

TR-H-115

物体の運動と形状の3次元復元と
操作のための体系的計算理論

志沢 雅彦

1994. 12. 26

ATR 人間情報通信研究所

〒619-02 京都府相楽郡精華町光台 2-2 ☎07749-5-1011

ATR Human Information Processing Research Laboratories

2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02 Japan

Telephone: +81-7749-5-1011

Facsimile: +81-7749-5-1008

物体の運動と形状の3次元復元と
操作のための体系的計算理論

志 沢 雅 彦

序文

本論文の目的は、運動視およびステレオ視によって、3次元世界の立体形状と物体あるいは観測者の3次元運動を計算・表現するための基本的な数学的枠組を体系的に与えることにある。カメラという人工の網膜に投影された画像から3次元物理世界を復元し、その復元データを効率的に操作することは、コンピュータ・ビジョンと幾何情報処理の中心的課題である。この問題に関しては、1960年代から人工知能研究の一環として多くの試みが行われてきた。しかし、現在の技術では、極めて限定された環境下の特殊な場合でなければ人間の持つ高度で柔軟な視覚機能には遠く及ばない。その理由として次の2点があげられる。

1. いままでのコンピュータ・ビジョンの理論は、極めて単純化した環境を前提条件にしていた。複雑な環境を扱う際には、多くのマジックナンバーを含む場当たりのアルゴリズムを用いざるをえない。そして、そのマジックナンバーは入力された画像に強く依存し、著しく柔軟性に欠けている。
2. 計算機がアルゴリズムのシミュレーションの手段としてしか利用されていない。3次元情報などを格納、変換、更新するための真に実地的なアルゴリズム、ハードウェアの設計原理がない。

一方、人間は、極めて巧妙な方法で、ビジョン以外の感覚との統合も含め、モーション、ステレオ、テクスチャ、カラー、陰影、表面復元などの複数の計算モジュールの協調的な動作によって、3次元物理世界の属性を推定している。また、3次元の物体情報を記憶、想起し、仮想的に運動させる能力(例えば、メンタルローテーション)は、人間の3次元情報表現が、実際に人間の知的能力の重要な部分を担っていることを示している。このような状況にかんがみ、筆者は、これらに、理論的な設計原理を与えるべく研究を行ない、それぞれに基本的な理論的貢献を行なった。本論文で述べる内容の従来理論に対する貢献と新規性は、次の2点に要約される。

1. 複数の運動、透明性をもつ物体による運動透明視・ステレオ視、複数の重なり合った物体の3次元構造、物体表面の復元に共通な数学的枠組を発見したこと。それを用いて既存理論を拡張し、理論的に複雑な環境を扱う手段を見出したこと。

2. 物体の形状データに対して運動あるいは座標変換をデジタル計算によって行なう際に、計算誤差の存在に関わらず、原データを厳密に復元可能な演算方式を提案したと。

以下に、本論文の構成を述べる。

第1章では、計算機に視覚の機能を与えるためのコンピュータビジョンの研究を概観する。まず、視覚研究の方法論として、Marrの提唱した3段階の水準を説明し、その中で本研究の位置付けを述べる。さらに、本研究を含む周辺技術について、その工学的応用の可能性について述べる。

第2章では、動画像処理の基本であるオプティカルフローの計算理論を演算子形式の重ね合わせの原理に基づいて拡張する。複数の運動を画像の各点に同時に検出する多重オプティカルフローの計算理論とアルゴリズムが展開される。次に、2眼ステレオ視の基本処理である視差検出の理論を拡張し、多重視差の計算理論を導く。理論を展開して、2重視差を一撃的に計算するアルゴリズムとそのネットワーク表現を導出する。最後に、多重情報に関する基本拘束方程式を導出・表現する統一的方法を提案する。これは、演算子形式の重ね合わせの原理の数学的一般化である。

第3章では、拡張された重ね合わせの原理を用いて、運動からの構造復元の基本問題を拡張し、複数運動を直接扱う計算理論を導く。この理論を用いて、2個の独立な物体の3次元運動と3次元構造を同時に数値的に計算するアルゴリズムを導く。次に、ビジョンにおける不良設定問題を解くための一般的枠組である標準正則化理論を重ね合わせの原理を用いて多価関数の正則化に拡張する。拡張された正則化理論を用いて、重なり合った複数枚の表面を同時に復元する超並列アルゴリズムを導く。

第4章では、計算機内部で、物体の形状や位置を表す座標値やベクトル量、テンソル量の座標変換（剛体運動、アフィン変換、射影変換）を量子化誤差の存在に関わらず、完全に可逆に行う理論を提案する。

第5章は、結論であり、本研究のまとめと今後の課題を述べる。

目次

序文	i
図一覧	vii
1 序論	1
1.1 研究の背景と目的	1
1.2 コンピュータ・ビジョンにおける3次元物理世界の復元とその応用分野	3
1.2.1 コンピュータ・ビジョンにおける3次元物理世界の推定・復元のシナリオ	3
1.2.2 コンピュータ・ビジョンの工学的情報処理技術としての応用分野	5
1.3 モジュールの統合から理論的統一への試み	6
1.3.1 視覚モジュールの統合	6
1.3.2 理論的統一への試み	8
1.4 研究成果の概要	9
2 重ね合わせの原理に基づいた多重オプティカルフロー・多重視差抽出の計算理論	11
2.1 多重オプティカルフローの計算理論	12
2.1.1 はじめに	13
2.1.2 オプティカルフローの基本拘束方程式と運動パタンの重畳	15
2.1.3 多重オプティカルフローの基本拘束方程式	20
2.1.4 多重オプティカルフローの計算アルゴリズム	23
2.1.5 多重度の決定	27
2.1.6 ガボール関数による画像表現と計算機シミュレーション	29
2.1.7 まとめ	38
2.2 多重ステレオ視差の計算理論	39
2.2.1 はじめに	40

2.2.2	ステレオ透明視のための多重視差推定の形式的計算理論	43
2.2.3	2 重視差の準線形アルゴリズムと一撃計算モデル	46
2.2.4	シミュレーション実験	50
2.2.5	まとめ	55
2.3	重ね合わせの原理の一般表現	56
2.3.1	はじめに	57
2.3.2	演算子とデータ分布関数による基本拘束方程式の一般的記述	57
2.3.3	振幅演算子を用いた重ね合わせの原理の数学的一般表現	60
2.3.4	属性値を計算するアルゴリズムの導出手順	64
2.3.5	まとめ	67
3	多重 3 次元運動・構造復元と多重 3 次元表面復元の計算理論	69
3.1	多重 3 次元運動・構造復元の計算理論	70
3.1.1	はじめに	71
3.1.2	複素運動の推定に関する従来の研究	71
3.1.3	多重 3 D 運動の基本拘束方程式の導出	73
3.1.4	2 重 3 D 運動の準最適な線形アルゴリズム	80
3.1.5	計算機シミュレーションによる検証	89
3.1.6	まとめ	95
3.2	多価関数への標準正則化理論の拡張となめらかな多重表面復元	96
3.2.1	はじめに	97
3.2.2	多価標準正則化理論による 2 重表面の復元	99
3.2.3	多価標準正則化理論の一般論	112
3.2.4	まとめ	116
4	物体運動の効率的計算機表現のための可逆座標変換法	117
4.0	背景	118
4.1	整数座標における可逆座標変換 (1) — 非同次座標における等積アフィン変換	
—	119
4.1.1	はじめに	120
4.1.2	物体形状の記述と座標変換	120
4.1.3	アフィン変換とその部分群	122
4.1.4	等積アフィン変換の可逆整数近似	124
4.1.5	n 次元における一般の場合の分解アルゴリズム	130

4.1.6	2次元1次変換の誤差最小分解	133
4.2	整数座標における可逆座標変換(2) — 同次座標における非特異射影変換 —	138
4.2.1	はじめに	139
4.2.2	3次元射影変換と4次元等積1次変換の関係	139
4.2.3	非特異射影変換に対する可逆整数近似変換の構成法	142
4.2.4	各種幾何要素の変換	144
4.3	まとめ	147
5	結論と展望	149
	謝辞	153
	参考文献	155
	序論に関連する参考書	155
	多重オプティカルフロー	156
	多重ステレオ視差	161
	多重3D運動・構造復元	162
	標準正規化理論の拡張による多重表面の復元	164
	可逆座標変換	165
	参考とした一般書籍	166
	原著論文目録	169
	学術論文誌	169
	国際会議(コンファレンス)	170
	国際会議(シンポジウム, ワークショップ)	171
	その他の共著論文	171
	研究会, 全国大会	172
A	第2章の付録	175
A.1	多重オプティカルフローに関する付録	175
A.1.1	多重オプティカルフローの基本拘束方程式の証明	175
A.1.2	多重オプティカルフロー推定のための連立非線形固有値方程式の導出	175
A.1.3	多重オプティカルフロー推定のための線形固有値方程式の導出	176
A.1.4	$n = 1$ および $n = 2$ の場合の線形固有値方程式の導出	176
A.1.5	混合フローテンソルの2個のフローベクトルへの分解	178

A.1.6	ガボールフィルタ出力からのエネルギーモーメントの計算	178
A.1.7	3 フレームからの 2 運動推定アルゴリズムとの関係	180
A.2	多重ステレオ視差に関する付録	183
A.2.1	単一視差の場合の拘束と従来法との関連	183
A.2.2	2 重視差の振幅演算子の近似演算子の導出	183
A.2.3	周波数領域における拘束方程式	184
B	第3章の付録	185
B.1	多重 3 次元運動・構造復元に関する付録	185
B.1.1	基本テンソル推定アルゴリズムの導出	185
B.1.2	基本行列の単位行ベクトル、単位列ベクトルの計算アルゴリズム導出	186
B.1.3	係数 ζ_i, η_j の計算アルゴリズム導出	187
B.1.4	基本行列の並進運動ベクトルと回転行列への分解	187
B.2	多価関数への標準正則化理論の拡張に関する付録	189
B.2.1	重ね合わせの原理に基づいた多価関数の直接表現の導出	189
C	第4章の付録	193
C.1	[定理 4.1 の証明]	193
C.1.1	[定理 4.1 の証明]	194
C.2	[命題 4.3 の証明]	195
C.3	[定理 4.3 の証明]	195
C.3.1	[不等式 (C.4) の証明]	196
C.3.2	[不等式 (C.5) の証明]	196
C.4	[命題 4.4 の証明 (略証)]	197
C.5	基本変換の交換関係	197

目 次

1.1	ビジョン研究の3つの水準とその理論的背景および技術要素	2
1.2	コンピュータビジョンの各種機能モジュールと相互関係	4
1.3	モジュール統合による複数画像生成要因の復元の概念 ([Aloimonos & Shulman 1989 [1]] から引用。)	6
1.4	モジュール統合の2つの形：弱結合 (weak coupling) と強結合 (strong coupling) ([Clark & Yuille 1990 [5]] から引用。)	7
1.5	モジュール統合から理論的統一へ	8
1.6	本論文の構成	9
2.1	和の形の重畳による運動透明視の例	17
2.2	積の形の重畳による運動透明視の例	18
2.3	フロー場の不連続は重畳として解釈できる	19
2.4	運動推定とセグメンテーションの新しい関係	20
2.5	2重オプティカルフローの時空間周波数領域における拘束条件	22
2.6	固有エネルギーを用いて最適化問題の安定性=非退化条件を検査する。	28
2.7	1-重、2-重オプティカルフローのためのアルゴリズム	29
2.8	実験1の結果	31
2.9	実験2の結果	32
2.10	実験3の結果	33
2.11	実験4の結果	34
2.12	超並列計算機による多重オプティカルフローのシミュレーション例	35
2.13	純粹の透明視をもたらすランダム輝度ステレオグラム (交差法)	40
2.14	平行2眼ステレオと透明視	41
2.15	ステレオ視差検出のための振幅演算子のはたらき	44
2.16	2重視差に対する一撃アルゴリズムのネットワーク表現	49
2.17	1次元信号を用いたシミュレーションにおける2つの視差マップ	51

2.18	シミュレーションのために生成された2重視差を含む1次元信号	51
2.19	一撃アルゴリズムによる2重視差推定結果(真の視差マップとの重畳表示) . .	52
2.20	2次元のランダム輝度ステレオグラムに対する推定結果。	53
3.1	3次元剛体運動と特徴点对応の投影像の幾何学(単一剛体運動の場合)	74
3.2	シミュレーションのために生成された特徴点对応データ	90
3.3	基本テンソル推定の際の36個の固有値の分布	90
3.4	基本行列の推定結果と真値	91
3.5	基本行列の分解によって得られた各物体の並進ベクトル・回転行列の候補 . .	91
3.6	物体構造パラメータの符号にもとづいて選択された運動パラメータの一意解 とその真値	92
3.7	3次元物体構造の復元結果と真値	93
3.8	方程式が退化する場合の固有値の分布	94
3.9	2重曲面(曲線)によってモデル化されるデータの例	99
3.10	2重曲面上のデータが満たす方程式	100
3.11	2重表面復元の緩和計算を行う超並列型ネットワーク	103
3.12	実験1の結果($\lambda_F = \lambda_G = 6.0$)	106
3.13	実験2の結果($\lambda_F = \lambda_G = 6.0$)	107
3.14	実験3の結果($\lambda_F = \lambda_G = 12.0$)	108
3.15	実験4の結果($\lambda_F = \lambda_G = 12.0$)	109
3.16	実験5の結果($\lambda_F = \lambda_G = 12.0$)	110
4.1	空間占有表現による、3次元物体形状表現法の例 ([Ballard & Brown 1982 [3]] より引用)	121
4.2	一般アフィン変換の部分群の幾何学的意味	123
4.3	基本変換の幾何学的意味(2次元の場合)	125
4.4	基本斜交軸変換の整数近似写像 $\hat{p}_{ij}[s]$ の例	127
4.5	2次元特殊1次変換 SL_2 の場合の例	135
4.6	2次元特殊直交変換 SO_2 の例	137

第 1 章

序論

1.1 研究の背景と目的

人間の視覚は、目というカメラの網膜に投影された 2 次元画像から、3 次元物理世界の幾何属性、物理属性を推測し、認識、行動に結び付ける役割を担っている。計算機に人間の視覚機能に相当する能力を与える研究は、コンピュータ・ビジョン (CV)、マシン・ビジョン、ロボット・ビジョンなどと呼ばれる。'60 年代から、人工知能研究の一分野として、信号処理、画像処理、記号処理、数値計算の手法を採り入れながら発展してきた。

'70 年代後期に、Marr[15] によって、人間の視覚の計算論的研究 (Computational Vision) の研究手段としてのコンピュータ・ビジョンの意味が明らかにされた。Marr は、この視覚機能を脳が行なう情報処理課題と位置付けた。そして、この情報処理課題を研究するために、3 つの相対的に独立な水準からなる方法論を提案した。

1. 情報処理課題を数学的に定式化し、基本的な性質を解明すること (計算理論 (*computational theory*) の水準)、
2. 問題を効率的、ロバストに解くための情報表現およびアルゴリズムを見い出すこと (アルゴリズムと表現 (*algorithm and representation*) の水準)、
3. アルゴリズムを効率的に実行するためのハードウェア構成を見い出すこと (インプリメンテーション (*implementation*) の水準)。

そして、これらの研究を行なう学問である、計算論的視覚 (*computational vision*) を定義した。Marr は、もともと、人間の視覚の研究手段として、この 3 つの水準を指摘した。それと同時に、計算機による視覚機能を工学的に利用するためにも、これらの 3 つの水準を区別して扱う重要性を指摘した。実際、視覚の様な極めて複雑で高度な機能を解明したり、それを計算機によって実現するためには、研究の方法論が重要である。場当たりの (*ad hoc*) な

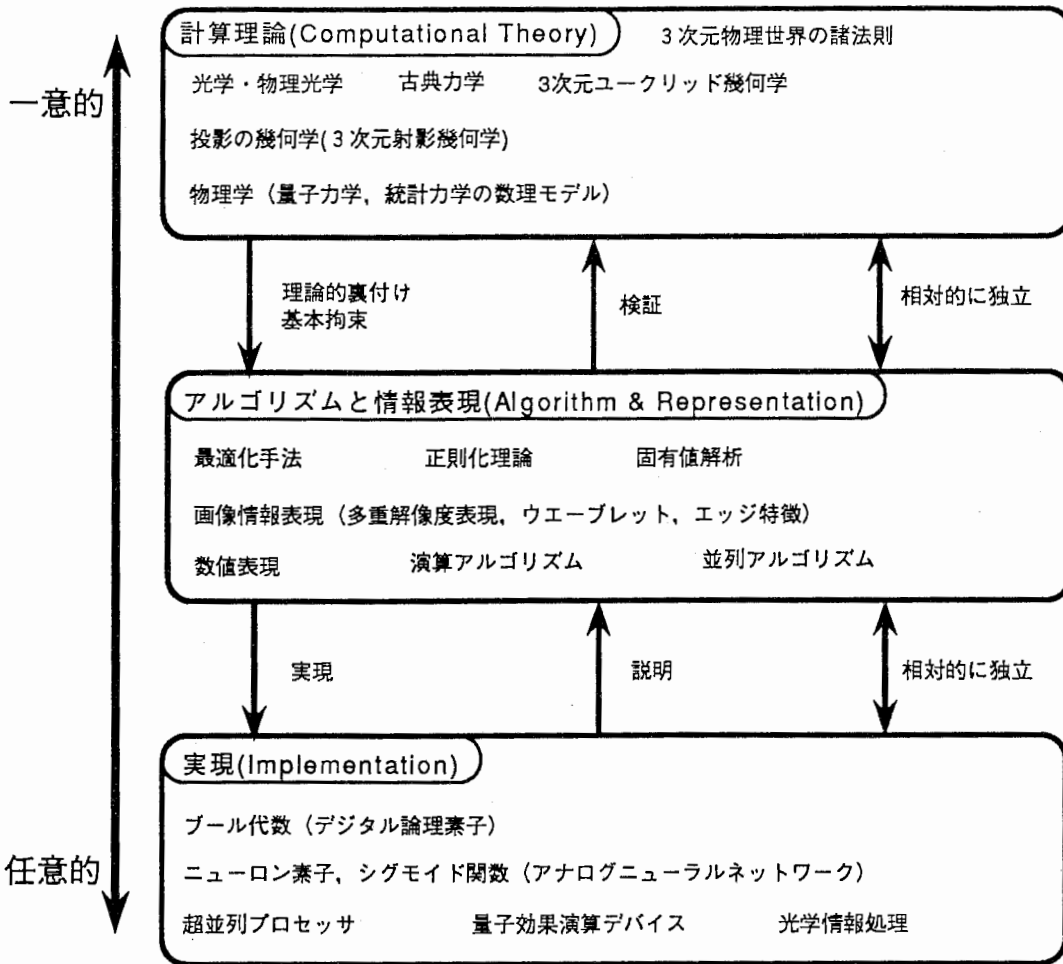


Figure 1.1: ビジョン研究の3つの水準とその理論的背景および技術要素

アプローチでは、複数の要因を統合しなければならない場合に、有効な設計指針が得られない。

人間の視覚系の目的は、網膜への投影像から、環境あるいは、物体の形状、表面属性を直接あるいは、間接的に推定することである。この人間の視覚機能と同等あるいは、それを超える能力を計算機上を実現することが、コンピュータビジョンの研究課題である。つまり、3次元世界の物理法則、光学法則、画像への投影の幾何学、物体運動の力学法則などの基本法則や、人間の視知覚の心理学的研究から得られた経験的拘束条件を用いて、無限に多くの解釈可能性の中から、もっともらしい解釈を抽出する方法を探求する。

図 1.1 は、これら3つの水準に必要な、理論背景、技術要素を概観した図である。

視覚という、網膜あるいはカメラから入力された輝度パターン (以下、これを画像と呼ぶ) から3次元物理世界を推測するという目的は共通であり、したがって、共通に議論できる問題も多い。特に、計算理論の水準では、ほとんど共通に論じることができる。この水準では、物理世界の基本法則によって、画像投影面に表れるパターンの基本的性質を表す拘束

(constraint) が導かれ、その基本的な性質が数学的に検討される。計算理論が、この様に、物理法則と極めて深い関係にあるということから、直接関係がうすい量子力学や統計力学などにおける基本原理や数学、数値計算法が基礎理論に用いられることが多い。本論文においても、波動力学や量子力学の基本原理の一つである重ね合わせの原理を、ビジョンの研究に世界で初めて導入した。

アルゴリズムと表現の水準においても、効率性とロバスト性を要求した場合、本質的に同じ計算原理や情報表現が有効である場合が多い。この水準では、いろいろな最適化手法や、統計的推定法、数値表現法、数値計算法、演算アルゴリズム、並列アルゴリズム、ネットワーク構成など、の設計理論が研究される。

インプリメンテーションの水準では、ブール論理素子(ロジック・ゲート)、ニューロン素子、超並列プロセッサ、光学情報処理、量子効果デバイスなどの計算素子のそれぞれの特長を活かしたハードウェア設計理論、あるいは、ソフトウェアによるコーディング法が研究される。

本論文の内容は、主に計算理論とアルゴリズムの水準における貢献である。

1.2 コンピュータ・ビジョンにおける3次元物理世界の復元とその応用分野

コンピュータ・ビジョンは、物体あるいはカメラ(観測者)の動き、物体の姿勢、物体の形状、物体の表面性状、光源情報などの物理世界の情報を、カメラによって撮影された画像列から推定、復元するための理論、アルゴリズム、ハードウェアを研究する分野である。

1.2.1 コンピュータ・ビジョンにおける3次元物理世界の推定・復元のシナリオ

図1.2は、このコンピュータビジョンの役割を簡単に示した図である。この図は、コンピュータビジョンのすべての側面を示していない。例えば、カメラを実時間で能動的に動かして情報を獲得するアクティブビジョンや、物体にパターンを投影して、その変形から3次元物体の形状情報を獲得するレンジセンサーなどのアクティブセンシングは省いた。コンピュータビジョンでは、カメラ(単眼、複眼)から入力された時系列画像から、3次元物理世界の幾何学的、光学的性質や表面の光学的特性を推定する。以下に、この機能を実現するいろいろな視覚モジュールを列挙する。これらの区分けは、説明の都合上行った。実際には、一つのアルゴリズムに統合されているものも含む。

1. 光源と物体表面の相対的位置関係によって決まる陰影情報から、表面の反射特性や形状情報を推定する「陰影からの表面復元(Shape from shading)」。

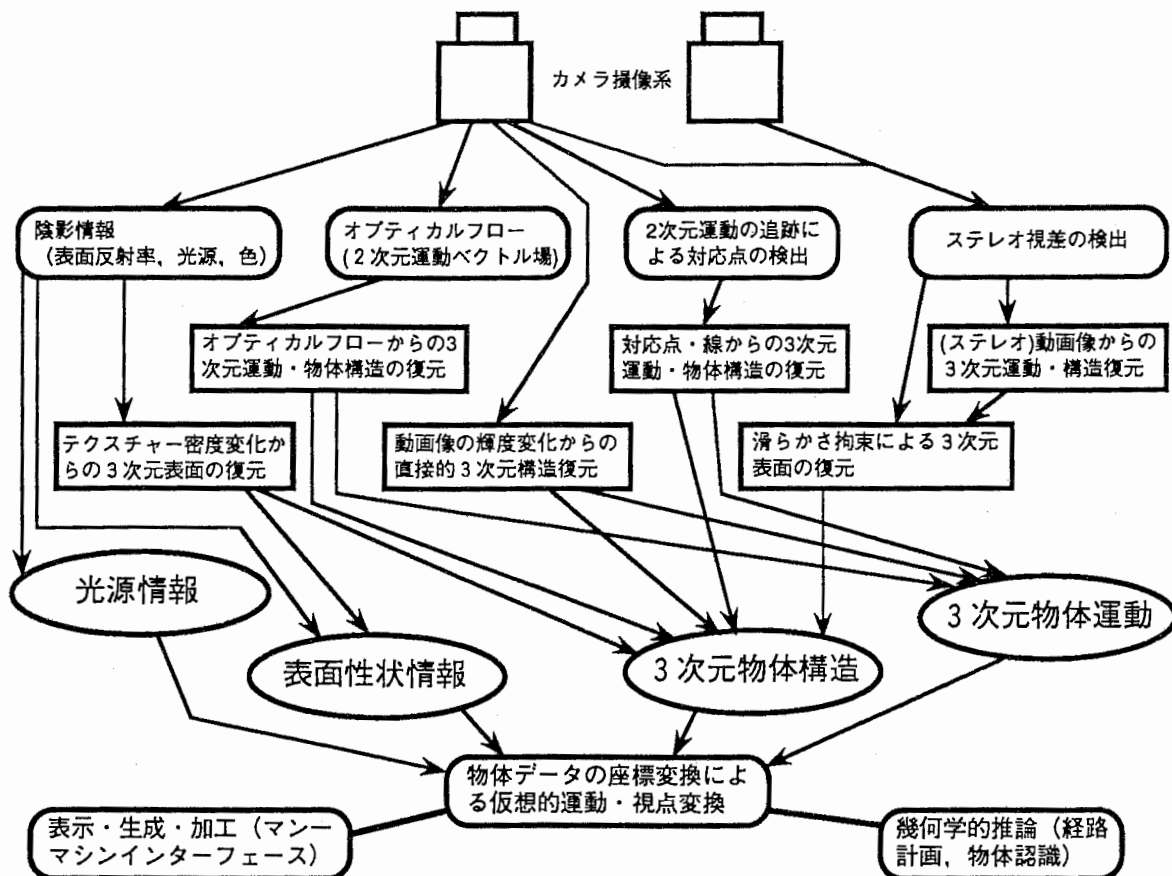


Figure 1.2: コンピュータビジョンの各種機能モジュールと相互関係

2. 物体の運動によって画像面上に生じる見かけの運動ベクトル場である「オプティカルフロー」の計算。
3. エッジ、かどなどの特徴抽出と、その画像フレーム間追跡と対応付け。
4. 両眼ステレオ視における、左右の画像の間のパターンのずれ量(視差)を計算するための対応付け(ステレオ整合)。
5. 点・線分などの特徴のフレーム間対応情報から、剛体の3次元運動と3次元構造を推定する「運動からの構造復元」。
6. オプティカルフローから、剛体の3次元運動と3次元物体構造を推定する「運動からの構造復元」。
7. 動画画像の輝度変化からの直接的3次元構造復元。

8. ステレオ視差データからの3次元構造復元。
9. 物体表面のテクスチャ密度変化からの3次元表面の形状復元。
10. 滑らかさ拘束を用いた疎データからの3次元表面の形状復元。

図 1.2 に、これら視覚モジュール間の相互関係を示した。

1.2.2 コンピュータ・ビジョンの工学的情報処理技術としての応用分野

コンピュータ・ビジョンの分野で培われた技術は、各種ロボットの視覚機能、製品組み立て・検査などの工業応用、仮想現実感 (Virtual Reality) の様な視覚機能を用いたマン・マシン・インタフェース、動画像符号化・圧縮などの画像通信、画像伝送、3次元環境認識・理解、パターン認識などの各分野に共通の基礎理論、基礎技術として応用可能である。

ロボットの視覚機能への応用も重要である。例えば、3次元環境の復元とそのナビゲーションへの応用がある。カメラの移動によって画像面上に生じる物体の運動から環境の構造情報を得たり、動物体の検出を行うことができる。

製品検査では、ベルトコンベアで流れてくる部品の形状と姿勢を復元、認識し、組み立てなどの作業を容易にしたり、データベースとの照合によって欠陥部品を検出する応用が考えられる。

仮想現実感では、3次元物体の形状および表面性状をデータ入力する際の基本技術として、また、物体の運動を仮想的に作り出すための、3次元物体運動の入力技術として、コンピュータビジョンのいろいろな3次元復元手法の応用が考えられる。

知的符号化の様な高度な画像通信のための符号化技術にもコンピュータ・ビジョンは応用される。物体の形状、表面特性および、力学的な性質(不変量)が復元できれば、あとは、運動パラメータなどの非常に少ない情報の伝送によって、自由な視点からの映像を復元することができる。

3次元物体あるいは環境のパターン認識に関しても、視点に依存しない認識が必要であるから、物体の姿勢や運動を求めるためのコンピュータ・ビジョンの理論が応用できる。

コンピュータ・ビジョン自身が、並列処理、特に、近年、注目されつつある超並列処理の大きな応用分野でもある。実際、ビジョンの計算処理には大量のデータのフィルタリングや大規模な繰り返し計算の様に計算量が膨大なものが多い。そのため、最終的には、ハードウェアによる実現が不可欠であり、そのためのインプリメンテーションのレベルの研究がますます重要になってきている。

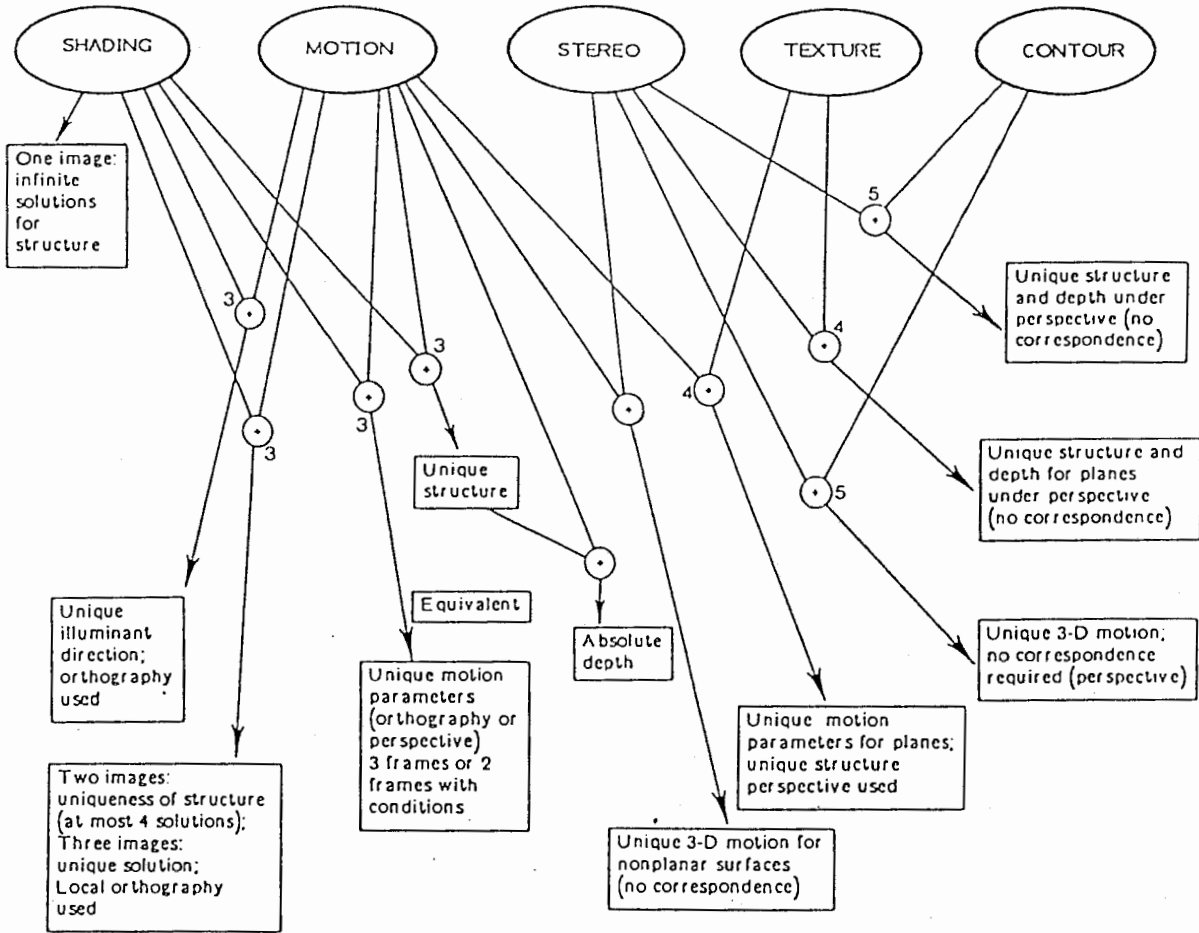


Figure 1.3: モジュール統合による複数画像生成要因の復元の概念 ([Aloimonos & Shulman 1989 [1]] から引用。)

1.3 モジュールの統合から理論的統一への試み

1.3.1 視覚モジュールの統合

従来、2次元画像に重畳された3次元物理世界のさまざまな属性を推定するために、視覚モジュールの統合 (integration) のアプローチが種々試みられてきた。図 1.3[1] は、その一例である。モジュールの結合には、図 1.4[5] に示す弱結合 (weak coupling) と強結合 (strong coupling) と呼ばれる2つの形態がある。弱結合は、比較的容易に実現できるが、複雑に撚り合わされた情報を解析しなければならないコンピュータビジョンでは、限られた有効性しかない。物理世界のいろいろな属性は有限の画像内に複雑に重畳されているため、その推定のためには強結合のモジュール統合が必要になる。しかし、この様に複雑に結合したモジュールの集合体が正しく動作するためには、モジュール間の結合の強度やタイミング調整など非常に多くのパラメータの調整が必要になる。現実には、これは極めて困難な問題

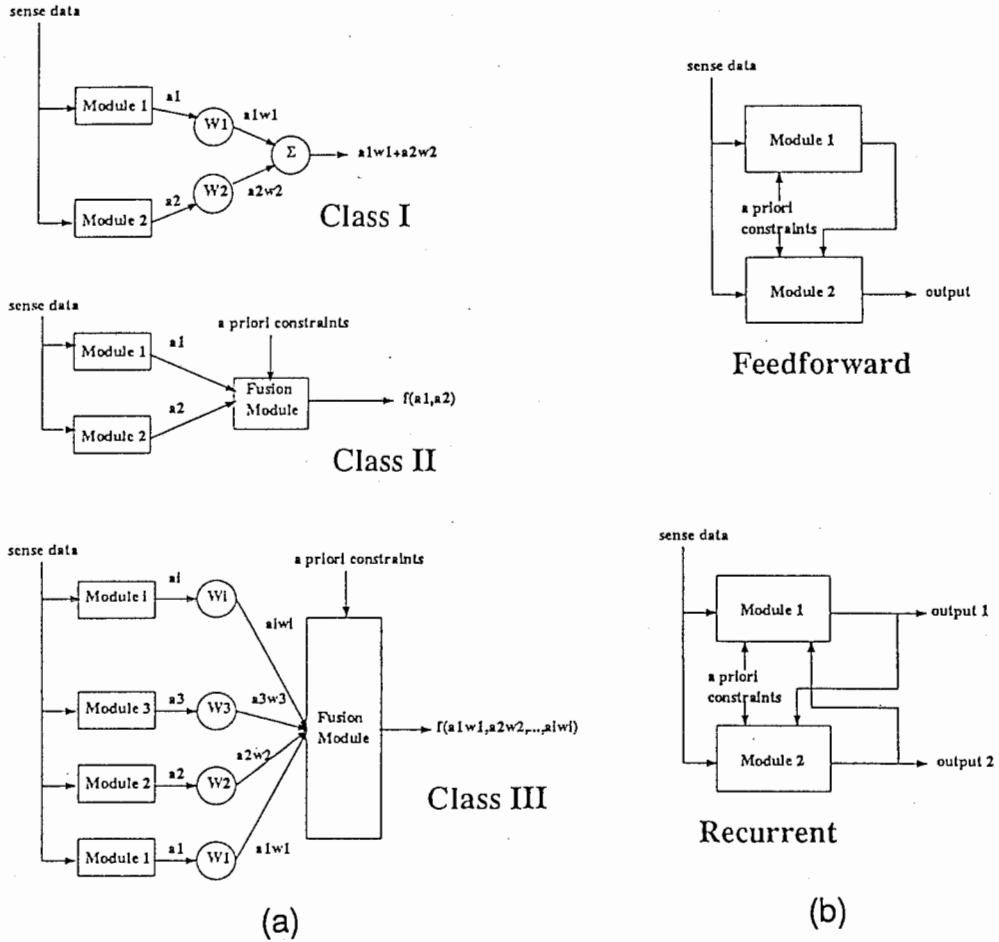


Figure 1.4: モジュール統合の2つの形：弱結合 (weak coupling) と強結合 (strong coupling) ([Clark & Yuille 1990 [5]] から引用。)

(a) モジュールの弱結合による統合。モジュール間の相互作用はない。設計は容易である。(b) モジュールの強結合による統合。モジュール間に相互作用が存在する。設計は容易でない。

である。強結合モジュールの統合理論を確立する必要がある。この様に、強結合のモジュール統合が困難なので、従来は、このアプローチを放棄し、比較的少数の特徴から、特定のタスク（例えば、衝突回避）だけを行う複数の独立なモジュールの集合体を構成することで十分であると主張する Qualitative Vision(定性的視覚), Purpositive Vision のパラダイムも提案されている [2]。しかし、このアプローチによって実現できるタスクは極めて限定されてしまう。最終的に推定の精度が要求される用途では、得られる推定値の精度を評価できなければならない。したがって、定性的視覚は限定された用途しか持たない。

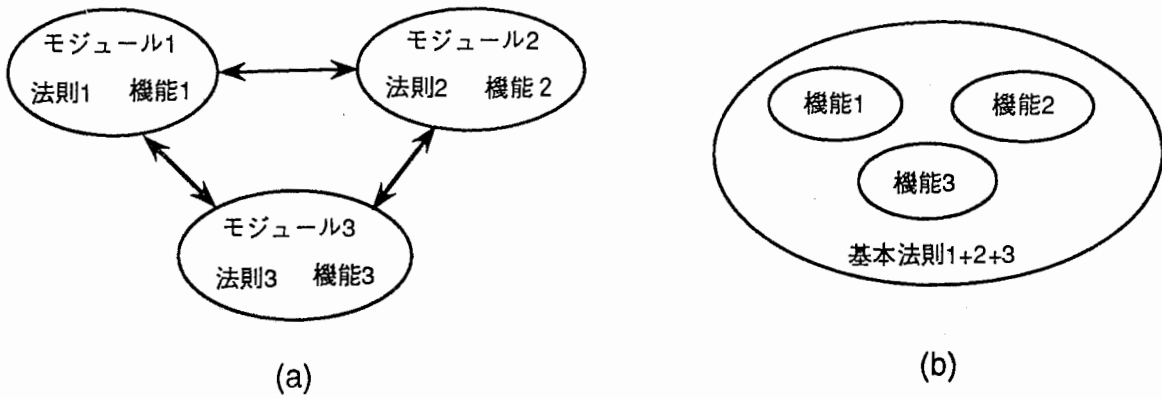


Figure 1.5: モジュール統合から理論的統一へ

(a) 従来のモジュール統合 (強結合) の概念図 (b) 本研究のアプローチ: 法則レベルで基本拘束方程式を統一、拡張する。

1.3.2 理論的統一への試み

本論文では、初期視覚と中期視覚において、既存モジュールの統合を目指すアプローチではなく、拘束方程式のレベルでの統一をはかることを試みた。これは、複数の要因をあらかじめ仮定する。その上で、画像が満たさなければならない直接的な拘束方程式をまず導出する。その方程式を複数の属性値に関して直接解くアルゴリズムを導く。この考え方によって、モジュールの概念を超えて一度に複数の画像生成要因を復元しようとするものである。図 1.5 にその概念図を示す。必要ならば、方程式を部分的に解くことによって、複数モジュール結合による実現方法も導き出すことができる。これは、インプリメンテーションのレベルでモジュール間の結合パラメータを調整する方法に比較して根拠のしっかりした方法であり、本当の意味のモジュール統合といえる。

本論文では、この考え方にもとづいて、運動視、ステレオ視、運動立体視、滑らかな表面復元のための正則化理論に関して、理論を拡張する。多重運動、多重表面の推定が可能にな

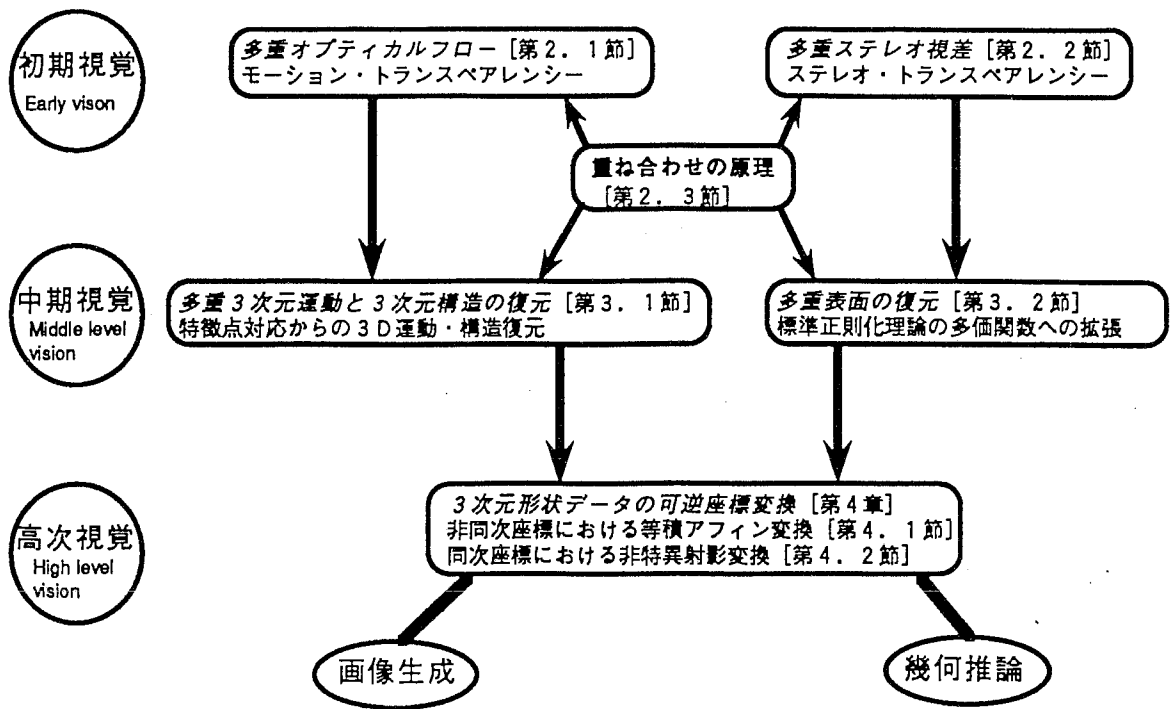


Figure 1.6: 本論文の構成

る。これは、同種の複数属性値の推定への理論の拡張である。したがって、統一の意味からすれば第一歩である。

1.4 研究成果の概要

本研究では、コンピュータビジョンと幾何情報処理に次の貢献を行った。

1. 重ね合わせの原理をビジョンの分野に初めて導入した。まず、初期視覚の基本理論であるオブティカルフロー、ステレオ整合の理論を拡張した。これによって、複数の物体が重なりあって撮像された画像から、複数の重なり合った2次元運動ベクトルとステレオ視差の計算を可能にした。
2. 次に、重ね合わせの原理を中期視覚の基本理論である3次元運動と3次元構造、物体表面復元に応用した。画像フレーム間で計算された特徴点の2次元運動情報から物体

の3次元運動と3次元構造を復元する計算理論を拡張した。これによって、複雑に重なり合った複数剛物体の運動復元を可能にした。さらに、疎らな3次元データから、滑らかな物体表面を復元する標準正則化理論を拡張した。これによって、重なり合った複数表面の復元を可能にした。

3. 物体の形状、構造データに対する仮想運動である、座標変換のデジタル演算法を改良した。計算誤差が存在するにも関わらず、前後で情報の損失の無い座標変換の演算原理を発見した。そのアルゴリズムの具体的構成法を提案した。

本論文は、本章も含め、全体で5章からなる。第2章では、重ね合わせの原理に基づいて初期視覚の基本理論を拡張する。第3章では、中期視覚の基本理論を拡張する。第4章では、幾何データの可逆的座標変換の理論を展開する。第5章は、全体のまとめと将来の課題を述べる。図1.6に本論文の章節構成を示す。

第 2 章

重ね合わせの原理に基づいた多重オプティカルフロー・多重視差抽出の計算理論

あらまし 運動視とステレオ視における初期視覚処理である、オプティカルフローと両眼視差の計算理論を拡張する。画像面の各点に複数のオプティカルフローベクトル、両眼視差を推定するための計算理論を展開する。そこで用いた「重ね合わせの原理」と名付けた数理技法を一般化する。これは、ビジョンの計算理論における多重情報の数理的表現法を与える。

2.1 多重オブティカルフローの計算理論

あらまし 従来のオブティカルフロー計算アルゴリズムは、画像面の各点に単一のフローベクトルが存在するという仮定のもとで成立する基本拘束方程式を基礎においていた。しかし、鏡面性や透明性を持つ物体によって複数の重なり合ったフローが生じる領域や、物体境界などのフローが不連続になる領域では、この基本拘束方程式が成立しない。そのため、従来は、正しいフロー推定が不可能であった。本節では、まず、従来のオブティカルフローの基本拘束方程式を拡張する。画像面の各点に n 個のフローが存在する場合に成立する基本拘束方程式の実領域と周波数領域における表現を導く。次に、多重度 2 の場合の基本拘束方程式を用いて、動画像の Gabor 関数展開表現から 2 組のフローパラメータを閉形式演算で求めるアルゴリズムを導く。さらに、各多重度についての最適化問題の良設定性を表す指標である余裕エネルギーを定義する。これを用いて、各画像点におけるフローベクトル数を決定する基準を与える。さいごに、運動透明視と運動境界を含む合成画像に対してこのアルゴリズムの有効性を確認する。本節で導く多重オブティカルフローの基本拘束方程式は、従来の種々の運動視モデルや動画像解析アルゴリズムを多重運動、運動透明視、運動境界を扱える様に拡張する際の基礎方程式として広く利用可能である。^a

^a本節に関連する内容は、文献 [177] [183] [185] [186] [193] [194] において発表された。

2.1.1 はじめに

人間の運動視・運動立体視の視覚系は、たとえ物体が透明性を持っていたり、複数の物体が重なり合っている場合にも、比較的容易に複数の運動速度、3次元運動と複数の3次元構造を推定することができる [25] [20] [75]。ところが、人間の運動検出モデル、および、工学的な動画像解析手法のいずれにおいても、従来は、画像面の各点に対応する運動が1個だけであるという仮定のもとで成立する拘束方程式をもとにして計算モデルと計算アルゴリズムを構成していた [26] [78] [38] [49] [56] [51] [79] [62] [53] [28]。このため、この仮定が成立しない上記の状況では正しい運動推定が不可能であった。

現実の画像では、以下に述べるさまざまな要因によって画像面上の各点における見かけの運動が単一であるという仮定が成立しない。

1. 不透明な物体表面における拡散性反射パターンと鏡面反射像の運動。
2. 複雑な投射パターンを持った光源に照射された動物体。
3. 窓ガラスなどの透明性物体における透過成分と鏡面反射成分。
4. 木、草、フェンスなどの複雑な物体によるオクルージョン。
5. 汚れた窓ガラスを通して景色を見ている場合 (c.f.[112])。

また、比較的単純な状況においても、物体の運動境界 (Motion Boundary) ではオプティカルフロー場が不連続となり、フロー推定に困難が生じていた。¹

このフロー場の不連続を含め、複数の運動を解析する問題に対して、今までに提案されてきたアルゴリズムを手法によって分類整理する。

1. 速度空間 (u, v) あるいは、3D 速度空間に対する投票あるいは Hough 変換、クラスタリングを用いて、運動境界の検出とセグメンテーションを行った例。[40] [72] [27] [41] [74] [52] [45] [77]。
2. 線過程 (line process) を場の不連続を表す状態変数として導入し、運動パラメータ推定過程との結合マルコフ場 (Coupled MRF) としてモデル化したもの [43] [61][4]、あるいは、このアプローチを簡略化し、輝度エッジ情報などと組み合わせてアナログハードウェアによって実現した例 [47]。

¹一般に、画像面あるいはその一部領域で複数の運動が知覚される現象を運動透明視・運動透明性 (Motion Transparency) という。本論文では、この運動透明視と、運動境界における様に一般にアルゴリズムの解析領域に複数の運動が存在する状況を併せて、多重運動 (Multiple Motion) と呼ぶことにする。

3. 統計的パラメータ推定の一手法であるロバスト統計 [159] にもとづき、複数の運動のうちで1個だけに同調し、他の運動を outlier として排除する重み付け最適化を推定法に導入した例 [32]。
4. 多重解像度空間において、単一運動の推定アルゴリズムが複数の運動のうち的一方に同調する傾向を用いて、シーケンシャルに複数の運動を抽出する方法 [34]。
5. 画像を複数のレイヤーに分割するマルチレイヤー表現を用いてセグメンテーションを確率的エネルギー最小化によって行なう方法 [129]。
6. 運動透明性や複数物体の運動などの2重運動に対し、2組の運動パラメータを割りあてて、一方の画像成分の消去を行った差分画像の生成と他方の運動パタンのパラメータ推定を交互に繰り返して収束値として3フレームから2組の運動パラメータを求める方法。 [30]

この様に多くの試みが行われている。しかし、これらの方法は、すべて、単一運動に関する基本拘束条件にもとづき、それになんらかのセグメンテーション手法を組み合わせて構成されたアルゴリズムである。また、6. 以外の手法は、時空間運動パタンの重ね合わせによって生じる純粋な運動透明視の場合には適用できない。これらの方法は、実験によって決定すべきパラメータが多く、最適化問題のエネルギー関数が非凸になって収束を保証することが困難である。より重要なことは、これらはすべてセグメンテーションというアルゴリズムの水準で複数運動を扱おうとする試みである。したがって、精度やノイズに対する最適性や振舞いを理論的に示すことが困難である。多くの場合に個々の画像に対して実験的にしか確認することができない。画像生成過程においては、和と積の型の(時)空間パタンの重ね合わせとそれらの組合せが生じる。一方のパターンが他方のパターンを被い隠すオクルージョンも重ね合わせの一種と考えられる。

以下は、人間・生物の視覚系における運動視モデル [25] [26] [78] [38] [49] と工学的な動画画像解析の初期処理アルゴリズム [56] [51] [62] [79] [53] [41] の共通の基礎方程式であるオブティカルフローの輝度勾配拘束方程式を多重運動に拡張する。画像面の各点に n 個のオブティカルフローが存在する場合に動画画像が満たさなければならない数学的拘束が導かれる。この基本拘束方程式は、時空間実領域では、単一フローの基本拘束方程式における輝度勾配の拘束を高階偏微分の拘束に拡張したものである。時空間周波数領域では単一フローの共平面拘束を複数平面の拘束に拡張したものである。

この基本拘束方程式を用いて多重フローの推定を行うアルゴリズムには、従来の単一フローについての方法と同様に多くのバリエーションが有り得る。本節では、そのうちの1ア

ルゴリズムを導く。時空間局所最小2乗最適化を用いた方法を示す。これは、動画像の Gabor 展開表現から2組のフローパラメータを同時に閉形式で計算する固有値計算アルゴリズムである。

オプティカルフローを多重運動に拡張した時に生じる問題に、フローの多重度 n の決定がある。これにも、数種の方法が考えられる。本節では、最小2乗最適化の安定性、すなわち、フローパラメータ推定問題の良設定性 (well-posedness)²を n 重フローモデルの妥当性と考える。これを表す指標である余裕エネルギーを定義する。この余裕エネルギーがもっとも大きいモデルの多重度 n をその点におけるフローの多重度として採用する。これにより、運動透明性の検出ができるばかりでなく、従来から困難な問題であった運動境界におけるフロー場の不連続の扱いを計算上の困難なく自然に行うことができる。

最後に、アルゴリズムの1実現方法として、画像の Gabor 関数展開を前提とした場合の計算手順を計算機上に実現する。ランダムテクスチャ画像を用いたシミュレーションによって、運動透明性と運動境界によって生じる2重運動が、それらの存在の検出とともに計算可能であることを示す。

2.1.2 オプティカルフローの基本拘束方程式と運動パタンの重量

2.1.2.1 オプティカルフローの基本拘束方程式

画像面上における見かけの速度場をオプティカルフローという [44][51]。オプティカルフロー $(u(x, y, t), v(x, y, t))$ の基本拘束方程式は、時空間画像 $f(x, y, t)$ に対する1階偏微分の拘束として次の様に表現される [56][81][51]。

$$u(x, y, t) \frac{\partial}{\partial x} f(x, y, t) + v(x, y, t) \frac{\partial}{\partial y} f(x, y, t) + \frac{\partial}{\partial t} f(x, y, t) = 0 \quad (2.1)$$

フローベクトルが定数ベクトルであるとみなすことができる領域において拘束方程式 (2.1) をフーリエ変換すると、時空間周波数領域におけるオプティカルフローの基本拘束方程式が得られる。

$$2\pi i(u\omega_x + v\omega_y + \omega_t)F(\omega_x, \omega_y, \omega_t) = 0 \quad (2.2)$$

$(\omega_x, \omega_y, \omega_t)$ は、時空間座標 (x, y, t) に対応する周波数軸の座標を表し、 $F(\omega_x, \omega_y, \omega_t)$ は、時空間画像 $f(x, y, t)$ のフーリエ変換を表す³。この拘束は、時空間周波数領域 $(\omega_x, \omega_y, \omega_t)$ において、原点を通り、法線ベクトルが $(u, v, 1)$ である平面上にフーリエ変換 $F(\omega_x, \omega_y, \omega_t)$ の成分が乗るという共平面拘束を表現している。

²視覚情報処理を例とする逆問題が良設定であるとは、データのノイズに対して解が滑らかに変化し、かつ、一意に確定することをいう [15]。良設定性は、ビジョンのアルゴリズムがもつべき重要な特性である。

³本論文の議論では、定数 $2\pi i$ には、実質的な効果はない。

本節では、次節で多重オプティカルフローの基本拘束方程式を導く準備をする。画像生成過程において多重運動がどのようにして生じうるかを考察する。まず、時刻 t の画像面上の点 (x, y) における輝度値 $f(x, y, t)$ が2つの速度場 $v_1(x, y, t) = (u_1(x, y, t), v_1(x, y, t))^T$ と $v_2(x, y, t) = (u_2(x, y, t), v_2(x, y, t))^T$ (T は転置を表す。) をそれぞれ持つパタン $f_1(x, y, t)$ と $f_2(x, y, t)$ の和あるいは積 (\oplus でその重畳の演算を表す) として $f(x, y, t) = f_1(x, y, t) \oplus f_2(x, y, t)$ と表される場合を考察する。 $f_1(x, y, t)$ と $f_2(x, y, t)$ は、輝度とは限らず透過率のパタン、あるいは、投影光の強度パタンであってもよい。これらは、単一オプティカルフローの基本拘束方程式(2.1)によって、それぞれ局所的に、

$$\begin{aligned} u_1(x, y, t) \frac{\partial}{\partial x} f_1(x, y, t) + v_1(x, y, t) \frac{\partial}{\partial y} f_1(x, y, t) + \frac{\partial}{\partial t} f_1(x, y, t) &= 0 \\ u_2(x, y, t) \frac{\partial}{\partial x} f_2(x, y, t) + v_2(x, y, t) \frac{\partial}{\partial y} f_2(x, y, t) + \frac{\partial}{\partial t} f_2(x, y, t) &= 0 \end{aligned} \quad (2.3)$$

と拘束されているものとする。

2.1.2.2 和の形の重畳による運動透明視

汚れていない透明な水面を通して、水面下の物体を見ている状況を考えよう。水面の鏡面反射によって、地上の情景と水面下の情景が和の形に重畳されて観察される。水面に波が生じると、地上の情景と水面下の情景はそれぞれ異なる運動として観察される (図 2.1 参照)。この運動の違いによって、鏡面反射による画像 $f_1(x, y, t)$ と透過してきた水面下の物体の像 $f_2(x, y, t)$ の2つの運動パターンを分離して知覚できる。

この場合、観測される画像は、

$$f(x, y, t) = f_1(x, y, t) + f_2(x, y, t) \quad (2.4)$$

と和の形の重畳になる。

2.1.2.3 積の形の重畳による運動透明視

暗い部屋の中から、汚れて透過率にゆらぎのある窓ガラスを通して景色を見ている場合 [112]、画像面上には、景色のパタンと窓ガラスの汚れの透過率パタンが積の形に重畳されて観察される⁴。観察者が運動すると、窓の外の情景像 $f_1(x, y, t)$ と窓ガラスの汚れによる透過率パタン $f_2(x, y, t)$ は異なる運動をするので、これらの運動パターンは分離して知覚される。

⁴部屋の中は暗いと仮定しているのでガラス表面における反射は無視できる。



Figure 2.1: 和の形の重畳による運動透明視の例

もう一つの例は、影である。例えば、木かげにおけるこもれびが地面に投射されている状況を考える。この場合には、画像は、こもれびの光源としての放射光強度パターン $f_1(x, y, t)$ と地面の反射率パターン $f_2(x, y, t)$ の積⁵である。木が、風によってゆらいだり、地面で動物が移動すると、これらのパターンは画像面上でみかけの運動が異なるので分離して知覚される (図 2.2 参照)。

これらの場合には、観測される画像は、

$$f(x, y, t) = f_1(x, y, t) \cdot f_2(x, y, t) \quad (2.5)$$

と積の形に重畳される。この式の両辺の対数をとることにより、和の型の重畳に変換できる。

$$\log f(x, y, t) = \log f_1(x, y, t) + \log f_2(x, y, t) \quad (2.6)$$

⁵これに環境光による反射パターンがさらに和の形に重畳されるであろうが、こもれびに比べて光強度は弱いので無視できるとした。

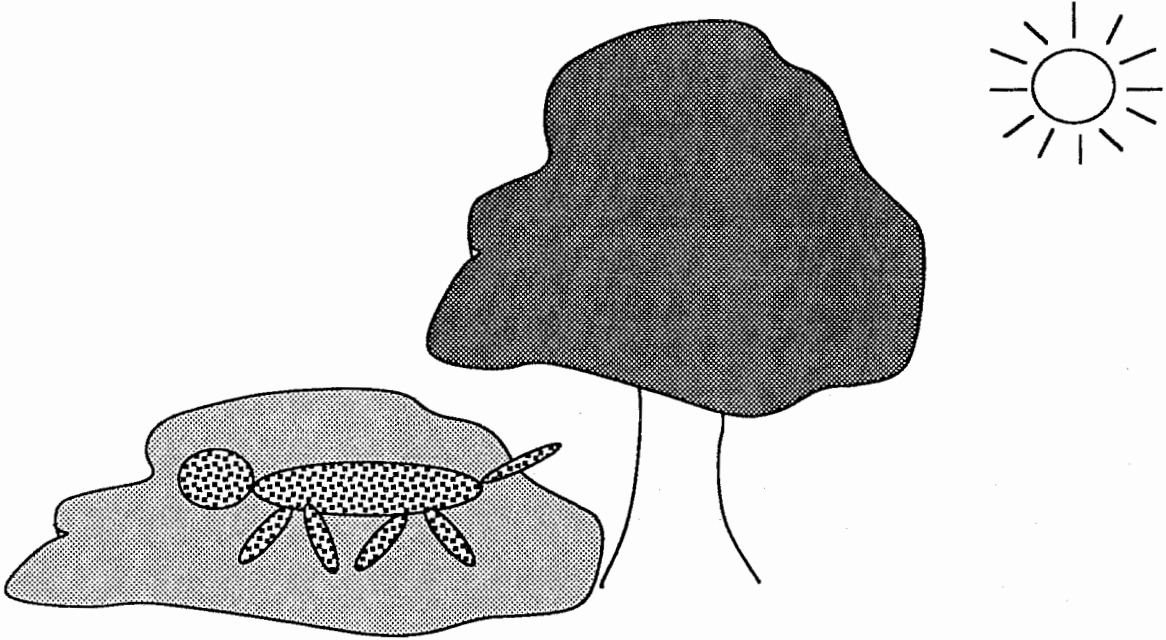


Figure 2.2: 積の形の重畳による運動透明視の例

一般に、適当な単調関数によって和の重畳に変換される重畳は、一般化重畳 (generalized superposition) と呼ばれる [66]。

2.1.2.4 フロー場の不連続の重畳としての解釈

オクルージョン、あるいは、フロー場の不連続は重畳とみなすことができる。図 2.3 は、フロー場の不連続を表した図である。図 2.3(a) は、セグメンテーションのパラダイムに基づく従来モデルである。フロー場の不連続部分を抽出することを目的とし、線過程 [43][61] を導入したマルコフ場モデル、拘束方程式のクラスタリング [72] などが試みられてきた。ところが、多重フローベクトル場を許すと、図 2.3(b) の様に、それぞれ v_1 、 v_2 なる連続なフロー場を持つ 2 個の時空間画像パターン $f_1(x, y, t)$ および $f_2(x, y, t)$ の和の型の重畳 $f(x, y, t) = f_1(x, y, t) + f_2(x, y, t)$ として不連続をモデル化できる。(次節で示すように、時空間画像の重畳はフローパラメータ空間におけるテンソル積に対応する。) このとき、パターンが存在しない領域では、 $f_1(x, y, t)$ または、 $f_2(x, y, t)$ が 0 であるとみなす。する

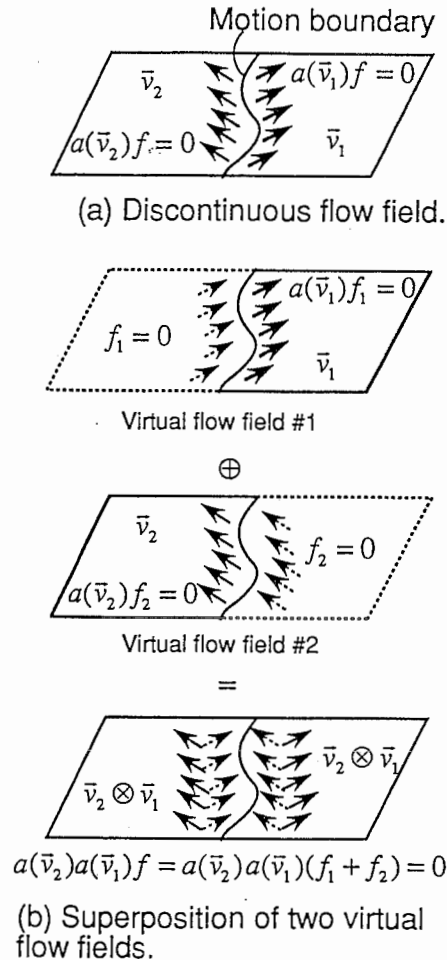


Figure 2.3: フロー場の不連続は重畳として解釈できる

(a) 不連続を含むフロー場、(b) 2 個の仮想連続フロー場の重畳を用いた不連続の表現。

と、拘束方程式と画像生成方程式がともに満たされる。これを積の型の重畳として考えることも可能である。この時は、パターンが存在しない領域では、 $f_1(x, y, t)$ または $f_2(x, y, t)$ が 1 をとると想定する。 $f_1(x, y, t)$ と $f_2(x, y, t)$ は、直接観測される量ではない。したがって、拘束方程式と画像生成方程式に矛盾しない限り、このような想定をすることは自由である。この様に、拘束方程式のレベルでは、フロー場の不連続に「不感」なモデルを想定できる。このことは、フロー推定アルゴリズムにおいて、フロー場の不連続を必ずしも陽に表現し検出する必要はないことを示している。多重フローのパラメータが得られれば、単一フローの拘束方程式を使って各画像点においてフローの選択を行うことができる。これによって運動情報に基づいた画像のセグメンテーションは極めて容易になる。必要ならば、不連続すなわち運動境界も検出できる。この事実は、セグメンテーションと運動パラメータ推定の関係に、根本的に新しい可能性を加える。つまり、セグメンテーションは、パラメータ推定の後で一撃 (one shot) 的に実行可能である (図 2.4 参照)。

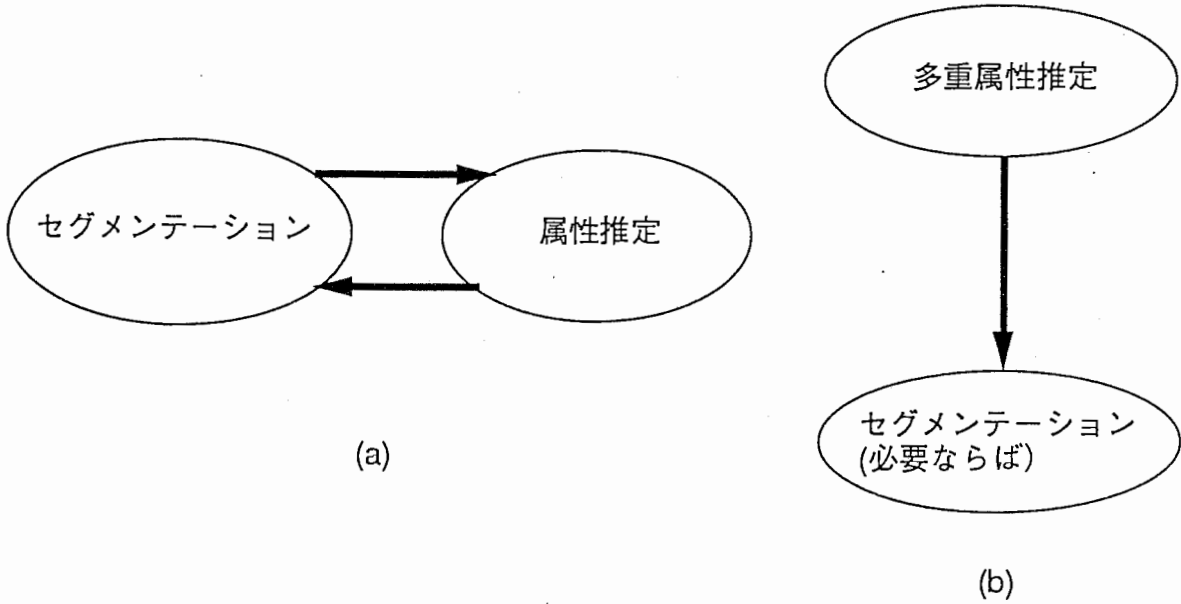


Figure 2.4: 運動推定とセグメンテーションの新しい関係

(a) 従来の考え方：セグメンテーションとパラメータ推定は「鶏と卵」の関係と似ている。従来は、二つのプロセスの協調的相互作用で解く方法が一般的であった。(b) 新しい考え方：セグメンテーションを経ずに複数組のパラメータ推定が可能である。パラメータが求めればセグメンテーションは容易である。

2.1.3 多重オブティカルフローの基本拘束方程式

以下では、基本拘束方程式 (2.1) および (2.2) を、フローパラメータ (u, v) によってパラメータづけられた線形演算子

$$a(u, v) \stackrel{\text{def}}{=} u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} + \frac{\partial}{\partial t} \tag{2.7}$$

とそのフーリエ変換

$$\tilde{a}(u, v) \stackrel{\text{def}}{=} 2\pi i(u\omega_x + v\omega_y + \omega_t) \tag{2.8}$$

によって、

$$a(u, v)f(x, y, t) = 0, \quad \tilde{a}(u, v)F(\omega_x, \omega_y, \omega_t) = 0 \tag{2.9}$$

と記述する。

このとき、オプティカルフローの基本拘束方程式は n 重のフローベクトル場を持つ時空間画像に拡張される。

定理 2.1 次の、3つの仮定をおく。(1) 時空間画像 $f(x, y, t)$ が、 n 個の時空間画像 $f_i(x, y, t)$ の和として、

$$f(x, y, t) = \sum_{i=1}^n f_i(x, y, t) \quad (2.10)$$

と表される。(2) 各 $f_i(x, y, t)$ は、単一オプティカルフローの基本拘束方程式、

$$a(\mathbf{v}_i) f_i(x, y, t) = 0 \quad (i = 1, \dots, n) \quad (2.11)$$

を満たす。(3) \mathbf{v}_i は、変数 x, y, t に関して定数である。

このとき、 $f(x, y, t)$ は、次式で拘束される。

$$a(\mathbf{v}_n) \dots a(\mathbf{v}_2) a(\mathbf{v}_1) f(x, y, t) = 0 \quad (2.12)$$

時空間周波数領域では次式である。

$$\tilde{a}(\mathbf{v}_n) \dots \tilde{a}(\mathbf{v}_2) \tilde{a}(\mathbf{v}_1) F(\omega_x, \omega_y, \omega_t) = 0 \quad (2.13)$$

□

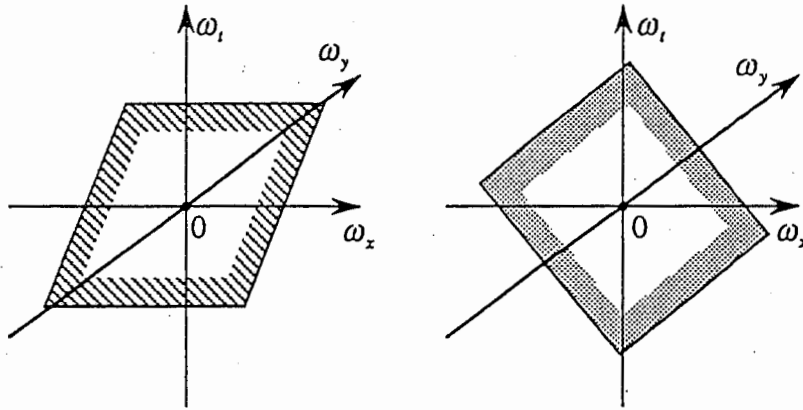
[証明] 付録 A.1.1 を参照。

この定理は、フローパラメータ \mathbf{v}_i が定数でないと厳密には成立しない。しかし、フローが時空間領域で十分なめらかで、局所的に定速度並進運動であれば、画像面の各点で近似的に成立する。第 2.1.2.4 節で述べたように、多重オプティカルフローを許容すると、物体境界などのフロー場が不連続になる部分でもこの滑らかさの仮定は成立する。フロー場の不連続が多い複雑な物体の重なり合いを撮像した動画像に対しても、原理的に、多重オプティカルフローの基本拘束方程式が成立する。

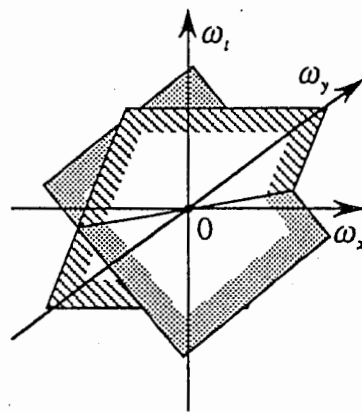
この拘束方程式は、 n 重オプティカルフローを仮定した場合に成立する必要条件である。 n 重フローの拘束方程式が成り立つ時は、 $m (> n)$ 重フローの拘束方程式も成立する。このとき、 m 組のフローパラメータのうち n 組は n -重の拘束方程式を適用した場合と同じ値をとる。残りの $m - n$ 組は任意の値をとることができる⁶。 n -重オプティカルフローの基本拘束方程式 (2.12) は、時空間画像 $f(x, y, t)$ の n 階偏微分、

$$\frac{\partial^n f(x, y, t)}{\partial x^p \partial y^q \partial t^r} \quad (p + q + r = n) \quad (2.14)$$

⁶実際には、第 6 章のシミュレーションで示す様に、画像の解像度の有限性とノイズの影響によってより複雑な振舞いをする。ここでの議論は、基本拘束方程式の前提条件が厳密に成立する場合のものである。



(a) Constraints of Flow #1 and Flow #2.



(b) Constraint of Double Flow.

Figure 2.5: 2重オブティカルフローの時空間周波数領域における拘束条件

(a) それぞれの単独運動に対する拘束条件 (原点を通る平面)、(b) 合成された2重運動に対する拘束条件 (原点を通る2平面)。

に関する拘束である。例えば、 $n = 2$ の場合の拘束方程式 $a(u_2, v_2)a(u_1, v_1)f(x, y, t) = 0$ は、時空間画像 $f(x, y, t)$ の2階偏微分の5個の成分に対する拘束である。

$$\begin{aligned}
 &u_1 u_2 \frac{\partial^2 f}{\partial x^2} + v_1 v_2 \frac{\partial^2 f}{\partial y^2} + (u_1 v_2 + v_1 u_2) \frac{\partial^2 f}{\partial x \partial y} \\
 &+ (u_1 + u_2) \frac{\partial^2 f}{\partial x \partial t} + (v_1 + v_2) \frac{\partial^2 f}{\partial y \partial t} + \frac{\partial^2 f}{\partial t^2} = 0
 \end{aligned} \tag{2.15}$$

時間方向に関しては1階および2階の偏微分の計算が必要である。したがって、最小3フレームの情報が必要である。

時空間周波数領域における2重オブティカルフローの拘束 $\tilde{a}(u_2, v_2)\tilde{a}(u_1, v_1)F(\omega_x, \omega_y, \omega_t) = 0$ は、原点を通る2枚の平面に時空間画像信号の周波数成分が乗るという拘束になる (図

2.5)。

$$(2\pi i)^2(u_1\omega_x + v_1\omega_y + \omega_t)(u_2\omega_x + v_2\omega_y + \omega_t)F(\omega_x, \omega_y, \omega_t) = 0 \quad (2.16)$$

2.1.4 多重オプティカルフローの計算アルゴリズム

オプティカルフローの計算アルゴリズムには、極めて多種多様なものがある。次にそれを分類する。

1. 空間領域、または、周波数領域における基本拘束方程式とフローの滑らかさ拘束を正則化あるいは弛緩法による画像全体の最適化を用いて解くもの [51] [79] [62]。
2. 局所的な拘束条件の当てはまり具合を表すエネルギー関数の最適化によって画像面の各点において独立にフローを求めるもの (局所最適化法) [53] [49]。
3. 小領域の相関計算を行なってピーク検出でフローベクトルを計算するもの [92] [28] [73] [45]。
4. 時空間画像のフーリエ位相を、ガボールフィルタなどの quadrature mirror pair をなすフィルタ⁷によって検出するもの [41]。

以下において、時空間局所最適化に基づく多重オプティカルフローの一般的な計算アルゴリズムを導く。

2.1.4.1 フローパラメータの同次座標表現とエネルギー積分

いままでの議論では、基本拘束方程式において、画像面の xy 平面上で定義されたフローベクトル (u, v) によってオプティカルフローを表現してきた。2 フレーム以上の時空間画像を対象とし、時空間フィルタを用いる場合には、空間方向 x, y と時間方向 t に関して平等な運動表現が望ましい。本論文で用いた量子力学における状態空間を表すヒルベルト空間とその積との形式的な一致をとるためにも (後述)、パラメータ空間を同次座標で表現しておくことが必要である。

そこで、以下では、 xyt 空間における方向ベクトル $(u, v, 1)$ を同次座標表現した $\mathbf{V} = (V_x, V_y, V_t)^T = (V_1, V_2, V_3)^T$ (T は転置を表す) をフローパラメータの表現として用いる。フローベクトル (u, v) との関係は $u = V_1/V_3, v = V_2/V_3$ である⁸。これは、一見パラメータのおきかえにすぎない。しかし、拘束方程式の 2 乗誤差を評価した時に、2 乗距離の定義に異

⁷位相が $\pi/2$ だけ異なるフィルタの組で、ヒルベルト変換によってたがいに交換される。

⁸ \mathbf{V} には符号の不定性が残るが、どちらの符号をとっても (u, v) に同じ値を与える。

なる効果がある。以下においては、 V の空間を \mathcal{H} で表す。同次フロー空間 \mathcal{H} のノルムを次式の一般の2次形式で定義する。

$$\|V\|_W \stackrel{\text{def}}{=} \sqrt{V^T W^T W V} \quad (2.17)$$

W は上三角行列で、 $W^T W$ は、正定値対称行列である。この同次座標表現を用いるとフローの振幅演算子は次式になる。

$$a(V) = V^T \nabla, \quad \tilde{a}(V) = V^T \omega \quad (2.18)$$

ここで、 $\nabla = (\partial_1, \partial_2, \partial_3)^T \stackrel{\text{def}}{=} (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial t})^T$, $\omega = (\omega_1, \omega_2, \omega_3)^T \stackrel{\text{def}}{=} (\omega_x, \omega_y, \omega_t)^T$ と置いた。

すると、 n -重フローの振幅演算子は、次の形式に書かれる。

$$\begin{aligned} a^{(n)}(V^{(n)}) &\stackrel{\text{def}}{=} a(V_n) \cdots a(V_2) a(V_1) \\ \tilde{a}^{(n)}(V^{(n)}) &\stackrel{\text{def}}{=} \tilde{a}(V_n) \cdots \tilde{a}(V_2) \tilde{a}(V_1) \end{aligned} \quad (2.19)$$

$V^{(n)}$ は、 $V_i (i = 1, 2, \dots, n)$ のテンソル積(クロネッカー積) $V^{(n)} = \bigotimes_{i=1}^n V_i = V_n \otimes \cdots \otimes V_2 \otimes V_1$ の形になる。このフローベクトルのテンソル積に伴うノルムは、次に定義される。

$$\|V^{(n)}\|_{W^{(n)}} = \prod_{i=1}^n \|V_i\|_{W_i} \quad (2.20)$$

$W^{(n)}$ は、上記関係を満たすように定義される上三角行列のテンソル積である。具体的には、

$$W^{(n)} = \bigotimes_{i=1}^n W_i = W_n \otimes \cdots \otimes W_2 \otimes W_1 \quad (2.21)$$

これらの定義は、線形ノルム空間あるいはヒルベルト空間同士のテンソル積空間の定義をフローベクトル空間に適用したものである⁹。

さらに、2乗残差を評価するために、 n 重フローのエネルギー密度演算子 $H^{(n)}(V^{(n)})$ と $\tilde{H}^{(n)}(V^{(n)})$ を次式で定義する。

$$\begin{aligned} H^{(n)}(V^{(n)})f(x) &\stackrel{\text{def}}{=} \overline{\{a^{(n)}(V^{(n)})f(x)\}} \{a^{(n)}(V^{(n)})f(x)\} \\ \tilde{H}^{(n)}(V^{(n)})F(\omega) &\stackrel{\text{def}}{=} \overline{\{\tilde{a}^{(n)}(V^{(n)})F(\omega)\}} \{\tilde{a}^{(n)}(V^{(n)})F(\omega)\} \end{aligned} \quad (2.22)$$

\bar{a} は、 a の複素共役を表す。振幅演算子の単なる2乗ではなく、複素共役との積として定義したのは、Gabor 関数の \cos 項と \sin 項の様に、quadrature pair をなす関数による画像表現を前提とする場合に、画像を複素画像として扱う必要があるためである。

⁹本来、テンソル $W^{(n)}$ は、データの持つ誤差に対して、推定値が不偏になる様に定めるべきである。しかし、本研究では、基本拘束方程式の検証が主眼である。したがって、この点に関しては論じない。近年、コンピュータビジョンにおける不偏最適推定の一般的方法が提案されている [13]。これを直接応用できる。

以上の定義を用いると、 n -重フローの推定問題を次のエネルギー関数 (以下、エネルギー積分と呼ぶ) $E^{(n)}$ の最小化問題として定義できる。

$$E^{(n)}(\mathbf{V}^{(n)}) = \frac{\int_{\mathcal{N}(x_c)} H^{(n)}(\mathbf{V}^{(n)}) f(x) dx}{\|\mathbf{V}^{(n)}\|_{W^{(n)}}^2 \int_{\mathcal{N}(x_c)} |f(x)|^2 dx} \quad (2.23)$$

$\mathcal{N}(x_c)$ は、時空間座標 $x_c = (x_c, y_c, t_c)^T$ を含み、フローに関する2次元並進運動の仮定が成立する時空間局所領域を表す。エネルギー積分 (2.23) の意味は次の通りである。このエネルギー積分の最小化は、分母の $\|\mathbf{V}^{(n)}\|_{W^{(n)}}^2$ が一定である条件 (計量) のもとで、分子の $\int_{\mathcal{N}(x_c)} H^{(n)}(\mathbf{V}^{(n)}) f(x) dx$ を最小化することと等価である。分母の $\int_{\mathcal{N}(x_c)} |f(x)|^2 dx$ なる項は、エネルギー $E^{(n)}(\mathbf{V}^{(n)})$ に、正規化された量としての意味を持たせるために置いた。この項は、パラメータの推定値には依存しない。画像 $f(x)$ が、あらかじめ $\int_{\mathcal{N}(x_c)} |f(x)|^2 dx = 1$ と正規化されていれば、この項を省くことができる。

2.1.4.2 固有値問題を用いた最適化

エネルギー関数 (2.23) の最小化問題を解いて、フローパラメータの推定を行なう。次の2通りの手順が考えられる。

1. フローベクトル $\mathbf{V}_i (i = 1, \dots, n)$ の各成分を $3n$ 個の未知変数とみなして直接最適化する。
2. テンソル積 $\mathbf{V}^{(n)} = \otimes_{i=1}^n \mathbf{V}_i$ の各成分を独立変数とみなすことにより、エネルギー積分を未知パラメータに関する2次形式の比 (レイリー商) に変形して最適化する。

第1の方法では、最小2乗の意味での最適推定が可能である。しかし、これは、非線形最適化問題である。このとき、初期値の選択と収束性の保証が問題となる。第2の方法では、実際に独立な変数よりも多い変数に関して最適化を行なう。したがって、原問題の最小2乗の意味での最適推定値は計算されない。しかし、線形固有値問題に帰着されるため、初期値は不要である。解は、閉形式の数値計算で求められる。第3の方法として、まず、第2の方法によって線形推定値をもとめ、それを初期値として、第1の方法で最適化を行なうことも考えられる。後に述べるシミュレーションでは、第2の方法によってフローベクトルを求めた。

非線形固有値方程式系による解法 各フローベクトル \mathbf{V}_k について個別に変分法を適用して得られるアルゴリズムを示す。次の n 個の固有値問題が導かれる (導出は、付録 A.1.2を参照)。

$$K^{(n,k)} \mathbf{V}_k = \lambda^{(n,k)} W_k^T W_k \mathbf{V}_k (k = 1, \dots, n) \quad (2.24)$$

行列 $K^{(n,k)}$ は、3 次の正定値対称行列で、 V_k 以外のすべてのフローパラメータに依存している。この連立非線形固有値問題に対して、次の反復解法が考えられる。

[アルゴリズム]

1. 各 k に対応する V_k に互いに異なる初期値を与える。
2. $k = 1, 2, \dots, n$ について、
 - (a) 現在の V_k の値を用いて、 $K^{(n,k)}$ ($k = 1, \dots, n$) を計算する。
 - (b) 固有値問題 (2.24) を解く。
 - (c) 収束を固有値やフローベクトルの変化の大きさなどで判定し、さらに繰り返しが必要ならば、2.(a) に戻る。

テンソル積パラメータについての線形固有値方程式による解法 エネルギー積分 $E^{(n)}(V^{(n)})$ において、フローベクトルのテンソル積 $V^{(n)} = \otimes_{i=1}^n V_i$ の 3^n 個の成分を独立変数とみなす。これは一種の近似である。エネルギー積分は、この 3^n 個のテンソル積フローパラメータに関して2次形式の商の形である。このとき、 n 重フローの最適推定問題は、次式の線形固有値問題に帰着される。

$$M_{(i,j)}^{(n)}(V^{(n)})_i = \lambda^{(n)} \{ (W^{(n)})^T W^{(n)} \}_{(i,j)} (V^{(n)})_j \quad (2.25)$$

$i = (i_1, i_2, \dots, i_n)$ と $j = (j_1, j_2, \dots, j_n)$ は、 $(V^{(n)})_i = (V_1)_{i_1} (V_2)_{i_2} \dots (V_n)_{i_n}$ 、 $(V^{(n)})_j = (V_1)_{j_1} (V_2)_{j_2} \dots (V_n)_{j_n}$ というテンソル積 $V^{(n)}$ のすべての成分を指定できるインデックスからなるベクトルを表す。 $M_{(i,j)}$ は、 $2n$ 次のエネルギーモーメントからなるテンソルである。

$$\begin{aligned} M_{(i,j)} &= \frac{\int_{N(x_c)} \overline{\partial_{i_1} \dots \partial_{i_n} f(x)} \cdot \partial_{j_1} \dots \partial_{j_n} f(x) dx}{\int f(x) \cdot f(x) dx} \\ &= \frac{\int_{N(x_c)} (4\pi^2)^n \omega_{i_1} \dots \omega_{i_n} \overline{F(\omega)} \cdot \omega_{j_1} \dots \omega_{j_n} F(\omega) d\omega}{\int \overline{F(\omega)} \cdot F(\omega) d\omega} \end{aligned} \quad (2.26)$$

フローのテンソル積空間の対称性に注意する必要がある。各フローに対応するノルム行列 W_i が等しい場合には、任意の異なる2つのフローパラメータ V_k と V_l を入れ換えてもエネルギー積分 $E^{(n)}(V^{(n)})$ の形は変化しない。この対称性によって、固有値問題 (2.25) は、最小固有値が $n!$ 重に縮退し、固有空間が2次元以上になる。つまり、最小固有値に対応する固有ベクトルは不定である。固有値が縮退した場合の数値計算は極めて困難である。テンソル積の対称化操作によって、この対称性をテンソル積空間から消去できる (付録 A.1.3、付録 A.1.4 参照)。

この対称化の操作によって得られる n 重フロー推定の線形固有値問題は、 n 重のフローが存在するという前提条件において最小固有値に縮退がない。したがって、最適化アルゴリズムとして、収束性や安定性の優れた固有値解析のための数値計算手法が利用できる。

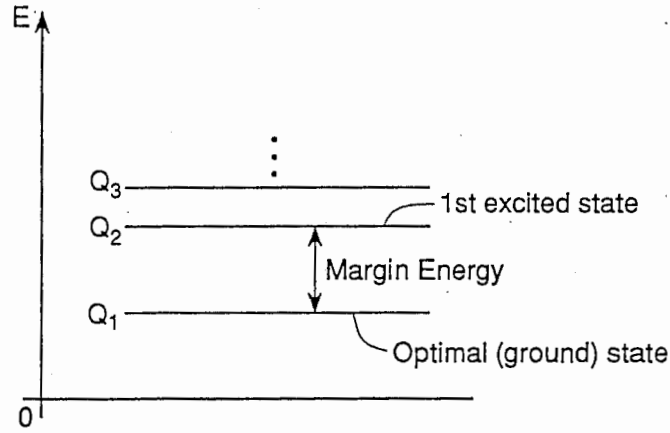
2.1.4.3 テンソル積パラメータの分解

線形固有値問題に帰着する多重フロー推定アルゴリズムでは、最後に推定結果のテンソル積パラメータを分解し、個々のフローパラメータを計算する必要がある。テンソル積パラメータの成分は、個々のフローパラメータの成分数の和よりも多い。したがって、過剰拘束の非線形方程式を解く必要がある。これは、有限回代数方程式を解き、方程式の冗長性を使って解を一義に決定するアルゴリズムで実現できる。ところが、2重フローのテンソル積に関しては、すべての成分に関する関係式を同時に最適に満たす解を求める閉形式の分解アルゴリズムが存在する。付録 A.1.4 にその導出を示す。

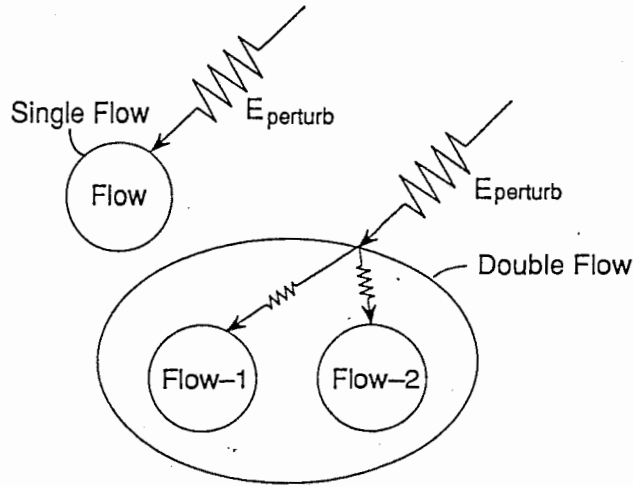
2.1.5 多重度の決定

2.1.3 で述べたように、 n 重フローの基本拘束方程式が満たされるとき、 n より大きい多重度の基本拘束方程式も成立する。従って、与えられたデータを満たす最も小さい多重度 n の選択が必要である。しかし、拘束方程式が厳密に成立しない実際の状況では、インプリメンテーションの上で困難である。そこで、固有値問題における最小固有値 (第1固有値) と最小から2番目の固有値 (第2固有値) の差に注目する。その理由は、多重フローのエネルギー積分から導かれる固有値問題が、量子力学における多粒子系のシュレーディンガー方程式から導かれる固有値問題に、数学的に類似した構造を持つことである (例えば、[158][161] を参照)。最適化問題の安定性、つまり良設定性を最小固有エネルギーに対応する基底エネルギー準位の外乱に対する安定性とみなす。このとき、アナロジーから基底エネルギーと第2固有エネルギー準位とのエネルギー差が大きい程安定であると考えられる。固有値問題に帰着されるエネルギー最小化問題では、最小固有値が縮退しないこと、つまり第1固有値と第2固有値が一致しないことが、解の一意性の必要十分条件である。この事実からも上記の指標が正当化できる。

ただし、我々の問題の場合には、異なる n に対応するモデル間でこの安定性を比較する必要がある。単一フローのエネルギー演算子 $H^{(1)}$ の単位をエネルギーの単位と考える。すると、 n -重フローのエネルギー演算子 $H^{(n)}$ は、このエネルギーの単位の n 乗の単位を持つ。これは、エネルギー積分 (2.25) を観察するとわかる。そこで、各固有値の n 乗根を



(a) Definition of Margin Energy.



(b) Use of Analogy to Energy Perturbation.

Figure 2.6: 固有エネルギーを用いて最適化問題の安定性=非退化条件を検査する。

(a) 余裕エネルギーの定義、(b) エネルギー擾動からの類推。

とった次の量 $Q_i^{(n)}$ を n -重フローモデルの固有エネルギーと定義する。

$$Q_i^{(n)} = (\lambda_i^{(n)})^{\frac{1}{n}} \tag{2.27}$$

差 $Q_2^{(n)} - Q_1^{(n)}$ を最適化問題の「余裕エネルギー」と呼ぶ。 n 個の粒子からなる系は、攪乱エネルギーが各粒子に分散される。したがって、 n -重のフローモデルは、単一フローモデルに比較して n 倍安定であると考え (図 2.6)。異なる多重度のモデル間で安定性を比較するための指標が必要である。以上の考察から、次式で定義される 1 フロー当たりの余裕エネルギーをこの指標として用いる。

$$J^{(n)} = \frac{1}{n}(Q_2^{(n)} - Q_1^{(n)}) = \frac{1}{n}\{(\lambda_2^{(n)})^{\frac{1}{n}} - (\lambda_1^{(n)})^{\frac{1}{n}}\} \tag{2.28}$$

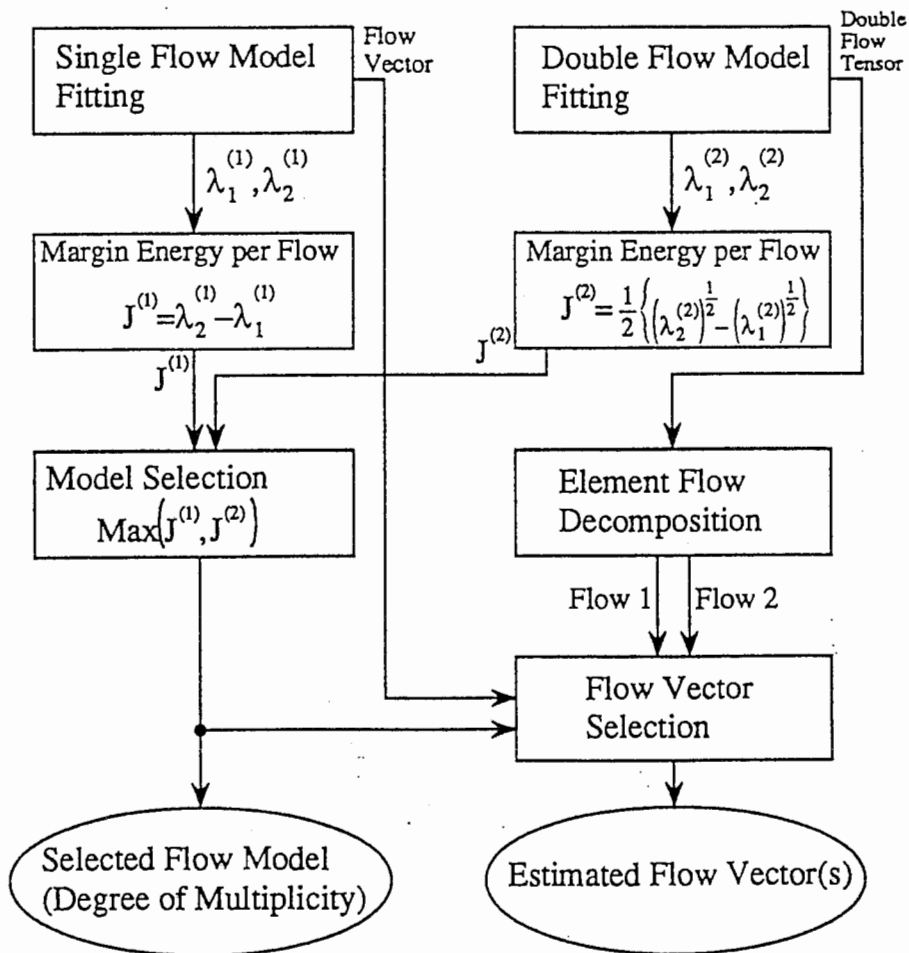


Figure 2.7: 1-重、2-重オブティカルフローのためのアルゴリズム

図 2.7 に、上記アルゴリズムを用いた 1-重、2-重フロー推定・多重度判定アルゴリズムの流れを示す。

ここでは、固有値の性質を用いた多重度の判定法を提案した。その他にも統計的仮説検定や情報量基準を用いたモデル選択法が考えられる。

2.1.6 ガボール関数による画像表現と計算機シミュレーション

2.1.6.1 ガボール要素関数の配置

3次元時空間ガボール要素関数は、複素正弦波とガウス関数の積で定義される [49] [70] [42]。

$$g(\sigma, \omega_c, x_c, x) = \frac{1}{\sqrt{2\pi}^{3/2} \sigma_x \sigma_y \sigma_t} \exp\left[-\left\{ \frac{(x - x_c)^2}{2\sigma_x^2} + \frac{(y - y_c)^2}{2\sigma_y^2} + \frac{(t - t_c)^2}{2\sigma_t^2} \right\}\right] \exp[2\pi i \{(x - x_c)\omega_{xc} + (y - y_c)\omega_{yc} + (t - t_c)\omega_{tc}\}] \quad (2.29)$$

$\sigma = (\sigma_x, \sigma_y, \sigma_t)^T$ は、実領域での関数の広がりを表し、 $\omega_c = (\omega_{xc}, \omega_{yc}, \omega_{tc})^T$ は、フィルタの中心周波数を表す。

このシミュレーションでは、ガボール要素関数のガウス窓の大きさを $\sigma_x = \sigma_y = \sigma_t = 4[\text{pixel}]$ と固定した。フィルタのカーネルを、 x, y, t 方向ともに 23 画素とした。これは、カーネルの端点に Gauss 関数の約 3σ の点に対応する大きさである。フィルタの中心周波数は、 $0.0[\text{pixel}^{-1}]$ から $0.4[\text{pixel}^{-1}]$ の範囲に各座標軸 $\omega_x, \omega_y, \omega_t$ にそって $0.07[\text{pixel}^{-1}]$ おきに設定した。これらの条件下で全フィルタ数は互いに原点对称になるものを同一視して 376 であった。3次元ガボール要素関数と時空間画像の畳み込みは、 x, y, t の各方向に対するカーネルの分離性 (separability) を利用して、 x, y, t 各方向への1次元の畳み込みのカスケードに分解して実現した。

$n = 1, 2$ の場合に、各ガボール要素関数と時空間画像の畳み込みによって得られる複素係数からのエネルギーモーメント (2.26) の計算式を付録 A.1.6 に示す。

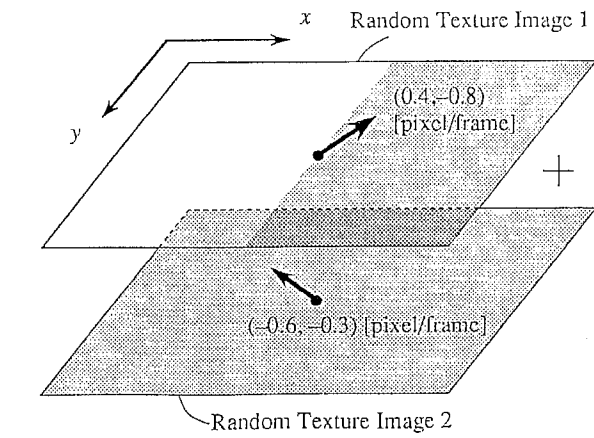
サンプル画像の構成法は以下の通りである。(1) 各画素に、それぞれ独立に、127 レベルの階調を乱数によって与えたランダムテクスチャ画像 (サイズ 512×480 画素) を 2 枚用意する。(2) それぞれを異なるパラメータで運動させた画像列を作る。(3) これらの画像列を各時空間座標点ごとに加算し、一つの画像列に合成する。サブピクセルの運動は、双一次関数による輝度補間を用いた。実験に使用した計算機の性能の限界から、 x, y 方向へは 32 画素おきに中心 x_c を設定し、これらの点だけでフロー計算を行なった。

2.1.6.2 シミュレーション：結果と考察

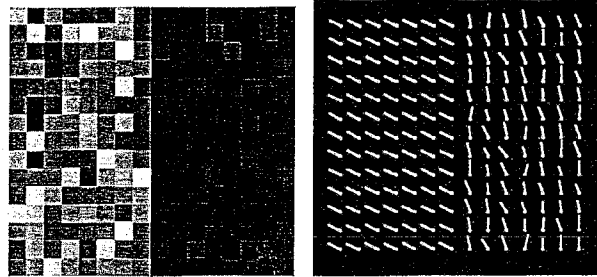
実験 1, 2, 3, 4 の全部で 4 通りのシミュレーションと結果を示し、考察を行なう。実験 1, 2, 3 では、運動透明視の検出を確認し、実験 4 では、運動境界の検出を確認する。

図 2.8, 図 2.9, 図 2.10, 図 2.11 は、サンプル画像生成法 (各図 (a))、単一フローモデルの余裕エネルギー (各図 (b))、単一フローモデルのフロー推定結果 (各図 (c))、2重フローモデルの1フロー当たりの余裕エネルギー (各図 (d))、2重フローモデルのフロー推定結果 (各図 (e))、フローの多重度 n の判定結果 (各図 (f))、各点で選択され出力されたフローの最終推定結果 (各図 (g)) を示す。

サンプル画像の生成法 図 2.8(a) は、実験 1 のサンプル時空間画像の生成法である。速度 $(-0.6, -0.3)[\text{pixel}/\text{frame}]$ で並進運動をするランダムテクスチャパターンと、速度 $(0.4, -0.8)[\text{pixel}/\text{frame}]$ で並進運動する画像面の右半分のみランダムテクスチャで左半分は透明 (輝度値 0) なパターンから輝度の加算によって合成した。合成された画像では、右半分に運動透明視が生じ、左半分は通常の単一運動が検出されなければならない。

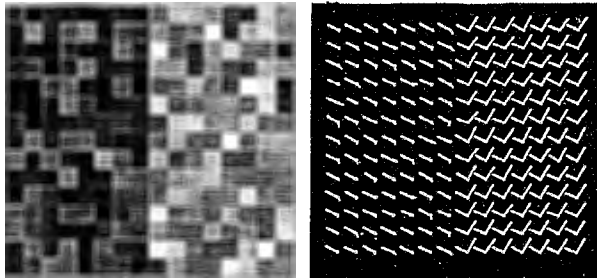


(a)



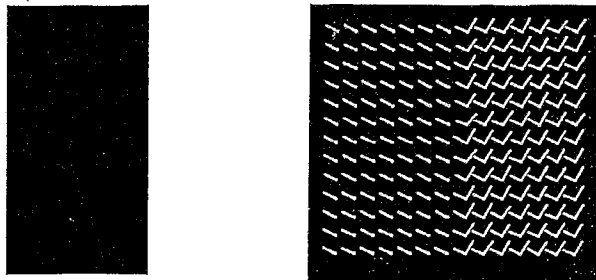
(b)

(c)



(d)

(e)



(f)

(g)

Figure 2.8: 実験1の結果

(a) サンプル画像の生成法、(b) 単一フローモデルの余裕エネルギー、(c) 単一フローモデルのフロー推定結果、(d) 2重フローモデルの1フロー当たりの余裕エネルギー、(e) 2重フローモデルのフロー推定結果、(f) フローモデル選択結果(黒: 単一フロー、白: 2重フロー)、(g) フロー推定の最終結果。

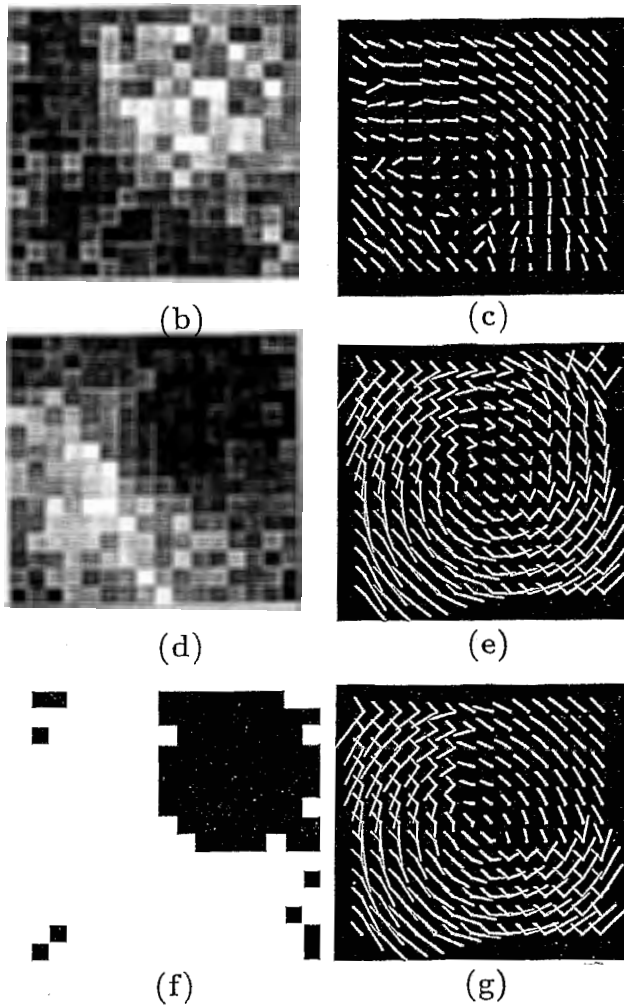
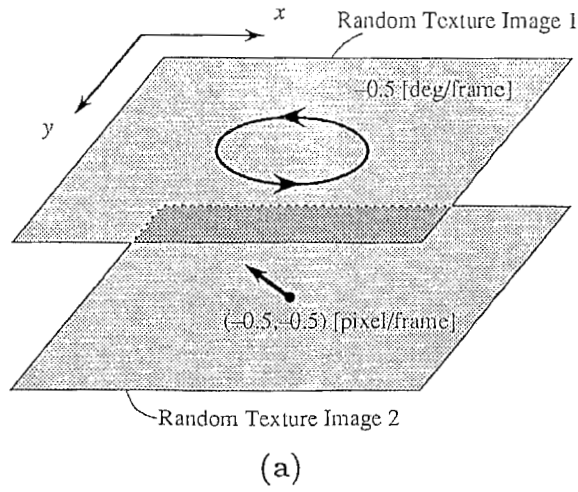
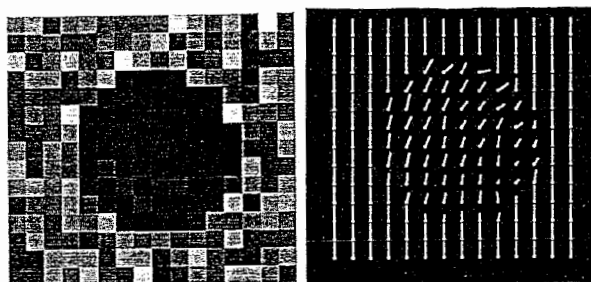
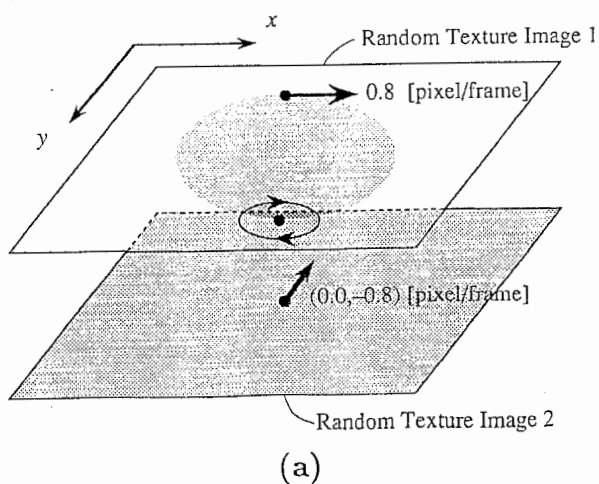


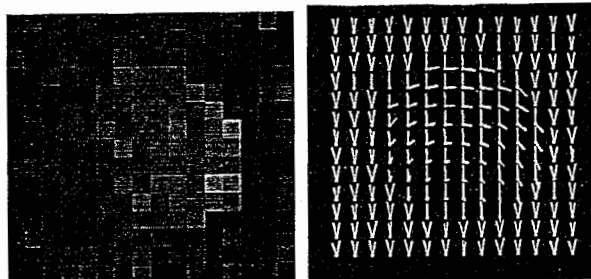
Figure 2.9: 実験2の結果

(a) サンプル画像の生成法、(b) 単一フローモデルの余裕エネルギー、(c) 単一フローモデルのフロー推定結果、(d) 2重フローモデルの1フロー当たりの余裕エネルギー、(e) 2重フローモデルのフロー推定結果、(f) フローモデル選択結果(黒: 単一フロー、白: 2重フロー)、(g) フロー推定の最終結果。



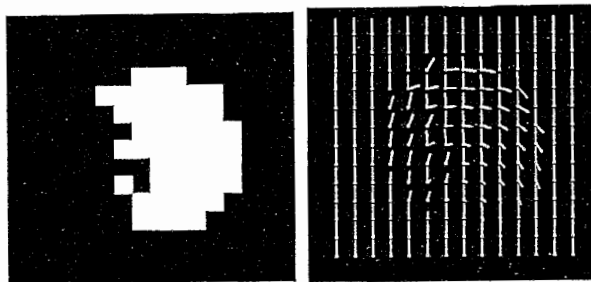
(b)

(c)



(d)

(e)

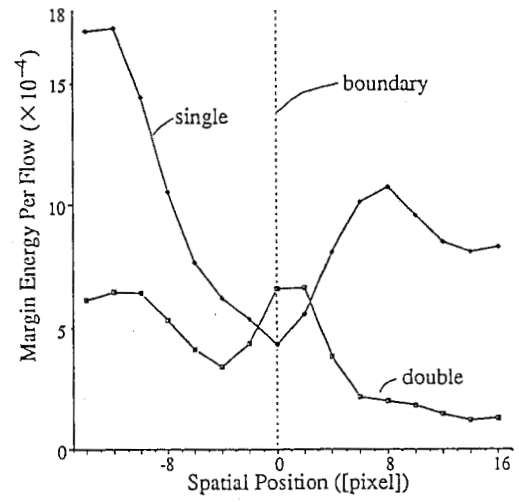
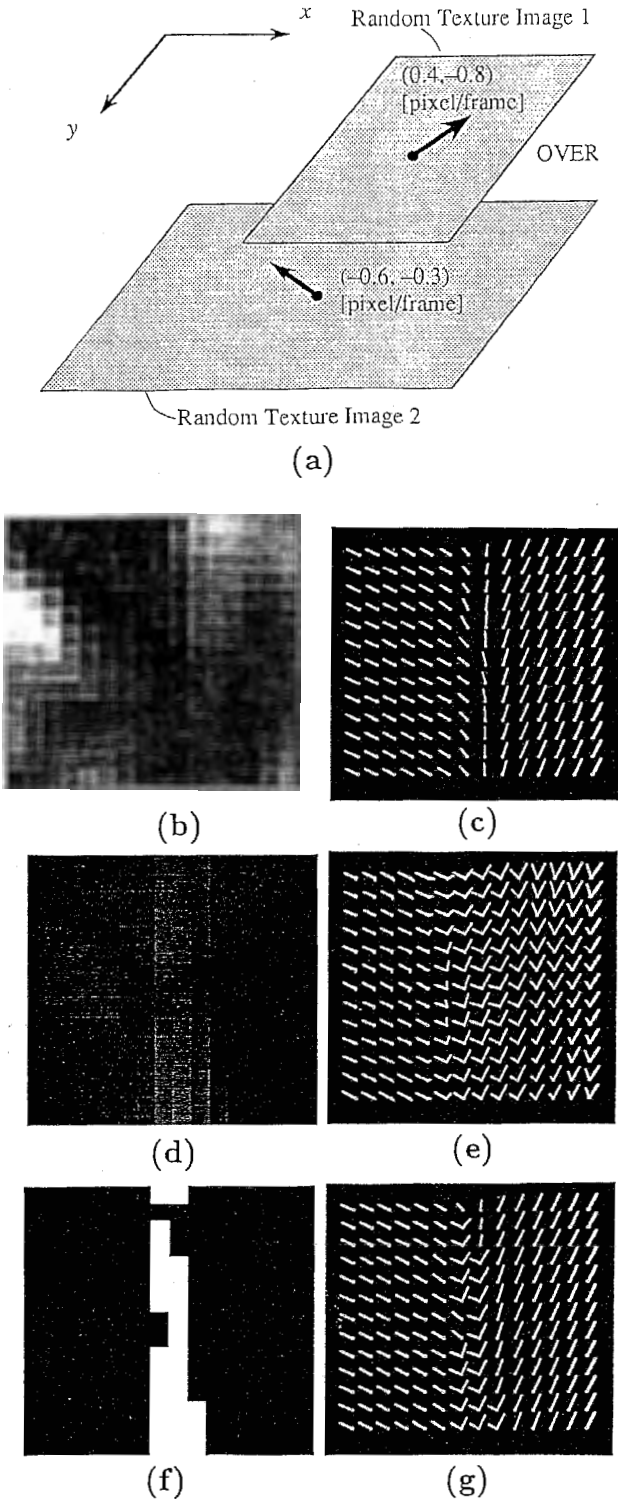


(f)

(g)

Figure 2.10: 実験3の結果

(a) サンプル画像の生成法、(b) 単一フローモデルの余裕エネルギー、(c) 単一フローモデルのフロー推定結果、(d) 2重フローモデルの1フロー当たりの余裕エネルギー、(e) 2重フローモデルのフロー推定結果、(f) フローモデル選択結果(黒: 単一フロー、白: 2重フロー)、(g) フロー推定の最終結果。



(h) Graphic Plot of Margin Energy per Flow Across Boundary.

Figure 2.11: 実験4の結果

(a) サンプル画像の生成法、(b) 単一フローモデルの余裕エネルギー、(c) 単一フローモデルのフロー推定結果、(d) 2重フローモデルの1フロー当たりの余裕エネルギー、(e) 2重フローモデルのフロー推定結果、(f) フローモデル選択結果(黒: 単一フロー、白: 2重フロー)、(g) フロー推定の最終結果、(h) フローの不連続点における余裕エネルギーの変化。

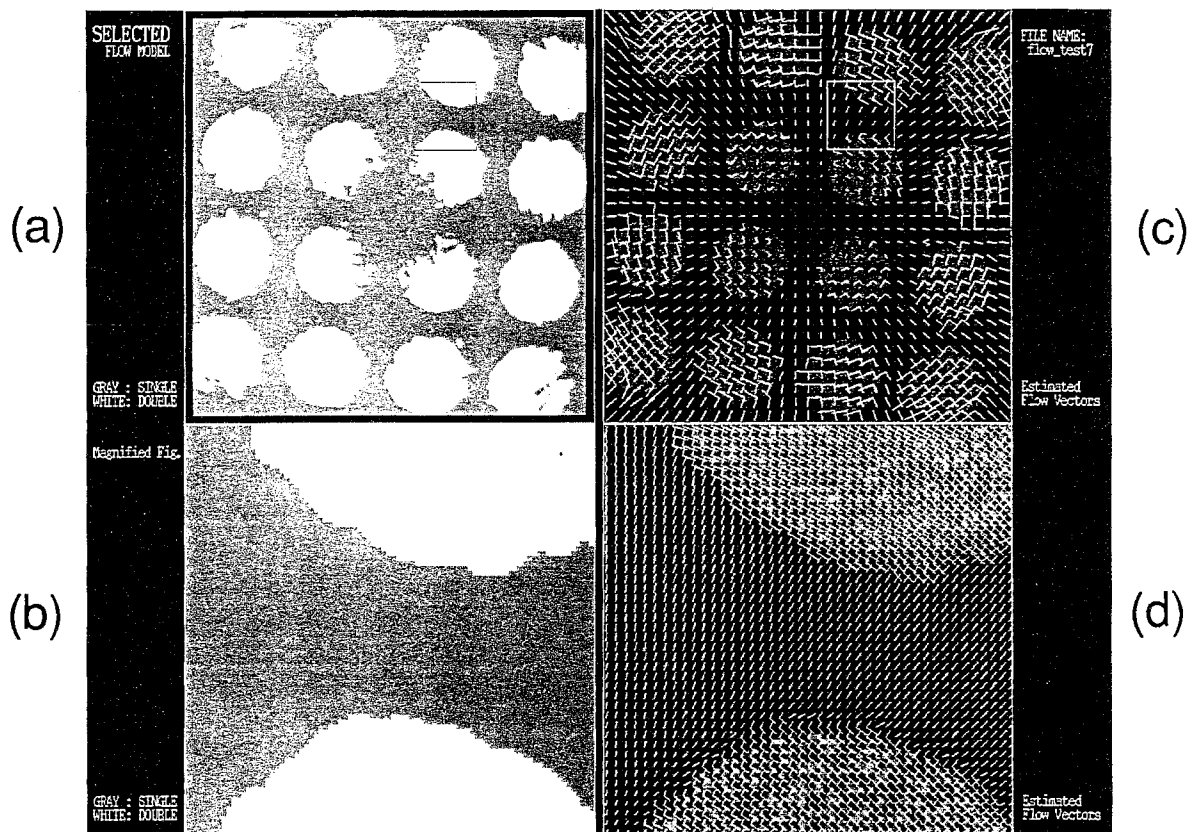


Figure 2.12: 超並列計算機による多重オブティカルフローのシミュレーション例
 (a) フロー多重度の判定結果、(b) フロー多重度の判定結果 (拡大図)、(c) 運動パターンとフロー推定結果、(d) 運動パターンとフロー推定結果 (拡大図)。

図 2.9(a) は、実験 2 のサンプル時空間画像の生成法である。速度 $(-0.5, -0.5)$ [pixel/frame] で並進運動をするランダムテクスチャパターンと、角速度 -0.5 [deg/frame] で回転するランダムテクスチャパターンを加算で合成した。

図 2.10(a) は、実験 3 のサンプル時空間画像の生成法である。速度 $(0.0, -0.8)$ [pixel/frame] で並進運動をするランダムテクスチャパターンと、画像の中心部分の円形領域にだけランダムテクスチャが存在し、その外部領域が透明なパターンを回転運動させたものを加算で合成した。円形領域は、下端を中心として、上端における速度が $(0.8, 0.0)$ [pixel/frame] となる回転運動をさせた。

図 2.11(a) は、実験 4 のサンプル時空間画像の生成法である。用意した 2 個の運動パターンは、実験 1 と同じである。右半分だけランダムテクスチャが存在するパターンが画像全体にテクスチャがあるパターンを覆い隠す様に重ね合わせた。画像面の中心部で、縦方向に運動境界が生じている。この場合、各運動パターンのテクスチャの性質はまったく同一なので運動以外の手がかりで境界を見い出すことは不可能である。

考察 実験 1 から、2.1.5 節で提案した多重度の判定が正しく行われていることがわかる。パターンのむらによって、余裕エネルギーそのものには空間方向にばらつきがある。それでも、多重度の判定が正確に行われている。これは、余裕エネルギーのロバスト性を示している。この実験結果では、単一フローの領域では、フローの推定値は、平均 $(-0.597, -0.293)$ 、標準偏差 $(0.0141, 0.0099)$ であった。2 重フローの領域では、推定されたフローベクトルは、平均が $(0.423, -0.624)$ と $(-0.572, -0.243)$ で、各座標値の標準偏差が $(0.0213, 0.0498)$ と $(0.0274, 0.0273)$ であった。精度に関しては、フィルタの設計法やパラメータの設定によって変化するので、この結果だけで評価できない。

実験 2 から、2 個のフローが近い値を示す画像の右上の部分が単一フローと判定されている。画像の 4 隅で正しくない判定がされているのは、回転パターンの速度が大き過ぎて、周波数領域における折り返しノイズ (aliasing) が生じているためである。

実験 3 では、円形領域のうち、速度ベクトルが近い値を持つ左下の部分が単一フローと判定されている。それ以外の部分は、多重度が正しく判定されている。

図 2.10(f) の背景領域に注目する。ここでは、フローが 2 重になっており、1 重に縮退しない。一方、図 2.8(f) では、その左半分で、2 個のフローのうち片方が正しいフローをあらわしている。もう一方は、ほぼ $(0.0, 0.0)$ の付近でばらついた値を示している。図 2.9(f) の右上の領域についても、結果が不安定になっている。図 2.10(f) の円領域の外では、フローは不安定である。しかし、これは、図 2.8(f) の左の領域における振舞いとは異なる。余裕エネルギーが小さいと、固有値問題の数値計算が不安定になり、最小固有値に対

応する固有ベクトルの計算が不良設定になることを示す。実際の画像では、2個のフローが両方とも正しいフローを表さない場合があることを示す。この事実は、2重フローのモデルだけで1重フローと2重フローを両方検出することが、実際上困難であることを示す。

この様に、2重フローモデルは、単一フローとみなされる画像に対しては不安定な振舞いをする。この不安定さが2重フローモデルの余裕エネルギーの値を小さくする。逆に、図 2.8(e)、図 2.9(e)、図 2.10(e) の2重フローに対応する領域を見ると、フローベクトルが不安定なばらつきを示している。この事実は、これらの領域で単一フローモデルの余裕エネルギーが小さいことに対応している。

図 2.12 は、超並列計算機である Thinking Machines 社 CM-2で行ったシミュレーション結果である。接近するランダムテクスチャの背景パターンと回転する水玉状のランダムテクスチャパターンとの加法的重ね合わせによって生成された動画像シーケンスをシミュレーションデータとして用いた。多重オブティカルフローが得られている。

2.1.7 まとめ

画像面の各点に複数のフローベクトルの存在を許容する多重オブティカルフローの計算理論を提案した。このアプローチは、アルゴリズムのレベルで複数運動を扱うセグメンテーション、クラスタリングの手法とは異なる。オブティカルフローの基本拘束方程式である輝度勾配拘束と、その周波数領域における表現である共平面拘束を拘束方程式レベルの数学的操作によって多重運動に拡張した。多重フローの基本拘束方程式を直接時空間画像にあてはめるアルゴリズムによって、複数運動に対応する複数組のフローパラメータを直接同時に推定できることを示した。この基本拘束方程式に基づいた方法は、画像生成過程(光学)における信号の和あるいは積という物理を正確に反映している。そのため、アドホックな特徴抽出過程を前置する必要はない。本アプローチは、物理的に透明性をもつ物体による純粋な透明視だけでなく、オクルージョンによって生じる透明視とみなすことができる状況にもそのまま適用できる。これは、従来、フロー場の不連続、あるいは一般に正則化問題における場の不連続を扱うために行なわれてきた多くの研究に対する強力なアンチテーゼとなっている。セグメンテーションを行なうために不連続を検出したり、不連続を表す補助的な場を用意したりする必要は、必ずしもない。直接複数の属性を計算可能である。

ここで述べたアルゴリズムおよびインプリメンテーションは、多重オブティカルフローの基本拘束方程式の一解法にすぎない。工学的には、従来の単一フローに対して行なわれてきた多くの計算手法、周辺技術を用いて、各アプリケーションに適した計算アルゴリズムの構築が必要である。

2.2 多重ステレオ視差の計算理論

あらかし 重ね合わせの原理に基づいて、両眼立体視における透明視(トランスペアレンシー)の計算理論と計算モデルを提案する。透明視とは、複数の重なり合った半透明な表面の視知覚である。従来の工学的なステレオビジョンのアルゴリズムは、全くこの状況を想定していない。輝度ベースのステレオ視のためのエピポーラ整合条件を透明視に拡張する。これは、画像面の各点に複数の視差が存在する場合に成立する基本拘束方程式である。この方程式は、左右の画像の輝度の差と、それぞれの画像の輝度変化の間の拘束である。その近似方程式に基づいた2重視差の準線形アルゴリズムと一撃計算モデルを提案する。前節の多重オブティカルフローの計算理論では、 n 重フローに対して、 n 階までの時間微分を要する。したがって、最低 $n+1$ 枚の画像が必要であった。そのため、2眼ステレオ視に応用できない。提案する計算モデルは、左右2枚からなるステレオ画像対の各点において、2個の視差を線形フィルタの出力から直接一撃計算できる。つまり、アドホックな特徴抽出過程をステレオ整合に前置する必要がない。^a

^a本節の内容は文献 [180] [186] [187] [195] において発表された。

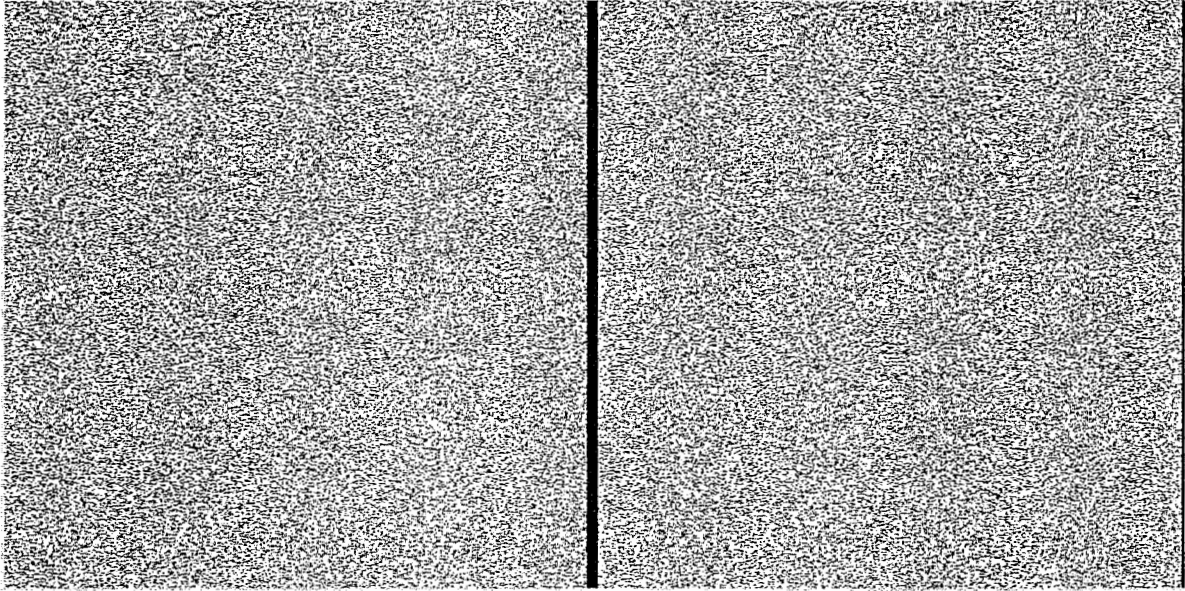


Figure 2.13: 純粹の透明視をもたらすランダム輝度ステレオグラム (交差法)

各画像の大きさは、 512×512 [pixel]、各面の視差は、それぞれ、背景： $D_1 = -5.0$ [pixel]、浮き出た中心部の矩形： $D_2 = 5.0$ [pixel]。

2.2.1 はじめに

両眼立体視 (ステレオ・ビジョン) のための両眼視差の検出は、視覚の計算理論において相当な注目を集めてきた [85] [89] [95] [97] [99]。ステレオ・ビジョンは、シーンの奥行き、あるいは3次元構造を受動的に獲得する方法として工学的にも盛んに研究されてきた [86] [92] [96]。しかし、これらのステレオ整合アルゴリズムは透明視 (第2.1節参照) を扱えない。ステレオ・ビジョンにおける透明視 (以下、ステレオ透明視と呼ぶ) とは、同一の網膜上の点あるいは、画像投影面上の点に複数の重なり合った表面を復元することである。近年、人間の両眼立体視における透明視の能力と限界が明らかにされてきた [83] [98] [100]。透明視は、オクルージョンや、物理的に光透過性をもつ物体、カメラの解像度以上に細かいオクルージョンが存在する複雑な環境を扱うコンピュータ・ビジョンの重要な課題でもある。これら複雑な環境は、フェンス、樹木、藪、水面、そして、窓ガラスなど、自然環境においても、人工環境においても共通に存在する。

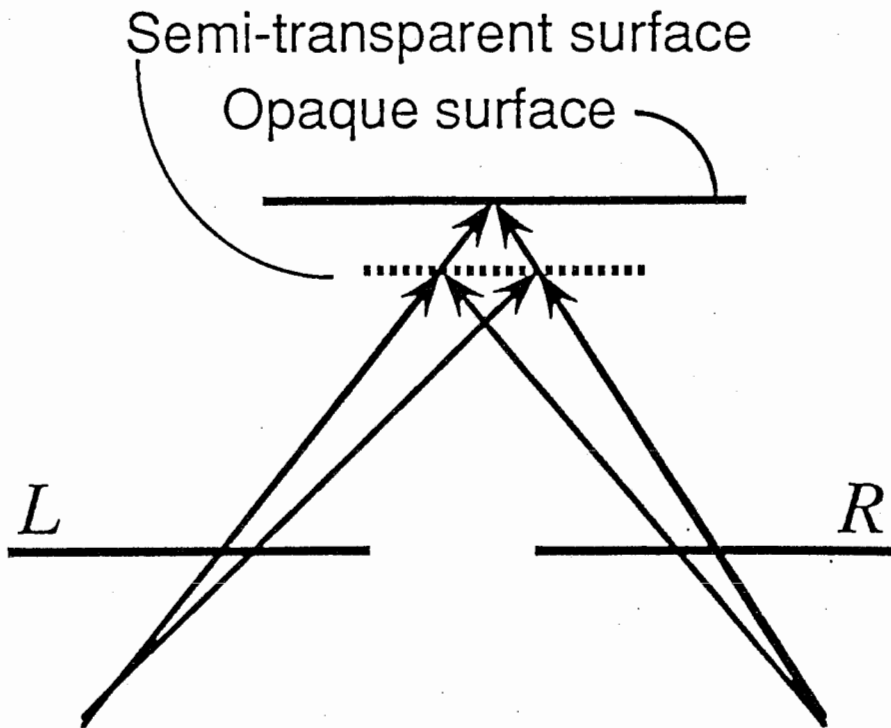


Figure 2.14: 平行2眼ステレオと透明視

近年、人間の両眼立体視において、パターンの繰り返し重畳によって作られる解釈が曖昧なステレオグラムに関する透明視の計算論的説明が、Weinshall [100] と Pollard et al.[98] によって論じられている。そこでは、ステレオ整合の一意性拘束 (uniqueness constraint) が透明視の場合に用いられているか否かが議論されている。いずれにしろ、まだ、十分に計算理論として深められているとはいえず、現象論的な説明にとどまっている。

Prazdny は、物体表面の不連続は、不連続の近傍に2枚の表面が存在するとみなせることを指摘した [99]。彼は、また、Burtらによって発見された人間の視差勾配限界 (disparity gradient limit) [85] の拘束に基づいた特徴点ベースのステレオ整合アルゴリズムを提案した。Marr & Poggioのアルゴリズム [95] は、近傍の視差に対する抑制的な相互作用を持つ。しかし、Prazdnyのアルゴリズムは、この相互作用を持たない。したがって、複数の重なり合う表面を扱うことができる。Pollardらも、これに類似したステレオ整合アルゴリズム (PMF アルゴリズム) を提案している [97]。Pollardらは、このモデルによって、人間の曖昧な多重表面の知覚現象をも説明できると主張している [98]。

しかし、以上のステレオ透明視アルゴリズムは、すべてランダムドットステレオグラムを想定した特徴点ベースのステレオ整合である。信号レベルでの透明視の生成過程をモデル化していない。したがって、純粋に信号の重畳によって生じる透明視（例えば、窓ガラスの透過像と反射像）に適さない。従来の特徴検出演算子は暗黙のうちに表面の非透明性を仮定しているからである。この場合には、各表面のパターンが画像生成過程で重畳されたあとで特徴抽出が行われる。このため、従来のエッジやかどの特徴検出演算子では、異なるの表面上の特徴間で干渉が起こり、誤った特徴が混入する。この場合は、画像生成過程の分析結果に基づいた輝度ベースの整合がより適切である。人間は、ランダムな輝度パターンの重畳によって作成された図 2.13 のステレオグラム（以下、ランダム輝度ステレオグラムと呼ぶ）からも多重表面を知覚できる。

従来、視覚の計算理論で取り上げられてきたステレオ透明視は、すべて密度の低いランダムドット（従って、ドット同士の重なりが少なく、空間的に特徴が分離されていて、特徴検出器の空間解像度の方がドット密度よりも十分高い）によって作られたステレオグラムを対象としており、ここで問題になる各表面パターン間の信号レベルの干渉作用は無視できた [99] [98]。一方、輝度ベースの整合 [92] [96] では、局所的に複数の視差を扱うことは、従来考えられていない。

図 2.14 は、もっとも簡単な平行ステレオの場合の透明視を説明している。手前の表面は、半透明なパターンを持つ面で、奥の表面は、不透明なパターンを持つ面である（一般に、人間は、もっとも奥の表面だけ不透明に感じ、その手前にある表面は透明に感じる）。左右の画像の各点に各表面に対応する 2 つの視差を検出することによって 2 枚の重なり合う表面の 3 次元奥行き情報を復元することができる。

本節では、セグメンテーションを経ずに、両眼の画像対から複数の視差を直接一撃的に復元するアルゴリズムを提案する。

まず、そのための計算理論を導く。画像の輝度が 2 枚の表面のパターンの加法的重ね合わせによって生成され、画像面の各点に 2 個の視差が存在するという仮定をする¹⁰。これを用いて、ステレオ画像対のエピポーラ線上の輝度変化に関する直接的拘束を表す線形基本拘束方程式を導出する。これは、前節の多重オプティカルフローと同様に、重ね合わせの原理を基礎にしている。

提案するアルゴリズムは、前節で述べた多重オプティカルフロー推定アルゴリズムの特殊な場合ではない。多重オプティカルフローの場合には、 n -重のフロー推定のためには n -階

¹⁰多重オプティカルフローの場合と同様に、一般化重畳 [66] に拡張すれば、乗法的重ね合わせの場合を含んだかなり広い範囲の重畳に対応できる。また、厳密には重畳の条件を満たさなくても、特定の周波数サブバンドでは、近似的に重畳の条件が成立する場合が多い。

の時空間偏微分が必要であり、 n -階の時間偏微分には、最小 $n + 1$ 枚の画像フレームが必要であった。ここで提案するアルゴリズムは、2枚の画像からエピポーラ線上の各点において複数の視差を計算できる。

以下、まず、第 2.2.2 節において、ステレオ透明視について説明する。次に、重ね合わせの原理にもとづいて、演算子形式によるステレオ透明視の形式的理論を展開して基本拘束方程式を導く。第 2.2.3 節で、基本拘束方程式の近似方程式を導出する。この近似拘束方程式は、ステレオ画像対の左右の画像の輝度微分に関して線形である。その係数は、複数視差の基本対称式に関して線形である。この近似拘束方程式の局所最適化問題を変分法で解くことによって、2個の視差を計算する閉形式アルゴリズムを導出する。このアルゴリズムに基づいて一撃計算ネットワークが得られる。第 2.2.4 節では、1次元信号およびランダム輝度ステレオグラムを用いたアルゴリズムの検証を行う。第 2.2.5 節で本研究をまとめる。

2.2.2 ステレオ透明視のための多重視差推定の形式的計算理論

本章では、まず、輝度ベースのステレオ整合拘束条件を、振幅演算子 (線形演算子) を用いて表記する。そして、重ね合わせの原理と合成演算子の対称化操作を用いて、ステレオ透明視の基本拘束方程式を導出する。なお、以下では、2重の透明視を述べる。3重以上の場合に理論を拡張することも原理的には可能である。

2.2.2.1 演算子形式によるステレオ整合の基本拘束条件の記述

不透明な表面に対する、輝度ベースのステレオ整合の拘束条件は次式である。

$$L(x, y) - R(x - D, y) = 0, \quad R(x, y) - L(x + D, y) = 0 \quad (2.30)$$

$L(x, y)$ と $R(x, y)$ は、それぞれ、左右の画像の輝度分布関数を表す。 D は x 方向の視差を表す。 x はエピポーラ線にそった座標である。 y はそれに垂直な方向の座標である。以下、平行2眼ステレオにおいて、視差は微小で、空間的に十分緩やかに変化し、局所的に定数と見なせると仮定する。式 (2.30) は、振幅演算子 $a(D)$ とデータ分布関数 $f(x, y)$ を用いて、次式に書き表わせる。

$$a(D)f(x, y) = 0 \quad (2.31)$$

ただし、

$$a(D) \stackrel{\text{def}}{=} \begin{bmatrix} 1 & -D(D) \\ -D(-D) & 1 \end{bmatrix}, \quad f(x, y) \stackrel{\text{def}}{=} \begin{bmatrix} L(x, y) \\ R(x, y) \end{bmatrix} \quad (2.32)$$

$D(D)$ は、 x 軸方向への D だけのずらし演算子であり、 $f(x)$ を $f(x - D)$ に変換する。図 2.15 に、ステレオ画像対に対する振幅演算子の働きを図示した。

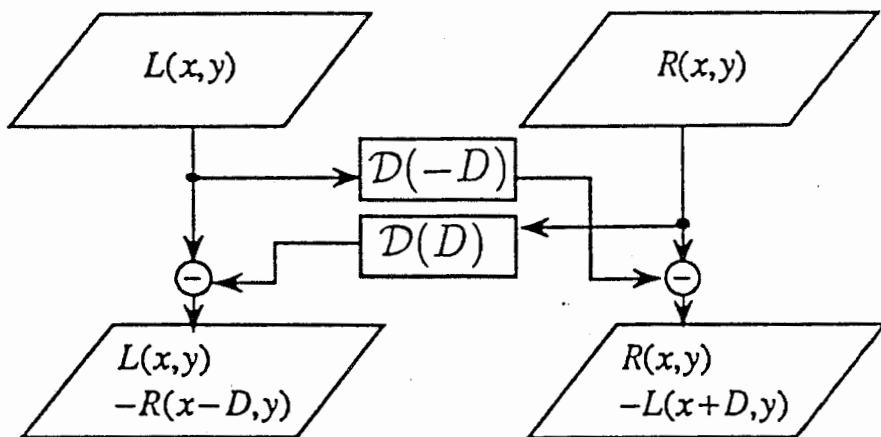


Figure 2.15: ステレオ視差検出のための振幅演算子のはたらき

$f(x - D)$ を x において D に関してテイラー展開すると次式が得られる。

$$f(x - D) = f(x) - D\partial_x f(x) + \frac{D^2}{2!}\partial_x^2 f(x) - \dots = \sum_{i=0}^{\infty} \frac{(-D)^i}{i!}\partial_x^i f(x) \tag{2.33}$$

$\partial_x^i f(x)$ は $f(x)$ の x に関する i 階偏導関数を表す。これから、ずらし演算子は擬微分演算子

$$\mathcal{D}(D) \stackrel{\text{def}}{=} \sum_{i=0}^{\infty} \frac{(-D)^i}{i!}\partial_x^i = \exp(-D\partial_x) = 1 - D\partial_x + \frac{D^2}{2!}\partial_x^2 - \frac{D^3}{3!}\partial_x^3 + \dots \tag{2.34}$$

として書かれることがわかる。したがって、式 (2.33) は $f(x - D) = \mathcal{D}(D)f(x)$ と書くことができる。このずらし演算子は線形演算子である。すなわち、 $\mathcal{D}(D)(f_1(x) + f_2(x)) = \mathcal{D}(D)f_1(x) + \mathcal{D}(D)f_2(x)$ および $\mathcal{D}(D)0 = 0$ が成立する。このずらし演算子はその合成積に関して閉じており、

$$\mathcal{D}(D_\alpha)\mathcal{D}(D_\beta) \equiv \mathcal{D}(D_\beta)\mathcal{D}(D_\alpha) \equiv \mathcal{D}(D_\alpha + D_\beta) \tag{2.35}$$

が成り立つ。

拘束方程式 (2.31) は、画像の輝度分布関数 $L(x, y)$ と $R(x, y)$ の高階の導関数に関する拘束に拡張することができる。 $L(x, y)$ と $R(x, y)$ の (p, q) 階偏導関数を $L^{(p,q)}(x, y) = \partial_x^p \partial_y^q L(x, y)$ および $R^{(p,q)}(x, y) = \partial_x^p \partial_y^q R(x, y)$ とし、これらを合わせて $f^{(p,q)}(x, y) = [L^{(p,q)}(x, y), R^{(p,q)}(x, y)]^T$ と書くことにする。すると、拘束方程式 (2.31) は $p, q = 0, 1, 2, \dots$ に対して、

$$a(D)f^{(p,q)}(x, y) = 0 \quad (2.36)$$

と拡張できる。ある画像点における (p, q) 階偏導関数のすべての集合はその点の周りの画像の局所構造を完全に特徴付ける [90]。この拡張によって、特徴ベースの整合と輝度ベースの整合が統一される。

ずらし演算子の線形性から振幅演算子 $a(D)$ もまた線形演算子である。

2.2.2.2 重ね合わせの原理を用いたステレオ透明視の基本拘束方程式

前節のステレオ整合の定式化に基づいて、加法的 2 重透明視の場合に重ね合わせの原理を形式的に適用すると、次の拘束方程式が得られる。

$$a(D_2)a(D_1)f^{(p,q)}(x, y) = 0 \quad (2.37)$$

ここで、 $f^{(p,q)}(x, y) = f_1^{(p,q)}(x, y) + f_2^{(p,q)}(x, y)$ であり、各 $f_i^{(p,q)}(x, y)$ は $a(D_i)f_i^{(p,q)}(x, y) = 0$ ($i = 1, 2$) と拘束されている。

この場合、振幅演算子の合成積 (行列としての積) が可換でない。実際、

$$a(D_1)a(D_2) \equiv \begin{bmatrix} 1 + \mathcal{D}(-D_1)\mathcal{D}(D_2) & -\mathcal{D}(D_1) - \mathcal{D}(D_2) \\ -\mathcal{D}(-D_1) - \mathcal{D}(-D_2) & 1 + \mathcal{D}(D_1)\mathcal{D}(-D_2) \end{bmatrix} \quad (2.38)$$

および

$$a(D_2)a(D_1) \equiv \begin{bmatrix} 1 + \mathcal{D}(-D_2)\mathcal{D}(D_1) & -\mathcal{D}(D_2) - \mathcal{D}(D_1) \\ -\mathcal{D}(-D_2) - \mathcal{D}(-D_1) & 1 + \mathcal{D}(D_2)\mathcal{D}(-D_1) \end{bmatrix} \quad (2.39)$$

であり、これらの差をとると、次式が得られる。

$$a(D_2)a(D_1) - a(D_1)a(D_2) = \begin{bmatrix} -K & 0 \\ 0 & K \end{bmatrix} \quad (2.40)$$

成分 K は次式で定義される。

$$K \stackrel{\text{def}}{=} \mathcal{D}(D_1 - D_2) - \mathcal{D}(D_2 - D_1) \\ = 2 \left\{ -(D_1 - D_2)\partial_x - \frac{(D_1 - D_2)^3}{3!}\partial_x^3 - \frac{(D_1 - D_2)^5}{5!}\partial_x^5 - \dots \right\} \quad (2.41)$$

したがって、演算子 $a(D_1)$ と $a(D_2)$ は、 $D_1 \neq D_2$ である限り、可換でない。振幅演算子の積の順序を変えると拘束方程式の意味が変化してしまい、かつ、基本拘束方程式 (2.37) の左辺は一般の信号に対して零にならない。前者の問題は、この振幅演算子の積に、視差 D_1 、と D_2 に関する対称化を施すことによって解決される。後者の問題についても、式 (2.37) の左辺の2乗エネルギーをとり、そのエネルギー最小化問題と捉えて回避する。

対称化された振幅演算子は次式で定義する。

$$\bar{a}^{(2)}(D_{(1), D_2}) \stackrel{\text{def}}{=} \frac{1}{2} \{a(D_1)a(D_2) + a(D_2)a(D_1)\} \quad (2.42)$$

添字に付けられた括弧は、それに囲まれた添字に関して対称化の操作がなされていることを表す。

2.2.3 2重視差の準線形アルゴリズムと一撃計算モデル

ステレオ透明視の基本拘束方程式は、視差 D_1, D_2 に関してきわめて非線形性が高い。そのままでは計算機による実現に適さない。本節では、振幅演算子の近似から導かれた両眼視差の準線形アルゴリズムと一撃計算モデルを述べる。

2.2.3.1 ステレオ透明視の近似拘束方程式と準線形アルゴリズム

本節では、2重視差の場合を述べる。単一視差の場合の近似方程式と従来手法との関係性を、付録 A.2.1 に示す。以下に述べる2重視差の拘束方程式は、この従来法の方程式の拡張である。

2.2.3.2 2重視差の近似拘束方程式

$|D| \ll 1$ の条件のもとで、ずらし演算子 $D(D)$ に対する一次近似

$$D(D) \approx 1 - D\partial_x \quad (2.43)$$

を考える。これを用いると2重視差の場合の対称化された振幅演算子は次の形に近似される(導出は付録 A.2.2 を参照)。

$$\bar{a}^{(2)}(D_{(1), D_2}) \approx \bar{\alpha}^{(2)}(D_{(1), D_2}) = \begin{bmatrix} 2 - D_1 D_2 \partial_x^2 & -2 + (D_1 + D_2) \partial_x \\ -2 - (D_1 + D_2) \partial_x & 2 - D_1 D_2 \partial_x^2 \end{bmatrix} \quad (2.44)$$

1階までの微分項では、係数に視差の関数として、 $(D_1 + D_2)$ しか現れない。これは、2重視差を推定するには不十分である。これが、2階までの微分項を用いる理由である。

以上から、残差関数は次式である。(0にならないので残差関数と呼ぶ。)

$$\begin{aligned} r_L^{(2)}[p, q](x, y) &\stackrel{\text{def}}{=} 2 \{L^{(p, q)} - R^{(p, q)}\} + (D_1 + D_2)R^{(p+1, q)} - D_1 D_2 L^{(p+2, q)} \\ r_R^{(2)}[p, q](x, y) &\stackrel{\text{def}}{=} 2 \{R^{(p, q)} - L^{(p, q)}\} - (D_1 + D_2)L^{(p+1, q)} - D_1 D_2 R^{(p+2, q)} \end{aligned} \quad (2.45)$$

ただし、 $L^{(p, q)} = L^{(p, q)}(x, y)$, $R^{(p, q)} = R^{(p, q)}(x, y)$ と省略した。これらの式を観察すると次のことがわかる。最初の式は、左右の輝度 (の (p, q) 階微分) の差 $L^{(p, q)}(x, y) - R^{(p, q)}(x, y)$ 、右画像 (の (p, q) 階微分) の x 方向 1 階微分 $R^{(p+1, q)}(x, y)$ 、左画像 (の (p, q) 階微分) の x 方向 2 階微分 $L^{(p+2, q)}(x, y)$ の間の近似的な線形従属関係を表している。また、2 番目の式は最初の式の左右の画像の役割を入れ換えたものである。

2.2.3.3 準線形アルゴリズム

点 (x_0, y_0) における視差 D_1 、 D_2 を推定する。そのために、次のエネルギー関数の局所最小化を行う。

$$E^{(2)}(D_1(x_0, y_0), D_2(x_0, y_0)) = \int_{\mathcal{R}^2} \{w_{p, q}(x - x_0, y - y_0)\}^2 \left[\{r_L[p, q](x, y)\}^2 + \{r_R[p, q](x, y)\}^2 \right] dx dy \quad (2.46)$$

$w_{p, q}(x - x_0, y - y_0)$ は、積分領域と画像微分量の統計的信頼度に応じて拘束式の寄与度を定める重み関数である。一般に高次の画像微分量は誤差が大きい。それに応じて重みは小さくする必要がある。後に述べるシミュレーションでは、微分の階数 $(p + q)$ を固定し、重み関数は矩形領域の定数関数とした。エネルギー関数 (2.46) は、2 つの視差に関する基本対称式 $s_1 = \frac{1}{2}(D_1 + D_2)$ 、 $s_2 = D_1 D_2$ に関して 2 次形式である。変数 D_1 、 D_2 に関するエネルギー最小化を行う代わりに、 s_1 、 s_2 を未知変数とみなして最小化を行う。エネルギー関数の平衡 (最小化) 条件 $\frac{\partial E^{(2)}}{\partial s_1} = 0$ 、 $\frac{\partial E^{(2)}}{\partial s_2} = 0$ から、次の連立 1 次方程式 (正規方程式) を得る。

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \quad (2.47)$$

係数は次式で定義される。

$$\begin{aligned} a_{11} &= 4 \sum_{p, q} \int_{\mathcal{R}^2} (w_{p, q})^2 \left\{ (R^{(p+1, q)})^2 + (L^{(p+1, q)})^2 \right\} dx dy, \\ a_{12} &= 2 \sum_{p, q} \int_{\mathcal{R}^2} (w_{p, q})^2 \left(L^{(p+1, q)} R^{(p+2, q)} - L^{(p+2, q)} R^{(p+1, q)} \right) dx dy, \\ a_{22} &= \sum_{p, q} \int_{\mathcal{R}^2} (w_{p, q})^2 \left\{ (R^{(p+2, q)})^2 + (L^{(p+2, q)})^2 \right\} dx dy, \end{aligned}$$

$$\begin{aligned}
b_1 &= 4 \sum_{p,q} \int_{\mathcal{R}^2} (w_{p,q})^2 (R^{(p+1,q)} + L^{(p+1,q)}) (R^{(p,q)} - L^{(p,q)}) dx dy, \\
b_2 &= 2 \sum_{p,q} \int_{\mathcal{R}^2} (w_{p,q})^2 (L^{(p+2,q)} - R^{(p+2,q)}) (L^{(p,q)} - R^{(p,q)}) dx dy
\end{aligned} \tag{2.48}$$

ただし、 $w_{p,q} = w_{p,q}(x - x_0, y - y_0)$ と省略した。

2つの視差 D_1 と D_2 ($D_1 \geq D_2$) は、 d に関する2次方程式の2実根である。

$$d^2 - 2s_1 d + s_2 = 0 \tag{2.49}$$

したがって、次式で2個の視差が求められる。

$$D_1 = s_1 + \sqrt{s_1^2 - s_2}, \quad D_2 = s_1 - \sqrt{s_1^2 - s_2} \tag{2.50}$$

根は、実数でなければならないので、

$$\Delta^{(2)} \stackrel{\text{def}}{=} s_1^2 - s_2 \geq 0, \tag{2.51}$$

が必要である。 $\Delta^{(2)}$ は、代数方程式 (2.49) の判別式である。したがって、本来は、式 (2.51) の条件のもとでエネルギー関数 (2.46) を最小化する必要がある。しかし、この不等式のもとでの最小化は解析解を持たない。そこで、計算の便宜上、条件 (2.51) を課さずに解を求める。判別式が負になる場合は解に信頼性が無いものと見なす。

以下、判別式 $\Delta^{(2)}$ に基づいた判定アルゴリズムを述べる。解が重根になる場合を単一視差として定義する。あるしきい値 $\epsilon_{\text{threshold}}$ によって、

$$|\Delta^{(2)}| \leq \epsilon_{\text{threshold}} \tag{2.52}$$

となる場合を検出する。この場合、単一視差は、 $D = s_1$ である。 $\Delta^{(2)} \leq -\epsilon_{\text{threshold}}$ の場合を計算に信頼性が無い場合、 $\Delta^{(2)} \geq \epsilon_{\text{threshold}}$ の場合を2重視差の場合と判定する。

2.2.3.4 局所ガウススケール微分を用いた多重スケール画像情報表現とステレオ透明視の計算モデル

前節では、画像の輝度微分を計算できることを前提としてアルゴリズムを導出した。しかし、微分演算は不良設定問題である。したがって、正則化手法を用いないとロバストではない。ここでは、ステレオ整合の局所画像表現として、局所ガウススケール微分 [90] を用いる。この方法は、線形基底フィルタ集合の出力をステレオ整合の基本要素とした Jones & Malik [88] のアプローチと近い。しかし、画像表現それ自身は、ステレオ透明視の計算理論とは独立である。本論文で、局所ガウススケール微分を表現として用いるのは、アルゴリズム

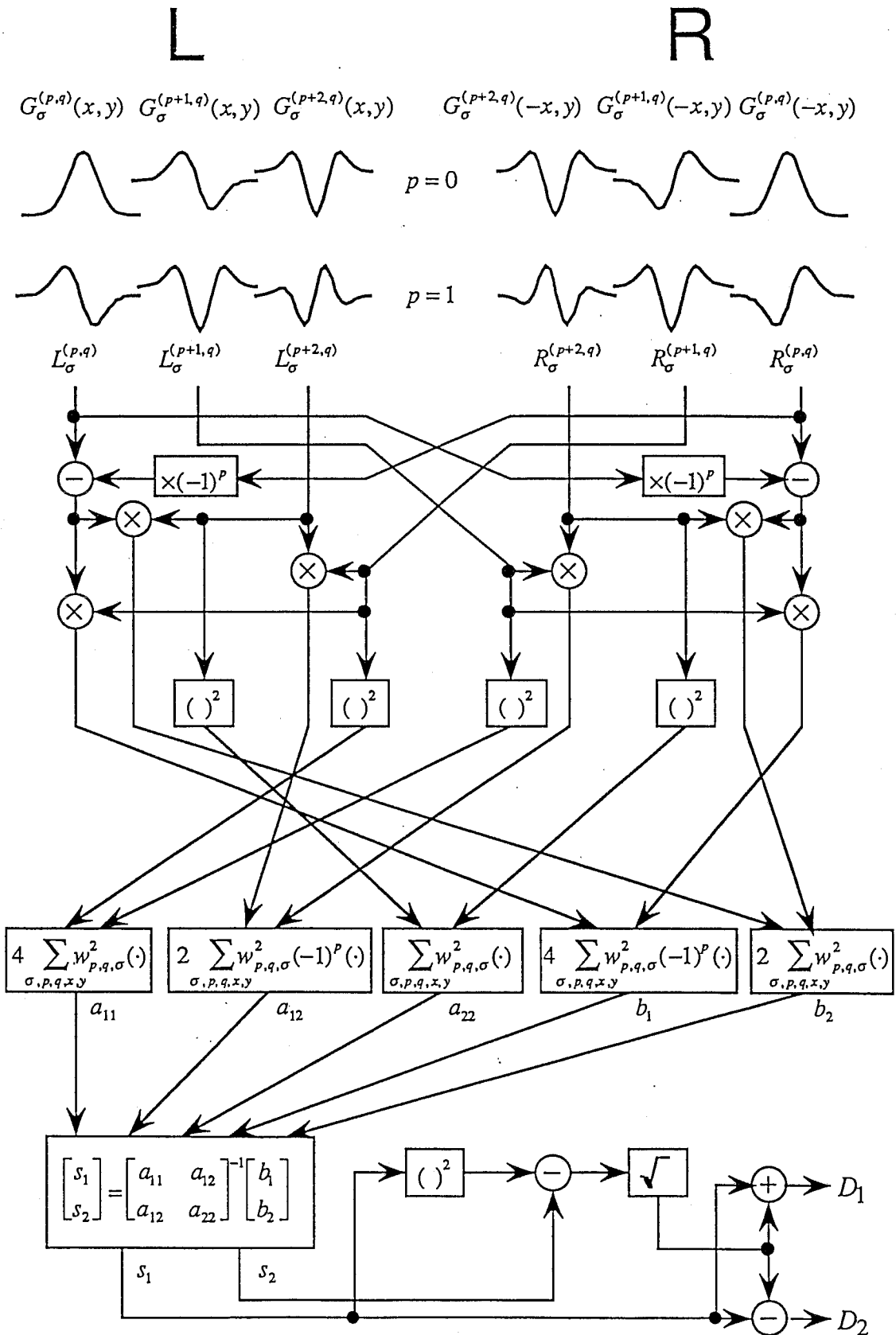


Figure 2.16: 2重視差に対する一撃アルゴリズムのネットワーク表現

ムに必要な情報である輝度微分を直接得られるからである。もちろん、第2.1.6節の多重オブティカルフローのアルゴリズムと同様、Gabor関数を用いることも可能である。

(p, q) 階の局所ガウススケール微分は、ガウス関数の (p, q) 階微分と画像の畳み込み積を計算して得られる。

$$\begin{aligned}
 L_{\sigma}^{(p,q)}(x, y) &= \int_{\mathcal{R}^2} G_{\sigma}(x-u, y-v) L^{(p,q)}(u, v) du dv \\
 &= \int_{\mathcal{R}^2} G_{\sigma}^{(p,q)}(x-u, y-v) L(u, v) du dv, \\
 R_{\sigma}^{(p,q)}(x, y) &= \int_{\mathcal{R}^2} G_{\sigma}(x-u, y-v) R^{(p,q)}(u, v) du dv \\
 &= \int_{\mathcal{R}^2} G_{\sigma}^{(p,q)}(x-u, y-v) R(u, v) du dv
 \end{aligned} \tag{2.53}$$

2次元ガウス関数 $G_{\sigma}(x, y)$ は、

$$G_{\sigma}(x, y) \stackrel{\text{def}}{=} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \tag{2.54}$$

と定義される。

図2.16は、導出されたアルゴリズムのネットワークモデルである。ここでは、微分階数 p, q と、スケール σ に関する統合も取り入れた。 $w_{p,q,\sigma}$ は、統合の際の重み係数を表す。生体情報処理では、Gabor関数の様に、quadrature pairを用いて画像を表現するモデルが普通である。ガウス関数の k 階微分と $(k+1)$ 階微分の間にはこのフィルタの組と類似な関係がある。図2.16の上部に示した例では、 $p=0$ と $p=1$ の場合について、そこで用いられるフィルタの x 方向の断面形状を示している。この場合は、ガウス関数の0階から3階までの微分を畳み込み演算として得る。

2.2.4 シミュレーション実験

本節では、まず、1次元信号によるシミュレーション結果を示す。さらに、ランダム輝度ステレオグラムから2重視差表面を復元した例を示す。

2.2.4.1 実験1：1次元信号を用いたシミュレーション

乱数とガウス関数による畳み込みによって生成した信号を用いた。この信号から図2.17に示す2通りの視差マップを持つ1次元ステレオ対を作る。次に、それらを左右独立に輝度の和をとって1/7にサブサンプルした。図2.18に生成した信号を示す。

実験には、 x 方向の5階までの局所ガウススケール微分を用いた。シミュレーションでは、ガウススケールパラメータは $\sigma = 4[\text{pixel}]$ とした。

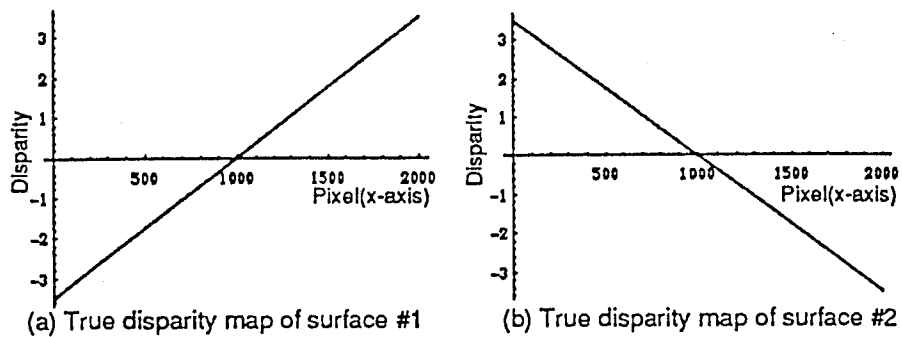


Figure 2.17: 1次元信号を用いたシミュレーションにおける2つの視差マップ

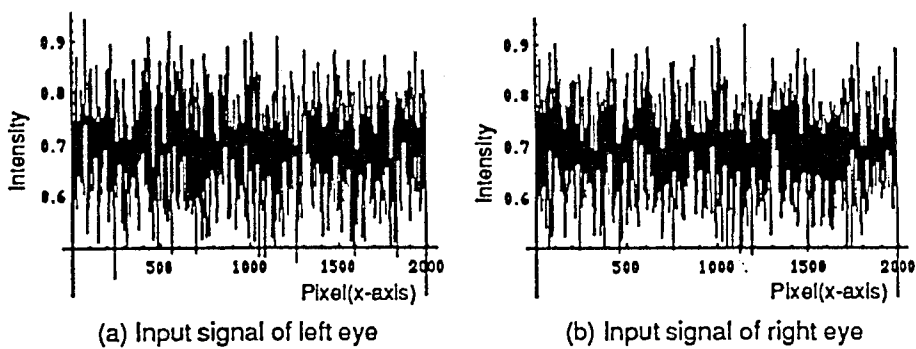


Figure 2.18: シミュレーションのために生成された2重視差を含む1次元信号

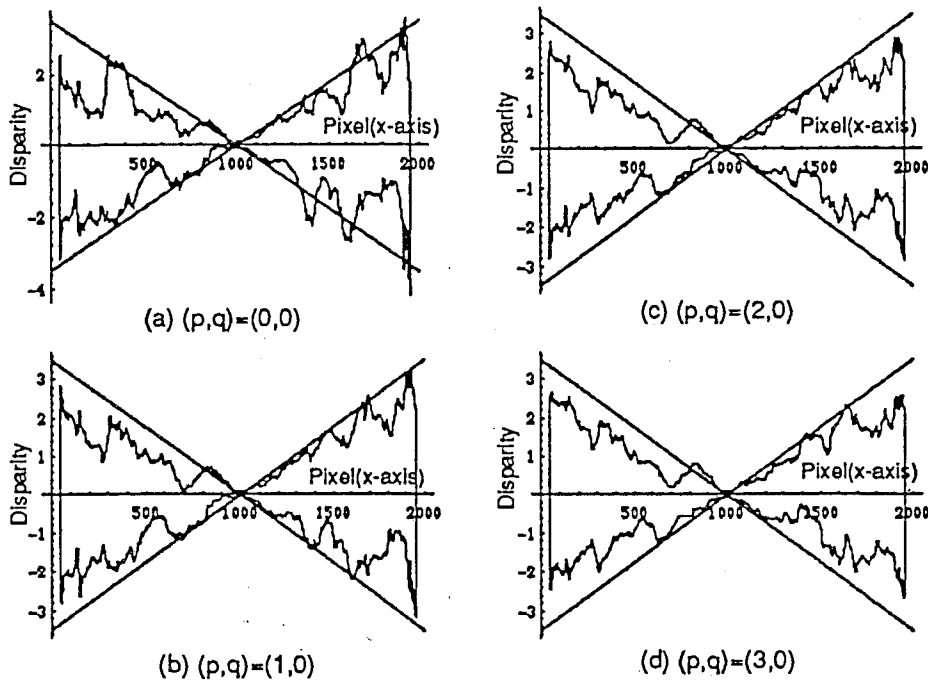


Figure 2.19: 一撃アルゴリズムによる2重視差推定結果(真の視差マップとの重畳表示)

エネルギー積分の重み関数 $w_{p,q}(x)$ は、幅 $M = 121[\text{pixel}]$ の矩形窓とした。計算は、 $p = 0, 1, 2, 3$ のそれぞれの場合に対して独立に行った。図 2.19 は、アルゴリズムの最終出力をそれらの真値とともに重畳したプロットである。

2.2.4.2 実験2：2次元信号を用いたシミュレーション

2次元信号を用いたシミュレーションは、超並列計算機 CM-2 上で行った。図 2.20 は、図 2.13 と同様のランダム輝度ステレオグラムに対する結果である。ただし、視差のパラメータは、 $D_1 = -0.5[\text{pixel}]$, $D_2 = 0.5[\text{pixel}]$ とした。フィルタは、 $p + q = 2$ 、 $\sigma = 1.6[\text{pixel}]$ のものだけを用いた。判別条件 (2.52) のしきい値は $\epsilon_{\text{threshold}} = 0.11$ とした。エネルギー関数の積分領域は、 $25 \times 25[\text{pixel}]$ の矩形領域とした。積分領域内で重み関数は、 $w_{2,0} = w_{0,2} = w_{1,1} = 1$ とすべて等しくした。図 2.20(a) は、計算された視差を高さで表した視差マップである。枠の領域では、正しく単一視差の判定が行われている。中

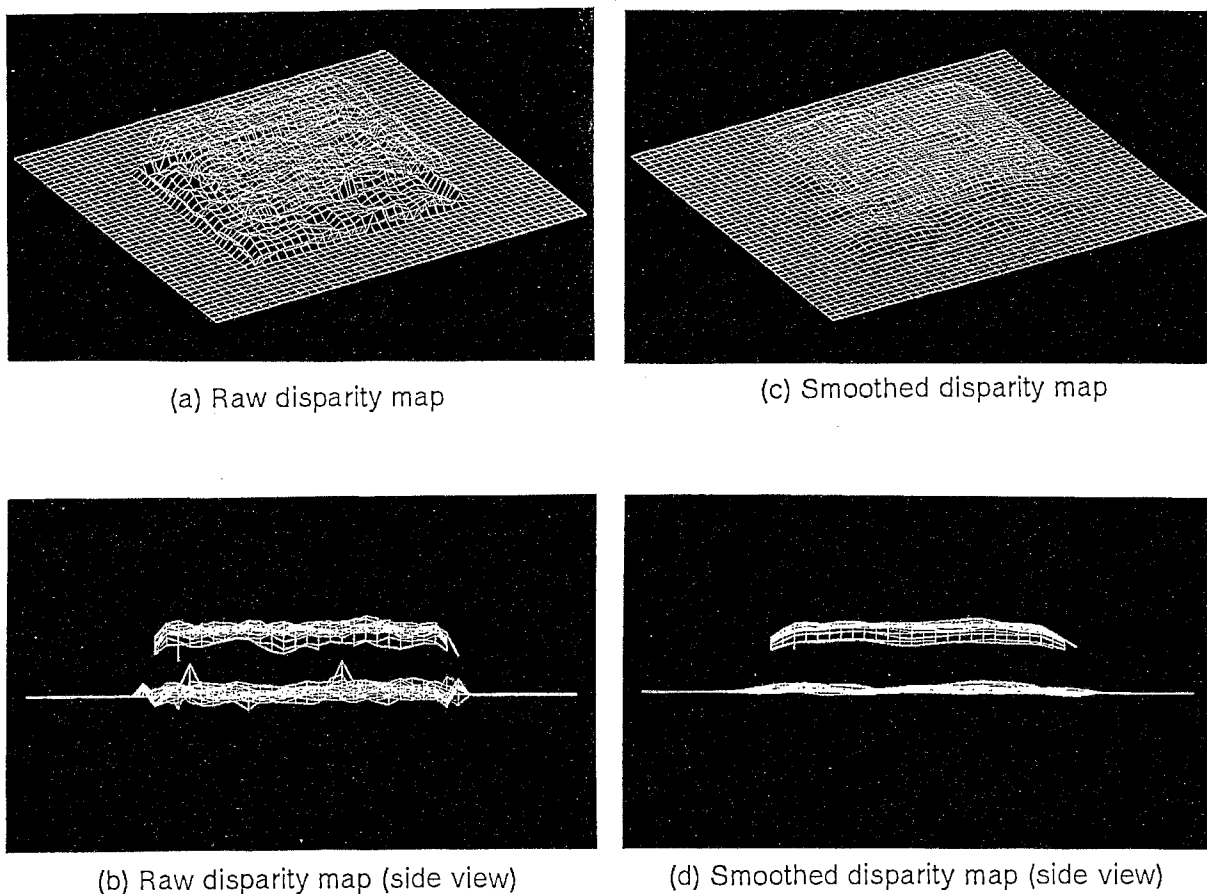


Figure 2.20: 2次元のランダム輝度ステレオグラムに対する推定結果。
(a) 推定された視差マップ、(b) (a)の側面図、(c) (a)のデータに移動平均を施したもの、(d) (c)の側面図。

心領域では、ノイズ成分は多くなるが、2重視差は計算されている。図 2.20(b) は、視差の平均化操作 (20×20 [pixel] の窓による移動平均) を施した結果である。2重視差の領域でノイズ成分が多い原因は、2枚の表面パターン間の干渉作用である。

2.2.5 まとめ

両眼立体視におけるトランスペアレンシーの計算理論とアルゴリズム、多重視差の一撃計算モデルを提案した。この計算理論は、演算子形式で表現された透明視の普遍的計算理論である重ね合わせの原理に基づいて導出された。導出された基本拘束方程式を各視差に対応する振幅演算子について1次近似した。局所エネルギー最小化問題を解いて、準線形一撃アルゴリズムを導いた。計算機で生成された1次元信号を用いたシミュレーションによってこの一撃アルゴリズムが2重視差を計算できることを示した。さらに、単一視差の領域と2重視差の領域を判定できることを示した。従来のステレオ透明視の計算理論は、特徴の空間的分離抽出が容易な密度の低いランダムドットステレオグラムを対象としていた。それに対して、本計算理論は、特徴抽出段階では空間的に分離不可能なランダムな輝度パターンの重畳によって生成されたステレオグラムに対しても、局所的なフィルタ群の出力から透明視が可能であることを計算論的に説明している。

ここで示したアルゴリズムは、視差がフィルタのスケール σ に比較して十分小さいことを前提としたアルゴリズムである。一撃法であるため精度もよくない。工学的により精密な視差の推定が欲しければ、繰り返し画像位置合わせ法 [92] の様な繰り返し修正アルゴリズムに発展させることが考えられる。大きな視差に対しては、局所ガウススケール微分のスケールパラメータの制御が重要になる。また、現在は、異なる微分階数と異なるスケール間での情報統合に必要な重み関数については理論的な可能性を示しただけである。具体的な手法には、例えば、統計モデルに基づいた適応的な窓の大きさの選択法 [96] などの発展的応用が考えられる。

また、ここでは、2重視差の基本拘束方程式とその検証のためのアルゴリズムの導出に重きをおいた。今後は、近年開発され洗練されてきたコンピュータビジョンにおける計算結果の信頼度を表す共分散の計算法などの誤差や偏差の一般的解析方法や、ステレオ整合に適した画像表現のための基底核関数の設計理論、非線形を重ね合わせや、表面の滑らかさ拘束の導入などについて検討することが必要である。また、視差情報に基づいた3次元表面の復元を行う際に問題となる表面の不連続やオクルージョン [91] について、本アルゴリズムがどの程度有効な情報を抽出しうるかを検証する必要がある。以上の意味で、本アルゴリズムは両眼立体視アルゴリズムの一要素である。

本節の計算理論とアルゴリズムは、2枚の画像フレームからの多重オプティカルフロー推定に発展させることもできる。これは、視差を x 方向だけでなく y 方向にも定義すればよい。この拡張によって、2重運動に関する3フレームアルゴリズム (付録 A.1.7、または、文献 [30] 参照) の欠点であるパターンの加速度に対する弱さを克服することが期待できる。

2.3 重ね合わせの原理の一般表現

あらまし 多重オブティカルフロー、多重ステレオ視差検出において、基本拘束方程式を導くために用いた「重ね合わせの原理」を数学的に一般化する。この一般化された重ね合わせの原理は、いろいろなビジョンの問題に適用できる普遍法則である。まず、ビジョンの既存理論を体系的に拡張するために、演算子形式を用いて基本拘束方程式を表現する方法を説明する。次に、この演算子の合成によって複数属性値、特に多重運動、多重表面の復元のための基本拘束方程式を全く形式的に得る方法を提案する。重ね合わせの原理自身は、物理学の量子力学や波動理論などで常識であるが、ここで述べる原理は、単なる関数の和を表すのではなく、重畳によって生じた新たな関数が満たす拘束方程式を自動的に導く方法を与えるものである。^a

^a本節の内容の一部は、文献 [193] [194] において発表された。

2.3.1 はじめに

画像生成過程においては、複数の画像生成要因が複雑に絡み合って作用するため、単一機能の視覚モジュールでは、これを正しく扱うことができない。2.1 節の多重オブティカルフロー、2.2 節の多重ステレオ視差の計算理論は、この典型的な場合であった。そこでは、複数運動パターンの和あるいは積の形の重ね合わせが画像生成過程における複数要因の合成過程となる場合について、基本拘束方程式を導く一般的方法を示唆している。本章で述べてきた多重オブティカルフローと多重視差の場合には、基本拘束方程式は、(擬)微分演算子の形式をしていた。本節では、これを微分とモーメントの任意の合成からなる演算子に拡張する。これによって、非常に広い範囲の拘束方程式を振幅演算子とデータ分布関数を用いて表現することができる。

以下、この手法を一般的に記述し、他のビジョンの計算理論への適用を可能にする。本節の結果は、次章で用いられる。

2.3.2 演算子とデータ分布関数による基本拘束方程式の一般的記述

多重オブティカルフローの計算理論では、時空間における微分演算子 (振幅演算子)

$$a(u, v) \equiv u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} + \frac{\partial}{\partial t} \quad (2.55)$$

を用いて、時空間画像 $f(x, y, t)$ に関する拘束を

$$a(u, v)f(x, y, t) = 0 \quad (2.56)$$

と表現することから始めた。

一方、多重ステレオ視差の計算理論では、次式に示す行列微分演算子の形をした振幅演算子 $a(D)$ と左右画像の輝度分布からなるベクトル値データ関数を用いて基本拘束方程式を記述した。

$$a(D)f(x, y) = 0$$

$$a(D) \equiv \begin{bmatrix} 1 & -\mathcal{D}(D) \\ -\mathcal{D}(-D) & 1 \end{bmatrix}, \quad f(x, y) \equiv \begin{bmatrix} L(x, y) \\ R(x, y) \end{bmatrix} \quad (2.57)$$

以下では、これらの定式化を一般化する。

2.3.2.1 基本拘束方程式の一般的記述

コンピュータ・ビジョンにおける種々の拘束方程式は、未知パラメータと、画像、あるいは、画像から2次的に得られる特徴量を含む拘束方程式で表現されているのが普通である。

多くの場合、その方程式は次の形に表現できる。

$$a(\mathbf{p})f(\mathbf{x}) = 0 \quad (2.58)$$

ここで、 $f(\mathbf{x})$ はデータ分布関数、 $\mathbf{x} = (x_1, x_2, \dots, x_m)^T$ は、データ空間の座標を表すベクトルである。 $a(\mathbf{p})$ は、振幅演算子、 $\mathbf{p} = (p_1, p_2, \dots, p_l)$ は、未知属性値を表すパラメータ空間のベクトルである。

ベクトル \mathbf{x} や \mathbf{p} は、関数空間の様に無限次元ベクトル空間であってもよい。ただし、実際のアルゴリズムでは、有限個の基底関数による展開を用いて有限次元ベクトルで記述されるのが普通である。これは、物体表面を表す関数が未知パラメータになっている場合などである。

データ分布関数の一般的性質 データ関数には、必要なだけの階数の微分と必要なだけの次数のモーメントの合成が2乗可積分で積分値が有限であるという制約を課する。

$$\int_{\mathcal{R}^m} \left| x_1^{s_1} x_2^{s_2} \cdots x_m^{s_m} \frac{\partial^{t_1+t_2+\dots+t_m}}{\partial x_1^{t_1} \partial x_2^{t_2} \cdots \partial x_m^{t_m}} f(\mathbf{x}) \right|^2 dx_m \cdots dx_2 dx_1 < \infty \quad (2.59)$$

ただし、 $s_i, t_i \geq 0$ である。この制約は、実際の画像において容易に満たしうる条件である。任意の x_j と $\frac{\partial}{\partial x_j}$ を交換した場合も含めて考える。ただし、同じ x_j 方向に関するモーメント演算子 x_j と偏微分演算子 $\frac{\partial}{\partial x_j}$ は交換可能でない。(これは、量子力学において不確定原理の数学的基礎である。)

この条件によって、基本拘束条件の2乗誤差の定義が可能である。2乗誤差が定義されると、最適化によるパラメータ推定が可能である。

振幅演算子の一般形 振幅演算子には、データ空間の座標に関するモーメント演算子と微分演算子の任意の合成からなる次のものを考えれば、多くの場合十分である。

$$a(\mathbf{p}) = \sum_{i=1}^l p_i \phi_i(x_1, x_2, \dots, x_m, \frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_m}) \quad (2.60)$$

$\phi_i(x_1, x_2, \dots, x_m, \frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_m})$ は、 x_j と $\frac{\partial}{\partial x_j}$ の多項式を表す。

この演算子には一般にはデータ関数に対する微分演算が含まれる。微分演算は、不良設定問題であるから、ノイズに対して敏感であり、正則化手法の導入が必要である。2.1.6 節で用いた Gabor 関数展開と 2.2.3 節で用いた Gauss 関数による畳み込み演算は、このための正則化手法の例である。

モーメント・微分演算子が振幅演算子になる例 例として、画像のアフィン運動モデルを考える。第 2.1 節で述べたオプティカルフローでは、局所的に運動が並進運動であると仮定しているが、これをアフィン運動としてみよう。次の置き換えをおこなう。

$$\begin{aligned} u &= a_{11}x + a_{12}y + b_1 \\ v &= a_{21}x + a_{22}y + b_2 \end{aligned} \quad (2.61)$$

$a_{11}, a_{12}, a_{21}, a_{22}, b_1, b_2$ は、アフィン速度の運動パラメータである。これを、オプティカルフローの基本拘束方程式 (2.1) に代入して、次式を得る。

$$a_{11}x \frac{\partial f}{\partial x} + a_{12}y \frac{\partial f}{\partial x} + b_1 \frac{\partial f}{\partial x} + a_{21}x \frac{\partial f}{\partial y} + a_{22}y \frac{\partial f}{\partial y} + b_2 \frac{\partial f}{\partial y} + \frac{\partial f}{\partial t} = 0 \quad (2.62)$$

次の記法を用いてまとめる。

$$\begin{aligned} \mathbf{p} &= (a_{11}, a_{12}, b_1, a_{21}, a_{22}, b_2, 1)^T \\ \Gamma &= \left(x \frac{\partial}{\partial x}, y \frac{\partial}{\partial x}, \frac{\partial}{\partial x}, x \frac{\partial}{\partial y}, y \frac{\partial}{\partial y}, \frac{\partial}{\partial y}, \frac{\partial}{\partial t} \right)^T \end{aligned} \quad (2.63)$$

すると、振幅演算子は、

$$a(\mathbf{p}) = \mathbf{p}^T \Gamma \quad (2.64)$$

と書くことができる。これは、1次微分と1次モーメントの積の線形結合からなる演算子である。

2.3.2.2 パラメータ推定の一般形

振幅演算子は線形演算子である。つまり、次式が成り立つ。

$$\begin{aligned} a(\mathbf{p})\{f_1(x) + f_2(x)\} &= a(\mathbf{p})f_1(x) + a(\mathbf{p})f_2(x), \\ a(\mathbf{p})0 &= 0 \end{aligned} \quad (2.65)$$

さらに、未知パラメータ \mathbf{p} は線形同次パラメータであると仮定する。もともとの形式が線形同次でなくても、パラメータの置き換えによって、線形同次形式に変形することができる。例えば、2.1.4.1 節において、オプティカルフローに対して、 $u = V_1/V_3, v = V_2/V_3$ という関係にある同次パラメータ (V_1, V_2, V_3) を用いて、フローを表現した。このとき、任意の定数 k に対して $a(k\mathbf{p}) = ka(\mathbf{p})$ が成立する。つまり、 $a(\mathbf{p})$ は、線形同次形式である。パラメータ \mathbf{p} のノルム $\|\mathbf{p}\|$ が与えられていると仮定する。基本拘束方程式 (2.58) から、未知パラメータ \mathbf{p} を求めるために、次の最適化問題を解く。

$$\text{Minimize } e(\mathbf{p}) = \int_{\mathcal{R}^m} |a(\mathbf{p})f(x)|^2 dx,$$

$$\text{Subject to } \|\mathbf{p}\| = c \quad (2.66)$$

この条件付き最適化問題は次の最適化問題と等価である。

$$\text{Minimize } E(\mathbf{p}) = \frac{\int_{\mathcal{R}^m} |a(\mathbf{p})f(\mathbf{x})|^2 dx}{\|\mathbf{p}\|^2 \int_{\mathcal{R}^m} |f(\mathbf{x})|^2 dx} \quad (2.67)$$

この最小化問題は、 $\|\mathbf{p}\|^2 = 1$ なる条件のもとで、分子 $\int_{\mathcal{R}^m} |a(\mathbf{p})f(\mathbf{x})|^2 dx$ を最小化することと等価である。 $\int_{\mathcal{R}^m} |f(\mathbf{x})|^2 dx$ による、パラメータの推定値そのものには直接影響しない。

2.3.3 振幅演算子を用いた重ね合わせの原理の数学的一般表現

拘束方程式が (2.58) の形に表現されさえすれば、あとは、演算子の形式的合成によって多重属性値に関する拘束方程式が自動的に得られる。

2.3.3.1 スカラー値の拘束方程式の場合

n 個のデータ分布関数 $f_i(\mathbf{x}) (i = 1, 2, \dots, n)$ が、振幅演算子 $a(\mathbf{p}_i) (\mathbf{p}_i \in \mathcal{H}_i, i = 1, 2, \dots, n)$ によって次式で拘束されていると仮定する。

$$a(\mathbf{p}_i)f_i(\mathbf{x}) = 0 \quad (2.68)$$

$f(\mathbf{x})$ は、パターン $f_i(\mathbf{x})$ の加法的重畳で表されると仮定する。

$$f(\mathbf{x}) = \sum_{i=1}^n f_i(\mathbf{x}) \quad (2.69)$$

この重畳は、次の様に重み付きの重畳に一般化できる。

$$f(\mathbf{x}) = \sum_{i=1}^n \alpha_i f_i(\mathbf{x}) \quad (2.70)$$

ここで、 α_i は、重み係数である。演算子 $a(\mathbf{p}_i)$ と $a(\mathbf{p}_j) (i \neq j)$ は、可換であると仮定する。つまり、 $a(\mathbf{p}_i)a(\mathbf{p}_j) = a(\mathbf{p}_j)a(\mathbf{p}_i)$ である。このとき、 $f(\mathbf{x})$ は、次の方程式で拘束される。

$$a(\mathbf{p}_n)a(\mathbf{p}_{n-1}) \cdots a(\mathbf{p}_1)f(\mathbf{x}) = 0 \quad (2.71)$$

合成演算子を

$$a^{(n)}(\mathbf{p}^{(n)}) = a(\mathbf{p}_1)a(\mathbf{p}_2) \cdots a(\mathbf{p}_n) \quad (2.72)$$

と定義する。すると、一個の演算子と見なすことができる。合成されたパラメータ $\mathbf{p}^{(n)}$ は、パラメータ \mathbf{p}_i の空間 $\mathcal{H}_i (i = 1, 2, \dots, n)$ のテンソル積空間の元である。

$$\mathbf{p}^{(n)} \in \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_n \quad (2.73)$$

この拘束方程式に基づいて次のエネルギー最小化問題を定式化する。

$$\text{Minimize } E^{(n)}(\mathbf{p}) = \frac{\int_{\mathcal{R}^m} |a^{(n)}(\mathbf{p}^{(n)})f(\mathbf{x})|^2 dx}{\|\mathbf{p}^{(n)}\|^2 \int_{\mathcal{R}^m} |f(\mathbf{x})|^2 dx} \quad (2.74)$$

2.3.3.2 ベクトル値の拘束方程式の場合

ベクトル値の拘束方程式の場合には、演算子の合成は、テンソル積（クロネッカー積）によって以下のように定義されなければならない。2個の2次元ベクトル値演算子

$$\begin{aligned} \mathbf{a}_1(\mathbf{p}_1) &= (a_1(\mathbf{p}_1), b_1(\mathbf{p}_1)), \\ \mathbf{a}_2(\mathbf{p}_2) &= (a_2(\mathbf{p}_2), b_2(\mathbf{p}_2)) \end{aligned} \quad (2.75)$$

の場合を以下に示す。拘束方程式は、

$$\begin{aligned} \mathbf{a}_1(\mathbf{p}_1)f(\mathbf{x}) &= 0 \\ \mathbf{a}_2(\mathbf{p}_2)f(\mathbf{x}) &= 0 \end{aligned} \quad (2.76)$$

である。このとき、合成された拘束方程式は、次の形である。

$$\begin{aligned} a_1(\mathbf{p}_1)a_2(\mathbf{p}_2)f(\mathbf{x}) &= 0, \\ b_1(\mathbf{p}_1)b_2(\mathbf{p}_2)f(\mathbf{x}) &= 0, \\ a_1(\mathbf{p}_1)b_2(\mathbf{p}_2)f(\mathbf{x}) &= 0, \\ b_1(\mathbf{p}_1)a_2(\mathbf{p}_2)f(\mathbf{x}) &= 0 \end{aligned} \quad (2.77)$$

これらをまとめて、

$$\mathbf{a}_1(\mathbf{p}_1) \otimes \mathbf{a}_2(\mathbf{p}_2)f(\mathbf{x}) = 0 \quad (2.78)$$

と書く。

2.3.3.3 演算子の対称化

一般に、スカラー値の演算子の合成演算子を群 G で対称化すると次式である。

$$\bar{a}^{(n)}(\bar{\mathbf{p}}^{(n)}) = \frac{1}{|G|} \sum_{\sigma \in G} a(\mathbf{p}_{\sigma(1)})a(\mathbf{p}_{\sigma(2)}) \cdots a(\mathbf{p}_{\sigma(n)}) \quad (2.79)$$

$|G|$ は、群 G の位数を表す。合成パラメータ $\bar{\mathbf{p}}^{(n)}$ は、テンソル積を用いた次式で定義する。

$$\bar{\mathbf{p}}^{(n)} = \frac{1}{|G|} \sum_{\sigma \in G} \mathbf{p}_{\sigma(1)} \otimes \mathbf{p}_{\sigma(2)} \otimes \cdots \otimes \mathbf{p}_{\sigma(n)} \quad (2.80)$$

n 次の置換群を P_n とし、 σ を添字の集合 $\{1, 2, \dots, n\}$ に作用する P_n の元とすると、合成演算子は次の形である。

$$\bar{a}^{(n)}(\bar{p}^{(n)}) = \frac{1}{n!} \sum_{\sigma \in P_n} a(p_{\sigma(1)}) a(p_{\sigma(2)}) \cdots a(p_{\sigma(n)}) \quad (2.81)$$

ここで、

$$\bar{p}^{(n)} = \frac{1}{n!} \sum_{\sigma \in P_n} p_{\sigma(1)} \otimes p_{\sigma(2)} \otimes \cdots \otimes p_{\sigma(n)} \quad (2.82)$$

である。これが、 n -重の場合の一般表現である。

もっとも簡単な2-重の場合、すなわち、 $G = P_2$ では、演算子は、

$$\bar{a}^{(2)}(\bar{p}^{(2)}) = \frac{1}{2} \{a(p_1)a(p_2) + a(p_2)a(p_1)\} \quad (2.83)$$

であり、合成パラメータは、

$$\bar{p}^{(2)} = \frac{1}{2} \{p_1 \otimes p_2 + p_2 \otimes p_1\} \quad (2.84)$$

となる。

2.3.3.4 データ関数と振幅演算子が共にベクトル値の場合

ステレオ透明視の計算理論では、データ関数がベクトル値で、振幅演算子が正方行列であった。この場合は、演算子の合成を行列の積で定義した。しかし、行列の積は一般に可換でないので、厳密な意味での拘束は得られない。この場合にもエネルギー最小化に基づいてアルゴリズムを導くことができる。

2.3.3.5 異なる型の演算子の合成

以上述べた場合は、すべて同じ型の演算子同士の合成であった。異なる型の演算子の合成も考えられる。ここでは、スカラー値の場合を述べる。ベクトル値の場合も同様に考えられる。

次の n 個の拘束方程式を考えよう ($i = 1, 2, \dots, n$)。

$$a_i(p_i) f_i(x) = 0 \quad (2.85)$$

式 (2.68) と異なるのは、演算子 $a_i(p_i)$ が、各 i ごとに異なることである。次式の合成演算子を定義する。

$$a_n(p_n) a_{n-1}(p_{n-1}) \cdots a_1(p_1) f(x) = 0 \quad (2.86)$$

この場合にも、演算子の可換性が要求される。可換でなくても、次の対称化の操作によって、合成演算子を定義することができる。

$$\bar{a}^{(n)}(\bar{\mathbf{p}}^{(n)}) = \frac{1}{n!} \sum_{\sigma \in P_n} a_{\sigma(1)}(\mathbf{p}_{\sigma(1)}) a_{\sigma(2)}(\mathbf{p}_{\sigma(2)}) \cdots a_{\sigma(n)}(\mathbf{p}_{\sigma(n)}) \quad (2.87)$$

異なる形の振幅演算子の合成の例 アフィン運動するパターンと並進運動するパターンの重ね合わせを例にあげる。

アフィン運動の振幅演算子を

$$a_a(\mathbf{p}_a) = \mathbf{p}_a^T \Gamma_a \quad (2.88)$$

とし、並進運動の振幅演算子を

$$a_t(\mathbf{p}_t) = \mathbf{p}_t^T \Gamma_t \quad (2.89)$$

とする。ベクトル $\mathbf{p}_a, \Gamma_a, \mathbf{p}_t, \Gamma_t$ は、次式で定義される。

$$\begin{aligned} \mathbf{p}_a &= (a_{11}, a_{12}, a_{21}, a_{22}, b_1, b_2, 1)^T, \\ \Gamma_a &= \left(x \frac{\partial}{\partial x}, y \frac{\partial}{\partial x}, x \frac{\partial}{\partial y}, y \frac{\partial}{\partial y}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial t} \right)^T \end{aligned} \quad (2.90)$$

$$\begin{aligned} \mathbf{p}_t &= (t_1, t_2, 1)^T, \\ \Gamma_t &= \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial t} \right)^T \end{aligned} \quad (2.91)$$

演算子 $a_a(\mathbf{p}_a)$ と $a_t(\mathbf{p}_t)$ の合成は、次式で定義される。

$$a_{at}(\mathbf{p}_{at}) = a_a(\mathbf{p}_a) a_t(\mathbf{p}_t) \quad (2.92)$$

次の形の振幅演算子になる。

$$a_{at}(\mathbf{p}_{at}) = \mathbf{p}_{at}^T \Gamma_{at} \quad (2.93)$$

ただし、

$$\mathbf{p}_{at} = (\mathbf{p}_a \otimes \mathbf{p}_t)_{\text{vec}} \quad (2.94)$$

$$\Gamma_{at} = (\Gamma_a \otimes \Gamma_t)_{\text{vec}} \quad (2.95)$$

である。 $(\)_{\text{vec}}$ は、テンソルの成分を縦ベクトルに並べ替えることを表す。具体的には、次のベクトルである。

$$\mathbf{p}_{at} = (a_{11}t_1, a_{11}t_2, a_{11}, a_{12}t_1, a_{12}t_2, a_{12}, a_{21}t_1, a_{21}t_2, a_{21},$$

$$a_{22}t_1, a_{22}t_2, a_{22}, b_1t_1, b_1t_2, b_1, b_2t_1, b_2t_2, b_2, t_1, t_2, 1)^T \quad (2.96)$$

$$\Gamma_{\text{at}} = \left(x \frac{\partial^2}{\partial x^2}, x \frac{\partial^2}{\partial x \partial y}, x \frac{\partial^2}{\partial x \partial t}, y \frac{\partial^2}{\partial x^2}, y \frac{\partial^2}{\partial x \partial y}, y \frac{\partial^2}{\partial x \partial t}, x \frac{\partial^2}{\partial y \partial x}, x \frac{\partial^2}{\partial y^2}, x \frac{\partial^2}{\partial y \partial t}, \right. \\ \left. y \frac{\partial^2}{\partial y \partial x}, y \frac{\partial^2}{\partial y^2}, y \frac{\partial^2}{\partial y \partial t}, \frac{\partial^2}{\partial x^2}, \frac{\partial^2}{\partial x \partial y}, \frac{\partial^2}{\partial x \partial t}, \frac{\partial^2}{\partial y \partial x}, \frac{\partial^2}{\partial y^2}, \frac{\partial^2}{\partial y \partial t}, \frac{\partial^2}{\partial t \partial x}, \frac{\partial^2}{\partial t \partial y}, \frac{\partial^2}{\partial t^2} \right)^T \quad (2.97)$$

演算子の対称化を行うともう少し複雑になるが、やはり、(2.93)の形式が得られる。

2.3.4 属性値を計算するアルゴリズムの導出手順

以上述べた様に、拘束方程式のレベルでは、複数の属性に関する拘束条件は比較的簡単に求められる。しかし、この拘束条件から属性値を実際に計算するアルゴリズムを求めるためには多少の数理的技術が必要である。その一般手順を示す。以下では、拘束条件のパラメータが同次形であると仮定する。非同次の場合にも同様にしてアルゴリズムを導ける。非同次の場合は、固有値問題の代わりにパラメータに関する正規方程式である連立一次方程式に帰着される。これらは、ちょうど、本章で述べてきた多重オプティカルフローと多重視差の場合にそれぞれ対応する。ただし、パラメータのとりかたは、問題の本質とは直接関係しない。

多重オプティカルフローで述べた2通りの方法(第2.1.4.2節)と同様にして、一般の場合にも2通りの計算方法が考えられる。

1. 要素となる各振幅演算子のパラメータに関して変分法を用いて、 n 個の相互に依存しあう固有値問題を導く。これは、パラメータがテンソル積の形になるので、多重線形空間におけるパラメータの推定問題になる。
2. テンソル積空間を一つの線形空間とみなして、未知パラメータに関して線形になるように置き換えを行う。2次元関数の最適化によって得られたテンソル積パラメータから、代数計算によってもとのパラメータを計算する。これを線形化手法と呼ぶ。同一種の2重属性推定の場合には、多重オプティカルフローで用いた分解手法の拡張を用いることができる。

2.3.4.1 テンソル積パラメータの分解手法

線形化手法を用いた場合は、テンソル積の形に撚り合わされた情報が得られる。このテンソル積の分解は、原理的には、連立代数方程式系の解法に帰着される。しかし、この連立代数方程式系は、未知数よりも成分方程式数の方が多い過剰拘束系である。そのため、データ

に誤差が存在する場合は、数学的には、解が存在しない。しかし、2つのベクトルのテンソル積の場合は、以下に示す最適解の計算が可能である。以下では、 r 次元と s 次元の単位ベクトル $\mathbf{p} = (p_1, p_2, \dots, p_r)^T$ ($\|\mathbf{p}\| = 1$)、 $\mathbf{q} = (q_1, q_2, \dots, q_s)^T$ ($\|\mathbf{q}\| = 1$) のテンソル積に関して述べる。

対称テンソル積の分解 対称テンソル積の場合は、付録 A.1.4 に次に示した公式を拡張して用いることができる。この場合は、次元が等しい2個のベクトル同士のテンソル積である ($r = s$)。この対称テンソル積は、次式で定義される対称行列 $\mathbf{B} = (b_{ij})$ で表現できる。

$$\mathbf{B} = (b_{ij}) = \frac{1}{2}(p_i q_j + q_i p_j) \quad (2.98)$$

一般のベクトル \mathbf{p} , \mathbf{q} に対して、行列 \mathbf{B} の階数は2である。固有値は、正と負が1個ずつで、他の $r-1$ 個の固有値はすべて零である。非零固有値を $\lambda_+ > 0$ 、 $\lambda_- < 0$ とする。すると、行列 \mathbf{B} の固有値展開は次式である。

$$\mathbf{B} = \lambda_+ \mathbf{e}_+^T \mathbf{e}_+ + \lambda_- \mathbf{e}_-^T \mathbf{e}_- \quad (2.99)$$

固有値 λ_+ 、 λ_- に対応する固有ベクトルをそれぞれ \mathbf{e}_+ 、 \mathbf{e}_- とする。すると、ベクトル \mathbf{p} 、 \mathbf{q} は、次式で計算される。

$$\begin{aligned} \mathbf{p} &= \sqrt{\frac{-\lambda_-}{\lambda_+ - \lambda_-}} \mathbf{e}_- + \sqrt{\frac{\lambda_+}{\lambda_+ - \lambda_-}} \mathbf{e}_+, \\ \mathbf{q} &= -\sqrt{\frac{-\lambda_-}{\lambda_+ - \lambda_-}} \mathbf{e}_- + \sqrt{\frac{\lambda_+}{\lambda_+ - \lambda_-}} \mathbf{e}_+ \end{aligned} \quad (2.100)$$

この公式は、固有値展開式 (2.99) と次の恒等式を比較することによって得られる。

$$\frac{1}{2}(p_i q_j + q_i p_j) = \frac{1}{4}(p_i + q_i)(p_j + q_j) - \frac{1}{4}(p_i - q_i)(p_j - q_j) \quad (2.101)$$

以上は誤差の存在しない場合であった。データに誤差が存在する場合にも、行列 \mathbf{B} の固有値展開において、最大固有値と最小固有値、および、それらに対応する固有ベクトルを用いて、公式 (2.100) を計算すれば、最適解が求められる。

対称でないテンソル積の分解 テンソル積が対称でない場合は、以下の様に最適分解できる。テンソル積を次の行列で表現する。

$$\mathbf{B} = (b_{ij}) = (p_i q_j) \quad (2.102)$$

この行列の階数は1である。次の恒等式が成立する。

$$\mathbf{B}\mathbf{B}^T = (p_i p_j) = \mathbf{p}\mathbf{p}^T$$

$$\mathbf{B}^T \mathbf{B} = (q_i q_j) = \mathbf{q} \mathbf{q}^T \quad (2.103)$$

これらの恒等式をそれぞれ行列 $\mathbf{B} \mathbf{B}^T$, $\mathbf{B}^T \mathbf{B}$ の固有値展開としてみる。すると、唯一の非零固有値 1 の固有ベクトルが、それぞれ、 \mathbf{p} , \mathbf{q} であることを示している。したがって、データに誤差が存在する場合は、行列 $\mathbf{B} \mathbf{B}^T$, $\mathbf{B}^T \mathbf{B}$ をそれぞれ固有値展開する。最大固有値に対応する固有ベクトルを求めれば、それらが、ベクトル \mathbf{p} , \mathbf{q} の最適解である。

2.3.5 まとめ

多重オプティカルフローと多重視差検出の計算理論として提案した演算子形式の重ね合わせの原理を拡張、一般化した。数学的には、和の型に重畳された分布関数によってデータを記述できれば、拘束条件を振幅演算子の合成操作によって統一的、自動的に導出できる。これによって、単一属性を推定するためにいままで導かれてきた種々の基本拘束方程式を、多重情報に関する拘束方程式に自動的に拡張することが可能になった。その意味で、本節で述べた重ね合わせの原理は、コンピュータビジョンにとって、いわば、「超」計算理論、あるいは、「メタ」計算理論といえる性格を持つ。もちろん、基本拘束方程式が求まるからといって、それからただちに適切なアルゴリズムが導かれるわけではない。問題によっては、解析的に厳密に解を求めようとする非現実的に計算手順が複雑になることも考えられる。データの誤差に対してロバストなアルゴリズムを導くことはさらに困難である。

第 3 章

多重 3 次元運動・構造復元と多重 3 次元表面復元の 計算理論

あらし オプティカルフローやステレオ視差などの初期視覚によって得られた情報から、3次元物体の運動パラメータ、物体構造パラメータ、表面形状情報を推定する中期視覚の問題に対して、重ね合わせの原理に基づいた拡張をおこなう。

- 2枚の画像フレーム間の特徴点对応からの物体の3次元剛体運動（3次元並進運動、3次元回転運動）の復元は、運動情報からの3次元情報処理において典型的な問題である。第3.1節ではこれを拡張し、複数の動剛体を直接扱う計算理論とアルゴリズムを導く。
- 3次元復元のアルゴリズムで得られる3次元データは、空間的にまばらで、かつ、ノイズを含む。この3次元データから視覚的に妥当な物体の表面形状を求めるためには、ある滑らかさをもった関数による近似手法が用いられる。この手法の1つは、標準正則化理論である。第3.2節では、この標準正則化理論による表面の復元理論を、複数の重なり合った表面の復元を可能にする枠組に拡張する。

3.1 多重3次元運動・構造復元の計算理論

あらまし 動画像の複数フレーム間の特徴の対応、あるいは、画像パターンの瞬時的な速度場であるオプティカルフローのいずれかが求められると、物体の3次元運動と3次元構造を復元することができる。この問題は運動からの構造復元 (structure from motion) と呼ばれる。

本節では、互いに重なりあって異なる運動をする2個の剛体の3次元運動パラメータと3次元構造パラメータを復元する計算理論とアルゴリズムを導く。物体表面のなめらかさ、あるいは、投影の近似モデルなどの付加的な前提条件を用いない。フレーム間の追跡などによってすでに得られた2画像間の特徴点对応情報から閉形式の数値計算と一撃クラスタリングで復元する。まず、2枚の透視投影像間の対応点集合が n 個の3次元剛体運動を含む場合の基本拘束方程式を、2.3 節で一般化した重ね合わせの原理を用いて導出する。次に、2重運動の場合について、2組の3次元剛体運動パラメータを、データの誤差を考慮して準最適に求める線形アルゴリズムを導出する。このアルゴリズムは、データにノイズが無い場合に、厳密解を与える。また、計算の閉形式性を保ちつつ、中間段階で得られる拘束方程式をできるだけ多く用いる最適化手法を取り入れている。そのため、誤差が存在する場合にも全体として準最適な解を求めることができる。さらに、この運動パラメータを用いて対応点の各物体へのクラスタリングと構造復元を行う一撃アルゴリズムを導く。特徴点の数が十分多く、それらが特殊な配置になれば、一般の運動パラメータに対して解が一意に定まることが構成的に証明される。最後に数値シミュレーションによって本アルゴリズムの正当性を示す。^a

^a本節の内容は、文献 [178] [189] [194] において発表された。

3.1.1 はじめに

従来の運動立体視の計算理論は、物体が単一の剛体であるという前提条件のもとで導かれた拘束方程式に基づいて展開されている。従って、視野内に複数の剛体が存在し、それらが互いに異なる 3 次元運動を行う場合には、弛緩法、線過程と結合されたマルコフ場、ハフ変換などの投票法、拘束方程式に対する適合性に基づいたデータのクラスタリングなどのいろいろなセグメンテーション手法と組み合わせたアルゴリズムが提案されてきた。

複数の物体領域に画像を分割するセグメンテーション手法は、その正常な動作範囲や、データの誤差に対する振る舞いなどの性質を正確に把握することが非常に困難であることが従来から問題点として指摘されてきた [10]。これは、緩和計算や、投票、クラスターの探索アルゴリズムの振る舞いが、数学的に厳密な解析手段に乗りにくいことがその理由である。もし、複数運動の推定問題が、純粋に数学的な手段だけで定式化できるならば、原理的に上記の問題点は生じない。数値的な計算プロセスだけでこの問題が解決される可能性が生まれる。

本節では、運動立体視において基本的代表的でかつ数学的に厳密な計算理論である、2 枚の透視投影像における特徴点の点对応情報から 3 次元運動と 3 次元構造の復元問題を対象とする。重ね合わせの原理を適用し、複数の剛体が独立な運動をする場合に理論とアルゴリズムを拡張する。

以下、まず、第 3.1.2 節において、従来の 3 D 運動立体視のアルゴリズムにおける複数運動の扱いをサーベイする。次に、第 3.1.3 節において、2 枚の透視投影像の特徴点对応から 3 次元運動および、3 次元構造を復元するための計算理論をその歴史的発展経緯とともに簡単に述べる。次に、演算子表現と重ね合わせの原理を導入して、多重 3 D 運動の基本拘束方程式を導出する。第 3.1.4 節では、2 重 3 D 運動の場合について、その線形アルゴリズムを述べる。そして、計算された運動パラメータを用いて特徴点データのクラスタリングを一撃的に行うアルゴリズムを提案する。第 3.1.5 節では、数値シミュレーションによって、実際に 2 組の 3 D 運動と 2 物体の 3 次元構造が求まることを検証する。

3.1.2 複数運動の推定に関する従来の研究

従来から、複数物体の運動を動画像から計算するための方法が試みられてきた。本節では、複数物体の複数組の 3 次元運動と 3 次元構造を復元する方法に限定してサーベイする。

1. 3 枚の画像の特徴点の直交投影像からの 3 次元運動復元 [122]。方程式数が運動パラメータの未知数よりも多いことを利用した。余剰な方程式を用いた剛体性の検査によって特徴点のクラスタリングを行った。

2. Hough 変換に代表されるパラメータ空間における投票法を用いたもの [40] [27]。オプティカルフロー場を異なる3次元運動パラメータを持つ物体領域に分割した。
3. オプティカルフロー場からの3次元復元に関して、物体の境界における不連続を表現する線過程を導入したエネルギー関数の、画像全体にわたる最適化を行ったもの [61]。画像領域の3次元運動によるセグメンテーションと3次元運動パラメータおよび3次元構造の復元を同時に行った。
4. ステレオ動画像における線特徴の間の剛体拘束を用いたもの [126]。カルマンフィルタを用いて情報統合を行うとともに、異なる3次元運動をする複数物体に線特徴をクラスタリングしている。
5. 複数枚の直交投影像から、特異値分解によって運動と構造を求める Factorization 法 [120] を拡張したもの [102]。画像行列の階数から、独立な3次元運動をする物体数に関する情報を知ることができる事実と、形状のなめらかさ拘束を用いて、クラスタリングをおこなっている。
6. 輝度変化から3次元運動を復元する直接法 [116] において、運動領域の分割と運動パラメータの推定を同時に行なったもの [68]。一つの運動が画像面で大部分を占めていることを前提とした。拘束方程式の2乗残差を用いて、領域分割をおこなった。
7. 直交投影下の動画像における特徴点对応からの3次元運動・構造復元問題において、剛体運動の基本拘束方程式を満たす特徴点の極大集合を探索によって求める方法を用いてセグメンテーションを行ったもの [142]。
8. 動画像を3次元時空間画像とみなして、剛体運動の運動モデルを用いた時空領域におけるセグメンテーションを行ったもの [110]。
9. 複数の3次元回転物体の3次元フローの直交投影像から、複数剛体の3次元回転ベクトルを Hough 変換で求めたもの [77]。

ここで列挙した例は、ごく一部に過ぎない。他の方法もこれらに近いものがほとんどである。これらいずれの方法においても、運動推定の基本となる拘束方程式はあくまでも単一運動のものである。クラスタリング、セグメンテーションとの協調的な計算過程によって、各物体への領域分割と運動推定を同時に実行している。アルゴリズムでは、透視投影を想定していなかったり、物体表面や運動のなめらかさを仮定している。つまり、物体の画像面上における表れ方に強い制約がおかれている。これらは、そもそも、セグメンテーションの問題

自体が数学的に明確に定義しにくい不良設定問題であることに起因する。この事実は、多くのコンピュータビジョンの研究者によって指摘されている。[10]。

3.1.3 多重3D運動の基本拘束方程式の導出

3.1.3.1 線形アルゴリズムの発展と現状

点对応からの3次元運動・構造復元に対する数学的に厳密なアルゴリズムを最初に導出したのは Longuet-Higgins である [113]。彼は、ベクトル代数を用いて8点からの線形解析解を導出した(8点アルゴリズム)。それとは独立に、Tsai と Huang は特異値分解を使った解を示した [121]。Zhuang et.al. は、これらの結果を統合し、8点アルゴリズムを後に述べる「基本行列」をもとにした形式に統一した [127]。Weng et.al. は、8点以上の対応点が与えられた時に、最小2乗法によって最適解を求める方法を提案した [124]。金谷は、特徴点座標を正規化された同次座標表現である N ベクトルを用いて表現し、これらの解に対して統一した形式で記述している [22] [12]。

本節では、単一運動の場合に金谷の定式化を準用する。さらに、これを振幅演算子表現に発展させ、重ね合わせの原理を適用して、多重運動の場合の計算理論と計算アルゴリズムを展開する。ただし、金谷は、カメラが運動し、物体は静止状態と仮定しているが、本理論では、カメラが静止し、複数物体が運動すると仮定する。しかし、本来、運動は相対的である。例えば、背景に対してカメラが運動し、背景中の動物体の運動とカメラ運動を推定する問題にも適用可能である。

3.1.3.2 特徴点の2透視投影像からの3次元運動および構造復元の計算理論

以下において、単一運動の場合の定式化と、その線形アルゴリズムを本論文の展開に必要な範囲で簡潔に結果のみを述べる。アルゴリズムの導出に関する詳細は、例えば、金谷 [22] [12], Weng et.al. [124] にある。

図 3.1 は、単一物体の剛体運動と点对応が与えられた時の投影の幾何学を示す。世界座標の原点をカメラのレンズ中心にとる。物体が剛体運動をした前後の2枚の透視投影像間で物体上の N 個の特徴点 $P_\alpha (\alpha = 1, 2, \dots, N)$ と $P'_\alpha (\alpha = 1, 2, \dots, N)$ の対応が与えられているとする。1枚目の画像(第1画像)を基準とし、その特徴点の画像面上の座標を (x_α, y_α) とする。これらに対応する2枚目の画像(第2画像)における座標を (x'_α, y'_α) とする。まず、 $\mathbf{m}_\alpha = (m_{\alpha(1)}, m_{\alpha(2)}, m_{\alpha(3)}) = \frac{1}{\sqrt{x_\alpha^2 + y_\alpha^2 + f^2}}(x_\alpha, y_\alpha, f)^T$, $\mathbf{m}'_\alpha = (m'_{\alpha(1)}, m'_{\alpha(2)}, m'_{\alpha(3)}) = \frac{1}{\sqrt{(x'_\alpha)^2 + (y'_\alpha)^2 + f^2}}(x'_\alpha, y'_\alpha, f)^T$, で定義される、大きさを1に正規化された同次3次元ベクトルで対応点データを表現する。これは、原点を中心とする単位球面に特徴点を投影することに

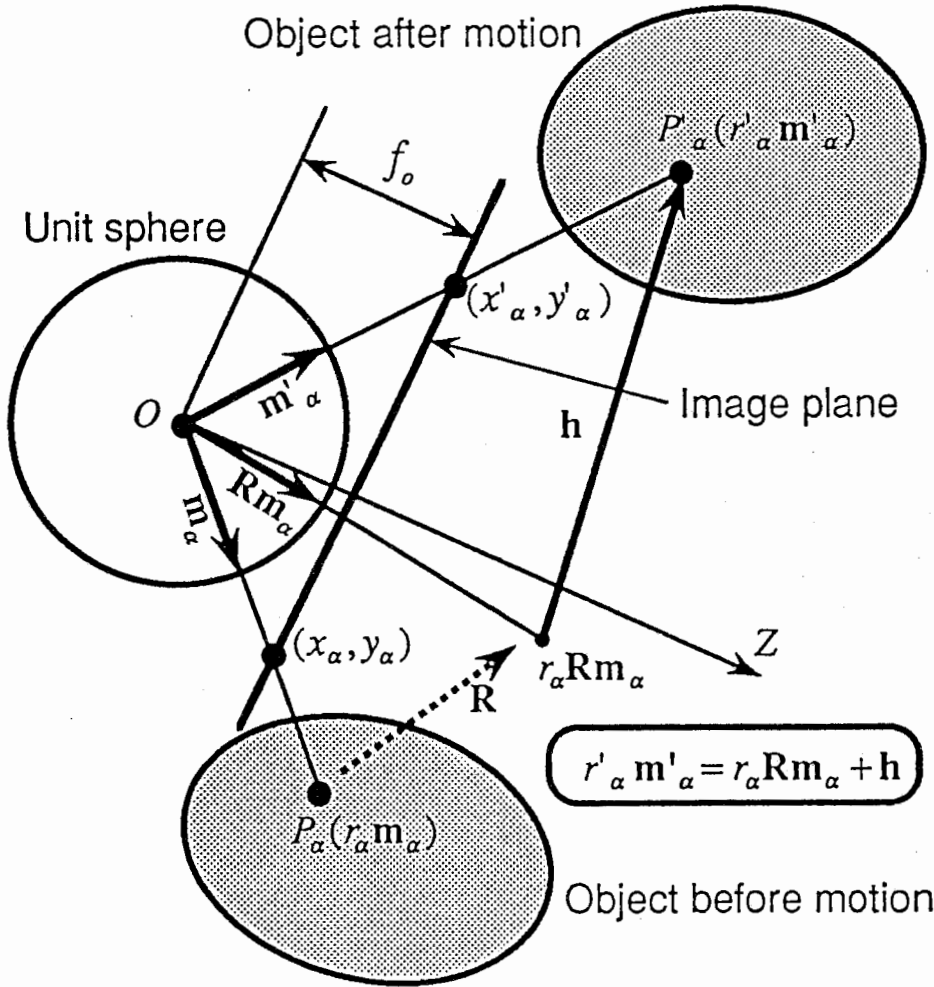


Figure 3.1: 3次元剛体運動と特徴点对応の投影像の幾何学 (単一剛体運動の場合)

相当する。 $f(>0)$ はカメラ系の焦点距離 (ピンホールの穴から投影面までの距離) である。同次ベクトル m_α と m'_α による特徴点の表現には符号に関して不定性が存在する。以下ではこの不定性を同一視して考える。

カメラのレンズ中心 O から点 P_α, P'_α までの距離をそれぞれ、 $r_\alpha(>0), r'_\alpha(>0)$ とする。すると、座標系 $\{O; (X, Y, Z)\}$ に対する P_α, P'_α の座標は、それぞれ、 $r_\alpha m_\alpha, r'_\alpha m'_\alpha$ である。

このとき、3次元空間における剛体運動の方程式は次式である。

$$r'_\alpha m'_\alpha = r_\alpha R m_\alpha + h \quad (3.1)$$

$R = (R_{ij})$ は3次元回転行列 (従って、 $RR^T = I, \det R = 1$)、 $h = (h_1, h_2, h_3)^T$ は、並進ベクトルである。原理的に、並進ベクトル h と構造パラメータ r_α および r'_α にはスケール倍の不定性がある。以下では、 $\|h\| = \sqrt{h_1^2 + h_2^2 + h_3^2} = 1$ と仮定する。 $h = 0$ の場合は、

3次元構造は原理的に得られない。以下、 $h \neq 0$ 。回転行列 \mathbf{R} は、回転軸を表す単位軸ベクトル $l = (l_1, l_2, l_3)$ と回転角 Ω (単位軸ベクトルの向きに対して右ねじの方向を正にとる) を用いて次の形に書くことができる (ロドリゲスの公式)。

$$\mathbf{R} = \begin{bmatrix} \cos \Omega + l_1^2(1 - \cos \Omega) & l_1 l_2(1 - \cos \Omega) - l_3 \sin \Omega & l_1 l_3(1 - \cos \Omega) + l_3 \sin \Omega \\ l_2 l_1(1 - \cos \Omega) + l_3 \sin \Omega & \cos \Omega + l_2^2(1 - \cos \Omega) & l_2 l_3(1 - \cos \Omega) - l_1 \sin \Omega \\ l_3 l_1(1 - \cos \Omega) - l_2 \sin \Omega & l_3 l_2(1 - \cos \Omega) + l_1 \sin \Omega & \cos \Omega + l_3^2(1 - \cos \Omega) \end{bmatrix} \quad (3.2)$$

剛体運動は、まず、回転運動 \mathbf{R} が行われ、次に、並進運動 h が行われると考える。方程式 (3.1) は、3本の3次元非零ベクトル $m'_\alpha, \mathbf{R}m_\alpha, h$ が、線形従属の関係にあることを表現しており、逆に、これら3本のベクトルが線形従属であれば適当な r_α, r'_α (ただし、正とは限らない) が存在して式 (3.1) の形に書くことができる (線形従属の定義)。これら3本のベクトルが線形従属になる必要十分条件は、それらのスカラー三重積が零になることである。

$$|m'_\alpha \quad h \quad \mathbf{R}m_\alpha| = (m'_\alpha, (h \times \mathbf{R})m_\alpha) = 0 \quad (3.3)$$

$h \times \mathbf{R}$ は、 h と \mathbf{R} の3個の列ベクトル r_1, r_2, r_3 ($\mathbf{R} = (r_1, r_2, r_3)$) との外積 (ベクトル積) から作られる3次行列 $[h \times r_1, h \times r_2, h \times r_3]$ をあらわす。 $\mathbf{G} = h \times \mathbf{R}$ は基本行列 (essential matrix) と呼ばれ、物体の運動情報をすべて含む。 \mathbf{G} は反対称行列 \mathbf{H} と回転行列 \mathbf{R} の積として

$$\mathbf{G} = \mathbf{H}\mathbf{R} \quad (3.4)$$

と書くことができる。ここで、

$$\mathbf{H} = \begin{bmatrix} 0 & -h_3 & h_2 \\ h_3 & 0 & -h_1 \\ -h_2 & h_1 & 0 \end{bmatrix} \quad (3.5)$$

である。この記法を用いると、特徴点データ m_α, m'_α に関する拘束である次式を得る。

$$(m'_\alpha, \mathbf{G}m_\alpha) = \sum_{i,j=1}^3 m'_{\alpha(i)} m_{\alpha(j)} G_{ij} = 0 \quad (3.6)$$

これが、点对応からの3D運動復元の基本拘束方程式である (運動立体視のエピポーラ拘束とも呼ばれる)。 \mathbf{G} には定数倍の不定性が残る。しかし、 $\|h\| = 1, \mathbf{R}\mathbf{R}^T = \mathbf{I}$ より $\|\mathbf{G}\| = \sqrt{\sum_{i=1}^3 \sum_{j=1}^3 G_{ij}^2} = \sqrt{2}$ となる。したがって、 \mathbf{G} は、大きさ $\sqrt{2}$ のベクトルとみなせる。ただし、 $\|\mathbf{G}\| = \sqrt{2}$ を満たす任意の行列 \mathbf{G} が式 (3.4) の形に分解できるわけではない。付加条件が必要である。この条件は次式である [109]。

$$\det(\mathbf{G}) = 0, \quad \|\mathbf{G}\mathbf{G}^T\| = \sqrt{2} \quad (3.7)$$

G の特異値 (すなわち、 GG^T の固有値) が $1, 1, 0$ であるという条件もこれと同値である。分解のための手法の一例を付録 B.1.4 に示す。拘束方程式 (3.6) からは G と $-G$ の2通りの解が得られる。さらに、 G を分解するとき、次式の多義性が存在する (twisted-pair と呼ばれる [115] [117])。

$$G = h \times R = (-h) \times (I_h R) \quad (3.8)$$

$I_h \equiv 2hh^T - I$ は、 h を方向ベクトルとする軸の周りの π だけの回転を表す。従って、拘束方程式 (3.6) からだけでは剛体運動パラメータに対して4重の多義性が残ることになる。この多義性は、剛体運動の方程式 (3.1) において、正值条件 $r_\alpha > 0, r'_\alpha > 0$ を用いると一義に絞れる。運動パラメータが $\{h, R\}$ であるとき、物体構造パラメータ r_α と r'_α は、次の公式から計算できる。

$$\begin{aligned} r_\alpha(h, R) &= \frac{-(h, m_\alpha) + (m'_\alpha, Rm_\alpha)(h, Rm'_\alpha)}{1 - (m'_\alpha, Rm_\alpha)^2}, \\ r'_\alpha(h, R) &= \frac{(h, m'_\alpha) - (m'_\alpha, Rm_\alpha)(h, Rm_\alpha)}{1 - (m'_\alpha, Rm_\alpha)^2} \end{aligned} \quad (3.9)$$

3.1.3.3 基本拘束方程式の演算子形式表現

本節では、データ分布関数と振幅演算子によって拘束方程式 (3.6) を表現する。まず、データ分布関数を定義する。この場合データは、ベクトル m_α, m'_α の積空間における分布関数として表現できる。

$$f(\xi', \xi) = \sum_{\alpha=1}^N \left\{ \prod_{i=1}^3 \mu(m'_{\alpha(i)} - \xi'_i) \prod_{j=1}^3 \mu(m_{\alpha(j)} - \xi_j) \right\} \quad (3.10)$$

$\xi = (\xi_1, \xi_2, \xi_3), \xi' = (\xi'_1, \xi'_2, \xi'_3)$ であり、 $\mu()$ は、通常のデルタ関数 $\delta()$ の平方根を表す。平方根をとるのは2乗エネルギーが意味を持つようにするためである。 $\mu()$ は次の性質を持つ関数として定義される。

$$\int_{-\infty}^{+\infty} \{\mu(\xi)\}^2 d\xi = 1 \quad (3.11)$$

$$\int_{-\infty}^{+\infty} \{\mu(x - \xi)f(\xi)\}^2 d\xi = \{f(x)\}^2 \quad (3.12)$$

より具体的には、 $\mu()$ は、平均0の正規分布関数の平方根に対して、分散 $\sigma \rightarrow 0$ なる極限をとったものと解釈できる：

$$\mu(x) = \lim_{\sigma \rightarrow 0} \frac{1}{(2\pi)^{\frac{1}{4}} \sqrt{\sigma}} \exp\left(-\frac{1}{4\sigma^2} x^2\right) \quad (3.13)$$

実際のデータに対しては、データの統計的信頼性に応じた重み付けが必要である。また、画像と特徴検出器の空間解像度は有限であるため、画像面上の位置の不確実性が存在する。式

(3.10) において、各 α に対するデルタ関数をそれぞれ異なる一般の振幅分布に一般化すれば特徴点の信頼性と不確定性の両方を直接表現可能である。本論文では、計算原理の確認という目的から、信頼性は一様で、空間解像度は無限であると仮定する。

こうして定義されたデータ関数に対して次の振幅演算子を定義する。

$$a(\mathbf{G}) = G_{ij}\xi'_i\xi_j \quad (3.14)$$

この演算子は、基本行列 \mathbf{G} によってパラメータ付けされた線形演算子である。 \mathbf{G} が定数倍を除いてしか定まらないことを反映して、この演算子 $a(\mathbf{G})$ は、 \mathbf{G} に関して同次である。

以上の定義を用いると基本拘束方程式は次式に書かれる。

$$a(\mathbf{G})f(\xi', \xi) = 0 \quad (3.15)$$

この方程式はかなり抽象的な表現である。以下に示す様に、これは、基本拘束方程式 (3.6) に等価である。

まず、エネルギー積分を次式で定義する。

$$E(\mathbf{G}) = \frac{\iint_{\mathcal{R}^3 \times \mathcal{R}^3} |a(\mathbf{G})f(\xi', \xi)|^2 d\xi' d\xi}{\|\mathbf{G}\|^2 \iint_{\mathcal{R}^3 \times \mathcal{R}^3} |f(\xi', \xi)|^2 d\xi' d\xi} \quad (3.16)$$

定義 (3.10) と関数 $\mu(\cdot)$ の性質 (3.11), (3.12) を用いると次式が導かれる。

$$E(\mathbf{G}) = \frac{\sum_{\alpha=1}^N \left(\sum_{i,j=1}^3 G_{ij} m'_{\alpha(i)} m_{\alpha(j)} \right)^2}{\|\mathbf{G}\|^2 N} = \frac{\sum_{\alpha=1}^N (\mathbf{m}'_{\alpha}, \mathbf{G} \mathbf{m}_{\alpha})^2}{\|\mathbf{G}\|^2 N} \quad (3.17)$$

これは、基本拘束方程式 (3.6) の2乗誤差の総和を \mathbf{G} のノルムと対応点の個数 N で割ったものの $1/2$ である。 $E(\mathbf{G})$ の最小化問題において誤差0の時に、式 (3.17) と式 (3.6) は等価である。

3.1.3.4 重ね合わせの原理に基づいた多重3D運動の基本拘束方程式

重ね合わせの原理に基づき、多重運動の基本拘束方程式を導出する。まず、2重運動の場合について基本拘束方程式を導く。 n -重運動の場合は、結果のみを示す。

2個の剛体に A, B とラベルをつける。まず、データ分布関数に関して考察する。剛体 A の N^A 個の対応点を $\{(m_{\beta}^A, m_{\beta}^A) | \beta = 1, 2, \dots, N^A\}$ とし、剛体 B の N^B 個の対応点を $\{(m_{\gamma}^B, m_{\gamma}^B) | \gamma = 1, 2, \dots, N^B\}$ とする。剛体 A, B に対応するデータ関数をそれぞれ $f^A(\xi', \xi)$ および $f^B(\xi', \xi)$ とすると、定義 (3.10) より、

$$f^A(\xi', \xi) = \sum_{\beta=1}^{N^A} \left\{ \prod_{i=1}^3 \mu(m_{\beta(i)}^A - \xi'_i) \prod_{j=1}^3 \mu(m_{\beta(j)}^A - \xi_j) \right\},$$

$$f^B(\xi', \xi) = \sum_{\gamma=1}^{N^B} \left\{ \prod_{i=1}^3 \mu(m_{\gamma(i)}^{B'} - \xi'_i) \prod_{j=1}^3 \mu(m_{\gamma(j)}^B - \xi_j) \right\} \quad (3.18)$$

である。各データ関数は、次式で拘束されると仮定する。

$$a(\mathbf{G}^A) f^A(\xi', \xi) = 0, \quad a(\mathbf{G}^B) f^B(\xi', \xi) = 0 \quad (3.19)$$

両方の物体の対応点集合を合成した対応点集合を $\{(m_\alpha, m'_\alpha) | \alpha = 1, 2, \dots, N, N = N^A + N^B\}$ とする。このとき、データ関数に関して次の加法的重ね合わせが成立する。

$$f(\xi', \xi) = f^A(\xi', \xi) + f^B(\xi', \xi) \quad (3.20)$$

従って、重ね合わせの原理から、 $f(\xi', \xi)$ に関する拘束方程式は次式となる

$$a(\mathbf{G}^A) a(\mathbf{G}^B) f(\xi', \xi) = 0 \quad (3.21)$$

\mathbf{G}^A および、 \mathbf{G}^B は、それぞれ物体 A および、物体 B の運動パラメータから定まる基本行列を表す。 n 重運動の場合も n 個の振幅演算子の積を用いて、

$$\overbrace{a(\mathbf{G}^A) a(\mathbf{G}^B) a(\mathbf{G}^C) \cdots}^n f(\xi', \xi) = 0 \quad (3.22)$$

という基本拘束である。明らかに左辺の演算子の積は可換である。したがって、誤差が無ければこの方程式は厳密に成立する。合成された振幅演算子を

$$a^{(2)}(\mathbf{G}^A, \mathbf{G}^B) = a(\mathbf{G}^A) a(\mathbf{G}^B) = G_{ij}^A \xi'_i \xi_j G_{kl}^B \xi'_k \xi_l = G_{ij}^A G_{kl}^B \xi'_i \xi'_j \xi'_k \xi'_l \quad (3.23)$$

とする。エネルギー積分は次式である。

$$E^{(2)}(\mathbf{G}^A, \mathbf{G}^B) = \frac{\iint_{\mathcal{R}^3 \times \mathcal{R}^3} |a^{(2)}(\mathbf{G}^A, \mathbf{G}^B) f(\xi', \xi)|^2 d\xi' d\xi}{\|\mathbf{G}^A\|^2 \|\mathbf{G}^B\|^2 \iint_{\mathcal{R}^3 \times \mathcal{R}^3} |f(\xi', \xi)|^2 d\xi' d\xi} \quad (3.24)$$

定義 (3.10), (3.11), (3.12) を用いると次式に等しいことが示される。

$$\begin{aligned} E^{(2)}(\mathbf{G}^A, \mathbf{G}^B) &= \frac{\sum_{\alpha=1}^N \left(\sum_{i,j,k,l=1}^3 G_{ij}^A m'_{\alpha(i)} m_{\alpha(j)} G_{kl}^B m'_{\alpha(k)} m_{\alpha(l)} \right)^2}{\|\mathbf{G}^A\|^2 \|\mathbf{G}^B\|^2 N} \\ &= \frac{\sum_{\alpha=1}^N \{(m'_\alpha, \mathbf{G}^A m_\alpha)(m'_\alpha, \mathbf{G}^B m_\alpha)\}^2}{\|\mathbf{G}^A\|^2 \|\mathbf{G}^B\|^2 N} \end{aligned} \quad (3.25)$$

このエネルギー積分は、各剛体 A および B に関する運動の拘束条件 $(m'_\alpha, \mathbf{G}^A m_\alpha) = 0$ と $(m'_\alpha, \mathbf{G}^B m_\alpha) = 0$ の辺々の積をとった方程式

$$(m'_\alpha, \mathbf{G}^A m_\alpha)(m'_\alpha, \mathbf{G}^B m_\alpha) = \sum_{i,j,k,l=1}^3 m'_{\alpha(i)} m_{\alpha(j)} G_{ij}^A m'_{\alpha(k)} m_{\alpha(l)} G_{kl}^B = 0 \quad (3.26)$$

の左辺の2乗誤差の和を対応点の数 N で割ったものの $1/4$ である。この拘束方程式は、各対応点对 $\{m_\alpha, m'_\alpha\}$ が $(m'_\alpha, G^A m_\alpha) = 0$ と $(m'_\alpha, G^B m_\alpha) = 0$ のうちの少なくともいずれか一方を満たす必要十分条件を表している。

2重運動の振幅演算子 (3.23) について考察する。この演算子は、未知パラメータ G^A 、 G^B に関して線形でない。直接これらのパラメータを推定するのは困難である。そこで、係数 $G^A_{ij} G^B_{kl}$ を、テンソル $G^{AB} = G^{AB}_{ijkl}$ で置き換える。

$$a^{(2)}(G^{AB}) = G^{AB}_{ijkl} \xi'_i \xi'_j \xi'_k \xi'_l \quad (3.27)$$

テンソル G^{AB} は、添字 i, j, k, l のそれぞれが、 $1, 2, 3$ をとる。合計 $3^4 = 81$ 個の成分を持つ。演算子にもともと存在する定数倍の不定性を除いて考えると、80次元の未知パラメータの推定問題である。各成分に関する方程式の従属性に考慮しなければならない。この従属性は、演算子が持つ添え字に関する対称性から生じる。演算子の変数部分 $\xi'_i \xi'_j \xi'_k \xi'_l$ に注目すると、添字 i と添字 k の交換を行なっても式は不変である。同様に、添字 j と l の交換を行なっても式は不変である。つまり、これらの添字の交換によって生じる変化はデータからは原理的に区別できない。従って、表現に不定性があるため、このままでテンソル G^{AB}_{ijkl} を推定することは困難である。そこで、これら交換可能な添字に関しての対称化 [165] によって、成分間の線形従属性を消去した演算子を用いる。対称化振幅演算子 $\hat{a}^{(2)}(\hat{G}^{AB})$ を次式で定義する。

$$\begin{aligned} \hat{a}^{(2)}(\hat{G}^{AB}) &= \hat{G}^{AB}_{(ik)(jl)} \xi'_i \xi'_j \xi'_k \xi'_l \\ &= \frac{1}{4} \left(G^{AB}_{ijkl} \xi'_i \xi'_j \xi'_k \xi'_l + G^{AB}_{kji l} \xi'_i \xi'_j \xi'_k \xi'_l + G^{AB}_{ilkj} \xi'_i \xi'_j \xi'_k \xi'_l + G^{AB}_{klij} \xi'_i \xi'_j \xi'_k \xi'_l \right) \end{aligned} \quad (3.28)$$

添字における丸括弧は、テンソル解析において対応する添字の間の対称性を表す [165]。テンソル \hat{G}^{AB} と基本行列 G^A 、 G^B の間には、次のテンソル方程式が成立する。

$$\hat{G}^{AB}_{(ik)(jl)} = \frac{1}{4} (G^A_{ij} G^B_{kl} + G^A_{kj} G^B_{il} + G^A_{il} G^B_{kj} + G^A_{kl} G^B_{ij}) \quad (3.29)$$

テンソル \hat{G}^{AB} (以下基本テンソルと呼ぶ) は、 $({}_3H_2)^2 = 36$ 個の独立な成分を持つ。正規化条件を考慮すると、35次元の未知パラメータをもつ。任意のテンソル \hat{G}^{AB} が式 (3.29) を満たす本来のパラメータである G^A 、 G^B に分解できるわけではない。付加条件が必要である。分解可能であるための必要十分条件は次式である。

$$\hat{G}^{AB}_{(ik)(jl)} \hat{G}^{AB}_{(i'k')(j'l')} = \hat{G}^{AB}_{(i'k)} \hat{G}^{AB}_{(j'l')} \quad (3.30)$$

これは、多重射影空間の代数多様体による特徴づけ [171] を、基本行列の2重テンソル積空間に適用したものである。方程式 (3.30) は、36次元空間における2次方程式群の拘束で

ある。添字 $i, j, k, l, i', j', k', l'$ のすべての値に対して、拘束方程式 (3.30) が意味があるわけではない。例えば、 $(i, j) = (i', j')$ または、 $(k, l) = (k', l')$ の場合は、恒等式である。このときは、拘束としての意味はない。方程式の対称性から、互いに同値な方程式も存在する。これらを考慮すると、テンソル方程式 (3.30) における独立な成分方程式の数は 19 である。この方程式によって、行列 G^A と G^B に分解可能な基本テンソル \hat{G}^{AB} は完全に規定される。

3.1.4 2重3D運動の準最適な線形アルゴリズム

方程式群 (3.30) の条件を満たすような最適解を直接求めることができれば、原理的には、各剛体に対する最小 8 点ずつの対応から基本テンソル \hat{G}^{AB} が決定できる。しかし、これは、その非線形性から極めて困難である。そこで、拘束式 (3.30) を無視し、基本テンソル \hat{G}^{AB} は、その 36 個の成分を $\|\hat{G}^{AB}\| = \sqrt{\sum_{i,j,k,l=1}^3 \left(\hat{G}_{\binom{ik}{jl}}^{AB} \right)^2} = 2$ と規格化された 36 次元の任意のベクトル値を取りうるものとして推定する。そして、得られた基本テンソルを出来る限り最適に 2 個の基本行列 G^A と G^B のテンソル積に分解する。この場合、対応点は最低 35 組必要である。このうち、最低 8 組ずつが物体 A と B に存在しなければならない。つまり、 $N^A \geq 8, N^B \geq 8$ および $N = N^A + N^B \geq 35$ が必要である。このとき、パラメータの推定は、エネルギー積分 (3.25) において、基本行列のテンソル積を基本テンソル \hat{G}^{AB} によって置き換えた次のエネルギー関数の最小化問題に帰着される。

$$E^{(2)}(\hat{G}^{AB}) = \frac{\sum_{\alpha=1}^N \left(\sum_{i,j,k,l=1}^3 \hat{G}_{\binom{ik}{jl}}^{AB} m'_{\alpha(i)} m_{\alpha(j)} m'_{\alpha(k)} m_{\alpha(l)} \right)^2}{\|\hat{G}^{AB}\|^2 N} \quad (3.31)$$

3.1.4.1 アルゴリズムの概略

この最小化問題にもとづいた線形アルゴリズムの全体の流れを次に示す。

1. 基本テンソル \hat{G}^{AB} の成分である 36 個の同次線形未知パラメータに関する過剰拘束の同次線形方程式系を解く。
 - (a) 対応点对集合から 4 次の相関テンソルを計算。
 - (b) 4 次の相関テンソルから 36 次元固有値方程式を構成。その数値解の計算により基本テンソルを推定。
2. 推定された基本テンソル \hat{G}^{AB} を 2 個の基本行列 G^A と G^B に分解し、さらに、それらを回転行列、並進ベクトルに分解する。

- (a) 基本テンソルから複数組の部分テンソルを取りだし、テンソル積に最適分解。異なる部分テンソルの分解結果の整合条件から 2 個の基本行列を推定。
- (b) 各基本行列を、回転行列と並進ベクトルに最適分解。
3. 物体構造パラメータ r'_α, r_α の計算、対応点对のクラスタリング、運動パラメータの多義性除去を行う。
- (a) 剛体運動の方程式を用いて対応点对を各動物体にクラスタリング。
- (b) 運動視差式を用いて物体構造パラメータを計算。
- (c) 物体構造パラメータの正值条件を用いて各物体の運動パラメータを一意に決定。

以下、このアルゴリズムの詳細を述べる。

3.1.4.2 第 1 段階：基本テンソルの推定

第 1 段階では、2 重運動の基本テンソル \hat{G}^{AB} の 36 成分を決定する。なお、導出は付録 B.1.1 を参照。

相関テンソルの計算 対応点对データ $\{(m_\alpha, m'_\alpha) | \alpha = 1, 2, \dots, N\}$ に対して、4 次の相関テンソルを計算する。

$$C_{ijklj'k'l'} = \frac{1}{N} \sum_{\alpha=1}^N m'_{\alpha(i)} m_{\alpha(j)} m'_{\alpha(k)} m_{\alpha(l)} m'_{\alpha(i')} m_{\alpha(j')} m'_{\alpha(k')} m_{\alpha(l')} \quad (3.32)$$

各添字 $i, j, k, l, i', j', k', l'$ は、それぞれが 1, 2, 3 という値をとる。このすべての組合せを計算すると $3^8 = 6561$ 個という膨大な計算をしなければならない。しかし、このテンソルの添字 $i, j, k, l, i', j', k', l'$ には、 i, k, i', k' の間の交換と j, l, j', l' の間の任意の交換に関する対称性がある。対称テンソルの記法を用いると $C_{\substack{(ik i'k') \\ (jl j'l')}}$ と書かれる。この対称性を考慮すると、実際に計算しなければならない独立な成分数は $({}_3H_4)^2 = 225$ である。

固有値方程式の係数行列の作成 固有値方程式は 36 次元一般化固有値問題である。

$$Mg = \lambda Dg \quad (3.33)$$

g は \hat{G}^{AB} の成分からなる 36 次元ベクトルである。M は、 36×36 の正定値対称行列で、その成分は相関テンソル $C_{\substack{(ik i'k') \\ (jl j'l')}}$ の成分の定数倍である。D は、 36×36 の正定値対称行列であり、その成分は、 \hat{G}^{AB} のノルムの定義によって決まる。以下では、基本テンソルのノルムを成分に関するユークリッドノルムと仮定した場合のノルム行列を用いた。λ は固有

値であり、最小固有値に対応する固有ベクトルが最適化問題 (3.31) の最適推定値である。まず、ベクトル g を \hat{G}^{AB} の成分を用いて定義する。 g は、次に定義する4個の9次元列ベクトル g_1, g_2, g_3, g_4 を $g = (g_1^T, g_2^T, g_3^T, g_4^T)^T$ と結合したものである。

$$\begin{aligned}
 g_1 &= \left(\hat{G}_{(ij)}^{AB} \right) = \left(G_{ij}^A G_{ij}^B \right) \quad (i = 1, 2, 3, j = 1, 2, 3) \\
 g_2 &= \left(\hat{G}_{(ij)}^{AB} \right) = \left(\frac{G_{ij}^A G_{il}^B + G_{il}^A G_{ij}^B}{2} \right) \quad (i = 1, 2, 3, (j, l) = (1, 2), (2, 3), (3, 1)) \\
 g_3 &= \left(\hat{G}_{(ik)}^{AB} \right) = \left(\frac{G_{ij}^A G_{kj}^B + G_{kj}^A G_{ij}^B}{2} \right) \quad (j = 1, 2, 3, (i, k) = (1, 2), (2, 3), (3, 1)) \\
 g_4 &= \left(\hat{G}_{(ik)}^{AB} \right) \\
 &= \left(\frac{G_{ij}^A G_{kl}^B + G_{kj}^A G_{il}^B + G_{il}^A G_{kj}^B + G_{kl}^A G_{ij}^B}{4} \right) \quad ((i, k), (j, l) = (1, 2), (2, 3), (3, 1))
 \end{aligned} \tag{3.34}$$

行列 M は、次の形に分割して定義する。

$$M = \begin{bmatrix} M_{11} & 2M_{12} & 2M_{13} & 4M_{14} \\ 2M_{21} & 4M_{22} & 4M_{23} & 8M_{24} \\ 2M_{31} & 4M_{32} & 4M_{33} & 8M_{34} \\ 4M_{41} & 8M_{42} & 8M_{43} & 16M_{44} \end{bmatrix} \tag{3.35}$$

各部分行列 M_{pq} は 9×9 で $M_{pq} = M_{qp}^T$ が成り立つ。これらの部分行列は相関テンソル $C_{(ikil'k'l')}$ の成分からなる。例えば、

$$M_{12} = \left(C_{(iii'j'j'l')} \right) \tag{3.36}$$

である。ここで、 $i, j = 1, 2, 3$ は、行列の行をインデックス付けし、 $i' = 1, 2, 3, (j', l') = (1, 2), (2, 3), (3, 1)$ は、列をインデックス付けする。成分の順序は、それぞれ対応する g_1, g_2 の成分の順序に対応させる。

定数行列 D は、次式で定義する。

$$D = \text{diag}[D_1, D_2, D_3, D_4] \tag{3.37}$$

ここで、 D_1, D_2, D_3, D_4 は 9×9 行列で、 I_9 を9次の単位行列としたとき、 $D_1 = I_9, D_2 = D_3 = 2I_9, D_4 = 4I_9$ と定義される。

固有値方程式の数値解の計算 固有値方程式 (3.33) を数値計算で解く。 $W = \sqrt{D} = \text{diag}[I, \sqrt{2}I, \sqrt{2}I, 2I]$ とし (D が一般の正定値対称行列の場合には、 $D = W^T W$ を満た

す上三角行列 W をとる)、 $g' = Wg$, $M' = (W^T)^{-1}MW^{-1}$ とおくと、固有値方程式は次の標準的な形に変換される。

$$M'g' = \lambda g' \quad (3.38)$$

以下では、ベクトル g_i の推定値を $\bar{g}_i (i = 1, \dots, 4)$ とし、それに対応する基本テンソルを $\bar{G}^{AB} = \bar{G}_{(ik)(jl)}^{AB}$ と記すことにする。推定された基本テンソルは、 $\|\bar{G}^{AB}\| = \sqrt{\sum_{i,j,k,l=1}^3 \left(\bar{G}_{(ik)(jl)}^{AB}\right)^2} = 2$ を満たすように規格化する。

3.1.4.3 第2段階：基本テンソルの分解

第2段階では、テンソル方程式 (3.29) を解いて、基本テンソルの推定値 \bar{G}^{AB} から基本行列 G^A と G^B を求める。この方程式は、36個の成分方程式をもつ同次連立代数方程式系である。基本テンソルの規格化条件を考慮するとパラメータ \hat{G}^{AB} の自由度は $36 - 1 = 35$ 次元である。一方、2個の基本行列の未知数は $9 + 9 = 18$ 個である。ただし、各基本行列の規格化条件を考慮すると自由度は $18 - 2 = 16$ 次元である。したがって、連立代数方程式系 (3.29) は、 $35 - 16 = 19$ 次元分が過剰な方程式系である。したがって、データにノイズが含まれる場合は解が存在しない。したがって、最適化手法をもちいて、これらの方程式をできるだけよく満たす解を求めることが必要である。なお、この余分な19次元の自由度は、分解可能条件 (3.30) を無視した結果である。式 (3.29) の2乗誤差を次式で定義する。

$$e(G^A, G^B) = \sum_{i,j,k,l=1}^3 \left[\bar{G}_{(ik)(jl)}^{AB} - \frac{1}{4} (G_{ij}^A G_{kl}^B + G_{kj}^A G_{il}^B + G_{il}^A G_{kj}^B + G_{kl}^A G_{ij}^B) \right]^2 \quad (3.39)$$

この誤差関数を最小化する基本行列 G^A 、 G^B を直接求めることができれば、基本テンソルの最適分解が可能になる。現在のところ、このエネルギー関数の最小化によって、テンソル方程式 (3.29) のすべての成分方程式を同時に最小2乗法の意味で最適に満たす解を求める解析的方法は見い出していない。以下では、式 (3.34) で定義した基本テンソルの成分からなるベクトルのうち、 g_4 を除き、 g_1, g_2, g_3 に含まれる27個の成分に関する方程式を平均的に満たす解を複数の固有値問題の組合せによって準最適に求める方法を示す。誤差に対する最適性を考慮しなければ g_4 の成分は基本行列の分解に関しては全く冗長な成分である。

基本行列の2通りの表現法 まず、任意の基本行列 G が、「単位行ベクトル」 $b_i^T = [b_{i1}, b_{i2}, b_{i3}]$ ($\|b_i\| = \sqrt{b_{i1}^2 + b_{i2}^2 + b_{i3}^2} = 1; i = 1, 2, 3$), 「単位列ベクトル」 $c_j = [c_{1j}, c_{2j}, c_{3j}]^T$ ($\|c_j\| = \sqrt{c_{1j}^2 + c_{2j}^2 + c_{3j}^2} = 1; j = 1, 2, 3$), それらの係数 ζ_i, η_j によって2

通りに表せることを指摘する。

$$\mathbf{G} = \begin{bmatrix} \zeta_1 \mathbf{b}_1^T \\ \zeta_2 \mathbf{b}_2^T \\ \zeta_3 \mathbf{b}_3^T \end{bmatrix} = [\eta_1 \mathbf{c}_1, \eta_2 \mathbf{c}_2, \eta_3 \mathbf{c}_3] \quad (3.40)$$

$\zeta = (\zeta_1, \zeta_2, \zeta_3)^T$ 、 $\eta = (\eta_1, \eta_2, \eta_3)^T$ とする。 $\|\mathbf{G}\| = \sqrt{2}$ から、 $\|\zeta\| = \sqrt{\zeta_1^2 + \zeta_2^2 + \zeta_3^2} = \sqrt{2}$ 、 $\|\eta\| = \sqrt{\eta_1^2 + \eta_2^2 + \eta_3^2} = \sqrt{2}$ なる拘束がある。以下では、 \mathbf{G}^A 、 \mathbf{G}^B に対応して、右肩に物体記号を付けた $(\mathbf{b}_i^A)^T$ 、 \mathbf{c}_j^B 、 ζ_i^A 、 η_j^B などの記法を用いる。

部分テンソルの最適分解による単位行ベクトルと単位列ベクトルの計算 基本テンソル $\hat{G}_{(ik)}^{AB}$ の部分テンソルである $\hat{G}_{(ii)}^{AB}$ ($i = k$ を満たす成分) と単位行ベクトルの間には次式の関係がある。

$$\hat{G}_{(ii)}^{AB} = \frac{1}{2} (G_{ij}^A G_{il}^B + G_{il}^A G_{ij}^B) = \frac{1}{2} \{ (\zeta_i^A b_{ij}^A) (\zeta_i^B b_{il}^B) + (\zeta_i^B b_{ij}^B) (\zeta_i^A b_{il}^A) \} \quad (3.41)$$

この部分テンソルは、ベクトル \mathbf{g}_1 および \mathbf{g}_2 の推定値だけから構成される。この関係に基づいて、 \mathbf{b}_i^A と \mathbf{b}_i^B を求めるための計算アルゴリズムが得られる。この計算アルゴリズムを以下に示す。

$i = 1, 2, 3$ のそれぞれの場合について、添字 j を行、添字 l を列とする 3×3 対称行列 \mathbf{S}_i を

$$\mathbf{S}_i = \begin{bmatrix} \bar{G}_{(ii)}^{AB} & \bar{G}_{(ii)}^{AB} & \bar{G}_{(ii)}^{AB} \\ \bar{G}_{(ii)}^{AB} & \bar{G}_{(ii)}^{AB} & \bar{G}_{(ii)}^{AB} \\ \bar{G}_{(ii)}^{AB} & \bar{G}_{(ii)}^{AB} & \bar{G}_{(ii)}^{AB} \end{bmatrix} \quad (3.42)$$

と定義する。その固有値展開を

$$\mathbf{S}_i = \sigma_{i+} \mathbf{s}_{i+} \mathbf{s}_{i+}^T + \sigma_{i0} \mathbf{s}_{i0} \mathbf{s}_{i0}^T + \sigma_{i-} \mathbf{s}_{i-} \mathbf{s}_{i-}^T \quad (3.43)$$

とする。固有値はすべて実数で、 $\sigma_{i+} \geq \sigma_{i0} \geq \sigma_{i-}$ とする。誤差と退化のない理想的な場合は、 $\sigma_{i+} > 0$ 、 $\sigma_{i0} = 0$ 、 $\sigma_{i-} < 0$ となる。誤差がある場合でも、次の公式によって \mathbf{b}_i^A 、 \mathbf{b}_i^B の最適推定値が得られる¹。

$$\bar{\mathbf{b}}_i^A, \bar{\mathbf{b}}_i^B = \pm \sqrt{\frac{\sigma_{i0} - \sigma_{i-}}{\sigma_{i+} - \sigma_{i-}}} \mathbf{s}_{i+} + \sqrt{\frac{\sigma_{i+} - \sigma_{i0}}{\sigma_{i+} - \sigma_{i-}}} \mathbf{s}_{i-} \quad (3.44)$$

¹ $\mathbf{S}_i \neq 0$ ($i = 1, 2, 3$) が必要である。 $\mathbf{S}_i = 0$ の場合は、このアルゴリズムは破綻するが、あらかじめ座標系を原点の周りに適当な角度だけ回転すればこれを避けることができる。

ベクトル \bar{b}_i^A, \bar{b}_i^B のそれぞれに符号の不定性が残る。右辺の2つの解を \bar{b}_i^A と \bar{b}_i^B のどちらかに割り振らなければならないかも知、この段階では決定できない。

単位列ベクトルも同様の計算で推定できる。この場合は、基本テンソル $\hat{G}_{(ik)}^{AB}$ の部分テンソルである $\hat{G}_{(jj)}^{AB}$ ($j = 1$ を満たす成分) と単位列ベクトルの間の次の関係を用いる。

$$\hat{G}_{(jj)}^{AB} = \frac{1}{2} (G_{ij}^A G_{kj}^B + G_{kj}^A G_{ij}^B) = \frac{1}{2} \{ (\eta_j^A c_{ji}^A) (\eta_j^B c_{jk}^B) + (\eta_j^B c_{ji}^B) (\eta_j^A c_{jk}^A) \} \quad (3.45)$$

この部分テンソルは、ベクトル g_1 および g_3 の推定値だけから構成される。

$j = 1, 2, 3$ のそれぞれの場合について、添字 i を行、添字 k を列とする 3×3 対称行列 T_j を

$$T_j = \begin{bmatrix} \bar{G}_{(11)}^{AB} & \bar{G}_{(12)}^{AB} & \bar{G}_{(13)}^{AB} \\ \bar{G}_{(21)}^{AB} & \bar{G}_{(22)}^{AB} & \bar{G}_{(23)}^{AB} \\ \bar{G}_{(31)}^{AB} & \bar{G}_{(32)}^{AB} & \bar{G}_{(33)}^{AB} \end{bmatrix} \quad (3.46)$$

と定義する。さらに、その固有値展開を

$$T_j = \tau_{j+} t_{j+} t_{j+}^T + \tau_{j0} t_{j0} t_{j0}^T + \tau_{j-} t_{j-} t_{j-}^T \quad (3.47)$$

とする。すると、次の公式で c_j^A と c_j^B の最適推定値が得られる (ただし、 $T_j \neq 0$ ($j = 1, 2, 3$) が必要)。

$$\bar{c}_j^A, \bar{c}_j^B = \pm \sqrt{\frac{\tau_{j0} - \tau_{j-}}{\tau_{j+} - \tau_{j-}}} t_{j+} + \sqrt{\frac{\tau_{j+} - \tau_{j0}}{\tau_{j+} - \tau_{j-}}} t_{j-} \quad (3.48)$$

単位行ベクトルの場合と同様に、各ベクトルに符号の不定性が残る。右辺の2つの解を \bar{c}_j^A と \bar{c}_j^B のどちらかに割り振らなければならないかも知、この段階では決定できない。

単位行ベクトル、単位列ベクトルの符号の不定性は、係数 ζ_i と η_j に吸収させる。したがって、以後、考慮しない。なお、本節のアルゴリズムの導出は付録 B.1.2 に述べた。

係数 ζ_i, η_j の計算と物体対応の決定 こうして計算された $\bar{b}_i^A, \bar{b}_i^B, \bar{c}_j^A, \bar{c}_j^B$ の候補から、式 (3.40) の関係を用いて ζ と η を求める。これらの候補について、物体 A または B への帰属もこの段階で決定する。そのために、2乗誤差関数

$$e(\zeta, \eta) = \left\| \begin{bmatrix} \zeta_1 \bar{b}_1^T \\ \zeta_2 \bar{b}_2^T \\ \zeta_3 \bar{b}_3^T \end{bmatrix} - [\eta_1 \bar{c}_1, \eta_2 \bar{c}_2, \eta_3 \bar{c}_3] \right\|^2 = \sum_{i,j=1}^3 (\zeta_i \bar{b}_{ij} - \eta_j \bar{c}_{ji})^2 \quad (3.49)$$

の最小化問題を考える。この式は、パラメータ ζ_i, η_j に関して2次形式である。これを活かした簡潔なアルゴリズムを導くため、本来の規格化条件 $\|\zeta\| = \|\eta\| = \sqrt{2}$ でなく、これら

2個のベクトルを合成した6次元ベクトル $\nu = (\zeta^T, -\eta^T)^T = (\zeta_1, \zeta_2, \zeta_3, -\eta_1, -\eta_2, -\eta_3)^T$ に関する1個の規格化条件 $\|\nu\| = \sqrt{\|\zeta\|^2 + \|\eta\|^2} = 2$ を用いる。これは、もとの正規化条件から見ると必要条件である。この条件で求められたベクトルの要素から ζ と η を取り出して各々規格化する。もとのデータに誤差が無ければ正しい値が得られるはずである。誤差があっても、近似解が得られる。

条件 $\|\nu\| = 2$ のもとで、式(3.49)を最小化する問題は、次の6次元固有値問題に帰着される。

$$U\nu = \rho\nu \quad (3.50)$$

ρ は固有値で、最小固有値の4倍が2乗誤差であり、最小固有値に対応する固有ベクトルが ν の推定値である。行列 U は、次式で定義される実対称行列である。

$$U \stackrel{\text{def}}{=} \begin{bmatrix} I_3 & V \\ V^T & I_3 \end{bmatrix} \quad (3.51)$$

I_3 は3次の単位行列で、 V は、

$$V \stackrel{\text{def}}{=} \begin{bmatrix} \bar{b}_{11}\bar{c}_{11} & \bar{b}_{12}\bar{c}_{21} & \bar{b}_{13}\bar{c}_{31} \\ \bar{b}_{21}\bar{c}_{12} & \bar{b}_{22}\bar{c}_{22} & \bar{b}_{23}\bar{c}_{32} \\ \bar{b}_{31}\bar{c}_{13} & \bar{b}_{32}\bar{c}_{23} & \bar{b}_{33}\bar{c}_{33} \end{bmatrix} \quad (3.52)$$

で定義される。固有値問題(3.50)を用いて運動 A, B への帰属決定とパラメータ ζ_i, η_j の推定を同時に行える。以下に、その手順を示す。

6個のベクトル $\bar{b}_1, \bar{b}_2, \bar{b}_3, \bar{c}_1, \bar{c}_2, \bar{c}_3$ のそれぞれに関して、物体への帰属対応に A, B 2通りの可能性がある。したがって、全部で、 $2^6 = 64$ 通りの組合せを検査する。このうちで、2通りのみが正しい組合せである。この2通りの正解は、互いに物体のラベル A, B を入れ換えたものに相当するはずである。誤差を考慮した検査を行うために、次の手順を実行する。

1. 64通りの候補に関して行列 U を作り、固有値問題(3.50)を解く。それぞれの候補に対する最小固有値とそれに対応する固有ベクトルを計算する。
2. 1. で求めた最小固有値を比較し、小さい方から2つ選び出す。この2つが、物体のラベルを互いに入れ換えた場合に相当することを検査する。
3. 2. で選び出した場合の固有ベクトル $\bar{\nu}$ から ζ と η の成分を取り出す。 $\|\zeta\| = \sqrt{2}$, $\|\eta\| = \sqrt{2}$ となるように再規格化された推定値 $\bar{\zeta}, \bar{\eta}$ を計算する。

最後の3. の規格化では符号に関して2通りの任意性が存在する。これは、基本行列の符号の任意性によるもので、後で考慮される。したがって、ここではどちらか一方を求めればよい(ただし、 $\bar{\zeta}$ と $\bar{\eta}$ の間で矛盾しない様に符号を選ぶ)。以上で、 $\bar{b}_i^X, \bar{\zeta}_i^X (i = 1, 2, 3; X = A, B)$ と $\bar{c}_j^X, \bar{\eta}_j^X (j = 1, 2, 3; X = A, B)$ が求められた。最後に、式(3.40)に基づき G^X の推定値 \bar{G}^X を計算する(以下、 X は、 A あるいは B を意味するものとする)。このとき、なるべく多くの情報を用いた方がロバストであると考え、行ベクトルに基づいた基本行列の推定値と列ベクトルに基づいた基本行列の推定値の平均をとった後、再規格化する。

本節のアルゴリズム導出は、付録B.1.3に述べた。

基本行列の並進ベクトル、回転行列への分解 基本行列 \bar{G}^X の分解は、従来の単一運動に関する方法がほぼそのまま利用できる。

並進ベクトルは、 $\bar{G}^X(\bar{G}^X)^T$ の最小固有値に対する単位固有ベクトル \bar{h}^X として求められる(付録B.1.4.1)。単一運動の場合には、拘束条件 $|\mathbf{h} \quad \mathbf{m}'_\alpha \quad \mathbf{G}\mathbf{m}_\alpha| \geq 0$ を用いて並進ベクトルの符号まで決定できる[22][12]。ところが、多重運動の場合には、各対応点对 $\{\mathbf{m}_\alpha, \mathbf{m}'_\alpha\}$ がどちらの物体に属するかがこの段階では決定されていない。したがって、この拘束条件を用いることはできない。この拘束条件は、もともと構造パラメータの正值性の拘束 $r_\alpha > 0, r'_\alpha > 0$ から帰結される条件である。そこで、後の第3段階において、点对対応のクラスタリング後に、直接この正值性の拘束条件を用いて並進ベクトルの符号を判定する。

回転行列は、極分解[108]、四元数[107]、特異値分解[101]などの方法によって計算できる。例えば、四元数による方法では、行列 $\bar{K}^X = \bar{G}^X \times \bar{h}^X$ の成分から作られる4次実対称行列の最大固有値に対応する単位固有ベクトルから回転行列の推定値 \bar{R}^X が閉形式で求められる(付録B.1.4.2にその手順を示す。)

\bar{G}^X の分解に際して、並進ベクトルの符号の不定性から、式(3.8)で説明した2重の多義性が存在する。さらに、 \bar{G}^X 自身に符号の不定性があるので、物体 A および B のそれぞれに4通りの多義性が残る。結局、この段階では、全体で $4^2 = 16$ 通りの多義性が残る。物体 X の運動パラメータ候補のうち1組を $\mathcal{M}_{++}^X = \{\bar{h}^X, \bar{R}^X\}$ とし、他の3組の運動パラメータ候補を、 $\mathcal{M}_{+-}^X = \{\bar{h}^X, \mathbf{I}_{\bar{h}^X} \bar{R}^X\}$, $\mathcal{M}_{-+}^X = \{-\bar{h}^X, \bar{R}^X\}$, $\mathcal{M}_{--}^X = \{-\bar{h}^X, \mathbf{I}_{\bar{h}^X} \bar{R}^X\}$ とする。

3.1.4.4 第3段階：一撃クラスタリングと偽の解の除去

運動パラメータが有限個の候補に絞られた。次は、剛体運動の方程式である式(3.1)とそれから導かれた公式(3.9)に基づいて物体構造パラメータ r'_α, r_α を計算し、特徴点データの

クラスタリングを行う。さらに、制約条件 $r_\alpha > 0$ 、 $r'_\alpha > 0$ を用いて解を一意的に決定する。クラスタリングは、式 (3.1) の2乗誤差

$$\mathcal{E}(\mathbf{h}, \mathbf{R}, r'_\alpha, r_\alpha, \mathbf{m}'_\alpha, \mathbf{m}_\alpha) = \|r'_\alpha \mathbf{m}'_\alpha - r_\alpha \mathbf{R} \mathbf{m}_\alpha - \mathbf{h}\|^2 \quad (3.53)$$

を用いて行う。

まず、公式 (3.9) にもとづいて運動パラメータが \mathcal{M}_{++}^X の場合の物体構造パラメータを $\alpha = 1, \dots, N$ にわたって計算し、それらを $\bar{r}'_\alpha^X(\mathcal{M}_{++}^X)$ 、 $\bar{r}_\alpha^X(\mathcal{M}_{++}^X)$ とする。

次に、このときの2乗誤差を式 (3.53) を用いて計算する。それを

$$\bar{\mathcal{E}}_\alpha^X = \mathcal{E}(\mathcal{M}_{++}^X, \bar{r}'_\alpha^X(\mathcal{M}_{++}^X), \bar{r}_\alpha^X(\mathcal{M}_{++}^X), \mathbf{m}'_\alpha, \mathbf{m}_\alpha) \quad (3.54)$$

とおく。次の集合を計算する。

$$\begin{aligned} \mathcal{I}^A &= \{\alpha | \alpha = 1, \dots, N, \bar{\mathcal{E}}_\alpha^A \leq \bar{\mathcal{E}}_\alpha^B\}, \\ \mathcal{I}^B &= \{\alpha | \alpha = 1, \dots, N, \bar{\mathcal{E}}_\alpha^A \geq \bar{\mathcal{E}}_\alpha^B\} \end{aligned} \quad (3.55)$$

\mathcal{I}^X は、物体 X 上の点であると推定された特徴点对応のインデクス集合を表す。

最後に、運動パラメータを一意的に決定する。 $\alpha \in \mathcal{I}^X$ に対して $\bar{r}'_\alpha^X(\mathcal{M}^X) > 0$ かつ $\bar{r}_\alpha^X(\mathcal{M}^X) > 0$ という条件を満たす運動パラメータ \mathcal{M}^X を $\mathcal{M}_{++}^X, \mathcal{M}_{+-}^X, \mathcal{M}_{-+}^X, \mathcal{M}_{--}^X$ の中から選び出す。ノイズの影響を小さくするため、次の量を最大にする運動パラメータを選択する。

$$J(\mathcal{M}^X) = \sum_{\alpha \in \mathcal{I}^X} \{sgn(\bar{r}_\alpha(\mathcal{M}^X)) + sgn(\bar{r}'_\alpha(\mathcal{M}^X))\} \quad (3.56)$$

関数 $sgn()$ は、次式で定義される。

$$sgn(x) = \begin{cases} 1 & \text{for } x > 0 \\ 0 & \text{for } x = 0 \\ -1 & \text{for } x < 0 \end{cases} \quad (3.57)$$

以上で、各物体の運動パラメータの計算と特徴点の一撃クラスタリングが終了し、それに対応する構造パラメータも計算された。

3.1.4.5 固有値問題における退化判定にもとづいた動剛物体数の推定

上記のアルゴリズムは、条件 $N^A \geq 8$ 、 $N^B \geq 8$ 、 $N = N^A + N^B \geq 35$ が満たされることが必要条件になっている。これらの条件のうちで、 $N \geq 35$ は、アルゴリズムの実行の前に検査することができる。しかし、条件 $N^A \geq 8$ と条件 $N^B \geq 8$ は、事前に知ることができな

い。しかし、基本テンソル推定のための36次元固有値方程式(3.33)の零固有値の数から、基本拘束方程式(3.26)の拘束としての退化次数を知ることができる。これから、画像に含まれる剛体運動の数を予測できる。次の表は、零固有値の数と剛体運動数の対応である。ただし、ここで書いたのは、必要条件であり、その数以上の剛体運動が存在する可能性を否定しない。極端な場合、各特徴点が、すべて異なる剛体に属すると考える事さえできる。つまり、ここで示した解釈は、もっとも「単純な解釈」である。

場合分け	零固有値の数	予測される剛体運動の数(もっとも単純な解釈)
(1)	0	2個以上の運動(拘束条件が充足されない場合)
(2)	1	2個の剛体運動(アルゴリズムの正しい前提条件)
(3)	2-8	2個のうちの1個の剛体が8個以上の特徴点を持たない場合
(4)	9	単一の剛体運動
(5)	> 9	単一の剛体運動の退化した場合

(2)の場合以外は、本節のアルゴリズムでは、正しい運動パラメータは求まらない。

3.1.5 計算機シミュレーションによる検証

アルゴリズムの正当性を確かめるための計算機シミュレーションを、乱数を用いて生成した合成データを入力として行った。物体A, Bの運動パラメータは、次の様に設定した。

$$\begin{aligned} \text{回転軸: } l^A &= (0.683763, -0.455842, 0.569803) \\ l^B &= (0.380143, -0.912343, 0.152057) \end{aligned} \quad (3.58)$$

$$\begin{aligned} \text{回転角: } \Omega^A &= 30^\circ \\ \Omega^B &= 20^\circ \end{aligned} \quad (3.59)$$

$$\begin{aligned} \text{並進ベクトル: } h^A &= (0.659062, 0.746937, 0.087875) \\ h^B &= (-0.666667, -0.333333, -0.666667) \end{aligned} \quad (3.60)$$

特徴点は、座標 X, Y, Z のそれぞれが、 $[-50, 50]$ 、 $[-50, 50]$ 、 $[50, 100]$ の範囲をもつ区間からなる直方体内に様に発生させた。特徴点数は、各物体に25個ずつとした($N^A = N^B = 25, N = 50$)。正規化ベクトルデータ m'_α, m_α には0.1%の一様なノイズを与え、再正規化したものをサンプルデータとして用いた。図3.2に生成データ、図3.3に基本テンソル推定時の固有値の分布、図3.4に基本行列の推定結果と真値、図3.5に並進ベクトル・回転パラメータ候補の推定結果、図3.6に選択された運動パラメータの一意解とその真値、図3.7に3次元構造の復元結果とその真値をそれぞれ示した。これらの計算結果から、正し

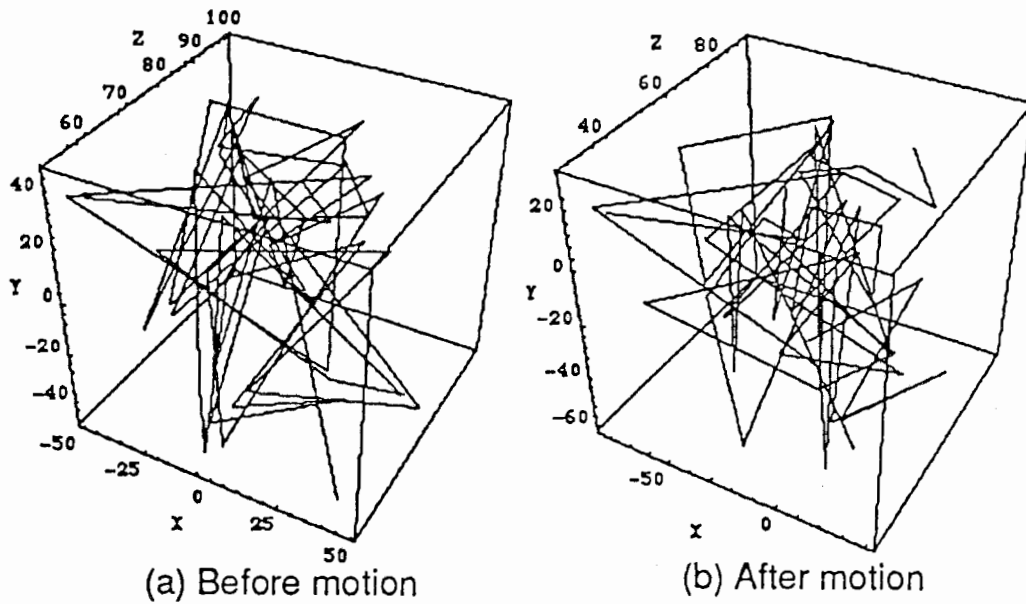


Figure 3.2: シミュレーションのために生成された特徴点对応データ

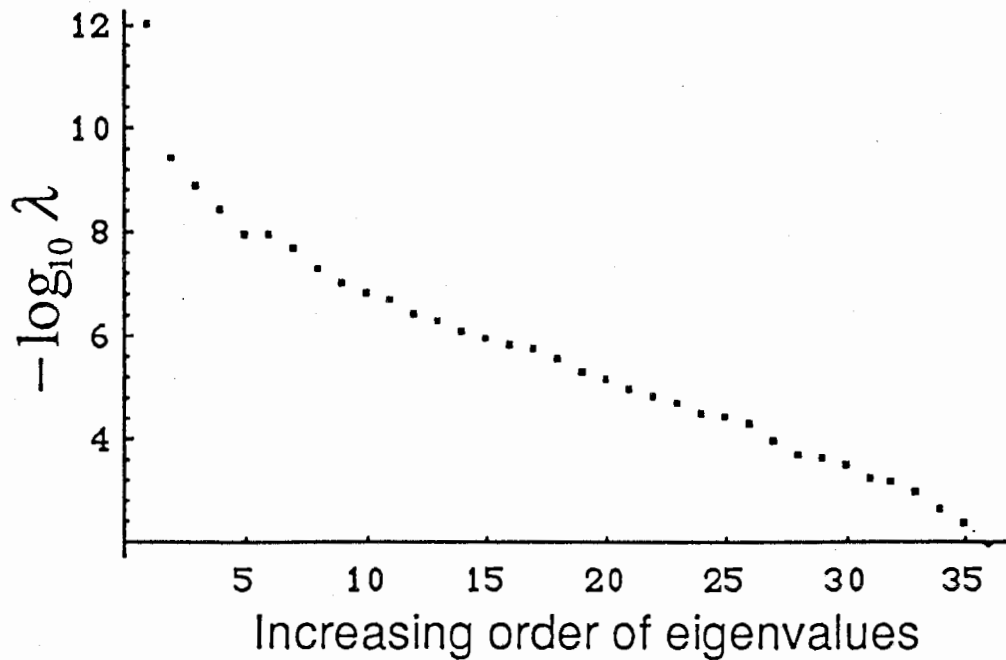


Figure 3.3: 基本テンソル推定の際の 36 個の固有値の分布

\bar{G}^A	\bar{G}^B																		
Estimated essential matrix of object A:	Estimated essential matrix of object B:																		
<table border="1"> <tr><td>-0.189113</td><td>-0.158261</td><td>-0.70989</td></tr> <tr><td>0.106949</td><td>0.225127</td><td>0.611715</td></tr> <tr><td>0.537424</td><td>-0.83453</td><td>0.1169</td></tr> </table>	-0.189113	-0.158261	-0.70989	0.106949	0.225127	0.611715	0.537424	-0.83453	0.1169	<table border="1"> <tr><td>0.0852496</td><td>-0.613755</td><td>0.405199</td></tr> <tr><td>0.411329</td><td>-0.127486</td><td>-0.836072</td></tr> <tr><td>-0.295747</td><td>0.692699</td><td>0.00963156</td></tr> </table>	0.0852496	-0.613755	0.405199	0.411329	-0.127486	-0.836072	-0.295747	0.692699	0.00963156
-0.189113	-0.158261	-0.70989																	
0.106949	0.225127	0.611715																	
0.537424	-0.83453	0.1169																	
0.0852496	-0.613755	0.405199																	
0.411329	-0.127486	-0.836072																	
-0.295747	0.692699	0.00963156																	
True essential matrix of object A:	True essential matrix of object B:																		
<table border="1"> <tr><td>0.187865</td><td>0.150824</td><td>0.712458</td></tr> <tr><td>-0.10301</td><td>-0.231092</td><td>-0.614874</td></tr> <tr><td>-0.533406</td><td>0.833106</td><td>-0.117002</td></tr> </table>	0.187865	0.150824	0.712458	-0.10301	-0.231092	-0.614874	-0.533406	0.833106	-0.117002	<table border="1"> <tr><td>-0.0844481</td><td>0.619377</td><td>-0.405951</td></tr> <tr><td>-0.421921</td><td>0.129715</td><td>0.833094</td></tr> <tr><td>0.295409</td><td>-0.684235</td><td>-0.010596</td></tr> </table>	-0.0844481	0.619377	-0.405951	-0.421921	0.129715	0.833094	0.295409	-0.684235	-0.010596
0.187865	0.150824	0.712458																	
-0.10301	-0.231092	-0.614874																	
-0.533406	0.833106	-0.117002																	
-0.0844481	0.619377	-0.405951																	
-0.421921	0.129715	0.833094																	
0.295409	-0.684235	-0.010596																	

Figure 3.4: 基本行列の推定結果と真値

\bar{h}^A	\bar{h}^B						
<table border="1"> <tr><td>-0.657565</td><td>-0.749269</td><td>-0.0787612</td></tr> </table>	-0.657565	-0.749269	-0.0787612	<table border="1"> <tr><td>-0.67403</td><td>-0.334205</td><td>-0.65878</td></tr> </table>	-0.67403	-0.334205	-0.65878
-0.657565	-0.749269	-0.0787612					
-0.67403	-0.334205	-0.65878					
$-\bar{h}^A$	$-\bar{h}^B$						
<table border="1"> <tr><td>0.657565</td><td>0.749269</td><td>0.0787612</td></tr> </table>	0.657565	0.749269	0.0787612	<table border="1"> <tr><td>0.67403</td><td>0.334205</td><td>0.65878</td></tr> </table>	0.67403	0.334205	0.65878
0.657565	0.749269	0.0787612					
0.67403	0.334205	0.65878					

(a) Candidate translation vectors

\bar{R}^A	\bar{R}^B																		
<table border="1"> <tr><td>0.929521</td><td>-0.325633</td><td>-0.173071</td></tr> <tr><td>0.243841</td><td>0.89481</td><td>-0.373975</td></tr> <tr><td>0.276644</td><td>0.305416</td><td>0.911147</td></tr> </table>	0.929521	-0.325633	-0.173071	0.243841	0.89481	-0.373975	0.276644	0.305416	0.911147	<table border="1"> <tr><td>0.207315</td><td>0.559199</td><td>0.802694</td></tr> <tr><td>0.541867</td><td>-0.748791</td><td>0.381697</td></tr> <tr><td>0.814494</td><td>0.355822</td><td>-0.458247</td></tr> </table>	0.207315	0.559199	0.802694	0.541867	-0.748791	0.381697	0.814494	0.355822	-0.458247
0.929521	-0.325633	-0.173071																	
0.243841	0.89481	-0.373975																	
0.276644	0.305416	0.911147																	
0.207315	0.559199	0.802694																	
0.541867	-0.748791	0.381697																	
0.814494	0.355822	-0.458247																	
$\bar{I}_T \bar{R}^A$	$\bar{I}_T \bar{R}^B$																		
<table border="1"> <tr><td>0.143247</td><td>0.9574</td><td>-0.25073</td></tr> <tr><td>0.978535</td><td>-0.174937</td><td>-0.108929</td></tr> <tr><td>-0.148151</td><td>-0.229745</td><td>-0.961909</td></tr> </table>	0.143247	0.9574	-0.25073	0.978535	-0.174937	-0.108929	-0.148151	-0.229745	-0.961909	<table border="1"> <tr><td>0.948516</td><td>-0.0724476</td><td>-0.308332</td></tr> <tr><td>0.0312298</td><td>0.990137</td><td>-0.136577</td></tr> <tr><td>0.315186</td><td>0.119916</td><td>0.941423</td></tr> </table>	0.948516	-0.0724476	-0.308332	0.0312298	0.990137	-0.136577	0.315186	0.119916	0.941423
0.143247	0.9574	-0.25073																	
0.978535	-0.174937	-0.108929																	
-0.148151	-0.229745	-0.961909																	
0.948516	-0.0724476	-0.308332																	
0.0312298	0.990137	-0.136577																	
0.315186	0.119916	0.941423																	

(b) Candidate rotation matrices

Figure 3.5: 基本行列の分解によって得られた各物体の並進ベクトル・回転行列の候補

$-\bar{h}^A$			\bar{h}^B		
0.657565	0.749269	0.0787612	-0.67403	-0.334205	-0.65878
True translation vector of object A:			True translation vector of object B:		
0.659062	0.746937	0.087875	-0.666667	-0.333333	-0.666667

(a) Selected translation vectors and their true values
(true vectors are normalized)

Selected rotation matrices for each object: $\bar{R}^A, \bar{I}_T, \bar{R}^B$

Estimated rotation axis vector of object A:	Estimated rotation axis vector of object B:
0.683464 -0.452411 0.572888	0.376014 -0.914064 0.151989
True rotation axis vector of object A:	True rotation axis vector of object B:
0.683763 -0.455842 0.569803	0.380143 -0.912343 0.152057
Estimated rotation angle: 29.803	Estimated rotation angle: 19.942
True rotation angle: 30.	True rotation angle: 20.

(b) Rotation axis vectors and rotation angles computed
from selected rotation matrices.

Figure 3.6: 物体構造パラメータの符号にもとづいて選択された運動パラメータの一意解とその真値

くもとの運動パラメータの推定が可能であることと、この運動パラメータにもとづく特徴点の一撃的クラスタリングと物体の3次元構造の復元が可能であることが確認された。

図3.8は、第3.1.4.5節で説明した方程式が退化する場合の結果を示す。固有値問題における零固有値の数が物体運動の数と退化条件を反映している。

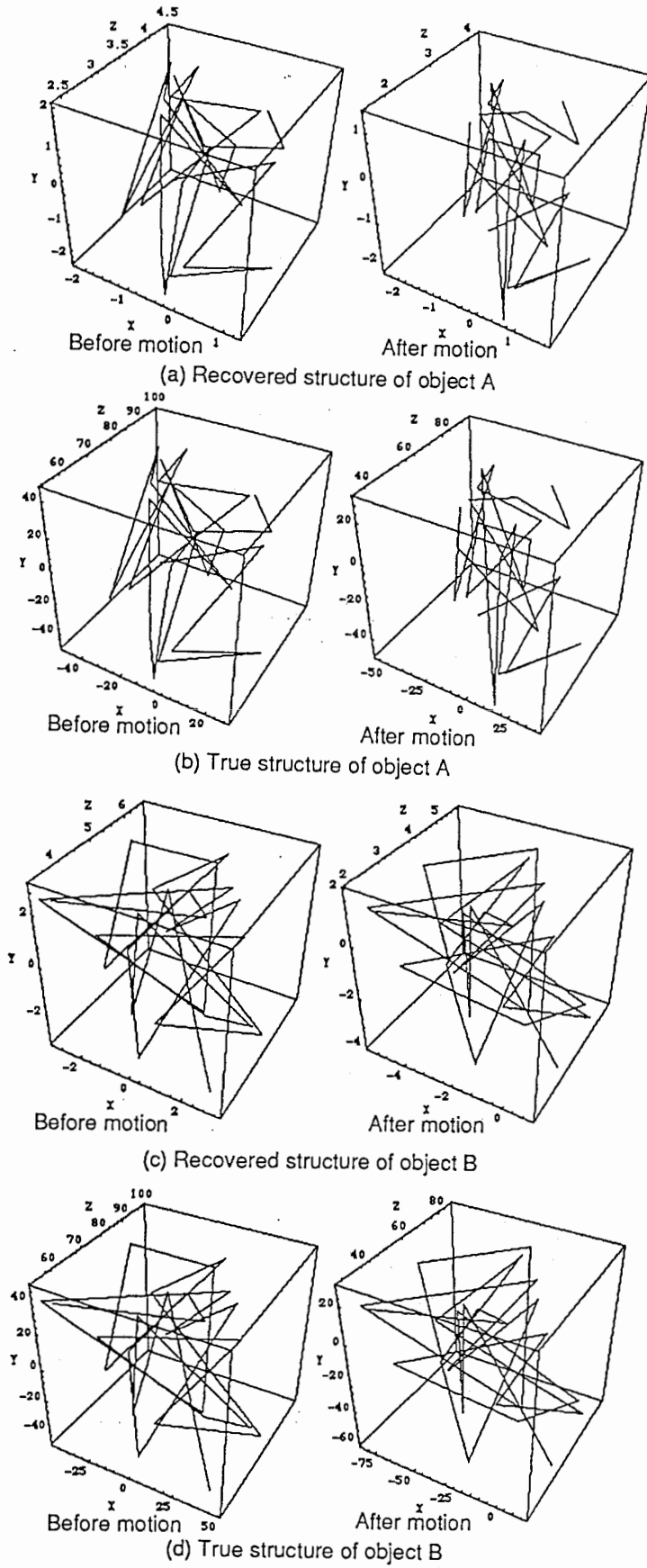
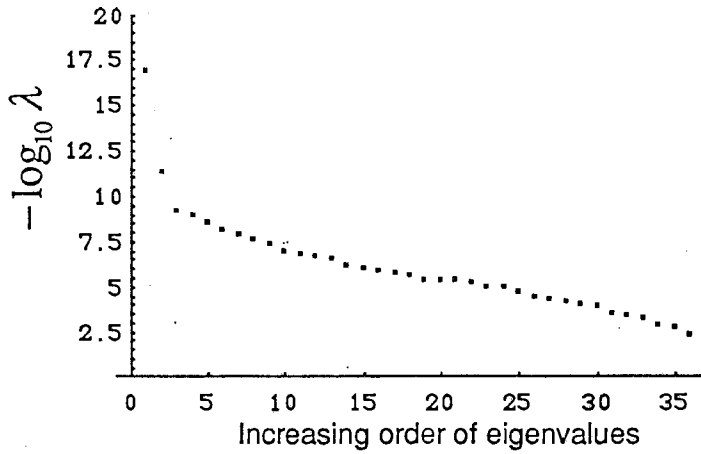
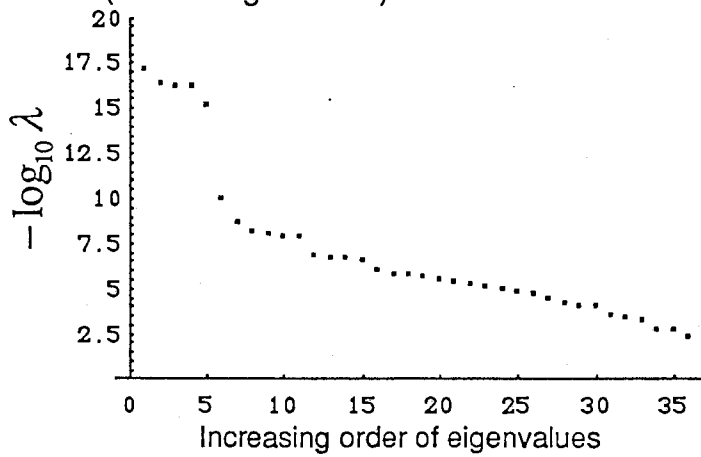


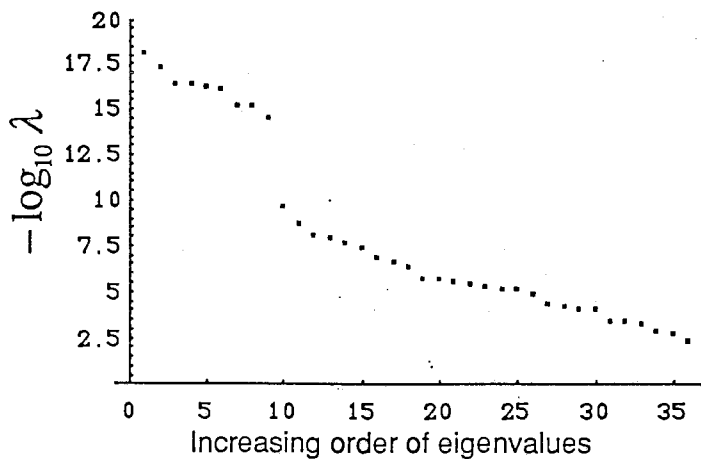
Figure 3.7: 3次元物体構造の復元結果と真値



(a) One of the objects has 8 feature points: proper condition for the algorithm (1 zero eigenvalue).



(b) One of the objects has 4 feature points (5 zero eigenvalues).



(c) Only one object exists (9 zero eigenvalues).

Figure 3.8: 方程式が退化する場合の固有値の分布

3.1.6 まとめ

多重3次元剛体運動による画像面上の特徴点の変換が満たす基本拘束方程式を、透明視の普遍的計算理論である重ね合わせの原理に基づいて導出した。それに基づいて、2組の3次元剛体運動パラメータを、2枚の画像フレーム間の特徴点对応からセグメンテーションやクラスタリングを経ずに、直接解析的に求める線形アルゴリズムを提案した。さらに、推定された運動パラメータに基づいた一撃クラスタリングによって、特徴点を各物体に瞬時に分離できることを示した。さいごに、数値シミュレーションによってその動作を確認した。本アルゴリズムは、計算全体の閉形式性を保持しつつ、中間段階で可能な限り多くの拘束条件が最適に満たされるパラメータを求める数値技法を取り入れた。したがって、これは、ノイズに対してある程度のロバスト性を持つ準最適な線形アルゴリズムである。

このアルゴリズムの存在は、十分多くの特徴点があり、それらが特殊な配置になれば、2個の剛体運動は、2フレームの特徴点对応情報から一意に定まる、という解の一意性の構成的証明にもなっている。しかし、従来、線形アルゴリズムにおいて臨界面面 (critical surface) と呼ばれる特殊な条件を満たす曲面上にすべての特徴点がある場合には、アルゴリズムが破綻することが知られている [114] [106]。それと同様な意味で、本論文のアルゴリズムが破綻する特別な配置とその検出・回避方法についての研究が将来必要になると考えられる。また、トランスペアレンシーによって単一運動の場合の多義解 [115] [117] に加えて新たな多義性が特殊な配置によって生じる可能性がある。さらに、特徴点の配置が特殊でなくても、運動が単一の剛体運動であったり、各物体に対応する特徴点の数が8点に満たない場合にも、本論文のアルゴリズムが破綻する可能性がある。これらの特殊な場合の研究もアルゴリズムの信頼性向上と適用可能領域の拡大のために必要である。

本節では、計算原理の確認という目的から、実際的な誤差モデルは採り入れずにアルゴリズムを導出した。しかし、従来、単一運動の線形アルゴリズムに対しては、推定の偏差に対する対策 [13] や、エネルギー関数のノルムに統計的な誤差モデルにもとづいた推定方法が研究されている。非線形最適化手法との統合も実用に耐えるアルゴリズムには必要である [118] [123] [105]。本アプローチが基本拘束方程式のレベルから展開されており、データに対する直接的な拘束条件となっているため、単一運動に対して考案されたこれらの手法が直接応用できる。

本節では、運動立体視のうちで、特徴点对応にもとづく理論を多重運動に拡張した。オプティカルフローからの運動立体視 [112] [103] や、線特徴のフレーム間対応からの運動立体視 [111] [119] [125] に関しても、多重運動への拡張を研究する価値がある。

3.2 多価関数への標準正則化理論の拡張となめらかな多重表面復元

あらまし 自然界、人工環境の多くの物体は、滑らかな表面を持っている。この滑らかな表面を復元するためには、運動立体視、ステレオ視などによって推定されたまばらでノイズを含む3次元構造情報から、物体表面を求める処理が必要である。これは、数学的には、まばらに特徴点として与えられた3次元データを、ある滑らかさ拘束を持った曲面で近似する問題である。本節では、視覚情報処理における複数の重なり合った表面の復元問題（表面復元問題における透明視（トランスペアレンシー））に代表される多価関数によるデータの関数近似を実現する。そのために、標準正則化理論を拡張する。標準正則化理論は、関数がある滑らかさをもった一価関数であるという前提条件のもとでデータの関数近似を行うために用いられてきた。しかし、物体境界などにおいて関数が不連続になる場合は、線過程と呼ばれる不連続を表現する補助的な場やそれと等価なメカニズムを導入する必要があった。近年、多重表面の復元のための多層表現法 (multi-layer representation) も提案されている。これもまた、データの所属を表す補助的な場を導入する必要があった。これらのアプローチのいずれも、エネルギー最小化問題が非2次かつ非凸の非線形最適化である。そのため、インプリメンテーションにかなりの工夫が必要であった。この分野の研究は、この最適化をいかにうまく行うかに関心が集まっていた。

本節では、テンソル積空間の概念に基づいた多価関数の直接的表現法を提案する。この表現法は、重ね合わせの原理から導出できる。この直接表現法を用いると、多価関数によるデータの関数近似問題を1個の2次元関数のエネルギー最小化に帰着できる。この場合には、Euler-Lagrange 方程式は線形である。そのため、従来の標準正則化理論のために用いられてきた最適値への収束が保証された最適化手法が、ほぼそのまま利用できる。この拡張された標準正則化理論を用いて、透明視モデルのための2重表面復元の緩和型超並列アルゴリズムを導出する。超並列計算機によるシミュレーション結果も報告する。^a

^a本節の内容は、文献 [179] [191] において発表された。

3.2.1 はじめに

正則化理論 (regularization theory) は、視覚情報処理をはじめとする不良設定の逆問題を解くための一般的な理論的枠組である [162] [69] [134]。その中でも、種々の視覚情報からの滑らかな物体表面復元は、3次元世界を推定するための重要な手段である。人間も初期視覚の問題を解くためにこの理論に基づいたメカニズムを用いているといわれている [69]。高次の視覚機能である学習や物体認識においても正則化理論に基づいたモデルが提案されている [136] [137]。

このように、正則化理論は、視覚情報処理の基本原理として極めて基本的な役割を果たしている。それと同時に、逆問題や関数近似のための基礎理論として工学的にも極めて重要な数理的方法を提供する。しかし、未知関数に関して2次元関数の最小化問題として定式化される標準正則化理論 (Standard Regularization Theory) では、近似関数は空間の至るところで滑らかな一価関数であるという前提条件があった。したがって、視覚情報処理において、物体の境界やかどにおける表面の不連続や、真に重なり合った複数表面の復元が必要な透明視 (トランスペアレンシー) を正しく扱うことができない。

近似関数の不連続を許容した正則化を行う手法には、Geman & Gemanによって提案された滑らかさの伝搬を不連続の位置で局所的に禁止する線過程 (line process) と呼ばれるメカニズム [43]、Terzopoulosによって提案された滑らかさ拘束の重みを適応的に制御 [76] する方法など、不連続の検出と区分的に滑らかな表面の復元を「協調的な」処理によって、同時に実現する方法が提案されている。また、これらの手法の変形や近似解法もいろいろ試みられてきた [4] [130]。そして、これらの方法は、人間の物体境界知覚やエッジ情報に基づいた視覚モジュール統合の基本メカニズムとしても用いられてきた [135]。

しかし、この様に滑らかさの程度を局所的に制御するだけでは、真に重なり合った複数の表面に対応するデータが混在する透明視の場合に、ほとんど至るところで表面が不連続となる。そこで、近年、コンピュータ・ビジョンや視覚の計算理論では多層表現 (multilayer representation) と呼ばれる手法が用いられている。これは、複数のレイヤーの集合として表面や運動を表現する手法である。例えば、Darrel & Pentland は、パターンが多層表現をモデルとして用いて動画像の直接セグメンテーションを行なった [129]。Wang & Adelson は、同様なメカニズムを用いてオプティカルフロー場の直接セグメンテーションを行った [138]。Kersten は、単眼の透明視において [54]、Madarasmiらは、両眼立体視において [132] それぞれ、物体表面の多層表現モデルを用いた多重表面の知覚モデルを提案している。しかし、これらの方法では、データが真に重なり合った場合を想定していない。つまり、表面のモデルは多重であるが、データは画像面の各点に1個ずつしか与えることができなかった。そこで、安藤 [128] は、各データに表面への所属を表す変数を割り当て、エネル

ギー汎関数のデータ項をその変数で切り替えるモデルを導入することにより、データの複数表面へのクラスタリングと滑らかな表面の復元を同時に行う方法を提案した。

これらの方法は、表面や運動のパラメータ以外に多数の補助パラメータを導入している。さらに、それらがエネルギー汎関数のデータ項に積の形にかかる。そのため、汎関数の最小化問題が非2次かつ非凸の一般の最適化となる。したがって、標準正則化理論の枠内では扱えず、局所最小解を避けるために、かなり複雑な最適化手法を導入しなければならなかった。

例えば、エッジを保存した画像復元や、不連続を保存した区分的に滑らかな表面復元の問題に対して、次の最適化手法が用いられてきた。

- Geman & Geman [43] は、線過程との結合マルコフ場モデルで不連続を表現した。それから導かれるエネルギー最小化問題に熱力学をモデルとした模擬焼きなまし (simulated annealing) 法 [131] を用いた。
- Blake & Zisserman [4] は、不連続を含むデータの滑らかな表面による近似に、エネルギー関数の凸関数による近似と最適化を繰り返し行う Graduated Non-Convexity (GNC) アルゴリズムを提案した。
- Geiger & Girosi [130] は、平均場近似 (mean field approximation) と呼ばれる統計物理学の近似手法を用いて、確率的緩和法によらずに決定論的な最適化を行う方法を提案した。

多層表現モデルにおいても、Darrel & Pentland [129]、Kersten [54]、Madarasmiら [132] は模擬焼きなまし法を用いており、安藤 [128] は、平均場近似を用いている。模擬焼きなましは、数学的には大局的最適解への収束が保証されているものの、計算量が膨大である。焼きなましの温度のスケジューリングが難しいという問題もある。平均場近似や GNC 法は、実用的には十分な場合もあるものの、大局的最適値が得られる保証はない。

本節では、エネルギー関数のレベルで補助的な場を一切導入しなくても、テンソル積を用いて、多価関数のための正則化問題を一価のベクトル値関数のための標準正則化問題に変換できることを示す。つまり、多価関数の正則化も、2次のエネルギー汎関数の最適化問題に帰着することができる。(この多価関数に拡張された標準正則化理論を多価標準正則化理論 (Multi-Valued Standard Regularization Theory, MVSRT) と呼ぶ。)

2次元平面上でスパースに配置されたデータの関数近似を行うための正則化理論に基づいた緩和アルゴリズムは、超並列計算機や生体情報処理を想定したマルコフ場モデル、そのアナログ抵抗ネットワークによる実現 [134] などの応用がある。

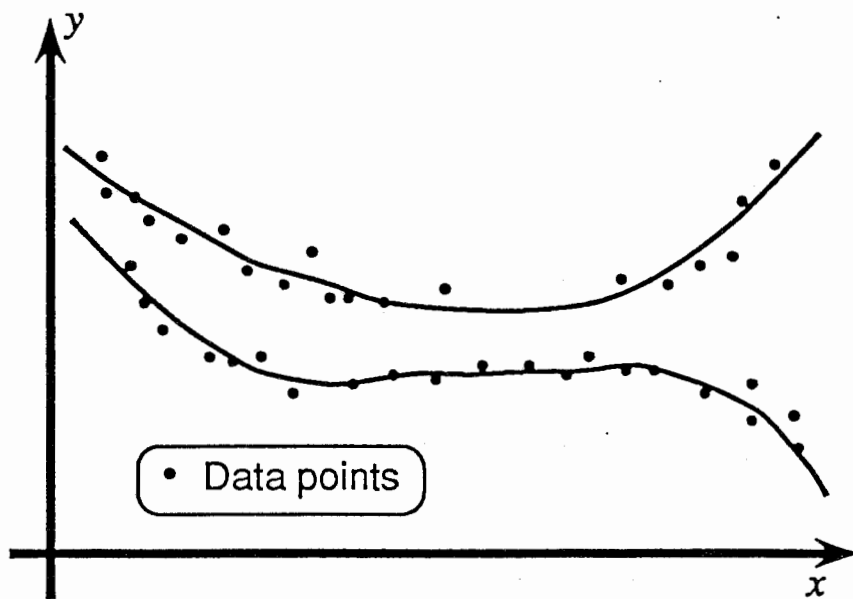


Figure 3.9: 2重曲面(曲線)によってモデル化されるデータの例

本節の構成を述べる。第 3.2.2 節において、2 価スカラー値関数の場合に相当する 2 次元平面上に疎に分布した奥行きデータからの 2 重表面復元について、多価正則化理論に基づいた一解法を示す。エネルギー汎関数の離散格子近似にもとづいて Gauss-Seidel 法による超並列繰り返し計算アルゴリズムを導き、シミュレーションを行う。第 3.2.3 節では、多価標準正則化理論の一般論を述べる。まず、 n 次元空間における h 価スカラー値関数に拡張した場合の定式化を述べる。次に、 n 次元空間における h 価 m 次元ベクトル値関数の場合について、テンソル積を用いた一般論を述べる。最後に、 $h = 2, m = 2$ の場合を例として展開する。

3.2.2 多価標準正則化理論による 2 重表面の復元

多価標準正則化理論の実現例を示す。2 次元空間に 2 層に分布した値を 2 枚の滑らかな曲面で近似する問題を解く標準正則化アルゴリズムを導出する。ビジョンにおいて透明視 (トランスペアレンシー) と呼ばれる問題では、図 3.9 に示した複数の重なり合った表面を復元する必要がある。図の上側の面の様に、いずれかの表面上の一部領域のデータが欠けている部分をも乗り越えて、空間全体を覆う複数枚の滑らかな表面で近似される必要がある。この復元が可能であるためには各点近傍の局所処理ではなく、正則化理論による大局的最適化のアルゴリズムが必要である。

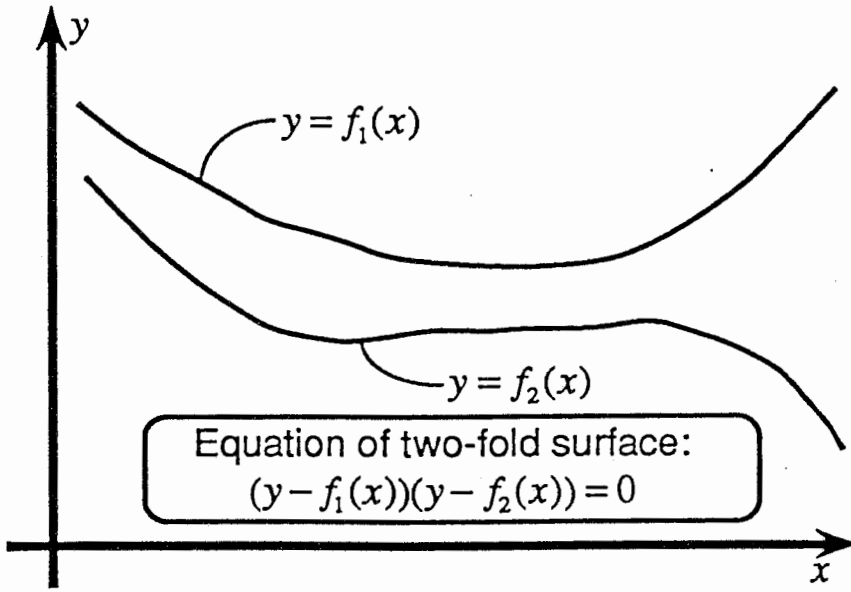


Figure 3.10: 2重曲面上のデータが満たす方程式

3.2.2.1 2重表面復元のためのエネルギー汎関数

一価関数によるデータ近似のための標準正則化問題は、次の汎関数のエネルギー最小化の形に表される [69]。

$$E^{(1)}[f] = \sum_{i=1}^N (y_{(i)} - f(x_{(i)}))^2 + \lambda \|Sf(x)\|^2 \quad (3.61)$$

S は正則化作用素、 $\lambda > 0$ は正則化パラメータと呼ばれる。 $\| \cdot \|$ は関数空間のノルムである。 $(x_{(i)}, y_{(i)})$ は、 i 番目のデータの座標と値を表し、 N は、総データ数を表す。右辺の初めの項はデータ項と呼ばれ、第2項は正則化項と呼ばれる。データ項は、近似関数とデータとの2乗誤差を表し、正則化項は、近似関数の滑らかさ拘束を表す。正則化パラメータは、これら2項の相対的な重みを決定する係数である。

式(3.61)のデータ項をみると曲面 $y = f(x)$ が、データが満たすべき拘束方程式になっている。そこで、2重表面の場合には、関数 $f_1(x)$ および $f_2(x)$ によって定義される2枚の表面 $y = f_1(x)$, $y = f_2(x)$ を考える(図3.10参照)。この2枚の表面が同時に空間 (x, y) に存在する条件を表す方程式は次式である。

$$(y - f_1(x))(y - f_2(x)) = y^2 - (f_1(x) + f_2(x))y + f_1(x)f_2(x) = 0 \quad (3.62)$$

このように、2枚の表面が1個の方程式で表される(図3.10)。この方程式は、各データ

が、曲面 $y = f_1(x)$ と $y = f_2(x)$ のうちの少なくとも一方の上に乗るための必要十分条件を数学的に厳密に表現している。実際には、データは誤差をもっているので、各データ点は、これらの曲面のどちらか一方によって近似されると考えられる。このようにすると、データが、どちらの曲面に属するかを明示することなく一個の拘束方程式で表現可能である。

この方程式の構成の仕方は、「重ね合わせの原理」(2.3)の応用とも考えられる(付録B.2.1)。

関数 $F(x)$ 、 $G(x)$ を次式で定義する。

$$\begin{aligned} F(x) &= f_1(x)f_2(x), \\ G(x) &= -(f_1(x) + f_2(x)) \end{aligned} \quad (3.63)$$

すると、データの拘束式(3.62)は、未知関数 $F(x)$ と $G(x)$ に関して線形である。もとの関数 $f_1(x)$ と $f_2(x)$ に正則化を行う代わりに関数 $F(x)$ と $G(x)$ に滑らかさの拘束を課す。すると、エネルギー汎関数はデータ項、正則化項共に2次形式となり、標準正則化理論が応用できる。変換(3.63)の逆変換は、次式で計算できる。

$$f_1(x), f_2(x) = \frac{1}{2} \left[-G(x) \pm \sqrt{\{G(x)\}^2 - 4F(x)} \right] \quad (3.64)$$

このように、これら2組の関数は、連続かつ微分可能な写像によって相互に対応付けられるので、 $f_1(x)$ と $f_2(x)$ のかわりに $F(x)$ と $G(x)$ に滑らかさの拘束を課す。ただし、 $f_1(x)$ 、 $f_2(x)$ と右辺の2個の値との対応は得られない。この不定性は、拘束式(3.62)が $f_1(x)$ と $f_2(x)$ に関して対称であるために避けることができない。そのため、以下では、条件 $f_1(x) \geq f_2(x)$ を付加して考える。変換(3.63)と逆変換(3.64)は、共に各点 x で完全に独立に閉形式で計算可能である。したがって、これにもなう計算量の増加は、正則化のための緩和計算に較べればほとんど無視できる。

以上の考察から、2重表面の復元は、次の2次エネルギー汎関数の最小化によって定義される標準正則化で実現できる。

$$\begin{aligned} E^{(2)}[F, G] &= \sum_{i=1}^N \left\{ F(x_{(i)}) + G(x_{(i)})y_{(i)} + (y_{(i)})^2 \right\}^2 \\ &\quad + \lambda_F \|S_F F(x)\|^2 + \lambda_G \|S_G G(x)\|^2 \end{aligned} \quad (3.65)$$

汎関数(3.65)の最小化によって得られる関数 $F(x)$ 、 $G(x)$ は、データの誤差の影響を受けるので、式(3.64)の平方根が実数になる保証はない。したがって、厳密には、

$$\{G(x)\}^2 - 4F(x) \geq 0 \quad (3.66)$$

という不等式制約条件が必要である。しかし、この条件を繰り返し計算の後で判定することになれば、最適化問題は単純な制約なしの2次汎関数である。

この標準正則化問題は、視覚情報処理におけるオプティカルフローの正則化 [51] と同様の、2次元ベクトル値関数 $(F(x), G(x))$ の標準正則化である。従って、従来から用いられてきた2次元ベクトル場の標準正則化についての緩和計算法やそのアナログ抵抗回路による実現がほぼそのまま有効である。

3.2.2.2 Gauss-Seidel 法による最適化アルゴリズム

本節では、式 (3.65) を $x = (x_1, x_2) \in Z^2$ の正方格子上で離散化して、Gauss-Seidel 法による超並列計算向きの繰り返しアルゴリズムを導く。2次関数の最適化アルゴリズムには、Gauss-Seidel 法以外にも多くの方法が利用可能である。ここでは、アルゴリズムの簡単さと局所並列演算による実現を考慮して、Gauss-Seidel 法を選んだ。本節の主な主張点は、特定の最適化アルゴリズムの提案ではなく、多重表面復元の一般的枠組の提案である。

正則化項には、もっとも簡単な薄膜 (membrane) モデルを用いた。薄膜モデルの正則化作用素は1階微分の $S_F = S_G = \nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2} \right)$ である。以下、変数 x_1, x_2 は格子点の整数座標値をとるものとする。するとエネルギー汎関数 (3.65) は、変数 $F_{x_1, x_2}, G_{x_1, x_2}$ に関する次のエネルギー関数で近似される。

$$\begin{aligned} \mathcal{E}^{(2)}(F, G) &= \sum_{i=1}^N \left\{ F_{x_1(i), x_2(i)} + G_{x_1(i), x_2(i)} y(i) + (y(i))^2 \right\}^2 \\ &+ \sum_{(x_1, x_2) \in Z^2} \lambda_F \left\{ (F_{x_1, x_2} - F_{x_1-1, x_2})^2 + (F_{x_1, x_2} - F_{x_1, x_2-1})^2 \right\} \\ &+ \sum_{(x_1, x_2) \in Z^2} \lambda_G \left\{ (G_{x_1, x_2} - G_{x_1-1, x_2})^2 + (G_{x_1, x_2} - G_{x_1, x_2-1})^2 \right\} \quad (3.67) \end{aligned}$$

$F_{x_1, x_2}, G_{x_1, x_2}$ は、それぞれ、関数 $F(x), G(x)$ の格子点 $(x_1, x_2) \in Z^2$ における値である。 $\{(x_1(i), x_2(i)), y(i)\} (i = 1, \dots, N)$ は、与えられたスパースなデータ点とその値である。このエネルギー関数の最小化問題を解くため、変数 $F_{x_1, x_2}, G_{x_1, x_2}$ に関する偏微分を零とおいて、次の平衡条件 (離散近似の Euler-Lagrange 方程式) を得る。

$$\begin{aligned} \frac{\partial \mathcal{E}^{(2)}(F, G)}{\partial F_{x_1, x_2}} &= 2 \left[\sum_{i=1}^N \delta_{x_1, x_1(i)} \delta_{x_2, x_2(i)} \left\{ F_{x_1(i), x_2(i)} + G_{x_1(i), x_2(i)} y(i) + (y(i))^2 \right\} \right] \\ &+ 2\lambda_F (F_{x_1, x_2} - \bar{F}_{x_1, x_2}) = 0, \\ \frac{\partial \mathcal{E}^{(2)}(F, G)}{\partial G_{x_1, x_2}} &= 2 \left[\sum_{i=1}^N \delta_{x_1, x_1(i)} \delta_{x_2, x_2(i)} \left\{ F_{x_1(i), x_2(i)} y(i) + G_{x_1(i), x_2(i)} (y(i))^2 + (y(i))^3 \right\} \right] \\ &+ 2\lambda_G (G_{x_1, x_2} - \bar{G}_{x_1, x_2}) = 0 \quad (3.68) \end{aligned}$$

$\delta_{a,b}$ は、Kronecker のデルタであり、 $a = b$ の時に 1 をとり、それ以外の場合に、0 をとる。 \bar{F}_{x_1, x_2} と \bar{G}_{x_1, x_2} は、それぞれ4近傍の値の平均であり、次式で定義される。

$$\bar{F}_{x_1, x_2} = \frac{1}{4} (F_{x_1+1, x_2} + F_{x_1-1, x_2} + F_{x_1, x_2+1} + F_{x_1, x_2-1}),$$

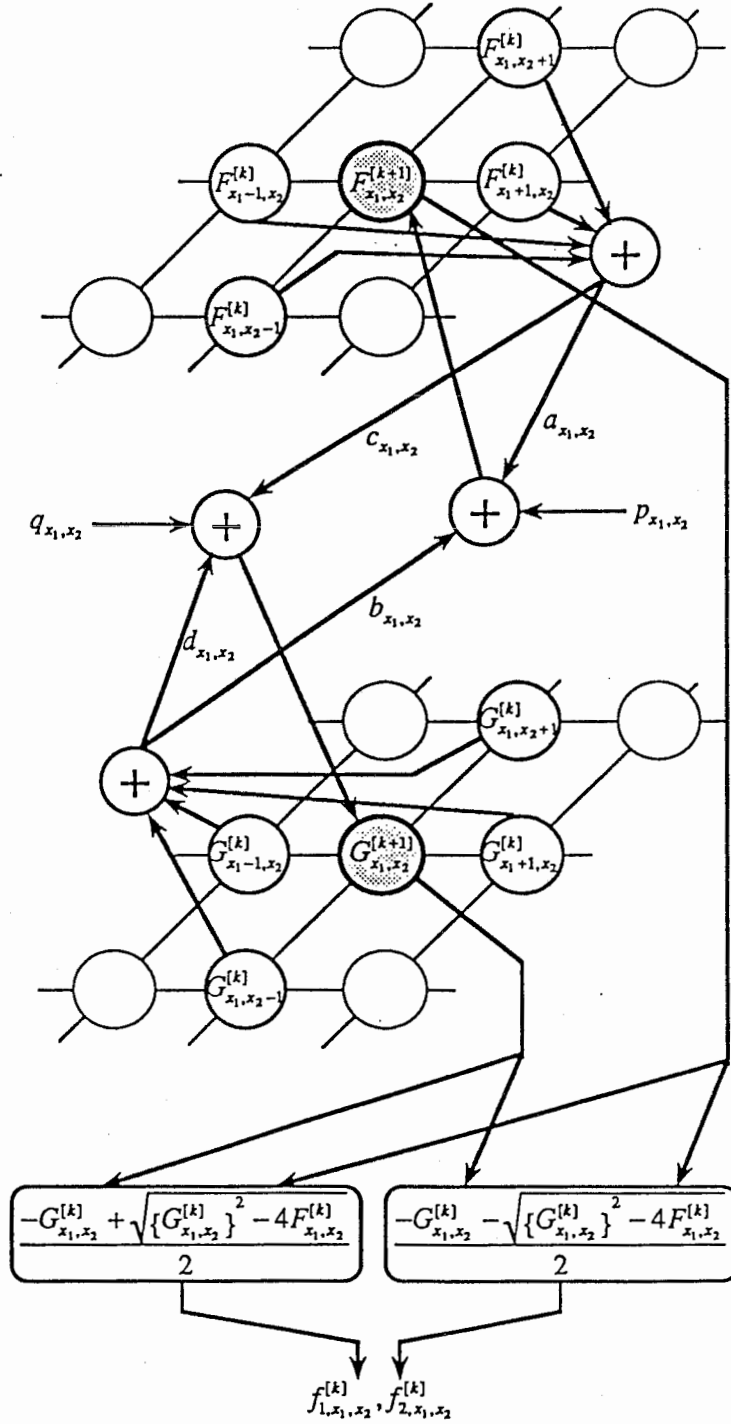


Figure 3.11: 2重表面復元の緩和計算を行う超並列型ネットワーク

$$\bar{G}_{x_1, x_2} = \frac{1}{4} (G_{x_1+1, x_2} + G_{x_1-1, x_2} + G_{x_1, x_2+1} + G_{x_1, x_2-1}) \quad (3.69)$$

データだけから計算できる次の量を定義する。

$$\begin{aligned} A_{x_1, x_2} &= \sum_{i=1}^N \delta_{x_1, x_1(i)} \delta_{x_2, x_2(i)}, \\ B_{x_1, x_2} &= \sum_{i=1}^N \delta_{x_1, x_1(i)} \delta_{x_2, x_2(i)} y(i), \\ C_{x_1, x_2} &= \sum_{i=1}^N \delta_{x_1, x_1(i)} \delta_{x_2, x_2(i)} (y(i))^2, \\ D_{x_1, x_2} &= \sum_{i=1}^N \delta_{x_1, x_1(i)} \delta_{x_2, x_2(i)} (y(i))^3 \end{aligned} \quad (3.70)$$

ここで定義した $A_{x_1, x_2}, B_{x_1, x_2}, C_{x_1, x_2}, D_{x_1, x_2}$ は、データが存在しない格子点 (x_1, x_2) では、すべて零である。この定式化では、同一点 x に複数のデータが与えられることも許される。したがって、入力データ自身が多価であることも当然許される。このとき、平衡条件 (3.68) は、次式に簡単化される。

$$\begin{aligned} (A_{x_1, x_2} + \lambda_F) F_{x_1, x_2} + B_{x_1, x_2} G_{x_1, x_2} &= \lambda_F \bar{F}_{x_1, x_2} - C_{x_1, x_2}, \\ B_{x_1, x_2} F_{x_1, x_2} + (C_{x_1, x_2} + \lambda_G) G_{x_1, x_2} &= \lambda_G \bar{G}_{x_1, x_2} - D_{x_1, x_2} \end{aligned} \quad (3.71)$$

この連立方程式を F_{x_1, x_2} と G_{x_1, x_2} について解くと、次式を得る。

$$\begin{aligned} F_{x_1, x_2} &= \frac{\lambda_F (C_{x_1, x_2} + \lambda_G) \bar{F}_{x_1, x_2} - \lambda_G B_{x_1, x_2} \bar{G}_{x_1, x_2} + B_{x_1, x_2} D_{x_1, x_2} - C_{x_1, x_2} (C_{x_1, x_2} + \lambda_G)}{(A_{x_1, x_2} + \lambda_F) (C_{x_1, x_2} + \lambda_G) - B_{x_1, x_2}^2}, \\ G_{x_1, x_2} &= \frac{-\lambda_F B_{x_1, x_2} \bar{F}_{x_1, x_2} + \lambda_G (A_{x_1, x_2} + \lambda_F) \bar{G}_{x_1, x_2} + B_{x_1, x_2} C_{x_1, x_2} - D_{x_1, x_2} (A_{x_1, x_2} + \lambda_F)}{(A_{x_1, x_2} + \lambda_F) (C_{x_1, x_2} + \lambda_G) - B_{x_1, x_2}^2} \end{aligned} \quad (3.72)$$

これから、Gauss-Seidel法による次の緩和アルゴリズムが導かれる。

$$\begin{aligned} F_{x_1, x_2}^{[k+1]} &= a_{x_1, x_2} \bar{F}_{x_1, x_2}^{[k]} + b_{x_1, x_2} \bar{G}_{x_1, x_2}^{[k]} + p_{x_1, x_2}, \\ G_{x_1, x_2}^{[k+1]} &= c_{x_1, x_2} \bar{F}_{x_1, x_2}^{[k]} + d_{x_1, x_2} \bar{G}_{x_1, x_2}^{[k]} + q_{x_1, x_2} \end{aligned} \quad (3.73)$$

k は、繰り返しの回数を表すインデクスである。 $a_{x_1, x_2}, b_{x_1, x_2}, c_{x_1, x_2}, d_{x_1, x_2}, p_{x_1, x_2}, q_{x_1, x_2}$ は、あらかじめ各点で独立に計算できる量であり、次式で定義される。

$$\begin{aligned} a_{x_1, x_2} &= \lambda_F (C_{x_1, x_2} + \lambda_G) / U_{x_1, x_2}, \\ b_{x_1, x_2} &= -\lambda_G B_{x_1, x_2} / U_{x_1, x_2}, \\ c_{x_1, x_2} &= -\lambda_F B_{x_1, x_2} / U_{x_1, x_2}, \end{aligned}$$

$$\begin{aligned}
d_{x_1, x_2} &= \lambda_G (A_{x_1, x_2} + \lambda_F) / U_{x_1, x_2}, \\
p_{x_1, x_2} &= \{B_{x_1, x_2} D_{x_1, x_2} - C_{x_1, x_2} (C_{x_1, x_2} + \lambda_G)\} / U_{x_1, x_2}, \\
q_{x_1, x_2} &= \{B_{x_1, x_2} C_{x_1, x_2} - D_{x_1, x_2} (A_{x_1, x_2} + \lambda_F)\} / U_{x_1, x_2}
\end{aligned} \tag{3.74}$$

ここで、 $U_{x_1, x_2} = (A_{x_1, x_2} + \lambda_F)(C_{x_1, x_2} + \lambda_G) - B_{x_1, x_2}^2$ とおいた。特に、データの存在しない点 (x_1, x_2) では、 $a_{x_1, x_2} = d_{x_1, x_2} = 1, b_{x_1, x_2} = c_{x_1, x_2} = p_{x_1, x_2} = q_{x_1, x_2} = 0$ であり、単なる充填過程である。このアルゴリズムは、図 3.11 に示した 2 重網状の超並列型ネットワークで実現できる。

3.2.2.3 最急降下法による最適化アルゴリズム

エネルギー汎関数 (3.65) の最小化問題の解は、次式で表される線形ダイナミカルシステムの平衡点を与える関数 $F(x), G(x)$ として計算することも可能である。

$$\frac{dF}{dt} = -\omega \frac{\delta E^{(2)}[F, G]}{\delta F}, \quad \frac{dG}{dt} = -\omega \frac{\delta E^{(2)}[F, G]}{\delta G} \tag{3.75}$$

記号 δ は変分、 ω は正定数を表す。これを離散化すると、最急降下法による緩和アルゴリズムが導かれる。

$$\begin{aligned}
F_{x_1, x_2}^{[k+1]} &= F_{x_1, x_2}^{[k]} - 2\omega \left\{ (A_{x_1, x_2} + \lambda_F) F_{x_1, x_2}^{[k]} + B_{x_1, x_2} G_{x_1, x_2}^{[k]} + C_{x_1, x_2} - \lambda_F \bar{F}_{x_1, x_2}^{[k]} \right\}, \\
G_{x_1, x_2}^{[k+1]} &= G_{x_1, x_2}^{[k]} - 2\omega \left\{ B_{x_1, x_2} F_{x_1, x_2}^{[k]} + (C_{x_1, x_2} + \lambda_G) G_{x_1, x_2}^{[k]} - D_{x_1, x_2} - \lambda_G \bar{G}_{x_1, x_2}^{[k]} \right\}
\end{aligned} \tag{3.76}$$

この方法は、収束が遅いのでデジタル計算には向かない。しかし、アナログ抵抗回路網による正則化 [134] の計算機シミュレーションとしての意味がある。

3.2.2.4 シミュレーション実験と結果

超並列計算機 CM-2 を用いて 2 重表面復元のシミュレーション実験を行った。x の空間を 512×512 の 2 次元格子で近似し、Gauss-Seidel 法による緩和計算を実現した。5 通りの実行結果を以下に示す。

実験 1 (図 3.12) では、図 3.12(a) に示した、2 枚の平面 $y = f_1(x) = 0.8$ 、 $y = f_2(x) = -0.8$ に、標準偏差 $\sigma = 0.05$ のガウシアンノイズを重畳し、 $f_1(x)$ のデータをくし状に取り除いたテストデータを用いた。データ数は、 $f_1(x)$ が約 6000 点、 $f_2(x)$ が 10000 点である。

実験 2 (図 3.13) では、図 3.13(a) に示した、2 枚の平面 $y = f_1(x) = 0.8$ 、 $y = f_2(x) = -0.8$ に、標準偏差 $\sigma = 0.05$ のガウシアンノイズを重畳し、 $f_1(x)$ と $f_2(x)$ のデー

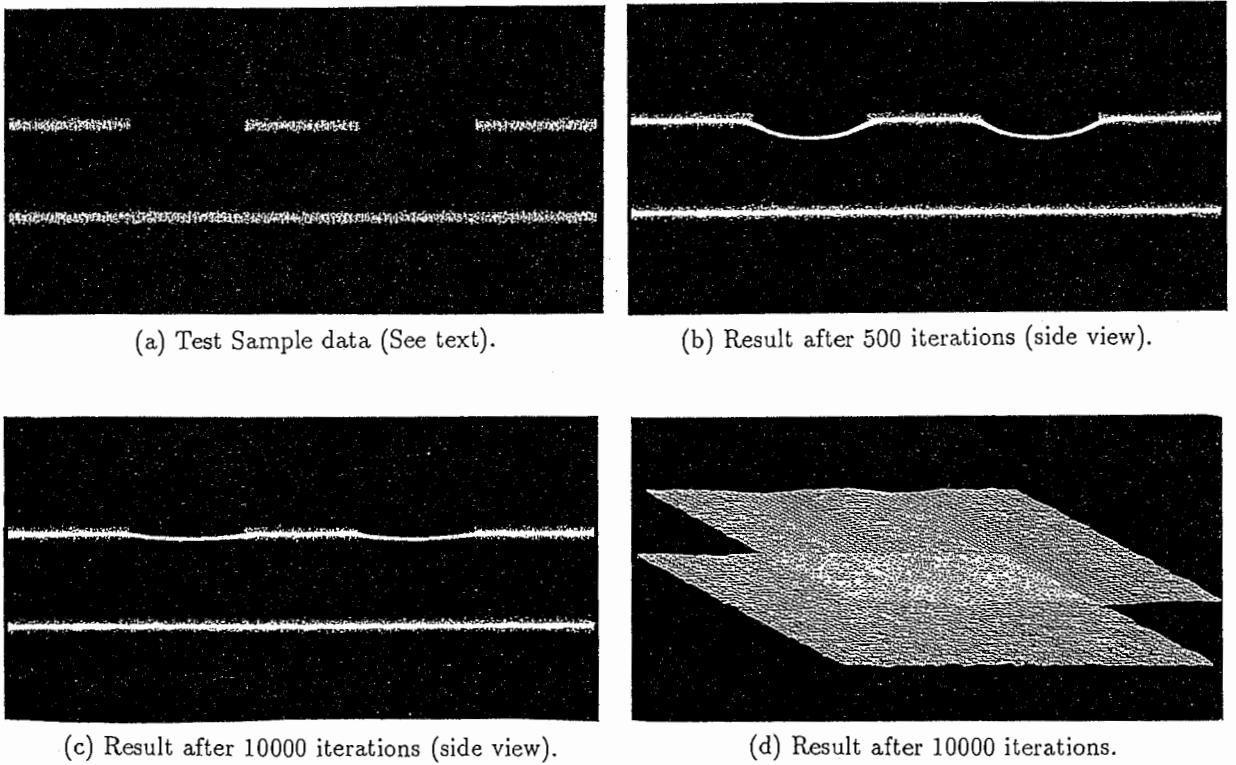
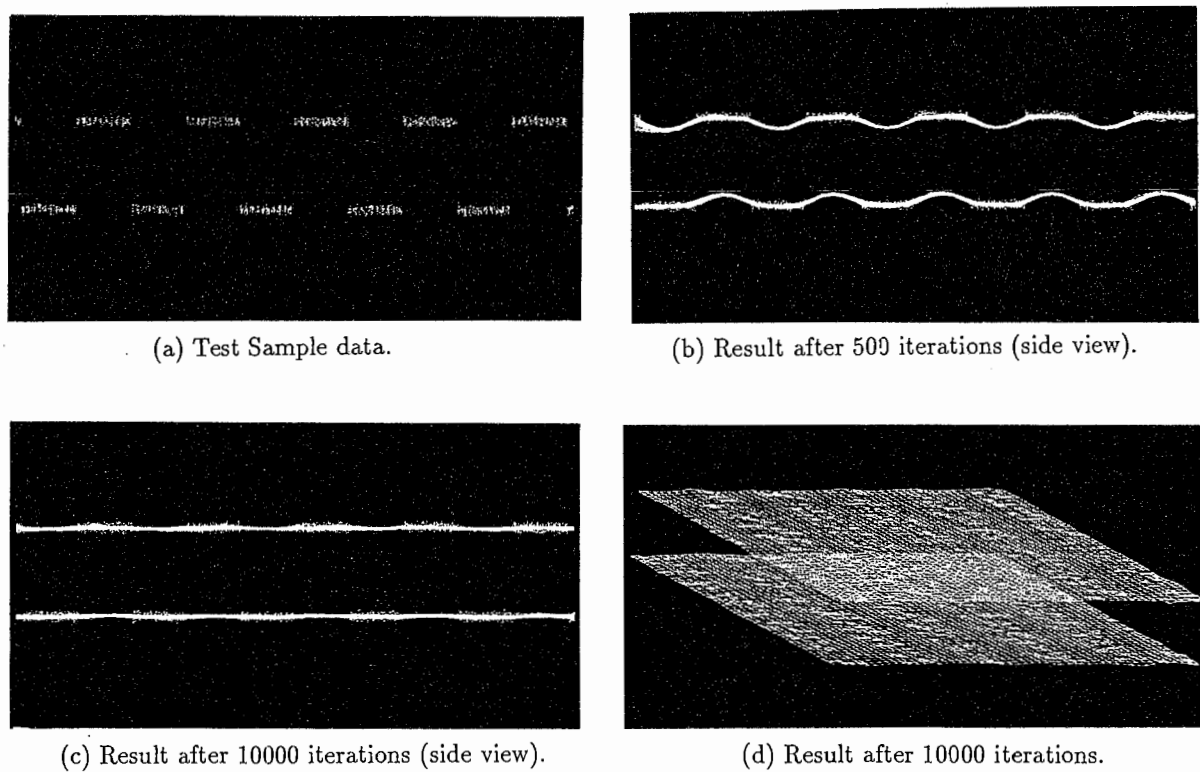


Figure 3.12: 実験 1 の結果 ($\lambda_F = \lambda_G = 6.0$)

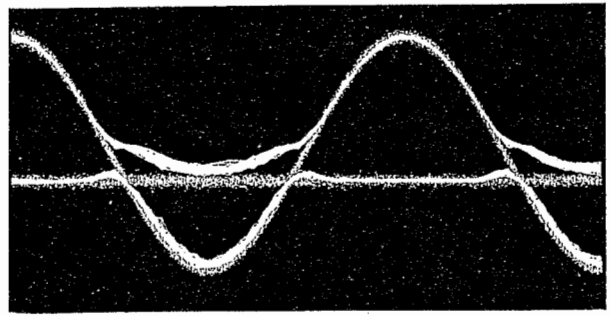
(a) テストデータ、(b) 500 回繰り返し後の結果 (側面図)、(c) 10,000 回繰り返し後の結果 (側面図)、(d) 10,000 回繰り返し後の結果。

Figure 3.13: 実験2の結果 ($\lambda_F = \lambda_G = 6.0$)

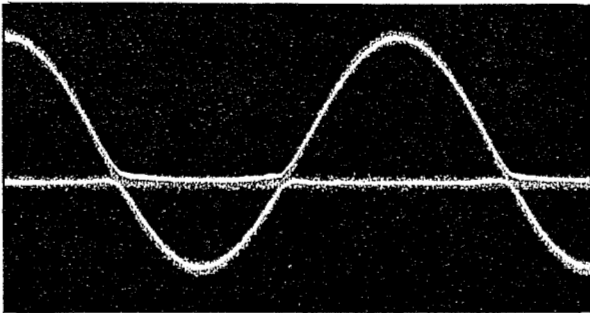
(a) テストデータ、(b) 500 回繰り返し後の結果(側面図)、(c) 10,000 回繰り返し後の結果(側面図)、(d) 10,000 回繰り返し後の結果。



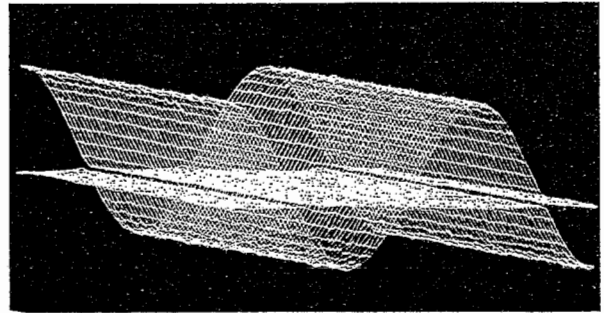
(a) Test Sample data (See text).



(b) Result after 500 iterations (side view).



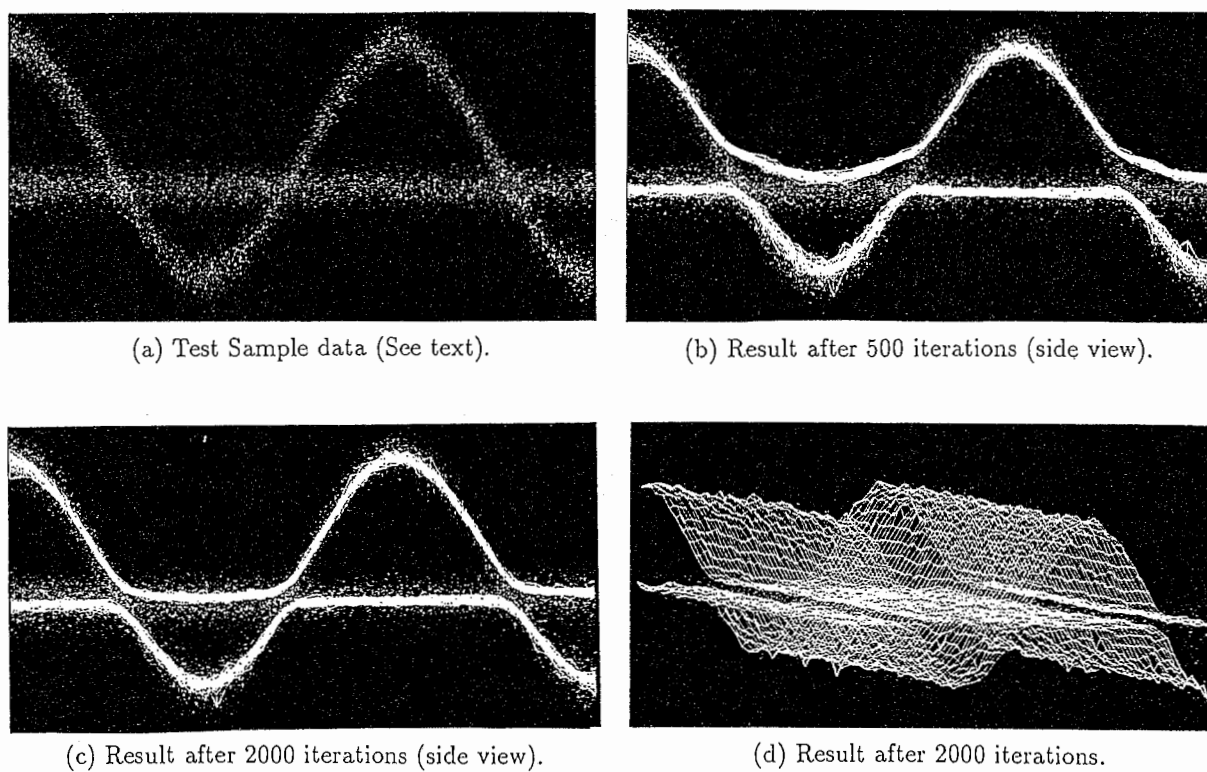
(c) Result after 2000 iterations (side view).



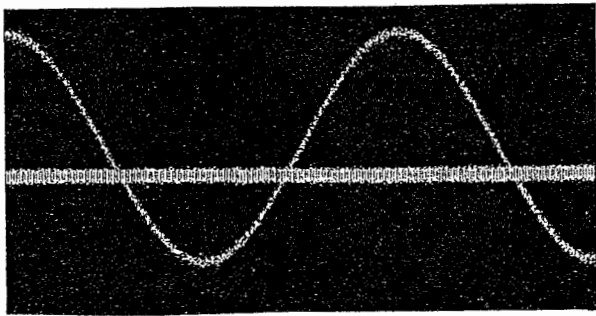
(d) Result after 2000 iterations.

Figure 3.14: 実験3の結果 ($\lambda_F = \lambda_G = 12.0$)

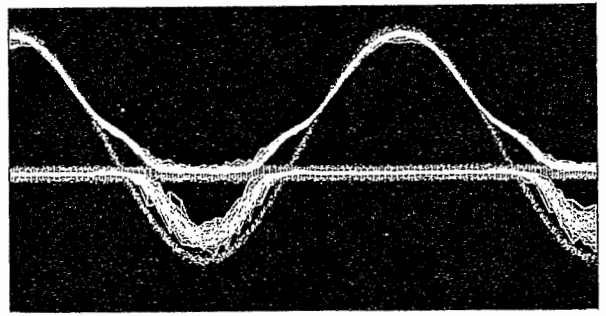
(a) テストデータ、(b) 500回繰り返し後の結果(側面図)、(c) 2,000回繰り返し後の結果(側面図)、(d) 2,000回繰り返し後の結果。

Figure 3.15: 実験4の結果 ($\lambda_F = \lambda_G = 12.0$)

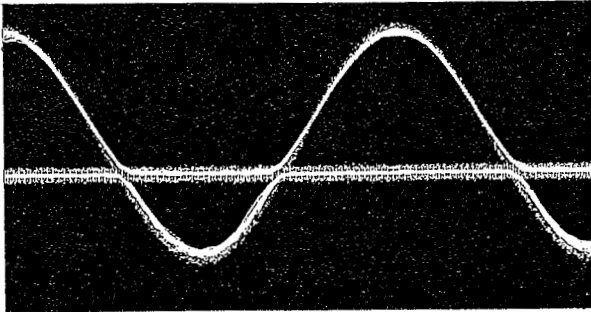
(a) テストデータ、(b) 500 回繰り返し後の結果 (側面図)、(c) 2,000 回繰り返し後の結果 (側面図)、(d) 2,000 回繰り返し後の結果。



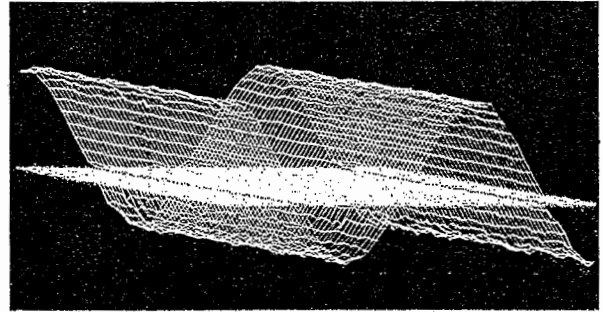
(a) Test Sample data (See text).



(b) Result after 500 iterations (side view).



(c) Result after 2000 iterations (side view).



(d) Result after 2000 iterations.

Figure 3.16: 実験5の結果 ($\lambda_F = \lambda_G = 12.0$)

(a) テストデータ、(b) 500回繰り返し後の結果(側面図)、(c) 2,000回繰り返し後の結果(側面図)、(d) 2,000回繰り返し後の結果。

タを交互にくし状に取り除いたテストデータを用いた。データ数は、 $f_1(x)$ 、 $f_2(x)$ 共に、約 5000 点である。

実験 3、4、5 (図 3.14、図 3.15、図 3.16) では、テストデータの生成に次の 2 個の式を用いた。

$$\begin{aligned} f_1(x) &: -0.5 + N(0, \sigma^2), \\ f_2(x) &: 2 \sin 3\pi \left(\frac{x_2 - 256}{512} \right) + N(0, \sigma^2) \end{aligned} \quad (3.77)$$

このように、定数関数と正弦関数からなる 2 価関数にそれぞれガウシアンノイズ $N(0, \sigma^2)$ を加えて生成した。 $N(m, \sigma^2)$ は、平均 m 、分散 σ^2 のガウシアンノイズを表す。関数 $f_1(x)$ 、 $f_2(x)$ に関して、それぞれ、 N_1 、 N_2 個の座標 x を乱数で生成し、ノイズを加えた関数値を計算して、合計 $N = N_1 + N_2$ 個のデータ $\{x_{(i)}, y_{(i)}\}$ ($i = 1, 2, \dots, N$) を生成した。

実験に用いたパラメータは次の通りである。

- 実験 3 : $N_1 = 10000$ 、 $N_2 = 20000$ 、 $\sigma = 0.05$
- 実験 4 : $N_1 = 10000$ 、 $N_2 = 10000$ 、 $\sigma = 2.0$
- 実験 5 : $N_1 = 10000$ 、 $N_2 = 1000$ 、 $\sigma = 0.05$

なお、初期値は、すべての実験で $F_{x_1, x_2}^{[0]} = G_{x_1, x_2}^{[0]} = 0$ とした。各実験に対して、平衡点へ達する前の過渡的状态 (各図 (b)) と、平衡点に達した後の結果 (各図 (c), (d)) を示した。

- 実験 1 では、上の面のデータが完全に欠けている部分でも自然に表面の補間が行われていることが注目される。
- 実験 2 では、不連続点を持つ関数も、計算上の多重表面として復元できることが示された。
- 実験 3 では、交点のところでは、曲面同士が近接するものの、接触せず、全体が上下に分割された曲面として表現された。
- 実験 4 では、ノイズをかなり大きくしても破綻せず、データの分布を正しく反映した結果が出ることを示された。
- 実験 5 では、データの密度が、2 個の関数の間でかなり異なる場合 (この例では 10 倍) でも、正しい表面が復元できることを示された。

3.2.3 多価標準正則化理論の一般論

本節では、多価標準正則化理論の一般論を述べる。まず、スカラー値 ($m = 1$) h 価関数に拡張した場合の一般論を述べ、次に、テンソル積を用いた m 次元ベクトル値 h 価関数の場合の一般論について、 $m = 2$ 、 $h = 2$ の場合を例にとり展開する。

3.2.3.1 スカラー値 h 価関数の場合

式 (3.62) と同様にして、 h 枚の曲面 $y = f_i(\mathbf{x})$ ($i = 1, 2, \dots, h$) を同時に表現する方程式は、次式である。

$$\prod_{i=1}^h (y - f_i(\mathbf{x})) = y^h - \pi_1^{(h)}(\mathbf{x})y^{h-1} + \pi_2^{(h)}(\mathbf{x})y^{h-2} - \dots + (-1)^h \pi_h^{(h)}(\mathbf{x}) = 0 \quad (3.78)$$

$\pi_k^{(h)}(\mathbf{x})$ は、 h 個の関数 $f_i(\mathbf{x})$ ($i = 1, \dots, h$) の基本対称式として次式で定義される関数である。

$$\begin{aligned} \pi_1^{(h)}(\mathbf{x}) &= \sum_{i=1}^h f_i(\mathbf{x}), \\ \pi_2^{(h)}(\mathbf{x}) &= \sum_{i_1=1}^h \sum_{i_2=i_1+1}^h f_{i_1}(\mathbf{x}) f_{i_2}(\mathbf{x}), \\ &\quad \dots, \\ \pi_h^{(h)}(\mathbf{x}) &= f_1(\mathbf{x}) f_2(\mathbf{x}) \cdots f_h(\mathbf{x}) \end{aligned} \quad (3.79)$$

2 価関数の場合と同様、 $f_i(\mathbf{x})$ の代わりに、 $\pi_k^{(h)}(\mathbf{x})$ を未知関数とした正則化問題を考える。このとき、 $\pi_k^{(h)}(\mathbf{x})$ から $f_i(\mathbf{x})$ への逆変換は、各点 \mathbf{x} において次に示す ϕ に関する h 次代数方程式の解として定義される。

$$\Lambda^{(h)}(\phi) = \phi^h - \pi_1^{(h)}(\mathbf{x})\phi^{h-1} + \pi_2^{(h)}(\mathbf{x})\phi^{h-2} - \dots + (-1)^h \pi_h^{(h)}(\mathbf{x}) = 0 \quad (3.80)$$

各点 \mathbf{x} において求まる h 個の解 ϕ_j ($j = 1, 2, \dots, h$) が、 $f_i(\mathbf{x})$ のどれに対応するかは原理的に決まらない。エネルギー汎関数は、次式で定義する。

$$\begin{aligned} E^{(h)}[\pi_1^{(h)}, \pi_2^{(h)}, \dots, \pi_h^{(h)}] &= \\ &= \sum_{i=1}^N \{y_{(i)}^h - \pi_1^{(h)}(\mathbf{x}_{(i)})y_{(i)}^{h-1} + \pi_2^{(h)}(\mathbf{x}_{(i)})y_{(i)}^{h-2} - \dots + (-1)^h \pi_h^{(h)}(\mathbf{x}_{(i)})\}^2 \\ &+ \sum_{k=1}^h \lambda_k \|S_k^{(h)} \pi_k^{(h)}(\mathbf{x})\|^2 \end{aligned} \quad (3.81)$$

この汎関数は、データ項、正則化項ともに未知関数 $\pi_k^{(h)}(\mathbf{x})$ に関して 2 次である。したがって、標準正則化理論の枠内で最適化問題の解が求められる。

この方法では、最後に h 次代数方程式を解かなければならない。しかし、5 次以上の代数方程式には解析解が存在しない。この場合は、Durand-Kerner 法として知られる次の反復法で、すべての解を同時に求められる。

$$\phi_j^{[k+1]} = \phi_j^{[k]} - \frac{\Lambda(\phi_j^{[k]})}{\prod_{i=1, i \neq j}^h (\phi_j^{[k]} - \phi_i^{[k]})} \quad (3.82)$$

これは、基本対称式の定義式 (3.79) に ϕ_j を代入したものを連立代数方程式とみて、それに Newton-Raphson 法を適用したものと等価である。2 位の収束をすることが知られている。この反復法は、フィードバック結合をもつネットワークで実現できる。

3.2.3.2 m 次元ベクトル値 h 価関数の場合: テンソル積表現

ベクトル値関数の場合には、各成分について独立にスカラー値の多価標準正則化理論を適用すればよい様に思われるが、これは正しくない。なぜならば、各成分の正則化の結果得られる複数のスカラー値同士を対応付けることができないからである。

この場合には、テンソル積(クロネッカー積)を用いて、正しく扱うことができる。 n 次元空間の 1 価 m 次元ベクトル値関数を $\mathbf{y} = \mathbf{f}(\mathbf{x})$ ($\mathbf{y} \in R^m, \mathbf{x} \in R^n$) とする。 h 価の場合、それぞれの値に対応する関数を $\mathbf{f}_k(\mathbf{x})$ ($k = 1, 2, \dots, h$) とすると、

$$(\mathbf{y} - \mathbf{f}_1(\mathbf{x})) \otimes (\mathbf{y} - \mathbf{f}_2(\mathbf{x})) \otimes \dots \otimes (\mathbf{y} - \mathbf{f}_h(\mathbf{x})) = \mathbf{0} \quad (3.83)$$

が正しい方程式である。前節までで述べてきた場合は、すべてこの方程式の特殊な場合である。この方程式の成分数は m^h 個である。しかし、方程式の順番を入れ換えても式の意味は同じでなければならない。つまり、この表現は冗長である。そこで、このテンソル積に対称化 [165] [173] の操作を施す。すると、成分方程式の数は、 ${}_m H_h$ 個 (m 個の集合から h 個取り出す重複組合せの数) である。それでも、依然、本来の自由度である mh よりは大きい。その代わり、表現は一意である。

$h = 2, m = 2$ の場合の例 この場合、基本方程式は、

$$\frac{1}{2} \{(\mathbf{y} - \mathbf{f}_1(\mathbf{x})) \otimes (\mathbf{y} - \mathbf{f}_2(\mathbf{x})) + (\mathbf{y} - \mathbf{f}_2(\mathbf{x})) \otimes (\mathbf{y} - \mathbf{f}_1(\mathbf{x}))\} = \mathbf{0} \quad (3.84)$$

である。2 価 2 次元ベクトル値関数を 2 個の 2 次元ベクトル値関数、

$$\begin{aligned} \mathbf{y} &= \mathbf{f}_1(\mathbf{x}) = (f_{1,1}(\mathbf{x}), f_{1,2}(\mathbf{x}))^T, \\ \mathbf{y} &= \mathbf{f}_2(\mathbf{x}) = (f_{2,1}(\mathbf{x}), f_{2,2}(\mathbf{x}))^T \end{aligned} \quad (3.85)$$

を用いて表現する。次の3個の方程式が導かれる。

$$\begin{aligned}
 y_1^2 - (f_{1,1}(\mathbf{x}) + f_{2,1}(\mathbf{x}))y_1 + f_{1,1}(\mathbf{x})f_{2,1}(\mathbf{x}) &= 0, \\
 y_2^2 - (f_{1,2}(\mathbf{x}) + f_{2,2}(\mathbf{x}))y_2 + f_{1,2}(\mathbf{x})f_{2,2}(\mathbf{x}) &= 0, \\
 2y_1y_2 - (f_{1,2}(\mathbf{x}) + f_{2,2}(\mathbf{x}))y_1 - (f_{1,1}(\mathbf{x}) + f_{2,1}(\mathbf{x}))y_2 \\
 + (f_{1,1}(\mathbf{x})f_{2,2}(\mathbf{x}) + f_{1,2}(\mathbf{x})f_{2,1}(\mathbf{x})) &= 0
 \end{aligned} \tag{3.86}$$

4個のスカラー値関数 $f_{1,1}(\mathbf{x}), f_{1,2}(\mathbf{x}), f_{2,1}(\mathbf{x}), f_{2,2}(\mathbf{x})$ を次の5個のスカラー値関数に置き換える。

$$\begin{aligned}
 F_1(\mathbf{x}) &= f_{1,1}(\mathbf{x})f_{2,1}(\mathbf{x}) \\
 F_2(\mathbf{x}) &= f_{1,2}(\mathbf{x})f_{2,2}(\mathbf{x}) \\
 G_1(\mathbf{x}) &= -(f_{1,1}(\mathbf{x}) + f_{2,1}(\mathbf{x})) \\
 G_2(\mathbf{x}) &= -(f_{1,2}(\mathbf{x}) + f_{2,2}(\mathbf{x})) \\
 H(\mathbf{x}) &= f_{1,1}(\mathbf{x})f_{2,2}(\mathbf{x}) + f_{1,2}(\mathbf{x})f_{2,1}(\mathbf{x})
 \end{aligned} \tag{3.87}$$

すると、3個の方程式 (3.86) は、関数 (3.87) に関してすべて線形になる。

$$\begin{aligned}
 y_1^2 + G_1(\mathbf{x})y_1 + F_1(\mathbf{x}) &= 0, \\
 y_2^2 + G_2(\mathbf{x})y_2 + F_2(\mathbf{x}) &= 0, \\
 2y_1y_2 + G_2(\mathbf{x})y_1 + G_1(\mathbf{x})y_2 + H(\mathbf{x}) &= 0
 \end{aligned} \tag{3.88}$$

F_1, F_2, G_1, G_2, H を未知関数とみなして正則化理論を適用する。すると、次のエネルギー汎関数の最小化問題に帰着される。

$$\begin{aligned}
 E^{(2,2)}[F_1, F_2, G_1, G_2, H] = \\
 \sum_{i=1}^N \left\{ (y_{1(i)})^2 + G_1(\mathbf{x}(i))y_{1(i)} + F_1(\mathbf{x}(i)) \right\}^2 + \sum_{i=1}^N \left\{ (y_{2(i)})^2 + G_2(\mathbf{x}(i))y_{2(i)} + F_2(\mathbf{x}(i)) \right\}^2 \\
 + \frac{1}{2} \sum_{i=1}^N \left\{ 2y_{1(i)}y_{2(i)} + G_2(\mathbf{x}(i))y_{1(i)} + G_1(\mathbf{x}(i))y_{2(i)} + H(\mathbf{x}(i)) \right\}^2 \\
 + \lambda_{F_1} \|S_{F_1} F_1\|^2 + \lambda_{F_2} \|S_{F_2} F_2\|^2 + \lambda_{G_1} \|S_{G_1} G_1\|^2 + \lambda_{G_2} \|S_{G_2} G_2\|^2 + \lambda_H \|S_H H\|^2
 \end{aligned} \tag{3.89}$$

$S_{F_1}, S_{F_2}, S_{G_1}, S_{G_2}, S_H$ は、正則化作用素、 $\lambda_{F_1}, \lambda_{F_2}, \lambda_{G_1}, \lambda_{G_2}, \lambda_H$ は正定数で、それぞれの関数の滑らかさの程度を与える。この汎関数は、未知関数に関して2次であるから、標準正則化理論の枠内で解くことができる。

変換式 (3.87) の逆変換は、例えば、以下のように実現できる。まず、次の関数を計算する。

$$f_{+,1}(\mathbf{x}) = \frac{1}{2} \left[-G_1(\mathbf{x}) + \sqrt{\{G_1(\mathbf{x})\}^2 - 4F_1(\mathbf{x})} \right] \tag{3.90}$$

$$f_{-,1}(x) = \frac{1}{2} \left[-G_1(x) - \sqrt{\{G_1(x)\}^2 - 4F_1(x)} \right] \quad (3.91)$$

$$f_{+,2}(x) = \frac{1}{2} \left[-G_2(x) + \sqrt{\{G_2(x)\}^2 - 4F_2(x)} \right] \quad (3.92)$$

$$f_{-,2}(x) = \frac{1}{2} \left[-G_2(x) - \sqrt{\{G_2(x)\}^2 - 4F_2(x)} \right] \quad (3.93)$$

2組の解、

$$\begin{aligned} \mathbf{f}_1(x) &= (f_{+,1}(x), f_{+,2}(x))^T \\ \mathbf{f}_2(x) &= (f_{-,1}(x), f_{-,2}(x))^T \end{aligned} \quad (3.94)$$

と

$$\begin{aligned} \mathbf{f}_1(x) &= (f_{+,1}(x), f_{-,2}(x))^T \\ \mathbf{f}_2(x) &= (f_{-,1}(x), f_{+,2}(x))^T \end{aligned} \quad (3.95)$$

が得られる。各点 x において、これらのうちから、

$$|H(x) - f_{1,1}(x)f_{2,2}(x) - f_{1,2}(x)f_{2,1}(x)| \quad (3.96)$$

が小さい方を選ぶ。

3.2.4 まとめ

従来の標準正則化理論を拡張し、多価関数を扱える多価標準正則化理論を提案した。正則化理論の1つの実現方法である、超並列計算あるいはアナログ回路網のデジタルシミュレーションのための緩和アルゴリズムを導出し、シミュレーション実験を行った。

本理論は、データに対するクラスタリングを行うことなく、複数の曲面を直接数学的にテンソル積を用いて表現した。そのため、複数曲面を一つの多様体として扱っている。これは、クラスタリングが不要であるという利点がある反面、最後に分解演算という付加処理が必要になる。

以下に、今後の課題をあげる。

1. 多価度 h をあらかじめ知ることが出来ないので、 h を1つずつ増やして、残差が充分小さくなる最小の表面数を決定するアルゴリズムを開発する必要がある。
2. データのクラスタリングを行わない代償として、データの誤差の影響が複数の曲面間で相互作用するので、正確な誤差モデルをたてにくい。推定された関数値を用いて、データを一撃的にクラスタリングし、それぞれに別の単一表面を当てはめる形への発展も考えられる。
3. エネルギー汎関数が2次形式になるように未知関数の置き換えを行う線形化手法を用いている。そのため、曲面とデータとの間の距離の定義が単純でない。今後、この線形化手法の得失を見極める必要がある。
4. 曲面同士が交差する場合を扱う方法を確立する必要がある。

第 4 章

物体運動の効率的計算機表現のための可逆座標変換法

あらまし 物体の形状を表す幾何情報を、デジタル計算機で効率的に運動させるための座標変換の理論を提案する。平行移動、回転、拡大縮小、透視変換などの座標変換は、3次元幾何情報を計算機で扱うために必要な基本演算であり、すでに無数の応用例が存在する。幾何要素の座標表現法は、デカルト座標（非同次座標）と同次座標が標準的である。この2つの座標表現に対して、変換前と変換後の状態が厳密に1対1に対応する離散近似演算を実現する基本理論を提案する。本理論によれば、演算が有限語長の整数近似演算であるにもかかわらず、厳密な意味で可逆性を持つ近似写像を実現する座標変換アルゴリズムが得られる。

このような可逆性という特長を持つ一般的な座標変換の手法は、従来は、その存在すら全く考えられていなかった。本理論は、基本理論なので、いろいろな利用方法、実現方法が可能である。例えば、ソフトウェアによってアルゴリズムを実現した場合、座標変換前の情報を記憶しておく必要がないという特長がある。これは、通常の計算機プログラミングでは常識であった変換前の情報の保持を不要にする。これによって、たとえば記憶領域の管理を簡略化できる^a。

^a本章の主要部分は、文献 [176]、[184] で発表された。

4.0 背景

計算機内部で、物体の3次元幾何形状とその運動を記述し、座標変換を実行することは、コンピュータビジョン、ロボティクス、CAD、グラフィクスにおける基本演算である。この演算は、いろいろな物理量に対する変換操作を頻繁に行なう必要がある自然現象のシミュレーションにおいても用いられる。コンピュータビジョンでは、物体の幾何情報を計算機内部で運動させ、幾何学的な推論をしたり、物体の運動やカメラの運動による変換を物体データに反映する時に、これらのデータに対する高速な座標変換が必要になる。計算機の能力の発展によって、大量の形状データを保持し、高速に座標変換を行なうことができるハードウェアが可能になってきた。

デジタル計算機では、演算の基本原則がデジタル論理演算に基づいている。そのため、変数は離散値しかとれない。有限固定長の数値表現を使う限り、変換に計算誤差が伴い、原データの持つ情報の一部が失われる。たとえ多値論理を用いたとしても、座標値の離散性から生じる計算誤差は避けられない。アナログ計算機においても、必ずノイズが存在するので同様の原理的問題がある。必要なだけ演算桁数を変化させる任意可変多倍長演算を用いたとしても、無理数を厳密に表現することはできない¹。そのため、施す座標変換が数学的には正則であっても、計算機上のインプリメンテーションにおいて一般には可逆でない²。これは、幾何情報を扱う計算機プログラミングにおいては常識である。

ところが、有限精度の計算であっても、変換前後の状態が厳密に1対1に対応するデータ構造および演算からなるアルゴリズムを用いると、逆演算を行うアルゴリズムによって、変換後のデータからもとの状態を完全に復元することができる。この様なアルゴリズムおよびその逆アルゴリズムが存在すれば、変換の履歴さえ保存しておけば、元のデータを保存する必要が無い。元のデータは逆変換によって完全に復元することができる。この演算方式を用いると、記憶領域の節約になるばかりでなく、その管理が簡易化される。さらに、この演算方式には、演算による情報の損失が原理的に無い。そのため、将来、「可逆論理コンピュータ」と呼ばれる新しい計算原理にもとづくコンピュータの基本演算方式となる可能性もある³。

¹無理数であっても例えば代数的整数（整数係数の代数方程式の根で定義される数）を表現することは可能である [14]。しかし、任意の実数は表現できない。

²数学用語では、「正則」あるいは「可逆」な変換とは、定義域の至るところでヤコビ行列式が零でなく、写像として全単射であることをいう。しかし、本章では、「可逆」という用語は、離散近似に対して用いる。数学における連続写像の可逆性に対しては、「正則」という用語を用いて区別する。

³「可逆論理コンピュータ」とは、量子力学にもとづいた新しいデジタル計算の原理である。可逆論理計算とは、計算をデータの変換とみなし、変換前後の状態が1対1に対応する論理回路を考える。このとき、変換前後の情報エントロピーは変化しない。量子力学によれば、原理的にエネルギーを消費しないで、この可逆

4.1 整数座標における可逆座標変換 (1) — 非同次座標における等積アフィン変換 —

あらまし 非同次座標表現において、任意の等積アフィン変換が、整数座標における可逆計算によって実現できることが示される。等積アフィン変換は、一般のアフィン変換の中で等積写像の条件を満たす変換である。表現行列の行列式の絶対値が1であることが、等積写像の必要十分条件である。等積アフィン変換は、任意の鏡映、回転、等積伸縮、せん断の合成を表すことができる。これらの変換は、計算機内で幾何情報を扱う際の基本的な変換である。等積アフィン変換の唯一の制約は、体積変化を伴う変換が含まれないことである。

論理演算を行う論理素子と論理回路の物理的存在可能性が示されている [153] [156] [157] [154] [155]。本論文の理論は、この可逆論理ゲートを用いてエネルギーを消費しない座標変換のデジタル演算を可能にする。しかし、この応用は論文の目的から外れるので機会を改めて述べる。

4.1.1 はじめに

本節では、有限固定長整数変数で表現される n 次元数空間において、この可逆性を満たす離散近似 (自己同型写像) が存在するアフィン変換のクラスを示し、この自己同型写像を構成するためのアルゴリズムを提案する。さらに、2次元において誤差が最小となる離散可逆写像の構成法を述べる。本方法は、コンピュータ・ビジョンの他にも、画像処理 [18]、図形処理 [17]、計算幾何学 [16]、画像パターン認識 [19]、数値計算 [14]、並列計算などの分野に広く応用可能であり、各応用分野における具体的実現方法についても研究の余地がある。しかし、本節では、与えられたアフィン変換を近似する格子点集合の自己同型写像を構成するアルゴリズムとして、一般的に論じる。

以下、本節の構成を述べる。

まず、第 4.1.2 節において、代表的な物体形状表現法についてまとめ、それに対する座標変換の意味を述べる。第 4.1.3 節で、アフィン変換の特殊な場合である回転変換、斜交軸変換、並進変換などについて整理する。第 4.1.4 節は、本章の中心であり、ここでは、等積アフィン変換を近似する離散可逆写像を構成するための基本原理を述べる。第 4.1.5 節では、実際に可逆写像を構成するための分解アルゴリズムを提案する。第 4.1.6 節では、2次元等積アフィン変換と2次元回転変換の場合について、誤差最小の最適分解を求める。

4.1.2 物体形状の記述と座標変換

3次元物体形状の表現方法には、次のものがあげられる。

1. 輪郭と表面パッチ (曲線、平面、曲面の集合)
2. 体積表現 (空間占有表現)
3. 掃引 (スイープ) による形状 (一般化円筒など)

第1の表現方法は、物体の表面を構成する平面あるいは曲面の集合とそれらの接続関係によって物体を記述する。(cf. [24])。曲面体同士のブール演算によって、複雑な物体を記述する CSG (Constructive Solid Geometry) も、CAD やグラフィクスではよく使われる。

第2の表現方法は、空間を格子状に分割し、物体の存在する領域上の格子点に、物体の有無の情報を与える。CT (Computer Tomography) 画像に代表される高次元デジタル画像について有効な表現方法である。

第3の表現方法は、2次元形状をスケールあるいは、断面形状を変化させながら、空間曲線に沿って掃引して曲面体を生成するもので、コンピュータ・ビジョンでは一般化円筒 (あ

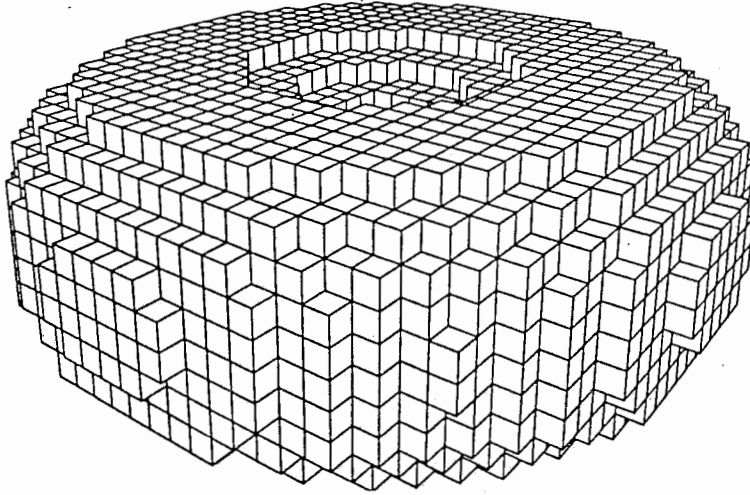


Figure 4.1: 空間占有表現による、3次元物体形状表現法の例 ([Ballard & Brown 1982 [3]]より引用)

るいは一般化円錐)と呼ばれる。これによって、動物、人間、木などの軸を持つ物体が効率的に表現できる。

以下、これらのうち、1. と 2. について以下に述べる。

4.1.2.1 点、直線、平面、代数曲面、代数曲線

前節で述べた表現のいずれにおいても、その幾何要素は、点、平面、代数曲面である。軸の表現法として直線や代数曲線も用いることがある。点、平面、曲面(体)、曲線などの各幾何要素は、3次元ユークリッド空間の座標ベクトルや、法線ベクトル、係数ベクトル、係数テンソルによって一意的に表現することができる。また、第4.2節で述べる様に、4次元の同次座標表現を用いれば、平面、直線、代数曲面、代数曲線は、定数倍の任意性をもつ非零テンソル量としてすべて統一的に扱うことができることが知られている [24]。本章の理論を用いれば、任意のテンソル量に対する可逆座標変換が常に可能である。したがって、これらテンソル量で表現される幾何要素はすべて可逆変換が可能である。

4.1.2.2 空間占有表現

空間占有表現は、3次元の配列あるいはそれと同等の記憶領域を用意し、その各点に物体への帰属の情報を与える。物体形状は、この配列に帰属の印が付けられた座標値の集合で表現される(図4.1)。この手法の変形として、八分木表現(オクトツリー)がある。これは、配列ではなく、木構造で階層的に表現された空間コードによって物体の空間的占有を表す手法であって、配列によって表現する方法に比較してデータ量は少なくて済む。CT像の様な3次元画像は配列表現である。いずれにしろ、これら空間占有表現は、3次元の各格子点上に属性を与えたものである。この物体表現を3次元剛体運動させることは、処理、表示のための基本的な操作である。これらの表現に対する3次元並進運動は比較的簡単である。ところが一般の3次元回転運動は難しい。そこで、八分木表現された3次元形状データに任意の剛体運動を施すアルゴリズムが研究されている [142] [143] [144]。これらは変換の可逆性を保証していない。データ量が大きい場合には、変換のために中間記憶領域を持たない方法が望ましい。本章で述べる可逆的座標変換の理論を用いれば、原データを保存する必要はない。しかも、原データを蓄えた記憶領域上での演算だけで、任意の剛体運動を含んだ任意の等積アフィン変換を行なうことができる。

4.1.3 アフィン変換とその部分群

本節では、等積アフィン変換のアフィン変換群における位置づけを明らかにする。そのため、非同次座標系におけるアフィン変換群の部分群をまとめる。第4.2節で述べる様に、同次座標系を用いると射影変換にまで対象を広げられる。アフィン変換は、射影変換の特殊な場合である。

一般アフィン変換とその部分集合を次の様に定義する。

定義 4.1 W は、次の写像 $f: \mathcal{R}^n \mapsto \mathcal{R}^n$ (R は実数体) の集合を表すものとする。

$$f(x) = Ax + b \quad (b \in \mathcal{R}^n, A = (a_{ij}) \in \mathcal{R}^{n \times n}) \quad (4.1)$$

W の部分集合を以下に定義する。

一般アフィン変換	$\mathcal{G} = \{f_{\mathcal{G}} \in W \det A \neq 0\},$
等積アフィン変換	$\mathcal{E} = \{f_{\mathcal{E}} \in W \det A = 1\},$
特殊アフィン変換	$\mathcal{S} = \{f_{\mathcal{S}} \in W \det A = 1\},$
一般1次変換	$\mathcal{GL} = \{f_{\mathcal{GL}} \in W \det A \neq 0, b = 0\},$
等積1次変換	$\mathcal{EL} = \{f_{\mathcal{EL}} \in W \det A = 1, b = 0\},$

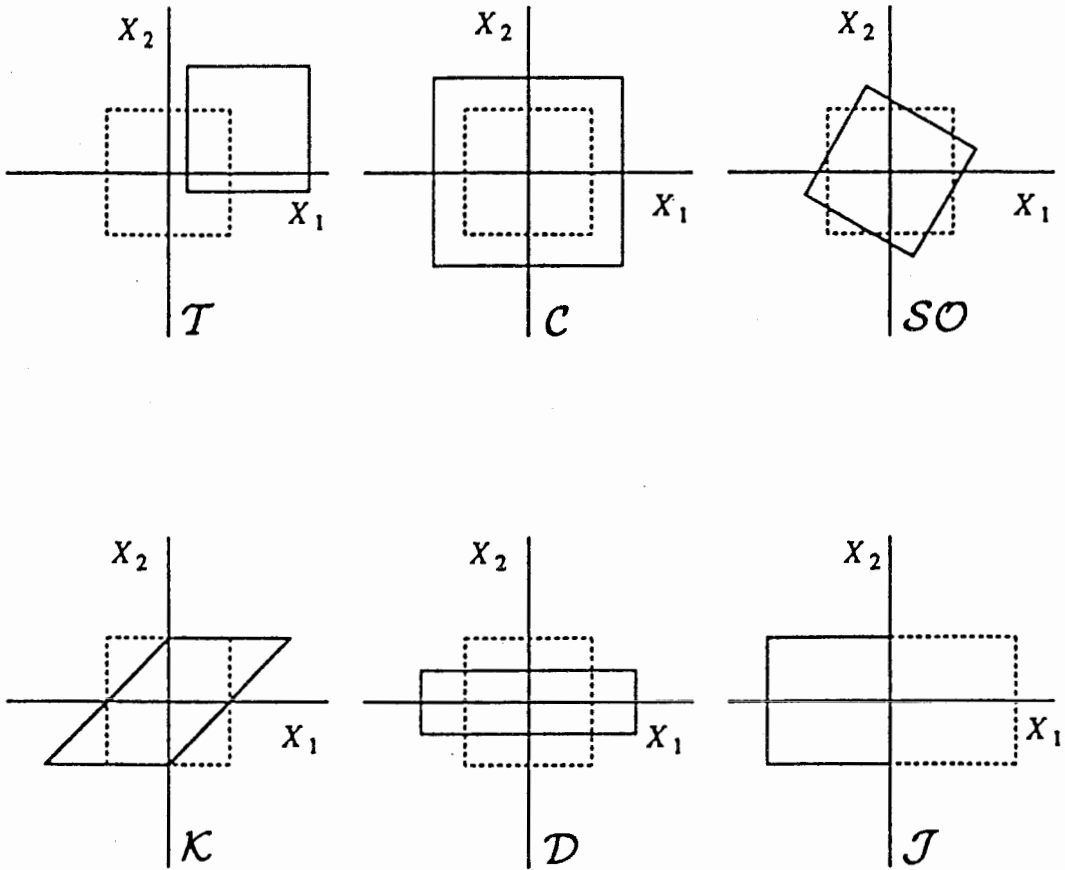


Figure 4.2: 一般アフィン変換の部分群の幾何学的意味

- 特殊1次変換 $SL = \{f_{sc} \in W \mid \det A = 1, \mathbf{b} = \mathbf{0}\},$
- 直交変換 $O = \{f_{so} \in W \mid A^T A = I, \mathbf{b} = \mathbf{0}\},$
- 特殊直交変換 $SO = \{f_{so} \in W \mid A^T A = I, \det A = 1, \mathbf{b} = \mathbf{0}\},$
- 並進変換 $T = \{f_T \in W \mid A = I\},$
- 斜交軸変換 $K = \{f_K \in W \mid a_{ii} = 1, a_{ij} = 0 (i < j), \mathbf{b} = \mathbf{0}\},$
- 等積スケール変換 $D = \{f_D \in W \mid a_{ii} > 0, \det A = \prod_i a_{ii} = 1, a_{ij} = 0 (i \neq j), \mathbf{b} = \mathbf{0}\},$
- 膨張変換 $C = \{f_c \in W \mid A = cI (R \ni c > 0), \mathbf{b} = \mathbf{0}\},$
- 反転変換 $J = \{f_J \in W \mid a_{11} \in \{-1, 1\}, a_{ii} = 1 (i \neq 1), a_{ij} = 0 (i \neq j), \mathbf{b} = \mathbf{0}\}$

これらの部分集合はすべて、写像の合成演算に関して閉じている。群の公理も満たす。□

以下では、特に指定しない限り、 n 次元空間の変換を考察する。なお、必要に応じて、定義空間の次元を下付き添字を用いて $\mathcal{EL}_2, SO_3, D_2$ の様に表す。

性質 4.1 等積アフィン変換 f_E は n 次元ユークリッド空間上の任意の有界な領域の n 次元体積を保存する。 $\det A = 1$ の場合には体積の符号も保存され、 $\det A = -1$ の場合には体積の符号が反転する。□

等積アフィン変換が持つこの体積保存の性質はアフィン変換に 1 対 1 整数近似が存在するための必要条件であることは明らかである。

一般アフィン変換 G は GL と T に一意分解可能である。 GL は、 SL と C に一意に分解できる。さらに、 SL は J, D, K, SO に一意に分解できる (岩沢分解)。つまり、次の命題が成立する。

命題 4.1 任意の一般アフィン変換 $f_G \in G$ は、次の形に一意分解できる。

$$f_G = f_T f_C f_J f_{SO} f_K f_D \quad (4.2)$$

任意の等積アフィン変換 $f_E \in E$ は、次の形に一意に分解できる。

$$f_E = f_T f_J f_{SO} f_K f_D \quad (4.3)$$

ここで、 $f_T \in T, f_J \in J, f_C \in C, f_{SO} \in SO, f_K \in K, f_D \in D$ であり、積は写像の合成である。□

例えば、2 次元 Euclid 空間において、分解 (4.2) を成分で書くと次式である。

$$f(\mathbf{x}) = c \begin{bmatrix} \delta & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \begin{bmatrix} d & 0 \\ 0 & d^{-1} \end{bmatrix} \mathbf{x} + \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}, \quad (4.4)$$

$$(\theta \in [0, 2\pi), c, d > 0, \delta \in \{-1, 1\}) \quad (4.5)$$

2 次元の場合に、これら各部分変換群の幾何学的解釈を図 4.2 に示した。

等積アフィン変換は、式 (4.4) において、制約条件 $c = 1$ を課した場合である。つまり、等積アフィン変換は、反転変換、回転変換、斜交軸変換 (せん断)、等積伸縮変換の任意の合成を表せる。命題 4.1 の表現は、アフィン変換と等積アフィン変換に一意の明確な幾何学的解釈を与える。本論文では、これを連続空間 R^n 上の表現として採用する。剛体運動を表すユークリッド変換は、特殊直交変換 SO_3 と並進変換 T_3 の合成変換である。

4.1.4 等積アフィン変換の可逆整数近似

前節で述べた連続座標における変換をデジタル計算機に実現する場合、実数型変数を用いたとしても、座標の精度は有限なので、定数倍によって整数変数とみなせる。したがって、以下では、変数はすべて整数変数と考える。

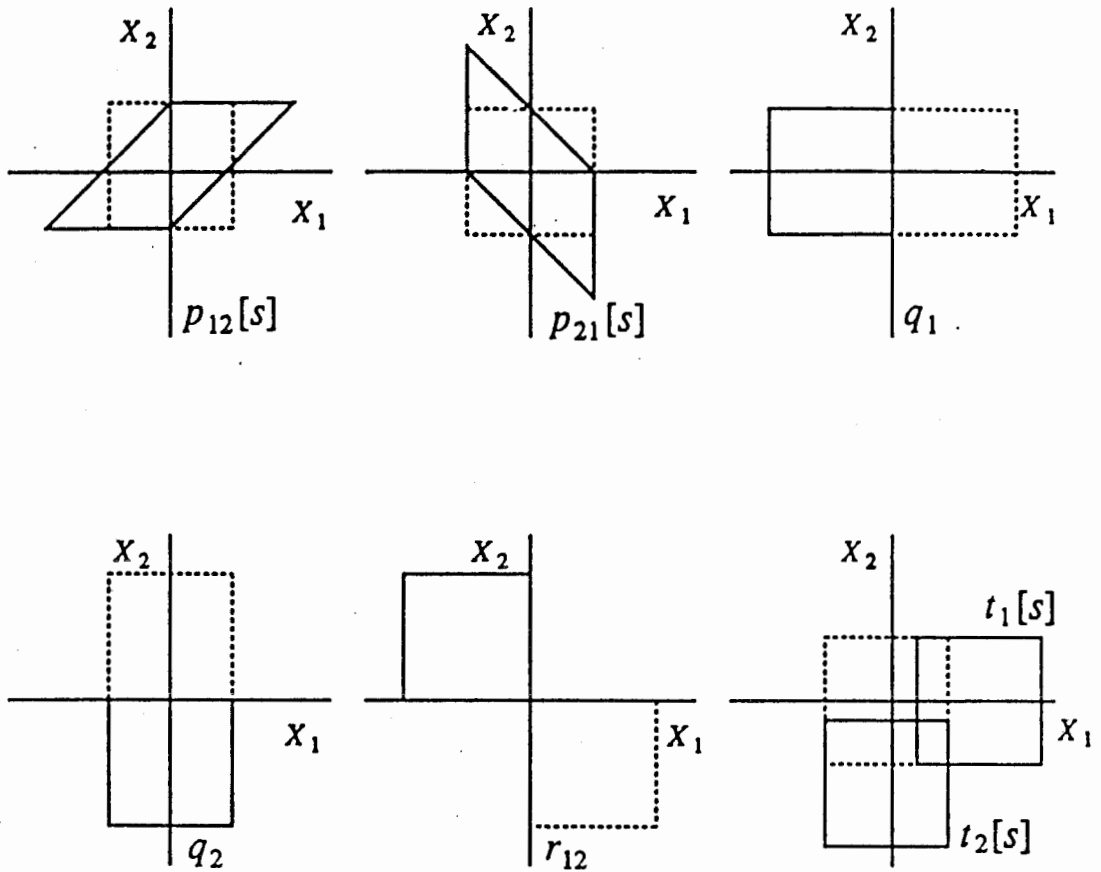


Figure 4.3: 基本変換の幾何的意味 (2次元の場合)

4.1.4.1 等積アフィン変換の等積基本変換への分解

等積基本変換 まず、本方法の基礎を準備するために、基本斜交軸変換、基本反転変換、基本並進変換を定義する。

定義 4.2 n 次の基本行列 $P_{ij}[s], Q_i, R_{ij}$, これらに対応する基本斜交軸変換 $p_{ij}[s] : \mathcal{R}^n \mapsto \mathcal{R}^n$ と基本反転変換 $q_i : \mathcal{R}^n \mapsto \mathcal{R}^n$, $r_{ij} : \mathcal{R}^n \mapsto \mathcal{R}^n$, 基本並進変換 $t_i[s] : \mathcal{R}^n \mapsto \mathcal{R}^n$ を以下に定義する。(I は単位行列、 I_{ij} は行列単位⁴、 u_i は第 i 単位ベクトルである。)

$$p_{ij}[s] : x \mapsto P_{ij}[s]x, \quad P_{ij}[s] = I + sI_{ij} (i \neq j) \tag{4.6}$$

これは、超平面 $x_j = 0$ に対する傾きが s となる様に、超平面 $x_i = 0$ にそって点を移動させる変換である。

$$q_i : x \mapsto Q_i x, \quad Q_i = I - 2I_{ii} \tag{4.7}$$

⁴ (i, j) 成分だけが 1 で、その他の成分がすべて 0 の行列。

これは、超平面 $x_i = 0$ に関する対称変換である。

$$r_{ij} : \mathbf{x} \mapsto \mathbf{R}_{ij}\mathbf{x}, \quad \mathbf{R}_{ij} = \mathbf{I} - \mathbf{I}_{ii} - \mathbf{I}_{jj} + \mathbf{I}_{ij} + \mathbf{I}_{ji} (i \neq j) \quad (4.8)$$

これは、超平面 $x_i = x_j$ に関する対称変換である。

$$t_i[s] : \mathbf{x} \mapsto \mathbf{x} + s\mathbf{u}_i \quad (4.9)$$

これは、 x_i 軸にそった s だけの平行移動を表す。□

通常の線形代数における基本行列は、 $\mathbf{P}_{ij}[s]$, \mathbf{R}_{ij} および、 $\mathbf{Q}_i[s] = \mathbf{I} - (1-s)\mathbf{I}_{ii}$ と定義される [169] [166]。定義 4.2 では、 $\mathbf{Q}_i[s]$ が $\mathbf{Q}_i[-1] = \mathbf{Q}_i$ に制限されている。これは、等積条件から要請される制限である。

$n = 2$ の場合、基本行列は、次に示す 5 種類である。

$$\begin{aligned} \mathbf{P}_{12}[s] &= \begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix}, & \mathbf{P}_{21}[s] &= \begin{bmatrix} 1 & 0 \\ s & 1 \end{bmatrix}, \\ \mathbf{Q}_1 &= \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, & \mathbf{Q}_2 &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, & \mathbf{R}_{12} &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \end{aligned} \quad (4.10)$$

2次元の場合の各基本変換の幾何学的意味を図 4.3 に示す。以下では、基本斜交軸変換と基本並進変換をまとめて 1 パラメータ基本変換と呼ぶことにする。

本論文で提案する方法は次に述べる定理に基礎をおく。

定理 4.1 等積 1 次変換 $f_{\mathcal{EL}} \in \mathcal{EL}$ は、 $(n^2 - 1)$ 個の基本斜交軸変換と有限個の基本反転変換の積に分解可能である。□

系 4.1 等積アフィン変換 $f_{\mathcal{E}} \in \mathcal{E}$ は、 $(n^2 - 1)$ 個の基本斜交軸変換と n 個の基本並進変換と有限個の基本反転変換の積に分解可能である。□

[証明] 付録 C.1 を見よ。

\mathcal{EL} の多様体としての次元は $(n^2 - 1)$ である。ゆえに、定理 4.1 における基本斜交軸変換の数は、 \mathcal{EL} の任意の要素を分解するために必要かつ十分な数である。この分解は一意ではない。したがって、変換全体の近似誤差を評価し、近似誤差ができるだけ小さい分解を選び出す必要がある。この問題は次節以降で論じる。

分解列における基本変換の順序は、ある程度変更可能である。この順序変更のための交換関係を付録 C.5 に示す。

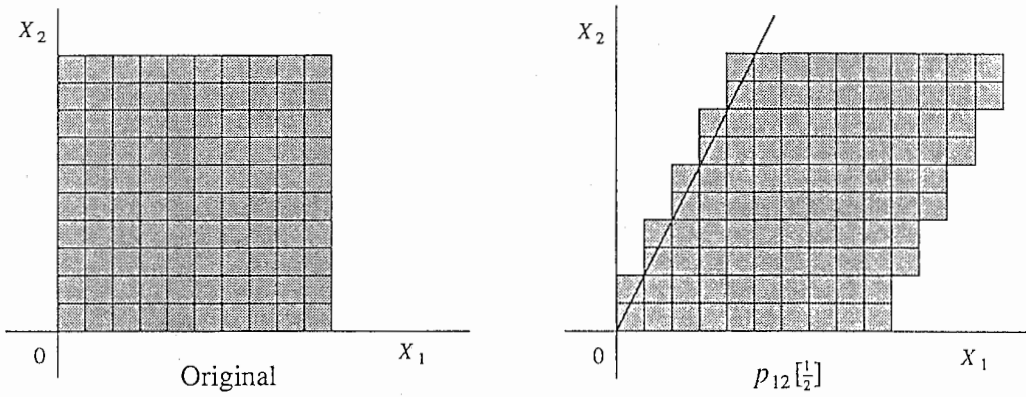


Figure 4.4: 基本斜交軸変換の整数近似写像 $\hat{p}_{ij}[s]$ の例

4.1.4.2 基本変換の整数近似

有限の格子と無限の格子を同時に扱うために、 \mathcal{Z} は、整数環 $\mathcal{Z}_0 = \{0, \pm 1, \pm 2, \dots\}$ 、奇数位数の剰余類環 $\mathcal{Z}_{2m+1} = \{0, \pm 1, \pm 2, \dots, \pm m\}$ 、偶数位数の剰余類環 $\mathcal{Z}_{2m} = \{0, \pm 1, \pm 2, \dots, \pm(m-1), m\}$ のうちのいずれかを表す。座標軸ごとに異なる位数を許すため、 \mathcal{Z}^n は、 $\mathcal{Z}_1, \mathcal{Z}_2, \dots, \mathcal{Z}_0$ の任意の形の積を表すことにする。有限位数の \mathcal{Z}_m が用いられた場合には、それぞれの近似基本変換の後に剰余演算を行うものとする。整数近似の方法は、次の関数 $int()$ によって定義される最近傍近似とする。

定義 4.3 関数 $int() : \mathbf{x} \mapsto \mathbf{z} (x = (x_1, x_2, \dots, x_n) \in \mathcal{R}^n, z = (z_1, z_2, \dots, z_n) \in \mathcal{Z}^n)$ は、次式を満たす整数値関数である。

$$x_i - \frac{1}{2} \leq z_i = int(x_i) < x_i + \frac{1}{2} \tag{4.11}$$

このとき、 \mathcal{Z}^n の各因子が $\mathcal{Z}_0, \mathcal{Z}_{2m+1}, \mathcal{Z}_{2m}$ のいずれかであることに対応して、 \mathcal{R}^n を定める。つまり、それぞれに、 $\mathcal{R}_0 = \mathcal{R}, \mathcal{R}_{2m+1} = [-m, m+1) \subset \mathcal{R}, \mathcal{R}_{2m} = [-(m-1), m+1) \subset \mathcal{R}$ を対応させる。それらの積を \mathcal{R}^n であるとする。□

基本反転変換は、変換の定義域と値域を \mathcal{Z}^n に制限することによって整数値ベクトルに対する変換に変更できる。この場合には誤差は発生しない。

基本並進変換と基本斜交軸変換は、 \mathcal{R}^n で計算されたベクトルの各成分を整数に丸めるとによって近似する。これらの整数近似変換は次式で定義される。

定義 4.4 $p_{ij}[s]$, q_i , r_{ij} , $t_i[s]$ を整数近似したものをそれぞれ、 $\hat{p}_{ij}[s]$, \hat{q}_i , \hat{r}_{ij} , $\hat{t}_i[s]$ とする。これらは、次式で定義される。

$$\begin{aligned}\hat{p}_{ij}[s] : & \quad \mathbf{z} \mapsto \text{int}(p_{ij}[s](\mathbf{z})) = \mathbf{z} + \text{int}(s\mathbf{I}_{ij}\mathbf{z}), \\ \hat{q}_i : & \quad \mathbf{z} \mapsto \text{int}(q_{ij}(\mathbf{z})) = \mathbf{Q}_i\mathbf{z} = \mathbf{z} - 2\mathbf{I}_{ii}\mathbf{z}, \\ \hat{r}_{ij} : & \quad \mathbf{z} \mapsto \text{int}(r_{ij}(\mathbf{z})) = \mathbf{R}_{ij}\mathbf{z} = \mathbf{z} - \mathbf{I}_{ii}\mathbf{z} - \mathbf{I}_{jj}\mathbf{z} + \mathbf{I}_{ij}\mathbf{z} + \mathbf{I}_{ji}\mathbf{z}, \\ \hat{t}_i[s] : & \quad \mathbf{z} \mapsto \text{int}(t_i[s](\mathbf{z})) = \mathbf{z} + \text{int}(s)\mathbf{u}_i\end{aligned}\quad (4.12)$$

図 4.4 は、2次元の基本斜交軸変換の整数近似の例である。以上の定義と議論から、次の命題が成り立つ。

命題 4.2 変換 $\hat{p}_{ij}[s]$, \hat{q}_i , \hat{r}_{ij} , $\hat{t}_i[s]$ は、デジタル格子 \mathcal{Z}^n の自己同型写像 (1対1写像) である。したがって、これらの変換には逆変換が存在する。□

命題 4.3 整数近似基本変換 $\hat{p}_{ij}[s]$, \hat{q}_i , \hat{r}_{ij} , $\hat{t}_i[s]$ の逆変換は、それぞれ、次式である。

$$\begin{aligned}\hat{p}_{ij}^{-1}[s] : & \quad \mathbf{z} \mapsto \mathbf{z} - \text{int}(s\mathbf{I}_{ij}\mathbf{z}), \\ \hat{q}_i^{-1} : & \quad \mathbf{z} \mapsto \mathbf{Q}_i\mathbf{z} = \mathbf{z} - 2\mathbf{I}_{ii}\mathbf{z}, \\ \hat{r}_{ij}^{-1} : & \quad \mathbf{z} \mapsto \mathbf{R}_{ij}\mathbf{z} = \mathbf{z} - \mathbf{I}_{ii}\mathbf{z} - \mathbf{I}_{jj}\mathbf{z} + \mathbf{I}_{ij}\mathbf{z} + \mathbf{I}_{ji}\mathbf{z}, \\ \hat{t}_i^{-1}[s] : & \quad \mathbf{z} \mapsto \mathbf{z} - \text{int}(s)\mathbf{u}_i\end{aligned}\quad (4.13)$$

[証明] 付録 C.2 を参照。

同型写像の合成則と命題 4.2 を用いると、次の主定理が成り立つ。

定理 4.2 $p_{ij}[s]$, q_i , r_{ij} および $t_i[s]$ の任意の合成によって得られる写像に対して、それを近似するデジタル格子上の自己同型写像 (1対1写像) が存在する。□

系 4.2 任意の等積アフィン変換 $f_{\mathcal{E}} \in \mathcal{E}$ に対して、それを近似するデジタル格子上の自己同型写像 (1対1写像) が存在する。□

4.1.4.3 演算子記法を用いた近似誤差の評価法

与えられた等積アフィン変換に対して、合成された写像全体がもつ近似誤差を評価する。この誤差評価においては、簡単のため、与えられた変換を分解する際の計算誤差を零と仮定する。これは、変換の分解によって近似的に得られた基本変換をふたたび合成したものを与

えられた変換行列であるとみなして誤差評価を行うことと等価である。明らかに、無限格子 Z_0^n 上で導いた誤差上界は任意の有限格子に対しても誤差上界となる。以下では、無限格子 Z_0^n 上で考える。

前節で示した結果によれば、1パラメータ基本変換の整数近似には、整数近似誤差が存在する。以下、 $\mathcal{R}_+^n = \{r \in R | r \geq 0\}^n$ の誤差ベクトルを用いて計算誤差を評価する。この誤差ベクトルは、無限精度の変数で計算した時に得られる座標値と、整数近似された1パラメータ基本変換を合成して得られる写像を用いた座標値の差の絶対値の上界をあらわすものとする。この誤差ベクトルの具体的構成法は以下に述べる。整数近似の良さを、この誤差ベクトルの成分から計算されるノルムによって測る。このノルムには、例えば、Euclid ノルム、 L_1 ノルム、 L_∞ ノルムなどが考えられる。各座標軸に関して対称でないノルムも用いることができるところが、この誤差ベクトルを用いた方法の利点である。誤差ベクトルのノルムと誤差ベクトルの初期値を仮定せずに以下の議論を進める。これらは、適用する問題に応じて決める必要がある。

まず、誤差伝搬演算子を定義する。

定義 4.5 各基本変換の整数近似 $\hat{p}_{ij}[s], \hat{q}_i, \hat{r}_{ij}, \hat{t}_i[s]$ にそれぞれ対応した誤差伝搬演算子 $\epsilon\{\hat{p}_{ij}[s]\}, \epsilon\{\hat{q}_i\}, \epsilon\{\hat{r}_{ij}\}, \epsilon\{\hat{t}_i[s]\}$ を、次式で定義する。

$$\begin{aligned} \epsilon\{\hat{p}_{ij}[s]\} &: e \mapsto P_{ij}[|s|]e + \frac{1}{2}u_j, \\ \epsilon\{\hat{q}_i\} &: e \mapsto e, \\ \epsilon\{\hat{r}_{ij}\} &: e \mapsto R_{ij}e, \\ \epsilon\{\hat{t}_i[s]\} &: e \mapsto e + \frac{1}{2}u_i \square \end{aligned} \tag{4.14}$$

次の定理によって、定義 4.5 の誤差伝搬演算子が、誤差上界を評価するために有効であることがわかる。

定理 4.3 $z \in Z^n, x \in R^n, e \in \mathcal{R}_+^n$ 、 $\alpha, \beta, \gamma, \dots$ は任意の基本変換で、 $\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \dots$ はそれらの整数近似変換とする。すると、次の不等式が成立する。

$$\begin{aligned} |z - x| \leq e &\implies \\ |\hat{\alpha}(z) - \alpha(x)| &\leq \epsilon\{\hat{\alpha}\}(e), \\ |\hat{\beta}\hat{\alpha}(z) - \beta\alpha(x)| &\leq \epsilon\{\hat{\beta}\}\epsilon\{\hat{\alpha}\}(e), \\ |\hat{\gamma}\hat{\beta}\hat{\alpha}(z) - \gamma\beta\alpha(x)| &\leq \epsilon\{\hat{\gamma}\}\epsilon\{\hat{\beta}\}\epsilon\{\hat{\alpha}\}(e), \\ &\dots \end{aligned}$$

絶対値記号 $|\cdot|$ と不等式は、引数のベクトルの各成分に独立に作用するものとする。□

[証明] 付録 C.3 を参照。

4.1.5 n 次元における一般の場合の分解アルゴリズム

n 次元の任意の等積アフィン変換を基本変換に分解するアルゴリズムを提案する。提案するアルゴリズムは、座標ベクトルに演算を行う順番に基本変換を決定する。したがって、前節の誤差評価の方法をアルゴリズム中に埋め込んで誤差最小分解を求めることが容易である。基本並進変換の分解は自明なので、以下では、等積1次変換 \mathcal{EL} について分解アルゴリズムを述べる。なお、整数近似変換全体の誤差をできる限り小さくするために、基本並進変換は、分解列の最後に置くことが望ましい。

\mathcal{EL} の表現行列 A ($|\det A| = 1$) をアルゴリズムの入力とする。

アルゴリズム 4.1

(1) $k = 0$, $B^{[k]} = (b_{ij}^{[k]}) := A$ と初期化する ($:=$ は代入を表す)。

(2) $k = 1, 2, \dots, (n-1)$ に対して、(2-1), (2-2) を順に実行する。

(2-1) もし、すべての $j(>k)$ に対して $b_{kj}^{[k-1]} = 0$ ならば、(a) を実行し、さもなければ (b) を実行する。

(a) 任意の $j(>k)$ を選んで、

$$\tilde{B}^{[k]} := B^{[k-1]} R_{kj} P_{jk} \begin{bmatrix} 1 \\ b_{kk}^{[k-1]} \end{bmatrix} \quad (4.15)$$

とする。

$$S^{[k]} \equiv P_{jk} \begin{bmatrix} -1 \\ b_{kk}^{[k-1]} \end{bmatrix} R_{kj} \quad (4.16)$$

と定義して (\equiv は分解列の定義を表す)、(2-2) に進む。

(b) $b_{kj}^{[k-1]} \neq 0$ を満たす任意の $j(>k)$ を選んで、

$$\tilde{B}^{[k]} := B^{[k-1]} P_{jk} \begin{bmatrix} 1 - b_{kk}^{[k-1]} \\ b_{kj}^{[k-1]} \end{bmatrix} \quad (4.17)$$

とする。

$$S^{[k]} \equiv P_{jk} \begin{bmatrix} b_{kk}^{[k-1]} - 1 \\ b_{kj}^{[k-1]} \end{bmatrix} \quad (4.18)$$

と定義して (2-2) に進む。

(2-2)

$$B^{[k]} := \tilde{B}^{[k]} \prod_{j \neq k} P_{kj} \begin{bmatrix} \hat{b}_{kj}^{[k]} \end{bmatrix} \quad (4.19)$$

とする。

$$T^{[k]} \equiv \prod_{j \neq k} P_{kj} \begin{bmatrix} \hat{b}_{kj}^{[k]} \end{bmatrix} \quad (4.20)$$

と定義する。

(3) 第 n 行に関しては、変換行列の等積条件から、 $b_{nn}^{[n-1]} = 1$ または -1 となる。そこで、 $b_{nn}^{[n-1]} = 1$ であるか、 $b_{nn}^{[n-1]} = -1$ であるかに応じて、

$$S^{[n]} \equiv I, \quad T^{[n]} \equiv \prod_{j < n} P_{nj} \left[b_{nj}^{[n-1]} \right] \quad (4.21)$$

または、

$$S^{[n]} \equiv Q_n, \quad T^{[n]} \equiv \prod_{j < n} P_{nj} \left[-b_{nj}^{[n-1]} \right] \quad (4.22)$$

と定義する。

(4) 行列 A の分解結果は、次式である。

$$T^{[n]} S^{[n]} T^{[n-1]} S^{[n-1]} \dots T^{[1]} S^{[1]} \square \quad (4.23)$$

このアルゴリズムは、連立1次方程式を解くためのガウス・ジョルダンの消去法に類似している。しかし、ピボット成分の正規化に相当するステップ (2-1)-(a) および (2-1)-(b) に違いがある。このアルゴリズムでは、第 k 行の成分がゼロでない列 (第 j 列) を決めて、基本行列 $P_{jk[s]}$ を右から作用させて第 k 対角成分を 1 に正規化している。

誤差最小分解の探索は、ガウス・ジョルダンの消去法におけるピボット選択問題と同様の考え方をを用いる。良いピボットの選択は、第 4.1.4.3 節で述べた意味で、より小さい誤差を与えるものとする。ガウス・ジョルダン法においては、ピボット選択は、列または行の置換によって実現される。本方法では、これらの操作に加えて、第 k 行の処理において、第 k 行または第 k 列の符号反転も誤差の改善のために用いることができる可能性がある (実際には、第 k 行の符号反転は、誤差ベクトルを変化させない。したがって、第 k 列の符号反転のみが誤差改善に意味のある操作である)。これらのピボット選択は、アルゴリズム 1 のステップ (2-1) において、行列 Q_i および R_{ij} を行列 $A^{[k]}$ の左右から乗じて実現できる。

命題 4.4 アルゴリズム 1 のステップ (2-1) における第 k 行についてのすべての可能なピボット選択と掃き出し操作、および符号反転において、誤差評価に影響を与えるものを実際に基本行列を用いた式で表す。

$v \neq k$ かつ $v \neq w$ の場合 (case I)、

$$A^{[k+1]} = R_{ku} A^{[k]} R_{kv} Q_{[s]} P_{wk} \left[\frac{1 - \delta a_{uv}^{[k]}}{a_{uw}^{[k]}} \right] P_{kv} \left[-a_{uk}^{[k]} \right] \prod_{j \neq v, j \neq k} P_{kj} \left[-a_{uj}^{[k]} \right] \quad (4.24)$$

$v = k$ の場合 (case II)、

$$\mathbf{A}^{[k+1]} = \mathbf{R}_{ku} \mathbf{A}^{[k]} \mathbf{Q}_{[\delta]} \mathbf{P}_{wk} \left[\frac{1 - \delta a_{uk}^{[k]}}{a_{uw}^{[k]}} \right] \prod_{j \neq k} \mathbf{P}_{kj} [-a_{uj}^{[k]}] \quad (4.25)$$

$v = w$ の場合 (case III)、

$$\mathbf{A}^{[k+1]} = \mathbf{R}_{ku} \mathbf{A}^{[k]} \mathbf{R}_{kw} \mathbf{Q}_{[\delta]} \mathbf{P}_{wk} \left[\frac{1 - \delta a_{uw}^{[k]}}{a_{uk}^{[k]}} \right] \mathbf{P}_{kw} [-a_{uk}^{[k]}] \prod_{j \neq w, j \neq k} \mathbf{P}_{kj} [-a_{uj}^{[k]}] \quad (4.26)$$

ただし、 $\mathbf{Q}_{[1]} = \mathbf{I}, \mathbf{Q}_{[-1]} = \mathbf{Q}_k$ であり、 $u \in \{k, \dots, n\}, v \in \{k, \dots, n\}, w \in \{k+1, \dots, n\}, \delta \in \{-1, 1\}$ は、最小誤差を与える分解を探索するためのインデクスパラメータである。□

[証明] 付録 C.4 を参照。

定理 4.3 における誤差伝搬演算子を用いて、命題 4.4 の操作による誤差ベクトルの変化を計算すると次の命題が導かれる。

命題 4.5 $\mathbf{e}^{[k]} = (e_1^{[k]}, e_2^{[k]}, \dots, e_n^{[k]})^T$ を命題 4.4 の第 k 行処理後の誤差ベクトルを表し、 $\mathbf{e}^{[0]} = (e_1^{[0]}, e_2^{[0]}, \dots, e_n^{[0]})^T$ を初期誤差ベクトルとする。すると、 $\mathbf{e}^{[k-1]}$ から $\mathbf{e}^{[k]}$ を計算する公式は次式で与えられる。

(case I)

$$e_j^{[k]} = e_j^{[k-1]} + |a_{uj}^{[k]}| e_v^{[k-1]} + \left| \frac{a_{uj}^{[k]} (a_{uv}^{[k]} - \delta)}{a_{uw}^{[k]}} \right| e_w^{[k-1]} + \frac{1}{2} (|a_{uj}^{[k]}| + 1) \quad (j \neq v, j \neq k) \quad (4.27)$$

$$e_v^{[k]} = e_k^{[k-1]} + |a_{uk}^{[k]}| e_v^{[k-1]} + \left| \frac{a_{uk}^{[k]} (a_{uv}^{[k]} - \delta)}{a_{uw}^{[k]}} \right| e_w^{[k-1]} + \frac{1}{2} (|a_{uk}^{[k]}| + 1) \quad (4.28)$$

$$e_k^{[k]} = e_v^{[k-1]} + \left| \frac{a_{uv}^{[k]} - \delta}{a_{uw}^{[k]}} \right| e_w^{[k-1]} + \frac{1}{2} \quad (4.29)$$

(case II)

$$e_j^{[k]} = e_j^{[k-1]} + |a_{uj}^{[k]}| e_k^{[k-1]} + \left| \frac{a_{uj}^{[k]} (a_{uk}^{[k]} - \delta)}{a_{uw}^{[k]}} \right| e_w^{[k-1]} + \frac{1}{2} (|a_{uj}^{[k]}| + 1) \quad (j \neq k) \quad (4.30)$$

$$e_k^{[k]} = e_k^{[k-1]} + \left| \frac{a_{uk}^{[k]} - \delta}{a_{uw}^{[k]}} \right| e_w^{[k-1]} + \frac{1}{2} \quad (4.31)$$

(case III)

$$e_j^{[k]} = e_j^{[k-1]} + |a_{uj}^{[k]}| e_w^{[k-1]} + \left| \frac{a_{uj}^{[k]} (a_{uw}^{[k]} - \delta)}{a_{uk}^{[k]}} \right| e_k^{[k-1]} + \frac{1}{2} (|a_{uj}^{[k]}| + 1) \quad (j \neq v, j \neq k) \quad (4.32)$$

$$e_w^{[k]} = e_k^{[k-1]} + |a_{uk}^{[k]}| e_w^{[k-1]} + |a_{uw}^{[k]} - \delta| e_k^{[k-1]} + \frac{1}{2} (|a_{uk}^{[k]}| + 1) \quad (4.33)$$

$$e_k^{[k]} = e_w^{[k-1]} + \left| \frac{a_{uw}^{[k]} - \delta}{a_{uk}^{[k]}} \right| e_k^{[k-1]} + \frac{1}{2} \square \quad (4.34)$$

(証明略)

これらの公式を用いて、最小誤差を達成する分解アルゴリズムを構成することができる。

4.1.6 2次元1次変換の誤差最小分解

本節では、以上の議論を、2次元特殊1次変換 SL_2 と2次元回転変換 SO_2 に適用した結果を述べる。2次元の可逆変換は、コンピュータディスプレイ表示のための自然画像や文字フォント画像の変換に応用できる。

4.1.6.1 2次元特殊1次変換 SL_2

命題 4.6 $f_{SL_2} \in SL_2$ は、次の (a), (b) 2通りの形式に分解される。

$$\begin{aligned} (a) \quad f_{SL_2} &= \sigma p_{21}[\alpha] p_{12}[\beta] p_{21}[\gamma], \\ (b) \quad f_{SL_2} &= \sigma p_{12}[\alpha] p_{21}[\beta] p_{12}[\gamma] \end{aligned} \quad (4.35)$$

$\sigma \in \{0, q_1 q_2, q_1 r_{12}, q_2 r_{12}\}$ である。これは、3個の基本斜交軸変換を用いた実質的に異なる分解⁵のすべての場合を尽くしている。□

SL_2 の表現行列 A を次式とする。

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad (4.36)$$

ただし、 $\det A = a_{11} a_{22} - a_{12} a_{21} = 1$ である。この行列 A は、式 (4.4) の結果を用いて、次の形に書ける。

$$A = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \begin{bmatrix} d & 0 \\ 0 & d^{-1} \end{bmatrix}$$

⁵付録 C.5に示した交換関係で互いに交換できるものを実質的に同じ変換と呼ぶことにする。実質的に同じ変換によっては、誤差ベクトルは変化しない。

$$= \begin{bmatrix} d \cos \theta & d^{-1}(k \cos \theta - \sin \theta) \\ d \sin \theta & d^{-1}(k \sin \theta + \cos \theta) \end{bmatrix} \quad (4.37)$$

各パラメータ θ, k, d は、次の範囲の実数をとる。

$$\theta \in [0, 2\pi), \quad k \in (-\infty, \infty), \quad d \in (0, \infty) \quad (4.38)$$

命題 4.7 SL_2 の表現行列 A は、次表に示す 8 通りの分解をもつ。

分解種類	分解形式	σ	α	β	γ
(1)	(a)	o	$(a_{22} - 1)/a_{12}$	a_{12}	$(a_{11} - 1)/a_{12}$
(2)	(a)	$q_1 q_2$	$(a_{22} + 1)/a_{12}$	$-a_{12}$	$(a_{11} + 1)/a_{12}$
(3)	(a)	$q_1 r_{12}$	$-(a_{12} + 1)/a_{22}$	a_{22}	$(a_{21} - 1)/a_{22}$
(4)	(a)	$q_2 r_{12}$	$-(a_{12} - 1)/a_{22}$	$-a_{22}$	$(a_{21} + 1)/a_{22}$
(5)	(b)	o	$(a_{11} - 1)/a_{21}$	a_{21}	$(a_{22} - 1)/a_{21}$
(6)	(b)	$q_1 q_2$	$(a_{11} + 1)/a_{21}$	$-a_{21}$	$(a_{22} + 1)/a_{21}$
(7)	(b)	$q_1 r_{12}$	$-(a_{21} - 1)/a_{11}$	$-a_{11}$	$(a_{12} + 1)/a_{11}$
(8)	(b)	$q_2 r_{12}$	$-(a_{21} + 1)/a_{11}$	a_{11}	$(a_{12} - 1)/a_{11}$

(4.39)

命題 4.8 初期誤差ベクトルを零ベクトルと仮定し、誤差ノルムを L_1 ノルムと仮定する。

命題 4.6 と命題 4.7 を用いて導かれる誤差評価式は、

$$\varepsilon = \frac{1}{2} (3 + |\alpha| + |\beta| + |\alpha\beta|) \quad (4.40)$$

である。□

2次元特殊1次変換 SL_2 の例 図 4.5 は、変換行列が座標軸方向への1対2の拡大縮小と30度の回転を合成した次の変換行列である場合に本結果を適用した例である。誤差最小の分解と誤差最大の分解と、それらから得られる可逆写像の実行例を示している。各格子点に文字を割り当てているため、写像が1対1であることを確認できる。

$$A = \begin{bmatrix} \cos 30^\circ & -\sin 30^\circ \\ \sin 30^\circ & \cos 30^\circ \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & \sqrt{2} \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{6}}{4} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{4} & \frac{\sqrt{6}}{2} \end{bmatrix} \quad (4.41)$$

4.1.6.2 2次元特殊直交変換 SO_2

1パラメータ群である SO_2 は、命題 4.7 において $k = 0$ 、 $d = 1$ とおいた場合である。このとき、分解 (1), (2), (3), (4) の誤差評価と、分解 (5), (6), (7), (8) の誤差評価はそれぞれ等しくなる。誤差最小の分解は、回転角度 θ に対して陽に求められる。

命題 4.9 SO_2 の誤差最小分解は、次表で与えられる。

θ の範囲	分解種類	$2(\varepsilon - 2)$
$[-\frac{\pi}{4}, \frac{\pi}{4}]$	(1), (5)	$ \sin \theta - \cos \theta + \left \tan \frac{\theta}{2} \right $
$[\frac{\pi}{4}, \frac{3\pi}{4}]$	(3), (7)	$ \cos \theta - \sin \theta + \left \tan \left(\frac{\theta}{2} - \frac{\pi}{4} \right) \right $
$[\frac{3\pi}{4}, \frac{5\pi}{4}]$	(2), (6)	$ \sin \theta + \cos \theta + \left \cot \frac{\theta}{2} \right $
$[\frac{5\pi}{4}, \frac{7\pi}{4}]$	(4), (8)	$ \cos \theta + \sin \theta + \left \cot \left(\frac{\theta}{2} - \frac{\pi}{4} \right) \right $

(4.42)

誤差が最大になるのは $\theta = \frac{n\pi}{4} (n \in \mathbb{Z})$ の時で、誤差の最大値は、 $\max \varepsilon = \frac{3+\sqrt{2}}{2} \approx 2.2$ である。

2次元特殊1次変換 SO_2 の例 図 4.6 は、変換行列が30度の回転変換である場合に対して本結果を適用した例である。誤差最小の分解と誤差最大の分解、および、それらから得られる可逆写像の実行例を示している。各格子点に文字を割り当てているため、写像が1対1であることを確認できる。

$$A = \begin{bmatrix} \cos 30^\circ & -\sin 30^\circ \\ \sin 30^\circ & \cos 30^\circ \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \quad (4.43)$$

4.2 整数座標における可逆座標変換 (2) — 同次座標における非特異射影変換 —

あらかし 同次座標表現において、任意の非特異射影変換が、整数座標における可逆計算によって実現できることが示される。射影変換には、任意のアフィン変換とともに、幾何情報の3次元表示で重要な役割を果たす透視変換が含まれる。

そのために、非同次整数座標における可逆座標変換の理論を拡張し、同次座標表現に適用する。特に、4次元同次座標表現された幾何データに対する任意の非特異射影変換が可逆離散近似可能であることを示す。議論を具体的にするため、本節では、4次元同次座標に対する3次元射影変換に限定して述べる。しかし、この理論は、任意次元の同次座標表現とその上での射影変換に適用可能である。

4.2.1 はじめに

4次元の同次座標を用いると、3次元空間の任意の射影変換が 4×4 の行列で表現される1次変換によって実現できる [17] [24]。3次元射影変換には、3次元幾何情報処理の基本変換処理である平行移動、回転、拡大縮小などの線形変換と射影変換の基本変換である透視変換の任意の組合せが含まれる。この方法は、グラフィクスやCADにおける幾何情報の計算機表現のための基本技術として用いられている。また、多関節ロボットの座標表現の標準的方法でもある。

本来、同次座標表現は、射影幾何学を展開するための舞台である射影空間を構成するための道具として、数学で用いられていたものである。同次座標表現は、1960年代前半にRogersとAdamsによって、幾何情報の計算機表現として整理された [17]。

同次座標表現によれば、任意の射影変換が、除算を含まない行列の乗法演算だけで表現できる。そのため、種々の幾何演算に、無限遠点などの特異な場合のための例外処理を設ける必要がない。幾何要素同士の交差判定や包含関係の推定も、除算を含まない行列式の符号判定だけによって実現できる方法が研究されている [24]。ここでは、幾何変換についてだけ論じ、交差判定などの幾何推論は扱わない。幾何推論の目的には、従来の方法をそのまま用いることができる。

以下、本節の構成を述べる。

第4.2.2節では、4次元同次座標を用いた3次元射影変換の表現法について、必要な事項をまとめる。第4.2.3節では、一般の非特異射影変換を4次元同次座標において可逆整数近似する方法を述べる。第4.2.4節では、点、平面、2次曲面、直線について、可逆整数近似変換の構成法を述べる。

4.2.2 3次元射影変換と4次元等積1次変換の関係

3次元射影空間の4次元同次座標による表現を説明する。特に、任意の3次元非特異射影変換が、4次元等積1次変換で表現可能であることを示す。

4.2.2.1 射影空間と射影変換の定義

n 次元の実射影空間 P^n は、 $n+1$ 次元実ベクトル空間 V^{n+1} において、互いに平行な0でないベクトル同士を同一視して構成される空間である [174]⁶。この射影空間 P^n 上で定義される射影変換 PGL_n とは、 V^{n+1} 上の1次変換で P^n 上の1対1変換である⁷。

⁶ $P^n = V^{n+1}/R - \{0\}$ と記す。

⁷ $PGL_n = GL_{n+1}/R - \{0\}$ と記す。

同次座標 (Homogeneous coordinates, 齊次座標とも呼ばれる) は、射影空間とその上の射影変換を表現する手法である。同次座標では、位置ベクトルは、零でない $n+1$ 次元実ベクトル $\mathcal{R}^{n+1} - \{0\}^{n+1}$ を用いて表現される。射影変換は、この同次ベクトルに対する $n+1$ 次元の正則 1 次変換を用いるのが普通である [17]。

以下、3次元ユークリッド空間における幾何情報の表現法を述べる。4次元ベクトル $\mathbf{X} = (X, Y, Z, W)^T$ で同次座標を表す。通常のユークリッド空間の座標 $\mathbf{x} = (x, y, z)^T$ との対応は、 $x = X/W, y = Y/W, z = Z/W$ によって定義する。射影変換は、4次の正則 1 次変換行列 \mathbf{A} で表現する。画像ディスプレイの2次元画像面への投影を表す射影変換の場合も、結果の視線方向の座標 (通常、 Z 座標が対応する) を無視することにすれば、正則変換を用いて実現可能である [17]。

以下では、表現行列 \mathbf{A} は、常に正則 ($\det \mathbf{A} \neq 0$) であると仮定する。表現行列が正則である射影変換は非特異射影変換と呼ばれる。

4次元同次座標ベクトル \mathbf{X} を \mathbf{X}' に移す射影変換は、1次変換行列 $\mathbf{A} \in \mathcal{GL}_4$ によって表される。

$$\mathbf{X}' = \mathbf{A}\mathbf{X} \quad (4.44)$$

変換行列 \mathbf{A} は、次に定義される4次の正方行列である。

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \quad (4.45)$$

射影変換は、

$$\begin{bmatrix} X' \\ Y' \\ Z' \\ W' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ W \end{bmatrix} \quad (4.46)$$

とかける。

射影変換は任意のアフィン変換を表現できる。実際、 $W = W' = a_{44} = 1, a_{41} = a_{42} = a_{43} = a_{44} = 0$ とおくと、

$$\begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (4.47)$$

であり、これは、3次元アフィン変換

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} a_{14} \\ a_{24} \\ a_{34} \end{bmatrix} \quad (4.48)$$

と等価である。数学的には、アフィン変換は射影変換の部分群である。

4.2.2.2 4次元等積1次変換による非特異射影変換の表現

同次座標系では、平行なベクトル同士は同一視されるため、各ベクトルを表現する上で定数倍の不定性がある。つまり、式(4.44)においてベクトル X, X' に任意の零でない定数を掛けてもそのベクトルがもつ意味が変化しない。変換行列 A に関しても定数倍の不定性がある。つまり、行列 A を任意の0でない定数 λ 倍した行列 λA も同一の射影変換を表現する。この事実から、変換を等積1次変換 \mathcal{EL}_4 に制限しても、任意の非特異射影変換を表すための表現の自由度は失われない。実際、定数 λ を次式とすれば、変換 A を等積1次変換 $\bar{A} = \lambda A$ に変換できる。

$$\lambda = \frac{1}{\sqrt[4]{|\det A|}} \quad (4.49)$$

このとき、 $|\det \bar{A}| = 1$ である。行列 A は、正則であると仮定しているので、この正規化は常に可能である。以上から、1次変換を等積1次変換に制限してもすべての非特異射影変換を表現可能であることが明らかになった。この事実を精密化すると、次の命題になる。

命題 4.10 3次元実射影空間 P^3 上の非特異射影変換 PGL_3 は、4次元同次座標表現における等積1次変換 \mathcal{EL}_4 によってちょうど2通りに表現される□

n 次元実射影空間の射影変換は、 $(n+1)$ 次元の等積1次変換 \mathcal{EL}_{n+1} と1対2に対応する。表現のこの2義性は、 \bar{A} に対して、 $-\bar{A}$ も同一の射影変換を表現することの反映である。4次元同次座標表現の場合は、 $\det(\bar{A}) = \det(-\bar{A})$ である。したがって、変換行列を特殊1次変換 SL_4 まで制限すると、 $\det(A) < 0$ である射影変換が表現出来ない。したがって、表現に2義性は残るものの、等積1次変換 \mathcal{EL}_4 が必要である。4次元等積1次変換は、一般の4次元1次変換を用いた従来の表現方法に比較して、3次元非特異射影変換に対する冗長度の少ない表現になっている。

この事実と、前節で展開した可逆アフィン変換の理論を用いると、任意の非特異射影変換が可逆演算によって実現できる。

定理 4.4 整数変数で表現された同次座標において、任意の非特異射影変換 PGL_n を近似する1対1写像(可逆写像)が存在する。□

4.2.3 非特異射影変換に対する可逆整数近似変換の構成法

本節では、一般の非特異射影変換と、その特殊な場合である並進、回転、透視変換などに対して、可逆整数近似写像の構成法を述べる。

4.2.3.1 一般の非特異射影変換に対する可逆整数近似変換の構成法

第4.1.4.1節に述べた等積1次変換の基本変換への分解理論を用いる。すると、

命題 4.11 一般の4次元等積1次変換 \mathcal{EL}_4 の表現行列 \bar{A} ($|\det \bar{A}| = 1$) は、15個の基本斜交軸変換 $P_{ij}[s]$ と有限個の反転変換 Q_i, R_{ij} の積に分解される。□

これと、命題 4.10 を用いると次の命題が成り立つ。

命題 4.12 一般の3次元非特異射影変換 PGL_3 は、その4次元同次座標表現において、15個の基本斜交軸変換 $p_{ij}[s]$ と有限個の反転変換 q_i, r_{ij} の積に分解できる。□

命題 4.11 と命題 4.12 は、一般のパラメータを持つ \mathcal{EL}_4, PGL_3 の変換に対して成り立つ。次節で例を述べるが、特殊な変換に対しては、基本斜交軸変換の数は、15以下にできる。

こうして得られた分解列を第4.1.4.2節に述べた整数近似の方法を用いて離散化すると、全体として可逆な整数近似写像が得られる。

4.2.3.2 代表的な非特異射影変換に対応する等積1次変換

本節では、読者の理解を助けるため、幾何情報処理に用いられる各種の要素的変換処理が、4次元等積1次変換によって具体的にどう実現されるのかをまとめる。

3次元並進運動 ベクトル $\mathbf{h} = (h_1, h_2, h_3)^T$ だけの3次元並進運動は、次の行列で表現される。

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & h_1 \\ 0 & 1 & 0 & h_2 \\ 0 & 0 & 1 & h_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.50)$$

この場合、 $\det \mathbf{A} = 1$ であるから、すでに4次元等積1次変換である。この行列は、3個の基本行列に分解できる。

$$\mathbf{A} = P_{14}[h_1]P_{24}[h_2]P_{34}[h_3] \quad (4.51)$$

3次元回転運動 非同次座標における3次元回転行列を

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad (4.52)$$

とすると $\mathbf{R}^T \mathbf{R} = \mathbf{I}$ かつ $\det \mathbf{R} = 1$ である。同次座標における表現行列は、次の形である。

$$\mathbf{A} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & 0 \\ r_{21} & r_{22} & r_{23} & 0 \\ r_{31} & r_{32} & r_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.53)$$

この行列は、 $\det \mathbf{A} = 1$ であるから、すでに4次元等積1次変換である。一般の回転変換に関しては、等積1次変換と同じく、最低15個の斜交軸変換が必要である。

各軸方向3次元スケーリング x, y, z 軸方向にそれぞれ s_1, s_2, s_3 倍するスケーリング変換の表現行列は、次の形である。

$$\mathbf{A} = \begin{bmatrix} s_1 & 0 & 0 & 0 \\ 0 & s_2 & 0 & 0 \\ 0 & 0 & s_3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.54)$$

これを等積1次変換に正規化すると次の形になる

$$\bar{\mathbf{A}} = \frac{1}{\sqrt[4]{|\det \mathbf{A}|}} \mathbf{A} = \begin{bmatrix} \frac{s_1}{\sqrt[4]{s_1 s_2 s_3}} & 0 & 0 & 0 \\ 0 & \frac{s_2}{\sqrt[4]{s_1 s_2 s_3}} & 0 & 0 \\ 0 & 0 & \frac{s_3}{\sqrt[4]{s_1 s_2 s_3}} & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt[4]{s_1 s_2 s_3}} \end{bmatrix} \quad (4.55)$$

3次元せん断変形 (斜交軸変換) 3次元空間の斜交軸変換は、幾何学的には各軸方向へのせん断変形である。4次元同次座標では、次の4次行列で表現される。

$$\mathbf{A} = \begin{bmatrix} 1 & K_{12} & K_{13} & 0 \\ 0 & 1 & K_{23} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.56)$$

この場合は、 $\det \mathbf{A} = 1$ であるからこのまま4次元等積1次変換として用いることができる。

3次元反転変換 座標平面 xy, yz, zx に関する3次元反転変換は、それぞれ次の4次行列によって表現される。

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.57)$$

これらは、それぞれ、等積基本変換 $\mathbf{Q}_3, \mathbf{Q}_1, \mathbf{Q}_2$ そのものである。

透視変換 透視変換は3次元の1次変換では得られない。透視変換の一般形は次の行列で与えられる。

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ p_1 & p_2 & p_3 & 1 \end{bmatrix} \quad (4.58)$$

この場合も $\det \mathbf{A} = 1$ であるから、このままで4次元等積1次変換の行列である。この場合は、3個の斜交軸変換の積に分解できる。

$$\mathbf{A} = \mathbf{P}_{41}[p_1]\mathbf{P}_{42}[p_2]\mathbf{P}_{43}[p_3] \quad (4.59)$$

以上述べてきたように、3次元スケール変換を除き、従来の表現方法を4次元等積1次変換による実現にそのまま移し換えることが可能である。3次元スケール変換の場合も、簡単な正規化の操作を加えれば4次元等積1次変換に変換できる。

以下、この分解とそれに基づいた可逆座標変換を3次元空間の点、平面、2次曲面、直線に対して実行するための具体的方法を述べる。

4.2.4 各種幾何要素の変換

点、平面、任意の代数曲面、直線、有理曲線などは、計算機で3次元情報を扱うための基本的な幾何要素である。これらの幾何要素は、3次元射影空間（4次元同次座標空間）の座標変換に関するテンソル量によって表現できる。任意のテンソル量に対して、任意の非特異射影変換の可逆整数近似が可能である。ここでは、幾何要素のうちで、点、平面、2次曲面、直線について、可逆整数近似変換の実現方法を述べる。

以下、次の記法を用いる。

行列 $\mathbf{A} \in \mathcal{EL}_4$ を基本変換へ分解した結果が次式で表されているものとする。

$$\mathbf{A} = \mathbf{E}_h \mathbf{E}_{h-1} \cdots \mathbf{E}_1 \quad (4.60)$$

行列 E_k は、基本行列 $P_{ij}[s]$, Q_i , R_{ij} のうちのいずれかを表す。

行列 E_k の定める写像を α_k とする。分解結果 (4.60) に対する写像列を次式で表す。

$$\alpha = \alpha_h \alpha_{h-1} \cdots \alpha_1 \quad (4.61)$$

写像 α_k の 1 対 1 整数近似を $\hat{\alpha}_k$ と記す。すると、分解列は次式である。

$$\hat{\alpha} = \hat{\alpha}_h \hat{\alpha}_{h-1} \cdots \hat{\alpha}_1 \quad (4.62)$$

この写像は、1 対 1 可逆写像である。写像 $\hat{\beta}$ の逆写像を $\hat{\beta}^{-1}$ と記す。すると、逆写像の分解列は、次式である。

$$\hat{\alpha}^{-1} = \hat{\alpha}_1^{-1} \hat{\alpha}_2^{-1} \cdots \hat{\alpha}_h^{-1} \quad (4.63)$$

点の変換 整数変数で表現された同次座標 $X = (X, Y, Z, W)^T \in Z^4$ を考える。

点 X の可逆整数近似変換は、次の写像の計算に帰着される。

$$X' = \hat{\alpha}_h \hat{\alpha}_{h-1} \cdots \hat{\alpha}_1 X \quad (4.64)$$

つまり、整数近似基本変換 $\hat{\alpha}_k$ を $k = 1, 2, \dots, h$ の順に座標ベクトルに施す演算である。

この逆変換は次式で計算される。

$$X = \hat{\alpha}_1^{-1} \hat{\alpha}_2^{-1} \cdots \hat{\alpha}_h^{-1} X' \quad (4.65)$$

平面の変換 3次元射影幾何学においては、平面は双対ベクトル空間の元で表現することができる。同次座標系における平面の表現は、整数係数ベクトル $D = (D_X, D_Y, D_Z, D_W)^T \in Z^4$ によって定まる同次 1 次式 f_1 である。

$$f_1(D; X) \stackrel{\text{def}}{=} D_X X + D_Y Y + D_Z Z + D_W W = 0 \quad (4.66)$$

この平面に対する射影変換は、係数ベクトル D に対する次の 1 次変換で実現できる。

$$D' = (A^{-1})^T D \quad (4.67)$$

平面の可逆整数近似変換は、係数ベクトル D に、整数近似基本変換を次式に基づいて施す。

$$D' = \hat{\alpha}_1^{-1} \hat{\alpha}_2^{-1} \cdots \hat{\alpha}_h^{-1} D \quad (4.68)$$

この逆変換は、次式で計算される。

$$D = \hat{\alpha}_h \hat{\alpha}_{h-1} \cdots \hat{\alpha}_1 D' \quad (4.69)$$

2次曲面の変換 2次曲面の方程式は、次に定義される2次形式である。

$$f_2(Q; X) \stackrel{\text{def}}{=} X^T Q X = 0 \quad (4.70)$$

行列 $Q = (Q_{ij}) \in Z^4 \times Z^4$ は、整数を成分とする対称行列である。このとき、射影変換は、この対称行列に対する次の変換である。

$$Q' = (A^{-1})^T Q A^{-1} \quad (4.71)$$

行列 Q に対する整数近似基本変換は、次式である。

$$Q' = \hat{\alpha}_1^{-1} \left(\hat{\alpha}_2^{-1} \cdots \left(\hat{\alpha}_h^{-1} Q \hat{\alpha}_h^{-1} \right) \cdots \hat{\alpha}_2^{-1} \right) \hat{\alpha}_1^{-1} \quad (4.72)$$

ただし、 $\hat{\alpha}_k^{-1}$ を左からかける時は、行列の各列ベクトルに独立に作用し、右からかける時は、行列の各行ベクトルに独立に作用すると定義する。

直線の表現と変換 2点 $X_1 = (X_1, Y_1, Z_1, W_1)^T$, $X_2 = (X_2, Y_2, Z_2, W_2)^T$ を通る直線は、次式で表現される [24]。

$$LX = 0 \quad (4.73)$$

L は、次の反対称行列である。

$$L = \begin{bmatrix} 0 & p & -q & r \\ -p & 0 & s & -t \\ q & -s & 0 & u \\ -r & t & -u & 0 \end{bmatrix} \quad (4.74)$$

$$\begin{aligned} p &= \begin{vmatrix} Z_1 & W_1 \\ Z_2 & W_2 \end{vmatrix}, & q &= \begin{vmatrix} Y_1 & W_1 \\ Y_2 & W_2 \end{vmatrix}, & r &= \begin{vmatrix} Y_1 & Z_1 \\ Y_2 & Z_2 \end{vmatrix}, \\ s &= \begin{vmatrix} X_1 & W_1 \\ X_2 & W_2 \end{vmatrix}, & t &= \begin{vmatrix} X_1 & Z_1 \\ X_2 & Z_2 \end{vmatrix}, & u &= \begin{vmatrix} X_1 & Y_1 \\ X_2 & Y_2 \end{vmatrix} \end{aligned} \quad (4.75)$$

この行列 L による直線表現に対する射影変換 A は、

$$L' = (A^{-1})^T L A^{-1} \quad (4.76)$$

で実行することができる。 L' は、変換後の直線の表現行列である。従って、この直線表現に対しても2次曲面と同様な方法で可逆整数近似変換が可能である。

$$L' = \hat{\alpha}_1^{-1} \left(\hat{\alpha}_2^{-1} \cdots \left(\hat{\alpha}_h^{-1} L \hat{\alpha}_h^{-1} \right) \cdots \hat{\alpha}_2^{-1} \right) \hat{\alpha}_1^{-1} \quad (4.77)$$

4.3 まとめ

整数座標で表現された3次元幾何要素に対する可逆座標変換を可能にする基本計算理論を提案した。その成果を以下にまとめる。

1. 非同次整数座標表現において、任意の等積アフィン変換が可逆計算可能なこと。その計算手順の構成アルゴリズムを示したこと。
2. 同次整数座標表現において、任意の非特異射影変換が可逆計算可能なこと。その計算手順の構成アルゴリズムを示したこと。
3. 同次整数座標表現において、任意のテンソル量が可逆計算可能なこと。
4. 幾何情報処理における基本プリミティブである点、平面、直線、代数曲面、有理曲線などが、4次元同次座標表現において任意の非特異射影変換が可逆計算可能なこと。

本章の理論は、数学的には、任意次元の射影空間におけるテンソル量の変換を、可逆性を保って離散近似する理論である。

本理論は基本計算原理とその構成アルゴリズムを提供する。その具体的な実現方法は、それをソフトウェアによって実現するのか、ハードウェアによる演算回路として実現するのかによって大きく異なる。今後は、これら個々の実現方法についての検討が必要になるであろう。

第 5 章

結論と展望

本論文の研究成果を要約し、今後の課題を述べる。

第 1 章では、計算機に 3 次元世界の、物理的、幾何的属性の推定を実現するコンピュータ・ビジョンと、推定された 3 次元幾何情報の認識に必要な 3 次元物体モデルの計算機表現、操作、推論などを含んだ「視覚情報処理」の体系化の必要性を強調した。Marr の述べた意味での計算理論、アルゴリズム、実現方法の 3 つの水準を意識した体系的な研究の重要性を述べた。特に、従来の個別機能モジュールの統合によって複雑な視覚機能を実現しようとするアプローチの問題点を指摘した。そして、アルゴリズムと実現方法のレベルでの方法論が先行している現状を指摘し、計算理論のレベルでの統合理論の必要性を強調した。それを踏まえて、従来のコンピュータビジョンの計算理論に欠けている機能として、複数物体、複数表面、複数運動などの多重属性の推定を指摘した。さらに、推定データあるいはモデルとして蓄積された、3 次元形状情報とその 3 次元運動情報を数学的な幾何変換としてデジタル計算機上に実現する時に問題となる、演算誤差による情報の損失現象を指摘した。

本論文では、画像面上の 2 次元的情報の推定を初期視覚、2 次元情報から、3 次元形状と 3 次元運動を推定する部分を中期視覚、復元されたデータに変換、推論操作をする段階を高次視覚と呼ぶことにした。この 3 つの段階のそれぞれに以下の基本的な貢献をおこなった。

第 2 章では、初期視覚の 2 種類の計算理論を多重属性に拡張した。一つは、動画像から複数の 2 次元運動の抽出に必要な多重運動視差 (多重オプティカルフロー) である。もう一つは、ステレオビジョンにおける互いに重なり合う多重表面の検出に必要な、多重ステレオ視差である。これらの基礎となる基本拘束方程式を導いた。この基本拘束方程式に基づいて、多重オプティカルフローと多重ステレオ視差の一撃計算アルゴリズムを導出し、シミュレーション実験を行った。ここで用いた技法は汎用の数理技法であり、数学的には、重ね合わせの原理と呼ばれる。コンピュータビジョンの種々の問題を一般的に記述、拡張するための演算子表記法も提案した。

第3章では、中期視覚の2種類の計算理論を多重属性に拡張した。一つは、2画像間の特徴点対対応集合から、画像中の複数剛体の3次元運動と3次元形状を復元するための計算理論である。もう一つは、距離データが与えられた時に、ある滑らかさ拘束を満たす複数の重なり合った表面を復元する計算理論である。これらの計算理論も第2章で提案した演算子表記法によって統一的に記述できる。複数の3次元運動復元に関しては、ノイズに対してロバスト性を保ちつつ、閉形式の数値計算で実現可能なアルゴリズムを導出した。多重表面復元に関しては、超並列型計算機に適した緩和型アルゴリズムを導出し、シミュレーション実験を行った。

第4章では、高次視覚をデジタル計算機に実現するための基本演算である座標変換法を改良した。従来の座標変換手法は、デジタル計算の離散有限語長の性質のために、変換を完全に可逆に行うことができなかった。この問題に対して、完全に可逆な変換を可能にする基本演算理論を構築した。この理論によれば、非同次座標表現における任意の等積アフィン変換が可逆的に実現できる。さらに、同次座標表現を用いれば任意の非特異射影変換が可逆的に実現できる。従来、計算機内の幾何情報に対して行われてきた主要な座標変換が、完全に可逆的に実行できることがこれによって明らかになった。

本研究では、問題を一貫して数理的に扱う計算論的アプローチをとった。また、従来の様な、いろいろな個々の問題を別個に解析するのではなく、多くの問題に共通に適用できる汎用の数理構造を見出し、それをいろいろな基本問題に適用し検証するという研究方法をとった。コンピュータビジョンにおける種々の基本拘束方程式を演算子表現法を用いて統一した形式で記述するという挑戦的な試みも行った。通常、この分野で行われる、特定の実画像列や画像集合に対して、性能を評価したりパフォーマンスを追求することは行わなかった。従来のコンピュータビジョンの研究では、本論文のアプローチはきわめて特異である。しかし、本研究は、これらの特徴をもった最初の試みである。

本論文の第2章と第3章で行った演算子を用いた体系化は、本論文のアイデアに少なからぬ影響を与えた量子力学を手本としている。これは、電気工学における Heaviside の演算子法と同様の役割を果たすことが期待される。画像処理の分野では、数理形態学 (mathematical morphology) と呼ばれる体系が、画像処理の種々のアルゴリズムを統一的に記述する方法として近年注目を集めている。しかし、コンピュータ・ビジョンにおいては、このような体系化の例は存在しなかった。いままでは、エッジ検出、方向検出、ステレオ視、運動視、運動からの3次元構造復元、物体認識など、個々の問題をそれぞれ独立に研究するというスタイルが一般的であった。その意味で本研究は、体系化の先鞭としての価値があると考えられる。

この演算子形式の理論は、まだその緒についたばかりである。以下に、今後の課題を列挙

する。

- ビジョンのより広範な問題に対する演算子形式表現の研究 ビジョンにおいては、ほとんどすべての問題が光学的、幾何学的、力学的拘束方程式と、これら拘束と画像データ（あるいはそれから計算される特徴点などの2次的データ）との適合性を表す関数の最適化問題として定式化されている。従って、本論文の個々の章において述べた様に、他の多くの基本拘束式を演算子形式で記述することを試みる必要がある。
- 統計的不偏最適性を持ったアルゴリズムの統一的導出法の研究 本論文では、理論の検証を主目的としたため、パラメータの推定法は単純な最小2乗法を用いたものが基本になっている。しかし、データの誤差分布を考慮した統計的に不偏で最適な推定方法を演算子形式と統合することが必要である。誤差に対する最適性に関しては、近年、一部の研究者によってその重要性が認識されはじめており、一般的統一的な試みも行われるようになってきた。これらとの統合を考える必要がある。
- 時間的統合に関する研究 時系列データのオンライン最小2乗法であるカルマンフィルタは、コンピュータビジョンの標準的な数値計算手法である。この手法に代表される時間を陽に含むオンラインの最適化を演算子形式で表現する方法を確立する必要がある。
- 異種視覚モジュール統合のための計算理論への発展 本論文では、演算子形式を用いて同種の複数属性の推定問題を定式化した。一方、異種の視覚モジュールを統合してより高性能な視覚システムを作り上げるというモジュール統合の必要性が従来から指摘されている。拘束方程式レベルで、この統合(統一)を行うための基本理論体系に発展させる必要がある。

本研究は、コンピュータビジョンと幾何情報処理を目的として行われた。しかし、ここで提示された数理的アイデアは、これらの分野にとどまらず、類似した計算手法を用いる他の分野にも貢献する可能性がある。例えば、重ね合わせの原理を用いると、関数近似やニューロコンピューティングにおいて、多価写像の学習が効率的に実現される可能性がある [181]。一般次元の離散可逆座標変換は、数値計算や、物理シミュレーションに新手法を提供する。

謝辞

本論文をまとめるにあたり、東京大学工学部 精密機械工学科 大園 成夫 教授には、終始懇切丁寧なご指導とご助言を賜りました。私の在学中の御教授も併せ、謹んで謝意を表します。また、東京大学工学部 計数工学科 杉原 厚吉 教授、同 電子情報工学科 原島 博 教授、同 高増 潔 助教授、同 精密機械工学科 鈴木 宏正 助教授 には、本論文の内容について御指導、御助言を賜りました。厚く御礼申し上げます。

本研究の内容と進め方について、次の方々から貴重な御教示とお励ましを頂きました。深く謝意を表します。米国マサチューセッツ工科大学 Tomaso Poggio 教授、Berthold K. P. Horn 教授、Edward Adelson 準教授、イリノイ大学 Thomas Huang 教授、同 Narendra Ahuja 教授、英国ケンブリッジ大学 John G. Daugman 教授、群馬大学 金谷 健一 教授。

本研究は、NTT (日本電信電話株式会社) 複合通信研究所、同 ヒューマンインタフェース研究所、ATR (国際電気通信基礎技術研究所) 通信システム研究所、同 視聴覚機構研究所、同 人間情報通信研究所において行われたものである。

本研究の機会を与えていただいた次の方々には深く感謝致します。NTT 複合通信研究所 元所長 松田 亮一 博士 (現 ソニー (株))、同 画像通信研究部 元部長 杉山 精 博士 (現 東京工芸大学教授)、同 画像処理研究室 室長 小杉 信 博士 (現 武蔵工業大学教授)、同 中項目リーダー 山田 豊通 博士 (現 NTT 研究開発本部)、NTT ヒューマンインタフェース研究所 元所長 塚田 啓一 博士 (現 松下電器産業 (株))、同 元所長 釜江 尚彦 博士 (現 ヒューレットパッカード (株) 日本研究所)、同 視覚情報研究部 元部長 小森 和昭 氏 (現 NTT インテリジェントテクノロジー (株) 社長)、同 元部長 小林 幸雄 博士 (現 千葉工業大学教授)、マルチメディア処理研究部 元部長 遠藤 隆也 氏 (現 NTT アドバンステクノロジー (株))、元視覚認識グループリーダー 酒井 高志 氏 (現 知能ロボット研究部長)、同 石井 健一郎 博士 (現 NTT 基礎研究所 情報科学研究部長)、ATR 通信システム研究所、視聴覚機構研究所、人間情報通信研究所 各会長 葉原 耕平 博士、ATR 通信システム研究所 社長 寺島 信義 博士、同 知能処理研究室 室長 岸野 文郎 氏、ATR 視聴覚機構研究所 前社長 淀川 英司 博士 (現 工学院大学教授)、同 認知機構研究室 前室長 下原 勝憲 博士 (現 人間情報通信研究所)、ATR 人間情報通信研究所 社長 東倉 洋一 博士、同 第2研究室室長 赤松 茂 博士。

次の諸氏には、直接の上司として、研究上の御教示を頂きました。謹んで謝意を表します。NTT ヒューマンインタフェース研究所 鈴木 智 博士、同 奥平 雅士 博士、同 間瀬 健二 博士、同 大谷 淳 博士(現 ATR 通信システム研究所)。

さらに、次の諸氏には、常日頃からご討論頂きました。心から謝意を表します。NTT ヒューマンインタフェース研究所 梅田 三千雄 博士(現 大阪電気通信大学教授)、同 金子 博 博士(現 東邦大学教授)、同 末永 康仁 博士、同 塩 昭夫 博士、同 武川 直樹 博士(現 NTT データ通信(株))、同 尺長 健 博士、同 上田 修功 博士(現 NTT コミュニケーション科学研究所)、同 前田 英作 博士(現 NTT 基礎研究所)、同 島谷 明 博士、ATR 通信システム研究所 田中 弘美 博士(現 立命館大学教授)、同 Jiang Yu Zheng 博士(現 九州工業大学)、同 境野 英朋 氏(現 NTT ヒューマンインタフェース研究所)、ATR 視聴覚機構研究所 西田 眞也 博士(現 NTT 基礎研究所)、同 有村 浩一 氏(現 NTT 基礎研究所)、ATR 人間情報通信研究所 川戸 慎二郎 博士(現 三菱電機(株))、同 佐藤 雅昭 博士、同 安藤 広志 博士、同 五十嵐 治一 博士(現 リコー(株))、同 早川 秀樹 博士(現 大阪ガス(株))、同 尾島 修一 博士、同 杉本 晃宏 氏、同 磯 俊樹 氏(現 NTT ヒューマンインタフェース研究所)、同 水谷 博之 氏(現 東芝(株))、同 小椋 毅 氏(現 アルプス電気(株))。

本研究で行ったシミュレーションにご協力いただいた次の諸氏に、心からお礼申し上げます。(株)CSK 宅見 正 氏、同 吉川 徹 氏。

参考文献

[序論に関連する参考書]

- [1] Aloimonos J.Y. and Shulman S.: Integration of Visual Modules —An Extension of the Marr Paradigm—, Academic Press, Boston, MA(1989).
- [2] Aloimonos J.Y.: Qualitative, purposive, and active vision, Proc. 10 th International Conference on Pattern Recognition, 1, pp.346-360 (1990).
- [3] Ballard D.H. and Brown C.M.: “Computer Vision”, Prentice-Hall, Englewood Cliffs(1982).
- [4] Blake A. and Zisserman A.: Visual Reconstruction, MIT Press, Cambridge, MA(1987).
- [5] Clark J.J. and Yuille A.L.: Data Fusion for Sensory Information Processing Systems, Kluwer Academic Publishers, Boston, MA(1990).
- [6] Gibson J.J.: The Perception of Visual World, Houghton Mifflin Co., Boston, MA (1950).
- [7] Gibson J.J.: The Ecological Approach to Visual Perception, Houghton Mifflin Co., Boston, MA (1979).
- [8] Grimson W.E.L.: From Images to Surfaces: A Computational Study of the Human Early Visual System. MIT Press Cambridge MA (1981).
- [9] Horn B.K.P.: Robot Vision, MIT Press, Cambridge, MA (1986).
- [10] Jain R. and Binford T.O.: “Ignorance, myopia, and naiveté in computer vision systems,” *CVGIP: Image Understanding*, **53**, 1, pp.112-117 (1991).
- [11] Kanatani K.: Group Theoretical Methods for Image Understanding, Springer, Berlin (1990).
- [12] Kanatani K.: Geometric Computation for Machine Vision, Oxford Univ. Press, Oxford (1993).

- [13] Kanatani K.: Statistical Optimization for Geometric Computation: Theory and Practice, Lecture Note, Artificial Intelligence Laboratory, Department of Computer Science, Gunma University (1994).
- [14] Knuth D.E.: The Art of Computer Programming 2nd ed., Volume I, II, Addison-Wesley, New York (1981).
- [15] Marr D.: Vision —A Computational Investigation into the Human Representation and Processing of Visual Information—, Freeman, San Francisco (1982).
- [16] Preparata F.P. and Shamos M.I.: “Computational Geometry —An Introduction—”, Springer-Verlag, New York (1985).
- [17] Rogers D.F. and Adams J.A.: “Mathematical Elements for Computer Graphics”, McGraw-Hill, London (1976).
- [18] Rosenfeld A. and Kak A.C.: “Digital Picture Processing, 2nd Ed.”, Academic Press, New York (1982).
- [19] Serra J.: “Image Analysis and Mathematical Morphology”, Academic Press, New York (1982).
- [20] Ullman S.: The Interpretation of Visual Motion, MIT Press, Cambridge MA (1979).
- [21] 大津展之: パターン認識における特徴抽出に関する数理的な研究、電子技術総合研究所研究報告第818号 (1981).
- [22] 金谷健一: “3次元認識の数理”、森北出版、東京 (1990).
- [23] 出口光一郎: “画像と空間 — コンピュータビジョンの幾何学 —”, 昭晃堂 (1991).
- [24] 山口富士夫: “コンピュータディスプレイによる形状処理工学 [III]”, 日刊工業新聞社 (1988).

[多重オプティカルフロー]

- [25] Adelson E.H. and Movshon J.A.: “Phenomenal Coherence of Moving Visual Patterns”, *Nature*, **300**, pp.523-525(1982).
- [26] Adelson E.H. and Bergen J.R.: “Spatiotemporal Energy Models for the Perception of Motion”, *J. Opt. Soc. Am., A*, **2**, pp.284-299(1985).
- [27] Adiv G.: “Determining Three-dimensional Motion and Structure from Optical Flow Generated by Several Moving Objects”, *IEEE Trans. Pattern Anal. & Mach. Intell.*, **PAMI-7**, 4, pp.384-401(1985).

- [28] Anandan P.: "A Computational Framework and an Algorithm for the Measurement of Visual Motion", *Int. J. Comput. Vision*, **2**, pp.283-310(1989).
- [29] Bergen J.R., Burt P., Hingorani R. and Peleg S.: "Transparent-Motion Analysis," *Proc. 1st European Conference on Computer Vision*, Antibes, France, pp.566-569(1990).
- [30] Bergen J.R., Burt P., Hingorani R. and Peleg S.: "Computing Two Motions from Three Frames," *Proc. IEEE 3rd Int. Conf. on Comput. Vision*, Osaka, Japan, pp.27-32, IEEE Computer Society Press(December, 1990).
- [31] Bergen J.R., Burt P., Hingorani R. and Peleg S.: "A Three-Frame Algorithm for Estimating Two-Component Image Motion", *IEEE Trans. Pattern Anal. & Mach. Intell.*, **PAMI-14**, 9, pp.886-896(1992).
- [32] Black M.J. and Anandan P.: "Robust Dynamic Motion Estimation Over Time", *Proc. IEEE Conf. on Comput. Vision and Pattern Recognition*, Maui, HI, pp.289-295, IEEE Computer Society Press(June,1991).
- [33] Bülthoff H., Little J.J. and Poggio T.: "A Parallel Algorithm for Real-time Computation of Optical Flow", *Nature*, **337**, pp.549-553(1989).
- [34] Burt P., Hingorani R. and Kolczynski R.J.: "Mechanisms for Isolating Component Patterns in the Sequential Analysis of Multiple Motion", *Proc. IEEE Workshop on Visual Motion*, Princeton, NJ, pp.187-193 (October,1991).
- [35] Darrel T. and Pentland A.P.: "Robust Estimation of a Multi-layered Motion Representation", *Proc. IEEE Workshop on Visual Motion*, Princeton, NJ, pp.173-178 (October,1991).
- [36] Darrel T., Simoncelli E., Adelson E.H. and Pentland A.P.: "A Model of Transparent Motion Perception Using Layers", *Investigative Ophthalmology and Visual Science Supplement(ARVO 1992)*, **33**, p.1142(April,1992).
- [37] Daugman J.G.: "Uncertainty Relation for Resolution in Space, Spatial Frequency, and Orientation Optimized by Two-dimensional Visual Cortical Filters", *J. Opt. Soc. Am.*, **A**, **2**, pp.1160-1169(1985).
- [38] Daugman J.G.: "Pattern and Motion Vision without Laplacian Zero Crossings", *J. Opt. Soc. Am.*, **A**, **5**, pp.1142-1148(1987).
- [39] Daugman J.G.: "Complete Discrete 2-D Gabor Transforms by Neural Networks for Image Analysis and Compression", *IEEE Trans. Acoust. Speech, and Signal Processing*, **36**, pp.1169-1179(1988).

- [40] Fennema C.L. and Thompson W.: "Velocity Determination in Scenes Containing Several Moving Objects", *Comput. Graphics & Image Process.*, **9**, pp.301-315(1979).
- [41] Fleet D.J. and Jepson A.D.: "Computation of Component Image Velocity from Local Phase Information", *Int. J. Comput. Vision*, **5**, 1, pp.77-104(1990).
- [42] Gabor D.: "Theory of Communication", *J. IEE*, Part 3, **93**, pp.429-457(1946).
- [43] Geman S. and Geman D.: "Stochastic Relaxation, Gibbs Distribution, and the Bayesian Restoration of Images", *IEEE Trans. Pattern Anal. & Mach. Intell.*, **PAMI-6**, pp.721-741(1984).
- [44] Gibson J.J.: "Optical Motions and Transformations as Stimuli for Visual Perception", *Psychol. Rev.*, **64**, 5 (1957).
- [45] Girod B. and Kuo D.: "Direct Estimation of Displacement Histograms", *Proc. Image Understanding and Machine Vision*, Cape Cod, MA, pp.73-76(1989).
- [46] Grzywacz N.M. and Yuille A.L.: "A Model for the Estimate of Local Image Velocity by Cells in the Visual Cortex", *Proc.R.Soc.London*, **B239**, pp.129-161(1990).
- [47] Harris J.G., Koch C., Staats E. and Luo J.: "Analog Hardware for Detecting Discontinuities in Early Vision", *Int. J. Comput. Vision*, **4**, 3, pp.211-223(1990).
- [48] Heeger D.J.: "A Model for the Extraction of Image Flow", *J. Opt. Soc. Am.*, **A**, **4**, pp.1455-1471(1987).
- [49] Heeger D.J.: "Optical Flow Using Spatiotemporal Filters", *Int. J. Comput. Vision*, **1**, pp.279-302(1988).
- [50] Heitz F. and Bouthemy P.: "Multimodal Motion Estimation and Segmentation Using Markov Random Fields", *Proc. Int. Conf. on Pattern Recognition*, **1**, Atlantic City, NJ, pp.378-383, IEEE Computer Society Press(June 1990).
- [51] Horn B.K.P. and Schunck B.G.: "Determining Optical Flow", *Artif. Intell.*, **17**, pp.185-203(1981).
- [52] Jasinschi R., Rosenfeld A. and Sumi K.: "Perceptual motion transparency: the role of geometrical information," *J. Opt. Soc. Am.*, **A**, **9**, 11, pp.1865-1879 (1992).
- [53] Kearney J.K., Thompson W.B. and Boley D.L.: "Optical Flow Estimation: An Error Analysis of Gradient-based Methods with Local Optimization", *IEEE Trans. Pattern Anal. & Mach. Intell.*, **PAMI-9**, pp.229-244(1987).

- [54] Kersten D.: "Transparency and the Cooperative Computation of Scene Attributes" in *Computational Models of Visual Processing* (Landy M.S. and Movshon J.A. eds.), MIT Press, Cambridge, MA(1991).
- [55] Langley K., Fleet D.J. and Atherton T.: "Multiple Motions from Instantaneous Frequency", *Proc. IEEE Conf. on Comput. Vision and Pattern Recognition*, Champaign, IL, pp.846-849, IEEE Computer Society Press (June,1992).
- [56] Limb J.O. and Murphy J.A.: "Estimating the Velocity of Moving Images in Television Signals", *Comput. Graphics & Image Process.*, **4**, pp.311-327(1975).
- [57] Longuet-Higgins H.C. and Prazdny K.: "The Interpretation of a Moving Retinal Image", *Proc.R.Soc.London*, **B208**, pp.385-397(1980).
- [58] Lucas B.D. and Kanade T.: "An Iterative Image Registration Technique with an Application to Stereo Vision", *Proc. Image Understanding Workshop*, pp.121-130(1981).
- [59] Marr D. and Ullman S.: "Directional Selectivity and Its Use in Early Visual Processing", *Proc. R. Soc. London*, **B211**, pp.151-180(1981).
- [60] Murase H.: "Surface Shape Reconstruction of a Nonrigid Transparent Object Using Refraction and Motion", *IEEE Trans. Pattern Anal. & Mach. Intell.*, **PAMI-14**, **10**, pp.1045-1052(1992).
- [61] Murray D.W. and Buxton B.F.: "Scene Segmentation from Visual Motion Using Global Optimization", *IEEE Trans. Pattern Anal. & Mach. Intell.*, **PAMI-9**, **2**, pp.220-228(1987).
- [62] Nagel H.-H.: "Displacement Vectors Derived from Second-order Intensity Variations in Image Sequences", *Comput. Vision, Graphics, & Image Process.*, **21**, pp.85-117(1983).
- [63] Negahdaripour S. and Horn B.K.P.: "Direct Passive Navigation", *IEEE Trans. Pattern Anal. & Mach. Intell.*, **PAMI-9**, **1**, pp.168-176(1987).
- [64] Noest A.J. and Koenderink J.J.: "Visual Coherence Despite Transparency or Partial Occlusion", *Perception*, **19**, **3**, p.384 (1990).
- [65] Ogata M. and Sato T.: "Motion-detection model with two stages: spatiotemporal filtering and feature matching", *J. Opt. Soc. Am.*, **A**, **9**, **3**, pp.377-387(1992).
- [66] Oppenheim A.V.: "Generalized Superposition", *Inf. & Control*, **11**, pp.528-536(1967).

- [67] Oppenheim A.V., Schafer R.W. and Stockham T.G.: "Nonlinear Filtering of Multiplied and Convolved Signals", *Proc. IEEE*, **56**, 8, pp.1264-1291(1968).
- [68] Peleg S. and Rom H.: "Motion Based Segmentation", *Proc. Int. Conf. on Pattern Recognition*, **1**, Atlantic City, NJ, pp.109-113, IEEE Computer Society Press (June 1990).
- [69] Poggio T., Torre V. and Koch C.: "Computational Vision and Regularization Theory", *Nature*, **317**, pp.314-319(1985).
- [70] Porat M. and Zeevi Y.: "The Generalized Gabor Scheme of Image Representation in Biological and Machine Vision", *IEEE Trans. Pattern Anal. & Mach. Intell.*, **PAMI-10**, 4, pp.452-468(1988).
- [71] Rosenfeld A., Jasinschi R.(Eds.): Special Issue on Perceptual Transparency. *Spatial Vision*, **7**(2), (1993).
- [72] Schunck B.G.: "Image Flow Segmentation and Estimation by Constraint Line Clustering", *IEEE Trans. Pattern Anal. & Mach. Intell.*, **PAMI-11**, 10, pp.1010-1027(1989).
- [73] Singh A.: "An Estimation-theoretic Framework for Image-flow Computation", *Proc. IEEE 3rd Int. Conf. on Comput. Vision*, Osaka, Japan, pp.168-177, IEEE Computer Society Press(December, 1990).
- [74] Spoerri A. and Ullman S.: "The Early Detection of Motion Boundaries", *Proc. IEEE 1st Int. Conf. on Comput. Vision*, London, England, pp.209-218, IEEE Computer Society Press(1987).
- [75] Stoner G.R., Albright T.D. and Ramachandran V.S.: "Transparency and Coherence in Human Motion Perception", *Nature*, **344**, pp.153-155(1990).
- [76] Terzopoulos D.: "Regularization of Inverse Visual Problems Involving Discontinuities", *IEEE Trans. Pattern Anal. & Mach. Intell.*, **PAMI-8**, pp.413-424(1986).
- [77] Tsukune H., Aggarwal J.K.: "Analyzing orthographic projection of multiple 3D velocity vector fields in optical flow", *CVGIP-42*, pp.157-191.
- [78] Watson A.B. and Ahumada A.J.: "Model of Human Visual-motion Sensing", *J. Opt. Soc. Am.*, **A**, **2**, pp.322-342(1985).
- [79] Yachida M.: "Determining Velocity Maps by Spatio-temporal Neighborhoods from Image Sequence", *Comput. Vision, Graphics, & Image Process.*, **21**, pp.262-279(1983).

- [80] Yuille A.L. and Grzywacz N.M.: "A Mathematical Analysis of the Motion Coherence Theory", *Int. J. Comput. Vision*, **3**, pp.155-175(1989).
- [81] 鈴木教洋、吹抜敬彦: "動画像の速度に関する二つの定理の等価性について," 信学論 D、**J70-D**, 9, pp.1828-1829(1987).
- [82] 大津展之: "不変特徴抽出の理論," [I-IV], 電子通信学会誌、**69**, 5, pp.469-475, 6, pp.585-590, 7, pp.722-727, 8, pp.831-837 (1986).

[多重ステレオ視差]

- [83] Akerstrom, R.A. and Todd, J.T.: "The perception of stereoscopic transparency," *Percept. & Psychophys.* 44(5): pp.421-432 (1988).
- [84] Baker, H.H. and Binford, T.O.: "Depth from edge and intensity based stereo," *Proc. 7th Int. Joint Conf. Artif. Intell.*, Los Altos. CA, pp.631-636 (1981).
- [85] Burt, P.J. and Julesz, B.: "A disparity gradient limit for binocular fusion," *Science*, 208: pp.615-617 (1980).
- [86] Dhond, U.R. and Aggarwal, J.K.: "Structure from stereo — a review," *IEEE Trans. Sys. Man, & Cybern.* 19(6): pp.1489-1510 (1989).
- [87] Hubel, D.H. and Wiesel, T.N.: "Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex," *J. Physiol.* 160: pp.106-154 (1962).
- [88] Jones, D.G. and Malik, J.: "Computational framework for determining stereo correspondence from a set of linear spatial filters," *Image and Vision Comput.* 10(10): pp.699-708 (1992).
- [89] Julesz, B.: "Binocular depth perception of computer generated patterns," *Bell Sys. Tech. J.* 38: pp.1001-1020 (1960).
- [90] Koenderink, J.J. and van Doorn, A.J.: "Receptive field families," *Biol. Cybern.* 63: pp.291-298 (1990).
- [91] Little, J.J., Gillett, W.E.: "Direct evidence for occlusion in stereo and motion," *Image and Vision Comput.* 8(4): pp.328-340 (1990).
- [92] Lucas, B.D. and Kanade, T.: "An iterative image registration technique with an application to stereo vision," *Proc. Image Understanding Workshop*, pp.121-130 (1981).
- [93] Madarasmı, S., Kersten, D., and Pong, T.-C.: "A multi-layer approach to segmentation and interpolation with application to stereo vision," *Investigative Ophthalmology and Visual Science Supplement (ARVO 1992)* 33: p.1371 (1992).

- [94] Marr, D. and Poggio, T.: "A cooperative computation of stereo disparity," *Science*, 194: pp.283-287 (1976).
- [95] Marr, D. and Poggio, T.: "A theory of human stereopsis," *Proc. R. Soc. London*, B204: pp.301-328 (1979).
- [96] Okutomi, M. and Kanade, T.: "A locally adaptive window for signal matching," *Int. J. Comput. Vision*, 7(2): pp.143-162 (1992).
- [97] Pollard, S.B., Mayhew, J.E.W., and Frisby, J.P.: "PMF: A stereo correspondence algorithm using a disparity gradient limit," *Percept.*, 14: pp.449-470 (1985).
- [98] Pollard, S.B. and Frisby, J.P.: "Transparency and the uniqueness constraint in human and computer stereo vision," *Nature*, 347: pp.553-556 (1990).
- [99] Prazdny, K.: "Detection of binocular disparities," *Biol. Cybern.* 52: pp.93-99 (1985).
- [100] Weinshall, D.: "Seeing "ghost" planes in stereo vision," *Vision Res.* 31(10): pp.1731-1748 (1991).
- [多重 3 D 運動・構造復元]
- [101] Arun K.S., Huang T.S. and Blostein S.D.: "Least-squares fitting of two 3-D point sets," *IEEE Trans. PAMI*, 9, pp.698-700 (1987).
- [102] Boulton T.E. and Brown L.G.: "Factorization-based Segmentation of Motions," *Proc. IEEE Workshop on Visual Motion*, Princeton, NJ, pp.179-186 (Oct.,1991).
- [103] Bruss A.R. and Horn B.K.P.: "Passive Navigation," *Comput. Vision, Graphics and Image Process.*, 21, 1, pp.3-20 (1983).
- [104] Chen H.H. and Huang T.S.: "Maximal matching of 3-D points for multiple-object motion estimation," *Pattern Recognition*, 21, 2, pp.75-90 (1988).
- [105] Daniilidis K. and Nagel H.-H.: "Analytical results on error sensitivity of motion estimation from two views," *Image and Vision Computing*, 8, 4, pp.297-303 (1990).
- [106] Horn B.K.P.: "Relative Orientation," *International Journal of Computer Vision*, 4, pp.59-78 (1990).
- [107] Horn B.K.P.: "Closed-form solution of absolute orientation using unit quaternions," *Journal of Optical Society of America*, A-4, pp.629-642 (1987).
- [108] Horn B.K.P., Hilden H.M., and Negahdaripour S.: "Closed-form solution of absolute orientation using orthonormal matrices," *Journal of Optical Society of America*, A-5, pp.1128-1135 (1988).

- [109] Huang T.S. and Faugeras O.D.: "Some Properties of the E Matrix in Two-view Motion Estimation," *IEEE Trans. PAMI*, **11**, 12, pp.1310-1312 (1989).
- [110] Liou S.-P. and Jain R.C.: "An approach to three-dimensional image segmentation," *CVGIP: Image Understanding*, **53**, 3, pp.237-252 (1991).
- [111] Liu Y. and Huang T.S.: "A linear algorithm for motion estimation using straight line correspondences," *Comput. Vision, Graphics and Image Process.*, **44**, pp.35-57 (1988).
- [112] Longuet-Higgins H.C. and Prazdny K.: "The Interpretation of a Moving Retinal Image," *Proc. Royal Society of London*, **B208**, pp.385-397 (1980).
- [113] Longuet-Higgins H.C.: "A Computer Program for Reconstructing a Scene from Two Projections," *Nature*, **392**, pp.133-135 (1981).
- [114] Longuet-Higgins H.C.: "The reconstruction of a scene from two projections — Configuration that defeat the 8-point algorithm," *Proc. 1st IEEE Conf. on Artif. Intell. Applications*, Denver, CO, pp.395-397 (December 1984).
- [115] Longuet-Higgins H.C.: "Multiple interpretations of a pair of images of a surface," *Proc. Royal Society of London*, **A418**, pp.1-15 (1988).
- [116] Negahdaripour S. and Horn B.K.P.: "Direct Passive Navigation," *IEEE Trans. PAMI*, **9**, 1, pp.168-176 (1987).
- [117] Negahdaripour S.: "Multiple interpretations of the shape and motion of objects from two perspective images," *IEEE Trans. PAMI*, **12**, 11, pp.1025-1039 (1990).
- [118] Spetsakis M.E. and Aloimonos J.Y.: "Optimal computing of structure from motion using point correspondences in two frames," *Proc. IEEE 2nd Int. Conf. Comput. Vision*, Tampa, FL, pp.449-453 (December, 1988).
- [119] Spetsakis M.E. and Aloimonos J.Y.: "Structure from Motion Using Line Correspondences," *Int. J. Comput. Vision*, **4**, pp.171-183 (1990).
- [120] Tomasi C. and Kanade T.: "Shape and motion from image streams under orthography: a factorization method," *Int. J. Comput. Vision*, **9**, 2, pp.137-154 (1992).
- [121] Tsai R.Y. and Huang T.S.: "Uniqueness and estimation of three-dimensional motion parameters of rigid objects with curved surfaces," *IEEE Trans. PAMI*, **6**, pp.13-27 (1984).
- [122] Ullman S.: "The Interpretation of Structure from Motion," *Proc. Royal Society of London*, **B203**, pp.405-426 (1979).

- [123] Weng J., Ahuja N., and Huang T.S.: "Optimal motion and structure estimation," *Proc. IEEE CVPR '89*, San Diego, CA, pp.144-152 (June, 1989).
- [124] Weng J., Huang T.S., and Ahuja N.: "Motion and Structure from Two Perspective Views: Algorithms, Error Analysis, and Error Estimation," *IEEE Trans. PAMI*, **11**, 5, pp.451-476 (1989).
- [125] Weng J., Huang T.S., and Ahuja N.: "Motion and structure from line correspondences: closed-form solution, uniqueness, and optimization," *IEEE Trans. PAMI*, **14**, 3, pp.318-336 (1992).
- [126] Zhang Z., Faugeras O.D., and Ayache N.: "Analysis of a Sequence of Stereo Scenes Containing Multiple Moving Objects Using Rigidity Constraints," *Proc. IEEE 2nd ICCV*, pp.177-186 (1988).
- [127] Zhuang X., Huang T.S., and Haralick R.M.: "Two-View Motion Analysis: A Unified Algorithm," *J. Opt. Soc. Am.*, **A-3**, pp.1492-1500 (1986).
- [標準正則化理論の拡張による多重表面の復元]
- [128] Ando H.: "A computational model for reconstructing multiple surfaces," *Investigative Ophthalmology & Visual Science (Supplement)*, 34(4): 1186 (1993).
- [129] Darrel T. and Pentland A.P.: "Robust estimation of multi-layered motion representation," *IEEE Workshop on Visual Motion*, Princeton, pp.173-178 (1991).
- [130] Geiger D. and Girosi F.: "Parallel and deterministic algorithms from mrfs: surface reconstruction," *IEEE Trans. PAMI*, **13**(5), pp.401-412 (1991).
- [131] Kirkpatrick S., Gelatt C.D., and Vecchi M.P.: "Optimization by simulated annealing," *Science*, **220**, pp.219-227 (1983).
- [132] Madarasmis S., Kersten D., and Pong T.-C.: "The computation of stereo disparity for transparent and for opaque surfaces," *In Advances in Neural Information Processing Systems*, **5**, Morgan Kaufman Publishers (1993).
- [133] Marroquin J.L.: "Random measure fields and the integration of visual information," *IEEE Trans. Sys., Man & Cybern.*, **22**, pp.705-716 (1992).
- [134] Poggio T. and Koch C.: "Ill-posed problems in early vision: from computational theory to analogue networks," *Proc. Royal Society of London*, **B226**, pp.303-323 (1985).
- [135] Poggio T., Gamble E., and Little J.J.: "Parallel integration of vision modules," *Science*, **242**, pp.436-440 (1988).

- [136] Poggio T. and Girosi F.: "Networks for approximation and learning," *Proceedings of IEEE*, 78(9), pp.1481-1497 (1990).
- [137] Poggio T. and Edelman S.: "A network that learns to recognize 3D Objects," *Nature*, 343, pp.263-266 (1990).
- [138] J.Y.A.Wang and E.H.Adelson: "Layered representation for motion analysis," *IEEE Conference on Computer Vision and Pattern Recognition*, New York, pp.361-366 (1993).
- [可逆座標変換]
- [139] Braccini C., Marino G.: "Fast Geometrical Manipulations of Digital Images," *Computer Graphics and Image Processing Vol.13*, pp.127-141 (1980).
- [140] Tanaka A., Kameyama M., Kazama S., Watanabe O.: "A Rotation Method for Raster Image Using Skew Transformation," *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Miami, pp.272-277 (1986).
- [141] Greene D.H. and Yao F.F.: "Finite-resolution Computational Geometry," *Proc. IEEE Annual Symposium on Foundations of Computer Science*, pp.143-152 (1986).
- [142] Chen H.H. and Huang T.S.: "A Survey of Construction and Manipulation of Octrees," *Computer Vision, Graphics, and Image Processing*, Vol.43, pp.409-431 (1988).
- [143] Weng J. and Ahuja N.: "Octree Representation of Objects in Arbitrary Motion," *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Francisco, pp.524-529 (1985).
- [144] Osse W.M. and Ahuja N.: "Efficient Octree Representation of Moving Objects," *7th IEEE International Conference on Pattern Recognition*, Montreal, Canada, pp.821-823 (1984).
- [145] Meagher D.: "Geometric Modeling Using Octree Encoding," *Computer Vision, Graphics, and Image Processing*, 19, pp.129-147 (1982).
- [146] Paeth A.W.: "A Fast Algorithm for General Raster Rotation," *Graphics Interface '86*, pp.77-81 (1986).
- [147] Weiman C.F.R.: "Continuous Anti-aliased Rotation and Zoom of Raster Images," *ACM SIGGRAPH '80*, pp.286-293 (1980).
- [148] Catmul E. and Smith A.R.: "3-D Transformations of Images in Scanline Order," *ACM SIGGRAPH '80*, pp.279-285 (1980).

- [149] Matula D.W. and Kornerup P.: "Finite Precision Rational Arithmetic: Slash Number Systems," *IEEE Trans. Computer*, Vol.C-34, pp.3-18 (1985).
- [150] Sugihara K. and Iri M.: "Two design principles of geometric algorithms in finite-precision arithmetic," *Applied Mathematics Letters* 2, pp.203-206 (1989).
- [151] 杉原厚吉, 伊理正夫: "計算誤差による暴走の心配のないソリッドモデラの提案," *情報処理学会論文誌*, 28, pp.962-974 (1987).
- [152] 杉原厚吉: "幾何アルゴリズムの数値的破綻とその対策," *応用数理*, 1, 4, pp.280-299 (1991).
- [153] 山本喜久, 上田正仁: "量子光学と新技術 [VII]," *電子情報通信学会誌*, 73, 1, pp.63-68 (1990-1).
- [154] 井宮淳: "計算の物理モデル," *情報処理学会誌*, 35, 4, pp.300-305 (1994).
- [155] 森田憲一: "計算における可逆性," *情報処理学会誌*, 35, 4, pp.306-314 (1994).
- [156] Fredkin E., Toffoli T.: "Conservative Logic," *International Journal of Theoretical Physics*, 21, 3/4, pp.219-253 (1982).
- [157] Feynman R.P.: "Quantum Mechanical Computers," *Opt. News*, 11, pp.11-20 (1985).
- [参考とした一般書籍]
- [158] Hartree D.R.: *The Calculation of Atomic Structures*, Wiley, New York(1957).
- [159] Huber P.J.: *Robust Statistics*, Wiley & Sons, New York(1981).
- [160] Penrose R.: *The Emperor's New Mind: Concerening Computers, Minds and the Laws of Physics*, Oxford Univ. Press, Oxford, U.K.(1989).
- [161] Sudbery A.: *Quantum Mechanics and the Particles of Nature*, Cambridge University Press (1986).
- [162] Tikhonov A.N. and Arsenin V.Y.: *Solutions of ill-posed problems*, W.H.Winston, Washington DC (1977).
- [163] クーラン、ヒルベルト (齊藤、銀林訳): *数理物理学の方法 2*、東京図書、東京 (1959).
- [164] 市田浩三、吉本富士市: *スプライン関数とその応用*, 教育出版 (1979).
- [165] 伊理正夫, 韓太舜: *テンソル解析入門*, 教育出版、東京 (1973).
- [166] 伊理正夫、阪田省二郎: *応用数学 3 = マトリックス*、倍風館、東京 (1982).

- [167] 近藤次郎: フーリエ変換とその応用、培風館、東京 (1975).
- [168] フォン・ノイマン (井上ほか訳): 量子力学の数学的基礎、みすず書房、東京 (1957).
- [169] 齊藤正彦: 線形代数入門、東京大学出版会 (1966).
- [170] 竹内外史: 線形代数と量子力学、裳華房、東京 (1971).
- [171] ファン・デル・ヴェルデン (前田訳): 代数幾何学入門、シュプリンガー・フェアラーク、東京 (1939,1991).
- [172] 山内恭彦, 杉浦光夫: 連続群論入門, 培風館 (1960).
- [173] 横沼健雄: テンソル空間と外積代数、岩波講座 基礎数学 3, 岩波書店 (1977)
- [174] 横田一郎: “群と位相”, 裳華房 (1971).

原著論文目録¹

学術論文誌

- [175] 志沢雅彦、加藤純一、大園成夫、高増潔: “デジタル PLL を用いたしま画像からの実時間形状復元処理,” 精密工学会誌、**53**, 2, pp.334-339(1987.2).
- [176] * 志沢雅彦: “整数座標系における可逆アフィン変換 — n 次元空間における一般理論—,” 電子情報通信学会論文誌 D-I, **J74-D-I**, 12, pp.805-813 (1991.12). (English translation: Shizawa M.: “Invertible affine transformations on integer coordinate system — General theory in n -dimensional space—,” Systems and Computers in Japan, 24(7), pp.1-12(1993).)
- [177] * 志沢雅彦, 間瀬健二: “多重オプティカルフロー: 基本拘束方程式と運動透明視・運動境界検出の統一計算理論,” 電子情報通信学会論文誌 D-II, **J76-D-II**, 5, pp.987-1005 (1993.5).
- [178] * 志沢雅彦: “運動立体視におけるトランスペアレンシー: 2 透視投影像における点対応からの準最適線形アルゴリズム,” 電子情報通信学会論文誌 D-II, **J77-D-II**, 2, pp.286-300 (1994.2).
- [179] * 志沢雅彦: “標準正則化理論の多価関数への拡張: なめらかな多重表面の復元,” 電子情報通信学会論文誌 D-II, **J77-D-II**, 6, pp.1146-1156 (1994.6).
- [180] * 志沢雅彦: “両眼立体視におけるトランスペアレンシーの計算理論と 2 重視差の一撃計算モデル,” 電子情報通信学会論文誌 D-II, **J77-D-II**, 7, pp.1245-1254 (1994.7).
- [181] 志沢雅彦: “多価正則化ネットワーク — 多対 h 写像を学習する多層ネットワークの理論 —,” 電子情報通信学会論文誌 A, 採録決定済.
- [182] 志沢雅彦、磯俊樹: “局所微分フィルタを用いた多重スケール・多重方向場の表現と検出,” 電子情報通信学会論文誌 A, 採録決定済.

¹文献番号の後に * を付した論文は、本論文の内容を含む。

国際会議 (コンファレンス)

- [183] * Shizawa M. and Mase K.: "Simultaneous Multiple Optical Flow Estimation," *Proc. 10th IAPR International Conference on Pattern Recognition (ICPR'90)*, **1**, Atlantic City, NJ, pp.274-278 (June,1990).
- [184] * Shizawa M.: "Discrete Invertible Affine Transformations," *Proc. 10th IAPR International Conference on Pattern Recognition (ICPR'90)*, **2**, Atlantic City, NJ, pp.134-139 (June,1990).
- [185] * Shizawa M. and Mase K.: "A Unified Computational Theory for Motion Transparency and Motion Boundaries Based on Eigenenergy Analysis," *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'91)*, Maui, HI, pp.289-295 (June,1991).
- [186] * Shizawa M.: "On Visual Ambiguities Due to Transparency in Motion and Stereo," *Proc. 2nd European Conference on Computer Vision (ECCV'92)*, Genova, Italy, pp.411-419 (May,1992).
- [187] * Shizawa M.: "Direct Estimation of Multiple Disparities for Transparent Multiple Surfaces in Binocular Stereo," *IEEE International Conference on Computer Vision (ICCV'93)*, Berlin, Germany, pp.447-454 (May,1993).
- [188] Shizawa M. and Iso T.: "Direct Representation and Detection of Multi-Scale, Multi-Orientation Fields Using Local Differentiation Filters," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'93)*, New York City, NJ, pp.508-514 (June,1993).
- [189] * Shizawa M.: "Transparent 3D Motions and Structures from Point Correspondences in Two Frames: A Quasi-optimal, Closed-form, Linear Algorithm and Degeneracy Analysis," *First Asian Conference on Computer Vision (ACCV'93)*, Osaka, Japan, pp.329-334 (November,1993).
- [190] Shizawa M.: "Regularization Networks for Approximating Multi-Valued Functions: Learning Ambiguous Input-Output Mappings from Examples," *IEEE International Conference on Neural Networks (ICNN'94)*, Orlando, FL, pp.137-142 (June-July,1994).
- [191] * Shizawa M.: "Reconstruction of multiple overlapping surfaces via standard regularization techniques," *The 12th International Conference on Pattern Recognition (ICPR'94)*, Jerusalem, Israel, (October,1994), in press.

国際会議 (シンポジウム, ワークショップ)

- [192] * Shizawa M.: "An Invertible Geometrical Transformation of n -dimensional Digital Pictures," *Proc. IEEE International Symposium on Information Theory (ISIT'88)*, Kobe, Japan, p.245 (June,1988).
- [193] * Shizawa M. and Mase K.: "Principle of Superposition —for Multiple Motion and Motion Transparency—,"(invited paper) *Proc. Workshop on Dynamic Scene Understanding (of International Joint Conference on Artificial Intelligence (IJCAI'91))*, Sydney, Australia (August,1991).
- [194] * Shizawa M. and Mase K.: "Principle of Superposition: A Common Computational Framework for Analysis of Multiple Motion," *Proc. IEEE Workshop on Visual Motion*, Princeton ,NJ, pp.164-172 (October,1991).
- [195] * Shizawa M.: "A Quasi-linear Single-shot Computational Model for Stereoscopic Pure Transparency," *Investigative Ophthalmology and Visual Science Supplement (ARVO'93 annual meeting)*, Sarasota, FL (May,1993).
- [196] Shizawa M.: "Networks for learning inverse models of nonlinear systems —Direct approximation of many-to- h input-output mappings—," The 2nd International Symposium on Inverse Problems in Engineering Sciences (国際研究集会「数理工学における逆問題 (IPES-94)」), Osaka, Japan pp.B38-B39 (1994.7).

その他の共著論文

- [197] 磯俊樹、志沢雅彦: "演算子形式を用いたウェーブレット表現からの多重方向検出の統一的方法," 画像の認識・理解シンポジウム (MIRU'92), 2, pp.29-36 (1992.7).
- [198] 磯俊樹、志沢雅彦: "重ね合わせの原理に基づく多重方向同時検出演算子," 電子情報通信学会秋季全国大会, D-291(1992.9).
- [199] 磯俊樹、志沢雅彦: "局所微分情報からの一撃的 L,T,X 交差検出," 電子情報通信学会パターン認識と理解研究会, PRU92-148 (1993.3).
- [200] 磯俊樹、志沢雅彦: "3次元物体投影像における交差形状解析 — 局所微分情報を用いた一撃的 Y 交差検出 —," 電子情報通信学会パターン認識と理解研究会, PRU93-41 (1993.7).
- [201] T. Iso and M. Shizawa: "Detecting L-, T-, and X-junctions from Low-order Image Derivatives," *Proc. SPIE Visual Communications and Image Processing '93*, 2094, Cambridge, MA (Nov. 1993).

- [202] 磯俊樹、志沢雅彦: “局所微分情報に基づく交差形状解析: 低次微分フィルターによる L, T, X 交差検出,” 電子情報通信学会論文誌 D-II, J76-D-II, 12, pp.2529-2536 (1993.12).
- [203] 加藤純一、藤田桂一、田中敏雄、志沢雅彦、古谷涼秋、大園成夫: “デジタル PLL を用いたしま画像からの実時間形状復元処理 (第 2 報),” 精密工学会誌、59, 9, pp.141-146 (1993.9).

翻訳、解説論文、書評

- [204] B.K.P.Horn 著 NTT ヒューマンインタフェース研究所プロジェクト RVT 訳 (共訳): ロボットビジョン — 機械は世界をどう見るか —, 朝倉書店 (1993).
- [205] 志沢雅彦: “書評: K.Kanatani 著 Geometric Computation for Machine Vision,” 情報処理学会誌、34, 8, pp.1098-1099 (1993.8).
- [206] * 志沢雅彦: “上手に拘束条件を設定するには: 多重視覚情報の抽出を例にして,” テレビジョン学会誌「最近の画像認識・理解技術」小特集、48, 2, pp.135-141 (1994.2).

研究会, 全国大会

- [207] * 志沢雅彦: “ n 次元デジタル画像の可逆的幾何変換法,” 情報処理学会グラフィクスと CAD 研究会資料, CAD27-4(1987.7).
- [208] * 志沢雅彦: “デジタル画像の可逆的アフィン変換 — 2 次元の最適な変換について —,” 情報処理学会第 35 回全国大会, 2G-1 (1987.9).
- [209] 志沢雅彦: “区分的等積 1 次変換 (PELT) を用いた不定形領域画像の格納方法,” 電子情報通信学会春季全国大会, D-223 (1988.3).
- [210] * 志沢雅彦: “4 次元同次座標系における可逆的座標変換,” 情報処理学会第 38 回全国大会, 4K-5 (1989.3).
- [211] 志沢雅彦, 奥平雅士: “3 次元運動視の基礎検討 — オプティカルフローの抽出 —,” 電子情報通信学会春季全国大会, D-555 (1989.3).
- [212] 志沢雅彦: “動画像からの物体の動き抽出法に関する検討,” 電子情報通信学会 画像理解の高度化と高速化シンポジウム, pp.83-88 (1989.4).
- [213] * 志沢雅彦, 間瀬健二: “時空間フィルタ法による多重オプティカルフローの抽出,” 電子情報通信学会秋季全国大会, D-161 (1989.9).

- [214] * 志沢雅彦, 間瀬健二: “時空間フィルタを用いた多重オブティカルフローの抽出法,” 情報処理学会コンピュータビジョン研究会資料, CV62-2 (1989.9).
- [215] 志沢雅彦, 間瀬健二: “可逆的座標変換の並列性に関する考察,” 情報処理学会第39回全国大会 (1989.10).
- [216] * 志沢雅彦, 間瀬健二: “多重オブティカルフロー抽出における多重度の判定,” 電子情報通信学会パターン認識と理解研究会, PRU89-96(1990.1).
- [217] * 志沢雅彦: “多重3次元剛体運動のパラメータ推定法,” 情報処理学会第41回全国大会, 4M-10 (1990.9).
- [218] 志沢雅彦: “アパーチャ問題を考慮した閉形式の時空間局所フロー抽出,” 電子情報通信学会秋季全国大会, D-439 (1990.10).
- [219] * 志沢雅彦: “多重3次元剛体運動の推定に向けて,” 情報処理学会コンピュータビジョン研究会資料, CV68-3 (1990.9).
- [220] 志沢雅彦, 他: “運動透過性と局所運動知覚モデルの妥当性,” 電子情報通信学会秋季全国大会, SD-10-3 (1991.9).
- [221] * 志沢雅彦, 間瀬健二, 他: “「重ね合わせの原理」に基づいた多重運動の解析,” 情報処理学会第43回全国大会, 7F-4 (1991.10).
- [222] * M.Shizawa: “Transparency in 3D Motion Vision: Fundamental Constraint Equations and Linear Analytical Solution”, 画像の認識・理解シンポジウム (MIRU'92), 1, pp.95-102.
- [223] 志沢雅彦, 他: “国際会議“ECCV'92”報告,” 電子情報通信学会パターン認識と理解研究会, PRU92-76 (1992.11).
- [224] * 志沢雅彦: “ステレオ視におけるトランスペアレンシーについて”, 電子情報通信学会パターン認識と理解研究会, PRU92-137 (1993.2).
- [225] * 志沢雅彦: “点对応からの運動立体視におけるトランスペアレンシー: 準最適かつ閉形式な線形アルゴリズム”, 電子情報通信学会パターン認識と理解研究会, PRU92-149 (1993.3).
- [226] 志沢雅彦, 他: “1993年コンピュータビジョンとパターン認識会議参加報告,” 情報処理学会コンピュータビジョン研究会, CV85-6 (1993.9).
- [227] * 志沢雅彦: “標準正則化理論の多価関数への拡張と多重表面の復元,” 電子情報通信学会ニューロコンピューティング研究会, NC93-34 (1993.9).

- [228] 志沢雅彦: “多価関数を近似する正則化ネットワークとその学習アルゴリズム,” 電子情報通信学会ニューロコンピューティング研究会, NC93-66 (1993.12).
- [229] * 志沢雅彦: “なめらかな多価関数によるデータ近似を行う標準正則化アルゴリズム,” 画像の認識理解シンポジウム (MIRU'94)、vol.I, pp.313-320 (1994.7).
- [230] * 志沢雅彦: “標準正則化手法による重なり合った複数表面の復元,” 電子情報通信学会秋季全国大会, D-335 (1994.9).
- [231] 志沢雅彦: “オプティカルフローからの多重運動立体視 — 重ね合わせの原理に基づいた基本拘束方程式の導出 —,” 情報処理学会第49回全国大会, 4F-09 (1994.9).

付録 A

第2章の付録

A.1 多重オプティカルフローに関する付録

A.1.1 多重オプティカルフローの基本拘束方程式の証明

簡単のため、多重度 $n = 2$ の場合だけを証明する。 n が 3 以上の場合についても同様に証明できる。

まず、演算子の可換性、 $a(\mathbf{v}_1)a(\mathbf{v}_2) = a(\mathbf{v}_2)a(\mathbf{v}_1)$ は、 \mathbf{v}_1 と \mathbf{v}_2 が定数ベクトルであるという仮定から容易に示される。この性質と演算子の線形性、 $a(\mathbf{v})\{f(x, y, t) + g(x, y, t)\} = a(\mathbf{v})f(x, y, t) + a(\mathbf{v})g(x, y, t)$, $a(\mathbf{v})0 = 0$ を用いる。

まず、演算子の線形性から

$$\begin{aligned} a(\mathbf{v}_2)a(\mathbf{v}_1)f(x, y, t) &= a(\mathbf{v}_2)a(\mathbf{v}_1)\{f_1(x, y, t) + f_2(x, y, t)\} \\ &= a(\mathbf{v}_2)a(\mathbf{v}_1)f_1(x, y, t) + a(\mathbf{v}_2)a(\mathbf{v}_1)f_2(x, y, t) \end{aligned} \quad (\text{A.1})$$

と展開できる。次に、演算子の可換性から、 $a(\mathbf{v}_2)a(\mathbf{v}_1)f_2(x, y, t) = a(\mathbf{v}_1)a(\mathbf{v}_2)f_2(x, y, t)$ である。これと、 $f_1(x, y, t)$ と $f_2(x, y, t)$ 、それぞれに関する拘束である $a(\mathbf{v}_1)f_1(x, y, t) = 0$ 、 $a(\mathbf{v}_2)f_2(x, y, t) = 0$ を用いる。

$$\begin{aligned} &a(\mathbf{v}_2)a(\mathbf{v}_1)f_1(x, y, t) + a(\mathbf{v}_2)a(\mathbf{v}_1)f_2(x, y, t) \\ &= a(\mathbf{v}_2)a(\mathbf{v}_1)f_1(x, y, t) + a(\mathbf{v}_1)a(\mathbf{v}_2)f_2(x, y, t) \\ &= a(\mathbf{v}_2)\{a(\mathbf{v}_1)f_1(x, y, t)\} + a(\mathbf{v}_1)\{a(\mathbf{v}_2)f_2(x, y, t)\} \\ &= a(\mathbf{v}_2)0 + a(\mathbf{v}_1)0 \\ &= 0 + 0 = 0 \end{aligned} \quad (\text{A.2})$$

したがって、定理の拘束方程式が成立する。

A.1.2 多重オプティカルフロー推定のための連立非線形固有値方程式の導出

エネルギー積分 (2.23) において、 \mathbf{V}_k 以外のフローベクトルを既知数とみなし、 \mathbf{V}_k だけに関して変分法を適用する。このとき、このエネルギー積分の最小化問題は次の関数の停

留点を求める問題と等価である [163][168]。

$$G^{(n,k)}(\mathbf{V}_k, \lambda^{(n,k)}) = \frac{\int_{\mathcal{N}(x_c)} H^{(n,k)}(\mathbf{V}_k) f(\mathbf{x}) d\mathbf{x}}{\int_{\mathcal{N}(x_c)} f(\mathbf{x}) \cdot f(\mathbf{x}) d\mathbf{x}} + \lambda^{(n,k)} (1 - \|\mathbf{V}_k\|_{\mathbf{W}_k}^2) \quad (\text{A.3})$$

エネルギー演算子 $H^{(n,k)}(\mathbf{V}_k)$ は、 $H^{(n)}(\mathbf{V}^{(n)})$ において \mathbf{V}_k 以外のフローパラメータを既知定数とした時のエネルギー演算子である。 $H^{(n,k)}(\mathbf{V}_k) f(\mathbf{x})$ は \mathbf{V}_k の成分に関して 2 次式である。式 $G^{(n,k)}(\mathbf{V}_k)$ を停留にする \mathbf{V}_k は、 $\frac{\partial G^{(n,k)}(\mathbf{V}_k)}{\partial \mathbf{V}_k} = 0$ 、 $\frac{\partial G^{(n,k)}(\mathbf{V}_k)}{\partial \lambda^{(n,k)}} = 0$ を解いて得られる。この連立式から固有値問題 (2.24) が得られる。この固有値問題の最小固有値が、エネルギー積分 (2.23) の最小値に対応し、そのときの固有ベクトルが \mathbf{V}_k 以外のフローベクトルを既知とみなしたときのフローベクトル \mathbf{V}_k の最適推定値である。 $K^{(n,k)}$ は一般的な表式が複雑なので省略する。

A.1.3 多重オプティカルフロー推定のための線形固有値方程式の導出

多重度が n の場合、エネルギー積分 (2.23) の最小化は次式の停留点を求める問題と等価である [163][168]。

$$G^{(n)}(\mathbf{V}^{(n)}, \lambda^{(n)}) = \frac{\int_{\mathcal{N}(x_c)} H^{(n)}(\mathbf{V}^{(n)}) f(\mathbf{x}) d\mathbf{x}}{\int_{\mathcal{N}(x_c)} |f(\mathbf{x})|^2 d\mathbf{x}} + \lambda^{(n)} (1 - \|\mathbf{V}^{(n)}\|_{\mathbf{W}^{(n)}}^2) \quad (\text{A.4})$$

$\lambda^{(1)}$ は、ラグランジュ乗数である。停留条件 $\frac{\partial G^{(n)}}{\partial \mathbf{V}^{(n)}} = 0$ 、 $\frac{\partial G^{(n)}}{\partial \lambda^{(n)}} = 0$ を計算して、固有値方程式 (2.25) と、モーメントテンソル (2.26) が得られる。この固有値方程式の最小固有値がエネルギー積分 $E^{(n)}$ の最小値を与え、その時の固有ベクトルが多重フローパラメータ $\mathbf{V}^{(n)}$ の推定値である。ただし、 $n = 1$ 以外の場合、このままでは、パラメータ空間の対称性のため、最小固有値が縮退し、対応する固有空間も 2 次元以上となる。このとき、固有ベクトルを求める数値計算は困難である。そこで、あらかじめ、固有値方程式とテンソル $\mathbf{V}^{(n)}$ を各 \mathbf{V}_k の置換群に関して対称化する。この対称化操作によって独立な変数の次元が、 3^n から $\frac{1}{2}(n+1)(n+2)$ になる (詳細は省略。 $n = 2$ の場合は次の付録 A.1.4 を参照。) ¹。

A.1.4 $n = 1$ および $n = 2$ の場合の線形固有値方程式の導出

インプリメンテーションで用いた $n = 1$ の場合と $n = 2$ の場合における線形固有値方程式を導出する。以下では、 $\mathbf{W}_k = I(k = 1, \dots, n)$ を仮定する。つまり、各フローベクトルに関して等方ノルムを仮定する。これは、フィルタが x, y, t の各方向に関して等方的に配置され、各運動パターンが等方的なパワーを持つ場合に対応する。

¹各フローに関するノルム $\|\cdot\|_{\mathbf{W}_k}$ が等しくない場合には、最小固有値は一般には縮退しない。しかし、あらかじめ各フローのノルムを知ることが困難である。初めはすべてのフローに関して同じノルムを仮定せざるを得ない。したがって、固有値の縮退の問題は避けられない。

$n = 1$ の場合 同次フローベクトル $\mathbf{V} = (V_1, V_2, V_3)$ に関する 3 次元の固有値方程式に帰着される。

$$A^{(1)}\mathbf{V} = \lambda^{(1)}\mathbf{V} \quad (\text{A.5})$$

$$A^{(1)} = \begin{bmatrix} M_{(11)} & M_{(12)} & M_{(13)} \\ M_{(21)} & M_{(22)} & M_{(23)} \\ M_{(31)} & M_{(32)} & M_{(33)} \end{bmatrix} \quad (\text{A.6})$$

各成分モーメント $M_{(ij)}$ は、次式で計算される。

$$\begin{aligned} M_{(ij)} &= \frac{\int_{\mathcal{N}(x_c)} \overline{\{\partial_i f(\mathbf{x})\}} \cdot \{\partial_j f(\mathbf{x})\} dx}{\int_{\mathcal{N}(x_c)} \overline{f(\mathbf{x})} \cdot f(\mathbf{x}) dx} \\ &= \frac{\int_{\mathcal{N}(x_c)} \overline{\omega_i F(\omega)} \cdot \omega_j F(\omega) d\omega}{\int_{\mathcal{N}(x_c)} \overline{F(\omega)} \cdot F(\omega) d\omega} \end{aligned} \quad (\text{A.7})$$

$\overline{\omega_i F(\omega)} \omega_j F(\omega) = \omega_i \omega_j |F(\omega)|^2$ および、 $\overline{F(\omega)} F(\omega) = |F(\omega)|^2$ であるから、時空間周波数領域におけるエネルギー分布 $|F(\omega)|^2$ の 0 次モーメントと 2 次モーメント (6 個の独立な成分をもつ) を計算すれば良い。

$n = 2$ の場合 まず、 $3^2 = 9$ 次元の固有値方程式 (2.25) を A.1.3 の方法で求める。次に、2 個の同次フローベクトル \mathbf{V}_1 と \mathbf{V}_2 の対称化されたテンソル積 $\mathbf{V}_s^{(2)} = \frac{1}{2}(\mathbf{V}_1 \otimes \mathbf{V}_2 + \mathbf{V}_2 \otimes \mathbf{V}_1)$ に関する 6 次元線形固有値問題に変形する (計算省略)²。このとき、 $\mathbf{V}_s^{(2)}$ を成分とする次の 6 次元ベクトルと同一視する。

$$\begin{aligned} \mathbf{V}_s^{(2)} &= (V_{s(xx)}, V_{s(yy)}, V_{s(tt)}, V_{s(xy)}, V_{s(yt)}, V_{s(tx)})^T \\ &\stackrel{\text{def}}{=} \left(V_{1x}V_{2x}, V_{1y}V_{2y}, V_{1t}V_{2t}, \frac{1}{2}(V_{1x}V_{2y} + V_{1y}V_{2x}), \frac{1}{2}(V_{1y}V_{2t} + V_{1t}V_{2y}), \frac{1}{2}(V_{1t}V_{2x} + V_{1x}V_{2t}) \right)^T \end{aligned} \quad (\text{A.8})$$

すると、線形固有値方程式は次式である。

$$A^{(2)}\mathbf{V}^{(2)} = \lambda^{(2)}D\mathbf{V}^{(2)} \quad (\text{A.9})$$

行列 $A^{(2)}$ 、 D は次式で定義される。

$$A^{(2)} = \begin{bmatrix} M_{(1111)} & M_{(1122)} & M_{(1133)} & 2M_{(1112)} & 2M_{(1123)} & 2M_{(1131)} \\ M_{(2211)} & M_{(2222)} & M_{(2233)} & 2M_{(2212)} & 2M_{(2223)} & 2M_{(2231)} \\ M_{(3311)} & M_{(3322)} & M_{(3333)} & 2M_{(3312)} & 2M_{(3323)} & 2M_{(3331)} \\ 2M_{(1211)} & 2M_{(1222)} & 2M_{(1233)} & 4M_{(1212)} & 4M_{(1223)} & 4M_{(1231)} \\ 2M_{(2311)} & 2M_{(2322)} & 2M_{(2333)} & 4M_{(2312)} & 4M_{(2323)} & 4M_{(2331)} \\ 2M_{(3111)} & 2M_{(3122)} & 2M_{(3133)} & 4M_{(3112)} & 4M_{(3123)} & 4M_{(3131)} \end{bmatrix} \quad (\text{A.10})$$

²係数 $\frac{1}{2}$ は、任意の 0 でない定数で置き換えても、エネルギー積分、固有値、固有エネルギーには全く影響を与えない。固有ベクトルに対しても定数倍の違いが生じるだけで、方向のみが意味を持つこの場合には影響ない。式表現を簡潔にするためにこの定数を選んだ。

$$D = \text{diag}[1, 1, 1, 2, 2, 2] \quad (\text{A.11})$$

ここで、diag は、[] 内の引数を対角成分を持つ対角行列を表す。モーメントテンソルの成分は、独立な 15 個の 4 次モーメントから構成される。

A.1.5 混合フローテンソルの 2 個のフローベクトルへの分解

対称テンソル積パラメータ $V^{(2)}$ を 2 つのフローベクトル V_1 、 V_2 に分解する計算は、固有値分解を用いて最適に実行できる。まず、次の 3 次元固有値方程式を解く。

$$Be = \mu e, \quad (\text{A.12})$$

$e = (e_1, e_2, e_3)^T$ は固有ベクトルであり、 B は、混合フローテンソル $V_s^{(2)}$ の成分から次式で構成される行列である。

$$B = \begin{bmatrix} V_{s(xx)}^{(2)} & V_{s(xy)}^{(2)} & V_{s(tx)}^{(2)} \\ V_{s(xy)}^{(2)} & V_{s(yy)}^{(2)} & V_{s(yt)}^{(2)} \\ V_{s(tx)}^{(2)} & V_{s(yt)}^{(2)} & V_{s(tt)}^{(2)} \end{bmatrix} \quad (\text{A.13})$$

固有値方程式 (A.12) は、3 個の実固有値をもつ。 $V^{(2)}$ が誤差を持たない場合、正負各 1 個ずつと 0 を固有値としてもつ。行列 B の階数は 2 である。行列 B の階数が 1 になるのは、 V_1 と V_2 が平行の場合である。このとき、2 個のフローベクトルは一致する (フローベクトルには同次座標を用いているのでベクトルの大きさは無視される)。 $V_s^{(2)}$ は固有ベクトルとして得られたので零ベクトルではありえない。したがって、行列 B の階数は 0 にはならない。

しかし、実際には、誤差が存在するため、一般には、階数は 3 である。計算された 3 個の実固有値を μ_a, μ_b, μ_c とおく。ただし、 $\mu_a < \mu_b < \mu_c$ とする。固有値を $-\mu_b$ だけずらして、 $\mu_- = \mu_a - \mu_b < 0$ 、 $\mu_+ = \mu_c - \mu_b < 0$ と置く。これは、固有値方程式 $(B - \mu_b I)e = \mu e$ を作ることに相当する。この修正された固有値方程式は、3 個の固有値 $\mu_-, 0, \mu_+$ をもち、固有ベクトルはもとの固有値方程式と同じである。対応する固有ベクトルを、それぞれ e_-, e_0, e_+ とおくと、求める 2 個のフローベクトル V_1 と V_2 は、次式である。

$$\begin{aligned} V_1 &= \sqrt{\frac{-\mu_-}{\mu_+ - \mu_-}} e_- + \sqrt{\frac{\mu_+}{\mu_+ - \mu_-}} e_+, \\ V_2 &= -\sqrt{\frac{-\mu_-}{\mu_+ - \mu_-}} e_- + \sqrt{\frac{\mu_+}{\mu_+ - \mu_-}} e_+ \end{aligned} \quad (\text{A.14})$$

V_1 と V_2 を入れ換えたものと V_1 と V_2 の符号を変えたものも解である。しかし、これらは本質的に同一である。

A.1.6 ガボールフィルタ出力からのエネルギーモーメントの計算

ガボールフィルタの出力からエネルギーモーメント (2.26) を計算する公式を導く。まず、エネルギー積分とガボール関数の関係について述べる。ガボール基底関数系を $\{g_i\} (i =$

$1, \dots, L$) と書く。基底関数としてのガボール関数は直交系ではない。これを近似的に直交系であるとみなして計算する。 $\|\sigma\|$ に比較して $\|\omega_c\|$ が小さくなければ良い近似である [70]³

$$\tilde{f}(x) = \sum_{i=1}^L \alpha_i g_i \quad (\text{A.15})$$

$\alpha_i (i = 1, 2, \dots, L)$ は、複素係数である。 g_i は、パラメータ $\sigma_i, \omega_{ci}, x_{ci}$ によって $g_i(\sigma_i, \omega_{ci}, x_{ci}; x)$ とパラメータづけされている。係数 α_i の求め方を以下に示す。関数近似の残差を次式で定義する。

$$e = \int_R |f(x) - \tilde{f}(x)|^2 dx = \int_R |f(x) - \sum_{i=1}^L \alpha_i g_i|^2 dx \quad (\text{A.16})$$

R は、 $f(x)$ の定義域全体をあらわす。停留条件 $\frac{\partial e}{\partial \alpha_i} = 0$ から、次の L 元連立 1 次方程式が得られる。

$$S\mathbf{a} = T \quad (\text{A.17})$$

$S_{ij} = \int_R \bar{g}_i g_j dx$, $\mathbf{a} = (\alpha_i)$, $T_i = \int_R f(x) g_i dx$ である。関数系 $\{g_i\}$ の直交性を仮定すると S_{ij} は単位行列の定数倍となり、係数 α_i はガボール関数 g_i と時空間画像 $f(x)$ の積和として直接求められる。次に、この係数 α_i から、エネルギーモーメントを計算する。パーセバルの定理 [167] からエネルギーモーメントの計算は、実領域と周波数領域のいずれで行なっても結果は等価である。以下では周波数領域で考察する。3次元ガボール関数 (2.29) のフーリエ変換は、 ω_c を中心とし、標準偏差が $(\frac{1}{2\pi\sigma_x}, \frac{1}{2\pi\sigma_y}, \frac{1}{2\pi\sigma_t})$ のガウス関数に位相項を乗じたものである。

$$\begin{aligned} G(\sigma, \omega_c, x_c, \omega) = & \\ & \exp[-2\pi^2 \{ \sigma_x^2 (\omega_x - \omega_{xc})^2 + \sigma_y^2 (\omega_y - \omega_{yc})^2 + \sigma_t^2 (\omega_t - \omega_{tc})^2 \}] \\ & \exp[2\pi i \{ (\omega_x - \omega_{xc}) x_c + (\omega_y - \omega_{yc}) y_c + (\omega_t - \omega_{tc}) t_c \}] \end{aligned} \quad (\text{A.18})$$

この関数のフーリエ・エネルギー $P(G(\sigma, \omega_c, \omega))$ は、原点对称に配置された 2 個の 3 次元時空間ガウス関数の和である [48]。

$$\begin{aligned} P(G(\sigma, \omega_c, \omega)) = & \\ & \frac{1}{G(\sigma, \omega_c, \omega) \overline{G(\sigma, \omega_c, \omega)}} \\ = & \frac{1}{4} \{ \exp[-4\pi^2 \{ \sigma_x^2 (\omega_x - \omega_{xc})^2 + \sigma_y^2 (\omega_y - \omega_{yc})^2 + \sigma_t^2 (\omega_t - \omega_{tc})^2 \}] \\ & + \exp[-4\pi^2 \{ \sigma_x^2 (\omega_x + \omega_{xc})^2 + \sigma_y^2 (\omega_y + \omega_{yc})^2 + \sigma_t^2 (\omega_t + \omega_{tc})^2 \}] \} \end{aligned} \quad (\text{A.19})$$

したがって、エネルギーモーメントは、この 2 個のガウス関数が、原点のまわりの偶数次モーメントへ与える寄与を全フィルタについて加算して得られる。

$$m_{pqr} = \sum_{i=1}^L \int_{\mathcal{R}^3} |\alpha_i|^2 P(G(\sigma, \omega_{ci}, \omega)) \omega_x^p \omega_y^q \omega_t^r d\omega_x d\omega_y d\omega_t \quad (p+q+r = \text{even}) \quad (\text{A.20})$$

³このガボール基底関数による展開を厳密に行なうためには Gabor 関数の陪直交関数系を用いる方法 [70] と、緩和手法 [39] がある。

このとき、展開係数 α_i からのエネルギーモーメント計算公式は次式である。

$$\begin{aligned}
 m_{000} &= \sum_{i=1}^L |\alpha_i|^2 \\
 m_{200} &= \sum_{i=1}^L |\alpha_i|^2 (\omega_{xci}^2 + c_x) \\
 m_{110} &= \sum_{i=1}^L |\alpha_i|^2 \omega_{xci} \omega_{y ci} \\
 m_{400} &= \sum_{i=1}^L |\alpha_i|^2 (\omega_{xci}^4 + 6c_x \omega_{xci} + 3c_x^2) \\
 m_{310} &= \sum_{i=1}^L |\alpha_i|^2 (\omega_{xci}^3 \omega_{y ci} + 3c_x \omega_{xci} \omega_{y ci}) \\
 m_{220} &= \sum_{i=1}^L |\alpha_i|^2 (\omega_{xci}^2 \omega_{y ci}^2 + c_x \omega_{y ci}^2 + c_y \omega_{xci}^2 + c_x c_y) \\
 m_{211} &= \sum_{i=1}^L |\alpha_i|^2 (\omega_{xci}^2 \omega_{y ci} \omega_{tci} + c_x \omega_{y ci} \omega_{tci})
 \end{aligned} \tag{A.21}$$

$c_x = \frac{1}{8\pi^2\sigma_x^2}$, $c_y = \frac{1}{8\pi^2\sigma_y^2}$, $c_t = \frac{1}{8\pi^2\sigma_t^2}$ である。この他のモーメント成分についても同様の計算公式が成り立つ。係数 c_x, c_y, c_t の項は、ガボール関数のガウス窓が有限であるために生じる偏りの補正項である。モーメント行列の成分は、 $M_{(12)} = m_{110}/m_{000}$, $M_{(1123)} = m_{211}/m_{000}$ と計算する。

A.1.7 3 フレームからの 2 運動推定アルゴリズムとの関係

本研究とほぼ同じ時期に、Bergen et.al. によって、2 つの運動を 3 枚の画像フレームから抽出するアルゴリズムが独立に提案された [29]。彼らのアルゴリズムは、オプティカルフローではなく、画像全体、あるいはある小領域で 2 組の運動パラメータを 3 枚の画像フレームから計算するアルゴリズムである。このアルゴリズムは、本論文 2.1 節のアルゴリズムと深い関係がある。以下では、この関係を解析する。

A.1.7.1 Bergen et.al. のアルゴリズム

Bergen et.al. のアルゴリズムを説明する。まず、彼らの記法を導入する。2 個の運動を p, q とし、対応する空間パターンを $P(x, y)$, $Q(x, y)$ とする。このとき、これら 2 個の運動パターンの重なりからなる画像列 $I(x, y, t)$ ($t = 1, 2, 3, \dots$) を重畳演算 (加算または乗算) \oplus によって、

$$I(x, y, t) = P^{tP} \oplus Q^{tq} \tag{A.22}$$

と記述した。 P^{tP}, Q^{tq} は、それぞれ、パターン P, Q に、運動 p, q を t 回適用したものを意味する。運動 p が既知であると仮定すると、次の差分画像列 D_n を計算できる。

$$\begin{aligned}
 D_1 &= I(x, y, 2) - I^P(x, y, 1) = (Q^q - Q^p)^q \\
 D_2 &= I(x, y, 3) - I^P(x, y, 2) = (Q^q - Q^p)^{2q}
 \end{aligned}$$

...

(A.23)

この差分画像列は、運動 p の成分を含まない。これは、パターン $Q^q - Q^p$ が定速運動 q をしている画像と見なすことができる。したがって、従来の運動推定アルゴリズムを用いて、運動 q を推定することができる。逆に、運動 q が既知であると仮定した時に、運動 p を推定することができる。Bergen et.al. は、この事実を用いて、次の交互推定アルゴリズムを提案した。

1. p に適当な初期値を与える。
2. 運動 p を既知と仮定して、運動 q を推定する。
3. 運動 q を既知と仮定して、運動 p を推定する。
4. 収束判定を行い、未収束ならば、2. に戻る。

このアルゴリズムには、次の問題点がある。

- 常に収束するとは限らず、振動する場合もある。
- 2個の運動が存在するという前提条件のもとでのアルゴリズムであり、運動の個数を判断する条件は与えられていない。
- ループの中に画像の変換が含まれるため、演算量が多い。

A.1.7.2 Bergen et.al. のアルゴリズムと本研究のアルゴリズムの間の関係

2.1 節における記法を用いて、Bergen et.al. のアルゴリズムとの関係を論じる。運動 p の速度を $\mathbf{v}_1(x, y, t) = (u_1(x, y, t), v_1(x, y, t))$ 、運動 q の速度を $\mathbf{v}_2(x, y, t) = (u_2(x, y, t), v_2(x, y, t))$ とする。それぞれの振幅演算子を $a(\mathbf{v}_1(x, y, t))$ 、 $a(\mathbf{v}_2(x, y, t))$ とする。画像 $f_1(x, y, t)$ をパターン P に、画像 $f_2(x, y, t)$ をパターン Q に対応させる。画像列 $I(x, y, t)$ には、時空間画像 $f(x, y, t) = f_1(x, y, t) + f_2(x, y, t)$ を対応させる。

次のテイラー展開を考える。

$$\begin{aligned} f_2(x + u_1(x, y, t)\delta t, y + v_1(x, y, t)\delta t, t + \delta t) = \\ f_2(x, y, t) + \left\{ u_1(x, y, t) \frac{\partial}{\partial x} f_2(x, y, t) + v_1(x, y, t) \frac{\partial}{\partial y} f_2(x, y, t) + \frac{\partial}{\partial t} f_2(x, y, t) \right\} \delta t \\ + O(\delta t^2) \end{aligned} \quad (\text{A.24})$$

$\delta t = -1$ とおき、高次項を省略すると、次式を得る。

$$a(\mathbf{v}_1(x, y, t))f_2(x, y, t) \approx f_2(x, y, t) - f_2(x - u_1(x, y, t), y - v_1(x, y, t), t - 1) \quad (\text{A.25})$$

$a(\mathbf{v}_1(x, y, t))f_1(x, y, t) = 0$ と $f(x, y, t) = f_1(x, y, t) + f_2(x, y, t)$ を用いて次のように変形する。

$$a(\mathbf{v}_1(x, y, t))f(x, y, t) = a(\mathbf{v}_1(x, y, t))\{f_1(x, y, t) + f_2(x, y, t)\}$$

$$\begin{aligned}
&= a(v_1(x, y, t))f_1(x, y, t) + a(v_1(x, y, t))f_2(x, y, t) \\
&= a(v_1(x, y, t))f_2(x, y, t)
\end{aligned} \tag{A.26}$$

式 (A.26) を時間軸方向に離散化したものを $\hat{f}_2(x, y, t)$ とする。これに、式 (A.25) を代入すると、

$$\hat{f}_2(x, y, t+1) = f_2(x, y, t+1) - f_2(x - u_1(x, y, t+1), y - v_1(x, y, t+1), t) \tag{A.27}$$

を得る。 $t = 1, 2, \dots$ とすれば差分画像列 D_n が得られる。

$$\begin{aligned}
D_1 &= \hat{f}_2(x, y, 2) = f_2(x, y, 2) - f_2(x - u_1(x, y, 2), y - v_1(x, y, 2), 1) \\
D_2 &= \hat{f}_2(x, y, 3) = f_2(x, y, 3) - f_2(x - u_1(x, y, 3), y - v_1(x, y, 3), 2)
\end{aligned} \tag{A.28}$$

Bergen et.al. の記法との対応を次表に示す。

本研究の記法	Bergen et.al. の記法
$f_2(x, y, 2)$	Q^{2q}
$f_2(x - u_1(x, y, 2), y - v_1(x, y, 2))$	Q^{q+p}
$f_2(x, y, 3)$	Q^{3q}
$f_2(x - u_1(x, y, 3), y - v_1(x, y, 3), 2)$	Q^{2q+p}

この対応では、速度一定の仮定 $v_1(x, y, 2) = v_1(x, y, 3)$ を用いている。つまり、加速度が零でなければならない。

以上から、次の事がわかる。Bergen et.al. の方法は、本研究における 2 重フローの基本拘束方程式において、 u_1 と u_2 のうちの片方を既知とし、残りの片方を推定するプロセスを交互に繰り返すこととほぼ等価である。ただし、Bergen et.al. は、微分演算の代わりに画像の変換を直接施している。これに対して、本研究では、両方のベクトルをともに未知ベクトルとして、同時かつ一撃的に求めている。したがって、これら 2 つの方法は、同じ基本拘束方程式を異なる方法で解いていると考えられる。

つまり、2 重オプティカルフローの計算理論は、Bergen et.al. の方法と密接な関係がある。しかし、Bergen et.al. のアルゴリズムは、画像列からの運動推定に限定されていて一般性がない。それに対し、本論文のアプローチは、一般的な基本原理から出発している。その基本原理の一応用として導かれたのが、第 2.1 節で述べた多重オプティカルフローの計算理論とアルゴリズムである。

A.2 多重ステレオ視差に関する付録

A.2.1 単一視差の場合の拘束と従来法との関連

近似 (2.43) を用いると単一視差 D の振幅演算子 $a(D)$ は、次式に示す演算子 $\alpha(D)$ である。

$$a(D) \approx \alpha(D) = \begin{bmatrix} 1 & -1 + D\partial_x \\ -1 - D\partial_x & 1 \end{bmatrix} \quad (\text{A.29})$$

すると、一般化基本拘束方程式 $a(D)f^{(p,q)}(x,y) = 0$ は、方程式 $\alpha(D)f^{(p,q)}(x,y) = 0$ で近似される。成分ごとに書くと次式である。

$$\begin{aligned} r_L^{(1)}[p,q](x,y) &\stackrel{\text{def}}{=} \{L^{(p,q)}(x,y) - R^{(p,q)}(x,y)\} + DR^{(p+1,q)}(x,y) = 0 \\ r_R^{(1)}[p,q](x,y) &\stackrel{\text{def}}{=} \{R^{(p,q)}(x,y) - L^{(p,q)}(x,y)\} - DL^{(p+1,q)}(x,y) = 0 \end{aligned} \quad (\text{A.30})$$

繰り返し修正によるステレオ整合法 [92][96] において、視差の修正量の計算に用いられた方程式は、上式において、 $p = q = 0$ とした場合である。ただし、論文 [92] [96] では、式 (A.30) の2式のうち、片方だけを用いている。本論文では、両方を同時に用いて、左右の対称性を持たせた。点 (x_0, y_0) の近傍で視差 D を推定するために、次のエネルギー関数を用いる。

$$E(D(x_0, y_0)) = \int_{\mathcal{R}^2} \{w[p,q](x-x_0, y-y_0)\}^2 [\{r_L^{(1)}[p,q](x,y)\}^2 + \{r_R^{(1)}[p,q](x,y)\}^2] dx dy \quad (\text{A.31})$$

$w[p,q](x,y)$ は、正值をとる重み関数である。視差 D に関する方程式 $\frac{\partial E(D)}{\partial D} = 0$ を解いて、点 (x_0, y_0) における視差の推定式を得る。

$$D(x_0, y_0) = \frac{b(x_0, y_0)}{a(x_0, y_0)} \quad (\text{A.32})$$

$b(x_0, y_0)$, $a(x_0, y_0)$ は、次式で定義される。

$$\begin{aligned} a(x_0, y_0) &= \sum_{p,q} \int_{\mathcal{R}^2} \{w[p,q](x-x_0, y-y_0)\}^2 \{(R^{(p+1,q)}(x,y))^2 + (L^{(p+1,q)}(x,y))^2\} d\mathbf{x} d\mathbf{y} \\ b(x_0, y_0) &= - \sum_{p,q} \int_{\mathcal{R}^2} \{w[p,q](x-x_0, y-y_0)\}^2 \\ &\quad \times \{L^{(p,q)}(x,y) - R^{(p,q)}(x,y)\} \{R^{(p+1,q)}(x,y) + L^{(p+1,q)}(x,y)\} dx dy \end{aligned} \quad (\text{A.34})$$

A.2.2 2重視差の振幅演算子の近似演算子の導出

視差 D_1 、 D_2 に対する振幅演算子 $a(D_1)$ 、 $a(D_2)$ の近似振幅演算子 $\alpha(D_1)$ 、 $\alpha(D_2)$ の合成演算子を計算する。

$$\alpha(D_2)\alpha(D_1) = \begin{bmatrix} 1 & -1 + D_2\partial_x \\ -1 - D_2\partial_x & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 + D_1\partial_x \\ -1 - D_1\partial_x & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 + (D_1 - D_2)\partial_x - D_1 D_2 \partial_x^2 & -2 + (D_1 + D_2)\partial_x \\ -2 - (D_1 + D_2)\partial_x & 2 - (D_1 - D_2)\partial_x - D_1 D_2 \partial_x^2 \end{bmatrix} \quad (\text{A.35})$$

D_1 と D_2 を交換すると次式を得る。

$$\alpha(D_1)\alpha(D_2) = \begin{bmatrix} 2 + (D_2 - D_1)\partial_x - D_1 D_2 \partial_x^2 & -2 + (D_1 + D_2)\partial_x \\ -2 - (D_1 + D_2)\partial_x & 2 - (D_2 - D_1)\partial_x - D_1 D_2 \partial_x^2 \end{bmatrix} \quad (\text{A.36})$$

$\frac{1}{2} \{ \alpha(D_1)\alpha(D_2) + \alpha(D_2)\alpha(D_1) \}$ を計算して、式 (2.44) を得る。

A.2.3 周波数領域における拘束方程式

ステレオ透明視の近似残差関数 (2.45) に、フーリエ変換を施す。すると、周波数領域における表現が得られる ($p = q = 0$ の場合を示す)。

$$\begin{aligned} \tilde{r}_L^{(2)}(\omega_x, \omega_y) &= 2(\tilde{L} - \tilde{R}) + (D_1 + D_2)(2\pi i)\omega_x \tilde{R} - D_1 D_2 (2\pi i)^2 \omega_x^2 \tilde{L} \\ \tilde{r}_R^{(2)}(\omega_x, \omega_y) &= 2(\tilde{R} - \tilde{L}) - (D_1 + D_2)(2\pi i)\omega_x \tilde{L} - D_1 D_2 (2\pi i)^2 \omega_x^2 \tilde{R} \end{aligned} \quad (\text{A.37})$$

ただし、 $\tilde{L} = \tilde{L}(\omega_x, \omega_y)$, $\tilde{R} = \tilde{R}(\omega_x, \omega_y)$ と省略した。この残差関数を用いて、例えば、Gabor 関数を用いたアルゴリズムが導かれる。

付録 B

第3章の付録

B.1 多重3次元運動・構造復元に関する付録

B.1.1 基本テンソル推定アルゴリズムの導出

エネルギー関数 (3.31) の最小化は、次のラグランジュ関数の停留化に帰着される [163]。

$$\mathcal{F}(\hat{\mathbf{G}}^{AB}, \lambda) = \frac{1}{N} \sum_{\alpha=1}^N \left(\sum_{i,j,k,l=1}^3 \hat{G}_{(ik)(jl)}^{AB} m'_{\alpha(i)} m_{\alpha(j)} m'_{\alpha(k)} m_{\alpha(l)} \right)^2 + \lambda (4 - \|\hat{\mathbf{G}}^{AB}\|^2) \quad (\text{B.1})$$

停留条件は、

$$\begin{aligned} \frac{\partial \mathcal{F}(\hat{\mathbf{G}}^{AB}, \lambda)}{\partial \hat{G}_{(i'k')(j'l')}^{AB}} &= \\ \frac{2}{N} \sum_{\alpha=1}^N \left(\sum_{i,j,k,l=1}^3 \hat{G}_{(ik)(jl)}^{AB} m'_{\alpha(i)} m_{\alpha(j)} m'_{\alpha(k)} m_{\alpha(l)} \right) m'_{\alpha(i')} m_{\alpha(j')} m'_{\alpha(k')} m_{\alpha(l')} - 2\lambda \hat{G}_{(i'k')(j'l')}^{AB} \\ &= 2 \sum_{i,j,k,l=1}^3 \hat{G}_{(ik)(jl)}^{AB} \frac{1}{N} \sum_{\alpha=1}^N m'_{\alpha(i)} m_{\alpha(j)} m'_{\alpha(k)} m_{\alpha(l)} m'_{\alpha(i')} m_{\alpha(j')} m'_{\alpha(k')} m_{\alpha(l')} - 2\lambda \hat{G}_{(i'k')(j'l')}^{AB} = 0 \end{aligned} \quad (\text{B.2})$$

i', j', k', l' は、それぞれ 1, 2, 3 を動く。 λ に関する停留条件は、

$$\frac{\partial \mathcal{F}(\hat{\mathbf{G}}^{AB}, \lambda)}{\partial \lambda} = 4 - \|\hat{\mathbf{G}}^{AB}\|^2 = 0 \quad (\text{B.3})$$

である。これは正規化条件に他ならない。

相関テンソルの定義 (3.32) を用いると、方程式 (B.2) は、次式に変形される。

$$\sum_{i,j,k,l=1}^3 \hat{G}_{(ik)(jl)}^{AB} C_{ijkl i' j' k' l'} = \lambda \hat{G}_{(i'k')(j'l')}^{AB} \quad (\text{B.4})$$

添字 i', j', k', l' の組合せは全部で 81 通りである。しかし、対称性を考慮すると、独立な方程式数は 36 個である。方程式 (B.4) を式 (3.34) の定義を考慮して整理すると式 (3.35) が得られる。

B.1.2 基本行列の単位行ベクトル、単位列ベクトルの計算アルゴリズム導出

式 (3.41) を考慮すると、式 (3.42) の行列 S_i は、単位行ベクトル b_i^A と b_i^B のテンソル積を用いて次式に変形される。

$$\begin{aligned} S_i &= \frac{1}{2} \zeta_i^A \zeta_i^B (b_i^A \otimes b_i^B + b_i^B \otimes b_i^A) \\ &= \zeta_i^A \zeta_i^B \left[\left\{ \frac{1}{2} (b_i^A + b_i^B) \right\} \left\{ \frac{1}{2} (b_i^A + b_i^B) \right\}^T - \left\{ \frac{1}{2} (b_i^A - b_i^B) \right\} \left\{ \frac{1}{2} (b_i^A - b_i^B) \right\}^T \right] \end{aligned} \quad (\text{B.5})$$

$\|b_i^A\| = \|b_i^B\| = 1$ から、 $(b_i^A + b_i^B, b_i^A - b_i^B) = \|b_i^A\|^2 - \|b_i^B\|^2 = 0$ である。したがって、 $b_i^A + b_i^B$ と $b_i^A - b_i^B$ は互いに垂直である。一方、行列 S_i の固有値展開は、(3.43) である。

誤差の無い理想的な場合、 S_i の階数は、2 以下である ($\det S_i = 0$ から帰結される)。 $\sigma_{i0} = 0$ であるから、固有値展開の一意性と固有ベクトル同士の直交性より、式 (B.5) の最後の辺は、この固有値展開に対応しなければならない。実際には、誤差のため、 $\sigma_{i0} \neq 0$ である。この誤差の影響をなるべく小さくとどめるため、 $\sigma_{i0} = 0$ となるように、固有値全体をずらす。 $\sigma_{i+} - \sigma_{i0}$, $\sigma_{i-} - \sigma_{i0}$ を、それぞれ σ_{i+} , σ_{i-} の代わりに用いる。このとき、次式の対応が成り立つ。 $\zeta_i^A \zeta_i^B > 0$ の場合:

$$\begin{aligned} \sqrt{\sigma_{i+} - \sigma_{i0}} s_{i+} &= \frac{\sqrt{\zeta_i^A \zeta_i^B}}{2} (b_i^A + b_i^B) \\ \sqrt{\sigma_{i0} - \sigma_{i-}} s_{i-} &= \frac{\sqrt{\zeta_i^A \zeta_i^B}}{2} (b_i^A - b_i^B) \end{aligned} \quad (\text{B.6})$$

この 2 式の和と差から、次式を得る。

$$\begin{aligned} \sqrt{\sigma_{i+} - \sigma_{i0}} s_{i+} + \sqrt{\sigma_{i0} - \sigma_{i-}} s_{i-} &= \sqrt{\zeta_i^A \zeta_i^B} b_i^A \\ \sqrt{\sigma_{i+} - \sigma_{i0}} s_{i+} - \sqrt{\sigma_{i0} - \sigma_{i-}} s_{i-} &= \sqrt{\zeta_i^A \zeta_i^B} b_i^B \end{aligned} \quad (\text{B.7})$$

$\zeta_i^A \zeta_i^B < 0$ の場合は、同様にして、

$$\begin{aligned} \sqrt{\sigma_{i+} - \sigma_{i0}} s_{i+} - \sqrt{\sigma_{i0} - \sigma_{i-}} s_{i-} &= \sqrt{-\zeta_i^A \zeta_i^B} b_i^A \\ \sqrt{\sigma_{i+} - \sigma_{i0}} s_{i+} + \sqrt{\sigma_{i0} - \sigma_{i-}} s_{i-} &= \sqrt{-\zeta_i^A \zeta_i^B} b_i^B \end{aligned} \quad (\text{B.8})$$

を得る。 $\|s_{i+}\| = \|s_{i-}\| = 1$ と $(s_{i+}, s_{i-}) = 0$ から

$$\left\| \sqrt{\sigma_{i+} - \sigma_{i0}} s_{i+} \pm \sqrt{\sigma_{i0} - \sigma_{i-}} s_{i-} \right\| = \sqrt{\sigma_{i+} - \sigma_{i-}} \quad (\text{B.9})$$

も成立する。式 (B.7), (B.8), (B.9) と、 b_i^A と b_i^B が単位ベクトルであることをもちいると、公式 (3.44) が得られる。単位列ベクトル c_j^A , c_j^B についても同様の考察から式 (3.48) が得られる。

B.1.3 係数 ζ_i, η_j の計算アルゴリズム導出

$\|\nu\| = \sqrt{\|\zeta\|^2 + \|\eta\|^2} = 2$ のもとで最小化する問題は、次のラグランジュ関数の停留化問題に帰着される。

$$\mathcal{L}(\zeta, \eta, \rho) = \sum_{i,j=1}^3 (\zeta_i b_{ij} - \eta_j c_{ji})^2 + \rho(4 - \|\zeta\|^2 - \|\eta\|^2) \quad (\text{B.10})$$

ρ は、ラグランジュ乗数である。停留条件は、

$$\begin{aligned} \frac{\partial \mathcal{L}(\zeta, \eta, \rho)}{\partial \zeta_{i'}} &= 2 \sum_{j=1}^3 b_{ij} (\zeta_i b_{ij} - \eta_j c_{ji}) + 2\rho \zeta_{i'} = 2 \left\{ \zeta_{i'} \|\mathbf{b}_{i'}\|^2 - \sum_{j=1}^3 \eta_j b_{ij} c_{ji} - \rho \zeta_{i'} \right\} = 0 \\ \frac{\partial \mathcal{L}(\zeta, \eta, \rho)}{\partial \eta_{j'}} &= 2 \sum_{i=1}^3 c_{ji} (\zeta_i b_{ij} - \eta_j c_{ji}) + 2\rho \eta_{j'} = 2 \left\{ \sum_{i=1}^3 \zeta_i b_{ij} c_{ji} - \eta_{j'} \|\mathbf{c}_{j'}\|^2 - \rho \eta_{j'} \right\} = 0 \\ \frac{\partial \mathcal{L}(\zeta, \eta, \rho)}{\partial \rho} &= 4 - \|\zeta\|^2 - \|\eta\|^2 = 0 \end{aligned} \quad (\text{B.11})$$

である。これらを $\|\mathbf{b}_{i'}\| = \|\mathbf{c}_{j'}\| = 1$ を用いて整理すると固有値問題 (3.50) が得られる。最小固有値に対応する固有ベクトルが ν の推定値である。

B.1.4 基本行列の並進運動ベクトルと回転行列への分解

基本行列 \mathbf{G} から、並進ベクトル \mathbf{h} と回転行列 \mathbf{R} を求める方法を述べる。なお、本付録の内容は、文献 [22] [12] から引用した。

B.1.4.1 並進運動ベクトルの計算

解法1 $\mathbf{G} = \mathbf{H}\mathbf{R}$ より、

$$\mathbf{G}\mathbf{G}^T = (\mathbf{H}\mathbf{R})(\mathbf{H}\mathbf{R})^T = \mathbf{H}(\mathbf{R}\mathbf{R}^T)\mathbf{H}^T \quad (\text{B.12})$$

となり、 \mathbf{R} は回転行列であるから、 $\mathbf{R}\mathbf{R}^T = \mathbf{I}$ 。したがって、

$$\mathbf{G}\mathbf{G}^T = \mathbf{H}\mathbf{H}^T = \begin{bmatrix} h_2^2 + h_3^2 & -h_1 h_2 & -h_3 h_1 \\ -h_1 h_2 & h_3^2 + h_1^2 & -h_2 h_3 \\ -h_3 h_1 & -h_2 h_3 & h_1^2 + h_2^2 \end{bmatrix} \quad (\text{B.13})$$

この等式から、 $\mathbf{h} = (h_1, h_2, h_3)$ が得られる。

解法2 式 (B.13) の成分方程式数は、未知数の数よりも多い。したがって、解法1では、すべての成分に関する方程式を同時に最適に満たす運動パラメータを計算できない。最適化手法を用いた解法を以下に示す。まず、定義式 (3.4) より、 \mathbf{G} の各列ベクトルは、 \mathbf{h} と直交する。つまり、 $j = 1, 2, 3$ に対して、

$$\sum_{i=1}^3 G_{ij} h_i = 0 \quad (\text{B.14})$$

が成立する。この方程式を最適に満たす解は、2 乗誤差

$$E(\mathbf{h}) = \sum_{j=1}^3 \left(\sum_{i=1}^3 G_{ij} h_i \right)^2 = (\mathbf{h}, \mathbf{G}\mathbf{G}^T \mathbf{h}) \quad (\text{B.15})$$

を条件 $\|\mathbf{h}\| = 1$ のもとで最小化して得られる。行列 $\mathbf{G}\mathbf{G}^T$ の最小固有値に対応する固有ベクトルが \mathbf{h} の最適推定値である。

B.1.4.2 回転行列の計算

基本行列 \mathbf{G} から 3 次元回転行列 \mathbf{R} を復元する方法には、行列の QR 分解による方法 [108]、特異値分解による方法 [101]、ハミルトンの四元数による方法 [107] がある。ここでは、ハミルトンの四元数を用いた方法を説明する。

四元数による 3 次元回転の表現 ハミルトンの四元数は、複素数の概念の拡張であって、次式で定義される。

$$\tilde{q} = q_0 + q_1 \tilde{i} + q_2 \tilde{j} + q_3 \tilde{k} \quad (\text{B.16})$$

ここで、 q_0, q_1, q_2, q_3 は実数で、 $\tilde{i}, \tilde{j}, \tilde{k}$ は、複素数における虚数単位に相当するものであり、以下の関係式を満たす数として定義される。

$$\begin{aligned} \tilde{i}^2 &= \tilde{j}^2 = \tilde{k}^2 = -1 \\ \tilde{j}\tilde{k} &= -\tilde{k}\tilde{j} = \tilde{i} \\ \tilde{k}\tilde{i} &= -\tilde{i}\tilde{k} = \tilde{j} \\ \tilde{i}\tilde{j} &= -\tilde{j}\tilde{i} = \tilde{k} \end{aligned} \quad (\text{B.17})$$

四元数の共役数（複素数における複素共役に相当する）と絶対値は、次式で定義される。

$$\begin{aligned} \tilde{q}^* &= q_0 - q_1 \tilde{i} - q_2 \tilde{j} - q_3 \tilde{k} \\ \|\tilde{q}\| &= \sqrt{\tilde{q}\tilde{q}^*} = \sqrt{q_0^2 + q_1^2 + q_2^2 + q_3^2} \end{aligned} \quad (\text{B.18})$$

3 次元座標ベクトルを $\mathbf{x} = (x, y, z)^T$ 、としたとき、対応する四元数を $\tilde{x} = x\tilde{i} + y\tilde{j} + z\tilde{k}$ と定義する。単位ベクトル $\boldsymbol{\omega} = (\omega_x, \omega_y, \omega_z)$ の回りの回転角 θ の回転を、単位四元数

$$\tilde{q} = \sin \frac{\theta}{2} + \cos \frac{\theta}{2} (\omega_x \tilde{i} + \omega_y \tilde{j} + \omega_z \tilde{k}) \quad (\text{B.19})$$

で表現する。ただし、 $\|\tilde{q}\| = 1$ である。回転変換 $\mathbf{x}' = \mathbf{R}\mathbf{x}$ に対応して、四元数 \tilde{x} は、次式で変換される。

$$\tilde{x}' = \tilde{q}\tilde{x}\tilde{q}^* \quad (\text{B.20})$$

四元数の成分で行列 \mathbf{R} を表現すると次式である。

$$\mathbf{R} = \begin{bmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & -2(q_0q_3 - q_1q_2) & 2(q_0q_2 + q_1q_3) \\ 2(q_0q_3 + q_1q_2) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & -2(q_0q_1 - q_2q_3) \\ -2(q_0q_2 - q_1q_3) & 2(q_0q_1 + q_2q_3) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{bmatrix} \quad (\text{B.21})$$

四元数による3次元回転の計算 並進ベクトルが B.1.4.1 の方法で求められていると仮定する。このとき、行列 $\mathbf{K} = (K_{ij})$ を次の様に計算する。

$$\mathbf{K} = (\mathbf{g}_1 \times \mathbf{h}, \mathbf{g}_2 \times \mathbf{h}, \mathbf{g}_3 \times \mathbf{h}) \quad (\text{B.22})$$

ここで、 \mathbf{g}_i は $\mathbf{G} = (\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3)$ なる縦ベクトルである。行列 \mathbf{K} の成分から、次の4次行列 \mathbf{N} を作る。

$$\mathbf{N} = \begin{bmatrix} K_{11} + K_{22} + K_{33} & K_{32} - K_{23} & K_{13} - K_{31} & K_{21} - K_{12} \\ K_{32} - K_{23} & K_{11} - K_{22} - K_{33} & K_{12} + K_{21} & K_{31} + K_{13} \\ K_{13} - K_{31} & K_{12} + K_{21} & -K_{11} + K_{22} - K_{33} & K_{23} + K_{32} \\ K_{21} - K_{12} & K_{31} + K_{13} & K_{23} + K_{32} & -K_{11} - K_{22} + K_{33} \end{bmatrix} \quad (\text{B.23})$$

行列 \mathbf{N} の最大固有値に対応する単位固有ベクトルを $\bar{\mathbf{q}} = (\bar{q}_0, \bar{q}_1, \bar{q}_2, \bar{q}_3)^T$ とする。このとき、回転に対応する四元数の最適推定値は、

$$\tilde{q} = \bar{q}_0 + \bar{q}_1 \tilde{i} + \bar{q}_2 \tilde{j} + \bar{q}_3 \tilde{k} \quad (\text{B.24})$$

である。この成分から、式 (B.21) を用いて、回転行列 \mathbf{R} が得られる。回転角 θ と回転軸ベクトル $(\omega_x, \omega_y, \omega_z)^T$ は、式 (B.19) から計算できる。

B.2 多価関数への標準正則化理論の拡張に関する付録

B.2.1 重ね合わせの原理に基づいた多価関数の直接表現の導出

多価関数の基本関係式 (3.62), (3.78), (3.83) は、第 2.3 節に述べた重ね合わせの原理から導出できる。

B.2.1.1 スカラー値多価関数の演算子表現

データの確率振幅分布関数を $\phi(\boldsymbol{\xi}, \nu)$ で表す。これは、次式で定義される。

$$\phi(\boldsymbol{\xi}, \nu) \stackrel{\text{def}}{=} \sum_{\alpha=1}^N \left\{ \mu_{(\alpha)}(\mathbf{x}_{(\alpha)} - \boldsymbol{\xi}) \rho_{(\alpha)}(y_{(\alpha)} - \nu) \right\} \quad (\text{B.25})$$

$\mu_{(\alpha)}(\mathbf{x})$ は、 α 番目のデータ点の、入力空間 \mathbf{x} における確率振幅分布である。 $\rho_{(\alpha)}$ は、 α 番目のデータ点の、出力値 y に関する確率振幅分布である。

単独データ点に正規分布を仮定すると、 $\mu_{(\alpha)}(\mathbf{x})$ と $\rho_{(\alpha)}(y)$ は、次式で定義される。

$$\mu(\mathbf{x}) = \mu_{(\alpha)}(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{n}{4}} \sigma_x^{\frac{n}{2}}} \exp\left(-\frac{1}{4\sigma_x^2} \mathbf{x}^T \mathbf{x}\right) \quad (\text{B.26})$$

$$\rho(y) = \rho_{(\alpha)}(y) = \frac{1}{(2\pi)^{\frac{1}{4}} \sigma_y^{\frac{1}{2}}} \exp\left(-\frac{1}{4\sigma_y^2} y^2\right) \quad (\text{B.27})$$

ただし、簡単のために、入力空間の確率分布に一様で等方的な標準偏差 σ_x を仮定した。出力値の標準偏差は、 σ_y と記した。関数 $\mu(x)$ と $\rho(y)$ は、 $\sigma_x \rightarrow 0, \sigma_y \rightarrow 0$ なる極限をとると、次の性質をもつ超関数とみなせる。

$$\int_{\mathcal{R}^n} \{\mu(\xi)\}^2 d\xi = 1, \quad (\text{B.28})$$

$$\int_{\mathcal{R}^n} \{\mu(x - \xi)g(\xi)\}^2 d\xi = \{g(x)\}^2. \quad (\text{B.29})$$

$$\int_{\mathcal{R}} \{\rho(\nu)\}^2 d\nu = 1, \quad (\text{B.30})$$

$$\int_{\mathcal{R}} \{\rho(y - \nu)h(\nu)\}^2 d\nu = \{h(y)\}^2 \quad (\text{B.31})$$

スカラー値関数の振幅演算子は、次の形式で表すことができる。

$$a(f) = \nu - f(\xi) \quad (\text{B.32})$$

以上の定義のもとで、演算子形式の方程式表現は次式である。

$$a(f)\phi(\xi, \nu) = 0 \quad (\text{B.33})$$

重ね合わせの原理を用いると、次式が h 価関数の基本関係式である。

$$a(f_1)a(f_2)\cdots a(f_h)\phi(\xi, \nu) = 0 \quad (\text{B.34})$$

この方程式が、式 (3.78) に等価であることを以下に示す。

まず、次の 2 乗誤差関数を考える。

$$\mathcal{E}(f_1, f_2, \dots, f_h) = \int_{\mathcal{R}} \int_{\mathcal{R}^n} \{a(f_1)a(f_2)\cdots a(f_h)\phi(\xi, \nu)\}^2 d\xi d\nu \quad (\text{B.35})$$

この式を $\sigma_x \rightarrow 0, \sigma_y \rightarrow 0$ のもとで、式 (B.28), (B.29), (B.30), (B.31) を用いて計算すると、次式を得る。

$$\mathcal{E}(f_1, f_2, \dots, f_h) = \sum_{i=1}^N \left\{ (y_{(i)} - f_1(x_{(i)})) (y_{(i)} - f_2(x_{(i)})) \cdots (y_{(i)} - f_h(x_{(i)})) \right\}^2 \quad (\text{B.36})$$

これは、方程式 (3.78) にデータを代入した結果の 2 乗誤差に他ならない。したがって、誤差が零の場合に方程式 (B.34) と (3.78) は等価である。

B.2.1.2 ベクトル値多価関数の演算子表現

前節の議論は、ベクトル値多価関数へ拡張できる。まず、データ分布関数は、次式に拡張される。

$$\phi(\xi, \nu) \stackrel{\text{def}}{=} \sum_{\alpha=1}^N \left\{ \mu_{(\alpha)}(x_{(\alpha)} - \xi) \rho_{(\alpha)}(y_{(\alpha)} - \nu) \right\} \quad (\text{B.37})$$

各データ点の存在確率に正規分布を仮定すると、 $\mu_{(\alpha)}(\mathbf{x})$ と $\rho_{(\alpha)}(\mathbf{y})$ は、次式で定義される。

$$\mu(\mathbf{x}) = \mu_{(\alpha)}(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{n}{2}} \sigma_x^{\frac{n}{2}}} \exp\left(-\frac{1}{4\sigma_x^2} \mathbf{x}^T \mathbf{x}\right) \quad (\text{B.38})$$

$$\rho(\mathbf{y}) = \rho_{(\alpha)}(\mathbf{y}) = \frac{1}{(2\pi)^{\frac{m}{2}} \sigma_y^{\frac{m}{2}}} \exp\left(-\frac{1}{4\sigma_y^2} \mathbf{y}^T \mathbf{y}\right) \quad (\text{B.39})$$

簡単のため、入力空間の確率分布に一樣で等方的な標準偏差 σ_x を仮定した。出力値の標準偏差は、 σ_y と記した。 $\mu(\mathbf{x}), \rho(\mathbf{y})$ は、 $\sigma_x \rightarrow 0, \sigma_y \rightarrow 0$ なる極限において、次の性質をもつ超関数と考えられる。

$$\int_{\mathcal{R}^n} \{\mu(\boldsymbol{\xi})\}^2 d\boldsymbol{\xi} = 1, \quad (\text{B.40})$$

$$\int_{\mathcal{R}^n} \{\rho(\mathbf{x} - \boldsymbol{\xi})g(\boldsymbol{\xi})\}^2 d\boldsymbol{\xi} = \{g(\mathbf{x})\}^2. \quad (\text{B.41})$$

$$\int_{\mathcal{R}^m} \{\mu(\boldsymbol{\nu})\}^2 d\boldsymbol{\nu} = 1 \quad (\text{B.42})$$

$$\int_{\mathcal{R}^m} \{\mu(\mathbf{y} - \boldsymbol{\nu})h(\boldsymbol{\nu})\}^2 d\boldsymbol{\nu} = \{h(\mathbf{y})\}^2 \quad (\text{B.43})$$

ベクトル値関数の振幅演算子は、次式である。

$$a(\mathbf{f}) = \boldsymbol{\nu} - \mathbf{f}(\boldsymbol{\xi}) \quad (\text{B.44})$$

このとき、一価関数の演算子形式表現は、次式である。

$$a(\mathbf{f})\phi(\boldsymbol{\xi}, \boldsymbol{\nu}) = 0 \quad (\text{B.45})$$

重ね合わせの原理を用いると h 価関数の方程式は、次式である。

$$a(\mathbf{f}_1)a(\mathbf{f}_2)\cdots a(\mathbf{f}_h)\phi(\boldsymbol{\xi}, \boldsymbol{\nu}) = 0 \quad (\text{B.46})$$

この方程式から、スカラー値の場合と同様にして、ベクトル値多価関数の基本関係式 (3.83) が導かれる。

付録 C

第 4 章の付録

C.1 [定理 4.1 の証明]

任意の正則行列が基本行列 $P_{ij}[s]$, R_{ij} , $Q_i[s] = I - (1-s)I_{ii}$ の積に分解できることは、線形代数で周知の事実である [166] [169]。しかし、変換行列と基本行列に等積条件が付いた、定義 4.2 と定理 4.1 の場合は自明でない。

以下では、正方行列の次元は () の付いた上つき添字で $A^{(n)}$ と表す。正方行列 $A^{(n)}$ の $(n+1)$ 次元への拡張行列 $(A^{(n)})^*$ を次式で定義する。

$$(A^{(n)})^* = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & A^{(n)} \end{bmatrix} \quad (C.1)$$

性質 C.1 行列 A の右 [左] から $P_{ij}[s]$, Q_i , R_{ij} を乗ずると、それぞれ次の効果が A に生じる。

1. $P_{ij}[s]$: 第 i 列 [第 j 行] に s を乗じたものが第 j 列 [第 i 行] に加算される。
2. Q_i : 第 i 列 [第 j 行] の符号が反転される。
3. R_{ij} : 第 i 列 [第 j 行] と第 j 列 [第 i 行] が交換される。□

性質 C.2 (1) $\det P_{ij}[s] = 1$, (2) $\det Q_i = \det R_{ij} = -1$ □

性質 C.3 (1) $P_{ij}[s]^{-1} = P_{ij}[-s]$, (2) $Q_i^{-1} = Q_i$, (3) $R_{ij}^{-1} = R_{ji}^{-1} = R_{ij} = R_{ji}$ □

性質 C.4 (1) $(P_{ij}^{(n)}[s])^* = P_{i+1, j+1}^{(n+1)}[s]$, (2) $(Q_i^{(n)})^* = Q_{i+1}^{(n+1)}$, (3) $(R_{ij}^{(n)})^* = R_{i+1, j+1}^{(n+1)}$ □

性質 C.5 (1) $(AB)^* = A^*B^*$, (2) $\det A^* = \det A$ □

以上の性質 C.1 ~ C.5 は、以下の証明で用いられる。

補題 C.1 $|\det A| = 1$ を満たす n 次行列 A は次の形に変形できる。

$$A = C_1 B^* C_2 \quad (C.2)$$

ここで、 B は $|\det B^*| = |\det B| = 1$ を満たす $(n-1)$ 次行列で、 C_1 と C_2 は、両方合わせて $2k-1$ 個以下の $P_{ij}^{(n)}[s]$ と 1 個以下の $R_{ij}^{(n)}$ の積からなる n 次行列である。

[証明]

行列 A の (i, j) 成分を a_{ij} とする。

- 場合 (1) もし、 $a_{1j} \neq 0$ と $1 < j \leq n$ を満たす j が存在するならば、 $G_1 = P_{j1}^{(n)} \left[\frac{1-a_{11}}{a_{1j}} \right]$ と置く。
- 場合 (2) もし、(1) の条件を満たす j が存在しなければ、 $a_{11} \neq 0$ でなければならない (もし、 $a_{11} = 0$ ならば $\det A = 0$ となる。これは、仮定に反する)。この場合は、 $G_2 = R_{1j}^{(n)} P_{j1}^{(n)} \left[\frac{1}{a_{11}} \right]$ と置く。

性質 C.1 を用いると、 AG_1 または AG_2 の $(1, 1)$ 成分は 1 である。この結果を、それぞれの場合 (1), (2) に応じて、 $\tilde{A} = [\tilde{a}_{ij}] = AG_1$ または $\tilde{A} = [\tilde{a}_{ij}] = AG_2$ とおく。

次に、

- 場合 (1) に対して：行列 H_1 を、 $H_1 = \prod_{p=2}^n P_{1p}^{(n)} [-\tilde{a}_{1p}]$ と定義し、 $\bar{A} = [\bar{a}_{ij}] = \tilde{A}H_1$ とする。
- 場合 (2) に対して：行列 H_2 を、 $H_2 = \prod_{q=2}^n P_{q1}^{(n)} [-\bar{a}_{q1}]$ と定義し、 $\bar{A} = [\bar{a}_{ij}] = \tilde{A}H_2$ とする。

ふたたび性質 C.1 を用いると、 \bar{A} は、式 (C.1) の形になる。

行列 B^* 、 C_1 は、 $B^* = H_2 \bar{A} = H_2 A G_1 H_1$ 、 $C_1 = H_2^{-1}$ 、と定義する。 C_2 は、場合 (1) のとき、 $C_2 = (G_1 H_1)^{-1} = H_1^{-1} G_1^{-1}$ 、とする。場合 (2) のとき、 $C_2 = (G_2 H_1)^{-1} = H_1^{-1} G_2^{-1}$ とする。性質 C.2 と性質 C.4 を用いると $|\det(B^*)| = |\det B| = 1$ である。

性質 C.3 を用いると、積 C_1 と積 C_2 に含まれる行列 $P_{ij}[s]$ の数は、それぞれ多くても $(k-1)$ と k である ($s=0$ の場合は、 $P_{ij}[0] = I$ (単位行列) なので、その分行列 $P_{ij}[s]$ の数は減る)。したがって、合計すると、最大 $(2k-1)$ 個の行列 $P_{ij}[s]$ が使われた。行列 R_{ij} は、多くても、1 回だけ用いられた。

C.1.1 [定理 4.1 の証明]

次元 n に関する数学的帰納法で証明する。

- (1) $n = 1$ の場合、 $|\det A^{(1)}| = |a_{11}| = 1$ であるから、 $A^{(1)} = (1) = I^{(1)}$ または、 $A^{(1)} = (-1) = Q_1^{(1)}$ でなければならない。したがって、定理は成り立つ。

- (2) $n = k$ の場合について定理が成り立つと仮定する。補題 C.1 を用いると、 $|\det \mathbf{A}^{(k+1)}| = 1$ を満たす $(k+1)$ 次行列 $\mathbf{A}^{(k+1)}$ は次の形に変形できる。

$$\mathbf{A}^{(k+1)} = \mathbf{C}_1^{(k+1)} (\mathbf{B}^{(k)})^* \mathbf{C}_2^{(k+1)} \quad (\text{C.3})$$

ただし、 $|\det \mathbf{B}^{(k)}| = |\det (\mathbf{B}^{(k)})^*| = |\det \mathbf{A}^{(k+1)}| = 1$ である。 $n = k$ の場合の仮定から、 $\mathbf{B}^{(k)}$ は、 $k^2 - 1$ 個の $\mathbf{P}_{ij}^{(k)}[s]$ と有限個の \mathbf{R}_{ij} の積に分解できる。補題 C.1 を用いると \mathbf{C}_1 と \mathbf{C}_2 に含まれる行列 $\mathbf{P}_{ij}^{(k+1)}[s]$ の数は $2(k+1) - 1$ であり、 \mathbf{R}_{ij} の数は 1 個以下である。ゆえに、行列 $\mathbf{A}^{(k+1)}$ の分解に含まれる行列 $\mathbf{P}_{ij}^{(k+1)}[s]$ の数は、 $\{k^2 - 1\} + \{2(k+1) - 1\} = (k+1)^2 - 1$ である。したがって、定理は $n = k+1$ の場合にも成立する。

(1)、(2) より、定理 4.1 は任意の次元 n に対して成り立つ

系 4.1 は、 \mathcal{E} が \mathcal{EL} と \mathcal{T} に分解されることから容易に示される。

C.2 [命題 4.3 の証明]

基本鏡映変換 \hat{q}_i, \hat{r}_{ij} については、命題の成立が明らかである。以下、 $\mathbf{z} = (z_1, z_2, \dots, z_n)^T$ 、 $\mathbf{z}' = (z'_1, z'_2, \dots, z'_n)^T$ とする。

$\hat{p}_{ij}^{-1}[s]$ について証明する。 $\mathbf{z}' = \hat{p}_{ij}[s](\mathbf{z})$ とおくと、 $\mathbf{z}' = \mathbf{z} + \text{int}(s\mathbf{I}_{ij}\mathbf{z})$ となる。これを成分で書くと、第 i 成分に関して、 $z'_i = z_i + \text{int}(sz_j)$ 、それ以外の成分 (第 k 成分) に関して、 $z'_k = z_k$ である。 $\hat{p}_{ij}^{-1}[s]$ においては $i \neq j$ であるから、 $z'_j = z_j$ が成り立つ。したがって、この関係を $z'_i = z_i + \text{int}(sz_j)$ に代入して移項すれば、 $z_i = z'_i - \text{int}(sz'_j)$ を得る。これと $z_k = z'_k$ から $\mathbf{z}' = \mathbf{z} + \text{int}(s\mathbf{I}_{ij}\mathbf{z})$ が成立する。

基本並進運動 $\hat{t}_i^{-1}[s]$ については、上と同様にして $\mathbf{z}' = \hat{t}_i[s](\mathbf{z})$ とおく。すると、 $\mathbf{z}' = \mathbf{z} + \text{int}(s)\mathbf{u}_i$ となる。移項して、 $\mathbf{z} = \mathbf{z}' - \text{int}(s)\mathbf{u}_i$ が得られる。

C.3 [定理 4.3 の証明]

基本斜交軸変換に関する命題

$$|\mathbf{z} - \mathbf{x}| \leq \mathbf{e} \implies |\hat{p}_{ij}[s](\mathbf{z}) - p_{ij}[s](\mathbf{x})| \leq \epsilon\{\hat{p}_{ij}[s]\}(\mathbf{e}) \quad (\text{C.4})$$

と 2 個の誤差演算子の合成に関する命題、

$$|\mathbf{z} - \mathbf{x}| \leq \mathbf{e} \implies |\hat{\beta}\hat{\alpha}(\mathbf{z}) - \beta\alpha(\mathbf{x})| \leq \epsilon\{\hat{\beta}\}\epsilon\{\hat{\alpha}\}(\mathbf{e}) \quad (\text{C.5})$$

だけを証明する。その他の基本変換に関する命題も同様に証明できる。

補題 C.2 定理 4.3 の斜交軸変換 $\hat{p}_{ij}[s](\mathbf{z})$, $p_{ij}[s](\mathbf{x})$ に関して次の不等式が成り立つ。

$$|\text{int}(sz_j) - sx_j| \leq |s||z_j - x_j| + \frac{1}{2} \quad (\text{C.6})$$

[証明]

定義 4.3 より、 $sz_j - \frac{1}{2} \leq \text{int}(sz_j) < sz_j + \frac{1}{2}$ が成り立つ。この不等式の各辺に $-sx_j$ を加えて整理すると、

$$s(z_j - x_j) - \frac{1}{2} \leq \text{int}(sz_j) - sx_j < s(z_j - x_j) + \frac{1}{2} \quad (\text{C.7})$$

となる。以下、 $\text{int}(sz_j) - sx_j < 0$ の場合と $\text{int}(sz_j) - sx_j \geq 0$ の場合に分けて考える。

- $\text{int}(sz_j) - sx_j < 0$ の場合：式 (C.7) の初めの不等式から、 $s(z_j - x_j) - \frac{1}{2} \leq \text{int}(sz_j) - sx_j (< 0)$ である。両辺の絶対値をとると、 $|\text{int}(sz_j) - sx_j| \leq |s(z_j - x_j) - \frac{1}{2}| \leq |s(z_j - x_j)| + \frac{1}{2} = |s||z_j - x_j| + \frac{1}{2}$ となる。したがって、命題が成立する。ここで、不等式 $|a + b| \leq |a| + |b|$ を用いた。
- $\text{int}(sz_j) - sx_j \geq 0$ の場合：式 (C.7) の 2 番目の不等式から、 $(0 \leq) \text{int}(sz_j) - sx_j < s(z_j - x_j) + \frac{1}{2}$ である。両辺の絶対値をとると、 $|\text{int}(sz_j) - sx_j| < |s(z_j - x_j) + \frac{1}{2}| \leq |s||z_j - x_j| + \frac{1}{2}$ となる。したがって、命題が成立する。

C.3.1 [不等式 (C.4) の証明]

定義 4.4 および定義 4.5 より、 $|\hat{p}_{ij}[s](z) - p_{ij}[s](x)| = |\{z + \text{int}(s\mathbf{I}_{ij}z)\} - (x + s\mathbf{I}_{ij}x)| = |(z - x) + \{\text{int}(s\mathbf{I}_{ij}z) - s\mathbf{I}_{ij}x\}| \leq |z - x| + |\text{int}(s\mathbf{I}_{ij}z) - s\mathbf{I}_{ij}x|$ 。第 i 成分以外は $|\text{int}(s\mathbf{I}_{ij}z) - s\mathbf{I}_{ij}x|$ が零である。第 k 成分 ($k \neq i$) に関して、次の不等式が成立する。

$$(|\hat{p}_{ij}[s](z) - p_{ij}[s](x)|)_k \leq |z_k - x_k| \leq e_k \quad (\text{C.8})$$

$(\)_k$ は括弧内のベクトルの第 k 成分を表す。命題の仮定 $|z - x| \leq e$ を用いた。

第 i 成分に関しては、補題 C.2 と $|z - x| \leq e$ を用いる。

$$(|\hat{p}_{ij}[s](z) - p_{ij}[s](x)|)_i \leq |z_i - x_i| + |\text{int}(sz_j) - sx_j| \leq |s||z_j - x_j| + \frac{1}{2} \leq |s|e_j + \frac{1}{2} \quad (\text{C.9})$$

である。不等式 (C.8) と (C.9) は、まとめて次式で表現できる。

$$|\hat{p}_{ij}[s](z) - p_{ij}[s](x)| \leq \mathbf{P}_{ij}[|s|]e + \frac{1}{2}\mathbf{u}_j \quad (\text{C.10})$$

この不等式の右辺は、定義 4.5 における $\epsilon\{\hat{p}_{ij}[s]\}$ の定義に他ならない。よって、不等式 (C.4) は証明された。

C.3.2 [不等式 (C.5) の証明]

任意の基本変換 α に関して、

$$|z - x| \leq e \implies |\hat{\alpha}(z) - \alpha(x)| \leq \epsilon\{\hat{\alpha}\}(e) \quad (\text{C.11})$$

が成立すると仮定して、不等式 (C.5) を証明する。

$x' \equiv \alpha(x)$, $z' \equiv \hat{\alpha}(z)$, $e' \equiv \epsilon\{\hat{\alpha}\}(e)$ と置き換えると、命題 (C.11) の右項は、 $|z' - x'| \leq e'$ となる。これに、ふたたび命題 (C.11) を α を β に替えて適用すると、 $|\hat{\beta}(z') - \beta(x')| \leq \epsilon\{\hat{\beta}\}(e')$ を得る。この不等式の x', z', e' を元に戻せば、命題 (C.5) の右項をえる。この議論を繰り返せば、3 個以上の演算子の合成に関する命題も証明できる。

C.4 [命題 4.4 の証明 (略証)]

アルゴリズム 4.1 において、第 k 行に対する処理を考える。ステップ (2-1), (2-2) の処理の前に、第 k 行から第 n 行までの各行の間の任意の交換と第 k 列から第 n 列までの各列の間の任意の交換が行える。そこで、第 k 行に関する処理の前に 1 に正規化する成分 (u, v) を決め、それが (k, k) の位置に来る様に適当に行と列を交換する。この交換は、基本鏡映変換行列 R_{ku} と R_{kv} をそれぞれ左と右から $A^{(k)}$ に乗ずることに対応する。さらに、第 k 列の対角成分を 1 に正規化する成分が属する列 (アルゴリズム 4.1 のステップ (2-1) で選択する第 j 列) を第 $k+1$ 列から第 n 列までの列から選べる (これは、添字 w の選択に対応する)。これらの選択操作は誤差評価に影響を与える。さらに、第 k 列の符号を反転させる (基本鏡映変換行列 Q_k を右から乗ずることに対応する)。すると、 (k, k) 成分を 1 に正規化するための基本斜交軸変換行列のパラメータを変化させる。しかし、第 k 行の符号を反転させてもそれ以降の誤差評価には影響がない。以上から、第 k 行では、(i) (k, k) に移動させる成分の選択、(ii) (k, k) 成分を 1 に正規化するために用いる列 (第 j 列) の選択、(iii) 第 k 列の符号を反転させるか否かの選択、を行える。これらを考慮して、アルゴリズム 4.1 のステップ (2) を実際に行うと命題の結果が得られる。

C.5 基本変換の交換関係

基本変換同士には、順序を交換可能なものや代替え可能なものがある。その規則を次にまとめる。

性質 C.6

1. $p_{ij}[s]p_{kl}[u] = p_{kl}[u]p_{ij}[s] \quad (i \neq l, j \neq k),$
2. $r_{ij}r_{kl} = r_{kl}r_{ij} \quad (i \neq k, i \neq l, j \neq k, j \neq l),$
3. $r_{ij}r_{ij} = r_{ij}r_{ji} = o \quad (o \text{ は恒等変換}),$
4. $r_{ij} = r_{ji},$
5. $q_iq_j = q_jq_i,$
6. $q_iq_i = o,$
7. $p_{ij}[s]r_{kl} = r_{kl}p_{ij}[s] \quad (j \neq k, j \neq l),$
8. $p_{ik}[s]r_{kl} = r_{kl}p_{il}[s] \quad (i \neq k, i \neq l),$
9. $p_{ij}[s]r_{ij} = r_{ij}p_{ji}[s] = p_{ij}[s]r_{ji} = r_{ji}p_{ji}[s],$
10. $r_{ij}q_k = r_{ji}q_k = q_kr_{ij} = q_kr_{ij} \quad (i \neq k, j \neq k),$

$$11. r_{ij}q_i = r_{ji}q_i = q_j r_{ij} = q_j r_{ji},$$

$$12. p_{ij}[s]q_k = q_k p_{ij}[s] \quad (j \neq k),$$

$$13. p_{ij}[s]q_j = q_j p_{ij}[-s],$$

$$14. r_{ij}q_j = p_{ij}[-1]p_{ji}[1]p_{ij}[-1] \quad \square$$