2枚の顔写真を用いた任意方向の顔画像の生成

向川康博[†] 中村裕一[††] 大田友一[††]

本論文では，2枚の顔写真のみを入力として，任意方向から見た顔画像を生成する手法を提案する．本手法の特徴は，2枚の入力画像のみを用いているが，いわゆる中割り法とは異なり3次元座標変換に忠実な見え方の変化を再現できること，ステレオ法による3次元復元とは異なり入力画像の撮影方向に関する情報を必要としないこと，structure-from-motion の原理を基礎としているが陽には3次元形状の復元を行わないことである．正射影モデル下で対象物体として剛体を仮定すれば，与えられた2枚の入力画像上の特徴点の2次元座標の線形結合で，任意方向の見え方画像上の特徴点座標が表現できる．線形結合の係数は生成画像についての指定方向に基づいて決定する．これらの特徴点を頂点とする2次元の3角形パッチモデルに対して，入力画像のテクスチャを適応的に張り付けることにより顔画像を生成する．

Synthesis of Arbitrarily Oriented Facial Views from Two Images

YASUHIRO MUKAIGAWA,[†] YUICHI NAKAMURA[††] and YUICHI OHTA[††]

In this paper, we propose a method for synthesizing arbitrarily oriented facial views from only two facial images. The method includes the following features: unlike the 2-D interpolation method, the views by true 3-D rotation can be generated; unlike the stereo, the camera parameters of the two input images are not necessary; although it is based on the structure-from-motion theory, the 3-D structure of a face is not explicitly reconstructed. Assuming the rigidity and the orthographic projection, the 2-D coordinates of feature points in any view can be expressed as a linear combination of those in the two input images. The coefficients of the linear combination are determined based on the facial orientation to be generated. The facial image is synthesized by adaptively mapping the texture of the input images to the 2-D triangular patch model whose vertices are the feature points.

1. はじめに

人間の顔画像は，特定個人の特徴だけでなく，表情表現など多くの情報を含ませることができるため，高度なマンマシンインターフェースに応用が期待されている．しかし，顔は，人間にとって最も親しみのある対象の1つであるため，実際の顔と合成された顔画像のわずかな違いでも，人間は大きな不自然さを感じてしまう場合がある．そのため，より自然な顔画像の生成法についての研究が進められている．

本来3次元物体である顔の様々な方向からの見え方を画像として生成するために，頭部を3次元計測して顔の3次元形状モデルを作成する手法が提案されている1)．任意方向からの見え方は，コンピュータ内部に持たせた3次元形状モデルを所定の向きに回転し2次元画像上に投影することによって生成できる．この場合，単一のモデルに基いてあらゆる方向から見た場合の見え方を生成することができ，蓄積すべきデータ量が少なくて済むという利点がある．しかし，この手法は各個人について正確な3次元情報を持ったモデルの作成を必要とし，レーザ光等を用いた特殊な計測装置を使わずに正確に3次元形状を計測することは一般には困難である．

また，標準的な頭部の3次元形状モデルを用意しておいて，各個人について適切な変形を行い，対象人物の3次元形状モデルとして利用する方法も提案されている2)．標準的なモデルを用いることで各個人について頭部形状を計測する必要はなくなるが，個人差を忠実に反映した形状の変形は困難である．

一方，対象物体の3次元形状モデルを持たずに，2枚の画像を合成して別の見え方画像を生成する手法として，中割り法3)がある．中割り法では2枚の画像間で対応付けられた点の座標や輝度値の内分により，2枚の画像の内挿画像を生成する．しかし，この方法で生成できるのは，一方の見え方からもう一方の見え方に2次元

[†] 筑波大学 工学研究科
Doctoral Program in Engineering, University of Tsukuba

[††] 筑波大学 電子・情報工学系
Institute of Information Sciences and Electronics, University of Tsukuba

的に変化させた中間の見え方だけであり，3次元回転による見え方の変化は本質的に表現できない．

このように従来の方法では，精度の高い顔画像を生成するためには精密な3次元形状モデルを必要とし，数枚の顔写真などのような簡単に得られる情報だけを利用する場合には，生成できる顔の向きに制限が多く，自然な顔画像の生成が困難であった．

この問題に対して，本研究では，撮影方向が未知の2枚の顔写真のみを入力として視覚的に満足できる顔画像を生成することを目指し，2枚の顔写真のみの組合せによって，任意方向からの見え方を生成する手法を提案する． 本手法は，structure-from-motion の原理を基礎としているが，陽には3次元復元を行わず，しかも，幾何学的に正確な3次元座標変換による見え方の変化を再現できる．

2. キーアイディア

Structure-from-motion の原理[4] によれば，対象物体を複数の方向から撮影した画像を用いて物体の3次元形状を復元することが可能であり，得られた3次元形状モデルを利用して様々な方向からの見え方画像を生成することができる．しかし，画像からの3次元形状復元は一般に不安定であり，復元精度が誤差などに大きく影響されることが多い．このため，時系列上で観測される非常に多数の画像を利用し，最小自乗法によって形状復元を安定化する枠組みが盛んに研究されている[5]．しかし，本研究の目標のように，処理の最終目的が2次元画像の生成の場合には，必ずしも中間段階で3次元形状を復元する必要はない．すなわち，入力として与えられた2次元画像から，見え方の異なる2次元画像を直接生成できれば十分である．

本研究では，対象物体は剛体を仮定し，正射影モデル下において与えられた2枚の顔写真から，3次元形状を陽に復元することなく任意方向の顔画像を生成する手法を明かにする．複数の画像を入力として，任意方向の見え方画像が生成可能であることの原理は，3枚の入力画像があれば3次元形状の復元が可能とする structure-from-motion の原理に関連している．Koenderink 等[6] は入力画像が3枚でなく2枚だけの時には3次元アフィン変換の自由度を残しての形状復元が可能であることを示している．また，Ullman 等[7] は複数枚の入力画像上の特徴点の2次元座標の線形結合により，任意方向の見え方の特徴点座標が表現可能であることを示し，2次元見え方モデルを用いた3次元剛体認識への応用について論じている．

本研究では，入力画像上の2次元座標の線形結合により任意方向の見え方の2次元座標が表現できるとする Ullman 等が示した原理を画像生成に応用し，3次元形状を陽に復元することなく，2枚の入力画像の組み合わせによって直接に任意方向の見え方画像を生成することを試みた．以下，3章ではその原理を説明し，指定された方向に対する線形結合の係数の決定法について述べる．4章では3章で得られた2次元の3角形パッチモデルに対して，入力画像のテクスチャを適応的に張り付けて画像生成する方法を述べる．5章に実験結果を示す．

3. 特徴点座標の推定

3.1 線形結合による座標算出の原理

3次元空間中の点 $\boldsymbol{p} = [x\ y\ z]^T$ を3次元空間中で回転・平行移動した座標 $\boldsymbol{q} = [x'\ y'\ z']^T$ は，回転成分を $r_{11\sim33}$，平行移動成分を t_x, t_y, t_z とすると，同次座標系では 4×4 の行列を用いて式 (1) のように求めることができる．(以下では座標を同次座標系で表す)

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \tag{1}$$

ここで3次元空間中に n 個の点があり，i 番目の点の座標を $[x_i\ y_i\ z_i\ 1]^T$ として，これらの座標を並べた行列を $\boldsymbol{P}$ とする．

$$\boldsymbol{P} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ y_1 & y_2 & \cdots & y_n \\ z_1 & z_2 & \cdots & z_n \\ 1 & 1 & \cdots & 1 \end{bmatrix} \tag{2}$$

点の集合 $\boldsymbol{P}$ を異なる5種類の 4×4 の行列 $\boldsymbol{R}_j$ $(j = 1, \ldots, 5)$ で座標変換して得られる点の集合を $\boldsymbol{B}_j$ とする．

$$\boldsymbol{B}_j = \begin{bmatrix} \boldsymbol{x}_j \\ \boldsymbol{y}_j \\ \boldsymbol{z}_j \\ \boldsymbol{1} \end{bmatrix} = \begin{bmatrix} x_{j1} & x_{j2} & \cdots & x_{jn} \\ y_{j1} & y_{j2} & \cdots & y_{jn} \\ z_{j1} & z_{j2} & \cdots & z_{jn} \\ 1 & 1 & \cdots & 1 \end{bmatrix} = \boldsymbol{R}_j \boldsymbol{P} \tag{3}$$

ここで $[x_{ji}\ y_{ji}\ z_{ji}\ 1]^T$ は i 番目の点を j 番目の行列で変換した座標であり，x, y, z 座標をそれぞれ一定の順に並べたベクトルを $\boldsymbol{x}_j, \boldsymbol{y}_j, \boldsymbol{z}_j$ とする．

ここで $\boldsymbol{R}_j$ の第1行ベクトルを $\boldsymbol{r}_j$ とすると，

$$\boldsymbol{x}_j = \boldsymbol{r}_j \boldsymbol{P} \quad (j = 1, 2, \cdots, 5) \tag{4}$$

であるが，$\boldsymbol{r}_j$ は4次のベクトルであるので，$\boldsymbol{r}_1 \sim \boldsymbol{r}_4$ が1次独立であれば

$$\boldsymbol{r}_5 = a_1 \boldsymbol{r}_1 + a_2 \boldsymbol{r}_2 + a_3 \boldsymbol{r}_3 + a_4 \boldsymbol{r}_4 \tag{5}$$

となる係数 $a_k(k=1,2,3,4)$ が一意に存在する．すなわち式 (4) ~ (5) により，$\boldsymbol{R}_5$ で変換した場合の全ての点の x 座標は次式のように線形結合で簡潔に表現できる．

$$\boldsymbol{x}_5 = a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + a_3\boldsymbol{x}_3 + a_4\boldsymbol{x}_4 \qquad (6)$$

同様に y 座標も係数 $b_k(k=1,2,3,4)$ を用いて次式のように表現できる．

$$\boldsymbol{y}_5 = b_1\boldsymbol{y}_1 + b_2\boldsymbol{y}_2 + b_3\boldsymbol{y}_3 + b_4\boldsymbol{y}_4 \qquad (7)$$

このことから剛体･正射影モデルを仮定すると，3次元物体を任意方向から見た場合の画像上の点の2次元座標は，基底となる4枚の画像上での2次元座標の線形結合で表現でき，陽に3次元情報を復元することは必要でないことがわかる．

3.2 2枚の顔写真を利用した座標推定

方向の異なる4枚の顔写真 B_1, B_2, B_3, B_4 が基底となる画像として与えられたとする．これらは3次元の頭部を回転・平行移動・スケール変換したものを正射影により2次元に投影したものである．顔表面上に特徴点を設定し，それぞれの基底となる画像上で n 個の特徴点が対応付けられているものとする．B_1, B_2, B_3, B_4 上での全ての特徴点の x, y 座標を一定の順に並べたベクトルをそれぞれ $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4, \boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{y}_3, \boldsymbol{y}_4$ とする．もし，$\boldsymbol{x}_i(i=1,2,3,4)$ が互いに1次独立であり，$\boldsymbol{y}_i(i=1,2,3,4)$ も互いに1次独立であれば，別の方向からの見え方 $\hat{B}$ 上における特徴点の座標を表すベクトル $\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}$ は適当な係数 a_i, b_i を用いて，次式のように線形結合により表現できる．

$$\hat{\boldsymbol{x}} = \sum_{i=1}^{4} a_i\boldsymbol{x}_i \quad , \qquad \hat{\boldsymbol{y}} = \sum_{i=1}^{4} b_i\boldsymbol{y}_i \qquad (8)$$

しかし，図1に示すように，利用できる顔写真が B_1, B_2 の2枚だけであっても，これらに加えて B_1 を画像面に垂直な軸回りに回転させた見え方 B_1^{Rotate} と，B_1 を画像面上で平行移動させた見え方 B_1^{Shift} を用意し，これら4枚の画像を基底画像とすることで，4枚の顔写真が与えられた場合と同様のことが可能である．なぜなら，B_1^{Rotate} に対応する変換行列の第1行ベクトルは，B_1 に対応する変換行列 $\boldsymbol{R}_1$ の第1行ベクトルと第2行ベクトルから計算され，$\boldsymbol{r}_1, \boldsymbol{r}_2$ の線形結合では表現できない．また B_1^{Shift} に対応する変換行列の第1行ベクトルも $\boldsymbol{r}_1$ の第4成分に平行移動量を加えたものであり，$\boldsymbol{r}_1, \boldsymbol{r}_2$ の線形結合では表現できない．したがって，この4枚の見え方に対応するアフィン変換の第1行ベクトルは互いに1次独立になり，基底画像として利用可能である．

2枚の顔写真から生成した4枚の基底画像上での全ての特徴点の x 座標を一定の順に並べたベクトルをそれぞれ $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_1^{Rotate}, \boldsymbol{x}_1^{Shift}$ とすれば，

$$\hat{\boldsymbol{x}} = a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + a_3\boldsymbol{x}_1^{Rotate} + a_4\boldsymbol{x}_1^{Shift} \qquad (9)$$

と表現できる．y 座標に関しても同様である．ただし，4つのベクトル $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_1^{Rotate}, \boldsymbol{x}_1^{Shift}$ が1次独立になるためには，2枚の顔写真 B_1, B_2 が互いに，画像面に垂直な軸回りの関係や，画像面上での平行移動の関係でないことが必要である．この条件を満たしていれば，撮影方向が未知の2枚の顔写真のみを利用した場合でも，適当な係数さえ決定できれば任意方向からの見え方における全ての特徴点座標を求めることが可能となる．

3.3 線形結合の係数の決定

任意方向から見た場合の特徴点座標は，適当な係数さえ決定できれば，線形結合で求められることがわかった．本節では生成したい画像の顔の向きに対応する線形結合の係数を求める方法について述べる．

係数は x, y 座標それぞれについて4個で合計8個あるが，顔の見え方の変化は回転3自由度，平行移動2自由度，スケール変換1自由度の合計6自由度を持つのみであり，すべての係数値の組合せが3次元的に妥当な見え方に対応するわけではない．

3.3.1 代表的な特徴点の3次元座標が既知の場合

顔表面上の特徴点のうち，最低4点の代表的な特徴点について3次元座標が既知であれば，これをもとに線形結合の係数を決定することができる．これらの代表的な特徴点を3次元空間中で回転させ，所望の顔の向きにした場合の2次元座標を求め，x 座標を一定の順に並べたベクトルを $\hat{\boldsymbol{x}'}$ とする．また，これらの点の基底画像上での x 座標を一定の順に並べたベクトルをそれぞれ $\boldsymbol{x}'_1, \boldsymbol{x}'_2, \boldsymbol{x}'^{Rotate}_1, \boldsymbol{x}'^{Shift}_1$ とし，

$$\hat{\boldsymbol{x}'} = a_1\boldsymbol{x}'_1 + a_2\boldsymbol{x}'_2 + a_3\boldsymbol{x}'^{Rotate}_1 + a_4\boldsymbol{x}'^{Shift}_1 \qquad (10)$$

となる係数 $a_1 \sim a_4$ を求める．y 座標に関しても同様である．代表的な特徴点が5点以上ある場合は，最小自乗法により求めることができる．

3.3.2 指標画像を用いる場合

指標画像は生成したい画像の顔の向きを示す顔画像であり，この画像と同一の向きになるように新しい顔画像を生成する．指標画像で，比較的安定に抽出できる最低4点の代表的な特徴点の画像上での2次元座標がわかっていれば，上記の場合と同様にして係数を求めることができる．

3.4 Structure-from-motion による方法との違い

前節で述べた指標画像を用いる場合には，2枚の入

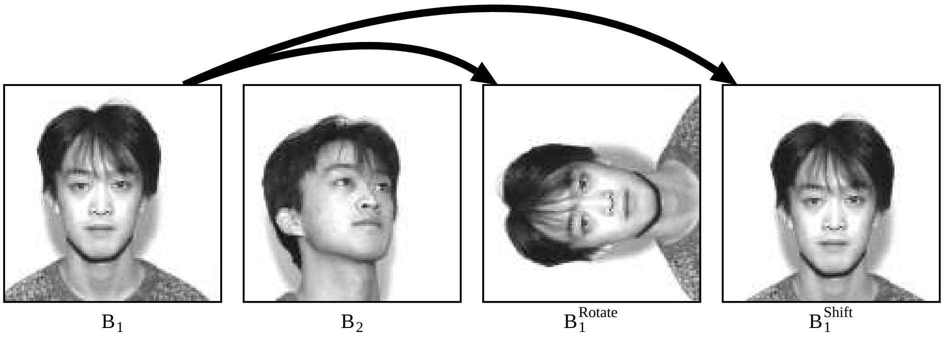

B_1　B_2　B_1^{Rotate}　B_1^{Shift}

図1　基底となる4枚の画像
Fig. 1 Four basis images

力画像と指標画像の3枚の画像上で最低4点の特徴点の対応付けを与える必要があり，これを用いればstructure-from-motion（SFM）によって，これらの特徴点の3次元座標と画像の撮影方向が計算できる．入力画像間の撮影方向が得られれば，ステレオ法により各点の3次元座標を計算し，任意方向から見た場合の2次元座標が推定できることになる．

しかし，本手法における特徴点の2次元座標の生成法と，SFMによる3次元座標の復元を介した2次元座標の生成では，本質的に生成できる画像の集合が異なる．本手法で生成できる画像は，剛体として考えた場合の形状を3次元的にアフィン変換した場合の見え方であり，線形結合の係数を変えることによって，この全ての見え方を生成することができる．ただし，人間が見た場合に適切な見え方を生成するためには，何らかの方法で線形結合の係数を適切に決めることが必要となる．その1つの方法が指標画像を用いることであり，もし指標画像として入力画像と同一人物の第3の見え方を与えれば，本手法とSFMとは同一の結果を生成する．

しかし，入力画像の特徴点座標に誤差が存在した場合は，SFMと本手法では算出される特徴点座標が異なる．誤差には，特徴点抽出の誤差などに加えて，入力画像と指標画像が他人であった場合の誤差が存在する．本手法では，問題設定として入力画像と指標画像は他人でなければ無意味であり，後者の場合の誤差は無視できない．

このような場合，SFMでは，異なる人物から得られる3枚の画像という，剛体仮説が成立していない入力から3次元形状復元を行うことになる．そのため，復元された形状は2人の人物の形状が入り混じったものとなり，これを3次元空間中でどのように回転しても入力画像の人物の見え方を正しく生成できない．つまり，生成される見え方は，剛体として正しくないばかりでなく，剛体として考えた場合のアフィン変換としても正しくない．

それに対し本手法では，指標画像で示された特徴点座標の集合に最も近い見え方を，2枚の入力画像上の特徴点座標の線形結合で算出する．いったん線形結合の係数を決めてしまえば，画像生成の際には指標画像での特徴点座標は使われない．そのため，生成される見え方は，剛体としては正しいという保証はないが，そのアフィン変換としては正しい見え方が生成される．また，何らかの付加的な方法で適切な線形結合の係数を拘束することができれば，剛体としても正しい見え方を生成できることになる．

4. 画像生成

顔画像上の特徴点の2次元座標を算出する方法について前節で述べた．しかし，顔画像を生成するためにはこれらの点座標をもとに，面を生成しなければならない．本節では，特徴点の間を埋める面をテクスチャマッピングにより生成する手法について述べる．

4.1 顔画像のためのパッチモデル

テクスチャマッピングにより面を生成するために，顔表面上の3個の特徴点の間に面を張った，図2に示す3角形パッチモデルを用意する．このモデルの特徴点数は45点であり，表面の曲率の大きい所には密に配置してある．各パッチはこれらの特徴点を頂点とする3角形であり，面数は80面である．このモデルに顔写真から得られるテクスチャをマッピングすることによって画像を生成する．

本研究のように，特徴点座標を2次元座標の線形結合

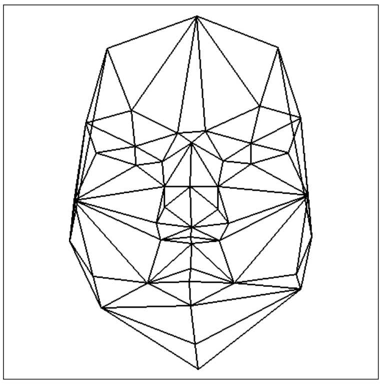


図2 3角形パッチモデル（2次元）
Fig. 2 Triangular patch model (two dimensional)

により求めた場合には，3角形パッチモデルは2次元平面上で定義されているだけであり，奥行き情報は欠落しているため，パッチどうしの正確な前後関係は分からない．しかし，3角形パッチモデル生成時に，各3角形の頂点を常に外から見て一定の順（例えば時計周り）に統一して記述しておけば，生成画像上での3角形パッチの頂点の出現順を調べることにより，各パッチの表裏の判定が可能となる．従って，裏側を向いている，すなわち完全に隠れているパッチにテクスチャがマッピングされることを防ぐのは容易である．人間の頭部のように凸多面体に近い形状の場合には，この程度の表裏の判定ができれば実用上十分である．

4.2 テクスチャの選択

テクスチャマッピングでは，通常1枚の画像からテクスチャを切り出すことが多い．しかし，裏側になって見えない部分があるため，1枚の顔写真から得られるのは顔表面上の部分的なテクスチャだけであり，あらゆる方向からの見え方の生成に対応することは困難である．本研究のように方向の異なる2枚の顔写真が利用できる場合には，これらを適切に組み合わせることで，生成できる見え方の範囲を広げることができる．本節では2枚の顔写真から得られるテクスチャをどのように選択し，組み合わせるかについて検討する．

各3角形パッチにマッピングするテクスチャを，2種類の顔写真から選択するための方法は，大きく分けて以下の2つが考えられる．

方法(1) 生成する顔画像と顔の向きが近い1枚の顔写真のみを用いる

方法(2) 各パッチごとに，写真上でのパッチ面積が大きい方の顔写真を用いる

方法(1)では1枚の画像生成には1種類の顔写真しか使用しないが，生成する顔の向きによって2種類の顔写真を切替えて使用する方法である．それに対し，方法(2)では1枚の画像生成時にも2種類の顔写真を用いて，各パッチごとにテクスチャを切り出す顔写真を切替える方法である．

しかし，2種類の顔写真からテクスチャを切り出して画像を生成する場合，照明条件の差などにより，テクスチャの不連続性が目立つ場合がある．そのため，本研究ではテクスチャの連続性を保つ方法を提案する． まず，テクスチャの連続性は次の2つに大別できる．

画像内連続性 隣接するパッチ間でのテクスチャの連続性．この連続性が保たれていないと，生成された画像は図3に例示するように，「つぎはぎ」が目立ち、不自然になる．

画像間連続性 少しずつ顔の向きを変化させて生成した動画像における画像間でのテクスチャの時間軸上での連続性．この連続性が保たれていないと，動画像上でテクスチャが突然に変化し，不自然になる．

テクスチャの選択方法(1)では，全てのテクスチャを同じ顔写真から切り出すため，画像内連続性は保たれるが，顔の向きを変化させるにつれて，ある向きを境に全てのパッチで一斉にテクスチャが変化し，画像間連続性は保たれない．逆に方法(2)では，パッチごとに異なる顔写真からテクスチャを切り出すため，画像内連続性は保たれないが，各パッチに張り付けるテクスチャは顔の向きに依らず一定であり，画像間連続性は保たれるという性質がある．これをまとめると表1のようになる．

また，各3角形パッチにマッピングされるテクスチャについては，方法(2)の方がテクスチャを切り出すパッチの面積が大きいため，テクスチャマッピングによる画像の劣化が少ないという利点がある．そこで本研究では方法(2)を改良し，画像内連続性も画像間連続性も保たれた画像を生成する手法を提案する．

方法(2)の画像内不連続は，主に以下の原因で起こると考えられる．

(a) 曲面を粗い平面で近似したために，パッチ上のテクスチャに歪みが生じ，隣接するパッチの境界で画素の対応がずれる

(b) 2枚の顔写真の照明条件の差により，テクスチャの明るさが異なる

(a)については境界線に沿って濃度断面の弾性マッチングを行い，適応的にパッチを再分割することで対応ずれの補正をする．(b)については，異なる顔写真からテクスチャを切り出したパッチが隣接する場所を探し，この境界のテクスチャが連続的につながるように濃度を補正する．これらの詳細については次節で述べる．

表1　テクスチャの連続性
Table 1　Texture continuity

	画像内連続性	画像間連続性
方法 (1)	保たれる	保たれない
方法 (2)	保たれない	保たれる

図3　つぎはぎの起こっている例
Fig. 3　Example of texture discontinuity between patches

4.3　適応的パッチ再分割

隣接するパッチ間の，画素の対応ずれを補正するために，本研究では図4に示すように，3角形パッチの境界の濃度断面の弾性マッチングを行い，対応点をもとにパッチを再分割してずれを補正する．このパッチ再分割の手順を以下に示す．

1. 隣接する2つのパッチの境界線上に沿って画素値を取り出し，一対の1次元波形を得る
2. 双方の波形上でエッジにはさまれた区間を対応単位とし，動的計画法（DP）により区間対応探索[8] を行う
3. 求められた対応区間ごとにパッチを再分割し，対応区間が一致するようにパッチを伸縮する

このようにテクスチャの不連続が激しい部分だけを抽出し，パッチの分割数を状況に応じて適応的に増加させることで，粗いパッチモデルであっても歪みの目立たない自然な画像の生成が可能になった．

4.4　濃度の補正

隣接するパッチ間で，全体的に明暗が不連続になる場合がある．これは2枚の顔写真の照明条件の差が主な原因である．このテクスチャの不連続をなくすために，本研究では隣接するパッチ間の境界での濃度値を連続的につながるように補正する．

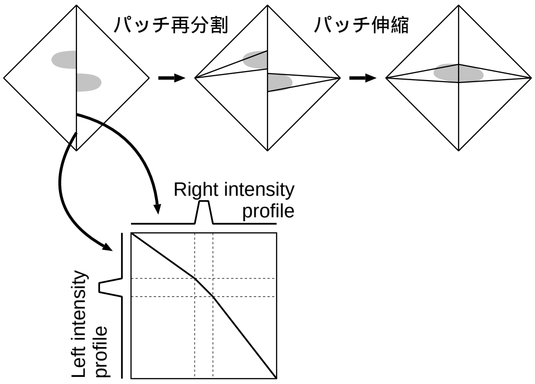


図4　動的計画法を用いたパッチ境界の整合とパッチ再分割
Fig. 4　Alignment of patch boundaries using dynamic programming and additional division of patches

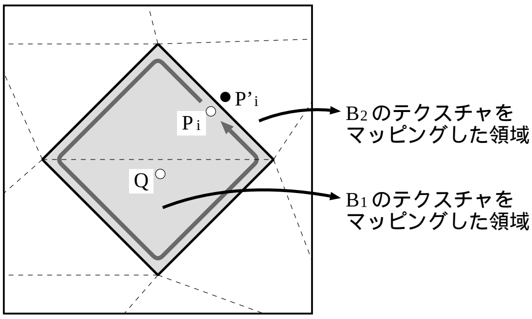


図5　パッチ内における濃度の補正
Fig. 5　Intensity adjustment within patches

図5に示すように，2枚の顔写真 B_1, B_2 からテクスチャを切り出して3角形パッチにマッピングした結果，生成された顔画像では顔写真 B_1 のテクスチャがマッピングされている3角形パッチと，B_2 のテクスチャがマッピングされている3角形パッチが存在する．ここで，同一の顔写真から得られたテクスチャが張られている隣接した3角形パッチの集合を「領域」と呼ぶ．テクスチャの不連続が問題となるのは，異なる顔写真から得られたテクスチャを持つ領域が接する境界である．

領域の境界上の画素については，領域をはさんで隣接する画素 P_i と P_i' の濃度をその平均値で置き換える．P_i と P_i' の濃度値をそれぞれ $I(P_i)$, $I(P_i')$ とすると，$I(P_i)$ に式 (11) のように補正値を加えることで不連続を解消する．画素 P_i に対する補正値 d_{P_i} は，式 (12) のように求められる．境界上の画素それぞれについて濃度の補正値を求める．

$$I(P_i) \leftarrow I(P_i) + d_{P_i} \tag{11}$$

$$d_{P_i} = \frac{I(P_i') - I(P_i)}{2} \tag{12}$$

領域内部の画素 Q に対する補正値 d_Q は，式 (13) のように領域の境界における補正値の重み付け平均とし

て求める．ただし w_i は画素 Q と P_i の距離に反比例した重み係数であり，$\sum w_i = 1$ とする．このようにして，それぞれの領域中の全ての画素に対する補正値を求め，濃度を補正する．

$$d_Q = \sum_i w_i d_{P_i} \tag{13}$$

この手法を用いることによって，光源の位置や頭部の立体構造が未知の場合でも，テクスチャの不連続がない自然な顔画像の生成が可能になった．

5. 実験結果

上記の手法によって顔画像を生成する実験を行った．実験では，顔の向きを示す指標画像を用い，基底画像の人物を指標画像の人物と同じ顔の向きにした場合の見え方を生成した．手順は，図**6**に示すように，まず指標画像から代表的な特徴点を抽出し，これをもとに線形結合の係数を求めて全ての特徴点座標を算出した．さらに，これらの特徴点を頂点とする3角形パッチモデルに2枚の顔写真から得られるテクスチャをマッピングし，最後にテクスチャの不連続の補正を行った．

なお，基底画像としては，2枚の顔写真のうち一方を，画像の左上の角を中心として時計周りに90° 回転させた画像と，x, y 方向にそれぞれ100画素平行移動した画像の合計4枚を用いた．

5.1　合成画像による実験

頭部の3次元形状と表面色を同時に計測できるレンジファインダ[1)] を用いて得た3次元形状モデルを用いることで，指定した角度に頭部を回転した場合の見え方を正確に合成できる．これを正解画像として，本手法で生成した顔画像との比較実験を行った．

まず，図**7**(a)(b) に示す2枚の顔画像を3次元形状モデルを用いて合成し，基底画像とした．(a) は正面向きから右に20° 上に20° の順に回転させた見え方であり，(b) は正面向きから左に20° 下に20° の順に回転させた見え方である．合成された2枚の顔画像上における45点の特徴点の2次元座標も，3次元座標の射影により算出した．またこれらの特徴点を頂点とする80面のパッチモデルをそれぞれ図7(c)(d) に示す．

図**8**上段は3次元形状モデルによって合成した正解画像である．これらの個々の画像を指標画像とし，目尻，口の両端，鼻の先端部の計5点の特徴点の2次元座標を用いて線形結合の係数を決定して，2枚の基底画像から生成した画像を図8下段に示す．

曲面を比較的大きな平面で近似しているために，顔表面上のテクスチャに若干のずれが生じているが，視覚的には満足できる画像が得られている．

5.2　実画像による実験

実画像を用いて，顔の向きを示す指標画像と同じ向きの顔画像を生成する実験を行った．図**9**(a)(b) に示す2枚の顔写真を基底画像とし，45点の特徴点の対応付けを手入力により与えた．頬などに配置した位置決めの困難な特徴点は，撮影時に顔にマーキングしておくことで位置決めを行った．

図9(c) に示す正面顔を指標画像として，目尻，口の両端，鼻の先端部の計5点の2次元座標を抽出し，線形結合によって45点の特徴点座標を求めて，テクスチャをマッピングした結果を図**10**に示す．(a) は顔の向きが近い方の顔写真のテクスチャのみを用いた結果，(b) はパッチの面積が大きい方の顔写真のテクスチャを用いた結果，(c) は (b) のテクスチャの不連続を補正した結果である．(a) では面積の小さなテクスチャを無理に引き延ばしているパッチがあり，例えば口の端などに歪みが生じている．また，前述したように動画像には向かない．(b) ではテクスチャの不連続が見られる．それに対し，(c) は自然な画像となっている． 通常の撮影環境では，2枚の顔写真で照明条件を完全に同一にすることは困難であるため，自然な画像を生成するためにはこのようなテクスチャの補正が必須であると言える．

様々な方向の指標画像を与えて顔画像を生成した結果を図11に示す．指標画像の人物は基底画像の人物と別人であり，代表的な特徴点として与えた5点の配置は指標画像と基底画像とで一致するものではないが，顔の向きを指定することは十分に可能であることが分かる．

6. むすび

本稿では，撮影方向が未知の2枚の顔写真のみを入力として，3次元構造を復元することなく任意方向の見え方を生成する手法を提案し，実験によりその有効性を確認した．

今回は人間の顔を剛体と見なしてシステムを構築したが，今後は生成する見え方に表情変化を持たせるために，様々な表情を持った顔写真を組み合わせる手法について検討する．

謝辞　本実験で生成結果の評価のために使用したレンジファインダによる3次元形状モデルは，ＮＴＴヒューマンインターフェース研究所で計測させて頂いたものである．

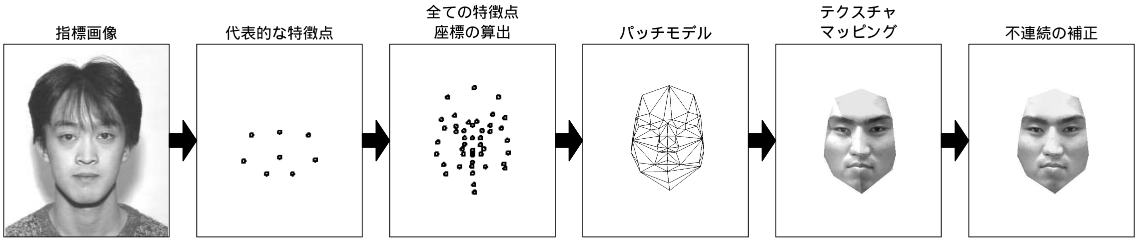


図6　顔画像生成処理の流れ
Fig. 6　Flow of the view generation process

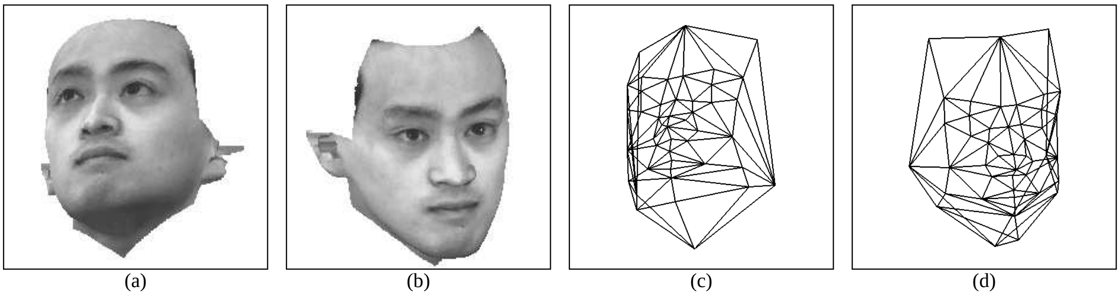


図7　入力顔画像 (a)(b) と対応するパッチ (c)(d)
Fig. 7　Input face images(a)(b) and corresponding patches(c)(d)

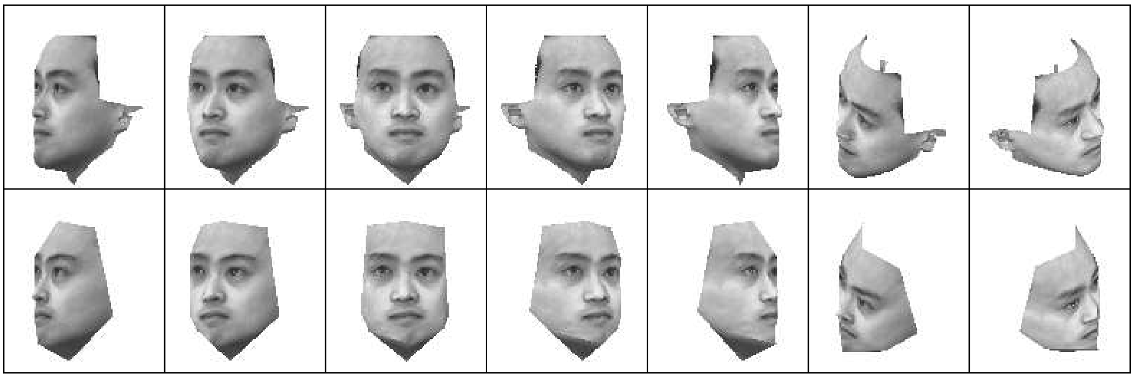


図8　生成画像の比較（上段：距離データから生成した画像，下段：２枚の顔写真から生成した画像）
Fig. 8　Comparison of synthesized images (upper row: images generated from range data, lower row: images synthesized by our method)

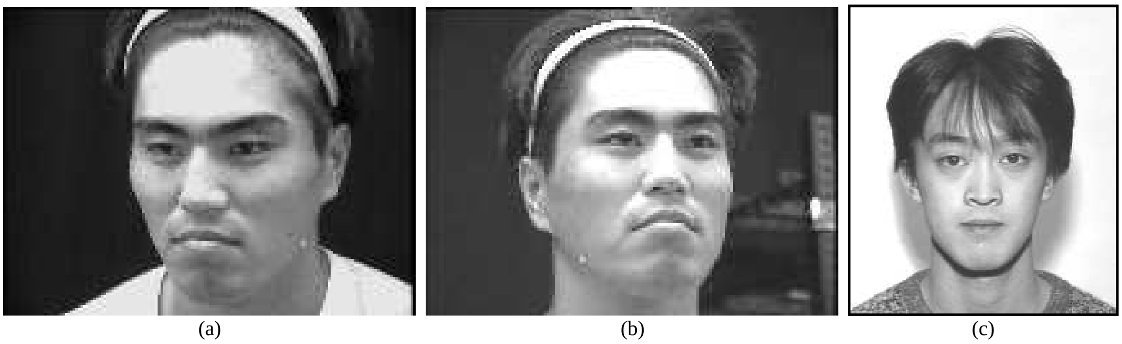
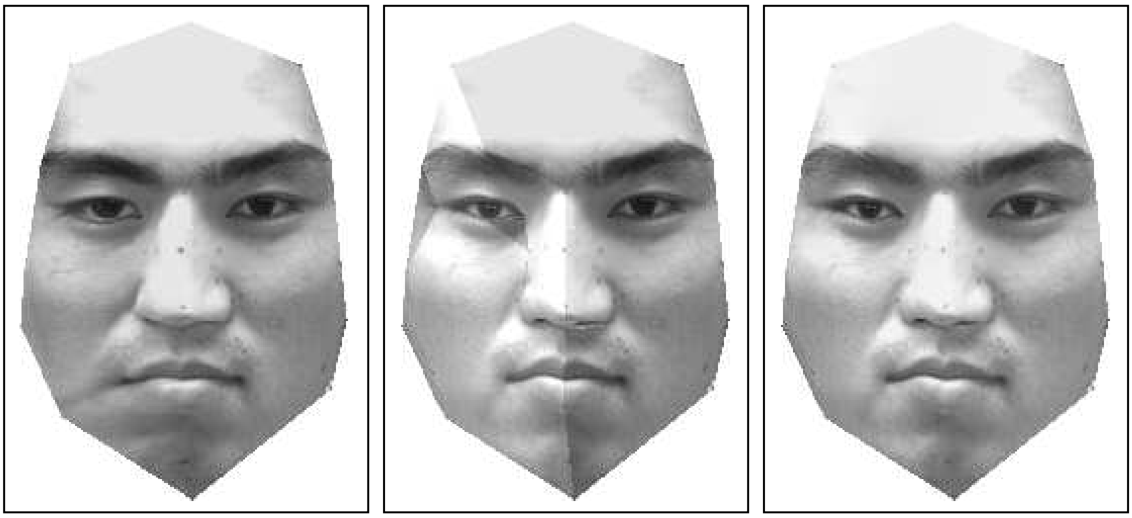


図9　入力画像 (a)(b) と指標画像 (c)
Fig. 9　Two input images(a)(b) and an image specifying facial orientation(c)



(a) Texture taken from one image

(b) Texture taken from two images

(c) Result of revising texture discontinuity

図10　生成された画像
Fig. 10　Synthesized images

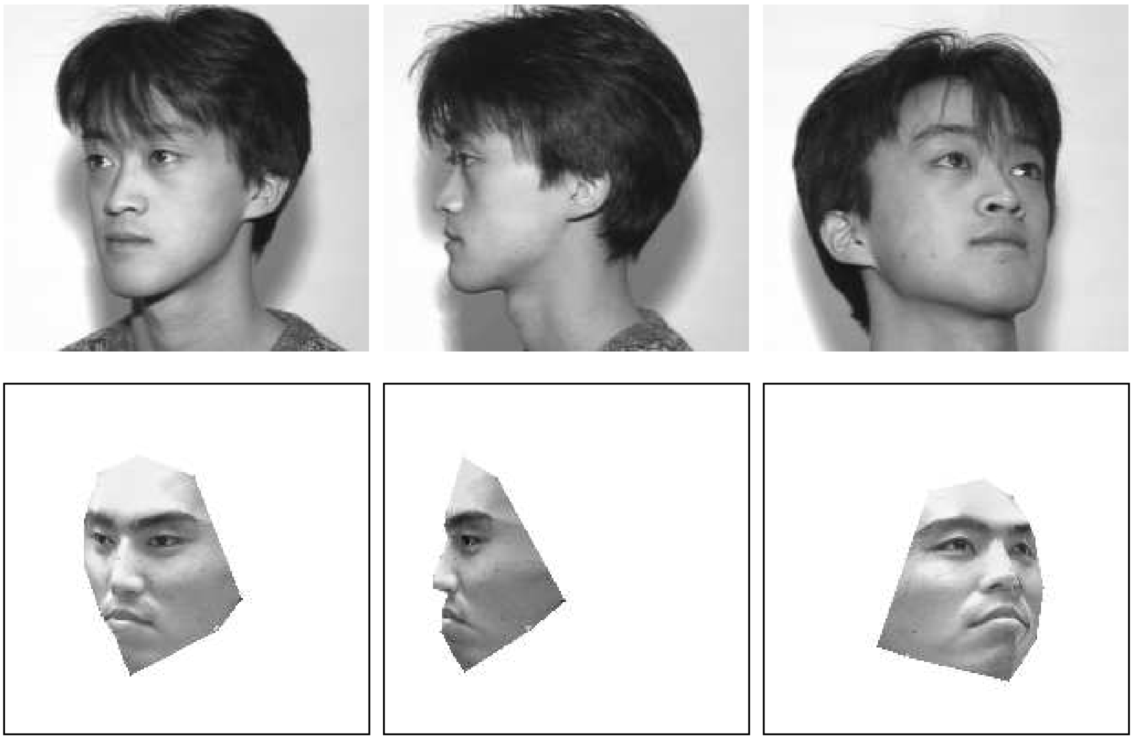


図11　様々な方向の顔画像の生成（上：指標画像，下：生成画像）
Fig. 11　Synthesis of various facial views
(top: images specifying facial orientation, bottom: synthesized views)

参考文献

1) 末永康仁, 渡部保日児: 3D形状と輝度(色)の同時計測が可能なスキャナとその顔画像入力への応用, 情処研報, CV 67-5 (1990).
2) 秋本高明, 末永康仁: 頭部の正面・側面像と基本3次元形状を用いた顔画像生成手法, 信学技報, PRU 88-47 (1988).
3) R.Skerjanc, and J.Liu: Computation of Intermediate Views for 3DTV, *Proc. 5th Workshop 1992 on Theoretical Foundations of Computer Vision*, pp. 190–201 (1992).
4) S.Ullman: The Interpretations of Visual Motion, MIT Press, Cambridge, MA (1979).
5) C.Tomasi, T.Kanade: The Factorization Method for the Recovery of Shape and Motion from Image Streams, *In Proceedings of Image Understanding Workshop*, pp. 459–472 (1992).
6) J.J.Koenderink, A.J.van Doorn: Affine structure from motion, J.Opt.Soc.Am.A, Vol. 8, No. 2, pp. 377–385 (1991).
7) S.Ullman, R.Basri: Recognition by Linear Combinations of Models, *IEEE Trans. PAMI*, Vol. 13, No. 10, pp. 992–1006 (1991).
8) 大田友一, 正井康之, 池田克夫: 動的計画法によるステレオ画像の区間対応法, 信学論, Vol. J68 No. 4, pp. 554–561 (1985).

(平成7年8月8日受付)
(平成8年2月7日採録)

向川 康博(学生会員)

平成4年筑波大学第3学群情報学類卒業.現在同大大学院博士課程工学研究科在学中.工学修士.コンピュータビジョンの研究に従事.

中村 裕一(正会員)

昭和60年京都大学工学部電気工学第二学科卒業.平成2年同大大学院博士課程修了.同年京都大学工学部助手.平成5年筑波大学電子・情報工学系講師.工学博士.画像理解,自然言語とパターン情報の統合に関する研究に従事.人工知能学会,電子情報通信学会各会員.

大田 友一(正会員)

昭和47年京都大学工学部電子工学科卒業.昭和52年同大大学院博士課程修了.同年学振奨励研究員.昭和53年京都大学情報工学科助手.昭和56年筑波大学電子・情報工学系講師.昭和57 ~ 58年カーネギーメロン大学計算機科学科客員研究員.昭和62年筑波大学電子・情報工学系助教授.平成4年同教授.工学博士.画像処理,コンピュータビジョンの研究に従事.人工知能学会,電子情報通信学会, IEEE各会員.