

全方位画像の特徴点と直線エッジ分布に基づく 基線長の最適化アルゴリズムを有する Structure from Motion による 3次元環境計測

川西 亮輔^{*1}, 山下 淳^{*1}, 金子 透^{*1}

Three-dimension Environment Measurement based on Structure from Motion with Baseline Optimization by using Distribution of Feature Points and Straight-lines in Omnidirectional Image

Ryosuke KAWANISHI^{*1}, Atsushi YAMASHITA^{*1} and Toru Kaneko^{*1}

^{*1} Department of Mechanical Engineering, Shizuoka University
3-5-1 Johoku, Naka-ku, Hamamatsu-shi, Shizuoka 432-8561, Japan

In this paper, we propose a method for 3D environment modeling based on structure from motion using an omnidirectional camera installed on a mobile robot. Our proposed method uses correspondent straight-lines in addition to feature points tracked on an omnidirectional image sequence. Straight-lines make it possible to estimate camera movement precisely. Baseline length have to be optimized for precise estimation of camera movement. We propose a baseline determination method by evaluated translation and distribution of these two features. The evaluation method is invariant for camera rotation, environments and the number of features. Experimental results show the effectiveness of our proposed method.

Key Words : Baseline optimization, Map construction, Structure from Motion, Omnidirectional camera

1. 緒 言

移動ロボットの自律活動を実現するための重要な課題として、搭載センサによる自己位置推定と地図構築が挙げられる。これら 2 つの課題を同時に解決する方法として、カメラ画像から得られる情報を利用した手法が研究されている。カメラ画像列から各撮影地点におけるカメラの位置姿勢推定と撮影物体の計測を同時に行う枠組みの従来研究としては、Structure from Motion に基づく手法⁽¹⁾⁽²⁾や SLAM による手法⁽³⁾などが提案されている。

広範囲にわたる環境の地図構築をすることを考えると、一般的な視野のカメラを用いた場合には非常に多くの撮影回数が必要となるため、広い視野を持つ撮像センサが有効である。広い撮像センサを利用した従来研究としては、全方位カメラ⁽⁴⁾⁽⁵⁾や魚眼カメラ⁽⁶⁾による手法などが挙げられる。本研究では、1枚の画像にロボットの周囲 360度を撮像できる、双曲面ミラーを利用した全方位カメラ(図1)を用いる。

本研究では、全方位カメラにより取得した画像列を

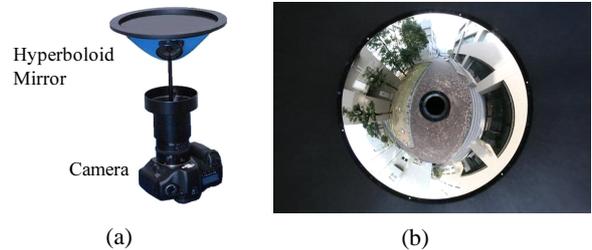


Fig. 1 Omnidirectional camera. (a) Camera with a hyperboloid mirror. (b) Acquired image.

用いて、Structure from Motion に基づく環境計測を行う。Structure from Motion は、1台のカメラを移動させながら撮影した画像列から、被写体の計測とカメラ運動(カメラの回転行列および並進ベクトル)の推定を同時に行う枠組みである。我々はずでに全方位カメラを用いた Structure from Motion に基づく環境計測手法を提案している⁽⁷⁾。我々の従来研究では、特徴点に加え直線エッジを画像中から抽出し、精度の良いカメラ運動推定と密な環境計測を実現した。エッジ情報を利用した計測手法が有効であることは、近年、多くの研究者によって示されている⁽⁸⁾⁽⁹⁾。

我々の従来手法における課題の1つとして、計測を

^{*1} 静岡大学工学部機械工学科 (〒 432-8561 浜松市中区城北 3-5-1) {f5945016, tayamas, tmtkane}@ipc.shizuoka.ac.jp

行うカメラ間の距離（基線長）の決定方法が挙げられる．一般に，画像計測は基線長が長いほど計測精度が良い．しかし実際の環境計測では，カメラと物体との位置関係によっては画像間の正確な対応づけが困難となるなどの原因で，基線長が長すぎると逆に計測精度が低下することがある．したがって，状況に応じて基線長を適切に選択する必要がある．

我々の従来研究⁽⁷⁾では基線長を手動で決定していた．しかし，ロボットが自律的に環境計測を行うためには，適切な評価に基づいて基線長を自動的に決定できる手法を適用することが望ましい．

基線長を自動的に決定するための指標を定義した従来研究として，Torrらの手法⁽¹⁰⁾がある．Torrらは，2枚の画像間の対応点群から算出される行列のランクを調べることで，画像を撮影したカメラ間の距離が十分であったかどうかを判定できることを示した．この指標をもとに基線長を自動的に決定したSLAM⁽¹¹⁾などが提案されている．

Torrらの手法はカメラ間の並進の有無を判定することが可能であるが，全方位カメラを用いた環境計測手法に適しているとはいえない．全方位カメラは環境を広く撮像するため，さまざまな方向・距離の物体が画像中に存在する．その結果，カメラ間距離がわずかであってもカメラ運動を推定するのに十分な情報が得られるケースが多い．また，通常の視野のカメラと比べ，計測のために必要十分な数の対応点を取得できるカメラ間距離が非常に長い．このような条件下においては，Torrらの手法のみでは安定した基線長決定は困難である．

そこで本研究では，全方位カメラによる環境計測に適用可能な基線長の自動決定手法を提案する．提案手法は，画像中における特徴の分布の偏りと，カメラ運動に対する画像特徴の相対的な並進移動量を評価し，基線長を自動的に決定する．このとき，画像間で対応付けられた特徴の絶対数やカメラ運動の回転，速度変化によらない評価方法を用いることで，安定して基線長を決定することが可能である．実験では，一旦停止やその場旋回を含んだカメラ運動において，提案手法が有効に機能することを示す．

2. 処理の概要

提案手法の処理概要を以下に述べる（図2）．

カメラの移動とともに画像を取得する．取得した画像の中から特徴点と直線エッジ（合わせて画像特徴と呼ぶ）を抽出し，前画像で取得した画像特徴と対応付ける特徴追跡を行う．

現在地点で取得した画像と，それより以前に通過し

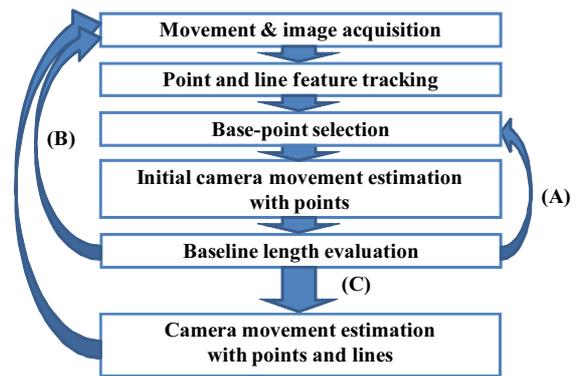


Fig. 2 Procedure of our proposed method. (A) The evaluation value does not satisfy the condition and base-point is not first observation point. (B) The value does not satisfy it and base-point is first observation point. (C) The value satisfies it.

た地点で取得した画像との間の対応点を用いて，8点法によりカメラ運動を推定する．このとき選択された移動前の地点を，計測の基準となる地点（以後，基点と呼ぶ）とする．推定されたカメラ運動における基線長の評価を行う．基線長が評価基準を満たさない場合は，さらに以前の地点を基点とし，現在画像と基点画像間で上記の処理を繰り返す（図2(A)）．ただし，移動開始直後などのように，カメラの移動距離が最大となる地点間を基線長としても評価基準を満たさず基線長が存在しない場合がある．そこで，基点が初期位置で，かつ基線長が評価基準を満たさない場合は，移動および画像取得のステップに戻る（図2(B)）．

基線長が評価基準を満たした場合（図2(C)），画像特徴の画像座標と算出された3次元座標との間の誤差を最小化するカメラ運動を推定する．ここで，直線エッジを計測するためには，少なくとも3枚の画像が必要となる．そこで，基点と現在地点に加え，これらの地点の中間に位置する地点（以後，中間点と呼ぶ）のカメラ運動も同時に推定する．以上の手順を繰り返し，環境の3次元計測を行う．

3. 画像特徴の抽出と追跡

提案手法では，物体のコーナー点などの追跡が容易な特徴点と，直線エッジを抽出し，画像間で対応付けることにより，画像特徴の追跡を行う．

3.1 特徴点の抽出と追跡 KLTトラッカ⁽¹²⁾による特徴点追跡により，画像間の対応点を得る．特徴点は画像中の物体のコーナーのように大きな輝度変化が発生している位置や，テクスチャが豊富な位置に抽出されやすい性質がある．また，特徴点追跡によって得

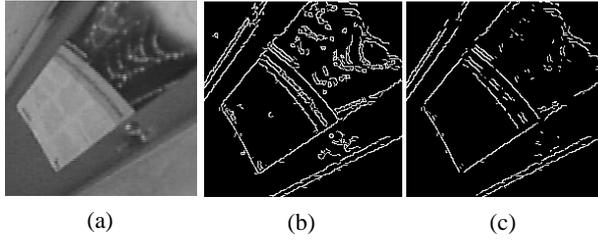


Fig. 3 Edge segment extraction. (a) Input image. (b) Extracted canny edge points. (c) Edge points without rejected corner points.

られる画像間の対応点は，線形解法による高速なカメラ運動推定を可能とする．

3.2 直線エッジの抽出と追跡 画像中から直線的なエッジを抽出し，画像列で追跡を行う．端点を必要とする直線エッジの抽出手法はノイズに弱く，全方位画像特有の歪みの影響を受けやすいため，我々は端点を必要としない直線エッジの抽出および追跡手法を提案している⁽⁷⁾．以下に直線エッジ追跡処理の概要を述べる．

3.2.1 エッジセグメントの抽出 全方位画像中からエッジセグメントを抽出する．まず，Cannyのエッジ点検出⁽¹³⁾によって，画像中の輝度勾配の大きい点がエッジ点として得られる（図3(b)）．エッジ点のうち，コーナ点と思われるエッジ点を判定し，除去することで，分離されたエッジセグメントを得る（図3(c)）．

3.2.2 直線エッジの判定 エッジセグメントの中から，直線を検出する．これには本研究で用いる全方位カメラを構成する双曲面ミラーの特性を利用する．3次元空間からカメラに向かう光線は，双曲面ミラーを反射してレンズ中心を通る．このとき，3次元空間中の点から双曲面ミラー上の反射点に向かう光線を延長すると，必ず双曲面の焦点を通る．したがって，直線のエッジセグメントに向かう光線は一平面上に存在する（図4）．そこでまず，同一のエッジセグメントに属するエッジ点座標 (u, v) から，ミラー焦点から反射点に向かう光線 \mathbf{r} を算出する（(1), (2) 式）．

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

$$s = \frac{\alpha^2 \left(f\sqrt{\alpha^2 + \beta^2} + \beta\sqrt{u^2 + v^2 + f^2} \right)}{\alpha^2 f^2 - \beta^2(u^2 + v^2)} \quad (2)$$

ここで， f はカメラの像距離， (c_x, c_y) は全方位画像の中心， p_x, p_y はピクセルサイズ， α, β, γ は双曲面のパラメータである．後の演算での光線ごとの偏差をなくすため \mathbf{r} を単位ベクトルに変換し，正規化する．

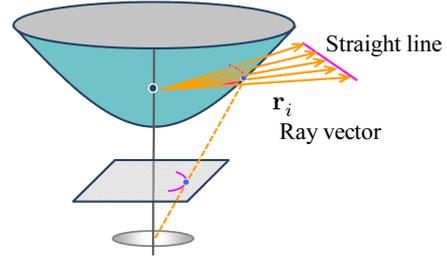


Fig. 4 Straight line and ray vector.

エッジセグメント i に属するエッジ点 j の光線ベクトル $\mathbf{r}_{i,j}$ を用いて，光線ベクトルのなす平面の法線ベクトル \mathbf{n}_i を算出する．ベクトル \mathbf{n}_i は，(3) 式を満たすエッジ点の数が最大となるときのベクトルとして算出される．(3) 式を満たすエッジ点の数がエッジセグメントに属する全エッジ点の過半数以上である場合，そのエッジセグメントを直線であると判定する．

$$\mathbf{r}_{i,j}^T \mathbf{n}_i < e_{th} \quad (3)$$

閾値 e_{th} は，推定された直線と画像エッジ点との誤差が0.5ピクセル以内となる値とする．また，明らかに直線に属さないセグメント中のエッジ点をRANSAC⁽¹⁴⁾によって除去することで，直線でないものの誤検出を低減する．

3.2.3 直線エッジの追跡 直線エッジを画像間に対応付け，画像列で追跡を行う．現画像から検出された直線エッジ i 上に，画像上で約1画素間隔になるようにサンプリング点を得る（図5(a)）．次画像中でエッジセグメントを抽出する（図5(b)）．サンプリング点と対応する次画像での点（以下，対応点）をKLT法により取得する（図5(c)）．次画像中で対応点とのユークリッド距離が最も短いエッジセグメント j を探索する．次画像中のエッジセグメント j の対応候補としてラベル i に投票する（図5(d)）．すべてのサンプリング点の投票が終了したときの最多投票数のラベル i_{max} の直線エッジが，次画像のエッジセグメント j と対応しているとみなす（図5(e)）．

ただし，投票数が現画像の直線エッジ i のサンプリング数の過半数に満たず，かつ次画像のエッジセグメント j のエッジ点数の過半数にも満たない場合には，次画像のエッジセグメント j に対応する現画像中の直線エッジ i は存在しないと判定する．以上の処理を画像間で繰り返すことで直線エッジを追跡する．

ここで，一般に直線エッジ上での対応付けには開口問題があるが，探索ウィンドウサイズを十分な大きさに設定することで，上記の処理において直線エッジ自体を誤ることは少なくなる．計測には直線エッジ上

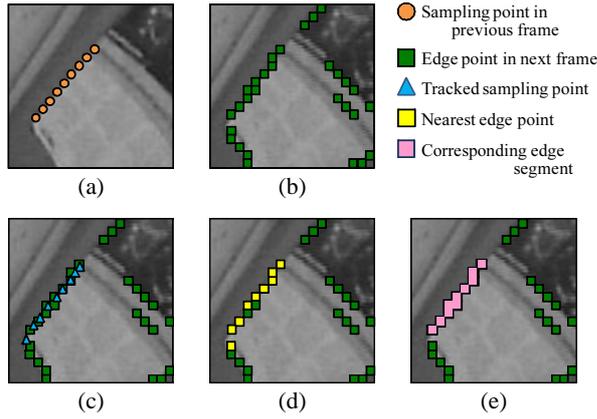


Fig. 5 Search for corresponding edge segment. (a) Sampling points in previous frame. (b) Edge points in next frame. (c) Tracked sampling points. (d) Points which are close to tracked sampling points. (e) Corresponding edge segment.

での点对点の対応ではなく、前述の法線ベクトル \mathbf{n}_i (3.2.2 項) を利用するため、開口問題による悪影響は少ない。

4. カメラ運動の初期推定

取得した画像間の対応点 (3.1 節) を用いて 8 点法によるカメラ運動推定を行う。同時に RANSAC⁽¹⁴⁾ により明らかに誤対応と判断される点を除去する。このとき用いる画像対は、基点で取得した移動前画像と現在の観測点で取得した移動後画像である。

5. 基線長の自動決定

抽出された画像特徴と、推定されたカメラ運動を用いて、基線長の評価を行う。提案手法では基線長の評価に、移動前後での画像特徴の分布の変化と、画像特徴の視差の 2 つの基準を用いる。画像特徴の分布の変化は以下の (4) ~ (6) 式で表わされる。

$$g_p = \left\| \sum_i^{n_p} (\mathbf{r}_{b,i} - \mathbf{R}_{b,h}^T \mathbf{r}_{h,i}) / n_p \right\|^2 \quad (4)$$

$$g_l = \left\| \sum_i^{n_l} (\mathbf{n}_{b,i} - \mathbf{R}_{b,h}^T \mathbf{n}_{h,i}) / n_l \right\|^2 \quad (5)$$

$$G = \frac{g_p n_p + g_l n_l}{n_p + n_l} \quad (6)$$

ここで、 b, h はそれぞれ基点と現在地点の画像番号、 n_p, n_l はそれぞれ画像間の対応点と対応エッジの数、 $\mathbf{r}_b, \mathbf{r}_h$ はそれぞれ基点と現在地点から対応点に向かう光線、 $\mathbf{n}_b, \mathbf{n}_h$ はそれぞれ基点と現在地点において

直線エッジから算出された法線ベクトル (3.2.2 項)、 $\mathbf{R}_{b,h}$ は基点に対する現在地点のカメラ運動の回転行列である。

g_p, g_l はそれぞれ画像間の特徴点と直線エッジの分布の重心変化を意味しており、移動前に対して移動後のカメラの並進量が大きいほど大きな値になる。この評価では、移動後ベクトルに基点に対する現在地点の回転行列 \mathbf{R}_{bh} の転置行列をかけることで、移動前後のカメラの回転運動の影響を排除している。また、対応点の光線ベクトルと対応エッジの法線ベクトルの画像間の変化量を評価に用いることで、カメラの絶対的な並進量ではなく周囲の物体に対するカメラの相対的な並進量を評価する。それぞれ (4), (5) 式に対応点数と対応エッジ数の比をかけて和をとった G (6 式) をカメラ運動の並進量の評価値とする。画像特徴の数で正規化することで、画像特徴の絶対数によらない評価となる。

画像特徴の視差は以下の (7) ~ (9) 式で表わされる。

$$m_p = \sum_i^{n_p} \|\mathbf{r}_{b,i} - \mathbf{R}_{b,h}^T \mathbf{r}_{h,i}\| \quad (7)$$

$$m_l = \sum_i^{n_l} \|\mathbf{n}_{b,i} - \mathbf{R}_{b,h}^T \mathbf{n}_{h,i}\| \quad (8)$$

$$M = \frac{m_p + m_l}{n_p + n_l} \quad (9)$$

m_p, m_l は、カメラ間の回転成分を排除した並進成分のみによる画像特徴の見かけの位置の変化、すなわち視差の総変化量を意味している。視差が大きいほど、理想的には計測精度が向上するため、 M (9 式) が大きいほど良い評価となる。

現在地点 h に対する基点 b における基線長の評価関数 $f(G_b)$ を、評価基準 G_b, M_b を用いて以下のように定義する。

$$f(G_b) = M_b \{1 - G_b\} \quad (10)$$

M_b によって画像特徴の視差量を評価する。また、精度の良いカメラ運動推定のためには、画像特徴が画像全体に均等に分布していることが望ましい。そこで視差値に $\{1 - G_b\}$ をかけることで、画像特徴の分布が偏りすぎないように制御する。評価関数 f をカメラ間の並進量の指標である G_b の関数とすることで、カメラの加減速による悪影響を低減した評価が可能となる。ここで、提案手法では直線エッジの並進量の評価として前述の法線ベクトル \mathbf{n} を用いているが、これは法線ベクトルの方向の変化が直線エッジの見え方の変化、すなわち視差と同義だからである。法線ベクトルの変

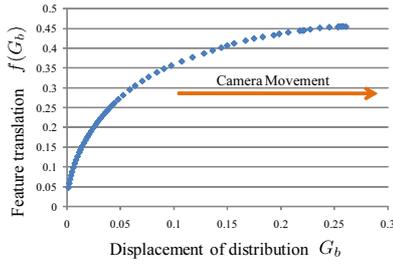


Fig. 6 Baseline evaluation.

化を評価に用いることで、点と同様の視差の評価が可能となる。

評価関数 f は、カメラの並進量が小さい場合には G_b による制御があまり働かないため、 M_b の増加が支配的となる。基線が長くなるにつれ、画像特徴の偏りによる G_b の制御が働くことで、評価関数 f の増加が緩やかになる。基線が一定以上の長さになると、評価関数 f は横ばいになる（ピークに達する）。実験で得られた評価関数 f のグラフの例を図 6 に示す。

評価関数がピークに達したとき、画像特徴の視差はそれ以降の画像とでほとんど変化しなくなっていることを意味する。したがって、そのときの画像特徴の見え方の違いは非常に大きく、その対応付けの精度は低下していると考えられる。以上のことから、対応付けの精度が保証できる基線長は、ピークに達する以前の、評価関数が増加傾向を示している範囲であるといえる。

一般的な環境において、厳密な意味で基線長を最適化することは非常に困難である。そこで提案手法では、評価関数 $f(G_b)$ を累乗近似 ($f(x) = ax^b$) し、その傾きが閾値以下になったときの基線長を自動的に選択する。

提案する評価関数は、周囲の物体とカメラとの相対的な並進量以外の変化に対して不変なため、近似曲線の傾きはカメラの並進量に依存する。閾値が小さいほど（評価関数の傾きが小さいほど）、周囲の物体とカメラとの位置関係において相対的に長い基線が選択されるようになる。ここで、安定して近似曲線を算出するためには、一定数以上の基点における基線長の評価結果が得られた段階で基線長の判定を開始する必要がある。その数は実験的に 5 点とした。

6. 特徴点と直線エッジを用いたカメラ運動の最適化

提案手法では、対応点だけでなく対応エッジも同時に用いてカメラ運動を最適化する。対応点から算出されたカメラ運動は、全方位画像に含まれる歪みによる特徴点の対応誤差などが含まれているが、提案手法による対応エッジは全方位歪みの影響を低減しているため、対応点を用いた非線形最小化手法よりも誤差の小さいカメラ運動推定が可能である⁽⁷⁾。しかし、対応

エッジのみを用いてカメラ運動の最適化を行うと、環境によっては十分な数の対応エッジが得られず、システムが不安定になることがある。そこで両方の特徴を同時に評価するカメラ運動推定の枠組みを提案する。

6.1 中間点の選出 提案手法では、対応エッジを用いてカメラ運動を推定するため、少なくとも 3 枚の画像情報が必要となる。そこで、基点と現在の観測点との中間に位置する観測点（中間点）を選出する。中間点は、基線長選択に用いる評価基準 M が、決定された基線長における評価値 M_b の半分の値になる地点とする。基点から中間点までのカメラ運動の初期値は 8 点法によって算出する。

6.2 回転運動の最適化 対応エッジを用いたカメラ運動では、回転成分のみを独立に算出することができる。そこでまずカメラの回転運動を最適化する。1 枚の画像から得られる対応エッジ i の情報は、観測点 o からエッジに向かう光線群のなす平面の法線ベクトル $\mathbf{n}_{i,o}$ として表わされる。 $\mathbf{n}_{i,o}$ は、対応エッジ i の 3 次元方向 \mathbf{d}_i と観測点ごとのカメラの回転行列 \mathbf{R}_o との間で以下の (7) 式が成り立つ。

$$(\mathbf{R}_o^T \mathbf{n}_{i,o})^T \mathbf{d}_i = 0 \quad (11)$$

3 観測点間の法線ベクトル \mathbf{n}_i を用い、(7) 式を 3 つ並べた連立方程式を解くことでエッジの 3 次元方向 \mathbf{d}_i を算出できる。これらのパラメータを用い、与えられた回転行列の誤差は以下のように算出される。

$$\sum_o \sum_i \left((\mathbf{R}_o^T \mathbf{n}_{i,o})^T \mathbf{d}_i \right)^2 \quad (12)$$

観測点 $o-1, o$ から対応点 i に向かう光線 $\mathbf{r}_{i,o-1}, \mathbf{r}_{i,o}$ と、観測点 o の $o-1$ に対するカメラ運動 \mathbf{R}, \mathbf{t} との間には以下の (9) 式が成り立つ⁽¹⁵⁾。

$$(\mathbf{r}_{i,o-1} \times \mathbf{R}^T \mathbf{r}_{i,o})^T \mathbf{t} = 0 \quad (13)$$

(9) 式には並進ベクトルが含まれているが、複数の光線対を用いて \mathbf{t} についての連立方程式として解くことで、光線と回転行列のみを用いて並進ベクトルを暫定的に算出できる。このようにして算出された並進ベクトルは、与えられた回転行列がカメラ運動と等しいとき、カメラ運動の並進ベクトルと一致する。ただし、各光線対とカメラ運動は以下の拘束条件 ((10), (11) 式) を満たす必要がある⁽¹⁵⁾。

$$\mathbf{q}_1 \cdot \mathbf{q}_2 < 0 \cap \mathbf{q}'_1 \cdot \mathbf{q}'_2 < 0 \quad (14)$$

$$\begin{aligned}
\mathbf{q}_1 &= \mathbf{R}^T \mathbf{r}_{i,o} - \mathbf{r}_{i,o-1} \\
\mathbf{q}_2 &= \mathbf{t} - \mathbf{r}_{i,o-1} \\
\mathbf{q}'_1 &= -\mathbf{q}_1 \\
\mathbf{q}'_2 &= -\mathbf{t} - \mathbf{R}^T \mathbf{r}_{i,o}
\end{aligned} \tag{15}$$

以上より，各パラメータを用いて，回転行列の誤差は以下のように算出される．

$$\sum_o \sum_i \left((\mathbf{r}_{i,o-1} \times \mathbf{R}^T \mathbf{r}_{i,o})^T \mathbf{t} \right)^2 \tag{16}$$

ただし，上記の拘束条件を満たす対応点の光線対には(12)式どおりの値を誤差値として与えるが，拘束条件を満たさない光線対の誤差値には1を与える．これにより，拘束条件を満たした解に収束させる．

(8)式と(12)式の和を最小化する回転行列を探索することで，カメラの回転運動を推定する．

6.3 並進運動の最適化 最適化された回転運動を用いて並進運動の誤差を最小化する．

対応エッジの3次元位置を表す位置ベクトル \mathbf{p}_i は，世界座標系から各観測点 o に向かう並進ベクトル \mathbf{t}_o と，対応エッジ i の方向ベクトル \mathbf{d}_i に対して垂直に向かうベクトル $\mathbf{g}_{i,o}$ から，三角測量の原理で算出することができる．具体的には，以下の(13)式を満たす \mathbf{p}_i を算出する．

$$(A_{i,o} \mathbf{g}_{i,o} + \mathbf{t}_o) + (B_{i,o} \mathbf{d}_i + \mathbf{p}_i) = 0 \tag{17}$$

ここで， $\mathbf{g}_{i,o}$ は $\mathbf{g}_{i,o}^T \mathbf{d}_i = 0$ かつ $\mathbf{g}_{i,o}^T (\mathbf{R}_o^T \mathbf{n}_{i,o}) = 0$ の単位ベクトルである．また， $A_{i,o}$ ， $B_{i,o}$ は $\mathbf{g}_{i,o}$ と3次元エッジが最短距離となるときに係数である．この最短距離が並進運動の誤差と見なすことができる．よって対応エッジの並進運動に対する誤差は以下のように表わされる．

$$\sum_o \sum_i (1 - \mathbf{g}_{i,o}^T \mathbf{p}'_i)^2 \tag{18}$$

$$\begin{aligned}
\mathbf{g}'_{i,o} &= (A_{i,o} \mathbf{g}_{i,o} + \mathbf{t}_o) / \|A_{i,o} \mathbf{g}_{i,o} + \mathbf{t}_o\| \\
\mathbf{p}'_i &= (B_{i,o} \mathbf{d}_i + \mathbf{p}_i) / \|B_{i,o} \mathbf{d}_i + \mathbf{p}_i\|
\end{aligned} \tag{19}$$

(14)，(15)式は，各観測点から見た対応エッジの画像位置と，対応エッジの3次元位置から画像に再投影した画像位置の誤差と同義である．すなわち，この誤差を最小化することは，対応エッジの再投影誤差の最小化を行うことに等しい．

対応点は再投影誤差を最小化するバンドル調整⁽¹⁶⁾に基づいて誤差の最小化を図る．本研究で用いる全方位画像で再投影誤差を最小化することは，各観測点から対応点の3次元座標に向かう光線の単位ベクトル \mathbf{r}'_i

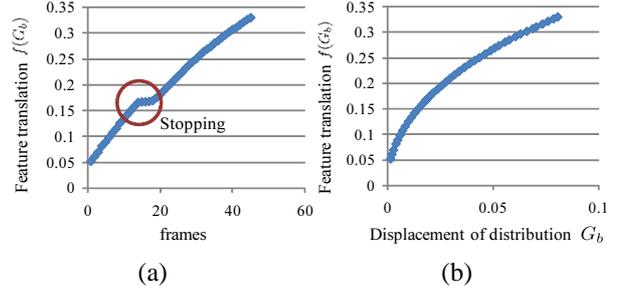


Fig. 7 Baseline evaluation with camera movement including stopping. (a) Evaluation value per frame. (b) Proposed method.

と対応点に向かう光線の単位ベクトル $\mathbf{R}_o^T \mathbf{r}_{i,o}$ の角度誤差を最小化することと同義である．そこで，以下の式で表わされる誤差を最小化する．

$$\sum_o \sum_i \left| \mathbf{r}'_i - (\mathbf{R}_o^T \mathbf{r}_{i,o} + \mathbf{t}_o) \right|^2 \tag{20}$$

(14)，(16)式の誤差和を最小化する並進ベクトルを探索することで，並進ベクトルの最適化を行う．

なお，本章での各処理ではRANSACにより誤差の大きい画像特徴を除外している．ここで，6章でRANSACなどにより誤差が大きいと判定された画像特徴は，その時点で追跡を中止し，破棄する．また画像特徴の中には，再投影誤差が小さくとも十分な基線長が得られなかったため計測精度が低いものが含まれる．計測精度が低い特徴は統合処理からは省くが，画像間での追跡は続行し，次回以降の処理にも用いる．次回以降で十分な基線長が得られる可能性があるからである．計測精度の評価は文献⁽¹⁵⁾の方法を用いる．

7. 実験

本論文の実験では，すべて 2496×1664 ピクセルの全方位画像列を使用した．また，本研究はすべての処理をオフラインで行っている．

7.1 基線長の自動決定 本研究が提案する基線長の自動決定手法の検証を行った．今回はすべての実験において基線長決定の閾値（累乗近似曲線の傾き）を2とした．まず，一旦停止を含むカメラ運動の場合の基線長決定の検証を行った．基線長が決定されるまでの評価関数の推移を図7に示す．画像特徴の視差による評価値を縦軸とし，横軸を移動開始時からのフレーム数としたグラフ（図7(a)）と，提案手法による評価関数のグラフ（図7(b)）を比較する．

視差の変化量のみを見た場合，カメラが一時停止するとフレーム数が増大しても視差が変化しないためグラフが横ばいとなる（図7(a)の円）．一方，提案手法ではグラフの横軸をカメラの並進運動の指標となる評

価値 G_b としているため、カメラ運動の速度変化の影響を受けず、等速直線運動の場合と同様な推移を示している。これは、画像特徴の視差変化量は、カメラの移動速度に依らずカメラの並進移動量に対して決まるからである。累乗近似曲線は $f(G_b) = 1.03 \times G_b^{0.459}$ となった。以上のことから、提案手法はカメラ運動の速度変化に対して不変であることが分かる。

次に、カメラが回転運動（約 90 度のその場旋回）を含む場合の基線長決定の検証を行った。基線長が決定されるまでの評価関数の推移を図 8 に示す。先程と同様に、フレーム数を横軸にとった場合のグラフ（図 8(a)）と、提案手法による評価関数のグラフ（図 8(b)）とを比較する。

カメラがその場旋回する場合、カメラは一時停止した状態になるため、旋回中はグラフが横ばいになる（図 8(a) の円）。また、カメラが回転運動をした場合、視差が変化しなくとも画像中の特徴分布の重心位置が変化（全方位画像の中心を軸とした回転移動）することが考えられる。しかし提案手法では、各評価値を算出する際に回転の影響を排除している。そのため結果的に提案手法のグラフ（図 8(b)）は等速直線運動の場合と同様な推移を示した。累乗近似曲線は $f(G_b) = 1.00 \times G_b^{0.450}$ となった。

上記の実験では、いずれも異なる環境、カメラ運動で行ったにもかかわらず、ほぼ同じ近似曲線が得られた。これは提案手法が、環境や並進以外のカメラ運動の変化に対して不変な評価値を用いることによって、常にカメラ運動の並進成分のみの等速直線運動として評価しているからである。したがって、基線長を選択する際の閾値（近似曲線の傾き）を事前に設定することによって、常に撮像対象との相対距離に対して一定の長さの基線を選択できる。逆に言えば、この閾値は撮像対象との距離に対する基線の相対的な長さを設定するものである。実際の環境ではカメラと撮像対象との位置関係は複雑であり、厳密な意味での最適な基線長を定義することは非常に困難である。提案手法は、撮像対象との距離に対して常に相対的に等しい長さとなる基線を最適な基線長として自動的に選択する手法である。

一旦停止を含むカメラ運動およびその場旋回を含むカメラ運動の実験における周囲環境の計測結果を図 9、10 に示す。地面に対して高い位置にある計測結果ほど赤く表示されている。いずれもカメラの並進量が各フレーム間で一定でないため、基線長の選択を誤ると計測が破たんする可能性のある状況である。一旦停止は図 9(c) の円内、その場旋回は図 10(c) の円内でそれぞ

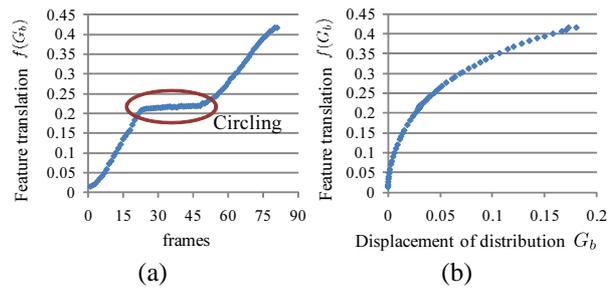


Fig. 8 Baseline evaluation with camera movement including circling. (a) Evaluation value per frame. (b) Proposed method.

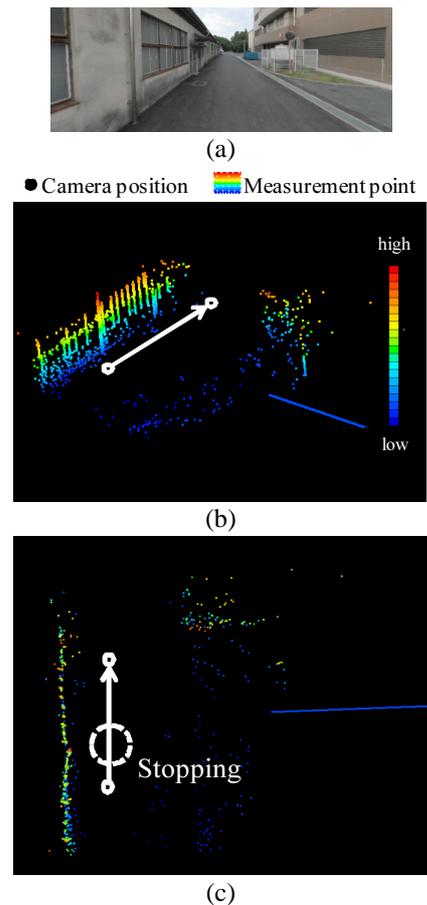


Fig. 9 Measurement results with camera movement including stopping. (a) Environment. (b) Bird's-eye view. (c) Top view.

れ発生している。しかし、提案手法により適切な基線長が選択されたため、安定してカメラ運動推定が行えた。撮像物体の計測結果を処理画像に投影した画像座標と、対応する特徴の画像座標の誤差は最大で 8 ピクセル以下であった。この結果からも、カメラ運動が精度良く推定できたといえる。

8. 結 言

本研究では、カメラ運動の変化に対応できる基線長の自動決定アルゴリズムを取り入れた環境計測手法を

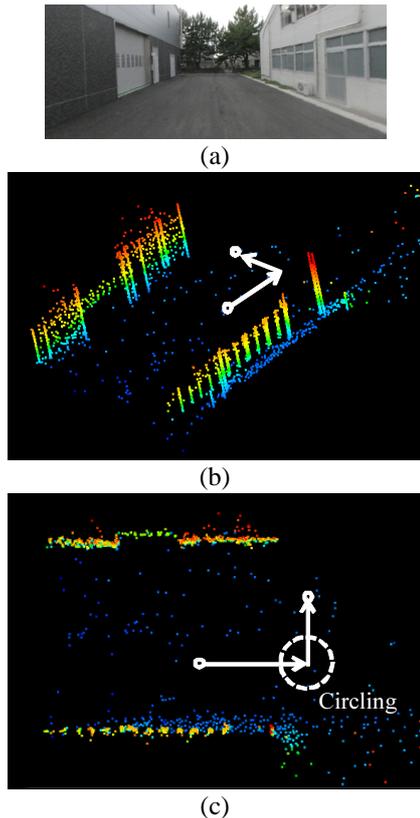


Fig. 10 Measurement results with camera movement including circling. (a) Environment. (b) Bird's-eye view. (c) Top view.

提案した . カメラ運動の加減速 , その場旋回などの計測が不安定になる状況下においても , 提案手法は安定して計測が行えることを示した .

今後の課題としては , より広範囲の環境計測を効率に行うための手法の構築が挙げられる . また , 提案手法による直線エッジ追跡では , 直線抽出の精度は高いが , 追跡のロバスト性が十分でないという問題がある . よりロバストに直線エッジの画像間の対応付けを行える手法を考案する必要がある .

参考文献

- (1) A. Bartoli and P. Sturm: "Multi-View Structure and Motion from Line Correspondences", *Proceedings of the 9th IEEE International Conference on Computer Vision*, pp. 207-212, 2003.
- (2) A. Murillo, J. Guerrero, and C. Sagues, "SURF Features for Efficient Robot Localization with Omnidirectional Images", *Proceedings of the 2007 IEEE International Conference on Robotics and Automation*, pp. 3901-3907, 2007.
- (3) A. J. Davison: "Real-Time Simultaneous Localisation and Mapping with a Single Camera", *Proceedings of the 9th IEEE International Conference on Computer Vision*, Vol. 2, pp. 1403-1410, 2003.
- (4) 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.
- (5) C. Geyer and K. Daniilidis: "Omnidirectional Video", *The Visual Computer*, Vol. 19, No. 6, pp. 405-416, 2003.
- (6) T. Nishimoto and J. Yamaguchi: "Three dimensional Measurement using Fisheye Stereo Vision", *Proceedings of the Society of Instrument and Control Engineers Annual Conference 2007*, 3A05-1, pp. 2008-2012, 2007.
- (7) 川西 亮輔, 山下 淳, 金子 透: "全方位画像列からの特徴点および直線エッジ情報を併用した Structure from Motion による 3 次元環境モデル生成", *電気学会論文誌 C*, Vol. 130-C, No. 9, pp. 1494-1503, 2010.
- (8) M. Chandraker, J. Lim and D. Kriegman: "Moving in Stereo: Efficient Structure and Motion Using Lines", *Proceedings of the 12th IEEE International Conference on Computer Vision*, pp. 1741-1748, 2009.
- (9) M. Tomono: "3D Object Mapping by Integrating Stereo SLAM and Object Segmentation Using Edge Points", *Lecture Notes in Computer Science*, Vol. 5875, pp. 690-699, 2009.
- (10) P. Torr, A. Fitzgibbon and A. Zisserman: "The Problem of Degeneracy in Structure and Motion Recovery from Uncalibrated Image Sequences", *International Journal of Computer Vision*, Vol. 32, No. 1, pp. 27-44, 1999.
- (11) M. Tomono: "3-D Localization and Mapping Using a Single Camera Based on Structure-from-Motion with Automatic Baseline Selection", *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, pp. 3353-3358, 2005.
- (12) 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.
- (13) J. F. Canny: "A Computational Approach to Edge Detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-8, No. 6, pp. 679-698, 1986.
- (14) 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.
- (15) 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.
- (16) 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.