



Improving realism in automated fingerspelling of American sign language

Souad Baowidan¹

Received: 24 June 2020 / Accepted: 6 June 2021 / Published online: 19 June 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract

Fingerspelling is a process of communicating letters of a spoken language alphabet using a person's hand or hands. Portraying animations of fingerspelling has proved surprisingly resistant to automation because of the collisions that arise from conventional interpolation of keyframes of individual manual letters. Previous methods have not been able to provide convincingly realistic fingerspelling due to the absence of effective collision avoidance in the underlying animation algorithms. This paper reports on the development and evaluation of a new collision avoidance algorithm that aids fingerspelling. Instead of analyzing letter transitions, the algorithm capitalizes on the transitions of individual fingers. The new strategy is efficient enough to support real-time fingerspelling while still maintaining a high level of predictive accuracy. Utilizing this strategy in signing avatars is expected to improve the current resources for deaf children, hearing teachers, hearing parents, and interpreting students who want to improve their fingerspelling comprehension. Future work will include testing the strategy's generality when applying it to other one-handed manual alphabets.

Keywords Sign language · Fingerspelling · American sign language · Collision avoidance · Sign language synthesis

1 Introduction

Fingerspelling is a process of communicating letters of a spoken language alphabet using a person's hand or hands. In American Sign Language (ASL), fingerspelling is used to spell people's names, technical terms, places without a lexical sign, and can also convey loan words from another language.

According to Calderon (2000), fingerspelling recognition is considered an important communication skill for deaf children, interpreting students, hearing teachers, and

✉ Souad Baowidan
sbaawidan@kau.edu

¹ King Abdulaziz University, Jeddah, Saudi Arabia

hearing parents. Whether used in a sign language or in a manual communication system, fingerspelling is an important communication technique for the members of the deaf and hard-of-hearing communities. This includes students requiring interpretation, parents and teachers of deaf children, as well as service providers who interact with the deaf and hard-of-hearing communities (Padden and Ramsey 1998). The potential benefits of fingerspelling skill include better deaf access to health, education, employment, and better interpreter training (Schick 2005).

Fingerspelling reception is notoriously difficult for hearing people to master (Mckee 1992), (Shipgood and Pring 1995). Fingerspelling is the first topic presented or taught during interpreter training (Padden and Gunsauls 2003). Nonetheless, it is the last topic learners master (Shaffer and Watson 2004).

Due to the challenges of acquiring fingerspelling recognition skills, most teachers who instruct deaf students and most parents of deaf children are not adequately skilled in fingerspelling, so many of them depend on interpreters. Another problem is under-qualified interpreters. It is easier to acquire fingerspelling production skills as opposed to fingerspelling reception skills (Patrie and Johnson 2011). According to Antona and Stephanidis (2015), even experienced interpreters mention fingerspelling as a top priority for more training. A survey of newly certified interpreters shows that they believe that the skill that they still need to improve is fingerspelling reception (Padden and Ramsey, 1998; Ebling et al., 2015).

Why is fingerspelling reception so hard? Basically, the reasons are classified into two major contributing factors: the nature of fingerspelling itself and the lack of practice opportunities (Antona and Stephanidis 2015; Toro et al. 2014).

The nature of fingerspelling itself acts as an obstacle to its reception. It is rare to perfectly produce the individual handshapes that comprise a fingerspelled word (Patrie and Johnson 2011). Fingerspelling is not simply formed as a series of static letters but as a smoothly changing movement in which the fingers do not stop while transitioning between letters (Wilcox 1992; Calderon 2000). Instead of being a sequence of static words, fingerspelling is a continuously flowing motion involving the constant movement of fingers, which does not pause following handshape formation (Akamatsu 1982). In fact, Akamatsu notes that when deaf children acquire fingerspelling production skills, they will mimic the motion of the fingerspelled word before they master the individual letters.

Therefore, it is difficult for signers to produce individual manual letters in a picture-perfect manner within fingerspelled words (Antona and Stephanidis 2015). Handshapes used to form manual letters are largely influenced by both succeeding and preceding letters, suggesting the presence of coarticulation. Because it is a continually flowing motion, fingerspelling involves continually changing hand movements and the signer's hand will not pause even after having formed a suitable handshape (van Zijl and Raitt 2004). Merely studying the stationary positions of manual letters does not guarantee word recognition (Antona and Stephanidis 2015).

2 Related work

The second contributing factor to the barrier to fingerspelling fluency is the lack of self-study materials and limited practical opportunities for students (Wolfe et al. 2015). According to Guillory (1966), there are numerous textbook materials recommending pair practice while learning to fingerspell. However, peer classmates frequently serve as partners during practice. Fellow students, unfortunately, cannot produce smooth fingerspelling (Antona and Stephanidis 2015). They can also not produce fingerspelling at fluent speeds (Guillory 1966). Instructors are not suitable practice partners. They have demanding schedules so face-to-face practice sessions become difficult to provide for individual students. With such barriers, automated software that improves fingerspelling reception serves as the best alternative to students for self-study (Antona and Stephanidis 2015). For these reasons, non-computerized approaches are unsuitable for providing fingerspelling practice.

Current computerized approaches for fingerspelling generation include video recordings, use of snapshots, video resynthesis, and three-dimensional animation via an avatar. Nonetheless, all of these approaches have limitations. In the 1990s, DVDs designed for practicing fingerspelling emerged (Jaklic et al. 1995), but adding new words to an extant DVD is not practical. In 2000, web-based applications appeared that uses sequences of static images (flashcards) of manual letters to spell words (Vicars 2005). The site software rearranges the images in any order and produces new words. Figure 1 an example of fingerspelling flashcards. It can spell any word, but it does not have the capability to produce the smoothly flowing motion of natural fingerspelling. Natural fingerspelling does not pause at each letter, so this method does not provide benefit to students who need to witness the flowing transitions between letters.

A third approach promises to retain the realism of prerecorded video while striving to achieve the flexibility of the flashcard approach. Stoll et al. (2018) are working to generate newly signed video sequences from previously recorded video via Neural Machine Translation. Currently, they are focusing on gross motor movement and are not considering the finer detail of fingerspelling. The work is preliminary as their generated sequences have not been evaluated by the Deaf community.

There is a fourth option, a three-dimensional avatar, that has the potential to combine the strengths of the three previously described approaches. It can portray the smoothly flowing motion of a fluent signer while offering the vocabulary



Fig. 1 Example of flashcards fingerspelling J-O-I-N

extensibility and flexibility of the flashcard approach. However, fingerspelling has proved surprisingly resistant to display by avatar because of the collisions that arise from conventional interpolation of the keyframes defining individual manual letters. Between many pairs of letters, straightforward interpolation may often lead to collisions of fingers, especially in cases of letters for which the handshape is in a closed posture. This often occurs when the thumb crosses the palm or the fingers are bent. Figure 2 shows three frames from a computer animation and three frames from a recording of a human performing fingerspelling. There are no collisions in the human performance, but a naïve computer animation does not prevent collisions. In this case, the avatar's thumb passes through, not around, the index finger.

Previously researchers attempted to address this problem through either pre-computed letter transitions or procedurally through the classification of manual letters. The pre-computed approach renders the transition between every letter pair (L_i, L_j), and stores it as a separate video clip. The clips are concatenated to produce a video of a fingerspelled word. Figure 3 demonstrates how video clips produce a fingerspelled word. The animation consists of three video clips depicting the transitions “ $T \geq U$ ”, “ $U \geq N$ ” and “ $N \geq A$ ” (Wolfe et al. 2015).

Unfortunately, prerendering every letter pair grows as the square of the number of letters in the alphabet. Even worse, this approach requires an expensive manual step where an artist inspects each letter-pair transition for collisions and adds animation keys to remove any collisions that occur. Thus, the approach does not scale well. From a complexity theory viewpoint, the approach is only processing a finite

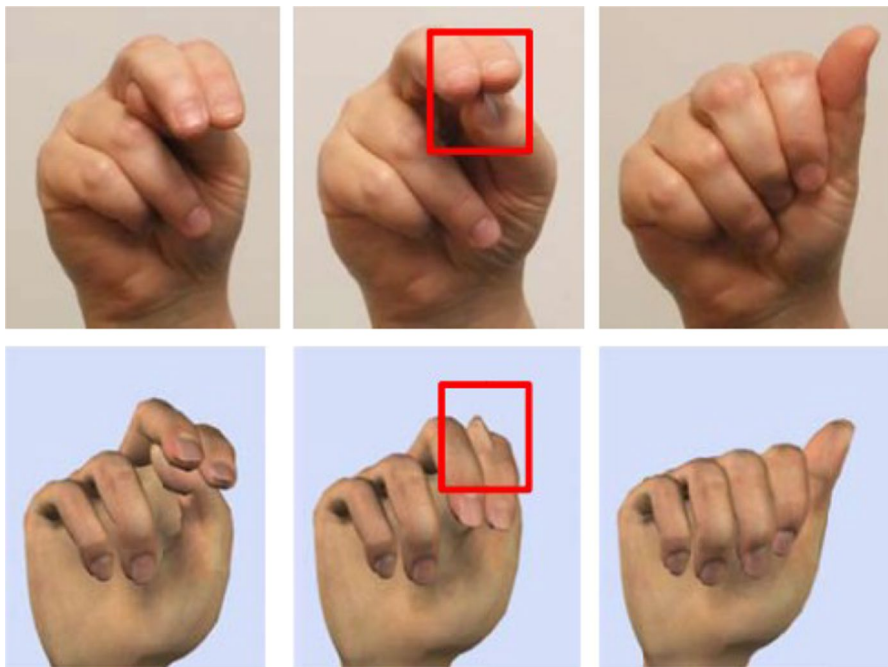


Fig. 2 Transition from N to A: comparison of an actual, virtual hand



Fig. 3 Transitions of a fingerspelling animation

number of letters, but the 26 letters of the English alphabet are at the practical limit of production. Other languages (German and Arabic) have more letters in their alphabet (Ebling et al. 2015). Because the transitions are all prerendered, there is limited flexibility in varying the fingerspelling coarticulation to simulate different fingerspelling styles. There is an additional weakness in producing repeated letters since the position of the hand changes requires more prerendering when repeating a letter. For example, to repeat the ASL letter ‘N’, the index and middle fingers are fused, held out straight, and then tapped against the thumb in a repeated motion. This necessitates an additional computation.

Based on the limitations of the previous approaches to displaying fingerspelling, there is a need for a procedurally based approach to address the collision avoidance. A procedurally based system has the potential for improved robustness. Carefully designed, a procedurally based approach could handle multiple one-hand manual alphabets without requiring the lengthy and expensive preprocessing step. One previous effort used this approach. The van Zijl and Raitt (2004) collision avoidance model for ASL used a procedural method that grouped the letters into sets and developed finite automata to govern the transitions between letters. However, this method has the same disadvantage of requiring a manual (human) analysis step, and as presented, only applied to the ASL manual alphabet. So, there was still a need for an approach that could generalize to other alphabets without requiring an initial step of manual analysis.

3 A new approach

This paper describes a new collision-avoidance approach to support the transitions between pairs of finger configurations rather than pairs of manual letters. It capitalizes on the physiology of the human hand to develop efficient and effective collision avoidance strategies automatically. The new strategy for collision avoidance is efficient enough to support real-time fingerspelling. The classification algorithm utilizes an intuitive set of geometric relationships between thumb and fingers. It forms the basis of applying an avoidance strategy. Although the current paper reports on the results of applying this method to the ASL manual alphabet, one of the advantages of using transitions between pairs of fingers instead of pairs of letters is that the approach has the potential of being language-neutral and can be applied to any one-handed manual alphabet.

4 Method

Two data sets and one corpus informed the creation of the new collision avoidance strategy. The data sets are from the data used to animate the American Sign Language Avatar at DePaul University and have undergone numerous evaluations with members of the Deaf community in the United States (Davidson 2000). The research could have used data sets from any signed language for this step, but at present the necessary data was only available in ASL. Descriptions of the two data sets are in the list that follows:

1. A data set of linguistic descriptors that have geometric interpretations as letters in the ASL manual alphabet (henceforth called “LettersDataset”). This data formed the basis for creating an animation containing a sequence of all $26^2 = 676$ transitions between two letters of the ASL manual alphabet. The animation is nothing more than interpolations through each letter and is unnatural in appearance because it contains many collisions. This gives us a baseline animation that proves useful for the analysis of collision detection. This animation is called “BaseAnimation”.
2. A data set of collision corrections (henceforth “CorrectionsDataset”). This data was created by artists who manually identified two-letter transitions containing collisions. They then added or modified the animation keys of the transition to remove collisions. This data is stored in the CorrectionsDataset. This data set is indexed by the two-letter transition containing a collision, such as “AN”, “AB”, etc. There may be multiple records for a single two-letter transition because multiple joints may require adjustments to avoid collisions. Each record has the following fields:
 1. The two-letter transition is containing a collision, as mentioned above.
 2. The name of the joint that needs to change the path to avoid a collision.
 3. Rotation key applied to the joint to avoid the collision.

When applied to the BaseAnimation, the motions in the CorrectionsDataset yield animations that have been consistently judged by Deaf and hearing observers as being error-free (Wolfe et al. 2006) and natural in motion (Ebling et al. 2015).

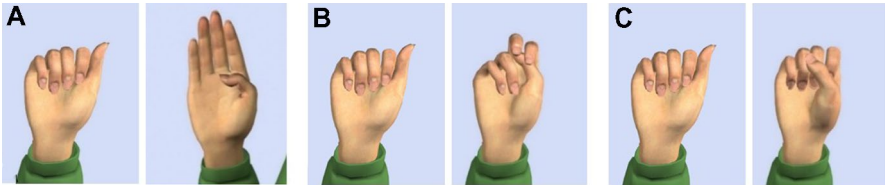
In addition to the two data sets, the method also used a corpus of previously-recorded videos. This corpus consists of fingerspelling sessions previously recorded and annotated at DePaul University. The goal of the annotations was to provide data to study potential collisions and the strategies that human takes to avoid finger collisions (Baowidan et al. 2017).

4.1 Preparatory data analysis

Instead of analyzing interactions between letters as a whole, the analysis focused on the interactions between individual fingers and the thumb. Reviewing the CorrectionsDataset reveals the two-letter transitions that created collisions and noted where

Table 1 Collision between individual fingers and the thumb

Finger	Collisions
Index	85
Middle	126
Ring	42
Pinky	3

**Fig. 4** **A** Radial \geq Ulnar + Above; **B** Radial \geq Ulnar + Under; **C** Radial \geq Ulnar + Over

the collisions took place. Since the index is the finger closest to the thumb and has the second-highest collision rate with the thumb (Table 1), I focused on the relation between the thumb and the index finger and then applied the results on the other three fingers. Although the thumb touches the pinky, it does not go past it the way that it can cross by the other three fingers. The angles in the LettersDataset determined the position of the thumb tip and the index finger's tip.

4.2 Classification scheme

Focusing on the thumb and index finger, several patterns emerged. Collisions between these two digits occur in three basic cases. For simplicity, assume the thumb tip is radial to the index finger in letter L_i :

- A. Ulnar + Above: Letter L_{i+1} has the thumb tip located ulnar to the index finger. The transition from A to B is an example. See Fig. 4A. In the DePaul avatar system, the coordinates used for Ulnar testing are the x-coordinates of the thumb tip and the index finger distal interphalangeal joint (DIP) joint in palm coordinates. The Above test compares the z-coordinate of the thumb tip and the index finger DIP in palm coordinates.
- B. Ulnar + Under: Letter L_{i+1} has the thumb tip located ulnar to the index finger and underneath it. The transition from A to N is an example. See Fig. 4B. In the DePaul avatar system, the Under test uses the y-coordinates of the thumb tip and index DIP joint in palm coordinates.
- C. Ulnar + Over: Letter L_{i+1} has the thumb tip located ulnar to the index

finger and covering it. The transition from A to S is an example. See Fig. 4C. The Over test uses the same information as the Under test.

4.3 Basic motions for collision avoidance

Studies of the CorrectionsDataset of manually created collision avoidance animations (source 2) and the annotated corpus (source 3) yields the following three basic motions for avoiding collisions:

1. **Thumb Delay:** It may be sufficient to delay the thumb's movement until the fingers have moved out of the thumb's path. The thumb will delay its movement to make way for the index finger movement. This basic motion will avoid a collision between the thumb and index finger for any letters where the thumb begins on the radial side of the hand and ends on the ulnar side of an extended index finger such as A to B, A to V, or L to H.
2. **Finger Flip:** To accommodate the progression of the thumb to an "under" position, the base joints of the covering fingers will rotate upward before moving into their final position. For example, consider the transition of the thumb as it moves from the radial side of the hand (as in the letter A) to a position that is under and to the ulnar side of the index finger as in the letter M. The index finger will straighten to make way for the thumb movement. This will avoid a collision between the thumb and index finger for any letters where the thumb begins on the radial side of the hand and ends on the ulnar side of the index finger, such as A to T, A to N, L to M.
3. **Thumb Swing:** To accommodate the progression of the thumb to an "over" position, it will need to move outwards from the palm to allow the fingers to move into their final position. For example, consider the transition of the thumb as it moves from the radial side of the hand (as in the letter A) to a position that is ulnar and covering the index finger as in the letter S. The thumb will move outward of the palm, so it does not collide with the index finger. This will work for avoiding a collision between the thumb and index finger for any letters where the thumb begins on the radial side of the hand and ends on the ulnar side of the index finger and covering it such as A to S, A to I.

Each of these basic actions adds intermediate keys between the two hand poses at letters L_i and L_{i+1} , similar to those that were manually added by the artist in the CorrectionsDataset database. The Thumb Delay function adds intermediate keys that are replicas of the thumb keys of letter L_i . The intermediate keys added by the Finger Flip function straighten the (index) finger. The Thumb Swing function adds keys that slightly straighten the thumb's proximal and distal interphalangeal joints and rotates the thumb's metacarpophalangeal joint (Figs. 5, 6, 7).

4.4 Combining basic motions

Using a single basic motion will address the transition types listed in the last section but will not solve the following three situations:

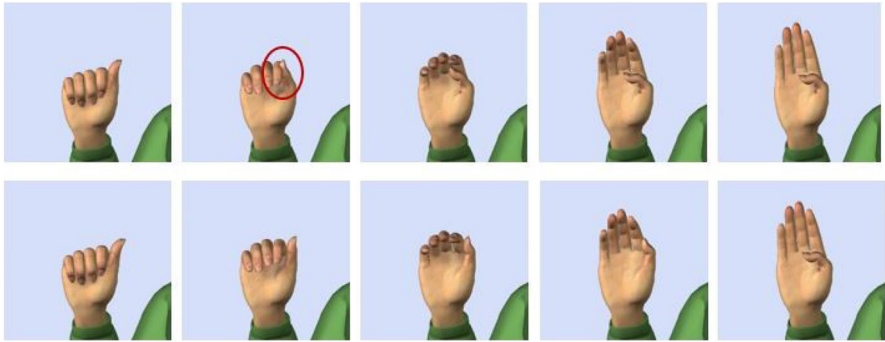


Fig. 5 Avoiding collision via thumb delay ($\text{Radial} \geq \text{Ulnar} + \text{Above}$)

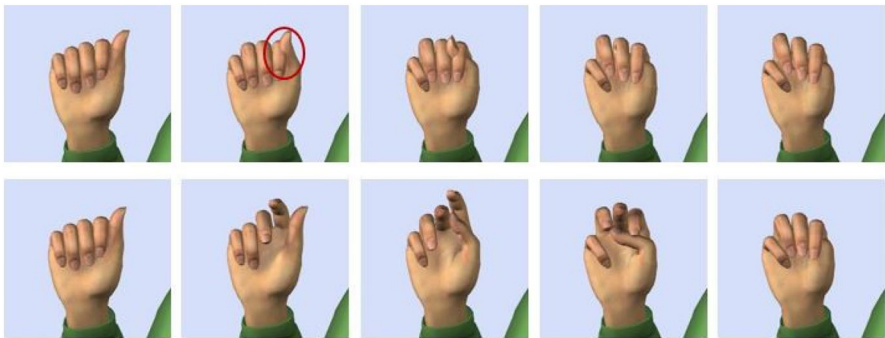


Fig. 6 Avoiding collision via finger flip ($\text{Radial} \geq \text{Ulnar} + \text{Under}$)

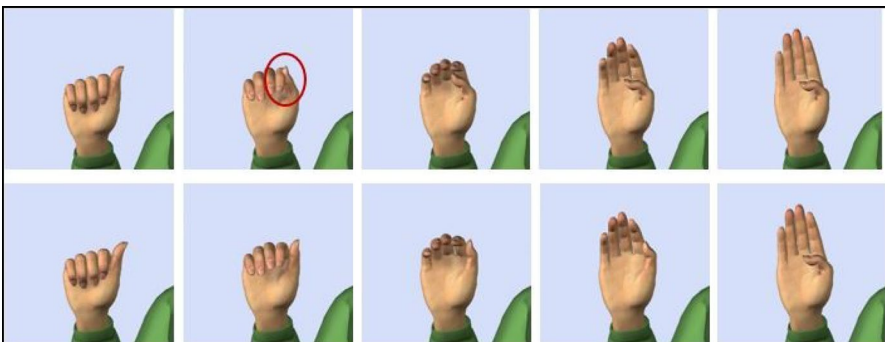


Fig. 7 Avoiding collision via thumb swing ($\text{Radial} \geq \text{Ulnar} + \text{Over}$)

1. The transition is from Ulnar + Above to Radial. This is simply the reverse of the Radial to Ulnar + Above transition. Instead of being delayed, the thumb must speed up. See Fig. 8. The keys inserted on the thumb between letters L_i and L_{i+1} now duplicate the thumb keys of L_{i+1} instead of L_i .
2. The transition from Ulnar + Under to Ulnar + Over. An example is M to



Fig. 8 Avoiding collision through thumb speed ($Ulnar + Above \geq Radial$)

S. Collisions of this type can be avoided by applying the basic motion Finger Flip followed by Thumb Swing. Figure 9 demonstrates the resulting motion.

3. The transition from Ulnar + Over to Ulnar + Under. This is the reverse of the previous case. An example is S to M. Collisions of this type can be avoided by applying Thumb Swing followed by Finger Flip. See Fig. 10.

Detecting collisions between the thumb and the middle finger, and collisions between the thumb and the ring finger follow a similar pattern. Instead of using the position of the index DIP joint and the thumb tip, the detection system for the thumb and middle finger utilizes the positions of the middle fingertip and the thumb tip. The difference in the finger joint comes from the shift to the hand's radial side that occurs when the middle finger curls toward the palm (McDonald et al. 2001). Applying the approach to the thumb and ring finger is similar.

This approach takes advantage of the physical limitations of the human hand, so the number of cases is manageable. The thumb is capable of crossing over only the index, middle, and ring fingers. Although outside the scope of this study, it is worth noting that few crossings between fingers are physically possible and are limited to



Fig. 9 Avoiding collision through finger flip followed by thumb swing ($Ulnar + Under \geq Ulnar + Over$)



Fig. 10 Avoiding collision through thumb swing followed by finger flip ($Ulnar + Over \geq Ulnar + Under$)

adjacent fingers. Since the new approach is only considering the physiology of the hand, it is handshape/alphabet independent. The approach is sufficiently general that it should be able to successfully avoid thumb-finger collisions in any one-handed manual alphabet.

The algorithm was prototyped using Maxscript in the 3d animation package 3dsmax (Murdock 2011) to classify the finger transitions and generate the fingerspelling animations. The new approach classifies the relationship between the two digits and takes evasive action as laid out in Table 2. The basic motions for avoiding a collision are simple and straightforward. Their effectiveness lies in predicting the need to apply them.

A video depicting these basic avoidance motions is available at <http://sltat.cs.dep-aul.edu/2019/baowidan.mp4>

5 Validation

5.1 Data source

- new approach (my algorithm)
- "conventional (automatic detection, Foley and Van Dam 1982).
- manual (artist visual inspections of animations for collisions/CorrectionsDataset)

Validation was not straightforward because there was no definitive ground truth. Instead, two sources of data were used to evaluate the new algorithm's predictive power. The first data set was generated by software written in Maxscript using a conventional Sutherland-Hodgman algorithm collision detection algorithm (Foley and Van Dam 1982). It tested transitions between letter pairs $\{(L_i, L_j) \mid L_i, L_j \in \{\text{Manual alphabet}\} \& L_i \neq L_j\}$ stored in the BaseAnimation file to identify the index finger transitions containing collisions with the thumb. When the software detected a collision, it added the letter pair to a list. The second source of data was already available in the Corrections- Dataset. This data set contains the animations previously created manually by artists to avoid collisions. Access to individual collision avoidance

Table 2 summarizes the collision avoidance necessary for each type of transition between the thumb and the index finger

	Radial	Ulnar + Above	Ulnar + Under	Ulnar + Over
Radial	No action	Delay thumb	Finger flip	Thumb swing
Ulnar + Above	Speed thumb	No action	No action	Thumb swing
Ulnar + Under	Finger flip	No action	Finger flip	Finger flip, then thumb swing
Ulnar + Over	Thumb swing	Thumb swing	Thumb swing, then finger flip	No action

The "Ulnar+Under" to "Ulnar+Under" case refers to transitions such as M to N. The label "No action" means that there will not be a collision, so no action is necessary

Table 3 Confusion matrix of conventional approach vs. manual approach

	Conventional: no collision	Manual: no collision	Row total
Manual: collision	78	178	256
Conventional: collision	8	386	394
Column total	86	564	650

Table 4 Manual approach type I errors

Letter	Count	Transitions
G	17	GA GB GD GE GF GI GJ GK GL GM GO GP GR GV GW GX GY
H	6	HA HL HN HP HT HY
P	13	PA PC PD PG PH PJ PL PO PQ PS PT PX PY
Q	9	QA QF QJ QK QL QM QI QP QX

animations in the CorrectionsDataset is through the letter pair (L_i, L_j) requiring the avoidance animation in its transition. A first computation of a confusion matrix compared the list of letter pairs identified by the new approach to the list generated by conventional detection, and a second computation of a confusion matrix compared the list of letter pairs of the new approach to the list of letter pairs stored in the CorrectionsDataset. The results of the two comparisons were sufficiently inconsistent as to warrant a comparison of the manual approach to the conventional approach.

5.2 Examining the data sources

The next step was to evaluate the consistency between the list of letter pairs generated by the conventional collision detection algorithm, and the list of letter pairs retrieved from the artist's manually created collision avoidance animation. The next computation of a confusion matrix compared the list of letter pairs generated by the conventional (automatic) approach to the list of letter pairs of the manually created avoidance animation. The confusion matrix is in Table 3. Again, the manual collision method is the approach that requires artists to visually inspect animations for collisions. It is notable that the number of collisions detected by the conventional detection method is nearly three times greater than the number of collisions detected by the manual method. In this confusion matrix, a Type I error corresponds to the case where the manual method predicted no collision, but the conventional method found a collision. This disparity warranted further analysis since the manual approach had been vetted in several evaluation studies involving users fluent in ASL.

A deeper investigation reveals that many of these Type I errors occurred in transitions involving the letters G, H, P, and Q, as can be seen in Table 4. When a signer produces these letters, the palm is either facing downwards (P, Q) or inwards toward the body (G, H). The artists missed these collisions. An additional set of Type I errors stems from the difference between the properties of

human flesh and the properties of polygon meshes. Polygons are rigid objects. Human muscle and skin are flexible. Figure 11 demonstrates the difference. Both pairs of images demonstrate the thumb/finger interaction that occurs when producing the letter F. When humans touch a thumb to an index finger, the pads of both digits flatten, but do not intersect. When a polygon mesh assumes an analogous position, the pad of the index finger will move through the surface of the thumb pad. Such a position will be recognized as a collision by a conventional collision detection algorithm.

There is one additional source of Type I errors. A further inspection of the geometry reveals that the poses for A, D, E, I, J, S, and Z have collisions that have been deemed undetectable by human test participants but are identified as collisions in the conventional method see Fig. 12.

In the confusion matrix, a Type II error corresponds to the case where the manual approach predicted a collision, but the conventional method did not detect one. In this case, the manual approach will be applying one of the evasive motions to avoid a collision that is not there.

The manual approach was previously judged as perceptually correct by human test participants. The conventional approach identifies collisions algorithmically through the application of mathematics to the geometric representation of the hand. However, the two approaches, deemed correct in their own context, do not produce consistent results due to the method used to detect collisions. This knowledge will help set the context for examining the performance of the new approach.

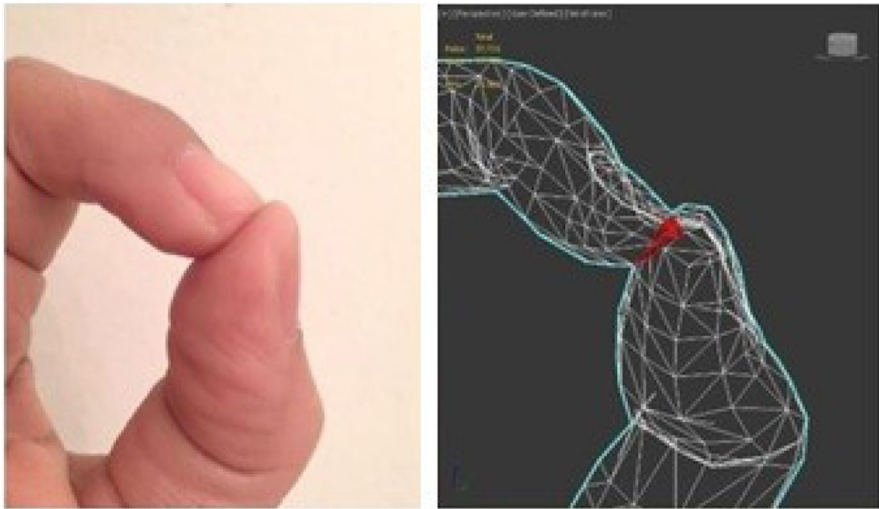


Fig. 11 Difference between the properties of human flesh and the properties of polygon meshes

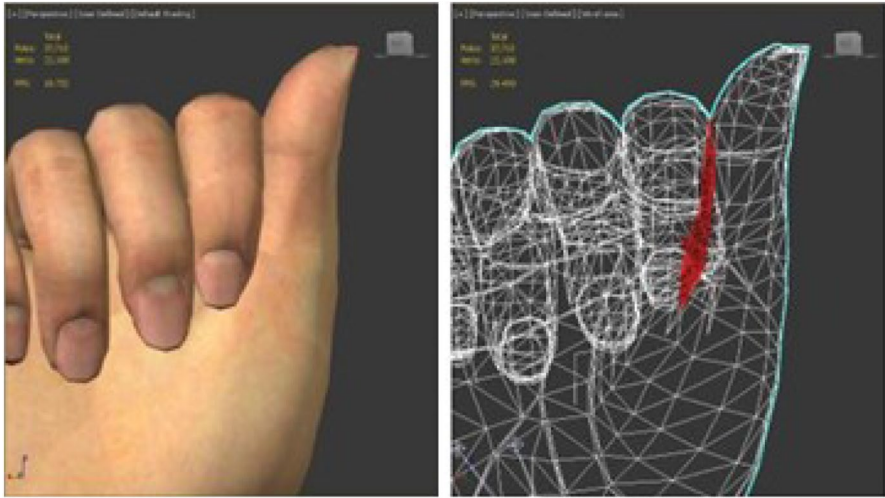


Fig. 12 Front close-in of letter A, in default shading and in wireframe

6 Effectiveness of the new approach

Table 5 contains the confusion matrix for the new approach as compared to the conventional approach. Type I errors correspond to the cases that are labeled by the conventional method as collision but were predicted by the new method as no collision. Many of these cases are caused by several letters (A, D, I, and J) that have a collision built into them, which was detected by the conventional method. The new approach did not classify them as a collision see Fig. 13.

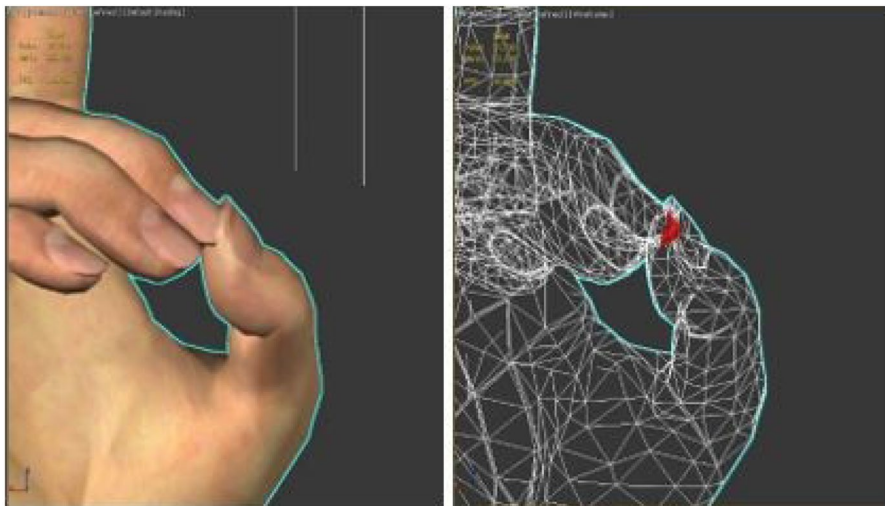


Fig. 13 Side close-in of letter D, in default shading and in wireframe

Table 5 Confusion matrix of conventional approach vs. new approach

	New approach: collision	New approach: no collision	Row total
Conventional: collision	209	47	256
Conventional: no collision	194	200	394
Column total	403	247	650

In this analysis, Type II errors correspond to the cases where the new approach predicts a collision, but the conventional approach says that there is no collision. This means that basic motions may be applied where there is no need to. Adding evasive motion where none is necessary can be as distracting to a viewer as a collision. See Table 6 for a list of these cases.

However, a deeper analysis of the Type II errors yields the following insights: of the 194 Type II errors, the majority (184 cases) are collisions classified as requiring a change in the thumb speed. Only 10 cases were classified as needing a Finger Flip. In Table 7, the first and second rows show the cases that contain transition from/to letter M or N. These two-letter transitions involve the thumb moving to/from a position underneath fingers which naturally requires flipping fingers as a human signer would produce it. In the transition from S to Q, the thumb needs to swing outwards first to allow the index to move to its final position. In the Q to Y transition, there is a slight brushing between the thumb and index finger as the index finger moves from an extended to a curled position.

In contrast to a finger flip, a change in thumb speed is not adding motion but is a perceptually subtle change, because only the timing of the thumb changes, not the positioning of the fingers.

Table 6 The 10 cases which were classified as index finger needing a Finger Flip

M ≥ E
M ≥ N
M ≥ T
T ≥ M
N ≥ E
N ≥ M
N ≥ T
E ≥ N
S ≥ Q
Q ≥ Y

Table 7 Confusion matrix of manual approach vs. new approach

	Manual: collision	New approach: no collision	Row total
New approach: collision	78	7	85
Manual: no Collision	325	240	565
Column total	403	247	650

Table 7 contains the confusion matrix comparing the new approach to the manual approach. Not surprisingly, these results are quite different from those in the confusion matrix comparing the new approach to the conventional approach.

The 7 cases of Type I error were checked closely. They fall into one of two categories: either above to covered or above to above. The two categories do not need correction even if they were marked as collisions. See Table 8. As expected, the number of Type II errors is much higher because there are fewer collisions detected overall with the manual method. Human artists visually detected collisions and missed those poses where the fingers were not facing the viewer's perspective, for instance, transitions involving letters H, G, P, and Q.

7 Discussion

There are challenges involved in analyzing the performance of the new approach, based on the fact that there is not an objectively perfect representation of actual conditions. The first basis for comparison is the conventional approach based on the avatar's hand geometry which does not accurately model the behavior of human muscle and skin. As demonstrated, the conventional geometry approach identifies multiple hand poses as having collisions where a human viewer has judged that none exist. In contrast, the second basis for comparison, the manually classified collisions, lacked a large number of collisions because the artists simply could not see them.

When compared with the conventional approach, the new approach has an accuracy of $(200 + 209)/650 = 63$ percent, and a misclassification rate of $(47 + 194)/650 = 37$ percent. However, a further examination of the 47 type I errors shows that although the conventional algorithm detected collisions in 44 of them, user studies have previously demonstrated that none of the 44 are detected as intersecting from a perceptual basis. Although there are 194 type II errors, only ten transitions predicted to collide by the new method will cause the addition of extraneous motion of finger flipping.

When compared to the manual approach, the new approach has an accuracy of $(78 + 240)/650 = 49$ percent. Of the misclassifications, the vast majority of them (325) are type II errors, where the new algorithm predicted a collision, but the manual method did not. This is attributable to a large number of collisions that were simply missed by the artists who visually inspected animations for the collision.

A compelling advantage of this new approach is that it is automated. It does not require manual intervention by artists to painstakingly identify each collision by inspection and design an animation to avoid it. A second compelling advantage is its

Table 8 Type I error transitions, classification, and avoidance actions

Transitions	Classification	Evasive action required
CT GN GT XN XT	Ulnar + Above to Ulnar + Over	No action
DR RD	Ulnar + Above to Ulnar + Above	No action

speed. Its calculation requires a maximum of 12 floating-point operations (flops) per finger in a manual letter, and this cost is spread out over the frames of a transition. For example, if the avatar is fingerspelling at a pace of three letters per second with a refresh rate of 30 frames/second, the computational cost per frame is $12 \text{ flop/finger} \times 4 \text{ fingers} = 48 \text{ flops}$ distributed over ten frames for an average per-frame cost of under five flops. Thus, it is sufficiently efficient to run in real-time.

8 Conclusion

This work has focused on implementing a real-time collision detection and avoidance algorithm for fingerspelling animation. The new algorithm can be integrated into a real-time avatar, suitable for use in fingerspelling learning and practice tools for interpreting students who are studying sign language. The algorithm was designed to be language agnostic. Because the algorithm tests transitions between pairs of fingers instead of pairs of letters, the hope is that it can accommodate any one-handed manual alphabet. The new approach is unlike the old video based or pre-rendered approach, which only had one view. It requires minimal computing resources to change an avatar from front to other perspectives. outcome of this work is expected to improve the current resources for deaf children, hearing teachers, hearing parents, and interpreting students for self-study and open the possibility of better fingerspelling comprehension. Future work will include creating a real-time platform to implement the algorithm and to evaluate its performance on other manual alphabets.

9 Future work

This study considered only the thumb interaction with the other fingers. To complete this approach requires one step beyond the scope of the current research. Future work also includes controlling the speed of the thumb delay and the size of the finger flip. Future researchers should incorporate the linguistic processes of coarticulation and deletion into this new collision avoidance approach for fingerspelling. This can help create different fingerspelling styles and allow students to watch/read different varieties of fingerspelling. This is analogous to viewing different styles of handwriting or listening to different speakers of a spoken language.

Moreover, future studies should investigate the viability of the collision avoidance algorithm in two-handed manual alphabets and other languages than ASL.

References

- Akamatsu CT (1982) The acquisition of fingerspelling in pre-school children. University Microfilms, Ann Arbor
- Antona M, Stephanidis C (2015) Universal access in human-computer interaction. In: Access to learning, health and well-being: 9th International Conference, UAHCI 2015, Held as Part of HCI International

- Baowidan S, Guo N, Johnson S, Moncrief R, Berke L (2017) A New N-gram analytics tool in ELAN and its application to improve automatic fingerspelling generation. DePaul University, Chicago
- Calderon R (2000) Parental involvement in deaf children's education programs as a predictor of child's language, early reading, and social-emotional development. *J Deaf Stud Deaf Educ* 5(2):140–155. <https://doi.org/10.1093/deafed/5.2.140>
- Davidson MJ (2000) Usability testing of computer animation of finger-spelling for American sign language. In: and others (ed) CTI Research Conference
- Ebling S, Wolfe R, et al. (2015) Synthesizing the finger alphabet of Swiss German Sign Language and evaluating the comprehensibility of the resulting animations. *Proceedings of SLPAT 2015: 6th Workshop on Speech and Language Processing for Assistive Technologies* pp 10–16
- Foley JD, van Dam A (1982) *Fundamentals of interactive computer graphics*, vol 2. Addison-Wesley Reading, MA
- Guillory LM (1966) Claitors publishing division
- Jaklic A, Vodopivec D, Komac V (1995) Learning sign language through multimedia. In: *International Conference on Multimedia Computing and Systems*, IEEE, pp 282–285
- Mcdonald J, Alkoby K, Carter R, Christopher J, Davidson M, Ethridge D, Wolfe R (2001) An improved articulated model of the human hand. *Vis Comput* 17(3):158–166
- Mckee RL (1992) What's so hard about learning ASL? Students' and teachers' perceptions. *Sign Language Studies* 75(1):129–157
- Murdock K (2011) *3Ds Max 2012 bible*, vol 783. John Wiley & Sons, New York
- Padden C, Gunksauls DC (2003) How the alphabet came to be used in a sign language. *Sign Language Studies* 4(1):10–33. <https://doi.org/10.1353/sls.2003.0026>
- Padden C, Ramsey C (1998) Reading ability in signing deaf children. *Top Lang Disord* 18:30–46
- Patric C, Johnson R (2011) *Fingerspelled word recognition through rapid serial visual presentation*. DawnSignPress, San Diego
- Schick B (2005) Look who's being left behind: educational interpreters and access to education for deaf and hard-of-hearing students. *J Deaf Stud Deaf Educ* 11(1):3–20. <https://doi.org/10.1093/deafed/enj007>
- Shaffer L, Watson W (2004) Peer mentoring: what is that. *Proceedings of the 15th National Convention Conference of Interpreter Trainers (CIT)*, Still shining after 25:77–92
- Shipgood LE, Pring TR (1995) The difficulties of learning fingerspelling: an experimental investigation with hearing adult learners. *Inter J Nal Lang Commun Disorders* 30(4):401–416. <https://doi.org/10.3109/13682829509087241>
- Stoll S, Camgoz NC, Hadfield S, Bowden R (2018) Sign language production using neural machine translation and generative adversarial networks. *Proceedings of the 29th British Machine Vision Conference*
- Toro JA, Mcdonald JC, Wolfe R (2014) Fostering better deaf/hearing communication through a novel mobile app for fingerspelling. In: and others (ed) *Computers helping people with special needs*, Springer International, pp 559–564
- Vicars B (2005) URL <http://asl.ms>
- Wilcox S (1992) *The phonetics of fingerspelling*. John Benjamins Publishing, Amsterdam
- Wolfe R, Alba N, Billups S, Davidson MJ, Dwyer C, Jamrozik D, (2006) An improved tool for practicing fingerspelling recognition. In: *Conference 2006 international conference on technology and persons with disabilities*, pp 17–22
- Wolfe R, Mcdonald J, Toro J, Baowidan S, Moncrief R, Schnepf J (2015) Promoting better deaf/hearing communication through an improved interaction design for fingerspelling practice. *International conference on universal access in human-computer interaction*, pp 495–505
- van Zijl L, Raitt L (2004) Implementation experience with collision avoidance in signing avatars. *Proceedings of the 3rd international conference on Computer graphics, virtual reality, visualisation and interaction in Africa*, pp 55–59

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.