

複数の画像を用いた頭部の形状推定とレンダリング

栗原恒弥 新井清志
日立製作所 中央研究所

複数の画像から人間頭部の3次元形状を推定して、画像を生成する方法を提案する。基本的な頭部の形状モデル（基本モデル）を用意し、これを入力された画像に適合するように変形する。具体的には、画像に対して特徴点の位置を指定し、これから特徴点の三次元座標を求める。得られた特徴点に適合するように基本モデルを変形する。次に複数の画像から得られるテクスチャ情報を精度を考慮して合成する。得られた形状モデルとテクスチャ情報を用いて、現実感のある画像を容易に生成できる。

Facial modeling and rendering using multiple images

Tsuneya Kurihara Kiyoshi Arai
Central Research Laboratory, Hitachi, Ltd.
Kokubunji Tokyo 185, Japan

This paper describes a method which reconstructs shape from 2D multiple images. In this method, a 3D canonical model of the human face is prepared, and it is transformed to be consistent with the images. Two dimensional coordinates of feature points are specified for multiple images, and three dimensional coordinates of feature points are estimated. The canonical model is transformed so that coordinates of its feature points coincide with those of the estimated feature points. In order to estimate its texture, images are composited taking into account its precision. Realistic images can be generated using reconstructed shape and texture.

1. はじめに

コンピュータグラフィックス（以下CGと略す）における重要な課題として人物の表現が残されている。CGで人物を表現できれば、ヒューマン・インタフェース、教育、プレゼンテーション、娯楽等への応用が考えられる。人物の表現の中でも頭部は非常に多くの情報を含んでおり、特に重要であると考えられる。

しかし、頭部の形状は複雑であり、データの入力に多くの工数がかかるという問題がある。現在、形状の入力方法としては、（1）3次元デジタイザを用いる方法、（2）対話的な入力（CAD等）の2つの方法がある。3次元デジタイザを用いる方法は特殊な装置が必要である[1, 2]。対話による方法も形状をユーザが逐一入力する必要があり、頭部のように複雑な形状の入力には多くの工数が必要となっている。

さらに肌や毛髪は非常に微細で複雑な形状を有しており、これらをリアリストックに表現することは困難である。肌に関して、現実感のある画像を生成する方法としては、現在テクスチャマッピング[3, 4, 5]が有効な方法である。しかし、通常、一枚の画像をテクスチャとしてマッピングすることが多いため、全方向から見た画像を生成することができない。

知的画像符号化の分野では、正面画像を用いて、顔の形状を推定し、テクスチャマッピングを用いて顔画像を生成している[6, 7]。このため、すべての方向から見て自然な画像を生成することが困難である。

本報告で提案する方法は、複数枚の画像から頭部の形状を推定するものである。この方法によれば、実在人物の頭部の形状入力が容易に実現できる。また特別な装置も不用である。架空の人物の頭部の形状を入力する場合には、そのスケッチ等を用意すればよい。

複数枚の画像を用いて、形状を推定する方法として、文献8、9の方法が提案されている。二方法ともに、頭部の正面画像と側面画像を用いてい

る。提案する方法は原理的には、画像は必ずしも正面・側面を向いている必要はない。文献8では4次のBezierパッチを用いて頭部の形状を生成し滑らかな形状を得ているが、処理が複雑となる。本方法では多面体モデルを用いており、形状の推定処理も単純かつ高速である。

得られた形状モデルにテクスチャマッピングを施して画像を生成する。全方向からの画像を生成するために、複数枚の画像を精度を考慮して合成してテクスチャを求める。

以上により、画像から頭部の形状モデルの推定及びレンダリングが実現される。簡単な例によってこの方法の有効性を示す。

2. 複数の画像からの頭部の形状の推定

頭部の形状は個人差はあるものの、おおまかには類似している。このため、頭部の基本的な形状を用意し、これを与えられた複数の画像に適合するように変形して頭部の形状を得る。

形状の推定は基本モデル上に指定された特徴点を用いて行なう。与えられた画像上での特徴点の位置を指定し、三角測量によって、特徴点の三次元座標を求める。得られた特徴点の三次元座標に適合するように、基本モデルを変形して、画像に適合する頭部の形状を得る。

2.1 基本モデル

人間の頭部の構造に関する知識として、基本的な形状モデル（以下、基本モデルと呼ぶ）を用意する。これは対話的に入力するかあるいは3次元デジタイザを用いて測定する。実験で用いた基本モデルを図1に示す。この基本モデルはマネキンの頭部を3次元デジタイザで測定して得た約5万8千点のデータから、ラブラシアンフィルタとドロネ網を用いて、約2700個の三角形で構成したものである[10]。頭部において、目や鼻は特に多くの情報を含んでおり、形状の推定において重要である。このため特徴的な点（40個）

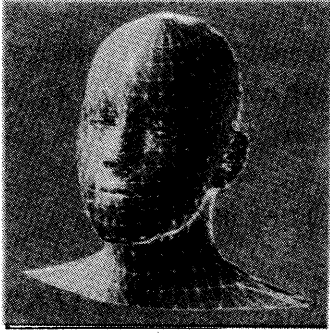


図1 頭部の基本モデル

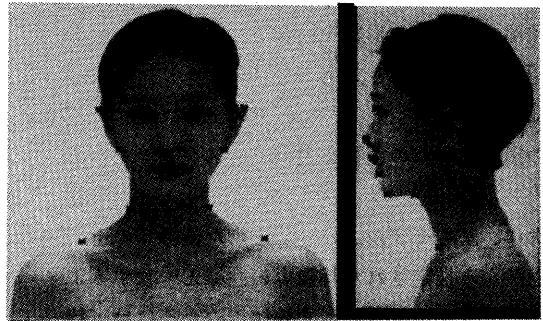


図3 特徴点の指定

を特徴点としてあらかじめ登録しておく。

2. 2 頭部画像の撮影

モデル化したい頭部を複数の方向から撮影する。撮影においては、照明条件ができるだけ一定であるようにする。また、画像の歪みを少なくするために十分に距離をおいて撮影する。撮影された画像は正射影されたものとして扱う。画像の例を図2に示す。これは頭部を前後左右から撮影したものである。頭部の形状推定には2枚の画像（正面・側面像）を用いる。他の2枚の画像はテクスチャマッピングに使用する。

2. 3 特徴点の指定

頭部の形状を複数の画像から推定する。このために、各画像に対して特徴点を指定する。特徴点の自動的な認識は困難であるため、特徴点の指定は対話的に行っている。特徴点の指定の例を図3に示す。

2. 4 特徴点の三次元座標の推定

複数の画像に対して指定された特徴点の位置から、特徴点の三次元座標を推定する。推定には三角測量を用いる。特徴点には、複数の画像上で位置が指定されているものと一枚の画像上でのみ位置が指定されているものがある。後者は三次元座標を推定することができない。このため、後述するように、まず三次元座標が推定できる特徴点を用いて、形状のおおまかな推定を行う。その後、推定された形状モデル上の位置と画像上の位置を用いて、一枚の画像だけで指定された特徴点の三次元座標を推定する。新たに推定された特徴点の座標を用いて再度、形状の推定処理を行なう。

なお、座標変換のパラメータが未知の場合には、基本モデル上の特徴点の座標と画像上の特徴点の座標を用いて、推定することが可能である [11]。



図2 入力画像

3.5 フィッティング

頭部の大局的な形状を適合させるために、頭部の幅、高さ、奥行き方向ごとにスケーリングする。頭部の中心を原点として、頭部の上下、前後、左右の端の特徴点の座標を用いて、スケーリングする。

次に、基本モデル上に与えられた特徴点の座標が画像から推定された特徴点の座標と一致するように、基本モデルを変形する。このためには、特徴点で与えられる変位を用いて、基本モデル上の各頂点での変位を求める必要がある。

ここでは、頭部の表面を2次元のパラメータ空間に対応付け、パラメータ空間での線形内挿によって、各頂点での変位を求めた。この方法は文献12で表情の制御に用いた内挿方法と同一である。頭部の形状は三次元空間上で定義されているが、表面を二次元空間と対応付けることによって、特徴点の変位の内挿問題は、二次元的な内挿問題に帰着できる。二次元のパラメータ空間を特徴点で構成される三角形に分割し、各頂点での変位は、その点が含まれる三角形の三頂点での変位を線形内挿して求める。最適な線形内挿を実現するために、三角形分割の方法として、ドロネ網 [13] を用いる。ドロネ網は、二次元空間上に不規則に分布している点を接続して、二次元空間を三角形領域に分割するものである。このとき個々の三角形はできるだけ正三角形に近くなるように分割される。二次元のパラメータ空間でのドロネ網を用いた線形内挿によって、分布が一様でない特徴点で与えられる変位を、計算に破綻を生ずることなく、高速に内挿することが可能である。

頭部の表面を二次元のパラメータ空間に対応付ける方法には特に制限はない。ここでは、円柱座標系での高さ成分と角度成分を用いた。すなわち、頭部の表面上の点を $A(x, y, z)$ とすると、二次元のパラメータ空間での座標 $A_s(\theta, y)$ は

$$(\theta, y) = (\tan^{-1}(z/x), y) \quad (1)$$

となる。

基本モデル上に定義されている特徴点を P とし、画像から推定された特徴点を Q とする。このとき点 P の変位 V は、

$$V = Q - P \quad (2)$$

となる。(1)式を用いて、 P をパラメータ空間に変換する。変換された点を、 $P_s(\theta, y)$ とする。 $P_s(\theta, y)$ を用いて、パラメータ空間を三角形分割する。三角形分割には前述のようにドロネ網を用いる。

基本モデル上の各頂点の変位は次のようにして求める。変位を求める頂点を $A(x, y, z)$ とし、この点に対応するパラメータ空間上での点を $A_s(\theta, y)$ とする。

パラメータ空間上で A_s が含まれる三角形を求め、三頂点での変位を線形内挿する。 A_s が含まれる三角形の三頂点を、 $C_1(\theta_1, y_1)$ 、 $C_2(\theta_2, y_2)$ 、 $C_3(\theta_3, y_3)$ とし、これらの点における変位を $V_1(\Delta x_1, \Delta y_1, \Delta z_1)$ 、 $V_2(\Delta x_2, \Delta y_2, \Delta z_2)$ 、 $V_3(\Delta x_3, \Delta y_3, \Delta z_3)$ とする。このとき、 $A_s(\theta, y)$ の重心座標、 k_1, k_2, k_3 ($k_1 + k_2 + k_3 = 1$) が次式より一意に求められる。

$$(\theta, y) = \left(\sum_{i=1}^3 k_i \cdot \theta_i, \sum_{i=1}^3 k_i \cdot y_i \right) \quad (3)$$

これを用いて、頂点 A における変位 V を求める。

$$V = \sum_{i=1}^3 k_i \cdot V_i \quad (4)$$

なお、特徴点には、複数の画像に指定されている特徴点と一枚の画像だけに指定されているとが存在する。前者は三次元座標を推定することが可能であるが、後者は三次元座標を推定することができない。このため、まず前者だけを用いて大まかなフィッティングを行い、その後、後者を用い

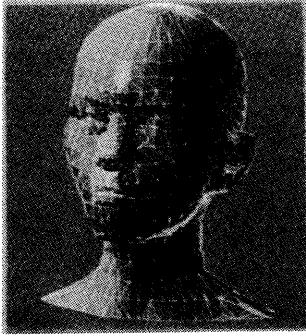


図4 推定された形状

て細部のフィッティングを行う。三次元座標が推定できる特徴点だけを用いて、推定された近似形状 S は求めたい形状を十分に近似しているとする。一枚の画像だけに指定されている特徴点 P については、近似形状 S 上の対応する特徴点 P' の座標と画像上の特徴点の位置から三次元座標を推定する。すなわち、 P' を通り、射影方向と垂直な面上で画像上の特徴点の位置と適合する点を求め、これを特徴点の座標の推定値とする。新たに三次元座標が推定された特徴点を加えて、再度形状の推定処理を行なって、最終的な推定形状とする。

画像上の特徴点の指定とフィッティングは、すべての特徴点を指定した後に行なう必要ない。特徴点の三次元座標の推定と、形状の推定は高速に実現できるため、形状を推定する上で重要と思われる点から順次指定していけばよい。また、画像情報だけでは推定できない凹凸等の形状も、濃淡等の情報からユーザは推定することが可能である。このように推定された形状に適合するように、特徴点を対話的に設定して、形状を修正することも可能である。

図4に推定された形状を示す。あらかじめ設定された40個の特徴点と対話的に指定した18個の特徴点から形状を推定した。

なお、この方法は頭部形状の対称性を仮定することによって、一枚の画像からの形状推定にも適用できる[11]。頭部が左右対称であるとの仮定から、例えば左斜めの方向から撮影した画像を

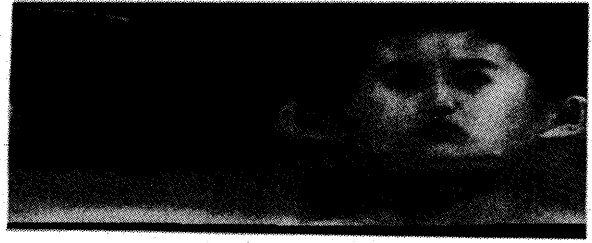


図5 合成されたテクスチャ

左右に反転すれば右斜めの方向から撮影した画像と等価になり、本方法を用いて頭部の形状を推定できる。

3. 画像生成

推定された形状に対して、与えられた画像をテクスチャマッピングして画像を生成する。全方向からの画像を生成するためには、複数の画像から得られるテクスチャを合成する必要がある。

二次元画像をテクスチャマッピングする場合には、物体表面の単位面積を画像上に投影したときの面積によって、投影されるテクスチャの精度が変化する。各画像の解像度が同一であれば、精度は表面の法線方向 N と視線方向 V との内積によって求められる。すなわち、表面の法線方向と画像の撮影方向のなす角が大きいほど、テクスチャの拡大率は大きくなり、精度は落ちる。この精度を考慮して、テクスチャの合成を行なう。

複数のテクスチャを合成するために、テクスチャは円柱座標系（あるいは極座標系）で表現する。各画像から得られる輝度情報と精度を円柱座標系に変換する。これらを精度を重みとして、合成する。画像 j から得られた輝度値を $C_j(\theta, y)$ 、精度を $W_j(\theta, y)$ とする。このとき、次式を用いてテクスチャを合成する。

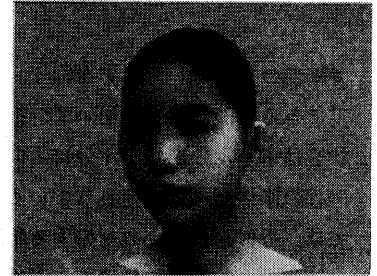
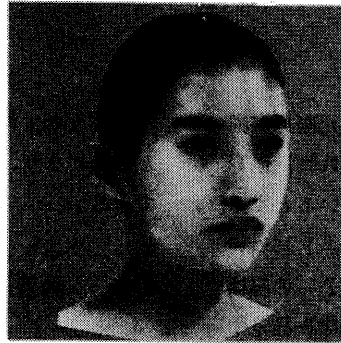
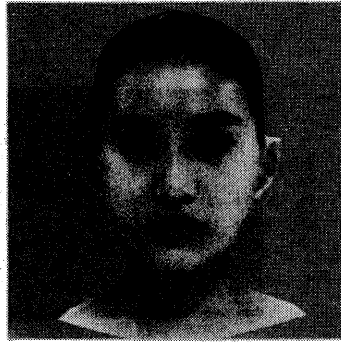


図6 生成された画像

図7 表情の生成例

$$C(\theta, y) = \frac{\sum C_j(\theta, y) \cdot f(W_j(\theta, y))}{\sum f(W_j(\theta, y))} \quad (5)$$

ここで、 $f(w)$ は、精度の低いテクスチャを無視するための関数であり、次のようなものを用いた。

$$\begin{aligned} f(W) &= (W-h) / (1-h) & (W \geq h) \\ f(W) &= 0 & (W < h) \end{aligned} \quad (6)$$

h は適当な定数であり、精度が h より低いテクスチャを無視する。ここでは $h = 0.5$ とした。

また凹凸が激しい部分では精度の変化が大きくなり、合成されるテクスチャが不連続になりやすい。このため、精度を求めるための法線は形状全体で滑らかになるように修正している。実験では、

頭部の形状を円柱で近似して、精度を推定するための法線を求めた。

以上で各テクスチャが合成される。合成されたテクスチャを用いて現実感のある画像を生成することが可能である。

4. 結果

(1) 4方向からの画像による生成例

図5に円柱座標系で合成されたテクスチャを示す。4方向からのテクスチャが滑らかに合成されていることがわかる。図6にテクスチャマッピングによって得られた画像を示す。図6に示したように、全方向からの人物画像が生成できる。

さらに、このモデルに対して、文献12の方法で表情を生成した表示例を図7に示す。目、鼻等がほぼ正確に適合しているため、自然な画像が生

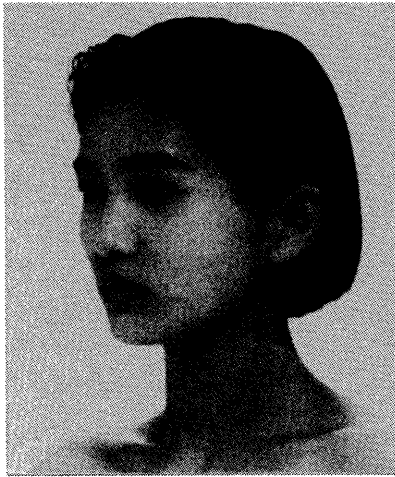


図8 入力画像

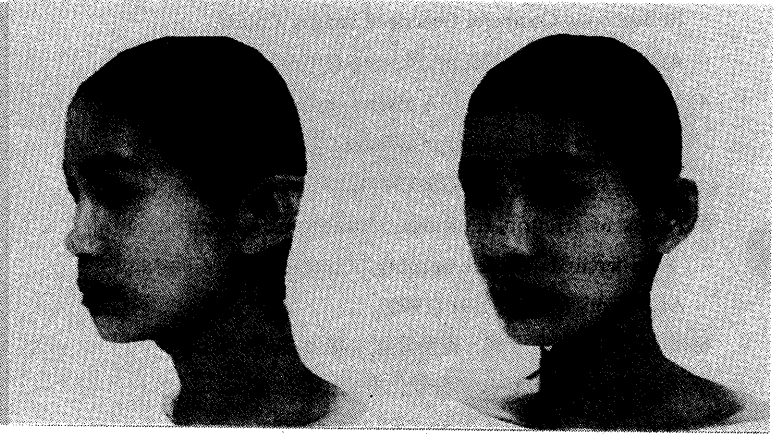
成されている。

(2) 一枚の画像を用いた生成例

前述のように、本方法は頭部の対称性を仮定することによって、一枚の画像からの形状推定にも適用できる。図8に入力画像を、図9に頭部の向きを15°変更した合成画像を示す。形状の推定には19個の特徴点を用いた。図9(a)において、耳の後ろの部分に耳のテクスチャがマッピングされている。この部分は入力画像では見えなかった部分である。この問題を解決するためには、マッピングするテクスチャを増やすか、境界部分の輝度を用いて、新たに見えるようになった部分の輝度を推定する必要がある。

写真にはすでに陰影がついているため、生成された画像は陰影のついたものとなっている。特に首の下の部分に影の領域がある。このため、光源の位置を変更した画像を生成するためには、写真撮影時の陰影を除去し、表面の反射率を推定する必要があると考えられる。

頭髪に関しては、モデル化していないため、不自然な印象を受ける。テクスチャマッピングによる頭髪の表現には限界がある。頭髪の表現は今後



(a)

(b)

図9 向きを変更した合成画像

の課題である。

5. おわりに

複数の画像と頭部の基本モデルを用いて、頭部のモデルを構成する方法について述べた。形状の生成には特徴点を用いた。複数の画像から得られるテクスチャ情報を精度を考慮して合成して、マッピングすることにより、現実的な画像を生成することができる。今後は、写真から生成したテクスチャから輝度を補正して反射率を推定し、光源の位置を変更できるように改良していく予定である。

参考文献

- 1) Magnenat-Thalmann, N. and Thalmann D., "The Direction of Synthetic Actors in the Film Rendezvous a Montreal," IEEE Computer Graphics and Applications, Vol. 7, No. 12, pp. 9-12 (1987).
- 2) Nahas, M., Huitric, H. and Saintouens, M., "Animation of a B-Spline figure," The Visual Computer, Vol. 3, No. 5, pp.272-276 (1987).
- 3) Blinn, J. F. and Newell, M. R., "Texture and

- Reflection in Computer Generated Images, Comm. ACM, Vol. 19, No. 10, pp.542-546 (1976).
- 4) Yau, J.F.S, Duffy N.D., "A texture mapping approach to 3D facial image synthesis," Computer Graphics Forum, 7, pp. 129-134 (1988).
 - 5) Nahas, M., Huitric, H., Rioux, M. and Domey, J., "Registered 3D-Texture Imaging," Computer Animation '90, pp. 81-91 (1990).
 - 6) 相沢清晴、原島 博、斉藤隆弘：構造モデルを用いた画像の分析合成符号化方式、信学論 (B-I), J72-B-I, 3, pp. 200-207 (1989).
 - 7) 金子正秀、羽島好律、小池 惇：形状変化の検出と3次元形状モデルに基づく顔動画の符号化、信学論 (B-I), J71-B, 12, pp. 1554-1563 (1988).
 - 8) 小松功児：アニメーション制作のための顔の曲面モデル、情処学論、30, 5, pp. 633-641 (1989).
 - 9) 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).
 - 10) 新井清志、栗原恒弥：ラプラシアンフィルタとドロネ網を用いた物体表面の再構成；90年春信学全大D-601 (1990).
 - 11) 小松功児：顔画像からの顔の3次元形状の推定とその表情の変更、信学論 (D-II), J73-D-II, 5, pp. 707-716 (1990).
 - 12) 新井清志、栗原恒弥：二次元パラメータ空間上での線形内挿による顔の三次元アニメーション；情処学会グラフィクスとCAD45-5(1990).
 - 13) 伊理正夫監修、腰塚武志編集：計算幾何学と地理情報処理；bit別冊、共立出版、第4章、pp. 126-148 (1986).