

疎な特徴点と顔変形モデルに基づく 動画像からの3次元顔モデル自動生成手法

原朋也[†] 前島 謙宣[†] 森島 繁生[†]

本稿では、1台のビデオカメラで撮影された顔を自由に振る動作の動画像から、対象人物の3次元顔モデルを高速自動生成する手法を提案する。著者らは先行研究において、顔変形モデルと顔形状の尤度制約を用いた顔の奥行推定に基づき、正面顔画像から高速に3次元顔モデルを生成する手法を提案している。しかしながら、従来手法は奥行推定に入力顔画像から得られる特徴点の2次元位置座標を用いており、3次元的な制約がないため、鼻の高さや頬部分の凹凸などの個人特徴を忠実に再現できないという問題点があった。そこで、提案手法では、一般的な3次元復元手法として知られる Structure-from-Motion のフレームワークと顔変形モデルに基づく手法を統合することで、より対象人物に近い3次元顔モデルの生成を実現した。

Automatic 3D Face Generation from Video based on Sparse Feature Points and Deformable Face Model

Tomoya Hara[†] Akinobu Maejima[†] Shigeo Morishima[†]

In this paper, we propose a 3D face reconstruction method with hybrid approach that combines Structure-from-Motion(SfM) approach based on "Factorization Method" to estimate an accurate 3D point depth information and generic-model approach based on "Deformable Face Model" to keep an appropriate local face shape. Unlike other methods, our method requires no manual operations from image capturing to 3D face model output and is executed quickly. In consequence, our method succeed to express more plausible facial geometry compared with previous method.

1. 序論

「顔」は人間の部位の中でも多くの情報を発信しており、この顔を対象とした CV・CG 技術に関する研究が盛んに行われている。中でも、個人の「3次元顔モデル」を構築する技術は、ゲームや映像制作をはじめとするエンタテインメントや医療、バイオメトリクス認証など様々な分野での応用がなされている[1,2]。

しかしながら、既存の3次元顔モデル構築手法の多くは、3次元形状情報を取得するためにレンジスキャナ[3,4]の様な大型の計測装置を使用している。これらは、非常に高価であり、撮影に手間がかかるという問題点がある。このような背景から、近年 CV・CG の研究分野において、一般的なデジタルカメラで撮影された2次元顔画像から対象人物の顔の3次元形状を推定して、3次元顔モデルを構築する手法が提案されている[5]-[7]。これらの手法により、3次元顔モデルは、大型な計測装置を必要としない3次元顔認証への応用や、ビデオチャット等のアバターへの利用など、更なる需要の増加と応用範囲の拡大が期待される。

大型の計測装置を使用せずに3次元顔モデルを構築するアプローチとして、ビデオカメラなどによって撮影された複数の顔画像を用いる方法[5]や、事前に保持している汎用の3次元顔モデルを利用する方法[6,7]などが挙げられる。

前者のアプローチは、一般に、ビデオカメラで撮影された画像列から3次元形状およびカメラ運動を推定する Structure-from-Motion(SfM)問題として知られている。Sugayaらは、単眼カメラで対象人物が顔を水平方向に振る動作を撮影し、得られた画像列から Active Appearance Model による特徴点追跡と因子分解法に基づき、3次元顔モデルを自動で構築する手法を提案している[5]。しかしながら、復元が可能な頂点数が少数であるため、鼻の先端や頬の丸みの表現が困難であるという問題がある（これらの部位は、復元された疎なモデルのメッシュを分割し、細分割曲面を作成することで表現されている）。また、処理時間に関する議論もなされていない。

一方、後者のアプローチは、事前に3次元顔形状計測装置によって獲得した様々な3次元顔モデルからデータベースを構築し、これらに含まれる3次元顔モデルの組み合わせにより汎用の顔モデルを表現し、入力された2次元顔画像にフィッティングさせ、汎用モデルを変形させるといった方法である。

[†] 早稲田大学先進理工学研究所
Department of Advanced Science and Engineering

Blanz らは、Morphable Model と呼ばれる形状・テクスチャ情報を有する汎用の顔モデルに基づき、正面および側面から撮影された顔画像から対象人物の3次元顔モデルを構築する手法を提案している[6]。この手法では、入力画像の特徴点と、その点に対応する Morphable Model の特徴点、および各々のテクスチャの残差が最小となる様な最適化を行うことで、精度の高い3次元顔モデルの構築を可能としているが、処理に必要な手間や計算時間に課題があると言える。

また、著者らは先行研究において、顔変形モデルと顔形状に対する尤度制約に基づき、正面顔画像から3次元顔モデルを構築する手法を提案している[7]。この手法は、Blanz らの手法に比べて、高速なモデリングが可能であるが、3次元形状の推定には、2次元的な特徴点の位置関係のみを用いているため、奥行き方向に対する制約がない。そのため、鼻の高さや頬部分の凹凸などの3次元的な個人特徴を忠実に再現できないという問題点があった。

そこで、本研究では、SfM のフレームワークと顔変形モデルに基づく手法を統合したアプローチを取ることにより、より個人に近い3次元顔モデルを高速自動生成する手法を提案する。具体的には、まず固定されたビデオカメラ1台で顔を自由に振る動作を撮影し、その動画画像から特徴点追跡と因子分解法に基づき疎な3次元点群を復元する。次にそれらの点群をターゲットとして密な顔変形モデルを顔形状尤度制約の下でフィッティングすることにより3次元顔モデルを生成する。提案手法の特徴は、(1)民生用ビデオカメラ1台以外に特別な計測装置を必要とせず、また、カメラキャリブレーションや手動操作が不要である点、(2)原理的に顔の動かし方が限定されないため、正面顔画像1枚のみを用いる手法や、水平方向のみの複数視点画像列を用いる手法に比べて、入力として様々な顔向きを取得でき、より3次元的な個人特徴を反映した3次元顔モデルを生成することができる点である。

2. 提案手法の概要

提案手法のワークフローを図1に示す。提案手法は、「SfM アルゴリズムに基づく3次元復元」と「顔変形モデルフィッティングに基づく3次元顔モデル生成」の2ステップから構成される。まず入力動画画像として、ビデオカメラに向かって顔を自由に振る動作を撮影し、得られた動画画像から入江らの手法[8]を用いて特徴点追跡を行う。次に、取得した特徴点列データから因子分解法[9]に基づき、疎な3次元点群を復元する。続いて、復元された疎な点群に対して、任意の変形が可能な顔変形モデルを顔形状尤

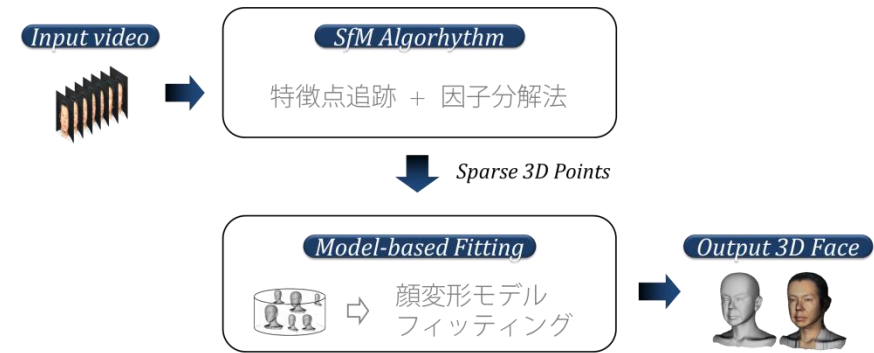


図1 提案手法のワークフロー

度制約の下でフィッティングすることにより、密な3次元顔モデルを生成する。以下では、各ステップの詳細について述べる。

3. SfM アルゴリズムに基づく3次元復元

本章では、提案手法の入力となる動画画像の取得方法と特徴点追跡手法、および3次元復元手法である因子分解法について述べる。

3.1 入力動画画像の取得

まず、動画画像の撮影を行う。以下、撮影環境について述べる。被験者の顔を正面から撮影できる位置にカメラを設置し、被験者が、固定されたビデオカメラに向かって、顔を無表情に保ちながら自由に振る動作を撮影し、動画画像を取得する。撮影に際して、カメラの透視投影効果をできる限り軽減させるため、被験者から4.0m離れた位置から撮影を行う。また、撮影に用いるカメラは、SONY製HDR-SR12であり、画像サイズは600*600 pixel、フレームレートは29.97fpsである。本手法は、Sugayaらの手法[5]とは異なり、任意の顔向きに対しても安定的に特徴点を追跡できるので、被験者に対して顔の動かし方を限定しない。したがって、水平方向だけでなく、垂直方向および斜め方向の動作も考慮することによって、より多くの顔情報を取得することが可能になると考えられる。図2に取得した動画画像の一部を示す。



図2 入力動画像列



図3 特徴点追跡結果

3.2 特徴点追跡

続いて、3次元復元に用いる特徴点の追跡を行う。特徴点は、入江らの手法[8]に基づき、動画中の各フレームに対して、顔の各部位(目・鼻・口・眉・顔輪郭)を含む計30点を自動で検出する。一般に SfM 問題では、時系列画像中の特徴点が全て正しく追跡できているという仮定に基づき3次元形状を推定する。しかしながら、実画像においてこの仮定が成り立つことは稀であり、パターン光やキャプチャマーカなどを使用しない限り、多少の追跡誤差が生じる。本研究で用いる手法[8]も、動画像に対してフレーム毎に特徴点検出処理を行うため、各フレーム間での特徴点位置の対応が正確に取れる保証はない。そこで、特徴点を検出するにあたり、現在注目しているフレームの前後フレームで検出される特徴点との平均位置を計算することで、対応誤差をできる限り軽減し、滑らかな特徴点追跡を実現する。図3に特徴点追跡結果の例を示す。

3.3 因子分解法

次に、取得した特徴点列データから3次元復元を行う。復元には、Tomasi と Kanade により提案された因子分解法[9]を使用する。因子分解法は、SfM 問題における代表的な解法として知られており、実際のカメラモデルである透視射影をアフィン射影に近似し、SfM 問題を線形最適化問題に帰着させることで、他の復元手法に比べて、極めて安定的に解を得ることができるという利点がある。本節では、その詳細について述べる。

まず、 F 枚の動画像列から、 P 個の特徴点追跡結果が得られているとする。この時、 f 番目のフレームにおける特徴点 p の画像上での座標を $(x_{f,p}, y_{f,p})$ とし、 $2F \times P$ の行列 \mathbf{W} を以下の式で定義する。

$$\mathbf{W} = \begin{bmatrix} x_{1,1} & \cdots & x_{1,P} \\ \vdots & \ddots & \vdots \\ x_{F,1} & \cdots & x_{F,P} \\ y_{1,1} & \cdots & y_{1,P} \\ \vdots & \ddots & \vdots \\ y_{F,1} & \cdots & y_{F,P} \end{bmatrix} \quad (1)$$

\mathbf{W} の各列はそれぞれの特徴点の追跡結果に対応し、各行は同一フレーム内の全特徴点座標に相当する。次に、 f 番目のフレームにおける特徴点座標の重心 \bar{x}_f と \bar{y}_f を求める。

$$\bar{x}_f = \frac{1}{P} \sum_{p=1}^P x_{f,p}, \quad \bar{y}_f = \frac{1}{P} \sum_{p=1}^P y_{f,p} \quad (2)$$

続いて、 \mathbf{W} の各要素から \bar{x}_f と \bar{y}_f を差し引いた行列 \mathbf{W}' を作る。 \mathbf{W}' は、特徴点の重心に対する相対位置を要素に持つ行列となり、本研究ではこれを「計測行列」と呼ぶ。

$$\mathbf{W}' = \begin{bmatrix} x_{1,1} - \bar{x}_1 & \cdots & x_{1,P} - \bar{x}_1 \\ \vdots & \ddots & \vdots \\ x_{F,1} - \bar{x}_F & \cdots & x_{F,P} - \bar{x}_F \\ y_{1,1} - \bar{y}_1 & \cdots & y_{1,P} - \bar{y}_1 \\ \vdots & \ddots & \vdots \\ y_{F,1} - \bar{y}_F & \cdots & y_{F,P} - \bar{y}_F \end{bmatrix} \quad (3)$$

ここで、Tomasi と Kanade によれば、対象物体が剛体であるという仮定のもとで、この計測行列 \mathbf{W}' は、カメラの運動を表す運動行列 \mathbf{M} と特徴点の 3 次元位置を表す形状行列 \mathbf{S} に分解することができる。

$$\mathbf{W}'_{(2F \times P)} = \begin{bmatrix} m_1^T \\ \vdots \\ m_F^T \\ n_1^T \\ \vdots \\ n_F^T \end{bmatrix} [s_1 \quad \cdots \quad s_P] = \mathbf{M}_{(2F \times 3)} \mathbf{S}_{(3 \times P)} \quad (4)$$

ただし、 m_f と n_f はそれぞれ f 番目のフレームにおけるカメラの姿勢ベクトルを表し、 s_p は特徴点 p の 3 次元座標位置ベクトルを表す。

本研究では \mathbf{W}' から \mathbf{M} と \mathbf{S} を分解するにあたり、特異値分解を用いる。特異値分解は数値計算的に安定な特性を有しており、ここに因子分解法が安定である要因がある。特異値分解の原理から、任意のランク r の $m \times n$ 行列 \mathbf{A} は、 $m \times m$ の列直交行列 \mathbf{U} 、 $m \times n$ の対角行列 Σ 、 $n \times n$ の列直交行列 \mathbf{V} の積に分解することができ、以下の式で表すことができる。

$$\mathbf{A} = \mathbf{U} \Sigma \mathbf{V}^T \quad (5)$$

行列 Σ の対角要素には特異値 $\sigma_i (i=0, \dots, m)$ が降順に並んでいるが、フレーム数 F と特徴点数 P に関係なく行列 \mathbf{W}' のランクは高々 3 であるので、 σ_4 以降は十分に小さな値となる。そこで、本研究ではこれらの値を 0 とみなして計算を行う。 \mathbf{W}' の特異値分解の結果に対して、

$$\hat{\mathbf{M}} = \mathbf{U} \quad , \quad \hat{\mathbf{S}} = \Sigma \mathbf{V}^T \quad (6)$$

と関係付けると、(4)式の形に分解することが可能である。しかし、一般に(4)式の \mathbf{M} 、 \mathbf{S} と、 \mathbf{W}' の特異値分解により得られる(6)式の $\hat{\mathbf{M}}$ 、 $\hat{\mathbf{S}}$ の間には、次の関係を満たす任意の 3×3 行列 \mathbf{Q} が存在し、 \mathbf{M} 、 \mathbf{S} が一意に定まらないことが知られている。

$$\mathbf{M} = \hat{\mathbf{M}} \mathbf{Q} \quad , \quad \mathbf{S} = \mathbf{Q}^{-1} \hat{\mathbf{S}} \quad (7)$$

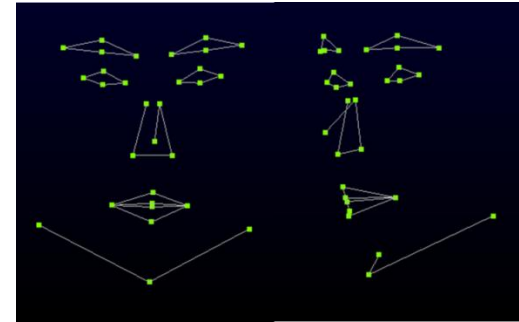


図 4 因子分解法による復元結果

そこで、本研究ではカメラモデルとして正射影モデルを仮定することで、行列 \mathbf{Q} および一意の分解行列 \mathbf{M} 、 \mathbf{S} を求める。具体的には、正射影モデルでは行列 \mathbf{M} の各行が正規直交ベクトルになるという性質を用いて、以下の拘束条件の下で計算を行う。

$$m_f m_f^T = n_f n_f^T = \mathbf{1} \quad , \quad m_f n_f^T = \mathbf{0} \quad (8)$$

4. 顔変形モデルフィッティングに基づく 3 次元顔モデル生成

本章では、復元された疎な 3 次元点群に対して、密な顔変形モデルを顔形状尤度制約の下でフィッティングさせ、最終的な 3 次元顔モデルを生成する手法について述べる。

フィッティングは、著者らが既に提案した手法[7]に基づき行う。従来手法では、データベース内の 3 次元顔モデルの集合から、任意の形状を表現することが可能な顔変形モデルを構築し、入力となる 2 次元の特徴点群とそれに対応する顔変形モデル上の頂点群との関係に基づき、顔変形モデルを変形させることで、3 次元顔形状を推定している。本研究ではこのフィッティングを 3 次元空間上で行えるように従来手法を拡張した。以下では、その詳細について述べる。

4.1 データベース

データベースは、老若男女 1153 名分の 3 次元顔モデルから構成される。3 次元顔モデルは、レンジスキャナにより獲得したデータから、[10]の手法を適用することで生成される。

4.2 顔変形モデル

顔変形モデルは、データベースに含まれる全3次元顔モデルの頂点座標群に対して、主成分分析(Principal Components Analysis : PCA)を行うことにより定式化され、データベース中の顔モデルの線形結合により任意の3次元顔形状を表現可能なモデルである。本節では、この顔変形モデルの定式化について述べる。

まず、データベース内の m 番目の3次元顔モデルの頂点座標を、列ベクトル

$$\mathbf{x}_m = (x_1, y_1, z_1, \dots, x_i, y_i, z_i, \dots, x_N, y_N, z_N)^T \quad (9)$$

として書き並べる。ここで、 N は3次元顔モデルの全頂点数を表し、本研究では $N=6586$ である。この表記を全3次元顔モデルに適用し、データ行列 \mathbf{X} に纏めると、

$$\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M) \quad (10)$$

となる。ここで、 M は3次元顔モデルの総数を表し、本研究では $M=1153$ である。行列 \mathbf{X} の分散・共分散行列を \mathbf{S} とすると、主成分分析により定式化を行うことは、行列 \mathbf{S} についての固有方程式

$$\mathbf{S}\mathbf{b} = \lambda\mathbf{b} \quad (11)$$

を解くことと等価である。ここで、 λ と \mathbf{b} はそれぞれ \mathbf{S} の固有値、および固有ベクトルを表す。(11)式を解くことにより顔変形モデルは、平均顔形状に相当する平均ベクトル $\bar{\mathbf{x}}$ と固有ベクトル行列 $\mathbf{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_N\}$ 、およびモデルパラメータに相当する主成分得点 $\mathbf{p} = \{p_1, \dots, p_N\}$ を用いて以下の式で表すことができる。

$$\tilde{\mathbf{x}}(\mathbf{p}) = \bar{\mathbf{x}} + \sum_{i=1}^{3N} p_i \mathbf{b}_i \quad (12)$$

\mathbf{p} の値を変化させることで、顔変形モデルは任意の顔形状を表現することが可能となる。また、(12)式において、 $H \ll N$ なる定数 H に注目し、上位 H 個までの主成分を利用すると考えれば、 $\tilde{\mathbf{x}}$ は、

$$\tilde{\mathbf{x}}(\mathbf{p}) \approx \bar{\mathbf{x}} + \sum_{h=1}^H p_h \mathbf{b}_h \quad (13)$$

と近似でき、次元圧縮による計算コストの削減が図れる。 H の値によって、次元圧縮率と情報量の間にトレード・オフが生じるが、本研究では寄与率が70%となる45個の主成分を使用した。

4.3 エネルギー関数

本研究では、顔変形モデルのフィッティングを行うにあたり、因子分解法により復元された頂点と対応する顔変形モデルの頂点との間の距離が最小となり、かつ、現在推定されている主成分に対する顔形状尤度が最大となるような、顔変形モデルのモデルパラメータ \mathbf{p} を推定するエネルギー最小化問題として定式化する。本節では、このエネルギー関数について述べる。以下に定式化したエネルギー関数を示す。

$$E(\mathbf{p}) = \alpha E_{fitness}(\mathbf{p}) - \beta E_{likelihood}(\mathbf{p}) \quad (14)$$

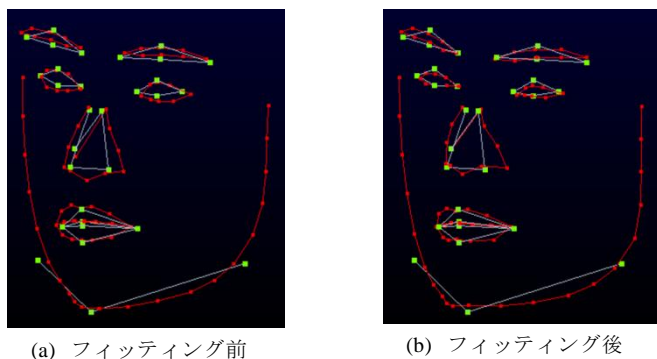
エネルギー関数は頂点制約項 $E_{fitness}(\mathbf{p})$ とモデルパラメータ \mathbf{p} に対する顔形状尤度制約項 $E_{likelihood}(\mathbf{p})$ の重み付き線形和で表され、 α と β は各項に対する優先度を決定する重み係数である。

まず、頂点制約項について述べる。頂点制約項は、因子分解法により復元された3次元頂点と、それに対応する顔変形モデルの頂点との間のユークリッド距離を計量する項であり、具体的には以下の式で表される。

$$E_{fitness}(\mathbf{p}) = \sum_{i=1}^N \|\mathbf{W}_i \{\mathbf{A}\tilde{\mathbf{x}}_i(\mathbf{p}) - \mathbf{z}_i\}\|^2 \quad (15)$$

ここで、 N は顔変形モデルの頂点数、 \mathbf{W} は頂点選択行列を表し、推定に用いる頂点を指定する。 \mathbf{A} は3次元アフィン変換行列を表し、(14)式を実際に解く前に、事前処理としてこの行列を求めることで、顔変形モデルが因子分解法により復元された頂点をターゲットとして粗く位置合わせされる。 \mathbf{p} は解くべきモデルパラメータ、 $\tilde{\mathbf{x}}_i(\mathbf{p})$ は現在推定されているモデルパラメータに基づき表現される顔変形モデルの i 番目の頂点座標ベクトルを表す。また、 \mathbf{z}_i は因子分解法により復元された点群のうち $\tilde{\mathbf{x}}_i(\mathbf{p})$ に対応する頂点座標ベクトルを表す。

次に、顔形状尤度制約項について述べる。顔形状尤度制約項は、データベースに含まれる顔形状の分布をモデル化し、それをモデルパラメータ \mathbf{p} に対する制約として定式化したものであり、具体的には以下の式で表される。



(a) フィッティング前 (b) フィッティング後

図5 顔変形モデルのフィッティング

(緑がターゲット点群, 赤が顔変形モデルの頂点群を表す.)

$$E_{likelihood}(\mathbf{p}) = \sum_{k=1}^K P(k) \Psi(\mathbf{p} | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \quad (16)$$

ここで, 本研究では顔形状分布のモデル化に, 混合ガウス分布モデル(Gaussian Mixture Model:GMM)を用いており, $P(k)$ は各ガウス分布における混合係数, K は GMM の混合数を表し, 本研究では実験的に $K=4$ と設定した. また, $\Psi(\mathbf{p} | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ は確率密度関数であり, 本研究では多変量ガウス分布を仮定している. $\boldsymbol{\mu}_k$, $\boldsymbol{\Sigma}_k$ は多変量ガウス分布のパラメータで, それぞれ各ガウス分布の平均ベクトル, および分散・共分散行列を表す. 尤度制約項は, 現在推定されているモデルパラメータに対する尤度を計算し, 推定結果の顔形状が人物の顔らしいかどうかを判断する. すなわち, 顔形状の破綻を防ぐ効果を持っている.

したがって, 頂点制約項をできる限り小さく, 顔形状尤度制約項をできる限り大きくする様なモデルパラメータが導出できれば, より自然で本人らしい顔形状の推定が可能となる. そこで, エネルギー関数 $E(\mathbf{p})$ を最小にする様なモデルパラメータの導出について考える. (14)式は解析的に解くことが困難であるため, 本研究では準ニュートン法の代表的な手法である BFGS 法により, 最適解 \mathbf{p} を求める. 最終的に得られたモデルパラメータに基づき, 顔変形モデルを変形することで, 目的の3次元顔形状を推定することが可能となる. 図5に因子分解法により復元された疎な3次元点群をターゲットとして, 顔変形モデルが変形する様子を示す.

5. 結果と考察

以上の手順に基づき, 顔変形モデルを変形させ, 入力動画像中の正面顔画像をテクスチャマッピングすることにより, その動画像中の人物らしい3次元形状を持つ顔モデルを生成した. 本章ではそれらの結果を示し, 考察を行う.

5.1 生成結果

図6に, 提案手法により生成された3次元顔モデルの結果を示す. また, 本手法の有効性を検証するため, データベース構築に用いたものと同様の手法[10]を用いて本人のレンジスキャンデータから構築した顔モデル (以後これを正解顔モデルと呼ぶ), および本研究と同様の環境下において撮影された同一人物の正面顔画像1枚から従来手法[7]を用いて生成された3次元顔モデルを併せて示す.

5.2 処理時間

提案手法の処理速度を評価するため, 提案手法を用いて3次元顔モデルを生成するのに要した平均処理時間を計測した. ここで, 特徴点追跡処理に要する時間は, 動画像列の長さに依存するため, 動画像を入力し終えてから各処理に要する平均処理時間に関して計測を行うものとした. また, 同様に従来手法を用いて正面顔画像から3次元顔モデルを生成するのに要した平均処理時間も計測した. 計測実験には, Intel® Core i5-750 (2.66GHz), 8GB メモリの PC を使用した. 表1に処理時間の計測結果を示す.

5.3 考察

図6から, 提案手法は, 入力として様々な顔向きを含んでいるため, 3次元的な特徴を反映した顔モデルを生成できていることが分かる. 特に, 鼻の高さや頬の凹凸感などに関して, 従来手法に比べて正解顔モデルに近い形状を推定できていることが見て取れる. ただし, 現状では従来手法と比べて実際にどれだけ精度が向上したのかが明確でない. そのため, 正解となる形状との誤差を実寸で計量する必要があると考える.

また, 表1から処理時間に関して, 提案手法は動画像を入力し終えてからおよそ7秒程度で3次元顔モデルを自動生成可能であることが分かる. これは従来手法[7]と比較すると多少劣るが, 最適化の際に逐一モデルのレンダリングが必要な手法[6]と比較すると十分高速であると言える.

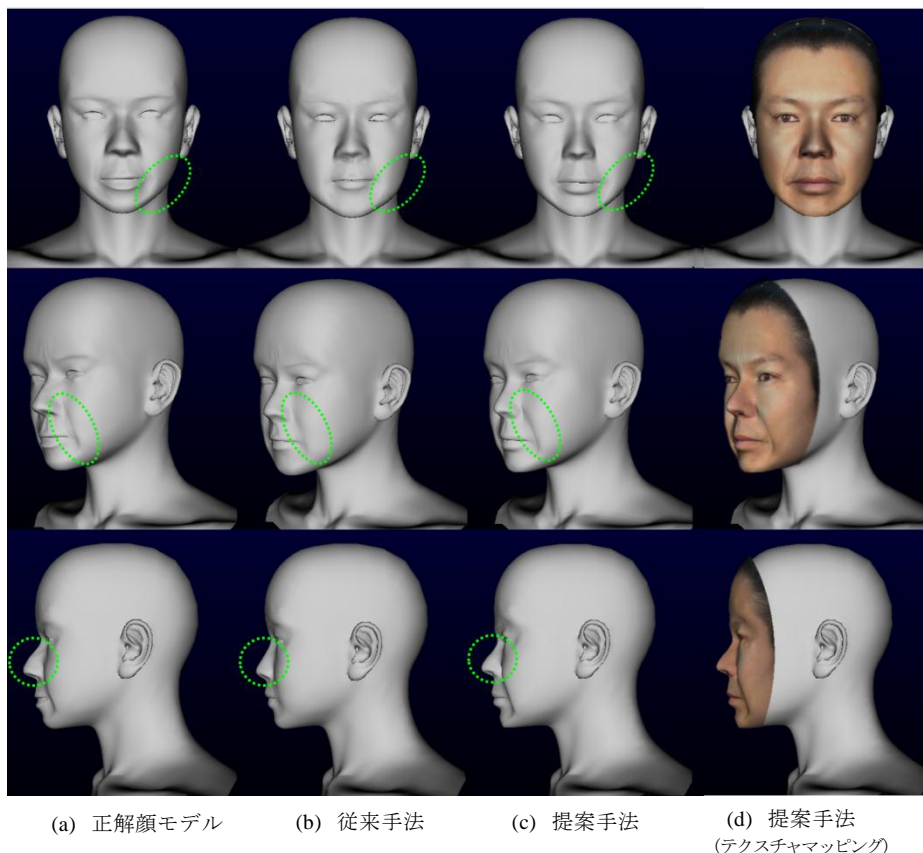


図 6 生成結果

表 1 平均処理時間

	従来手法	提案手法 (内訳)
平均処理時間[秒]	2.7	6.4 (因子分解法: 3.1 モデルフィッティング: 3.3)

6. 結論

本研究では、顔を自由に振る動作を撮影した動画像から 3 次元顔モデルを高速自動生成する手法を提案した。様々な顔向きを含む画像列を入力とすることで、正面顔画像 1 枚のみを入力とする従来手法と比較して、特に鼻や頬領域においてより対象人物に近い顔モデルを生成することが可能となった。提案手法により、ユーザーはカメラに向かって顔を動かすシーンを撮影するだけで自身の 3 次元顔モデルを生成することができ、ゲームやビデオチャットのためのアバター作成などのホームエンタテインメントへの応用が期待できる。

今後の課題として、まず、提案手法の有効性を定量的尺度に基づく評価実験により検証する必要がある。加えて、異なる人物に対する検証や顔変形モデルの表現能力向上のためのデータベースの拡充も重要な課題であると言える。また、今後は、透視投影モデルへの拡張や顔の表情変化などにも対応した復元手法の導入など SfM アルゴリズムの改良についても検討していく方針である。

謝辞 本研究は文部科学省の平成 23 年度科学技術戦略推進費による「安全・安心な社会のための犯罪・テロ対策技術等を実用化するプログラム」の一環として実施されたものである。また、顔特徴点検出に関して技術支援を頂いたオムロン株式会社の川出雅人氏に深謝致します。

参考文献

- 1) S.Morishima : Dive into the Movie -Audience-driven Immersive Experience in the Story-, IEICE TRANSACTIONS on Information and Systems, Special Section on Human Communication III, Vol.E91-D, No.6, pp.1594-1603 (2008)
- 2) 宮澤一之, 野畑幸里, 青木孝文 : 3 次元医用データとカメラ画像の融合手法, 画像の認識・理解シンポジウム(MIRU2009), IS3-35 (2009)
- 3) K.Fujimura, Y.Matsumoto, E.Tetuichi : Multi-camera 3D modeling system to digitize human head and body, Proc. of SPIE Vol.4298, pp.40-47 (2001)
- 4) NEC エンジニアリング(株), "Danae", <http://www.nec-eng.co.jp/pro/3d/introduction.html>
- 5) Y. Sugaya, S.Ando, I.Miyagawa, A.Suzuki, H.Koike : Factorization-based Automatic 3D Face Modeling from Turntable Image Sequence using Monocular Camera, The Journal of the IIEEJ, Vol.40, No.1, pp.114-123(2011)
- 6) V.Blanz, T.Vetter : Face Recognition Based on Fitting a 3D Morphable Model, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.25, No.9, pp.1063-1074 (2003)

- 7) A.Maejima, S.Morishima : Fast Plausible 3D Face Generation from a Single Photograph, ACM SIGGRAPH ASIA 2008, Poster, ISBN 978-1-60558-388-4, maejima.pdf (2008)
- 8) 入江淳, 高際睦起, 森山孝三, 山下隆義 : 回帰を用いた階層的フィッティングによる顔輪郭検出の高精度化, 第 17 回画像センシングシンポジウム(SSII2011), IS1-06 (2011)
- 9) C. Tomasi, T. Kanade : Shape and Motion from Image Streams under Orthography: a Factorization Method, International Journal of Computer Vision 9, pp.137-154 (1992)
- 10) 前島謙宣, 森島繁生 : 最適化局所アフィン変換に基づく正面顔レンジスキャンデータからの頭部モデル自動生成, 画像電子学会論文誌, Vol.38, No.4, pp.404-413 (2009)