

パンチルトカメラにより自動追跡撮影された 顔画像列からの3次元顔形状復元

鈴木 誠司[†] 齋藤 英雄[†] 持丸 正明[‡]

[†] 慶應義塾大学理工学部情報工学科
[‡] 産業技術総合研究所デジタルヒューマン研究センター

本稿では3次元顔形状データベースを用いて、監視カメラ映像から顔の姿勢および3次元形状を復元するシステムを提案する。パンチルトカメラを監視カメラとして用い、自動的に顔を追跡撮影することによって、低解像度のカメラでも、復元に必要な顔画像を大量に撮影可能とする。そして、得られた画像列から撮影された人物の画像を複数枚選択し、選択された複数の画像から対象人物の顔画像の3次元形状を復元する。まず、データベースの平均顔を初期形状として、各画像における顔姿勢を推定する。その後、この姿勢推定された顔画像群にフィットするように3次元顔形状をデータベースの主成分パラメータを変化させることにより変形させて行く。これらの姿勢推定および形状推定の操作を収束するまで繰り返す。提案システムを用いた実験を行った結果、入力画像の輪郭や骨格に沿った3次元顔形状を復元することができた。

Reconstruction of 3D Facial Shape from Pan-Tilt Camera Image Sequence with Automatic Face Tracking

Seiji Suzuki[†], Hideo Saito[†], and Masaaki Mochimaru[‡]

[†] Department of Information and Computer Science, Keio University
[‡] Digital Human Research Center, AIST

We propose a system which reconstruct both of facial pose and 3D shape from a surveillance camera image sequence by using a 3D facial shape database. A pan-tilt camera is used as the surveillance camera, and it tracks a face automatically. It enables to record a lot of facial images even if the surveillance camera is low resolution. A user chooses several images from the image sequence for facial shape reconstruction. The facial pose in each image is estimated by using the facial shape. (The initial shape is average of the database.) Then the facial shape is estimated to be fitted into the pose estimated images by changing the PCA parameter. These pose and shape estimation are repeated. This proposed system can reconstruct the facial shape which is proper to the input image sequence.

1 はじめに

監視カメラの映像は事件発生時に犯人逮捕の有力な手がかりとなる。監視カメラで捉えられた人物の映像は様々な視点から人物を撮影したものと異なるため、非同期で撮影されているものの、同一人物に対して撮影された多視点画像とみなすことができる。これらの非同期多視点画像をもとに、撮影された人物の3次元顔形状を復元することができれば、様々な応用が期待できる。

非同期多視点画像からの3次元形状復元の一例と

して、ビデオ等の連続画像から3次元形状を復元する Structure from Motion という手法がある [1]。これは、1台のカメラで撮影された同一人物の画像列から顔の3次元形状を復元しようとするものである。しかし、この手法は復元に際し、各フレーム間の対応点を求めるために Optical Flow を用いている。顔は頬などテクスチャの均一な部分が大部分を占めており、Optical Flow による対応点算出が正確に行われず、復元精度が悪くという問題があった。

そこで、単に同一人物に対して撮影された画像を利用するのではなく、あらかじめ収集しておいた人間の頭

部 3次元形状情報を持つデータベースを主成分分析し、低次元の固有空間から形状を推定する Active Shape Model という手法が提案されている。Active Shape Model の例として、顔の 3次元形状・テクスチャ情報を主成分分析し推定する手法 [2-4] がある。これらの手法はテクスチャを貼り付けた顔の 3次元モデルを高々 1枚の入力画像 (単視点画像) にフィッティングさせることによって復元を行っている。すなわち別視点から捉えた画像を用いておらず、幾何学的情報を無視してしまっている。一方、多視点画像に対して Active Shape Model を適用した手法 [5, 6] が提案されている。しかし、この手法は顔のカメラに対する姿勢が既知であるという条件を用いており、監視カメラ映像のような顔の姿勢が未知である場合に用いることができなかつた。

これに対し、顔の姿勢、3次元形状が共に未知の状態からその両方を復元する手法 [7] が提案されている。しかし、この手法で想定している監視カメラシステムにおいては、高解像度で顔画像を得ようとカメラを設置すると、すぐに対象がフレームアウトしてしまうという問題点があった。反対に、様々な角度で顔画像を大量に得ようとカメラを遠方に設置した場合、顔画像の解像度が下がり、復元に用いることができなかつた。また、監視カメラの性質上、高解像度のカメラを遠方に設置するという方法では、長時間の録画が困難であった。

そこで本稿では、パンチルトカメラを用いて人物を追跡撮影することで前述の問題を解決可能なシステムを構築し、このシステムにより撮影された多視点画像を入力とし、前述した竹内らの手法 [7] と同様に解剖学的データベースを主成分分析して顔の 3次元形状を復元する手法を提案する。なお、本手法は顔形状復元に用いられる輪郭評価関数に外輪郭という概念を取り入れることにより復元精度の向上も図る。さらに最適化の行き過ぎによる顔形状の過変形を防ぐため、顔尤度評価関数を導入することで安定して復元を行うことができるよう改善する。

2 提案システムの概要

本稿で提案するシステムは監視カメラによるサーベイランスシステムである。本システムは通常のサーベイランスシステムと同様に、平常時は監視カメラ映像を録画保存しておく。事件発生時、この映像から顔の 3次元形状を復元し、不審人物の顔の自由視点画像を生成することにより捜査を支援する。ただし、顔の 3次元形状を復元するためには多くの高解像度顔画像が必要となるため、本システムではパンチルトカメラを

監視カメラとして用いて自動追尾することにより、これを解決する。

まず、監視カメラから得られた画像に人物の顔が検出された場合には、その人物が常に画像中心に写るようにパンチルトカメラを制御する。このとき同時に人物の顔を輪郭に沿って自動的に切り出しておく。不審人物が現れた場合には、その人物が撮影された映像から復元に向く画像を数枚選択する。選択された画像は、なるべく顔の向きが異なる画像同士であること、鼻尖点・目尻点・口角点が撮影されていること、を満たしていなければならない。次に、選択された画像それぞれにおいて、切り出された顔の輪郭を手動で修正する。この操作は首などが輪郭として切り出されてしまった場合に必要となる。また、鼻尖点・目尻点・口角点の画像上の座標および顔の外輪郭を手動で入力する。顔の外輪郭とは撮影された顔と背景との境界のことを指す。一方、顔と首・髪の毛・耳などの境界のことを内輪郭と呼ぶことにする。このようにして得られた入力画像群と、解剖学的データベースをもとに顔の 3次元形状復元を行う。

3 顔追跡撮影手法

まず、パンチルトカメラからキャプチャされた画像に対して、Haar-Like 特徴量に基づく顔検出を行う。顔検出に用いる学習データは正面向きのものを用いる。顔が検出されるとその矩形 $R_{\text{face}} = [x, y, w, h]$ が得られる。このとき、 x, y は矩形の左上の座標、 w, h はそれぞれ矩形の幅と高さを表す。なお、画像座標系は左上を原点とし、右方向に x 軸、下方向に y 軸が定義されているものとする。

これに対し、

$$R_{\text{nose}} = \left[x + \frac{w}{4}, y + \frac{h}{2}, \frac{w}{2}, \frac{h}{5} \right] \quad (1)$$

という矩形を鼻領域として定義する。

この領域の明度値の平均値を対象人物の肌色 I_{skin} として登録する。Haar-Like 特徴量は、検出にあたり明度値の濃淡のみを情報として用いている。そのため、人物の顔検出に色の情報を用いていない。以下の操作において色情報を用いるが、このようにして肌色を取得することで対象人物の肌の色に対して頑強性を保障する。

次に、矩形

$$R_{\text{search}} = \left[x - \frac{w}{4}, y - \frac{h}{4}, \frac{3}{2}w, \frac{3}{2}h \right] \quad (2)$$

を探索範囲として設ける。

対象人物の肌色 I_{skin} と、この探索範囲内の各明度値 $I_{i,j}$ との正規化相互相関 (NCC) を取ることで、 $I_{i,j}$ が肌色かどうかを判定する。本手法では、

$$\frac{I_{i,j} \cdot I_{skin}}{\|I_{i,j}\| \|I_{skin}\|} > 0.995 \quad (3)$$

なる $I_{i,j}$ を肌色と判定した。

NCC は色を RGB 空間のベクトル $C = [r, g, b]$ と見立てて、そのベクトル同士の $\cos \theta$ を相関値として用いている。つまりベクトルの向きだけを用いているため、ベクトルの大きさ (すなわち明るさ) に依らず、鼻などによってできた影の部分も肌色と判定することができるという利点がある。

さて、続いて肌色と判定された画素を白画素、それ以外を黒画素とした 2 値画像を用意し、その輪郭抽出を行う。輪郭として得られたもののうち、内側が白画素、外側が黒画素で、かつ面積が最大のものを選出する。これが顔の輪郭となる。

このようにして得られた顔輪郭を用いて、パーティクルフィルタによる予測追跡を行う。各パーティクルの尤度は輪郭からのユークリッド距離を d として、

$$L(d) = \frac{1}{\sqrt{2\pi}\sigma} \left(-\frac{d^2}{2\sigma^2} \right) \quad (4)$$

という正規分布の確率密度関数として定義した。なお、本手法では $\sigma = 10$ としている。

次フレームではパーティクルフィルタによる予測値が推定される。本手法では再分布されたパーティクルを囲う矩形を新たな探索範囲 R_{search} とし、この領域内のみで、Haar-Like 特徴量による顔検出を行う。Haar-Like 特徴量に基づく顔検出アルゴリズムは検出に時間がかかるため、このように探索範囲を限定することは、リアルタイム追跡を行う上で有用である。

この結果、顔が検出された場合は上記の手法を繰り返し適用する。顔が検出されなかった場合は、探索矩形内に顔がないのではなく、顔が斜めや横といった方向を向いている可能性が高い。そこで、再び肌色領域を検出し、輪郭を得る。

しかし、この場合は顔が横を向いていることから、眉や目、唇といった肌色でない顔のパーツが輪郭の外側と判定されてしまうことがある。モルフロジー処理でこのノイズを取り除くことも可能であるが、これは輪郭を鈍らせてしまう効果がある。本手法ではここで得られた輪郭を顔の 3 次元復元で用いるため、この手法は用いない。得られた輪郭の内、眉や目、唇といった部分は凹状欠損となっていることに注目して、この“くびれ”を取り除くことを考える。

まず、Sklansky のアルゴリズム [8] を用いて輪郭の凸包を見つける。この凸包を再帰的に縮退することにより凹状欠損のノイズ除去を行う。図 1 に輪郭の凸包の再帰的縮退の概念図を載せる。

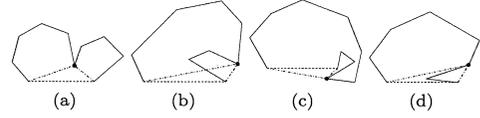


図 1: 凸包の再帰的縮退

黒の実線で示された線が輪郭の一部を表す線である。これと破線で囲われる領域が凹状欠損である。このとき、以下のアルゴリズムで再帰的に凸包を縮退して、ノイズとなる凹状欠損を埋める。

1. 凹状欠損が直線の場合、縮退を終了する
2. 凹状欠損の面積が一定以下の場合、縮退を終了する
3. 凹状欠損間の距離 (破線の長さ) が一定以下の場合、凹状欠損を構成する輪郭上の点すべてを除去することで欠損を埋めて縮退を終了する
4. 凹状欠損が三角形の場合、縮退を終了する
5. 上記の操作で縮退が終了しなかった場合、凹状欠損を 2 分割する
 - i. 凹状欠損を表す輪郭上の点を順にチェックし、注目点とする
 - ii. 注目点で分割する場合に分割線 (一点鎖線) が 1 つでも輪郭線と交差する場合 (図 1(b)), 無視して次の注目点に移る
 - iii. 注目点が凹状欠損の逆側にある場合 (図 1(c)), 無視して次の注目点に移る
 - iv. 注目点で分割する場合に破線と分割線で囲われる三角形内に 1 つでも輪郭上の点がある場合 (図 1(d)), 無視して次の注目点に移る
 - v. 以上の操作で無視されなかった注目点のうち、長さが最小の分割線を持つ点で分割する (図 1(a))

以上の操作を経て、眉や目、唇といった肌色でない顔のパーツを輪郭内に納める。このようにしてノイズ除去された輪郭からの距離を用いて再びパーティクルフィルタを用いた追跡を行っていく。

4 3次元顔形状復元手法

本手法では、暫定的な 3 次元形状を投影することにより画像を作成し、入力画像と比較する。入力画像と

投影画像との輪郭評価、特徴点評価、テクスチャ評価、顔尤度評価の値(エネルギー評価値)が最も低くなるような、入力画像の顔の姿勢と3次元形状を推定する。

まず、入力画像に対して暫定的な3次元顔形状を用いることにより、顔の姿勢推定を行い、大まかな姿勢を推定する。次に顔の姿勢が推定された入力画像を用いて3次元形状推定を行う。この顔の姿勢推定、3次元形状推定の処理を繰り返すことで、誤差を減少させる。

姿勢推定、3次元形状推定の処理で計算される評価値が収束するまでこれを行い、収束したときの顔の姿勢、3次元形状を復元結果とする。

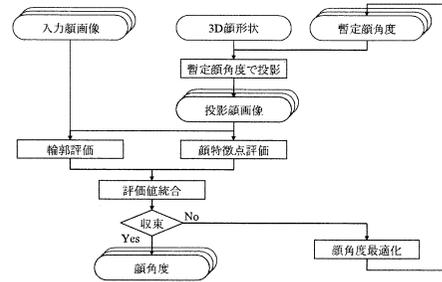


図 2: 顔角度推定の流れ

4.1 3次元顔形状パラメータ数

本手法は Levenberg-Marquardt 法により、顔の3次元データを最適化する手法をとる。この最適化においてパラメータ数(最適化目的関数の引数の数)はできる限り少ないことが望ましい。パラメータ数が多いと目的関数が複雑化し、局所解に陥りやすくなる。一方で、パラメータ数が少なくなれば、顔形状の表現度が減少してしまう。すなわち、最適化と顔形状の表現はトレードオフの関係にある。

本手法で用いる顔形状データベースの頂点数は260点である。この260点それぞれについて、 (x, y, z) の3つの次元があることから、顔を表現するために必要なパラメータ数は、 $260 \times 3 = 780$ である。このパラメータ数を減少させるため、主成分分析を用いる。データベースを主成分分析し、累積寄与率がほぼ累積寄与率90%を超え、頭打ちとなる上位20成分で顔形状を表現することとした。

4.2 顔姿勢推定

顔姿勢推定においては、推定パラメータ数を減らすため、カメラの画角は60度であると仮定し、カメラの顔に対する並進ベクトルは顔輪郭の面積と画像上の平行移動により近似する。すなわち、カメラに対する回転角度のみを推定パラメータとして用いる。

図2に顔角度推定の流れを示す。

使用する3次元顔形状は初回のみデータベースの平均3次元形状を使用し、2回目以降は顔の3次元形状推定によって得られた形状を使用する。初期暫定顔角度は初回のみ正面向きであり、2回目以降は前回の推定で得られた角度を用いる。

まず、顔の3次元形状を暫定顔角度で投影する。これにより、シルエット画像、顔特徴点(鼻尖点・目尻点・口角点)を投影した2次元座標が得られる。入力

画像とこの投影画像を2次元画像上でシルエットの面積が一致するように正規化し、鼻尖点が一致するように位置合わせを行う。これにより、カメラの顔に対する3次元的な並進ベクトルを近似する。

位置合わせされた入力画像と投影画像において、輪郭評価、顔特徴点評価を行い、評価値が減少するように Levenberg-Marquardt 法により顔角度を最適化する。

4.3 3次元顔形状推定

顔姿勢推定によって得られた顔の角度を用いて顔の3次元顔形状推定を行う。

図3に3次元顔形状推定の流れを示す。

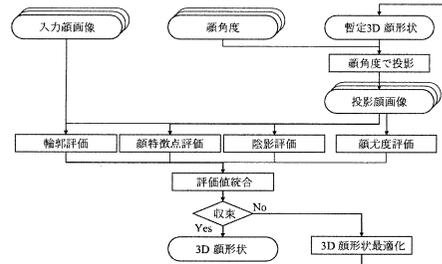


図 3: 3次元顔形状推定の流れ

使用する顔角度は顔姿勢推定によって得られた推定値である。初期暫定3次元顔形状は初回のみデータベースの平均顔を、2回目以降は前回の推定で得られた形状を用いる。

まず、入力画像の内、顔角度が最も正面向きに近いと推定されたものをテクスチャ画像として用いる。テクスチャ画像として用いられる入力画像の向きに一度暫定的な3次元顔形状を投影する。次にテクスチャ画像と投影画像を姿勢推定のとときと同様の方法で位置合わせする。位置合わせを行ったら、3次元顔形状を構成する各3次元点を投影した2次元座標と、テクスチャ

画像の2次元座標との対応関係をそれぞれ求める。ただし、投影された点が自己オクルージョンを起こしている陰点の場合、対応させない。このようにしないと、陰面にまでテクスチャが貼られてしまう。

次に暫定的な3次元顔形状をそれぞれ入力画像の推定された顔角度で投影する。すなわち、入力画像群と同じ向きの画像を作成する。このようにして投影することにより、シルエット画像、顔特徴点(鼻尖点・目尻点・口角点)を投影した2次元座標、テクスチャマッピング画像、テクスチャマッピングを行ったパッチかどうかを表すマスク画像が得られる。さらに入力画像と得られた投影画像とを、姿勢推定のととき同様の方法で位置合わせを行う。

位置合わせされた入力画像と投影画像において、輪郭評価、顔特徴点評価、テクスチャ評価を行う。また、暫定的な3次元顔形状の顔尤度評価を行う。これらの評価値が減少するように Levenberg-Marquardt 法により3次元顔形状を最適化する。

4.4 評価関数

顔の姿勢推定、3次元形状推定において用いられている評価関数について述べる。姿勢推定において用いられる評価関数は、輪郭評価関数と特徴点評価関数である。また、3次元形状推定においては、これらに加えてテクスチャ評価関数と顔尤度評価関数を用いる。なお、それぞれ経験的に求められた係数で重み付けを行い評価値を統合する。

4.4.1 輪郭評価関数

輪郭評価では Boundary-weighted XOR [9](以下、B-XOR) による評価と、外輪郭による評価を行う。

図4に B-XOR による評価の概念図を示す。

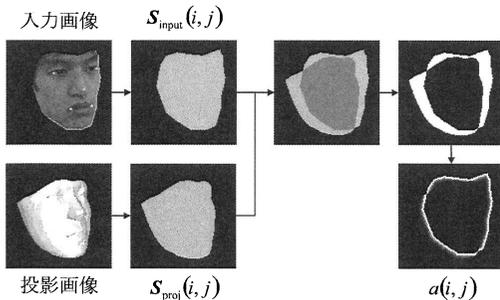


図4: B-XOR による評価

B-XOR による評価は2つの画像の XOR に輪郭から

の距離で重み付けしたものである。位置合わせされた入力画像、投影画像それぞれに対して、そのシルエット画像をそれぞれ S_{input} , S_{proj} としたとき、B-XOR 評価関数は

$$f_{B-XOR}(S_{input}, S_{proj}) = \sum_{i=1}^h \sum_{j=1}^w a(i, j) \quad (5)$$

$$a(i, j) = \begin{cases} \frac{1}{d_{input}(i, j)^2} & \text{if } S_{input}(i, j) \neq S_{proj}(i, j) \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

$$d_{input}(i, j) = D(S_{input})(i, j) + D(\tilde{S}_{input})(i, j) \quad (7)$$

として定義される。 w, h はそれぞれ画像の幅と高さである。また、 \tilde{S} は2値画像 S の明度反転画像であり、 $D(S)$ は S のユークリッド距離変換である。この関数は入力画像の輪郭から離れれば離れるほどその距離の2乗のオーダーで減少するような重み付けをしていることになる。

次に図5に外輪郭による評価の概念図を示す。

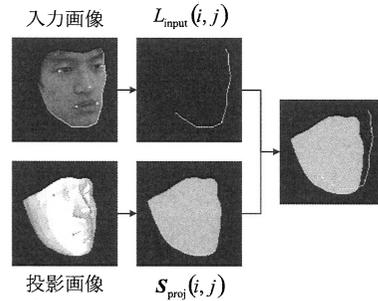


図5: 外輪郭による評価

外輪郭評価関数は、入力画像の外輪郭を L_{input} としたとき、

$$f_{contour}(L_{input}, S_{proj}) = \sum_{i=1}^h \sum_{j=1}^w b(i, j) \quad (8)$$

$$b(i, j) = \begin{cases} d_{proj}(i, j) & \text{if } L_{input}(i, j) = \text{white} \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

$$d_{proj}(i, j) = D(S_{proj})(i, j) + D(\tilde{S}_{proj})(i, j) \quad (10)$$

と定義される。すなわち、入力画像の外輪郭から投影画像のシルエットまでの距離の総和を評価値として用いる。この評価値により、内輪郭が最適化されることによって、真に最適化されるべき外輪郭の最適化が行われないことを防ぐことができる。

4.4.2 特徴点評価関数

特徴点評価関数は入力画像と投影画像の特徴点(鼻尖点・目尻点・口角点)座標の類似度を評価する。位置合わせされた入力画像、投影画像の特徴点をそれぞれ $\mathbf{P}_{\text{input}}, \mathbf{P}_{\text{proj}}$ とする。この行列はそれぞれの行に鼻尖点・左目尻点・右目尻点・左口角点・右口角点の画像座標ベクトルを格納した行列である。例えば、 $\mathbf{P}_{\text{input } 1}$ は入力画像の鼻尖点の座標ベクトルを表し、 $\mathbf{P}_{\text{input } 2}$ は入力画像の左目尻点の座標ベクトルを表す。

このとき、評価関数は

$$f_{\text{feature}}(\mathbf{P}_{\text{input}}, \mathbf{P}_{\text{proj}}) = \sum_{i=1}^5 \|\mathbf{P}_{\text{input } i} - \mathbf{P}_{\text{proj } i}\| \quad (11)$$

と定義される。ただし、鼻尖点で位置合わせを行っているため、鼻尖点の座標ベクトルは必ず一致し、 $\mathbf{P}_{\text{input } 1} - \mathbf{P}_{\text{proj } 1} = 0$ である。

図6にこの特徴点による評価の概念図を示す。

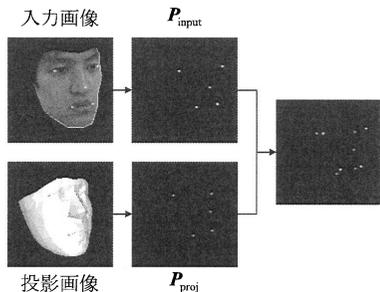


図6: 特徴点による評価

式(11)は、図6で表されたそれぞれの点同士のユークリッド距離の総和を計算している。

4.4.3 テクスチャ評価関数

位置合わせされた入力画像、投影画像それぞれに対して、その画像とテクスチャマッピングされた投影画像を $\mathbf{I}_{\text{input}}, \mathbf{I}_{\text{proj}}$ とする。このとき、テクスチャ評価関数は、

$$f_{\text{brightness}}(\mathbf{I}_{\text{input}}, \mathbf{I}_{\text{proj}}) = \sum_{i=1}^h \sum_{j=1}^w c(i, j) \quad (12)$$

$$c(i, j) = \begin{cases} \|\mathbf{I}_{\text{input}}(i, j) - \mathbf{I}_{\text{proj}}(i, j)\| & \text{if } \mathbf{S}_{\text{input}}(i, j) = \mathbf{T}_{\text{proj}}(i, j) = \text{white} \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

と定義される。ただし、 \mathbf{T}_{proj} はテクスチャマッピングを行ったパッチかどうかを表すマスク画像である。

すなわち、この関数はテクスチャが貼れている比較可能な領域に対して、差の絶対値の総和(SAD)を評価値とする。これにより、輪郭と特徴点だけでは評価し切れなかった細かな部分まで評価することができる。

図7に陰影評価の概念図を示す。

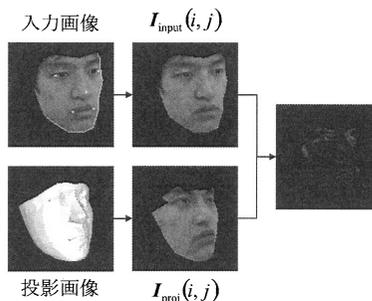


図7: 陰影による評価

最も右の図が差の絶対値を表す画像であり、これが黒一色に近づくとも評価値が下がることになる。

4.4.4 顔尤度評価関数

顔尤度評価関数は顔形状の20次元ベクトルを入力として、データベース上の顔形状集合との相関値を評価値とする。引数の顔形状ベクトルを $\mathbf{s} = [s_1, s_2, \dots, s_{20}]^T$ とすると、顔尤度評価関数は

$$f_{\text{likelihood}}(\mathbf{s}) = 1 - \exp\left(-\frac{1}{2} \sum_{i=1}^{20} \frac{s_i^2}{\lambda_i}\right) \quad (14)$$

と定義される。このとき、 $\lambda_i (1 \leq i \leq 20)$ は各主成分を表す固有ベクトルに対応する固有値で、その主成分における学習データの分散を表す。この関数は人物の顔形状が平均顔形状からの正規分布確率密度で存在すると仮定して近似したものである。すなわち、引数の顔形状がデータベースに登録されている顔形状から離れると評価値が上がっていく関数である。

5 実験

5.1 顔追跡撮影実験

実際に顔を追跡撮影した結果を図8に示す。矩形が Haar-Like 特徴量に基づく顔検出結果を表している。この矩形が表示されていない場合は Haar-Like 特徴量に基づく顔検出が行えなかったことを示す。パンチルト

カメラが自動的に首を振り、常に顔が画面中央に来るように追跡を行っていることがわかる。顔が横向きの場合でも、肌色情報に基づいて輪郭を抽出し追跡することができる。

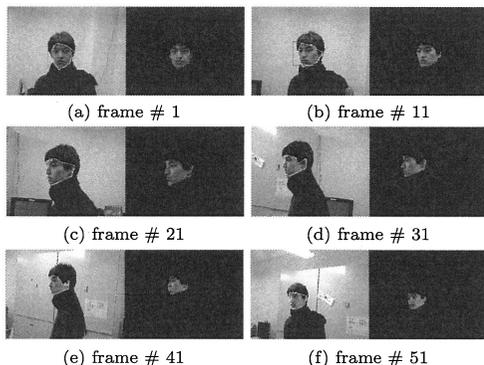


図 8: 顔追跡撮影の結果

5.2 3次元顔形状復元実験

次に入力画像群から3次元顔形状を復元する実験を行った。人物Aの10枚の入力顔画像それぞれに対して、3次元顔形状復元結果を復元された姿勢で重畳したものを図9に示す。図上の点は顔の特徴点を表す。投影画像の輪郭が入力画像の外輪郭にフィッティングし、顔特徴点も近い位置に最適化されていることが確認できる。また、メッシュを見ると鼻筋などの骨格にも沿うように復元されていることがわかる。

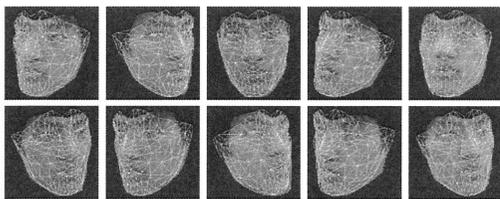


図 9: 3次元顔形状重畳結果(人物A)

図10に復元結果を用いた自由視点画像の生成例を示す。上段は復元された顔の3次元形状を表示したものである。下段は復元された顔姿勢から、最も正面向きに近いとされた入力画像をテクスチャとして貼り付けたものである。このように違和感のない画像が生成されることが確認できる。

人物B, Cについても同様の実験を行った。それぞれを図11-図14に示す。なお、人物Cは5枚の入力画像から復元した。

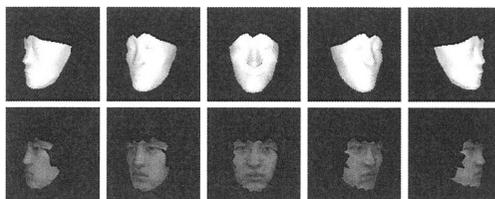


図 10: 自由視点画像(人物A)

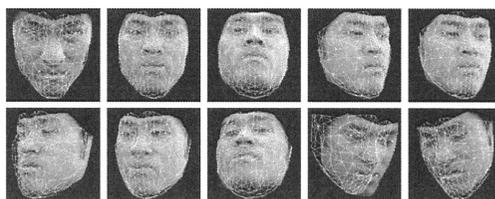


図 11: 3次元顔形状重畳結果(人物B)

5.3 凸包の再帰的縮退によるノイズ除去の有効性

人物の顔が横を向いている場合、肌色情報を用いて顔検出を行うと、眉、目、口といった肌色でない場所が欠けてしまう場合がある。提案する凸包の縮退を用いることで、このノイズ除去を行った。その様子を図15に示す。図15(a)は入力となるノイズを含む顔シルエット画像である。図15(b)に示したものは、元の輪郭と凸包の縮退によるノイズ除去後の輪郭である。出力となるノイズ除去後のシルエット画像が図15(c)となる。凹状欠損のみを埋めており、ノイズとして認識されない部分は全く輪郭が鈍っていないことがわかる。

5.4 復元された3次元顔形状の妥当性

復元された3次元顔形状の妥当性を定量的に評価するため、復元で用いなかった画像に対して評価値の計算を行った。その結果を表1に示す。平均顔形状を用いて、顔の姿勢を推定し、すべての評価関数の評価値

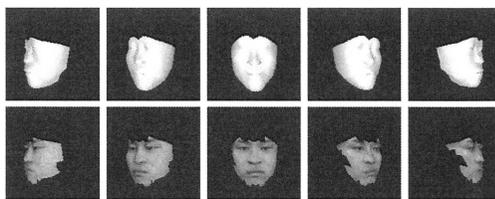


図 12: 自由視点画像(人物B)



図 13: 3次元顔形状重畳結果 (人物 C)

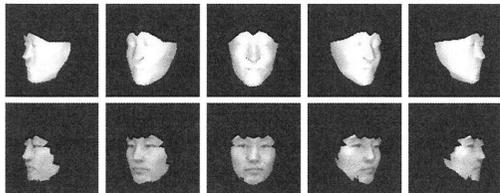


図 14: 自由視点画像 (人物 C)

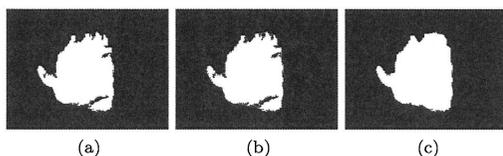


図 15: 凸包の縮退によるノイズ除去

を重み付けして統合した値が最適化前の評価値である。一方、復元された3次元顔形状を用いて顔の姿勢を推定し、同様に評価値を求めたものが最適化後の評価値である。この操作を人物 A, B, C それぞれに対して行った。

表 1: 誤差の評価

	人物 A	人物 B	人物 C
最適化前	545	692	736
最適化後	444	582	317

すべての人物に対して評価値が一樣に減少している。これにより、平均顔に比べ本手法で復元された3次元顔形状の方が撮影対象の人物に最適化された形状となっていることが確認できた。

6 まとめ

本論文ではパンチルトカメラを用いて自動追尾撮影された画像列から顔の3次元形状を復元し、その形状を用いて自由視点画像を生成することによって、不審者の捜査を支援するサーベイランスシステムを提案した。また、輪郭評価関数に外輪郭という概念を導入することで復元精度を上げ、新たに顔尤度評価関数を導入

することにより安定した復元を行えるように改善した。

参考文献

- [1] Amit K. Roy Chowdhury and Rama Chellappa: "Face reconstruction from monocular video using uncertainty analysis and a generic model", Computer Vision and Image Understanding, Vol. 91, Issues. 1-2, pp. 188-213, 2003.
- [2] Volker Blanz and Thomas Vetter: "A Morphable Model for the Synthesis of 3D Faces", Proc. of ACM SIGGRAPH'99, pp187-194, 1999
- [3] Sami Romdhani, Volker Blanz and Thomas Vetter: "Face Identification by Fitting a 3D Morphable Model Using Linear Shape and Texture Error Functions", Proc. of ECCV 2002, Vol. 4, pp. 3-19, 2002.
- [4] V. Blanz and T. Vetter: "Face recognition based on fitting a 3D morphable model", IEEE Trans. on PAMI, Vol. 25, No. 9, 2003.
- [5] Yuxiao Hu, Dalong Jiang, Shuicheng Yan, Lei Zhang and Hongjiang zhang: "Automatic 3D Reconstruction for Face Recognition", Proc.of Int'l Conf. on FGR'04, pp. 843-848, 2004.
- [6] Dalong Jiang, Yuxiao Hu, Shuicheng Yan, Lei Zhang, Hongjiang Zhang and Wen Gao: "Efficient 3D reconstruction for face recognition", Pattern Recognition, Vol. 38, Issue. 6, pp. 787-798, 2005.
- [7] 竹内敏夫, 斎藤英雄, 持丸正明: "非同期多視点カメラからの顔の3次元復元", MIRU2007, pp. 1289-1294, 2007.
- [8] Jack Sklansky: "Finding the convex hull of a simple polygon". Pattern Recognition Letters, Vol. 1, Issue. 2, pp. 79-83, 1982.
- [9] Baback Moghaddam, Jinho Lee, Hanspeter Pfister and Raghu Machiraju: "Model-Based 3D Face Capture with Shape-from-Silhouettes". Proc. of Int'l Workshop on AMFG'03