Ontology and Rule based Retrieval of Sound Objects in Augmented Audio Reality System for Museum Visitors

Hatala, Marek; Kalantari, Leila; Wakkary, Ron, & Newby, Kenneth

Abstract. *ec(h)o* is an "augmented reality interface" utilizing spatialized soundscapes and a semantic web approach to information. The initial prototype is designed for a natural history and science museum. The platform is designed to create a museum experience that consists of a physical installation and an interactive virtual layer of three-dimensional soundscapes that are physically mapped to the museum displays. The source for the audio data is digital sound objects. The digital objects originate in a network of object repositories that connect digital content from one museum with other museums collections on the network. The interface enables people to interact with the system by movement and object manipulation-based gestures *without* the direct use of a computer device. The focus of this paper is the retrieval mechanism for the sound objects for the museum visitor. The retrieval mechanism is built on the user model and conceptual descriptions of the sound object and museum artifacts in the form of ontologies for sound and psychoacoustics, topic ontology and Conceptual Reference Model for museum information. The retrieval criteria are represented as inference rules that represent knowledge from psychoacoustics, cognitive domain and composition aspects of interaction. The system will be demonstrated in exhibition space in Nature Museum in Ottawa in December 2003.

Introduction

Audio museum guides have existed for some time. For example, museum docents guide museum visitors in groups through collections providing contextual and supplemental information for what is on display. Within the limits of a group activity and the knowledge level of the individual docent, a modicum of interactivity and further knowledge exploration is possible. Audio guides in the form of audio-cassettes have also been available as a means of overcoming the scheduling inflexibility of group tours by docents. While beneficial in many respects, the audio guides are limited by their linear sequence and non-interactive structure. Bedersen [3] developed a prototype utilizing portable mini-disc players and an infra-red system to allow museum visitors to explore at their own pace and sequence. As museum visitors approached artifacts on display, relevant audio information would be triggered on the mini-disc player and heard through headphones. *Hyperaudio* [14] provided visitors with palmtop computers and developed specific user models for adaptive systems within a museum setting. *MEG* [2] is a portable digital museum guide for the Experience Music Project in Seattle that allows visitors 20 hours of audio and video on demand. Visitors make their selections either by use of the keyboard within the PDA device or by pointing the device at transmitters located adjacent to artifacts.

In the previous works, the relationship of the digital content to the artifacts is either pre-planned and fixed, or the digital content is not networked and limited to the local device, in some cases both limits are true. *ec(h)o* employs a semantic web approach to the museum's digital content thus it is networked, dynamic and user-driven. The interface of *ec(h)o* does not rely on portable computing devices, rather it utilizes a combination of gesture and object manipulation recognized by a vision system.

The dynamic and user-driven nature of the ec(h)o requires a highly responsive retrieval mechanism with a criteria defined by psychoacoustics, content and composition domains. The retrieval mechanism is based on user model that is continually updated as user moves through the exhibition and listens to the audio objects. The criteria a represented by rules operating on the ontological descriptions sound object, museum artifacts and user interests.

The paper is organized as follows. We first present a user scenario and analyze ec(h)o's system needs. This is followed by a description of the overall ec(h)o architecture. Next we describe the retrieval mechanism in ec(h)o and provide details of the semantic object descriptions, retrieval criteria, user model and describe retrieval mechanism. Finally, we provide details on the current stage of the system implementation and conclude by highlighting our contributions.

Analysis of ec(h)o needs

The platform for ec(h)o is an integrated audio, vision and location tracking system installed as an augmentation of an existing exhibition installation. The platform is designed to create a museum experience that consists of a physical installation and an interactive layer of three-dimensional soundscapes that are physically mapped to museum displays and the overall exhibition installation.

Each soundscape consists of zones of ambient sound and "soundmarks" generated by dynamic audio data that relates to the artifacts the visitor is experiencing. The soundscapes change based on the position of the visitor in the space, their past history with viewing the artifacts, and their individual interests in relation to the museum collection. By way of a gesture-based interaction, the visitor can interact with a single artifact or multiple artifacts in order to listen to related audio information. The audio delivery is dynamic and generated by agent-assisted searches inferred by past interactions, histories and individual interests.

The source for the audio-data is digital objects. In the case of ec(h)o, we developed a large sample set of digital objects that originated from the partner museums. These digital objects were used to populate the network of object repositories. This allows a visitor in the Canadian Museum of Nature in Ottawa to seamlessly access data from the digital object repositories of the Nature Museum in Ottawa and the Museum of Anthropology at UBC in Vancouver. The two institutions are physically several thousand kilometers apart.

Scenario

The following is a script for a scenario of the user experience we produced while planning the design of the ec(h)o system:

> Visitors to the ec(h)o environment wear wireless headphones for audio, are given an ec(h)o object interaction, and a digital tag for position tracking.
>
> ec(h)o is based on two modes of interaction. In the **first mode** the visitor interacts by moving through the exhibition space. This movement creates a soundscape of ambient sounds and sound marks. A zone of ambient sound could be the sound of celebration as one passes a collection of potlatch artifacts. Sound marks are sounds particular to a single artifact. A sound mark for a bird totem might be the sound of an eagle.
>
> A visitor becomes immersed in ambient sounds that provide an audio context for the collection of objects nearby. A sound mark would beckon the visitor to investigate an artifact in greater depth.

The **second mode** of interaction occurs when a visitor enters the space immediately surrounding an artifact. A visitor is presented with audio cues to their left side, center, and right side. The audio cues are questions that correspond to audio objects.

A visitor selects one of the three audio cues by rotating the ec(h)o-object. This gesture selects and activates an audio object that enriches the visitor's experience of the artifact. For example, the visitor chooses the audio cue on the left. He hears a story from the Haida oral tradition about the seawolf.

After listening to the left audio object, the visitor is presented with three new audio cues.

These audio cues lead to further exploration related to the information heard previously. He then chooses the object on the right. He hears a story about the raven.

Ec(h)o requirements

As is evident in the scenario, one of the main goals of ec(h)o is to achieve an enhanced experience for the museum visitors without inserting an extra layer of technology between the visitor and the museum exhibit. In order to facilitate ec(h)o's minimal interface we have streamlined the navigational choices and relied on simple mapping of the interaction options to a three-sided object[1]. In addition to achieving our goal of not intervening with an additional layer of technology, this interaction approach dramatically reduces the cognitive load for the user.

The user's movement through the exhibition space has a direct influence on a soundscape presented to the user. As a user moves through the space the soundmarks create a fluid environment inviting people to explore artifacts in the exhibition. To achieve this type of audio experience the overall system must be integrated with a position tracking system that has a frequent update cycle and a high level of spatial resolution. The user movement through the space is also indicative of the user interests and user behavior. When user slows down or stops in front of the artifact it can be considered as an indication of user interest in the topics presented in the artifact. A pattern of the user movement can indicate the type of the museum visitor [18] as well as user intentions [15].

We have identified the following requirements for the appropriate sound objects:

C1. Content-relevant to the viewed artifact

C2. Content-relevant to the user interests

C3. Inviting to explore other areas

C4. Plausible from the psychoacoustics perspective

In addition to the criteria for an individual objects the following criteria apply to the sequence of the objects offered to the user:

C5. Provide for exploration of a subject in depth

C6. Provide for the fluidity in experience both in content and sound experience

---

[1] The choice of an object have refined through series of workshops with attendees from academic community, students and general public.

C7. Provide a mix of informational and entertaining objects

To be able to satisfy the above criteria a rich semantic model is needed. We have identified the following information as essential for the ec(h)o:

- the content description of the user interests (user model), sound objects and museum artifacts
- psychoacoustics and sound characteristics of the sound objects
- sequencing models of an interaction

In ec(h)o we have opted for a combination of ontologies and forward chaining rules to support the retrieval mechanism. In the core of the retrieval mechanism is an inference engine that applies rules to the comprehensive user model of the user and ontological descriptions of sound objects and museum artifacts.

Ec(h)o architecture

The ec(h)o architecture consists of four independently functioning modules: position tracking module, vision module, sound delivery module, and reasoning module. Two main types of events trigger the communication between the modules: user's movement through the exhibition space and user's explicit selection of the sound objects. The ec(h)o high level architecture is shown in Figure 1.

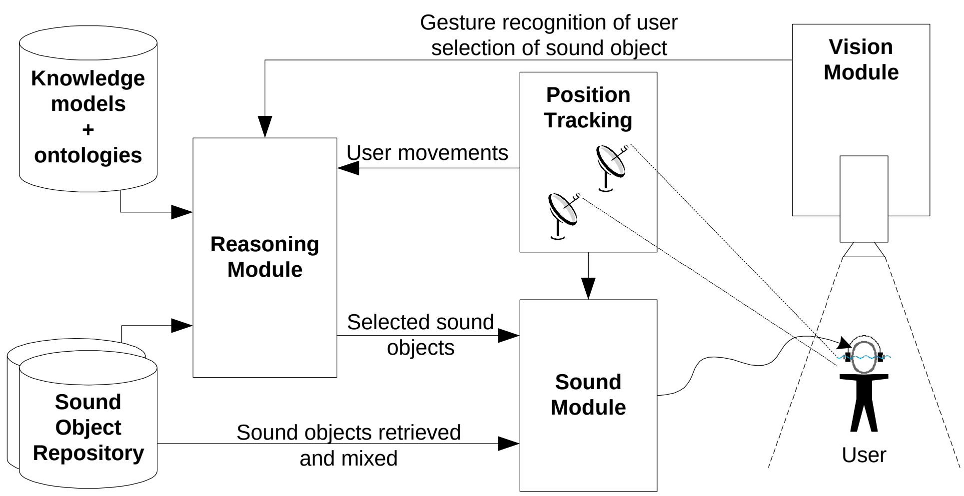

Figure 1 ec(h)o high level architecture

Position tracking module. The users in ec(h)o environment wear a tag and their movement is tracked by a radio frequency position tracking system. The user position information is fed to the reasoning module which infers intentions (such as 'crossing the room') and profile of the museum visitor. The sound system uses position information to smoothly change the sounds attached to the artifacts as user moves through the exhibition.

Sound module. Two types of sound are delivered to the user: soundscapes and sound objects. A soundscape is a sound attached to the exhibition artifact and is selected based on user interests. Once selected, the soundscape does not change during the user's visit but the radius (reach) of the sound can change based on the changing user interests. By increasing the radius users are driven to the artifacts they may find interesting. The sound module uses the user's position information to present appropriate and possible overlapping soundscapes. Sound objects represent a second type of sound delivered by the sound module. When the user stops in front of the artifact the reasoning module selects the three 'most relevant' sound objects based on the user interests. The sound system presents the available objects with audio icons and when the user makes a selection (via visual gesture recognition) the sound module retrieves and delivers the sound object to the user.

Visual module. The user selects the sound object by gesturing with an ec(h)o object (currently a cube with colored sides). By using a visual module we enable the user to freely move around an artifact and still be able to interact with the system.

Reasoning Module. The reasoning module is responsible for retrieval of the sound objects offered to the user. The sound objects are selected based on the user interests, user intentions and user's interaction history. The knowledge base is encoded in forms of the rules and sound objects are richly described by several ontologies, including topic ontology, Conceptual reference model (CRM), psychoacoustics ontology and sound ontology.

Sound Object Retrieval Mechanism

Two mechanisms contribute to an accurate retrieval of sound objects in ec(h)o: the user model and ontology descriptions of objects. As mentioned in the analysis part of the paper user's interaction space is limited to three sound objects. This poses extreme requirements on the retrieval mechanism as there is no recourse once the 'bad' choices are made. For the reasoning module to make the best suggestions of sound objects, the objects have to be richly annotated with the semantic information.

Semantic Description of Objects

To achieve plausible experience for the museum visitor it is mandatory that sound objects presented to the user satisfy several criteria. Domains we used to specify criteria for the plausible experience are those of narrative (storytelling), psychoacoustics (sound perception), augmentation (continuity with the physical environment) and sound.

Ontologies for describing content

The interaction model is based on the semantic description of the content of the objects. We have developed an ontology where a sound object is described using several properties. As an ability to link to other museums is an important feature of ec(h)o our ontology builds significantly on the standard Conceptual Reference Model (CRM) for heritage content developed by CIDOC [5]. The CIDOC Conceptual Reference Model (CRM) provides definitions and a formal structure for describing the implicit and explicit concepts and relationships used in cultural heritage documentation. To describe sound objects we use CRM `TemporalEntity` concept for modeling periods and events and `Place` for modeling locations. We describe museum artifacts using full CRM model.

The second important component of our ontology is the topic ontology. The topic ontology is a concept map of the domain that is captured in the sound objects available in the museum collection. Since the collections in individual museums are different so are the concept maps describing these collections. To facilitate the mappings between topic ontologies in individual museums we have mapped the concepts to the Dewey Decimal Classification [7] whenever possible. We discuss the networking between the museums in the section below.

Table 1 shows content related properties[2] with their domains and ranges.

Table 1 Content related properties

Property	Domain	Range
hasTopicFrame	SoundObject	Topic from Topic ontology attached to the sound keyframe
hasTopic	MuseumArtifact	Topic from Topic ontology demonstrated by a museum object
relatesToTemporalEntity	SoundObject MuseumArtifact	CRM_TemporalEntity (can be Period or Event described by object)
relatesToPlace	SoundObject MuseumArtifact	CRM_Place
describesArtifact	SoundObject	MuseumArtifact (Refers to artifact iff sound object is specific to the artifact)

Psychoacoustics and sound characteristics

The auditory interface of ec(h)o follows an ecological approach to the sound composition. It provides the basic mechanisms of navigation and orientation within the information space. These mechanisms form the key components of a modeled acoustic ecology that takes into account the variety, complexity and balance of the informational soundscape. Three areas are taken into account: psychoacoustic, cognitive, and compositional problems in the construction of a meaningful and engaging interactive audible display. Psychoacoustic characteristics of the ecological balance include spectral balancing of audible layers. Cognitive aspects of listening contribute to the design and effective use of streaming mechanisms allowing segmentation, selection and switching among audible semantic objects within the soundscape. Compositional aspects are addressed in the form of the orchestration of an ambient informational soundscape of immersion and flow that allows for the interactive involvement of the visitor.

The psychoacoustics and sound ontology define the characteristic of the sound objects that are used by the composition rules. Table 2 and Table 3shows properties defined in the psychoacoustics and sound ontologies for the sound object.

Table 2 Psychoacoustic properties defined for the Sound Object

Property	Range Values
actorAge	Child | Young | Adult | Old | VeryOld
actorGender	Female | Male | Neutral
actorVoicePitch	Deep | Normal | High
hasTempo	Slow | Medium | Fast
hasTone	Disturbing | Active | Neutral | Calm | Passive
hasMotif	Values specific to the sound collection
hasPerceptableLoudness	VeryQuiet | Queit | Normal | Loud | VeryLoud
hasSource	SourceTypeValue (i.e. AnimalSound, HumanEnvironmnetSound)

[2] To enable the system to relate sound objects to exhibition artifacts exhibition ontology defines exhibition artifacts as a subclass of an content object. Effectively this provides an exhibition object with the same content descriptive properties as sound objects.

Table 3 Properties defined for the Sound concept in the sound ontology

Property	Range Values
`isSoundEffect`	`Yes | No`
`isForegroundSound`	`Yes | No | Maybe`
`hasDirectionality`	`Ambient | PointSource | Stereo`
`spectralDensityCenter`	<number>
`spectralDensityWidth`	<number>

Composition Criteria

The composition criteria are embedded in the system in the form of an interaction model. The interaction model defines how sound objects are selected for the user based on object content and level in the interaction tree. As we want to achieve a plausible experience we have based the composition criteria in ec(h)o on Bordwell's narrative and nonnarrative formal systems of film forms [4]. The overarching formal system applicable to the museum experience is a nonnarrative categorical form, but we have included elements of nonnarrative associational and narrative variation, similarity and repetition forms.

The model can be summarized as follows:

1. Select top level sound objects from the categories with the highest user interest values [categorical]

2. Position sound objects from the same categories at the same location (left, center, right) [repetition]

3. At next levels select objects that expand at least one more category in addition to the top level category with the next highest user interest values [categorical, variation, similarity]

4. After reaching certain level (currently set to 4) select an object from category with low user interest value to support exploration [categorical, variation, association]

Above principles are implemented as forward chaining rules in the reasoning module.

The user model

In the core of the ec(h)o's reasoning module is a user model [19] that is continually updated as user moves through the exhibition and listens to the audio objects.

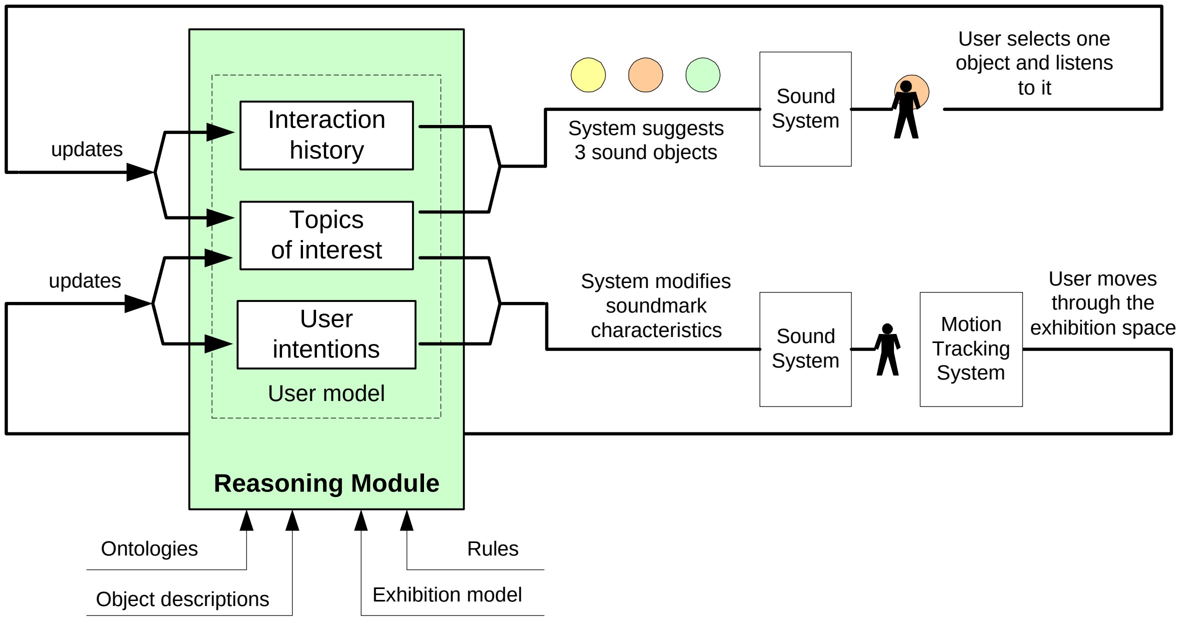

Figure 2 Interaction of user model with other modules

Figure 2 shows an interaction schema of the user model with other modules. There are two main update sources in the system. First, as the user moves through the exhibition the speed of the movement and stops or slowing down at different artifacts provide updates to the user model. The user behaviors are computed based on the speed and homogeneity of the user movement. The stops and slowing down in front of an artifact are interpreted as an interest in topics represented by the artifact. The user interests and intentions influence the presentation of soundmarks. For example, soundmark radius and volume is increased for those artifacts that correspond with current user interests. Another example can be reducing the number of soundmarks in the exhibition if user's recognized intention is to quickly cross the room.

The second source of updates to the user model considers user's direct interaction when user selects a sound object to listen to. In the model this maps to an increased user interest in topics presented by the sound object and updates to the user's interaction history. We describe the user model and retrieval mechanism in detail below.

As mentioned above, the user model consists of three main parts: *Interaction history*, *User Interest*, and *User Behavior*.

Interaction History

Interaction History is a record of how the user interacts with the ec(h)o-augmented museum environment. Two types of events are stored in the interaction history: the user's movement and user's selection of objects. The user path through the museum is stored as discrete time-space points of locations on the path. A time-space point is represented as a fact:

```
(user-location (user-id john) (x-position 10.7) (y-position 11.5) (time 172.0))
```

The user model correlates the user locations with the exhibition physical model to calculate the relative location of the user to the artifacts/exhibitions. Also, the speed of the user and how much time the user spends in front of the artifact is determined and used to infer a type of user's behavior.

Second type of information stored in Interaction History is user's selections in the form of URLs of sound objects selected by the user.

```
(user-selection (user-id john) (sound-object http://echo/narratives/123.mp3)
(in-front-of artifact-1) (time 184))
```

This information is essential for several tasks ranging from simple avoidance of the delivery of redundant narratives to updating user interests.

User Behavior

The user behavior in the museum context is well studied in curatorial science [6] and used in several systems personalizing the user experience [17, 18]. Several categorizations were used, for example one user may go through almost every artifact that is one his/her way, and another user may be more selective and chooses artifacts that have certain concepts. Our categorization of user types is based on Sparacino's work [18] and it classifies users to three main categories. These categories were validated by our own research of site studies and interviews with staff at our partner museums:

- The *greedy* type wants to know and see as much as possible. He is almost sequential, and does not rush.
- The *selective* type explores artifacts which represent certain concepts, and wants to dig into those concepts only.
- The *busy* type does not want to spend much time on a single artifact and wants to stroll through the museum to get a general idea.

The user type is inferred from interaction history every 1 minute from the location data accumulated within 3 minute interval and topics of previously selected sound objects.

User Interest

The User Interest element of the user model captures concepts that the system realizes the user may want to further explore. The capturing of the user interests is in the center of the research of several disciplines such as information retrieval, information filtering and user modeling [19]. Most of the systems were developed for retrieval of documents where document content is analyzed and explicit user feedback is solicited to learn or infer the user interests. In the context of ec(h)o there is no direct feedback from the user. Ec(h)o can be categorized as the personalized system as it must be able to watch the user's behavior and make generalizations and predictions about the user based on their observations [9,16]. Our approach uses an unobtrusive observation of user behavior and it is similar to the approaches monitoring user browsing patterns [10,12] or user mouse movement and scrolling behavior [8].
The interests for the user are represented as a set of facts where each fact represents a single interest and its relative level

```
(user-interest (user-id john) (concept evolution) (level strong))
```

As described in the previous sections, each artifact/exhibition is associated with a set of concepts. The sound objects address a set of particular concepts as well. The interaction of the user and artifacts and sound objects is stored in the Interaction History that together with the user behavior type are used to infer the user's interests. The following principles for the user interest inference are implemented using the reinforcement learning approach [11]:

- If a greedy type user slows/stops in front of an artifact, we can infer that the user is interested in any of general concepts represented by the artifact. If the user continues with his *greedy* behavior in front of that artifact, his interests is updated with related concepts from sound objects selected (not necessarily closely related).

- Interests of a *selective* user do not get easily overwritten. If a *selective* user is moving slowly in front of an artifact he is not interested in[3], one of his previous interests is overwritten by a concept that is 'close' to his previous interests. If a *selective* user stops in front of an artifact he is not interested in, one of his previous interests is overwritten by a concept that is represented strongest by the artifact.
- If a *busy* user slows/stops in front of an artifact, several of his interests are overwritten by general concepts that are also represented strongly by the artifact.
- If a user's behavior is not categorized yet, *User Interest*s can be any general concepts that are strongly represented by the artifact the user slows/stops in front of.

We limit the number of concepts represented in the user model as user interests to 6 to reduce the error in retrieved objects [13].

This is an example of a rule that computes interests of a *greedy* user:

```
(defrule get-user-interest1
       (time ?current-time)
       (is-slow ?user ?current-time)
       (in-front ?user ?object ?current-time)
       (user-behavior ?user greedy)
       (object-model (object-id ?object) (concept ?i) (level ?l) (x-position ?x)
(y-position ?y) (radius ?r))
=>
(assert (user-interest (user-id ?user) (concept ?i))))
```

Sound Object Retrieval Mechanism

The retrieval process in ec(h)o can be broken into four steps as illustrated in Figure 3.

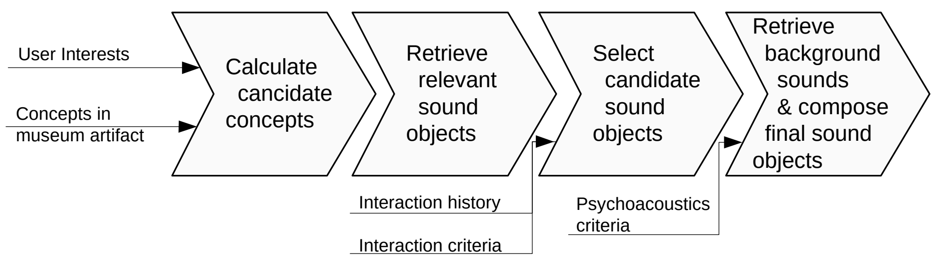

Figure 3 Retrieval process

First, the system determines the candidate concepts as an overlap between user interests and concepts represented by the museum artifacts. The candidate concepts are ranked by a combination of the level of the interest of the user and how strongly they are represented by the artifact. In the second step the candidate concepts are used to retrieve semantic descriptions of the information sound objects. The temporal and location properties of the artifact are used to narrow the search to sound objects that are closely related to the presented artifacts. Next, ec(h)o uses interaction history as another input and applies interaction criteria to select candidate information sound objects. Finally, the background sound objects

[3] There is no overlap between artifact concepts and user interests.

for each information objects are retrieved using psychoacoustics criteria and the psychoacoustics rules are applied to finalize the choice of the sound objects.

Network of museums

One of the main features of the ec(h)o system is that it enables the user to experience the richness of the museum collections located not only in the visited museums but also from the other linked museums. For example, a visitor standing in front of a bear specimen in Nature Museum in Ottawa can listen to the sound object about the role of the bear in the mythology of aboriginal tribes on the West Coast[4] retrieved from the Museum of Anthropology in Vancouver.

Two aspects are critical for fluid retrieval and access of sound objects from other museums: protocol compatibility and semantic mapping between conceptual structures. We addressed the protocol compatibility issue by reusing the infrastructure and protocols our group developed for connecting learning object repositories [21, 22]. The only difference is that instead of learning object metadata we share the sound object semantic descriptions.

As different museums can have different conceptualization of the topics covered by their stories the problem of mapping between these conceptualizations need to be addressed. First, we use the standard Conceptual Reference Model for describing temporal and spatial entities which allows us to relate sound objects to time and space. Second, we use Dewey Decimal Classification as an intermediary for mapping between museum specific topic maps. Although this does not provide for an exact mapping our solution is acceptable in the museum setting where the exploration aspect[5] of the user experience dominates the in-depth learning aspect.

Implementation

[To reviewer: This is a status as of Sept 6, 2003. The ec(h)o has a scheduled in-situ demonstration in Nature Museum in Ottawa in December 2003 in "Finders and Keepers" exhibit. By the time of a deadline for camera-ready version we expect the system to be fully implemented and going through evaluations before the final demonstration.]

Currently we are exploring several possible solutions to the global tracking of users throughout the space. There are three basic types of tracking systems available: vision based, audio based, and RF based. Currently, RF based tracking hasn't yielded high enough resolutions to be useful (resolution is at the 6 foot range). However, we are hopeful with the new developments with blue tooth tracking systems that these problems will be addressed. Visions systems are limited by occlusion and camera geometry, but are sufficient for certain kinds of exhibits. Currently, the audio based solution provides the most flexibility and accuracy of the currently available systems. This system is accurate within centimeter accuracy at a 30 frames per second update rate.

The system uses the "eyes" vision system within the "Max/MSP" environment to allow for quick experimentation and adaptation to many circumstances. The vision module uses two color cameras connected to the Apple PowerBook laptops to cover two exhibits. The cameras are positioned on the

[4] Assuming the user model indicates the user is also interested in history.

[5] Providing that information is relevant to the temporal and spatial aspects of a museum artifact.

ceiling above the artifacts and they are able to recognize several people gesturing with an ec(h)o object within one camera zone[6] in combination with the global positioning system.

The sound module consists of a soundfile playback and mixing system driven by the position tracking module. User position information is provided by the position tracking system and used to dynamically mix the soundscapes the user is immersed in. The sound module uses a custom-designed software mixing system implemented on an Apple computer. We have developed an authoring environment for mapping soundmark design to the physical topology of an exhibition. The delivery of the sound object is through the stereo audio interface using FM radio transmission to wireless headphones. In our testing environment the system is serving four users and scales simply by adding more hardware audio channel pairs.

As a part of the project we are producing over 600 reusable sound objects at the low level of granularity and annotating them with the ontological information. A majority of informational and narrative sound objects were produced from the interviews with curators and experts from the nature Museum in Ottawa. We have also digitized some of the recordings from the museum archives, mainly animal sounds.

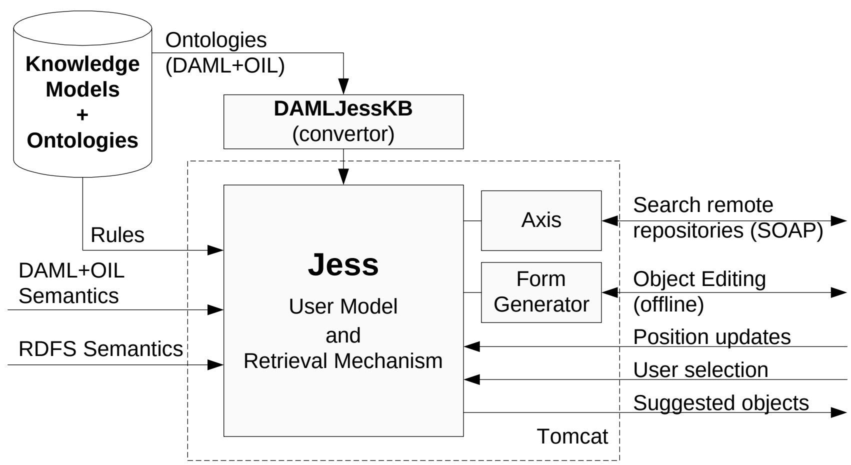

Figure 4 Implementation schema of the reasoning module

The reasoning module is fully implemented with all features described in the previous section and embedded into the Tomcat environment. The Figure 4 shows the implementation schema with Jess Inference engine in the center of the reasoning module. DAMLJessKB[7] converts DAML+OIL ontologies to Jess facts. Reasoning module is connected with other modules through the socket connections and communicates with other 'museums'[8] via SOAP based protocol [22]. We have developed a web-based tool for editing of ontological descriptions of sound objects that generates forms based on the ontology definition by direct querying of the ontologies loaded into the Jess inference engine.

Due to the unresolved tracking module we have not been able to test a fully integrated system up to this point. We did testing on partially integrated modules with emulated input from the tracking module.

---

[6] The zone for the camera depends on the height of the mount and height of the hand handling an object. For example, the zone diameter for the camera mounted at 4m can be as wide as 15m with a wide angle lens.

[7] http://plan.mcs.drexel.edu/projects/legorobots/design/software/DAMLJessKB/

[8] Currently we emulate a museum network by seeding independently operating repositories on separate computers. As the project is funded by the Canarie who operates the broadband internet in Canada our assumption is that the connectivity between museums will be at the superior level.

Conclusions

In this paper we presented retrieval mechanism used in an Augmented Audio Reality System for Museum Visitors named ec(h)o. Each visitors experience is tailored to the user interests. The user interests are inferred from the user movement through the exhibition as well as from the visitor's interaction with the sound objects. The sound objects are retrieved based on their relevance to the user interests, narrative criteria and psychoacoustic criteria. Ec(h)o uses ontologies to describe concepts, temporal and spatial characteristics, psychoacoustic and sound characteristics of sound objects. The retrieval mechanism is represented in form of the rules that capture contextual, sound, psychoacoustic and composition criteria for plausible user experience.

The system is a result of convergent research streams from research in object repositories, interaction design, auditory display, knowledge representation, and information retrieval. The project builds on current research in each of these areas, and focuses on an interdisciplinary and integrated approach to design, prototyping and evaluation. An ecological model allows for a strategy to work across the problem domains, and helps manage the design and use of information and structures through all layers of the project. We believe this has enabled us to extend works cited through the paper in several directions. First, it extends the work of the Alfaro et al. work [1] by building rich model of the concepts represented by the sound objects. In ec(h)o, the content presented to the user is not pre-processed for possible linkages as in the systems using Rhetorical Structure Theory [20]. Our approach replaces pre-processed linkages with a retrieval mechanism based on composition and interaction criteria formulated in the form of the rules and applied to semantically described independent objects. This allows ec(h)o to create a network of museums sharing objects and providing richer user experience.

Acknowledgements

Work presented in this paper is supported by Canarie Inc. grant under E-Content program. Authors would especially thank to the Mark Graham and his colleagues in the Nature Museum in Ottawa for the enthusiastic support to this project. We would also like to thank our colleagues and participants in several workshops that contributed to the development of the project, namely Dale Evernden, Doreen Leo, Gilly Mah, Robb Lowell, Mark Brady and Phil Thomson. We are grateful to Jim Bizzochi for his exciting introduction to the film theory.

References

1. Alfaro, I., Zancanaro, M., Cappelletti, A., Nardon, M., Guerzoni, A. Navigating by Knowledge. In Proceedings of International Conference on Intelligent User Interfaces. Miami, Florida. January 12-15, 2003.

2. Andolesk, Diane, Michael Freedman. "Artifact As Inspiration: Using Existing Collections And Management Systems To Inform And Create New Narrative Structures." *Archives and Museum Informatics Museums and the Web 2001*, 2001.

3. Bederson, B. "Audio Augmented Reality: a prototype automated tour guide." *Conference companion on Human factors in computing systems*, 1995.

4. Bordwell, D.: Film Art: An Introduction (5th edition). McGraw Hill, 1997.

5. Crofts, N., Dionissiadou, I., Doerr, M., Stiff, M. (eds.) Definition of the CIDOC object-oriented Conceptual Reference Model , July 2001 (version 3.2.1)

6. Dean, D. (1994). Museum Exibition: Theory and Practic. London, Routledge.

7. Dewey Decimal Classification, http://www.oclc.org/dewey/

8. Goecks J and Shavlik J (2000) Learning users' interests by unobtrusively observing their normal behavior. In: Proceedings of the 5th International Conference on Intelligent User Interfaces, New Orleans, Louisiana, January 2000, pp. 129–132.

9. Kobsa, A., Fink, J. User Modeling for Personalized City Tours. Artificial Intelligence Review 18, 2002, pp. 33-74, Kluwer Academic Publishers.

10. Liberman, H. Letizia: An Agent that Assists Web Browsing. IJCAI-95, pp. 924-929.

11. Mitchell, T. (1997). Machine Learning. New York, NY, McGraw-Hill.

12. Mladenic, D. Personal WebWatcher: Implementation and Design, Technical Report IJS-DP-7472, Department for Intelligent Systems, J.Stefan Institute, October, 1996.

13. Mostafa,J., Mukhopadhyay, S,.and Palakal, M. Simulation Studies of Different Dimensions of Users' Interests and their Impact on User Modeling and Information Filtering, Information Retrieval 6 (2): 199-223, April 2003

14. Sarini, M., Strapparava, C. Building a User Model for a Museum Exploration and Information-Providing Adaptive System. *Proceedings of the 2nd Workshop on Adaptive Hypertext and Hypermedia, HYPERTEXT'98*, 1998.

15. Schlieder, Ch., Vogele, t., Werner, A. Location Modeling for Intentional Behavior in Spatial Partonomies Ubicomp 2001, Workshop "Location Modeling for Ubiquitous Computing", Atlanta, Georgia, USA, pp.63-70

16. Seo Y-W and Zhang B-T (2000) A reinforcement learning agent for personalized information filtering. In: Proceedings of the 5th International Conference on Intelligent User Interfaces, New Orleans, LA, January 2000, pp. 348–351.

17. Serrell, B. (1996). The question of visitor styles. In S. Bitgood (Ed.). Visitor Studies: Theory, Research, and Practice, Vol. 7.1. Jacksonville AL: Visitor Studies Association, pp. 48-53.

18. Sparacino, F. (2002). The Museum Wearable: real-time sensor-driven understanding of visitors' interests for personalized visually-augmented museum experiences. In: Proceedings of Museums and the Web (MW2002), April 17-20, Boston.

19. Wahlster, W. and A. Kobsa (1989). User Models in Dialog Systems. In: A. Kobsa and W. Wahlster, eds.: User Models in Dialog Systems. Heidelberg Œ Berlin, Springer Verlag.

20. Zancanaro, M., Stock, O., Alfaro, I. Using Cinematic Techniques in a Multimedia Museum Guide. Using Cinematic Techniques in a Multimedia Museum Guide. In Proceedings of Museums and the Web 2003. Charlotte, North Carolina. March 19-22, 2003.

21. Hatala, M., Richards G. Global vs. Community Metadata Standards: Empowering Users for Knowledge Exchange, in: I.Horrocks and J.Hendler (Eds.): International Semantic Web Conference 2002, Springer, LNCS 2342, pp. 292-306, 2002.

22. Hatala, M., Richards G. Semantic Cobblestones: An Interoperability Mechanism for Learning Object Repositories. In McGreal, R.(ed): Online Education Using learning Objects, Kogan Page, London, 2003 (in press)