

A Model to Design Multimedia Software for Learners with Visual Disabilities

Jaime Sánchez

Department of Computer Science, University of Chile
Blanco Encalada 2120, Santiago, CHILE
jsanchez@dcc.uchile.cl

Abstract. Current interactive multimedia learning software can not be accessed by learners with disabilities. This is the case for students with vision disabilities. Modeling techniques are necessary to map real world experiences to virtual worlds by using 3D auditory representations of objects for blind people. In this paper we present a model to design multimedia software for blind learners. The model was validated with existing educational software systems. We describe the modeling of the real world including cognitive usability testing tasks by considering not only the representation of the real world but also modeling the learner's knowledge of the virtual world. Finally, we analyze critical issues in designing software for learners with visual disabilities and propose some recommendations and guidelines.

1 Introduction

A great amount of educational software has been developed for supporting learners with disabilities. Software for blind people aims to increase their access to current computing materials based on Graphic User Interfaces, GUIs, such as games, educational software, and Web navigation systems.

Clearly the task of developing software for people with visual disabilities has some complexities. For blind learners we face the problem of constructing interfaces that do not rely on graphics. However, we can find some similarities in the process of designing and constructing computer software for people with different types of disabilities. This is the case when developing cognitive systems aiming at modeling and implementing the real world by a computer system.

Problems in constructing learning software for people with disabilities arise when the model is presented to the user to interact with. Multimedia software for sighted learners transmit this model to the learners by using graphics (with or without animation), sounds and text, taking advantage of a wide spectrum of the computer's multimedia capabilities. People with visual disabilities have shrunk this spectrum. This forces the software designer to project the information kept in the model provided by the student through the available auditory channel. Additionally, non-traditional interaction modes such as haptic devices can be used. The same considerations are valid for the construction of the learner's model.

Several virtual reality systems and virtual environments combined with appropriate human-computer interfaces have been used to enhance sensual capabilities of people with sensory disabilities. This is the case of presenting graphic information by text-to-speech and 3D auditory navigation environments to construct spatial mental representations and to assist users in acquiring and developing cognitive skills [20].

A sonic concentration game described in Roth, Petrucci, Assimacopoulos & Pun [15] consists of matching different pair levels of basic and derived geometric shapes. To represent geometric shapes it is necessary to build a two-dimensional sound space. The concept allows the shape to be rendered by the perception of moving sound in a specific plane. Each dimension corresponds to a musical instrument, and raster points correspond to pairs of frequencies on a scale.

VirtualAurea by Sánchez [16] was developed after interactive sound-based virtual environments proved to trigger the development of cognitive spatial structures in blind children. VirtualAurea is a set of spatial sound editors that can be used by parents and teachers to design an ample variety of spatial maps such as the inner structure of the school and rooms, corridors, and other structures of a house.

This paper proposes a model for developing multimedia learning systems for learners with visual disabilities. The model includes various steps and recommendations by considering key issues for conceptualization and implementation. Special attention is paid to the feedback issue considered to be a critical point in existing software.

2 Model

2.1 Development of educational software

We propose a model for creating educational software for people with visual disabilities. The modeling process starts with the definition of desired cognitive skills. Then we create a virtual environment that includes a navigable world by using adequate modeling languages, dynamic scene objects, and acting characters. Scenic objects are characterized by graphic and acoustic attributes; character's actions are based on deterministic and non-deterministic plans in the same way as in interactive hyperstories described by Lumbreras and Sánchez [11]. The learner explores the virtual world by interacting with appropriate interfaces and obtains

interactive feedback. The learner's actions, such as sound reproductions, are collected, evaluated, and classified based on student's modeling and diagnostic subsystems. The modeling process follows the steps illustrated in Figure 1 [19].

We define cognitive skills in a real world situation, for example self-motivating activities, drill and practice applications, problem solving, and leisure time occupations. Objects in fictitious world scenarios are constructed of geometric primitives. They are characterized by acoustic attributes and grouped into components with input and output slots. Control elements of the virtual world are represented by acoustic elements, known as icons and earcons.

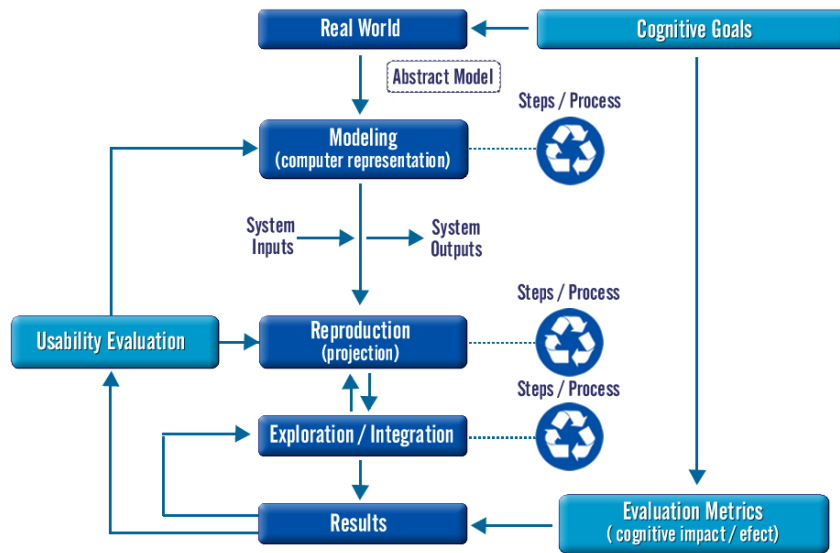


Figure 1. A model for designing educational software for children with visual disabilities.

We develop an internal computer representation and define a geometric environment containing a 2D or 3D visual and acoustic model without considering the limitations of potential users. Modern object-oriented modeling languages are powerful tools for building virtual worlds using scene graphs, interaction, and animation facilities. We insist on the necessity of special editors for teachers and learners to create synthetic models independently. According to the problem and target users, problem-specific correspondences between graphic and acoustic attributes or properties can be used to reduce the model to its acoustic projection. The resulting model meets the requirements in terms of certain information channels that should not be used in such a way that cannot be explored by people with visual disabilities. Another way is to generate the model directly by using a special editor for visually-impaired learners.

The acoustic representation of the model for learners with visual disabilities uses spatial sound. Interaction and navigation are based on acoustic control elements.

The learner explores the object space by navigating, interacting with suitable interfaces, interpreting, and reproducing the structures. This can be done through navigating without changing viewpoints or by using internal representations of users

by giving them the illusion of being part of the virtual scene. A blind person explores neighboring models by grasping them, tracking objects or listening to typical sounds. Sound-emitting objects help them to build a mental model of real or fictitious worlds. Additionally, the learner may build an external reconstruction of the mental model or try to rebuild the acoustic model as it was perceived and imagined after exploring the model space. We must be sure that conditions during the reconstruction process are always the same. Therefore, the interfaces involved are calibrated accordingly.

Depending on the particular parameters and entities of the model, error measures between internal representations and the model reproduced by the learner are defined. Learner's actions are collected, evaluated, and classified. The outcome is transformed to a user-adapted aural output.

2.2 Workflow

AudioDoom was used as an exemplary application for blind children. Thus, we validated our model by describing in more detail the modeling workflow. The whole testing is described in Baloian, Luther, and Sánchez [3]. AudioDoom allows blind children to explore and interact with virtual worlds created with spatial sound [10] [17]. AudioDoom has been usability tested with more than forty Chilean children aged 7-12 in school settings for blind children [16].

We followed the modeling steps introduced above. First, a computer-based representation is derived from the real world scenario by means of abstraction and reduction without considering the limitations of potential users. Then, the computer-based model is reproduced (projected) to an appropriate acoustical internal model explored by people with sensory disabilities through the use of available communication channels. Appropriate editors support the modeling process. Important model entities and parameters must be identified at this stage. By interacting with the model, the learner makes an internal or external reconstruction, which is later evaluated (see Figures 2a and 2b). This can be done by using appropriate error measures depending on the learner, the computer-based (internal) representation, and the reconstructed model. Finally, the degree of similitude is derived from the error measure and the result is displayed on a learner-adapted output and used for modeling the learner's knowledge.

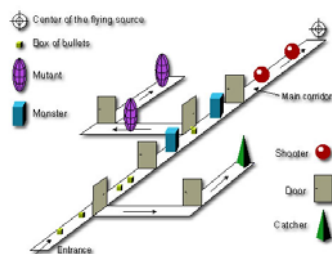


Figure 2a. Virtual game world in AudioDoom



Figure 2b. Reconstructed external model

The modeling workflow in AudioDoom starts with the computer representation of the virtual game world (a simple labyrinth with one main corridor and two secondary corridors including entities and objects) which is projected to an internal model consisting only of sounds. At this stage, the role of volume, frequency, melody, and rhythm in representing different forms, volumes, and distances is analyzed. Learners interact with this model by ‘virtually walking’ through the labyrinth with a keyboard, mouse, and ultrasonic joystick. Sound-emitting objects help them to build a mental model of the labyrinth. Finally, they make concrete mental modeling with Lego blocks and try to rebuild the internal model as it was perceived and imagined after exploring the spatial structure. Different types of blocks represent objects of the virtual world in the computer representation.

The concrete reconstruction is checked by a human tutor against the game world or by a camera and image processing algorithms to look for any spatial correspondence with the computer representation of the original world model by evaluating the error measure.

As we can see, despite object representation, interfaces, perception modes, and error measures, there are important tasks which should be undertaken when developing systems for blind people. A critical task in this modeling process is the reduction of the original model to only acoustic output.

2.3 Testing

We can describe in more detail two steps of the modeling process. To process an external model and to evaluate error measures, we assume that a blind child reconstructs AudioDoom’s maze structure by using Lego blocks. Each block is individually marked by black bars and dots. A blind user perceives different blocks, selects the appropriate ones, and rebuilds the mental image by using the blocks one by one.

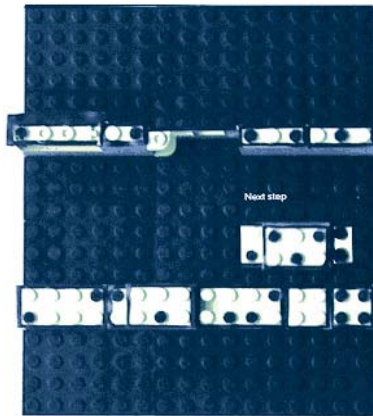


Figure 3. Reconstructed LEGO-model

After each step a picture is taken by a digital camera placed on a fixed position over the scene. We highlight a typical state of the reconstruction process and indicate the next step by adding a new Lego block on the lower wall. Figure 3 shows

the situation before and after the next step. Reducing the colors to black and white makes it possible to apply low level image processing routines to detect the new Lego block. After a calibration of the two pictures, we can localize the new block through an XOR-operation on both images resulting in one or two dashes, otherwise in new dots. Starting from these new picture elements we calculate the position and type of the new Lego block. Finally, the learner's model representation is updated.

Thus it is proved that by certain low-level operations on succeeding pictures the external model can be transformed into an internal representation which is used to feed an error measure function. A degree of fidelity is derived and displayed by text-to-speech. For a more sophisticated approach using Lego RCX robots see Ressler & Antonishek [14].

In AudioDoom the differences concerning the parameters volume, articulation, stress, and intonation must be visualized. From these visual patterns important parameters are detected by means of picture processing and in connection to the stored reference patterns. An intelligent component of the system interprets these results, displays a visualization of the error function and transforms them into instructions for the learner. Whereas the wave amplitude image provides in the xy-plane any feedback concerning volume and rate, omitted syllables are characterized by a lack of colored areas, intonation and articulation by the parameters hue and saturation as well as the shape of the curves. Details are described in Hobohm [8] and Gräfe [7].

For designing and developing information-equivalent interfaces for people with visual disabilities we followed a similar model work flow as described above in two new systems. AudioBattleShip [18] is an interactive version of the board game Battleship, providing different interfaces for both sighted and blind people to enhance collaboration and cognition (see Figure 4). Ebbinghaus's forgetting function experiments were implemented in a project concerning the historical replication of key experiments in psychology [5].



Figure 4. AudioBattleShip input device for blind players [18]

3 Guidelines and recommendations

Virtual environments (VE) can be used to simulate aspects of the real world not physically available to users due to a wide variety of reasons. They become more realistic through multimedia displays which include haptic and auditory information. According to Colwell et al. [6] and Paciello [12], there are several domains in which VEs can be used to build educational software for people with visual disabilities:

A. In education a virtual laboratory assists students with physical disabilities in learning scenarios. Possible applications concern problem solving, strategic games, exploring spaces or structures, and working with concrete materials. Special VE interfaces such as head-mounted devices, the space mouse, and gloves are often included.

B. Training in virtual environments deals with mobility and cognitive skills in spatial or mental structures.

C. Rehabilitation is possible in the context of physical therapy - a recovery of manual skills or learning how to speak and listen to sound can be targeted.

D. Access to educational systems is facilitated via dual navigation elements such as earcons, icons, and haptic devices.

Our idea is to support learners with visual disabilities in building conceptual models of real world situations as sighted users do. Our approach is comparable to the one introduced by Zajicek et al. [26].

We can identify four important common elements and aspects in modeling for software for people with visual disabilities:

1. The conceptual model is a consequence of mapping the real or fictitious world situation into a computer model using digital media by applying adequate modeling languages.

2. The perceptual model is created by developing a perceptual model and a script for the dynamic changes of the model. It can be perceived by the learner using only available information channels and considering the type of disabilities of end-users. It is important to provide surprising elements to call attention in order to enhance the perception process. The computer model description should be based on text. Explanation of graphic objects should be given in caption form. This text can be presented to the blind by using a text-to-speech plug-in and a Braille display. Intuitive correspondences between graphic and aural objects must be defined. Attention must be paid to the fact that only a small number of sounds can be memorized. Also, melodies that help to identify objects should be used.

3. The implementation design provides icons and earcons in parallel. If there are animated image sequences or videos with sound, subtitles and moving text-banners should be used. Sound provides a rich set of resources which complement visual access to a virtual world. The four types of audio examined by Ressler and Wang

[13] are ambient background music, jingles, spoken descriptions, and speech synthesis. Ambient music can play and vary as the user moves from one room to another, providing an intuitive location cue. Jingles or small melodies should characterize special objects and control elements. Spoken descriptions of objects can play as the viewer moves closer to an object. Speech synthesizers can read embedded text from the nodes in a scene graph representation. Recent Web-languages provide anchor node descriptions, EnvironmentNodes or WorldInfo nodes. Internet accessible speech synthesizers supply easy access to text-to-speech technology.

4. The implementation tools concerning special editors or languages such as Java, Java3D, VRML, OpenGL, and DirectX should be used. VRML defines a standard language for representing a directed acyclic graph containing geometric information nodes that may be communicated over the Internet, animated, and viewed interactively in real-time. VRML also provides behavior, user interaction, sensors, events, routes and interpolators for interactive animation. User interaction, time dependent events or scripts can trigger dynamic changes in some properties of an object. VRML viewers are available not only as plug-ins to Internet browsers, but also as interactive objects that may be embedded into standard Office documents. However, the actual version does not yet support collaborative scenarios. As shown in Sánchez, Baloian, Hassler & Hoppe [18], a special platform for replicated collaborative applications such as MatchMaker (TNG) together with Java was successfully applied. This is a summarizing description of some tools that are currently available. We have focused on those that are platform independent and based on international standards.

4 Conclusion

This paper presents an integrated model to design multimedia for learners with disabilities. We provide an overview of the state of the art in the field of existing educational software for people with visual disabilities. We focus on transferring the real world into appropriate computer representations by introducing an integrated methodology for modeling the real world for these people. The model starts with the definition of targeting cognitive skills. Then a virtual environment is created including a navigable world through the use of a modeling language, dynamic scene objects, and acting characters. The learner explores the virtual world through appropriate interfaces and obtains immediate feedback. Actions such as sound reproductions are collected, evaluated, and classified based on student's modeling and diagnostic subsystems. The usability of the model is illustrated with software for blind people. We believe that by using models such as the one presented here the process of software design can be facilitated and improved.

Some educational software has already been developed for people with visual disabilities. Then by following our model we can make generalizations and recommendations for designing and improving educational software. The development of computer systems for people with visual disabilities should no longer appear as isolated handcrafted efforts. Rather, efforts should be made to

systematize the construction of these systems. Recent advances in hardware and software design support our idea and provide hope that the technological and educational foundation for such systems has already been laid.

5 Acknowledgement

This report was funded by the Chilean National Fund of Science and Technology, Fondecyt, Project 1030158.

References

1. Alonso, F., de Antonio, A., Fuertes, J. L., and Montes, C. (1995). Teaching Communication Skills to Hearing-Impaired Children. *IEEE Multimedia*, 2, (4), 55-67.
2. Baloian, N., Luther, W. (2002). Visualization for the Mind's eye. Workshop on Software Visualization, Dagstuhl, 20-25, May 2001, Software Visualization, State-of-the-Art Survey LNCS 2269, (St. Diehl ed.) Springer, 354-367.
3. Baloian, N., Luther, W. and Sánchez, J. (2002). Modeling Educational Software for People with Disabilities: Theory and Practice. Proceedings ASSETS 2002 conference, July 8-11, 2002, Edinburgh, Scotland, 111-118.
4. Baloian, N., Luther, W. (2003). Various modeling aspects of tutoring systems for people with auditory disabilities. IFIP SECIII conference, July 22-26, 2002, Dortmund, in Informatics and the digital society, social, ethical and cognitive issues, eds. Tom J. van Weert, Robert K. Munro, Kluwer, 197-206.
5. Biella, D., Luther, W., Mietzel, G., Musahl, H.-P., Wormuth, L. (2003). Replication Of Classical Psychological Experiments In Virtual Environments (Replicave). Proceedings ECEL03, 6.-7. Oct. 2003, Glasgow, UK.
6. Colwell, Ch., Petrie, H., Kornbrot, D., Hardwick, A., Furner, S. (1998). Haptic Virtual Reality for Blind Computer Users. The Third Annual ACM Conference on Assistive Technologies, April 15-17, Marina del Rey, California, USA, 92-99.
7. Gräfe, J. (1998). Visualisierung von Sprache und Erkennung sprechtypischer Parameter und ihre Veränderung bei Spätertaubten. Diploma-Dissertation, Duisburg.
8. Hobohm, K. (1993). Verfahren zur Spektralanalyse und Mustergenerierung für die Realzeit-visualisierung gesprochener Sprache. PhD-Dissertation, TU Berlin.
9. ISAEUS. (1997). Speech Training for Deaf and Hearing-Impaired People, http://www.ercim.org/publication/Ercim_News/enw28/haton.html.
10. Lumbreras, M., Sánchez, J. (1998). 3D aural interactive hyperstories for blind children *International Journal of Virtual Reality* 4 (1), 20-28. .
11. Lumbreras, M., Sánchez, J. (1999). Interactive 3D Sound Hyperstories for Blind Children. Proceedings of ACM CHI '99, Pittsburg PA, USA, 318-325.
12. Paciello, M. (1998). Making the Web accessible for the deaf, hearing and mobility impaired. Yuri Rubinsky Insight Foundation . <http://xml.coverpages.org/yuriInsightFound.html>

13. Ressler, S. and Wang, Q. (1998). Making VRML Accessible for People with Disabilities. In Proceedings of ACM ASSETS '98, the Third Annual ACM Conference on Assistive Technologies, Marina del Rey, California, USA, 144-148, <http://www.itl.nist.gov/iaui/ovrt/people/sressler/vrmlsigcaph.pdf>.
14. Ressler, S., Antonishek, B. (2001). Integrating Active Tangible Devices with a Synthetic Environment for Collaborative Engineering. Proceedings of the 2001 Web3D Symposium, Paderborn, Germany, Febr. 19-22, 93-100, <http://www.itl.nist.gov/iaui/ovrt/people/sressler/tangible3.pdf>
15. Roth, P., Petrucci, L., Assimacopoulos, A., Pun, Th. (2000). Concentration Game, an Audio Adaptation for the blind. CSUN 2000 Conference Proceedings, Los Angeles, USA.
16. Sánchez, J. (2001). Interactive virtual acoustic environments for blind children. Proceedings of ACM CHI '2001, Seattle, Washington, April 2-5, 2001, 23-25.
17. Sánchez, J., Lumbreras, M. (1999). Virtual Environment Interaction through 3D Audio by Blind Children. Journal of Cyberpsychology and Behavior, CP&B 2(2), 101-111.
18. Sánchez, J., Baloian, N., Hassler, T. Hoppe, U. (2003). AudioBattleship : Blind Learners Collaboration through Sound. Proceedings of ACM CHI 2003, April 5-10, Ft. Lauderdale, Florida, USA, 798-799.
19. Sánchez, J., Baloian, N., Flores H. (2004). A methodology for developing audio-based interactive environments for learners with visual disabilities. Proceedings of The World Conference on Educational Multimedia, Hypermedia & Telecommunications ED-MEDIA 2004. June 21-26, 2004. Lugano, Switzerland, pp. 124
20. Savidis, A., Stephanidis, C., Korte, A., Crispie, K., Fellbaum, K. (1996). A generic direct-manipulation 3D-auditory environment for hierarchical navigation in non-visual interaction. Proceedings of ACM ASSETS 96, 117-123.
21. SpeechViewer III (2003). http://www-306.ibm.com/able/solution_offerings/snsspv3.html
22. VRML Consortium (1997). VRML97 International Standard Specification (ISO/IEC14772-1:1997), <http://www.web3d.org/Specifications>.
23. Visual Talker (1999), http://www.nttc.edu/resources/funding/awards/doed/phase1/sbir99/doed_sbir991.asp
24. Wans, Cl. (1999). Computer-supported hearing exercises and speech training for hearing-impaired and postlingually deaf. Assistive Technology Research Series, Vol 6 (1), IOS Press, 564-568.
25. Willett, D. (2000). Beiträge zur statistischen Modellierung und effizienten Dekodierung in der automatischen Spracherkennung. PhD-Dissertation, GMU Duisburg.
26. Zajicek M., Powell C., and Reeves C. (1998). A Web Navigation Tool for the Blind. Proceedings. of ACM ASSETS'98, the 3rd ACM/SIGCAPH on Assistive Technologies, Los Angeles, USA, 204-206.