全方位画像列からの周囲環境モデル構築

川西亮輔(静岡大学)山下 淳(静岡大学)金子 透(静岡大学)

Construction of Environment Model from an Omni-Directional Image Sequence

*Ryosuke KAWANISHI (Shizuoka University), Atsushi YAMASHITA (Shizuoka University), Toru KANEKO (Shizuoka University)

Abstract— Map information is important for path planning and self-localization when mobile robots accomplish autonomous tasks. In unknown environments, however, mobile robots should generate an environment map by themselves. We propose a method for 3D environment modeling by a mobile robot. A 3D environment model can be generated from the result of 3D measurement using image data. To realize a 3D measurement of objects more efficiently, the robot uses an image sequence acquired by an omni-directional camera which has a wide field of view. The measurement method is based on structure from motion (SFM). A triangular mesh is constructed from measurement data. Experimental results showed the effectiveness of the proposed method.

Key Words: Omni-Directional Image, Structure from Motion, Triangular Mesh, Environment Model

1. 序論

移動ロボットが経路生成や自己位置同定を行うため には地図情報が必要である.しかし,ロボットが未知 環境で自律的に行動する場合にはロボット自身が周囲 環境を計測し,地図を作成する必要がある.このとき, 未知環境下における自律移動ロボットの活動のための 地図情報としては3次元の幾何学的な情報を持つ環境 モデルが有効である.また,テクスチャを含むことで, より有用な情報を持つことができる.

地図生成にはカメラによって視野内の物体の計測を 行うアプローチがある.これには魚眼レンズを用いた カメラによる計測[1]や周囲 360 度の視野を持つ全方 位カメラによる計測[2]がある.ロボットに搭載するこ とを考慮すると,ロボットの前後左右を一度に撮影で きる全方位カメラが有効である.

周囲環境の計測には,対象の計測とロボットの位置姿勢推定を同時に行うことができるStructure from Motion(SFM)と呼ばれる手法を用いる.本手法では,動 画像中から物体の角などの特徴的な点(特徴点)を追 跡し,ロボットの移動前後の2画像間で対応する特徴 点(対応点)を取得する.対応点情報を元に,ロボッ トの位置姿勢推定と対象の計測を行う.

環境モデルの生成には,得られた計測点群から三角 網を構築し,三角網に周囲環境のテクスチャマッピン グを行い,色情報を持つ環境モデルを生成する.

SFM による計測の精度向上のためには,より精度良 く位置・姿勢を推定することが重要である.また三角 網を構築することによるモデリングでは,計測対象の 形状に合った三角網を構築する必要がある.そこで本 研究では,より精度良く位置・姿勢を推定するための 手法の提案および計測対象の形状に対して三角網を最 適化し実環境に即した環境モデルを生成することを目 的とする.

2. 処理概要

実験装置は移動ロボットに,ビデオカメラに双曲面 ミラーを取り付けた全方位カメラを搭載したものを使 用する(図1).移動ロボットを環境中で走行させ,走 行中に撮影した動画像を用いて計測を行う.

本研究の処理手順を以下に示す.まず動画像を取得 する.最初の画像中から特徴点を抽出し,以後のフレー ムで追跡を行う.追跡を行った画像間で対応点が取得 できる.対応点情報を用いて線形推定により位置・姿 勢を推定し,推定された位置・姿勢を用いて対応点の3 次元座標を算出する.その後,非線形推定によって位 置・姿勢をより精度良く推定する.こうして最終的に 得られた位置・姿勢を用いて対応点の3次元座標を再 度算出する.最後に複数回の計測結果を合成する.

- 3. 環境計測
- 3·1 対応点抽出

ロボットの移動前後の2画像間における対応点を取得 するため,移動前の画像から特徴点を抽出,移動後の画 像まで追跡する.本手法では,Lucas Kanade Tracker と画像をピラミッド構造化して探索する手法を組み合 わせた特徴点追跡手法[3]を用いる.

特徴を定量的に評価したものを特徴量と呼ぶ.特徴 量が大きいほど追跡が容易な点で画像間の対応がとり



Fig.1 移動ロボットと全方位画像

やすい.逆に特徴量が小さい点ほど対応が取りにくいため追跡に失敗する可能性が高い.そこで,本手法では特徴量が大きい順に特徴点を抽出する.ここで抽出する特徴点の総数を抽出数と呼ぶ.

3·2 基本行列 E の算出

カメラから 3 次元空間中の点に向かう光線を光線ベ クトル $\mathbf{r} = [x, y, z]^T$ とする.本研究で用いる全方位カ メラはカメラ前方に双曲面ミラーを装着したものであ り,この光線ベクトル \mathbf{r} は,双曲面のミラー側の焦点 から 3 次元空間中の点へ向かう光線が双曲面上で反射 する点へのベクトルで表せる(図2).

特徴点の画像座標 $[u, v]^T$ から式 (1)(2) により r を求める.ここでa, b, cは双曲面のパラメータであり, fは焦点距離である.

$$\mathbf{r} = \begin{bmatrix} su\\ sv\\ sf - 2c \end{bmatrix} \tag{1}$$

$$s = \frac{a^2 \left(f \sqrt{a^2 + b^2} + b \sqrt{u^2 + v^2 + f^2} \right)}{a^2 f^2 - b^2 (u^2 + v^2)}$$
(2)

そして,後の演算で特徴点ごとの偏差を生じさせな いように ||r||の大きさを単位ベクトルに変換し,正規 化する.

計測に用いる 2 画像間の対応点の光線ベクトル $\mathbf{r}_i = [x_i, y_i, z_i]^T$, $\mathbf{r}'_i = [x'_i, y'_i, z'_i]^T$ を用いて幾何関係により カメラ間の位置・姿勢情報からなる行列を求める.こ の行列は基本行列 E と呼ばれ,式(3)で表される.式 (3)を変形すると式(4)が得られる.

$$\mathbf{r}^{\prime T} \mathbf{E} \mathbf{r} = 0 \tag{3}$$

$$\mathbf{u}^{\mathrm{T}}\mathbf{e} = 0 \tag{4}$$

ただし,

$$\mathbf{u} = [x_i x'_i, y_i x'_i, z_i x'_i, x_i y'_i, y_i y'_i, z_i y'_i, x_i z'_i, y_i z'_i, z_i z'_i]^T$$

$$\mathbf{e} = [e_{11}, e_{12}, e_{13}, e_{21}, e_{22}, e_{23}, e_{31}, e_{32}, e_{33}]^T$$

(*e_{jk}*はEの*j*行*k*列成分)

基本行列 E は 8 点以上の対応する光線ベクトル対に 対する連立方程式, すなわち式 (5) を解いて求める.こ こで, n は特徴点の数とする.

$$\min \| \mathbf{U} \mathbf{e} \|^2 \tag{5}$$

ただし, $\mathbf{U} = [u_1, u_2, ..., u_n]^T$ である. e は $\mathbf{U}^T \mathbf{U}$ の最 小固有値に対応する固有ベクトルとして与えられ,基 本行列 e が求まる.

3·3 外れ値除去

特徴点追跡によって得られた対応点は全てが正しく 対応しているとは限らない.対応を誤っている点が計 測に含まれると位置・姿勢の推定に悪影響を及ぼす. そこで,位置・姿勢に悪影響を及ぼす点を外れ値とし て除去する.これには RANSAC (RANdom SAmple Consensus) [4] を用いる.

まず基本行列 E が決定する最小の対応点数である 8 個の特徴点をランダムで抽出し \mathbf{E}_{rand} を求める.そし て光線ベクトル \mathbf{r}_i , \mathbf{r}'_i に対して式 (6)を満たす特徴点 の数を k とする.



Fig.2 全方位カメラの座標系

$$|\mathbf{r}'_{i}^{T}\mathbf{E}_{rand}\mathbf{r}_{i}| < q \tag{6}$$

ただしq は閾値である.この \mathbf{E}_{rand} , k を求める処理を あらかじめ設定した回数繰り返す.そしてk が最大と なったときの式(6)を満たさない特徴点を外れ値とす る.その後,外れ値を除いた特徴点を用いて式(5)よ り基本行列 E を算出する.

3.4 抽出数の自動決定

一般に8組の特徴点があれば位置・姿勢の推定が行え るが、実際は各種の誤差の影響があり、抽出数が少ない と正確な位置・姿勢の推定は困難である.環境の詳細な モデル化のためには抽出数が多いほど良いが、抽出数 が多すぎると特徴量の小さい点が抽出され、位置・姿勢 推定に悪影響を及ぼす(3.1節).そのため、RANSAC によって外れ値とされる特徴点の数(外れ値数)が増 加する.外れ値数が多いほど位置・姿勢の推定精度が 悪いと考えられる.したがって、本手法では位置・姿 勢の推定精度が十分に期待できる最大の外れ値数とな るときの抽出数を最適な抽出数と考え、これを自動的 に決定する.ここで以下の式(7)(8)を定義する.

$$\zeta(k) = \left| \frac{w \sum_{i=k}^{k+w} i \cdot O_i - \sum_{i=k}^{k+w} i \sum_{i=k}^{k+w} O_i}{w \sum_{i=k}^{k+w} i^2 - \left(\sum_{i=k}^{k+w} i\right)^2} \right|$$
(7)
$$\zeta(k) < G$$
(8)

ここで O_i は抽出数iのときの外れ値数,G は閾値, $\zeta(k)$ は抽出数がkから(k + w)まで変化するときの外れ値数の増加量を最小二乗法で求めたものである.wを適切に設定することで安定して外れ値数の増加量を求めることができる.

追跡が容易で画像間の対応がとりやすい特徴点の数 は環境ごとに異なるが、いずれにしてもその数は限ら れている.したがって、適切な抽出数を超えると外れ値 数が急激に増加するという現象が起こる.これを利用 し、式(8)を満たす最大の k を抽出数として算出する.

3.5 位置・姿勢の初期推定

基本行列 E は回転行列 R と並進移動ベクトル t = $[t_x, t_y, t_z]^T$ により式 (9) で表される.

$$\mathbf{E} = \mathbf{RT} \tag{9}$$

ただし,T =
$$\begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix}$$
である.
E に特異値分解を行い,R と T を求める.

3·6 対応点の3次元計測

算出された回転行列 R と並進移動ベクトルt を用い て,対応点の3次元計測を行う.移動前と移動後のそ れぞれの観測点のカメラ位置から3次元空間中の特徴 点へ向かう光線ベクトルが交わる点を計測結果とする (図3).

3.7 計測結果の限定

計測精度は一般にカメラの移動方向の延長線付近ほ ど,あるいはカメラから遠いものほど悪い.そこで計 測精度の良い対応点を使用するため,計測結果 $\mathbf{p}_{m,i}$ を 2 画像上の特徴点座標 $[u_{m,i}, v_{m,i}]^T$, $[u'_{m,i}, v'_{m,i}]^T$ で偏 微分したものから式 (10) で誤差を評価するベクトル g を求める.そしてベクトル g のノルムが式 (11) を満た す計測結果を以後の処理に用いる.ただし,h は閾値 である.

$$\mathbf{g} = \left| \frac{\partial \mathbf{p}_{m,i}}{\partial u_{1,m,i}} \right| + \left| \frac{\partial \mathbf{p}_{m,i}}{\partial v_{1,m,i}} \right| + \left| \frac{\partial \mathbf{p}_{m,i}}{\partial u_{2,m,i}} \right| + \left| \frac{\partial \mathbf{p}_{m,i}}{\partial v_{2,m,i}} \right| \quad (10)$$

$$||\mathbf{g}|| < h \tag{11}$$

3.8 バンドル調整

3.5節で推定した位置・姿勢は特徴点に含まれる誤差 を考慮していないため,必ずしも良い推定結果である とは限らない.そこで各特徴点の計測誤差を考慮して 位置・姿勢を推定する.これには3.4節で推定した位 置・姿勢を初期値とし,特徴点の再投影誤差の二乗和 を最小化するバンドル調整法[5]と呼ばれる手法を用い る.再投影誤差和を計算する際,特徴点ごとに計測誤 差を評価し適切に重みをつけることで,より精度の良 い推定をすることができる.

ここで再投影誤差とは,特徴点の3次元座標値を画 像上に再投影したときの座標と,元の画像座標との差 である.これが小さいほど位置・姿勢の推定精度が良 い.特徴点の再投影されたときの座標は式(12)~(14) より求められる.



Fig.3 対応点の3次元計測

観測点2

観測点1

$$d' = \frac{f}{dz' + 2c} \tag{14}$$

画像の再投影誤差和を以下のように定義する.

$$E_f = \sum_p r_{fp} \left| \mathbf{x}_{fp} - \mathbf{x}'_{fp} \right|^2 \tag{15}$$

ここで, E_f は第 f フレームでの再投影誤差和, r_{fp} は 第 f フレームの p 番目の特徴点に対する重み係数, \mathbf{x}_{fp} は特徴点の元の画像座標, \mathbf{x}'_{fp} は再投影したときの画像 座標である.重み係数は特徴点ごとに誤差を評価する ことにより決定する.誤差の評価には 3.7 節の式 (10) で表されるベクトルgのノルムを用い,以下の式で定 義する.

$$\mathbf{r}_{fp} = \frac{1}{\left\| \mathbf{g}_{fp} \right\|} \tag{16}$$

ここで, g_{fp} は第 f フレームの p 番目の対応点の誤差 ベクトルである.これにより対応点の計測誤差を考慮 し,適切に重みをつけることができる.

3.9 計測結果の合成・統合

各観測点間で得られた計測結果を,各観測点の位置・ 姿勢推定結果を用いて重ね合わせ,合成する.同じ特 徴点を示す計測結果でも,含まれる誤差の違いから観 測点間ごとに計測結果は異なる.したがって,各観測 点間の計測結果を重ね合わせるとずれが生じ,同じ特 徴点を示す計測結果が複数存在することになる.そこ で本手法では空間をボクセルで区切り,各特徴点の計 測誤差を考慮した評価値(式(16))で重み付けをした 投票を行うことで,特徴点の3次元位置を推定し計測 結果を統合する.

3·10 三角網の生成

得られた計測点群から3次元のドロネー分割により 三角網を構築する.しかし,ドロネー分割によって得ら れる多面体構造は計測対象の形状を考慮していないた め,物理的な辺に矛盾する三角網が生成されることが ある.そこで本手法では,矛盾を検出するテンプレー トを導入して三角網を物理的な形状に適合するように 最適化する手法[6]を適用する(図4).これにより実 際の環境と矛盾しないモデルを生成することができる.

3.11 テクスチャの貼り付け

各観測点で撮影した画像の中でテクスチャの解像度 が最大となる画像を選び,三角網の各面に貼り付ける. しかし,隣り合った三角形に貼り付けるテクスチャが, それぞれ異なる地点で撮影された画像から選択されて いる場合,境界が不自然になる.そこで,境界付近で は異なる複数の画像をアルファブレンディングにより 重ね合わせることで,境界の不自然さを解決する.



Fig.4 三角網の最適化

4. 実験

屋内環境(図5(a)L字廊下,(b)部屋)でそれぞれ計 測を行った.L字廊下における計測では,全方位カメ ラを搭載した移動ロボットを約10cm/sの速度で走行 させながら撮影を行った.部屋における計測では人が 全方位カメラを手に持って移動しながら撮影を行った.

廊下および部屋の計測結果を上から見た図6,8に示 す.青い点がロボットの軌跡,赤い点が計測点である. 提案手法(3.3,3.4,3.7,3.8節)を適用しない場合(a) に比べ,提案手法を適用した場合(b)の結果は,計測 結果のばらつきが少なく環境を正確に表しており,精 度の良い計測ができていることが分かる.

廊下および部屋のモデリング結果を図7,9に示す. 生成された環境モデルに貼り付けられたテクスチャの 歪みが少ないことから,実環境と矛盾しない三角網が 構築できているといえる.

5. 結論

本論文では,より精度良く位置・姿勢推定をするた めの手法の提案をした.また,計測対象の形状に対し て三角網を最適化する手法を適用した.その結果,計 測精度および環境モデルの精度が向上し,本手法の有 効性が示された.

今後の課題としては,より正確にカメラの位置・姿勢を推定するための手法として,複数の観測点の位置・ 姿勢を同時に推定する手法の構築が考えられる.

謝辞

本研究の一部は, 文部科学省 科学研究費補助金 若 手研究 (B) 20700184, および財団法人 立石科学技術振 興財団の補助を受けた.

- [1] 宮川 勲, 若林 佳織, 荒川 賢一: "魚眼投影モデルに基づく 画像系列からの平面運動と3次元形状の復元", 電子情報 通信学会論文誌, Vol. J87-D-II, No. 5, pp. 1120-1132, 2004.
- [2] R. Bunschoten and B. Krose: "Robust Scene Reconstruction from an Omnidirectional Vision System ", IEEE Transactions on Robotics and Automation, Vol. 19, No. 2, pp. 351-357, 2003.
- [3] J. Y. Bouguet: "Pyramidal Implementation of the Lucas Kanade Feature Tracker Description of the Algorithm ", OpenCV, Intel Corporation, 2000.
- [4] M. A. Fischler and R. C. Bolles: "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography", Communications of the ACM, Vol. 24, No. 6, pp. 381-395, 1981.
- [5] B. Triggs, P. McLauchlan, R. Hartley and A. Fitzgibbon: "Bundle Adjustment -A Modern Synthesis ", Vision Algorithms: Theory & Practice, Springer-Verlag LNCS 1883, 2000.
- [6] 中辻 敦忠, 菅谷 保之, 金谷 健一 "ビデオ画像からの形 状復元のための三角網の最適化", 情報処理学会研究報 告 CVIM-149-2, pp. 7-14, 2005.







ミ手法なし (b) 提案手法あり Fig.6 廊下の計測結果(上視点)



(a) 壁 1



(b) 壁 2Fig.7 廊下の環境モデル



(a) 提案手法なし
 (b) 提案手法あり
 Fig.8 部屋の計測結果(上視点)









Fig.9 部屋の環境モデル(左:実際の環境,右:環境モデル)