

3次元顔モデルを用いた顔認識システム —ステレオ画像からの3次元形状の復元と認識—

千葉 紀之[†] 花泉 弘[†]

[†]法政大学情報科学部 〒184-8584 東京都小金井市梶野町 3-7-2

E-mail: † {i02t9011, hana}@k.hosei.ac.jp

あらまし 近年、セキュリティへの関心が高まり、広い分野で個人認証が導入されている。特に、バイオメトリクスと言われる人間の身体的特徴を用いた生体認証が一般的になりつつある。その中で顔認識は、認識対象に対する心理的抵抗感が少ないという長所がある。我々は、3次元顔モデルを用いた顔の認識システムの開発を目標としている。従来の顔認識手法には、撮影状況(照明環境、画像内の顔の大きさ、向き)に対して影響されやすいという問題を含む手法が多く存在する。本システムでは、3次元顔モデルの形状情報を積極的に用いた認識を行うことで、撮影状況に依らない頑強なシステムを構成することが可能である。本システムにおける、撮影状況の違いに対しての有効性を実験により示す。

キーワード 顔認識, 3次元顔モデル, ステレオマッチング, 3次元形状復元

Three-Dimensional Face Recognition System — 3D Shape Reconstruction from Stereo Pairs —

Noriyuki CHIBA[†] Hiroshi HANAIZUMI[†]

[†] Faculty of Computer and Information Sciences, Hosei University

E-mail: † {i02t9011, hana}@k.hosei.ac.jp

Abstract Increasing popularity of security of prompted interests in person identification using face recognition. However 2-D (image-based) face recognition methods appear to be sensitive to illumination conditions, head-pose, scale and etc. In this paper, we propose a new 3-D face recognition system using 3-D face shapes obtained from stereo pair face images. The 3-D face shapes are rendered under various illumination conditions, head-pose, scale and occlusion. The effectiveness of our system is demonstrated by the experimental results.

Keyword Face Recognition, Stereo Matching, 3-D Reconstruction, 3-D Face Shape

1. はじめに

近年、国内外の社会不安や企業の IT(情報技術)活用の広がりとともに、セキュリティへの関心が高まっている。特にバイオメトリクスと言われる、人間の身体的特徴を用いた生体認証が一般的になりつつある。主な身体的特徴として、指紋、掌紋、手形、手の甲の静脈、虹彩、顔、DNA、音声などが挙げられる。現状、コストや簡便性などに問題は残るものの、究極の本人認証という点から、バイオメトリクス技術を用いた本人認証が最も有効であると言われている。

身体的特徴の中で顔認識は、人間が行っている個人認証の基本であり、指紋などと比較し心理的抵抗感が少なく、最も親しみやすい個人認証である。現状の顔認識システムは 90%以上の高い認識が可能である。ただし、この高い認識率を実現するためには、一定の良好な照明環境下と、ほぼ正面を向いた顔という制約がある。顔認識システムの実用化にあたり、いくつかの課題が存在する。耐照明環境性、顔の向きへの対応、眼鏡等への対応、表情変化への対応、化粧、髪、眉、頬にかかる髪等への対応、なりすまし防止などが挙げられる^{[1][2]}。我々の目標とする顔認識システムは、利用者に対する制約が緩い認識システムである。つまり利用者が何も身につけることなく、かつ意識して認識されるよう行動しなくてもよいシステムである。そこで、ステレオ視により復元された 3 次元顔モデルを用いることで、上記課題に対し、より頑強なシステムの開発を目指す。

従来の認識手法は、主に顔画像から抽出した特徴ベクトル(顔画像からの輝度値(色)、顔形状、オプティカルフローなど)をもとに行われている。しかし、特徴ベクトルが輝度値(色)から求められた場合には、照明環境や顔の向きに対し影響を受けることとなる。そこで、本システムでは、認識に 3 次元形状を用いることで、照明環境性や顔の向きに対し、より頑強なシステムになると考える。ステレオ画像からの対応点検出では色情報を用いることとなるが、取得されるステレオ画像はすべてが同じ照明環境下で撮影されたものであるため、対応点検出精度に与える照明の影響は小さいものだと考える。さらに、認識には色情報を用いないため、3 次元形状が正確に復元された場合において、照明の影響を受けることはなく認識が可能である。また、ステレオ視による 3 次元形状の復元は、カメラパラメータが既知の場合、不定性なく 3 次元形状が復元可能であり^[3]、画像内の顔領域の大きさの違いに対し対応が可能である。3 次元顔モデルは回転変換を行うことで顔の向きに対しても対応することが可能である。

本論文では、3 次元顔モデルを用いた顔認識システ

ムを提案する。なお、本システムにおける 3 次元顔モデルとは、人頭部の耳より前面部分(顔)のサーフェスモデルを表す。まず、ステレオ画像から 3 次元形状を復元するための画像間の対応付けを行うステレオマッチング手法について述べる。ステレオマッチングは、精度はもちろん、多くの対応点を検出する必要がある。続いて、復元された 3 次元顔モデルをもとにデータベースへ登録される登録 3 次元顔モデルの生成について述べる。登録 3 次元顔モデルは、広範囲の顔の向きに対応させるため、複数視点から復元された 3 次元顔モデルを統合することにより生成する。本論文での一つの視点とは、本システムを構成している 3 台のカメラからなる。最後に 3 次元顔モデルの形状情報を利用した認識手法について述べ、本システムの有効性を実験により示す。

2. システム構成

本システムの構成を図 1 に示す。デジタルカメラより取得されたステレオ画像からマッチング処理を行い、対象物の 3 次元形状(3 次元顔モデル)を復元する。3 次元顔モデルより特徴量を抽出し、データベースに登録されている登録 3 次元顔モデルとの類似度を計算することで、顔認識を行う。

本システムは 3 台のデジタルカメラから構成される。3 台のデジタルカメラから取得された 3 枚の画像すべての組み合わせで対応点が検出できた場合を真の対応点とすることで、ミスマッチを軽減し、3 次元形状復元精度の向上を図る。正面から見たおおよその配置を図 2 に示す。上段中央に 1 台、下段左右に 1 台ずつ配置する。なお、各カメラは基準立方体を用いて校正済みであるものとする。カメラを校正することで、透視カメラ行列 p を一般化した 3×4 の射影カメラ行列 p_a が求められている。

図 3 は、3 台のカメラから得られたステレオ画像を、それぞれ対応する位置に配置して示したものである。1 枚の画像サイズは 1408×1058 画素である。

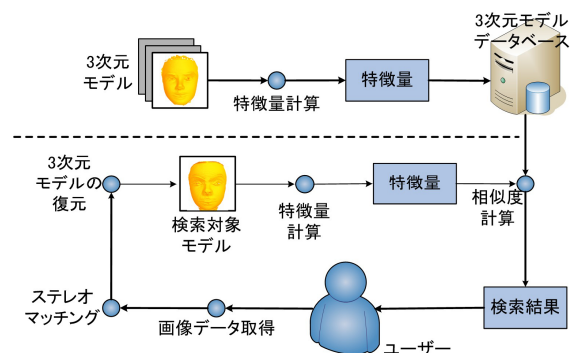


図 1 システム構成

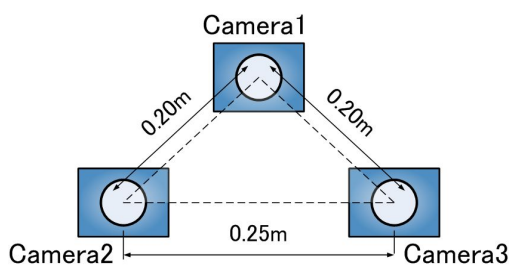


図2 3台のカメラ配置

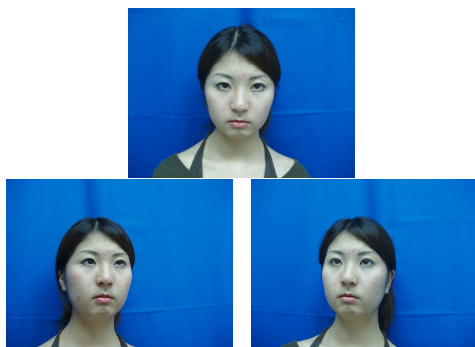


図3 3眼ステレオ画像

3. 原理および処理手順

3.1. ステレオマッチング

本システムの認識手法は3次元顔モデルの3次元形状をもとに行われるため、3次元顔モデルの復元精度が、認識精度に大きく影響を及ぼすこととなる。さらに、正確な認識を行うために、より多くの対応点を全画素マッチングにより検出する必要がある。そのため、そのためのマッチング手法について述べる。本手法では、処理時間を短縮するためにマッチング処理の階層化を行う^[4]。つまり、ステレオ画像間での対応点の検出は、高度差による画像上での位置ずれが大きいため、探索範囲が広がってしまい処理時間が莫大になってしまう。また、処理の階層化により全画素マッチングでの探索範囲が狭くなるため、対応点のミスマッチを抑制する効果もある。

3.1.1. プレマッチング処理

全画素マッチング処理を行うにあたり、まず探索を行うための走査開始点を設定する必要がある。プレマッチングを行い、顔領域上に走査開始点を設定する。処理手順として、画像内に一様に対応点对を配置する。探索領域内の対応点对に対し正規化相関(Normalized Correlation Coefficient : NCC)法を用いたマッチングを行い、求められた対応点を、全画素マッチングのための走査開始点とする。

3.1.2. 4 近傍領域への走査による全画素マッチング

画像間の全画素を効率よくマッチングするために、探索マスクを4近傍領域へ走査し全画素マッチングを

行う(図4)。プレマッチングにより求められた走査開始点より走査を開始する。マッチング処理には、プレマッチングと同様に NCC 法を用いた。本手法では、一つの走査開始点に対し、複数の走査経路が存在する。そのため、一つの走査経路で走査が終了した場合でも、他の走査経路では走査が続いている。走査開始点が少ない場合でも、より多くの画素を走査することが可能である。また、この走査手法は、既にマッチング処理済の画素に対しても、別の走査経路から進入し新たにマッチング処理を行うといった処理を簡単に実現することが可能である。改めてマッチング処理を行うことで、対応点が検出できなかった画素に対し対応点を検出することや、ミスマッチであった画素に対して、正確な対応点を検出することが可能である。ただし、一つの画素へ同様の処理を何度も繰り返し、計算時間が増大する可能性が考えられる。そこで、以下に示す制約条件を課すことで、無駄な処理を回避している。

- 一つの画素に対し、同一の走査開始点から走査が開始された走査経路からの進入を禁止する。ただし、処理から一定の期間が過ぎた走査開始点から開始された走査経路からの進入は認めるものとする。
- 異なる走査開始点からの走査経路による進入でも、同一の進入経路のものが処理済である場合、その進入を禁止する。

続いて、一つの走査経路に対する走査終了条件を以下に示す。

- NCC 法より求められた相関係数が閾値以下である。
- 走査を行う画素が存在しない。
- 処理画素とその画素への進入前の画素において、それぞれ復元された3次元座標間の距離を3次元空間上での移動量とし、その値が閾値以上である。

3次元空間上での移動量を計算するために、対応点検出ごとに3次元形状の復元を行う。この移動量は、ミスマッチの判定基準としてだけではなく、顔領域のみの探索を行うことを可能としている。つまり、顔領域から探索領域が外れた場合、移動量が大きく変化するため、顔領域外への探索を行うことを抑制する。

すべての走査経路において、走査を行う画素が存在しない場合、全画素マッチングの処理を終了する。

なお、対応点はすべての画像の組み合わせにおいて相関係数の合計が最も高いものとする。また、すべての画像間における対応点が一対一の対応関係であるものとする。基準画像内の複数の画素が参照画像内の、ある一つの画素に対し相関係数が最大を示した場合、相関係数の合計が最も高い組合せのものを対応がある

こととする。

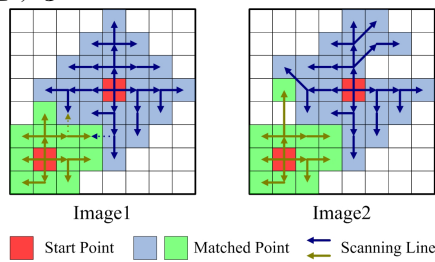


図4 4近傍領域への走査による
全画素マッチング

3.2. 登録3次元顔モデルの生成

本システムでは、顔の向きに対し頑強なシステムとするために、より広範囲に対応した3次元顔モデルを、データベースへ登録する必要がある。一度の撮影で取得されたステレオ画像より復元された3次元顔モデルは、顔の一部のみのサーフェスモデルとなる。そこで、複数の3次元顔モデルを統合することで、より広範囲に対応した3次元顔モデルを生成する。また、一視点から復元された3次元顔モデルは、オクルージョン等の問題により、鼻の周囲などに穴が開く場合がある。複数の3次元顔モデルを統合することで、その穴をなくすことが可能である。本システムの認識手法は、基準となる登録3次元顔モデルに穴が開いていた場合、認識結果に影響を及ぼす。そのために登録3次元顔モデル内の穴をなくすことも重要となってくる。図5は一組のステレオ視におけるローカル座標系、図6はそれらの統合座標系である。

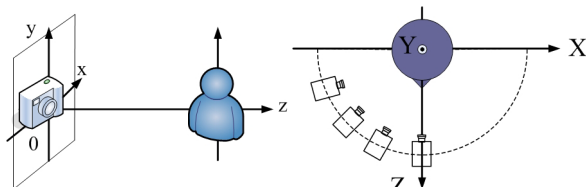


図5 ローカル座標系

図6 統合座標系

3.2.1. 座標系の統合

ここでは、簡単のため異なる2視点から取得された3次元顔モデルの統合手法について述べる。それぞれの3次元顔モデルを M_1 、 M_2 とする。 M_1 の座標系を基準座標系とし、 M_2 の座標系を基準座標系へ変換することとする。 M_1 、 M_2 にはそれぞれ射影カメラ行列 P_{a1} 、 P_{a2} が存在する。つまり、 P_{a2} を基準座標系となる3次元空間中へ復元可能となるように再計算することで M_2 を M_1 へ統合する。

まず、 M_1 、 M_2 のそれぞれに対応するステレオ画像を用いて、モデル間の対応関係をマッチング処理により求める。この処理により、 M_2 のステレオ画像上の座標と M_1 の3次元上の座標との対応関係がわかる。

この対応関係をもとに射影カメラ行列 P_{a2} の再計算を行う。求められた P_{a2} より M_2 の3次元形状を復元することで、 M_1 の座標系への統合が行われる。

以上の計算をすべての視点で行ない、座標系を統合する。

3.2.2. 3次元顔モデルの統合

複数視点からの3次元顔モデルを統合する場合の問題点として、ある部分について、3次元座標がわずかに異なる複数の復元結果(ずれ)ができてしまう場合がある(図7(a))。そのため、それぞれの3次元座標を統合する必要がある(図7(b))。ここで、統合する2つの3次元顔モデルを、それぞれ M_1 、 M_2 とする。基準を M_1 とし、 M_2 はすでに座標系の統合は行われているものとする。3次元顔モデル間の対応関係は、座標の統合の際のマッチング処理により既知である。本システムの3次元顔モデルは、通常、3枚のステレオ画像より3次元形状が復元される。しかし、3次元顔モデル間のステレオ画像同士で対応点が存在する画像上の座標においては、そのすべてのステレオ画像を用いて3次元座標を復元することが可能である。つまり、 M_1 と M_2 、2つの3次元顔モデルを統合する場合、それぞれの3枚のステレオ画像を合わせた6枚のステレオ画像より3次元形状を復元することとなる。ただし、3次元顔モデル間における対応関係が存在しない画像上の座標においては、それぞれの3次元顔モデルにおける3枚のステレオ画像から、3次元形状を復元する。

複数視点の統合モデルを生成する場合、視点数を N とすると、最大で $3N$ 枚のステレオ画像を用いて3次元形状が復元されることとなる。

以上の処理を行うことで、統合3次元顔モデル上に存在した3次元座標のずれをなくすことが可能である。

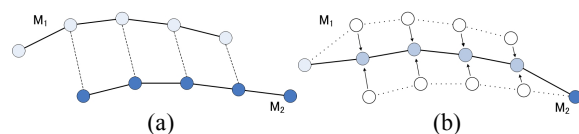


図7 モデルの統合

(a)3次元座標のずれ、(b)3次元座標の統合

3.3. 3次元顔モデルの形状情報を利用した顔認識

3.3.1. ICPアルゴリズム

本システムにおける顔認識は、色情報を用いず、形状情報を用いた認識を行う。本認識手法はICP(Iterative Closest Point)アルゴリズム^[5]を利用する。ICPアルゴリズムとは、特徴点の抽出といった前処理などを必要とせず、より高精度に重ね合わせを行う手法である。図8にICPアルゴリズムの処理手順を示す。

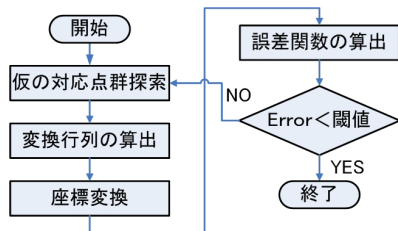


図8 ICP アルゴリズムの処理手順

ICP アルゴリズムは、仮の対応点を求め、それをもとに変換行列を算出し処理を行っていく。3次元空間中の2点 \vec{r}_1, \vec{r}_2 間のユークリッド距離 d は以下のよう表すことができる。

$$\vec{r}_1 = (x_1, y_1, z_1), \vec{r}_2 = (x_2, y_2, z_2) \quad (1)$$

$$d(r_1, r_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \quad (2)$$

ここで、 A を N_a 個の点 \vec{a}_i からなる点群とし、

$$A = \{\vec{a}_i\} \quad i=1 \dots N_a \quad (3)$$

任意の点 \vec{p} と点群 A の距離を次のように定義する。

$$d(\vec{p}, A) = \min_{i \in N_a} d(\vec{p}, \vec{a}_i) \quad (4)$$

式(4)からわかるように、点群 A に含まれる点の中で最も距離の近い点との距離(最近傍距離)が \vec{p} と点群 A との距離となる。

ここで、 N_t 個の点群からなる点群 $T = \{t_i | i \in N_t\}$ に対し、異なる視点から計測された N_s 個の点からなる点群 $S = \{s_i | i \in N_s\}$ の重ね合わせを行う場合、点群 S の各点 s_i において点群 T との最近傍距離は以下のよう求める。

$$d(s_i, T) = \min_{j \in 1 \dots N_t} d(s_i, t_j) \quad (5)$$

$$= d(s_i, m_i) \quad (6)$$

s_i に対応する点 $m_i \in T$ とすると、 s_i の対応点集群 M は

$$M = C(S, T) \quad (7)$$

と表せる。ここで C は最近傍点を求める関数である。本手法では、仮の対応点となる M との最近傍距離を用いた認識を行う。

3.3.2. 形状情報を利用した認識手法

本認識手法では、まずICPアルゴリズムにより、登録3次元顔モデルに対して、認識計算を行う対象である3次元顔モデルを重ね合わせる。反復計算が終了した時点での、式(5)、(6)における点群 T を登録3次元顔モデル、点群 S を認識対象3次元顔モデルとした最近傍距離 d を求める。この d を類似度として認識に用いる。つまり、 d が0に近いほど形状が類似している。類似度の計算には、3次元顔モデル内のすべての3次元上の点における最近傍距離を用いるのではなく、距離の小さい上位半分を用いることで、3次元復元の際の誤差やノイズの影響を軽減した。また、表情の変化

や眉、頬にかかる髪に対しても有効であると考えられる。つまり、距離の大きい部分が表情の変化や髪にかかる部分である可能性が高いためである。

4. 結果および議論

4.1. ステレオマッチング

図9に対応点の検出結果を示す。白色部分が対応点の検出された部分を表す。85601点の対応点検出されている。非常に密に対応点検出されていることがわかる。ただし、鼻の周りなどオクルージョン等の問題により検出されなかった部分は存在するが、良好な結果と言える。

図10に上段中央のカメラにおける対応点検出の遷移を示す。今回の計算には、410ステップを要した。各図の右下部分にステップ数を示す。走査開始点毎に色を変化させている。各走査開始点からの走査で検出された対応点が複雑に絡み合っていることがわかる。このことから、本手法が mismatches やある走査経路では検出ができなかった画素などに対し、再処理を行うことでより多くの対応点検出されていることがわかる。

図11に復元された3次元顔モデルを示す。左より正面、右側、上部を視点とした画像を示す。対応点検出できなかった部分は穴が開いているが、3次元形状が正確に復元されていることがわかる。



図9 対応点検出結果

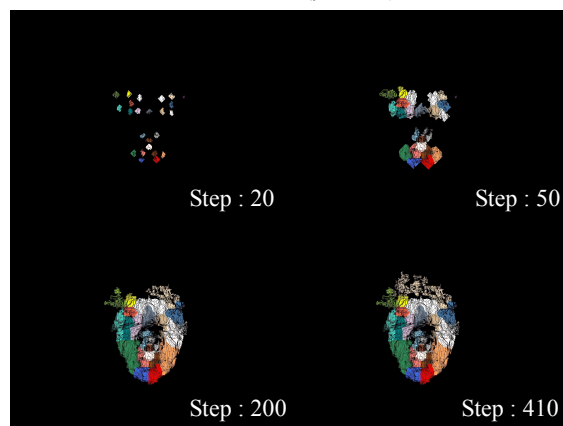


図10 対応点検出の遷移

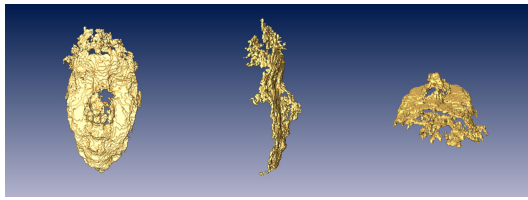


図 11 復元された 3 次元顔モデル

4.2. 登録 3 次元顔モデルの生成

図 12 に生成された登録 3 次元顔モデルを示す。本実験では、図 11 に示した被験者が正面を向いて撮影されたステレオ画像より復元された 3 次元顔モデルの他に、左右へ約 20° 向いて撮影したステレオ画像から復元された 3 次元顔モデルの 3 つの 3 次元顔モデルを用いて、登録 3 次元顔モデルを生成した。図 11 では、オクルージョンの問題等で開いていた鼻の周りの穴がなくなっていることがわかる。生成された登録 3 次元顔モデルは精度良く生成されている。

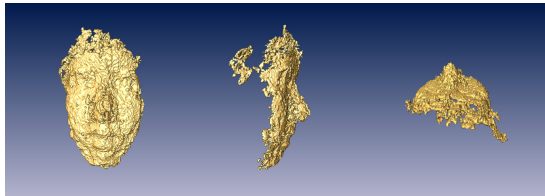


図 12 生成された登録 3 次元顔モデル

4.3. 3 次元顔モデルの形状を用いた認識

3 人の被験者 (A ~ C) を対象に認識実験を行った。それぞれの被験者に対して、正面、左右へ約 20° 向いた計 3 組のステレオ画像より復元された 3 次元顔モデルを統合し、登録 3 次元顔モデルとした。続いて、被験者を別途撮影し、復元された 3 次元顔モデルを認識対象とした。なお、撮影は登録 3 次元顔モデルと同じ 3 方向から行い、それぞれの 3 次元顔モデルに対しての類似度計算を行った。今回の実験では、被験者に対し撮影時の表情は統一し、画像内の顔向きの違いに対しての有効性の確認を行った。被験者 (A ~ C) の認識対象となる、上段中央のカメラで撮影された正面向きの画像を図 13 に示す。



図 13 認識実験の被験者(左より A, B, C)

実験結果を表 1 に示す。すべての被験者に対して、同一人物が最も良い結果を示している。また、顔の向きの違いに対して、類似度がほとんど影響されていないことがわかる。ただし、本人と別人との明確な差が

あるとも言い難い。被験者の人数が少ないこともあるが、今回の結果だけでは、個人認証のための条件を決定することは難しいと考える。

表 1 形状情報を用いた認識結果

		登録モデル			
		A0	B0	C0	
被験者	A1	正面	0.039	2.035	1.583
		左	0.154	0.938	1.245
		右	0.238	1.764	1.784
	B1	正面	1.239	0.583	1.453
		左	1.983	1.113	2.654
		右	1.881	1.093	1.993
	C1	正面	2.135	2.005	0.099
		左	2.349	1.668	0.563
		右	1.989	1.732	0.873

5. まとめ

本論文では、3 次元顔モデルを用いた顔認識は、画像内の顔の向きの違いに対して有効であることを示した。認識には 3 次元形状を直接用いることで、比較的シンプルな手法で良好な結果を得ることができた。

また、3 次元形状の復元手法も、顔のみではなく様々な対象物に利用することが可能である。本システムは、特別な装置を使用せず、デジタルカメラから取得された画像より 3 次元形状の復元を行っているため、比較的容易に導入することができる。

複数視点からのステレオ画像から復元された 3 次元顔モデルを統合することで、人の顔の大部分の形状データを取得することができた。このデータは認識だけではなく、多方面への応用が期待できる。

現時点では、少ない被験者に対しての実験しか行っていない。今後、被験者を増やし、表情の変化や眉、頬にかかる髪等に対しても、本認識手法が有効であることを示す必要がある。

文 献

- [1] 瀬戸洋一:サイバーセキュリティにおける生体認識技術, 共立出版株式会社, 2002.5.
- [2] (社)日本自動認識システム協会:これでわかったバイオメトリクス, オーム社, 2001.9.
- [3] 佐藤淳:コンピュータビジョン-視覚の幾何学-, コロナ社, 1999.5.
- [4] H.Hanaizumi, T.Takesaki, S.Fujimura, "AN AUTOMATED METHOD FOR ESTIMATING TERRAIN HEIGHT FROM SPOT STEREO PAIR IMAGE", Proc IGARSS'95, 2167/2169, (1995)
- [5] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes", IEEE Transactions on Pattern Analysis and Machine Intelligence, pp.239-256, 14 Feb. 1992