

キャラクタ特有の特徴再現を考慮した リアルな表情リターゲティング手法の提案

加藤卓哉^{†1} 川井正英^{†1} 斉藤隼介^{†1} 岩尾知頼^{†1} 前島謙宣^{†1} 森島繁生^{†2}

CGキャラクタの表情作成に用いられるブレンドシェイプでは、それぞれのキャラクタ特有の表情変化を再現したキーシェイプ作成に多大な労力を要していた。そこで本研究では、人間の表情を転写することによって生成されたキーシェイプからアーティストが作成したキーシェイプへの写像を、少数の表情を用いて学習し、人間の表情が転写された他の表情のキーシェイプに適用する手法を提案する。本手法で生成したキーシェイプを用いることで、アーティストが少数の表情で定義した、キャラクタ特有の表情特徴を考慮したリターゲティングを表現した。

Realistic Facial Retargeting with Individual Characteristics

TAKUYA KATO^{†1} MASAHIDE KAWAI^{†1} SHUNSUKE SAITO^{†1}
TOMOYORI IWAO^{†1} AKINOBU MAEJIMA^{†1} SHIGEO MORISHIMA^{†2}

Although facial retargeting using *Blendshape Animation* has been a major method to create CG facial animations, scalping its *Blendshapes* has been a drawback since it causes enormous labor. In this paper, we propose a method to create mapping between Blendshapes without individual characteristics and the training examples, and apply them onto the input facial model created by transferring geometry of other characters. Our system transfers the individual characteristics defined on couple of training examples to efficiently achieve facial retargeting for Blendshape Animation with individual characteristics.

1. あらまし

映画やゲームなどのCG映像コンテンツにおいて、キャラクタの表情は、シーンのクオリティを高める上で、最も重要な要素の一つである。近年では、作品中の多くのキャラクタがCGで製作されたコンテンツも多く、またその一つ一つのキャラクタの表情を細部まで精細に作成しているものも増加している。

このようなキャラクタの表情を精細に作成する手法として、リターゲティングと呼ばれる手法が用いられている。リターゲティングとは、モーションキャプチャなどの機器を用いて得られた役者の表情データから、高精細なアニメーションを作成する手法である。その中でも、ブレンドシェイプと呼ばれる手法を用いたリターゲティングは多くの作品に用いられている。ブレンドシェイプとは、キャラクタの基本表情であるキーシェイプの線形結合によって任意の表情を表現する手法である。リターゲティング対象のキャラクタのキーシェイプと計測された表情データとを対応付けることによって、役者の表情に適したフェイシャルアニメーションを作成することができる。一方で、ブレンドシェイプには、手動によるブレンド率設定とキーシェイプ作成という2つの要点が存在する。ブレンド率設定では、表情データを高精度に計測する手法や、ブレンド率を自動かつ正確に推定する研究など様々なアプローチ

がとられてきた[1, 2, 3, 4, 5]。この点に関しては、様々な研究がなされており、その精度は大きく向上した。一方、キーシェイプ生成に関しては、あるキャラクタの表情を他のキャラクタに転写することでキーシェイプを自動生成する手法が提案されている[6, 7, 8]。しかしながら、別のキャラクタの表情を転写する手法では、それぞれのキャラクタで少しずつ異なるキャラクタ特有の表情変化の特徴が再現できないという問題があった。例えば、あるキャラクタの笑いと、別のキャラクタの笑いでは、笑いという意味合いが同じでも、キーシェイプとしての形状としては異なり、この差分を再現することは困難であった。

そこで本研究では、表情特徴が再現されていないリターゲティング対象のキャラクタのキーシェイプに、キャラクタ特有の表情変化の特徴を再現することで、リアルなアニメーションを作成する手法を提案する。本手法では、まずアーティストによって作成されたリターゲティング対象のキャラクタの表情モデルをトレーニングサンプルとして作成しておく。次に、リターゲティング元のキャラクタの表情を転写して作成された表情モデルと、意味合いが同様の表情のトレーニングサンプルとの形状の差分を、写像として抽出する。そして、顔をいくつかの領域に分割し、抽出された表情変化の特徴を、キーシェイプの各領域に適切に選択し、再現させる。この際、表情変化の特徴を適切にブレンドする方法を用いることによって、少数のトレーニングサンプルからでも、効率的にキャラクタ特有の特徴を再現したフェイシャルアニメーションの生成を実現した。また、従来のキーシェイプ自動生成手法では困難であった、

^{†1} 早稲田大学

^{†1} Waseda University

^{†2} 早稲田大学理工学術院理工学総合研究所

^{†2} Waseda Research Institute for Science and Engineering

メッシュ構造の異なるキャラクタに対するリターゲッティングにも、対応することが可能となった。

2. 関連研究

2.1 ブレンドシェイプによるリターゲッティングに関する研究

リターゲッティングによるフェイシャルアニメーションの生成手法には、様々な方法がある。その中でも、ブレンドシェイプを用いたリターゲッティングは、ある2つの表情の中間の表情を、キーシェイプの線形和で作成できることから、アニメータにとって扱いやすいという利点がある。

ブレンドシェイプによるリターゲッティングに関する研究には主に2つのアプローチがある。まず、顔に貼ったマーカーの座標値などの表情データを取得することで、ブレンド率推定の精度向上を図るアプローチがある。このアプローチをとっている研究には、デプスデータと2次元画像の両方を用いて役者の表情を高精度にトラッキングする手法[1]、2次元画像のみから表情データを高精度に取得する手法[2]などがある。また、表情を合成する際のブレンド率に着目して、高精度かつ効率的にフェイシャルアニメーションを作成するアプローチも提案されている。このアプローチをとっている研究としては、フレーム間の各キーシェイプのブレンド率の連続性を考慮する手法[3]や、同時に使用するキーシェイプの数を減らして効率的に表情を作成する方法[4]、顔領域を分割して高精度に非対称な表情を作成する方法[5]などが提案されている。しかしこれらの手法は、予めキーシェイプを作成する必要があるため、合成された表情が、キーシェイプをどの程度精細に作成したかとその個数に大きく依存する。よって、精細に作成されたキーシェイプを用いていない場合、合成された表情には表現力が欠けてしまう問題があった。

2.2 キーシェイプ生成に関する研究

本手法のようにキーシェイプの生成や形状の最適化を目的とした手法は、いくつか提案されている。

Sumnerらは、キーシェイプを自動生成する手法として、あるキャラクタの変形を別のキャラクタに転写する手法を提案している[6]。しかし、この手法は、部位同士の接触を考慮しない手法であるため、部位によって不自然な変形が転写される問題があった。Saitoらは、顔に特化した表情の転写手法として、目と口に現れる不自然な変形を、より自然に見えるように変形させる手法を提案した[7]。この手法では、目や口に対して疑似的なポリゴンを作成することで、目や口が閉じきらないなどの不自然さを取り除くことができた。しかし、彼らの手法は、人同士などの、顔の形状が似たキャラクタ同士の変形転写を想定していた。そのため、転写元のキャラクタの表情変化をそのまま転写先のキャラ

クタに適用しようとするために生じる、転写先のキャラクタ特有な表情変化の特徴が再現できない問題は解決されなかった。また顔の各部位の大きさやメッシュ構造が著しく異なるキャラクタ間では、目や口の中のポリゴンの数や、大きさが異なってしまい、不自然さが残ってしまう問題があった。

また、予め作成されたキーシェイプの形状を最適化する手法も提案されている。Liらの手法では、まずアーティストが作成したトレーニングサンプルに最も近い表情のモデルを、予め用意されたキーシェイプのブレンドシェイプによって作成する。次に、そのときのブレンド率を固定して、トレーニングサンプルを表現するために、予め用意されたキーシェイプの形状を最適化する手法を提案した[8]。この手法は、キャラクタ特有の表情変化の特徴を持ったトレーニングサンプルを表現するために適切な形状にキーシェイプを変形することで、キャラクタ特有の表情変化の特徴をキーシェイプに反映することができた。しかし、彼らの手法は、予め用意されたトレーニングサンプルの表情のみから特有の表情変化の特徴を学習するため、トレーニングサンプルにないような未知の表情に対して適切な形状に最適化されているという保証はされていない。また、同様の理由から、任意の表情に対応するためには、多数のトレーニングサンプルを必要とすることが問題であった。

2.3 本手法の新規性

関連研究(2.1, 2.2 章)の問題点を踏まえて、本手法のゴールについて述べる。まず、Liらの手法は、トレーニングサンプルに存在しない表情に対して適切な形状になっている保証がない上、この手法は多くのトレーニングサンプルが必要であった。そこで、本手法では、トレーニングサンプルに存在しないような表情に対しても、少数のトレーニングサンプルから適切な表情変化の特徴を作成し、再現させることを目標とした。また、Saitoらの手法では為しえなかった、顔の各部位やメッシュ構造が大きく異なるキャラクタ間でも、表情変化の特徴を再現できることを目指した。以上を要約すると、以下の3つの目標にまとめることができる。

1. トレーニングサンプルに存在しない表情に対しても、適切な表情特徴を再現させる
2. 従来手法と比べて、トレーニングサンプルを少なくする
3. メッシュ構造が大きく異なるモデルにも適用できる

この3点を達成することで、効率的かつ汎用的にキーシェイプを生成する手法を提案する。

3. キャラクタ特有の表情変化の特徴を表す写像の作成

本研究では、あるキャラクタにおいて作成された表情のモデルを用いて、その表情をリターゲッティング対象のキャラ



図1. 本手法の流れ

ラクタに転写することでキーシェイプの生成を行う。しかし、リターゲット元キャラクターの表情をそのまま転写すると、リターゲット対象キャラクターに対してキャラクター特有の表情変化の特徴が再現されず、もっともらしい表情にならない。よって、キャラクター特有の表情変化の特徴は、その特徴が欠如した表情と適切に再現された表情との間の差分として表現されると考えられる。例えば、ある人間のキャラクターの表情をサルキャラクターにリターゲットする場合、そのまま表情を転写すると、各部位の大きさや動かし方が異なることで、人間の表情を持ったサルの表情モデルができてしまう。しかし、一般的にこの表情モデルの表情は、このサルに適した表情ではない。つまり、この差分を埋めることが、キャラクター特有の表情変化の特徴を再現するということになる。

そこで、本手法では、まずアーティストが作成したトレーニングサンプルと、その表情と意味合いが同様であるリターゲット元のキャラクターの表情を転写して作成した表情モデルを用意する。そして、そのトレーニングサンプルと表情モデルとの、無表情からの変形の様子を比較し、その間の写像を作成する。この写像を、同様の方法でそれぞれのトレーニングサンプルから抽出する。本手法の流れを図1に示す。

3.1 Deformation Transfer

本手法では、まずリターゲット元の表情を転写して、リターゲット対象キャラクターのキーシェイプを作成するため、Sumnerらの提案した、“Deformation Transfer”を用いる[6]。Deformation Transferは、ソースとなるメッシュモデル(以下、ソースモデル)が変形した時、対応する各ポリゴンがどのように変形したかを転写することで、ターゲットとなる別のメッシュモデル(以下、ターゲットモデル)に同様の変形を加える手法である。

顔の例で考えてみると、無表情のソースモデルから任意の表情のソースモデルにおいて各ポリゴンでの変形の様子を転写することに相当する。ここで、変形の様子を測るために変形勾配テンソルと呼ばれる物理量を導入する。変形勾配テンソルとは、拡大縮小要素と回転要素を統合的に表す $R^{3 \times 3}$ のテンソル量である。

まず、頂点数 n 、ポリゴン数 m の無表情のソースモデルと

任意の表情のソースモデルの間で、各ポリゴンについての変形勾配テンソルを求める。そこで、 j 番目のポリゴンの一つの頂点からポリゴンに向かって垂直な単位ベクトルを作った先の点を新たな頂点 v_{j4} とするため、以下の式(1)から v_{j4} を計算する。

$$v_{j4} = v_{j1} + \frac{(v_{j2} - v_{j1}) \times (v_{j3} - v_{j1})}{\sqrt{(v_{j2} - v_{j1}) \times (v_{j3} - v_{j1})}} \quad (1)$$

ここで $v_{jk} \{k = 1, 2, 3\}$ は無表情のソースモデルにおける j 番目のポリゴンの各頂点である。この v_{j4} を用いて、(2)式より、 i 番目のポリゴンの変形勾配テンソル J_i を求める。

$$J_i = V_i' V_i^{-1} \quad (2)$$

$$V_i = [v_{i2} - v_{i1} \quad v_{i3} - v_{i1} \quad v_{i4} - v_{i1}]$$

$$V_i' = [v'_{i2} - v'_{i1} \quad v'_{i3} - v'_{i1} \quad v'_{i4} - v'_{i1}]$$

ここで $v'_{jk} \{k = 1, 2, 3, 4\}$ は変形後のソースモデルにおける j 番目のポリゴンの各頂点である。この変形勾配テンソルを行列として以下のようにまとめる。

$$F = [J_1, J_2, \dots, J_M]^T$$

またターゲットモデルに関しては、変形後の頂点を格納する行列 $X' = [x'_1, \dots, x'_N]^T$ と積を計算したときに変形勾配テンソルとなるような行列 $A \in R^{3m \times n}$ を作成する。この行列 A の作成方法についての詳細は[Sumner et al. 2005]を参照されたい。

このようにして得られた F, A を用いて、式(3)より、変形後のターゲットモデルの頂点 X' を求める。

$$\min \|F - AX'\| \quad (3)$$

この式を解くことで、ソースモデルの変形勾配テンソルと、ターゲットモデルの変形勾配テンソルを等しくするような頂点の座標が得られる。人間のメッシュモデルをソースモデル、サルのメッシュモデルをターゲットモデルとしてDeformation Transferを行った結果を図2に示す。

3.2 トレーニングサンプルの準備

本手法では、キャラクター特有の表情変化の特徴を抽出するため、すでにキャラクター特有の表情変化の特徴が再現されているような表情モデルを用意して、トレーニングサン

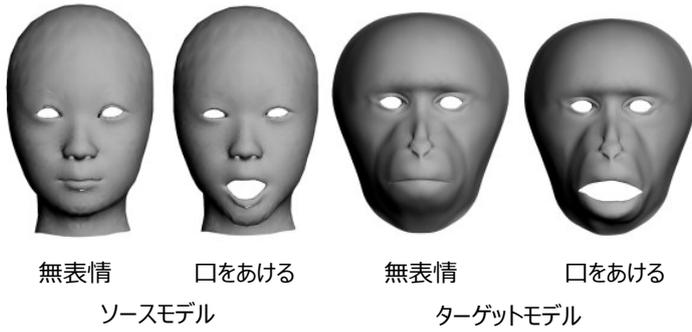


図 2. Deformation Transfer[Sumner et al. 2004]の結果

ルとして用いる．キャラクタ特有の表情変化の特徴は，そのモデルが潜在的に持つものではなく，一般的にアニメータが，そのキャラクタにどのような表情をしてほしいかによって決定されるものである．そこで本稿では，リターゲット対象のキャラクタに対して，いくつかの特徴的な表情をアニメータに作成してもらい，その表情モデルをトレーニングサンプルとした．

3.3 キャラクタ特有の表情変化の特徴を表す写像の学習

転写先のキャラクタ特有の表情特徴を再現する写像を，トレーニングサンプルから学習する．

まず，ある表情 M のトレーニングサンプル t_M と，Deformation Transfer を用いて作成されたリターゲット対象のキャラクタの表情 M のキーシェイプ s_M を用意する．ここで，用意されたそれぞれの表情モデルの一例を図 3 に示す．それぞれの表情モデルを比較してみると，その形状には差があることがわかる．そこで， s_M と t_M の i 番目のポリゴンに対し，無表情時からの変形を表すそれぞれの変形勾配テンソル $s_M^i \in \mathbb{R}^{3 \times 3}$ ， $t_M^i \in \mathbb{R}^{3 \times 3}$ を，式(2)と同様の式である式(4)を用いて計算する．

$$\begin{aligned} s_M^i &= V_{s_M^i} V_{N^i}^{-1} \\ t_M^i &= V_{t_M^i} V_{N^i}^{-1} \end{aligned} \quad (4)$$

$$\begin{aligned} V_{s_M^i} &= [v_{s_M^i 2} - v_{s_M^i 1} \quad v_{s_M^i 3} - v_{s_M^i 1} \quad v_{s_M^i 4} - v_{s_M^i 1}] \\ V_{t_M^i} &= [v_{t_M^i 2} - v_{t_M^i 1} \quad v_{t_M^i 3} - v_{t_M^i 1} \quad v_{t_M^i 4} - v_{t_M^i 1}] \\ V_{N^i} &= [v_{N^i 2} - v_{N^i 1} \quad v_{N^i 3} - v_{N^i 1} \quad v_{N^i 4} - v_{N^i 1}] \end{aligned}$$

ここで $v_{s_M^i k}$ ($k = 1, 2, 3, 4$) は，ソースモデルにおける頂点を表し， $v_{t_M^i k}$ ($k = 1, 2, 3, 4$) は，ターゲットモデルにおける頂点を表す．

そして， i 番目のポリゴンにおいて，生成された表情 M のキーシェイプの変形勾配テンソル $s_M^i \in \mathbb{R}^{3 \times 3}$ と，表情 M のトレーニングサンプルの変形勾配テンソル $t_M^i \in \mathbb{R}^{3 \times 3}$ を等しくするような写像 $f_M^i \in \mathbb{R}^{3 \times 3}$ を式(5)によって求める．

$$f_M^i = t_M^i s_M^i^{-1} \quad (5)$$

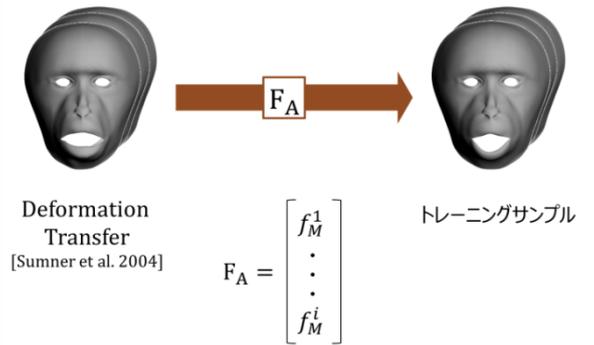


図 3. トレーニングサンプルからの写像の抽出

この写像を，トレーニングサンプルが存在する全て表情の全ポリゴンについて求める．この写像は，無表情時の表情モデルと s_M との変形勾配テンソルに適用し，式(3)の F として解くことで， t_M の頂点を得るような写像である．つまりこの写像は，Deformation Transfer で作成された表情モデルを，キャラクタ特有の表情変化の特徴が再現されるように修正するような写像であるといえる．トレーニングサンプルと似た別の表情を Deformation Transfer で作成した表情モデルにこの写像を適用することで，キャラクタ特有の表情変化の特徴を再現することができる．

4. 領域毎の写像の適用

前章で生成した写像は，無表情の表情モデルから，Deformation Transfer を用いて生成されたトレーニングサンプルと意味的に同様な表情の表情モデルに対しての変形勾配テンソルを，顔全体として修正するようなものであった．しかし，顔の各部位は独立に変形するため，部位によって適用すべき写像が異なる場合がある．そこで，本手法では，顔をいくつかの領域に分割し，その各領域に適した写像を新たに生成し，適用する．本稿では，写像の変形の様子から経験的に 18 個の領域に分割を行った．その結果を図 4 に示す．

4.1 領域毎の写像の生成

生成された各写像は，その写像を作った際に用いたトレーニングサンプルの表情に対してのみ有効な写像であるといえる．しかし，キャラクタ特有の表情変化の特徴を再現させたい入力表情が，トレーニングサンプルの表情とある程度異なることを想定しておく必要がある．そこで本手法では，入力表情モデル，各写像を生成するとき用いた表情の中でどの表情と形状に近いのかを各領域で判定し，その類似性の度合に応じて，新たな写像を領域毎に生成する手法を提案する．具体的には，入力表情がどのトレーニングサンプルの形状に近いのかを，Deformation Transfer を用いて作成されたリターゲット対象のキャラクタの表



図 4. 本稿での分割領域

情 M のキーシェイプ s_M をキーシェイプとして、ブレンドシェイプを行う。そこで得られたブレンド率に応じた変形の写像を新たに生成し、適用する。

4.2 極分解と Exponential Map を用いた変形勾配テンソルのブレンド法

変形勾配テンソルのブレンド方法には様々な方法が考えられる。最もシンプルな方法は、各要素の値を、ブレンド率にしたがって線形補間する方法である。しかし、この方法では、各要素の値によってポリゴンがつぶれる変形や、ポリゴンが反転する変形を表すような変形勾配テンソルが得られてしまう可能性がある。このような変形勾配テンソルをそのままポリゴンに適用すると、不自然な形状の表情モデルが得られてしまう問題が発生する。

そこで本手法では、Kaji らが提案したつぶれる変形や反転が起こらない 2 つの勾配変形テンソル同士の補間方法 [21] を参考として、新しい変形勾配テンソルのブレンド方法を提案する。この手法を用いることで、見た目として自然な変形勾配テンソルのブレンドを実現した。

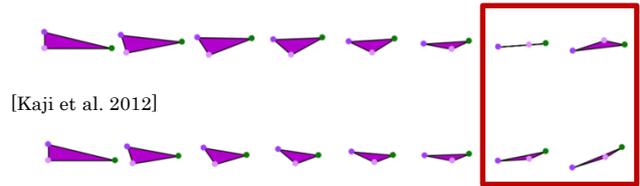
Kaji らは、Alexa らが提案した 2 つの変形勾配テンソルの補間法 [20] を用いると、つぶれる変形や反転といった、人間が見て直感的に不自然と感じるような変形を含んだ補間になってしまうことを指摘した。その様子を図 5 に示す。そこで、まず変形勾配テンソルを極分解し、回転行列とそれ以外の変形を表す行列に分解して、それぞれの要素で別々の補間法を用いる手法を提案した。具体的には、回転行列は角度に関して線形な補間方法、それ以外の変形成分は、行列対数と行列の指数関数を用いた補間方法の積として計算する。単位行列 $I \in R^{3 \times 3}$ と任意の変形勾配テンソル $d \in R^{3 \times 3}$ との間で、重み係数 $w \in [0, 1]$ で補間された変形勾配テンソル $interp_d(w) \in R^{3 \times 3}$ は、以下の式によって得ることができる。

$$interp_d(w) = R_\theta^w \exp(w \log S)$$

ここで、 d の回転行列を R_θ^w とし、それ以外の変形を表す行列を S とする。また、重み係数 w は、 $w = 0$ のとき $interp_d = I$ 、 $w = 1$ のとき $interp_d = d$ となるような値である。このように補間を行うことで、つぶれる変形や反転を含まないような補間を実現するような条件、 $\det(interp_d) > 0$ を常に成り

[Alexa et al. 2000]

[Kaji et al. 2012]



Mathematical analysis on affine maps for 2D shape interpolation [Kaji et al.] 2012

図 5. 変形勾配テンソルの補間方法の違いによる差

立たせることができる。

以上の Kaji らの補間法を参考として、以下の方法で変形勾配テンソルをブレンドする。まず、 M 個の写像 f_M^i に対して極分解を行い、回転行列 $Rot f_M^i \in R^{3 \times 3}$ と回転以外の変形を表す行列 $Sym f_M^i \in R^{3 \times 3}$ に分解する。次に、 $Rot f_M^i$ と $Sym f_M^i$ の、ブレンド率に応じた変形量を表す行列を、それぞれ別の方法で計算する。 i 番目のポリゴンにおける回転行列 Rot_i については、クォータニオンを用いた球面線形補間を適用する。クォータニオンを用いた球面線形補間を行うことで、回転行列を、重み係数に応じた角度の回転量で表す回転行列にすることができる。具体的には、 f_M^i のブレンド率に応じた回転量の回転行列を式(6)によって生成する。

$$Rot_i = \prod_{m=0}^M slerp(q_I, q_{Rot f_M^i}, w_{jm}) \quad (6)$$

ここで $slerp$ は球面線形補間を表し、 q_I 、 $q_{Rot f_M^i}$ はそれぞれ単位行列と $Rot f_M^i$ のクォータニオン、 $w_{jm} \in [0, 1]$ は領域 j における m 個目の写像のブレンド率である。また、 i 番目のポリゴンにおける回転以外の変形を表す行列 Sym_i については、行列対数と行列の指数関数を用いた方法を適用する。具体的には、 M 個の写像 f_M^i のブレンド率に応じた変形量の和を表す回転以外の変形を表す行列を式(7)によって生成する。

$$Sym_i = \exp(\sum_{m=0}^M w_{jm} \log(Sym f_M^i)) \quad (7)$$

こうして得られた回転行列と回転以外の変形を表す行列の積を式(8)によって計算することで、ブレンド率に応じた写像 $Blend f^i$ を生成する。

$$Blend f^i = Rot_i \cdot Sym_i \quad (8)$$

この手法を用いることで、ポリゴンが反転したり、つぶれたりすることなく、写像をより自然に写像をブレンドすることが実現できる。この新しい写像を各領域でのブレンド率に応じてブレンドすることによって、その領域に適した変形を持つ写像を作成することができる。

4.3 領域毎のブレンド率の推定

s_M をキーシェイプとしたブレンドシェイプのブレンド率を求める。この際、領域毎にどのキーシェイプの形状が近いかを判定するため、領域毎にブレンド率の推定を行う。

各キーシェイプの r 番目の領域の線形和によって任意の表情モデルの r 番目の領域の形状を表現することを考える. この線形問題は, 式(9)のように表すことができる.

$$\mathbf{v}_r = \mathbf{B}_r \mathbf{w}_r \quad (9)$$

ここで, $\mathbf{v}_r \in \mathbb{R}^{3n \times 1}$ は領域内の n 個の頂点の各座標値を要素に持つベクトル, $\mathbf{B}_r \in \mathbb{R}^{3n \times M}$ は M 個のキーシェイプの領域内の n 個の頂点の各座標値からなる行列, $\mathbf{w}_r \in \mathbb{R}^{M \times 1}$ は推定するブレンド率である. これより, $\mathbf{v}_r, \mathbf{B}_r$ が既知である時, 次のエネルギー関数 E を最小化することで \mathbf{w}_r を求めることができる.

$$E(\mathbf{w}_r) = \|\mathbf{v}_r - \mathbf{B}_r \mathbf{w}_r\| \quad (10)$$

このエネルギー最小化問題は, 次のように解くことができる.

$$\mathbf{w}_r = (\mathbf{B}_r^T \mathbf{B}_r)^{-1} \mathbf{B}_r^T \mathbf{v}_r \quad (11)$$

この式を領域毎に解いていくことで, 領域毎のブレンド率を独立に推定していく.

4.4 写像の合成方法

前章までで得られた写像は, 変形勾配テンソルにかけられることで, その変形勾配テンソルを修正し, キャラクター固有の表情変化の特徴を再現させるようなものであった. そこで, 本手法では, Deformation Transfer において用いた, 変形後のターゲットモデルの頂点を得る式である式(3)を利用して変形後の頂点座標を求める.

まず, 各領域のブレンド率に応じた写像を, 各ポリゴンに対して用意する. 次に, リターゲットング対象のキャラクターのメッシュモデルのポリゴン数を m とすると, ブレンド率に応じた変形の写像を \mathbf{B} として以下のように並べる.

$$\mathbf{B} = [\text{Blend}f^1, \dots, \text{Blend}f^i, \dots, \text{Blend}f^m]^T$$

この \mathbf{B} を用いて, Deformation Transfer における式(6)と同義である式(12)を解くことでリターゲットング対象のキャラクターのメッシュモデルの頂点を得る.

$$\min \|\mathbf{B}\mathbf{F} - \mathbf{A}\mathbf{X}'\| \quad (12)$$

この最小化問題は, 式(13)のように解くことが可能であり, \mathbf{X}' を求めることができる.

$$\mathbf{X}' = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{B}\mathbf{F} \quad (13)$$

このように各頂点の座標を直接計算することで, 領域間の境界等で変形の様子にある程度の差があっても, 与えられた変形勾配テンソルに対するリターゲットング対象のキャラクターの頂点の最小二乗解が得られるため, メッシュのポリゴン間の連続性はおおよそ保たれることになる.

5. 結果

本手法の有効性を確かめる為に, 以下のような実験を行った. まず, トレーニングサンプルと意味合いが同様のリターゲットング元のキャラクターの表情モデルをキーシェイ

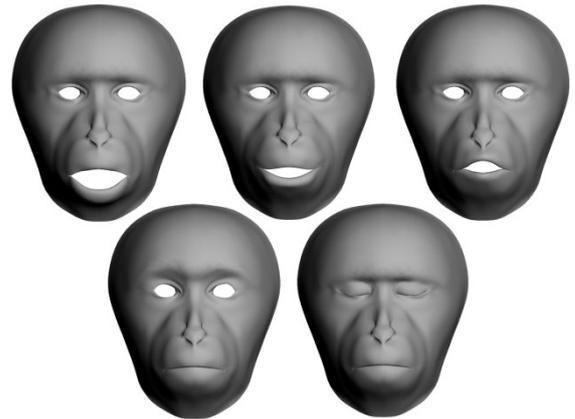
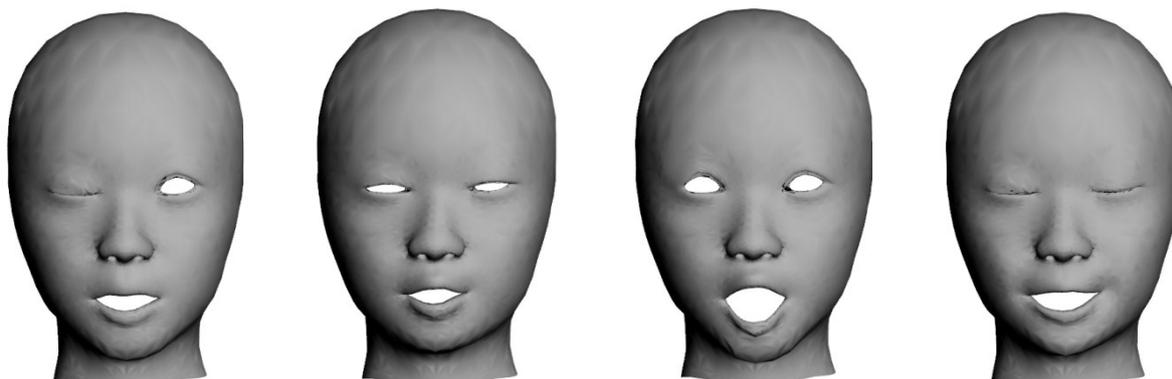


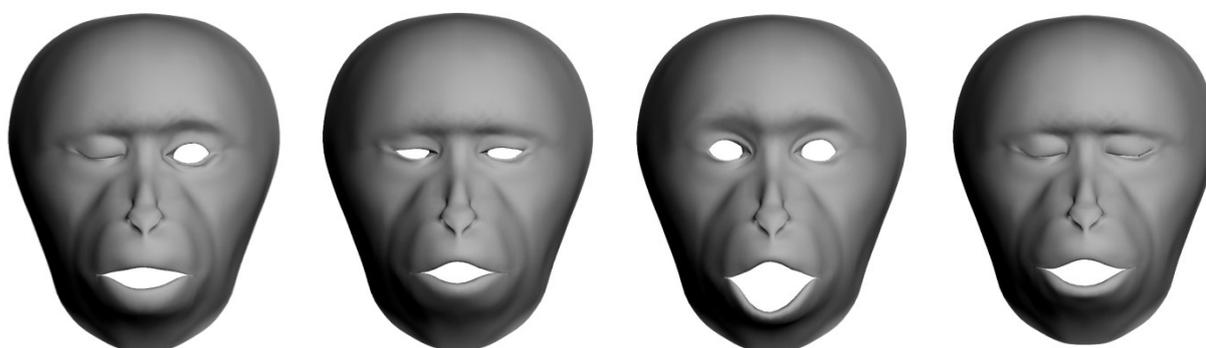
図 6. トレーニングサンプル

プとしてブレンドシェイプを行って任意の表情モデルを作成する. この時のブレンド率は経験的に決定した. このリターゲットング元のキャラクターの表情モデルを入力表情として, 図6のようなトレーニングサンプルを用いて本手法を適用した. その結果を図7に示す. 従来法である Sumnerらの手法と比較してみると, 口の動かし方がトレーニングサンプルに近づき, リターゲットング対象のキャラクターに特有の表情変化の特徴が再現できていることがわかる. また, 目の開き方が, 従来法では完全に閉じきらなかったり, 不自然さが残ったりしているが, 提案手法で作成された表情モデルでは自然に表情がリターゲットングされていることがわかる. 本稿では, 人の表情モデルをリターゲットング元のモデル, サルの表情モデルをリターゲットング対象のモデルとして用いた. これらのモデルは顔の各部位の形状や, ポリゴン数, 頂点数がそれぞれ異なるモデルであるが, 違和感なく再現できていることがわかる. よって, 本手法は, Saito らの手法が有効でないようなメッシュ構造や各部位の大きさが異なる表情モデルに対しても, 有効であることがわかる. また本稿では, トレーニングサンプルの個数を5つという限られた個数に抑えることを達成した.

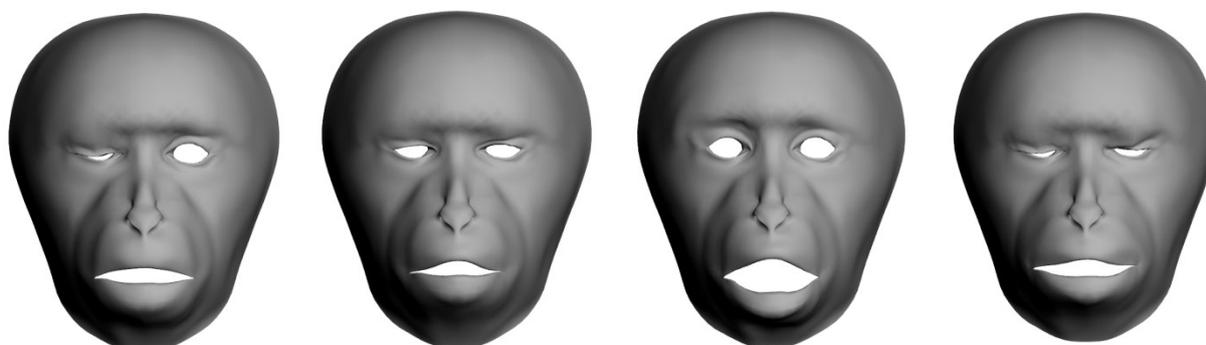
トレーニングサンプルの表情と著しく異なる表情を持つ表情モデルを入力表情として本手法を適用した結果を, 図8に示す. また, アーティストが作成した意味合いが同様であるサルの表情モデルも同様に図8に示す. それぞれを比較してみると, 従来法と比較して, 下唇の動かし方や口の全体的な形状が自然に変形していることがわかる. 一方で, アーティストが作成したキーシェイプと比較すると, 口の両端の閉じ方などが大きく異なっているといえる. このように本手法は, トレーニングサンプルから著しく異なる表情モデルに対しては, 有効でないといえる. ただ, このような特異な表情モデルがトレーニングサンプルとして用意されていれば, その中間的な表情には対応できることになる. よってトレーニングサンプルに, 表情変化を持った表



リターゲット元表情モデル



提案手法



Deformation Transfer [Sumner et al. 2004]

図7 提案手法とDeformation Transferの比較

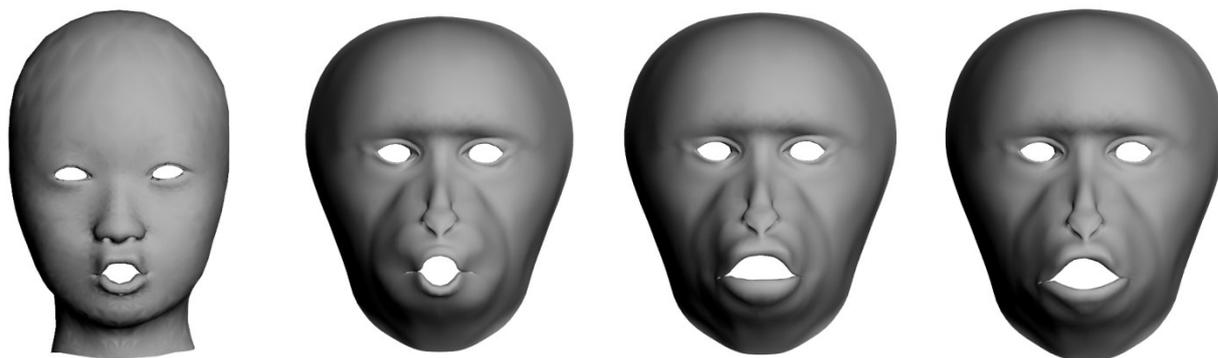
情モデルを用意することで、より本手法が有効になると考えられる。

6. まとめと今後の課題

本研究では、表情特徴が再現されていないリターゲット対象のキャラクターのキーシェイプに、キャラクター特有の表情変化の特徴を再現することで、リアルなアニメーションを作成する手法を提案した。この手法を用いることによって、従来法の抱えていた多くのトレーニングサン

ルを必要とする問題や、メッシュ構造が大きく異なるキャラクターに適用できなかった問題が解消された。

この手法は、精細なキーシェイプを生成する研究であるため、表情データの取得や、表情の合成に着目した他のブレンドシェイプによるリターゲット手法と併用することによって、より高精度なリターゲットが可能となると考えられる。また、他の手法におけるキーシェイプ生成の労力を減らすことができ、より簡便にフェイシャルアニメーションを作成することが可能になる。また本手法は、従来のキーシェイプ生成手法とも併用が可能であるた



リターゲティング元の
 表情モデル

アーティストが作成した
 表情モデル

Deformation Transfer
 [Sumner et al. 2004]

提案手法

図 8.特異な表情変化を持つ表情モデルへの適用例

め、本手法によりキーシェイプを生成した後、さらに形状を最適化する手法を適用することでキーシェイプの精度を改善することができる。

今後の課題としては、まず本稿で最も一般的な方法を用いたブレンド率推定方法の改善が挙げられる。本手法は、結果が推定されたブレンド率に大きく依存する手法である。そこで、適切なブレンド率を推定するために、ブレンド率の上下限に一定の幅を持たせて解く手法の考案を行っていく必要があると考えられる。また、本稿では領域分割を経験的に行ったが、自動で分割領域決定する手法を提案することで、よりユーザーの負担を減らすことができる。そのため、初期状態の分割領域から、よりそのモデルに適した分割領域へと変化させて分割領域を決定する手法の提案を行っていく必要があると考えられる。また、トレーニングサンプルとして推奨される表情の選び方に関しては言及しなかった。そこで今後は、あらかじめ役者の表情データから、高頻度で現れる表情を分析し、トレーニングサンプルとして採用することを検討する。

謝辞 本研究の一部は JST 「OngaCREST プロジェクト」の支援を受けた。

参考文献

- 1) Hao Li, Jihun Yu, Yuting Ye, Chris Bregler: Realtime facial animation with On-the-fly correctives, ACM Transactions on Graphics (TOG) SIGGRAPH 2013 Conference Proceedings, Volume 32, Issue 4, July, Article No.42(2013)
- 2) Chien Cao, Yanlin Weng, Stephen Lin, Kun Zhou: 3D shape regression for real-time facial animation, ACM Transactions on Graphics (TOG) SIGGRAPH 2013 Conference Proceedings, Volume 32, Issue 4, July, Article No. 41(2013)

- 3)Yeongho Seol, J. P. Lewis, Jaewoo Seo, Byungkuk Choi, Ken Anjyo, Junyong Noh, Spacetime expression cloning for blendshapes. ACM Transactions on Graphics . Volume 31 Issue 2, April Article No. 14(2012)
- 4) Sofien Bouaziz, Yangang Wang, Mark Pauly, Online modeling for realtime facial animation. ACM Transactions on Graphics (TOG) SIGGRAPH 2013 Conference Proceedings, Volume 32 Issue 4, July, Article No.40(2013)
- 5) J. Rafael Tena, Fernando De la Torre, Iain Mathews: Interactive Region-Based Linear 3D Face Models. ACM SIGGRAPH2011, Vol.30, Article 76 (2011)
- 6) Robert Sumner, Jovan Popovic: Deformation transfer for triangle meshes. ACM SIGGRAPH 2004 Papers. ACM, Los Angeles, California. 1015736 399-405 (2004)
- 7) Jun Saito: Smooth contact-aware facial Blendshapes transfer. Digipro '13 Proceedings of the Symposium on Digital Production Pages 7-12 (2013)
- 8) Hao Li, Thibaut Weise, Mark Pauly: Example-based facial rigging. ACM Trans. Graph. 29, 4, 1-6(2010)
- 9) Robert Sumner: Mesh modification using deformation transfer. PhD thesis, Massachusetts Institute of Technology(2005)
- 10) Marc Alexa, Daniel Cohen-Or, David Levin: As-Rigid-As-Possible shape interpolation, SIGGRAPH, 157-164(2000)
- 11) Shizuo Kaji, Sampei Hirose, Shigehiro Sakata, Yoshihiro Mizoguchi, Ken Anjyo: Mathematical analysis on affine maps for 2D shape interpolation, Proceeding SCA '12, Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation Pages 71-76(2012)