

テリトリベース 3D-2D レジストレーション手法の拡張とその応用

喜多 泰代[†] DaleL. Wilson^{††} J.Alison Noble^{†††} 喜多 伸之[†]

Extension of territory-based 3D-2D registration method for a wide-range of applications

Yasuyo KITA[†], Dale L. WILSON^{††}, J. ALISON NOBLE^{†††}, and Nobuyuki KITA[†]

あらし 3次元自由形状剛体の位置・姿勢をその2次元投影図から高速に算出する手法[1]をより汎用的に活用するための拡張を提案する。文献[1]においては、対象の3次元モデルとその位置・姿勢に関する大まかな初期推定が与えられるという前提条件のもと、頭部血管の中心軸を3次元モデルとし、1枚の観測X線像に実時間で位置合せする手法を提案した。本研究では、この手法をより多くの局面に適用するため、次の2点の拡張を行う。まず第一に、複数観測画像情報の統合や移動するカメラ座標系を取扱えるようにするため、3次元モデルの位置・姿勢算出式を任意個の観測カメラ座標系を想定した形へと一般化する。第二に、モデルと観測画像間の対応特徴として遮蔽輪郭線も用いられるようにする。観測画像上の遮蔽輪郭に対応する3次元モデル上の点を高速に得るために、OpenGLに代表されるようなグラフィックスシステムが算出する奥行き画像(各画素に対象までの距離が記述された画像)を効果的に利用する。これら2点の拡張による汎用性の向上と手法の有用性を、それぞれ1)複数X線像を用いた3次元血管モデルの位置・姿勢決め、2)プラント環境内のカメラの位置・姿勢検出の二つの応用実験により示す。

キーワード 3D-2D レジストレーション、自由形状剛体位置・姿勢検出、医用ビジョン、ロボットビジョン。

1. ま え が き

3次元物体の位置・姿勢をその投影画像から決定する課題は、コンピュータビジョン研究に於ける基本的で重要な一課題であり、その応用も機械部品の位置あわせ、医用画像における人体・臓器位置あわせ、顔の向き検出など広範囲に及ぶ。これまで多くの手法が開発されてきたが、そのアプローチは大きく、1)画像ベースと2)特徴ベースに分類できる。前者は、あらゆる位置姿勢からの対象物体の見え方をモデルとして、画像レベルで観測画像との類似性が最大となる位置・姿勢を求める。よって、対象の形状によらず適用

できるが、モデルが膨大となる問題がある。固有空間を用いた効果的な探索空間の圧縮[2]などにより高速化が進んだが、複雑なシーンのすべての見え方の算出にはやはり時間がかかる上、あらかじめ考慮していない画像の変化に対処できないという短所がある。

後者は、対象の3次元モデルを作成し、観測画像から特徴(エッジやコーナーなど)の抽出を行い、3次元モデル上の特徴との対応を求めて、この関係を満たす対象の適切な位置・姿勢を算出する。対象が明らかな特徴点を持ち3D-2Dの対応点ペアが容易に得られる場合には、対象の位置・姿勢は高速に算出できる[3]。当初、機械部品など明らかな特徴点を持つ物体を対象に研究が進められたが、最近では、自由形状物体のように明確な特徴を持たない場合でも、対象の位置・姿勢の初期値が与えられることを前提として、研究が行われている。具体的には、2次元観測特徴点と3次元モデルの投影点との距離[4]や、2次元観測特徴点への投影線と3次元モデル点との距離[5][6]を基準として対応付けし位置・姿勢を求める手法が提案されている。ただし、自由形状物体の場合、遮蔽輪郭

[†] 産業技術総合研究所 知能システム研究部門, つくば市
R.I. of Intelligent Systems, N. I. of Advanced Industrial Science and Technology, 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568 Japan

^{††} CSIRO, シドニー
Division of Mathematics and Information Sciences, CSIRO, Locked Bag 17, North Ryde NSW 1670, Australia

^{†††} オックスフォード大学, オックスフォード
Dept. of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PJ, England

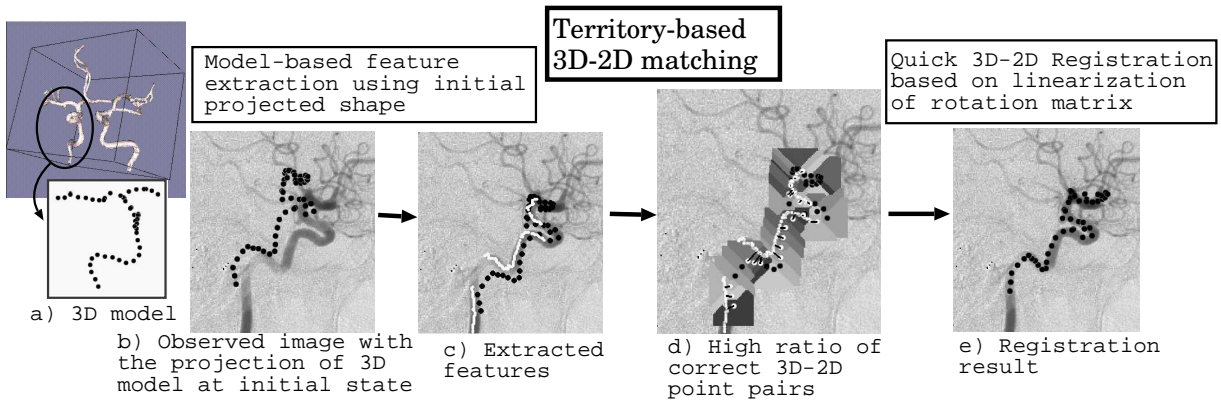


図1 テリトリベース 3D-2D レジストレーション手法 [1] の基本手順
Fig. 1 Basic scheme of territory based 3D-2D registration method [1]

などの観測特徴を画像から頑健に抽出すること自体が難しく、完全な自動化はできていなかった。

これに対して、文献 [1] では、抽出すべき特徴が欠けていても信頼性高く対応付けを行えるテリトリベース 3D-2D マッチング手法を開発し、これにモデルドリブな特徴抽出を組み合わせることで完全な自動化を可能にし、血管の位置あわせに適用した。ただし、この手法では血管が管状物体であることを利用し、その中心軸を 3 次元モデルとして利用しているため、このままでは管状物体以外の対象に適用できない。また、3D-2D 点対応から対象の正しい位置・姿勢を算出する式も、1 枚の観測 X 線像に対する状況しか想定していなかった。

本研究論文ではより適用範囲を広げるために次の 2 つの拡張を行う。まず、3D-2D 対応点ペアからモデルの 3 次元位置・姿勢を算出する式を、任意個の観測カメラ座標系を想定した、より一般的な形に整理する。これにより、複数の観測画像情報の統合や動くカメラ座標系の取り扱いも簡潔に行えるようになる。次に、より一般的な形状の対象に適用するため、観測画像上の遮蔽輪郭を 3 次元モデルとの対応特徴として用いられるようにする。観測遮蔽輪郭に対応するモデル表面上の輪郭生成曲線を効率よく算出するために、OpenGL などに代表されるグラフィックスシステムを手法に取り込み、それが供給する予測シーンの奥行き画像 (各画素に対象までの距離が記述された画像) を利用する。

以下、2 章で論文 [1] で提案した基本となる 3D-2D レジストレーション手法を簡単に説明し、3 章で第一

の拡張を述べ、4 章でその実応用における有効性を示す。さらに、5 章で第二の拡張を述べ、6 章ではこの拡張を活かした、プラント環境内のカメラの位置・姿勢検出手法および実験結果を示す。最後に、7 章でまとめを述べる。

2. テリトリベース 3D-2D レジストレーション手法 [1]

本章では、文献 [1] で提案した、基本となる 3D-2D レジストレーション手法を簡単に説明する。図 1 に、血管の 3 次元モデルとその 2 次元投影像を位置あわせする例を用いて、全体の処理の流れを示す。図 1a に示すように、対象の 3 次元モデルは血管の 3 次元再構成データの中心線から等間隔で抽出した代表点で表されている。図 1b は、図 1a の 3 次元モデルに対応する造影 X 線像である。観測 X 線像に対する 3 次元モデルの位置・姿勢は X 線装置の目盛から得られるが誤差を含むので、この情報に基づき算出される 3 次元モデルの投影像は図 1b の黒い点に示されるようにずれる。この初期状態から、処理は次のように行われる。

1) モデルの投影像を観測画像上で 2 次元的に平行移動して、暗い陰影 (血管候補領域) に最も重なるように移動する (図 1c の黒い点)^(注 1)。

2) 画像上の対応特徴、ここでは 2 次元血管の中心線をこの投影像の近傍でモデル駆動型処理により抽出

(注 1) : 文献 [1] では、3 次元モデル自身もこの平行移動に見合うように初期移動させていたが、画像上での最短距離に基づいた 3D-2D 点対応関係から、正しい位置・姿勢への初期状態からの移動量が直接算出できるため実際には必要なく、現在は行っていない。

する (図 1c の白い線)。

3) 画像上の各画素を最も近い 3 次元モデル投影点のテリトリに分割することにより、モデルの投影点の形状に応じた不等方性のテリトリ領域を設定し、その領域内で最短の観測特徴点をモデル点の対応位置とする。この処理をテリトリベース 3D-2D マッチングと呼び、テリトリ領域への探索限定により、抽出すべき観測特徴の欠けに対して頑健な処理となっている。

4) 得られた 3 次元モデル点 $\mathbf{X}_i(X_i, Y_i, Z_i)^\top$ と観測点 $\mathbf{x}_i(x_i, y_i, f)^\top (i = 1 \sim m)$ の m 個のペアを満たす、3 次元モデルの回転運動 \mathbf{R} を式 (1) の評価式を用いて算出し、その平行移動 $\mathbf{T} = (t_x, t_y, t_z)$ を式 (2) の $2m$ 連立方程式を解いて算出し、新しい位置・姿勢を求め。

$$\min_{\mathbf{q}} \sum_{i=1}^{m-1} \sum_{j=i+1}^m ((\mathbf{x}_i \times \mathbf{x}_j) \cdot (q_0(\mathbf{X}_i - \mathbf{X}_j) + 2\mathbf{q}' \times (\mathbf{X}_i - \mathbf{X}_j)))^2 \quad (1)$$

$$\frac{f(X'_i + t_x)}{Z'_i + t_z} = x_i \quad \text{and} \quad \frac{f(Y'_i + t_y)}{Z'_i + t_z} = y_i \quad (2)$$

ただし、 f は焦点距離、 $\mathbf{q} = (q_0, q_1, q_2, q_3)^\top$ は \mathbf{R} を表す 4 元数で、 $\mathbf{q}' = (q_1, q_2, q_3)^\top$ 。また、 $\mathbf{X}'(X'_i, Y'_i, Z'_i)$ は式 (1) で得られた \mathbf{R} を用いて $\mathbf{X}' = \mathbf{R}\mathbf{X}$ として算出する。対応ペアが厳密に正しいものではないこと、回転行列の線形化を行っていることにより、正しい位置・姿勢は 1 回の演算では算出されないが、3D-2D マッチングとそれに基づくモデルの移動を繰り返し処理することにより図 1e のような矛盾のない陰影を与える 3 次元位置・姿勢に収束する。

3. 観測カメラ座標系の一般化

上述した式 (1),(2) を、ワールド座標系に任意個数の観測カメラが存在する場合に拡張する。まず、カメラ座標系 c のワールド座標系における原点、各軸方向余弦、そして焦点をそれぞれ、 $(X_{0,c}, Y_{0,c}, Z_{0,c})$ 、 $(\lambda_{x(y,z),c}, \mu_{x(y,z),c}, \nu_{x(y,z),c})$ 、 f_c と表すこととする。また、そのカメラ座標系における座標値を表すために、上付き文字 “c” を用いる。3 次元モデル点 $\mathbf{X}_i(X_i, Y_i, Z_i)^\top$ と観測点 $\mathbf{x}_i^c(x_i^c, y_i^c, f_c)^\top$ の m 個のペアから 2 組、 i, j を選び、文献 [3] と同様に、図 2 に示すような位置関係に着目する。そして、

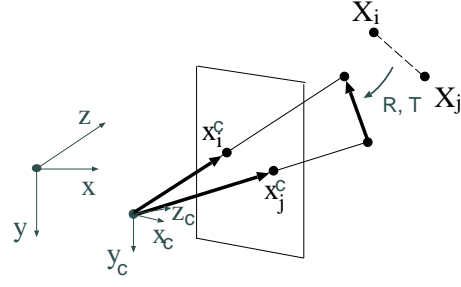


図 2 3D モデル点と観測点の位置関係
Fig.2 Geometrical relation between 3D model points and observed points

$(\mathbf{R}\mathbf{X}_i + \mathbf{T}) - (\mathbf{R}\mathbf{X}_j + \mathbf{T}) = \mathbf{R}(\mathbf{X}_i - \mathbf{X}_j)$ が $(\mathbf{x}_i^c \times \mathbf{x}_j^c)$ に垂直でなければならない条件より、平行移動線分を除去した回転成分にだけ依存する評価式、

$$\sum_{i=1}^{m-1} \sum_{j=i+1}^m (F_c(\mathbf{x}_i^c \times \mathbf{x}_j^c) \cdot \mathbf{R}(\mathbf{X}_i - \mathbf{X}_j))^2$$

を得る。ただし、 F_c は、カメラ座標系での値をワールド座標系での値に変換する関数、すなわち、 $\mathbf{X} = F_c(\mathbf{X}^c)$ である。

ここで、 n 個の観測カメラ座標系があるとして、添字 c を $c_k (k = 1 \sim n)$ に拡張し、各系で m_{c_k} 個の 3D-2D 点対応が得られているとする。回転行列 \mathbf{R} を前章で記述したのと同様に 4 元数で表し、線形近似すると次の評価式が得られる。

$$\min_{\mathbf{q}} \sum_{k=1}^n \sum_{i=1}^{m_{c_k}-1} \sum_{j=i+1}^{m_{c_k}} (F_{c_k}(\mathbf{x}_i^{c_k} \times \mathbf{x}_j^{c_k}) \cdot (q_0(\mathbf{X}_i - \mathbf{X}_j) + 2\mathbf{q}' \times (\mathbf{X}_i - \mathbf{X}_j)))^2 \quad (3)$$

平行移動成分は、各カメラ座標系で式 (2) と同様な関係式を作成し、これに $t_{x(y,z)}^{c_k} = \lambda_{x(y,z),c_k} t_x + \mu_{x(y,z),c_k} t_y + \nu_{x(y,z),c_k} t_z$ を代入することにより得られる、次のような $2 \sum_{k=1}^n m_{c_k}$ 個の連立方程式を解けば求める。

$$\begin{aligned} (f_{c_k} \lambda_{c_k,x} - x_i^{c_k} \lambda_{c_k,z}) t_x + (f_{c_k} \mu_{c_k,x} - x_i^{c_k} \mu_{c_k,z}) t_y + \\ (f_{c_k} \nu_{c_k,x} - x_i^{c_k} \nu_{c_k,z}) t_z = x_i^{c_k} Z_i'^{c_k} - f_{c_k} X_i'^{c_k} \\ \text{and} \\ (f_{c_k} \lambda_{c_k,y} - y_i^{c_k} \lambda_{c_k,z}) t_x + (f_{c_k} \mu_{c_k,y} - y_i^{c_k} \mu_{c_k,z}) t_y + \\ (f_{c_k} \nu_{c_k,y} - y_i^{c_k} \nu_{c_k,z}) t_z = y_i^{c_k} Z_i'^{c_k} - f_{c_k} Y_i'^{c_k} \end{aligned} \quad (4)$$

ここで、 $\mathbf{X}_i'^{c_k}(X_i'^{c_k}, Y_i'^{c_k}, Z_i'^{c_k})$ は、式 (3) で得られた \mathbf{R} を用いて $\mathbf{X}_i'^{c_k} = F_{c_k}^{-1}(\mathbf{R}\mathbf{X}_i)$ から算出される。

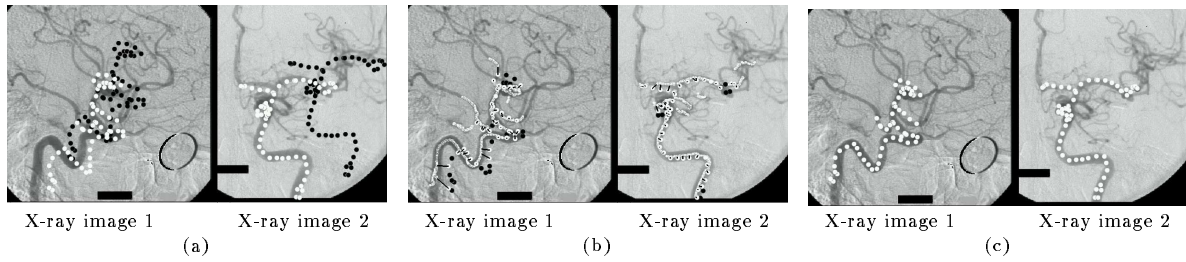


図3 図1aの血管モデルの左内頸動脈の2枚のX線画像を用いた実験: (a) 初期状態における3次元モデル投影像; (b) テリトリベース探索領域を用いた対応付け(繰り返し処理第1回め); (c) 得られた位置・姿勢における3次元モデルの投影像

Fig. 3 Experiment using two views of the left internal carotid artery in Fig. 1a; (a) projections of the 3D model at the initial state; (b) correspondences using territory-based search regions; (c) projections of the 3D model at the resultant position and pose.

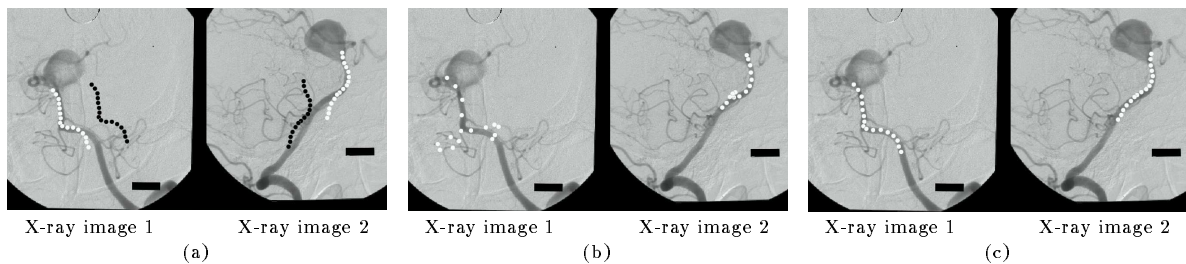


図4 単純形状の脳動脈の2枚のX線画像を用いた実験: (a) 初期状態における3次元モデル投影像; (b) 各X線像を別々に用いた場合の結果; (c) 2枚のX線像を同時に用いた結果

Fig. 4 Experiment using two views of a basilar artery: (a) projections of the 3D model at the initial state; (b) results using one image respectively; (c) result using the two images simultaneously.

4. 複数観測画像を用いた3次元血管モデルの位置・姿勢検出

3で導出した式を用いて、2枚の観測X線画像から3次元血管モデルの位置・姿勢を検出する実験を行った。X線装置の目盛から算出される2枚のX線像間の相対的な位置関係には多少の誤差が含まれるが、X線装置と血管モデルの相対位置関係に関する初期誤差、姿勢方向 ± 10 度程、位置 ± 50 mm程と比較し十分小さいので、ここではないと想定して実験を行っている。図3は、図1aと同様な3次元血管モデルのレジストレーション結果を示す。図3a内の黒および白い点は、それぞれ初期状態、および最も暗い領域に重なるように画像上で平行移動した後の3次元モデル点の投影像である。図3bには、第1回目のテリトリベース3D-2Dマッチングによって得られたペアを示す。テリトリベース3D-2Dマッチングと式(3),(4)を用

いたモデル移動算出処理を27回繰り返した後、3次元モデルは図3c内の白い点に示すような投影図を生じる位置・姿勢に収束した。3次元モデルの初期状態よりの移動量は、 13.0 度の回転移動、 22.3 mmの平行移動であった。

この例では、対象が複雑な形状であるため1枚の観測画像だけからでも血管モデルの位置・姿勢を検出でき、複数枚の観測画像を用いる大きな違いはない。しかし、図4に示すように対象が単純形状の場合、複数の観測画像を用いることが非常に効果的である。図4a内の黒および白い点は、図3a同様、それぞれ初期状態、および平行移動後の3次元モデル点の投影像である。それぞれの観測画像ごとに3次元モデルの位置・姿勢を検出した結果の投影像を図4bに示す。図4bの左画像では、 63.6 度回転して、大きく陰影からずれた結果となった。図4bの右画像では投影像が観測陰

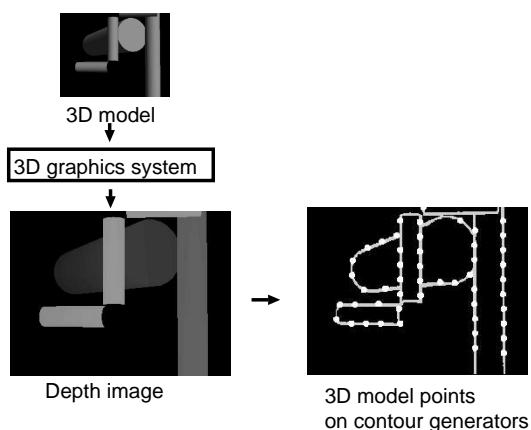


図5 グラフィックスシステムの提供する奥行き画像を利用した3次元モデル点算出

Fig.5 Calculation of 3D model points using the depth image supplied by graphics system

影に重なってはいるが、62.5度回転しており、初期値の回転のずれが10～20度ぐらいである実際の値から大きく外れている。一方、図4cに示す、2枚を同時に用いて位置・姿勢を算出した結果は16.9度の回転という、実際の回転誤差10～20度内の値で、両画像において妥当な投影を与えている。残念ながら、この実験画像に関する血管の正しい位置・姿勢情報は測定されていないため、その精度は検討できていない。

5. 遮蔽輪郭線利用への拡張

3次元の血管モデルの例では、視線方向とは関係なくその中心線が対応特徴として利用できた。一般的には遮蔽輪郭が典型的な観測特徴であり、これに対応するモデル表面上の3次元曲線、輪郭生成曲線は視線方向によって大きく変化し、3次元モデルの位置・姿勢が変化するごとに算出し直さなければならない。複雑な3次元シーンにおいても観測特徴に対応する3次元モデル点を効率よく算出するために、我々はOpenGLのような3次元グラフィックスシステムがハードウェアを用いて高速に算出するシーンの奥行き画像(各画素に対象までの距離が記述された画像)を利用する。具体的には、図5に示すように、奥行き画像を微分処理して奥行きが大きく変るオクルーディングエッジ(図中のグレイ線)を求め、そのエッジ上から等間隔で抽出した代表点(図中の白点)の画素の奥行き情報よりその点の3次元座標値を得る。

これにより、次のような利点がある。

(1) 隠れ線処理も含む複雑な輪郭生成曲線の算出が容易に行える。

(2) 既存の専用ハードウェアで高速化できる。

(3) 予測される見え方情報を同時に利用できる。

三点目は、同時に作成する見え方画像上で濃淡差が少ないと予測されるモデル点は除去するなど、処理の頑健性向上に活かすことができる。その利用の仕方は、対象の形状情報だけでなく、その表面の反射特性、照明環境まで含めて、モデルがどこまで正確に構築されているかによって大きく変わる。本論文では、6.3において、反射特性や照明環境が特に与えられていない場合の利用例を示す。

6. プラント環境内のカメラの位置・姿勢検出

6.1 アルゴリズム

カメラを取り巻く環境モデルが与えられる場合、その3次元モデルをカメラで観測した画像にレジストレーションすることで、カメラの位置・姿勢を検出することができる。この原理に基づいて、拡張した3D-2Dレジストレーション手法をアクティブカメラの位置・姿勢検出のためのビジュアルフィードバックに適用する。アクティブカメラは、その制御コントロールモジュールなどから位置・姿勢を把握することが可能であるが、通常これらの値はなんらかの誤差を含み、動いている間に誤差が累積されていく。提案手法で、この誤差を求め、修正することを考える。図6に、5に述べた遮蔽輪郭に対応する3次元モデル点の算出機構を組み込んだカメラの位置・姿勢検出アルゴリズムの流れ図を示す。処理の手順は次のようになる。

i) グラフィックスシステムにおいて初期位置に設定した仮想カメラにおける予測奥行き画像の微分画像を求め、微分値の大きい点の奥行き情報より輪郭生成曲線上の3次元モデル点を算出する。同じく仮想カメラにおける予測見え方画像を利用し、頑健に抽出される可能性の低いモデル点は省く。

ii) テリトリベースマッチングを行い、上記3次元モデル点と観測画像特徴間で3D-2D点対応ペアを得る。式(3)(4)を $n=1$ の条件で用いて、3D-2D点対応ペアから3次元モデルの移動量を算出する。

iii) 仮想カメラの位置・姿勢をii)で得た移動量に対して相対的に移動して更新する。

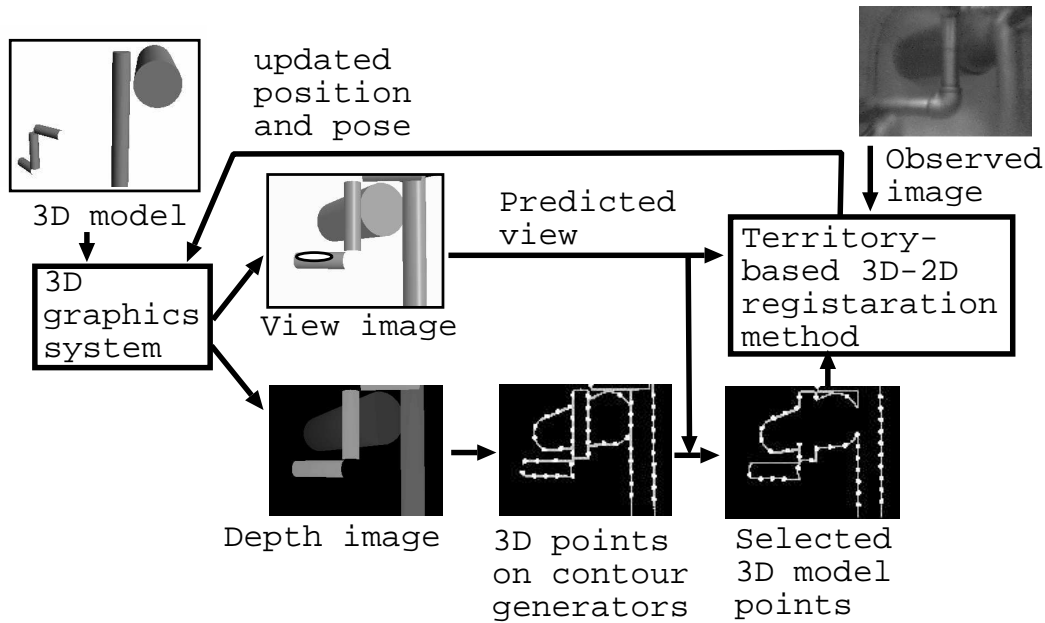


図6 遮蔽輪郭を用いたカメラの位置・姿勢検出アルゴリズム

Fig. 6 Scheme of determination of the position and pose of a camera using occluding edges

i) から iii) までの処理を繰り返すことにより、予測画像が観測画像に収束していき、正しい位置・姿勢が得られる。

図 1 に示した基本手順との大きな相違点は、1) 3次元モデル点を各繰り返し毎に算出し直す、2) 対象物体の代わりに撮影系の位置・姿勢を変更する、の 2 点である。

6.2 合成画像を用いた実験

図 7 に合成画像を用いた実験結果を示す。図 7a はカメラをとりまく世界を上から見た図で、画像右下にある白い丸がカメラの位置を、そこからでている 2 本の線がカメラの視線と上方向を示している。図 7b がこの状態の観測画像である。もし、カメラの初期位置・姿勢の推定に誤りがあれば、予測画像は観測したものとはずれる。図 7c は、その状態を観測画像上に予測画像を重ねて表している。図 7a, c 内の白い点は処理 i) で得られる、3次元モデル点を示している。図 7d 内の白い点は観測画像上におけるその投影位置を表す。この投影形状を、黒い線で示される観測エッジに最も重なるように、画像上で 2 次元的に平行移動した結果が灰色の点で示す位置である。処理 ii) では、テリト

リーベース 3D-2D マッチングを用いて図 7d に細い線分で示す、3次元モデル点の観測画像上の対応位置が決定される。図 7e, f は処理 i) から iii) を 11 回繰り返した後の、収束結果である。実際には、平行移動誤差なしの軸 $(0.89, 0.0, -0.45)$ の回りに 11.5 度の回転誤差であるのに対して、 $(-0.06, -0.13, -0.05)$ の平行移動誤差、軸 $(0.89, -0.0, -0.45)$ の回りに 11.4 度の回転誤差が検出された。平行移動誤差の大きさの目安として、手前の円柱の高さが 30.0 である。

許容可能な初期見積り誤差に関する検討に関しては [7] に詳しくあるが、許容誤差はモデルの形状などに依存して変動する。本実験においては、任意の回転軸まわりに 15 度以下の回転誤差で 10 以下の平行移動誤差であれば、正しく収束する傾向が見られた。

6.3 実データを用いた実験

パイプが複雑に配置されているために観測画像上の特徴の多くが遮蔽輪郭であり、マークとなるような特徴的な点がほとんど存在しない原子力プラントの実物大モックアップ環境内 [8] で、移動ロボットに積載したアクティブカメラヘッドを動かし、実験を行った。この実験では、対象環境の色情報・反射特性や照明環

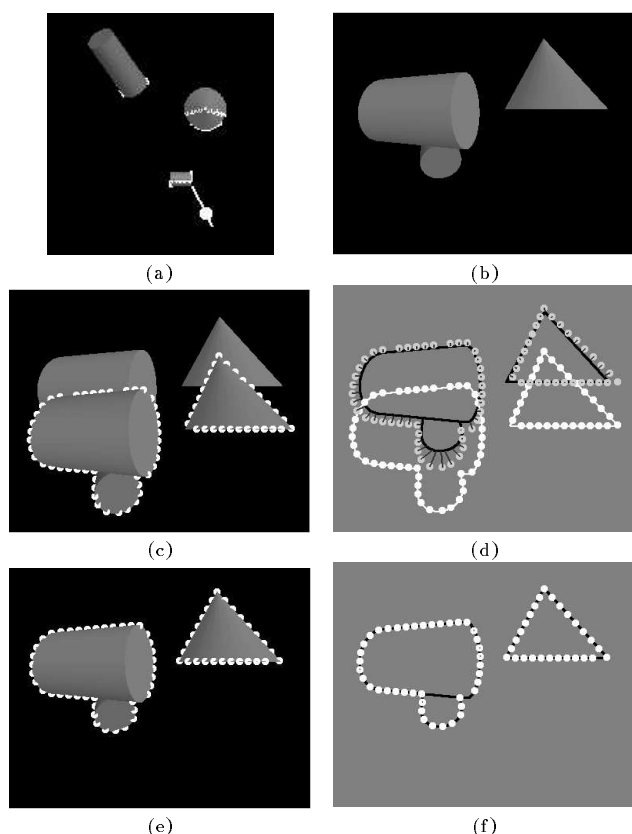


図7 人工データを用いたビジュアルフィードバックの実験: (a) 上から見たカメラ位置 (画像右下の白い丸); (b) 観測画像; (c) 観測画像へのモデル投影; (d) 観測エッジ (黒い線) への3次元モデル点の投影 (白が初期状態、灰色は初期平行移動後); (e) 収束後のモデル投影画像; (f) 収束後の3次元モデル点の投影

Fig. 7 Experiment on visual feedback using synthetic data (a) top view of the 3D model; (b) observed image; (c) projection of the model at the initial estimated state; (d) projection of the 3D model points on the observed edge (black lines); white and grey points are respectively at the initial state and after initial translation; (e) projection of the 3D model after convergence; (f) projection of the 3D model points after convergence.

境は特に与えられない場合として、「すべてのパイプは背景の壁より暗い同一色、天井のあたりに照明」という状況で予測濃淡画像を作成し、5の(3)予測見え方画像の利用を以下のように行っている。

(1) 頑健に抽出可能なモデル点だけを選択する。遮蔽輪郭上のモデル点に対応する予測濃淡画像上の画素のエッジ強度を調べ、その強度が弱く観測されにく

いと予測されるモデル点は除去する。

(2) 点对応ペア決定の属性として利用する。

具体的に、観測画像からは各エッジ近傍の、予測濃淡画像からは各モデル点近傍の濃淡値最大変化方向を算出して8方向に分類し、観測エッジ点、モデル点の属性として記録する。3D-2D 点对応処理において、異なる属性を持つモデル点、観測点是对応ペアになれないとする。

図8aは観測画像の一例である。図8bはこの画像のエッジ画像(Canny operator使用)で、各エッジ近傍の濃淡値最大変化方向を8方向に分類して、カラーで分けて表示してある。図8cにOpenGLで作成した円柱5本からなる部分モデルの側面図と上から見た図を示す。上面図中の白丸はカメラ位置、そこから出ている線は視線方向を示す。ワールド座標系は、図8cの上面図に示すように、 x 、 z 軸を水平床面上にとり、 y 軸は鉛直高さ方向にとる。側面図は、観測画像を得たカメラの初期位置・姿勢をマニュアルで計測し、その状態における予測画像となっている。図8dは、これを観測画像に重ねて黒く表示したもので、初期位置・姿勢計測誤差のため実際に各パイプが観測される位置からずれている。図8eに、予測奥行き画像のエッジ上に求めた3次元モデル点(白い点)を側面図、上面図に示す。ただし、予測濃淡画像情報を用いて明るさ変化が少ない点は省かれるため、同色のパイプ同士が重なっているところは選択されていない。図8fに、モデル点の投影位置を観測エッジ画像に重ねて表示する。各モデル点のカラーは、その点の濃淡値最大変化方向属性を表わす。

カメラ角度のわずかなずれが、画面上では大きな平行移動を生むので、第1回目の初期対応時だけはこれを考慮する。具体的に、投影モデル点群を画面上で2次的に平行移動し、モデル点と同じ方向性を持つエッジに最もよく重なる位置を探索する。図8gが移動後の位置である。この位置においてテリトリ3D-2D マッチングを行い、得られた対応付が同図中に白線で示されている。式(3),(4)を用いて得られた位置・姿勢変化量を相対的にカメラの位置・姿勢に反映し、その新しい位置において同様の処理を繰り返す。この例では、15回繰り返し後に収束条件を満たし、カメラ位置を(28.6, -4.8, -57.9)(mm)平行移動修正し、軸(-0.72, 0.10, -0.69)回りに3.9度回転修正して、図8h,iのような予測画像と観測画像の重なりを得た。パイプが複雑に遮蔽しあい、正反射領域のエッジが存在

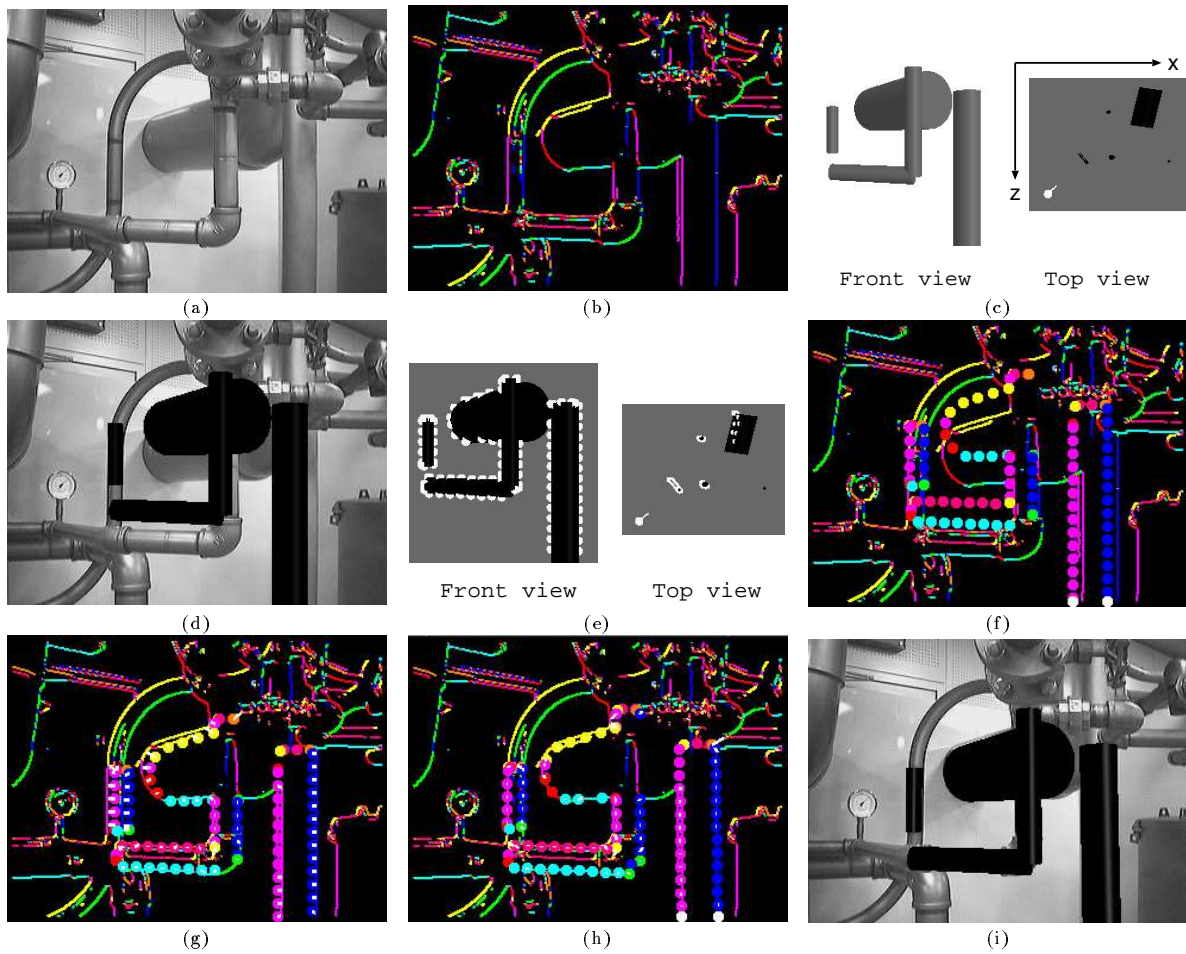


図8 実画像を用いたビジュアルフィードバックの実験 1: (a) 観測画像; (b) 観測画像のエッジ (濃淡値の最大変化方向別に濃度で表す); (c) 対象の3次元モデルの側面図および上面図; (d) カメラの初期推定位置・姿勢におけるモデル投影; (e) 3次元モデル点 (白い点) の正面図、側面図; (f) 3次元モデル点の観測エッジ画像での投影位置; (g) 初期平行移動後の3次元投影モデル点; (h) 収束後のカメラ位置・姿勢におけるモデル点投影 (エッジ画像上); (i) 収束後のカメラ位置・姿勢におけるモデル投影。

Fig. 8 Experiment on visual feedback using real data 1: (a) observed image; (b) edges in the observed image (classified by different grey levels); (c) front and top views of the 3D model; (d) projection of the model at the initial estimated state; (e) front and top views of the 3D model points (white points); (f) projection of the 3D model points on the observed edge image; (g) projection of the 3D model points after initial 2D translation; (h) projection of the 3D model points after convergence; (i) projection of the 3D model after convergence.

する難しい環境において、良好な位置・姿勢検出が行えている。多少のずれは、与えた3次元モデルが厳密に正確でないことが一因であると思われる。処理時間は、Pentium II(333MHz)マシン上で、画像上の初期平行移動に320msec、3Dモデル点を選出し、その

対応観測特徴点を求め、これに基づきカメラの位置・姿勢を変更する1ループに約200msecかかり、トータルで3.8秒であった。

同様に、カメラを移動させながら10ヶ所で本手法を適用した。その結果例を図9に示す。各コラムは、左か

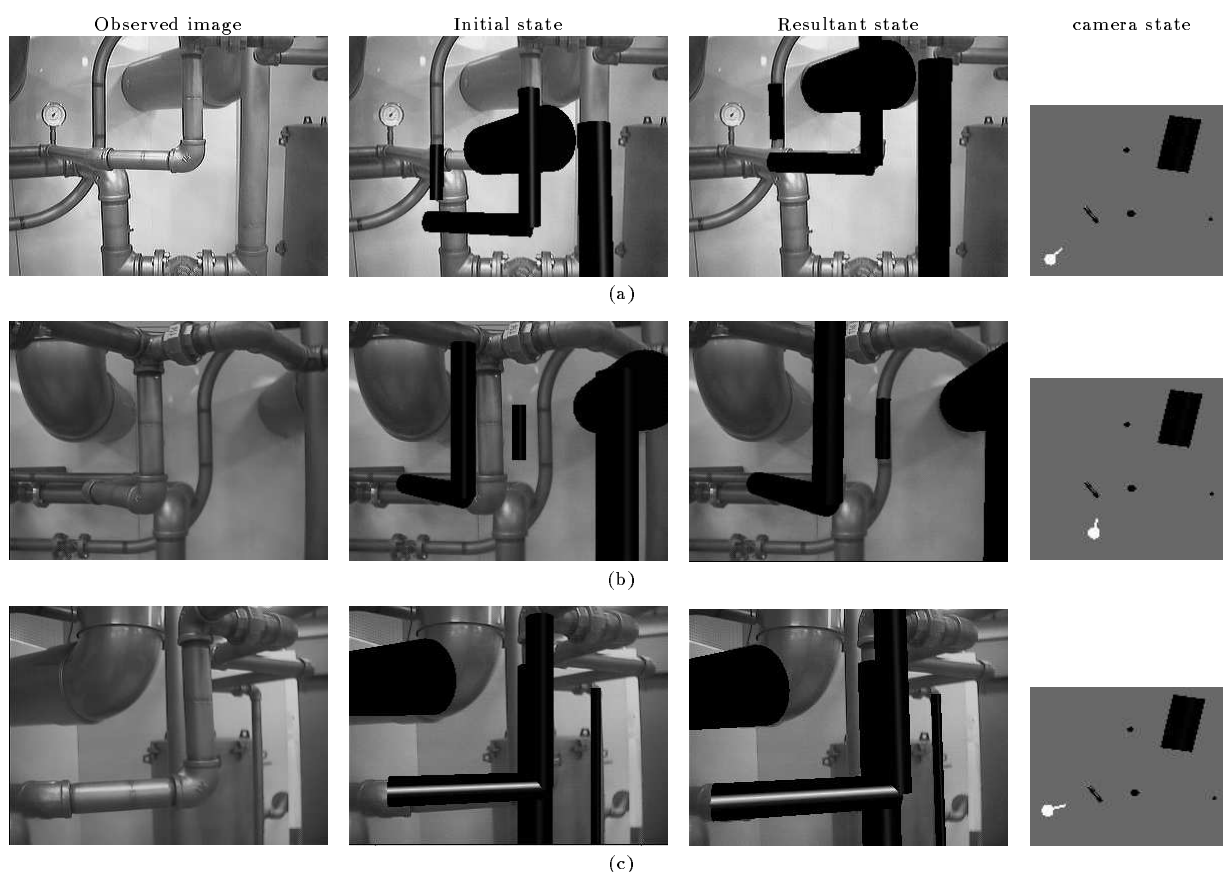


図9 実画像を用いたビジュアルフィードバックの実験 2
Fig. 9 Experiment on visual feedback using real data 2

ら観測画像、初期状態、収束結果、カメラの位置を示す上面図である。10ヶ所中8ヶ所で、図9a,bのように良好に収束した。図9aは、先の実験の状態から、カメラをチルト方向に10度下げて入力した画像であり、先の実験における収束時のカメラ位置・姿勢を初期状態として与えた。この例では10回繰り返し後に収束条件を満たし、カメラの位置・姿勢を $(-1.2, 1.8, 0.5)$ (mm)平行移動修正し、軸 $(0.69, 0.01, 0.72)$ 回りに10.2度回転修正した。これは、マニュアルで計測したチルト軸方向 $(0.64, 0.0, 0.77)$ の回りに10.0度の回転、 $(0.0, 0.0, 0.0)$ の平行移動というアクチュエータの値にととても近い値である。

残り2ヶ所では、図9cのようなずれを残した。この例では、中央部の2本の垂直なパイプによる遮蔽輪郭がカメラのわずかな位置の変動でその見え方の順序を変えるため、得られた観測特徴がどちらのパイプのも

のが紛らわしく、これがずれの原因であると思われる。他の失敗例も同様な状況であった。予測画像は、このような紛らわしい特徴の存在をあらかじめ予測するために利用することも可能であり、こうした利用方法を付加して、紛らわしい特徴に対応するモデル点を取り除くことを検討中である。

7. ま と め

3次元自由形状剛体の位置・姿勢を、その2次元投影図から高速に算出する手法の拡張について述べた。具体的に、

- 1) 3D-2D 対応点ペアからモデルの3次元位置・姿勢を算出するための式を任意個数の観測カメラが存在する場合に拡張。
- 2) グラフィックスシステムが供給する予測シーンの奥行き画像を利用し、遮蔽輪郭に対応する3次元モデル

点を効果的に算出。

の二点の拡張を行い、現実的に必要性の高い応用課題を用いて、それぞれの拡張の有効性を示した。

血管の3次元モデルの位置・姿勢検出課題においては、特に対象が単純形状の場合に複数観測画像を利用することの有効性を示した。また、3次元環境モデルの相対的な移動を利用した、アクティブカメラの位置・姿勢検出が同じ戦略で実現できることを示した。総合すると、拡張以前の問題点を克服し、より幅広い対象形状、観測状況に対応できる手法となった。

特に、グラフィックスシステムが提供する予測奥行き画像と予測見え方画像の利用の導入は、効果が大きい。今回示した実験においても、複雑な環境である上に金属表面による正反射による影響が強く、今まで対処が難しかった状況において、位置合わせを実現した。あらかじめ環境モデルが構築できる前提が必要であるが、そのモデルの情報量によっては、より難しい環境に対応できる発展性を秘めている。環境モデルからよりリアルな予測画像が生成され活用できるほど、処理の頑健性が増すので、照明環境なども考慮したより現実に近い人工画像を作成する view シミュレータ [9] を組み込んで使用することを検討している。

<謝辞> 臨床に関する助言、および実験に必要なデータをいただいた Dr J. Byrne と Dr D. Royston に感謝する。筆頭著者は、平井成興元電総研知能システム部長、産総研の視覚情報研究グループの人々に感謝する。

文 献

- [1] 喜多 泰代, Dale L. Wilson, J. Alison Noble: “3次元血管モデルのX線画像への実時間位置合わせ手法”, 信学論 (D-II), J83-DII, pp.254-262, 2000.
- [2] H. Murase and S.K. Nayar: “Visual learning and recognition of 3-D objects from appearance”, *International Journal of Computer Vision*, Vol. 14, pp.5-24, 1995.
- [3] J.J. Heuring and D.W. Murray: “Visual head tracking and slaving for visual telepresence”, In *Proc. of IEEE Int Conf. on Robotics and Automation*, 1996.
- [4] J. Feldmar, G. Malandain, N. Ayache, S. Fernandez-Vidal, E. Maurincomme and Y. Troussel: “Matching 3D MR angiography data and 2D X-ray angiograms”, In *Proc. of CVRMed-MRCAS'97*, pp.129-138, 1997.
- [5] S. Lavallée, and R. Szeliski: “Recovering the position and orientation of free-form objects from image contours using 3D distance maps”, *IEEE trans. Pattern Anal. & Mach. Intell.*, 17, 4, pp. 378-390, 1995.
- [6] P. Wunsch and G. Hirzinger: “Registration of CAD-Models to Image by iterative inverse perspective matching”, In *Proc. of 13th International Conference on Pattern Recognition*, pp. 78-83, 1996.
- [7] 喜多 泰代, Dale L. Wilson, J. Alison Noble, 喜多 伸之: “汎用 3D-2D レジストレーション手法とその応用例”, 情処研資, 00-CVIM-120, pp. 73-80, 2000.
- [8] 喜多 伸之: “移動ロボット群による原子力プラント巡回点検”, 電総研ニュース, Vol. 607, pp.2-7, 2000.
- [9] 比留川 博久, 喜多 伸之, 喜多 泰代, 末広 尚士, 平井 成興, 小原 一太郎, 河村 進: “HRP 仮想プラットフォームにおける View シミュレータ”, 日本ロボット学会第 17 回学術講演会, 3, 1217-1218, 1999.