

画像から復元された3次元特徴点マップ群の世界座標系への統合 Integration of Reconstructed 3D Feature Maps into World Coordinate

林将之 北原格 亀田能成 大田友一
Masayuki Hayashi Itaru Kitahara Yoshinari Kameda Yuichi Ohta

筑波大学 大学院システム情報工学研究科
Graduate School of Systems and Information Engineering, University of Tsukuba.

1. はじめに

現実空間と仮想空間を融合して表示する複合現実感 (Mixed Reality: MR) を実現するために、画像からカメラの位置・姿勢を推定する手法が研究されている。そのアプローチには、予め与えられた環境の3次元モデルに基づいて用いてカメラトラッキングを行うモデルベースの手法と、未知の環境の3次元特徴点マップを復元し、それに基づいてカメラトラッキングを行う手法がある。環境の正確な3次元モデルを作成するのは容易ではないため、後者の方が手軽に利用できるという利点がある。

後者の手法で Castle ら[1]は、広範囲の環境を複数のマップに分けて事前に復元しておき、必要に応じて利用するマップを切り替えることで広範囲でのトラッキングを実現している。しかし、この手法ではマップは復元時のカメラの初期位置・姿勢から決まる相対座標系に基づいており、マップ毎に異なる座標系で構築されているため、表示させる仮想物体の位置・姿勢を単一の世界座標系上で定義することが出来ない問題がある。

本稿では、異なる相対座標系で復元されたマップ群を事前に定義した世界座標系へ統合する手法を提案する。この手法では、事前にマップを復元し、その中の3点以上に対し世界座標系における座標値を与えておく。提案手法は未知の環境では利用出来ないが、モデルベースの手法に比べて必要な点数が大幅に少なくて済むという利点がある。マップの復元とカメラトラッキングは文献[1]と同じく、Klein らの PTAM (Parallel Tracking And Mapping) [2]を用いて実験を行った。

2. 相似変換行列の推定

ここでは、ある座標系 C で表されたマップを世界座標系 W に変換する手法を述べる。2つの座標系の関係は相似変換で表現できるものとする。また、両方の座標系で対応の与えられた N 個の点があり、 i 番目の点を座標系 C で表したものを \mathbf{X}_i 、座標系 W で表したものを \mathbf{Y}_i とすると、その関係は相似変換 \mathbf{H} を用いて以下のように表せる。

$$\mathbf{Y}_i = \mathbf{H}\mathbf{X}_i \quad (i=1, \dots, N), \quad \mathbf{H} = \begin{bmatrix} \mathbf{S}\mathbf{R} & \mathbf{T} \\ \mathbf{0} & 1 \end{bmatrix}$$

ここで、 $\mathbf{S}, \mathbf{R}, \mathbf{T}$ はそれぞれ、3次元のスケール、回転、並進を表す行列である。相似変換 \mathbf{H} はスケール \mathbf{S} とユークリッド変換 \mathbf{He} に分解できるため、本稿ではまず行列 \mathbf{S} を求め、続いてユークリッド変換 \mathbf{He} を特異値分解を用いた手法[3]によって求める。

行列 \mathbf{S} は3次元の単位行列 \mathbf{I} と拡大率 s を用いて $\mathbf{S} = s\mathbf{I}$ で表せ、拡大率 s は任意の2点 $i, j (i \neq j)$ 間の距離を用いて以下のように表せる。

$$s = \frac{|\mathbf{Y}_i - \mathbf{Y}_j|}{|\mathbf{X}_i - \mathbf{X}_j|}$$

本稿では、 \mathbf{Y}_i もしくは \mathbf{X}_i の値にノイズが含まれることを想定し、 s は全ての組み合わせで計算した値の中央値を取る。また、[2]では、並進 \mathbf{T} を点群 \mathbf{X} と \mathbf{Y} の重心の距離としているが、本稿ではこれを \mathbf{X}_i と \mathbf{Y}_i の距離として \mathbf{He}_i を計算し、4節に示す3次元再投影誤差が最も小さくなる \mathbf{He}_i を \mathbf{He} とする。

3. 世界座標系への統合手順

PTAM[2]が生成するマップを現実空間のスケールを持つ世界座標系へ統合する具体的手順を述べる。マップは、ある時点のカメラ画像 (キーフレーム) とそこから抽出された特徴点の3次元座標 (推定値) によって構成される。マップを復元した後、特徴点の世界座標系における3次元座標を測量機を用いて計測する。こうして得られた推定値と計測値を対応点として、マップを世界座標系に変換する相似変換行列を推定する。

4. 再投影誤差の評価

推定した相似変換行列 \mathbf{H} を評価するため、以下の式で3次元再投影誤差の平均値を計算した。

$$E = \frac{1}{N} \sum_i (\mathbf{Y}_i - \mathbf{H}\mathbf{X}_i)$$

室内を撮影して復元した19キーフレーム、1871点からなるマップのうち、10点の世界座標を測量機 Leica TCR805 を用いて計測し、相似変換行列を算出した。その結果、3次元座標の再投影誤差は 9.0[cm] となった。誤差の原因には、PTAM が推定した3次元位置推定誤差の影響が考えられる。より高い精度で変換行列を求める手法として、今回のように世界座標系で計測した3次元座標が信頼できる場合、これを固定点としてマップの3次元座標の再推定を行う方法が考えられる。

参考文献

- [1] Castle, Klein and Murray, "Video-rate Localization in Multiple Maps for Wearable Augmented Reality," IEEE ISWC, 2008
- [2] Klein and Murray, "Parallel Tracking and Mapping for Small AR Workspaces," IEEE and ACM ISMAR, 2007
- [3] 玉木, "姿勢推定と回転行列," 信学会技報. SIS, 2009