Using Web3D Technologies for Visualization and Search of Signs in an International Sign Language Dictionary

Fabio Buttussi*
HCI Lab
Dept. of Math and Computer Science
University of Udine
via delle Scienze, 206
33100 Udine, Italy

Luca Chittaro†
HCI Lab
Dept. of Math and Computer Science
University of Udine
via delle Scienze, 206
33100 Udine, Italy

Marco Coppo‡
Italian Deaf Association (ENS)
LIS Working group - Udine
via del Pozzo, 36
33100 Udine, Italy

Abstract

Sign languages are visual languages used by deaf people to communicate. As with spoken languages, sign languages vary among countries and have their own vocabulary and grammar. Therefore, the different deaf communities need a dictionary that associates signs to the words of the spoken language of their country as well as dictionaries which translate signs from a sign language to another. Several researchers proposed multimedia dictionaries for sign languages of specific countries, but there are only a few proposals of multilanguage dictionaries. Moreover, current multimedia dictionaries suffer from serious limitations. Most of them allow only for a word-to-sign search, while only a few of them exploit sign parameters (i.e., handshape, orientation, location, and movement) to allow for a sign-to-word search. Current solutions also commonly use pictures or videos to represent signs and their parameters, but 2D images are often misleading for a correct identification (e.g., recognizing an handshape can be very difficult due to occlusions). This paper aims at facing the above described issues, exploiting Web3D technologies such as X3D and H-Anim humanoids to better understand signs and to simplify sign-to-word and sign-to-sign search, by proposing an online international sign language dictionary, called 3DictSL. The paper presents the client-server architecture of 3DictSL and authoring tools which allow deaf communities to extend the dictionary with their own language. As a practical case study, the paper discusses the implementation of Italian Sign Language (LIS).

CR Categories: K.4.2 [Computers and Society]: Social Issues—Assistive technologies for persons with disabilities; H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Animations and Artificial, augmented, and virtual realities; I.3.7 [Computers Graphics]: Three-Dimensional Graphics and Realism—Animation; I.3.6 [Computers Graphics]: Methodology and Techniques—Interaction techniques; H.5.2 [Information Interfaces and Presentation]: User Interfaces—Graphical user interfaces (GUI) and User-centered design.

Keywords: sign languages, online dictionaries, X3D, H-Anim, virtual humans, visualization, search

*e-mail: buttussi@dimi.uniud.it
†e-mail: chittaro@dimi.uniud.it
‡e-mail: udine@ens.it

1 Introduction

Sign languages are visual languages used to communicate by moving hands and body. Sign and spoken languages rely on different communication channels: sign languages rely on the visual channel, while spoken languages rely on the sound channel. Figure 1 provides an example of how the same concept can be communicated on the two channels: in a spoken language, such as Italian, the sender says a word which is associated to a concept, the receiver ears that word and associates it to the intended concept; in a sign language, such as Italian Sign Language (LIS), the sender performs a sign by moving hands and body, the receiver watches the movements and associates them to the concept. Since sign languages rely on the visual channel, they can be used by people who suffer from total or partial hearing loss.

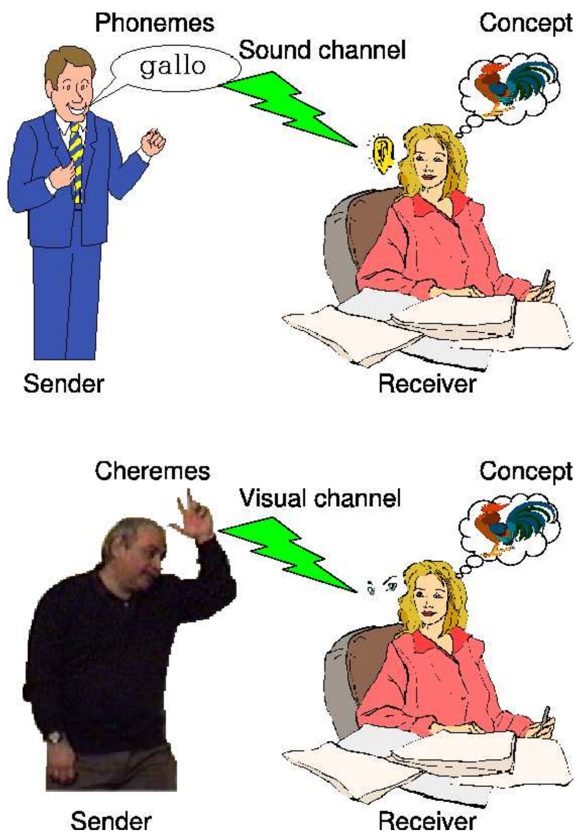

Figure 1: Example of communication through the visual and the sound channel.

While a word in a spoken language is composed by phonemes and two different words differ by at least one phoneme, a sign is composed by cheremes, which are the basic units of signed communication, and two different signs differ by at least one chereme[1]. The American linguist William Stokoe was the first to work on the cheremes of American Sign Language (ASL). Stokoe pointed out that a sign consists of and is uniquely identified by four cheremes, one for each of the following parameters:

- handshape, i.e. the position of fingers and/or their movement;
- palm orientation, e.g., hand upturned, facing the sender;
- location, i.e. the part of the body or a place close to it where the sign starts to be performed;
- movement, i.e. the sequence of positions of the hands in space during performance of the sign.

The above described parameters apply to all sign languages, not only ASL, but the set of cheremes for each parameter varies among countries, as it happens for some letters of the alphabet: for example, a particular handshape may be used in LIS, but not in ASL. Recently, some sign languages have introduced facial expression as an additional parameter.

There are several sign languages in the world and each of them has its own cheremes, vocabulary, and grammar rules. Therefore, deaf people of different countries need dictionaries that associate signs to the words of the spoken language of their country as well as dictionaries which translate signs from a sign language to another. For example, an Italian deaf may need both a LIS-Italian dictionary to write a document and a LIS-ASL dictionary to learn ASL signs and thus be able to communicate with deaf people who use that language.

Unfortunately, traditional paper dictionaries [Stokoe et al. 1976; Radutzky 1992] do not provide the best solution for sign languages, since it is very difficult to represent movements with static images or textual descriptions. Therefore, since the early '80s, several researchers [Wilcox et al. 1994; Montagnana et al. 2005; Herreweghe et al. 2004; BritishSignLanguage.com 2000; Michigan State University Communication Technology Laboratory 2000; VCOM3D 2004] have proposed multimedia dictionaries for the sign language of specific countries, but, unfortunately, there are only a few and very preliminary proposals of a multilanguage dictionary [Waterfall Rainbows and Solarstorm Design 2006].

Moreover, current multimedia dictionaries suffer from serious limitations. Most of them allow for a word-to-sign search, but only a few exploit sign parameters to allow for a sign-to-word search. They also commonly use pictures or videos to represent signs and their parameters, but 2D images may be particularly misleading for a correct identification: for example, recognizing an handshape from a 2D image may be very difficult because of finger occlusions, and movement projection on a 2D plane may be ambiguous. Finally, technologies employed by most projects are proprietary and available only on particular platforms.

Therefore, after analyzing the needs of deaf people and considering potentials and limitations of current proposals, we identified the following goals for the design of a multimedia sign language dictionary:

- providing deaf people with a multilanguage dictionary, so that they can learn how to perform a sign in foreign sign languages;
- allowing deaf people to better understand how to perform signs by means of 3D animations;
- providing easy-to-use word-to-sign, sign-to-word, and sign-to-sign search functionalities;
- supporting different platforms;
- allowing deaf communities to easily extend the dictionary with their own languages.

To achieve these goals, this paper proposes an architecture for an online international sign language dictionary, called 3DictSL. To better visualize signs and to simplify sign-to-word and sign-to-sign search, the architecture exploits standard Web3D technologies such as X3D and H-Anim, which are available on different platforms. To allow deaf communities to add the support for their own sign languages, the proposed architecture includes the H-Animator modeling tool [Buttussi et al. 2006] for the authoring of signs and cheremes, while other dedicated tools simplify localization and the definition of associations between words and signs.

The paper is organized as follows. Section 2 surveys related work on sign language animation and sign language dictionaries, especially focusing on 3D and online proposals. Section 3 describes the client-server architecture of 3DictSL, its user interface, and some implementation issues. Section 4 analyzes in detail a case study, i.e. the implementation of Italian Sign Language support. Section 5 provides conclusions and outlines future research directions.

2 Related work

Research on sign languages has been particularly active over the last years. This section discusses related work on sign language animation and sign language dictionaries, particularly considering 3D and online approaches.

2.1 Sign language animation

Computer graphics research on sign language began in the early '80s. A first architecture to visualize schematic representations of signs was proposed by Loomis et al. [1983]. Results were limited by the hardware of that time, but the idea of using a skeleton to constrain movements is still used nowadays for sign language as well as generic humanoid animation.

About ten years later, Wilcox et al. [1994] proposed the first multimedia sign language dictionary for the American Sign Language (ASL), pointing out that animations are fundamental to correctly understand signs, since movement is one of the sign parameters. Besides searching signs by typing the associated American English word, this dictionary allows one to look for a sign by describing it. However, the proposal relies on videos for sign language animations and it is thus affected by the limitations described in Section 1.

A preliminary 3D arm for sign language animation was proposed by Gibet [1994]. The arm has only two joints, which are not enough to correctly discriminate among signs, but was useful to test the potential of 3D animation as a mean to visualize sign language sentences.

Geitz et al. [1996] focused on fingerspelling, i.e. showing a sequence of finger configurations to describe a word letter by letter. Since the system aimed at allowing deaf people who live in isolated areas to learn ASL finger configurations for letters, authors proposed a VRML animated hand that ran on a website.

[1] Recently, some linguists have proposed to use the term *phoneme* also for cheremes. However, other linguists prefer to keep the original term *chereme* which is etymologically and functionally more correct (from the Greek word *cheir* that means hand). In this paper, we use the term *chereme*, which also avoids ambiguity.

Sagawa and Takeuchi [2002] proposed a system to learn Japanese Sign Language. By using a dataglove, the student can try to perform the signs proposed by a 3D virtual teacher. The teacher then visualizes both the student sign and the correct sign, helping her to immediately notice errors. The system has been evaluated on users, proving the validity of the proposal.

To support Greek Sing Language (GSL) teaching in schools, Karpouzis et al. [2007] proposed a translation architecture which visualizes GSL signs and sentences by means of an H-Anim virtual character. Teachers can provide input in written Greek, which is transformed into GSL, animated using a VRML plug-in and the STEP scripting language [Huang et al. 2003], and then projected on a screen.

A machine translation system from American English to ASL has been proposed by Zhao et al. [2000]. To obtain realistic ASL sentence animations, they suggest to avoid using intermediate neutral postures between a sign and the next one in a sentence, since deaf people pause between signs only at the end of a sentence. Therefore, the system automatically builds transitions between signs by using Parallel Transition Networks [Badler et al. 1999].

ViSiCAST [Elliott et al. 2000] is a project that aims at providing improved access to services and facilities to deaf people through sign language presented by a virtual human. For example, the system translates English sentences commonly used in UK Post Offices into British Sign Language (BSL) signs, to simplify deaf customers' interaction with clerks. The virtual human can also perform European Sign Language animation of written English sentences, it is based on VRML and H-Anim and intended to be suitable for web as well as broadcast TV standards (e.g., the Multimedia Home Platform which is available on set-top boxes for terrestrial digital television).

Yi et al. [2005] proposed an interface which allows users to model sign language gestures such as hand, arm, or head movements, and then reuse these gestures to create signs. Gestures and signs are stored in a database. The stated final goal for the tool is to allow deaf people to "write" in their own language by composing sentences using the stored signs.

Sign Smith [VCOM3D 2004] is a suite of sign language products developed by VCOM3D. The suite consists of a 3D illustrated dictionary of ASL and Signed English (SE), an educational software for kids, four ASL sign libraries, and an authoring tool which allows users to create ASL content even if they had no previous programming and animation experience.

Sign language animations can also be modeled with the general H-Animator tool we described in [Buttussi et al. 2006], freely downloadable from [HCI Lab - University of Udine 2006]. H-Animator is a visual tool which allows the user to model, reuse and share any X3D animation for H-Anim humanoids. In particular, users can create standard H-Anim humanoids starting from existing X3D meshes, visually model simple X3D humanoid animations, without knowing the X3D language, and compose complex animations by visually concatenating simple ones.

2.2 Sign language dictionaries

As mentioned in the previous sections, paper dictionaries are not well-suited for sign languages, since it is very difficult to represent movement, which is one of the fundamental parameters of a sign, with 2D static figures. Anyway, paper dictionaries such as the one proposed by Stokoe [1976] for ASL and the one proposed by Radutzky [1992] for LIS are still used by deaf people and by researchers working on sign languages. Stokoe himself was aware of the limitations of paper dictionaries, and thus worked with Wilcox et al. [1994] to design the first ASL multimedia dictionary we described in Section 2.1.

Since the early '90s, several researchers and deaf communities have published multimedia dictionaries for the sign language of a specific country on CDs/DVDs or online, but most of these dictionaries (e.g., [Michigan State University Communication Technology Laboratory 2000; BritishSignLanguage.com 2000]) rely on videos and allow users to search for signs only by typing or selecting spoken language words.

A basic search functionality is available in [Waterfall Rainbows and Solarstorm Design 2006], an online dictionary which supports three different languages (ASL, BSL, and Australian Sign Language) and grows thanks to the cooperation of users, who can send videos of new signs. After choosing a language, the user can select a picture depicting an handshape chereme and then select another picture or a word associated to a sign with the selected handshape chereme. However, some handshapes cheremes belong to a large amount of signs, and thus, even using a database with less than a thousand signs, after selecting the handshape chereme the user may have to choose among more than 250 alternatives.

Since a sign can be described by specifying the cheremes for the four parameters (i.e. handshape, orientation, location, and movement), some recent dictionaries (e.g., [Herreweghe et al. 2004; Montagnana et al. 2005]) allow the user to specify the cheremes through a web interface and then find the word corresponding to the sign. Unfortunately, current solutions use static images to represent handshapes, and symbols or simple animated images to represent movements. Since some configurations and movements are very similar, even the deaf author of this paper, who is a sign language teacher, found the mentioned interfaces very difficult to use and error prone.

Other proposed solutions to input signs are transcription systems (e.g. Stokoe notation [Stokoe et al. 1976], HamNoSys [Prillwitz et al. 1989], SignWriting [Sutton 1981]) and visual recognition (e.g., by using hidden Markov models [Starner and Pentland 1995]). However, transcription systems require their users to learn several symbols (e.g., about 200 for HamNoSys), while visual recognition is natural for deaf users, but it has been effectively employed only to recognize a sign from a small set of alternatives and it is thus not well-suited to retrieve a sign from a dictionary.

To the best of our knowledge, VCOM3D [2004] illustrated dictionary, mentioned in Section 2.1, is the only one which exploits 3D technologies to better display signs. However, this dictionary provides only some ASL and SE signs, and does not offer a sign-to-word search functionality.

3 Our proposal: 3DictSL

To overcome the limitations of current solutions and meet the needs of deaf people, 3DictSL exploits standard Web3D technologies such as X3D and H-Anim to model and visualize 3D animations for a better understanding of signs. Moreover, H-Anim humanoid animations are used also to show the cheremes to simplify sign-to-word and sign-to-sign searches, because a guided exploration of the 3D model can avoid occlusions and ambiguous 2D projections. The following subsections will introduce the 3DictSL architecture, user interface, and some implementation issues.

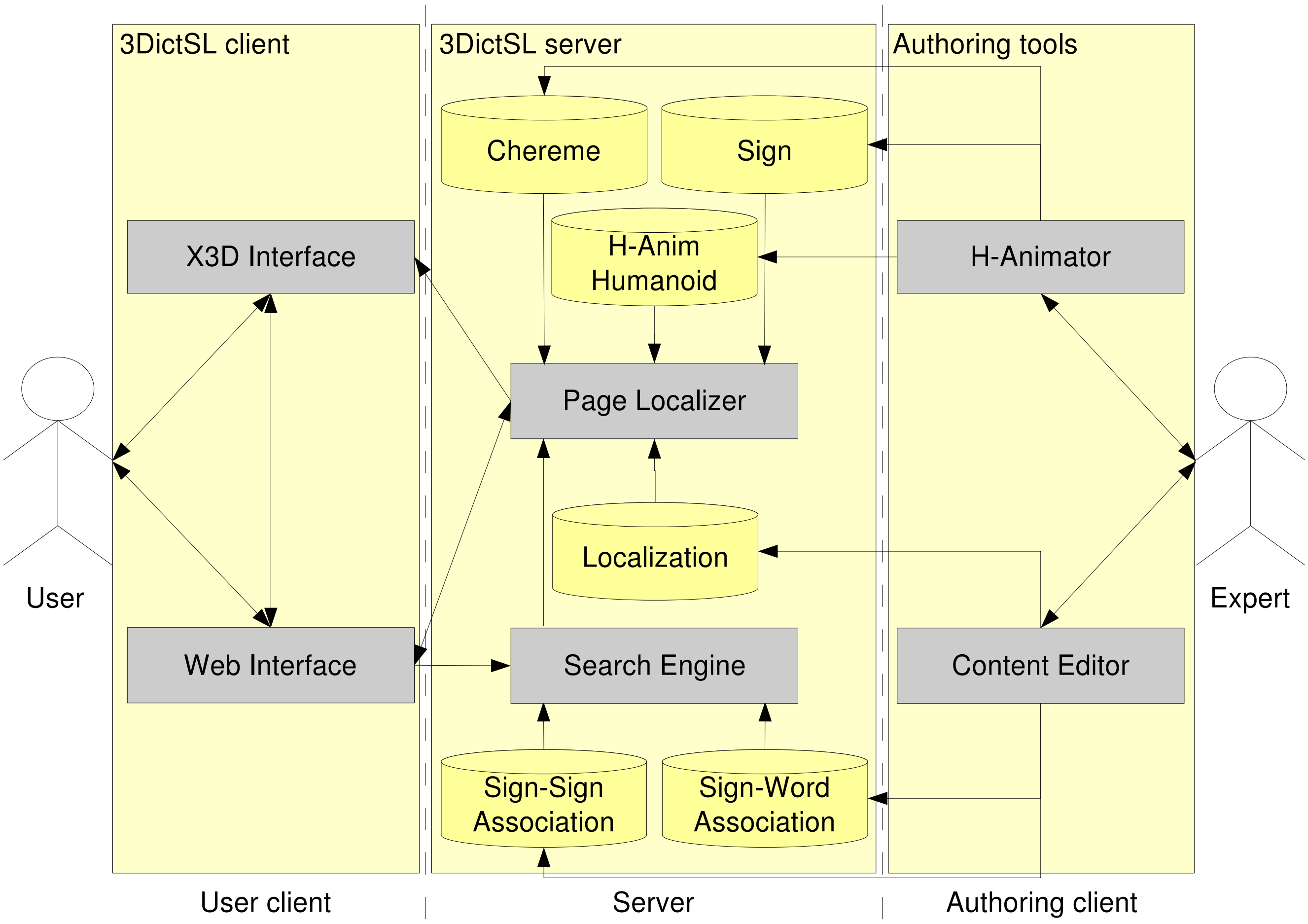

Figure 2: Client-server architecture of 3DictSL.

3.1 3DictSL architecture

The client-server architecture of 3DictSL is illustrated in Figure 2. The 6 modules are distributed on a server and on authoring and user clients. On the server side, the *Page Localizer* dynamically generates webpages and 3D content required by the user, localizing them for the language she chooses, while the *Search Engine* maps choices made by the user on the *Web Interface* into database queries and provides the Page Localizer with the results.

Internationalization and localization in the sign language domain are new and interesting topics, since adding a sign language locale means providing the textual content (e.g., webpage titles, help, button captions) as well as cheremes and signs for the specific sign language. Moreover, the textual content for a particular sign language can be provided in the language spoken in the same area, but deaf people might be only partly familiar with spoken language and thus choosing short, simple and meaningful words and sentences is more important than in other domains.

Page Localizer retrieves textual content for the chosen language from the *Localization* database, and X3D animations and H-Anim humanoids from *Chereme*, *Sign*, and *H-Anim Humanoid* databases. The Search Engine can find words and signs in the *Sign-Word Association* and in the *Sign-Sign Association* databases.

On the 3DictSL client, the *Web Interface* module handles user choices on the webpage by locally updating the page, by requiring new pages to the Page Localizer, or by updating the X3D Scene. Events in the X3D scene are handled by the *X3D Interface*.

To extend the dictionary with the support for new languages, sign language experts can provide textual content for the localization and populate the databases with sign-word and sign-sign associations by using the *Content Editor*. Moreover, since introducing a new sign language requires to model animations for cheremes and signs, users can exploit the functionalities of the *H-Animator* tool [Buttussi et al. 2006] to that purpose.

The overall usage of the website is illustrated with a flow chart in Figure 3. The user can select a language from the homepage, which provides welcome and getting started information in all the supported languages, and her choice is sent to the Page Localizer. If the chosen language is a spoken one, information on the following webpages is provided in that language, and the Page Localizer will provide a typical search form, where the user can type a word and have it searched in the Sign-Word Association database by the Search Engine. If the word is not in the database, the Page Localizer will notify the user about it and then redirect her to the search page, otherwise it will provide the user with a webpage which displays the result with an X3D H-Anim animation. The user can rotate the humanoid, get closer or farther, and replay the animation by means of controls on the webpage (free navigation in the 3D scene is disabled, since users without previous experience on 3D navigation may get lost). Actions on these controls are mapped into rotations, translations and timer changes by client-side scripts

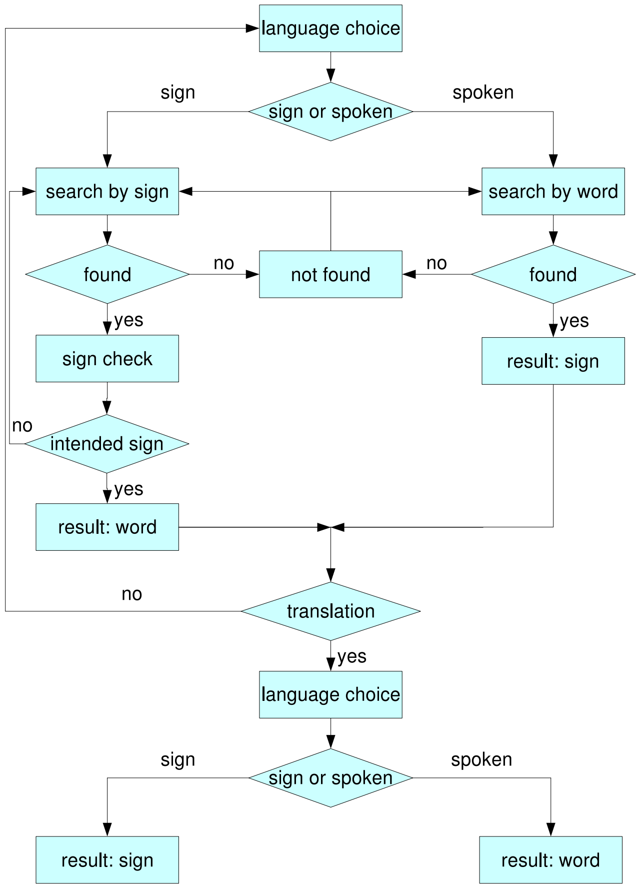

Figure 3: Overall functioning of the 3DictSL website.

through the X3D Scene Access Interface (SAI). The user can also have the word translated into other spoken or sign languages.

If the user selects a sign language from the homepage, then the Page Localizer will build webpages in the spoken language associated to the chosen sign language in the Localization database. The first webpage it builds allows the user to specify a sign by choosing a chereme for each parameter. When the user has chosen all the four cheremes, she is provided with a new webpage where the sign characterized by those cheremes is visualized, if it is found in the Sign database. The user can interact with the 3D scene and decide whether the sign is the one she intended or not. Indeed, some cheremes are very similar, thus if the sign is not the intended one, the user can return to the search page and try another chereme combination. When the user confirms the sign, the associated word is returned and the user, as mentioned above, can look for a translation into other spoken or sign languages.

3.2 User interface

The user interface of most 3DictSL webpages is very simple, therefore we will focus on the *search-by-chereme* webpage, i.e. the webpage where the user can look for the translation of a sign, by choosing its four cheremes. The search-by-chereme webpage is particularly interesting, since it provides different views (Figure 4, 5, 6, and 7) to choose the chereme for the different parameters.

As shown in the figures, the search-by-chereme webpage is vertically split into two areas: the left one contains the controls to choose the cheremes, while the right one contains an embedded X3D player displaying a H-Anim humanoid, and some controls to interact with the 3D scene. To simplify the choice of the cheremes, as the user selects one of them, the H-Anim humanoid posture or movement are updated to preview the resulting sign. The buttons below the embedded X3D player allow the user to rotate the humanoid and get closer or farther, so that she can check whether the cheremes she has chosen are the intended ones or not, while the buttons in the lower left area allow to switch among the views. Some views (Figure 4 and 7) provide also additional buttons to switch among different categories of cheremes concerning the same parameter (e.g., handshape cheremes can be grouped by considering the number of open fingers).

In the *handshape* view (Figure 4), the user is provided with the controls to select the chereme for the handshape. Other dictionaries which allow for a search by handshape [Montagnana et al. 2005; Herreweghe et al. 2004; Waterfall Rainbows and Solarstorm Design 2006] use 2D drawings or schematic representations of an hand, while we decided to take thumbnails of the humanoid hand. Moreover, since there are usually several cheremes for the handshape, some of the current solutions [Montagnana et al. 2005; Herreweghe et al. 2004] split the handshape cheremes into categories by considering the number of open fingers. However, the deaf author of this paper pointed out that it is very difficult to associate each chereme with a particular number of open fingers, since some cheremes require to open or close some fingers during the performance of the sign. Therefore, we proposed a deterministic criteria to split LIS handshape cheremes (see Section 4) and the Content Editor allows other deaf communities to freely choose the handshape categories for their language. Another issue with handshape cheremes is that some of them differ only for the order in which the same or similar finger configurations occur: for example, a chereme may start with all the fingers closed and finish with them open, while another may start with all the fingers open and finish with them closed. Therefore, with current video-based solutions the user may select an handshape chereme and find out she has selected a wrong one only when she plays the video of the resulting sign, while in 3DictSL the H-Anim humanoid hand is immediately animated and the user can test different handshape cheremes before proceeding with the other parameters.

The choice of the orientation (Figure 5) is usually simpler than the handshape one, since there are usually few cheremes for this parameter: the palm is facing the left or facing the right of the person who is performing the sign, upturned or downturned, facing the person who performs the sign or facing the person who looks at it. To simplify the choice, we use thumbnails of the right hand, the right arm and the upper torso of the humanoid. In the thumbnails, we use an handshape which simplifies orientation understanding, and as the user selects one of these thumbnails, the humanoid shows the selected orientation with the handshape chereme chosen in the previous view.

To choose the location where the sign is performed, current solutions which allow for a search by location [Montagnana et al. 2005; Herreweghe et al. 2004] use a 2D picture of a human, and the user can click on the different parts of the human body or the nearby space to select the location. Since 3DictSL uses a 3D humanoid, we allow the user to directly click on it or select a thumbnail of the location (Figure 6). For both actions, the interface provides the user with immediate feedback by moving the humanoid hand in the selected location.

The last parameter, i.e. movement, is the most critical for solutions relying only on images and thus most of the proposals do not support it, or have recently added some simple animated images or

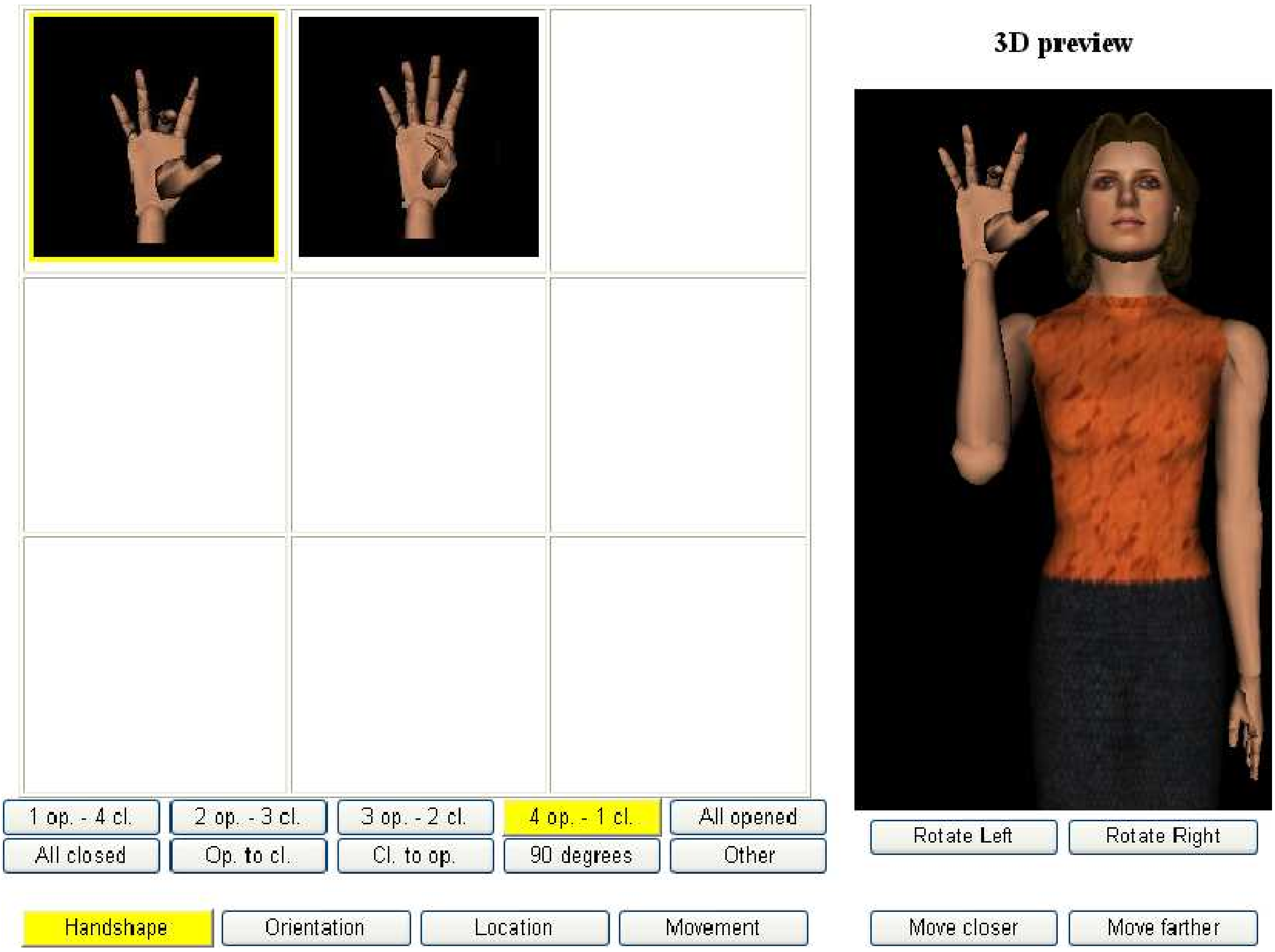

Figure 4: Handshape choice in the search-by-chereme webpage. Although the textual content in the webpages was originally in Italian, for this paper it has been translated into English to ease comprehension.

schematic representations of the movement. Since displaying several animated images in the same webpage may confuse the user, we decided to use schematic representations (e.g., straight or curved arrows), while the movement in 3D is shown by the humanoid, as the user selects one of the schematic representations (Figure 7).

3.3 Implementation issues

Since the user may need to rapidly test different cheremes until she finds the intended ones, we looked for approaches which allow for a fast response. Therefore, we decided to implement 3DictSL webpages in ASP.NET AJAX, which allows us to update only some portions of the webpage by using the *UpdatePanel* controls. In particular, to implement the search-by-chereme webpage, we split the page into two panels: one for the controls to select the cheremes, and one for the X3D Player and the controls to interact with the 3D scene. To allow the user to switch among views without refreshing the page, we used the *MultiView* and the *View* controls provided by ASP.NET 2.0. The Page Localizer module builds the webpages using ASP.NET, while the behavior of the controls that need information from the server is implemented in C#. For example, C# code is used to change the active *View* of the *MultiView* control, and to retrieve localization information, X3D H-Anim animations, and associations between words and signs from the databases.

Databases were implemented with Microsoft Access to reuse humanoids, animations and additional information from H-Animator databases. In particular, H-Anim Humanoid database consists in a subset of the tables of an H-Animator database, Sign database is an H-Animator database with additional metadata (e.g., information on the cheremes which identify the sign), and Chereme database contains information from the H-Animator database (e.g., the handshape cheremes are H-Animator simple animations), but also other information. The additional information is acquired with ad hoc functionalities that we added to the general purpose H-Animator tool. For example, location cheremes, which consist in *rotation* values for the right shoulder and the right elbow of the humanoid, and orientation cheremes, which consist in *rotation* values for the right wrist, are specified by modeling a single posture. Movements can be modeled as simple animations, but without specifying an handshape chereme and assuming to perform the sign in the space nearby. Therefore, they cannot be saved as H-Animator simple animations (i.e. a set of *PositionInterpolator* and *OrientationInterpolator* nodes with the routings to *HAnimJoint* nodes), but as *Script* nodes which change shoulder and elbow rotation after considering their current values (due to the location).

Client-side code is written in Javascript and was used, in particular, to access the 3D scene. In most of the webpages, the interaction between the webpage and the 3D scene is limited to changes of the *rotation* and the *translation* values of the whole *H-AnimHumanoid*

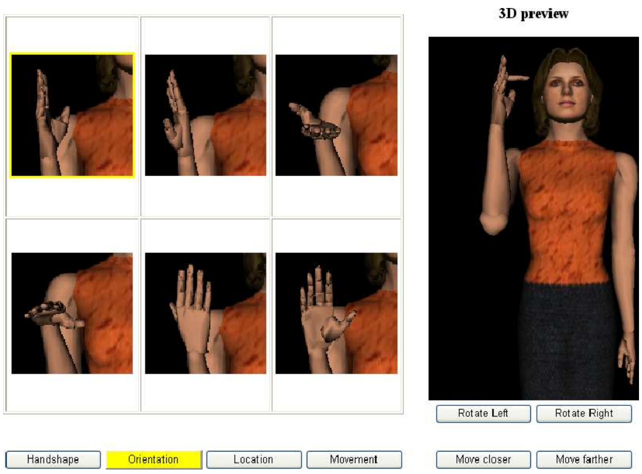

Figure 5: Orientation choice in the search-by-chereme webpage.

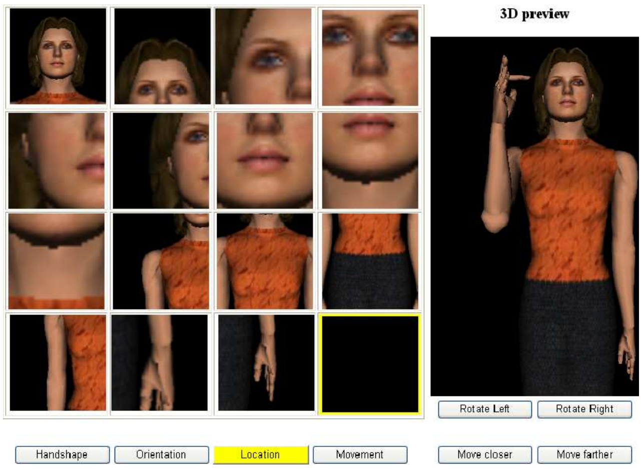

Figure 6: Location choice in the search-by-chereme webpage.

node, but the search-by-chereme webpage requires more interesting interaction, implemented with both Javascript and X3D code. For example, to select the location the user can both choose a thumbnail and click on the humanoid. Therefore, we added a field event listener for each of the *TouchSensor* nodes on the humanoid body to update the selected picture in the webpage, as the user touches a sensor. User interaction with both the thumbnails and the humanoid body causes also *rotation* changes in the right shoulder and right elbow *HAnimJoint* nodes. *Rotation* values are provided by the Page Localizer, which automatically generates Javascript code with the values when it builds the webpage considering the particular sign language. Since the values are written in Javascript, they can be immediately set as the user performs an action, without the need for a server request.

The same implementation strategy has been used for orientation cheremes, but movements and handshape cheremes required a different strategy, since there are usually a lot of these cheremes and it is thus very inefficient to download the X3D code for all of them when the search-by-chereme webpage is loaded. Therefore, when some X3D code is needed, it is written in a hidden label without refreshing the whole page by using ASP.NET AJAX C# code. When the X3D code is written in the label, it is added to the 3D scene with the SAI *createX3DFromString* function. The general idea of combining AJAX and X3D into the so-called Ajax3D has been proposed by Parisi [2006].

4 Italian Sign Language case study

In this section, we will focus on Italian Sign Language (LIS) to provide an example of how deaf communities can author the content needed to extend 3DictSL with the support for their own sign language. LIS has 56 handshape cheremes, 6 orientation cheremes, 16 location cheremes, and 48 movement cheremes, besides other cheremes which are used more rarely.

Since there are several handshape and movement cheremes, we split them into categories. As mentioned in Section 3.2, some of the current solutions (e.g., [Montagnana et al. 2005]) split handshape cheremes by considering the number of open fingers, but this criteria is not well-suited for LIS, since some handshape cheremes require to open or close fingers. Therefore, we propose a deterministic criteria which allows to split handshape cheremes into ten categories: 6 categories contain cheremes which do not move fingers and have a fixed number of fingers fully open or fully closed (i.e. all closed, 1 open and 4 closed, 2 open and 3 closed, and so on), a category contains cheremes which require to open fingers, another category contains cheremes which require to close fingers, a category contains cheremes which have fingers not fully open nor closed, and the last category contains all the other cheremes (i.e. cheremes which require a particular movement of the fingers). We also split movement cheremes into 4 categories: straight movements of the hand toward a particular direction, movements which require to draw arcs with the hand, movements which require to bend the wrist, and all the other movements.

LIS cheremes can be combined in more than 50000 LIS signs, but only a subset of them (about 3000) is frequently used and understood by almost all Italian deaf people who use the language. However, it is clearly a huge work even to model the basic signs for a single language, thus the need for easy-to-use authoring tools. The current version of H-Animator was intended for general purpose animation modeling and it is thus not sufficient to significantly speed up the modeling of signs, but is quite good to model handshape cheremes: an untrained computer science student modeled all the most used LIS handshape cheremes spending about 10 minutes for each. While handshape cheremes and some test signs were modeled with the released H-Animator tool, to model other LIS cheremes we developed a preliminary new version of the tool specifically focused on chereme modeling (e.g., it allows to store a single posture as a location chereme).

Since accuracy is critical to discriminate among cheremes, the deaf author of this paper (who is also a LIS teacher) was filmed from several viewpoints while performing the cheremes. Moreover, for each chereme, we filmed also a test sign which employs it. Figure 8 shows a frame of one of these videos and the H-Anim humanoid performing the same handshape chereme. We checked the accuracy of the modeled cheremes on some deaf users by asking them to look at the humanoid performing a chereme and then choose the word associated to the test sign for that chereme.

Besides modeling cheremes and signs, deaf communities can use the Content Editor to enter textual localization content as well as sign-word and sign-sign associations. We have already entered localization information in Italian and American English, and we have created associations for a few test signs.

Figure 7: Movement choice in the search-by-chereme webpage.

5 Conclusions and future work

In this paper, we described how standard Web3D technologies such as X3D and H-Anim can be exploited to visualize signs and to simplify sign-to-word and sign-to-sign search in an online international sign language dictionary (3DictSL). To achieve these goals, we proposed a client-server architecture, which includes authoring tools to extend the dictionary with the support for new sign languages. The paper provided details on the user interface of the website and on how AJAX, ASP.NET, Javascript, and X3D technologies were combined to develop our proposal.

With respect to the goals stated in Section 1, we provided the architecture for a multilanguage dictionary with 3D animations and advanced search functionalities, and authoring tools to extend the dictionary with the support for new languages. However, the dictionary can be visualized only by recent browsers with an X3D player supporting H-Anim nodes and Javascript SAI. Moreover, current tools are not sufficient to simplify and significantly speed up the authoring task, since sign languages have several thousands of signs.

Motion capture and computer vision may seem promising techniques to semi-automatically populate sign database, but, unfortunately, since sign languages require very accurate animations to discriminate among similar signs with different meanings, expensive equipment is needed. As a result, these technique may be employed in a laboratory to populate the database with an initial set of signs, but deaf communities will need cheaper tools to define further signs at home. Therefore, we are working on a new version of H-Animator specifically designed for sign languages. This version should run on a website, directly store data on 3DictSL databases, integrate the functionalities of the Content Editor, and provide modeling functionalities which exploit the structure of the signs, i.e. it should allow the user to model a sign by reusing existing cheremes. This new version will potentially allow all the members of a deaf community to augment the dictionary with new signs. By adding these new features, we will also have to deal with security issues (e.g., preventing guests from editing content provided by experts) and design a functionality which allows users to rate the quality of proposed signs.

3DictSL has been constantly evaluated by the deaf author of the paper. A formal user evaluation of the system on several deaf users is planned. Other interesting future research may concern facial expression, since it is increasingly used in sign languages and has been recently introduced as a fifth parameter.

6 Acknowledgements

Tamara Trippardella modeled the Italian Sign Language cheremes for the handshape with the H-Animator tool.

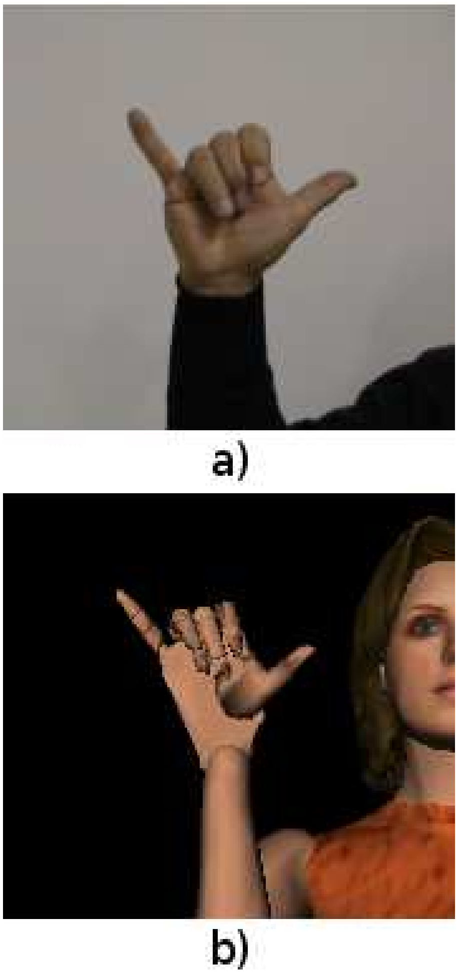

Figure 8: A deaf sign language teacher performing a LIS handshape chereme (a) and the same chereme performed by the humanoid (b).

References

BADLER, N. I., PALMER, M. S., AND BINDIGANAVALE, R. 1999. Animation control for real-time virtual humans. *Commun. ACM 42*, 8, 64–73.

BRITISHSIGNLANGUAGE.COM, 2000. BritishSignLanguage.com website. http://web.ukonline.co.uk/p.mortlock/.

BUTTUSSI, F., CHITTARO, L., AND NADALUTTI, D. 2006. H-animator: a visual tool for modeling, reuse and sharing of X3D humanoid animations. In *Web3D '06: Proceedings of the eleventh international conference on 3D web technology*, ACM Press, New York, NY, USA, 109–117.

ELLIOTT, R., GLAUERT, J. R. W., KENNAWAY, J. R., AND MARSHALL, I. 2000. The development of language processing support for the ViSiCAST project. In *Assets '00: Proceedings of the fourth international ACM conference on Assistive technologies*, ACM Press, New York, NY, USA, 101–108.

GEITZ, S., HANSON, T., AND MAHER, S. 1996. Computer generated 3-dimensional models of manual alphabet handshapes for the World Wide Web. In *Assets '96: Proceedings of the second annual ACM conference on Assistive technologies*, ACM Press, New York, NY, USA, 27–31.

GIBET, S. 1994. Synthesis of sign language gestures. In *CHI '94: Conference companion on Human factors in computing systems*, ACM Press, New York, NY, USA, 311–312.

HCI LAB - UNIVERSITY OF UDINE, 2006. H-Animator website. http://hcilab.uniud.it/h-animator.

HERREWEGHE, M. V., VERMEERBERGEN, M., WEERDT, K. D., VANHECKE, E., MULDERS, K. V., AERTS, S., BRAEM, B., PAREDAENS, J., MICHIELS, P., DEXTERS, N., AND ADRIAENSSENS, J., 2004. Flemish Sign Language dictionary. http://gebaren.ugent.be.

HUANG, Z., ELIËNS, A., AND VISSER, C. 2003. Implementation of a scripting language for VRML/X3D-based embodied agents. In *Web3D '03: Proceeding of the eighth international conference on 3D Web technology*, ACM Press, New York, NY, USA, 91–100.

KARPOUZIS, K., CARIDAKIS, G., FOTINEA, S.-E., AND EFTHIMIOU, E. 2007. Educational Resources and Implementation of a Greek Sign Language Synthesis Architecture. *Computers & Education 49*, 1, 54–74.

LOOMIS, J., POIZNER, H., BELLUGI, U., BLAKEMORE, A., AND HOLLERBACH, J. 1983. Computer graphic modeling of american sign language. In *SIGGRAPH '83: Proceedings of the 10th annual conference on Computer graphics and interactive techniques*, ACM Press, New York, NY, USA, 105–114.

MICHIGAN STATE UNIVERSITY COMMUNICATION TECHNOLOGY LABORATORY, 2000. American Sign Language Browser. http://commtechlab.msu.edu/sites/aslweb/browser.htm.

MONTAGNANA, S., STANIZZI, I., VETTORI, C., ZANONI, C., BAJ, C., BATTAGIN, A., CONSOLATI, M., AND VALENTE, M., 2005. e-LIS. http://elis.eurac.edu.

PARISI, T., 2006. Ajax3D White Paper. http://www.ajax3d.org/whitepaper.

PRILLWITZ, S., LEVEN, R., ZIENERT, H., HANKE, T., AND HENNING, J. 1989. *HamNoSys. Version 2.0; Hamburg Notation System for Sign Languages. An introductory guide*. Signum, Hamburg, Germany.

RADUTZKY, E., Ed. 1992. *Dizionario bilingue elementare della Lingua Italiana dei Segni*. Edizioni Kappa, Rome, Italy.

SAGAWA, H., AND TAKEUCHI, M. 2002. A teaching system of Japanese sign language using sign language recognition and generation. In *MULTIMEDIA '02: Proceedings of the tenth ACM international conference on Multimedia*, ACM Press, New York, NY, USA, 137–145.

STARNER, T., AND PENTLAND, A. 1995. Real-time American Sign Language recognition from video using hidden Markov models. In *ISCV '95: Proceedings of the International Symposium on Computer Vision*, IEEE Computer Society Press, Los Alamitos, CA, USA, 265–270.

STOKOE, W., CASTERLINE, D., AND CRONEBERG, C. 1976. *A dictionary of American sign language on linguistic principles*. Linstok Press, Silver Spring, MD, USA.

SUTTON, V. 1981. *Sign writing for everyday use*. Sutton Movement Writing Pr., Boston, MA, USA.

VCOM3D, 2004. Sign smith products. http://www.vcom3d.com.

WATERFALL RAINBOWS, AND SOLARSTORM DESIGN, 2006. Dictionaryofsign.com. http://www.dictionaryofsign.com.

WILCOX, S., SCHEIBMAN, J., WOOD, D., COKELY, D., AND STOKOE, W. C. 1994. Multimedia dictionary of American Sign Language. In *Assets '94: Proceedings of the first annual ACM conference on Assistive technologies*, ACM Press, New York, NY, USA, 9–16.

YI, B., FREDERICK C. HARRIS, J., AND DASCALU, S. M. 2005. From creating virtual gestures to "writing" in sign languages. In *CHI '05: Extended abstracts on Human factors in computing systems*, ACM Press, New York, NY, USA, 1885–1888.

ZHAO, L., KIPPER, K., SCHULER, W., VOGLER, C., BADLER, N. I., AND PALMER, M. 2000. A Machine Translation System from English to American Sign Language. In *AMTA '00: Proceedings of the 4th Conference of the Association for Machine Translation in the Americas on Envisioning Machine Translation in the Information Future*, Springer-Verlag, London, UK, 54–67.