

屋外環境下での広範囲における強化現実端末の自己位置・姿勢検出 精度向上法の基礎的検討

A Study of Improving the Estimation Accuracy of the Position and Orientation of an Augmented Reality Terminal in a Wide Area

藤原 裕也†
Yuya Fujiwara

大谷 淳†
Jun Ohya

1. まえがき

古代遺跡等において過去の建造物のCG映像と現在の風景とを強化現実 (AR) 端末を用いて重畳表示するためには、端末の位置と向きを正確に計測する必要がある。既存のARシステムの多くはGPSとジャイロを利用しているが、位置と向きの計測精度に課題を残してきた。

筆者らはGPS、ジャイロ、CCDカメラと赤外発光ダイオードによるマーカーでの画像処理モジュールを搭載した携帯端末を提案し検討を進めてきた[1]。以前に筆者らが提案したIRマーカーを用いた屋外環境下での自己位置・姿勢検出精度向上法[1]によってGPSとジャイロによる検出結果の精度を向上させることに成功した。しかし、その手法ではIRマーカーに対して3m以内の範囲でしか推定できないのでマーカーを多数配置しなければならないという問題点が今後の課題として挙げられた。

本報告ではこの問題点を解決するための手法を提案する。具体的には、フィールド上に複数配置されているIRマーカーは十分離れた位置においては画像に対するマーカーの大きさが非常に小さくなる。よってマーカー1つ1つを特徴点としてとらえることができる。カメラによって撮影される動画像からこの特徴点を追跡することにより、Tomasi-Kanadeによる因子分解法[2]を用いてそれぞれの特徴点(IRマーカー)とカメラとの相対関係を求め、カメラの位置と角度を推定する。

2. システム

2.1 ARBoard

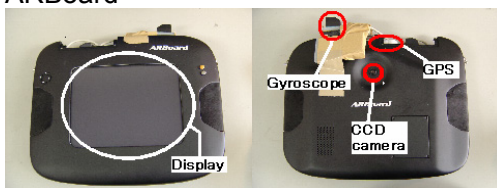


図1 ARBoard

筆者らが研究しているARBoardの外観を図1に示す。ARBoardの内部にはTablet PC、GPS、ジャイロおよびCCDカメラを内蔵している。カメラはボードの裏面に搭載されており(図1右)、カメラから取得した映像は一度PCに取り込まれた後、前面のディスプレイ(図1左)に表示される(ビデオスルー方式)。そのため、PCにより生成したCGの画像と周囲の風景を重畳表示することが可能になる。

2.2 IRマーカー

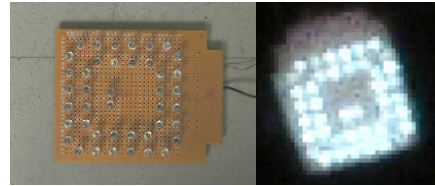


図2 IRマーカー

画像処理により端末の位置と向きを推定する場合、目印となるものを利用する手法が最も一般的である。しかし、目印となる建造物等が残されていないような遺跡においては新たに目印となるマーカーを設置することが必要となる。その場合、このようなマーカーは極力目立たないことが望ましい。そのため、従来手法では図2に示すように、赤外LEDを複数用いて正方形の既知パターン(IRマーカー)を作成し環境中に設置した。図2左はIRマーカーの外観であり、右はLEDを点灯させIRパスフィルタを通し、マーカーをCCDカメラで撮影したものである。ここで、CCDカメラは赤外光の波長にも感度があるので、図2右では点灯している様子が観測されるが、人の目には見えない。

3. 改良手法

3.1 従来法とその問題点

取得画像から抽出したIRマーカーの外側の境界を最小二乗法により各辺を近似した直線から交点を求める。画像中におけるマーカーの4つの頂点の座標が得られることによりARToolKitの計算アルゴリズム[3]を利用してカメラを基点としたマーカーの3次元座標と、カメラに対する傾きを算出することができる。

従来法の問題点としてマーカーから3m以上離れるとマーカーを認識することが不安定になった。これはCCDカメラからの取得画像に対してマーカーの大きさが極端に小さくなってしまふからであり、そのため形状がつぶれてしまうのでマーカーと認識できなくなることが原因であったと考えられる。

3.2 前処理

本実験では赤外LED以外の光(可視光)をノイズとして扱う。IRパスフィルタは赤外光以外に赤色に近い波長の光をわずかに透過する。そこで前処理として可視光の影響を低減させるため、画像における赤色成分の大きさを下げる。

3.3 特徴点追跡

取得した画像における各マーカーの重心を特徴点とみなす。前フレームに対して四近傍の点を現在のフレームにおいて同じ特徴点として認識し、ラベリングを行う。このことにより画像中の各特徴点座標を得ることができる。

3.4 因子分解法によるカメラの傾きと位置の推定

連続した f フレーム目の画像中における特徴点 P の x 座標を u_{fp} とし、 y 座標を v_{fp} とする。このとき $2F \times P$ の観測行列 W を定める。

$$W = \begin{bmatrix} U \\ V \end{bmatrix} = \begin{bmatrix} u_{11} & \cdots & u_{1p} \\ \vdots & & \vdots \\ u_{F1} & \cdots & u_{FP} \\ v_{11} & \cdots & v_{1p} \\ \vdots & & \vdots \\ v_{F1} & \cdots & v_{FP} \end{bmatrix} \quad (1)$$

各フレームに対して特徴点全体の x 座標の平均値 a_f と y 座標の平均値 b_f より \tilde{u}_{fp} と \tilde{v}_{fp} を

$$\tilde{u}_{fp} = u_{fp} - a_f \quad (2)$$

$$\tilde{v}_{fp} = v_{fp} - b_f \quad (3)$$

とする。ただし、

$$a_f = \frac{1}{P} \sum_{p=1}^P u_{fp} \quad b_f = \frac{1}{P} \sum_{p=1}^P v_{fp} \quad (4)$$

ここで $F \times P$ の行列 $\tilde{U} = [\tilde{u}_{fp}]$ と $\tilde{V} = [\tilde{v}_{fp}]$ を用いて計測行列 \tilde{W} は以下のように表せる。

$$\tilde{W} = \begin{pmatrix} \tilde{U} \\ \tilde{V} \end{pmatrix} \quad (5)$$

この計測行列 \tilde{W} はフレーム数 F と特徴点の数 P が変化しても行列の階数は高々3であるため、次のようにカメラ回転行列 M と形状行列 S に分解できる。

$$\tilde{W} = MS = \begin{bmatrix} i_1^T & \cdots & i_F^T \\ j_1^T & \cdots & j_F^T \end{bmatrix}^T \begin{bmatrix} s_1 & \cdots & s_P \end{bmatrix} \quad (6)$$

実際には計測行列にノイズが含まれるため、行列の階数が3であるとは限らない。このとき特異値分解を用いて大きな3つの特異値を持つように分解すると、二乗誤差の最小化により最適分解を得ることができる。

$$W = U\Sigma V \quad (7)$$

ここで U 、 V はそれぞれ $2F \times P$ 、 $P \times P$ の直交行列であり、 Σ は $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0$ を対角要素とする対角行列である。これらより

$$\hat{M} = U\Sigma^{\frac{1}{2}} \quad \hat{S} = \Sigma^{\frac{1}{2}}V^T \quad (8)$$

のように $\tilde{W} = \hat{M}\hat{S}$ となるような \hat{M} と \hat{S} を定める。ここで M の行列の成分である i_f と j_f において以下のような拘束条件が成り立つ。同時に、 $i_f^T = \hat{i}_f^T A$ とおくと以下のような式が成り立つ。

$$\begin{cases} i_f^T i_f = \hat{i}_f^T A A^T \hat{i}_f = 1 \\ j_f^T j_f = \hat{j}_f^T A A^T \hat{j}_f = 1 \\ i_f^T j_f = \hat{i}_f^T A A^T \hat{j}_f = 0 \end{cases} \quad (9)$$

このときの行列 A を用いると \hat{M} によって M を求めることができる。

$$M = \hat{M}A \quad (10)$$

式(6)、式(10)より \hat{S} によって S を以下のように表せる。

$$S = A^{-1}\hat{S} \quad (11)$$

式(9)より $L = AA^T$ とおくと行列 L を求めることができ、 L をコレスキー分解することにより行列 A を求めることができる。

式(10)よりカメラの回転行列 M が求まり、形状行列 S とカメラで撮影した特徴点の座標からカメラとマーカの相対的な位置座標を推定することができる。

4. 実験結果と考察

精度を検証するため実験を行った。カメラとマーカとの距離、およびカメラの角度を認識したときの誤差を表1に示す。

表1 マーカとの距離とカメラの角度の認識結果

距離[cm]	誤差[cm]	誤差[deg]		
		yaw	roll	pitch
30.0	2.7	0.38	-0.13	3.75
40.0	1.2	-0.98	2.71	-1.79
50.0	-1.4	-4.26	0.52	-0.70

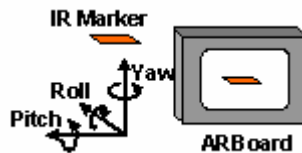


表1より、誤差の大きさは一般的なGPSとジャイロの誤差よりも小さくなった。また、今回の手法ではマーカの形状や姿勢を認識する必要がない。

よって、マーカとカメラの距離が大きくなり画像の大きさに対するマーカの大きさが小さくなったとしても、認識する精度には影響がないと考えられる。以上の理由から、従来法[1]よりもマーカに対して広範囲にわたりカメラの位置と姿勢を測定することが可能になることが期待できる。

参考文献

- [1] 藤原 裕也, 大谷 淳, “赤外発光ダイオードを用いた屋外環境下での強化現実端末の自己位置・姿勢検出精度向上法の基礎的検討” 電子情報通信学会総合大会, 2004
- [2] C. Tomasi and T. Kanade, Shape and Motion from Image Streams: a Factorization Method Full Report on the Orthographic Case, Cornell TR 92-1270 and Carnegie Mellon CMU-CS-92-104, March 1992.
- [3] 加藤 博一, “マーカ追跡に基づく拡張現実感システムとそのキャリブレーション” 日本バーチャルリアリティ学会論文誌, Vol.4, No.4, pp.607-616, 1999