Analysis in Support of Realistic Timing in Animated Fingerspelling

Nkenge Wheatland* UC Riverside
Ahsan Abdullah† UC Davis
Michael Neff‡ UC Davis
Sophie Jörg§ Clemson University
Victor Zordan¶ Clemson University

Figure 1: The ASL letters of the word HEAVEN. According to our analysis, the four middle letters of this word are signed faster than the first and last letter, with the last letter pose being held for the longest amount of time.

Abstract

American Sign Language (ASL) fingerspelling is the act of spelling a word letter-by-letter when a specific sign does not exist to represent it. Synthesizing intelligible ASL, which includes fingerspelling as an integral part, is important to create signing virtual characters for training and communicating in virtual environments or further applications. The rhythm and speed of fingerspelling play a large role in how well fingerspelling is understood. Using motion capture technologies, we record fingerspelling and analyze timing information about letters in the words. Our goal is to identify fingerspelling timing information and use it to create fingerspelling animations that are natural and understandable.

Index Terms: I.3.0 [Computer Graphics]: General—; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Animation; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual Reality; I.3.8 [Computer Graphics]: Applications—

1 Introduction

There are many cases in American Sign Language (ASL) where words do not have a pre-defined sign. In these instances, words are spelled out using fingerspelling with the individual letter signs of the ASL alphabet. ASL fingerspelling has been described as a "signed representation of written English" [8]. Language elements often fingerspelled include proper nouns, acronyms, and technical terms [7].

An important step in analyzing how fingerspelling is performed is understanding how speed and timing play a part in the process. Patrie and Johnson use the following terminology to describe three different forms of fingerspelling [5]:

- Careful fingerspelling - slower spelling where each letter pose is formed
- Rapid fingerspelling - quick spelling where letter poses are often not completed and signs/letters contain remnants of other signs/letters in the word
- Lexicalized fingerspelling - spelling that often uses no more than two hand shapes to convey the meaning of a word; looks more like a sign than fingerspelling [2]

Those who communicate in ASL are sensitive to the rhythm of fingerspelling. The rhythm and speed can affect how intelligible fingerspelling motions are to a fellow communicator [8]. For the current project, we extract timing features of fingerspelling such as letter speed and letter hold length from recorded data. These findings inform a synthesis system to animate the timing of letters and transitions that occur between letter poses. Our goal is to fit the speed and rhythm of natural fingerspelling to produce comprehensible animations. The timings can be used to animate signing full body avatars for training or communication in virtual environments.

Fingerspelling is a well structured space of 26 standalone characters that combine in different ways to convey a particular meaning. Previous models have been developed to animate fingerspelling. Though these systems attempt to model some of the parameters that inform the speed and rhythm of fingerspelling, they do not account for how speed varies throughout the spelling of the word.

2 Related Work

In linguistic research of ASL, studies on fingerspelling speed have focused on rapid fingerspelling as it is the more natural way of fingerspelling for those fluent in the practice. Researchers have noticed interesting features in the timing of fingerspelling. Quinto-Pozos' experiments look at letter speed in relation to the length of words [6]. The author analyzes letter speed for short words (3 or fewer letters) and long words (4 or more letters) in a speech-like setting to determine if the length of the word affects the speed at which the letters are signed. He finds that short words are signed at an average rate of 7.08 letters/second (141 ms/letter) and long words are signed at a rate of 7.65 letters/second (130 ms/letter) meaning that longer words are fingerspelled at a faster speed than short words. All of the words in the study had an average speed of 5-8 letters/second (125 - 200 ms/letter).

Recent fingerspelling animation methods have focused on creating anatomically accurate hand shapes for each letter and interpolating between these shapes. Often these letters are presented at a constant rate [7], which is not consistent with how words are fingerspelled [8]. Adamo-Villani and Beni interpolate between keyframed letter poses [1], which are each held for the same amount of time with exceptions that include pauses after the last letter of a word and speeding up the signing of syllables and certain double letters. Huenerfauth uses a commercial sign language software and applies a parameter of 4.1 letters/second (243 ms/letter) to all of the letter poses except the last letter which is held for an extra 0.8 seconds for emphasis [4]. Even at this speed, which is slower than

*e-mail: wheatlan@cs.ucr.edu
†e-mail: aabdullah@ucdavis.edu
‡e-mail: mpneff@ucdavis.edu
§e-mail:sjoerg@clemson.edu
¶e-mail:vbz@clemson.edu

Number of Letters	**Careful**		**Normal**	
	Avg Hold Time	**Avg Speed**	**Avg Hold Time**	**Avg Speed**
Three Letters	0.16 sec	0.80 sec/letter	0.10 sec	0.54 sec/letter
Four Letters	0.14 sec	0.71 sec/letter	0.09 sec	0.48 sec/letter
Five+ Letters	0.13 sec	0.63 sec/letter	0.08 sec	0.41 sec/letter

Table 1: The average amount of time spent in letter holds and the average rate of speed of each letter depending on the number of letters in the fingerspelled words. This information was extracted from motion capture data recordings.

the speeds noted as being natural in the above studies, participants found these fingserspelling animations difficult to understand.

We believe that by using timing based on recorded variations in fingerspelling, we can produce animations that are more comprehensible than these proposed models.

3 Synthesis Procedure

To synthesize fingerspelling, we are developing a data-driven statistical timing model to inform an animation system. The timing data is determined from motion capture recordings. Our current animation system uses straightforward interpolation.

We motion captured a person proficient in ASL fingerspelling. Our actor was asked to sign the alphabet, a collection of letter pairs, a collection of words of varying lengths, and series of sentences that included terms that would most likely be fingerspelled normally. The hand motion was recorded using gloves with bend sensors called CyberGloves [3] and set of motion capture markers.

We recorded two motion databases. Database 1 contains signs of individual letters as well as signs of pairs of letters. These signs may be used to produce the final animations. Database 2 contains the fingerspelled words for timing analysis. We plan to compare our final animations to these recordings.

To determine information about speed from our recordings, we extract information such as the number of seconds per word, number of letters per second in words of differing lengths, number of seconds per letter hold, and the number of seconds per inter-letter transition. Words of similar length have been grouped together to perform analysis on how the lengths of words affect the speed at which the letter of the word is signed. When labeling letter hold frames, a hold is defined to begin when we can see a letter pose without any joint movement, and to end when a joint or multiple joints begins to move out of a pose. The current analysis is being performed using this labeling.

4 Assessment

Table 1 and Figure 2 show our current timing results. From these timing results and letter poses from Database 1, we have produced animations of various words such as CAB and HEAVEN. Transitions between letter poses is achieved using straightforward interpolation. Our timing results in Table 1 match results previously presented by Quinto-Pozos that state longer words are signed at a faster speed than shorter words [6]. Initial comments from a few sign language experts include that animations with our timing look more "human" than animations with a constant rhythm, which supports our findings. We have also been able to produce timing information for example words left out of our analysis and it correlates to the timing extracted from our recordings.

5 Conclusion

In this poster, we present results of extracting timing information from motion capture recordings of ASL fingerspelling. We use this timing information to build a fingerspelling animations that emulates the way humans fingerspell. Our goal is to produce animations that are natural and can be understood by a native signer. Our

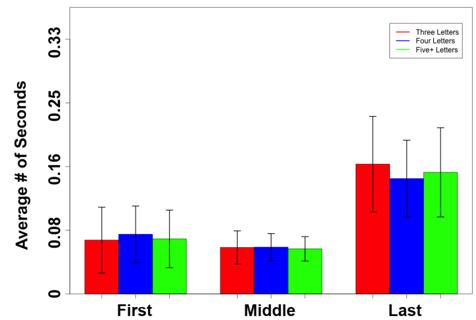

Figure 2: The average amount of time spent holding the first, middle, and last letters of the fingerspelled words. Red bars are for three letter words, blue bars for four letter words, and green bars for words with five or more letters. Less time is spent on middle letters and the last letter pose is held on average longer than the other letters in the word.

current results show that our timing analysis is consistent with previous linguistics research into fingerspelling speed and our animations are favorable to animations using constant timing. There are many steps to produce animations of high quality, such as a more sophisticated transition method, but currently we have a result that can represent the natural timing found in fingerspelling.

Acknowledgements

References

[1] N. Adamo-Villani and G. Beni. Automated finger spelling by highly realistic 3D animation. *British Journal of Educational Technology*, 35(3):345–362, 2004.

[2] R. Battison. *Lexical Borrowing in American Sign Language*. Linstok Press, Silver Spring, Maryland, 1978.

[3] CyberGlove Systems. Motion capture system website, http://www.cyberglovesystems.com/products/cyberglove-ii/overview, August 2014.

[4] M. Huenerfauth. A linguistically motivated model for speed and pausing in animations of american sign language. *ACM Trans. Access. Comput.*, 2(2):9:1–9:31, June 2009.

[5] C. J. Patrie and R. E. Johnson. *Fingerspelled word recognition through rapid serial visual presentation : RSVP*. DawnSignPress, San Diego, CA, 2011.

[6] D. Quinto-Pozos. Rates of fingerspelling in american sign language. *Poster at the Theoretical Issues in Sign Language Research conference, West Lafayette, Indiana*, 2010.

[7] E. Sedgwick, K. Alkoby, M. J. Davidson, R. Carter, J. Christopher, B. Craft, J. Furst, D. Hinkle, B. Konie, G. Lancaster, S. Luecking, A. Morris, N. Tomuro, J. Toro, and R. Wolfe. Toward the effective animation of american sign language. In *8th International Conference in Central Europe on Computer Graphics, Visualization and Interactive Digital*, 2001.

[8] S. Wilcox. *The Phonetics of Fingerspelling*. Studies in Speech Pathology and Clinical Linguistics. John Benjamins Publishing Company, 1992.