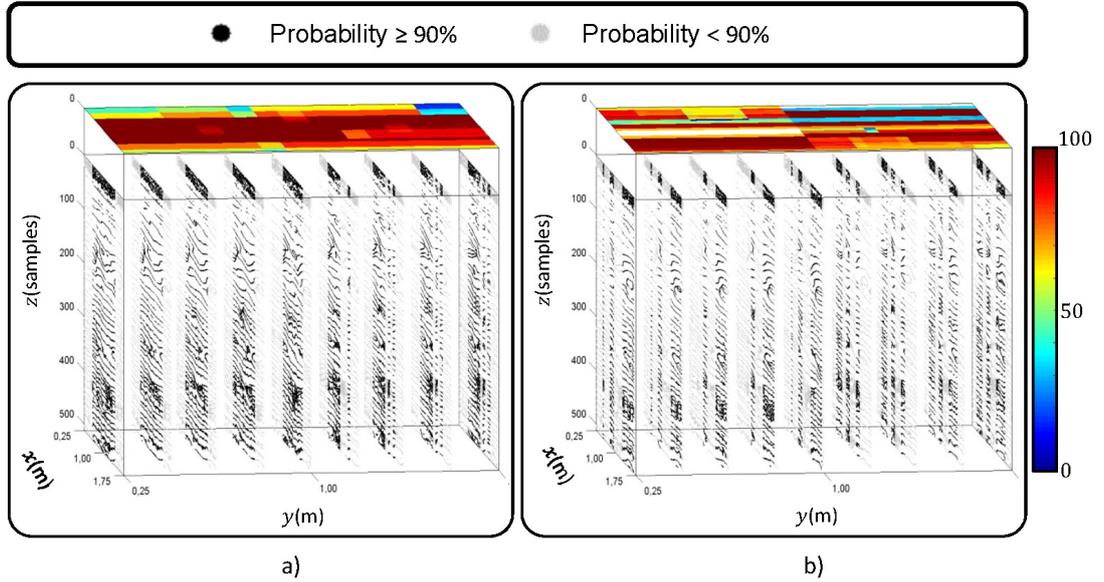


Figure 6.16: Combination of pre-processed profiles and probability area. (a) set A, and (b) set B.

It can be observed in Figure 6.16 that the pre-processing consolidates the data, thus allowing interpretation, when compared with data in Figure 6.13,b. Additionally, it is observed in both sets that the proposed method reduces the amount of data necessary to interpret the interest area. This is the result from the fact that 100% of data from horizontal profiles and over 50% of data from vertical profiles are of no interest to the agents. Taking advantage of the consolidation and reduction of data, and the enhancement of the interest data provided by the method in both sets, we now select the areas prone to contain the sought pipe. This interpretation, combined with elements of Figure 6.16, is presented in Figure 6.17.

6.6. Analysis of the proposed system for detecting the piping layout in a case of an urban WSS

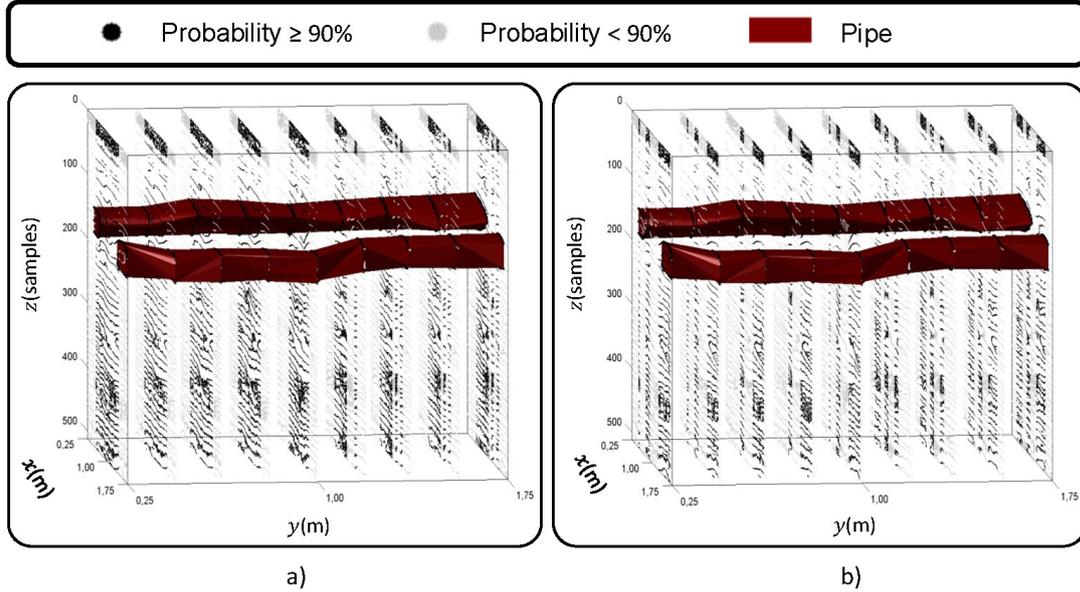


Figure 6.17: Merging Figure 6.16 with the interpretation of likely location of the pipe. (a) set A, and (b) set B.

Figure 6.17 presents the path of the pipe about we had *a priori* knowledge of existence in the interest area. However, a new surface appears that could be related to another object within the same area with 90% of probability of corresponding (partially or totally) to a pipe; this happens in both sets. To observe more clearly this finding, we present in Figure 6.18 the plan view of Figure 6.17.

By contrasting the plan view of Figure 6.17, given by Figure 6.18, with the location of the pipe of which we had previous knowledge, we can see that the work of the agents was not only correct but also broader than expected, since the study area contained not only one pipe, but a second one that was discovered by the agents.

In Figure 6.18 it can be observed that the area prone to be interpreted as a pipe (from the area information given by the points, marked in red), when compared with the field observation (for the known pipe, marked in blue), exhibits remarkable quality both in terms of location and diameter of the pipe.

6.6. Analysis of the proposed system for detecting the piping layout in a case of an urban WSS

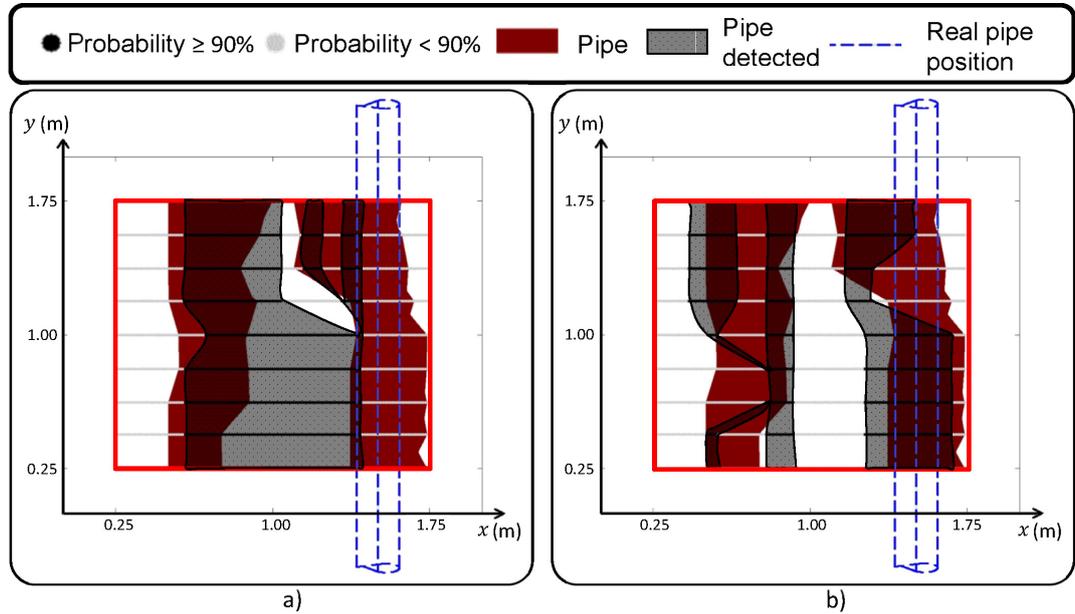


Figure 6.18: Plan view of Figure 6.17. (a) set A, and (b) set B

Moreover, we note that for this pipeline the path obtained with the probability maps (shown in black), presents a better approximation in set B, because the steps of the agents are smaller. This same fact can be observed overall for the two sets of Figure 6.18, where set A combines the two interpretations in one area, showing a single pipe that combines the two diameters and includes the area in between, while for set B, because of the smaller step size of the agents, the two parallel pipes are clearly observed.

6.7 Summary and comments

This chapter has evaluated the feasibility of generating a methodological tool to build the cadastre of the pipes of a WSS by applying the non-destructive technique GPR in an automatic way. The aim was that the proposed tool was easy to use and allow that operators with low experience in handling GPR could infer the layout of the pipes. In this regard, a system has been proposed in two approaches determined by profile capturing (Approach 1) and by profile analysis of captured profiles (Approach 2) to discover pipes in GPR images.

The main feature of Approach 1 lies in the feasibility of implementing an automatic detection system of piping layout in real time, provided that it is possible to set appropriate conditions and relationships between components, such as GPR equipment and capture system, that support the proposed detection system.

Approach 2's main feature is the feasibility of implementing an automatic detection system of piping layout from data previously obtained by mapping the probability that, besides showing the piping layout, enable us to better approach the area affected by the object in the recorded images. This improved view gives us an idea of the diameter of the material contained in the sampled space.

After having applied the proposed methodology in this chapter (for both approaches) in an urban area, it can be concluded that the approach facilitates the analysis and interpretation of the images obtained with GPR to locate paths of pipes in WSSs, as the methodology generates mapping routes that reduce the time of data capture, minimize the amount of data to analyze, and define the most relevant data. These results favor more complete analyses, since areas with little or low interest are discarded, making it feasible the study of WSS buried pipes by non-destructive methods by non highly skilled personnel in GPR image analysis.

6.7. Summary and comments

This chapter has presented the importance of the tools developed to help carry out dynamic management of water supply systems. Additionally, it has shown the important facts that such tools may be obtained from non-intrusive techniques, are easy to use, and are able to use real data. The results obtained with the system of automatic detection of pipes proposed under the two approaches are promising in the sense that the system presents interesting possibilities of adaptation to detect other features that may be of interest, such as water leaks. This prospective adaptation is based on the improved pre-processing of the data we have presented. However, it will require prior extraction of patterns to search for these features.

6.7. Summary and comments

Chapter 7

Water supply system component evaluation from GPR radargrams using a multi-agent approach

The work performed up to this point is as follows. In Chapter 3 we have evaluated the feasibility of improving the visualization of pipes of different materials used in WSSs in GPR images. Later, we have found that improving the visualization can promote automatic detection of objects (pipes) in complex cases, as is the case of plastics (Chapter 4). We have seen in Chapter 5 that, through the analysis using the Hough transform, the visualization of objects can be improved further by finding the peaks and valleys of the wave traces contained in the images. We have also observed that it is possible to automatically detect WSS piping layouts from these peaks and valleys (Chapter 6). This chapter attempts to improve image pre-processing by providing agility for obtaining data containing information on the objects captured by GPR images. This chapter presents a multi-agent approach to be used as a quick and easy tool for the interpretation and analysis of the characteristics of WSS components when working on a collection of GPR survey files. The general scope of this chapter, in the framework of the thesis is presented in Figure 7.1.

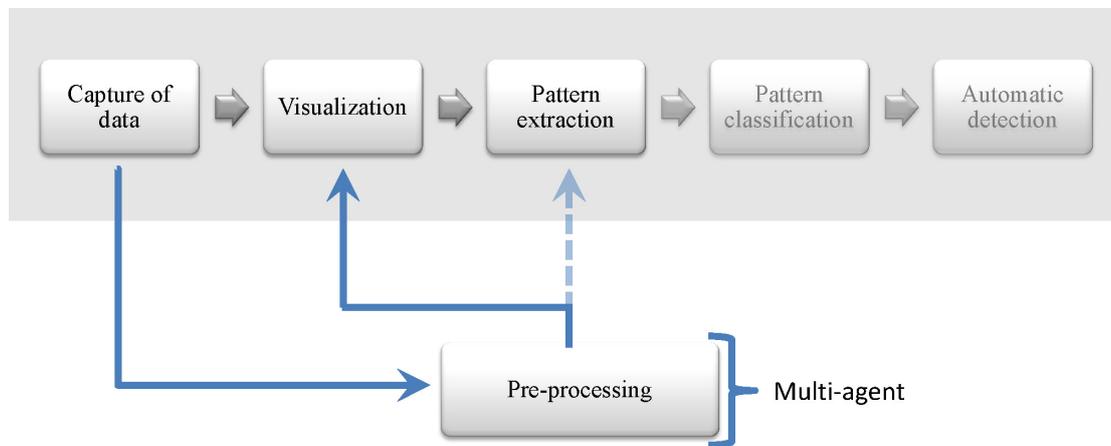


Figure 7.1: Chapter 7 scope

A first approximation to the content of this chapter was presented in the Congress *Mathematical Modelling in Engineering & Human Behaviour 2011*. It was also published as a short version in [Ayala-Cabrera *et al.* (2011c)]. Afterward, and improved approach was published as an extension in the high impact Journal *Mathematical and Computer Modelling* [Ayala-Cabrera *et al.* (2013b)].

7. WATER SUPPLY SYSTEM COMPONENT EVALUATION FROM GPR RADARGRAMS USING A MULTI-AGENT APPROACH

The importance of the approach presented in this chapter resides in that the improved processes herein described reduce the cleaning data time, and preserve data that provide valuable information on the objects in the GPR images, in relation to the transformation presented in Chapter 3 (matrices T_{14} and T_{15}), while the applications and approaches outlined in Chapters 5 and 6 may still be applied. So, this chapter, using a multi-agent approach, builds an easy and quick tool that improves the interpretation and further analysis of the characteristics of WSS components working on a collection of GPR survey files. The multi-agent algorithm proposed in this chapter has been developed in MatLab and is based on Game Theory. The input is the result of the survey GPR radargram and the output consists of the agents' scores in the game, a race. The columns of the output matrix, describing the agents' movements, together with the associated racing times can be interpreted in a meaningful way. As a result, this analysis enables easy determination of the electromagnetic properties of the prospected underground, and provides an accurate classification of underground properties. The results of this agent racing algorithm are promising, since it groups and, consequently, reduces the amount of points constituting the initial radargrams, while at the same time preserves the objects main properties, and promotes clearer views of pipes and better identification of the various components of WSSs.

The multi-agent paradigm is used in this chapter to evaluate components of WSSs from GPR radargrams. The aim of this chapter is to provide non-highly qualified technicians with non-destructive, easy-to-use and computationally efficient procedures for interpreting GPR survey files. These procedures enable technicians to gain insight into the layouts of the systems, and uncover various concealed characteristics of WSS components. Following the same line of research on GPR image processing discussed in Chapter 3, this chapter takes the matter further by presenting a new multi-agent algorithm.

This chapter is organized as follows. In this section we have given a general outline of the work developed in the chapter. In Section 7.1, we present the basis of the multi-agent methodology proposed. Then, a new section develops the proposed methodology. Some comments on the methodology implementation

and the corresponding pseudo-code are presented in Section 7.3. An experimental case already studied in previous chapters is revisited to exemplify the proposed methodology in Section 7.4. Finally, a section of conclusions closes the chapter.

7.1 Principles of the proposed pre-processing – Agents race

A radargram (GPR images) is the result of a GPR survey, a matrix A of size $m \times n$. The dimension m is the volume of signal data each trace records, which depends on the characteristics of the equipment used. The sample rate is an equipment parameter, coming generally in sets of 512, 1024, and 2048 samples/trace for commercial equipments. Typical analyses carried out in raw GPR images (location and identification of hyperbolas) are based on intensity differences demarcated by the wave amplitude generated in the images after the passage of the signal through various subsurface strata. The analysis we present in this chapter is based on time characteristics. In this analysis, the peaks and valleys (maxima and minima) of the waves generated are extracted. The trend of the path of each trace and the average value of peak-to-peak time are studied. The basic principle assumes that the field is homogeneous and, thus, there should be a clear correspondence between the various peaks obtained. However, it should be noted that although the material is homogeneous, in practice, the measured values are different, even though very close together. Consequently, very different values demarcate anomalies in the image. These peaks were first extracted and numbered according to their occurrence in the trace, by using the matrix $A'f$ presented in Chapter 5 (Section 5.1). They were then placed in their respective positions in distance, and finally, the last value was used to fill an array of size m (number of rows of matrix A – see Figure 7.2).

A fast calculation of this matrix (Figure 7.2,e) is given by the algorithm proposed in this chapter. In this algorithm, we obtain the above-mentioned values by means of a system based on multiple entities (agents) that search for trace peaks by simulating a race. A multi-agent system consists of a population of autonomous entities (agents) situated in a shared structured framework (environment) [Weyns & Holvoet (2005)] (see Chapter 4, Section 4.2). Agents operate independently (in our case) but they can also interact with their environment and coordinate with other agents [Herrera *et al.* (2010)]. However, even if the

7.1. Principles of the proposed pre-processing – Agents race

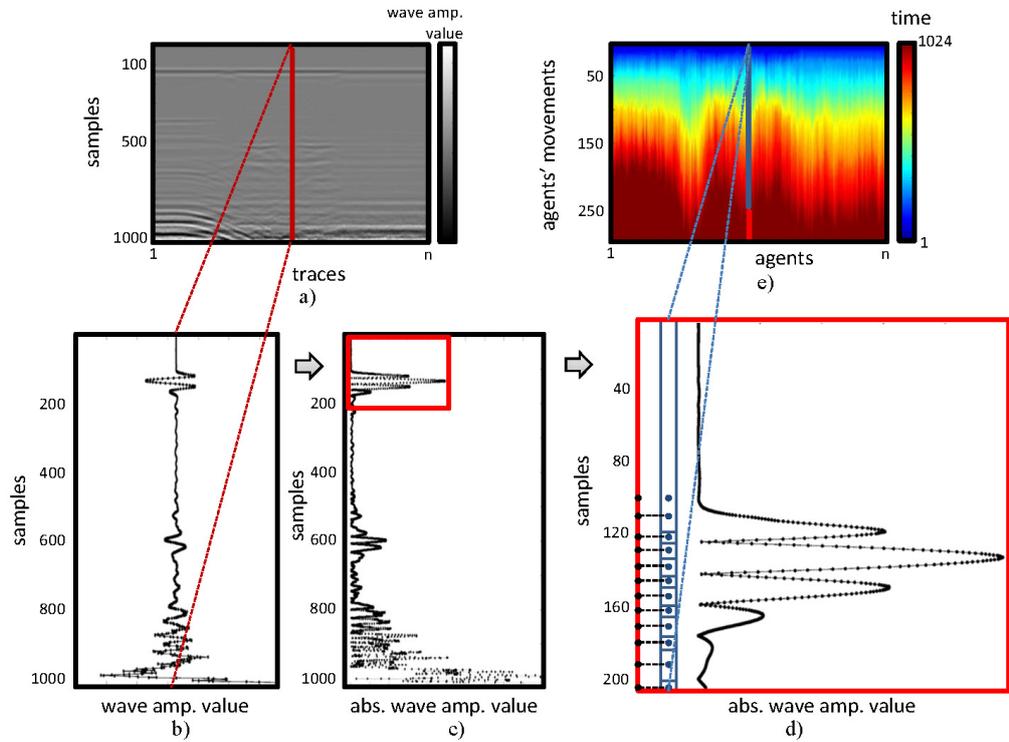


Figure 7.2: Principles of the GPR image pre-processing. Source: Ayala-Cabrera *et al.* (2013a)

agents are able to act and achieve their goals by themselves, it may be beneficial to (partially) cooperate for better performance, thereby forming coalitions. Turning to coordinating activities, either in a cooperative or a competitive environment, one basic way to solve the potential conflicts that may arise between agents is by means of negotiation. Negotiation may be seen as the process of identifying interactions based on communication and reasoning regarding the state and intentions of other agents. All the agent operations: coordination, cooperation, competition, and negotiation, are based on Game Theory [Neumann & Morgenstern (1947)], which origins the basic principles underlying the multi-agent systems [Leyton-Brown & Shoham (2008)].

7.1. Principles of the proposed pre-processing – Agents race

The reliability of finding buried objects is maintained even after reducing the amount of data (objective of pre-processing). This is because only a small rate of the agent population (proportional to the width in pixels of the image under study) will keep enough information for decision-making. A number of studies [Ayala-Cabrera *et al.* (2013c); Ocaña-Levario *et al.* (2013)] have shown that it is possible to reconstruct images of buried objects even by reducing the data by more than 50%. Note that the available data depends on the material in which the wave is propagated, together with the type of material of the buried objects.

The pre-processing of GPR images proposed in this chapter was termed as *Agent race*. This algorithm was proposed to quickly obtain matrices $T14$ and $T15$ (Chapter 3). The algorithm has been developed in MatLab, is based on game theory and uses the multi-agent paradigm. In this algorithm, we obtain the above-mentioned values by means (among others) of a system based on multiple entities (agents) that search for peaks by simulating a race.

In the racing game, agents compete through their interaction with the environment. As a result, each agent is rewarded by a movement at every iteration of the game. The number of game iterations or “racing time” should be tuned to preserve reliability in the approaches. Note that a low number of iterations allows anyone to win by chance. Nevertheless, if this number is large enough, the winners of the races will be those agents that have competed selecting better strategies. The “racing time” has been tuned by multiple laboratory tests (such as those presented herein) in which the buried object is known. From a visualization point of view, we can check that the final MAS (Multi-Agent System) based patterns, provided that a suitable racing time has been considered, offer images perfectly explaining the actual layouts.

The *Agent race* algorithm will provide an interpretation and a grouping method for data from GPR radargrams. In this pre-process we try to reduce the number of data integrating the initial radargram, while preserving its initial properties and all the most relevant data, so that its ability to identify buried objects through suitable visualizations is preserved. The multi-agent approach enables significant reduction of the time needed for the analysis.

7.1. Principles of the proposed pre-processing – Agents race

The input of the *Agent race* algorithm is the raw material obtained from the GPR survey radargram (see Figure 7.2,a). The signals received in GPR explorations are stored in a matrix, A (radargram), that is made up of m -vectors, b_k , $k = 1, \dots, n$, (traces), that represent the variation of the soil's electromagnetic properties in terms of depth. Let us represent this matrix by columns $A = [b_1, b_2, \dots, b_{n-1}, b_n]$. The length, m , of vectors b_k , corresponds to the volume of signal data recorded in each trace.

In the current chapter, agents run a race (an obstacle course in fact) focused on the study of the environment that these agents explore. The individual success of an agent depends just on the environmental characteristics and its ability to handle the posed difficulties. Thus, each agent is a single player playing against itself and “Nature”, in the sense that “Nature” [Szép & Forgó (1985)] continually poses various difficulties the agent has to cope with. The communication among agents consists in continuously observing the others and outlining the paths already explored by the agent for excluding them from other agents' strategies. The environment keeps track of this communication. Then, the race is an endurance test for the agents, with a prize consisting in advancing one position (movement) depending on the effort made. Efforts are based on wave amplitude values in each column of A (see Section 7.2.2). The movement of agents during the racing will be conditioned by the changing trends of the traces they travel through. The racing will end once time t has elapsed. The winner(s) are the agent(s) that, according to the race conditions, have performed a bigger amount of movements. The output of this competition among the agents is a matrix, R , with columns $m1$ -vectors X_s , $s = 1, \dots, n$, n being the number of participating agents. X_s represents the race time variation, which collects the various movements performed by agent s . The length $m1$ is the number of movements performed by the winner(s). As a non-winner agent, s , has managed to fill a lower number of coordinates in X_s , value m is used to complete its missing coordinates till the value $m1$ accomplished by the winner(s). This pre-processing work demarcates two spaces, the original space and the pre-processed space. The first one is characterized by the size of the matrix of raw data ($m \times n$). The second space, in its turn, is characterized by the size of R ($m1 \times n$).

7.2. Proposed pre-processing algorithm – Agents race

Game theory uses models to study formalized interactions between incentive structures (the called games) and carry out the decision processes. Thus, the optimal strategies, and the expected and observed behavior for the agents (players) are studied. The proposed game (in the current chapter) is a payoff function specified for each player. So, the game is a function $\pi : \prod_{s \in P} \Sigma^s \rightarrow \mathbb{R}^n$, [Shoham & Leyton-Brown (2009)], where P is the set of agents (players). It is a finite set, which we label $\{1, 2, \dots, n\}$. Each agent s in P has a finite number of strategies making up a strategy profile set, Σ . The n traces generated by the GPR survey (columns of matrix A) are used as pseudo-parallel tracks for the n agents to compete.

During the race, each agent s in P builds its vector of strategies, whose i -th coordinate is the strategy taken by the agent at time i . A strategy for a player is a function that maps sequences of states to a natural number, corresponding to a move available to the player at the end of the sequence [Lomuscio & Raimondi (2006)]. Here, the vector of strategies for an agent is determined by its respective column (vector b) in matrix A .

The agents' competition evolves in time from $i = 1$ till $i = m$. In the competition each agent s in P has four properties: (a) interpretation, (b) decision to move, (c) movement time, and (d) the race phases.

7.2 Proposed pre-processing algorithm – Agents race

In this section, we explain the *Agent race* algorithm as a multi-agent activity and describe aspects of the multi-agent behavior programming. Once, the *Agent race* principles have been stated (previous section), the schematic proposed algorithm is shown in Figure 7.3.

7.2. Proposed pre-processing algorithm – Agents race

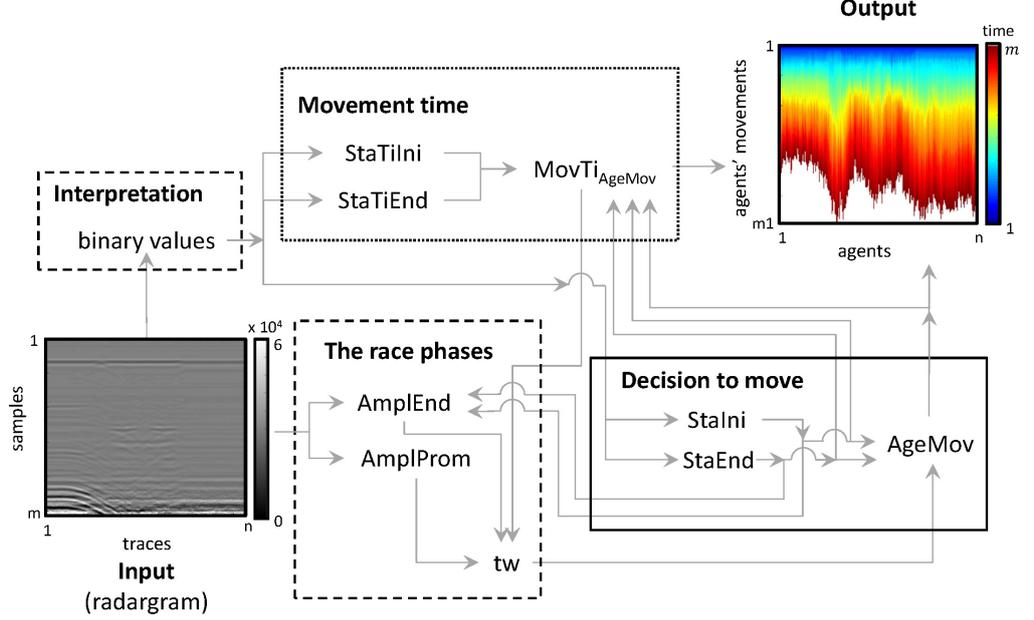


Figure 7.3: Schema of the proposed multi-agent algorithm. Source: Ayala-Cabrera *et al.* (2014a)

The four properties of agents are explained next.

7.2.1 Interpretation

For each time during the race, an agent takes one value of the trace, b_i ; and then this value is compared with two more signal values, the before-value b_{i-1} and the next-value b_{i+1} ; and a binary value ($bin_i \in \{0, 1\}$) is generated as a result.

$$bin_i = \begin{cases} 1, & \text{if } i = 1 \\ 1, & \text{if } b_{i-1} < b_i < b_{i+1} \vee b_{i-1} > b_i > b_{i+1} \\ 0, & \text{if } b_{i-1} > b_i < b_{i+1} \vee b_{i-1} < b_i > b_{i+1} \\ 0, & \text{if } i = m \end{cases} \quad (7.1)$$

- **Interpretation exceptions.** The exceptions for the interpretation property are related to the equalities between the current value and the contrast values (before and next values). Thus, one or both contrast values can equal

7.2. Proposed pre-processing algorithm – Agents race

the value of the current time. The equalities can be due to causes such as: (a) wave amplitude values being too small to be differentiated among themselves; (b) application of filters; (c) failure to emit antenna signal because of internal faults; (d) highly reflective soils; and (e) others.

The agent will look back and forward in search of times, for which the contrast values (before and next, respectively) are different to the current time value. These searches enable an agent to interpret the $bin = 0$ position. This value is compared with the current position and as a result the binary value for the current position is generated. The exception interpretation pseudo-code is shown in Table 7.1.

Table 7.1: Interpretation rule exceptions pseudo-code

```

... looking back
     $i1 = 1$ 
    while  $(b_{i-i1} = b_i \vee i1 \leq i - 2)$ ; do  $i1 = i1 + 1$ ; end
... looking forward
     $i2 = 1$ 
    while  $(b_i = b_{i+i2} \vee i2 \leq m - i - 1)$ ; do  $i2 = i2 + 1$ ; end
... searching the  $bin_{i3} = 0$  position
     $i3 = \text{round}((2 \cdot i - i1 + i2) / 2)$ 

```

where $i1$: is the number of before times the agent looks back to get a value different from the value for the current time.

$i2$: is the number of next times the agent looks forward to get a value different from the value for the current time.

$i3$: is the time when the agent assumed that the binary value is zero.

After time $i3$ is obtained, the interpretation rule for exceptions is determined by Equation 7.2.

$$bin_i = \begin{cases} 0, & \text{if } i3 = i \\ 1, & \text{otherwise} \end{cases} \quad (7.2)$$

7.2.2 Decision to move

An agent's decision to move is based on the binary value variation. According to this variation, a property called stamina varies positively (variable $StaIni$, Equation 7.3) or negatively (variable $StaEnd$, Equation 7.4).

$$StaIni = \begin{cases} 1, & \text{if } i = 1 \\ StaIni + 1, & \text{if } bin_{i-1} = 0 \wedge bin_i = 1 \\ StaIni, & \text{otherwise ,} \end{cases} \quad (7.3)$$

$$StaEnd = \begin{cases} 0, & \text{if } i = 1 \\ StaEnd + 1, & \text{if } bin_{i-1} = 1 \wedge bin_i = 0 \\ StaEnd, & \text{otherwise .} \end{cases} \quad (7.4)$$

When the total stamina is zero, that is $StaIni$ equals $StaEnd$, the agent receives its payoff for the effort performed. This is accomplished by the variable $AgeMov$. As explained in Section 7.2.4, this is applied during the “official” race, just after the *warming-up*.

$$AgeMov = \begin{cases} 0, & \text{if } i = 1 \\ AgeMov + 1, & \text{if } StaIni = StaEnd \wedge tw \neq 0 \\ AgeMov, & \text{otherwise .} \end{cases} \quad (7.5)$$

7.2.3 Movement time

Each effort developed by an agent happens between a start time and an end time. These values, associated to the agent movement ($AgeMov$) are stored in two agent personal vectors, namely, $StaTiIni$ (Equation 7.6) and $StaTiEnd$ (Equation 7.7), respectively.

$$StaTiIni_{AgeMov+1} = \begin{cases} 1, & \text{if } i = 1 \\ i, & \text{if } bin_{i-1} = 0 \wedge bin_i = 1 \end{cases} , \quad (7.6)$$

7.2. Proposed pre-processing algorithm – Agents race

$$StaTiEnd_{AgeMov} = \begin{cases} 1, & \text{if } i = 1 \\ i, & \text{if } bin_{i-1} = 1 \wedge bin_i = 0 \end{cases} . \quad (7.7)$$

Also, every agent movement (*AgeMov*), has one movement time associated *MovTi* that we define as the average time between the stamina's time start (*StaTiIni*), and the stamina's time end (*StaTiEnd*). A component of *MovTi* is defined every time the difference between these stamina values is 0:

$$MovTi_{AgeMov} = \frac{StaTiIni_{AgeMov} + StaTiEnd_{AgeMov}}{2} . \quad (7.8)$$

7.2.4 The race phases

On the surface, the first signals that are recorded by the GPR come from direct transmission between the transmitting and receiving antenna through the air at light speed (air wave). This is followed by the direct signal through the floor or along the surface, the surface wave. These signals form a pair of black and white lines on top of all GPR images. These lines cover any reflection at an equivalent depth of half the antenna separation [Robinson *et al.* (2013)]. These opening lines divide the space (in depth), such that data above the lines are considered not susceptible to analysis and data below them are the data of interest for the analysis. Consequently, in this section we have used these two zones and divided the race into two phases (a) warming-up, and (b) racing, corresponding to the area of little or no interest, and the area of interest in the analysis, respectively. The phases are characterized by two times: a warming-up time (*tw*, Equation 7.11), and a racing time (*tr*), totaling a time $t = tw + tr$, where the *tw* time corresponds to the time for the agent to overcome the end wave amplitude value (*AmplEnd*, Equation 7.9) in some percentage of the average wave amplitude value for before values for the current time (*AmplProm*, Equation 7.10).

7.3. Implementation issues and recommendations

$$AmplEnd = \begin{cases} 1, & \text{if } i = 1 \\ b_i, & \text{if } StaIni = StaEnd \\ AmplEnd, & \text{otherwise} \end{cases}, \quad (7.9)$$

$$AmplProm = \begin{cases} b_i, & \text{if } i = 1 \\ \sum_{j=1}^{i-1} \frac{b_j}{(i-1)}, & \text{otherwise} \end{cases}, \quad (7.10)$$

$$tw = \begin{cases} 0, & \text{if } i = 1 \\ MovTi_{AgeMov}, & \text{if } |AmplEnd| > x \cdot |AmplProm| \wedge tw = 0 \\ 0, & \text{otherwise} \\ tw, & \text{if } tw \neq 0 \end{cases}, \quad (7.11)$$

where $x = 1.1$, this being an experimental value.

7.3 Implementation issues and recommendations

The algorithm, *Agent race*, proposed in this chapter has been implemented in the Windows environment, using MatLab for matrix manipulations. The pseudo-code corresponding to the agent race is developed in Table 7.2.

In the proposed method (Section 7.2), the raw traces can be used. However, we recommend data interpolations so that the use of interpretation exceptions (Section 7.2.1) is minimized. In addition, in this work we use the absolute wave amplitude values, which improves the final visualization, because the agents' stamina amount increases during the competition, as it happened in the trace standardization process developed in Chapter 5 (Section 5.2.1.1).

With the interpolation, the clipped wave parts have been corrected. We also use interpolation to obtain a finer data discretization. Thus, we take the trace value from the original amount to a constant value (4096 samples/trace), which enables comparison between radargrams with different rates of capture (samples per trace).

7.3. Implementation issues and recommendations

Table 7.2: The payoff function for each agent – Agent race pseudo-code

```

1  ... variables definition for first race time
2    bin1 = 1; StaIni = 1; StaEnd = 0; AgeMov = 0;
3    StaTiIniAgeMov+1 = 1; StaTiEndAgeMov = 1;
4    AmplEnd = 1; AmplProm = b1; tw = 0; x = 1;
5  ... payoff iterative function for each agent
6    For i = 2 : m;
7      ... interpretation property
8        if (bi is monotone); bini = 1;
9        elseif (bi has reached an extreme); bini = 0;
10       elseif (i = m); bini = 0;
11       else; interpretation rules exceptions are applied (Table 7.1);
12       end;
13     ... staminas and their respective times
14       if (bini-1 < bini);
15         StaIni = StaIni + 1; StaTiIniAgeMov+1 = i;
16       end;
17       if (bini-1 > bini);
18         StaEnd = StaEnd + 1; StaTiEndAgeMov = i;
19       end;
20     ... decision to move and movement time properties
21       if (StaIni = StaEnd);
22         if (tw ≠ 0); AgeMov = AgeMov + 1; end;
23         MovTiAgeMov = (StaTiIniAgeMov + StaTiEndAgeMov) / 2;
24         AmplEnd = bi;
25       end;
26     ... the race phases property
27       AmplProm =  $\sum_{j=1}^{i-1} \frac{b_j}{(i-1)}$ ;
28       if (|AmplEnd| > x · |AmplProm| ∧ tw = 0);
29         tw = MovTiAgeMov;
30       end;
31     end;

```

7.4 Experimental study – *Agent race*

This section provides the implementation of the proposed method of WSS component evaluation from GPR radargrams using a multi-agent approach, as described in Section 7.2. This section first evaluates the implementation in a particular case study under laboratory conditions. The viability of application of the method is also assessed. Then, various cases (all of them laboratory centered) help complete the evaluation of the effectivity of the method as far as data reduction concerns.

7.4.1 Case study

The case study corresponds to GPR images taken from a plastic pipe commonly used in WSSs. The pipe material tested was PVC with a diameter of 0.10 m. The GPR image was obtained by burying the pipe in dry soil in the test tank. The following task consists in pre-processing the captured GPR radargram using the proposed method. The equipment parameters used in this captures correspond to the assay configuration number 70 (see Appendix B). In Figure 7.4, some competition times are shown.

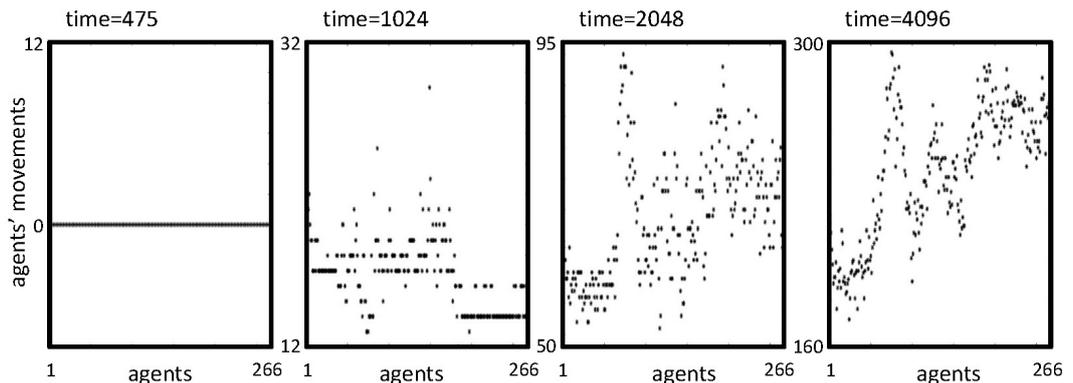


Figure 7.4: Some competition times. Source: Ayala-Cabrera *et al.* (2013b)

7.4. Experimental study – Agent race

In Figure 7.4, we can observe the temporal agent activity. Thus, for the analyzed radargram the warming-up phase is not finished until time 475, and the first movement for the racing phase takes place at time 476. Similarly, the grouping of agents in areas is observed after the competition is finished, and this corresponds to the proposed test configuration. For a better interpretation, in Figure 7.5 we represented the end of the race with the schematic configuration test proposed.

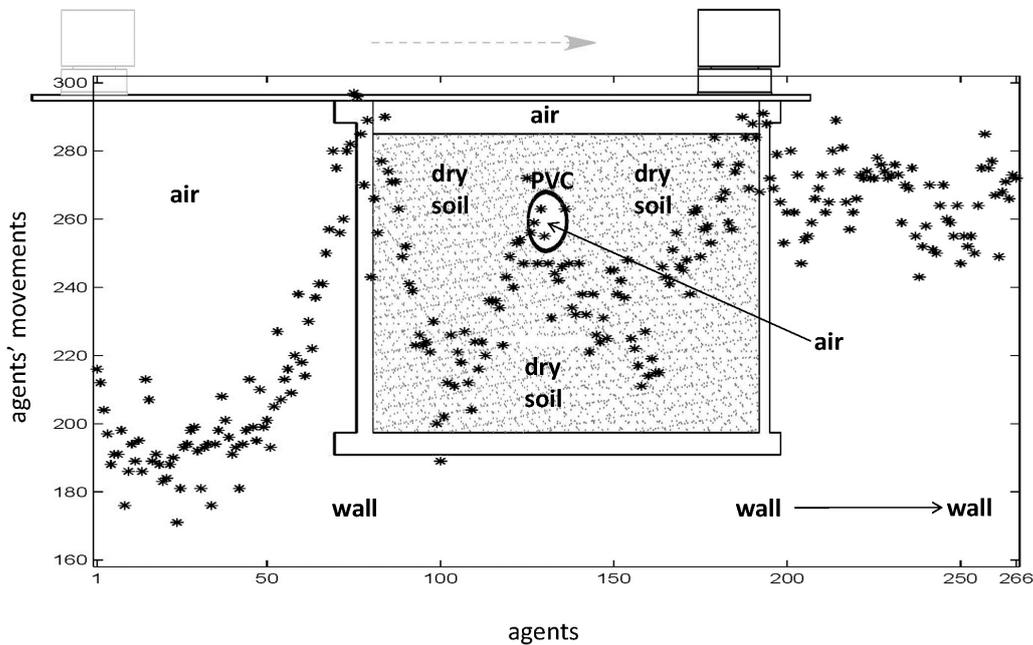


Figure 7.5: Last time race vs schematic configuration test. Source: Ayala-Cabrera *et al.* (2013b)

In the last time step of the race, the agent movements indicate, for the proposed test configuration, different velocity areas (Figure 7.5). In addition, the marked areas correspond to the soil velocities for the test: the materials for the tested soil being air, wall, dry soil, mixed area, dry soil and wall (from left to right). Moreover, the mixed area is the accumulation of velocities, since the air, dry soil, PVC, air, PVC and dry soil compose the mixed area (from top downwards). The movement times ($MovTi$) for each agent are rendered graphically and the result is shown in Figure 7.6,c. The input for the race (radargram) and

the schematic test configuration are shown in Figures 7.6,a and d, respectively.

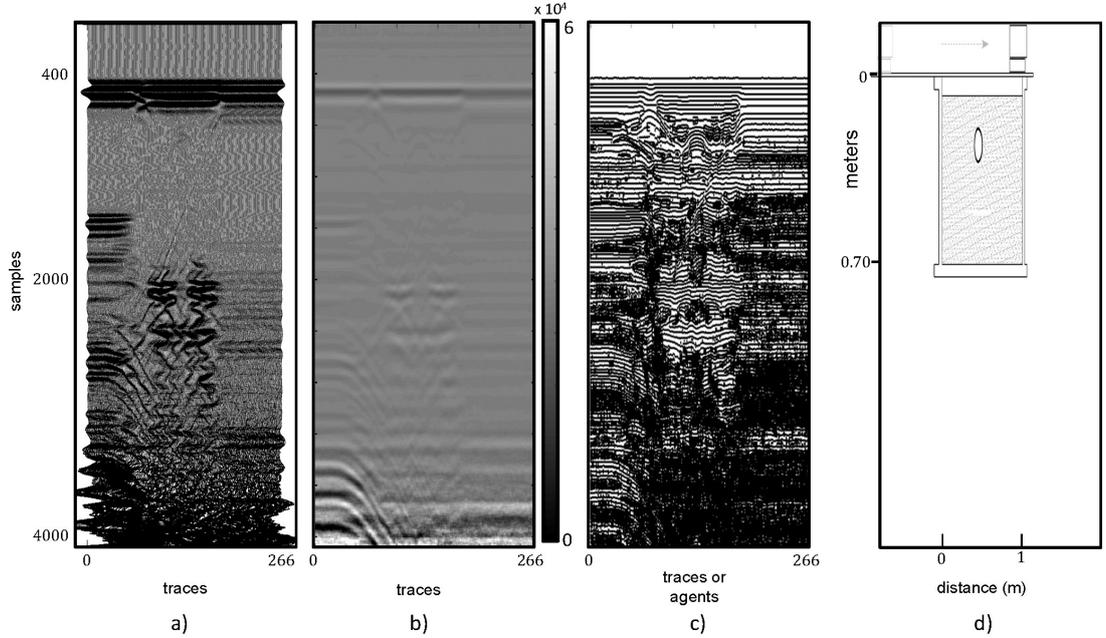


Figure 7.6: Pre-processing results. (a) Radargram, (b) GPR image (c) projection into original space of agent movements; final image, and (d) schematic configuration for test. Modified: Ayala-Cabrera *et al.* (2013b)

Each image in Figure 7.6,c presents a smaller number of points than the corresponding image in Figure 7.6,a thus enabling easier interpretation. It can also be observed, when comparing Figure 7.6,c and d, that obtained points with similar configurations also produce similar images. As a result, we can demonstrate that the application of the proposed multi-agent method improves insight into the subsoil properties.

7.4.2 Data reduction percentages in GPR matrix data

As already seen when implementing the method proposed in this chapter, in the case study, the *Agent race* algorithm minimizes the amount of data to be analyzed, while preserving the characteristics of the objects (if any) in the images. Then, we have used different configurations of trials (001-114, see Appendix B), with

7.4. Experimental study – *Agent race*

different objects (pipes of different materials) and without objects (tank with dry soil without pipes) to determine the effectiveness in data reduction and the ability of the method to preserve object features. The results are presented in Figure 7.7. It is worth to mention here that the case study presented in Section 7.4.1 has been highlighted with a green dot in the figure (test 24; 1024 samples). The corresponding results for this configuration, for 512 and 2048 samples, are given in test #24 with their respective curves.

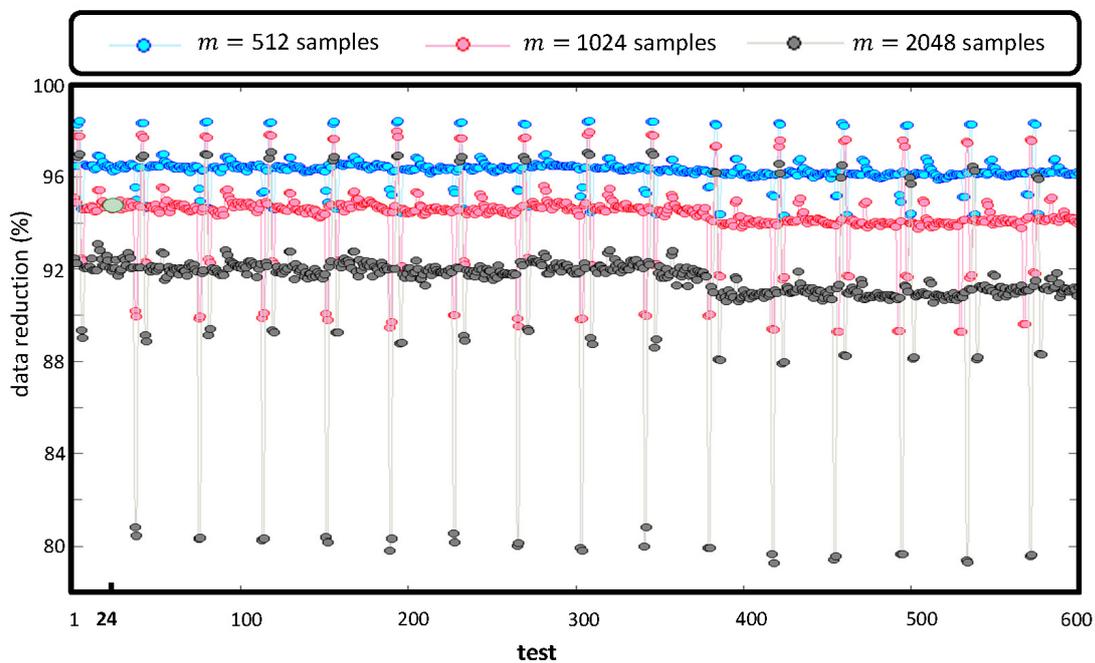


Figure 7.7: Data reduction percentages through the multi-agent pre-process developed

After having analyzed a total of 1800 images captured under different characteristics and configurations (see Figure 7.7), we can say that data reduction is higher than 78% in all the cases. This data can relate to the densities obtained in Chapter 5 in the analysis of density (Section 5.3). This relationship can be expressed by the expression $\{\text{density} + \text{data reduction} = 100\}$. By contrasting the data for matrices Ub and Db in this chapter with the obtained data we can see that the percent reduction, greater than 75% in both cases, is consistent; as a result, the similarity of the processes is shown.

Returning to the analysis of this chapter, it can further be seen in Figure 7.7 the similarity between data of different configurations for the same samples. For the different samples we have on average a percentage of data reduction of: 96% for 512-samples, 94% for 1024-samples and 91% for 2048-samples. Recall that the data has been taken at a constant rate of 4096 samples / trace and yet significant differences were obtained depending on the number of samples captured, because the greater the number of samples taken, the greater the amount of detail that is captured in the images. Data reduction is representative, since data reduction and preservation of the characteristics of the objects contained in the images of GPR favors more detailed and comprehensive analysis which can be of interest in WSSs, such as inspections of WSS sections to locate pipes, water leaks and other features or events.

7.5 Summary and comments

In this chapter we propose a tool for WSS component evaluation from GPR radargrams using a multi-agent approach. In the raw captured radargram without pre-processing, we can see how the weakly reflective plastic pipe materials (PVC) are difficult to identify. The transformation of the raw data based on the proposed multi-agent method improves the visualization of plastic pipe images by producing better representation of the signal characteristics. This procedure reduces the amount of points that comprises the radargrams, and provides clear information for further intelligent processes. The points can be visualized in binary scale of colors (white and black), and thereby the subjectivity in the choice for color scale is eliminated.

Finally, it should be noted that this is a simple process that does not depend on specialist skills (thus being a non-subjective process) and is repeatable. The proposed multi-agent method is efficient regarding computational resources (even in the complicated case of plastic pipes). The amount of information dealt with has been reduced, while reliability is preserved. Moreover, the proposed method

7.5. Summary and comments

offers the possibility of more detailed analyses in terms of time with the movements of agents, and this creates the possibility of better interpretations that could serve as a basis for intelligent training systems. This approach would help give WSS managers a more accurate vision of the systems they operate and, as a result, offer better service to users.

Chapter 8

Hierarchical agglomerative clustering analysis of GPR data

In this chapter, a combination of the multi-agent paradigm, which uses the pre-processing algorithm (*Agent race*) presented in Chapter 7, and a very well known clustering technique, namely Hierarchical agglomerative clustering, is used for unsupervised classification of subsoil characteristics working on a collection of GPR survey files. The main objective is to assess the feasibility of a new way of extracting features and patterns from radargrams. By optimizing both the field work and the interpretation of the raw images our target is to obtain visualizations that are automatic, fast, and reliable so to suitably assess the characteristics of the prospected areas and extract relevant information. The scope of this chapter, in the framework of the thesis, is shown in Figure 8.1.

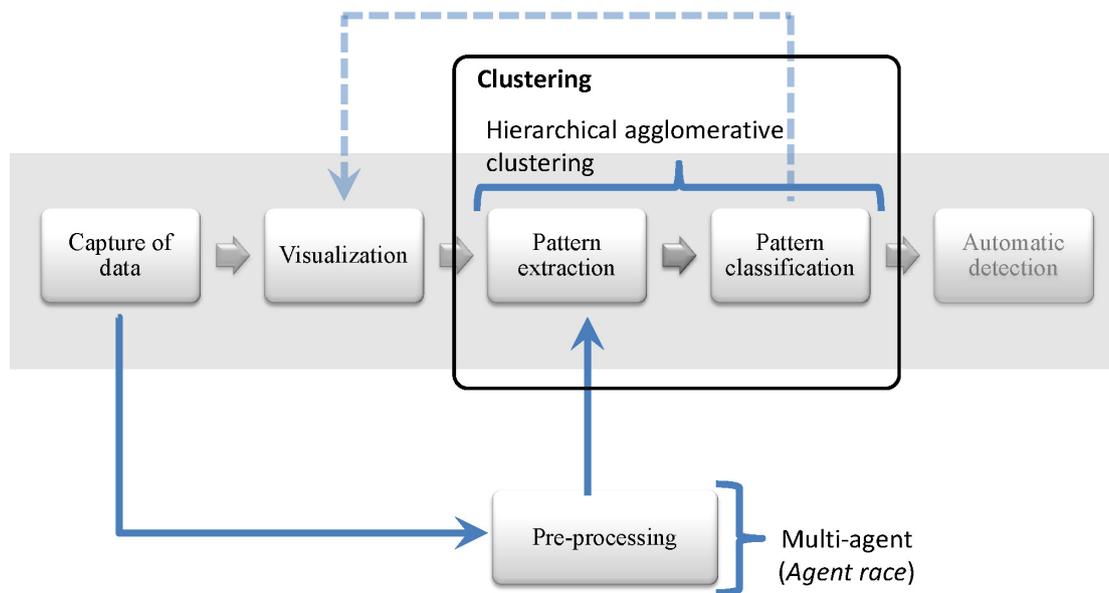


Figure 8.1: Chapter 8 scope

A first contribution with materials from this chapter was presented in the Congress *VII SELASI* [Ayala-Cabrera *et al.* (2011e)]. Afterward, a substantial improvement was published as an extension in the high impact Journal *Digital Signal Processing* [Ayala-Cabrera *et al.* (2014a)]. The importance of the approach developed in this chapter within the thesis hinges on the fact that the proposed system shows the viability of arranging GPR data from survey files into clusters, thus reducing the amount of information to be dealt with, while preserving its

8. HIERARCHICAL AGGLOMERATIVE CLUSTERING ANALYSIS OF GPR DATA

reliability. This chapter aims at generating a tool for analysis and simplification of GPR databases that could help decision-making in WSS management. The specific objective is to obtain a reduced number of clusters capturing the more relevant subsoil characteristics. The main idea behind the process boils down to gathering into clusters objects or anomalies within the inspected area, in a natural and fast way. As a result, the searching spectrum is enlarged and interpretation may be achieved in a fast way without requiring high levels of skill and experience. In addition, the proposed system does not require any assumptions about subsoil parameters. A powerful tool to analyze underground components in water supply systems is thus generated that acts in a non-destructive way and supports decision-making in water supply management. The implementation of the processes described in this chapter has been performed in MatLab in its entirety.

This chapter is organized as follows. This section has presented an outline of the work developed for the chapter. Section 8.1 presents the motivation and the principles of the proposal, which uses Hierarchical agglomerative clustering of pre-processed GPR images. Then, Section 8.2 presents the architecture of the proposal. Section 8.4 presents an experimental layout made out of a number of arrangements where the system has been tested. Also, a number of sensitivity analyses on various candidate metrics and linkage procedures has helped in the process of fine-tuning the proposed system architecture by selecting the most suitable combination of metrics and linkage procedures. Finally, a closing section includes the conclusions.

8.1 Motivation and principles of the proposed clustering method – Hierarchical agglomerative clustering

GPR offers a non-destructive way of exploring the immediate subsurface to detect buried objects such as pipes, cables, ducts and drains. Such buried objects cause reflections in the form of hyperbolas in the images (radargrams) that are perceived by GPR signals. Originally, these images are interpreted manually by human experts in costly and time consuming processes. Speeding up these processes for automating the interpretation is highly desirable [Janning *et al.* (2014)]. To this purpose, reducing data in the arrays of images, for better understanding and analysis is of paramount importance. Currently, it is well known that the decomposition of data in the image traces provides efficient signal representation. Also, discriminating characteristics are further provided for pattern classification [Shao *et al.* (2013)]. In addition, methodologies based on the analysis and attribute extraction of the so-called twin-peaks of the traces enable size estimation and combination of attributes of the object in the time-domain form. These reductions of the original matrix give way to more advanced processing such as clustering, allowing to cope with the amount of noise that frequently does not allow proper identification of buried objects using GPR images. Various initiatives using clustering methods have been used recently. In works such as Janning *et al.* (2014), clustering methods based on detection of hyperbolas to estimate the corresponding position of objects in the images of GPR are used. In Uchikado *et al.* (2012), clusters in which the reflected waves can be identified in the pipeline are generated. However, the main drawback for this initiative is the variability of results with the trials. To cope with this difficulty, an additional step to improve the accuracy is proposed. Firstly, we consider that those methods are based particularly in the selection and search of hyperbolas in the GPR images. Secondly, it has also been observed in other studies, such as Ayala-Cabrera *et al.* (2013c) and Ocaña-Levario *et al.* (2013), that GPR images, with appropriate pre-processing, enable us to reveal other features, along with the hyperbolas, to allow for better

8.1. Motivation and principles of the proposed clustering method – Hierarchical agglomerative clustering

comprehension of buried objects. In these works, objects and anomalies such as pipes and water leaks are identified; also the structure that contains the objects is delimited. The results of these studies are based on the pre-processing of images proposed in Chapter 7. This motivates the use proposed in this chapter of new pre-processing tools that enable to synthesize relevant data and to reliably identify buried objects. The pre-processing tools used in this chapter are based on multi-agent systems, which in turn are based on the Game Theory [Neumann & Morgenstern (1947)]. Additionally, it is intended that the information analyzed (regarding buried objects) facilitates the work of more powerful classification techniques. In order to encourage this type of new techniques, this chapter proposes using cluster analysis for classification purposes.

Cluster analysis is the organization of a set of patterns (represented as vectors of measurements or, in general, points in a multidimensional space) into groups based on some similarity measure. Patterns within a given group are more similar to each other than they are to a pattern belonging to a different group. In the unsupervised case, the problem is to group a given collection of unlabeled patterns into meaningful groups. In a sense, the labels are associated with the groups, but the labels are based on categorical data, i.e., the labels are obtained by inspecting the data [Abbas (2008)]. To classify data, hierarchical clustering has several advantages over other procedures. First, it is a fully unsupervised method. In the case of networks, this allows to cluster all units without having to specify a priori the number of clusters. In addition, the generation of a hierarchical tree provides not only partitions of the network (either based on how units are grouped in agglomerative clustering, or how units are divided into groups, in the case of divisive clustering), but also allows to visualize how the basic, first-order clusters are combined into higher-level groups [Alcodea & Marín (2010)]. Thus, in this chapter we have proposed the use of Hierarchical clustering for classification of buried objects through GPR surveys, in order to take advantage of performing unsupervised classification. The key idea is to promote intensive searches of more specific objects (pipes, for example), through more robust techniques.

8.2 Proposed system architecture

The architecture of the proposed system (Figure 8.2) may be split into three interrelated processes: I – pre-processing, II – hierarchical agglomerative clustering, and III – information retrieval and visualization. The first process uses a methodology denoted *Agent race* (Chapter 7). It is a multi-agent process to develop a pre-analysis of the signals in GPR survey files. This technique builds two spaces, named warming-up and racing, using the agents' behavior into the world where the agents evolve. The output of this process is used as input for the clustering process. This process embodies an unsupervised technique to cluster the survey data in a natural way that, at the same time, allows easy and reliable interpretation. The methodology used in this chapter is the so-called *Hierarchical Agglomerative Clustering* (HAC). To implement HAC we have evaluated some of the most common procedures and have chosen the one that best represents the various clusters showing the soil variability. Finally, the last process retrieves the data constituting the developed clusters and places everything back into the original space. The visualization herein obtained is the final objective since the sought results are easily obtained from the new images.

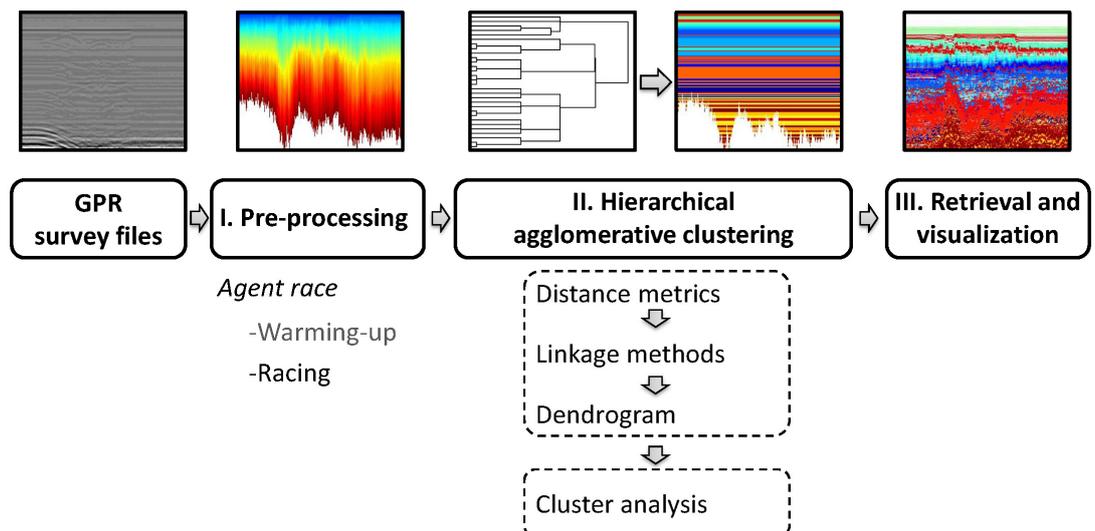


Figure 8.2: Architecture of the proposed system. Source: Ayala-Cabrera *et al.* (2014a)

8.3. Hierarchical Agglomerative Clustering

This architecture is run repeatedly by varying the metrics and the linkage methods used in process II. In this chapter, process III is applied to all the cases described in Section 8.4. The obtained images through these runs are crucial to determine the most suitable procedures to use in process II. These processes and a discussion about the performed selection are presented in detail below.

8.3 Hierarchical Agglomerative Clustering

Hierarchical agglomerative clustering is a frequently used clustering technique for making groups in various databases. In our case, it attempts to group signal traces into small clusters and to join clusters into higher-level systems [Ward (1963)].

The deterministic nature of HAC is strictly based on local decisions and a single-step analysis (not iterative procedures) [Hastie *et al.* (2001)]. Nevertheless, it can represent an advantage for our analysis because of its speed and its possible inclusion into an automatic process for a real-time process (for instance, a fast technique as HAC is crucial to include clustering into a multi-agent process). Furthermore, the hierarchical tree is easily viewed as a dendrogram and can be considered in a post-processing phase for explaining the way in which the clusters have been established and obtain a number of conclusions through it. The usual approach of k -means [Hartigan & Wong (1979)] has been rejected because, as the number of clustering centers is increased, initial placement of the centers becomes critical. Thus, for analyses that involve large numbers of clustering centers, it makes sense to use more structured techniques such as the HAC proposed.

The resulting hierarchical tree is a nested set of partitions represented by a tree diagram or dendrogram. Sectioning a tree at a particular level produces a partition into disjoint groups. If two groups are chosen from different partitions (the results of partitioning at different levels) then either the groups are disjoint or one group wholly contains the other, [GhasemiGol *et al.* (2010)]. This hierarchical clustering process can be visualized in a dendrogram form, where each

8.3. Hierarchical Agglomerative Clustering

step in the clustering process is illustrated by a join in the tree (see Figure 8.2 – process II). This process is divided into two parts: (1) dendrogram construction, and (2) cluster analysis. In the first phase, a dendrogram is built according to two aspects, which we want to evaluate: (a) distance metrics, and (b) linkage methods. These issues are related through the *cophenetic* correlation, which gives the goodness of the classification by comparing distances among input data and distances among output data. The second phase trims the dendrogram according to a criterion previously determined. The criterion used in this chapter pursues a trimming point such the data are divided in a natural way. To this purpose, the so-called *inconsistency* coefficient is used. We explain now the elements involved in these two phases.

8.3.1 Distance metrics

We use three distance metrics in our system, which are common in agglomerative hierarchical clustering. Given an $m1 \times n$ data matrix X , which is treated as $m1$ ($1 \times n$) row vectors x_1, x_2, \dots, x_{m1} , the different distances between vectors x_s and x_t are defined as follows:

1. *Euclidean distance*:

$$D_{Eu}^2(x_s, x_t) = (x_s - x_t)(x_s - x_t)' . \quad (8.1)$$

2. *Seuclidean distance*. Each coordinate difference between rows in X is scaled by dividing by the corresponding element of the standard deviation:

$$D_{Seu}^2(x_s, x_t) = (x_s - x_t)V^{-1}(x_s - x_t)' , \quad (8.2)$$

where V is the $n \times n$ diagonal matrix whose j -th diagonal element is $S(j)^2$, where S is the vector of standard deviations.

8.3. Hierarchical Agglomerative Clustering

3. *Cosine distance*. One minus the cosine of the angle between points (treated as vectors):

$$D_{Co}(x_s, x_t) = 1 - \frac{x_s x_t'}{\sqrt{(x_s x_s') (x_t x_t')}}. \quad (8.3)$$

8.3.2 Linkage methods

Once the proximity between objects in the data set has been computed, we can determine how objects in the data set should be grouped into clusters, using the linkage methods. The linkage methods take the distance information generated by metric measures and link pairs of objects that are close together into binary clusters (clusters made up of two objects). The linkage methods then link these newly formed clusters to each other and to other objects to create bigger clusters until all the objects in the original data set are linked together in a hierarchical tree. There are many possible choices in updating the similarity values. Among them, the most common linkage methods are (1) single, (2) average, and (3) complete [Lampropoulos & Tsihrintzis (2004)]. The following notation is used to describe the linkages used by these methods:

r and s are two clusters.

n_r and n_s are the number of object in clusters r and s , respectively.

x_{ri} is the i -th object in cluster r .

x_{sj} is the j -th object in cluster s .

The linkage methods we use here are the following:

1. *Single linkage*: also called *nearest neighbor clustering*; it is based on the minimum distance between clusters:

8.3. Hierarchical Agglomerative Clustering

$$Z(r, s) = \min (\text{dist} (x_{ri}, x_{sj})), \quad (8.4)$$

where $i \in (1, \dots, n_r), j \in (1, \dots, n_s)$.

2. *Average linkage*: also called *unweighted average distance* (UPGMA); it is calculated as the average distance between members of a pair of clusters. Average linkage tends to join clusters with small variances, and it is slightly biased towards producing clusters with the same variance:

$$Z(r, s) = \frac{1}{n_r n_s} \sum_{i=1}^{n_r} \sum_{j=1}^{n_s} \text{dist}(x_{ri}, x_{sj}). \quad (8.5)$$

3. *Complete linkage*: also called *furthest neighbor clustering*; its distance is based on the points in all the clusters that are farthest apart:

$$Z(r, s) = \max (\text{dist} (x_{ri}, x_{sj})), \quad (8.6)$$

where $i \in (1, \dots, n_r), j \in (1, \dots, n_s)$.

8.3.3 Cophenetic correlation

After linking the objects in a data set into a hierarchical cluster tree, we might want to verify that the distances (that is, heights) in the tree reflect the original distances accurately. The cophenetic correlation for a cluster tree is defined as the linear correlation coefficient between the cophenetic distances obtained from the tree, and the original distances (or dissimilarities) used to construct the tree. Thus, it is a measure of how faithfully the tree represents the dissimilarities among observations [Carr *et al.* (1999); Wedel & Bijmolt (2000)]. The output value, c , is the cophenetic correlation coefficient. The magnitude of this value should be very close to 1 for a high-quality solution. This measure can be used to compare alternative cluster solutions obtained using different algorithms. The cophenetic correlation is defined by

8.3. Hierarchical Agglomerative Clustering

$$c = \frac{\sum_{i < j} (Y_{ij} - y)(Z_{ij} - z)}{\sqrt{\sum_{i < j} (Y_{ij} - y)^2 \sum_{i < j} (Z_{ij} - z)^2}}, \quad (8.7)$$

where

Y_{ij} is the distance between objects i and j given by Equations (8.1), (8.2) or (8.3);

Z_{ij} is the cophenetic distance between objects i and j given by Equations (8.4), (8.5) or (8.6);

y and z are the average of Y_{ij} and Z_{ij} , respectively.

8.3.4 Cluster analysis

The hierarchical cluster tree may naturally divide the data into distinct, well-separated clusters. This can be particularly evident in a dendrogram, a diagram created from data where groups of objects are densely packed in certain areas and not in others. The inconsistency coefficient of the links in the cluster tree can identify these divisions where the similarities between objects change abruptly. We can use this value to determine where the cluster function creates cluster boundaries. In this way we define a label for each link in our dendrogram. This label shows to what extent two clusters are similar. With this measure, we can join clusters if the inconsistency value is less than certain specific threshold. The inconsistency coefficient characterizes each link in a cluster tree by comparing its length with the average length of other links at the same level of the dendrogram. The higher the value of this coefficient, the less similar the clusters connected by the link. For each link, k , the inconsistency coefficient is calculated as:

8.3. Hierarchical Agglomerative Clustering

$$IC(k) = \frac{L(k) - W1(k)}{W2(k)}, \quad (8.8)$$

where

IC and L are $(m - 1) \times 1$ vectors.

IC : Inconsistency coefficient.

$W1$: Mean of the lengths of all the links included in the calculation.

$W2$: Standard deviation of all the links included in the calculation.

L : Vector of the lengths of the links.

For leaf nodes, nodes that have no further nodes under them, the inconsistency coefficient is set to 0. The threshold used in this chapter is 1.

8.4 Experimental Study – HAC

In this section we evaluate the sensitivity of the proposed system to subsoil variations and determine the most suitable combinations of procedures. We have prepared a laboratory layout with 16 types of configurations. In a schematic way, these configurations are presented in Figure 8.3.

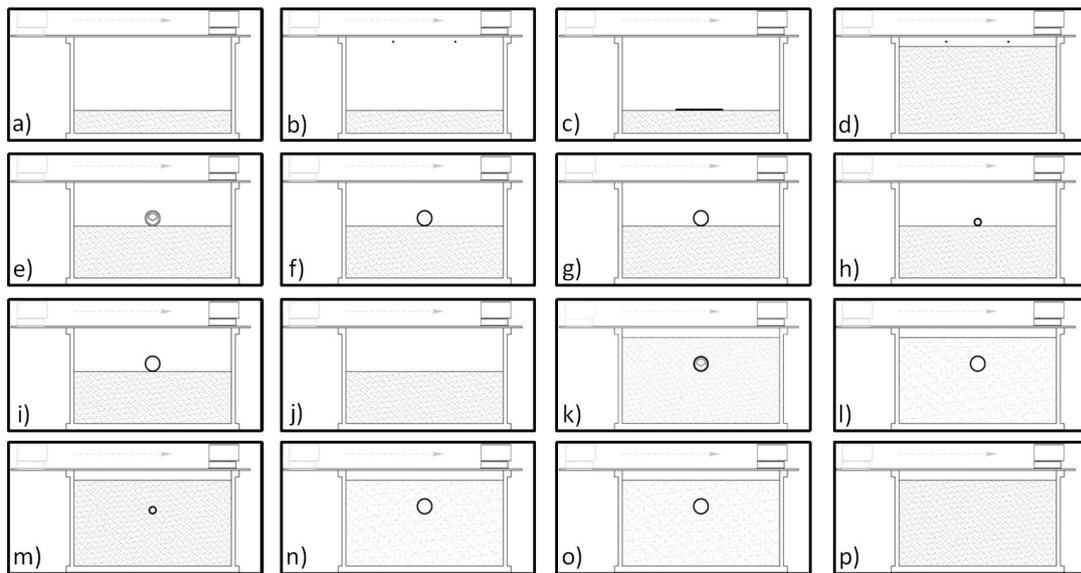


Figure 8.3: Layouts of the prepared configurations. Source: Ayala-Cabrera *et al.* (2014a)

In Figure 8.3, configurations from e) to i) and from k) to o) correspond to the soil configurations j) and p), respectively. Those configurations differ from j) and p), respectively, due to the addition to the layout of one of the four different pipe materials commonly used in WSS. A description of the characteristics of the inserted pipes and the configuration of the ensemble is presented in Table 8.1. All the arrangements in Figure 8.3 were tested with the GPR. In order to observe the responses of the various configurations, the same array of parameters was used in the equipment for all the surveys. The GPR equipment used in each exploration corresponds with a monostatic commercial antenna, with a control unit GSSI SIR 3000 and an antenna central frequency of 1.5 GHz, able to capture

Table 8.1: Characteristics of the buried pipes used for testing

Material	Index	Index Figures 8.3, 8.4, 8.6	Index Figure 8.7	Diam. (mm)
PVC	PVC	i), o)	e), k)	100
Poliethylene	PE1	h), m)	d), i)	35
Poliethylene	PE2	g), n)	c), j)	76
Asbestos cement	Fib	e), k)	a), g)	80
Cast iron	Fund	f), l)	b), h)	86

16 bits/trace, 60 traces/second, 1024 samples/trace and 20 η s/1024 samples. The parameters used by the equipment for these captures correspond to the assay configuration 70 (see Appendix B). Image capturing consisted in sliding the antenna on a flat surface, under the established configurations. We also have to mention that, even though the captured samples amount to 1024 samples/trace, the images in this chapter only show till sample number 600.

In Figure 8.4 we present the obtained raw results of the performed surveys by using a gray scale.

In the images corresponding to raw data shown in Figure 8.4, it can be easily observed that for PVC, PE1, and PE2 materials borderlines are not clear-cut and, as a result, poor visualization properties are obtained. This is attributed to the very low permittivity exhibited by these materials. This causes a low color intensity and, consequently demarcation of these materials is very difficult, making them almost invisible. Even being more easily observed, configurations for the other materials (Fib and Fund) present some degree of difficulty. Clearly, these images in Figure 8.4, which have undergone no processing at all, are really hard to understand and interpret, and even to discriminate. This fact underlines the need of methodologies enabling suitable interpretation.

8.4. Experimental Study – HAC

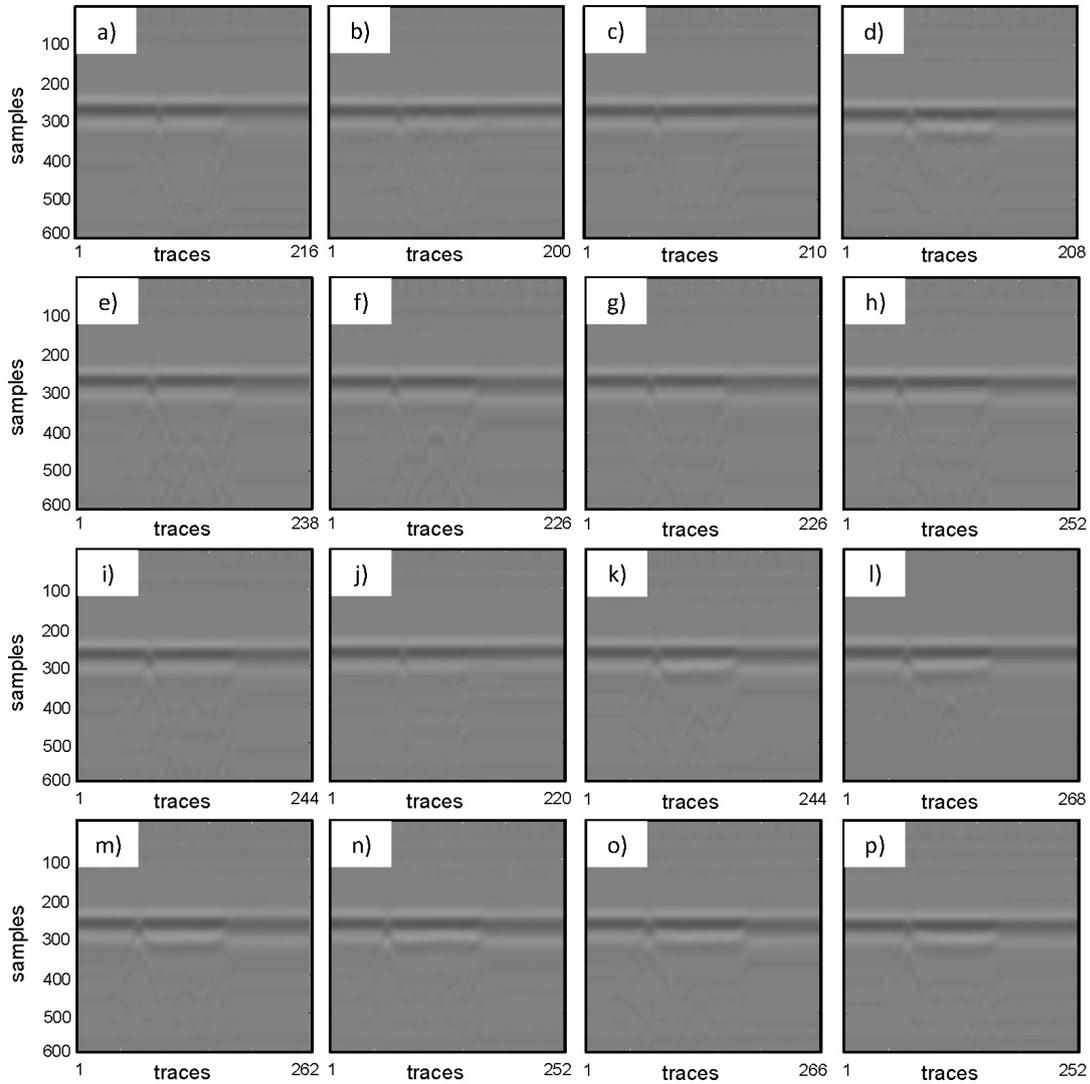


Figure 8.4: GPR survey files for the subsoil configurations in Figure 8.3. Source: Ayala-Cabrera *et al.* (2014a)

Images in Figure 8.4 have been treated by the system described in the previous section. We have used three different metrics and another three linkage methods. Table 8.2 presents the nine combinations among these metrics and linkage methods. The entries of the matrix symbolize the different combinations, and are used below.

Table 8.2: Combinations of metrics and linkage methods used within HAC

	Single	Average	Complete
Euclidean	SEu	AEu	CEu
Seuclidean	SSEu	ASEu	CSEu
Cosine	SCo	ACo	CCo

After evaluating these alternatives within the proposed system for the arrangements defined above, the *cophenet* correlation coefficient for each configuration was obtained. These coefficients are shown in Figure 8.5.

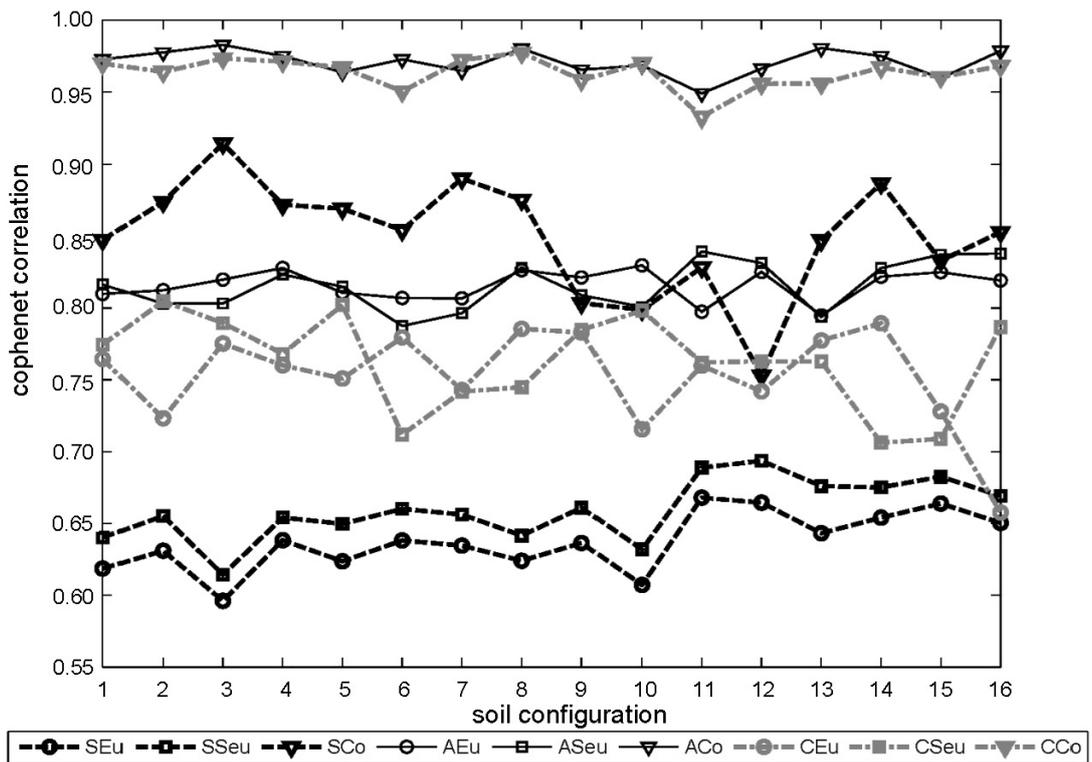


Figure 8.5: Obtained *cophenet* correlation coefficients. Source: Ayala-Cabrera *et al.* (2014a)

8.4. Experimental Study – HAC

It can be easily observed that the use of *cosine* increases the *cophenet* correlation among the proposed subsoil configurations. We can also observe that the combinations ACo and CCo present the highest values. Finally, it is ACo the combination that provides the highest and most constant *cophenet* correlation coefficients. In view of these results, we have adopted the combination of *cosine*, as a

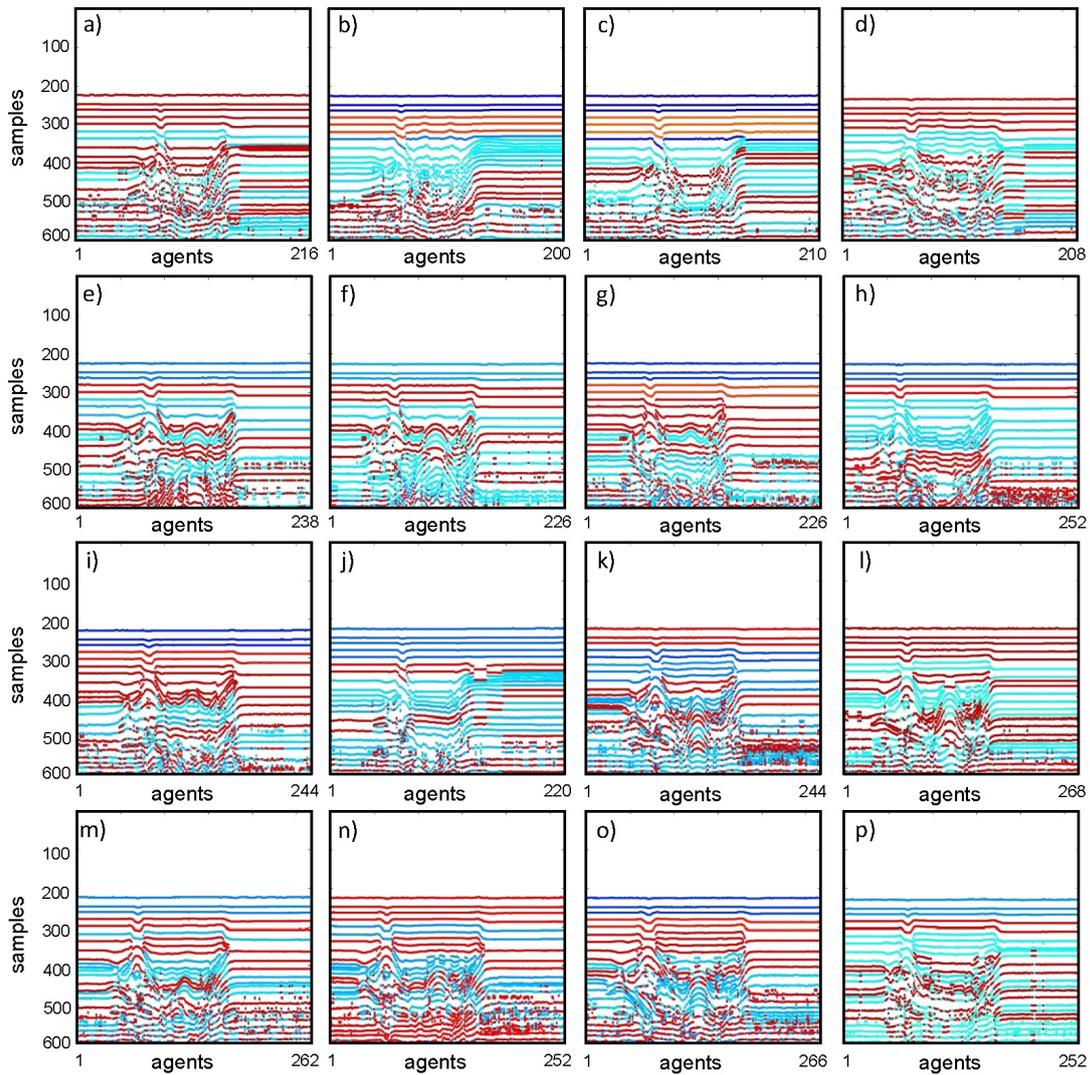


Figure 8.6: Clusters obtained for the configurations considered in Figure 8.3 placed back in the original space. Source: Ayala-Cabrera *et al.* (2014a)

distance metric, and *Average Linkage*, as the most suitable alternative for GPR images to obtain clusters through HAC. In Figure 8.6 we show the clusters (taken back to the original space) obtained by the use of ACo.

Images shown in Figure 8.6 present a lower amount of points than their corresponding images in Figure 8.4, thus enabling easier interpretation. It can also be observed when comparing Figures 8.3 and 8.6, that layouts with more similar configurations also produce more similar images. As a result, we claim that the application of HAC with the right combination of procedures improve the insight into the subsoil properties since these techniques manage to cluster objects in a more natural and more identifiable way. The developed clusters can be seen as layers that can be added or removed at one's convenience. For example, in Figure 8.7 we show insets e) to p) of Figure 8.6 selecting the layers where the pipe is located. Insets j) and p) are included (with all their layers) with the purpose of reference since they embody the soil configuration (without pipe).

Thus, in Figure 8.7 the location of the various added pipes can be clearly observed when comparing with the pipe-less configurations. Specifically, in Figure 8.7 comparisons between parts a) to e) and parts g) to k) against their corresponding pipe-less parts, f) and l), respectively, are performed. In these images PVC and PE pipes are much better visualized, since solid and well defined hyperbolae can be easily observed. This is the result of a clear reduction of the inherent noise. For example, the asymptotes of the hyperbolae are clearly observed now. After pre-processing with the proposed methodology the various radargrams corresponding to various types of pipes, it can be stated the following, regarding the pipe materials.

- **PVC pipes.** A clear improvement of their visualization and an increase of the number of observed characteristics are obtained when compared with the raw data.
- **PE1 and PE2.** This material is clearly visualized and pipe demarcation is clearly improved with regard to raw data. It is worth to mention that for this material, even pipes of very small diameter can be clearly observed.

8.4. Experimental Study – HAC

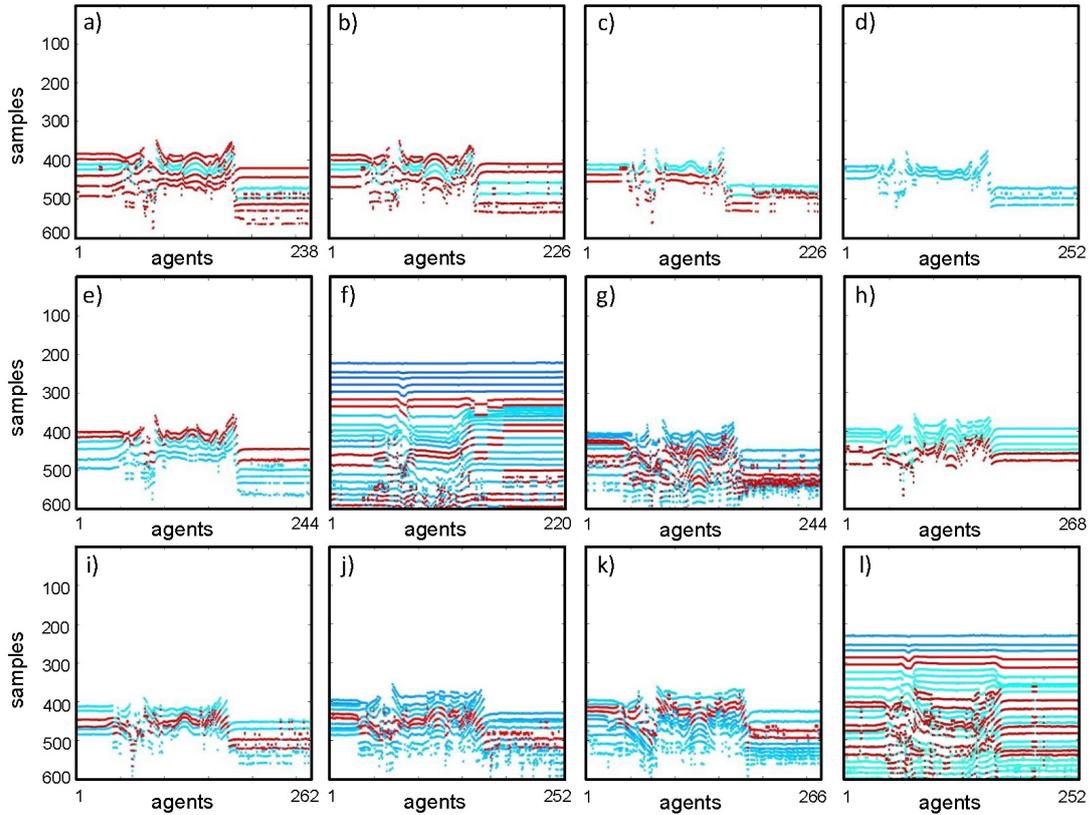


Figure 8.7: Images resulting from the removal of pipe-free layers. Source: Ayala-Cabrera *et al.* (2014a)

- **Fib and Fund.** Of course, for these materials also the visualization of the various characteristics is improved. In the case of the concrete (Fib) pipe a number of unexpected hyperbolae are observed. The extra hyperbolae correspond to intrusion layers developed within the pipe, since an old pipe was used in this experiment. This results is worth mentioning since it opens an interesting possibility of exploring the interior of the pipes, what will be of paramount importance in water system management.

Although most of the procedures are performed automatically, the removal of layers was performed manually in this study. Nevertheless, this process could also be performed automatically provided that suitable processes for pattern recognition would have been developed.

In addition to the results shown in this study, the proposed methodology was tested with 1140 images of pipes of various materials through suitable GPR inspections. By applying this methodology the identification was successful in 95% of the cases, since cleaner images with a significant reduction of points were rendered. Also, the decrease regarding the number of points constituting the processed image is worth mentioning: the processed images had on average less than 60% of information compared with the raw images. This represents an important step forward regarding posterior classification using intelligent methodologies, since a lower amount of information is considered, while simultaneously preserving the main characteristics of the image and obtaining more clear soil profiles. These profiles would ease 3D images interpolation and posterior production of maps of isoclines.

8.5 Summary and comments

In this chapter, a tool for unsupervised classification of soil characteristics from GPR surveys, based on multi-agent and clustering approaches, has been presented. We have specifically focused on the identification of pipes of various materials in the prospected area. The proposed methodology is able to suitably group data into clusters, thus suitably classifying the information for better interpretation. As a result, soil is characterized in a natural and quick way, what favors the interpretation of GPR files performed by non-highly specialized personnel. Also, there is no need of any a priori parameter assumption. The results are promising, since they help reduce the amount of information needed to be dealt with, while preserving the essential data for materials identification. This produces more clear visualization of pipes in water distribution systems, thus favoring better identification of the system components. The tool developed, in conjunction with the non-destructive nature of a powerful technique as the GPR, can be used to create soil profiles with automatic procedures that will favor WSS technical management.

8.5. Summary and comments

Chapter 9

GPR-based water leak models in water supply systems

Motivated by the interesting progress with visualization of buried pipes from GPR images we have achieved, which has been presented in the previous chapters, in this chapter we extend the developed knowledge and build useful tools to consider another omnipresent, although undesirable, elements in WSSs. This chapter addresses the problem of leakage in water distribution systems through the use of GPR as a non-destructive method. Laboratory tests are performed to extract features of water leakage from the obtained GPR images. Moreover, a test in a real-world urban system under real conditions is performed. Feature extraction is performed by interpreting GPR images with the support of a pre-processing methodology based on an appropriate combination of statistical methods and multi-agent systems. Specifically, the *Agent race* algorithm (see Chapter 7) was used in this chapter. The results of these tests are presented, interpreted, analyzed and discussed in this chapter.

Identifying leaks using GPR images is not an easy task and requires a high level of expertise by the operator. Added complications include the complex spatial arrangement of many networks, along with the steady growth in the supply infrastructure of cities. These aspects greatly increase the difficulty in using and interpreting data obtained with GPR in detecting leaks and analyzing the results, thus reducing the potential for solving problems and increasing the need for highly qualified personnel. This chapter attempts to address this issue with the extraction of features visually and numerically. To this purpose, laboratory tests are performed in which we seek to extract features of water leakage from GPR images. Feature extraction is performed by interpreting the GPR images with the support of a pre-processing methodology based on an appropriate combination of statistical methods and multi-agent systems. Subsequently, these features are observed in a field test on a real water leak-case in a WSS. The ultimate goal of these processes is to extract features to feed intelligent automatic mechanisms for (automatic) detection of leaks in WSSs using GPR images. This chapter thus seeks to encourage the use of these tools in finding leaks by non-highly qualified personnel and, thus, promoting improved management of WSSs.

obtained in Sections 9.4 and 9.5 are then presented in Section 9.7. The eighth section presents the identification of a leak in fieldwork. Finally, a section of conclusions closes the chapter.

9.1 Short introduction to the role of GPR to tackle leaks in WSSs

Water is valuable, but challenging to manage. It has been calculated that WSSs around the world lose more than 40 percent of the clean water pumped into the distribution system because of leaks before water reaches end consumers [GrowingBlue (2013)]. By reducing the amount of water leaked, WSS managers can reduce the amount of money and energy spent on producing and pumping water, increase system reliability and more easily satisfy present and future consumer needs. Having access to sufficient information regarding leaks is a complex task. Many water utilities struggle to measure and locate leaks in their distribution networks.

Improved leakage management in WSSs is one of the intelligent solutions that can make a difference. The use of different types of smart sensors to gather data and the application of advanced analytics could provide valuable information on the location of leaks in the network. Specifically, non-destructive methods, such as ground penetrating radar (GPR), can help locate primordial leaks and, so, help resolve the problem, while avoiding social and economic costs.

In Hao *et al.* (2012), a review of the various pipeline inspection techniques most commonly used in water supply and wastewater collection systems is performed. These techniques are divided into four groups: (a) visual techniques; (b) electromagnetic and radio frequency techniques; (c) acoustic and vibration techniques; and (d) other techniques. Closed-circuit television (CCTV) and sewer scanner and evaluation technology (SSET) are highlighted in the first group (visual techniques). The second group (electromagnetic and radio frequency techniques) consists of magnetic flux leakage (MFL), eddy current technique, hydroscope technology (HT), rapid magnetic permeability scan (RMPS), low frequency electromagnetic field (LFEM), passive magnetic fields (PMFs), time domain ultra wide band (UWB) and ground penetrating radar (GPR). The third group (acoustic and vibration techniques) includes sonar, vibro-acoustics, impact

9.1. Short introduction to the role of GPR to tackle leaks in WSSs

echo/spectral analysis of surface waves and correlator and listening sticks for leaks. The last group (other techniques) includes infrared thermography, continuous wave Doppler sensing technique, laser surveys, combined techniques (broad-band electromagnetics/wave impedance probe (WIP), pipe inspection real-time assessment technique (PIRAT) and the Sahara Project.

Among these techniques, the most popular for locating leaks in water supply systems are those included in the acoustic and vibration technique group, infrared thermography and GPR [Demirci *et al.* (2012); Dong *et al.* (2011)]. Acoustic methods detect the acoustic wave generated by the leak based on correlation analyses of the wave velocity of the sound emitted by the pipe being inspected. Such methods are widely used to identify leaks in fluid-filled metal pipes [Juliano *et al.* (2013)]. The main drawback of the aforementioned methods is their inefficiency in detecting leaks in non-metallic pipes (e.g., polyvinyl chloride PVC pipes) [Bimpas *et al.* (2010)]. Infrared methods detect thermal contrasts caused by the difference of temperature between ground and water. However, even though easy to implement, these methods produce errors when there are considerable differences in temperature. Furthermore, it is not possible to use these techniques in summer and winter, due to the absence of significant differences between ground and water [Dong *et al.* (2011)]. GPR is shown as an effective non-destructive tool that favors inspection of WSSs by demarcating on GPR image (radargrams) contrasts between the leaked water and the surrounding soil that are caused by differences in dielectric characteristics [Crocco *et al.* (2010); Charlton & Mulligan (2001)].

The use of GPR as a method for locating leaks in WSSs has become more widespread in recent years. In this sense, there is fieldwork, such as Stampolidis *et al.* (2003), performed on urban pipe sections. Pre-processing of the obtained images is performed by using low-pass filters to identify leaking PVC pipes. Another representative fieldwork is reported in Hunaidi & Giamou (1998). In this case, a plastic pipe (PVC) was drilled and buried in the ground; and, an analysis was made using raw images. Likewise, there is fieldwork using a combination of methods. Such is the case of Liu *et al.* (2002), which combines GPR assays

9.2. Data capturing: design and layout of the laboratory tests

with electrical potential and geochemical assays to detect leaks in non-pressurized non-metallic pipes. In this case, leaks are identified from raw GPR images. Laboratory tests are also employed in finding leaks using GPR. Works, such as Hyun *et al.* (2007) and Hasan (2012), concentrate on plastic pipes. In these cases, pre-processing includes background removal and image filtering, respectively. A combination of survey work conducted both in the field and in the laboratory is presented in Tavera (2008). In this work, various tests on leaks in plastic and metallic pipes were performed. In this work, Kirchhoff migration and the Hilbert transform were used as pre-processing methods. These assays are promising with respect to the use of GPR in finding leaks in WSSs. However, most of these assays are based only on the location and interpretation of hyperbolas generated either in raw or pre-processed images.

9.2 Data capturing: design and layout of the laboratory tests

This section presents the layout of the laboratory tests. In the set of tests performed, a pipe commonly used in WSSs is buried in dry soil in a tank (see Figure 9.2,a). The characteristics of the buried pipe are: (a) PVC; (b) diameter of 100 mm; (c) length of 0.95 m; (d) hole drilled for simulating the leak in the central part of the pipe; (e) two points for water input (herein termed WI) and water output (herein termed WO) with connections at the ends. A wooden tank measuring $1.0 \text{ m} \times 1.0 \text{ m} \times 0.70 \text{ m}$ was used. After the pipe was positioned, its supports were removed, and it was then covered with dry soil. The surface of the tank was covered with a polypropylene plate. Eleven paths parallel to the x -axis and another eleven paths parallel to the y -axis were marked on this plate. These 22 paths were spaced 0.10 m apart (see Figure 9.2,b) and, so, formed the sampling grid (see Figure 9.2,c). Onto each line of this mesh, the GPR antenna was slipped. The image produced by sliding the antenna is termed a profile in this chapter (see, for example, s5 in Figure 9.2,c). Additionally, we distinguish

9.2. Data capturing: design and layout of the laboratory tests

between horizontal and vertical profiles when referring to the profiles parallel to the x - and the y -axes, respectively.

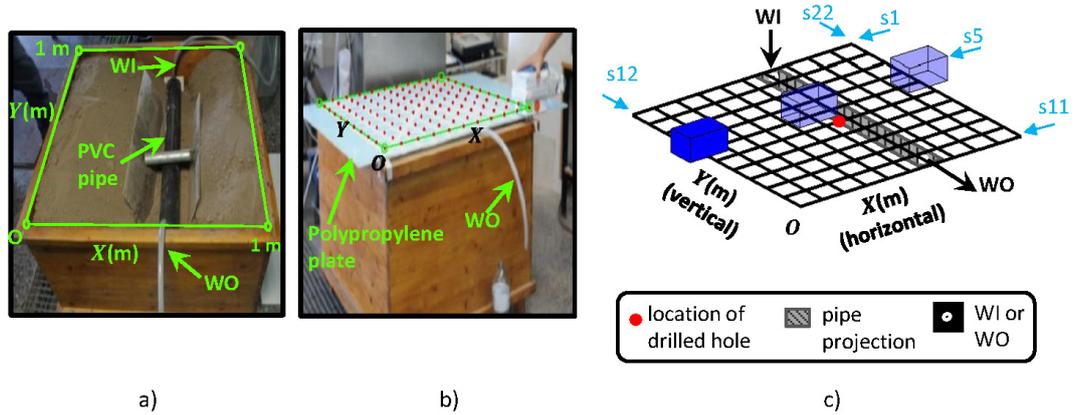


Figure 9.2: Laboratory layout. Source: Ayala-Cabrera *et al.* (2013a)

The GPR equipment used in each survey corresponds to a commercial mono-static antenna with a central frequency of 1.5 GHz. The parameters of the equipment used for these captures correspond to 120 traces/s, 512 samples/trace and 20 η_s /512 samples. These parameters correspond to the assay configuration 80 (see Appendix B). Two tests with the aforementioned features of 22 profiles were performed. These two tests are differentiated one from the other in that the first was performed without water inside the pipe (and, consequently, without leaked water), while the second test had water inside the pipe and leaked water. These two situations are called initial and final state, respectively, in this chapter.

9.3. Analysis of the location and identification of hyperbolas from the raw images

9.3 Analysis of the location and identification of hyperbolas from the raw images

This section presents a typical analysis to identify anomalies (leaks in this case) in the raw images. This analysis consists in locating and identifying hyperbolas in the GPR images, according to the schema presented in the previous section. These tests are performed on the horizontal and vertical profiles for both the initial and the end states. To facilitate the interpretation of raw images in Sections 9.3.1 and 9.3.2, various examples of target forms are sketched in Figure 9.3. Incidentally, Figure 9.3,a corresponds to Figure 9.6,a, Figure 9.3,b to Figure 9.6,c and Figure 9.3,c to Figure 9.6,g.

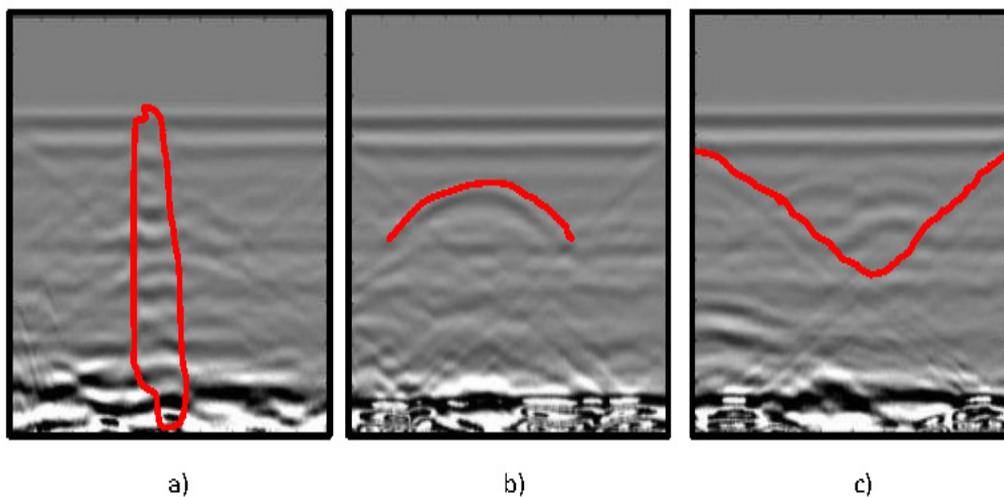


Figure 9.3: Identifiable forms in raw images in Sections 9.3.1 and 9.3.2. (a) distorting vertical strip, (b) hyperbola, and (c) triangles. Source: Ayala-Cabrera *et al.* (2013a)

9.3.1 Initial state: raw images

The resulting images from the horizontal and vertical profiles of the tests performed in the laboratory for the initial state are presented in Figures 9.4 and 9.5, respectively.

9.3. Analysis of the location and identification of hyperbolas from the raw images

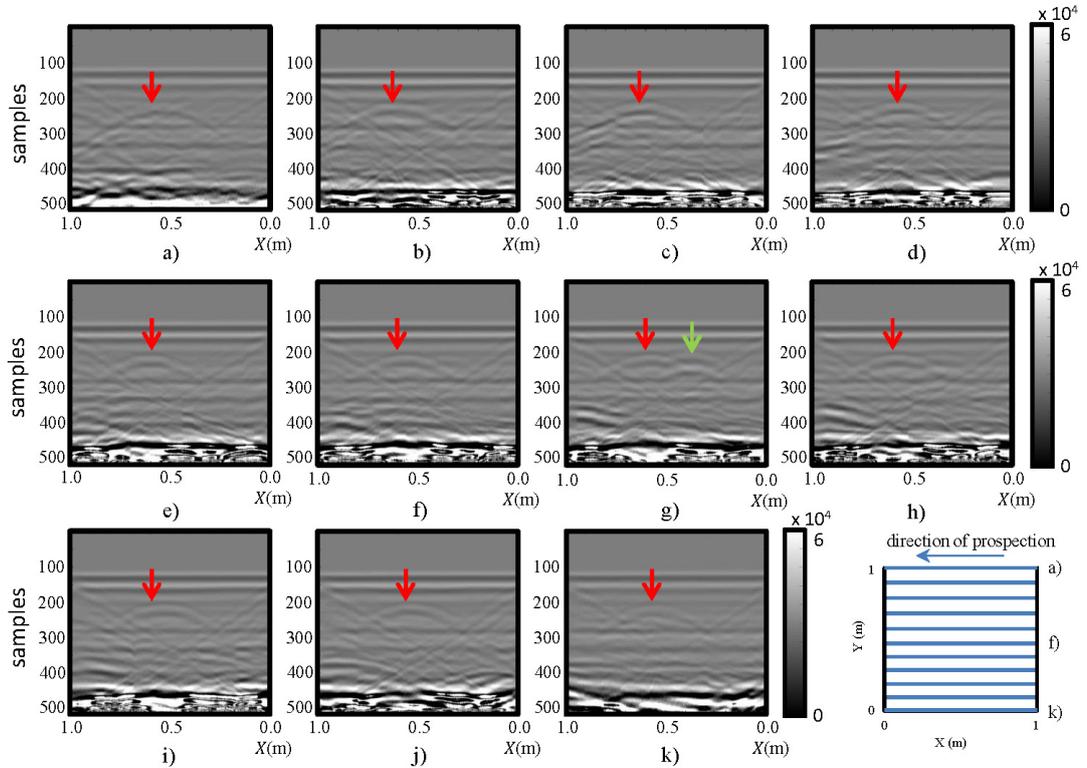


Figure 9.4: Horizontal profiles – Raw images for the initial state. (a–k) are profiles s1 to s11, respectively. Based on Ayala-Cabrera *et al.* (2013a)

All the images presented in Figure 9.4 exhibit a hyperbola, whose center is located at approximately 0.6 m in all cases (red arrows). This location coincides with the location of the PVC buried pipe. In this same figure, in parts (f), (g) and (h), a new anomaly, a new hyperbola located to the right of the hyperbola associated with the buried pipe, can be identified (green arrow). The new hyperbola increases in intensity in Figure 9.4,g. In this case, it is known that this anomaly cannot represent the leak, since we are still analyzing the initial state. However, in an uncontrolled case, this feature could lead to confusion.

All the images depicted in Figure 9.5 show how the hyperbola representing the pipe in Figure 9.4 no longer appears. This is consistent with the schematic configuration proposed for the test, given that vertical profiles are being considered. However, formations (triangles) inclined from the tank walls, which were also

9.3. Analysis of the location and identification of hyperbolas from the raw images

seen in the horizontal profiles (Figure 9.4), can be observed. In all the images in Figure 9.5, a gradually spanning structure can be observed (between red arrows): for parts from (a) to (d) between samples 275–325; for part (e) between samples 250–325; for (f) to (h) between samples 225–275; for part (i) between samples 275–325; and for (j) to (k) between samples 300–325. This pattern corresponds to the signal response for the now longitudinally placed pipe.

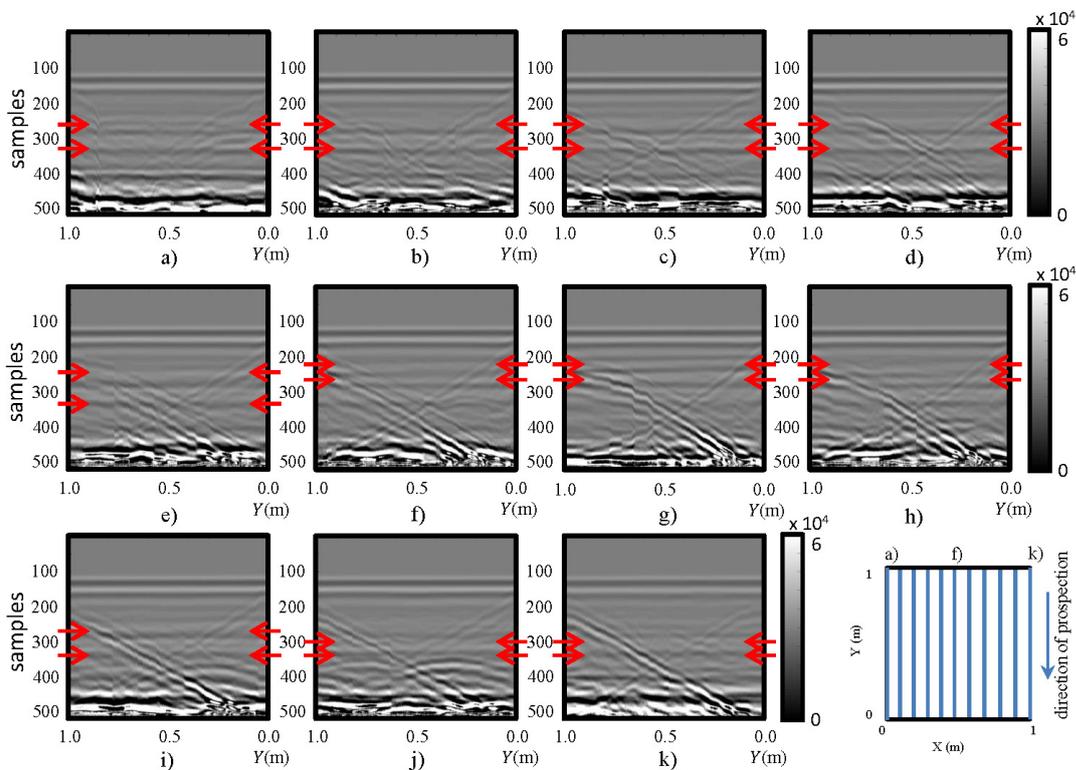


Figure 9.5: Vertical profiles – Raw images for the initial state. (a–k) are profiles s12 to s22, respectively. Based on Ayala-Cabrera *et al.* (2013a)

9.3.2 Final state: raw images

The resulting images of horizontal and vertical profiles of the tests performed in the laboratory for the final stage are presented in Figure 9.6 and Figure 9.7, respectively.

9.3. Analysis of the location and identification of hyperbolas from the raw images

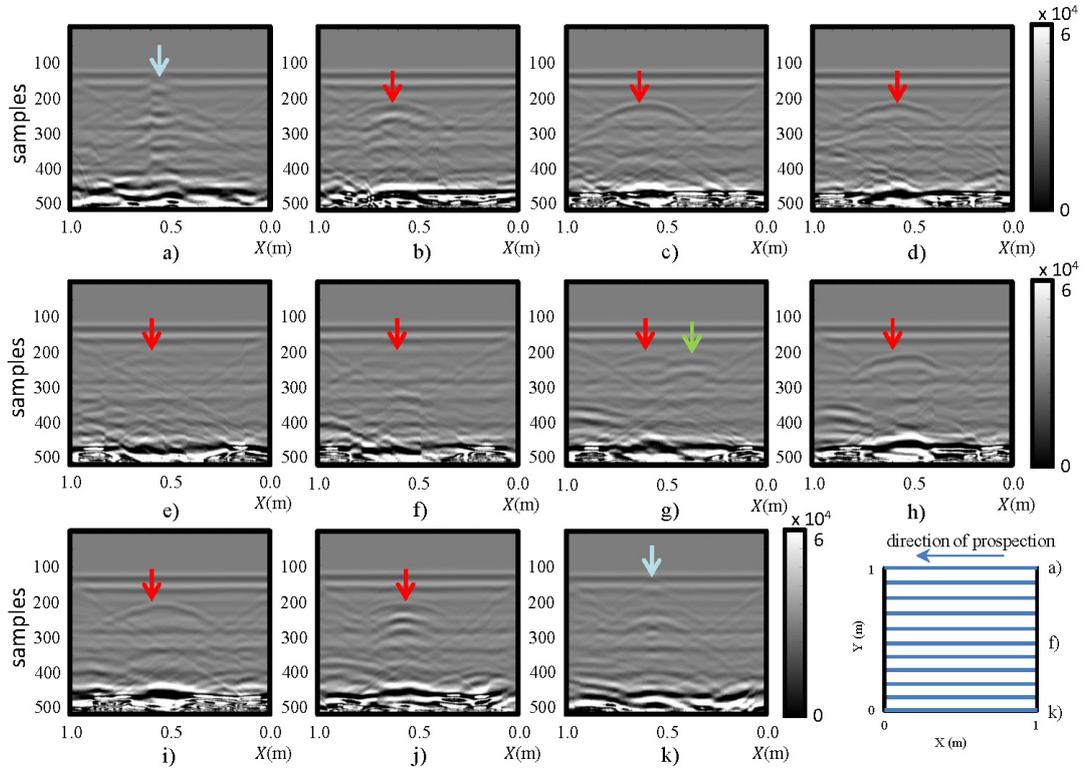


Figure 9.6: Horizontal profiles – Raw images for final state. (a–k) are profiles s1 to s11, respectively. Based on Ayala-Cabrera *et al.* (2013a)

In Figure 9.6, in images (a) and (k), a vertical strip that distorts both images can be seen between 0.5 and 0.6 m (blue arrows). This effect is attributed to the presence of water in WI and WO. This effect, which is more accentuated in (a), is more clearly seen when contrasting these images with their corresponding images in Figure 9.4: the area containing the deformed strip in Figure 9.6 contained (in Figure 9.4) the shape of the hyperbola previously identifying the buried pipe (red arrows). In images (b) and (j) of Figure 9.6, a new hyperbola surrounded by another upper hyperbola (pipe initial hyperbola) can be identified (red arrows). By observing profiles from (b) to (j), we can say that the hyperbola we had considered to represent the pipe decreases in intensity as the position of the profile nears the point of the leak. In effect, the hyperbola is almost invisible in images (e) and (f). We can also observe how in image (f), two hyperbolas are identified (very faintly), where, initially, only one hyperbola appeared. In (g), this second

9.3. Analysis of the location and identification of hyperbolas from the raw images

hyperbola as obtained in the initial state (close to the hyperbola of the pipe) is observed (green arrow). If we had not previously had the image without water, this second hyperbola would surely have been interpreted as a leak. This shows the need to extract patterns for interpretation and leak detection in water supply systems and, so, avoid common mistakes in field interpretation. This additional hyperbola may result from any object (such as stones) near the pipeline.

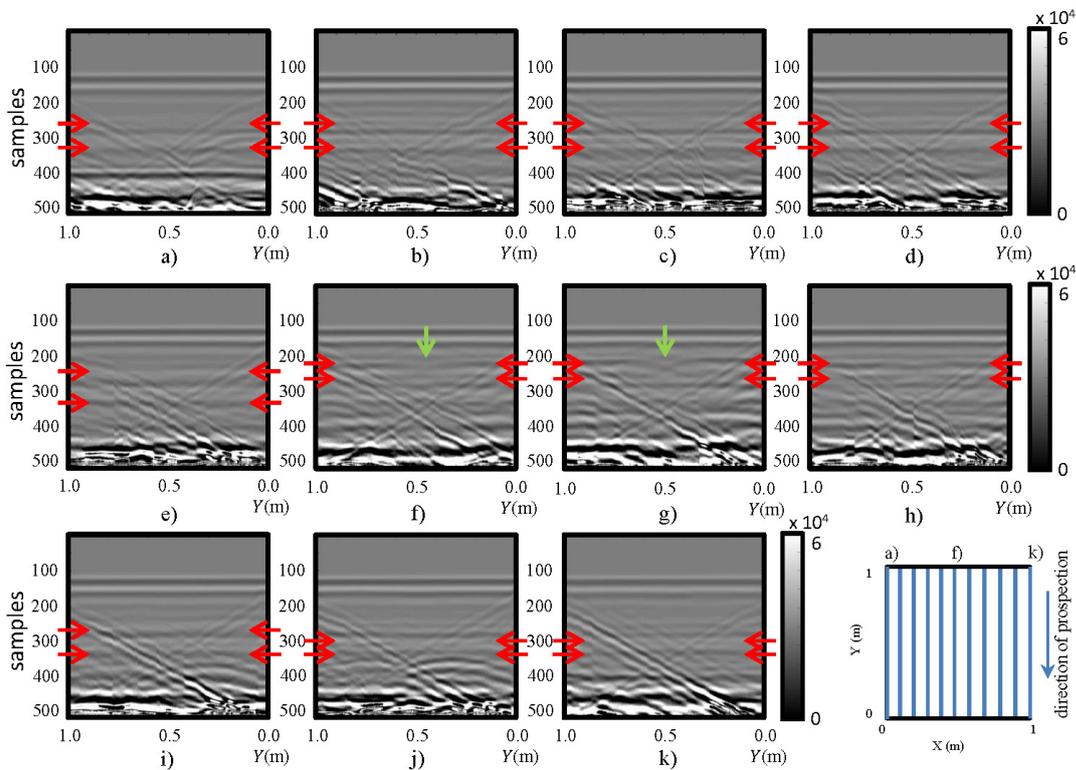


Figure 9.7: Vertical profiles – Raw images for final state. (a–k) are profiles s12 to s22, respectively. Based on Ayala-Cabrera *et al.* (2013a)

Images in Figure 9.7, in contrast, show similar characteristics to the respective images of the initial state. However, in (f) and (g), a contraction (green arrows) of the area we had identified as the pipe in Figure 9.5 can be observed between 0.5 and 0.4 m (distance axis) and samples 225 and 275 (depth axis). This contraction is better observed in image g), which corresponds to the vertical section closer to the pipe. Furthermore, the position of the contraction is consistent with the location of the leakage point in the test.

9.4. Analysis: contrast between raw images for the initial and final states

After these various analyses of the different horizontal and vertical profiles corresponding to the leakage, we observe that identifying leaks directly from raw images is difficult. In fact, even considering the *a priori* information about the location of the leak, it is a challenge to distinguish all the above-mentioned features.

9.4 Analysis: contrast between raw images for the initial and final states

In this section, since a suitable metric is available, we present a comparison between the initial and the final states of the tests performed for each raw image. Figures 9.8 and 9.9 show the differences in horizontal and vertical raw images, respectively.

The images in Figure 9.8 show the presence of water, which is the differentiating factor between the images of the initial and the final state. In parts (a) and (k) of the figure, the water in WI and WO can be easily seen (blue arrows). The water contained in the pipe can be observed in pictures from (b) to (j) (red arrows). The brightest part now corresponds to the water that has replaced the air contained in the pipe in the initial state. This shows that the difference in intensity identifies the specific fluid (air or water, in our case) contained in the pipe. Similarly, it is observed that the color intensity decreases as the location of the images nears profile (f) (the profile closest to the leak). Moreover, in images (e) to (g), between 0.4 and 0.6 m and between samples 150 and 200, a hyperbola that was not readily detectable in the respective raw images (and was missed in the interpretations) is now easily observed. This hyperbola is defined, according to the observation made *in situ*, by the rise in water leaked by capillarity. For images (e) and (f), the water contained in the pipe, on the right side (of the image), is demarcated with correspondingly lower intensity. This would indicate that the leak is running from left to right. This is because in that area, there is a mixture of soil and water, and this causes lower color intensity. A decrease in

9.4. Analysis: contrast between raw images for the initial and final states

color intensity in the demarcated area of the pipe when approaching the leakage point (profile f) can also be observed. Additionally, some deformation is also observed below all the areas demarcated by the contrast, which is the result of the expected convolution (bearing in mind that the measurements are performed in depth). Additionally, it can be observed that the anomaly observed in Figure 9.4,g (green arrow) is not any more visible in Figure 9.8,g (green arrow). This, effectively, shows that the anomaly corresponds to an object not related with the case study although very close to the pipe under study.

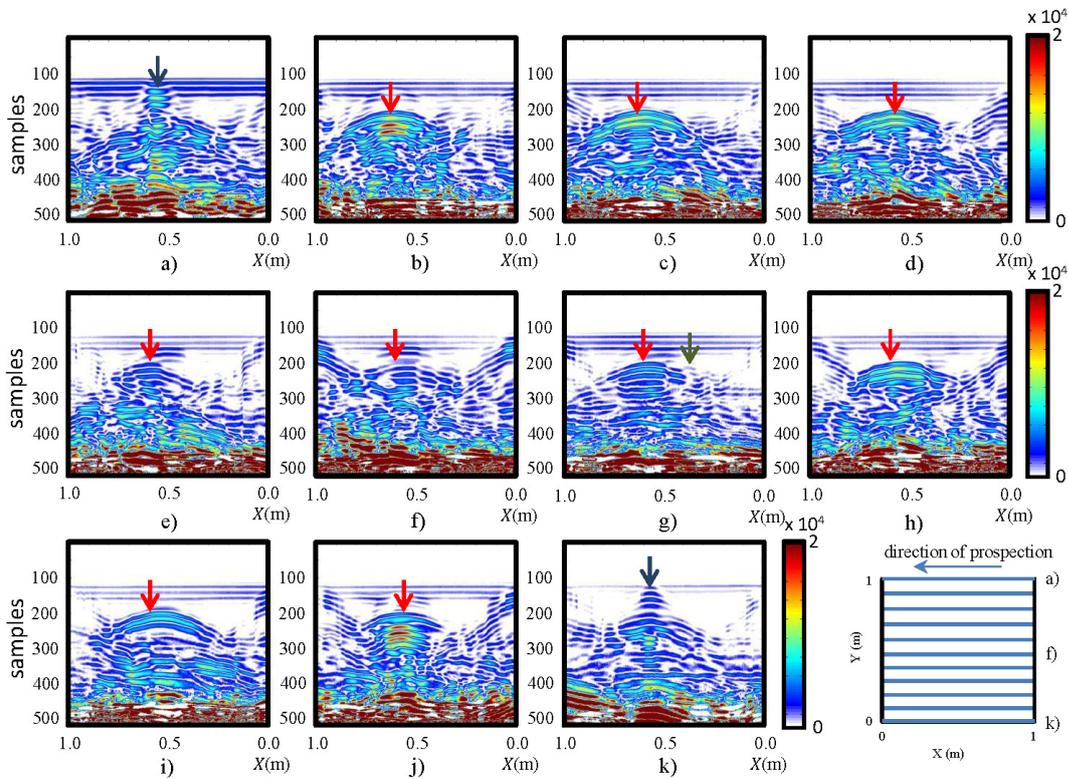


Figure 9.8: Horizontal profiles – Contrasting raw images for initial and final states. (a–k) are profiles s1 to s11, respectively. Based on Ayala-Cabrera *et al.* (2013a)

In Figure 9.9, image (g) (which is the closest to the pipe and to the leakage point), an increase in the color intensity in the strip running almost through the entire profile can be observed. This strip, between samples 200 and 250, shows the addition of water in the pipe. In this same image and for the same strip,

9.4. Analysis: contrast between raw images for the initial and final states

the effect described regarding the same image for the final state can also be seen. This is the effect of the contraction of the pipe at the point where the leak is located. In addition, there is a hyperbola at the top (a feature not seen in its respective final state image). Similarly, in (h) (the profile closest to the pipe and on the path where the pipe is located), the strip has lower color intensity than in (g), while being higher than in the closer profiles. Observe that this strip fades and diminishes in height, in the direction from (i) to (j) and from (f) to (e).

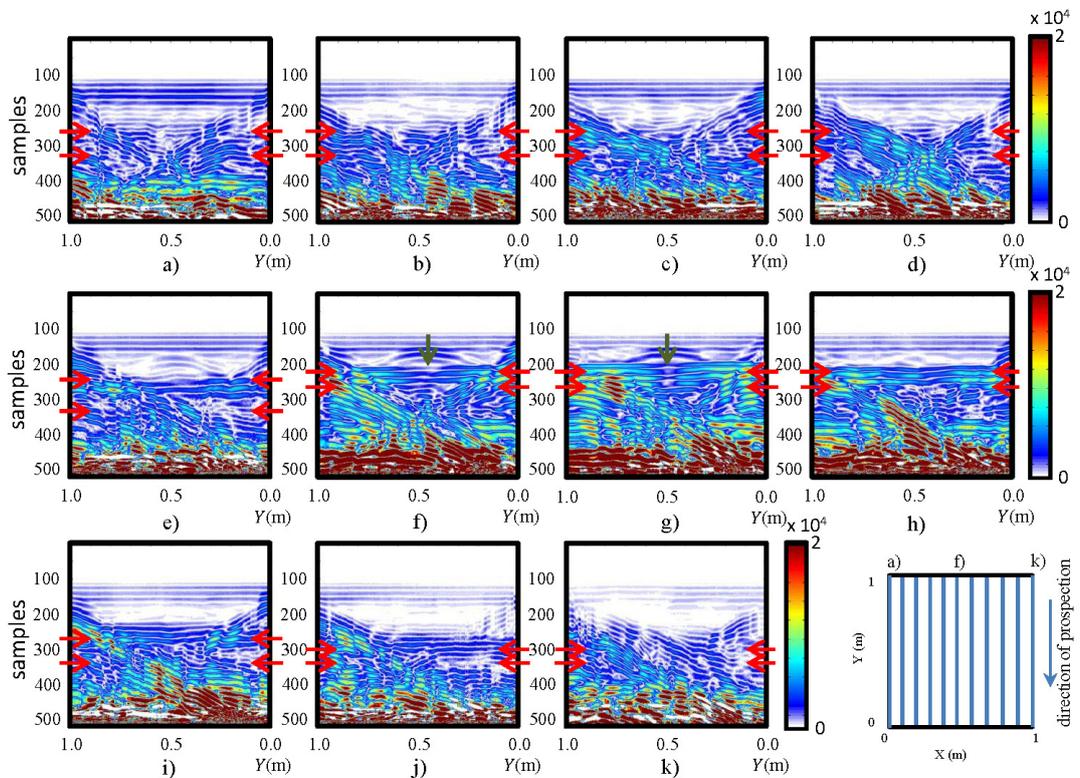


Figure 9.9: Vertical profiles – Contrasting raw images for initial and final states. (a–k) are profiles s12 to s22, respectively. Based on Ayala-Cabrera *et al.* (2013a)

Likewise, a hyperbola in (e), (f) and (g) (green arrows), which gains color intensity as the point of leakage is approached (image g) can be seen. From (f) to (e), the hyperbola loses strength and increases in size as a result of the increasing distance from the point of the leak.

9.5 Analysis of location and identification of anomalies in pre-processed images

In this section the used pre-processing principles and data ordering (including the posterior normalization) are presented. Subsequently, this algorithm is applied to the images obtained in the laboratory for the initial and final states. The results will be compared and analyzed later.

9.5.1 Data pre-processing and ordering algorithm

The pre-processing of GPR images used in this chapter was proposed in Chapter 7 and was termed an *Agent race*. The *Agent race* algorithm provides an interpretation and a grouping method for data from GPR radargrams. In this pre-process, we reduce the amount of data in the initial radargram, while preserving its initial properties and the most relevant data, so that its ability to identify buried objects through suitable visualizations is preserved. The input to this algorithm is the resulting radargram of the GPR exploration, which consists of an $m \times n$ -sized matrix. The n -traces, of length m , that are generated are used in this work as parallel tracks for the n -agents to run. The race is an endurance test for the agents with a prize consisting in advancing one position depending on the effort made. Efforts are based on wave amplitude values in each column of the matrix (radargram). The agent race is comprised of two phases: (a) warming-up; and (b) competition. The race lasts a total time of $t = tw + tr = m$, where tw is the warming-up time and tr the competition time. The movements of an agent in tr are conditioned by the changing trend of the wave amplitude of the trace the agent runs. The race ends when time t has elapsed. The winner is the agent who manages to obtain more movements (reward) during this time. The output (Output1), termed R matrix, of this process is an array of size $m1 \times n$, where $m1 =$ the maximum number of movements. The missing values of the rest of agents are filled with the maximum value of t (512 in this chapter) to complete the matrix. The columns of this matrix describe the movement of the agents in

9.5. Analysis of location and identification of anomalies in pre-processed images

relation to the competition. In this chapter, we restrict the output (Output2) to the matrix of the size $tr \times n$ that collects the movements performed by each agent during the competition (see Figure 9.10). In this chapter, the various movements developed by the agents are termed timelines.

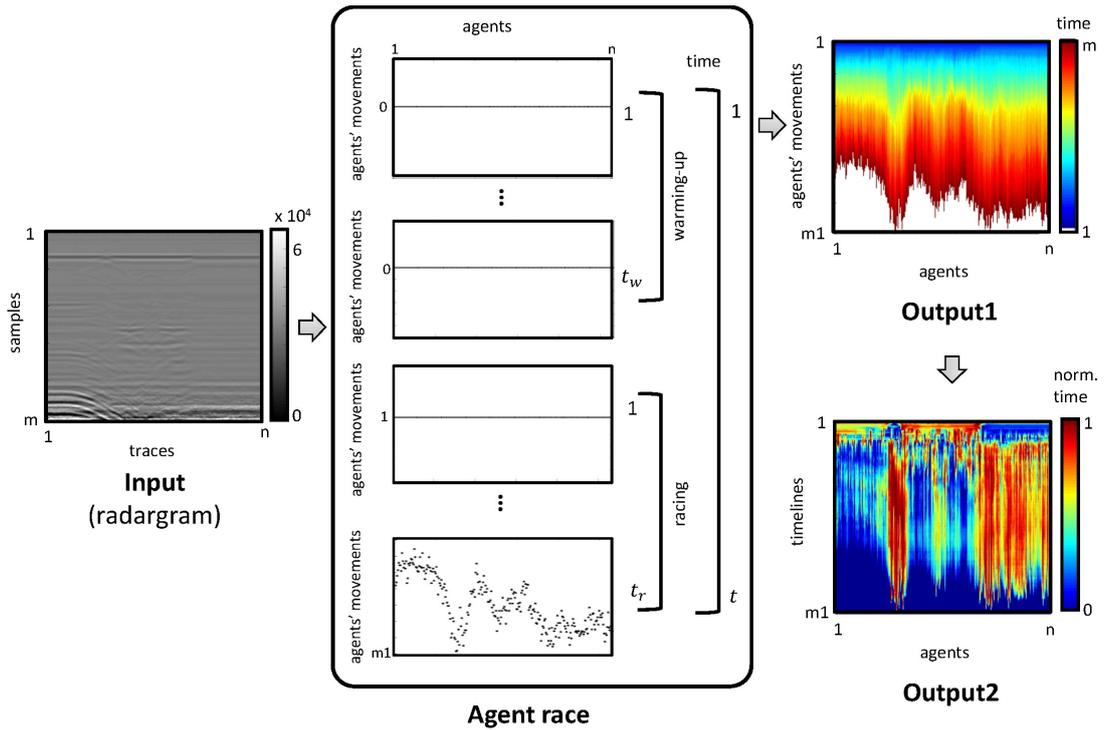


Figure 9.10: Scheme for the agent-race algorithm – Generation of Output1, and Output2. Source: Ayala-Cabrera *et al.* (2013a)

On each timeline, times obtained in the competition for each agent are ordered decreasingly. These timelines are then normalized. This produces Output2 (Figure 9.11), which is the matrix we use herein. This eliminates the delays caused by the gap between rows and enables more intense highlighting of anomalies that are very small and difficult to observe in raw images. However, care must be taken with this regularization, because although the anomalies are highlighted, their intensity is determined in relation to the ground prospected in each profile. This can cause visual errors in interpretation. Yet, even if this occurs, the interpretation based on the identification of various forms will enable zones of interest to be delimited, thus facilitating more complex analyses.

9.5. Analysis of location and identification of anomalies in pre-processed images

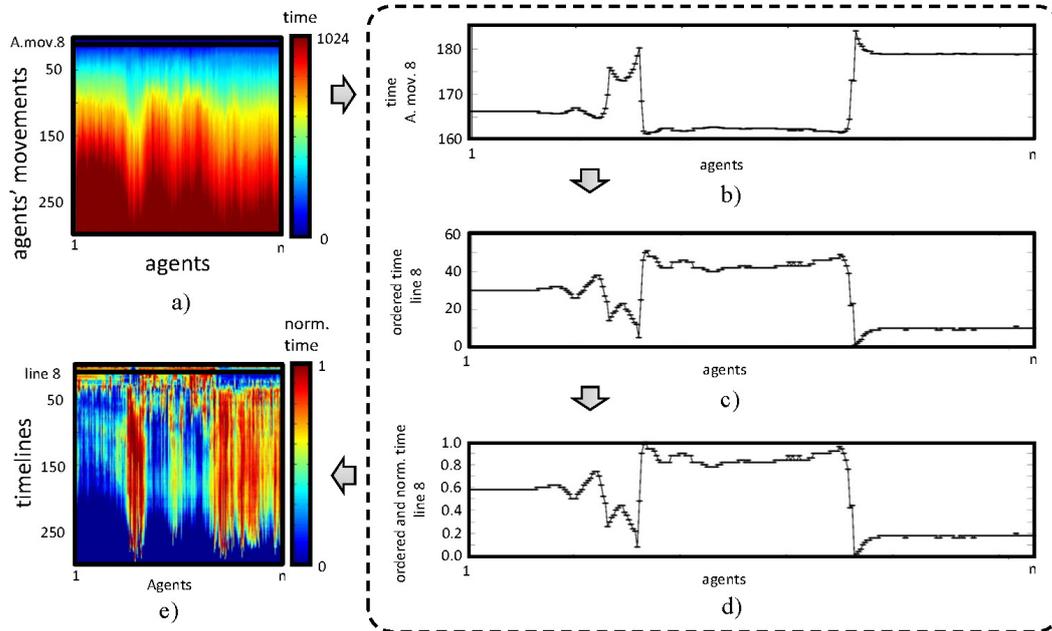


Figure 9.11: Output 2 building. (a) Output 1, (b) timeline, (c) ordered time line, (d) ordered and normalized timeline, and (e) Output 2. Source: Ayala-Cabrera *et al.* (2013a)

Again, to facilitate the interpretation of pre-processed images in Sections 9.5.2 and 9.5.3, examples of the more relevant forms are sketched in Figure 9.3. In this case, Figure 9.12,a corresponds to Figure 9.15,a, Figure 9.12,b to Figure 9.15,c, Figure 9.12,c to Figure 9.15,g and Figure 9.12,d to Figure 9.16,f.

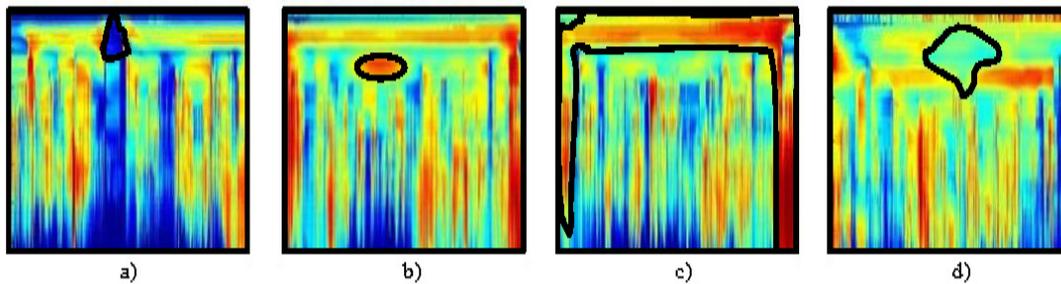


Figure 9.12: Identifiable target forms in pre-processed images in Section 9.5.2 and 9.5.3. (a) conical shape for water input (WI) and output (WO), (b) ellipse, (c) tank and measuring plate contours, and (d) water leak shape. Source: Ayala-Cabrera *et al.* (2013a)

9.5. Analysis of location and identification of anomalies in pre-processed images

9.5.2 Initial state: pre-processed images

The pre-processed images derived from the tests performed in the laboratory for the initial state show the horizontal and vertical profiles presented in Figures 9.13 and 9.14, respectively.

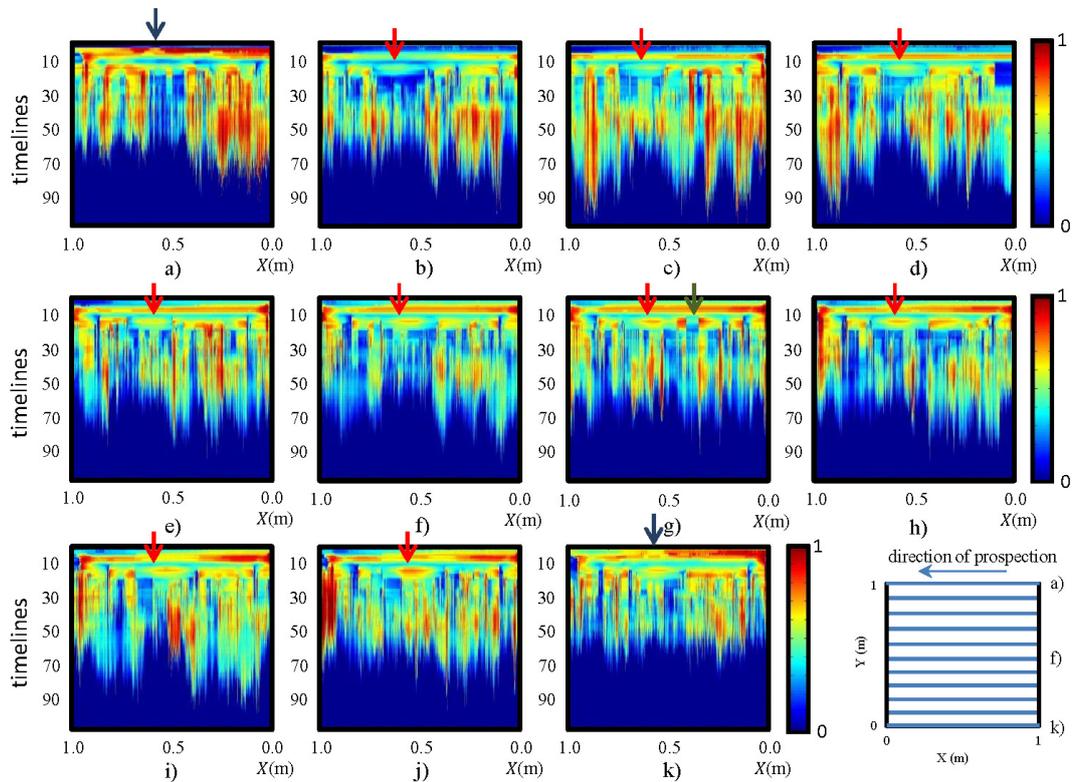


Figure 9.13: Horizontal profiles – Pre-processed images for the initial state. (a–k) are profiles s1 to s11, respectively. Based on Ayala-Cabrera *et al.* (2013a)

In Figure 9.13, an ellipse-shaped object is observed exactly in the place where the pipe was placed (red arrows). In each image, it can be observed how the color intensity decreases from the center of the ellipse to its border. In images (a) and (k), an additional conical shape may be observed above this ellipse, which is not noticeable in the corresponding raw images (blue arrows). This image corresponds to WI and WO. Additionally, at about timeline 10 in all the images

9.5. Analysis of location and identification of anomalies in pre-processed images

(with variable intensity), the development of a horizontal figure can be observed that crosses all the images, which, at each of its two ends, joins an intense zone that demarcates a vertical figure. This horizontal formation is the polypropylene plate used for sliding the antenna during measurement, and the vertical forms correspond to the walls of the tank. It can also be observed the persistence of the anomaly close to the pipe in Figure 9.13,g (green arrow).

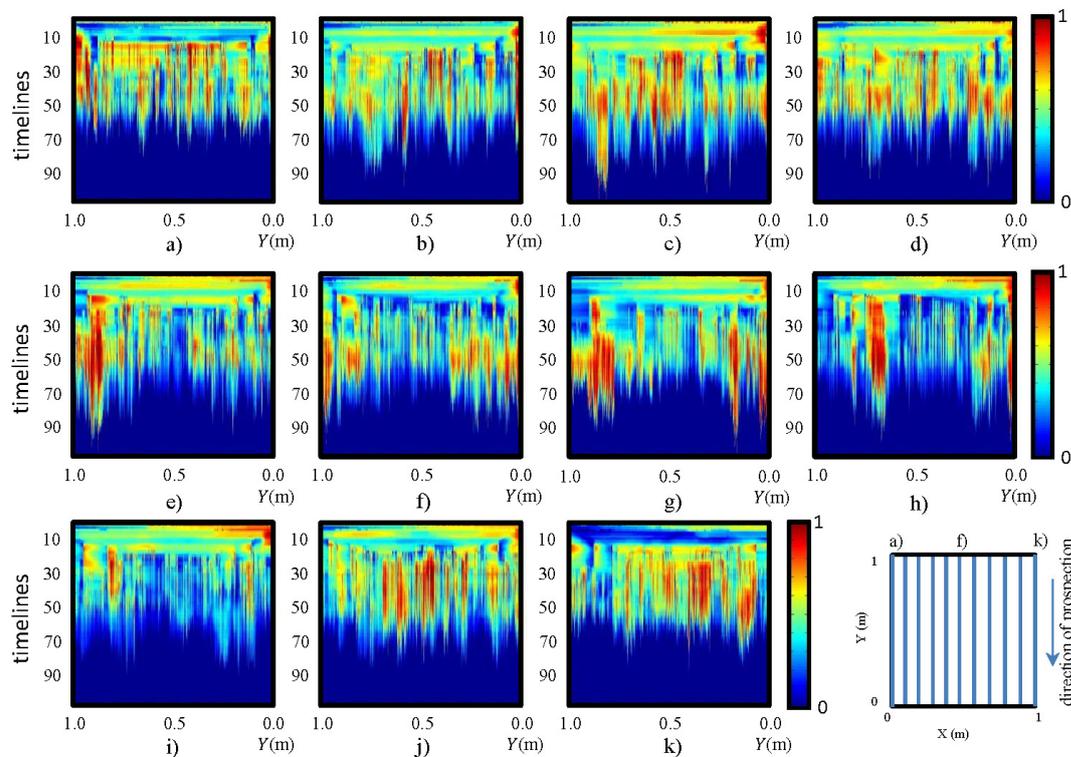


Figure 9.14: Vertical profiles – Pre-processed images for the initial state. (a–k) are profiles s12 to s22, respectively. Source: Ayala-Cabrera *et al.* (2013a)

In the images in Figure 9.14, a certain relationship between the different color intensities of each profile can be observed. The same characteristic areas of the tank and the measuring plate, already observed in the pre-processed horizontal profiles, are demarcated with greater intensity. However, the inclination of this area, which was previously horizontal, can be observed from images (f) to (h).

9.5. Analysis of location and identification of anomalies in pre-processed images

Moreover, in the central part, we note the non-appearance of the ellipse that appeared in the pictures above.

9.5.3 Final state: pre-processed images

The pre-processed images derived from the tests performed in the laboratory for the final state, showing the horizontal and vertical profiles, are presented in Figure 9.15 and Figure 9.16, respectively.

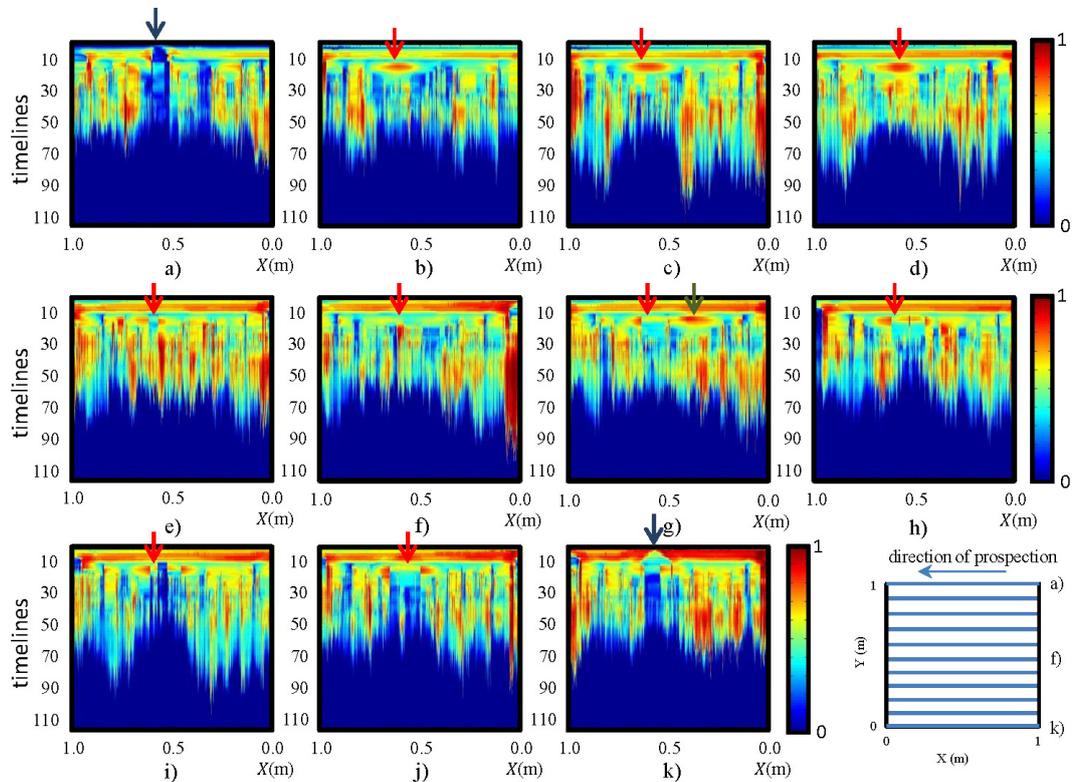


Figure 9.15: Horizontal profiles – Pre-processed images for the final state. (a–k) are profiles s1 to s11, respectively. Based on Ayala-Cabrera *et al.* (2013a)

In Figure 9.15, it can be seen how the ellipse observed in the corresponding initial state image gains color intensity with the addition of water in the pipe for all the images (red arrows). Likewise, in images (e) and (f), such intensity

9.5. Analysis of location and identification of anomalies in pre-processed images

decreases. This feature was observed in the profile of the raw images and can be seen more clearly in the pre-processed images. Likewise, it can be observed in (a) and (k) that the intensity of the structure representing WI and WO increases with the addition of water to the system (blue arrows). It should be noted that this feature is clearly observed only after contrasting the initial and final states and not directly in the raw images. However, in the pre-processed image, this feature is easily identifiable. In images (g) to (k), the central ellipse loses intensity toward 0 m and fades in its central part. The contours generated by the tank and measuring plate are better delineated in these images. We have to mention here that the anomaly that persisted along all the images (Figure 9.15 – green arrow) gains color intensity and elliptic shape consistency after water addition to the system.

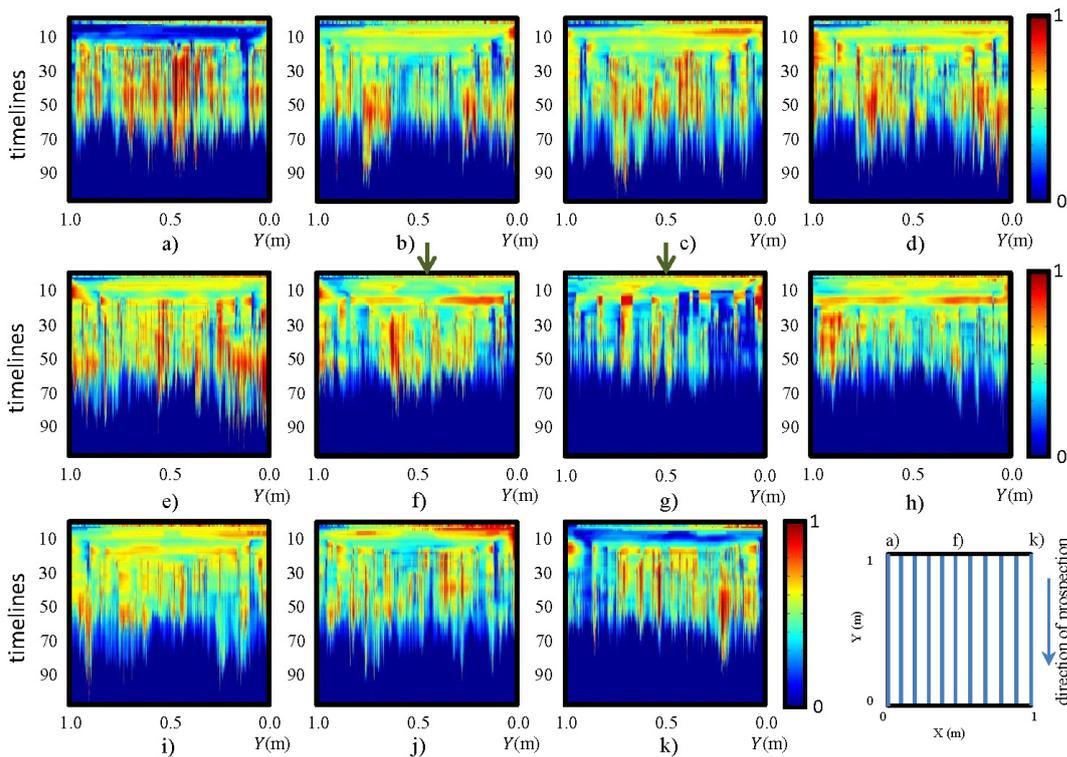


Figure 9.16: Vertical profiles – Pre-processed images for the final state. (a–k) are profiles s12 to s22, respectively. Based on Ayala-Cabrera *et al.* (2013a)

9.6. Analysis: contrast between pre-processed images

In Figure 9.16, a discontinuity of color intensity in the form of an ellipse between images (e), (f) and (g) (green arrows) can be observed. This discontinuity grows as it approaches profile (g) (location of the leak). It should be mentioned that this was not at all observable in the corresponding raw image. Now, in the pre-processed images, it becomes apparent.

9.6 Analysis: contrast between pre-processed images

In this section, we present a number of contrasts between the pre-processed profiles for the initial and final states. In Figure 9.17 and Figure 9.18, contrasts for the horizontal and vertical profiles, respectively, are presented.

In Figure 9.17, images (a) and (k), the mark that was observed in the interpretations already made can be easily seen (blue arrows). This mark corresponds to the presence of water at WI and WO. Additionally, in the same pictures, an anomaly in the form of an ellipse at the bottom can be appreciated (red arrows). This shape corresponds to the water in the pipe. Moreover, various profiles reveal the formation of the ellipse whose center is highly intense and which fades toward the ends. In these profiles, we see how the color intensity decreases as we near the central profile (leak location profile). Similarly, in images (e) and (g), one can observe the formation of a new ellipse that is located just above the ellipse corresponding to the water in the pipe. We also note that in image (f), this form is a hyperbola enveloping the ellipse corresponding to the area of water in the pipe. This form corresponds to leaked water, part of which has risen by capillarity towards the surface.

9.6. Analysis: contrast between pre-processed images

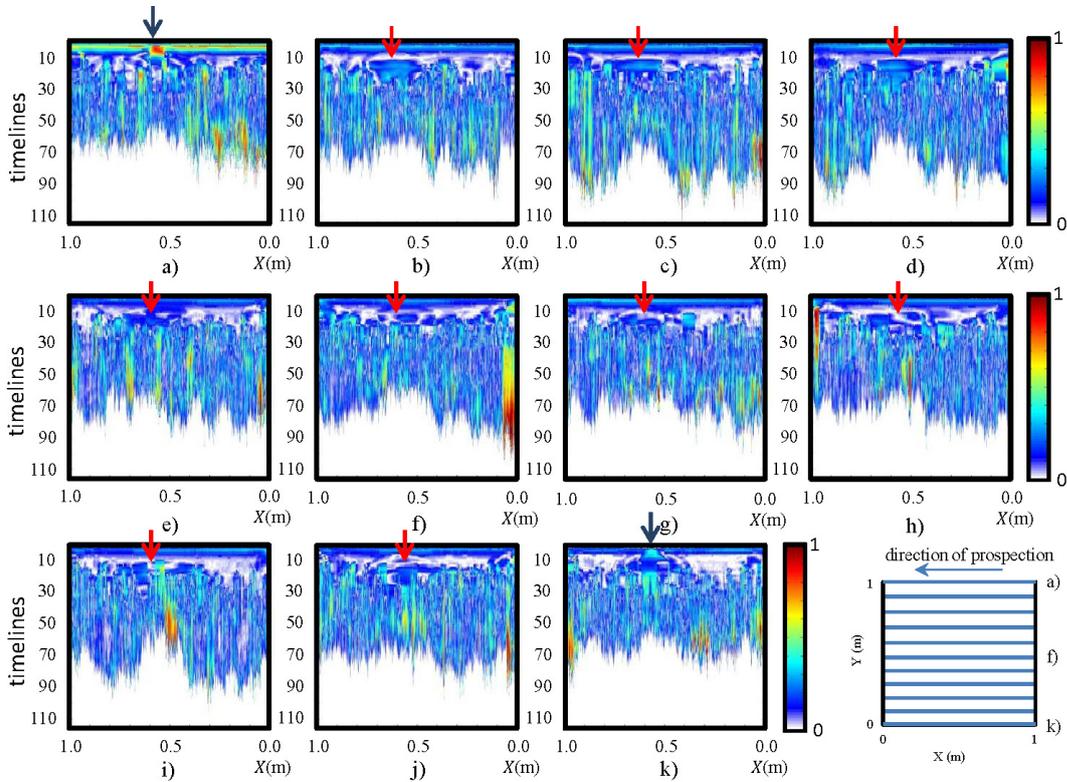


Figure 9.17: Horizontal profiles – Contrast between pre-processed images. (a–k) are profiles s1 to s11, respectively. Based on Ayala-Cabrera *et al.* (2013a)

In Figure 9.18, images (e), (f) and (g) (green arrows), one can observe the generation of an ellipsoidal shape that increases in size and consistency, as it travels from (e) to (g). Note that profile (g) is closest to the location of the leak.

9.7. 3D comparison of the analysis of contrasts between raw and pre-processed images

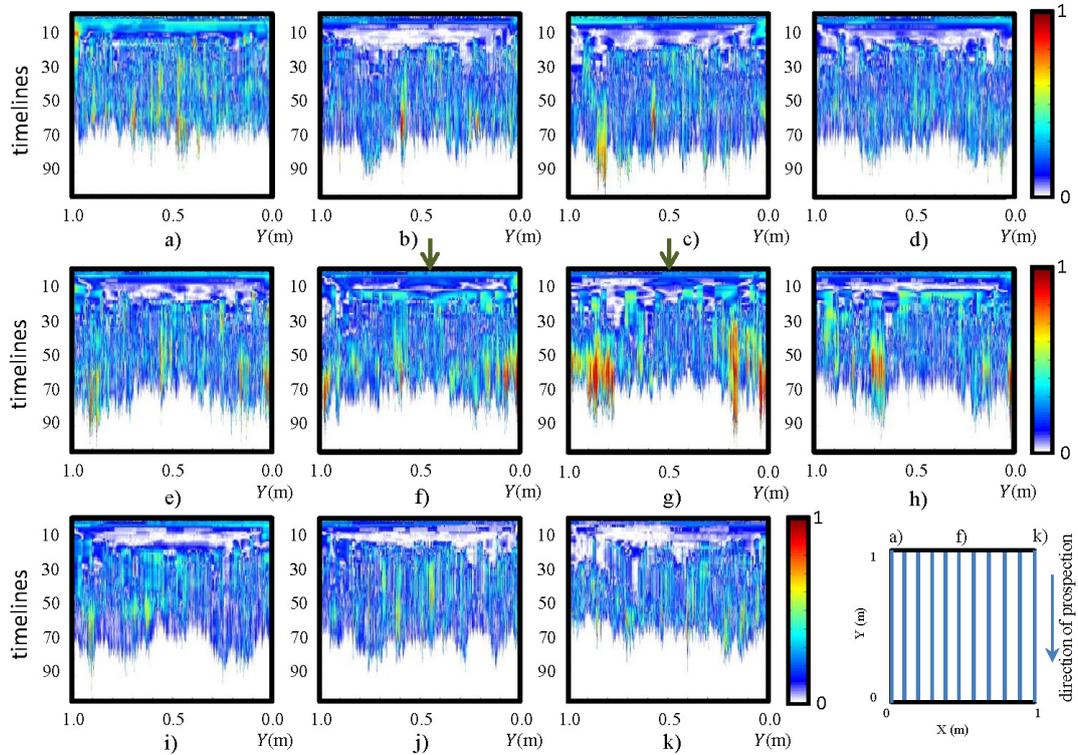


Figure 9.18: Vertical profiles – Contrast between pre-processed images. (a–k) are profiles s12 to s22, respectively. Based on Ayala-Cabrera *et al.* (2013a)

9.7 3D comparison of the analysis of contrasts between raw and pre-processed images

In this section, we perform the extraction of the contours of the images resulting from the previously conducted contrasts. With these contours, two 3D models of contrasting performances in raw images and pre-processed images of the water in the pipe and leaked water are obtained. Contour extraction and the corresponding 3D representation have been developed in MatLab. In Figure 9.19, we present a simple schema of one of the processes, whose two main steps are described as follows:

9.7. 3D comparison of the analysis of contrasts between raw and pre-processed images

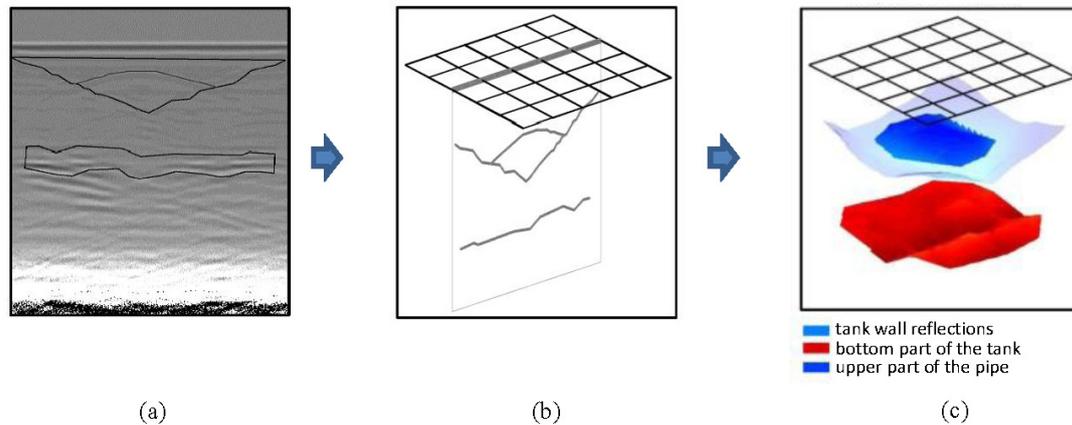


Figure 9.19: Example of 3D model construction. (a) GPR interpretation of raw data, (b) contour extraction, and (c) 3D model rendering. Source: Ayala-Cabrera *et al.* (2013a)

GPR interpretations. Contours that can represent features of interest in the prospected underground are extracted either from the raw or the pre-processed images. When interpreting images of radargrams, the features most commonly analyzed are hyperbolas. They are demarcated by the color intensity within the image. A non-automatic process is proposed in this chapter with the aim of determining the feasibility of interpretation of GPR data. In addition, we are interested in observing if these interpretations are reliable and can gain relevance for understanding features of leaks in WSS. However, if this process is effective, it could be implemented automatically.

Fusion process and 3D model. The fusion of the GPR image interpretations is achieved by their projection onto a common space. Firstly, a classification of common zones in the GPR interpretations that may correspond to specific sections of the area or volume is performed. Then, a mesh in the 3D space is constructed. Each of these volumes or surfaces is then added to a common space, shrinking, thus, into the 3D model. There are a number of applications to achieve this kind of fusion employed in photography or painting [Russell *et al.* (2011)], which can be adapted to generate 3D models of the GPR interpretations.

9.7. 3D comparison of the analysis of contrasts between raw and pre-processed images

In Figure 9.20, we present the 3D models.

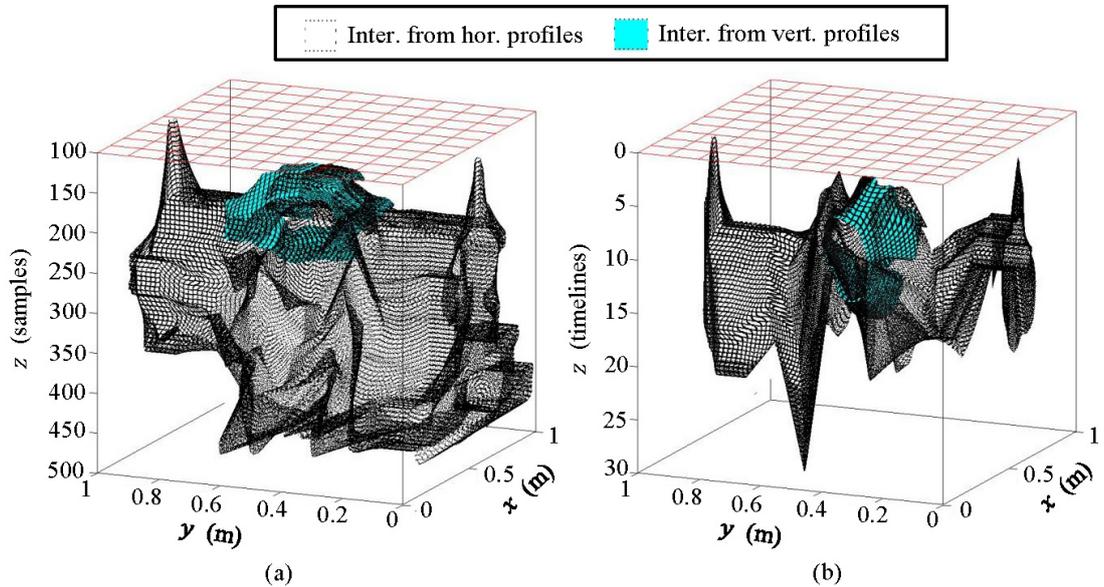


Figure 9.20: Comparison between 3D models generated from the interpretation of the horizontal and the vertical profiles contrasting the initial and the final states. (a) 3D model obtained from raw images; original space, and (b) 3D model obtained from pre-processed images; pre-processing space. Source: Ayala-Cabrera *et al.* (2013a)

The 3D models presented in Figure 9.20 show clear agreement with the schematic approach performed. It is also clear that the models are related, although their depth axis differs. Obviously, this is because they come from different images. In both cases, the water in WI and WO and the water in the pipe can be clearly observed. The leakage is also identified, but more consistently in (b) than in (a). It can be seen that 3D models produce more measurable results (numerically) and more comprehensible interpretation than flat images.

9.8 Analysis of field images: a case study

This section discusses the use of the multi-agent methodology, proposed in the previous section, in a real case. The fieldwork was conducted in an urban water supply system on a section of pipe, whose location was known. Using a geophone on an area where it was supposed that there was a water leak, turbulence points were detected. Also, tests on the area near the pipe were performed using GPR. The roadway is hydraulic concrete and the pipe is 400 mm cast iron. A structure of paths similar to those in the laboratory tests was demarcated on the roadway (Figure 9.21,a). The grid-mesh was 0.50 m. The area was 4.0 m by 2.0 m long in x and y -axes, respectively. The profiles obtained with the survey were named sf1 to sf8 and sf9 to sf13 for horizontal and vertical lines, respectively (Figure 9.21,b). It should be mentioned that the pipe is located in the sf11 profile. The GPR equipment used and the capture parameters correspond to the same features of the equipment used in the laboratory tests. However, in this case, the antenna used had a central frequency of 400 MHz. The equipment configuration used for these captures corresponds to configuration 116 (see Appendix B).

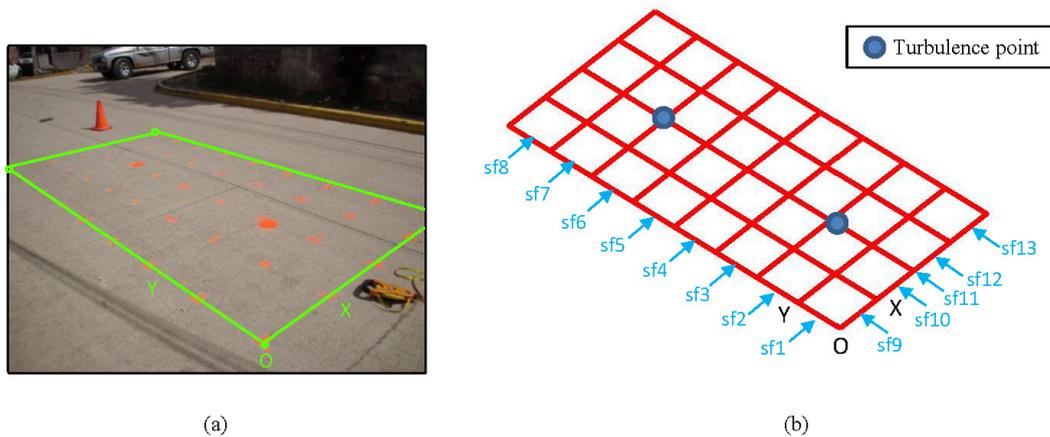


Figure 9.21: Layout for the fieldwork. Source: Ayala-Cabrera *et al.* (2013a)

In Figure 9.22 and Figure 9.23, the raw images for the horizontal and vertical profiles, respectively, are presented.

9.8. Analysis of field images: a case study

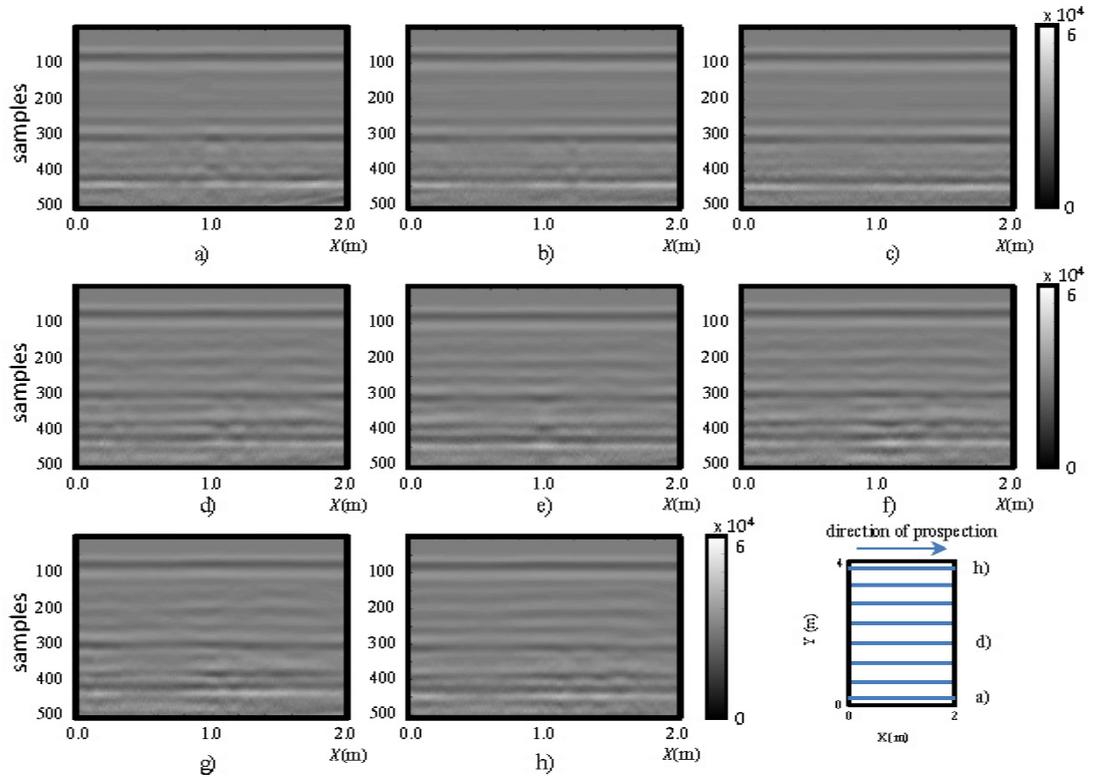


Figure 9.22: Horizontal profiles – Raw images. (a–h) are profiles sf1 to sf8, respectively. Source: Ayala-Cabrera *et al.* (2013a)

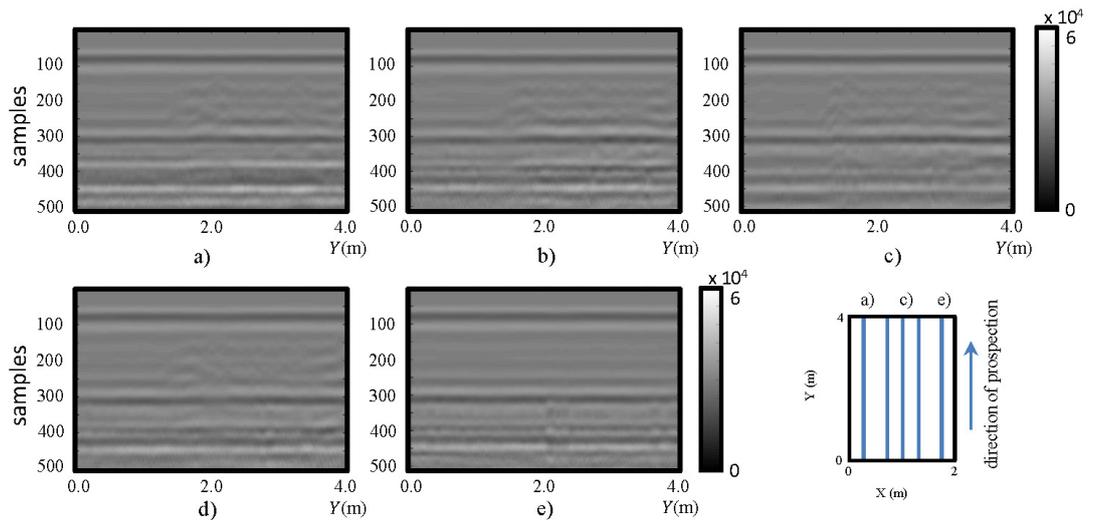


Figure 9.23: Vertical profiles: raw images. (a–e) are profiles sf9 to sf13, respectively. Source: Ayala-Cabrera *et al.* (2013a)

9.8. Analysis of field images: a case study

In Figures 9.22 and 9.23, the observation of anomalies is deemed very complex. Only a small perturbation from (a) to (d) in Figure 9.23 can be observed.

In Figures 9.24 and 9.25, we present the fieldwork pre-processed images corresponding to the horizontal and vertical profiles, respectively.

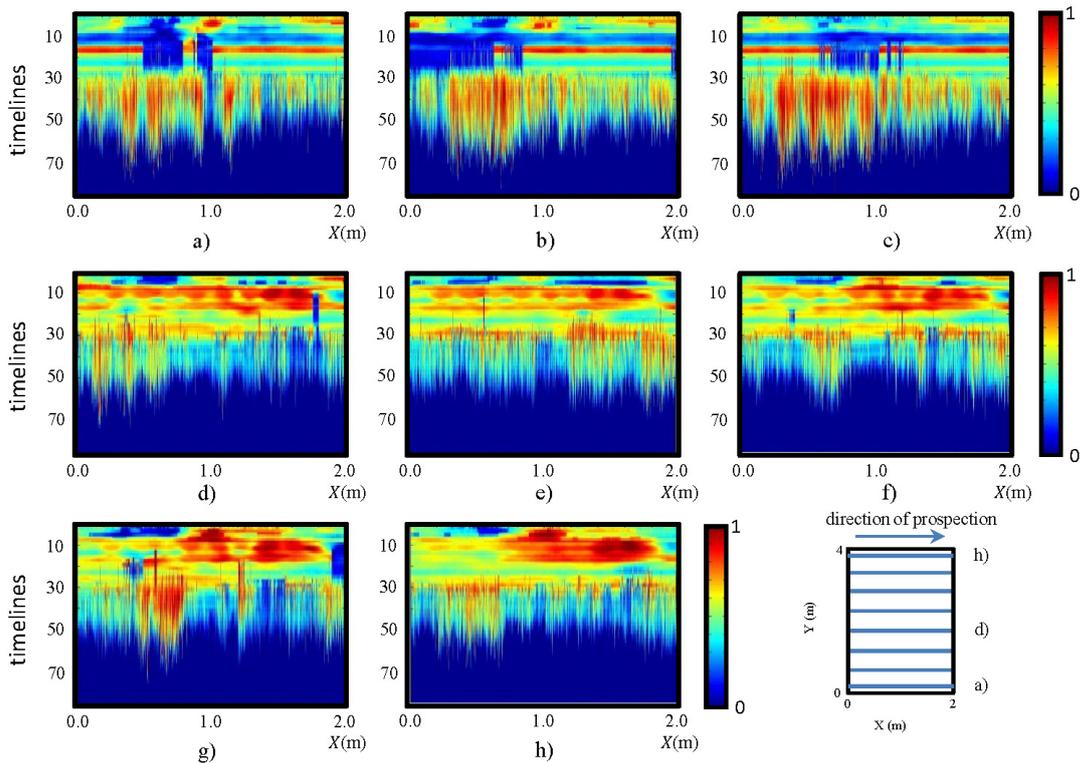


Figure 9.24: Horizontal profiles: pre-processed images. (a–h) are profiles sf1 a sf8, respectively. Source: Ayala-Cabrera *et al.* (2013a)

In Figures 9.24 and 9.25, the visualization improvement generated with the pre-processing can be clearly seen, making it possible to identify the elliptical shape of the pipe in the horizontal profiles (Figure 9.24), centered approximately on the position (1 m, sample 5). Moreover, in this same figure, in pictures (d) to (h), a structure can be seen that starts at the bottom right of the ellipse corresponding to the pipe. This structure gains strength and is clearly demarcated as an ellipse in the direction (d) to (h). In Figure 9.25, two vertical structures at 1.5 m and 3.0 m, in sample 10, are also seen. These structures start from

9.8. Analysis of field images: a case study

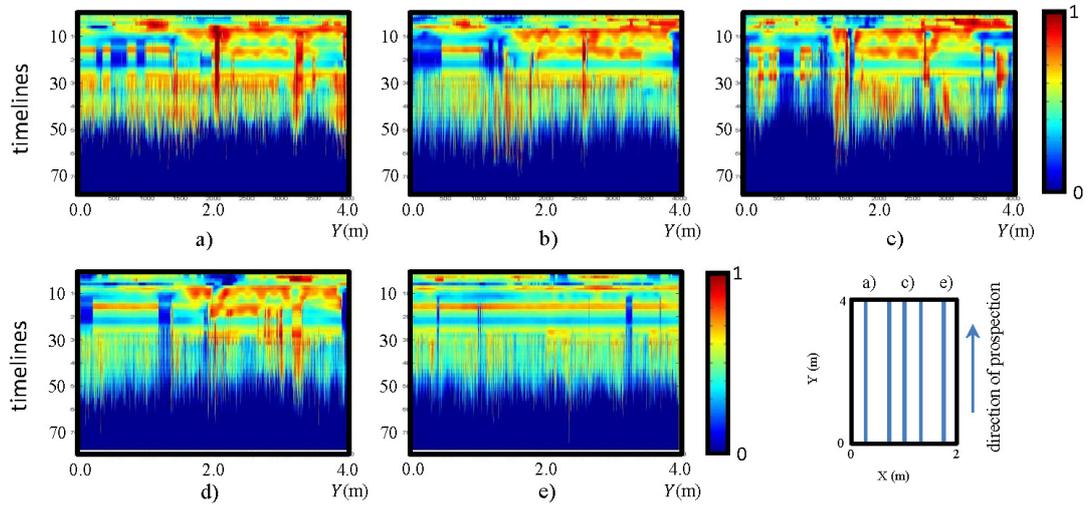


Figure 9.25: Vertical profiles – Pre-processed images. (a–h) are profiles sf9 to sf13, respectively. Source: Ayala-Cabrera *et al.* (2013a)

image (a), intensify towards image (c) and reduce their intensity in d), until finally disappearing in (e). This effect corresponds to the convolution observed in Figure 9.8.

Figure 9.26 shows the conjugation of these interpretations as 3D models, both for the pipe and likely leaks.

In Figure 9.26, water spots generated by the water leak in the pipe can be observed. These spots coincide with the turbulence points detected with the geophone.

9.9. Summary and comments

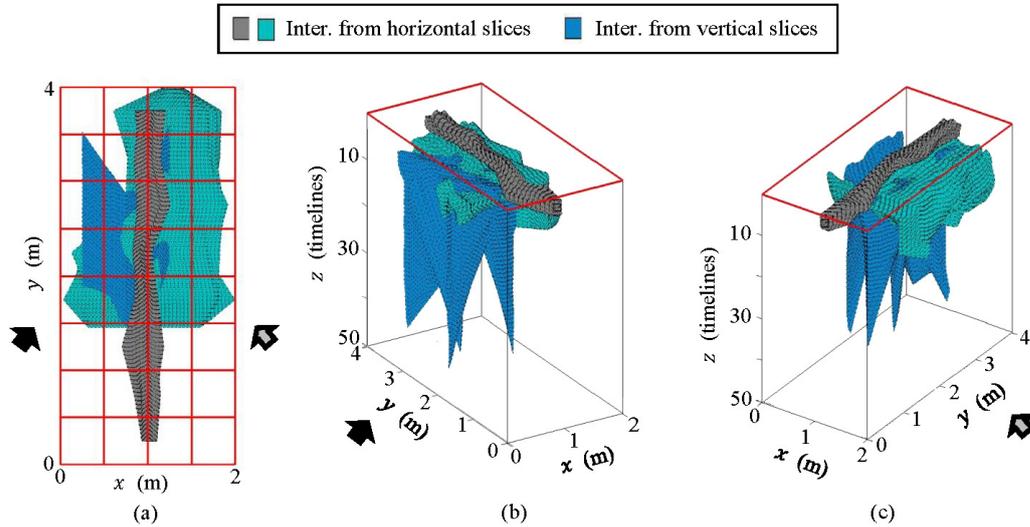


Figure 9.26: Interpretation from pre-processed images, 3D model. (a) azimuthal view, (b) lateral view #1, and (c) lateral view #2. Source: Ayala-Cabrera *et al.* (2013a)

9.9 Summary and comments

By reducing the amount of water leaked, WSS managers can reduce the amount of money and energy spent on producing or purchasing water. The use of various types of smart sensors to gather data and the application of advanced analytics, such as pattern detection, could provide valuable information on the location of leaks in the network. GPR, among other various sensors, can help detect potential leaks and abnormalities within WSSs.

In this chapter, two GPR tests have been conducted in laboratory and field conditions. One test employed a PVC pipe with no water inside and, obviously, without leaks, and the other test used the same pipe with water inside and some leakage. Additionally, we have conducted a field test in a leaking pipe of a real WSS. The different analyses show the difficulty of making interpretations from raw images. Indeed, it follows from these tests that introducing water into the system causes significant (numeric) differences that are not easily discernible from typical interpretations of the raw data obtained.

Moreover, this chapter shows that the application of appropriate pre-processing methodologies facilitates the visualization of features that are not reflected in the raw images. This is true both in the case of laboratory testing and in field cases. Consequently, we argue that the pre-processing used in this chapter facilitates the work of interpretation for non-highly qualified personnel in the use of GPR.

Additionally, it should be mentioned that in this type of pre-process, characteristics are obtained that are quantifiable (numerically). These characteristics can be the basis for further processing for automatic classification of leaks in WSSs.

Finally, from these tests, we can conclude that the interpretations obtained from GPR images are not merely subjective results. This is supported by the fact that there is consistency for all the results in 3D models between what is captured by GPR and what is expected. Thus, the reconstruction of 3D models from GPR analyses facilitates the interpretation of results.

9.9. Summary and comments

Chapter 10

Conclusions and future developments

10.1 Summary

This doctoral work is focused on the design of tools, together with the assessment of their application feasibility, for the characterization of components of WSSs in use, by using non-destructive methods. In this thesis we have used the GPR because its manifest feasibility of detecting objects of various materials buried underground. The idea was to advance in the parametrization of the records obtained from GPR surveys to substantiate the basis for determining features of WSS components through the use of intelligent classification methodologies. The ultimate goal is to create tools that enable effective and efficient performance of the networks integrating the WSS.

The inclusion of advanced technologies in service systems, such as WSSs (as in the case of this thesis), is a challenge today. The data obtained with these new technologies should, in principle, facilitate decision-making processes regarding the actions to be taken over the assets of these systems. There are several tools that promote the study of WSS components without altering the environment (non-destructive methods). However, the large volume of information they generate and the difficulty of interpreting these data minimize the potential use of these tools. In this thesis we have used the GPR as a non-destructive technique that allows suitable visualization of pipes and water leaks in WSSs. However, the interpretation of the data is an additional problem that, as with other non-destructive methods, also applies to the GPR. The main reason is that, usually, the analysis and interpretation of the images obtained with this method (radar-grams) are carried out by non-highly qualified personnel in the use of GPR, since it would be unworkable and dysfunctional that such work would be carried out by skilled staff.

The main assets of WSSs are the pipes, due to their larger impact in relation with other WSS components. Precisely for this reason, tighter control of their paths and characteristics should be applied. The logically obvious idea pervading any work related to WSS management stems on the knowledge of layouts and characteristics of the system. However, the reality is different. The population

10.1. Summary

steady growth, the consequent extension of supply networks, the new continuous and constant maintenance and operational activities needed, coupled with the lack of adequate means for referencing and information updating has produced over time a set of system buried assets integrating a real submerged city, and sometimes a big unknown.

In the course of this work, after a deep analysis of the literature, it has been found that currently records obtained using GPR are analyzed using visualization techniques, and offer many and varied processing methodologies (pre and post-processing) of such records, which improve the quality of the captured images. However, it is noteworthy that these methods have a high dependency on specialist skills for application, presenting great subjectivity in interpretation.

The general proposal of this doctoral study raises the use of GPR images (input) as a tool for the characterization of WSS components, the pipes being the main focus. The main difficulties detected in the course of this study in the case of a single image analysis, according to the review of works by other authors, are:

Difficulties

- The absence of clear, precise and extrapolable protocols for various cases and study conditions that allow to carry out interpretations of GPR images, particularly in the search for buried objects, pipes in our case, so that the interpretation is based on criteria and analyses carried out by highly qualified personnel, what reduces the potential for application of GPR as a non-destructive tool.
- The difficulty of object visualization in the GPR raw images, necessitating interpretation by expert staff. In general, using non-destructive methods, more effective and efficient detection is obtained for metal objects. Although adequate records of other materials, such as plastics (e.g. polyethylene - PE, polyvinyl chloride - PVC), are obtained in the GPR images, visualization is

complex because of the low contrast obtained between the environment and the plastic materials. This work and others such as Ocaña-Levario (2014) have shown the difficulty of visualizing plastic objects in GPR images, resulting in a clear difficulty for visualizing materials commonly employed for WSS pipes, such as polyethylene (PE), PVC, asbestos cement (Fib) and cast iron (Fund). This difficulty may be ordered for these various materials as {PE>PVC>Fib>Fund}.

- Interferences that cause blurring of objects in the GPR images. Aspects such as direct reflection of waves (antenna crosstalk and ground bounce) substantially affect the auto-focus metric of the objects embedded in the GPR images [Wei & Zhang (2015)].
- The problem of excessive noise and disorder (clutter problem) of data in GPR images. The first is due to multiple spurious signal sources, during the formation of the initial wave, because of noise added by external sources (e.g. mobiles), or caused by poor high-frequency electronic design, among others. The second is caused by certain events due to reflections from the initial wave, which damp the signal of the objective reflections [Harrison (2005)]. These problems mask the objects contained in the images, thereby hampering identification. Another noise also typically found in GPR images is the speckle noise, which is given by the random fluctuations of the signal. This type of noise is very critical in GPR image analyses, and can lead to misinterpretation [El-Mahallawy & Hashim (2013)].
- Deformation of the prospected objects due to equipment characteristics, and electromagnetic characteristics of the environment, among others.
- Others.

Among the advantages for WSSs obtained with proper analysis and interpretation of GPR images and, in general, with the inclusion of these (non-destructive) technologies, we can mention

10.1. Summary

Advantages

- Identification, cadastre and control of WSS assets with lower social and economic cost, which provides knowledge of the layouts and characteristics of the components, and can be performed by non-highly qualified staff in GPR image analysis.
- Favoring implementations that improve technical management of WSSs, such as sectorization, design and control optimization, network vulnerability, management of intermittent systems, among others. These activities should obviously be based on detailed and comprehensive knowledge of the components of the system.
- Additional contributions to good technical management of WSSs in such activities as: identification of illegal connections, detection and control of water leaks, simulation and operation of networks, study of the evolution of pollutants in the networks, planning of maintenance activities, rehabilitation and renewal of supply components, applications of geographic information systems (GIS), among others.
- Favoring studies of crucial phenomena, such as water leaks.
- Extrapolation to other systems, such as irrigation, sewer and drain networks.

10.2 Conclusions

This thesis is integrated by a number of approaches based on techniques of intelligent data analysis, such as multi-agent systems, segmentation, clustering, and also other processes raised by the author. The basic principle of this work was to develop suitable data treatment through different methodologies, with the aim of finding stable and repeatable configurations, and to obtain, using the least possible amount of data, useful information for WSSs with respect to their components, with special emphasis on the pipes. The results derived from the conducted research are summarized below, the items corresponding to the respective Chapters.

1. **Transformations into matrices $T14$ and $T15$.** Through various laboratory tests, it has been proposed, designed and proven a method employing the wave amplitude as a differential parameter for the characterization of different objects. This process, called transformation into matrices $T14$ and $T15$ in this work, is based on a procedure that allows the identification of those buried objects contained in GPR images. This procedure consists basically in the identification of shapes and forms, which are delimited by the contrasts in color intensity due to differences in the electromagnetic properties of each of the materials contained in the images, provided that the contrast is sufficient to demarcate such contours. The raw images are subject to a uniform color scale variation so that the objects are highlighted. This makes it possible to extract characteristics in the images, to be subsequently interpreted. This action, although useful, is not easily repeatable in other case studies and can only be performed by highly skilled personnel in the analysis and interpretation of GPR data.

The construction process of the matrices $T14$ and $T15$ uses the matrix of the initial GPR image (raw images) as input and subjects the data to a series of matrix transformations, allowing better visualization of the different objects embedded in the images. These transformations are carried out by the proposed algorithm, so that the analysis is facilitated by not highly qualified

10.2. Conclusions

personnel. Also, these changes provide classified data which can be a good starting point for more robust procedures for automatic detection of objects. Regarding this proposed transformation it can specifically be concluded the following.

- In the proposed processes for building these matrices we take as starting points the minimum (valleys) and maximum (crests) values of the wave amplitude, and study the traces bottom-up and top-down in both cases. As a result, these matrices allow visualization of features that are demarcated more solidly than in the images obtained using a ranking method. These matrices suitably preserve the limits and group the amplitude values so that the respective object material conformation is best shown.
- The amplitude extraction procedures and the compilation of matrices lead, via a cumulative process, to obtain the matrices $T3$, $T5$, $T9$, $T12$, $T14$ and $T15$, as have been termed in this work. These matrices present compactly the characteristics of each material and are a good starting point for consolidated visualization of the information resulting from the conducted surveys. In addition, especially $T14$ and $T15$ serve to parametrize GPR records, since these matrices provide better consolidation of forms.
- The layered approach of matrices $T14$ and $T15$ additionally revealed a natural grouping tendency of data depending on the material. In both case studies, the central and upper part of the hyperbola, and the value of hyperbola development (maximum and minimum layer), clearly showed both a space between plastic elements and another space between these and non-plastic elements clearly demarcated.
- In view of the results obtained for pipes buried in dry soil, as a result of the application of the experimental methodological tool (the transformation into matrices $T14$ and $T15$), it can be stated that, in general, the application of this tool significantly improves the occurrence of features in the images. This improvement is illustrated through the

application of matrices $T3$, $T5$, $T9$ and $T12$, and subsequently, in the consolidation of the images and their visualization achieved with the transformation into the matrices $T14$ and $T15$. Thus, a substantial visualization improvement of the characteristics of buried piping is achieved, which allows visualization of plastic elements (PE and PVC), which are not detectable in raw data. This represents a breakthrough of great interest in the detection of pipes by non-destructive methods for WSSs, because an important part of these networks have elements of these materials installed.

2. **Automatic location of plastic pipes in GPR images.** An automatic detection system of plastic pipes in GPR images based upon transformation matrices $T14$ and $T15$ was implemented. The system was implemented using a multi-agent approach and the conclusions we can draw from this work are the following.

- Viability of automatic identification, by means of a suitable pre-processing, of plastic pipes in GPR images, so that a plastic pipe can be identified not only when there is no other nearby object (simple case), but also when nearby objects (complex case), more easily identifiable (e.g. metal piping materials), distort the image. The proposed approach enable us to clearly obtain the first hyperbola of the plastic pipe.
- The proposed methodology (based on multi-agent systems) is efficient, since the activities are performed by multiple distributed entities, the agents. This optimizes the time and effort of the work, which turns out to be of low computational cost, and allows adaptation to different case studies (pipes of different types of materials). This methodology has proven to be efficient for detection of plastics pipes in GPR images. The configuration of this system consists of three different phases which were designated: (a) roll-on; (b) termite-eating; and (c) final detection. These three phases were performed by three different breeds of agents, respectively. The premise of these three types of agents was to detect images not corresponding to pipes, so that eventually the space of the object prone to be a plastic pipe is clearly delimited.

10.2. Conclusions

- The behavior and configuration of the most successful agents in the detection of plastic pipes (95% in simple cases and 75% in complex cases) was: eight roll-on agents that run the image horizontally, 125 termites for 400 iterations, and 4 square agents for 200 iterations.

3. **Segmentation and cleaning of GPR images.** In this section we sought to produce more understandable images through better views of the objects in the GPR images, by simplifying the image and replacing it with a new image more meaningful and easier to analyze, by grouping data with similar attributes. In this sense, the implementation of segmentation and cleaning proposal developed two important aspects: (a) improvement of visualization of objects in GPR images, and (b) definition of relevant data for automatic detection of objects (pipes in this thesis).

- The use of isolines in the raw GPR images (matrix A) corresponding to controlled assays with soil homogeneity characteristics (in distance) and using objects of plastic and metal materials produced a matrix representation that allows understanding of how registration of objects in GPR images are performed. The approach to this interpretation is based on the achievement of the matrix $A'f$ (Equation 5.1). This matrix represents the behavior of the data as a collection of functions (families), belonging to three different spaces, the medium (fl), the objects (g) and the noise (r), wherein the relationship among these spaces constitutes the initial GPR image, while each space has its own characteristics that differentiate it from each other. The family fl is unique for all the images and is integrated by individuals whose data correspond to straight horizontal lines. Families g are typical of each individual and, in general, are constituted by hyperbolas packed in depth. Each individual of this family springs out from an individual of fl . The noise elements do not exhibit continuity in depth and are lonely individuals without any relation to fl .

- The approach looks for a way of separating the spaces in the matrix $A'f$. To this purpose, we proposed a segmentation methodology based on the Hough transform, to detect and eliminate horizontal straight lines (family fl) and analyze the two remaining spaces ($g + r$). The analysis was based on the values of the wave amplitude, starting from the transformation of raw images given by matrices $T14$ y $T15$. The separation of those spaces through the proposed pre-process improved the visualization of pipes buried in dry soil. In general, after applying this pre-process a considerable visualization improvement is achieved that strongly consolidates the pipe characteristics with respect to the raw images. As a result, various characteristics not observable in the raw images are now apparent (in matrix A'), such as the deformation of the space in depth, what can be used to extract information about the pipe diameter. This approach enables us to observe not only the first hyperbola, but also a whole family of hyperbolas providing a more direct approximation to the object of interest in the image, what allows more in-depth analyses.
- This data classification into families allows various kinds of analysis. For example, by analyzing data densities (amount of data produced by the process divided by the amount of total data), it becomes apparent that the process is consistent, and that there is a certain relationship between the process phases and the objects of interest. For example, in homogeneous conditions in distance, it can be shown that the amount of data necessary to get relevant conclusions from the images is about 8% with respect to the initial amount of data. Also, in this scenario, the noise represents 4% of the data initial volume.
- The density analysis performed under those homogeneity conditions in distance also revealed a clear and natural separation among the various materials, plastic and non-plastic pipes, providing the following qualitative relationship for the density values: {no object > PVC > PE > Fib > Fund}.

10.2. Conclusions

- Using the images obtained after the elimination of the horizontal straight lines, the following facts were obtained. For pipe depth of 0.16 m, the deformation ranged between 2.8 and 3.4 times the outer pipe diameter; this value was 2.8 for PVC, 2.9 for PE, 3.2 for Fib and 3.4 for Fund, what shows a clear dependence among pipe material, depth and object deformation in the GPR image.
- Also, using the density analysis it can be shown that it is possible to bound images for non-homogeneous case in distance. As a result, it is possible, through the division of the original space into smaller images, to bound the space influenced by the families (g), so that this relation allows approximate understandable images (schematic configuration). We have shown, through Equation 5.9 that there exists at least a division into quadrants that allows better visualization of the images.

4. **Automatic generation of pipe layout maps.** We have prepared a tool that enable us to automatically detect WSS pipe layouts from GPR images, after image pre-processing consisting in the segmentation and cleaning processes we have described in this work. The proposed system was based on the idea that the tool is easy to use and enable operators with low experience in GPR to decide upon pipe paths. To this purpose, two approaches were proposed: Approach 1 is devoted to details on suitable profile capturing, and Approach 2 focus on the analysis of already captured profiles.

The main goal of Approach 1 is to study the feasibility of implementing an automatic system for pipe layout detection in real time, stating the necessary conditions and relationships between the two intervening components, namely the own GPR equipment and the capturing protocol to build the proposed system.

Approach 2 addresses automatic detection of pipes, and evaluates the possibility of getting this automatic detection from data previously obtained. The main objective is to produce probability maps that, in addition to visualize pipe layouts, enable us to approach the objects contained in the interest area from where images have been captured. These approach allows

to get other information, for example, related to the diameter of the pipe contained in the explored area.

- After application of both approaches in an urban area, it can be concluded that the joint approach facilitates the analysis and interpretation of the images obtained with GPR, favoring the location of paths of pipes in WSSs. This is achieved because the joint approach generates mapping routes that reduce the time of data capturing, minimize the amount of data to analyze, and clearly define the most relevant data. These results favor more thorough analyses, leaving aside areas with little or low interest, making it feasible the study of WSS pipes by non-destructive methods by not highly qualified personnel in GPR image analysis.
- The selection of the starting point (initial capture) by Latin square sampling provided a good basis for analyzing the data under both approaches, increasing the success of the detection system under the two proposed approaches.
- A traffic-light-like indicator (aci , Equation 6.3), which allows analyzing the data from the various tests performed, was defined. This indicator is able to show the plausibility of existence of the pipe in the tested sector and is built using autoregressive models of order 4, which were analyzed by several assays.
- The importance of generating tools to carry out dynamic management of water supply systems is shown in this work. The fact that these tools are obtained through non-intrusive methods, are as continuous as possible, and use real data is of great importance for these systems, which allows for more dynamic management at different levels. The results obtained with the automatic system of pipe detection constituted by the two described approaches are promising in the sense that the system presents possibilities of adaptation to detect other features that may be of great interest, such as leaks, since the data pre-processing system is improved and allows extraction of patterns to search for these objects.

10.2. Conclusions

5. **Evaluation of WSS components in GPR images through a multi-agent approach.** In this chapter we have developed a data pre-processing approach based on multi-agent systems. The main algorithm termed *Agent race* (Table 7.2) gathers the basic principles of the transformations into matrices T_{14} and T_{15} and, as a result, builds a matrix containing the average values between peaks (maximum and minimum values of the wave amplitude), which are shown as agent movements. The matrix therein obtained closely resembles matrix $A'f$ (Equation 5.1). As a consequence, the previously defined segmentation and cleaning approaches are still valid and applicable to this matrix. As a result, the binarization and refining processes used in that segmentation, and the identification of borders are greatly simplified.

- The multi-agent algorithm proposed reduces the amount of data to be analyzed while the properties of objects in the images are preserved. It is thus obtained that the data reduction is generally higher than 78%, relative to the amount of original data. It was also obtained that data reduction is further related to the amount of captured samples, thus obtaining average reductions of 96% for 512 samples, and 94% to 91% for 1024 and 2048 samples.

6. **Application of unsupervised clustering to pre-processed GPR images.** The possibility of grouping into clusters data from the GPR pre-processed images (using the *Agent race* algorithm) was raised and their viability was studied. A significant reduction in the amount of information to be analyzed is obtained while the reliability thereof is preserved, thus characterizing the soil naturally and quickly according to its composition. This natural classification was possible thanks to the implementation of an unsupervised classification system, the well known Hierarchical agglomerative clustering, favoring the interpretation of GPR files by personnel not highly qualified without the need to assume soil parameters. The results are promising, since the number of points that make up the initial radar-gram are grouped and reduced, while retaining the main properties and

favoring more clear identification of WSS pipes, including better identification of their components. The proposed tool, coupled with the GPR as a non-destructive tool, favors the creation of soil profiles that allow the development of automatic procedures that favor technical management of WSSs.

- The cophenetic correlation coefficient was used as the evaluation parameter, in order to evaluate the relationships among various metric distances and cluster linking methods. The evaluated distance metrics were: (a) Euclidean distance, (b) pseudoeuclidean distance, and (c) cosine distance. The evaluated linking methods were: (a) method of the minimum (single linkage), (b) method of the mean (average linkage), and (c) method of the maximum (complete linkage). Out of the nine possible combinations, precisely those employing the cosine distance gave better results in terms of increased cophenetic coefficient. The two best combinations were cosine distance + method of the mean, and cosine distance + method of the maximum, with respective cophenetic coefficients above 0.93 and 0.95, the latter becoming the best of all combinations, since the domain of the cophenetic coefficient is $[0, 1]$, with 1 being the best value).

7. **WSS leak identification from GPR images.** The problem of water leakage in WSSs is addressed through the use of GPR as a non-destructive method. For this purpose, laboratory trials were performed to extract characteristics of water leaks from the GPR images. Feature extraction is performed by interpreting GPR images, with the support of the pre-processing system based on the *Agent race* algorithm and an approach using contrasts. The approach was applied to a real case, showed its effectiveness to make leaks visible and thereby allowed the detection of both the area affected by the leak and the leaking pipe.

- A contrast technique was used in laboratory tests, to observe and define the characteristics of the leaked water and how this was registered in the GPR images. The contrast consisted in taking profiles for an initial

10.2. Conclusions

time (before the leak) and for a final time (after the leak). To contrast both profiles we calculated $\text{abs}(\text{profile}_{(\text{initial time})} - \text{profile}_{(\text{final time})})$. This approach allowed us to observe the registration of the entire leak process in GPR images. It was also possible to identify the elements integrating the system as well as the point from which water was leaking.

- A variant of the matrix obtained with the *Agent race* pre-processing algorithm was proposed. This variation consisted in ordering (decreasingly) the time data contained in each row, and number them consecutively. Subsequently, the numbered data were normalized. This modification proposed a new scenario that highlighted the contours of the area affected by the water leak and the leaking pipe. After applying this approach, a similar contrast to that performed with the raw images of laboratory tests was obtained. The results showed great similarity with the contrasts performed with the raw data, regarding representation and merging to provide 3D representations in their respective spaces.
 - Since uncontrolled laboratory conditions do not provide an initial image that allows the contrast of images, the pre-processing and its variant was applied to profiles obtained in field test, showing very good results, both for leaked water and leaking pipe. In addition, the results were shown as 3D representations, which allowed improved understanding of the information obtained from 2D images. This clearly favors understanding by not highly skilled personnel in GPR image analysis.
8. To analyze the different tools proposed, several GPR surveys by using a laboratory tank, with just dry soil and with five pipes of four different commercial materials (PVC, PE, Fib, Fund) commonly used in WSSs were performed. Real data from field cases were also obtained. In all the cases we presented the results of these tests in terms of raw data and the analyses performed. When comparing with the raw results we find the following.

- When displaying raw data, using a gray scale, pipes of weakly reflective materials (PVC, PE) present considerable visualization difficulties, for the different cases studied. In the case of PE, as weak wave amplitude values occur in the reflections, boundaries are poorly demarcated. Moreover, pipes Fib and Fund, having high values of wave amplitude compared with plastics, allow high color intensity, thus protruding their resulting forms.
- After analyzing the raw images, it can be said that the visual identification of forms can be realized due to the stark contrast of color intensity of reflective materials, and also with little reflective materials, allowing *a priori* identification of pipes in GPR radargrams. However, this method of identification is based on subjective parameters that do not allow repeatability of the process, resulting in infeasibility of parametrizations leading to classification of WSS components by intelligent systems.

Finally, we indicate that the wave amplitude, which has been a routinely dismissed parameter for use in the processing of records obtained by GPR prospecting, has revealed as a parameter of great interest and use in the identification and classification of buried objects, as we have seen in the development of this work. Such is the case of the procedures outlined in this doctoral work, which are characterized by their simplicity of application and for not depending on expert judgment (not subjective processes). This makes these procedures repeatable and, through appropriate adjustments, offer the possibility of WSS radargram parameterizations aimed to serve as a basis for training intelligent systems.

10.3 Future developments

The methodologies developed and presented in this doctoral work, and the subsequent application in laboratory and field tests, open lines of research leading to suitable classification of WSS components through intelligent classification systems. Specifically, we can mention the following.

- Application of automatic systems for generation of synthetic radargrams, in order to generate different and varied models that simulate various conditions of WSS components, to be compared with simulations in laboratory tests, and also with the aim of generating databases to be employed in classification tasks.
- Use of other statistical tools and feature detection algorithms, designed to serve as a basis for training intelligent classification systems.
- Development of a base to feed more advanced visualization systems, as is the case of augmented reality, serving (once automated) for interpretations of GPR images; these interpretations may provide the visualization systems with the virtual components they need, which can be updated by specific GPR images; this seeks to revitalize technical management of WSSs in such activities as maintenance and implementation of action plans.
- Automatic generation of GPR image interpretation to detect both pipes of different materials and water leaks in GPR images.
- Comparison of the cluster generation method proposed in this document (HAC) with methods adapted, for example, from K-means or support vector clustering to evaluate the effectiveness and reliability of those methods in underground natural classification. The most efficient algorithm will be incorporated into more complex systems of pattern recognition of data in order to generate automatic tools for technical management of WSSs by non-destructive methods like GPR.

10.3. Future developments

- Study of other phenomena that occur in WSSs and are of great importance, such as water leaks through GPR images.
- Finally, implementation, training and activation of an intelligent system for detection and characterization of pipes.

10.3. Future developments

Appendices

Appendix A

File structure: *.dzt

The GPR equipment used in this doctoral work is GSSI 3000. The information is stored as binary data blocks. The data contained in the structure is preceded by a header that contains information about various characteristics of data acquisition, such as location (if a supplementary GPS system is available), system configuration, and some equipment pre-processing parameters. The extension used by the system to store data is *.dzt. The detailed format of this extension is presented in Tables A.1 and A.2. The file header is designed to occupy 1,024 bytes (1 kilobyte) of memory. To facilitate reading this Appendix this information is herein provided. Alternatively, the same information can be found with more detail in Al-Nuaimy (1999) and Lucius & Powers (2002).

Table A.1: *.dzt date structure

	Data type	Field name	Description
{	unsigned	sec2 : 5;	/* second/2 (0-29) /*
	unsigned	min : 6;	/* minute (0-59) /*
	unsigned	hour : 5;	/* hour/2 (0-23) /*
	unsigned	day : 5;	/* day (1-31) /*
	unsigned	month : 4;	/* month (1-12; 1=Jan, 2=Feb, etc.) /*
	unsigned	year : 7;	/* year-1980 (0-127; 1980-2107) /*
}			/* 4 bytes (32 bits) /*

A. FILE STRUCTURE: *.DZT

Table A.2: *.dzt header structure

Data type	Field name	Description
{	unsigned short	h_tag : 1; /* OxO0FF if header */
	unsigned short	h_data : 1; /* offset to data in bytes (1024) */
	unsigned short	h_nsamp : 1; /* samples per scan */
	unsigned short	h_bits : 1; /* bits per data word (8 or 16) */
	short	h_zero : 1; /* binary offset */
	float	h_sps : 1; /* scans per second */
	float	h_spm : 1; /* scans per meter */
	float	h_mpm : 1; /* meters per mark */
	float	h_position : 1; /* position (ns) */
	float	h_range : 1; /* range (ns) */
	unsigned short	h_npass : 1; /* scans per pass for 2D files */
	struct	create_date /* date created (Table A.1) */
	struct	modified_date /* date modified (Table A.1) */
	unsigned short	h_rgain : 1; /* offset to range gain function */
	unsigned short	h_nrgain : 1; /* size of range gain function */
	unsigned short	h_text : 1; /* offset to text */
	unsigned short	h_ntext : 1; /* size of text */
	unsigned short	h_proc : 1; /* offset to processing history */
	unsigned short	h_nproc : 1; /* size of processing history */
	unsigned short	h_nchan : 1; /* number of channels */
	float	h_epsr : 1; /* average dielectric constant */
	float	h_top : 1; /* top position in meters */
	float	h_depth : 1; /* range in meters */
	char	reserved : 31; /* reserved */
	char	h_dtype : 1; /* data type */
	char	h_antena : 14; /* antenna name */
	unsigned short	h_chanmask : 1; /* active channel mask */
	char	h_name : 12; /* initial file name */
	unsigned short	h_checksum : 1; /* checksum for header */
	char	variable; /* range gain, comments & history */
}		/* 1024 bytes if tightly packed */

Appendix B

Equipment configuration for the
performed tests

B. EQUIPMENT CONFIGURATION FOR THE PERFORMED TESTS

Observations	027	029	031	033	035	037	039	041	043	045	047	049	051
File	027	029	031	033	035	037	039	041	043	045	047	049	051
Duplicate file	028	030	032	034	036	038	040	042	044	046	048	050	052
COLLECT													
Radar													
Antenna	1.5 GHz												
Scan													
Samples	1024	2048	512	1024	2048	512	1024	2048	512	1024	2048	512	1024
Format (bits)	16	16	16	16	16	16	16	16	16	16	16	16	16
Range (ns)	20	20	20	20	20	20	20	20	20	20	20	20	20
Diel	3	3	5	5	5	7	7	7	9	9	9	9	9
Rate	32	32	32	32	32	32	32	32	10	10	10	20	20
SCN/UNIT	64	64	64	64	64	64	64	64	64	64	64	64	64
Gain													
Points	Manual												
GP1 (dB)	5	5	5	5	5	5	5	5	5	5	5	5	5
GP2 (dB)	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20
GP3 (dB)	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20
GP4 (dB)	10	10	10	10	10	10	10	10	10	10	10	10	10
GP5 (dB)	20	20	20	20	20	20	20	20	20	20	20	20	20
GP5 (dB)	50	50	50	50	50	50	50	50	50	50	50	50	50
Position													
Offset	Auto												
Surface(%)	55.15	55.15	55.15	55.15	55.15	55.15	55.15	55.15	55.15	55.15	55.15	55.15	55.15
Filters	10	10	10	10	10	10	10	10	10	10	10	10	10
LP_IRR	2500	2500	2500	2500	2500	2500	2500	2500	2500	2500	2500	2500	2500
HP_IRR	200	200	200	200	200	200	200	200	200	200	200	200	200
LP_FIR	0	0	0	0	0	0	0	0	0	0	0	0	0
HP_FIR	0	0	0	0	0	0	0	0	0	0	0	0	0
STACKING	3	3	3	3	3	3	3	3	3	3	3	3	3
BGR_RMVL	0	0	0	0	0	0	0	0	0	0	0	0	0

B. EQUIPMENT CONFIGURATION FOR THE PERFORMED TESTS

Observations	079	081	083	085	087	089	091	093	095	097	099	101	103
File	079	081	083	085	087	089	091	093	095	097	099	101	103
Duplicate file	080	082	084	086	088	090	092	094	096	098	100	102	104
COLLECT													
Radar													
Antenna	1.5 GHz												
Scan													
Samples	512	1024	2048	512	1024	2048	512	1024	2048	512	1024	2048	512
Format (bits)	16	16	16	16	16	16	16	16	16	16	16	16	16
Range (ns)	20	20	20	20	20	20	20	20	20	20	20	20	20
Diel	9	9	9	9	9	9	9	9	9	9	9	9	9
Rate	120	120	120	32	32	32	32	32	32	32	32	32	32
SCN/UNIT	64	64	64	70	70	70	80	80	80	60	60	60	50
Gain													
Points	Manual												
GP1 (dB)	5	5	5	5	5	5	5	5	5	5	5	5	5
GP2 (dB)	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20
GP3 (dB)	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20
GP4 (dB)	10	10	10	10	10	10	10	10	10	10	10	10	10
GP5 (dB)	20	20	20	20	20	20	20	20	20	20	20	20	20
GP5 (dB)	50	50	50	50	50	50	50	50	50	50	50	50	50
Position													
Offset	Auto												
Surface(%)	55.15	55.15	55.15	55.15	55.15	55.15	55.15	55.15	55.15	55.15	55.15	55.15	55.15
Filters	10	10	10	10	10	10	10	10	10	10	10	10	10
LP_IRR	2500	2500	2500	2500	2500	2500	2500	2500	2500	2500	2500	2500	2500
HP_IRR	200	200	200	200	200	200	200	200	200	200	200	200	200
LP_FIR	0	0	0	0	0	0	0	0	0	0	0	0	0
HP_FIR	0	0	0	0	0	0	0	0	0	0	0	0	0
STACKING	3	3	3	3	3	3	3	3	3	3	3	3	3
BGR_RMVVL	0	0	0	0	0	0	0	0	0	0	0	0	0

B. EQUIPMENT CONFIGURATION FOR THE PERFORMED TESTS

Appendix C
Contributions

Publications

Journal papers

1.  Ayala-Cabrera, D.; Herrera, M.; Izquierdo, J.; Pérez-García, R. (2014) GPR data analysis using multi-agent and clustering approaches: a tool for technical management of water supply systems. *Digital Signal Processing*, Vol. 27, pp. 140-149.
2.  Ayala-Cabrera, D.; Campbell, E.; Carreño-Alvarado, E.P; Izquierdo, J.; Pérez-García, R. (2014) Water leakage Evolution Based on GPR interpretations. *Procedia Engineering*, Vol. 89, pp. 304-310.
3.  Ayala-Cabrera, D.; Ocaña-Levario, S.J.; Izquierdo, J.; Pérez-García, R. (2014) Improving water supply system components visualization into GPR images. *Athens Journal of Technology & Engineering*, Vol. 1 (4), pp. 253-261.
4.  Ayala-Cabrera, D.; Izquierdo, J.; Pérez-García, R.; Ocaña-Levario, S.J. (2014) Probabilistic maps for buried pipes location based on GPR images. *International Journal of Complex Systems in Science*, Vol. 4 (1), pp. 27-30.
5.  Ayala-Cabrera, D.; Herrera, M.; Izquierdo, J.; Pérez-García, R.; Ocaña-Levario, S.J. (2013) Dynamic management of water supply systems: A tool to build scenarios by merging GPR surveys and augmented reality. *Water Utility Journal*, Vol. 6 (1), pp. 3-8.
6.  Ayala-Cabrera, D.; Herrera, M.; Izquierdo, J.; Ocaña-Levario, S.J.; Pérez-García, R. (2013) GPR-based water leak models in water distribution systems. *Journal Sensors*, Vol. 2013, pp. 15912-15936.
7.  Ayala-Cabrera, D.; Izquierdo, J.; Montalvo, I.; Pérez-García, R. (2013) Water supply system component evaluation from GPR radargrams using a multi-agent approach. *Mathematical and Computer Modelling*, Vol. 57 (7-8), pp. 1927-1932.

C. CONTRIBUTIONS

8.  Ayala-Cabrera, D.; Herrera, M.; Izquierdo, J.; Pérez-García, R. (2011) Location of buried plastic pipes by multi-agent support based on GPR images. *Journal of Applied Geophysics*, Vol. 75 (4), pp. 679-686.
9.  Ayala-Cabrera, D.; Herrera, M.; Montalvo, I.; Pérez-García, R. (2011) Towards the visualization of water supply system components with GPR images. *Mathematical and Computer Modelling*, Vol. 54 (7-8), pp. 1818-1822.

Chapters of books

1.  Ayala-Cabrera, D.; Izquierdo, J.; Pérez-García, R.; Carreño-Alvarado, E.P.; Ocaña-Levario, S.J.; Campbell, E. (2014) Multi-agent and clustering in data analysis of GPR images, Ch. 23 in *Mathematical Modeling in Social Sciences and Engineering*, Nova Science Publisher Inc. Eds., pp. 241–251 (ISBN:978-1-63117-335-6).
2.  Ayala-Cabrera, D.; Herrera, M.; Izquierdo, J.; Pérez-García, R. (2012) Adaptive mapping routes of pipes in water supply systems using GPR and multi-agent approach, Ch. 3 in *Modelling for Engineering & Human Behaviour 2012*, L. Jódar, Luis Acedo, Juan Carlos Cortés and F. Pedroche, Instituto Universitario de Matemática Multidisciplinar – Universitat Politècnica de València (IMM-UPV) Eds., pp. 14-18 (ISBN:978-84-695-6701-2).
3.  Ayala-Cabrera, D.; Herrera, M.; Montalvo, I.; Pérez-García, R. (2011) Towards the visualization of water supply components with GPR images, Ch. 16 in *Applications of Intelligent Data Analysis in Urban Water Systems*, Universitat Politècnica de València (UPV) Eds., pp. 251–260 (ISBN: 978-84-89487-33-8).
4.  Ayala-Cabrera, D.; Izquierdo, J.; Montalvo, I.; Pérez-García, R. (2011), Water supply system component evaluation from GPR radargrams using a multi-agent approach, Ch. 7 in *Modelling for Engineering & Human*

Behaviour 2011, L. Jódar, Luis Acedo and Juan Carlos Cortés, Instituto Universitario de Matemática Multidisciplinar – Universitat Politècnica de València (IMM-UPV) Eds., pp. 42-45 (ISBN:978-84-695-2143-4).

5. *im²* Ayala-Cabrera, D.; Herrera, M.; Izquierdo, J.; Pérez-García, R. (2010) Towards the visualization of water supply system components with GPR images, Ch. 5 in Modelling for Addictive Behaviour, Medicine and Engineering 2010, L. Jódar, Instituto Universitario de Matemática Multidisciplinar – Universitat Politècnica de València (IMM-UPV) Eds., pp. 17-20 (ISBN:978-84-693-9537-0).
6.  Ayala-Cabrera, D.; Pérez-García, R.; Herrera, M.; Izquierdo, J. (2010) Imágenes de GPR como herramienta para la visualización de tuberías en sistemas de abastecimiento de agua, Ch. 2 in Retos tecnológicos y metodológicos en la gestión técnica de los sistemas urbanos de agua, Universitat Politècnica de València (UPV) Eds., pp. 61–76 (ISBN:978-84-89487-32-1).
7.  Ayala-Cabrera, D.; Pérez-García, R.; Izquierdo, J.; Herrera, M. (2010) Visualización de componentes de sistemas de abastecimiento mediante imágenes de georadar, Ch. 19 in Planificación, proyecto y operación de sistemas de abastecimiento de agua, Universitat Politècnica de València (UPV) Eds., pp. 217-230. (ISBN 978-84-89487-31-4).

Conference papers

2015

1. Ayala-Cabrera, D.; Pérez-García, R.; Izquierdo, J.; Herrera, M. Reconstrucción 3D de elementos de redes de agua mediante la interpretación de imágenes de GPR en un entorno urbano, in XIV Seminario Iberoamericano de Redes de Agua y Drenaje, SEREA 2015 (Sep. 2015, Guanajuato, México).
2. Ayala-Cabrera, D.; Ocaña-Levario, S.J.; Izquierdo, J.; Pérez-García, R. 3D representation of (Buried) water supply elements using pre-processed GPR

C. CONTRIBUTIONS

images, in 3rd Annual International Forum on Water (July 2015, Athens, Greece).

2014

1. Ayala-Cabrera, D.; Herrera, M.; Izquierdo, J.; Pérez-García, R. Representaciones 3D del seguimiento de la evolución de fugas de agua mediante la interpretación de imágenes de georadar, in XIII Simpósio Iberoamericano de Redes de Água, Esgoto e Drenagem, SEREA 2014 (Nov. 2014, Fortaleza, Brazil).
2. Ayala-Cabrera, D.; Campbell, E.; Ocaña-Levario, S.J.; Izquierdo, J.; Pérez-García, R. Reconstrucción 3D de tuberías a partir de mediciones de GPR, in X Seminario Euro Latinoamericano de Sistemas de Ingeniería, X SELASI 2014 (Nov. 2014, La Habana, Cuba).
3. Ayala-Cabrera, D.; Carreño-Alvarado, E.P.; Izquierdo, J.; Pérez-García, R. Visualización de la evolución de fugas de agua en tuberías plásticas mediante imágenes de georadar, in XXVI Congreso Latinoamericano de Hidráulica (August. 2014, Santiago, Chile).
4. Ayala-Cabrera, D.; Ocaña-Levario, S.J.; Izquierdo, J.; Pérez-García, R. Improving water supply system components visualization into GPR images, in 2nd Annual International Forum on Water (July 2014, Athens, Greece).
5. Ayala-Cabrera, D.; Campbell, E.; Carreño-Alvarado, E.P.; Izquierdo, J.; Pérez-García, R. Water leakage evolution based on GPR interpretations, in 16th Water Distribution System Analysis Symposium, WDSA 2014 (July 2014, Bari, Italy).
6. Ayala-Cabrera, D.; Izquierdo, J.; Ocaña-Levario, S.J.; Pérez-García, R. 3D model construction of water supply system pipes based on GPR images, in 7th International Congress on Environmental Modelling and Software, iEMSs 2014 (June 2014, San Diego, USA).

2013

1. Ayala-Cabrera, D.; Ocaña-Levario, S.J.; Pérez-García, R.; Izquierdo, J.; Herrera, M. Visualización de fugas de agua en sistemas de abastecimiento de agua mediante imágenes de GPR, in XII Simposio Iberoamericano sobre Sistemas de Abastecimiento de Agua y Drenaje Urbano, SEREA 2013 (Nov. 2013, Buenos Aires, Argentina).
2. Ayala-Cabrera, D.; Ocaña-Levario, S.J.; Pérez-García, R.; Izquierdo, J.; Herrera, M. Georadar y realidad aumentada: una aproximación a la visualización automática del trazado y las características de las tuberías en los sistemas de abastecimiento de agua, in IX Seminario Euro Latinoamericano de Sistemas de Ingeniería. IX SELASI 2013 (Nov. 2013, La Victoria, Venezuela).
3. Ayala-Cabrera, D.; Herrera, M.; Izquierdo, J.; Pérez-García, R.; Ocaña-Levario, S.J. Dynamic management of water supply systems: A tool to build scenarios by merging GPR surveys and augmented reality, in 8th International Conference of EWRA - Water Resources Management in an Interdisciplinary and Changing Context 2013 (June 2013, Porto, Portugal).

2012

1. Ayala-Cabrera, D.; Herrera, M.; Izquierdo, J.; Pérez-García, R. Catastro de tuberías en sistemas de abastecimiento de agua mediante georadar. Establecimiento adaptativo de rutas de mapeo, in VIII Seminario Euro Latinoamericano de Sistemas de Ingeniería. VIII SELASI 2012 (Nov. 2012, La Habana, Cuba).
2. Ayala-Cabrera, D.; Herrera, M.; Izquierdo, J.; Pérez-García, R. Adaptive mapping routes of pipes in water supply systems using GPR and multi-agent approach, in Mathematical Modelling in Engineering & Human Behaviours 2012 (Sep. 2012, Valencia, Spain).

C. CONTRIBUTIONS

3. Ayala-Cabrera, D.; Herrera, M.; Izquierdo, J.; Pérez-García, R.; Detección de tuberías mediante georadar en sistemas de abastecimiento de agua, in XI Seminario Iberoamericano sobre Sistemas de Abastecimiento de Agua y Drenaje, SEREA 2012 (Jul. 2012, Coímbra, Portugal).

2011

1. Ayala-Cabrera, D.; Pérez-García, R.; Izquierdo, J.; Herrera, M. Análisis de datos de GPR para la visualización de tuberías de sistemas de abastecimiento de agua, in VII Seminario Euro Latinoamericano de Sistemas de Ingeniería. VII SELASI 2011 (Dec. 2011, Valencia, Spain).
2. Ayala-Cabrera, D.; Izquierdo, J.; Montalvo, I.; Pérez-García, R. Water supply system component evaluation from GPR radargrams using a multi-agent approach, in Mathematical Modelling in Engineering & Human Behaviours 2011 (Sep. 2011, Valencia, Spain).
3. Ayala-Cabrera, D.; Pérez-García, R.; Herrera, M.; Izquierdo, J. Segmentación y limpieza de imágenes de GPR de tuberías enterradas, in X Seminario Iberoamericano sobre Planificación, Proyecto y operación de Sistemas de Abastecimiento de Agua, X SEREA (Jan. 2011, Morelia, Mexico).

2010

1. Ayala-Cabrera, D.; Pérez-García, R.; Izquierdo, J.; Herrera, M. Imágenes de GPR como herramienta para la visualización de tuberías en sistemas de abastecimiento de agua, conference in I Ibero-american Seminar of IT-AGUA network - CYTED (Dec. 2010, Valencia, Spain).
2. Ayala-Cabrera, D.; Pérez-García, R.; Herrera, M.; Izquierdo, J. Visualización de tuberías plásticas en sistemas de abastecimiento de agua con imágenes de GPR, in XXIV Congreso Latinoamericano de Hidráulica (Nov. 2010, Punta del Este, Uruguay).

-
3. Ayala-Cabrera, D.; Herrera, M.; Izquierdo, J.; Pérez-García, R. Towards the visualization of water supply system components with GPR images, in *Mathematical Models of Addictive Behaviour, Medicine & Engineering 2010* (Sep. 2010, Valencia, Spain).

2009

1. Ayala-Cabrera, D.; Pérez-García, R.; Izquierdo, J.; Herrera, M. Visualización de componentes de sistemas de abastecimiento mediante imágenes de georadar, in *IX Seminario Iberoamericano sobre Planificación, Proyecto y Operación de Sistemas de Abastecimiento de Agua, IX SEREA* (Nov. 2009, Valencia, Spain).

Poster communications

1. Ayala-Cabrera, D.; Ocaña-Levario, S.J.; Izquierdo, J.; Pérez-García, R.; Herrera, M. Analysis of GPR data through interpretation of pre-processed images obtained by a multi-agent approach to identify pipes in water supply systems, in *Congrés Internacional de l'Associació Catalana d'Intel·ligència Artificial 2013 - CCIA'13* (Oct. 2013, Vic, Spain).
2. Ayala-Cabrera, D.; Pérez-García, R.; Izquierdo, J.; Herrera, M. Aplicación de técnicas de análisis inteligente de datos para el catastro de sistemas de abastecimiento de agua mediante georadar (GPR), in *II Jornadas de Ingeniería del Agua, JIA 2011* (Oct. 2011, Barcelona, Spain).

Visits to research labs

- Microwave and Fiber Optics Laboratory (MFOL) group, Department of Electrical and Computer Engineering, National Technical University of Athens (NTUA), (July – Nov. 2011, Athens, Greece).

C. CONTRIBUTIONS

Prizes and awards

- Finalist with the FluIng-IMM-UPV innovation project: “Updating Water Supply Networks Management (UWSNM)”, in The 17th edition of Innovact Campus Award - The European forum for innovative Start-Ups, (March. 2013, Reims, France).

Invited talks

- Ayala-Cabrera, D.; Ocaña-Levario, S.J.; Pérez-García, R.; Izquierdo, J.; Herrera, M. Métodos no destructivos en la localización de tuberías y la detección de fugas en sistemas de abastecimiento de agua. Caso específico, georadar, invited conference, Universidad del Valle, (2013, Cali, Colombia).
- Ayala-Cabrera, D.; Pérez-García, R. A tool for GPR images visualization of water supply system components, invited conference, WATERPIPE: Come l’innovazione può assistere nell’ottenere la Gestione Economica dei sistemi idrici - ECOMONDO, (2009, Rimini, Italy).

References

- ABBAS, O.A. (2008). Comparisons between data clustering algorithms. *The International Arab Journal of Information Technology*, **5** (3), pp. 320–325. 192
- AHMADI, R., FATHIANPOUR, N. & NOROUZI, G.H. (2015). Detecting physical and geometrical parameters of some common geotechnical targets through their effects on GPR responses. *Arabian Journal of Geosciences*, **8** (7), pp. 4843–4854. 31, 39
- AL-NUAIMY, W. (1999). *Automatic feature detection and interpretation in Ground Penetrating Radar data*, PhD. Thesis, University of Liverpool, Liverpool (UK). 38, 41, 48, 113, 114, 269
- AL-NUAIMY, W., HUANG, Y., NAKHKASH, M., FANG, M.T.C., NGUYEN, V.T. & ERIKSEN, A. (2000). Automatic detection of buried utilities and solid objects with GPR using neural networks and pattern recognition. *Journal of Applied Geophysics*, **43** (2–4), pp. 157–165. 54
- AL-SHAYEA, Q.K. & BAHIA, I.S.H. (2010). Ground penetrating radar slice reconstruction for embedded object in media with target follow. *WSEAS Transactions on Computers*, **9** (5), pp. 496–505. 135
- ALCODEA, R. & MARÍN, I. (2010). Jerarca: efficient analysis of complex networks using hierarchical clustering. *PLoS ONE*, **5** (7): e11585, pp. 01–07. 192
- AYALA-CABRERA, D. (2009). *Caracterización de tuberías enterradas para redes de abastecimiento en servicio mediante el análisis de imágenes obtenidas con*

REFERENCES

- radar de subsuelo (Ground Penetrating Radar – GPR)*, Master Thesis, Universitat Politècnica de València, Valencia (Spain). 51, 60
- AYALA-CABRERA, D., PÉREZ-GARCÍA, R., IZQUIERDO, J. & HERRERA, M. (2009). Visualización de componentes de sistemas de abastecimiento mediante imágenes de georadar. In: *IX SEREA, Seminario Iberoamericano sobre Planificación, Proyecto y Operación de Abastecimiento de Agua*, Valencia (Spain), pp. 1–19. 51
- AYALA-CABRERA, D., PÉREZ-GARCÍA, R., HERRERA, M. & IZQUIERDO, J. (2010). Visualización de tuberías plásticas en sistemas de abastecimiento de agua con imágenes de GPR. In: *XXIV Congreso Latinoamericano de Hidráulica*, Punta del Este (Uruguay). 79
- AYALA-CABRERA, D., HERRERA, M., IZQUIERDO, J. & PÉREZ-GARCÍA, R. (2011a). Location of buried plastic pipes using multi-agent support based on GPR images. *Journal of Applied Geophysics*, **75** (4), pp. 679–686. 18, 60, 79, 81, 84, 86, 87, 90, 91
- AYALA-CABRERA, D., HERRERA, M., MONTALVO, I. & PÉREZ-GARCÍA, R. (2011b). Towards the visualization of water supply system components with GPR images. *Mathematical and Computer Modelling*, **54** (7–8), pp. 1818–1822. 18, 26, 51, 59, 71
- AYALA-CABRERA, D., IZQUIERDO, J., MONTALVO, I. & PÉREZ-GARCÍA, R. (2011c). Water supply system component evaluation from GPR radargrams using a multi-agent approach. In: *Mathematical Modelling in Engineering & Human Behaviours 2011*, Valencia (Spain), pp. 42–45. 167
- AYALA-CABRERA, D., PÉREZ-GARCÍA, R., HERRERA, M. & IZQUIERDO, J. (2011d). Segmentación y limpiado de imágenes de GPR de tuberías enterradas. In: *X SEREA, Seminario Iberoamericano de Planificación, Proyecto y Operación de Abastecimiento de Agua*, Morelia (Mexico), pp. 55–66. 98
- AYALA-CABRERA, D., PÉREZ-GARCÍA, R., IZQUIERDO, J. & HERRERA, M. (2011e). Análisis de datos de GPR para la visualización de tuberías de sistemas

REFERENCES

- de abastecimiento de agua. In: *VII SELASI, Seminario Euro Latinoamericano de Sistemas de Ingeniería*, Valencia (Spain), pp. 31–42. 189
- AYALA-CABRERA, D., HERRERA, M., IZQUIERDO, J. & PÉREZ-GARCÍA, R. (2012a). Adaptive mapping routes of pipes in water supply systems using GPR and multi-agent approach. In: *Mathematical Modelling in Engineering & Human Behaviours 2012*, Valencia (Spain), pp. 14–18. 131
- AYALA-CABRERA, D., HERRERA, M., IZQUIERDO, J. & PÉREZ-GARCÍA, R. (2012b). Detección de tuberías mediante georadar en sistemas de abastecimiento de agua. Establecimiento adaptativo de rutas de inspección. In: *XI SEREA, Seminário Ibero-Americano sobre Sistemas de Abastecimento e Drenagem*, Coímbra (Portugal). 131
- AYALA-CABRERA, D., HERRERA, M., IZQUIERDO, J., OCAÑA-LEVARIO, S.J. & PÉREZ-GARCÍA, R. (2013a). GPR-Based water leak models in water distribution systems. *Sensors*, **13** (12), pp. 15912–15936. 18, 171, 212, 217, 218, 219, 220, 221, 222, 224, 225, 227, 228, 229, 230, 231, 232, 234, 235, 236, 237, 238, 239, 240, 241, 242
- AYALA-CABRERA, D., IZQUIERDO, J., MONTALVO, I. & PÉREZ-GARCÍA, R. (2013b). Water supply system component evaluation from GPR radargrams using a multi-agent approach. *Mathematical and Computer Modelling*, **57** (7–8), pp. 1927–1932. 18, 167, 181, 182, 183
- AYALA-CABRERA, D., OCAÑA-LEVARIO, S.J., IZQUIERDO, J., PÉREZ-GARCÍA, R. & HERRERA, M. (2013c). Analysis of GPR data through interpretation of pre-processed images obtained by a multi-agent approach to identify pipes in water supply systems. In: *Congrés International de l'Associació Catala d'Intelligència Artificial*, Vic (Spain), *Frontiers in Artificial Intelligence and Applications*, **256**, pp. 253–256. 172, 191
- AYALA-CABRERA, D., OCAÑA-LEVARIO, S.J., PÉREZ-GARCÍA, R., IZQUIERDO, J. & HERRERA, M. (2013d). Visualización de fugas de agua en sistemas de abastecimiento de agua mediante imágenes de GPR. In:

REFERENCES

- XII SEREA, Simposio Iberoamericano sobre Sistemas de Abastecimiento de Agua y Drenaje Urbano*, Buenos Aires (Argentine). 212
- AYALA-CABRERA, D., HERRERA, M., IZQUIERDO, J. & PÉREZ-GARCÍA, R. (2014a). GPR data analysis using multi-agent and clustering approaches: a tool for technical management of water supply systems. *Digital Signal Processing*, **27**, pp. 140–149. 18, 175, 189, 193, 200, 202, 203, 204, 206
- AYALA-CABRERA, D., IZQUIERDO, J., PÉREZ-GARCÍA, R. & OCAÑA-LEVARIO, S.J. (2014b). Probabilistic maps for buried pipes location based on GPR images. *International Journal of Complex Systems in Science*, **4** (1), pp. 27–30. 131
- BAILEY, R.A. (2008). *Design of comparative experiments*, Cambridge University Press, Cambridge (UK). 141
- BELOTTI, V., DELL'ACQUA, F. & GAMBA, P. (2002). Denoising and hyperbola recognition in GPR data. In: *VII Image and Signal Processing for Remote Sensing*, Toulouse (France), SPIE, **4541**, pp. 1–11. 39, 40
- BENÍTEZ, J., DELGADO-GALVÁN, X., IZQUIERDO, J. & PÉREZ-GARCÍA, R. (2015). Consistent completion of incomplete judgments in decision making using AHP. *Journal of Computational and Applied Mathematics*, **290**, pp. 412–422. 8
- BIMPAS, M., AMDITIS, A. & UZUNOGLU, N. (2010). Detection of water leaks in supply pipes using continuous wave sensor operating at 2.45 GHz. *Journal of Applied Geophysics*, **70** (3), pp. 226–236. 215
- BIMPAS, M., AMDITIS, A. & UZUNOGLU, N.K. (2011). Design and implementation of an integrated high resolution imaging ground penetrating radar for water pipe rehabilitation. *Water Resources Management*, **25** (4), pp. 1239–1250. 3
- BIRKENFELD, S. (2010). Automatic detection of reflexion hyperbolas in GPR data with neural networks. In: *2010 World Automation Congress (WAC)*, Kobe (Japan), IEEE Conference Publications, pp. 1–6. 55

REFERENCES

- BONOMO, N. & DE LA VEGA, M. (2010). El método de georadar. Available online URL <http://www.gaia.df.uba.ar/Radar.pdf> (last accessed in January 2010). 26, 28
- BRÁS, C.S. (2012). Levantamento e atualização da informação cadastral dos sistemas de abastecimento de água e drenagem de água residuais no concelho de coimbra. uma perspectiva de sustentabilidade para as gerações vindouras. In: *XI SEREA, Seminário Ibero-Americano sobre Sistemas de Abastecimento e Drenagem*, Coímbra (Portugal). 4
- BUSQUET, E., CASAS, A., PINTO, V., RIVERO, L. & SABADÍA, J.B. (1996). Geofísica ambiental: técnicas no destructivas para el reconocimiento de zonas contaminadas por vertidos. *Acta Geológica Hispánica*, **30** (1–3), pp. 73–82. 10, 25, 28, 43
- CAMPBELL, E., AYALA-CABRERA, D., IZQUIERDO, J., PÉREZ-GARCÍA, R. & TAVERA, M. (2014). Water supply network sectorization based on social networks community detection algorithms. *Procedia Engineering*, **89**, pp. 1208–1215. 8
- CARR, D.B., YOUNG, C.J., ASTER, R.C. & ZHANG, X. (1999). Cluster analysis for CTBT seismic event monitoring. In: *21st Seismic Research Symposium: Technologies for Monitoring the CTBT*, Las Vegas, NV (USA). 197
- CATAPANO, I., AFFINITO, A., BERTOLLA, L., PORSANI, J.L. & SOLDOVIERI, F. (2014). Oil spill monitoring via microwave tomography enhanced GPR surveys. *Journal of Applied Geophysics*, **108**, pp. 95–103. 27
- CHAN, T., SO, H. & HO, K. (2007). A study on two-sided linear prediction approach for land mine detection. In: *15th European Signal Processing Conference*, Poznan (Poland), pp. 2414–2418. 53
- CHARLTON, M.B. & MULLIGAN, M. (2001). Efficient detection of mains water leaks using ground penetrating radar. In: *III Subsurface and Sensing Technologies and Applications*, San Diego, CA (USA), SPIE, **4491**, pp. 375–386. 215

REFERENCES

- CHEN, H. & COHN, A.G. (2011). Buried utility pipeline mapping based on multiple spatial data sources: a bayesian data fusion approach. In: *IJCAI'11, 22nd International Joint Conference on Artificial Intelligence*, Barcelona (Spain), pp. 2411–2417. 135
- CIST, D.B. & SCHUTZ, A.E. (2001). State of the art for pipe & leak detection DE-FC26-01NT41317. A low-cost GPR gas pipe & leak detector. *Geophysical survey systems, Inc.* 5
- CIU, Y., WANG, L. & XIAO, J. (2010). Automatic feature recognition for GPR image processing. *World Academy of Science, Engineering and Technology*, **61**, pp. 176–179. 55
- CROCCO, L., SOLDVIERI, F., MILLINGTON, T. & CASSIDY, N.J. (2010). Bistatic tomographic GPR imaging for incipient pipeline leakage evaluation. *Progress In Electromagnetics Research*, **101**, pp. 307–321. 5, 215
- CUKAVAC, M., KLEMČIĆ, G. & Č. LAZOVIĆ (2008). Reconstruction of buried objects by implementation of ground penetrating radar technique: example on Roman tomb in Brestovik (Serbia). In: *International Conference on Geoarchaeology and Archaeomineralogy*, Sofia (Bulgaria), pp. 333–338. 53
- DE LA VEGA, M. (2001). *Radar de subsuelo. Evaluación para aplicaciones en arqueología y en patrimonio histórico-artístico*, PhD Thesis, Universitat Politècnica de Catalunya, Barcelona (Spain). 32
- DELBO, S., GAMBA, P. & ROCCATO, D. (2000). A fuzzy shell clustering approach to recognize hyperbolic signatures in subsurface radar images. *IEEE Transactions on Geoscience and Remote Sensing*, **38** (3), pp. 1447–1451. 55
- DELL'ACQUA, A., SARTI, A., TUBARO, S. & ZANZI, L. (2004). Detection of linear objects in GPR data. *Signal Processing*, **84** (4), pp. 785–799. 135
- DEMIRCI, S., YIGIT, E., ESKIDEMIR, I.H. & OZDEMIR, C. (2012). Ground penetrating radar imaging of water leaks from buried pipes based on back-projection method. *NDT&E International*, **47**, pp. 35–42. 215

- DONG, L., CARNALLA, S. & SHINOZUKA, M. (2011). Experimental and analytical study of detecting leakage of water pipes using ground-penetrating radar. In: *9th International Workshop on Remote Sensing for Disaster Response*, Stanford, CA (USA), pp. 1–10. 215
- DOOLITTLE, J.A. & BELLANTONI, N.F. (2010). The search for graves with ground-penetrating radar in Connecticut. *Journal of Archaeological Science*, **37** (5), pp. 941–949. 5
- DUTTA, R., COHN, A.G. & MUGGLETON, J.M. (2013). 3D mapping of buried underworld infrastructure using dynamic bayesian network based multi-sensory image data fusion. *Journal of Applied Geophysics*, **92**, pp. 8–19. 135
- EL-MAHALLAWY, M.S. & HASHIM, M. (2013). Material classification of underground utilities from GPR images using DCT-based SVM approach. *IEEE Geoscience and Remote Sensing Letters*, **10** (6), pp. 1542–1546. 40, 249
- FISHER, R.A. (1926). The arrangement of field experiments. *Journal of the Ministry of Agriculture of Great Britain*, **33**, pp. 503–513. 141
- GADER, P.D., NELSON, B.N., FRIGUI, H., VAILLETTE, G. & KELLER, J.M. (2000). Fuzzy logic detection of landmines with ground penetrating radar. *Signal Processing*, **80** (6), pp. 1069–1084. 53
- GAMBA, P. & LOSSANI, S. (2000). Neural detection of pipe signatures in ground penetrating radar images. *IEEE Transactions on Geoscience and Remote Sensing*, **38** (2), pp. 790–797. 55, 69, 88
- GHASEMIGOL, M., SADOOGHI, H. & MONSEFI, R. (2010). A new hierarchical clustering algorithm on fuzzy data (FHCA). *International Journal of Computer and Electrical Engineering*, **2** (1), pp. 134–140. 194
- GIANETTI, L., MATURANA, F.P. & DISCENZO, F.M. (2005). *Agent-based control of a municipal water system*, Lecture Notes in Computer Science, Springer-Verlag, pp. 500–510. 82

REFERENCES

- GÖTTSCHE, F.M. (1997). *Identification of cavities by extraction of characteristic parameters from ground probing radar reflection data*, PhD Thesis, Christian-Albrechts Universität, zu Kiel (Germany). 26, 37
- GROWINGBLUE (2013). Available online URL <http://growingblue.com/wp-content/uploads/2011/04/Growing-Blue.pdf> (last accessed in 29 Sept. 2015). 214
- GURBUZ, A.C., MCCLELLAN, J.H. & SCOTT, W.R. (2010). Detection of linear and planar structures in 3D subsurface images by iterative dimension reduction. *Digital Signal Processing*, **20** (2), pp. 391–400. 135
- GUTIÉRREZ-PÉREZ, J.A., HERRERA, M., PÉREZ-GARCÍA, R. & RAMOS-MARTÍNEZ, E. (2013). Application of graph-spectral methods in the vulnerability assessment of water supply networks. *Mathematical and Computer Modelling*, **57** (7–8), pp. 1853–1859. 8
- HAO, T., ROGERS, C.D.F., METJE, N., CHAPMAN, D.N., MUGGLETON, J.M., FOO, K.Y., WANG, P., PENNOCK, S.R., ATKINS, P.R., SWINGLER, S.G., PARKER, J., COSTELLO, S.B., BURROW, M.P.N., ANSPACH, J.H., ARMITAGE, R.J., COHN, A.G., GODDARD, K., LEWIN, P.L., ORLANDO, G., REDFERN, M.A., ROYAL, A.C.D. & SAUL, A.J. (2012). Condition assessment of the buried utility service infrastructure. *Tunnelling and Underground Space Technology*, **28**, pp. 331–344. 214
- HARRISON, D. (2005). *Threshold estimation using wavelets and curvelets on ground penetrating radar data for noise and clutter suppression*, Master Thesis, University of British Columbia, South Dakota (USA). 40, 41, 249
- HARTIGAN, J.A. & WONG, M.A. (1979). Algorithm AS136: a K-means clustering algorithm. *Applied Statistics*, **28**, pp. 100–108. 194
- HASAN, A.E. (2012). The use of ground penetrating radar with a frequency 1 GHz to detect water leaks from pipelines. In: *IWTC16, 16th International Water Technology Conference*, Istanbul (Turkey), pp. 1–6. 216

REFERENCES

- HASTIE, T., TIBSHIRANI, R. & FRIEDMAN, J. (2001). *The elements of statistical learning*, Springer New York. 194
- HE, X.Q., ZHU, Z.Q., LIU, Q.Y. & LU, G.Y. (2009). Review of GPR rebar detection. In *PIERS 2009, Progress In Electromagnetics Research Symposium*, Beijing (China), pp. 804–813. 53
- HERRERA, M., IZQUIERDO, J., PÉREZ-GARCÍA, R. & AYALA-CABRERA, D. (2010). Water supply clusters by multi-agent based approach. In: *WDSA2010, Water Distribution System Analysis*, Tucson, Arizona (USA), pp. 861–869. 83, 170
- HERRERA, M., IZQUIERDO, J., PÉREZ-GARCÍA, R. & AYALA-CABRERA, D. (2012). El sistema de abastecimiento de agua en ciudades inteligentes. In: *XI SEREA, Seminário Ibero-Americano sobre Sistemas de Abastecimiento e Drenagem*, Coímbra (Portugal). 9
- HERRERA, M., RAMOS-MARTÍNEZ, E., IZQUIERDO, J. & PÉREZ-GARCÍA, R. (2015). Graph constrained label propagation on water supply networks. *AI Communications*, **28** (1), pp. 47–53. 8
- HUNDAIDI, O. & GIAMOU, P. (1998). Ground-penetrating radar for detection of leaks in buried plastic water distribution pipes. In: *7th International Conference on Ground Penetrating Radar*, Lawrence, Kansas (USA), pp. 783–786. 215
- HYUN, S.Y., JO, Y.S., OH, H.C., KIM, S.Y. & KIM, Y.S. (2007). The laboratory scaled-down model of a ground-penetrating radar for leak detection of water pipes. *Measurement Science and Technology*, **18** (9), pp. 2791–2799. 216
- ILAYA-AYZA, A.E., CAMPBELL, E., PÉREZ-GARCÍA, R. & IZQUIERDO, J. (2014). Optimización multiobjetivo de horarios de suministro en sistemas de abastecimiento intermitente. In: *XIII SEREA, Simpósio Iberoamericano de Redes de Água, Esgoto e Drenagem*, Fortaleza (Brasil). 8

REFERENCES

- IZQUIERDO, J., HERRERA, M., MONTALVO, I. & PÉREZ-GARCÍA, R. (2009). Agent-based division of water distribution systems into district metered areas. In: *4th International Conference on Software and Data Technologies*, Sofia (Bulgaria). 82
- IZQUIERDO, J., MONTALVO, I., HERRERA, M. & PÉREZ-GARCÍA, R. (2010a). Multi-agent applications in urban hydraulics. In: *16th European Conference on Mathematics for Industry*, Wuppertal (Germany). 82
- IZQUIERDO, J., MONTALVO, I., PÉREZ-GARCÍA, R. & IZQUIERDO, F.J. (2010b). Hydraulic transient simulation using a multi-agent based approach. In: *WDSA2010, Water Distribution System Analysis*, Tucson, Arizona (USA), pp. 96–111). 82
- IZQUIERDO, J., CAMPBELL, E., MONTALVO, I. & PÉREZ-GARCÍA, R. (2015). Injecting problem-dependent knowledge to improve evolutionary optimization search ability. *Journal of Computational and Applied Mathematics*, **In press**. 8
- JANNING, R., BUSCHE, A., HORVÁTH, T. & SCHMIDT-THIEME, L. (2014). Buried pipe localization using an iterative geometric clustering on GPR data. *Artificial Intelligence Review*, **42** (3), pp. 403–425. 39, 40, 42, 191
- JOL, H.M. (2009). *Ground penetrating radar: Theory and applications*, ELSEVIER, Amsterdam (Netherlands). 30, 34, 108
- JULIANO, T., MEEGODA, J. & WATTS, D. (2013). Acoustic emission leak detection on a metal pipeline buried in sandy soil. *Journal Pipeline Systems Engineering and Practice*, **4** (3), pp. 149–155. 215
- LAMPROPOULOS, A.S. & TSIHRINTZIS, G.A. (2004). Agglomerative hierarchical clustering for musical database visualization and browsing. In: *3rd Hellenic Conference on Artificial Intelligence*, Samos Island (Greece), pp. 5–8. 196

- LANE, J.W., BUURSINK, M.L., HAENI, F.P. & VERSTEEG, R.J. (2000). Evaluation of ground-penetrating radar to detect free-phase hydrocarbons in fractured rocks—results of numerical modeling and physical experiments. *Ground Water*, **38** (6), pp. 929–938. 29
- LEYTON-BROWN, K. & SHOHAM, Y. (2008). *Essentials of game theory: A concise, multidisciplinary introduction (synthesis lectures on artificial intelligence and machine learning)*, Morgan and Claypool Publishers, San Francisco, CA (USA). 171
- LI, S., CAI, H. & KAMAT, V.R. (2015). Uncertainty-aware geospatial system for mapping and visualizing underground utilities. *Automation in Construction*, **53**, pp. 105–119. 23, 24, 28, 29, 33
- LIU, C. & SHEN, L.C. (1991). Numerical simulation of subsurface radar for detecting buried pipes. *IEEE Transactions on Geoscience and Remote Sensing*, **29** (5), pp. 795–798. 54, 55
- LIU, G., JIA, Y., LIU, H., QIU, H., QIU, D. & SHAN, H. (2002). A case study to detect the leakage of underground pressureless cement sewage water pipe using GPR, electrical and chemical data. *Environmental Science & Technology*, **36** (5), pp. 1077–1085. 215
- LOMUSCIO, A. & RAIMONDI, F. (2006). Model checking knowledge, strategies, and games in multi-agent systems. In: *5th International Joint Conference on Autonomous Agents and Multiagent Systems*, Hakodate (Japan), pp. 161–168. 174
- LOULIZI, A. (2001). *Development of Ground penetrating radar signal modeling and implementation for transportation infrastructure assessment*, PhD. Thesis, Virginia Polytechnic Institute and State University, Blacksburg, Virginia (USA). 54
- LU, Q., PU, J. & LIU, Z. (2014). Feature extraction and automatic material classification of underground objects from ground penetrating radar data. *Journal of Electrical and Computer Engineering*, **2014** ID 347307, pp. 1–10. 38, 41

REFERENCES

- LUCIUS, J. & POWERS, M.H. (2002). GPR data processing computer software for the PC. *USGS U.S. Geological Survey*, Report 02–166. 269
- MALVAR, L.J. & CLINE, G.D. (2004). Detecting voids under asphalt and concrete pavements. *Transportation Systems, Fort Lauderdale, FL*, Report. 29
- MARTÍNEZ, R.M. & RODRÍGUEZ, I. (2011). Comparación de la anisotropía dieléctrica de la madera en las coníferas y frondosas por medio de la técnica no destructiva del georadar. Bachelor Thesis, Universitat Pòlitécnica de València, Valencia (Spain). 32
- MONTALVO, I., IZQUIERDO, J., CAMPBELL, E. & PÉREZ-GARCÍA, R. (2014a). Cloud-based decision making in water distribution systems. *Procedia Engineering*, **89**, pp. 488–494. 8
- MONTALVO, I., IZQUIERDO, J., HERRERA, M. & PÉREZ-GARCÍA, R. (2014b). Water supply system computer-aided design by agent swarm optimization. *Computer-Aided Civil and Infrastructure Engineering*, **29** (6), pp. 433–448. 8
- NAGASHREE, R.N., ASWINI, N., DYANA, A. & SRINIVAS, C.H. (2014). Detection and classification of ground penetrating radar image using textural features. In: *International Conference on Advances in Electronics, Computers and Communications*, Bangalore (India), pp. 1–5. 40, 41
- NEUMANN, J. & MORGENSTERN, O. (1947). *Theory of games and economic behavior*, Princeton University Press. 171, 192
- OCAÑA-LEVARIÓ, S.J. (2014). *Reconocimiento de patrones para la identificación de tuberías enterradas en sistemas de abastecimiento de agua a partir de imágenes de GPR*, Master Thesis, Universitat Politècnica de València, Valencia (Spain). 249
- OCAÑA-LEVARIÓ, S.J., AYALA-CABRERA, D., PÉREZ-GARCÍA, R., IZQUIERDO, J. & HERRERA, M. (2013). Reconocimiento de patrones para la localización y caracterización de elementos plásticos de los sistemas de

REFERENCES

- abastecimiento de agua mediante GPR. In: *XII SEREA, Simposio Iberoamericano sobre Sistemas de Abastecimiento de Agua y Drenaje Urbano*, Buenos Aires (Argentine). 172, 191
- OLHOEFT, G.R. (2000). Maximizing the information return from ground penetrating radar. *Journal of Applied Geophysics*, **43** (2–4), pp. 175–187. 54
- PASOLLI, E., MELGANI, F. & DONELLI, M. (2008). Automatic detection and classification of buried objects in GPR images using genetic algorithms and support vector machines. In: *IGARSS 2008, International Geoscience and Remote Sensing Symposium*, Boston, MA (USA), pp. 525–528. 55
- PASOLLI, E., MELGANI, F. & DONELLI, M. (2009a). Automatic analysis of GPR images: A pattern-recognition approach. *IEEE Transactions on Geoscience and Remote Sensing*, **47** (7), pp. 2206–2217. 5, 55
- PASOLLI, E., MELGANI, F. & DONELLI, M. (2009b). A pattern recognition system for extracting buried object characteristics in GPR images. In: *IGARSS 2009, International Geoscience and Remote Sensing Symposium*, Cape Town (South Africa), pp. 430–433. 53, 55
- PETRINOVIC, D. (2009). Continuous time domain properties of causal cubic splines. *Signal Processing*, **89** (10), pp. 1941–1958. 108
- PUEYO, Ó., ZAMORA, S.A. & POCOVÍ, A. (2005). Aplicación de la prospección geofísica por georradar (GPR) a la delimitación y caracterización de estructuras arqueológicas. Ejemplo de las ruinas del convento Agustino de Fraga (Huesca). *Geogaceta*, **38**, pp. 131–134. 25, 26, 27, 28, 43
- QUEIROZ, F.A.A., VIEIRA, D.A.G., TRAVASSOS, X.L. & PANTOJA, M.F. (2012). Feature extraction and selection in ground penetrating radar with experimental data set of inclusions in concrete blocks. In: *11th International Conference on Machine Learning and Applications*, Boca Raton, Fla (USA), pp. 48–53. 27, 29

REFERENCES

- RAHKAR-FARSHI, T., KESEMEN, O. & BEHJAT-JAMAL, S. (2014). Multi hyperbole detection on images using modified artificial bee colony (ABC) for multimodal function optimization. In: *22nd Signal Processing and Communications Applications Conference*, Trabzon (Turkey), pp. 894–898. 114
- RAMOS-MARTÍNEZ, E., HERRERA, M., IZQUIERDO, J. & PÉREZ-GARCÍA, R. (2014). Ensemble of naïve bayesian approaches for the study of biofilm development in drinking water distribution systems. *International Journal of Computer Mathematics*, **91** (1), pp. 135–146. 8
- RIZO, P. (2010). Water and wastewater pipe health monitoring: a review. *Advances in Civil Engineering*, **2010** ID 818597, pp. 1–13. 9, 24
- ROBINSON, M., BRISTOW, C., MCKINLEY, J. & RUFFELL, A. (2013). *Ground penetrating radar*, Part 1, Sec: 5.5. In Clarke, L.E & Nield, J.M. (Eds.) *Geomorphological Techniques* (online Edition). British Society for Geomorphology, London (UK), pp. 1–26. 25, 38, 39, 40, 41, 42, 178
- RODRIGUEZ, J.B., PANTOJA, M.F., TRAVASSOS, X.L., VIERA, D.A.G. & SALDANHA, R.R. (2015). A prediction algorithm for data analysis in GPR-based surveys. *Neurocomputing*, **168**, pp. 464–474. 27
- ROYAL, A.C.D., ATKINS, P.R., BRENNAN, M.J., CHAPMAN, D.N., CHEN, H., COHN, A.G., FOO, K.Y., GODDARD, K.F., HAYES, R., HAO, T., LEWIN, P.L., METJE, N., MUGGLETON, J.M., NAJI, A., ORLANDO, G., PENNOCK, S.R., REDFERN, M.A., SAUL, A.J., SWINGLER, S.G., WANG, P. & ROGERS, C.D.F. (2011). Site assessment of multiple-sensor approaches for buried utility detection. *International Journal of Geophysics*, **2011** ID 496123, pp. 1–19. 4, 134
- RUSSELL, B.C., SIVIC, J., PONCE, J. & DESSALES, H. (2011). Automatic alignment of paintings and photographs depicting a 3D scene. In: *3rd International IEEE Workshop on 3D Representation and Recognition*, Barcelona (Spain), pp. 1–8. 236

- SCHROTT, L., OTTO, J.C., GÖTZ, J. & GEILHAUSEN, M. (2013). *Fundamental classic and modern field techniques in geomorphology: An overview*, Ch. 14.2. in: Shroder, J. (Editor in Chief), Switzer, A.D., Kennedy, D.M. (Eds.), *Treatise on Geomorphology*. Academic Press, San Diego, CA (USA), pp. 6–21. 5, 22, 26, 31, 33
- SHAO, W., BOUZERDOUM, A. & PHUNG, S.L. (2013). Sparse representation of GPR traces with application to signal classification. *IEEE Transactions on Geoscience and Remote Sensing*, **51** (7), pp. 3922–3930. 191
- SHAW, M.R., MILLARD, S.G., MOLYNEAUX, T.C.K., TAYLOR, M.J. & BUNGEY, J.H. (2005). Location of steel reinforcement in concrete using ground penetrating radar and neural networks. *NDT & E International*, **38** (3), pp. 203–212. 53
- SHOHAM, Y. & LEYTON-BROWN, K. (2009). *Multiagents systems: algorithmic, Game-Theoretic, and logical foundations*, Cambridge University Press, Cambridge (UK). 82, 174
- SIMI, A., BRACCIALI, S. & MANACORDA, G. (2008). Hough transform based automatic pipe detection for array GPR: algorithm development and on-site test. In: *RADAR'8, IEEE Radar Conference*, Rome (Italy). 38, 54, 114
- SINGH, N.P. & NENE, M.J. (2013). Buried object detection and analysis of GPR images: using neural network and curve fitting. In: *International Conference on Microelectronics, Communications and Renewable Energy*, Kanjirapally (India), pp. 1–6. 10, 38, 40, 41, 42, 44, 45, 46
- SOLLA, M., LORENZO, H., RIAL, F.I. & NOVO, A. (2012). Ground-penetrating radar for the structural evaluation of masonry bridges: Results and interpretational tools. *Construction and Building Materials*, **29**, pp. 458–465. 27
- STAMPOLIDIS, A., SOUPIOIS, P., VALLIANATOS, F. & TSOKAS, G.N. (2003). Detection of leaks in buried plastic water distribution pipes in urban places – a case study. In *2nd International Workshop on Advanced Ground Penetrating Radar*, Delft (Netherlands), pp. 120–124. 215

REFERENCES

- SYCARA, K.P. (1998). Multiagent systems. *American Association for Artificial Intelligence – AI Magazine*, **2** (2), pp. 79–92. 82
- SZÉP, J. & FORGÓ, F. (1985). *Introduction to the theory of games*, Mathematics and Its Applications, Springer Netherlands, pp. 230–236. 173
- TAKAHASHI, K. (2006). *Detection and localization of subsurface objects by ground penetrating radar*, PhD. Thesis, Tohoku University, Sensai (Japan). 5
- TAVERA, M. (2008). *Aplicación de georadar para la mejora del rendimiento de una red hídrica*, Master Thesis, Universitat Politècnica de València, Valencia (Spain). 216
- TAVERA, M. (2013). *Metodología para la gestión y planificación de un sistema de agua potable con suministro intermitente: Aplicación a la ciudad de Tegucigalpa (Honduras)*, PhD Thesis, Universitat Politècnica de València, Valencia (Spain). 4
- TAVERA, M., PÉREZ-GARCÍA, R., IZQUIERDO, J. & MONTALVO, I. (2008). Métodos no-destructivos para la mejora del rendimiento de una red hidráulica: casos de estudio con el georadar. In: *XXIII Congreso Latinoamericano de Hidráulica*, Cartagena de Indias (Colombia). 33, 36
- THOMSON, J., WANG, L. & ROYER, M. (2009). State of technology review report on condition assessment of ferrous water transmission and distribution systems. *EPA United States Environmental Protection Agency*, Report EPA/600/R-09/055. 5, 23, 53
- TOSTI, F. & UMILIAO, A. (2014). FDTD simulation of the GPR signal for preventing the risk of accidents due to pavement damages. *International Journal of Interdisciplinary Telecommunications and Networking*, **6** (1), pp. 1–9. 27
- TOSTI, F., PATRIARCA, C., SLOB, E., BENEDETTO, A. & LAMBOT, S. (2013). Clay content evaluation in soils through GPR signal processing. *Journal of Applied Geophysics*, **97**, pp. 69–80. 27, 40

REFERENCES

- UCHIKADO, T., KOBAYASHI, M., NAKANO, K. & SHIN, S. (2012). Emphasis of GPR signals reflected from buried pipes using statistical clustering techniques. In: *International Conference on Wavelet Analysis and Pattern Recognition*, Xian (China), pp. 410–416. 191
- ULRIKSEN, P. (1982). *Application of impulse radar to civil engineering*, PhD Thesis, Lund University of Technology, Lund (Sweden). 42
- URREA, J.P. & OSPINA, E. (2004). Implementación de la transformada de hough para la detección de líneas para un sistema de visión de bajo nivel. *Scientia et Technica*, **X** (24), pp 79–84. 113
- WARD, J.H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, **58** (301), pp 236–244. 194
- WEDEL, M. & BIJMOLT, T.H.A. (2000). Mixed tree and spatial representations of dissimilarity judgments. *Journal of Classification*, **17** (2), pp. 243–271. 197
- WEI, X. & ZHANG, Y. (2015). Interference removal for autofocusing of GPR data from RC bridge decks. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **8** (3), pp. 1145–1151. 249
- WEYNS, D. & HOLVOET, T. (2005). On the role of environments in multiagent systems. *Informatica*, **29**, pp. 409–421. 82, 170
- WILENSKY, U. (1999). NetLogo. Center for Connected Learning and Computer-Based Modeling, North-western University, Evanston, IL. 79, 85