JAIST Repository

https://dspace.jaist.ac.jp/

Title	時空間の特徴を用いたステレオ動画像による物体の3次 元軌跡推定
Author(s)	三宅,章太郎
Citation	
Issue Date	1999-03
Туре	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1263
Rights	
Description	Supervisor:阿部 亨, 情報科学研究科, 修士



Japan Advanced Institute of Science and Technology

修士論文

時空間の特徴を用いたステレオ動画像

による物体の3次元軌跡推定

指導教官 阿部亨 助教授

北陸先端科学技術大学院大学 情報科学研究科情報処理学専攻

三宅章太郎

1994年2月15日

Copyright \bigodot 1999 by Shotaro MIYAKE

要旨

本稿では,時系列ステレオ画像から物体の三次元運動を推定するために,左右画像間での対応づけと,時間的に連続するフレーム画像間での対応づけとの統合について検討を行う.ここでは,空間・時間方向に広がりを持つ特徴窓を対応づけの単位とし,画像上の輝度勾配だけでなく輝度の時間変化も考慮に入れた特徴窓の設定箇所選択法,特徴窓のサイズを空間・時間方向に適応的に変化させる手法を提案する.

目 次

1	はじ	うめに しんしん しんしん しんしん しんしん しんしん しんしん しんしん しん	1
	1.1	研究の目的と背景・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	1
	1.2	本論文の構成	3
2	対応	「点問題とローカルサポート	4
	2.1	対応点問題	4
	2.2	ローカルサポート領域の拡張	5
	2.3	ステレオ視と軌跡推定	7
		2.3.1 ステレオ視	7
		2.3.2 軌跡推定	11
	2.4	3次元軌跡の推定	12
		2.4.1 対応づけの流れ	12
3	Π-	·カルサポート 設定箇所の決定法	17
3	□- 3.1	·カルサポート 設定箇所の決定法 従来法:輝度変化からの選択法	$\frac{17}{17}$
3	□- 3.1 3.2	·カルサポート 設定箇所の決定法 従来法 : 輝度変化からの選択法 提案手法 1 : 輝度変化と時間変化からの選択法	17 17 18
3	□ - 3.1 3.2 3.3	 ・カルサポート設定箇所の決定法 従来法:輝度変化からの選択法 提案手法1:輝度変化と時間変化からの選択法 提案手法2:時間変化とエピポーラ拘束からの選択法 	 17 17 18 21
3	□ - 3.1 3.2 3.3 3.4	 ・カルサポート設定箇所の決定法 従来法:輝度変化からの選択法 提案手法1:輝度変化と時間変化からの選択法 提案手法2:時間変化とエピポーラ拘束からの選択法 カラー画像への対応法 	 17 17 18 21 22
3	□ - 3.1 3.2 3.3 3.4 □ -	 ・カルサポート設定箇所の決定法 従来法:輝度変化からの選択法 提案手法1:輝度変化と時間変化からの選択法 提案手法2:時間変化とエピポーラ拘束からの選択法 カラー画像への対応法 ・カルサポートの大きさの決定法 	 17 17 18 21 22 27
3	□ - 3.1 3.2 3.3 3.4 □ - 4.1	 ・カルサポート設定箇所の決定法 従来法:輝度変化からの選択法 提案手法1:輝度変化と時間変化からの選択法 提案手法2:時間変化とエピポーラ拘束からの選択法 カラー画像への対応法 ・カルサポートの大きさの決定法 大きさ固定の場合 	 17 17 18 21 22 27 27
3	□ - 3.1 3.2 3.3 3.4 □ - 4.1 4.2	 ・カルサポート設定箇所の決定法 従来法:輝度変化からの選択法 提案手法1:輝度変化と時間変化からの選択法 提案手法2:時間変化とエピポーラ拘束からの選択法 カラー画像への対応法 ・カルサポートの大きさの決定法 大きさ固定の場合 ・2次元の adaptive window 	 17 17 18 21 22 27 27 27 27
3	$ \Box - \\ 3.1 \\ 3.2 \\ 3.3 \\ 3.4 \\ \Box - \\ 4.1 \\ 4.2 \\ 4.3 \\ $	 カルサポート設定箇所の決定法 従来法:輝度変化からの選択法 提案手法1:輝度変化と時間変化からの選択法 提案手法2:時間変化とエピポーラ拘束からの選択法 カラー画像への対応法 カルサポートの大きさの決定法 大きさ固定の場合 2次元の adaptive window 3次元の adaptive window 	 17 17 18 21 22 27 27 27 27 30
3 4	□- 3.1 3.2 3.3 3.4 □- 4.1 4.2 4.3 シミ	 カルサポート設定箇所の決定法 従来法:輝度変化からの選択法 提案手法1:輝度変化と時間変化からの選択法 提案手法2:時間変化とエピポーラ拘束からの選択法 カラー画像への対応法 カラー画像への対応法 カルサポートの大きさの決定法 大きさ固定の場合 2次元の adaptive window 3次元の adaptive window ニレーションによる実験 	 17 17 18 21 22 27 27 27 30 32

5.2	ローカルサポート設置箇所選択に関する実験	33
	5.2.1 従来法と提案手法の比較	33
	5.2.2 提案手法1に対する提案手法2の有効性の実験	37
5.3	ローカルサポートの大きさ決定法に関する実験	42
	5.3.1 ローカルサポートの大きさによる測定精度の変化	42
	5.3.2 大きさ固定の窓と adaptive window の比較実験	43
5.4	ステレオマッチングに関する実験	44
	5.4.1 ステレオマッチング法の精度比較	44
	5.4.2 同一パターン区別の実験	46
5.5	特徴領域追跡の実験・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	50
	5.5.1 2次元と3次元の対応づけの比較	50
	5.5.2 比較実験 2	52
5.6	軌跡推定の実験・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	53
まと	b	57
6.1	結論	57
6.2	今後の課題	58
謝辞	5	59
参考	之献 (1997年19月1日)	30
研究	美績 6	32

6

図目次

1.1	3次元的にボリュームを持った画像	2
2.1	対応点問題	5
2.2	従来のローカルサポート	6
2.3	ローカルサポートの拡張	6
2.4	ステレオ視	7
2.5	エピポーラ拘束	9
2.6	ステレオマッチング法 (提案手法)	10
2.7	運動推定(提案手法)	13
2.8	時間方向への追跡・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	14
2.9	対応づけの流れ..................................	15
3 1	開口問題	19
3.2	固有値と領域内の輝度値の分布(従来法)	19
3.3	固有値と領域内の輝度値の分布(提案手法)	20
3.4		23
3.5		23
3.6	提案手法による固有値の画像上の分布	23
3.7	手法1,2,3による固有値のヒストグラムの比較	24
3.8	手法1,2,3による固有値のヒストグラムの比較	25
4.1	2D adaptive window system	29
4.2	3D adaptive window system	30
5.1	実験環境	32
5.2	従来法によるローカルサポート設置箇所	34
5.3	従来法による固有値のヒストグラム	34

5.4	提案手法1によるローカルサポート設置箇所	35
5.5	提案手法1による固有値のヒストグラム	35
5.6	提案手法2によるローカルサポート設置箇所	36
5.7	提案手法2による固有値のヒストグラム	36
5.8	実験画像1の配置図	37
5.9	実験画像 1(左画像)	38
5.10	実験画像 1(右画像)	38
5.11	提案手法1による選択領域	38
5.12	提案手法 1 による対応づけ成功部	38
5.13	提案手法 2 による選択領域	38
5.14	提案手法 2 による対応づけ成功部	38
5.15	提案手法1による固有値のヒストグラム	39
5.16	提案手法 2 による固有値のヒストグラム	40
5.17	実験画像2の配置図	43
5.18	実験画像2(左画像)	44
5.19	実験画像2(右画像)	44
5.20	adaptive window における性能向上	44
5.21	選択領域	45
5.22	従来法による対応づけ成功部・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	45
5.23	提案手法による対応づけ成功部	45
5.24	実験画像3の配置図	46
5.25	実験画像 3(左画像)	47
5.26	実験画像 3(右画像)	47
5.27	従来法で選択された特徴領域	47
5.28	提案手法で選択された特徴領域	47
5.29	従来法による固有値のヒストグラム	48
5.30	提案手法1による固有値のヒストグラム	49
5.31	エッジのある領域...............................	50
5.32	従来法による運動推定成功部・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	50
5.33	提案手法による運動推定成功部・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	50
5.34	実験画像 (左画像:1フレーム目)	53
5.35	実験画像 (右画像:1フレーム目)	53
5.36	実験画像 (左画像:10 フレーム目)	53

5.37	実験画像 (右画像:10 フレーム目)	53
5.38	推定軌跡 (2 次元)	54
5.39	推定軌跡 (2 次元:xy 方向)	54
5.40	推定軌跡 (3 次元)	54
5.41	推定軌跡 (3 次元: xy 方向)	54
5.42	変位ベクトルの平均値	55
5.43	変位ベクトルの全体の分散 (x 方向)	56
5.44	変位ベクトルの全体の分散 (y 方向)	56
5.45	変位ベクトルの全体の分散 (z 方向)	56

表目次

5.1	提案手法1と2の特徴選択数の比較	38
5.2	提案手法1と2のステレオマッチング精度比較	41
5.3	窓の大きさと測定精度の関係	42
5.4	ステレオマッチング法の比較	45
5.5	同一パターンの識別性能比較	48
5.6	運動推定性能の比較	51
5.7	手法の組み合わせ、、、、、、、、、、、、、、、、、、、、、、、、、、、、、、、、、、、、	52
5.8	手法1と手法2の性能比較	52
5.9	従来法と提案手法・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	53

第1章

はじめに

1.1 研究の目的と背景

異なる視点や異なる時刻における画像間での対応点を正確かつロバストに求める研究は 精度の高いステレオ視や3次元運動推定を実現し,自動車のナビゲーション,運動予測, 弾道計画,自動監視システムなどへの幅広い応用が期待される.

ステレオ視や時系列画像から3次元運動を解析する際の最大の問題点は,2つの画像間 (左右,時間)で正しい対応点を見つけることである.

従来の対応づけ方法は,小領域内の画素の輝度の差から対応づけを行う領域法 (areabased matching)[1],[2],[3],[4] とエッジなどの特徴を抽出して対応づけを行う特徴法 (featurebased matching) [5],[6],[7] の 2 つに大別される.前者は密な視差マップを求めることができるが,精度に欠け計算コストがかかることが多い.一方,後者は疎な視差マップしか求めることができず精度が特徴抽出の結果に依存するが,計算コストが低く画像間の幾何学的な歪みの影響を受けにくい.もし対応づけが正確に得られるならば,領域法のほうが密に視差マップを得られるため応用範囲が広いと考えられる.領域法では,精度を向上させるために物体の角のような特徴的な領域を選択し,対応を求める.この特徴的な領域のことを以下ではローカルサポート (local support)領域と称する.

提案手法では,同じ大きさの従来のローカルサポートよりも豊富な情報を用いること によって正確な対応づけを行うために,図1.1のように時系列の連続画像を3次元的にボ リュームを持った画像としてとらえ,従来は注目画像平面上のみで定義されていたローカ ルサポート領域を時系列画像の中から立体的にとる.このような時空間の特徴を用いたブ ロックマッチングの研究としては文献[8]がある.これは時系列画像の中から斜方体のブ ロックを取り出し,精度の高い変位ベクトルの推定を行うというものである. ローカルサポートを画像の設置箇所は,文献[1]の固有値による領域の選択法を3次元 方向に拡張した手法によって選択する.従来法では画像上から一様にローカルサポート領 域が選択されていたが,この手法では時間変化している領域からしか選択されないため運 動追跡に適している.このローカルサポート領域は輝度の時間変化を含むことからオプ ティカルフロー的な性質も帯びることになる.文献[9]ではオプティカルフローと距離情 報から移動物体の追跡を行っている.これはフロー情報と距離情報の一方では追跡が困難 なところを,もう一方を用いて追跡するという相互補完の性格が強いものである.

このローカルサポートの大きさをどのようにとるかについては,文献[1],[2] などで「ウィ ンドウの大きさはマッチングに対して輝度変化を含むのに十分大きく,しかし,射影の歪 みの効果を避けるのに十分小さくなくてはならない」と指摘されている.また,各ローカ ルサポートの適切な大きさは周囲の状況にもよるため,全てのローカルサポートの大きさ を均一にするのは対応づけの精度を下げる要因となりうる.そのため文献[2]の適応型の ウィンドウを提案したローカルサポートに応用する.

本稿ではこの特徴領域を用いて対応づけを行い,時系列のステレオ画像から運動物体の 正確な3次元軌跡を復元するために,精度の高い対応づけが行えるようなローカルサポー ト領域の設定方法を提案する.



時系列画像とローカルサポート ローカルサポートの拡張

図 1.1: 3 次元的にボリュームを持った画像

1.2 本論文の構成

本論文の構成は次の通りである.第2章では,ステレオ視や運動推定における対応点問 題について述べる.第3章では,ローカルサポートとしてふさわしい部分を画像上から どのようにして選択し,選択されたローカルサポートの大きさをどのように設定するかに ついて述べる.また,第4章には,シミュレーション画像を用いた実験結果を示す.第5 章では,本研究の結論について述べる.

第2章

対応点問題とローカルサポート

2.1 対応点問題

本稿の目的は,左右のステレオマッチングの問題も動画像解析の問題もブロックマッチングによる対応づけによって求めることにより図2.1のような時系列のステレオ動画像から3次元軌跡を推定することである.

ブロックマッチングによる軌跡推定の原理は,左画像中のある画素が右画像中のどこの 画素に対応しているのか,また,その画素が次の時刻におけるフレームでどの画素に対応 しているのかを求めることである.つまり,フレーム L₁, R₁, L₂, R₂の4点で対応がとれ たら,2フレームの3次元軌跡を推定することができる.

本来ならば対応点探索は対象上の観測点の3次元位置が自由に変化するような場合は, 画素ごとに行わなければならない.しかし,画像間の対応づけには本質的に「あいまい さ」が存在する.具体的には,シーンによっては似たようなパターンが随所に存在するこ ともあり,そうでなくても視点の異なる投影による幾何学的な歪み,カメラ特性の違い, ランダムノイズ,オクルージョン,光源や物体の運動による見えかたの変化などの原因に より左右の対応点は完全に一致しない.このように画素単位での対応づけを正確に行うこ とは複雑で多様な現実のシーンでは非常に困難で,不可能といってよい.

そのため多くの文献では対応点問題に対し、

1. 注目点の周囲に窓を設定し,その窓内の局所的なパターンから対応点を探索する. (領域法)

2. 画像上からエッジなどの特徴を抽出し,両画像の特徴間で対応づけを行う.(特徴法) といった方法がとられてきた.これらの方法は注目点と対応点の間で対応づけを行う際 に,1 では局所的な部分の輝度値のパターンに基づいて.2 では濃淡エッジなどの特徴に 基づいて行っている.特徴法では密な対応づけが可能で,正確に対応づけられるなら広範 囲に応用できるため,以下では領域法を用いる.



図 2.1: 対応点問題

2.2 ローカルサポート領域の拡張

画素単位での対応づけのあいまいさを解消するために,より多くの情報源を利用して, 対応づけの信頼性を高めるという方策がある.この1つの具体例として,注目点の周囲の 画素を情報源として用いるローカルサポート領域がある.ローカルサポート領域とは注目 点の周りの局所的な輝度パターンをテンプレートとしてとったものであり,画像中の隣接 した画素間には通常何らかの関連があることを利用し,図2.2のようにローカルサポート 領域の対応によって対応点を求める.

さらに,画素間の関連は同じ画像上だけではなく,近い瞬間に撮影された時系列画像同 士にも非常に密接な相関関係が存在することに注目する.ローカルサポート領域を例えば 注目画像とその前後の瞬間に撮影された3枚の画像から選択することによって,より大 きくローカルサポートをとることができる.そのため提案手法では,ローカルサポート領 域を注目点のある画像上だけではなく連続した複数枚の時系列画像から立体的にとって



図 2.2: 従来のローカルサポート



図 2.3: ローカルサポートの拡張

いる.図 2.3にその様子を示した.文献 [8] では,時系列画像画像から斜方体のローカル サポートをとり,運動推定を行っているが,本稿で提案するローカルサポートは直方体に なる.

どのようなローカルサポートを良い特徴と考えるかについては,文献 [1],[3] で次のよう に述べられている「良い特徴というものは正しい数学的な意味においてうまく探索され うるものであると定義できる.」以下ではこの定義に従って,対応づけがうまくいくよう な特徴を「良い特徴」としている.つまり,ローカルサポートの性能をステレオマッチン グ,運動推定における対応づけの精度で計ることにする.第4章でマッチングの成功率を ローカルサポートの評価に用いているのもこの定義による.

2.3 ステレオ視と軌跡推定

2.3.1 ステレオ視



図 2.4: ステレオ視

ステレオ視の基本的な原理は三角測量であり,左右一対のカメラでとられた画像上で シーンの中で同じ点に対応する写像の左右の位置を求めれば,シーンの中での3次元位置 情報が得られるというものである.この原理は古くから知られ,写真測量技術として確立 しているが,対応探索,つまり左右の画像上での対応点を求める操作が,自動化にあたっ て非常に困難な問題として残っている.

ブロックマッチングと呼ばれる対応づけ手法ではテンプレートとして対応づけに用いられるローカルサポート領域を一方の画像からとり,他方の画像の探索範囲内に同じ領域をとり,その2つの領域間で輝度値の正規化相互相関 (normalized cross correlation) や差の二乗和 (SSD: Sum of Squared Differences)を計算し,対応を求めることになる. 図 2.4に基本的なステレオ視の配置を示す.この図では水平方向に平行に左右一対の同じカメラを置いて,前方3次元空間内の点P(X,Y,Z)の像を左画像,右画像上でそれそれ $P_l(X_l,Y_l), P_r(X_r,Y_r)$ に結像させている.ここで,両カメラのレンズの焦点距離をf,左右のカメラのレンズ位置間の距離 (基線長)をbとすると以下の関係式が成立する.

$$X = \frac{b(X_l + X_r)}{2d} \tag{2.1}$$

$$Y = \frac{b(Y_l + Y_r)}{2d} \tag{2.2}$$

$$Z = \frac{bf}{d} \tag{2.3}$$

ただし,

$$d = X_l - X_r \tag{2.4}$$

である.これらの式から分かるように,いちばん重要な値は *d* であり,この *d* のことを視差 (disparity) と呼ぶ.また,図 2.4のように焦点距離が等しい光軸が平行なカメラが 2 台 水平に並ぶような配置では常に,

$$Y_l = Y_r \tag{2.5}$$

という関係が成立し,対応点を見つける際には同じY座標の線を探索すればよい.この ような,直線のことをエピポーラ線と(epipolar line)と呼び,ステレオ画像の対応点が満 たすべきこのような拘束条件をエピポーラ拘束(epipolar constraint)と呼ぶ.図2.5はエ ピポーラ幾何の概念を表している.

2つのカメラが3次元空間中の同一点を射影しているとすると,このとき3次元空間中 の注目点と両カメラの撮影中心によって,3次元空間中に1つの平面が定まる.この平面 をエピポーラ平面と呼び,エピポーラ平面と左右の画像平面上の交線をエピポーラ線と呼 ぶ.つまり,左画像平面の撮影中心点の右画像平面における対応点は必ずエピポーラ平面 と右画像平面の交線であるエピポーラ線に乗るということを表わしている.よって,左画 像平面のある1点の対応点を右画像平面から探索するには,右画像平面を全部探索する必



図 2.5: エピポーラ拘束

要はなく,右画像平面上のエピポーラ線上のみを探索すればよい.この拘束条件を用いる ことによって,探索範囲を2次元から1次元に落とすことが可能になり,探索を効率的に 行うことができる.

時間方向に拡張した立体的なローカルサポートの差の二乗和を計算することによって, ステレオフレーム間の対応づけを行なう.ここで,t = Tでの(x, y, t)における左右のロー カルサポートの画像輝度をそれぞれ L(x, y, t), R(x', y', t + 1)とおおくと,従来は距離は 式 (2.6) で定義されていた.

距離
$$(SSD) = \sum_{W} (L(x, y) - R(x', y'))^2$$
 (2.6)

これを時間方向に拡張し,図2.6に示すように,ローカルサポートを立体的にして式(2.7) によって探索する.

$$\mathbf{ER}(SSD) = \sum_{H} \sum_{W} (L(x, y, t) - R(x', y', t))^2$$
(2.7)

ただし,Nは計算に用いるフレーム数,Wはローカルサポート領域の1フレームあたりの画素数である.(x',y')を変化させ,距離が最小になったところを対応点とする.このようにローカルサポート領域を時間方向に拡張し,立体的にステレオマッチングを行うことによる利点としては,



図 2.6: ステレオマッチング法 (提案手法)

- 同一パターンを持ち,異なる速度で変化する J(x, y, t) と K(x, y, t) のような領域を
 区別できる.
- ノイズなどに対する耐性が向上する.

ということが挙げられる.

2.3.2 軌跡推定

時系列の2フレームの間で選択された特徴領域を追跡するには,いくつかの方法が考 えられる.

- 1.2フレーム間でそれぞれ特徴点をとり,比較する.
- 2. 輝度勾配の拘束式から移動ベクトルを推定する.(勾配法)
- 3.1 フレームの特徴点を求め,次フレームで差の二乗和 (SSD) を最小にする点を探索 する.(ブロックマッチング法)

1は,一般の画像では2フレーム間で必ず同じ箇所が選択されるという保証がないため, 用いることができない.

2の勾配法では,オプティカルフローの拘束式から変位ベクトルを計算する.この手法 では変位ベクトルを一意に決定するときに「ローカルサポート内の速度分布が滑らかに変 化する」という仮定が用いられる.物体と背景の境界付近などの領域では,この仮定が成 り立たないためこの手法を用いることができない.

以上より本稿では,3の手法を用いる.時刻t = Tにおいて,選択された3次元ローカ ルサポート領域と時刻t = T + 1における同じ大きさの領域とで差の二乗和をとり,その 距離が最小になる領域を時刻t = T + 1の画像中から探索し,それを時刻t = T + 1にお ける対応点とする.

ここで,t = T, T + 1における領域の中心座標をそれぞれI(x, y, t), I(x', y', t + 1)とすると,従来は距離は式 (2.8) で定義されていた.

距離
$$(SSD) = \sum_{W} (I(x, y, t) - I(x', y', t+1))^2$$
 (2.8)

これを時間方向に拡張し,図2.7に示すように,ローカルサポートを立体的にして式(2.9) によって探索する.

距離
$$(SSD) = \sum_{H} \sum_{W} (I(x, y, t) - I(x', y', t+1))^2$$
 (2.9)

ただし,Nは計算に用いるフレーム数,Wはローカルサポート領域の1フレームあたりの画素数,I(x, y, t)は(x, y, t)における画像輝度である.(x', y')を変化させ,距離が最小になったところを対応点とする.時間方向に拡張する利点としては,

- 物体が等速運動しているときは (1) に比べて (2) のほうがより多くの一致を得るので,安定した対応づけが可能になる.
- 同じパターンを持ち,異なる速度で運動しているローカルサポートの区別が可能になる.

が, 欠点としては,

● (3)のように物体が非等速運動をしている場合には,誤探索が多くなる可能性がある.

が挙げられる.

また,時刻t = T + 1以降の軌跡は,図2.8 に示すようにt = T + 1で提案手法により 新たに特徴領域を選択しt = T + 2へその領域を追跡するという手順を繰り返すことに よって行う.この方法では追跡している特徴領域が同じ部分であるということは,保証さ れない.しかし,剛体運動している物体から移動軌跡を求める際には,運動物体上からど の点を選択しても変位はほぼ一定であると仮定できる.運動物体が自転しながら移動し ており自転速度が移動速度に対して無視できないほど大きい場合などはこの仮定は成り 立たないが,通常の移動軌跡を求める際にはこの仮定を用いることができる.そのため, 2フレーム間の対応づけがうまくとれていれば,特徴領域の連続性が保証できなくてもそ れほど問題はない.

2.4 3 次元軌跡の推定

2.4.1 対応づけの流れ

3次元軌跡推定は図 2.9の流れで行う.ここでは 3次元軌跡をよりロバストがつ正確に 推定するために次のような方針に従って議論を進める.

- 画像中からできるだけ対応づけの成功率が高い箇所を特徴領域として選択する.
- 選択した領域の大きさをできるたけ対応づけの成功率が高くなるように設定する.
- 得られた領域をできるだけ誤探索が少なくなるように対応づける.



図 2.7: 運動推定(提案手法)



図 2.8: 時間方向への追跡



図 2.9: 対応づけの流れ

ローカルサポート設定箇所の選択

正確に軌跡推定を行うことが目的なので,対応づけのとりやすい領域のみを選択する. そのため,ステレオマッチングおよび時間方向への追跡が容易であると考えられるローカ ルサポートを画像中のどこから選択するかについて考慮する必要がある.

画像中からローカルサポートのような特徴的な点を選択し,それを用いて対応づけを行う利点としては,次の2つが挙げられる.

- あまり特徴的ではない点は対応づけの精度が低く、これを使わないことによって全体的な対応づけの精度が向上する。
- 一般的に特徴点はそれ以外の点に比べて数が少ないので,計算量が少なくてすむ.

各ローカルサポートが確実に対応づけできるという保証がない以上,ある程度の割合で マッチングが成功するだけの領域数を選択しないと正確なマッチングは望めない.しか し,逆に領域数を増やしすぎるとマッチングや追跡にあまりふさわしくない特徴領域まで 含んでしまうため,結果として推定精度は下がる.いくつの領域を選択するかを決定する ときには,このトレードオフを考えなくてはならない.

ローカルサポートの大きさの決定

画像中から選択されたローカルサポート領域の大きさは,小さすぎると十分な輝度変化 をカバーできず,ノイズの影響を受けやすくなる.一方,大きすぎると視差の変化,左右 の画像の射影による歪みによって最大の相関や最小の SSD が正しい対応を示さないこと がある.また,適切な領域の大きさは各領域によって異なるため,窓の大きさを固定する と誤対応が多くなることがある.そのため,adaptive window と呼ばれる適応型のウィン ドウを用い,各ローカルサポートの周辺状況に応じたウィンドウの大きさを決定する.

ローカルサポートに基づく対応づけ

このようにして設定されたローカルサポートを用いてステレオフレーム間および時系 列画像間の対応づけを行う.

第3章

ローカルサポート 設定箇所の決定法

3.1 従来法:輝度変化からの選択法

Tomasi らは軌跡推定を行うために,画像中から物体の角のような特徴的な箇所を選択 するために以下の方法を用いている[1].

近い瞬間に撮られた画像間には、通常は強い相関関係があることから,それらの画像間の輝度が保存されているとみなし,式(3.1)のように表すことができる.

$$I(x, y, t + \tau) = I(x - \xi, y - \eta, t)$$
(3.1)

$$I(\boldsymbol{x}) = I(x, y, t + \tau), \quad I(\boldsymbol{x} - \boldsymbol{d}) = I(x - \xi, y - \eta, t)$$
(3.2)

ノイズ $n(\boldsymbol{x})$ によって輝度が完全に保存されない場合を考え,画像 $I(\boldsymbol{x}), I(\boldsymbol{x}-\boldsymbol{d})$ 間の関係を,

$$I(\boldsymbol{x}) = I(\boldsymbol{x} - \boldsymbol{d}) + n(\boldsymbol{x}) \tag{3.3}$$

と表わす.

ここで $I(\boldsymbol{x} - \boldsymbol{d})$ を,

$$I(\boldsymbol{x} - \boldsymbol{d}) = I(\boldsymbol{x}) - \boldsymbol{g} \cdot \boldsymbol{d}$$
(3.4)

と Taylor 展開し, あるローカルサポート領域 W全体における輝度保存を考える.領域全体のノイズを くすると, eは,

$$\epsilon = \sum_{W} [I(\boldsymbol{x}) - \boldsymbol{g} \cdot \boldsymbol{d} - I(\boldsymbol{x} - \boldsymbol{d})]^2 \omega = \sum_{W} [h - \boldsymbol{g} \cdot \boldsymbol{d}]^2 \omega$$
(3.5)

のように表すことができる.ここで,ωは領域の重み関数である.ただし、

$$\boldsymbol{g} = (I_x, I_y) \tag{3.6}$$

 $I(\boldsymbol{x}) - I(\boldsymbol{x} - \boldsymbol{d})$ は同じ点における輝度変化を表すので,

$$h = I(\boldsymbol{x}) - I(\boldsymbol{x} - \boldsymbol{d}) = I_t \tag{3.7}$$

である.ノイズ ϵ を最小にする変位ベクトル $d = (\xi, \eta)$ を求めたいので、式 (3.5) を ξ, η で それぞれ偏微分することによって, 2 つの式 (3.8) が得られる.

$$(\sum_{W} I_{x} I_{x} u + \sum_{W} I_{x} I_{y} v)\omega = (\sum_{W} I_{x} I_{t})\omega$$

$$(\sum_{W} I_{y} I_{x} u + \sum_{W} I_{y} I_{y} v)\omega = (\sum_{W} I_{y} I_{t})\omega$$
(3.8)

これより,行列式(3.9)が得られる。

$$\begin{pmatrix} \sum_{W} I_{x} I_{x} \omega & \sum_{W} I_{x} I_{y} \omega \\ \sum_{W} I_{y} I_{x} \omega & \sum_{W} I_{y} I_{y} \omega \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \sum_{W} I_{x} I_{t} \omega \\ \sum_{W} I_{y} I_{t} \omega \end{pmatrix}$$
(3.9)

この行列式 (3.9) の左辺の 2 × 2 行列

$$\begin{pmatrix} \sum_{W} I_{x} I_{x} \omega & \sum_{W} I_{x} I_{y} \omega \\ \sum_{W} I_{y} I_{x} \omega & \sum_{W} I_{y} I_{y} \omega \end{pmatrix}$$
(3.10)

の2つの固有値 λ_1, λ_2 と閾値 λ の関係から時系列画像の追跡に用いるための特徴領域を選 択する.図3.1で示されるような開口問題 (aperture problem) などを考慮すると,特徴と して選択するのがふさわしいのは(3)のような領域内で輝度勾配が多方向に分散している 部分である.ここで(3.10)の固有ベクトルは, $I_x = I_y$ 空間での輝度勾配の分布が最大の 広がりを見せる方向とそれに直行する方向のベクトル(図3.1では赤い矢印で示してある) を表し,固有値はその大きさを表しているため,固有値 λ_1, λ_2 と閾値 λ の関係が、

$$\min(\lambda_1, \lambda_2) > \lambda \tag{3.11}$$

となる領域を選択する.

3.2 提案手法1:輝度変化と時間変化からの選択法

提案手法では対象とする一画像だけではなく、その前後に撮られた画像からも情報を得て、特徴点を選択する際に時間軸(t軸)を加え、ローカルサポート領域 W'(時間方向に連



- 1. 領域内にエッジがない:対応づけが困難
- 2. 領域内に1次元のエッジ:一意に対応が決まらない
- 3. 領域内に2次元のエッジ:一意に対応が決定

図 3.1: 開口問題



- 1. 領域内で輝度勾配が一定 $\rightarrow \lambda_1, \lambda_2 < \lambda$
- 2. 領域内で輝度勾配が1つの方向に集中 $\rightarrow \lambda_1 < \lambda < \lambda_2$
- 3. 領域内で輝度勾配が多方向に分散 $\rightarrow \lambda < \lambda_1, \lambda_2$

図 3.2: 固有値と領域内の輝度値の分布(従来法)



1. 時間的に輝度変化がない領域

2. 時間的に輝度変化があり輝度変化が1つの方向に集中

3. 時間的に輝度変化があり輝度変化が多方向に分散

図 3.3: 固有値と領域内の輝度値の分布(提案手法)

続した N 枚のローカルサポート領域 W) で,拡張した行列式

$$\begin{pmatrix} \sum_{W'} I_x I_x \omega & \sum_{W'} I_x I_y \omega & \sum_{W'} I_x I_t \omega \\ \sum_{W'} I_y I_x \omega & \sum_{W'} I_y I_y \omega & \sum_{W'} I_y I_t \omega \\ \sum_{W'} I_t I_x \omega & \sum_{W'} I_t I_y \omega & \sum_{W'} I_t I_t \omega \end{pmatrix}$$
(3.12)

の3つの固有値を $\lambda_1, \lambda_2, \lambda_3$ とし,閾値を λ とすると,

$$\min(\lambda_1, \lambda_2, \lambda_3) > \lambda \tag{3.13}$$

となる領域は,領域内で輝度勾配が多方向に分散していて,輝度の時間変化もあるという 特徴的な領域であるため,図3.2の(3)の部分に相当するこの領域を選択する。*I*_tは対象 となる画像の点の時間軸方向の偏微分を表す。 このローカルサポート領域は,内部に輝度の変化と時間変化を持つ非常に特徴的な領域 であるので,ステレオマッチングにも運動推定にも用いることが可能である.

3.3 提案手法2:時間変化とエピポーラ拘束からの選択法

提案手法1では,ローカルサポート空間内で3方向のエッジを持つ必要がある.これ は極めて厳しい条件のため,選択される領域はどうしても少なくなってしまう.そこで, より多くの領域を選択するために,ステレオマッチングでは第2章で述べたようにエピ ポーラ線上を探索すれば対応点が求まるので対応づけの決め手となるのはエピポーラ線 に平行なエッジであることに着目し,エピポーラ線に平行なエッジを持ち且つ時間変化を している領域を選択するという手法を考える.

ここで,エピポーラ線の方向ベクトルを (e_x, e_y) とすると,

$$\begin{pmatrix} \sum_{W'} I_e I_e \omega & \sum_{W'} I_e I_t \omega \\ \sum_{W'} I_e I_t \omega & \sum_{W'} I_t I_t \omega \end{pmatrix}$$
(3.14)

ただし,

$$I_x e_e + I_y e_y = I_e, e_x^2 + e_y^2 = 1$$
(3.15)

行列 (3.14) の 2 つの固有値を λ_1, λ_2 とし, 閾値を λ とすると,

$$\min(\lambda_1, \lambda_2) > \lambda \tag{3.16}$$

となる領域は時間変化するエピポーラ線に平行なエッジを持つ特徴的な領域であるので選 択する。

例えば,図2.4のようなカメラ配置の場合はエピポーラ線がx軸に平行となるので,エ ピポーラ線の方向ベクトルは

$$(e_x, e_y) = (1, 0)$$

であり,行列式(3.14)は,

$$\begin{pmatrix} \sum_{W'} I_x I_x \omega & \sum_{W'} I_x I_t \omega \\ \sum_{W'} I_x I_t \omega & \sum_{W'} I_t I_t \omega \end{pmatrix}$$
(3.17)

となる.

一般に特徴点の選択数を増やしてゆくことは,標本数が増えるという点では好ましいが,選ばれた特徴点の信頼性が下がってゆくので一概に良いとは言えない.しかし,提案
 手法1で選択された箇所の他に多くの特徴点が選択したいときなどは,提案手法2を用いることによって対応づけに有効な特徴点の選択数を増やすことができる.

3.4 カラー画像への対応法

論文 [1],[2] 等ではグレイスケールの画像を対象にしているため,画像輝度がI(x, y, t), f(x, y)のように表されている.提案手法もこれにならって,画像の輝度をI(x, y, t)のように表記したが,実際には輝度はR[0, 255], G[0, 255], B[0, 255]のように RGB の3つの輝度を持つ.カラー画像を対象にする場合はこの $R, G, B \ge I$ の関係をどのようにして表すか考慮する必要がある.ここで,R,G,B の輝度を $I_R, I_G, I_B \ge U$,式 (3.10),(3.12),(3.14) によって実際に計算される固有値を λ とする.このとき $I_R, I_G, I_B \ge I$ の間の関係をどのように表すかについて,いくつかの方法が考えられる.

- 1. $\sum_{W} I_x^2 \omega = \sum_{W} (I_{Rx}^2 + I_{Gx}^2 + I_{Bx}^2) \omega \dots$ として固有値を取り評価する。
- 2. $I = 0.299I_R + 0.587I_G + 0.114I_B$ として固有値を取り評価する。
- 3. $I = \sqrt{I_R^2 + I_G^2 + I_B^2}$ として固有値を取り評価する。

以上3つの手法を評価するために図3.4に示す実画像を用い,以下の実験を行った.実画 像は以下の条件で撮影した.

- デジタルビデオカメラで撮影.
- 照明 複数の室内蛍光灯.
- 640 × 480 画素の PPM 画像,フレーム間隔 1/30sec.

この実画像を上の3つの手法を用い式(3.10)で表される手法で固有値を計算したときの固有値の画像上の分布と固有値のヒストグラムをそれぞれ図3.5,図3.7に示す.図3.5 は固有値の大きさによって,画像平面を10段階に色分けして固有値の画像上の分布を示したものである.明るい部分が小さな固有値で,暗いほとが大きな固有値を表している.

また,式(3.12)で表される提案手法で固有値を計算したときの固有値の画像上の分布 と固有値のヒストグラムを図(3.4),図(3.8)に示す.





図 3.4: 実画像(時系列)





図 3.5: 従来法による固有値の画像上の分布



図 3.6: 提案手法による固有値の画像上の分布



図 3.7: 手法1,2,3 による固有値のヒストグラムの比較



図 3.8: 手法1,2,3 による固有値のヒストグラムの比較

従来法,提案手法のどちらの方法でも,ヒストグラムは手法1,2,3 に対して,同じよう な挙動を示している.また,固有値の画像上の分布にもほとんど変化がみられない.し かし,方法1は方法2,3 に比べて標本数が3倍あり,ヒストグラムも方法2に比べて広 がっている.よって,閾値の微少な変化に対して選択される部分が大きく変化しないので 閾値が決定しやすく,また標本数が多いことから他の手法よりも閾値の信頼性が高いと考 えられる.ゆえに以下では,手法1を採用した.

第4章

ローカルサポートの大きさの決定法

4.1 大きさ固定の場合

前章で相関関係,SSD を用いたマッチングの問題点としては,窓はマッチングに対し て輝度の変化を含むのに十分大きくかつ射影の歪みの効果を避けるのに十分小さくなく てはいけないということを述べた.あらかじめ大きさを固定した窓では,対応する画像フ レームとステレオマッチングを行って視差を求め,その視差から窓の大きさを求めるとい うステップが省けるため,全体の処理の実行時間は早い.しかし,各ローカルサポートに 対して窓の大きさは均一になり,周囲の状況に対する考慮が全くなされていない.

窓の大きさを固定してしまうことの大きな欠点は、ウィンドウを複数とったとき、ウィ ンドウにはそれぞれふさわしい大きさがあるということを全く考慮に入れていないこと である.例えば3×3の窓だと精度がよくないからといって5×5に拡張すると今度は一 部で窓が視差の異なる物体を跨がったりすることがある.つまり、適切な窓の大きさは選 択された特徴点の周囲状況によって決まるので、一意には定まらない.そのため、個々の 特徴点に対して窓の大きさが決まるような適応型の窓(adaptive window)のほうが固定式 の窓よりも精度が向上すると考えられる.

4.2 2次元の adaptive window

Kanade ら [2] はローカルサポート内の輝度変化と視差変化を次のようにモデル化した. まず,基線がx軸に平行な左右のステレオ画像平面を $f_1(x, y), f_2(x, y)$ と定義する.(x, y)での視差を $d_r(x, y)$ とすると, $f_1(x, y) \ge f_2(x, y)$ の間の関係は次のように表すことがで
きる.

$$f_1(x,y) = f_2(x + d_r(x,y), y) + n(x,y)$$
(4.1)

 $n(x,y) \sim N(0, 2\sigma_n^2)$ は白色ノイズである.

ローカルサポート内の $f_1 \ge f_2 \ge f_1(\xi, \eta), f_2(\xi, \eta) \ge f_2(\xi, \eta) \ge f_2(\xi, \eta)$ の間の輝度の差は,

$$f_1(\xi,\eta) - f_2(\xi + d_r(0,0),\eta) \approx (d_r(\xi,\eta) - d_r(0,0)) \frac{\partial}{\partial \xi} f_2(\xi + d_r(0,0),\eta) + n(\xi,\eta)$$
(4.2)

と近似できる.式(4.2)の右辺をnsとして,次のようにガウス分布で近似する.

$$ns(\xi,\eta) \approx f_1(\xi,\eta) - f_2(\xi + d_r(0,0),\eta) \sim N(0, 2\sigma_n^2 + \alpha_f \alpha_d \sqrt{\xi^2 + \eta^2})$$
(4.3)

ただし,

$$\alpha_f = E\left[\left(\frac{\partial}{\partial\xi}f_2(\xi + d_r(0,0),\eta)^2\right)\right]$$
(4.4)

$$\alpha_d = \frac{E\left[\left((d_r(\xi,\eta) - d_r(0,0))^2\right)\right]}{\sqrt{\xi^2 + \eta^2}}$$
(4.5)

である.

また,領域の相関関係,エッジの方向性などの画像特性が純粋にそれぞれの点から独立 して与えられるときにはたいていよい精度の対応づけが得られる.しかし,実際には反射 光など画素は近隣の画素からの影響を受けている.そのため, $d_r(\xi, \eta)$ 内の視差の変化量 を次のようにモデル化した.

$$d_r(\xi,\eta) - d_r(0,0) \sim N(0, \alpha_d \sqrt{\xi^2 + \eta^2})$$
(4.6)

(0,0)はローカルサポートの設置された座標である.これは (ξ,η) の(0,0)からの視差の 差がそれら距離に比例したゼロ平均のガウス分布になっていることを表している.

以上のモデル化によりローカルサポート内の輝度変化と視差変化によって,最も確から しい視差推定値と視差推定の信頼性(不確かさ)を記述することが可能になる.

adaptive window は下の式によって制御される.

視差の増加分の推定値

$$\widehat{\Delta d} = \frac{\sum_{i,j \in W} \frac{(f_1(\xi_i, \eta_j) - f_2(\xi_i + d_0(0, 0), \eta_j)) \frac{\partial}{\partial \xi} f_2(\xi_i + d_0(0, 0), \eta_j)}{2\sigma_n^2 + \alpha_f \alpha_d \sqrt{\xi_i^2 + \eta_j^2}}}{\sum_{i,j \in W} \frac{(\frac{\partial}{\partial \xi} f_2(\xi_i + d_0(0, 0), \eta_j))^2}{2\sigma_n^2 + \alpha_f \alpha_d \sqrt{\xi_i^2 + \eta_j^2}}}$$
(4.7)



🛛 4.1: 2D adaptive window system

推定の不確かさ

$$\sigma_{\widehat{\Delta d}}^2 = \frac{1}{\sum_{i,j \in W} \frac{\left(\frac{\partial}{\partial \xi} f_2(\xi_i + d_0(0,0),\eta_j)\right)^2}{2\sigma_n^2 + \alpha_f \alpha_d \sqrt{\xi_i^2 + \eta_j^2}}}$$
(4.8)

ただし, $\sigma_n, \alpha_d, \alpha_f$ は,

ノイズの量

$$\sigma_n = \frac{1}{N_\omega} \sum_{i,j \in W} (f_1(\xi_i, \eta_j) - f_2(\xi_i + d_0(0, 0), \eta_j))^2$$
(4.9)

窓内の視差の変動量

$$\alpha_d = \frac{1}{N_\omega} \sum_{i,j \in W} \frac{(d_0(\xi_i, \eta_j) - d_0(0, 0))^2}{\sqrt{\xi_i^2 + \eta_j^2}}$$
(4.10)

輝度の変化量

$$\alpha_f = \frac{1}{N_{\omega}} \sum_{i,j \in W} (\frac{\partial}{\partial \xi} f_2(\xi_i + d_0(0,0), \eta_j))^2$$
(4.11)

ここで, *N*_ωはローカルサポート領域 *W*の中の標本数である.

これらの結果から,画像上のどの画素の周囲からも最小の不確かさで視差の推定を生成 する窓を探索することができる.

adaptive window は原理的には任意の形状をもつが,以下のアルゴリズムでは長方形の 窓になるように制限されている.その幅と高さは-x,+x,-y,+yの4方向に独立してコン トロールされている.

1. マッチングを行ない, 視差の初期推定 d₀(x, y) を求める

(a) 点 (x, y) を中心に 3 × 3 の窓を設定して (4.8) を求める

- (b) *x*+, *x*-, *y*+, *y*-の各方向に窓を伸ばし,それに対する(4.8)を計算
- (c) (4.8) を最小にする方向に窓を伸ばし,(4.8) を増やす方向には拡張を禁止
- (d) これを全ての方向が拡張を禁止されるか,規定の大きさに達するまで繰り返す
- 2. (4.7) によって推定視差の増加分を計算し,

$$d_{i+1}(x,y) = d_i(x,y) + \Delta d(x,y)$$

の式で推定視差を更新する

3. 以上のプロセスを d_iが収束するか、ある繰り返し数に達するまで繰り返す

4.3 3次元の adaptive window

ここでは,ローカルサポート $d_r(\xi, \eta, \zeta)$ 内の近隣の画素からの影響力を次のようにモデル化した.

$$d_r(\xi,\eta,\zeta) - d_r(0,0,0) \sim N(0,\alpha\sqrt{\xi^2 + \eta^2 + \zeta^2})$$
(4.12)

(0,0,0)はローカルサポートが設置された座標である.これは (ξ,η,ζ) の(0,0,0)からの視差の差がそれらの距離に比例したゼロ平均のガウス分布になっていることを表している.

窓を3次元方向に拡張する利点としては,ローカルサポートとして定義される時系列 画像の枚数を適応的に決定できるということがある.

adaptive window を次のように 3 次元に拡張する.

視差の増加分の推定値

$$\widehat{\Delta d} = \frac{\sum_{i,j,k\in W'} \frac{(f_1(\xi_i,\eta_j,\zeta_k) - f_2(\xi_i + d_0(0,0,0),\eta_j,\zeta_k)) \frac{\partial}{\partial \xi} f_2(\xi_i + d_0(0,0,0),\eta_j,\zeta_k)}{2\sigma_n^2 + \alpha_f \alpha_d \sqrt{\xi_i^2 + \eta_j^2 + \zeta_k^2}}}{\sum_{i,j,k\in W'} \frac{(\frac{\partial}{\partial \xi} f_2(\xi_i + d_0(0,0,0),\eta_j,\zeta_k))^2}{2\sigma_n^2 + \alpha_f \alpha_d \sqrt{\xi_i^2 + \eta_j^2 + \zeta_k^2}}}$$
(4.13)

推定の不確かさ

$$\sigma_{\widehat{\Delta d}}^{2} = \frac{1}{\sum_{i,j,k \in W'} \frac{(\frac{\partial}{\partial \xi} f_{2}(\xi_{i} + d_{0}(0,0,0),\eta_{j},\zeta_{k}))^{2}}{2\sigma_{n}^{2} + \alpha_{f} \alpha_{d} \sqrt{\xi_{i}^{2} + \eta_{j}^{2} + \zeta_{k}^{2}}}}$$
(4.14)

ただし, $\sigma_n, \alpha_d, \alpha_f$ は,



 \blacksquare 4.2: 3D adaptive window system

ノイズの量

$$\sigma_n = \frac{1}{N_{\omega'}} \sum_{i,j,k \in W'} (f_1(\xi_i, \eta_j, \zeta_k) - f_2(\xi_i + d_0(0, 0, 0), \eta_j, \zeta_k))^2$$
(4.15)

窓内の視差の変動量

$$\alpha_d = \frac{1}{N_{\omega'}} \sum_{i,j,k \in W'} \frac{(d_0(\xi_i, \eta_j, \zeta_k) - d_0(0, 0, 0))^2}{\sqrt{\xi_i^2 + \eta_j^2 + \zeta_k^2}}$$
(4.16)

輝度の変化量

$$\alpha_f = \frac{1}{N_{\omega'}} \sum_{i,j,k \in W'} \left(\frac{\partial}{\partial \xi} f_2(\xi_i + d_0(0,0,0),\eta_j)\right)^2 \tag{4.17}$$

 $N_{\omega'}$ はローカルサポート領域W'の中の標本数である.

3次元の adaptive window は以下のアルゴリズムでは立方体の領域になるように制限されている.その幅と高さは-x, +x, -y, +yの 4方向に独立してコントロールされている.

1. マッチングを行ない, 視差の初期推定 d₀(x, y, t) を求める

- (a) 点 (x, y, t) を中心に 3 × 3 の窓を設定して (4.14) を求める
- (b) x+, x-, y+, y-各方向に窓を伸ばし, それに対する (4.14) を計算
- (c) (4.14) を最小にする方向に窓を伸ばし, (4.14) を増やす方向には拡張を禁止
- (d) これを全ての方向が拡張を禁止されるか,規定の大きさに達するまで繰り返す

2. (4.13) によって推定視差の増加分を計算し,

$$d_{i+1}(x, y, z) = d_i(x, y, z) + \Delta d(x, y, z)$$

の式で推定視差を更新する

3. 以上のプロセスを *d_i*が収束するか、ある繰り返し数に達するまで繰り返す

第5章

シミュレーションによる実験

これまでの提案の確認をするため,ローカルサポート設定箇所選択,ローカルサポート の大きさ決定法,ステレオマッチング,特徴領域追跡,軌跡推定の実験を行った.

5.1 シミュレーション環境



図 5.1: 実験環境

シミュレーション環境は図 5.1のようになる.座標軸は左手系に従って設定されている.

- カメラは x 軸上に平行に配置, 光軸は z 軸に平行
- カメラ間距離 (基線長) b

- 焦点距離 *f*
- 画像平面: カメラから仰角 60 度, 縦横比 3:4

以下で特に言及がないときには,基線長b = 4,焦点距離f = 1のカメラ配置とした.

5.2 ローカルサポート設置箇所選択に関する実験

5.2.1 従来法と提案手法の比較

輝度変化のみからローカルサポートの設置箇所を決定する従来法,輝度変化と輝度の 時間変化から設置箇所を決定する提案手法1,輝度の時間変化とエピポーラ線から設置 箇所を決定する提案手法2をそれぞれ図3.4の時系列画像に適用した.提案手法2では, $(e_x, e_y) = (1.0, 0.0)$ とした.ローカルサポートの大きさは 5×5 の固定とし,従来法では 閾値 500000,提案手法1,2では閾値 50000以上の領域を選択した結果,それぞれ4993 個,270個,4058 個の領域を得た.同じ閾値をとっても,提案手法2では提案手法1より も多くの領域を選択する.

従来法によるローカルサポートの設置箇所と固有値のヒストグラムを図 5.2,5.3に,提 案手法1による設置箇所とヒストグラムを図 5.4,5.5に,提案手法2による設置箇所とヒ ストグラムを図 5.6,5.7に示した.図 5.2,5.4,5.6において,白い四角形で囲まれた領域が 選択されたローカルサポート領域である.

従来法では,画像中からエッジが交差している部分やウィンドウ内の輝度勾配が井戸型 になっている部分が一様に選択されるのに対して,提案手法1,2では移動物体からしか選 択されていない.図5.5,5.7に示す.提案手法1,2のヒストグラムを見ると,提案手法2 のヒストグラムのほうが提案手法1のヒストグラムより広がりが大きなため,提案手法2 からのほうがより多くのローカルサポートが選択される.

34



図 5.2: 従来法によるローカルサポート設置箇所







図 5.4: 提案手法1によるローカルサポート設置箇所







図 5.6: 提案手法 2 によるローカルサポート設置箇所



図 5.7:提案手法2による固有値のヒストグラム

5.2.2 提案手法 1 に対する提案手法 2 の有効性の実験

提案手法 2 のほうが提案手法 1 よりも多くのローカルサポートを選択するのは前節の 実験からも分かる.ここでは,ローカルサポートの設定箇所選択法として提案手法 1,2 を用いて,対応づけを行う実験を行い,その精度を比較した.図 5.9,5.10に示す実験画像 1 は図 5.8 のような配置で,カメラが一定速度 $(v_x, v_y, v_z) = (0.2, 0, 0)$ /frame で右方向に 進んでいるシーンを撮影したステレオ動画像の一部である.



図 5.8: 実験画像 1 の配置図

この実験画像1を用いて,提案手法1と提案手法2で選択された特徴領域を対応づけ させることによって選択手法を比較した.提案手法1,2では小さな固有値はどちらも領 域内の輝度の時間変化がないときには0になり,時間変化があるときはそれに応じた大き さをとる.そのため,提案手法1,2で閾値を0として特徴領域を選択した.提案手法1, 2によって選択されたローカルサポートをそれぞれ図5.11,5.13に白い四角形で示した.ま た,選択されたローカルサポートを右画像と対応づけ,視差の誤差が±5 画素以下のとき 対応づけ成功とし,その部分をそれぞれ図5.12,5.14に白い四角形で示した.



図 5.9: 実験画像 1(左画像)



図 5.11: 提案手法 1 による選択領域



図 5.10: 実験画像 1(右画像)



図 5.12: 提案手法 1 による対応づけ成功部



図 5.13: 提案手法 2 による選択領域



図 5.14: 提案手法 2 による対応づけ成功部

画素数	提案手法 1	提案手法 2
選択箇所	1080(0.36%)	14088(4.7%)
非選択箇所	296121	283113
探索範囲全体	297201	297201

表 5.1: 提案手法 1 と 2 の特徴選択数の比較



図 5.15: 提案手法 1 による固有値のヒストグラム



図 5.16: 提案手法 2 による固有値のヒストグラム

固有値のヒストグラムを示した図 5.15, 5.16から,提案手法2のほうがより多くの領域 で閾値以上の固有値を示すため,提案手法1より多くの領域を選択している.表 5.1,5.2

画素数	提案手法 1	提案手法 2
選択箇所	1080	14088
一致箇所	926	12038
成功率	85.7%	85.4%

表 5.2:提案手法1と2のステレオマッチング精度比較

より,提案手法2は提案手法1の10倍以上の点を選択しているのにもかかわらず,ステレオマッチングに関しては提案手法1と同程度の性能を得た.

5.3 ローカルサポートの大きさ決定法に関する実験

5.3.1 ローカルサポートの大きさによる測定精度の変化

図 5.9,5.10に示す実験画像 1 を用いて,固定サイズの窓の大きさを変化させ,それによる測定精度の変化を調べる実験を行った.大きさは $3 \times 3,5 \times 5,7 \times 7,9 \times 9,11 \times 11$ として実験を行なった.これらの実験では,全て閾値 0 以上の領域を選択した.ここでは,ローカルサポート設置箇所の選択は提案手法 1 で行い,ローカルサポートの大きさは 5×5 の固定型とした.対応づけは提案手法で行った.その結果を表 5.3に示す.

特徵領域	選択箇所	一致箇所	成功率
3×3	584	474	81.2%
5×5	1080	926	85.7%
7×7	1696	1480	87.3%
9×9	2440	2181	89.4%
11×11	3246	2450	75.5%

表 5.3: 窓の大きさと測定精度の関係

ただし,視差の誤差が±5 画素以下のとき対応づけ成功とした.

これらのローカルサポートは同じ特徴を指していると考えられる.選択数が異なるの は、同じ1つの特徴からでも窓を大きくしていくと、より多くのローカルサポートを選 択することができるからである.対応づけの精度は窓が小さ過ぎでも大き過ぎても低下す る.このことから各ローカルサポートにふさわしい大きさを適応的に設定する必要性があ ることがわかる.

5.3.2 大きさ固定の窓と adaptive window の比較実験

図 5.17のような配置の図 5.18, 5.19に示す実験画像を用いて, adaptive window と固定 式ウィンドウの精度比較を行なった.

adaptive window は窓の中で視差のばらつきのある実画像を想定しているため,背景が 均一の場合はあまり有効ではない.また,背景方向へウィンドウが広がると,背景とミス マッチングを起こすこともあった.そのため,ここでは左右画面背景に [0,255] の異なる 一様乱数をかけ,背景方向へのウィンドウの拡張を抑えた.

特徴点の選択手法は式 (3.12) から最小の固有値を計算し, 閾値が0以上になるところ を全て選択し, 固定式と adaptive window によってウィンドウの大きさを設定した.固定 式のウィンドウの大きさは11×11でとり, adaptive widow は画像上で初期状態3×3か ら最大11×11の大きさをとり, 用いるフレームは初期状態の3枚から最大5枚になるよ うにした.



図 5.17: 実験画像 2 の配置図

ただし,視差の誤差が±5 画素以下のとき対応づけ成功とした.



図 5.18: 実験画像 2 (左画像)



図 5.19: 実験画像 2 (右画像)

	固定式	adaptive window
選択箇所	2512	2512
一致箇所	1202	1488
成功率	47.9%	59.2%

図 5.20: adaptive window における性能向上

5.4 ステレオマッチングに関する実験

従来法では,選択された特徴領域と近いパターンを持つ領域を対応する画像上からの SSD を計算することによって求める.提案手法ではローカルサポート領域を時間方向に 拡張し,立体的なローカルサポート領域を体積で比較することによって輝度パターンが類 似した運動物体と背景の間での誤対応が回避でき,対応づけの精度が向上すると考えら れる.

5.4.1 ステレオマッチング法の精度比較

従来のステレオマッチング法と提案したステレオマッチング法について,比較実験を行 なった.

ステレオマッチングの性能を比較するため領域の選択は画像 5.9からエッジを含む領域 を全て選択するという同じ方法で行い,従来法と提案手法でステレオマッチングを行なっ た.選択された領域は図 5.21に白い四角形で示した.

ただし,視差の誤差が±5 画素以下のとき対応づけ成功とした.

図 5.225.23表 5.4に示したとおり対応づけの成功率は明らかに向上している.従来法で は区別できなかった深さの差を提案したステレオマッチング法では区別し,ミスマッチン グを減少させることが可能となった.



図 5.21: 選択領域





図 5.22: 従来法による対応づけ成功部 図 5.23: 提案手法による対応づけ成功部

	従来法	提案手法
選択箇所	14880	14880
一致箇所	7631	12278
成功率	51.2%	82.5%

表 5.4: ステレオマッチング法の比較

5.4.2 同一パターン区別の実験

索範囲であるエピポーラ線上に視差の異なる同様なパターンが存在するときにそれを区別 できるかの実験を行なった.図5.25,5.26に示す実験画像3は図5.24のような配置で,Y軸に 水平に配置された左右2つの物体がそれぞれ速度 $(v_x, v_y, v_z) = (0.3, 0, 0)/\text{frame}, (v_x, v_y, v_z) = (0.1, 0, 0)/\text{frame}$ で右方向に進んでいるシーンを撮影したステレオ動画像の一部である.2 つの球は同じ物体色であるが,異なる速度で画面右方向へ移動している.これは,近景と 遠景に同一パターンが存在するときの状況を想定している.



図 5.24: 実験画像 3 の配置図



図 5.25: 実験画像 3(左画像)



図 5.26: 実験画像 3(右画像)



図 5.27: 従来法で選択された特徴領域



図 5.28: 提案手法で選択された特徴領域

従来法と提案手法1による固有値のヒストグラムを図 5.29,5.29に示した.ここで,従 来法では固有値が1000000,提案手法1では固有値が50000以上の領域を選択した.ロー カルサポートの大きさは5×5の固定とした.ただし,視差の誤差が±5画素以下のとき 対応づけ成功とした.

図 5.27,5.27 と表 5.5より,従来法では異なる速度の2物体を区別することができず.右の球と左の球の間で誤対応を生じているのに対し,提案手法では右の球と左の球の区別が可能になっていることがわかる.



図 5.29: 従来法による固有値のヒストグラム

画素数	従来法	提案手法
探索範囲全体	297201	297201
選択箇所	1980	2023
一致箇所	1194	1988
成功率	60.3%	98.2%

表 5.5: 同一パターンの識別性能比較



図 5.30: 提案手法 1 による固有値のヒストグラム

5.5 特徴領域追跡の実験

5.5.1 2次元と3次元の対応づけの比較

2次元のローカルサポート領域と3次元のローカルサポート領域を用いた運動推定の比較を行なった.運動追跡の性能を比較するため領域の選択は実験画像15.9からエッジを含む領域を全て選択するという同じ方法で行い,従来法と提案手法で運動推定を行なった. ローカルサポートは 5×5 の固定型そのうち変位の真値と変位の推定値が誤差なく一致した部分を図5.32,5.33に白い四角形で表し,その結果を表5.6に示した.ここで,対応づけの誤差 Δx , Δy が5以下になったとき「一致箇所」とし,誤差なく一致したものを「完全一致箇所」とした.



図 5.31: エッジのある領域



図 5.32: 従来法による運動推定成功部



図 5.33:提案手法による運動推定成功部

また,

誤差
$$1 = \frac{1}{N} \sum_{N} \sqrt{(x_t - x_e)^2 + (y_t - y_e)^2}$$

誤差 $2 = \frac{1}{N} \sum_{N} \frac{\sqrt{((x_t - x_e)^2 + (y_t - y_e)^2)Z_p^2}}{Z_c^2}$

画素数	従来法	提案手法
選択箇所	14880	14880
完全一致箇所	16	570
一致箇所	272	1901
誤差 1	69830	1799
誤差 $2(Z_c = 20)$	202650	6126

表 5.6: 運動推定性能の比較

対応づけの誤差: $(\Delta x, \Delta y) = (x_t - x_e, y_t - y_e)$ 標本画素数: N 変位の真値: $v_t = (x_t, y_t)$ 変位の推定値: $v_e = (x_e, y_e)$ 深さ: Z_p ある深さ: Z_c

とした.誤差1は,1画素あたりの対応づけの誤差の二乗和を表している.しかし,視差 が大きい画像手前側と視差の小さな画像奥側では同じ対応づけの誤差でも推定に与える 影響が異なる.そのために誤差2では,ある一定の深度 Z_cにならした対応づけの誤差を 示した.これらの誤差を対応づけの精度の目安とした.表 5.6,図 5.32,5.33 よりローカル サポートを拡張したほうが対応づけの精度が向上していることがわかる.

5.5.2 比較実験 2

従来法の組み合わせと提案手法の組み合わせを比較するために,図 5.9,5.10に示される 実験画像1において,図 5.7に示した選択手法,対応づけ手法の組み合わせで特徴領域追 跡の実験を行った.表 5.8より提案手法の組合せのほうが従来法の組み合わせよりも精度 の高い対応づけができていることがわかる.

	特徴点選択	対応づけ
手法 1(従来法)	従来法	2次元
手法2(提案手法)	提案手法 1	3次元

表 5.7: 手法の組み合わせ

画素数	手法 1	手法 2
選択箇所	988	988
完全一致箇所	14	503
一致箇所	61	880
誤差 1	213	17
誤差 $2(Z_c = 20)$	7526	465

表 5.8: 手法 1 と手法 2 の性能比較

5.6 軌跡推定の実験

図 5.8の配置に従う図 5.34~5.37 で表される実験画像から従来法と提案手法で軌跡の推定を行なった.実験に用いた従来法と提案手法の手法の組み合わせは表 5.9に示す.これ

	特徵点選択	対応づけ
手法 1(従来法)	従来法	2次元
手法2(提案手法)	提案手法 1	3次元

表 5.9: 従来法と提案手法





図 5.34: 実験画像 (左画像:1フレーム目) 図 5.35: 実験画像 (右画像:1フレーム目)





図 5.36: 実験画像 (左画像: 10 フレーム目) 図 5.37: 実験画像 (右画像: 10 フレーム目)

らの手法によって得られた対応づけから各フレームでの3次元位置をプロットして,3次 元軌跡を推定した.従来法によるものを図5.38,5.39に提案手法によるものを図5.40,5.41 に示す.





図 5.38: 推定軌跡 (2 次元)

図 5.39: 推定軌跡 (2 次元: xy 方向)



図 5.40: 推定軌跡 (3 次元)



図 5.41: 推定軌跡 (3 次元: xy 方向)

評価法

この例の場合は 360 個程度の特徴領域を選択しているので,復元した 3 次元位置から軌跡を描いても差がわかりにくい.そこで,この例の場合は画面上の物体は全て等しい速度 $(v_x, v_y, v_z) = (-0.1, 0.0, 0.0)$ で動いているので,全特徴点の変位ベクトルの平均値をとって,変位の真値と比較し,図 5.42に示した.選択された特徴点の中で誤対応が多くなれば,真の変位ベクトルとの差が大きくなってゆく.また,変位ベクトル全体の分散 V についても比較した.変位ベクトルの(x, y, z)成分の分散を図 5.43,5.44,5.45にそれぞれ示した.分散 Vは,

$$V = \sum_{N} \frac{(\bar{v} - v_e)^2}{N} \tag{5.1}$$

である.図 5.42に変位ベクトルの真値と従来法,提案手法で推定した変位ベクトルの平均値を示した.提案手法のほうが,真値に近い推定を行っている.



図 5.42: 変位ベクトルの平均値



図 5.43: 変位ベクトルの全体の分散 (x 方向)



図 5.44: 変位ベクトルの全体の分散 (y 方向)



図 5.45: 変位ベクトルの全体の分散 (z 方向)

第6章

まとめ

6.1 結論

本稿では,時間方向へのローカルサポート領域の拡張し,画像中から拡張されたローカ ルサポート領域の設置箇所を選択する手法とそのローカルサポート領域の大きさを設定 する方法について提案した.シミュレーション画像を用いて実験を行い,以下に示すよう な結果を得た.

- 従来は1つの画像平面上にとられていたローカルサポート領域を,時間的に連続した画像上に拡張し,それを立体で対応づけることによって,異なった速度で移動する同様なパターンなどの誤対応を減少させ,対応づけの精度を向上させることができる.
- その際に,画像上でどの位置にローカルサポート領域を置くとよいかについては第 3章で述べた,行列式(3.12)による選択法は非常に厳しい条件のため選択数が少な くなるが,行列式(3.14)による選択法だと行列式(3.12)よりも選択数を多くし,な おかつほぼ同程度の測定精度が得られる.
- 選択されたローカルサポート領域の大きさをどのように設定するかについては第3 章で述べた.この方法を用いることによって,適応的に窓の大きさを変化させることができる.
- ステレオマッチングの対応づけの精度が向上した要因としては、
 - 類似したパターンで,異なる速度で動いている領域の区別が可能.

• ノイズに対する耐性が向上.

が挙げられる.また,運動追跡の対応づけの精度が向上した要因としては,

• 特に等速運動する物体に対しては対応づけが安定する.

• ノイズに対する耐性が向上.

が挙げられる.

6.2 今後の課題

adaptive window によるローカルサポート領域の設定法は,領域内の視差のギャップを 避ける方向で動くため,画像全体から depth map を作るなど画像全体から視差を求める ときには誤差を減少させることができる.しかし,本研究の特徴領域選設置箇所選択手法 は視差のギャップがある領域を特徴領域として選ぶため,ウィンドウの初期状態が物体側 に寄っているときは物体側にウィンドウが拡張し,ウィンドウの初期状態が背景側に寄っ ているときは背景側にウィンドウが拡張するという動きを見せた.そのため,背景が一 様あるいは一定のパターンを持つような状況でウィンドウが背景側に拡張したときには, ウィンドウ内に背景側の画素を多く含むことになり,背景と誤対応を起こすという現象が 見られた.このような一様あるいは一定のパターンを持つ背景では,固定型のウィンドウ よりもマッチングに対する精度が落ちることもあった.

この問題の解決策としては,

1. 特徴領域の設置箇所選択の段階で,背景側に寄った部分を選択しないようにする.

2. ウィンドウが背景側に拡張しないように, adaptive window を制御する.

という方法が考えられる.

また,3次元軌跡を正確に求めるために,運動物体からある程度の数の特徴領域を選択 し,その運動ベクトルの平均と分散から精度の評価を行ったが,複数個の特徴領域から運 動ベクトルの平均と分散を求める方法は,画面内の運動物体の運動ベクトルが一定のとき には有効だが複数の運動物体が異なる運動ベクトルで運動している場合は得られた複数 の運動ベクトルを運動物体ごとにクラスタリングする必要が生じてくる.

謝辞

本研究を行うにあたり,日頃から全般に渡り熱心な御指導,御鞭撻を頂きました阿部 亨助教授に深く感謝致します.

本研究に関して,多くの御討論,御助言を頂きました堀口 進教授に心から感謝致します. サブテーマにおいて,熱心に御指導を頂きました下平 博教授に深く感謝致します. 最後にゼミは多くの御討論をして頂き,また日常においてもいろいろとお世話になりました堀口・阿部研究室の皆様に感謝致します.

参考文献

- Calro Tomasi, Takeo Kanade, "Detection and Tracking of Point Features", CMU-CS-91-132, April 1991.
- [2] Takeo Kanade, Masatoshi Okutomi, "A Stereo Matching Algorithm with an Adaptive Window", IEEE PAMI, VOL.16, NO.9, September 1994.
- [3] Jianbo Shi, Carlo Tomasi, "Good Features to Track" IEEE CVPR94, June 1994.
- [4] Toshihiko Motita, Takeo Kanade, "A Sequential Factrization Method for Recovering Shape and Motion From Image Streams" IEEE PAMI, VOL. 19, NO.8, August 1997.
- [5] Jen-Yu Shieh, Hanqi Zhuang, "Motion Estimation from a Sequence of Stereo Images : A Direct Method" IEEE TRANSACTIONS ON SYSTEMS, VOL.24,NO.7,July,pp.1044-1053 1994.
- [6] Zhengyou Zhang, Oliber D.Faugeras, "Finding Planes and Clusters of Objects from 3D Line Segments with Application to 3D Motion Determination", CVGIP, VOL.60,NO.3,November,pp.267-284,1994.
- [7] Behzad Kamgar-Parsi, Behrooz Kamgar-Parsi, "Matching Sets of 3D Line Segments with Apprication to Polyagonal Arc Matching", IEEE PAMI, VOL.19, NO.10, October, pp.1090-1099, 1997.
- [8] 鄭 文涛,鹿喰 善明,田中 豊,湯山 一郎,"時空間 3 次元ブロックマッチングに おける動画像の動きべ クトル場測定"電子情報通信学会論文誌 (D-II), vol.J81-D-II,NO.8,pp1736-1743,Aug 1998.
- [9] 岡田隆三,白井喜明,三浦純,久野義徳"オプティカルフローと距離情報に基づ く動物体追跡"電子情報通信学会論文誌 A,vol.J79-A,NO.10,pp1-8,Oct 1996.

- [10] 田縁 英治, "複数のカメラを用いた3次元移動物体追跡に関する研究" 北陸先端科 学技術大学院大学修士論文,1997.
- [11] Jakie Neider, Tom Davis, Mason Woo, "Open GL Programming Guide" Addison-Wasley Publishing Company (1995).
- [12] Open GL Archtecture Review Board, "Open GL Reference Manual" Addison-Wasley Publishing Company (1993).

研究業績

[1] 時空間の特徴を用いたステレオ動画像による物体の3次元軌跡推定
 三宅章太郎,阿部亨,堀口進
 電気関係学会北陸支部連合大会,F-50,pp.354 (1998).