

表情変動を許容した実時間頭面部姿勢推定のための 個人間および個人内変動に対する顔形状推定

菅野 裕介[†] 佐藤 洋一[†]

[†] 東京大学 生産技術研究所
〒 153-8505 東京都目黒区駒場 4-6-1
E-mail : {sugano, ysato}@iis.u-tokyo.ac.jp

あらまし

本論文では、任意の人物を対象として変形を含む頭面部姿勢推定を実現するための顔形状推定手法について提案する。顔形状に含まれる個人間変動と個人内変動を二つの独立したパラメータにより記述したモデルを用いることで、時間依存の変形パラメータに対象を絞った頭面部姿勢推定が可能になる。しかし、時間に依存しない個人差パラメータを複数フレームの情報を元に決定することに関しては、これを実時間処理の枠組みで実行する手法は提案されていなかった。本手法では Model-based Bundle Adjustment の枠組みをパラメータ分離モデルの個人差調整に用いることにより、実時間システムにも適用可能な形で個人差パラメータの調整を実現する。さらに、この調整処理を実時間の連続フレームに対して適用するための手法について検討し、その効果に関して評価実験を行った。

Estimation of Interpersonal and Intrapersonal Shape Deformations for Real-Time Monocular 3D Face Tracking

Yusuke Sugano[†] Yoichi Sato[†]

[†]Institute of Industrial Science, The University of Tokyo
4-6-1 Komaba, Meguro-ku,
Tokyo 153-8505, JAPAN
E-mail : {sugano, ysato}@iis.u-tokyo.ac.jp

Abstract

For the purpose of building a real-time 3D face tracking system, we propose a new method for shape estimation of human faces. Using face models with separate parameters for shape (interpersonal differences) and action (intrapersonal ones), it is possible to estimate user's facial expression with time-varying action parameters. However, there is no real-time technique for resolving time-invariant shape parameters in multiframe processing. To achieve this, we use a framework of model-based bundle adjustment with a dual parameter model. And we demonstrate its accuracy with some experiments.

1 はじめに

コンピュータビジョンの分野において、3次元空間内における人間の頭面部姿勢を実時間で追跡・測定するための技術は古くから重要な課題の一つとなっている。その目的の一つとして人

の顔に関する情報を利用したヒューマン・コンピュータ・インタラクション技術の構築が挙げられるが、近年では、ビデオチャットやテレビ電話などの普及とカメラの低価格化に伴い、大規模な設置型システムに留まらず、個人のデスクトップ環境やモバイルデバイスなど、インタ

ラクション技術をより多様な場面に適用することが可能になってきている。こうした中で幅広いアプリケーション応用を想定した頭部姿勢推定システムを構築する上では、次のような点が重要となる。

- 個人毎の事前準備などを必要とせず、任意のユーザに対してすぐに推定を開始できる
- 発話や表情変化に伴う顔変形が起こった場合でも、安定して推定を継続できる
- 設置コストが低く、1台のカメラのみを用いてシステムを構築できる

頭部姿勢推定技術に関しては現在も多くの手法が提案されているが [1, 6, 2, 5, 9, 11, 15], 変形推定のためにユーザ毎の事前準備が必要となるものや複数台のカメラを必要とするもの、対象が剛体運動に限られるものなど、実用性という意味では不十分な点が残る手法がほとんどである。

本稿では、上で述べた3つの条件を満たす頭部姿勢推定を実現するための枠組みとして、顔形状の個人間変動と個人内変動を分離して表現するパラメータ分離モデルを用いた頭部姿勢推定手法を提案する。変形パラメータの逐次推定と個人差パラメータの解析を実時間処理の中で統合するために、従来は剛体モデルの枠組みで用いられてきた Model-based Bundle Adjustment の枠組みをパラメータ分離モデルに導入する。さらに、パラメータ制約を効果的に用いることにより、動画フレームに対する実時間処理に適した調整を実現する。そして、本手法を用いた姿勢推定の精度を評価実験により検証する。

以降、まず第2章では関連研究とそれに対する本研究の位置付けについて述べる。さらに第3章では具体的な本手法の構成について、第4章ではそれにもとづく評価実験の結果について述べ、最後に第5章で全体のまとめと今後の課題について述べる。

2 関連研究

本章では、提案手法の関連研究について述べる。まず第2.1節では、パラメータ分離モデルの

利用について、実時間頭部姿勢推定と顔形状構築の二つの観点から整理する。そして第2.2節では本手法で用いる Model-based Bundle Adjustment について述べ、提案手法の位置付けを整理する。

2.1 パラメータ分離顔形状モデルの利用

安定した頭部姿勢推定を実現する上でユーザの顔変形への対応は重要な課題となるため、複雑なバリエーションを含む人間の顔をいかにモデル化するかが非常に重要な問題となる。

顔形状の変動における要因は、大きく次の二つに分けて捉えることができる。

- 目の位置や鼻の高さといった、個人差に対応する要素
- 眉の上下や口の開閉といった、個人内の変形に対応する要素

これらを全て単一のパラメータで表現することで特定のユーザに依存しない汎用性の高いモデルを作成する場合、パラメータの次元数を低く抑えることは困難である。これに対し二つの要素をそれぞれ異なるパラメータにより記述したモデルは、近似精度を保ったまま個々のパラメータの次元を低く抑えることが可能になる。さらに形状変化の要因を分けて捉えることができるため、表情認識などの応用を考える上でもメリットが大きい。

2.1.1 実時間頭部姿勢推定

頭部姿勢推定のための手法では、Matthewsら [9]による複数人のデータを含む Active Appearance Model のように単一パラメータのモデルを用いて汎用的な顔モデルを構築する例も多く見られる。しかし、実時間処理を行う際、処理時間や安定性の上でパラメータの次元数が大きいことは大きなデメリットになる。

一方、パラメータ分離モデルを想定した場合、実時間追跡の問題において個人内で変動するパラメータは変形パラメータに限られる。Dornaikaらの手法 [5]では個人差パラメータと変形パラメータを線形分離したメッシュモデルが用

いられており、変形パラメータに対象を絞った状態推定を実現している。しかし、このとき個人差パラメータは各ユーザに対して追跡ごとに決定する必要がある。

パラメータ分離モデルを用いた手法を含め、これまでに提案されてきた変形を伴う頭部姿勢推定手法は個々のユーザ毎にモデルの調整が必要となるものがほとんどである。実時間システムの中で、不特定多数のユーザに対して個人差パラメータを決定するための手法に関しては十分な議論がなされていない。

2.1.2 ビデオ入力にもとづく顔形状構築

一方、時系列のビデオ入力を元に3次元の顔形状を構築するための技術においても、パラメータ分離モデルを用いた手法がいくつか提案されている。時間的に変化する変形パラメータとは異なり、ユーザごとに一定の値となる個人差パラメータは、複数フレームの情報を用いて解析的に決定することができる。

DeCarloらはモデルベースオプティカルフローの残差成分を利用することでフレームに共通の形状パラメータを調整する手法を提案している[3]。また、VlasicらはN-mode SVDにもとづく解析によりパラメータが分離された多重線形モデルを用いることで、手作業でパラメータを分離したアニメーション用のモデルを用いる手法と比べてより洗練された形で顔変形のモデリングを行う手法を提案している[12]。しかし、これらの手法は撮影した動画の後処理を前提としており、実時間処理にそのまま適用することはできない。

2.2 Bundle Adjustment の枠組みによるパラメータ調整

前節で述べた二つの観点による手法を実時間処理として統合するために、本手法ではModel-based Bundle Adjustmentの枠組みを利用する。

Bundle Adjustmentは最尤推定手法の一つであり、複数フレームに跨る誤差関数を最小化することでカメラ姿勢とモデル形状を同時に調整する手法である。映像から3次元モデルを復元

するStructure from Motion技術などにおいて、おおまかな推定形状を精緻化するための手法として広く用いられている。近年、Bundle Adjustmentの枠組みにパラメータ表現された事前構築モデルを取り入れるModel-based Bundle Adjustment手法が注目されている[4, 10]。Xinらはこれを利用して、安価なカメラからのビデオ入力を元に効率よく頭部の3次元形状を自動構築する手法を提案している[13]。また、Vachettiらの手法[11]ではこれを実時間の頭部姿勢追跡に応用することで、任意のユーザに対して頭部の剛体運動推定を可能にしている。

これまでの例ではBundle Adjustmentは基本的に剛体の枠組みで用いられており、XinらやVachettiらの手法も同様に、対象は剛体運動に限られていた。本手法ではパラメータ分離モデルにこの枠組みを導入することにより、姿勢変動と同時に変形を行うオブジェクトに対するパラメータ調整を実現する。

すなわち、次の2つの観点から本研究の主要な貢献を挙げることができる。まず第一に、パラメータ分離モデルにおいて複数フレームの処理による個人差成分の決定を行う手法を実時間処理の枠組みの中に統合すること。そして第二に、Model-based Bundle Adjustmentの枠組みを変形を含む形状取得に応用し、連続フレームの処理に適した調整手法を提案することである。

3 システムの概要

本章では、提案手法の具体的な処理の流れについて述べる。まず3.1節ではパラメータ分離顔形状モデルの作成手法について、次に3.2節ではそのモデルを用いた頭部姿勢推定の枠組みについて述べる。

3.1 パラメータ分離顔形状モデルの作成

本研究では、パラメータ分離モデルとしてN-mode SVDにもとづく多重線形モデルを利用した。サンプルとして用意した形状データをもとに、任意のユーザに対応できる汎用的な顔モデルを作成する。

本手法で用いる頭部モデルは K 個の特徴点を持っており、各々の特徴点はユーザ頭部に固定されたモデル座標系内での 3次元座標と特徴点のテンプレート画像により表現される。ここでは、 M を K 個の特徴点の 3次元座標で構成された $3K$ 次元形状ベクトルとする。本論文で使用するモデルは、 $K = 10$ で固定されており、各点は両目の両端と口の両端、両鼻孔、そして両眉の内側の端点に対応する。図 1 中、写真内の + マークが各特徴点を示している。本手法では、この形状ベクトル M を以下のような形で事前にモデル化している。

モデル構築の段階では、先行研究 [15] にもとづく頭部姿勢推定を利用した。これは複数台のカメラを利用したシステムになっており、ステレオマッチングにより取得した顔形状をもとにユーザ独自の变形モデルを逐次的に構築する。これにより、任意のユーザに対する頭部姿勢及び顔形状の推定を実現している。

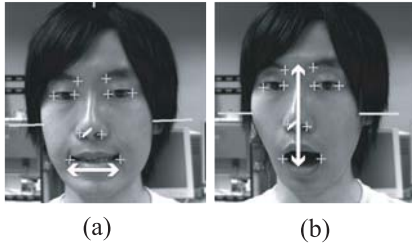


図 1: モデル化の対象となる顔変形の例

4章で述べる本論文の実験では、モデル化の対象とする変形は次の二種類に絞った。一つは図 1(a) のような口の横方向の運動であり、もう一つは同図 (b) のような眉の上下を含む顔全体の縦方向の運動である。モデル構築のサンプルとして用いる S 人の人物それぞれについてこれらの運動を行う様子を一定時間記録し、ここから自動的に形状データを抽出する。それぞれの運動には基準となる距離 l (図 1 中の矢印に対応) を定め、まず人物毎に l の最大値 l_{max} と最小値 l_{min} を計測する。これにより定義される変形率 $r = (l - l_{min}) / (l_{max} - l_{min})$ を基準にして、 $r = 0.0 \sim 1.0$ の範囲で等間隔に、合計 A 通りの形状データを取得する。

さらに、取得した形状データは顔中心で左右

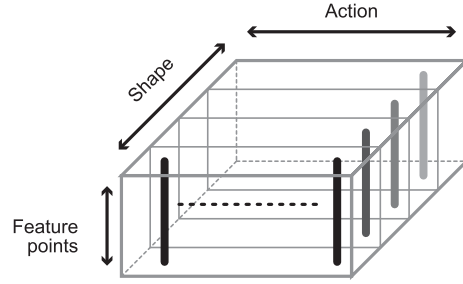


図 2: モデル構築に用いるデータテンソル

に分割し、それぞれを反転させて生成した形状を二つの独立したサンプルとして扱う。これによりモデルを左右対称なものとし、不要な回転成分が含まれるのを防いでいる。

このようにして得られたデータを元に、図 2 のような $3K \times 2S \times A$ 次元のデータテンソル \mathcal{T} を作成する。図中、Feature points 方向には顔形状ベクトル M が対応し、Action 方向には同一人物のデータが、Shape 方向には r の値が等しい変形が、それぞれ対応付けられた形で格納される。

N-mode SVD を用いてテンソルを分解することで、データテンソル \mathcal{T} は次のようにコアテンソル \mathcal{C} と各モード i の基底行列 U_i とのテンソル積として記述できる [12].

$$\begin{aligned} \mathcal{T} &= \mathcal{C} \times_{\text{feature}} U_{\text{feature}} \\ &\quad \times_{\text{shape}} U_{\text{shape}} \times_{\text{action}} U_{\text{action}} \\ &= \mathcal{M} \times_{\text{shape}} U_{\text{shape}} \times_{\text{action}} U_{\text{action}} \quad (1) \end{aligned}$$

このとき、 U_i の列方向が各要素の基底ベクトルに対応している。さらに、 U_i の次元数を削減した低次元の基底行列 \check{U}_i を用いることで、次のようにデータテンソルの近似表現が得られる。

$$\mathcal{T} \approx \check{\mathcal{M}} \times_{\text{shape}} \check{U}_{\text{shape}} \times_{\text{action}} \check{U}_{\text{action}} \quad (2)$$

このとき、 \check{U}_i の各行はデータテンソルに含まれる特定の変形や個人に対応する係数となっており、個人差パラメータベクトル \mathbf{s} と変形パラメータベクトル \mathbf{a} によって次のように任意の顔形状ベクトル M を記述することができる。

$$M = \check{\mathcal{M}} \times_{\text{shape}} \mathbf{s}^T \times_{\text{action}} \mathbf{a}^T \quad (3)$$

3.2 制約を利用した逐次的パラメータ調整

以上の手順により作成した顔形状モデルに対して、連続フレームの情報を利用した個人差パラメータの調整を実行する。本章ではその具体的な手順について述べる。

まず、Bundle Adjustment は一般的に次のような誤差関数の最小化問題となる。

$$\min_{\mathbf{P}_i, \mathbf{X}_k} \sum_{i=1}^x \sum_{k=1}^n D(\mathbf{x}_{ik}, \mathbf{P}_i \mathbf{X}_k)^2 \quad (4)$$

ここで \mathbf{X}_k はモデル中 k 番目の点の 3 次元座標であり、 \mathbf{P}_i は i 番目のフレームにおける投影行列、 \mathbf{x}_{ik} は、フレーム i の入力から画像特徴により得られた点 k の正確な 2 次元座標を示す。 $D(\mathbf{x}_{ik}, \mathbf{P}_i \mathbf{X}_k)^2$ は、画像中の特徴点 \mathbf{x}_{ik} とモデルにもとづく投影点 $\mathbf{P}_i \mathbf{X}_k$ とのユークリッド距離である。すなわち式 (4) は、動画フレームから得られた特徴点の 2 次元軌跡を元にして、投影点が軌跡に最も近くなるようなモデルの 3 次元形状とカメラ姿勢を推定する式になっている。

本手法で用いる顔形状モデルを用いた場合、Bundle Adjustment によるパラメータ調整は次のようになる。まず、フレーム i におけるユーザ頭部の状態は、世界座標系からモデル座標系への並進と回転を表す 6 次元頭部姿勢ベクトル \mathbf{p}_i と変形パラメータベクトル \mathbf{a}_i 、そして時間に依存しない個人差パラメータベクトル \mathbf{s} によって記述できる。このとき、次のような投影関数 \mathcal{P} を定義することができる。これは、 \mathbf{a}_i 、 \mathbf{s} から式 (3) によって決まる顔形状ベクトル \mathbf{M}_i に頭部姿勢 \mathbf{p}_i に応じた並進や回転を適用した上で、 \mathbf{M}_i の各特徴点を画像平面に投影する関数である。

$$\mathbf{m}_i = \mathcal{P}(\mathbf{p}_i, \mathbf{M}_i) \quad (5)$$

すなわち、 \mathbf{m}_i は K 個の投影点の 2 次元座標で構成される $2K$ 次元ベクトルとなる。一方、フレーム i においてカメラ画像 \mathbf{I}_i が得られたとき、各点のテンプレート画像を元に \mathbf{I}_i 中から探索した真の特徴点 2 次元座標を $\hat{\mathbf{m}}_i$ とすると、第 t フレームにおいて次のような過去 n フレーム分の誤差関数が定義できる。

$$F_t = \sum_{i=t-n+1}^t D(\hat{\mathbf{m}}_i, \mathcal{P}(\mathbf{p}_i, \mathbf{M}_i))^2 \quad (6)$$

本手法では Zhang らの手法 [14] と同様に、新たな観測 \mathbf{I}_i が得られるごとに F_t の最小化を行うことで姿勢 \mathbf{p}_t と変形 \mathbf{a}_t を推定すると共に、同時に個人差 \mathbf{s} を更新していく。一般的に、形状に比べて姿勢はより安定した推定が可能であることが知られており、本手法ではこれを利用した 2 段階の最小化を行う。さらに、パラメータの調整範囲に対する制約条件を伴う Levenberg-Marquardt アルゴリズム [7, 8] を用いてこの最小化を行うことで、フレームの連続性を考慮したパラメータの調整を行っている。本手法で用いる調整条件の詳細は次のようになる。以下、 $\mathbf{p} = (\mathbf{p}_{t-n+1}^T, \dots, \mathbf{p}_t^T)^T$ 、 $\mathbf{a} = (\mathbf{a}_{t-n+1}^T, \dots, \mathbf{a}_t^T)^T$ とし、あるパラメータ \mathbf{x} に関する制約を \mathbf{C}_x 、調整における初期値を $\hat{\mathbf{x}}$ と表記する。

第一段階

まず、次のように顔形状は固定として姿勢のみに関する最適化を行う。

$$\min_{\mathbf{p}} F_t, \quad \mathbf{p} \in \mathbf{C}_p \quad (7)$$

このとき、 $\hat{\mathbf{p}}$ および \mathbf{C}_p は次のように設定する。

$$\hat{\mathbf{p}}_i = \begin{cases} \mathbf{p}_i^{(t-1)} & (i < t) \\ \mathbf{p}_{i-1}^{(t-1)} & (i = t) \end{cases} \quad (8)$$

$$\mathbf{C}_p = \{\mathbf{p} \mid \hat{\mathbf{p}}_i - \lambda_p \leq \mathbf{p} \leq \hat{\mathbf{p}}_i + \lambda_p\} \quad (9)$$

ここで $\mathbf{p}_i^{(j)}$ はフレーム j までの観測にもとづくフレーム i の調整結果を示し、すでに調整が行われているフレーム $t-1$ まではこれを初期値として用いる。パラメータが未知である新たなフレーム t に関しては、隣接するフレーム $t-1$ の値を初期値として用いる。また、 λ_p は定数ベクトルであり、初期値から一定の範囲内に収まるような制約を与えている。

第二段階

次に、形状を含めた全てのパラメータに対して最適化を行う。

$$\min_{\mathbf{p}, \mathbf{a}, \mathbf{s}} F_t, \quad \mathbf{p} \in \mathbf{C}_p, \quad \mathbf{a} \in \mathbf{C}_a, \quad \mathbf{s} \in \mathbf{C}_s \quad (10)$$

まず、姿勢パラメータに関しては初期値として第一段階の調整結果 \mathbf{p}' を用いる。

$$\hat{\mathbf{p}} = \mathbf{p}' \quad (11)$$

C_p は式 (9) に従うが、 λ_p は第一段階よりも小さな値を新たに設定する。また、変形パラメータに関しても姿勢と同様の考えにもとづき次のように \hat{a} , C_a を設定する。

$$\hat{a}_i = \begin{cases} a_i^{(t-1)} & (i < t) \\ a_{i-1}^{(t-1)} & (i = t) \end{cases} \quad (12)$$

$$C_a = \{a \mid \hat{a}_i - \lambda_a \leq a \leq \hat{a}_i + \lambda_a\} \quad (13)$$

ただし、変形パラメータにはモデルにもとづく最大値 a_{max} , 最小値 a_{min} を定め、 $\hat{a}_i + \lambda_a$ が a_{max} を超える場合は a_{max} を調整の上限としている（下限も同様）。

一方、個人差パラメータは式 (14) のように、初期値として過去 $N (> n)$ フレーム分の調整結果の平均を与える。

$$\hat{s} = \frac{1}{N} \sum_{i=t-N}^{t-1} s^{(i)} \quad (14)$$

$$s \in \{s \mid \hat{s} - \lambda_s \leq s \leq \hat{s} + \lambda_s\} \quad (15)$$

さらに、式 (15) における制約範囲 λ_s は姿勢や変形と比べて小さい値を設定する。これにより、個人差パラメータに関してはフレーム間の細かい変動を制限しながら長期的な観測にもとづく調整を行う。

4 評価実験

3章で述べた枠組みを元に、評価実験を行った。

本実験では、ユーザが頭部の移動と顔の変形を同時に行っている 1800 フレーム分の画像列を対象に、頭部姿勢及び顔形状の推定を行う。まず、誤差関数 F_t における観測ベクトル \hat{m} を取得するために、モデル構築の際に用いたものと同じ複眼システムによって顔形状、頭部姿勢を計測する。この段階では 2 台のカメラを利用した。これによって得られるカメラ 1 台分の顔特徴点二次元軌跡を元に提案手法によるパラメータ推定を行い、複眼システムによる推定の結果と比較することで提案手法の精度を評価する。

実験では Intel Pentium4 3.0GHz と Windows XP を搭載した汎用 PC を 1 台使用した。入力画像としては、IEEE1394 カメラにより取り込まれた 640×480 画素の画像を使用した。また、

	x	y	z
平均 [mm]	5.21	5.35	23.45
標準偏差 [mm]	2.48	2.87	12.48

	roll	pitch	yaw
平均 [degree]	0.45	3.81	1.90
標準偏差 [degree]	0.33	2.36	1.48

表 1: 提案手法による頭部姿勢の推定誤差

実験は 11 人のユーザから取得した 10 個の変形（図 1 の二つの変形について、それぞれ 5 個）をもとに作成したモデルを使用し、パラメータの次元数はそれぞれ個人差 7, 変形 5 とした。なお、姿勢推定の実験対象としたユーザの情報はモデルに含まれている。調整に用いるフレーム数は $n = 5$, また個人差パラメータの初期値を算出するためのフレーム数は $N = 50$ とした。1~5 フレームの調整における初期値はある程度現実的なものとして、姿勢については画面のほぼ中央に顔が位置するような値を、また形状パラメータはほぼ 0 に近い値を与えている。Levenberg-Marquardt アルゴリズムにおける繰り返しの回数は、第一段階と第二段階それぞれ 5 回、合計 10 回に制限した。このとき、実時間処理に適用可能な平均 20.6[ms/frame] の計算速度を実現している。

図 4 はそれぞれの手法による推定結果のグラフを示す。この図において、細い線は複眼システムによる推定結果であり、太い線は提案手法を用いた単眼による推定結果である。推定値の平均誤差及び標準偏差は表 1 のようになっている。並進・回転ともに精度の良いパラメータ推定を実現しており、単眼では推定が難しい奥行き方向の変動に対しても高い安定性を示している。また、図 3 は頭部姿勢推定の結果画像を示す。この図には推定された頭部姿勢に対応するモデル座標軸が描かれている。また、図中の細い×マークは観測座標 \hat{m} , 太い+マークは推定結果にもとづく座標 m に対応する。

一方、図 5 は顔形状の推定値に関して、頭部座標系における特徴点ごとの平均誤差を示す。変形を伴う顔形状に対しても、一定の誤差範囲内

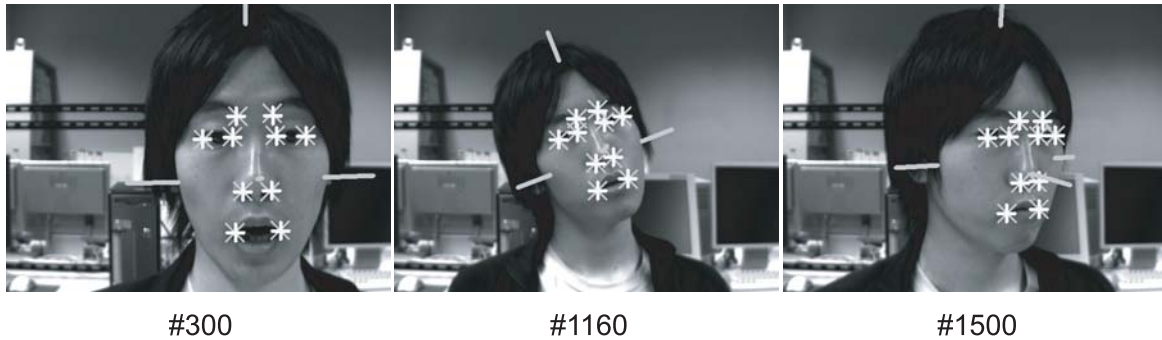


図 3: 頭部姿勢推定の結果画像

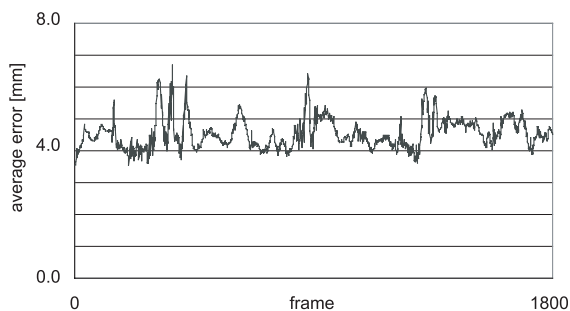


図 5: 推定された顔特徴点座標（頭部座標系）の平均誤差

で形状を推定できていることがわかる。

5 おわりに

本論文では、任意の人物を対象として顔変形を含む頭部姿勢推定を実現するための手法について述べた。本手法の貢献としては、次の2点が挙げられる。

1. パラメータ分離顔形状モデルに関して、複数フレームの情報をを用いた個人差決定を実時間処理に統合した。
2. Model-based Bundle Adjustment の枠組みを、変形を含む対象のパラメータ推定に応用した。

提案手法により高い精度で顔形状および頭部姿勢の推定が可能になることを、評価実験により示した。

今後の課題として、現在は複眼システムを用いて実現している特徴点軌跡の追跡処理を実時

間で実装することで、実際に単眼カメラによる頭部姿勢推定システムを構築することが挙げられる。パーティクルフィルタを用いた追跡や顔検出処理を組み合わせることで、安定した単眼推定を実現することを目標としている。

また、推定の精度向上を考える上ではモデル自体の改善も重要な課題になる。評価実験において、特に *pitch* 方向の推定などに比較的大きな誤差が見られるが、こうした誤差にはモデル自体に含まれる回転の要素が大きく影響しているものと考えられる。このような姿勢変動との切り離しが難しい変形要素は、人間の顔形状モデルから完全に切り除くことはできないが、提案手法の枠組みにおいては、これらの要素を可能な限り個人差パラメータで表現することにより姿勢変動への影響を軽減できる。今後は、より姿勢推定に適したモデルの構築についても検討を行う。さらに、パラメータ分離モデルのさらなる応用として、より詳細なモデルを用いた顔変形の取得と表情解析などの課題にも取り組んでいく予定である。

参考文献

- [1] Black, M. and Yacoob, Y.: Tracking and recognizing rigid and non-rigid facial motions using local parametric models of image motion, *Proc. IEEE Int. Conf. Computer Vision*, pp.374-381 (1995).
- [2] DeCarlo, D. and Metaxas, D.: Optical flow constraints on deformable models with applications to face tracking, *Int. J. Computer Vision*, Vol. 38, No. 2, pp. 99-127 (2000).
- [3] DeCarlo, D. and Metaxas, D.: Adjusting shape parameters using model-based optical flow residuals, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 24, No. 6, pp. 814-823 (2002).

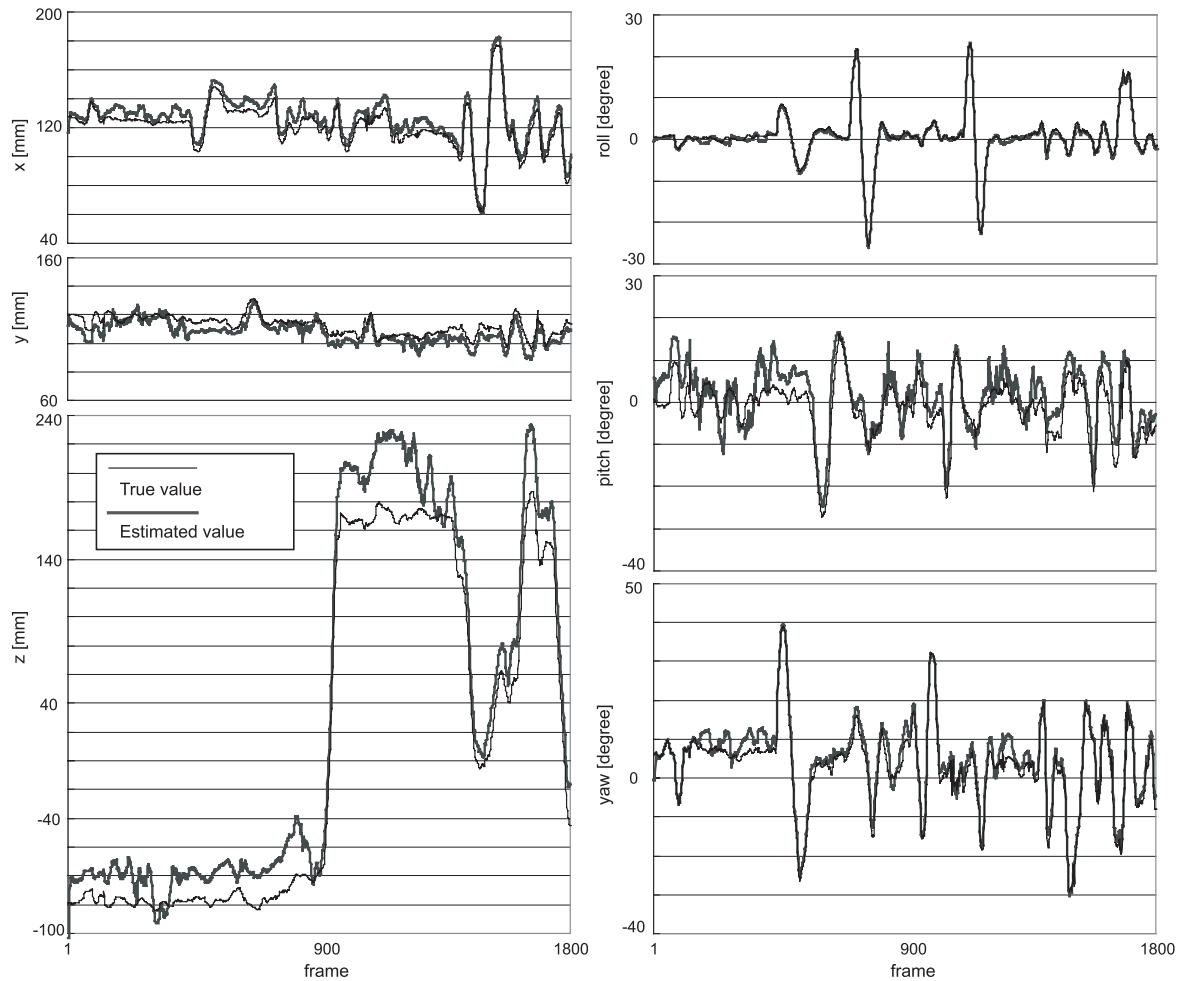


図 4: 提案手法を用いた頭部姿勢推定の結果. x は水平方向, y は鉛直方向, z は深さ方向の運動であり, $roll$ は z 軸回り, yaw は y 軸回り, $pitch$ は x 軸回りの回転である.

- [4] Dimitrijevic, M., Ilic, S. and Fua, P.: Accurate face models from uncalibrated and ill-lit video sequences, *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition* (2004).
- [5] Dornaika, F. and Davoine, F.: Simultaneous facial action tracking and expression recognition using a particle filter, *Proc. IEEE Int. Conf. Computer Vision* (2005).
- [6] Jebara, T. and Pentland, A.: Parametrized structure from motion for 3D adaptive feedback tracking of faces, *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*, pp.144–150 (1997).
- [7] Kanzow, C., Yamashita, N. and Fukushima, M.: Levenberg-Marquardt methods for constrained nonlinear equations with strong local convergence properties, *J. Computational and Applied Mathematics*, Vol. 172, pp. 375–397 (2004).
- [8] Lourakis, M.I.A.: levmar: Levenberg-Marquardt nonlinear least squares algorithms in C/C++, <http://www.ics.forth.gr/~lourakis/levmar/> (2004).
- [9] Matthews, I. and Baker, S.: Active appearance models revisited, *Int. J. Computer Vision*, Vol. 60, No. 2, pp. 135–164 (2004).
- [10] Shan, Y., Liu, Z. and Zhang, Z.: Model-based bundle adjustment with application to face modeling, *Proc. IEEE Int. Conf. Computer Vision*, pp. 644–651 (2001).
- [11] Vacchetti, L., Lepetit, V., and Fua, P.: Stable real-time 3D tracking using online and offline information, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 26, No. 10, pp. 1380–1384 (2004).
- [12] Vlasic, D., Brand, M., Pfister, H., and Popovic, J.: Face transfer with multilinear models, *ACM Transactions on Graphics (Proc. ACM SIGGRAPH 2005)*, Vol. 24, No. 3, pp. 426–433 (2005).
- [13] Xin, L., Wang, Q., Tao, J., Tang, X., Tan, T. and Shum, H.: Automatic 3D face modeling from video, *Proc. IEEE Int. Conf. Computer Vision*, pp. II-1193–1199 (2005).
- [14] Zhang, Z. and Shan, Y.: Incremental motion estimation through modified bundle adjustment, *Proc. IEEE Int. Conf. Image Processing*, pp. II-343–346 (2003).
- [15] 岡兼司, 菅野裕介, 佐藤洋一: 頭部変形モデルの自動構築を伴う実時間頭部姿勢推定, *情報処理学会論文誌* Vol. 47 No. SIG 10 (CVIM 15) pp. 185–194 (2006).