Doctoral

Applied Arts

2015-05-01

Multimodal Content Delivery for Geo-services

Keith Gardiner
Technological University Dublin, keith.gardiner@tudublin.ie

Multimodal Content Delivery for Geo-services

A Thesis submitted in partial fulfilment of the requirements for the Degree of

Doctor of Philosophy

Keith Gardiner

M.Phil. Dublin Institute of Technology (2004)
B.Sc. Waterford Institute of Technology (2001)

School of Media
Dublin Institute of Technology

Supervised by: Dr. Charlie Cullen & Dr. James Carswell

May 2015

Abstract

This thesis describes a body of work carried out over several research projects in the area of multimodal interaction for location-based services. Research in this area has progressed from using simulated mobile environments to demonstrate the visual modality, to the ubiquitous delivery of rich media using multimodal interfaces (geo-services). To effectively deliver these services, research focused on innovative solutions to real-world problems in a number of disciplines including geo-location, mobile spatial interaction, location-based services, rich media interfaces and auditory user interfaces. My original contributions to knowledge are made in the areas of multimodal interaction underpinned by advances in geo-location technology and supported by the proliferation of mobile device technology into modern life. Accurate positioning is a known problem for location-based services, contributions in the area of mobile positioning demonstrate a hybrid positioning technology for mobile devices that uses terrestrial beacons to trilaterate position. Information overload is an active concern for location-based applications that struggle to manage large amounts of data, contributions in the area of egocentric visibility that filter data based on field-of-view demonstrate novel forms of multimodal input. One of the more pertinent characteristics of these applications is the delivery or output modality employed (auditory, visual or tactile). Further contributions in the area of multimodal content delivery are made, where multiple modalities are used to deliver information using graphical user interfaces, tactile interfaces and more notably auditory user interfaces. It is demonstrated how a combination of these interfaces can be used to synergistically deliver context sensitive rich media to users - in a responsive way - based on usage scenarios that consider the affordance of the device, the geographical position and bearing of the device and also the location of the device.

Thesis Declaration

I certify that this thesis which I now submit for examination for the award of PhD, is entirely my own work and has not been taken from the work of others, save and to the extent that such work has been cited and acknowledged within the text of my work.

This thesis was prepared according to the regulations for postgraduate study by research of the Dublin Institute of Technology and has not been submitted in whole or in part for another award in any other third level institution.

The work reported on in this thesis conforms to the principles and requirements of the DIT's guidelines for ethics in research.

DIT has permission to keep, lend or copy this thesis in whole or in part, on condition that any such use of the material of the thesis be duly acknowledged.

Signature _____ Date _____

Acknowledgements

I would like to thank my supervisors Dr. Charlie Cullen and Dr. James Carswell for their continued support throughout this process. I would like to thank my wife (to be) Lisa for her patience; this wouldn't have been possible without her. I would also like to thank all my colleagues, in particular Dr. Junjun Yin and Dr. Seamus Rooney with whom I have worked with on many projects throughout the process; your expertise and banter made it an extremely enjoyable experience.

List of Acronyms

2D	(2 Dimensional)
2.5D	(2 Dimensional With Height Information)
3D	(3 Dimensional)
3DQ	(3 Dimensional Query)
ACM	(Association of Computer Machinery)
AR	(Augmented Reality)
AUI	(Auditory User Interface)
BRETAM	(Breakthrough Replication Empiricism Theory Automation Maturity)
BT	(Bluetooth)
CH	(Cultural Heritage)
CHI	(Cultural Heritage Interfaces)
CM	(Content Modelling)
COTS	(Commercial Off-The-Shelf)
CS	(Computer Science)
DFO	(Department of Fisheries and Oceans)
DIT	(Dublin Institute of Technology)
EJB	(Enterprise Java Bean)
ESA	(European Space Agency)
EU	(European Union)
FOV	(Field-Of-View)
GIS	(Geographic Information Systems)
GLLFAS	(Great lakes laboratory for Fisheries and Aquatic Sciences)
GM	(Geo-data Modelling)
GPRS	(General Packet Radio Service)
GPS	(Global Positioning System)
GSM	(Global System for Mobile Communications)
GUI	(Graphical User Interface)
HCI	(Human Computer Interaction)
HD	(High Definition)
HQR	(Hidden Query Removal)
HTML	(Hypertext Markup Language)
ICARE	(Interaction Complementarity Assignment Redundancy Equivalence)
ICING	(Innovative Cities For The Next Generation)
ID	(Identifier)
ILC	(ICiNG Location Client)
ISPRS	(International Society for Photogrammetry and Remote Sensing)
IT	(Information Technology)
IUI	(Intelligent User Interface)
LAN	(Local Area Network)
LBS	(Location-Based Services)
LNCS	(Lecture Notes in Computer Science)
LOS	(Line-Of-Sight)
MEMS	(Mobile Environmental Management System)
MMT	(Mobile Mapping Technology)
MSI	(Mobile Spatial Interaction)
MR	(Mixed Reality)
NASA	(National Aeronautics and Space Administration)

NL	(Natural Language)
OAA	(Open Agent Architecture)
OS	(Operating System)
PC	(Personal Computer)
PDA	(Personal Digital Assistant)
POI	(Point Of Interest)
RF	(Radio Frequency)
RFID	(Radio Frequency Identification)
RMS	(Record Management System)
SFX	(Sound Effects)
SOAP	(Simple Object Access Protocol)
SOTA	(State-Of-The-Art)
SQL	(Structured Query Language)
TLX	(Task Load Index)
UI	(User Interface)
UM	(User Modelling)
VR	(Virtual Reality)
VRML	(Virtual Reality Modelling Language)
WAP	(Wireless Application Protocol)
WIFI	(Wireless Fidelity)
WIMP	(Windows, Icons, Menus, Pointer)
WIP	(Knowledge Based Presentation System)
WLAN	(Wireless Local Area Network)
WML	(Wireless Markup Language)
WWW	(World Wide Web)
XML	(Extensible Markup Language)
XSLT	(Extensible Stylesheet Language Transformations)

List of Publications

1. Gardiner, K., C. Cullen and J. D. Carswell. (2015) "MobiLAudio - A Multimodal Content Delivery Platform for Geo-Services", Journal of Location Based Services, Taylor and Francis. (Accepted – In Press)

2. Gardiner, K., C. Cullen and J. D. Carswell. (2015) "mobiSurround: An Auditory User Interface for Geo Service Delivery", 26th International Conference on Database and Expert Systems Applications (DEXA) Valencia, Spain. (Accepted – In Press)

3. Whipple, J., C. Cullen, K. Gardiner and T. Savage. (2015). "Syllable Circles for pronunciation learning and teaching." ELT Journal 69(2): 151-164.

4. Carswell, J., K. Gardiner and J. Yin. (2010) "3DQ: Threat Dome Visibility Querying on Mobile Devices." GIM International, Vol.24, (8), 24, August 2010.

5. Carswell, J. D., K. Gardiner and J. Yin. (2010) "Mobile Visibility Querying for LBS", Transactions in GIS, 14(6): 791–809, DOI: 10.1111/j.1467- 9671.2010. 01230.x.

6. Carswell, J.D., K. Gardiner, A. Rizzini, M. Bertolotto, and N. Mandrak. (2010) "A Web and Mobile System for Environmental Decision Support" in Decision Support Systems in Agriculture, Food and the Environment: Trends, Applications and Advances by Basil Manos, Konstantinos Paparrizos, Nikolaos Matsatsinis and Jason Papathanasiou (editors), IGI Global.

7. Gardiner, K., J. Yin and J. D. Carswell. (2009) "EgoViz –A Mobile Based Spatial Interaction System": The 9th International Symposium on Web and Wireless Geographical Information Systems (W2GIS), Maynooth, Ireland. Springer-Verlag, Lecture Notes in Computer Science, Volume 5886, pp 135-152, DOI: 10.1007/978-3-642-10601-9_10.

8. Carswell, J. D., K. Gardiner, A. Rizzini, M. Bertolotto, and N. Mandrak. (2008) "A Web-Based and Mobile Environmental Management System". Journal of Environmental Informatics 12(1) 9-20, DOI: 10.3808/jei.200800119.

9. Rooney, S., K. Gardiner and J. D. Carswell. (2007) "An Open Source Approach to Wireless Positioning Techniques". In Proceedings of the 5th International Symposium on Mobile Mapping Technology (MMT), Padua, Italy. ISPRS Archives – Volume XXXVI-5/C55.

10. Kilfeather, E., J. Carswell, K. Gardiner and S. Rooney. (2007) Urban Location Based Services using Mobile Clients: The ICiNG Approach, GISRUK, Maynooth, Ireland.

11. Rizzini, A., K. Gardiner, M. Bertolotto and J. Carswell. (2006) "MEMS Mobile GIS: A Spatially Enabled Fish Habitat Management System". The 6th International Symposium on Web and Wireless Geographical Information Systems (W2GIS), Hong Kong, China.

12. Gardiner, K., A. Rizzini, J. Carswell, M. Bertolotto. (2005) "MEMS - Mobile Environmental Management System". GIS Research UK, 13th Annual Conference, University of Glasgow.

13. Carswell, J., K. Gardiner, M. Bertolotto, N. Mandrak. (2004) "Applications of Mobile Computing for Fish Species at Risk Management", International Conference on Environmental Informatics, August 25-27, Regina, Saskatchewan, Canada.

14. Gardiner, K. (2004) "A Framework for Processing Viewer-based Directional Queries in a Simulated Mobile Environment." Masters Thesis. Dublin Institute of Technology.

15. Gardiner, K. and J. D. Carswell. (2003) "Viewer-Based Directional Querying for Mobile Applications", Third International Workshop on Web and Wireless Geographical Information Systems W2GIS, Roma, Italy. In Proceedings of the Fourth international conference on Web information systems engineering workshops Pages 83-91, IEEE CS Press, ISBN:0-7695-2103-7.

16. Carswell, J., K. Gardiner, M. Neumann (2003) "Wireless Spatio-Semantic Transactions on Multimedia Datasets", 19th Annual ACM Symposium on Applied Computing, Nicosia, Cyprus.

17. Carswell, J. D., A. Eustace, K. Gardiner, E. Kilfeather and M. Neumann. (2002) "An Environment for Mobile Context-Based Hypermedia Retrieval." DEXA Workshops 2002: 532-536.

List of Selected Papers

I. Gardiner, K., C. Cullen and J. D. Carswell. (2015) "MobiLAudio - A Multimodal Content Delivery Platform for Geo-Services", Journal of Location Based Services, Taylor and Francis.

II. Gardiner, K., C. Cullen and J. D. Carswell. (2015) "mobiSurround: An Auditory User Interface for Geo Service Delivery", 26th International Conference on Database and Expert Systems Applications (DEXA) Valencia, Spain.

III. Carswell, J. D., K. Gardiner and J. Yin. (2010) "Mobile Visibility Querying for LBS", Transactions in GIS, 14(6): 791–809, DOI: 10.1111/j.1467- 9671.2010. 01230.x.

IV. Gardiner, K., J. Yin and J. D. Carswell. (2009) "EgoViz –A Mobile Based Spatial Interaction System": The 9th International Symposium on Web and Wireless Geographical Information Systems (W2GIS), Maynooth, Ireland. Springer-Verlag, Lecture Notes in Computer Science, Volume 5886, pp 135-152, DOI: 10.1007/978-3-642-10601-9_10.

V. Carswell, J. D., K. Gardiner, A. Rizzini, M. Bertolotto, and N. Mandrak. (2008) "A Web-Based and Mobile Environmental Management System". Journal of Environmental Informatics 12(1) 9-20, DOI: 10.3808/jei.200800119.

VI. Rooney, S., K. Gardiner and J. D. Carswell. (2007) "An Open Source Approach to Wireless Positioning Techniques": In Proceedings of the 5th International Symposium on Mobile Mapping Technology (MMT), Padua, Italy. ISPRS Archives – Volume XXXVI-5/C55.

VII. Gardiner, K. and J. D. Carswell. (2003) "Viewer-Based Directional Querying for Mobile Applications", Third International Workshop on Web and Wireless Geographical Information Systems W2GIS, Roma, Italy. In Proceedings of the Fourth international conference on Web information systems engineering workshops Pages 83-91, IEEE CS Press, ISBN:0-7695-2103-7.

Table of Contents

List of Figures

List of Tables

1. Introduction

The use of geo-location for informing information retrieval has become more pervasive in society in recent times. Mobile devices are now equipped with multiple sensors to track movement in order to improve and deliver services. From simple mapping to immersive audio tours, the technology that provides these services is rapidly becoming more ubiquitous as consumers demand more complex task relevant services over basic functionality. Reliance on mobile devices has also changed due to the proliferation of mobile applications (apps) that provide a vast range of services, where the importance of the delivery of the service is on par with the usability and user experience of the app.

Building apps that effectively deliver these services requires an understanding of the underlying technologies and poses a number of technical problems. The first and most fundamental problem is how accurately a user can be located. GPS is a powerful technology, however even now (unassisted on mobile devices), it is accurate to only 10 -30m. This is an on-going problem on commercial off-the-shelf (COTS) mobile devices. However, assuming that we have accurate positioning, the second issue is how do we structure and query the data. There are two aspects to this problem and these are (1) how do we define our interaction with the physical environment and (2) how do we query our virtual environment effectively. To do this, accurate virtual representations of the environment have to be produced and a precise geo-location is essential. In addition, modelling a user's interaction with a space in terms of their egocentric parameters has to be considered. Finally, given that we have precisely defined the space and the user interaction, effective information delivery is the third problem. Specifically, the issue of information overload while using the visual modality poses many problems when trying to convey information effectively (with auditory/haptic interfaces not very prevalent).

In this Thesis, the first issue is investigated from an assisted point of view. While GPS suffers from the urban canyon effect due to multi-storey buildings reducing satellite visibility and is effectively unavailable indoors, the idea of a hybrid-positioning model is investigated where the use of additional beacons such as using Wi-Fi, GSM and Bluetooth are used to assist in determining a more accurate user position. The detection of known situated beacons is used alongside the GPS signal to provide additional information that can assist in trilaterating a more accurate position. This approach proves to be highly effective with the availability of an optimum density of beacons.

The second issue is investigated from two perspectives. Firstly, to query the physical environment, an accurate representation has to be created and secondly the mode of interaction with the virtual representation is a complex task and requires a solid understanding of spatial interaction and the parameters required to effectively do this. The creation of accurate geo-referenced 3D models is described as a foundation for performing directional queries. The orientation of the user on the 2D plane is taken into account when querying these models. Extending this, the utilisation of a user's 3D perspective is investigated leading to extensive investigations in the area of mobile spatial interaction (MSI).

The third issue relates to content delivery and the research focuses on effective means of delivering content to users using multiple modalities. This multimodal approach begins with the use of traditional methods using graphical user interfaces and extends to complex methods of data and content modelling to create a platform for survey, preparation and delivery of immersive audio content in the form of speech, background music and sound effects. Using a unique 3-tiered listening model, issues relating to information overload and content redundancy are addressed. In addition, novel

approaches to the use of 3D audio as a delivery mechanism are tested and the implementation of a natural directions model is described.

The remainder of this section will outline the Thesis aims and describe the overall structure of the document in the context of the research carried out.

1.1. Thesis Aims

The delivery of geo-services in a multimodal environment is incumbent on several systemic parts working together to provide the service. There are minimum sets of requirements that need to be present for this to be achieved. (1) Accurate positioning, although difficult to guarantee, is a fundamental part of the process on which all other parts heavily rely on. This needs to be provided and managed effectively. (2) A geometrically accurate (and ideally photo realistic) virtual representation of the physical environment is required in order to perform spatial queries; therefore a 3D model is required. (3) A 3D spatial query processor is required to determine the spatial interaction between the user and the physical environment representation. This relies on mobile device hardware to provide accurate compass (direction) and gyroscopic (tilt) data in real-time. (4) An effective content delivery engine for delivering visual content using a graphical user interface (GUI) and audio content using an auditory user interface (AUI). This requires that a stringent data and content modelling strategy be in place in order to achieve this effectively. (5) The system needs to be tested on real users to determine how effective the use of multiple modalities for content delivery is in different contexts. Therefore, the aims of the work examined in this Thesis are summarised as the following:

Aim 1: Develop systems for media retrieval based on mobile spatial interaction (MSI) and contextual data sets.

Aim 2: Multimodal content delivery with a main focus on the visual and aural modalities based on innovative applications of content modelling and phrase synthesis.

Aim 3: Application of methods developed in multiple domains to validate the effectiveness of the multimodal content delivery.

1.2. Thesis Structure
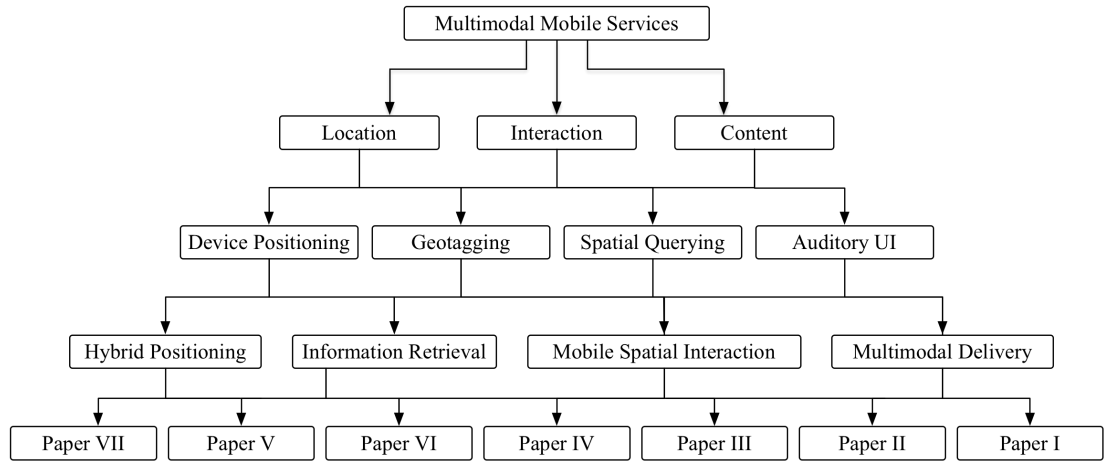
Figure 1: Thesis structure. *Overview of Thesis subject/research areas and their linkages to academic research outputs. This illustration is used throughout the document to provide context and aid navigation.*

Figure 1 illustrates the linkages between subject areas, research areas and relevant publications that relate to the research presented in this Thesis. The overarching theme of multimodal mobile services splits into three central components, which are location, interaction and content. This leads into the subject areas of device positioning, geotagging, spatial querying and auditory user interfaces (AUIs). Following this, the specific research areas of hybrid positioning, information retrieval, MSI and multimodal delivery are linked directly to the relevant publications in these areas.

1.3. Thesis Contributions

This section summarises the contributions made in the three main areas of research presented in this Thesis. Significant contributions were made in the areas of MSI and mobile device positioning leading to further contributions in the area of multimodal content delivery.

1.3.1. Mobile Spatial Interaction

The main contributions in this area involved the simulation of a mobile environment and the development of a directional query processor. This was extended with the development of a web application and an offline mobile application for the collection of context sensitive geo-information in the field. Further contributions demonstrated the use of MSI to query multimedia datasets and the development of a 3D querying platform for location-based services (LBS).

1.3.2. Mobile Device Positioning

Contributions in this area involved the development of a hybrid positioning module and the production of a beacon database for Dublin that demonstrated accurate position and orientation determination on mobile devices, vital for effective LBS.

1.3.3. Multimodal Content Delivery

Initial contributions in this area involved the demonstration of effective UI for visualizing spatial queries and results. Extending this, a data/content modelling strategy for LBS was established that supported the development of an AUI and platform for geo-services production. Further contributions in the form of a user trial tested the validity and effectiveness of multimodal content delivery for geo-services in a mobile context.

2. Methodologies

In this Chapter a description of the subject areas and methodologies that pertain to the body of research carried out is given. Each paper included in this Thesis relates to one of these areas but is not necessarily restricted to it. In some cases, there are multiple linkages - mainly where research is extended into new fields. This section describes the methods used for each branch of research and references contributions made in order to progress beyond the state-of-the-art and create new knowledge.

Section 2.1 describes the research area of LBS and in particular research carried out by the Author in simulated mobile environments. Section 2.2 outlines research in the area of hybrid positioning and how this is a fundamental part of LBS. Section 2.3 describes work carried out in the area of MSI with a particular focus on contextualized information delivery. Section 2.4 is an evaluation of the area of AUIs identifying the relevance of this area in terms of location-based information delivery, revealing some of the problems that exist in the field. Finally, Section 2.5 reviews the methodologies employed in the area of multimodal interaction with a focus on input and output modalities relevant to mobile content delivery.

2.1. Location-Based Information Retrieval

Location-based information retrieval has been a very active area of research for many years (Salton and McGill 1983, Abowd et al. 1997, Egenhofer 1999, Schiller and Voisard 2004, Chan et al. 2005, Jiang and Yao 2006, Frohlich, Simon, and Baillie 2009, Dey et al. 2010, Rahman and El Saddik 2013). Since the first methods of categorizing data based on its geographic locality, the idea of providing context sensitive information based on location has become increasingly relevant on the Web (Buyukkokten et al. 1999). Defining the fundamental parameters for this process is key to its effectiveness,

as the data needs to be accurate to provide relevant results. In (Gravano, Hatzivassiloglou, and Lichtenstein 2003), the approach to location-based Web search queries is to automatically define the geographical scope of Web resources (Ding, Gravano, and Shivakumar 2000) and compare this against an explicit user location, to provide location sensitive results. The goal of the SPIRIT project (Heinzle, Kopczynski, and Sester 2003) was to enrich web resources with spatial information and make it accessible via search engines on the web. A further extension of this work was the idea of a sketch-based approach that enabled the user to sketch out their query, essentially defining a spatial query window[1]. This provided a unique method of localized search on the web using spatial relations. The results of this research demonstrated successful classification of resources and showed how queries could be performed to determine resources in the locality based on a fixed location. These innovations in the Web domain proved successful and with advances in mobile technology, led to possibilities in the mobile domain. The idea of querying datasets (with a spatial component) based on a mobile users location or indeed updating the dataset was explored. With advances in mobile device technology, this idea had become feasible in many ways primarily due to device positioning systems. However, before such devices were commercially available, significant research was carried out in the area of spatial information retrieval using geo-location. In (Persson et al. 2003), the GeoNotes system allows users to post notes to a database with a spatial component for the purposes of discovery by other users of the system. Laptops/PDAs with GPS receivers and an extensive WLAN were used to demonstrate the system and the exciting possibilities of such technologies.

Research in (Paper VII) describes work that simulated the concept of mobile context-

---

[1] A spatial query window is a 2d bounding box that defines a geographical area used to query a spatial datasets.

based information retrieval by using a user's position to query an underlying map and POI database in order to deliver text and images about the location. To achieve this, the user's location, a digital map and access to contextually relevant data was essential. A 3D virtual map was developed to demonstrate and monitor the user's physical interaction within the environment. The map was stored in a spatial database and registered to real-world coordinates together with a number of points-of-interest (POIs) representing culturally relevant locations (Gardiner 2004). Using these three pieces of structured data, early examples of location-based information delivery were demonstrated which outlined the potential for LBS and identified areas that required significant additional research for LBS to be considered viable in real-world environments (Carswell et al. 2002, Schiller and Voisard 2004, Wakkary et al. 2004).

This idea was extended in the Agamemnon project (Ancona, Scagliola, and Traverso 2006) to demonstrate the delivery of content in a real world context to provide focus-attention guided visits of archaeological sites and museums. This work demonstrated the viability of these types of applications in the field. In (Simon, Fröhlich, and Anegg 2007) an application framework for building spatially aware mobile applications was proposed. This approach included a responsive XML based data exchange format that describes the geographic vicinity of the user enabling novel applications to be developed using gesture based interaction. More recently, the idea of mobile spatial search has gained significant attention due to improvements in device sensing technologies. The point-to-query concept has been investigated in (Jacob, Mooney, and Winstanley 2012, Carswell and Yin 2012, Yin and Carswell 2013) with a focus on spatial interaction and haptic feedback, enabling users to define a search space and receive feedback to aide their interaction.

Extending into the area of AR, an interesting approach is taken in (Lee, Kwon, and Sumiya 2009) with a layer-based model for information delivery. In this research the idea of presenting information using mixed-reality is explored using object identification and a layer-based approach to information display based on proximity. Studies by (Geiger et al. 2014, Schickler et al. 2015) supports this idea and investigate the design of AR applications for LBS with the implementation of a mobile augmented reality engines for multiple platforms. A novel approach is taken in (Balduini et al. 2012) where AR is used to deliver personalised location-based recommendations of social media streams. This idea is further explored in (Lin et al. 2014) with a location aware AR application that enables users to discover social data and share their own AR imagery with others in their social network. The concept of using wearable technologies to provide location-based information has started to emerge. In (Rogers et al. 2015), the concept of using visual or auditory barcodes with Google Glass is described to provide users with location specific information. To assist in social interactions, research in (Mandal et al. 2014) proposes the use of facial recognition to provide information on people in face-to-face scenarios. Social interaction is also explored in (Xu et al. 2015) where the focus is on user attitudes towards social interaction assistance and highlighting related issues.

2.2. Hybrid Positioning Techniques

To provide LBS effectively, a number of technologies are required to work together successfully (Schiller and Voisard 2004, Hightower and Borriello 2001). Two of the most essential aspects of this process are location accuracy and access to reliable contextual information (Dey et al. 2010). Location accuracy refers to the accuracy of a user's acquired geographical position and reliable contextual information refers to other

data about the user's current context and how they interact with a space, i.e. their orientation, their FOV, their personal preferences or the time of day.

To obtain a user's geographical position, the now standard and arguably most reliable way to achieve this is to use the global positioning system (GPS). GPS uses signals transmitted from satellites orbiting the planet to trilaterate a position on the surface of the earth (El-Rabbany 2006), with the receiver being a mobile device for example. It generally requires at least 4 satellites to be in view of the receiver in order to resolve a position. However, the visibility and consequently the accuracy of the position can be affected by environmental conditions such as bad weather and tall buildings. In addition, reflection of signals off buildings or compromised signals can also have a huge impact on the accuracy of the position calculation. This "urban canyon" effect means that the location-based application can assume a positional error along the street of ~28 meters for 95% of the time (Modsching, Kramer, and ten Hagen 2006). This is a major drawback as the primary requirement for these applications is location.

Fortunately, the same trilateration principles used in GPS can also be used to accurately locate users based on terrestrial radio frequencies such as GSM, Wi-Fi and Bluetooth signals (LaMarca et al. 2005). These radio frequency transmitters are considered because mobile devices can detect them and they usually have a fixed location, i.e. a beacon. Using information about the locations of these beacons combined with signal strengths, it is possible to incorporate these into position calculations to improve the accuracy (Hightower and Borriello 2004). In (LaMarca et al. 2005), research was conducted that developed a hybrid positioning module for mobile devices (Place Lab) that monitored GPS, GSM, Wi-Fi radio signals to deliver a more accurate position than GPS alone.

The goal of Place Lab was to use the availability of additional radio signals (GSM, Wi-Fi, and Bluetooth) to improve position accuracy. By exploiting the use of fixed reference points in the form of a database of known beacon locations, a more accurate position could be calculated using the principles of trilateration. Tracker and spotter modules are used to detect the fixed beacons and this information combined with signal strength and historical data results in more accurate position estimations. Unfortunately, at the time, certain technologies were not yet advanced enough to take Place Lab fully mobile as some of the signal spotters worked on mobile phones (i.e. Bluetooth and GSM) while others required a laptop (i.e. Wi-Fi). Furthermore, Place Lab also supported passive monitoring giving users control over disclosure of their location; meaning correction values over GSM were not considered, laying the foundation for privacy-observant location-based applications (Hightower, Lamarca, and Smith 2006).

A more open approach is considered in (Bruning et al. 2007), where the idea of a cooperative hybrid position system is proposed. MagicMap is a cooperative positioning system that uses a combination of various wireless networking technologies to determine the client's position. The unique aspect is the utilisation of any available radio signal for position calculations (WLAN, RFID and ZigBee). This generic approach leads to better spatial positioning coverage, higher dependability and increased accuracy in places where multiple radio signals are available. This approach relies on communication with other mobile clients via a Web service interface, with each client able to calculate its own position as well as the position of any other client within the MagicMap environment, resulting in a 33% improvement on average in positional accuracy when compared to GPS alone. Nevertheless, the drawbacks of this technology are that each client requires a network connection in order to make its

position public and consequently, the privacy is affected due to its open nature.

With the possibility of calculating a user's position accurately (in absolute terms) in an urban environment meant that the concept of effective LBS was becoming more plausible. However, there were still some barriers to successfully achieving this in a truly mobile context, one being the availability of Wi-Fi hardware on mobile devices and the other being access to relevant geospatial data over mobile data network connections (Sharif et al. 2006, Roos, Schwarzbacher, and Wieland 2006). The former was the motivation for research carried out in the area of network independent hybrid positioning techniques outlined in (Paper VI). In this research a number of innovations were made enabling effective demonstration of LBS applications. The primary objective was the development of privacy sensitive positioning module that relied solely on the detection of situated beacons without the requirement of network access. With the addition of GSM beacon data and a weighted centroid approach to position determination (based on signal strength), positioning accuracy was improved considerably in urban areas with an accuracy of 15-20m in optimum conditions (3 802.11x access points during 10sec window). A related approach is taken in (Zirari, Canalda, and Spies 2010) where Wifi is used in conjunction with GPS to support trilateration when the optimum number of satellites is unavailable, resulting in a significant improvement in accuracy over GPS alone. A Wifi only approach to indoor positioning is considered in (Bell, Jung, and Krishnakumar 2010) with the development of the *SaskEPS* system. The focus here is on the mapping, calibration and classification of Wifi beacons to provide GPS-like accuracy. Work in this area is further extended by (Yapeng et al. 2013), where the use of Bluetooth beacons is the main focus due to improvements in Bluetooth 2.1 that permits the detection of radio signal strength (RSS)

without the time-consuming pre-connection process of previous versions. This makes it a more viable addition to hybrid positioning systems compared to previous implementations, helping improve position estimations considerably.

The use of radio frequency identification (RFID) and near field communication (NFC) as an addition to hybrid positioning systems has gained considerable attention. RFID in particular has gained significant recognition for indoor positioning with multiple applications described in (Retscher 2012, Bai et al. 2012), however complex infrastructure and unstable RSS signals are amongst on-going issues with the technology when used for hybrid positioning. In (Seo and Ahn 2013), NFC is used in conjunction with Wifi in a gallery context, where Wifi level positioning provides viewing path information and NFC provides exhibit specific information (4-10cm range). This type of combined approach improves the accuracy and effectiveness of content detection and delivery in mobile applications.

More recently, the use of image recognition is used to support indoor/outdoor positioning. In (Marimon Sanjuan et al. 2011, Werner, Kessel, and Marouane 2011), AR applications that suffer due to imprecise GPS readings use image recognition techniques to supplement GPS to provide more accurate positioning. This approach is further extended in (Guan et al. 2014) with the use of panoramic imagery database and head tracking approach for the detection and determination of position calculations. An inertial navigation system for indoor positioning is described in (Diaz 2015) where dead reckoning is used to estimate position. In this research a novel step-and-heading approach is used to provide 3D positioning using a step detector that detects steps up and down and a novel vertical displacement estimator based on an in-pocket scenario. The algorithms employed outperform the previous attempts at 3D positioning using

similar approaches.

2.3. Mobile Spatial Interaction

Mobile spatial interaction (MSI) is the study of the spatial interaction that occurs between mobile nodes (e.g. GPS enabled mobile devices) and the environment (Frohlich, Simon, and Baillie 2009). This interaction is effectively a virtual interaction, where the parameters required to detect the interaction are absolute node location, a spatial data model of the environment and direction data (Strachan and Murray-Smith 2009). Therefore MSI is the detection of spatial interaction using spatial relations.

The approach of using direction to query spatial data is the focus of substantial research efforts within the spatial database community (Goyal and Egenhofer 2001, Liu, Shekhar, and Chawla 2000, Chan et al. 2005). Direction relations therefore represent an important class of queries in spatial databases and their applications to geographic information systems. To make sense of direction in this context, a reference frame must first be established. The study of object-orientation-based direction queries in spatial databases in (Liu, Shekhar, and Chawla 2000) focuses on an intrinsic frame of reference, where the reference frame is in respect to the orientation of an object: front or back, left or right of a building for example and queries like "Are there any artefacts in front of the post office?" can be answered. In this instance, object-based directional queries are performed that determine the direction predicates[2] of a given dataset in relation to the reference object. For configurations of spatial objects, in a GIS or digital image, that represent real positions and orientations of the environment for example, it has been customary to use extrinsic reference systems where the reference frame is established independently of the orientation of the features or the observers, e.g. north,

---

[2] Directional predicates refer to the directional relationship between two spatial objects.

south, east, west. In (Goyal and Egenhofer 2001), an extrinsic reference system is used to develop a computational model for determining the directional similarities between spatial objects where a (cardinal) direction-relation matrix is successfully used to assess similarities in spatial scenes. In contrast, an example of a query within a deictic [3]reference frame would be "Are there any artefacts contained within my FOV?" where the reference frame is relative to each individual looking at the scene, e.g. what is "in front" for me might be "to the left of" someone else. Therefore, MSI is the study of the interaction between the egocentric view (based on a deictic reference frame) of a mobile user and a world of spatial information.

This egocentric approach to LBS was first investigated in (Paper VII), where directional queries were performed on multimedia datasets in a simulated mobile environment. This research identified issues concerned with positioning and mobile hardware availability which were investigated in (Paper III) and (Paper VI). With these barriers removed, research in the area of MSI was a distinct possibility. Research in (Simon et al. 2007) describes the use of MSI to create a smart pointer that gives access to digital information attached to POIs. This popular approach is also used in (Lei and Coulton 2009) to perform mobile spatial search and content generation in the form of directional photo uploads, where the direction the photo was taken from in relation to the POI is taken into account. In (Frohlich, Simon, and Baillie 2009, Strachan and Murray-Smith 2009), comprehensive studies are also carried out in this area with novel applications proposed that identify the possibilities of this technology for information discovery, way-finding and VR using customised mobile devices.

2.3.1. Egocentric Visibility in LBS

---

[3] A deictic frame of reference refers to observation from a user point-of-view or egocentric position.

Visibility modelling can be used to determine regions that can be viewed from a given observation viewpoint. This has been an active topic of research in GIS for many years with the isovist [4](Benedikt 1979, Batty 2001) and viewshed analysis (O'Sullivan and Turner 2001, Fisher 1995) amongst the most popular approaches used to determine the limits of view. More recently the use of visibility modelling has gained popularity in the LBS space (Bartie, Mills, and Kingham 2008) as a means of filtering data based on the visibility of the user. LBS applications have had a slow uptake mainly due to privacy concerns, poor user experiences and a lack of service dependability (Chincholle et al. 2002), however the use of LBS is on the rise (Zickuhr and Smith 2010) and intelligent methods of information filtering are becoming increasingly important. With the availability of accurate location data and widespread access to context sensitive data the issue of information overload has never been more prevalent.

Egocentric visibility can be described as the filtering of data based on visibility from a user's egocentric perspective. Early examples of visibility determination using ray tracing is examined in (Beeharee and Steed 2005) and demonstrated in a simulated mobile environment in (Gardiner and Carswell 2003). In (Bartie, Mills, and Kingham 2008, Bartie et al. 2010), research shows that a filter based on the visibility of features is a useful additional capability for such services. This is made possible through the use of digital surface models where a number of visibility metrics are suggested to provide visibility information so that points of interest may be ranked according to a meaningful priority. Research in (Maierhofer, Simon, and Tobler 2007) describes a solution that uses a 3D-geometry database of a city and uses a guided visibility sampling algorithm to determine object visibility by starting with a number of random visibility samples,

---

[4] An isovist is the volume of space visible from a given point in space, together with a specification of the location of that point.

and then extending the set of visible objects, by sampling in the vicinity of them. A similar approach that further extends this idea is described in (Paper IV). This approach takes into account the user's actual field of view from an egocentric perspective. Various information filters are used to represent line-of-sight, viewshed and isovist queries. The application operated on current COTS mobile devices and thus was capable of being used in real-world situations. By using field-of-view (FOV) as the query "window" for spatial queries, results returned were restricted to what was in their 2D FOV and helped to assist in the reduction of the associated problem of information overload. This approach was extended to 2.5D spatial search when height was considered as part of the spatial query however this was not truly representational of a users actual FOV. In (Yin and Carswell 2011) a case study using Open Street Map (OSM) and iOS to perform *Touch2Query* operations is presented. The mobile prototype allows users to perform spatial queries using customisable query windows to extract relevant information about the environment. The *Zapp* application described in (Meek et al. 2013) enables geo-tagging of any distant point on the visible landscape in order to record observations from multiple viewpoints. An interesting approach is also taken in (Moniri, Feld, and Muller 2012, Moniri et al. 2012) to personalised user interaction in urban environments. In this research the use of visibility models is used in an automotive context to reduce the reference set for voice-based queries about the physical environment. Research in (Ballatore et al. 2015, Pham Thi et al. 2013) describes the use of egocentric visibility models for the development of personal geo-services for Universities, where multiple forms of user input, based on visibility can be used to query an underlying 3D representation of the campus (eCampus model). These works outline the relevance of visibility modelling for contemporary geo-services.

2.3.2. 3D Spatial Search

Performing 3D spatial queries presents a number of challenges for LBS (Zlatanova and Verbree 2003, Yin and Carswell 2011). Although there is still a large gap between developments on mobile platforms and the state-of-the-art (SOTA) in relation to 3D GIS and Spatial Data Infrastructures, applications are starting to emerge that can demonstrate services effectively (Goetz and Zipf 2010). To perform such queries a 3D dataset of the environment plus an accurate user position and 3D spatial query operators are required to accurately determine the spatial interaction in 3D (Schön et al. 2009, Wang and Liu 2008). In (Simon et al. 2007, Wilson and Shafer 2003, Lei and Coulton 2009), the idea of a customised mobile GeoWand/XWand was introduced that enables users to point at a building or object and obtain information about it. Taking this a step further and extending the idea of egocentric visibility, the development of the 3DQ processor is described in (Paper III), which allows the user to perform point-to-select and visibility queries on COTs devices. The 3DQ processor enabled the user to define dynamic query shapes based on a user's viewshed, a 360° isovist and a threat dome geometry (Carswell, Gardiner, and Yin 2010) to limit the amount of data returned to the user. This solution used a 3D city model stored in Oracle Spatial 11g and interrogated it with position and orientation data from an iPhone 4s using 3D operators in Oracle 11g. This approach is also detailed in (Yin 2013) where the use of 2d and 3d spatial queries are used to query 3D city models. The approach uses 3D vectors models in combination with MSI data to perform 3D spatial search of urban environments. A more recent approach to 3D spatial search is to use AR or Mixed Reality (MR) interfaces to query a 3D space. In (Nurminen, Järvi, and Lehtonen 2014) the use of a MR interface is used to demonstrate real-time tracking of public transportation. The system provides a real-time 3D map interface and a MR interface of the bus network. In the context of tourist

applications, research in (Tatzgern et al. 2015) describes the use of AR to query distant real-world objects. This innovative approach demonstrates the applicability of AR to geo-services by overlaying a virtual copy of the environment on the real environment to support the continuous exploration from different viewpoints of real-world objects. In (Altwaijry, Moghimi, and Belongie 2014) the use of AR is used to provide a tour guide experience using Google Glass. This approach demonstrates the effective use of wearable technology for mobile geo-services.

2.4. Auditory User Interfaces

An AUI is defined as the use of sound to communicate information about the state of an application or computing device (McGookin 2004, Shneiderman 2005). AUIs have traditionally been used for accessibility in desktop applications for the visually impaired with the main objective of transforming a graphical representation to an auditory one (Edwards 1989, Gaver 1989, Mynatt 1990, Donker, Klante, and Gorny 2002, Jagdish et al. 2008, Shinohara and Tenenberg 2009). Auditory non-speech audio has also been used to provide additional information to users about navigating a GUI (Kramer 1993, Absar and Guastavino 2008). This idea has gained popularity in the LBS and Augmented Reality (AR) space and has been used to provide navigational information and/or content to users of mobile applications (Baus et al. 2007, Walker and Lindsay 2006, Stahl 2007).

In (Holland, Morse, and Gedenryd 2002), the use of tones and audio panning is used in the AudioGPS application to communicate direction and distance when directing a user through a space. User trials show that the use of simple tones and the placement of them on the stereo stage are effective for indicating direction and that a simple oscillating tone was effective for communicating distance. Related work by (Strachan et al. 2005,

Strachan, Williamson, and Murray-Smith 2007) investigated the use of music for the purposes of navigation where a listener's music is modulated according to the changing predictions of a user position to guide them through a space.

The use of audio for content delivery is described in (Bartie and Mackaness 2006) where a combination of visibility querying and audio augmented reality is used to create a CityGuide application. The application proved to be an effective, non-invasive approach to augmenting the user's reality as it offered hands-free and eyes-free usage, enabling the user to keep focussed on the environment. The use of spatialised soundscapes to convey contextual information has been used in many cases in an attempt to deliver an immersive user experience (Wenzel 1994, Misra, Cook, and Wang 2006, Vazquez-Alvarez, Oakley, and Brewster 2012, Stahl 2007). The LISTEN project (Eckel 2001) provides users with intuitive access to personalized and situated audio information to enhance the sensual impact of applications ranging from art installations to entertainment events. This is achieved by augmenting the physical environment using immersive, spatial audio and advanced user modelling methods.

An interesting approach is taken by (Heller et al. 2009) with the development of an interactive virtual tour where users can remain independent or can interact with other users on the tour, essentially "checking in" with them. The ec(h)o system described in (Wakkary et al. 2004, Droumeva 2005) uses an *audio augmented reality interface* that is used to deliver exhibition data based on a semantic structure and utilises spatialised soundscapes to provide an immersive experience. Work carried out in (McGee and Cullen 2009a, b) suggested the usefulness of audio as a delivery modality for LBS. In particular, the advantages identified were both its hands-free and eyes-free nature, plus fast neural processing rate. An experimental case study in (Vazquez-Alvarez, Oakley,

and Brewster 2012) tests the use of earcons, proximity and spatial audio queues for effective information discovery using audio. A combination of 3D spatial audio techniques together with earcons proved to the most effective audio display. This work also suggests that the location and orientation sensing technologies available in COTS devices are effective when used to create rich and compelling outdoor augmented-audio environments. All these works point to the benefits of the hands-and-eyes-free nature of non-visual modalities and to the potential the aural modality has for the provision of navigational and context relevant information in a geo-services context.

In (Paper II), the development of an AUI for geo-services is detailed. The idea of presenting geo-referenced information both visually and aurally to the user in a collective way is the primary focus. The visual modality is presented using a graphical user interface (GUI) in a state-of-the-art mobile app. The aural modality is presented using the AUI, which is the main innovation in terms of delivering context sensitive information using focus independent (eyes free) means. The AUI is intended to be a non-visual interface that can be used in combination with a GUI, or not. The idea is that all information in the app, including content, directions, advertisements, etc. can be delivered without the user ever needing to touch the mobile device (McGee and Cullen 2009a, Alvarez and Brewster 2010, Vazquez-Alvarez and Brewster 2009). A comparable approach to auditory content delivery is taken in the *Spacebook* project (Mackaness et al. 2014, Janarthanam et al. 2013). The system described is capable of both guiding and describing the urban landscape as the tourist navigates around a city using synthesised speech. Automatic speech recognition, interaction management, and language generation allow the user to request more detailed information. The system is underpinned by a 2.5D model of the city, supporting real time modelling of landmarks

in the users FOV. In (Niculescu et al. 2014) the development of a conversational agent for the tourist domain called *SARA* is described. This multimodal application receives instructions using a speech interface and presents information using the visual and auditory modalities.

2.5. Multimodal Interaction

Multimodal interaction refers to interaction where a user is provided with multiple ways to communicate with a system using natural modes (Oviatt 2003), which can include graphics, speech and gesture. Multimodal interfaces enable this to happen and can combine two or more input modes in a coordinated manner to produce multimedia output. Modality refers to the type of communication channel used to convey or acquire information and the manner in which this is performed or perceived (Nigay and Coutaz 1993). From the perspective of multimodal systems, human output channels are referred to as *modality inputs*, and computer output channels are referred to as *media outputs* (Klemettinen 2007). For our purposes, there are two significant aspects to this research that relate to LBS.

2.5.1. Input Modalities

Input modalities can be described from two main perspectives, the human and the computer. Human input modalities refer to our sensory abilities such as visual, auditory and touch (Blattner and Glinert 1996, Turk 2014). Computer input modalities refer to the capability of a device to handle input from a user through a number of communication channels (i.e. a direct link from a user output modality to a computer input modality). These can include amongst others, text, sound, speech recognition, image/video and mouse pointing and clicking (Turk 2014).

The most noteworthy example of a multimodal interface is described in (Bolt 1980), which enabled a user sitting in a chair in a media room to interact with a map display to perform actions like "put that there" and "make that smaller" using speech and gestures. The system required both speech and gestural input to interpret commands properly demonstrating a simple, natural interaction by the user. A similar approach is described in (Neal et al. 1989) where natural language (NL) processing is used with deictic and graphic gestures for user inputs and combinations of language, maps and graphics as system-generated outputs. An innovative implementation is described in (Koons, Sparrell, and Thorisson 1998) where the use of gaze is combined with speech and gesture to generate deictic instructions in order to manipulate a map interface. The idea of combining pen-based interaction with other modes such as speech, non-linguistic vocalisations and touch gained significant attention in the research community (Cohen et al. 1997, Oviatt et al. 2000, Oviatt 2002, Harada, Saponas, and Landay 2007, Hinckley et al. 2010, Tian et al. 2006). These systems supported synergistic usage of one or more modalities where the user could use modalities at the same time.

More recent innovations in mobile technology in the *post-WIMP*[5] era have led to significant research in the area of multimodal input (Dam 1997) with modes like touch, gesture and speech opening up new possibilities in the form of perceptual user interfaces which are more natural and try to emulate how users interact with each other and their environment (Crowley, Coutaz, and Bérard 2000, Iannizzotto et al. 2005, López 2009, Oviatt and Cohen 2000, Ruddarraju et al. 2003, Turk and Robertson 2000). In these systems, the term fusion engine is used to describe the components of a system that interprets multiple input streams, which can vary, based on contextual differences

---

[5] post-WIMP interaction refers to work on user interfaces that goes beyond the traditional windows, icons, menus and pointing devices paradigm.

(Nigay and Coutaz 1993, Pfleger 2004, Dumas, Lalanne, and Oviatt 2009). A study outlined in (Lalanne et al. 2009) describes the development stages of fusion engines with 60% considered to be in the maturity phase based on the BRETAM model (Gaines and Shaw 1986)[6]. This means "theories have been assimilated and are routinely used without questions and have been deployed in large practical applications".

In multimodal applications an active event is an asynchronous occurrence of a mouse click or speech recognition results where the user is the input source. This is most common type of input event and has been the main focus of research in this area (Klemettinen 2007). A passive input event is one that has not been directly initiated by the user but is a representation of a physical environment property translated through the user to the computer with examples being GPS signals and compass headings. It has been shown that spatial domains are particularly suited to multimodal interactions (Schutte, Kelleher, and MacNamee 2009, Cheyer and Julia 1998) however of the 60% of fusion models believed to be at the maturity level, only 2 considered GPS and/or magnetometer input as an input device (Bouchet and Nigay 2004, Bouchet, Nigay, and Ganille 2004, Mansoux, Nigay, and Troccaz 2007). This low number may be the result of the slow adoption of LBS and the sometimes secondary nature of location in mobile applications, however applications where location is the primary input, fusion models need to consider geo-information as a primary input device.

A potential application of passive input events in multimodal systems is in defining context through the use of user position, orientation and trajectory in a 3D space. This approach identifies an area of overlap between the fields of perceptual multimodal interaction and MSI. In multimodal systems, communication channels usually map from

---

[6] The BRETAM model is used to categorise information technologies based on the logical progression of developments in the area in terms of Breakthrough, Replication, Empiricism, Theory, Automation and Maturity.

human output to application input. However, in mobile multimodal interaction this can be reversed. In (Paper IV, Paper VII and Paper III), the use of human visual perception is used to query a virtual space. The use of a user's view frustum or viewport is considered when performing spatial queries on multimedia datasets. In this case the user's visual input modality is mapped directly to the applications input modality in a passive manner. The use of visibility as an input modality has been described in section 3.1 in the context of MSI and has been used as a query mechanism with demonstrable benefits over traditional methods (Frank, Caduff, and Wuersch 2004, Gardiner 2004, Beeharee and Steed 2005, Maierhofer, Simon, and Tobler 2007, Simon and Fröhlich 2007, Bartie, Mills, and Kingham 2008, Bartie et al. 2010, Zaid 2010, Meek et al. 2013).

In AR applications the idea of multimodal, non-tactile interaction is gaining popularity where spatial aspects of human perception are used as interaction modalities (Elepfandt and Sünderhauf 2011, Heidemann, Bax, and Bekel 2004, Lee, Billinghurst, and Kim 2004, Bimber and Raskar 2005, Kölsch et al. 2006, Santos, Lamounier, and Cardoso 2011). In a multimodal context, human visibility is also used as a passive input modality in (Paper I and Paper II) in combination with the active touch modality. The applications described use a user's viewport to interrogate a 3D city map for POIs while enabling users to navigate through content using the touch screen interface. An AUI is used to deliver information in a synergistic way based on touch and visibility-based spatial interaction. Extending this approach, the idea of adaptive multimodal input has been explored (Gentile et al. 2009), where the selection of modality adjusts dynamically based on environmental conditions and thereby improves the user experience. In (Reithinger et al. 2003, David et al. 2011) a conceptual framework and rule-based

approach are used in the automatic adaptation of input modalities based on a user's context. Combinations of user input modalities are selected based on a user's level of attention in different settings (e.g. in car, walking, cycling, etc.). Interesting approaches are also taken in (Zaguia et al. 2010, Kong et al. 2011) where user preferences in different environments and the availability of web-services dictate the active input modalities for a particular situation. This type of adaptive multimodal input selection is also considered in (Paper II) and is the topic of future work outlined in Section 3.9.1. In (Dumas, Solórzano, and Signer 2013), a set of eight design guidelines are presented for adaptive multimodal mobile input solutions. The use of these guidelines is further illustrated through the design and development of an adaptive multimodal calendar application. An interesting approch is also taken in (Ghiani et al. 2014), where an adaptive multimodal solution is presented that provides automatic augmentation of Web applications in such a way as to enable them to exploit various combinations of graphical and vocal modalities. In (Dumas, Signer, and Lalanne 2014), the results of an investigation into software support for the SMUIML multimodal user interaction description language (UIDL) outlines the importance of multimodal interface design and results show that using a triad of tools is effective in the development of multimodal interfaces, addressing the modelling, framework and visual editing levels of the process.

2.5.2. Output Modalities

Output modalities can be described from two main perspectives, the human and the computer. Human output modalities refer to aural, touch, gaze and kinaesthetic[7]. Computer output modalities refer to the ability of a device to deliver output to the user

---

[7] Kinaesthesia refers to the physical position and movements of the body.

using a number of communication channels. These include graphics, language, film and music and are directly targeted at the main human input modalities of vision, auditory and haptic (Bernsen 1997).

With output modalities the original focus was on the presentation of multimedia content via multiple modalities such as graphics and audio. Formative work in (Andre et al. 1993), describes the architecture of a knowledge-based presentation system (WIP) that demonstrated the use of synergistic graphical and textual output for instructional direction. The idea of the automatic presentation of media content based on the coordination of two modalities and a user's context has been a very active area of research for multimodal output (Bouchet, Nigay, and Ganille 2004, Rousseau et al. 2006b, Sinha and Landay 2003). The ICARE (Mansoux, Nigay, and Troccaz 2007) and CrossWeaver (Sinha and Landay 2003) systems take a similar approach to output modalities with multistage processes for the design and production of content and Open Agent Architectures (OAA) to facilitate cost effective modifications of modalities and combinations of them.

More recently, the concept of multimodal fission has been used coordinate multiple output modalities (Rousseau et al. 2006a, Jaimes and Sebe 2007, Perakakis and Potamianos 2008, Costa and Duarte 2011, Honold, Schüssel, and Weber 2012). Fission techniques enable a multimodal system to generate an adequate message in the correct form based on user profile and context. There are usually three stages to the process, content selection and structuring, output channel selection and output coordination. The output channels are based on the affordance of the computing device and can include text, graphics, speech synthesis or embodied conversational agents (Foster 2002).

In a mobile context, fission techniques are very useful since modern mobile devices

provide multiple output modalities such as visual, auditory and haptic. Interestingly, the output channel selection can be based on device affordance and can also be affected by changes in a user's context (i.e. providing audio while driving). Visual interfaces are used effectively to provide text, image and video data on mobile devices (Chittaro 2008, Church, Smyth, and Oliver 2009). The coordination of these modes (audio, video and/or graphics) using the visual modality has been demonstrated effectively in many cases (Chittaro 2008, Jaimes and Sebe 2007, Sarter 2006, Sebe 2009). The auditory modality on the other hand has seen a number of applications mainly related to synthesised speech or recorded speech[8] (Nepper, Treu, and Küpper 2008a, Rajput and Nanavati 2012). Bartie (Bartie and Mackaness 2006) describes the development of a (synthesised) speech based AR system that delivers contextual geospatial information to assists a user in locating landmarks, demonstrating the benefits of focus independent directional instructions. Research in (Doyle, Bertolotto, and Wilson 2009, Kurkovsky 2009) describes the CoMPASS mobile application for the elderly and identifies that natural speech is preferred over synthesised speech, the down side being a significant overhead producing recorded speech assets. The haptic modality has gained significant attention with a myriad of useful applications demonstrated in a multimodal context. Related work by (Robinson, Eslambolchilar, and Jones 2008, Williamson, Murray-Smith, and Hughes 2007, Jacob et al. 2012) investigated haptics (touch) for the purposes of navigation where users are directed through a space using only vibrotactile (the perception of vibration through touch) feedback as a guide.

The benefits of combining output modalities has been identified in many cases (Mousavi, Low, and Sweller 1995, Akatsu et al. 2009, Cao, Theune, and Nijholt 2009)

---

[8] Synthesised speech is artificially produced human speech and recorded speech is the use of natural speech recordings.

as a means of reducing cognitive load on the user. Research has shown that the use of a mixed audio-visual modality is better than the audio modality alone (Hooten, Hayes, and Adams 2013). This indicates that the audio modality combined with the visual modality can be highly effective particularly in a mobile context, where the visual modality alone can be distracting. This type of mixed modality is useful for LBS where users are on the move and in most cases require hands-and-eyes free information delivery (Kaasinen 2003, Dowell and Shmueli 2008).

(Paper I) describes the research and development of a novel Geo-Services platform for the multimodal delivery of high quality and task relevant content to constrained mobile devices (e.g. spatially enabled smartphones). Importantly, it investigates the delivery of location-based content to the user by means of the phone's aural modality as the primary delivery mechanism (Fitch and Kramer 1994, Nepper, Treu, and Küpper 2008b). The visual modality is also presented using a graphical user interface (GUI) that incorporates and builds on many of the ideas developed in previous work on location-based services. A primary objective of this research was to perform a live user trial to evaluate the effectiveness of multimodal content delivery in the form of media content and navigational directions. Results show that multimodal delivery is effective and that user choice in terms of modality combined with contextual instruction (informing them when something is wrong) improves the user experience considerably. Extending this, the idea of adaptive multimodal output was also considered for this application by using the light sensor to detect when the device was "in hand" or "in pocket" to automatically control the primary modality, however this was not employed in the user trial. Similar approaches are taken by the AdaptO (Teixeira et al. 2011) and ProFi (Honold, Schüssel, and Weber 2012) systems where the output modality is determined from user

preferences, environmental factors and device affordance. This work is extended in (Teixeira et al. 2014) with the development of a multimodal personal life assistant that adapts to real-life scenarios to provide multimodal output based on the users needs. This automatic adaptation of multimodal output based on environmental conditions, usage patterns and context has gained considerable attention in recent times, has significant potential and is the topic of future work in the area of adaptive multimodal output, which is outlined in Section 3.9.2.

This Chapter detailed the significant advancements made in the areas of location based information retrieval, hybrid positioning, MSI, AUIs and multimodal interaction. The methodologies described play a significant role in the effective multimodal delivery of content in a mobile context. The main objective was to set the published work in the context of existing literature and to stress the coherence of the publications linking them to the methodologies adopted. The following Chapter will critically describe the published work by establishing how its fits into the overarching theme of multimodal mobile services and evaluate the contributions made in related discipline areas.

3. Results and Discussion

This Chapter covers the results of the research carried out in each publication. It considers the problem, the proposed solution, the actual solution and gives a critical analysis of the result. Finally, the linkages with further associated research are identified.

3.1. Paper VII: Viewer-Based Directional Querying for Mobile Applications

This paper addressed problems outlined in the area of location-based information retrieval using methodologies described in Sections 2.1 and 2.3.
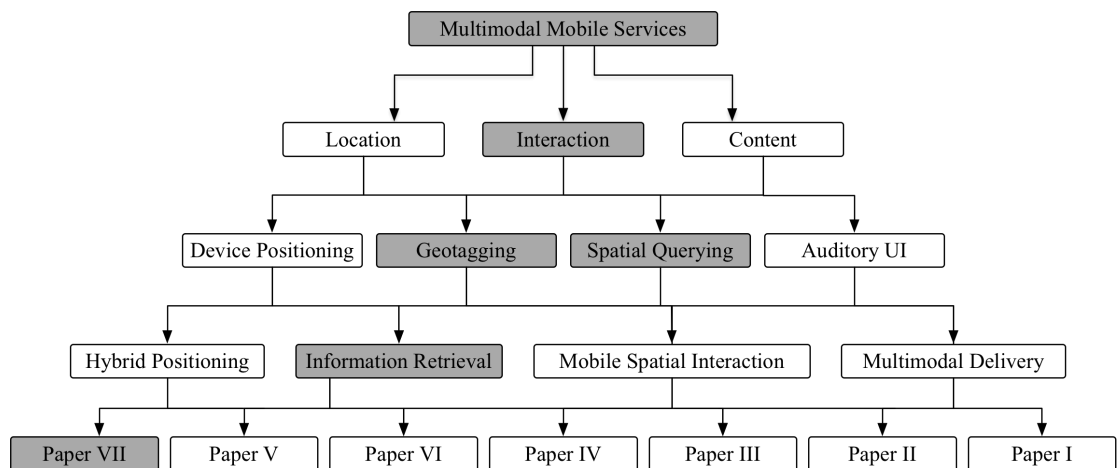
Early LBS applications faced numerous challenges in order to become viable and operate in real-world environments. These challenges ranged from accurately determining a user's location to effectively delivering data to a host device. The need to do research in this area surpassed the emergence of these technologies at a significant pace. In order to perform this work advances in research were required at two levels, at the simulation level and at the mobile technologies level. The latter required advances in infrastructure and mobile technologies, which were very slow to progress, so the former was the focus of the research. In this paper there were two main problems:

Problems

1. Simulation of a mobile environment

2. Development of a directional query processor to function in this environment

Proposed Solution

The solution in this instance was to create a technology demonstrator (3D model of Dublin City) using a virtual reality (VR) modelling language to simulate the real-world environment. Conceptually, the user could navigate through a virtual 3D Dublin Street on a desktop computer and view data relevant to their location in this space. The virtual street contains the facades of buildings constructed and rendered in real-time. As the user moves within this 3D world, their location or virtual space coordinates are concurrently transformed into real world geographic coordinates (latitude/longitude). These lat/long coordinates are then passed as parameters for the construction of a query to the spatial database. The query retrieves all data from the database, within a certain tolerance, of the user's position in virtual 3D space, e.g. retrieve all data within +/-10m of these coordinates. The result of the query is returned to the user's hand held (mobile) device, which is a series of simulated devices (e.g. cell phone, WAP phone, Pocket PC, PDA, etc.), displayed alongside the 3D model on a standard desktop computer.

Problem 1 Contribution

To realise the technology demonstrator introduced above, the main components were implemented in a three-tier web-based architecture typical of spatially enabled enterprise applications, i.e. client layer, application server layer, and database layer.

The client layer consists of a standard web browser to display the VR model and various simulated mobile devices are also displayed but run as separate applications.

The application server layer contains 2 Enterprise Java Beans (EJBs). The session EJB is responsible for all communication with the client. As it is a stateful session EJB, an object of this type is instantiated for each user session. This EJB is responsible for monitoring changes in the user's context. For example, if the position of the user's location in virtual space changes, this EJB transforms the new coordinates into real-world (lat/long) space before passing it to the entity beans. Other relevant context data is also be passed at this time including the type of mobile device currently being simulated on the client and any user profiles that may exist. The SQL EJB is responsible for formulating and executing queries to the database based on the current position of the user. The query results are returned to the session EJB, which organizes the results into (simulated) device specific format. The SQL EJB also formats the query results for a selection of platforms as already described, including WAP, PDA, and a standard web browser. XML is used in conjunction with XSLT (extensible stylesheet language transformations) to convert XML data into the various required formats of the time e.g. HTML, WML, etc.

The database layer (Oracle 9i) provides spatial object type storage, SQL access, spatial operations, and indexing as well as map projections and coordinate systems support. Through this functionality, spatial queries are efficiently executed without the additional overhead of maintaining coordinate information separate from the attribute data. This is accomplished by defining the attribute information as a spatial data type, e.g. a point, line, or area. The primary advantage of spatial data types is that queries can be restricted to a pre-defined geographical area, e.g. within a 10m radius of a given location. By exploiting the spatial indexing mechanisms, which essentially organise the information within the database tables according to their geographic location, all location relevant

data can be retrieved efficiently.

Problem 2 Contribution

A traditional range query is the recognized standard operation for spatial querying based on a window. The query window is of a specified width and height centered on the user's location, which is represented as an optimised rectangle and defined by a minimum of two points *p1* and *p2*. The window is used as a parameter to a spatial query (sdo_relate) and the query is performed against a table of Cultural Heritage (CH) artefacts, which also have a spatial component, typically a point. If an interaction is detected between the window and any points in the database, the id is returned with any relevant data.

In the enhanced implementation, the method of querying the CH database is more sophisticated. In this approach orientation is a necessary parameter so that the user's view-port can be dynamically constructed, resulting in only data contained within the viewers FOV being returned.
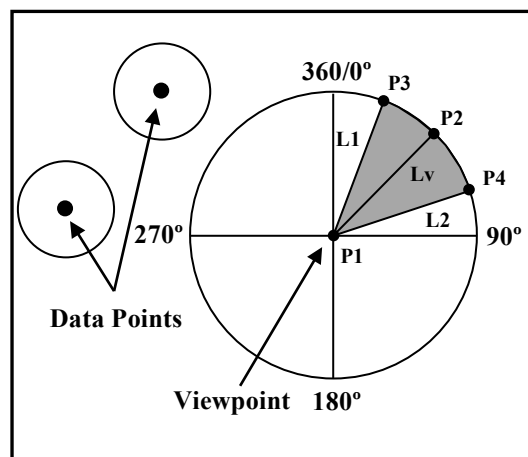
Figure 2: Directional query processor window. *The illustration shows how a user's location and viewpoint are used to create a directional query window for querying cultural heritage (CH) datasets.*

In the directional query model, three points define the triangular query window

34

representing the extents of the user's FOV. The user's viewpoint *p1* is always one of the vertices of the triangle. The points *p3* & *p4* are calculated by first attaining the azimuth from the 0° North direction to the line-of-sight (LoS) of the user, i.e. to point *p2*. (Figure 2) .The azimuth of the line-of-sight (*Lv*) is obtained from a VRML browser. To determine the azimuths to points *p3* and *p4* a specified fraction of an angular FOV value is subtracted from the user's azimuth *Lv* to get the azimuth to point *p3* and by adding the same fraction of the FOV value to *Lv* to get the azimuth to point *p4*. These FOV extents (*L1* & *L2*) are then used to calculate the positions of vertices *p3* and *p4* on a query buffer of specified radius thus giving the view-port a finite distance. Together, the three vertices are used to produce an optimised spatial query shape that will only select data that is inside a triangle oriented in the same direction as the user's line-of-sight.

The next objective was to investigate and develop a line-of-sight algorithm to determine if data contained within the view-port is actually in the viewers line-of-sight. This problem is illustrated in Figure 3. The large triangular area in the diagram represents the user's view-port in 2 dimensions. The brick filled shapes B1, B2, B3, and B4 represent building blocks and D1, D2, D3, and D4 represent CH data points. The enclosed white space in the diagram highlights the desired shape that the line-of- sight algorithm should identify as the query area. The light grey sections represent the areas that should be excluded from the query space, as they are not visible from the user's viewpoint.
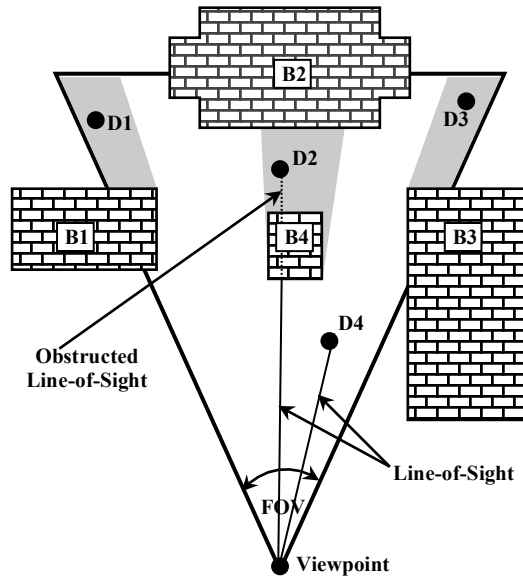
Figure 3: Directional query processor with line-of-sight detection. *The illustration shows how a user's location and field-of-view are used to determine line-of-sight when querying cultural heritage (CH) datasets.*

To determine the LoS for both the CH layer and block layer, a combination of Oracle Spatial operators and a LoS algorithm is required. Our solution was to take a well-known algorithm in Computer Graphics and apply it to the area of Spatial Databases. The *scanline* algorithm was chosen because the topological and Boolean operations needed to process the algorithm are already inherent in the Oracle database schema. The algorithm works by making a progressive scan of the area in question (FOV) to determine whether there are any objects in the scan line path. If so the point at which the scan line intersects the object is recorded. A series of scans is carried out. The end product is the coordinates of a polygon object that represents the search space minus the surrounding building object geometries.

Implementing the *scanline* algorithm (McDonell 1994) for the CH layer is accomplished using the *sdo_intersection* operator in Oracle. First, a procedure was developed to create a series of lines between the viewpoint of the user and each data

point present inside the view-port. For the CH layer, these lines are considered as the *scanlines*. In turn each of the *scanlines* are used as the input parameters to the *sdo_intersection* operator, to determine if they interact with any of the objects in the block layer. If there is any interaction between any of the objects in the block layer and the *scanline*, the CH artefact is not visible to the user from that viewpoint and it will not be placed in the LoS resultset. If there are no interactions between the CH data point and the block layer along that *scanline*, the data point is considered to be visible and is placed in the LoS resultset.

Discussion

The implementation of the simulated environment was a success overall. It enabled the demonstration of location-based information retrieval albeit in a virtual environment as limitations with mobile device hardware of the day could not compute sophisticated LBS processes. Key elements associated with the process such as using a user's location to query a spatial database in real-time were successfully demonstrated. Furthermore, testing the directional query processor and line-of-sight processor wouldn't have been possible in a live situation without accurate position and azimuth variables. Coincidently, this is also one of the drawbacks of this research. The level of accuracy provided by the simulated environment was unrealistically high and this level of accuracy was unachievable from GPS and digital compasses. The demonstrator performed very well as a result of this but this was not applicable in a real-world context.

Linkages

This research identified 3 main areas of interest that required further investigation in order to deliver mobile LBS in real-time. Research in the area of mobile connectivity

and synchronisation, hybrid positioning for improved positioning accuracy and MSI techniques for improved spatial query results, were being considered. The most critical one at that time was research in the area of mobile connectivity and spatial data management due to the necessity to demonstrate LBS in a real-world context.

1. **Mobile Connectivity and Synchronisation (Paper V)**

2. Hybrid Positioning (Paper VI)

3. Mobile Spatial Interaction (Paper IV)

The following section describes research work carried out in the area of mobile connectivity and synchronisation with a focus on spatial data management and information retrieval.

3.2. Paper V: A Web-Based and Mobile Environmental Management System

This paper addressed problems outlined in the area of location-based information retrieval using methodologies described in Section 2.1.
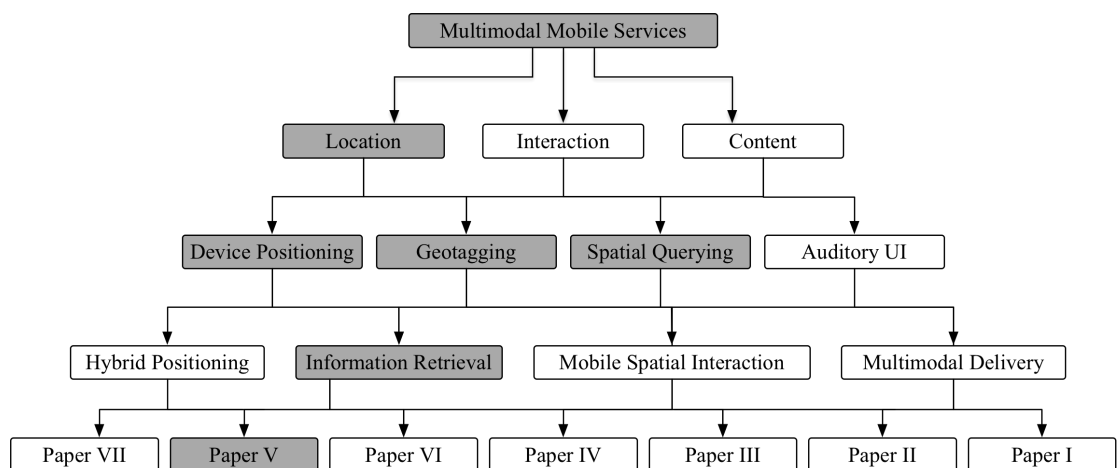
Following the successful demonstration of location-based information retrieval in a simulated mobile environment, the need to demonstrate these techniques in a real-world context was important. In addition, another central aspect of LBS research that was

important at that time was spatial data management. Therefore, the decision was taken to carry out research in the area of mobile application development with a focus on the effective use of geo-positioning for LBS. Fortunately the opportunity arose to apply some of the techniques demonstrated in (Paper VII) in a real-world environment. The decision was made to apply these techniques by replacing a paper-based data collection and analysis system for the Great Lakes Laboratory for Fisheries and Aquatic Sciences (GLLFAS) Fish Habitat Management Group at the Canadian Department of Fisheries and Oceans (DFO). This enabled research to be carried out in area of geo-positioning, mobile connectivity and spatial data management for LBS using a real-world data set.

In order to address the issues surrounding mobile connectivity and real time positioning, research was also carried out in the area of offline caching and synchronization. The focus in this instance was effective data collection in a physical environment. Due to the lack of reasonable mobile Internet connectivity at the time, alternative methods of recording data in the field were investigated. In parallel, state of the art mobile devices did not contain the requisite hardware to provide enough power to run a GIS type application or GPS hardware to provide accurate positioning. Therefore, the most suitable platform for this task was a tablet PC. This paper details the two most relevant research problems associated with this task:

Problems

1. Development of a web application for the collection of context sensitive geo-information

2. Development of an offline mobile application for the collection of context sensitive geo-information

Proposed Solution

The proposed solution to these problems centered on the use of a tablet PC as the client that communicated with a middleware server application hosted on the Internet. This application would serve as the interface between the client and database layers of the system. The requirement of a tablet PC was twofold. Firstly, the application would be used on boats; a rugged tablet is more suitable in this environment and secondly, a GPS receiver was needed to provide positioning in order to locate the device and geotag data for upload. In addition, an offline mode was proposed that would record and cache data in areas with no network connectivity (a common problem), that could be synchronized once Wi-Fi connectivity was available.

Problem 1 Contribution

The solution was the development of a Web-based data management system tailored to deliver context-aware functionality aided by visualization, analysis and manipulation of spatial and attribute datasets. The datasets were provided by the Canadian Department of Fisheries and Oceans (DFO) and the prototype customized to the specific needs of the Great Lakes Laboratory for Fisheries and Aquatic Sciences (GLLFAS) Fish Habitat Management Group. The functionality required by GLLFAS biologists included access to geo-referenced maps and imagery, to overlay the current position on a map and to manipulate attribute data in the field while wirelessly connected (where possible) to the office database. Additional functionality also required was the ability to record, edit and view multimedia annotations, perform scientific/common name conversion and graph generations of results. The prototype uses a typical three-tier architecture composed of the client layer, application server layer, and the database layer. The client is a tablet PC with GPRS (General Packet Radio Service) data connectivity and GPS module for positioning. The application server was an *iSmart* application server with offline

synchronization capabilities and the database was an Oracle Spatial database. The client application provides a real-time, geo-referenced mapping interface that enables the user to view, input and modify data as well as to annotate maps and photos in a variety of ways. It also provided an advanced query interface to enable the user to query the database in a number of ways.
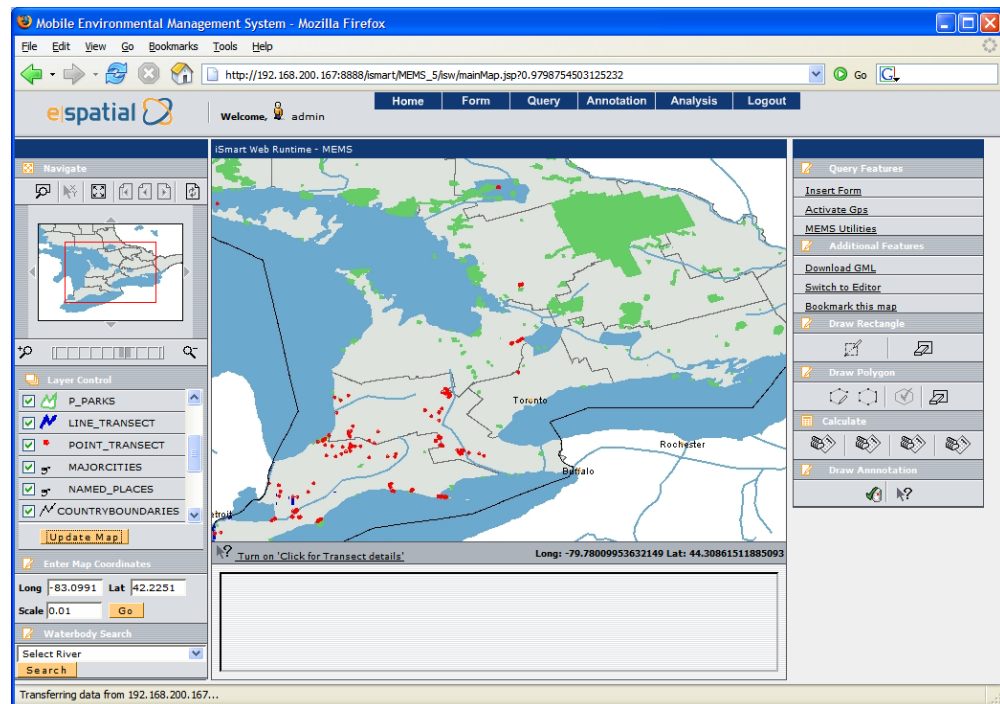
Figure 4: Web-based interface of the MEMS system. *The MEMS Web interface was a dual-purpose application designed to display/add maps of catch data and to perform analysis in the field using a tablet PC.*

Problem 2 Contribution

It was observed that some areas the biologists sampled in had only intermittent cellular signal, if at all. As it was difficult to detect or predict network availability before heading to the field, the offline module was designed. The module is a standalone application that connects directly to the online application server back in the office where the biologist is asked to select an area of interest. The selected area is compressed into a zip file that contains all the data required by the field biologist and the application (Figure 5). This offline application enables the user to work remotely when a GPRS

connection is not available. In the field, the offline application reads in the compressed file and displays the map, enabling the biologists to perform spatial queries, insert new forms, look at previously recorded forms, display GPS feeds and navigate the map. It also offers restricted multimedia annotation functionality. Data that is changed or added during the sampling session is stored and saved in the offline file. When the application establishes a network connection and can re-connect to the application server, the data is synchronized with the server and any additions are added to the database.
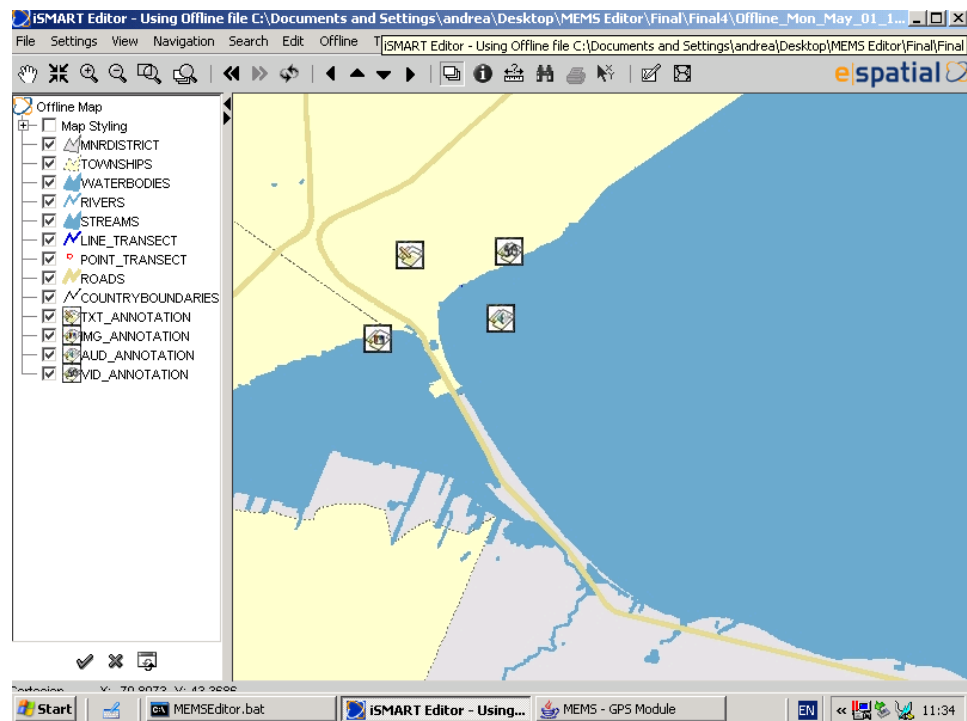
Figure 5: Offline interface of the MEMS system. *The offline interface is a Java application used for surveying areas without GSM network reception. Specific areas can be downloaded prior to sampling and then synchronised with the server afterwards.*

Discussion

The MEMS project was a good opportunity to take an operational paper based system and develop a digital solution using state-of-the-art mobile technology in the form of a tablet PC and web-based management system. The application was very successful and could be used online and offline to record and analyse spatial datasets in real-time.

Unfortunately, one of the major drawbacks of this technology was the technology itself. GLLFAS scientists were given the option to use either system and most defaulted back to the paper based system. The reason for this was mainly the inability to use the tablet PC. In fact, users were not only expected to familiarize themselves with the concept of a tablet PC but also learn how to use advanced GIS functionality in the field, in sometimes-harsh conditions. This was a drawback that had not been planned for in the projects design phase that identified for the first time issues with usability in mobile applications. Further research was considered to solve some of these issues, mainly in the area of familiar mobile interfaces. Another drawback of the solution was the GPS. In open areas, the positioning accuracy was acceptable enabling the GLLFAS scientists to record and analyse data in the field. However, in built-up urban areas such as a city river, this was not the case. The derived location became erratic presenting significant difficulties while using the software. This was mainly due to the urban canyon effect, reducing visibility of the sky and hence blocking the satellite signal essential for accurate position determination.

Linkages

This research identified 2 main areas of interest that required further investigation in order to deliver effective services in both urban and rural areas. Given the issues identified in urban areas with GPS, hybrid positioning was a potential solution to this problem because of the increasing density of Wi-Fi beacons in these areas. Another focus was the usability of mobile applications and how this could be improved.

1. **Hybrid positioning for mobile devices (Paper VI)**
2. Usable interfaces for mobile devices (Paper III)

The following section describes research work carried out in the area of hybrid

positioning for mobile devices with a focus on harnessing existing RF beacons to improve positional accuracy.

3.3. Paper VI: An Open Source Approach to Wireless Positioning Techniques

This paper addressed problems outlined in the area of hybrid positioning using methodologies described in section 2.2.
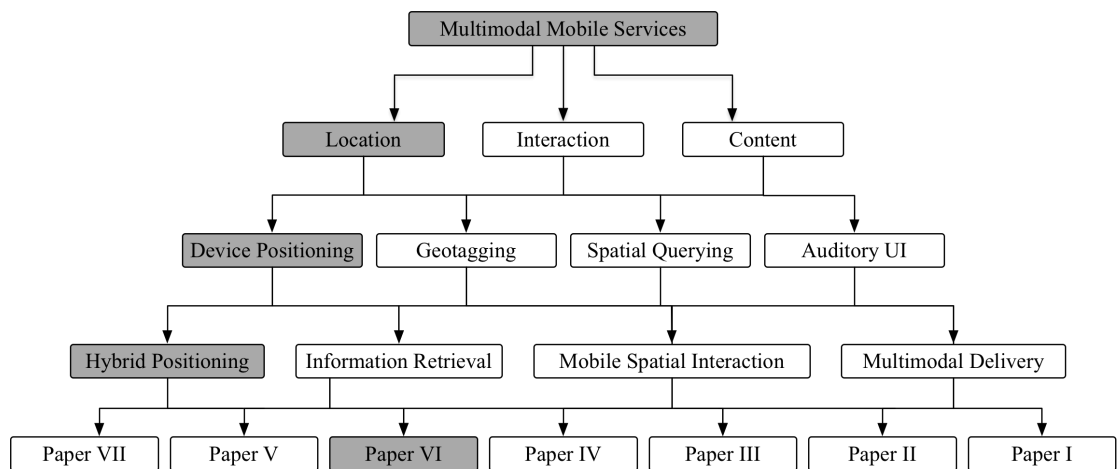
Several problems exist when trying to determine accurate positioning on a mobile device. The first issue and somewhat significant one is the location itself. Unlike various other forms of wireless communications, GPS actually performs better in rural areas. In urban areas with a high density of buildings, the GPS signals required by a mobile device to trilaterate a position are partially blocked, affecting the accuracy of the position calculations and in some cases (less than 3 satellites visible), the GPS receiver is unable to determine a position at all. This is called the urban canyon effect (Hofmann-Wellenhof, Lichtenegger, and Collins 2001) and in 2006 was a major barrier to mobile app development. Smartphone technology was starting to emerge with the communications and user interface hardware required to provide LBS, however GPS was still an issue with no real solutions apart from Galileo's promise of sub-meter

accuracy over a decade away (EU/ESA 2006). There were some solutions such as assisted GPS (Djuknic and Richton 2001) but this was not a possibility on COTS devices, so the focus turned to other radio frequencies that could be utilized. Mobile devices had GSM, Bluetooth and Wi-fi hardware that could be utilized to support GPS. This lead to research in the area of Hybrid positioning, where the use of a combination of GPS, GSM, Bluetooth and Wi-fi signals were used to accurately determine a mobile devices location in urban areas.

Problems

1. Development of a Hybrid Positioning module
2. Production of a beacon database for Dublin

Proposed Solution

The problems described above were part of a larger undertaking that involved the implementation of a multi-modal, multi-access concept for e-Government (ICiNG 2006). The idea of a *Thin-Skinned City* that was sensitive to both the citizen and the environment through the use of mobile devices, universal access gateways, social software and environmental sensors was the focus of the research. The idea of intelligent infrastructure that communities interact with to avail of services created by the administration, and to create their own information-based services was investigated. Dublin, Barcelona and Helsinki were set up as *City Laboratories* for researching, evaluating and demonstrating technologies and services using intelligence in the environment.

One of the fundamental requirements for this type of service interaction was location. In Dublin, one service to be provided was an *issue tracker* that enables citizens to report accessibility issues (e.g. lack of wheelchair access) to Dublin City Council. This

required location as part of the report and the module that would provide this location is the ICiNG Location Client (ILC).

Problem 1 Contribution

The ILC was by design an open source, network independent, location determination mobile application that can utilise GPS, Wi-Fi, GSM, Bluetooth and Semacode information or any combination of them, to calculate location. The research task was to address positioning issues and introduce the next logical step for freely available mobile positioning, advancing the pioneering work done by Place Lab at Intel (LaMarca et al. 2005). The ILC integrated all the above location finding technologies into one positioning module. With all these technologies finally available on one device, it was now possible to employ a personal positioning system that could work effectively in any environment. Another important advantage of the ILC was its ability to do this without any direct communication with outside sources, resulting in reduced communication costs and increased privacy because it did not actively connect to any external network or other services to trilaterate its position. The ILC was developed as a standalone module to run on a Series 60 (3rd Edition) mobile phone running the Symbian operating system (version 9.x), although other platforms and operating systems could be accommodated with relatively minor changes.

Two of the key components of the ILC module were the RMS database (WiGLE 2007) of known beacon locations (updated through a process called Wardriving) and the Tracker module. The ILC used a series of spotters (one for each RF technology) to scan for active RF signals. On detection, the hardware ID for the beacon was used to search the RMS database for the last known location of the beacon. The Tracker module was responsible for collecting beacon information from the four spotters, organizing the

beacon information into local databases, trilaterating the current position of the mobile device based on this data and responding to location requests from other mobile device applications.

When collecting the beacon data, a set of rules quickly identified the degree of accuracy to be expected depending on the beacon type. Although all beacons in the RMS Database were read, it first looked for Bluetooth beacons and if any were found, the degree of accuracy would be <10 meters and if there were no Bluetooth beacons, GPS was used and hence the degree of accuracy would be >10 meters.
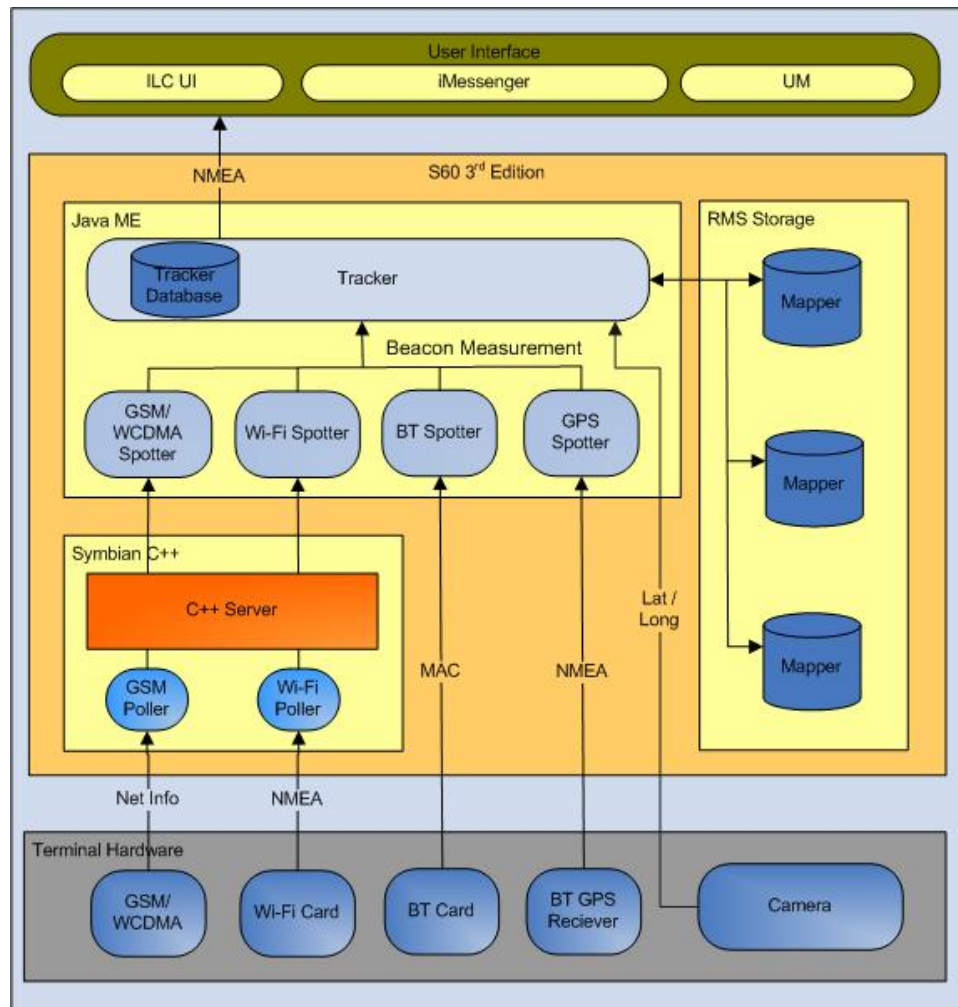
Figure 6: The ILC module design for hybrid positioning on Symbian S60 OS. *The ILC uses C++ spotters for GSM and Wi-Fi detection and Java spotters for BT and GPS detection. Known beacon databases (Mappers) are used to match hardware identifiers.*

Problem 2 Contribution

The operation of the ILC was solely dependent on the use of an RMS database of known beacon locations. At time of writing, a database existed for Dublin that was 60% accurate at best (Hightower, Lamarca, and Smith 2006). This was an issue but fortunately it was possible to contribute to the database using an activity called Wardriving. Wardriving involves driving around the affected area with a laptop running Wardriving software that detects RF signals and tags the hardware IDs with GPS locations. These locations are then uploaded to a central database online for others to download and use (www.wigle.net). As part of the ILC development, significant periods of Wardriving were undertaken to improve the coverage for the centre of Dublin hence increasing the accuracy of the position calculations provided by the ILC. This was an essential activity that contributed to the effectiveness of the applications using the ILC as a source for accurate location information.

Discussion

Development of the ILC was an essential step in the advancement of mobile device software for the support of LBS. Location dependent apps are essentially non-functional without the provision of accurate positioning. The ILC utilized existing RF infrastructure to improve the accuracy of device positioning by incorporating land based beacon locations into position calculations. Although effective at doing this (after a considerable contribution to the beacon database as part of the research), there is one caveat, and that is the reliance on the beacon database as a whole. Because the beacon data are collected by Wardriving using techniques based on signal strength, a number of issues arise due to multipath propagation where the actual beacon position determination process is affected by signals bouncing of buildings, etc. This means that the source of the position data can be inaccurate which affects the ILC calculation. In

addition, the mobility of beacons can also be an issue. As most of the beacons are actually Bluetooth and Wi-Fi means that they can be moved. Although the ILC can detect the type of Bluetooth device spotted, be it a phone or situated device, it is not possible to determine when a Wi-Fi antenna or GSM tower has been moved. This requires a new Wardriving session to re-scan the area and update the RMS database. This is a typical example of how open source software can be difficult to rely on because of the ambiguous way it is managed. Detecting and dealing with these types of issues is fundamental to the success of open source hybrid positioning technologies.

Linkages

The developments in the area of hybrid positioning for location-based information retrieval opened up a host of possibilities for research that relied upon accurate positioning to become viable. Two such areas were MSI and mobile visibility querying. In terms of MSI there was significant potential in this area because of the improved positioning accuracy achieved in (Paper VI), which would improve the accuracy of spatial query results also. There was also considerable interest around the area of visibility querying which would help increase the relevance of data returned to the user.

1. **Mobile Spatial Interaction (Paper IV)**

2. Mobile Visibility Querying (Paper III)

The following section describes research work carried out in the area of MSI with a focus on the accurate detection of POIs using a range of visibility techniques to improve spatial query search results.

3.4. Paper IV: EgoViz –A Mobile Based Spatial Interaction System

This paper addressed problems outlined in the area of MSI using methodologies described in Section 2.3.
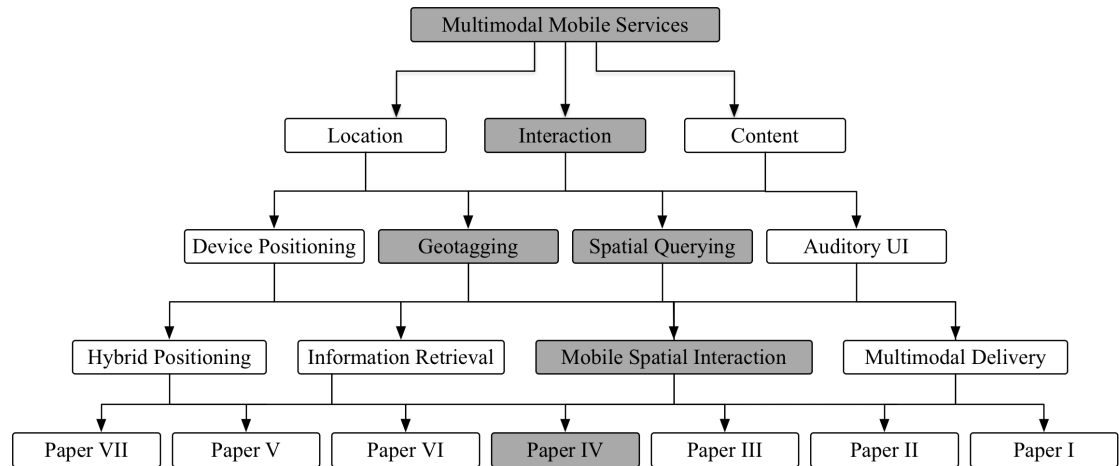
With advances in mobile technology and improvements in hardware such as GPS, digital compasses and gyroscopes, the possibilities for mobile applications that could utilize these technologies became apparent. Ideas developed in previous research in the areas of directional querying (Paper VII) and hybrid positioning (Paper VI) were now a possibility in real-world terms without the previous barriers meaning applications could now harness this data to provide innovative applications based on position and orientation. This advancement also meant that prior research in the area of mobile application development (Paper V) using multi-layered architectures could be progressed. Using a combination of expertise gained from prior research and with the significant advances in mobile devices and communications, early examples of mobile LBS could now be demonstrated in a real-world setting.

The *EgoViz* system was developed to demonstrate these ideas working together using a typical 3-tier architecture. The idea was to query a search space based on egocentric visibility. The client (mobile device) would supply the position and orientation of the

device to the application server layer, which would in turn query a spatial database based on these values. These queries had only been demonstrated in a simulated mobile environment before (Paper VII) using very accurate parameters. The challenge in this instance was to provide accurate position and orientation values to a query processor and to efficiently query the underlying spatial database for points/buildings that interacted with the user FOV. There were two main research problems that needed addressing in order for these queries to become a possibility.

Problems

1. Demonstrate accurate position and orientation determination on mobile device

2. Demonstrate mobile spatial interaction (MSI)

Proposed Solution

To perform directional queries in a mobile environment requires accurate location and orientation values, a directional query processor and a spatial database. The solution was to develop a mobile application that would comprise two main components, a positioning component for calculating and delivering accurate position estimates and an orientation module responsible for managing and delivering accurate orientation estimates. These modules would provide these values to a directional query processor (middleware) component of the system hosted on an application server. The directional queries would be processed at the database level and the results of any interacting points or building would be identified and returned.

Therefore, given that we had an underlying 2D map available stored in an Oracle database, which included POI information, the two main areas of interest were the mobile app development and the query processor.

Problem 1 Contribution

In the case of the *EgoViz* system the proposed solution was to use experience gained in prior research (Paper VI) in hybrid positioning to build a positioning module. The locator component or *LocateMe* module was an open source, network independent mobile location determination application that can utilise GPS, Wi-Fi and GSM beacon information or any combination of them, to trilaterate location estimates.

The orientation component or *DirectMe* module uses data from a number of sensors including a GPS sensor, a magnetometer sensor (digital compass) and an accelerometer sensor (tilt sensor). Its function is to determine the direction that the mobile device is currently pointing by using a digital compass and tilt sensors. This data is then collected on request from the *TellMe* application and synchronised with the location data coming from the *LocateMe* module. The architecture of the *DirectMe* module is similar to the *LocateMe* module where each technology (i.e. - compass and tilt sensors) has a native hardware spotter that relays data to a higher-level component synchronising data from each spotter. This architecture is illustrated in Figure 7.

The module described uses a native Symbian S60 component to interface directly with the device hardware using a local sensor server. This server then feeds the PyS60 component, which synchronizes the data and makes it available to the *TellMe* application.
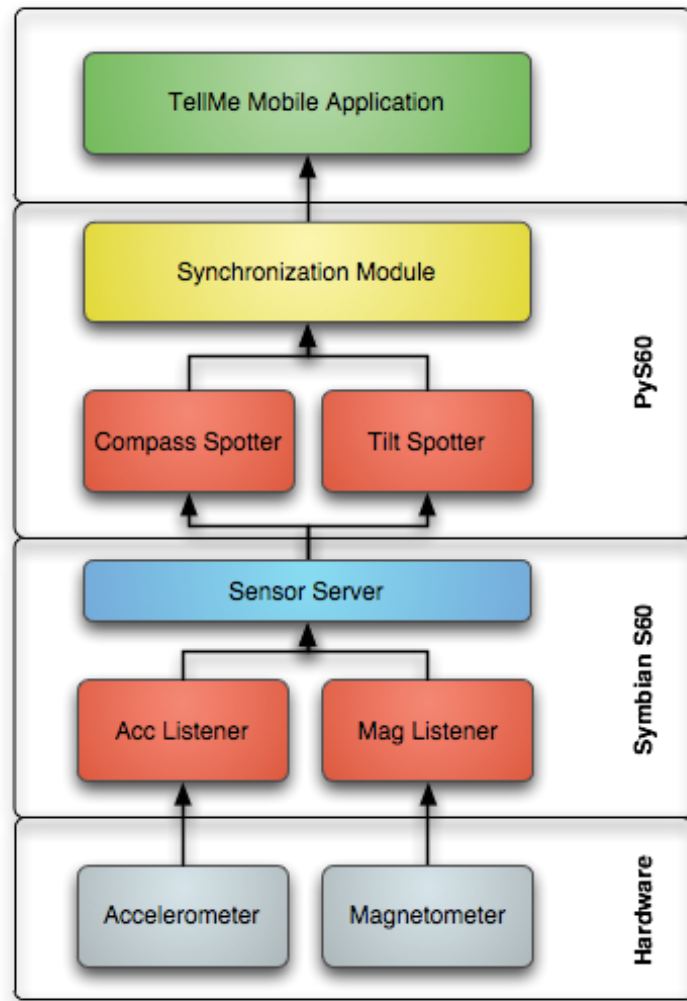
Figure 7: The architecture of the DirectMe mobile module for orientation. *The DirectMe module consists of a Symbian S60 hardware interface component and a PyS60 synchronisation module that communicates with the TellMe application.*

Problem 2 Contribution

To demonstrate MSI, the *TellMe* server is used to perform all the complex spatial queries in the system. It is responsible for communicating with the *TellMe Mobile* application, which collects data (including location, direction and orientation) from the *LocateMe* and *DirectMe* modules on the mobile device. This data is communicated wirelessly to the spatial application server and used to perform spatial queries against the *Oracle Spatial* 3D Database.

The *TellMe Server* is based on ESRIs *ArcGIS Server* platform and is used to perform

the complex queries that are required to determine the MSI between the user's line-of-sight and the 3D database. This platform provides server extensions for many of the traditional spatial query functions found in most GIS. Using these extensions, it is now possible to perform these types of spatial queries in a mobile context that were previously only possible in a desktop setting. Another important reason for this choice of server is based around the issue of scalability; using this software ensures that the system is scalable and can manage a large number of users efficiently (assuming optimal network conditions). This is illustrated in Figure 8.
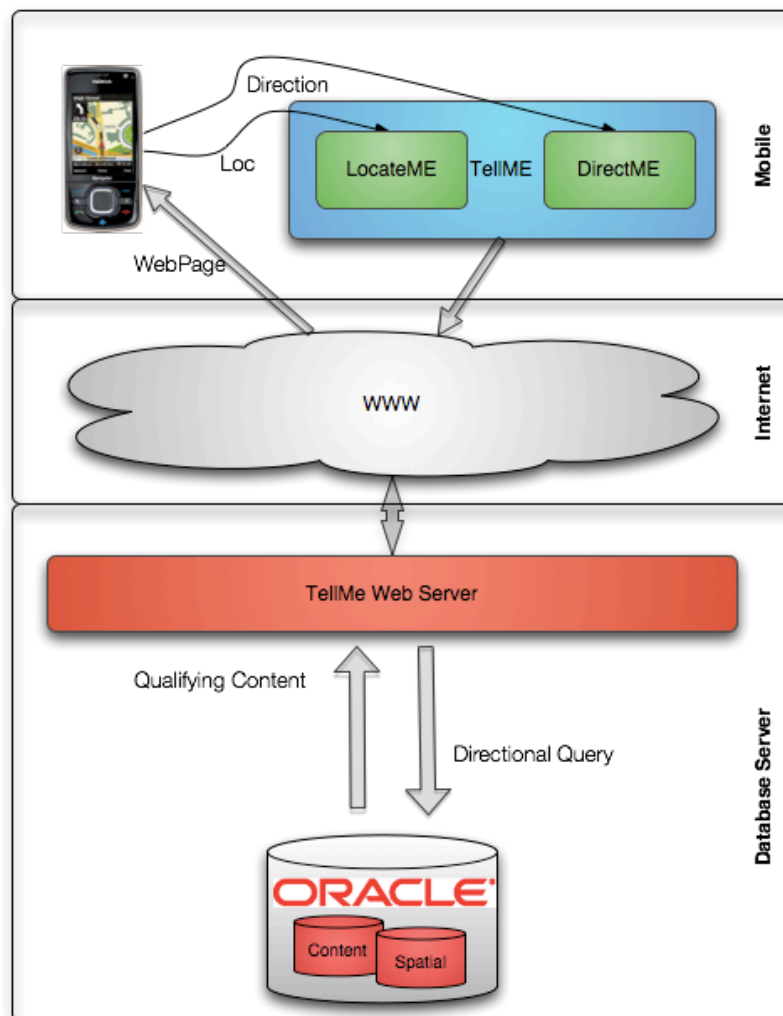
Figure 8: The architecture of the TellMe system. *The diagram shows the 3-tier architecture used with the LocateMe and DirectMe modules on the client and the TellMe Server arrangement.*

The *TellMe* Web Server is responsible for performing all queries in the system, all of which are based on the user's egocentric point-of- view. Egocentric visibility refers to the portion of a search space that is visible to a user at a particular time based on their location, direction and orientation. For example, in the case of the *TellMe* application, the user's visible query space acts as a secondary filter on data that is returned by the query processor by restricting it to contain only objects that are in the user's FOV. The FOV therefore excludes the portions of the dataset obscured by buildings. This method is primarily used to identify POIs other than buildings that are in the user's FOV within a predefined distance from the user. This method can be used to identify objects in the distance that may be too small to point at directly but are still in the user's FOV nonetheless, like a monument or statue. The algorithm to determine the searchable space in this instance builds on previous work outlined in (Paper VII) where a 2D directional query processor was developed and used in a virtual environment for similar purposes.

Discussion

The *TellMe* system successfully demonstrated egocentric visibility in a MSI context. The system delivered accurate results when tested dependant on optimal network conditions and minimal hardware interference. Tests demonstrated that directional queries performed best overall (~1 sec), with FOV queries, isovist queries and frustum queries taking significantly longer to process (~3 sec). This was due to the complexity of the back end processing required for the latter.

Although the *TellMe* system performed well in relatively protected environments, one of the major drawbacks of the mobile application was the effect electrical devices had on the accuracy of the hardware readings. In an area where there is significant electrical

interference such as air conditioning units or server rooms, the reliability of the mobile devices hardware was compromised. The noise detected on the signal was extremely high, making the reading unusable.

In addition to these problems, the *EgoViz* application user interface was minimal with no use of a map to display results; this again was a significant drawback of the mobile application with only text-based data being returned from the server. Significant improvements were required in this area, which would require more network bandwidth and an improved user interface design.

Linkages

This research was a milestone in the area of MSI. With improved positioning accuracy and the ability to accurately determine the interaction between a user and their environment using a COTS device opened up possibilities for research in two main areas. The architecture developed paved the way for further research to be performed in the areas of mobile visibility querying and multimodal user interface design. Mobile visibility querying had potential in dealing with the issue of information overload by considering FOV as an input mode. Multimodal user interfaces were also of interest in terms of content delivery as the issue of information overload could also be addressed by using the aural and visual modalities together.

1. **Mobile Visibility Querying (Paper III)**

2. Multimodal User Interfaces (Paper II)

The following section describes research work carried out in the area of mobile visibility querying for LBS. The focus is on accurately determining users FOV in 2D and extending it to 3D in order to filter spatial query results based on a user's egocentric perspective.

3.5. Paper III: Mobile Visibility Querying for LBS

This paper addressed problems outlined in the area of MSI using methodologies described in sections 2.3.1 and 2.3.2.
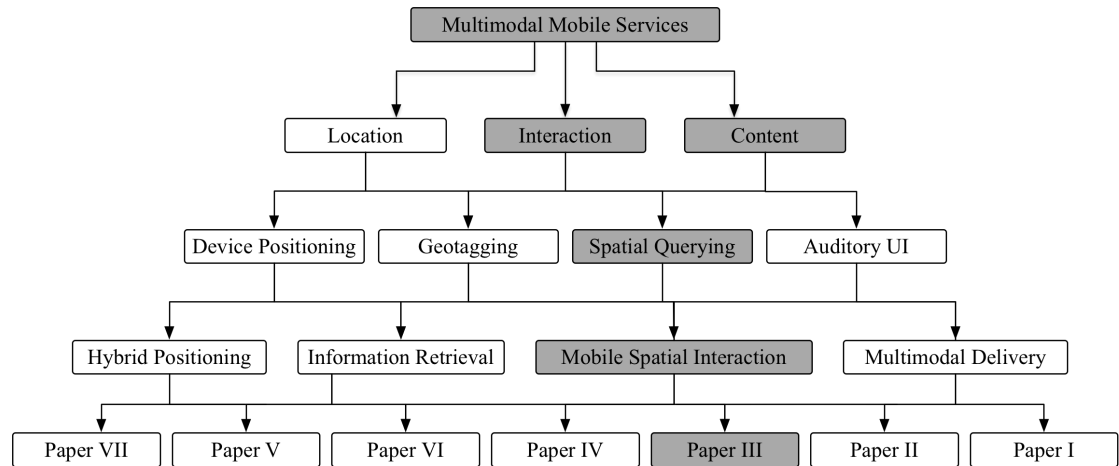
With advancements in mobile technology and the increased access to mobile operating system APIs that provided hardware access to third party applications, a new wave of mobile applications started to emerge. These applications harnessed the availability of location data, cardinal directions (compass headings) to provide essential location based services primarily focused on the 2D plane (Paper IV). These applications were capable of locating a device, performing 2D range queries and calculating a user FOV for visibility querying. The natural progression from the 2.5D approach taken with these applications was to extend to 3D, where the user can point a mobile device at a building or a particular floor for example and retrieve information. The subsequent availability of accelerometer hardware (that could also provide g-force acceleration values) meant that 2D orientation data could be extended to the 3D domain. This meant that a more effective spatial query could be performed and point-to query operations could be realized. The nature of these queries also meant that a more effective means of representing these queries and the results was necessary.

This research was primarily concerned with the effective representation of visual perception in the form of visibility analysis. To achieve this, two main problem areas were identified that required further investigation.

Problems

1. Development of a 3D querying platform for LBS
2. Demonstration of an effective UI for visualizing queries/results

Proposed Solution

The approach in this instance was to extend previous research carried out the area of MSI (Paper IV). There were 2 main aspects to this research that progressed the state-of-the-art. The first was the investigation of 3D queries as a method interacting with a space and retrieving relevant information based on this interaction. The primary objective here was to effectively create a dynamic 3D vector that could be used to query a 3D dataset based on the movements of the user device. This could then be extended to include view shed and threat dome configurations. The underlying aspect of this research that enabled these types of queries was the ability of the spatial database to process them. Up to this point, spatial databases could store 3D data, but the capacity the query in 3D was not possible. With the release of Oracle 11g, the storage, indexing and querying of 3D spatial objects using the *sdo_geometry* data type was now possible for the detection of 3D topological relationships. To perform these queries from the client, the proposed solution was to develop SOAP and Restful web services that would expose this functionality to supported mobile devices.

The second aspect to this research was to develop a usable graphical user interface in the form of a 2D map for visualization of queries and the results, i.e. the building/floor that the device was pointed at or the POI's currently visible to the user based on

egocentric visibility. This would take the form of an interactive map with dynamic vector overlays that would allow the user to see and control the extent of the user's view shed for example, and view the results consisting of building names or identifiers.

Problem 1 Contribution

The 3DQ prototype for location and orientation aware mobile devices was developed as part of this research. The entire query/retrieval process was performed on the server-side, with the primary parameters required; location, direction, and tilt provided by the mobile device. The types of 2D queries available were: standard range queries (*all neighbours* and *nearest neighbours*); single ray directional queries (*point-to-select*); full 360° queries (*Isovist*) and directionally constrained (*field-of-view*) queries with hidden query removal (HQR) functionality. In the 3D domain, *point-to-select 3D* tells a user, not only which building the device is pointing at but also which floor or even the particular window on that floor. In 3D, the elevation of the user is taken into account along with the heights of the buildings. If a user is pointing over a lower building to a higher one behind, the 3DQ processor is capable of recognizing this difference. If a user requires detailed information about individual objects in their 3D FOV, a *frustum view* query is available to generate the required query shape and retrieve the corresponding data to the device. A frustum view query can be thought of as a "squared" flashlight beam scanning the wall of a building to get information about whatever gets "illuminated". Finally, a *threat dome* query provides a 360° isovist view in three dimensions out to a specified radius. The described query processes are all implemented by generating the respective query shapes as 3D objects in a spatial database and then utilizing inherent 3D query operators to identify topological relationships.

The data exchange between the client and the server is based on either a SOAP or RESTful style web-service. A WebLogic server is used for deploying the services and providing the interfaces for the mobile devices where the request from a mobile device and response returned are wrapped in an XML document – thus allowing portability across mobile platforms. The architecture of the system is illustrated in Figure 9.
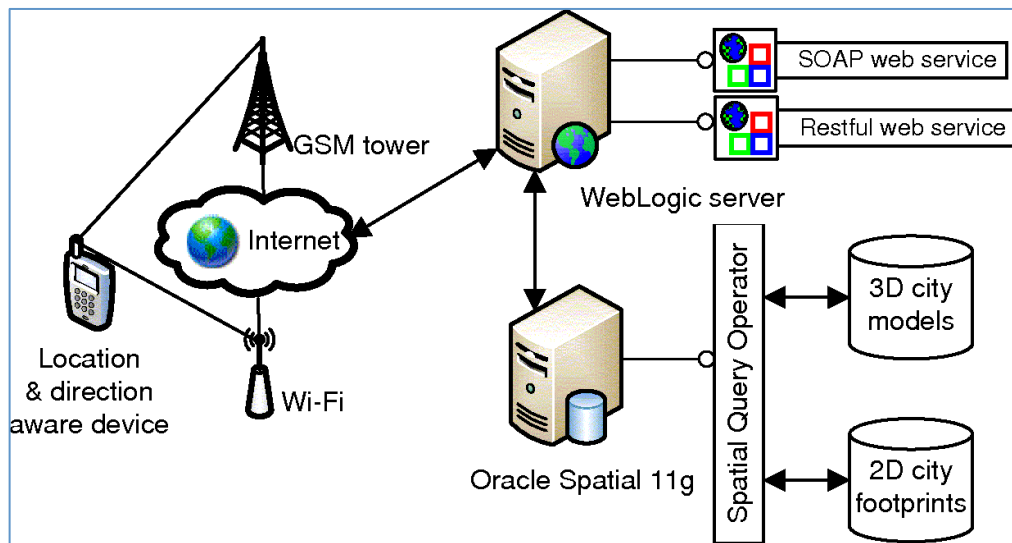
Figure 9: 3DQ Architecture diagram for egocentric visibility querying. *The 3-tier architecture used focuses the load on the server side to increase performance and improve mobile application usability.*

Problem 2 Contribution

Egocentric visibility excludes the portions of the dataset obscured by buildings and is primarily used to identify POIs other than buildings that are in a user's FOV within a predefined distance. These types of query are used to identify objects that are too small to point at precisely from a distance but are still in the user's FOV, like a monument or other tourism artefact. The algorithm to determine the query space in this instance builds on our previous work in (Paper VII) where a 2D directional query processor was developed for an on-line virtual environment. To successfully visualize this Google Maps was used as the background map for the query results. The returned result are

overlaid on the map in the form of building outlines or POIs to offer the user an opportunity to adjust the pointing gestures and re-query if necessary. The mapping interface on the mobile device is shown in Figure 10, where the arrow indicates the position and direction a user is pointing and the colour represents the status of the calibration level of the sensor signals.

Figure 10: Mapping Interface for 3DQ processor on Nokia 6210. *The Google Maps interface helps users to perform spatial search queries using visual indicators such as position and orientation and visualises query results improving usability in both cases.*

This dynamic interface was used to visualize the spatial queries in order to accurately perform mobile spatial search in 2D and 3D. The results were then subsequently displayed using the native UI elements. This was a significant advancement from the basic interface of the *EgoViz* mobile app, which delivered text based results with no visualization of the query.

Discussion

The 3DQ prototype demonstrated effective visibility based queries for mobile devices. The innovative line-of-sight and hidden query removal functionality produced a more

accurate query result by only returning information about objects visible within a user's 3D FOV. The innovative user interface was the first step towards delivering a rich user experience on a mobile device. However, the overall process was server heavy so that the thin client could function well with the limited responsibility of collecting sensor information and visualizing the results. The main drawback was query speed, which varied based on the quality of the data connection and the complexity of the queries themselves. This is due to the fact that all query processing and content is carried out on the server, requiring multiple requests over a network. This could be improved by performing some queries on device but would affect the client performance. Although state-of-the art, the Nokia device was not a very effective when it came to usability. The Symbian operating system was awkward and difficult to program with restricted access to hardware requiring multiple modules to integrate to deliver the required level of functionality. In addition, the delivery modality was mainly restricted to visual mode with limited scope for other forms of content delivery.

Linkages

This research was a milestone in the area of 3D querying of spatial datasets and the visualizing of content on a mobile device. The need for a more effective way of delivering information using multimodal interfaces was evident based on the limitations of the Symbian platform as a whole. The need for more innovative ways of conveying information based on a user's interaction with a space was also pertinent. This led to further research in the area of AUIs and multimodal content delivery.

1. **Auditory User Interfaces (Paper II)**

2. Multimodal Content Delivery (Paper I)

The following section describes research carried out in the area of AUIs with a focus on

content delivery for geo-services. The aural modality is used in combination with the visual modality to deliver rich media content. Innovative methods of delivering content with a minimal footprint enable all content to be stored on device. All spatial interaction queries are also performed on device minimising the requirement for mobile data connectivity.

3.6. Paper II: mobiSurround: An Auditory User Interface for Geo Service Delivery
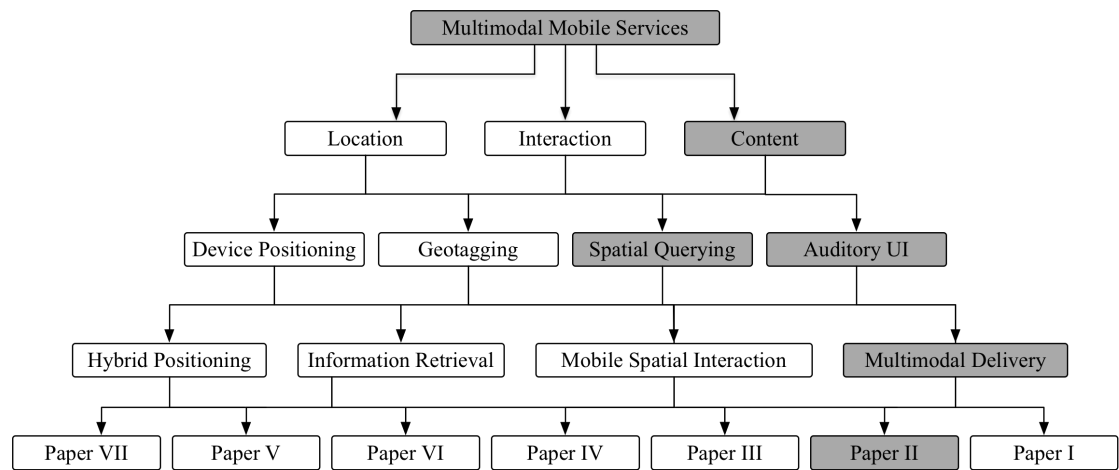
This paper addressed problems outlined in the area of AUIs using methodologies described in Section 2.4.

Following a considerable body of work investigating novel applications of LBS and MSI, the need for an increased level of quality of media content and some new ways of delivering this content became apparent. This led the research in a new direction and the idea of an efficient media content delivery engine that could deliver geo-referenced media in an immersive way, using multiple modalities while reducing data redundancy, was the new focus. Two of the most pertinent problems identified within this area were information overload at UI level and database redundancy. The use of a restricted GUI in the form of a map and native interface components had been used before (Paper III) for navigation and general information delivery with limited scope for the latter. One

approach to alleviating this issue is to separately visualize map and content data although this is not always ideal and can lead to confusion. Another approach is to use multi-modal content delivery where the GUI is supported by other types of interfaces such as haptic or aural.

After some investigation and an analysis of human perception (McGee and Cullen 2009a), the aural modality was identified as the best approach to take, as it was focus independent and could compliment the GUI. Other options such as haptic were also considered but were found to be more focus dependent and hence a less suitable approach. One significant drawback to the aural modality however was content size. To produce audio that could be used to effectively deliver content for all eventualities is a substantial undertaking and would create a sizeable amount of redundancy. To address this problem, the idea of using phrase synthesis to alleviate this problem was considered based on previous work in the area of speech analysis and avatars (Cullen et al. 2009). Using a real-speech phrase synthesizer would considerably reduce redundancy, enable content storage on device thus reducing network overhead in the process.

Problems

1. Development of a data/content modelling strategy for LBS

2. Development of an Auditory User Interface (AUI)

Proposed Solution

The platform chosen for this research was the iOS platform. This was a significant upgrade in terms of hardware quality (GPS, Compass, Accelerometer), processing power and screen size. The proposed solution was based around the idea of presenting geo-referenced information both visually and aurally to the user in a collective way. The visual modality would be presented using a graphical user interface (GUI) in a state-of-

the-art mobile app that incorporated many of the ideas developed in previous work in the area of MSI (Paper VII, Paper IV and Paper III). The aural modality would be presented using an AUI, the main innovation in terms of delivering context sensitive information using focus independent (eyes free) means. The idea of adaptive multimodal selection would also be explored based on device affordance and environmental conditions.

There were two main parts to this solution. Firstly, the idea of the AUI, which is instrumental to delivering information to users in a clear and natural way and secondly, the structure of the content itself. In many cases, the amount of redundant data required by the system was very high. Because the content required by the system is mainly audio, the opportunity to work on reducing some of this redundancy was considered by introducing the concept of phrase synthesis. Phrase synthesis is based on the idea of taking a dictionary of recorded words and creating phrases from them by combing them into sentences *on-the-fly*. This helps maintain a high quality of audio using real-speech assets with the potential to deliver the same content but with a significantly reduced footprint, thus reducing the overhead of the app considerably. Further research would be carried out in the area of multi-channel content delivery to produce effective background audio and soundscaping. Combining these, the AUI for geo-services would deliver an immersive experience using real speech assets, background music generated on-the-fly, transient sound effects and natural directions.

Problem 1 Contribution

The *mobiSurround* engine is an AUI for LBS that models context and location to provide a number of exciting new innovations in the areas of content delivery, mobile navigation, and narrative context. The central component of the engine is based on a 3-

stage data modelling process that informs the various parts of the application that feed both the visual and aural presentation modalities. The Geodata Modelling (GM) task is a process by which all geographic data that defines POIs, navigational waypoints and regions are collected. The User Modelling (UM) task is where user preferences are modelled, essentially defining the type of user and their context. This data has a scaling effect on the content modelling strategy, where varying amounts of content is presented based on a user's profile. The Content Modelling (CM) task is where the structure of the media content is defined for each node based on the geo-data (GM) and application preferences data (UM). It is this strictly controlled process of data modelling from multiple perspectives that ensures that data in the system is structured correctly, enabling the dynamic loading of media content in a coherent manner (Figure 11).
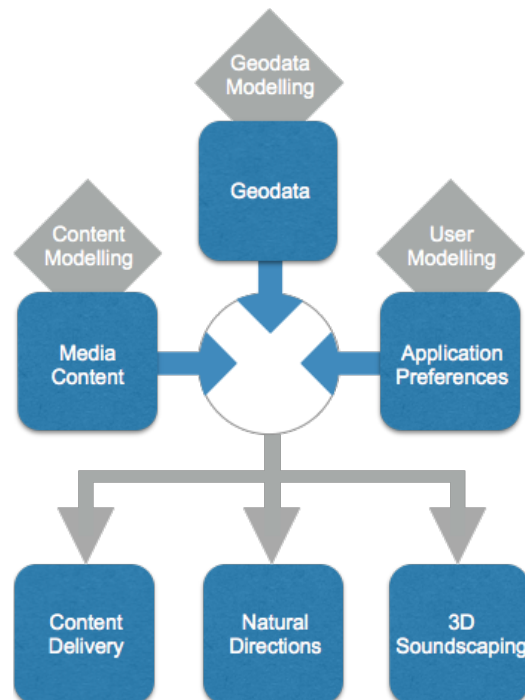
Figure 11: mobiSurround 3-Stage Data Modelling Process. *Geodata, media content and user preferences are used to structure data for delivery by the AUI.*

The *mobiSurround* concept focuses on the idea of adaptive narrative that may be used in (though not limited to) tours and exhibitions, where the user may experience a space in

their own way, in any direction and in any order. This has been achieved in *mobiSurround* with the development of a node/waypoint/region model for the mapping of data in a given environment (Figure 12).
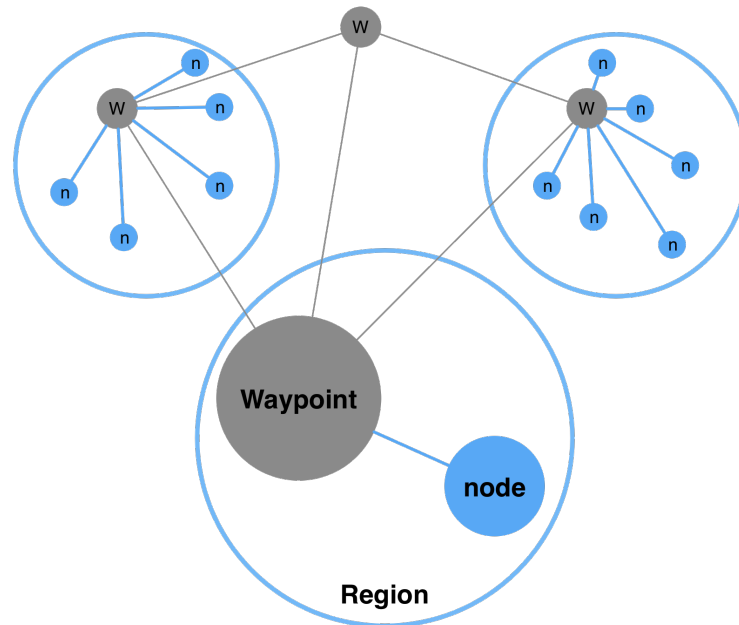
Figure 12: mobiSurround node/waypoint/region hierarchical structure. *This structure is used to organise content in order to effectively deliver audio content, soundscaping and natural directions via the AUI.*

A node is defined as a logical POI, service, or facility, with waypoints being used for navigation between points in the space. Regions represent logical shifts in the narrative of the space, such as moving between groups of exhibits in a tour. These basic elements are the foundation of all data modelling activities in the system. They are explicitly linked to the processes that construct the audio for the purposes of content delivery, navigation and narrative context.

Problem 2 Contribution

The *mobiSurround* engine models context and location for content delivery, mobile navigation, and narrative context. A demonstrator application was developed (Dublin Zoo) that uses an intelligent content phrase synthesiser, algorithmic music sequencer

and adaptive virtualised 3D soundscaping for content delivery and a natural directions model for real-time navigation. Content delivery is based on a background to foreground listening paradigm where speech is prioritized over soundscaping and music and is always foreground content (Figure 13). A spatialised 3D soundscape is used to simulate environmental aspects of the current region in the background. This is overlaid with algorithmic music composed from melodies, rhythms, and harmonics that are randomly selected and sequenced at runtime.
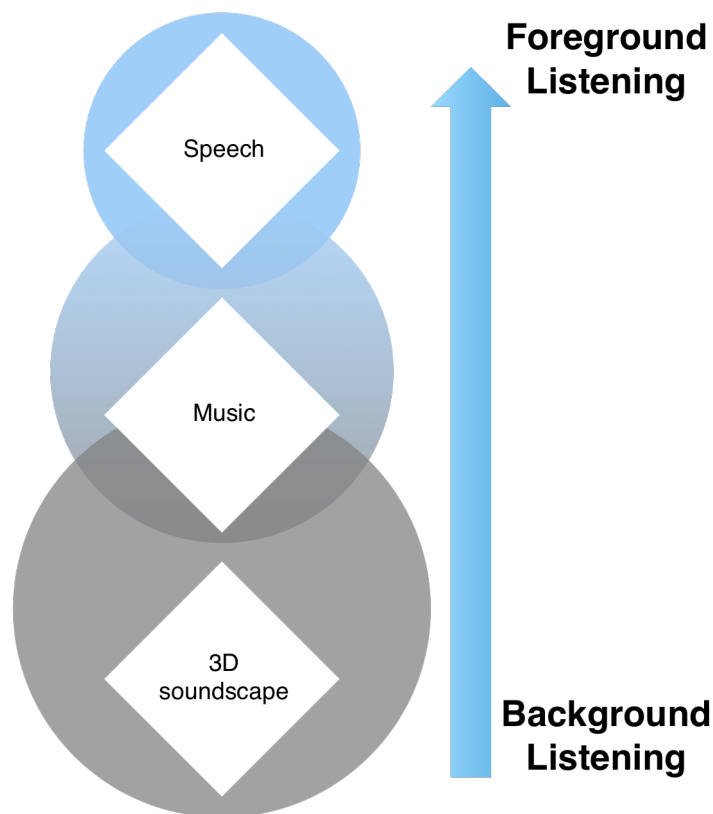
Figure 13: mobiSurround background to foreground delivery model. *The listening model shown illustrates how content is delivered to the user prioritising speech over music and 3D soundscaping.*

One of the main contributions is the intelligent phrase synthesizer used to deliver node content. The synthesizer was developed both for natural directions and due to application size considerations that preclude large amounts of audio content being

included in a download. The phrase synthesizer is used in all node content delivery to allow a relatively small set of commentary phrases (delivered by the narrator) to be reused for a given node based on a set of linguistic rules. In this manner content for each node is assembled at runtime to give the impression of listening to a documentary interview between a narrator and expert.
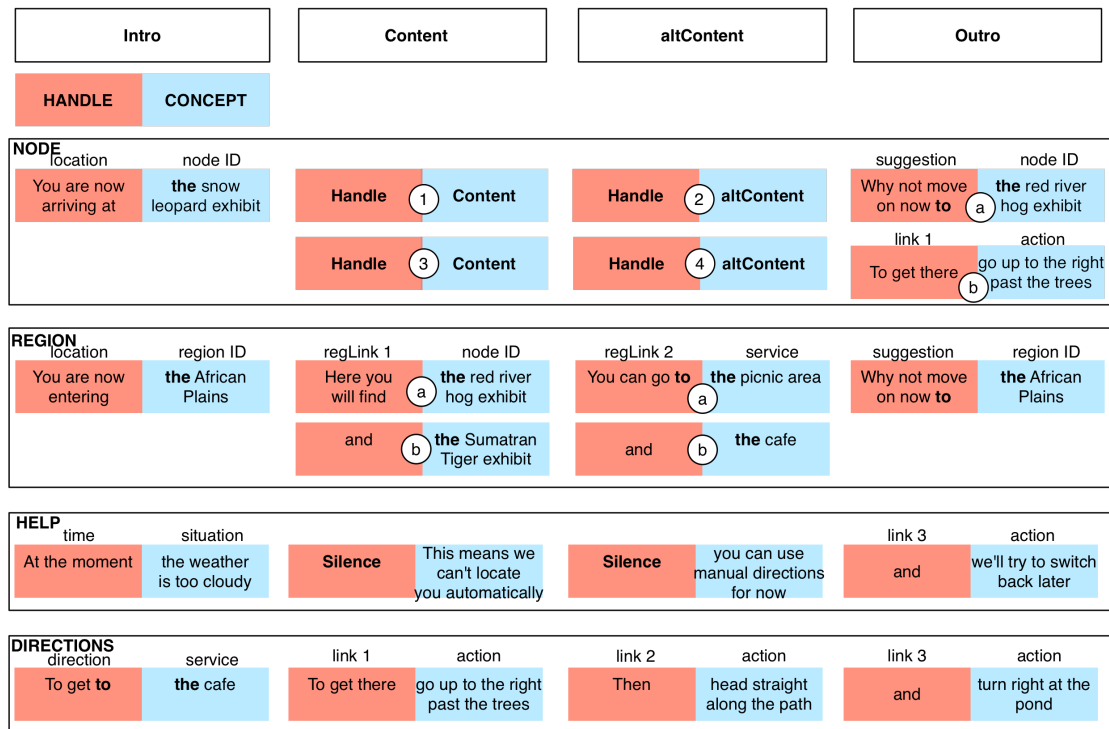
Figure 14: Phrase synthesis structure for node, region, help and directions content. *The illustration shows how the phrase synthesizer assembles node, region, help and direction content sentences in real time for playback by the AUI.*

Each node has two narrators associated with it, the main narrator and a domain expert. Facts 1 and 2 are initially loaded into the Content and altContent busses, and then they are swapped out for Facts 3 and 4 (Figure 14). The phrases are constructed in real-time and loaded into the *AUGraph* where they are organised sequentially, initialised with the required bus control data from the database and played in order (Table 1).

Table 1: Phrase structure for content delivery. *This structure is used to deliver audio content to the user and is based on a narrator led interview with a domain expert.*

[You are now arriving at] + [the Snow Leopard exhibit] + [Lets ask john to give us some information] + [Snow Leopards are large cats that live way up high in the mountains ranges of central Asia] + [Wow that's interesting] + [FACT2] + [Cool can you give us another fact] + [FACT3] + [Interesting, what else can you tell us about it] + [FACT4] + [Well now you have heard about it lets move on to] + [the red river hog exhibit] + [DIRECTIONS]

The conditional logic that controls the *MixerBus* is described in the decision table below (Table 2). When a new event is triggered (e.g. new poi detected), a set of conditions is checked and the associated actions are performed to construct and deliver the audio.

Table 2: mobiSurround engine decision table. *Decision table describing the rules, conditions and actions that control loading, management and playback of the MixerBus AUGraph.*

		Rules									
Conditions	New POI Detected	Y		Y			Y		Y		Y
	New Region Detected		Y	Y			Y		Y		Y
	New Event Occurring				Y						
	Directions Request					Y	Y				
	Random SFX Change							Y	Y		
	Random Music Change									Y	Y
Actions	Prepare speechURLArray	X		X	X		X		X		X
	Prepare musicURLArray		X	X					X	X	X
	Prepare SFXURLArray		X	X				X	X		X
	Prepare naviURLArray		X	X		X	X		X		X
	FadeOut Channel 12-15	X			X	X					
	FadeOut Channel 0-3										
	FadeOut Channel 4-11										
	FadeOut MixerBus (0-16)		X	X			X		X		X
	FadeOut Channel Index							X		X	
	swapSpeechStructArray	X		X	X				X		X
	swapMusicStructArray		X	X					X		X
	swapSFXStructArray		X	X					X		X
	swapNaviStructArray		X	X	X	X	X		X		X
	swapRandomSFXStructElement							X			
	swapRandomMusicStructElement									X	
	FadeIn Channel 12-15	X			X	X	X				
	FadeIn Channel 0-3										
	FadeIn Channel 4-11										
	FadeIn MixerBus (0-16)		X	X					X		X
	FadeIn Channel Index							X		X	

Natural directions are a recent innovation in navigation technology, using visible points in a space to give directions rather than providing distances and cardinal points. The

advantages of natural directions are speed (humans do not need to translate data) and adaptability (there is no need to hold a compass in a certain way to obtain a heading) that allow them to be delivered hands free. The disadvantage of natural directions is the size of the content model required, where each direction is a unique narrative in itself.

To address this, the natural directions model uses a combination of node and waypoint information to control the phrase synthesizer, providing a set of natural directions (which can also be displayed on a map if required) to any given point in the space. The model uses the concept of waypoint linkage to create a lookup table of directions between waypoints that is then completed by directions from the last waypoint to the node required. In this manner, natural points of focus within the space are first defined as waypoints and used to script and create the directions. When a specific node is queried, the only information required is its nearest waypoint, which allows a lookup table to complete the navigation path to it and subsequently configure the phrase synthesizer.
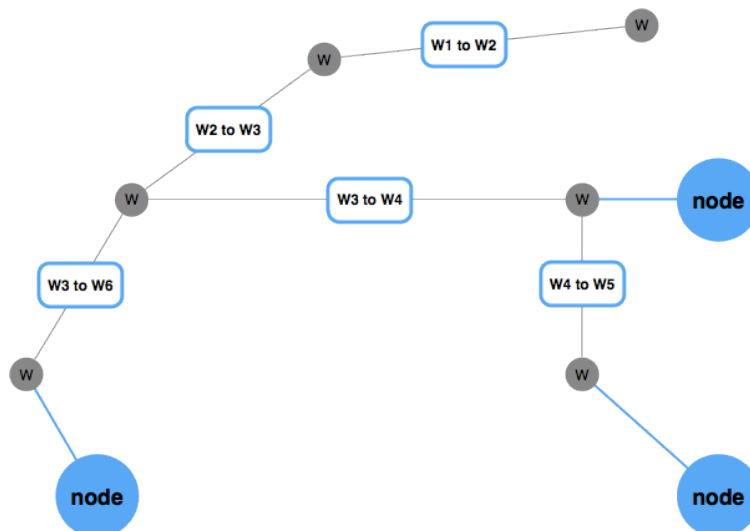
Figure 15: Waypoint Linkage Example. *In the illustration, each node is owned by a waypoint; directions are given between waypoints and each waypoint contains direction to each of its nodes.*

Figure 15 demonstrates how waypoint linkages are used to provide directions to any node in the space. For the three destination nodes (e.g. zoo exhibits), a common set of waypoint linkage directions are used to get to waypoint 3, where different waypoint combinations then lead to the specific nodes. Each node is associated with one waypoint only. When a direction query is performed, three pieces of information are needed.

(1) Directions from the current note to its waypoint
(2) Directions from the current waypoint to the destination waypoint
(3) Directions from the destination waypoint to the destination node

By building a set of waypoint linkages in both directions, (getting from W2 to W1 is not the same as going from W1 to W2) each node is linked to its nearest waypoint, which is also linked in a lookup table to every other waypoint. The phrase synthesizer is then loaded with the entries from the lookup table and provides the directions to the user as required (Table 3).

Table 3: Natural directions waypoint lookup table. *This table is used to determine the files required by the MixerBus to assemble and play natural directions between destination nodes.*

Waypoint	1	2	3	4	5	6
1		WP1 to WP2	WP1 to WP2 WP2 to WP3	WP1 to WP2 WP2 to WP3 WP3 to WP4	WP1 to WP2 WP2 to WP3 WP3 to WP4 WP4 to WP5	WP1 to WP2 WP2 to WP3 WP3 to WP6
2	WP2-WP1		WP2 to WP3	WP2 to WP3 WP3 to WP4	WP2 to WP3 WP3 to WP4 WP4 to WP5	WP2 to WP3 WP3 to WP6
3	WP3 to WP2 WP2 to WP1	WP3 to WP2		WP3 to WP4	WP3 to WP4 WP4 to WP5	WP3 to WP6
4	WP4 to WP3 WP3 to WP2 WP2 to WP1	WP4 to WP3 WP3 to WP2	WP4 to WP3		WP4 to WP5	
5	WP5 to WP4 WP4 to WP3 WP3 to WP2 WP2 to WP1	WP5 to WP4 WP4 to WP3 WP3 to WP2	WP5 to WP4 WP4 to WP3	WP5 to WP4		WP5 to WP6
6	WP6 to WP3 WP3 to WP2 WP2 to WP1	WP6 to WP3 WP3 to WP2	WP6 to WP3	WP6 to WP3 WP3 to WP4	WP6 to WP3 WP3 to WP4 WP4 to WP5	

Table 3 illustrates the structure of the lookup table used to determine directions between waypoints. The variable lists of linkages (step 2) are used in combination with steps 1 and 3 outlined above, to deliver effective directions using the phrase synthesiser. For example, if the user wants to navigate from node 6 to node 5, they first need directions from node 6 to waypoint 6, then directions from waypoint 6 to waypoint 5, and finally directions from waypoint 5 to node 5. Directions between nodes and their waypoints are named accordingly and are loaded automatically. The structure used to construct the audio for directions between the 2 nodes is shown in Table 4.

Table 4: Phrase structure for directions delivery. *This structure is used to deliver directions to the user and is based on directions from nodes to waypoints and a lookup table of directions between waypoints.*

[To get to] + [Node 6] + [N6 to WP6] + [Then] + [WP6 to WP3] + [Then] + [WP3 to WP4] + [Then] + [WP4 to WP5] + [and] + [WP5 to N5] + [I hope this is clear, if not request directions again]

The example above illustrates how the phrase synthesiser dynamically constructs directions using randomly selected handles combined with a series of directions segments. The result is loaded and rendered by the AUI, and animated visually on the map. The GUI can be used to help aide navigation in the event the user feels lost. The location of the user and the location of the closest node is visualised on the map. The user is also able to manually select nodes, ask for directions and control playback using this interface. Using multimodal interfaces ensures that the user can navigate a space effectively, avoid uncertainty and have a good user experience. For example, when the user is interacting with the GUI, the AUI content is controlled to avoid delivering redundant information and to reduce cognitive load. The GUI for the *mobiSurround* prototype app is shown in Figure 16.

Figure 16: mobiSurround Graphical User Interface. *The graphical user interface is used to visualise the exhibits and to manually control the app in poor geo-positioning conditions if necessary.*

Discussion

This new work produced a spatialised AUI developed for LBS (*mobiSurround*). It was achieved after several advances upon the current state-of-the- art, notably in the areas of location/data modelling and adaptive content production.

As a technology demonstrator, the app delivered a high quality user interface for tour delivery. The methodologies employed in the data modelling process and the development of the AUI served only as a foundation for further work. The most significant drawback to the approach taken was the level of manual data modelling and content modelling that was required. To replicate the app for a different context would require an equal amount of time to model the data and content. This was a major issue meaning tours took a significant amount of time to produce. In addition a

comprehensive user trial is needed to test and provide feedback in terms of functionality and usability of the AUI for delivering content. Development of an authoring tool for planning, structuring, and combining content is also required. This tool will help streamline the process of audio production as outlined in the data-mapping matrix and listening model.

Linkages

With the AUI engine developed, the focus turned to the concept of producing tours using an authoring tool, where the geo-data could be used to effectively define the tour and content could be attached using a predefined structure. This led to the concept of the 3-stage process, survey, author, and deliver for effective tour production.

1. A Platform for Multi-modal Content Delivery (Paper I)

The following section describes work carried out in the area of multimodal content delivery with a focus on content organisation and a user trial for geo-service delivery.

3.7. Paper I: MobiLAudio - A Multimodal Content Delivery Platform for Geo-Services
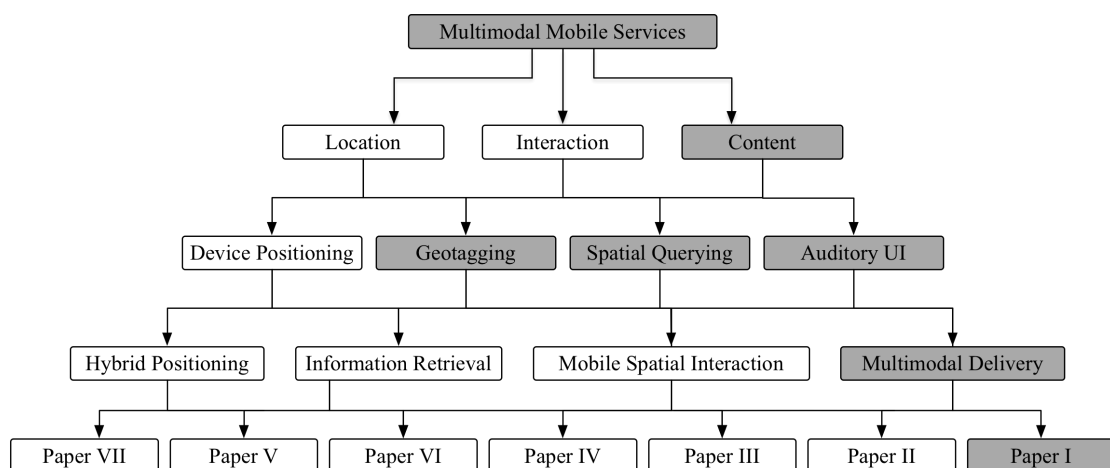
This paper addressed problems outlined in the area of multimodal interaction and AUIs using methodologies described in Sections 2.4 and 2.5.

With major advancements in the area of mobile and communications technologies in the form of powerful devices such as the iPhone and Samsung Galaxy series, effective LBS is now a reality. A growing number of services are available that provide context relevant information on demand. Given the power and capabilities of these devices the ability to deliver these services in a multi-modal context is also becoming more pervasive.

Previous research (Paper II) demonstrated this multimodality in the form of a GUI and AUI for LBS. The challenges encountered were significant, most notably in the area of data modelling and content management. This was due to the fact that the content was authored on the fly from a corpus of audio assets and rendered by the AUI phrase synthesizer or music sequencer. Although the AUI functioned correctly and performed well, delivering audio content, natural directions and synthesized music in real-time, formal testing was not carried out. To determine if this novel approach of an AUI for LBS was a success, a user trial was required to test the application in terms of functionality and user experience. To perform the user trial, a second app would be required that was independent of the initial implementation. This would again require significant work in the area of data modelling. To address this issue, the idea of an authoring tool was explored that would help minimize the amount of manual modelling required. This was also reinforced by the results of a feasibility study carried out in the area of mobile tours, which indicated that there was significant interest from the business community about the concept and in particular the idea of user generated tours (Brennan 2012).

Problems

1. Development of platform for geo-services production

2. User trial for multimodal geo-service delivery

Proposed Solution

The proposed solution in this instance was the development of a novel Geo-Services platform for mobile multimodal content delivery. The visual modality would be presented using a graphical user interface (GUI) that incorporates and builds on many of the ideas developed in previous work on LBS (Paper III, Paper IV, Paper V and Paper VII). The aural modality would incorporate an AUI for delivering context sensitive information to the user using focus independent (eyes free) means. The idea was that all information in the application (including media content, directions, advertisements, etc.) would be delivered without the user ever needing to touch the device. An adaptive approach would also be taken to output modality selection by using the light sensor to detect the location of the device (i.e. in hand or in pocket) to control content and provide a better user experience. There are four main parts to this research, the three app development stages and the user trial.

Survey Stage – uses a mobile survey tool for collecting region specific geo-information.

Author Stage – an authoring process that focuses on the structure of the content and how the overall process can be optimized.

Delivery Stage – uses the mobile app for delivery of content to users in a clear and natural way using a GUI and/or an AUI.

The fourth and arguably most significant part of this research was the live user trial that

was performed to test the functionality, user experience, and overall usability of the app.

Problem 1 Contribution

The multimodal content delivery platform (*MobiLAudio*) introduced in this paper is a toolset for audio tour production that extends the work carried out by the *mobiSurround engine* (Paper II). The *MobiLAudio* platform provides a toolset where professional HD tours can be planned, authored and delivered. The following three sections describe the tools and processes used in each of these stages.

Survey Stage

The core concept of the *MobiLAudio* platform is the synergy of location data and professionally produced media content. To achieve this, a mobile survey tool was developed to provide a full LBS map of a given space and to geo-tag content for effective tour production. The tool allows a tour creator to annotate specific locations within a space by attaching rich media in the form of text, images, audio and video (Figure 17). The geo-data collected by the survey tool are used to define the three core elements of a tour space, and are referred to as nodes, waypoints, and regions.
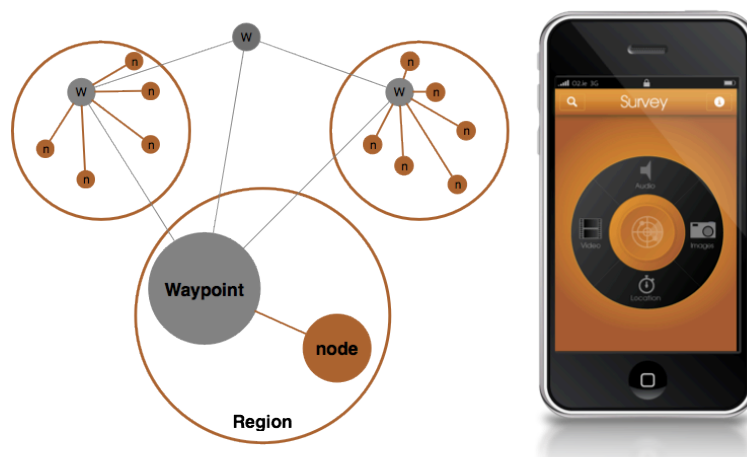
Figure 17: MobiLAudio surveying tool. *A mobile app used to survey a space before the content production stage; it provides the geo-data for nodes, waypoints and regions.*

Authoring Stage

With a content map in place, the 3D audio soundscape, algorithmic music and interactive narrative content for nodes and regions within a given tour must be composed and structured to trigger accurately as the user navigates the space. This is done using the *Authoring Tool*, which is used to view the survey information for a given tour, helping to inform aesthetic decisions on soundscape, music, and narrative for all points in the tour (Figure 18).
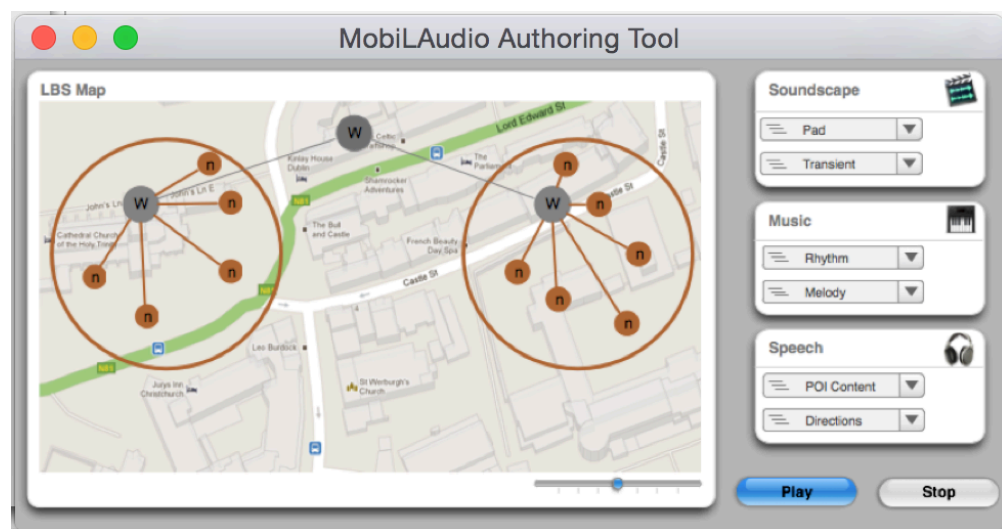
Figure 18: MobiLAudio Authoring Tool. *This tool is used to organize audio assets for nodes, waypoints, and regions by linking the assets for each node to geo-data collected by the survey tool.*

A data delivery matrix is used to structure the data where a node is defined as a logical POI, service or facility. Waypoints are defined as navigational reference points in the space and regions are logical boundaries that define a space. These basic elements are the spatial foundation of all functionality in the system. The translation of the required functionality into a content model for authoring is undertaken for each of the three elements in the matrix in terms of content delivery, navigation and narrative context.

Delivery Stage

The *Delivery Engine* is an AUI for LBS that models both context and location to provide a number of exciting innovations in the areas of content delivery, mobile navigation, and narrative context. The most notable contributions made in this research are based around the development of:

- An intelligent content phrase synthesiser
- A natural directions model for real time navigation
- An algorithmic music sequencer
- Adaptive virtualised 3D soundscapes

This work is described in detail in the Problem 2 Contribution subsection of section 3.6.

Problem 2 Contribution

A central goal of this research was to test the usability and user experience of the *MobiLAudio* prototype. This was achieved in the form of a detailed user trial to determine specific user responses to the functionality, usability and content provided by the iOS audio tour developed. The trial focused extensively on user feedback to determine which elements of the proposed functionality are most important to the end user. Specific levels of detail were of interest (e.g. no music or SFX, voice only, etc.) alongside additional functions utilising a variety of media to try and accurately model the user experience.

The user trial was carried out in central Dublin based around the theme of James Joyce's *Ulysses* and its central character Leopold Bloom. The tour tested the functionality of the survey and authoring tools to determine how effectively multiple tours could be produced (Figure 19). A total of 31 participants took part in the trial, and all participants completed the trial fully.
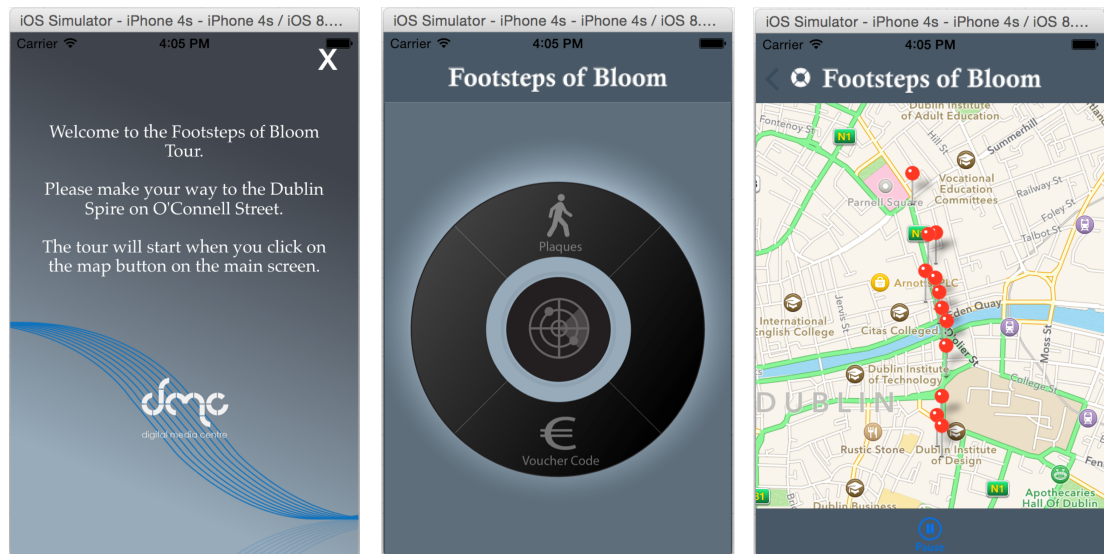
Figure 19: Bloom Tour graphical user interface. *The illustration shows the mobile app that was used for the Footsteps of Bloom user trial in Dublin City Centre.*

The trial focused on the execution of several location based tasks, which were evaluated as part of functionality testing. Quantitative questions relating to the user experience of the application were given alongside post-test NASA Task Load Index (TLX) tests to assess how much of a cognitive load the application places on the user. The testing schedule defined three separate test areas; functionality, user experience, and post-test TLX evaluation. The results provided both quantitative (functionality and TLX) and qualitative (user experience) data for analysis. This combination provided a comprehensive picture of the performance of the application, within the defined testing space of the *Footsteps of Bloom* audio tour.

Functionality Testing. After using the application, participants were asked a series of questions relating to the performance and functionality of the application. These included questions on the music, SFX, speech and directions provided, alongside the general operation of the application on the device.

User Experience Testing. After using the application, each participant was asked to evaluate the experience by answering a short set of qualitative questions. Questions included ease of use, evaluation of audio tours, and whether the app is worth paying for. These questions were intended to be short scenario specific questions with yes/no answers. Subsequently, an improved approach for measuring usability called the system usability scale (SUS) has been recommended, which consists of a 10-item questionnaire with five response options for respondents from strongly agree to strongly disagree. This scale will be used in future application testing.

NASA TLX Testing. After the user experience testing was completed, a brief TLX test was carried out. The smallest TLX test contains a series of five scales that can be marked by either the tester or participant relating to metrics such as mental, physical and temporal demand. These metrics allow the participant to express how much cognitive demand the app placed upon them, which is crucial to the analysis of both functionality and user experience results.

Footsteps of Bloom Tour

The *Footsteps of Bloom* tour was constructed specifically for the purposes of the user trial and commercial evaluation and validation. The tour comprises 10 points of interest on a linear narrative beginning at the O'Connell Street Spire and ending at Davy Byrne's Pub on Grafton Street. The tour points are based on various sidewalk plaques found in central Dublin that punctuate Leopold Bloom's journey in Ulysses, and provides information based on the novel at each point. The tour took approximately 35-40 minutes to complete and utilizes all the major functional elements of the AUI.

Pre-test Questions

Each participant was asked several pre-test questions to determine whether they had any

hearing problems (which would invalidate their test) and to obtain general information about their language skills, familiarity with the subject matter (and Irish culture) and their familiarity with the technologies involved. The results of each question are shown below (Figure 20).
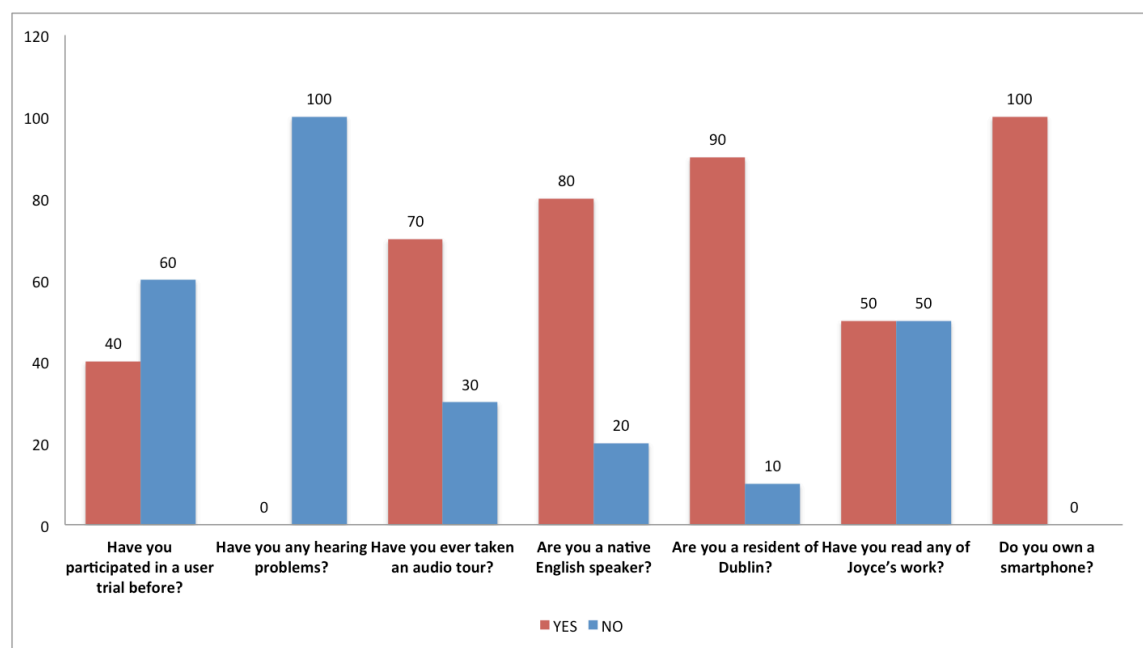
Figure 20: MobiLAudio user trial pre-test question results. *The chart shows answers to yes/no questions relating to the trial and its content.*

Of the participants questioned, none had any hearing difficulties that would preclude them from taking part in the trial. A total of 70% had previously taken an audio tour, which was considered useful in providing a benchmark of familiarity for the application. Most participants were native speakers (80%) living in Dublin (90%) though only 50% were familiar with the work of James Joyce. All participants own a smartphone, and as the test was designed to be largely hands free this was considered to be a sufficient level of prior exposure to the technologies involved in the tour.

Functionality Questions

Each participant was asked a series of questions relating to the functionality of the application. These questions related to the SFX, music, speech, and directions

modelling provided as part of the *Footsteps of Bloom* tour. Questions were also included to determine the general operation of the application, notably its performance as a hands free audio tour (Figure 21).
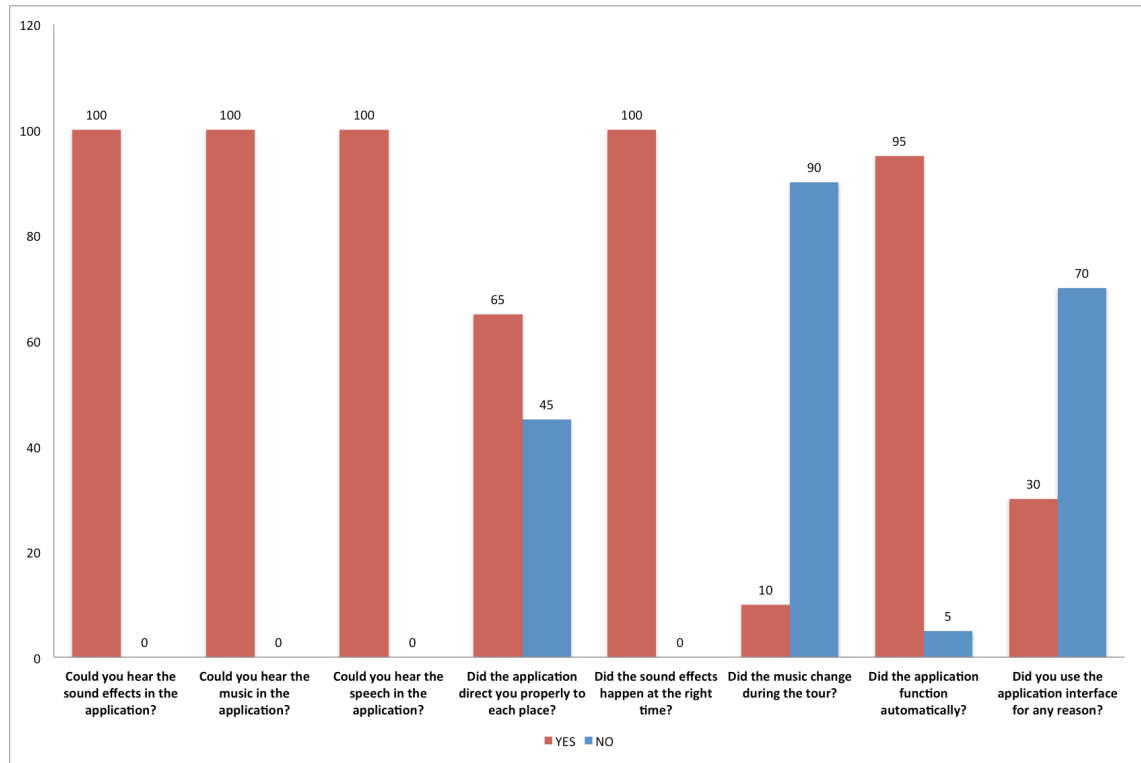
Figure 21: MobiLAudio user trial functionality question results. *The chart shows answers to yes/no questions relating to the operation of the application and its component functionality.*

All participants could distinguish between the music, SFX and speech within the application and considered the balance between the three to be correct. Similarly, they considered the application to have functioned automatically and as required.

Directions Modelling

Only 65% of participants felt the application had directed them correctly, though in many cases this was due to physical occlusion of several Leopold Bloom sidewalk plaques (e.g. due to outside tables at Davy Byrne's Pub, or telecom repair work on Aston Quay for 2 days of the trial). Of those who were misdirected in another way, two participants, after given the first instruction to proceed down O'Connell Street towards

the bridge (from the Dublin Spire), walked instead in the opposite direction towards Parnell Street. In another three instances, participants struggled to find a plaque on Westmoreland Street because the directions to Conways Pub were not strictly accurate (as the sidewalk sign is actually outside Charlie's Chinese Takeaway).

Auditory User Interface Performance

All participants indicated that they could hear sound effects (SFX), music, and speech correctly, and in particular the SFX were considered to be a really good addition to the experience. All users considered the timing of the SFX to be correct (100%) and again this was backed up by qualitative comments indicating that SFX were a significant positive element of the tour. The determination of all three audio elements (SFX, music, speech) validated the design of the listening mode model employed by the application, but there was an issue indicated with repetition of music during the tour. Only 10% of participants felt the music changed or varied during the tour, with many commenting that the music was either too much or too distracting.

This indicates that the use of immersive background music may not be as important or necessary as was initially thought, with listeners indicating a preference for a music 'mute' button in the application. Indeed, the use of the application interface (participants were told the application was automatic) by 70% of participants was mainly to change the volume between POIs, with only 2 participants stating they had used the application's map out of personal interest in the tour.

User Experience Questions

Each participant was also asked a series of qualitative questions relating to the user experience of the *Footsteps of Bloom* tour. These questions focused on the engagement of the application, with questions on user tours and social media integration being

included alongside the opportunity to make more general comments (Figure 22).

All participants found the application easy to use and most found it engaging (95%) and would be interested in other similar tours (95%). Comments on the general concept and content of the tour were positive, with many participants providing suggestions for other tours:

'...easy to use, liked the directions between points...'
'...liked the voice and the sounds, the barge going past sounded real!'
'...storytelling was great, would have liked more content though...'
'...could you have running tours...or cycling tours...'
*'An Easter Rising tour would be brilliant with the bombs going off around you!'*
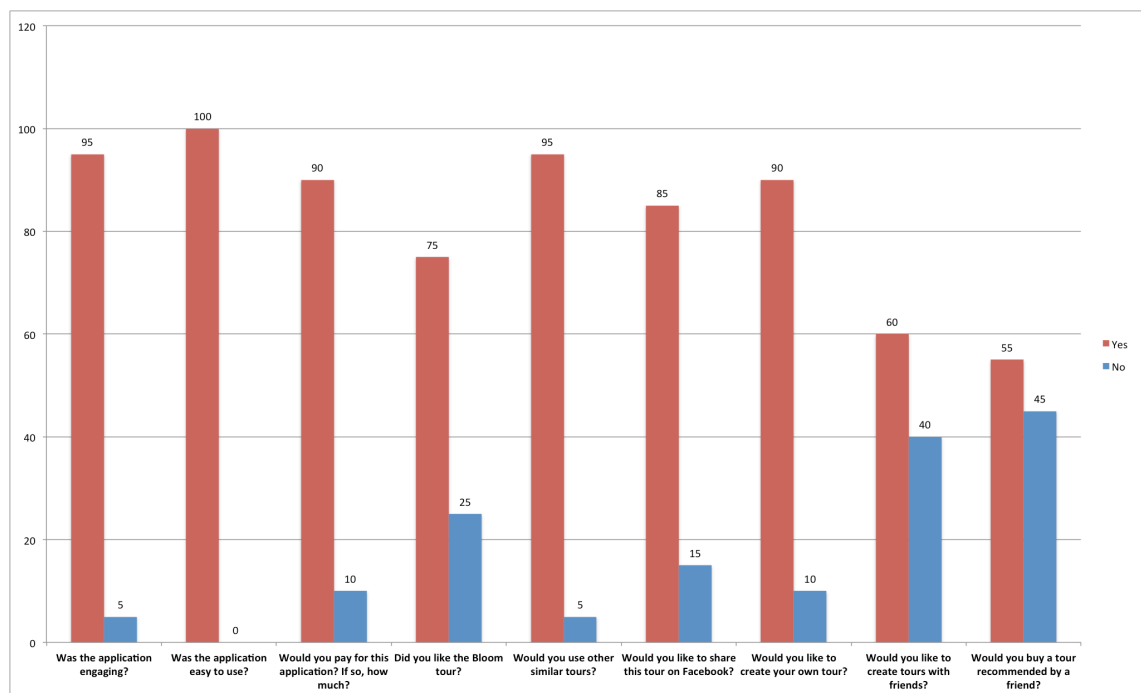
Figure 22: MobiLAudio user trial user experience question results. *The chart shows answers to yes/no questions relating to the engagement of the application, its potential for self generated tours and social media integration opportunities.*

Post Test TLX Results

After completing the tour and answering all test questions, a set of post-test NASA TLX tests were carried out to determine the load on the participant. NASA TLX tests were

originally developed to evaluate astronaut performance and behaviour under demanding or stressful conditions. Under normal circumstances, the tests are useful in assessing the mental and physical load that a certain operation or task places on a participant and are a useful means of comparing how participants actually performed relative to how well they *thought* they performed (Figure 23).
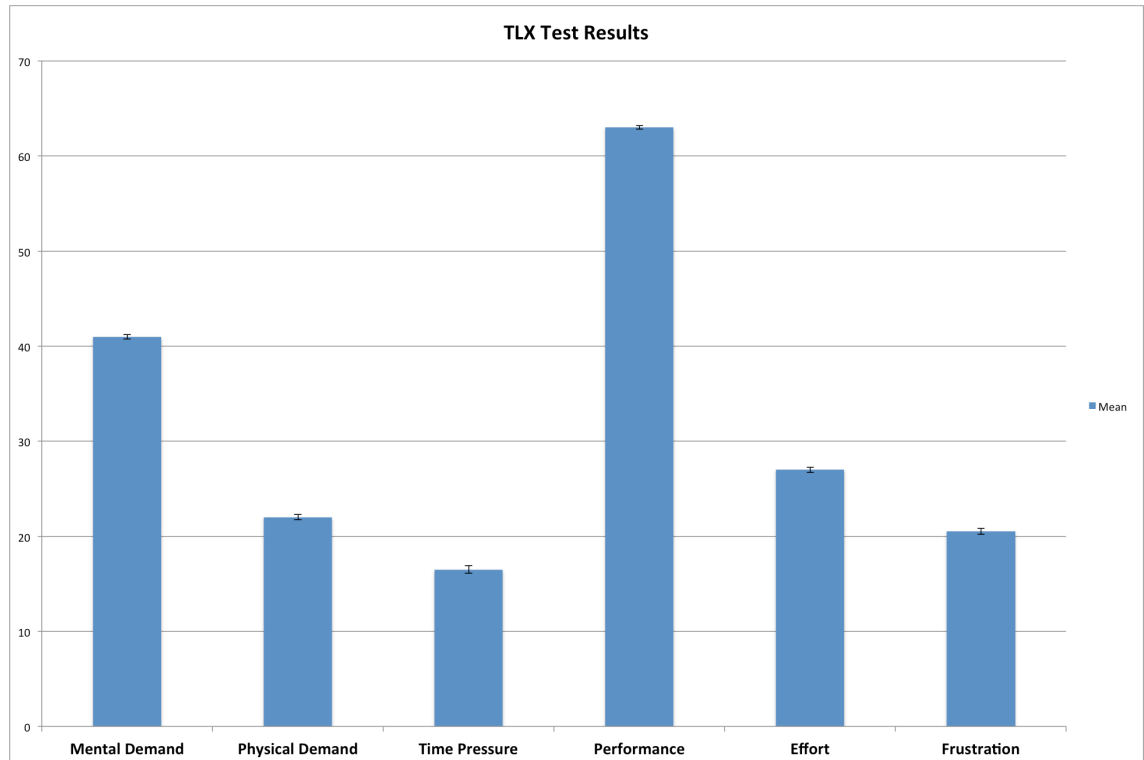
Figure 23: MobiLAudio user trial post test TLX results. *The chart shows the mean and standard deviation (in percent) for each of the six TLX scales. TLX scales cover mental and physical demand, time pressure, performance, effort and frustration.*

The overall results were low (as desired), with physical demand and frustration having values of 22% and 20.5% respectively, suggesting that the user experience results for engagement and ease of use (100%) were a fair reflection of the performance of the application. The important value of time pressure was also very low at 16.5%, suggesting that the application did not impinge upon the tour experience in its function or delivery. Mental demand was higher at 41% and this is potentially explained by the requirement for participants to listen to directions as they navigate the tour (no

participant had any prior idea of the direction the tour would take). The effort score of 27%, though still very low, can perhaps be attributed to the combination of walking and listening to a narrative at the same time. It is also interesting to consider the performance value of 63%, which may indicate in some respects difficulties with the directions modelling provided. In several instances, participants indicated that the directions were too long or confusing and expressed concern that they had not '*done well enough*' even though they themselves were not directly being evaluated. It is also significant that percentile deviations were under 0.4% for all scales, indicating that the TLX values obtained were representative of the overall participant population.

Discussion

The overall reaction to the *Footsteps of Bloom* tour was very positive, and the technical elements of the application were all evaluated to function properly (though directional modelling did cause problems in its current delivery format). In particular, SFX was found to be a very positive addition to the experience of the application, though background music proved to be distracting for some users. The automated nature of the application was also considered a significant feature, with most people only using the graphical interface out of interest rather than necessity. The narrative and conversational flow of the tour was very well received, and the vast majority of participants would be interested in other audio tours. The concept of user-generated tours was strongly endorsed by participants, with integration into social networking being another positive element. Overall, participants reacted very positively to the tour and validated both the design and technical and directions taken in *MobiLAudio*. The results of the Post Test TLX survey were mainly positive with significantly low scores for frustration, time pressure, physical demand and effort. The most interesting part of these results is the

high value of 63% for performance. This could be attributed to the use of a combination of map based visual directions and egocentric based audio directions, which have been shown to cause confusion when users switch between frames of reference. Another reason for this could be that the audio directions were too long as some participants did express concern about the length and complexity of the directions.

3.8. Summary

The work presented in this section has demonstrated consistent contributions to the field of multi-modal geo-services. Initial work in a simulated mobile environment demonstrated effective content delivery using spatial querying techniques, emphasising the viability of location-based information retrieval (Paper VII). The techniques developed in this research demonstrated the potential for LBS applications and created a foundation for further work in this area. This work led to a significant investigation in the area of mobile connectivity and data management in the form of the MEMS system (Paper V). The MEMS project provided the opportunity to take an operational paper based system and develop a digital solution using state-of-the-art mobile technology in the form of a tablet PC and web-based management system. The application was very successful and could be used online and offline to record and analyse spatial datasets in real-time. Issues identified with GPS in urban areas identified considerable challenges for LBS and led to further research in the area of mobile positioning (Paper VI) where a hybrid positioning system was developed to provide application with accurate positioning. Development of the ILC was an essential step in the advancement of mobile device software for the support of LBS. Location dependent apps are essentially non-functional without the provision of accurate positioning. The ILC utilized existing RF infrastructure to improve the accuracy of device positioning by incorporating land

based beacon locations into position calculations. Given accurate positioning, the next challenge was to improve the relevance of data returned to the user and to address the problem of information overload. This led to an investigation into 2D egocentric querying, with the development of an advanced query processor that considered the direction of a user in a real-world context (Paper IV). The *TellMe* system successfully demonstrated egocentric visibility in a MSI context. The system delivered accurate results when tested dependant on optimal network conditions and minimal hardware interference. Tests demonstrated that directional queries performed best overall, with FOV queries, isovist queries and frustum queries taking significantly longer to process. This work was extended in the area of visibility analysis, where multiple visibility approaches were investigated with a focus simulating a user's FOV (Paper III). The 3DQ prototype demonstrated effective visibility based queries for mobile devices. The innovative line-of-sight and hidden query removal functionality produced a more accurate query result by only returning information about objects visible within a user's 3D FOV. The innovative user interface was the first step towards delivering a rich user experience on a mobile device. This research was a milestone in the area of 3D querying of spatial datasets and identified the need for a more effective way of delivering information using multimodal interfaces based on the limitations of the Symbian platform. Work in the area of multimodal interfaces was undertaken, to explore the possibility of delivering information using visual and aural modalities. The *mobiSurround* prototype demonstrated multimodal content delivery using phrase synthesis, 3D soundscaping and natural directions (Paper II). This new work produced a spatialised AUI developed for LBS. It was achieved after several advances upon the current state-of-the- art, notably in the areas of location/data modelling and adaptive content production. The app delivered a high quality AUI for tour delivery. The

methodologies employed in the data modelling process and the development of the AUI served as a solid foundation for further work. This was followed by the development of a 3-stage mobile deployment model for geo-services and a user trial to test the effectiveness of multimodal applications in the field (Paper I). The solution in this instance was the development of a novel Geo-Services platform for mobile multimodal content delivery. The visual modality was presented using a graphical user interface (GUI) that incorporates and builds on many of the ideas developed in previous work on LBS (Paper III, Paper IV, Paper V and Paper VII). The aural modality incorporated an AUI for delivering context sensitive information to the user using focus independent (eyes free) means. The idea was that all information in the application (including media content, directions, advertisements, etc.) would be delivered without the user ever needing to touch the device. Another significant aspect of this research was a live user trial that was performed to test the functionality, user experience, and overall usability of an app created by the platform.

3.9. Conclusions and Future Work

The use of geo-location for information retrieval has become more pervasive in society in recent times (Balduini et al. 2012, Jacob, Mooney, and Winstanley 2012, Geiger et al. 2014). Mobile devices are equipped with multiple sensors to track our every movement in order to improve and deliver services. From simple mapping to immersive audio tours, the technology that provides these services is becoming more ubiquitous as consumers demand more complex services over basic functionality (Rahman and El Saddik 2013, Vazquez-Alvarez, Oakley, and Brewster 2012, Mackaness et al. 2014). Reliance on mobile devices has also changed due to the proliferation of mobile apps that provide a vast range of services, where the importance of the delivery of the service

is on par with the usability and user experience of the application.

Research in this area has progressed from using simulated mobile environments to demonstrate the visual modality to the ubiquitous delivery of rich media using multimodal interfaces (geo-services). To effectively deliver these services, research focused on innovative solutions to real-world problems in a number of disciplines including geo-location, MSI, LBS, rich media interfaces and AUIs. The results show that significant advances have been achieved in the areas of device positioning and MSI underpinned by improvements in communications technologies. Results also show that the concept of multimodal geo-service content delivery is starting to gain traction, with positive results indicating that mobile users are becoming familiar with synergistic approaches to multimodal content delivery and the demand is increasing.

The techniques developed in this research demonstrate the viability of multimodal geo-services in a mobile context. The next logical step for this research is in the area of multimodal content delivery using AR and wearable devices. The use of AR for geo-service delivery has gained significant attention in recent time with the use of Google Glass as a demonstrator platform (Altwaijry, Moghimi, and Belongie 2014, Geiger et al. 2014, Schickler et al. 2015). Applications in the area tourism, transport and biomedical monitoring have been developed that demonstrate the significant potential of this technology (Nurminen, Järvi, and Lehtonen 2014, Rodgers, Pai, and Conroy 2015, Tatzgern et al. 2015). More specifically, research in the area wearable devices as facilitators of health and lifestyle change has huge potential to support medical monitoring and observation (Miguez-Burbano and Ergon 2015, Patel, Asch, and Volpp 2015). With the advent of the Internet of things and intelligent devices such as eyewear, watches and clothing, the potential to develop effective geo-services is also significant

(Farmer and Tarassenko 2015, Meurer et al. 2015). These services will be delivered using multiple devices across multiple platforms that will communicate using multiple RF technologies such as WiFi, Bluetooth, Zigbee and NFC. To provide these services effectively across multiple platforms, additional research is required in the areas of multimodal interaction for mobile services and the extension of multimodal fusion and fission techniques for wearable devices (Yeung 2015).

Therefore, future work will focus on multimodal input/output for mobile/wearable devices building on multimodal fusion/fission approaches that leverage ideas developed in the area of perceptual user interfaces (Turk 2014). Proposals are being considered in the following areas.

3.9.1. Adaptive Multimodal Input

With the advancement of mobile/wearable technology and a host of new interaction modalities, mobile user interfaces are a growth area for future applications (Page 2015, Castillo and Thierer 2015). This combined with the ability to effectively track personalised contextual information opens up new possibilities in the area of adaptive multimodal input for mobile services. Future work in this area will focus on the automatic adaptation of multimodal input based on device capability and contextual information. The idea of a user driven rule-based to context mapping will be investigated to support the automatic profile selection and adaptation of multimodal input (Zaguia et al. 2010, Kong et al. 2011). Applications will be developed based on guidelines proposed in (Dumas, Solórzano, and Signer 2013) and the effectiveness of adaptive multimodal input will be tested based on the CARE properties (Lalanne et al. 2009), with the aim of improving usability and overall user experience of multimodal interfaces. This approach will help reduce cognitive load on the user and will extend the

state-of-the-art by taking a responsive approach to multimodal input for mobile applications.

3.9.2. Adaptive Multimodal Output

Familiarity with multimodal output modalities is a barrier to effective mobile content delivery. Multimodal output has considerable benefits over modal interfaces when perceptual rules are considered (Turk 2014). Future work in the area of multimodal output will investigate the idea of adaptive multimodal content delivery based on device capabilities and contextual appropriateness of output modalities. Adaptive multimodal output selection will be based on the current input modalities being used, usage patterns and environmental conditions (temperature, weather, noise, etc.) (Dumas, Solórzano, and Signer 2013, Dumas, Signer, and Lalanne 2014). Applications will be developed based on guidelines proposed in (Dumas, Solórzano, and Signer 2013) and the effectiveness of adaptive multimodal output will be tested based on the CARE properties (Lalanne et al. 2009), with the aim of improving usability and overall user experience of multimodal interfaces. Mobile applications output modes will adapt to a user's context and available interfaces, to provide a rich user experience based on contextual information (Teixeira et al. 2014).

3.9.3. Content Aggregation for Geo-services

The Social Web is becoming a reliable source of information for mobile applications. Adaptive content delivery based on contextual information can provide dynamic personalised user experiences based on social profiles. This type of content aggregation is a complex task given the scale of the data available. Future work in this area will investigate the aggregation of content for geo-services, with a focus on the social media APIs. Selection of APIs will be primarily considered based on availability, rating and

device affordance. Further work will also consider the dynamic usage of social media

APIs based on social interactions and proximity (Yu, Zimmermann, and Tang 2015).

4. References

Abowd, Gregory D., Christopher G. Atkeson, Jason Hong, Sue Long, Rob Kooper, and Mike Pinkerton. 1997. "Cyberguide: A mobile context‐aware tour guide." *Wireless Networks* 3 (5):421-433. doi: 10.1023/A:1019194325861.

Absar, R, and Catherine Guastavino. 2008. "Usability of non-speech sounds in user interfaces." International Conference on Auditory Display.

Akatsu, Hiroko, Hiroko Akatsu, Akinori Komatsubara, and Akinori Komatsubara. 2009. "Auditory and Visual Guidance for Reducing Cognitive Load." *Lecture Notes In Computer Science; Vol. 5619.*

Altwaijry, H., M. Moghimi, and S. Belongie. 2014. "Recognizing locations with Google Glass: A case study." Applications of Computer Vision (WACV), 2014 IEEE Winter Conference on, 24-26 March 2014.

Alvarez, Yolanda Vazquez, and Stephen A. Brewster. 2010. "Designing spatial audio interfaces to support multiple audio streams." Proceedings of the 12th international conference on Human computer interaction with mobile devices and services, Lisbon, Portugal.

Ancona, M, N Scagliola, and A Traverso. 2006. "Application of 3G cellular phones to cultural heritage: the Agamemnon project." *Recording, Modeling and Visualization of Cultural Heritage*:217-226.

Andre, Elisabeth, Wolfgang Finkler, Winfried Graf, Thomas Rist, Anne Schauder, and Wolfgang Wahlster. 1993. "WIP: the automatic synthesis of multimodal presentations." In *Intelligent multimedia interfaces*, edited by T. Maybury Mark, 75-93. American Association for Artificial Intelligence.

Bai, Yuntian Brian, Suqin Wu, Hongren Wu, and Kefei Zhang. 2012. "Overview of RFID-based indoor positioning technology." Proceedings of the Geospatial Science Research Symposium. Melbourne: RMIT University.

Balduini, Marco, Irene Celino, Daniele Dell'Aglio, Emanuele Della Valle, Yi Huang, Tony Lee, Seon-Ho Kim, and Volker Tresp. 2012. "BOTTARI: An augmented reality mobile application to deliver personalized and location-based recommendations by continuous analysis of social media streams." *Web Semantics: Science, Services and Agents on the World Wide Web* 16:33-41.

Ballatore, Andrea, Thoa Pham, Junjun Yin, Linh Truong-Hong, and JamesD Carswell. 2015. "Design and Development of Personal GeoServices for Universities." In *Information Fusion and Geographic Information Systems (IF&GIS' 2015)*, edited by Vasily Popovich, Christophe Claramunt, Manfred Schrenk, Kyrill Korolenko and Jérôme Gensel, 3-26. Springer International Publishing.

Bartie, Phil J., and William A. Mackaness. 2006. "Development of a Speech-Based Augmented Reality System to Support Exploration of Cityscape." *Transactions in GIS* 10 (1):63-86. doi: 10.1111/j.1467-9671.2006.00244.x.

Bartie, Phil, Steven Mills, and Simon Kingham. 2008. "An Egocentric Urban Viewshed: A Method for Landmark Visibility Mapping for Pedestrian Location Based Services." In *Geospatial Vision*, edited by Antoni Moore and Igor Drecki, 61-85. Springer Berlin Heidelberg.

Bartie, Phil, Femke Reitsma, Simon Kingham, and Steven Mills. 2010. "Advancing visibility modelling algorithms for urban environments." *Computers, Environment and Urban Systems* 34:518-531. doi: 10.1016/j.compenvurbsys.2010.06.002.

Batty, M. 2001. "Exploring isovist fields: space and shape in architectural and urban morphology." *Environment and Planning B: Planning and Design* 28 (1):123-150.

Baus, Jörg, Rainer Wasinger, Ilhan Aslan, Antonio Krüger, Andreas Maier, and Tim Schwartz. 2007. "Auditory perceptible landmarks in mobile navigation." IUI '07: Proceedings of the 12th international conference on Intelligent user interfaces.

Beeharee, Ashweeni, and Anthony Steed. 2005. "Filtering Location-Based Information Using Visibility." In *Location- and Context-Awareness*, edited by Thomas Strang and Claudia Linnhoff-Popien, 306-315. Springer Berlin Heidelberg.

Bell, Scott, Wook Rak Jung, and Vishwa Krishnakumar. 2010. "WiFi-based enhanced positioning systems: accuracy through mapping, calibration, and classification." Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness, San Jose, California.

Benedikt, M L. 1979. "To take hold of space: isovists and isovist fields." *Environment and Planning B: Planning and Design* 6:47-65. doi: 10.1068/b060047.

Bernsen, Niels Ole. 1997. "Defining a taxonomy of output modalities from an HCI perspective." *Comput. Stand. Interfaces* 18 (6-7):537-553. doi: 10.1016/s0920-5489(97)00018-4.

Bimber, Oliver, and Ramesh Raskar. 2005. "Spatial Augmented Reality Merging Real and Virtual Worlds." *Scientist* 6:83-92. doi: 10.1260/147807708784640126.

Blattner, M. M., and E. P. Glinert. 1996. "Multimodal integration." *MultiMedia, IEEE* 3 (4):14-24. doi: 10.1109/93.556457.

Bolt, Richard A. 1980. "Put-that-there: Voice and gesture at the graphics interface." *SIGGRAPH Comput. Graph.* 14 (3):262-270. doi: 10.1145/965105.807503.

Bouchet, Jullien, and Laurence Nigay. 2004. "ICARE: a component-based approach for the design and development of multimodal interfaces." *CHI04 extended abstracts on Human factors in*:1-4. doi: 10.1145/985921.986055.

Bouchet, Jullien, Laurence Nigay, and Thierry Ganille. 2004. "ICARE software components for rapidly developing multimodal interfaces." ICARE software components for rapidly developing multimodal interfaces.

Brennan, Emma. 2012. mobiSurround Market Feasability Report. Enterprise Ireland.

Bruning, S., J. Zapotoczky, P. Ibach, and V. Stantchev. 2007. "Cooperative Positioning with MagicMap." *2007 4th Workshop on Positioning, Navigation and Communication*. doi: 10.1109/WPNC.2007.353606.

Buyukkokten, O, J Cho, H Garcia-Molina, L Gravano, and N Shivakumar. 1999. "Exploiting Geographical Location Information of Web Pages." In Proc. of the ACM SIGMOD Workshop on the Web and Databases (WebDB '99).

Cao, Yujia, Mariet Theune, and Anton Nijholt. 2009. "Modality effects on cognitive load and performance in high-load information presentation." Proceedings of the 14th international conference on Intelligent user interfaces, Sanibel Island, Florida, USA.

Carswell, J. D, K Gardiner, and J Yin. 2010. 3DQ: Threat Dome Visibility Querying on Mobile Devices. *GIM International* Vol.24 (8).

Carswell, J.D., A. Eustace, K. Gardiner, E. Kilfeather, and M. Neumann. 2002. "An environment for mobile context-based hypermedia retrieval." *Proceedings. 13th International Workshop on Database and Expert Systems Applications*. doi: 10.1109/DEXA.2002.1045952.

Carswell, James D, and Junjun Yin. 2012. "Mobile spatial interaction in the Future Internet of Things." Geoinformatics (GEOINFORMATICS), 2012 20th International Conference on.

Castillo, Andrea, and Adam Thierer. 2015. "Projecting the Growth and Economic Impact of the Internet of Things." *Available at SSRN 2618794.*

Chan, L.-W, Y.-Y Hsu, Y.-P Hung, and J.Y.-J Hsu. 2005. "Orientation-Aware Handhelds for Panorama-Based Museum Guiding System." UbiComp 2005 Workshop: Smart Environments and their Applications to Cultural Heritage.

Cheyer, Adam, and Luc Julia. 1998. "Multimodal maps: An agent-based approach." *Multimodal Human-Computer Communication*:111-121.

Chincholle, Didier, Mikael Goldstein, Marcus Nyberg, and Mikael Eriksson. 2002. "Lost or Found? A Usability Evaluation of a Mobile Navigation and Location-Based Service." In *Human Computer Interaction with Mobile Devices*, edited by Fabio Paternò, 211-224. Springer Berlin Heidelberg.

Chittaro, Luca. 2008. "Interacting with Visual Interfaces on Mobile Devices." In *Human-Computer Interaction Symposium*, edited by Peter Forbrig, Fabio Paternò and AnneliseMark Pejtersen, 1-5. Springer US.

Church, Karen, Barry Smyth, and Nuria Oliver. 2009. "Visual interfaces for improved mobile search." CEUR Workshop Proceedings.

Cohen, Philip R., Michael Johnston, David McGee, Sharon Oviatt, Jay Pittman, Ira Smith, Liang Chen, and Josh Clow. 1997. "QuickSet: multimodal interaction for distributed applications." Proceedings of the fifth ACM international conference on Multimedia, Seattle, Washington, USA.

Costa, David, and Carlos Duarte. 2011. "Adapting multimodal fission to user's abilities." *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 6765 LNCS:347-356. doi: 10.1007/978-3-642-21672-5_38.

Crowley, James L., Joëlle Coutaz, and François Bérard. 2000. "Perceptual user interfaces: things that see." *Communications of the ACM* 43:54-ff. doi: 10.1145/330534.330540.

Cullen, Charlie, Carl Goodman, Paula McGloin, Anna Deegan, and Evin McCarthy. 2009. *Reusable, interactive, multilingual online avatars.*

Dam, Andries van. 1997. "Post-WIMP user interfaces." *Commun. ACM* 40 (2):63-67. doi: 10.1145/253671.253708.

David, Lincoln, Markus Endler, Simone D.J. Barbosa, and Jose Viterbo Filho. 2011. "Middleware Support for Context-Aware Mobile Applications with Adaptive Multimodal User Interfaces." *2011 Fourth International Conference on Ubi-Media Computing*:106-111. doi: 10.1109/U-MEDIA.2011.50.

Dey, Anind, Jeffrey Hightower, Eyal de Lara, and Nigel Davies. 2010. "Location-Based Services." *Pervasive Computing* 9:11-12. doi: 10.1109/MPRV.2010.10.

Diaz, Estefania Munoz. 2015. "Inertial Pocket Navigation System: Unaided 3D Positioning." *Sensors* 15 (4):9156-9178.

Ding, Junyan, Luis Gravano, and Narayanan Shivakumar. 2000. "Computing Geographical Scopes of Web Resources." Proceedings of the 26th International Conference on Very Large Data Bases (VLDB 00) %@ 1-55860-715-3.

Djuknic, G.M., and R.E. Richton. 2001. "Geolocation and assisted GPS." *Computer* 34. doi: 10.1109/2.901174.

Donker, Hilko, Palle Klante, and Peter Gorny. 2002. *The Design of Auditory User Interfaces for Blind Users*, *NordiCHI*.

Dowell, John, and Yael Shmueli. 2008. "Blending speech output and visual text in the multimodal interface." *Human factors* 50:782-788. doi: 10.1518/001872008X354165.

Doyle, J., Michela Bertolotto, and David C Wilson. 2009. "Evaluating the benefits of multimodal interface design for CoMPASS - a mobile GIS." *GeoInformatica*. doi: 10.1007/s10707-009-0079-2.

Droumeva, Milena. 2005. "Understanding Immersive Audio: A Historical and Socio-Cultural Exploration of Auditory Displays." Eleventh Meeting of the International Conference on Auditory Display (ICAD'05).

Dumas, Bruno, Denis Lalanne, and Sharon Oviatt. 2009. "Multimodal interfaces: A survey of principles, models and frameworks." *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 5440 LNCS:3-26. doi: 10.1007/978-3-642-00437-7_1.

Dumas, Bruno, Beat Signer, and Denis Lalanne. 2014. "A graphical editor for the SMUIML multimodal user interaction description language." *Science of*

Computer Programming 86 (0):30-42. doi:
http://dx.doi.org/10.1016/j.scico.2013.04.003.

Dumas, Bruno, María Solórzano, and Beat Signer. 2013. "Design guidelines for
adaptive multimodal mobile input solutions." *Proceedings of the 15th
international conference on Human-computer interaction with mobile devices
and services - MobileHCI '13*:285-294. doi: 10.1145/2493190.2493227.

Eckel, G. 2001. "Immersive audio-augmented environments: the LISTEN project."
Proceedings Fifth International Conference on Information Visualisation. doi:
10.1109/IV.2001.942112.

Edwards, Alistair. 1989. "Soundtrack: An Auditory Interface for Blind Users." *Human-
Computer Interaction* 4:45-66. doi: 10.1207/s15327051hci0401_2.

Egenhofer, Max. 1999. "Spatial Information Appliances: a next generation of
Geographic Information Systems." in: Proceedings of the 1st Brazilian
workshop on geoinformatics.

El-Rabbany, Ahmed. 2006. *Introduction to GPS : the Global Positioning System*. 2nd
ed, *Artech House mobile communications series*. Boston, MA: Artech House.

Elepfandt, Monika, and Marcelina Sünderhauf. 2011. "Multimodal, Touchless
Interaction in Spatial Augmented Reality Environments." In *Digital Human
Modeling*, edited by VincentG Duffy, 263-271. Springer Berlin Heidelberg.

EU/ESA. 2006. "Galileo: Satellite launches."
http://ec.europa.eu/enterprise/policies/satnav/galileo/satellite-
launches/index_en.htm.

Farmer, Andrew, and Lionel Tarassenko. 2015. "Use of Wearable Monitoring Devices
to Change Health Behavior." *JAMA* 313 (18):1864-1865.

Fisher, P. F. 1995. "An exploration of probable viewsheds in landscape planning."
Environment and Planning B: Planning and Design 22 (5):527-546.

Fitch, T., and G. Kramer. 1994. *Sonifying the Body Electric: Superiority of an Auditory
over a Visual Display in a Complex Multivariate System*. Edited by G. Kramer,
Auditory Display: Sonification, Audification and Auditory Interfaces: Addison-
Wesley.

Foster, M. E. 2002. State of the art review: Multimodal fission. In *COMIC project
Deliverable 6.1*.

Frank, Christopher, David Caduff, and Markus Wuersch. 2004. "From GIS to LBS: An Intelligent Mobile GIS." *IfGI Prints* 22:277-287.

Frohlich, P., R. Simon, and L. Baillie. 2009. "Mobile Spatial Interaction." *Personal and Ubiquitous Computing* 13 (4):251-253. doi: Doi 10.1007/S00779-008-0208-1.

Gaines, Brian R, and Mildred L G Shaw. 1986. "A Learning Model for Forecasting the Future of Information Technology." *Future Computing Systems* 1:31-69.

Gardiner, K. 2004. "A Framework for Processing Viewer-based Directional Queries in a Simulated Mobile Environment." Masters, School of Media, Dublin institute of Technology.

Gardiner, K., and J.D. Carswell. 2003. "Viewer-based directional querying for mobile applications." Third International Workshop on Web and Wireless Geographical Information Systems W2GIS.

Gaver, William. 1989. "The SonicFinder: An Interface That Uses Auditory Icons." *Human-Computer Interaction* 4:67-94. doi: 10.1207/s15327051hci0401_3.

Geiger, Philip, Marc Schickler, Rüdiger Pryss, Johannes Schobel, and Manfred Reichert. 2014. "Location-based mobile augmented reality applications: Challenges, examples, lessons learned."

Gentile, A, A Santangelo, S Sorce, A Augello, G Pilato, A Genco, and S Gaglio. 2009. "Exploiting Multimodality for Intelligent Mobile Access to Pervasive Services in Cultural Heritage Sites." *Multimodality in Mobile Computing and Mobile Devices: Methods for Adaptable Usability*:217–240.

Ghiani, Giuseppe, Marco Manca, Fabio Paternò, and Claudio Porta. 2014. "Beyond Responsive Design: Context-Dependent Multimodal Augmentation of Web Applications." In *Mobile Web Information Systems*, edited by Irfan Awan, Muhammad Younas, Xavier Franch and Carme Quer, 71-85. Springer International Publishing.

Goetz, M, and A Zipf. 2010. "OPEN ISSUES IN BRINGING 3D TO LOCATION BASED SERVICES (LBS)-A REVIEW FOCUSING ON 3D DATA STREAMING AND 3D INDOOR NAVIGATION." *Remote Sensing and Spatial ...* XXXVIII:121-124.

Goyal, R., and Max J. Egenhofer. 2001. "Similarity of cardinal directions." *Advances in Spatial and Temporal Databases* 2121:36-55.

Gravano, L, V Hatzivassiloglou, and R Lichtenstein. 2003. "Categorizing Web Queries According to Geographical Locality." In Proc. of the 12th ACM Conference on Information and Knowledge Management (CIKM 2003).

Guan, Tao, Yin Fan, Liya Duan, and Junqing Yu. 2014. "On-Device Mobile Visual Location Recognition by Using Panoramic Images and Compressed Sensing Based Visual Descriptors." *PLoS ONE* 9 (6):e98806. doi: 10.1371/journal.pone.0098806.

Harada, Susumu, T. Scott Saponas, and James A. Landay. 2007. "VoicePen: augmenting pen input with simultaneous non-linguisitic vocalization." Proceedings of the 9th international conference on Multimodal interfaces, Nagoya, Aichi, Japan.

Heidemann, Gunther, Ingo Bax, and Holger Bekel. 2004. "Multimodal interaction in an augmented reality scenario." Proceedings of the 6th international conference on Multimodal interfaces - ICMI '04.

Heinzle, F, M Kopczynski, and M. Sester. 2003. "Spatial Data Interpretation for the Intelligent Access to Spatial Information in the Internet." In Proc. of 21st International Cartographic Conference, Durban, South Africa., August 10-16.

Heller, Florian, Thomas Knott, Malte Weiss, and Jan Borchers. 2009. "Multi-user interaction in virtual audio spaces." Proceedings of the 27th international conference extended abstracts on Human factors in computing systems - CHI EA '09.

Hightower, Jeffrey, and Gaetano Borriello. 2001. "Location Systems for Ubiquitous Computing." *Computer %@ 0018-9162* 34 (8):57-66. doi: 10.1109/2.940014.

Hightower, Jeffrey, and Gaetano Borriello. 2004. "Particle Filters for Location Estimation in Ubiquitous Computing: A Case Study." In *UbiComp 2004: Ubiquitous Computing*, edited by Nigel Davies, ElizabethD Mynatt and Itiro Siio, 88-106. Springer Berlin Heidelberg.

Hightower, Jeffrey, Anthony Lamarca, and Ian E. Smith. 2006. "Practical lessons from Place Lab." *IEEE Pervasive Computing* 5:32-39. doi: 10.1109/MPRV.2006.55.

Hinckley, Ken, Koji Yatani, Michel Pahud, Nicole Coddington, Jenny Rodenhouse, Andy Wilson, Hrvoje Benko, and Bill Buxton. 2010. "Manual deskterity: an exploration of simultaneous pen + touch direct input." CHI '10 Extended Abstracts on Human Factors in Computing Systems, Atlanta, Georgia, USA.

Hofmann-Wellenhof, B, Herbert Lichtenegger, and James Collins. 2001. *Global positioning system : theory and practice / B. Hofmann-Wellenhof, H. Lichtenegger, J. Collins., GPS.*

Holland, Simon, David R. Morse, and Henrik Gedenryd. 2002. "AudioGPS: Spatial Audio Navigation with a Minimal Attention Interface." *Personal and Ubiquitous Computing* 6 (4):253-259. doi: 10.1007/s007790200025.

Honold, Frank, Felix Schüssel, and Michael Weber. 2012. "Adaptive Probabilistic Fission for Multimodal Systems." Proceedings of the 24th Australian Computer-Human Interaction Conference.

Hooten, Eli R., Sean T. Hayes, and Julie A. Adams. 2013. "Communicative modalities for mobile device interaction." *International Journal of Human Computer Studies* 71:988-1002. doi: 10.1016/j.ijhcs.2013.05.003.

Iannizzotto, Giancarlo, Carlo Costanzo, Francesco La Rosa, and Pietro Lanzafame. 2005. "A multimodal perceptual user interface for video-surveillance environments." Proceedings of the 7th international conference on Multimodal interfaces - ICMI '05.

ICiNG. 2006. "Innovative Cities for the Next Generation." http://cordis.europa.eu/project/rcn/79299_en.html.

Jacob, Ricky, Peter Mooney, and Adam Winstanley. 2012. *What's up that Street? Exploring Streets Using a Haptic GeoWand*: Springer.

Jacob, Ricky, Adam Winstanley, Naomi Togher, Richard Roche, and Peter Mooney. 2012. "Pedestrian navigation using the sense of touch." *Computers, Environment and Urban Systems* 36 (6):513-525. doi: http://dx.doi.org/10.1016/j.compenvurbsys.2012.10.001.

Jagdish, Deepak, Rahul Sawhney, Mohit Gupta, and Shreyas Nangia. 2008. "Sonic Grid: an auditory interface for the visually impaired to navigate GUI-based environments." IUI 08 Proceedings of the 13th international conference on Intelligent user interfaces.

Jaimes, Alejandro, and Nicu Sebe. 2007. "Multimodal human-computer interaction: A survey." *Computer Vision and Image Understanding* 108:116-134. doi: 10.1016/j.cviu.2006.10.019.

Janarthanam, Srini, Oliver Lemon, Phil Bartie, Tiphaine Dalmas, Anna Dickinson, Xingkun Liu, William Mackaness, and Bonnie Webber. 2013. "Evaluating a city

exploration dialogue system combining question-answering and pedestrian navigation."

Jiang, Bin, and Xiaobai Yao. 2006. "Location-based services and GIS in perspective." *Computers, Environment and Urban Systems* 30:712-725. doi: 10.1016/j.compenvurbsys.2006.02.003.

Kaasinen, Eija. 2003. "User needs for location-aware mobile services." *Personal and Ubiquitous Computing* 7:70-79. doi: 10.1007/s00779-002-0214-7.

Klemettinen, Mika. 2007. *Enabling Technologies for Mobile Services: The MobiLife Book*: Wiley Publishing.

Kölsch, Mathias, Ryan Bane, Tobias Höllerer, and Matthew Turk. 2006. "Multimodal interaction with a wearable augmented reality system." *IEEE Computer Graphics and Applications* 26:62-71. doi: 10.1109/MCG.2006.66.

Kong, J., W. Y. Zhang, N. Yu, and X. J. Xia. 2011. "Design of human-centric adaptive multimodal interfaces." *International Journal of Human Computer Studies* 69:854-869. doi: 10.1016/j.ijhcs.2011.07.006.

Koons, David B., Carlton J. Sparrell, and Kristinn Rr. Thorisson. 1998. "Integrating simultaneous input from speech, gaze, and hand gestures." In *Readings in intelligent user interfaces*, edited by T. Maybury Mark and Wahlster Wolfgang, 53-64. Morgan Kaufmann Publishers Inc.

Kramer, Gregory. 1993. *Auditory Display: Sonification, Audification, and Auditory Interfaces*: Perseus Publishing.

Kurkovsky, Stan. 2009. *Multimodality in Mobile Computing and Mobile Devices: Methods for Adaptable Usability*: Information Science Reference - Imprint of: IGI Publishing.

Lalanne, Denis, Laurence Nigay, Philippe Palanque, Peter Robinson, Jean Vanderdonckt, and Jean-Francois Ladry. 2009. "Fusion engines for multimodal input: a survey." International Conference on Multimodal Interfaces.

LaMarca, Anthony, Yatin Chawathe, Sunny Consolvo, Jeffrey Hightower, Ian Smith, James Scott, Timothy Sohn, James Howard, Jeff Hughes, Fred Potter, Jason Tabert, Pauline Powledge, Gaetano Borriello, and Bill Schilit. 2005. "Place Lab: Device Positioning Using Radio Beacons in the Wild." *Pervasive Computing* 3468:116-133. doi: 10.1007/11428572_8.

Lee, Gun A., Mark Billinghurst, and Gerard Jounghyun Kim. 2004. "Occlusion based interaction methods for tangible augmented reality environments." *Virtual Reality Continuum And Its Applications*:419. doi: 10.1145/1044588.1044680.

Lee, Ryong, Yong Jin Kwon, and Kazutoshi Sumiya. 2009. *Layer-based media integration for mobile mixed-reality applications*.

Lei, Zhang, and Paul Coulton. 2009. "A mobile geo-wand enabling gesture based POI search an user generated directional POI photography." Proceedings of the International Conference on Advances in Computer Enterntainment Technology, Athens, Greece.

Lin, Pei-Jung, Sheng-Chang Chen, Yi-Hsung Li, Meng-Syue Wu, and Shih-Yue Chen. 2014. "An Implementation of Augmented Reality and Location Awareness Services in Mobile Devices." In *Mobile, Ubiquitous, and Intelligent Computing*, edited by James J. Park, Hojjat Adeli, Namje Park and Isaac Woungang, 509-514. Springer Berlin Heidelberg.

Liu, Xuan, Shashi Shekhar, and Sanjay Chawla. 2000. "Processing object-oientation-based direction queries: a summary of results." Proceedings of the 8th ACM international symposium on Advances in geographic information systems, Washington, D.C., USA.

López, Francisco J Perales. 2009. "Advanced perceptual user interfaces: Applications for disabled and elderly people." *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 5727 LNCS:930-931. doi: 10.1007/978-3-642-03658-3_116.

Mackaness, W.A, P Bartie, T Dalmas, S Janarthanam, O Lemon, X Liu, and B Webber. 2014. "Talk the Walk and Walk the talk: Design, Implementation and Evaluation of a Spoken Dialogue System for Route Following and City Learning'." Annual Conference of the Association of American Geographers.

Maierhofer, Stefan , Rainer Simon, and Robert F.   Tobler. 2007. "Simplified Guided Visibility Sampling for Location Based Services." In Proceedings of the 12th Int. Conference on Urban, Regional Development and Information Society (CORP2007), Vienna, Austria.

Mandal, Bappaditya, Shue-Ching Chia, Liyuan Li, Vijay Chandrasekhar, Cheston Tan, and Joo-Hwee Lim. 2014. "A wearable face recognition system on google glass for assisting social interactions." Computer Vision-ACCV 2014 Workshops.

Mansoux, Benoît, Laurence Nigay, and Jocelyne Troccaz. 2007. "Output Multimodal Interaction: The Case of Augmented Surgery." In *People and Computers XX — Engage*, edited by Nick Bryan-Kinns, Ann Blanford, Paul Curzon and Laurence Nigay, 177-192. Springer London.

Marimon Sanjuan, David, Tomasz Adamek, Arturo Bonnin, and Tomasz Trzcinski. 2011. "Enhancing global positioning by image recognition." IEEE International Symposium on Mixed and Augmented Reality (ISMAR), Workshop on Enabling Large-Scale Outdoor Mixed Reality and Augmented Reality.

McDonell, M. 1994. "Scan-Line Methods in Spatial Data Systems." Proceedings of the 4th International Symposium on Spatial Data Handling 1990, Zurich, Switzerland.

McGee, J, and C Cullen. 2009a. "Vocate: Auditory Interfaces for Location-based Services." Proceedings of the 23rd Conference on Computer Human Interaction, Cambridge, United Kingdom.

McGee, J, and C Cullen. 2009b. "Vocate: auditory interfaces for the LOK8 project." In Proceedings of the 9th IT & T Conference, Dublin Institute of Technology, Dublin, Ireland.

McGookin, D. 2004. "Understanding and Improving the Identification of Concurrently Presented Earcons." PhD, University of Glasgow.

Meek, Sam, Gary Priestnall, Mike Sharples, and James Goulding. 2013. "Mobile capture of remote points of interest using line of sight modelling." *Computers and Geosciences* 52:334-344. doi: 10.1016/j.cageo.2012.09.028.

Meurer, Johanna, Julian Dax, Martin Stein, Thomas Ludwig, and Volker Wulf. 2015. "Bridging Location-based Data with Mobile Practices." *MOBILITY 2015*:22.

Miguez-Burbano, Maria Jose, and Emma Ergon. 2015. "Use of Wearable Monitoring Devices to Change Health Behavior." *JAMA* 313 (18):1865-1865.

Misra, Ananya, Perry R. Cook, and Ge Wang. 2006. "A New Paradigm for Sound Design." Proceedings of the International Conference on Digital Audio Effects.

Modsching, M, R Kramer, and K ten Hagen. 2006. "Field Trial on GPS Accuracy in a medium size city: The influence of built-up." *3 rd Workshop on Positioning*:1-10.

Moniri, Mohammad Mehdi, Michael Feld, and C. Muller. 2012. "Personalized In-Vehicle Information Systems: Building an Application Infrastructure for Smart

Cars in Smart Spaces." Intelligent Environments (IE), 2012 8th International Conference on, 26-29 June 2012.

Moniri, Mohammad Mehdi, Christian M, #252, and ller. 2012. "Multimodal reference resolution for mobile spatial interaction in urban environments." Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Portsmouth, New Hampshire.

Mousavi, Seyed Yaghoub, Renae Low, and John Sweller. 1995. "Reducing cognitive load by mixing auditory and visual presentation modes." *Journal of Educational Psychology* 87:319-334. doi: 10.1037/0022-0663.87.2.319.

Mynatt, Elizabeth D. 1990. *Auditory Presentation of Graphical User Interfaces*. Edited by G. Kramer, *Auditory Display: Sonification, Audification and Auditory interfaces*: Addison-Wesley.

Neal, J. G., C. Y. Thielman, Z. Dobes, S. M. Haller, and S. C. Shapiro. 1989. "Natural language with integrated deictic and graphic gestures." Proceedings of the workshop on Speech and Natural Language, Cape Cod, Massachusetts.

Nepper, Patrick, Georg Treu, and Axel Küpper. 2008a. "Adding Speech to Location-based Services." *Wireless Personal Communications* 44 (3):245-261. doi: 10.1007/s11277-007-9378-9.

Nepper, Patrick, Georg Treu, and Axel Küpper. 2008b. "Adding speech to location-based services." *Wireless Personal Communications* 44:245-261. doi: 10.1007/s11277-007-9378-9.

Niculescu, A. I., Yeo Kheng Hui, L. F. D'Haro, Kim Seokhwan, Jiang Ridong, and R. E. Banchs. 2014. "Design and evaluation of a conversational agent for the touristic domain." Asia-Pacific Signal and Information Processing Association, 2014 Annual Summit and Conference (APSIPA), 9-12 Dec. 2014.

Nigay, Laurence, and Joelle Coutaz. 1993. "A design space for multimodal systems: concurrent processing and data fusion." Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems, Amsterdam, The Netherlands.

Nurminen, Antti, Juha Järvi, and Matti Lehtonen. 2014. "A Mixed Reality Interface for Real Time Tracked Public Transportation."

O'Sullivan, David, and Alasdair Turner. 2001. "Visibility graphs and landscape visibility analysis." *International Journal of Geographical Information Science* 15 (3):221-237. doi: 10.1080/13658810151072859.

Oviatt, Sharon. 2002. "Advances in the robust processing of multimodal speech and pen systems." In *Multimodal interface for human-machine communication*, 203-218. World Scientific Publishing Co., Inc.

Oviatt, Sharon. 2003. "Multimodal interfaces." In *The human-computer interaction handbook*, edited by A. Jacko Julie and Sears Andrew, 286-304. L. Erlbaum Associates Inc.

Oviatt, Sharon, Phil Cohen, Lizhong Wu, John Vergo, Lisbeth Duncan, Bernhard Suhm, Josh Bers, Thomas Holzman, Terry Winograd, James Landay, Jim Larson, and David Ferro. 2000. "Designing the user interface for multimodal speech and pen-based gesture applications: state-of-the-art systems and future research directions." *Hum.-Comput. Interact.* 15 (4):263-322. doi: 10.1207/s15327051hci1504_1.

Oviatt, Sharon, and Philip Cohen. 2000. "Perceptual user interfaces: multimodal interfaces that process what comes naturally." *Communications of the ACM* 43:45-53. doi: 10.1145/330534.330538.

Page, Tom. 2015. "A Forecast of the Adoption of Wearable Technology." *International Journal of Technology Diffusion (IJTD)* 6 (2):12-29.

Patel, Mitesh S, David A Asch, and Kevin G Volpp. 2015. "Wearable devices as facilitators, not drivers, of health behavior change." *Jama* 313 (5):459-460.

Perakakis, M, and Alexandros Potamianos. 2008. "Multimodal system evaluation using modality efficiency and synergy metrics." *... international conference on Multimodal ...*:9-16. doi: 10.1145/1452392.1452397.

Persson, P, F Espinoza, V Fagerberg, A Sandin, and R Cöster. 2003. "GeoNotes: A Location-based Information System for Public Spaces." *Designing Information Spaces: The Social Navigation Approach Computer Supported Cooperative Work* pp 151-173. doi: 10.1007/978-1-4471-0035-5_7.

Pfleger, Norbert. 2004. "Context based multimodal fusion." *Proceedings of the 6th international conference on Multimodal interfaces - ICMI '04*:265. doi: 10.1145/1027933.1027977.

Pham Thi, ThanhThoa, Linh Truong-Hong, Junjun Yin, and JamesD Carswell. 2013. "Exploring Spatial Business Data: A ROA Based eCampus Application." In *Web and Wireless Geographical Information Systems*, edited by SteveH L. Liang, Xin Wang and Christophe Claramunt, 164-179. Springer Berlin Heidelberg.

Rahman, A. S. M. M., and A. El Saddik. 2013. "Mobile PointMe-based spatial haptic interaction with annotated media for learning purposes." *Multimedia Systems* 19 (2):131-149. doi: Doi 10.1007/S00530-012-0285-X.

Rajput, Nitendra, and Amit Anil Nanavati. 2012. *Speech in Mobile and Pervasive Environments*: Wiley Publishing.

Reithinger, Norbert, Jan Alexandersson, Tilman Becker, Anselm Blocher, Ralf Engel, Markus Löckelt, Jochen Müller, Norbert Pfleger, Peter Poller, Michael Streit, and Valentin Tschernomas. 2003. *SmartKom: adaptive and flexible multimodal access to multiple applications*.

Retscher, G, Zhu, M and Zhang, K 2012. "RFID positioning." In *Ubiquitous Positioning and Mobile Location-Based Services in Smart Phones*, edited by Ruizhi Chen, 69-95. Information Science: Hershey.

Robinson, Simon , Parisa Eslambolchilar, and Matt Jones. 2008. "Point-to-GeoBlog: gestures and sensors to support user generated content creation." Proceedings of the 10th international conference on Human computer interaction with mobile devices and services, Amsterdam, The Netherlands.

Rodgers, M. M., V. M. Pai, and R. S. Conroy. 2015. "Recent Advances in Wearable Sensors for Health Monitoring." *Sensors Journal, IEEE* 15 (6):3119-3126. doi: 10.1109/JSEN.2014.2357257.

Rogers, Eliot, Tony Rodriguez, John Lord, and Adnan Alattar. 2015. "Digimarc Discover on Google Glass." IS&T/SPIE Electronic Imaging.

Roos, A, A.Th. Schwarzbacher, and S. Wieland. 2006. "Broadband Wireless internet access in public transportation." In Proc. of VDE Kongress - Innovations in Europe, Aachen, Germany, October

Rousseau, Cyril, Yacine Bellik, Frédéric Vernier, and Didier Bazalgette. 2006a. "A framework for the intelligent multimodal presentation of information." *Signal Processing* 86:3696-3713. doi: 10.1016/j.sigpro.2006.02.041.

Rousseau, Cyril, Yacine Bellik, Fredric Vernier, and Didier Bazalgette. 2006b. "A framework for the intelligent multimodal presentation of information." *Signal Process.* 86 (12):3696-3713. doi: 10.1016/j.sigpro.2006.02.041.

Ruddarraju, Ravikrishna, Antonio Haro, Kris Nagel, Quan T. Tran, Irfan a. Essa, Gregory Abowd, and Elizabeth D. Mynatt. 2003. "Perceptual user interfaces using vision-based eye tracking." *Proceedings of the 5th international conference on Multimodal interfaces - ICMI '03*:227. doi: 10.1145/958468.958475.

Salton, Gerard, and Michael J. McGill. 1983. *Introduction to modern information retrieval, McGraw-Hill computer science series*. New York: McGraw-Hill.

Santos, E S, E A Lamounier, and A Cardoso. 2011. "Interaction in Augmented Reality Environments Using Kinect." Virtual Reality (SVR), 2011 XIII Symposium on.

Sarter, Nadine B. 2006. "Multimodal information presentation: Design guidance and research challenges." *International Journal of Industrial Ergonomics* 36:439-445. doi: 10.1016/j.ergon.2006.01.007.

Schickler, Marc, Rüdiger Pryss, Johannes Schobel, and Manfred Reichert. 2015. "An Engine Enabling Location-based Mobile Augmented Reality Applications."

Schiller, Jochen, and Agnes Voisard. 2004. *Location Based Services*: Morgan Kaufmann Publishers Inc.

Schön, Bianca., Bianca. Schön, Debra F Laefer, Debra F Laefer, Sean W Morrish, Sean W Morrish, Michela. Bertolotto, and Michela. Bertolotto. 2009. "Three-Dimensional Spatial Information Systems: State of the Art Review." *Recent Patents on Computer Science* 2:21-31.

Schutte, N, J Kelleher, and B MacNamee. 2009. "A mobile multimodal dialogue system for location based services." IT & T Conference, Dublin, Ireland., 22 - 23 October.

Sebe, Nicu. 2009. "Multimodal interfaces: Challenges and perspectives." *Journal of Ambient Intelligence and Smart Environments* 1:23-30. doi: 10.3233/AIS-2009-0003.

Seo, Yoondeuk, and Jinho Ahn. 2013. "Hybrid Contents Recommendation Service Using LBS and NFC Tagging." *International Journal of Smart Home* 7 (5):251-262.

Sharif, Masoud, Babak Hassibi, Babak Hassibi, and California Institute of Technology. Division of Engineering and Applied Science. 2006. "Broadband wireless broadcast channels : throughput, performance, and PAPR reduction." Caltech,. http://resolver.caltech.edu/CaltechETD:etd-08292005-100440.

Shinohara, Kristen, and Josh Tenenberg. 2009. "A blind person's interactions with technology." *Communications of the ACM* 52:58. doi: 10.1145/1536616.1536636.

Shneiderman, Ben. 2005. "Designing the user interface." *Designing the User Interface* 19:624. doi: 10.1049/tpe.1975.0323.

Simon, R, P Fröhlich, G Gerhard Obernberger, and E Wittowetz. 2007. "The Point to Discover GeoWand." In Proceedings of the 9th International Conference on Ubiquitous Computing, Innsbruck, Austria, September 16-19.

Simon, Rainer, and Peter Fröhlich. 2007. "A mobile application framework for the geospatial web." *Proceedings of the 16th international conference on World Wide Web - WWW '07*:381. doi: 10.1145/1242572.1242624.

Simon, Rainer, Peter Fröhlich, and Hermann Anegg. 2007. "Enabling Spatially Aware Mobile Applications." *Transactions in GIS* 11 (5):783-794. doi: 10.1111/j.1467-9671.2007.01072.x.

Sinha, Anoop K., and James A. Landay. 2003. "Capturing user tests in a multimodal, multidevice informal prototyping tool." Proceedings of the 5th international conference on Multimodal interfaces, Vancouver, British Columbia, Canada.

Stahl, Christoph. 2007. "The roaring navigator: a group guide for the zoo with shared auditory landmark display." Proceedings of the 9th international conference on Human computer interaction with mobile devices and services, Singapore.

Strachan, S., and R. Murray-Smith. 2009. "Bearing-based selection in mobile spatial interaction." *Personal and Ubiquitous Computing* 13 (4):265-280. doi: Doi 10.1007/S00779-008-0205-4.

Strachan, S., J. Williamson, and R. Murray-Smith. 2007. "Show me the way to Monte Carlo: density-based trajectory navigation." Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, San Jose, California, USA.

Strachan, Steven, Parisa Eslambolchilar, Roderick Murray-Smith, Stephen Hughes, and Sile O'Modhrain. 2005. "GpsTunes: controlling navigation via audio feedback."

Proceedings of the 7th international conference on Human computer interaction with mobile devices & services, Salzburg, Austria.

Tatzgern, Markus, Raphael Grasset, Eduardo Veas, Denis Kalkofen, Hartmut Seichter, and Dieter Schmalstieg. 2015. "Exploring real world points of interest: Design and evaluation of object-centric exploration techniques for augmented reality." *Pervasive and Mobile Computing* 18 (0):55-70. doi: http://dx.doi.org/10.1016/j.pmcj.2014.08.010.

Teixeira, António, Annika Hämäläinen, Jairo Avelar, Nuno Almeida, Géza Németh, Tibor Fegyó, Csaba Zainkó, Tamás Csapó, Bálint Tóth, and André Oliveira. 2014. "Speech-centric multimodal interaction for easy-to-access online services– A personal life assistant for the elderly." *Procedia computer science* 27:389-397.

Teixeira, António J. S. , Carlos Pereira, Miguel Oliveira e Silva, Osvaldo Pacheco, António J. R.   Neves, and José Casimiro. 2011. "AdaptO - Adaptive Multimodal Output." 1st International Conference on Pervasive and Embedded Computing and Communication Systems,, Vilamoura, Algarve, Portugal, 5-7 March.

Tian, Feng, Tiegang Cheng, Hongan Wang, and Guozhong Dai. 2006. "Research on user-centered design and recognition pen gestures." Proceedings of the 24th international conference on Advances in Computer Graphics, Hangzhou, China.

Turk, Matthew. 2014. "Multimodal interaction: A review." *Pattern Recognition Letters* 36:189-195. doi: 10.1016/j.patrec.2013.07.003.

Turk, Matthew, and George Robertson. 2000. "Perceptual User Interfaces." *Communications of the ACM* 43:32-34. doi: 10.1145/330534.330535.

Vazquez-Alvarez, Yolanda, and Stephen Brewster. 2009. "Investigating background & foreground interactions using spatial audio cues." CHI '09 Extended Abstracts on Human Factors in Computing Systems, Boston, MA, USA.

Vazquez-Alvarez, Yolanda, Ian Oakley, and Stephen A. Brewster. 2012. "Auditory display design for exploration in mobile audio-augmented reality." *Personal and Ubiquitous Computing* 16:987-999. doi: 10.1007/s00779-011-0459-0.

Wakkary, Ron, Kenneth Newby, Marek Hatala, Dale Evernden, and Milena Droumeva. 2004. "Interactive Audio Content : An Approach to Audio Content for a

Dynamic Museum Experience through Augmented Audio Reality and Adaptive Information Retrieval." *Museums and the Web Conference*.

Walker, Bruce N, and Jeffrey Lindsay. 2006. "Navigation performance with a virtual auditory display: effects of beacon sound, capture radius, and practice." *Human factors* 48:265-278. doi: 10.1518/001872006777724507.

Wang, F, and J Liu. 2008. "Towards 3D LBS-challenges and opportunities." *Proceedings of ISPRS Congress*.

Wenzel, Elizabeth M. 1994. *Spatial sound and sonification*. Edited by G. Kramer, *Auditory display: sonification, audification, and auditory interfaces*: Addison Wesley.

Werner, M., M. Kessel, and C. Marouane. 2011. "Indoor positioning using smartphone camera." Indoor Positioning and Indoor Navigation (IPIN), 2011 International Conference on, 21-23 Sept. 2011.

WiGLE. 2007. "WiGLE (Wireless Geographic Logging Engine) ". https://wigle.net/.

Williamson, John, Roderick Murray-Smith, and Stephen Hughes. 2007. "Shoogle: excitatory multimodal interaction on mobile devices." Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, San Jose, California, USA.

Wilson, Andrew , and Steven Shafer. 2003. "XWand: UI for intelligent spaces." Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Ft. Lauderdale, Florida, USA.

Xu, Qianli, Michal Mukawa, Liyuan Li, Joo Hwee Lim, Cheston Tan, Shue Ching Chia, Tian Gan, and Bappaditya Mandal. 2015. "Exploring users' attitudes towards social interaction assistance on Google Glass." Proceedings of the 6th Augmented Human International Conference, Singapore, Singapore.

Yapeng, Wang, Yang Xu, Zhao Yutian, Liu Yue, and L. Cuthbert. 2013. "Bluetooth positioning using RSSI and triangulation methods." Consumer Communications and Networking Conference (CCNC), 2013 IEEE, 11-14 Jan. 2013.

Yeung, WL. 2015. "Work in Progress Modality Selection Based on Agent Negotiation for Voice Search on Smartphones Paired With Wearable Accessories: A Preliminary Architectural Design."

Yin, Junjun. 2013. "Mobile 2D and 3D Spatial Query Techniques for the Geospatial Web." PhD, School of Media, Dublin Institute of Technology.

Yin, Junjun, and James D Carswell. 2013. "Spatial Search Techniques for Mobile 3D Queries in Sensor Web Environments." *ISPRS International Journal of Geo-Information* 2 (1):135-154.

Yin, Junjun, and James D. Carswell. 2011. "Touch2Query enabled mobile devices: A case study using OpenStreetMap and iPhone." *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 6574 LNCS:203-218. doi: 10.1007/978-3-642-19173-2_16.

Yu, Yi, Roger Zimmermann, and Suhua Tang. 2015. "Social Interactions over Location-Aware Multimedia Systems." In *Multimedia Data Mining and Analytics*, 117-146. Springer.

Zaguia, A, M Hina, C Tadj, and A Ramdane-Cherif. 2010. "Interaction Context-aware Modalities and Multimodal Fusion for Accessing Web Services. ." *Ubiquitous Computing and Communication* 5 (4).

Zaid, Farid. 2010. "Towards robust and efficient visibility-aware sensor-based mobile spatial interaction." 2010 8th IEEE International Conference on Pervasive Computing and Communications Workshops, PERCOM Workshops 2010.

Zickuhr, Kathryn , and Aaron Smith. 2010. "4% of online Americans use location-based services." Accessed 6th March. http://www.pewinternet.org/2010/11/04/4-of-online-americans-use-location-based-services/.

Zirari, Soumaya, P. Canalda, and F. Spies. 2010. "WiFi GPS based combined positioning algorithm." Wireless Communications, Networking and Information Security (WCNIS), 2010 IEEE International Conference on, 25-27 June 2010.

Zlatanova, Siyka, and Edward Verbree. 2003. "Technological developments within 3D location-based services." International symposium and exhibition on geoinformation (ISG03), Shah Alam, October 13-14.

5. Appendix

5.1. Paper I: MobiLAudio - A Multimodal Content Delivery Platform for Geo-Services

Gardiner, K., C. Cullen and J. D. Carswell. (2015) "MobiLAudio - A Multimodal Content Delivery Platform for Geo-Services", Journal of Location Based Services, Taylor and Francis. (Accepted - In Press)

MobiLAudio – a multimodal content delivery platform for geo-services

Keith Gardiner, Charlie Cullen and James D. Carswell

Digital Media Centre, Dublin Institute of Technology, Dublin, Ireland

ABSTRACT

Delivering high-quality context-relevant information in a timely manner is a priority for location-based services (LBS) where applications require an immediate response based on spatial interaction. Previous work in this area typically focused on ever more accurately determining this interaction and informing the user in the customary graphical way using the visual modality. This paper describes the research area of multimodal LBS and focuses on audio as the key delivery mechanism. This new research extends familiar graphical information delivery by introducing a geoservices platform for delivering multimodal content and navigation services. It incorporates a novel auditory user interface (AUI) that enables delivery of natural language directions and rich media content using audio. This unifying concept provides a hands-free modality for navigating a mapped space while simultaneously enjoying rich media content that is dynamically constructed using such mechanisms as algorithmic music and phrase synthesis to generate task-relevant content based on the path taken. This paper outlines the innovative ideas employed in the design of the AUI and details the geoservices platform developed for facilitating the authoring and delivery of multimodal LBS applications. The paper concludes with a discussion on the results of a live user trial. The results are analysed and presented to validate the original hypothesis for this approach, address the research questions outlined and to inform further research directions. The results show that the proposed solution significantly progresses the state-of-the-art in terms of mobile tour production. The results also show that an AUI is an effective modality for the delivery of audio content and natural directions when used in combination with a graphical user interface, producing significantly reduced overheads in terms of content size and network usage. The results also indicate that the AUI provides a good overall user experience, performing well in the user trial.

1. Introduction

This paper describes research and development of a novel geoservices platform for multimodal delivery of high-quality and task-relevant content to constrained mobile devices

CONTACT   Keith Gardiner   ✉ keith.gardiner@dit.ie

(e.g. spatially enabled smartphones). Importantly, it investigates delivering location-based content to the user by means of the phone's aural modality as the primary delivery mechanism. The visual modality is also presented using a graphical user interface (GUI) that incorporates and builds on many of the ideas developed in previous work on mobile spatial interaction for location-based services (LBS) (Carswell, Gardiner, and Neumann 2003; Gardiner and Carswell 2003; Gardiner 2004; Gardiner et al. 2005; Rizzini et al. 2006; Kilfeather et al. 2007; Rooney, Gardiner, and Carswell 2007; Gardiner, Yin, and Carswell 2009; Carswell, Gardiner, and Yin 2010a,b).

The aural modality incorporates an auditory user interface (AUI), which is the main innovation in terms of delivering context-sensitive information to the user using focus independent (eyes-free) means. The AUI is intended to be a non-visual interface that can be used in combination with a GUI, or not. The idea is that all information in the application (including media content, directions, advertisements, etc.) can be delivered without the user ever needing to touch the device.

There are two main motivations for this work. Firstly, the concept of multimodality is investigated in terms of both multimodal input and output. Historically, mobile applications used just one form of multimodal input or output. More recently, the idea of combining multiple modalities at the input and output stages has gained significant traction in the mobile spatial interaction research community. Combining touch/gesture interaction with spatial navigation can provide a unique method of information discovery to enhance the overall user experience. Similarly, combining modalities at output level by using the visual and aural modalities together can provide a synergistic method of content delivery. The second motivation for this work is the development of an AUI for geoservice content delivery. The idea of combining mobile spatial interaction data with high-quality media content in the form of audio assets is explored where the goal is to enhance the user experience using multiple modalities while achieving a significant reduction in processing/data overheads. By using a phrase synthesizer to dynamically construct all content delivered by the system, physical storage requirements for the content is significantly reduced, resulting in the possibility to store all content data in-app. This further reduces network overhead as the requirement for external data calls is minimized.

Considering the above motivations, there are two main research questions related to this work. The first of these (Q1) is concerned with issues surrounding mobile tour production. To date, producing mobile tours necessitates a significantly arduous manual process, where geodata is collected and combined/linked with business data (content) to produce the point-of-interest (POI) objects used by the application. This manual process is repeated for each tour and requires significant resources. One important question is whether this process can be improved by reducing this manual overhead. Our work proposes a solution by developing a mobile tour production platform for audio tours, where each stage feeds data to the next using a structured workflow. In traditional GUI-based applications, this has been achieved already; however in our case, the intended output is a multimodal application that also delivers content using an AUI. A second important research question (Q2) is whether the use of an AUI is even effective in the delivery of content for geoservices. The hands-free and eyes-free nature of audio as a delivery mode has significant advantages over the focus dependant nature of GUIs; however, GUIs do provide more control. One of our aims therefore is to measure the overall user experience while using the AUI as the primary modality for content delivery.

The proposed approach to addressing the research questions outlined above is the development of a platform for multimodal geoservice delivery that will incorporate an AUI for multimodal content exploration. There are four main parts to this research.

- Survey Stage – Development of a mobile app for collecting region-specific geo-information.
- Author Stage – Implementation of an authoring tool that focuses on the structure of the content and how the overall process can be optimized.
- Delivery Stage – Development of a mobile app for delivery of content to users in a clear and natural way using an AUI and/or GUI.
- User Trial Stage – Testing the effectiveness of the multimodal tour guide in terms of functionality, usability and user experience.

The fourth and arguably most significant part of this research is the live user trial that was performed to test the functionality, usability, and overall user experience of the app. The user trial was carried out in Dublin City where a James Joyce Tour was created using the three-stage process outlined above and released as part of James Joyce week. The trial included 31 participants who used the app to navigate through the City Centre, visiting 12 plaques on route relating to James Joyce's book *Ulysses*. The app provided information on each plaque, directions between plaques and information on local amenities and commercial offers.

In Section 2, related work is presented with a focus on four of the most relevant aspects of previous research in this area, LBS, mobile spatial interaction and AUIs and multimodal content delivery. In Section 3, the geoservices platform is introduced which describes the three-stage process based on the idea of a survey, author and delivery methodology for app development, and deployment. Section 4 gives a comprehensive account of the live user trial carried out on the AUI including a detailed analysis of the results, outlining relevant observations, outcomes and recommendations. In Section 5, conclusions are drawn and further work is proposed.

2. Background

Location-based information retrieval has been a very active and pertinent area of research for many years (Bellotti et al. 2002; Persson et al. 2003; Wilson and Shafer 2003; Shi et al. 2010). Since the first methods of digitally creating maps, POI locations and geotags, the idea of providing context-sensitive services based on a user's location supported by an underlying geographic map have become increasingly important. These developments form the basis for previous research in this area and in particular work in the areas of location-based information retrieval, mobile spatial interaction, AUIs and multimodal content delivery.

2.1. Location-based information retrieval

With advances in mobile device technology, the idea of LBS has become a reality in many ways. However, even before such devices were commercially available, initial research presented by (Carswell et al. 2002; Gardiner 2004), describes work that simulated the concept of LBS by querying an underlying map on a desktop PC. Using a user's position and orientation in the map, a content delivery service was provided in the form of simple text and images. A 3D virtual map of Dublin City was developed to demonstrate and monitor the user's

physical interaction within the environment. The map was stored in a spatial database and registered to real-world coordinates together with a number of POIs representing notable locations relevant to Ireland's 'Easter Rising' in 1916. Using these three pieces of structured data, an early example of the potential of LBS was demonstrated.

Subsequent research focused on the accurate detection of POIs in a 2D space (Gardiner and Carswell 2003) based on the senses, i.e. how we identify and consume information presented to us. The visual modality was studied as the primary means of consuming information and more importantly the restrictions that line-of-sight (LOS) presented (Gardiner 2004). Defining actual visibility presented varying degrees of complexity because the location and definition of the interaction have to be accurate (Carswell et al. 2002; Rooney, Gardiner, and Carswell 2007), to effectively deliver location-based information in the correct context.

2.2. Mobile spatial interaction

Mobile spatial interaction is the study of the interaction between a mobile user and a world of spatial information. In (Gardiner and Carswell 2003; Gardiner 2004; Chan et al. 2005; Fröhlich et al. 2006; Simon and Fröhlich 2007a, 2007b; Simon, Fröhlich, and Grechenig 2008; Gardiner, Yin, and Carswell 2009; Carswell, Gardiner and Yin 2010a), the idea of an egocentric view of the world is investigated in a mobile context. In (Gardiner and Carswell 2003), the idea of using a triangular 2D shape to represent a user's field-of-view (FOV) to query POIs stored in a spatial database was introduced. By using a user's actual FOV as the 'query window' in the database, the query results are limited to what is actually visible to the user in a particular direction. This research proved successful in reducing the associated problem of information overload.

To increase query relevance and reduce information overload further, in (Strachan and Murray-Smith 2009; Gardiner, Carswell and Yin 2009), the idea of generating a more accurate FOV was demonstrated. This research focused on determining the exact shape of the user's viewport. Extending this an Isovist algorithm (Benedikt 1979) was used to determine 360° LOS. This resulted in a more representative viewport shape, minus the geometry of any buildings or objects that interrupt a user's LOS.

In (Strachan, Williamson, and Murray-Smith 2007; Fröhlich, Simon, and Baillie 2009; Gardiner, Yin, and Carswell 2009; Carswell, Gardiner and Yin 2010b), the focus shifted to the area of 3D directional queries, where the idea of 'point-to-query' was investigated. Before this, the idea of performing this specific type of spatial search was not feasible for the following reasons. Firstly, although storing 3D data was possible in some commercial database management systems at the time, performing 3D queries at database level was not. This was due to a lack of built-in 3D spatial query operators. Secondly, the availability of mobile devices that contained the sensor hardware necessary to measure both the horizontal and vertical orientations of the device was still limited to high-end or customized research devices. However, with the introduction of smartphones and improvements in spatial database technology, these barriers diminished and new research in the area of mobile 3D queries was now possible (Robinson, Eslambolchilar, and Jones 2008; Gardiner, Yin, and Carswell 2009).

A practical example of mobile egocentric-visibility querying was successfully demonstrated in a simulated *Internet of Things* environment by the *3DQ* query processor developed in (Carswell, Gardiner and Yin 2010a; Yin and Carswell 2013). *3DQ* enabled a user to point a smartphone anywhere at a building to obtain information about any visible feature (e.g. window, door, floor, environmental sensors, etc.).

2.3. Auditory user interfaces

An AUI is defined as 'the use of sound to communicate information about the state of an application or computing device' (McGookin 2004; Shneiderman 2005). AUIs have traditionally been used for accessibility by the visually impaired in desktop applications with the main objective of transforming a graphical representation to an auditory one (Edwards 1989; Gaver 1989; Mynatt 1994; Donker, Klante, and Gorny 2002; Jagdish et al. 2008; Shinohara and Tenenberg 2009). Auditory non-speech audio has also been used to provide additional information to users about navigating a GUI (Kramer 1993; Absar and Guastavino 2008). This idea has gained popularity in the LBS and augmented reality (AR) space to provide navigational information to users of mobile applications (Walker and Lindsay 2006; Baus et al. 2007; Stahl 2007).

In (Holland, Morse, and Gedenryd 2002), the use of tones and audio panning is used in an AudioGPS application to communicate direction and distance to direct a user through a space. User trials show that the use of simple tones and the placement of them on the stereo stage are effective for indicating direction and that a simple oscillating tone was effective for communicating distance. Related work by (Strachan et al. 2005; Strachan, Williamson, and Murray-Smith 2007) investigated the use of music for the purposes of navigation where a listener's music is modulated according to the changing predictions of a user's position to guide them through a space.

In the context of multimodal geoservice delivery, the use of AUIs as an output modality has shown significant potential in the area of multimodal content delivery. In addition to providing instructional and navigational content, AUIs can deliver audio content in the form of real or synthesized speech to inform users about the surrounding environment (Abowd et al. 1997). This approach was previously used to augment the environment by providing context-sensitive information with applications in the field of mobile tour guides (Bederson 1995).

2.4. Multimodal content delivery

The concept of mobile multimodal content delivery therefore has been the topic of considerable research in recent times with multiple output modalities now available on today's off-the-shelf devices (Chittaro 2010). Combining output modalities to provide information in a synergistic manner by augmenting a space has also gained attention (Lalanne et al. 2009; Turk 2014). The use of audio for content delivery is described in (Bartie and Mackaness 2006) where a combination of visibility querying and audio AR is used to create a CityGuide application. The application proved to be an effective, non-invasive approach to augmenting the user's reality as it offered hands-free and eyes-free usage, enabling the user to keep focused on the environment.

The use of spatialized soundscapes to convey contextual information has been used in many cases in an attempt to deliver an immersive user experience (Wenzel 1994; Misra, Perry Cook, and Wang 2006; Stahl 2007; Vazquez-Alvarez, Oakley, and Brewster 2012). The LISTEN project (Eckel 2001) provides users with intuitive access to personalized and situated audio information to enhance the sensual impact of applications ranging from art installations to entertainment events. This is achieved by augmenting the physical environment using immersive, spatial audio and advanced user modelling methods.

An interesting approach was taken by (Heller et al. 2009) with the development of an interactive virtual tour where users can remain independent or can interact with other users on the tour, essentially 'checking in' with them. The ec(h)o system described in (Wakkary et al. 2004; Droumeva 2005) uses an audio AR interface that is used to deliver exhibition data based on a semantic structure and utilizes spatialized soundscapes to provide an immersive experience. Work carried out in (McGee and Cullen 2009a, 2009b) suggested the usefulness of audio as a delivery modality for LBS. In particular, the advantages identified were both its hands-free and eyes-free nature, plus fast neural processing rate. An experimental case study in (Vazquez-Alvarez, Oakley, and Brewster 2012) tests the use of earcons, proximity and spatial audio queues for effective information discovery using audio. A combination of 3D spatial audio techniques together with earcons proved the most effective audio display. This work also suggests that the location and orientation sensing technologies available in off-the-shelf devices are effective when used to create rich and compelling outdoor augmented-audio environments. All these works point to the benefits of the hands-and-eyes-free nature of non-visual modalities and to the potential the aural modality has for the provision of navigational and context-relevant information in a geoservices context.

3. MobiLAudio

The *MobiLAudio* system described in this paper extends the above related work by proposing a mobile audio tour guide production platform that incorporates a structured workflow for the collection of data, the production of media assets and the effective multimodal delivery of region specific information all in one application.

There are two significant innovations that differentiate the MobiLAudio platform from previous work, the first is the workflow itself which uses a unique survey–author–deliver metaphor for tour production, and secondly is the method by which the content is synthesized and delivered using an AUI at the delivery stage. The *MobiLAudio* platform builds on the work carried out in the development of the *mobiSurround* engine (Gardiner, Cullen, and Carswell 2015) with an emphasis on dynamic tour production and multimodal content delivery. In earlier research (Carswell et al. 2002, 2008; Carswell, Gardiner, and Neumann 2003; Gardiner and Carswell 2003; Rizzini et al. 2006, 2010), numerous issues with the amount and size of the media content required were encountered. In many cases, the amount of redundant data in the system is very high, where multiple files contained much the same content with small variations. In MobiLAudio, because a sizable amount of the content is now audio, the opportunity to reduce data redundancy using phrase synthesis is exploited.

The concept of phrase synthesis is based on taking a dictionary of recorded words and phrases, and combing them into sentences. This technique can deliver the required content via the AUI using a set of pre-recorded words and phrases that are synthesized at runtime using a set of production rules, which significantly reduces the required content in terms of size, thus reducing the overall size of the app. Taking advantage of this reduction in overhead, further research was carried out in the area of multichannel content delivery via the AUI. The multichannel AUI can provide up to 32 channels of audio information on a mobile device (Apple Developer Library 2014). In addition to the multiple channels of speech audio, multiple channels of background audio are used in the form of algorithmic music and 3D soundscaping to provide ambient audio and other context-relevant sound effects.
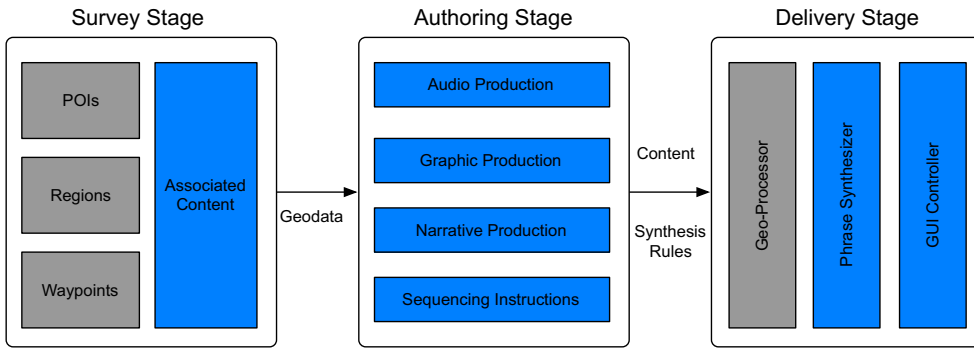
Figure 1. MobiLAudio system architecture diagram.
Note: This diagram describes the overall system architecture, which includes the survey, author and delivery stages of the tour production process.

Combining these, the AUI can deliver an immersive experience using real speech assets, background music generated on-the-fly, transient sound effects and natural directions. *MobiLAudio* developed a number of these innovative features, all of which improve and augment the viability of the AUI for real-world applications. Our focus is on the delivery of high-quality authored media content that is organized in such a manner that location events trigger timely media transitions. The *MobiLAudio* platform provides a toolset where professional HD tours (using high fidelity content) can be planned, authored and delivered. The production of a tour involves a three-stage process and begins with the use of a mobile survey tool to collect geodata. These data along with associated field-based content are passed to the content authoring tool, which models the geodata/content for the tour for delivery via multiple mobile platforms. The architecture of the *MobiLAudio* platform is illustrated in Figure 1.

The following three sections outline the three-stage process describing the novel aspects of tour production and multimodal content delivery that differentiates it from other approaches. The main focus is on the authoring and delivery stages of the process as this is where the main advances beyond the state-of-the-art are made (research question Q1).

3.1. Survey stage

The core concept of the *MobiLAudio* platform is the synergy of location data and professionally produced media content. To achieve this, a mobile survey tool was developed to provide a full LBS map of a given space and to geotag content for effective tour production. The tool allows a tour creator to annotate specific locations within a space by attaching rich media in the form of text, images, audio and video (Figure 2). The geodata collected by the survey tool are used to define the three core elements of a tour space, and are referred to as nodes, waypoints and regions.

3.1.1. Nodes
Nodes are considered a basic unit of data in the *Delivery Engine* and are used to represent a basic POI, service or facility (Figure 2). Each node contains a number of pieces of data that firstly identifies the node, the four neighbouring nodes, the closest waypoint and the region
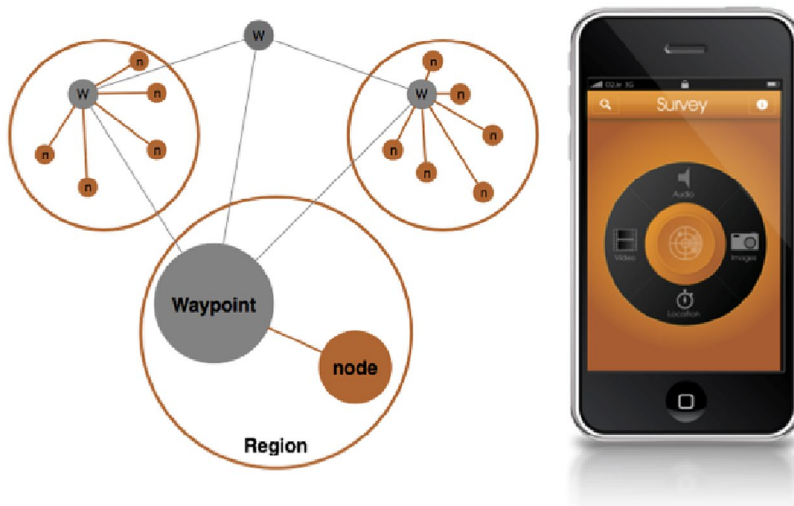
Figure 2. MobiLAudio surveying tool.
Note: The mobile app used to survey a space before the content production stage; it provides the geodata for nodes, waypoints and regions.

the node belongs to. Using the node ID, the engine determines what content to load on arrival, whether to change region and what node is next based on the direction the user has arrived from.

Each node has a number of associated media assets; these include images, text and audio. The audio is split into parts, each containing separate facts about a POI which are dynamically loaded at runtime by the phrase synthesizer. The phrase synthesizer is used in all node content delivery to allow a relatively small set of commentary phrases (delivered by the narrator) to be reused for a given node based on a set of linguistic rules. In this manner, node content is assembled at runtime to give the user the impression of listening to a documentary-style interview between narrator and expert.

3.1.2. Waypoints
Waypoints are the main navigational unit in the *Delivery Engine*. They can be considered as reference or control points for all other nodes in a space, where each node essentially belongs to or is governed by a waypoint such that navigation around a space is achieved by specifying directions between waypoints (Figure 2). Once navigation between waypoints has been achieved, directions from the destination waypoint and all of its associated nodes are accessible. Directions to the destination node are then provided. More significantly, waypoints are defined as natural landmarks that exist in an area making them perfectly suitable for delivering natural language directions. For example, 'walk down as far as the Big Tree and take a left; the shop is on your right hand side'. In this instance, 'the Big Tree' is the destination waypoint and 'the shop' is the destination node.

3.1.3. Regions
Regions are defined as functional areas that divide a space. Each region has a set of physical boundaries that give it a unique geospatial identity. Regions can encompass multiple

waypoints and nodes, essentially helping to define the context in which they exist (Figure 2). Navigating between regions triggers a region change, which in effect signifies a change in context. The result is a managed fade-out of all contextual audio and sound effects for that region, followed by an introduction to the new region and a gradual fade-in of all contextual audio for this region. This subtle effect is a key feature of the adaptive virtualized 3D soundscaping described more in Section 3.3.

With the geodata mapping completed, the data are then used to design a brief for content authoring, indicating where trigger points lie, where waypoints should be and the boundary lines for each region. This helps to define exactly what needs to be rendered by the AUI and when (i.e. content modelling). This is done at the *Authoring Stage* where a set of production guidelines for speech audio, soundscaping, SFX and adaptive music are produced.

3.2. Authoring stage

With the geodata collected for nodes, regions and waypoints, the audio soundscape, adaptive music and narrative content for nodes and regions within a given tour must be composed and structured to trigger accurately as the user navigates the space. This is done using the *Authoring Tool*. The Authoring Tool is used to load the survey information for a given tour, and inform the design and production of soundscape, music, and narrative assets for all points and regions in the tour (Figure 3).

The data collected by the survey tool are raw geodata that needs to be structured and augmented so that it can be used. A data delivery matrix is used to structure the data where a node is defined as a logical POI, service or facility. Waypoints are defined as navigational reference points in the space and regions are logical boundaries that define a space. These basic elements are the spatial foundation of all functionality in the system. The translation of the required functionality into a content model for authoring is undertaken for each of the three elements in the matrix. These elements can be broadly defined in terms of:
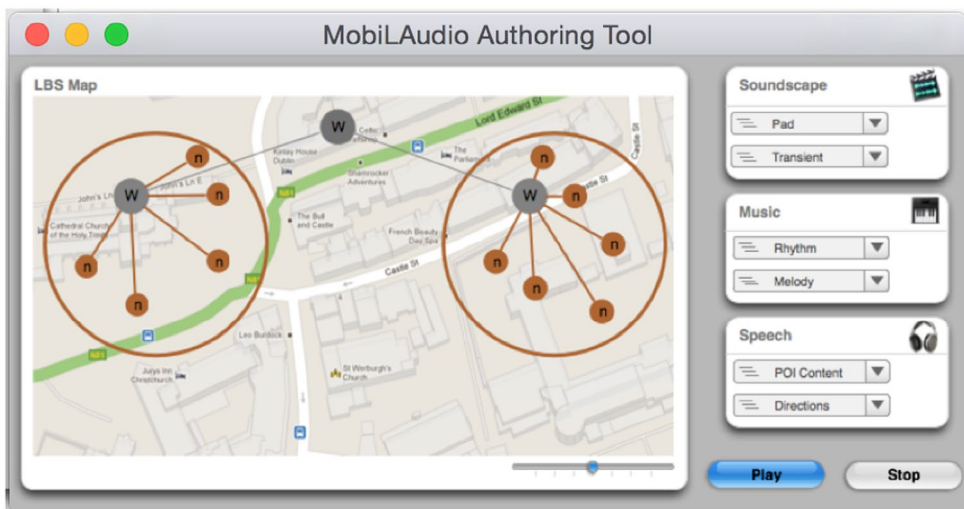

Figure 3. MobiLAudio authoring tool.
Note: Desktop application used to organize audio assets for nodes, waypoints and regions by linking the assets for each node to geodata collected by the survey tool.

- Content Delivery – information relating to a specific node in the space.
- Navigation – information needed to traverse between points in the space using waypoints.
- Narrative Context – changes between defined regions allow the narrative to adapt to the new context of the user.

In order to collect and model tour data for effective delivery of content, navigation instructions, and narrative context, a content modelling strategy is used. This process (outlined below) helps ensure that the AUI is provided with accurate sequencing instructions at runtime.

3.2.1. Content modelling and selection

The central component of the *Authoring Tool* is based on a three-stage data modelling process that informs the various parts of the application that feed both the visual and aural presentation modalities (Figure 4).

3.2.1.1. Geodata modelling.   The geodata modelling (GM) task is a process by which all geographic data that defines POIs, navigational waypoints and regions is modelled. These data inform the rest of system and is the basis of the content modelling strategy.

3.2.1.2. User modelling.   The user modelling (UM) task is where user preferences are modelled, essentially defining the type of user and their context. These data have a scaling effect on the content modelling strategy where varying amounts of content is presented based on a user's profile (Figure 4).

3.2.1.3. Content modelling.   The content modelling task is where the structure of the media content is defined for each node based on the geodata (GM) and application preferences data (UM). It is this strictly controlled process of data modelling from multiple perspectives that ensures data in the system is structured correctly, enabling the dynamic loading of media content to give the perception of an adaptive coherent narrative with real speech content and natural directions, and algorithmic music and 3D soundscaping (Figure 4).

The following section outlines the innovative approach to providing this experience and describes how they integrate to deliver the functionality that forms the basis of the Delivery Engine. In particular, aspects relating to phrase synthesis, adaptive 3D soundscapes, algorithmic music sequencing and natural directions are outlined.

3.3. Delivery stage

Following considerable investigations into the area of contemporary LBS coupled with an increasing focus on the affect that quality media assets have on providing a good user experience; the need for improved, innovative methods of delivering this content became apparent. This led to the development of a content *Delivery Engine* capable of delivering an immersive experience using multiple modalities while minimizing data redundancy using a reusable content approach. The research focused on several distinct areas of functionality necessary to deliver the required content, navigational instructions, music and soundscaping. This was achieved through several advances on the state-of-the-art, notably in the areas of
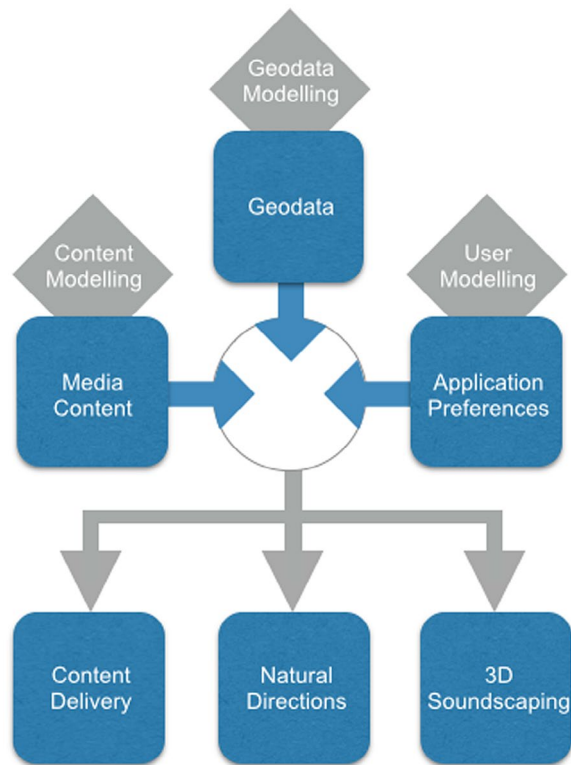
Figure 4. Three-stage data modelling process.
Note: This diagram describes the structure of the data required to build a mobile audio tour based on survey data, contextual information and media content assets.

multimodal content delivery and AUIs. The architecture of the *MobiLAudio Delivery Engine* is illustrated in Figure 5.

The concept of using an AUI for content delivery in a mobile context has been investigated before (Wang et al. 2008; Magnusson et al. 2009); however, in most cases, pre-recorded audio is used without consideration for a user's limited ability to perceive multiple audio sources at once (Potamianos and Perakakis 2008). For example, audibility problems can arise when speech and background music or sound effects are playing together. A model where multiple channels of audio are used to deliver speech, music and soundcaping separately based on priority is a more effective approach and is the main motivation for the development of the *MobiLAudio AUI*. The main problem here (relating back to research question Q2) is whether the use of an AUI based on a prioritized listening model can provide an immersive user experience, while effectively delivering speech content and navigational instructions to the user.

Our approach uses a multichannel audio mixer and phrase synthesizer to deliver the required music, speech and soundscaping based on a background to foreground listening model. This approach advances the state-of-the-art in the area of mobile multimodal interfaces with the development of an AUI that models content and narrative context to effectively deliver speech, navigation instructions, algorithmic music and soundscaping in a LBS context. More specifically the engine comprises:
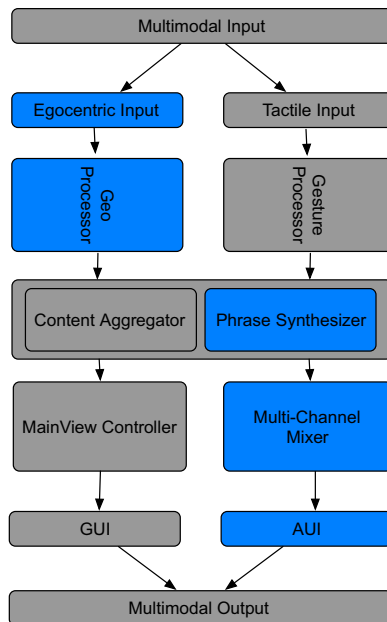
Figure 5. Delivery engine architecture.
Note: The diagram illustrates the architecture of the mobile delivery app, which includes the geoprocessor, phrase synthesizer and GUI controller.

- A background to foreground listening model.
- An intelligent content phrase synthesizer.
- A natural directions model for real-time navigation.
- A location manager with adaptive resolution monitoring.
- Adaptive virtualized 3D soundscapes.
- An algorithmic music sequencer.
- A multichannel auditory mixer.

The delivery platform chosen is an iOS application that uses a GUI in conjunction with an AUI for content delivery using earphones. The GUI for the iOS application is shown in Figure 6.

The most fundamental part of the AUI design is the underlying listening model. The following section describes the listening model and how it forms the foundation of each component in the system. The subsequent five subsections describe the main research contributions made to their respective areas during the development of the AUI. The final subsection describes the multi-channel audio mixer, which details how all components of the AUI integrate to produce an immersive user experience.

3.3.1. Listening model

The AUI delivery model is based on a background to foreground listening paradigm (Truax 2001). Research has shown that using this approach helps users to delineate foreground speech from background audio resulting in improved auditory perception (Beauchamp 2007; Darwin 2008). Aligning the audio features of the AUI with this listening paradigm prioritizes speech over background audio and music. This ensures that speech content and directional instructions are always in the foreground ensuring they are heard and coherent.

Figure 6. Delivery engine graphical user interface (GUI).
Note: This interface is used in manual mode to view data on the map and trigger POI playback.

The listening model illustrated in Figure 7 shows how content within the AUI is delivered to the user. A spatialized 3D soundscape is overlaid with algorithmic music that the user can detect but not prioritize directly. Speech content (either node information or directions) will always be foreground content. The prioritization of node content as the primary function of the AUI is defined as part of the speech model, with waypoint navigation as a specific subset of natural speech directions within this model. Region content is required to give change of context, and so a combination of adaptive virtualized 3D soundscaping and specialized algorithmic music is implemented to convey broader conceptual variations within the narrative of the space.

3.3.2. Intelligent content phrase synthesizer

Concatenative synthesis has been the subject of considerable research in the area of speech and music sound for many years (Schwarz 2005; Neustein and Markowitz 2013). Traditionally the technique synthesizes sounds by concatenating short sounds of up to 1 s, into sounds or words. In speech synthesis, concatenative synthesis methods are most widely used and result in more natural-sounding and intelligible speech (Schwarz 2004). The *MobiLAudio AUI* uses the more simplified approach of unit selection synthesis, where pre-recorded words and phrases are assimilated at runtime by the AUI using a process called phrase synthesis.

The intelligent phrase synthesizer developed delivers speech and natural direction content for mobile tours. The primary motivation for this approach was the reduction in mobile application size, which is an important consideration in mobile app development, and the significant reduction in data redundancy that this method achieves. By using this approach, all content is stored in-app, reducing the cost/reliance on mobile network data connectivity, avoiding issues experienced in previous projects with unreliable data connectivity. The phrase synthesizer is used in all node content delivery to allow a relatively small set of commentary phrases (delivered by the narrator) to be reused for a given node based on a set of linguistic rules. In this manner content for each node is assembled at runtime to give the impression of listening to a documentary-style interview between a narrator and a domain expert.
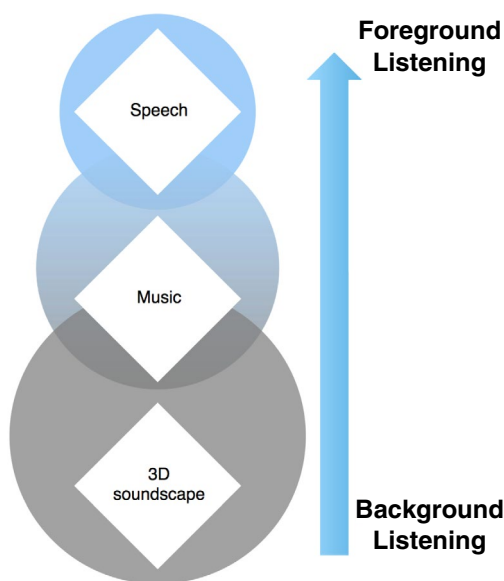
Figure 7. Delivery engine listening model.
Note: This model is based on a background to foreground listening paradigm, prioritizing speech over background and music.

Figure 8 illustrates how the phrase synthesizer assembles node, region, help and direction content sentences in real-time. Each channel (Intro, Content, altContent and Outro) is loaded with audio using a handle and a content phrase. The handle is randomly selected from a predefined library and the content relates to the nodeID of the current active node. An example of a synthesized sentence is as follows:

> [You are now arriving at] + [the Snow Leopard exhibit] + [Lets ask john to give us some information] + [Snow Leopards are large cats that live way up high in the mountains ranges of central Asia] + [Wow that's interesting] + [FACT2] + [Cool can you give us another fact] + [FACT3] + [Interesting, what else can you tell us about it] + [FACT4] + [Well now you have heard about it lets move on to] + [the red river hog exhibit] + [DIRECTIONS]

Each node has two narrators associated with it, the main narrator and a domain expert. Facts 1 and 2 are initially loaded into the Content and altContent channels, and then swapped out for Facts 3 and 4. The phrases are constructed in real-time, loaded into the audio graph (MixerBus) and played. All node content and directions are constructed in this manner, significantly reducing the overhead of recording and playing large static files.

3.3.3. Natural directions model for real-time navigation

The area of personal way finding has gained significant attention in recent years due to the proliferation of GPS enabled mobile devices and personal navigation systems (Klippel et al. 2005; Kainulainen et al. 2007; Modsching et al. 2007; Stark, Riebeck, and Kawalek 2007). The augmentation of traditional map-based approaches with other forms of navigation such as the use of directional arrows and photographs of landmarks has been widespread (Burnett 2000; Chittaro and Burigat 2005; Millonig and Schechtner 2007; Hile et al. 2008).
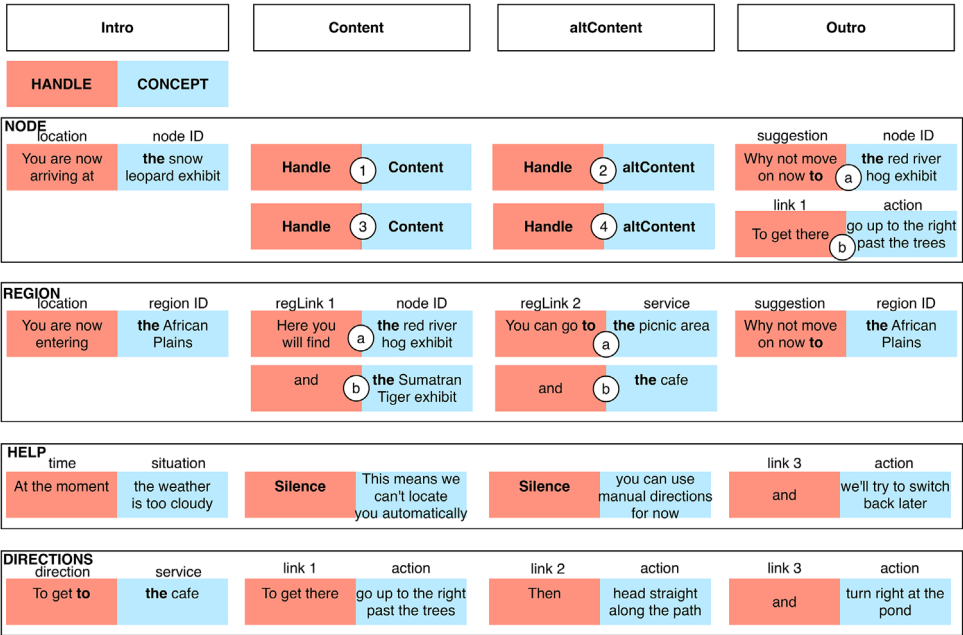
Intro	Content	altContent	Outro

HANDLE	CONCEPT

NODE

location	node ID					suggestion	node ID
You are now arriving at	the snow leopard exhibit	Handle (1) Content		Handle (2) altContent		Why not move on now to	the red river (a) hog exhibit
						link 1	action
		Handle (3) Content		Handle (4) altContent		To get there	go up to the right (b) past the trees

REGION

location	region ID	regLink 1	node ID	regLink 2	service	suggestion	region ID
You are now entering	the African Plains	Here you will find	the red river (a) hog exhibit	You can go to	the picnic area (a)	Why not move on now to	the African Plains
		and	the Sumatran (b) Tiger exhibit	and	the cafe (b)		

HELP

time	situation					link 3	action
At the moment	the weather is too cloudy	Silence	This means we can't locate you automatically	Silence	you can use manual directions for now	and	we'll try to switch back later

DIRECTIONS

direction	service	link 1	action	link 2	action	link 3	action
To get to	the cafe	To get there	go up to the right past the trees	Then	head straight along the path	and	turn right at the pond

Figure 8. Phrase synthesis structure for node, region, help and directions content.
Note: The phrase synthesizer requires 4 MixerBus channels to facilitate the assembly and playback mechanism employed by the AUI.

This has also resulted in the use of audio as an aide to personal way-finding with voice guidance gaining particular interest. In (Rehrl et al. 2014), the use of voice, map and AR are used in the context of pedestrian navigation. Test results show that the use of audio-only guidance leads to significantly better navigation and user experience amongst users. A study in (Rehrl, Häusler, and Leitinger 2010) also indicates that voice-only navigation is preferable over other modalities and that users preferred landmark-based navigation instruction over metric based distances, mainly due to a preference for the more natural eyes-free and hand-free approach of egocentric audio directions (Chittaro and Burigat 2005). However, this approach also presents some difficulties mainly in the area of audio production and landmark surveying, where both aspects require significant manual effort.

These difficulties are the primary motivation to our approach taken in *MobiLAudio*, where a phrase synthesizer is used to reduce the amount of audio required to provide directions, and the initial survey stage of the three-stage tour production model is used to collect information on landmarks. The main research question here is whether natural directions can be constructed and delivered effectively using the survey–author–deliver methodology employed by the *MobiLAudio* platform.

In *MobiLAudio*, natural directions are defined as the use of audio to deliver directions based on visible points or landmarks in a space rather than providing metric distances and cardinal points. The advantages of natural directions are speed (humans do not need to translate data) and adaptability (there is no need to hold a compass in a certain way to obtain a heading) that allow them to be delivered hands free. The disadvantage of natural directions is the size of the content model required, where each direction is a unique narrative in itself.

To address this, the *Delivery Engine* natural directions model uses a combination of node and waypoint information to control the phrase synthesizer, providing a set of natural directions (which can also be displayed on a map if required) to any given point in the space. The model uses the concept of waypoint linkage to create a lookup table of directions between waypoints that is then completed by directions from the last waypoint to the node required. In this manner, natural points of focus within the space are first defined as waypoints using the survey tool and then the authoring tool is used to script and create the directions needed to link them together. When a specific node is queried, the only information required is its nearest waypoint, which allows a lookup table to complete the navigation path to it and subsequently configure the phrase synthesizer (Figure 9).

The example above demonstrates how waypoint linkages are used to provide directions to any node in the space. For the three destination nodes (e.g. zoo exhibits) in Figure 9, a common set of waypoint linkage directions is used to get to waypoint 3, where different waypoint combinations then lead to the specific nodes. Each node is associated with one waypoint only. When a direction query is performed, three pieces of information are needed.

(1) Directions from the current node to its waypoint.
(2) Directions from the current waypoint to the destination waypoint.
(3) Directions from the destination waypoint to the destination node.

By building a set of waypoint linkages in both directions, (getting from W2 to W1 is not the same as going from W1 to W2), each node is linked to its nearest waypoint, which is then linked by a lookup table to every other waypoint. The phrase synthesizer is then loaded with the entries from the lookup table and provides the directions to the user as required (Table 1).
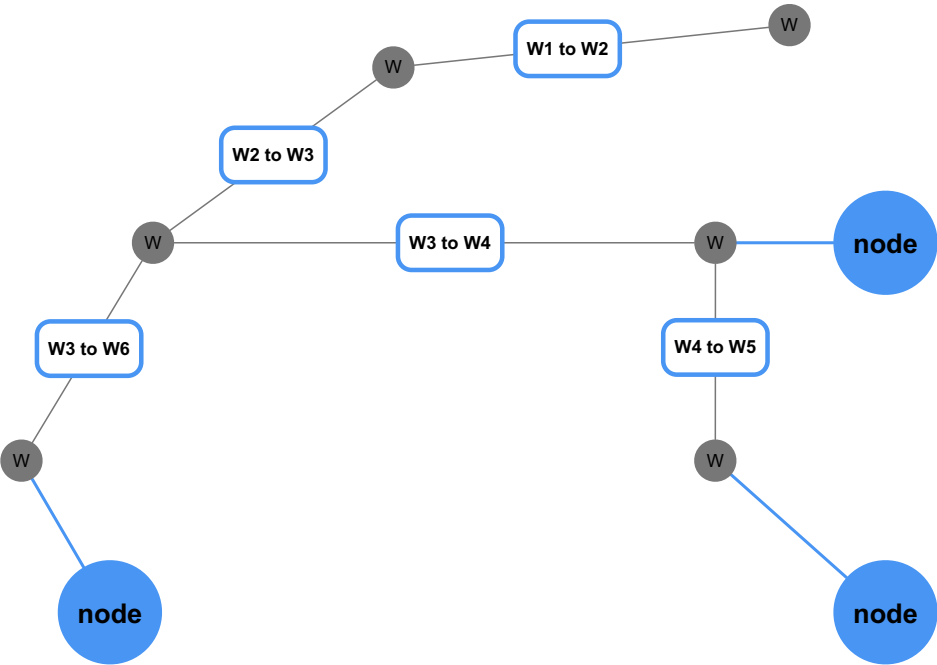

Figure 9. Waypoint linkage example.
Note: Each node is owned by a waypoint; directions are given between waypoints and each waypoint contains direction to each of its nodes.

Table 1. Waypoint lookup table.

Waypoint	1	2	3	4	5	6
1		WP1 to WP2	WP1 to WP2 WP2 to WP3	WP1 to WP2 WP2 to WP3 WP3 to WP4	WP1 to WP2 WP2 to WP3 WP3 to WP4 WP4 to WP5	WP1 to WP2 WP2 to WP3 WP3 to WP6
2	WP2-WP1		WP2 to WP3	WP2 to WP3 WP3 to WP4	WP2 to WP3 WP3 to WP4 WP4 to WP5	WP2 to WP3 WP3 to WP6
3	WP3 to WP2 WP2 to WP1	WP3 to WP2		WP3 to WP4	WP3 to WP4 WP4 to WP5	WP3 to WP6
4	WP4 to WP3 WP3 to WP2 WP2 to WP1	WP4 to WP3 WP3 to WP2	WP4 to WP3		WP4 to WP5	
5	WP5 to WP4 WP4 to WP3 WP3 to WP2 WP2 to WP1	WP5 to WP4 WP4 to WP3 WP3 to WP2	WP5 to WP4 WP4 to WP3	WP5 to WP4		WP5 to WP6
6	WP6 to WP3 WP3 to WP2 WP2 to WP1	WP6 to WP3 WP3 to WP2	WP6 to WP3	WP6 to WP3 WP3 to WP4	WP6 to WP3 WP3 to WP4 WP4 to WP5	

Note: This lookup table is used to determine the files required by the phrase synthesizer to assemble and play natural directions between destination nodes.

Table 1 illustrates the structure of the lookup table used to determine directions between waypoints. The variable lists of linkages (part 2) are used in combination with parts 1 and 3 outlined above, to deliver effective directions using the phrase synthesizer. For example, if the user wants to navigate from node 6 to node 5, they first need directions from node 6 to waypoint 6, then directions from waypoint 6 to waypoint 5 and finally directions from waypoint 5 to node 5. Directions between nodes and their waypoints are named accordingly and are loaded automatically. The format of the directions between waypoints determined using the waypoint lookup table is shown below:

[To get to] + [Node 6] + [N6 to WP6] + [Then] + [WP6 to WP3] + [Then] + [WP3 to WP4] + [Then] + [WP4 to WP5] + [and] + [WP5 to N5] + [I hope this is clear, if not request directions again]

The example illustrates how the phrase synthesizer dynamically constructs directions using randomly selected handles combined with a series of directions segments. The result is rendered by the AUI, and animated visually on the map (Figure 10).

3.3.4. Location manager with adaptive resolution monitoring

The *Delivery Engine* is a spatialized AUI, but practical LBS require accurate positioning in order to deliver a satisfactorily immersive and hands-free user experience. The engine is by design fully automated in its response to changes in location, so a *Location Manager* was developed to manage this process (Figure 10).

The *Location Manager* detects both current region and nearest node in real time, with a location resolution monitor used to determine when the engine should deliver content for a given node. In addition, the manager can distinguish between different types of node (e.g. POI, facilities, events), as demonstrated in the Dublin Zoo tour (Figure 10).

In any situation where positioning accuracy is too low to deliver content effectively, an intelligent audio housekeeping manager updates the user in real time by informing them of their positioning conditions and instructs them to switch to manual mode, where all nodes

Figure 10. The delivery engine location manager interface.
Note: The location manager manages the visual representations of natural directions (demonstration using Dublin Zoo tour).
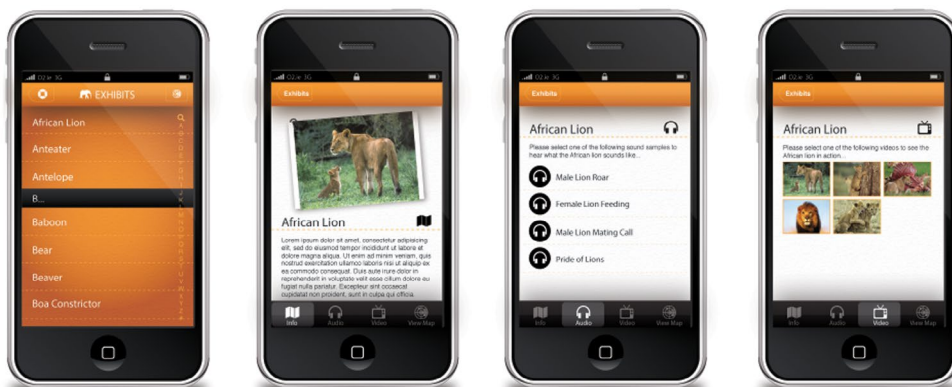

Figure 11. Delivery engine graphical user interface for manual content delivery.
Note: This mode is used when GPS is unavailable or the user requires a visual aide.

can be triggered by manual selection (Figure 11). In this manner, the engine ideally never delivers inaccurate content to the user providing them with a higher degree of usability. As an extension to this, the location manager incorporates live weather updates into its positioning accuracy assessment (accuracy is often linked to cloud cover) that is also reported to the user by the housekeeping manager.

3.3.5. Adaptive virtualized 3D soundscapes

The *Delivery Engine* is capable of delivering immersive soundscaping based on the region context of the user. Transitions between regions are dealt with automatically as fades between atmospheric 3D background sounds, as the user moves between different regions in a space. Fades are timed to function at walking pace so the user perceives a natural spatial transition between one context and another.

Each region has a specifically authored soundscape comprised background synth pad sounds and foreground transient sound effects (SFX), where a transient sounds (such as a bird singing high in a tree behind your right shoulder) are delivered using real-time transient positioning to give the listener a sense of a live environment. This is achieved using the DTS Audio Virtualizer to give the impression of movement within the soundscape, allowing sounds relating to non-stationary objects (e.g. flying birds) to be delivered algorithmically as part of the narrative. The engine automatically adapts to changes in location and context from a top down perspective, so transient SFX never clash with more important foreground sounds such as node content or directions.

The Virtualizer contains a low bitrate version of the codec for potential implementation on constrained devices. The virtualization algorithm takes a head related impulse response for a given azimuth (measured using a binaural dummy head) and convolutes this with the incident signal to deliver an output signal which is perceptually similar to one from that azimuth. In the case of the DTS Virtualizer, azimuths are provided for a common 5.1 multichannel audio setup that defines rear left (rL) and rear right (rR) positions. This Virtualizer was performance tested using the iOS audioGraph render model to determine whether a real-time fast Fourier transform convolution could be performed using the iOS render callback model.

The results (Figure 12) define a cardioid response pattern for rear content when delivered using the Virtualizer. This leads to a practical constraint of positioning where idealized radius r (translating to distance to source) is replaced by $r1$ in the rear condition and $r2$ for frontal cues in the virtualized case. Taking the desired effect into account, these results proved sufficient to produce a satisfactory result.

3.3.6. Algorithmic music sequencer

The *Delivery Engine* contains a full multipart algorithmic music sequencer, where regions are given a specific musical score containing several parts that vary automatically over time. The sequencer swaps content based on compositional rules provided in a similar manner to algorithmic music found in video games to ensure that the user never becomes bored with a continuous loop of music. The *Delivery Engine* sequencer also responds to contextual changes, swapping all music for each region to compliment the underlying soundscape using the same fading strategy. In this manner, the user may move wherever they wish, with SFX and music seamlessly adapting to their current context. Music and background pads are juxtaposed by placing them at the front and rear of the virtualized soundscape, respectively.
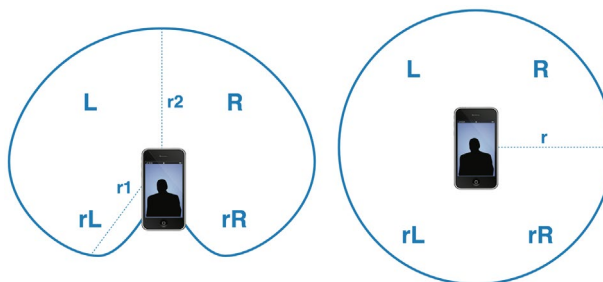

Figure 12. Virtualizer perception test results.
Note: This figure shows the comparison between the actual and ideal cardioid response pattern for delivering rear content using the DTS Virtualizer.

This provides the maximum spatial distance possible between content elements, providing the user with a much wider auditory vista within the interface.

To provide all functionality described in the previous five subsections, a very flexible approach to content delivery is required, where each element of the content can be controlled and managed separately. A multichannel audio mixer (*MixerHost*) was used to achieve this.

3.3.7. Multi-channel auditory mixer

The audio mixer component of the iOS Core Audio Framework was selected as it provides up to 32 independently controllable channels of audio. For our purposes, we used 16 channels of audio for phrase synthesis (content, help and directions), algorithmic music sequencing and 3D soundscaping. This requires a rigid approach to file management based on the linguistic rules and sequence instructions modelled at the authoring stage. This fundamental aspect of the entire process is really only evident at this stage in the process. The breakdown of the assigned channels for each aspect of audio content is illustrated in Figure 13.

The diagram above illustrates the structure of the *MixerBus* and the allocation of the channels for SFX (3D soundscaping), Music (algorithmic music) and Speech (content, directions, region changes and help).

The first four channels are reserved for SFX, which are a combination of ambient synth pads and virtualized transient sound effects. These are loaded at runtime from the *swapSFXStructArray*, which is a buffer array of randomly selected files for the current region. When the user enters a new region, the SFX for the current region are gradually faded-out on the *MixerBus*. The *swapSFXStructArray* is loaded with the new region's synth pads and transient sounds, ready to be loaded in the *MixerBus*. When the fade-out is complete, the new files are loaded into the *MixerBus* and faded back in. The next eight channels are reserved for the
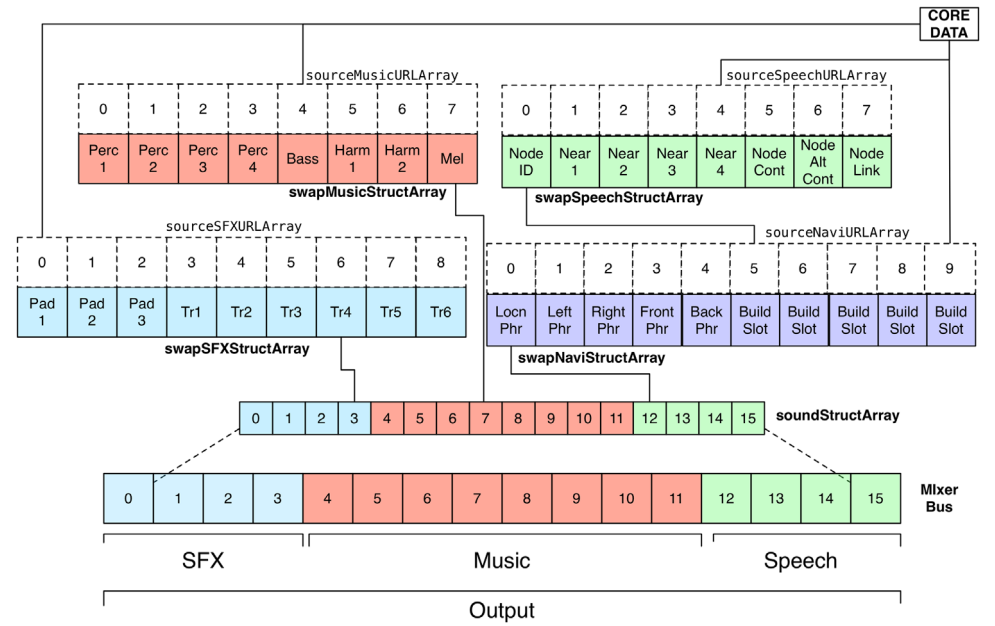

Figure 13. MixerBus logic diagram. The MixerBus provides 32 independently controllable channels of audio, which are used to configure speech, music and SFX.

construction of algorithmic music composed of melodies, rhythms and harmonics randomly selected and organized at runtime (Schoenberg 1967). Similar to the SFX, the *swapMusic-StructArray* is a buffer array that feeds the *MixerBus* with audio files. The eight channels are carefully loaded when a region change occurs (fading in and out) and are algorithmically modified over time so that the music does not seem repetitive.

The last four channels are reserved for phrase synthesis where node, region, help or natural directions content is constructed at runtime. These four channels are the most active part in the overall process and require careful planning. Both the *swapSpeechStructArray* and the *swapNaviStructArray* use these channels for the production of phrases that will inform the user about their current node or region, give them directions to the next node or provide help. These buffers are managed using a prioritization matrix that only plays the audio with the highest priority first and queues the rest. The values that control this are assigned at the authoring stage of the process. One of the most significant features of the *MixerBus* is the ability to manage each channel or set of channels independently, which supports the dynamic loading of audio separately for speech content, music and background SFX.

This section described the main features of the *MobiLAudio* platform, how they function in isolation and how they integrate to deliver an immersive user experience using multimodal interfaces. The main component of the platform is the AUI, which can only operate with the support of the preceding survey and authoring process stages. Considerable contributions to the state-of-the-art in the area of mobile multimodal content delivery have been presented, next they are validated in the form of a user trial.

The following section details the user trial carried out to answer the research questions presented in Section 3, and to assess the effectiveness of the approach taken. Details of the trial are given, followed by the results of functionality and usability tests and concluded by a discussion of the results.

4. User trial

A central goal of this research was to test the usability and user experience of the *MobiLAudio* prototype. This was achieved in the form of a detailed user trial to determine specific user responses to the functionality, usability and content provided by the iOS audio tour developed. The trial focused extensively on user feedback to determine which elements of the proposed functionality are most important to the end user. Specific levels of detail were of interest (e.g. no music or SFX, voice only, etc.) alongside additional functions utilizing a variety of media to try and accurately model the user experience.

The user trial was carried out in central Dublin based around the theme of James Joyce's *Ulysses* and its central character Leopold Bloom. The tour tested the functionality of the survey and authoring tools to determine how effectively multiple tours could be produced. A total of 31 participants took part in the trial, and all participants completed the trial fully (Figure 14).

The trial focused on the execution of several location-based tasks, which were evaluated as part of functionality testing. Quantitative questions relating to the user experience of the application were given alongside post-test NASA Task Load Index (TLX) tests to assess how much of a cognitive load the application places on the user. The testing schedule defined three separate test areas; functionality, user experience and post-test TLX evaluation. The
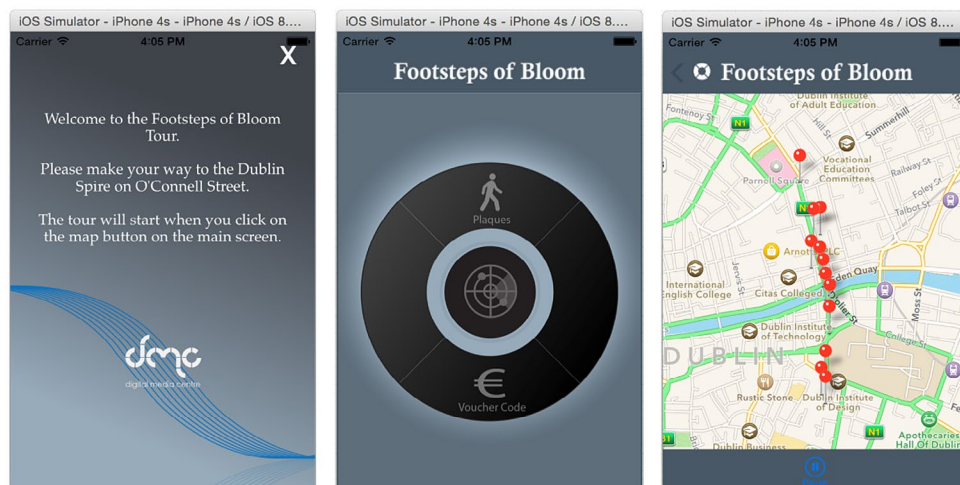
Figure 14. Bloom tour graphical user interface.
Note: This interface was used for the user trial in Dublin City centre for use in the event that manual control was necessary.

results provided both quantitative (functionality and TLX) and qualitative (user experience) data for analysis. This combination provided a comprehensive picture of the performance of the application, within the defined testing space of the *Footsteps of Bloom* audio tour.

Functionality Testing: After using the application, participants were asked a series of questions relating to the performance and functionality of the application. These included questions on the music, SFX, speech and directions provided, alongside the general operation of the application on the device.

User Experience Testing: After using the application, each participant was asked to evaluate the experience by answering a short set of qualitative questions. Questions included ease of use, evaluation of audio tours, and whether the app is worth paying for.

NASA TLX Testing: After the user experience testing was completed, a brief TLX test was carried out. The smallest TLX test contains a series of five scales that can be marked by either the tester or participant relating to metrics such as mental, physical and temporal demand. These metrics allow the participant to express how much cognitive demand the app placed upon them, which is crucial to the analysis of both functionality and user experience results as an increased cognitive load can negatively affect user experience.

4.1. Footsteps of bloom tour

The *Footsteps of Bloom* tour was constructed specifically for the purposes of the user trial and commercial evaluation and validation. The tour comprises 10 points of interest on a linear narrative beginning at the O'Connell Street Spire and ending at Davy Byrne's Pub on Grafton Street. The tour points are based on various sidewalk plaques found in central Dublin that punctuate Leopold Bloom's journey in Ulysses, and provides information based on the novel at each point. The tour takes approximately 35–40 min to complete and utilizes all the major functional elements of the AUI.
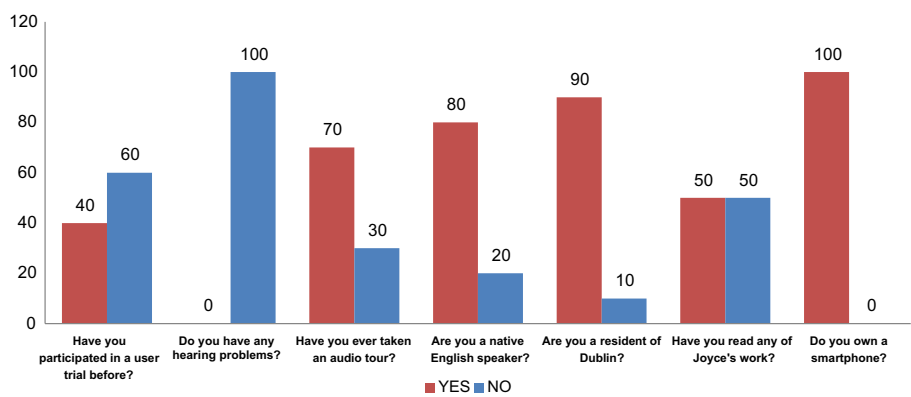
Figure 15. MobiLAudio user trial pre-test question results.
Note: The table shows answers to yes/no questions relating to the user and their prior experience.

4.1.1. Pre-test questions

Each participant was asked several pre-test questions to determine whether they had any hearing problems (which would invalidate their test) and to obtain general information about their language skills, familiarity with the subject matter and their familiarity with the technologies involved. The results of each question are shown in Figure 15.

Of the participants questioned, none had any hearing difficulties that would preclude them from taking part in the trial. A total of 70% had previously taken an audio tour, which was considered useful in providing a benchmark of familiarity for the application. Most participants were native speakers (80%) living in Dublin (90%) though only 50% were familiar with the work of James Joyce. All participants own a smartphone, and as the test was designed to be largely hands free this was considered to be a sufficient level of prior exposure to the technologies involved in the tour.

4.1.2. Functionality questions

Each participant was asked a series of questions relating to the functionality of the application. These questions related to the SFX, music, speech and directions modelling provided as part of the *Footsteps of Bloom* tour. Questions were also included to determine the general operation of the application, notably its performance as a hands-free audio tour (Figure 16).

All participants could distinguish between the music, SFX and speech within the application and considered the balance between the three to be correct. Similarly, they considered the application to have functioned automatically and as required.

4.1.3. Directions modelling

Only 65% of participants felt the application had directed them correctly, though in many cases this was due to physical occlusion of several Leopold Bloom sidewalk plaques (e.g. due to outside tables at Davy Byrne's Pub, or telecom repair work on Aston Quay for 2 days of the trial). Of those who were misdirected in another way, two participants, after given the first instruction to proceed down O'Connell Street towards the bridge (from the Dublin Spire), walked instead in the opposite direction towards Parnell Street. In another three instances, participants struggled to find a plaque on Westmoreland Street because the directions to
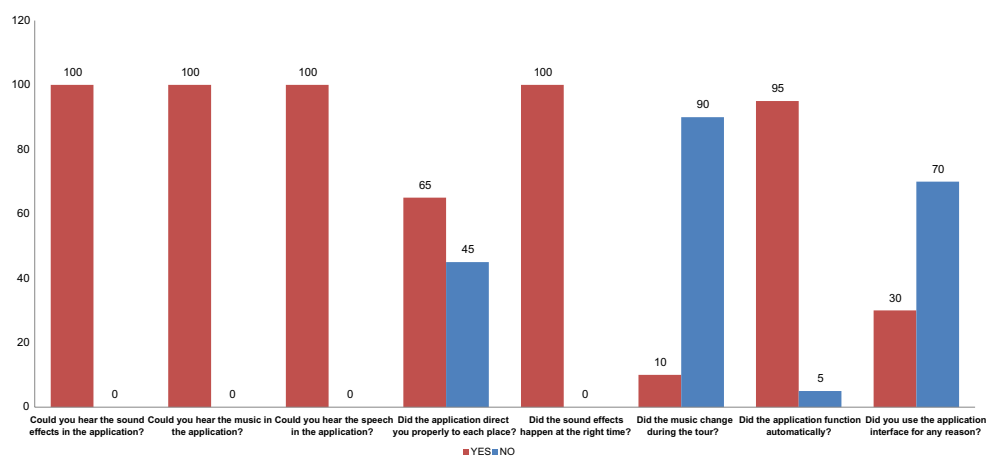
Figure 16. MobiLAudio user trial functionality question results.
Note: The table shows answers to yes/no questions relating to the operation of the application and its component functionality.

Conways Pub were not strictly accurate (as the sidewalk sign is actually outside Charlie's Chinese Takeaway).

The most significant directional issues occurred at the second POI on Middle Abbey Street, as there were several components making up the navigation directions from Joyce's Statue on Earl Street. Many participants reported having difficulty keeping up with the tour whilst also trying to remember the series of steps needed to get to the Easons building. Some participant comments on the application (taken post-test) included:

… directions were a great idea, but difficult in places …

… (I) had only one chance to listen to the directions …

Staged directions would have been a lot more useful …

… Street names are no good for tourists … landmarks are better …

… a wrong turning sound would have been helpful around Easons … otherwise it was grand

This feedback suggests that the directions component, though considered very useful by participants, did not function to the level of accuracy and detail that they would have liked. Future work will have to consider directions modelling not just in terms of node/hub networks (as performed in *Delivery Engine*) but also in terms of the complexity of the route. In addition, the use of street names in tandem with visual landmarks proved not to provide any useful additional information, and indeed was considered slightly confusing in some cases.

4.1.4. AUI performance

All participants indicated that they could hear sound effects (SFX), music and speech correctly, and in particular the SFX were considered to be a really good addition to the experience. All users considered the timing of the SFX to be correct (100%) and again this was backed up by qualitative comments indicating that SFX were a significant positive element of the tour.

The determination of all three audio elements (SFX, music, speech) validates the design of the listening mode model employed by the application, but there was an issue indicated

with repetition of music during the tour. Only 10% of participants felt the music changed or varied during the tour, with many commenting that the music was either too much or too distracting:

> … voice could be a bit louder relative to the music …
>
> I didn't like the music, it didn't fit in with Joyce …

This indicates that the use of immersive background music may not be as important or necessary as was initially thought, with listeners indicating a preference for a music 'mute' button in the application. Indeed, the use of the application interface (participants were told the application was automatic) by 70% of participants was mainly to change the volume between POIs, with only two participants stating they had used the application's map out of personal interest in the tour. Interestingly, the expectation of traffic noise as a distraction to users did not arise. This may have been due to the level of the background music or the overall immersion experienced by users.

4.1.5. User experience questions

Each participant was also asked a series of qualitative questions relating to the user experience of the *Footsteps of Bloom* tour. These questions focused on the engagement of the application, with questions on user tours and social media integration being included alongside the opportunity to make more general comments (Figure 17).

All participants found the application easy to use and most found it engaging (95%) and would be interested in other similar tours (95%). Comments on the general concept and content of the tour were positive, with many participants providing suggestions for other tours:

> … easy to use, liked the directions between points …
>
> … liked the voice and the sounds, the barge going past sounded real!
>
> … storytelling was great, would have liked more content though …
>
> … could you have running tours … or cycling tours …
>
> An Easter Rising tour would be brilliant with the bombs going off around you!

One of the more significant comments made here aside from the high level of engagement experienced was that of the changes to the sound stage. Users reported on a noticeable change in the spatial image due to the use of sound effects (barges passing by and trams in the distance), with some reporting that they sounded 'very real'.

4.1.6. Tour pricing

Ninety per cent of participants said they would pay to use a mobile tour application, and gave an average price value of €2.95 for the amount they would be willing to pay. This figure is interesting, as 90% of the participants were Dublin residents (with 80% being native English speakers) who would have had some previous exposure to elements of the tour, its concept and content. Additionally, though only 50% said they had read Joyce's work, this does not seem to preclude interest in audio tours in general with 75% indicating they enjoyed the *Footsteps of Bloom* tour.

4.1.7. Social networking integration

One of the most interesting aspects of the results was the high values for social networking integration with tours (85%), alongside the possibility of user tour creation (90%). In more
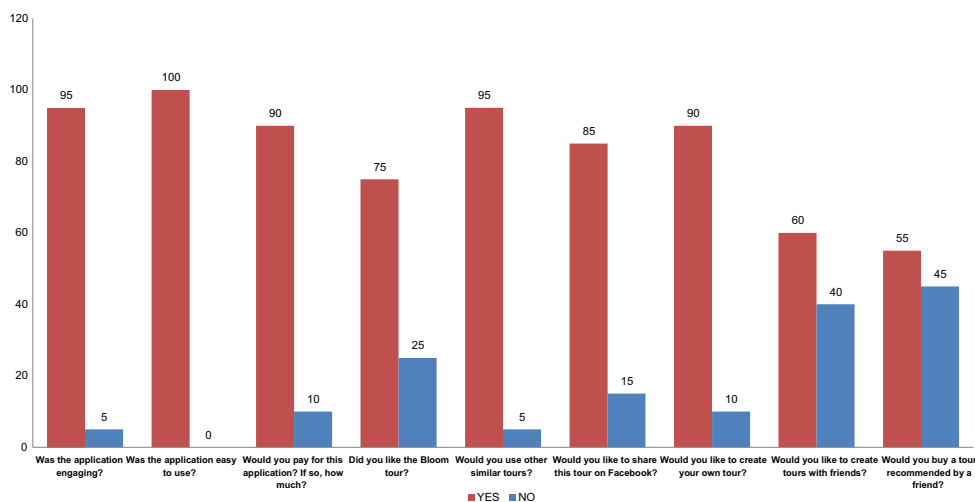
Figure 17. MobiLAudio user trial user experience question results.
Note: The table shows answers to yes/no questions relating to the engagement of the application, its potential for self-generated tours and social media integration opportunities.

detail, 60% considered the possibility of creating tours in a group to be interesting, though the social marketing potential of tour recommendations did not provide a definitive answer (55%). Comments about user tour creation included:

> … I'd love to build a tour for visiting family, it would give them something to do while I'm at work during the day …

> … I'd build a tour for my boyfriend, linking all the places we've been since we met …

> … when friends come to visit it would be great to give them a pub tour …

> Is there a web interface for building the tours? I'd love to put one together …

> … my own voice would be nice for family, they would like the personal touch …

4.1.8. Post test TLX results

After completing the tour and answering all test questions, a set of post-test NASA TLX tests were carried out to determine the load on the participant. NASA TLX tests were originally developed to evaluate astronaut performance and behaviour under demanding or stressful conditions. Under normal circumstances, the tests are useful in assessing the mental and physical load that a certain operation or task places on a participant and are a useful means of comparing how participants actually performed relative to how well they *thought* they performed (Figure 18).

The overall results were low (as desired), with physical demand and frustration having values of 22 and 20.5%, respectively, suggesting that the user experience results for engagement and ease of use (100%) were a fair reflection of the performance of the application. The important value of time pressure was also very low at 16.5%, suggesting that the application did not impinge upon the tour experience in its function or delivery. Mental demand was higher at 41% and this is potentially explained by the requirement for participants to listen to directions as they navigate the tour (no participant had any prior idea of the direction the
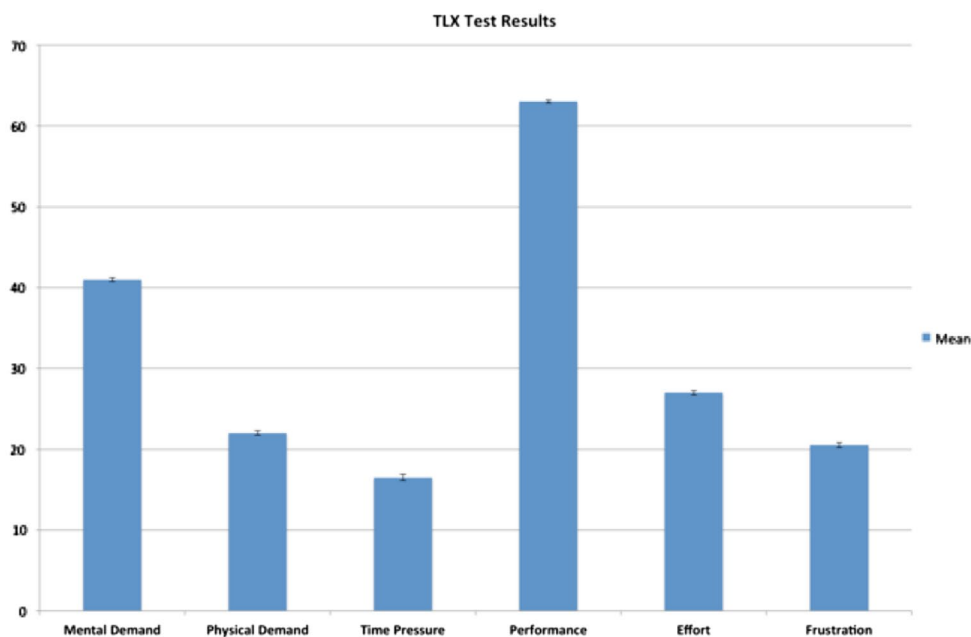
Figure 18. MobiLAudio user trial post-test TLX results.
Notes: The table shows the mean and standard deviation (in per cent) for each of the six TLX scales. TLX scales cover mental and physical demand, time pressure, performance, effort and frustration.

tour would take). The effort score of 27%, though still very low, can perhaps be attributed to the combination of walking and listening to a narrative at the same time. It is also interesting to consider the performance value of 63%, which may indicate in some respects difficulties with the directions modelling provided. In several instances, participants indicated that the directions were too long or confusing and expressed concern that they had not 'done well enough' even though they themselves were not directly being evaluated. It is also significant that percentile deviations were under 0.4% for all scales, indicating that the TLX values obtained were representative of the overall participant population.

4.1.9. Discussion

The overall reaction to the *Footsteps of Bloom* tour was very positive, and the technical elements of the application were all evaluated to function properly (though directional modelling did cause problems in its current delivery format). In particular, SFX was found to be a very positive addition to the experience of the application, though background music proved to be distracting for some users. The automated nature of the application was also considered a significant feature, with most people only using the graphical interface out of interest rather than necessity. The narrative and conversational flow of the tour was very well received, and the vast majority of participants would be interested in other audio tours. The concept of user-generated tours was strongly endorsed by participants, with integration into social networking being another positive element. Overall, participants reacted very positively to the tour and validated both the design and technical and directions taken in *MobiLAudio*.

The results of the Post Test TLX survey were mainly in line with expectations with significantly low scores for frustration, time pressure, physical demand and effort. The most interesting part of these results is the high value of 63% for performance. This could be attributed to the use of a combination of map-based visual directions and egocentric-based audio directions, which have been shown to cause confusion when users switch between frames of reference. Another reason could be that the audio directions were too long as some participants did express concern about the length and complexity of the directions and that they had not 'done well enough'.

5. Conclusions and future work

The *MobiLAudio* platform demonstrates significant advances in the area of multimodal content delivery. In addition to a traditional GUI, an AUI is used to deliver rich, interactive content using audio as the primary modality. The AUI comprises novel innovations in the areas of natural directions (providing directions using landmarks as opposed to cardinal directions) and content delivery using an intelligent content phrase synthesizer, controlled by a users location, orientation and previous visits to the space. In addition, an immersive audio experience is delivered using region detection and an algorithmic music sequencer. The result is an audio tour that provides timely information about an environment, 3D soundscaping that delivers immersive background audio and transient sounds based on location and natural directions between nodes on the tour. All of the content in the app is assembled on the fly with music, sound effects and speech audio being constructed dynamically using the audio sequencer and phrase synthesizer. This approach has proven to significantly reduce the amount of data redundancy in the system, enabling a far greater degree of efficiency overall.

Taking the results of the user trial into account, future work will focus on refining the delivery of directions by considering directions modelling not just in terms of node/hub networks but also in terms of the complexity of the route. In addition, the use of street names in tandem with visual landmarks proved not to provide any useful additional information, and was considered slightly confusing in some cases. Indeed, staged natural directions based on a user's current location are a more obvious choice in an urban environment.

Further testing will be carried out in order to fine-tune the variances in the listening model to determine a more optimum balance of audio in the AUI. In addition, an adaptive volume control based on external noise levels will be considered to alleviate the issues identified in the user trial around volume levels and algorithmic music.

Although the usability testing identified some interesting results regarding user engagement and level of immersion, this was somewhat inconclusive. Future usability testing will use the standardized system usability scale to determine the usability score for applications.

Extending the use of the authoring tool to include user-created tours will also be investigated with a view to testing the idea of social media integration enabling users to create, share and interact with users of a tour.

Disclosure statement

References

Abowd, G. D., C. G., Atkeson, J. Hong, S. Long, R. Kooper, and M. Pinkerton. 1997. "Cyberguide: A Mobile Context-aware Tour Guide. *Wireless Networks* 3 (5): 421–433.

Absar, R., and C. Guastavino. 2008. "Usability of Non-speech Sounds in User Interfaces." *Proceedings of the 14th International Conference on Auditory Display (ICAD2008)*, June 24-27, 2008, Paris: International Community for Auditory Display.

Apple Developer Library. 2014. "iOS – Core Audio Essentials." Accessed November 16. https://developer.apple.com/library/ios/documentation/MusicAudio/Conceptual/CoreAudioOverview/CoreAudioEssentials/CoreAudioEssentials.html

Bartie, Phil J., and William A. Mackaness. 2006. "Development of a Speech-based Augmented Reality System to Support Exploration of Cityscape." *Transactions in GIS* 10 (1): 63–86. doi:10.1111/j.1467-9671.2006.00244.x.

Baus, Jörg, Rainer, Wasinger, Ilhan, Aslan, Antonio, Krüger, Andreas, Maier and Tim Schwartz. 2007. "Auditory Perceptible Landmarks in Mobile Navigation." In *Proceedings of the 12th International Conference on Intelligent User Interfaces (IUI '07)*, 302–304. New York: ACM. doi:10.1145/1216295.1216352

Beauchamp, J. W., ed. 2007. *Analysis, Synthesis, and Perception of Musical Sounds*. New York: Springer-Verlag .

Bederson, B.B. 1995. "Audio augmented reality: a prototype automated tour guide." In *Conference Companion on Human Factors in Computing Systems (CHI '95)*, edited by Katz, I., R. Mack, and L., Marks, 210–211. New York: ACM. 10.1145/223355.223526

Bellotti, F., C. Berta, A. de Gloria, M. Margarone. 2002. "User Testing a Hypermedia Tour Guide." *IEEE Pervasive Computing* 1 (2): 33–41.

Benedikt, M. L. 1979. "To Take Hold of Space: Isovists and Isovist Fields." *Environment and Planning B: Planning and Design* 6 (1): 47–65.

Burnett, G. E. 2000. "'Turn Right at the Traffic Lights': The Requirement for Landmarks in Vehicle Navigation Systems." *Journal of Navigation* 53 (3): 499–510.

Carswell, J. D., A. Eustace, K. Gardiner, E. Kilfeather, and M. Neumann. 2002. "An Environment for Mobile Context-based Hypermedia Retrieval." *13th International Conference on Database and Expert Systems Applications (DEXA)*, IEEE CS Press, Aix en Provence, 532–536.

Carswell, J., K. Gardiner, and M. Neumann. 2003. "Wireless Spatio-Semantic Transactions on Multimedia Datasets." 19th Annual ACM Symposium on Applied Computing, Nicosia, Cyprus.

Carswell, J. D., K. Gardiner, A. Rizzini, M. Bertolotto, N. Mandrak. 2008. "A Web-based and Mobile Environmental Management System." *Journal of Environmental Informatics* 12 (1): 9–20.

Carswell, J. D., K. Gardiner, and J. Yin. 2010a. "Mobile Visibility Querying for LBS." *Transactions in GIS* 14 (6): 791–809.

Carswell, J. D., K. Gardiner, and J. Yin. 2010b. "3DQ: Threat Dome Visibility Querying on Mobile Devices." *GIM International* 24 (8): 24–26.

Carswell, J. D., K. Gardiner, A. Rizzini, M. Bertolotto, and N. Mandrak. 2010. "A Web and Mobile System for Environmental Decision Support." In *Decision Support Systems in Agriculture, Food and the Environment: Trends, Applications and Advances*, edited by Basil Manos, Konstantinos Paparrizos, Nikolaos Matsatsinis, and Jason Papathanasiou, 317–338. Hershey, PA: IGI Global.

Chan, L.-W., Y.-Y. Hsu, Y.-P. Hung, and J. Y.-J. Hsu. 2005. "Orientation-Aware Handhelds for Panorama-Based Museum Guiding System." UbiComp 2005 Workshop: Smart Environments and their Applications to Cultural Heritage, September 11–14 , Tokyo, Japan.

Chittaro, L. 2010. "Distinctive Aspects of Mobile Interaction and Their Implications For The Design of Multimodal Interfaces." *Journal on Multimodal User Interfaces* 3 (3): 157–165.

Chittaro, L., and S. Burigat. 2005. "Augmenting Audio Messages with Visual Directions in Mobile Guides: An Evaluation of Three Approaches." In *Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices & Services*, 107–114. New York: ACM.

Darwin, C. J. 2008. "Listening to Speech in the Presence of Other Sounds." *Philosophical Transactions of the Royal Society B: Biological Sciences* 363 (1493): 1011–1021. doi:10.1098/rstb.2007.2156.

Donker, Hilko, Palle Klante, and Peter Gorny. 2002. "The Design of Auditory User Interfaces for Blind Users." In *Proceedings of the Second Nordic Conference on Human-Computer Interaction (NordiCHI '02)*, 149–156. New York, NY: ACM.

Droumeva, Milena. 2005. "Understanding Immersive Audio: A Historical and Socio-Cultural Exploration of Auditory Displays." Eleventh Meeting of the International Conference on Auditory Display (ICAD'05), July 6-9, Limerick, Ireland.

Eckel, G. 2001. "Immersive Audio-Augmented Environments: The LISTEN Project." In *Proceedings Fifth International Conference on Information Visualisation* 571–573. IEEE. doi:10.1109/IV.2001.942112.

Edwards, Alistair. 1989. "Soundtrack: An Auditory Interface for Blind Users." *Human-Computer Interaction* 4: 45–66. doi:10.1207/s15327051hci0401_2.

Fröhlich, P., R. Simon, L. Baillieand H. Anegg. 2006. "Comparing Conceptual Designs for Mobile Access to Geo-Spatial Information." 8th International Conference on Human- Computer Interaction with Mobile Devices and Services (MobileHCI 06), Helsinki, Finland.

Fröhlich, P., R. Simon, and L. Baillie. 2009. "Mobile Spatial Interaction." *Personal Ubiquitous Computing* 13 (4, May 2009): 251–253. doi:10.1007/s00779-008-0208-1.

Gardiner, K. 2004. "A Framework for Processing Viewer-based Directional Queries in a Simulated Mobile Environment." Masters diss., Dublin Institute of Technology.

Gardiner, K., and J. D. Carswell. 2003. "Viewer-Based Directional Querying for Mobile Applications." In *Proceedings of the Fourth International Conference on Web Information Systems Engineering Workshops*, edited by Giuseppe Santucci, Wolfgang Klas, Michela Bertolotto, Coral Calero and Luciano Baresi, 83–91. Washington, DC: IEEE Computer Society.

Gardiner, K., A. Rizzini, J. D. Carswell, and M. Bertolotto. 2005. "MEMS – Mobile Environmental Management System" 13th Annual Conference on GIS Research UK, University of Glasgow, Scotland.

Gardiner, K., J. Yin, and J. D. Carswell. 2009. "EgoViz –A Mobile Based Spatial Interaction System." In *9th International Symposium on Web and Wireless Geographical Information Systems (W2GIS)*, LNCS 5886, edited by James D. Carswell, A. Stewart Fotheringham, and Gavin McArdle, 135–152. Berlin Heidelberg: Springer. doi:10.1007/978-3-642-10601-9_10

Gardiner, K., C. Cullen and J. D. Carswell. 2015. "MobiSurround: An Auditory User Interface for Geo Service Delivery." In *26th International Conference on Database and Expert Systems Applications (DEXA)*, Part II, LNCS 9262, edited by Q. Chen, A. Hameurlain, F. Toumani, R. Wagner, and H. Decker, 57–72. Cham: Springer International Publishing.

Gaver, William. 1989. "The SonicFinder: An Interface That Uses Auditory Icons." *Human-Computer Interaction* 4: 67–94. doi:10.1207/s15327051hci0401_3.

Heller, Florian, Thomas Knott, Malte Weiss, and Jan Borchers. 2009. "Multi-user Interaction in Virtual Audio Spaces." In *Extended Abstracts on Human Factors in Computing Systems (CHI EA '09)*, 4489–4494. New York, NY: ACM.

Hile, H., R. Vedantham, G. Cuellar, A. Liu, N. Gelfand, R. Grzeszczuk, and G Borriello. 2008. "Landmark-based Pedestrian Navigation from Collections of Geotagged Photos." In *Proceedings of the 7th International Conference on Mobile and Ubiquitous Multimedia (MUM '08)*, 145–152. New York, NY: ACM.

Holland, Simon, David R. Morse, and Henrik Gedenryd. 2002. "AudioGPS: Spatial Audio Navigation with a Minimal Attention Interface." *Personal and Ubiquitous Computing* 6 (4): 253–259. doi: 10.1007/s007790200025.

Jagdish, Deepak, Rahul Sawhney, Mohit Gupta, and Shreyas Nangia. 2008. "Sonic Grid: An Auditory Interface for the Visually Impaired to Navigate GUI-based Environments." In *Proceedings of the 13th International Conference on Intelligent User Interfaces (IUI '08)*, 337–340. New York, NY: ACM.

Kainulainen, A., M. Turunen, J. Hakulinen, and A Melto. 2007. "Soundmarks in Spoken Route Guidance." Proceedings of the 13th International Conference of Auditory Display (ICAD 2007), June 26-29, Montreal, Canada.

Kilfeather, E., J. Carswell, K. Gardiner, and S. Rooney. 2007. "Urban Location Based Services Using Mobile Clients: The ICiNG Approach." GISRUK, Maynooth, Ireland.

Klippel, A., H. Tappe, L. Kulik, and P. U. Lee. 2005. "Wayfinding Choremes—A Language for Modeling Conceptual Route Knowledge." *Journal of Visual Languages & Computing* 16 (4): 311–329.

Kramer, Gregory. 1993. *Auditory Display: Sonification, Audification, and Auditory Interfaces*. New York: Perseus Publishing.

Lalanne, Denis, Laurence Nigay, Philippe Palanque, Peter Robinson, Jean Vanderdonckt, and Jean-Francois Ladry. 2009. "Fusion Engines for Multimodal Input: A Survey." In *Proceedings of the 2009 International Conference on Multimodal Interfaces (ICMI-MLMI '09)*, 153–160. New York, NY: ACM.

Magnusson, Charlotte, Konrad Tollmar, Stephen Brewster, Tapani Sarjakoski, Tiina Sarjakoski, and Samuel Roselier. 2009. "Exploring Future Challenges for Haptic, Audio and Visual Interfaces for Mobile Maps and Location Based Services." In *Proceedings of the 2nd International Workshop on Location and the Web (LOCWEB '09)*, Article 8. New York, NY: ACM. doi: 10.1145/1507136.1507144.

McGee, J., and C. Cullen. 2009a. "Vocate: Auditory Interfaces for the LOK8 Project." Proceedings of the 9th IT & T Conference, October 22-23, Dublin Institute of Technology, Dublin.

McGee, J., and C. Cullen. 2009b. "Vocate: Auditory Interfaces for Location-Based Services." Physicality Workshop at 23rd Conference on Computer Human Interaction (HCI 2009), September 23, Cambridge, UK.

McGookin, D. 2004. "Understanding and Improving the Identification of Concurrently Presented Earcons." PhD diss., University of Glasgow, 155–159.

Millonig, A., and K. Schechtner. 2005. "Developing Landmark-based Pedestrian Navigation Systems." *IEEE Transactions on Intelligent Transportation Systems* 8(1): 43–49. doi:10.1109/TITS.2006.889439

Misra, Ananya, R. Perry Cook, and Ge Wang. 2006. "A New Paradigm for Sound Design." Proceedings of the International Conference on Digital Audio Effects (DAFx-06), September 18-20, Montreal, Quebec, Canada.

Modsching, M., R. Kramer, M. Riebeck, A. Stark, K. ten Hagen, and J. Kawalek. 2007. "Field Trial on the Efficiency and User Experience of GPS Based State of the Art Navigational Systems for Pedestrians." In *4th Workshop on Positioning, Navigation and Communication (WPNC '07)*, 129–134. IEEE. doi: 10.1109/WPNC.2007.353623

Mynatt, Elizabeth D. 1990. "Auditory Presentation of Graphical User Interfaces." *Auditory Display: Sonification, Audification and Auditory Interfaces*, edited by G. Kramer, 533–556. Reading, MA: Addison-Wesley.

Neustein, A., and J. Markowitz, eds. 2013. *Mobile Speech and Advanced Natural Language Solutions*. New York, NY: Springer. 10.1007/978-1-4614-6018-3

Persson, P., F. Espinoza, P. Fagerberg, A. Sandin, and R. Cöster. 2002. "GeoNotes: A Location-based Information System for Public Spaces." In *Designing Information Spaces: The Social Navigation Approach, Part I*, edited by Kristina Höök, David Benyon, and Alan J. Munro, 151–173. London: Springer.

Potamianos, A., and M. Perakakis. 2008. "Multimodal Processing and Interaction." *Multimedia Systems and Applications Series* 33: 1–18. doi:10.1007/978-0-387-76316-3.

Rehrl, K., E. Häusler, and S. Leitinger. 2010. "Comparing the Effectiveness of GPS-Enhanced Voice Guidance for Pedestrians with Metric-and Landmark-Based Instruction Sets." In *Geographic Information Science*, 189–203. Berlin Heidelberg: Springer.

Rehrl, K., E. Häusler, S. Leitinger, D. Bell. 2014. "Pedestrian Navigation with Augmented Reality, Voice and Digital Map: Final Results from an *in Situ* Field Study Assessing Performance and User Experience." *Journal of Location Based Services* 8 (2): 75–96.

Rizzini, A., K. Gardiner, M. Bertolotto, and J. D. Carswell. 2006. "MEMS Mobile GIS: A Spatially Enabled Fish Habitat Management System." 6th International Symposium on Web and Wireless Geographical Information Systems (W2GIS); Springer LNCS vol.4295; Hong Kong, China.

Robinson, S., P. Eslambolchilar, and M. Jones. 2008. "Point-to-GeoBlog: Gestures and Sensors to Support User Generated Content Creation." Proceedings of the 10th Conference on Human-Computer Interaction with Mobile Devices and Services, Mobile HCI 2008, September 2-5, Amsterdam, The Netherlands. doi: 10.1145/1409240.1409262

Rooney, S., K. Gardiner, and J. D. Carswell. 2007. "An Open Source Approach to Wireless Positioning Techniques." In Proceedings of the 5th International Symposium on Mobile Mapping Technology (MMT), Padua, Italy.

Schoenberg, A. 1967. *Fundamentals of Music Composition*. Boston, MA: Faber and Faber.

Schwarz, D. 2004. "Data-driven Concatenative Sound Synthesis." PhD Thesis, Ircam Centre Pompidou, Paris, France.

Schwarz, D. 2005. "Current Research in Concatenative Sound Synthesis." Proceedings of the International Computer Music Conference (ICMC), September 5-9, Barcelona, Spain.

Shi, X., T. Sun, Y. Shen, K. Li, and W Qu. 2010. "Tour-guide: Providing Location-Based Tourist Information on Mobile Phones." In *Computer and Information Technology (CIT), 2010 IEEE 10th International Conference on*, 2397–2401. IEEE.

Shinohara, Kristen, and Josh Tenenberg. 2009. "A Blind Person's Interactions with Technology." *Communications of the ACM* 52(8): 58–66. doi:10.1145/1536616.1536636

Shneiderman, S. B., and C. Plaisant. 2005. *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. 4th ed. Boston: Pearson Addison Wesley.

Simon, R., and P. Fröhlich. 2007. "A Mobile Application Framework for the Geospatial Web." Proceedings of the 16th international conference on World Wide Web, Banff, Alberta, Canada.

Simon, R., and P. Fröhlich. 2007. "The Point to Discover GeoWand." 9th International Conference on Ubiquitous Computing (UbiComp 07), Innsbruck, Austria.

Simon, R., P. Fröhlich, and T. Grechenig. 2008. "GeoPointing: Evaluating the Performance of Orientation-Aware Location-Based Interaction under Real-World Conditions." *Journal of Location Based Services* 2 (1): 24–40.

Stahl, Christoph. 2007. "The Roaring Navigator: A Group Guide for the Zoo with Shared Auditory Landmark Display." In *Proceedings of the 9th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '07)*, 383–386. New York, NY: ACM.

Stark, A., M. Riebeck, and J. Kawalek. 2007. "How to Design an Advanced Pedestrian Navigation System: Field Trial Results." In *4th IEEE Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS 2007)*, 690–694. IEEE. doi:10.1109/IDAACS.2007.4488511

Strachan, S., R. Murray-Smith. 2009. "Bearing-based Selection in Mobile Spatial Interaction." *Personal and Ubiquitous Computing* 13 (4): 265–280.

Strachan, Steven, Parisa Eslambolchilar, Roderick Murray-Smith, Stephen Hughes, and Sile O'Modhrain. 2005. "GpsTunes: Controlling Navigation via Audio Feedback." Proceedings of the 7th international conference on Human computer interaction with mobile devices & services, Salzburg, Austria.

Strachan, S., J. Williamson, and R. Murray-Smith. 2007. "Show Me the Way to Monte Carlo: Density-based Trajectory Navigation." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*, 1245–1248. New York, NY: ACM. doi:10.1145/1240624.1240812

Truax, B. 2001. *Acoustic Communication*. 2nd ed. Westport, CT: Ablex Publishing.

Turk, Matthew. 2014. "Multimodal Interaction: A Review." *Pattern Recognition Letters* 36: 189–195. doi:10.1016/j.patrec.2013.07.003./.

Vazquez-Alvarez, Yolanda, Ian Oakley, Stephen A. Brewster. 2012. "Auditory Display Design for Exploration in Mobile Audio-Augmented Reality." *Personal and Ubiquitous Computing* 16: 987–999. doi:10.1007/s00779-011-0459-0.

Wakkary, R, K Newby, M Hatala, D Evernden, and M Droumeva. 2004. "Interactive Audio Content : An Approach to Audio Content for a Dynamic Museum Experience through Augmented Audio Reality and Adaptive Information Retrieval." In *Museums and the Web 2004 Selected Papers*, edited by D. Bearman and J. Trant, 55–60. Arlington, VA: Archives & Museum Informatics.

Walker, Bruce N., and Jeffrey Lindsay. 2006. "Navigation Performance with a Virtual Auditory Display: Effects of Beacon Sound, Capture Radius, and Practice." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 48: 265–278. doi:10.1518/001872006777724507.

Wang, Lei, Paul Roe, Binh Pham, and Dian Tjondronegoro. 2008. "An Audio Wiki Supporting Mobile Collaboration." In *Proceedings of the 2008 ACM Symposium on Applied Computing (SAC '08)*, 1889–1896. New York, NY: ACM. doi:10.1145/1363686.1364145. http://doi.acm.org/10.1145/1363686.1364145

Wenzel, Elizabeth M. 1994. "Spatial Sound and Sonification." In *Auditory Display: Sonification, Audification, and Auditory Interfaces*, edited by G. Kramer, 127–150. Reading, MA: Addison Wesley.

Williamson, J., R. Murray-Smith, and S. Hughes. 2007. "Shoogle: Excitatory Multimodal Interaction on Mobile Devices." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 121–124. New York, NY: ACM.

Wilson, A., and S. Shafer. 2003. "XWand: UI for Intelligent Spaces." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '03.*, 545–552. New York, NY: ACM.

Yin, Junjun, and J. D. Carswell. 2013. "Spatial Search Techniques for Mobile 3D Queries in Sensor Web Environments." *ISPRS International Journal of Geo-Information* 2 (1): 135–154.

5.2. Paper II: *mobiSurround*: An Auditory User Interface for Geo-Service Delivery

Gardiner, K., C. Cullen and J. D. Carswell. (2015) "mobiSurround: An Auditory User Interface for Geo Service Delivery", 26th International Conference on Database and Expert Systems Applications (DEXA) Valencia, Spain. (Accepted - In Press)

mobiSurround: An Auditory User Interface for Geo-Service Delivery

Keith Gardiner, Charlie Cullen and James D. Carswell

Digital Media Centre, Dublin Institute of Technology, Ireland
(keith.gardiner, charlie.cullen, jcarswell)@dit.ie

Abstract. This paper describes original research carried out in the area of Location-Based Services (LBS) with an emphasis on Auditory User Interfaces (AUI) for content delivery. Previous work in this area has focused on accurately determining spatial interactions and informing the user mainly by means of the visual modality. *mobiSurround* is new research that builds upon these principles with a focus on multimodal content delivery and navigation and in particular the development of an AUI. This AUI enables the delivery of rich media content and natural directions using audio. This novel approach provides a hands free method for navigating a space while experiencing rich media content dynamically constructed using techniques such as phrase synthesis, algorithmic music and 3D soundscaping. This paper outlines the innovative ideas employed in the design and development of the AUI that provides an overall immersive user experience.

Keywords: auditory user interface, phrase synthesis, content delivery, geo-services, navigation.

1   Introduction

In this paper the delivery of high quality media content and navigational information via multiple modalities is investigated. The idea of presenting geo-referenced information both visually and aurally to the user in a collective way is the primary focus. The visual modality is presented using a graphical user interface (GUI) in a state-of-the-art mobile app that incorporates and builds on many of the ideas developed in previous works in the area of mobile spatial interaction [1], [2], [3], [4], [5], [6], [7], [8]. The aural modality is presented using an auditory user interface (AUI), which is the main innovation in terms of delivering context sensitive information using focus independent (eyes free) means. The AUI is intended to be a non-visual interface that can be used in combination with a GUI, or not. The idea is that all information in the app, including content, directions, events, etc. can be delivered without the user ever needing to interact with the mobile device [9].

There are two main parts to this research. First, the use of an AUI for the delivery of context sensitive geo-tagged content is investigated and secondly, the contextual data modelling process used to structure the data for effective delivery is described. In previous work [10], [11], [12], [13], several difficulties were encountered because of

the large quantities of content required to achieve this task and as a direct result, the size of the content was an issue. In many cases, the level of redundancy in these systems is very high. In this research, because the content used by the system is primarily audio, the approach taken aims to prevent this level of redundancy by introducing the concept of phrase synthesis to location-based services (LBS). Phrase synthesis is the process by which a dictionary of recorded words is used to create phrases or sentences "on-the-fly" based on linguistic rules. Using this technique we attempt to deliver content and navigational instructions via the AUI resulting in significant reductions in size and redundancy, thus reducing the overheads of the app. This approach is then extended to produce soundscaping where multiple audio channels are sequenced to provide ambient background audio and sound effects that contribute to an overall immersive experience.

In Section 2, previous work in the areas of location-based services and mobile spatial interaction is outlined. Section 3 describes the *mobiSurround* AUI and the innovations that provide natural directions for real-time navigation, intelligent phrase synthesis, algorithmic music sequencing and adaptive virtualized 3D soundscapes. In Section 4, conclusions and proposed further work is described.

2 Previous Work

Research into location-based information retrieval and delivery has been on going for more that 4 decades now [14], [15], [16]. Since the first methods of creating digital maps and cityscapes, points-of-interest (POI) locations and geo-tags, the idea of providing information based on a users location and supported by an underlying geographic map have become increasingly common and sophisticated. It is this work that forms the foundation for research in this area and is the basis for our contributions in the areas of location-based information retrieval and mobile spatial interaction.

2.1 Location-Based Services

With advances in mobile computing technology (e.g. spatially enabled smartphones), implementation of effective location-based services has become a reality in many ways. Initial research in [13] describes work that simulated LBS in a 3D environment using a users virtual position and orientation for the discovery of cultural heritage artefacts. Further research focussed on demonstrating the concept of LBS in real-world environments [17], [18] identifying several fields that required further investigation. One of these areas was visibility and the exact nature of human information retrieval based on the senses, i.e. how we retrieve and process information presented to us; essentially looking at the different modalities for content consumption. The visual modality was studied and in particular the accurate representation of a user's line-of-sight was investigated. In [12], [19], the horizontal field-of-view was considered as the search space, representing what the user can see. This required significant accuracy in terms of positioning and orientation, which was not yet available on mobile devices. Subsequent research in [20], demonstrated the use of location and orien-

tation for mobile spatial search. The point-to-discover service used a customised mobile device to perform spatial selection queries of cultural artefacts demonstrating mobile spatial interaction.

2.2 Mobile Spatial Interaction

Mobile spatial interaction in this context is the study of the interaction between a mobile user and a world of spatial information, i.e. both the physical built environment and all it's connected sensors and related attributes. In [5], [7], [8], [21], [22], the effect direction has on query results presented to users on mobile devices was studied and the idea of an egocentric view of the world was investigated in a mobile context. The applications developed as part of this research operated on contemporary COTS (commercial-off-the-shelf) mobile devices and thus were capable of being used in real-world situations. The results of this research proved successful in addressing the problem of information overload, where too much information gets returned to the mobile device causing display clutter and confusion.

Extending this research further, in [23], [24], the idea of creating a more accurate search space is investigated. The horizontal position and orientation of the user is used to determine the interaction between a ground map of buildings and a user's position and orientation; resulting in a more realistic representation of the user's viewport. In addition, an Isovist [25] was used to create a 360-degree visibility window. In [26] the idea of 3D directional queries was introduced and resulted in the development of the *3DQ* query processor which enabled a user to point a mobile device at a building, for example, to obtain information about the building or more specifically at a particular floor.

Following considerable work investigating novel applications for location-based services and mobile spatial interaction, the need for an improved approach to data modelling and organisation [16] for the delivery of high quality content was apparent [3], [6]. This led to the present approach described in Section 3 of an efficient media content delivery engine that can deliver geo-referenced media in an immersive way, using multiple modalities while reducing data redundancy [21], [27]. This new work produced a spatialised auditory user interface (AUI) for LBS with contributions in the areas of location/data modelling and adaptive content production and delivery [16].

3   *mobiSurround*

The *mobiSurround* engine is an auditory user interface (AUI) for location-based services (LBS) that models context and location to provide a number of exciting new innovations in the areas of content delivery, mobile navigation, and narrative context. More specifically the engine provides the following features:

- An intelligent content phrase synthesiser
- A natural directions model for real time navigation
- An algorithmic music sequencer

- Adaptive virtualised 3D soundscapes
- A full demonstration tour application based on Dublin Zoo

The central component of the *mobiSurround* engine is based on a 3-stage data modelling process that informs the various parts of the application that feed both the visual and aural presentation modalities. The geodata-modeling (GM) task is a process by which the geographic data that defines POIs, navigational waypoints and regions are collected. This data informs the rest of system and is the basis of the content modelling strategy. The user modeling (UM) task is where user preferences are modelled, essentially defining the type of user and their context. This data has a scaling effect on the content modelling strategy, where varying amounts of content is presented based on a user's profile. The content modeling (CM) task is where the structure of the media content is defined for each node based on the geo-data (GM) and application preferences data (UM). It is this strictly controlled process of data modelling from multiple perspectives that ensures that data in the system is structured correctly, enabling the dynamic loading and delivery of media content in a coherent manner (Figure 1).
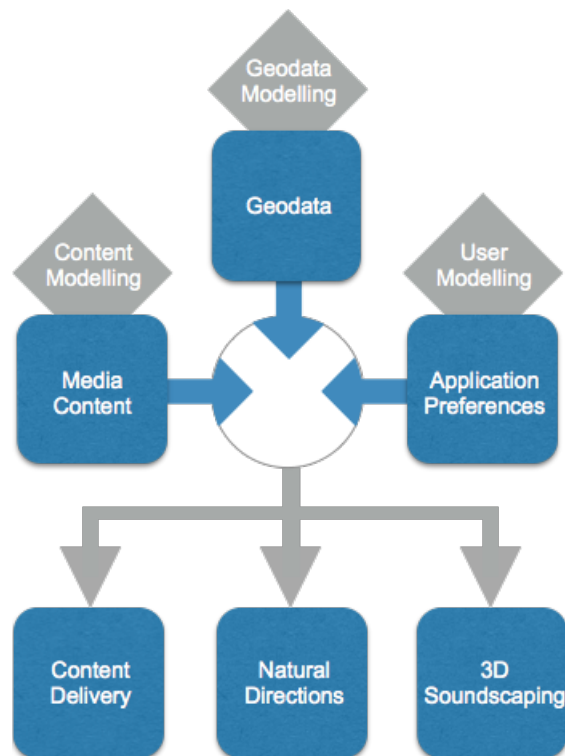

Fig. 1. 3-stage data modeling process includes user, context and content modeling for effective data management and content organization

The following sections outline our innovative approach to modeling for geo-services and describe how the 3-stage process integrates to deliver the functionality described. In particular, aspects relating to content modeling and delivery are outlined that describe; how a space is mapped, how content is modelled around the geo-data, and finally how the content is delivered based on spatial interactions.

3.1 Geodata Mapping and Delivery Matrix

To provide these services a model for the structure and narrative of the content to be delivered was developed. The *mobiSurround* concept focuses on the idea of adaptive narrative that may be used in (though not limited to) tours and exhibitions, where the user may experience a space in their own way, in any direction and in any order. This has been achieved in *mobiSurround* with the development of a node/waypoint/region model for the mapping of data in a given environment (Figure 2).
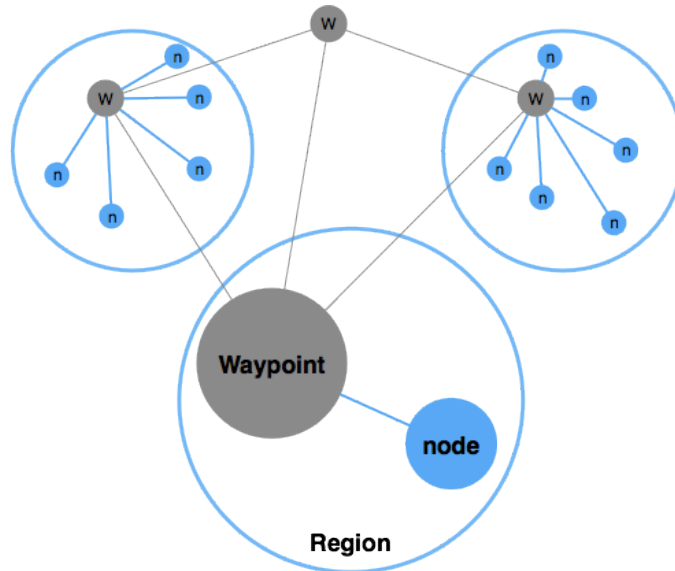
Fig. 2. Data mapping structure describing node, waypoint and region nodes for navigation

Nodes.
A node is considered a basic unit of data in the *mobiSurround* engine and is used to represent a POI, service, or facility. Each node has a nodeID and contains information about its 4 neighbouring nodes, its closest waypoint and the region it belongs to. Using the nodeID, the engine determines what content to load on arrival, whether or not to change region and what the next node is based on the users path.

Each node has a number of associated media assets that includes images, text and audio. The phrase synthesizer dynamically loads segments of recorded speech, each containing separate facts about the POI. The phrase synthesizer is used in all node

content delivery to allow a relatively small set of commentary phrases (delivered by the narrator) to be reused for a given node based on a set of linguistic rules. In this manner, the content for each node is built at runtime to give the user the impression of listening to a dynamic documentary type interview between narrator and expert. This is illustrated in Figures 4 and 5.

Waypoints.

Waypoints are the main navigational unit in the *mobiSurround* engine. They can be considered as reference points for all other nodes in a space, where each node essentially belongs to or is governed by a waypoint such that navigation around a space is achieved by specifying directions between waypoints. Once navigation between waypoints has been achieved, directions from the destination waypoint to all of its associated nodes are accessible. Specific directions are then provided to a destination node. Significantly, waypoints are defined as natural landmarks that exist in the area making them perfectly suitable for delivering natural directions. For example: "*walk down as far as the Big Tree and take a left; the shop is on your right hand side*". In this instance the Big Tree is the destination waypoint and the shop is the destination node.

Regions.

Regions are defined in this context as functional areas that divide a space, for example, Dublin Zoo is divided into four characteristic regions representing different "corners" of the world. Each region has a set of boundaries that give it a unique geospatial identity. Regions can encompass multiple waypoints and nodes essentially helping to define the context in which they exist. Navigating between regions triggers a region change, which in effect signifies a change in context. The result is a managed "outro" transition of all contextual audio and sound effects for that region followed by an introduction to the new region and a managed intro transition of all contextual audio for that region. This subtle effect is a key feature of the adaptive virtualised 3D soundscaping described in the following section.

3.2   Content Modelling

With a delivery matrix in place, the translation of the required functionality into a content model for authoring is undertaken for each of the 3 elements in the matrix (nodes, waypoints and regions). These 3 elements are then used to define and structure the content for the following areas:

- **Content Delivery**- the information relating to a specific node in the space
- **Navigation**- the information needed to traverse between points in the space using waypoints
- **Narrative Context**- changes between defined regions allow the narrative to adapt to the context of the user

Listening Model. The *mobiSurround* content delivery model is based on a background to foreground listening paradigm [28], [29]. Aligning these elements based on the listening paradigm prioritises speech over background audio and music. This ensures that content delivery and directions are always heard and coherent. This is illustrated in Figure 3.
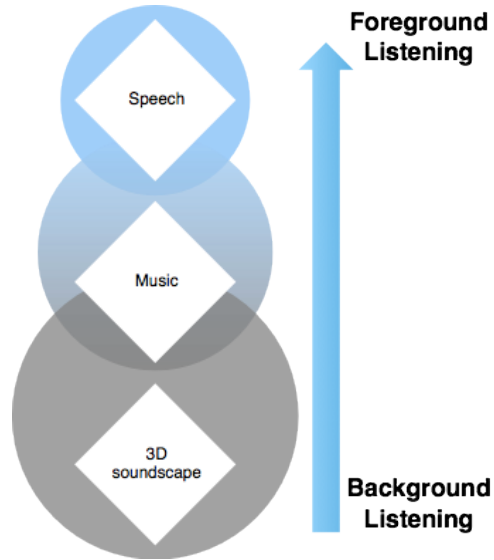

Fig. 3. *mobiSurround* background to foreground listening model that prioritises speech over background audio and music

The listening model illustrates how content is delivered to the user. A 3D soundscape is produced to recreate environmental aspects of the current region. The 3D soundscape is primarily composed of ambient synth pad sounds and virtualised transient sound effects (SFX). Together, they produce a background harmony and atmosphere as a result of very long attack and decay times with extended sustains [30]. This is layered with algorithmic music composed from melodies, rhythms, and harmonics that are randomly selected and organised at runtime and that the user can detect but not prioritize directly. Speech content (node information or directions) is then layered and will always be foreground content. This layered approach of using 3D Soundscaping, algorithmic music and prioritised speech helps convey broader conceptual variations within the narrative of the space.

The following sections describe the key components and software processes that enable the *mobiSurround* engine to deliver the innovative features described.

Intelligent Content Phrase Synthesizer. One of the major contributions of the *mobiSurround* engine is the intelligent phrase synthesizer implemented to deliver speech content. The synthesizer was developed to provide multimodal content delivery on constrained devices that preclude large amounts of audio content being included in a

wireless network download. The phrase synthesizer is used for all node content delivery to allow a relatively small set of commentary phrases (delivered by the narrator) to be reused for a given node based on a set of linguistic rules. In this manner, node content is dynamically assembled at runtime to give the user the impression of listening to a documentary interview between narrator and expert. A multi-channel audio graph (AUGraph) is used to provides 16 individually controlled busses, much like a traditional mixing desk [31]. This is illustrated in Figure 4.
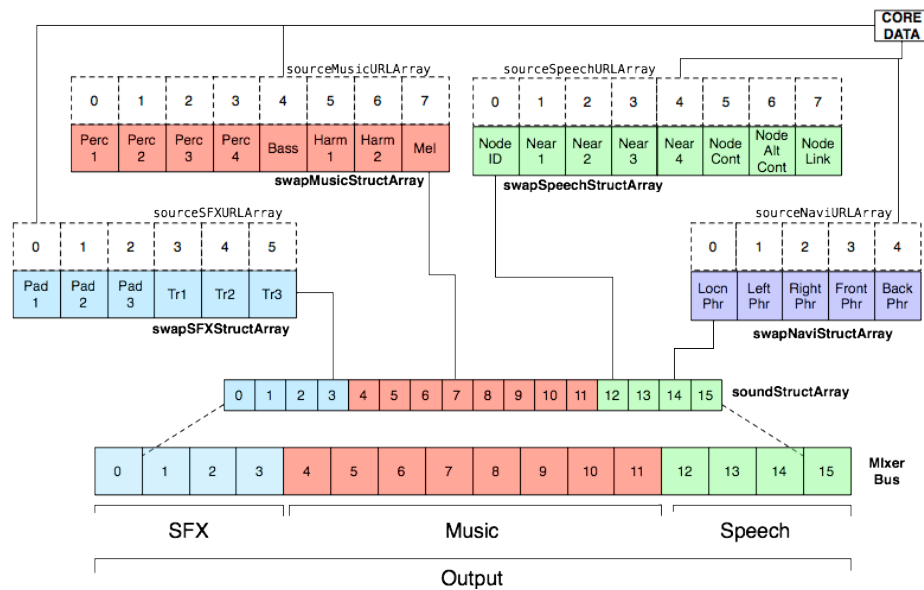
Fig. 4. Mixerbus multichannel audio graph controls all audio playback in the application based on a set of rules

The phrase synthesiser has the exclusive use of 4 busses in the AUGraph with each bus capable of managing 2 sets of files at once. Full control of each bus is maintained as data is swapped in and out in real-time. This is all managed dynamically by a set of rules determined during the data-modelling phase. The geodata and content for each node is combined with a narration template defining the overall structure of a node. When the context changes, the graph is dynamically controlled by a set of decision rules that are stored on Core Data. These rules are described in Table 1.

Figure 5 describes how the phrase synthesizer constructs node, region, help and direction content sentences in real time using busses 1-4. Each bus (Intro, Content, altContent, Outro) is dynamically loaded with audio using a handle and a content phrase. The handle is randomly selected from a predefined library and the content relates to the nodeID of the currently selected node. An example of a synthesized sentence is as follows:

[You are now arriving at] + [the Snow Leopard exhibit] + [Lets ask john to give us some information] + [Snow Leopards are large cats that live way up high in the mountains ranges of central Asia] + [Wow that's interesting] + [FACT2] + [Cool can you give us another fact] + [FACT3] + [Interesting, what else can you tell us about it] + [FACT4] + [Well now you have heard about it lets move on to] + [the red river hog exhibit] + [DIRECTIONS]
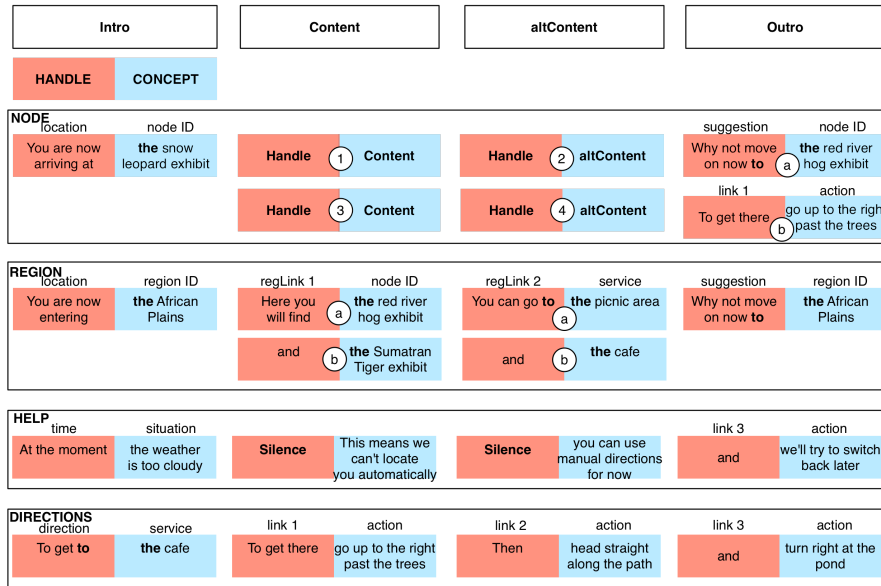

Fig. 5. Phrase synthesis structure for node, region and directions content

Each node has two narrators associated with it, the main narrator and a domain expert. Facts 1 and 2 are initially loaded into the *Content* and *altContent* busses, and then they are swapped out for Facts 3 and 4. The phrases are constructed in real-time and loaded into the AUGraph where they are organised sequentially, initialised with the required bus control data from the database and played in order. The result is a seamless set of sentences played back in real-time describing the current node (POI). When the node information is finished, all 4 busses are faded out and reset awaiting the discovery and subsequent delivery of new node information. All node, region information, help and directions content are constructed in this manner, significantly reducing the overhead of recording and playing large static files.

The conditional logic that controls the mixerbus is described in the decision table below. Each time conditions change; the graph control data is loaded from core data and applied.

Table 1. Decision table describing the rules, conditions and actions that control loading, management and playback of the mixerbus graph

		Rules									
Conditions	New POI Detected	Y		Y			Y		Y		Y
	New Region Detected		Y	Y			Y		Y		Y
	New Event Occuring				Y						
	Directions Request					Y	Y				
	Random SFX Change							Y	Y		
	Random Music Change									Y	Y
Actions	Prepare speechURLArray	X		X	X		X		X		X
	Prepare musicURLArray		X	X						X	X	X
	Prepare SFXURLArray		X	X				X	X		X
	Prepare naviURLArray		X	X		X	X		X		X
	FadeOut Channel 12-15	X			X	X					
	FadeOut Channel 0-3										
	FadeOut Channel 4-11										
	FadeOut MixerBus (0-16)		X	X			X		X		X
	FadeOut Channel Index							X		X	
	swapSpeechStructArray	X		X	X				X		X
	swapMusicStructArray		X	X					X		X
	swapSFXStructArray		X	X					X		X
	swapNaviStructArray		X	X	X	X	X		X		X
	swapRandomSFXStructElement							X			
	swapRandomMusicStructElement									X	
	FadeIn Channel 12-15	X			X	X	X				
	FadeIn Channel 0-3										
	FadeIn Channel 4-11										
	FadeIn MixerBus (0-16)		X	X					X		X
	FadeIn Channel Index							X		X	

Natural Directions Model for Real-Time Navigation. Natural directions are a recent innovation in navigation technology. Instead of providing measurements and compass headings, landmarks are used to give directions. The advantages of natural directions are speed (humans do not need to translate data) and adaptability (there is no need to hold a compass in a certain way to obtain a heading) that allow them to be delivered hands free. The disadvantage of natural directions is the size of the content model required, where each direction is a unique narrative in itself.

To address this, the *mobiSurround* natural directions model uses a combination of node and waypoint information to control the phrase synthesizer, providing a set of natural directions (which can also be displayed on a map if required) to any given point in the space. The model uses the concept of waypoint linkage to create a lookup table of directions between waypoints that is then completed by directions from the last waypoint to the node required. In this manner, natural points of focus within the space can be defined as waypoints that are used to script and create the directions needed to link them together. When a specific node is queried, the only information required is its nearest waypoint, which in turn allows a lookup table to be used to configure the phrase synthesizer.

Figure 6 demonstrates how waypoint linkages can be used to provide directions to any node in the space. For the 3 example nodes shown, a common set of waypoint linkage directions is used to get to waypoint 3, where different waypoint combinations then lead to nodes 3, 4 and 5. By building a set of waypoint linkages in both directions (getting from W2 to W1 is not the same as going from W1 to W2) each node is linked to its nearest waypoint, which is then linked by a lookup table to every other waypoint. The phrase synthesizer is then loaded with the entries from the lookup table and provides the directions to the user as required.
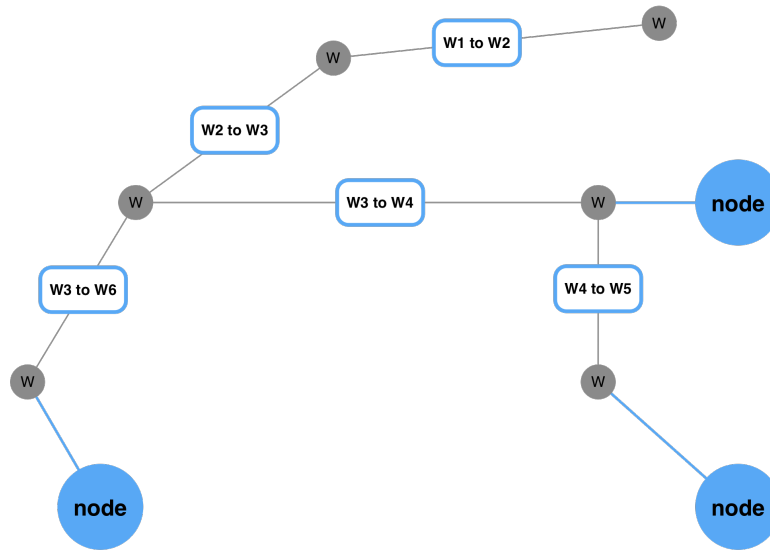
Fig. 6. Waypoint linkage example for node and waypoint navigation

Location Manager with Adaptive Resolution Monitoring. The *mobiSurround* engine considers spatialised auditory user interfaces, but the concept of LBS requires suitably accurate positioning in order to deliver an immersive and hands free user experience. The engine is designed to be fully automated in its response to changes in location so a Location Manager was developed (Figure 7).

The manager defines both current region and nearest node in real-time, with a location resolution or proximity monitor being used to determine when the engine can deliver content for a given node. In addition, the manager can determine different types of node (e.g. POI, facilities, events) at once, as demonstrated in the Dublin Zoo demo tour (Figure 7).

Accurate positioning (particularly in urban areas) is still an issue with COTS devices, with changes in weather conditions and urban canyons due to the presence of buildings and other large objects preventing a GPS receiver from delivering an accurate reading in real-time. Although *mobiSurround* does not directly seek to innovate in positioning accuracy, the location manager employs an adaptive resolution monitor

to specify reasonable bounds for positioning accuracy. This ensures that content delivered by the engine is always appropriate, and also that future advances in positioning (such as EGNOS or NFC) to provide sub-meter accuracy are compatible with the current engine.

Fig. 7. *mobiSurround* Location Manager (demonstration using Dublin Zoo tour).

In any condition where positioning accuracy is too low to deliver content effectively, an intelligent audio housekeeping manager updates the user in real-time by informing them of the condition and asking them to switch to manual mode, where all nodes are listed in the app for triggering by manual selection (Figure 8). In this manner, the engine will ideally not deliver the wrong content to the user and so provide them with a far higher degree of usability and flexibility. As an additional development to this, the location manager has incorporated live weather updates into its accuracy assessment (GPS accuracy is often linked to cloud cover) that can be broadcast to the user by the housekeeping manager.
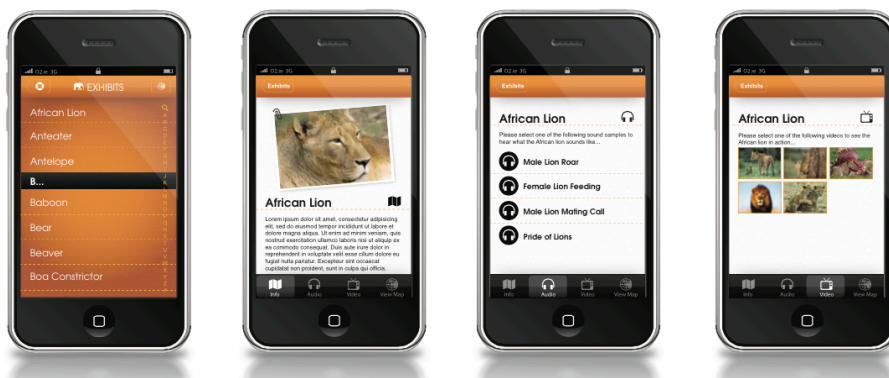
Fig. 8. mobiSurround GUI used in manual mode enables access to POI, facility and event information when accurate positioning data is unavailable

Adaptive Virtualized 3D Soundscapes. The *mobiSurround* engine is capable of delivering immersive soundscaping based on the region context of the user. Transitions between regions are dealt with automatically as fades between atmospheric 3D background sounds as the user moves between different regions in a space. Fades are timed to function at walking pace (though this can be varied programmatically) so the user perceives a natural spatial transition between one context and another.

Each region has a specifically authored soundscape comprising of background synth pad sounds and foreground transient SFX, where a transient (such as a bird singing) is delivered using real-time transient positioning. This gives a listener the sense of being in a live environment, giving the impression of hearing a lion roar just over your right shoulder, for example. This effect is achieved using a Digital Theatre Systems (DTS) virtualiser to give the impression of movement within the soundscape, allowing sounds relating to non-stationary objects to be delivered algorithmically as part of the narrative. The engine automatically adapts to changes in location and context from a top down perspective, so transient sound effects (SFX) never clash with more important foreground sounds such as node content or directions.

The virtualiser provided by DTS Audio contains a low bitrate version of the code for potential implementation on constrained devices. A virtualization algorithm takes a head related impulse response (HRIR) for a given azimuth (measured using a binaural dummy head) and convolutes this with the incident signal to deliver an output signal which is perceptually similar to one as heard coming from that azimuth. In the case of the DTS virtualiser, azimuths are provided for a common 5.1 multichannel audio setup that defines rear left (rL) and rear right (rR) positions. This virtualiser was performance tested using the iOS audioGraph render model to determine whether a real-time Fast Fourier Transform (FFT) convolution could be performed using the iOS render callback model.
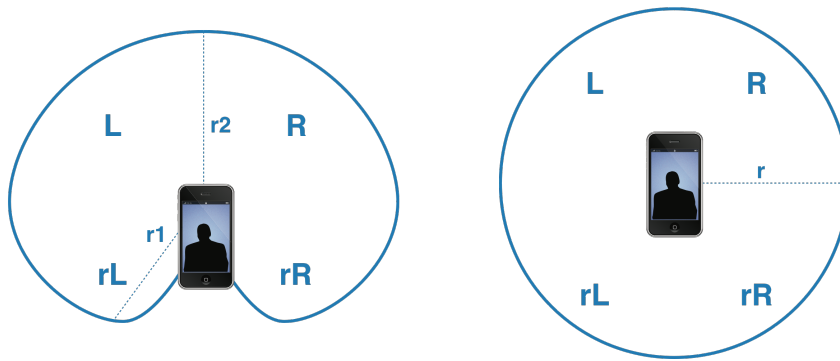

Fig. 9. Visual Summary of Virtualiser Perception Test Results (Comparing Ideal and Actual)

The results (Figure 9) define a cardiod response pattern for rear content when delivered using the virtualizer. This leads to a practical constraint of locationing where idealised radius r (translating to distance) is replaced by r1 in the rear condition and r2

for frontal cues in the virtualised case. Taking the desired effect into account, these results are satisfactory and produce a realistic directional listener experience.

Algorithmic Music Sequencer. The *mobiSurround* engine contains a full multipart algorithmic music sequencer, where regions are given a specific musical score containing several parts that vary automatically over time. The sequencer swaps content based on certain compositional rules provided to ensure that the user never becomes bored by a continuous loop of music in a similar manner to algorithmic music found in video games. In *mobiSurround* the sequencer also responds to contextual changes, swapping all music for each region to compliment the underlying soundscape using the same fading strategy (Table 1). In this manner, the user may move wherever they wish, with SFX and music seamlessly adapting to their current context. Music and background synth pads are juxtaposed by placing them at the front and rear of the virtualised soundscape respectively. This provides the maximum spatial distance possible between content elements, providing the user with a much wider auditory vista within the interface.

4 Conclusions and Future Work

The *mobiSurround* prototype demonstrates significant advances in many areas of spatialised multi-modal content delivery. In addition to a traditional graphical user interface (map), an auditory user interface (AUI) is used to deliver rich, interactive content using audio as the primary modality. The AUI comprises novel work in the areas of natural directions (providing directions using landmarks as opposed to cardinal directions) and content delivery using an intelligent content phrase synthesiser controlled by user location, orientation, and previous visits to the space. In addition, an immersive audio experience is delivered using region detection and an algorithmic music sequencer. The result is an audio tour that provides timely information about the current environment (i.e. Dublin Zoo), 3D soundscaping that delivers immersive background audio and transient sounds based on location and natural directions between nodes on the tour. All of the media content in the app is assembled on-the-fly, with music, sound effects and speech audio being constructed dynamically using the music sequencer and phrase synthesizer. This approach has proven to significantly reduce the amount of redundancy in the system enabling a far greater degree of efficiency overall.

Future work (as a result of testing) will also consider how best to further reduce the application footprint. This will include the manipulation of non-interleaved audio files (all virtualised content is stereo) to better function with the iPhone auGraph. In addition a comprehensive user trial is planned to test and provide feedback in terms of functionality and usability of the AUI for delivering content. Development of an authoring tool for planning, structuring, and combining content is also underway. This tool will help streamline the process of audio production as outlined in the data-mapping matrix and listening model.

5   References

1. Fröhlich, P., R. Simon, L. Baillie, and H. Anegg, *Comparing conceptual designs for mobile access to geo-spatial information.* Proceedings of the 8th conference on Human-computer interaction with mobile devices and services - MobileHCI, 2006: p. 109.

2. Persson, P., F. Espinoza, V. Fagerberg, A. Sandin, and R. Cöster, *GeoNotes: A Location-based Information System for Public Spaces.* Designing Information Spaces: The Social Navigation Approach Computer Supported Cooperative Work, 2003. **pp 151-173**.

3. Robinson, S., P. Eslambolchilar, and M. Jones. *Point-to-GeoBlog: gestures and sensors to support user generated content creation.* in *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services.* 2008. Amsterdam, The Netherlands: ACM.

4. Simon, R. and P. Frohlich, *A mobile application framework for the geospatial web*, in *Proceedings of the 16th international conference on World Wide Web.* 2007, ACM: Banff, Alberta, Canada. p. 381-390.

5. Simon, R., P. Fröhlich, G. Gerhard Obernberger, and E. Wittowetz. *The Point to Discover GeoWand.* in *In Proceedings of the 9th International Conference on Ubiquitous Computing.* 2007. Innsbruck, Austria.

6. Simon, R., P. Frohlich, and T. Grechenig, *GeoPointing: evaluating the performance of orientation-aware location-based interaction under real-world conditions.* J. Locat. Based Serv., 2008. **2**(1): p. 24-40.

7. Frohlich, P., R. Simon, and L. Baillie, *Mobile Spatial Interaction.* Personal and Ubiquitous Computing, 2009. **13**(4): p. 251-253.

8. Strachan, S. and R. Murray-Smith, *Bearing-based selection in mobile spatial interaction.* Personal and Ubiquitous Computing, 2009. **13**(4): p. 265-280.

9. McGee, J. and C. Cullen. *Vocate: Auditory Interfaces for Location-based Services.* in *Proceedings of the 23rd Conference on Computer Human Interaction.* 2009. Cambridge, United Kingdom.

10. Carswell, J.D., K. Gardiner, M. Bertolotto, A. Rizzini, and N. Mandrak, *A web-based and mobile environmental management system.* Journal of Environmental Informatics, 2008. **12**: p. 9-20.

11. Carswell, J.D., K. Gardiner, and M. Neumann, *Wireless spatio-semantic transactions on multimedia datasets*, in *Proceedings of the 2004 ACM symposium on Applied computing.* 2004, ACM: Nicosia, Cyprus. p. 1201-1205.

12. Gardiner, K. and J.D. Carswell. *Viewer-based directional querying for mobile applications.* in *Third International Workshop on Web and Wireless Geographical Information Systems W2GIS.* 2003. Springer.

13. Carswell, J.D., A. Eustace, K. Gardiner, E. Kilfeather, and M. Neumann, *An environment for mobile context-based hypermedia retrieval.* Proceedings. 13th International Workshop on Database and Expert Systems Applications, 2002.

14. Salton, G. and M.J. McGill, *Introduction to modern information retrieval.* McGraw-Hill computer science series. 1983, New York: McGraw-Hill. xv, 448 p.

15. Schiller, J. and A. Voisard, *Location Based Services.* 2004: Morgan Kaufmann Publishers Inc.

16. Jiang, B. and X. Yao, *Location-based services and GIS in perspective.* Computers, Environment and Urban Systems, 2006. **30**: p. 712-725.

17. Mohapatra, D. and S.B. Suma, *Survey of location based wireless services.* IEEE International Conference on Personal Wireless Communications (ICPWC). 2005.

18. Rao, B. and L. Minakakis, *Evolution of mobile location-based services.* Communications of the ACM - Mobile computing opportunities and challenges, 2003. **46**(12): p. 61-65.

19. Chan, L.-W., Y.-Y. Hsu, Y.-P. Hung, and J.Y.-J. Hsu, *Orientation-Aware Handhelds for Panorama-Based Museum Guiding System*, in *UbiComp 2005 Workshop: Smart Environments and their Applications to Cultural Heritage*. 2005.

20. Simon, R., H. Kunczier, and H. Anegg, *Towards Orientation-Aware Location Based Mobile Services.* Information Systems Journal, 2007: p. 1-8.

21. Wilson, A. and S. Shafer, *XWand: UI for intelligent spaces*, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2003, ACM: Ft. Lauderdale, Florida, USA. p. 545-552.

22. Strachan, S., J. Williamson, and R. Murray-Smith, *Show me the way to Monte Carlo: density-based trajectory navigation*, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2007, ACM: San Jose, California, USA. p. 1245-1248.

23. Gardiner, K., J. Yin, and J.D. Carswell. *EgoViz - A mobile based spatial interaction system*. in *The 9th International Symposium on Web and Wireless Geographical Information Systems (W2GIS)*. 2009. Maynooth, Ireland: Lecture Notes in Computer Science.

24. Carswell, J.D., K. Gardiner, and J. Yin, *Mobile Visibility Querying for LBS.* Transactions in GIS, 2010. **14**: p. 791-809.

25. Benedikt, M.L., *To take hold of space: isovists and isovist fields*, in *Environment and Planning B: Planning and Design*. 1979. p. 47-65.

26. Carswell, J.D., K. Gardiner, and J. Yin *3DQ: Threat Dome Visibility Querying on Mobile Devices*. GIM International, 2010. **Vol.24**.

27. Williamson, J., R. Murray-Smith, and S. Hughes, *Shoogle: excitatory multimodal interaction on mobile devices.* Conference on Human Factors in Computing Systems, 2007: p. 121.

28. Truax, B., *Soundscape, Acoustic Communication and Environmental Sound Composition*, in *Contemporary Music Review*. 1996. p. 49-65.

29. Misra, A., P.R. Cook, and G. Wang, *A New Paradigm for Sound Design*, in *Proceedings of the International Conference on Digital Audio Effects*. 2006. p. 1-6.

30. Schoenberg, A., *Fundamentals of Music Composition*. 1967, London: Faber and Faber.

31. Apple. *iOS Developer Library (Core Audio)*. Available from: https://developer.apple.com/library/ios/documentation/MusicAudio/Conceptual/ CoreAudioOverview/CoreAudioEssentials/CoreAudioEssentials.html.

5.3. Paper III: Mobile Visibility Querying for LBS

Carswell, J. D., K. Gardiner and J. Yin. (2010) "Mobile Visibility Querying for LBS", Transactions in GIS, 14(6): 791–809, DOI: 10.1111/j.1467- 9671.2010. 01230.x.

Research Article

Mobile Visibility Querying for LBS

James D. Carswell
Digital Media Centre
Dublin Institute of Technology

Keith Gardiner
Digital Media Centre
Dublin Institute of Technology

Junjun Yin
Digital Media Centre
Dublin Institute of Technology

Abstract

This article describes research carried out in the area of mobile spatial interaction (MSI) and the development of a 3D mobile version of a 2D web-based directional query processor. The TellMe application integrates location (from GPS, GSM, WiFi) and orientation (from magnetometer/accelerometer) sensor technologies into an enhanced spatial query processing module capable of exploiting a mobile device's position and orientation for querying real-world spatial datasets. This article outlines our technique for combining these technologies and the architecture needed to deploy them on a sensor enabled smartphone (i.e. Nokia Navigator 6210). With all these sensor technologies now available on off-the-shelf devices, it is possible to employ a mobile query system that can work effectively in any environment using location and orientation as primary parameters for directional queries. Novel approaches for determining a user's visible query space in three dimensions based on their line-of-sight (ego-visibility) are investigated to provide for "hidden query removal" functionality. This article presents demonstrable results of a mobile application that is location, direction, and orientation aware, and that retrieves database objects and attributes (e.g. buildings, points-of-interest, etc.) by simply pointing, or "looking", at them with a mobile phone.

1 Introduction

This article focuses on directional querying and the development of algorithms for mobile 2D and 3D spatial data interaction. In conjunction with location, an orientation module provides angular parameters to a spatial query processor, making it possible to

Address for correspondence: James D. Carswell, Digital Media Centre, Dublin Institute of Technology, Aungier St., Dublin, Ireland. E-mail: jcarswell@dit.ie

perform directional queries on a mobile device in a real world environment. The fundamental requirements for this type of service interaction are location, direction and orientation.

Some typical applications from current literature enable a user to point a custom-built mobile device at a building and, using position and direction, determine the building's address/identity (Simon and Fröhlich 2007a). This requires both accurate location and orientation as query parameters and the sw components that gather and provide this input data in our case are the LocateMe and DirectMe modules. The locator component, or LocateMe module, is by design an open source, network independent mobile location determination application that can utilize GPS, Wi-Fi and GSM beacon information, or any combination of them, to trilaterate for cellphone position estimates (Kilfeather et al. 2007, Rooney et al. 2007a).

The orientation component, or DirectMe module, is the primary focus of this article. It uses data from a number of sensors including a GPS sensor, a magnetometer sensor (digital compass) and an accelerometer sensor (tilt sensor). Using these data, the TellMe application formulates directional queries that are performed in a spatial database to determine if any spatial interaction exists between the query "window" and any of the buildings or Point-of-Interest (PoI) objects in our 3D model testbed. The shape of this query window takes on a variety of 2D and 3D forms from a simple ray, to a polygon, to a volume (see Figure 14). The query results of this interaction are subsequently returned to the user in the form of building address/details plus web-links to more information (e.g. classroom timetables, office hours, etc.).

Essentially, the TellMe application provides a framework for processing two dimensional queries in an open, non-directional query-space (e.g. standard range query), but with 2D hidden query removal. This allows for limiting query results to only what is actually visible to the user and subsequently addresses some of the major information overload issues in the field of MSI research within cityscape environments.

In relation to sensor data quality (e.g. signal-to-noise ratio), tests to compare the data gathered from a higher quality external sensor pack to that of integrated phone sensors are ongoing. These findings will ultimately help determine the suitability/ accuracy of current mobile device technology for providing meaningful mobile spatial interaction in LBS applications. Following this, our attention shifts to exploring three dimensional visibility with the development of an "Ego-Visibility" query processor that further confines the query space to simulate a user's view field-of-view frustum. Results from this will enable us to incorporate full 3D hidden query removal functionality.

The remainder of the article is organized as follows. Section 2 describes some related work in the area of directional querying. Section 3 outlines some methods used in current GIS to perform these queries using visibility analysis. Section 4 describes the design of our TellMe application which includes the LocateMe and DirectMe modules, the TellMe Server, and other hardware considerations. Section 5 demonstrates our directional query approach on a Nokia Navigator cell phone with five different types of possible queries, and Section 6 concludes with a summary and future work.

2 Related Work

There have been a number of applications recently that utilize compasses and tilt sensors attached to custom built mobile devices. Some early work by Wilson and Pham (2003)

and Wilson and Shafer (2003) introduced the idea of an XWand pointing device that queries an intelligent environment using a variety of sensors in a combination that supports pointing and gesture recognition inputs. A museum guide is described in Chan et al. (2005) to help visitors quickly locate exhibits using orientation aware handheld devices. In Baillie et al. (2005), an MSI method is presented where a user points a mobile device at a building to view a virtual representation of it at different times in the past. In Essl and Rohs (2007), an approach to use sensor data for turning the device into a musical instrument using shaking and sweeping gestures is reported. The Shoogle project (Williamson et al. 2007) aims to use inertial sensing to provide an eyes-free vibrotactile display that mimics objects rattling around inside the device for a rich multimodal interaction without any visual attention.

The majority of current research focuses on providing enhanced navigation capabilities in the area of MSI. The Point-to-Discover GeoWand (Simon and Fröhlich 2007a, 2008) is a system and application framework for orientation-aware location-based mobile services. This application demonstrates how their custom-built device can be used as a pointing tool to display web-based information about bars and restaurants on the device. A very similar approach is taken by Intelligent Spatial Technologies with the iPointer application (see http://www.i-spatialtech.com/ipointer.html for additional details). It is based on an augmented reality engine and a thin client API that provides a local mobile search based on GPS and an eCompass to deliver content, such as pictures, menus, and audio overviews streamed back to the user's phone. A rather different approach is taken by Robinson et al. (2008) and their Point-to-GeoBlog application, where users are able to select landmarks by using the point and tilt functionality of yet another custom built device. No content is provided up front but is later when the user logs onto a computer with more visually adequate display capabilities.

A detailed usability study in Robinson et al. (2008) reports that the most user friendly approach was to provide a simple point and tilt interface over a visual map and also a preference for remote tagging that allows users to select landmarks beyond their line-of-sight. These results confirmed a comparative outdoor study by Fröhlich et al. (2006) that tested conceptual designs for four interaction areas considered important for spatial information applications. The information Pull technique, where the user decides what information to view and has control over it, was the most intuitive and preferred data interaction approach tested by users (Persson et al. 2002, Strachan et al. 2007, Strachan and Murray-Smith 2009). The Push technique, where all information is automatically presented to the user, is not easily managed on the variously display (and memory and processor and network connection) constrained mobile devices.

An alternative approach that combats problems related to information overload is to restrict the search space based on certain criteria. In Gardiner and Carswell (2003), the approach is to restrict the search space to a user's field-of-view (FoV) using the concept of an observer's two dimensional query frustum to determine what the user can actually see from their location in the environment. In contrast, the use of the Point-to-Discover block model and visibility computation algorithm in (Simon and Fröhlich 2007a) uses a rather different approach to determine what buildings the observer can actually see from a single vantage point. In (Simon and Fröhlich 2007b), this idea is extended with the development of a local visibility model that introduces the concept of "billboards" as a mechanism to identify what buildings a user can actually see. Determining this type of egocentric visibility is a primary focus of this

article. We consider only possibilities in relation to what a user can physically see from their current position by performing visibility analysis calculations. Utilizing this approach to visibility querying on today's "off-the-shelf" mobile phones is a new concept and a key aspect of our research.

3 Visibility Analysis

There are a number of methods used in current research for performing line-of-sight queries. This section gives an overview of the technologies and methods used by a number of them, and describes some emerging possibilities.

3.1 Ray Tracing

The ray tracing process is fundamental in this area and works by simulating the light that travels within a given space. In many applications, this process is performed backwards to determine the visibility from one or multiple points (e.g. user positions), which instead of perceiving the light emitted from the objects themselves (e.g. buildings or other infrastructure), the ray is transmitted from a given user position outwards in all directions. By retrieving all the intersections between any generated ray with objects, a polygon in a 2D environment and a volume in a 3D environment can be constructed to represent the visibility query space from a view point (Figure 1). This approach is termed Line-of-Sight (LoS) in Geographic Information Systems (GIS).

3.2 Line-of-Sight (LoS)

The LoS function determines the visibility from the observer's position to a target point considering DTM fluctuations and other obstructions. The path from the view-
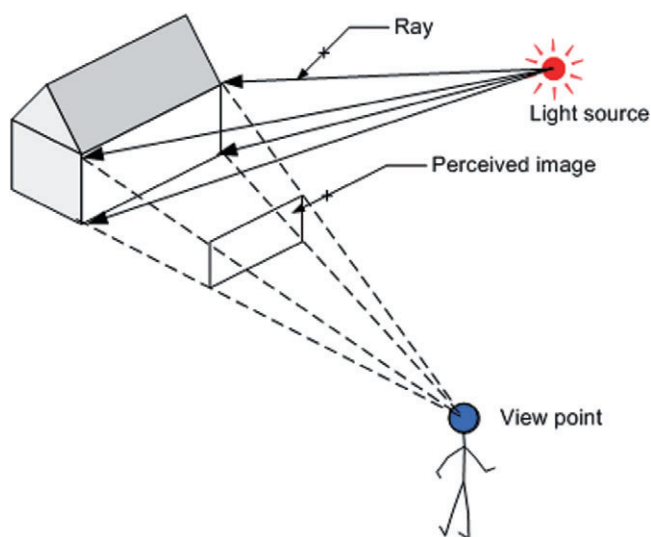
Figure 1 Ray tracing process
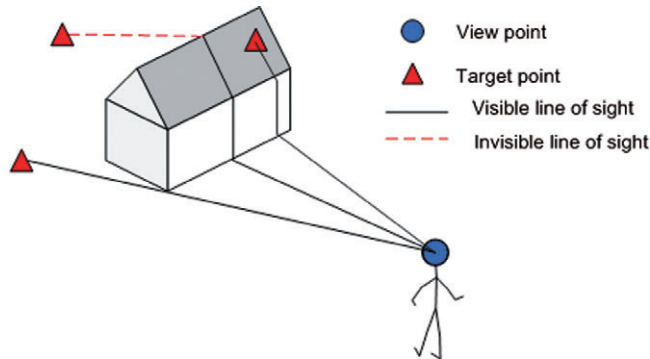
Figure 2 Line-of-sight process

point to the target point is one of the rays emitted from the viewpoint. In effect, some of the projected rays are truncated where they intersect with obstacles (e.g. building blocks, hills, etc.) while on route to the target. This collection of interaction points forms a convex polygon or volume representing the visibility area of a specific viewpoint. In a 2D environment, where the elevation of the objects is not taken into account, the line-of-sight calculation can be simplified by recording all the intersection points between the rays and obstacles surrounding it in the horizontal plane (Figure 2).

In a three dimensional space, however, the ray is not just projected in the horizontal plane but also in the vertical plane using the tilt angle in 360 degrees. In space syntax terminology, this particular shape is referred to as an Isovist (Benedikt 1979). A 2D Isovist is a visibility polygon comprised of the set of all points visible from a given vantage point in space within an environment. In Jiang and Liu (2009) this concept has been extended recently with the automatic generation of an axial map which in turn can be used to generate the Isovist.

3.3 Viewshed Analysis

A similar function to Isovist is available in GIS and is known as Viewshed Analysis. It adopts the same approach as line-of-sight analysis; however instead of examining the single path from the viewpoint to the target point, a beam of multiple rays is generated from the viewpoint in all directions in the horizontal plane. Concurrently, a beam of rays is generated vertically along each horizontal ray path and the azimuth of each ray is taken into consideration within a certain range. For instance, in Figure 3 the viewshed is based on the viewpoint and the range combined with the horizontal, vertical, and tilt angles, where the actual visible space is a volume excluding any section of this 3D query shape that intersects with a building or other object.

Adopting such an approach to measure the visible space within a certain radius can significantly reduce the number of calculations required to determine the possible intersection shape. This strategy is similar to radar scanning the environment with a limited signal distance (strength) and is essentially an extreme case of the radial line-of-sight approach.
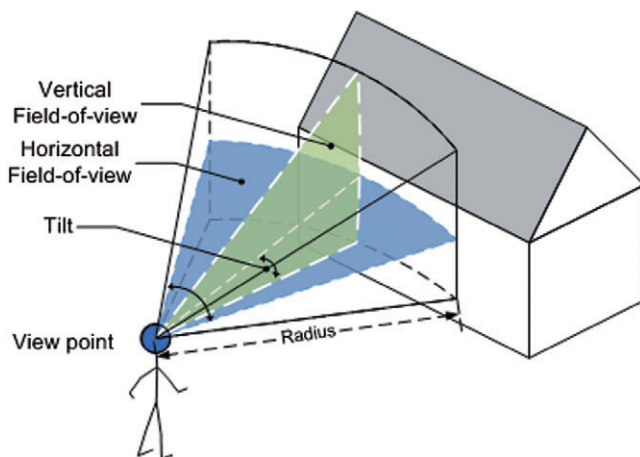
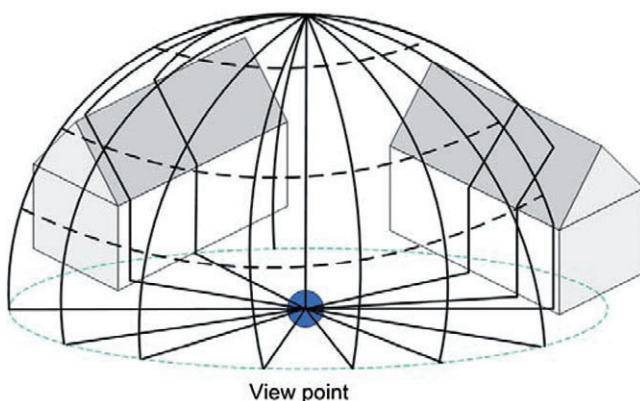**Figure 3** Viewshed analysis in GIS

Figure 4 Example of a Threat Dome query space

3.4 *Threat Dome Analysis*

To estimate the visibility shape in all directions from the viewpoint using viewshed analysis, the azimuth in the horizontal plane and the view angle in the vertical plane can be extended to a full range, which generates rays between 0 and 360 degrees to form a dome shaped viewshed. An example of this idea is the Threat Dome from military applications, which is used for determining all possible locations in space that are visible to/from a given position (ESRI 2009, Skyline 2009).

In contrast to viewshed analysis, capturing the visible volume from a point taking into account all existing obstacles (buildings, etc.) in the environment can be very computation intensive. The threat dome approach specifies a radius defining the view distance from a point and hence the rays emitted from the viewpoint form a sphere. By determining the intersections of the rays with obstacles at different levels of elevation within the sphere, the actual visible space within this specified radius can be determined (Figure 4). Decreasing the distance between the levels of calculation will improve the granularity/accuracy of the query shape.

4 TellMe

Our mobile TellMe application requests information from two modules, namely the LocateMe and DirectMe modules, to perform directional queries using the TellMe Server against various spatial data sets. Two main approaches were investigated in relation to implementing the TellMe Server.

One approach was to host the spatial dataset on the mobile device itself. In this case the data retrieved from LocateMe and DirectMe is used to perform spatial queries locally on the mobile device. The results of the queries are displayed to the user by the TellMe Mobile application. Because of limited memory and the processing speeds required to perform spatial queries on large geographic data sets, this type of architecture was deemed unsuitable for unconstrained real-time spatial querying.

The alternative approach (i.e. thin client/server approach) is to request the required data from LocateMe and DirectMe and perform spatial queries on a dataset hosted on an external server. The parameters are passed to a web application, which in turn carries out the queries on a spatial database to determine any spatial interaction between the user's location and orientation with the dataset. The results are subsequently displayed to the user in a web browser on the mobile device. This architecture is illustrated in Figure 5.
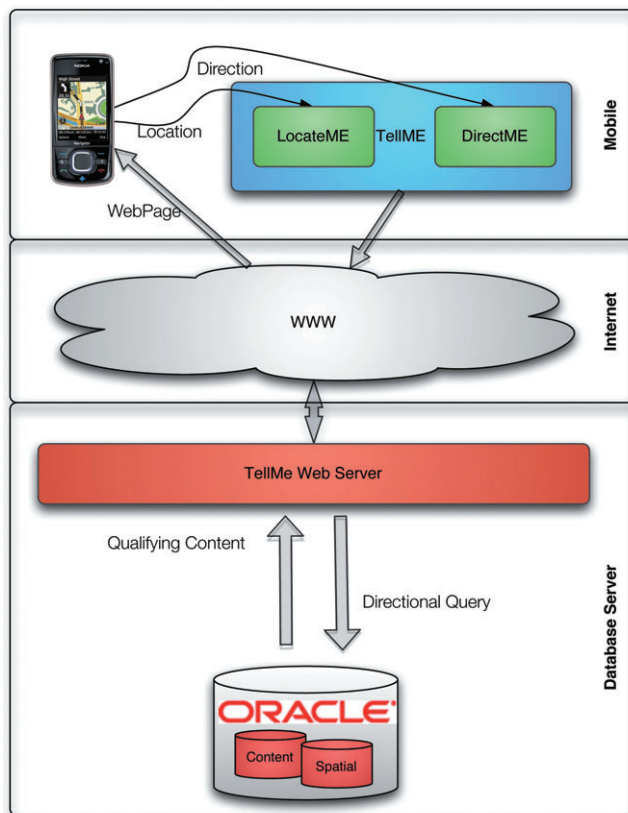
Figure 5   Overall TellMe architecture

This approach of using a web server to query and display the returned data in a web browser follows current trends because of the increasing complexity associated with developing applications to run on many different devices (W3C 2008). To avoid developing a different application for each mobile operating system, a more attractive solution is to develop applications using a web scripting language capable of producing highly interactive applications that can perform the same functions as desktop applications without the concern about underlying operating system functions or restrictions.

Taking this approach a step further, it is now becoming more acceptable to allow the browser access to more different types of data as well. For example, using a set of JavaScript APIs, it is possible to access some of the hardware on a mobile device directly from the browser. This is one of the newest features of Google Gears where the Geolocation API allows an application access to the GPS hardware on the device (Google 2008). Using this approach, if all sensor data were made available to a web application, it would be possible to eliminate the need for an on-device application altogether, essentially making the TellMe application entirely web-browser based.

4.1 LocateMe

The LocateMe module is used in conjunction with the DirectMe module to gather information about a user's current position and orientation. These data are then used to execute directional queries against various data sets both internal and external to the phone. The LocateMe module is based on a hybrid positioning calculation that utilizes GSM, Wi-Fi and Bluetooth radio signals in addition to GPS to determine location. A more comprehensive description of this technology can be found in Rooney et al. (2007b).

4.2 DirectMe

The DirectMe module is one of two modules that provide data to the TellMe application. Its function is to determine the direction that the mobile device is currently pointing by using a digital compass and tilt sensors. These data are then collected on request from the TellMe application and synchronized with the location data coming from the LocateMe module. The architecture of the DirectMe module is similar to the LocateMe module where each sensor has a native hardware spotter that relays data to a higher-level component that synchronizes data coming from each spotter. This architecture is illustrated in Figure 6.

This type of architecture is required to overcome the frustrating restrictions imposed on would-be developer access to various mobile device hardware components (e.g. compass and tilt sensors) found within some proprietary operating system APIs.

4.3 Mobile Device Hardware/Sensors

There are currently three main mobile platforms that contain the required sensor hardware for our TellMe system. The Nokia Navigator 6210 has a digital compass and tilt sensors and the API providing access to these sensors has just (Q4, 2009) been back ported from Symbian S60 5th Edition to Symbian S60 3rd Edition FP2 (Nokia 2008). Using this version of the API, the DirectMe module can ascertain the current heading and orientation of the device. However, there have been recent reports that the
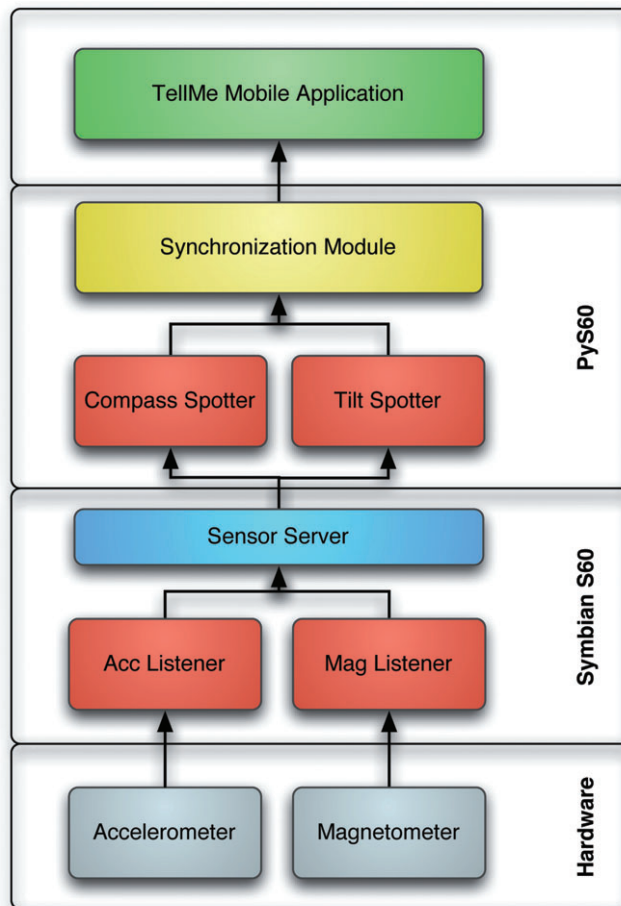
Figure 6 The design of the DirectMe module

performance of these integrated sensors is of poor quality (Essl and Rohs 2007). Another hardware option is the HTC Dream (Google phone), also with integrated digital compass and accelerometers. This device runs the Android operating system from Google and at time of writing is not yet available for sale in Ireland (HTC 2008). The most recent possibility in terms of hardware suitability is the Apple iPhone 3GS (2009). In addition to providing WiFi and accelerometer access there is now a compass integrated as well.

Considering today's limited number of off-the-shelf mobile devices available for acquiring orientation type data, another option is to use external sensors packs such as the SHAKE. The SHAKE SK7 is an external sensor pack with digital compass, accelerometer, and tactile feedback, among others. This sensor pack communicates with a recording device (e.g. laptop, cellphone) via Bluetooth. Unlike current cellphone sensors, the SHAKE device has better quality sensors and a number of filters on board to reduce any noise introduced by, for example, the mobile phone's antenna. However, following a review of these options, it was decided that the DirectMe module would collect data from a number of integrated MEMS (Micro Electro-Mechanical Systems) sensors on a
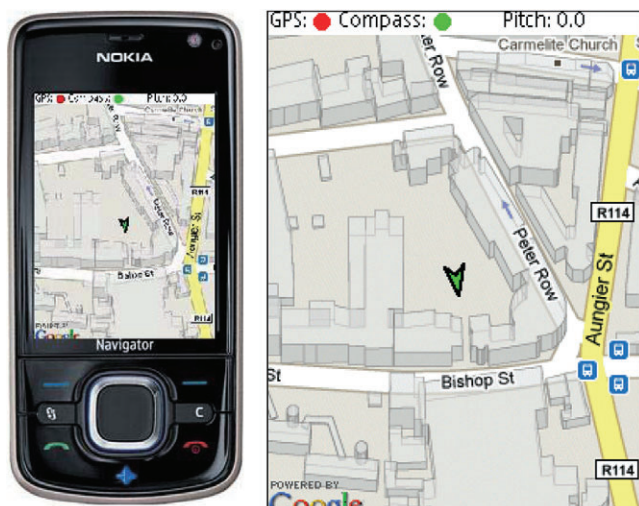
Figure 7   Nokia Navigator 6210 with directional map display

Nokia Navigator 6210 (Figure 7) instead of the SHAKE to demonstrate the state-of-the-art application development potential of today's mobile devices.

4.4 TellMe Server

The TellMe Server is essentially a spatial application server that performs all the complex spatial queries in the TellMe system. It is responsible for communicating with the TellMe Mobile application, which collects data (including location, direction and orientation) from the LocateMe and DirectMe modules on the mobile device. These data are communicated wirelessly to the spatial application server and used to perform spatial queries against an Oracle Spatial database.

The TellMe Server is based on ESRIs ArcGIS Server platform and is used to perform the complex queries required for determining the mobile spatial interaction between a user's line-of-sight and the 3D database. This platform provides server extensions for many of the traditional spatial query functions found in most GIS. Using these extensions, it is now possible to perform these types of spatial queries in a mobile context that were previously only possible in a desktop setting. Another important reason for this choice of server is based around the issue of scalability; using this software ensures that the TellMe system is scalable to manage a large number of users efficiently.

4.5 3D Database

To perform directional queries, a 2D spatial database is the minimum requirement for supporting extensive feature types, indexing techniques, and for performing 2D spatial queries effectively to identify objects (e.g. buildings, PoIs) that intersect with a direction vector. Some types of queries possible in 2D are illustrated in Figures 10–12.

However, to perform 3D spatial queries this process becomes somewhat more complex. In comparison, a 3D spatial query should be able, for example, not only to

identify which building a direction vector is intersecting with but also the floor (or window, door, etc.) of the building it is directed at. This means a true 3D database should support three-dimensional data types such as point, line, surface and volume in its geometric data model, be capable of indexing the data and must also offer functions and operations embedded in its spatial query language that can operate on these data types (Schön et al. 2009). In fact, these requirements significantly reduce the number of database options of enterprise level capabilities that are available to us – there are two main options here.

4.5.1 Oracle 11g

Beginning with the 11g version of Oracle, it is now possible to store, index and query 3D spatial objects using the sdo_geometry data type. Using this datatype, it is now possible to store point, line, polygon, polygon with hole, and collection data types in 2D and 3D (Schön et al. 2009). An example of the types of data that can be stored is illustrated in Figure 8b.

4.5.2 ESRI ArcGIS

To support this rising trend in 3D spatial data storage, ESRI developed a native volumetric geometry feature type called the Multipatch feature supported by its geodatabase models. The Multipatch is treated like any other geometry type in the database and is constructed of triangle strips and fans to define object boundaries using triangular faces (Figure 8a).

   In our case, we chose Oracle 11g for its 3D forms: simple solid, composite surface, composite solid and collection to represent our cityscape data. Our 3D data is based on Ordinance Survey Ireland's (OSi) 2D vector footprint data overlaid on a 10 m DTM and extruded using height values taken from airborne LiDAR scans (see http://www.osi.ie/ for additional details). The result is a building block model of the Dublin city centre. This
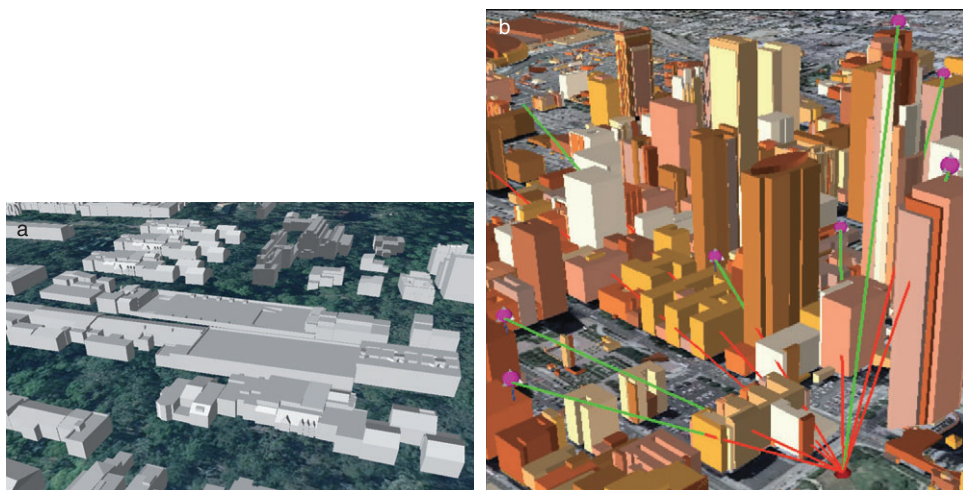

Figure 8   (a) ArcGIS multipatch 3D objects (ESRI 2009) and (b) Oracle SDO_Geometry
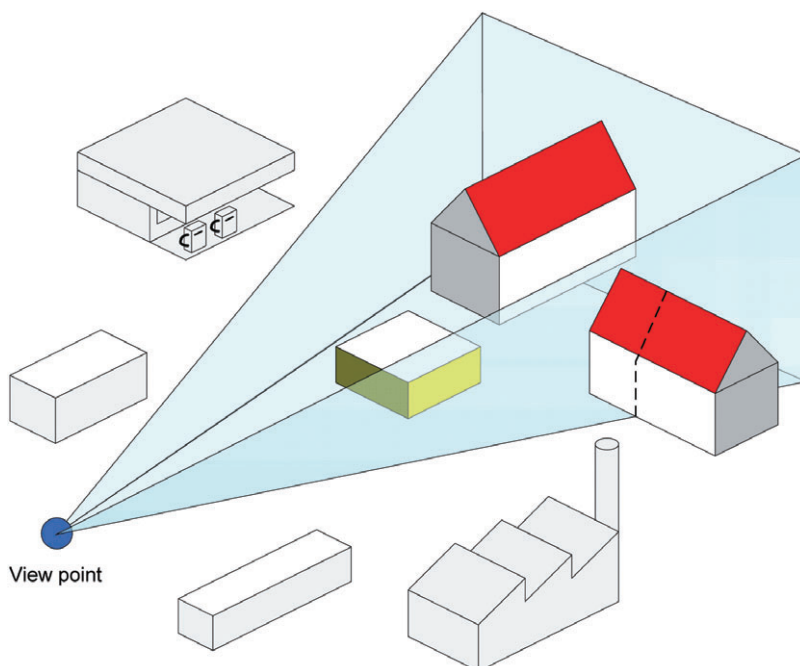
Figure 9 3D spatial query in Oracle 11g using field-of-view frustum as query window shape

data are spatially indexed and queried using a 3D query window generated by the sensor data collected from the mobile device. All attribute data (building name, class timetables, etc.) presented to the user are also stored in the Oracle database. To perform 3D queries, the set of operators is restricted to SDO_Filter, SDO_Anyinteract, SDO_Within_Distance, and SDO_NN (nearest neighbour) using the Geographic-3D coordinate system. An example of a typical 3D frustum query shape is shown in Figure 9.

5 Demonstration on Nokia Navigator 6210

In this section, the different types of queries that are performed by the TellMe application are discussed, all of which are based on the user's egocentric point-of-view. Egocentric visibility refers to the portion of a search space that is visible to a user at a particular time based on their location, direction and orientation. For example, in the TellMe application, a user's visible query space in 2D acts as a secondary filter on data returned by the query processor by further restricting it to contain only objects that are in the user's field-of-view.

Ego-visibility therefore excludes the portions of the dataset obscured by buildings and is primarily used to identify points-of-interest (PoI) other than buildings that are in a user's FoV within a predefined distance (Figure 11). This query method could be used to identify objects that may be too small to point at precisely from a distance but are still in the user's FoV nonetheless, like a monument or other tourism artefact. The algorithm

to determine the query space in this instance builds on our previous work in Gardiner and Carswell (2003) where a 2D directional query processor was developed for an on-line virtual environment.

Although Nokia mapping is based on Navtec maps, for ease of implementation (i.e. avoiding Symbian C++ programming issues) and interaction we use Google Maps as the background map for the query results. The returned result is overlaid on the map in the form of building outlines or PoIs to offer the user an opportunity to adjust the pointing gestures and re-query if necessary. The mapping interface on the mobile device was shown in Figure 7, where the arrow indicates the position and direction a user is pointing and the colour represents the status of the calibration level of the sensor signals (i.e. green = good, red = poor). The following sections present the types of queries that can be performed by the TellMe application.

5.1 Point-to-Select Query

The Point-to-Select query function allows the user to point the device in a particular direction and using this direction vector, retrieve information about a particular object. All intersecting objects are presented initially to the user for refining the selection. When the user selects one from the list, an outline of it (outline coordinates taken from the Oracle database) gets displayed on the mobile device with the direction vector highlighted. An example of this is shown in Figure 10.

5.2 Field-of-View Query

The Field-of-View query simulates the user's actual visual field in a particular direction and generates a 2D view shape as a query window in the spatial database. Only database objects that intersect the FoV shape get returned to the user. By default a 120° angle is
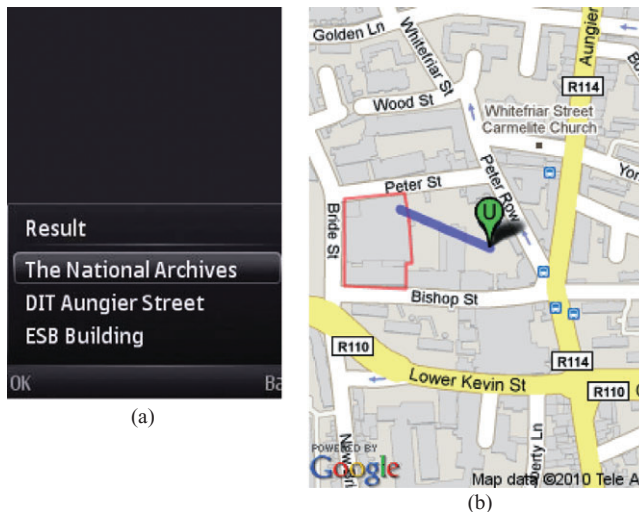

(a)

(b)

Figure 10 (a) All buildings that intersect with the direction vector and (b) Selected building and direction vector from user's location displayed on map
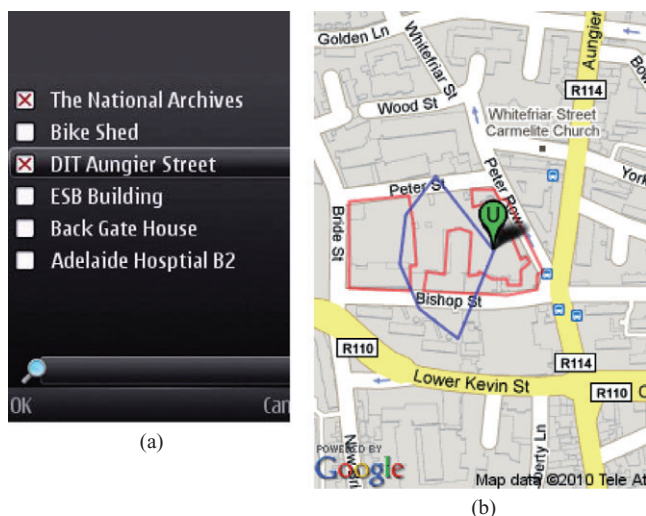
(a)

(b)

Figure 11   (a) All buildings present in the user's field-of-view and (b) Selected buildings and FoV window displayed on the map

chosen for the view field to mimic the natural view angle of human vision. The radius of the query shape (visible distance) is dynamic and varies based on the tilting angle of the device. For example, if the device is pointing vertically (i.e. to the sky) the radius is 0 m, and if the device is pointing horizontally the radius is set to 200 m, with variable distances in between. An example is shown in Figure 11.

5.3 Isovist Query

Recent work in this area by Simon and Fröhlich (2007b) describes a local visibility model (Lvis) that uses the concept of billboards to determine what buildings are in the user's FoV in a 2.5D environment. To achieve this in our case, a different approach is taken based on work carried out by Jiang and Liu (2009) into the concept of Isovists and medial axes. Using Isovists, we attempt to automatically generate a panoramic (360°) line-of-sight search space. This method has shown to be very effective as the fundamental aim is firstly, to determine precisely the geographic area visible to the user in all directions and secondly, to determine what objects inside this area are of interest to the user. Considering the dense distribution of buildings in a city, by default we only consider buildings within a 300 m radius in our calculations. Figure 12a shows the result of the query displayed on the server, and Figure 12b shows the results overlaid on the map of a mobile device indicating the user's visibility convex polygon (Isovist). Figure 12c illustrates how the Isovist can be used to identify all visible buildings or PoIs that interact with a user's complete visibility polygon and Figure 12d overlays them on the phone's map display.

5.4 Frustum Query

The Frustum View query simulates a user's three dimensional field-of-view. It essentially allows the user to utilize the mobile device like a flashlight beam scanning
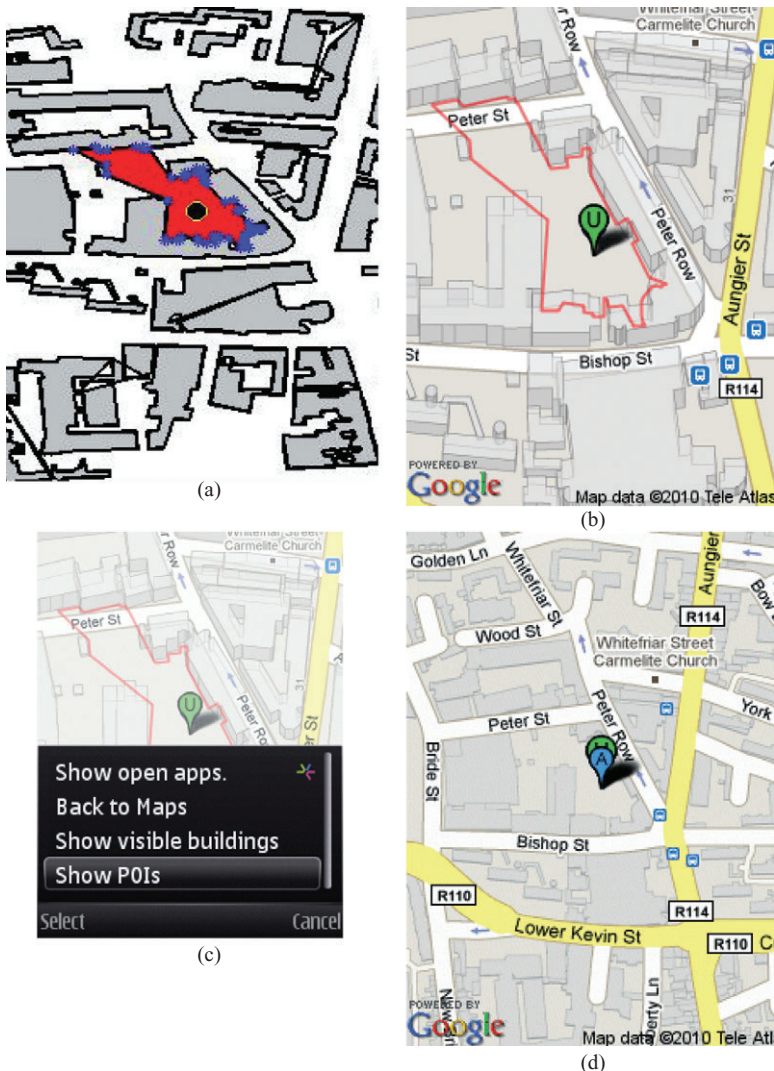
Figure 12 (a) Isovist calculation as displayed on the server; (b) The shape of Isovist overlaid on the client-side map; (c) Further request to return all intersecting PoIs to the user; and (d) A PoI query result using the Isovist shape as the query window

the building to retrieve any information it illuminates. An example is shown in Figure 13.

In this case, a 3D Isovist is represented by a 3D object that represents a user's actual LoS omitting the geometry of any other objects that interact with it. Building this type of 3D Isovist query requires a detailed 3D database of buildings and building "furniture" (e.g. windows, doors, etc.) at floor/room level in order to calculate the query frustum geometry accurately. With today's usage of Lidar scans for 3D cityscape construction, the availability of such geometrically detailed datasets is becoming more and more common. Examples of possible query shapes used for the frustum are illustrated in Figure 14.
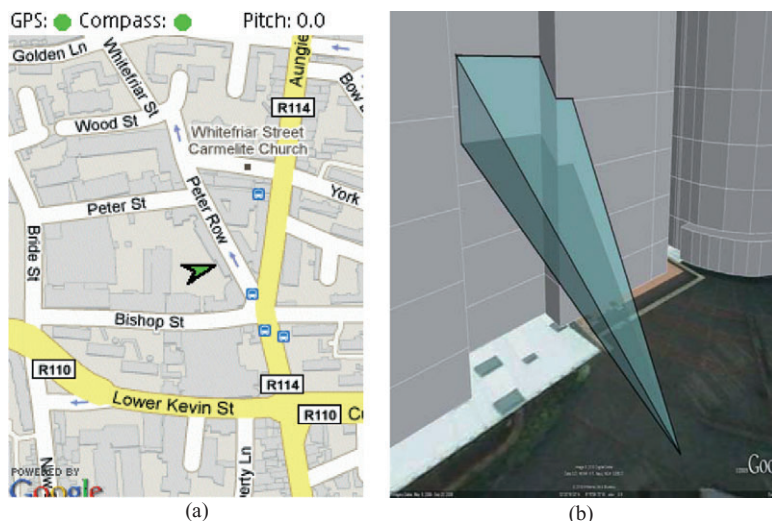
Figure 13 (a) Frustum pointing direction in 2D and (b) Visualization on the server of the frustum query shape interacting with 3D buildings in the database

5.5 *Ego-Dome*

The Ego-Dome is essentially an extension of the 2D Isovist described previously except the visible open space around a user at a particular location is reproduced in 3D. The concept of querying all open space in this manner has been explored previously in military applications but usually in standalone desktop situations only. In the case of our TellMe application, an Ego-Dome is created using the current location of the user (Figure 15) and any objects that interact with the dome volume are detected and presented using visual, auditory, or tactile interfaces. This type of spatial interaction is very useful for a host of applications in terms of real-time feedback, allowing the user an ability to interact with and have knowledge about their surrounding environment in real-time.

6 Conclusions

The TellMe MSI application is designed to allow users to interact in a context sensitive way with environmental information using off-the-shelf mobile phone technology. Today's mobile devices are beginning to offer integrated hardware such as digital compasses and tilt sensors combined with the now almost ubiquitous GPS sensor. Through these technologies, a wide spectrum of mobile spatial applications is emerging. Further development and testing of the DirectMe module and its synchronization with the LocateMe module is ongoing. Concurrently, spatial database development for refining/ testing our directional query techniques in a comprehensive TellMe mobile application is taking place.

In relation to performing directional queries there are a number of fundamental issues, highlighted during our research, regarding the quality of the data produced from
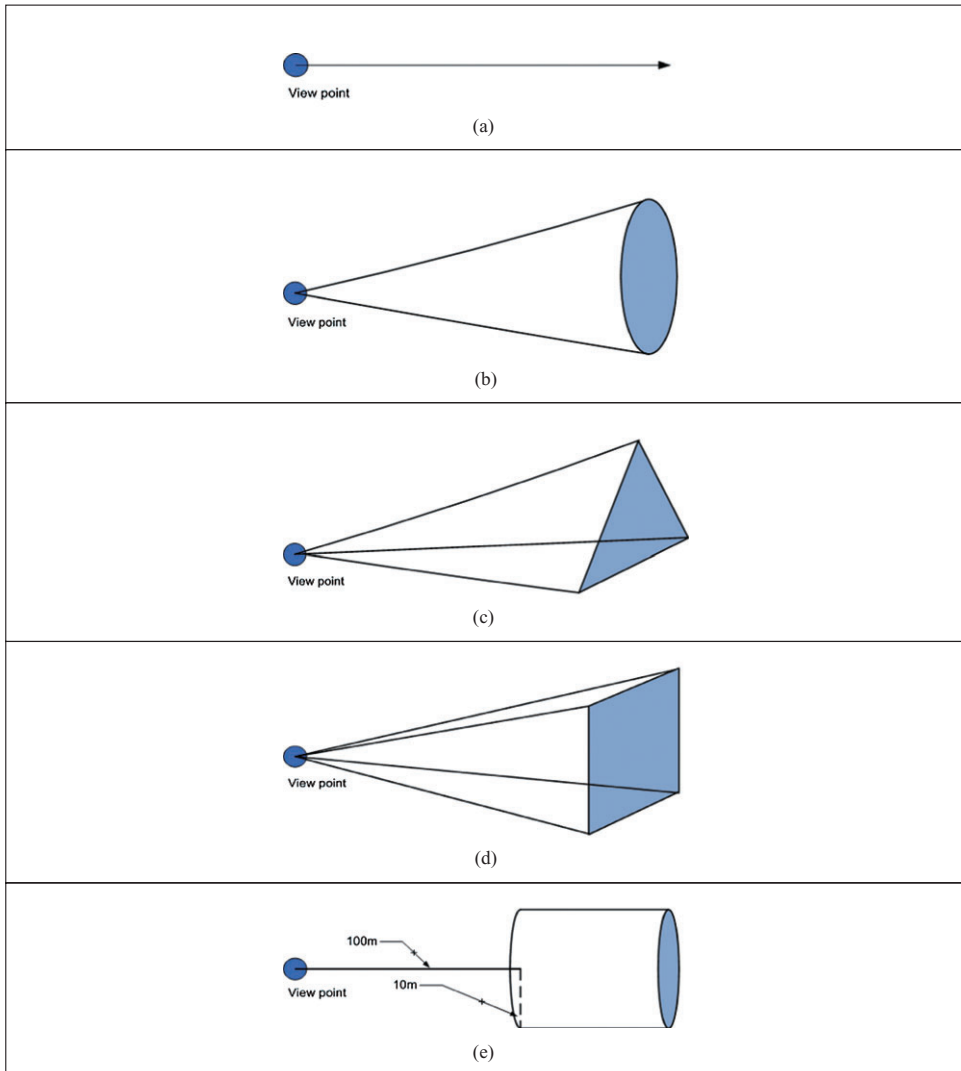
Figure 14   Examples of possible 3D visibility frustum query shapes

the sensors. In Simon and Fröhlich (2008) the quality of the sensor data from a custom built sensor pack is analysed and compared to actual data showing good results for performing directional queries. With the quality of the sensor data being an important aspect of directional queries, in terms of the relevance of the results returned and the move towards making sensor interaction available to developers on mainstream mobile devices, we intend to carry out more sets of quality assurance testing on real-world datasets. We have already investigated the quality of selected external sensors (i.e. SHAKE and Nokia LD-4W GPS) and will compare these results with data quality returned from the integrated sensors in our off-the-shelf mobile phone (i.e. Nokia Navigator 6210).
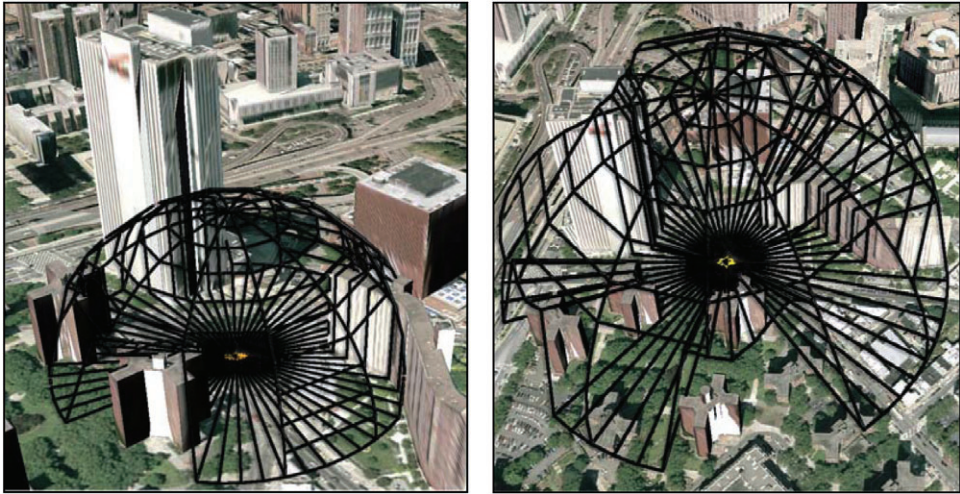
Figure 15   Ego-dome for determining the visible query space within a certain radius

Further testing will be carried out on a detailed NUI Maynooth 3D Campus model constructed from LiDAR data. In relation to 3D construction of the query shape, a new approach will be introduced that uses the FoV parameters from the built-in cell phone camera for the query frustum. This will combine reality with the query results in the same display. In addition, and in keeping with current trends of publishing mobile apps on the web, we intend to investigate making the TellMe technology available to a larger audience through various online app stores (e.g. iPhone App Store, Android Market, Nokia OVI Store).

Acknowledgements

References

Baillie L, Kunczier H, and Anegg H 2005 Rolling, rotating and imagining in a virtual mobile world. In *Proceedings of the Seventh International Conference on Human Computer Interaction with Mobile Devices and Services*, Salzburg, Austria: 111: 283–6

Benedikt M L 1979 To take hold of space: Isovists and Isovist fields. *Environment and Planning B* 6: 47–65

Chan L-W, Hsu Y-Y, Hung Y-P, and Hsu J Y-J 2005 Orientation-Aware handhelds for panorama-based museum guiding system. In *Proceedings of the 2005 UbiComp Workshop on Smart Environments and their Applications to Cultural Heritage*, Tokyo, Japan: 43–6

ESRI 2009 ESRI globe. WWW document, http://www.esri.com/products/index.html#4

Essl G and Rohs M 2007 ShaMus: A sensor-based integrated mobile phone instrument. In *Proceedings of the International Computer Music Conference (ICMC)*, Copenhagen, Denmark: 27–31

Fröhlich P, Simon R, Baillie L, and Anegg H 2006 Comparing conceptual designs for mobile access to geospatial information. In *Proceedings of the Eighth International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI 06)*, Helsinki, Finland: 159: 109–12

Gardiner K and Carswell J 2003 Viewer-based directional querying for mobile applications. In *Proceedings of the Third International Workshop on Web and Wireless Geographical Information Systems (W2GIS)*, Rome, Italy: 83–91

Google 2008 Google Gears API. WWW document, http://code.google.com/apis/gears/api_geolocation.html

HTC 2008 HTC: Dream. WWW document, http://htcdream.com/

Jiang B and Liu X 2009 AxialGen: A research prototype for automatically generating the axial map. In *Proceedings of the Eleventh International Conference on Computers in Urban Planning and Urban Management*, Hong Kong

Kilfeather E, Carswell J, Gardiner K, and Rooney S 2007 Urban location-based services using mobile clients: The ICiNG approach. In *Proceedings of the 2007 GISRUK Conference*, Maynooth, Ireland: 227–30

Nokia 2008 Nokia: Sense your location. WWW document, http://www.nokia.com/A41229034

Persson P, Espinoza F, Fagerberg P, Sandin A, and Cöster R 2002 GeoNotes: A location-based information system for public spaces. In Höök K, Benyon D, and Munro A (eds) *Readings in Social Navigation of Information Space*. Berlin, Springer: 151–73

Robinson S, Eslambolchilar P, and Jones M 2008 Point-to-GeoBlog: Gestures and sensors to support user generated content creation. In *Proceedings of the Tenth International Conference on Human Computer Interaction with Mobile Devices and Services*, Amsterdam, The Netherlands: 197–206

Rooney S, Gardiner K, and Carswell J 2007a An open source approach to wireless positioning techniques. In *Proceedings of the Fifth International Symposium on Mobile Mapping Technology (MMT 07)*, Padua, Italy

Rooney S, Gardiner K, and Carswell J 2007b Wireless positioning techniques: A developer's update. In Ware J M and Taylor G E (eds) *Web and Wireless Geographical Information Systems: Proceedings of the Seventh International Symposium (W2GIS 2007)*. Berlin, Springer Lecture Notes in Computer Science Vol. 4857: 162–74

Schön B, Laefer D F, Morrish S W, and Bertolotto M 2009 Three-dimensional spatial information systems: State-of-the-art review. *Recent Patents on Computer Science* 2: 21–31

Simon R and Fröhlich P 2007a The point to discover GeoWand. In *Proceedings of the Ninth International Conference on Ubiquitous Computing (UbiComp 07)*, Innsbruck, Austria: 3–10

Simon R and Fröhlich P 2007b A mobile application framework for the geospatial Web. In *Proceedings of the Sixteenth International Conference on the World Wide Web*, Banff, Alberta: 381–90

Simon R and Fröhlich P 2008 GeoPointing: Evaluating the performance of orientation-aware location-based interaction under real-world conditions. *Journal of Location Based Services* 2: 24–40

Skyline 2009 Skyline Terraserver. WWW document, http://www.skylinesoft.com/skylineglobe/corporate/Products/TerraExplorer.aspx

Strachan S and Murray-Smith R 2009 Bearing-based selection in mobile spatial interaction. *Personal and Ubiquitous Computing* 13: 265–80

Strachan S, Williamson J, and Murray-Smith R 2007 Show me the way to Monte Carlo: Density-based trajectory navigation. In *Proceedings of the FP 205 SIGCHI Conference on Human Factors in Computing Systems*, New York: 1245–48

Williamson J, Murray-Smith R, and Hughes S 2007 Shoogle: Excitatory multimodal interaction on mobile devices. In *Proceedings of the ACM SIG CHI Conference*, San Jose, California: 121–4

Wilson A and Pham H 2003 Pointing in intelligent environments with the WorldCursor. In *Proceedings of the Interact 2003 Conference*, Zurich, Switzerland: doi:10.1.1.88.4174

Wilson A and Shafer S 2003 XWand: UI for intelligent spaces. In *Proceedings of the 2003 Conference on Human Factors in Computing Systems*, Fort Lauderdale, Florida: 545–52

5.4. Paper IV: EgoViz – A Mobile Based Spatial Interaction System

Gardiner, K., J. Yin and J. D. Carswell. (2009) "EgoViz – A Mobile Based Spatial Interaction System": The 9th International Symposium on Web and Wireless Geographical Information Systems (W2GIS), Maynooth, Ireland. Springer-Verlag, Lecture Notes in Computer Science, Volume 5886, pp 135-152, DOI: 10.1007/978-3-642-10601-9_10.

EgoViz – A Mobile Based Spatial Interaction System

Keith Gardiner*, Junjun Yin, and James D. Carswell

Digital Media Centre,
Dublin Institute of Technology, Ireland
{keith.gardiner,junjun.yin,jcarswell}@dit.ie

Abstract. This paper describes research carried out in the area of mobile spatial interaction and the development of a mobile (i.e. on-device) version of a simulated web-based 2D directional query processor. The *TellMe* application integrates location (from GPS, GSM, WiFi) and orientation (from digital compass/tilt sensors) sensing technologies into an enhanced spatial query processing module capable of exploiting a mobile device's position and orientation for querying real-world 3D spatial datasets. This paper outlines the technique used to combine these technologies and the architecture needed to deploy them on a sensor enabled smartphone (i.e. Nokia 6210 Navigator). With all these sensor technologies now available on one device, it is possible to employ a personal query system that can work effectively in any environment using location and orientation as primary parameters for directional queries. In doing so, novel approaches for determining a user's query space in 3 dimensions based on line-of-sight and 3D visibility (ego-visibility) are also investigated. The result is a mobile application that is location, direction and orientation aware and using these data is able to identify objects (e.g. buildings, points-of-interest, etc.) by pointing at them or when they are in a specified field-of-view.

Keywords: MSI, Directional Query, Isovist, Radial Query.

1 Introduction

This paper focuses on mobile directional querying and the development of algorithms for 2D and 3D spatial data processing. In conjunction with location, an orientation module provides angular data to a spatial query processor making it possible to perform directional queries from a mobile device in a real world environment.

The fundamental requirements for this type of service interaction are location, direction and orientation. Some typical applications in the current literature enable a user to point a mobile device at a building and, using position and direction, determine the building's address/identity [1]. This requires both accurate location and orientation as part of the query and the modules that provide this data in our case are the *LocateMe* and *DirectMe* modules. The locator component or *LocateMe* module is by design an open source, network independent mobile location determination application that can utilise GPS, Wi-Fi and GSM beacon information or any combination

---

* Corresponding author.

of them, to trilaterate location estimates [2, 3]. The orientation component or *DirectMe* module and a primary focus of this paper uses data from a number of sensors including a GPS sensor, a magnetometer sensor (digital compass) and an accelerometer sensor (tilt sensor). Using these data, the *TellMe* application formulates directional queries that are performed by a spatial database to determine if any spatial interaction exists between the query "window" and any of the buildings in our 3D university campus model. The shape of this query window takes on a variety of 2D and 3D forms from a simple ray to a polygon to a volume (see Figure 14). The results of this interaction are subsequently presented to the user in the form of a building address/details plus web-links to more information (e.g. classroom timetables, lab opening hours, etc.).

Primarily, the *TellMe* application provides a framework for processing 2-dimensional queries in an open, non-directional query-space (e.g. range query). This subsequently enables us to investigate some of the major issues in the area of mobile spatial interaction within 3D environments (e.g. 3D Visibility).

In relation to sensor data quality (e.g. noise), tests to compare the data gathered from a higher quality (i.e. more stable) external sensor packs with that of the integrated sensors on our mobile device are ongoing and will help to determine the suitability of current mobile devices for exploitation in the area of mobile spatial interaction. Following this, our attention shifts to exploring 3-dimensional visibility with the development of an *"Ego-Visibility"* query processor that further confines the query-space to simulate a user's view frustum. Results from this will enable us to investigate the possibility of providing full *"Hidden Query Removal"* functionality where only what the user can actually physically see gets returned by the directional query processor.

The remainder of the paper is organised as follows: Section 2 describes some related work in the area of directional querying. Section 3 outlines some methods used by current GIS to perform these queries using visibility analysis. Section 4 describes the design of the *TellMe* system which includes the *LocateMe* and *DirectMe* modules, the *TellMe Server* and hardware considerations. Section 5 details our approach to egocentric visibility and describes five different types of queries and Section 6 concludes with a summary and future work.

2   Related Work

There have been a number of different applications proposed recently that utilise compasses and tilt sensors in current state-of-the-art mobile devices. Some early work by [4, 5] introduced the idea of an *XWand* which is a custom built pointing device that controls electronic devices in an intelligent environment using a variety of sensors in combination to support pointing and gesture recognition tasks. A museum guide is described in [6] to help visitors quickly locate exhibits using orientation aware handheld devices. [7] reports on an interaction method where a user points a mobile device at a building to view a virtual representation of it at different times in the past. In [8], their approach is to use the sensor data to turn the device into a musical instrument using shaking and sweeping gestures. The *Shoogle* project [9] aims to use inertial sensing to provide an eyes-free vibrotactile display that mimics objects rattling around

inside the device. Such active use of sensors provides for a rich multimodal interaction that can be used without any visual attention.

The majority of current research focuses on providing enhanced navigation capabilities in the area of mobile spatial interaction. The *Point-to-Discover* GeoWand [1, 10] is a system and application framework for orientation-aware location-based mobile services. This application demonstrates how their custom-built device can be used as a pointing tool to display web-based information about bars and restaurants on the device. A very similar approach is taken by [11] with the *iPointer* application which is based on an augmented reality engine and a thin client api that provides a local mobile search based on GPS and an eCompass and delivers content, such as pictures, menus, and audio overviews, which are streamed back to the user's phone. A rather different approach is taken by [12]. Using their *Point-to-GeoBlog* application, users are able to select landmarks by using the point and tilt functionality of a custom built device. No content is provided up front but later when the user logs onto a computer with more visually adequate display capabilities.

A detailed usability study by [12] reported that the most intuitive approach was to provide a simple point and tilt interface over a visual map and also a preference for remote tagging that allows users to select landmarks beyond their line-of-sight. These results confirmed a comparative outdoor study by [13] that tested conceptual designs for 4 interaction areas considered important for spatial information applications (SIA). In addition, the information *Pull* technique, where the user decides what information to view and has control over it [14-16] was the most intuitive and preferred data interaction approach by users. Whereas the *Push* technique, where all information is automatically presented to the user, is not easily managed on constrained mobile devices.

An alternative approach that combats the types of problems related to too much data being presented to the user is to restrict the search space based on certain criteria. In [17], the approach is to restrict the search space to a users field-of-view using the concept of an observer's 2-dimensional query frustum to determine what the user can actually see from their position in the environment. The use of the *Point-to-Discover* block model and visibility computation algorithm in [1] uses a rather different approach to determine what buildings the observer can actually see from a single vantage point. In [18], this idea is extended with the development of a local visibility model that introduces the concept of "billboards" as a mechanism to identify what buildings the user can see.

This type of egocentric visibility will be a primary focus of this paper. We consider only possibilities in relation to what a user can physically see from their current position by using visibility analysis to do so. Utilizing this type of visibility shape on a mobile platform is a new concept and will be a key aspect of the research.

3　Visibility Analysis

There are a number of methods used in the current research described above for performing line-of-sight queries. This section gives an overview of the technologies and methods used by a number of them and describes some emerging possibilities.

3.1 Ray Tracing

The ray tracing process is a fundamental process in this area and works by simulating the light that travels within a space. In many applications, this process is performed backwards to determine the visibility from one or multiple points (e.g. user positions), which instead of perceiving the light emitted from the objects (e.g. buildings or other infrastructure), the ray is transmitted from a given point outwards in all directions. By retrieving all the intersections from a generated ray and the objects, a polygon in a 2D environment and a volume in a 3D environment can be constructed that represents the visibility from the point in 2D and 3D respectively (Figure 1). This approach is termed *Line-of-Sight* (LOS) in Geographic Information Systems (GIS).
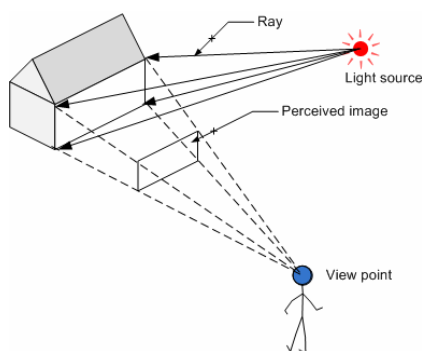
Fig. 1. Ray Tracing Process

3.2 Line-of-Sight

The *Line-of-Sight* function has been integrated in most commercial GIS software. It determines the visibility from the observer's position to a target point considering DTM fluctuations. The path from the viewpoint to the target point is one of the rays emitted from the viewpoint. In effect, some of the projected rays are truncated where they intersect with obstacles (e.g. building blocks, etc) while on route to the target. This collection of interaction points forms a convex polygon or volume representing the visibility area of a specific viewpoint. In a 2D environment, where the elevation of the objects is not taken into account, the line-of-sight function to establish the visibility of a particular point in space can be simplified by recording all the intersection points between the rays and obstacles surrounding it in the horizontal plane (Figure 2). However, in a 3 dimensional space, the ray is not just projected in the horizontal plane but also in the vertical plane using the tilt angle from the horizontal plane that ranges from 0 to 360 degrees. This particular object is referred to as an *Isovist* in a 3D environment [19]. An isovist is described as a visibility polygon comprised of the set of all points visible from a given vantage point in space with respect to an environment. Isovists were first introduced as a method for analysing space in space syntax research by [19]. In [20] this concept has been extended recently with the automatic generation of an axial map which can be used to generate an isovist. Fortunately, a similar function available in GIS is known as a *Viewshed* for geodetic survey applications.
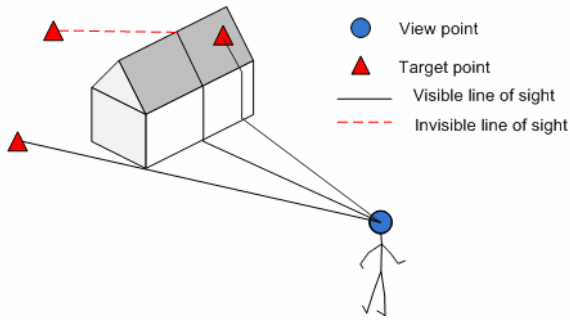
Fig. 2. Line-of-Sight Process

3.3 Viewshed Analysis

Viewshed analysis adopts the same approach as line-of-sight analysis. However, instead of examining the single path from the viewpoint to the target point, a beam of rays is generated from the viewpoint in the horizontal plane. Concurrently, a beam of rays is generated vertically along each ray path in the horizontal plane and the azimuth is taken into consideration within a certain range. For instance, in the example in Figure 3 the viewshed is based on the viewpoint and the range (based on the azimuth) combined with the horizontal angle or tilt, where the actual visible space is a volume excluding the section of the 3D object that intersects with the surface or building.

To estimate the visibility in all directions from the viewpoint using viewshed analysis, the azimuth in the horizontal plane and the view angle in the vertical plane can be extended to a full range, which generates the rays from a point between 0 and 360 degrees, forming a frustum shaped viewshed. One existing application of this idea is the *Threat Dome,* which is used for estimating all possible locations in space that are visible from a given point and is utilised mainly in military defence situations [21, 22].
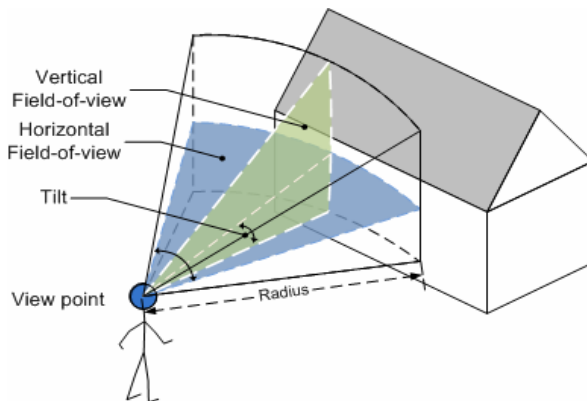

Fig. 3. Viewshed analysis in geodetic surveying

Adopting such an approach to measure the visible space within a certain radius can significantly reduce the number of calculations required to determine the possible intersection points. This strategy is similar to radar scanning the environment with a limited signal distance (strength). Essentially, it is an extreme case of the radial line-of-sight, which is derived by calculating the basic line-of-sight.

3.4 Threat Dome Analysis

In contrast to viewshed analysis, capturing the visible volume from a point taking into account all existing obstacles (objects) in the environment can be very computationally intensive. The threat dome approach specifies a radius defining the view distance from a point and hence the rays emitted from the viewpoint form a sphere. By determining the intersections of the rays with obstacles in different levels of elevation within the sphere, the actual visible space within this specified radius can be determined (Figure 4).
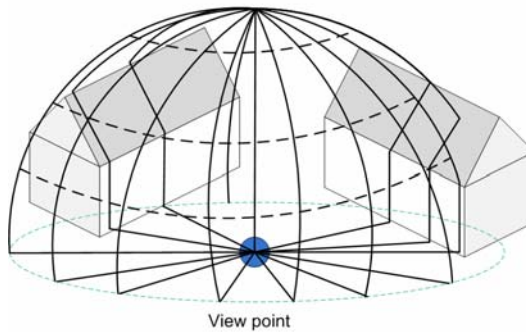
Fig. 4. Example of Threat Dome

4 TellMe

The *TellMe* application is the mobile application that requests information from two modules, namely the *LocateMe* and *DirectMe* modules and uses this information to perform directional queries using the *TellMe Server* against various spatial data sets. Two main approaches are investigated in relation to implementing the *TellMe Server*.

One approach is to host the spatial dataset on the mobile device itself. In this case the data retrieved from *LocateMe* and *DirectMe* is used to perform spatial queries locally on the mobile device. The results of the queries are displayed to the user by the *TellMe Mobile* application. Because of limited memory and the processing speeds required to perform spatial queries on large geographic data sets, this type of architecture was deemed unsuitable.

The alternative approach (i.e. thin client/server approach) is to request the required data from *LocateMe* and *DirectMe* and perform spatial queries on a dataset hosted on an external server. The parameters are passed to a web application, which in turn carries out the queries on a spatial database to determine any spatial interaction

between the users location and orientation and the dataset. The results are subsequently displayed to the user in the web browser on the mobile device. This architecture is illustrated in Figure 5.

This approach of using a web browser to display the returned data follows current trends in this area. The argument for using web browsers for as many on-device functions as possible is because of the increasing complexity associated with developing applications to run on many different devices [23]. To avoid developing a different application for each operating system, a more attractive solution is to develop applications using a web scripting language capable of producing highly interactive applications that perform the same functions as desktop applications without the concern about underlying operating system functions and restrictions.

Taking this approach a step further, it is now becoming more acceptable to let the browser access more different types of data as well. Using a set of JavaScript APIs, it is even possible to access some of the hardware on a mobile device directly from the browser. For example, this is one of the newest features of *Google Gears* [24] where the *Geolocation* API allows an application access to the GPS hardware on the device. Using this approach, if all sensor data was made available to a web application, it would be possible to eliminate the need for an on-device application altogether, essentially making the *TellMe* application entirely web-based.
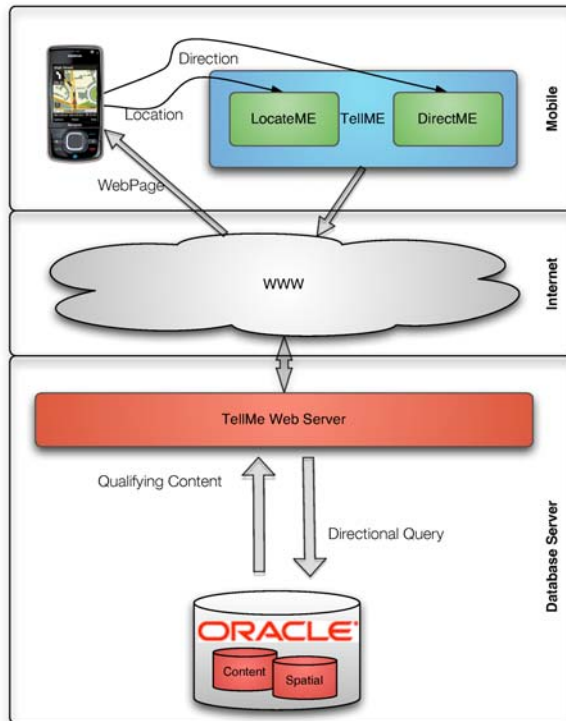

Fig. 5. Overall *TellMe* Architecture

4.1 LocateMe

The *LocateMe* module is used in conjunction with the *DirectMe* module to gather information about a user's current position and orientation. This data is then used to execute directional queries against various data sets both internal and external to the phone. The *LocateMe* module is based on a hybrid positioning system that utilises GSM, Wi-Fi and Bluetooth radio signals in addition to GPS to determine location. As this is not the focus of this paper, a more comprehensive description of this technology can be found in [25].

4.2 DirectMe

The *DirectMe* module is one of two modules that provide data to the *TellMe* application. Its function is to determine the direction that the mobile device is currently pointing by using a digital compass and tilt sensors. This data is then collected on request from the *TellMe* application and synchronised with the location data coming from the *LocateMe* module. The architecture of the *DirectMe* module is similar to the *LocateMe* module where each technology (i.e. - compass and tilt sensors) has a native hardware spotter that relays data to a higher-level component synchronises data from each spotter. This architecture is illustrated in Figure 6.
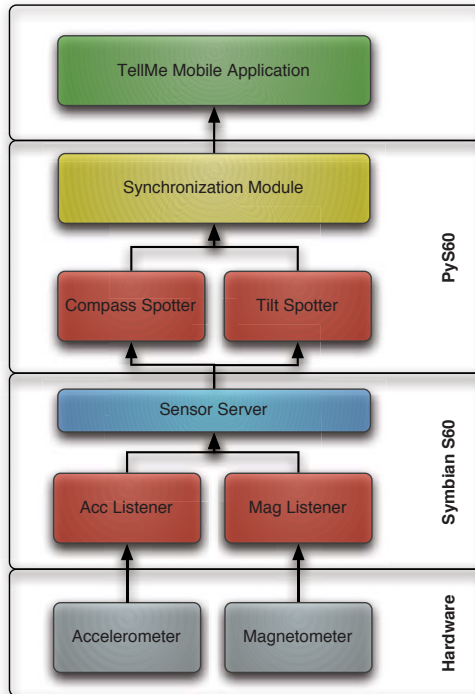

Fig. 6. DirectMe Design

This type of architecture is required in order to overcome some restrictions imposed in relation to access to various mobile device hardware components (e.g. - compass and tilt sensors) from particular APIs.

4.3 Mobile Device Hardware/Sensors

There are three main mobile platforms that currently provide mobile devices that contain the required sensor hardware for the *TellMe* system. The *Nokia 6210 Navigator* [26] has a digital compass and tilt sensors and the API providing access to these sensors has just (winter '09) been back ported from Symbian S60 5th Edition to Symbian S60 3rd Edition FP2. Using this version of the API, the *DirectMe* module can ascertain the current heading and orientation of the device. However, there have been recent reports that the performance of these integrated sensors is of poor quality [8]. Another hardware option is the *HTC Dream* (a.k.a. *Google phone*), with integrated digital compass and accelerometers. This device runs the Android operating system from *Google* and is currently not available in Ireland [27]. The most recent possibility in terms of hardware suitability is the *Apple iPhone 3GS* (Apple, 2009). In addition to providing WiFi and accelerometer access there is now a compass available making it suitable also. With this limited number of options available in relation to acquiring orientation data from mobile devices integrated or "on-board" sensors, another option is to use external sensors packs such as the *SHAKE*. The *SHAKE SK6* is an external sensor pack with digital compass, accelerometer, and tactile feedback. The even smaller SK7 has been released in Q1 2009. This sensor pack communicates with a mobile device via Bluetooth. Unlike current cellphone sensors, the *SHAKE* device has better quality sensors and a number of filters on board to reduce any noise introduced by the mobile phones antenna. The *SHAKE SK7* is shown in Figure 7.

However, following a review of these options, it was decided that the *DirectMe* module would collect data from a number of MEMS (Micro Electro-Mechanical Systems) sensors on a *Nokia 6210 Navigator* (Figure 7) in favor of the SHAKE as these sensors are integrated. In particular, the sensors that we use are magnetometers, which are capable of sensing the magnetic field surrounding the device. Using the magnetic field to determine magnetic north, it is possible to calculate compass bearing [8]. Furthermore, accelerometers are used to measure acceleration or the rate of change of velocity with respect to time.
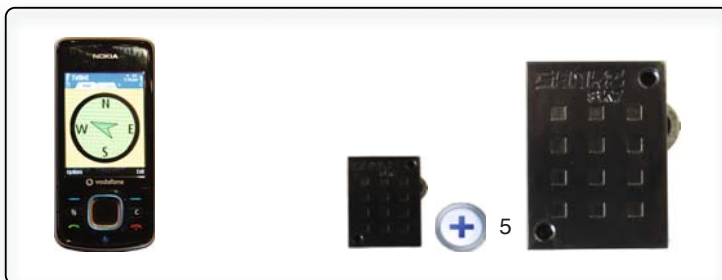

Fig. 7. Nokia 6210 Navigator and SHAKE SK7 sensor pack

4.4 TellMe Server

The TellMe server is essentially a spatial application server that is used to perform all the complex spatial queries in the system. It is responsible for communicating with the *TellMe Mobile* application, which collects data (including location, direction and orientation) from the *LocateMe* and *DirectMe* modules on the mobile device. This data is communicated wirelessly to the spatial application server and used to perform spatial queries against the *Oracle Spatial* 3D Database.

The *TellMe Server* is based on ESRIs *ArcGIS Server* platform and is used to perform the complex queries that are required to determine the mobile spatial interaction between the users line-of-sight and the 3d database. This platform provides server extensions for many of the traditional spatial query functions found in most GIS. Using these extensions, it is now possible to perform these types of spatial queries in a mobile context that were previously only possible in a desktop setting. Another important reason for this choice of server is based around the issue of scalability; using this software ensures that the system is scalable and can manage a large number of users efficiently.

4.5 3D Database

To perform directional queries, a 2D spatial database can be used and is the minimum requirement in order to do so. These databases support extensive 2D feature types and indexing techniques for performing spatial queries. 2D spatial queries can be performed effectively with a standard spatial database in an efficient manner and can identify objects (i.e. buildings) that intersect with a direction vector for example. Some types of queries possible are illustrated in Figures 10, 11 & 12.

However, to perform 3D spatial queries this process becomes somewhat more complex. In comparison, a 3D spatial query should be able, for example, to not only identify what building a direction vector is intersecting with but also the floor of the building it is directed at in a 3D Euclidean space. This means a true 3D database should support three-dimensional data types such as point, line, surface and volume in its geometric data model, be capable of indexing the data and must also offer functions and operations embedded in its spatial query language that can operate on these data types [28]. In fact, these requirements significantly reduce the number of options that are available to us in terms of being able to perform these types of queries. There are two main options here.

4.5.1 Oracle 11g

Beginning with the 11g version of Oracle it is now possible to store, index and query 3D objects using the *sdo_geometry* data type. Using this type, it is now possible to store point, line, polygon, polygon with hole and collection data types in 2D and 3D [28]. An example of the types of data that can be stored is illustrated in Figure 8 (b).

4.5.2 ESRI ArcGIS

To support this rising trend in 3D data storage, ESRI developed a native volumetric geometry feature type called the *Multipatch* feature supported by its geo-database

models that is treated like any other geometry type in the database. The *Multipatch* is constructed of triangle strips and fans and defines objects boundaries using triangular faces. This is shown in Figure 8(a)
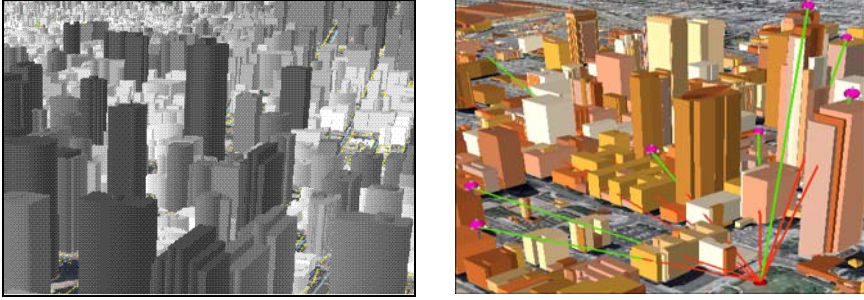

Fig. 8. (a) ArcGIS Multipatch 3d objects [21] (b) Oracle SDO_Geometry

In our case, we use Oracle 11g and its exclusively 3D forms: simple solid, composite surface, composite solid and collection to represent our data in the database. The data is comprised of 3d data that is based on Ordinance Survey Ireland's [29] 2d vector data that has extruded using height values from airborne LiDAR scans. The result is a block model of the NUI campus and Dublin city centre. There are also some detailed building furniture models of the NUI campus buildings, illustrated in Figure 8b. This data is spatially indexed and is queried using a 3D query window generated by the data collected from the mobile device sensors. All attribute data (building name, class timetables, etc.) presented to the user is also stored in the Oracle database. To perform 3D queries an extended set of operators can be used as support for 3D data is restricted to the SDO_Filter, SDO_Anyinteract, SDO_Within_Distance, and SDO_NN (nearest neighbour) operators using the Geographic-3D coordinate system. An example of a typical 3D query or frustum query is shown in Figure 9.
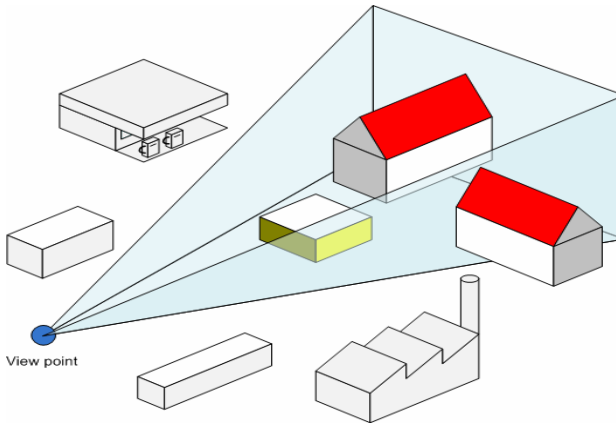

Fig. 9. Spatial query using view frustum

5 Ego-Visibility

In this section, the different types of queries that are performed by the *TellMe* application are discussed, all of which are based on the users egocentric point-of-view. Egocentric visibility refers to the portion of a search space that is visible to a user at a particular time based on their location, direction and orientation. For example, in the case of the *TellMe* application, the user's visible query space acts as a secondary filter on data that is returned by the query processor by restricting it to contain only objects that are in the users field-of-view (FOV). The FOV therefore excludes the portions of the dataset obscured by buildings. This method is primarily used to identify points-of-interest (POI) other than buildings that are in the users FOV within a predefined distance from the user and is illustrated in Figure 11. This method can be used to identify objects in the distance that may be too small to point at directly but are still in the users FOV nonetheless, like a monument or statue. The algorithm to determine the searchable space in this instance builds on previous work outlined in [17] where a 2D directional query processor was developed and used in a virtual environment for similar purposes. The following sections outline the queries that can be performed using the *TellMe* system.

5.1 Directional Query

Directional querying in terms of Mobile Spatial Interaction (MSI) is a method by which a device's position and orientation along 2-axis (horizontal and vertical) can be determined with the use of GPS and compass/tilt sensors. This data is then used to build a "query space" in the database to identify what object(s) the device is pointing at - in our case a campus building and any relevant information about it (Figure 10). When the user points a device at a building (creating a query vector), the device is able to identify what it is pointing at by determining the interaction between the query vector and the building data model stored in the database. The returned information can be communicated back to the user using visual, auditory or tactile interfaces.

Fig. 10. Identifying buildings using directional query vectors

This initial scenario is the simplest possible example of what we aimed to achieve in our research and it assumes that a users LOS is a straight-line vector. As this is not the case, we investigated other possibilities in terms of what the user can actually see. We look at ways of representing 2D and 3D query frustums that interact with objects based on LOS (i.e. direction) and visible open space (i.e. everything a user can see from a single point in all directions) or Egocentric Visibility (EgoViz).

5.2 Field-of-View Query

The field-of-view query maintains the use of the direction vector in the query process by producing a query frustum that closely represents a users actual field-of view. In contrast to the directional query, performing this type of query returns a list of results that identify not only what the user is pointing at but also everything in the users field-of-view (Figure 11). It is also possible to pre-select the layers that the user requires information about. For example, a user may only want information about the buildings that are in their field-of-view or alternatively they may want information about points-of-interest only. This categorisation of the required data helps to speed up the query process by eliminating sets of data not required.

Taking this concept a step further, the idea of the visibility polygon or Isovist is introduced. A visibility polygon is the portion of open space a user can see in all directions. Being able to determine a users visibility polygon in real-time as they navigate through the campus or city streets enables the *TellMe* system to deliver a much richer experience in terms of the relevance of the data to the user and the speed at which the data is delivered. As the visibility polygon identifies the area the user can actually see, this generates a much smaller search space reducing the size and complexity of the spatial queries to be performed.

Fig. 11. Identifying POI objects in a user's field-of-view

5.3 Isovist Query

Recent work in this area by [18] describes the local visibility model (Lvis) that uses the concept of billboards to determine what buildings are in the users FOV in a 2.5D

environment. To achieve this in our case, a different approach is taken based on work carried out by [20] into the concept of Isovists and medial axes. Using Isovists, we attempt to automatically generate a panoramic (360º) line-of-sight search space for performing spatial queries on. This method should prove to be very effective as the fundamental aim is firstly to determine precisely what area in geographical terms is visible to the user in all directions and secondly to determine what objects inside this area is of interest to the user. This idea is illustrated in Figure 12.

Fig. 12. Egocentric Visibility using IsovistExplorer 360°

5.4 Frustum Query

Extending the 2D Isovist to work in 3D space is a primary objective of this work. In the case of the frustum query, information about relevant artifacts that lie within the sections of open space (i.e. the space between the buildings in all directions up/down/ sideways) that has been identified as visible to the user are queried and returned.

Fig. 13. 3D Visibility Query
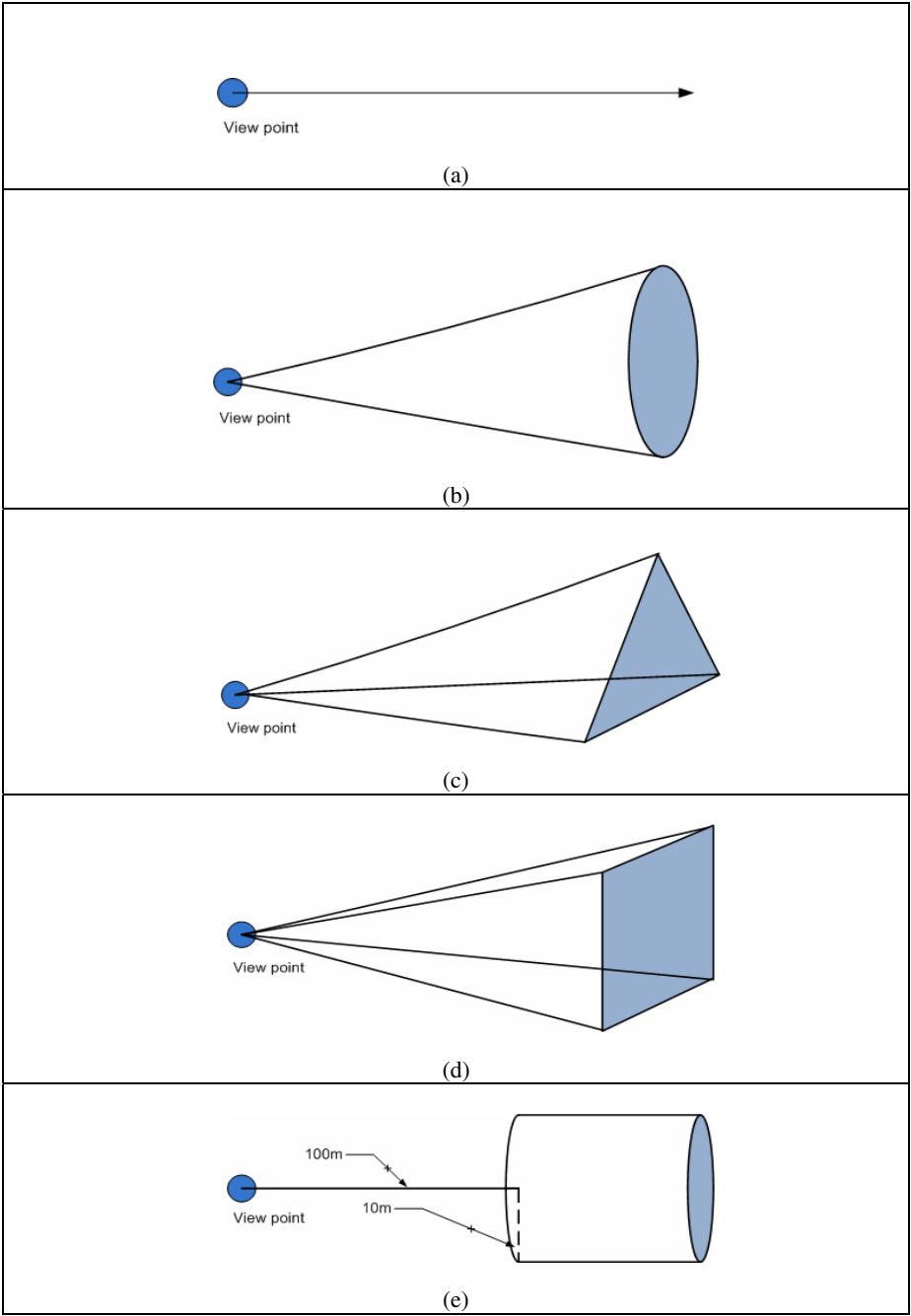
(a)

(b)

(c)

(d)

(e)

Fig. 14. Examples of 3D Visibility Queries

In this case, the second generation of the 3D Isovist is represented by a 3D object that represents a users actual LOS omitting the geometry of any other objects that interact with it (illustrated in Figure 13). Building this type of 3D Isovist query requires an extensive 3D database of buildings and building furniture at floor and room detail in order to utilise the query frustum geometry. Examples of the types of geometries used are illustrated in Figure 14.

5.5 Ego-Dome

The Ego-Dome is essentially an extension of the 2D Isovist described above except the visible open space of the user from a particular location is reproduced in 3D. This concept of querying all open space in this manner has been explored previously but usually only using standalone desktop applications to do so, offline. In the case of the *TellMe* system, the *Ego-Dome* is created using the current location of the user (Figure 15). Any objects that interact with the dome are detected and the required results are presented to the user using visual, auditory or tactile interfaces. This type of spatial interaction is very useful for a host of applications in terms of real-time feedback that gives the user the ability to interact and have knowledge about an environment in real-time.
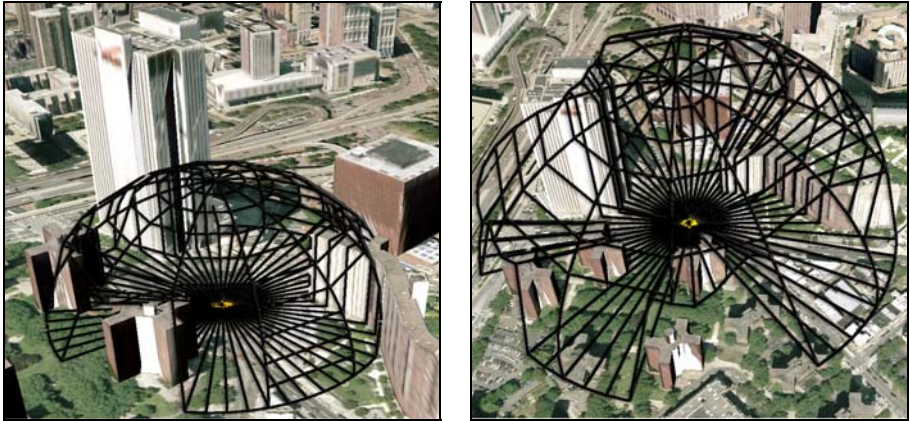

Fig. 15. Ego-dome being used to determine the visible space within a certain radius

6 Conclusions

This mobile based spatial interaction system (Egoviz), which is based on the *TellMe* application is designed to allow users to interact in a context sensitive way with information using current mobile phone technology. Current devices are beginning to offer hardware such as digital compasses and tilt sensors. With the use of these technologies, a wide spectrum of applications is emerging. The development of the *DirectMe* module and its synchronisation with the *LocateMe* module is ongoing. Concurrently, the development of the complete *TellMe* mobile application is taking place and provides a framework for testing our directional query techniques.

In relation to performing directional queries, there are a number of fundamental issues that have been highlighted in terms of the quality of the data that is retrieved from the sensors. In [10] the quality of the sensor data from a custom built sensor pack is analysed and compared to actual data showing good results for performing directional queries. With the quality of the sensor data being a pivotal aspect of directional queries in terms of the relevance of the data returned and the move towards making this type of interaction available on mainstream mobile devices, we intend to carry out two important sets of tests. Firstly we investigated the quality of selected external sensors (i.e. SHAKE and Nokia LD-4W GPS) and secondly we will compare these results with the quality of the data from the integrated sensors in the "off-the-shelf" phone (i.e. Nokia 6210 Navigator).

Regarding egocentric visibility, research will be carried out to help us better identify what comprises exactly a user's FOV. Initially we have studied work by [20] in the area of spatial planning and in particular Isovists by developing a 2D EgoVis filter to determine a users true LoS in 360°. In due course, the technical feasibility of extending to a true 3D directional query processor will be investigated using the idea of the threat dome as an example.

Acknowledgments

References

1. Simon, R., Frohlich, P.: The Point to Discover GeoWand. In: 9th International Conference on Ubiquitous Computing (UbiComp 2007), Innsbruck, Austria (2007)
2. Kilfeather, E., Carswell, J., Gardiner, K., Rooney, S.: Urban Location Based Services using Mobile Clients: The ICiNG Approach. GISRUK, Maynooth (2007)
3. Rooney, S., Gardiner, K., Carswell, J.: An Open Source Approach To Wireless Positioning Techniques. In: The 5th International Symposium on Mobile Mapping Technology (MMT 2007), Padua, Italy. (2007)
4. Wilson, A., Pham, H.: Pointing in Intelligent Environments with the WorldCursor. Interact (2003)
5. Wilson, A., Shafer, S.: XWand: UI for Intelligent Spaces. CHI, Fort Lauderdale, Florida (2003)
6. Chan, L.-W., Hsu, Y.-Y., Hung, Y.-P., Hsu, J.Y.-J.: Orientation-Aware Handhelds for Panorama-Based Museum Guiding System. In: UbiComp 2005 Workshop: Smart Environments and thier Applications to Cultural Heritage (2005)
7. Baillie, L., Kunczier, H., Anegg, H.: Rolling, Rotating and Imagining in a Virtual Mobile World. In: 7th international conference on Human computer interaction with mobile devices & services, Salzburg, Austria (2005)
8. Essl, G., Rohs, M.: ShaMus - A Sensor-Based Integrated Mobile Phone Instrument. In: Proceedings of the International Computer Music Conference (ICMC), Copenhagen (2007)

9. Williamson, J., Murray-Smith, R., Hughes, S.: Shoogle: Excitatory Multimodal Interaction on Mobile Devices. In: Proceedings of ACM SIG CHI Conference, San Jose (2007)
10. Simon, R., Frohlich, P.: GeoPointing: Evaluating the Performance of Orientation-Aware Location-Based Interaction Under Real-World Conditions. In: 4th International Conference on LBS and TeleCartography. Taylor & Francis, Hong Kong (2008)
11. Intelligent Spatial Technologies (2009),
    `http://www.i-spatialtech.com/ipointer.html`
12. Robinson, S., Eslambolchilar, P., Jones, M.: Point-to-GeoBlog: Gestures and Sensors to Support User Generated Content Creation. In: 10th international conference on Human computer interaction with mobile devices and services, Amsterdam, The Netherlands (2008)
13. Frohlich, P., Simon, R., Baillie, L., Anegg, H.: Comparing Conceptual Designs for Mobile Access to Geo-Spatial Information. In: 8th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI 2006), Helsinki, Finland (2006)
14. Persson, P., Espinoza, F., Fagerberg, P., Sandin, A., Cöster, R.: GeoNotes: A Location-based Information System for Public Spaces. Readings in Social Navigation of Information Space, 151–173 (2002)
15. Strachan, S., Williamson, J., Murray-Smith, R.: Show me the way to monte carlo: density-based trajectory navigation. In: Proceedings of the FP 205 SIGCHI conference on Human factors in computing systems, New York, USA, pp. 1245–1248 (2007)
16. Strachan, S., Murray-Smith, R.: Bearing-based selection in mobile spatial interaction. Personal and Ubiquitous Computing 13 (2009)
17. Gardiner, K., Carswell, J.: Viewer-Based Directional Querying for Mobile Applications. In: Third International Workshop on Web and Wireless Geographical Information Systems W2GIS, Rome, Italy (2003)
18. Simon, R., Frohlich, P.: A Mobile Application Framework for the Geospatial Web. In: Proceedings of the 16th international conference on World Wide Web, Banff, Alberta, Canada (2007)
19. Benedikt, M.L.: To take hold of space: isovists and isovist fields. Environment and Planning B 6, 47–65 (1979)
20. Jiang, B., Liu, X.: AxialGen: a research prototype for automatically generating the axial map. In:11th International Conference on Computers in Urban Planning and Urban Management, Hong Kong (submitted, 2009)
21. ESRI Globe (2009), `http://www.esri.com/products/index.html#4`
22. Skyline Terraserver (2009),
    `http://www.skylinesoft.com/skylineglobe/corporate/Products/`
    `TerraExplorer.aspx`
23. Mobile Web Application Best Practices (2008),
    `http://www.w3.org/TR/2008/WD-mwabp-20080729/`
24. Google Gears API (2008),
    `http://code.google.com/apis/gears/api_geolocation.html`
25. Rooney, S., Gardiner, K., Carswell, J.: Wireless Positioning Techniques – A Developers Update. In: Ware, J.M., Taylor, G.E. (eds.) W2GIS 2007. LNCS, vol. 4857, pp. 162–174. Springer, Heidelberg (2007)
26. Sense Your Location (2008), `http://www.nokia.com/A41229034`
27. HTC - Dream (2008), `http://htcdream.com/`
28. Schön, B., Laefer, D.F., Morrish, S.W., Bertolotto, M.: Three-Dimensional Spatial Information Systems: State of the Art Review. Recent Patents on Computer Science 2, 21–31 (2009)
29. Osi Website (2009), `http://www.osi.ie`

5.5. Paper V: A Web-Based and Mobile Environmental Management System

Carswell, J. D., K. Gardiner, A. Rizzini, M. Bertolotto, and N. Mandrak. (2008) "A Web-Based and Mobile Environmental Management System". Journal of Environmental Informatics 12(1) 9-20, DOI: 10.3808/jei.200800119.

A Web-Based and Mobile Environmental Management System

J. D. Carswell[1,*], K. Gardiner[1], M. Bertolotto[2], A. Rizzini[2], and N. Mandrak[3]

[1]*Digital Media Centre, Dublin Institute of Technology, Ireland*
[2]*School of Computer Science and Informatics, University College Dublin, Ireland*
[3]*Dept. of Fisheries and Oceans, Canada Centre for Inland Waters, Burlington, Ontario, Canada*

ABSTRACT. This paper describes a Web-based and mobile system specifically developed to monitor fish species at risk. Such a system integrates spatial functionality to allow users not only to visualise maps and metadata of the area of concern but also to perform context-aware queries and updating of spatial datasets. The spatial datasets are provided by the Canadian Department of Fisheries and Oceans (DFO) and the prototype is customised to the specific needs of the Great Lakes Laboratory for Fisheries and Aquatic Sciences (GLLFAS) Fish Habitat Section requirements for fish species at risk assessment. Currently, researchers, habitat biologists and enforcement officers have access to the fisheries database, containing layers of biological information solely from the office. Delivering these data overlaid on base maps of the Great Lakes region to a GPS-enabled hand-held device and linking it to each task currently being investigated allows for mobile DFO biologists and enforcement officers in the field to make informed decisions immediately. In this paper we describe the system and demonstrate how it is used by the DFO in practice.

Keywords: spatially enabled mobile computing, web-based GIS, fish species at risk

1. Introduction

The diffusion of the Internet and mobile technologies has allowed for the development of many applications which can be accessible anytime and anywhere (Hinze et al., 2003; Schiller, 2000; Talukder et al., 2006). Many such applications can nowadays provide information based on the context in which they are used, including their geographic context. The advantages of these technologies are especially important for those professions that involve field work. Many such professions require the manipulation of spatial datasets, including maps of the areas visited. Examples of professional users that avail of spatial data during their field work are environmental scientists, surveyors, geographers.

This paper describes the development of a Web-based and Mobile Environmental Management System (MEMS) tailored to deliver context-aware functionality aided by visualization, analysis and manipulation of spatial and attribute datasets. The MEMS datasets are provided by the Canadian Department of Fisheries and Oceans (DFO) and the prototype is customized to the specific needs of the Great Lakes Laboratory for Fisheries and Aquatic Sciences (GLLFAS) Fish Habitat Management Group requirements for fish species at risk

assessment. Currently, biologists have only access to the fisheries data from their office. This greatly prevents them from interacting with the data in a real-time environment, reducing their productivity and effectiveness in the field. Spatially enabling a mobile device allows mobile GLLFAS biologists to make informed decision immediately.

This research concerns DFO priorities specifically to administer the fish habitat provision of the Fisheries Act, in particular those that are aimed at preventing the harmful alteration, disruption or destruction of fish habitat. This is done to conserve, restore and develop the productive capacity of habitats for recreational, commercial and subsistence fisheries both in the freshwater and marine environments (Minns, 1997, 2001). The functionality required by GLLFAS biologists includes access to geo-referenced maps and imagery, to overlay the current position on a map and to manipulate (e.g. input/ edit/query) attribute data in the field while wirelessly connected (where possible) to the office database. Additional functionality also required is the ability to record, edit and view multimedia annotations, perform scientific/common name conversion and graph generations of results. The traditional "fish species at risk" work-flow, whereby scientists enter textual/ pictorial information on paper field data sheets is inefficient, has potential for inaccuracies during both initial recording and subsequent data entry phases, and does not facilitate knowledge-sharing between staff. Also, different types of information may be stored in different locations and valuable time can often be lost trying to correlate data in order to make decisions. The proposed MEMS system has the following advan-

tages over such practice:

- Facilitates knowledge-sharing and data analysis/ synthesis;
- Supports effective communication between different staff at different physical locations (e.g. scientists in the laboratory or office, and colleagues in the field);
- Allows important multimedia data and associated annotations to be combined with text-based records;
- Standard web-browser interface allows staff to input and access information anywhere at any time without having to return to dedicated access points (i.e. in the office);
- Reduces errors by eliminating time elapsed between data collection and data entry, saving time as well as paperwork.

The contribution of our work is the development of a system that facilitates the work of the GLLFAS biologists and makes their collected data more reliable and accessible. Such a system has been developed using the most recent GIS and mobile technologies and therefore provides a high level of functionality not just for collection and manipulation but also for spatial analysis of the data that was not available to these scientists before. This powerful system demonstrates the potential of integrating spatial functionality within environmental applications.

This paper describes the development of the MEMS system. In Section 2 the system architecture and the development environment are described. The system relies on the Oracle Spatial Database Management System that guarantees interoperability at the data level (Rigaux et al., 2002). Section 3 details the system requirements, while Section 4 discusses the functionality implemented with practical examples of its use through the Graphical User Interface (GUI). In particular, we describe the development of both the Web-based and mobile functionality. An on-line and an off-line version of the mobile system have been developed to allow for use in areas where wireless coverage is not available. Section 5 concludes with final remarks and future research directions.

2. System Architecture and Development Environment

The MEMS prototype relies on a typical three-tier architecture for enterprise information systems, composed of the client layer, application server layer, and the database layer (Figure 1).

This architecture focuses on the development of services for a versatile, extendible (J2EE) application server, instead of giving GIS capabilities to a large monolithic application. The communication between the client layer and the database are conducted through the application server layer. With this type of architecture, the processing load is balanced, as each tier of the system resides on a separate computer (Carswell et al., 2004; Rizzini et al., 2006).

The MEMS prototypes were developed using eSpatial's

iSmart Suite application development environment (eSpatial Solutions, 2004). This software is a collection of tools that enables developers to build and deploy spatial applications using sets of standard procedures. It offers developers a high-level development environment which is several times faster than developing the application from base Java. It reduces the need for major software engineering prerequisites that developers typically must have in order to understand and deploy a full spatial application. For the implementation and deployment parts of the system, *iSmart Eclipse* was used along with *iSmart Application Sever*.
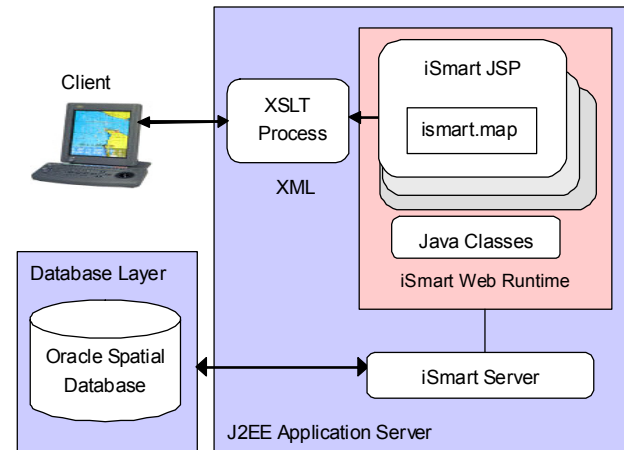

Figure 1. Three-tier architecture.

iSmart technology adopts the strategy of focusing the services on the application server layer of the system. This architecture allows for the development of individual components of the system separately, thus maintaining component independence. In this way, different parts of the system can be developed at different stages, some more than others, without affecting the entire system each time a change is made. For example, this architecture has proven ideal for developing Extensible Markup Language (XML)-based applications because all XML/XSLT processing is carried out on the middle tier of the system, without affecting client and/or database tier manipulation or development (Figure 2). The following sections describe each level of the three-tier architecture, and includes a brief description of the technologies that are required for each layer.

2.1. Client Layer

The client layer of the system consists of a Web-based (i.e. standard web-browser) interface, that enables data-input and analysis functionality to provide biologists with the ability to input, edit and annotate data over the Internet. As the user navigates within the system, the position and orientation of the user is displayed on a geo-referenced map of the area. This information is used when loading data into the database and for contextually querying the database in real-time. Client hardware consists of a Tablet PC or Personal Digital Assistant (PDA) equipped with a Global Positioning System (GPS) re-

ceiver to determine the position of the device. The device also has a General Packet Radio Service (GPRS) network connection for requesting and transmitting data to and from the application server over a wireless phone network. All mobile aspects of the system require a streamlined design, as current speed and bandwidth capabilities of hand-held devices are not yet comparable to those of desktop PCs. Therefore, to deliver content efficiently to these devices it is scaled down using a device-specific content adaptation approach.
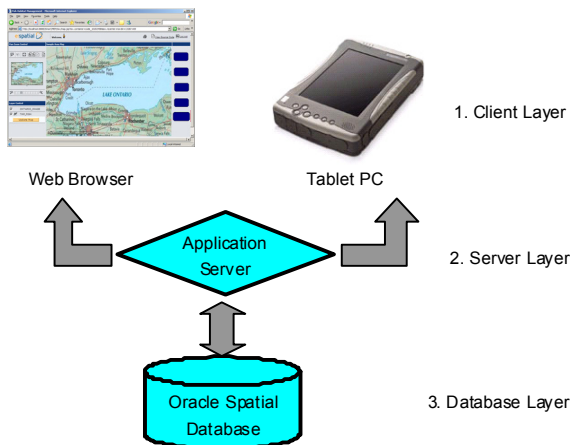
Figure 2. MEMS architecture.

2.2. Application Server Layer

The application server in Figure 2 acts as the main hub between the client and database layers of the MEMS system and consists of a J2EE Application Server made up of two main components - the *iSmartWeb Runtime* and the *iSmart Server*. These two components are used to run applications that have been developed using the *iSmart Web Designer*.

The *iSmart Web Runtime* is a component integrated with the *iSmart Server* that executes the web application. When a web page is requested, this component executes the corresponding JSP pages and returns a HTML page back to the client. Within this component, there is a Servlet called the *iSMARTWebServlet*. This Servlet acts as an interceptor to all incoming HTTP requests and uses control logic specified by the developer to decide what actions to take when, for example, a button is pressed on the application web page.

The *iSmart Server* contains all the functions and controls in the application server layer and handles all requests made to the application server automatically with the help of an XML file called *isw_application.xml*. This file is created at deployment time and is used to process requests to the application. This layer of the system is also responsible for determining device-specific capabilities. Using XML and XSLT, the *iSmart Server* adapts all returned content based on the type of device that makes the initial request. This is achieved by creating an XML file based on the client device request. An existing XSLT stylesheet is then applied to the XML file to create the device specific output (Figure 3).

2.3. Database Layer

The spatial database layer of MEMS is responsible for processing all queries, both spatial and transactional, in the system. Oracle Spatial is used as the database platform and includes the Spatial Data Option (SDO), a spatial extension to SQL introducing new spatial data capabilities, e.g. geocoding and topological queries.

Oracle Spatial provides a platform that supports a wide range of applications from automated mapping/facilities management and Geographic Information Systems (GIS), to wireless location services and location-enabled e-business. Oracle Spatial is integrated into the extensible Object Relational Database Management System (ORDBMS), which allows access to the full functionality and security of the underlying DBMS (Lopez, 2003; Sharma, 2001). Along with the database, the Oracle proprietary OC4J application server is used. The application server is a component of Oracle, and is installed automatically. The application server acts as deployment platform for Oracle applications.

This tier stores all spatial and non-spatial data together, including raster (map/remotely sensed image) data and any metadata as well as the topological properties of these data. It enables Spatial Data Types (SDT) to be inserted, stored, manipulated and queried in the database as they are represented in physical space.
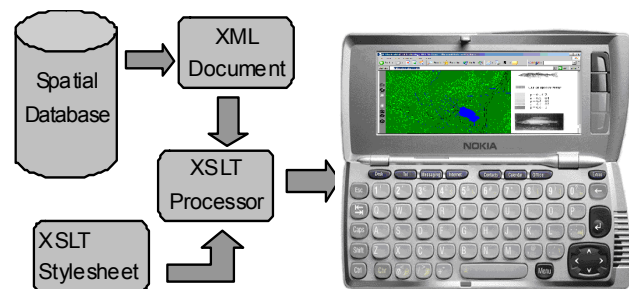

Figure 3. Device adaptation workflow using XML/XSLT.

3. System Requirements

The traditional workflow of the biologists of the DFO includes different phases, i.e. the collection phase, the input phase, analysis phase. The collection phase entails recording the data on paper in the field. Subsequently, at a later stage, the data is inserted manually into the database residing in the office. This process is time consuming and can cause inconsistencies in the data inserted due to human interpretations and typos (e.g. "n.a", "N/A" or "St.", "Saint"). In addition, the data is usually not validated, thus allowing these errors to be entered into the database.

During a field trip, biologists obtain samples from fish caught and then record the data collected. For example, they retrieve a caught fish from the storage tank in the vessel, identify it, measure the size, and record the number caught for

that particular species. This inevitability causes the biologist to navigate through many forms, editing some, until all fish are recorded.

In order to identify their requirements, there was extensive interactions with the targeted end users of our system (the biologists of the DFO). This meant participating on DFO field trips and observing their work, as well as trying to identify any additional functionality that could facilitate their work. We describe the MEMS functional requirements in the following sections both for the web-based version and for the mobile version.

3.1. Web-Based Prototype

The functional requirements for the web-based prototype are outlined based on the layer of the system with which they are associated.

Client Layer

- A real-time, geo-referenced mapping interface that enables the user to view, input and modify data as well as to annotate maps and photos in a variety of ways. The option to add, delete and query annotations will be available;
- An advanced query interface to enable the user to query the database in a number of ways (e.g., through frequently-used static queries along with the functionality to create and store additional queries);
- A query interface that enables the user to create com-

plex queries using a series of drop-down menus and also the facility for more experienced users to enter SQL to query the database directly;
- A context-aware MBR query that enables the user to query using a selected area.

Server Layer

- The ability to process form-based data and map data into the required format for use in the client layer;
- The ability to transform coordinate data from a GPS receiver into a format that can be utilized by the application;
- The ability to construct and execute all queries in the system based on the users location and requirements;
- An annotation processor that manages all annotations in the system. This is advanced functionality that requires an individual processing module.

Database Layer

- A multimedia component management module that stores and processes all the multimedia components in the database.

3.2. Mobile-Based Prototype

The following is a list of the functional requirements of the mobile prototype based on the layer with which they are associated. Note that the database layer is shared by the two versions of the system.
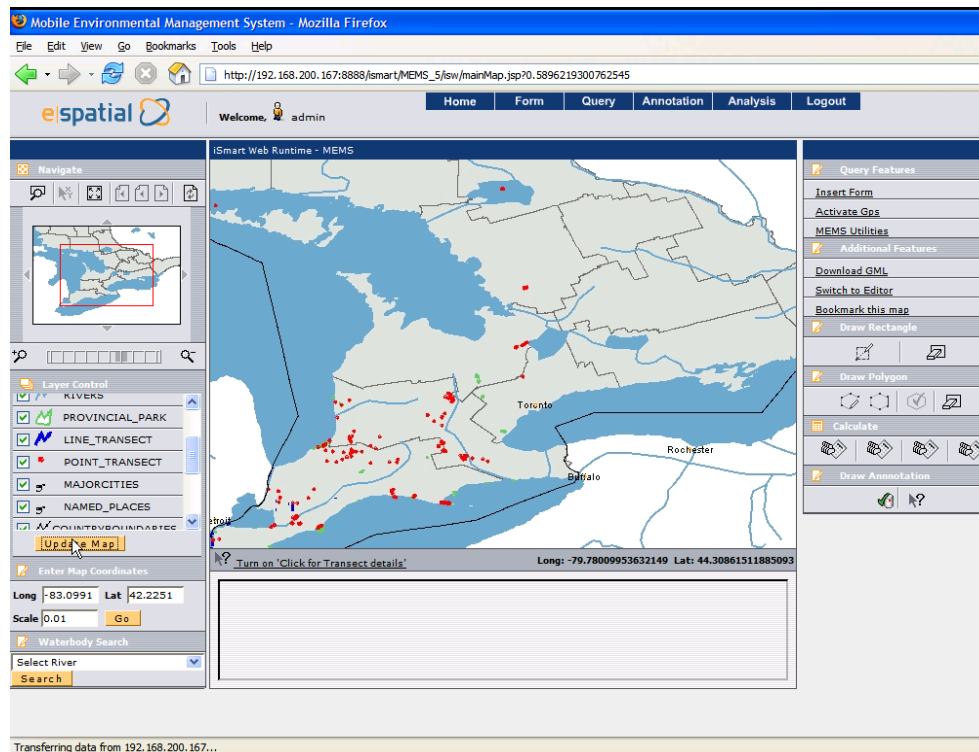

Figure 4. MEMS GUI.

Figure 5. Data display form.

Client Layer

- A Tablet PC or PDA, used as the client; GPRS mobile phone network connection for transmitting data where a wireless connection is possible;
- A GPS receiver connected to the mobile device to determine the current position of the device.

Server Layer

- The ability to process data in online and offline modes, depending on network coverage;
- The ability to transmit data over a GPRS network to a central server;
- The ability to determine the position of the mobile device and deliver it to the application.

4. System Functionality

In this section we describe the implemented system functionality with practical examples of its use through the Graphical User Interface (GUI). The MEMS GUI is divided into three main parts (Figure 4).

Thee left-most panel is the Navigation panel. It contains the functions used to navigate the map (zoom to area, click x/y coordinates, zoom to extent, back to original view, previous view, refresh map). The navigation map contains a red rectangle which represents the area displayed by the main map. This is designed to help the user navigate more effectively when zoomed into a small area on a map. The Layer

Control is located below the navigation panel and is used to select layers to display on the main map. The last two controls on the navigation panel contain Map Coordinates, used to centre the main map with any given coordinates and the Search function control, used for searching previously recorded habitats.

In the centre of the application the Map Panel contains the main map. The main map offers clickable map functionality that enables the user to select features on the map and displays attribute information about them. The right-most Tools panel contains controls for functions which are explained in the following sections.

4.1. Forms Module

An important objective of our work is to minimise the work for the biologist in the office in or order to maximize the hours spent recording new data in the field. To this end, the forms module of the system enables the biologists to view and input data directly while in the field thus reducing the time spent post-processing field data in the office.

Specifically, the forms module enables the biologists to display, insert and edit data using dynamic forms at different stages of the filed sampling process. Validation constraints were implemented so that incorrect data is not inserted into the database (for example, drop down lists are used). Biologists can also semi-fill forms and revisit them later in the process to make changes. Figure 5 shows the Data Display Form as an example.

4.2. Utilities Module

This module hosts the following functionality.

- Name Lookup: biologists can identify different types of fish species with their common name, but they are required to record their scientific name in the form. In order to help the biologist remember the scientific names of 269 different fish, this utility offers a quick translation mechanism. The user can select the common name from a drop down menu and instantly the biologist obtains the scientific name (Figure 6).

- Transect Tool: used in order to calculate the distance that a biologist needs to travel in a particular river or water body in order to achieve a satisfactory representative sampling result.

- Accumulation Curve Tool: a graphical generation tool to display the accumulation curve for biologists. The accumulation curve is used by biologists to calculate and display how many fish have been caught for a given set of transects and for a particular species (Figure 7).
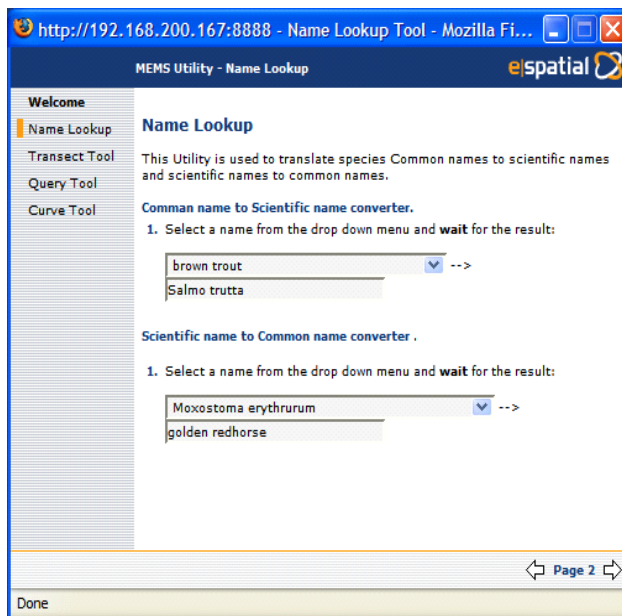
Figure 6. Name lookup.

The Accumulation Curve tool dynamically generates two possible types of graph. The bar chart is displayed by clicking the first icon on the left of the "*graph toolbar*". There is also a line chart generated by clicking on the second icon on the left of the "*graph toolbar*". The types of graph are illustrated in Figure 8.

4.3. Spatial Queries Tool

The Spatial Queries Tool is an advanced querying tool which enables biologists to quickly and easily query the database without requiring them to have any understanding of Oracle Spatial or the Sequential Query Language (SQL).

Users draw a polygonal area on the map on where it is then possible to execute a number of predefined spatial queries on the selected area. The polygon highlighted on the map is the actual spatial component and the highlighted buttons are the predefined queries requested by the GLLFAS biologists. The procedure starts by drawing a polygon or a rectangle on the map. The dots on the map represent locations where the biologists have previously recorded. The result of any given query is parsed and displayed in a pop-up window (Figure 9).
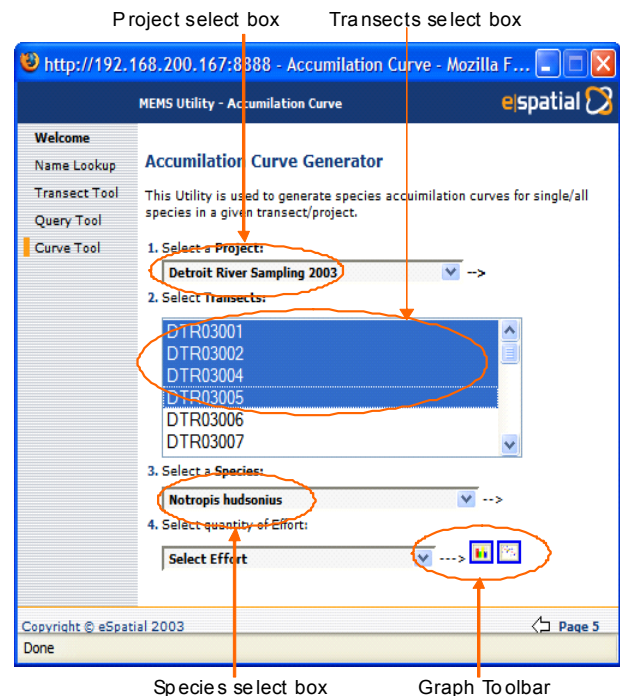

Figure 7. Accumulation curve tool.

Spatial queries are a particular type of query that use location as their primary search parameter. In the MEMS system, Oracle stores non-spatial data relating to a spatial object such as the MNRDistrict and Province of the spatial location (e.g. spawning sites). Although the application will offer advanced non-spatial query functionality, spatial queries can often give a clearer and more simplistic answer to question. Spatial queries are executed by executing Oracle SQL queries against the database and the result is returned as a hybrid ResultSet Java object called GeoResultSet. The GeoResultSet is an iSmart implementation of the ResultSet, which is capable of storing both the spatial and non-spatial attribute ResultSets. The iSmart application can subsequently extract the information from the ResultSet and displays it in a simple visual format. Currently, in the MEMS prototype, there are three main spatial queries possible. The Range or Buffer query, the Locate on Map query and the Clickable Map query. In this section a brief description of each is given.

4.3.1. Clickable Map Query

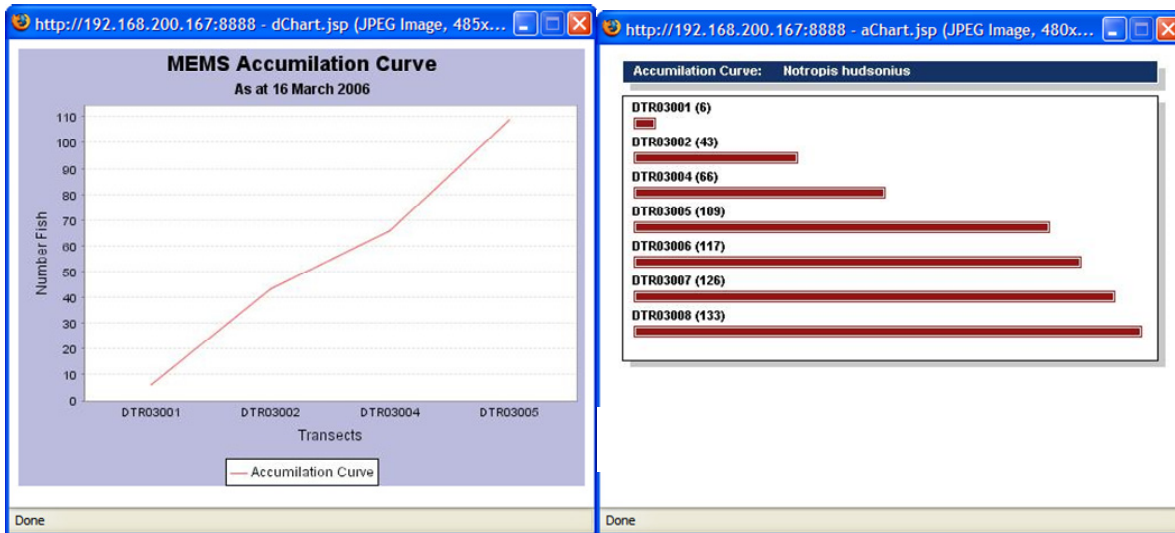The *Clickable Map* is an extension of the normal map-

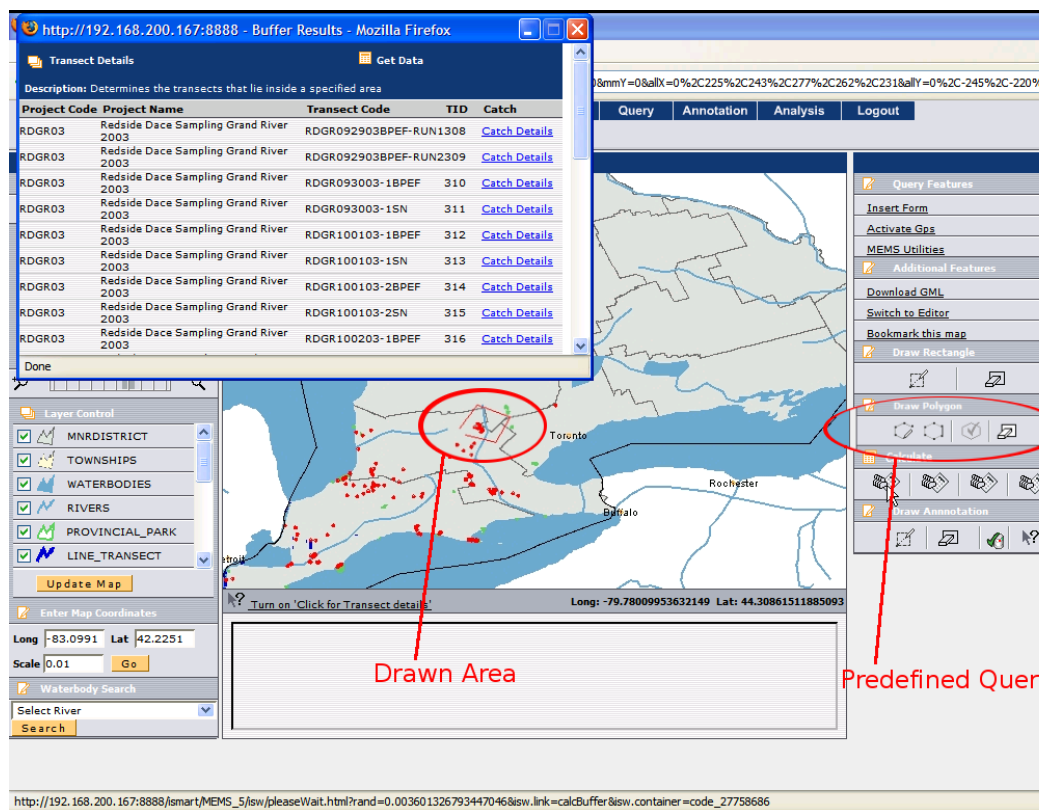**Figure 8**. Graphs generated by the accumulation curve tool.

Figure 9. Spatial queries tool.

ping interface, where the particular layer to be queried needs to be specified. The purpose of this query is to enable the *DFO* biologists in the field to click on objects on the map interface and extract information about the object and display it on the device. This helps the biologists to get familiar with the surrounding environment and quickly target areas of interest. The spatial query is implemented using *JavaScript, JSTL* (JSP Standard Tag Library), and the *iSmartWeb*. The map, which is displayed on the web application, is a *JPEG* plain image and is refreshed each time the user request additional
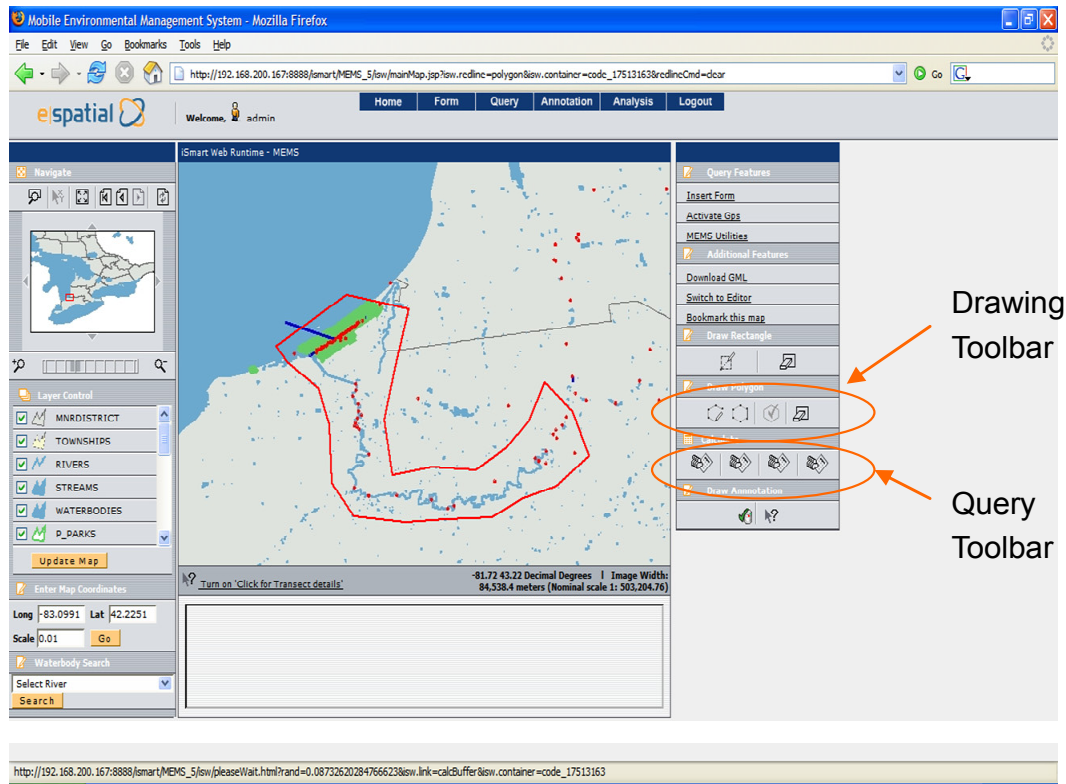
Figure 10. Drawn query.

information.

4.3.2. Locate on Map Query

The Locate on Map query enables the user to locate and display a specific transect on the map. The query consists of two parts. The first part is a text-based search where the user searches for a specific water body. The results of which are displayed with a link to the location of the water body on the map and a display link which displays more information about the transect. The hyperlinks are linked to iSmart actions, which carry out a query when selected. The locate action is the second part of the query and the result of this query centres the Viewport on the selected transect. This query is particularly useful to biologists in the field when looking for a particular site in a water body. In this way, the biologists can locate the transect, and from analysing other details about the position, determine if a sample session is required for the area or not.
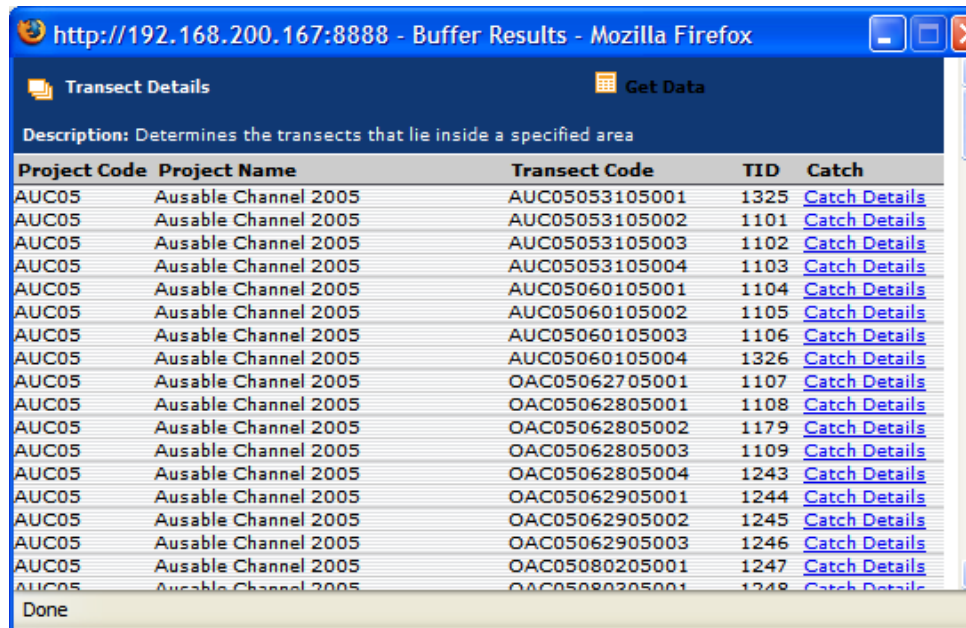
4.3.3. Range Query

Up to this point, the types of queries explained are not spatial queries by definition. They are presented as spatial queries because they interact with the applications spatial component but in theory they are not spatial because the query itself does not actually interrogate the spatial component. The Range Query is a common spatial query and is extremely useful and will enable the biologists to interrogate the

database to gather information about sites within a given area. Once the Clickable Map is activated the user is given the option to search the surrounding area for sites. The user specifies a distance in meters from their current location and the query is performed. The result returned is a list of sites that are included in this area.

4.3.4. Drawn Query

This type of query is required when wanting to quickly select a custom area of interest and apply a series of predefined queries. The selection of the area is achieved by drawing the area out on the Map Interface. Once the area of interest is created a series of "Spatial queries" can be executed against it. The results will be parsed and displayed in appropriate windows within the web application. This module offers a significant advantage over the other query modules: the ability to effectively draw the area in which the user wants the query to be performed. If the query was to be processed without being able to draw the area of interest, the user would need to know the coordinates which make up the area of interest and feed them into the spatial query manually, making the query very impractical.

Looking at Figure 10, this module is instantiated by selecting the drawing tool which corresponds to the first button on the left of the "Drawing Toolbar". The drawing tool uses a complex JavaScript in order to draw and record the points which constitute the polygon. Once the user is satisfied with

Figure 11. Popup result from the drawn query.

the polygon, s/he must press the accept button (third button from the left) on the "Drawing Toolbar". The accept button is used to store the geometry of the polygon in the user session. The process starts by reading the points collected in the JavaScript and passing them to a server side component which creates an Oracle Geometry Object called "JGeometry".

In order to execute the query independently of any knowledge of SQL programming, a series of spatial queries required by the GLLFAS where statically embedded in the application. These queries are shown in the "Query Toolbar" in Figure 10. One example of the query is to return the "Project Name", "Transect Code" and "TID" for all the "Transects" in the drawn polygon. A popup window contains the results of the query and for each entry in the form there are links that generate further queries. An example of the popup windows displayed by this function is displayed in Figure 11.

4.4. GPS Module

This is a core module as the distributed MEMS application forces the client to acquire GPS coordinates first and then send them to the application server. As the client was not designed to be installation independent it is necessary to execute the Java application remotely.

One of the technologies that Java offers is JNLP (Java Network Launching Protocol). JNLP applications are launched using Java Web Start as part of the Java Runtime Environment. The corresponding GPS application registers each client IP address and host-name with the application server so that multiple feeds can be displayed on the map. The screenshot in Figure 12 shows the GPS acquisition overlayed on the MEMS map. In the top left corner the GPS module control is shown and is launched by the java Web Start while in the centre the black square is a zoomed image of the GPS feed.

This module enables the user to record location data and store it as a spatial component in the database. Once activated the GPS module opens a "GPS Toolbar" and a "GPS Application". This Java application is used to start and stop the GPS receiver. When the GPS receiver has locked onto a steady signal the "GPS Application" displays the current "Latitude" and "Longitude" position, a green light indicating the status and also the signal strength which is the accuracy of the position in meters.

Subsequently, the current location is displayed on the map interface and is symbolized by an arrow pointing in the direction the user is moving and a trail showing where the user has been. On the tool bar the user has two buttons, the start and stop recording buttons. These buttons are used to save the start and stop position of a transect and the start and stop time for the duration of a particular sampling session (see Figure 12).

4.5. Multimedia Annotation

Multimedia annotation is another advanced feature of MEMS that allows data to be recorded on-the-fly. For example, a new species of fish could be encountered and visual evidence would be of great assistance. This functionality enables the user to embed video, audio, text and image annotations on the map. These annotations are uploaded to the database as BLOB data along with their associated coordinates (Figure 12).

Using JSP and Servlet technology, the user is required to enter some text describing the annotation. If the annotations are video, audio or image annotations a file is also required to be uploaded. When this procedure is completed the map is refreshed and an icon representing the annotation is displayed. The user can view the annotations by clicking on the corres-
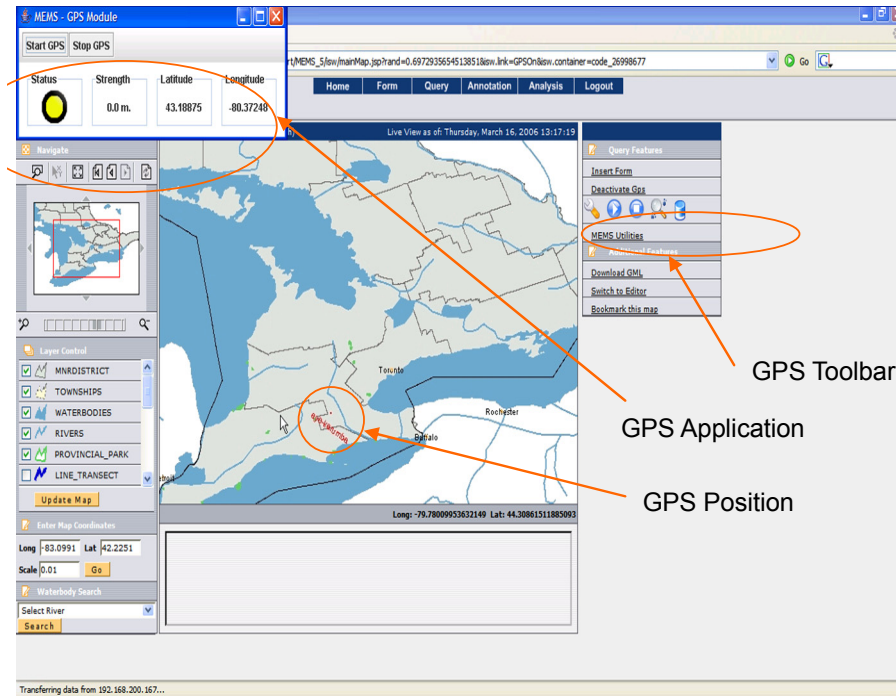
Figure 12. GPS module.
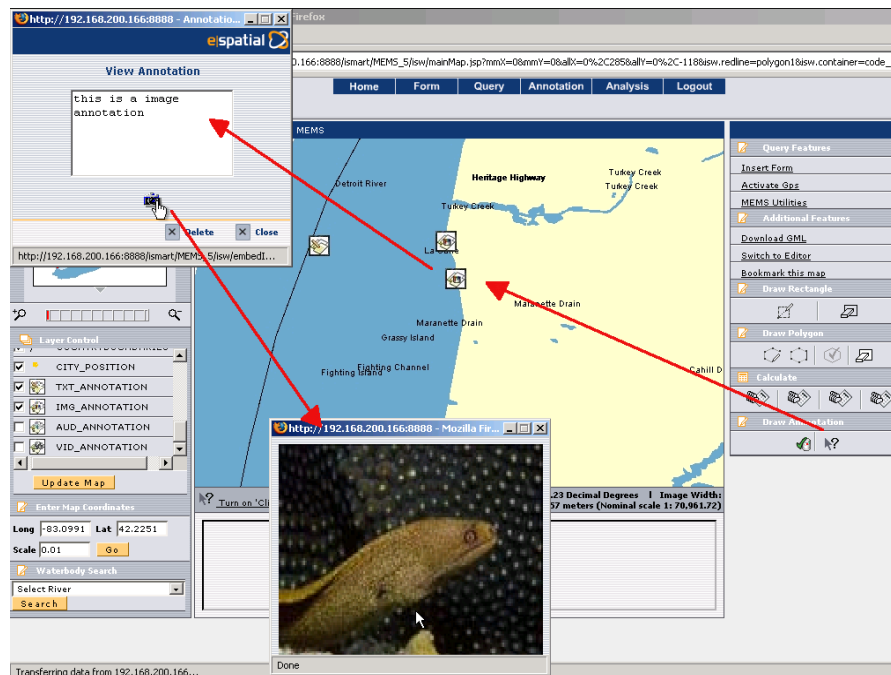
Figure 13. Multimedia annotation.

ponding icon. This action opens a popup window that displays any text associated with the annotation and a link to any multimedia data.

4.6. Offline Module

The offline module was only considered after our first field visit to the DFO. It was observed that some areas the biologists sample in had only intermittent cellular signal, if at all. As it was difficult to detect or predict network availability before heading to the field, the offline module was designed. The module is implemented using stand-alone Java and developed using the iSmart technology. The Offline Module, also

Figure 14. MEMS offline application.

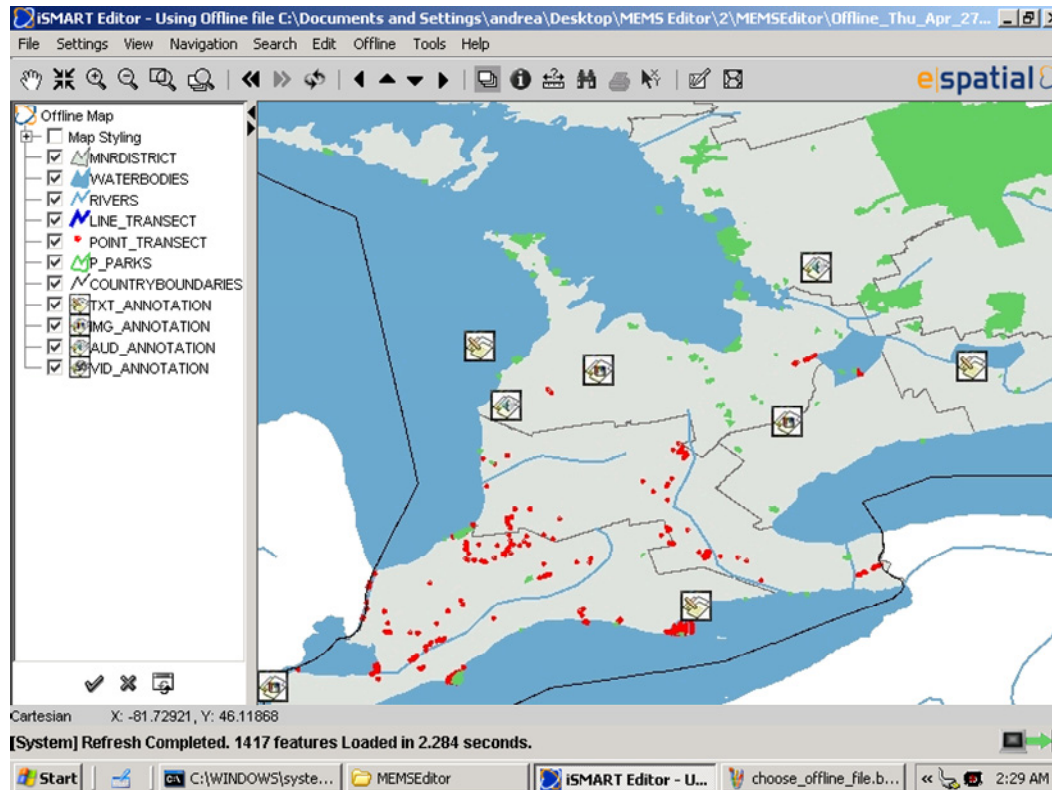called MEMSOffline, is a standalone application that connects directly to the online application server back in the office where the biologist is asked to select an area of interest. The selected area is compressed into a zip type file and saved locally on the tablet PC. The MEMSOffline compressed offline file contains all the data required by the field biologist and the application (Figure 14).

This offline application enables the user to work remotely when a GPRS connection is not available. In the field, the offline application reads in the compressed file and displays the map using the iSmart Editor application, enabling the biologists to perform spatial queries, insert new forms, look at previously recorded forms, display GPS feeds and navigate the map. It also offers restricted multimedia annotation functionality. Data which is changed or added during the sampling session is stored and saved in the offline file. When the application establishes a network connection and can re-connect to the application server, the data is synchronized with the server and any additions are added to the database.

5. Conclusions

In this paper we have described the development of a Web-based and mobile environmental management system called MEMS. Such a system was implemented to facilitate the work of biologists in the GLLFAS Fish Habitat Management Group of the Canadian Department of Fisheries and Oceans.

The MEMS system has been successfully implemented from the technical specifications design. The MEMS system relies on a Three-Tier Architecture that offers good distribution of processor load, ideal for the development of applications on low-spec devices such as PDAs, and Tablet PCs. Most of the computation is done on the Application Server Layer, and some on the Database Layer, minimizing as much as possible computation on the client layer. The result is a lightweight web application with a map interface that enables the user to display fish-sampling transect information based on a water body name, a location, or by clicking on an element on the map. The user is also able to insert new transect information using an intuitive interface.

The system includes an annotation module which enables the users to annotate points and areas on the maps and also add, edit, delete and query annotations in the database. This functionality has been fully implemented on the web-based application. A GPS module enabling the application to automatically acquire global position has also been fully implemented for the web-based application. In addition a pure Java standalone offline application was developed to mimic the functionality and the operability of the web-based application. Therefore, the same functionality of GPS module, Form Module, Annotation Module and Synchronization Module, along with all the mapping functionality is also available in the offline version and is fully operational.

MEMS is currently deployed at the GLLFAS headquarters in Burlington Ontario, where extensive field testing has

been carried out. The system offers biologists a spatially-enabled mobile and adaptable service that ensures better utilization of resources. However, the advantages of MEMS go beyond the system's functionality to improve both the biologists working environment and in-house data management through automating fisheries data entry.

Overall response from these groups was generally very favourable once the initial learning curve for a new data collection/entry methodology was overcome. However, this would not be an issue for biologists using the system daily. In terms of cost savings, as the filling out of existing field data sheets was generally mimicked on the MEMS tablet PC (i.e. the same biological measurements need to be recorded), its use in the field typically took a similar amount of time for data recording. However, the real cost savings would be given by not having to manually enter the data into the office database after returning from the field. This would save up to 3 months of time every year, depending on the number of sites sampled in the Great Lakes region every summer, and the number of biologists assigned to data entry tasks.

In the future we intend to investigate the adaptability of the MEMS architecture and GUI to different applications and departments within the DFO and elsewhere.

References

Carswell, J.D., Gardiner K., Bertolotto M. and Mandrak N., (2004). Applications of mobile computing for fish species at risk management. Proceedings International Conference on Environmental Informatics of International Society of Environmental Information Sciences (ISEIS2004), Regina, Canada, August 2004.

eSpatial Solutions, (2004). iSmart Web Developers Guide. http://www.espatial.com.

Hinze, A. and Voisard, A., (2003). Locations and time-based information information delivery in tourism. Proceedings 8th International Symposium on Advances in Spatial and Temporal Databases, Santorini Island, Greece, pp. 489-507.

Lopez, X., (2003). Oracle Spatial and Oracle Locator, An Oracle technical white paper, Oracle Corporation, Redwood City, CA, http://www.oracle.com/technology/products/spatial/pdf/10g_spatial_locator_twp.pdf.

Minns, C.K. (1997). Quantifying "No Net Loss" of Productivity of Fish Habitats, Can. J. Fish. Aquat. Sci., 54, 2463-2473, doi:10.1139/cjfas-54-10-2463.

Minns, C.K., (2001). Science for Freshwater Fish Habitat Management in Canada: Current Status and Future Prospects, Aquat. Ecosyst. Health Manage., 4, 423-436, doi:10.1080/146349801317276099.

Rigaux, P., Scholl, M. and Voisard, A., (2002). Spatial Databases, Academic Press, Morgan Kaufmann Publishers.

Rizzini, A., Gardiner, K., Bertolotto, M. and Carswell, J., (2006). MEMS Mobile GIS: A Spatially Enabled Fish Habitat Management System. Proceedings W2GIS'06 (6th International Symposium on Web and Wireless Geographical Information Systems), Hong Kong, pp. 113-122.

Schiller, J., (2000). Mobile Communications, Addison Wesley Publishers.

Sharma, J., (2001). Oracle Spatial , an Oracle technical white paper, Oracle Corp., Redwood City, CA, USA.

Talukder, A. and Yavagal, R., (2006). Mobile Computing, McGraw Hill.

5.6. Paper VI: An Open Source Approach to Wireless Positioning Techniques

Rooney, S., K. Gardiner and J. D. Carswell. (2007) "An Open Source Approach to Wireless Positioning Techniques": In Proceedings of the 5th International Symposium on Mobile Mapping Technology (MMT), Padua, Italy. ISPRS Archives – Volume XXXVI-5/C55.

AN OPEN SOURCE APPROACH TO WIRELESS POSITIONING TECHNIQUES

S. Rooney *, K. Gardiner , J.D. Carswell

Digital Media Centre, Dublin Institute of Technology, Aungier St., Ireland
(seamus.a.rooney, keith.gardiner, jcarswell)@dit.ie

ABSTRACT:

There are several problems encountered when trying to determine the location of a mobile phone, including whether you are in an urban or rural environment. Also, it is well known that some positioning technologies work better than others depending on the environment they are in. For example, GPS works well in rural areas but not as well in urban areas, GSM positioning accuracy can be acceptable in urban areas with the right triangulation technology, but is less accurate in rural areas. Positioning with other technologies such as WiFi, Bluetooth, and Semacode all have their own advantages and disadvantages as well, depending on the overall environment in which they are used. One research task of the ICiNG project is to address these issues and introduce the next logical step for freely available mobile positioning, advancing the pioneering work done by Place Lab at Intel. The EU-FP6 ICiNG project component that initiates this advance is called the *ILC (ICiNG Location Client)*. The ILC integrates all the above location finding technologies into one positioning module. This paper outlines the technique we developed to combine these technologies and the architecture used to deploy them on a mobile phone. With all these technologies finally available on one device, it is now possible to employ a personal positioning system that can work effectively in any environment. Another important advantage of the ILC is its ability to do this without any direct communication with outside sources, so users need not worry about "big brother" tracking their every movement. The ILC only "listens" for, and makes use of, radio signals that are freely available in the current environment, and does not actively connect to any external network or other services to triangulate its position.

1. INTRODUCTION

The ICiNG project is about researching a multi-modal, multi-access concept of e-Government (DIT, 2006). It develops the notion of a 'Thin-Skinned City' that is sensitive to both the citizen and the environment through the use of mobile devices, universal access gateways, social software and environmental sensors. Intelligent infrastructure enables a Public Administration Services layer and a Communities Layer.

Communities interact with the infrastructure to avail of services created by the administration, and can also create their own information-based services. The ICiNG project will set up test-beds in high-profile European locations such as Dublin, Barcelona and Helsinki to act as 'City Laboratories' for researching, evaluating and demonstrating technologies and services using intelligence in the environment.

In the ICiNG project, the fundamental requirement for this type of service interaction is location. For example: in Dublin, one service to be provided is an *issue tracker* that enables citizens to report accessibility issues (e.g. lack of wheelchair access) to Dublin City Council. This requires location as part of the report and the module that will provide this location is the ILC (ICiNG Location Client). The ILC is by design an open source, network independent, location determination mobile application that can utilise GPS, WiFi, GSM, Bluetooth and Semacode information or any combination of them, to calculate location. (Kilfeather et al., 2007)

The paper is organised as follows: Section 2 gives a brief summary of currently available location determination techniques. Section 3 focuses on the ICiNG Location Client and Section 4 covers the conclusions for the project so far and

comments on the expectations for the ILC location determination methodology implemented.

2. LOCATION DETERMINATION TECHNIQUES

2.1 Radio Signal Strength: Signal Strength is the measure of how strongly a transmitted signal is being received or measured at a reference point that is a particular distance from the transmitting antenna. By using the signal strength of a particular radio wave and the position of the transmitting antenna, it is possible to calculate the location of the receiving device. The degree of accuracy for this location can vary greatly using this method depending on the environment that the signals have to travel through. (Cheng et al., 2005) For example, in an urban environment, the signals may have buildings and even traffic (e.g. double-decker bus) to travel through before reaching the receiving antenna. In such a dynamic environment, it is found that estimating a position based on signal strength alone can be prone to significant error in accuracy.

2.2 Radio Time-of-Flight: Radio Time-of-Flight is measuring the time needed for a radio wave to travel from a transmitter to the receiving antenna. However, since radio waves travel at different speeds depending on the atmospheric conditions they travel through, this also needs to be taken into account. For example, sound waves have a velocity of approximately 344 meters per second in 21°C air. Therefore, an ultrasound pulse sent by an object and arriving at a point 14.5 milliseconds later calculates out to the object being 5 meters away from the point. (Hightower and Borriello, 2001) Using this method to determine location can be very useful but it does require that certain extra hardware be installed in the transmitting tower and therefore rules out the possibility of

calculating the position of a mobile device independently of such infrastructure.

2.3 Proximity: Proximity location determination is a technique whereby a location is determined when an object is near a known location. Of the different proximity techniques such as: physical contact (e.g. pressure sensors); monitoring radio access points (e.g. determine where a device is connected to a physical network) or; automatic ID systems (e.g. last login on a computer terminal), only monitoring radio access points would offer a benefit to the ILC as it is designed to listen for radio signals, check its own database of beacon locations and then triangulate a position.

2.4 GPS: The Global Positioning System (GPS) is a constellation of satellites of known position operated by the U.S. military and is made accurate to use for all, providing its *selective availability* (SA) feature is set to introduce zero error. It uses the Time-Of-Flight positioning technique and is currently the oldest fully functional satellite navigation system in operation, followed by GLONASS and two additional systems, Galileo and Beidou, which are proposed to come on-line over the next few years.

2.5 GLONASS: The Russian GLONASS system is a counterpart to the U.S. GPS and is managed by the Russian Space Forces. Both systems share the same principles in the data transmission and positioning methods. However, due to the recent economic situation in Russia the system became almost obsolete, but following a joint venture with the Indian Government it is hoped to have the system fully operational again by 2008 with 18 satellites, and by 2010 with 24 satellites.

2.6 Galileo: Galileo is an E.U. led project to make GPS style positioning available to its civilian population with higher accuracy than current GPS and GLONASS, and without intentional error added to the signal in times of political strife. It is expected to be operational by 2011-2012, three or four years later than was originally expected.

The accuracy of the three systems above varies. GPS and GLONASS have similar accuracy of ~20-100 meters unassisted and with SA switched off, but both systems can increase their accuracy to sub-meter using a technique called differential positioning (El-Rabbany, 2006). Although Galileo is not complete, it is expected to provide greater accuracy (down to ~1m) than the previous two systems with greater penetration in urban canyon type environments and a faster fix. Additionally, it will not suffer from one disadvantage of the current GPS system in that it is still possible that the U.S. could at any time reintroduce a selective availability error to intentionally reduce the accuracy of the positioning signal.

2.7 Beidou: Even though China has a €200M stake in Galileo, Beidou is their contribution to the choice of Satellite Positioning System. The main difference between this system and the others is that Beidou will use a circular geostationary orbit where each satellite appears to remain at a fixed point in the sky over a fixed point on the earth. Although this means that the system does not require a large constellation of satellites, it will have limited use in the higher latitudes as the coverage area is reduced.

Overall, these are the four main satellite systems that will be used in the near future for location determination and although

newer technology will increase the accuracy, such systems also suffer from varying degrees of similar signal fix problems in urban environments. Also, even though GPS chips are getting cheaper, in mobile phones they will continue to suffer more in urban environments than bulkier, albeit more robust, professional survey standard receivers.

2.4 Semacode: Semacode is a relatively new technology that uses print tags to provide location to mobile devices with a camera and Semacode software installed. The location of the tag is encoded on a 2D barcode and when imaged, is decoded and made available to LBS applications on the device. This is a very accurate positioning technique (+/- 1m) and can be used to correct the position estimates of less accurate techniques like stand-alone GPS readings.

2.5 Place Lab: Place Lab was an Intel research project from 2003 to 2006. The goal of Place Lab was to try and use available radio signals (GSM, WiFi, and Bluetooth) by building a database of their locations, and then use these radio beacon locations to triangulate for the users location. However, at the time, certain technologies were not yet advanced enough to take Place Lab fully mobile as some of the signal spotters worked on mobile phones (i.e. Bluetooth) while others needed a laptop (i.e. WiFi). Like the ICiNG Location Client, Place Lab also wanted to determine position using passive monitoring and gives the user control over when their location is disclosed, laying the foundation for privacy-observant location-based applications. (LaMarca et al., 2005)

2.6 Privacy: One of the main concerns during the ILC design phase was that of user privacy. We did not want to design a system that would or could be used to track a citizen's location without their knowledge or prior consent. Instead, we wanted to develop a system where all the location determination could be done on the mobile device itself and then, only if the user wanted, it would be possible to inform the rest of the ICiNG system of their location. In this way, the user will have full control over their location and would not have any concerns about their movements being tracked without their knowledge.

For users that allow the ILC to disclose their location to ICiNG services, we also wanted a system that would address any other privacy concerns they might have. Of these issues, we identified the following to be particularly important; location information retention, location information use, and location information disclosure.

For ICiNG, it was decided at an early stage that location information retention, if it needed to happen at all, would be only for a short task-specific time frame depending on whether the user was partaking in specific studies or if they had signed up to a service that required their movements to be tracked. For example, a parent could register their child's mobile phone to such a service to monitor their child's whereabouts. Another issue that needed to be addressed was that of disclosing movements of users to 3rd parties, which the ILC never directly does.

So far we have discussed some of the different positioning techniques available, and some of their advantages and disadvantages. We identified concerns that most people would have about their privacy being infringed upon and what the possible solutions to these issues are. We also noted, due to

technology limitations of the day, what Place Lab was not able to do in bringing a fully mobile location based system to fruition, and how ICiNG would take the next step and extend their work by designing and developing such a system. The next part of this paper gives a more detailed overview of the ICiNG system, focusing on the ICiNG Location Client.

3. ICiNG

3.1 ICiNG Test Bed

The ICiNG system is designed to help bring communities closer to information about their environment with the use of mobile phones, universal access gateways, social software and environmental sensors. It will be deployed in three cities namely, Dublin, Barcelona and Helsinki. As illustrated in Figure 1, the proposed system architecture can be seen as a layered structure of services, technologies, and networks that allow ICiNG services to reach the citizens and vice versa. The bottom layer is composed of the two main sources of data for the ICiNG project: the city and its citizens. In ICiNG we generalize the concept of "Citizen Sensor" in its objective of creating an 'always on' channel of available and context rich bi-directional communications between the city and its citizens across a broad range of technologies, content formats and interaction schemes.
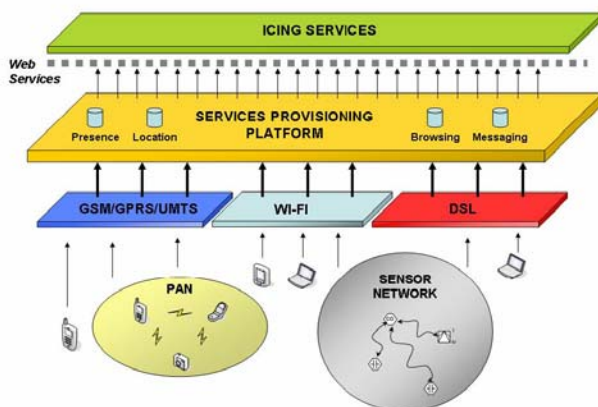

Figure 1: Test bed Layered Architecture (Telefonica, 2007)

Thus, in ICiNG we see citizens not only as consumers of the provided services but also as an active part in the creation of them. Furthermore, if we consider the capillarity of citizens, and more concretely of their personal devices, we can see them as a source of passive information of paramount importance for the construction of the 'Thin-Skinned City' model.

The general approach of the Dublin test bed specifically is to bring existing wireless and wired infrastructure belonging to the Dublin Institute of Technology and Dublin City Council together within an experimental on-street wireless network (Figure 2). Access to the new wireless network will, in the first instance, be open, free and supported by the Dublin ICiNG project team.

The ICiNG street signs will be the WiFi/Bluetooth Access points and the ILC will use its internal logic to determine a best guess location from the known positions of the street signs and other radio signals in the area.
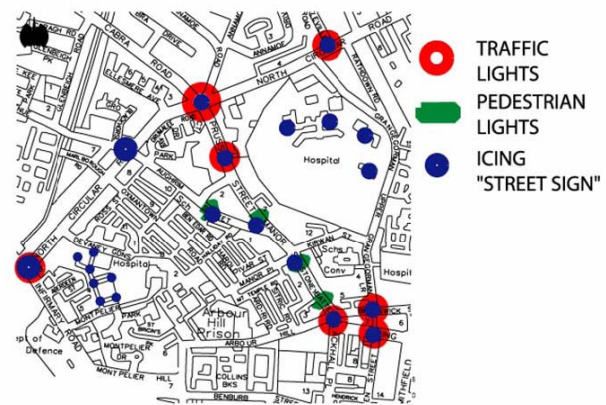

Figure 2: Dublin Test Bed

3.2 Specific Terminology

The *Icing Mobile Client (IMC)* refers to the complete set of application components on the mobile device. The IMC is comprised of:

- The *ICING Location Client (ILC)* whose purpose is to calculate and make available the device location to the MDA based on GPS, Semacode, and Wireless Beacon information.

- A number of *Mobile Device Applications (MDA)*. An example of an MDA is an accessibility application that enables users of the ICiNG system to report accessibility issues to the City Council using a Jabber client[1] extension.

3.3 Location Considerations

There are many Location Based Services (LBS) identified in the ICiNG project. These range from providing a location tracking sensor network to retrieving metadata based on a mobile device's location. While these services are heterogeneous in nature they all require a method of determining the location of a particular device or sensor. There are many existing systems available to provide this location information, some using cell services provided by mobile telecoms providers or others using satellite technology such as GPS.

However, as discussed previously each of these technologies and services have inherent advantages and disadvantages. Some services operate well in urban areas and in areas of high cellular radio density while others perform well where line of sight to satellites in the GPS system is established. Also, beyond the purely technical or technological considerations to be taken into account in location determination are issues of privacy and safety which location technologies raise. (Vossiek et al., 2003)

Fortunately, the issue of deciding which of these technologies to use is being somewhat mooted by the increasing trend of mobile devices to incorporate multiple access technologies in

---

[1] An open, secure, ad-free alternative to consumer IM services like AIM, ICQ, MSN, and Yahoo. Jabber is a set of streaming XML protocols and technologies that enable any two entities on a network to exchange messages, presence, and other structured information in close to real time.

the same platform. The availability of GSM, WiFi, Bluetooth and GPS on the same device offers the possibility of intelligently using all these technologies in combination to improve location availability and accuracy.

3.4 ILC – ICiNG Location Client

For the ILC, we use a combination of technologies to develop a location determination system that integrates the best features of all technologies available. By using all these technologies together, any disadvantage that each individual technology has diminishes. The ILC is designed as provider-network independent, privacy sensitive and zero cost (in terms of network resource usage) software component that allows mobile devices to determine location by a "best guess" methodology. The prototype ILC is designed to run on a Series 60 (3rd Edition) mobile phone running the Symbian operating system (version 9.x), although other platforms and operating systems could be accommodated with relatively minor changes. Figure 3 shows how this architecture works together.

There is a set of rules that the ILC will follow when searching for beacons in its RMS Database. It will be dictated by the degree of accuracy that should be expected depending on the beacon. Although all beacons in the RMS Database will be read, it will first look for Bluetooth beacons and if any are found it will discard the other beacons for determining its location; hence the degree of accuracy should be <10 meters. Next, if there are no Bluetooth beacons, it will look for a GPS reading, if there is no GPS reading then it will look for WiFi beacons and if there are no WiFi beacons, it will look for GSM beacons. As the technologies used changes, so will the degree of accuracy the tracker is providing. Even though beacons with a lesser weight for accuracy are disregarded in the location determination returned to the MDA, this beacon data does not get totally discarded. If the more accurate beacons become unavailable and the ILC switches to the less accurate beacons in its Database for triangulating position, then the lesser accurate location get a correction applied based on its proximity to the last known more accurate beacons.
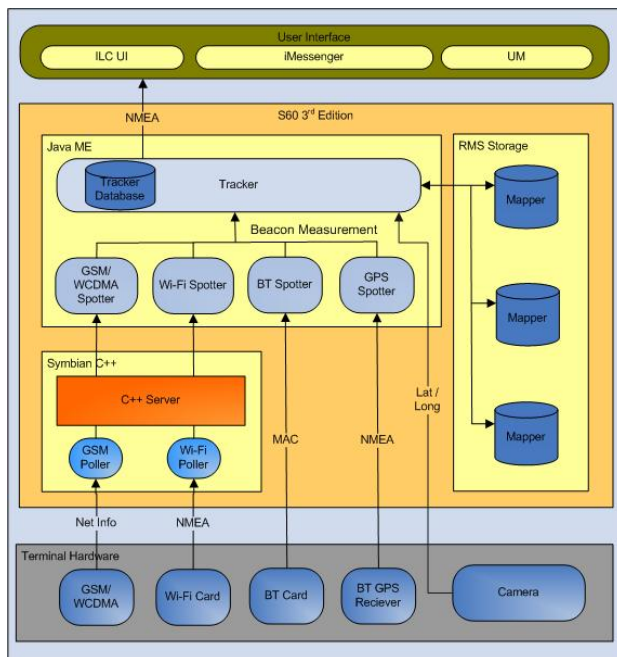

Figure 3: ILC Architecture

3.4.1   Java Bluetooth Spotter: The Bluetooth spotter will poll the Bluetooth Terminal hardware to scan for any Bluetooth devices in range, and any devices found in range will be returned to the tracker module.

3.4.2   Java GPS Spotter: The GPS Spotter communicates with the external GPS receiver to get Lat/Long coordinates and returns these coordinates to the tracker module.

3.4.3   Java GSM Spotter: The Java GSM spotter sends requests to the C++ GSM spotter to retrieve the current GSM tower information and returns it to the tracker module.

3.4.4   Java WiFi Spotter: The Java WiFi spotter sends requests to the C++ WiFi spotter to retrieve information about the WiFi access points present in the area and returns it to the Tracker module.

3.4.5   Semacode Spotter: The Semacode module is responsible for taking photos of 2D barcodes and translating them into Lat/Long coordinates. The spotter then returns these accurate coordinates to the tracker module.

3.4.6   Tracker Database: This is a list of the beacons that are currently in range of the mobile device. This list is compared against the known beacon database to get the beacon locations.

3.4.7   RMS Database: This is a database of known beacons and their locations. This list is used by the tracker to get beacon locations which are used to triangulate the current position of the mobile device. Initially this database has only data we inputted into it manually, but future versions of the ILC will use databases created directly from Wardriving and that can be downloaded from websites like [**www.wigle.net**].

3.4.8   C++ GSM Spotter: The C++ GSM spotter polls the GSM terminal hardware to determine the cell information about the currently active cell tower and returns this data back to the GSM spotter.

3.4.9   C++ WiFi Spotter: The C++ WiFi Spotter polls the WiFi terminal hardware to determine the MAC addresses of any WiFi access points in range and returns this data to the WiFi spotter.

3.4.10   Tracker Module: The Tracker module is responsible for collecting beacon information from the four Java Spotters, organizing the beacon information into local databases, triangulating the current position of the mobile device based on the data it has and responding to location requests from Mobile Device Applications. The tracker works as follows,

1. The tracker module waits for location requests from Mobile Device Applications.
2. Upon receiving these requests, the tracker module polls four Java spotters for any beacons in range of the phone.
3. The Java spotters return information about any beacons in range.
4. The tracker module then stores this information in the tracker database.

5. The tracker module then checks to see what beacons are in the tracker database.

6. Depending on the beacons available, the tracker chooses which beacons to use and then compares them against known beacons in the RMS database.

7. If the current beacons are unknown, then the tracker module will revisit the tracker database to select the next viable beacons to check against the RMS database.

8. Once the tracker module's beacons have been checked against the RMS Database, it then attempts to triangulate the best possible position based on the current beacon information.

9. When the position has been determined, it is returned to the Mobile Device Application that sent the request, along with a degree of accuracy.

3.5 Location Calculation

The ILC will implement and test several different location triangulation algorithms to determine the best possible location. After initial testing we will look at improving upon one or more of these algorithms as there might be a performance price that needs to be paid on a mobile device to achieve increased accuracy. Only after careful testing will we make these determinations. Of the algorithms we have decided to implement and initially test, a brief description of each one follows.

3.5.1 Centroid Location Determination: This is one of the simplest algorithms that can be implemented. It involves taking into account the locations of all known beacons in the area and then positioning the user mathematically at the centre of them. This approach ignores many things that could improve the location determination including beacon signal strength, confidence in the beacons location and environmental issues (tall buildings, hills, buses). (Hightower et al., 2006)

3.5.2 Weighted Centroid Location Determination: This is very similar to the Centroid algorithm above but it takes into account other values in the RMS database when calculating a location. For example the signal strength of each beacon can be taken into account to further determine if the ILC is nearer one beacon or another. In our test bed, we have fixed beacons with exact known positions. Initially we will only be using these beacons but during later testing, it is possible to do Wardriving in and around the test bed area and add unknown beacons to the RMS database that might have only an estimated location. In this case, we give our known beacons a greater weight when calculating location than beacons that are found through Wardriving.

3.5.3 Assisted ILC: Much like assisted GPS, assisted ILC will use an Assistance Server but on the phone. For example, if a mobile device was within signal distance of a known Bluetooth beacon but then moves out of range of that Bluetooth beacon. This information would not be discarded straight away. Assisted ILC will use this information to help correct the location determination for a period of time/distance after the user has moved away from the more accurate beacons. As the time/distance passes so the weight of the correction will decrease. Also depending on the type of beacon being used for the correction, there will be an error attached to the expected accuracy it can correct for.

Initially we have decided to implement and test these three algorithms in our test bed environment. There are other techniques that were considered, e.g., Particle Filters (Hightower and Borriello, 2004) and Fingerprinting like what is used in RADAR (Bahl and Padmanabhan, 2000) but initially it was decided that these might be too costly in terms of performance on a mobile device. Another possible technique is Gaussian Processes for Signal Strength-Based Location Estimation (Ferris et al., 2006) which appears more optimal for mobile devices. After thoroughly testing our first three Location Algorithms we will then decide if a different approach is required.

4. CONCLUSIONS

The ICiNG system is designed to allow citizens of the city to interact with Government Departments and City Councils in a context sensitive way using 3rd generation mobile phone technology. In this paper, the initial attempts at designing and developing an open source location determination approach has been described. This approach has been taken to provide the next step for context sensitive interaction through the ILC. The ILC is developed as fully open source and is aimed at providing a location determination component to provide a quick to market solution for LBS applications. The ILC is a zero cost, privacy sensitive location determination application.

The development of the ILC is ongoing. The next step in the development will involve investigating increasing the accuracy of the calculated location, based on integrating a number of triangulation algorithms mentioned and testing them in environments that have a sectorized and structured infrastructure and in environments that do not have a structured infrastructure. We will also test how each of the triangulation algorithms performs on a mobile device to verify any performance issues and what alterations are needed.

5. REFERENCES

Bahl, P. and Padmanabhan, V.N., 2000. RADAR: An In-Building RF-based User Location and Tracking System.

Cheng, Y.-C., Chawathe, Y., LaMarca, A. and Krumm, J., 2005. Accuracy Characterization for Metropolotan-scale Wi-Fi Localization.

DIT, 2006. ICiNG Project website; http://www.fp6-project-icing.eu/.

El-Rabbany, A., 2006. Introduction to GPS: The Global Positioning System, Second Edition Artech House Publishers.

Ferris, B., Hahnel, D. and Fox, D., 2006. Gaussian Processes for Signal Strength-Based Location Estimation.

Hightower, J. and Borriello, G., 2001. Location systems for ubiquitous computing. Computer, 34(8): 57-66.

Hightower, J. and Borriello, G., 2004. Particle filters for location estimation in ubiquitous computing: A Case Study. Springer Berlin / Heidelberg, 3205: 88-106.

Hightower, J., LaMarca, A. and Smith, I., 2006. Practical Lessons from Place Lab. IEEE Pervasive Computing, vol. 5, no. 3(Jul-Sept): 32-39.

Kilfeather, E., Carswell, J., Gardiner, K. and Rooney, S., 2007. Urban Location Based Services using Mobile Clients: The ICiNG Approach, GISRUK, Maynooth, Ireland.

LaMarca, A. et al., 2005. Place Lab: Device Positioning Using Radio Beacons in the Wild.

Telefonica, 2007. SIXTH FRAMEWORK PROGRAMME - Innovative Cities for the Next Generation.

Vossiek, M. et al., 2003. Wireless local positioning. Microwave Magazine, IEEE, 4(4): 77-86.

5.7. Paper VII: Viewer-Based Directional Querying for Mobile Applications

Gardiner, K. and J. D. Carswell. (2003) "Viewer-Based Directional Querying for Mobile Applications", Third International Workshop on Web and Wireless Geographical Information Systems W2GIS, Roma, Italy. In Proceedings of the Fourth international conference on Web information systems engineering workshops Pages 83-91, IEEE CS Press, ISBN:0-7695-2103-7.

Viewer-Based Directional Querying for Mobile Applications

Keith Gardiner, James D. Carswell

Digital Media Centre, Dublin Institute of Technology, Ireland
{keith.gardiner, jcarswell}@dit.ie

Abstract

With the steady and fast advancements in the integration of geographic information systems and mobile location-based services, interest in exploiting this technology for Cultural Heritage (CH) data sharing has become apparent. In this area there has been an increasing need to integrate positional information with non-positional data and add a spatial dimension to the definition of a users "context". In this paper we describe an implementation of a viewer-based directional query processor that operates on an Oracle Spatial database. The spatial position and orientation are taken from the viewer's perspective. Using this frame of reference a view-port is defined in real time as the viewer progresses through the space and used as the primary filter to query an R-tree spatial index. Finally, an experimental implementation shows how the query processor performs within a VRML model of Dublin linked to a spatially enabled CH dataset.

1.   Introduction

The approach of using direction to query spatial data is the focus of substantial research efforts within the spatial database community [6,9,10,12,13,18,19,20]. Direction relations therefore represent an important class of user queries in spatial databases and their applications to geographic information systems. To make sense of direction, a reference frame must first be established, where in general there are three possible options:

- *Intrinsic*, where the reference frame is in respect to the orientation of an object, e.g. front or back, left or right of a building;
- *Deictic*, where the reference frame is relative to each individual looking at the scene, e.g. what is "in front" for me might be "to the left of" someone else and;
- *Extrinsic*, where the reference frame is established independently of the orientation of the features or the observers, e.g. north, south, east, west.

For configurations of spatial objects, in a GIS or digital image, that represent real positions and orientations of the environment, it has been customary to use extrinsic reference systems [6]. However, although in [19] the direction from fixed objects in 2D space has a profound but highly static affect on the objects relevance in a context-aware environment, for our purposes, where each individual has their own personal LoS and therefore dynamic, personalised search space, a deictic reference frame is what we consider.

For example, when working within an intrinsic reference frame, queries like "Are there any CH artefacts in front of the post office?" can be answered, where the post office is the Object. In contrast, an example of a query within our viewer-based or deictic reference frame would be "Are there any CH artefacts contained within my view-port in the direction that I am facing?", i.e. a view-port virtually constructed along my LoS.

Therefore, for our purposes the direction that the viewer is facing will be used as the selection condition for queries to a spatial database. In this paper we show how the position of the viewer combined with the direction of his/her LoS is used to develop a viewer-based directional query processor that utilises an oriented, bounded object together with standard Oracle Spatial topological and metric operations. Our approach uses the LoS direction vector to represent orientation and constructs a view-port of varying, user-defined dimensions as the primary filter when querying the database. As the direction of the user changes, the view-port is reconstructed in real time to reflect the users new LoS search space. This method of querying the database is similar to a standard range query except that the shape of the view-port window is user defined and has orientation. This approach does not include new indexing data structures or access methods, instead utilises the already well-known R-tree index data-structure to perform the spatial queries [7,17].

2.   Non-Directional Queries

A traditional, non-directional range query is the recognized standard operation to query our database for any Cultural Heritage artefacts that are present in the location of the query window. The window is of a specified width and height centred on the user's location and is represented in Oracle Spatial as an optimised rectangle that is defined by a minimum of two points $p1$ and $p2$. An advantage of using an optimised rectangle (or any shape) is that it does not have to be inserted into an Oracle table before it can be used as a query window

(Figure 1). The following code excerpt shows how the *sdo_relate* operator is used to achieve this:

```
SELECT   A.ID, STREET, BUILDING
FROM     CHI.CHI_CONTENT_BUFFER A
WHERE    SDO_RELATE (A.POSITION,
         MDSYS.SDO_GEOMETRY (2003,NULL, NULL,
         MDSYS.SDO_ELEM_INFO_ARRAY (1,1003,3),
         MDSYS.SDO_ORDINATE_ARRAY (X1,Y1, X2,Y2)),
         'mask = anyinteract querytype=window') = 'TRUE' ORDER
         BY ID";
```

This example shows the *sdo_relate* operator being used to compare the query window with the CH dataset to determine weather they interact in any way. The *sdo_relate* operator accepts three parameters, an *sdo_geometry* column in a table that must be spatially indexed, the query window *sdo_geometry* and a *param* list that determines the behaviour of the operator. The query window geometry is defined in the SQL statement string and contains two arrays. The *sdo_elem_info_array* contains the values that define the type of geometry that is to be queried against the dataset. In this instance it is an optimised rectangle that is defined by a triplet value (1 (offset), 1003 (outer polygon), 3 (optimised rectangle)). The *sdo_ordinate_array* contains the coordinate values of the rectangle. The mask is set to *"anyinteract"* which means if any of the topological Boolean predicates return true (or interact) then adds the geometry to the resultset.

This SQL string is used to query the spatial database for interaction between the query geometry and the CH data geometries. The query is processed every 5 seconds. If the user is on the move and the position of the user is the same or less than 5m away from the position that the last query was processed, the query will not be run again. When the user's position is outside this threshold, the query is run again. This significantly reduces the computational cost of redundant queries to the database.
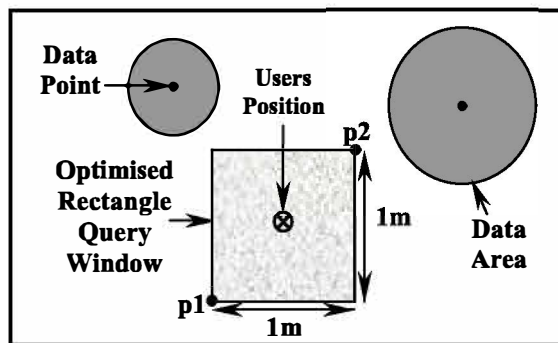

Figure 1: Optimised Rectangle (Window) Query

Each CH *data point* in the database represents a CH artefact and is surrounded by a buffer (the extents of which are also explicitly stored by Oracle Spatial) of varied radius depending on the size and/or significance of the artefact. Taken together the data point and surrounding buffer represent a *data area*. Justification for placing a buffer around the individual data points rather than around the viewer's dynamic location in space was one of maximising query optimisation as one of the most important aspects of the query process is the speed at which it is executed. If instead the circular buffer (or indeed any other complex shape) were dynamically generated around the mobile user, the points needed to specify the buffer extents would have to be recalculated each time the user's context changes in either time or space. An optimised rectangle therefore is the most favourable geometry to query the database against.

When the user's query window intersects a data area in any way, e.g. touch, overlap, etc., the relevant data is placed into a resultset and displayed to the user in the form of text, images, audio, and video files. In the initial implementation it was sufficient to collect data in this manner. The problem here is that the orientation or direction that the viewer is facing is not taken into account when formulating the query.

An enhanced implementation extends the initial attempt by adding functionality that allows the query window to be of any required shape. Unlike the initial approach of using an optimised rectangle, the users position shifts to point *p1* (Figure 2) and it is now required that all the points required to construct the shape-query are specified. Therefore, each point that makes up the optimised shape is now calculated relative to point *p1*, also called his *viewpoint*.

Although this adds flexibility to the dimensions of the users relevant query space, the problem with this approach is that the query window's orientation is static, i.e. by changing the orientation of the viewer, simply by rotating about the z-axis, the query window's position is not affected. This is not an optimal condition as the user is receiving information about data that may be behind them and not what is in their direct Field-of-View (FoV), i.e. along their LoS.

3.   Directional Queries

In the enhanced implementation, the method of querying the Cultural Heritage (CH) database is more sophisticated. In this approach orientation is a necessary parameter so that the users view-port can be dynamically constructed, resulting in only data contained within the viewers FoV being returned. This makes querying the data more adaptive; as the user progresses through the VRML world their view-port is being continuously updated with respect to the direction they are facing. The query that is formulated in this manner is similar to a standard range query with an optimal shape. As with the standard range query described previously, the query window is compared to an R-tree index. If the query

window comes in contact with any of the Minimum Bounding Rectangles (MBRs) of the data areas, the data is said to be intersecting in some way.

In our Directional Query Model there are three points defining the triangular query window representing the extents of the user's FoV. The user's viewpoint *p1* is always one of the vertices of the triangle. The points *p3* & *p4* are calculated by first attaining the azimuth from the 0° North direction to the Line-of-Sight (LoS) of the user, i.e. to point *p2*. (Figure 2)
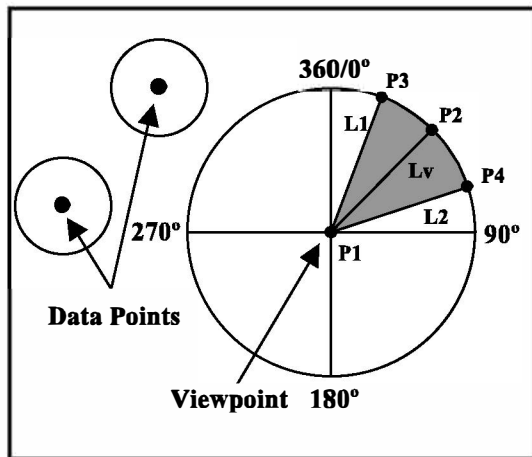

Figure 2: Directional Query

The azimuth of the LoS (*Lv*) could be obtained either from a VRML browser (as is our case) or from a digital compass embedded within a spatially enabled PDA. To determine the azimuths to points *p3* and *p4* a specified fraction of an angular FoV value is subtracted from the user's azimuth *Lv* to get the azimuth to point *p3* and by adding the same fraction of the FoV value to *Lv* to get the azimuth to point *p4*. These FoV extents (*L1* & *L2*) are then used to calculate the positions of vertices *p3* and *p4* on a query buffer of specified radius thus giving the view-port a finite distance. Together, the three vertices are used to produce an optimised spatial query shape that will only select data that is inside a triangle oriented in the same direction as the users LoS.

The orientation and position coordinates that are delivered to the java application are obtained by using the External Authoring Interface (EAI) Java Application Programming Interface (API). The EAI is a programming interface for communication between VRML and external programs and allows the developer to register some of the VRML events and properties to the Java programming environment [5,21].

The location of the user's viewpoint, while navigating within the VRML model, is used to simulate the user's position in the real-world streets of Dublin. The virtual 3D coordinates (x,y,z) are transformed into geographic coordinates (φ,λ), the initial interest of the

context-based query to the spatial database. In addition to the position, the orientation of the user's LoS can also be obtained in the same way. The orientation field values provide a rotation axis about which to rotate the viewpoint and a rotation angle specifying the amount by which to rotate around that axis. The first three values in the field specify the X, Y and Z components of the 3D direction vector. The fourth value in the orientation field specifies the positive or negative rotation angle measured in radians [1].

In VRML the z-axis is the vertical axis and is used to specify the orientation of the user. *R* is the angle of orientation of the user around any given axis. The *R*-values scale from 0 to +/-3.14 and are subsequently converted into degrees. In the case of the z-axis, this angle is the azimuth and is used to calculate the intersection of the FoV bounding lines and the query buffer. In the VRML data model 0° is North, in Java 0° is east (Figure 3). This discrepancy was adjusted for in the application so that all azimuths are synchronised both in the VRML implementation and in the Java implementation.
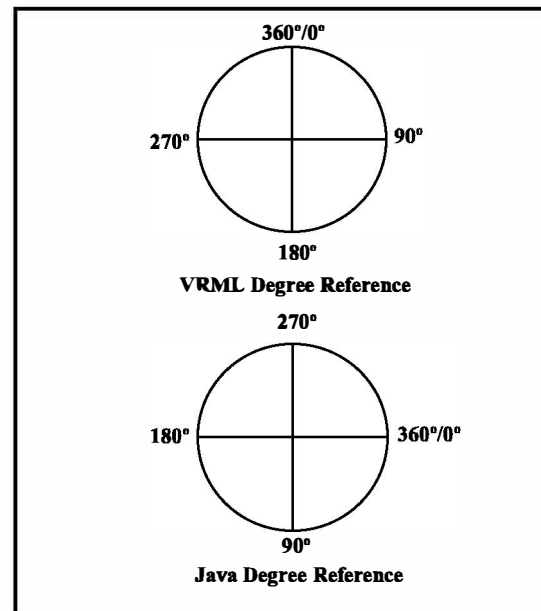

Figure 3: Degree Reference Systems

The java class that constructs the query window accepts five parameters and returns three points, those used to implement the window. The five values are those obtained from the EAI.

The three points returned from the algorithm are passed into the SQL query string used to query the spatial database. In this example the *sdo_elem_info_array* is modified to represent a polygon, i.e. the query window.

```
SELECT  A.ID, STREET, BUILDING
FROM    CHI.CHI_CONTENT_DATA A
WHERE   SDO_RELATE (A.POSITION,
        MDSYS.SDO_GEOMETRY (2003, NULL, NULL,
        MDSYS.SDO_ELEM_INFO_ARAY (1,1003,1),
        MDSYS.SDO_ORDINATE_ARRAY (X1, Y1, X2, Y2, X3,
        Y3, X1, Y1)), 'mask=anyinteract querytype=window') =
        'TRUE';
```

The *sdo_elem_info_array* contains the values that define the type of geometry that is to be queried against the dataset. In this instance it is a polygon that is defined by a triplet value (1 (offset), 1003 (outer polygon), 1 (points are connected by straight lines)). In the *sdo_ordinate_array* the three points are specified and the first again to close the polygon.

3.1.   Line-of-Sight

The next objective was to investigate and develop a LoS algorithm to determine if data contained within the view-port is actually in the viewers LoS. This problem is illustrated in Figure 4. The large triangular area in the diagram represents the user's view-port in 2 dimensions. The brick filled shapes B1, B2, B3, and B4 represent building blocks and D1, D2, D3, and D4 represent CH data points. The enclosed white space in the diagram highlights the desired shape that the LoS algorithm should identify as the query area. The light grey sections represent the areas that should be excluded from the query space, as they are not visible from the user's viewpoint.
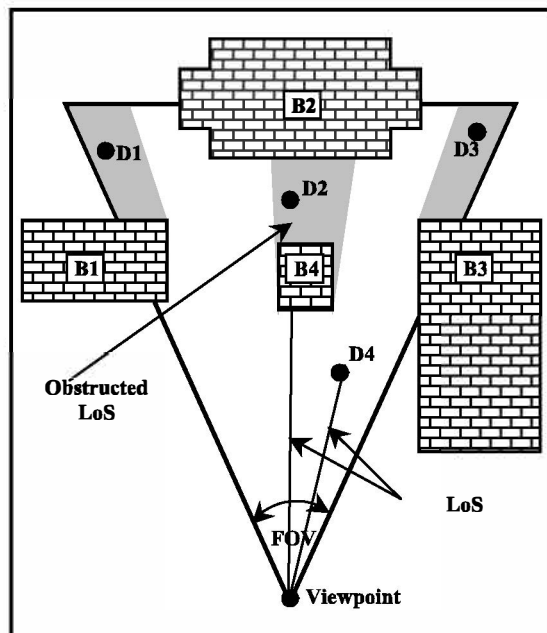

Figure 4: Optimised View-Port

When the CH database is queried using an oriented view-port, the query simply checks to determine weather any of the CH dataset is in contact with the larger triangular shaped view-port. If there is any topological relationship detected the data is returned to a resultset and subsequently displayed to the user. This querying of the database does not take into account the fact that the triangular query window is also interacting with other layers in the database (e.g. the building layer).

In reality, if the user is standing on the outside looking into a building, they cannot see what is inside. In our initial implementation of the query model, if the user is standing on the outside they indeed could retrieve data that is unseen to them. As a primary filter this condition is unacceptable when in fact the user only wishes to receive data about objects that they can actually see. Therefore, an option was added that checks to determine if the view-port interacts with any of the building blocks in the block layer of the database. If so, the sub-area of the triangular shaped view-port that overlaps the building polygon should be removed from the query window. This is illustrated in Figure 4 where the shape of the triangular query window has been reduced to the enclosed white and grey space only.

The information obtained by performing this check eliminates the building blocks from the view-port while at the same time gives us data on what buildings the view-port is actually intersecting with. This data can then be used to determine if the building blocks involved in the intersection are in the users LoS to other data points in the resultset.

Additionally, in comparison with the CH layer, the building block layer also contains attribute information about the individual polygon objects stored in the layer. In the case of the block layer therefore, attributes like the name, address, purpose and associated history of the building are linked to each building object. As such this metadata may be as useful to the user as any other CH artefact and so the LoS algorithm must also be applied to the building block layer.

The solutions to these two problems however are slightly different. In the case of the CH layer, the solution to LoS determination is relatively less complicated than LoS determination of the block layer because the LoS between a viewpoint and a data point requires querying against only a single line. The LoS between the viewpoint and a polygon is more complicated to determine because the number of intersection possibilities are far greater.

To determine the LoS for both the CH and block layer a combination of Oracle Spatial operators and a LoS algorithm is required. Our solution was to take a well-known algorithm in Computer Graphics and apply it to the area of Spatial Databases. The *scanline* algorithm was chosen because the topological and Boolean operations

needed to process the algorithm are already inherent in the Oracle database schema.

- *Scanline Algorithm* – The algorithm works by making a progressive scan of the area in question (FoV) to determine weather there are any objects in the scan line path. If so the point at which the scan line intersects the object is recorded. A series of scans is carried out. The end product is the coordinates of a polygon object that represents the search space minus the surrounding building object geometries.

Implementing the scanline algorithm involves using the *sdo_intersection* operator in Oracle and is as follows: First, a series of lines are created between the viewpoint of the user and each data point present inside the view-port. For the CH layer, these lines are considered as the scanlines. In turn each of the scanlines are used as the input parameters to the algorithm, to determine if they interact with any of the objects in the block layer. If there is any interaction between any of the objects in the block layer and the scanline, the CH artefact is not visible to the user from that viewpoint and it will not be placed in the LoS resultset. If there are no interactions between the CH data point and the block layer along that scanline, the data point is considered to be visible and is placed in the LoS resultset.

Implementing the scanline algorithm for the block layer is more computationally expensive than the CH data layer. The LoS between a point and a polygon has to be determined, in which case there may be many possible LoS and until there is a positive LoS detected every possibility has to be checked. The process involves initially testing if the view-port interacts with one or more of the building objects in the block layer by using the *sdo_intersection* operator. If so, the IDs of the objects are put into an array. Oracle's *sdo_difference* operator is then used to compute the (spatial) difference between the view-port and the intersecting block object. The *sdo_difference* operator returns an *sdo_geometry* object that represents the difference polygon between the two geometries. If there are multiple intersecting objects, the next block object in the array is compared with the result of the previous *sdo_difference* operation to determine the new difference polygon. This procedure continues until all objects in the array have been processed. The result of this procedure is the view-port in Figure 4 with building blocks B1, B2, B3 and B4 removed, i.e. the shape of the enclosed white and grey space only.

The next step in the LoS algorithm is to connect each point of the resulting query space polygon object (each having an associated intersecting object ID) to the viewpoint. In turn each of these scanlines are tested until at least one scan proves negative for intersection with any other object, if so it is evident that there is indeed an

unobstructed LoS to that block object and therefore the processing of the remaining points relating to that particular object is stopped, the ID of the object is placed in the LoS resultset, and the next object is tested. The information on that block therefore is added to the list of objects in the users LoS.

In the case where no points of an intersecting object are in the users LoS an additional scan test of the object has to be made to determine for sure that the object is not currently in the user's FoV. A series of points (e.g. 1m apart) around the perimeter of the block object are calculated. Each point is connected to the viewpoint with a straight line. A scan is run with each of these scanlines to determine if there is any that has no intersection with any other block object, if they all intersect with other objects the block object is clearly not visible to the user. If there is one scan that doesn't intersect other block objects, the building object is said to be in view and is added to the users LoS resultset.

Only after all the data points and block objects are checked for LoS can the list of objects in the users LoS (objects interacting with white space in Figure 4) be supplied to the narrative engine of the system for metadata processing, to create a hyperlinked, digital story based on the events surrounding what the user can see [4].

4.   3 Dimensional Queries

Adding orientation and LoS functionality to the query greatly increases the relevance of the data being returned to the user. Even though this greatly increases the accuracy of the query to the database, the fact that it is still a 2D horizontal query leaves room for enhancement with regards to what the user can actually see in their vertical FoV.

At present the data that is contained in the CH database has two coordinates associated with it, x and y. This is sufficient because the queries that are being generated only require a 2D point set to query the data. This means that any data that is present within the query view-port will be passed back in the resultset regardless of the height of the users vertical FoV.

The human FoV spans approximately 200° horizontally taking into account for both eyes and 135° vertically [2]. This limits the amount of data that can be seen at any one time. The normal binocular FoV is 120° with and extra 70° of monocular vision (35° each side). The default angle in the our query model is 60°. The user can modify this value interactively as well as the viewing distance to whatever distance they want. The human FoV also has an angle of 60° above the direct line of sight and 75° below it. This means that the height of data in the model has to be taken into account as well as the vertical area that is being searched.

This can be partially achieved by adding an additional coordinate to each data point in the database giving it height. The z-value is then used as a clause in the spatial query to determine what data is to be subtracted from the initial result set. The view-port can also have on *offset* height value off the ground ensuring that the space being searched is a true 3 dimensional volume. Adding depth allows queries like "Are there any cultural heritage artefacts contained within the view-port in *front* of the viewer up to a height of 10 meters off the ground?" and in the second example where the view-port has a height offset, queries like "Are there any cultural heritage artefacts contained within the view-port in *front* of the viewer that are between 10 and 15 meters off the ground?".

This approach extends the query model by adding the ability to construct essentially a 3D viewer-based directional query to the search space. The data is then searched by using only topological and metric operations to do so [17]. A second order constraint is then applied to check for height and further reduce the resultset to only the data that satisfies the constraint.

4.1.   View-port Query Control

To give the users control of the desired view-port angle, radius and direction a View-port Query Control was developed (Figure 5). This small frame developed in Java gives the user total control of view-port dimensions. The user has the option to change the FoV angle at which the view-port can expand and also the radius it can be extended. The option to modify the orientation of the view-port with reference to the users orientation is also available. Useful while walking down a street but with (query) "eyes" in the back or side of your head for example.
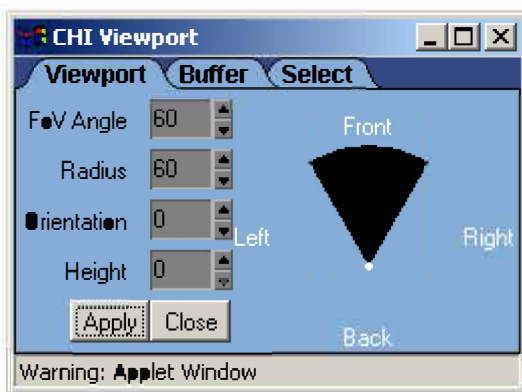

Figure 5: View-port Query Control

It was realised during the development that even though a pie shape view-port is useful it might not always

be the preferred option for the user. For example a user might want to query all around them and not just in the FoV.

To cater for this need a series of tabs were added to the View-port Control. One of these tabs is the buffer tab. There are two spinner controls on this tab that adjust the radius of the view-port to the desired size and the height. (Figure 6)
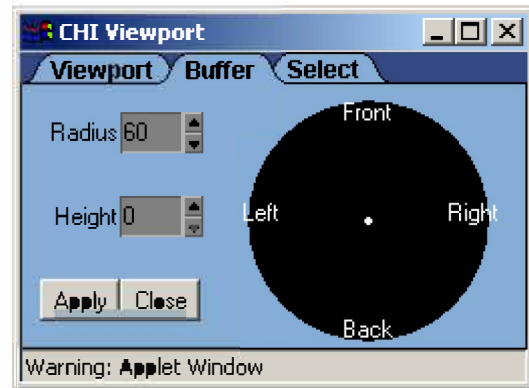

Figure 6: Buffer Control

The next tab to be added to the view-port was a static selection control that allows the user to change the orientation of the view-port in relation to the orientation of the user (Figure 7). On this tab the user has a selection of different fields of vision. The default is a human field of view at a height of 2m but a cat (.3m), dog (.5m) or rabbit (.2m) could also be selected to experience these other FoV realities.
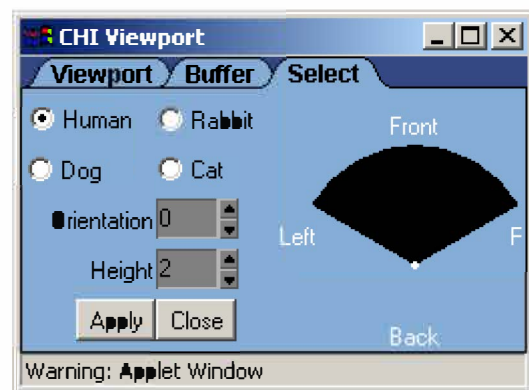

Figure 7: Select Control

5.   Implementation

An overview of the Cultural Heritage Interface (CHI) workflow, developed by the Digital Media Centre of the Dublin Institute of Technology, is illustrated in Figure 8.
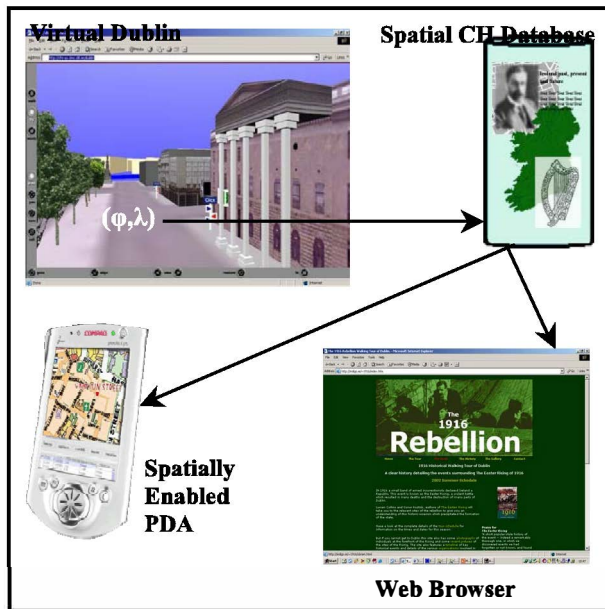
Figure 8: Conceptual Overview of CHI System

The main technical components are implemented in a three-tier web-based architecture typical of spatially enabled enterprise applications [3], i.e. it comprises three layers, namely the Client Layer, the Application Server Layer, and the Database Layer. All communications between the client layer and the database are conducted through the application server layer. All query building and query results formatting, is executed on the application server. The client communicates with the application server using the existing HTTP networking protocols.

5.1. Database Layer

The database implemented for the CHI project (Oracle 9i) provides spatial object type storage, SQL access, spatial operations, and indexing as well as map projections and coordinate systems support [15]. This is accomplished by defining the attribute information (CH hypermedia) as a spatial data type (which implies associated coordinate data). In the example below the *position* field is the spatial data type:

```
CREATE   TABLE CHI_CONTENT_DATA (
         ID NUMBER (20) PRIMARY KEY NOT NULL,
         POSITION MDSYS.SDO_GEOMETRY NOT NULL,
         STREET   VARCHAR2 (20) NOT NULL,
         BUILDING VARCHAR2 (20) NOT NULL);
```

The spatial extension to SQL allows us to use this syntax to create the above table with a spatial data type of type *sdo_geometry*. It also allows us to insert positional data into the table as follows:

```
INSERT   INTO CHI_CONTENT_DATA VALUES (4,
         MDSYS.SDO_GEOMETRY (2,NULL, NULL,
         MDSYS.SDO_GEOMETRY (2001, NULL,
         MDSYS.SDO_POINT_TYPE (919.0, 513.0), NULL,
         NULL),'OCONNELL', 'GPO');
```

The code excerpt above identifies the object geometry as a 2D Point "2001", 2 meaning 2D and 1 meaning a point datatype. The coordinates of the point are then specified in the *sdo_point_type* array. The data in the table is then indexed using the R-tree index data structure that is implemented using the extensible indexing framework of Oracle Spatial [14].

One advantage of spatial data types is that subsequent queries can be restricted to a pre-defined geographical area, e.g. within a 10m radius of a given location. By exploiting the spatial indexing mechanisms inherent to Oracle 9i, which essentially organises the information within the database tables according to their geographic location, all location relevant data is retrieved most efficiently.

The hypermedia CH objects stored in the CHI project database (together with their spatial component) comprise an "historic walking tour of Dublin". Such a tour can begin and end at specific times and places and pass specific landmarks along the way. As each of the landmarks is encountered in turn, a particular "story object" will be retrieved about its historical significance. It is the text of this "story" that will comprise the bulk of the data stored in the CHI database layer. A challenging aspect to this research is the investigation of the methodologies for retrieving these story objects both automatically and coherently as their positions in space are approached. To accomplish this task successfully, the causality of the localised series of events is considered [4]. A simple experiment compares the performance of the classical range query with the view-port directional query strategy.
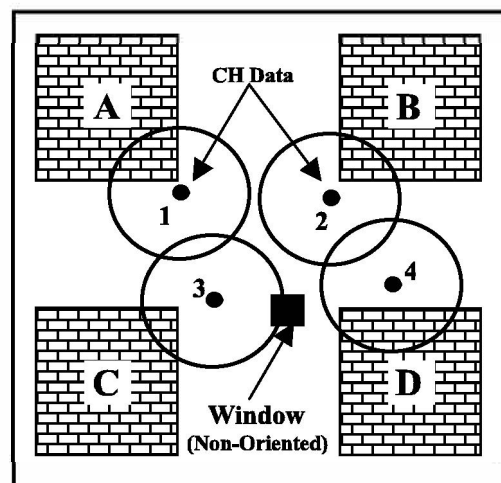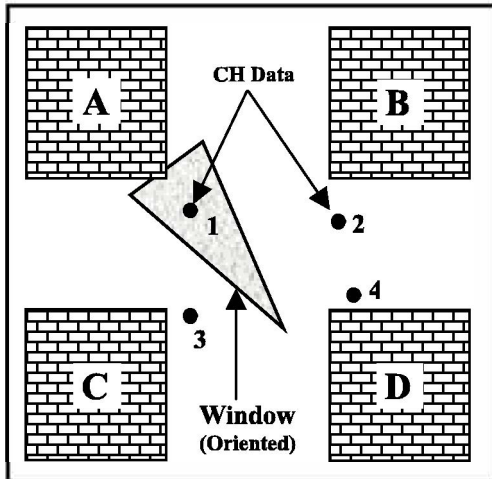

Figure 9: Non-Directional Query

Figure 10: Comparison of Query Methods

Figure 9 illustrates how the CHI spatial database was organised previously and how the data was queried using a range query. Objects A, B, C and D represent street building blocks in the system. The points surrounded by circles represent cultural heritage artefacts (CH data) within the database, each of which are represented by a 2 dimensional point with a buffer. If the Query window interacts with any of the buffers the resulting data set is extracted from the database and presented to the user. A problem with this method is as follows: if the query window is situated in data area 3 and the orientation of the user is northwest the returned data is data point 3, when the actual data should be data point 1. Similarly if the user is facing east the returned data is data 3 when it should be data 4.

Our novel solution to this problem is illustrated in Figure 10. In this example the query window is an oriented and dynamically generated triangle and a 3D point represents each CH data point. The buffers around the data points are no longer necessary because the oriented view-port is being used. If the user's viewpoint is situated in the same position as the previous example and is also facing northwest, the data that is returned is data point 1 together with Building A metadata (if any) and if the orientation is changed to be east, data point 4 is returned (if visible) plus Building D metadata.

The next test considered how accurate the queries would be when querying for height along with horizontal intersection. The SQL query tested for data that was contained within the view-port window and had a height of less than 20m and greater than 10m. This addition to the system means that layers of data can be added to the database with the same X and Y coordinates but a different Z coordinate to distinguish it from data positioned, for example, on different floors of the same building.

6. Conclusion

We have introduced a directional method of querying a spatial database system that considers the user's LoS in the context of cultural heritage information retrieval. Tests show the enhanced demonstrator performs as expected, with the relevance of the data greatly improved compared to the initial non-directional querying prototype.

The determination of the LoS of the user is only a small step in the direction of realistically querying the spatial database. The approach of utilising a scan line intersection algorithm delivers the desired results needed to determine the line of sight but a limitation is that it is specific to 2 dimensional data.

The cost of the queries in the revised demonstrator is slightly more than that of the initial prototype because there are more points needed to create the directional window. In the enhanced system the query window is constructed in real-time every five seconds if the viewpoint differs more than 5m in position or more than 30° in direction from the last query processed. The user's FoV is adjustable in horizontal angle, height and range to accompany many varied points of view

7. Future Work

The next phase of the research is to implement a perspective query *frustum* (view-pyramid) that will mimic the human FoV more accurately, i.e. in 3 dimensions. Determining the LoS of the user in 3 dimensions involves using 3D spatial indexes on 3D objects to determine if 3D data points lie inside the objects. To achieve this, Voxels [16] and Octrees [8] will be considered for indexing the 3D Objects.

We plan therefore to develop a perspective query [11] processor that will use a 3D view-pyramid to query the data taking into account the vertical FoV angle. This approach introduces the concept of querying the CH dataset based on the idea of a "birds-eye-view" of the data. Achieving this in the VRML world will be relatively straightforward, as the viewers direction rotations about the 3 axes is known, however in reality the introduction of 3D Compasses and tilt sensors embedded in next generation PDA's is needed.

The approach considers retrieving all data interacting with the projected footprint of the "floor+base" of the frustum up to any height as a primary filter and then further processing this resultset against a 3D polygonal sweep of specified dimensions from left to right to complete the query and effectively build the view-pyramid in real-time.

Acknowledgements

References

[1] Ames, A. L., Nadeau, D. R., and Moreland, J.L.: "VRML 2.0 Sourcebook": 1997: John Wiley and Sons, Inc. pp. 519-531

[2] Barfield, W., Hendrix, C., Bjorneseth, O., Kaczmarek, K. A., and Lotens, W.: 1995: "Comparison of human sensory capabilities with technical specifications of virtual environment equipment," Presence: Teleoperators and Virtual Environments.

[3] Bertolotto, M., Carswell, J. D., McGeown, L., and McMahon, J., "iSmart+iSIS: 2001: "Deploying Integrated Web-Based Spatial Applications Within an Oracle Database Environment", *International Workshop on Web Geographical Information Systems (WGIS2001)*, IEEE CS Press, Kyoto, Japan.

[4] Carswell, J. D., Eustace, A., Gardiner, K., Kilfeather, E., and Neumann, M.: 2002: "An Environment for Mobile Context-Based Hypermedia Retrieval." DEXA Workshops, 532-536.

[5] Diehl, S.: 2001: Distributed virtual worlds: foundations and implementation techniques using VRML, Java, and CORBA, Berlin, New York, Springer.

[6] Goyal, K. G., and Egenhofer, M. J.: "Similarity of Cardinal Diractions" in: Jensen, C., Schneider, B. S., Seeger. B., Tsotras, V. (eds.): 2001: *Seventh International Symposium on Spatial and Temporal Databases Lecture notes in Computer Science* Vol. 2121, Springer-Verlag, pp. 36-55.

[7] Guting, R.H.: 1994: "An Introduction to Spatial Database Systems", *VLDB*, 3:357-399.

[8] Jackins, C., Tanimoto, S. L.: 1980: "Oct-Trees and Their Use in Representing Three-Dimensional Objects", *CGIP*, 14(3), , 249—270.

[9] Liu, X., Shekhar, S., and Chawla, S.: 2003: "Object-based Directional Query Processing in Spatial Databases", *IEEE Transactions of Knowledge and Data Engineering*

[10] Liu, X., Shekhar, S., and Chawla, S.: 2000: "Processing Object-orientation-based Direction Queries", in the proceeding of the *Eighth International Symposium on Advances in Geographic Information Systems*, 69-76, Washington D.C.

[11] Li, J., Jing, N., and Sun, M.: "Spatial Database Techniques Oriented to Visualisation in 3D GIS." School of Electronic Science and Engineering, National University of Defense Technology.http://www.digitalearth.ca/html_papers /DE_A_064.htm

[12] Papadias, D., Theodoridis, Y., and Sellis, T.: 1994: The Retrieval of Direction Relations Using R-trees. Proc. *5th Int'l Conference on Database and Expert Systems Applications, DEXA'94*, Athens, Greece. Springer - Verlag LNCS #856.

[13] Pfoser, D., Jensen, C. S., and Theodoridis, Y.: 2000: "Novel Approaches in Query Processing of Moving Objects." *In proc. Intl. Conf. on Very Large Databases (VLDB)*.

[14] Ravi, K. V. K., Ravada, S., Abugov, D.: 2002: "Quadtree and R-tree Indexes in Oracle Spatial: A Comparison using GIS Data": 2002: Proceedings of the ACM SIGMOD *international conference on Management of data*: Pages: 546 – 557.

[15] Rigaux, P., Scholl, M., and Voisard, A.: 2002: "Spatial Databases With Application to GIS", *Academic Press, Morgan Kaufmann Publishers*. pp. 352-360

[16] Samet H.: 1990: "Application of Spatial Data Structures". *Addison-Wesley*.

[17] Shekhar, S., and Chawla, S.: 2003: "Spatial Databases: A Tour", *Prentice Hall Publishers*: pp. 28.

[18] Shekhar, S., Chawla, S., Ravada, S., Fetterer, A., X. Lui, and Lu, C.T.: 2003: "Spatial Databases: Accomplishments and Research Needs", *IEEE Trans. on Knowledge and Data Engineering*.

[19] Shekhar, S., Liu, X., and Chawla, S.: 2003: "An Object Model of Direction And Its Implications", GeoInformatica, 3(4), 357-379, Kluwer Academic Publishers.

[20] Theodoridis, Y., Papadias, D., and Stefanakis, E.: 1996: "Supporting Direction Relations in Spatial Database Systems." *Proc. 7th Int'l Symposium on Spatial Data Handling, SDH'96*, Delft, The Netherlands.

[21] VRML Consortium: http://www.vrml.org/

5.8. Co-Author Declarations

I attest that PhD candidate Keith Gardiner was a main contributor to the paper/publication entitled - *MobiLAudio - A Multimodal Content Delivery Platform for Geo-Services*.
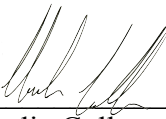
_____   _____   02/07/2015
Charlie Cullen   James D. Carswell   Date

I attest that PhD candidate Keith Gardiner was a main contributor to the paper/publication entitled - *mobiSurround: An Auditory User Interface for Geo Service Delivery*.

_____   _____   02/07/2015
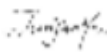Charlie Cullen   James D. Carswell   Date

I attest that PhD candidate Keith Gardiner was a significant contributor to the paper/publication entitled - *Mobile Visibility Querying for LBS.*

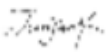_____   _____   02/07/2015
James D. Carswell   Junjun Yin   Date

I attest that PhD candidate Keith Gardiner was a main contributor to the paper/publication entitled - *EgoViz – A Mobile Based Spatial Interaction System.*

_____   _____   **02/07/2015**
James D. Carswell   Junjun Yin   **Date**

I attest that PhD candidate Keith Gardiner was a significant contributor to the paper/publication entitled - *A Web-Based and Mobile Environmental Management System.*
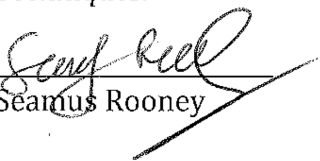
_____   _____   _____
James D. Carswell   Andrea Rizzini   Michela Bertolotto

_____   02/07/2015
Nick Mandrak   Date

237

I attest that PhD candidate Keith Gardiner was a significant contributor to the paper/publication entitled - *An Open Source Approach to Wireless Positioning Techniques.*

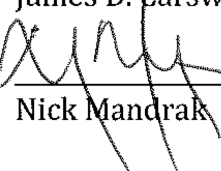Séamus Rooney   James D. Carswell   02/07/2015
Date

I attest that PhD candidate Keith Gardiner was a main contributor to the paper/publication entitled - *Viewer-Based Directional Querying for Mobile Applications.*

James D. Carswell   02/07/2015
Date

5.9. User Trial Questionnaire

Digital Media Centre

User Trial Footsteps of Bloom

Prepared for: MobilAudio
Prepared by: Keith Gardiner

Digital Media Centre   Aungier Street, Dublin 2, Ireland   **T** +353 (0)1 4023268   **E** keith.gardiner@dit.ie   Work URL: www.dmc.dit.ie

240

Participant Entry Form

You have been asked to be a participant in a user trial, which is part of an Enterprise Ireland funded project in the Dublin Institute of Technology (DIT). The DMC research team will ask you questions for research purposes and the collected data and findings may possibly feature in future research publications. As well as this, the team may ask your permission to use images or video of you carrying out the task for dissemination purposes relating to the project. If you are asked, the content will be shown to you in advance for your approval.

The DMC research team will endeavour to do the following:

- To protect the welfare and dignity of the participant.
- To respect the individual's freedom to decline participation.
- To maintain confidentiality of the research data.
- To be responsible for maintaining ethical standards.
- To NOT specifically identify individuals with their data unless it is necessary, and then only after the individual has given consent.
- To take every precaution and make every effort to minimise potential risk to participants.
- To only use the data supplied by the participant with their full consent.

I herby give my consent to the DMC research team to use any data supplied by me on the accompanying form for research purposes and possible further publications:

Signature	Date

Pre-test Questions

The following short questions may be of relevance to the results of the trial, You can choose whether or not you answer any/all of them.

Question (tick as appropriate)	Yes	No
Have you participated in a user trial before?		
Have you any hearing problems?		
Have you ever taken an audio tour?		
Are you a native English speaker?		
Are you a resident of Dublin?		
Have you read any of Joyce's work?		
Do you own a smartphone?		

Overview

Outline

Thank you for taking part in the MobilAudio User Trial. You will be asked to participate in a short audio walking tour of Dublin City Centre, using the mobile device provided. At the end of the tour, you will be asked some short questions about your experience. If you have any problems, you can stop the tour at any time and get in direct contact with the trial supervisors using the number provided on the mobile.

About the tour

The "Footsteps of Bloom" tour is a short interactive audio tour through the Centre of Dublin, starting at the O'Connell Street spire. The application uses headphones to deliver an audio tour using a mobile device, and will provide you with sound effects, music and spoken information from Leopold Bloom about some of the sights (and sounds) of Ulysses Dublin. The tour is fully automated, so you are asked to start the application near the O'Connell Street Spire and then keep the mobile in your pocket or bag for the duration of the tour.

Questions

At the end of the tour, you will be met by a trial supervisor who will take the equipment from you and ask you a few quick questions about your experience. Thank you for assisting us in our research.

Functionality Questions

The following short questions relate to your use of the application, please answer yes or no, and comment where you think necessary:

Question	Yes	No	Comments
Could you hear the sound effects in the application?			
Could you hear the music in the application?			
Could you hear the speech in the application?			
Did the application direct you properly to each place?			
Did the sound effects happen at the right time?			
Did the music change during the tour?			
Did the application function automatically?			
Did you use the application interface for any reason?			

User Experience Questions

The following short questions relate to your experience of the application, please answer yes or no, and comment where you think necessary:

Question	Yes	No	Comments
Was the application engaging?			
Was the application easy to use?			
Would you pay for this application? If so, how much?			
Did you like the Bloom tour?			
Would you use other similar tours?			
Would you like to share this tour on Facebook?			
Would you like to create your own tour?			
Would you like to create tours with friends?			
Would you buy a tour recommended by a friend?			

Question	Comments
What didn't you like about the application?	
What did you like about the application?	
Is there anything else you want to add?	

Post-Test Questions

Please mark a point on each scale:

How much **Mental Demand** did you experience during the trial?

Low								High

How much **Physical Demand** did you experience during the trial?

Low								High

How much **Time Pressure** did you experience during the trial?

Low								High

How well do you think you **Performed** during the trial?

Good								Bad

How much **Effort** did it take to do the trial?

Low								High

How **Frustrating** did you find the trial?

Low								High