ICTA'07, April 12-14, Hammamet, Tunisia

Towards Web-Based automatic interpretation of written text to Sign Language

Mohamed JEMNI Oussama ELGHOUL
Research Unit of Technologies of Information and Communication UTIC
Ecole Supérieure des Sciences et Techniques de Tunis,
5, Av. Taha Hussein, B.P. 56, Bab Mnara 1008, Tunis, TUNISIA
Mohamed.jemni@fst.rnu.tn oussama.elghoul@gmail.com

Abstract

The extraordinary progresses of new technologies related to data processing and to Internet, offer real opportunities to a better life quality and in particular, to those that endure handicaps. In this context, our project aims to improve accessibility for deaf persons by using the technologies of information and communication. The objective was to design and to develop a tool facilitating communication through the Web for deaf. This tool allows automatic interpretation of texts in visual-gestural-spatial language by using avatar technology. The architecture is constructed around a virtual animated person who automatically interprets a text within a sign language. This system uses a dynamic data base for sign dictionary that contains gesture codes for words. This tool is presented as a multilingual Web application that offers a full set of interfaces for real-time interpretation, collaborative alimentation of the dictionary and creation of deaf communities in addition to administration interfaces.

1. Introduction

Disabled persons face non-ending difficulties when they want to deal with the new technologies: the use of a computer, the access to Internet, the edition and the impression of a text, etc…. Reading a document can be extremely complex tasks in spite of their simplicity for normal user.

Today, the extraordinary progress of the new technologies, bound to the data processing and Internet, offers remarkable opportunities to bring a better quality of life to those who endure handicap and disabilities.

In this context, we target a specific category of disabled person: the deaf.

The objective of our project is to develop a Web-based interpreter of Sign Language (SL). This tool would enable people who do not know SL to communicate with deaf individuals. Therefore, contribute in reducing the language barrier between deaf and hearing people.

Our secondary objective is to distribute this tool on a non-profit basis to educators, students, users, and researchers, and to disseminate a call for contribution to support this project mainly in its exploitation step and to encourage its wide use by different communities.

The automatic translation from the written language to the sign language is a promising way that requires the mastering of several domains covering the field of the data processing, of linguistics, mathematics, etc... A sign language is a language which uses manual communication instead of sound to convey meaning - simultaneously combining hand shapes, orientation and movement of the hands, arms or body, and facial expressions to fluidly express a speaker's thoughts. The sign language remains nevertheless a fully-fledged language, with its own constructional method of the sentences.

Research on automatic interpretation of written text in sign languages and generation of gestural component of spoken languages has been held back by the unavailability of open source computational tools that can be widely used by users. The challenge is to develop computational application that can be deployed via the Web and internet and that gathers two main properties: efficiency and ease of use. Efficiency reports mainly to quick response time and low bandwidth requirements. However the ease of use reports to fullness and conviviality of user interfaces. Here we describe WebSign, a Web based application that we have designed to facilitate the creation and the

use of an-online dictionary in order to interpret at real time texts in sign language.

This paper is organized as follow: the next section is devoted to present the state of the art related to three main alternatives for sign language interpretation. In section 3, we describe the different functionalities offered by our tool. In section 4, we present the general approach we have adopted to develop our tool. Finally, we draw a conclusion and some perspectives.

2. State of the art

Only a minority of the deficient auditory people masters the written language. Meanwhile, most communication systems for this community exploit textual information, badly adapted to their real needs. Thus, to develop a communication system essentially founded on the sign language becomes a real socioeconomic stake.

This section exposes the state of the art of the computer and technological tools that can be considered as the most representative systems of the communication in sign language through the whole world, in different environments. Three main approaches exist in the literature: the first is based on writing or drawing symbols, the second approach is based on video; however the third is based on 3D and animation of a virtual person according to a standard. Another approach is constituted by systems that are oriented dynamic Web.

2.1. Writing based systems

The first transcription of signs languages is drawing, it represents the means generally used to replace the absence of writing. Several systems of transcription appeared like HamNoSys and SignWriting in spite of the difficulty to encode sign language in a linear way.

The first version of HamNoSys (Hamburg Notation System) was defined in 1984. It was developed as a linear phonetic transcription tool for sign languages. This transcription should be possible for virtually all sign languages in the world, but it is not easily usable by the people for the catch of notes or the reading.

SignWriting is developed by Valerie Sutton for the Center of Sutton Movement Writing, in 1974 [15]. It is based on graphical, bidimensional representations, using graphical symbols. This system is easy to encode in computers in a linear way, by assigning numeric codes to each special character.

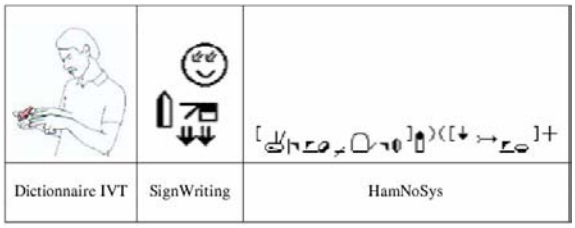

Figure 1. Transcriptions of word [read] in French sign language (LSF) [16].

2.2. Video based systems

The video based systems consist in encrusting a video sequence corresponding to a human interpreter in the initial information. In this case, the performances are bound closely to the diagram of coding put in work for the video in sign language. Today, we can offer ourselves many normalized tools, outclassing of very far our predecessors in term of debit/quality. Some current projects based on video are: “The Personal Communicator” [1], “LSF Lexique” [2].

To remedy the main problem of these oriented video systems, a first technique consists in coding the plots video in their totality, whereas a second solution consists in coding a selective objects video of arbitrary shapes. Essential objects are the hands, face, bust and bottom of the picture. Teherfore, an automatic procedure of segmentation has to be used. The quality, the integrity and the intelligibility of the sign language content, versus the compressive debit, depend on the following parameters: number of plots in the diagram of coding not on quantification, considered cadence video and resolution of the video.

2.3. Avatar based systems

In this new technological context, the modeling of a virtual character can be achieved either according to a segmented model, gotten by an hierarchical graph of the anatomical segments 3D, or according to a model seamless. In both cases, the surfaces are represented either with the help of polygonal stitch, either by a mathematical analyzing, and either again according to implicit functions to skeleton. The creation of the virtual character is achieved either from a modeler of geometric primitives, or with the help of a scanner 3D. In new standards, the approach of segmented virtual character exists in the H|Anim specifications [6], as well as in those of MPEG-4 FBA (Face & Body Animation) [7], whereas the representation by virtual character is processed in MPEG-4 BBA (Bone-Based

Animation). Some current projects based on video are: Signeuse Virtuelle 3D, eSign and VSigns [3,4,5].

2.4. Conclusion

If the coding of sequences videos in sign language conducts currently to performances in the limit of the constraints of the low debit and the requisite quality, its limitations are more due to the fact that the contents are non easily reusable and therefore a human translation in language of signs is necessary and this remains very costly in real applications. These reasons leaded us to be more interested in communication systems in sign language by animation of virtual 3D persons, more precisely an avatar in VRML in conformity with the H|Anim. Furthermore, our system must be more efficient than the existing ones and has to innovate mainly in designing a solution that combines the Web, not consuming of transfer bandwidth, a friendly use solution to ensure the evolution of the system and its use in real time.

3. Description of the functionalities of our system

Our tool called WebSign is a Web application. It is based on the technology of avatar (animation in virtual world). The input of the system is a text in natural language. The output is a real-time and on-line interpretation in sign language. This interpretation is constructed thanks to a dictionary of word and signs. The creation of this dictionary can be made in an incremental way by users who propose signs corresponding to words. A word and its corresponding sign interpretation are added effectively to the dictionary only after its verification by an expert administering the system.

However, contrary to popular belief, sign language is not universal. Wherever communities of deaf people exist, sign languages develop, but as with spoken languages, these vary from region to region. Hundreds of sign languages are in use around the world and are at the core of local deaf cultures [9]. Some sign languages have obtained some forms of legal recognition, while others have no status at all.

For this reason, we implemented the notion of virtual communities. A community is a group of users that can build and share a common dictionary of sign language. A dictionary can be created totally by a specific community or can be just an instance of an existent dictionary where some specific words are interpreted differently to respect the intrinsic specification of the concerned community.

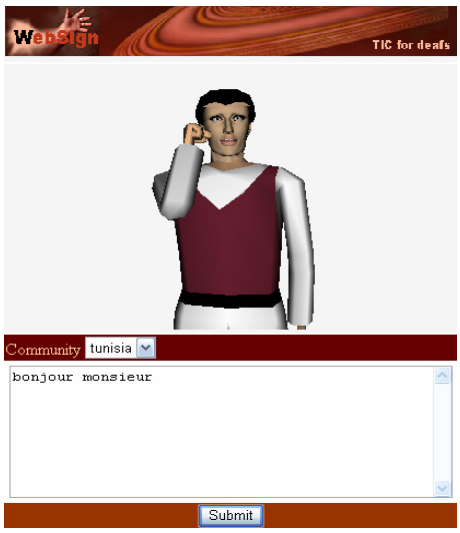

Figure 2. Interface of interpretation.

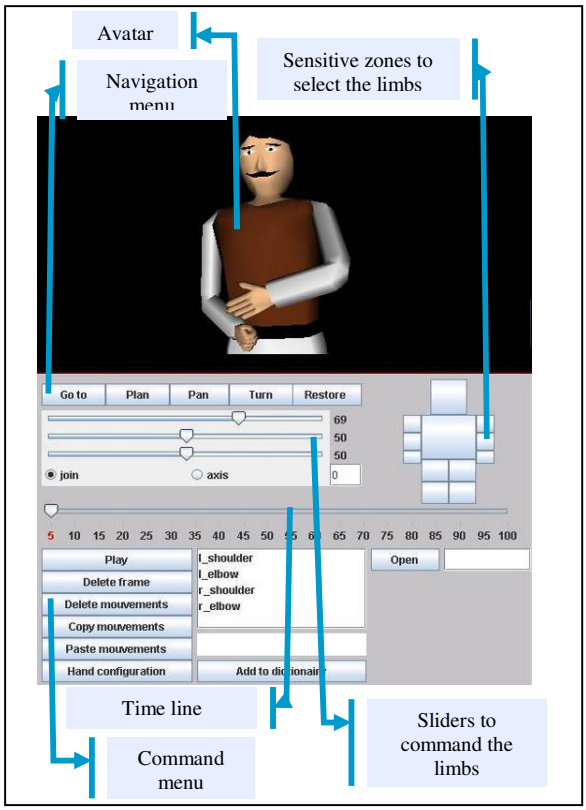

Figure 3. Interface for the construction of signs

There are three main actors in our WebSign application. The general administrator of the system plays a primordial role. He provides the basic information for the treatment of our tool. He also updates the system and consults it from the global administration interface. He communicates with all others actors. The local administrator of a community: his role is restricted to the administration of his community and the verification of the proposed signs before their integration in the corresponding dictionary. The third actor is the final user. He benefits from the main functionalities: the real-time interpretation of texts in sign language (in on-line way), the possibility to propose signs and words by the use of graphical dedicated interface, the use of a messenger tool to communicate with deaf persons by sending texts to be interpreted automatically in sign language. The interpretation can be in synchronous or asynchronous way.

Figure 4. Real time interpretation

Figure 5. Asynchronous interpretation

Websign is multilingual; In fact the administrator of community can personalize the interface language of his community. To create a community the default (the initial) language used is English. Using a specific interface the local administrator of a community can translate the interface to another language.

4. Technical approach

4.1. Modelling the avatar

To create our avatar, we have realized the following steps: the generation of a skin for the avatar, the carving of the skin in small surfaces, the creation of the skeleton and the creation of the texture.

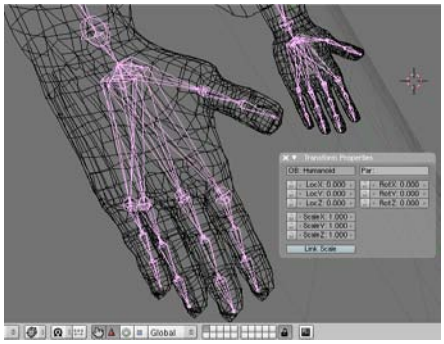

Figure 6. Hand modeling

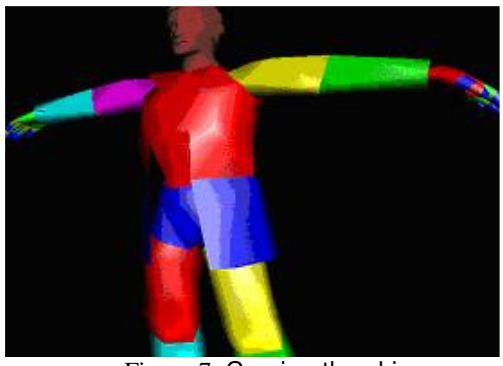

Figure 7. Carving the skin

4.2. Personalization of avatar

Now we are working on ameliorating the graphical user interface of Websign by allowing the user to customize his avatar. This can be done through the new module of websign. The object of this new module, is to the user to submit his photo to our web-server, in order to generate automatically a personalized avatar resembling to the submitted user's photo. Such that every user can use his personalize avatar when he communicates with our Websign based messenger tool. Basically, this new module estimates the suitable 3D rendering of the avatar based on the 2D user's photo.

4.3. Construction of the player module

We integrated a VRML player and a Java applet that command it in a Web page. We implemented the algorithm of animation using SML: Sign Modeling Language, the language that we constructed for the modeling of signs.

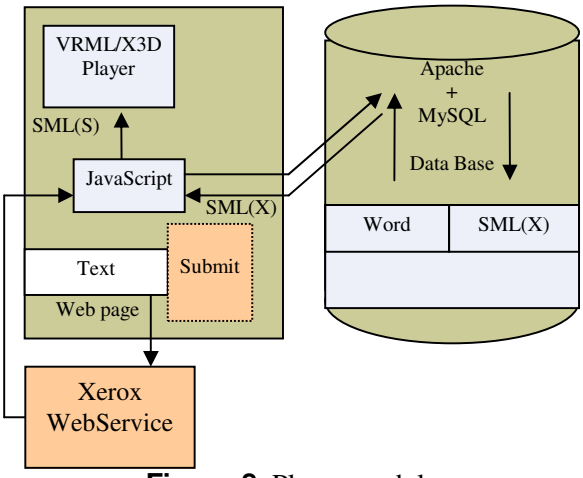

Figure 8. Player module

For linguistic processing, the system use Xerox WebService which used to tokenize and doing the morphological analysis.

4.4. Construction of the interpretation module

Our system is based on Client/Server architecture. To implement this architecture, we have realized the following steps: first, we create the clients, the server and the data base. Second, we saved a set of SML codes to initiate the data base and finally we have implemented the interpretation algorithm via the service of messaging.

4.5. SML

SML is an XML based descriptive language, which we have developed to codify gestures in Websign application. SML is used to store description of signs, which can be generated by a virtual human, or an avatar.

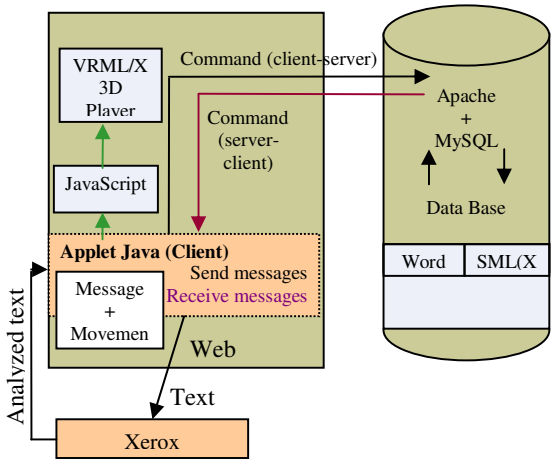

Figure 9. Interpretation module

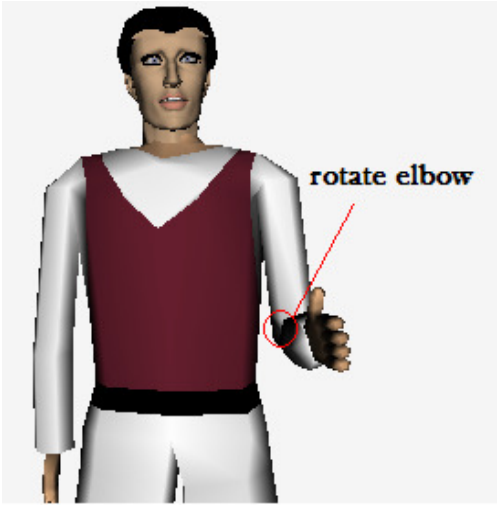

Figure 10. Rotation of l_elbow join using SML(X)

```
<Sentence>
   <word>
      <mouvement>
         <duration>0.5</duration>
         <join name="l_elbow">
            <rotation type="euler">
               <heading>80</heading>
               <attitude>50</attitude>
               <bank>50</bank>
            </rotation>
         </join>
      </mouvement>
   </word>
</Sentence>
```

SML considers that a sign is an animation of many joins. The animation is a successive movement of a group of joins. Every movement has a fixed time, during it, the rotation of every join of group is done.
The armature description respects the H|Anim specification, in which each join has a specific name and specific initial orientation.

In “Figure 9” the avatar rotates the left elbow join (l_elbow) to the rotation with heading=80, attitude=50 and bank=50 using Euler angle. The duration of movement is 0,5 seconds.

4.6. Tools and technologies

For the 3D animation on the Web, we used the Virtual Reality Modeling Language (VRML) that is a programming language specialized in the representation of virtual universes in 3D. This interpreted language is an ISO international norm. X3D, based on XML syntax, has been created by the Web3D consortium in the goal to follow VRML. It was normalized by the ISO in 2005. We used a specific editor X3D-Edit.

To ensure an instantaneous communication, we developed a messaging server in PHP (since we used the PHP technology and Apache as application server) and a client in Java (a Java applet).

5. Conclusion and perspectives

We presented in this paper a tool that aims to enhance communication for deaf, hard-of-hearing and speech disabled individuals. The originality of this tool, in addition to be an open source, consists on two points: first it combines the advantages of different computer techniques and recent technologies; second its collaborative and incremental approach to create dictionaries based on virtual communities. Moreover, thanks to coding gestures in SML Language (XML based language), the code is independent of the language (words or texts). Therefore, it is possible to share same signs and some parts of dictionaries among multilingual communities.

A first version of WebSign is finalized and completely functional. We plan to make it available very soon in the Web. We already made some contacts with specialized deaf associations to train users in dictionary alimentation. As perspective of WebSign, we plan to develop a module of movement capture as a new component of the system in order to generate automatically sign code.

6. References

[1] http://commtechlab.msu.edu/.
[2] http://www.lsf.univ-nancy2.fr/.
[3] http://www.cnrs.fr/.
[4] http://www.visicast.sys.uea.ac.uk/.
[5] http://vsigns.iti.gr/.
[6] http://www.h-anim.org/.
[7] http://www.mpeg.org/.
[8] http://www.esstt.rnu.tn/utic/.
[9] Neidle, C., S. Sclaroff, and V. Athitsos, “SignStream™: A Tool for Linguistic and Computer Vision Research on Visual-Gestural Language Data”, Behavior Research Methods, Instruments, and Computers, 33:3, 2001, pp. 311-320
[10] Guido Gybels, Deaf and Hard of Hearing Users and Web Accessibility, white paper of The Royal National Institute for Deaf People, October 2004.
[11] Mike Kemp, Fundamentals of Evaluating Sign Language Programs: Checklists for Program Assessment, ISBN: 1-58121-008-6, DawnSignPress.
[12] R. Parent, Animatique : Algorithmes et techniques, Paris : Vuibert, 2003.
[13] S. Kiss, Computer animation for articulated 3D characters, Technical Report TR-CTIT-02-45, Université de Twente, Hollande, 2002.
[14] S. Redon, N. Galoppo, M.C. Lin, Adaptive dynamics of articulated bodies. ACM Trans. Graph., 2005.
[15] Official SignWriting site, http://www.signwriting.org/
[16] L. Kervajan, Problèmes de représentation de la Langue des Sig Française en vue du Traitement Automatique, avril 2006