Gesture synthesis from SignWriting notation

Yosra Bouzid (yosrabouzid@hotmail.fr)
Research Laboratory LaTICE
University of Tunis

Mohamed Jemni (Mohamed.jemni@fst.rnu.tn)
Research Laboratory LaTICE
University of Tunis

Abstract

Sign language synthesis has seen a large increase in applications over the past few decades, as it represents a potential solution to communication problem for the deaf community. All that is needed is to convert a writing form (books, newspapers, e-mails, internet pages…) or speech into sign. Most works in this area focus on the translation of a spoken language text into a fluid signing by using machine translation (MT), while others attempt to create synthetic animation from a sign language notation. We introduce in this paper a new method for the automatic generation of signed gestures from SignWriting notation using a 3D avatar. The SW notation is provided as input in an XML based format called SWML (SignWriting Markup Language). Such tool would help deaf readers to grasp and interact with the signed transcription through a more user-friendly interface.

Keywords: Gesture synthesis; Sign Language; SignWriting; SWML; 3D signing animation; avatar

Introduction

Sign languages have been demonstrated to be true languages at par with spoken languages that have their own independent vocabularies and their own grammatical structures, but do not have, until now, a widely established writing system. Since the 60's, several attempts have been made to create writing systems for sign language. These systems are expected to meet several objectives including teaching, the compilation of sign dictionaries, and especially the provision of a detailed phonetic transcription for research purposes (Lewis, 2009). But, ultimately, none of them have succeeded in becoming widely used by the deaf community for many reasons; the most important ones are their inability to capture the three dimensional quality of SLs, and their use of special symbols that are not easy to learn.

In order to make these notations more accessible to signers, a number of researchers have shown increased interest in synthesizing the transcribed data into articulated movements. The visualization of its content in real time will make a major difference in the usability of such notations since the information will be processed through visual-gestural modality. Some efforts have been oriented towards using recorded videos to display signing gestures. Although this approach provides a higher degree of accuracy, it implies considerable production cost as it requires the use of specific capture equipments and trained people with deep knowledge of the sign language. Others efforts have chosen to create animations of a 3D human like character since the content production is potentially easy and cost-effective.

In this context, we present in this paper a new system for automatically generating virtual reality animation sequences from a sign language notation in the well-known SignWriting system. The proposed system will process the SW notation, which is provided as input in an XML based format called SWML (SignWriting Markup Language), to extract the significant features of signing gestures and reformulate them afterward in an accurate gesture description language with additional details. The resulting script will be converted then in a sign modeling language (SML) which is interpreted automatically by the WebSign player (ElGhoul & Jemni, 2007) to animate the avatar.

Sign Language Notation Systems

At present, there are many notation systems used to transcribe sign languages. The three discussed here are those already used as a base in sign synthesis applications.

Stokoe Notation

Stokoe notation is an abstract phonemic script created for writing ASL. It defines fifty-five symbols divided into three groups of cheremes each representing a parameter for the sign formation: configuration, location and movement. The location and movement symbols are iconic, while hand shapes are represented by units taken from the Latin letters and numerals. These cheremes were written in a strict order with meaning dependent on placement within the string (Stokoe, Casterline & Cronenberg, 1976).

The use of this script is mostly restricted to linguists and academics for encoding singular signs in a dictionary or the study of sign linguistics, but using it to actually write signed languages is near impossible because it was designed specifically for manual gestures and has no way for indicating facial expressions and other non-manual features.

SignSynth is a sign language synthesis application which was developed in the University of New Mexico to generate many signs using Stokoe parameters. It takes as input a sign language text in ASCII-Stokoe notation and converts it to an internal feature tree. This underlying linguistic representation is then converted into a three-dimensional animation sequence in Virtual Reality Modeling Language (VRML) which is automatically rendered by a Web3D browser (Grieve-Smith, 2001).

HamNoSys

Hamburg Notation System is a well established phonetic transcription system that could be used to transcribe any sign language in the world (Hanke, 2004). It has its roots in the Stokoe notation but it has attempted to be more accurate and more phonetic. HamNoSys has an alphabet of about 200 pictographic characters that are notated sequentially. Generally, reading or writing a sign in a linear structure may cause greatest problems for deaf learners because the sign language phonemes cannot only be occurred sequentially, as in spoken languages, but also co-temporally.

ViSiCAST is a European project aiming to provide improved access to services and facilities for deaf citizens through sign language presented by a virtual human, or avatar (Kennaway, 2003). This project has developed an XML application language called Signing Gesture Markup Language basing on HamNoSys notation. SiGML allows SL sequences to be defined in a form suitable for performance by an avatar on computer screen. SiGMLSigning is the software system that has been used to generate 3D animations from a SiGML description.

SignWriting

SignWriting (SW) is a writing system invented by Valerie Sutton for communication purposes rather than linguistic purposes. Unlike the first two systems, SW transcribes the signed gestures spatially, in two-dimensional canvas, as they are visually perceived (Slevinski, 2012).

The International SignWriting Alphabet ISWA 2010 defines 30 groups of highly iconic symbols including 639 base symbols to represent the different phonological aspects of SL. This script is intuitive for new learners, but nevertheless requires mastery of a set of conventions different from those of the other transcriptions systems to become a proficient reader or writer (Miller, 2001).

The projects VSign (Papadogiorgaki et al., 2005) and SASL (Moemedi, 2010) provide two examples of SL visualization systems which use SignWriting notation as input. They interpret and convert the SWML (SignWriting Markup Language) format, as a computer encoding of SW, to MPEG-4 BAP (Body Animation Parameters) sequences which are used then to animate a virtual avatar.

SWML

SWML (SignWriting Markup Language) is an XML-based format that was developed for the storage and processing of SW texts and dictionaries, allowing the interoperability of SW applications and promoting web accessibility to deaf people in their own natural languages (Da Rocha, 2003). Each sign encoded in SWML corresponds to a signbox comprising the set of symbols that together represent the notation. To identify the aspect of sign language to which it corresponds and the variation to which was subjected, each symbol is specified by a unique ID. Further, the coordinates concerning the relative position of each symbol in the bi-dimensional canvas are also denoted. The representation of the sign "salute" in American Sign Language and its SWML translation are given in table 1.

Table 1: SWML encoding of sign "salute".

```
<signbox>
 <seq>01-04-004-01-05-02</seq>
 <seq>02-03-001-01-01-08</seq>
 <seq>02-01-001-01-01-01</seq>
 <seq>04-02-001-01-01-01</seq>
 <seq>04-04-002-01-01-01</seq>
 <sym width="36" height="35" left="-18" top="-18">04-02-001-01-01-01</sym>
 <sym width="36" height="35" left="-18" top="-18">04-04-002-01-01-01</sym>
 <sym width="21" height="21" left="16" top="-18">01-04-004-01-05-02</sym>
 <sym width="31" height="31" left="25" top="-54">02-03-001-01-01-08</sym>
 <sym width="13" height="14" left="7" top="-29">02-01-001-01-01-01</sym>
</signbox>
```

Indeed, there are two main drawbacks, in the SWML encoding, which make difficult to drive a signing avatar automatically. On the one hand, there is no temporal order in which symbols are written to interpret correctly the sign. A SignSpelling Sequence (Slevinski, 2012) can be proposed by the creator of each sign for internal ordering the symbols, but it should be notated that this sequence is an optional prefix, it cannot be found in all signboxes. For this reason, it is not possible to adopt this sequence for arranging the notation elements. In other hand, the SWML format encodes just the original glyphs and this means that a certain amount of information will be omitted. For instance, SW does not denote a symbol for location, so is SWML.

Proposed Approach

The input of our sign synthesis system is the SWML signbox of the SW notation to be visualized. To render content in sign language, the set of symbols found in the signbox will be processed and converted into 3D animation sequences by implementing four steps, as shown in Fig. 1.

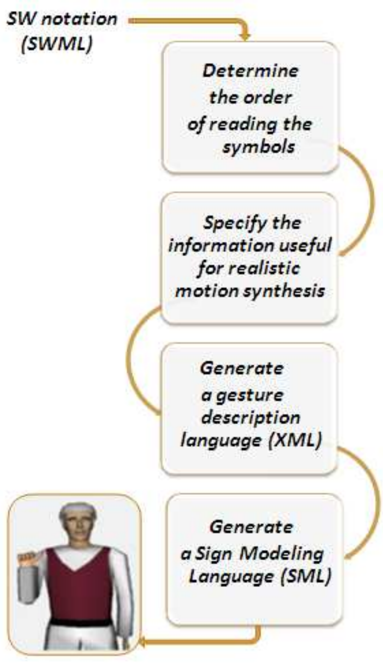


Figure 1: Overview of the proposed architecture.

The first step is dedicated to determine the correct order of symbols in the signbox, while the second is devoted to provide the missing information needed to describe the avatar movements. The third step ensures the automatic generation of a gesture description language which explicitly specifies how sign is articulated. Finally, the last step is devoted to transform the obtained sign description to SML for rendering animations.

Sorting Procedure

An SWML signbox can have two types of elements: "sym" as spatial symbol and "seq" as sequential symbol. The spatial symbols include 2-dimensional positioning and are considered unordered, while sequential ones form an ordered list of symbols IDs organized by the SignSpelling Sequence rules.

For the sorting process, we have defined the set of spatial tokens representing hands, directional movement, finger movement and contact as an underlying structure to the sign through which we can identify its type: static or dynamic, one-handed or two-handed, symmetrical (in which both hands have the same shape and make the same movement) or asymmetrical, in unit (both hands move as a group to the same direction) or not in unit, simple or compound. The compound sign here includes at least two hand symbols and two movement symbols for each articulator (right or left). Our proposed taxonomy for SW signs is given in Fig. 2.

According to the type and the underlying structure of a sign (USS), a set of precedence rules will be applied appropriately to arrange the list symbols. For example, the sign ”salute” mentioned above, is a simple sign articulated with one hand, whose USS includes one configuration, a directional movement and a touch contact. In this case, a precedence rule based on the direction of the movement arrow is used to determine if the touch contact occurs before or after the motion. This kind of rules serves essentially to identify the predecessors of a directional movement that are located at the tail of the arrow, and its successors that are located at the arrow head.

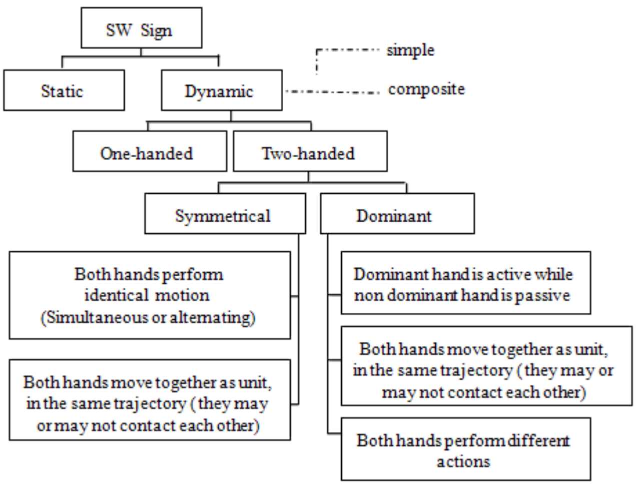


Figure 2: A taxonomy for SW signs.

But, when there is another type of movement (instead of directional arrow) in this USS, we should need to adopt another series of precedence rules, for example:

- if the movement symbol describes an arm circle or wrist circle (e.g. politics), the movement occurs always before the touch contact.
- if the movement symbol is an axial arrow or indicates a wrist flex (e.g. suffer) , the movement occurs always after the touch contact
- if the movement symbol describes a finger movement (e.g. bird), the finger movement occurs always after the touch contact.

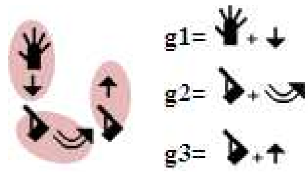


Figure 3: The ASL sign “dream”

Consider now the sorting procedure of a composite sign. The idea consists in partitioning the notation elements into subgroups where each one constitutes a simple sign, so, the problem of sorting amounts to identifying the order of symbols in each group. The sign “import” presented in Fig. 3, is a composite sign which includes three hand shapes and three directional movements. To determine the first group of this notation, we would firstly select the movement symbol that has no movement symbol which precedes it, and this by using the precedence rules. This token is actually the first movement that will be occurred in the sign. All the glyphs that precede this token are included also in the first group, while its successors form another group in the sign. If the group of successors represents a composite sign, the partition procedure will be applied again on these symbols.

Specifying Missing Information

The hand location is an important aspect in SLs that was not explicitly expressed in SignWriting and thereby in SWML. It can be deduced only from the visual image, with the physical arrangement of symbols within the sign. So, to determine this feature, we must always consider the relative position of each hand symbol from other symbols placements, especially those representing the head, shoulders, torso and limbs, and this by relying on the horizontal and vertical coordinates of the upper left corner of their glyphs, as well as on their predefined sizes (width and height). But sometimes, it will be more suitable to employ the detailed location symbols, which are not used in general writing and can be presented in the SignSpelling sequence, in order to be more precise in describing the hand location. For example, the ASL sign “improve” must involve these symbols to specify the different contact points on the passive arm.

Gesture Description

The purpose of the gesture description is to support the representation of signs at the phonetic level, and this by representing phonetically the significant features of signing.

The gesture description we present includes essentially two segments: posture and movement. The hand posture is defined by the shape of the hand, its orientation, its degree of rotation, and also, where necessary, its position in the signing space surrounding the signer's body. The movement segment can be used to represent the global movement, local movement or contact. To provide great precision, the dynamics and relevant temporal-spatial properties of movements are defined. We have used also a repeat attribute to indicate how many times the movement or contact is repeated. The "part" and "loc" attributes are used only for contact to specify respectively the parts of the hand and body that are in touch.

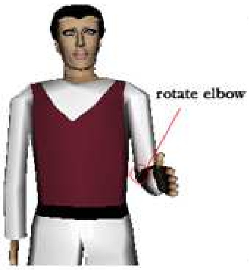
```
<sign hands=single inUnity=false symmetry=false >
  <mouth shape=smile />
  <eyebrows shape=straight_up />
<Posture>
  <Right-hand shape=H72 orientation=WP-Side rotation=45 />
<Posture/>
<Movement>
  <Right-hand contact=touch repeat=1 part=r_index_tip_outside
              loc=r_upper_face />
<Movement/>
<Movement>
  <Right-hand globalMovement=WP_straight joint=elbow
   repeat=1 size=small direction=up_right speed=normal />
<Movement/>
<sign/>
```

Figure 4: Gesture description of "salute"

The definition of non-manual components are also provided in the gesture description to specify the behavior of facial articulators such as raised eyebrows, puffed cheeks, eyes blinking, and bodily movements such as tilting the head, body and shoulders. The definition of these features is based on the corresponding definition in the SignWriting alphabet (ISWA 2010). Fig. 4 gives the description of the sign "salute" where the tip of the right index touches the upper right side of the face and then moves upward to the right in straight movement. The eyebrows of the signer are straight up and his mouth smiles.

Animating the avatar

The WebSign kernel (El Ghoul & Jemni, 2007) developed by our research laboratory LaTICE, integrates a 3D rendering module for generating 3D animations from an SML description of the sign to be visualized. The SML (Sign Modeling Language) is an XML based language designed to provide an extra layer around X3D for facilitating the manipulation of the virtual avatar. The SML animations are stored as a successive movement or rotation of groups of joints. Every movement has a fixed time during it the rotation of every joint in the group is done. Fig. 5 shows how left elbow joint can be rotated using SML. So, what we need to do in this phase is to convert of the obtained sign description into SML in order to be interpreted automatically by the WebSign player.




```
<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>
```

Figure 5: Rotation of elbow joint in SML

Conclusion

We have presented in this paper a new approach to synthesize signing gestures from SignWriting notation using a 3D avatar. Our proposal aims to offer additional support for the use of formal notation of SL to deaf users.

References

Costa, A. C. R. & Dimuro, G. P. (2003). SignWriting and SWML: Paving the way to sign language. *Proceedings of TALN* (pp. 193-202), Batz-sur-Mer

Grieve-Smith, A. B. (2001). SignSynthesis: A Sign Language Synthesis Application Using Web3D and Perl. *Proceedings of the 4th International Workshop on Gesture and Sign Language Based Human-Computer Interaction* (pp. 134-145), London

Hanke, T. (2004). HamNoSys - representing sign language data in language resources and language processing contexts. *Workshop on the Representation and processing of sign languages, LREC 2004* (pp. 1-6), Paris

Jemni, M., & El Ghoul, O. (2007). A system to make signs using collaborative approach. *Proceedings of the 13th International Conference on Computers Helping People with Special Needs* (pp. 670-677), Linz, Australia

Kennaway, R. (2004). Experience with and requirements for a gesture description language for synthetic animation. *Proceedings of 5th International Gesture Workshop* (pp. 300-311), Genova, Italy

Lewis, M. P. (2009). *Ethnologue: Languages of the World*. Sixteenth edition. Dallas, Tex: SIL International

Miller, C. (2001). Some reflections on the need for a common sign notation. Sign Language and Linguistics.

Moemedi, K. (2010). *Rendering an Avatar from SignWriting Notation for Sign Language Animation*. Master thesis, University of the Western Cape

Papadogiorgaki, M., Grammalidis, N., Makris, L. (2005). Sign synthesis from SignWriting notation using MPEG-4 standard, H-Anim and inverse kinematics techniques. International Journal of Disability & Human Development 2005;4(3), 191-203

Slevinski, S (2012). The SignPuddle Standard for SignWriting Text.

Stokoe, W.C., Casterline, D.C., & Cronenberg, C.G. (1976). Dictionary of American Sign Language on Linguistic Principles. (new edition). Silver Spring: Linstok Press.