SYNTHETIC SIGNING

I. Zwitserlood

Viataal, The Netherlands

ABSTRACT

This paper introduces a new technique for storing and representing sign language, which has been developed in the European projects *ViSiCAST* and *eSIGN*. The signed information created with this technique looks like a movie of a person signing, but it has the flexibility of a writing system. It is (almost) as easy to construct a signed "text" as a written text, provided that the author has sufficient sign language skills. The technique, described below, offers the possibility to provide much larger information access for Deaf[1] people.

INTRODUCTION

A large amount of information is provided in writing. In order to be able to read, one must have command of the language. Many deaf people's spoken language skills are insufficient for good comprehension of a written text. In the Netherlands e.g., the reading comprehension of only 25% of the deaf in the age of 7 - 20 equals that of a hearing 7-year old child (Wauters (1)). Most deaf adults are functionally analphabetic. As a result, deaf people have little access to information from the start, and the information deprivation will only grow in the following years.

In order for deaf people to have full access to information, it should be provided in a sign language. Sign languages are genuine, natural languages, that have their own grammar and lexicon. Sign languages are not universal, nor derived from spoken languages. Neither are they invented languages. Since there is no accepted writing system for sign languages, the only way to store and represent sign language information were movies of persons signing. However, movies have some serious disadvantages, because of which signed movies are used only scantly. First, recording a video of sufficient (signed) quality is expensive and time consuming. Second, mistakes and later changes in the text cannot be easily repaired but the movie (or parts of it) in which they occur have to be rerecorded. Third, for many people the required bandwith necessary for downloading movies from the Internet is still problematic. Fourth, movies are signer dependent, which is not always desirable, and fifth, movies are inherently two-dimensional, for which reason signed information (that is inherently three-dimensional) may be lost. Sign synthesis, in contrast, overcomes these disadvantages and provides an extra way of information provision in sign language.

ANIMATION: AVATARS

In sign synthesis, sign language content is expressed by a virtual human or avatar. An avatar is an abstraction of a human being, that consist of a skeleton and skin ("mesh"). The

---

[1] Following the conventions, I use the term "deaf" to refer to people with a hearing deficit and the term "Deaf" for people that belong to a Deaf Community, that is, people who share Deaf Culture and a sign language.

skeleton consists of a set of "joints". Such a joint has markers that mark its beginning and end and markers that define its orientation. The skeleton and mesh of an avatar are illustrated in Figure 1.

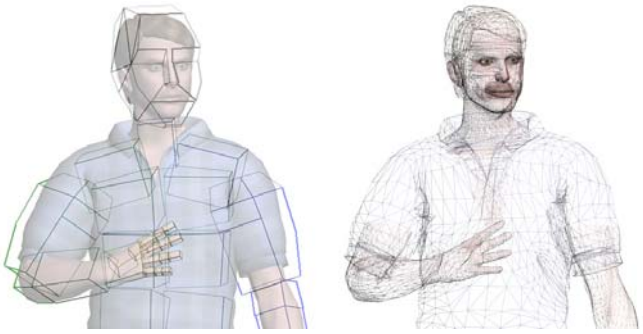


Figure 1: Avatar skeleton and mesh

Additionally, the signing avatars have a set of elementary deformations of the face, called morphs. Combinations of morph trajectories allow the avatar to perform the fine-tuned face movements that are a necessary part of understandable signed utterances. By their nature, although necessarily represented two-dimensionally, avatars are three-dimensional: a viewer has control over the avatar and can move (turn, pivot) the avatar in order to have a better view on the signing.

SIGN SYNTHESIS

Sign synthesis is not new. The most commonly used technique for sign language animation is through motion capturing. This technique has also been used in the EU project *ViSiCAST*, as described by Elliot *et al.* (2) and Kennaway (3). A user evaluation of captured signing on web pages with Deaf signers (Verlinden *et al.* (4)) showed reasonably good comprehension of the signed content and subjective approval of the avatar and the technique. However, motion capturing is not without drawbacks: expensive equipment is needed and it requires labour-intensive calibration of the equipment and post-editing of the captured data. An alternative to motion capturing is the synthetic creation of sign, which was developed in the EU projects ViSiCAST and eSIGN by a consortium of German, British and Dutch universities, companies and deaf institutions.

The general outline of the system is as follows (as illustrated in Figure 2). In brief, the shape of a sign is described in a phonetic transcription system for sign languages developed at the university of Hamburg: HamNoSys (Prillwitz *et al.* (5), Hanke (6)). Since HamNoSys was not originally designed for computer processing, a new version of the system was encoded in XML: SiGML (Signing Gesture Markup Language) and a translator from HamNoSys to SiGML was written.

A SiGML string enters the animation synthesizer (Animgen), where it is combined with the avatar descriptions (i.e. skeleton, mesh and morphs). Here, details that are missing in the SiGML string (such as default locations) are filled in and the movement of each avatar joint is calculated at intervals of typically 1/25 of a second. Finally, the avatar renderer displays the avatar on the screen, in the specified postures at the specified times.

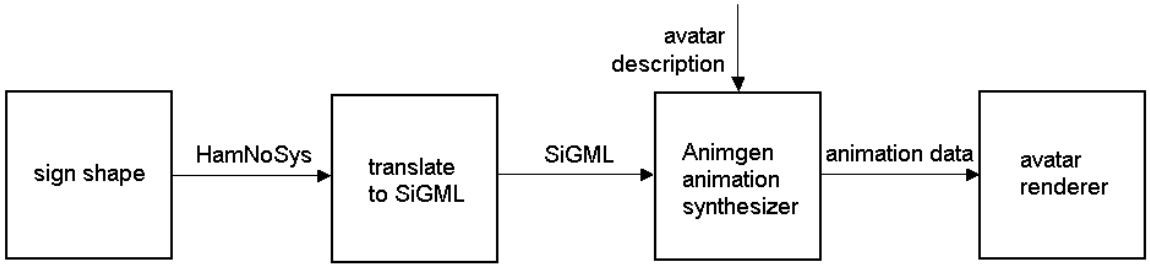


Figure 2: Block diagram of the generation of synthetic signing animation

HamNoSys - SiGML

HamNoSys was developed as a phonetic transcription tool for sign languages. It is language-independent and describes the shape of signs in components (such as handshape, hand orientation, location and movement) into detail. Being a linguistic system, predictable information (such as the fact that when the hands move to contact, contact will be at the outside of the hands and at the sides that are closest) is omitted in the transcription. However, what is predictable in human signing is not necessarily also predictable in a computer program. As said before, some of the omitted information is filled in by the animgen component. Nevertheless, the notation of synthetic signs often needs to be very explicit in order for the avatar to be able to perform signs correctly.

Originally, HamNoSys focuses on the manual part of signs. However, signing involves more than movements of the hands: non-manual components, such as movements of the torso, shoulders, head and mouth, and facial expressions are essential to comprehensible signing. Therefore, HamNoSys was extended with codes for non-manual parts (6). Furthermore, in some sign languages (e.g. Sign Language of the Nederlands [Nederlandse Gebarentaal: NGT] and German Sign Language), the mouth is very important in that many signs are accompanied with mouthings of words of the spoken language (Dutch and German respectively). So far, this is covered by adding the SAMPA code (a phonetic transcription in ASCII code) of such a spoken word to the HamNoSys transcription. An example of the HamNoSys transcription and SiGML translation for the sign for "to visit" (which has the mouthing *zuk*@) from NGT is shown in Table 1.

The handshape of this sign (flat hand with spread thumb) is expressed by in HamNoSys, the extended finger and palm orientation of the fingers by the respective symbols and . The initial hand position is near the chest ( ) and the movement consists of a curved forward movement ( ), with a simultaneous change in extended finger orientation ( ). The diaeresis ( ) indicates that it is a two-handed sign, in which both hands have the same configuration and perform the same action.

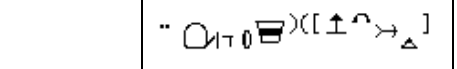
HamNoSys		**SiGML**
	→	<hamgestural_sign gloss="NGT_VISIT"> <sign_nonmanual> <mouthing_tier> <mouth_picture picture="zuk@"/> </mouthing_tier> </sign_nonmanual> <sign_manual both_hands="true" lr_symm="true"> <handconfig handshape="flat" thumbpos="out"/> <handconfig extfidir="uo"/> <handconfig palmor="l"/> <location_bodyarm location="chest" contact="close"/> <par_motion> <directedmotion direction="o" curve="u"/> <tgt_motion> <changeposture/> <handconfig extfidir="o"/> </tgt_motion> </par_motion> </sign_manual> </hamgestural_sign>
Sampa coding		
zuk@		

Table 1: HamNoSys transcription and SiGML translation of NGT sign for "to visit"

Avatar renderer

Figure 3: Avatar 'vGuido' makes the sign for "to visit"

As described above, SiGML sequences are combined with the description of a particular avatar in the Animgen component. The position of each of the avatar's joints are calculated for every time interval; this consecutive string of positions is shown as a smooth movement. The avatar's mesh is constantly and automatically adjusted to fit the joint constellation. The same holds for the 'deformations' of the face that are connected with the non-manual part of the signing. The animation data are sent to the avatar window on the computer screen. Figure 3 shows three frames (frame 5, 13, and 18) of the performance of the NGT sign for "to visit" by eSIGN avatar 'vGuido'. (Note that the avatar mouthes the Dutch word *bezoeken*.)

CONTENT CREATION

Sign animation becomes only interesting for the target group if it can be used in a purposeful way: to construct signed content. In many cases a written text is available, which needs to be translated into the target sign language(s). For various reasons, development of programs for the automatic translation between sign and spoken languages is still in its infancy. First, the current knowledge of the grammar of sign languages is insufficient. Second, even automatic translation of the written form of one language into the written form of another language is problematic still. Third, translation in this case is not from text to text, but from text to sign synthesis (comparable to translating written English into spoken Russian), which requires information about prosody (e.g. pace, "intonation") which is insufficient in the text.

Instead of automatic translation, two ways have been developed to create signed content. First, the partners developed semi-automatic translation tools, such as a weather forecast creation tool (ViSiCAST) and a tool for constructing job vacancy information (eSIGN). For obvious reasons these concern limited domains. Such tools can be used by anybody who commands the spoken language but not necessarily the target sign language. With these tools structured signed content is made on the basis of a lexicon and a set of preconstructed phrase and sentence models of the spoken language, that are linked to a lexicon and a set of preconstructed phrase and sentence models of the sign language.

Figure 4: question form of the sign for "to visit"

Second, the University of Hamburg has developed an editing program, in which sign transcriptions can be stored in a database. With this program a sign language user can retrieve the synthetic signs from the database and put

them in the desired order. The signs can then be provided with non-manual information that is necessary for a correct understanding. E.g. non-manual information has to be added to distinguish questions, negations and declarative sentences. This is illustrated in Figure 4, where the NGT sign for "to visit" is taken from a sentence meaning "Shall I come and visit you?". The signs in this NGT sentence need a slight forward tilt of the head and raised eyebrows in order to form a comprehensible question.

Creating content in this way is comparable to writing a text on a word processor in that all that is needed are the language skills, the necessary software and the skills to use the software. The person creating the signed content can do this at his/her own pace. S/he is not dependent on the availability of a studio or a camera person. Repairing mistakes and adapting signed text is rather easy in comparison to making and editing movies of signed text.

APPLICATIONS

The partners of the eSIGN consortium have constructed several applications with the new technique, besides translations of existing text, notably of web pages. In the Netherlands and in Germany, forms have been provided with signed support. A semi-automatic translation tool for job vacancies has been developed in the Netherlands. Partners in the United Kingdom developed a system with which non-signing office clerks can communicate in British Sign Language with Deaf citizens and the Deaf citizen can communicate in writing with the clerk. A consecutive project in the Netherlands *Gebarennet* (Sign Web) introduced the technique and possibilities via a bilingual (Dutch - NGT) web page: *www.gebarennet.nl*. The software was made available to the Deaf public via this web page. In the Netherlands, the attention is currently on the development of various educational materials and support for Deaf schools in the Netherlands.

CONCLUDING REMARKS

The techniques described in this paper can be used in a wide range of multimedia applications for Deaf people. Although their succes can undoubtedly be improved by further technical development, the techniques are sufficiently mature to move from demonstrators to real-life implementation, possibly starting with a pilot phase. They can help to make text-based information more accessible for Deaf people, and, furthermore, provides an extra, flexible way for Deaf people to inform other Deaf people that is not necessarily based on written text.

REFERENCES

1. Wauters, L., 2005. Reading Comprehension in Deaf Children: the Impact of the Mode of Acquisition of Word Meanings. Ph.D. Thesis. Nijmegen: Radboud University Nijmegen.

2. Elliot, R., Glauert, J.R.W., Kennaway, J.R., and Marshall, I., 2000. The Development of Language Processing Support for the ViSiCAST Project. 4th International ACM SIGCAPH Conference on Assistive Technologies (ASSETS 2000). pp. 101 to 108.

3. Kennaway, R., 2004. Experience with and requirements for a gesture description language for synthetic animation. Lecture Notes in Computer Science, Vol. 2915. pp. 300 to 311.

4. Verlinden, M., Tijsseling, C. and & Frowein, H., 2001. Sign Language on the WWW. Proceedings of the 18th International Symposium on Human Factors in Telecommunications (HFT2001) in Bergen, Norway, 5-7 November 2001. pp. 197 to 204.

5. 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. Hamburg, Signum.

6. Hanke, T., 2002. ViSiCAST Deliverable D5-1: Interface Definitions. Manuscript.

ACKNOWLEDGEMENTS

The development of the technique described in this paper was made possible by EU funding under the Framework V IST Programme (Grant IST-1999-10500) and the eContent Programme (EDC-22124 ESIGN/27960). I am grateful to my collegues at Viataal and to my collegues at the partner institutions (Televirtual Ltd., University of East Anglia, Institut für Deutsche Gebärdensprache at the University of Hamburg, Norfolk County Council, Royal National Institute for the Deaf, and Systematics Integrations) for their support.