

モンテカルロ法とベイズ推定に基づく動的カメラからの三次元形状復元を利用したカメラの移動方向評価尺度に関する検討

成澤 祐*1 大谷 淳*1 小方 博之*2

*1 早稲田大学大学院国際情報通信研究科

*2 成蹊大学理工学部

概要- 移動ロボットに搭載されたカメラにより獲得される動画像から物体全体の形状を三次元復元するためには、任意の一視点方向からの復元だけでは不十分であり、移動経路において適宜最適な移動方向を選択しつつ、局所ごとの形状復元を行い、最終的に複数の局所的な復元結果を統合する必要がある。本報告では、移動カメラ画像からの三次元復元法として、モンテカルロ法とベイズ推定を用いる手法を用いて対象物体の局所的な三次元形状を復元する方法を検討する。更に、三次元形状復元を行った視点（動画像の初期フレームに対応）からの対象物体の見え方に関する情報量の取得を行い、ロボットの移動方向の選択をするための評価尺度としての有用性を検討する。

Study of Evaluating the Camera's Moving Direction in 3D Object Shape Reconstruction from a Dynamic Camera's Image Based on Monte Carlo Method and Bayes Estimation

Yu NARISAWA*1 Jun OHYA*1 Hiroyuki OGATA*2

*1 Waseda University, Graduate School of Global Information and
Telocommunication Studies

*2 Seikei University, Faculty of Science and Technology

Abstract- In order to reconstruct the entire 3-D structure of an object from the video sequence acquired by the camera attached to a mobile robot, one view direction is not sufficient; it is necessary to select the optimal direction at each spot on the robot's path, reconstruct local structures, and integrate the multiple local structures. This paper studies a Monte Carlo method and Bayes Estimation based method for the local 3-D reconstruction. In addition, we study the volume of information from a view point (the initial frame of the view sequence) is useful for evaluating the robot's moving direction

1.はじめに

近年、ロボットの研究開発が活発になっており、その適用範囲も、流れ作業的な工場などロボットが作業をすることが前提となっている環境から家庭、オフィスなど人間が生活する一般的な環境へと広がる兆候が見られるようになってきた。このような一般的な環境中を自律的に移動して知的な作業を行うロボットを実現するためには、ロボットによる動的な物体の認識が必要となってくる。例えば、家庭内においてロボットに指示した物体を持ってくるように命令したとする。工場など予め整備された環境に比べると家庭内は指示した物体以外の対象物が数多く存在することだろう。

そこで、未知なる物体に対するロボットの対象物の認識技術が必要になる。

そのためには、ロボットが移動している間に、ロボットに搭載したカメラにより獲得される時系列画像を自動的に解析し、ロボットが移動する環境の3次元構造を復元する技術の実現が必須と言っても過言でない。このような技術は **Structure from Motion (SfM)** と呼ばれ、数多く研究されてきており、ロボットだけでなく、デジタルアーカイブ、仮想空間インタフェースなど様々な分野への応用が期待されている。

さて、未知なる物体に対し、ロボットまたはコンピュータが物体を認識する為に対象物体の三次元形状復元を行う際に問題となってくる点は、物

体に対し、限定された数の視点方向からの復元は曖昧さを含んでいる、ということである。

ある方向からはただの面にしか見えなかった物体が、異なる方向から観測すると別の物体が隠れていたり、凹凸がある複雑な形状を含んでいるかもしれない。そこで、移動ロボットにより物体全体の形状を三次元復元するためには、移動経路において適宜最適な移動方向を選択しつつ、局所ごとの形状復元を行い、最終的に複数の局所的な復元結果を統合する必要がある。

このために必要となる技術は、移動カメラからの三次元復元法とロボットの移動方向の選択となる。本論文では移動ロボットに搭載されたカメラにより獲得される動画像から物体の三次元形状を復元する方法と、物体全体の形状復元を行う際に必要となるロボットの移動方向を選択するための評価尺度を検討する。

2. 関連研究

以下に従来手法による三次元形状復元技術と視点の評価尺度に関する研究例を示す。

2.1. SfM に関する従来技術

物体の三次元形状復元技術として SfM と呼ばれる手法があげられることは 1.1 節で述べたが、SfM の代表的手法として因子分解法[2]があげられる。この手法はカメラを移動させることで得られる連続画像から対象物の形状復元とカメラの運動をロバストに推定する手法である。しかし、因子分解法の主な問題点の一つとして中心射影の線形近似の問題点がある。因子分解法では計算の際に特異値分解を使用することから計算が線形である必要があり、カメラの射影モデルとして中心射影の線形近似モデルを用いる。中心射影の線形近似の問題とは、この近似をすることで復元結果に悪影響が出てしまう問題である。

この問題を解決するべく研究された手法がベイズ更新に基づく SfM の手法である[1]。ここで用いられるベイズ更新とは、ロボットの位置情報を走行誤差が生じることを考慮し、移動命令とセンサ値から確率分布を考え、物体の距離情報の事後分布を更新していくものである。

復元手順としては、画像上の特徴点をフレーム毎にトラッキングし、モンテカルロ法によりリサンプリングを行っていくものである。この手法の問題点として、特徴点のトラッキングが失敗した場合、形状復元結果に深刻な影響を与えたり、形状復元そのものが困難になることがあげられる。

2.2. 視点の評価尺度に関する従来技術

コンピュータビジョンの分野では、物体像の持つ情報量を定義し、評価する研究が行われている。

例えば、形状認識のために物体に対して視点毎に物体のどの視点から物体を観測すべきか、とい

う情報量を定義し、定量的に評価する手法について研究例が見られる[4][5][6]。

佐藤らは、多面体像を対象に、視点によって得られる情報量を定義し評価している[4]。

石川らは、佐藤らが定義した情報量[4]に基づき、情報量と視点選択における心理尺度の関連性について心理実験を行ない、物体の見え方と知覚のしやすさを測定している[5]。大場らは、三角パッチ表現をされた CG(ComputerGraphics)を用い、パッチの平面と稜線より形状情報量を定義し、物体の形状特徴をよりよく把握できる視点の候補を決定する手法について述べている[6]。

また、情報量を計算するために、必要となる物体の形状復元に関しては、佐藤らは、実際の対象物に対し、レンジファインダを用い、カメラと物体との距離を計測することで形状の把握を行い、定義した勾配空間より全体に対する面の比率から情報量を計算している[4]。

石川、大場らは共に、予め設定した対象物の三角パッチを使用した CG を用いて、三角パッチからの情報量計算の有効性を評価している[5][6]。

本論文の目的の一つである物体全体の三次元形状復元のための観点からの「視点の評価尺度」に関する研究はほとんど見られない。

3. 提案手法

この章では本論文における提案手法を説明する。

3.1. 概要

Fig.1 に示すように、ロボットは初期位置から移動を行い、複数の視点から対象物に対して移動経路上の複数の視点から形状復元を行い、最終的にその結果を統合することで対象物体に対しての三次元復元を行うものとする。

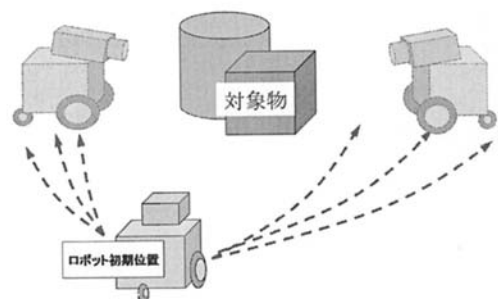


Fig.1. 提案手法イメージ図

実際の環境として、まず Fig.2 のように対象物体に対し、移動ロボットを想定したデジタルカメラを用い、並進運動をさせながら複数の画像を取得することを前提とする。この際に、 $t=0$ の位置で取

得した画像を"初期フレーム"と定義し、この初期フレームにおいて撮影した対象物の形状に対して復元を行っていくものである。

また、 $t=1$ の位置で取得した画像を"第一フレーム"。 $t=2$ の位置で取得した画像を"第二フレーム"と定義する。

この初期フレームを、特徴点を抽出するフレームとし、情報量を取得するロボットの視点位置として定義する。

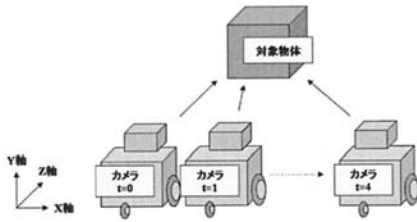


Fig2. 実験環境図

また、形状復元を行う際のカメラの投影モデルとして、Fig.3 のような中心射影により、画像平面に投影される三次元空間中の点を関係づける。

Fig.3 はカメラ原点から空間上の特徴点を見た図であり、Image Plane は取得する画像を表す。Fig.3 のように、3次元空間中の点(X,Y,Z)が、中心射影によりカメラ平面に点(u,v)として投影されると、(X,Y,Z)と(u,v)との関係は次式で表される。

$$u = f \frac{X}{Z} \quad v = f \frac{Y}{Z} \quad (3.1)$$

f はカメラの焦点距離である。

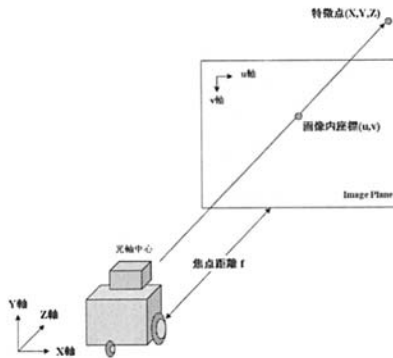


Fig3. 中心射影

次に、Fig.4 を用いて提案手法の説明をする。

まず、ロボットは対象物体に対して最初の形状復元を行う。次に、形状復元の結果から情報量を計算し、次に移動する方向の選択を行う。

この際に対象物体に対して十分な復元ができたかを判断し、十分な復元が行えていないと判断されれば、次の視点に移動して形状復元を行い、十分な復元が行えているようであれば、局所的に取得した三次元形状復元の結果を統合し、最終的な物体の三次元復元結果を出力する。

なお、本論文の検討対象は Fig.4 において*のついた処理であり、3.2 節以降で詳しい説明を行う。

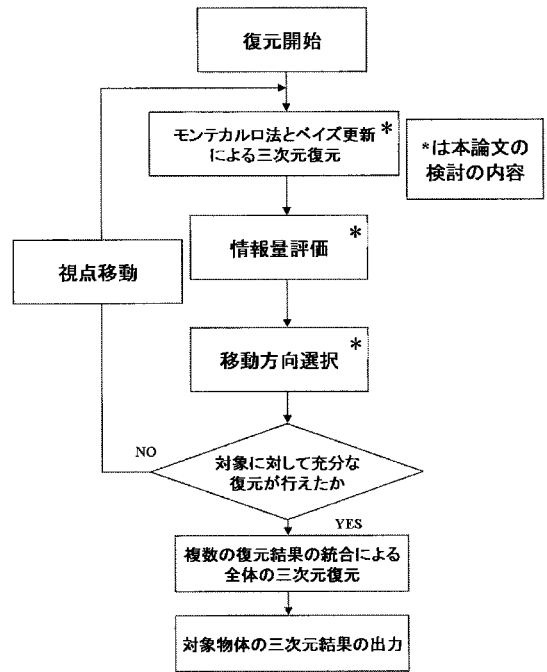


Fig4. 提案手法

3.2. モンテカルロ法とベイズ更新に基づく三次元形状復元法

Fig.5 に本手法における三次元形状復元のフローチャートを示す。この部分は、Fig.4 における"モンテカルロ法とベイズ更新による三次元復元"に対応している。

また、本提案手法において、従来の手法[1]と最も異なる部分は、特に 2)~4)の部分であり、サンプルをロボットの移動確率密度に従い移動させる部分、画像内で特徴点の確率密度表現をする部分である。

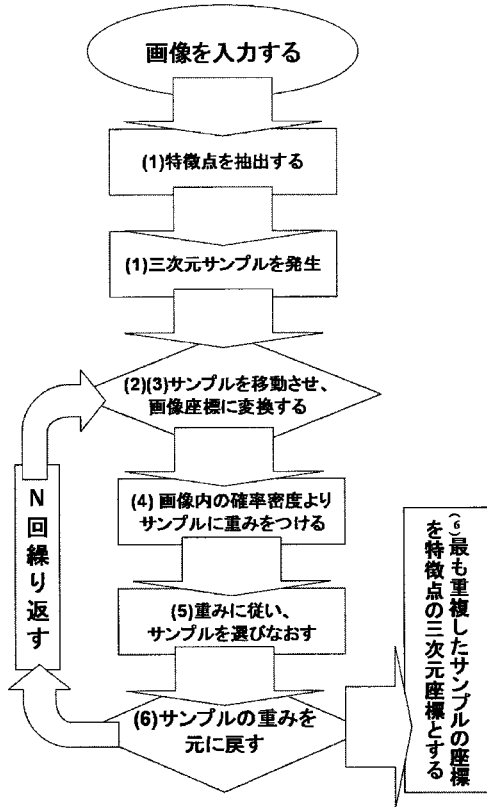


Fig5.三次元形状復元法のフローチャート

1).まず、時系列画像における初期フレームにおいて特徴点を検出する。この特徴点は初め、画像座標系において 3.1 節の中心射影に従って取得し、この座標を (u, v) とする。画像座標系から三次元座標系への変換(キャリブレーション)は 3.1 節の式(3.1)を使う。次に特徴点に対するサンプルを N 個発生させる。具体的には、まずサンプルの Z 座標にあたる一様乱数を発生させる。

サンプルはカメラ光軸上に発生させるので、 X 座標、 Y 座標は式(3.1)から求められる。即ち、

$$X = \frac{uZ}{f} \quad Y = \frac{vZ}{f} \quad (3.2)$$

である。ここで発生させたサンプルは実際の特徴点の三次元座標の解に当たるものを含んでいる。また、座標系については、Fig.3 に示した三次元の世界座標、同じく三次元のカメラ中心を原点と置く光軸方向へのカメラ座標系、画像内の二次元である画像座標系を用いる。

簡単の為、カメラ座標系と世界座標系を本手法では一致させる。具体的には、カメラの光軸中心を世界座標系の原点とする。

2).1)で発生させたサンプル群をロボットの移動に従い移動させる。初期フレーム以降、画像内において特徴点は移動しているように見えるためである。サンプルの三次元座標をロボットの移動に基づき移動させることにより、特徴点の追跡を行わずに表現する。また、ロボットの移動誤差は正規分布に基づくとし、サンプルのカメラ座標系に正規乱数を加えることでロボットの移動を表現する。詳しくは 3.3 節で述べるものとする。

3).2)で移動させたサンプルを次フレームにおける画像内のどの座標に相当するか、三次元座標から画像座標に変換する。

この変換も 3.1 節の式(3.1)に従って行う。

4).変換した画像座標において、特徴点らしさの確率をそのサンプルの重みに掛け合わせる。画像内における特徴点らしさの確率に関しては、3.4 節で説明する。この処理を行うことで、リサンプリングの際により確率の高いサンプルが選ばれることになる。

5).重みに従い、リサンプリングを行う。リサンプリングにおいて同じサンプルを選ぶことにより、そのサンプルが特徴点の三次元座標の解である確率が高まっていく。

6).2)~5)までの動作を画像の各フレーム毎に行い、重みを初期値に戻しながらサンプルを更新していく。最終的にリサンプリングで最も確率の高くなったサンプルを解として出力する。

3.3.ロボットの移動モデルに関する確率分布表現

3.2 節のフローチャート 2)で記述した「ロボットの位置の確率密度表現」について論じる。本手法においてサンプルの三次元座標系に誤差を含めたロボットの移動を足し合わせることで確率密度表現を行う。あるサンプルの三次元座標を $\text{Sample}(X, Y, Z)$ とおくと、そのサンプルは次フレームにおいて、 $\text{Sample}(X+e, Y+e, Z)$ がある。

これは、問題設定としてロボットが上下の誤差を含むとしたためである。なお、 e はロボットの移動を表現した正規乱数であり、ロボットの移動誤差が平均 0、標準偏差 1 の正規分布に従うとして扱った。サンプルが次フレームにおいて $\text{Sample}(X+e, Y+e, Z)$ にあるとした場合、中心射影の式(3.1)より、サンプルの画像座標は (u', v') と変換できる。さて、サンプルの重みは画像内座標 (u', v') の特徴点らしさの確率密度となる。この部分については 3.4 節で詳しく説明する。

3.4. 各画像における特徴点の確率密度表現

3.2節のフローチャート4)に対応する画像中の特徴点らしさの確率密度表現に関して説明をする。特徴点らしさの確率密度表現として、特徴点を中心とするテンプレートマッチングを行う。

本論文においては簡単のため、ロボットの移動は水平方向のみに限定するが、誤差を全く含まずに移動することは困難である。

そこで、ロボットは上下左右の誤差を予め含むものとして、マッチングをエビポーラ線を中心とする上下数十画素について行ない、確率密度表現を行う。Fig.6において、左図中の赤い丸を初期フレームにおける特徴点とする。次のフレームでは、この特徴点が右図のように上方に移動してしまっていると想定する。

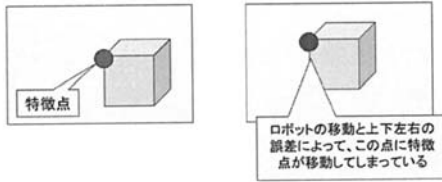


Fig6.特徴点の移動のイメージ

そこで、7×7のテンプレートを想定し、赤い丸の存在する画像内座標が(u,v)であるとするれば、テンプレートは(u,v)を中心とする7×7のピクセルの色情報を保存する。

次に、次フレームにおいてエビポーラ線を中心とする数十画素について走査する。(実験では、320×240画素を1フレームとしたため、上下30画素ずつ、計60画素について走査した。)テンプレートマッチングによる走査の結果、Fig.7 徴点らしさの確率密度表現ができる。

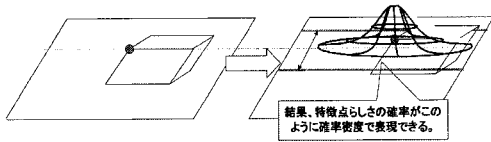


Fig7 特徴点らしさの確率密度表現

n番目のサンプルの重みを w_n として、ある時刻におけるロボットの移動確率を e_n とする。初期フレームの特徴点(u,v)に対する次フレームで移動したサンプルの画像内座標(u',v')とのテンプレートマッチングによる確率を $p(u',v')$ とすると

$$w_n = e_n \times p(u',v') \quad (3.3)$$

という式によってサンプルの重みを表現できる。

3.5. 情報量取得法

3.4節まで、対象物体に対する三次元座標の取得法を論じてきた。ここまでで得られた三次元座標は点群のデータとなっている。情報量の取得を行うためには、この点群データを連結し、面を生成する必要がある。そこで、三角形分割の手法の中で最も一般的に用いられるドロネー分割を用いて三角パッチの生成を行う。三角パッチからの情報量取得法として、大場らの物体の形状特徴をよりよく把握できる視点の候補を決定する手法 [6]を参考にした。

各三角パッチを Fig.8 のように定義し、単位法線ベクトル $n(x_n, y_n, z_n)$ とする。単位法線ベクトル n よりX方向の傾き u' 、Y方向の傾き v' を式(3.4)のように計算でき、この傾き(u',v')をラジアンで表現することで勾配空間(u,v)を式(3.5)のように定義する。

$$u' = \frac{x_n}{z_n} \quad v' = \frac{y_n}{z_n} \quad (3.4)$$

$$u = \arctan u' \quad v = \arctan v' \quad (3.5)$$

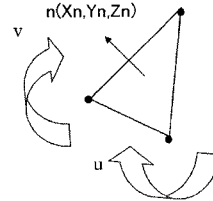


Fig8.三角パッチの法線ベクトル

三角パッチの勾配(u,v)に対し、視野角に対して $(-\pi/2 < (u,v) < \pi/2)$ の範囲にあるかを判定し、パッチがこの範囲内にあるとすれば、次に面積Sの計算を行う。本手法では、形状復元から三角パッチを作成しており、パッチ毎の面積は異なっている。

また、世界座標系の原点をカメラの光軸中心と一致させており、カメラの視点方向ベクトルは世界座標系(X,Y,Z)=(0,0,1)となる。よって、パッチに対してX-Y平面への正射影を行い、ヘロンの式を用いて面積の計算を行なう。面積Sの総和を S' とおくと、全体の面積に対するn番目のパッチの比率は

$$r_n = S_n / S' \quad (3.6)$$

となる。

文献[4][5][6]より情報量 $H(N)$ は

$$H(N | r_1, r_2, \dots, r_N) = - \sum_{i=1}^N r_i \log r_i \quad (3.7)$$

と計算できる。次に、以上の三次元復元と情報量をロボットの移動の評価尺度として有用か検討を行った。

4.実験

移動ロボットの移動経路上の一点(例えば、既に獲得され、三次元復元に用いられた動画像の最終フレームに対応)から次にどの方向に移動するべきかを決定する状況を想定し、24方向からの式(1)の値を比較する。即ち、形状復元対象となる物体を中心に Fig.9 のような空間を設定し、角度変化は Fig.10 のように水平面に対してつける。また、カメラの移動に対する各視点は Fig.11 のようになっている。

各視点毎に式(3.7)の情報量を計算し、対象物体に対する最適視点を決定する。本論文における"最適視点"とは、最も効果的な形状復元を行なう為の視点と定義する。その結果、Fig.12 のように、水平方向との角度が 60° であるような斜め上方からの、特徴点が多数存在するような視点が得られており、評価尺度として有効であることが確認できた。また、その視点からの形状復元を Fig.13 のように得ることができた。

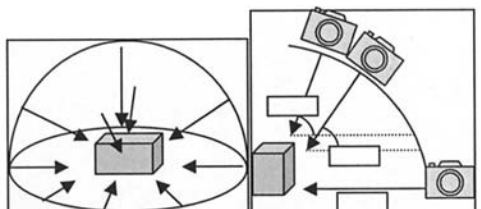


Fig.9 実験空間

Fig10 空間内の角度変化



Fig.11 実験空間上図

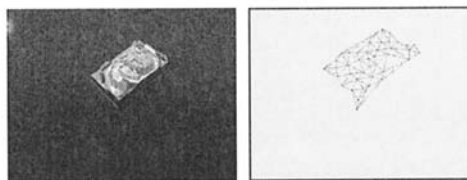


Fig.12 最適視点 Fig.13 最適視点からの形状復元

5.考察

本論文では、特徴点追跡が不要な三次元形状復元と、物体の三次元復元を行う為に必要な、ロボットの移動方向を選択する為の評価尺度の検討を行なった。

三次元復元の結果として、対象物体に対して 0° 、 45° 、 60° と角度をつけた周囲 24 方向からの形状復元結果が得られた。

また、得られた三次元復元の結果から、どの視点から物体を観測すべきかの情報量を計算し、高い情報量を持つ視点は特徴点を多く持つような視点という結果が得られた。よって、本論文によって提案された手法はロボットの移動方向選択の際の評価尺度として有用であると言える。

参考文献

- [1]大井 康義、池田 光司、小方 博之、"移動ロボットによる逐次の3次元形状復元", 日本機械学会ロボティクス・メカトロニクス講演会'05, CD-ROM 収録, 2005. 6
- [2]C. Tomasi and T. Kanade: "Shape and Motion from Image Streams Under Orthography: A Factorization Method", International Journal of Computer Vision, Vol. 9, No. 2, pp. 137-154 (Nov. 1992)
- [3]成澤 祐、小方 博之、大谷 淳、"カメラの上下動を考慮するモンテカルロ法とベイズ推定に基づく物体の三次元形状復元法の検討", FIT2007 講演論文集, pp.133-134, 2007.9
- [4]佐藤 幸男、加藤 哲孝、"多面体像の情報量と視方向評価", 電子情報通信学会論文誌, D-II, pp1346-1352, (1992)
- [5]石川 尋代、佐藤 幸男、"多面体認識における心理量と物体像のエントロピーの関係について", 電子情報通信学会信学技報, PRMU2000-113 (2000-11)
- [6]大場 聡司、猪飼 武夫、山下 拓哉、福永 邦雄、"三角パッチ表現向き形状情報量による三次元物体の視点選択", 第 4 回情報科学技術フォーラム, pp. 89-90, (2005)