

# 輝度拘束に基づく校正指標を用いないステレオカメラ校正

若井信彦<sup>1,a)</sup> 登一生<sup>1</sup> 佐藤智<sup>1</sup> 吾妻健夫<sup>1</sup>

**概要:** 車載やロボット等の移動体向けステレオカメラの校正を専用の校正指標を用いず実施するため、対応点探索を用いない一般風景からステレオカメラを校正する手法を提案する。提案法は、カメラパラメータが正しい場合、被写体の任意の3次元点をステレオ画像に投影した画素値が一致する輝度拘束を用いることで、ステレオカメラ間の対応点を求めることなく、直接、カメラパラメータを推定する。提案法はまず、初期カメラパラメータを用いて投影した画素間の画素値差を評価する。次に、反復計算で画素値差を最小化することで、カメラパラメータを更新する。シミュレーション実験によって、提案法は、駐車場の一般風景から校正でき、さらに、典型的な校正条件において Tsai の手法と比べ、根 2 乗平均再投影誤差に関して 42%高精度であることを示す。

**キーワード:** カメラ校正, ステレオカメラ

## 1. はじめに

ステレオカメラは3次元測距センサとして、自動運転車、先進運転支援システム、および、ロボットに広く商業利用されている。ステレオカメラによる被写体の3次元座標の取得には、カメラの幾何を定義する内部、外部カメラパラメータが必要である。これらのカメラパラメータを推定することをカメラ校正と呼び、通常、工場出荷時に実施される。そのため、経年変化や振動、衝撃等の影響によってカメラパラメータが変化した場合、修理工場等にカメラを輸送して再校正する必要がある、稼働率が低下することが課題であった。

カメラ校正法としては多くのものが提案されており、校正指標を使用する手法[1, 2]と、使用しない手法[3-11]に大別される。

校正指標を使用する典型的な手法は Tsai の手法である[1]。この手法は世界座標の3次元座標と画像上の2次元座標の組を使用し、世界座標をカメラパラメータで画像上に投影した座標と対応する画像座標の距離(再投影誤差)の総和を最小化することによりカメラを校正する。この手法は誤差の少ない前述の3次元点と2次元点の組を利用することで、高い校正精度が得られる。

一方、校正指標を用いない手法として、フレーム間の対応点を使用する方法[3]、ステレオ画像を使用する手法[4-8]、または、そのどちらも使用する手法[9, 10]などが提案されている。Maybank ら[3]は、エピポーラ幾何を表す2つの画像間の内部パラメータの基礎行列に関する Kruppa 方程式を利用した先駆的な自己校正法を提案した。Xu ら[4]は、自動車のステレオカメラを校正する手法を提案した。この手法は横断歩道の角のステレオ画像対応点を検出し、これらの対応点を用いて外部パラメータを推定する。Yu ら[9]は、ステレオカメラを構成する2つのカメラが同じ(内部パラメータの設計値が等しい)場合に拘束条件を追加することで、移動するステレオカメラのフレーム間の対応付けを用いたカメラパラメータの推定精度を向上さ

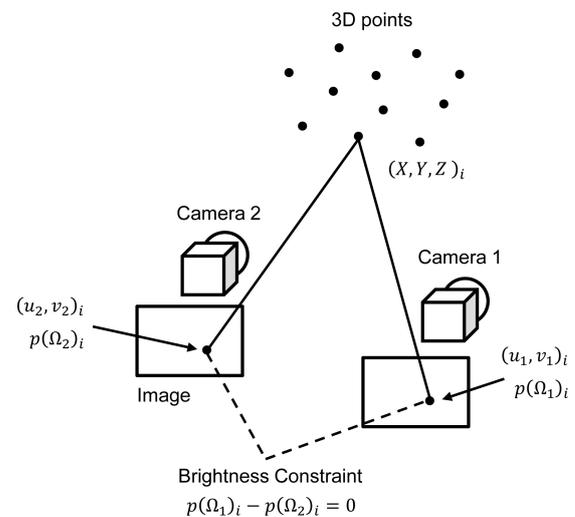


Figure 1. Conceptual image of brightness constraint of a stereo camera.

せる。これらの手法は校正指標を必要としないが対応点探索に基づく誤差の影響を受けるため、校正指標を用いる手法より精度が低い。

本稿では、校正指標を用いず、校正指標を用いる手法以上の精度が得られる校正法を提案する。提案法では、同一被写体をステレオカメラで撮像した場合、ランバート反射を仮定すると、オクルージョンでない領域の画素値がステレオ画像で一致することを表す拘束(輝度拘束)を用いる。提案法は誤差の原因である対応点探索を行わず、レーザーレンジファインダで取得した3次元点をステレオ画像に投影した画素値の差である輝度拘束を評価関数として最小化することでカメラパラメータを推定するため、高精度なカメラ校正が可能である。

提案法は、明確な特徴がなく対応点探索が困難なテクスチャに乏しい画素も評価に利用できる。これにより、提案法はマーカーや対応点探索を用いる従来手法に比べ 10 倍以上の校正点を使用できる。その結果、提案法は高精度なカメラパラメータを取得できる。

以下、輝度拘束を用いたステレオカメラ校正の原理説明を2章、一般風景を用いてステレオカメラを校正できることを3章、校正データにノイズを含む場合の提案法の精度評価を4章、結論を5章に述べる。

## 2. 輝度拘束を用いたステレオカメラ校正

本章では、提案法である輝度拘束を用いたカメラ校正法を説明する。輝度拘束とは、同一被写体をステレオカメラで撮像した場合、ランバート反射を仮定すると、オクルージョンでない領域の画素値はステレオ画像で一致する拘束である。本方式は最も輝度拘束を満たすカメラパラメータを推定する。

### 2.1 ステレオカメラの輝度拘束

本節では、ステレオカメラの輝度拘束について述べる。図1は提案法のステレオカメラにおける輝度拘束の概念図である。図1において、 $\Omega_1, \Omega_2$ はカメラ1とカメラ2のカメラパラメータであり、 $i$ 番目の世界座標点を $\Omega_1, \Omega_2$ でステレオカメラの各画像に投影した点を $(u_1, v_1)_i, (u_2, v_2)_i$ とする。 $p(\Omega_1)_i, p(\Omega_2)_i$ はそれぞれカメラ1とカメラ2の投影した画像座標の画素値である。ここで、被写体表面上にランバート反射を仮定すると、式(1)が成り立つ。

$$p(\Omega_1)_i - p(\Omega_2)_i = 0, \quad (1)$$

これを輝度拘束式と呼ぶ。カメラパラメータが真値から離れるに従い、 $p(\Omega_1)_i$ と $p(\Omega_2)_i$ の輝度差は増加する。したがって、この輝度差を評価値とし、最小化することによりカメラパラメータを得る。式(2)に、 $p(\Omega_1)_i$ と $p(\Omega_2)_i$ の輝度差の根2乗平均(RMS)である評価値を示す。

$$J(\Omega_1, \Omega_2) = \sqrt{\frac{1}{N} \sum_{i=1}^N (p(\Omega_1)_i - p(\Omega_2)_i)^2}, \quad (2)$$

ここで、 $J$ は評価値、 $N$ は3次元点の点数である。

### 2.2 輝度拘束を用いたカメラパラメータ推定

本節では、提案法である輝度拘束を用いたステレオカメラのカメラパラメータ推定について説明する。図2は、提案法のカメラ校正のフローチャートである。まず、ステレオ画像、初期カメラパラメータ、および、被写体の3次元点座標を入力する。次に、式(2)で示した評価関数を最小化するカメラパラメータを、階層型グリッドサーチで探索する。この階層型グリッドサーチは1つのカメラパラメータ（焦点距離等）ごとに、収束するまでグリッドサーチを順に反復する手法であり、この手法の処理フローを図3に示す。

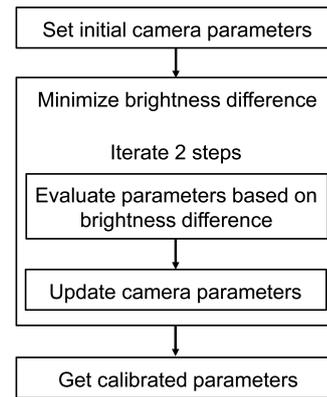


Figure 2. Camera calibration process flow.

提案法は任意のカメラに適用できるが、ここでは、単純なカメラモデルであり、本実験で扱うピンホールカメラモデルを利用する。

$$h \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f/d_u & 0 & -C_u & 0 \\ 0 & f/d_v & -C_v & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_1 & r_2 & r_3 & T_x \\ r_4 & r_5 & r_6 & T_y \\ r_7 & r_8 & r_9 & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}, \quad (3)$$

ここで、 $(X, Y, Z)$ は3次元点の世界座標、 $(u, v)$ は3次元点を投影した2次元画像座標、 $f$ は焦点距離、 $d_u, d_v$ はそれぞれイメージセンサ1画素の $u, v$ 方向の長さ、 $C_u, C_v$ は画像中心の $u, v$ 座標、 $r_1-r_9$ はカメラ姿勢の回転を表す行列、 $T_x, T_y, T_z$ は世界座標系からカメラ位置を基準とする3次元座標系に変換する並進の $x, y, z$ 成分、 $h$ は自由度のない媒介変数である。3変数( $f, d_u, d_v$ )については、自由度が2しかないため、 $d_v$ を任意の値で固定する。したがって、提案法は10個のパラメータ( $f, d_u, C_u, C_v, T_x, T_y, T_z, \theta$ : ロール角,  $\phi$ : ピッチ角,  $\varphi$ : ヨー角)を推定する問題に帰着する。

## 3. 実験 I: 原理実証

この章では、一般風景を用いてステレオカメラを校正できることを計算機シミュレーションで示す。すなわち、評価値を最小化することで、真値と異なる初期カメラパラメータから真値に十分近いカメラパラメータ (RMS 再投影誤差 0.1 pixel 以下) を推定できることを示す。

### 3.1 実験条件

実験に用いた駐車場風景の3次元データを図4に示す。このシーンではオクルージョンは生じない。また、2つのカメラは光軸平行に水平方向に15 mm 離して配置した。

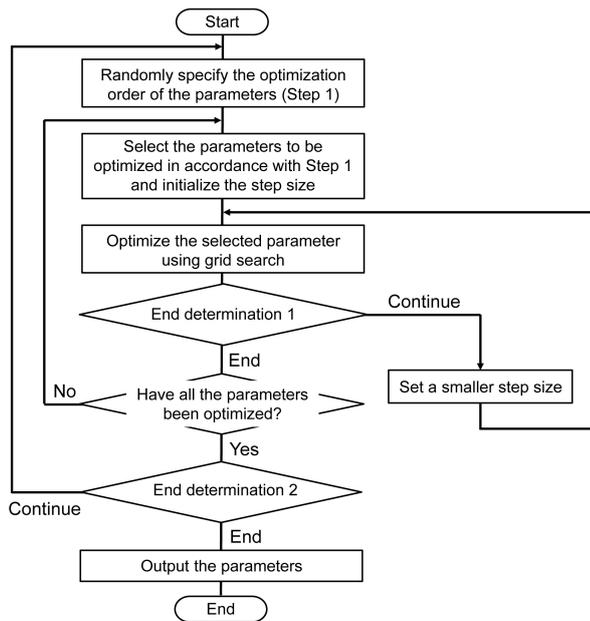


Figure 3. A flow of hierarchical grid search process.

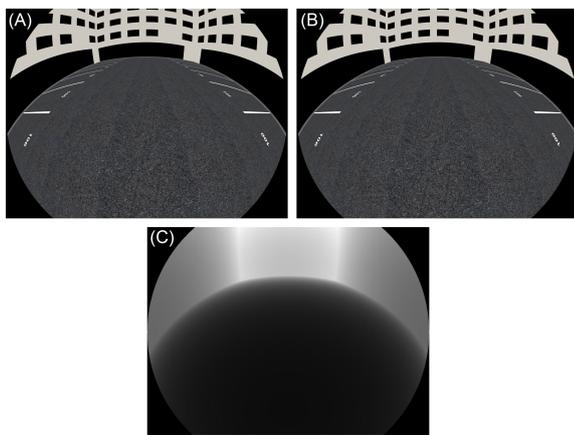


Figure 4. A parking lot scene used in this experiment. (A) Camera 1 (256-gradation RGB, 1280 × 960 pixel) (B) Camera 2. (C) A depth map of Camera 1.

カメラパラメータの真値を表 1 に示す。  $d_v$  の値は設計値 (0.00275 mm) を用いた。

提案法はステレオカメラのカメラ 1 とカメラ 2 を同時に校正できるが、実験においてカメラ 1 は真値で固定して校正せず、カメラ 2 のみを校正する。このように片方のカメラのみ校正することは経年変化やカメラへの接触時に必要とされる。

初期カメラパラメータは真値に正規分布のノイズを付加して生成した。正規分布の平均は 0 とし、標準偏差は内部パラメータ ( $f, d_u, C_u, C_v$ ) で真値の 0.5%、外部パラメータの各回転 ( $\theta, \phi, \psi$ ) で  $1^\circ$ 、各並進 ( $T_x, T_y, T_z$ ) で 5 mm とした。これらの標準偏差の大きさは組み立て誤差を考慮して決定

Table 1. True values of camera parameters

Parameter	True value
$f$	1.12 [mm]
$d_x$	0.00275 [mm]
$C_u$	640 [pixel]
$C_v$	480 [pixel]
$\theta, \phi, \psi$	0 [deg]
$T_x, T_y, T_z$	0 [mm]

Table 2. Criteria for hierarchical grid search

	Criterion
Number of searches	20
Initial search ranges	2 <sup>nd</sup> significant digit
End determination 1	Step size is smaller than the 6 <sup>th</sup> significant digit.
End determination 2	Number of iterations is 500.

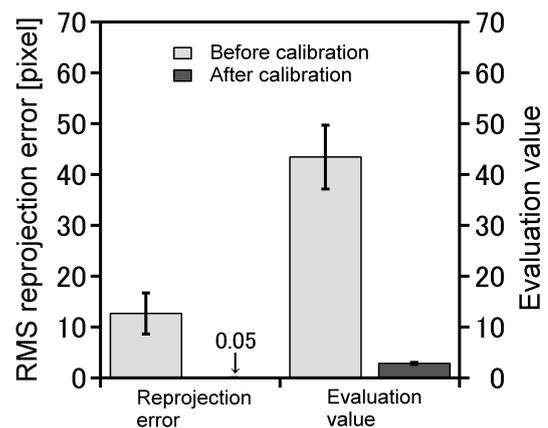


Figure 5. RMS reprojection errors and evaluation errors before and after camera calibration.

した。この初期カメラパラメータを、表 2 で示す階層型グリッドサーチの閾値を用いて評価値を最小化した。

### 3.2 実験結果

本節では、階層型グリッドサーチにより評価値が収束した時に、推定したカメラパラメータが真値に収束したことを示す。推定したカメラパラメータの真値との誤差は RMS 再投影誤差で表すことができる。

10,000 点の 3 次元点を用いて 100 回試行した校正の前後における RMS 再投影誤差と評価値を図 5 に示す。この 100 回の試行は、正規分布のノイズを用いた 100 個の初期カメラパラメータに対して、それぞれのカメラパラメータを推定することを意味する。

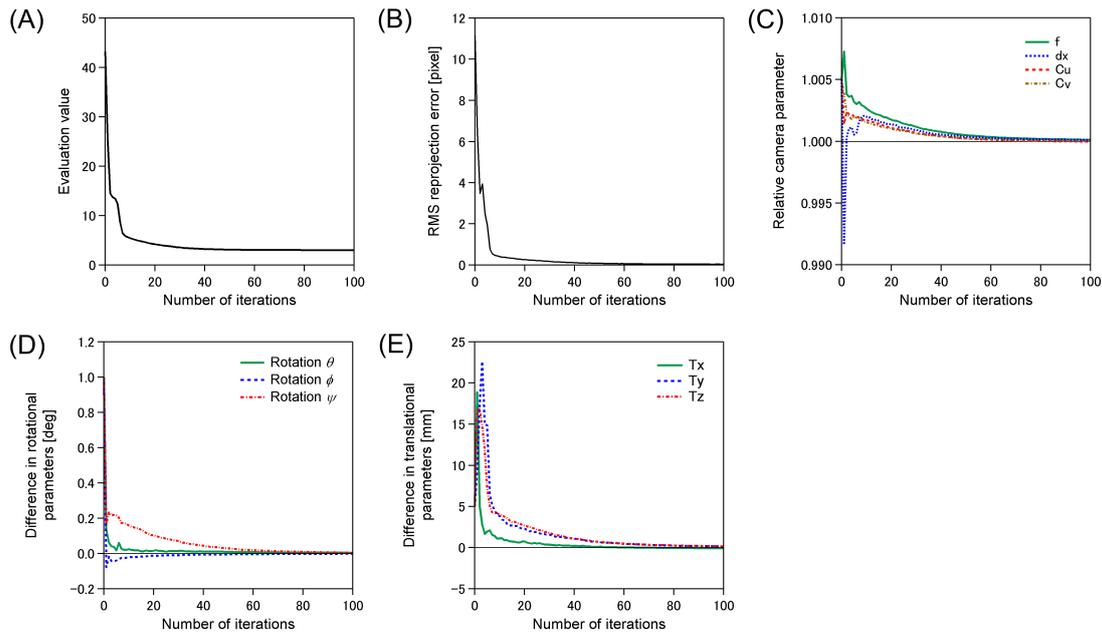


Figure 6. Convergence characteristics using this method. (A) Evaluation value (B) RMS reprojection error (C) Ratio of intrinsic camera parameters to true values (Estimation values / true values) (D) Difference between rotational camera parameters and true values (Estimation values - true values) (E) Differences between translational parameters and true values (Estimation values - true values).

校正後の RMS 再投影誤差（全校正点における RMS）と評価値はそれぞれ、 $0.05 \pm 0.01$  pixel (100 回の RMS 計算における Mean  $\pm$  S.D.)と  $3.03 \pm 0.00$  となった。この RMS 再投影誤差 0.05 pixel は 1 画素より十分小さく、カメラを校正できたと思わせる。また、評価値は階調誤差を表しており、収束時に得られた 3.03 は入力画像の階調 (256 階調) の ~1% であり、評価値は十分最小化されたと思わせる。これらの結果より、輝度拘束に基づく評価値を最小化することで、カメラパラメータに正規分布のノイズがあるカメラを校正できることが分かる。

提案法の収束特性を図 6 に示す。初期評価値 43.0 は、100 回の反復計算により 3.0 に収束した (図 6A)。同様に、RMS 再投影誤差は初期値 14.59 から 0.05 pixel に収束した (図 6B)。カメラパラメータの各要素も収束した (図 6C, D, and E)。この反復計算の時間はおおよそ 3 次元点数に比例し、実行時間は 10,000 点の 3 次元点に対する 1 回の反復計算あたり 11 秒であった。但し、Intel Xeon E5-1620 3.5 GHz PC を用いた。

これらの結果より、輝度拘束を最小化する提案法は、真値から一定の範囲内の初期カメラパラメータに対し、ステレオカメラを校正できることを示した。

#### 4. 実験 II: 精度評価

本章では校正データにノイズを含む場合の提案法の

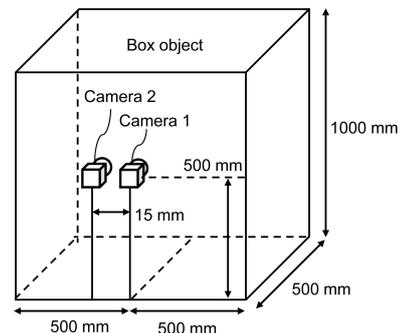


Figure 7. Positional relationship between the subject and the cameras.



Figure 8. Image used as subject texture (256-gradation RGB,  $1024 \times 1024$  pixel).

精度を示す。比較対象として広く利用されている Tsai の手法を選択した。また、Tsai の手法で校正することで、3 次

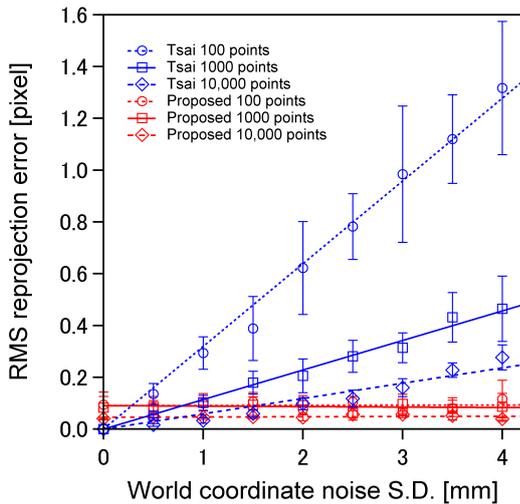


Figure 9. Comparison of RMS reprojection errors when 3D coordinates include noise.

元測距センサとして利用可能な精度が得られるので、Tsai の手法は校正精度の基準となる手法の一つとして考えられる。

校正データとして、提案法では、世界座標系の 3 次元点とステレオ画像を用いる。一方、Tsai の手法では世界座標系の 3 次元点と画像上の 2 次元座標の組を用いる。また、正規分布のノイズを校正データに加え、RMS 再投影誤差で評価した。

#### 4.1 実験条件

実験に用いた被写体とカメラの配置を図 7 に示す。提案法と Tsai の手法では取り扱うデータが異なるが、同一の条件で比較するために、箱状の被写体を共通に使用した。箱の大きさは  $1000 \times 500 \times 500$  mm であり、カメラ 1 を正方形の面の中心に配置し、カメラ 2 をカメラ 1 の左 15 mm に配置した。

ノイズ条件ごとに 10 回試行し、各試行において、入射角  $70^\circ$  以内の 100–10,000 点の校正点を箱の表面からランダムに選択し、校正点とした。これらの世界座標の 3 次元点と画像上の 2 次元座標を Tsai の手法に用いた。また、提案法では、Tsai の手法に使用した 3 次元点とカメラ 1 とカメラ 2 の画像を用いた。カメラ 1 とカメラ 2 の画像は内側にテクスチャ (図 8 [11]) を張った被写体の箱を撮影することにより取得した。

カメラ校正の精度を評価するために、入射角  $70^\circ$  以内の全画素 (764,882 点) に対する RMS 再投影誤差を用いた。カメラパラメータ推定時に、計算が不安定となる校正点を除くため、入射角  $70^\circ$  以内の画素を用いた。この計算の不安定性は、入射角  $90^\circ$  付近の校正点の世界座標を、カメラを基準とする 3 次元座標にカメラパラメータで変換した場合、変換後の校正点の 3 次元座標がカメラの後方となり、校正点の 3 次元座標をカメラパラメータで画像上に投影で

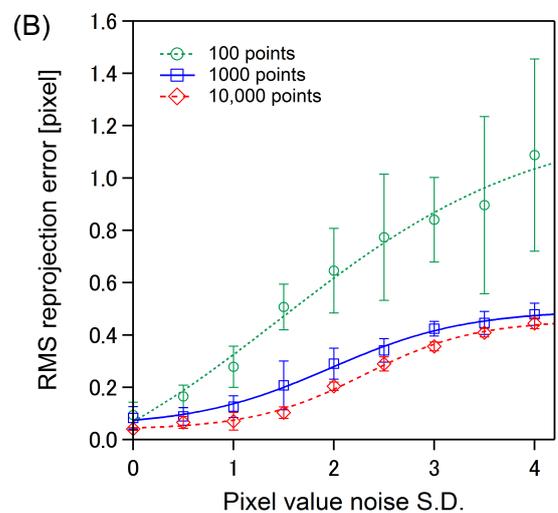
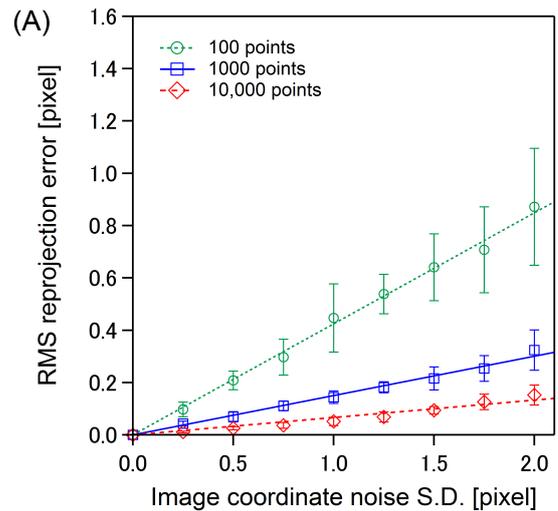


Figure 10. Calibration results of Tsai's method and our proposed method. (A) Calibration according to Tsai's method (Noise added to image coordinates) (B) Calibration according to our proposed method (Noise added to pixel values).

きないためである。

実験 II におけるカメラモデルと校正手順は実験 I と同じである。

#### 4.2 実験結果

**世界座標にノイズを含む場合の精度比較.** 世界座標に平均 0、標準偏差 0.0–4.0 mm の正規分布のノイズを加えた校正点を用いてカメラを校正した。この時、画像上の 2 次元座標と画像にはノイズを加えない。Tsai の手法と提案法の RMS 再投影誤差を図 9 に示す。

世界座標にノイズを含む場合、標準偏差 0.5 mm 以下を除き提案法の RMS 再投影誤差は Tsai の手法より小さい結果であった。10,000 点の校正点を使用した時、提案法の RMS 再投影誤差は 0.06 pixel 以下であった。一方、Tsai の手法の RMS 再投影誤差は、4.00 mm の世界座標のノイズに

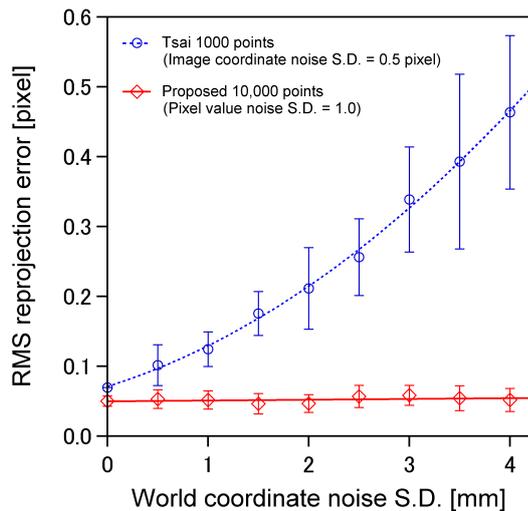


Figure 11. Accuracy comparison between Tsai's method and our proposed method under typical conditions.

対して 0.28 pixel であった。したがって、世界座標にノイズを含む条件において、提案法は Tsai の手法より校正精度が高い。

また、図 9 において、提案法の RMS 再投影誤差は、ノイズ増加した場合においてもおおよそ一定の値になっている。一方、Tsai の手法の RMS 再投影誤差はノイズの増加に伴い増加する。このことは、世界座標に正規分布のノイズを付加した場合、Tsai の手法の評価値の空間（画像座標）で誤差の偏りが生じ、一方、提案法の評価値の空間（画素値）では誤差の偏りが生じなかったためと考えられる。

**Tsai の手法における画像座標ノイズの影響。** 平均 0、標準偏差 0.0–2.0 pixel とする正規分布のノイズを画像座標に付加し、カメラを校正した。ここで、世界座標にノイズは付加していない。RMS 再投影誤差を図 10A に示す。画像座標にノイズを含む場合、10,000 点の 3 次元点を用いても、RMS 再投影誤差は 0 でなかった。これは、前述の世界座標にノイズを付加した場合と同様に、評価値の空間（画像座標）で誤差の偏りが生じたと考えられる。

**提案法における画素値ノイズの影響。** 平均 0、標準偏差 0.0–4.0 階調（256 階調画像）とする正規分布のノイズを画像に付加し、校正点の数を 100, 1,000, 10,000 と変化させ、カメラを校正した。校正点が 100 点の場合を除き、画素値ノイズ 4.0 階調以下の RMS 再投影誤差は 0.48 pixel 以下であった（図 10B）。校正点が 100 点の場合の RMS 再投影誤差は、校正点を 1000 点使う場合に対して 3 倍以上となった。この 100 点の校正点は全画素の 1% 以下であり、提案法に対して不十分なデータ数と考えられる。一方、10,000 点の校正点を用いた時、1.5 階調以下のノイズにおいて、RMS

再投影誤差は 0.13 pixel 以下となった。以上より、校正には、1000 点以上の校正点が必要である。

**典型的な条件における精度比較。** 提案法の実環境での有効性を確認するため、校正データにノイズを付加し、典型的な条件でカメラを校正した。校正条件を以下に示す。

**Tsai の手法：** 画像座標ノイズの標準偏差 0.5 pixel，校正点数 1000 点，世界座標ノイズの標準偏差 0.0–4.0 mm。

**提案法：** 画素値ノイズの標準偏差 1.0 階調，校正点数 10,000 点，世界座標ノイズの標準偏差 0.0–4.0 mm。

また、平均 0 の正規分布ノイズを用いた。提案法では、レーザースキャナ等で容易に多くの校正点を取得できるため、提案法では Tsai の手法より多くの校正点を使用した。

RMS 再投影誤差の比較を図 11 に示す。世界座標ノイズの標準偏差が 0.5 mm において、提案法は Tsai の方式に比べ RMS 再投影誤差に関して 42% 高精度であった。この結果より、典型的な条件において、提案法は Tsai の手法より校正精度が高い。これは、提案法がより多くの校正点を使用できるため、校正精度が高くなったと考えられる。

## 5. 結論

車載やロボット等の移動体向けステレオカメラの校正を専用の校正指標を用いず実施するため、校正指標を用いる手法以上の精度が得られる、校正指標を用いないステレオカメラ校正法を提案した。提案法は、対応点探索をせずにステレオ画像間の輝度拘束を用いてカメラパラメータを推定することで、一般風景からステレオカメラを校正できる。第一の実験により提案法が駐車場の一般風景からカメラを校正できることを示した。第二の実験により提案法は、典型的な条件において Tsai の手法より RMS 再投影誤差に関して 42% 高精度であることを示した。また、第二の実験において、3 次元点におけるノイズに対して頑健であることを示した。今後は、オクルージョン、鏡面、および、透過等に起因する輝度拘束を満たさない領域に対する手法の確立に取り組む。

## 参考文献

- [1] 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*, 3(4):323–344, 1987.
- [2] Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000.
- [3] S. J. Maybank and O. D. Faugeras. A Theory of Self-calibration of a Moving Camera. *International Journal of Computer Vision*, 8(2):123–151, 1992.

- [4] D. Xu, Q. Zeng, H. Zhao, C. Guo, K. Kidono, and Y. Kojima. Online stereovision calibration using on-road markings. *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pages 245–252, 2014.
- [5] Q. Yao, H. Sankoh, K. Nonaka, and S. Naito. Automatic camera self-calibration for immersive navigation of free viewpoint sports video. *2016 IEEE 18th International Workshop on Multimedia Signal Processing (MMSP)*, pages 1–6, 2016.
- [6] G. Xu, L. Chen, and F. Gao. Study on binocular stereo camera calibration method. *2011 International Conference on Image Analysis and Signal Processing*, pages 133–137, 2011.
- [7] Y. Ling and S. Shen. High-precision online markerless stereo extrinsic calibration. *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1771–1778, 2016.
- [8] Z. Zhang and Q. Tang. Camera Self-Calibration Based on Multiple View Images. *2016 Nicograph International (NicoInt)*, pages 88–91, 2016.
- [9] H. Yu and Y. Wang. An Improved Self-calibration Method for Active Stereo Camera. *2006 6th World Congress on Intelligent Control and Automation*, volume.1, pages 5186–5190, 2006.
- [10] T. Dang, C. Hoffmann, and C. Stiller. Continuous Stereo Self-Calibration by Camera Parameter Tracking. *IEEE Transactions on Image Processing*, volume 18(7), pages 1536–1550, 2009.
- [11] A. Ess and B. Leibe and K. Schindler and L. van Gool, A Mobile Vision System for Robust Multi-Person Tracking, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'08)*, 2008.