

# 顔情報処理のための 共通プラットホームの構築

八木康史（大阪大学） 森島繁生（成蹊大学） 金子正秀（KDD）  
原島 博（東京大学） 谷内田正彦（大阪大学）  
原 文雄（東京理科大学） 橋本周司（早稲田大学）

顔画像処理に対する様々な分野での関心の高まりや、工学分野における顔画像処理技術の研究成果の蓄積を背景にして、顔画像処理に関する共通ソフトウェアのツールの作成に向けた活動が進められている。この活動は、「感性擬人化エージェントのための顔情報処理システムの開発」（略称、アドバンストエージェントプロジェクト）と呼ばれ、情報処理振興技術協会（IPA）における独創的情報技術育成事業に関わる開発テーマの一つとして、平成7年度より3年間の計画で精力的に活動を行ってきた。擬人化エージェント技術はさまざまな技術要素から構成されているが、本プロジェクトでは、この中で特に「顔」の役割に着目し、顔画像の認識・合成に関わる顔情報処理システムの開発に主眼をおいた。これと同時に、本システムでは工学のみならず心理学や医学などの分野も含めた顔関連分野における共通の実験用ツールを広く提供することも目標としている。本稿では、平成10年3月で終了するこのプロジェクトの概要と、共通ソフトウェアの紹介を行う。

## Facial Image Processing Environment

Yasushi Yagi (Osaka University)  
Shigeo Morishima (Seikei University)  
Masahide Kaneko (KDD)  
Hiroshi Harashima (The University of Tokyo)  
Masahiko Yachida (Osaka University)  
Fumio Hara (Science University of Tokyo)  
Shuji Hashimoto (Waseda University)

The aim of Advanced Agent Project, supported by Information Technology Promotion Agency (IPA), is to develop the image processing environment for analysis and synthesis of human facial images. This report is mainly concerned with an introduction of an overview of the project and the developed environment for the facial image processing.

## 1. はじめに

顔画像処理に対する様々な分野での関心の高まりや、工学分野における顔画像処理技術の研究成果の蓄積を背景にして、顔画像処理に関する共通ソフトウェアのツールの作成に向けた活動が進められている。本稿では、平成10年3月で終了するこのプロジェクトの概要と、共通ソフトウェアの紹介を行う。

## 2. Advanced Agent Project

### 2.1 プロジェクト概要

顔画像の処理に関する共通ソフトウェアツール作成に向けたこの活動は、「感性擬人化エージェントのための顔情報処理システムの開発」(略称、アドバンストエージェントプロジェクト)と呼ばれ、情報処理振興技術協会(I.P.A.)における独創的情報技術育成事業に関わる開発テーマの一つとして、平成7年度より3年間の計画で精力的に活動を行ってきた。そして現在そのプロジェクトのファイナルステージを迎えている。

擬人化エージェント技術はさまざまな技術要素から構成されているが、本プロジェクトでは、この中で特に「顔」の役割に着目し、顔画像の認識・合成に関わる顔情報処理システムの開発に主眼をおいた。これと同時に、本システムでは工学のみならず心理学や医学などの分野も含めた顔関連分野における共通の実験用ツールを広く提供することも目標としている。

### 2.2 研究開発体制

東京大学、大阪大学、東京理科大学、早稲田大学、成蹊大学、〔財〕イメージ情報科学研究所を中心として研究グループを組織し、感性擬人化エージェントの実現に向けての研究方針の策定、研究動向調査等から開始し、実際に顔情報処理システムの開発を行うワーキンググループを組織して研究を進め、平成10年3月をもって標準ソフトウェアとしての開発を完了する。

本プロジェクトは、顔認識、顔合成、顔印象の3つのサブテーマから構成される。特に、顔認識、顔合成は基本ソフトウェアシステムを構成し、顔印象はこの基本ソフトを利用した応用分野として位置づけできる。

## 3. 顔の認識・合成システムの概要

### 3.1 顔認識システム

顔認識システムの目的は、顔の形状や表情・動きを認識するための基本ツールを提供することにある。す

なわち顔の3次元形状モデルを入力画像中の人物の顔部分に整合させるために必要な顔特徴の高精度な抽出である。これは顔の静的な特徴のうち、顔造作の位置や輪郭特徴をほぼ実用に耐えるレベルで抽出することを可能にする。

以下に処理内容の概要を示す。

#### 1) 顔の切り出し

色情報に基づいて肌色領域と髪領域を抽出する。

#### 2) 瞳中心点の算出

まばたきに伴う瞳部分の色の変化から瞳の位置を確定する方法と、瞳・目蓋・眉にかけての鉛直線上における輝度値分布を利用するこにより、瞳中心点の安定した抽出を実現している。

#### 3) 顔造作の概略位置決定

瞳中心点の位置を基準にして、眉、目、鼻、口の各顔造作が含まれる方形領域を推定する。さらに、各方形領域においてエッジ強度分布を調べ、各造作位置の概略位置を決定する。

#### 4) 顔造作の輪郭抽出

動的輪郭モデル(スネーク)を利用して、各顔造作の輪郭を抽出する。他に瞳を中心として放射状に目付近の画像を展開する方法がある。

#### 5) 顔輪郭の抽出

左右の瞳の位置関係を基準として、数十人の顔画像から得られた顔輪郭の平均位置と分散を利用して顔輪郭が存在する候補領域を設定する。この候補領域内で動的輪郭モデルを適用して顔輪郭の抽出を行う。

### 3.2 顔合成システム

顔画像をもつ擬人化エージェントを実現する際には、表現力豊かな顔動画像を合成できることが重要である。リアリティを持たせるためには、実写画像を基にしてかつGeometry構造変形を伴う顔画像の生成が優れている。そこで、多数の小さな三角形パッチ群から構成される人物頭部の3次元形状モデルを標準モデルとして利用する。またリアルな質感は、実写の写真から得られるテクスチャを3次元形状モデルの表面にテクスチャマッピングすることにより表現する。

しかし、異なる人物の顔画像合成する際に、個々の人物に最適な頭部モデルを毎回作成するわけにはいかないため、標準モデルをGeometry整合することにより個人のモデルを近似的に作り出す。厳密に言えば2次元の情報のみ個人適応して、奥行き情報は標準値のままとなるが、大きく向きを変えない擬人化エージェントとしての利用には十分耐えると思われる。もちろん、横方向の画像を用いて奥行き整合したり、3次元レンジデータを利用する等によって奥行きの正確な推定も不可能ではないが、今回のツールにはこの機能は

組み込まれていない。

顔合成システムは3つのサブシステムから成り立っている。この3つとは、顔整合ツール、表情編集ツール、アニメーション編集ツールである。

### 1) 顔整合ツール

カメラ入力された顔正面画像に対して標準ワイヤフレームモデルを整合させ、個人モデルを作成するツールである。

### 2) 表情編集ツール

表情を記述する基本単位(Expression Controller)の強度を変化させることによって、望みの表情をマウスを用いて編集するツールである。この表情記述ユニットは Ekman による FACS と類似しているが、顔部位の分類によって独自に設定している。

唇部分に関しては、別途口形パラメータを定義しており、この組み合わせによって任意の口形を編集することができる。

### 3) アニメーション編集ツール

表情編集ツールで編集された顔画像をキーフレームに配置して、表情アニメーションを作成するツールである。

## 4 顔認識システムの紹介

本システムは、顔画像処理のための標準ライブラリ並びにグラフィカルユーザインターフェイス(GUI)からなる。標準ライブラリには、顔の発見、その輪郭抽出並びに、眉・目・鼻・口といった顔器官の発見とそれらの輪郭抽出のための各種関数が用意されている。本システムは、画像処理研究者のような基礎研究者から、自動的に顔の特徴が計測できればよい応用・開発研究者に至るまで、幅広いユーザを想定し、各ユーザのレベルに応じた開発環境をインテラクティブに操作できるGUI上で実現している。

### 4.1 GUI

本システムでは、4つのレベルのユーザを想定している。第1は、自動的に顔の特徴が計測できればよい応用・開発研究者を対象とした、簡易自動顔認識手法である。この手法を用いることで、画像処理に関する知識はなくとも、顔器官並びに顔輪郭を自動で抽出することができる。なお、自動認識の全体概要に関しては、後述する。

第2レベルは、カスタマイズ機能の実現で、ユーザが各ステップごとで標準ライブラリから適当な処理方法を選択し、自由にユーザの目的に合わせた顔画像処理方法を構築できる。図1にしめすように入出力関係から選択可能な次の処理方法の一覧を提示、それらか

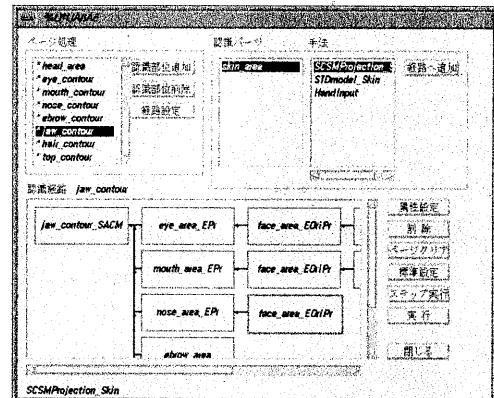


図1 認識システムのGUI画面

ら逐次、処理方法を選択することで、ユーザの希望する処理シーケンスを造りだすことができる。さらに第3レベルでは、顔画像処理の基礎研究者のために各関数の内部パラメータを自由に変更できるパラメータチューニングウインドウを用意している。カスタマイズ機能における処理手順の選択では、ライブラリ内のすべての関数間の入出力関係を処理経路定義ファイルで管理している。従って、基礎研究者を対象とした第4レベルでは、ユーザが新たに作成した関数を組み込みたい場合、この処理経路定義ファイルを更新し、再コンパイルによりユーザ定義関数をリンクすることでも容易に実現できる。

### 4.2 標準ライブラリ

顔認識システムの最終的な目的は、どんな画像中からでも顔が抽出できると共に、眉・目・鼻・口といった顔器官と顔の輪郭が確実に抽出できることである。そこで、本ライブラリでは、顔の発見、瞳の抽出、各顔器官の検出に加え、顔の輪郭並びに各顔器官の輪郭形状を抽出する関数が32個用意されている。図2にこのうちの4つの処理結果を示す。

ところで、一般に研究者のニーズは多岐に渡り、利用環境は一様とならない。従ってすべての利用環境で、動作可能な統一的手法を用意することは不可能と言える。そこで本ライブラリでは、同一の目的に対し、動作環境など仕様が異なる複数の手法を用意することで、広いニーズへの対応を意識している。以下、本ライブラリで用意した代表手法について説明する。

#### 1) 顔領域の切り出し

画像からの顔の抽出に関しては、肌色情報を利用することにより顔領域を切り出す方法を提案した。まず、色情報を人間の視覚特性と一致した色空間で表現し、

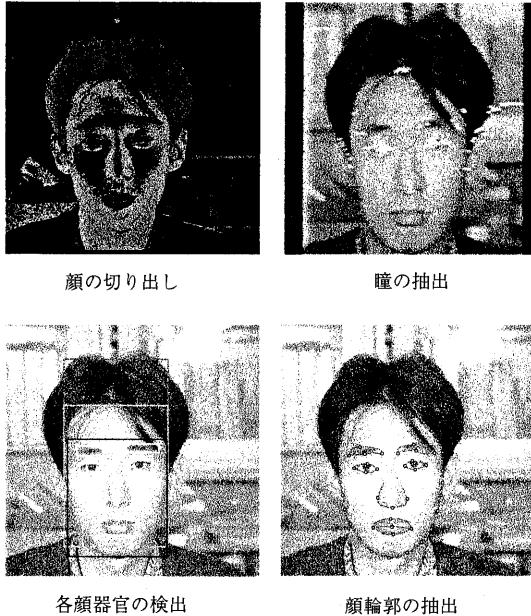


図2 顔画像処理標準ライブラリの例

その色空間内における肌色の分布領域を参考にして画像中から肌色領域を抽出する。次に、安定した顔領域の抽出のために、同様にして髪領域を抽出する。画像中の顔のサイズと位置が未知なので、前もって用意された正面、右向き、左向きの3種類の顔形状のモデルを大きさを変化させながら画像にマッチングさせ、ファジィパターンマッチングにより肌色領域と髪領域をもとに最終的に顔領域を抽出する。この結果、複雑な背景を持つ屋外シーンの場合でも、また、画像中に顔が二つ以上ある場合でも顔を正確に抽出できる。

## 2) 瞳中心点の抽出

続いて、顔の造作である顔器官を抽出する方法として、左右の瞳を抽出した後、その瞳の位置関係から各顔器官が確実に含まれる領域を特定し、それぞれの領域内から顔器官を抽出する。瞳の抽出に関しては、2つの手法を用意している。

1番目は、目の動きの時間的特徴である瞬きを利用する方法である。黄色人種の場合、瞳部分は黒色のに対し、瞼は肌色である。このことを利用し、瞬きをしたときの瞳部分の色の違いから、差分処理により瞳の位置を推定する方法である。

2番目は、目の空間的特徴を利用する方法で、顔の造作において、眉と瞳は明らかに肌色とは異なる。そこで、眉と瞳の色と位置関係を用いて画像中から瞳を

抽出する手法を確立した。すなわち、瞳から鉛直線上に眉に至る画像における瞳の黒、瞼の肌色、眉の黒という輝度値の分布は、顔が多少回転したり多少大きさが変わっても変化しない。そこで、この瞳-瞼-眉の輝度値の分布パターンと画像中の輝度値分布の相互相關値を計算することにより瞳の位置を確定した。顔画像の大きさに合わせて瞳-瞼-眉の輝度値の分布パターンの大きさを変化させることにより、照明条件がとのえば、背景に依存せず最終的に左右瞳を抽出できる。

## 3) 各顔造作の輪郭抽出

そして左右の瞳を抽出した後、この左右瞳の位置を基準にして、被験者に依存せず、かつどんな表情をしても確実に顔の各器官が含まれる領域を、30人の被験者の6基本表情をもとに算定し、それぞれの領域内から顔器官を抽出する方法を用意した。具体的には、眉・目・鼻・口の各顔器官が含まれる領域から、それぞれの器官の特徴に応じた手法により輪郭を抽出した。

眉は、その生え際が鮮明でないことが多く、肌色と眉の黒色との境をエッジ検出処理だけで獲得することは困難である。そこで、まず横方向のエッジ検出を行って眉のおおよその上下端を決定し、その内部を2値化処理することにより眉の輪郭を決定した。目と肌の境は、そのつながりが連続でないために明らかに暗くなっている、輝度値の低い部分が輪郭に相当する。

瞳を中心放射状に目の画像を開くと、目の輪郭である暗い部分は下に凸の2つの3角形の辺として現れるので、それぞれの辺を放物線で近似することにより目の輪郭を抽出した。

鼻は、小鼻の輪郭に沿って暗い部分が確実に現れるので、エッジ強調とテンプレートマッチングにより鼻の位置を限定し、相互関係により鼻の輪郭である暗い部分を抽出した。口の場合は、エッジ処理した画像からおおよその位置を限定し、その中を2値化することにより口の中心を求める。

次に、口中心から放射状に画像を開くと、目と同様に輪郭部分が下に凸の2つの3角形の辺として現れるので、それを2次近似して口の輪郭を抽出した。

他の手法としては、動的輪郭モデルを利用する方法で、曲率変化の大きな部分で各顔器官の輪郭形状を分割し、分割された滑らかな輪郭部々々に対し、滑らかさを基準とする動的輪郭モデルを用意することで、顔器官の輪郭形状を抽出した。

次に各顔器官の位置関係を用いて顔の輪郭抽出も行った。一般に顔は曲面物体と考えることができるため、鮮明な顔の輪郭線がいつも現れるとは限らない。また皺やほくろなどの顔器官以外の特徴があり、一般

にそれらの特徴は、個人同定などの目的では有用であるが、顔輪郭や顔器官の抽出には難音になる。従つて、顔輪郭検出においては、顔の特徴をとらえた何らかのモデルを利用する方法が有用で、本手法では、Kass らの提案した動的輪郭モデルを利用し顔輪郭並びに顔器官輪郭を自動抽出する。動的輪郭モデルは、モデル変形のための拘束条件とモデルの初期位置により収束状態が変わる。ここでは、顔器官を囲む多角形を動的輪郭モデルのための初期位置とする。

そしてモデル変形のための拘束条件に関しては、平均輪郭モデルを用意する方法と左右対称性を利用する方法を検討した。平均輪郭モデルを用意する方法では、予め人の輪郭形状を統計的に求め、平均輪郭モデルを用意し、平均輪郭モデルからのずれをモデル変形のための制約として利用する。左右対称性を利用する方法では、人間の顔が有する大局的な特徴である左右対称性を利用し、モデル変形のための拘束条件に非軸固定型の軸対称性拘束をいれる。そして、本手法では、これらの条件を利用し、動的輪郭モデルへ変形することで、最終的な顔輪郭線を抽出できた。

#### 4.3 簡易自動顔認識手法

本手法は、標準的には照明条件のもとで顔の輪郭形状並びに目・鼻・口といった顔器官の輪郭形状を、入力顔画像（正面顔）から自動的に抽出する。基本的には、上記ライブラリ関数の中なら、動作環境等を考慮し、一連の処理を選択したものである。図3に処理の流れを示す。第1ステップでは、顔並びに顔器官の概略位置の基準となる瞳の発見を行う。この瞳発見は、瞳-瞼-眉の輝度値の分布パターンを用いる方法により行う。第2ステップでは、各器官の標準存在領域の情報を用いて推定する。次に第3ステップでは、顔並びに顔器官の輪郭形状を抽出する。眉並びに

目の輪郭は、動的輪郭モデルにより抽出する。鼻に関しては、相互相關による鼻輪郭抽出方法を用いる。口に関しては、口中心から放射状に画像を展開する手法を用いる。上記の処理は、標準ライブラリに含まれている処理をそのまま利用している。最後に、顔輪郭に関しては、標準ライブラリの2つの輪郭モデル（軸対称モデルと平均輪郭モデル）を選択的に利用する手法で、正確な輪郭形状を抽出することができる。以上の一連の処理を経て、顔器官や顔の輪郭などを抽出できれば、顔合成システムにおいて必要とされる顔画像の3次元モデルのフィッティングを自動的に行うことができると共に、ヒューマンインターフェースの分野における重要な課題である個人識別や、表情・感情認識等の実用化を可能にするものと考えられる。

### 5. 表情合成システムの紹介

本システムは、顔合成システムとグラフィカルユーザインターフェース(GUI)から構成されている。顔合成システムは3つのサブシステム、顔整合ツール、表情編集ツール、アニメーション編集ツールから構成される。シリコングラフィックス上で開発されていたシステムは、Windows上にも移植され、身近なパソコン環境でも手軽に利用することができるようになった。本システムは、主としてアプリケーション分野の研究者の使い勝手を想定し、マウス1つで容易に操作できる環境を実現した。さらに表情記述パラメータを独自のものに入れ換えることによって、表情のクオリティを好み合わせて適応化させることも可能で、汎用性とともに拡張性も考慮したシステム構成となっている。

#### 5.1 整合ツール

カメラ入力された顔正面画像に対して標準ワイヤフレームモデルを整合させ、個人モデルを作成するツールである。このツールでは顔の特徴点をマウスによって移動させて、序々にモデルを整合してゆく。まず顔の輪郭上の4点を合わせ、次に目、眉、口、鼻の輪郭の4点を合わせ、さらに細かく任意の輪郭を合わせるというように、マクロな整合からミクロな整合へと進行させることができる。点の移動はマウスによって行い、ワイヤフレームの変形結果および合成画像は画面上でインタラクティブに確認できるので、整合の進行具合を確認しながら操作は容易に行える。またミクロな整合を行う際には、画面を拡大表示してさらに精度を上げることも可能である。

この整合ツールは、前記の認識システムによる特徴点抽出結果を受け取り、微調整および修正をマニュアルで行う環境をも提供している。

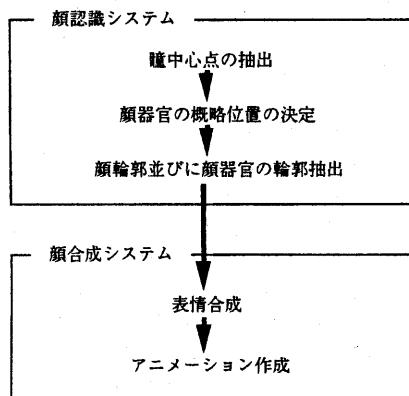


図3 簡易自動顔認識手法

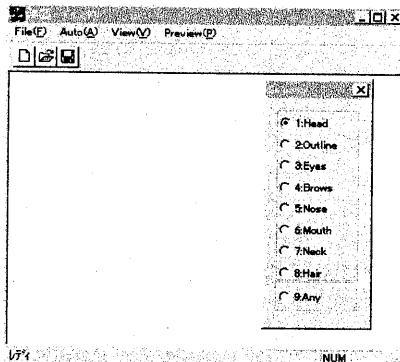


図4 整合ツール起動画面

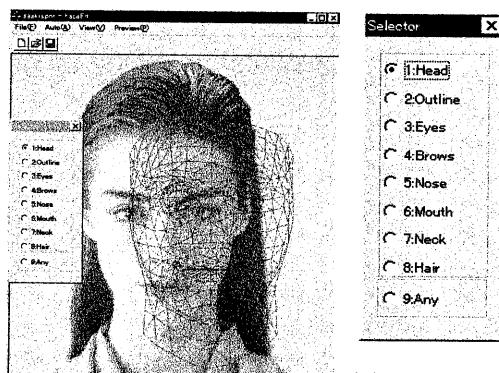


図5 顔画像の読み込み

このツールを使用するためには、まず、ビットマップ(拡張子：bmp)形式のファイルフォーマットで、二次元顔画像を用意する。次に、整合ツール(Facefit.exe)を起動する。起動すると、図4が表示される。メニューに従ってビットマップ形式の顔画像を読み込むと同時に、図5のように、3次元顔モデルの標準データが顔画像に重なって表示される。その後、図6のセレクタを用い、整合したい顔の各部位(1.頭、2.顔の輪郭、3.目、4.眉、5.鼻、6.唇、7.首、8.髪)を選択し、部位毎に定められた3次元モデルの制御点を移動させることで、顔画像にモデルを整合させる。

この整合の作業の最中に、図7に示すような別のPreview Window上で3次元モデルへテクスチャマッピングした様子を表示することができる。この画像は移動、拡大、縮小、回転を行なうことができるので、整合の具合をインタラクティブに確認することが可能である。この移動、拡大、縮小、回転の操作はマウスにより行なうことが出来る。

整合は、はじめ各部位の4点程度を決定してから、さらに輪郭の細かい点を決定する。顔の各部位の整合

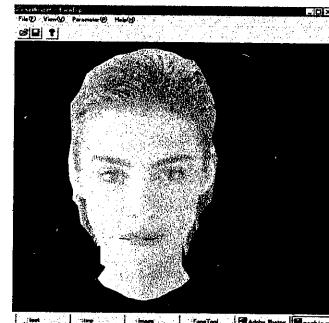


図7 Preview 画面

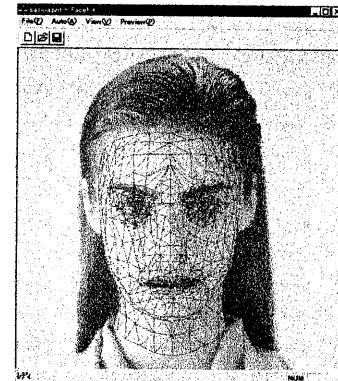


図8 整合結果表示

が終了したらメニュー選択とファイル名の入力により、モデルデータ(拡張子:pnt)を保存する。図8に整合結果を示す。

## 5.2 表情合成ツール

表情を記述する基本単位(Expression Controller)の強度を変化させることによって、表情を編集するツールである。このユニットは、顔部品毎に移動させる部位と方向、強度を指定する。強度の指定はコントロールバーをマウスで操作することにより行える。また合成画像はインタラクティブに画面で確認できるため、所望の表情を容易に編集することができる。また標準的なデータとして6基本表情のパラメータも提供する。

唇部分に関しては、別途口形パラメータを定義しており、この組み合わせによって任意の口形を編集することができる。これによりたとえば音声の研究者が日本語や英語をはじめとする言語発声の際の基本口形を編集することも可能である。基本データとして、日本語の母音、子音データも含まれている。

表情合成ツール(Faceexp.exe)を起動し、整合ツールで作成した顔モデルのデータ(拡張子:pnt)を読み込み

表情編集操作を行う。整合ツールの Preview の際と同様、マウス操作により移動、拡大、縮小と回転を行って合成された表情を任意の方向と大きさでインテラクティブに確認することが出来る。

図 9 に示す Expression Controller の個々の強度をコントロールして、顔の各部位を動かし、様々な表情を作成することが出来る。また、作成した Expression パラメータの値を保存して後に述べるアニメーション作成ツール等で再利用することができる。また、6つの基本感情(anger, disgust, fear, happiness, sadness, surprise)の Expression パラメータの組み合わせを標準値として用意している。図 10 にこの表情合成ツールにより作成された怒りの表情合成結果と各ユニットの強度を示す。

口形に関しては図 11 の Mouth Controller を用いて、唇の周辺の 9 個の制御点を動かして口形を作成する。Mouth パラメータの標準値として、英語の発音記号に基づいた 24 種類の口形(9 母音、15 子音)が用意されている。

### 5.3 アニメーション作成ツール

表情編集ツールで編集された顔画像をキーフレームに配置して、表情アニメーションを作成するツールである。アニメーションシナリオはスクリプト形式で記述しており、時間軸上のキーフレーム位置の変化によ



図 9 Expression Controller

り表情の変化のスピードをコントロールすることができる。また表情と口形は独立に取り扱うことができるため、会話のシーンに表情を挿入するようなアニメーションも可能である。このアニメーションシーンは画面上で確認することができる。

このアニメーション作成ツールでは、まず、スクリプトファイルを読み込み、アニメーションの各フレームの 3 次元顔モデルデータとして出力する。この全フレームのデータを「Facemovie.exe」を用いてテクスチャマッピングを行なうことで、アニメーションが完成する。

簡単なシナリオ記述により、基本的な表情の組み合わせでアニメーションを作成することができ、また、発話時のテキスト情報をシナリオに組み込むことで、会話のシーンの口形変化も自動作成することが出来る。基本的な表情や口形はあらかじめ標準値として用意されている。それら以外に表情合成ツールを用いて

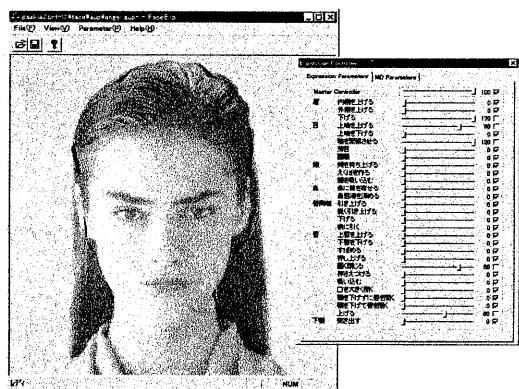


図 10 表情合成結果「怒り」とパラメータ

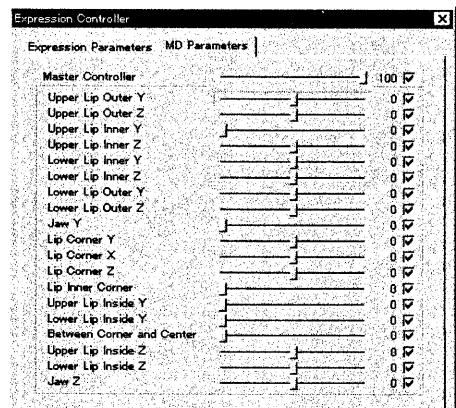


図 11 Mouse Controller

新たな表情や口形を作成し、シナリオ用のデータとして用いることは可能である。

シナリオを記述するスクリプトファイルのサンプルを以下に示す。

```
1:debug=off;
2:
3:text = "ko(happiness:1.0:25:25)nnnitha.";
4:#text = ",,(happiness:1.0:15:15),,,,";
5:
6:outfile="/home/eishi/aawg/script/demo/thank";
7:modelfile="/home/eishi/tmp/saskia.pnt"
8:linkfile="/home/eishi/aawg/script/model/face.lnk";
9:gmodelfile="/home/eishi/aawg/script/model/face.pnt";
```

1行目：デバッグモードの制御

2行目：空白

3行目：シナリオの本体

4行目：コメント文

5行目：空白

6～9行目：各データファイルの指定

このシナリオの中身は、幸福の表情で「こんにちは」と発話するシーンを合成する。”こ”の発声直後が表情のピークとなり、この時の幸福の表情の強度が1で、立ち上がり25フレーム、たち下がり25フレームであることを意味する。

実際の動画を確認するには、Facemovie.exeを起動し、シナリオに基づいて作成されたアニメーションデータを読み込む。その時の状態を図12に示す。その後、図13のPlayer Controllerを用い、アニメーションの再生や停止などを行なう。

整合ツールや表情合成ツールのときと同様にして、マウスを用い、拡大や、回転を行なうことで、様々な大きさや角度で見ることが出来るように配慮されている。

## 6. まとめ

本稿では、アドバンストエージェントプロジェクトの成果である、顔認識・合成ソフトウェアについて紹介した。このプロジェクトは平成10年3月に終了し、完成したソフトウェアはフリーウェアとして公開される予定である。

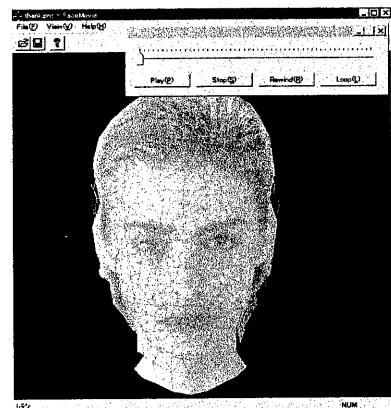


図12 Animation の読み込み

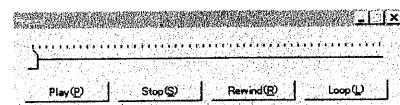


図13 Animation Controller

## 参考文献

- [1] 原島 博 他17名，“感性擬人化エージェントのための顔画像情報処理システムの開発”，平成8年度テーマ別年次総括報告書，IPA, March 1997.
- [2] 長谷川 修、森島繁生、金子正秀，“顔”的情報処理”，電子情報通信学会論文誌A, VOL.J80-A, No.8, pp.1231-1249, August 1997.