

モデル・テクスチャ・皺のブレンディングによる顔表情の表現

Facial Expression Synthesis by Blending Geometry, Texture and Wrinkles of Base Models

高橋 光紀

森島 繁生

Mitsunori TAKAHASHI

Shigeo MORISHIMA

成蹊大学工学部

Faculty of Engineering, SEIKEI University

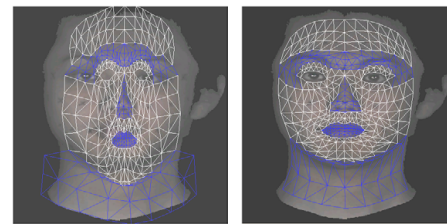
1. 概要

顔表情の表現法として、一般に FACS(Facial Action Coding System)が用いられるが、本稿では皺などの詳細な表現を、効果的に行う方法を提案する。皺などの情報を持った顔モデル群(基本顔)を用意し、それらのブレンディングによって任意の顔表情を構築する。さらに、テクスチャについてもブレンディングすることで、顔表面の詳細な特徴の表現力を飛躍的に向上させた。また応用として、VISEMEの口形モデルを構築し、それらのブレンディングによって、発話口形の表現を行うことも試みている。

2. 顔面モデルの作成

人物の顔は、3D ワイヤフレームモデルに、顔画像をテクスチャとしてマッピングすることで表現する。ここでは、この論文における顔面モデルの作成方法について説明する。

まず、一般的な顔形状をした標準ワイヤフレームを用意する。また、テクスチャとして使用するための顔画像を、デジタルカメラを用いて撮影する。この顔画像に標準ワイヤフレームを整合することで、個人の顔形状を持つワイヤフレームを作成する。また、この整合によって、テクスチャ座標を取得できるので、個人の顔面モデルを作成することが可能となるが、この顔モデルの奥行き情報は標準ワイヤフレームのものであり、個人のものではない。そこで、NEC製のレンジファインダ Danae-R を用いて、その人物の顔のレンジデ



整合前

整合後

図1 ワイヤフレーム整合

ータを取得する。取得したレンジデータとワイヤフレームとの対応を取るために、図1のように、レンジデータとワイヤフレームの整合を行い、ワイヤフレームの各頂点における奥行き情報を取得する。これで、完全な個人の顔モデルを作成することができた。

3. 顔面モデルのブレンディング

顔表情を作成する研究として、Ekmanら^[1]が提唱したAUを用いた研究が多く行われている。AUは任意の人物に適用することができるので、汎用性に優れている。例えば、顔モデルにAUを適用して変形させるような場合、AUによる顔モデルの変形ルールを作成してしまえば、任意の人物の顔モデルに変形ルールを適用し、表情モデルを作成することができる。しかし、そのために、個人に依存するような特徴を表現することは困難であった。特に皺の表出の仕方などは個人によって大きく異なるものである。そこでここでは、特に皺のような詳細な特徴を失うことなく、顔表情を作成する手法を提案する。

その手法とは、様々な簡単な表情変形をした顔

モデルを複数用意し、それらをブレンドすることによって、顔表情を持ったモデルを作成する方法である。ここで用意する顔モデル群は基本顔モデルと呼ばれる^[2]。基本顔モデルには皺などの詳細な特徴も含まれている。

ここでは、基本顔モデルとして、11種類の顔を定義している。これらの基本顔は、AUを参考にしながら、皺の表出しやすさを考慮して選ばれたものである(図2)。基本顔ひとつひとつについて、前節で説明した方法で顔モデルを作成し、無表情の顔モデルとブレンドをすることで表情変形を表現する。ブレンドとは、各基本顔にブレンド率という重みを0.0~1.0の範囲で定義し、それに従い顔モデルを混ぜ合わせることである。実際には下式を用いて顔面ワイヤフレームモデルの頂点ベクトル v を計算する。

$$v = v_n + \sum (v_i - v_n) R_i \quad (\text{式 1})$$

ここで、 v_n は無表情、 v_i は各基本顔の頂点ベクトルである。 R_i は各基本顔のブレンド率である。

これにより、定義した基本顔に任意のブレンド率を設定することで、任意の表情を表現できるようになった。ブレンドにより作成された顔面表情モデルは、皺などの詳細な特徴も失われることなく表現されているはずである。

4. テクスチャのブレンド

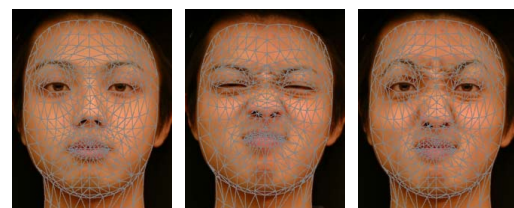
個人に依存するような詳細な特徴は、皺などに多く含まれていると考えられる。しかし、使用している標準ワイヤフレームの頂点数では、皺によってできる顔表面上の凹凸を表現することができないのが現状である。そこで、そういった細かい情報については、テクスチャによって補うこととする。従来の表情作成手法の多くは、表情を変形しても、無表情の顔画像のみをテクスチャとして用いている。しかし、皺などは表情を作った場合に表出する場合が多いので、ここでは、全ての



図2 11種類の基本顔と無表情

基本顔の画像を使用することとする。これにより、表情変形後の顔表面の情報も、失うことなく利用することができるようになる。

テクスチャについても、顔面モデルと同様に、ブレンドによって任意の表情に対応したテクスチャを作成する。各基本顔画像の画素値をブレンドするのだが、撮影した顔画像は顔の大きさや向きなどが一定でないだけでなく、基本顔はそれぞれ顔の一部が変形してしまっている。このままでは単純に画素値のブレンドを行うことができない。そこで、整合されたワイヤフレームモデルを用いて、全ての基本顔について、無表情顔テクスチャと各顔部位の画素位置が同じになるように正規化を行う(図3)。(A)は無表情顔のテクスチャ画像にワイヤフレームを重ねたものである。(B)は基本顔の一つであるが、このテクスチャを無表情のテクスチャに対して、正規化したものが(C)である。



(A) (B) (C)

図3 テクスチャの正規化

これでテクスチャのブレンドが可能となったが、ここではブレンド手法として 3 種類の方法を説明する。

まず単純に RGB 値を用いたブレンド方法を説明する。ブレンド後の画素値 r は、各基本顔の画素値にその基本顔のブレンド率を掛け合わせたものの総和を、ブレンド率の総和で割ったものとなる。これは下式を用いて算出する。

$$r = \frac{\sum r_i R_i + r_n (1 - R_{MAX})}{\sum R_i + (1 - R_{MAX})} \quad (式 2)$$

r_n は無表情画像の画素値、 r_i は各基本顔画像の画素値、 R_i は各基本顔のブレンド率である。また最大のブレンド率を R_{MAX} とし、無表情のブレンド率を $(1 - R_{MAX})$ と定義している。

しかし、この方法を用いると、複数の基本顔のブレンドを行った場合に、各基本顔の特徴が薄れてしまうという問題が起きてしまう。これは例えば、皺がないという特徴すらもブレンドされてしまうという理由から起こるものである。

そこで、次に、RGB 差分値を用いたブレンド方法を説明する。複数の基本顔のブレンドを行った場合でも皺などの特徴が薄くならないようなブレンド手法である。下式を用いて画素値を算出する。

$$r = r_n + \sum (r_i - r_n) R_i \quad (式 3)$$

各基本顔画像において、無表情画像との差分画像を作成しておき、その差分にブレンド率を掛け合わせたものを無表情画像に加算する。こうすることで、複数の基本顔のブレンドを行った場合でも、それぞれの特徴を失うことなく顔のテクスチャ画像を作成することができる。

さらに、3 つ目の手法として、皺を強調したブレンドについて説明する。表情を変化させ

たときに特に変化するのは、皺のある部分のみであろうと考え、皺のある部分のみを強調してブレンドを行う。各基本顔の画像において、無表情画像との輝度値の差分が、ある一定値より大きくなっている画素があれば、その画素は皺の表出する画素であると考え。そして、その皺の表出する画素群のみに、前述の RGB 値を用いたブレンドを行う。こうすることにより、皺の表出を強調させることができる。しかし、皺の表出する画素と、そうでない画素の境界が目立ち、多少不自然な印象を与えてしまっている。

ここで説明した、3 つのテクスチャブレンド手法による結果の例を図 4 に示した。RGB 値のブレンド(A)では、複数の基本顔をブレンドした際に皺が薄くなっている。差分値によるブレンド(B)では、皺がうすくなる問題を解決している。皺を強調したブレンド(C)では、皺が強調されたが、皺以外の場所で色が不自然に見えるようになってしまった。

どの手法においても、現在の整合の精度では画素の対応を正確にとることができないため、皺や眉などがずれて二重になってしまうという問題が起きている(図 5)。

5. 皺に特化したテクスチャブレンド

ただ単にテクスチャ画像全体をブレンドするだけでは、皺や眉が二重になってしまうという問題が起きる。これを解決するために、テクスチャブレンドのもうひとつのアプローチとして、皺の位置を整合して皺のみをブレンド



図 4 テクスチャブレンド手法の比較



図5 二重になった眉毛

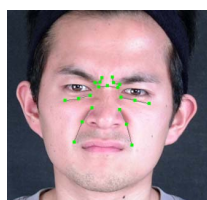


図6 皺の整合

ィングするという手法を提案する。ここでは皺ひとつひとつの位置を整合することによって、皺がずれる問題を回避する。また、テクスチャの皺の部分以外は、表情変形時にも大きく変化しないと考え、皺の周辺のみをブレンディングさせることによって計算量を減少させる。

まず、個人個人の皺の出る数に応じて 10 ~ 20 程度の皺領域を定義する。全ての基本顔のテクスチャにおいて、定義した皺領域上に 3 点程度の頂点からなる折れ線を整合する(図 6)。この折れ線によって、複数の基本顔上にある皺の位置を対応づける。

テクスチャブレンディングは、これらの皺の周辺のみで行う。各皺上に整合した折れ線を軸として幅 10pixel 程度の領域にある画素を、皺テンプレート画像として保存する(図 7)。次に、各基本顔から作られたこれらの皺テンプレート画像を、前述の RGB 差分値を用いた手法でブレンディングする。これにより、全ての皺において、任意のブレンド率での皺テンプレート画像が作成される。最後に、作成された皺テンプレート画像を、無表情顔のテクスチャに貼り付けることで、顔のテクスチャを完成させる。この際に、皺テンプレート画像は中心軸からの距離によって、アルファ値を変化させ、皺テンプレートの境界をぼかしている。

6. 口形発話アニメーション

これで、基本顔のブレンド率を変化させることで、皺などの詳細な情報を持った任意の顔表情を作ることが可能となったが、ここでは、さらなる

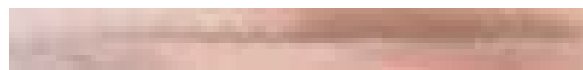


図7 口の右上から作成した皺テンプレート画像

応用例として、口形発話アニメーションについて簡単に説明する。

表情合成時には 11 種類の基本顔を用意したが、口形発話アニメーションを作成する場合は、さらに VISEME の口形モデル群を用意する。VISEME を作った顔を、基本顔と同様にデジタルカメラ、レンジファインダを用いて作成する。作成した VISEME の口形モデル群を、任意のブレンド率でブレンディングすることで、任意の口形を再現することができる。ある口形モデルから別の口形モデルへ、次々と変化するようにブレンド率を変化させながら、随時ブレンディングして顔モデルを作成することによって、口形発話アニメーションを作成することができる。

7. 結論

顔表情の構築にあたって、AU による顔モデル変形ルールでは表現しきれない個人に依存するような特徴の表現を、ブレンディングという方法で可能にした。また、ワイヤフレームモデルだけでなく、3 種類のテクスチャブレンディング手法や、皺のブレンディングを試み、より詳細な特徴の表現を可能にした。さらに応用例として、口形発話アニメーションの作成が可能であることを示した。

参考文献

- [1] P.Ekman and W.Friesen : " Facial Action Coding System " Consulting Psychologists Press、1977
- [2] 柳澤, 高橋, 森島:“テクスチャブレンディングによる皺の表現と基本顔モデルによる感情空間の構築” 電気情報通信学会総合大会(2002-03)