An Architecture for Sign Language Synthesis

Maxime Delorme[1], Michael Filhol[1] and Annelies Braffort[1]

[1] LIMSI-CNRS, BP 133 F-91403 Orsay cedex FRANCE

Keywords: Sign Language, Signing Avatars, Sign Synthesis.

1 Introduction

Deaf people are not always able to read plain text. In order to give them access to every piece of audio information, we choose to display sign language (SL) on screens. We could use videos of people signing isolated signs to make complete utterances but such a method lacks in flexibility. Thus, we generate animations for a dedicated virtual humanoïd called a signing avatar. The synthesis is completely automatic and based on a description model of signs. We have chosen to use a geometric model developed at LIMSI. In this paper, we present a process to generate standard animation data from sign descriptions of the model.

2 Animation generation process

When it comes to sign generation, lots of approaches use parametric models [3], that is to say models that describe a sign as a tuple of parameters whose values are taken in finite sets. A discussion of the drawbacks of such models can be found in [1]. Combining a parametric approach with temporal structures, Liddell and Johnson [4] propose a hybrid description model. Although it has the same drawbacks as typical parametric models, it regards a sign as a sequence of timing units, which makes is adequate to describe dynamics. Following the same principle, Filhol developed a sign description model based on a geometric approach [2]. This temporal approach divides signs in *timing units*, being either *key postures* (when the body has reached a stable state) or *transitions* (when the body moves from a stable state to the next). The model provides three constraints to build timing units: placement of an articulation in space, orientation of a bone along a vector or in a plane.

The process we developed builds an animation in BioVision Hierarchical data (BVH) from the geometric description of the signs. The construction is made in a four-step procedure : 1- set up; 2- postures resolution; 3- transitions computation; 4- export. Each of these steps is composed of one or more sub-treatments (see fig.1).

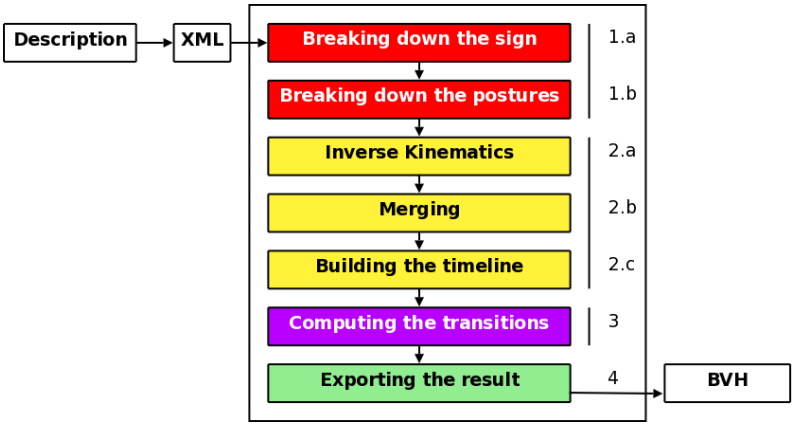

Fig. 1. The animation synthesis process in four steps.

Set up divides the sign in sub-problems that will be solved by the next steps. The first division step cuts the sign in timing units, separating key postures from transitions. Then, key postures are divided in separate kinematic problems. *Posture resolution* begins by solving each kinematic problems. Then, solutions are merged into complete postures. Finally, postures are placed according to the animation timing on a timeline. From this timeline, the *transition computation* phase builds an animation, from each posture to the next. Finally, the process *exports* the resulting animation to a standard format (BVH data file). This format is readable by most animation programs and specific plateforms dedicated to signing avatars.

4 Conclusion

Set up phase and inverse kinematics have been successfully developed and tested. In February we expect the posture resolution phase to be complete. However, the merging step (2.b) may have to deal with conflicting situations (e.g. placing the wrists too far apart). This treatment will need deeper investigation to solve correctly those problems.

References

1. Filhol M., Braffort A., "A Sequential Approach to Lexical Sign Description", *In proceedings LREC 2006,* 2006.
2. Filhol M., "Modèle descriptif des signes pour un traitement automatique des langues des signes" (to be translated), *PhD Thesis,* LIMSI/CNRS, France 2008.
3. Kennaway R., "Synthetic Animation of Deaf Signing Gestures". *Gesture and Sign Language in Human-Computer Interaction*, LNCS/LNAI vol. 2298, Springer, 2001.
4. Liddell S.K., Johnson R.E, "Amecian Sign Language : The phonological base", *Sign Language Studies vol 64,* Gallaudet University Press, USA, 1989.