

マルチカメラを用いた人体頭部 3D モデリング システムとその応用

藤村恒太[†] 大上靖弘[†] 寺内智哉[†]

概要 : 複数台のデジタルカメラを用いて、人体頭部の三次元形状と色彩 (テクスチャ) 情報を一度に入力・モデリングする 3D モデリングシステム「ピェリモ™」について、技術と応用の両面から紹介する。本システムの特徴は次の 3 つである。1) モデリングに必要な全画像を短時間 (3 秒以内) で撮影する。2) Shape-from-Silhouette 手法とマルチカメラアクティブステレオ手法を併用した独自のハイブリッドモデリング手法により、頭髪を含んだ頭部全体をモデリングするとともに、目鼻口耳などの細部を高精度に再現する。3) 全周囲の形状と写真画質のテクスチャからなるモデルをパソコンで一度に生成する。このため本システムは、テクスチャ情報とシルエット情報を取得するために通常光のもとで撮影するカメラ 16 台と、詳細形状を取得するためにパターン光を投影して撮影するカメラ 12 台の合計 28 台のカメラと 2 台のスライドプロジェクタ、および制御・計算用のパソコンからなる。本稿では、これらの技術について紹介すると共に、彫像切削加工やクリスタル加工、各種シミュレーションなどへの応用展開についても紹介する。

Systems and Applications based on Human-head 3D Modeling Technology using Multiple Cameras

Kouta Fujimura[†] Yasuhiro Oue[†] Tomoya Terauchi[†]

Abstract : A multi-camera 3D modeling system named “Pierimo” to digitize a human head is presented in this paper. The main features of this system are as follows: 1) Fast capturing: Both of texture images and pattern images can be taken within a few seconds using multiple digital still cameras which are set around the target human. Slide projectors are also set to provide a color line patterned light on the target for pattern image capturing, 2) Hybrid Algorithm: Our modeling algorithm is based on a hybrid method where the Shape-from-Silhouette technique and the Active-Stereo technique are combined. In the first step, the rough shape of the target is estimated in a voxel space using our Extended Shape-from-Silhouette method. In the next step, the shape is refined based on the depth-map data that is calculated using a multi-camera active stereo method. This combination makes up for the shortcomings of each method, and 3) Realistic Shape and Texture: The whole shape and photo-realistic textures of the human head including hair can be digitized at a time on a personal computer. So, This system consists 28 digital cameras, which include sixteen cameras for capturing under normal lighting and twelve cameras for capturing projected pattern, two slide projectors and a personal computer for controlling them and generating 3D data. These algorithms are described in this paper. And some business applications are also introduced, for example statues carved from wood, crystal engraved by laser and some software simulations.

[†] 三洋電機株式会社 技術開発本部 デジタルシステム技術開発センタービジネスユニット
Digital Systems Development Center, Technology R&D Headquarters, SANYO Electric Co., Ltd

1. はじめに

近年、三次元グラフィックスは、ゲームや映画などの制作現場で広く利用されるとともに、一般ユーザにとってもより身近な存在になり、よりリアリティの高いものが求められるようになってきた。特に人体・顔の三次元の情報は、映像制作だけでなく、計測、認証などのさまざまな応用展開も期待されている。一方、政府の推進するデジタルアーカイブ構想¹にみられるように、有形・無形文化財のデジタル化の動きが、世界的規模で始まっている。しかし、実在する三次元物体を、ありのままにデジタル化することは、装置の価格、使い勝手、対象への制限などから、容易ではないのが現状である。このため、高性能・高機能で安価な3Dスキャナ・デジタイザの実現が望まれている。

現在市販されている3Dスキャナ・デジタイザは、次の2つに分類できる。一つは接触式と呼ばれるもので、アーム型の装置を用いて対象物に触れ、各点の三次元座標を取得することを繰り返すことにより、対象物表面の三次元情報を点群として獲得する。もう一つは非接触式と呼ばれるもので、主には光をアクティブに照射し、照射された光が物体にあたっている様子を、別の角度からセンサで観測することにより、三角測量の原理で三次元情報を算出する。これもまた物体表面の三次元情報を点群として獲得する。なお、ここで照射される光には、レーザー光、パターン光、モアレ縞光、構造化光、正弦波光など^{2,3,4}がある。これらの3Dスキャナ・デジタイザは、精度が高く、計測・検査の分野で広く普及している。しかし、いくつかの問題を抱えているのも事実である。特に人体や顔を対象とした場合には、その問題が重要になってくる。その問題を以下に示す。

- 1) 接触式は、時間と労力が相当必要で、長時間の静止が困難な対象には不向きである。
- 2) 非接触アクティブ方式は、光を反射、透過、吸収する対象(黒髪など)は計測できない。
- 3) 全周囲の三次元データを一度に取得できない。もしくはノウハウや労力が必要である。
- 4) 高品位な色彩(テクスチャ)を取得できない。

しかしながら、非接触式には、特別な光を照射することなく、撮影された画像のみから三次元情報を獲得するパッシブ方式もある。これは、一般的にShape-from-X(Xには、Shading, Silhouette, Stereo, Motion などがある⁵⁻⁹)と呼ばれている。このパッシブ方式は、対象表面の材質に影響されにくいという特徴がある。我々は、この特徴を生かして、より多

くの物体を入力可能なShape-from-Silhouette(SfS)手法を拡張したアルゴリズムを採用した3Dモデリングシステム”サイバーモデラー™”を開発した¹⁰。このシステムは、デジタルカメラと自動回転台を使用した非常にコンパクトで簡便なシステムになっており、全周囲の形状とテクスチャを一度に獲得することを特徴としている。しかしながら、SfS手法では、シルエットの輪郭に現れない細部の凹凸が再現できないという問題があった。そこでさらに、画像相関ベースのパッシブステレオ手法を併用したShape-from-Silhouette/Stereo(SfS²)手法を開発した¹¹。この手法では、SfSにより求められる形状を基にステレオ対応付けを制限することにより、実用的な時間で細部の再現が可能となった。しかし、人体や顔を対象とした場合には、物体表面のテクスチャ特徴が不十分なため、パッシブなステレオは不向きである。そこで本システムでは、SfSにアクティブ方式のステレオ手法を併用したハイブリッドモデリング手法を開発し、採用した¹²。本手法では、カラーのストライプパターンを対象物体に投影して撮影した画像を、コード化してステレオ処理することにより、計算時間短縮を実現している。また今回さらに、その結果をベースに、サブピクセルレベルのステレオ処理を追加した2段階ステレオ処理を開発し、高精度化を実現したので合わせて紹介する。

本稿では、このハイブリッドモデリング手法により、人体頭部の三次元形状とテクスチャを一度に入力・モデリングことを可能にした3Dモデリングシステム「ピエリモ™」について、技術と応用の両面から紹介する。本システムの特徴は次の3つである。

- 1) モデリングに必要な全画像を短時間(3秒以内)で撮影する。
- 2) 独自のハイブリッドモデリング手法により、頭髪を含んだ頭部全体をモデリングするとともに、目鼻口耳などの細部を高精度に再現する。
- 3) 頭髪を含む全周囲の形状と写真画質のテクスチャをパソコンで一度に生成する。

以下、第2章では、システムの概要を、第3章では、3Dモデリングアルゴリズムの概要を、そして第4章では、本アルゴリズムの中核をなすハイブリッドモデリング技術の詳細について紹介する。続いて、第5章では、モデリング結果の例を示し、第6章では、彫像切削加工やクリスタル加工、各種シミュレーションなどへの応用展開について紹介する。

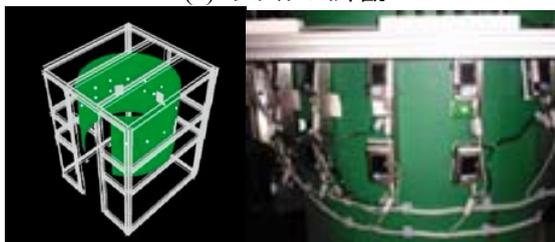
2. システム概要

本ピエリモ™システムは、28台のデジタルカメラ、2台のスライドプロジェクタ、照明装置、および1台

のパソコンからなる。デジタルカメラは、テクスチャ情報とシルエット情報を取得するために通常光のもとで撮影されるカメラが16台と、詳細形状を取得するためにパターン光を投影して撮影されるカメラ12台である。各カメラは、パソコンにUSB接続されており、複数台が同時に制御される。また、スライドプロジェクタと照明装置は、それぞれRS-232Cでパソコンに接続された専用コントローラを介して自動制御される。本システムの外観を、図1(a)に示す。また、システムの内部の様子を図1(b)に示す。



(a) システム外観



(b) システム内部の様子

図1 システム概要

撮影は、通常光撮影とパターン光撮影の右、左の順で行われ、3秒以内で完了する。撮影画像の例を図2に示す。(a)は、通常光撮影の画像例で、(b)は、左側のパターン光撮影の画像例である。



(a) 通常光撮影 (b) パターン光撮影

図2 撮影画像例

3. アルゴリズム概要

本システムでは、まず、カメラのキャリブレーション処理として、カメラの内部パラメータの算出と外部パラメータである各カメラの位置・姿勢計算を行う。そしてこれらの情報を基に、入力対象を撮影した画像に対し、以下の3つのステップで三次元モデルを生成する(図3参照)。本章では、カメラ位置・姿勢計算と各ステップの概略について述べる。

ボクセル空間での三次元形状生成
ポリゴン生成
テクスチャ生成

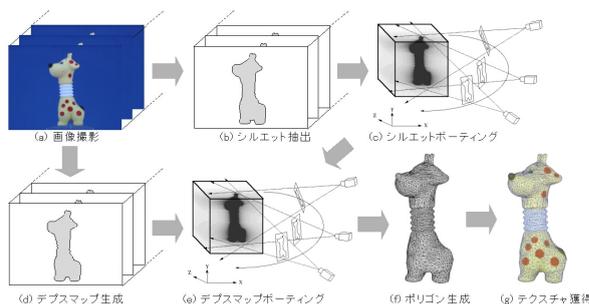


図3 アルゴリズムの流れ

3.1. カメラ位置・姿勢計算

各カメラ位置は、図4に示すように6色の円マークを不規則に配した参照物体を撮影した画像から、以下のステップで計算される。

撮影画像を領域分割し、領域ごとの円らしさ、円周長、色を基準にして、円マークを抽出する。

抽出された円マークの任意の三つとシステムに予め登録されている同じ色に分類される参照円マーク三つを用いてカメラ位置パラメータ(平行移動成分 x, y, z , 及び回転成分 θ, ϕ, ψ)を計算する。この計算をすべての組み合わせについて行い、それぞれの結果を六次元 Hough 空間に投票する。Hough 空間からピーク値を抽出し、カメラ位置パラメータを獲得する。

得られたカメラ位置パラメータは、Hough 空間での投票分解能に精度が制限されるため、Levenberg-Mardquart 法により最適解を求める。

以上の手順で、カメラ位置と姿勢、つまり参照物体座標系とカメラ座標系との変換パラメータを獲得する。図5に Hough 空間への投票の様子を示す。



図 4 カメラ位置計算用参照物体

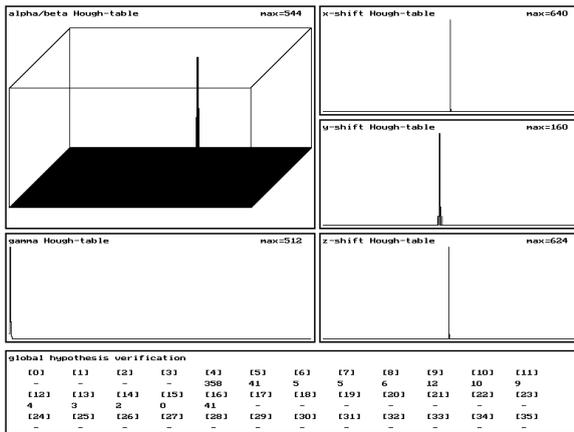


図 5 Hough 空間への投票

3.2. ボクセル空間での三次元形状生成

取得した各カメラのキャリブレーション情報と通常光撮影の画像とパターン光撮影の画像から三次元形状を生成する。本システムでは、図 6 に示すように、Shape from Silhouette 手法と Multi-camera Active Stereo 手法を併用した当社独自のハイブリッドモデリングアルゴリズムを採用し、ボクセル空間で物体形状を再現する。本アルゴリズムについては、次章で詳しく述べる。



図 6 ハイブリッドモデリングアルゴリズム

3.3. ポリゴン生成

ボクセル空間で生成された三次元形状を、ポリゴン(三角パッチ)を用いた表面モデルで表現する。ここでは、表現に必要なポリゴン数を少なく抑え、かつ形状表現の精度を維持することが要求される。本アルゴリズムでは、以下に示す方法によってポリゴンを生成することでこれらの要求を満たしている。

物体表面を構成するボクセルの中心を頂点として、隣接頂点を接続し、初期ポリゴンを生成する。

隣接するポリゴンを統合する前後の形状の変化を評価し、変化が小さい隣接ポリゴンを順に統合することで、ポリゴン数を目標値まで削減する。

3.4. テクスチャ生成

全周囲の通常光撮影画像から、全ポリゴンのテクスチャ情報を獲得する。具体的には、各ポリゴンのテクスチャ情報を与える画像(以下参照画像と呼ぶ)を決定した後、ポリゴンを参照画像に逆投影し、その投影部分の色彩をテクスチャ情報として獲得する。しかし、実際には、ポリゴンを逆投影可能な参照画像は複数存在する場合がほとんどである。そこで本アルゴリズムでは、ポリゴンの参照画像への投影面積の大きさと、隣接するポリゴンが参照する画像のカメラ位置の近さをあらかず評価関数を設定し、エネルギー最小化問題として解くことにより、参照画像を決定している。さらに、参照画像だけではなく、その隣接カメラ画像のテクスチャ情報を用いた平均化処理と、ポリゴン境界での距離重み付けスムージング処理を施すことにより、ポリゴン境界での画像の連続性を高めている。

4. ハイブリッドモデリングアルゴリズム

本章では、ハイブリッドモデリングアルゴリズムの詳細について紹介する。

4.1. 拡張 Shape-from-Silhouette 法

通常光撮影の画像から、あらかじめ指定した背景領域の各画像の色情報(クロマキー色)を用いて、物体と背景を分離し、シルエット情報を求める。次に、各カメラのシルエット情報を、カメラキャリブレーション情報を用いて、オブジェクト座標系に設定された三次元ボクセル空間への投票(ポーティング)を行い、その結果を閾値処理することにより、三次

元形状を獲得し、これを概略形状とする。この投票・閾値処理により、多少のエラーがあっても、形状精度の低下が抑えられる。シルエットポータリング処理の様子を図 7 に示す。

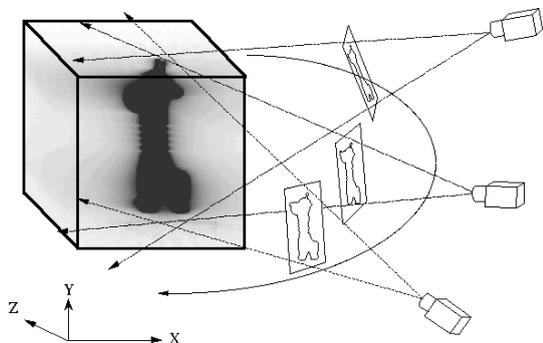


図 7 シルエットポータリング処理

4.2. マルチカメラアクティブステレオ

シルエットの輪郭に現れない細部の凹凸形状は、概略形状には再現されない。そこで、複数台のカメラで撮影されたパターン光画像を用いてステレオ処理を行い、人体頭部で重要な目鼻口耳などの詳細な形状を取得し、概略形状に統合する。本アルゴリズムは、以下の3つのステップからなる。

- 1) カラーコード3眼ステレオ処理
- 2) 相関ステレオサブピクセル処理
- 3) デプスマップ統合処理

以上のように、本アルゴリズムでは、2段階のステレオ処理により、実用的な時間での高精度な形状獲得を可能にしている。以下、各ステップについて詳しく述べる。

4.2.1. カラーコード3眼ステレオ処理

本ステップでは、以下のサブステップでパターン画像をコード化し、ステレオ処理を行う。

概略形状をパターン光撮影画像面へ逆投影し、ステレオマスク画像を生成する。
 マスク画像内のパターン画像を領域分割し、領域境界のエッジ点の両側の色情報を獲得し、カラーコード化する。
 エッジ点のステレオ対応候補を見つける。
 3眼ステレオ処理により対応を絞り込む。
 概略形状を基準にした奥行き(デプス)制約により、ステレオ対応候補を絞り込む。
 オクルージョン制約により、対応を絞り込む。
 対応結果によるデプスを計算する。
 エッジ点のデプスから、それ以外の点のデプスを補間し、デプスマップを生成する。

図 8 に、カラーパターンの様子を示す。次に、この3眼ステレオ処理について以下に詳しく述べる。

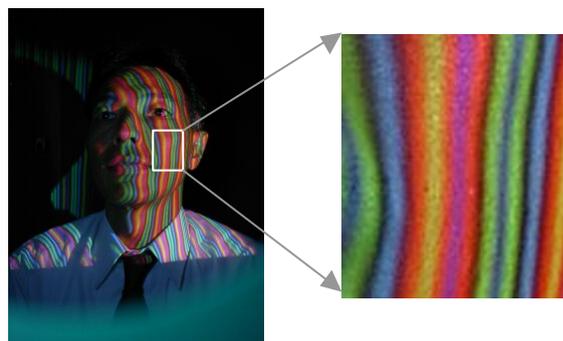


図 8 カラーパターンの様子

図 9 に示すように、現在ステレオ処理している基準画像をマスタ画像、対応候補を探索する画像をスレブ画像とする。また、スレブ画像のステレオ対応候補を絞り込むための、もう一方のステレオ対応処理画像をチェック画像とする。3眼ステレオ処理としては、以下のとおりである。

- i. スレブ画像、チェック画像ともに全てのエッジ点において得られる対応候補について、デプスを算出する。
- ii. スレブ画像の各対応候補について、デプスが最も近いチェック画像の対応点とのデプスの差を求め、それが閾値以外のものは、正しい対応でないとは判断し、対応候補から除く。

以上により対応候補が絞られ、そして次のサブステップでさらに絞り込まれる。最終的に複数の候補が得られた場合は、概略形状から得られるデプスに近い候補が選択される。以上によりマスタ画像上のパターン境界エッジ点のデプスが求まる。そしてサブステップで、エッジ点以外について、距離の重み付けられた補間処理によりデプスを計算し、全体のデプスマップを生成する。

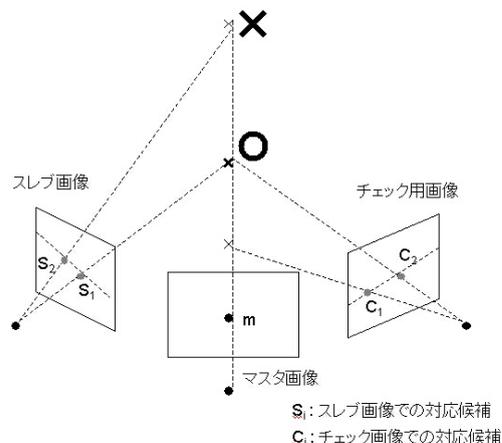


図 9 3眼ステレオ処理

4.2.2. 相関ステレオサブピクセル処理

さらに高精度に奥行き値を算出するために、マスタ画像のエッジ点について、先のピクセルレベルのステレオ処理で得られたデプスマップからスレブ画像上での対応点位置を逆投影により求め、これを基準に画像相関ベースのサブピクセル処理を行う。まず、デプスマップから逆投影で得られた対応点 i とエピポーラ線上で隣接する2点 $i-1, i+1$ について、それぞれ相関値 $C(i), C(i-1), C(i+1)$ を求める。その3点を放物線近似し、最大値が得られる座標をサブピクセルレベルでの対応点とする。なお、相関には正規化相互相関を用いている。

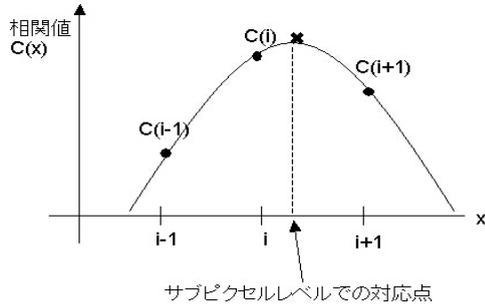


図 10 サブピクセル相関ステレオ処理

4.2.3. デプスマップ統合

複数のカメラ視点でのデプスマップを、概略形状に統合し、最終形状を獲得する。ここでも三次元ボクセル空間への投票方式を採用することによりデプスマップに含まれるエラーを軽減しながら、詳細形状を反映した全体形状を獲得する。図 11 にカメラ視点 i のデプスマップをボクセル空間のある平面へポーティング処理している様子を示す。 $L(v, i)$ は、ボクセル v と視点 i との距離を示す。 $dixy$ は、デプスマップと視線の交点 (x, y) のデプス値を示す。ポーティング処理では、 $L(v, i)$ が $dixy$ 以上のボクセル v に 1 を加える。この処理を三次元ボクセル全体に対して処理することで、視点 i でのポーティング処理が終了する。そして、すべての視点からのポーティング処理した結果を閾値処理することにより、最終形状を確定する。

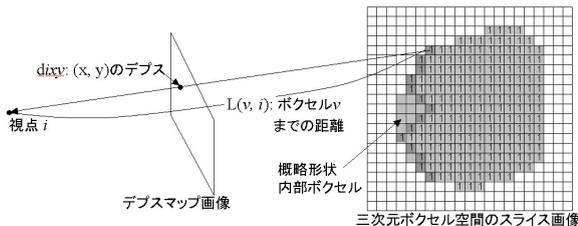


図 11 デプスマップポーティング処理

5. モデリング結果例

モデリング結果の例として、男性と女性に適用した結果を、それぞれ図 12、図 13 に示す。

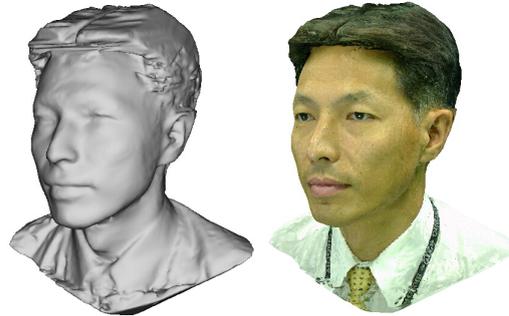


図 12 モデリング結果例 1



図 13 モデリング結果例 2

図 12 は、図 2 で例示した撮影画像から、10 万ポリゴンでモデリングした結果についての、シェーディング表示 (左) とテクスチャ表示 (右) のスクリーンショットである。また、図 13 の左の画像は、通常光撮影の画像例であり、右側は 3 万ポリゴンのモデリング結果をワイヤーフレーム表示したスクリーンショットである。いずれも、頭髪を含む頭部全体の形状を、目鼻口耳などの細部の形状とともに良好に再現している。また、図 12 では、写真画質のリアルなテクスチャを全周囲で獲得できているのが分かる。

6. 応用展開

本章では、3Dモデリングシステム「ピエリモ™」を活用した応用展開について、いくつか紹介する。

6.1. 彫像・人形切削加工

本システムは頭髪を含んだ頭部全体のモデルを生成可能であるという特徴を生かして、3Dデータから切削加工用データを生成し、従来にないリアルな彫像・人形を自動造形するシステムへ展開している。図 14 に示した彫像例は、図 13 に示した結果から加工した例で、左から 1/8, 1/2, 1/4 サイズとなっている。材料は主に木を使用している。



図 14 彫像切削加工例

その他の例として、図 15 に、加工後にブロンズ色などの着色を施した例を示す。



図 15 その他の彫像加工例

6.2. クリスタル加工

切削加工と同様に、獲得した3Dデータから、点群データを生成し、クリスタル中にレーザで自動加工するシステムへも展開している。ここでも、360度全周囲の形状とテクスチャを活用し、従来にないリアルかつ全方向から観賞できるクリスタル製品に仕上がっている。図 16 に、そのクリスタル加工の例を示す。左側は、図 13 の結果から加工した例である。



図 16 クリスタル加工例

6.3. 各種シミュレーション

本システムでは、頭部全体の形状とリアルなテクスチャが獲得できるため、ソフトウェア上では、様々なシミュレーションへの応用が考えられる。ここでは、

その例として、眼鏡試着のシミュレーション、ヘアスタイルシミュレーション、ウエディングドレス試着シミュレーションについて、それぞれの様子を、図 17、図 18、図 19 に示す。

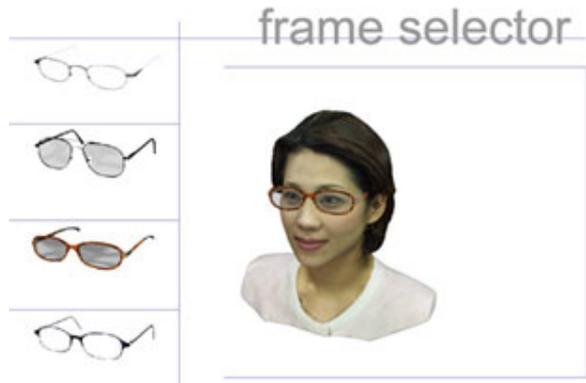


図 17 眼鏡試着シミュレーション



図 18 ヘアスタイルシミュレーション



図 19 ウエディングドレス試着シミュレーション¹⁷

これらの他に、コンピュータグラフィックス^{14、15}や、似顔絵システム¹⁶などへも応用展開を図っている。

7. まとめ

本稿では、複数台のデジタルカメラを用いて、人体頭部の三次元形状とテクスチャ情報を一度に入力・モデリングする3Dモデリングシステム「ピエリモTM」について、技術と応用の両面から紹介した。本システムの特徴は次の3つである。1)モデリングに必要な全画像を短時間(3秒以内)で撮影する。2) Shape-from-Silhouette 手法とマルチカメラアクティブステレオ手法を併用した独自のハイブリッドモデリング手法により、頭髪を含んだ頭部全体をモデリングするとともに、目鼻口耳などの細部を高精度に再現する。3)全周囲の形状と写真画質のテクスチャからなるモデルをパソコンで一度に生成する。

本システムは、テクスチャ情報とシルエット情報を取得するために通常光のもとで撮影するカメラ16台と、詳細形状を取得するためにパターン光を投影して撮影するカメラ12台の合計28台のカメラと2台のスライドプロジェクタ、および制御・計算用のパソコンからなる。そして、ハイブリッドモデリングアルゴリズムでは、カラーのストライプパターンを対象物体に投影して撮影した画像を、コード化してステレオ処理することにより、計算時間短縮を実現している。また、今回さらに、その結果をベースに、サブピクセルレベルのステレオ処理を追加した2段階ステレオ処理を開発し、高精度化を実現した。本稿では、これらの技術を紹介するとともに、その応用展開として、本システムの特長である360度全周囲の形状とテクスチャを活用した、彫像切削加工やクリスタル加工、各種シミュレーションなどについても紹介した。

今後の課題としては、3Dモデリングシステムとしては、人体全身への拡張や動物体への対応が挙げられる。また、様々な応用展開においては、目的に合わせたシステムカスタマイズの効率化、モデリング計算の高速化、加工用データへの連携などが挙げられる。今後も、技術と応用の両面から研究開発を進めていく予定である。

謝辞

日頃から「ピエリモTM」システムの応用展開に活躍され、本稿作成においても資料提供など多大なご協力を頂いた、三洋電機(株)ニュービジネスセンタービジネスユニットピエリモプロジェクトの石川猶也リーダー、およびメンバー全員に深く感謝する。

参考文献

1. デジタルアーカイブ推進協議会, “デジタルアーカイブ白書2003,” 2003.
2. R. B. Lewis and A. R. Johnston, “A scanning laser range finder for a robotic vision,” *Int. Joint Conf. Artificial Intelligence*, pp. 962-968, 1972.
3. G. T. Agin and T. O. Binford, “Computer description of curved objects,” *IEEE Trans. Computer C-25 (4)*, pp. 439-449, 1976.
4. M. Idesawa et al., “Scanning moire method and automatic measurement of 3-d shapes,” *Appl. Opt.*, pp. 2151-2162, 1977.
5. P. Giblin and R. Weiss, “Reconstruction of surfaces from profiles,” *Int. Conf. Computer Vision*, pp. 136-144, 1987.
6. B. Horn, “Obtaining shape from shading information,” *Psychology of Computer Vision*, P. H. Winston ed., McGraw-Hill, pp. 115-155, 1975.
7. D. Marr, “Vision,” W. H. Freeman and Co., 1982.
8. S. Ullman, “The interpretation of structure from motion,” *Proc. the Royal Society of London*, pp. 405-426, 1979.
9. A. Witkin, “Recovering surface shape and orientation from texture,” *Artificial Intelligence* **17**, pp. 17-45, 1981.
10. Y. Matsumoto, K. Fujimura, and T. Kitamura, “CyberModeler: A compact 3D scanner based on monoscopic camera,” *Proc. SPIE, Three-Dimensional Image Capture and Application II*, pp 2-10, 1999.
11. Y. Matsumoto, K. Fujimura and T. Kitamura, “Shape-from-Silhouette/Stereo and Its Application to 3-D Digitizer”, *Proc. DGCI, LNCS 1568*, Springer-Verlag Berlin Heidelberg, pp. 177-188, 1999.
12. K. Fujimura, et al, “Multi-camera 3D modeling system to digitize human head and body,” *Proc. of SPIE Vol.4298*, pp.40-47, 2001.
13. D. Ritter and Y. Matsumoto, “3d-surface reconstruction with a hand held video camera”, *Image Multidimensional Digital Signal processing '98 (IMDSP '98)*, pp. 5-8, 1998.
14. CGWORLD編集部：“多眼3次元モデリングシステム「ピエリモ」”, *CGWORLD*, Vol.40, pp.128-129, 2001
15. CGWORLD編集部：“多眼3次元モデリングシステム「ピエリモ」”, *CGWORLD*, Vol.41, pp.78-81, 2001
16. T. Fujiwara et al, “3D似顔絵フィギュア制作の実用化への試み,” *情報処理学会論文誌：コンピュータビジョンとイメージメディア*, Vol43, No. SIG 4 (CVIM 4), pp.85-94, 2002
17. <http://www.violette.jp/>