

1. はじめに

3次元シーンフローとは3次元空間中の物体の動きを表したものであり、認識や3次元形状復元、シーン構造の時間的内挿等への応用が期待できる。従来、3次元シーンフローを推定するには、強カメラ校正を必要とする。しかし、強カメラ校正は各カメラに対して行うため、非常に手間のかかる作業である。そこで、本研究では未校正のマルチカメラを用いたシーンフローの推定を目的とする。

2. 未校正マルチカメラによる 3 次元空間の定義

強校正を行ったカメラでは、画像と実空間の直接的な関係を知ることができる。しかし、未校正(弱校正)のカメラでは、エピポラ幾何と呼ばれるカメラ間の相対的な関係しか知ることができない。本研究ではエピポラ幾何情報のみから定義される3次元空間を用いる。

エピポラ幾何 2台のカメラで3次元空間中の同じ点 M を観測すると、点 M と2台のカメラの両方の画像にその点を投影した点が1つの平面上にある。画像1の観測点 \tilde{m}_1 と画像2上の \tilde{m}_2 との対応点 \tilde{m}_2 は式(1)を満たす。ここで F を基礎行列、(1)式をエピポラ方程式という。

$$\tilde{m}_1^T F \tilde{m}_2 = 0 \quad (1)$$

Projective Grid Space エピポラ幾何を利用して定義される3次元空間としてProjective Grid Space(以下PGS)が提案されている。PGSは図1に示すように、2つのBase Viewの p 軸、 q 軸、 r 軸により定義される空間である。PGSでは空間中に格子が作成され、この各格子点は (p, q, r) として表される。

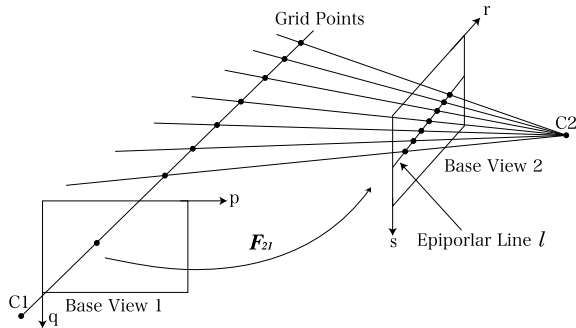


図1: Projective Grid Space

3. エピポラ幾何を利用したシーンフロー推定

3次元空間中の点 M にシーンフローが存在する場合、その点を投影した各カメラ画像上の点にオプティカルフローが存在するはずである。点 M をBase View 1に投影した点を m_1 とし、その点のフローの終点を m_1' とする。また、点 M をBase View 2に投影した点 m_2 にもフローが存在するはずである。この場合、図2に示すように、Base View 2における点 m_2 のフローの終点 m_2' は、点 m_1' におけるBase View 2上のエピポラ直線 l' 上に存在する。

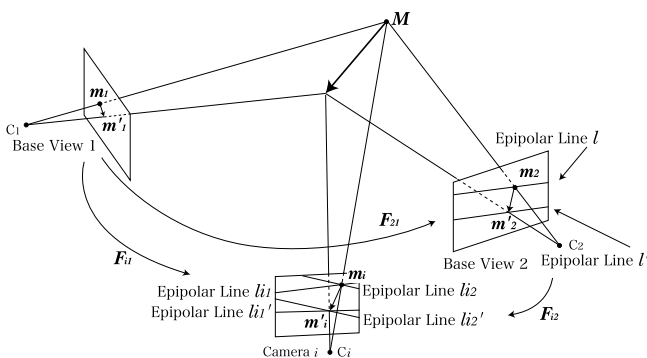


図2: 画像間のフローの一致性

任意のカメラ C_i の画像においては、2つのBase Viewとの間の基礎行列から求めることができる。点 M にシーンフローが存在する場合、カメラ C_i に投影されるフロー

の終点の位置は、Base View 1におけるフローの終点 m_1' と、Base View 1とカメラ i との間の基礎行列 F_{i1} から求められるエピポラ直線 l'_{i1} と、同様にBase View 2から求められるエピポラ直線 l'_{i2} の交点となる。上記のフローの一致性を用いた3次元シーンフロー推定アルゴリズムを以下に示す。

Step1 各入力カメラ画像に対し時刻 t と $t-1$ のオプティカルフローを計算

Step2 PGSのある格子点についてBase View 1とBase View 2においてオプティカルフローの一致性を計算

Step3 フローが一致する場合、任意のカメラ C_i においてもオプティカルフローの一致性を計算

Step4 全てのカメラにおいてオプティカルフローが一致した場合、シーンフローの大きさと方向を計算

PGSにおける全ての格子点においてStep2 ~ Step4を繰り返す。

4. シーンフロー推定実験

天井に設置された3台のカメラを用いて人物が両手を上に振り上げている映像を撮影し、この映像に対し3次元シーンフローを推定した。図3に各カメラ映像におけるフレーム t と $t-1$ 間のオプティカルフローを、図4に推定した3次元シーンフローを示す。

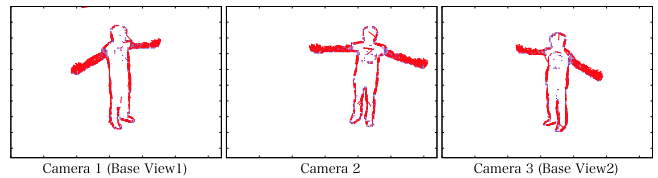


図3: 各カメラ画像のオプティカルフロー

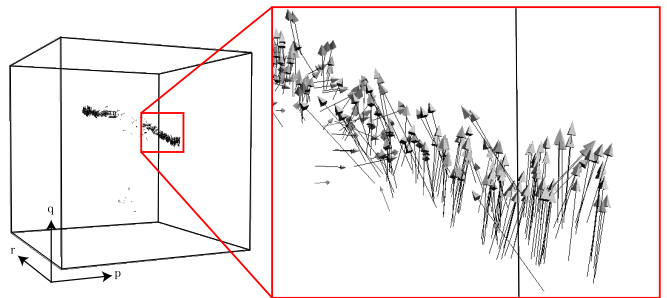


図4: 3次元シーンフロー

図4より、腕部のフローは上方向に向いている事がわかる。また腕の先の方が長いフローが存在していることから、求められた3次元シーンフローは正しいと言える。定量的な評価として、推定された全てのフローに対する方向の分布を表1に示す。表中の p, q, r はPGSにおける軸を、+, - はそれぞれの軸に対する正負の方向を表している。

表1: 3次元シーンフローの方向の分布 [%]

	$p+$	$p-$	$q+$	$q-$	$r+$	$r-$
2台	17.1	7.5	42.9	9.5	10.4	12.6
3台	13.3	4.5	62.1	0.5	12.3	7.2

表1より、カメラの台数が2台の場合より、3台の方が方向の分布が小さく、正しいフローが多く推定されている事がわかる。これは、カメラの台数を増やす事により、2次元オプティカルフロー算出のノイズ等の影響を減少させる事ができたためである。

5. おわりに

本研究では、未校正のマルチカメラを用いた3次元シーンフローの推定手法を提案した。提案手法は、未校正のカメラであるため、実空間でのフローの長さは測定できない。しかし、相対的なフローの長さや方向が測定できることを確認した。今後はジェスチャー認識や、シーン構造の時間的内挿等への応用を検討する予定である。