Effects of character geometric model on perception of sign language animation

Nicoletta Adamo-Villani*, Ronnie Wilbur♦, Petra Eccarius§, Laverne Abe-Harris♣
*♣Department of Computer Graphics Technology, Purdue University, USA
♦§ Department of Speech, Language and Hearing Sciences, Purdue University, USA
{nadamovi@purdue.edu; wilbur@purdue.edu; peccarius@purdue.edu ; harris1@purdue.edu }

Abstract— **This paper reports a study that aimed to determine whether character geometric model (i.e. segmented vs. seamless) has an effect on how animated signing is perceived by viewers. Additionally, the study investigated whether the geometric model affects perception at varying degrees of linguistic complexity--specifically handshape complexity. We modeled and animated two polygonal 3D characters: Torrents, one seamless mesh, and Robby, a fully segmented avatar. Both characters had similar geometrical proportions, identical skeletal systems, similar visual styles and color schemes, and met standards of good character design. Each signed 60 stimulus signs, divided into three groups—those with simple (group I), moderately complex (group II), and complex (group III) handshapes according to factors established in the linguistic literature. 53 participants, who learned ASL by age 5, viewed animated clips in random order via web survey. They (1) identified the sign (if recognizable), and (2) rated its realism using a 5-point Likert scale. Findings show that the seamless avatar (Torrents) was rated highest, and simple handshapes were rated higher than moderately complex and complex ones. The interaction between character and handshape complexity was also significant. For Robby (more than for Torrents), ratings decreased as handshape complexity increased. The lower ratings for Robby could indicate a preference for seamless, deformable characters over segmented ones, especially in signs with complex handshapes.**

Keywords: 3D Animation; Sign Language; 3D Characters

I. INTRODUCTION

Computer animation of American Sign Language (ASL) has the potential to remove many of the barriers to education for deaf students because it provides a low-cost and effective means for adding sign language translation to any type of digital content. Like a video of a human ASL signer, computer animation technology allows for direct communication of ASL in a dynamic visual form that eliminates the need for closed captioning text, symbolic representations of the signs [1] or static sign images.

The benefits of rendering sign language in the form of 3D animations have been investigated by several research groups [2-6] and commercial companies [7-9] during the past decade and the quality of 3D animation of ASL has improved significantly. However, its effectiveness and wide-spread use is still precluded by one major limitation: low realism of the signing motions of the characters- we define realism of animated signing as 'perceived closeness to real signing'. Low realism results into poor legibility of the animated signs and low avatar appeal.

The general goal of our research is to advance the state-of-the-art in ASL animation by improving the quality of the characters' signing motions. Our first step towards this objective is to determine whether certain characteristics of a 3D avatar, such as character geometric model (i.e. segmented versus seamless) and character visual style (i.e. stylized versus realistic), have an effect on the way its signing motion is perceived by the viewer. Specifically, the objective of the research reported in this paper was to determine whether the presence of organic deformations of the character skin during motion has a significant effect on perceived realism of the animated signs. Additionally, the study aimed to determine whether the complexity of the handshape is a significant factor affecting sign identification for each character.

The paper is organized as follows. In section 2 we present the 3D characters and in section 3 we describe the stimulus items used in the study. Experiment design is described in section 4 and findings are reported in section 5. Discussion of results and conclusive remarks are included in section 6.

II. THE ANIMATED CHARACTERS

For the purpose of this study we designed, modeled, textured, rigged, and animated two polygonal 3D characters using MAYA 2008 software [10]. One character, Torrents, is constructed as one seamless polygonal mesh, the other one, Robbie, is a fully segmented avatar comprised of 130 rigid polygonal components. Both characters (shown in Figure 1) have similar geometrical proportions, are rigged using the same skeletal system, and present the same color scheme. Torrents' mesh was skinned to the skeleton using a smooth bind with a maximum number of influences equal to 4, whereas the segments that make up Robby's body were parented to the appropriate skeleton joints. We could have duplicated Torrents and created a segmented version of him so the two characters would have been perfectly identical. We chose not to because the objective is to determine whether a visually appealing segmented character is able to convey the signs as realistically as a character that shows organic deformations of the skin. By segmenting Torrents we would have produced a 'stand-in character' rather than a segmented avatar that meets the standards of good character design.

All signing animations were keyframed on Robby and the animation data were exported and applied to Torrents; thus, the animated signing motions are identical for both characters. Videos of a native signer performing the signs were used as reference footage for the creation of the clips. The signer was actively involved in the animation process.

978-0-7695-3734-4/09 $25.00 © 2009 IEEE
DOI 10.1109/VIZ.2009.23

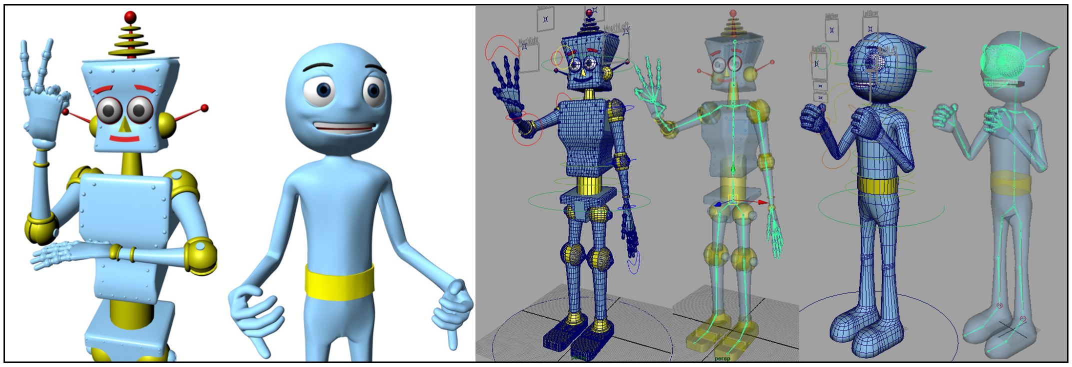


Figure 1. From the left: rendering of Robby; rendering of Torrents; Robby's geometry and skeletal system; Torrents' geometry and skeletal system.

III. THE STIMULUS ITEMS (SIGNS)

Words in ASL (or any sign language) are built up of different kinds of visual components (called 'parameters'), much like spoken language words are built up of different kinds of sounds. Specifically, signs are made up of different handshapes, places of articulation (locations) in relation to the body, and movement types, with hand orientation and other body movements playing lesser roles. Of these parameters, handshape is the most studied and best understood in terms of linguistic complexity. For this reason, (and because of the perceptually complicated sets of joint configurations involved), the 60 signs used as stimulus items for this project were chosen primarily based on their handshape complexity. The varying levels of complexity were determined using factors established in the sign language literature, such as anatomical ease and frequency of occurrence both within ASL and across other known sign languages [11]. The stimulus set itself was divided into three complexity groups: 20 signs with linguistically Simple handshapes (group I), 20 signs with Moderately Complex handshapes (group II), and 20 signs with Complex handshapes (group III) (see Table 1 for illustrations).

- The Simple handshapes consisted of 5 configurations with anatomically 'easy' joint positions (maximally extended or maximally flexed) and sets of 'selected'-- i.e. prominent-- fingers (all fingers together, or the index finger or thumb). These handshapes are also most common within and across sign languages.

- The Moderately Complex handshapes consisted of 5 configurations with either 'easy' joint positions or selected finger sets, but not both. These handshapes are fairly common within and across sign languages, but less so than those in the Simple group.

- The Complex handshapes consisted of 5 configurations in which neither the joints nor the selected finger sets were anatomically simple. These handshapes are quite rare within and across sign languages.

Other linguistic parameters of the stimulus items were controlled as much as possible across the three handshape groups, given the relatively sparse phonological distribution of signs in ASL. The general location of each sign's articulation (i.e. where the hands were located during the sign's main action) was balanced across the three handshape groups: 25% of the signs in each group were located on the character's head/face, and 75% were located in the 'neutral space' in front of the character. (More specific locations, e.g. cheek vs. chin, varied only slightly across groups.) Movement types (e.g. arc-shaped, circular, straight), and the presence or absence of an orientation change at the wrist joint were also distributed as evenly as possible, differing by no more than 15% between handshape groups.

IV. EXPERIMENT DESIGN

Subjects

Data came from 53 participants (50 Deaf and 3 Hearing), ranging in age from 18 to 63 (Median: 47). All learned ASL by the age of 5, and 32 learned from birth from Deaf family members (including the 3 Hearing participants).

Stimuli

120 animation clips (60 signs per character) rendered to a resolution of 640x480 pixels using Maya software rendering algorithm; lighting setup, camera framing, and background were maintained the same in all clips. The animations were output to Quick Time movies with Sorensen3 compression and a frame rate of 24 fps. The clips were embedded in a web survey consisting of 3 screens per animated clip with a total of 360 screens (3x120). The animated sequences were presented in random order; data collection was embedded in the survey.

Procedure

Subjects were sent an email containing a brief summary of the research and its objectives, an invitation to participate in the study, and the http address of the web survey. Participants completed the on-line survey using their own computers and the survey remained active for 1 week. For each clip, subjects were asked to (1) play the animation clip and input whether they recognized or did not recognize the sign- subjects were allowed to playback the clip as many times as needed; (2) enter the word corresponding to the

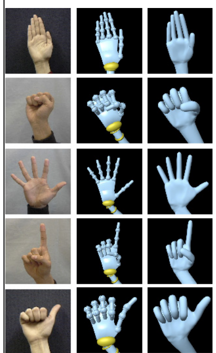
Handshape complexity groups	**ASL Sign Stimuli**
Simple	 CHILDREN, SUNDAY, NICE, ALL-DAY, SOCIETY, WORK, YEAR, SYSTEM, FIFTH, FIVE-O'CLOCK, MOTHER, FARM, ONE-YEAR-OLD, ONE, WONDER, CANDY, PAGES, SURGERY, ABLE, FAR
Moderately complex	 SECOND, TWO, KIND/TYPE, TWO-DAYS, SMOKE, HONEST, SIGN-UP, EVERY-THURSDAY, DEAF-SCHOOL, DRAW, JUICE, STRING, WISE, APPLE, CELEBRATE, TEXAS, NINE-YEARS-OLD, NINE-O'CLOCK, FRENCH-FRY, FAMILY
Complex	 DORMITORY, DEPARTMENT, DICTIONARY, DOCTOR, MONDAY, MATH, MUSEUM, EVERY-MONDAY, SIX-YEARS-OLD, SIX-O'CLOCK, WATER, WORLD, RESTROOM, REASON, RESTAURANT, ROPE, SEVENTH, SEVEN, SEVEN-WEEKS, SEVEN-DAYS

Table 1. Stimulus items divided in the three handshape complexity groups.

sign in a text box, if recognized, or leave the text box blank, if not recognized; and (3) rate the realism of the animation using a 5-point Likert scale. Figure 2 shows 4 screens extracted from the web survey. The full survey is available at: http://web.ics.purdue.edu/~mpetkov/survey.html

V. Findings

Overall, Torrents (seamless avatar) was rated higher (mean=4.18) than Robby (segmented avatar) (mean=3.95), $F(1,5734)=105.41$, $p=.000$. Overall, handshape group I (lowest complexity) was rated higher (4.11) than group II (4.05) and group III (4.05), which did not differ from each other, $F(2, 5734)=3.48$, $p=.031$. The interaction between character and handshape complexity was also significant, $F(2,5734)=4.78$, $p=.008$. For group III, the difference in rating between Torrents and Robby increased to .31, nearly double the difference for group II (.15) and 55% larger than the difference for group I (.20). For Robby, ratings decreased as handshape complexity increased.

VI. DISCUSSION AND CONCLUSION

The lower ratings for Robby as compared to Torrents indicate a general preference for seamless, deformable characters over segmented ones. While the difference in ratings between the two characters is not very evident for simple handshapes, it becomes more pronounced for moderately complex and complex handshapes. These results suggest that segmented avatars, which are easier to set up and much less computationally expensive for real time interaction, might be acceptable whenever the signed content includes primarily signs with linguistically simple handshapes. However, when the signed content includes complex handshapes, the presence of organic deformations of the character skin during motion contributes to increasing the legibility of the signs.

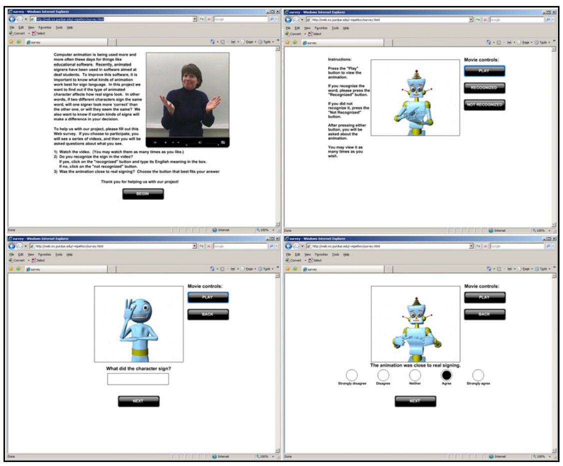


Figure 2. Four screenshots extracted from the web survey

We note that, in our study, the differences in ratings could have been influenced by the slightly different finger shapes between the two characters. Robby's fingers are thinner and have spheres placed at the joints (which is often the case for segmented avatars). The thinner fingers create additional areas of negative space in complex handshapes and the spheres can detract from the clarity of the hand silhouette; both factors may lead to misperceptions. (In simpler handshapes where fingers are maximally extended or flexed, the shape of the negative space surrounding a handshape and the hand silhouette are less important to identification.) Finger shape, as a factor, will be explored in future work.

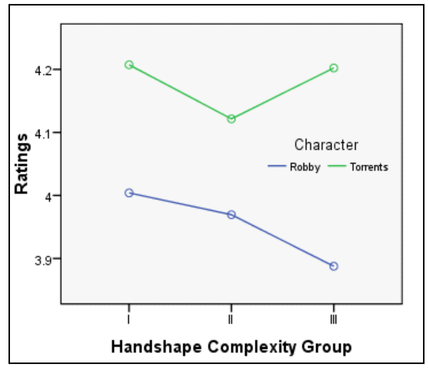


Figure 3. Mean ratings for the two characters

ACKNOWLEDGMENT

This research is supported in part by NSF-RDE grant 0622900. We are grateful to all the signers who participated in the study. Special thanks to Jason, Jill and John Lestina.

REFERENCES

[1] SIGNWRITING. Retrieved February 20th, 2009 url: http://www.signwriting.org/ .

[2] Adamo-Villani, N., Doublestein, J. & Martin Z. 2005. Sign Language for K-8 mathematics by 3D interactive animation. Journal of Educational Technology Systems, 33, 3, pp. 241-257.

[3] Adamo-Villani, N. & Wilbur, R. 2008. Two Novel Technologies for Accessible Math and Science Education. IEEE Multimedia – Special Issue on Accessibility, October-December 2008, pp.38-46.

[4] Whitney, S. 2004. Mocap ASL for the Sciences. National Science Foundation RDE-DEI: award #HRD-0435679.

[5] Vesel, J. 2005. Signing Science. Learning & Leading with Technology, 32, 8, pp. 30-31.

[6] DePaul University American Sign Language Project. Retrieved February 17th, 2009 url: http://asl.cs.depaul.edu/.

[7] Vcom3D. Retrieved February 18th, 2009 url: http://www.vcom3d.com/.

[8] Sims, E. 2000. SigningAvatars. Final Report for SBIR Phase II Project, U.S. Department of Education.

[9] eSIGN. Retrieved February 18th, 2009 url: http://www.visicast.cmp.uea.ac.uk/eSIGN/index.html.

[10] AUTODESK MAYA. Retrieved February 17th, 2009 url: http://www.resources.autodesk.com/med/Autodesk_Maya.

[11] Brentari, D. 1998. A Prosodic Model of Sign Language Phonology. Cambridge, MA: MIT Press.