

全方位カメラ搭載移動ロボットによる Structure from Motion を用いた 3 次元環境モデリング*

山下 淳^{*1,*2}, 原田 知明^{*1}, 金子 透^{*3}

Three-Dimensional Environment Modeling from Images Acquired with an Omni-Directional Camera Mounted on a Mobile Robot by Structure from Motion

Atsushi YAMASHITA^{*4,*5}, Tomoaki HARADA and Toru KANEKO

^{*4} Department of Mechanical Engineering, Shizuoka University
3-5-1 Johoku, Hamamatsu-shi, Shizuoka 432-8561, Japan

^{*5} Department of Mechanical Engineering, California Institute of Technology
1200 E. California Blvd. MC104-44, Pasadena, CA 91125, U.S.A.

Measurement and modeling of surrounding environments are important for mobile robots to move autonomously. This paper describes a new method for environment measurement using an omni-directional camera on a mobile robot. Our method is based on a structure-from-motion technique under the assumption that the environment is static. The method measures surrounding environments at the same time as estimating the robot's motion. The validity of our method is verified through experiments.

Key Words : structure from motion, omni-directional camera, 3-D modeling, single camera, estimation of position and posture

1. 序 論

本研究では、全方位カメラ搭載移動ロボットによる画像列を用いた 3 次元環境モデリング手法を提案する。

環境地図は、移動ロボットが移動している最中の自己位置同定や経路計画などに必要不可欠な情報である。よって未知環境で作業を行う必要がある場合、移動ロボット自身が環境地図を生成することが重要となる⁽¹⁾。

地図生成には、レーザレンジファインダを用いて周囲の物体までの距離を計測する手法⁽¹⁾や、カメラを用いる手法がある。カメラを用いた地図生成では、物体までの距離を計測できることに加え、物体表面のテクスチャを同時に取得できることが大きな利点の 1 つである。その際、通常のカメラと比較して視野が広い全方位カメラを利用することが有効である⁽²⁾。全方位カメラは、特殊な形状のミラーをカメラの前方に設置することにより、360 度の視野を有するカメラである⁽³⁾。

一般的に 3 角測量の原理を用いて物体までの距離を計測する際、全方位カメラを用いる場合にも 2 台以上

のカメラが必要である^{(4)~(6)}。ステレオ計測では基線長が長いほど精度が良いが、これらの手法では全方位カメラをロボット上に設置しているため基線長をある程度以上長くすることは困難である。このため計測対象が遠距離に存在する場合には、精度の良い計測ができないことがある。

これに対して、1 台のカメラを移動させることにより移動前後の画像からステレオ計測を行うモーションステレオ法が提案されている。移動ロボットにカメラを搭載する場合、基線長がロボットの移動距離となるため、移動距離を調節することにより計測対象が遠距離に存在する場合にも対応できる。ただしこの手法を用いるためには、カメラ間、すなわちロボットの移動前後の相対的な位置姿勢情報が必要となる。移動による位置姿勢変化をデッドレコニング（オドメトリ）により推定することが考えられるが、デッドレコニングは走行距離が長くなるに従い誤差が蓄積することが知られており精度の良い計測は望めない。また、GPS を用いて位置姿勢を推定する方法⁽⁷⁾があるが、屋内の計測には向かない。

これに対する手法として、画像情報のみからロボットの移動前後の相対的な位置姿勢を推定する手法が提案されている⁽⁸⁾⁽⁹⁾。この手法は SFM（Structure from Motion）と呼ばれ、カメラ以外のセンサ情報を必要と

* 原稿受付 2006 年 9 月 6 日

^{*1} 正員、静岡大学工学部機械工学科（〒 432-8561 静岡県浜松市城北 3-5-1）

^{*2} カリフォルニア工科大学機械工学科（〒 91125 米国カリフォルニア州パサデナ市）

^{*3} 静岡大学工学部

Email: yamashita@ieee.org

せずに周囲環境の計測が可能である⁽¹⁰⁾⁽¹¹⁾。

以上、移動ロボットが周囲環境のセンシングを行う際には、視野の広い全方位カメラを用い、誤差の蓄積しない SFM により画像情報から位置・姿勢を推定し環境計測を行うことが有効である^{(12)~(14)}。そこで本研究では、1 台の全方位カメラをロボットに設置し、ロボットの移動前後の 2 画像間で対応する特徴点の画像座標の関係からロボットの相対的な位置・姿勢を推定し、周囲環境の計測を行う手法を提案する。そして得られたカメラ間の位置姿勢情報を用いて複数の計測結果を組み合わせることで広範囲な環境計測データを取得し、3 次元モデリングを行う手法を構築する。

提案手法では、全方位のデータを 1 回の計測で同時に取得できることから、通常のカメラと比較して位置・姿勢推定精度の向上が期待できる。また、通常の画像よりも分解能が低く計測に適さない場所が全方位画像中には存在するという問題に対しては、外れ値や低精度の計測点の除去を行い、高精度の対象のみを計測することで対応する。

2. 提案手法の概要

1 台の全方位カメラを搭載したロボットを用い、ロボットの移動によって得られた撮影位置の異なる 2 画像に対してステレオ計測を行う。これには移動前後のカメラ間の相対的な位置関係を必要とする。すなわちカメラ間の回転行列 $\mathbf{R}$ と並進移動ベクトル $\mathbf{t}$ が必要である。本研究では、 $\mathbf{R}$ と $\mathbf{t}$ は画像間の対応点を用いて weakly calibrated stereo の原理により求められる⁽¹⁵⁾。

提案手法では、まず時系列画像間の対応点を取得する。具体的には、ロボットの移動中に動画を撮影し、画像中で特徴があり対応の取りやすい点群を特徴点として抽出し以後のフレームで追跡を行う。

そしてカメラ間の相対的な位置姿勢である $\mathbf{R}$, $\mathbf{t}$ を推定し、推定された $\mathbf{R}$, $\mathbf{t}$ を用いて特徴点の 3 次元計測を行う。ただし環境は全て静止しているものとする。また、カメラの内部パラメータは事前に取得しておき計測中に変化しないものとする。

次に、複数回の計測結果を合成する。合成方法の概念図を図 1 に示す。まず観測点 1 と 2 の画像を用いて、観測点 1 から 2 へのロボットの回転行列 $\mathbf{R}_1$ と並進移動ベクトル $\mathbf{t}_1$ を推定し画像中の対応点の 3 次元座標 $\mathbf{p}_{1,i}$ を計測する。ただし i は対応点の番号とする。次に観測点 2 と 3 の画像用いて $\mathbf{R}_2$ と $\mathbf{t}_2$ を推定し $\mathbf{p}_{2,i}$ を計測する。これにより得られた $\mathbf{R}_1$, $\mathbf{R}_2$, $\mathbf{t}_1$, $\mathbf{t}_2$ から観測点 1, 2, 3 の相対的な位置関係が決定する。ただし、この計測は画像の対応点情報のみを用いている

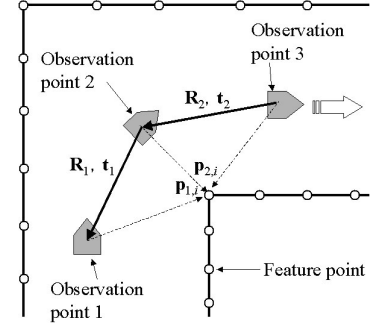


Fig. 1 Measurement data combination.

ため観測点間の距離は決定できず、計測結果は $\|\mathbf{t}_i\|$ に対して相対的なものとなる。このため観測点 1, 2 間の距離と観測点 2, 3 間の距離が異なる場合、 $\|\mathbf{t}_1\|$ と $\|\mathbf{t}_2\|$ を観測点間の距離の関係になるようにスケールを一致させる必要がある。スケール合わせでは、それぞれの計測で共通して使用した対応点の 3 次元座標 $\mathbf{p}_{1,i}$, $\mathbf{p}_{2,i}$ が一致するように $\|\mathbf{t}_i\|$ を求める。これにより計測間のスケールを合わせることができ、計測結果の合成が可能となる。上記の計測と合成を繰り返すことにより広範囲な計測結果が得られる。

最後に、計測で得られた点群と取得した全方位画像を用いてテクスチャ付き環境モデルを生成する。

3. 環境センシング手法

3.1 特徴点の取得 取得した動画画像から画像間で対応している点を取得するため、画像中の角や点など輝度値に特徴がある点を特徴点として抽出し、それらの点を動画画像中で追跡する。本手法では KLT (Kanade-Lucas-Tomasi) Tracker を用いる⁽¹⁶⁾⁽¹⁷⁾。

まず初期画像において特徴点を指定した l 個抽出し、その後の動画画像に対して特徴点を追跡する。対応が取れないことにより追跡できなかったものは逐次消去し、残った特徴点を対応点とする。次回の計測での初期画像には特徴点数が l 個となるように前回の計測で残った特徴点に加えて新しく特徴点を抽出する。

3.2 特徴点方向の算出 KLT Tracker により得られた特徴点の 3 次元空間上の点からカメラに対して入射する光線の逆方向の単位ベクトルを $\mathbf{r} = [x, y, z]^T$ とする¹。本研究で用いる全方位カメラはカメラ前方に (1) 式で表現される双曲面ミラーを配置したもの⁽¹⁸⁾であり、この光線ベクトル $\mathbf{r}$ は、全方位カメラの双曲面の焦点から特徴点の光線が双曲面上で反射する点へのベクトルとして表すことができる (図 2)。

$$\frac{X^2 + Y^2}{a^2} - \frac{Z^2}{b^2} = -1, \quad (1)$$

¹ベクトルまたは行列右肩の T は転置を表す。

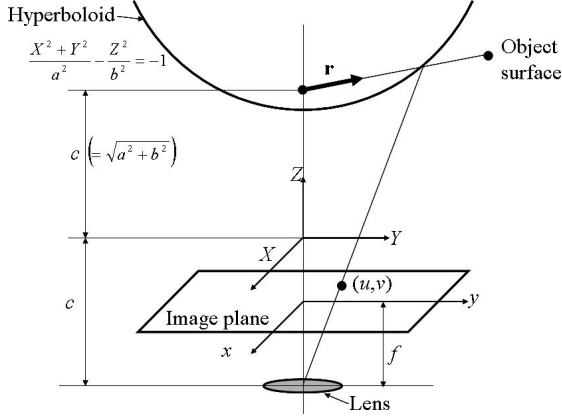


Fig. 2 Ray vector.

$$c = \sqrt{a^2 + b^2}, \quad (2)$$

ただし, a, b は双曲面のパラメータであり, c は原点と焦点の距離である.

特徴点の画像座標 $[u, v]^T$ から (3)(4) 式により $\mathbf{r}$ を求める.

$$\mathbf{r} = \frac{1}{(su)^2 + (sv)^2 + (sf - 2c)^2} \begin{pmatrix} su \\ sv \\ sf - 2c \end{pmatrix}, \quad (3)$$

$$s = \frac{a^2(f\sqrt{a^2 + b^2} + b\sqrt{u^2 + v^2 + f^2})}{a^2 f^2 - b^2(u^2 + v^2)}, \quad (4)$$

ここで, f は像距離 (カメラのレンズ中心から結像面までの距離) である.

3.3 基本行列 $\mathbf{E}$ の算出 計測に用いる 2 つの画像における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・姿勢の情報からなる行列を求める. この行列は基本行列 $\mathbf{E}$ と呼ばれ, (5) 式で表される.

$$\mathbf{r}'_i^T \mathbf{E} \mathbf{r}_i = 0. \quad (5)$$

(5) 式を変形すると (6) 式が得られる.

$$\mathbf{u}_i^T \mathbf{e} = 0, \quad (6)$$

ただし,

$$\mathbf{u}_i = [x_i x'_i, y_i x'_i, z_i x'_i, x_i y'_i, y_i y'_i, z_i y'_i, x_i z'_i, y_i z'_i, z_i z'_i]^T, \quad (7)$$

$$\mathbf{e} = [e_{11}, e_{12}, e_{13}, e_{21}, e_{22}, e_{23}, e_{31}, e_{32}, e_{33}]^T, \quad (8)$$

である (e_{ab} は $\mathbf{E}$ の a 行 b 列成分).

基本行列 $\mathbf{E}$ 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を解くことにより求められる. すなわち (9) 式を解くことになる. ただし, n は特徴点の数とする.

$$\min_{\mathbf{E}} \sum_{i=1}^n (\mathbf{r}'_i^T \mathbf{E} \mathbf{r}_i)^2. \quad (9)$$

これを变形すると,

$$\min_{\mathbf{E}} \|\mathbf{U} \mathbf{e}\|^2, \quad (10)$$

ただし, $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n]^T$ である.

$\mathbf{e}$ 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 $\mathbf{E}$ が求まる.

3.4 外れ値除去 特徴点追跡は全てが正しく対応しているとは限らない. そこで, 誤った追跡を行ったものを外れ値として除去する. これには RANSAC⁽¹⁹⁾ を用いる.

具体的には, まず基本行列 $\mathbf{E}$ 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. そして全特徴点の光線ベクトル $\mathbf{r}_i$, $\mathbf{r}'_i$ に対して (11) 式を満たす特徴点の数を k とする.

$$|\mathbf{r}'_i^T \mathbf{E}_{rand} \mathbf{r}_i| < q, \quad (11)$$

ただし q は閾値である.

この $\mathbf{E}_{rand}$, k を求める処理を繰り返す. k が最大になり以後の繰り返し回数が予め設定した数 w に達したとき, k が最大となったときの (11) 式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 (10) 式より基本行列 $\mathbf{E}$ を算出する.

3.5 位置・姿勢推定 基本行列 $\mathbf{E}$ は回転行列 $\mathbf{R}$ と並進移動ベクトル $\mathbf{t} = [t_x, t_y, t_z]^T$ により (12) 式で表される.

$$\mathbf{E} = \mathbf{R} \mathbf{T}, \quad (12)$$

ただし,

$$\mathbf{T} = \begin{pmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{pmatrix}. \quad (13)$$

本手法では画像のみを入力としスケール情報を得ずに計測を行うためカメラ間の距離 $\|\mathbf{t}\|$ が決定できない. そこでカメラ間の距離を 1 として計測を行うこととし, このために $\|\mathbf{t}\| = 1$ なるように $\|\mathbf{E}\|$ を変える. $\|\mathbf{t}\| = 1$ より $\mathbf{T}$ のフロベニウスノルムは $\sqrt{2}$ となり, (12) 式で $\mathbf{T}$ に回転行列 $\mathbf{R}$ を掛けてもノルムは変化しないため, $\|\mathbf{E}\| = \sqrt{2}$ とする ((14) 式).

$$\mathbf{E}' = \frac{\sqrt{2}}{\|\mathbf{E}\|} \mathbf{E}. \quad (14)$$

$\mathbf{E}'$ を $\mathbf{R}$ と $\mathbf{T}$ に分解するために特異値分解を行う.

$$\mathbf{E}' = \mathbf{U} \mathbf{\Sigma} \mathbf{V}, \quad (15)$$

ただし, $\mathbf{\Sigma} = \text{diag}(r, s, 0)$ である.

r, s は $\mathbf{E}'$ の特異値であり共に 1 付近の値となる. $\mathbf{R}$ を回転行列とするため $r = s = 1$ とする. これを (16) 式のように変形させることにより (17)(18) 式から $\mathbf{R}, \mathbf{T}$ を求める.

$$\mathbf{E}' = \mathbf{U}\mathbf{Y}\mathbf{V}^T\mathbf{V}\mathbf{Z}\mathbf{V}^T, \quad (16)$$

$$\mathbf{R} = \mathbf{U}\mathbf{Y}\mathbf{V}^T, \quad (17)$$

$$\mathbf{T} = \mathbf{V}\mathbf{Z}\mathbf{V}^T, \quad (18)$$

ただし,

$$\mathbf{Y} = \begin{pmatrix} 0 & \pm 1 & 0 \\ \mp 1 & 0 & 0 \\ 0 & 0 & \det \mathbf{U}\mathbf{V}^T \end{pmatrix}, \quad (19)$$

$$\mathbf{Z} = \begin{pmatrix} 0 & \pm 1 & 0 \\ \mp 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (20)$$

そして $\mathbf{T}$ の成分から $\mathbf{t}$ を作成する. $\mathbf{R}$ と $\mathbf{t}$ の解は複数存在するが, 特徴点の計測結果が光線ベクトルの向きと一致するものを正しい解とする.

上記ではロボットの 3 次元移動を許容して位置姿勢推定を行うが, ロボットが平面移動を行う場合は 2 次元移動に限定する必要がある. 2 次元移動に限定する場合, 前提としてロボットは床面に対して傾かず, カメラの光軸は床面に対して垂直に設置するものとする. このとき回転行列 $\mathbf{R}$ と並進移動ベクトル $\mathbf{t}$ は (21)(22) 式で表され, 基本行列 $\mathbf{E}$ は (23) 式で表される.

$$\mathbf{R} = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (21)$$

$$\mathbf{T} = \begin{pmatrix} 0 & 0 & t_y \\ 0 & 0 & -t_x \\ -t_y & t_x & 0 \end{pmatrix}, \quad (22)$$

$$\mathbf{E} = \begin{pmatrix} 0 & 0 & -t_x \sin \theta + t_y \cos \theta \\ 0 & 0 & -t_x \cos \theta - t_y \sin \theta \\ -t_y & t_x & 0 \end{pmatrix}. \quad (23)$$

$\mathbf{E}$ の可変成分は 4 つであり, (6) 式の $\mathbf{u}_i$ と $\mathbf{e}$ を

$$\mathbf{u}_i = [z_i x'_i, z_i y'_i, x_i z'_i, y_i z'_i]^T, \quad (24)$$

$$\mathbf{e} = [e_{13}, e_{23}, e_{31}, e_{32}]^T, \quad (25)$$

として解くことでロボットの移動前後の位置姿勢を推定することができる.

3.6 特徴点の計測 $\mathbf{R}_m, \mathbf{t}_m$ を m 回目の計測における 2 画像から推定された回転行列と並進移動ベクトルとすると, $\mathbf{R}_m, \mathbf{t}_m$ の位置関係にカメラが存在する

として 3 角測量の原理を用いて特徴点の計測を行う. すなわち 2 つのカメラからの光線ベクトルが交わる点を計測結果とする. 実際には画像ノイズなどの影響により光線ベクトルはねじれの関係にあることが多いため, 光線ベクトル間の距離が最短となる時の中点を計測結果とする. 特徴点 i の計測結果 $\mathbf{p}_{m,i}$ は (26) 式で表される.

$$\mathbf{p}_{m,i} = \frac{1}{2} \left\{ \frac{A_{m,i} \mathbf{r}_{m,i} + B_{m,i} \mathbf{r}'_{m,i}}{(\mathbf{r}_{m,i}^T \mathbf{r}_{m,i})(\mathbf{r}'_{m,i}^T \mathbf{r}'_{m,i}) - (\mathbf{r}_{m,i}^T \mathbf{r}'_{m,i})^2} + \mathbf{t}_m \right\}, \quad (26)$$

ただし,

$$A_{m,i} = (\mathbf{r}_{m,i}^T \mathbf{t}_m)(\mathbf{r}'_{m,i}^T \mathbf{r}'_{m,i}) - (\mathbf{r}_{m,i}^T \mathbf{R}_m^T \mathbf{r}'_{m,i})(\mathbf{r}'_{m,i}^T \mathbf{R}_m^T \mathbf{t}_m) \quad (27)$$

$$B_{m,i} = (\mathbf{r}_{m,i}^T \mathbf{t}_m)(\mathbf{r}'_{m,i}^T \mathbf{R}_m \mathbf{r}_{m,i}) - (\mathbf{r}_{m,i}^T \mathbf{r}_{m,i})(\mathbf{r}'_{m,i}^T \mathbf{R}_m^T \mathbf{t}_m). \quad (28)$$

3.7 低精度の計測点の除去 計測精度は一般にカメラの移動方向の延長線付近ほど悪く, カメラから遠いものほど悪い. このため計測結果は精度の良いものと悪いものが存在することになる. こ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 (29) 式で誤差を評価するベクトル $\mathbf{g}$ を求める. そしてベクトル $\mathbf{g}$ のノルムが (30) 式を満たす計測結果を以後の処理に用いる.

$$\mathbf{g} = \left| \frac{\partial \mathbf{p}_{m,i}}{\partial u_{m,i}} \right| + \left| \frac{\partial \mathbf{p}_{m,i}}{\partial v_{m,i}} \right| + \left| \frac{\partial \mathbf{p}_{m,i}}{\partial u'_{m,i}} \right| + \left| \frac{\partial \mathbf{p}_{m,i}}{\partial v'_{m,i}} \right|, \quad (29)$$

$$\|\mathbf{g}\| < h, \quad (30)$$

ただし, h は誤差評価のための閾値である.

3.8 計測結果の合成 上記により動画像中の 2 画像における計測結果が得られ, 同時にカメラ間の位置関係も得られる. そこで次の計測には前回の計測で使った最終画像を初期画像として以後の数フレームを用いて同様な計測を行う. これにより, 得られたカメラ間の位置関係を用いて計測結果を合成できる.

しかし個々の計測結果はカメラ間の距離を 1 としており, 実際の観測点間の距離が異なる計測結果をそのまま合成すると図 3 の点線で示す今回の計測結果が実線で示す前回の計測結果と一致しない. そこで, 2 つの計測で共通して用いた画像の観測点 $\mathbf{c}$ を中心として, 同じ特徴点における前回の計測結果 $\mathbf{P}_{m,i} = [x_{m,i}, y_{m,i}, z_{m,i}]^T$ と今回の計測結果 $\mathbf{P}_{m+1,i} = [x_{m+1,i}, y_{m+1,i}, z_{m+1,i}]^T$ が一致するようにスケールを合わせる.

ここで, $\mathbf{P}_{m,i}$ と $\mathbf{P}_{m+1,i}$ の座標間の距離を最小にするようなスケールにすると $\mathbf{c}$ から遠い計測結果の影響を

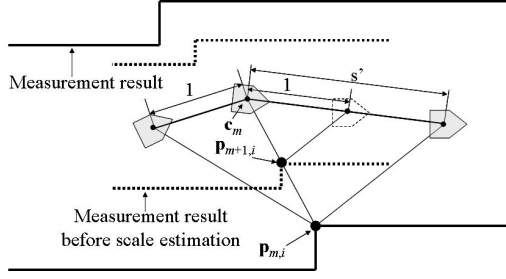


Fig. 3 Scale mismatching.

大きく受けるため、計測結果の各成分に対数をとったものの2乗距離が最小となるようにスケールを合わせる．すなわちスケール s' は (31) 式から求める．

$$\min_{s'} \sum_{i=1}^n \|\log(\mathbf{p}_{m,i} - \mathbf{c}) - \log(s' \mathbf{p}_{m+1,i} - \mathbf{c})\|. \quad (31)$$

3.9 モデリング 上記の処理を繰り返して得られた計測点群を用いてモデリングを行う．

具体的には、点群から各点を結ぶ3角形（3角パッチ）を構成し、テクスチャを貼り付ける平面を作成する．まず、3次元のドロネー分割を用いて点群から凸形状の3角錐群を生成する．ただし、カメラ（観測点）から観測される計測点（可視点）は3角錐の内部に存在し、表面に現れない可能性がある．そこで、観測点と可視点を結ぶ線分の間には物体が存在しないことを利用し、観測点と可視点を結ぶ線分が3角錐と交差する場合、線分が交差した面を削除する．この処理により、3角パッチを生成する．

次に、得られた画像から最大の面積を持つテクスチャを各3角パッチに貼り付けることで、高解像度の3次元環境モデルを生成する．

4. 実験

4.1 実験環境 SONY 製ハイビジョンカメラ HDR-HC1 に末陰産業製双曲面ミラー SOIOS70-SCOPE を装着した全方位カメラ（図4(a)）を、Mobile Robots Inc. 製 Pioneer3 上に設置し（図4(b)）、2次元平面上で移動ロボットを移動させながら画像を取得した．また、人間が全方位カメラを把持して位置と姿勢を3次元的に変化させながら画像を取得した．いずれの場合も、画像サイズは 1920×1080 pixels である．

4.2 カメラキャリブレーション カメラとミラーを分離し、カメラ単体での内部パラメータのキャリブレーションを行った後、双曲面ミラーを装着して双曲面パラメータのキャリブレーションを行う．

カメラの内部パラメータのキャリブレーションは、円形パターンを既知の一定間隔で描画した平面板を複数視点から撮影した画像を用いて求める⁽²⁰⁾．



(a) Omni-directional camera. (b) Mobile robot.

Fig. 4 Experimental system.

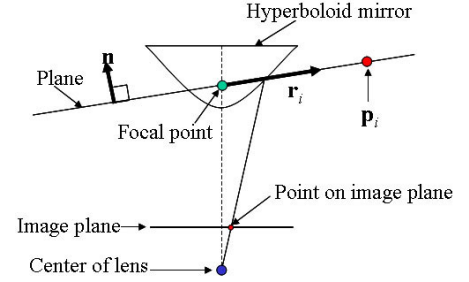


Fig. 5 Omni-directional camera calibration.

次に、カメラに双曲面ミラーを装着してキャリブレーションを行う．画像上の点から得られた光線ベクトル $\mathbf{r}$ は双曲面の形状に依存し、双曲面の形状は (1) 式の a , b により決定する．本手法では $\mathbf{r}$ の方向のみを計測に用いるため、 $\mathbf{r}$ の大きさに関わる s を無視することができ、 $\mathbf{r}$ の方向は b/a の関数として表される．よって、双曲面パラメータを b/a として推定する．

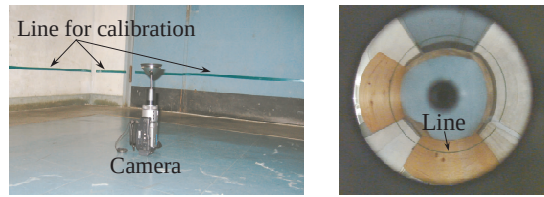
まず、ミラー側の双曲面の焦点を含む平面上に複数のマークをつけた環境で全方位画像を撮影する．そして画像中のマークから得られる光線ベクトル $\mathbf{r}$ が全て平面上のベクトルとなるように双曲面パラメータを求める（図5）．

任意の平面の法線ベクトルを $\mathbf{n} = [n_x, n_y, 1]^T$ とし、複数のマークにおける光線ベクトル $\mathbf{r}_i$ とすると、 b/a が (32) 式となる時光線ベクトル $\mathbf{r}_i$ は最も平面をなしていると考えられる．これを満たすような双曲面パラメータおよび法線ベクトルを Levenberg-Marquardt 法⁽²¹⁾を用いて推定する．

$$\frac{b}{a} = \arg \min_{b/a, \mathbf{n}} \sum_i \mathbf{n}^T \mathbf{r}_i. \quad (32)$$

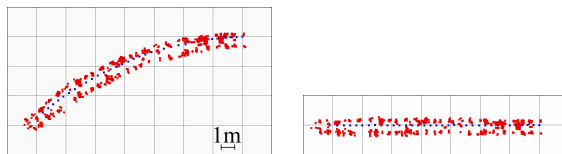
双曲面パラメータ b/a と共に法線ベクトル $\mathbf{n}$ を推定することで、撮影環境で設定した平面に対してカメラが多少傾いても双曲面パラメータを推定することが可能である．

4.3 キャリブレーション結果 双曲面の焦点と同じ高さに線を引いた環境で全方位画像を撮影し



(a) Line for calibration. (b) Acquired image.

Fig. 6 Omni-directional camera calibration.



(a) Without calibration. (b) With calibration.

Fig. 7 Result of camera calibration.

(図 6(a)), 取得画像 (図 6(b)) の線上の点を用いて双曲面パラメータ b/a を推定した。ただし、カメラのレンズ中心は双曲面のミラー側でない焦点と一致していると仮定し、全方位画像の中心は手で与えた。

双曲面ミラー購入時の仕様書に書かれたカタログ値の双曲面パラメータ b/a は 1.576 であったことに対して、キャリブレーションを行った結果 $b/a = 1.663$ と推定された。直線廊下の形状を提案手法で計測した結果、カタログ値を用いた場合には実際の形状と大きく異なった歪んだ結果になったことに対して (図 7(a)), キャリブレーションで推定した b/a を用いた場合にはほぼ直線形状で正しく計測できた (図 7(b))。

以上の結果より、提案した双曲面パラメータのキャリブレーション手法が有効であることが確認できた。

4.4 2次元移動の結果 カメラの移動が2次元の実験を行った。L字に曲がった廊下においてロボットを約 40cm/s の速度で移動させ、移動中に全方位カメラで撮影した 15fps の動画を計測処理の入力とした。個々の計測の観測点は 20frame ごとに設定した。これにより基線長は約 530mm となる。

全方位画像に抽出された特徴点を表示した例を図 8 に示す。◇は RANSAC により除去された点、△は精度が悪いと判定された点、○は結果の合成に使用し計測結果を表示した点を示す。初期画像における特徴点の抽出数は試行錯誤的に $l = 200$ 個と決定した。また、RANSAC の閾値は $q = 0.01$ とし、繰り返し回数を $w = 5000$ とした。計測精度の良い点を判別する閾値としては、 $h = 0.25$ を用いた。更に、○以外で示した特徴点を計測終了後に除去して次回の計測には画像全体から再抽出した。

処理時間は 2.53GHz の CPU で、特徴点の追跡が 1

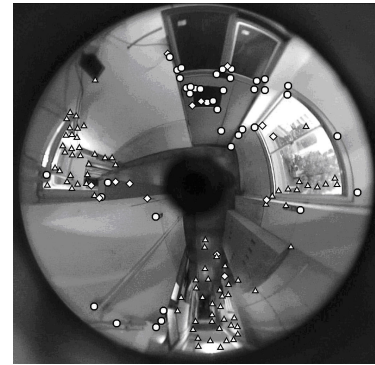
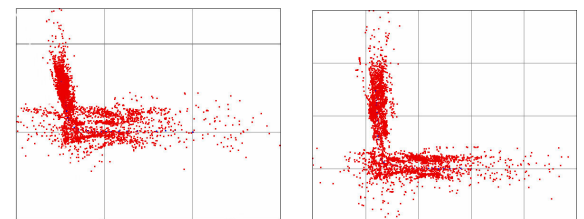
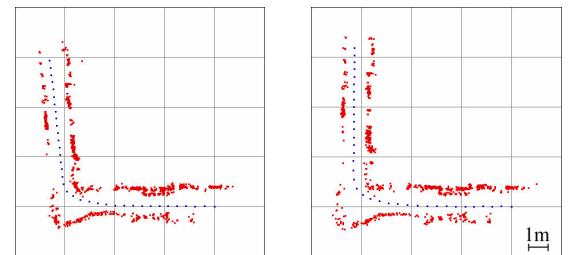


Fig. 8 Feature points in an omni-directional image.



(a) Result 1.

(b) Result 2.



(c) Result 3.

(d) Result 4.

Fig. 9 Outlier and low accuracy points rejection.

枚の画像ごとに約 2.0s であり、対応点が得られてからの計測時間は約 0.6s であった。

周囲環境を計測した結果を上から見た図を図 9 に示す。(a) は外れ値の除去及び精度の悪い結果の除去を行わずに計測した結果、(b) は外れ値除去のみを行った結果、(c) は低精度除去のみを行った結果、(d) は外れ値除去と低精度除去の両方を行った結果である。L 字型廊下状に分布している点は特徴点の計測結果を示し、廊下中央に点在する点は計測に用いたカメラの位置の推定値を示す。

(a) では外れ値により観測点の位置関係の推定に誤差を生じ、廊下の角が直角に計測されていない。また誤差の大きい点を計測しているため、計測結果がばらついている。これに対して (b) では外れ値除去により観測点の位置関係の推定精度が向上したため廊下の角は直角に計測されているが、ばらつきは大きい。(c) では低精度の除去により計測精度の良い点のみを計測しているため計測結果は廊下壁面にまとまっているが、

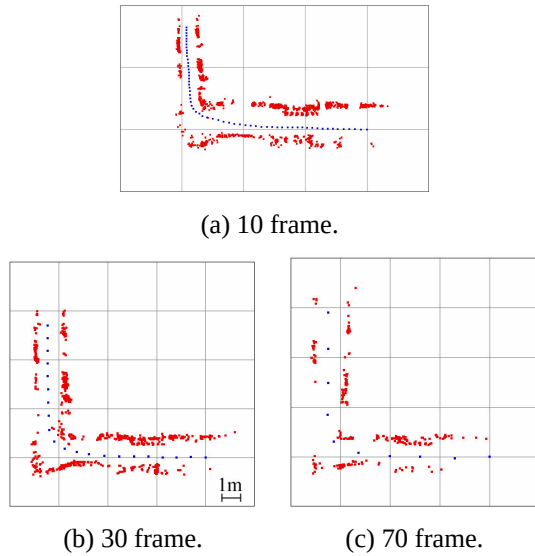


Fig. 10 Influence of baseline length.

廊下の角が直角に計測されていない。(d)では外れ値除去により観測点の位置関係の推定精度向上し、低精度除去により計測精度が向上しており、直角に曲がった廊下の環境を正しく計測されている。

次に、観測点間の距離である基線長を変化させ、基線長別の計測結果を作成した。入力する動画は図9と同じものを使用し、基線長の変化には1回の計測に用いるフレーム数を変化させることによって行った。10frame毎、30frame毎、70frame毎の計測結果を図10に示す。(a)の10frame毎の計測では廊下を曲がる前と後で大きくスケールが変化している。これは基線長が短いため計測結果の誤差が大きく、これを用いたスケール合わせに大きく影響しているためであると考えられる。(b)の30frame毎の計測ではスケールの変化がほとんど無く、また計測点の数も十分得られ、環境を正しく計測されている良好な結果であると考えられる。(c)の70frame毎の計測では計測点の密度が荒い結果となった。これは計測回数が少ないことに加え、対応点抽出時における特徴点の追跡で多くの特徴点が追跡できずに失われ、対応点の数が減少したためである。

最後に、ロボットのデッドレコニングの値を観測点間の位置姿勢として用いた場合と、本手法の画像のみから推定した観測点間の位置姿勢を用いた場合の計測結果を図11に示す。デッドレコニングと本手法の入力には同じ動画を使用した。(a)にデッドレコニングによる計測結果を、(b)に本手法による計測結果を示す。(a)ではロボットの車輪に偏りがあるため、計測結果が曲がっている。それに対し(b)の本手法による計測では、直角に曲がった廊下の形状が正しく計測された結果が得られた。定量的には、(b)では廊下の長

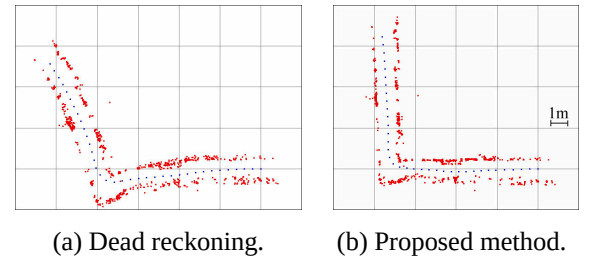


Fig. 11 Dead reckoning and proposed method.

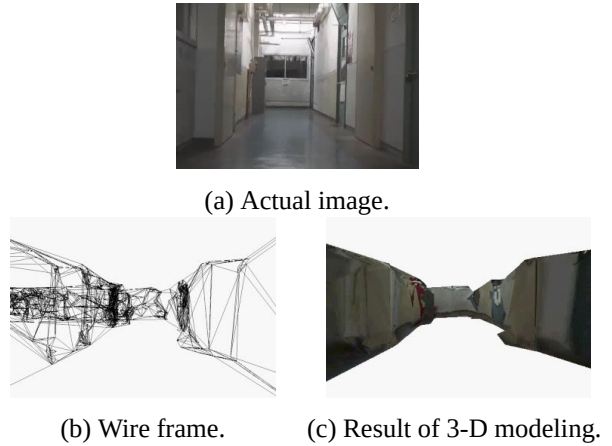


Fig. 12 Result of 3-D modeling.

さの計測誤差は平均2.6%であった。これにより、本手法はデッドレコニングと比べて十分精度の良い位置姿勢の推定が可能であることが分かる。

また、3次元環境モデリングの結果を図12に示す。(a)は通常のカメラで廊下を撮影した結果、(b)は計測結果から作成した環境モデルの3角パッチをワイヤフレームで表示した結果、(c)はこれにテクスチャを貼り付けた結果である。3次元モデリング結果と通常のカメラで撮影した画像と比較すると、物体の位置関係が一致しているため、実環境を表す環境モデルが適切に作成できていると考えられる。

4.5 3次元移動の結果 カメラを3次元的に移動させた実験を行った。人間が全方位カメラを把持して、位置・姿勢を変化させながら4角い部屋の中を移動させ取得した画像を用いた結果を図13に示す。(a)は鳥瞰図3次元計測結果であり、矢印の始点がカメラの位置、矢印の方向がカメラの姿勢を示す。(b)はテクスチャを貼り付けた結果である。部屋の形状及びカメラの位置・姿勢が正しく推定されており、3次元的に位置・姿勢が変化した場合にも適切な結果が出ていることが分かる。

5. 結 論

双曲面ミラーによる全方位カメラの移動中の動画をを用いてSFMにより周囲環境を計測し、複数の結果

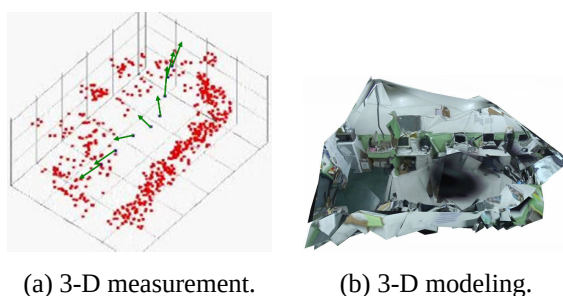


Fig. 13 Result of 3-D movement.

を合成して広範囲な3次元計測結果を得ると同時に、環境モデルを生成する手法を提案した。形状情報とテクスチャ情報の両方を持つ環境モデルは、ロボットの自己位置推定や人間への情報提示など様々な用途で用いることができる。カメラの画像のみを入力情報としていることによりシステム構成が単純であり、姿勢変動のあるカメラの移動にも対応できる点も提案手法の利点である。実験では多数の計測結果の合成に対して誤差の累積が小さいという結果が得られ、本手法の有効性が示された。

今後の課題としては、対応点の検出精度の向上や、3箇所以上の観測点の幾何学的関係により複数の観測点の位置関係を同時に推定することなどが挙げられる。

謝 辞 本研究の一部は、文部科学省大都市大震災軽減化特別プロジェクトの一環として、国際レスキューシステム研究機構、防災科研との協力の下になされた。

文 献

- (1) S. Thrun: "Robotic Mapping: A Survey," *Technical Report CMU-CS-02-111*, Carnegie Mellon University (2002)
- (2) J. Gaspar, N. Winters and J. Santos-Victor: "Vision-Based Navigation and Environmental Representations with an Omnidirectional Camera," *IEEE Transactions on Robotics and Automation*, Vol.16, No.6, pp.890–898 (2000)
- (3) Y. Yagi: "Omnidirectional Sensing and Its Applications," *IEICE Transactions on Information and Systems*, Vol.E82-D, No.3, pp.568–579 (1999)
- (4) H. Ishiguro, M. Yamamoto and S. Tsuji: "Omnidirectional Stereo," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.14, No.2, pp.257–262 (1992)
- (5) J. Takiguchi, M. Yoshida, A. Takeya, J. Eino and T. Hashizume: "High Precision Range Estimation from an Omnidirectional Stereo System," *Proceedings of the 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.263–268 (2002)
- (6) H. Koyasu, J. Miura and Y. Shirai: "Mobile Robot Navigation in Dynamic Environments Using Omnidirectional Stereo," *Proceedings of the 2003 IEEE International Conference on Robotics and Automation*, pp.893–898 (2003)
- (7) J. Meguro, J. Takiguchi, Y. Amano and T. Hashizume: "Omni-directional Motion Stereo Vision based on Accurate GPS/INS Navigation System," *Proceedings of the 2nd Workshop on Integration of Vision and Inertial Sensors* (2005)
- (8) K. Yamazaki, M. Tomono, T. Tsubouchi and S. Yuta: "Object Shape Reconstruction and Pose Estimation by a Camera Mounted on a Mobile Robot," *Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.4019–4025 (2004)
- (9) 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)
- (10) O. D. Faugeras: "What can be seen in three dimensions with an uncalibrated stereo rig?," *Proceedings of European Conference on Computer Vision*, pp.563–578 (1992)
- (11) R. Hartley, R. Gupta and T. Chang: "Stereo from Uncalibrated Cameras," *Proceedings of the 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.761–764 (1992)
- (12) J. Gluckman and S. K. Nayar: "Ego-Motion and Omnidirectional Cameras," *Proceedings of 6th International Conference on Computer Vision*, pp.999–1005 (1998)
- (13) P. Chang and M. Hebert: "Omni-directional Structure from Motion," *Proceedings of the 2000 IEEE Workshop on Omnidirectional Vision*, pp.127–133 (2000)
- (14) M. Oe, T. Sato and N. Yokoya: "Estimating Camera Position and Posture by Using Feature Landmark Database," *Proceedings 14th Scandinavian Conference on Image Analysis*, pp.171–181 (2005)
- (15) G. Xu and Z. Zhang: *Epipolar Geometry in Stereo, Motion and Object Recognition: A Unified Approach*, Kluwer Academic Publishers (1996)
- (16) C. Tomasi and T. Kanade: "Detection and Tracking of Point Features," *Technical Report CMU-CS-91-132*, Carnegie Mellon University (1991)
- (17) 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)
- (18) K. Yamazawa, Y. Yagi and M. Yachida: "HyperOmni Vision: Visual Navigation with an Omnidirectional Image Sensor," *Systems and Computers in Japan*, Vol.28, No.4, pp.36–47 (1997)
- (19) 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, pp.381–395 (1981)
- (20) R. Y. Tsai: "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses," *IEEE Journal of Robotics and Automation*, Vol.RA-3, No.4, pp.323–344 (1987)
- (21) W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling: *Numerical Recipes in C*, Cambridge University Press (1988)