

入力画像と生成画像の差の最小化に基づく 実時間運動追跡

日浦 慎作^{†1} 森谷 貴行^{†1} 佐藤 宏介^{†1}

本論文では計算機内のモデルから生成した画像と、センサから新たに得られた画像との差を最小化するような運動パラメータを求めることにより、画像からの特徴抽出を行うことなく物体の位置・姿勢を実時間的に追跡する手法について、これまで我々がやってきた複数の研究成果に他の手法との比較を交えながら概観する。濃淡画像を用いた運動の追跡については、1) 多視点映像を用いることによる安定な 6 自由度運動推定、2) ステレオカメラによる運動追跡と対象モデリングの並列化、3) 環境マッピング技術を用いた鏡面反射物体の運動追跡、について述べる。さらに同じ手法を距離画像に適用した例として、4) 実時間距離画像センサを用いた運動と全周形状の同時獲得と、5) 距離画像と濃淡画像の融合による運動推定手法についても紹介する。これらの手法は古くから知られている Horn らの勾配法や直接法のほかに近年注目を集めている Yang らの微分カメラに関する研究、さらには距離画像の位置合わせのための ICP 法などとも深く関連しており、それらとの関連性についても述べる。

Real-time Object Tracking Based on the Minimization between Input and Generated Images

SHINSAKU HIURA,^{†1} TAKAYUKI MORITANI^{†1}
and KOSUKE SATO^{†1}

In this paper, we describe a method to estimate the pose and position of the moving object by minimizing the difference between generated images and input images from sensors. At first, we listed several examples; 1) stable motion estimation using multiple viewpoint images, 2) simultaneous localization and modeling using stereo camera, 3) tracking of the object with mirrored surface based on environment mapping. Then we also introduce two research achievements, 4) simultaneous tracking and modeling using real-time range sensor, 5) object tracking using the fusion of range and intensity images. Furthermore, we will discuss about the relationships among ours, gradient, direct method, ICP, differential camera, and so on and tried to make it clear what is the originality and characteristic of our method.

1. まえがき

カメラから連続的に得た画像から、シーン中の物体の位置・姿勢を連続的に推定する研究は従来より広く行われている¹⁾。その理由として、工業生産の自動化やロボットの自己位置同定、さらには複合現実感やメディア応用など多くの分野において運動推定の実用的手法が求められていることのほかに、問題を純粋に幾何学的に取り扱うことができ、それゆえ美しい定式化が可能であることがあげられる。なぜならシーンからエッジや特徴点を検出・追跡できれば、それ以降はシーン表面の 3 次元座標とカメラ画像上の 2 次元座標との幾何学的関係だけを取り扱えばよいからである。

しかし画像とはそもそも、光源から発した光が物体表面で反射されて生じる輝度分布を記録したものであり、物体の運動や形状のような幾何学的要素と、物体表面の輝度分布のような測光学的要素は不可分なものである。それゆえ、なだらかに変化するテクスチャを持つ物体や表面につやを有する物体、またはそれらが混合されて観測される物体のように、物体表面に固定された特徴点が得にくい対象は幾何学的枠組みだけでは取扱いが容易ではない。そのため、このような物体の運動解析手法は十分研究されているとはいえない。いい換えると、運動と画像の関係を幾何学的な側面だけからとらえた場合についての数学的枠組みの構築に成功した代わりに、これらの手法は特徴抽出が難しい物体に対する汎用性を獲得することができなかったといえる。

他方、このような幾何学的枠組みに立脚した動物体の運動解析手法の欠点を補完すべく、パラメトリック固有空間法²⁾など物体のアピアランスに基づく手法が提案されている。アピアランスに基づく手法では様々な位置・姿勢に関する対象物体の画像を準備する必要があるが、それには対象物体を実際に撮影する手法以外に、対象物体の CG モデルからレンダリングにより画像を生成する手法が考えられる。特に時系列的な運動の追跡では物体の現フレームの位置・姿勢は前フレームまでに得た値の近傍にあると仮定することができるため、あらかじめ生成した画像を用いて解析するのではなく、追跡時に必要な画像を生成しながら運動を推定することも可能である。このような考え方にに基づき、我々は対象のモデルからレンダリングした画像と入力画像との差の最小化に基づく動物体追跡手法を提案してきた(図 1)。

^{†1} 大阪大学大学院基礎工学研究科

Graduate School of Engineering Science, Osaka University

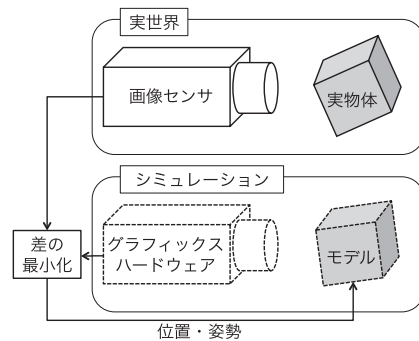


図 1 画像の差の最小化に基づく動物体追跡

Fig. 1 Tracking of moving object based on the minimization of image difference.

この手法では対象物体モデルからの画像生成さえ正確にできれば、その後の運動推定処理は数値解析における最適化問題以外のなにものでもなく、そのためシーン中で生じる様々な光学的現象や多様な画像センサに対して一般性を有する。実際に我々はこれまで単眼カラーカメラによる運動の推定手法³⁾を基本とし、さらに多視点画像を用いた 6 自由度運動推定の安定化⁴⁾や互いに隠蔽を生じる複数物体の同時追跡などを実現した。また本論文では新たに、環境マッピングによる鏡面物体のレンダリングに基づく鏡面反射物体の追跡手法について詳述する。

本論文で提案する手法では実画像と比較しうる品質の CG を生成する必要があり、そのための対象物体のモデリングのコストが高いのではないかと指摘が多くの研究者からなされてきた。しかしステレオ法やレンジセンサなど対象物体の形状計測技術は近年大幅に発展しており、画像からの特徴選択や手動によるモデリングを要する特徴点・エッジベースの手法に比べ対象物体への依存性が低く、自動化も容易である。実際に我々はステレオカメラによるシーンの形状計測と、それにより得たモデルを用いたカメラ自己運動推定を高精度かつ実時間的に並列処理する手法⁵⁾を提案した。また先述のように本手法はセンサの種類にかかわらず適用することができるため、距離画像における奥行き値をそのまま濃淡画像の輝度と見なすことで運動推定を行うこともできる。実際に我々は、実時間型レンジファインダから出力される距離画像を用いた 6 自由度運動推定と対象物体形状の同時全周モデリング⁶⁾や、この手法に濃淡画像の情報を融合することで見かけの形状に変化がないような運動の追跡手法⁷⁾を示した。

提案手法は各画素の輝度値の差を最小とするような対象物体の運動を求めるものであり、画素の隣接関係はいっさい用いられない。その点で提案手法は古くから知られている Horn らの勾配法⁸⁾に関連するが、画像上の局所領域の動きではなく、対象物体の形状情報に基づき物体全体の動きを直接的に求めるという意味で Horn らの直接法⁹⁾により近い手法である。しかし直接法は物体の動きとフローの関係(画像ヤコビアン)を求める必要があるのに対し、提案手法ではこれが CG レンダリングにおけるジオメトリ計算に置き換えられるため、CPU 負荷が小さくなるという利点がある。さらに提案手法は現在位置・姿勢近傍の画像を多眼カメラによって得る Yang らの微分カメラを用いる手法¹⁰⁾と最適化の原理に共通点を持ち、一方距離画像の位置合わせにおいて多用されている ICP 法とも深く関連を持つ。

本論文ではまず 2 章において、入力画像と生成画像の差の最小化に基づく運動推定の原理について述べた後、3 章でこの手法を用いた実装例を 4 件示す。さらに 4 章では原著論文としては未発表の鏡面反射物体の追跡手法について詳述する。5 章ではそれらをもとに、勾配法や直接法、ICP 法などとの関連について多面的に考察する。

2. 運動推定の原理

本論文で述べる運動推定手法は、1 章で述べたように入力画像と生成画像の差の数値解析的な最小化と見ることができ一方で、勾配法の拡張としても解釈することができる。そこでこの章では勾配法についてまず述べ、これを物体の位置・姿勢に直接関連づけることで本論文で提案する運動追跡手法の原理を示す。

2.1 勾配法に基づく運動パラメータ推定

実空間中での対象物体の運動にともない、撮影された時系列画像間で各画素に輝度変化が生じる。ここで対象物体表面の輝度分布が時間経過によって変化せず、各画素の輝度が運動のみにより変化するという仮定に基づき、輝度の変化からオプティカルフローを推定する手法が勾配法である。今、時刻 t における画像上の点 (x, y) の輝度を $I(x, y, t)$ とし、微小時間 δt 経過後にこの点が $(x + \delta x, y + \delta y)$ に移動したとすると以下の式が成り立つ。

$$I(x, y, t) = I(x + \delta x, y + \delta y, t + \delta t) \quad (1)$$

この式をテーラー展開によって 1 次近似し、整理することによって以下のオプティカルフロー拘束式を得る⁸⁾。

$$\frac{\partial I}{\partial x} \frac{\delta x}{\delta t} + \frac{\partial I}{\partial y} \frac{\delta y}{\delta t} + \frac{\partial I}{\partial t} = 0 \quad (2)$$

この式は画像平面上の各点の動き、すなわちオプティカルフローと輝度変化の関係を表し

た式であり、求めるべきフローが画像上各点につき 2 パラメータであるのに対し輝度変化は 1 パラメータであるために、何らかの拘束条件がなければ解くことができないことが知られており⁸⁾、これをアパーチャ問題と呼ぶ。しかし剛体の運動を考えた場合、各点の見かけの運動は同一ではないが、対象物体の剛体性に基づき互いに拘束されているはずである。そこで以下では、より直接的に物体の運動と輝度変化の関係について考察する。

照明条件など周囲の環境に変化がないと仮定すると、物体の見えは 3 次元空間中の剛体の位置・姿勢 6 パラメータによって一意に決定され、それに対応する 1 枚の濃淡画像が得られる。ここで、画像空間を N 次元空間、また N 画素を持つ 1 枚の画像^{*1}を N 個の輝度値を要素とするベクトルで表し、それぞれの画像をこの画像空間内の 1 点として表現することとする。すなわち、対象物体の 6 つの位置・姿勢パラメータ p_1, \dots, p_6 から画像空間内の 1 点が決定される。今、時刻 t における画像空間内の画像ベクトルを $I(p_1, \dots, p_6)$ とする。ここで物体の位置・姿勢は連続的であり表面のテクスチャも滑らかであると仮定すると、微小時間変化に対する画像ベクトルの軌跡 (マニフォールド) は連続かつ滑らかとなる。そこでこの軌跡を時刻 t の画像ベクトルの近傍で線形近似することで、式 (2) のオプティカルフロー拘束式と同様に次式が成り立つ。

$$\frac{dI(p_1, \dots, p_6)}{dt} = \sum_{i=1}^6 \frac{\partial I}{\partial p_i} \frac{dp_i}{dt} \quad (3)$$

ここで時間軸について離散化し、差分により近似する。つまり、微小時間の経過によって対象物体の位置・姿勢を表すパラメータが $\Delta p_1, \dots, \Delta p_6$ だけ変化し、それにもとない画像の変化が ΔI だけ生じたとすると、次の近似式が得られる。

$$\Delta I = \sum_{i=1}^6 \frac{\partial I}{\partial p_i} \Delta p_i \quad (4)$$

ΔI は連続画像間の明度変化より得られ、一般に画素数 N は運動の自由度よりもはるかに大きいので、 $\partial I / \partial p_i$ が分かれば、6 つの運動パラメータを直接的に求めることができる。

ここで式 (2) と式 (3) を比べてみると符号が異なることが分かるが、これは式 (2) において対象が x 軸方向に微小移動したときに生じる明度の変化率と、画像の x 軸方向の空間微

分の符号が逆であることによる。式 (2) において画像上いたるところでフローが同一であると仮定して計算することは、式 (3) において対象物体が光軸に垂直な平面上で、かつその面内の平行移動しか行わないと仮定した場合と等価であることから、本手法の与える拘束条件が非常に強力であることが分かる。

2.2 CG 画像を利用した運動パラメータの決定

明度変化から物体の位置・姿勢パラメータを求めるためには、式 (4) の右辺の偏微分 $\partial I / \partial p_i$ を数値的に求める必要がある。Horn らの直接法⁹⁾ ではこれを画像面上の輝度勾配と画像ヤコビアン¹⁰⁾の積により求めることに帰着する。しかし後に述べるように、各画素における画像ヤコビアンを求めることは簡単ではない。そこで我々は、対象モデルから CG 画像が任意の位置・姿勢に対して正確に生成可能であること、つまり任意のパラメータを与えることにより画像空間内の画像ベクトル $I(p_1, \dots, p_6)$ を自由に求めることができることを利用する。すなわち、微小の有有限値 δ を用いて生成した 2 枚の CG 画像から、以下の近似式により偏微分係数を求める。

$$\frac{\partial I}{\partial p_i} \simeq \frac{I(\dots, p_i + \frac{1}{2}\delta, \dots) - I(\dots, p_i - \frac{1}{2}\delta, \dots)}{\delta} \quad (5)$$

具体的に、対象物体の 6 自由度の運動を求める手法 (図 2) について以下で述べる。まず、時刻 t での対象物体の位置・姿勢について生成した CG 画像を $\hat{I}_{[t]}$ とする。さらにこの位置・姿勢を基準とし、世界座標系において仮想的に X 軸方向の正・負の両方向に微小量 $\frac{1}{2}\delta_t$ だけモデルを平行移動させることにより生成した CG 画像をそれぞれ \hat{I}_{+Xt} , \hat{I}_{-Xt} とする。同様に、 Y 軸、 Z 軸に関しての平行移動によって生成された CG 画像をそれぞれ \hat{I}_{+Yt} , \hat{I}_{-Yt} , \hat{I}_{+Zt} , \hat{I}_{-Zt} 、各軸を回転軸として正・負の両方向に微小量 $\frac{1}{2}\delta_r$ だけ回転させることにより生成した CG 画像をそれぞれ \hat{I}_{+Xr} , \hat{I}_{-Xr} , \hat{I}_{+Yr} , \hat{I}_{-Yr} , \hat{I}_{+Zr} , \hat{I}_{-Zr} とする。ここで、微小時間 δt 後にカメラから入力される画像を $I_{[t+\delta t]}$ とすると、式 (4) に基づき次式を得る。

$$\begin{aligned} I_{[t+\delta t]} - \hat{I}_{[t]} &= \frac{\hat{I}_{+Xt} - \hat{I}_{-Xt}}{\delta_t} \Delta X_t + \frac{\hat{I}_{+Yt} - \hat{I}_{-Yt}}{\delta_t} \Delta Y_t + \frac{\hat{I}_{+Zt} - \hat{I}_{-Zt}}{\delta_t} \Delta Z_t \\ &+ \frac{\hat{I}_{+Xr} - \hat{I}_{-Xr}}{\delta_r} \Delta X_r + \frac{\hat{I}_{+Yr} - \hat{I}_{-Yr}}{\delta_r} \Delta Y_r + \frac{\hat{I}_{+Zr} - \hat{I}_{-Zr}}{\delta_r} \Delta Z_r \end{aligned} \quad (6)$$

このとき、 $\Delta X_t, \Delta Y_t, \Delta Z_t, \Delta X_r, \Delta Y_r, \Delta Z_r$ は推定すべき各軸に対する平行移動量と回転移動量である。ここで、

*1 実際には、画像から対象物体の領域を切り出し、その領域内の画素のみを用いる。対象物体の領域は、生成されたすべての CG 画像において対象物体を描画した領域の共通領域として容易に求められる。

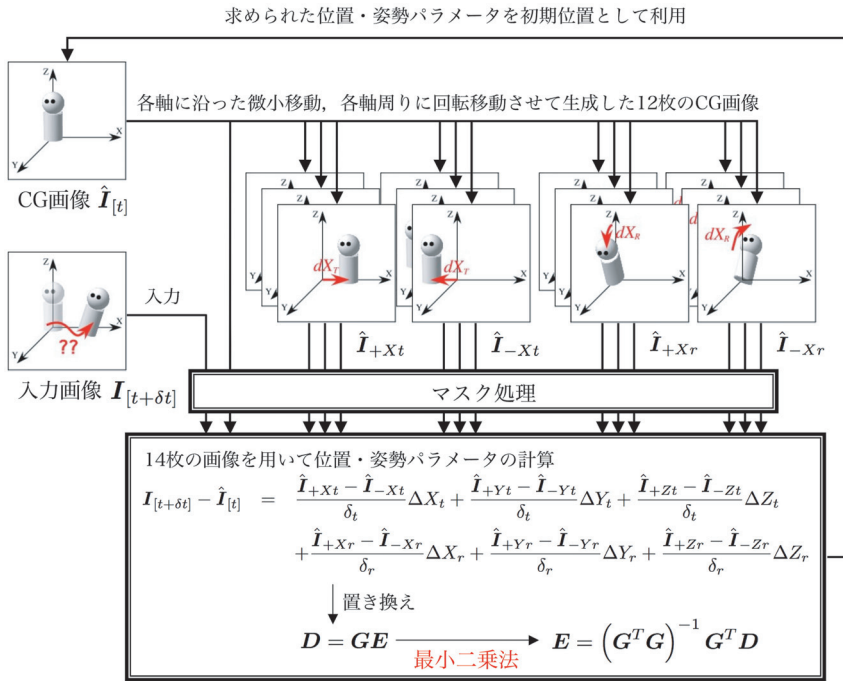


図2 追跡処理の流れ

Fig.2 Flowchart of tracking process.

$$D = I_{t+\delta t} - \hat{I}_t$$

$$G = \left(\frac{\hat{I}_{+Xt} - \hat{I}_{-Xt}}{\delta_t} \quad \dots \quad \frac{\hat{I}_{+Zr} - \hat{I}_{-Zr}}{\delta_r} \right)$$

$$E = (\Delta X_t \ \Delta Y_t \ \Delta Z_t \ \Delta X_r \ \Delta Y_r \ \Delta Z_r)^T$$

とくと、式(6)は

$$D = GE \tag{8}$$

と表すことができる。ここで行列 G は $N \times 6$ 行列であり過制約となるため、以下の最小二乗法によって E を求める。

$$E = (G^T G)^{-1} G^T D \tag{9}$$

これにより、各軸に関する運動パラメータを直接的に推定することができる。求められた運動パラメータは前フレームの位置・姿勢に合成され、対象の現在位置・姿勢を表すとともに次回のCG生成に用いられる。1回の演算により求められる対象物体の回転量は微小であると見なせるうえ、以下で述べるように繰り返し演算により入力画像と生成画像の差を最小化しているため、回転の合成順序を考慮する必要はない。

ここで、画像の差のベクトル D は連続した2フレームの入力画像どうしの差ではなく、生成画像 \hat{I}_t に対する入力画像 $I_{t+\delta t}$ の差としている。そのため、画像上のノイズや線形性の仮定からのずれにより推定運動値に誤差が含まれた場合も誤差が蓄積することなく、この演算を繰り返すことによりCGモデルは対象の位置・姿勢に近づいていく。局所解の存在については対象物体表面のテクスチャや初期位置・姿勢に依存するため一般的に論じることは難しいが、初期値が真値に対し十分近い場合は問題とはならないことが実験的に確かめられている。

以下にあげる実装例では必ずしも6自由度の運動を求めているわけではないが、一般に n 自由度の運動を求めるときには、各自由度ごとに正と負の微小変化をそれぞれ与えるために計 $2n$ 枚、さらに入力画像とモデル画像の差 D を求めるために微小変化を与えていない場合の画像を1枚の、合計 $2n+1$ 枚のCG画像を1フレームの追跡処理につき生成することになる。

3. 実装例

この章では我々がこれまでに提案手法を用いて実装した運動追跡の例について概要を示す。

3.1 多視点画像による自由曲面物体の6自由度運動追跡⁴⁾

2章で述べた動物体追跡原理は単一のカメラにおいても機能する手法であるが、対象物体の大きさや撮影条件によっては6自由度の運動を安定に行うことができない場合がある。たとえば物体が光軸方向に遠ざかったり近づいたりする場合、もし対象物体が小さく焦点距離の長いレンズを用いたならば、運動にともなう画像上の変化はわずかなものになってしまう。また対象物体の形状によっては、対象物体の縦軸周りの回転と水平方向の平行移動のオプティカルフローが非常に似通ったものになることがある。そこで我々は提案手法を多視点観測に拡張することで、6自由度の位置・姿勢の推定を安定に行うことができることを示した⁴⁾。

3.1.1 多視点画像による6自由度運動追跡結果

図3のように対象物体を取り囲むように4台のカメラを設置し、それにより対象物体を

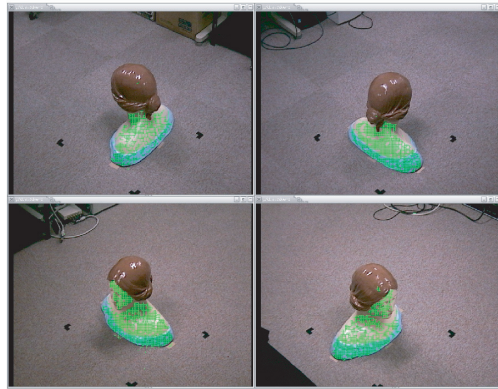


図3 4方向の視点から得られる画像
Fig.3 Multiple viewpoint images.

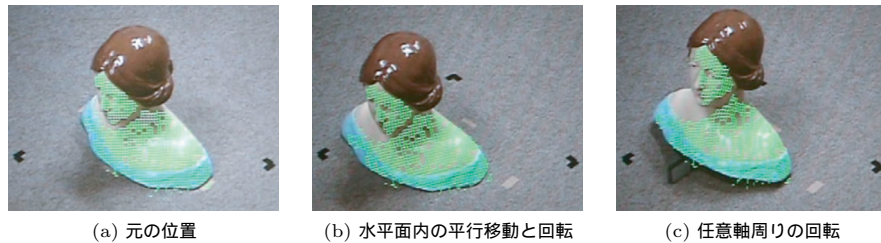


図4 多視点画像による6自由度運動の追跡結果
Fig.4 Results of tracking of 6DOF motion using multiple camera.

撮影する．対象物体の運動を推定するには、これら4枚の画像のすべてについて、生成画像と入力画像の差が小さくなることを望ましい．そこですべての入力画像から対象物体領域内の画素の輝度値を取り出し、それを単純に並べたベクトルを用いて式(9)の計算を行うことで追跡を行う．追跡結果を図4に示す．また、この対象物体をスライドステージまたはターンテーブルに載せ、平行移動量または回転量の真値と推定値の比較を行ったときの結果を図5に示す．この図のように対象物体の高さが42cm、カメラから対象物体までの距離がおおよそ2.4mという条件において、平行移動の推定誤差は±1.5cm以内、回転の推定誤差は±0.3度以内であった．それぞれのカメラ画像間でエッジ検出や対応点の計算を行う必要はなく、カメラの配置は完全に自由である．しかし前述のように、単眼視では安定

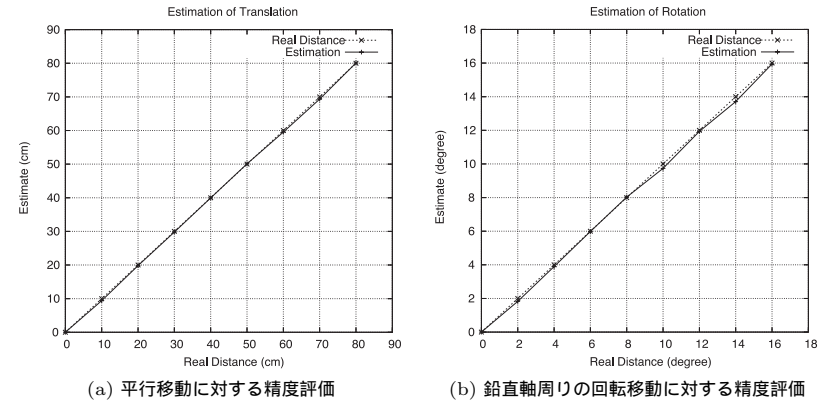


図5 精度評価
Fig.5 Evaluation of accuracy.

な推定が困難な運動が他の視点から観察されることで6自由度の運動推定を安定に行うことができ、これは式(9)における行列 $G^T G$ の条件数*1からも示されている⁴⁾．

3.1.2 PC クラスタを用いた並列処理による高速化

この研究では同時に、PC クラスタを用いた演算処理の並列化についても検討している．すべての入力画像をホスト計算機に転送し、集中的に処理する方式はネットワークやCPU負荷の観点で望ましいとはいえない．そこで式(9)を多視点画像について以下のように展開し、並列処理を行う．

$$E = \left(\begin{bmatrix} G_1 \\ \vdots \\ G_n \end{bmatrix}^T \begin{bmatrix} G_1 \\ \vdots \\ G_n \end{bmatrix} \right)^{-1} \begin{bmatrix} G_1 \\ \vdots \\ G_n \end{bmatrix}^T \begin{bmatrix} D_1 \\ \vdots \\ D_n \end{bmatrix} = (G_1^T G_1 + \dots + G_n^T G_n)^{-1} \cdot (G_1^T D_1 + \dots + G_n^T D_n) \quad (10)$$

ここで n はカメラもしくは計算機の台数を表す．この式から、入力画像と生成画像の画素を集めたベクトルを構成する代わりに、それぞれのカメラもしくは計算機で独立に撮影・生成し、負荷の高い計算を分散して行うことができることが分かる．この式において $G_i^T G_i$ は6次の正方行列、 $G_i^T D_i$ は6次のベクトルであり、これをホスト計算機に集めてから姿

*1 最大固有値と最小固有値の比．

勢 E を計算するが、このときのネットワーク伝送負荷およびホスト計算機の CPU 負荷は無視できる程度に小さいことが分かる。

図 3 から分かるように、対象物体は頭部から肩部への影が生じ、また塗装のために鏡面反射が観察される。しかし運動推定に多くの画素からの情報を統合していることなどから、それらによる顕著な悪影響は見られない。

3.2 ステレオカメラを利用した自己位置・姿勢の実時間推定とシーンの同時モデリング⁵⁾

本論文で提案する動物体の運動推定手法では、対象物体やシーンのモデルがあらかじめ得られていると仮定している。このモデルから生成された CG 画像は実画像と比較され、その差が最小となるように運動パラメータが推定されるため、CG 画像には見た目に自然であることが要求されるのではなく、実画像と CG 画像の間で幾何学的および光学的整合性が満たされていなければならない。我々が提案してきたこの動物体追跡手法に対してはこれまで、そのような CG 生成に耐えうる対象物体の正確なモデリングが困難であろうという予想に基づく疑義が多く呈されてきた。そこで我々は、ステレオカメラを用いた簡便な対象のモデリング手法を動物体追跡手法に組み合わせ、シーンに対するカメラ位置・姿勢の実時間推定と、シーンのモデル化を同時に行う、いわゆる SLAM (Simultaneous Localization and Mapping) を実現する手法を提案した⁵⁾。この手法の特徴は以下のとおりである。

- 単眼による自己位置・姿勢推定を行いながら、必要に応じて自動的にステレオ計測を起動することで間欠的にシーンのモデル獲得を行う。
- マルチプロセッサシステム上で追跡処理とステレオ処理を並列実行することで、ステレオ処理に起因する追跡処理のフレームレートの低下を引き起こさないようにする。
- ステレオ計測のための画像の取得から、それにより生成されるシーンのモデルが使用可能になるまでの間のカメラ運動を求めることで、モデル入れ替え時の追従失敗を防止する。
- ステレオ計測のために取得した画像に対し位置合わせ処理を繰り返すことで、モデルの位置合わせ精度を向上させ、それにより誤差の蓄積を抑制する。

3.2.1 モデリングと追跡の同時処理の概要

モデリングと追跡の同時処理の概要を以下に示す。まず、ステレオカメラを用いてシーンのステレオ計測を行う。これにより得た距離画像と濃淡画像はそれぞれ、シーンの 3 次元形状モデルとテクスチャとして CG 画像の生成に用いられる。その後カメラ運動の計測には、ステレオカメラのうち代表となる 1 台のカメラからのみ画像を入力し用いることでデータ量を削減し、計測を高速化する。

カメラが移動を続けると、最初にモデリングされた領域から視野がずれていくため、追跡に用いることができる画像上の領域が減少する。これを自動的に判定し、モデリング処理を起動することで新たなモデルを作成し、計測を続行する。モデリングのつど得られた 3 次元形状モデルはそれぞれファイルに記録され、またそれらの間の位置関係はモデリング時のカメラ姿勢により求められているため、それらを統合してシーン形状とすることができる。もちろん ICP¹¹⁾ などの手法により、得られたモデル間の位置関係を高精度化することも可能であると考えられるが、このようなオフライン処理はこの研究では扱っていない。

ステレオ計測には市販のステレオカメラを用いた。そのため、各カメラから得た複数の画像間の対応点探索や 3 次元座標の計算はステレオカメラに付属の API を用いた。この API から得たパラメータにより、入力画像と形状データの幾何学的整合性は容易に満たされる。また同じくステレオカメラから得た濃淡画像をそのままテクスチャとして用いることで光学的整合性が満たされると仮定した。これはシーン表面で鏡面反射が観察されないと仮定することにほぼ等しい。

ステレオ法では対象物体上のすべての点にわたり、安定に奥行き値が求められるとは限らない。そこで奥行き値の信頼度が低い点に関してはモデル生成から除外し、そのような点により追跡処理が乱されることがないようにする。具体的には、ステレオカメラに付属の API の機能のうちテクスチャ検証 (texture validation) と唯一性検証 (uniqueness validation) を用い、これらにより得た信頼度をしきい値で判定し頂点を取り除く。テクスチャ検証は画像から切り出した小ブロックの内部の明度値が平坦で、テクスチャが不十分な領域を検出するための機能である。また唯一性検証は対応点を探索したときに、最も良く一致した点とその次に良く一致した点との一致度の差が小さい場合に信頼度が低いと見なし取り除くための機能である。これらの手法を用いると、対象物体のモデルは多数の穴を持つものとなる。これは見た目には不自然であるが、運動追跡にはかえって好都合である。なぜなら、たとえば明度一定の平坦な壁面は微小移動による明度変化がなく、運動追跡には寄与しない。そこでそのような画素を省くことで運動追跡に有用である明度変化を豊富に持つ部分のみを用いた処理が行われ、計算量が削減される。

先に述べたように、追跡に用いる 3 次元形状モデルと入力画像のずれが大きくなるとシーンの再計測が起動され、シーンの新たなモデルが得られる。しかし再計測が起動されてからモデルが得られるまでの間追跡処理が停止されていると、2 章で述べた微小運動の仮定が崩れ、追跡に失敗することがある。そこで我々はシーンのモデリングの最中にも更新前のモデルを用いて追跡を続行することで、モデリング後に過大なずれが生じないようにした。

3.2.2 モデル更新時の誤差蓄積の抑制

モデルを更新すると、それ以降のカメラの位置・姿勢は更新されたモデルからの相対運動として求められる。そのため、モデル更新時のカメラの位置・姿勢の誤差、すなわち更新前と更新後のモデルの間の相対的な位置関係の誤差が大きいと、それがモデル更新のたびに蓄積され、長時間の追跡の後には大きな誤差を生じるおそれがある。したがって、モデル更新時のカメラの位置・姿勢は特に精度良く推定する必要がある。そこでモデル更新時のステレオ計測に用いる複数の濃淡画像については 2 章で述べたカメラの位置・姿勢の推定処理を複数回繰り返すことで精度を向上させる。

以上の 2 点について図 6 に沿って詳説する。図中 (a) や (b) では追跡処理のみが行われている。ここでの追跡処理の 1 サイクルは 2 章で述べた計算を 1 回行うことを示している。しかし (b) の処理において、モデリング処理を起動する条件が満たされたとすると、次の画像入力 (c) のように単一の濃淡画像ではなくステレオ計測のための複数画像が入力されることになる。このステレオ画像を用いて対応付け計算 (f) が行われる。また同時に、このステレオ画像が入力されたときのカメラの位置・姿勢は (d) の追跡処理により求められる。このときのカメラの位置・姿勢は更新前後のモデルの相対的な位置関係となるため、特に誤差を抑制する必要がある。そこで (d) から (e) までの間、2 章で述べた運動推定処理を繰り返すことで誤差を減少させ、通常時よりも高い精度でカメラ位置を求める。この繰り返し回数は任意であるが、我々の用いたシステムでは画像間の対応付けにより距離画像が計算されるまでは新たな画像の入力が不可能であるため、その間繰り返すこととした。

距離画像がカメラのステレオ計算 API から出力されると、その後画像の入力を再開し、(h) のように新たな画像の入力を行う。しかしこの時点ではまだ、ステレオ計測により求めた新たなモデルは使用できない。なぜなら、得られた距離画像を CG モデルに変換する処理 (g) が必要だからであり、その間の (i) の追跡処理では (h) で得られた新たな画像に対し、

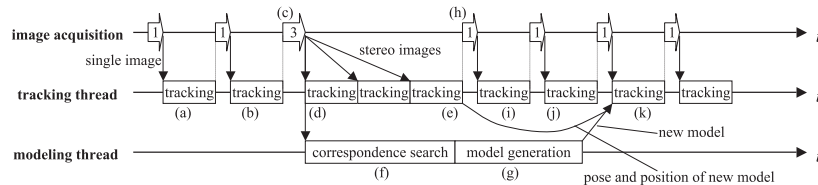


図 6 追跡とモデリングの並列処理

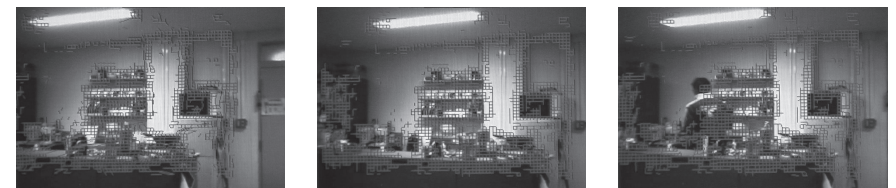
Fig. 6 Parallel execution of tracking and modeling.

古い 3 次元形状モデルを用いて追跡を行う。これにより新たなモデルが利用可能になるまでの間のカメラ運動を求めることができ、モデル更新時に CG 画像と入力画像の差が過大になることによる追跡の失敗を防ぐことができる。(g) により生成した新たなモデルは追跡処理 (k) から用いるが、その際に更新前と更新後のモデルの相対位置・姿勢には (e) で求められた高精度な姿勢情報を用い、他方、追跡処理 (k) におけるカメラ姿勢の初期値には (k) の直前の追跡処理 (j) により求められた最新の値を用いる。このように、ステレオ計測が開始されてから新たなモデルが使用可能になるまでの間に、誤差の蓄積を防ぐための処理とモデルを更新するときの追跡の不安定化を避けるための処理を行う。

3.2.3 実験結果

リアルタイム追跡の評価には安定した繰り返し運動が必要であるため、ロボットアーム PA10-6C (三菱重工業製) の先端にカメラを固定して実験を行った。また処理に用いる画像の解像度を 320×240 (QVGA) に設定した。実験に先立ち、追跡とモデリングに要する計算処理時間をそれぞれ調べたところ、運動の追跡はモデリング処理の間に少なくとも 5 回繰り返すことができることが分かった。

実験ではまず、2 点間の平行移動を繰り返したときに生じる視野の変化を、モデルの更新処理を行うことなくリアルタイムで追跡した (図 7 (a)). ステレオ計測によって生成された 3 次元モデル (重畳されたワイヤフレーム表示) が視野の変化に追従しており、カメラ運動の追跡に成功していることが分かる。このとき 2 点間の距離は約 600 mm であり、カメラが 1 往復にかかる時間は約 30 秒であった。また、フレームレートは約 9 frame/sec であった^{*1}。この際、図 7 (b) に示すように人物などが視野内に侵入した場合においても、本手法



(a) カメラ運動より変化する視野の追跡

(b) 人物が視野内に入ったときの追跡

図 7 リアルタイム追跡結果

Fig. 7 Results of real-time tracking.

*1 現在の計算機では 2 倍以上高速化されている。

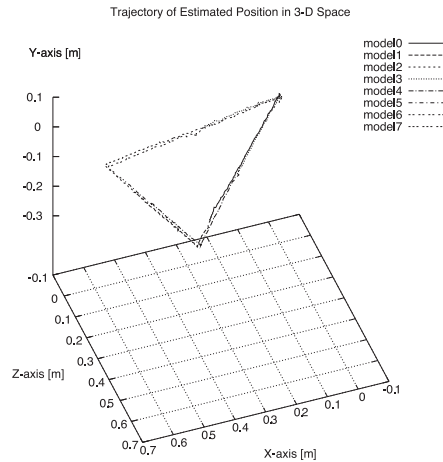


図 8 カメラ位置の推定軌跡

Fig. 8 Trajectory of estimated position.

ではその侵入領域以外にも運動推定に利用できる十分な画素が存在するため、大きな影響を受けることなく比較的安定した追跡が実現されている。

続いて、辺の長さが約 600 mm, 500 mm, 400 mm の三角形の軌跡を描くようにカメラを平行移動させ、前説で述べたモデルの更新処理を実装した場合でのリアルタイム追跡を行った。このときカメラが 1 周する間にかかる時間は約 34 秒であり、これを 2 周させた場合に推定されたカメラ位置の軌跡を図 8 に示す。2 周する間に 7 回のモデリングが行われているが、1 周目と 2 周目の軌跡がほぼ重なっており、カメラの位置が精度良く推定されていることが分かる。ほかに定量的な実験を行った結果、モデル更新時の誤差の蓄積を抑制できていることが分かっている⁵⁾。またカメラ運動に回転が含まれる場合についても実験を行ったが、ほぼ同程度のカメラ運動の位置軌跡を得ることができた。

3.3 距離画像を用いた実時間動物体追跡と全周形状モデリング

本論文で述べる動物体の位置・姿勢の追跡法は、入力画像と生成画像の差の最小化に基づく。そのため、入力画像や生成画像は通常の濃淡画像である必要はなく、同じ位置・姿勢の物体について取得または生成したものが一致するならばどのようなデータでもよいことになる。そこで以下この章では、距離画像について提案手法を適用し、動物体の 6 自由度運動を実時間的に追跡した例を示す。なお、2 章で述べた運動推定の原理は、もともとはこの章

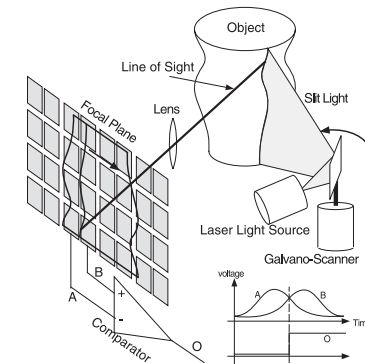


図 9 シリコンレンジファインダの計測原理
Fig. 9 Principle of Silicon Range Finder.

で述べる距離画像を用いた位置合わせのために我々が考案したものである。

我々は 1990 年代初頭から通常の画像センサでなく、専用のフォトセンサを用いて高速に距離画像を計測する方式の開発を進めてきた。その成果の 1 つとして、縦・横各 24 画素と画素数は少ないものの、毎秒 30 フレーム以上のフレームレートでかつ奥行き精度の高い距離画像を実時間的に取得することができるスリットレーザ光投影方式のレンジファインダ「シリコンレンジファインダ」を開発した¹²⁾。シリコンレンジファインダの動作原理を図 9 に示す。センサは専用に開発製造されたフォト LSI を用いており、2 つのフォトセルと増幅器などの回路を各画素ごとに備えている。図のようにスリットレーザ光はガルバノスキャナにより偏向され、その像をセンサが観測するが、2 つのフォトセルの境界をスリット像が通過する時刻を各画素の比較型アンプが高精度に検出し、センサからは列並列に読み出すことで高速な距離画像計測を可能としていた。

我々はこのセンサを用い、自由に運動する不定形物体の運動を実時間的に計測しつつ、得られた距離画像を統合することで対象物体の全周形状を求める手法を開発した⁶⁾。処理の概略を図 10 に沿って述べる。図 10(a) は実験に用いた対象物体である。これをシリコンレンジファインダにより計測し、モデル化したものが図 10(b) である。メッシュモデルの頂点がシリコンレンジファインダにより計測された対象物体上の計測点であり、この間にポリゴンを張ることで対象物体の初期形状モデルを得る。次に対象物体の姿勢が変化すると、図 10(c) のように新たに得られた距離画像(メッシュモデルにより表示)と計算機内の形状モデル(陰影付けされたポリゴンモデルにより表示)の差が最小化され、計算機内形状

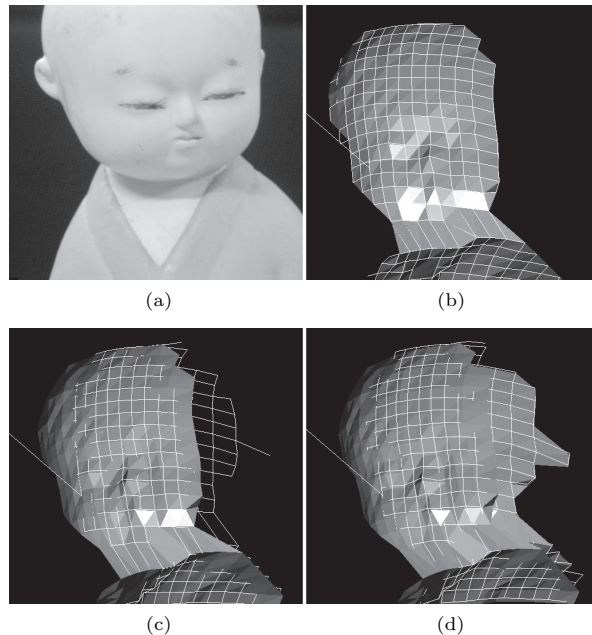


図 10 距離画像による運動推定と全周形状計測
Fig. 10 Simultaneous tracking and modeling using range image sequence.

モデルの位置・姿勢が推定・更新される。対象物体の回転により初期姿勢では観測されなかった領域が可視となり、その部分の形状が計測されるが、そのような領域の大きさが一定以上になるとその部分をモデルへ統合し、新たな形状モデルを得る(図 10(d))。なお、図 10(b)–(d)において左上へ引かれた直線は、センサの光軸を表している。対象物体を鉛直軸周りに 1 回転させ、運動追跡と同時に復元した全周形状モデルを図 11 に示す。位置の推定精度(標準偏差)は 0.17 mm、また初期形状モデルに対して対象物体を 30 度回転させたときの各点の奥行き差の標準偏差は 1 mm 以下であった。

この研究で用いた運動推定の原理は 2 章で述べた手法とまったく同じであるが、利用している画像が濃淡画像ではなく距離画像であるという点のみが異なる。つまりそれぞれの画素において、センサから得られた奥行き値と、形状モデルのレンダリングによって求めた奥行き値の差を求め、これを最小化する。形状モデルからの距離画像のレンダリングには CG

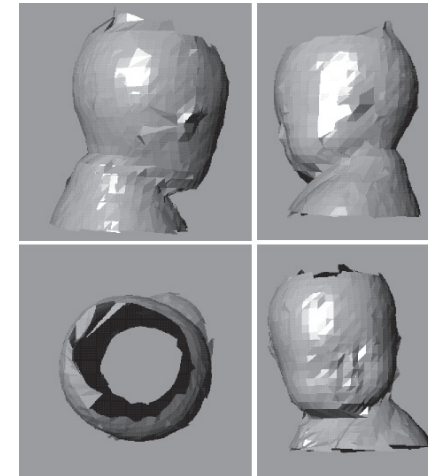
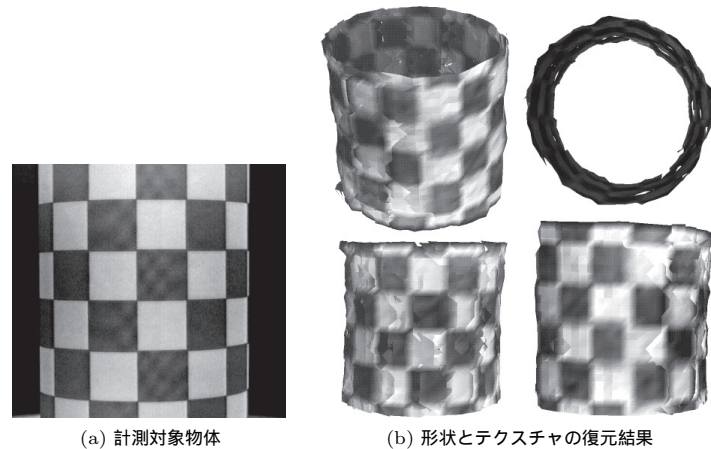


図 11 復元された全周形状
Fig. 11 Reconstructed shape of the object.

ハードウェアに搭載されている Z バッファ値を利用することで高速化を図っており、GPU を用いた画像処理の高速化という観点でも萌芽的な研究例の 1 つである¹³⁾。それにより、CPU のクロック周波数が 100 MHz 前後であった 1996 年当時の計算機 (SiliconGraphics Iris Indigo2 Extreme) により毎秒 7 フレームのフレームレートで 6 自由度運動の追跡が実現されていた。

3.4 距離画像と濃淡画像の融合による実時間動物体追跡と全周形状モデリング

シリコンレンジファインダの光学系にはプリズムが搭載されており、距離画像を計測するためのフォト LSI と同一視点からの濃淡画像を取得することができる。そこで我々は、このセンサから得られる距離画像と濃淡画像の双方を用い、物体の運動追跡と同時に全周形状とテクスチャのモデリングを行う研究を行った⁷⁾。ここでは 3.1.1 項で述べた、複数カメラから得た画像の統合方法と同様に、距離画像と濃淡画像の各画素値を単に順に並べたデータを用い、その差が最小となるような位置・姿勢を追跡する方法とした。ただし距離値と輝度値はスケールが異なるため、一方の画像をスカラー倍することで両画像の影響度を重み付けした。また物体表面上の輝度は物体の回転にともない変化するが、距離画像が得られていることを用いて簡易的に反射率の推定を行った。すなわち対象物体を拡散反射であると見なし、既知の光源方位と距離画像から得た法線方向から拡散反射率を推定する。また画



(a) 計測対象物体

(b) 形状とテクスチャの復元結果

図 12 距離画像と濃淡画像の融合による運動追跡と全周形状モデリング

Fig. 12 Simultaneous tracking and modeling using range and intensity image fusion.

像のレンダリング時には、同じ光源方位から照明したような画像を CG レンダリング API (IrisGL) により生成することで入力画像と生成画像の光学的整合性を満たした。なお反射率の推定にも IrisGL を用いており、いったん、反射率を一定として CG をレンダリングし、これと入力画像との比を各画素について求めることで反射率を得た。

実験に用いた対象物体の写真を図 12 (a) に示す。円筒はその回転軸に沿った平行移動と、回転軸周りの回転については形状が不変であるため、距離画像のみでは運動の復元が不可能である。そこで図に示すように対象物体にチェッカボード状のテクスチャを与え、これをターンテーブル上で回転させることで 6 自由度の運動推定とモデリングを同時に行った。これにより復元された対象物体のモデルを図 12 (b) に示す。濃淡画像は距離画像センサの計測点 (縦・横各 24 画素) と一致した画素の値しか用いていないため、復元されたモデルのテクスチャの解像度は低いものとなっているが、チェッカボード状のテクスチャの接続関係に破綻はなく、正常にモデルが復元されたことが分かる。このときの追跡速度は毎秒約 5 フレームであった。

4. 鏡面反射物体の追跡

鏡面反射は表面を高精度に仕上げた金属製品などに生じるため、応用上の観点では、特に

工業分野において計測手法の確立が求められている¹⁴⁾。しかし前述のように鏡面仕上げが施された物体は、その表面を介して周囲環境が映り込むため、物体そのものに関する情報のみを独立に画像から抽出することは容易ではない。特に、鏡面反射を対象とした形状計測手法はいくらか見られるが^{15)–18)}、鏡面物体の位置・姿勢の推定や追跡に関する研究は少ない¹⁹⁾。その理由の 1 つとして、曲面からなる鏡面反射物体上の反射像の移動は非常に複雑になる¹⁸⁾ ことがあげられる。しかし本論文で述べる運動追跡手法を用いることで、この逆問題を解析的に取り扱うのではなく、順問題である画像のレンダリングと最適化により解くことができる。

鏡面反射は磨かれた金属や鍍金が施された物体に生じる正反射のほか、表面が滑らかに仕上げられた非金属物体に生じる 2 色性反射において観察される。このような物体では画像として観測される輝度は拡散反射成分と鏡面反射成分の混合となり、前述のように特徴点追跡などによる幾何学的な枠組みでの厳密な取扱いはいっそう困難となる。我々は、このような物体についても本論文で提案するレンダリングと最適化の枠組みにより同様に運動追跡が可能になると考えている。この章では 2 色性反射物体ではなく鏡面反射のみが観察される物体を対象としているが、その理由は以下のとおりである。第 1 に、3.1 節で述べた単純なテクスチャマッピングを用いている従来研究⁴⁾ ではテクスチャに乏しく、また表面につやを有するような物体における追跡成功例が示されていることがあげられる。そのため、鏡面反射と拡散反射が重畳されて観測される対象について実験を行う方法では、鏡面反射成分を利用したことの効果が分かりにくいと考えた。第 2 の理由として、鏡面反射と同様により高度な光学現象に関する議論は異方性反射物体や近接光源、影のレンダリングなどにおいても可能であるが、本章の主旨はこのような高度なモデリング・レンダリングの追求により追跡可能な物体の種類や状況をできるだけ多く示すことではないからである。本論文の目的は鏡面反射成分をはじめとする高度な光学現象を外乱的な要因として取り扱うのではなく、より高度なレンダリング技法を提案手法に組み合わせることで、これらの光学現象を物体の位置・姿勢推定の手がかりとして利用できることを示すことである。

以下では、動物体の運動追跡の原理として、CG 画像と入力画像の差の最小化に基づく手法について述べた後、4.3 節において鏡面反射物体の運動追跡について述べる。

4.1 環境マッピング

物体表面で鏡面反射しているように CG 環境で周囲の映り込みを再現するために、本論文では環境マッピングを行う。環境マッピングは鏡面反射物体の表面から周囲環境までの距離が十分遠いという前提に基づいた近似であり、図 13 に示すように視点から表面における

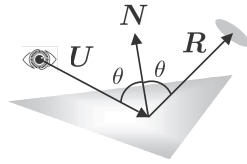


図 13 鏡面反射特性
Fig. 13 Specular reflection.

頂点への単位ベクトル U 、頂点における物体表面の単位法線ベクトル N により以下の式で計算されるベクトル R の方向からの入射光を求めることによって反射は表現される。

$$R = U - 2(N \cdot U)N \quad (11)$$

この環境マッピングを実現する方法としてスフィアマッピングやキューブマッピングなどの手法が知られているが、本論文で扱う動物体追跡では視点の位置が環境に対して変わらないため、高速に演算でき、環境マップの獲得も簡単であるスフィアマッピング²⁰⁾により物体表面の映り込みを計算することとする。

スフィアマッピングのための周囲の環境マップを得るためには、テレセントリックレンズのような平行射影により球面鏡を撮影することが望ましいが、球面鏡の見かけの画角が過大でなければ中心射影による誤差は無視できるため、我々は物体の追跡に利用するカメラの画像中心付近に配置した球面鏡を撮影してこれを得た。この環境マップを利用し、OpenGL に組み込まれた機能によりレンダリングすることで CG 環境での鏡面反射による周囲環境の映り込みを高速かつ容易に再現できる。スフィアマッピングの問題点として、環境マップ画像上においてカメラからの勾配が大きい部分の分解能が低くなることがあげられるが、このような領域が用いられるのは追跡対象のうえでも遮蔽輪郭に近い部分であるため問題とはならない。

4.2 追跡実験と評価

この章に示す実験では、SONY 製カメラ EVI-G20 (768 × 494 画素) と PLUS 製プロジェクタ V-339 を用いた。カメラとプロジェクタによってレンジファインダが構成されており、追跡対象の CG モデルを生成するための形状計測を行う。対象物体の形状 (奥行き) だけが必要であった従来研究⁴⁾とは異なり、鏡面物体のレンダリングには正確な法線情報が必要であるため、我々は従来法で用いていたグレイコード法による物体形状計測ではなく、スリット光を順次投影する手法により物体形状を計測した。画像上の各点においてスリット通過時の明度変化を記録し、その明度を重みとしたスリット番号の重心を求めることでサブ



図 14 鏡面半球と鏡面反射物体
Fig. 14 Specular hemisphere and object.

ピクセル単位でのスリット位置を求めた。これにより得た法線ベクトルは、さらに近傍画素との間で平滑化を行い、滑らかに変化する法線ベクトルを得た。

プロジェクタは形状計測のみに用いられ、物体の運動を求めるときには用いられない。カメラからの画像の入力と CG 画像の生成はともに VGA 解像度 (640 × 480 画素) で行う。スフィアマッピングに用いる環境マップの取得には、画像の中心付近に球面鏡を設置し、画像を 1 枚撮影することにより取得する。球面鏡の領域の切り出しは手動で行った。

鏡面反射物体の追跡実験に用いる対象物体を図 14 右下に示す。この対象物体の大きさは約 160 × 35 × 80 mm (W × H × D) である。ただし表面が鏡面である物体の形状は、通常の光投影法では計測が困難であるため¹⁴⁾、図 14 の右上のように白色に塗装した同形状の物体を利用して 3 次元形状の計測を実施した。これにより得た形状モデルにスフィアマッピングを施した結果を図 15 に示す。(a) は鏡面半球 (図 14 左) を利用して取得したスフィアマッピング用の環境マップ画像であり、(b) は実物体を鏡面半球の設置場所に置いて撮影した画像、(c) は同一の位置・姿勢に形状モデルを仮想的に配置して生成した CG 画像である。テクスチャの解像度が低いものの、写り混んだ物体の配置や位置は一致していることが分かる。

続いて、鏡面反射物体の追跡実験について述べる。ここでは物体を微小に移動させながらカメラで撮影した画像をハードディスクに保存しておき、それを読み出しながら追跡を行うオフライン処理によって実験した。

まず、スライドステージを利用して平行移動に関する実験を行った。X-Y 平面内のある

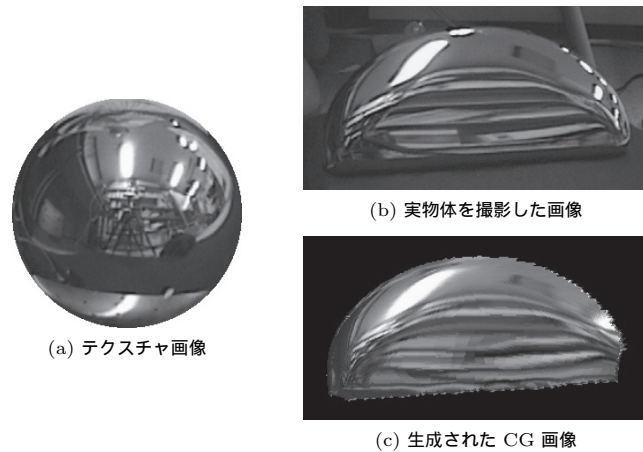


図 15 スフィアマッピングの結果
Fig. 15 Results of sphere mapping.

1 軸に沿って物体を初期位置から 5 mm 間隔で 200 mm まで移動させ、それぞれの位置で CG モデルの位置・姿勢パラメータが収束したときの値を記録する。同様に、ターンテーブルを利用して回転運動に関する実験を行い、Z 軸周りに初期姿勢から 0.5 度間隔で 40 度回転させて各角度での推定された回転角を記録した。どちらの場合も、物体の運動は平面内の運動（平行移動 2 軸と、その面に直交した軸周りの回転運動 1 軸）の 3 自由度と仮定し、その空間内での運動を算出している。オフライン処理ではあるが、フレームレートを計測したところ平均して約 15.5 frame/sec であった。

追跡結果を図 16 に、また精度評価結果を図 17 に表す。図 16 から、物体の移動にともない物体表面に映り込んだ周囲環境が物体に対しずれている様子が分かる。(a) の平行移動の場合では、空間に対する法線ベクトルの方位は不変であるが、カメラの画角効果によりカメラから物体への方角が変動することで、映り込みが移動する。また (b) の回転の場合では、最終的な回転角度が 40 度と大きいこともあり、映り込みには大きな変化が見られる。図 16 では、推定された物体の位置・姿勢をもとに赤色のメッシュ状の画像を合成表示している。これによると、物体の見え方の大きな変動にかかわらず物体に対し計算機内モデルが適切に追従していることが分かる。

精度評価のグラフ（図 17）では、横軸は物体の実際の移動量を、縦軸は推定値を表して

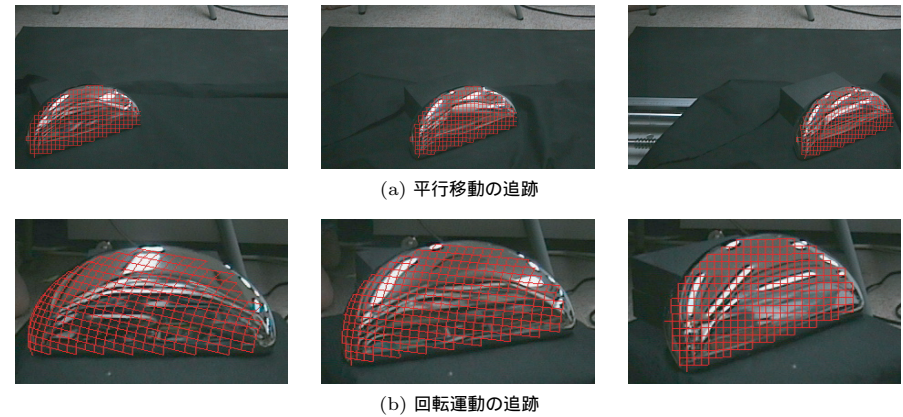


図 16 鏡面反射物体の追跡
Fig. 16 Tracking of specular object.

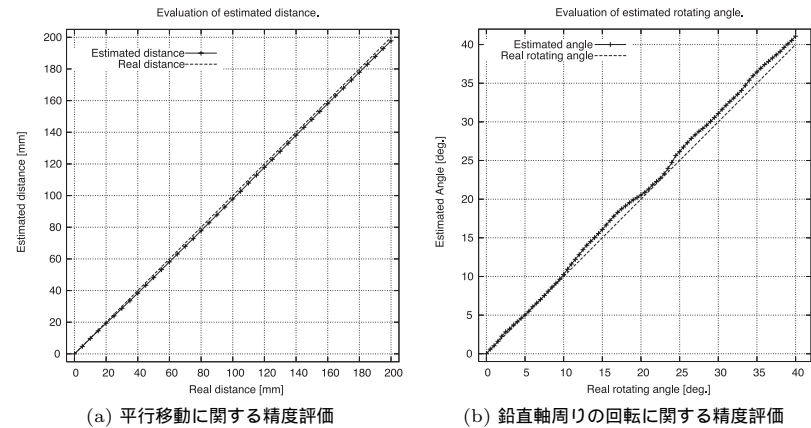


図 17 精度評価
Fig. 17 Evaluation of translation and rotation.

いる。(a) の平行移動に関する評価では、推定値は最大誤差で 3 mm 以内であり、物体の総移動量が 200 mm、物体の最大長が 160 mm であることや、図 16 から分かるように見かけのテクスチャが大きく変動していることを勘案すると、提案手法の効果が十分認められる。

(b) の回転運動に関する評価では、(a) の平行移動の結果に比べるとグラフに凹凸が見られるが、最大誤差は 1.5 度以内であった。どちらの場合も、対象物体の見かけの向きの変化に対し、鏡面反射方向はその倍だけ変化することに注意されたい。

4.3 鏡面物体追跡に関するまとめ

2 章で述べたように、本論文で提案する手法では動物体の 6 自由度剛体運動を求めることが理論上可能である。しかしこの章では 3 自由度の運動推定の結果を示した。その理由として、単眼視における追跡の安定度の問題があげられる。実験では周囲の照明環境を取り込むために球面鏡を用い、これを環境マッピングすることにより画像のレンダリングを行った。そのため、鏡面物体と環境の間の距離は物体の大きさや物体の移動量に対し十分大きいと仮定していることになる。この場合、カメラ光軸方向の対象物体の平行移動を精度良く求めることが困難になる。なぜなら物体の移動にともなう画像の変化がごくわずかになるからであり、特にカメラから対象物体までの距離が大きく、シーンから画像への射影が平行投影に近くなる場合は顕著である。このような場合、従来研究⁴⁾において指摘したように、複数のカメラにより異なる方位から対象物体を観察することが有効であると考えられる。また対象物体と周囲環境の距離が小さい場合は、単純な環境マッピングの代わりに、より厳密な反射計算を行う方法が考えられる。なぜなら環境マッピングは反射の計算において方位のみを扱っており、近接物体の反射を正確に再現することができないからである。CG のレンダリング技術は長足の進歩をとげており、GPU の機能を用いることで近接物体の反射を高速かつ正確に再現することも可能となっている。そこで環境の 3 次元形状をモデリングすることができれば、レンダリング手法を更新するだけで本手法はそのまま近接物体の反射が生じる場合の運動追跡に利用することができる。

シーンの照明条件の取得、およびそれを用いた画像のレンダリングについて取り扱うとき問題になる要素としてカメラやレンダリング段階のダイナミックレンジがあげられる。実際に、本論文のシステムにおいてレンダリングされた画像では、輝度の正確性を欠いている。しかし本論文の手法では入力画像とレンダリング画像の一致を基本としており、照明のようにきわめて輝度の高い部分では画素値の飽和が生じて、照明環境取得時に利用した球面鏡と追跡対象との反射特性に差がなければ、大きな問題とはならない。なぜならどのような曲率の球面鏡を撮影しても、反射により観察された対象物体の輝度は不変であり、環境マップ取得時に飽和した画素値は、追跡時にも同じように飽和するからである。ただしこの章で扱うような完全な鏡面反射物体とは異なり、滑らかなプラスチックのような表面では鏡面反射の強度が弱くなるため、照明環境取得時のダイナミックレンジや飽和について配慮する

必要があると考えられる。

この章では対象物体の位置・姿勢のみが変化することを前提に述べたが、同様の手法はカメラの移動についても適用することができると考えられる。ただし本論文で用いたスフィアマッピング法は、カメラと光源の位置関係が一定である場合に効率的なレンダリングが可能となる手法であり、カメラの運動を取り扱う場合には他のレンダリング手法を用いる必要があると考えられる。照明環境の運動推定や、鏡面物体とカメラの位置関係が固定されているような場合についても同様である。たとえば曲面ミラーを用いた全方位カメラ²¹⁾において自己姿勢（動揺）の変動を求めるとき、通常は射影の変換と特徴点の追跡により行われるが、本手法によれば特徴点の抽出や、通常の透視投影画像への変換を行うことなく姿勢推定を行うことが可能であると思われる。

5. 考 察

この章ではこれまでに述べた計測原理とその実装例をもとに、他の研究との関係やそれに対する独自性などについて考察する。

5.1 勾配法⁸⁾ との関係

本論文で述べた計測原理は、2.1 節の議論からも分かるように、明らかに勾配法との強い関連性を持つ。勾配法では対象物体の移動にともなう各点の輝度の変化から対象物体の運動を求めようとする手法であり、本質的に画素の並び順は大きな意味を持たない。ただし勾配法ではその点の勾配を求めるために隣接画素との間の差分を求める必要があり、また滑らかさの拘束などでも画素の近接性は重要な意味を持つ。しかし前者はそもそも、「対象が微小移動したときの各画素の輝度変化率」が分かるならば必要のない演算であり、現に我々の提案手法では隣接画素間の差分ではなく、CG モデルの微小移動により、運動にともなう輝度変化率を求めている。また後者の画素間の拘束条件は、そもそも勾配法がアパーチャ問題により不良設定問題であるから導入されたものであり、問題を正則化するための経験則的な細工にすぎない。また、もし対象物体の動きが画像全体にわたり一定であると分かっているならば、その知識を用いてより強力な拘束を与えることができる。

2.1 節で述べた我々の手法は、対象物体を剛体であると仮定し、それにより 6 個に限定された運動パラメータを全画素の情報から求める手法である。このような視点から見ると、提案手法は勾配法により現実的かつ強力な拘束を与えていることになり、したがって勾配法よりも安定なフロー推定を行うことができることを示唆する。

勾配法で求められるオプティカルフローは前後のフレーム間の運動に関するものであり、

長時間にわたる運動推定では誤差の蓄積によるドリフトが発生する．それに対し提案手法は生成された画像と入力画像の差を最小化するため，原理的にドリフトは生じないという利点も持つ．

5.2 直接法⁹⁾ との関係

Horn らは勾配法をより発展させた手法として，画像上の輝度の変化から各点のオプティカルフローを介さずに，直接的に対象物体の運動を求める直接法⁹⁾ を提案した．直接法ではカメラが回転運動のみで移動しない場合など，解を求めることができるいくつかの場合をあげているが，そのうちの 1 つに奥行きが既知である場合があげられている．これをビジュアルフィードバック分野で広く用いられている画像ヤコビアンに基づいて整理すると，直接法は以下のような数式で表すことができる．

$$\Delta I = \begin{bmatrix} \frac{dI}{du} & \frac{dI}{dv} \end{bmatrix} \begin{bmatrix} \frac{du}{dX} & \frac{du}{dY} & \frac{du}{dZ} \\ \frac{dv}{dX} & \frac{dv}{dY} & \frac{dv}{dZ} \end{bmatrix} \begin{bmatrix} \frac{dX}{dp_1} & \dots & \frac{dX}{dp_6} \\ \frac{dY}{dp_1} & \dots & \frac{dY}{dp_6} \\ \frac{dZ}{dp_1} & \dots & \frac{dZ}{dp_6} \end{bmatrix} \begin{bmatrix} \Delta p_1 \\ \vdots \\ \Delta p_6 \end{bmatrix} \quad (12)$$

ここで ΔI は近似されるべき画像の差を表し，また Δp_i は求める微小移動量である．右辺の行列は左から画像の勾配，カメラパラメータの微分係数，および姿勢の微小変動に関する特徴点の動きを表す行列である．特にこれらの行列のうち後 2 者をまとめたものを画像ヤコビアンといい，以下のようにまとめられる．

$$\begin{aligned} J &= \begin{bmatrix} \frac{du}{dX} & \frac{du}{dY} & \frac{du}{dZ} \\ \frac{dv}{dX} & \frac{dv}{dY} & \frac{dv}{dZ} \end{bmatrix} \begin{bmatrix} \frac{dX}{dp_1} & \dots & \frac{dX}{dp_6} \\ \frac{dY}{dp_1} & \dots & \frac{dY}{dp_6} \\ \frac{dZ}{dp_1} & \dots & \frac{dZ}{dp_6} \end{bmatrix} \\ &= \begin{bmatrix} \frac{du}{dp_1} & \dots & \frac{du}{dp_6} \\ \frac{dv}{dp_1} & \dots & \frac{dv}{dp_6} \end{bmatrix} \end{aligned} \quad (13)$$

Horn らの手法はもとの勾配法から出発しているため，画像面内の勾配 $\left[\frac{dI}{du} \frac{dI}{dv} \right]$ を隣接画素間の差分から求めている．それにあわせ，物体の微小姿勢変動に関する画像面上の特徴点の移動量を求め，この積により画像の変動が近似できると仮定することにより，微小移動量 Δp_i を求めている．しかしこの手法において，問題となるのは画像ヤコビアン J の計算である．画像ヤコビアンを構成する 2 つの行列のうち，前者はカメラパラメータから求めることができるが，実際にはこれは定数ではなく，物体表面上の各点の座標 (X, Y, Z) に依存した値となる．たとえば Z がカメラ光軸方向の奥行きである場合， $\frac{du}{dX}$ の絶対値は奥行き

値 Z の絶対値が大きいほど小さな値となる．また後者の行列の計算も容易ではない．求めるべきパラメータ p_i が平行移動に関するものであれば簡単であるが，回転に関する値であるとき，この行列の各成分もまた，座標 (X, Y, Z) に応じて変化する値となる．たとえばある点が回転中心の手前にあるときと奥にあるときとでは，回転にともなうその点の移動は逆向きとなる．また，画像面内の点の座標 (u, v) と世界座標 (X, Y, Z) との対応付けも簡単ではない．特に物体の形状が複雑でありオクルージョンを生じる場合には非常に複雑となるため，実装例としては対象物体の形状を平面と仮定したもの²²⁾ が多い．また Tamaki らは CG ハードウェアを用いて画像ヤコビアン²³⁾ の算出に必要な画像各点の奥行き²²⁾ の算出や，点の可視性の判定を行っているが²³⁾，依然として画像の勾配とヤコビアンを独立に算出している．それに対し我々の手法は CG ハードウェアにより多数の画像を描画し，それらの差分により直接的に，式 (4) を行列表記した以下の式

$$\Delta I = \begin{bmatrix} \frac{dI}{dp_1} & \dots & \frac{dI}{dp_6} \end{bmatrix} \begin{bmatrix} \Delta p_1 \\ \vdots \\ \Delta p_6 \end{bmatrix} \quad (14)$$

における姿勢の微小変動に関する輝度の変化量を求めている．これにより CPU では複雑な実数演算をする必要が大幅に減少し，高速化を達成することができる．現在は GPU により 13 枚の画像を生成し，それを CPU が読み出すことで後の演算を行っているが，近年は GPU の性能向上にともない，GPU を汎用的に利用するための環境が整ってきていることから，画像どうしの差分をとることはもちろん，さらに式 (9) における共分散行列や逆行列の計算も GPU で可能になりつつあると思われる．さらに提案手法では，従来の直接法では取り扱うことが困難であった鏡面反射物体の運動を求めることもできる．

5.3 テンプレートマッチングとの関係

テンプレートマッチングは様々な分野で実用化されている代表的な画像処理手法であり，運動の推定にも用いられる．テンプレートマッチングにおいてよく問題とされるパラメータの 1 つにテンプレートのサイズがある．一般に，より大きなテンプレートを用いると精度が向上するとされており，また局所解に陥る危険性も減少する．しかし対象物体の位置や姿勢が変化すると，テンプレートに対し対象物体の見かけの大きさが変わるために照合に失敗することがある．そこで対象物体の姿勢推定を行う場合，非常に小さなテンプレートにより特徴点を追跡し，その軌跡の統合により対象物体全体の位置と姿勢を推定することが多い．このとき，テンプレートによる特徴点の追跡に失敗するとその後の運動復元に悪影響がある

ため、運動復元において誤対応に対するロバスト性を有する手法を用いたり、姿勢パラメータからテンプレート位置を予測したりすることで誤対応を防ぐ手法などが提案されている。しかしこのような手法も先に述べた勾配法による運動推定と同様に、2段階に処理が分かれていることで問題を不必要に複雑化させているきらいがある。それに対し、我々の提案手法では対象物体全体の形状をモデリングし、それにより運動復元を行っているため、姿勢の変動によるテンプレートの見かけの大きさの変動の問題は自ずと解決されている。対象物体の姿勢の変化によりテンプレートの見かけの傾きが大きくなったときは、CGのレンダリングにより自動的にテンプレートが圧縮されることと同等の効果を持つ。

テンプレートを用いた運動追跡に比べ、対象物体のモデリングに要する手間やコストが高いことを問題点として指摘する読者もあろう。しかし3.2節で述べたように、ステレオカメラを用いることでオンライン処理として対象物体のモデリングを行うことが可能であることも示されている。ステレオカメラによるシーンのモデリングは、ちょうど、テンプレートの切り出しとそのモデル内の位置を求めることと等価である。図7に示したように、対象物体のモデルは密に得られている必要もなく、テクスチャなど特徴に乏しいところは穴であってもよく、また、通常の特徴点追跡では使いにくい直線や曲線状のエッジやテクスチャも運動追跡に寄与するテクスチャとして利用することもできる。また物体間や物体内のオクルージョンによりテンプレートが不可視となるという問題や、4章で示したような複雑な反射モデルに基づく物体への対応などの点でも提案手法に有効性が認められると考えている。

対象物体の回転にともない、照明条件によっては生成画像と入力画像との間で光学的整合性が満たされなくなるという問題点も指摘されてきている。しかしテンプレートマッチングに依存したほとんどの手法では、このような照明条件の変化によるテンプレートの輝度の変化は取り扱っておらず、一定輝度のテンプレートをそのまま用いている。テンプレートマッチングでは正規化相関など大局的な輝度変化に対して頑健な類似度尺度も存在する²⁴⁾が、それに対応するような画像処理については今後の課題としたい。

テンプレートマッチングでは多くの場合、探索空間の全探索により最も適合度の高い位置を探す。それに対し提案手法は現在位置の近傍の適合度から線形補間により位置を求めていることに相当する。しかし提案手法によって6自由度の運動を求めるとき、仮に全探索をするならば探索空間の次元が大きすぎるために実用的な速度での実行が難しいであろうと考えられる。

5.4 微分カメラ¹⁰⁾との関係

Yangらはカメラが微小移動したときの画像を生成する代わりに複数のカメラを用い、そ

のうち2台のカメラから得た画像の差を、そのカメラの位置・姿勢の変動にともなう画像の変化と見なすことでカメラの自己位置を復元する手法を提案した¹⁰⁾。実際にはカメラを4台用い、基準位置を表すカメラに対して並進3成分のために残りの3台を用いる。カメラの回転については、画像のホモグラフィ変形により等価な画像を得ることができる。Yangらは、この手法では鏡面物体や透過物体など、シーンの反射特性に依存せずにカメラの自己位置が追跡できると述べており、実際にそのような実験例も示しているという点で特に興味深い。しかしカメラの位置・姿勢同定の原理そのものは本論文で述べた我々の手法とまったく同じであり、特に位置・姿勢の微小変動に対する画像の変動の線形性をManifoldの性質から説明している点が興味深い。彼らの手法ではシーンをモデリングしないため、我々の手法との間で性能を比較することは必ずしも適切ではないが、フレーム間の移動量の積算により移動を求めているため誤差の累積の問題があることはYangら自らが指摘している。

勾配に基づく運動推定では、勾配は局所的に線形であると仮定しているが、この仮定が崩れると推定された位置・姿勢がずれる。そのためYangらは画像をあらかじめぼかして用いているが、これはカメラ運動が比較的速い場合も含め、追跡の安定化に寄与する。それに対し我々は画像のぼかしを行っていないが、これは生成画像と入力画像の差分を繰返し演算により最小化しており、ぼかさないう画像のほうが最終的な誤差は減少すると考えているからである。もちろん物体の運動が速い場合にはぼかしたほうが追従性能は向上すると思われる。

5.5 距離画像の位置合わせ問題として

同一の物体を異なる視点から計測した一対の距離画像から、計測視点の相対位置と姿勢を求める問題^{*1}を距離画像の位置合わせ(registration)という。距離画像の位置合わせにおいて最も広く用いられている手法はICP(Iterative Closest Point)法¹¹⁾である。しかしICPには、最近傍点の探索という非常に計算量の大きい処理が含まれる。この計算をより高速に行うため、距離画像を階層表現する手法など多くの手法が提案されているが、依然としてその後の距離最小化演算よりも計算量が大きく、リアルタイム処理ができる状況とはいえない。最近傍点を求める代わりに、一方のセンサから見た視線方向に点どうしを対応付ける方法も考えられているが、この方法では距離最小化において光軸に対して垂直方向の移動や光軸周りの回転が求められないという問題がある。またICP法では図10(c)にあげたような、ほぼ回転対称であり奥行き変動が浅い物体では収束が遅く、多数の繰返し演算が必要

*1 固定されたレンジファインダに対し、物体の移動前後の距離画像が得られたときの移動量の推定問題と同じである。

である。なぜなら物体表面の勾配が小さいため、探索の結果得た最近傍点が、実際の対応点よりも近くにとられてしまうからである。それに対し我々の手法では、物体の形状モデルを微小移動することによりこのような運動が生じたときの奥行きの変化量を求め、それにより運動推定するため高速であるという利点がある。

本論文で述べた手法はもともと、リアルタイム距離画像センサ「シリコンレンジファインダ」から得た距離画像により実時間的に対象物体の位置・姿勢を追跡するために考案されたものである。この手法では ICP で必要であった最近傍点決定のための探索が必要ないだけでなく運動推定のための演算も単純であり、レンダリングハードウェアが利用できることも相まって 10 年前の計算機でも実時間処理が可能であったことは特筆できるであろう。ただし姿勢が大きく変動したときの収束性や、純粋に 3 次元形状どうしの演算である ICP に比べ処理が本質的に視点依存性を持つという問題は未検討である。

5.6 最適化問題として

我々が提案する手法は、入力画像と生成画像の差を単に最小化することに基づく。そのため、同様の原理は単に対象物体やカメラの位置・姿勢の追跡だけでなく、他の変動要因の追跡にも利用することができると考えられる。たとえば対象物体を照らす照明の位置・強度の変動や、物体のパラメトリックな変形の追跡にも応用可能であろう。また鏡面反射物体の運動推定と同様に、鏡面物体そのものと、それに映り込んだ物体との双方の運動復元も可能ではないかと考えている。ただしパラメータの変動に関する画像の変化量が小さい場合や、異なる要因に関する画像の変化が似通ったものである場合、退化により逆行列が安定に求められないということは考えられる。そのような場合には、3.2 節で述べたような複数視点からの観察が有効である場合もある。

同様に、画像生成の高度化も興味深いテーマである。たとえば対象物体の反射モデルとして拡散反射や鏡面反射だけでなく、2 色性反射や異方性反射、さらには透明物体の屈折現象のシミュレーションや影のレンダリングにより、複雑な光源環境での運動の追跡も可能であると思われる。またプロジェクタなどを用いたアクティブライティングのシミュレーションができるならば、能動計測との融合も考えられよう。

6. ま と め

本論文では計算機内のモデルから生成した画像と、センサから新たに得られた画像との差を最小化するような運動パラメータを求めることにより、画像からの特徴抽出を行うことなく物体の位置・姿勢を実時間的に追跡する手法について述べた。我々は 10 年以上にわたっ

て様々な条件における運動追跡問題へこの手法を実装し、高速かつ高精度に運動の推定が行えることを示した。特に、1) 多視点映像を用いることにより自由度の高い運動の推定を安定化すること、2) ステレオカメラや距離画像センサを用いることで対象物体のモデリングと運動の追跡を同時に行うこと、3) 鏡面反射物体の追跡により物体表面に固定されたテクスチャに頼らなくても運動の復元ができること、などを示した。また距離画像センサから得た距離画像と濃淡画像を融合することで運動を推定する例も示した。

これらの手法や実装例について、これまで発表のたびに様々な有用な意見や議論をいただき、そのつど回答するとともに他の研究との関連や理論的背景についても思索してきた。その結果、あまり直接的には関係しないと思われていた勾配法・直接法とテンプレートマッチング、ICP などの手法と本手法が関連することが分かり、また近年注目されている微分カメラや GPU ベース画像処理などとも共有できる知見が見いだされた。本手法は、逆問題を解くことを徹底的に嫌うことから出発しており、そのあまりの単純さが、ややもすると理論面での底の浅さと見られることもあろう。しかし近年興隆を見せている事例ベース画像処理においても、画像の生成過程には踏み込まずに様々な成果を収めている。そのようなアピアランスベース画像処理の一手法として本論文でまとめた手法が位置づけられ応用されていくことを期待したい。

参 考 文 献

- 1) 加藤文和, 深尾隆則, 羽下哲司: 対象追跡—フレーム間の類似度に着目した手法から動きのモデルに着目した手法まで, 情報処理学会研究報告コンピュータビジョンとイメージメディア, Vol.2005, No.88, pp.185-198 (2005).
- 2) Murase, H. and Nayar, S.K.: Visual Learning and Recognition of 3-D Objects from Appearance, *International Journal of Computer Vision*, Vol.14, pp.4-24 (1995).
- 3) 森谷貴行, 日浦慎作, 井口征士: CG 画像の生成と比較に基づく動物体の運動計測, 情報処理学会研究報告 2002-CVIM-133, pp.87-94 (2002).
- 4) 森谷貴行, 日浦慎作, 佐藤宏介: 多視点画像と CG 画像の比較に基づく動物体の運動計測, 信学論 D-II, Vol.J88-D-II, No.5, pp.876-885 (2005).
- 5) 森谷貴行, 日浦慎作, 佐藤宏介: ステレオカメラを利用した自己位置・姿勢の実時間推定とシーンの同時モデリング, 電子情報通信学会論文誌, Vol.J90-D, No.6, pp.1579-1591 (2007).
- 6) 日浦慎作, 山口 証, 佐藤宏介, 井口征士: 動距離画像の計測と生成による任意形状物体の実時間追跡, 信学論 D-II, Vol.J80-D-II, No.6, pp.1539-1546 (1997).
- 7) 日浦慎作, 山口 証, 佐藤宏介, 井口征士: 距離画像と濃淡画像の融合による任意形状物体の実時間追跡, 信学論 D-II, Vol.J80-D-II, No.11, pp.2904-2911 (1997).

- 8) Horn, B.K.P. and Schunck, B.G.: Determining Optical Flow, *Artificial Intelligence*, Vol.17, No.1-3, pp.185-203 (1981).
- 9) Horn, B.K.P. and Weldon, E.J.: Direct Methods for Recovering Motion, *International Journal of Computer Vision*, Vol.2, No.1, pp.51-76 (1988).
- 10) Yang, H., Pollefeys, M., Welch, G., Frahm, J.M. and Ilie, A.: Differential Camera Tracking through Linearizing the Local Appearance Manifold, *Proc. CVPR07*, pp.1-8 (2007).
- 11) Besl, P.J. and McKay, N.D.: A method for registration of 3-D shapes, *IEEE Trans. Pattern Analysis and Machine Intell.*, Vol.14, No.2, pp.239-256 (1992).
- 12) 横山 敦, 佐藤宏介, 芦ヶ原隆之, 井口征士: 無調整型フォト VLSI センサを用いた実時間距離画像計測—シリコンレンジファインダ, *信学論 D-II*, Vol.J79-D-II, No.9, pp.1942-1500 (1996).
- 13) Hiura, S., Sato, K. and Inokuchi, S.: Range Image Processing using Graphics Accelerator, *Proc. Japan-USA Symposium on Flexible Automation*, pp.955-958 (1996).
- 14) 馬場 充, 成田大助, 坂本 明: 光線追跡法による鏡面物体の 3 次元形状計測, 第 11 回パターン計測シンポジウム予稿集 (2006).
- 15) 馬場 充, 小西忠孝, 半田久志: スリット光投影法による柱状鏡面物体の形状測定法, *電子情報通信学会論文誌*, Vol.J83-D-II, No.8, pp.1773-1782 (2002).
- 16) 汪 増福, 加藤博一, 佐藤宏介, 井口征士: 鏡面物体の 3 次元形状計測, *信学論 D-II*, Vol.J75-D-II, No.7, pp.1177-1186 (1992).
- 17) 汪 増福, 加藤博一, 佐藤宏介, 井口征士: 単眼観測画像による鏡面形状の復元・推定, *信学論 D-II*, Vol.J76-D-II, No.3, pp.575-585 (1993).
- 18) Oren, M. and Nayar, S.K.: A Theory of Specular Surface Geometry, *Int. Journal of Computer Vision*, Vol.24, No.2, pp.105-124 (1997).
- 19) 大場光太郎, 池内克史: 局所固有空間手法による金属物体の安定認識, *信学論 D-II*, Vol.J80-D-II, No.12, pp.3147-3154 (1993).
- 20) Shreiner, D., Woo, M., Neider, J. and Davis, T.: *OpenGL Programming Guide 4th Edition: The Official Guide to Learning OpenGL, Version 1.4*, Addison Wesley (2004).
- 21) 山澤一誠, 八木康史, 谷内田正彦: 移動ロボットのナビゲーションのための全方位視覚系 HyperOmni Vision の提案, *電子情報通信学会論文誌 D-II*, Vol.J79-D-II, No.5, pp.698-707 (1996).
- 22) 杉本茂樹, 奥富正敏: ステレオ画像を用いた高速な平面パラメータ推定法, *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*, Vol.48, No.SIG1 (CVIM17), pp.24-34 (2007).
- 23) Tamaki, T. and Yamamoto, M.: Calibration Method by Image Registration with Synthetic Image of 3D Model, *IEICE Trans. Inf. Syst.*, Vol.E86-D, No.5, pp.981-985 (2003).
- 24) Viola, P. and Wells III, W.M.: Alignment by Maximization of Mutual Information, *International Journal of Computer Vision*, Vol.24, No.2, pp.1573-1405 (1997).
- 25) 井口征士, 佐藤宏介: 三次元画像計測, 昭晃堂 (1990).
- 26) Baker, S. and Matthews, I.: Lucas-Kanade 20 years on: A unifying framework, *International Journal of Computer Vision*, Vol.56, No.3, pp.221-255 (2004).
- 27) Armstrong, M. and Zisserman, A.: Robust object tracking, *Proc. ACCV95*, Vol.1, pp.58-62 (1995).
- 28) Drummond, T. and Cipolla, R.: Real-time visual tracking of complex structures, *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol.24, No.7, pp.932-946 (2002).
- 29) 角 保志, 石山 豊, 富田文明: Hyper Frame Vision: 移動物体の 6 自由度位置決めシステム, *信学論 D-II*, Vol.J86-D-II, No.9, pp.1320-1328 (2003).

(平成 20 年 5 月 10 日受付)

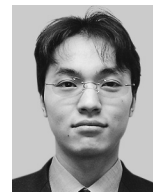
(平成 20 年 7 月 30 日採録)

(担当編集委員 加藤 丈和)



日浦 慎作 (正会員)

1993 年大阪大学基礎工学部制御工学科飛級中退。1995 年同大学大学院基礎工学研究科修士課程修了, 1997 年同博士課程短期修了。同年京都大学工学部リサーチアソシエイト。1999 年大阪大学大学院基礎工学研究科助手, 2003 年同助教授, 現, 准教授。2008 年マサチューセッツ工科大学メディアラボ客員准教授。3 次元動画像計測・処理とその VR・コミュニケーション応用の研究に従事。1993 年電気関係学会関西支部連合大会奨励賞, 2000 年画像センシングシンポジウム優秀論文賞受賞。電子情報通信学会, VR 学会各会員。博士 (工学)。



森谷 貴行

2002 年大阪大学基礎工学部システム科学科卒業。2004 年同大学大学院基礎工学研究科博士前期課程修了。2007 年同博士後期課程修了, 同年マツダ株式会社に入社, 現在に至る。博士 (工学)。



佐藤 宏介（正会員）

1983年大阪大学基礎工学部制御工学科卒業．1985年同大学大学院基礎工学研究科修士課程修了．1986年大阪大学基礎工学部助手．1988年カーネギーメロン大学ロボット工学研究所客員研究員．1992年大阪大学基礎工学部講師．1994年奈良先端科学技術大学院大学情報科学研究科助教授．1999年大阪大学大学院基礎工学研究科助教授．2003年同教授．3次元画像計測，仮想現実感，デジタルアーカイブ等の映像情報メディアの研究に従事．1987年篠原記念学術奨励賞受賞．電子情報通信学会，色彩学会，映像情報メディア学会，VR学会，IEEE等各会員．工学博士．
