Personal style and non-negative matrix factorization based exaggerative expressions of face

Seongah Chin, Chung-yeon Lee, and Jaedong Lee
Division of Multimedia, Sungkyul University, Anyang, Republic of Korea

Abstract -*An intriguing tactic for achieving facial expressions beyond realistic approaches has gradually drawn lots of attention in the fields of video games, avatars, teleconference, human-computer interface and computer animation. However, most researches on facial expressions tend to remain faithful to the trend in realistic expressions rather than exaggerating expressions to take into account personal styles. In the work presented here, we propose a method for exaggeration of facial expressions created by exaggeration mapping that transforms facial motions into exaggerated motions. The exaggeration mapping is derived from non-negative matrix factorization. As if each individual has an identical personality, a conceptual mapping of personal styles for exaggeration of facial expressions needs to be considered. By conducting experiments, we have shown the validity of the exaggeration mapping and simulations of facial expressions.*

Keywords: facial expression, animation, exaggeration, avatars, emotion

1 Introduction

The human face plays a significant role in non-verbal communication with a plethora of expressions varied by personality. Facial expressions, physically formed by the motions of the eyes, eyebrows, mouth, and many specific muscles on the face, are crucial in determining the emotional state of the characters. Most researchers in the area of facial expression synthesis seem to focus on realistic expressions, or retargeting rather than exaggerating expressions [1][2]. But interested observers are also curious about exaggerated expressions suitable for the personality of characters depicted in applicable areas. To draw attention to subjects, it is necessary to exaggerate facial expression, but this is a labor-intensive process. The magnitude of an expression's intensity causes different interpretations, even in only one facial expression (e.g. anger, happy, etc.). This principle lets us intuitively recognize that diverse levels of intensity exist for even a single expression [5]. Understanding such an abstract message tends to be easier in real life than it is in cyber space or digital media because of awareness of atmosphere as like reading subtle nuances without verbal expression. So, a methodology for exaggerating facial expression is necessary to help people communicate efficiently in digital space.

In cognitive science, several researchers specializing in emotional intelligence have studied emotion decoding ability by exploring correlations between personal properties and facial expressions. For example, an individual who smiles frequently is seen as more sociable and competent than someone who rarely smiles, whereas an expression of fear can alert others to potential dangers in a situation. Thus, each expression represents the potential behavior of the individual. However, very few research papers have reported on the exaggeration of facial expressions that depends largely on the personality of each individual.

In this paper, we propose a method to exaggerate facial expressions by taking personal styles into consideration. To decompose facial motions acquired by motion capture into a form of transformation, we employ non-negative matrix factorization [6][7]. Differences between mean and variance-based transformation is applied to exaggerate facial motions. Myers-Briggs Type Indicator [8] classification is used to determine a personal style in order to define style rates. By conducting experiments, we show the validity of the proposed method by implementing facial expressions.

2 Personal Style mapping

Facial motion can be likened to a form of media that carries the internal emotions of human beings. Hence, to find common features of facial motion behaviors, we need to classify individuals into meaningful and coherent homogeneous groups. To classify individuals into coherently meaning groups which mostly show similar personality, we refer to previous research indicating that emotional intelligence is broadly defined as references of the 'ability to monitor one's own and others' feelings and emotions'. The previous research derives implications between facial measures and the Myers-Briggs feeling or thinking score. The MBTI, Form G, is a self-report, personality type indicator widely used in behavioral science. In our approach, EF (Extroversion and Feeling), ET (Extroversion and Thinking), IF (Introversion and Feeling) and IT (Introversion and Thinking) called 'styles' are used to classify personal styles of facial motions. Emotional perception affects mostly Extroversion or Introversion and Thinking or Feeling dimensions [9][10].

To derive proper ‘personal style mapping’ denoted by the Styled Exaggeration Rate (SER), we first need to analyze feature movements of facial expressions. The ultimate goal is to construct SER, which represents similar behaviors with respect to exaggerated motions. Defining personal style mapping begins with gathering both facial motion samples and MBTI information from around 30 participants. They are asked to respond to MBTI GS form and perform six universal expressions taken as realistically as possible considering three expression phases (i.e. an initial attack phase, a sustaining phase, and a relaxation phase). Sufficient training with pictures of professional actors was conducted before taking their performance pictures. In the process of taking the pictures of the performers, we used a Nikon D80 supporting 800×600 pixel resolution captured from 1m distance. For the purpose of pre-processing, we conducted transformation of the samples, including skew correction and scale, using Adobe Photoshop in order to minimize errors. The facial feature points were defined while referencing MPEG-4. The positioning of the markers was determined by carefully considering muscle movements. Feature movements of facial expressions need to be analyzed in order to make each personal style be applied to exaggeration rate. A proper exaggeration rate dependent on MBTI is required to apply SER.

Figure 1. A sample of pictures taken with denoted markers with index 8 and 35 and six universal expressions sadness, happiness, anger, surprise, fear and disgust from the left.

Given an expression sample obtained from pictures shown in Figure 1, we locate 30 marker positions on the facial muscles. We then compute a distance δ between six expressions and neutral expressions. A distance function δ between the neutral marker of expression and a marker vector of facial expressions is defined in Eq. (1)

$$\delta = \left(\sum_{s=1}^{k} \lVert m_s - m_{Ns} \rVert^p\right)^{\frac{1}{p}}, \qquad (1)$$

where an index for markers $1 \leq s \leq 30$. m_s is the component of marker vectors for an expression and m_{Ns} is the component of marker vectors for neural expression.

The average distances of six expressions associated with four styles appear in Table 1. For example, surprise implies the largest average value 100.14 at style EF meaning that surprise at EF style is rather intensive in the expression than other styles.

$$\varphi_i = \frac{\Delta_{max} - \Delta_i}{\Delta_{max}} + \frac{\Delta_i}{\Delta_{max} \cdot g}, \qquad (2)$$

where $\Delta_{max} = \delta_{max} - \delta_{min}$ and $\Delta_i = \delta_{max} - \delta_i$,

The initial style rates with g=1.5 and expression index i shown in Table 2 are calculated using Eq. (2), the values of which imply that the expression is intensive at a style. For example, happiness, surprise and sadness are the most intensive expressions at EF. So, we need to rate this property according to how styles relate to intensiveness of expressions. Each style and expression in Figure 2 appears on a radial shape graph as well. It illustrates relative positions of four styles with six expressions. Surprise is located on the outside, which accounts for the fact that it has larger movements than the other expressions.

The style mapping noted by SER is defined in Eq. (3).

$$\Theta_s(\omega) = \begin{cases} C & \omega \geq 1 \\ k\left(\omega \cdot \frac{1}{\varphi}\right)^n \cdot e^{-\frac{1}{\varphi}(\omega+1) \cdot n} + 1 & 0 \leq \omega < 1 \end{cases}, \qquad (3)$$

where style index s, exponent parameter n, scale parameter k, exaggeration rate ω (0–1), and C is convergent value.

If ω is in $0 \leq \omega < 1$ then SER is gained by the second definition in Eq. (3). Provided that $\omega \geq 1$, SER returns the convergence C. In addition, we calculate the scaling constant k in order to keep $1 \leq \Theta(\omega) \leq 1.5$ (k = 5.5100e-004). Users are asked to input MBTI style and exaggeration rate ω in which $\omega = 0$ means no exaggeration is applied while $\omega = 1$ indicates the highest rate is demanded. Figure 3 plots the SER for surprise with EF, ET, IF and IT styles. It shows that the circle means the SER $\Theta_{s=IT}(\omega) = 1.2$ at exaggerate rate $\omega = 0.6$ with MBTI style IT.

Table 1. Average distance δ of styles

Expression\Style	EF	ET	IF	IT
Happiness	68.82	64.45	50.90	43.65
Surprise	100.14	94.73	64.05	55.82
Sadness	60.19	52.17	43.39	34.69
Fear	62.78	66.91	52.20	41.86
Anger	67.10	72.80	56.94	58.99
Disgust	64.81	70.88	49.65	39.42

Table 2. Style rates φ_i

Expression\Style	EF	ET	IF	IT
Happiness	1.00	0.97	0.88	0.83
Surprise	1.00	0.98	0.86	0.83
Sadness	1.00	0.95	0.89	0.83
Fear	0.97	1.00	0.90	0.83
Anger	0.94	1.00	0.83	0.85
Disgust	0.97	1.00	0.89	0.83

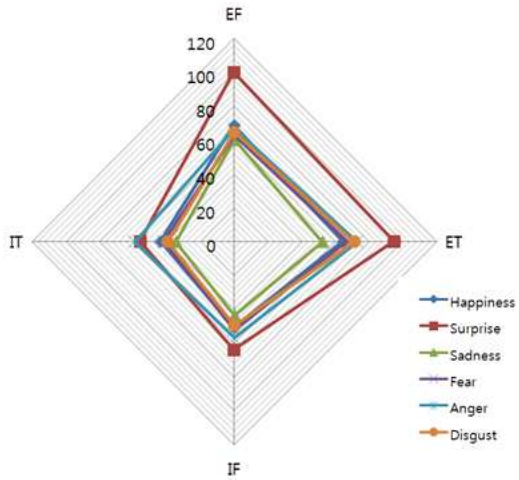

Figure 2. Radial shape graph of average distance of styles

Horizontal axis indicates exaggeration rate ω selected by a user and vertical axis account for the SER, $\Theta(\omega)$. Once $\omega \geq 1$ all the SER stay on the different convergent values 1.5 for EF, 1.47 on ET, 1.31 for IF and 1.27 for IT.

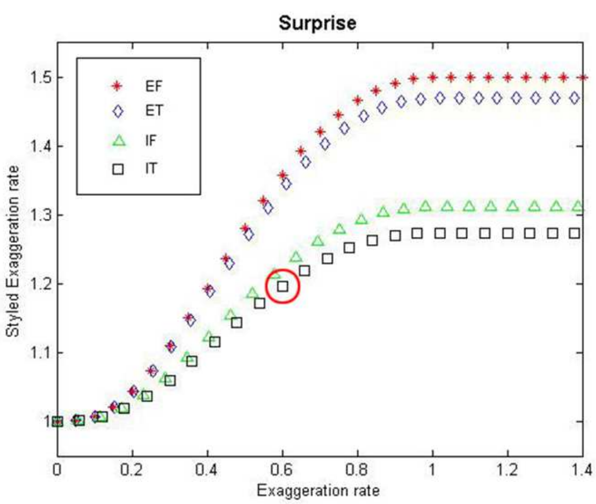

Figure 3. The SER, $\Theta_{s=IT}(\omega) = 1.2$ at exaggerated rate ω=0.6 with MBTI style IT.

3 Exaggeration Mapping

A fundamental idea of the approach to exaggeration mapping of facial motions is to put more emphasis on the features where movements are relatively greater than others. The first step to making this work is to seek proper data factorization. A task to acquire the optimal structure of feature analysis is thought to be a crucial step. How do we exaggerate facial motions? First, we realize that facial motions need to be factorized into two matrices. Once decomposition matrices are obtained, a transformation on the matrices should be applied. Non-negative matrix factorization (NMF) is successfully applied to a variety of data set to find part-based linear representations of non-negative data. Our approach is to employ NMF to represent the feature of motion data obtained by using optical motion capture system. Cross-cultural research presents six universal expressions: surprise, fear, disgust, anger, happiness, and sadness [4] of which a sequence of facial markers called a facial motion is factorized in order to exaggerate individual facial markers of the facial motion. Each motion consists of 3-D movements of markers attached to the principal facial muscles of the actor.

The original motion data has the movements of markers on the facial muscles and it consists of either positive values or negative values, so we first translate by adding the minimum value of the data to make all its values non-negative applicable to NMF. The facial motions are regarded as $n \times m$ matrix M, each column of which consists of x, y and z coordinates of markers. Given a facial motion denoted by M, NMF decomposes M into two matrices B and E as Eq. (4) to approximate the facial motions.

$$M_{i\mu} \approx (BE)_{i\mu} = \sum_{a=1}^{r} B_{ia} E_{a\mu} \tag{4}$$

Each column of matrix B contains a basis vector while each column of E includes the weights corresponding to the measurement column in M using the bases from B. The dimensions of the factorized matrices B and E are $n \times r$ and $r \times m$, respectively with r satisfying that $(n + m)r < nm$. To estimate the factorization matrices, an objective function has to be defined. This objective function works out the likelihood of computing the facial motions in M from the basis B and encodings E. A possible objective function we have used is given in Eq. (5)

$$H = \sum_{i=1}^{n} \sum_{\mu=1}^{m} [M_{i\mu} \log(BE)_{i\mu} - (BE)_{i\mu}] \tag{5}$$

Solutions for NMF begin with initializing non-negative conditions for B and E. Continuing the iteration of the update rules in Eqs. (6), M finds an approximate factorization $M \approx BE$ by converging to a local maximum of the objective function given in Eq. (5)

$$B_{ia} \leftarrow B_{ia} \sum_{\mu} \frac{M_{i\mu}}{(BE)_{i\mu}} E_{a\mu}$$

$$E_{a\mu} \leftarrow E_{a\mu} \sum_{i} B_{ia} \frac{M_{i\mu}}{(BE)_{i\mu}} \tag{6}$$

A fundamental idea of exaggerating facial motions is likely to employ exaggeration of difference from the mean in which it emphasizes movements for the distinctive feature points of a facial motion (Mo et al., 2004). We use the method to generate caricatures based on both feature DFM (difference from mean) and variance in order to compute exaggerated facial motions because it is superior to those that only concerns DFM.

Provided that the objective function stops after iterations E is properly acquired, E needs to be divided again into the mean mi and the deviation d_i of each column of E for its exaggeration. Each dimension is composed of a basis vector b_i

(the i-th column of matrix B) and its weights, including expectation mi and deviation d_i of the i-th row in Matrix E. The residual $\vec{r}$ can be thought of as distinctive facial motions that are not expressed in the facial motions. The facial motion $\vec{f}$ in Eq (7) is in the form of a non-negative linear combination of the basis and a residual.

$$\vec{f} = \sum_i e_i \cdot \vec{b_i} + \vec{r} = \sum_i (m_i + d_i) \cdot \vec{b_i} + \vec{r} \qquad (7)$$

The exaggerated facial motion $\vec{f}'$ called exaggeration mapping in Eq. (8) is calculated by scaling the deviation d_i and residual $\vec{r}$ with an exaggeration factor $\Theta(\omega)$.

$$\vec{f}' = \sum_i (m_i + t \cdot d_i) \cdot \vec{b_i} + 0.5 \cdot \Theta(\omega) \cdot \vec{r} \qquad (8)$$

where $t_i = 1$ if $|d_i| < 2 \cdot s_i$ and $t = \Theta(\omega)$ if $|d_i| \geq 2 \cdot s_i$ with $|d_i| = |e_i - m_i|$ and standard deviation s_i.

If a style is defined, $\Theta(\omega) = \Theta_S(\omega)$, and if not, $\Theta(\omega) = \Theta_e(\omega)$, $1 \leq \Theta_e(\omega) \leq 1.5$. The more detailed definition of $\Theta(\omega)$ will be shown in the next section.

4 Experiments

4.1 Validation of Exaggeration of Facial Motion

We asked a participant to visualize six universal expressions as realistically as possible for three expression phases (i.e. an initial attack phase, a sustaining phase, and a relaxation phase) in order to properly capture expression data and analyze it. The captured data was then analyzed to understand the facial muscle behaviour of the individual for six universal facial expressions and motions; the data provides the core input of exaggerating facial motions.

Facial motions defined by a sequence of set marker locations at a time frame needs to be validated with respect to relative motions with various exaggeration rates. Each marker shows distinctive movements with different exaggeration rates. We display the motion gaps on the clearly characteristic marker (8th) as shown in Figure 4. The distinctive differences appear on both larger and smaller values.

4.2 Exaggerating Facial Expressions

In this research, we employ a muscle based method [13] to simulate exaggerated facial expressions. The muscles need to be embedded into the 3-D face model in order to realize exaggerating facial expressions. For linear muscle, we employ 18 linear muscles of which contraction and expansion movement can be explained by a fan-shaped geometry. By using cosine functions, the movement of vertices is determined by the distance from the beginning point and the degree of angular changes.

Three sphincter muscles are used to express eye blinking and mouth movement by using the ellipsoidal principles. However, an eye is not ideally shaped as ellipsis. However, an eye is not ideally shaped as ellipsis. Accordingly, some variation in terms of distortion of the eye shape should be allowed.

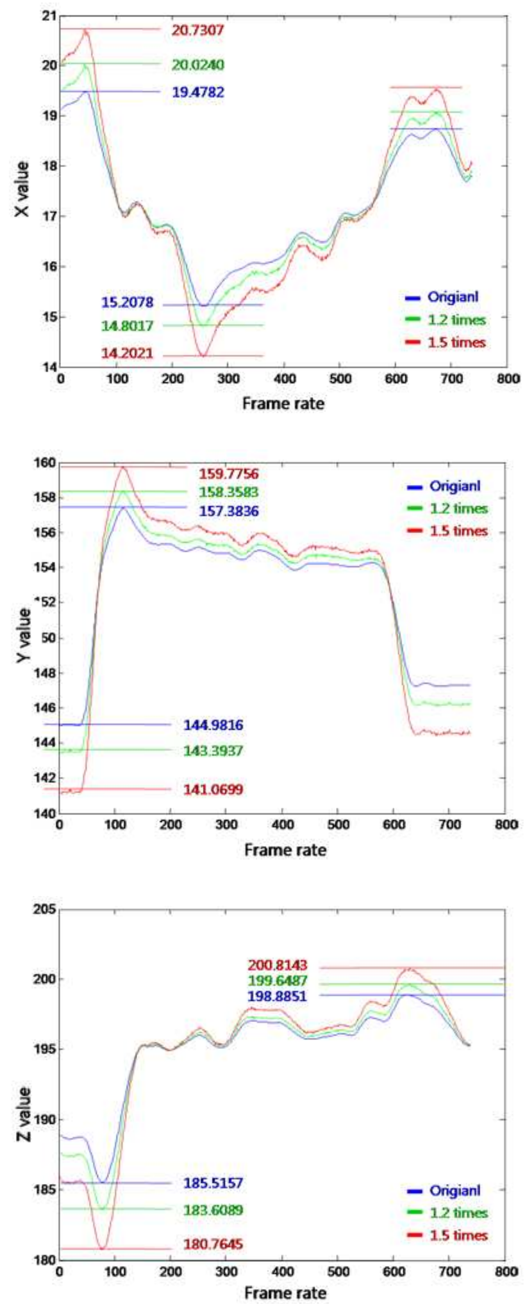

Figure 4. Exaggeration in difference on the 8th marker as denoted by the upper circle in Figure 1. for surprise motion with blue line (Θ_e (ω)=1.0), green line (Θ_e (ω)=1.2), red line (Θ_e (ω)=1.5) viewing distinctive gaps on large and small values.

To realize all details for making expressions when facial motions composed of markers are given, we follow all the details by adjusting muscle contractions.

Accordingly, some variation in terms of distortion of the eye shape should be allowed. To realize all details for making expressions when facial motions composed of markers are given, we follow all the details reported in the reference [5][12].

As previously noted, exaggerating facial motions is the first step to making exaggerating facial expressions. Based on exaggerating facial motions, we manipulate muscles by selecting proper contraction values to reach ultimate facial expressions. Here we show not only validation of exaggeration of facial motions, but also that the equivalent exaggeration of facial expression is identical to that of facial motion. The locations of the markers vary slightly because of changes of exaggeration rates.

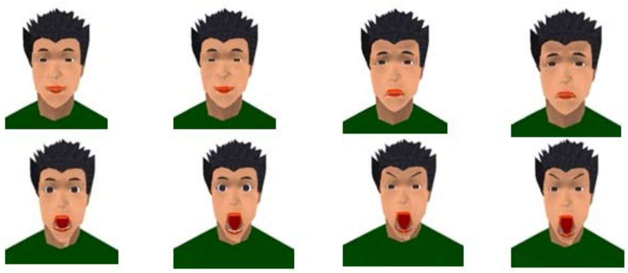

Figure 5. Happiness (s=IF, ω=0.55), sadness (s=EF, ω=0.41), surprise (s=IT, ω=0.73) and anger (s=IT, ω=0.57) with styled exaggeration rate 1.0 (original expressions) and 1.2 showing slight exaggerations

To simulate personally styled exaggeration of facial expressions, we first have to collect MBTI from participants. Fortunately, it is not complicated to gather personal style information MBTI from participants because it is mandatory for all freshman students to carry out MBTI tests conducted from the General Manpower Development Centre at our institute during their orientation course. As described earlier, style mapping formula is introduced in the previous section on personal style mapping.

To follow the proposed algorithm, the study participants were initially asked to respond to the MBTI information. Once the MBTI was obtained, the participants needed to input an initial exaggeration rate, which is computed by the personal style mapping to get the styled exaggeration rate. Provided that the styled exaggeration rate is determined, we apply the exaggeration mapping described in the previous section on exaggeration mapping on facial motions. If the styled exaggeration rate is not determined, the methodology ignores a personal style so that it simply asks an exaggeration rate without personal styles ranging from 1.0 to 1.5. By empirical experiments, we finally range the exaggeration rate from 1.0 to 1.5. Finally, the exaggerated facial expressions are simulated by manipulating muscles. Additional user surveys will be reported in the future. Figure 5 shows six slightly exaggerated expressions with styled exaggeration rate 1.0 (original) and 1.2 corresponding to various MBTI and exaggeration rate ω.

5 Conclusion

Exaggeration is considered to be similar to placing more of an emphasis on the key features of an object. The main features need to be highlighted more than the lesser ones. Current studies on facial expression synthesis tend to concentrate on realistic expression or retargeting, but observers are expressing more of an interest in characters showing exaggerated expressions, which is a labour intensive process.

In this paper, we focus on a methodology for automatic exaggeration of the facial expression considering personal styles, which has attracted significant attention in the fields of entertainment in digital media, avatars in the virtual world, and characters in video games, movies and commercials. To decompose facial motion acquired by motion capture into a form of transformation, we employ non-negative matrix factorization (NMF). Differences between mean and variance-based transformation are applied to exaggerate facial motions. Myers-Briggs Type Indicator (MBTI) classification is used to determine a personal style in order to define style rates. By conducting experiments, we show the validity of the proposed method by implementing facial expressions based on muscles. Finally, the proposed method has been validated by showing experimental results.

Acknowledgments

This research was partially supported by the Korea Research Foundation Grant funded by the Korea Government (KRF-2008-521-D00398).

6 References

[1] Q. Zhang, Z. Liu, B. Guo, D. Terzopoulos, H. Shum, “Geometry-driven photorealistic facial expression synthesis”; *IEEE Transactions on Visualization and Computer Graphics*, vol.12, no.1, pp. 48-60, 2006.

[2] I. Kotsia, I. Pitas, “Facial expression recognition in image sequences using geometric deformation features and support vector machines”; *IEEE Transactions on Image Processing*, vol.16, no.1, pp. 172-187, 2007.

[3] E.S. Chuang, “Analysis, synthesis, and retargeting of facial expressions”; Ph.D. dissertation, Stanford University, Palo Alto, CA, USA, 2004.

[4] P. Ekman, W.V. Friesen, “In Unmasking the face: a guide to recognizing emotions from facial clues”; Malor Books. Cambridge, MA, USA, 2003.

[5] S. Chin, K.Y. Kim, “Emotional intensity-based facial expression cloning for low polygonal applications”; *IEEE Transactions on Systems Man and Cybernetics – part C*, vol.39, no. 3, pp. 315-330, 2009.

[6] D.D. Lee, H.S. Seung, “Learning the parts of objects by non-negative matrix factorization”; *NATURE*, vol.401, pp. 788-791, 1999.

[7] P . Paatero, U. Tapper, “Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values”; *Environmetrics*, vol.5, no.2, pp. 111–126, 1994.

[8] I.B. Myers, M.H. McCaulley, N.L. Quenk, A.L. Hammer, “In MBTI manual: a guide to the development and use of the Myers-Briggs Type Indicator”; Consulting Psychologists Press. Palo Alto, CA, USA, 1998.

[9] D.A. Morand, “The emotional intelligence of managers: assessing the construct validity of a nonverbal measure of people skill”; *Journal of Business and Psychology*, vol.16, no.1, pp. 21-33, 2001.

[10] R.A. Martin, G.E. Berry, T. Dobranski, M. Horne, “Emotion perception threshold: individual differences in emotional sensitivity”; *Journal of Research in Personality*, vol.30, no.2, pp. 290-305, 1996.

[11] Z. Mo, J.P. Lewis, U. Neumann, “Improved Automatic Caricature by Feature Normalization and Exaggeration”; *In Proceedings of ACM SIGGRAPH 2004 Sketches International Conference on Computer Graphics and Interactive Techniques*, pp. 57-59, 2004.

[12] S. Chin, K.Y. Kim, “Expressive 3D Face in Mobile Devices”; *IEEE Transactions Consumer Electronics*, vol.54, no.3, pp. 1294-1302, 2008.

[13] F.I. Parke, K. Waters, “In Computer Facial Animation”; MA: A K Peters. Wellesley, 2004.