

輪郭によるフェイスアライメントにおける姿勢変化への対応 のための顔輪郭の利用について

高野 博幸^{1,a)} 出口 光一郎^{1,b)}

概要：画像上の顔を認識し各部位の詳細な位置決めを行うフェイスアライメントでは、横顔になると正面顔では見えていた顔の特徴が遮蔽されて見えなくなり、顔の方向の推定が不安定になる。顔の輪郭は、顔と背景の境界として重要な特徴であるが、顔の側面の遮蔽輪郭は角度によって対応する顔の部位が異なり使うことが難しかった。だが顔の遮蔽輪郭を積極的に使って顔の角度をより正確に推定し得る可能性がある。本稿では顔の多様な角度に対し遮蔽輪郭を利用するために、角度毎に輪郭の形状を調べ、また個人の顔の形状に依存する差を調べた。次に遮蔽輪郭から他の顔の部位の位置を推定し、その精度を評価する。そして、推定された部位を使って Constrained Local Model の枠組みでパラメータ最適化を行いフェイスアライメントの精度を改善する。

1. はじめに

本論文ではカメラに対して正面顔から横顔まで、様々な向きの顔に対して、画像上の目、口、鼻、眉そして輪郭などの位置決め（以降、フェイスアライメントと呼ぶ）する手法を提案する。フェイスアライメントを行うための代表的な手法は、Point Distribution Model(PDM) [1] で表現される顔のモデル（図 1 にモデルを構成する点を白点で示す）と顔画像の濃淡パターンなどで作られるアピアランスで特徴検出器を作成し、入力画像とマッチングするものである。AAM(Active Appearance Model) [2], [3], [4], ASM(Active Shape Model) [5], そして CLM(Constrained Local Model) [6], [7] などが提案されており、顔がカメラに対して正面付近を向いている範囲で精度よく計測できる。

しかし、大きく横を向いた場合、それらの性能は著しく低下してしまうことが知られている [8]。従来法ではモデルの輪郭（以下、モデル輪郭と呼ぶ）は正面顔の輪郭で作成しているが（図 1 において輪郭は耳からあごの辺りの点）、図 2 に示すように正面顔と横顔の輪郭は顔上の異なる特徴位置を指し示している。特徴点の顔上の位置がポーズによって異なるため、モデル輪郭と特徴点が対応しなくなりパラメータ最適化に失敗してしまう。その結果、アライメント

できる角度範囲が正面周辺に制限される問題があった。このような理由から輪郭をフェイスアライメントに使うことが難しかった。

人間の顔の形状において、奥行きである Z 方向の長さは個人毎にあまり変わらないと言われている。人間の頭部を円筒形の剛体とみなして高い精度でポーズを推定する手法 [9], [10], [11] も提案されている。人間の頭部を単純な形状に近似してもポーズ推定を行うことができる。このことから図 3 の右下に示すように、画像上に現れた横顔の輪郭（以降、遮蔽輪郭と呼ぶ）を補正してモデルに使われている部位は推定できると考えられる。その補正が実用上必要とされる精度を満たすなら、従来の遮蔽輪郭の問題を解決できる可能性がある。

本論文では顔の遮蔽輪郭を積極的に使って顔の角度をより正確に推定する。顔の多様な角度に対し遮蔽輪郭を利用するために、角度毎に輪郭の形状を調べた。また個人の顔の形状に依存する差を調べた。そして図 3 の右上に示すように遮蔽輪郭から他の顔の部位の位置を推定し、その精度を評価した。そして、推定された部位を使って Constrained Local Model の枠組みでパラメータ最適化を行いフェイスアライメントの精度を改善する。

性能評価により横向きの角度 90° までフェイスアライメントを行うことができることを確認した。本手法を顔の追跡に適用することにより、激しく様々なポーズに動く人物の頭部を安定して追跡することができた。

本論文では、第 4 章にて提案手法を示すが、それが必要な理由を明確にするためにまず第 2 章にて CLM フレーム

¹ 東北大学 情報科学研究科 システム情報科学専攻
6-6-01 Aza Aoba, Aramaki, Aoba-ku, 980-8579 Sendai,
Japan

^{f1} 現在、東北大学
Presently with Tohoku University

^{a)} takano@fractal.is.tohoku.ac.jp

^{b)} kodeg@fractal.is.tohoku.ac.jp

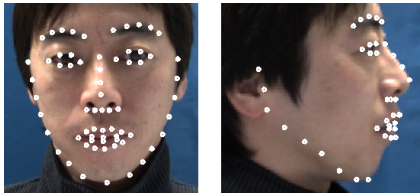


図 1 モデルの特徴点位置 (出典 [12])
Fig. 1 The positions of model points.

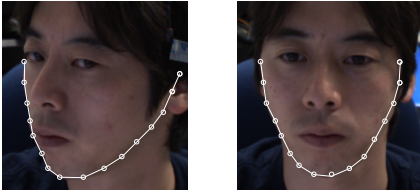


図 2 ポーズによる遮蔽輪郭の変化 (出典 [12])
Fig. 2 The position of an image contour changes depending on a head pose.

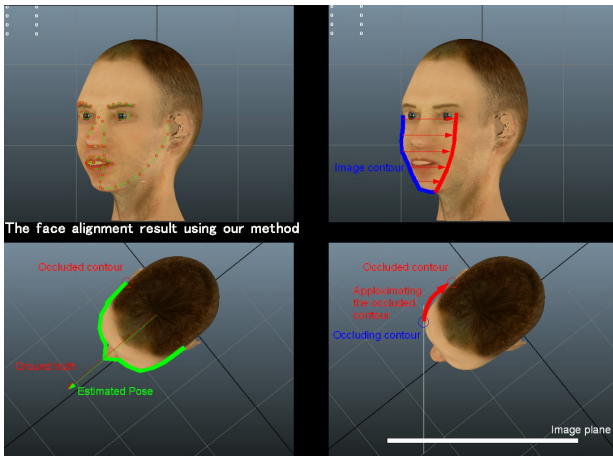


図 3 遮蔽輪郭の補正 左上:提案手法によるアライメント 右上:遮蔽輪郭の補正 左下:正しいポーズ 右下:真上から見た遮蔽輪郭の補正

Fig. 3 Compensation of the occluding contour. Left top image:FaceAlignment with our method, Right top image:Compensation of the occluding contour, Left bottom image:Correct pose, Right bottom image:Compensation of the occluding contour from a top view.

ワークでのフェイスアライメントの概要を示し、第 3 章で遮蔽輪郭の問題点を議論する。

2. CLM によるフェイスアライメント

本論文のフェイスアライメントは顔モデルを PDM で表現し、CLM の枠組みで計算される。CLM の基本処理は第一に特徴点検出器により入力画像から顔の特徴点を検出するステップ、次に顔モデルと特徴点との距離が最小になるようにモデルパラメータを最適化するステップの 2 つのステップからなる。

2.1 PDM と特徴検出器

CLM は顔モデルである PDM と濃淡パタンのアピアランスから作られた特徴点検出器で成る。PDM は次元数に拘ないが、提案手法では三次元座標で顔モデルを表現した。

2.1.1 PDM

PDM は複数人の表情から 3 次元座標データ (3D シェイプ) を取得しモデル化される。図 1 に顔モデルを構成する特徴点の位置の例を示す。PDM の各点を示す 3D 位置ベクトルは

$$\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_M)^\top, \quad (1)$$

で表し、画像上の点 \mathbf{x} は \mathbf{X} を弱透視投影して求める。 $\mathbf{X}_i = (x_i, y_i, z_i)^\top$ は PDM の i 番目の特徴点を示す。PDM は、

$$\mathbf{X} = \bar{\mathbf{X}} + \Phi \mathbf{q}, \quad (2)$$

と表し、 Φ は表情の動きと個人毎の違いを PCA でモデル化した行列、 \mathbf{q} はそのパラメータ、そして $\bar{\mathbf{X}}$ は 3 D シェイプの平均を表す。モデルの i 番目の画像上の特徴点 $\mathbf{x}_i(\mathbf{p})$ は以下のように画像上に弱透視投影される。

$$\mathbf{x}_i(\mathbf{p}) = s \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \mathbf{R} [\bar{\mathbf{X}}_i + \Phi_i \mathbf{q}] + \mathbf{t}. \quad (3)$$

パラメータ $\mathbf{p} = \{s, \mathbf{R}, \mathbf{t}, \mathbf{q}\}$ の要素の s はスケール、 \mathbf{R} は角度ピッチ α 、ヨー β 、ロール γ でなる回転行列、 \mathbf{t} は平行移動、そして \mathbf{q} は変形パラメータである。 Φ_i は Φ の i 番目の要素である。フェイスアライメントとはこの \mathbf{p} を求めることである。

2.1.2 アピアランスと特徴検出器

顔モデルを構成するそれぞれの点は、特徴検出器を持っており、入力画像から対応する特徴点を検出する。検出器は様々な種類があるが、例えばサポートベクターマシンを使い複数人の特徴点位置のアピアランスを学習し作成される。

2.2 パラメータ \mathbf{p} の最適化

$\mathbf{x}_i(\mathbf{p})$ と入力画像から検出された特徴点 $\hat{\mathbf{x}}_i$ を用いて

$$Q(\mathbf{p}) = \sum_{i=1}^M \|\hat{\mathbf{x}}_i - \mathbf{x}_i(\mathbf{p})\|^2 + \|\mathbf{q}\|_{\Lambda^{-1}}^2. \quad (4)$$

を最小化することで \mathbf{p} を求める。 \mathbf{q} が大きくなった時に明らかに人間の顔の形から外れたフェイスアライメント結果となる場合があるので、パラメータ \mathbf{q} は平均 0、分散 Λ の正規分布に従うと仮定し、右項の第二項において \mathbf{q} が極端な値をとらない様に制約を加えている。第 4 章で遮蔽輪郭からモデル作成に使用した正面顔の輪郭”本当の輪郭”を推定する手法を導入するため式 (4) は変更されるが基本的な考え方は変わらない。

Φ に角度変化を加え人物の表情と個人毎のシェイプをモ

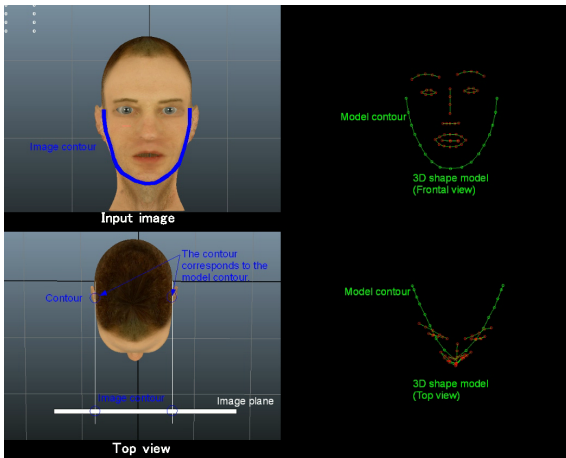


図 4 輪郭とモデル輪郭 左上:輪郭 (青線) 右上:モデルの輪郭 (緑点) 左下:真上から見た輪郭の位置 (青丸) 右下:真上から見たモデルの輪郭 (緑点)

Fig. 4 A facial contour and a model contour. Left top image: A contour (Blue points), Right top image: Model contours (Green points), Left bottom image: the contour from a top view (Blue circles), Right bottom image: Model contours from the top view (Green points).

デル化すると, q は高次元のパラメータになってしまい安定して解く事が難しくなる. 提案手法では Φ では角度変化分を表現せず, モデルを三次元で表現し弱透視投影により $x(p)$ を表現した.

3. 従来の問題点

3.1 正面顔と横顔の輪郭

正面顔における画像上の輪郭の位置とモデル輪郭を図 4 に示す. 左上の青線が輪郭を示す. 右の点群がモデルを表し, 緑点が輪郭の点である. 図の下は真上からこれらを表したものである. モデル点はパラメータ p が正確に得られたとき, 画像上の各特徴点上に投影される. しかし, 横顔の輪郭はモデル点とは別の部位となる. 図 5 の右上に横顔の輪郭を青線で示す. 図の右下に真上から表すと, 横顔では画像上に現れた輪郭 (遮蔽輪郭) は青丸で示す部位だが, 正面顔では赤丸で示す部位 (本当の輪郭) である. モデルの輪郭部分は赤丸の部位の位置で作成されているため, 横顔に対してフェイスアライメントを実行すると図 6 の左側のように遮蔽輪郭 (青丸) の影響を受け失敗する. 図 6 の緑色の線はモデルを表し, 左側の図は遮蔽輪郭を用いてフェイスアライメントを実行した例であり, 右側は正面顔で示された輪郭の部位を使用した例である. 左の図ではモデル化に使用された特徴点の部位 (本当の輪郭) ではない部位にモデルを強引にアライメントしようとしたため実際のポーズよりも小さい角度のポーズが推定された. 図 7 に横顔に対して CLM を実行した例を示す. 顔を横に向けていくと遮蔽輪郭によりフェイスアライメントに失敗する. 多様な顔の向きに対してフェイスアライメントするには遮蔽輪郭の影

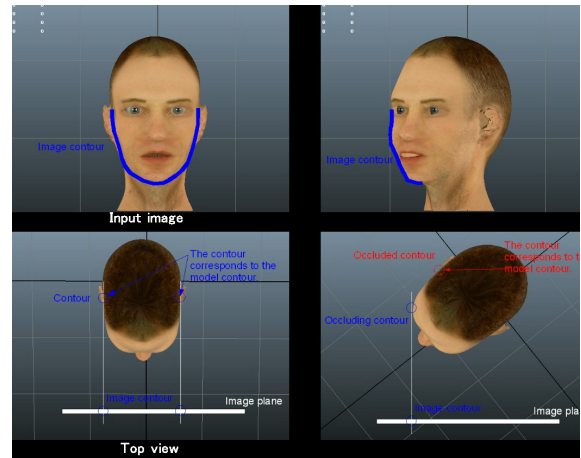


図 5 正面顔と横顔の輪郭 右上:遮蔽輪郭 (青線) 右下:真上から見た遮蔽輪郭 (青丸)

Fig. 5 A contour of both a frontal face and a profile face. Right top image: Occluding contour (Blue line), Right bottom image: the position of both the occluding contour and the occluded contour.

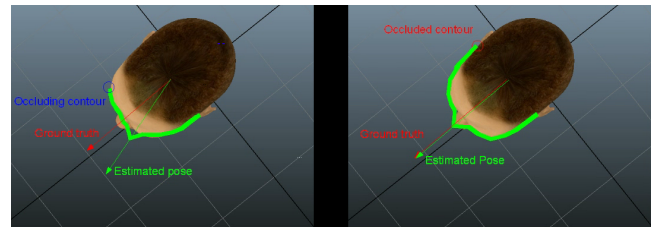


図 6 遮蔽輪郭によるポーズ推定失敗 左:遮蔽輪郭によるポーズ推定失敗 右:正確なポーズ

Fig. 6 Pose estimation error by an occluding contour. Left image: Error of the pose estimation, Right image: Correct pose.

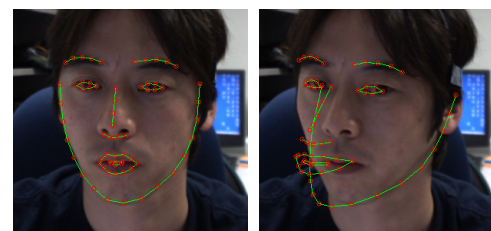


図 7 ポーズ推定失敗によるフェイスアライメント失敗

Fig. 7 Face alignment error by a pose estimation error.

響を考慮した手法が必要である.

4. 遮蔽輪郭の補正によるポーズ推定

遮蔽輪郭の問題を考慮した新しいモデルを構築するため, 画像から検出した遮蔽輪郭の位置を補正する方法を検討した. 図 3 に本提案手法のアイデアを示す. 正確にフェイスアライメントをするには図左下のように本当の輪郭を使用しなければならない. しかし, その赤丸の部位は顔の向きによりセルフオクルージョンが発生して画像から検出できない. 画像から検出できる部位は右下の図において青丸で

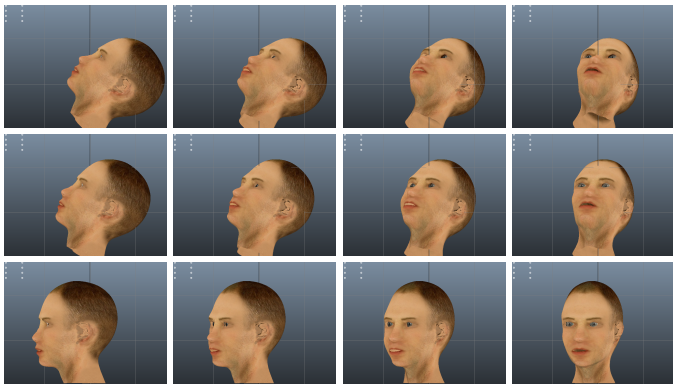


図 8 補正值 ($\Delta\hat{x}_{xi}(\mathbf{p})$, $\Delta\hat{x}_{yi}(\mathbf{p})$) の作成に使用した CG 顔
Fig. 8 CG face to compute the value of ($\Delta\hat{x}_{xi}(\mathbf{p})$, $\Delta\hat{x}_{yi}(\mathbf{p})$).

示される遮蔽輪郭である。画像上では、遮蔽輪郭は図右上の青線であり、本当の輪郭の赤線とは距離がある。そこで画像から検出した遮蔽輪郭から本当の輪郭に補正する手法を提案する。

4.1 遮蔽輪郭の補正したパラメータ最適化

4.1.1 輪郭補正テーブル

遮蔽輪郭の補正は、検出した遮蔽輪郭の座標に本当の輪郭までの画像上の距離を加算することで行う。この距離は予め 3D シェイプを得ている CG の顔画像を使って角度毎の遮蔽輪郭から本当の輪郭までの距離を計測して求めた。距離の計測にあたり図 8 のように CG 顔をヨーとピッチ方向に回転させて角度毎に画像を作る。-90 度から 90 度までの範囲で 5 度刻みで測定し、ピッチ角 α とヨー角 β についてのテーブルを作成する。アライメントではこのテーブルを参照して使用する。本稿ではこのテーブルを輪郭補正テーブルと呼ぶ。

4.1.2 遮蔽輪郭の補正

本当の輪郭の位置 $\{\hat{x}'_i(\mathbf{p})\}_{i=1}^k$ は

$$\hat{x}'_i(\mathbf{p}) = \hat{x}_i + \mathbf{D}_i(\mathbf{p}), \quad (5)$$

として、遮蔽輪郭の位置に補正值

$$\mathbf{D}_i(\mathbf{p}) = s \begin{bmatrix} \cos(\gamma) & -\sin(\gamma) \\ \sin(\gamma) & \cos(\gamma) \end{bmatrix} \begin{bmatrix} \Delta\hat{x}_{xi}(\mathbf{p}) \\ \Delta\hat{x}_{yi}(\mathbf{p}) \end{bmatrix}, \quad (6)$$

を加算して表される。輪郭補正テーブルは 5° 刻みの離散化された値なので、 $\Delta\hat{x}_{xi}(\mathbf{p})$, $\Delta\hat{x}_{yi}(\mathbf{p})$ は輪郭補正テーブルをバイリニア補間した形で表現する。そしてスケールと回転を考慮して、遮蔽輪郭から本当の輪郭までの補正值とする。式 (5), 式 (6) により式 (4) は

$$Q(\mathbf{p}) = \sum_{i=1}^M \|\hat{x}_i + \mathbf{D}_i(\mathbf{p}) - \mathbf{x}_i(\mathbf{p})\|^2 + \|\mathbf{q}\|_{\Lambda^{-1}}^2, \quad (7)$$

と書き換えられる。パラメータ \mathbf{p} は式 (5), 式 (6) の制約の下に遮蔽輪郭を補正しながら最適化される。

4.1.3 パラメータ最適化

式 (7) は PDM 式 (3), および特徴点 式 (5) の \mathbf{p} のテラー展開の一次までの項

$$\mathbf{x}_i(\mathbf{p}) \approx \mathbf{x}_i^c + \mathbf{J}_i \Delta\mathbf{p}, \quad (8)$$

$$\hat{x}'_i(\mathbf{p}) \approx \hat{x}'_i^c + \mathbf{G}_i \Delta\mathbf{p}, \quad (9)$$

を用いて表す。適当な初期値 \mathbf{p} を設定し $\mathbf{p} \leftarrow \mathbf{p} + \Delta\mathbf{p}$ のように繰り返し更新しながら最小化する。そして収束する \mathbf{p} を解として求める。ここで $\mathbf{J} = [\mathbf{J}_1; \dots; \mathbf{J}_n]$, $\mathbf{G} = [\mathbf{G}_1; \dots; \mathbf{G}_n]$ はヤコビアン, $\mathbf{x}^c = [\mathbf{x}_1^c; \dots; \mathbf{x}_n^c]$, $\hat{x}^c = [\hat{x}'_1^c; \dots; \hat{x}'_n^c]$ は繰り返し処理の現在の顔モデルと特徴点を表す。繰り返しの度に輪郭補正テーブルを新たな \mathbf{p} で参照して $\mathbf{D}_i(\mathbf{p})$ を更新する。画像から正確に特徴点を検出できるなら、初期値として使う \mathbf{p} が期待値と大きく異なっている場合でも正確に期待値に収束できる。だが実際に特徴検出器で得られた位置には誤差が含まれているので、適切な初期値が必要である。提案手法では $\pm 30^\circ$ 程度以内であれば期待値に収束する。繰り返し数は 10 回程度であった。

アライメントする入力された顔の精密な 3D シェイプが得られるならば正確な $\mathbf{D}_i(\mathbf{p})$ を求められる。だが不特定の顔の精密な 3D シェイプを予め得ることは難しいため、CG 顔を使用した。この顔はアライメント対象本人のものではないため出来上がった補正值には誤差がある。しかし PDM で作る顔モデルは複数人のシェイプから作成しているため計測対象本人以外のシェイプを使ったとしてもその誤差は吸収できると考える。

4.2 遮蔽輪郭の補正精度の評価

CG 顔で計測した補正值により、本当の輪郭が正確に計算できるか評価を行った。図 9 に評価結果を示す。評価は顔画像データベース CMU pose, illumination and expression database (Multi-PIE) [13] の 4 人に対して行った。図の左列はヨー角度 75 度と右列は 60 度の顔の向きである。画像から検出された特徴点から遮蔽輪郭を補正した輪郭の位置を赤線で、正解値を緑線で示す。提案手法は計測対象本人の輪郭補正データを使わずに CG 顔の本当の輪郭を計算したが、どの結果も正解値と比較して大きな誤差は無いことがわかる。

4.3 提案手法の有効性検討

4.3.1 ポーズ推定精度の評価

提案手法は、モデルとは異なる不適切な遮蔽輪郭を本当の輪郭に補正するので、正確にポーズを推定できる。本当の輪郭が正しく計算されフェイスアライメントできるか確認するために、従来の遮蔽輪郭を補正しないモデル (PDM1), 提案手法 (Our model) と遮蔽輪郭を使わず画像上から観測できる特徴点のみを使うモデル (PDM2) を比較した。遮蔽輪郭の補正值を計測した CG モデル図 8 をヨー角度 10° か

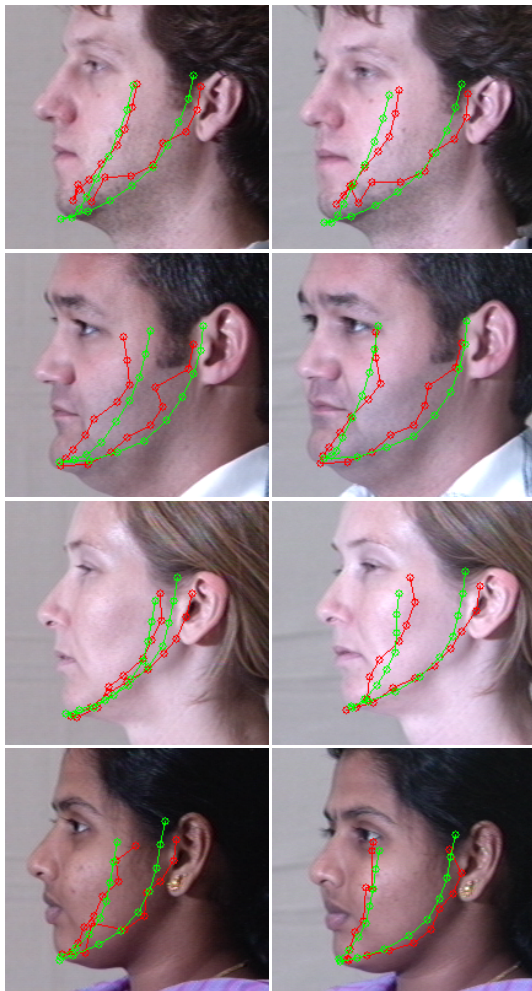


図 9 個人毎の遮蔽輪郭の補正精度 補正後の輪郭(赤), 正解値(緑)
左:ヨー角度 75° 右:ヨー角度 60°

Fig. 9 Correctness of the occluding contour compensation in individual. Red lines are compensated contours. Green lines are ground truths. Left column: 75° of a head rotation, Right column: 60° of a head rotation.

ら 90° まで 10° 毎に回転して, 画像上の特徴点座標を角度毎に計測し, それぞれのモデルを使ってヨー角度を計算した. 図 10 に計測結果を示す. PDM1 は遮蔽輪郭が悪影響し角度が大きくなるにつれ誤差が大きくなる. Our model と PDM2 は大きな誤差が発生しなかった. 遮蔽輪郭を推定した場合と PDM1 を比較すると, 提案手法の誤差が小さいので, 提案手法はポーズ推定精度をあげられることがわかる. 次に, 提案手法は PDM2 とほぼ同じ結果である. 画像に見える特徴のみを使った PDM2 はほぼ期待値であることから, 提案手法は本当の輪郭を正確に推定できることがわかる.

4.3.2 ポーズ推定の安定性評価

提案手法はモデルを画像上に投影した位置と画像から検出された特徴点の位置の差の合計が小さくなるようにパラメータ最適化処理していく際に, 顔の輪郭から推定される顔の向きを制約条件として加えることにより, 誤ったパラ

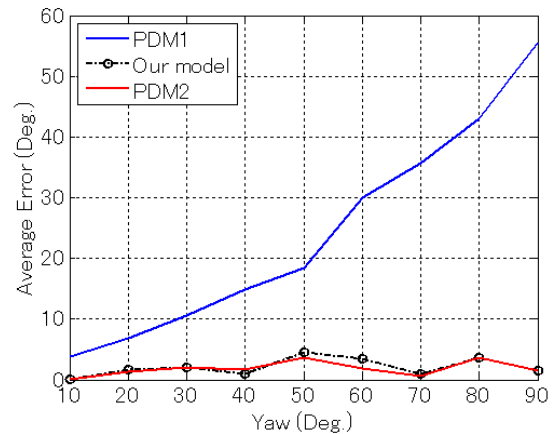


図 10 ヨー角度の推定誤差 PDM1:従来手法によるモデル, Our model:提案手法, PDM2:画像から検出される特徴点のみを使う場合の従来モデル, Noise 10:±10 画素のランダムノイズ付加, Noise 20:±20 画素のランダムノイズ付加

Fig. 10 Error of the yaw angle of a head rotation. PDM1:PDM, Our model:Proposed model, PDM2:PDM constructed with only visible features. Noise 10:Added ±10 pixels random noise, Noise 20:Added ±20 pixels random noise.

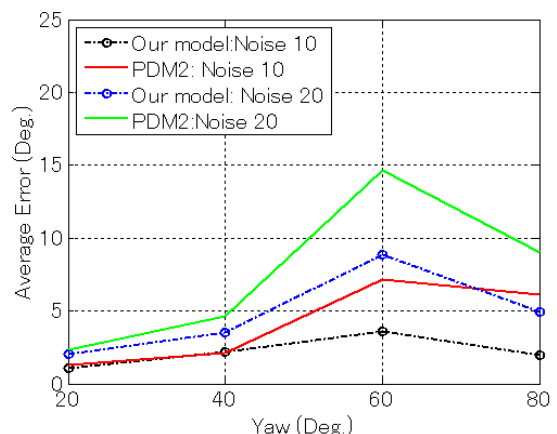


図 11 ポーズ推定の安定性評価

Fig. 11 Stability evaluation of the pose estimation.

メータが推定されるのを抑制する. しかし提案手法が求めた本当の輪郭は, p の関数であり, 計算によって自由に補正值を作り出されるので, 本当にこの制約が機能しているか確認する必要がある. そこで遮蔽輪郭を使わず画像上から観測できる特徴点のみを使ってフェイスアライメントを行う PDM2 と提案手法を比較した. 第 4.3.1 章の評価で使った座標にランダムノイズを加えたとき, Our model と PDM2 のヨー角度推定精度の評価を行った. ±10 画素と ±20 画素の 2 種類のランダムノイズで評価した. 図 11 に評価結果を示す. 一つの角度に対し 50 回試行を行い縦軸に正解値からの平均誤差を示す. 顔のサイズはおよそ縦 300 画素横 300 画素である. 提案手法は PDM2 に比べ平均誤差をより抑えることができた. この結果から, 提案手法は遮

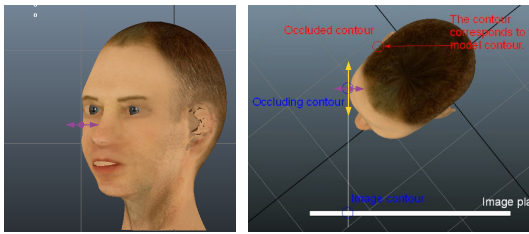


図 12 遮蔽輪郭検出における誤差 奥行き方向の誤差 (黄矢印), X 軸方向の誤差 (紫矢印).

Fig. 12 Error of the occluding contour detection. Error of Z-axis direction (Yellow arrow), Error of X-axis direction (Purple arrow).

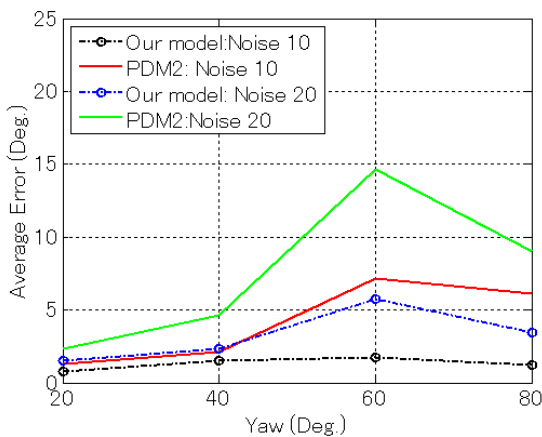


図 13 Y 軸方向のノイズに対するポーズ推定の安定性評価

Fig. 13 Stability evaluation of the pose estimation with noise of Y dimension.

遮蔽輪郭から推定されるポーズをパラメータ最適化の制約条件とすることで、パラメータが誤った値に推定されることを抑えて、推定の精度を上げることがわかる。

4.3.3 輪郭を使うことによる利点

輪郭が位置する顔の頬は滑らかで円筒状の曲率を持つ。図 12 の右図のように奥行き方向の検出誤差 (黄色の矢印) と X 軸方向の検出誤差 (紫矢印) を考えたとき、奥行き方向の誤差は画像に表れにくい。また、顔と背景の境界線上に位置しているため、背景のテクスチャにもよるがテクスチャの違いを利用して判別しやすい場合が多いので、X 軸方向の誤差を抑えて検出しやすい利点がある。第 4.3.2 章の評価条件に対し遮蔽輪郭へのノイズを Y 軸方向のみに加えた場合の評価を行った。図 13 に評価結果を示す。提案手法はノイズを ± 20 画素加えた場合も、PDM1 とほぼ同じかそれ以上の推定精度をなした。この結果から、実应用到提案手法を適用した場合、遮蔽輪郭は X 軸方向に少ない誤差で検出できる場合が多いと考えられるので、その場合よりポーズ推定の誤差を抑えることができることがわかる。

これらの実験結果から提案手法はフェイスアライメントのためのポーズ推定に有効に機能すると考えられる。

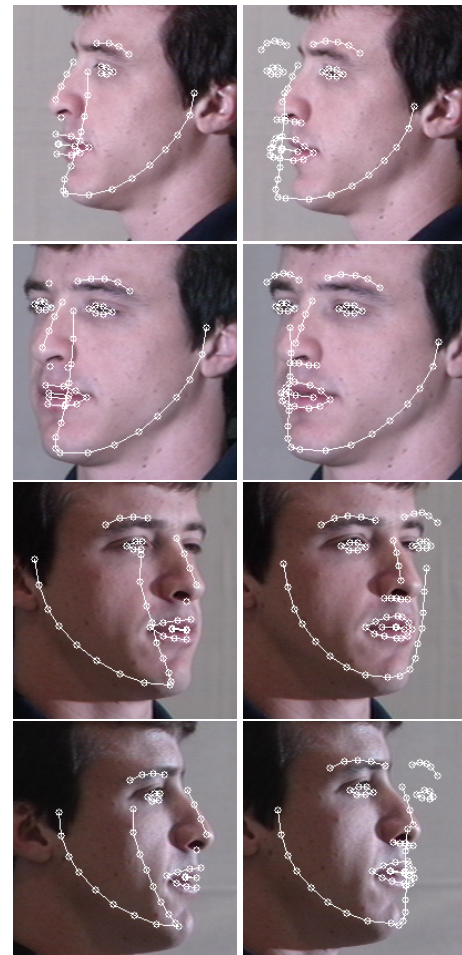


図 14 Multi-PIE を使った提案手法 (左) と CLM (右) によるアライメント結果の比較

Fig. 14 Left image: the results using CLM, Right image: the results using our method.

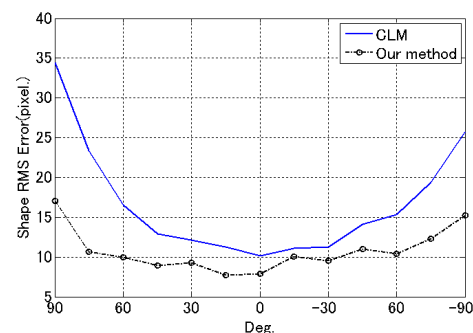


図 15 提案手法 (黒の破線), および CLM (青線) によるアライメント結果と正解値との RMS エラー (出典 [12]).

Fig. 15 RMS error between the ground truth and the resulting alignments by our method and those by the CLM.

5. 性能評価

5.1 静止画を使ったアライメントの精度評価

アライメント精度を Multi-PIE を使って、提案手法と PDM2 を比較した。両手法ともに第 2 章で説明した CLM

の枠組みでフェイスアライメントを行った。Multi-PIE は異なる角度から同時に撮影された顔画像で成るデータベースである。そのため、ヨー方向の回転に対するアライメント精度の評価に適している。本評価では、15人の計測対象の15°刻みで撮影された195枚の画像を使用した。目、口、鼻、眉、そして画像輪郭と遮蔽輪郭を正確に縁取って白線が表示されるほど精度が高い。図14に提案手法とPDM2を用いたCLM(以下CLMと呼ぶ)でのフェイスアライメント結果(以下、すべての評価で見ると判定された特徴と遮蔽輪郭を白で描画する)を示す。左列が提案手法で右列がCLMである。角度が大きくなるに従いPDM2はアライメントの精度が悪化したが、提案手法の位置決め精度は大きく悪化しなかった。図15にそれぞれ正解値とのroot-mean-square(RMS)を示す。提案手法のRMSは角度が大きくなってもCLMよりも低く抑えられおり、提案手法はCLMに比べ良い精度を示した。その差は角度が大きくなったとき大きくなり、90°の時顔サイズ約幅170pixel高さ200pixelに対し10pixel以上改善した。

5.2 大きな動きのある動画によるトラッキング精度評価

大きなポーズの変化、移動、セルフオクルージョンを含む動画を使って評価を行った。提案手法とXiangxinらの手法[14]の結果を目視によって比較した。Xiangxinらの手法はモデルをTree structured part modelを使って表現する。遮蔽輪郭は補正せずそのまま使用する。この手法はCVPR2012で提案された手法で、現在のフェイスアライメント分野においてstate-of-the-artの一つと考えられることから提案手法と比較した。しかしながら、提案手法は前フレームで計算したパラメータを初期位置として利用するが、Xiangxinらの手法は前フレームの情報を使用しないため、本評価ではXiangxinらの手法に若干不利になる要素がある。評価結果を図16に示す。図の左列が提案手法であり赤丸がアライメント結果で青丸がセルフオクルージョンにより画像上から見えなくなったと判定された特徴である。右列がXiangxinらの手法であり、赤点のアライメント結果である。図から提案手法は大きな顔の向きに対しても顔の輪郭位置を推定しアライメントできることが分かる。この動画は約600フレームあり大きなポーズの変化、移動、セルフオクルージョンの発生があるが、提案手法は精度良くアライメントすることができた。

6. 結論

遮蔽輪郭を補正して本当の輪郭を計算し、それを用いてフェイスアライメントすることで、大きな顔の向きに対してアライメントできる手法を提案した。本論文ではCLMの枠組みで提案手法で作成したモデルを評価し、アライメント精度を確認した。遮蔽輪郭を補正した場合と補正しない場合、画像上に観測できる特徴点のみを使う場合で顔の

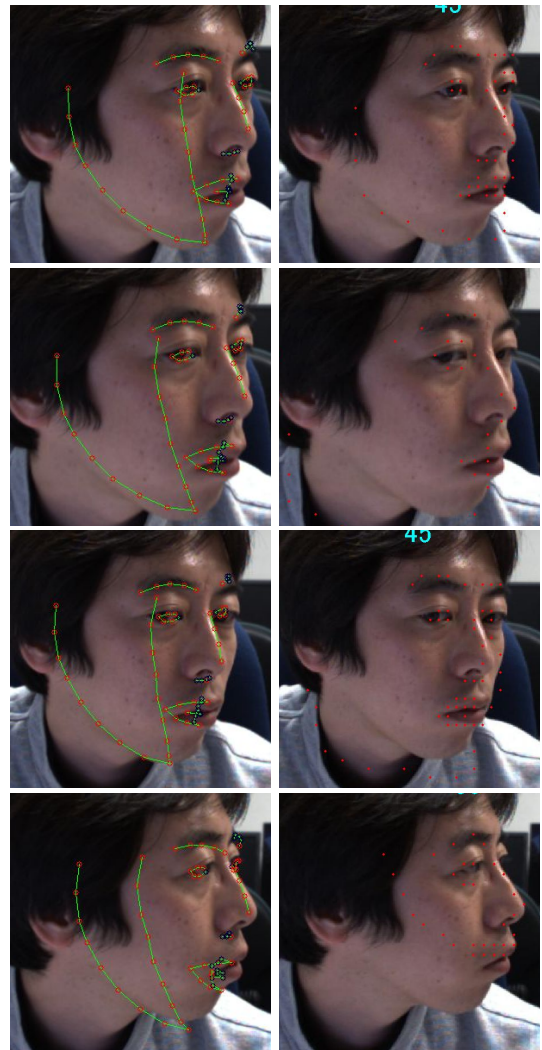


図16 提案手法とXiangxing Zhuらの手法の評価。左が提案手法、右がXiangxing Zhuらの手法。

Fig. 16 Left image: the results using our method, Right image: the results using XianxingZhu method.

ポーズ推定精度の評価を行い、遮蔽輪郭の補正がポーズ推定精度を向上させることを確認した。そして、CG顔から計測した遮蔽輪郭の補正値を使用しても様々な人に対しても補正を比較的正確に行えることを確認し、輪郭の補正によりパラメータ最適化において誤ったパラメータが推定されないように制約を加えることができることを確認した。輪郭補正テーブルによって画像輪郭から検出した特徴点の位置を補正し遮蔽輪郭に変換することでモデル輪郭に正しく対応した特徴点でパラメータ最適化に使えるようにし、多様な角度に対してフェイスアライメントを行うことができた。

評価結果から、提案手法は角度が大きく変化したときでも、従来手法に比べより正確に安定してアライメントできることを確認した。提案手法を用いることで、動画中の激しく動く顔追跡できることを確認した。

提案手法は輪郭を計算するので、実応用を考えたとき、ア

バターなどのアニメーションを制御するには顔の輪郭が必要であり、応用の範囲を広げることができる。

本稿では、画像上から特徴点を検出する手法については考察していない。今後は顔の角度の変化やアピランスの変化に対応した特徴点検出手法を検討する。また、View-based AAMのような複数のモデルを用意し角度によって切り替えてアライメントを行う手法と比較し、提案手法の有効性を評価したい。

7. 謝辞

本研究はカーネギーメロン大学ロボティクス研究所の金出武雄教授、山崎彬人博士、Junsik Kim 博士に熱心なご指導を頂きました。ここに感謝の意を表します。

参考文献

- [1] T. Cootes, C. Taylor, D. Cooper and J. Graham: "Training models of shape from sets of examples", Proc. British Machine Vision Conference, Vol. 9 Citeseer, p. 18 (1992).
- [2] S. Baker and I. Matthews: "Equivalence and efficiency of image alignment algorithms", Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, Vol. 1, pp. 1090–1097 (2001).
- [3] T. Cootes, G. Edwards and C. Taylor: "Active appearance models", IEEE Transactions on Pattern Analysis and Machine Intelligence, **23**, 6, pp. 681–685 (2001).
- [4] T. Cootes, G. Wheeler, K. Walker and C. Taylor: "View-based active appearance models", Image and Vision Computing, **20**, 9-10, pp. 657–664 (2002).
- [5] T. Cootes and C. Taylor: "Active shape models—smart snakes", Proc. British Machine Vision Conference, pp. 266–275 (1992).
- [6] D. Cristinacce and T. Cootes: "Feature detection and tracking with constrained local models", Proc. British Machine Vision Conference, Vol. 3, pp. 929–938 (2006).
- [7] J. Saragih, S. Lucey and J. Cohn: "Deformable model fitting by regularized landmark mean-shift", Int. Journal of Computer Vision, **91**, 2, pp. 200–215 (2011).
- [8] X. Zhang and Y. Gao: "Face recognition across pose: A review", Pattern Recognition, **42**, 11, pp. 2876–2896 (2009).
- [9] J. Jang and T. Kanade: "Robust 3d head tracking by online feature registration", 8th IEEE Int. Conf. on Automatic Face and Gesture Recognition (2008).
- [10] M. La Cascia, S. Sclaroff and V. Athitsos: "Fast, reliable head tracking under varying illumination: An approach based on registration of texture-mapped 3d models", IEEE Transactions on Pattern Analysis and Machine Intelligence, **22**, 4, pp. 322–336 (2000).
- [11] J. Sung, T. Kanade and D. Kim: "Pose robust face tracking by combining active appearance models and cylinder head models", Int. Journal of Computer Vision, **80**, 2, pp. 260–274 (2008).
- [12] 高野, 出口: "多様な顔の向きに対して画像上の顔部位を位置決めする手法", 第 86 回パターン計測部会研究会資料 SICE, pp. 13–18 (2012).
- [13] R. Gross, I. Matthews, J. Cohn, T. Kanade and S. Baker: "Multi-pie", Image and Vision Computing, **28**, 5, pp. 807–813 (2010).
- [14] Z. Xiangxin and R. Deva: "Face detection, pose estimation, and landmark localization in the wild", in pro-

ceedings of Computer Vision and Pattern Recognition (2012).