

ツビューからの顔画像からの 3D モデルの自動化 Request of the Manuscript for IPSJ-Kansai 2012

圓城 達也† 吳 海元†
Tatuya Enjo Haiyuan Wu

概要

顔の 3 次元モデルの自動生成システムを実現するための第一歩として、本論文では顔パーツのコーナー検出に特化している FAST (Features from Accelerated Segment Test) を提案する。局所領域内の画素単位の明度差の情報を用いたコーナー抽出アルゴリズムは、他の手法（例えば、Harris）よりもしっかりと顔パーツのコーナーを特徴点として検出できる。本論文では、FAST の閾値問題を解決するために、OpenCV のライブラリで検出される顔領域内で明度の正規化を行う。顔画像から検出される顔パーツ特徴点の特徴量を記述するのはマッチング精度の高い SIFT (Scale-invariant feature transform) の特徴記述法を採用する。数人の顔画像セットを用いた実験結果より、同じ人物の 2 枚の姿勢が異なる画像から正しく顔パーツの特徴点を抽出し、マッチングすることができることを確認した。

1. はじめに

1.1 研究目的と背景

個人の顔の 3 次元モデルの生成ができれば、顔の表情や、頭部姿勢など自由に生成でき、個人識別、表情認識などに貢献できるので、さまざまな分野で応用することが期待できる。

例えば、映画やゲームなどの映像製作においては、人間の顔を表現した作品が数多く作られており、3 次元顔モデルの需要も高まってきている。しかし、3 次元顔モデルを作る作業は一般的に簡単ではなく、膨大な時間やコストを必要とする。これらのことを踏まえ、もしデジカメで異なる 2 つの視点から撮影されたある人の顔写真からその人の個人 3 次元顔モデルを自動的に生成することができれば、膨大な時間をかけずに、コストも抑えて、比較的簡単にたくさんの個人の 3 次元顔モデルを作成することが可能となる。これにより、ユーザが自分で自由に表情・姿勢変化、顔パーツの位置・大きさ調整などの応用が可能になる。また、各個人の顔のデータベース化が実現でき、犯罪者データベースの 3 次元化、顔の骨格に関する種々の研究、アニメーションの作成など数多くの応用分野への活用が広がる。

図 1 に示すように、フォトカルクツビューというソフトウェアは、角度をつけて撮影された 2 枚の画像から 3 次元的な計測ができるソフトウェアであり、その中の機能の一つ「3D 点」を使い、顔のポリゴンモデルの作成を行うことが可能である。「3D 点」とは、その名の通り、異なる視点から撮影された 2 枚の画像間において、対応付

けた点の 3D 座標値を得ることができる機能である。得られた点群の 3 次元座標値から、ポリゴンを作成、修正を行い、テクスチャを貼り、顔のポリゴンモデルを作成できる。

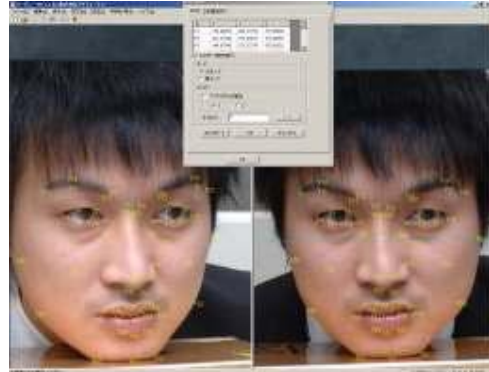


図 1 フォトカルクツビューで顔のポリゴンモデルを作成するための画像セット例と手動で指定された特徴点群（黄色の点群）

しかし、2枚画像間の顔特徴点の対応付けは手作業で行い、手間がかかる問題が残っている。また、より正確な顔表面を作るため、顔パーツの特徴点だけではなく、頬あたりにも対応付けた点ができるだけ多く取れる方が望ましい。この作業自身が手作業でも容易ではない。

図 1 に示すような特徴点を自動的に取得できるようにすることが本研究の目的と位置付ける。この目標を達成するため、

1) 顔パーツ(目、鼻、口)の両端の正確なコーナー特徴点を自動的に抽出すること

2) 頬の特徴点を自動的に抽出することが必要となる。本論文ではこのような目標を実現するための第一歩として、FAST (Features from Accelerated Segment Test) [1][5]とSIFT(Scale-invariant feature transform)[6]を用いた顔パーツの特徴点の自動抽出と対応付け方法を提案する。

1.2 関連研究

顔特徴点や特定部位の抽出や追跡に関する研究は、従来から活発に発表されている。

たとえば、東芝研究開発の湯浅真由美らは複数の顔特徴点を自動的に検出する方法[2]を提案している。湯浅らはあらかじめ定めた14個の顔特徴点について、それぞれの特徴点の性質に基づいた特徴点候補検出、パターンによる検証、近傍特徴点に対する幾何学的な位置条件を利用して特徴点の候補を検出する。さらに検出された特徴点候補全体に対する幾何学的な整合性を3D標準顔形状モデルを利用して判定することより、顔らしい配置のもの

を選択する処理を行なっている。

他にも、法政大学の小谷秀輝らは、シーン中から顔を検出する方法として、まず、カラー画像の領域分割処理によって顔面内部に位置する左右の目と唇の候補を多数抽出した後、それら3つの特徴点を基準としたアフィン変換で切り出される矩形領域の濃淡画像が顔パターンとみなしうるか否かを部分空間類別法を用いて判定する方法[4]をとりあげている。カラー画像処理については画像処理ソフトHALCONを使用して入力画像から顔を自動的に検出するシステムを構築し、さまざまな画像入力条件のもとでのそのシステムによる顔検出の安定性を評価している。

華春生らは眼鏡の有無と顔特徴点の自動検出方法を提案している[3]。彼らはほぼ正面を向いている顔画像から、色情報(肌色と髪色)を抽出し、エッジ情報(強度と方向)を抽出する。この情報より、顔と顔器官を含む領域を決定する。頬部分の肌らしさの平均値と目尻周辺のそれと比、頬部分のエッジ強度の平均値と目尻周辺のそれとの比から、眼鏡の有無を検出する。検出された顔器官を含む領域から、口、鼻、目の位置を予測し、エッジ情報より確定する。最後に、SUSAN オペレータを用いて顔特徴点を抽出している。

1.3 論文の構成

本論文では、第2章で第3章～4章で用いられる関連アルゴリズムを紹介し、原理や問題点などについて説明する。第3章では提案する手法について述べる。第4章では実験の条件と結果を、第5章ではまとめと今後の課題を述べる。

2. 関連アルゴリズム

2.1 Harris コーナー検出アルゴリズム

Harris オペレータは、画像中からエッジやコーナーを抽出できるアルゴリズムである。エッジとコーナーについて、次のように考えることができる。

エッジ … 画像中で輝度値が急激に変化する点。

コーナー … 2つのエッジの交差、またはエッジ上で勾配の強い点。

グレースケールの2次元画像を I とし、画素 (u, v) 上に画像の局所パッチを置き、 (x, y) ずつずらすことを考える。2つのパッチの差分の重みつき2乗和 $S(x, y)$ は以下で与えられる。

$$S(x, y) = \sum_u \sum_v w(u, v) (I(u, v) - I(u - x, v - y))^2$$

$I(u + x, v + y)$ はテイラー展開により近似できる。 I_x, I_y は I を x 方向と y 方向でそれぞれ偏微分するものとする、

$$I(u + x, v + y) \approx I(u, v) + I_x(u, v)x + I_y(u, v)y$$

この近似式から次の近似式が導かれる。

$$S(x, y) \approx \sum_u \sum_v w(u, v) (I_x(u, v)x + I_y(u, v)y)^2$$

これを行列で書くと次のようになる。

$$S(x, y) \approx \begin{pmatrix} x & y \end{pmatrix} A \begin{pmatrix} x \\ y \end{pmatrix}$$

ここで、

$$A = \begin{bmatrix} I_x I_x & I_x I_y \\ I_x I_y & I_y I_y \end{bmatrix}$$

この行列 A から求めることができる2つの固有値を λ_1, λ_2 とする。図2に示すように、固有値によって以下の3つのパターンに分けて、注目画素がエッジ、あるいはコーナーであるかどうかを判断して検出することができる。

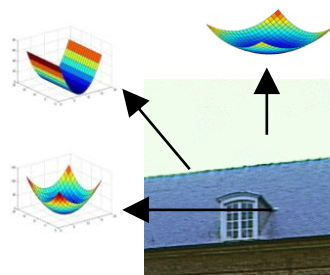


図2 Harris コーナー検出アルゴリズムのイメージ

- (1) $\lambda_1 \approx 0$ かつ $\lambda_2 \approx 0$ のとき、平坦 (エッジでもコーナーでもない)
- (2) $\begin{cases} \lambda_1 \approx 0 \text{ かつ } \lambda_2 \text{ が正の大きな値} \\ \lambda_2 \approx 0 \text{ かつ } \lambda_1 \text{ が正の大きな値} \end{cases}$ のいずれかであればエッジ
- (3) λ_1, λ_2 が大きく、異なる正の値であればコーナー

opencv の `cvGoodFeaturesToTorack` という関数を用い、最小固有値が十分に大きいところをコーナーとして検出できる。図3には検出されたコーナー点の画像例を示す。



図3 Harris コーナー検出アルゴリズムで検出されたコーナー点の画像例

2.2 FAST コーナー検出アルゴリズム

FAST (Features from Accelerated Segment Test) はS.M. Smithらによって開発されたSUSAN (Smallest Univalve Segment Assimilating Nucleus) というアルゴリズムに基づいて開発されたコーナー検出アルゴリズムである。

SUSANコーナー検出法の原理のイメージは図4に示し、計算式が以下になる。

$$c(i, p) = \begin{cases} 1 & \text{if } |I(i) - I(p)| \leq t \\ 0 & \text{if } |I(i) - I(p)| > t \end{cases}$$

$$n(p) = \sum_{i=1}^N c(i, p)$$

$C(i, p)$ は画像内の画素(P)を中心とし、半径 r とした円内の画素 $i(N$ 個) と P の明度の差を比較し、差が閾値 t より小さい場合は 1 とし、閾値より大きい場合は 0 とする。 $n(P)$ が円内の $C(i, p)$ の合計であり、 $3N/4$ より少ない場合、コーナー(特徴点)として検出する。また、 $n(P)=N/2$ の場合、エッジとして検出できる。

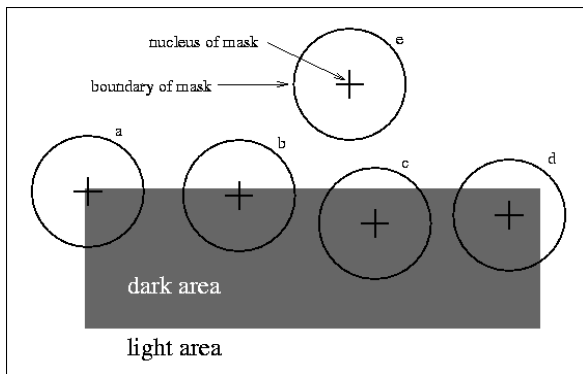


図4 SUSAN コーナー検出アルゴリズムのイメージ

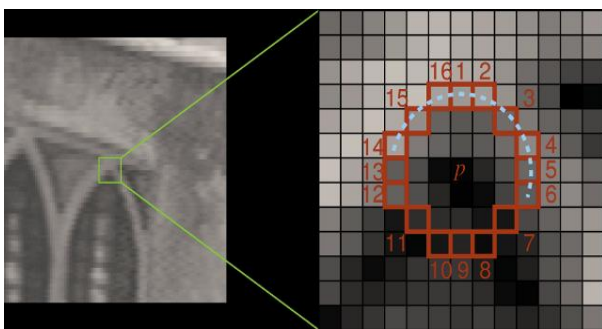


図5 FAST コーナー検出アルゴリズムのイメージ

E. RostenらはSUSANの高速化を行い、FASTを提案している。FASTの原理は図5に示す。この場合、画像内の画素(P)を中心とし、半径 r とした円周上の画素(N個)とPの

明度の差を比較し、差が閾値より大きい場合は+1、合計数が $3N/4$ より多い場合、コーナー(特徴点)として検出する。さらにスピードアップするために、図5の右側のように、1(上)と9(下)の比較、1(上)と9(下)と5(右)の比較、1(上)と9(下)と15(左)の比較で処理を行う。図6から、SUSANより、FASTは画素間の比較数を減らしていることが分かる。図6にFASTコーナー検出アルゴリズムで検出された画像例を示す。



図6 FAST コーナー検出アルゴリズムで検出されたコーナー点の画像例

2.3 SIFTによる特徴量記述とその対応付け

SIFT(Scale-invariant feature transform)の処理は、特徴点(あるいは、キーポイントと呼ぶ)の検出(detection)と特徴量の記述(description)の2段階からなる。SIFTは、スケール変化、回転変化に不変な特徴量を記述することができるため、近年、物体認識・追跡、画像探索など様々なCVの分野で、幅広く利用されている。

2.3.1 特徴量の記述

本論文では、固有値あるいはFASTによるコーナー(特徴点)を検出し、SIFTの特徴量記述部分のみ利用する。特徴量の記述には特徴点の周辺領域が持つ勾配情報に基づいて128次元の特徴量を求める。

特徴点の周辺領域に注目し、勾配方向×Gaussian重みのヒストグラムを求め、ピークに対応する方向をオリエンテーションとして割り当てることにする(ただしピークは常時1つとは限らない)。また、オリエンテーションとは各特徴点を中心とした、領域内の最も頻度の高い勾配を示す。図7(a)に示すように、特徴点を持つオリエンテーションの向きに座標軸を合わせて特徴量を記述するため、回転に対する不変性が得られることになる。

次に、図7(b)に示すように、回転した領域の各マスに位置する画素の輝度勾配を得て、そこに特徴点を中心とするガウスフィルタを畳みこみ、輝度勾配に重み付けをする。このガウスフィルタの標準偏差の大きさは、SIFTによる特徴点を検出する場合、特徴点を検出されたDoG画像の平滑化スケールによって決定される。平滑化スケールに応じてフィルタの大きさを変えることで、構造的に

同一の範囲の特徴量を記述することになり、これによりスケール変化に対する不変性が得られる。本論文では、検出された顔のサイズによってガウスフィルタの標準偏差の大きさを推定する。

最後に、図 7(c)に示すように、16×16 の領域を 4×4 ごとに区切る。区切ったそれぞれの領域について、領域にある各点の輝度勾配を 8 方向に分割し、方向ヒストグラムを作成する。区切られた 16 の領域について、それぞれ 8 方向のヒストグラムが作成されるため、合わせて 128 (=4×4×8) 次元のベクトルが作られることになる。また、128 次元の特徴ベクトルをノルムで正規化することにより、照明変化に対して頑強な特徴量を得ることができる。

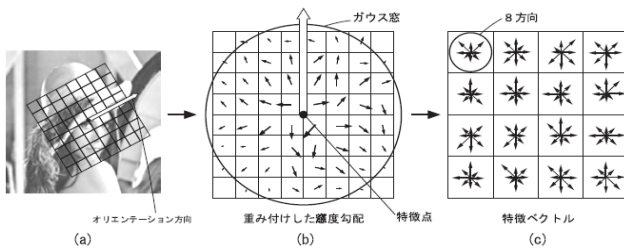


図 7 SIFT による特徴量記述

2.3.2. SIFT特徴量による2枚画像間の特徴点の対応付け

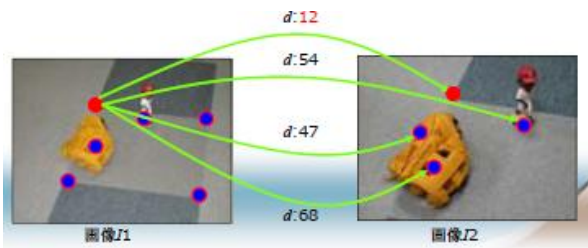


図 8 特徴量による特徴点の対応付け例

画像 A と画像 B 同士の特徴点 K_a と K_b をマッチングす

るには、特徴点の特徴量 V をもとに、ユークリッド距離 d を以下の式で計算する。

$$d(v^{Ka}, v^{Kb}) = \sqrt{\sum_{i=1}^{128} (v_i^{Ka} - v_i^{Kb})^2}$$

そして、この値が最小になるペア同士を対応づいた特徴点ペアとする。図 8 は実際の画像ペア間の特徴点のユークリッド距離による対応付けの例を示し、2 枚の画像内の特徴点同士は最小距離を取っている場合、3 次元空間内の同じ点を映っている可能性が高い。

3. 顔パーツのコーナー検出に特化している FAST

3.1. 部分領域の明度正規化によるFASTの安定化

FASTは円形領域の中心画素の明度と円周上の画素の明度間の差によって中心画素がコーナーか否かを判定する手法である。画素単位でコーナー特徴点を検出できるメリットがある一方、その差を評価する閾値が入力画像の撮影状況によって調整する必要があり、正しく設定できない場合、予想通りのパフォーマンスができない可能性がある。

予備実験に、照明条件の影響で画像内の顔領域が暗くなったり、白とびになったりすると、固定な閾値で顔パーツのコーナー点を全部検出できなくなる。また、閾値を入力画像ごとに調整すれば、顔パーツのコーナー点を全部検出できることが確認できている。

本論文では、FASTで顔パーツのコーナー検出処理の安定化と汎用化を図るために、FASTの閾値を入力画像によらず最適な固定値で決める必要があると考え、OpenCVのライブラリで検出される顔領域内で動的に局所領域の明度の正規化を行う工夫を加える。

具体的に、検出される顔（図9を参照）の中心のx,y座標と中心からの半径情報を用いて顔を含む矩形領域に対して、cvsetImageROIで顔領域内のみで明度の正規化を行う。正規化処理にはcvNormalizeを用いる。

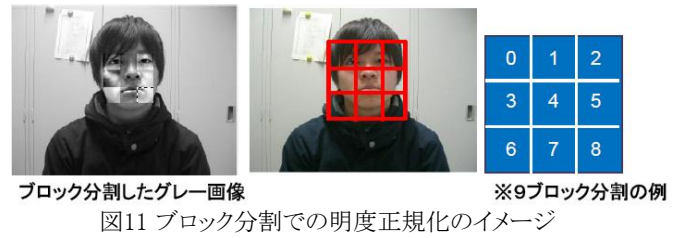


図9 カスケード分類器(harcascade_frontalface_default)による顔検出

図10には、(a)Harrisで検出されたコーナー点、(b)従来のFASTで検出されたコーナー点、(c)検出された顔領域全

体として明るさを正規化した後、FASTで検出されたコーナー点の結果をそれぞれ示す。Harrisを用いて顔画像内での特徴点の検出を行ったがコーナーでの特徴点検出においてFASTのほうが高いパフォーマンスを示していることが分かる。また、提案手法の方が右目や口元での特徴点が新たに検出され、より多くのコーナーでの特徴点が抽出されている。すなわち、局所領域内で正規化を行うと、固定閾値でも特徴点が多く検出できていることが分かる。また、特徴点検出領域内で明度の差が小さいほど精度の高い検出が可能となることが分かる。

は、検出された顔領域をブロック分割して局所的に明度を正規化する方法を提案する。提案手法イメージについては、9ブロック分割した場合を例として図11に示す。



本論文では、顔矩形を3×3の9ブロック、4×4の16ブロックの各ブロック内で明度の正規化を行ったあと、FASTで顔パーツのコーナー特徴点を検出する。図12には、顔画像でブロック分割無し、9ブロック分割、16ブロック分割で明るさの正規化を行い、同じ閾値を用いてFASTで検出された特徴点の画像例を示す。



(a) Harrisで検出されたコーナー点



(b) FASTで検出されたコーナー点(明度正規化前)



(c) FASTで検出されたコーナー点(提案手法:明度正規化後)

図10 各手法によるコーナー特徴点の検出比較例

しかし、顔画像から検出された顔の範囲に多少のずれ誤差がある。たとえば顔の検出範囲が広く、領域内に白と黒の画素が存在する場合であれば、明度の正規化の効果があまり得られない問題が残っている。

3.2 ブロック分割での明度正規化によるFAST特徴点検出

上記の問題を解決し、より精度を向上するために、本論文で



(a) ブロック分割無し



(b) 9ブロック分割



(c) 16ブロック分割

図12 ブロック分割によるFAST特徴点の検出例

本論文では、FASTによる顔パーツのコーナー特徴点の検出率を評価する基準として眉毛の両端、目の両端、鼻の穴の下、唇の両端の計12個のコーナーパーツ特徴点を定める。図12に示す3種類の方法で検出されたコーナーパーツ点検出数比較グラフは図13に示す。図13から16ブロック分割の場合最も良いパフォーマンスが得られている。

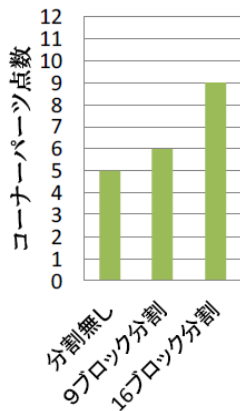


図13 コーナーパーツ点検出数比較グラフ

3.3 各ブロック内での明度ヒストグラムの線形変換 (Linear Stretch)

本論文では明るさやコントラストの調節は、線形変換 (Linear Stretch) 法を用いる。

$$Y = \frac{255}{b-a} * (x-a)$$

ただし、画像には輝度の低いノイズや高いノイズ等を含むので、ノイズを考慮して設定しなければならない。したがって、図14に示すように、ヒストグラムの両端をブロック矩形のピクセル数(縦ピクセル×横ピクセル)×0.02の数を閾値として、明度値の個数がその閾値を超えている明度値を新たな最小値a, 最大値bとする。

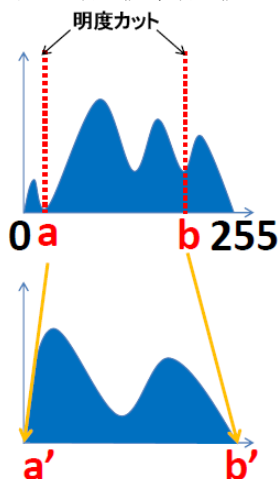


図14 ヒストグラムの画素値カットの明度線形変換のイメージ

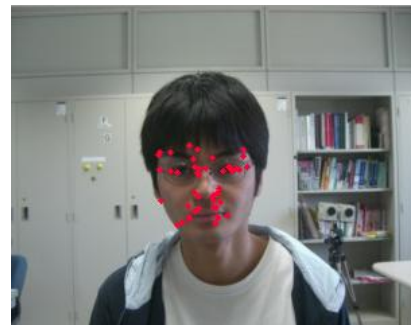
3.4 FAST特徴点検出の問題と解決策

3.2で上述したように、検出された顔矩形に基づいて固定の配置で $n \times n$ ブロックに分割し、その矩形内で適度な閾値から明度の最大値、最小値を算出し明度の線形変換を行った後 FAST 処理を行うと、コーナー検出結果は顔領域の検出精度に依存し、オーバーラップの問題もある。

本提案手法では、まず予備実験による、一番よい結果が得られた 16 ブロック分割の方法を採用し、動的な配置法を採用する。具体的に、その 16 ブロック矩形サイズの縦横の長さを N として、検出された顔矩形内を縦横ともに $N/2$ ずつ計 81 個の矩形をかぶせながらその領域内部で明度線形変換と FAST 処理を繰り返し行った。そのイメージについては以下に(図 15 参照)示す。



図 15 FAST 処理におけるブロック移動のイメージ



(a) 提案手法(動的に 16 ブロック分割の結果)



(b) 固定の 16 ブロック分割の結果

図 16 動的と固定ブロック分割の比較結果

図 16.(a)に動的に 16 ブロック分割で明度正規化をしながら FAST でコーナー検出の結果を示す。図 16.(b)に固定の 16 ブロック分割で明度正規化をしながら FAST でコーナー検出の結果を示す。

一般的に、FAST 処理に用いられる明度の閾値は大きくすると(厳しくすると)抽出できる特徴点の数は少なくなる。図 16 には、(a)では閾値を 80 とし、(b)では閾値を 50 として FAST 処理をしているものの、どちらもほぼ同数の特徴点が検出できている。提案手法の方がより顔パーツのコーナー特徴点を検出できる。これは、提案手法ではしっかりと明度のヒストグラムの両端の明度値をカットして明度の正規化を行えているためである。

4. 実験

提案手法の有効性と汎用性を確認するため、複数人の異なる姿勢の顔画像を撮影し、異なるツービューの顔画像を用いて、顔画像間の顔パーツのキーポイント特徴点の対応付けの実験を行った。

提案手法の顔パーツのコーナー検出に特化している FAST で検出されたコーナー点周辺の特徴量を SIFT の特徴量記述のアルゴリズムを用いて記述し、同じ顔パーツのコーナー特徴点間の対応付けを行った。

FAST 中の明度の閾値については全て同じパラメータで実験を行った。SIFT で記述される特徴量はアフィン変換に対して不変ではないことを考慮し、今回の実験では、左右の首の回転についての画像は首軸まわりの回転を小さい画像セットを撮影した。実験結果例は図 17 に示す。



(a)



(b)

図17 SIFTによるFAST特徴点の対応付けの結果

5. 結論

本論文では顔3次元モデルを生成するための第一歩として、顔パーツのコーナー検出に特化しているとされている FAST (Features from Accelerated Segment Test) を提案した。

実画像を用いた実験より、顔パーツのコーナー特徴点の検出と対応付けを考察した。まず、特徴点抽出において OpenCV ライブラリの Harris と FAST を用いたが、FAST の方が比較的正しい位置に抽出されることを確認できた。また、FAST については顔の領域を抽出し、ブロック単位の局所的な明度の正規化を行ってから FAST 処理を行うことにより正規化前よりも多くの特徴点が得られることも確認できた。

一方、SIFT で特徴点の特徴量を記述し、対応づけを行えることも確認できたが、SIFT が奥行方向の回転に対して不変できないので、正面顔との対応付けは首軸まわりで回転した顔に比べて、顔の傾きによる回転をした顔のほうが対応付けのパフォーマンスが高いことが分かった。

今後、アフィン変換に強い新たな特徴量記述の考察をする必要がある。さらに、コーナーパーツ点の情報を元にして頬からの特徴点を検出する方法も考察していかなければならない。

謝辞

本研究の一部は、文部科学省科学研究費補助金基盤研究(c) 24500205 の補助を受けている。

参考文献

- [1]FAST Corner Detection:
<http://www.edwardrosten.com/work/fast.html>
- [2] 湯浅真由美、武口智行、小坂谷達夫、山口修:静止画顔認証のための自動顔特徴点検出、信学技報、PRMU2006-222, 2007

- [3] 華春生、呉海元、塩山忠義: 眼鏡の有無と顔特徴点の自動検出, 信学技報、PRMU2002-184, 2003
- [4] 小谷秀輝, 関口利彦, 真壁弘樹, 赤松茂: カラー画像処理と部分空間類別法を用いたシーン中の顔検出、信学技報、PRMU2003-278, 2004
- [5] E. Rosten and T. Drummond, Machine learning for high-speed corner detection, European Conference on Computer Vision. 2006
- [6] Lowe, David G. Object recognition from local scale-invariant features, Proceedings of the International Conference on Computer Vision. **2**. pp. 1150–1157, 1999