MAGA: A Mobile Archaeological Guide at Agrigento

Agnese Augello[1], Antonella Santangelo[1], Salvatore Sorce[1], Giovanni Pilato[2], Antonio Gentile[1], Alessandro Genco[1], and Salvatore Gaglio[1,2]

[1] DINFO - Dipartimento di ingegneria INFOrmatica, University of Palermo
Viale delle Scienze - 90128 Palermo – Italy
{augello,santangelo,sorce}@csai.unipa.it
{gentile,genco,gaglio}@unipa.it

[2] ICAR - Istituto di CAlcolo e Reti ad alte prestazioni, Italian National Research Council
Viale delle Scienze - 90128 Palermo - Italy
g.pilato@icar.cnr.it

Abstract. The use of a PDA with ad-hoc built-in information retrieval and auto-localization functionalities can help people in visiting a museum in a natural manner instead of traditional audio/visual pre-recorded guides. The goal of this paper is to build an user-friendly virtual-guide system adaptable to the user needs of mobility and therefore usable on different mobile devices (e.g. PDAs, Smartphones). An information retrieval service is included which is easily accessible through a spoken language interaction or an auto-localization service. The system takes the advantages of Chatbot, speech recognition technologies and RFID detection, allowing a natural interaction with the user.

Keywords: Multimodal Interfaces, Chat-bots, PDAs, Information Retrieval, Cultural Heritage, Radio Frequency Identification, Ubiquitous Computing, Pervasive Systems.

1 Introduction

In the ubiquitous computing vision, technology is seamlessly integrated into the environment and provides useful services to humans in their everyday lives [15]. The advent of small, mobile devices, the new possibilities offered by third-generation wireless and Bluetooth communication, localization technologies such as Radio Frequency Identification (RFID) and Global Positioning System (GPS) can significantly contribute to the realization of such a vision, and they can be exploited to reach the user needs of mobility and focused information access.

In particular RFID [10] technology allows to sense changes in the environment in which the user is located and to automatically adapt and act accordingly to these changes based on user needs. The use of pervasive technology in the cultural heritage area has led to the creation of cultural mobile guides that put the user and his needs of mobility at the focus of attention, and that can be easily consulted through a more natural interaction [5], [13].

This technology can be exploited to improve the visitor experience in a cultural heritage environment, allowing the access to relevant cultural heritage resources in a context-dependent fashion. However, pervasive devices suffer from the limited dimensions of their screens. Users are compelled to use special styli/touch pens, with consequential uncomfortable interaction experiences[3]. Vocal interaction is a more direct approach [12] but it comes with a different set of constraints such us needed processing power, speakers adaptation, narrow dictionaries and bounding grammars.

In this paper, we propose a versatile, user-friendly, virtual guide for cultural heritage museums, called "MAGA". To achieve a more natural interaction and enable site fruition also by inexpert and/or disabled users, a conversational agent is designed based on a conversational agent (chatbot) [1] integrated with reasoning capabilities, speech recognition/synthesis module, and RFID-based auto-localization module.

Smart interaction is therefore enabled by making deductions about the visitor information requests. An OpenCyc [11] ontology has therefore been created for the specific domain, and it has been combined with the agent dialogue capabilities. The system is equipped with a multimodal interface using the HTML+Voice language. The vocal technologies embedded in a multimodal browser allow the information access through spoken commands in addition to traditional visual and keyboard (or stylus) based commands.

The application scenario is assisting visitors at the "*Museo Archeologico Regionale di Agrigento*", located in the Valle dei Templi area. The proposed system can replace the traditional audio/visual pre-recorded guides implying a more natural and enjoyable fruition of the wealth of information about the site. A system prototype has been implemented on a Asus A730W PDAs equipped with CF-RFID readers.

2 The MAGA System

The proposed system is an intelligent museum guide. It is accessible using a PDA equipped with an auto-localization module and provides users a natural language speech interface for chatbot interaction. An overview of the overall architecture, based on a client-server paradigm, is shown in Fig.1.

The key feature of MAGA guide is that the user can interact with the system by means of natural language sentences through a chatbot which has been implemented using the ALICE technology [1]. The knowledge base of the chatbot is composed of question-answer modules, called categories and structured with AIML, an XML-like language designed for creating stimulus-response chat robots. The question is called the "pattern" while the answer is called the "template". The template can be composed by other AIML tags, that can transform the response in a software that can save data, can start other programs, can give conditional answers and can recursively call other categories. Besides an ad hoc micro-theory (i.e. a collection of concepts and facts usually regarding one specific knowledge domain) has been created and hooked up to some of the existing microtheories present in the OpenCyc [11] commonsense Knowledge Base.

The created microtheory contains cyc-collections defining specific concepts for the Museum, new predicates, assertions and inference rules. The knowledge base of the

chatbot has been therefore enriched with new AIML tags that enable the ontology querying, executing and asserting Cyc-oriented statements. The Cyc responses are therefore automatically composed in a natural language sentence according to the rules inside the AIML categories.

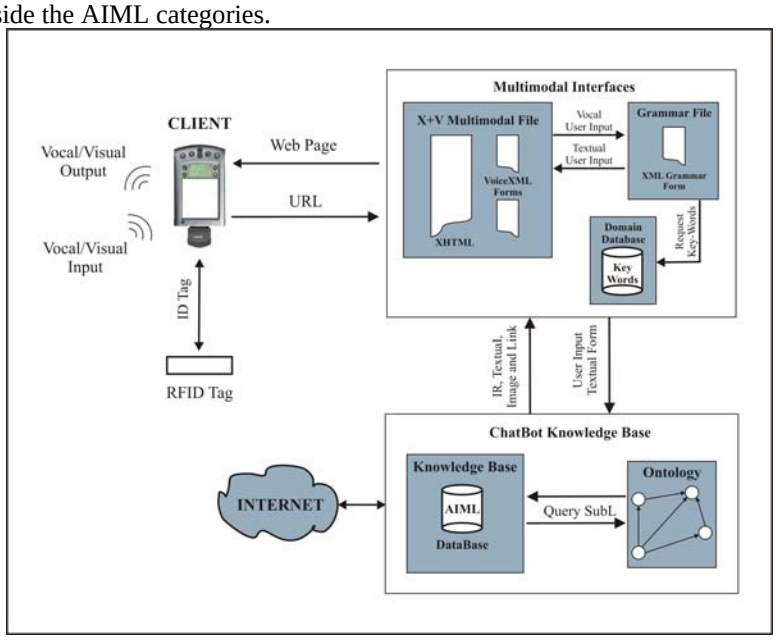

Fig. 1. The proposed system architecture

The Cyc tags enable the composition of answers that are not present in the traditional AIML knowledge base. This feature allows to enhance the chatbot knowledge in two ways: a) extending the number of AIML categories to obtain a more fluent dialogue, and b) extending the ontology with new concepts, facts and relations of the specific domain to provide the chatbot a smarter behavior.

The dialogue between chatbot and user is carried on verbally, using spoken natural language. The multimodal interface is an X+V [4], [8] component loaded by the multimodal browser. The interface is developed with the IBM Multimodal Tools 4.1.2.2 [7] for WebSphere Studio V5.1.2, that includes the developed environment and the technology for the application execution. The control of interaction between vocal and visual parts is made through XML Event, so that voice handlers can be invoked through a standard EventListener interface.

The speech recognition process is carried out through an ad-hoc built-in speech grammar. An ad-hoc grammar for the specific domain has been realized, which is a set of rules that specifies utterances that a user may pronounce. This allows the improvement of the automatic speech recognition process: in fact a small domain context allows to build a richer and thorough set of pre-defined sentences. Besides, a particular domain-restricted cleaver grammar increases the user freedom of expression, enhancing the naturalness of the dialogue.

The grammar has been built using XML Form [2] from a set of key-words about the domain. These key-words are classified in categories and stored in a database. The

set of key-words are all names related to the application domain. An ad-hoc algorithm gets the key-words from the database and puts them into the grammar to build spoken user utterance. The grammar is easy-fitting and can be improved with low effort because it suffices changing the key-words database to induce a grammar for the new context domain. When the speech matching occurs, the grammar returns in textual form what has been recognized and the system submits the textual query to the chatbot.

The interaction between the application running on the PDA and the system is also triggered by the detection of a RFID tag, which is used to estimate the PDA position within the environment [6], [10]. RFID tags are categorized as either active or passive. Active RFID tags are powered by an internal battery, passive RFID tags operate without a separate external power source and obtain operating power generated from the reader.

We have passive tags each holding an unique ID, and each attached to an item of interest. Once the RFID reader on the PDA detects a tag, its unique ID is passed to the server application via the Wi-Fi connection. This allows the chatbot to start the interaction with the PDA by providing some basic information related to its current position. The link between the information sent to the PDA and the user's position is given by the detected ID.

According to this feature, people can go on asking questions about the current object to the chatbot with vocal queries, or they can discard the information and continue their tour.

3 Application Description

The system is accessible through a web page in a multimodal browser from the handheld device. The interaction occurs through the loading of X+V pages, that can be triggered by user vocal and visual command or RFID detection; in every page the user can have a vocal dialogue with the chatbot. Whenever the user pronounces his request, the multimodal browser looks for a match in the grammar file.

If a match is found the application submits the recognized query to the chatbot waiting for the answer. In the same manner the application submits the query associated to the ID tag detected from the RFID reader. The chatbot searches for the best match rule into its knowledge base. The AIML rule can directly produce an answer, or it could be necessary to query Cyc for building an answer that fits better with the user request. The chatbot answer could also be the result of a query to standard search engines that will search for local or remote documents that are related to the user query.

The analyzed domain concerns the "*Museo Archeologico Regionale di Agrigento*". The museum was built in 1960, by the architect Franco Minissi and is located in the Valle dei Templi area. The building contains 18 rooms that document the history of Agrigento. The rooms are divided in two main sections. The former is dedicated to the objects discovered in the ancient city. The latter is dedicated to the archaeology of the area.

A system screen shot is shown in Fig. 2.a, while a sample of AIML category is illustrated in Fig. 2.b. In what follows a few examples of interaction are reported.

(a)

```
<category>
 <pattern>I need information about
          the showcase * </pattern>
 <template> In this showcase you can see
   <cycsystem>(cyc-query'
  (#$Contains <star/> ?X))
   </cycsystem>
   belonging to<cycsystem>(cyc-query
   '(#$CreatedDuring <star/> ?X))
   </cycsystem>
   attributed to <cycsystem>(cyc-query
  '(#$CreatedBy <star/>?X))
  </cycsystem>
    representing <cycsystem>(cyc-query
   '(#$Representing <star/>?X))
  </cycsystem>
 </template>
</category>
```

(b)

Fig. 2 ***a)*** System screen shot ***b)*** An example of AIML category

The user can query the chatbot for specific information:

User: I need information about the Archeological Museum in Agrigento.
Chatbot: It is one of the most modern museums in Sicily and contains archeological finds from the provinces of Agrigento and Caltanisetta, extending from prehistoric to Roman times.

The chatbot can evaluate the information obtained from the RFID tag about the art object nearest to the user:

Chatbot: You are in the room 3, in front of the showcase 17. Do you want more information?

The chatbot explores the ontology with the aim of extracting information about the objects inside the showcase, such as the historical period, the subject and the author of each work:

Chatbot: In this showcase you can see an attic amphora belonging to 520-500 a.C. attributed to Dikaios painter and representing Athena on a horses quadriga.

The chatbot can store the number of the current room allowing the user to ask information about the next room:

User: What is in the next room?
Chatbot: In the room 4 there are Capitals and gargoyles with lion's heads from various temples.

4 Conclusions

A multimodal pervasive user-friendly system has been presented to assist visitor aimed to enable a more natural and enjoyable fruition of the wealth of aggregated information about the “Museo Archeologico Regionale di Agrigento”

The system takes the advantages of chatbot, speech recognition and RFID technologies embedded in a Personal Digital Assistant, which can be easily carried during the site tour. The proposed system is a first prototype intended as a smart replacement of the outdated traditional audio/visual pre-recorded guides, future work will regard the enhancement of the system improving the interaction naturalness.

References

1. Alice: http://www.alicebot.org
2. Andrew, H. and McGlashan, S.: Speech Recognition Grammar Specification Version 1.0, W3C Raccomandation, (2004)
3. Aridor, Y., Carmel, D., Maarek, Y. S., Soffer, A., and Lempel, R. 2002. “Knowledge encapsulation for focused search from pervasive devices.” ACM Trans. Inf. Syst. 20, 1 (Jan. 2002), 25-46.
4. Axelsson, J.,Cross, C.,Ferrans, J.,McCobb, G.,Raman, T. and Wilson, L.: XHTML+Voice profile 1.2, W3C Note. (2004)
5. Figa, E. and Tarau, P. 2002. The VISTA architecture: experiencing stories through virtual storytelling agents. SIGGROUP Bull. 23, 2 (Aug. 2002), 27-28. (2002)
6. Genco, A.,Sorce, S. and Varrica, E.:“Combined People Position and Compass Detection for Context-aware Service Provision”, proc. of IEEE 3rd International Workshop on Embedded Computing IWEC-06, Columbus, Ohio, USA, 14 Aug 2006, pp 368-375
7. IBM Multimodal Tools: http://www-306.ibm.com/software/pervasive/multimodal/
8. McGlashan, S., Burnett, D. C.,Carter, J., Danielsen,P., Ferrans, J.,Hunt, A.,Lucas, B.,Porter, B., Rehor, K. and Tryphonas, S.:Voice Extensible Markup Language (VoiceXML) Version 2.0, W3C Recommendation, (2004)
9. MobileHCI '05, vol. 111. ACM Press, New York, NY, 243-246. (2005)
10. Ni, L.M.; Yunhao Liu; Yiu Cho Lau; Patil, A.P.: LANDMARC: indoor location sensing using active RFID, Proceedings of the First IEEE International Conference on Pervasive Computing and Communications (PerCom 2003), 23-26 March 2003 Page(s): 407-415
11. OpenCyc: http://www.opencyc.org
12. Pilato, G.,Augello, A., Santangelo, A., Gentile, A. and Gaglio, S.: An Intelligent Multimodal Site-guide for the “Parco Archeologico della Valle dei Templi in Agrigento, Proc. of First European Workshop on Intelligent Technologies for Cultural Heritage Exploitation at The 17th European Conference on Artificial Intelligence, Riva del Garda, Italy, August, 45-49 (2006)
13. Stock, O.: ALFRESCO: Enjoying the combination of natural language processing and hypermedia for information exploration. In intelligent Multimedia interfaces, M. T. Maybury, Ed. American Association for Artificial Intelligence, Menlo Park, CA, 197-224, (1993).
14. Tarau, P. and Figa, E.: Knowledge-based conversational agents and virtual storytelling. In Proceedings of the 2004 ACM Symposium on Applied Computing (Nicosia, Cyprus, March 14 - 17, 2004). SAC '04. ACM Press, New York, NY, 39-44. (2004)
15. Weiser M.: The Computer for the Twenty-First Century, Scientific American, pp. 94-100, (1991).