

Combination of facial movements on a 3D talking head

The Duy Bui

Dirk Heylen
University of Twente
Department of Computer Science
The Netherlands
{theduy,heylen,anijholt}@cs.utwente.nl

Anton Nijholt

Abstract

Facial movements play an important role in interpreting spoken conversations and emotions. There are several types of movements, such as conversational signals, emotion displays, etc. We call these channels of facial movement. Realistic animation of these movements would improve the realism, liveliness of the interaction between human and computers using embodied conversational agents. To date, no appropriate methods have been proposed for integrating all facial movements. We propose in this paper a scheme of combining facial movements on a 3D talking head. First, we concatenate the movements in the same channel to generate smooth transitions between adjacent movements. This combination only applies to individual muscles. The movements from all channels are then combined taking into account the resolution of possible conflicting muscles.

1 Introduction

Facial movements play an important role in interpreting spoken conversations and emotions. They occur continuously during social interactions and conversations. They include lip movements when talking, conversational signals, emotion displays and manipulators to satisfy biological needs. Unfortunately when and how a movement appears and disappears, and how co-occurrent movements are integrated (co-articulation effects, for instance) are difficult to quantify [14]. In addition, the problem of overlaying and blending facial movements in time, and the way felt emotions are expressed in facial activity during speech, has not received much attention [20].

In the field of embodied agents, facial animation has received quite a lot of attention. Realistic animation of faces would improve the realism, liveliness of the interaction between human and machine. To create realistic facial animation, many 3D face models have been proposed (see [22] for a summary). Recently, more and more attention has

been paid to facial animation in synchronization with synthesized/natural speech. Many talking faces have been developed. Examples include [1], [6], [19] and [23]. These systems combine facial movements by just adding them together without taking into account the resolution of conflicting muscles/parameters temporally. Specifically, significant attention has been paid to visual speech [5, 19]. Some systems are also able to generate facial expressions as conversational signals during speech [2, 24]. However, no appropriate methods have been proposed for integrating all these facial movements.

The activity of human facial muscles is far from simply additive. A typical example would be smiling while speaking. The Zygomatic Major and Minor muscles contract to pull the corner of lip outward, resulting in a smile. The viseme corresponding to the phoneme “@U” in the word “Hello” requires the contraction of the lip funneler Orbicularis Oris, which drives the lips into a tight, pursed shape. However, the activation of the Zygomatic Major and Minor muscles together with the lip funneler Orbicularis Oris would create an unnatural movement. We call these “conflicting” muscles. The activation of a muscle may require the deactivation of other muscles [10]. Depending on the priority of the tasks to be performed on the face, appropriate muscles are selected to activate. In most of the cases, the visual speech has higher priority than the smile. The smile may also have higher priority than the visual speech when the subject is too happy to utter the speech naturally.

In this paper, we propose a scheme of combining facial movements on a 3D talking head. There are several types of movements, such as conversational signals, emotion display, etc. We call these channels of facial movement. We concentrate on the dynamic aspects of facial movements and the combination of facial expressions in different channels that are responsible for different tasks. First, we concatenate the movements in the same channel to generate smooth transitions between adjacent movements. This combination only applies to individual muscles. The movements from all channels are then combined taking into ac-

count the resolution of possible conflicting muscles.

Section 2 gives an overview of the system. The 3D face model for the talking head is discussed in Section 3. This section also presents a summary of conflicting muscles on the face. We break facial movements into so-called ‘atomic’ movements as described in Section 4. Section 5 explains how facial movements inside a channel are combined while Section 6 explains the combination of movements from all channels.

2 System overview

Our talking face takes as input the text to be pronounced marked up with ‘atomic’ facial movements other than lip movements when talking. We define an ‘atomic’ movement as a group of muscle contractions that share the same function (e.g. conversation signal, emotion display), start time, end time, and onset and offset duration. A simple example of marked up text looks like this

```
<PHONEME time="0" text="Oh really? I like it very much." />  
<EMOTIONDISPLAY time="0" duration="2.0" Surprise="0.7" />
```

Each ‘atomic’ facial movement belongs to a specific channel, which contains only non-conflicting movements. Atomic movements within a channel occur sequentially, although they may overlap each other at their beginning and ending. This classification is also based on the function of the movements [13]. It is similar to Pelachaud et al.’s one [23]. Movements from different channels can happen in parallel and can involve conflicting muscles.

In our system, we distinguish six channels:

- Channel 1 contains **manipulators**, which are movements to satisfy biological requirements of the face. In our system, we consider eye blinking to wet the eyes as manipulators. These movements are random rather than repeated with fixed rate as in [24]. The random eye blinking is generated based on the algorithm proposed in [17].
- Channel 2 contains **lip movements** when talking (represented as viseme segments). Lip movements are generated from the text that is going to be spoken by the talking head. The text is converted to phoneme segments (phoneme with temporal information - starting and ending time) [30]. The phonemes are converted to corresponding visemes. Each viseme is equipped with a set of dominance functions of parameters participating in the articulation of the speech segment. We use dominance functions from [6] for each viseme segment.
- Channel 3 contains **conversational signals**. These are movements to accentuate or emphasize speech. They

can occur on pauses due to hesitation or to signal punctuation marks (such as a comma or an exclamation mark). They are used to help the interaction between the speaker and the listener or to provide feedback from a listener. The generation of conversational signals can be done by analyzing the text [24] or speech [2].

- Channel 4 contains **emotion displays** and **emotion emblems**, which are movements to express emotions that are currently being experienced by the speaker or being referred, respectively. We have proposed a fuzzy rule based system to generate emotion displays from emotions [3].
- Channel 5 contains **gaze movements** and Channel 6 contains **head movements**. Gaze and head movements are generated to support eye contact or to point to something during conversation. Head movements are also used to replace verbal content (e.g. nodding the head for saying yes). As the eyes and the head do not stay in the same place all the time, we use noise function to create random subtle movements to make the talking head more lively.

An overview of our system can be seen in Figure 1. From text input, the text to phoneme module [30] generates phoneme sequences, which are used to synthesize speech [8]. They are also used to generate lip movements when talking.

The system also takes other facial movements as input. Facial movements are then combined in two stages: internal channel combination and cross channel combination. The former concatenates movements in the same channel and creates smooth transitions between them. The latter combines the movements from all channels taking into account the muscle conflicting resolution. The result is displayed on a 3D face model to create the final animation in synchronization with the synthesized speech.

3 The 3D face model

Our 3D face model is described in [4] in some details. It is a simple muscle-based 3D face model that can realize both of the following objectives: producing realistic facial expressions and real-time animation on a regular personal computer. The face model, which is also not too complicated so as to keep the animation realtime, allows high quality and realistic facial expressions. The face is equipped with a muscle system that produces realistic deformation of the facial surface, handles multiple muscle interaction correctly and produces bulges and wrinkles in real-time.

In our face model, there are 35 muscles that are responsible for visual speech (lip movements) and facial expressions. The muscles are shown in Table 1. Note that for

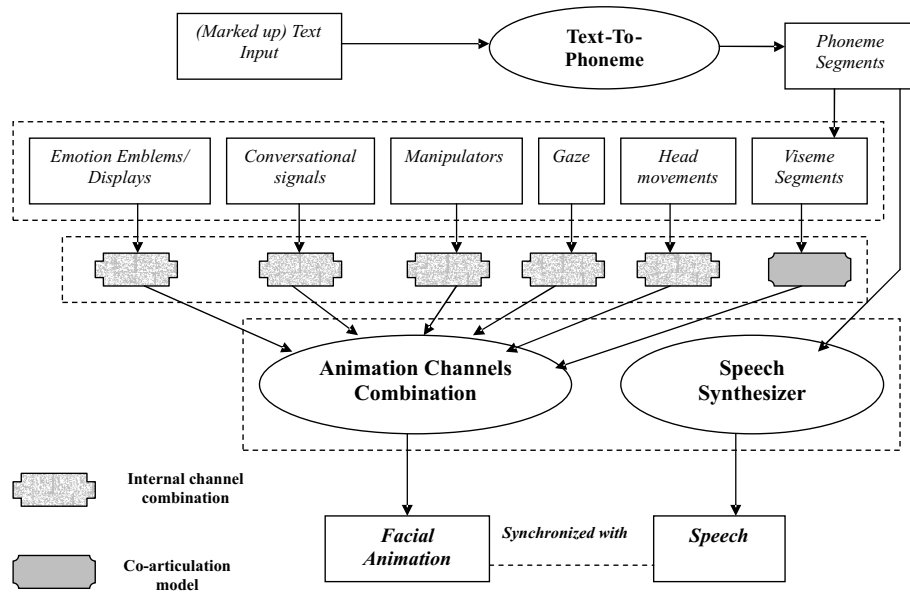


Figure 1. Overview of the system

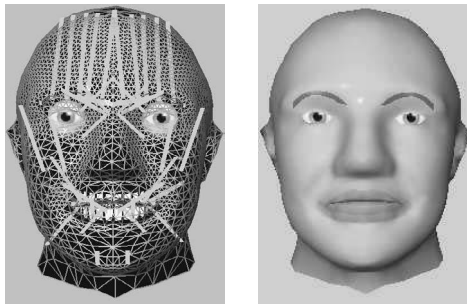


Figure 2. Our 3D face model

every muscle in that table, there is a left one and a right one except the Orbicularis Oris and Masseter. We also have parameters for eye and head movements.

We have paid attention to the combination of multiple muscle contractions [4]. However in the face, not every muscle can contract at the same time. For example, the lip protrusion muscles Levator Labii Superioris, which pull the upper lip apart from the lower lip, and the lip pressor muscle Orbicularis Oris, which presses the lips together, cannot contract together. The simple additive combination of their contraction would result in unnatural movement, which can be seen in Figure 3. Thus, some muscle actions require the deactivation of other muscles. Ekman and Friesen [10] have discussed possible conflicting movements on the face in term of Action Units (summarized in [29]). Based on this information, we have generated the list of all conflicting pairs of muscles, shown in Table 2. As can be seen from the table, most of the pairs of conflicting muscles in-

volve the mouth region. Other ones include the opening and closing of the eyelids, the Frontalis Medialis and the nose wrinkling muscles.

For convenience, when referring both to facial muscle and parameters for eye and head movements, from now on, we will use the same term “parameter”.

4 Atomic facial movements

Although facial movements happen continuously, most of them are known from electromyography (EMG) studies to occur in distinct phases [16]. The flow of movements can then be broken up into so-called ‘atomic’ movements. Kalra et al. [18] break each movement into four phases: attack (onset), decay, sustain (apex), and release (offset). We follow Pelachaud et al. [24] to synthesize facial movements in three phases: onset, apex, and offset.

Basically, each facial movement in our system is defined as a triple:

$$\Gamma^m = \{(PV_i^m), Ts^m, Te^m\}$$

where the set of parameters values $(PV_1^m, PV_2^m, \dots, PV_n^m)$ defines the target state of the movement (n is the number of deformation parameters in the face); Ts^m and Te^m are the starting and ending time of the movement, respectively.

Lip movements when talking:

The lip movements when talking are also associated with an apex duration Da^m :

$$\Gamma^l^m = \{(PV_i^m), Ts^m, Te^m, Da^m\}$$

The activity of these lip movements is created based on dominance functions, which are used for producing co-

articulation effect [5]. This will be discussed in Section 5.1.

Other movements:

Other movements are associated with two more temporal parameters: the **onset duration**, Do^m , determines how long the facial movement takes to appear; the **offset duration**, Dr^m , determines how long the facial movement takes to disappear:

$$\Gamma_o^m = \{(PV_i^m), Ts^m, Te^m, Do^m, Dr^m\}$$

The activity of a parameter involved in the creation of a single facial movement is described as a function of time:

$$F_p^m(t) = \begin{cases} PV_p^m \times \phi_+(t - Ts^m, Do^m) & \text{if } (Ts^m < t < Ts^m + Do^m) \\ PV_p^m & \text{if } (Ts^m + Do^m \leq t \leq Te^m - Dr^m) \\ PV_p^m \times \phi_-(t - Te^m + Dr^m, Dr^m) & \text{if } (Te^m - Dr^m < t < Te^m) \\ 0 & \text{if } (t \leq Ts^m \text{ or } t \geq Te^m) \end{cases}$$

where ϕ_+ and ϕ_- are the functions that describe the onset and offset phase of the parameter activity.

We used Essa's work [16] on analysis, identification and synthesis of facial expressions to design the temporal pattern of facial movements. Essa used exponential curves to fit the onset and offset portions of each deformation parameter. A function of the form $(e^{bx} - 1)$ is suggested for the onset portion, while a function of the form $(e^{c-bx} - 1)$ is suggested for the offset portion. We derived the following function to describe the onset portion of the parameter activity:

$$\phi_+(x, Do^m) = (e^{\frac{\ln 2}{Do^m} x} - 1)$$

The offset portion is described by the following function:

$$\phi_-(x, Dr^m) = (e^{\ln 2 - \frac{\ln 2}{Dr^m} x} - 1)$$

The activity of a muscle in a movement with 7s in duration, 2s in onset duration, and 3s in release duration can be seen in Figure 4.

5 Combination of movements in one channel

We combine movements in one channel by modulating the activity of each muscle involved the movements in that channel to create transition effects between movements. Recall from Section 2 that movements in one channel occur in sequence. However, they can be specified to overlap each other. When there are two overlapping movements, there will be a transition from the preceding one to the following one.

We use the dominance model [5] to create the co-articulation effect of lip movements when talking. Co-articulation is the blending effect that surrounding phonemes have on the current phonemes.

For the combination of movements from other channels, we propose an algorithm in Section 5.2 to produce smooth transition between movements.

5.1 Combination of lip movements when talking

Cohen and Massaro [5] have adapted Löfqvist's [21] gestural production model to drive their synthetic visual speech. In this model, a lip movement corresponding to a speech segment is represented as a viseme segment. It has dominance over the vocal articulators that increase and decrease over time during articulation. This dominance function determines how close the lips come to reaching their target value of the viseme. Adjacent segments will have overlapping dominance functions which lead to a blending over time of the articulatory commands related to these segments. Each segment has not only a single dominance function but rather a set of such functions, one for each articulator. The general form for dominance suggested by Cohen and Massaro is given by the negative exponential function $e^{-\theta\tau^c}$. In this function, dominance falls off according to the time distance τ from the segment center, to the power c modified by the rate parameter θ . Different dominance functions can overlap for a given segment. The weighted average of all the co-occurrent dominance functions produce the final lip shape.

5.2 Other movements

We propose here an algorithm for the concatenation of facial movements other than lip movements. When there are two overlapping movements, we create the transition from the preceding movement to the following one based on their activity.

For every two subsequent movements:

$$\Gamma_o^1 = \{(PV_i^1), Ts^1, Te^1, Do^1, Dr^1\} \text{ and}$$

$$\Gamma_o^2 = \{(PV_i^2), Ts^2, Te^2, Do^2, Dr^2\}$$

The two movements must occur sequentially, which requires:

$$Ts^2 > Ts^1 \quad \text{and} \quad Te^2 > Te^1$$

They are overlapping when:

$$Ts^1 < Ts^2 < Te^1$$

The activity of the combined movement follows the first movement until time Ts^2 . It then increases/decreases to reach the target of the second movement and then follows the normal second movement. The activity of a parameter p of the combined movement is described as follows:

$$(1) \text{ if } F_p^1(Ts^2) < PV_p^2$$

$$F_p^{12}(t) = \begin{cases} F_p^1(t) & \text{if } t \leq Ts^2 \\ F_p^1(Ts^2) + \phi_+(t - Ts^2, Do^2) & \text{if } Ts^2 < t < Ts^2 + \xi \\ PV_p^2 & \text{if } Ts^2 + \xi \leq t \leq Ts^2 + Do^2 \\ F_p^2(t) & \text{if } t > Ts^2 + Do^2 \end{cases}$$

where ξ is the time that $\phi_+(\xi, Do^2) = PV_p^2 - F_p^1(Ts^2)$,

No.	Muscle name	No.	Muscle name
01	Zygomatic Major	11	Levator Labii Nasi
02	Zygomatic Minor	12	Levator Labii Superioris
03	Triangularis	13	Depressor Supercilli
04	Risorius	14	Corrugator Supercilli
05	Depressor Labii	15	Depressor Glabellae
06	Mentalis	16	Levator Palebrae Superioris
07	Orbicularis Oris (Lip Funneler)	17	Orbicularis Oculi Palebralis (Eye Closing)
08	Orbicularis Oris (Lip Pressor)	18	Orbicularis Oculi Orbitalis
09	Frontalis Medialis	19	Masseter (jaw rotation)
10	Frontalis Lateralis		

Table 1. Implemented muscles in the system

	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19
01			x	x		x	x					x							
02			x	x		x	x					x							
03	x	x					x												
04	x	x					x												
05						x		x											
06	x	x			x														
07	x	x	x	x				x											
08					x		x				x	x							
09											x								
10																			
11								x	x							x			
12	x	x						x											
13																			
14																			
15																			
16											x								
17																x			
18																			
19																			

Table 2. Conflicting muscles in the system

(2) if $F_p^1(Ts^2) \geq PV_p^2$

$$F_p^{12}(t) = \begin{cases} F_p^1(t) & \text{if } t \leq Ts^2 \\ F_p^1(Ts^2) - PV_p^2 + \phi_+(t - Ts^2, Dr^1) & \text{if } Ts^2 < t < Ts^2 + \xi \\ PV_p^2 & \text{if } Ts^2 + \xi \leq t \leq Ts^2 + Do^2 \\ F_p^2(t) & \text{if } t > Ts^2 + Do^2 \end{cases}$$

where ξ is the time that $\phi_-(\xi, Dr^1) = F_p^1(Ts^2) - PV_p^2$.

The combined movement of Γo^1 and Γo^2 is then further

combined with next movements in the same manner to create the final facial movement of the channel. An example of combining the Zygomatic Major of two movements in the same channel is shown in Figure 5. The Zygomatic Major's activity of the combined movement follows the first movement until time 3, when there is a desire for the second movement. It then releases to the target value of the Zygomatic Major in the second movement (0.5), followed by the second movement.

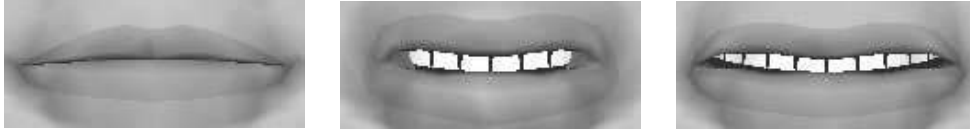


Figure 3. The addition of conflict muscles: lip pressor (left) and lip protrusion (middle) results in unnatural movement (right).

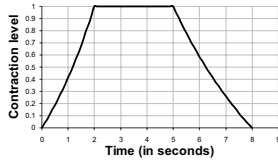


Figure 4. The activity of a facial movement

6 Combination of movements in different channels

In this section, we consider the combination of facial movements from different channels. To combine them together, we first resolve the conflicts between the parameters involved in different movements. After that, the activities of parameters are combined by taking the maximum values.

At a certain time, when there is a conflict between parameters in different animation channels, the parameters involved in the movement with higher priority will dominate the ones with lower priority. The activity of that muscle around that time is also affected so that the parameter cannot activate or release too fast. A parameter's activity $aF_p(t, c)$ at a time t in channel c is modified as follows:

- $aF_p(t, c) = 0$ if there is an activated conflicting parameter involved in a higher priority movement at that time.
- $aF_p(t, c) = \min(F_p(t, c), \Psi_{\leftarrow}(t), \Psi_{\rightarrow}(t))$ if there is no activated conflicted parameter involved in a higher priority movement at that time.

where

$$\Psi_{\leftarrow}(t) = \phi_{+}(t - T_{\leftarrow}, Dmin_{+}),$$

T_{\leftarrow} is the last time before t there exists an activated conflicting parameter involved in a higher priority movement, $Dmin_{+}$ is the minimum onset duration for a parameter to activate completely;

$$\Psi_{\rightarrow}(t) = \phi_{-}(Dmin_{-} - (T_{\rightarrow} - t), Dmin_{-}),$$

T_{\rightarrow} is the next time after t there exists an activated conflicting parameter involved in a higher priority movement, $Dmin_{-}$ is the minimum offset duration for a parameter to release completely.

The final activity of a parameter is obtained by taking the maximum value of this parameter from all channels c :

$$F_p(t) = \max(aF_p(t, c))$$

Figure 6 illustrates this algorithm. It shows the Orbicularis Oris muscle involved in speech. The Orbicularis Oris muscle conflicts with the Zygomatic major muscle and has higher priority. When the Orbicularis Oris is activated (at time 3), the Zygomatic major is inhibited. The activity of the Zygomatic major before that time is adjusted so it does not release too fast, which creates unnatural movement.

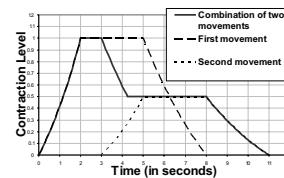


Figure 5. Combination of the Jaw Rotation of the two movements in the same channel.

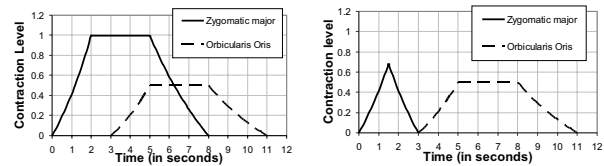


Figure 6. The activity of Zygomatic major and Orbicularis Oris before (left) and after (right) implementing our combination algorithm

Figure 7 shows (frame by frame) our face model uttering the sentence "Oh, really? I like it very much". Two emotion displays happen during the utterance. The first one is a full surprise display, which start at time 0 and lasts 2 seconds. The second one is a full happiness display, which start at 1 second later and lasts 3 seconds.

The figure shows the smooth and natural combination of visual speech and emotion displays. First, the transition

from the surprise display to the happiness is shown from frame 16 to frame 23. The happiness display is then combined with the visual speech for “I like it” as there are no conflicting parameters in these movements. The conflict starts at frame 34, where the pronunciation of phoneme “v” in the word “very” requires the activation of the lip funneler Orbicularis Oris. This muscle conflicts with the smiling muscles Zygomatic Major and Minor. Because the visual speech movement has higher priority than the emotion display, the activation of the lip funneler Orbicularis Oris results in the deactivation of the Zygomatic Major and Minor. This deactivation, however, starts earlier from frame 30 so they do not release too fast to prevent unnatural movement. After finishing the sentence, the Zygomatic Major and Minor gradually reach the full contraction (frame 47).

7 Conclusion

In this paper, we have discussed the problem of combining facial movements on a 3D talking head. We concatenated the movements in the same channel to generate smooth transitions between adjacent movements. The movements from all channels were then combined taking into account the resolution of possible conflicting muscles. The activity of all muscles were adjusted so they do not contract or release too fast. We have succeeded in creating natural facial animation of a talking head uttering while displaying other facial movements such as conversational signals, manipulators and emotion displays.

In the future, we want to test out other functions that simulate the onset and offset portion of a muscle’s activity. While this is not very important for most of the facial movements, it may affect the believability of felt emotion displays. Finding the appropriate values of the onset and offset durations for felt emotion displays also plays an important role in making these display natural, which is also an issue need to be investigated.

References

- [1] I. Albrecht, J. Haber, K. Khler, M. Schrder, H.P. Seidel(2002), “May I talk to you? :-) Facial Animation from Text”, *Proc. Pacific Graphics 2002*.
- [2] I. Albrecht, J. Haber, H.-P. Seidel (2002), “Automatic Generation of Non-Verbal Facial Expressions from Speech”, *Proc. CGI 2002*.
- [3] Bui,T.D., Heylen,D., Poel, M. and A., Nijholt(2001), “Generation of facial expressions from emotion using a fuzzy rule based system”, *Proc. AI 2001*.
- [4] Bui, T.D., D. Heylen & A. Nijholt (2003), “Improvements on a simple muscle-based 3D face for realistic facial expressions”, *Proc. CASA’2003*.
- [5] Cohen, M. M., & Massaro, D. W. (1993), “Modeling coarticulation in synthetic visual speech”, In N. M. Thalmann & D. Thalmann (Eds.) *Models and Techniques in Computer Animation*, Tokyo: Springer-Verlag.
- [6] DeCarlo, D. C. Revilla, M. Stone and J. Venditti (2002), “Making discourse visible: Coding and animating conversational facial displays”, In *Computer Animation 2002*, pp. 11-16
- [7] Ira Cohen, Nicu Sebe, Larry Chen, Ashutosh Garg, Thomas S. Huang (2002), Facial Expression Recognition from Video Sequences: Temporal and Static Modeling, submitted to CVIU special issue on Face recognition.
- [8] T. Dutoit, V. Pagel, N. Pierret, O. van der Vreken, F. Bataille (1996), “The MBROLA Project: Towards a Set of High-Quality Speech Synthesizers Free of Use for Non-Commercial Purposes”, *Proc. ICSLP 96*, Philadelphia.
- [9] Ekman, P. and W. Friesen(1975). *Unmasking the Face*. Prentice Hall.
- [10] Ekman, P., and W. Friesen (1978), *Facial Action Coding System*, Consulting Psychologists Press, Inc.
- [11] P. Ekman (1982), *Emotion in the human face*, Cambridge university Press.
- [12] P. Ekman (1984), “Expression and the nature of emotion”, in *Approaches to emotion*, K. Scherer, P. Ekman ed.
- [13] P. Ekman (1989), “The argument and evidence about universals in facial expressions of emotion”, in H. Wagner and A. Manstead, eds., *textitHandbook of Social Psychophysiology*, Wiley, Chichester, New-York.
- [14] P. Ekman, T.S. Huang, T.J. Sejnowski, and J.C. Hager (Eds) (1993), “Final Report to NSF of the Planning Workshop on Facial Expression Understanding,” Human Interaction Laboratory, University of California, San Francisco, March, 1993.
- [15] Irfan A. Essa and Alex Pentland (1994), “A Vision System for Observing and Extracting Facial Action Parameters”, *Proc. of IEEE Computer Vision and Pattern Recognition Conference*.
- [16] Irfan A. Essa (1994), “Analysis, Interpretation, and Synthesis of Facial Expressions”, Ph. D. Thesis, Massachusetts Institute of Technology, Cambridge, MA.
- [17] L. Itti, N. Dhavale, F. Pighin (2003), “Realistic Avatar Eye and Head Animation Using a Neurobiological Model of Visual Attention”, *Proc. SPIE 48th Annual International Symposium on Optical Science and Technology*.
- [18] Prem Kalra (1993), “An Interactive Multimodal Facial Animation System”, PhD Thesis, University of Geneva and EPFL.
- [19] Scott A. King, Richard E. Parent, and Barbara Olsafsky (2000), “An Anatomically-Based 3D Parametric Lip Model to Support Facial Animation and Synchronized Speech”, *Proc. Deform 2000*.

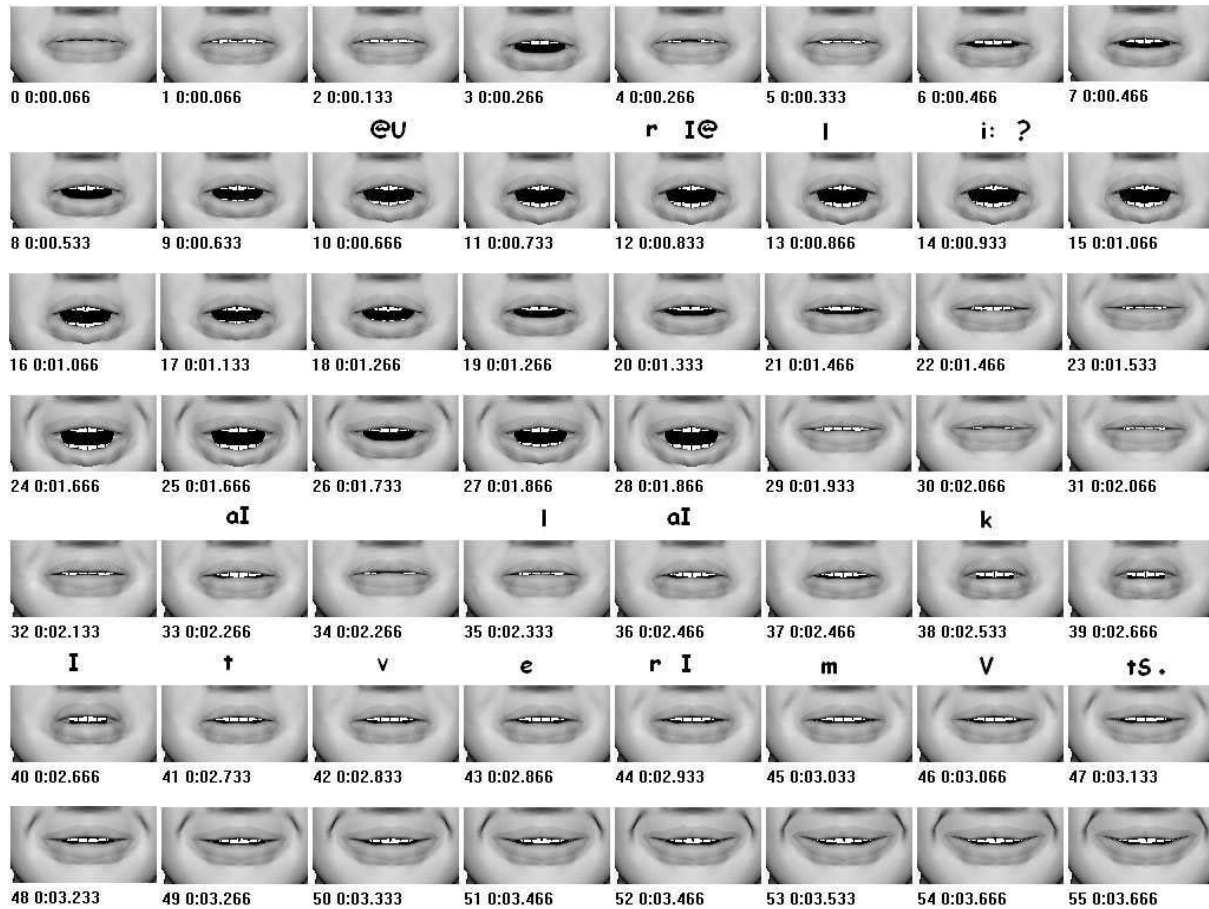


Figure 7. Our talking head utters the sentence “oh, really? I like it very much.” while displaying surprise followed by happiness (frame by frame).

- [20] Latta, C., Alvarado, N., Adams, S.S., & Burbeck, S. (2002), “An expressive system for animating characters or endowing robots with affective displays”, *Proc. AECSI - AISB'02*.
- [21] A. Löfqvist (1990), “Speech as audible gestures”, *Speech Production and Speech Modeling*.
- [22] Parke, F. I. and Waters, K. (1996), *Computer Facial Animation*, AK Peters.
- [23] C. Pelachaud, M.L. Viaud, H. Yahia (1993), “Rule-Structured Facial Animation System”, *IJCAI*, vol.2.
- [24] Pelechaud, C., Badler, N. I., & Steedman, M. (1996), “Generating facial expressions for speech”, *Cognitive Science*, 20, 1-46.
- [25] Prevost, S. and C. Pelachaud (1995), “Talking Heads: Physical, Linguistic and Cognitive Issues in Facial Animation”, *Course Notes for Computer Graphics International 1995*, Leeds, UK.
- [26] K.L. Schmidt (2000), “Variation in the timing and display of the human smile”, *Am J Phys Anthropol [Suppl]* 30:272.
- [27] K.L. Schmidt (2001), “Dynamic modeling of human facial expression”, *Am J Phys Anthropol [Suppl]* 31:132.
- [28] K.L. Schmidt and J. Cohn (2001), “Dynamics of facial expression: Normative characteristics and individual differences”, *Proc. ICME2001*.
- [29] A. Wojdel, L.J.M. Rothkrantz & J.C. Wojdel (2003), “Fuzzy-logical implementation of co-occurrence rules for combining AUs”, *Proc. CGIM 2003*.
- [30] <http://www.anthemion.co.uk/speech.htm>