

ビジョンベース MR トラッキングのための エピポーラ拘束を利用した特徴点優先順位決定方法

Determining Reliability of Feature Points Using Epipolar Constraint for Vision Based MR Tracking

豊原由規¹⁾, 森尚平²⁾, 柴田史久¹⁾, 木村朝子¹⁾, 田村秀行¹⁾
Yuki Toyohara, Shohei Mori, Fumihisa Shibata, Asako Kimura, and Hideyuki Tamura

1) 立命館大学 大学院情報理工学研究科 (〒525-8577 滋賀県草津市野路東 1-1-1)

2) 立命館大学 大学院理工学研究科 (〒525-8577 滋賀県草津市野路東 1-1-1)

Abstract: This paper presents a method of determining reliability of feature points for vision-based mixed reality (MR) camera tracking using epipolar constraint of a stereo-camera. While robust estimation methods such as random sample consensus (RANSAC) algorithm are known as effective methods of eliminating outliers of feature points, these may possibly cause jittering of superimposed virtual objects. In this paper, epipolar constraint of a stereo-camera is used to determine reliability of feature points for camera pose estimation scheme such as perspective-n-point (PnP) problem to present stable results without jittering. The experimental results showed this algorithm effective and efficient.

Key Words: Mixed Reality, Vision-Based Tracking, Computational Stereo Vision

1. はじめに

現実世界と仮想世界を継ぎ目なく融合する複合現実感 (Mixed Reality; MR) は、従来の人工現実感 (Virtual Reality; VR) の発展形であり、本学会の主要研究分野の 1 つとして進化している。我々は、この MR 技術を映画制作のプリプロダクションで用いる PreViz (Pre-Visualization, 事前可視化) に活用し、映画人の創造性発揮を支援する MR-PreViz を提唱し、その研究開発を推進してきた[1]。また、近年の S3D (Stereoscopic 3D, 両眼立体視) 映画ブームに伴い、MR-PreViz に映画撮影用ステレオカメラ (ステレオリグ, 図 1) を導入する、MR-PreViz システムの S3D 拡張版である S3D MR-PreViz の開発にも着手している[2]。この S3D 拡張では、単なるシステムの機能追加に留まらず、MR-PreViz で用いられてきたビジョンベース・トラッキング法にステレオリグのエピポーラ拘束を積極利用することでの性能向上も進めている[2]。

S3D MR-PreViz では、SfM による特徴点の 3 次元位置推定とその画像平面での 2 次元位置を対応づけることで Perspective-n-Point (PnP) 問題[3]を解き、カメラ位置姿勢を推定する。この時、特徴点の 3 次元位置の推定やその画像平面上での位置とのマッチングの誤差により外れ値を含むことがある。こうした外れ値は、カメラ位置姿勢推定時に悪影響を与えるため、PnP 問題を解く上で排除されることが望ましい。こうした、外れ値を排除する手法としてロバスト推定法の一つである Random Sample Con-

sensus (RANSAC) 法[4]が有効であることが知られている。しかし、選択される特徴点にばらつきが発生し得るため、用いる特徴点数が統計処理するには不十分な場合、重畳する仮想物体が揺れて見える、ジッタが発生することがある。

そこで、本論文では、この仮想物体のジッタを軽減させるために、ステレオリグのエピポーラ拘束を利用し、左右のカメラからの画像で別々に追跡した特徴点の追跡誤差を活用して、特徴点にスコアを付け、カメラ位置姿勢推定に用いる特徴点の優先順位決定方法を提案する。

2. 提案手法

2.1 アルゴリズムの概要

本手法では、RANSAC 法のように多くの特徴点を用いて統計的に尤もらしいカメラ位置姿勢を計算することはせず、継続的に追跡した特徴点の追跡誤差を、エピポーラ



図 1 基線長と輻射角を変更可能な映画撮影用ステレオリグ (左: ビームスプリッターリグ, 右: 平行リグ)

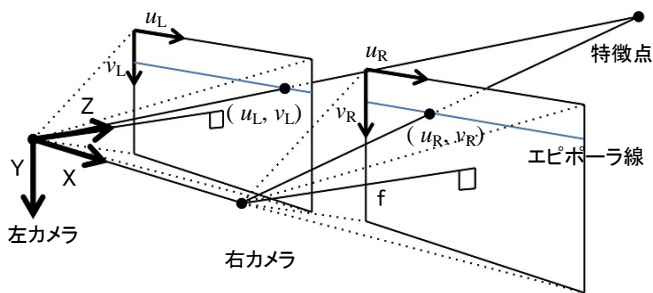


図2 カメラ座標系と特徴点の関係

拘束を満たすかどうかにより、用いるべき特徴点の信頼度を計算する。エピポーラ拘束を満たすかどうかの確認はステレオマッチングで行う。よって、利用するステレオカメラにおいて、高精度なステレオマッチングを可能にするため、また、問題の簡単化のため、基線長はシーンに応じて変更可能であると、輻輳角は0°に設定可能とする。これらに加え、左右のカメラは事前にキャリブレーションされ、平行化されているものとする。この条件における、カメラ座標系と特徴点の関係を図2に示す。

このとき、エピポーラ線は両カメラ座標系の u 軸方向に水平である。即ち、エピポーラ拘束より、左カメラからの画像（左画像）での特徴点の位置 $\mathbf{p}_L=[u_L, v_L]^T$ は右カメラからの画像（右画像）において v_L を通る u 軸方向に平行な直線上に存在する。よって、右画像上での対応点は $\mathbf{p}_R=[u_R, v_L]^T$ として表される。同様に、左画像と右画像とで別々に検出した特徴点 $\mathbf{p}_L, \mathbf{p}_R$ を別々に追跡した結果である $\mathbf{p}'_L=[u'_L, v'_L]^T, \mathbf{p}'_R=[u'_R, v'_R]^T$ もエピポーラ拘束を満たす。しかしながら、実画像上での追跡結果は、以下に例を挙げるような場合にエピポーラ拘束を満たす位置に存在しないことがある。

- ・検出した特徴点がコーナーでなくエッジやフラット
- ・カメラ位置や方向によって特徴点周りの見え方が変わる

特徴点ベースのカメラ位置姿勢推定法においては、このように誤って検出及び追跡される特徴点は悪影響を及ぼすため、可能な限り利用しないことが望ましい。よって、本手法では、継続的に特徴点を追跡した結果と先のエピポーラ拘束を利用して、各特徴点にスコア付けを行い、カメラ位置姿勢推定時に利用する特徴点の優先順位を決定する。図3に提案手法のアルゴリズムのフロー図を示す。後続の節では、このフローに則って詳細を述べる。

2.2 特徴点優先順位決定方法

初期フレームとそれ以外のフレームで処理が異なるため、フレーム番号を $i (0, 1, 2, \dots, n)$ として、アルゴリズムの説明を行う。

2.2.1 初期化処理 ($i=0$)

まず、入力されたステレオ画像に対して、事前にキャリブレーションされた結果を用いて平行化と歪み除去を行う (1-1)。次に、高速かつ再現性の高いとされる Harris オペレータ [5] により左画像における特徴点 \mathbf{p}_L を検出する (1-2)。続いて、ステレオマッチングを行う (1-3)。右画像での \mathbf{p}_L を起点とし、エピポーラ線上を u 軸の負の方向に

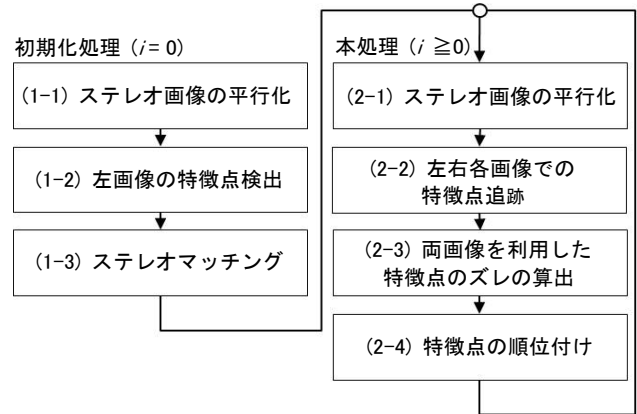


図3 提案手法の流れ

テンプレートマッチングを行って探索し、類似度が最大となる位置 \mathbf{p}_R を探索する。左右画像間での輝度値の違いを吸収するため、本手法では、類似度として正規化相互相関 (Normalized Cross Correlation; NCC) を採用する。また、エピポーラ線上での探索範囲 w はシーンによって変化するためユーザが任意に決定するものとする。検出された全特徴点に対して行われるこのステレオマッチングは、高速化とロバスト性向上を目的として、ピラミッドレベル lv の画像ピラミッドを用いて行う。

2.2.2 本処理 ($i \geq 0$)

初期化処理と同様に、入力されたステレオ画像に対して平行化と歪み除去を行う (2-1)。次に、初期化処理にて左画像上で検出した特徴点 \mathbf{p}_L と右画像上での対応点 \mathbf{p}_R を、左画像と右画像にて別々に追跡する (2-2)。追跡手法には KLT (Kanade-Lucas-Tomasi) トラッカ [6] を採用する。これは、KLT トラッカによって得られる追跡結果 $\mathbf{p}'_L, \mathbf{p}'_R$ が、初期化処理で用いた Harris オペレータによって検出される特徴点と等価であるためと、KLT トラッカは高速かつロバストな手法として知られているため、左右の画像で別々に追跡処理を必要とする本手法に適しているからである。

続いて、エピポーラ拘束を利用して、追跡結果の誤差 $\boldsymbol{\varepsilon}=[\varepsilon_u, \varepsilon_v]^T$ を求める (2-3)。まず、左右画像上での追跡結果の正誤に関して、以下のように仮定を立てることができる。**【追跡結果が正しい場合】** 追跡結果はエピポーラ拘束を満たす ($v'_L=v'_R$, そして、左画像-右画像と右画像-左画像のステレオマッチングの結果が常に正しいとし、それぞれを $\mathbf{p}'_{L-R}=[u'_{L-R}, v'_{L-R}]^T, \mathbf{p}'_{R-L}=[u'_{R-L}, v'_{R-L}]^T$ とした時、 $u'_L=u'_{R-L}$ かつ $u'_R=u'_{L-R}$ 。

【追跡結果が誤っている場合】 追跡結果はエピポーラ拘束を満たさない ($v'_L \neq v'_R, u'_L \neq u'_{R-L}, u'_R \neq u'_{L-R}$)。

追跡結果が誤っている場合において、左画像-右画像と右画像-左画像のステレオマッチングのそれぞれの結果で得られる最大の類似度を S_{L-R}, S_{R-L} としたとき、誤差 $\boldsymbol{\varepsilon}$ を以下の式で定義する。尚、図4は、左画像と右画像上での各パラメータの関係を示している。

ただし、左画像-右画像と右画像-左画像のステレオマッチングを行った際に、両結果において対応点が得られない

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_u \\ |v'_L - v'_R| \end{bmatrix} \quad (1)$$

$$\varepsilon_u = |u'_L - u'_{R-L}| \text{ if } S_{L-R} \leq S_{R-L} \text{ otherwise } |u'_R - u'_{L-R}|$$

場合がある。この時は、誤差を求めないものとする。尚、これまでにこの誤差の計算を行った回数を参照回数 ref として保存しておき、後続の処理に用いる。

これまでに算出した誤差の平均と分散を算出する (2-4)。ある特徴点の誤差 ε_j ($j = 0, 1, 2, \dots, ref$) の u , v 軸方向それぞれの平均と分散である, m_u , m_v , δ_u , δ_v は以下の式で表される。尚、この計算は n フレーム毎に行う。

$$m_u = \frac{1}{ref} \sum_j \varepsilon_{u,j}, \quad m_v = \frac{1}{ref} \sum_j \varepsilon_{v,j} \quad (2)$$

$$\delta_u = \sqrt{\frac{1}{ref} \sum_j (m_u - \varepsilon_{u,j})^2}, \quad \delta_v = \sqrt{\frac{1}{ref} \sum_j (m_v - \varepsilon_{v,j})^2} \quad (3)$$

ある特徴点に関して得られたこの誤差の平均と分散を用いて、以下の式に示すように優先順位を決定するためのスコアを算出する。

$$Score = \frac{1}{\sqrt{(m_u + \delta_u)^2 + (m_v + \delta_v)^2}} \quad (4)$$

この式では u 軸, v 軸方向のズレの平均と分散を足し合わせ、ユークリッド距離を計算し、その逆数を取っている。そして、この式で求められたスコアが高い特徴点は継続的に正確に追跡されていると判断する。

3. 実験

3.1 実験目的と実験方法

提案手法の有効性を実証するために、以下に詳細を述べる。カメラ位置姿勢を推定する簡単なプログラムを実装し、提案手法、RANSAC 法、ランダムに特徴点を選択する方法 (以下、ランダム法) との比較を行う。その際、RANSAC 法やランダム法と比較して本手法において仮想物体の重畳時のジッタが軽減されたかどうかを確認する。

比較方法は、仮想物体 (ユタ・ティーポット) を重畳することによる目視での確認だけでなく、レールに載せたステレオリグを規定の距離だけ移動させたときのその距離を真値として、各方法で PnP を解くことで得られるカメラ位置を比較する (図 5)。以下は、本実験で用いるカメラ位置姿勢推定を行うプログラムの処理手順である。

1. Harris オペレータで特徴点を検出 (検出閾値 : 0.01, 特徴点間の距離 : 15 [pixels])
2. 検出した特徴点を 2.2.1 項で述べたステレオマッチングにより 3 次元位置推定
3. 提案手法、RANSAC 法、ランダム法で特徴点を選択
4. 選択した特徴点の 2D-3D の対応関係から PnP[3] を解き、カメラ位置姿勢を計算 (提案手法とランダム法では 6 点の特徴点を用い、RANSAC 法ではそのフレームで得

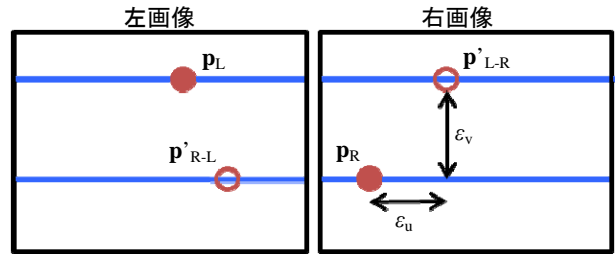


図 4 エピポーラ拘束と追跡誤差

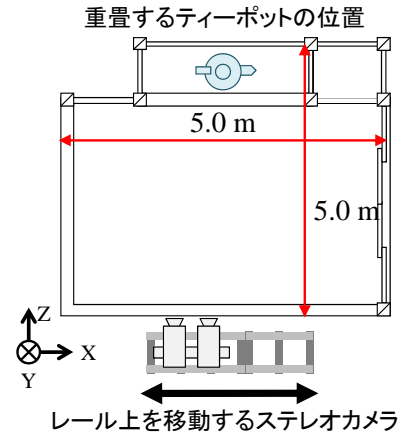


図 5 実験シーンでの機材配置

られたコンセンサス・セットを用いる。また、提案手法のスコアの更新は 5 フレーム毎 ($n = 5$) に行う)

5. KLT トラッカで特徴点を追跡し、3, 4 の処理を繰返す
上記の条件で、以下の 2 つの実験を行う。

【実験 1】レール上のステレオリグを X 軸方向に 2m 移動させ、先の 3 つの手法を比較する

【実験 2】レール上のステレオリグが止まった状態から X 軸方向に 1m 移動させ、しばらく止めてから、X 軸方向に -1m 移動させ、提案手法と RANSAC 法の比較を行う

3.2 実験 1 の結果と考察

提案手法による仮想物体の重畳結果を図 6 に示す。図 6 より、違和感なく、またジッタが少ない状態で MR 合成を行うことができていることが分かる。この時に推定された移動量は、提案手法、RANSAC 法、ランダム法それぞれで 2.193m, 2.113m, 2.159m と、0.100m 程度の差しか見られず、ほぼ横並びであった。図 7 に示すのは、3 手法での移動の推移を表すグラフである。ランダム法の移動推移の結果から分かる通り、フレームが進むにつれて、特徴点の誤追跡による特徴点の 2 次元位置と 3 次元位置の対応関係が取れていない特徴点が登場する。その特徴点をカメラ位置姿勢の計算に使用してしまうため、カメラ位置姿勢の誤差が大きくなっていくことが分かる。しかし、提案手法では、フレームが進んでも、そのような結果は得られず、RANSAC 法と同様に安定した結果となった。これらの結果から、提案手法では、フレームが進むにつれて、安定して追跡できる特徴点とそうでない特徴点とを上手く選別できていることが分かる。



図 6 MR 合成結果

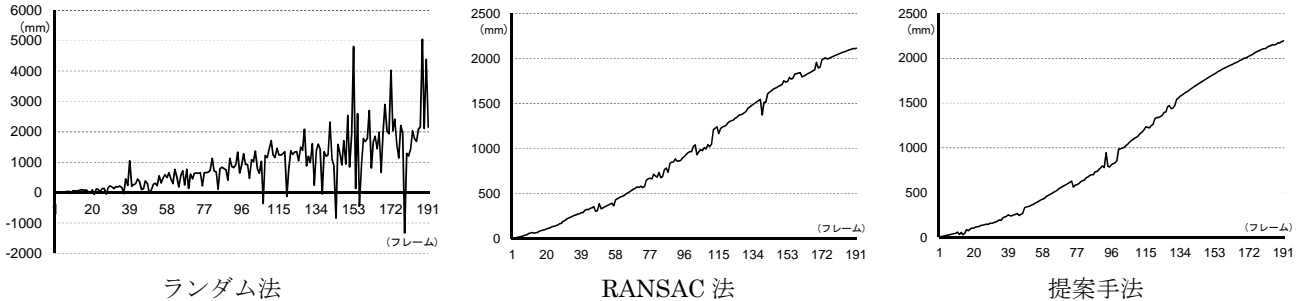


図 7 実験 1 におけるステレオリグの移動の推移

3.3 実験 2 の結果

図 8 に示すのは、提案手法と RANSAC 法での移動推移の結果を表したグラフである。1 - 50 フレームと 200 - 350 フレームにかけて、カメラが静止しているが、RANSAC 法では、後者において、大きなところでは 0.200m 程度のジッタが発生してしまっている。一方で、提案手法では、そのようなジッタが軽減され、フレームが進んでも安定した結果が得られていることが分かる。

最初の静止状態においては、RANSAC 法、提案手法ともに、ほとんどジッタが発生しなかった。これは特徴点の見え方が変化しないため、初期フレームでの特徴点 2 次元座標と 3 次元座標の対応関係に誤差が生じていないためである。一方で、提案手法は信頼できる特徴点の選別に、見え方の違う数フレームを与える必要があるため、その間では特徴点の評価結果が入れ替わる状態にあるため、50 - 100 フレームでの挙動のようにジッタを生じさせてしまう。しかし、移動していくにつれ、誤差の平均値と分散値が収束していくため、それ以降はジッタが目に見えて生じることはなかった。

4. むすび

本論文では、特徴点を用いたビジョンベースの MR トラッキング法において使用する特徴点に優先順位を付けることで、選択する特徴点の切り替わりを防ぎ、重畳する仮想物体のジッタを軽減する手法を提案した。提案手法では、ステレオカメラのエピポーラ拘束を利用して、特徴点の追跡誤差を計算し、スコアを算出することで優先順位を付与した。実験では、MR 合成による目視での確認と、RANSAC 法を含む他の手法と本手法の比較評価を行い、本手法の有効性を確認した。具体的には、RANSAC 法と同等の精度を得ながらも、ジッタの軽減が可能であった。

本論文では、カメラ位置姿勢推定法や特徴点追跡法等を具体的に示したが、ホモグラフィ行列の推定等、同様の問題に対しても本手法が有効に機能すると考えられる。

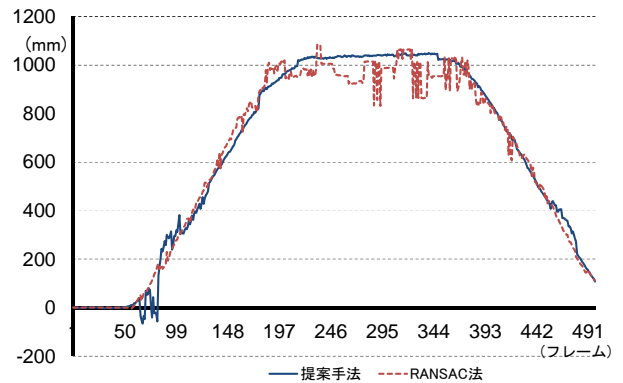


図 8 実験 2 におけるステレオリグの移動の推移

参考文献

- [1] 田村秀行, 柴田史久: “可視化技術で創造力を高める映画制作支援”, 情報処理, Vol. 48, No. 12, pp. 1365 - 1372, 2007.
- [2] S. Mori, R. Ichikari, F. Shibata, A. Kimura, and H. Tamura: “Enabling on-set stereoscopic MR-based previsualization for 3D filmmaking,” Proc. SIGGRAPH Asia, 2011.
- [3] F. Moreno-Noguer, V. Lepetit, and P. Fua: “Accurate non-iterative $O(n)$ solution to the PnP problem,” Proc ICCV '07, pp. 1 - 8, 2007.
- [4] M. A. Fischler, and R. C. Bolles: “A paradigm for model fitting with applications to image analysis and automated cartography,” Comm. of the ACM, Vol. 24, pp. 381 - 395, 1981.
- [5] C. Harris and M. Stephens: “A combined corner and edge detector,” Proc. Alvey Vision Conf., pp. 147 - 151, 1988.
- [6] C. Tomasi and T. Kanade, “Detection and Tracking of Point Features,” Carnegie Mellon University Technical Report, CMU-CS-91-132, Apr. 1991.