Chapter 1

Facial Expressions

Facial expressions are a crucial element of communication in both spoken and ??s. In ??s, they serve a dual purpose: conveying the emotional state of the signer and encoding grammatical information such as negation and emphasis. While spoken languages rely on tone and intonation for these functions, ??s depend heavily on ??, particularly facial expressions, to fully convey the meaning of an utterance. Therefore, creating accurate and expressive facial animations in signing avatars is essential for producing realistic and comprehensible ?? content.

However, synthesizing facial expressions for signing avatars presents significant challenges due to the complexity and subtlety of these expressions. Unlike manual signs, which involve distinct hand and arm movements, facial expressions require the coordinated movement of numerous facial muscles, each contributing to the overall expression. Different facial features, such as eyebrows, eyes, and mouth, often move independently yet in harmony to produce coherent expressions. Previous chapters focused on the synthesis of skeletal features in ??. In contrast, this chapter centers on the synthesis of facial expressions from the AZee model, emphasizing their critical role in the conveyance of grammatical information in ??s.

The chapter is structured as follows: Section **??** describes the process of extending the AZee model to include facial expressions. Section **??** details the creation of blendshapes for facial expressions, while Section **??** explains how motion curves are used to animate these blendshapes. Finally, Section **??** presents the results of the facial expression synthesis and evaluates the accuracy and expressiveness of the synthesized expressions.

1.1 Defining facial expressions in AZee

The AZee model does have a specification for *morph* constraints. However, a proper specification for facial expressions was missing. Thus, our first tasks was to extend the set of *morphs* in the AZee. We used the 40 brèves corpus **??** as a reference for our study. The 40 brèves corpus consists of 23 facial expressions in the form of AZee rules (see section **??** for number of occurrences).

We discussed the ?? earlier in section **??**. ?? is a comprehensive system for describing facial expressions in terms of action units (AUs), which represent the activation of specific facial muscles. AUs can also be combined to create facial expressions. Thus, our first goal is to break down each AZee rule into its constituent AUs.

We use the AU detector by **?** on mediapipe extractions of still images for each of the facial expressions from the corpus (figure **??**).

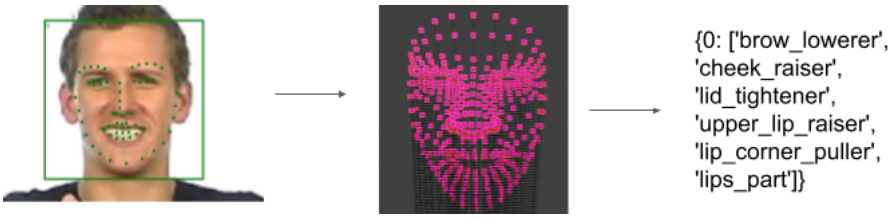


Figure 1.1: Facial action unit detection using mediapipe and the AU detector by **?** for *big-threatening*

While it is a good starting point, manual adjustments were required by a linguist to ensure accuracy. This can be seen in figure **??** for the expression *big-threatening*.

In total, we extended the low-level AZee language morph set with 94 morphs corresponding to most of the action units in the ?? system (figure **??**) along with some additional blendshapes for the tongue. A few action units were not added **??** since they were controlled by other skeletal constraints. Some morphs are also alternative blendshapes for the same action units (figure **??**) but with a different effect on the face.

1.2 Modeling Blendshapes for Facial Expressions

Once the AUs were identified, we used FACSHuman **?** blendshapes as reference to create shapekeys on our BAZeel avatar that correspond to each AU (figure **??**). Shape keys (figure **??**) are essentially different versions of a 3D model, each rep-

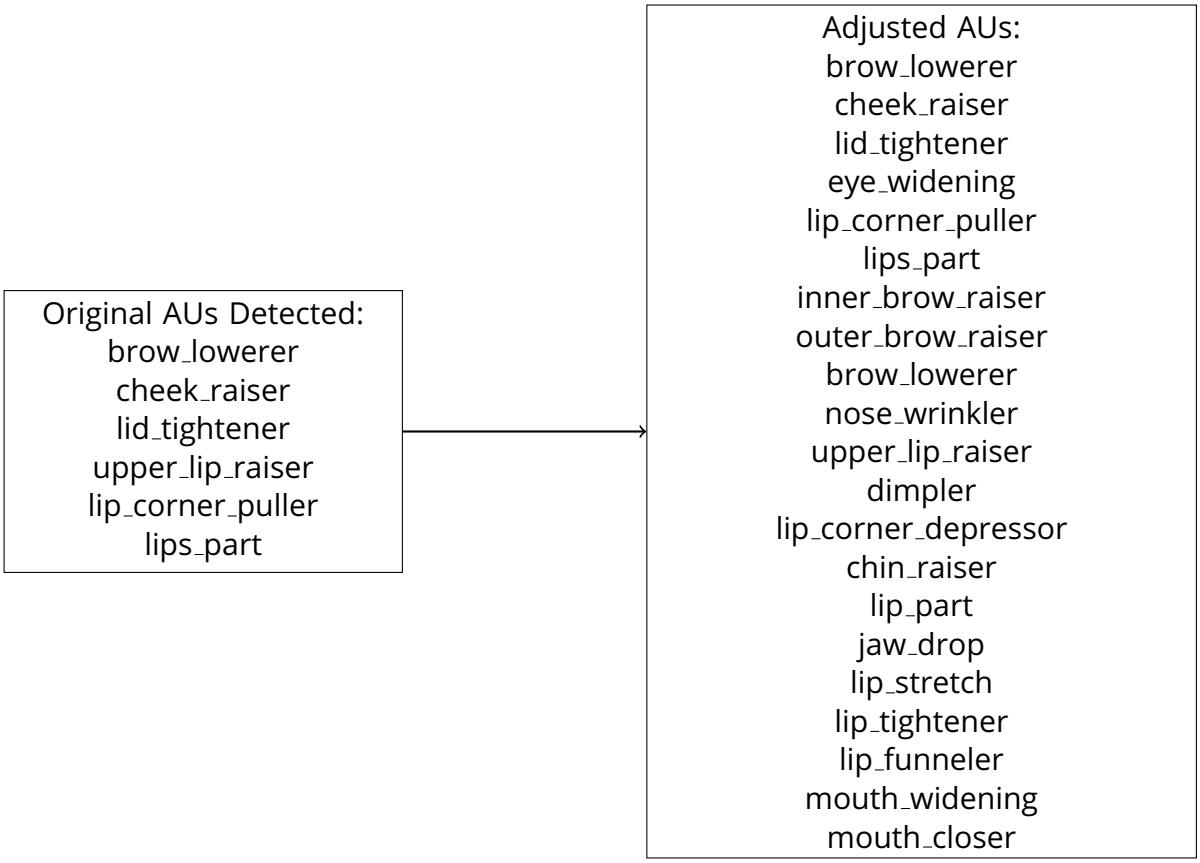


Figure 1.2: Manual adjustment for *big-threatening*

resenting a specific change in vertices. By blending and interpolating these keys together, we can create a wide range of mesh shapes.

Since the FACSHuman blendshapes work on any MakeHuman template mesh, this technique of synthesis is avatar-independant. However, since the original model did not cover all the blendshapes for every facial mesh, some were created manually. The creation of blendshapes also involved ensuring that they could be seamlessly combined to create complex expressions. For example, the blendshape for AU4 (Brow Lowerer) was designed to work in conjunction with AU6 (Cheek Raiser) and AU12 (Lip Corner Puller) to create expressions of anger or determination.

Let V be the set of vertices of the 3D facial mesh, where $V = \{v_1, v_2, \ldots, v_n\}$, and each vertex v_i has a position in 3D space $v_i = (x_i, y_i, z_i)$. Each facial Action Unit (AU) A_j modifies the positions of the vertices based on predefined shape key transformations.

Shape Key Definition For each AU A_j, let S_j be the shape key associated with that AU. S_j defines a vector of vertex displacements for extreme positions. Specifically, for a vertex v_i, the shape key specifies a displacement $\Delta v_i^j = (\Delta x_i^j, \Delta y_i^j, \Delta z_i^j)$, where:

$$\Delta v_i^j = (x_i^j - x_i, y_i^j - y_i, z_i^j - z_i)$$

Here, (x_i^j, y_i^j, z_i^j) are the extreme position coordinates of vertex v_i under AU A_j.

AU Activation Let $\alpha_j \in [0, 1]$ be the activation level of AU A_j, where:

- $\alpha_j = 0$ means the AU is not activated (neutral position),
- $\alpha_j = 1$ means the AU is fully activated (extreme position).

Vertex Position Update For a given activation level α_j, the new position $v_i' = (x_i', y_i', z_i')$ of vertex v_i under the influence of AU A_j is computed as:

$$v_i' = v_i + \alpha_j \Delta v_i^j$$

This can be expanded as:

$$x_i' = x_i + \alpha_j \Delta x_i^j$$

$$y_i' = y_i + \alpha_j \Delta y_i^j$$

$$z_i' = z_i + \alpha_j \Delta z_i^j$$

1.2.1 Combined Effect of Multiple AUs

If multiple AUs $\{A_1, A_2, \ldots, A_k\}$ are active simultaneously, the final position of vertex v_i is determined by the weighted sum of the displacements for each AU:

$$v_i' = v_i + \sum_{j=1}^{k} \alpha_j \Delta v_i^j$$

Figure 1.3: FACSHuman blendshapes used as reference for creating shapekeys in blender.

Figure 1.4: Blender interface. (a) Shape Key properties (b) 3D Viewport (c) Non-linear Editor (d) Action Editor (e) AZee editor

1.3 Motion Curves for Blendshapes

For shape keys, the motion curves (discussed earlier in section **??** of chapter **??**) represent time in x axis, and the weight of the shape key in the y axis. Each shape key has one motion curve. Figure **??** shows how motion curves can be used for shape keys for the facial expressions of the above example.

After creating the blendshapes, we used motion curve based templates (as discussed earlier in chapter **??**) to control how these shapes are animated over time. Motion curves define the changes in the blendshape's influence over the course of an animation, allowing for smooth transitions between different facial expressions.

We used the still images from the corpus as the apex of our curves (figure **??**). Next, we extended our existing intermediate block generation algorithm to in-

AU Code	Description	AU Code	Description
AU1	Inner Brow Raise	AU1_L	Inner Brow Raise (Left)
AU1_R	Inner Brow Raise (Right)	AU2	Outer Brow Raise
AU2_L	Outer Brow Raise (Left)	AU2_R	Outer Brow Raise (Right)
AU4	Brow Lowerer	AU5	Upper Lid Raise
AU5_L	Upper Lid Raise (Left)	AU5_R	Upper Lid Raise (Right)
AU6	Cheek Raise	AU6_L	Cheek Raise (Left)
AU6_R	Cheek Raise (Right)	AU7	Lids Tight
AU7_L	Lids Tight (Left)	AU7_R	Lids Tight (Right)
AU8	Lips Toward Each Other	AU9	Nose Wrinkle
AU9_L	Nose Wrinkle (Left)	AU9_R	Nose Wrinkle (Right)
AU10	Upper Lip Raiser	AU10_L	Upper Lip Raiser (Left)
AU10_R	Upper Lip Raiser (Right)	AU11	Nasolabial Furrow Deepener
AU11_L	Nasolabial Furrow Deepener (Left)	AU11_R	Nasolabial Furrow Deepener (Right)
AU12	Lip Corner Puller	AU12_L	Lip Corner Puller (Left)
AU12_R	Lip Corner Puller (Right)	AU13	Sharp Lip Puller
AU13_L	Sharp Lip Puller (Left)	AU13_R	Sharp Lip Puller (Right)
AU14	Dimpler	AU14_L	Dimpler (Left)
AU14_R	Dimpler (Right)	AU15	Lip Corner Depressor
AU15_L	Lip Corner Depressor (Left)	AU15_R	Lip Corner Depressor (Right)
AU16	Lower Lip Depress	AU17	Chin Raiser
AU18	Lip Pucker	AU19	Tongue Show
AU20	Lip Stretch	AU20_L	Lip Stretch (Left)
AU20_R	Lip Stretch (Right)	AU21	Neck Tightener
AU22_25_up_down	Lip Funneler (Both Lips)	AU23	Lip Tightener
AU24	Lip Presser	AU25	Lips Part
AU25_L	Lips Part (Left)	AU25_R	Lips Part (Right)
AU26	Jaw Drop	AU27	Mouth Stretch
AU28	Lips Suck	AU29	Jaw Thrust
AU30_L	Jaw Sideways (Left)	AU30_R	Jaw Sideways (Right)
AU31	Jaw Clencher	AU32	Bite
AU33	Blow	AU34	Puff
AU35	Cheek Suck	AU36	Tongue Bulge
AU37	Lip Wipe	AU38	Nostril Dilate
AU39	Nostril Compress	AU43	Eye Closure
AU43_L	Eye Close (Left)	AU43_R	Eye Close (Right)

Table 1.1: Added blend shapes

clude motion curves for facial morphs. This involved creating additional curves that specify the timing and intensity of facial movements based on the template. By controlling the acceleration and deceleration of these movements, we were able to create more naturalistic animations that reflect the dynamic nature of facial expressions.

For example, in the expression *big-threatening*, the motion curves were designed to gradually increase the influence of the blendshapes corresponding to AU10 (Upper Lip Raiser) and AU25 (Lips Part) while simultaneously decreasing the influence of AU4 (Brow Lowerer) as the expression transitions from a neutral state to one of aggression (figure **??**).

AU Code	Description	AU Code	Description
AU40	Sniff	AU41	Lid Droop
AU42	Slit	AU44	Squint
AU45	Blink	AU46	Wink
AU51	Head Turn Left (IK controlled)	AU52	Head Turn Right (IK controlled)
AU53	Head Up (IK controlled)	AU54	Head Down (IK controlled)
AU55	Head Tilt Left (IK controlled)	AU56	Head Tilt Right (IK controlled)
AU57	Head Forward (IK controlled)	AU58	Head Back (IK controlled)
AU61	Eyes Turn Left (lookat constraint)	AU62	Eyes Turn Right (lookat constraint)
AU63	Eyes Up (lookat constraint)	AU64	Eyes Down (lookat constraint)

Table 1.2: Action units skipped

AU Code	Description	AU Code	Description
AU1_a	Inner Brow Raise (Alternative)	AU2_a	Outer Brow Raise (Alternative)
AU4_a	Brow Lowerer (Alternative A)	AU4_b	Brow Lowerer (Alternative B)
AU6_a	Cheek Raise (Alternative A)	AU6_b	Cheek Raise (Alternative B)
AU9_a	Nose Wrinkle (Alternative A)	AU12_a	Lip Corner Puller (Alternative A)
AU12_b	Lip Corner Puller (Alternative B)	AU17_a	Chin Raiser (Alternative A)
AU18_a	Lip Pucker (Alternative)	AU25_a	Lips Part (Alternative A)
AU22_25_upper	Lip Funneler (Upper Lip)	AU22_25_down	Lip Funneler (Bottom Lip)
AU26_lip_down	Jaw Drop Bottom Lip Down (Alternative)	AU26_tongue_down	Jaw Drop Tongue Down (Alternative)
AU26_tongue_out	Jaw Drop Tongue Out (Alternative)	AU26_a	Jaw Drop (Alternative)
AU28_a	Lips Suck (Upper Lip)	AU28_bottom	Lips Suck (Lower Lip)
tongue_back_up	Tongue Back Up	tongue_out	Tongue Out
tongue_up	Tongue Up	tongue_wide	Tongue Wide

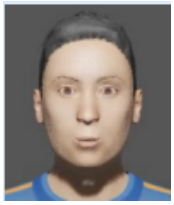
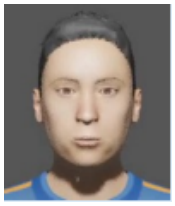
Table 1.3: Alternative blend shapes

1.4 Results and Evaluation

Synthesis of 22 facial expressions from the 40 brèves corpus can be seen in table **??**. The expressions were created by combining the relevant blendshapes to produce realistic and expressive facial animations.

Table 1.4: Synthesized Expressions

AZee rule	Original Expression	Synthesized Expression
almost-reaching:		

AZee rule	Original Expression	Synthesized Expression
big-threatening:		
closer-look:		
continuously:		
do-you-realise:		
impressive-grandiose:		
inter-subjectivity:		
it-is-a-shame:		

AZee rule	Original Expression	Synthesized Expression
most-probably:		
much-almost-too-much:		
nothing-sticks-out:		
peacefully:		
something-sticks-out:		
takes-a-while:		
too-scared-to-look:		

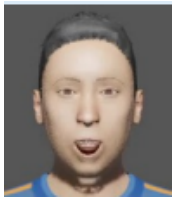
AZee rule	Original Expression	Synthesized Expression
trouble-disturbance:		
uneasy-awkward:		
with-chaos:		
with-no-precision:		
with-surprise:		
with-uncertainty:		
with-worry:		

Figure 1.5: Motion Curves for *big-threatening*

Video [1] shows the range of all the blendshapes in AZee's morph set. Video [2] shows all the synthesized facial expressions with their interpolations.

Table **??** shows the subjective evaluation by the linguist who created the corpus of the facial expressions using AZee. The expressions were assessed on isolated still images of the apex of their blendshapes based on their accuracy, expressiveness, and effectiveness in conveying the intended emotional and grammatical cues.

We observe that expressions that involved subtle mouth movements, such as "it-is-a-shame" or "something-sticks-out," were more challenging to model accurately, highlighting areas for further refinement.

We also compare the synthesized utterance *big-threatening(hot())* (with facial expressions) and the utterance *hot()* (without facial expressions) to evaluate the impact of facial expressions on ?? comprehension along with the original video. The results can be seen in the supplementary video[3]. We observe how non-manuals can enhance the meaning of the sign.

[1] https://github.com/Paritosh97/phd/raw/master/supplementary_material/ch6allaus.mp4

[2] https://github.com/Paritosh97/phd/raw/master/supplementary_material/ch6allexpressions.mp4

[3] https://github.com/Paritosh97/phd/raw/master/supplementary_material/ch6_comparison.mp4

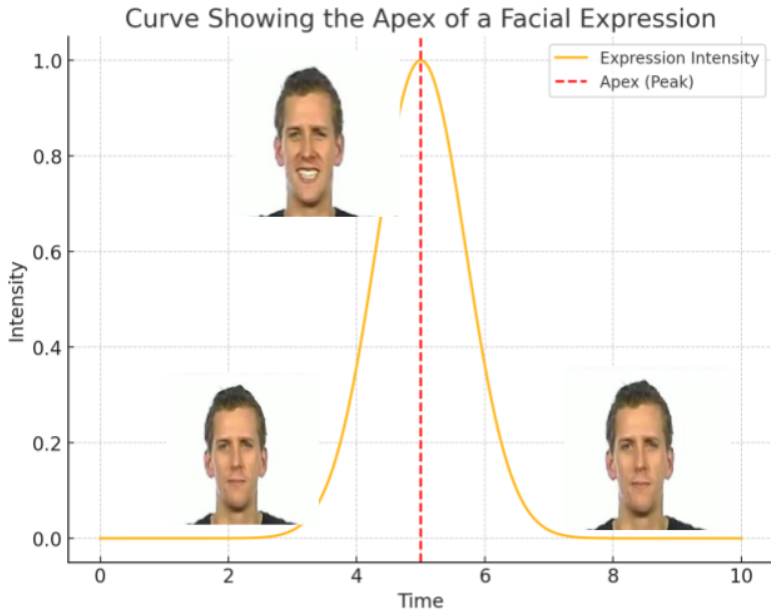


Figure 1.6: Apex of *big-threatening*

1.5 Conclusion

In this chapter, we presented a method for synthesizing facial expressions using the AZee framework, focusing on the creation of blendshapes from action units (AUs) and the generation of motion curves to control these shapes over time. Our method involved analyzing AUs from the ?? system and extending the morph set in the AZee language with 94 blendshapes. Next we created those blendshapes on our avatar based on a MakeHuman template mesh. We also used motion curves to animate these blendshapes. The synthesized facial expressions were evaluated subjectively by the linguist who created them. The results show that even though some of the expressions were challenging to model accurately, our method was able to generate some expressions which were effective in conveying the intended form of the corresponding AZee rule.

For future, we plan to use more modern systems to refine the blendshapes and motion curves to improve the accuracy and naturalness of the facial expressions. A good starting point for this would be to animate these rules using a point cloud based capturing systems and then study the capture data to accurately define blendshapes as well as motion templates. Use of emotion recognition systems such as **?** can be used on the existing dataset to define the flame **?** principle components. However, establishing a link between the AZee rule and these principle components which drive the mesh will be a challenge. Some syn-

Expression	Limitations
almost-reaching	Mouth modeling unconvincing.
continuously	”Pffff” air and cheek puff difficult, neutral eyebrows.
do-you-realise	Thick eyebrow issue.
it-is-a-shame	Mouth expression not quite real.
most-probably	Less visible teeth preferred, thick eyebrow issue.
much-almost-too-much	Frowning eyebrows and lack of eye wrinkles not convincing.
nothing-sticks-out	Tucked lips difficult to model.
something-sticks-out	Interpreted as confusion, mouth modeling limitation.
trouble-disturbance	Frowning eyebrows difficult, mouth ”rising” hard to model, result not convincing.
uneasy-awkward	Tongue tip out with slightly open mouth hard to model, unconvincing.
with-chaos	Single cheek blow/puff and alternating eye blinks hard without animation.
with-no-precision	Upper lip over lower and mouth near nose unmodellable.
with-surprise	Cannot lower lower eyelid fully, thick eyebrow issue.
with-uncertainty	Appears sadder than uncertain, thick eyebrow issue.
with-worry	Lack of wrinkles around nose/forehead.

Table 1.5: Limitations for the synthesized blendshapes

thesized facial expressions using the FLAME model on a default SMPL-X avatar can be seen in the annex section **??**.