Fine Tuning Dynamics in Contextualized Classifier Constructs from Linguistic Descriptions

John McDonald
jmcdonald@cs.depaul.edu
DePaul University
Chicago, IL, USA

Michael Filhol
michael.filhol@limsi.fr
LIMSI, CNRS, Université Paris Saclay
Orsay, France

KEYWORDS

Sign Language, Avatar, Classifiers, Proforms, Human Motion

1 INTRODUCTION

The movements of the human body during signing are filled with subtleties that contribute to the richness of sign language in both linguistic and biomechanical capacities. The pacing and dynamics of the body's motions can be influenced not only by the emotions communicated by the signer but also by the linguistic (grammatical) structure of the discourse [3, 6]. Reproducing such subtleties in the dynamics is critical for sign synthesis via a signing avatar, and both the avatar's linguistic input and animation system must be capable of representing and reproducing such variations in movement.

In particular, many proform constructs such as classifier placements and size and shape specifiers have highly context sensitive and flexible movements. Prior work has concentrated on difficulties in synthesizing simple positional variability in classifier constructs [1, 2, 4]. However, in different contexts, similar constructs may exhibit variation in motion styles that avatars have heretofore struggled to capture. For example, the motion of placing a classifier in space will vary in the following circumstances:

(1) anchoring a single object in space;
(2) placing multiple but separate items in space consecutively;
(3) placing the multiple items of a set in space consecutively.

In cases (1) and (2), the signer tends to deliberately move the hand down into place with a short pause at the end of each movement, whereas in case (3), signers have been observed to make quicker motions for each object placement without pausing between them. Additionally, there are other accompanying subtleties that must be present in the avatar motion to produce natural signing. Faithfully producing those subtleties in form is vital to creating natural avatar motions.

2 TUNING MOTION FROM LINGUISTIC DESCRIPTIONS

Our prior work [1] explored the functionality necessary for the Paula avatar to use artist templates to allow natural placement and movement of classifier constructs based on a structured linguistic description of the motion. The technique effectively handled the motion for case (1) above. However, for supporting the dynamics differences for cases (2) and (3), additional levels of control are necessary, similar to those animators have been manually using for a long time [5], including:

Figure 1: Single Frame from the placement of four plates in a single set

- the shape of the motion path;
- the abruptness or ease in which a body part approaches or leaves a target;
- the synchronization the timing of torso movement with the rest of the motion;
- other coordinated body motions that affect the perception of the movement.

Tuning such features automatically from linguistic description requires that the avatar use both knowledge of human movement and linguistic structure to trigger changes in these parameters automatically.

Our experiments use the AZee description model for the linguistic input, as it provides the required flexibility. The reason is that AZee was precisely designed to formalize identified form–meaning associations, and therefore allows short-cutting on things that have meaning and capture contrasts between the numbered items above as described in [1].

2.1 Placing a single object

Case (1) can be exemplified with the placement of, say, a plate, using a two-handed classifier for flat round objects as in Figure 1 in the middle of the signing space. An AZee expression such as (E1) can be compiled to render the corresponding animation. It applies one rule "place-classifier", which carries the meaning "object of the classifier's type *be at* the location". It anchors the object in space, and specifies both the settling (downward) movement of the hands into the right position and the synchronized eye gaze towards the same location.

E1 place-classifier(class-flat-round-large, midssp)

SLTAT 2019, September 29, 2019, Hamburg, Germany

In this situation, the avatar system can shortcut directly on the known place-classifier process while filling in the necessary body postures and motions from an artist template. It then can use the timing and duration information directly from AZee to coordinate the manual and gaze processes. As described in the previous work, the artist template provides several important cues for producing a natural posture and settling motion. A schematic diagram for the motion is displayed in Figure 2 where the horizontal axis is time and the vertical axis is the height of the signer's wrists.

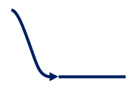

Figure 2: Diagram of a settle movement

2.2 Placing multiple separate items

Case (2) can be exemplified with a set of four plates on a table, say at points $p_1..p_4$, each anchored in its own position. AZee not allowing for signs in sequence with no justification in meaning, the way to string the four consecutive classifier placements is to use the rule "each-of", as in expression (E2) below[1].

E2 each-of(for p in p_1, p_2, p_3, p_4: place-classifier(class-flat-round-large, p))

This expression takes a list of signed productions as an argument, and conveys the fact that each placement is applied in space, with no order of importance or precedence. The expression specifies the resulting form to render, which consists of the expected sequence, with a specific holding time at the end of each item, allowing the interpretation of the above meaning. Expression (E2) compiles into the animation available at:

http://asl.cs.depaul.edu/proforms/Plates-EachOf.mp4

Since this process is a repeated application of the same movement at different spatial locations, the avatar system can simply apply the artist template as before several times with the additions of the holds specified by AZee. Figure 3 shows a schematic diagram of the resulting motion in this case.

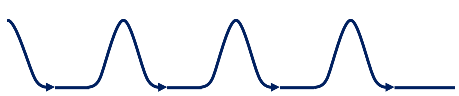

Figure 3: Diagram of movement for "Each-of"

2.3 Placing multiple items from a set

We can illustrate case (3) with a similar set of four plates, this time producing quicker movements back to back. In terms of meaning, the focus shifts from the individual items to the set formed by all of them together. Replacing the rule "each-of" with "all-of" captures this change in meaning, using the same argument list. The resulting form specified by this new rule is a shorter duration or squeeze for each of the items, and does not specify hold blocks between them.

E3 all-of(for p in p_1, p_2, p_3, p_4: place-classifier(class-flat-round-large, p))

In this case, analysis of corpus examples shows that the motion is altered in more ways than those provided by AZee. The downward placement actually ceases to "settle" and becomes a distinct bouncing between the placements. The top-down short-cutting system allows the avatar to distinguish the difference between the each-of and all-of. So, Paula is free to alter the motion within the bounds of the linguistic constraints to produce this bounce. This application of the "coarser the better" principle is in fact necessary here to provide the correct motion allowing the avatar to:

(1) cause the arm's approach to the target point to be more abrupt instead of easing-in;
(2) start the next cycle abruptly to complete the bouncing of the arm at the target point;
(3) depending on the geometry of the classifier and the amount of arm motion involved, to shorten the stroke of the cycles to compensate for the squeezed timing;
(4) alter the timing of the signer's eye and head movement in synchronization with the actions on the hands, with a more continuous progression.

The effects of all of these can be seen in figure 4 where the path bounces instead of coming in tangentially and the heights of the cycles are somewhat shorter than before. The motion produced can be viewed in the animation at the following site:

Figure 4: Diagram of movement for "All-of"

http://asl.cs.depaul.edu/proforms/Plates-AllOf.mp4

This presentation will demonstrate these linguistic rules in detail, discuss the output from AZee received by the avatar, and explore the ways in which the Paula system uses short-cutting to produce the required dynamics. Finally, it will provide a discussion of the motion controls triggered by the linguistic output and show the resulting rendered animations.

REFERENCES

[1] Filhol, M., and Mcdonald, J. Extending the azee-paula shortcuts to enable natural proform synthesis.
[2] Huenerfauth, M., Marcus, M., and Palmer, M. *Generating American Sign Language classifier predicates for English-to-ASL machine translation.* PhD thesis, University of Pennsylvania, 2006.
[3] Johnson, R. E., and Liddell, S. K. A segmental framework for representing signs phonetically. *Sign Language Studies 11*, 3 (2011), 408–463.
[4] López-Colino, F., and Colás, J. The synthesis of lse classifiers: From representation to evaluation. *Journal of Universal Computer Science* (2011).
[5] Thomas, F., Johnston, O., and Thomas, F. *The illusion of life: Disney animation.* Hyperion New York, 1995.
[6] Wilbur, R. The linguistic description of american sign language. In *Recent perspectives on American sign language.* Psychology Press, 2017, pp. 7–31.

[1] As is standard in most programming languages, the "for" operator in (E2) iterates a variable over values to repeat a portion of code, in this case to generate a list of similar expressions, where only the plate's location changes.