

3次元顔頭部への高速テクスチャマッピング

依岡 圭亮 岩切 裕哉 青木 公也 金子 豊久
豊橋技術科学大学 情報工学系

3次元顔頭部CTから得られた3次元形状モデルにデジタルカメラから得られた画像を用いて高速にテクスチャをマッピングする手法を提案する。既に我々は、写真画像を3次元顔頭部にマッピングするのにソフトウェアのみで処理をする手法を提案している。しかし、この方法では6時間の処理時間を要し実用性に乏しい。そこで、ハードウェアを利用することや、マッチング回数を減らす手法を用いて処理時間を大幅に短縮する手法を提案する。

High-speed Texture Mapping on 3D Human Face

Keisuke Yorioka Yuya Iwakiri Kimiya Aoki Toyohisa Kaneko
Dept. of Information and Computer Sciences,
Toyohashi University of Technology

This paper is concerned with a fast computing method for mapping photographic images to a 3D model extracted from a CT image. Our previous method based upon software alone requires about six hours, which is not practical. Our new method which utilizes a new matching technique in addition to a graphic board with a geometry engine now requires about several minutes of computation time.

1. はじめに

本論文は、手術シミュレーションなどの応用を考えて、CT画像から得られた人間の顔頭部3次元形状モデルに、デジタルカメラにより得られた画像を高速かつ正確にマッピングする手法の提案を行う。

顔頭部のテクスチャを得る問題に対しては過去に多くの論文がある。秋元らは、2枚の写真からプレートマッチングによる顔頭部の作成を行った[1]。Yauらはデプスマップとカラー写真からテクスチャマッピングを行う手法を提案した[7]。Koch[5]らは、手術シミュレーションのために、患者のCT画像にテクスチャをマッピングする問題にCyberwareにより得られた3次元テクスチャを用いている。

我々は既に、3次元顔頭部モデルの2次元投影像からエッジを抽出し、3次元モデルに逆投影することにより3Dエッジを定義し、それを移動、回転させて写真との対応づけをとることによりカメラパラメータを求める手法を提案した[2],[3]。しかし、3Dエッジの抽出、写真とモデルとの対応づけをすべて

ソフトウェアで行っているため、一連の処理におよそ6時間もかかり実用性に乏しい。

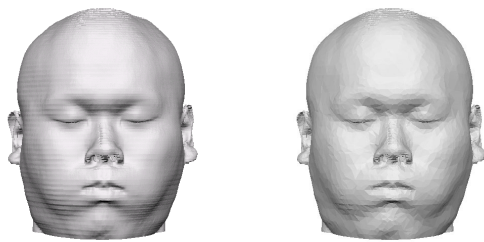
また我々は、高速に3次元形状のテクスチャを得る手法として、物体のシルエット画像によるマッチングを行う手法を提案した[4]。しかしこの手法は、顔頭部のような視点によるシルエット形状の変化が乏しいモデルへの適用は困難である。

本論文では、このような問題を解決するために以下のようなアプローチをとる。

1.1. アプローチ

使用するデータは顔頭部のCTと、正面、右、左から撮影した3枚の写真画像である。まず、CTデータからMarching Cubes法[6]を用いて3次元モデルの構築を行い(図1(a))、表面形状が変わらない程度にポリゴン数の低減(約3万)を行った3次元モデル(図1(b))を使用する。

写真画像を3次元モデルにテクスチャマッピングするには、写真画像を撮影した時の撮影位置と仮想カメラの位置を一致させ、写真画像に対応する3次元モデルの表面に逆投影することにより行う。



(a)元のモデル (b)ポリゴンリダクション後

図 1: 3次元モデル

本論文では、デジタルカメラを用いて撮影を行うため撮影位置は未知である。そのため、写真画像と3次元モデルのマッチング処理により、位置を推定する。マッチング処理にはエッジ間の最短距離にある点同士を対応点とするICP法[2]を用いる。

本論文は、以下のようなアプローチにより高速に3次元モデルと写真画像の対応付けを行う手法を提案する。

- 3次元モデルから顔表面の3Dエッジを得る。この処理はマッチング処理の前に1回だけ行う。
- 3次元モデルと3Dエッジは、ジオメトリエンジンを搭載したビデオカードを用いることにより高速に2次元投影される。
- 初期位置決定、移動と回転のマッチング、視野角のマッチングを独立に行うことによりマッチング回数を大幅に削減する。

2. 提案手法

本論文におけるテクスチャマッピングまでのプロセスは以下の通りである。

- (1) 写真画像, 3次元モデルからの投影像のエッジ抽出
- (2) エッジ同士のマッチング
- (3) テクスチャマッピング

以下に、それぞれの手法について述べる。

2.1. エッジ抽出

マッチングを行うために、顔の特徴のみを残したエッジ画像の作成法について述べる。

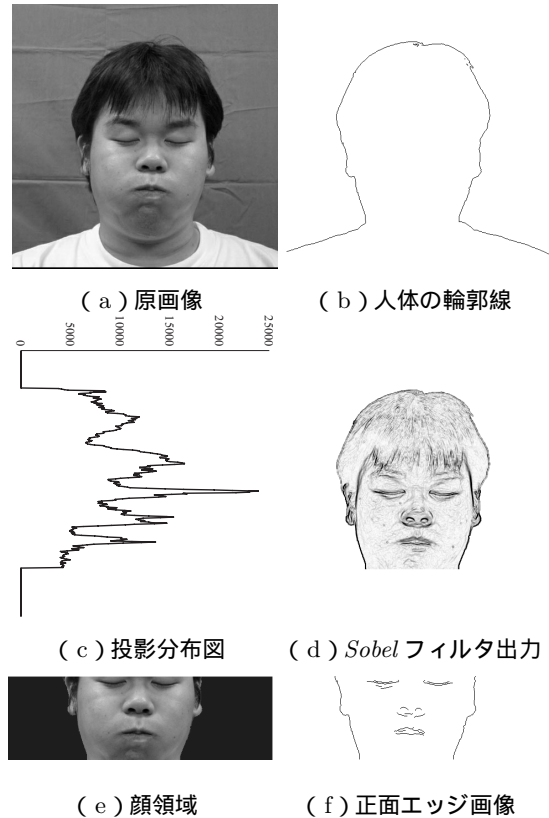


図 2: 正面顔画像

2.1.1. 写真画像のエッジ抽出

本論文では、耳を特徴点として用いるため、撮影画像には耳が露出している必要がある。また、青色のスクリーンを背景にして撮影し背景除去を行う。写真画像には髪、衣服など、3次元モデルには存在しない部分があるが、これらから抽出されるエッジはすべて除去しておく。以下にその手法について述べる。

(1) 正面顔のエッジ抽出

正面顔画像(図2(a))では、青画像でしきい値を用いて背景と人体とを分離し、輪郭線を抽出する(図2(b))。さらに輪郭が狭くなっている首部分を抽出し、首から下をカットした画像に対してSobelフィルタにより、エッジ画像を得る(図2(d))。そこから水平方向のエッジ強度の投影分布(図2(c))を基に目を検出(ピーク点)した後、目から上をカットする(図2(e))[8]。最後に目、鼻、口などの顔表面のエッジを抽出し、2値化、細線化を経た画像が正面顔画像のエッジである(図2(f))。

(2) 横顔のエッジ抽出

横顔画像(図3(a))から正面顔と同様に輪郭線

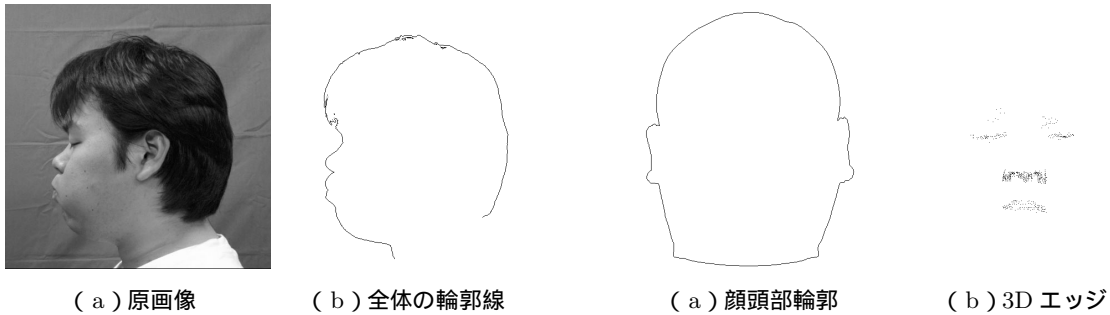


図 4: 顔頭部エッジ抽出

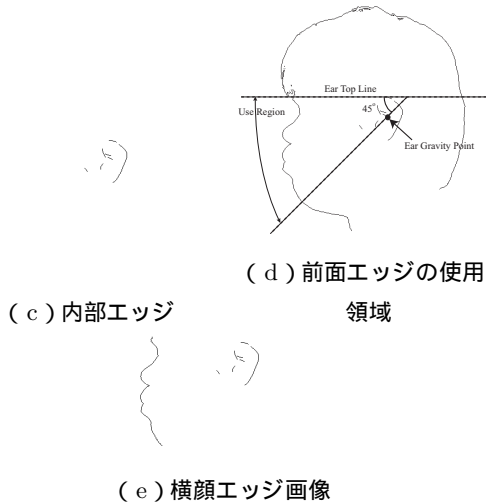


図 3: 横顔画像

を抽出した画像(図 3(b))を作成する。さらに図 3(b)の輪郭線の重心を中心とした半径 100[pixel]の内部のエッジを横顔画像から抽出し、2 値化、細線化する(図 3(c))。次に、図 3(d)のように耳の重心から 45 度の傾斜した線から耳の上側の領域で顔の前面のエッジを抜き出し、図 3(e)の横顔のエッジを作成する。

(3) 顔画像の部位ラベル付け

正面顔画像は左目、右目、鼻、口、左耳、右耳の 6 個、横顔画像は顔の前面、耳の 2 個の部位が存在する。これらのエッジにそれぞれ番号を割り振り、後述するマッチングの重み付けに使用する。

2.1.2. 3次元モデルのエッジ抽出

まず、実空間をほぼ再現するように仮想空間の照明を配置する。3次元モデルは、ジオメトリエンジンが搭載されたビデオカードを用いることにより高速に 2次元投影される。この顔頭部の投影画像は移動、回転の影響で形状が変化する。エッジ検出を行うと多大の計算時間を要する。一方、顔頭部の輪郭を得る処理は高速に行える。そこで我々は、顔頭部

エッジを以下のステップで抽出することによりエッジ検出を行う回数を 1 回にする手法をとる。

(1) 顔頭部の輪郭作成

顔頭部画像の背景色とモデルの色の差で外輪郭を抽出すると図 4(a)が得られる。図 4(a)では、2D エッジに対応点のない頭部、首からのエッジが多く含まれているが、提案手法ではマッチングの初期位置はほぼ解に近い値に決められるため影響はない。

(2) 3D エッジ抽出

まず正面の 3D エッジは、顔頭部画像のうち、図 4(a)の内部領域に Sobel フィルタをかけ、2 値化、細線化した画像を得る。その画像をモデルに逆投影し、エッジピクセルが重なったポリゴンのみを点で描画する(図 4(b))。横画像からの 3D エッジは耳だけが必要であるため、顔頭部の輪郭画像の重心から半径 100[pixel]の円の内部で、見えるポリゴンと見えないポリゴンが隣接しているポリゴンを点で描画する。これらを 3D エッジと定義する。3D エッジは 3次元点集合であり、一度抽出すれば 3次元モデルと同様に移動、回転させることができる。

顔頭部の輪郭は視点から見て前面と背面の境界に存在し、視点にかなり依存するため、3D エッジを用いるのは困難である。

(3) エッジの合成

まず顔頭部モデルをレンダリングし、輪郭を抽出する。次にレンダリングするモデルを 3D エッジに切り替えて、内部領域を得る。これらをあわせたものが実際にマッチングに使われるエッジ(モデルエッジと呼ぶ)となる。

2.2. マッチングアルゴリズム

3次元モデルと写真画像の位置合わせを正確に行うためのマッチングは、写真画像から抽出したエッジ(2D エッジと呼ぶ)と 3次元モデルから抽出し

たエッジ間の距離により評価を行う ICP 法 [2] を用いる。

平行移動 (3 個), 回転移動 (3 個), 視野角 (1 個) をパラメータとし, この 7 個のパラメータの値を変えながらモデルエッジを作成し, エッジ間距離が最小となるパラメータを求める。マッチングは初期位置決定, 移動回転マッチング, 視野角マッチングの 3 ステップで構成される。

(1) モデルの初期位置決定

ICP 法は初期値が的確に定まっていなければ局所解に収束しやすい。まず, 平行移動 $T(T_x, T_y, T_z)$, 回転移動 R_x を用い刻み幅を大きく取り, ICP 法を用いておおよそのカメラパラメータを推定する。

(2) 移動と回転によるマッチング

平行移動 $T(T_x, T_y, T_z)$, 回転移動 $R(R_x, R_y, R_z)$ をパラメータとしてマッチングを行うアルゴリズムは以下のとおりである。ここでは視野角は一定にする。

1. 平行移動の刻み幅を s_t , 回転移動の刻み幅を s_r とし, 6 個のパラメータの初期探索範囲を $T - s_t \leq T \leq T + s_t$, $R - s_r \leq R \leq R + s_r$ とする。
2. まず移動パラメータ T のみ (R は固定) でモデルを移動させ, モデルエッジを得る。
3. エッジ間の距離が最小になる T を求める。
4. 3. で求めた T を固定したまま, 回転パラメータ R を変更し, モデルエッジを得る。
5. エッジ間の距離が最小になる R を求める。
6. 刻み幅が規定値以下であれば終了。そうでなければ刻み幅をそれぞれ小さくし, 5. で求めた R を固定して 2. に戻る。

(3) 視野角マッチング

視野角が変わると, 視点ベクトルと垂直な平面以外は見た目が変わってくる。顔の場合では, 例えば正面からのマッチングは耳のラインの形状が特に影響を受ける。このため, より正確なカメラパラメータを求めるためには視野角マッチングが必要になる。視野角マッチングのアルゴリズムは以下のとおりである。

1. 視野角の刻み幅を s_f とし, 視野角の初期探索範囲を $F - s_f \leq F \leq F + s_f$ とする。
2. 視野角が変わってもモデルの大きさが同じに見えるように, 3 次元モデルと仮想カメラの距離

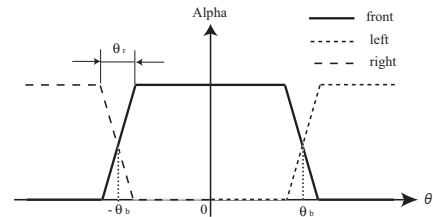


図 5: ブレンディング関数

を計算しモデルエッジを得る。

3. エッジ間の距離が最小になる F を求める。
4. 刻み幅が規定値を下回れば終了。そうでなければ刻み幅を小さくし, 2. に戻る。

2.3. テクスチャマッピング

撮影画像を 2.2 で述べたマッチング処理により, カメラの位置および向きを修正を行い, 3 次元モデルに撮影画像をマッピングする。

テクスチャマッピングは, モデルをマッチングした位置に移動させた後に各ポリゴンに行く。マッピングされるポリゴンの条件は,

- (1) ポリゴンが視点から可視,
- (2) 視点からの視線ベクトルと表面ポリゴンの法線ベクトルのつくる角が 80 度以下,

の 2 つである。

この条件ではマッピングできないポリゴン (ホールと呼ぶ) が存在する場合, これを埋める必要がある。1 つのポリゴンのみのホールであれば周囲 3 ポリゴンのテクスチャデータを補完するような処理を追加する。

また, 撮影されたテクスチャ画像は 3 枚であるため, すべてをマッピングすると境界面で視覚的に不連続になる。良好なモデルを作成するために, 図 5 の関数を使ってブレンディングを行う。 $x-z$ 平面における角度が, モデルの原点から $\pm(\theta_b - \frac{\theta_r}{2}) \sim \pm(\theta_b + \frac{\theta_r}{2})$ の角度内の位置の表面ポリゴンに対してブレンディング処理を行う。

本論文では θ_b を 45 度, θ_r を 20 度とした。

3. 実験

実験環境は Pentium4 2[GHz], ビデオカードは nVIDIA Quadro2 EX である。デジタルカメラは

表 1: 処理時間 [sec]

処理内容	正面	右	左
エッジ抽出	1	0.5	0.5
初期値決め	15	24	24
エッジマッチング	45	44	44

Sony Cyber-shot DSC-F505 を使用し，撮影時の解像度は 640×480 [pixel] である．

CT データの平面解像度は 0.468×0.468 [mm]，スライス間隔は 2.5 [mm]，スライス枚数は 103 枚である．実際に使用する 3 次元モデル(図 1(b))は 31676 ポリゴンである．

デジタルカメラで顔頭部の正面，左右の 3 方向から撮影した画像を用いる(図 6(a)，図 7(a)，図 8(a))．写真画像のサイズは不必要な左右の領域を取り除いた 480×480 [pixel] である．

まず，写真画像から 2D エッジを抽出し，マッチングの初期値を設定した(図 6(b)，図 7(b)，図 8(b))．求めた初期値からマッチングを行った結果を図 6(c)，図 7(c)，図 8(c)に示す．マッチング処理での各部位の重みは，正面では左目:0.2，右目:0.2，鼻:0.2，口:0.2，左耳:0.1，右耳:0.1 とし，側面では顔前面:0.4，耳:0.6 とした．

求めたカメラパラメータを用いて 3 枚の撮影画像を 3 次元モデルにテクスチャマッピングを行った結果を図 6(d)，図 7(d)，図 8(d)に示す．図 9 は撮影方向以外から見た結果である．図 10 に，実際の顔写真を示す．図 9 と図 10 は極めて似ている．

マッチングの各処理に要した時間を表 1 に示す．処理時間は従来手法が写真画像 3 枚を処理するのに約 6 時間かかっていたのに対し，本手法では約 200 秒で行うことができた．

顔頭部モデルには髪が存在しないため，頭の部分は不自然であるが，顔領域は実物に近いモデルを作成できた．また，写真を撮影していない方向からモデルを見た場合でも元の人物の顔を再現することができた．

4. まとめ

顔頭部の CT データから再構築された 3 次元モデルに対してデジタルカメラで撮影された顔画像との位置合わせを行い，高速にテクスチャマッピングする手法を提案した．本手法により，3 枚の写真画

像を 3 次元モデルにテクスチャマッピングする一連の処理を，約 200 秒で実行することができた．

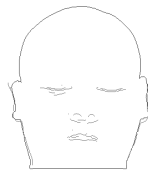
顔頭部を CT により測定するときと，写真撮影を行う時は姿勢が異なっているため，顔の形状が若干変化している．今後の課題としては，写真画像もしくは 3 次元モデルを変形させることにより，より正確にマッチングを行うことである．また，現在の 3 次元モデルに頭髪モデルを加え，よりリアルな顔頭部の CG を作成することである．

参考文献

- [1] T. Akimoto and Y. Suenaga. Automatic creation of 3D facial model. *IEEE Comput. Graphics & Appl.*, Vol. 13, No. 4, pp. 16–22, Sep. 1993.
- [2] Ali Md. Haider and Toyohisa Kaneko. Automated 3D-2D projective registration of human facial images using edge features. *Accepted for publication in the International Journal of Pattern Recognition and Artificial Intelligence(IJPRAI)*.
- [3] Ali Md. Haider and Toyohisa Kaneko. Automatic reconstruction of 3D human face from CT and color photographs. *IEICE transactions on Information and Systems*, Vol. E82-D, No. 9, pp. 1287–1293, Sep. 1999.
- [4] Yuya Iwakiri and Toyohisa Kaneko. PC-based real-time texture painting on real world objects. *EURO-GRAPHICS*, Vol. 20, No. 3, pp. C105–113, 2001.
- [5] R. M. Koch, M. H. Gross, F. R. Carls, D. F. von Büren, G. Fankhauser, and Y. I. H. Parish. Simulating facial surgery using finite element models. In *Proc. of SIGGRAPH '96*, pp. 421 – 428, 1996.
- [6] W. E. Lorensen and H. E. Cline. Marching cubes: A high resolution 3D surface construction algorithm. *Computer Graphics*, Vol. 21, No. 4, pp. 163–169, Jul. 1987.
- [7] J.F.S. Yau and N.D. Duffy. A texture mapping approach to 3-D facial image synthesis. *Computer Graphics Forum*, Vol. 7, pp. 129–134, 1988.
- [8] 山戸一宏, 寺田賢治, 上田哲史, 大橋剛介. フラクタル符号を用いた顔の自動識別. 第 6 回画像センシングシンポジウム, pp. 429 – 434, June 2000.



(a) 写真画像



(a) 初期値決め

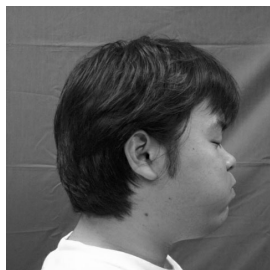


(b) マッチング結果



(c) マッピング結果

図 6: 結果 (正面)



(a) 写真画像



(a) 初期値決め



(b) マッチング結果



(c) マッピング結果

図 7: 結果 (右側面)



(a) 写真画像



(a) 初期値決め



(b) マッチング結果



(c) マッピング結果

図 8: 結果 (左側面)



(a) right 45°



(b) left 45°

図 9: 撮影点以外の視点のマッピング結果



(a) right 45°



(b) left 45°

図 10: 撮影点以外の視点の写真画像