

筋肉モデルによる表情合成システムと筋肉パラメータの決定

下田 宏*¹ 楊 大昭*² 石井 裕剛*¹ 吉川 榮和*¹

A Facial Expression Synthesis System by using Muscle Model and Decision of Muscle Parameters

Hiroshi Shimoda*¹ Dazhao Yang*² Hirotake Ishii*¹ and Hidekazu Yoshikawa*¹

Abstract - A prototype system of facial expression synthesis has been developed by using facial muscle modeling method. The system can synthesize dynamic process of expressing facial expressions and display it as computer graphics. When synthesizing facial expressions by muscle modeling method, the parameters such as the muscle contraction rates and time variations should be decided beforehand. In this study, these parameters were decided by two subject experiments. The first experiment was conducted to decide muscle contraction rate, and the second one was to decide time variations of the muscle contractions. By using these parameters, it was confirmed that the prototype system could dynamically synthesize facial expressions and display them in real time.

Keywords : Facial Expression Synthesis, Muscle Model, Affective Interface, Appearance Time, Time Variation Pattern

1. はじめに

人間の脳の活動は、理性的な思考活動だけでなく感情を併せ持っている。近年の研究では、意思決定、理解力、学習などの認知活動において、感情が非常に重要な役割を果たしていることがわかってきた^[1]。そのため、ヒューマンインタフェースの分野においても、感情を扱うアフェクティブインタフェースの概念が導入され、その実現や応用に向けた研究が盛んになってきている。アフェクティブインタフェースはコンピュータ側から人間に感情として情報を伝える出力チャンネルと、コンピュータが人間の感情を認識・推定する入力チャンネルから構成される。筆者らは、コミュニケーションメディアの中でも、感情に直接関連のある表情に着目し、これまで後者の入力チャンネルとしての動的表情認識法の研究を行ってきた^[2]。本研究では前者の出力チャンネルとしての動的顔表情合成を研究対象とする。人間同士のコミュニケーションにおいては、表情は非常に重要な役割を果たしており、コミュニケーションの文脈の中での微妙な表情の変化やそのタイミングによって、感情を主とする多くの情報を伝達しているものと考えられる。本研究では、アフェクティブインタフェースの出力チャンネルとしての表情合成手法を対象としており、表情合成に際しては、表情の変化過程の合成、リアルタイム性、感情との対応付け易さ、自然さなどが重要である。

表情合成とは、表情の表出による顔の変形をコンピュータグラフィックス(CG)などの手段で人間に見える形として表現することである。この分野では、主に、(1)内挿法、(2)パラメータ法、(3)Performance Driven法、(4)筋肉モデル法、および(5)擬似筋肉法、の5つの基本手法^[3]があるが、本研究では、上記の観点から筋肉モデル法を用いることとした。ただし、筋肉モデル法により表情の表出過程を合成するためには、表情が自然に見えるような表情筋の収縮量やその時間変化を決めなければならない。従来の研究では、表情の動的表出に主眼が置かれており、表情筋の収縮量やその時間変化は表情合成システム設計者個人の主観により決められていた。そこで、本研究では、筋肉モデル法による表情合成システムを実際に試作するとともに、複数の被験者による主観評価実験により、表情が自然に見えるような表情筋の収縮量や時間変化を決定することを目的とする。

2. 筋肉モデルによる表情合成手法

本章では、まず、表情に関する心理学の分野における知見を述べ、次に顔表情を作り出す表情筋の概要を説明する。そして、これらの知見に基づき、人間の顔表情をCGとして表現するための「表情筋モデル」を用いた顔表情合成手法について説明する。

2.1 表情に関する心理学的知見

2.1.1 基本6表情

基本6表情は、基本6感情、すなわち驚き(surprise)、恐怖(fear)、嫌悪(disgust)、怒り(anger)、喜び(happiness)、悲しみ(sadness)の6つの感情により引き起こされる表情であり、表情に関連する研究において最も広く

*1: 京都大学大学院エネルギー科学研究科

*2: 京都大学大学院エネルギー科学研究科、現在(株)ミノルタ

*1: Graduate School of Energy Science, Kyoto University

*2: Graduate School of Energy Science, Kyoto University Present; Minolta Co., Ltd.

用いられている表情の分類法である^[4]。各表情にはそれぞれに特徴的な表情動作が存在する。なお、表情と感情は混同されやすいが、本研究では、感情は内面の心理状態であるのに対し、表情は感情によって引き起こされる顔面の変形とする。

2.1.2 表情記述法 FACS と AU

FACS(Facial Action Coding System)^[4]は Ekman らによって提案された表情記述法である。FACS では、表情を AU(Action Unit) と呼ばれる解剖学的に独立し、かつ視覚的に識別可能な表情動作の最小単位によって表情を記述する。

2.2 表情筋

表情変化は、顔面皮膚下にある表情筋、および口の閉閉を行う咀嚼筋と呼ばれる筋肉の収縮・弛緩により引き起こされる顔面形状変化である。各表情筋の配置と機能を、前述の AU に対応させると Table 1 のようになる。

表1 Action Unit と表情筋の対応関係
Table 1 Relationship between AUs and muscles.

AU	AU を引き起こす表情筋
1:眉の内側を上げる	前頭筋(内側)
2:眉の外側を上げる	前頭筋(外側)
4:眉を下げる	皺眉筋、鼻根筋
5:上脣を上げる	眼輪筋(上眼瞼部)
6:頬を持ち上げる	眼輪筋(眼窩部)
7:脣を緊張させる	眼輪筋(下眼瞼部)
9:鼻に皺を寄せる	上唇鼻翼挙筋
10:上唇を上げる	上唇挙筋
12:唇端を引張り上げる	大頬骨筋
15:唇端を下げる	口角下制筋
16:下唇を下げる	下唇下制筋
20:唇を横に引張る	笑筋
23:唇を固く閉じる	口輪筋
26:顎を下げて唇を開く	咀嚼筋*

*厳密には表情筋ではない

2.3 表情合成手法

上述の知見を利用して、感情を表現するための表情を CG として合成する過程は次の4つの段階(部分)から構成される。

- (1) 感情から顔の3領域における基本表情の組み合わせを決定する(感情-表情変換部)。
- (2) 顔の3領域における基本表情の組み合わせから各表情筋の収縮・弛緩を決定する(AU/筋肉運動生成部)。
- (3) 表情筋の収縮・弛緩によって引き起こされる顔表面の変形を計算する(筋肉モデルシミュレーション部)。
- (4) 顔表面の変形を CG として提示する(グラフィックエンジン)。

以下では、この各部について説明する。

2.3.1 感情-表情変換部

感情-表情変換部では、Table 2 に基づき、入力された感情を顔の3領域の基本表情に変換する。例えば、入力された感情が「軽蔑」の時には、顔の上部の表情が「中

立」、中央部が「中立」、下部が「嫌悪」となる。入力できる感情は34種類であり、感情の強さは考慮していない。

表2 感情から顔の3領域の表情の組合せへの変換
Table 2 Conversion from emotions to combinations of basic expressions in three regions on the face.

	感情	顔の領域		
		上部	中央部	下部
1	驚き	驚き	驚き	驚き
2	疑問	驚き	中立	中立
3	興味	中立	驚き	中立
	:	:	:	:
14	軽蔑	中立	中立	嫌悪
	:	:	:	:
33	決断	悲しみ	悲しみ	怒り
34	隠された悲しみ	悲しみ	悲しみ	驚き

2.3.2 AU/筋肉運動生成部

ここでは、感情-表情変換部から出力される顔の3領域での表情から合成すべき顔表情を表現する AU の組み合わせを生成し、さらに、これらの AU を表出するため表情筋の種類、その収縮の量と時間変化を筋肉データベースを参照して決定する。なお、筋肉データベースに格納するデータは主観評価実験によって決定するが、その詳細は次章に譲る。

2.3.3 筋肉モデルシミュレーション部

筋肉モデルシミュレーション部では、Waters によって提案された筋肉モデル法^[5]を用い、その収縮や弛緩によって引き起こされる皮膚表面の移動を模擬することによって、3次元ポリゴンモデルで示された皮膚表面のポリゴン頂点の移動方向や移動量を算出する。

表情筋は、線形筋と括約筋に分けることができる。また、表情筋ではないが、表情を表出するための重要な要素である口の閉閉によって引き起こされる下顎部皮膚表面の移動も顎の回転モデルにより算出する。Fig.1 に本研究で用いる3次元顔ポリゴンモデル上の線形筋モデルと括約筋モデルの種類と配置を示す。

線形筋モデル

線形筋とは、一方の端が骨に付着し、もう一端が皮膚に付着している線状の筋肉のことである。実際には、筋肉と骨、筋肉と皮膚とは複数の接点で付着しているが、線形筋モデルでは簡略化のためそれぞれ1点で付着しているとしてモデル化する。すなわち、線形筋 $\overrightarrow{V_1V_2}$ は骨との接点 V_1 が固定されており、その収縮や弛緩によって皮膚との接点 V_2 が V_2' へ移動し、それに影響されて周囲の皮膚表面上の点 P が P' へ変位する量を以下の式により計算する。

$$\overrightarrow{V_1P'} = \overrightarrow{V_1P} - k \cdot r \cdot \cos\left(\frac{A}{B} \cdot \frac{\pi}{2}\right) \cdot \frac{\overrightarrow{V_1P}}{|\overrightarrow{V_1P}|} \quad (1)$$

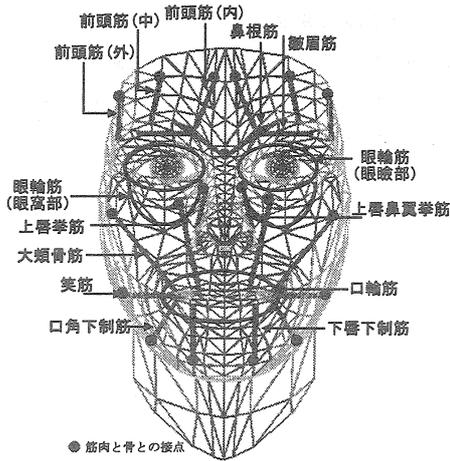


図1 3次元顔モデル上の筋肉モデルの配置
Fig.1 Arrangement of muscle models on 3D facial polygon model

$$\begin{cases} r = \cos \left[\left(1 - \frac{d}{r_s} \right) \cdot \frac{\pi}{2} \right] (P \in V_1 P_m P_n) \\ r = \cos \left(\frac{d - r_s}{r_f - r_s} \cdot \frac{\pi}{2} \right) (P \in P_n P_r P_s P_m) \end{cases} \quad (2)$$

ここで、 k は筋肉収縮率であり、以下のとおりである。

$$k = \frac{|\vec{V}_1 \vec{V}_2'|}{|\vec{V}_1 \vec{V}_2|} \quad (3)$$

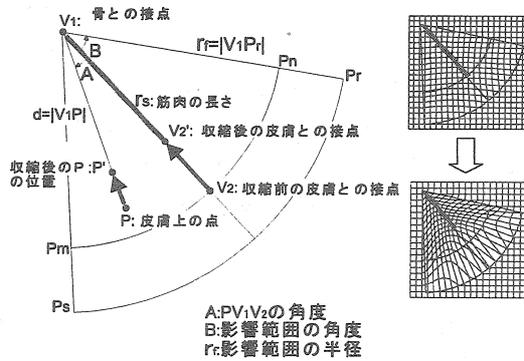


図2 線形筋モデル
Fig.2 Linear muscle model.

括約筋モデル

括約筋は皮膚との接点を持たない環状の筋肉であり、その収縮や弛緩によってその周囲の皮膚表面が移動する。本研究では、括約筋を楕円としてモデル化し、その収縮によって筋肉の周辺の皮膚の変位を求める。具体的には、Fig.3に示すように括約筋の収縮により楕円形の影響範囲内の皮膚上の点 P が中心に向かって P' に変化する量を以下の式により計算する。

$$\vec{OP}' = \vec{OP} - 2 \cdot k \cdot \left[1 - \sqrt{\left(\frac{P_x}{l_x} \right)^2 + \left(\frac{P_y}{l_y} \right)^2} \right] \cdot \frac{\vec{OP}}{|\vec{OP}|} \quad (4)$$

なお、本研究では、AU5, AU6, AU7を独立に動作させるために、眼輪筋を、上眼瞼部、眼窩部、下眼瞼部の3

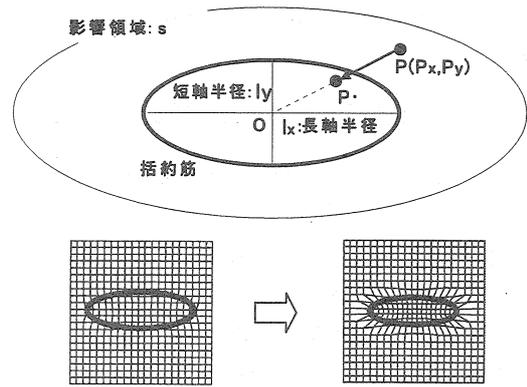


図3 括約筋モデル
Fig.3 Sphincter muscle model.

つの半楕円でモデル化している。

顎の回転モデル

口の開閉は、表情筋の収縮による顔表面皮膚の変位ではないので、前述のような筋肉モデルを用いた線形筋や括約筋をモデル化する方法では、それによる皮膚表面の移動を合成できない。そこで、下顎骨がその最上点を支点として回転することで口が開閉するという点に着目し、Fig.4に示すように、パラメータ法を用いて皮膚表面の動きを模擬する。具体的には、左右の顎関節を結ぶ線を回転中心として、以下の式により回転角 α を求めて、皮膚上の点の移動量を計算する。

$$\alpha = \begin{cases} \theta \cdot \sqrt{1 - \frac{8d^2}{9l^2}} (0 \leq d \leq l) \\ (d - L) \cdot \frac{\theta}{3(l - L)} (l < d \leq L) \end{cases} \quad (5)$$

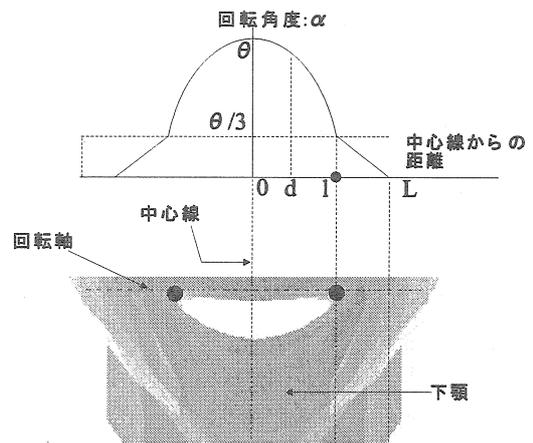


図4 顎の回転モデル
Fig.4 Jaw rotation model.

なお、皮膚上の点が多数の筋肉モデルや顎の回転モデルから影響を受ける場合には、それぞれの移動量の和を計算する。

2.3.4 グラフィックエンジン

グラフィックエンジンは、筋肉モデルシミュレーション部で算出したポリゴンモデルの頂点の座標変化を実際にCGとして表示する。

また、以上の方法に基づき表情合成システムを試作した。

3. 表情筋の収縮率と時間変化の決定

前章で述べたように、自然な顔表情の表出過程を合成するためには、その表情ごと、表情筋ごとに、その収縮率と時間変化を決定しなくてはならない。そこで、被験者による主観評価実験によりそれらを決定する。

3.1 実験1: 表情筋の収縮率と顎の回転角度の決定

本実験では、各基本表情を表出する際の各表情筋の収縮率と顎の回転角度を決定することを目的として、複数の被験者を対象に、表情が自然に見えるための表情筋の収縮率や顎の回転角度の値を決めてもらい、その代表的な値を求める。

3.1.1 実験の方法

実験には、Fig.5に示すGUIを持つ顔表情合成システムを使用した。



図5 ユーザインタフェース
Fig. 5 Graphical user interface (GUI).

手順は以下の通りである。

- (1) まず、被験者に実験の概要を説明し、次いで顔ポリゴンモデルにおける各表情筋の位置、その収縮による顔表情の変化、口の開閉等を説明する。さらに、実験で用いるシステムを事前に使用してもらい、その操作法に習熟してもらう。
- (2) 次に、被験者に基本6表情の各表情ごとに表情筋の収縮率と顎の回転角度をコントロール部のスライドバーで調整し、自然に見える表情を作成してもらう。
- (3) 被験者が十分自然に見える表情を作成したと判断した時点で、そのときの各表情筋の収縮率と顎の回転角度を記録する。

被験者は20代の男性8名および女性1名の計9名である。

3.1.2 実験の結果

実験の結果の一例として、被験者が「驚き」の表情を作成した際の筋肉の収縮率をFig.6に示す。図中のグラフは、横軸が表情筋の種類あるいは顎の回転を表し、縦軸が表情筋の収縮率(グラフ左縦軸)あるいは顎の回転角

度(グラフ右縦軸)を表している。グラフ中には各表情ごとに各被験者が設定した表情筋の収縮率と顎の回転角度を×印でプロットしてあり、その右側に平均値とその平均値からの標準偏差をエラーバーにより示している。

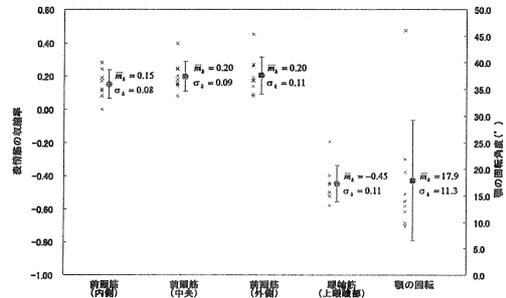


図6 「驚き」の表情時の筋肉収縮率
Fig. 6 Muscle contraction rates for expression "Surprise".

3.1.3 結果の考察

Fig.6からもわかるように、本実験で得られた表情筋の収縮率や顎の回転角度の値は、被験者によってばらつきが大きい。被験者には、実験前の説明の際に「漫画のように表情を誇張せず、人間として自然に見える」顔表情を作成するように指示していたが、実際には、被験者によっては、漫画的に誇張した表情を作成する傾向があったためである。そのため、被験者実験で得られた表情筋の収縮率や顎の回転角度を単純に平均して各表情における代表値を求めるのではなく、平均から標準偏差内の収縮率のみを用いて、それらの平均を算出することで誇張された表情の表情筋の収縮率や顎の回転角度を省いてから、それぞれの代表値を算出した。結果をTable 3に示す。

3.2 実験2: 表情筋収縮率と顎の回転角度の時間変化の決定

本実験では、基本6表情および中立の計7種類の顔表情が表出される際の各表情筋の収縮率と顎の回転角度の時間変化を決定することを目的とする。具体的には、表情表出時の時間変化パターンと表情表出時間を求める。

3.2.1 実験の方法

実験には、GUIにより表情表出時間と後述の時間変化パターンを設定できる顔表情合成システムを使用した。手順は、以下の通りである。

- (1) まず、被験者に実験の概要を説明する。ここでは、被験者に5種類の異なる時間変化パターンがあることだけを知らせ、それぞれのパターンの詳細については説明しない。用意した5種類の時間変化パターンをFig.7に示す。さらに、実験で用いるシステムを事前に使用してもらい、その操作法に習熟してもらう。

表3 基本6表情の筋肉収縮率

Table 3 Muscle contraction rates for basic six expressions.

筋肉	喜び	驚き	恐れ	悲しみ	嫌悪	怒り
前頭筋(外側)		0.17				
前頭筋(中央)		0.19				
前頭筋(内側)		0.14		0.29		
皺眉筋			0.33	0.22		
鼻根筋			0.58	0.14		
眼輪筋(上眼瞼部)		-0.48	-0.18		0.26	0.36
眼輪筋(眼窩部)	0.31					
眼輪筋(下眼瞼部)			0.65	0.38	0.43	
上唇鼻翼拳筋					0.19	
上唇拳筋					0.08	
大頬骨筋	0.19					
口角下制筋				0.28		
下唇下制筋						0.23
笑筋						0.13
口輪筋						0.46
顎の回転(°)	8.1	14.4	3.9	0.6	0.7	3.5

- (2) 被験者に実験システムを操作してもらい、表情が自然に見える表情筋収縮率と顎の回転角度の時間変化パターンと表情表出時間の上限値・下限値を決めてもらう。

被験者は20代の男性5名(A~E)である。

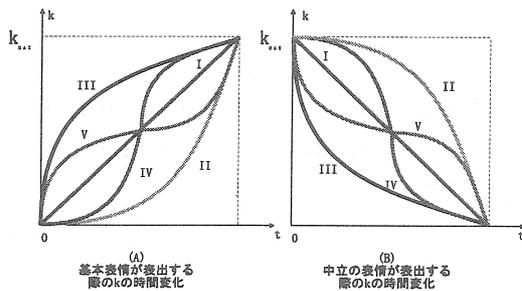


図7 筋肉収縮率の変化パターン
Fig. 7 Variation patterns of muscle contraction.

3.2.2 実験の結果

実験結果の一例として、「中立」から「喜び」の表情を表出する際の各時間変化パターンと各被験者が自然に見えると答えた表情表出時間の上限値と下限値をFig.8に示す。図では、自然に見える表情表出時間の下限値から上限値までを太線で示している。また、0.04~5秒までのいかなる表情表出時間においても表情表出が自然に見えなかった時間変化パターンについては、太線を記載していない。

3.2.3 結果の考察

実験の結果より、中立の表情から基本6表情が表出される場合は、パターンI,II,Vが不自然に見えるという被験者がいた。一方、基本表情から中立の表情が表出される場合は、パターンVが不自然に見えるという被験者がいた。このパターンVは、表情が中立に戻る際にその速度が中間で0になるものであり、実験後の報告から、「顔表情が停止する段階があるから自然」と「停止する段階があるから不自然」と逆の結果が出ている。一方、

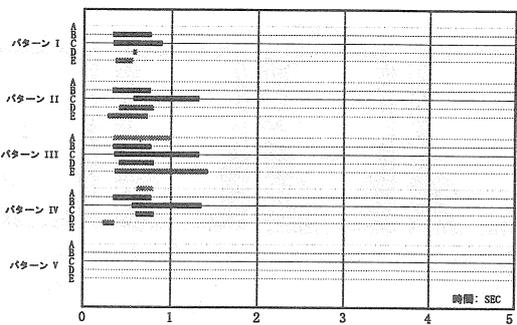


図8 喜びの表情表出時間の下限値と上限値
Fig. 8 Shortest and longest appearance time from "Neutral" to "Happiness" expression.

すべての被験者が不自然とは感じなかったパターンは、パターンIIIとIVであった。本研究では、このうち計算の容易さから、パターンIIIを時間変化パターンの代表として採用する。一方、表情表出時間の範囲(上限値と下限値)および代表値を決定する際には、各表情ごとに5人の被験者の設定した上限値と下限値について、それぞれメディアン値を求めた。また、代表値としては上限値と下限値の対数平均を求めた。結果の一部をTable 4に示す。

表4 基本表情の表出時間
Table 4 Appearance time from "Neutral" to basic expressions.

基本表情	上限値(秒)	代表値(秒)	下限値(秒)
喜び	1.00	0.60	0.36
驚き	1.00	0.40	0.16
怒り	0.88	0.50	0.28
悲しみ	1.52	0.78	0.40
恐怖	1.00	0.63	0.40
嫌悪	1.00	0.63	0.40

4. 表情合成例

以上の実験により決定した筋肉パラメータを用いて、基本6表情を合成したものをFig.9に示す。

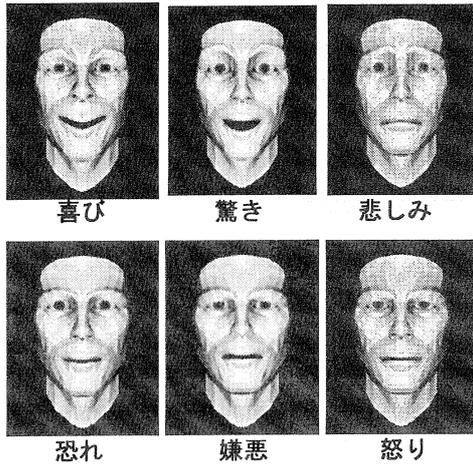


図9 基本6表情の合成

Fig. 9 Synthesis of basic six expressions.



図10 表情合成例

Fig. 10 An example of facial expression synthesis.

また、表情合成システムの使用例として、童話「やまんばの木」^[6]の内容を合成音声により音読し、それにあわせて表情を合成した例の一部を Fig.10に示す。この例では、童話の内容を再生するのに要した時間は63秒であり、その間に7回の表情変化がある。また、数名の男子学生に、童話を音声で再生しながら、顔表情合成の経過を観察してもらい、その感想を聞いたところ、以下のような感想が寄せられた。

- とても面白かった。
- 気持ち悪いほどよく動く。
- 怒りなどの感情は、童話の音声再生がなければ、理解できなかったと思われる。童話の音声再生を伴わせることで、その表情から感情が伝わってきた。
- 1分間に顔表情が7回も変化して、おかしい。
- 視線の移動がなく、見つめられている感じがして、不自然に感じた。

なお、これらの例では、表情表出後の顔画像しか示していないが、実際にはその表出過程や消失過程もリアルタイムに合成されていることを確認した。また、SGIのO2ワークステーションを用いて表情合成を行う際のフレームレートは25フレーム/秒以上であり、十分滑らかな動画像として表示できている。

5. ま と め

本研究では、顔表情を用いたアフェクティブインタフェースの出力チャンネルとして、顔の筋肉モデルに基づいて、顔表情の表出過程をCGの動画として合成する表情合成システムを試作し、被験者実験により表情表出時の表情筋の収縮率と時間変化パターンを決定する方法を提案した。

表情の筋肉モデルは、3次元顔ポリゴンモデル上に配置された18本の線形筋モデル、7個の括約筋モデルおよび顎の回転モデルから構成される。表情を表出する際には、各モデルにより筋肉の収縮を模擬し、ポリゴンモデルの頂点の変位を計算し表示する。さらに、自然な表情の表出過程を実現するために2つの被験者実験を行った。最初の実験では、基本6表情を表出した際の筋肉の収縮率と顎の回転角度を求めた。次の実験では、表情表出時あるいは消失時の時間や収縮率の変化パターンを求めた。そして、実験で求めた筋肉パラメータを用いて表情を合成した例を示し、本研究で試作した表情合成システムがリアルタイムで自然な表情を合成できることを確認した。

今後の課題としては、(1)より複雑な感情や感情の強さを表情として表現する方法の検討、(2)瞬目や視線の動きを合成する方法の検討、などが挙げられる。そして、さらに表情を用いたアフェクティブインタフェースの実現のために、人間とコンピュータとの感情交流について研究を進めていきたい。

謝 辞

本研究は、日本学術振興会、平成9年度未来開拓学術研究推進事業(JSPS-RFTF97I00102)の一環として遂行したことを附記し感謝の意を表する。

参考文献

- [1] Rosalind W. Picard: Affective Computing, The MIT Press (1997).
- [2] 下田宏, 國弘威, 吉川榮和: 動的顔画像からのリアルタイム表情認識システムの試作, ヒューマンインタフェース学会論文誌, Vol.1, No.2, pp.25-32 (1999).
- [3] Frederic I. Parke: Techniques for facial animation, New Trends in Animation and Visualization, pp.229-241 (1991).
- [4] Ekman and Friesen 著, 工藤訳編: 表情分析入門, 誠信書房 (1987).
- [5] K. Waters: A Muscle Model for Animating Three-dimensional Facial Expression, Computer Graphics, SIGGRAPH'87, Vol.2, No.4, pp.17-24 (1987).
- [6] 木暮正夫: やまんばの木, 佼成出版社 (1989).