

2次元画像マーカのパターンマッチングによる ARのためのカメラトラッキングの高精度化

植松 裕子[†] 齋藤 英雄[†]

[†] 慶應義塾大学大学院理工学研究科

E-mail: †{yu-ko,saito}@ozawa.ics.keio.ac.jp

あらまし 複合現実感などに用いられる2次元矩形マーカを用いたカメラトラッキングにおける問題点として、カメラの投影面がマーカ平面に対してほぼ平行であるとき、つまりマーカ平面に対するカメラの投影軸が垂直に近いときに、マーカ平面(X-Y)に対する垂直軸(Z軸)方向の推定精度が低下し、トラッキングが不安定になることが挙げられる。本研究では、パーティクルフィルタを利用してカメラの動きを推定し、推定したパラメータによって投影されるモデルパターンと、実際に撮影された入力画像が最も一致するようなパラメータを探索する。本手法ではARToolkitの2次元マーカを発展させ、2値パターンではなく、階調を持ったグレースケールパターンを用い、外枠のエッジとともにモデルとのパターンマッチングを行う。そして、マッチング結果をパーティクルフィルタの評価値として最適なパラメータを推定する。モデルパターン中からあらかじめ特徴点を抽出しておき、その注目点周辺でマッチングを行うことによって、計算コストを減少させる。その結果、ARの幾何学的位置合わせにおける重要な問題であるフレーム間のジッタや揺らぎを軽減する。

キーワード 複合現実感, カメラトラッキング, 2次元矩形マーカ, パーティクルフィルタ

Accuracy Improvement of Camera Tracking for AR by Pattern Matching of 2D Fiducial Marker

Yuko UEMATSU[†] and Hideo SAITO[†]

[†] Graduate School of Science and Technology, Keio University

E-mail: †{yu-ko,saito}@ozawa.ics.keio.ac.jp

Abstract This paper presents a method for improving accuracy of marker-based tracking using a 2D marker for Augmented Reality. We focus on that camera tracking becomes unstable when the image plane of the camera is almost parallel to a marker plane. Especially, tracking of Z axis which is perpendicular to the marker plane (X-Y) becomes unstable. For improving tracking accuracy in this case, we search rotation parameters that are the fittest to projected pattern based on the particle filter. Our method introduces a grayscale pattern instead of a binary pattern of ARToolkit and uses the marker's edge and pattern matching of the inside pattern. Feature points are detected from the model pattern and are projected onto the input image. By matching the regions around the projected points and detected feature points, our method reduces computation cost. This method can reduce jitters between frames, which is a big problem in AR. In the experiment, we demonstrate that our method can improve the tracking accuracy of the 3D coordinate system compared with just using ARToolkit.

Key words Augmented Reality, camera tracking, 2D rect markers, particle filter

1. はじめに

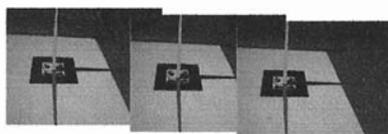
複合現実感とは、現実世界に仮想世界（仮想物体）を融合することにより、現実世界へ付加情報を与えることができる技術である。現実世界を撮影した映像（画像）の上に、別途作り出した仮想物体を重ねて表示することによって実現される。

コンピュータで作り出された仮想物体を現実世界に違和感なく重ね合わせるために重要となるのが、現実世界の3次元座標系と入力画像と2次元座標系との幾何学的な位置合わせ問題である。仮想物体と現実世界を正しく位置合わせするためには、映像を撮影しているカメラの動きを正しく追跡し、その動きに合わせて仮想物体を表示することが必要となる。

複合現実感におけるカメラトラッキングでは、ARToolkitなどで用いられる2次元（平面）の矩形マーカが広く利用されている[1]。マーカの輪郭情報を利用し、その位置および方向を画像から抽出することで、カメラの位置・姿勢をリアルタイムで安定に推定することが可能である。マーカの輪郭である4本のエッジを抽出し、その交点を矩形の頂点として算出することで3次元位置（並進）を決定し、またそれらの平行な直線の方向から3次元元的な姿勢（回転）を推定している。

このようなマーカを用いてカメラの動きを追跡するとき、マーカ平面に対するカメラの投影軸の角度が、その追跡精度に影響するという問題がある。特にマーカ平面（ X - Y 平面）に対して垂直な軸（ Z 軸）の方向の追跡精度は、マーカ平面に対するカメラの投影軸の角度が大きく影響する。

図1(a)のように、マーカに対して斜めの方向からカメラ



(a) マーカ平面に対して斜めの方向から撮影された画像列



(b) マーカ平面に対してカメラの投影軸がほぼ垂直となるような画像列

図1 3次元座標系追跡における画像列の例。

Fig. 1 Example images on tracking 3D coordinate system.

ラが撮影している場合には、既存の ARToolkit のアルゴリズムを用いてカメラの位置・姿勢を表すパラメータを正しく推定することができ、マーカ上の3次元座標系が正しく追跡できている。一方、図1(b)のように、マーカ平面に対してカメラの投影軸が垂直に近い場合、つまりカメラの投影面とマーカ平面がほぼ平行になっているような場合には、フレーム間でカメラがほとんど動いていないにも関わらず、軸の追跡に失敗してしまうことがある。これは、マーカの輪郭線である平行な直線の消失点を安定して求めることができないため、その推定結果の不安定さが Z 軸の方向の決定にも影響しているのだと考えられる。

しかしこの結果より、軸の方向の推定に関しては不安定であるが、3次元座標軸の原点はほぼ正しく求められていることが分かる。つまり、カメラ位置・姿勢のうち、位置を表す並進パラメータに関しては、正しく推定が行われているのではないかと考えられる。そこで本論文では、2次元の矩形マーカを用いたカメラトラッキングにおいて、マーカ平面に対するカメラの3次元姿勢を表す回転パラメータの推定精度を向上させる手法を提案する。特に図1(b)のような、既存の ARToolkit のアルゴリズムを用いるだけでは正しいトラッキングができないような画像列に対して、回転パラメータの推定精度を向上させることを目的とする。

最適な解決法としては、全てのパラメータ候補を全探索することが考えられるが、当然ながらオンラインシステムに適用することは計算量の面から不可能である。そこで我々は、時系列フィルタリング手法の一つであるパーティクルフィルタの枠組みを利用する。オンラインで動作可能な数のパラメータの仮説群をパーティクルフィルタに保持させておき、パーティクルフィルタによる仮のパラメータによって生成される投影パターンと、実際に撮影された画像とを比較することで、最適なパラメータを探索していく。事前の情報からの確率のみに基づいてパラメータを推定するだけではなく、実際に撮影された画像と推定したパラメータによる結果画像を逐次的に比較することができるため、常に最適なパラメータを探索することが可能になる。その結果、マーカ平面上の座標系（特に Z 軸）を正しく追跡できることになり、フレーム間のジッタや揺らぎも軽減させることができる。

本手法では、パーティクルフィルタで利用する重みを決定する評価基準としては、マーカの輪郭線であるエッジと、内部に描かれたパターンを用いる。エッジによる評価では、あらかじめマーカの輪郭線上で数点を等間隔でサンプリングし、それらの点を仮のパラメータで入力画像上へ投影し、実際のエッジとの距離を求めて重みを

算出する。また、内部のパターンによる評価では、あらかじめパターンから特徴点を抽出し、仮のパラメータでそれらの点が投影された入力画像上の点周辺の画素とマーカパターンとのマッチングをとって重みを算出する。従来のARToolkitのアルゴリズムでは、マーカの内部のパターンは、各マーカの識別のためと4方向の回転の向きを認識するためにだけ利用されており、カメラの位置姿勢推定には直接用いられていなかった。よって、内部のパターンをトラッキングに用いることは、精度向上につながると考えられる。

本手法では、評価値の計算にマーカ内部のパターンマッチングを用いるために、図2(a)のような従来のARToolkitに用いられている2値のマーカパターンではなく、図2(b)のようなグレースケールのパターンを利用する。これは、2値パターンを利用する場合に比べて、入力画像中のパターンと投影されるマーカパターンとの間のずれの度合いの差が、パターンマッチングの評価値の違いに、より表れやすいと考えられるからである。つまり、2値パターンの場合は、両者が非常に離れている場合の評価値と、わずかにずれた場合との評価値に、ほとんど違いが表れない。一方、グレースケールのパターンの場合は、そのずれの度合いに応じて、評価値が連続的に変化することが期待できる。そのために本手法では、グレースケールパターンを利用する。

本論文では、時系列フィルタを用いた関連研究について2章で述べた後、本手法の詳細および3次元軸を定義するためのパラメータの推定精度向上について3章で述べる。そして、実画像を利用した実験を行い、本手法の性能を評価する。

2. 関連研究

ここでは、時系列フィルタに基づくトラッキング手法の関連研究について述べる。時系列フィルタを用いた物体追跡手法は、過去の状況を基に確率的に推定を行うため、入力画像中の雑音などに対して頑健であることが知られており、近年、コンピュータビジョンを利用した人物追跡や頭部姿勢推定に応用した研究などが行われてい

る。時系列フィルタの中でも特に、パーティクルフィルタを利用した研究が多数行われている。岡らの手法では、頭部の3次元姿勢を6個のパラメータで表し、それらを状態ベクトルとするパーティクルフィルタの仮説群を適応的に拡散させることによって、高い精度で人物の頭部姿勢を追跡することに成功した[2]。

これらの手法は、対象物体の動きを追跡することが主目的であったが、一方で、撮影しているカメラの動きを追跡するためにパーティクルフィルタを利用した研究も行われている。

2次元マーカを用いたカメラトラッキングにパーティクルフィルタを適用した関連研究では、Marimonらの手法がある[3]。彼らは、1枚のマーカを用いたカメラトラッキングにおいて、カメラがマーカ周辺を動くのに従って、マーカがフレームから外れたり他の物体で隠されたりすることが起こり、結果として仮想物体の位置合わせができなくなる、という問題を解決するためにパーティクルフィルタを利用している。具体的には、マーカが見えているときにはマーカを用い、利用できなくなったときには、マーカの頂点などの特徴点を用いたトラッキング手法を用いることで、マーカが利用できなくなったフレームでの位置合わせを補っている。この両者の手法を組み合わせるためにパーティクルフィルタを用いている。

一方、我々の提案する手法では、マーカを用いたトラッキングにおける精度向上に焦点を当てている。2次元矩形マーカを利用したカメラトラッキングにおいて、カメラの3次元的な動きの変化と、マーカの2次元的な見え方の変化は、マーカに対するカメラの角度によって大きく異なることが知られており、その角度によっては、マーカ平面上に定義される3次元座標系の軸の追跡に大きく失敗することがある。そこで本手法は、軸の追跡に失敗するようなマーカの見え方の変化を入力画像中の雑音と同様に考え、その影響を受けずに安定に対象を追跡できるパーティクルフィルタを導入する。

3. パーティクルフィルタによるカメラトラッキング手法

本章では、提案手法で用いるパーティクルフィルタと、カメラ位置・推定手法について説明する。手法の全体の流れを図3に示す。本手法では、2次元矩形マーカを1台のカメラで撮影した画像を入力とする。まず、ARToolkitのアルゴリズムを用いて、画像中からマーカを抽出してカメラの並進パラメータを求める。次に、パーティクルフィルタを構成する仮説群が持つ回転パラメータと、マーカより求めた並進パラメータを用いて、各仮説に対する評価値を算出する。そして、評価値が最大となった仮説



図2 2次元矩形マーカ。

Fig. 2 2D rectangular marker.

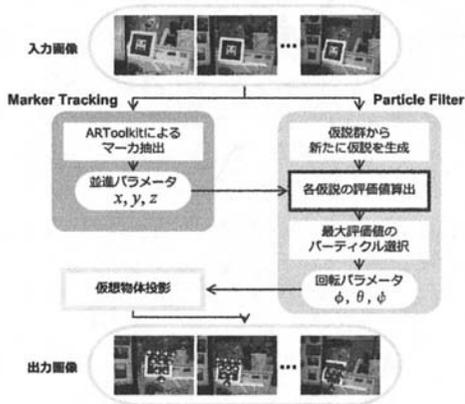


図3 提案カメラトラッキング手法の流れ.

Fig.3 Flow of our camera tracking method.

の回転パラメータを、現在のフレームにおける回転パラメータとして採用する。また、ここで求めた評価値を基にして仮説群を伝播し、次のフレームにおける仮説群を新たに生成する。なお、仮説群の初期値には、ARToolkitで求めた回転パラメータを用い、その周辺に仮説群を分散させる。

3.1 パーティクルフィルタの枠組み

本手法では、時刻 t におけるマーカ平面に対するカメラの3次元位置・姿勢のうち、回転成分を3次元状態空間中のベクトル $\mathbf{p}_t = (\phi_t, \theta_t, \psi_t)^\top$ として表現する。位置を表す並進パラメータ $(x_t, y_t, z_t)^\top$ は、前述のようにARToolkitによって求める。またパーティクルフィルタでは、3次元状態空間中の確率密度関数を、 N 個の離散的な仮説 $s_t^{(i)}$ とそれに対応する重み $\pi_t^{(i)}$ ($i = 1, \dots, N$) によって表現する。

はじめに、追跡開始時に仮説群の初期値を決定する。初期化の際には、ARToolkitによって求めた回転パラメータの値を利用し、その初期値近傍にランダムに仮説 $s_0^{(i)}$ を生成する。そして各仮説に一定の重み $\pi_0^{(i)} = 1/N$ を与え、仮説群の初期値を決定する。

その後、各入力画像列をもとに逐次的に推定を行う。まず、直前の時刻 $t-1$ で推定した仮説群 $\{s_{t-1}^{(i)}; \pi_{t-1}^{(i)}\}$ から、現時刻 t における仮説 $\{s_t^{(i)}; \pi_t^{(i)}\}$ を生成する。具体的には、次式のようにして各仮説を $s_{t-1}^{(i)}$ から $s_t^{(i)}$ へ伝播する。

$$s_t^{(i)} = s_{t-1}^{(i)} + v_{t-1} + \mu \quad (1)$$

ここで、 v_{t-1} はカメラの移動速度であり、時刻 $t-2$ から $t-1$ での移動量を表している。また μ はランダム雑音を表す。つまり、カメラが等速運動を行うという仮定に基づいて各仮説を動かした上で、各仮説にランダム雑

音を加えて拡散させることにより、時刻 t での仮説を生成することになる。

このようにして時刻 t における N 個の仮説が生成されたら、入力画像をもとに、各仮説 $s_t^{(i)}$ に対する重み $\pi_t^{(i)}$ を算出する。そして、最大の重みを持つ仮説 $s_t^{(i)}$ が時刻 t におけるパラメータであると考え、 $\mathbf{p}_t = (\phi_t, \theta_t, \psi_t)^\top$ とする。

4. 評価値の算出

本手法では、重みを決定する評価基準として、マーカの輪郭線であるエッジと、内部に描かれたパターンを用いる。そして、それぞれで計算された重みを合計した値が、各パーティクルに対する重みとなる。

4.1 エッジベースの評価値の算出

エッジによる評価値算出の流れを図4に示す。ここではあらかじめ、2次元矩形マーカの輪郭上で K 個の点をサンプリングしておく。本研究で行う実験では $K = 12$ とし、マーカの外側の輪郭線の各辺を3点ずつサンプリングする。次に、仮説が持つ回転パラメータとARToolkitから算出した並進パラメータを用いて、エッジ抽出した入力画像上へサンプリング点をそれぞれ投影する。そして、図5のように、各投影点を結んだ投影輪郭線(黄矩形)から垂直方向にエッジを探索していき、最も近いエッジまでの距離を算出する。

これらの距離を全てのサンプリング点に対して計算し、その値を d_j ($j = 1, \dots, K$) とする。さらに、算出された距離の合計が-1から1の値になるように、式(2)によって正規化を行う。つまり、正しいパラメータに近い仮説ほど、1に近い $score_t$ を持つようになる。なお max とは、図5に示すように、投影輪郭線から垂直にエッジを

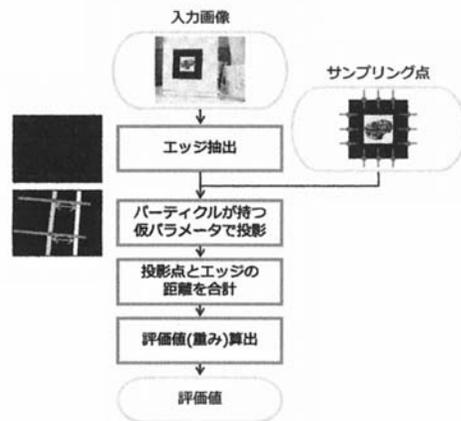


図4 エッジベースの評価値算出の流れ.

Fig.4 Flow of computing evaluation values based on edges.

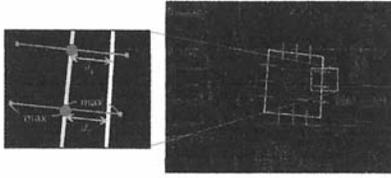


図5 投影されたサンプリング点からエッジまでの距離の算出
Fig.5 Computing distance between projected sampled points and edge.

探索するときの探索範囲の最大値である。

$$score_t^{(i)} = 1 - \frac{2 \sum_{j=1}^K d_j}{max} \quad (-1 \leq score_t^{(i)} \leq 1) \quad (2)$$

N 個の仮説に対して $score_t^{(i)}$ を求めたら、全ての仮説に対する $score_t$ の合計が 1 となるように、つまり最も正しい仮説が最大の値を持つように、ガウス関数を用いて式 (3) のようにそれぞれの仮説に対する $score_t$ を正規化する。この値を、エッジベースで評価したときの重み $\pi_t^{(i)}$ とする。

$$\pi_t^{(i)} \propto e^{-\frac{(1 - score_t^{(i)})^2}{2\sigma^2}} \quad \left(\sum_{i=1}^N \pi_t^{(i)} = 1 \right) \quad (3)$$

4.2 パターンベースの評価値の算出

パターンによる評価値算出の流れを図 6 に示す。計算時間に制限が無ければ、マーカの内部全体をパターンマッチングするのが望ましいが、本研究ではオンラインで動作可能なことが目的であるため、パターン中の特徴的な部分のみをパターンマッチングすることによって、計算時間を削減する。

本手法では、あらかじめモデルパターンから特徴点を抽出しておき、その特徴点を各仮説のパラメータによって入力画像上へ投影する。よって、毎フレーム入力画像から特徴点を抽出したり追跡する必要はない。本手法では、OpenCV に実装されている cvCornerMinEigenVal 関数を用い、画像中から大きな固有値を持つ角を特徴点として抽出する。実際に抽出した特徴点群を図 6 中に示す。

評価の流れはエッジベースの評価と同様に、仮説が持つ回転パラメータと ARToolkit から算出した並進パラメータを用いて、モデルパターンから抽出した特徴点を入力画像上に投影する。次に、投影点近傍でテンプレートサイズ $m \times m$ の領域を入力画像から図 7(a) のように切り出す。さらに、入力画像中で切り出された領域と同じ領域をモデルパターンから抽出するため、切り出された領域をモデルパターンへ逆投影し、特徴点を中心とし

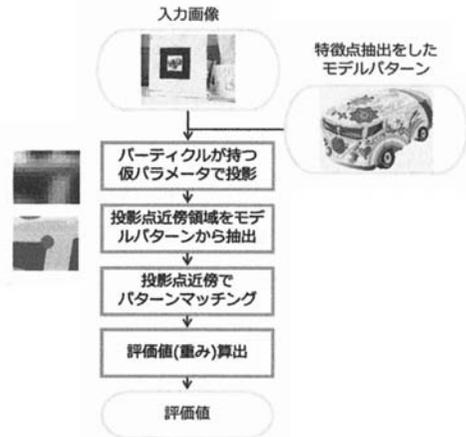


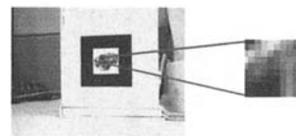
図 6 パターンベースの評価値算出の流れ。

Fig.6 Flow of computing evaluation values based on pattern.

た同じ領域をモデルパターンから図 7(b) のように切り出す。そして、入力画像の投影点近傍の領域と、モデルパターンの特徴点近傍領域とのパターンマッチングを行う。

入力画像とモデルパターンとのパターンマッチングを行うとき、図 7(a) から分かるように、入力画像は手ブレや焦点ボケなどの影響から、モデルパターンのような鮮明な画像が得られることは期待できない。そのため、入力画像から切り出された領域とモデルパターンのテンプレートをそのまま比較しても、良好なマッチング結果が得られるとは考えにくい。

そこで本手法では、あらかじめモデルパターンの方をある程度ぼかしてからパターンマッチングを行うことにする。図 7(c) に、図 7(b) のテンプレートをぼかした例を示す。ここでは一定量のボケを与えているだけである



(a) 入力画像から切り出したパターン



(b) モデルパターン



(c) ぼかしたモデルパターン

図 7 入力画像中から切り出した領域とモデルパターンのテンプレート。

Fig.7 Extracted region of input image and template of model pattern.



図 8 マーカ平面に対して斜めから撮影した入力画像列

Fig. 8 Input images captured from slanted view points with respect to marker's plane.

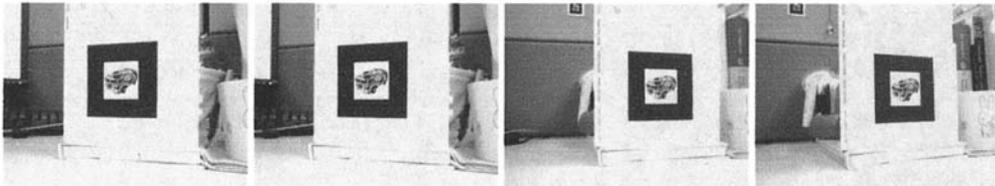


図 9 マーカ平面に対して垂直に撮影した入力画像列

Fig. 9 Input images captured from perpendicular view points with respect to marker's plane.

が、入力画像からそのボケの量を推定し、マーカパターンをボケさせるということも考えられる。

5. 実 験

本論文で提案した追跡手法を用いて、カメラトラッキングの実験を行う。本実験では、2次元の矩形マーカを1台のUSBカメラで撮影し、その画像(640×480 pixel)を汎用PC(OS:Windows XP, CPU: Intel Pentium 4.3 GHz)へと入力する。使用したマーカの矩形のサイズは80×80 mmであり、パーティクルフィルタの仮説数は300とする。

まず、図8のように、カメラをマーカ平面に対して斜めの方向から撮影した画像列に対して本手法を適用し、ARToolkitのアルゴリズムを用いて全てのパラメータを推定した場合と、本手法によって回転パラメータを推定した場合の比較を行う。1章でも述べたように、このような画像列に対しては、ARToolkitを用いて各パラメータが比較的精度良く推定できることが知られている。従って、まずこのような画像列に対して本手法を適用することで、本手法による回転パラメータの推定が正しく行われているかを示す。

次に、図9のように、カメラの光軸がマーカ平面に対して垂直になるように撮影した画像列に対して、同様に本手法とARToolkitを適用してその結果を比較する。こ

のような画像列に対しては、ARToolkitを用いるだけではマーカ平面に垂直なZ軸の追跡が不安定になるため、本手法によってその追跡精度が向上することを示す。

実験結果については、本研究会発表当日に紹介する。

6. ま と め

本論文では、2次元矩形マーカを用いたカメラトラッキングにおいて、時系列フィルタであるパーティクルフィルタを利用することで、カメラの3次元的な姿勢を表すパラメータの推定精度を向上させるための手法を提案した。本手法を用いることで、マーカ平面とカメラが正対したときに起こる、マーカ平面に対する垂直方向の軸の定義の不安定さを解消し、安定した3次元カメラトラッキングを実現する。今後の課題として、パターンマッチングを行う際にモデルパターンをぼかすボケの量を、入力画像に合わせて適宜変化させ、より安定したマッチングを行うことを検討する。

文 献

- [1] M. Billinghurst and H. Kato: "Collaborative mixed reality", Proc. of ISMR, pp. 261-284 (1999).
- [2] 岡兼司, 佐藤洋一, 中西泰人, 小池英樹: "適応的拡散制御を伴うパーティクルフィルタを用いた頭部姿勢推定システム", 電子情報通信学会論文誌, 8, pp. 1601-1613 (2005).
- [3] D. Marimon, et al.: "Particle filter-based camera tracker fusing marker and feature point cues", Proc. of IS&T/SPIE Conf. on Visual Communications and Image Processing (2007).