

全方位画像列を用いた Line-based Structure from Motion

川西 亮輔 山下 淳 金子 透

静岡大学工学部 〒432-8561 静岡県浜松市中区城北 3-5-1

E-mail: {f5945016, tayamas, tmtkane}@ipc.shizuoka.ac.jp

あらまし 本論文では、1台の全方位カメラから得られる画像列のみを入力とし、撮影時のカメラ運動を推定するとともに撮影環境の3次元計測およびモデリングを行う手法を提案する。提案手法では、画像中から抽出・追跡される点と直線の2つの特徴を利用する。点と直線の2つの特徴を3次元復元することで、テクスチャの豊富な物体とテクスチャレスな物体の両方の形状を計測することが可能である。

キーワード 全方位カメラ, Structure from Motion, 3次元環境モデル

1. 序論

移動ロボットが未知環境で自律的に活動する際には、ロボット自身が周囲環境を計測し、地図を生成する機能を持つことが重要である。環境計測を行うためのセンサとして、広い視野を持つ全方位カメラが考えられる。全方位カメラは、周囲360度を1枚の画像に撮像できるという特性により、環境計測に有効であることが示されている[1]。本研究は、1台の全方位カメラを用いて撮影した画像列から Structure from Motion の枠組みで環境計測を行い、環境の3次元モデルを構築する手法を提案する。

我々は文献[2]において全方位カメラを用いた Structure from Motion による3次元環境モデリング手法を提案した。しかし、文献[2]は画像中の特徴点のみを用いた手法であるため、特徴点が得られにくい人工物のような物体が多く存在する環境では疎な計測点群しか得られないという問題がある。

そこで本研究では、特徴点に加え、直線も同時に利用する計測手法を提案する。直線を利用した計測の従来手法としては、ビルの直線形状と消失点を利用した手法[3]や、線分を利用した単眼カメラによる SLAM [4]などが挙げられる。これらの従来研究における直線検出は円状に歪んだ全方位画像に直接適用することが困難であり、また限定された環境のみにしか有効でない[3]、直線を抽出するためにその両端が常に見えている必要がある[4]などの問題がある。

そこで本研究では、直線の端点情報を利用することなく歪んだ全方位画像から直線を判別・抽出する手法を提案する。特徴点に加えて直線の3次元計測を行うことで、様々な物体が存在する環境下におい

て、より密な3次元計測および環境モデル生成を実現する。

2. 環境計測

2.1. 全方位カメラ座標系

本研究では全方位カメラとして、通常のカメラの先端に双曲面ミラーを取り付けたものを使用する(図1)。カメラを移動させながら取得した全方位画像列を用いて以降の処理を行う。

全方位カメラの座標系を図2に示す。提案手法におけるカメラ座標系は、双曲面ミラーの焦点を投影中心とするモデルとして考えることができる。画像座標 $[u, v]^T$ は、以下の式(1)、(2)により投影中心から3次元空間に向かうベクトル \mathbf{r} (以下、光線ベクトルと呼ぶ) に変換される。

$$\mathbf{r} = \begin{bmatrix} \lambda(u - c_x)p_x \\ \lambda(v - c_y)p_y \\ \lambda f - 2\gamma \end{bmatrix} \quad (1)$$

$$\lambda = \frac{\alpha^2(f\gamma + \beta\sqrt{u^2 + v^2 + f^2})}{\alpha^2 f^2 - \beta^2(u^2 + v^2)} \quad (2)$$

ここで、 $[c_x, c_y]^T$ は全方位画像の中心座標、 α 、 β 、 γ は双曲面ミラーの形状パラメータ、 f はカメラの焦点距離である。



図1 全方位カメラ(左)と取得画像(右)

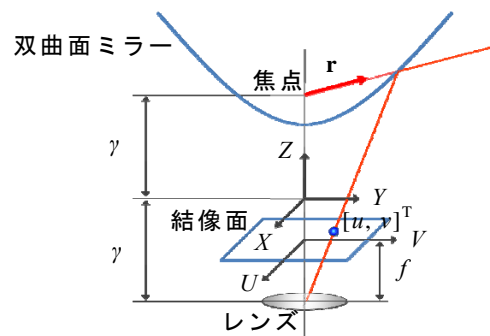


図2 全方位カメラ座標系

2.2. 特徴追跡

KLT 法[5]により全方位画像列で特徴点の抽出・追跡を行う。これにより、離れた視点の画像間における対応点が取得できる。

提案手法では、特徴点と同時に直線の抽出・追跡も同時に行う。全方位画像は円状に歪んでいるため、3次元空間においては直線であっても画像中では曲線として結像されるという問題がある。提案手法ではこの問題を解決する直線検出手法として、以下のアルゴリズムを考案した。

全方位画像から Canny のオペレータによってエッジ点を検出する (図 3(a))。コーナー点を除去することで、エッジ点を分離する (図 3(b))。分離された各エッジセグメントに対し直線判定を行う。具体的には、以下の式(3)を満たすエッジ点が過半数となるエッジセグメントを直線とする。

$$\left(\mathbf{r}_{i,j}^T \mathbf{n}_i \right)^2 < l_{th} \quad (4)$$

ここで、 $\mathbf{r}_{i,j}$ はエッジセグメント i に含まれるエッジ点 j に向かう光線ベクトル、 \mathbf{n}_i はエッジセグメントに向かう光線群から算出した最小二乗平面の法線ベクトル、 l_{th} は閾値である。光線ベクトル $\mathbf{r}_{i,j}$ は、3次元直線と双曲面ミラー焦点を通る平面上に存在する。したがって、十分に小さい閾値 l_{th} において上記の条件を満たすエッジセグメントは直線であるといえる (図 3(c))。これにより、直線が画像中で曲線として結像される場合においても、その直線性を正しく判定することが可能である。

直線エッジを画像間で対応付け、画像列で追跡を行う。現フレームから検出された直線エッジに、画像上で等間隔 (実験では 1 ピクセル程度とした) になるようにサンプリング点を抽出する (図 4(b))。次フレーム中でエッジセグメントを抽出する (図 4(c))。サンプリング点と対応する次フレームでの点 (以下、対応点) を探索する (図 4(d))。次フレーム中で対応点とのユークリッド距離が最も短いエッジ点を探索し (図 4(e))、そのエッジ点の属するエッジセグメントに現フレームの直線ラベルを投票する。すべてのサンプリング点の投票が終了したとき、最多投票数で、かつ投票数とその直線のサンプリング数が過半数であるものを、次フレームのエッジセグメントと対応している直線であるとみなす (図 4(f))。



(a)エッジ点 (b)コーナー点除去 (c)直線

図 3 直線検出

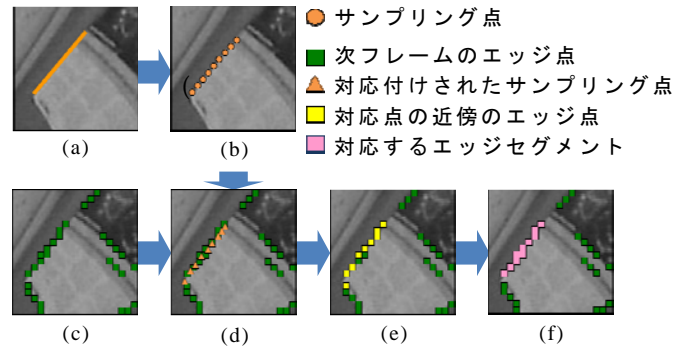


図 4 直線追跡

2.3. カメラ運動の初期推定

対応点を用いてカメラ運動の初期推定を行う。点対応情報を利用することで、線形解法による高速なカメラ運動推定が可能である。本研究では、8点法[6]と、RANSAC[7]による誤対応点の除去を組み合わせた手法によりカメラ運動を推定する。この推定結果は、次節の直線対応を用いたカメラ運動の最適化における初期値として利用される。

2.4. カメラ運動と 3次元計測の最適化

直線対応からのカメラ運動推定では、少なくとも 3 視点からの画像と、互いに平行でない 3 本以上の直線の対応関係が必要である。ここでは特徴点を、ランダムに選ばれた 2 つの特徴点を結んだ線分として扱うことで、直線と同じように評価することができる。

直線対応によるカメラ運動推定では、回転運動と並進運動を独立して推定することができる。提案手法では、以下の式(5)、(6)で定義される誤差関数を最小化する回転行列 \mathbf{R} と直線の 3次元方向ベクトル \mathbf{d} を推定する。

$$E_R = \sum_c \sum_i \left(\left(\mathbf{R}_c^w \mathbf{n}_i^c \right)^T \mathbf{d}_i^w \right)^2 \quad (5)$$

$$\sum_c \left(\mathbf{R}_c^w \mathbf{n}_i^c \right)^T \mathbf{d}_i^w \rightarrow \min \quad (6)$$

ここで、 \mathbf{R}_c^w は世界座標系 w からカメラ座標系 c への回転行列、 \mathbf{n}_i^c はカメラ座標系 c における直線 i の法線ベクトル (2.2 節)、 \mathbf{d}_i^w は直線 i の世界座標系における 3次元方向ベクトルであり、式(6)を満たす。誤差関数の最小化にはレーベンバーグ・マーカート法を用いる。

次に、並進運動を推定する。並進運動の最適化は、バンドル調整[8]に基づく再投影誤差の最小化によって行う。具体的には、以下の誤差関数を最小化する並進ベクトル \mathbf{t}_c^w と直線の 3次元位置ベクトル \mathbf{l}_i^w を推定する。

$$E_i = \sum_c \sum_i \left(\mathbf{l}_i^c - \mathbf{l}_i^{cT} \mathbf{g}_i^c \right)^2 \quad (7)$$

ここで、 \mathbf{l}_i^c は各カメラ座標系 c から直線 i に対して垂直に向かうベクトルで、式(8)より算出される。

$$\mathbf{l}_i^c = \mathbf{R}_c^w \left(B_{i,c} \mathbf{d}_i^w + \mathbf{l}_i^w - \mathbf{t}_c^w \right) / \left\| B_{i,c} \mathbf{d}_i^w + \mathbf{l}_i^w - \mathbf{t}_c^w \right\| \quad (8)$$

$$\left\| \left(A_{i,c} \mathbf{R}_c^{wT} \mathbf{g}_i^c + \mathbf{t}_c^w \right) - \left(B_{i,c} \mathbf{d}_i^w + \mathbf{l}_i^w \right) \right\| \rightarrow \min \quad (9)$$

ただし、 $A_{i,c}$ 、 $B_{i,c}$ は式(9)から算出される定数である。また、 \mathbf{g}_i^c は $\mathbf{g}_i^c \perp \mathbf{n}_i^c$ 、 $\mathbf{g}_i^c \perp (\mathbf{R}_c^w \mathbf{d}_i^w)$ の単位ベクトルである。直線の3次元位置ベクトル \mathbf{l}_i^w は式(10)を満たすベクトルとして算出される。

$$\sum_c \left\| \left(A_{i,c} \mathbf{R}_c^{wT} \mathbf{g}_i^c + \mathbf{t}_c^w \right) - \left(B_{i,c} \mathbf{d}_i^w + \mathbf{l}_i^w \right) \right\| \rightarrow \min \quad (10)$$

誤差関数の最小化には、回転運動の推定と同様に、レーベンバーク・マーカート法を用いる。また、カメラ運動の最適化ステップにおいても RANSAC によりカメラ運動推定に悪影響のある（すなわち対応精度が悪いと思われる）直線を除去することで、カメラ運動推定の安定化を図る。

最終的に得られたカメラ運動推定結果を用いて、特徴点と直線の3次元計測結果を得る。直線は点群の集合として表現される。

2.5. 3次元モデル生成

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。そこで提案手法では、計測対象の物理的な形状に適合するように三角網を最適化する手法[9]を適用する。

各観測点で撮影した画像のテクスチャを三角網の各面に貼り付けることで、モデルに色情報を付加する。ここで、全方位画像からそのまま三角形を生成すると、テクスチャに全方位画像特有の歪みが生じる。そこで、全方位画像の歪みの影響を低減したテクスチャ生成手法を適用し、より実際の環境に忠実なモデルを生成する。

3. 実験

実験では、画像サイズ 2496×1664 ピクセルの全方位画像を用いた。処理はすべてオフラインで行った。

提案手法によるカメラ運動推定の評価を図4に示す環境で行った。カメラは図4中の小円で示した位置から移動を開始し、4回の90度回転を経て出発地点に戻る運動とした。走行距離はおよそ60mで、移動中に得られた約1000枚の全方位画像を入力とする。ここで、出発地点の画像を入力画像列の最初と

最後の画像とした。これにより、理想的には出発地点と到着地点が一致する。これを真値として評価を行う。

出発地点の3次元座標を(0,0,0)とした場合の到着地点の座標は(-0.008,-0.016,0.021)で、出発地点と到着地点の距離誤差は0.028であった。ただし Structure from Motion はカメラ運動の絶対スケールが未知であるため、これは最初に計測を行った基線長を1とした場合の値である。このスケールでの走行距離は11.8であり、実際の走行距離（約60m）から実スケールに換算すると、誤差は約14cmとなった（図5）。移動距離との比で0.4%となり、非常に精度の良いカメラ運動推定が行えたといえる。また、今回はクロージング処理による最適化を行っていないため、今後さらなる安定化および精度向上が見込める。

より広範囲での環境計測およびモデリングの実験を行った。実験環境は屋外で、テクスチャの豊富な物体とテクスチャの少ない物体が混在している（図6）。走行距離は約150mで、取得した約2500枚の全方位画像を入力とした。

3次元計測の鳥瞰図を図7(a)に、上視点からみた図を(b)にそれぞれ示す。建物の輪郭などからは直線が計測され、その他テクスチャのある部分は特徴点によって計測されており、特徴点と直線の両方の特徴を用いることで密な3次元計測が行えた。

生成された環境モデルを図8((a)鳥瞰図、(b)上視点図)に示す。3次元構造とテクスチャとの整合性が取れており、実際の環境に近いモデルが生成できている。また、建物の角など色合いの変化に乏しく特徴点を得られにくい部分に関しては直線エッジが計測されているため、より建物の形状を再現したモデルとなっている。



図4 カメラ運動推定の評価実験環境

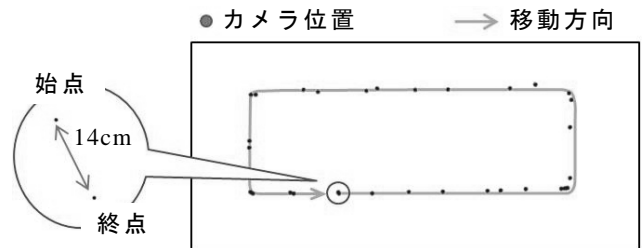
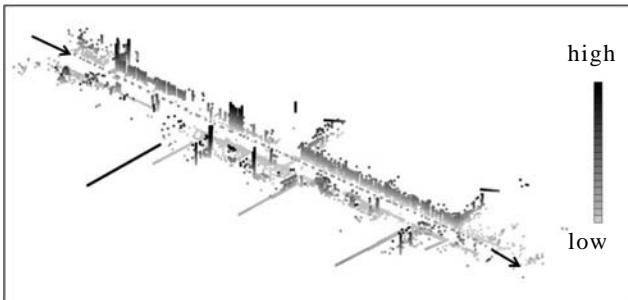


図5 カメラ運動推定結果

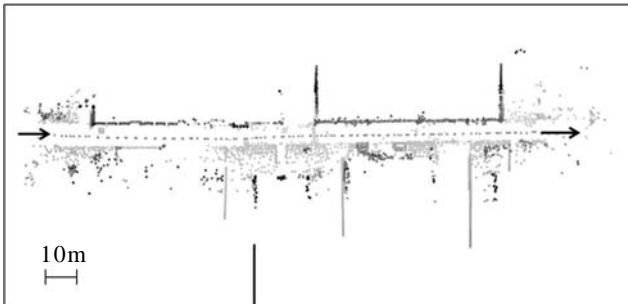


図 6 環境モデリングの実験環境

→... カメラ位置 ■ 計測結果



(a) 鳥瞰図

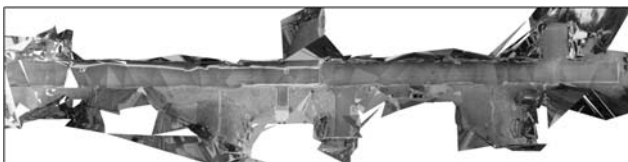


(b) 上視点図

図 7 環境の3次元計測結果



(a) 鳥瞰図



(b) 上視点図

図 8 環境の3次元モデル生成結果

4. 結論

本研究では、特徴点と直線の2つの特徴を用いたカメラ運動推定および計測手法を提案した。実験により、提案手法の有効性を示した。

今後の課題としては、より広範囲の環境計測を効率的に行うための手法の構築が挙げられる。具体的には、視野外に外れた画像特徴が再び視野内に現れる際の対応付けや、計測のクロージング処理などである。

文 献

- [1] J. Gluckman and S. K. Nayar: "Ego-motion and Omnidirectional Cameras", *Proceedings of the 6th International Conference on Computer Vision*, pp. 999-1005, 1998.
- [2] R. Kawanishi, A. Yamashita and T. Kaneko: "Estimation of Camera Motion with Feature Flow Model for 3D Environment Modeling by Using Omni-Directional Camera", *Proceedings of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.3089-3094, 2009.
- [3] G. Schindler, P. Krishnamurthy and F. Dellaert: "Line-Based Structure from Motion for Urban Environments", *Proceedings of the 3rd International Symposium on 3D Data Processing, Visualization, and Transmission*, pp. 846-853, 2006.
- [4] P. Smith, I. Reid and A. Davison: "Real Time Monocular SLAM with Straight Lines", *Proceedings of the 17th British Machine Vision Conference*, pp. 17-26, 2006.
- [5] J. Shi and C. Tomasi: "Good Features to Track", *Proceedings of the 1994 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 593-600, 1994.
- [6] R. Hartley: "In defense of the eight-point algorithm", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 6, pp. 580-593, 1997.
- [7] 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.
- [8] B. Triggs, P. McLauchlan, R. Hartley and A. Fitzgibbon: "Bundle Adjustment -A Modern Synthesis", *Proceedings of the International Workshop on Vision Algorithms: Theory and Practice*, Springer-Verlag LNCS 1883, pp.298-372, 1999.
- [9] A. Nakatsuji, Y. Sugaya, and K. Kanatani: "Optimizing a Triangular Mesh for Shape Reconstruction from Images", *IEICE Transactions on Information and Systems*, Vol. E88-D, No. 10, pp. 2269-2276, 2005.
- [10] 川西 亮輔, 瀨瀬 理志, 山下 淳, 金子 透: "単眼全方位カメラを用いた環境復元における2視点からの見えの違いを利用した物体メッシュモデル最適化", 第17回画像センシングシンポジウム講演論文集, IS2-06, pp.1-7, 2011.