

入力映像とパノラマ画像の射影変換を用いた 位置合わせパラメータの高速推定手法

— パノラマベースド注釈提示手法の拡張として —

興梠 正克 蔵田 武志 坂上 勝彦 村岡 洋一

早稲田大学大学院 理工学研究科

〒 169-8555 東京都新宿区大久保 3-4-1

TEL:03-3209-5198

E-mail:kourogimuraoka@waseda.ac.jp

本研究の目的は、利用者視点のライブ映像上に実環境に関する注釈情報を重ね合わせ表示することである。筆者らは、入力映像とパノラマ画像の位置合わせにより、パノラマ画像上に配置された注釈情報を入力映像へマップするアプローチを提案し、アフィン変換による近似的な位置合わせパラメータの高速推定手法を示した。しかしながら、アフィン変換による近似には限界があり、利用可能な局面が制限される問題点があった。そこで、本研究では、位置合わせ手法を改良して、射影変換による位置合わせパラメータの高速推定を可能とする。提案手法は、アフィン変換では近似が難しい条件下でも安定して位置合わせが実現でき、改良前の手法と比べて計算コストが10-20%しか増加せず、ソフトウェアによる実時間処理を実現できることを実験的に検証した。

注釈付き映像, パノラマ画像, 射影変換, 位置合わせ, 実時間処理

Improved method of image registration between frames and panoramic images using projective transform

— for panorama-based annotation overlay —

Masakatsu Kouroggi Takeshi Kurata Katsuhiko Sakaue Yoichi Muraoka

School of Science and Engineering, Waseda University

3-4-1 Ohkubo Shinjuku-ku, Tokyo 169-8555, JAPAN.

In this paper, we propose an improved method of annotation overlay on live video which is one of the most promising applications in augmented reality and wearable computer. In our previous work, we proposed the panorama-based method of annotation overlay which aligns an input video frame and a panoramic image using affine transform. However, affine transform is generally not capable of image registration between a frame and a panorama. We improved the previous method so that it can estimate projective transform parameters without severely increasing computational cost. We confirmed that improved method could estimate image registration parameters under conditions that hindered the previous method. Its computational cost increased only by 10-20% and its real-time processing with software implementation was achieved.

annotation overlay, panoramic image, projective transform, image registration, real-time processing

1 はじめに

本研究の目的は、利用者視点から得られるライブ映像上に実環境に関する注釈情報を重ね合わせ提示することである。これは、着用型計算機(ウェアラブルコンピュータ)の最も有力な応用例の一つであり、拡張記憶や道案内、アミューズメントなど様々な応用分野に対して有効な拡張現実感の要素技術の一つである [1][2][3]。

筆者らはパノラマ画像群を情報源として用いるパノラマベースド注釈提示手法を提案し、高速な画像間の位置合わせ手法を用いて、ライブ映像上への注釈提示の実時間処理を実現してきた [4][5]。提案手法は、実環境中の複数地点で撮影されたパノラマ画像群とその画像上に配置された注釈情報を事前知識として用いる。入力画像とパノラマ画像の位置合わせ結果に基づいて、ライブ映像上の適切な位置に注釈情報を重ね合わせ提示する。図 1 にその概要を示す。提案手法は、画像間の位置合わせにはアフィン変換による高速位置合わせパラメータ推定手法 [7] を用いている。

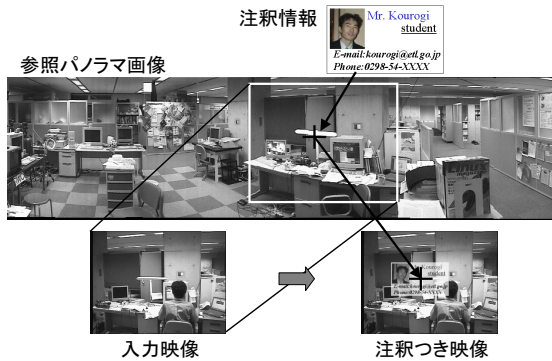


図 1: パノラマベースド注釈提示手法の概要

しかしながら、一般的には、円筒面に投影されているパノラマ画像とフレーム画像平面に投影されている入力映像をアフィン変換で位置合わせすることはできない。そのため、アフィン変換により近似的に位置合わせできる範囲でしか、正しく動作しない問題点があった。実験では、利用者の視点が水平方向から概ね ± 30 度以内でしか、安定して動作しないことが確認されている。本研究では、射影変換を用いた位置合わせに基づく映像上への注釈提示を実現する。しかしながら、射影変換パラメータ推定は非線形な最小化問題となるため、解が不安定となる。そこで、

本研究では、アフィン変換を用いてパノラマ画像と位置合わせできる円筒面にフレーム画像を再投影する。そのためには、フレーム画像の中心点がパノラマ画像上のどの点と対応するか仮定を立てる。この仮定は、初期状態においては、パノラマ画像全体を一定間隔で網羅する点の集合となる。それ以降については、前回の推定結果とその周辺を局部的に網羅する点の集合となる。次に、円筒面に再投影されたフレーム画像とパノラマ画像をアフィン変換を用いて位置合わせする。各仮定に基づいて得られた結果から、画像間の輝度差分の二乗平均を最小とするものを選び、これを最終的な推定結果とする。

本論文の構成は以下の通りである。2節では、射影変換を用いてフレーム画像とパノラマ画像の位置合わせする手法について述べ、3節で提案手法の評価実験を行い、位置合わせの精度と処理の計算コストについて評価する。4節ではまとめを述べる。

2 画像間の射影変換パラメータ推定手法

アフィン変換を用いたパノラマ画像とフレーム画像の近似的な位置合わせには限界がある。そこで円筒面に投影されたパノラマ画像とアフィン変換を用いて位置合わせ可能にするために、フレーム画像を円筒面に再投影する。フレーム画像の中心点に対応するパノラマ画像上の点の位置について仮定を立て、各仮定より得られる推定結果のうち、もっとも良好な位置合わせを与えるものを選ぶ。

2.1 フレーム画像の円筒面への再投影

フレーム画像は、その撮像系の光軸の仰角が既知であれば、円筒面へと再投影することができる。そこで、パノラマ画像との位置合わせに先だって、フレーム画像の中心点(光軸とフレーム画像平面の交点)がパノラマ画像上で対応する点について仮定を立てる。対応する点の位置から、パノラマ画像の撮影条件を用いてフレーム画像撮像系の光軸の仰角を求めることができる。なお、議論の簡単化のため、光軸周りの回転はないものとする。パノラマ画像の焦点距離を f_P 、対応点の座標を (x_P, y_P) 、光軸の仰角を α とおくと、

$$\alpha = \tan^{-1} \left(\frac{y_P}{f_P} \right) \quad (1)$$

が成り立つ。フレーム画像上の各点 (x, y) を仰角だけ X 軸方向に光軸を回転して射影された点を (x', y') とし、フレーム画像の撮像系の焦点距離を f_c とすると、以下の式が成り立つ。

$$x' = \frac{f_c x}{y \sin \alpha + f_c \cos \alpha} \quad (2)$$

$$y' = \frac{f_c (y \cos \alpha - f_c \sin \alpha)}{y \sin \alpha + f_c \cos \alpha} \quad (3)$$

次に、仰角 α だけ回転後のフレーム平面画像上の点を (x', y') とし、この点在同一の焦点で焦点距離 $f = f_c$ の円筒面へ投影される点を (x'', y'') とすると、以下の式を満たす。

$$x'' = f_c \tan^{-1} \frac{x'}{f_c} \quad (4)$$

$$y'' = f_c \cos \left(\tan^{-1} \frac{y'}{f_c} \right) \quad (5)$$

中心点に対応するパノラマ画像上の点の位置（仮定）の y 座標値が十分に真値に近ければ、焦点を共有する円筒面間の位置合わせはアフィン変換で実現可能であるため、円筒面に再投影されたフレーム画像とパノラマ画像はアフィン変換により位置合わせできると考えられる。図2にフレーム画像の円筒面への再投影の概要を示す。

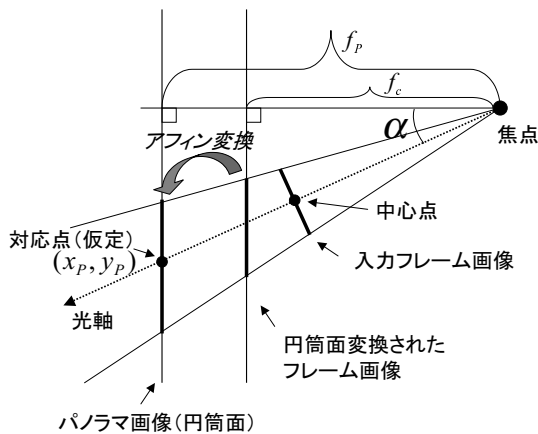


図2: 円筒面への再投影

2.2 位置合わせパラメータ推定

初期状態では、フレーム画像とパノラマ画像の対応付けに関する事前知識を持たない。そこで、提案手

法は、パノラマ画像全体を一定間隔で網羅するように対応点を仮定を初期条件として与える。各初期条件に基づいて推定された位置合わせパラメータのうち、最も良好な位置合わせ結果を与えるものを、最終的な位置合わせ結果として出力する。以降の処理については、前回の推定結果によりフレーム画像中心点がパノラマ画像上で対応する点およびその周辺4点を初期条件として与えて、円筒面への再投影とアフィン変換による位置合わせ計算を行う。なお、位置合わせ結果の評価基準として、画像間の輝度差分の二乗平均を用いる。

2.3 光軸周りの回転補正

円筒面の画像間の位置合わせは、原理的にはスケールと平行移動により実現可能である。しかし、フレーム画像の光軸周りの回転のため、推定されたアフィン変換パラメータには回転成分が現れる。この光軸周りの回転成分が大きいと、円筒面間をアフィン変換で正しく位置合わせすることはできない。そこで、アフィン変換による位置合わせ結果に現れる回転成分は、次回の推定では、フレーム画像を円筒面に変換する際に、光軸周りの回転として事前に補正する。

円筒面に射影されたフレーム画像の点を (x''_c, y''_c) 、対応するパノラマ画像上の点を (x_p, y_p) 、得られたアフィン変換パラメータを a_{ij} とおくと、以下の式が成り立つ。

$$\begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x''_c \\ y''_c \\ 1 \end{bmatrix} \quad (6)$$

ここで、フレーム画像の円筒面とパノラマ画像の円筒面間のスケール係数を (s_x, s_y) 、平行移動成分を (t_x, t_y) とすると、 $(a_{11}, a_{22}) = (s_x, s_y)$ 、 $(a_{13}, a_{23}) = (t_x, t_y)$ が成り立ち、回転成分 $(r_x, r_y) = (a_{12}, a_{21})$ である。この回転成分を次回の位置合わせ計算の前処理で補正する。したがって、第 $n+1$ 回目の位置合わせ計算で補正される回転成分は、第 n 回目の補正に用いられた回転成分と位置合わせ結果から得られる補正回転成分の積となる。

2.4 提案手法の処理の流れ

提案手法の処理の流れは以下の通りである。

1. フレーム画像の中心点に対応するパノラマ画像上の点（複数個）を仮定として選択する。

2. (初期状態でなければ) 前回の推定結果と前回の補正回転成分に基づいて, 光軸周りの回転成分を補正する.
3. パノラマ画像の撮影条件から, フレーム画像の水平方向に対する仰角を算出する.
4. 算出された仰角に基づいてフレーム画像を焦点距離 $f = f_c$ の円筒面に射影変換する.
5. 円筒面に射影されたフレーム画像とパノラマ画像(焦点距離 $f = f_p$ の円筒面)とアフィン変換を用いた位置合わせ計算を行う.
6. 推定されたアフィン位置合わせパラメータを用いて, フレーム画像中心点のパノラマ画像上での対応点を計算し, この対応点とその周辺4点を次回の位置合わせ計算における仮定として用いる.
7. ステップ1で与えられた複数個の仮定に基づいて推定された各位置合わせパラメータから, 画像間の輝度差分の二乗平均を最小とするものを選び, これを最終的な位置合わせ結果として出力する.

3 評価実験

2節で述べた, 射影変換による位置合わせ手法を評価する. 画像間の正しい位置合わせ結果がすでに得られているフレーム画像列とパノラマ画像を実験に用いて, 提案手法による画像間の位置合わせ精度を評価して, 改良前のアフィン変換による位置合わせ手法による精度との比較評価を行う.

3.1 実験条件

利用者視点で撮影されたフレーム画像列 100 枚(每秒 15 フレーム, 画像サイズ: 320×240 画素)に対して, 情報源となる 1 枚のパノラマ画像(画像サイズ: 2560×480 画素)を用いて評価実験を行う. パノラマ画像上で特徴が著しい点(人間が視認により確実に識別できるもの)を取り, これらの点が各フレーム画像上で対応する点を人手により選び出して正解データとして作成した. 各フレーム画像中には少なくとも 5 点の正解データが含まれている. 情報源として用いるパノラマ画像は, 全方位撮像系 HyperOmni Vision(アコウル社製, 仰角 15 度 ~ -45 度まで撮影可

能)[10]で撮影した全方位画像を円筒面に展開して生成した. 全方位画像の撮影には, デジタルカメラ(ソニー社製, DSC-F505V, 334 万画素)を用いた. 生成されたパノラマ画像を図3に示す. 入力映像の撮影には, 超小型 CCD カメラ(東芝製, IK-SM43H, レンズ $f = 4mm$)を用いた. 提案手法は, PC 上のソフトウェアとして実装され, IBM-PC 互換機(CPU: Intel PentiumII-450MHz, OS: Linux-2.2.14)上で実行, 評価される.



図3: 評価実験に用いたパノラマ画像

3.2 射影変換を用いた位置合わせ結果

提案手法による位置合わせ精度を正解データを用いて評価した. 各フレーム画像に対するアフィン変換および提案手法による位置合わせ結果の誤差グラフを図4に示す. 本実験におけるフレーム画像の仰角 α の推定結果のグラフを図5に示す.

これらの結果より, ほとんどすべての状況下において, 提案手法による位置合わせは, アフィン変換による位置合わせ手法よりも高い精度を実現できることが確認され, 5-10 ピクセル程度の位置合わせ結果の精度向上が見られた. また, 仰角が ± 30 度の範囲を越えると, アフィン変換による位置合わせは不安定となったが, 提案手法による位置合わせは安定して良好な位置合わせ結果を得ることができた.

また, アフィン変換による位置合わせ結果と提案手法による位置合わせ結果の一例を, それぞれ図6と図7に示した. 図中で示されている枠は, フレーム画像の4隅の点がパノラマ画像上で対応しているものである. この結果より, アフィン変換では正しく位置合わせができない条件下においても, 提案手法では正しく位置合わせできていることが読みとれる.

提案手法による位置合わせ結果に基づいて生成された注釈付き映像の一例を図8に示す.

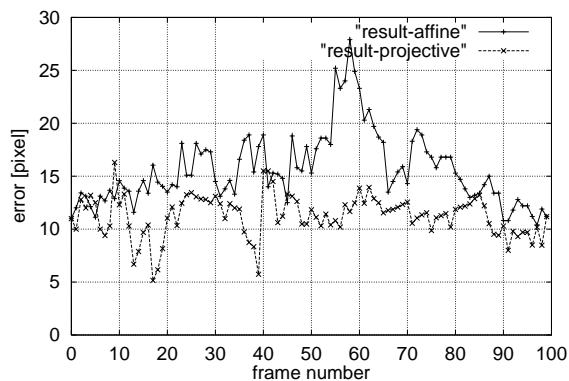


図 4: 位置合わせ結果の誤差

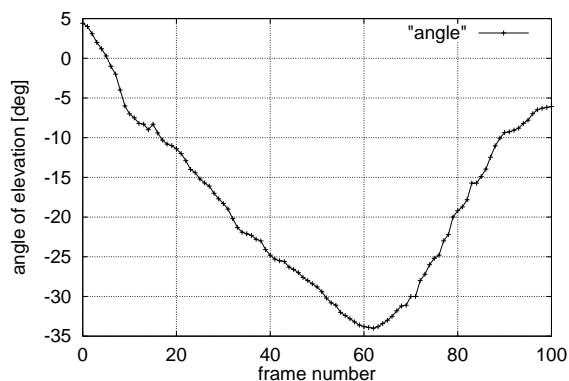


図 5: フレーム画像の仰角

しかしながら、情報源となるパノラマ画像と入力映像の撮影時における照明条件が大きく異なる場合、提案手法は正しく位置合わせ結果を与えることができない場合があった。その原因は、複数個の各仮定から得られる推定結果から最良のものを選び出す過程が照明条件の変化に対して弱いことにある。選別の評価基準として用いている輝度差分の二乗平均は照明条件の局所的な変化に対してロバストではなく、良好な位置合わせ結果に低い評点を与えることがあったためである。村瀬ら [9] の提案している増分符号相関など、照明条件の変化に対してロバストな手法を取り入れる必要があると考えられる。

3.3 処理の計算コスト

前述した計算機環境において、提案手法が一つの初期条件から射影変換パラメータを推定するまでに要した計算時間は、50-100 ミリ秒（毎秒 10-20 フレーム）であった。



図 6: アフィン変換を用いた画像間位置合わせの一例



図 7: 提案手法による画像間位置合わせの一例

ム）であった。提案手法の処理に要する計算時間の内訳を表 1 に示す。これより、アフィン変換による位置合わせと比べて、処理時間は 10-20% しか増大せず、容易に実時間処理を達成できることが確認された。

4 まとめ

本稿では、パノラマ画像群を情報源として用いるパノラマベースドアノテーション手法の拡張として、



図 8: 提案手法による注釈付き映像の一例

	処理時間
円筒面への再投影	5-10 ミリ秒
位置合わせ計算 (アフィン変換)	40-80 ミリ秒
輝度差分の二乗平均の算出	5-10 ミリ秒
全処理の合計	50-100 ミリ秒

表 1: 本手法の処理時間

射影面の異なるフレーム画像とパノラマ画像を、射影変換を用いて位置合わせする手法を提案した。提案手法は、フレーム画像を円筒面に再投影して円筒面間のアフィン変換を用いた位置合わせ計算により、画像間の射影変換による位置合わせを実現した。提案手法を評価して、その有効性および処理の実時間性について良好な結果を得た。

今後の課題としては、照明条件の変化に対する提案手法のロバスト性が挙げられる。そのためには、照明条件の変化に影響を受けにくい位置合わせ結果の評価基準を模索する必要がある。また、提案手法は、輝度勾配に基づく位置合わせ計算を行うため、画像全体の輝度勾配が平坦な画像に対しては、正しい位置合わせ結果が得られない。ジャイロセンサーと組み合わせることで、射影変換による位置合わせパラメータ推定のロバスト性と精度を高めることができると期待される。

謝辞 本研究はリアルワールドコンピューティング (RWC) プログラムの一環として行われた。

参考文献

- [1] R. Azuma, "A survey of augmented reality," in *Presence: Teleoperators and Virtual Environments*, vol. 6, no. 4, pp. 355-385, 1997.
- [2] J. Rekimoto and K. Nagao, "The world through the computer: Computer augmented interaction with real world environments", in *Proc. of UIST '95*, pp. 29-36, 1995.
- [3] T. Starner, S. Mamm, B. Rhodes, J. Levine, J. Healy, D. Kirsh, R. Picard and A. Pentland, "Augmented reality through wearable computing," in *Presence: Teleoperators and Virtual Environments*, vol. 6, no. 4, pp. 386-398, August

1997.

- [4] 興枙, 蔵田, 坂上, 村岡, "パノラマ画像群を位置合わせに用いたライブ映像上への注釈提示とその実時間システム," 第5回知能情報メディアシンポジウム論文集, pp. 17-24, 1999.
- [5] M. Kourogi et al., "A real-time panorama-based technique for annotation overlay on video frames," in *Proc. of ICME'2000*, TA. 2.05, 2000.
- [6] 蔵田, 興枙, 栗田, 村岡, 坂上, "気の利いたウェアラブルビジョンシステムの実現にむけて: 顔検出・識別とハンドマウスの実装", 信学技報 PRMU99-270, pp. 65-72, 2000.
- [7] 興枙, 村岡, "グローバルなアフィン動きパラメータの実時間推定手法", 信学論 (D-II), vol. J82-D-II, No.7, pp. 1161-1170, 1999.
- [8] R. Szeliski, "Video mosaics for virtual environments," *IEEE Computer Graphics and Applications*, vol. 16, no. 2, pp. 22-30, 1996.
- [9] 村瀬, 金子, 五十嵐, "増分符号相関にロバスト画像照合", 信学論 (D-II), Vol. J83-D-II, No. 5, pp. 1323-1331, 2000.
- [10] 山澤, 八木, 谷内田, "移動ロボットのナビゲーションのための全方位視覚系 HyperOmni Vision の提案", 信学論 (D-II), Vol. J79-D-II, pp. 698-707, 1996.