

広域人物検出・追跡システムのための 線形回帰木を用いたカメラキャリブレーション

大倉 有人¹ 和田 俊和¹

概要: 本報告では、人物検出システムのためのカメラキャリブレーション法について述べる。近年では、単独で動作する比較的安価な人物検出用の組み込みシステムが普及しており、これを複数台用いた広域監視システムの構築が求められている。画像上の人物位置を地図上の位置に変換し、その位置の近接性から、異なるシステムで検出された同一人物の判定を行うことでシステム間の連携を行うことができる。先行研究では座標の変換を単一の線形回帰木で行うことで、高速かつ省メモリな座標変換を実現していたが、人物検出システム設置時のキャリブレーションコストが非常に大きくなるという問題点があった。本報告では、線形回帰木で内部パラメータの補正を事前に行い、カメラ設置後に Homography を用いて外部パラメータの補正を行う。これら 2 つの変換を合成することで単一の線形回帰木を得ることができ、メモリ消費量、計算速度・精度等でのデメリットは生じない。

キーワード: カメラキャリブレーション, 線形回帰木, 分散協調システム

Camera calibration using linear-regression tree for wide-area surveillance system

YUTO OKURA¹ TOSHIKAZU WADA¹

Abstract: This report presents a camera calibration method for visual people detection devices. Recently, several embedded devices for visual people detection are commercially available. By combining those devices, a wide-area surveillance system can be constructed. In such systems, detected people by different devices have to be corresponded for tracking people across different surveillance areas. By transferring the people positions on the image to those on surveillance-area map, we can correspond proximal people on the map. In our previous research, we realized this transfer by linear-regression tree for fast and memory efficient computation. For constructing the tree, however, several point correspondences spread on the image and the map are necessary. This condition requires much labor. In this report, we first construct a linear-regression tree for removing lens-distortions and estimate the Homography between undistorted image to the map. By employing this strategy, we can reduce the number of corresponding points between the image and the map. Further, we can construct a compiled linear-regression tree representing the learned undistortion and the Homography. This compiled tree guarantees almost the same speed and the memory efficiency as the previous method.

Keywords: Camera calibration, Linear-regression-tree, Cooperative Distributed System

1. はじめに

近年、画像を用いた人物検出システムが普及しており、

これらのシステムはビルの入退室管理など、マーケティングやセキュリティの用途で広く用いられている。特に、デパートやチェーン店などででは、来客パターンの分析の需要が高く、このような人物検出システムはマーケティングの分野で積極的に用いられている。しかし、カメラ単体で

¹ 和歌山大学
Wakayama University

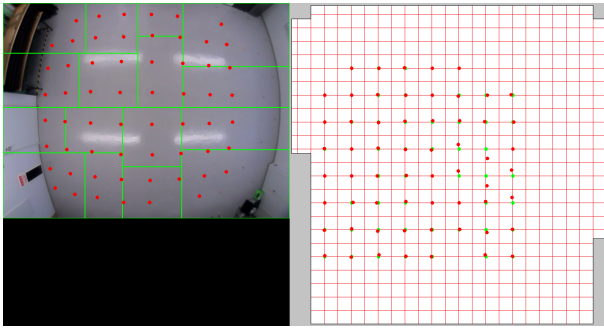


図 1 従来手法でのキャリブレーション

観測できる範囲には限界があり、カメラ 1 台のシステムでは広い店舗内全ての客の移動を把握することはできない。より広角のカメラで広範囲をカバーする視野を持たせるような人物検出システムの開発も行われているが、それにも限界がある。したがって、広い店舗全体を撮影するためには、複数台のカメラを協調動作させる必要がある。

複数台の人物検出システムを協調動作させるためには、異なるカメラで検出された同一人物を対応付ける必要がある。本研究では各人物検出システムで検出された画像中の人物位置を、単一の地図上の位置座標へと変換し、その上での近接性に基づいて、異なるカメラ間での同一人物の判定を行う。このとき、座標変換を行うアルゴリズムは追跡システムのスループットを落とさないように計算コストが低く、消費メモリの少ないアルゴリズムおよびデータ構造を採用すべきであるため、線形回帰木を用いて変換を行う。

このような手法は、後述する先行研究ですでに採用されていたが、カメラの画像から地図上の座標への変換を行う線形回帰木を構築する際に、カメラ画像に分布させた大量の点の座標と地図上の座標の対応付けを手動で行っていたため、システム設置時のキャリブレーションのコストが非常に大きくなるという問題点があった(図 1)。

本研究では、先にカメラごとにレンズ歪みの除去を行い透視投影モデルにあった投影像を作成しておき、カメラ設置後に歪みを除去した画像と地図との間での透視投影変換を求める方法をとる。こうすることで、カメラ設置時に人手で行う作業は透視投影変換の変換行列を求めるために必要な対応点を最低 4 点求めるだけになり、大幅な労力の削減につながる。

各カメラのキャリブレーションは、大型ディスプレイにグレイコードを表示することによって実現することもできるが、その代わりにチェッカーパターンを用いることができる。チェッカーパターンを用いるほうが安価であり、表示の手間も無いため、高速にキャリブレーションが行える。しかし、広い範囲を観測する人物検出システムのカメラは、広角であり、画像の周辺部ではチェッカーパターンが小さく撮影されるため、コーナ点の検出が困難になる。このため、本報告では、グレイコードによるキャリブレーション

の結果から、広角カメラで撮影した画像上で、ほぼ均一に写るチェッカーパターンを生成する方法についても検討する。

2. 関連研究

1990 年代後半から 2000 年初期にかけて、国内外で自動視覚監視に関する研究プロジェクトが実施されてきた。日本での分散協調視覚システム [1] や、米国での Visual Surveillance and Monitoring(VSAM) プロジェクト [2] などが代表的なものである。これらのプロジェクトでは複数のカメラを分散配置させたシステムを用いて、現実の空間に起きる様々な変化を監視する研究が行われてきた。

複数台のカメラを連携させると、1 台のカメラだけでは監視できない広範囲を監視したり、1 台のカメラだけでは遮蔽物で隠れて撮影できないような箇所を監視できるというメリットがある。

2.1 複数カメラの協調動作による人物追跡に関する研究

ここでは、複数台のカメラ同士を連携させることで単一、もしくは複数の人物の検出や追跡を行っている研究に関して述べる。

Particle Filter(CONDENSATION と呼ばれる)によって 3 次元空間中の頭部位置に関する仮説を生成し、仮説の尤度を各カメラで撮影した画像上で評価し、再統合する際の計算方法を工夫することで、オクルージョンに対して頑健な追跡を実現した方法 [3] がある。他にも画像中の人物位置を表す Particle と、床面の人物の足元位置を表す Particle を複数台のカメラ間で協調させることによって、人物の追跡を行っている研究 [4] もある。

これらの研究では、いずれもカメラで撮影した画像をコンピュータで処理し、処理した結果から得られた情報を様々な方法で統合することによって、人物の検出と追跡が行われている。そのため、2.2 で述べるような人物位置のみを用いた人物追跡を行うことはできなかった。

2.2 地図を用いた人物検出システムの連携に関する研究 [5]

先述した研究とは異なり、カメラそのものが人物検出の機能を備えた人物検出システムを用いて、各カメラから送られてくる人物の位置情報を処理するだけで広域人物追跡を実現している。

この研究では、人物検出器から得られた画像中の人物の足下の位置を、広域人物追跡を行う単一の地図上の位置に変換して、地図上の位置の近接性に基づいた人物追跡および異なるカメラ間での同一人物の判定を行っている。

人物検出システムでは実時間性が求められるため、この研究では線形回帰木の一種である PaLM-Tree[6] を用いて効率よく座標変換を行っている。画像から地図への変換を

行う PaLM-Tree を構築するためには画像と地図での複数点の対応付けが必要になるが、この研究ではこの対応点を求めるために、撮影した画像の全体に点が分布するよう床面にマーカを設置してから、それらの点が地図上のどこに対応するかという関係を1つずつ手で求めていた。複数のカメラを設置する際にこれらの作業を行うことは、実運用上、大きなコストになるということが論文中で課題として挙げられている。

3. 線形回帰木と Homography によるキャリブレーション

本章では人物検出器を設置する際のカメラキャリブレーションを最小限のコストで行う方法について述べる。本手法では設置時にカメラ画像と地図を直接対応付けすることで、内部パラメータと外部パラメータのキャリブレーションの両方をしてきた先行研究 [5] の方法とは異なり、従来のカメラキャリブレーションに方法に則り、カメラの内部パラメータと外部パラメータを別々に補正する方法で地図へのマッピングを行う。レンズ歪み等の内部パラメータは事前に補正しておき、カメラの設置位置や角度といった外部パラメータはカメラ設置後に補正する。最後に、2つのキャリブレーションに使用した変換を1つの線形回帰木にまとめて、これを実際のマッピングに用いる。

3.1 PaLM-Tree

本手法では線形回帰木の一種である PaLM-Tree[6] を非線形写像の学習に用いる。

まず、 p 次元ベクトルの説明変量を $\mathbf{x}_k = (x_{1k}, x_{2k}, \dots, x_{pk})^\top$ 、 q 次元ベクトルの目的変量を $\mathbf{y}_k = (y_{1k}, y_{2k}, \dots, y_{qk})^\top$ と表す。また、 p 次元のベクトル空間 R^p と q 次元のベクトル空間 R^q をそれぞれ入力空間、出力空間と呼ぶ。PaLM-Tree では説明変数と目的変数の間の非線形写像関係 $F: R^p \mapsto R^q$ を推定する問題を扱う。

PaLM-Tree では、通常の線形回帰木と同じく、入力空間を写像の推定誤差に基づいて複数の定義域に分割し、各定義域ごとに対応する目的変量に対する線形写像を推定する。具体的には、説明変量と目的変量の間の非線形写像 F を複数の線形写像 $\phi_i: i \geq 1$ に分解する。また、PaLM-Tree では入力空間の分割を2分木により表現している。これを2分木として用いることで、新たな入力を与えられたときに、与えられた入力に属する定義域を高速に求めることができる。

PaLM-Tree の特殊性は以下のように Split and Merge の要領で構築されることである。はじめに、線形回帰である Principal Component Regression(PCR) によってノード τ 内のデータ集合 $D(\tau)$ に対する線形写像を推定し、回帰行列 $M(\tau)$ 、推定誤差 $err(D(\tau))$ を求める。 $D(\tau)$ 内の個々のデータに対する目的変数の推定誤差が閾値以上になった

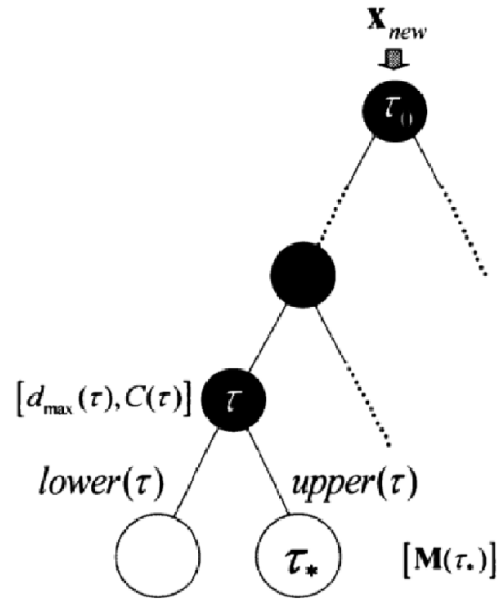


図 2 PaLM-Tree の探索と回帰計算 ([6] より引用)

場合、入力空間を2つの超直方体領域に分割し、新たな子ノード $upper(\tau)$ と $lower(\tau)$ を生成し、これらに対して同様に、線形写像の推定と推定誤差の計算を行い、推定誤差が閾値以下になるまで再帰的に分割を繰り返す。そして、 τ 内のすべてのデータの推定誤差が閾値を下回った場合、ノード τ を終端ノードとし、回帰係数を保存する。

次に汎化性能を向上させるために、先ほど分割した領域の併合を行う。まず、任意の葉ノード τ_f と隣接する葉ノード τ_n に格納されているデータ集合 $D(\tau_f)$ と $D(\tau_n)$ の和集合 D' について PCR によって線形写像を求める。このとき推定誤差が閾値以下になった場合は、新たに求められた回帰係数を共有することで、一度分割した入力空間を併合する。PaLM-Tree では、この併合を行うことによって、過学習の問題を解決している。

PaLM-Tree 構築後は、新たなデータ \mathbf{x}_{new} が来てきたときに、2分木を探索することで \mathbf{x}_{new} が入力空間のどの定義域に属するかを調べる。そして、2分木の葉ノードに到達した際には、PCR で求めた回帰行列 $M(\tau_*)$ を用いて出力 $\hat{\mathbf{y}}$ は

$$\hat{\mathbf{y}} = M(\tau_*)(\mathbf{x}_{new} - \bar{\mathbf{x}}) + \bar{\mathbf{y}}$$

として計算される。(図 2) このときの $\bar{\mathbf{x}}, \bar{\mathbf{y}}$ はノード τ_* に属する、説明変量と目的変量の平均ベクトルである。

本手法ではカメラ画像上の座標 ($p = 2$) から歪みの無い平面上の座標 ($q = 2$) への写像を学習し、レンズ歪みの除去を行う。

3.2 外部パラメータのキャリブレーション

カメラ設置後に撮影を行い PaLM-Tree を用いて歪み除去を行うことで、歪みを除去した画像を得ることができる。Homography 行列 H はこの画像中の床面と地図との対応

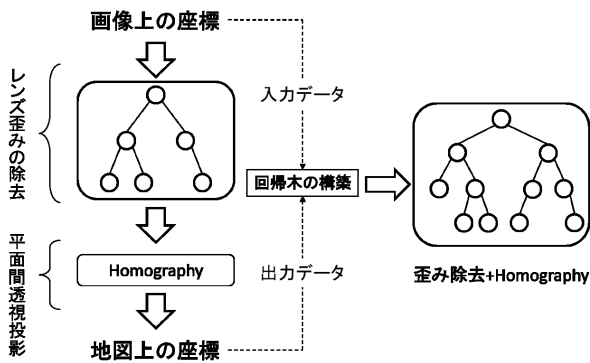


図 3 変換の合成の流れ

点を求めることで推定できる．新たに検出された画像中の人物位置を p_{in} とすると，地図上の人物位置を p_{out} は，

$$p_{out} = H\{M(\tau_*)(p_{in} - \bar{x}) + \bar{y}\} \quad (1)$$

のように求められる．Homography は最低でも 4 点の対応点を与えることで変換行列を推定できるので，先行研究のようにカメラ設置後に，画面全体に分布する点と地図との大量の対応点を手で求める必要はなく，キャリブレーションのコストを抑えることができる．

3.3 変換の合成

式 (1) によって，カメラ画像から地図への人物位置のマッピングが行えるようになった．しかし，このままでは人物位置を地図にマッピングする際に 2 度の座標変換を行う必要があるため，計算速度の面でデメリットが生じる可能性がある．そこで本手法では，これら 2 つの変換を 1 つの線形回帰木に統合し，それを用いて画像から地図への座標変換を行うことで，コンパクトかつ高速な座標変換を実現する．

カメラ画像上に一様に点を分布させた点群の座標を入力し，それらの点群を地図上の座標へと変換したものを出力とする学習データで PaLM-Tree を再度構築する (図 3)．こうすることで，実際に人物が検出された際には PaLM-Tree の探索と 1 回の回帰計算のみで地図上の人物位置を求めることができる．

4. レンズ歪みにあわせたチェッカーパターンの作成

本章では，内部パラメータのキャリブレーションを行う際に必要な対応点を容易に得ることができるようなチェッカーパターンを作成する方法について説明する．

通常，レンズ歪みの除去を行う際にはチェッカーパターンをカメラで撮影し，撮影した画像上のコーナ点と実世界上のコーナ点の対応関係からレンズ歪みの推定をするという方法がとられる．しかし，本手法で用いるカメラは広い範囲を撮影できるようにするために広角レンズを使用し

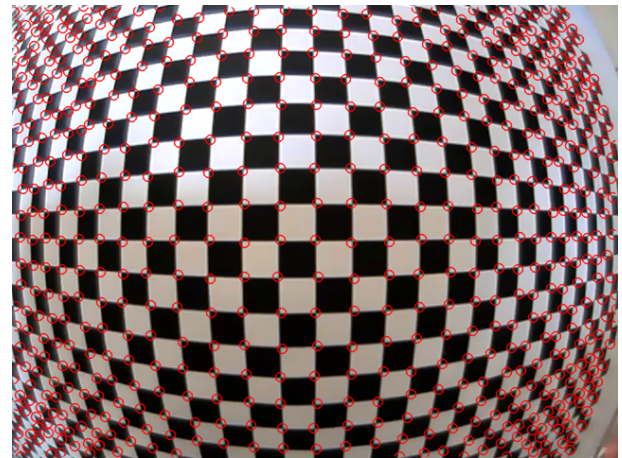


図 4 チェッカーパターンを撮影しコーナ検出を行った結果

ているため，レンズ歪みが非常に強い．そのため，チェッカーパターンを撮影すると図 4 ように撮影した画像の 4 隅にコーナ点が大量に出現してしまい，コーナ点の検出と対応付けが困難になる．

そこで，本研究ではグレイコードを用いてカメラ画像上の画素とディスプレイ上の画素の対応関係を求めておき，求めた対応関係に基づいて，カメラ画像上で歪みの少ないチェッカーパターンとして写るような，歪みを有するチェッカーパターンを作成する．こうすることで，キャリブレーションに必要な対応点を容易に求めることができる．この方式のメリットは，一旦，歪みチェッカーパターンを作成してしまえば，紙などに印刷して使用することが可能であるため，カメラキャリブレーションのたびにディスプレイを用意する必要がないということである．

4.1 グレイコードによる画素ごとの対応付け

画素ごとの対応関係を求めるためには，まず，ディスプレイを水平方向と垂直方法に分割し，それぞれの領域に対してグレイコードを割り当てた縞模様の画像を作成し，bit の数だけ用意する．それらを順に表示し，撮影した画像をデコードすると，画素ごとの対応関係を求めることができる．この対応関係に基づいてチェッカーパターンの画像を参照することで歪みチェッカーパターンを作成することができる．

グレイコードには隣り合う数字同士のハミング距離が常に 1 であるという性質があるので，デコードの際に縞模様の境界付近で起きる輝度値の判定ミスによる座標のズレが少ないという利点がある．さらに，図 5 からわかるように，グレイコードでは縞の境界がある場所で 1 度現れると，それ以外の bit の画像では同じ場所に現れないことが保証されているので，判定ミスは 1 回しか起こりえないというメリットもある．

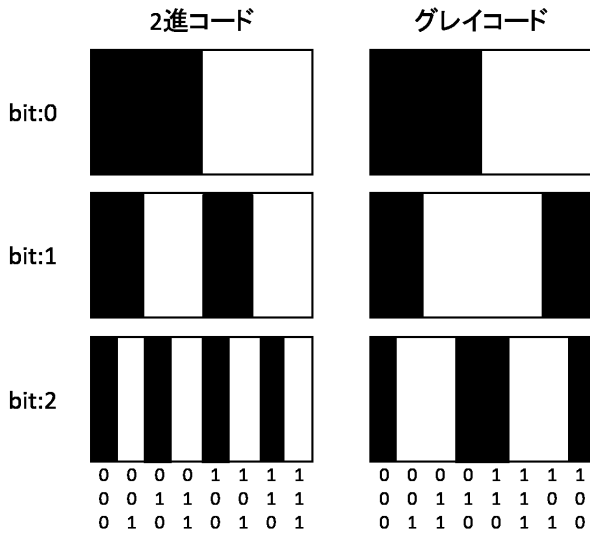


図5 2進コードとグレイコードの比較

5. 実験

本研究の有効性を確認するためにカメラキャリブレーションの実験と人物追跡に関する実験を行った。実験には人物検出システムとして技研トラステム株式会社のPALOSSIEAIOを使用した。この人物検出システムはカメラ画像に人物モデルを当てはめることで人物の位置を検出するものであり、1台で画像中の人物を最大100人まで検出可能である。

5.1 グレイコードに関する実験

ディスプレイにグレイコードを表示させて画素ごとの対応付けを行い、カメラに適したチェッカーパターンが作成できるか確認した。実験には解像度1980×1080のディスプレイを使用し、1画素ごとにグレイコードを割り当てて、画素ごとの対応付けを行った。グレイコードをカメラで撮影し、対応付けを行った結果図6のようなチェッカーパターンが得られた。

このチェッカーパターンをカメラで撮影すると図7のコーナー点の検出が容易な画像が得られる。このとき、画像の所々に誤った対応付けが原因で発生したジャギーが見られ、誤ってコーナー点として検出されてしまう問題があったが、コーナー検出の際にガウシアンフィルタを適用し画像をぼかすことで解決した。

5.2 キャリブレーションの精度に関する実験

キャリブレーションの精度を確かめるために実験を行った。実験は単体の人物検出システムで行い、カメラ画像中で検出された人物位置を線形回帰木でマッピングし地図の上での位置を求めた。実際に人物が移動した経路が図8であり、マッピングの結果が図9となっている。

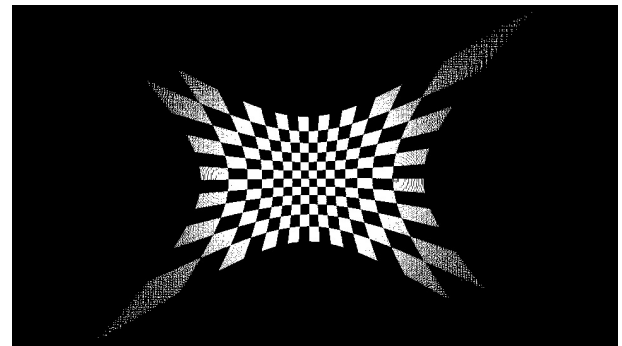


図6 作成したチェッカーパターン

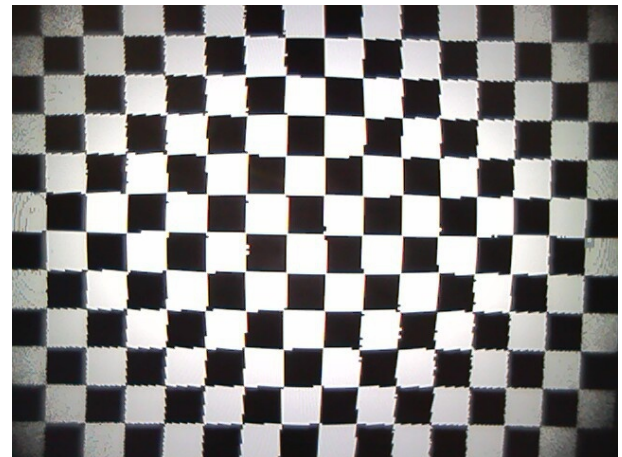


図7 図6を撮影した画像

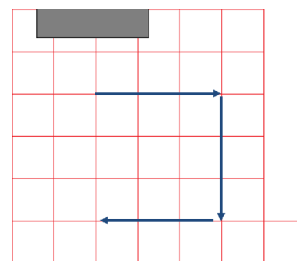


図8 人物が実際に歩いた経路

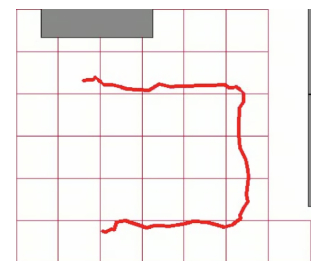


図9 マッピングの結果

実験の結果から、カメラ設置時に求める対応点を減らしても、検出した人物位置を正しく地図上にマッピングできることが確認できた。若干の誤差は見られるが、これは先行研究[5]でも挙げられている、人物モデルと実際に検出した人物との身長差によって検出位置がずれる性質が原因であると考えられる。

5.3 キャリブレーションの速度に関する実験

次に、マッピングに要する時間を確認するための実験を行った。この実験では、カメラ画像と同じ大きさの640×480の2次元空間に点をランダムに発生させ、それらの点をすべて地図にマッピングした際にかかる時間について調べた。

速度の計測は、PaLM-TreeとHomographyによるマッピングと、これら2つの変換を合成したPaLM-Treeによる変換の2つについて、データ数を変化させて行った。

Number of data	compiled tree[ms]	PaLM-Tree & Homography[ms]
10	0.051	0.059
100	0.460	0.468
1000	4.737	4.909
10000	45.098	46.176

表 1 マッピング速度の比較

2つの方法について、データ数を変化させ処理で速度を計測した結果を表 1 になる。この表からわかるとおり、2つの変換を合成した回帰木を用いる方が、わずかであるが、速度が早いと言える。

また、検出システムが検出できる最大の人物数である 100 個の点でマッピングを行っても、速度は 4.737[ms] であることから実時間で処理が行えていると言える。

6. おわりに

本報告では、広域人物検出・追跡システムを設置時にカメラキャリブレーションのコストを大幅に削減する方法を示した。

実験ではレンズ歪みの強いカメラに適したチェッカーパターンを撮影することで、求めたチェッカーパターンから線形回帰木を用いてレンズ歪みが除去できていることを確認した。また、外部パラメータに Homography 変換を用いて対応点の数を最も少ない 4 点にしても、マッピングの精度に問題がないことを示した。最後に、構築した線形回帰木の探索と回帰計算にかかる時間を計測し、実時間でマッピングを行えることを示した。

今後の発展としてカメラ設置後に歪みを除去した画像同士間の共通視野内で対応点 4 点を求め、Homography を求めることで、複数のカメラ画像を張り合わせて地図を作成することで、地図を作成するコストも削減するということが考えられる。

謝辞 この研究は技研トラステム株式会社からの支援を受けて実施された。

参考文献

- [1] 松山隆司: 分散協調視覚-研究成果と今後の展望-, *CVIM121-6*, pp. 41-48 (2000).
- [2] R.T.Collins, A.J.Lipton and T.Kanade: Introduction to the Special Section on Video Surveillance, *IEEE Transactions on Pattern Analysis and Machine Recognition*, Vol. 22, No. 8, pp. 745-746 (2000).
- [3] 松本郁祐, 加藤丈和, 和田俊和: 複数カメラを用いた CONDENSATION のための隠れに対して頑健な重み統合, *情報処理学会論文誌, CVIM*, Vol. 48, No. SIG9(CVIM18), pp. 100-114 (2007).
- [4] W.Du and J.Piater: Multi-camera People Tracking by Collaborative Particle Filters and Principal Axis-Based Integration, *ACCV* (2007).
- [5] 森 敦, 和田俊和: 地図を用いた人物追跡システムの連携法に関する研究, *情報処理学会研究報告*, Vol. 2015-CVIM-195, No. 39, pp. 1-7 (2015).

- [6] 中村恭之, 加藤丈和, 和田俊和: 非線形写像のための PaLM-Tree の提案, *日本ロボット学会誌*, Vol. 6, No. 12, pp. 1043-1048 (2004).