An Advanced Multimodal Platform for Educational Social Networks

Maria Chiara Caschera, Arianna D'Ulizia, Fernando Ferri, and Patrizia Grifoni

Institute of Research on Population and Social Policies (IRPPS)
National Research Council (CNR)
00185, Rome, Italy
{mc.caschera,arianna.dulizia,fernando.ferri,
patrizia.grifoni}@irpps.cnr.it

Abstract. Multimediality and multimodality have demonstrated significant potential supporting students in learning activities; providing them with multimodal interaction can be a crucial issue for improving accessibility to the multimedia contents in learning environments. In fact, involving multiple input and output modalities enables a broader spectrum of users with different ages, skill levels and abilities to access these contents. In this paper, we present AMPLE, an advanced multimodal platform for e-learning that provides users with a multimodal access to an educational social network. The platform is based on a distributed architecture that integrates social networking technologies and multimodal facilities, enabling a collaborative learning experience to learners.

Keywords: Learning environment, Multimodal interaction, Web 2.0 networking technologies, Social networks.

1 Introduction

The use of Web 2.0 social networking technologies, such as chatting, blogging and text messaging, enables users to share personal information, connect themselves with other users, edit and upload multimedia contents including video, audio and even 3D data. These technologies have the great potential to support formal learning contexts, as they provide a new form of engagement that is participatory and collaborative. They can be used to facilitate collaborative problem solving, to share course materials, to follow interactive lessons, and to promote peer-to-peer interaction. A series of experiments [1] [2] demonstrated the usefulness of multimediality in supporting students in their learning activities. Multimedia learning contents are more easily and intuitively accessed if available over several channels, and offered in a multimodal fashion.

This paper describes our effort in improving the access to multimedia learning contents through an Advanced Multimodal Platform for e-LEarning (AMPLE). The platform supports multimodal access to a multimedia repository, enabling the retrieval, editing and sharing of multimedia contents in an educational social network.

The remainder of the paper is structured in four main sections. Section 2 summarizes related works on social networking technologies, learning and multimodality.

R. Meersman et al. (Eds.): OTM 2010 Workshops, LNCS 6428, pp. 339–348, 2010.

The third section provides details of the proposed advanced multimodal learning environment. In Section 4, a case study is described, which explains the functioning of the system. Section 5 concludes the paper.

2 Related Work

The social networking technologies in educational environments can support learning and exchanging ideas, knowledge, skills, experience, and competencies. The communication and the collaboration are efficiently supported by the use of technology results and the development of more efficient forms of education. The use of technological media and tools provides learning environment where teachers and learners are separated by time and/or space. In this environment education takes place in a virtual learning environment that can be viewed as a software system designed to support teaching and learning by tools, such as discussion forums, blogs and whiteboards.

Several works in the literature investigate the use of social networking technologies in educational environments [3] [4] [5] [6], proving the growing interest in this research area. As Zane Berge [7, p.28] argued, "the trend in online education is toward a Web-based, desktop, virtual classroom - the result of text-based e-mail, mailing lists, conferencing and chat functions as well as the video, graphics, and audio channels that deliver interactive multimedia over the Internet". An example of educational environment, offering several learning tools to the community members, such as synchronous chat rooms, threaded discussion boards, whiteboards, file and public link sharing, is Educational MUVES [8]. Another example is the educational social network proposed by Varlamis and Apostolakis [9], which provides a knowledge base for the educational material, a profile base for the storage of learners' history and a collaboration environment for the communication and participation in synchronous activities.

In the attempt to improve accessibility to multimedia learning contents, multimodal interfaces are a powerful tool that facilitates the access to the different contents. Many research studies [10] [11] [12] emphasize the advantages of multimodal interfaces, compared to traditional graphical user interfaces, since they make human-computer communication more intuitive, natural and efficient, enabling a broader spectrum of users with different ages, skill levels and abilities to access to computational systems, and increasing the level of freedom offered to users. In the literature, various multimodal learning environments have been proposed [13] [14] [15]. MultiLezi [13] allows users to access to teaching materials across various channels, devices (the Web, the telephone, and hand-held devices) and contexts and with different modalities (speech, mouse and keyboard). It provides, indeed, a multimodal interface for accessing learning contents using a standard point-and-click interaction paradigm integrated with vocal commands. AmbiLearn [14] is another example of multimodal learning environment devoted to provide educational contents to children by using a combination of speech and pen input. Client server architecture has been used in AmbiLearn for supporting multiple users within the learning environment. Differently from AmbiLearn, the multimodal learning environment, called LEMMA [15], provides also an authoring system that enables teachers to develop and evolve multimodal learning content, in addition to an interactive environment that enables students to access these contents by speech, written text, and 2D/3D imagery.

The examined multimodal learning environments have a set of advantages and drawbacks, which are summarized in Table 1. In MultiLezi, a personalization of the multimodal interaction occurs, since multimedia content can be navigated and accessed, according to the characteristics of the user detailed in the user profile. However, it cannot be exploited in a social networking environment since it does not support multiple users. Moreover, it is addressed mainly to learners, since it does not provide an authoring environment that allows teachers to edit learning material, but only an interactive learning environment that enables students to access to learning contents. This drawback can also be found in AmbiLearn. Moreover, this system relies on the kind of supported modalities, which are only speech and pen-based gesture, without the possibility to visually interact with the system. However, AmbiLearn is the only examined system that provides a client server architecture that enables multiple users to access and collaborate within the learning environment. LEMMA, indeed, is the only examined environment that provides an authoring environment, which allows teachers to edit learning material. Moreover, similarly to MultiLezi, it enables a personalization of the multimodal interaction allowing users to adjust presentations according to their preferences. The main disadvantage of LEMMA relies on its inability to support multiple users and, consequently, collaborative learning tasks.

Table 1. Advantages and drawbacks of the examined multimodal learning environments

	MULTIMODAL LEARNING ENVIRONMENTS		
	MultiLezi	**AmbiLearn**	**LEMMA**
ADVANTAGES	• It allows to personalize the multimodal interaction to the features of the user	•It supports collaborative learning tasks	• It provides an authoring environment for editing multimedia contents • It allows to personalize the multimodal interaction to the features of the user
DRAWBACKS	• It does not support collaborative learning tasks • It does not provide an authoring environment for editing multimedia contents	•It does not provide an authoring environment for editing multimedia contents • It does not support the visual input modality, but only speech and touch	• It does not support collaborative learning tasks

The learning environment proposed in this paper joins together the advantages of the examined multimodal learning environments, trying to overcome most of their drawbacks. In particular, the proposed environment, similarly to AmbiLearn, follows a distributed architecture where learners and teachers can access the server from their own devices (laptops, as well as mobile devices), supporting collaborative learning tasks. Moreover, analogously to MultiLezi, our environment provides authoring functionalities and the possibility to personalize the multimodal interaction to the features of the user. A detailed description of the proposed learning environment and its functionalities are given in the following sections.

3 The Advanced Multimodal Platform for e-LEarning

The investigation of existing learning environments, provided in the previous section, leads us to believe that a distributed architecture, networking technologies, and multimodal facilities have to be integrated to enable a collaborative learning experience to learners. Therefore, in this paper an Advanced Multimodal Platform for e-LEarning (AMPLE) is presented, which enables the retrieval, editing and sharing of multimedia contents in an educational social network through a multimodal interface. This platform allows efficiently managing multimodal communication between people, participating in an educational social network.

In the following sections, the architecture of the proposed platform and its functionalities are described.

3.1 AMPLE Architecture

AMPLE is based on client-server architecture as depicted in Figure 1. A similar architecture has been proposed in our previous work [16], more focused on supporting the interaction in a gamed-based learning environment.

Each person (i.e. learner and teacher) can access to AMPLE from its own device that is equipped with a *multimodal interface*. Therefore, an AMPLE client includes specific I/O devices, such as, for example, display, cameras, microphone, and loudspeakers, as well as the components for extracting features from the received signals. The feature extraction occurs on the client side, since it requires limited amount of memory and computational power, whilst the recognition processes, which consist in matching the extracted features with a predefined set of patterns, are executed on the server. The *multimodal interaction management* and the *educational social network* are provided on the server side.

The *multimodal interaction management*, whose architecture has been proposed in our previous work [17], is responsible for recognizing unimodal input coming from the features extractors of each modality, appropriately interpreting these inputs, integrating these different interpretations into a semantic interpretation, and understanding which is the better way to react to the interpreted multimodal request by activating the most appropriate output devices. This component includes:

- the unimodal input recognizers, such as, for example the Automatic Speech Recognizer and the gesture recognizer, and the output generators, such as the Speech Synthesizer;
- the multimodal interpreter that integrates the recognized inputs, and applies the production rules stored in the Multimodal Grammar Repository, to parse the multimodal input;
- the modeling components, that are aimed at capturing some information used during the interpretation phase for leading up to the most probable interpretation of the user input (e.g. user, content and context modeling components);
- the multimodal output manager for generating appropriate output information, through the available output modalities (multimodal fission).

A user can interact with AMPLE for retrieving, editing and sharing multimedia contents within the educational social network. For instance, the user can ask information about the topic of a multimedia document by using speech in combination with sketch, handwriting, or pointing gesture that complete the meaning of the speech sentence. When the multimodal message is interpreted, it is sent to the educational social network in order to accordingly activate the AMPLE learning functionalities. Analogously, the multimedia information required by the user is sent to the multimodal output manager that allows visualizing it by using the appropriate output devices.

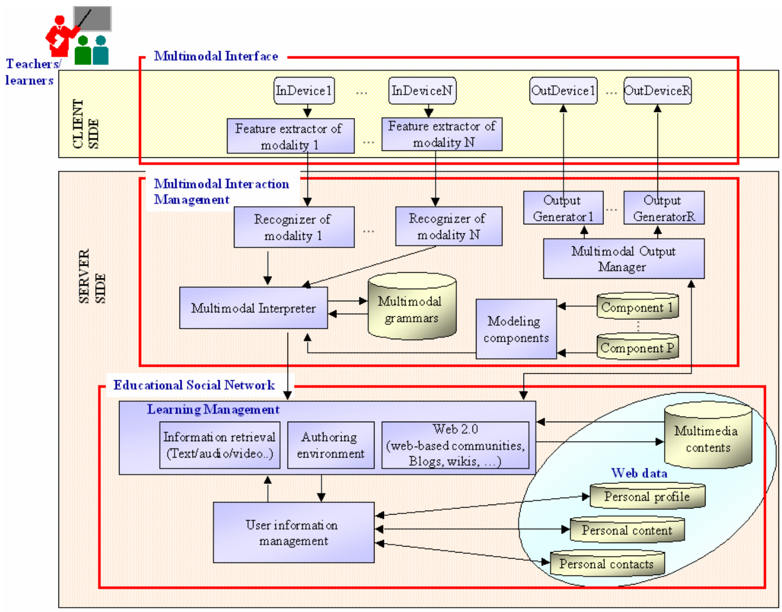

Fig. 1. Architecture of the Advanced Multimodal Platform for e-LEarning

The *educational social network* consists of two main components: the learning management system and the user information management. The learning management system has three components:

- Information retrieval Module that supports advanced indexing, search and retrieval methodologies for all content types based on semantic tagging. The multimedia learning contents are stored in a networked repository and are accessible by the multimodal interface. This networked repository provides facilities to store the following types of multimedia content: text, images, different formats of audio content, different forms of video content, Macromedia flash movies, Power-point presentations and others.
- Authoring environment that supports multimedia data management, including multimedia content editing, integration, and synchronization. The contents are edited, uploaded and managed by the multimodal interface. Once created or edited, multimedia contents are stored into the networked repository.
- Web 2.0 Module that provides social networking services, such as blogs, wikis, and web-based communities, for supporting students in the learning activities.

In particular, these services can be used for stimulating social real time interaction among students but also between teachers and students. For instance, blogs and wikis are useful tools for online discussions and forums, while video-sharing websites provide the possibility to upload video-based tutorials that can be shared with other students.

The user information management is another module of the educational social network, devoted to store and manage personal data of network members. In particular, it provides controlled access to the network and to user information, such as personal profile (user's interaction behaviors), contents and contacts. These data are contained in three networked repositories. The user information management cooperates with the learning management system, particularly, when new multimedia learning contents have to be edited from personal contents, as well as the user searches for information contained in the personal profile.

3.2 AMPLE Functionalities

AMPLE combines the features of the multimodal systems with the e-learning environments functionalities. The multimodal interaction environment allows easily accessing contents using different interaction modalities according to the contexts and user knowledge contained in the *Modeling components*.

The client side consists of the *Multimodal Interface* that was designed for allowing the use of different interaction channels in order to edit and retrieve educational materials and to post messages. Information, which are extracted by the modal features extractor of the client, are captured and recognized by the *Recognizers of modalities,* and combined using the *Multimodal interpreter* employing the *multimodal grammar*.

In the server side, the *multimodal interaction management*, combined with the *education social network* (see Figure 1), provides several functionalities to support the learning process. The retrieval of multimedia contents is enabled by the *information retrieval* module and it can be performed using different modal inputs thanks to the functionalities of the modules that compose the *multimodal interaction management*. This process is supported by information about personal profile and contents obtained by the *user information management* module. Moreover, comments and messages can be posted using different interaction modalities and the *Web 2.0* module. The interaction process is customized adapting interaction to the user profile, context and available modalities, which are information contained in the *modeling components*. The *authoring environment* manages data formats and low-level representations, and it ensures consistency of the learning contents. Moreover, it enables the author of learning content to easily write and modify contents through a multimodal user interface, and the learner cannot use contents before the author has given permission. The fruition and the visualization of information are provided by the *multimodal output manager* that supports several views of the contents, as described in Section 4.

4 A Scenario to Access Multimedia Learning Contents

AMPLE can be usefully applied to support the fruition of multimedia learning materials by multimodal interaction, and it is based on the visualization system proposed in

[18]. In AMPLE, the multimedia contents are collected in order to offer and integrate different visualization views by using the *multimodal output manager* according to the user's purpose. In particular, it provides the possibility: to cluster multimedia contents, depending on media types in which they are available, through the *authoring environment*; to visualize them on a map according to the landscape metaphor, and to provide a visual representation of the temporal evolution of the media types of the multimedia contents through the *multimodal output manager*. In particular, let us suppose that risk management materials concern: *forests fire, extreme floods events in large river basins, earthquake, governance for hospital* and *health due to e-commerce*. Those materials are available in different media types: *text, image, sound* and *video*. The user accesses the AMPLE system by using the functionalities of the *multimodal interface*.

Let us suppose that the user says by speech:

"Show me information about this topic"

and she/he selects the icon connected with the risk management materials by gesture. The *multimodal interaction management* recognizes unimodal inputs, which come from the modality *features extractors,* and it reacts to the interpreted input (by the *multimodal interpreter*) providing the visualization of the multimodal learning contents considering media types by which they are provided according to the landscape metaphor (see Figure 2.a). In the example shown in Figure 2.a, material about *forest fire* risk management is available by *video*, *image* and *sound,* while learning contents about *earthquake* risk management is available by *video*, *text* and *sound*.

In order to analyze the accessibility levels of the multimedia learning contents, the visualization focuses on class of media types showing them as 3D half-ovoid solids having circular horizontal sections. The visualization allows displaying colored areas that represent classes of media types and a specific learning content is placed in the landscape. The width of each area of class of media types is proportional to the number of multimedia learning contents that are provided by the specific media type. The height of the position, associated with each multimedia content on the ovoid, reflects the *media access* level of the multimedia content according to the media types. Therefore, isolines in the landscape identify levels of media access of multimedia contents in the classes of media types. The multimedia content that has the higher value of media access level according to media type is located in the higher position on the solid visualizing the specified class of media types. Moreover, each multimedia content can be provided by more than one class of media type. For example, Figure 2.a shows three multimedia contents (*Forest fire, Hospital,* and *Health due to e-commerce*) that can be provided by *image* (yellow ovoid). In this figure, the multimedia content *Forest fire* is also provided by *video* and *sound* (visualized using black edges). This visualization of the educational learning contents considering media types can support the analysis of available learning contents about risk management materials in an educational social network. It enables users to detect which contents are available through specific media types in order to evaluate which channels can be used to access to the document. This functionality is originated from the fact that, frequently, the needed contents are not available by different media types and only specific channels can be used to access them.

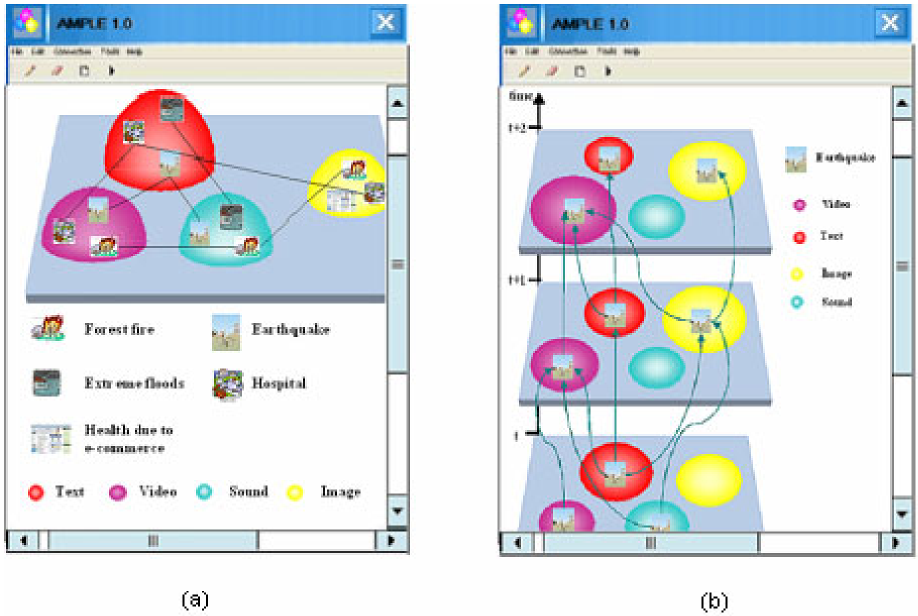

Fig. 2. Possible visualizations offered by the AMPLE system: (a) Multimedia contents, visualized according to the media types, and their access level; (b) Visualization of the evolution of the *earthquake* contents according to the media types at time t, t+1, t+2

The presented visualization can be also usefully applied to find out which are the people who have edited the multimedia learning contents, and to establish a direct connections with them creating shared awareness by the *Web 2.0* module of the AMPLE system. In this case, the dimensions are: 1) the 3D virtual representation of colored ovoids that represents the available multimedia learning contents; 2) people who have edited content, which are positioned in different points of the ovoid according to the number of multimedia learning contents they have edited in the network. The person who has inserted the higher number of multimedia learning contents is positioned in the higher position on the ovoid.

Moreover, AMPLE offers the possibility to visualize the evolution of the media types of the multimedia contents during the time (see Figure 2.b). In fact, a further dimension, which can be visualized alternatively to the media access level, is the *time*. For example, let us suppose that a user says by speech:

"Visualize the temporal evolution of the kinds of media by which I can access this material"

and she/he simultaneously points the icon of the content about *earthquake* on the touch screen display. The *multimodal interaction manager* interprets and recognizes the input, and the *educational social network* retrieves the information requested by the user. This information is presented by the *multimodal interface,* displaying the evolution over time of *earthquake* content according to the media types by which it has been provided. Figure 2.b presents the transformations of the media types at time t, t+1, t+2. This figure shows the evolution of multimedia *earthquake* contents, which at the time *t* are provided by *video*, *text* and *sound*, and at the time *t+1* by *video*, *text* and *image* due to the fact that *sound* contents have been changed into *video* and *image* contents, and some *text* contents have been also changed into *video* contents. Considering time *t+2,* the figure shows that the *video* contents increase because both *text* and *image* contents have changed into *video*.

Information obtained from the sequence of temporal layers allows analyzing how the media types of the learning contents are evolving during time. The evolution of the media types of learning contents is due to the feedback from users' interactions, and it points at fitting the users' needs [19]. Therefore, analyzing the changing of media types allows understanding the transformation of the user interaction behaviors in the fruition of learning contents.

5 Conclusion and Future Work

The spreading of the Web 2.0 social networking technologies has offered great support to formal learning contexts improving the processes of knowledge building, working together, information sharing and problem solving. In this paper, social networking technologies for learning environments have been combined with multimodal interaction functionalities in order to access multimedia learning contents. Starting from an overview of the main features of social networking technologies and an analysis about how multimodal interaction supports the access to multimedia learning contents, the AMPLE has been designed in order to efficiently support multimodal communications among people in a community and the participation in educational social networks. The client-server architecture of the AMPLE system has been depicted describing the client side composed of the *multimodal interface*, and the server side made up of the *multimodal interaction manager* and the *educational social network*. Examples of the functionalities provided by this system have been described focusing on the visualizations offered by this system.

The presented system has been tested and evaluated in the scenario of scientific seminars participation, organization and management involving 77 people, with ages from 25 to 65. All the participants answered to a questionnaire to investigate the users' satisfaction on the proposed system in order to express a qualitative evaluation. It consisted of a short interview that involved all the participants. The evaluation of the proposed system has produced very satisfactory results both, in term of stimulating social networking and in term of users' satisfaction. An interesting task for future work is the evaluation of the system by comparing the usability of AMPLE with other existing learning environments. Such a comparison will provide possible improvements on the system.

References

1. Mayer, R.E., Anderson, R.B.: The instructive animation: Helping students build connections between words and pictures in multimedia learning. Journal of Educational Psychology 84, 444–452 (1992)
2. Mayer, R.E., Sims, V.K.: For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. Journal of Educational Psychology 86, 389–401 (1994)

3. De Wever, B., Mechant, P., Hauttekeete, L., Veevaete, P.: E-learning 2.0: social software for educational use. In: Proceedings of Ninth IEEE International Symposium on Multimedia Workshops (ISMW 2007), pp. 511–516 (2007)
4. Lockyer, L., Patterson, J.: Integrating Social Networking Technologies in Education: A Case Study of a formal learning environment. In: Proceedings of 8th IEEE International Conference on Advanced Learning Technologies, Santander, Spain, July 1-5, pp. 529–533 (2008)
5. Collins, D., Schwarte, A.R., Alkire, K., Adkinson, P., Rosichelli, M., Kallish, A.: Social Networking for Learning Communities: Using e-portfolios, blogs, wikis, pod casts, and other Internet based tools in the foundation art studio. FATE in Review 29, 4–9 (2008)
6. Jin, L., Wen, Z.: An Augmented Social Interactive Learning Approach through Web2.0. In: Proceedings of 33rd Annual IEEE international Computer Software and Applications Conference (COMPSAC 2009), Seattle, Washington, pp. 607–611 (2009)
7. Berge, Z.: Components of the online classroom. In: Weiss, R., Knowlton, D., Speck, B. (eds.) Principles of effective teaching in the online classroom, San Francisco, Jossey Bass, pp. 23–28 (2000)
8. Cooper, J.: Educational MUVES: Virtual Learning Communities. Journal of Education, Community and Values (2003) , `http://bcis.pacificu.edu/journal/2003/09/cooper/cooper.php` (accessed Septmeber 2009)
9. Varlamis, I., Apostolakis, I.: A Framework for Building Virtual Communities for Education. In: EC-TEL 2006, pp. 165–172 (2006)
10. Oviatt, S.L., DeAngeli, A., Kuhn, K.: Integration and synchronization of input modes during multimodal human-computer interaction. In: Proceedings of Conference on Human Factors in Computing Systems, pp. 415–422
11. Oviatt, S.L., Cohen, P.R.: Multimodal interfaces that process what comes naturally. Communications of the ACM 43(3), 45–53 (2000)
12. Oviatt, S.L., Cohen, P.R., Wu, L., Vergo, J., Duncan, L., et al.: Designing the user interface for multimodal speech and pen-based gesture applications: State-of-the-art systems and future research directions. Human-Computer Interaction 15, 263–322 (2000)
13. Sbattella, L., Mainetti, L., Barbieri, T., Bianchi, A., Bruna, S., Pernici, B.: MAIS MultiLezi: an adaptable multichannel and multimodal environment for e-learning. Rivista del Politecnico di Milano 10, 42–47 (2006)
14. Hyndman, J., Lunney, T., Mc Kevitt, P.: AmbiLearn: Ambient Intelligent Multimodal Learning Environment for Children. In: Proceedings of the 10th Annual PostGraduate Symposium On The Convergence of Telecommunications Networking & Broadcasting PG Net, June 22 – 23, pp. 277–282. John Moores University, Liverpool (2009)
15. Breisinger, M., Höllerer, T., Ford, J., Folsom, D.: Implementation and Evaluation of a 3D Multi Modal Learning Environment. In: Kommers, P., Richards, G. (eds.) Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications, pp. 2282–2289 (2006)
16. Caschera, M.C., D'Ulizia, A., Ferri, F., Grifoni, P.: Multimodality in Game-based Learning Environments. IEEE Learning Technology Newsletter 12(1), 20–22 (2010)
17. D'Ulizia, A., Ferri, F., Grifoni, P.: Toward the Development of an Integrative Framework for Multimodal Dialogue Processing. In: Meersman, R., Tari, Z., Herrero, P. (eds.) OTM-WS 2008. LNCS, vol. 5333, pp. 509–518. Springer, Heidelberg (2008)
18. Caschera, M.C., Ferri, F., Grifoni, P.: SIM: A dynamic multidimensional visualization method for social networks. PsychNology Journal 6(3), 291–320 (2008)
19. Leung, C.H.C., Chan, W.S., Liu, J.: Collective Evolutionary Indexing of Multimedia Objects. In: Gervasi, O., Taniar, D., Murgante, B., Laganà, A., Mun, Y., Gavrilova, M.L. (eds.) ICCSA 2009. LNCS, vol. 5592, pp. 937–948. Springer, Heidelberg (2009)