Multimodal Mood-based Annotation

BARBARA RITA BARRICELLI, MARIO MALCANGI, PIERO MUSSIO
Department of Computer Science and Communication
Università degli Studi di Milano
Via Comelico 39/41, 20135 Milano
ITALY
{barricelli, malcangi, mussio}@dico.unimi.it

Abstract: - The paper presents an architecture for multimodal mood-based annotation systems. The architecture aims at the implementation of interactive multimodal systems to support communities of users in the creation and management of annotations in locative media projects. The annotations are multimodal in that they can be created and accessed through visual and audio interaction. The annotations are mood-based in that they reflect the mood of the user respect to the point of interest s/he is commenting. The paper presents a definition of multimodal mood-based annotation and a description of the architecture, illustrating in particular the interaction process between users and systems through the audio interface. A concrete application of the architecture is presented: an annotative locative media project aimed at supporting tourists in creating annotations related to the Valchiavenna valley in Italy.

Key-Words: - Communication, Multimedia, Multimodal, Audio, Annotation, Mood, Speech recognition

1 Introduction

The increasing popularity of mobile devices and the growing availability of Wi-Fi hotspots on territory, lead to the continuous development of new technologies called locative media. Locative media, together with pervasive computing, are technologies and processes that promise to reconfigure our understandings and experiences of space and culture [1]. They support people to collaborate in common activities related to real or virtual spaces and enable the creation and grow of social communities.

To this aim, digital maps are used to support communication and sharing of knowledge and are used as boundary objects for cooperative activities.

The use of digital maps capitalizes on the long tradition of using paper maps as a mean of thought and communication. However, to be used digital maps requires users to develop new skills and determines new relations between map authors and users [2].

Two main categories of locative media projects have been identified in [3]: phenomenological and annotative. In phenomenological locative media projects the actions of people acting in the world are traced and from their behavior information about the environment are derived. On the other hand, annotative locative media projects enables the users to change the world in which they live by adding data to it and are made possible by the use of Web 2.0 technologies that support cultures of participation [4]. In this way, the evolution of the maps from static objects to new media of knowledge creation, accumulation, distribution and sharing is fostered. The maps evolve becoming digital interactive, pro-active tools, which content changes in time through the use of annotation tools. In fact, users can annotate the maps by enriching the information embedded in them and can explicit their tacit knowledge by annotating them. Annotation tools allow users in a) creating and managing a shared knowledge base, and b) using digital maps as locative media to support knowledge sharing and access [5][6].

In human-to-human communication multiple signal are active at the same time. Written words are just a little part of communication. Utterances are richer in communication content as it integrates intonation, emotion, gender, and so on. Several researches indicate that voice annotation helps to express the more complex and social aspects of a collaborative writing task. In their study comparing text and voice annotation in papers reviewing, Neuwirth et al. [7] demonstrated that there is greater expressivity of the voice modality and that it produces annotations that writers also find usable.

Audio offers also additional advantages such as the hand-free interaction (speech recognition, speech synthesis), so a higher degree of automation can be achieved in the related application. The growing diffusion of multimedia and multimodal systems opens to new possibilities of interaction with software systems and offers new methods of documentation in different application domains, both technical and scientific.

In this paper, we describe an architecture that allows to implement interactive multimodal mood-based annotation systems for locative media projects. The architecture enables the collaborative creation and management of a shared knowledge base through the use of multimodal annotation tools. Particularly, the annotations created by the users reflect the mood of the user her/himself in browsing the digital map and its points of interest. This specific type of annotation is defined in the paper as mood-based annotation and its use is illustrated through the presentation of a case study.

The paper is organized as follows. Section 2 presents some annotative locative media projects. In Section 3 multimodal annotation in locative media projects is presented and the evolution of digital documents is discussed. The formal definition of multimodal mood-based annotation is given in Section 4, while in Section 5 the architecture is presented. Finally, the case study in Section 6 and Conclusions close the paper.

2 Annotative Locative Media Projects

Several experiences of locative media projects give indications on the use of multimedia annotation tools to support knowledge gathering and management.

WikiMapia [8] combines Google Maps with a wiki system and permits to create and maintain a free, multilingual, editable interactive map, which allows actors to add information, notes and photos related to points of interest. WikiMapia could also be accessed through the use of mobile devices via another system called WikiPlaces [9].

Hide&SEEK project [10] faces the challenge of exchanging cultural knowledge through the creation of a collaborative street game. Users can publish annotations on a map through the use of any device with wireless networking capabilities.

Analogously, Radio aporee [11] is a Web application that supports the collaborative creation and exploration of public soundscapes. Actors can upload and manage audio clips associated with locations around the world.

The project Silence of the Lands [12] allows users to use ambient sounds as conversation pieces of a social narrative. Using mobile devices, users can geo-reference their walk on the territory and the sounds they record on the field. The recorded data are associated to GPS data and stored on the Web server as well as visualized on the Web site as individual soundscapes.

3 Interactive Systems, Digital Maps and Multimodal Annotation in Locative Media

The increasingly improvement of interaction technologies has fostered the use of the Web as a remote software interface [13]: it must be designed consciously thinking and planning the experiences that the application will create for its users. To support such an approach in interactive systems design, Borchers [14] evolves the Alexandrian concept of “quality without a name” which characterize environments that people enjoy living in. This quality cannot be reduced to one dimension, but it can be only described by the patterns of events that frequently happen there or in our language by patterns describing user experiences. In the case of interactive software systems, these experiences are determined by the user interpretations of the messages they exchange with the system and which create the environment as perceived by users [15].

We give a name to the “quality without a name”, calling it “habitability”, thus extending the common meaning associated to this word in Human-Computer Interaction [16]. In our view, habitability refers not only to the match between the language people employ in their activities and the language that the system can accept [17][18], but requires the match of the whole system of symbols used in the application domain (words, graphic symbols, images, sounds) to describe the activities to be performed with the system of symbols (e.g. words, icons, widgets) the application can accept [19].

According to Naeve, recording, transmission and annotation of information are activities at the base

of the Human Semantic Web – i.e. of the conceptual interface through which humans can access Web resources [20].

Digital maps are particular types of e-documents that act as medium to permanently record and communicate concepts, thoughts, and data related to territory. e-documents and e-maps management derives from the traditional documentation techniques: professional communities developed and shaped customized documentation styles, notations and annotation procedures to better express and record their specific knowledge on a permanent physical media. However, the evolution from traditional documents to e-documents passed through six main stages:

1. Traditional paper-based documents are written by one or more authors but once they are released their content is definitive and could not be changed. The distribution of these documents takes place in the real world with the physical release of the documents to the interested people. The readers of the documents may annotate them using for example pens, pencils or sticky notes but to share their annotations with other people, they have to exchange them physically since the annotations on paper-based documents affect only the copy of the document on which the annotations are added.
2. e-documents created by one or more authors, who are the only ones allowed to change their content. This type of document is exchangeable among interested people by e-mail or other digital communication channels.
3. Multimodal e-documents (textual, audio, and video) modifiable only by their authors but usable by other people. Also these documents are exchangeable among a community through e-mails and/or other digital communication channels.
4. Annotable e-documents created by one or more authors and annotable by other persons. The annotations are available to all the documents' readers that become therefore producers and not only consumers of the documents and of the annotations.
5. Multimodal annotable e-documents created by one or more authors, distributed and used through different digital communication channels and annotable by all their readers (who become producers and not only passive consumers).
6. Multimodal annotable e-documents created collaboratively in time by a group of authors, materialized on different channels and annotable by all the users (who become producers and not only passive consumers). This specific type of e-document can be defined as perpetual beta in that it does not reach a final and definitive version but continuously evolves in time.

As emerges clearly from the sixth stage, the evolution of Web 2.0 technologies leads to a fuzzy distinction between documents creators and users, and opens to new possibilities of collaboration and work organization. A further stage could be identified in which the e-document becomes a Web document, i.e. an element "consisting of dynamic, flexible, nonlinear content, represented as a set of linked information items, stored in one or more physical media or networked sites" [21]. This shift determines a loss of stability of the document in that its content is built "on demand" [22]. The Web document is the representation of data and structure that are stable and persistent in time; however, their presentation changes according to the chosen use modality (text, video, audio). Multimodality allows the access to the same knowledge (having the same meaning) in different ways without the loss of information. The process of creation of the content of a document is distinct from the process of its presentation: it can be adapted to the culture, skills and abilities of the current user, without altering document or annotation content but adapting them to the access modality in use [23].

4 Multimodal Mood-based Annotation

As highlighted by the previous section, annotation is a typical tool used to add information on documents: an annotation can be a note created by someone and kept private or shared with others and used to start discussions and/or to exchange and accumulate knowledge [2]. In this section, we first provide a formal definition of annotation that is independent by the modality by which it is accessed (text, video, audio). Then, the formal definition is adapted to the various access and use modalities and finally the multimodal mood-based annotation definition is given (see Figure 1).

4.1 Annotation Formal Definition

An annotation is referred to a document and in particular to a part of it (e.g. a word, a sentence, a

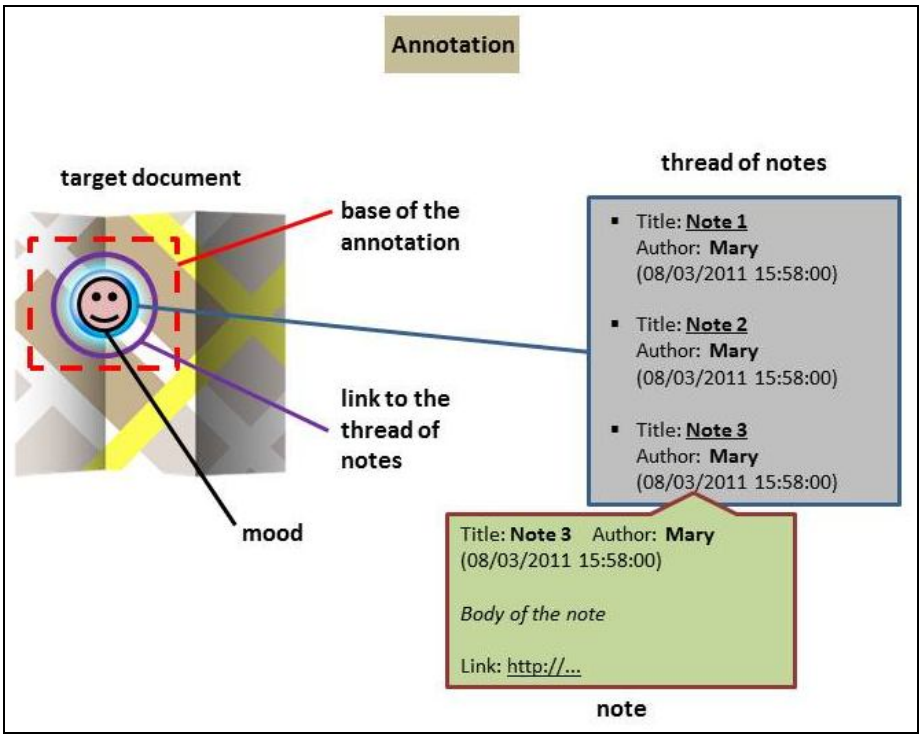

Figure 1. The definition of annotation given in Sections 4.1 and 4.2.

paragraph, an image). The document to which an annotation is added is called target document (td) and the specific fragment of the document to which the annotation is linked is called base of the annotation (base).

We do not consider an annotation as a single comment but as a thread of one or more comments that contribute to develop a discussion among document's readers. An annotation is therefore defined as:

$$annotation = <td, base, thread, link>$$

where thread is the set of comments and link is the element that is associated to the first comment in the thread and that allows to access to the thread itself.

A thread is defined as:

$$thread = \{note_1 , \dots , note_n\}$$

where $note_i$ is a comment created by a single author. To a first note further notes can be added, not necessarily created by the same author of the first one. The notes that belong to the same thread are ordered on the base of their creation date and time.

A note is defined as a quintuple:

$$note = <date\_time, author, title, body, reference>$$

Each note is composed by: date and time of creation (date_time), author's name (author), title chosen by the author (title), the text that was typed by the author of the note (body), a reference to further multimedia resources associated to the note (reference).

4.2 Multimodal Mood-based Annotation

Access and representation of annotations are strictly dependent on the device used for the system access.

Our proposal in this paper presents a solution to annotation management that aims at making annotations accessible through both visual and audio interactive system. In what follows, the definitions given in the previous section are adapted to the two interaction modalities.

4.2.1 Visual Interactive Systems

In case of visual interactive systems, the presence of an annotation on a map is made through the use of visual links that are graphical placeholders for the related thread of notes. A visual link could represent points of interests, areas or paths.

When a user selects a visual link, the related thread of notes is presented on the screen and the user is able to browse the notes and to add new ones to the thread.

4.2.2 Audio Interactive Systems

In case of audio interactive systems, the same threads of notes that are made available on a visual interactive system are accessible, but through a different interaction modality. Audio interactive system could be used in several cases, e.g. by cyclists, visually impaired persons, car drivers. This type of systems is thought for an on-site use: the presence of annotations in the area in which the user is, is signaled with a sound. This sound is a placeholder that, according to the definitions given in the previous section, is called audio link. The user can annotate the place where s/he is by recording a message through a voice-guided procedure. The annotation is made available both for the users of the audio interactive system and for the users of the visual interactive system. Using a visual system, the user will find on the map a visual link that represents the presence of an audio annotation and the content registered by the creator of the note could be heard or read through a speech-to-text feature execution.

4.2.3 Mood-based annotation

In locative media projects, the users play a fundamental role. Their contribution to the creation of a shared knowledge base about the places they visit influences the other people interested in visiting the same locations.

In order to enable the users to better explain their feelings in relation to a specific point of interest they visited, each annotation (i.e. the thread of notes and the link to it) should be connected to a *mood*.

Mood-based annotation extends the definition given in Section 4.1 to the following:

annotation = <td, base, thread, link, mood>

The user who creates the annotation is able to associate a mood to the annotation s/he is creating and it will be used to represent the link to the annotation. In visual interactive systems, the visual link should have a shape and a color that recall the associated mood, while in audio interactive systems, the synthetic voice that guides the interaction should alert the user of the mood related to the existing annotation.

In the following Section, an architecture for multimodal mood-based annotation systems that implements the definitions presented in this Section is proposed.

5 Architecture for Multimodal Mood-based Annotation Systems

In this Section the architecture we implemented on the basis of the formal annotation definition given above is described. The architecture (see Figure 2) supports the creation of multimodal mood-based annotation systems to be used in annotative locative media projects.

On client side, the users access the interactive system through visual or audio interface are able to browse the existing annotations or to create new ones.

On server side, the repository of annotations and maps is stored and shared among all the users of the interactive systems connected to the current locative media project. The annotation engine allows both the visual and audio interface to access the repositories.

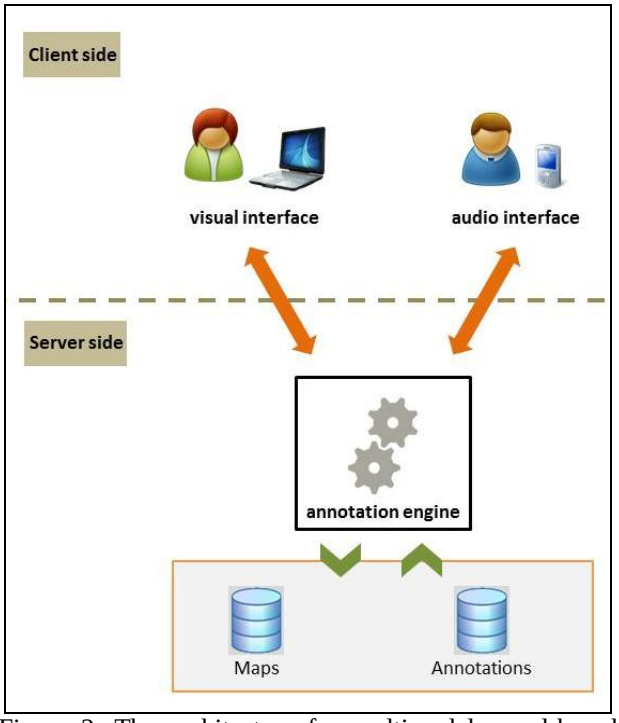

Figure 2. The architecture for multimodal mood-based annotation systems.

The repository of maps and allows the users to explore the location in real time with a mobile device or from home by using a desktop PC.

Through a visual interactive system, the user can create new annotations and to browse the existent annotations, both audio and visual ones, by using a

point-and-click interaction style. The annotations are stored in the repository using RDF language.

The main interesting aspect of the architecture is that the contents created using a visual interactive system can be accessed by an audio system and vice versa. To better describe how the audio interface works, the next subsections are presented.

5.2 Audio Interface

The system is based on an audio interface that enables the user to execute the annotation. The audio interface consists of voice command recognition, answering functions, and a set of embedded functions.

Voice command recognition has been implemented as speaker independent – vocabulary limited [24][25]. A set of command words has been recorded. The uttered command words are recognized by the speech recognition kernel and speech-to-text conversion is then executed.

Voice answering is used to drive the user during its annotation activity. It consists of a set of pre-recorded messages that are played according to the annotation activity flow.

The audio interface includes a set of embedded functions that implements some important tasks such as background noise measurement and voice-activity detection [26]. Background noise measurement is important for automatic equalization of the microphone chain to ensure high reliability of the speech recognition function. Voice activity detection is used to automatically identify an incoming voice command and to identify the starting and the ending of a voice annotation during the recording phase.

Text-to-speech (TTS) [27], speech emotion recognition [28], and environment audio recognition [29] will be implemented in the future developments to full automate the annotation process. TTS will be used to utter a text, speech emotion recognition to activate the related mood, and environment audio recognition to identify the surrounding context during the annotation.

5.3 Constraints and Silence Detection

Some constraints arise in audio-mode operation. Unlike text operation, audio operation requires to supervise the way data and commands are inputted.

The most important problem arises when an audio input is furnished because the communication interface needs to understand if it is an effective input or an artifact. Silence detection and end-point detection (EPD) is fundamental for efficient and effective audio interface functionality.

Silence detection is necessary for two main reasons:

- Hand-free and unattended operation
- Storage optimization

Silence detection requires a very smart signal processing and pattern matching algorithm because the high variability of environmental noise audio level and of microphone chain signal. A smart audio (speech) EPD has been developed to satisfy the above requirements based on fuzzy logic [26].

EPD is able to distinguish which part of the signal captured by the microphone is uttered speech mixed to background noise and which one is only background noise. So, effective utterance can be captured and speech recognition can be applied when voice command function is enabled. If voice annotation function is enabled, then it does not contain silence. Also empty recordings are then avoided.

A constraint is also that voice annotation can be no longer than 60 seconds. Such constraint is necessary to ensure EPD fault tolerance in detecting end-points of utterance during voice annotation.

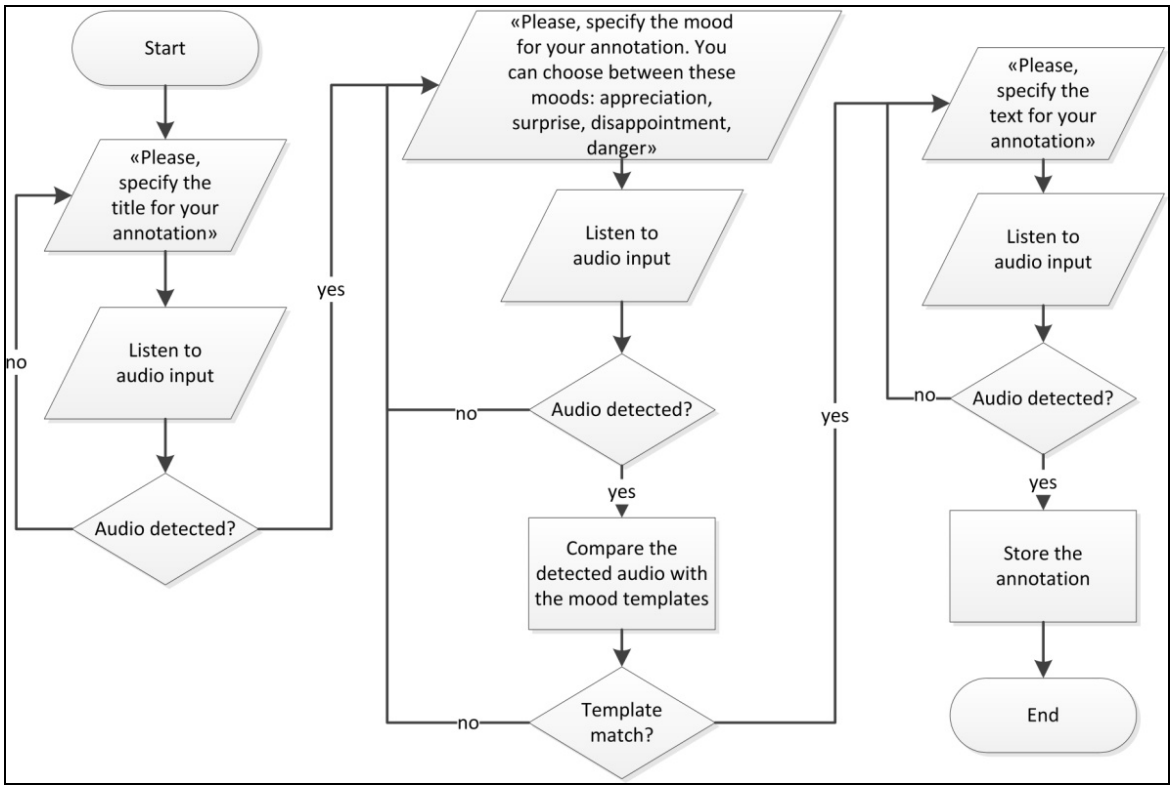

Figure 3. The three steps of interaction between the user and the system.

5.4 The Three Steps of Interaction

The audio interaction between the user and the system consists in a dialogue between the user and the system made of three sequential steps depicted in Figure 3: 1) specification of the title; 2) specification of the mood; 3) specification of the body of the note.

The first step consists of the specification of the title for the current annotation. The title can be chosen freely and the user has 8 seconds to express his choice. In the second step, the user can choose one mood to be associated to the current annotation. S/he has 6 seconds to specify the mood chosen among four available moods: appreciation, surprise, disappointment, and danger. For each of these moods, a template message is stored in the system and is compared with the detected audio, to check the validity of the user input. If the detected audio matches one of the mood templates, the user input is accepted as valid, otherwise the user is required to specify another choice.
The last step is the one in which the user can specify the body of the annotation. The time available for this operation is 60 seconds.

6 A Concrete Application

A concrete application of the architecture described in the previous section, is the one developed in the frame of the SCV (Sistema Culturale Valchiavenna) project [30][31]. The project is funded by Fondazione Cariplo and is carried out by the Territorial Community of Valchiavenna valley together with Università degli Studi di Milano and other partners.

The project aims to coordinate, plan and promote tangible and intangible assets related to artistic, historical and cultural heritage of the Valchiavenna valley in north of Italy. The goal of the development of a multimodal mood-based annotation system in this project is the exploitation of interdisciplinary capabilities and experiences developed by lecturers and researchers of several institutions to collaborative create a continuously updated knowledge base. The contents stored in the knowledge base are made available for dissemination and for guiding tourists that wish to visit the valley. The tourists may participate by accessing the domain expert knowledge base and

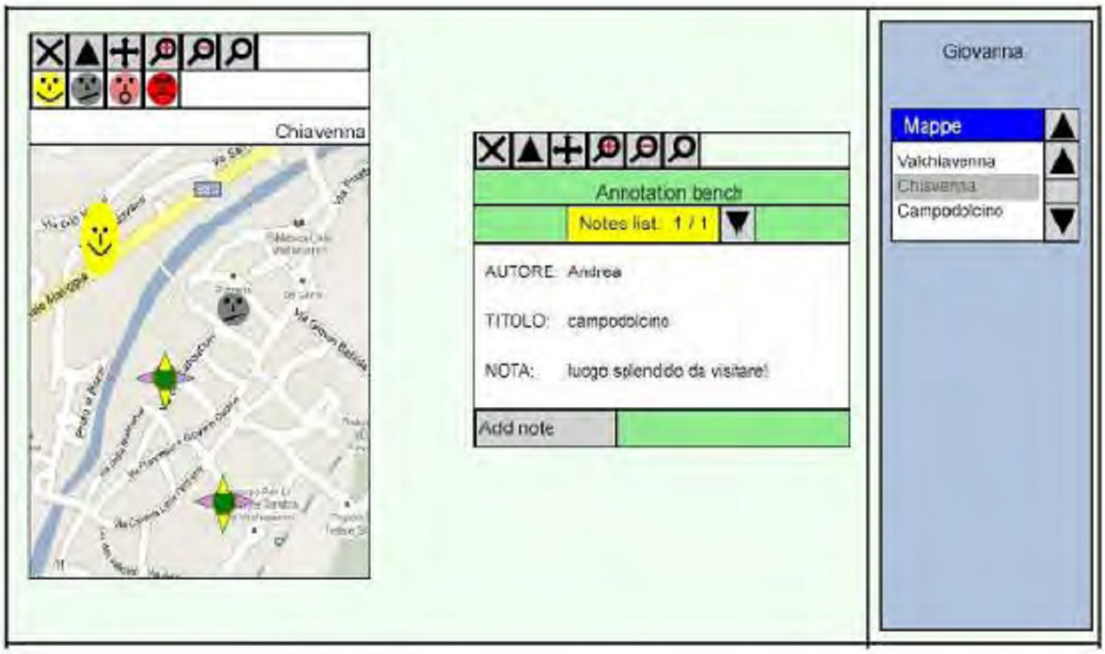

Figure 4. The visual interface of the SCV multimodal mood-based system.

collaboratively commenting the maps and the points of interest of the Valchiavenna valley. The visual interface of the system (Figure 4) allows the users to access the knowledge base (the maps and the annotations on them) through interaction with digital maps.

The tourists access the same shared knowledge base and exchange their impressions on the region by annotating the map using a set of moods (represented by emoticons). The moods represent the tourists' impressions about the places they visited and express appreciation, surprise, disappointment, and sense of danger.

These moods have three roles: i) they are tags, created by tourists to express their appreciation on the tagged element on the map, ii) they are links to the annotations left by the tourists, and iii) they are indexes because the information they convey can be used as criteria for future annotations retrieval.
VCP is accessible in two modalities, visual and audio. The moods associated by the tourists are visually represented by emoticons.

The audio interface of the system has been designed to meet the needs of tourists who want to annotate a map of a region they are visiting, by using vocal interaction instead of the visual one. This kind of interaction is useful in several cases in which the tourists are not able to interact with direct manipulation with the system, for example while they are driving a car, a motor bike or a bicycle.

The system is able to recognize the vocal commands that the tourist uses in order to express both the annotation text and the mood to be assigned to it. The audio annotation procedure is fired by using vocal commands and the base of the annotation on the e-document is detected by the use of GPS component available on the mobile device in use.

Once the user creates the audio annotation, it is stored in the knowledge base and through the visual interface is shown as an emoticon on the map that resumes the mood expressed by the user in the mood specification step.

Figure 5 illustrates the part of the visual interface in which the map is visualized.

On the map, together with an emoticon (red) that represents the link to a textual annotation and an icon (yellow) that is the link to an audio annotation. The color of the icon is representative of the mood the user specified in the creation of the audio annotation. The users can access both the textual annotation, by opening them like shown in Figure 4, and the audio annotation, by the opening of the audio file with the default media player installed on their devices.

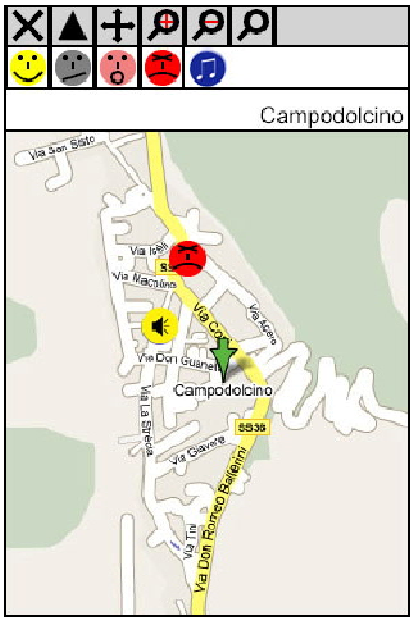

Figure 5. Part of VCP interface with two visual links: red emoticon to textual annotation, yellow to audio annotation.

7 Conclusions

This paper introduces an architecture for multimodal mood-based annotation systems for locative media projects. The architecture allows to implement interactive multimodal systems that support the creation and management of annotations made by communities of users. The annotations are created and accesses through visual and audio interaction and are mood-based in that they reflect the mood of the user respect to the point of interest they comment on.

A formal definition of multimodal mood-based annotation is given and the architecture is illustrated, considering both the interaction process between the user and the system and the audio interface of the architecture. A concrete application of the architecture in an annotative locative media project is also presented.

As future developments, the architecture will be extended with the implementation of a new feature: Text-to-speech (TTS) will be used to utter a text and to recognize the emotion expressed by the recorded speech. The detected emotion will be used to express the mood related to an audio annotation. Furthermore, the environmental sounds detected from the speech recorded are used to identify the surrounding context in which the user was during the annotation creation.

Autors' Note

Shortly after this paper was ready for submission, our department suffered the tragic loss of the colleague whose ideas had been the main inspiration behind this work. Piero Mussio was a tireless thinker whose intellectual generosity was known to students and colleagues alike. His creativity, drive, and demanding leadership will be sorely missed.

References:

[1] Gallow, A., Ward, M., Locative Media as Socialising and Spatializing Practice: Learning from Archaeology, *Leonardo*, Vol. 14, No. 3, 2006.
[2] Marcante, A., Mussio, P., Electronic Interactive Documents and Knowledge Enhancing: a Semiotic Approach, *Document Academy*, 2006, Berkeley, CA, USA.
[3] Tuters, M., Kazys, V., Beyond Locative Media, In *Networked Publics*, MIT press, 2008.
[4] Fischer, G., Cultures of Participation and Social Computing: Rethinking and Reinventing Learning and Education. In *Proc. Of ICALT*, 2009, pp. 1-5.
[5] Drew, H., Guest Editorial: Locative Media, *LEA*, Vol. 14, No. 3/4, 2006, pp. 82-84.
[6] Happel, H.-J., Towards Need-driven Knowledge Sharing in Distributed Teams. In *Proc. I-KNOW 2009*, 2009, pp. 128-139.
[7] Neuwirth, C. M., Chandhok, R., Charney, D., Wojahn, P., Kim, L., Distributed Collaborative Writing: A Comparison of Spoken and Written Modalities for Reviewing and Revising Documents. In *Proc. of CHI94*, 1994, pp. 51-57.
[8] Wikimapia. http://wikimapia.org
[9] Wikiplaces. http://www.mifki.com/wikiplaces/
[10] Giles, T., Marienek, M., Willis, K.S., Geelhaar, J., Hide&SEEK: Sharing Cultural Knowledge. In *Proc. of MULTIMEDIA '07*, 2007, pp. 481-484.
[11] Radio aporee. http://aporee.org/maps/
[12] Giaccardi, E., Eden, H., Fischer, G., The Silence of the Lands: Promoting the Virtual Museum as a Place of Cultural Negotiation. In *Proc. of New Heritage Conference*, 2006, pp. 94-114.
[13] Garrett, J.J., Customer Loyalty and the Elements of User Experience. *Design Management Review*, Vol. 17, No. 1, 2006, pp. 34-39.
[14] Borchers, J., A Pattern Approach to Interaction Design, *John Wiley & Sons*, 2001.
[15] Norman, D.A., Signifiers, not affordances. *Interactions*, Vol. 15, No. 6, 2008, pp. 18-19.

[16] Reenskaug, T., The Model-View-Controller (MVC) Its Past and Present, 2003. http://heim.ifi.uio.no/~trygver/themes/mcv/mvc-index.html
[17] Watt, W.C., Habitability. *American Documentation*, Vol. 19, 1968, pp. 338-351.
[18] Hone, K.S., Barber, C., Designing habitable dialogues for speech-based interaction with computers. *International Journal of Human-Computer Studies*, Vol. 54, No. 4, 2001, pp. 637-662.
[19] De Souza, C.S., Cypher, A., Semiotic engineering in practice: redesigning the CoScripter interface. In *Proc. of AVI2008*, 2008, pp. 165-176.
[20] Naeve, A., The Human Semantic Web Shifting from Knowledge Push to Knowledge Pull, *Int. J. Semantic Web Inf. Syst.*, Vol. 1, No. 3, 2005, pp. 1-30.
[21] Shamber L., What is a document? Rethinking the concept in uneasy times, *J. of the ASIS*, Vol. 47, No. 9, 1996, pp. 669-671.
[22] Pédauque, R.T., *Document: form, sign and medium as reformulated for electronic Documents*, 2003.
[23] Mussio P., E-Documents as tools for the humanized management of community knowledge. In *Constructing the Infrastructure for the Knowledge Economy: Methods and Tools; Theory and Practice*, 2004, pp. 27-43.
[24] Malcangi, M., Soft-computing Approach to Fit a Speech Recognition System on a Single-chip. In *Proc. of International Workshop System-on-Chip for Real-Time Applications*, 2002.
[25] Parson W.T., *Voice and Speech Processing*, Mc Graw-Hill Book Company ED, 1986.
[26] Malcangi, M., Improving Speech Endpoint Detection Using Fuzzy Logic-based Methodologies. In *Proc. of the Thirteenth Turkish Symposium on Artificial Intelligence and Neural Networks*, 2004, pp. 139-146.
[27] Malcangi, M., Grew, P., Toward Language-Independent Text-to-Speech Synthesis, *WSEAS Transactions on Information Science and Applications*, Vol. 7, No. 3, 2010, pp. 411-421.
[28] Seungmin, R., Sang-Soo, Y., Bridging the semantic gap in multimedia emotion/mood recognition for ubiquitous computing environment, *The Journal of Supercomputing*, 2010, pp.1-13.
[29] Clarkson, B., Pentland, A., Extracting context from environmental audio. *Wearable Computers, Second International Symposium on Digest of Papers*, 1998, pp.154-155.
[30] Marcante, A., Parasiliti Provenza, L., Social Interaction through Map-based Wikis, *PsychNology Journal*, Vol. 6, No. 3, 2008, pp. 247-267.
[31] Barricelli, B.R., Iacob, C., Zhu.L., Map-Based Wikis as Contextual and Cultural Mediators. In *Proc. of Community Practices and Locative Media workshop at Mobile HCI09*, 2009. http://www.uni-siegen.de/locatingmedia/workshops/mobilehci/