Conference papers

Digital Media Centre

2015-09-01

mobiSurround: An Auditory User Interface for Geo-Service Delivery

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

Charlie Cullen
Technological University Dublin, charlie.cullen@tudublin.ie

James Carswell
Technological University Dublin, james.carswell@tudublin.ie

mobiSurround: An Auditory User Interface for Geo-Service Delivery

Keith Gardiner[✉], Charlie Cullen, and James D. Carswell

Digital Media Centre, Dublin Institute of Technology,
Dublin, 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 [1–8]. 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 [10–13], 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 Sect. 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 Sect. 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 [14–16]. 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 orientation 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 Sect. 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 modelling (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 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 and delivery of media content in a coherent manner (Fig. 1).
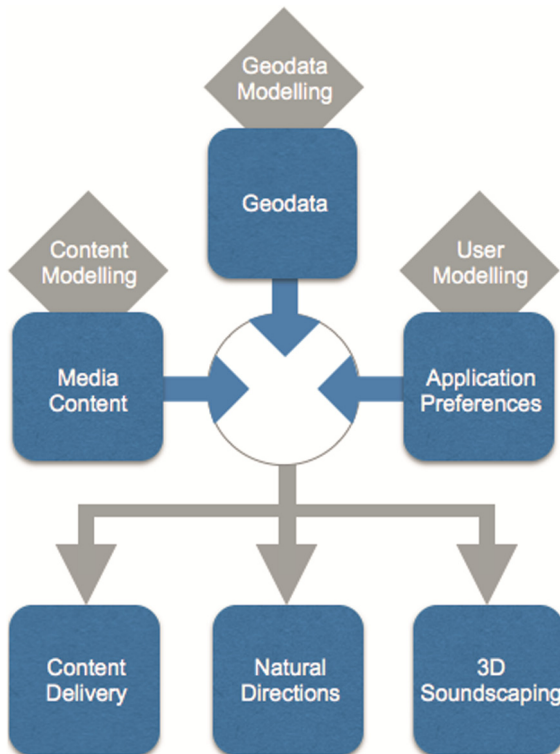

Fig. 1. 3-stage data modelling process includes user, context and content modelling for effective data management and content organization

The following sections outline our innovative approach to modelling for geo-services and describe how the 3-stage process integrates to deliver the functionality described. In particular, aspects relating to content modelling 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 (Fig. 2).
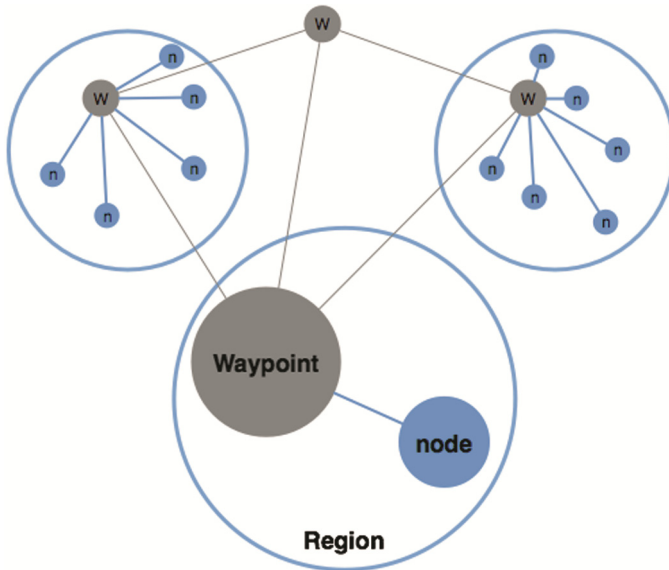

Fig. 2. Data mapping structure describing node, waypoint and region nodes for navigation

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 Figs. 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 Fig. 3.
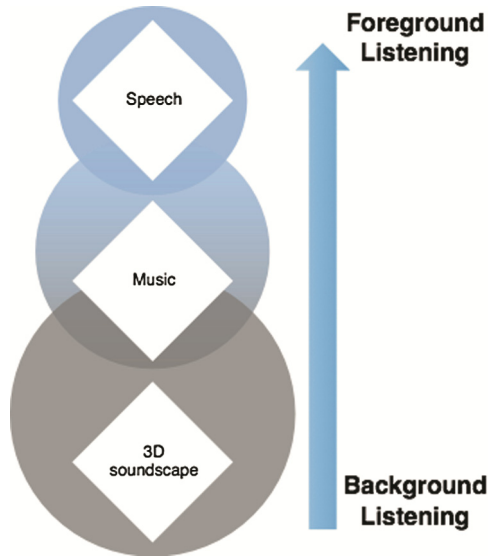

Fig. 3. mobiSurround background to foreground listening model that prioritises speech over background audio and music

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 Fig. 4.
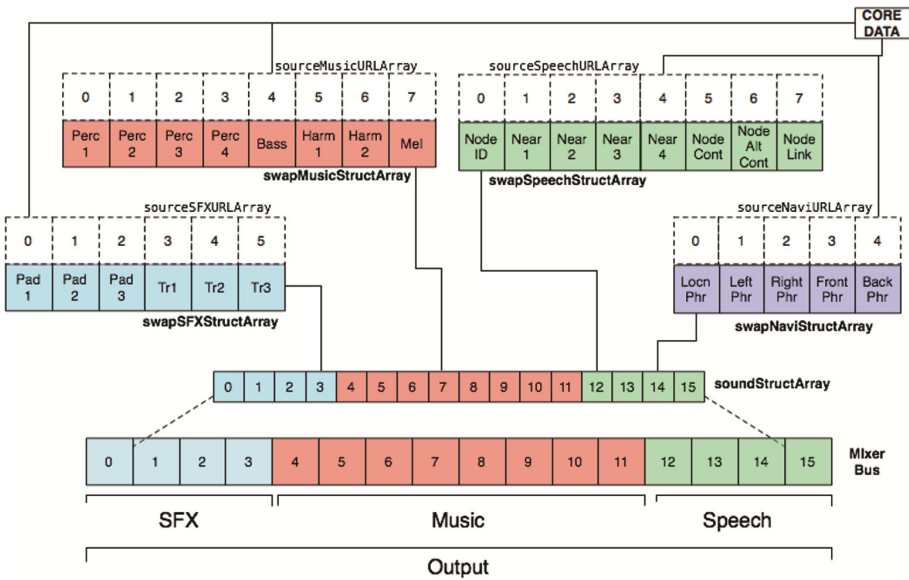

Fig. 4. Mixerbus multichannel audio graph controls all audio playback in the application based on a set of rules

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]
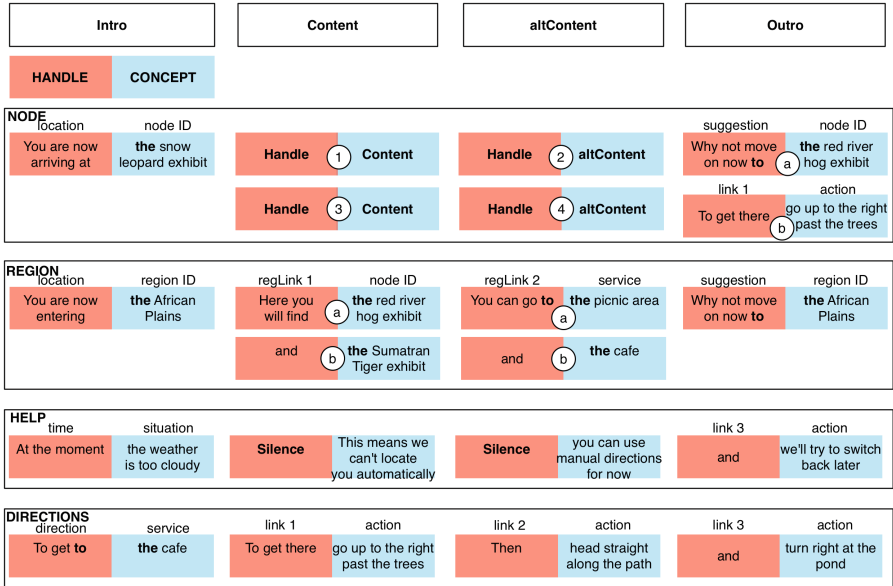

Fig. 5. Phrase synthesis structure for node, region and directions content

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 AUGraph

		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.
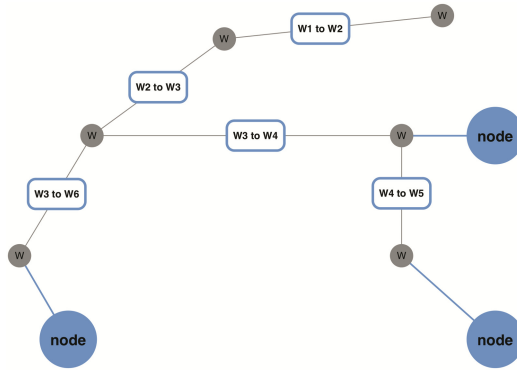
Fig. 6. Waypoint linkage example for node and waypoint navigation

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.

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 (Fig. 7).

Fig. 7. *mobiSurround* location manager (demonstration using Dublin Zoo tour).

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 (Fig. 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.

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 (Fig. 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.
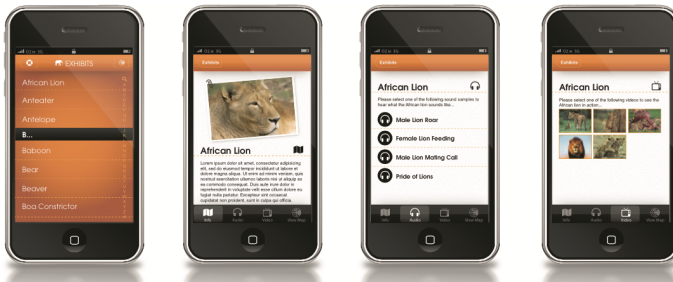

Fig. 8. mobiSurround GUI used in manual mode enables access to POI, facility and event information when accurate positioning data is unavailable

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 *AUGraph* render model to determine whether a real-time Fast Fourier Transform (FFT) convolution could be performed using the iOS render callback model.

The results (Fig. 9) define a cardioid 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.
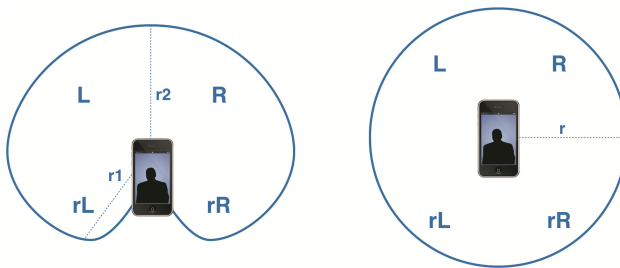

Fig. 9. Visual summary of virtualiser perception test results (Comparing ideal and actual)

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 a 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.

References

1. Fröhlich, P., Simon, R., Baillie, L., Anegg, H.: Comparing conceptual designs for mobile access to geo-spatial information. In: Proceedings of the 8th Conference on Human-Computer Interaction with Mobile Devices and Services – MobileHCI, p. 109 (2006)
2. Persson, P., Espinoza, F., Fagerberg, V., Sandin, A., Cöster, R.: GeoNotes: a location-based information system for public spaces. In: Höök, K., Benyon, D., Munro, A.J. (eds.) Designing Information Spaces: The Social Navigation Approach Computer Supported Cooperative Work, pp. 151–173. Springer, London (2003)
3. Robinson, S., Eslambolchilar, P., Jones, M.: 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. ACM, Amsterdam, The Netherlands (2008)
4. Simon, R., Frohlich, P.: A mobile application framework for the geospatial web. In: Proceedings of the 16th International Conference on World Wide Web, pp. 381–390. ACM, Banff, Alberta, Canada (2007)
5. Simon, R., Fröhlich, P., Gerhard Obernberger, G., Wittowetz, E.: The point to discover GeoWand. In: Proceedings of the 9th International Conference on Ubiquitous Computing, Innsbruck, Austria (2007)

6. Simon, R., Frohlich, P., Grechenig, T.: GeoPointing: evaluating the performance of orientation-aware location-based interaction under real-world conditions. J. Locat. Based Serv. **2**(1), 24–40 (2008)
7. Frohlich, P., Simon, R., Baillie, L.: Mobile spatial interaction. Pers. Ubiquit. Comput. **13**(4), 251–253 (2009)
8. Strachan, S., Murray-Smith, R.: Bearing-based selection in mobile spatial interaction. Pers. Ubiquit. Comput. **13**(4), 265–280 (2009)
9. McGee, J., Cullen, C.: Vocate: auditory interfaces for location-based services. In: Proceedings of the 23rd Conference on Computer Human Interaction, Cambridge, United Kingdom (2009)
10. Carswell, J.D., Gardiner, K., Bertolotto, M., Rizzini, A., Mandrak, N.: A web-based and mobile environmental management system. J. Environ. Inform. **12**, 9–20 (2008)
11. Carswell, J.D., Gardiner, K., Neumann, M.: Wireless spatio-semantic transactions on multimedia datasets. In: Proceedings of the 2004 ACM Symposium on Applied Computing, pp. 1201–1205. ACM, Nicosia, Cyprus (2004)
12. Gardiner, K., Carswell, J.D.: Viewer-based directional querying for mobile applications. In: Third International Workshop on Web and Wireless Geographical Information Systems W2GIS. IEEE CS Press, Rome, Italy (2003)
13. Carswell, J.D., Eustace, A., Gardiner, K., Kilfeather, E., Neumann, M.: An environment for mobile context-based hypermedia retrieval. In: Proceedings 13th International Workshop on Database and Expert Systems Applications (2002)
14. Salton, G., McGill, M.J.: Introduction to Modern Information Retrieval. McGraw-Hill Computer Science Series, xv, p. 448. McGraw-Hill, New York (1983)
15. Schiller, J., Voisard, A.: Location Based Services. Morgan Kaufmann Publishers Inc., San Francisco (2004)
16. Jiang, B., Yao, X.: Location-based services and GIS in perspective. Comput. Environ. Urban Syst. **30**, 712–725 (2006)
17. Mohapatra, D., Suma, S.B.: Survey of location based wireless services. In: IEEE International Conference on Personal Wireless Communications (ICPWC) (2005)
18. Rao, B., Minakakis, L.: Evolution of mobile location-based services. Commun. ACM Mob. Comput. Oppor. Chall. **46**(12), 61–65 (2003)
19. 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 their Applications to Cultural Heritage (2005)
20. Simon, R., Kunczier, H., Anegg, H.: Towards orientation-aware location based mobile services. Inf. Syst. J. p. 1–8 (2007)
21. Wilson, A., Shafer, S.: XWand: UI for intelligent spaces. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 545–552. ACM, Ft. Lauderdale, Florida, USA (2003)
22. Strachan, S., Williamson, J., Murray-Smith, R.: Show me the way to Monte Carlo: density-based trajectory navigation. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1245–1248. ACM, San Jose, California, USA (2007)
23. Gardiner, K., Yin, J., Carswell, J.D.: EgoViz – a mobile based spatial interaction system. In: Carswell, J.D., Fotheringham, A., McArdle, G. (eds.) W2GIS 2009. LNCS, vol. 5886, pp. 135–152. Springer, Heidelberg (2009)
24. Carswell, J.D., Gardiner, K., Yin, J.: Mobile visibility querying for LBS. Trans. GIS **14**, 791–809 (2010)
25. Benedikt, M.L.: To take hold of space: isovists and isovist fields. Environ. Plan. **6**, 47–65 (1979)

26. Carswell, J.D., Gardiner, K., Yin, J.: 3DQ: threat dome visibility querying on mobile devices. GIM Int. **24**(8), 24 (2010)
27. Williamson, J., Murray-Smith, R., Hughes, S.: Shoogle: excitatory multimodal interaction on mobile devices. In: Conference on Human Factors in Computing Systems, p. 121 (2007)
28. Truax, B.: Soundscape, acoustic communication and environmental sound composition. In: Contemporary Music Review, pp. 49–65 (1996)
29. Misra, A., Cook, P.R., Wang, G.: A new paradigm for sound design. In: Proceedings of the International Conference on Digital Audio Effects, pp. 1–6 (2006)
30. Schoenberg, A.: Fundamentals of Music Composition. Faber and Faber, London (1967)
31. Apple. iOS developer library (Core audio), 2015 February. https://developer.apple.com/library/ios/documentation/MusicAudio/Conceptual/CoreAudioOverview/CoreAudioEssentials/CoreAudioEssentials.html. Accessed 2015