# TEXT-TO-SIGN LANGUAGE SYNTHESIS TOOL

*Maria Papadogiorgaki, Nikos Grammalidis, Dimitrios Tzovaras and Michael G. Strintzis*

Informatics and Telematics Institute, Centre for Research and Technology Hellas
1st Km Thermi Panorama str., 57001, Thermi Thessaloniki, Greece
phone: +(30) 2310464160, fax: +(30) 2310464164, email: mpapad@iti.gr;ngramm@iti.gr
web: http://www.iti.gr/

## ABSTRACT

This document presents an approach for generating VRML animation sequences from Sign Language notation, based on MPEG-4 Face and Body Animation. The proposed application aims in providing a computer-based sign-language synthesis output for the deaf and the hearing impaired. Moreover the application may be used as a teaching tool for relatives of deaf people as well as people interested in learning the sign language. The application receives text sentences as input and provides as output 3D animated VRML sequences able to be visualised in any VRML-compliant browser.

## 1. INTRODUCTION

The SignWriting system is a writing system for deaf sign languages developed by Valerie Sutton for the Center of Sutton Movement Writing, in 1974 [1]. Almost all international sign languages, including the American Sign Language (ASL) and the Brazilian Sign Language (LIBRAS), can be represented in the SignWriting system. Each sign-box (basic sign) consists of a set of graphical and schematic symbols that are highly intuitive (e.g. denoting specific head, hand or body postures, movements or even facial expressions). The rules for combining symbols are also simple, thus this system provides a simple and effective way for common people with hearing disabilities that have no special training in sign language linguistics, to write in sign languages.

An efficient representation of these graphical symbols in a computer system should facilitate tasks as storage, processing and even indexing of sign language notation. For this purpose, the SignWriting Markup Language (SWML), an XML-based format, has recently been proposed [6]. An online converter is currently available, allowing the conversion of sign-boxes in SignWriting format (produced by SignWriter, a popular SignWriting editor) to SWML format.

Another important problem, which is the main focus of this paper, is the visualization of the actual gestures and body movements that correspond to the sign language notation. Traditionally, dictionaries of sign language notation contain videos (or images) describing each sign-box, however the production of these videos is a tedious procedure and has significant storage requirements. On the other hand, recent developments in computer graphics and virtual reality, such as the new Humanoid Animation (H-Anim) [8] and MPEG-4 SNHC [3] standards, allow the fast conversion of sign language notation to Virtual Reality animation sequences, which can be easily visualized using any VRML-enabled Web browser.

In this document, we present the design, implementation details and preliminary results of a system for performing such a visualization of sign-boxes, available in SWML.

After the application of the proposed technique, the resulting sequences of MPEG-4 Face and Body Animation Parameters can be used to animate any H-anim-compliant VRML avatar using MPEG-4 SNHC BAP and FAP players, provided by EPFL [4].

The proposed technique has significant advantages:

- Web- (and Internet-) friendly visualization of signs. No special software has to be installed except a VRML plugin to a Web browser,
- Allows almost real-time visualization of sign language notation, thus enabling interactive applications,
- Avatars can easily be included in any virtual environment created using VRML, which is useful for a number of envisaged applications, such as TV newscasts, automatic translation systems for the deaf, etc.
- Efficient storage and communication of animation sequences, using MPEG-4 coding techniques for BAP/FAP sequences.

Significant similar work for producing VRML animations from signs represented in the HamNoSys transcription system to VRML has been carried out by the EC IST ViSiCAST project[10], and its follow-up project "E-Sign[11]. Current extensions of HamNoSys are able to transcribe all possible body postures, movements and facial expressions [12] and significant work towards supporting MPEG-4 BAPs has been made. The main contribution of the proposed approach in this paper is the attempt to work towards the same direction for the most common and popular representation of Sign Languages, which is the SignWriting notation system.

## 2. APPLICATION DESCRIPTION

The "EPFLBody" BAP player [4], developed by the École Polytechnique Fédérale Lausanne (EPFL) for the Synthetic and Natural Hybrid Coding (SNHC) subgroup of MPEG-4 was used to animate H-anim-compliant avatars

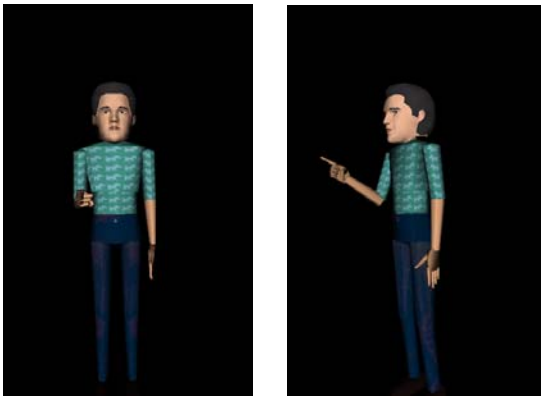

**Figure 1 Animation of the “YouÓsign in ASL using an H-anim avatar.**

using the generated BAP sequences. Since most BAPs represent rotations of body parts around specific body joints, this software calculates and outputs these rotation parameters as animation key-frames to produce a VRML (“animation descriptionÓ) file that can be used for animating any H-anim-compliant VRML avatar. The “MirafaceÓ FAP player, also developed for MPEG-4 SNHC, by MIRALab, University of Geneva and LIG, EPFL was used for Facial Animation. This software had to be modified so that:

a) VRML animation output is produced using one CoordinateInterpolator node per face model vertex. A problem with the chosen implementation is that the computational demands for the hardware that is reproducing these animations are increased. A possible solution for this problem that should be investigated in the future is to add CoordinateInterpolator nodes only for the points that have actually been moved.

b) The face model to be animated using the FAP frame sequence was attached to the body to be animated using the BAP frame sequence. Some slight modifications of the VRML face model were also required (e.g. addition of teeth).

Two frames from resulting animations are illustrated in Figure 1.

By including a VRML TouchSensor Node within the VRML file describing the H-anim avatar, the viewer can interactively start and/or replay the animation sequence, by clicking on the avatar. The viewer can also interact by zooming in and out to any specific body region and/or by rotating and translating the model within the 3-D space, in order to fully understand the represented sign.

The application, whose interface is illustrated in Figure 2, is currently based on a 3200-word SWML dictionary file, obtained by the SWML site, which has been parsed and inserted into a relational database. The user is allowed to enter one or more words, which are looked up in this dictionary. If more than one entry is found, all possible interpretations are presented to the user, so that he can choose the desired one. On the other hand, if no entries are found for a specific word, the word is decomposed using its letters (finger-spelling). In any case, the user may choose whether to include a particular term to the selected terms to be used for sign synthesis or not.

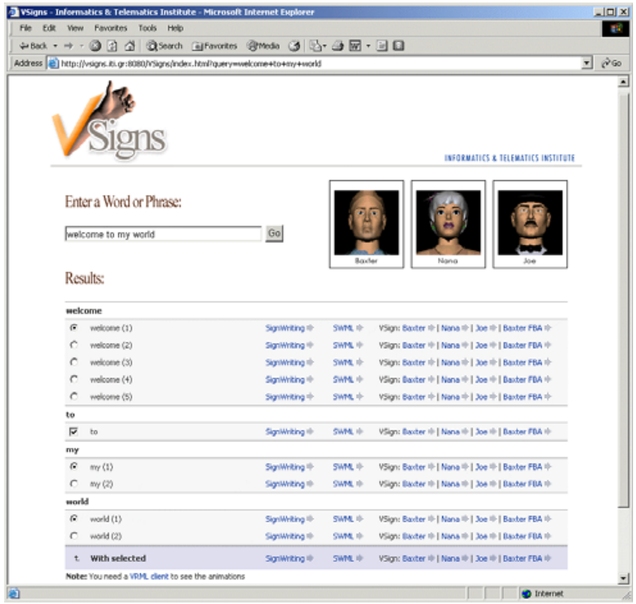


**Figure 2 Example query: “Welcome to my worldÓ The user may then select the desired terms and then produce and display sign synthesis results using the selected words or the entire phrase, using any of the available H-anim avatars.**

The user then selects a column corresponding to an H-anim compliant avatar, which is used for sign synthesis of the selected term or terms. A fourth column (“Baxter FBAÓ) allows the user to observe facial animation in addition to body animation, using the modified “Baxter avatarÓ Furthermore, the user may produce and display the corresponding sign(s) in SignWriting format (in PNG format) and SWML for a specific term or the selected terms [9].

## 3. SYSTEM ARCHITECTURE

The input for the sign synthesis system consists of the SWML entries of the sign boxes to be visualized. For each sign box, the associated information corresponding to its symbols is parsed.

Currently, symbols from the 1995 version of the Sign Symbol Sequence (SSS-1995) are supported. This sequence comprises an "alphabet" of the SignWriting notation system, while true images (in gif format) of each symbol contained in this sequence are available in [2].

The proposed technique first converts all individual symbols found in each sign box to sequences of MPEG-4 Face and Body Animation Parameters. The resulting sequences are used to animate a H-anim-compliant VRML avatar using MPEG-4 SNHC BAP and FAP players, provided by EPFL.

The system is able to convert all hand symbols as well as the associated movement, contact and movement dynamics symbols contained in any ASL sign-box. Manual (hand) gestures and facial animations are currently supported.

More specifically, the conversion of the symbols contained in a SWML sign box to BAP sequences starts by first examining the symbols contained within the input sign box. If no symbols describing dynamic information such as hand movements, contact or synchronization exist, the resulting BAP sequence corresponds to just one frame (i.e. a static gesture is reproduced). Information provided by the fields of the (one or two) hand symbols, contained in the sign box, is used to specify the BAPs of the shoulder, arm, wrist and finger joints. On the other hand, if symbols describing dynamic information exist, the resulting BAP sequence contains multiple frames, describing animation key-frames (i.e. a dynamic gesture is reproduced). The first key-frame is generated by decoding the existing hand symbols, as in the case of static gestures. Since the frame rate is constant and explicitly specified within a BAP file, the number of resulting frames may vary, depending on the complexity of the described movement and its dynamics. Synchronization symbols and contact also affect the represented movement and in some cases require special treatment.

When a signbox contains facial expression or animation symbols, the corresponding FAP frame(s) are determined by predefined lookup tables, which provide the FAP values defining one or more FAP frames per facial animation symbol. When two or more facial expression symbols co-exist within the same sign-box, these may either define an animation sequence or have to be combined all together (if each symbol activates different FAPs). The latter case, which is more common, is currently supported by the proposed system.

Smooth and natural-looking transitions between the Face and Body Animation parameters corresponding to each signbox is achieved by generating additional intermediate frames using a FAP/BAP interpolation procedure. Various interpolation functions can be selected in order to improve results. Since Body Animation Parameters represent rotations around specific joints, quaternion interpolation was seen to provide good results [7], but the complexity of the method is increased. For this reason, a linear interpolation function is used to generate additional FAP/BAP frames to implement:

a) The transition between the neutral face/body position and the first frame of the first sign-box

b) The transition between the end frame of one signbox and the start frame of the next signbox

c) The transition between the end frame of the last sign-box and the neutral body position.

Furthermore, in order to achieve Face/Body synchronization:

a) The frame rates defined for the FAP and BAP sequences should be equal

b) The number of generated FAP frames generated for each sign-box should be always equal to the corresponding number of BAP frames. In order to achieve this goal, the BAP frame sequence is first generated and then specific linear interpolation procedures are used to generate the FAP frame sequence.

The system architecture is presented in the following diagram (Figure 3).

End users under any platform can use the application. The hardware requirements are limited to a graphics card that can allow the animation of a 3D avatar used to perform the signs and an Internet connection. No special software has to be installed except a Web browser with a VRML97 plug-in.

## 4. SYSTEM EVALUATION

The system was evaluated using the proposed on-line system. This experimental Web application has already allowed us to identify problems with the synthesis of static and dynamic gestures, which have to be solved in the future, e.g. when contacts and complex movements are involved. A major problem that has to be solved occurs when the sign-box contains contact symbols. In that case the touch between the hands, or the hand and the face is difficult to be achieved. Problems may also occur for complex movements, when the inclinations of the hand joints, which have been estimated in each key frame, are not accurate enough for the exact description of the movement. In the future, improved reproduction of difficult movements (e.g. touching) will be made using inverse kinematics techniques as in [5]. There was no systematic evaluation of the system with real users till now. Further evaluation is planned for the future, using Greek and International SignWriting users, and attempts will be made to solve possible problems in the reproduction of specific signs.

## 5. CONCLUSIONS

A demonstrator for generating VRML animation sequences from Sign Language notation, based on MPEG-4 Body Animation has been developed. The system is able to convert almost all hand symbols as well as the associated movement, contact and movement dynamics symbols contained in any ASL sign-box. Furthermore, most facial expression and animation symbols are also supported, while torso movements will be also supported in the near future. Some facial expressions, e.g. cheek wrinkles, have not been implemented, since no FAPs exist to produce such movements. Results are currently being evaluated by SignWriting users and experts so that problems associated with specific SignWriting symbols are identified and solved. In the future, improved reproduction of difficult movements (e.g. touching) will be made possible using inverse kinematics techniques.

A short-term goal is to design other practical applications of the proposed system, either as a “plug-inÓto existing applications (e.g. sign language dictionaries) or as a stand-alone tool for creating animations for TV newscacts (e.g. weather reports). Particular emphasis will be given in applications that can be used and evaluated by national Sign Language communities (e.g. the Greek Sign language), thus many dictionaries of Sign languages, in SignWriter notation, are planned to be supported in the future.

## 6. ACKNOWLEDGEMENTS

This work has been conducted in conjunction with the “SIMILARÓEuropean Network of Excellence on Multimodal Interfaces of the IST Programme of the European Union (www.similar.cc).

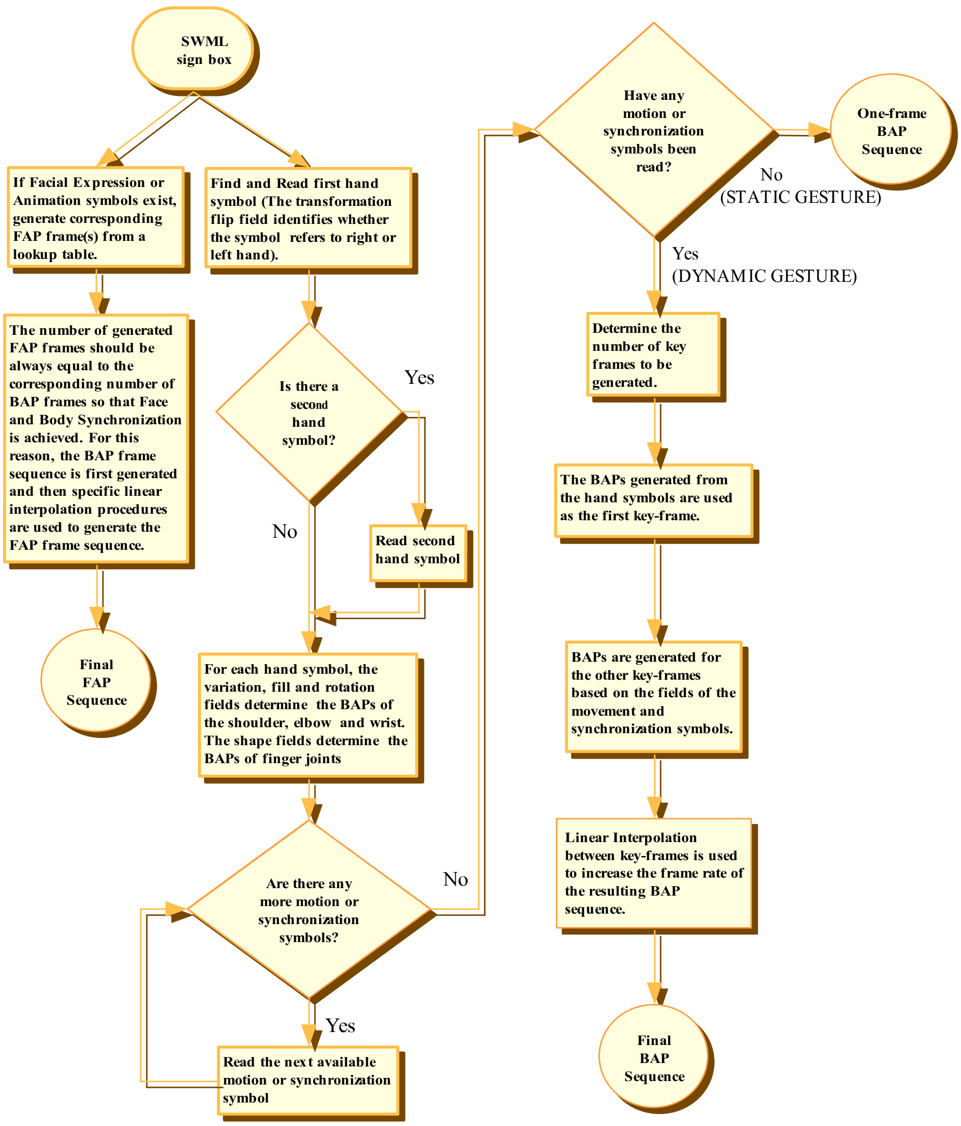


**Figure 3 Detailed system architecture diagram.**

## REFERENCES


[1] Official SignWriting site, http://www.signwriting.org/

[2] Official Site of SWML, http://swml.ucpel.tche.br/

[3] Moving Pictures Experts Group, (1999). Generic coding of audio-visual objects - Part 2: Visual, MPEG Document ISO/IEC JTC1/SC29/WG11 N3056. Maui.

[4] F. Vergnenegre, T.K. Capin, and D. Thalmann (1999). Collaborative virtual environments-contributions to MPEG-4 SNHC. ISO/IEC JTC1/SC29/WG11 N2802, http://coven.lancs.ac.uk/mpeg4/

[5] A. B. Grieve-Smith (2001). SignSynth: A Sign Language Synthesis Application Using Web3D and Perl. Gesture Workshop, London, pp. 134-145.

[6] A. Carlos da Rocha Costa, C. Pereira Dimuro (2001). Supporting Deaf Sign Languages in Written Form on the Web. The SignWriting Journal, Number 0, Article 1, July. http://gmc.ucpel.tche.br:8081/sw-journal/number0/article1/index.htm.

[7] M. Preda, F. Preteux (2001). Advanced virtual humanoid animation framework based on the MPEG-4 SNHC standard. Proc. EUROIMAGE ICAV3D'01, Mykonos, Greece, pp. 311-314.

[8] Humanoid Animation Standard Group. Specification for a Standard Humanoid: H-Anim 1.1. http://h-anim.org/Specifications/H-Anim1.1/

[9] Vsigns page, http://vsigns.iti.gr

[10] R. Kennaway (2001). Synthetic Animation of Deaf Signing Gestures. Gesture Workshop, pp. 146-157, London.

[11] Official site of E-sign (Essential Sign Language Information on Government Networks) project. http://www.visicast.sys.uea.ac.uk/eSIGN/

[12] Th. Hanke (2002). iLex - A tool for sign language lexicography and corpus analysis. In: Proceedings of the Third International Conference on Language Resources and Evaluation, Las Palmas de Gran Canaria, Spain., pp. 923–926.