

エッジベーストステレオによる三次元情報の
獲得とそれを用いた認識に関する研究

杉本 和英

システム情報工学研究科
筑波大学

2008年3月

梗概

3次元視覚機能の実用化は、産業界からその実現が強く望まれている。3次元情報を入力するためのセンシング手法に始まり、獲得された距離情報に基づく物体の3次元形状の再構成や幾何モデルの生成、対象の認識・識別、運動追跡等、果たすべき役割に対する期待に応えるべく、解決すべき課題に対して積極的な取り組みが多くなされている。センシングに関しては、自ら光を照射してその反射光を利用するレンジファインダに代表される能動法と、撮像系のみからなるステレオ（両眼立体視）に代表される受動法に大別でき、レンジファインダは、用途を限定すれば安定して距離データを獲得できるため商品化もなされている。一方ステレオは、人間の眼と同様にパッシブなセンシング手法であり、その役割に対する期待は非常に大きく、3次元再構成のための重要な手段として、これまで対応問題を解決するアルゴリズムをはじめ、テクスチャやシェーディングのある曲面の復元に関する手法が多数提案されてきたが、実用化に向けた標準的な手法の確立が求められている。また、任意の3次元物体の認識も重要な課題の一つであり、モデルに対応する物体を任意のシーン中より検出し、その位置姿勢を実時間で決定する問題は、特に顧客の製品ニーズの多様化に対応する多品種少量生産を必要とする多くの現場から、その実時間処理の実現が望まれている。

本研究では、3次元視覚機能の実用化に向けた汎用性促進への貢献を目指し、その適用範囲の拡大ならびに実時間処理の観点から、エッジベースステレオによる、一様な平面からなる多面体に関する3次元再構成と、モデルベースの3次元物体認識の高速化および一般物体認識に着目し、前者については、単一視点からの平面推定と視点の移動による逐次的モデルの更新による復元を、後者については、ネットワーク計算機による並列分散処理評価に関して研究を進めると共に、モデル爆発の問題回避に向けた定性的な認識手法について、いくつかの応用例による検討を行った。

本論文では、第1章にて研究背景と課題を明確化し、第2章にて、エッジベースステレオの対応単位として用いるセグメントを得るための特徴点検出手法について述べ、第3章にて、表面に模様やシェーディング等の手がかりのない一様な平面からなる多面体に対して、面ベースの記述を生成する3次元再構成手法について述べる。第4章にて、視点の移動により、各視点より得られる面ベースの記述を逐次解釈・統合してより完全な幾何モデルを生成する手法について述べる。第5章にて、構築された幾何モデルを用いた観測データとの照合による3次元物体認識処理の高速化について述べ、第6章にて、トップダウンのモデル駆動型物体認識処理におけるモデル爆発を回避するための、定性的物体認識に関するアプローチとして共通モデルのコンセプトとその応用について言及し、実験結果や応用事例と合わせてその有効性を示す。第7章にて、本研究を総括すると共に、今後の展望について述べる。

キーワード：

特徴点検出、エッジベースステレオ、面の復元、境界表現、幾何モデリング、動的負荷分散、定性的物体認識

SYNOPSIS

It is strongly expected to realize practical use of the three-dimensional (3D) computer vision techniques for object recognition in the industrial field such as localization and pose determination for object manipulation and inspection.

As for the input methods for obtaining 3D information, they can be roughly classified into the following two approaches. One is with an active sensing method and the other is based on a passive way. As for the former, a range finding system such as a laser range finder becomes popular. And for the limited use, such active sensing methods have commercially released and the stable dense depth information can be obtained. On the other hand, as for the latter, a stereo method is well known. It uses only cameras and the depth information can be obtained based on a triangulation principle in a passive way like human being does. So, it is expected that stereo plays the significant role for 3D information acquisition. Based on a stereo vision technique, many algorithms have been proposed to solve a correspondence problem and reconstruction of object surfaces with textures and shading patterns.

In addition, 3D object recognition is one of the important issues to be solved. In a model-based approach, system first detects objects in a scene. Then it determines the 3D position and orientation of the object by matching the shape data with the shape model registered in the database. These should be processed in real time, and realization of such system is strongly expected in the field of production.

In this study, the following two issues are focused and discussed. One is for 3D reconstruction of objects with planar surfaces which have neither texture nor shading information by using edge-based stereo vision. And the other is for the robust and high speed processing for a model based 3D object recognition under the parallel networked computing environment, and the qualitative object recognition to solve so-called model explosion problems.

In chapter 1, the background and issues to be solved are clarified. In chapter 2, a method for dominant point detection on the contour boundary of each region in the image is introduced to obtain components such as a line segment and a curved one (convex or concave) which are used as units for stereo correspondence search. In chapter 3, a method for 3D reconstruction of objects with planar surfaces is presented. In chapter 4, a method for geometric modeling by combining the surface based description which are obtained from different view points by stereo motion is presented. In chapter 5, high speed processing by using networked computer cluster with dynamic load balancing techniques for 3D model-based object recognition is presented. In chapter 6, an approach for qualitative object recognition to solve problems so-called model explosion is presented with some experiments and applications. Chapter 7 summarizes main contributions of this study and introduces future works.

Keywords :

Dominant point detection, Edge-based stereo vision, Surface reconstruction, Boundary representation, Geometric modeling, Dynamic load balancing, Qualitative object recognition

目次

第1章	序論	1
1.1	本研究の背景と目的	1
1.2	本研究の具体的課題	2
1.3	本論文の構成	3
第2章	特徴点検出による境界線の分割	5
2.1	具体的課題と関連研究ならびに特徴点検出の概要	5
2.2	処理の流れと諸定義	7
2.2.1	処理の流れ	7
2.2.2	点列の獲得	7
2.2.3	特徴点の種類	8
2.2.4	曲線近似誤差	8
2.3	曲率と法線方向ならびに近似誤差算出	9
2.3.1	曲線近似と誤差の算出	9
2.3.2	曲率と法線方向の算出	11
2.4	最適近似区間の決定と曲率・法線方向の更新	11
2.4.1	最適近似区間の決定	11
2.4.2	曲率・法線方向の更新	12
2.5	特徴点検出	13
2.5.1	屈曲点の検出	13
2.5.2	遷移点の検出	15
2.5.3	変曲点の検出	15
2.6	パラメータ k の動的適応	17
2.6.1	雑音点の検出	17
2.6.2	パラメータ k の適応	17
2.6.3	検証	17
2.7	実験	19
2.8	第2章のまとめ	20
第3章	エッジベースステレオによる平面からなる多面体の三次元再構成	23
3.1	具体的課題と関連研究ならびに概要	23
3.2	ステレオカメラ構成と入力画像	24
3.2.1	輻輳カメラと標準カメラモデル	24

3.2.2	カメラキャリブレーション	26
3.3	画像の境界表現と対応探索	29
3.3.1	画像の境界表現 (データ構造)	29
3.3.2	ステレオ対応探索	30
3.3.3	多重対応の除去	31
3.4	ステレオデータ	32
3.5	面の復元	34
3.5.1	オクルージョンの検出	34
3.5.2	実在テスト	36
3.6	境界線の分類と解釈不可能なループ	41
3.6.1	境界線の分類	41
3.6.2	仮セグメント	43
3.6.3	解釈不可能なループ	43
3.7	第3章のまとめ	44
第4章	移動ステレオによる統合と解釈	47
4.1	具体的課題と関連研究ならびに概要	47
4.2	世界と記述	48
4.3	データの統合と再解釈	50
4.3.1	セグメントの信頼性	50
4.3.2	実在テスト2	51
4.3.3	データの統合	52
4.3.4	対応探索	53
4.3.5	再解釈	55
4.3.6	検証	56
4.3.7	考察	59
4.4	第4章のまとめ	60
第5章	物体認識処理の高速化	65
5.1	具体的課題と関連研究ならびに概要	65
5.2	エッジベースステレオによる物体認識	66
5.3	微調整処理の並列化	69
5.4	並列性能評価	71
5.4.1	プロセッサの性能比	72
5.4.2	分配タスク数による性能評価	72
5.4.2.1	初期分配タスク数による性能評価	72
5.4.2.2	再分配タスク数による性能評価	73
5.4.3	プロセッサ数による性能評価	74
5.4.4	動的負荷分散の効果	76
5.4.5	他のモデルとシーンにおける実験	77

5.5	第5章のまとめ	78
第6章	定性的物体認識へのアプローチ	81
6.1	具体的課題と関連研究ならびに概要	81
6.2	共通モデル	83
6.3	ロゴマーク検出への応用	84
6.3.1	教示システム	85
6.3.2	実験	86
6.4	斜め対称パターンの検出	87
6.4.1	候補軸の検出1 (2対の直線セグメントより)	88
6.4.2	候補軸の検出2 (1対の直線セグメントと1直線セグメントより)	89
6.4.3	候補軸の検出3 (1対のコニックより (1))	90
6.4.4	候補軸の検出4 (1対のコニックより (2))	91
6.4.5	候補軸の検出5 (1非コニックより)	91
6.5	斜め対称プリミティブの統合	92
6.5.1	斜め面对称プリミティブの検出	94
6.5.2	未知対象物体把持動作計画への利用	96
6.6	手指の検出と運動追跡への応用	97
6.6.1	手指領域の検出	98
6.6.2	手指領域の運動追跡	100
6.6.3	運動情報の算出	102
6.7	第6章のまとめ	104
第7章	結論	107
付録A	光感応蛋白質を利用する インテリジェントイメージセンサ	111
	付録	111
A.1	概要	111
A.2	光感応蛋白質による移動物体の特徴抽出	112
A.2.1	特徴抽出原理	112
A.2.2	実験システム	113
A.3	光感応蛋白質イメージセンサの出力特性	114
A.3.1	光感応蛋白質イメージセンサ受光部	114
A.3.2	画素電極アレイにおけるBR応答出力	114
A.3.3	移動物体の特徴量分離抽出処理	116
A.4	まとめ	120
付録B	平面拘束定理	121

謝辞	123
参考文献	123
公表論文リスト	143

目次

2.1	特徴点の種類	8
2.2	円および直線のあてはめと誤差, 曲率, 法線	10
2.3	屈曲点近傍における曲線近似と近似誤差	12
2.4	k -曲率	13
2.5	曲率の更新	14
2.6	法線方向の更新	15
2.7	曲線における屈曲点検出	16
2.8	特徴点の検出結果	16
2.9	k の動的適応による特徴点再検出	18
2.10	特徴点抽出結果 1	19
2.11	入力原画像および特徴点検出結果 2	20
3.1	輻輳カメラモデル	25
3.2	標準カメラモデル	25
3.3	ステレオ原画像と輻輳補正後の境界線セグメント画像	27
3.4	キャリブレーションパターン	28
3.5	画像の境界表現	30
3.6	Winged-edge 構造	31
3.7	3次元セグメント 1	33
3.8	ブロック	33
3.9	3次元セグメント 2	34
3.10	オクルージョン	34
3.11	明らかな実セグメント (太線) と明らかな虚セグメント (点線)	35
3.12	オクルージョン点 (○) の検出	36
3.13	重なったブロック	37
3.14	実在テストの結果	38
3.15	面の候補	38
3.16	実在テストの結果	38
3.17	穴のあいたブロック	39
3.18	面の候補	39
3.19	実在テストの結果	40
3.20	面の外周ループと穴のループの統合	40
3.21	稜線の両側の法線	41

3.22	境界線の分類 (1)	41
3.23	境界線の分類 (2)	42
3.24	境界線の分類 (3)	43
3.25	仮セグメント (点線)	44
3.26	解釈不可能なループ	44
3.27	幾何モデル	46
4.1	シーン中に存在する物体	49
4.2	物体の境界表現	49
4.3	紙の裏と表の面を表す二つのループ	50
4.4	解釈不可能なループ 1	51
4.5	解釈不可能なループ 2	51
4.6	第 1 の視点における世界地図	52
4.7	第 2 の視点における局所地図	53
4.8	セグメントの対応	54
4.9	初期世界地図 (実線) と第二の局所地図 (破線) との対応	55
4.10	世界地図を用いた局所地図の再解釈	56
4.11	異なる視点から見た同一のループ	57
4.12	第二の観測後に更新された世界地図	58
4.13	第 3 の視点での局所地図との対応探索と統合	62
4.14	エッジベースステレオによる面ベースの記述と統合・更新された世界地図	63
5.1	初期照合と微調整	66
5.2	頂点を用いた初期照合	67
5.3	対応点探索	68
5.4	モデルデータ 1	68
5.5	認識結果 1	69
5.6	動的負荷分散	70
5.7	ワークステーションクラスタ	71
5.8	プロセッサの性能比	72
5.9	初期分配タスク量と実行時間の関係	73
5.10	再分配タスク数と実行時間の関係	74
5.11	プロセッサ数に対する実行時間	75
5.12	台数効果	76
5.13	各プロセッサにおいて処理されたタスク数	76
5.14	モデルデータ 2	77
5.15	モデルデータ 3	77
5.16	認識結果 2	78
5.17	認識結果 3	79
6.1	共通モデル	84

6.2	教示システム	85
6.3	一般的な関係	86
6.4	ロゴマーク認識結果（アピアランス変化への対応）	86
6.5	文字”S”の記述例と画像シーケンスからの対象フレームの抽出例	87
6.6	T交差	89
6.7	2対の直線セグメント	89
6.8	1対の直線セグメントと1直線セグメント	90
6.9	1対のコニック	91
6.10	1非コニック	92
6.11	入力原画像とエッジ画像	93
6.12	局所斜め軸対称パターンの検出結果	93
6.13	斜め軸対称プリミティブの検出結果	94
6.14	斜め対称軸端点の決定	95
6.15	斜め軸対称図形の統合	95
6.16	斜め面対称図形の検出	95
6.17	把持要素例	96
6.18	把持動作計画への利用例	97
6.19	親指の記述例と手指検出結果	99
6.20	手指の教示サンプル	100
6.21	追跡点と（○）と仮の対応点（+）	101
6.22	開口問題	102
6.23	手指の運動追跡	102
6.24	手指の動きに連動した仮想空間中の物体操作例	103
6.25	仮想空間中の物体操作デモシステム	104
A.1	特徴抽出原理	112
A.2	実験システム	113
A.3	入力動画像、背景画像および2D画素電極アレイ	115
A.4	電流時系列データ	116
A.5	画素電極（横一列分）における電流応答	117
A.6	特徴量分離抽出処理	118
A.7	(a)入力動画像, (b)1次データ, (c)輪郭・移動方向抽出結果, (d)物体領域抽出結果	119

表目次

4.1 各観測点データ中の3次元セグメント数, ループ数等	59
5.1 初期照合と微調整の実行時間と対応候補数	69
5.2 使用プロセッサ数と認識処理時間	77

第1章 序論

1.1 本研究の背景と目的

コンピュータによる視覚情報処理—コンピュータビジョン—は2次元の図形を認識する2次元視覚と、立体を認識する3次元視覚に大別でき、前者は文字認識や検査等、特定の条件下で実用化されているのに対し、後者の実用化例は少なく、産業界からその実現が強く望まれている。3次元視覚の最初の処理である3次元情報の入力手法は、自ら光を照射してその反射光を利用するレンジファインダ [1] に代表される能動法と、撮像系のみからなるステレオ（両眼立体視）に代表される受動法に大別できる [2]。レンジファインダは、用途を限定すれば安定して距離データを獲得できるため商品化もなされている。一方ステレオは、人間の眼と同様にパッシブなセンシング手法であり、知能ロボットに共通的な眼として、その役割に対する期待は非常に大きく、3次元計測・再構成・認識のための重要な手段として、これまで対応問題を解決するアルゴリズムをはじめ、テクスチャやシェーディングのある曲面の復元に関する手法が多数提案されてきたが、実用化に向けた標準的な手法の確立が求められている。

特にステレオの課題の一つである対応探索手法については、古くは局所的相互相関を計算する手法から、弛緩法・粗密解析・動的計画法等によりエッジ点の対応を求める手法が提案されているが、対応付けの曖昧さをできる限り排除するために対応の単位を大きくし、付随する情報量を増加させる観点から、本研究では、画像の境界表現から得られる領域とその境界線を用いる手法に基づき、境界線の連続性を対応の強力な拘束として利用することとした。画像中の領域の境界線セグメントを単位として対応探索する際、境界線を分割してセグメント化するために、境界線上の屈曲点・変曲点・遷移点といった特徴点の検出が重要となる。前記特徴点により境界線を分割することで、直線セグメントと凹もしくは凸の曲線セグメントが得られる。得られた境界線セグメントを用いてステレオ対応を求めることにより、エッジ部分に関して3次元情報を獲得することができる。しかしながら、単一の視点から観測されるシーンより3次元再構成を行う際、照明条件・雑音に起因する誤対応等により正しく再構成できない領域や、隠れ等により観測できない領域等については、別の視点より獲得される3次元情報との統合による再解釈が必要となる。

また、3次元情報の獲得・再構成と並び、3次元物体の認識はコンピュータビジョンの当初からの重要な課題の一つであり、モデルに対応する物体を任意のシーン中より

検出し、その位置・姿勢を実時間で決定する問題は、生産ラインの自動化のみならず、画像内容検索等幅広い応用が見込まれる。特に顧客からの製品ニーズの多様化に対応する多品種少量生産を必要とする多くの現場から、その早期実現が望まれている。

本研究では、人間の目と同様に、受動法による3次元視覚に着目し、単一の視点より得られるステレオ画像に基づくエッジベースステレオにより、テクスチャやシェーディングの無い一様な平面から構成される対象に対して、各面の輪郭境界線の3次元情報を獲得し、前記境界線により囲まれた領域に対して偽の面を生成することなく平面を内挿する手法、各視点にて再構成された3次元情報を、視点の移動により逐次新たに得られる3次元情報と統合・再解釈することにより対象のより完全な幾何モデルを生成する手法、3次元モデルと観測データとの照合により対象の位置姿勢を推定する認識処理の高速化に向け、ネットワーク接続された性能の異なるワークステーションクラス上で台数効果を得るための動的負荷分散処理に関する手法、また、モデルベースの認識処理における組み合わせ爆発を回避するための定性的物体認識手法に関して研究すると共に、運動追跡といった実時間動画処理の実現に向けて、高効率の動画特徴抽出処理を行うインテリジェントイメージセンサ構築の可能性についても検討を行う。これら適用範囲の拡大と処理の高速化により、コンピュータビジョン手法の汎用性の促進に貢献することを目的とする。

1.2 本研究の具体的課題

境界線セグメントを対応単位とするステレオの出力より、対応の求まるエッジ部分に関しては3次元情報を獲得することができる。このとき、対応探索空間の不要な増大を防ぐ観点からも、境界線は、滑らかな曲線部分が多角形近似されたりすることなく、直線と凹もしくは凸の曲線セグメントに正しく分割されることが重要である。また、ステレオにより、テクスチャやシェーディングのある曲面に関しては、密な3次元情報を獲得することにより復元が可能であるが、多くの人工物がそうであるように、一様な平面で構成された多面体に対しては、境界線内部の特徴のない部分を直接対応によって復元することができない課題がある。従って、境界線の3次元復元の後、面を復元する処理が必要となる。しかし、物体の輪郭線（奥行きの不連続部分）と稜線（面の法線方向の不連続部分）を区別しない単純な面内挿アルゴリズムでは、偽の面を生成してしまう。また、照明条件や雑音の影響により面の推定ができなかった領域については、視点の移動による再観測・再解釈が必要となる。移動により新たに観測された面との統合も含め、異なる観測点から得られる部分的なステレオデータ間での対応探索や、不完全なステレオデータを逐次的に統合、更新して、面表現による記述を生成できることが重要である。

エッジベースステレオによる物体認識においては、シーンから検出された物体の

特徴点と、あらかじめ構築された対象物体の幾何モデルとの照合により対応する部分を探索し、その位置・姿勢を決定するトップダウンのモデル駆動型アプローチが一般的である。このとき、認識対象とする物体上の特徴点数や照合すべきモデル数の増加に比例して多くの処理時間を要することとなるため、処理の高速化が重要となる。更に、実世界の認識対象すべてに対して詳細な幾何モデルを用意することは難しく、この所謂モデル爆発の問題も解決されるべき課題である。

また、物体の運動追跡といった実時間動画像処理においては、各画像フレームにおける特徴抽出や、フレーム間での対応探索処理等に関する処理の効率化・高速度化が重要となる。このとき、イメージセンサの応答速度の向上のみならず、高効率の動画像特徴抽出処理機能を具備したインテリジェント化への期待も大きい。

1.3 本論文の構成

本論文は、第1章序論、第2章特徴点検出による境界線の分割、第3章ステレオによる平面からなる多面体の再構成、第4章移動ステレオによる統合と解釈、第5章物体認識処理の高速化、第6章共通モデルに基づく定性的な認識、第7章結論、から構成される。尚、インテリジェントイメージセンサの構築可能性の検討に関しては別添とした。

第2章では、特徴点検出による境界線のセグメンテーションについて述べる [3]。ここでは、境界線上の各点における曲率ならびに法線方向といった局所特徴算出の際の、最適近似区間の決定ならびに近似誤差に応じた近傍画素決定パラメータの動的適応手法について述べる。提案手法により、屈曲点・変曲点・遷移点等の特徴点を検出できることを示す。

第3章では、エッジベースステレオにより、表面に模様やシェーディング等の手がかりが存在しない様な平面からなる多面体に関して、面ベースの記述を生成する3次元再構成手法について述べる。ここでは、面を推定する際の重要な手がかりとなるオクルージョン（隠蔽箇所）検出手法、平面の推定手法、一意な解釈が可能な平面の内挿による多面体の再構成手法について述べる [4]。これにより、面ベースの記述が生成できることを示す。

第4章では、移動ステレオによる逐次的な幾何モデリングとして、視点の移動により各視点より得られる面ベースの記述を逐次解釈・統合する手法について述べる [5]。提案手法により、単一視点から得られたデータでは解釈できなかった部分を、新たに得られたデータを用いて再解釈・統合することにより、より完全なモデルを逐次構築できることを示す。

第5章では、構築された幾何モデルを用いた観測データとの照合による3次元物体認識に関して、処理の高速化に重点を置き、性能の異なるネットワーク計算機（ワー

クステーションクラスタ) を用いて台数効果を得るために, マスタ/スレーブによるタスクプール方式を用いた動的負荷分散により, モデル照合処理における並列処理性能を評価 [6] し, その有効性を示す.

第6章では, トップダウンのモデル駆動型物体認識処理におけるモデル爆発の課題を解決するために, 特徴点により分割された境界線セグメントを一般的な関係によりボトムアップ並列的に組み合わせることで, 対象物を定性的に記述する「共通モデル」のコンセプト [7] と, 共通モデルに基づく認識・識別手法について述べる. ロゴマーク検出や対称性のある局所特徴抽出への適用により, アピアランスの変化への対応可能性を, ハンド・アイシステムによる対象物把持動作計画への応用 [8] や人の手指の運動追跡による立体モデルのハンドリング [9] への適用により, 探索空間の削減や個体差の吸収といったモデル爆発問題の解決可能性に関して, 応用事例と合わせて提案手法の有効性を示す.

第7章では, むすびとして, 本研究で得られた結果を総括する.

また, 動画像処理の高速化へのアプローチとして, 光感応蛋白質を利用したハードウェアとしてのインテリジェントイメージセンサ構築を通じ, 設計と試作 [208], 移動体検出実験・評価も行った. これらに関しては, 付録Aとして添付しているので合わせて参照されたい.

第2章 特徴点検出による境界線の分割

2.1 具体的課題と関連研究ならびに特徴点検出の概要

画像上の点や直線等の幾何学的プリミティブに基づいた解析は、3次元視覚情報処理における3次元復元・認識の際に重要な技術であり、画像よりそれらプリミティブをどう抽出するかが課題といえる。ステレオの対応探索の際にも点や直線を単位としてそれらの対応を如何に決定するかが重要となる。画素を単位として対応を求める際、計算コストや対応の取り易さを考慮して色や濃度変化の激しい点に着目し、角点 (corner)、関心点 (interest point)、特徴点 (feature point) と呼ばれる点を選ぶことが多い。そして、それらを抽出するために特徴点抽出作用素 (operator) を用い、それらが極値を与える点を抽出する手法が一般的である。微分幾何学に基づくアプローチとして、一次微分とガウス平滑化の組合せにより評価する Harris 作用素 [10] が最も良く用いられ、動画像上での特徴点追跡の為の応用としての KLT 法 [11] や、直線の交点を検出するための変形による手法 [12] も提案されている。その他、微分を用いず、テンプレートとの局所相関による手法 [13]~[15] も提案されているが、それらは多くのテンプレートや高次特徴空間 [16][17] を必要とする。注目点近傍の輝度値の8方向の分散を直接評価し、一つのマスクで済ませる作用素を用いる手法 [18] やその改良手法 [19] も提案されているが、それらは輝度値の変動に基づくものといえる。それらに対して、本章にて述べる特徴点検出手法は、2次元画像中もしくはそれら进行处理することにより抽出される物体の輪郭 (線図形) やその離散曲率から、頂点や輪郭線上の特徴点を検出するアプローチといえる。

エッジベースステレオにおいては、対応探索の際の対応単位として用いる直線や曲線セグメントといった幾何学的プリミティブを正確に抽出することが重要となる。そのために、2次元画像中もしくはそれら进行处理することにより得られる領域の輪郭境界線 (以下、境界線とする) や線図形の特徴点を精度良く抽出し、直線や凹もしくは凸の滑らかな曲線からなるセグメントに分割できることが重要となる。ここで、特徴点としては、端点、分岐点、屈曲点 (角点) のみならず、凹凸の反転する変曲点、直線から滑らかに曲線に移行する (直線と曲線が接続し、かつ曲率変化が滑らかな) 遷移点まで検出できる必要がある。端点、分岐点は連結数により容易に検出できるため、本章では、残りの連結数が2となる境界線上の特徴点を検出する手法について述べる。

与えられた画像データから得られる輪郭線や線図形から特徴点を正確に検出する処理は、図面の認識 [20] および文字のベクトル化 [21] や、画像の圧縮・符号化 [22] といった、線図形を自動的に少ないデータ量で忠実に近似表現する処理 [23]~[26] のみならず、波形解析や形状の認識・記述 [27]~[29] においても大変重要な役割を持つといえ、それらを正確に検出する処理の開発が強く望まれている。

また、多面体のような人工物等に関しては直線部分が多く見られるため、直線部分は1本の直線セグメントとして抽出できることが好ましい。更に、人間の手といった自由曲面物体に関しては変曲点や遷移点といった特徴点を検出できることが、効率的な形状特徴抽出ならびにそれらに基づく3次元位置・姿勢の認識や運動追跡処理において重要となる。

特徴点検出に関する従来手法の多くは、 k -曲率 [30] を用いた屈曲点検出に関するものがほとんどであった [31]~[33]。しかしながら、 k -曲率は、雑音の影響を受けやすく、ガウス関数等を用いた低域フィルタによるぼかし処理を施されることが多い。これは、形状の階層表現や曲率波形のふるまいを求める手法においても一般的に用いられている [34]~[37]。ぼかし処理を施した曲率や、その曲率を微分して得られるゼロ交差から屈曲点を検出する手法は多いが [25][38]~[40]、変曲点や遷移点の検出について明確に言及しているものは少ない。これらは、曲率の変化が滑らかな点であり、極大・極小といった特徴としてとらえられないことによる。これらを精度良く検出するためには、曲率を算出する際の接線方向の差分をとるとききざみ幅 (k -曲率の k に相当) を対象の図形の大きさや輪郭形状の細かさに応じて変えなければならない。あまり小さいきざみ幅で差分をとると、曲線のデジタル化による凹凸を拾ってしまい特徴が埋もれてしまう。ぼかし処理と同様な効果を得、雑音の影響を受け難くするには、大きな k 値が好ましい。しかしながら、大きな k 値では曲率がぼかされ正しく算出することができないため、 k 値よりも細かな形状特徴については正しく検出することができなくなる。細かな輪郭形状部分においても、曲率や法線方向を正しく算出するためには、輪郭形状に応じて k 値を動的に適応させる必要がある。

特に曲線図形の場合には、多角形近似による処理手法も提案されているが、近似による誤差の影響を無視することはできず、より詳細な情報を得るのには適さない。また、2次曲線をあてはめ、パラメータ表現可能な区間を抽出して曲線を分割する場合、最適なあてはめ区間の決定は難しく、あてはめ誤差の大きい点で分割するといった処理となる。その場合、曲率の符合 (凹凸) が等しく、かつ、接線 (法線) 方向が滑らかに変化する点においても過剰に分割されてしまう。

本章では、これらの問題点を克服した特徴点検出手法について述べる。まず、各点における曲率、法線ベクトルといった局所特徴と近似誤差の算出手法を紹介し、次に、近傍画素間での近似誤差の比較による最適近似区間の決定と、前記最適近似区間において算出される特徴量に基づいて各点における局所特徴を更新する手法を紹介する。

続いて近似誤差の大きい点に対して、近傍画素を決定するパラメータ k を狭めた局所近似により局所特徴と近似誤差を再算出する、パラメータ k の動的適応手法を紹介する。これにより、屈曲点のみならず、凹凸の反転する変曲点や直線から滑らかに曲線に移行する遷移点についても検出可能となることを示す。

2.2 処理の流れと諸定義

2.2.1 処理の流れ

以下の手順に従い、2.2.2に述べる点列を対象とした特徴点検出・境界線の分割手法について述べる。

- 1) 点列の獲得
- 2) 曲率，法線（接線）方向，誤差の算出（法線から接線への変換，接線から法線への変換は容易なため，これ以降は法線と統一する）
- 3) 曲率，法線方向の更新
- 4) 特徴点の検出
- 5) パラメータ k の動的適応
- 6) 検証

2.2.2 点列の獲得

入力画像より抽出された領域の境界線（線図形）の各画素は，格子間隔を1とした正方格子状に配置され，その8連結で接続する点列として与えられるものとする。与えられた入力画像より順序（方向）付けされた点列を得る方法の具体例を以下に述べる。

- 1) TVカメラやスキャナ等によりデジタル画像として計算機に入力する。
- 2) 入力画像に対してしきい値処理にてエッジ検出を施し，更に延長処理・細線化により，一様な明るさの画素からなる領域を囲む境界線である8連結で接続する画素列を獲得する。
- 3) 上記境界線を領域毎に求め，境界線上の画素列を，例えば領域を右に見るような向きに順序付けしたものを点列として抽出する。
- 4) 連結数により，端点・分岐点を検出する。

このようにして得られたすべての点列に対して、本手法を適用する。このほかにも、与えられた線図形を、順序付けされた8連結で接続する画素列に変換さえすれば、本手法の適用が可能である。

2.2.3 特徴点の種類

本手法にて検出対象とする特徴点は以下の通り。

屈曲点：法線方向が急激に変化する点（角点，図2.1(a)参照）

変曲点：曲率の正負（凹凸）の反転する点（図2.1(b)参照）

遷移点：曲率が零から非零（非零から零も同義）へ移行する点（直線と曲線が滑らかに接続する点）（凹凸属性が等しく、曲率の異なる曲線同士が滑らかに接続するような点は、遷移点とは考えない）（図2.1(c)参照）

雑音点：曲線近似誤差（2.2.4参照）がしきい値を越える点

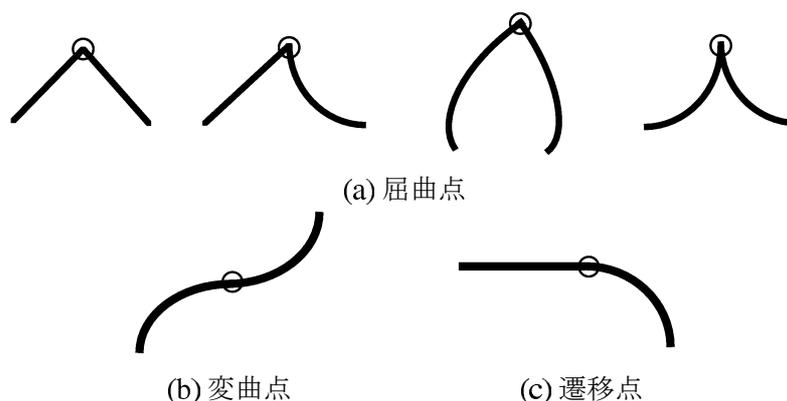


図 2.1: 特徴点の種類

2.2.4 曲線近似誤差

点列の各画素における曲率・法線方向を算出する際に、注目画素の k 近傍画素を用いて局所的に2次曲線（直線もしくは円）をあてはめる。各近傍画素とあてはめられた近似曲線上の最近傍点間のユークリッド距離を、注目画素の k 近傍画素すべてに対して求め、その最大値を注目画素における曲線近似誤差とする。曲線近似誤差として、最大値のほかに、各 k 近傍画素と、近似曲線上の最近傍点間のユークリッド距離の総和を用いる方法も試みたが、近傍画素各点の曲線近似誤差が全体的に小さく、少数の

近傍画素でスパイク状の雑音が含まれる場合と、近傍画素各点に一樣に誤差が含まれる場合を区別することができないため、最大値を採用することとした。

2.3 曲率と法線方向ならびに近似誤差算出

2.3.1 曲線近似と誤差の算出

画像処理の沿革から見れば、 k -曲率は計算量を低減する手法として位置付けられるが、比較的少ない計算量と引きかえに、以下の課題がある。

- 雑音に敏感である。すなわち、注目点の1画素のずれが計算される曲率に大きく影響し得る。そして、この傾向は k が小さいほど強い。この場合、なんらかの形でぼかし処理を施さざるを得ない場合が多い。
- k が大きい場合、ある注目点付近の曲率を遠くの点を参照して求めることとなるため、輪郭線細部の形状を反映しにくい。

そこで、領域境界線上の点列の各点について、注目画素の k 近傍画素を用いて、次式で表される2次曲線（直線もしくは円）を局所的に最小二乗近似によりあてはめることにより、曲率・法線方向を算出することとする。

$$a \cdot (x^2 + y^2) + b \cdot x + c \cdot y + d = 0 \quad (2.1)$$

しかし、直線に近くなるほどパラメータ a は小さくなる。近似誤差を小さくするためには、 a に対して、直線判定のためのしきい値を設定しなければならない。これを避けるために、円と直線を独立にあてはめ、それぞれに対するあてはめ誤差を比較し、その小さい方を選択することにより注目画素の仮の曲率と仮の法線方向を求める。以下にその手順について詳しく述べる。

図2.2に示すように、注目画素 P_i における仮の曲率と法線方向を算出するために、 P_i の k 近傍画素 P_j ($i - k \leq j \leq i + k$)（ここで k は近傍画素を決定するためのパラメータ）を用いて局所的に、次式で表される直線 l_{p_i} と円 c_{p_i} を最小二乗近似によりあてはめる。

$$\begin{aligned} l_{p_i} &: a \cdot x + b \cdot y + c = 0 \\ c_{p_i} &: x^2 + y^2 + a \cdot x + b \cdot y + c = 0 \end{aligned} \quad (2.2)$$

l_{p_i} に対する近似誤差 L_{e_i} の算出方法を以下に示す。

- 1) P_j より l_{p_i} へ下ろした垂線の足 F_j と P_j とのユークリッド距離を E_j とする。
- 2) あてはめに用いたすべての k 近傍画素 P_j に対する誤差 E_j の最大値 L_{e_i} を求める。
- 3) L_{e_i} を l_{p_i} に対する近似誤差とする。

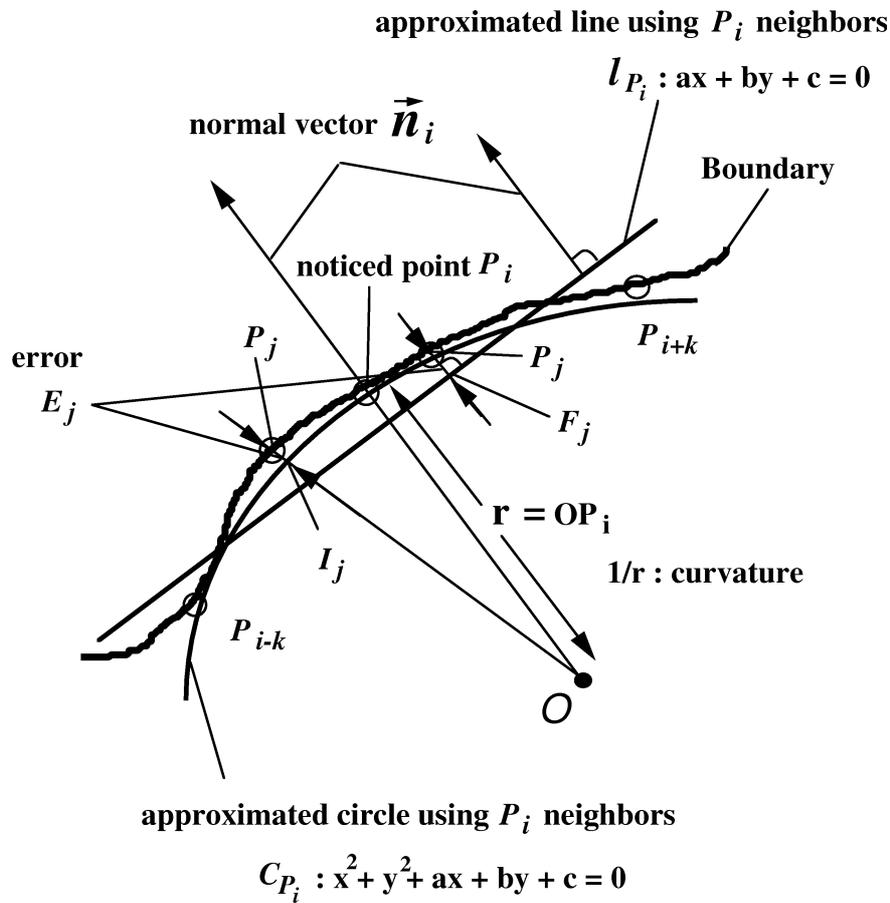


図 2.2: 円および直線のあてはめと誤差, 曲率, 法線

次に, c_{p_i} に対する近似誤差 C_{e_i} の算出方法を以下に示す.

- 1) c_{p_i} の中心を O とする.
- 2) O と P_j を通る直線と c_{p_i} との 2 交点のうち, P_j に近い方を I_j とする.
- 3) P_j と I_j 間のユークリッド距離を E_j とする.
- 4) あてはめに用いたすべての k 近傍画素 P_j に対する誤差 E_j の最大値 C_{e_i} を求める.
- 5) C_{e_i} を c_{p_i} に対する近似誤差とする.

しかし, k 近傍画素を用いた最小二乗近似による直線と円のあてはめを点列の全点に対して施す処理の計算コストは高い. そこで, 注目画素 P_i に対して, k 近傍画素 P_j の両端の 2 点 P_{i-k}, P_{i+k} を用いた局所的な直線近似と, この 2 点に加え P_i 自身とを用いた円近似を採用した.

まず、2点 P_{i-k}, P_{i+k} を通る直線 l_i を求め、 l_i に対して注目画素 P_i の k 近傍画素 P_j ($i-k \leq j \leq i+k$) 各点から l_i へ下ろした垂線の足 F_j までのユークリッド距離 d_j を計算し、その最大値を近似誤差 L_{e_i} とする。直線性を判定するためのしきい値 d_{th} を定め、 $L_{e_i} \leq d_{th}$ であれば直線と判定する。次に、 $L_{e_i} > d_{th}$ の場合のみ曲線と判定し、注目画素 P_i および P_{i-k}, P_{i+k} の3点を通る円を求める。 C_{e_i} の算出方法は、最小二乗近似の場合と同様のため省略する。

2.3.2 曲率と法線方向の算出

次に、点列の各点について、仮の曲率と仮の法線ベクトルを計算する。直線と円を独立にあてはめた場合には、それぞれに対する近似誤差を比較し、 $L_{e_i} \leq C_{e_i}$ なら直線上の点とみなして仮の曲率を0とし、 l_{p_i} に直交し、点列の方向と90度半時計回りの（領域から外を向く）単位ベクトルを仮の法線ベクトル \vec{n}_i とする。 $L_{e_i} > C_{e_i}$ の場合は、 c_{p_i} の半径を r とした場合、 $\frac{1}{r}$ を仮の曲率とする。但し、円の中心が点列の向きに対して左右どちら側にあるかにより符号付けをする（右にある場合を+（凸）、左にある場合を-（凹）としている）。そして、円の中心から注目画素へ向かう単位ベクトルを、仮の法線ベクトルとする。

2.4 最適近似区間の決定と曲率・法線方向の更新

しかしながら、屈曲点近傍や雑音の大きい部分では、近似誤差が大きくなり、曲率・法線ベクトル共に正しく算出することができない。そこで、注目画素の k 近傍画素に対する近似誤差を比較し、近似誤差が最小となる点を求め、この近似誤差最小点における仮の曲率と仮の法線ベクトルを、注目画素におけるそれとして採用することにより、曲率と法線ベクトルをそれぞれ更新する。

2.4.1 最適近似区間の決定

いま、図2.3に示す屈曲点 P_i における k 近傍画素を用いた曲線近似の例を考える。注目画素 P_i における近似直線と近似円をそれぞれ l_i, c_i とする。

屈曲点近傍においては、 l_i に対する近似誤差 L_{e_i} 、 c_i に対する近似誤差 C_{e_i} のいずれもが大きな値となる。そして、直線部分であるにもかかわらず円の近似誤差が小さくなる場合やその逆の場合もあり、正しい曲率も法線方向も求めることができない。そこで、 P_i の k 近傍画素 (P_j ($i-k \leq j \leq i+k$)) に対する近似誤差 E_j ($i-k \leq j \leq i+k$) を比較し、近似誤差が最小となる点 P_{i-min} を求める。なぜなら、 P_j における近似直線 l_i もしくは近似円 c_i の算出に P_i が貢献している (P_i が l_i もしくは c_i 上に乗っている) と考えられるからである。

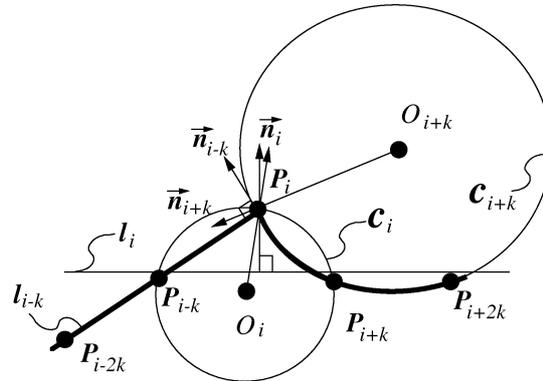


図 2.3: 屈曲点近傍における曲線近似と近似誤差

同図の例で，仮に P_{i-k} におけるあてはめ誤差 $L_{e_{i-k}}$ が最小となるものとするれば， P_i は直線 l_{i-k} 上の点とみなすことができ，曲率は0， P_i における法線ベクトルは， P_{i-k} における仮の法線ベクトル \vec{n}_{i-k} に等しくなるものとする．このとき， P_{i-k} を中心として， P_j ($i-2k \leq j \leq i$) を， P_i における最適近似区間とみなすことができる．

また，仮に P_{i+k} におけるあてはめ誤差 $C_{e_{i+k}}$ が最小となるとき， P_i を円 c_{i+k} 上の点とみなし，曲率は円 c_{i+k} の半径の逆数とし， P_i における仮の法線ベクトルを，円 c_{i+k} の中心 O_{i+k} から P_i へ向かう単位ベクトル \vec{n}_{i+k} とする．同様に， P_{i+k} を中心とする， P_j ($i \leq j \leq i+2k$) が， P_i における最適近似区間とみなされる．

2.4.2 曲率・法線方向の更新

近似誤差が最小となる点 P_{i-min} の仮の曲率と仮の法線ベクトルを，注目画素 P_i におけるそれとして採用することにより，曲率と法線ベクトルをそれぞれ更新する．このとき， P_i の誤差 (L_{e_i} もしくは C_{e_i}) を最小の近似誤差値 ($L_{e_{i-min}}$ もしくは $C_{e_{i-min}}$) と置換する．この処理を，点列の全点に施すことにより，各点において局所的に最も妥当と思われる曲率と法線方向に更新する．

図 2.8 に示した図形について，各画素において計算された曲率を図 2.4 に示す．同図中， k -曲率 ($k=10$) を破線で，ガウシアンフィルタ ($\sigma=10$) にてスムージングした曲率を実線で示している．前者は，雑音に敏感であること，後者は，スムージングにより曲率の変化の小さい点 (150 から 500 の区間) ではほとんど 0 となり，また，屈曲点近傍ではなまっている．(図 2.4 中の横軸の Position は，図 2.8 中の図形の左上角を起点 (Position=0) に，反時計回りに追跡している．画像サイズは， 512×512 である．)

図 2.5 に，同図形に対して，本手法により計算された曲率を示す．実線と破線はそれぞれ，更新前の各画素の仮の曲率，およびあてはめ誤差が最小になる点により更新

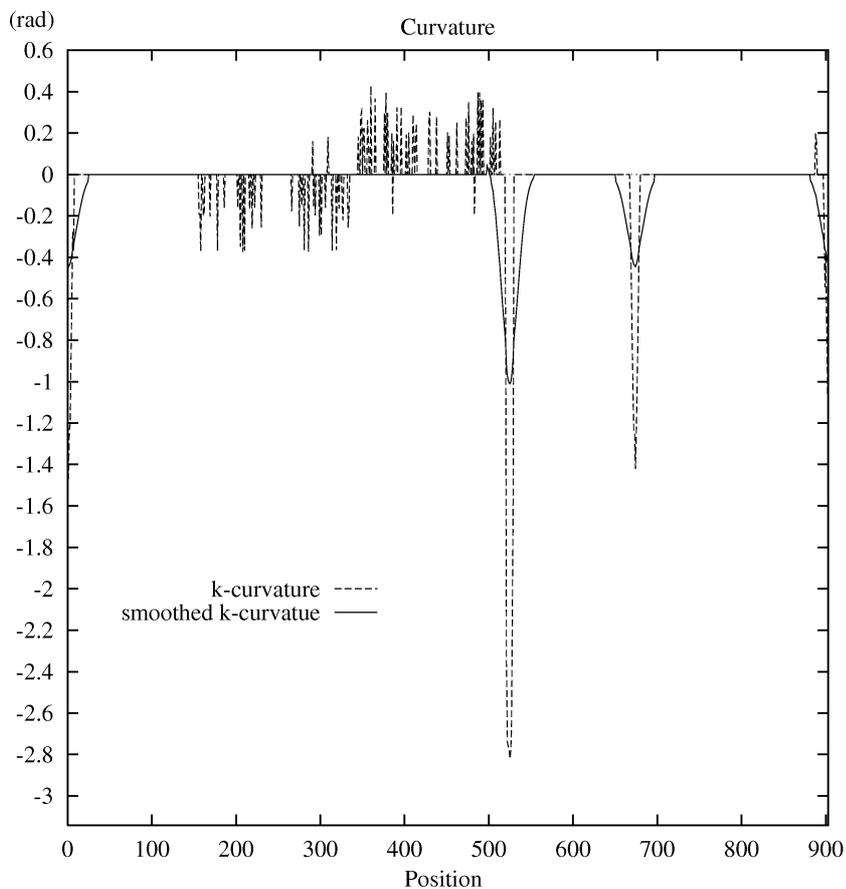


図 2.4: k-曲率

されたそれを示している。直線部分の曲率は0となり，凹や凸の曲率がほぼ一定の曲線部分においても良好な結果が得られているといえる。

また，同図形に対して，本手法にて求めた法線方向を図 2.6 に示す。実線と破線はそれぞれ，更新前の各画素の仮の法線方向，およびあてはめ誤差が最小になる点により更新されたそれを示している。

2.5 特徴点検出

2.5.1 屈曲点の検出

屈曲点では，法線ベクトルの方向が急激に変化する。そこで，注目画素と隣接する画素の法線ベクトルのなす角がしきい値 θ_{th} 以上となる点を屈曲点とする。但し，図 2.7 に示すように曲線同士が接する場合には，更に以下の手順により屈曲点の判定を行う。

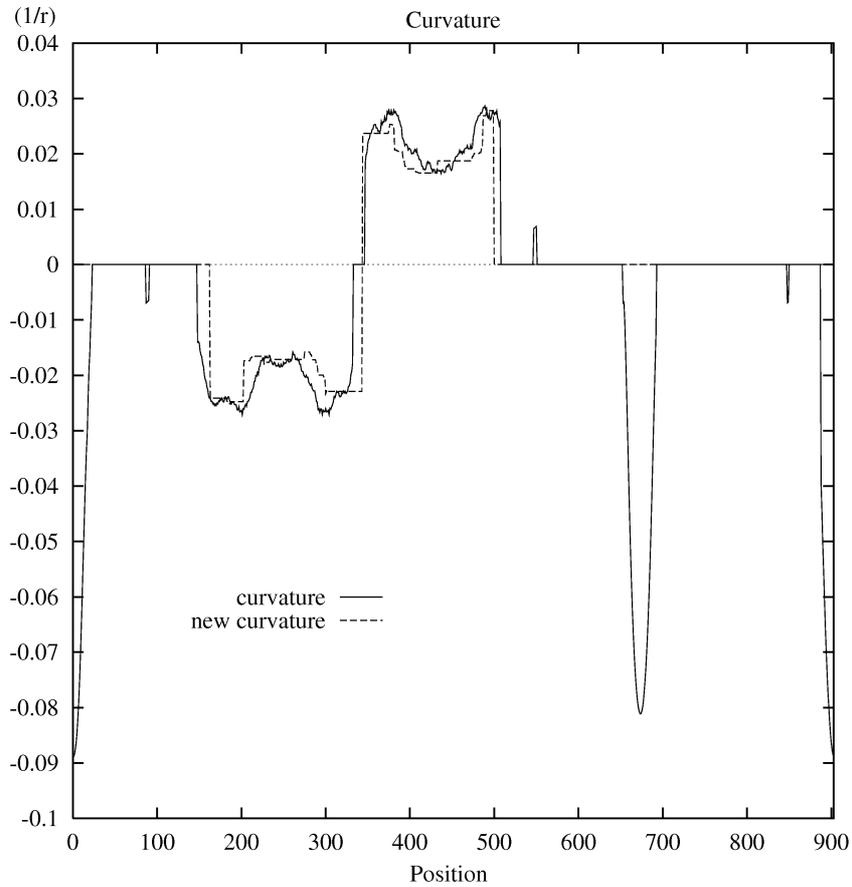


図 2.5: 曲率の更新

- 1) 注目画素 P_i の法線ベクトル \vec{n}_i の算出に用いた円 c_i の曲率から、円周の単位長あたりの中心角の変化量として、 P_i と P_{i+1} 間のユークリッド距離に相当する円 c_i の中心角の変化量 $d\theta_i$ を計算する.
- 2) P_{i+1} の法線ベクトル \vec{n}_{i+1} の算出に用いた円 c_{i+1} の曲率から、円周の単位長あたりの中心角の変化量として、 P_i と P_{i+1} 間のユークリッド距離に相当する円 c_{i+1} の中心角の変化量 $d\theta_{i+1}$ を計算する.
- 3) P_i と P_{i+1} 間の法線方向の変化量を $d\theta$ とするとき、以下の条件を満たす場合、 P_i を屈曲点とする.

$$|d\theta - d| > \theta_{th}, \quad d = \begin{cases} d\theta_i & (d\theta_i \geq d\theta_{i+1}) \\ d\theta_{i+1} & (d\theta_i < d\theta_{i+1}) \end{cases}$$

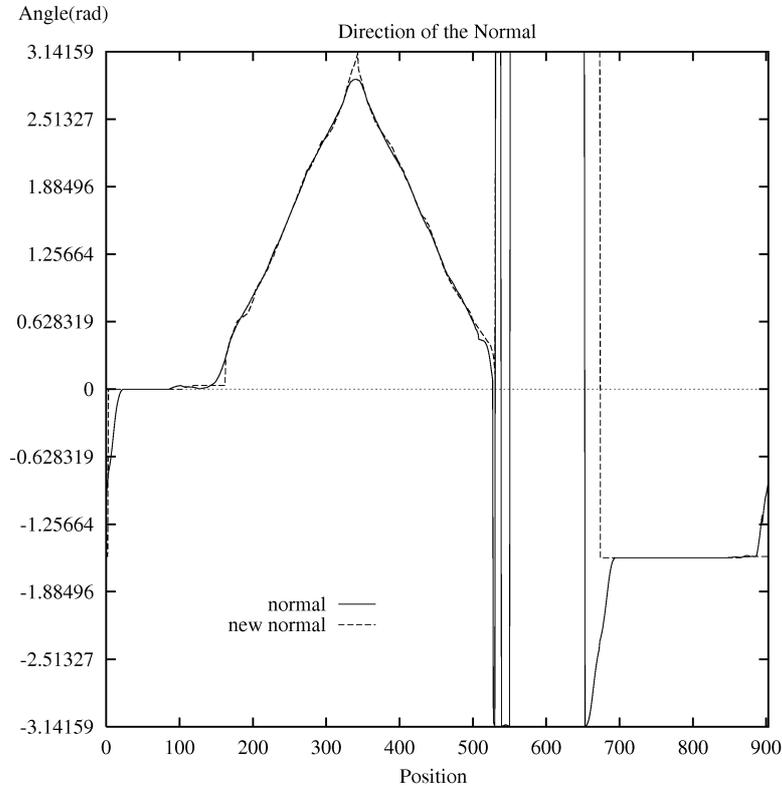


図 2.6: 法線方向の更新

2.5.2 遷移点の検出

遷移点は、曲率が零から非零へ移行する点である。曲率を用いて検出する場合には、対象毎に曲率に関する微妙なしきい値設定が必要となるため、一般的な実用性を考え、対象毎にパラメータ調整を必要としない手法の実現を目指し、曲率を用いず、各点における直線・曲線の判定結果を利用することとした。直線上の点と曲線上の点が滑らかに接続する（角度差が θ_{th} 以下となり、屈曲点とならない）点のみを遷移点として検出する。

2.5.3 変曲点の検出

変曲点は、曲率の符号が異なる2曲線が滑らかに接続する点である。遷移点検出同様、各点における曲線の判定結果の利用に加えて凹凸属性を利用する。そして、凹凸属性の異なる曲線が滑らかに接続する（角度差が θ_{th} 以下となる）点のみを変曲点と

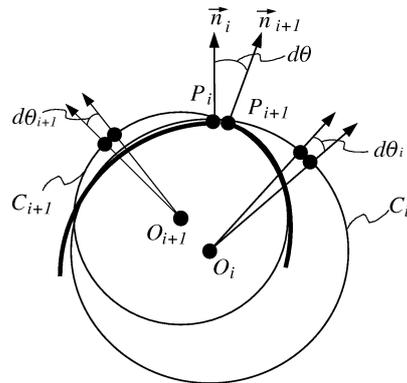


図 2.7: 曲線における屈曲点検出

して検出する. 図 2.8 に特徴点検出結果を示す. □, ×, △ は, それぞれ, 屈曲点, 遷移点, 変曲点を示している.

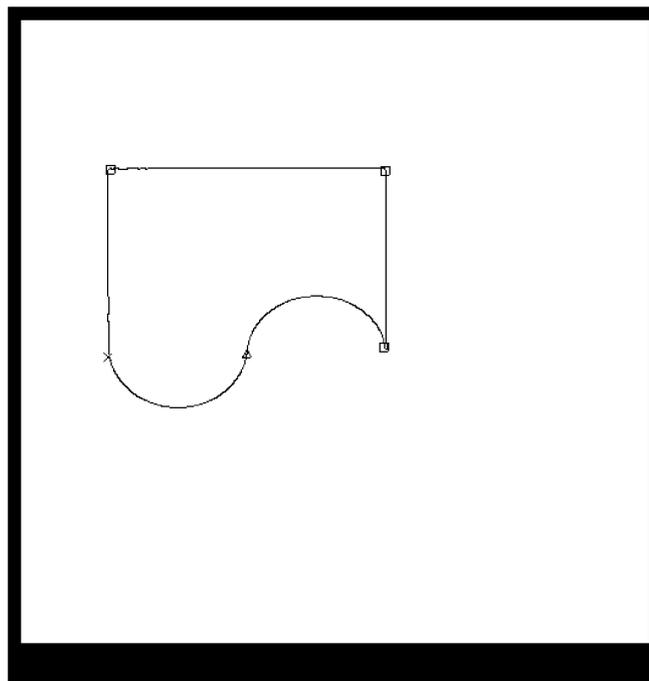


図 2.8: 特徴点の検出結果

2.6 パラメータ k の動的適応

2.6.1 雑音点の検出

あてはめ誤差の更新にもかかわらず、依然として誤差の大きい点は、雑音点として扱う。曲率の変化が緩やかで長く滑らかな曲線や、若干の雑音を含む図形に対しては、 k を大きく設定することによりスムージングの効果が得られ、過分割されず良好な結果が得られる。それ故、パラメータ k を大きく設定するほど雑音の影響を受けにくくなるといえる。しかしその反面、設定されたパラメータ k が、対象とする境界線の形状に比べて大き過ぎる場合には、あてはめ誤差が大きくなり、境界線の微小構造に対応した特徴点検出ができなくなる。このような箇所においては、細部の形状に忠実な特徴点検出ができるよう、パラメータ k を小さくする必要がある。

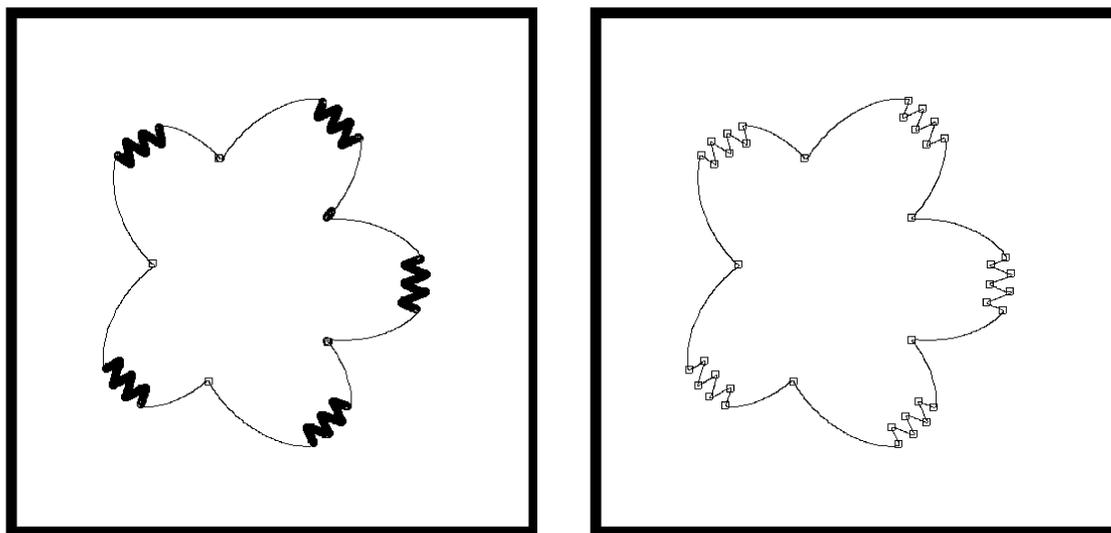
2.6.2 パラメータ k の適応

そこで、雑音点については、距離が k 以下の近傍画素すべてについて、パラメータ k を小さく設定し直し（例えば初期設定の $\frac{1}{2}$ 等）、曲率・法線方向・あてはめ誤差を再計算し、特徴点を再検出する。この処理を、雑音点がなくなるまで繰り返す。図 2.9(a) は、境界線の細かい形状に対して大きめの固定の値 ($k = 20$: 画像の大きさは 512×512) による特徴点検出の結果を示している（実際は、雑音点を \odot で示しているが、表示上太線状に見える）。同図 (b) は、雑音点の k 近傍点についてのみ、パラメータ k を毎回小さく変化させ（初期値の $\frac{1}{4}$ まで）、特徴点を再検出した結果を示している。

2.6.3 検証

境界線は、検出された特徴点により直線と凹もしくは凸の滑らかな曲線セグメントに分割される。形状の認識・記述といった処理においては、セグメント数の増加は、記憶領域や処理時間の増大につながり、できるだけ避けることが望ましい。そこで、以下の処理手順により、遷移点では含まれる一定長未満（抽出するセグメント長の下限值、実際には k を適応させる際の終了条件の値と等しくすればよい）の微小セグメントについては、隣接するセグメントの近似曲線に統合し、統合した部分の曲率・法線方向、特徴点のラベルを更新する。

- 1) 微小セグメントが直線であり、かつ隣接する曲線セグメントの凹凸属性が異なる場合、微小セグメント両端の特徴点ラベルを消去し、セグメントの midpoint に変曲点のラベル付けをする。そして、変曲点の両側の微小セグメント上にあった

(a) $k = 20$ 固定(b) k を 20 から 5 まで適応図 2.9: k の動的適応による特徴点再検出

点については、隣接する曲線セグメントの最近傍点の近似曲線を用いて、曲率ならびに法線方向を更新する。

- 2) 微小セグメントが直線であり、かつ隣接する曲線セグメントの凹凸属性が等しい場合、微小セグメント両端の特徴点ラベルを消去し、セグメントの中点を境に隣接する曲線セグメントの最近傍点の近似曲線を用いて曲率を更新する。
- 3) 微小セグメントが曲線であり、かつ隣接する直線セグメントの法線方向のなす角度がしきい値を越える場合、微小セグメント両端の特徴点ラベルを消去し、セグメントの中点に屈曲点のラベル付けをする。そして、屈曲点の両側の微小セグメント上にあった点については、隣接する曲線セグメントの最近傍点の近似直線を用いて、曲率ならびに法線方向を更新する。
- 4) 微小セグメントが曲線であり、かつ隣接する直線セグメントの法線方向のなす角度がしきい値を越えない場合、微小セグメント両端の特徴点ラベルを消去し、セグメントの中点を境に隣接する直線セグメントの最近傍点の近似直線を用いて、曲率ならびに法線方向を更新する。

また、各点の直線・曲線の判定は、局所近似曲線によるため、セグメントに分割された後に、各セグメントを単位として、大局的にその直線性を検証する。

各セグメントの両端点を通る直線を求め、この直線に対してセグメント上の各点から垂線を下ろし、その足までの各点からのユークリッド距離を計算し、その最大値が直線性を判定するためのしきい値 d_{th} を越えなければ直線、それ以外は曲線と判定す

る。次に、セグメントの直線・曲線の属性が、この判定と異なる場合は、判定結果に応じてセグメント両端点と中点を用い、直線もしくは円弧近似を行い、曲率ならびに法線方向を再計算し、セグメント上の各点の属性をと合わせて更新すると共に、特徴点のラベルも更新する。

2.7 実験

本手法を実施するにあたり、あらかじめ設定すべきパラメータとして、屈曲点とみなす角度差 (0.3 radian), 雑音点検出のためのあてはめ誤差判定値 (正方格子の一边強に相当する, 1.1 pixel), k の初期値 (20 pixel), 分割すべき区間長の最小値 (5 pixel) を用いている。これらを設定するだけで (システム実装上は, 対話的に可変とすることも可能である), その他のパラメータや対象毎のしきい値調整は基本的に必要としない。 k に関しても, 先に述べた動的適応により, 対象画像毎に異なる設定をする必要はない。

図 2.10 に, 従来の屈曲点検出手法に基づく直線のみからなる多角形近似結果 (同図 (a)) と, 本手法による特徴点抽出に基づく直線・曲線抽出結果 (同図 (b)) の比較を示す。TVカメラにて撮像した画像 (512 × 512) に対して, 本手法を適用した実験結果であり, 図中の□, ×, △はそれぞれ, 屈曲点, 遷移点, 変曲点を示す。

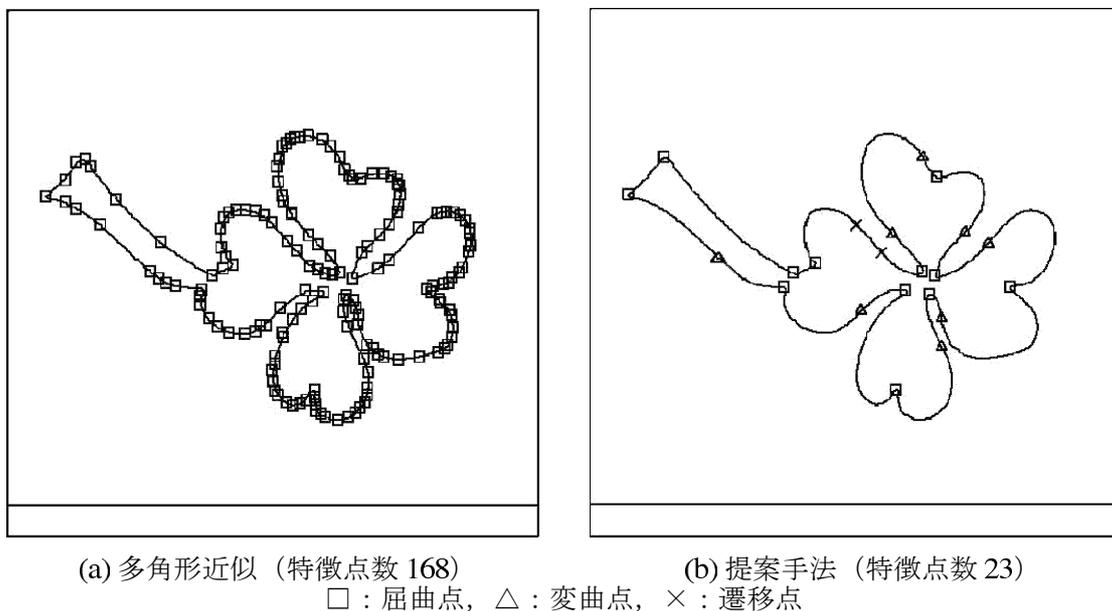


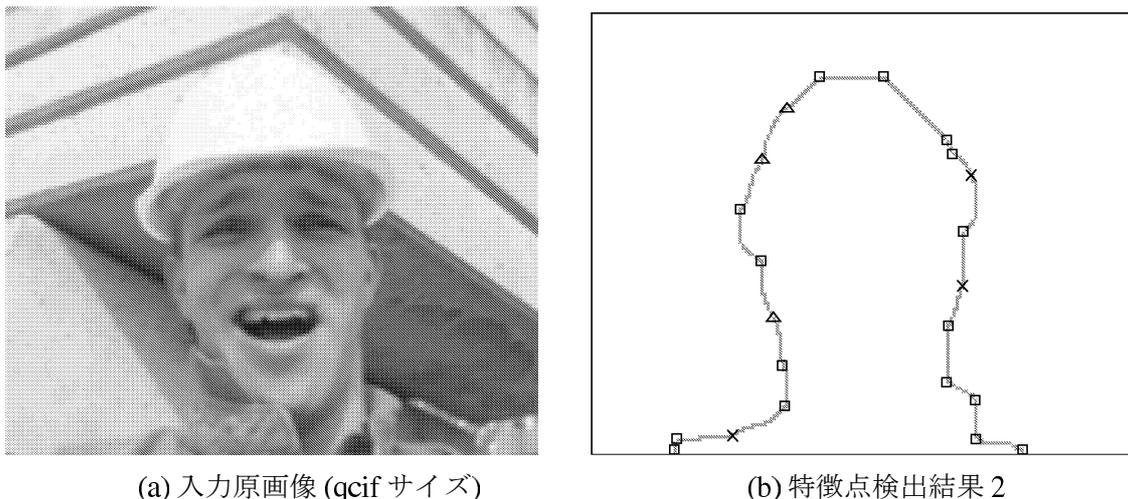
図 2.10: 特徴点抽出結果 1

図 2.10 の結果比較から, 本手法により検出された特徴点に基づいて, 境界線を直線セグメントと凹もしくは凸の曲線セグメントに分割することで, 少ないセグメント数

にて効率的なセグメント化がはかれていることがわかる。

また、図2.11(b)は、同図(a)の ITUT 標準画像 (qcif サイズ) を濃淡画像に変換した原画像から獲得した人物領域の境界線に対して、本手法を適用した実験結果である。

この実験結果からも、入力画像サイズや対象の大きさが異なる場合にも、本節冒頭に記載のパラメータのみを用い (しかもそれらに対話的に調整することなく)、屈曲点のみならず、変曲点、遷移点も検出できていることがわかる。



(a) 入力原画像 (qcif サイズ)

(b) 特徴点検出結果 2

図 2.11: 入力原画像および特徴点検出結果 2

2.8 第2章のまとめ

境界線上の特徴点検出手法について述べた。屈曲点のみならず、変曲点、遷移点といった特徴点を検出するためには、境界線上の各点における曲率、法線ベクトルといった局所特徴を精度良く求める必要がある。そこでまず、注目画素の近傍画素を用いた近似曲線により各点の局所特徴を求める。次に、近傍画素間での近似誤差比較により最適近似区間を決定し、局所特徴を更新する。続いて近似誤差の大きい点に対して、近傍画素を決定するパラメータ k を小さく設定し直し、局所近似により局所特徴と近似誤差を再計算する。この最適近似区間の決定と、パラメータ k の動的適応により、屈曲点のみならず、凹凸の反転する変曲点や直線から滑らかに曲線に移行する遷移点をも検出することが可能となった。

また、本手法は、検出に必要なしきい値を対象毎に調整する必要がないばかりでなく、境界線形状の微小構造を自動的に検出し、その近傍部分に対してのみパラメータ k を動的に適応させ、曲率・法線方向を再計算することにより、細かい形状の部分についても特徴点を検出することができる (一種の粗密解析による動的適応ということ

もできる)。これにより、適応の度合を変えることで、階層的な符合化や記述の生成も可能となる。例えば、動画像を符合化する際、微小構造部分に関しては、出現したフレームでのみデータを符合化し、連続して出現する以降のフレームでは、微小構造のラベルと対応する前フレームにおける微小構造を指すポインタといった形で符合化できる。

また、計測によりデジタル化された3次元点群からなる物体形状の3次元モデル化等への応用を念頭に、3次元拡張も行った [41]。

しかしながら、本手法にて検出される特徴点は、第2.6.3節で述べた検証の操作を加えない場合にも、同一画像の縮小といったスケール変化に対する不変性が保証されるものではない。3次元空間中の平面図形を任意の方向から観測した場合を考えると、変曲点や遷移点は、数画素のずれが発生する可能性があり不変的なものとは言い難い。屈曲点に関しても、画像の量子化の影響を受け、真（サブピクセルレベル）の屈曲点位置の近傍の8連結で接続される2画素に等しい特徴量をもたらすようにデジタル化される（縮小の際にはサンプリングとなる）ことも多く、セグメントの端点1画素については、そのずれについて誤差を許容すべきであり、セグメント化した後に、セグメント画像上でのサブピクセル化による補正を行うこととしている。

本手法における境界線上の特徴点は、境界線（線図形）をより少ないセグメントに分割することを目的として検出されており、形状記述の効率の向上にも効果的であるばかりでなく、次章以降で述べるエッジベースステレオにおける対応探索や再構成、物体認識におけるモデルとの照合時の探索空間の削減、共通モデルにおける形状プリミティブ抽出の前処理として、非常に重要な役割を果たしている。

第3章 エッジベースのステレオによる平面からなる多面体の三次元再構成

3.1 具体的課題と関連研究ならびに概要

ステレオ（両眼立体視）は、シーンの距離情報を三角測量の原理で得る基本的な方法である。人間の眼と同様に非接触かつ受動的に距離測定できる技術であり、将来の知能ロボットに共通的な眼として、汎用性と融通性の観点から期待され、これまでステレオの対応問題を解決する多くのアルゴリズムが提案されてきた。対象とするシーンの画像特徴に着目し、特徴点 [42]、エッジ [43]～[45]、セグメント [46]、面 [47] を単位とした対応法が提案されている。対応探索のアプローチとしては、粗密解析法 [44][48][50]～[52]、弛緩法 [49]、動的計画法 [53][57][58]、構造解析法 [46][54][63] に大別することができる。また、異種類の画像特徴を併用して対応探索を行う手法 [55]～[57] や、多眼化により誤対応除去と測定精度向上をはかる手法 [59][60] も提案されている。しかし、特徴の多種類化と共に、対応探索空間は指数的に増大してしまうため、並列処理機構 [61] を用いた協調的統合に関する手法も提案されている [62]。また、物体の形状認識への利用を考えた場合、精度ならびに出力形式も重要となり、ステレオ画像の対応の単位として、画像の境界表現から得られる領域の境界線を構成するセグメントを利用する方法が提案されている [63]。点単位の方法と比較して、画像全体のデータ量が大幅に圧縮されかつ個々のセグメントの持つ情報量が増大することと、あいまいな対応を排除することができるため対応の探索空間が小さく、正しい対応を容易に得ることができることから、本研究ではセグメントベースのアプローチを採用した。また、領域の特徴に加え、セグメントの連続性を対応の強力な拘束として利用することとした。

しかし、ステレオにはさらに、テクスチャーのある面やシェーディングのある曲面を復元することはできるが [56][64]、多くの人工物がそうであるように、一様な平面で構成された物体に対しては、その境界線を検出できても特徴のない面の内部を直接対応によって復元することはできない問題がある。これは、未知環境を自律的に移動するロボットにとっては好ましくない。すなわち、障害物の境界線を検出できても、境界線のどちら側に面があるのか、つまり、境界線のどちら側を通過できるのわからないので立ち往生してしまう。また、ハンドアイロボットにとっても、同じ理由で、物体のどの部分を掴むことができるのかわからない。従って、境界線の復元の後、面

を復元する処理が必要となる。しかし、物体の輪郭線（奥行きの不連続部分）と稜線（面の法線方向の不連続部分）を区別しない単純な面内挿アルゴリズムでは、偽の面を生成してしまう。Waltzの線画のラベル付けの方法[65]も、面を推定する一つの方法であると考えられるが、物体が3面頂点の物体に限定されるほか、本来完全な線画を仮定したものであるため、実データを用いた場合には、要である頂点タイプの決定が容易でないし、線画の一部の欠損が他のすべての部分の解釈を不可能にしてしまう。Herman[66]らは、ステレオ航空写真から市街のビルディングのワイヤフレームを生成した後、面を生成する方法を開発したが、ビルディングは3面頂点物体で水平面上の線と垂直面上の垂直線で構成されているという特殊な条件を利用したものである。また、Bras-Mehlman[67]らは、ステレオによって得られる3次元エッジセグメントから3次元空間のドロネー三角網（4面体パッチ）を構成して面を生成する方法を開発している。しかし、計算コストが非常に高く、また偽の面も生成される。富田ら[63]は、ステレオ対応問題に対して、平面パターンのアフィン変換可能性を拘束条件として利用する方法を述べているが、厳密な面の復元問題については論じていない。また、対象とする物体面上の点群から三角パッチによるサーフェスモデルを獲得する手法[68]も提案されている。しかし、網羅的探索による一貫性のある解釈を生成するための計算コストは大きく、また、物体上の点以外が混入した際には解釈も難しくなる。

そこで、本章では、この面の復元問題を解決するために、ステレオ画像の境界表現の対応から得られる物体の3次元境界線の構造から、一意的に存在可能な面を決定する方法について紹介する。本手法は、平面で構成された物体すべてに対して有効であり、面上に模様や影があっても構わない。また、欠損があるために面を一意的に決定できない境界線を局所的に区別することができ、それを補償するように後の観測に委ねることが可能となる。そして、結果として、物体/環境の境界表現（立体を表現する最も一般的な幾何モデルで、通称**b-rep**と呼ばれる）が得られる。この出力形式は、前処理をすることなく、そのまま物体認識システム[69]の入力として利用することができる。

3.2 ステレオカメラ構成と入力画像

3.2.1 輻輳カメラと標準カメラモデル

カメラの輻輳運動により、注視点 F の距離を変えることができる。カメラ間の基線長を長くして距離測定精度を高めるため、撮像環境として輻輳カメラモデルを用いている（図3.1）。

しかしながら、輻輳カメラモデルにおいては、カメラのレンズ中心を結ぶ線分 $\overline{O_L O_R}$ と注視点 F により張られる平面と画像面との交線（これをエピポーラ線と呼ぶ）は、

左右各画像面の水平走査線と平行にはならない．左画像上の点に対応する右画像上の点は，このエピポーラ線上で求めることとなり，これを「エピポーラ拘束」と呼ぶ．右画像上の点に対する対応点を左画像中より求める際も同様にこのエピポーラ拘束に従う．ここで，ステレオ画像の扱いを容易にするために，撮像系として，図 3.2 に示すように，カメラの光軸が平行である標準カメラモデル [63] を用いる．

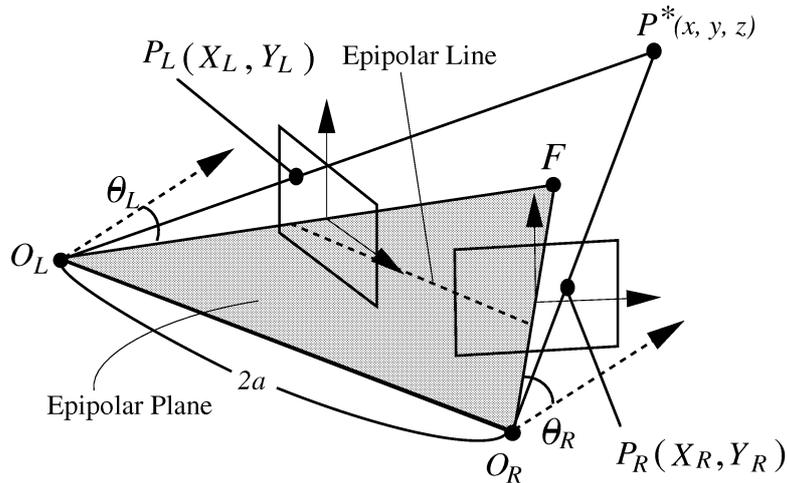


図 3.1: 輻輳カメラモデル

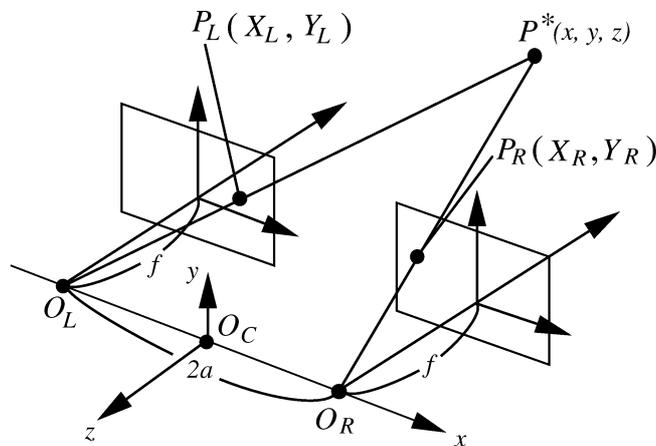


図 3.2: 標準カメラモデル

カメラのレンズ中心 O_L , O_R の中点 O_C を原点とし， $\overrightarrow{O_LO_R}$ を x 軸，光軸方向を z 軸とし， x 軸， z 軸に直交する方向を y 軸とする．焦点距離 f ($=1$) とし，画像平面が基線に平行にレンズ中心から f の距離にあるカメラ構成を考える．3次元空間中の点 $P^*(x, y, z)$ の左右画像における像を $P_L(X_L, Y_L)$, $P_R(X_R, Y_R)$ ， $\overline{O_LO_C} = \overline{O_CO_R} = a$ と

すると、視差 $d = X_L - X_R$ から、点 P^* の距離 z は次式により得られる.

$$z = \frac{2af}{d} \quad (3.1)$$

次式の幾何学的な拘束に基づいて左右の画像で対応点を求める.

$$X_L > X_R \quad (3.2)$$

$$Y_L = Y_R \quad (3.3)$$

このように、標準カメラモデルでは、エピポーラ線が画像面の x 軸と平行となるため、左画像上の点に対応する右画像上の点は、常に同一走査線の半直線上で求めればよいこととなり探索処理を簡便化できる. 右画像上の点に対する対応点を左画像中より求める際も同様である. カメラの輻輳角を θ とした場合、輻輳カメラの画像上の一点の座標 (X, Y) は、次式 (3.4), (3.5) によって標準カメラ画像上の座標 (X', Y') に変換することができる (ステレオ対応探索処理においては、原画像そのものを補正するのではなく、画像より獲得した境界表現 (次節 3.3.1 参照) に対して補正を行っている (図 3.3(b) 参照)).

$$X' = \frac{\sin \theta + X \cos \theta}{\cos \theta - X \sin \theta} \quad (3.4)$$

$$Y' = \frac{Y}{\cos \theta - X \sin \theta} \quad (3.5)$$

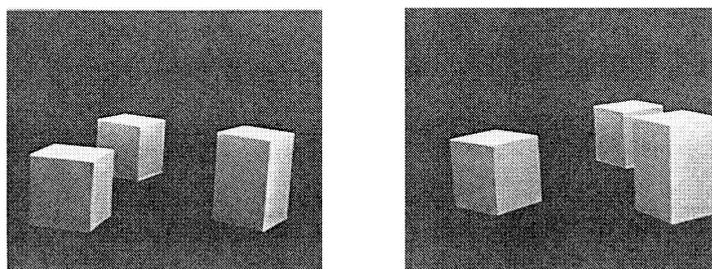
図 3.3(a) に、実験に用いた入力画像例を示す. この例において、基線長は約 30cm であり、対象物と各カメラとの距離は、100cm 程度である. 焦点距離 25mm のレンズを使用し、対象物が視野の中央に納まるように、輻輳角をつけて撮像している.

カメラ座標系における点 $P(x_c, y_c, z_c)$ は、次式 (3.6) で表される.

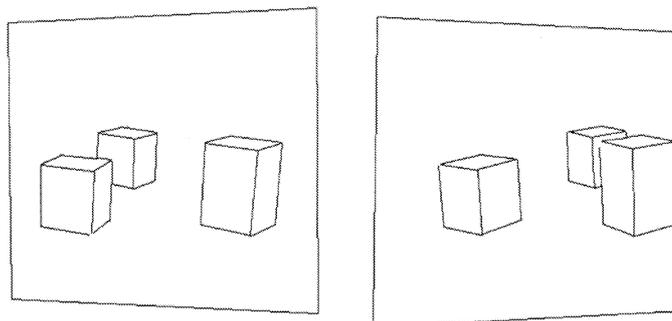
$$\begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} = \frac{2a}{X_L - X_R} \begin{pmatrix} \frac{X_L + X_R}{2} \\ Y_L \\ 1 \end{pmatrix} \quad (3.6)$$

3.2.2 カメラキャリブレーション

ステレオ対応が得られれば、対応点の視差情報を得ることができる. ステレオビジョンによる観測データから対象物体を 3 次元的に再構成し、立体幾何モデルを生成するためには構造に関する情報まで獲得する必要がある. ここで、視差分布図の生成に留まることなく距離測定を行うためには、前述の標準カメラモデルへの変換も含め、カメラパラメータを正しく推定しなければならず、これまで、カメラキャリブレーションに関する多くの研究がなされてきた [70]~[77].



(a) ステレオ原画像



(b) 輻輳補正後の境界線セグメント画像

図 3.3: ステレオ原画像と輻輳補正後の境界線セグメント画像

カメラパラメータとして、カメラ座標系の原点（レンズ中心）の世界座標系における位置（3自由度）、カメラ座標系の方向（3自由度）といったカメラの外部パラメータ（計6自由度）と、焦点距離（1自由度）及び撮像面から画像バッファへのアフィン変換（4自由度）からなるカメラ内部パラメータの計11自由度のパラメータを決定しなければならない。このカメラキャリブレーションには、3次元位置が既知の参照点を多数観測させ、その3次元座標と投影像の2次元座標との対応関係から計算する古典的なアプローチがある [78]～[80]。また、単一カメラのキャリブレーションとして、カメラ内部パラメータを求める手法 [81]～[83] や、2台のカメラに特化し、高精度な結果を得ようとする手法 [84]、ステレオカメラ対のエピポラ幾何条件のみを求める校正法 [85]～[87] も提案されている。

参照点の3次元座標を正確に設定できる場合、カメラパラメータを高精度かつ安定に推定できる。カメラが複数台ある場合にも、それらに同一の参照点を見せて個別にキャリブレーションを行うことにより、全カメラの位置と姿勢を共通の座標系上で求めることができる（カメラの外部パラメータは参照点の位置を記述した座標系上で得られるため）。ステレオビジョンでは、複数のカメラ間の相対位置関係が重要となるため、上述した古典的なアプローチにより、カメラキャリブレーションを行った。図 3.4 に、利用したパターンを示す。同図 (a) は、マーカを装着したマニピュレータであ

る。ビジョンベーストマニピュレーションを行わせる際には、ビジョンの座標系とロボット座標系を統合する必要があるため、マニピュレータ上のマーカを複数位置に移動させ、それらをステレオカメラにより撮影することにより、キャリブレーションを行う。同図(b)(c)は参照点を多数設定したものの、(d)は立方体ブロックで、画像より各頂点を検出し参照点として用いる。

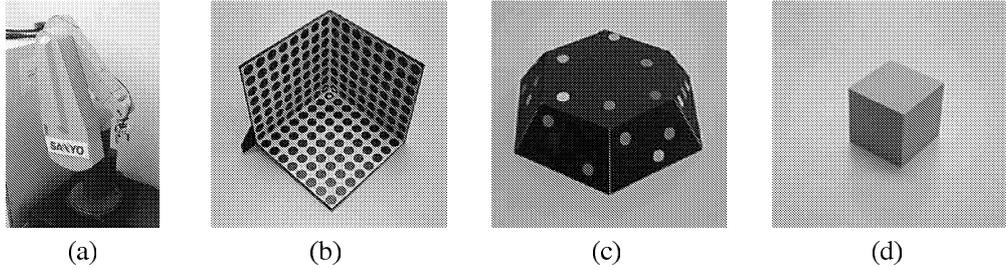


図 3.4: キャリブレーションパターン

世界座標系の1点 $P(x, y, z)$ のカメラ座標系での位置 (X, Y) への変換は、次式 3.7 のような1次分数式で表される。

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \frac{1}{h_{31}x + h_{32}y + h_{33}z + 1} \times \begin{pmatrix} h_{11}x + h_{12}y + h_{13}z + h_{14} \\ h_{21}x + h_{22}y + h_{23}z + h_{24} \end{pmatrix} \quad (3.7)$$

実際のカメラモデルでは、ピンホールカメラモデルにあてはまらないレンズ等による非線形な光学歪みが生じるため、別途補正を行う必要がある [88]~[92]。また、各カメラの光学的特性のばらつきも吸収する必要がある。同じ光の強度に対して同じ輝度が得られるような輝度の補正である。

点 $P(x_c, y_c, z_c)$ の世界座標系の標記 (x_w, y_w, z_w) への変換式は、 R を回転移動成分、 t を並進移動成分とすると、次式 (3.8) で表される。

$$\begin{pmatrix} x_w \\ y_w \\ z_w \end{pmatrix} = R \times \begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} + t \quad (3.8)$$

カメラキャリブレーションの結果を用いれば、画像の座標 (col, row) との関係は、次式 (3.9) で表される。

$$\begin{pmatrix} col \\ row \end{pmatrix} = \frac{1}{h_{31}x_w + h_{32}y_w + h_{33}z_w + h_{34}} \times \begin{pmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \end{pmatrix} \begin{pmatrix} x_w \\ y_w \\ z_w \\ 1 \end{pmatrix} \quad (3.9)$$

(3.6), (3.8), (3.9) 式より, 画像座標 (col, row) から標準カメラモデル座標系 (X, Y) への変換は, 次式 (3.10) で与えられる.

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \frac{1}{m_{31}col + m_{32}row + m_{33}} \times \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \end{pmatrix} \begin{pmatrix} col \\ row \\ 1 \end{pmatrix} \quad (3.10)$$

3.3 画像の境界表現と対応探索

3.3.1 画像の境界表現 (データ構造)

画像の中間表現として, 立体の (3次元) 境界表現の仕様に従う画像の (2次元) 境界表現を用いる. 画像の境界表現は, 画像を再構成できる程度に情報を圧縮する画像の中間表現である. 画像の境界表現を得るための処理の概略を以下に述べる.

1) エッジ検出と境界線の抽出 :

まず, 一次微分オペレータにより画像の各点についてエッジ強度と勾配方向を求め, エッジの細線化—しきい値—延長処理により, 領域を囲む閉じた一画素幅のエッジを得る.

2) 境界線上の特徴点検出によるセグメント化 :

次に, 領域を右手に見る方向にエッジを追跡し, 領域の境界線を求める. したがって, 外周境界は右回り, 内部の領域 (穴) に対する内周境界は左回りとなる. 2章で述べた手法により, この境界線上の特徴点として, 分岐点, 屈曲点, 変曲点, 遷移点の検出を行い, これら特徴点にてセグメントに分割する.

3) セグメントの分類 :

分割されたセグメントは, 直線セグメントと曲線セグメントに分類され, 更に曲線セグメントは, 凹凸属性も合わせて保持し, 凹曲線もしくは凸曲線に分類される.

画像の境界表現では, 画像は明るさが一様な領域の集合で表される. 図 3.5 に示すように, 領域—境界線—セグメント—境界点の四階層のデータ構造によって表現される. 各領域 (**R**) は, 一つの外周境界 (**B**) と, いくつかの穴が存在する場合には, 穴との内周境界 (**H**) のリストによって表される.

各境界は, それを構成する境界セグメント (**S**) のリストによって表される. 各セグメント (実線) は, 右の領域の境界であることを示す方向を持ち, 更に, セグメント上のエッジ点 (**P**) に対するポインタも合わせ持つ. また, 各セグメントは, 図 3.6 に示すように同一エッジ上に存在する反対の向き (隣接領域の境界を表す) セグメ

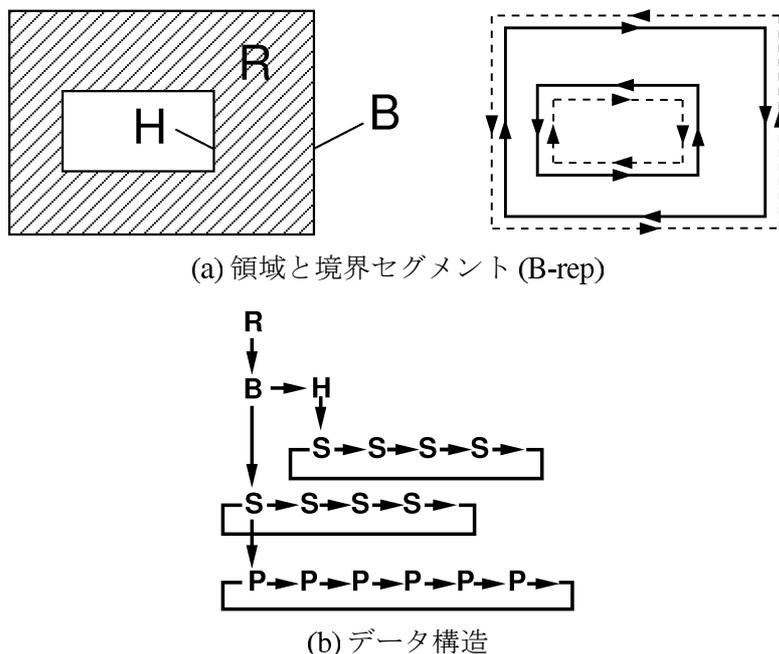


図 3.5: 画像の境界表現

ント（点線）へのポインタをもつことによって、いわゆる翼状エッジ（winged edge）を構成している（同図中、各セグメントはエッジ上に乗っているが、説明上離して表示している）。各境界点は、法線方向や領域の輝度情報も保持している。画像の境界表現は、このようなトポロジカルなグラフ構造をなしており、エッジベースステレオにおいて、有益な情報を効率良く蓄積することができる。

ステレオ画像の対応探索はこのデータ構造上で行なわれる [63]。例として、図 3.3(b) に、同図 (a) のステレオ画像から生成した画像の境界表現のセグメントを示す。

3.3.2 ステレオ対応探索

ステレオの対応単位としてセグメントを用い、対応探索の際には、エピソード拘束のほかに、領域の明るさ、セグメントの方向、他のセグメントとの連結性を利用する [93]。以下に、この概要を述べる。

- 1) まず、領域間で対応の候補を求める。対応の必要条件は以下の通りである。
 - 領域の明るさが類似している。
 - 領域内にエピソード拘束を満足する対応部分が存在する。
- 2) 次に、対応する領域の境界線を構成するセグメント間で対応の候補を求める。対応の必要な条件は以下の通りである。

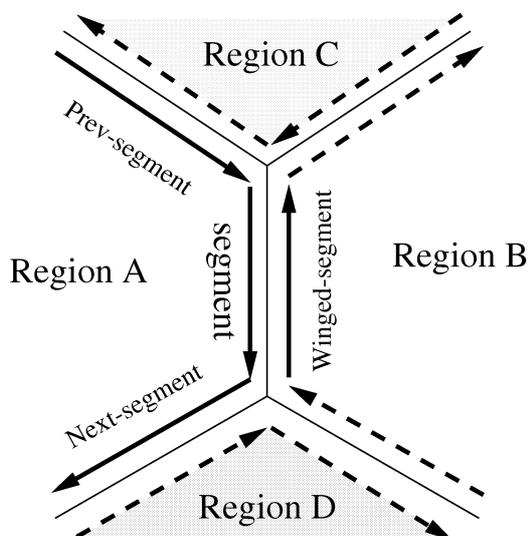


図 3.6: Winged-edge 構造

- セグメントの方向が類似している.
 - セグメント上にエピポーラ拘束を満足する対応部分が存在する.
- 3) 一つのセグメントが同一のエピポーラ線上で複数のセグメントと多重対応する場合には, 対応するセグメントの空間的な連続性を調べ, その結果得られるセグメントの系列 (ワイヤフレーム) のなかで, 最も長くかつ形状が最も類似する系列を選択する. 特に, 以下の平面拘束定理に基き, 平面を構成するアフィン変換可能なセグメントの系列を優先する.

[平面拘束定理]

3次元空間中の任意の平面図形について, 標準カメラモデルにおけるステレオ画像上の一方への投影像と, 他方のそれとは互いにアフィン変換可能である [94] (詳細については, 付録 B を参照されたし).

そして, 対応する一対のセグメントにより, 一つの 3次元セグメントが生成される. 3.4 に, この対応後の処理について述べる.

3.3.3 多重対応の除去

エピポーラ条件を満足した対応候補セグメント対について, 対応部分の始点, 終点位置を求める. 対応部分の特徴から, 明らかに対応しない候補を, 以下の特徴の相違度に基づき判定する.

輝度：対応部分の始点と終点の中間点における輝度値

方向：対応部分の始点と終点を結んだ弦の方向

形状：対応部分の直線，曲線の属性

(曲線に関しては，凹凸属性の異なる組合せを対応候補より削除する)

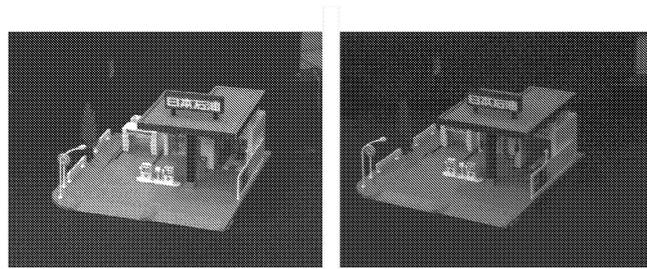
また，ステレオカメラペア以外に，検証用のカメラが利用可能な場合には，検証用カメラ画像を用いて，対応候補の更なる絞込みを行うことができる [95][96]．これは，正しい対応であれば，その対応により求まる3次元空間の点を，検証用カメラ画像上へ逆投影した際の位置に，同じ性質を有する点が存在しなければならないという拘束条件を利用する．以下に，検証方法の概略を述べる．

- 1) 検証用カメラから生成された境界表現の点列を画像面上にプロットする
- 2) ステレオ対応候補における対応部分において，式 (3.6) より，3次元空間座標を求める
- 3) 得られた3次元空間座標を，式 (3.9) にて検証用画像へ逆投影する
- 4) 検証用画像において，対応部分の midpoint の逆投影点座標近傍（上下左右）に境界点（検証点）が存在するか調べる
- 5) 探索範囲に境界点が存在しない対応候補は，偽の対応であると考えられるため対応候補より削除する

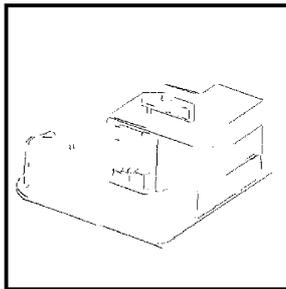
図 3.7 に，同図 (a) のステレオ画像より対応の得られたセグメントを3次元再構成し，複数の視点より眺めた図を同図 (b)~(d) に示す．

3.4 ステレオデータ

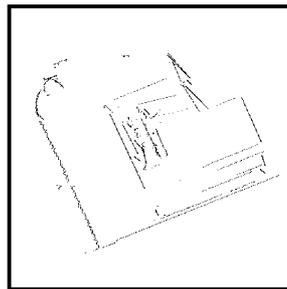
ステレオの直接の出力は，左若しくは右画像の領域の（外周および内周）境界を構成していた連結セグメントの系列を3次元空間に逆射影して得られる3次元セグメントの系列（以下，ループと呼ぶ）の集合である．この際，一方の画像で不連続であっても，他方の画像で連結しているセグメントは，一つのループを形成する．また，ループには，画像の一つの領域の外周境界と内周境界に相当するBループとHループがあり，それらは結合されて一組のループとして扱われる．例として，図 3.8(a) の3次元ブロックを考える．ブロックの境界線として，輪郭線 (→) と凸稜線 (+) が観測される．ステレオは，このブロックを同図 (b) の四つのBループによって表現する．ルー



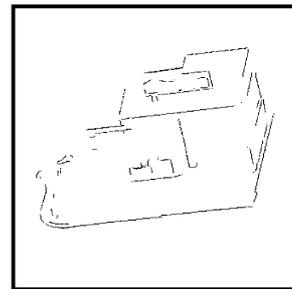
(a) ステレオ原画像



(b)



(c)

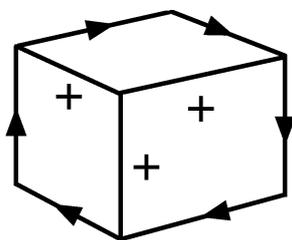


(d)

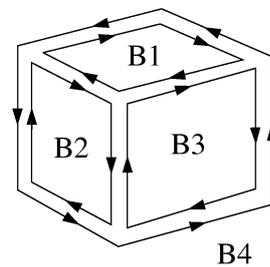
図 3.7: 3次元セグメント 1

プ上の各セグメントは、右の面の境界であることを示す方向を持ち、また、同一境界線上に存在する反対向きの（隣接面の境界を表す）セグメントへのポインタを持つことによって、翼状エッジ（図 3.6 参照）を構成している。

ここで、物体の輪郭線上には、実在する面の境界を表す「実セグメント」と、図 3.8(b)のループ B 4 上のセグメントのように、実在しない（現在の視点からは見えない）面の境界を表す「虚セグメント」が共存する。虚セグメントは、偽の面を作る原因となる。従って、正しい面の復元には、物体の輪郭線と稜線を区別し、輪郭線から虚セグメントを除去することが課題となる。



(a) 輪郭線と凸稜線



(b) 境界セグメントのループ

図 3.8: ブロック

3.5 面の復元

3.5.1 オクルージョンの検出

ステレオにより，左右両画像で観測できるエッジの距離を測定することができる．対応の求まったエッジ部分に関しては，3次元セグメントとして3次元情報が得られる(図3.7参照)．図3.9に，図3.3のステレオデータより得られた3次元セグメントを画像面に逆投影したものを示す．

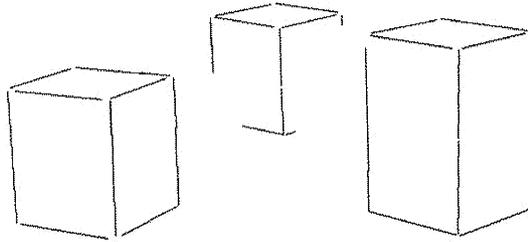


図 3.9: 3次元セグメント 2

しかし，3次元世界では，手前（視点側）にある面が後方の面を隠すこと（オクルージョン）により，一方の画像中では見えていても他方の画像中では見えないエッジが少なからず存在する．

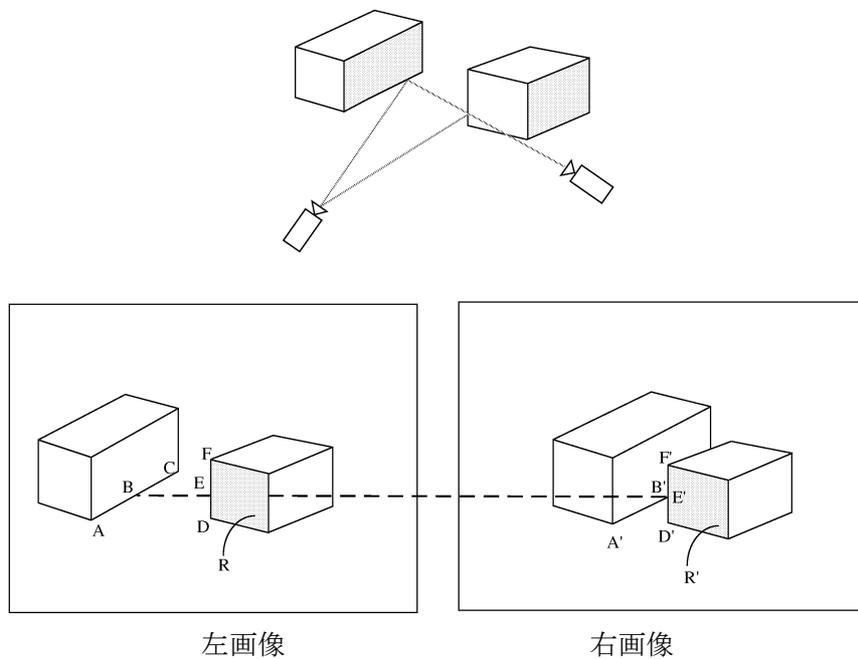


図 3.10: オクルージョン

オクルージョンが生じると，図3.10のセグメント(AC, A'B')のように（以下、左画像の特徴を示す記号に対応する右画像の記号には'を付ける），対応の不完全なセグメント対が出現する．すなわち，左画像のセグメントのABの部分，右画像のセグメントのA'B'の部分に対応しているが，左画像の領域RがセグメントのBCの部分より手前にあると，右画像では領域Rに対応する領域R'によってBCに対応する部分が隠蔽されて見えなくなる．図3.9の結果にも，オクルージョンにより一方の画像中で観測されずに不完全な対応となる部分が見られる．

オクルージョンが生じている個所では，図3.11のように，D'E' (E'F')を「明らかな実セグメント」（太線）に分類でき，D'E' (E'F')と翼状エッジを構成する逆向きのセグメントF'E' (E'D')を「明らかな虚セグメント」（点線）に分類できる．

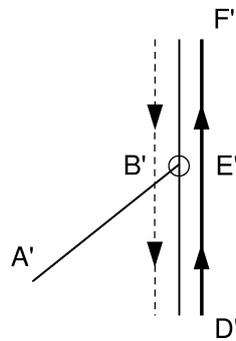


図3.11: 明らかな実セグメント（太線）と明らかな虚セグメント（点線）

このように，オクルージョンは，面の復元に重要な手掛かりを与える．そこで，以下に右画像でのオクルージョン検出の手順を述べるが，左画像の場合も同様である．

- 1) 対応の不完全なすべてのセグメント対を選ぶ．いま，左画像のセグメントACが，右画像のセグメントA'B'とABの部分で対応するが，BCの部分に対応するセグメントが存在しないとする．
- 2) セグメントA'B'の端点B'を通り，A'B'を左側に見る他のセグメントD'E' (E'F')を探す．但し，画像上では，B'とE'は同じ座標である．
- 3) DE (EF) がD'E' (E'F')に対応しているとする．
- 4) ここで，BとB'間の視差を $d(B, B')$ 、EとE'の間の視差を $d(E, E')$ とする．もし $d(B, B') < d(E, E')$ ならば，セグメントのBCに対応する部分は，右画像ではB'からセグメントD'E' (E'F')を境界とする面によって隠蔽されていると判定する．

図3.12では，図3.3のシーンで検出されたオクルージョンの生じている点（E'に相

当する) を○印で示している. そして, これらの点のまわりでは, 明らかな実セグメントと明らかな虚セグメントが決定される.

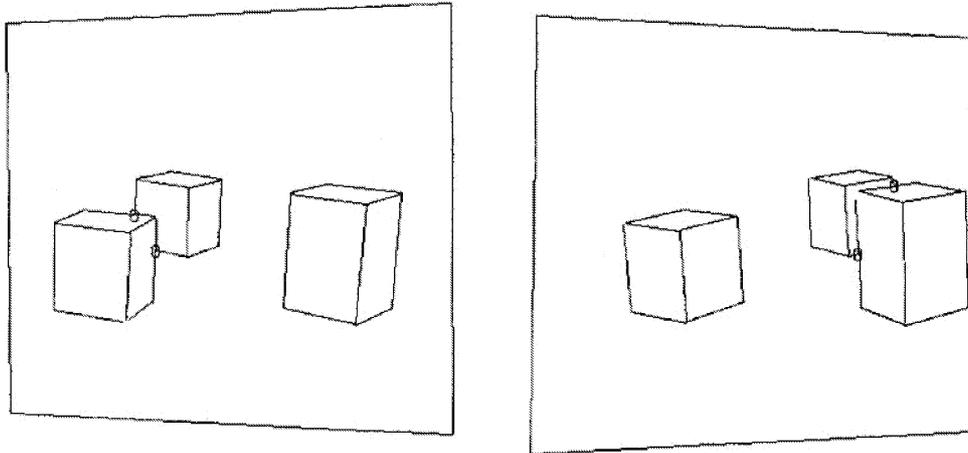


図 3.12: オクルージョン点 (○) の検出

3.5.2 実在テスト

次に, 図 3.7 および図 3.9 の結果に示した 3 次元空間中の針金細工のように粗く得られた 3 次元情報とそれらの接続関係に基づいたループを用いて, 各ループに面を張ると同時に, ループ上でまだ分類されていない他のセグメントを, 実セグメントまたは虚セグメントに分類する実在テストについて述べる. 一つのループが「解釈可能」である, すなわち, そのループに面を張ることができる条件は以下の通りである.

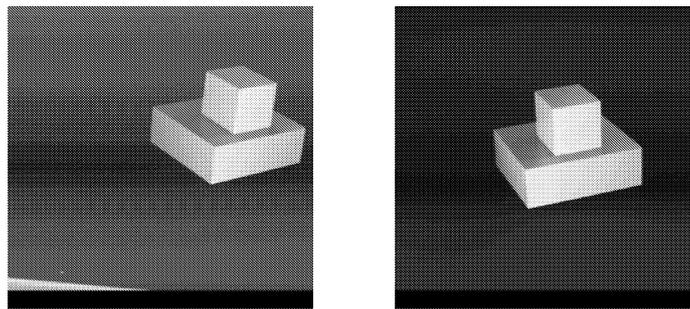
- ①**可能性:** Bループのセグメントを含むセグメントの組により, 平面が定義できなければならない. 但し, 明らかな実セグメントが存在する場合は必ずそれを含め, 明らかな虚セグメントが存在する場合はそれを含まない.
- ②**最奥性:** 同一ループ上のセグメントは, ①で定義される平面上または平面より手前 (視点側) に位置し, 平面より後方 (視点の反対側) に位置してはならない.
- ③**一意性:** 最奥性を満足する平面は唯一つでなければならない.

そして, 一意的に定義された面上に存在するセグメントを実セグメントに, また, 面より手前にあるセグメントを虚セグメントに分類する. なぜなら, 虚セグメントは, 実在する面が手前の面によって一部隠蔽されたとき, 手前の面上に発生するからである.

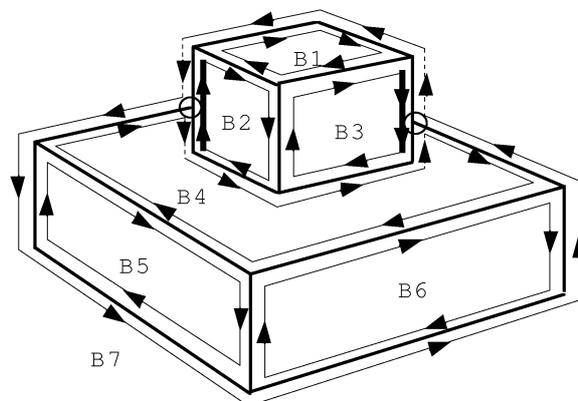
一方、上記の条件を満足しないループについては、次の空間性を調べる。すなわち、ループ上の各セグメントと翼状エッジを構成するセグメントのすべてが、実セグメントに分類されている場合には、ループは自由空間の境界を表しているとみなし、ループ上のすべてのセグメントを虚セグメントに分類する。

以下に、この実在テストの適用例を示す。

事例 1 一つのブロックが他の一つのブロックの上にあるシーンで、七つのループと二つのオクルージョン点が発見されている (図 3.13)。



(a) ステレオ原画像



(b) セグメントのループ

図 3.13: 重なったブロック

- (1) **ループ B 1, B 2, B 3, B 5, B 6**: 各ループ上のすべてのセグメントは、ただ一つの平面を構成するので、実セグメントに分類される。
- (2) **ループ B 4**: 上のブロックが下のブロックに接しているときは、すべての未分類セグメントは、ただ一つの平面を構成し、また最奥性を満足する (明らかな虚セグメントは面より手前に存在する) ので、実セグメントに分類される (図 3.14(a))。一方、上のブロックが下のブロックから離れてい

るときは，図 3.15 のように，複数の面が構成可能となるが，同図 (c) の面だけが最奥性を満足する．従って，面の手前にある上のブロックの下辺の二つのセグメントは虚セグメントに分類される (図 3.14(b)) ．

(3) ループ B7：結合する B ループの存在しない孤立 H ループは面を決定することはできないが，空間性により，図 3.16 のように，ループ上のすべてのセグメントは虚セグメントに分類される．

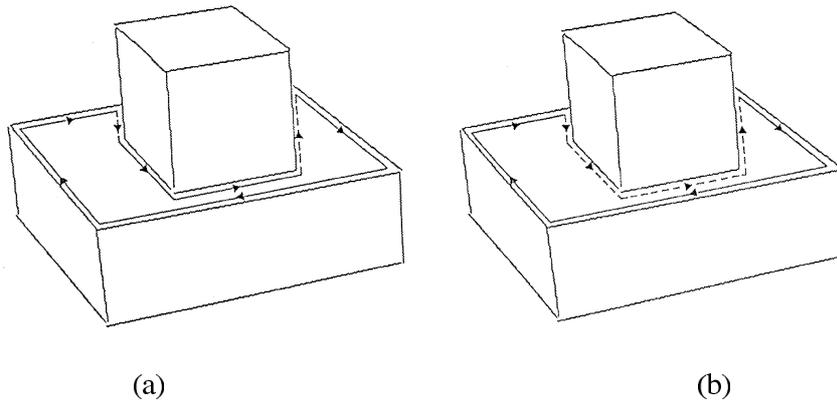


図 3.14: 実在テストの結果

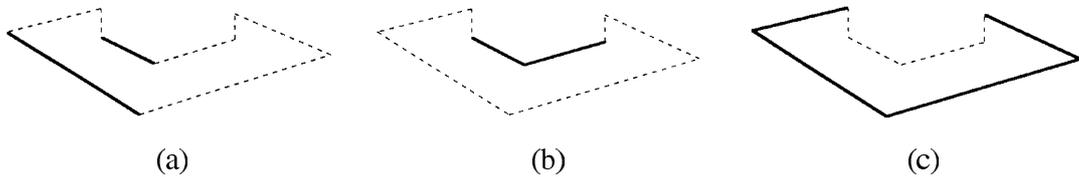


図 3.15: 面の候補

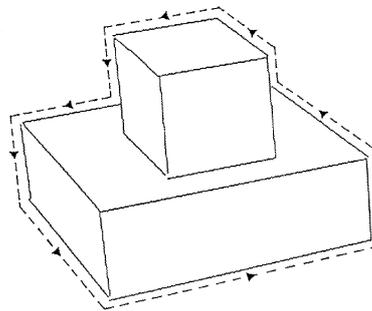
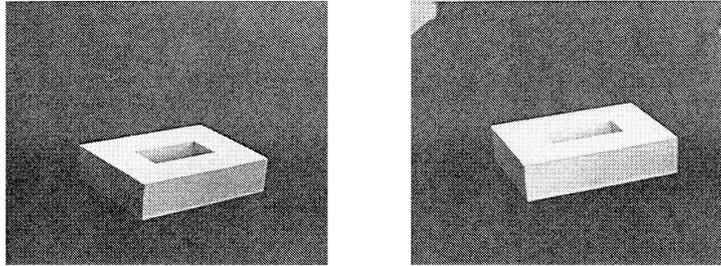
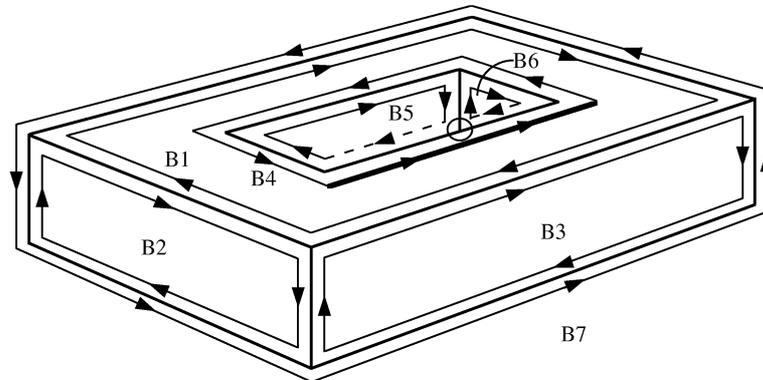


図 3.16: 実在テストの結果

事例 2 穴のあいたブロックのシーンで、七つのループと一つのオクルージョン点が発見されている (図 3.17) .



(a) ステレオ原画像



(b) 境界線セグメントのループ

図 3.17: 穴のあいたブロック

- (1) ループ **B 1 + B 4** (Bループ **B 1** と Hループ **B 4** は結合されている) : すべてのセグメントは、唯一つの平面を構成するので、実セグメントに分類される.
- (2) ループ **B 5** : 図 3.18 のように、二つの面が構成可能となるが、同図 (a) の面だけが最奥性を満足する. 従って、奥の二つのセグメントが実セグメントに、手前の二つのセグメントは虚セグメントに分類される (図 3.18(a)) .

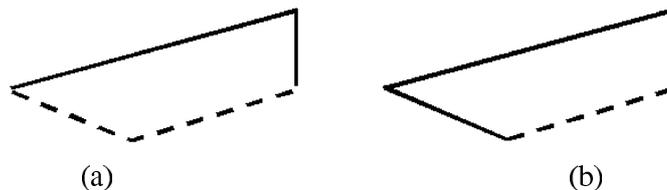


図 3.18: 面の候補

(3) **ループB 6**: ループB 5と同様に, 奥の二つのセグメントが実セグメントに, 他の一つのセグメントは虚セグメントに分類される (図 3.19).

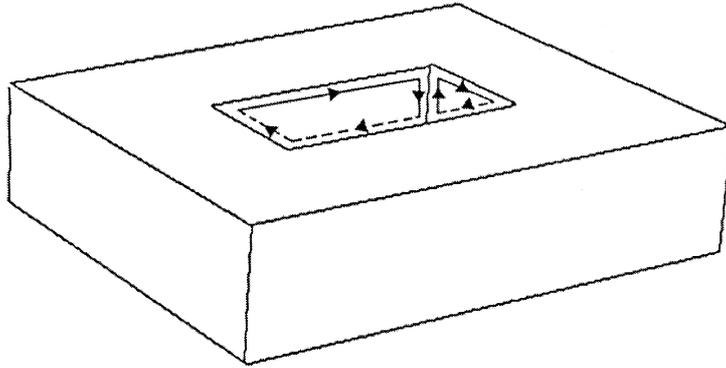
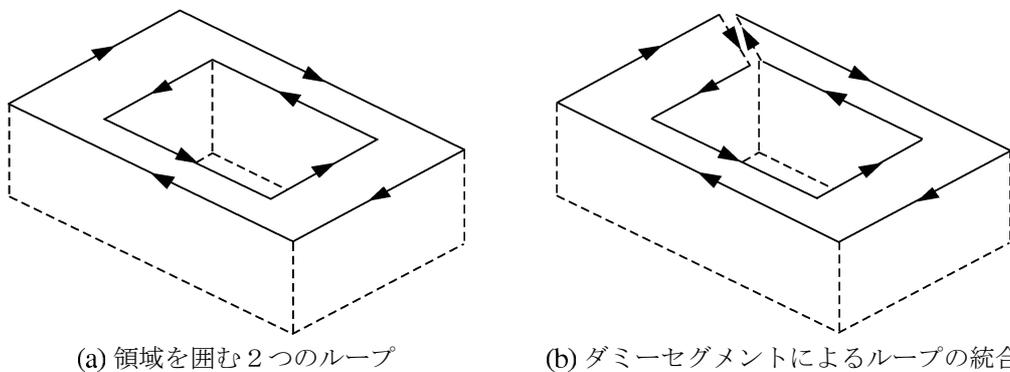


図 3.19: 実在テストの結果

左回りのHループ (例えば穴を表すループ) が存在する場合, 図 3.20 に示すように2本のダミーセグメントを作ることによってその外側の右回りのBループと統合する. 何故なら, 面が存在するとすれば同一の面を共有するからである.



(a) 領域を囲む2つのループ

(b) ダミーセグメントによるループの統合

図 3.20: 面の外周ループと穴のループの統合

以上の例は, いずれも物体の背景に何もないシーンであるが, 一般的な環境では, 物体が置かれている床などの境界のBループが通常観測される. 従って, Hループが孤立して, 空間性によりHループのすべてのセグメントが虚セグメントに分類されることは稀である. 床のBループと結合するHループの場合は, 床面上のセグメントが実セグメントに分類されることになる.

また, 視野内に物体の一部しか観測されない場合でも, オクルージョンの場合と同様に扱うことにより, 面を復元することができる. すなわち, 画像を四角い穴を通し

て見たものとして考えることができ、画像の枠上のセグメントをすべて明らかな虚セグメントに分類した後、各ループに実在テストを施せばよい。

3.6 境界線の分類と解釈不可能なループ

3.6.1 境界線の分類

以上の処理の結果、物体の境界線を、その境界線上の実セグメントの個数により、輪郭線と稜線に分類することができる。すなわち、輪郭線は一つの実セグメントしか持たないのに対して、稜線は互いに反対向きの二つの実セグメントを持つ。更に、稜線は、図 3.21 のように、その両側の面の法線方向のなす角度が正、負、ゼロによって、凸、凹、平（模様あるいは影による）稜線にそれぞれ分類することができる。



図 3.21: 稜線の両側の法線

例えば、図 3.22 では、図 3.3(b) のステレオセグメント画像よりセグメントの分類ならびに面の推定を行い、各境界線に対して輪郭線 (→), 凸稜線 (+), 凹稜線 (-) のラベル付けを行なっている。

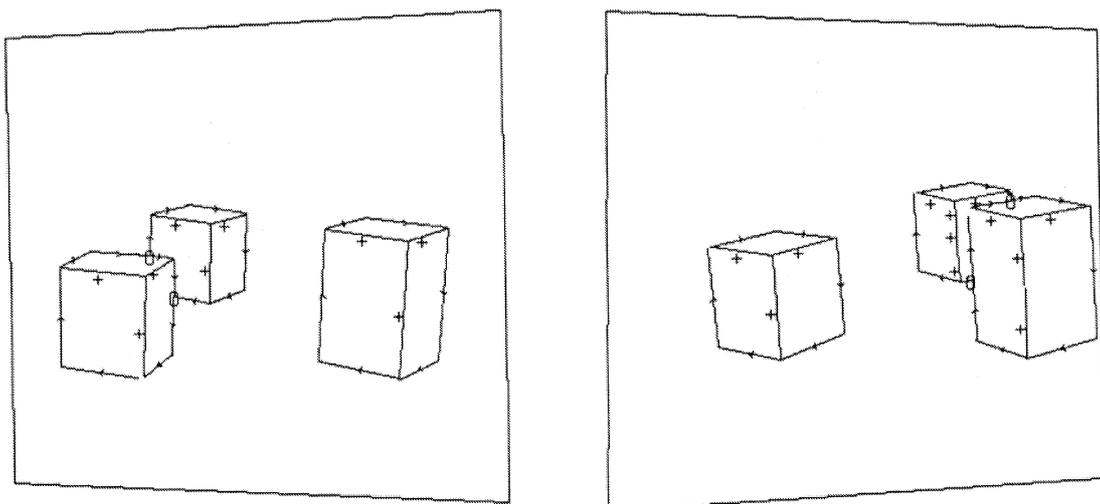
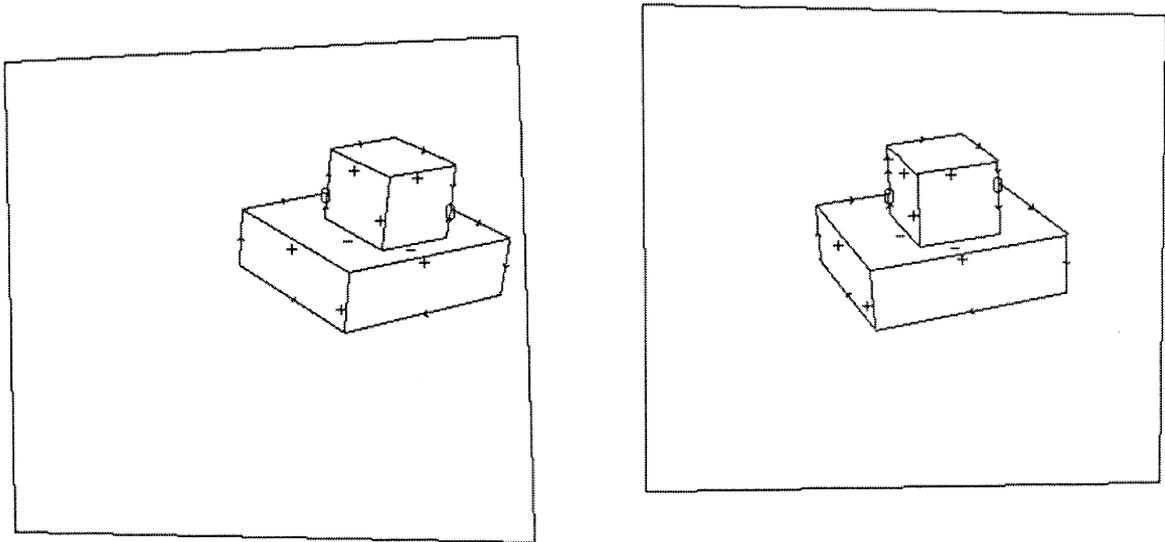
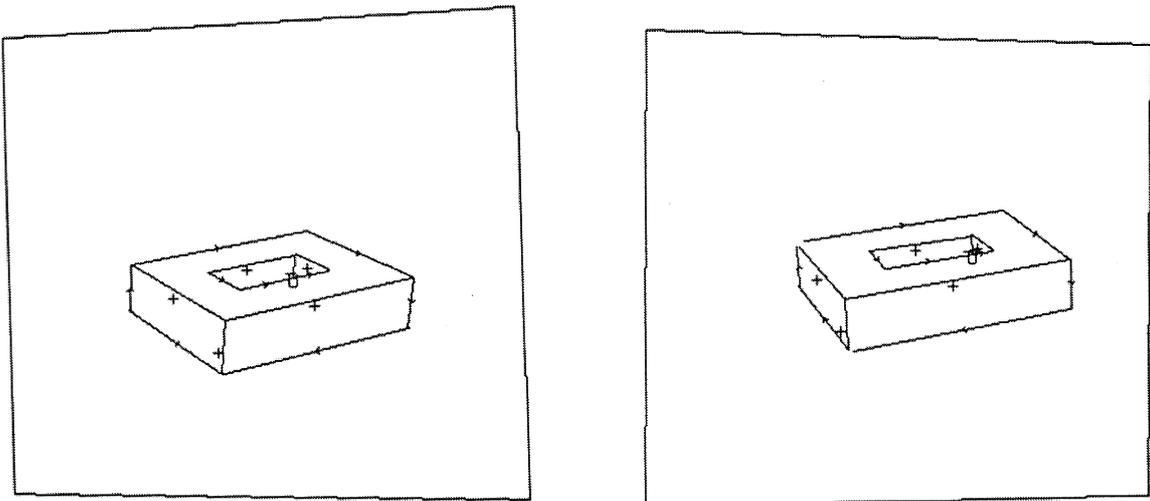


図 3.22: 境界線の分類 (1)

図 3.23(a) には, 図 3.13 のステレオ画像より, 同図 (b) には, 図 3.17 のステレオ画像より, セグメントの分類ならびに面の推定を行い, 各境界線に対してラベル付けを行なった結果を示す.



(a)重なったブロック



(b)穴のあいたブロック

図 3.23: 境界線の分類 (2)

図 3.24(a) の模様のあるブロックからは, 同図 (b) のように, 平稜線 (P) が検出されるが, それらを除去することによって, 同図 (c) のように, 物体の純粹の立体形状だけを得ることも可能となる.

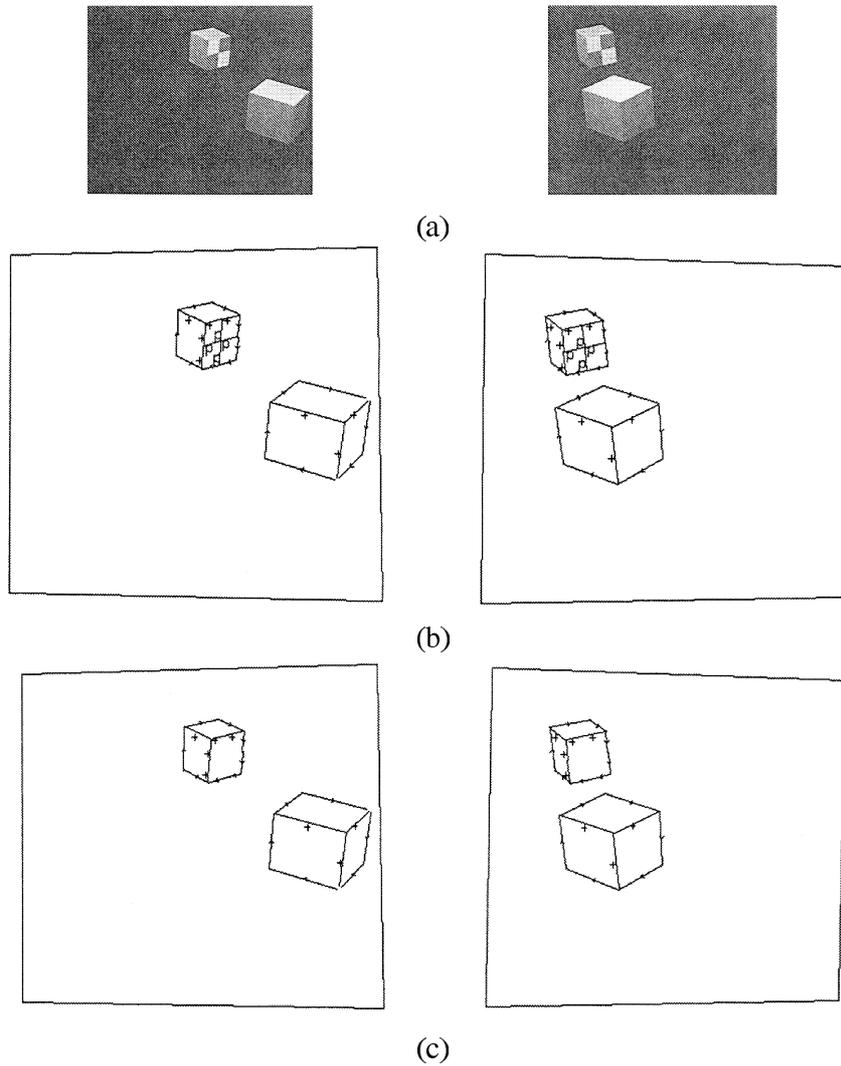


図 3.24: 境界線の分類 (3)

3.6.2 仮セグメント

最後に、一部隠蔽された面の境界を一時的に補間する（虚セグメントの代わりとなる）仮セグメントを生成する。3.25 に点線で示すように、仮セグメントは、ループ上の虚セグメントへの視線を定義された面に投影することによって得られる。

図 3.27 に、各シーンの物体の面の復元後の幾何モデル（境界表現）を示す。

3.6.3 解釈不可能なループ

しかし、観測される明るさの変化が小さいために、存在すべき境界線の一部が検出できていない場合には、関連するループは実在テストを満足せずに「解釈不可能」に

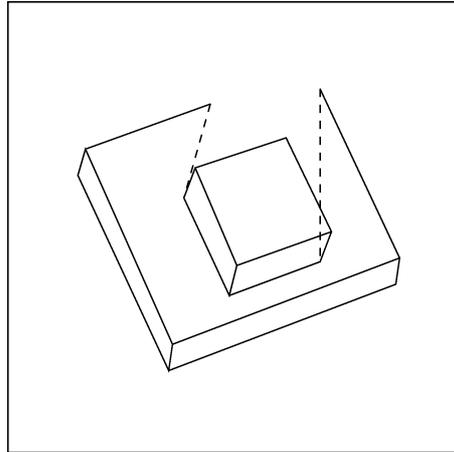


図 3.25: 仮セグメント (点線)

なる。つまり、ループに対して面あるいは空間を定めることができない。

例えば、図 3.26(a) の家の屋根の凸稜線（尾根線）が、同図 (b) のように、検出されなかった場合のループを考える。このループに対しては、同図 (c), (d) に太線で表示された各二本のセグメントによって定義される二つの面（網掛け表示部）が最奥性を満足する。しかし、これは一意性の条件に違反するので、このループは解釈不可能となる。

このように、解釈不可能なループは、その内部に境界線の欠損があることを示しており、現時点で無理矢理面を生成することはしないで、解釈不可能なループとして保存し、欠損を補う後の観測（よく見直す、あるいは、別の方向から見るなど）にゆだねることになる。

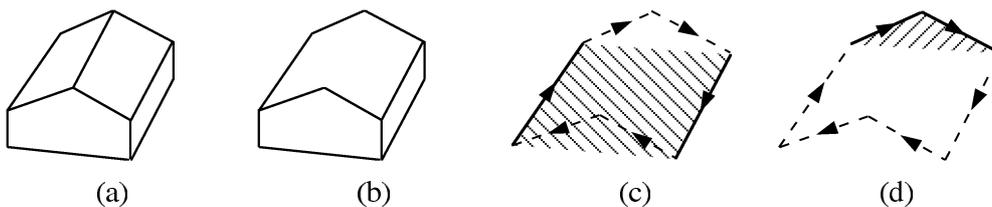


図 3.26: 解釈不可能なループ

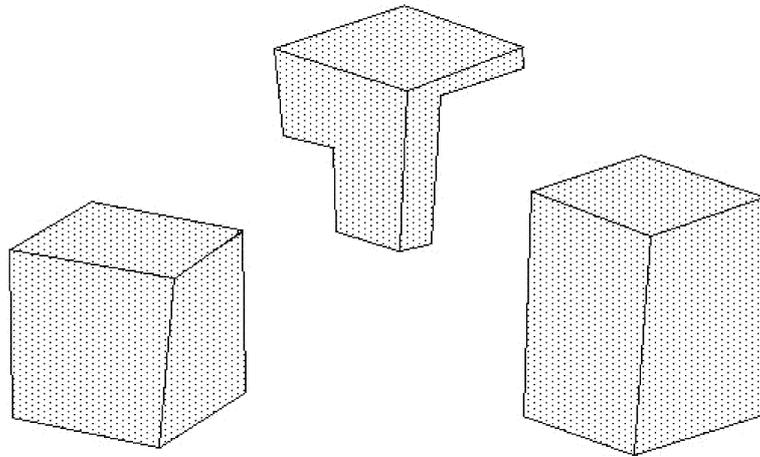
3.7 第3章のまとめ

本章では、単一視点からのステレオ視によって得られる物体の境界線のワイヤフレームから正しく平面を推定するアルゴリズムにより物体の面を復元し、内挿された平面に基づいて観測シーンの境界表現を生成する方法について述べた。物体の境界線

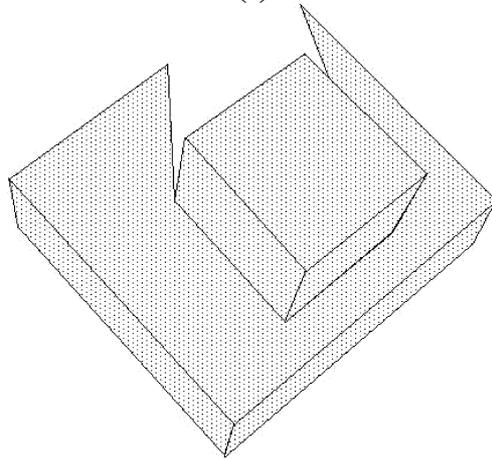
は、右の面の境界であることを示す方向を持つセグメントのループによって表現される。最初に、オクルージョンの生じている個所で、明らかな実セグメントと明らかな虚セグメントを求める。次に、実在テストによって、各ループ上の他の未分類セグメントを実セグメントまたは虚セグメントに分類し、結果として、実セグメントによって各ループに張るべき面が定義される。また、境界線を、一つの実セグメントしか持たない輪郭線と互いに反対の向きの二つの実セグメントを持つ稜線に分類し、更に、稜線を、その両側の面の法線方向のなす角度が、正、負、ゼロによって、凸、凹、平稜線にそれぞれ分類する。実在テストにより決定された平面を内挿することにより、面ベースの記述が可能となる。最後に、虚セグメントの代わりに仮セグメントを生成して各面の境界を決定し、境界表現を完成する。但し、欠損があるために実在テストを満足しない境界線ループには、それを補償する後の観測のために解釈不可能なループとしてラベルを付与し保存する。これにより、境界線の間を単純に線形補間することにより不確かな（ともすれば実際には存在しない偽の）面を生成してしまうといった問題を回避できるとともに、面の内挿を行う際にデータのどの部分が不完全であるために解釈不可能となるのか検出可能な局所性を備えている。また、環境の記述としては、CADの分野で開発された最も一般的な幾何モデルであるいわゆる境界表現(b-rep)に基づいている。

この単一視点のステレオによって得られる物体の境界線のワイヤフレームから正しく平面を生成するアルゴリズムにより物体の面を復元し、内挿された平面に基づいて観測シーンの境界表現を生成する方法により、エッジベースステレオによる適用範囲の拡張をはかることで、汎用性の向上に貢献できるものと考えている。

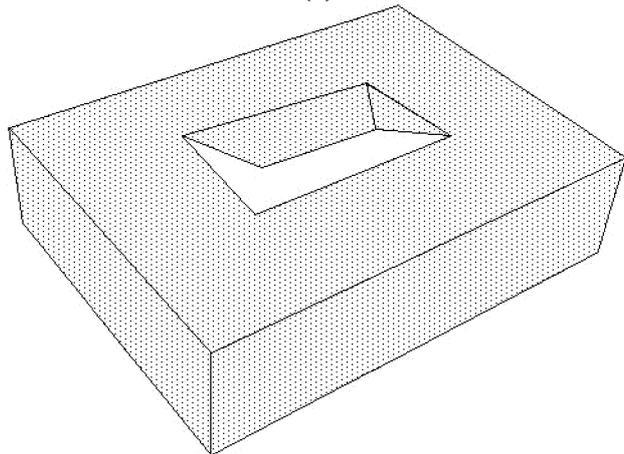
しかしながら、単一視点より観測可能な視野には限界があり、照明条件・雑音等によるエッジ検出の失敗以外にも観測できない部分が残る。より完全な幾何モデルを生成するためには、視点の移動により新たに観測される部分の統合・更新を行う必要がある。この統合・更新手法については、次章にて紹介する。



(a)



(b)



(c)

図 3.27: 幾何モデル

第4章 移動ステレオによる統合と解釈

4.1 具体的課題と関連研究ならびに概要

第3章にて、単一視点からのステレオ視によって得られる物体の境界線のワイヤフレームから平面を推定し、内挿された平面に基いて観測シーンの境界表現を生成する方法について述べた。しかしながら、単一視点より観測可能な視野には限界があり、観測できない部分も残ってしまう。シーンのより完全な幾何モデルを生成するためには、照明条件・雑音等によりエッジ検出を失敗した部分の再観測による情報獲得に加え、視点の移動により新たに観測される部分の統合・更新も行う必要がある。

モデリングに関しては多視点画像や全周画像を用いた手法も多く提案されているが、蓄積された全多視点画像を用いる因子分解法 [110] や Structure-from-motion [111] といったアプローチによる形状の再構成、視体積交差法 [112][113] によるモデリングに関するアプローチ等とは異なり、提案手法ではステレオカメラを移動することにより、各視点にて生成した環境の境界表現を、既に構築している環境の幾何モデルと逐次統合・更新する。

ロボットは不完全な世界地図（環境の幾何モデル）を携えながら移動可能な空間を移動し、任意の地点における観測データに基づき、各地点における環境の局所地図を作成する。そして、新たに得られた局所地図と世界地図間で対応を求め、対応の求まる部分については精度向上をはかり、新たに得られた部分については、追加統合することにより世界地図を更新する。

移動ロボットの距離センサとしては超音波やレーザを利用する能動センサが一般的に用いられているが [196]、視覚センサとしてステレオを用いたものも見られるようになってきた [114]。しかし、ステレオには、対応問題に加えて測定精度、出力の形式、面の内挿といった問題があり、複雑な一般シーンの解析が難しく、実用化に際しては残念ながら能動センサに遅れをとっている。対応問題を解くことに関するだけでも、多くのアルゴリズムが提案されていることは前章にて紹介した通りである。その多くは、対応の単位として特徴点、特にゼロクロッシングを用いている。しかしながら、点単位のステレオでは、視差を距離に変換すると誤差や雑音が増幅されるためか、出力のほとんどは視差分布図にとどまり、距離測定まで行っているものは少ない。

また、対象とする画像から特徴点を密に抽出し、移動前後の画像間でそれらの可能な組合せを網羅的に検証しながら外れ値を除外 [115] しつつ解釈可能な組合せを探索

するといったアプローチもあるが、特徴点数が限られる場合には適用できない。そこで、エッジ部分の限られた情報に基いて面の推定をしなければならないといった場合、例えば、一様な平面により構成される多面体世界における対応探索と得られたエッジ情報からの平面の推定のための課題解決に関する手法については、前章にて述べた通りである。

視点の移動にともない、異なる視点においても新たな境界表現（局所地図）が得られる。各視点にて得られる局所地図を逐次統合することにより、作成した世界地図をより完全なものへと更新することが可能となる。例えば、移動前の世界地図の座標系を基準とするならば、移動後の視点より得られたデータを移動前の世界座標系に変換した後照合を行い、移動前後で対応するセグメント対を決定する必要がある。ここで、対応が得られたセグメントに関しては、より視点に近い位置で観測されるといった、信頼性の高い方のデータを用いて高精度化がはかられるべきである。移動後の新たな観測により獲得された部分に関しては、移動前の境界表現には対応する部分は見つからないため、他の対応部分との整合性をとりながら追加統合されなければならない。しかしながら、特に、データの一部が不完全であるために解釈不可能となっているループに関しては、視点の移動により、欠損していた境界線について再観測することで検出し直し、既に得られている境界線と統合することにより正しく面の内挿を行わねばならない。

そこで、本章では、ステレオビジョンによって未知の環境を認識しながら自律的に移動するロボットへの応用を想定し、異なる観測点から得られる部分的なステレオデータ間での対応探索や、不完全なステレオデータを逐次的に統合・更新して、面表現による記述を生成する方法について述べる。以降の各節にて、処理対象とするシーン中の物体のクラス、データ構造、統合の際のセグメントの信頼性と移動前後のセグメントの対応探索、統合後の再解釈と検証について実データを用いて紹介する。

4.2 世界と記述

ロボットは、事前に自分の置かれた環境内に存在する物体に関する知識を持たず、むしろどんな物体がどこに存在するかを知ることを目的とする。ここでは物体は平面のみで構成されたものに限定する。図4.1に示すように、物体はエッジが凸または凹の多面体 (a) で、3面（以下）頂点に限定しない (b)。何枚かの紙で構成されていてもよく (c)、穴のある物体でもよい (d)。また、物体の表面には影や模様があってもよい (e)。さらに、物体はどのような姿勢でもよく、面やエッジはどのような方向を向いていてもよい。そして、複数の物体が重なっていてもよい (f)。ただし、透明な面は検出できないため、物体は不透明な面のみによって記述され、そのためのデータ構造として境界表現を採用した。これは、図4.2に示すように物体の面 (F) と

エッジ (E) の関係を表すグラフ構造となっている。各エッジは互いに反対の向きを持つ2本のセグメントによって表され、いわゆる翼状エッジ (winged edge) 構造となっている。各セグメントの方向は、セグメントが構成する面を右に見るように設定されている。したがって、紙は表裏二つの面をもつことになり、各面は、図4.3に示すように、互いに反対方向をもった異なるセグメントのリストによって構成される。

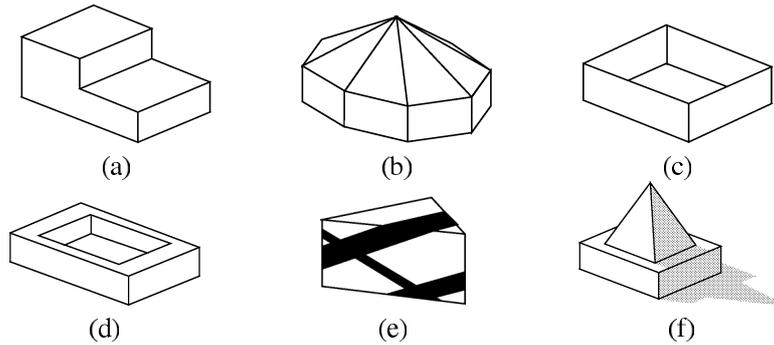
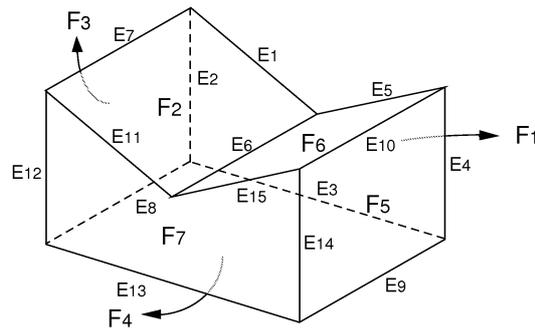
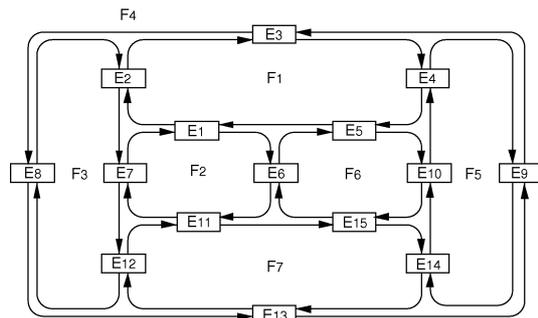


図 4.1: シーン中に存在する物体



(a) 物体の構成要素



(b) データ構造

図 4.2: 物体の境界表現

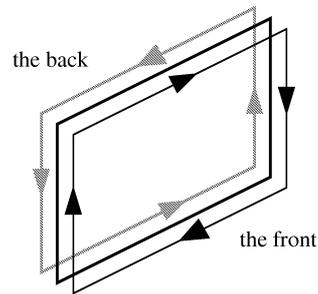


図 4.3: 紙の裏と表の面を表す二つのループ

4.3 データの統合と再解釈

第3章にて述べた手法に基づいて、セグメントベースのステレオ解析と実在テスト (3.5.2 参照) により、面の内挿ならびに境界線の分類 (3.6.1 参照) を行う。しかしながら、照明条件や雑音等の影響により、理想的にエッジ情報が得られるとは限らず、一意に解釈できない場合がある。そこで、その様な解釈不可能な領域に関しては、観測点を若干変更することにより得られるエッジ情報に基づき、新たに解釈可能となった部分や既に解釈されている領域に関しても、より対象に近い視点から観測された画像データより再構成された精度の高い記述が得られた部分を逐次統合することにより、対象物のより詳細なモデルを生成できることが好ましい。

4.3.1 セグメントの信頼性

ステレオ観測データより再構成されたセグメントの3次元位置の正確さは、視点からの距離と傾きによる。各セグメントには、この正確さの情報ももたせてある。ステレオの各画像は、カメラの光軸が平行で、画像面は光軸に直交し、基線方向と画像横軸が平行となるような標準カメラモデル座標系 (図 3.2 参照) に変換されている。

ステレオ画像中で対応する2次元のセグメントを考える。ここで、 (x_s, y_s) , (x_e, y_e) は左画像中のセグメントの両端点とし、 (x'_s, y'_s) , (x'_e, y'_e) を右画像中の両端点とする。対応の結果得られる3次元セグメントの精度を、二つの対応するセグメントの間の面積：

$$\frac{1}{2} \{ (x_s - x'_s) + (x_e - x'_e) \} \cdot |y_s - y_e|$$

で定義する。したがって、視点により近くより垂直な3次元セグメントほど大きな値をもつこととなる。

4.3.2 実在テスト2

3.6.3 で述べたように，見えている面上に実セグメントが存在しない場合や，オクルージョンや照明の具合により境界線の一部が検出できない場合には，関連するループは実在テストを満足せずに「解釈不可能」となる．つまり，そのループに対して唯一の面を内挿することができない．例えば，図4.4(a)に示すような，本のように凹に折った紙を考える．同図(b)に示すように凹稜線が検出できなかったとすると，内側のループに対して，同図(c)，同図(d)の影をつけた二つの面が候補となり，一意性を満足しないためにこのループは解釈不可能となる．

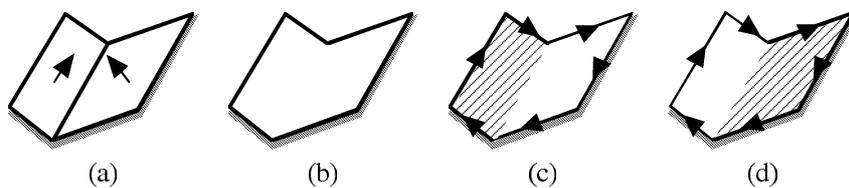


図 4.4: 解釈不可能なループ 1

また，注目しているループは，隣接するループの解釈の結果によっても，解釈不可能になる場合がある．図4.5(a)に示すように，屋根のように凸に折った紙を考える．そして，同図(b)のように凸稜線が検出できなかった場合を考える．内側の領域のループに対して，同図(c)の影をつけた面が実在テスト1を満足する．したがって，その面より手前にあるセグメント（薄い線）は虚セグメントであると判定される．一方，同図(d)の外側の領域のループに対しては，結合するBループの存在しない孤立Hループであり，自由空間との境界を表していると考えられ，ループ上のすべてのセグメントは虚セグメントであると判定される(3.5.2参照)．その結果，定義された面より手前の境界線上には実セグメントは存在しないこととなる．つまり，境界線が存在しないことと同義となってしまう．しかし，実際に境界線は検出されているので矛盾する．このような場合には，両方のループは解釈不可能となる．虚セグメントと解釈不可能なループ上のセグメントには，「未定」の印と共に，移動後の処理において利用するために，接続関係を記録する．

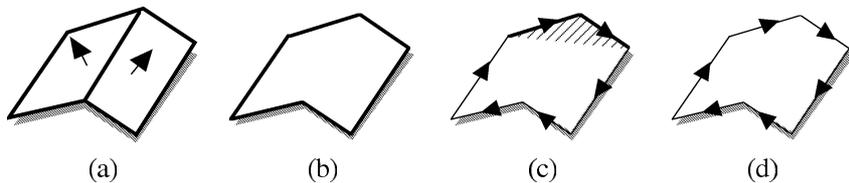
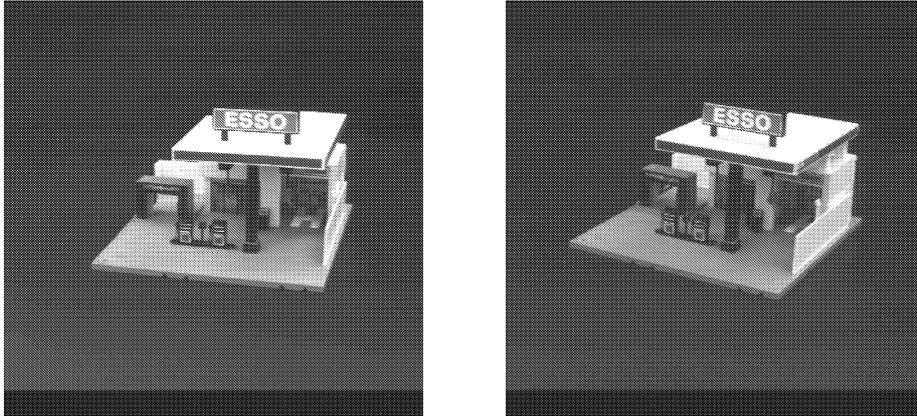


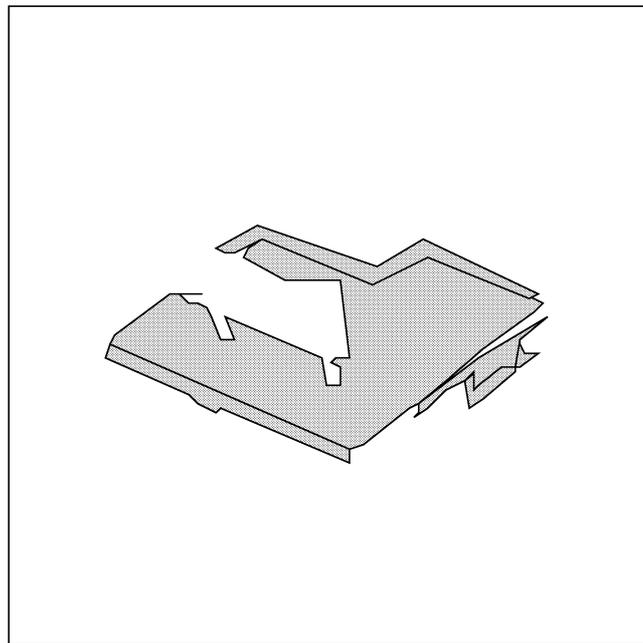
図 4.5: 解釈不可能なループ 2

4.3.3 データの統合

移動ロボットは、初期位置での環境の観測から、図4.6に示すような初期の世界地図を持つものと仮定する（定義された面に対しては陰影が施されている）。そして、ロボットが移動することによって、図4.7に示すように、新たな位置から観測した環境の局所地図を作成する。

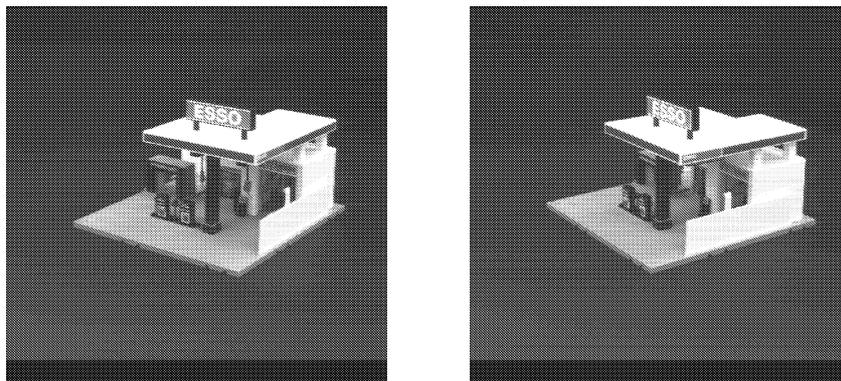


(a) 初期位置におけるステレオ原画像

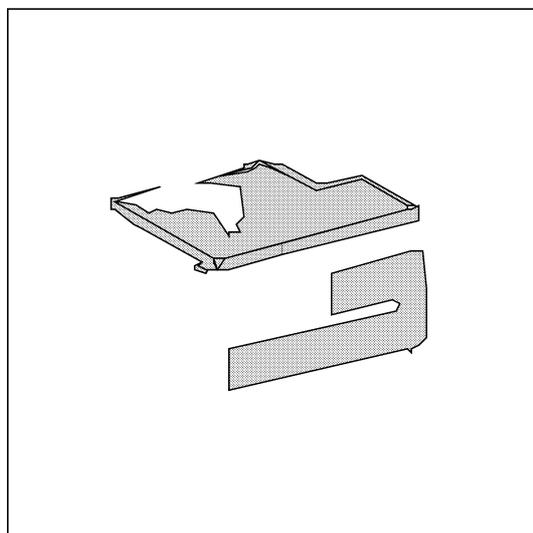


(b) 初期の世界地図

図4.6: 第1の視点における世界地図



(a) 第2の視点からのステレオ原画像



(b) 局所地図

図 4.7: 第2の視点における局所地図

4.3.4 対応探索

まず初めに，移動前後の両地図間での対応を求める．移動前の座標系と移動後の座標系の移動量を回転行列 R ，平行移動ベクトル t とし， x_i と x'_i をそれぞれ移動前後のカメラで観測された同一点の座標とすると，次の関係が成り立つ．

$$\mathbf{x}_i = R^{-1}\mathbf{x}'_i - \mathbf{t}$$

移動パラメータはロボットの駆動系から大まかに求めることができる．これを移動パラメータの初期値として用いる．正確なパラメータは，世界地図と局所地図との対応を求めることにより，以下のように計算により求めることができる．

- 1) 移動パラメータの初期値を用いて局所地図中のセグメントを座標変換して，世

界地図と局所地図を重ね合わせる。移動パラメータには誤差が含まれるため、対応するセグメントも正確には重ならない。

- 2) 図4.8に示すように、局所地図中のセグメントの両端点 (A', B') から世界地図中のセグメントを含む直線上の点に対して垂線を下ろすことにより、局所地図中の各セグメントに対して同じ方向を持つセグメントで最も近いものを選び出す。すなわち、世界地図中のセグメントの CD の部分に対して、局所地図中のセグメントの $C'D'$ の部分に対応すると仮定した場合に、 $CC' + DD'$ が最小となるセグメントを選ぶ。

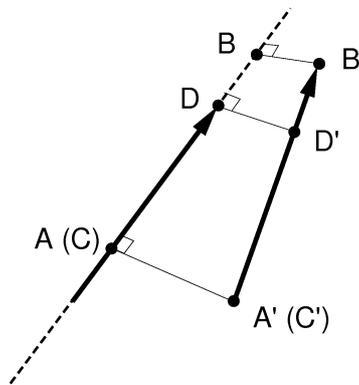


図 4.8: セグメントの対応

- 3) しかし、 CC' か DD' がしきい値 τ より大きい場合には、両セグメントは対応するとはみなされない。しきい値以下の場合には、両セグメントは対応するとみなされ、対応の誤差を次式によって定義する。

$$\frac{1}{2}(|CC'| + |DD'|) \cdot |CD|$$

- 4) 局所地図中のすべてのセグメントの誤差の総和を S とし、最小二乗法によりこの値 S を最小にするように移動パラメータを計算し直す。

誤対応を取り除くために τ を小さくして行き、誤差が十分小さくなるまで、上述した処理を繰り返す。オクルージョン等により各視点においてセグメントの対応する部分が常に観測されるとは限らないため、対応探索には頂点情報は用いない。図4.9は世界地図（第一の局所地図、図4.6(b)参照）中のセグメント（実線）に対して、第二の局所地図（図4.7(b)参照）中のセグメント（破線）を重ね合わせた結果である。

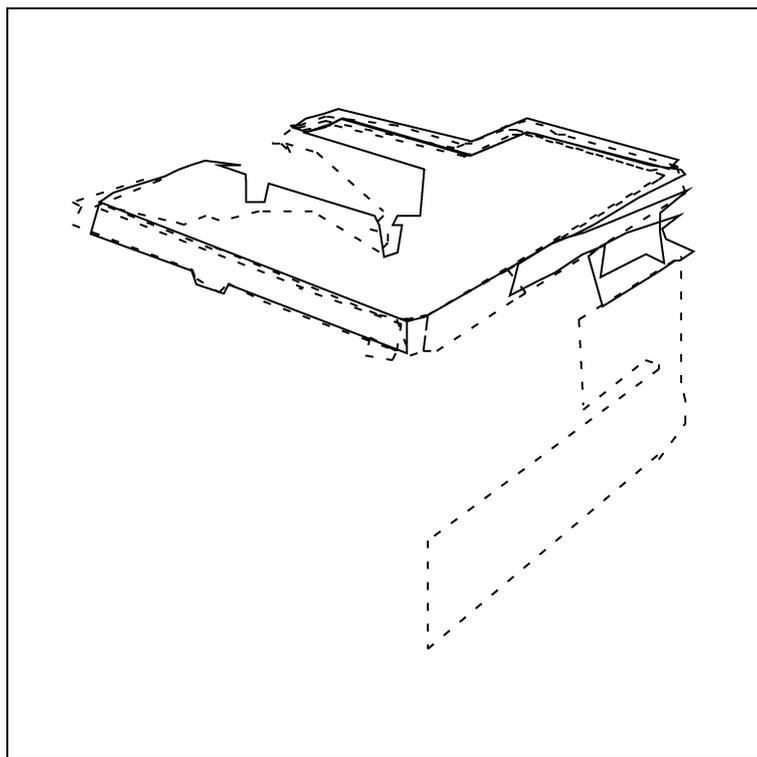


図 4.9: 初期世界地図（実線）と第二の局所地図（破線）との対応

4.3.5 再解釈

世界地図と局所地図間で対応の求まった部分の更新や、新たな観測点における局所地図中にて新規に解釈可能となった部分を逐次統合・更新する際、単純な統合ではなく、世界地図中の解釈不可能な部分の再解釈と合わせ、全体の解釈の整合性を検証しながら進めなければならない。

図 4.10 (a) の世界地図および図 4.10 (b) の局所地図を考える。陰影をつけた部分は面が内挿されていることを示している。一方、陰影の付いていないループは解釈不可能で、面が内挿されるのを待っている部分である。局所地図中で解釈不可能なループを解釈するために、欠けているセグメントを世界地図中より検出し統合することにより、新たなループを生成し再解釈する必要がある。

- 1) 局所地図中の解釈不可能なループ L' 上のセグメントに対応するセグメントが世界地図中の解釈可能なループ L 上に存在する場合には、ループ L 上のすべてのセグメントと、翼状エッジを構成している反対向きのセグメントとをループ L' に統合する。その結果、ループ L' 中に別の新たなサブループが生成される。このとき、少なくともループ L に対応するサブループは図 4.10 (c) のループ A のように、解釈可能となる。

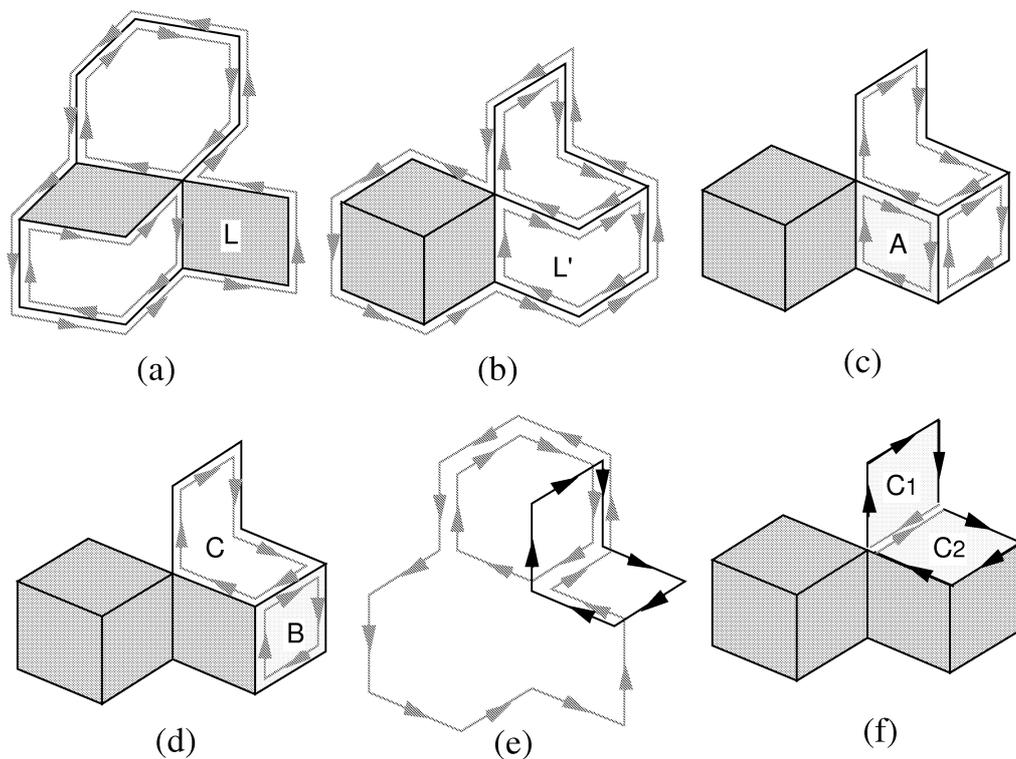


図 4.10: 世界地図を用いた局所地図の再解釈

- 2) 残りの各サブグループに対して実在テスト 1 を適用し、内挿可能な場合には、図 4.10 (d) のループ B のように面を内挿する。
- 3) 図 4.10 (d) のループ C のように依然解釈不可能なループが残っている場合、図 4.10 (e) のように、世界地図中の未定のセグメントのなかから、ループの異なる 2 点を結ぶようなセグメントの系列をすべて求める。
- 4) 新たに生成された各サブグループに対して実在テスト 1 を適用し、内挿可能な場合には、図 4.10 (f) のループ C₁ と C₂ のように面を内挿する。
- 5) そして、すべての解釈可能なループに対して実在テスト 2 を適用する。

4.3.6 検証

次に、異なる視点から同じ面を見た場合について考える (図 4.11) . 同図 (b) (c) にそれぞれ示すように、いま、それぞれ独立に解釈すると以下ようになる。ループ L₁ の内側後方には背景以外の物体が見えるので、このループは穴であると判断され、面は内挿されない (同図 (b)) .

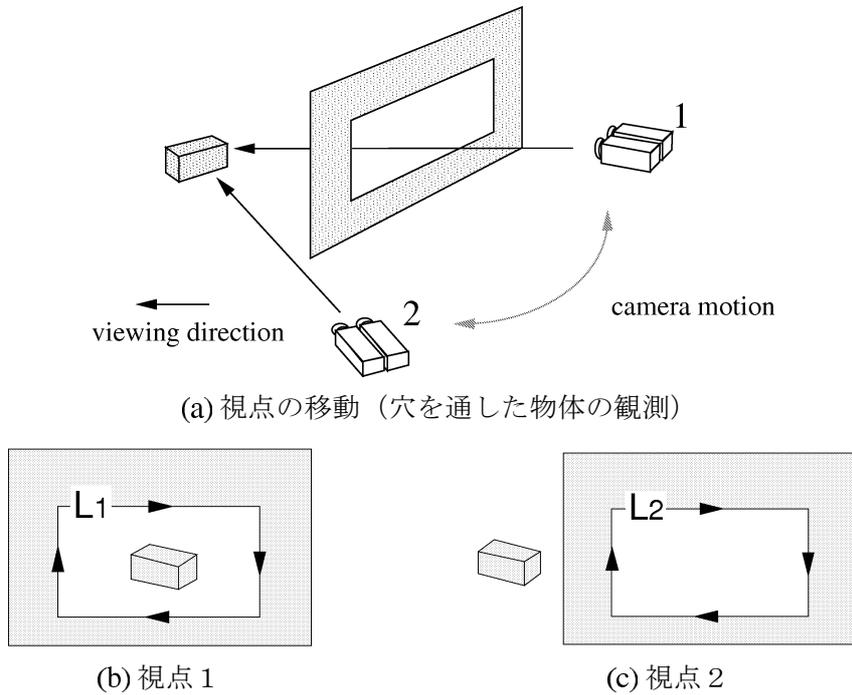


図 4.11: 異なる視点から見た同一のループ

一方、ループ L_2 の内側には何も見えていないので、このループには面を内挿する (同図 (c))。したがって、視点 1 が先で視点 2 が後ならば、先に穴の情報が得られるので、この内側のループに対して面を内挿することは避けられる。しかし、視点 2 が先で視点 1 が後の場合には、後の視点で初めて穴であったことが分かるので、前の視点で内挿されていた面は誤りであり取り除かねばならない。このように、新たな局所データを用いて世界地図を更新する際、その視点で見えている世界地図中の面に対して、局所データ中のセグメントを隠すような面があるかどうか常に調べなければならない。そして、そのような面が存在すれば世界地図から取り除かねばならないし、その面を構成していたセグメントは、現在の視点からは見えない面上のセグメントでなければならない。

最初の視点 1 では、何らかの都合により解釈不可能なループが、次の視点 2 で解釈可能となる場合には、問題はより難しくなる。このループが穴であるという情報はどこにもないので、面が内挿されてしまう。したがって、以前の視点 1 に戻り、この面が他のいかなるセグメントをも隠さないことを調べる必要がある。このように、移動ロボットは、それまでの世界地図中に存在しなかった新たな面を生成する際には常に、今までのすべての視点位置において、その面が何も隠さないことを調べなければならない。このことは、すべての視点情報とすべてのセグメントがどの視点から見えたかという情報を持っていないなければならないことを意味する。しかしながら、実際の移動

ロボットへの応用を考えた場合，すべてのデータを持つ必要はないと考える．なぜなら，ロボットが連続的に環境を観測しているのであればこのようなことが起こるのはいずれであるし，起きたとしても最新の数カ所での視点でのデータを持っていれば実用上は十分であると考えられるからである．

世界地図は新たに得られた局所地図を統合することにより更新される．局所地図中のセグメントで世界地図中に対応するセグメントが存在しない場合には，このセグメントは世界地図中に追加される．世界地図中に対応するセグメントが存在する場合には，精度の良いものを世界地図中に統合することにより，世界地図を更新する．

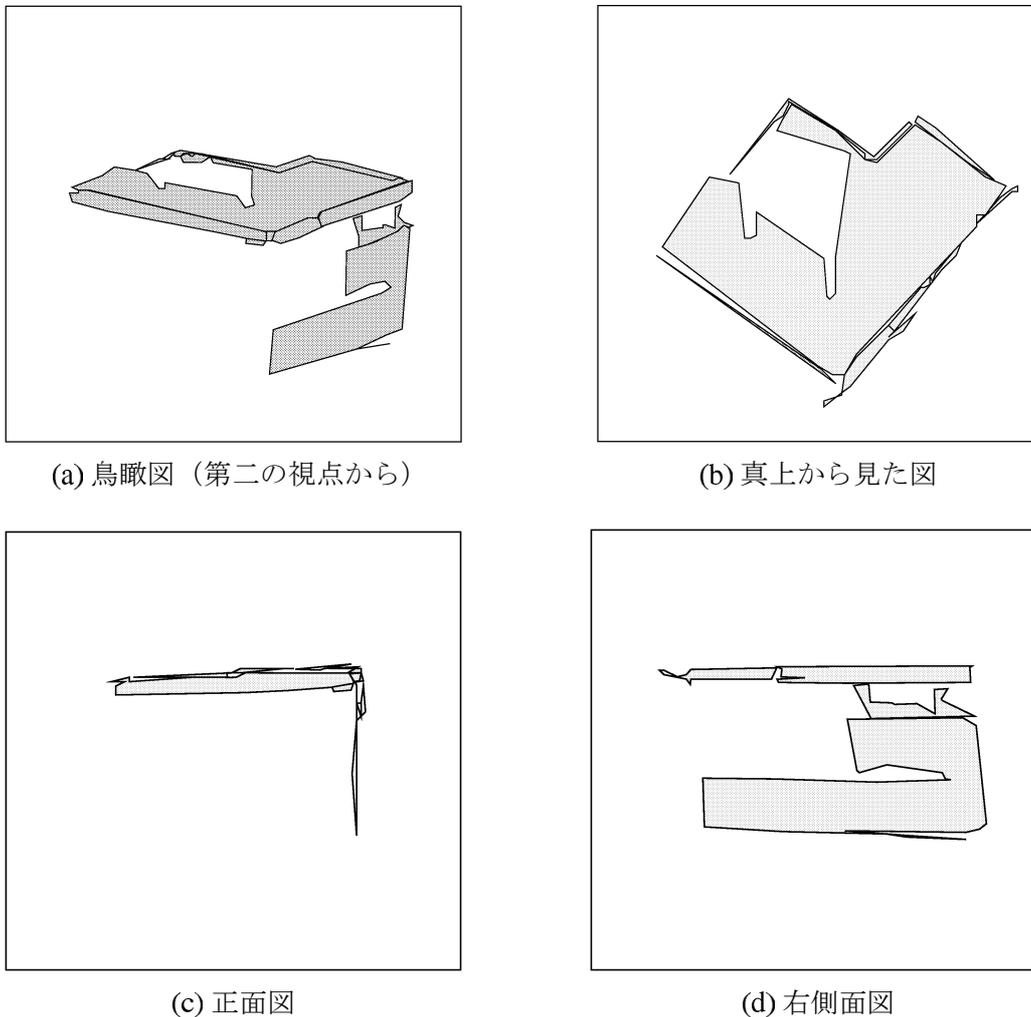


図 4.12: 第二の観測後に更新された世界地図

4.3.7 考察

図4.12では、第二の観測後に更新された世界地図を示している。図4.6 (b) で内挿された面のうち、天板下の右側面で、重なっている二つの面の一方は、第二の観測後の検証によって内挿されてはならない穴として更新されている。

ここでの検証は、移動前後で対応の求まったすべてのセグメントと世界モデル中の面に対して、面がセグメントを隠蔽していないか調べることにより行った。セグメントの各端点から、移動後の視点へ向かう半直線と各面との交点が、内挿された面領域中に含まれているかにより判定している。また、第一、第二両観測データを用いた再解釈により、新たに出現した面と合わせて世界地図中で記述の更新がなされ、例えば天板およびそれと接している右側面等は、視点の移動により得られていなかったエッジが新たに検出され、一つのエッジ上に二つの実セグメントをもつ稜線を構成する面として正しく接続できている。

図4.13 (a) は、第三の視点から観測した環境であり、図4.13 (b) は、世界地図中のセグメント（実線）と第三の視点で得られた局所地図中のセグメント（破線）との対応を、比較のため第二の視点からの鳥瞰図で示している。

図4.14は、第三の観測後に、エッジベースステレオによる面ベースの記述に基づいて統合・更新された世界地図を示したものである。第二、第三の両観測データ中で、唯一天板上部奥の小さな面で対応する部分が求まり更新されるにとどまっている。新たに得られた他の面は、世界地図中に追加されているが、再解釈には至っていない。

こうして得られた世界地図中の面の復元形状には、エッジの検出が不十分なことに起因した観測誤差は含まれるものの、実際の撮像対象の実形状と比較して矛盾はなく、良好な結果が得られているといえる。しかし、各観測で得られるループ数が少ない点が問題である。特に対象シーン中の給油器付近や広告看板部のデータは、視点の移動によっても解釈不可能なループのままである。

なお、各観測点データ中の3次元セグメント数、解釈不可能なループ数（孤立セグメントおよび非閉ループを除く）、内挿済みループ数は表4.1の通りである。

表 4.1: 各観測点データ中の3次元セグメント数, ループ数等

	第一の視点	第二の視点	第三の視点
3次元セグメント数	261	286	248
解釈不可能なループ数	15	8	12
内挿済みループ数	5	9	5

4.4 第4章のまとめ

本章にて、未知の環境における移動ロボットのためのステレオビジョンへの応用を念頭に、異なる観測位置から得られる不完全なステレオデータを逐次統合しながら、境界表現一面表現による環境の記述を生成する手法について述べた。複数地点から獲得されるデータを逐次統合解釈する際、各視点において、セグメントの欠けがなくすべてのループの解釈が可能となる完全なステレオデータを得ることは難しいが、検出できないセグメントをできるだけ検出することにより、解釈不可能なループを減らす努力をしていかねばならない。

本章で述べたアルゴリズムによれば、面を内挿する際にデータのどの部分が不完全であるか知ることができるので、各観測において注視の手法を用いることができ、さらに次なる観測位置の計画に利用することもできる。また、曲面物体の復元手法との融合や移動物体を含むより複雑な環境下でのステレオの応用を行なっていかなければならない。

提案手法ではステレオカメラを移動することにより、各視点にて生成した環境の境界表現に基づく局所地図（環境の幾何モデル）を世界地図（既に構築されている環境の幾何モデル）と逐次統合・更新するアプローチをとっているが、移動に伴うカメラの外部パラメータの変動を自動補正するには、セルフキャリブレーション [76] が必要となる。移動の都度パラメータ補正をするコストを鑑みた場合、1台のカメラを移動し、任意に選択した複数の画像からカメラの外部パラメータを算出し、環境の3次元再構成（モデリング）をするアプローチとの比較検討も必要であろう。その際、カメラが1台で済む構成の簡便さはあるものの、任意に複数枚の画像を選択可能とするための画像蓄積保持が必要となる点、初期位置（移動前の撮影状態）では局所地図の生成ができない点等のトレードオフもあるため、用途に応じてシステム構成を検討する必要がある。しかしながら、第3章にて述べた、獲得された3次元セグメントのループから面を内挿するアルゴリズムや、本章にて述べた面ベースの環境の幾何モデルの逐次的統合・解釈のアルゴリズムは、カメラ構成が1台のアプローチに対しても、3次元セグメントのループが獲得された後にはそのまま適用することができる。

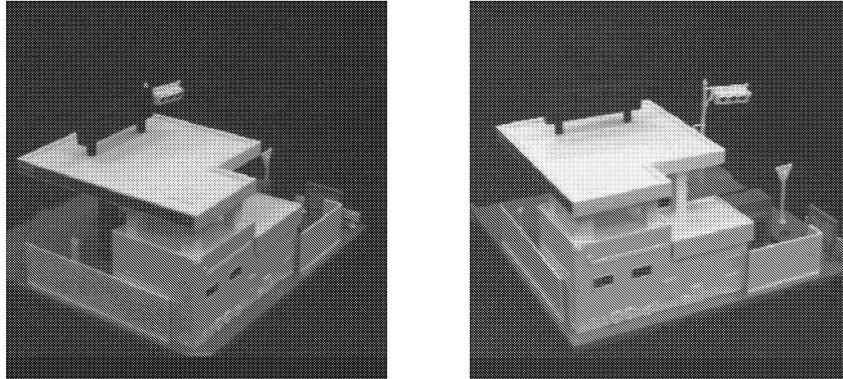
また、セグメントの欠けがなくすべてのループの解釈が可能となる完全なステレオデータを獲得することが難しい要因として、撮像された画像データの輝度分解能の粗さや、撮像対象までの距離と比較した画像の解像度による空間分解能の不足があげられる。

前者に関しては、本研究にて用いた撮像系では、手動による絞りの調整のみ行い、自動制御（アイリスコントロール）機能は用いておらず、輝度分解能は8bitである（AGC:オートゲインコントロール機能も用いていない）。輝度分解能に関しては、近年10bitに拡張されつつあり、検査用途では12bitのものも見られるようになってきた。このハードウェアの進歩に伴い、課題解決につながることに期待する。また、アイリ

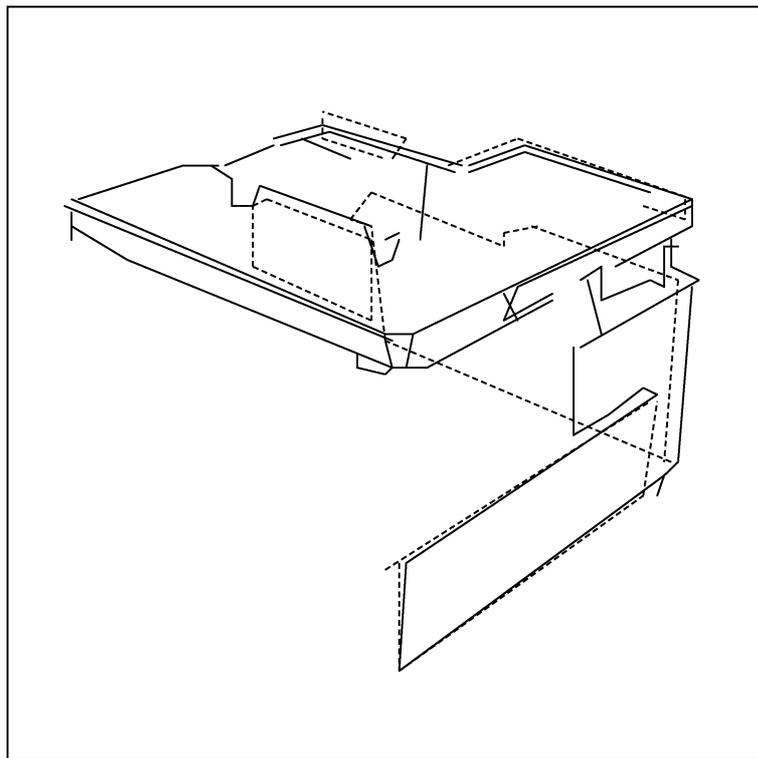
スの積極的な制御手法の導入による効果に関する検証についても行う必要がある。

後者に関しては、解釈不可能なループに対する注視点制御として、撮像系全体の移動・ズームやフォーカス制御、超解像画像生成技術 [97]~[109] の導入によるより詳細な画像情報の獲得も有効と考える。その際、それに伴うカメラパラメータ算出・当該ループ部分の幾何モデルの3次元復元ならびに世界地図との統合・再解釈等、本手法の有効性の更なる検証に向けた実証実験も合わせて行う必要がある。

更に、ステレオの誤対応により、実際には3次元空間中に存在しないセグメントが再構成されてしまう場合についても考慮する必要がある。誤対応が起きた場合には、面積0のループ（自身とそのwinged-edgeセグメントからなる孤立セグメント）として検出される場合が多い。孤立セグメントは、他のセグメントとの連結性が確保され、解釈不可能なループの解釈に貢献し統合されるまでは、解釈不可能なループ同様に扱われる。しかしながら、誤対応による場合にはそもそも存在してはならないセグメントであるから、何らかの方法により除去されなければならない。孤立セグメントが実際に面が存在する空間よりも視点側で観測される場合には、3次元空間中に針金が浮いているのと同じ状況と考えられ、面が内挿されるべきループに対して実在テストにより問題なく面の推定が行われる。しかしながら、誤対応セグメントが奥行き方向に延び、実際に存在する面に対して視点と反対側に位置するセグメントとして検出される場合には、孤立セグメントに対して視点側に存在する面については、実際に面が存在するにもかかわらず実在テストにより解釈不可能なループと判定されてしまい解釈に対して悪影響を及ぼす。それゆえ、孤立セグメントに関しては、誤対応による可能性も鑑み、解釈不可能なループと同様、視点の移動による検証が必要である。そのために、誤対応による孤立セグメントである可能性について、観測されるべき視点からどれだけ観測されたかといった観測回数の指標の導入等、対応の不確かさについての検証も有効と考える。

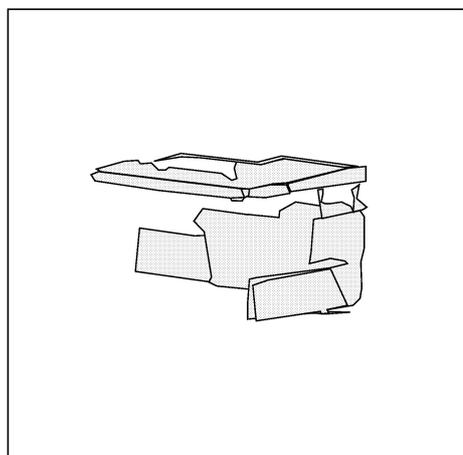


(a) 第3の視点からのステレオ原画像

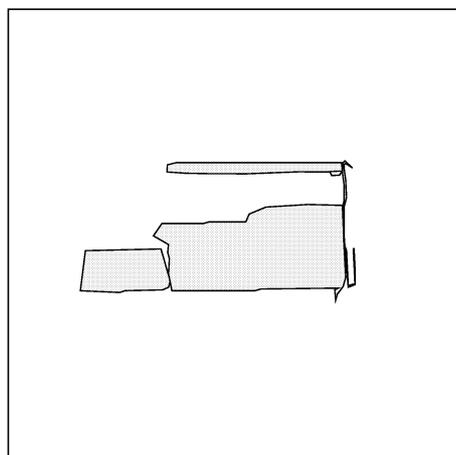


(b) 第3の視点での対応探索と統合

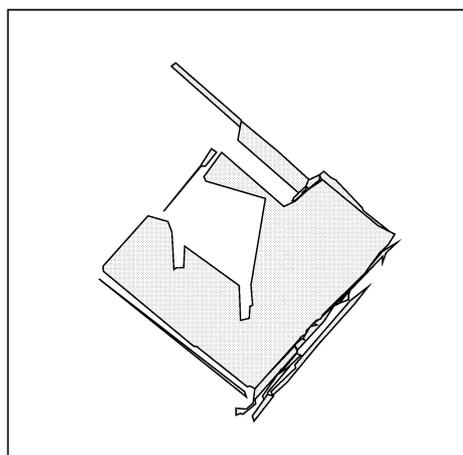
図 4.13: 第3の視点での局所地図との対応探索と統合



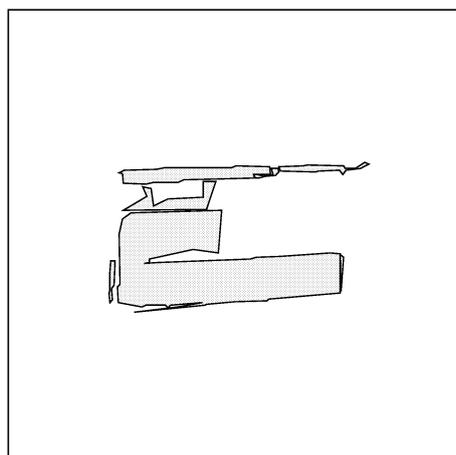
鳥瞰図



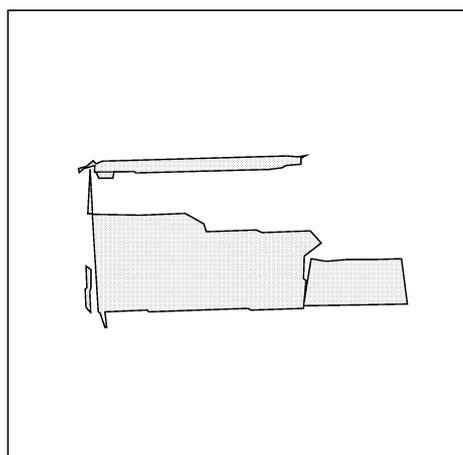
正面図



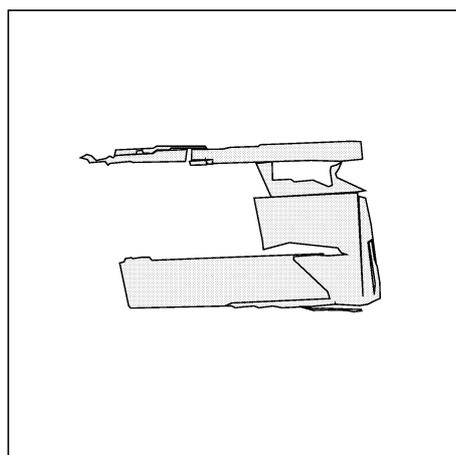
天頂図



左側面図



背面図



右側面図

図 4.14: エッジベースステレオによる面ベースの記述と統合・更新された世界地図

第5章 物体認識処理の高速化

5.1 具体的課題と関連研究ならびに概要

任意の3次元物体の認識は、コンピュータビジョンの当初からの重要な課題の一つである。モデルに対応する物体を任意のシーン中より検出し、その位置姿勢を実時間で決定する問題は、生産ラインの自動化のみならず、画像内容検索など幅広い応用が見込まれる。実際、特に顧客の製品ニーズの多様化に対応する多品種少量生産を必要とする多くの現場からその早期実現が求められている [116][117]。

本章では、処理の高速化に焦点を当て、単なる処理のハード化だけでなく、高速な物体認識を実現するための手段として並列分散処理が有効であると考え、セグメントベースステレオにより復元されたシーン中の物体の境界線に関する3次元情報から、物体の位置姿勢を認識する処理を、ネットワーク型並列計算機環境で高速化する方法について述べる。

後述するように、並列化を行う認識アルゴリズムにおいて実行時間の90%を占める処理は、シーン中の不特定多数の物体とモデルの対応候補の妥当性を検証する処理である。この処理は、各対応候補に対して独立に行うことができるため、並列化に適した処理であるといえる。また、認識対象とするモデルやシーンの複雑さによって異なるが、対応候補の数は通常数万にもなるため、これらを並列に処理することによって、並列度(=対応候補数)に応じた性能向上を得ることができれば、実用化可能な数秒から1秒以内の実時間認識を実現することも可能であるといえる。

従来から並列画像処理に関する研究は数多く行われている。しかし、並列化が行われているのは、画像から線や領域などの特徴を抽出する画像解析処理が多く [118][119]、抽出した特徴からモデルや知識を用いて認識を行う画像認識処理における並列処理の研究はあまり行われていない。また、画像認識処理における並列化の手法も、対象となる画像を分割してプロセッサに割り当てる単純なデータ分散処理が主である。Rygolらはトランスピュータを用いた並列計算機上に3次元ビジョンシステムを実装している。これは、画像データを水平方向に分割してプロセッサに割り当て、並列に処理を行うものである [120]。Modayurらは画像中の直線とモデルを構成する直線の組合せによって得られる対応候補をプロセッサに分散してその妥当性を並列に検証している。実装している環境は1024プロセッサの並列計算機である [121]。

これに対して、実際の応用を考えた場合にコストや拡張性の観点から、専用の並列

計算機ではなくパソコンやワークステーションによる並列計算機環境が有効であると考へ、システムを実装する計算機環境としてワークステーションクラスタを選択した。このような環境では、性能の異なる計算機が混在することが想定され、システムがこれに対応できる必要がある。そこで、構成する計算資源の性能差を吸収するための動的な負荷分散機能を実装し、性能の異なるワークステーションクラスタ上で台数効果を得るための、マスタ/スレーブによるタスクプール方式を用いた動的負荷分散を行い、モデル照合処理の並列処理性能を評価し、3倍以上の性能差のあるプロセッサが混在した環境において検証した有効性について述べる。

5.2 エッジベースステレオによる物体認識

エッジベースステレオによって復元したシーンの輪郭線に関する3次元情報と、あらかじめ構築しておいた物体の幾何モデルとを照合し、その物体の位置と姿勢を決定する。

ここで、認識アルゴリズムは、観測データ中の特徴点と幾何モデルの特徴点との対応関係から物体の局所的な特徴による対応候補を作成する「初期照合」と、各対応候補について物体の全体的な特徴を用いて詳細に照合を行う「微調整」の2段階で実行される(図5.1)。

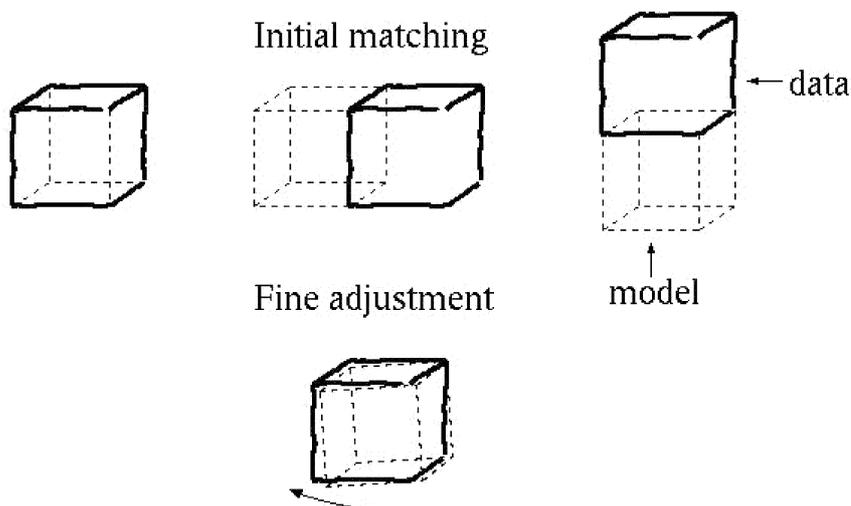


図 5.1: 初期照合と微調整

初期照合処理は、モデルと観測データから対応する特徴点の組を抽出し、各特徴点对から物体の位置姿勢を求める処理である。ただし、特徴点の正しい対応を前もって

知ることはできないので、可能性が高いと推定されるすべての組合せを対応候補とする。モデル数を M 、モデル i の頂点数を m_i 、データの頂点数を n とすると、調べるべき対応点の組合せは (5.1) 式で表せる。認識対象とするモデル数やデータ数の増加に伴い、対応候補とする組合せ数も増加する。これら対応候補の検証は独立に処理することが可能なため、実時間認識を実現するために並列処理が有効といえる。

$$\sum_i^M (m_i C_3 \cdot n C_3) \quad (5.1)$$

物体の位置姿勢は、モデルを観測データ中の3次元構造に一致させるように移動させる回転行列 \mathbf{R} と平行移動ベクトル \mathbf{t} により記述できる。以下に、頂点による初期照合について述べる (図 5.2)。

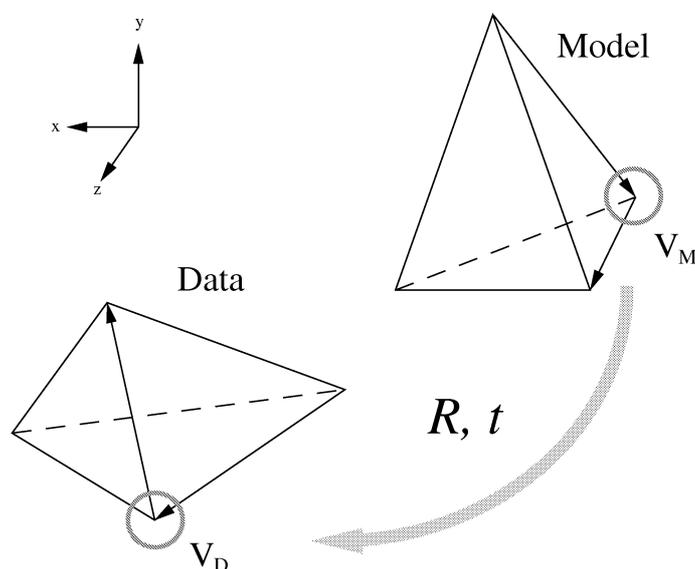


図 5.2: 頂点を用いた初期照合

まず、モデルの頂点 \mathbf{V}_M と観測データ中の頂点 \mathbf{V}_D を比較し、頂点を構成する二つのベクトルのなす角が大きく異なる場合は、正しい対応である可能性が低いと考えられるので候補から外す。次に、 \mathbf{V}_M と \mathbf{V}_D が一致するように物体の位置姿勢を決定する。頂点の位置から \mathbf{t} が得られ、各頂点を構成する二つのベクトルの方向から \mathbf{R} が得られる。これを、モデルと観測データ中のすべての頂点の組合せについて行う。ここで得られるのは、物体の局所的な特徴を用いた大まかな対応候補であり、図 5.1 に示すように正解以外のものも多く含まれる。

微調整処理は、初期照合の結果得られた対応候補の妥当性を判別すると同時に、物体の位置姿勢を示す \mathbf{R} と \mathbf{t} に含まれる誤差を小さくする処理である。まず、初期照合で得られた対応候補の位置姿勢に基づいてモデルのすべての点を移動し、観測データ中

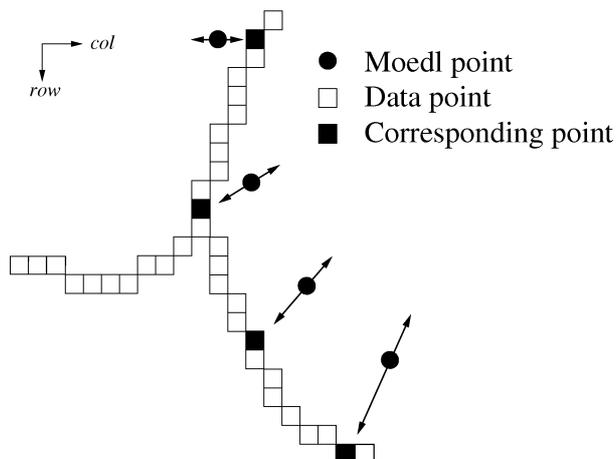


図 5.3: 対応点探索

でこれらに対応する点を探す(図 5.3)。ここで、対応点探索は、モデル点の法線方向で一定の距離範囲内にある点を対象とする。十分な数の対応点が見つからない候補は誤りであったとみなして棄却する。次に、見つかった対応点に一致するようにモデルを移動させる最適な R と t を最小二乗法により算出し、これを更新する。以上の処理を、モデルと観測データの平均二乗誤差が収束するまで繰り返す。すべての対応候補に対して上記微調整処理を行い、平均二乗誤差が最小のものを認識結果として出力する。

図 5.4 のモデルに相当する物体認識結果を図 5.5 に示す。また、この処理を Sun Ultra2(300MHz) で行った際の初期照合と微調整の実行時間、及び初期照合の結果得られる対応候補の数を表 5.1 に示す。

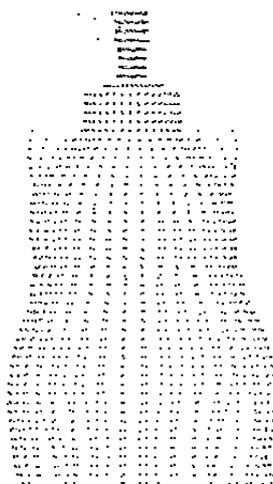


図 5.4: モデルデータ 1



図 5.5: 認識結果 1

表 5.1: 初期照合と微調整の実行時間と対応候補数

初期照合	14 秒
微調整	2 分 39 秒
候補数	15,531

この表からわかるように、処理量としては、初期照合に比べて微調整に要する処理時間がはるかに多く、全体の約 90%を占めている。ここで、微調整処理においては各対応候補に対して独立に処理が可能であるため、その並列化により処理の高速化を効率的に図ることができる。

5.3 微調整処理の並列化

微調整によって最終的な対応候補を求めるまでの処理は、初期照合によって得られたすべての対応候補に対して同じ処理が行われる。また、これらの処理はそれぞれ独立に実行できるため、対応候補の数だけ処理に並列性があることになる。そこで、一つの対応候補に対する微調整処理（以下、これをタスクと呼ぶ）を単位として並列化を行う。

処理の流れは以下ようになる。まず、1台の計算機において画像データの入力から初期照合までの処理を行う。次に、対応候補をすべての計算機に分散して配布し、微調整処理を並列に行う。各計算機で得られた結果は再度1台の計算機に集められ、最終的な認識結果として出力される。

ここで、並列処理の効果を得るために、各計算機に対してどのようにタスクを割り当てるかが問題となる。最も単純な方法は、すべての計算機に同じ数のタスクを割り当てるものである。しかし、タスクの内容が、照合処理がしきい値以下に収束するまで繰り返し処理を行うものであることから、タスクごとに必要とする処理量が異なることが予想されるため、すべての計算機が同時に処理を終了するとは限らない。この場合、処理を終了した計算機は他の計算機の終了を待つことになるので、十分な並列処理の効果が得られない。タスクごとに必要な処理量を事前に見積もることは不可能であるため、各計算機の処理時間が等しくなるようにタスクを割り当てることは困難である。また、計算機の能力が異なる場合には、処理能力の低い計算機が全体の処理時間を決定することになり、システム全体の処理性能を有効に使用することができない。そこで、このような静的なタスクの分配に起因する性能低下を避けるために、動的な負荷分散が必要となる。

そこで、各タスクの処理量の違いを吸収し、また計算能力が均等でない実行環境においても効率良く実行するために、マスタ/スレーブによるタスクプール方式を用いた動的負荷分散を導入した。これは、全計算機上で起動されたスレーブプロセスと一つのマスタプロセスによって実行される。マスタプロセスは、スレーブプロセスへのタスク割当てを、各計算機の負荷状況に応じて行うプロセスであり、スレーブプロセスは実際のタスク処理を行うプロセスである。

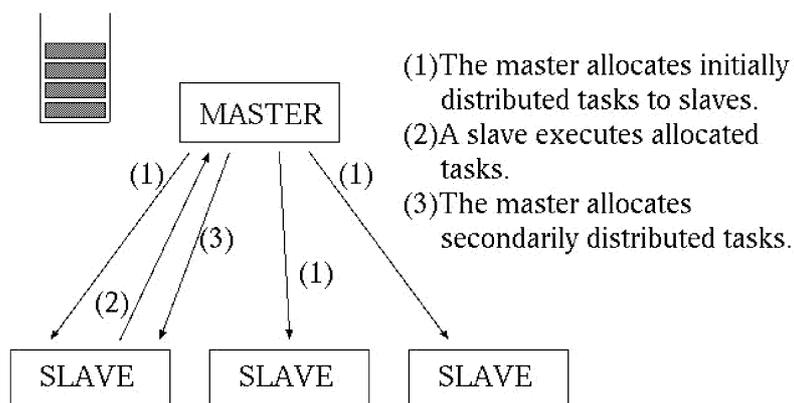


図 5.6: 動的負荷分散

マスタプロセスはすべてのタスクを一つのキューで管理し、スレーブプロセスに対してこれを供給する (図 5.6 の (1))。スレーブプロセスは割り当てられたタスクの処

理を行い、処理が終了すると、結果をマスタープロセスに返す（図 5.6 の (2)）。マスタープロセスは、このスレーブプロセスに対して、再度タスクの割り当てを行い（図 5.6 の (3)）、キューが空になるまで、これを繰り返す。

マスタープロセスがスレーブプロセスに対して、一度に割り当てるタスクの数は少ない方が粒度の細かい負荷分散が可能になる。一度に割り当てるタスクの数を 1 にすれば、最も速くすべての処理を終了する計算機と最も遅いものとの差は、タスク一つ分の処理時間以内になる。しかし、割り当てる際に通信などのオーバーヘッドが生じるため、必ずしも性能の向上につながるとは限らない。そこで、最初に分配するタスク（初期分配タスク）の数と、2 回目以降に割り当てるタスク（再分配タスク）の数を各々設定できるようにし、両者を動的負荷分散のパラメータとして用いることとした。

5.4 並列性能評価

図 5.5 の画像に対して、図 5.4 のモデルに相当する物体の認識を行い、モデル照合処理の並列性能の評価を行った。評価実験は、図 5.7 のようなイーサネットで接続された 5 台のワークステーションによって構成されるワークステーションクラスタ上で行った。プロセスの起動とプロセス間の通信には PVM[122] を用い、ワークステーションクラスタを構成する各ワークステーションがマルチプロセッサ構成であるため、プロセッサ数は全部で 12 である。

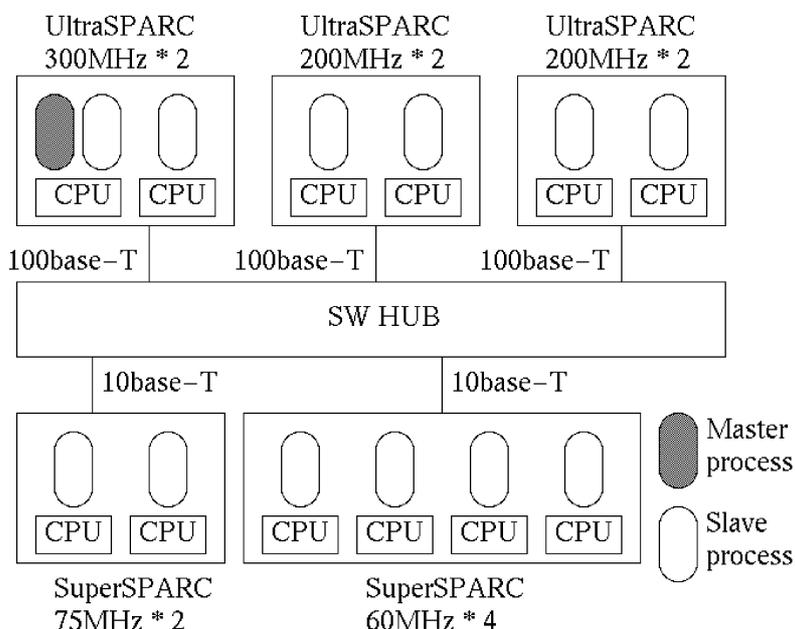


図 5.7: ワークステーションクラスタ

5.4.1 プロセッサの性能比

図 5.7 に示すように、ワークステーションクラスタを構成するプロセッサは4種類であり、各々の性能に差があるため、図 5.5 の画像における物体認識を逐次行うことによって、各プロセッサの性能比を調べた。

図 5.8 に、各プロセッサの性能比を示す。最も高速なプロセッサの性能を1とすると、他のプロセッサの性能は、0.65, 0.31, 0.27となる。ワークステーションクラスタは、性能1のプロセッサを二つ、性能0.65のプロセッサを四つ、性能0.31のプロセッサを二つ、性能0.27のプロセッサを四つで構成されているので、全体の性能は最も高速なプロセッサ性能の6.3倍ということになる。

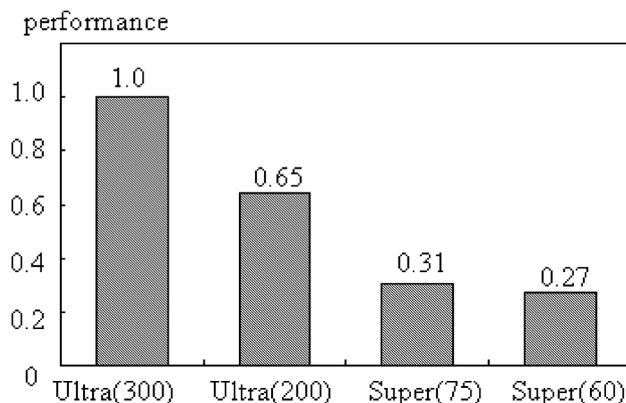


図 5.8: プロセッサの性能比

5.4.2 分配タスク数による性能評価

前述したように、最初に各スレーブプロセスに与える初期分配タスクの量と、残りのタスクを順次割り当てる再分配タスクの数がパラメータとなる。これらパラメータが性能にどのように影響するかを以下に述べる。

5.4.2.1 初期分配タスク数による性能評価

図 5.9 は、図 5.5 の画像における物体認識において、初期分配タスクの数を変えて並列処理を行い、その実行時間を示したものである。横軸は各プロセッサに対して割り当てられた初期分配タスク数（括弧内は初期分配タスクの総数とタスクの総数の比）を示し、縦軸は認識に要した時間を示す。使用したプロセッサ数は12であり、再分配タスク数は32である。

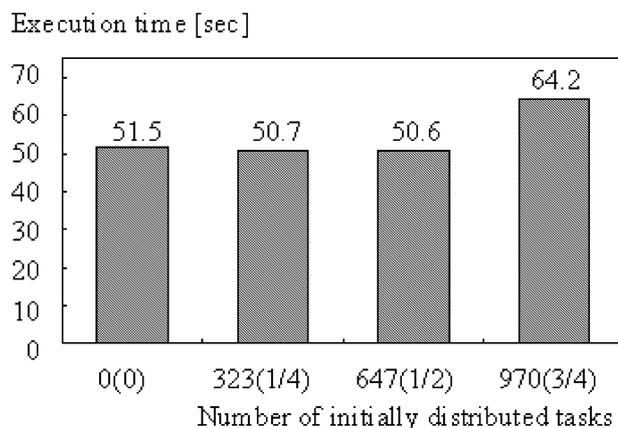


図 5.9: 初期分配タスク量と実行時間の関係

初期分配タスク数としては、0の場合と、全タスクの1/4、1/2、3/4の場合の4種類について実験を行った。0の場合、初期分配は行われず、すべてのタスクは再分配タスクとして32ずつプロセッサに割り当てられる。1/4の場合は、全タスク（15,531）の1/4である3,876（すべてのプロセッサに同じ数のタスクを割り当てるために12の倍数にしている）が初期分配タスクとなる。各プロセッサに初期分配タスクとして割り当てられるのは、その1/12である323となる。1/2、3/4の場合も同様である。

初期分配タスク数が970の場合に、大きく性能が劣化している。これは、初期分配タスクの数が多すぎるために、すべてのタスクの割り当てが終了し、高速なプロセッサでのタスク実行が終了してアイドル状態になった後も、最も低速なプロセッサでは初期分配タスクの実行がまだ終了していないという現象が生じていると考えられる。これに対して、初期分配タスク数が0、323、647の性能は良いが、その中でも0の場合が他よりも少し性能が悪くなっている。これは、すべてのタスク（15,531）を32という小さな単位で割り当てるため、割り当てのための通信が増えることが原因であると考えられる。結果として、初期分配タスク数は、323、647の場合に良い結果が得られていると言える。

5.4.2.2 再分配タスク数による性能評価

図 5.10 は、前節と同一の物体認識処理において、再分配タスクの数を変えて並列処理を行い、その実行時間を示したものである。横軸は再分配タスクの数、縦軸は認識に要した時間を示している。使用したプロセッサは12、初期分配タスク数は全タスクの1/4とした。

再分配タスク数は32の場合に最も良好な性能が得られ、それ以上でもそれ以下でも性能が劣化するという結果となった。再分配タスク数が少ない場合は、処理が終了

するまでに必要となるタスクの分配回数が増え、分配のために生じる通信のオーバーヘッドが大きくなったために性能が劣化していると考えられる。また、再分配タスク数が多い場合には、負荷分散の粒度が粗いためにすべてのスレーブプロセスに負荷を均等に分配することができず、初期分配タスク数が $3/4$ の場合と同様の状況が発生しているものと考えられる。

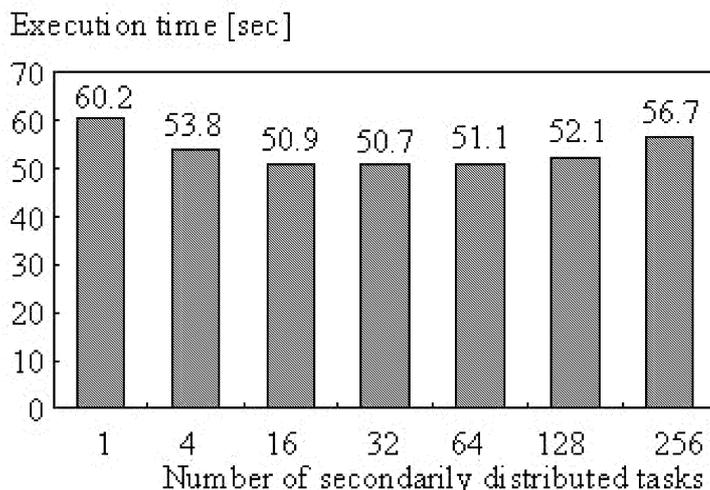


図 5.10: 再分配タスク数と実行時間の関係

5.4.3 プロセッサ数による性能評価

前節と同一の物体認識処理において、プロセッサの数を変えて性能評価を行った結果を図 5.11 に示す。横軸は使用したプロセッサ数、縦軸は認識に要した時間を示している。初期分配タスク数は全タスクの $1/4$ 、再分配タスク数は 32 とした。

評価実験では、最も高速なプロセッサから使用し、プロセッサ数を増やすに従って追加されるプロセッサは低速なものとしている。図 5.11 ではプロセッサ数が 6 を超えたあたりから性能向上の度合いが大きく低下しているように見えるが、これは使用しているプロセッサの平均性能が低下していることが影響している。1 プロセッサ使用時の実行時間が 171.7 秒で、12 プロセッサ使用時の実行時間が 50.7 秒である。5.4.1 で述べたように、12 プロセッサ全体の計算上の性能は、最も高速なプロセッサの性能の 6.3 倍である。したがって、ワークステーションクラスタの計算上の性能 6.3 に対して、実行時間における並列処理による性能向上は 3.4 倍という結果が得られた。

ワークステーションクラスタの計算上の性能が 6.3 であることから考えると、理想的な並列処理による実行時間は、1 プロセッサの場合の処理時間である 171.7 秒に対し ($171.7/6.3 =$) 27.3 秒となるが、実際には 50.7 秒であり、23.4 秒の差が生じた。こ

れには二つの原因があると考えられる。認識処理中に逐次処理が含まれることと、スレーブプロセスの起動及びマスタ/スレーブ間の通信オーバーヘッドが発生することである。5.2で述べたように、認識処理全体に対し並列化した部分は90%であり、残りの10%は逐次処理である。このため、171.7秒の10%である約17秒の処理は並列に実行されない。90%の部分が理想的に並列実行されると41.7秒*になることが予想され、(50.7-41.7)=9秒の差は、スレーブプロセスの起動及びマスタ/スレーブ間の通信に要した時間であると考えられる。

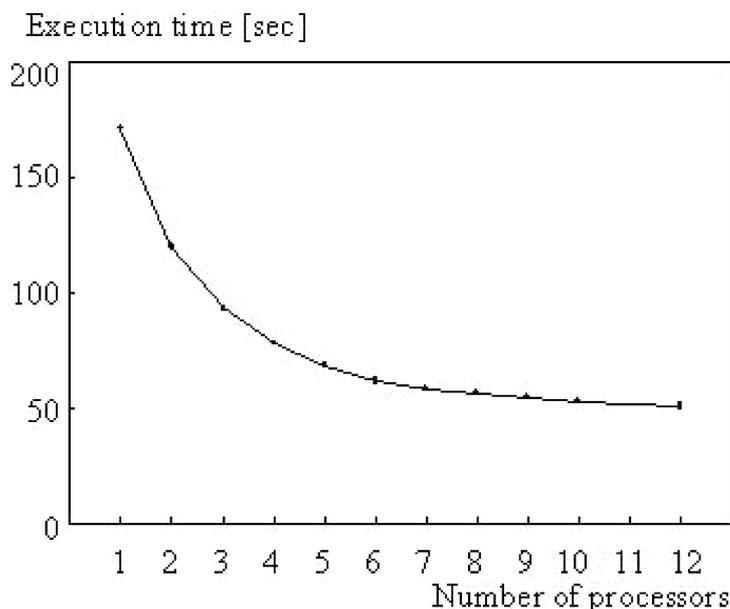


図 5.11: プロセッサ数に対する実行時間

図 5.11 の結果から、台数効果を求めたのが図 5.12 である。横軸が使用したプロセッサの全体性能、縦軸が台数効果であり、実験結果を実線で、理想曲線を点線で示している。ここで、並列化を行ったのは認識処理の90%であることから、理想曲線は台数効果を E 、使用したプロセッサの全体性能を N とすると、 $E = \frac{10N}{9+N}$ † となる。

プロセッサの性能が2以上の部分は並列処理が行われている部分であるので、通信などのオーバーヘッドによって実験結果が理論値を下回っているが、使用するプロセッサの数が多くなり、プロセッサ性能の合計が大きくなっても、実験結果と理論値の差は大きく開くことはない。これは、5.3で述べた初期分配タスクと再分配タスクの導入により、使用するプロセッサの数が増えても通信オーバーヘッドの増加を抑えることができているためであると考えられる。

* $(171.7 \cdot 0.9) / 6.3 + 171.7 \cdot 0.1 = 41.7$

† $E = \frac{1}{\frac{1}{10} + \frac{1}{10 \cdot N}} = \frac{10N}{9+N}$

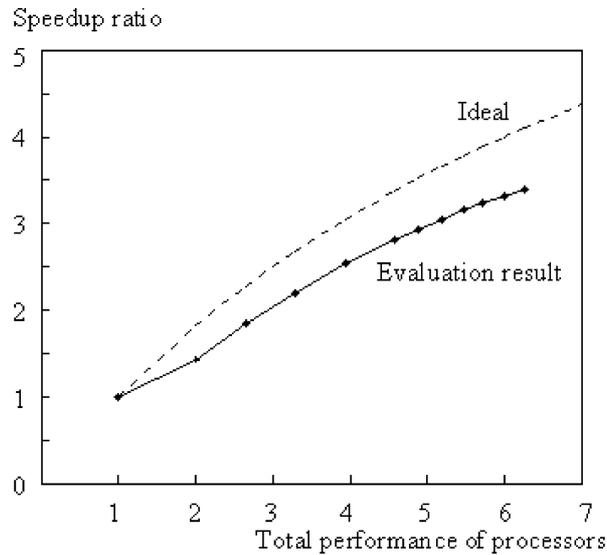


図 5.12: 台数効果

5.4.4 動的負荷分散の効果

動的負荷分散の効果を見るために、図 5.11 の 12 プロセッサ使用時に各プロセッサで処理されているタスク数を図 5.13 に示す。

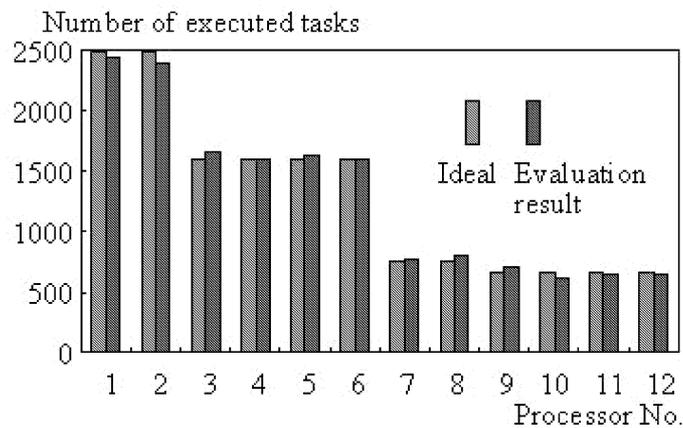


図 5.13: 各プロセッサにおいて処理されたタスク数

横軸はプロセッサの番号であり、高速なものから順に、プロセッサの性能比から算出した理論値と実験値を表示している。多少のばらつきはあるものの、理論値と実験値の差は0.1~7.5%であり、動的負荷分散によって、各プロセッサの性能に応じた数のタスクが割り当てられていることがわかる。したがって、3倍以上の性能差のある

プロセッサが混在した環境においても、システム全体の性能を引き出すことができているといえる。

5.4.5 他のモデルとシーンにおける実験

平面のみで構成されるモデル（図 5.14）と、比較的単純な自由曲面体によって構成されるモデル（図 5.15）の認識例について、使用したプロセッサ数と認識処理時間の関係を表 5.2 に示す。

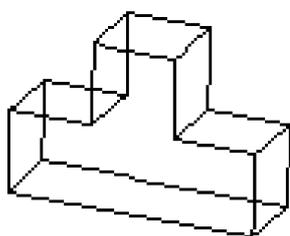


図 5.14: モデルデータ 2



図 5.15: モデルデータ 3

表 5.2: 使用プロセッサ数と認識処理時間
(初期分配タスク数:全タスクの 1/4, 再分配タスク数:8)

使用プロセッサ数	1	6	12
図 5.16 の認識処理時間 (秒)	7.3	3.8	4.4
図 5.17 の認識処理時間 (秒)	11.1	4.6	4.4

図 5.16 の認識の場合は、1 プロセッサ使用時が 7.3 秒、6 プロセッサ使用時が 3.8 秒、12 プロセッサ使用時が 4.4 秒であった。

同様に、図 5.17 の認識に要した時間は、1 プロセッサ使用時が 11.1 秒であった。初期分配タスク数を全タスクの 1/4、再分配タスク数を 8 とし、UltraSPARC のみ 6 プロセッサで並列処理を行った場合が 4.6 秒、全 12 プロセッサを用いると 4.4 秒という結果となった。

並列処理によって 2 倍から 2.5 倍の性能向上が得られた。しかしながら、図 5.4 のような複雑なモデルの認識時に比べて、並列処理の効果は小さくなっている。また、12 プロセッサ使用時の実行時間が 6 プロセッサ使用時の実行時間とあまり変わらない。若しくは上回っていることから、低速なプロセッサを使用することで全体の性能に悪影響を与えていることがわかる。

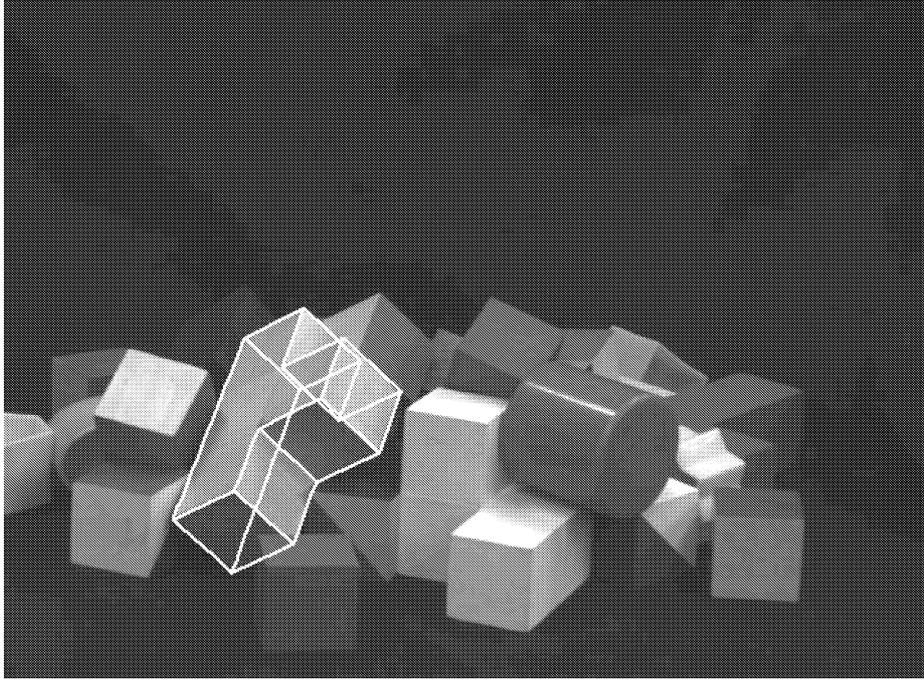


図 5.16: 認識結果 2

処理の並列度を示す対応候補の数は、図 5.14 のモデル認識時が 749、図 5.15 のモデル認識時が 963 であった。図 5.4 のモデル認識時の 15,531 に比べて並列度が 1/16 から 1/20 であるために、プロセス間の通信などの影響によって低速なプロセッサがシステムの性能に寄与しないという結果となっていると思われる。

5.5 第5章のまとめ

3次元物体認識の並列処理による高速化について述べた。エッジベースステレオによるシーン中の物体認識において、シーンから抽出された物体の特徴点と、あらかじめ構築された物体の幾何モデルとの照合に並列処理を適用した。この物体認識では、特徴点の組合せから物体の位置姿勢の候補を選定し、その妥当性を検証するため、特徴点を多く含む物体の認識には多くの処理時間を必要とするが、各候補の照合処理は独立に実行できるため、並列処理による大幅な高速化が可能となる。

今回、初期分配タスク数による性能評価ならびに、再分配タスクによる性能評価と合わせ、プロセッサ数による性能評価ならびに動的負荷分散の効果に関する評価を行った。

初期分配タスク（最初にすべてのスレーブプロセスに分配するタスク）数が多く、高速な CPU でのタスク終了後、低速な CPU で処理が継続する場合に大きく性能が低

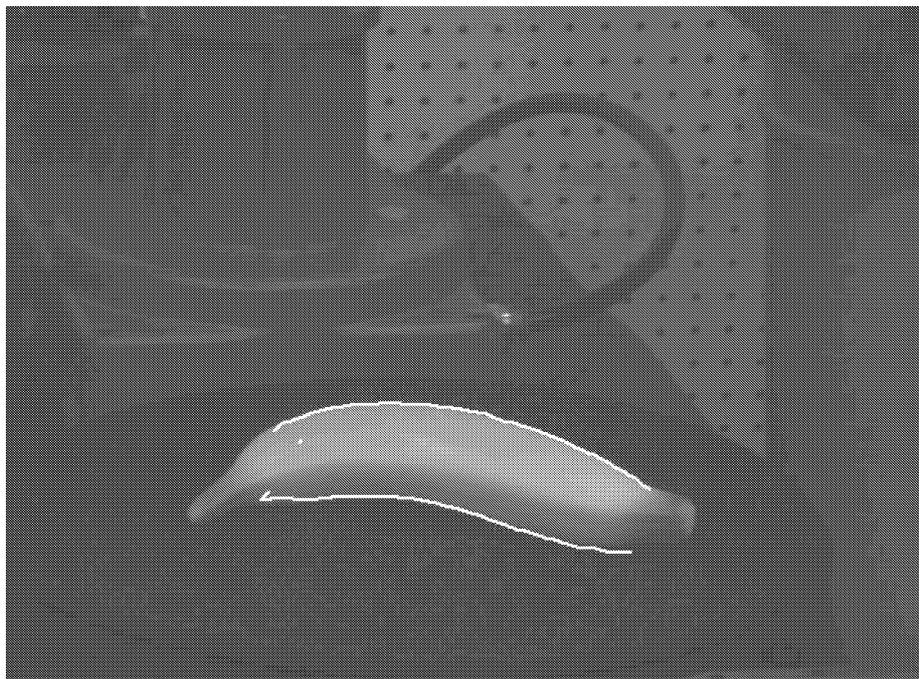


図 5.17: 認識結果 3

下する。また、再分配タスク（2回目以降に割り当てるタスク）数を小さくし過ぎる場合には、負荷分散粒度が粗くなり、通信オーバーヘッドの影響も受け、スレーブプロセスへの負荷が均等に配分できなくなる。そこで、性能の異なるワークステーションによって構成されるワークステーションクラスタ上で台数効果を得るために、マスター/スレーブによるタスクプール方式を用いた動的負荷分散を行い、モデル照合処理の並列処理性能を評価した。初期分配タスク数を大きな値に、再分配タスク数を小さな値に設定することによって、良好な並列性能を得ることができている。

今回並列化を行った部分はモデル照合処理全体の90%であり、更に並列性能を向上させるためには、残りの10%の逐次処理部分の並列化が必要である。また、高速ネットワークの利用による通信オーバーヘッドの削減や、使用プロセッサ数の増加による並列性能の向上も重要と考える。

ヘテロジニアスなワークステーションクラスタ（3倍以上の性能差のあるプロセッサが混在し、性能にばらつきがある計算機で構成されており、計算能力が均等でない）環境下にて並列処理することにより、12プロセッサ使用時（計算上の合計性能は最も高速な1プロセッサの性能の6.3倍）に逐次処理と比較して最大4.3倍の性能向上が得られ、動的負荷分散によってシステム全体の性能が引き出せることを示すとともに、その有効性が確認できた。

本研究により得られたタスクプール方式を用いた動的負荷分散による並列分散手法

は、ホモジニアスな並列計算機環境でも、通信負荷や各プロセッサエレメントの負荷のばらつきが発生する際には有効であり、PC クラスタ環境においても性能評価が待たれるところである。

また、分散コンピューティングプロジェクトにおいては、古くは暗号解読を目的とした **distributed.net** や、地球外生命体からの電波を探索する **SETI@home**[123] 等、インターネットを通じて家庭のパソコンの空いている CPU パワーを集め、暗号解読や医療研究などの複雑な処理を行わせるアプリケーションへの取組みも見られる。

近年、こうした技術を含むより包括的な概念としてグリッドコンピューティングと呼ばれるシステムの実現に向けてさまざまな試みが行なわれている。処理能力や記憶容量を増大させるためにネットワークを介して複数のコンピュータを接続し並列処理を行わせることで、一台一台の性能は低くとも仮想的に高性能コンピュータを作るというものである。高速に大量の処理を実行できるようになることで、大規模科学技術計算といった学術研究用途はもとより処理の高速性を重視したビジネス利用にも多くの可能性が期待されている。グリッドコンピューティングシステムにおいても、本手法の有効性が期待できるため、様々なアプリケーションによる性能評価を通じて有効性の実証を行いたい。

第6章 定性的物体認識へのアプローチ

6.1 具体的課題と関連研究ならびに概要

コンピュータビジョンの重要な課題の一つである3次元物体認識において、既知の物体を観測データ（画像）から見つけ出し（物体の3次元幾何モデルと対応する部分を探索し）、その位置姿勢を決定する座標変換を求める、同定（identification）に関する研究においてはトップダウンのモデル駆動型アプローチが一般的といえる [124]～[128]。しかしながら、実世界の認識対象すべてに対して詳細な幾何モデルを用意することは難しい。この所謂モデル爆発の問題は、制約のない実世界シーンの画像において、シーン中に含まれる物体を一般的な名称で表す際の同一クラスの範囲の広さや同一クラスに属する対象のアピランス変化による、対象の特徴抽出や認識のためのモデル構築ならびに学習データセットの構築が困難なため、解決するのが難しい課題となっている。人間は、数万種類の対象を認識可能 [129] であるのに対し、計算機にとっては、これまで1種類の対象ですらアピランスの大きな変化により認識が難しかった。また、画像と言語の対応付けを目指した一般物体認識に関する研究は古い [130] が、人間の正面顔の自動検出を除いて、実用的な精度で認識可能な対象がほとんどないのが現状である。

古くは線画の解釈 [131] から、分割した領域に対する緩和法に基くラベリングによる認識 [132]、シーン理解システム [133]、知識ベース型の画像理解システム [134][135] が登場したが、人手によるルールベースの手法であり、知識獲得のボトルネックという人工知能研究の課題を解決できなかった。その後、これら2次元的な物体認識手法は3次元情報の復元を重視する手法 [136][137] に移行し、model-basedの手法 [138] が多数提案された。また、物体の機能から認識を試みた function-based recognition [139] や context-based recognition [140]、画像エキスパートシステム [141]～[143] 等が提案されたが、これらもルールベースの手法であり一般化できなかった。その他、カラーインデクシング [144] による画像検索への応用 [145][146] や、固有空間法 [147] による appearance-based の手法も提案されているが、学習画像の用意という観点から対象の種類を増やすことは容易ではない。

近年、計算機の高速度大容量化や機械学習手法の進歩、part-based と呼ばれる新しい物体表現法 [149]～[153] の登場により、未知の物体をそれが属すべき既知のクラスに分類する認識手法（classification） [148] が進歩してきている。画像中の局所特徴の組合

せにより照合を行う手法 [154] は、画素値やその微分値を特徴ベクトルとし、高次特徴ベクトルを用いた点の集合による画像表現に基づく手法といえ、特徴点抽出とその記述法を組合せた SIFT(Scale Invariant Feature Transform)[155] の登場によりオクルージョンのあるシーンの認識が実現され、分類 (classification) においても有効性が示されている [156]. また、位置情報を用いずに多数の局所特徴量の集合 (数百個のオーダー) から対象を表現する Bag-of keypoints[157] も、部分的な情報を用いる part-based approach といえる. 高次自己相関特徴量 [158][159] と同様の着想であるが、確率的な応用が可能で、一般物体認識に多数応用されている [160]~[167]. この part-based approach においても、認識カテゴリーの多様化と同一クラス内の変化に対して、典型的ビュー [168] から作成したモデルのみを用いて対応できるようにすることが解決されるべき課題として残されている.

そこで、特徴点ベースの確率的アプローチとは異なり、観測画像中の対象領域の境界線の特徴点にて分割すること (第2章参照) によって得られるセグメント (直線や凹もしくは凸の曲線セグメント) を単位として、複数の基本形状特徴 (以下、形状プリミティブと呼ぶ) を一般的な関係 (例えば、平行である・凸角をなす・同心円状に配置する・滑らかな法線変化を伴い接するといった関係等) で組み合わせることにより、複雑な3次元物体の認識の際にしばしば起こる所謂モデル爆発問題の回避を念頭に、画像中の物体を定性的かつ階層的に記述するコンセプトとして”共通モデル”を提案する.

本章では、まず、この共通モデルのコンセプトについて紹介する. 次に、応用事例としてロゴマーク検索への適用例を通じ、画像中での最下層の形状プリミティブとして抽出したセグメントを単位とし、あらかじめ用意した形状プリミティブ間の一般的な関係を用いて対話的に教示することにより、ボトムアップかつ階層的に記述を生成するための教示システムを紹介するとともに、観測方向の違いに起因する画像中での対象の変形 (アピアランス変化) に対する頑健性を備えた記述について述べる.

しかしながら、人間が詳細な幾何モデルを用意することなく、実際に世の中の対象を認識しているのと同様の機能 (対象の概念的把握に基づく定性的な認識機能) を実現するために、どれだけの形状プリミティブやそれらの間の一般的な関係を用意すれば十分であるのかは明らかにできていない. そこで、人工物の多くが持つ特徴の一つである対称性に着目し、それらが観測された際に画像中に現れる特徴である斜め対称性 [172] を一般的な関係の一つとして取り上げ、上位の形状プリミティブとして画像中より斜め対称パターンとして抽出 [173][7] することで認識の際の有力な手がかりとすることができると考え、形状プリミティブとしての斜め対称プリミティブの検出方法を開発した. 斜め対称プリミティブの検出アルゴリズムについて述べるとともに、マンピュレータにより未知の対象物を操作する際の効率的な把持姿勢計画への応用事例 [8] と合わせて紹介する.

更に, **appearance** の違いへの対応のみならず, クラス内変化への対応可能性の検証として, 個体差のある人の手指を認識対象として記述を生成し, 剛体運動追跡への応用 [9] 例と合わせ, 共通モデルを用いた認識処理の有効性を示す.

6.2 共通モデル

実世界においては, モデルベースのトップダウン的物体認識のアプローチ [169][170] のように, すべての物体に対して詳細な幾何モデルを用意することは難しい. 一方, 我々が実世界で車を認識するといった状況では, 車の詳細な幾何モデルを必要とするのではなく, 車の存在を知るだけで十分である場合が多い.

認知科学の見地から, 対象物体をそれらに共通する基本形状の組合せによって階層的に構成する定性的アプローチ [129] が提案されている. 素立体より上位のレベルでの概念ではあり, コンセプトとして賛同できるものであるが, 画像から抽出されるパターンから素立体までの間を埋める, パターン認識技術との融合が待たれる. 反対に, 観測データよりボトムアップ的にパターン統合する技術として, 画像から基本形状パターンをボトムアップ的に組み合わせて記述を生成する知覚統合に関する手法 [171] が提案されている.

そこで, 両者を包含するコンセプトとして共通モデルを提案する. 図 6.1 に示すように, 画像中の直線もしくは曲線セグメントを単位として, それら複数の形状プリミティブを, 「平行である」, 「凸角をなす」といった一般的な関係で組合せることにより, 認識のためのモデルを定性的かつ階層的に記述する. あるクラスの物体は同じ記述により表現され, 認識の際には, 画像からボトムアップ並列的な特徴抽出により認識される.

ここで, 形状プリミティブは, 構造体—素立体—面—辺といった階層構造をもつ. 例えば, 同図中に示すようなポットを考えてみる. ポット (構造体) の外形は, 円柱 (素立体) や円錐台 (素立体) からなる. 更に円柱は, 楕円 (面) や円柱側面 (面) からなる. 円柱側面は, 平行線 (辺) とコニック (辺) より構成されるといった具合である. これらはすべて形状プリミティブであり, これらプリミティブは, 絶対的な寸法を考慮することなく定性的に表現され, 下位のプリミティブは, 上位のプリミティブによって共有される. 最下層のプリミティブは, 画像より抽出された直線セグメントと凹もしくは凸の曲線セグメントである.

形状プリミティブを用いて, 階層的にボトムアップに記述を生成することにより, 観測データの一部が, 照明条件やオクルージョン等により欠落していても, また, 物体表面に模様や影・雑音があっても, 観測方向に対して不変的な形状特徴の部分照合により, 対応箇所を抽出できる頑健性を備えることができる. また, 拘束の大きな特徴 (より上位階層の形状プリミティブ) を効果的に記述し, それらに基づいて初期照

合を行うことで（トップダウン的探索との組合せともいえるため），効率的かつ高速な探索が可能となる．

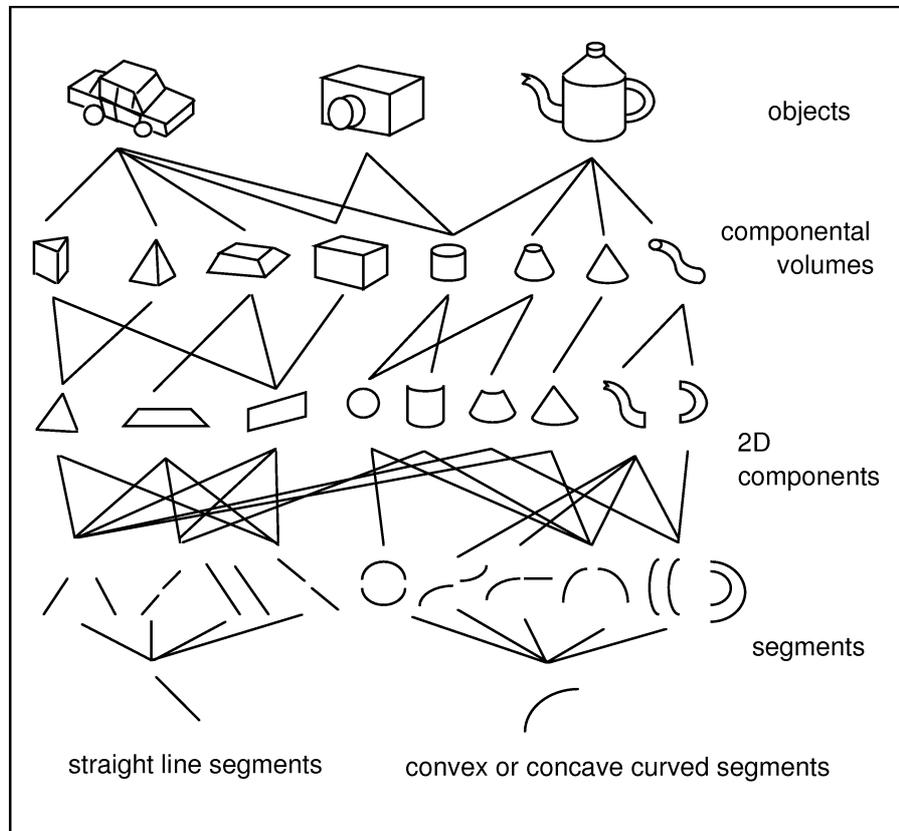


図 6.1: 共通モデル

6.3 ロゴマーク検出への応用

物体を定性的に認識するためには，包括的に物体の概念を表現できるモデルが重要となる．例えば，平行四辺形のモデルは，2組の平行線と4組の凸角の組み合わせにより記述されるといった具合である．そこで，定性的かつ簡素な記述を生成するための教示システムを構築するとともに，ロゴマーク検索を対象として，観測方向の違いに起因する画像中での対象の変形に対する認識の頑健性に関して検証実験を行った．認識に用いるモデルは，形状プリミティブとそれらの一般的な関係を表すリストによって記述されている．認識の際には，画像より，この記述に適合する直線および曲線の組み合わせを抽出することによって行なわれる．モデルの記述には，物体のサイズといった定量的な情報は含まれないため，隠れのみならず変形にも頑健な認識が可能となる．

6.3.1 教示システム

図 6.2 に、教示システム例を示す。ロゴマークの一部（”Y”の文字部分）についてのモデル作成例である。画像よりロゴマークの領域輪郭線が抽出され、9本の直線セグメントに分割されている。セグメント間の関係として、図 6.3 に示す関係を用意した。図 6.2 の例では、凸角をなす（angle）、平行である（parallel）、一直線上に乗る（colinear）といった関係により、”Y”の文字に関する記述が生成されている。同図中の文字の境界線セグメントは、左上より時計回りに、0番から8番まで番号が与えられ、0番と1番が右回りの angle、0番と3番が colinear、といった具合に、抽出結果に対して、AND もしくは OR により関連付けができる。そして、これらの関係付けされたプリミティブは、1次プリミティブとして登録される。更に、これら1次プリミティブを複数組み合わせることで、2次プリミティブを定義することができる。同様に、2次プリミティブを複数組み合わせることにより、ユーザ定義のプリミティブを階層的に教示できるように設計されている。

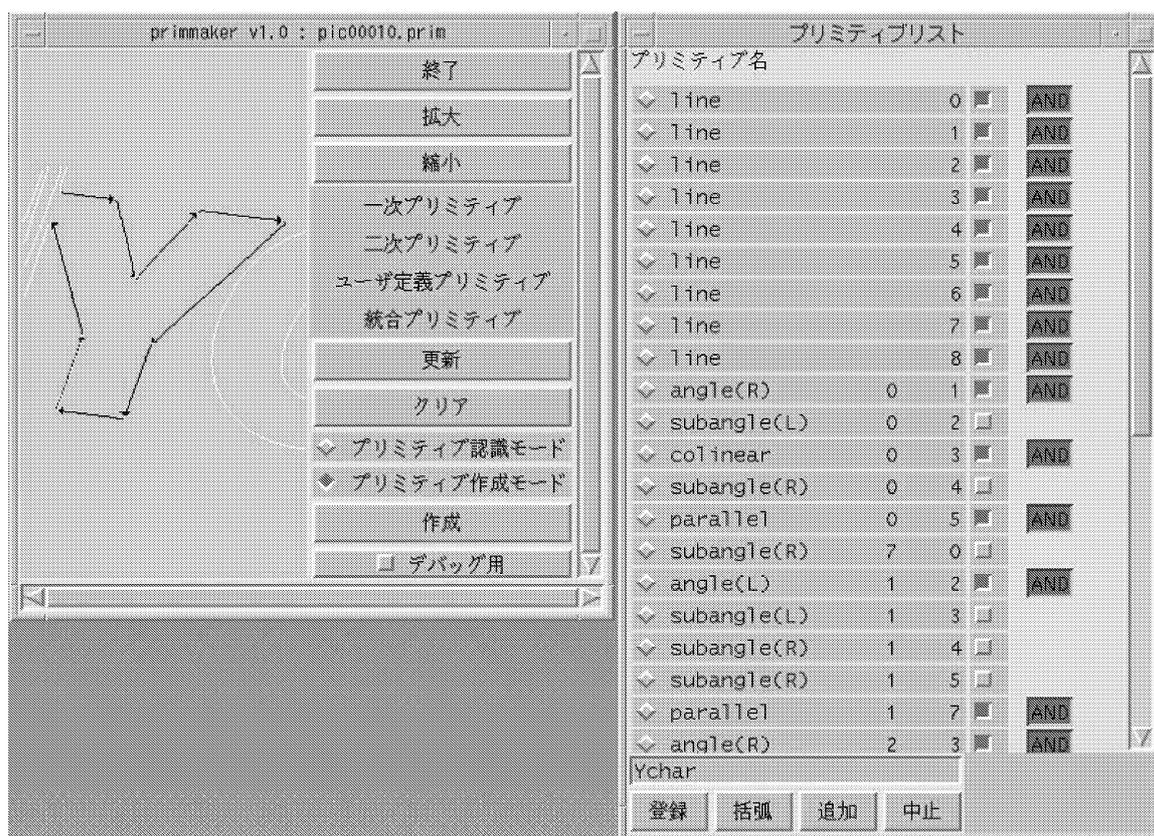


図 6.2: 教示システム

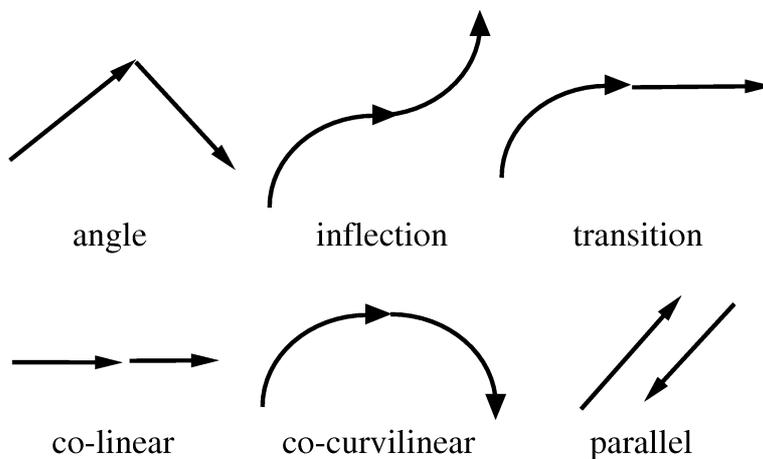


図 6.3: 一般的な関係

6.3.2 実験

図 6.2 の教示システム例の画像より生成した ”Y” の文字に関する記述に基づき、異なる角度から撮影した画像からの認識結果を図 6.4 に示す。さまざまな角度より撮影された画像中の変形の大きな対象についても検出ができていていることから、アピアランスの変化に頑健な記述が生成できているといえる。

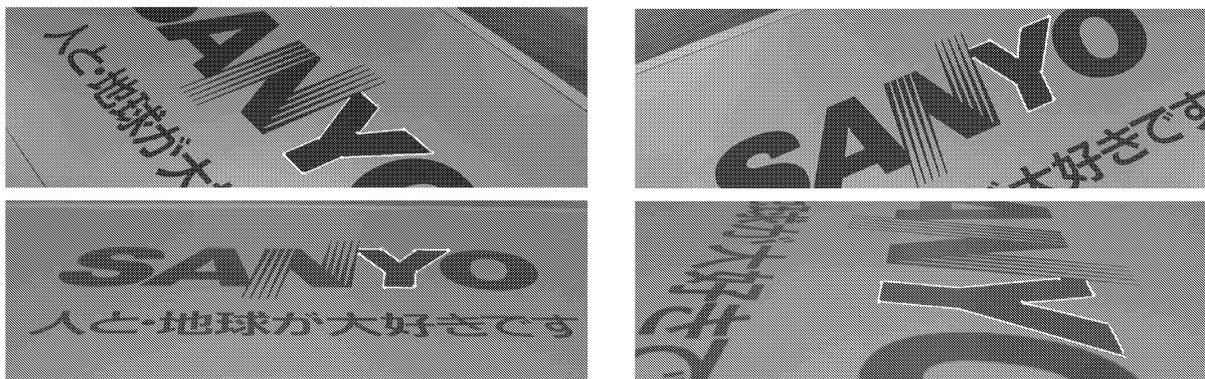


図 6.4: ロゴマーク認識結果 (アピアランス変化への対応)

また、広告宣伝効果の自動集計システムへの応用を念頭に置き、画像シーケンスからのロゴマークを含む対象フレームの抽出へも適用した (図 6.5 参照)。

同図 (a) に ”S” のプリミティブの記述例を示す。図 6.2 に示した教示システムを用いて、”A”, ”O” の各文字についても同様にプリミティブを記述し、更に、”S”, ”A”, ”Y”, ”O” の文字からなるプリミティブを ”co-linear” の関係によって統合したユーザ定義プリミティブを教示した。前記ユーザ定義プリミティブの検出処理を

画像シーケンスの全フレームに対して適用する (同図 (b)). 各フレームのエッジ画像 (同図 (c) 参照) より同図 (d) に示ように前記ユーザ定義プリミティブが検出されたフレームを計数する. 同図 (b) 最前部に, 前記ユーザ定義プリミティブを含むフレームを示す.

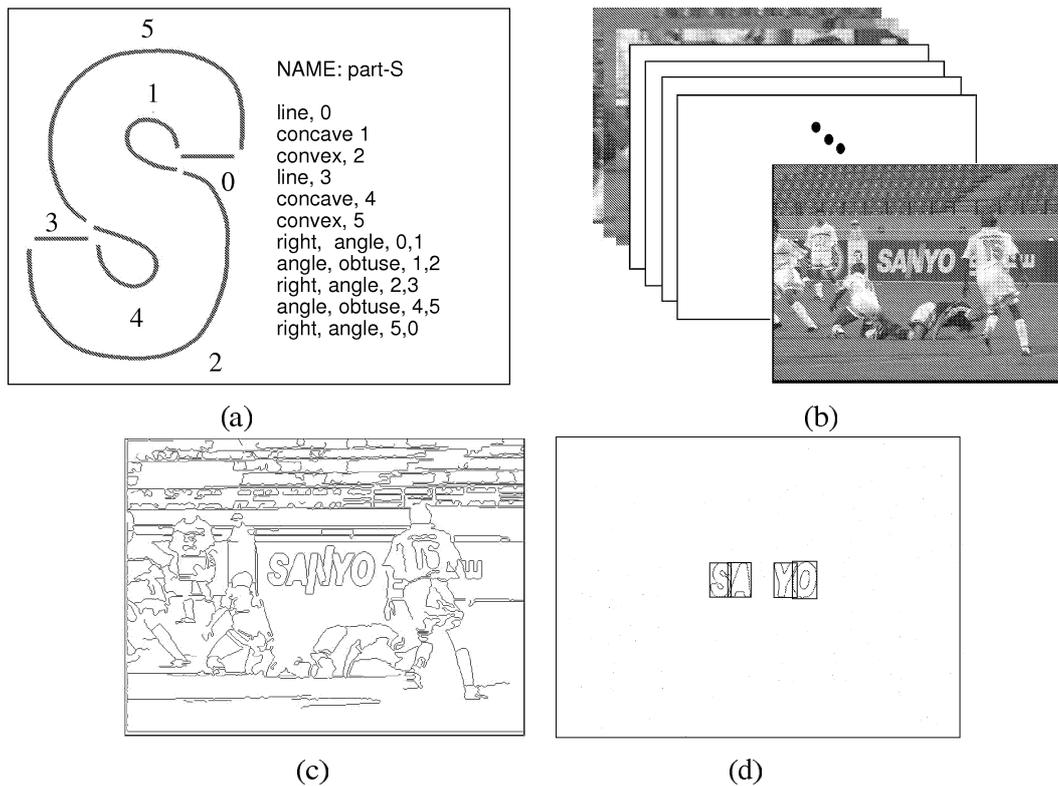


図 6.5: 文字”S”の記述例と画像シーケンスからの対象フレームの抽出例

6.4 斜め対称パターンの検出

空間中の対称パターンを直交射影により画像面に投影した図形は, 斜め対称の性質をもつ. 画像中の斜め対称パターンは, 3次元空間中の面の向きに拘束を与えることから, 物体認識において有力な手掛りとなることが知られている [172]. また, 人工物の多くが対称パターンを有することからも, 画像からの斜め対称パターンの検出は重要といえる.

対称パターンの輪郭線の持つ, 対象軸の両側においてその曲率や長さ, モーメントが等しいといった性質は, 斜め対称パターンにはそのままあてはまらないため, 対称パターン抽出アルゴリズムをそのまま斜め対称パターン抽出に適用することはできない.

対称パターンのモーメント特徴から、斜め対称軸に対する両側の面積が等しくなるという性質を用いた検出手法 [174] も提案されているが、オクルージョンを含むパターンに適用することはできない。また、斜め対称軸の両側に位置する対応点を結ぶ斜め横軸の角度が等しくなる性質を利用して、すべてのセグメントの両端点の組合せから、Hough 変換を用いて斜め横軸および斜め対称軸を決定する手法 [173] も提案されている。オクルージョンは克服できているが、輪郭線を直線セグメントにより多角形近似してしまっているために、検出精度に課題が残っていた。

そこで、上位階層のプリミティブとして斜め対称パターンを取り上げ、その検出手法を提案した [7]。直線セグメントに加え、凹もしくは凸の曲線セグメントも単位とすることで、多角形パターンのみならず曲線パターンに対しても、オクルージョンに頑健な検出を実現している。以下にその検出手順を述べる。

まず、入力画像より、第2章にて述べた境界線分割手法により、領域の輪郭境界線セグメントを得る。得られたセグメントを単位として、セグメントもしくはセグメント対からなる、以下の4種類の構成要素（局所的斜め対称パターンと呼ぶ）より、斜め対称軸、斜め横軸の候補を求め、構成要素のセグメント、斜め対称軸と斜め横軸を斜め対称構成要素として検出する。検出された複数の斜め対称構成要素より、それらの斜め対称軸・斜め横軸が等しいものを統合することにより斜め軸対称パターンとして抽出する。更に、斜め軸対称パターンを統合することにより、面对称パターンとして抽出する。

- 2対の直線セグメント
- 1対の直線セグメントと1直線セグメント
- 1対のコニック曲線セグメント
- 1つの非コニック曲線セグメント

セグメント対は、図 6.6 に示すようなT交差をしないものを対象とする。直線セグメントの場合には、セグメントが乗る直線を、曲線セグメントの場合には、両端点を通る直線をT交差検査用直線とし、両端点で挟まれる区間をT交差検査用線分とする。候補とするセグメント対のT交差検査用直線の交点が、いずれのT交差検査用線分上にも乗らないもののみを候補リストに追加し、T交差する組合せは除外する。また、曲線セグメントについては、コニックを最小二乗近似した際の誤差に基づき、コニックと非コニックのクラスに分類する。

6.4.1 候補軸の検出 1 (2対の直線セグメントより)

まず、2対の直線セグメントより候補軸を求める手順を以下に示す。

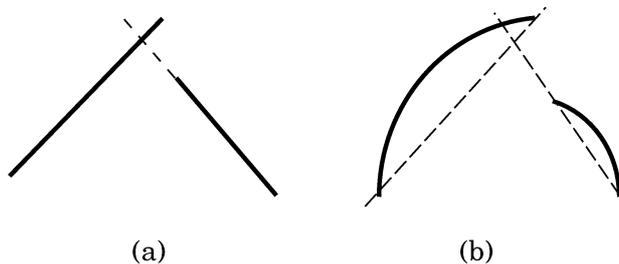


図 6.6: T 交差

1. 図6.7に示すようなT交差をしない2対の直線セグメント, (S_1, S_2) および (S_3, S_4) を選ぶ. ここで, C_1 および C_2 をそれぞれ (S_1, S_2) と (S_3, S_4) の交点とし, C_3 および C_4 を同様に (S_1, S_4) と (S_2, S_3) の交点とする.
2. C_1 と C_2 を通る直線 l_s が $\angle C_3C_1C_4$ と $\angle C_3C_2C_4$ の範囲内にあり, C_3C_4 の中点が l_s 上にある場合, l_s を斜め対称軸に, C_3C_4 を通る l_t を斜め横軸とする.

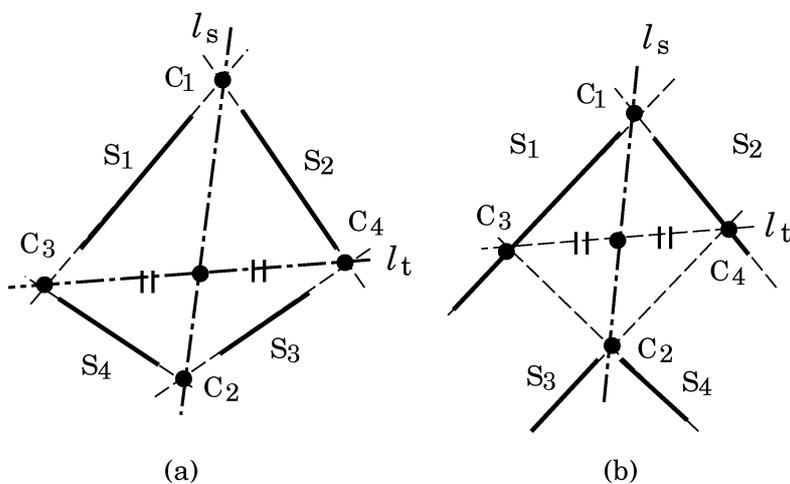


図 6.7: 2 対の直線セグメント

6.4.2 候補軸の検出 2 (1 対の直線セグメントと 1 直線セグメントより)

1 対の直線セグメントと 1 直線セグメントより候補軸を求める手順を以下に示す.

1. 図 6.8(a) に示すように, T 交差をしない 1 対の直線セグメント (S_1, S_2) を選ぶ.
2. 次に, (S_1, S_2) のいずれとも T 交差をしない 1 本の直線セグメント S_3 を選ぶ.

3. C_1 を (S_1, S_2) の交点とする.
4. C_2 および C_3 は, それぞれ S_1, S_3 および S_2, S_3 の交点とする.
5. M_1 を C_2C_3 の中点とする.
6. C_1M_1 を通る直線 l_s を斜め対称軸の候補とする. S_3 の乗る直線 l_t 斜め横軸とする.
7. 同図 (b) に示すように (S_1, S_2) が平行の場合には, C_1 を求めることはできない. その際には, (S_1, S_2) のいずれとも T 交差をしない 1 本の直線セグメント S_3 を選び, 斜め横軸とする. M_1 を C_2C_3 の中点とすると, M_1 を通り, (S_1, S_2) に平行な直線 l_s を斜め対称軸に, S_3 の乗る直線 l_t を斜め横軸とする.

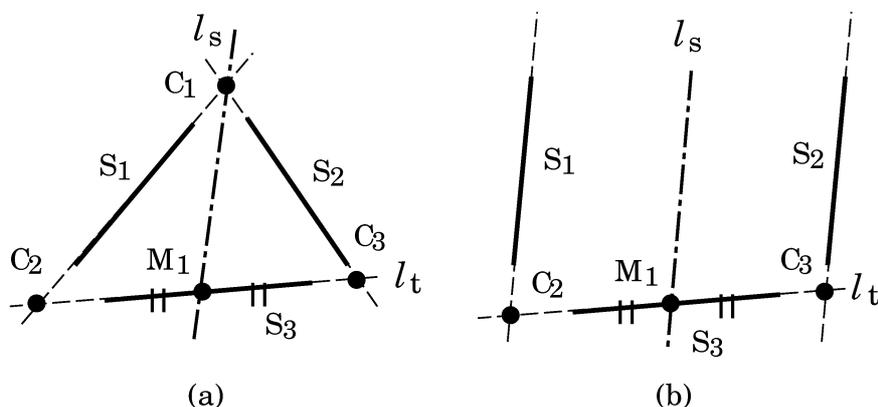


図 6.8: 1 対の直線セグメントと 1 直線セグメント

6.4.3 候補軸の検出 3 (1 対のコニックより (1))

1 対のコニックより候補軸を求める手順を以下に示す.

1. T 交差をしない 1 対のコニック (S_1, S_2) を選ぶ.
2. O_1 および O_2 を, それぞれコニック S_1 と S_2 の中心とし, 図 6.9(a) に示すように, O_1 および O_2 を通る直線 $l_t(i, j)$ を斜め横軸とする.
3. 点 $P_1(i)$ および $P_2(j)$ を, それぞれ $l_t(i, j)$ とコニック S_1 及び S_2 との交点とし, $M(i, j)$ を $P_1(i)$ と $P_2(j)$ の中点とする.
4. コニック S_2 上の点 $P_2(k)$ を通り, $l_t(i, j)$ に平行な直線 $l_t(k)$ とコニック S_1 との交点を $P_1(k)$ とし, $M(k)$ を $P_1(k)P_2(k)$ の中点とする.

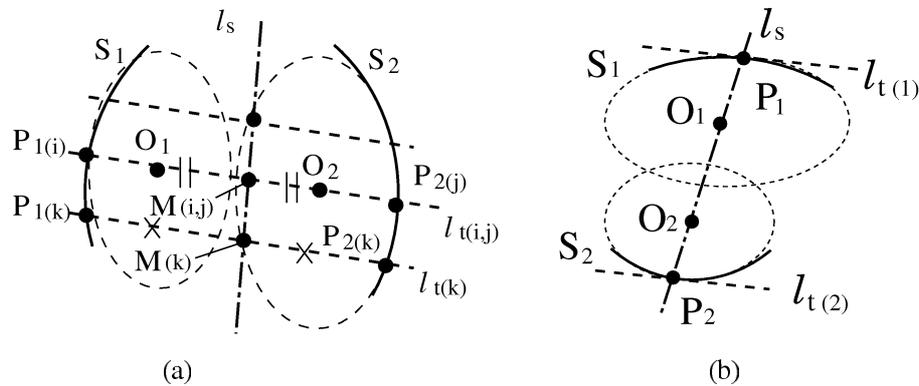


図 6.9: 1 対のコニック

5. $P_1(k)P_2(k)$ をコニック (S_1, S_2) の端点より複数点選び, 求まる複数点の $\{M(k)\}$ に対して直線近似し, これを斜め対称軸とする.

6.4.4 候補軸の検出 4 (1 対のコニックより (2))

しかしながら, 第 6.4.3 節に述べた手順は, 斜め対称軸が両コニック (S_1, S_2) と交差する場合 (図 6.9(b) 参照) には適用できない. そのような場合には, 以下の手順にて候補軸の検出を試みる.

1. 第 6.4.3 節に述べた手順同様, O_1 および O_2 を, それぞれコニック S_1 および S_2 の中心とする. 図 6.9(b) に示すように, O_1 と O_2 を通る直線 l_s を, 斜め対称軸とする.
2. $l_t(1)$ および $l_t(2)$ を, それぞれコニック S_1 および S_2 の主軸とする.
3. $l_t(1)$ と $l_t(2)$ の傾きが等しければ, この傾きを斜め横軸のそれとし, l_s を斜め対称軸とする.

6.4.5 候補軸の検出 5 (1 非コニックより)

非コニックと判定された 1 本の曲線セグメント S からの斜め対称軸検出手順は以下の通りである.

1. 非コニックと判定された 1 本の曲線セグメント S を選ぶ (図 6.10 参照).
2. セグメント S 上の任意の異なる 2 点 ($P(i)$ および $P(j)$) を選び, それらを通る直線 $l_t(i, j)$ を斜め横軸の候補とする.

3. 点 $P(k)$ を通り, $l_t(i, j)$ と同じ傾きの直線を $l_t(k)$ とし, $l_t(k)$ とセグメント S が2交点を持つ場合のみ, 2交点の中点を M_k として対象軸決定のための点列リストに登録する.
4. 対象軸決定のための点列リストに登録されたすべての $\{M_k\}$ を用いて最小二乗近似にて直線をあてはめ, 近似誤差の最大値がしきい値以下の場合に, 近似直線を斜め対称軸とする.

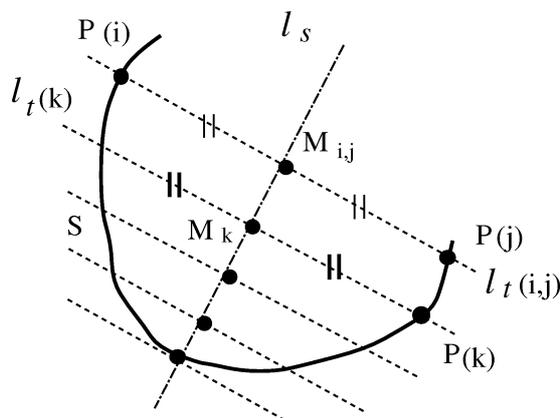


図 6.10: 1 非コニック

6.5 斜め対称プリミティブの統合

前節の手順にて検出された局所的斜め対称パターンのうち, 同じ斜め対称軸・斜め横軸をもつもの同士を複数統合することにより, より大局的な斜め軸対称パターンとして検出することができる. また, コニックに関しては, 複数のコニックがより大きな一つのコニックを構成する場合には, 統合することによってより信頼できる構成要素とすることができる.

以下に, 斜め対称パターンの統合基準を示す.

1. 斜め軸対称パターンの横軸と同じ方向をもち, かつ斜め対称軸と交差する直線セグメント
2. T交差しない (いずれのセグメント上にも互いに交点を持たない) 直線セグメント対のうち, 交点が斜め対称軸上にあり, かつ前記交点と反対側の端点を結ぶ線分の中点も斜め対称軸上にあるもの
3. その中心が斜め対称軸上にあるコニック

図 6.11 に、斜め軸対称パターンの検出実験に用いた入力原画像ならびにエッジ画像を示す。図 6.12 に、図 6.11 の入力画像からの斜め軸対称パターンの検出結果を示す。

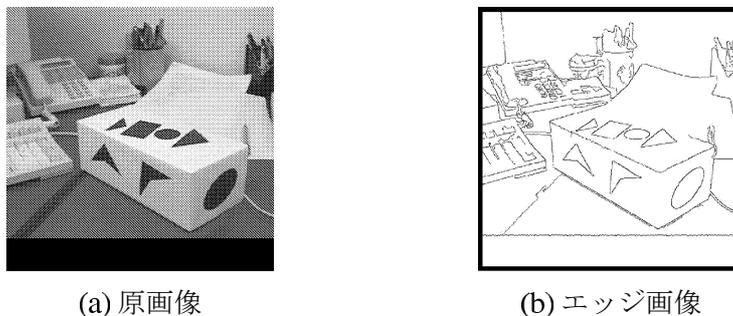


図 6.11: 入力原画像とエッジ画像

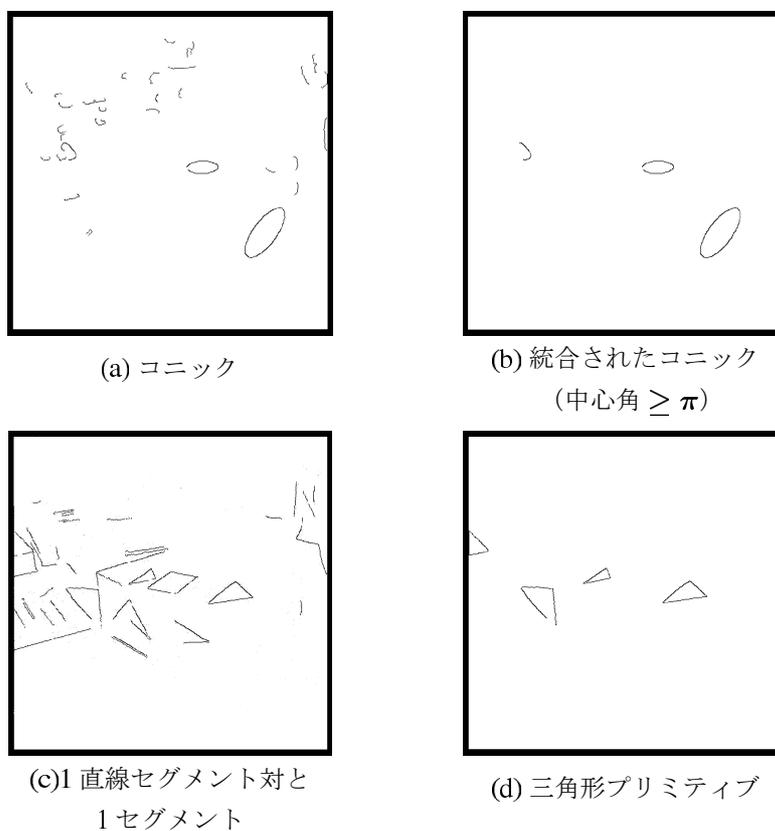


図 6.12: 局所斜め軸対称パターンの検出結果

同図 (a) には、1 対のコニックに基づいて検出された局所斜め対称パターンを、同図 (b) には、それらの中から統合可能なコニック同士を統合し、中心角が π 以上のものを抽出した結果を示す。更に、同図 (c) に、1 対の直線セグメントと 1 セグメントに基づいて検出された局所斜め対称パターンを、同図 (d) に、三角形の辺の長さの合計の 80% 以上がパターンとして含まれているものを示す。

図 6.13 に、局所斜め対称パターンを複数統合することにより検出された斜め対称パターン（斜め軸対称プリミティブとする）を示す（2対の直線セグメントからなる局所斜め対称パターンや、統合されたコニックのみからなる局所斜め対称パターンについては含めていない）。

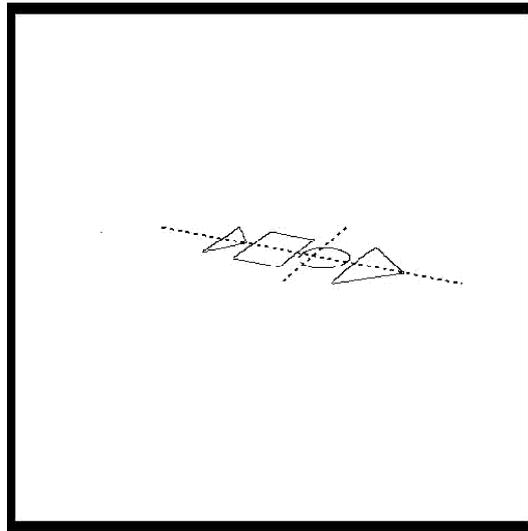


図 6.13: 斜め軸対称プリミティブの検出結果

6.5.1 斜め面对称プリミティブの検出

3次元空間中で、ある面に対して対称の位置にある3次元セグメントを、任意の視点から観測した画像中には、それらセグメントは、斜め面对称パターンとして投影される。斜め面对称性は、人工物等に多くみられる対称性に基づいて立体を認識する際の手掛かりとして利用することができる。斜め軸対称プリミティブを複数統合することにより、斜め面对称プリミティブを検出することができる。形状プリミティブをボトムアップに統合し、斜め面对称プリミティブという強い拘束のパターンとして検出することで、モデルとの照合の際の探索空間の削減に大きく貢献できるものとする。斜め面对称プリミティブの手順を以下に示す。

- 図 6.14 に示すように、各斜め軸対称プリミティブに対して、斜め横軸 l_t と等しい傾きをもち外接する2直線（中の (l_1, l_2) ）と、斜め対称軸 l_s との2交点（同図中 P_1, P_2 ）を求め、斜め対称軸の端点とする。
- 注目する斜め軸対称プリミティブと斜め横軸の傾きが等しい斜め軸対称プリミティブをすべて選び出す。

- 図 6.15 に示すように、互いの斜め対称軸 (l_{s1} および l_{s2}) の交点 C が、互いの斜め対称軸の端点 ((P_{11}, P_{12}) および (P_{21}, P_{22})) の一方に近い場合 (同図の例では、 P_{12} と P_{21} が C に十分近いことを示している)、これら斜め軸対称プリミティブ (同図中の (G_1, G_2)) を統合し、斜め面对称プリミティブとする。
- 斜め面对称プリミティブの各斜め対称軸に対して、斜め対称軸の交点が斜め対称軸の端点の一方に近い斜め軸対称プリミティブをすべて統合する。

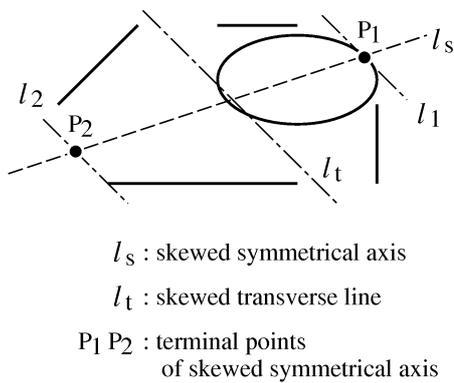


図 6.14: 斜め対称軸端点の決定

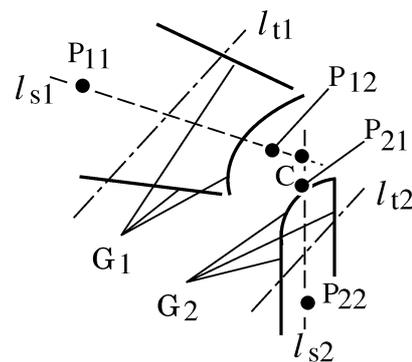
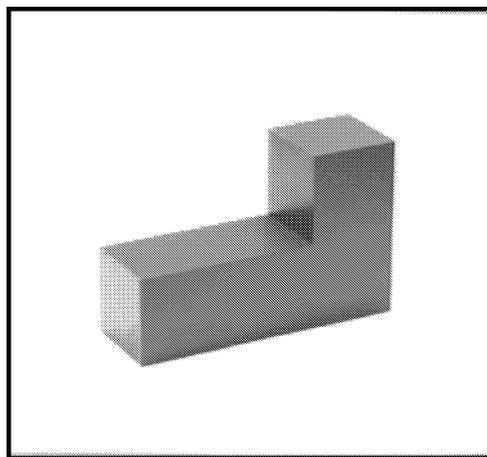
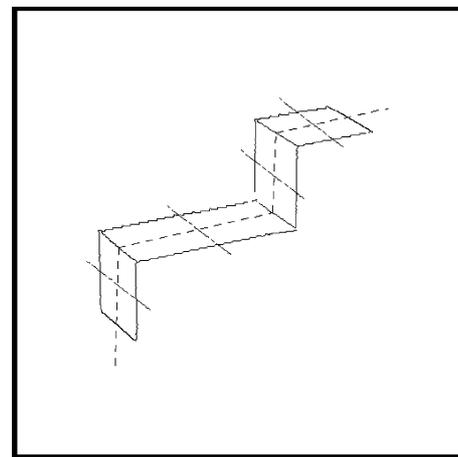


図 6.15: 斜め軸対称図形の統合

図 6.16(b) に、同図 (a) のブロックの画像からの斜め面对称図形の検出結果を示す。



(a) 原画像



(b) 検出結果

図 6.16: 斜め面对称図形の検出

6.5.2 未知対象物体把持動作計画への利用

ビジョンベースのマニピュレーションシステムに用いられる物体認識機能の多くは、対象物や作業環境に関する幾何モデルを必要とする。平行二指ハンドによる pick-and-place のみならず、回転や滑らし動作といった器用なハンドリングによる知的物体操作のための動作計画を行うシステム [176] においても、基本的にはモデルベースの動作計画システムであるため、通常利用可能な対象の幾何モデルが利用できない、未知の対象物の操作を考えた場合、把持といった単純な動作に関してもその計画は難しくなる。

平行二指ハンドによる物体操作を行う際の安定把持について、図 6.17 に示すような、各把持要素の対象物の動きに与える拘束を考える。ここで、同図 (a) に示すグリップパに関して、軸 A_1 周りの回転の自由度は考えない（物体とグリップ間に十分な摩擦力が発生し、滑りによる回転動作はしないものとする）こととする。同図 (b) のような点一点把持の場合には、 A_2 および A_3 の自由度が残る。同図中の軸周りの円表記は、把持される対象物の動きに与える拘束が弱いことを示している。ここで、各軸 A_1 , A_2 , A_3 は、互いに直交している。同図 (b) から (h) に向かうにつれ、把持の安定性が増すことを示す。但し、摩擦や重力は考慮しないものとする。動作計画をする際には、より安定した把持が可能な要素を選択する必要がある。同図 (h) に示す、面一面把持が最も安定していることとなる。

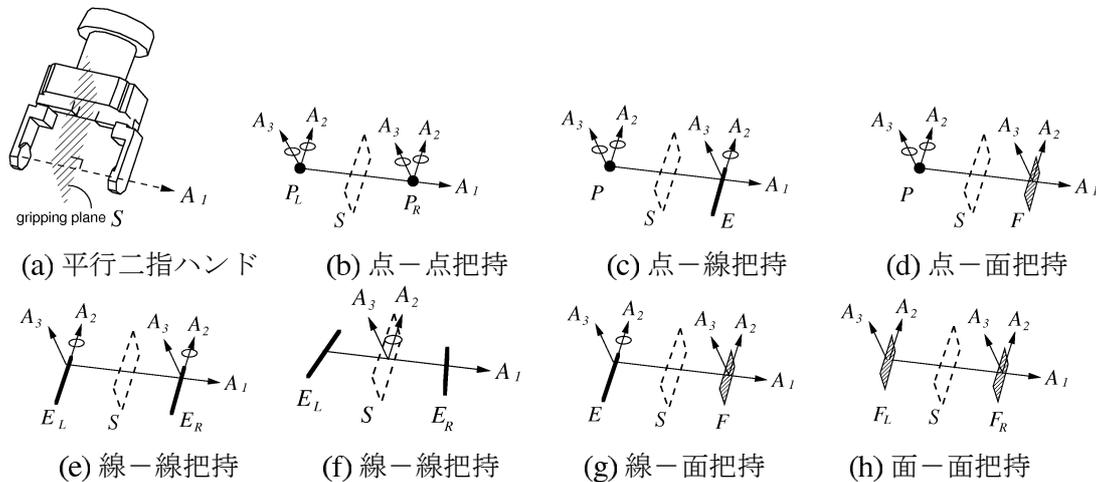


図 6.17: 把持要素例

そこで、ハンド・アイシステムによる未知対象物に関する把握動作計画において、対象物をハンドリングするための把持面を決定する際に、画像より抽出した斜め面对称プリミティブの利用を試みた [175]。

しかしながら、2次元画像中の斜め面对称プリミティブは、必ずしも3次元空間中

で面対称の関係にあるとはいえないため、ステレオによって再構成された3次元の面やエッジの情報により検証する必要がある。

一方、各画像中で検出された斜め面対称プリミティブ同士をステレオ対応の単位とすることにより、大きな拘束として利用することができ、探索空間の低減や信頼性の向上が期待できる。

図6.18に、作業を平行2指ハンドによる多面体物体の把持に限定した場合の、斜め面対称性を利用した未知対象物把持の模擬動作計画例を示す。

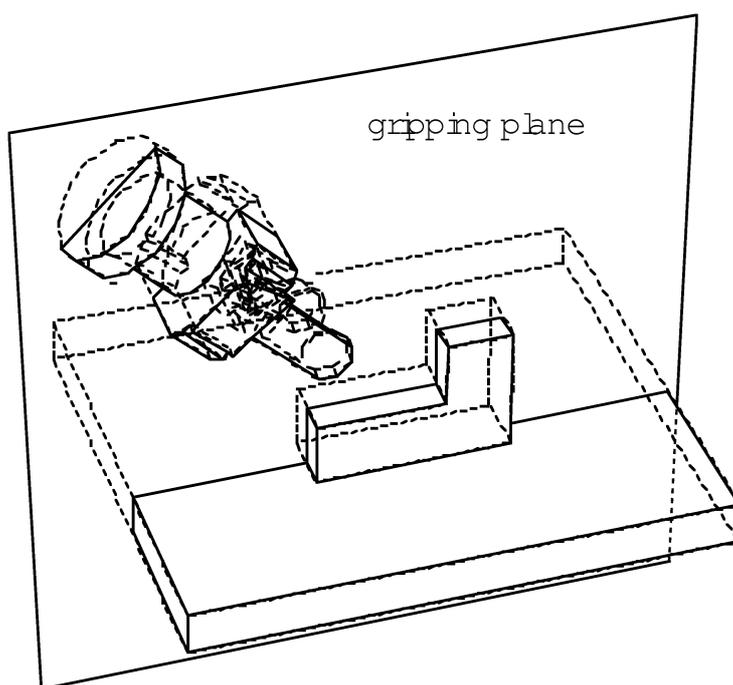


図 6.18: 把持動作計画への利用例

斜め対称パターンは、正射影の仮定の下で成り立つ関係である。第6.4節及び第6.5節にて用いた画像の撮像条件は、第3.2節にて言及した条件に等しく、透視投影の影響を受けているといえる。そのため実験においては、平行性に関して11度程度の誤差をマージンとすることにより影響を吸収した。

6.6 手指の検出と運動追跡への応用

本節では、クラス内変化への対応可能性の検証として、自由曲面体であり、また、個体差のある人の手指を認識対象とし、共通モデル(6.2節参照)に基いて人の手指の記述を生成することによる認識処理の有効性を示すと共に、剛体運動追跡への応用例を紹介する。

マルチメディア産業の急速な伸展により、立体モデルの利用に対する需要が増加している。それに伴い、実世界の対象物を3次元データに基づく現実感溢れるモデルとして生成する取り組みも見られるようになってきた[177]~[179]。これらのデータの仮想環境における利用は、今後益々盛んになると思われる[180]。例えば、ヴァーチャル・モールにおけるオンライン・ショッピング等においては、仮想空間中のショウケースから商品を取り出し、あたかも手にとって見ているかのごとく様々な方向から眺めるといった場面を想定できる。従来のリモコンやマウスと比較して、より容易な操作環境と、ユーザの感覚に合ったハンドリングを提供することは大変重要であると同時に期待も大きい。また、そのための次世代インタフェースの実現へ向けての取り組みも進められている[181]。これまで、立体モデルの操作方法に関しては、データ・グローブ、手袋、マーカー等を装着する方法が提案・利用されてきた。しかしながら、ユーザに対する負担の軽減や、より自然な対話システム実現の観点からは、非接触かつ受動的なセンシングによる手法が好ましい。そこで、1台のモノクロカメラにより手指の動きを撮影することにより得られる時系列の濃淡画像から、手指の位置を認識し、その運動追跡を行い、算出した運動パラメータに基づいて、仮想空間中の立体モデルを3次元的に操作するアプリケーションに適用した。

従来手法の多くは、対象物の幾何モデルを用いたり、屈曲点といった特徴を用いている。しかし、手指を対象とする場合、自由曲面物体でありかつ個体差もあるため、精密な幾何モデルを個々に用意することは難しいばかりでなく、特徴点に関して検出位置の不変性が保てないために特徴点検出に基づく手法も適用できない。

以下では、まず、共通モデルに基づく手指領域の検出について述べ、続いて、初期フレームの画像から検出された手指領域の輪郭線上に定めた追跡点に対する次フレーム画像中での対応点探索手法について述べる。次に、検出した対応点对の座標を用いて3次元空間中の平面の有限運動仮定に基づき[182]、手指の動きに応じて計算機内部の立体モデルを操作する応用例[9]を紹介する。

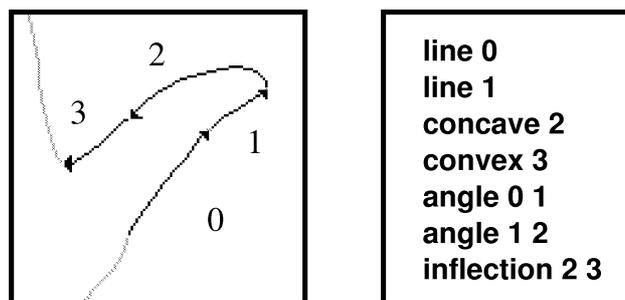
本手法の特長として、1枚の白黒濃淡画像から手指の位置を認識可能なこと、およびその運動を追跡可能なことがあげられる。

6.6.1 手指領域の検出

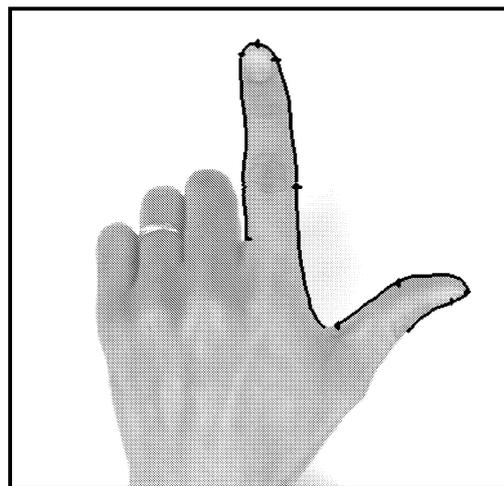
次世代の知的インタフェースを実現するにあたり、手や指の位置を認識し、それらの動きを追跡する処理は、大変重要な役割を果たすといえる。手の位置・姿勢の推定に関して、これまで多くの手法が提案されている[183]~[188]。また、手の領域検出に関しては、色情報は大変強力な拘束となるので[189]、それらを用いた手法が多く提案されている[190]。しかしながら、これら色情報を用いる手法は、もちろん白黒濃淡画像に適用することはできない。運動追跡に関しては、モデル・ベースの手法[191]

やステレオによる手法 [192] は強力であるが，単眼時系列画像に適用できるとは限らない．また，モデルベースのアプローチを実世界のすべての物体に適用する際には，先に述べたとおり，**appearance** の違いのみならず，クラス内変化への対応も合わせモデル爆発の問題に直面する．特に個体差がある対象について，個々に精密な幾何モデルを用意することは難しく，この問題を解決するために，第 6.2 節で述べた共通モデルのコンセプトを導入する．

図 6.19(a) に手指部位（親指）の記述例を示す．記述の生成には，6.3.1 節で述べた教示システムを用い，あらかじめ数枚の教示画像（図 6.20 参照）を用いて，人さし指，親指の記述をそれぞれ生成した．手指についても，直線セグメントおよび凹もしくは凸の曲線セグメントといった形状プリミティブの組合せにより記述することができる．親指ならびに人差し指部位の抽出は，この記述に基づいて行われる．この形状プリミティブの一般的な関係（6.3 参照）による定性的な記述により，手指部位の個体差を吸収することができる．



(a) 親指の記述例



(b) 手指の検出結果

図 6.19: 親指の記述例と手指検出結果



図 6.20: 手指の教示サンプル

まず，入力画像の初期フレームより，手指領域の輪郭線のエッジを検出し，屈曲点，変曲点，遷移点といった特徴点により，形状プリミティブとしてのセグメントを得る．しかしながら，手指領域の輪郭線上に現れるこれら特徴点は不変的なものではない．いいかえれば，いつも同じ部位に出現するとは限らないので，特徴点ベースの認識手法を適用すべきではない．追跡処理に関しても同様である．動的輪郭モデルを用いた輪郭線追跡に関する手法 [193][194] も提案されているが，短時間で収束させるには，収束時の輪郭に十分近い初期値を与える必要がある．しかしながら，初期位置や姿勢が未知の場合には，初期値を正しく与えることは難しい．

得られたセグメントの中から，記述と合致するプリミティブを抽出することにより手指部位を検出する．生成された記述は，見え方に不変的な特徴により記述される．そのため，フレームにて位置・姿勢が未知の手指部位の検出に適用できる．初期前記動的輪郭モデルに基づく運動追跡処理を行う際に利用可能な初期値を得るために利用することもできる．

それらの記述に基いて，未知の入力画像からの検出を行う．図 6.19(b) に，親指と人指し指の検出結果を示す．

6.6.2 手指領域の運動追跡

検出された手指の輪郭線上に追跡点（フレーム間で追跡対象とする参照点参照点）を決定し，隣接するフレームより，各追跡点に対する対応点を ICP(Iterative Closest Point) アルゴリズム [195] に基き，以下の手順により探索する．

- (1) 初期フレームにおいて，追跡点の法線方向を輪郭線上の近傍点を用いて算出する．

- (2) 隣接するフレームにおいて，前フレームの追跡点座標より法線方向にエッジ点を探索し，これを仮の対応点とする．図 6.21 に前フレームの輪郭線と追跡点 (○)，及び仮の対応点 (+) を示す．

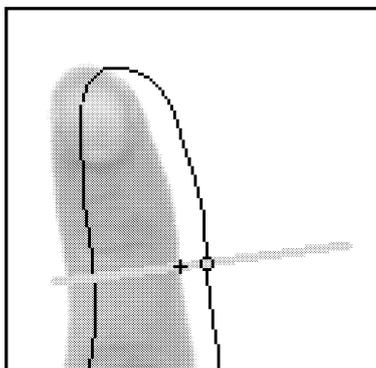


図 6.21: 追跡点と (○) と仮の対応点 (+)

- (3) 対応点の座標は，追跡点全体（ここでは指上の点）の動きを表現できる多項式 6.1 を解くことにより求める．ここで，すべての追跡点と仮の対応点は遮蔽輪郭上に位置するもの（前記各点が 3 次元空間中の同一平面上にのるもの）と仮定し，仮の対応点の座標を用いて，最小二乗法によりアフィン変換パラメータを求める．

$$\begin{pmatrix} X' \\ Y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} \quad (6.1)$$

a, b, c, d, e, f は，アフィンパラメータ．

- (4) 得られたアフィン変換パラメータを用い，追跡点座標から対応点座標を算出する．同様に，追跡点の法線方向より，対応点の法線方向を算出する．
 (5) 得られた対応点と法線を，各々新たな追跡点ならびに法線とする．

対応点の座標が安定するまで，(2) から (5) の処理を繰り返す．しかし，図 6.22 に示すように，対応点が正しく求まらない場合がある．この例では，フレーム t_i における追跡点 T が，隣接するフレーム t_{i+1} の $C_j (j = 1, 2, 3)$ いずれの点に対応するのか決定できない．

このような，いわゆる開口問題を回避するために，追跡点の法線方向が一方向に偏らず，90 度の角度差のある二方向以上の法線方向を有するように，図 6.23(a) に示すように，親指と人さし指の輪郭線上に追跡点を設定した．

同図中 (b), (c) に，後続の各フレームの画像において検出された対応点 (+) を示す．

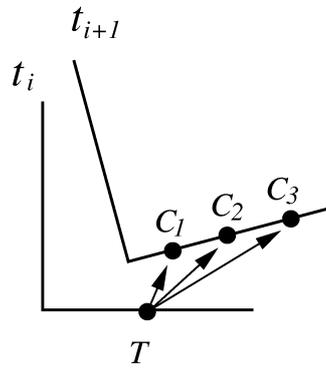


図 6.22: 開口問題

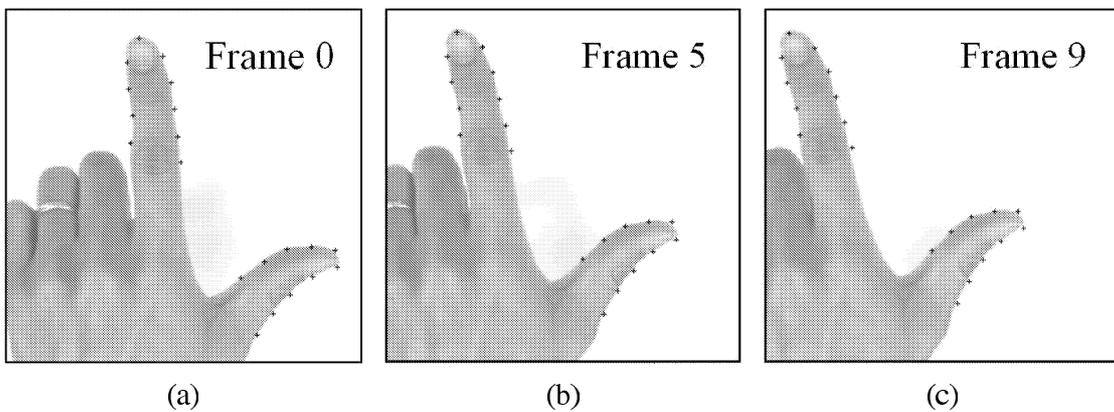


図 6.23: 手指の運動追跡

6.6.3 運動情報の算出

追跡結果として得られた複数（少なくとも4対以上）の対応点对の座標を用いて、手指の剛体運動を3次元空間中の平面の有限運動と仮定して、その運動パラメータとして、回転行列 \mathbf{R} 及び並進ベクトル \mathbf{t} を求める。

追跡点と対応点の3次元座標をそれぞれ、 $\mathbf{X}(x, y, z)$, $\mathbf{X}'(x', y', z')$ とし、 $\mathbf{X}_0(x_0, y_0, z_0)$ を回転中心とすると、手指の運動は次式で表される。

$$\mathbf{X}' = \mathbf{X}_0 + \mathbf{R}(\mathbf{X} - \mathbf{X}_0) + \mathbf{t} \quad (6.2)$$

そして、面の勾配を表すベクトルを (p, q) 、光軸に沿ってカメラのレンズの中心から平面へ下ろした垂線の足までの距離を r とする時、手指ではられる平面は次式で表される。

$$z = px + qy + r \quad (6.3)$$

ここで、カメラの焦点距離を f とし、3次元空間中の点 X , X' の画像面上への投影点を各々 (X, Y) , (X', Y') とする時、次式が成り立つ。

$$(X, Y) = (fx/z, fy/z), (X', Y') = (fx'/z', fy'/z') \quad (6.4)$$

これらの式を用いて、最小二乗法により、運動パラメータ R と t 並びに平面のパラメータ (p, q, r) が求まる。ここで、未知数は9であるが、並進ベクトル t と面迄の距離 r は、その比が求まることとなり、自由度は8となる。図 6.24 は、本手法を用いて、手指の動きに応じて算出された運動パラメータを用いて、計算機内部の仮想空間中に用意された立体モデルを3次的に操作した例を示している。また、図 6.25 にデモシステムの外観を示す。よりスムーズなインタラクションを実現するには、高速な運動追跡が重要となる。センサーレベルで移動体の輪郭や領域の抽出、移動方向の検出等ができれば格段の性能向上が期待できるとの観点から、インテリジェントイメーセンサの実現可能性についても検討した。詳しくは、付録 A を参照されたい。

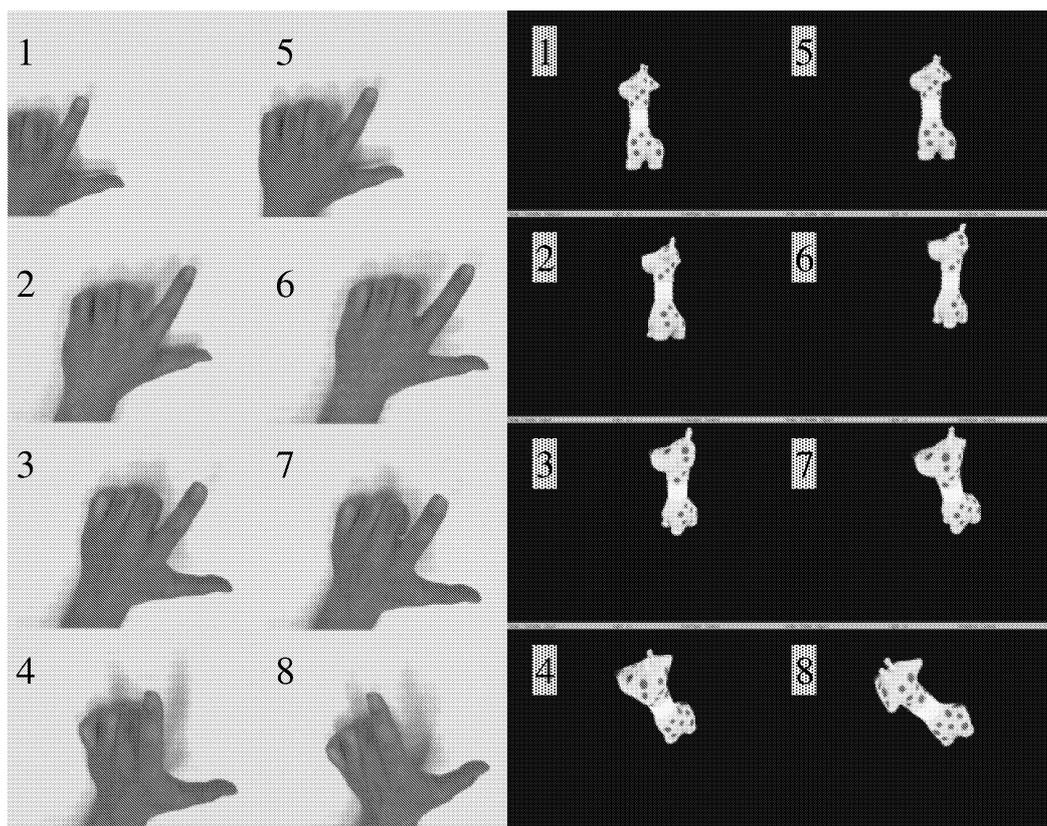


図 6.24: 手指の動きに連動した仮想空間中の物体操作例

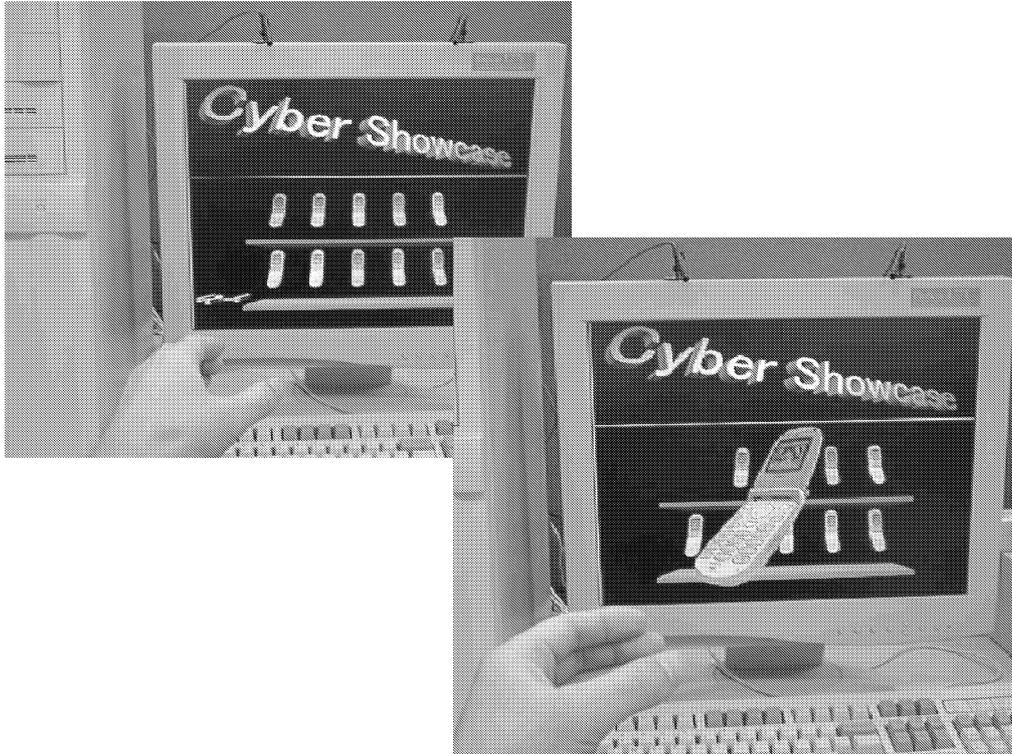


図 6.25: 仮想空間中の物体操作デモシステム

6.7 第6章のまとめ

appearance の違いへの対応の応用例として、ロゴマーク検出ならびに斜め対称性の抽出による軸対称図形および面对称物体の認識への応用について述べた。また、クラス内変化への対応可能性の検証として、個体差のある人の手指を認識対象として記述を生成し、剛体運動追跡への応用について述べた。

定性的な物体認識の実現に向け、基本形状特徴とそれらの一般的な関係により、対象を定性的かつ階層的に記述する共通モデルのコンセプトを紹介した。人工物の多くが持つ対称性に着目し、画像中の斜め対称の関係をを用いた認識への有効性を確認するとともに、観測方向の違いに起因する画像中での対象の変形や、個体差のある人の手指といった対象の認識にも適用し、概念的な記述の生成について可能性を示した。今後は、画像中のパターンからの概念形成の手法確立に向け、形状プリミティブを組み合わせる際の一般的な関係の整理と有効な応用を念頭に置いた適用範囲の拡大が重要と考える。

観測データの一部が、照明条件やオクルージョン等により欠落していても、また、物体表面に模様や影・雑音があっても、観測方向に対して不変的な形状特徴の部分照合により、対応箇所を抽出できる頑健性を備えることができる [171]。また、拘束の

大きな特徴（より上位階層の形状プリミティブ）を効果的に記述し、それらに基づき、トップダウン的探索との組合せにより初期照合を行うことで効率的かつ高速な探索が可能となると考える。

また、1台のモノクロカメラにより撮影された時系列濃淡画像を入力として、その輪郭線情報を用いて手指の位置・姿勢を認識し、運動を追跡する手法を提案した。手指は、自由曲面で構成されるため不変的な特徴点がなく個人差も大きいため、従来の幾何モデルベースのアプローチでは難しかった認識に関する問題を、定性的な認識手法により解決できることを示すとともに、手指の動きを剛体運動に限定して運動追跡を行い、親指と人さし指ではられる平面の3次元空間中での有限運動仮定に基づいて、固定カメラに対する相対的な6自由度の運動パラメータ算出を行い、算出されたパラメータを用いた仮想空間中の立体モデル操作への応用も試みた。今後は、非剛体運動（指の曲げ伸ばし等）の認識を行うためにステレオ化による拡張をはかり、手指の動きによる仮想空間における立体モデル操作や、ジェスチャによる利用者認証への応用を検討している。

しかしながら、人間が詳細な幾何モデルを用意することなく、実際に世の中の対象を認識しているのと同等の機能（対象の概念的把握に基づく定性的な認識機能）を実現するために、どれだけの形状プリミティブやそれらの間の一般的な関係を用意すれば十分であるのかは明らかにできていない。一般物体認識へのアプローチとして、形状プリミティブとそれらの間の一般的な関係についての検討を継続して進めたい。

第7章 結論

3次元視覚機能の汎用性の向上を念頭に、エッジベースステレオによる3次元情報の獲得と認識に関する研究として、その適用範囲の拡大と処理の高速化に重点を置き研究を行った。本論文では、基盤技術としての境界線分割のための特徴点抽出手法、テクスチャやシェーディングのない様な平面より構成される多面体の解釈・再構成手法、単一視点からの観測では解釈不可能な領域の特定と別視点からの観測データとの統合による再解釈ならびに記述の更新手法、モデルベースの物体認識処理の高速化ならびにモデル爆発問題の解決に向けた定性的物体認識手法に関して述べた。

第2章では、境界線上の特徴点検出手法について述べた。境界線上の特徴点は、境界線をより少ないセグメントに分割することを目的として検出される。特徴点により分割されて得られる直線セグメントや凹もしくは凸の曲線セグメントは、エッジベースステレオにおける対応探索や再構成、物体認識におけるモデルとの照合の際の単位として用いられると共に、共通モデルにおける最下層の形状プリミティブとしても用いられる。そのため、特徴点検出は、前処理として非常に重要な役割を果たしている。特徴点としては、屈曲点のみならず、凹凸の反転する変曲点や、直線から滑らかに曲線に移行する遷移点をも検出することができる。検出にあたり、最適近似区間の決定と近傍画素を決定するパラメータの動的適応をはかり、境界線上の各点における曲率・法線ベクトルといった局所特徴を精度良く算出した。本手法によれば、検出に必要なしきい値を対象毎に調整する必要がないばかりでなく、境界線形状の微小構造を自動的に検出することにより、細かい形状の部分についても特徴点を検出することができると共に、緩やかな曲線部分が多角形近似により過分割されることも回避できる。

提案手法の確立により、ステレオ対応や物体認識におけるモデルとの照合における探索空間の削減や、共通モデルの生成が実現でき、エッジベースのアプローチの効率化・高速化に貢献できたと考える。

第3章では、単一視点からのステレオ視によって得られる物体の境界線のワイヤフレームから正しく平面を推定するアルゴリズムにより物体の面を復元し、内挿された平面に基いて観測シーンの境界表現を生成する方法について述べた。実在テストにより、各ループ上のセグメントを可視面の輪郭境界を表す実セグメントと手前の面と背景との境界を表す虚セグメントに分類し、各ループに張るべき面を決定すると共に、境界線を輪郭線と稜線に分類し、更に、稜線を、その両側の面の法線方向のなす角度

によって、凸、凹、平稜線に分類できる。

また、欠損等によりデータが不完全となる場合に偽の面を生成することなく、実在テストを満足しないために解釈不可能となる境界線ループに対しては、それを補償する後の観測のためにラベルを付与し保存することができる局所性を備えている。本手法により、エッジベースステレオの適用範囲を拡張でき、汎用性の向上に貢献できたと考える。

第4章では、異なる観測位置から得られる不完全なステレオデータを逐次統合しながら、境界表現一面表現による環境の記述を生成する手法について述べた。単一視点より観測可能な視野には限界があり、照明条件・雑音等によるエッジ検出の失敗以外にも観測できない部分が残る。より完全な幾何モデルを生成するために、視点の移動により新たに観測される部分の統合・更新を行う。

提案手法ではステレオカメラを移動することにより、各視点にて生成した環境の境界表現に基づく局所地図（環境の幾何モデル）を世界地図（既に構築されている環境の幾何モデル）と逐次統合・更新する。両地図間の対応部分より移動パラメータを算出し直すと共に、対応の求まる部分に関しては各セグメントの信頼性に基き精度向上をはかる。また、解釈不可能なループに関しては、統合後に接続可能なセグメントがあれば新たなループの構成可能性を調べ、新規のループに対して実在テストにより面の推定（再解釈）を行い、解釈可能となった面については、新たに観測された面と合わせて世界地図に統合する。この統合と再解釈により、逐次的環境モデルの構築が可能となった。

第5章では、3次元物体認識の並列処理による高速化について述べた。エッジベースステレオによるシーン中の物体認識において、シーンから抽出された物体の特徴点と、あらかじめ構築された物体の幾何モデルとの照合に並列処理を適用した。各候補の照合処理は独立に実行できるため、並列処理による大幅な高速化が期待できると共に、特徴点を多く含む物体の認識には多くの処理時間を必要とするために、高速化に対する期待が大きい。

今回、ネットワーク接続された性能の異なるワークステーションによって構成されるワークステーションクラス上で、モデル照合処理の並列処理性能に関して、初期分配タスク数による性能評価、再分配タスクによる性能評価、プロセッサ数による性能評価ならびにマスタ／スレーブによるタスクプール方式を用いた動的負荷分散の効果に関する評価を行い、初期分配タスク数を大きな値に、再分配タスク数を小さな値に設定することによって、台数効果ならびにシステム全体の性能の引き出しといった良好な並列性能によりその有効性が確認できた。本手法は、認識処理に限定されることはなく、一般的な並列分散処理に応用できるものであり、汎用性の高い手法として広く貢献できるものと期待する。

第6章では、一般物体認識へのアプローチとして、appearanceの違いとクラス内変

化への対応を目的として、対象を定性的かつ階層的に記述するコンセプトとして共通モデルを提案した。appearanceの違いへの対応に関しては、ロゴマーク検出ならびに斜め対称性の抽出による軸対称図形および面对称物体の認識への応用例と合わせその有効性を示した。また、クラス内変化への対応に関しては、個体差のある人の手指を認識対象として記述を生成し、剛体運動追跡への応用を紹介した。観測データの一部が、照明条件やオクルージョン等により欠落していても、また、物体表面に模様や影・雑音があっても、観測方向に対して不変的な形状特徴の部分照合により、対応箇所を抽出できる頑健性と合わせて共通モデルの可能性を示した。

本研究により、3次元情報の獲得に関しては、一様な平面のみからなる多面体から構成される環境下での幾何モデル生成に関して、面の推定や解釈不可能な領域の特定が可能となり、エッジベーストステレオの適用範囲の拡大がはかかれているが、より一層の汎用性の向上に向け、テクスチャやシェーディングを含む領域や曲面物体の復元手法といった、他のパッシブなセンシング手法との融合をはかる必要がある。更に、注視点制御の導入や観測位置の計画、移動物体を含む環境下での応用に関しても、性能や頑健性の向上と適用範囲の拡大に向けて取り組んでいかねばならない。合わせて、ナビゲーションやセキュリティ応用も鑑み、SLAMのアプローチ [213]~[216] といった研究との融合も考慮しながら、マーカ等も用いない自己位置同定と環境モデリングにも取り組み、本研究の有効性の更なる検証をする必要があると考える。

認識に関しては、処理の高速化については、通信負荷や各プロセッサエレメントの負荷のばらつきが発生するホモジニアスな並列計算機環境やPCクラスタ環境、グリッドコンピューティングシステムにおいても、本手法の有効性が期待できるため、様々なアプリケーションによる性能評価を通じて有効性の実証を行いたい。また、一般物体認識の実現に向けた一アプローチとしての取り組みとして、共通モデルのコンセプトに基づいて形状プリミティブを組み合わせる際の一般的な関係の整理と有効な応用を念頭に置いた適用範囲の拡大が重要と考える。加えて、画像中のパターンからの概念形成の手法確立の可能性についても検討が必要と考えている。

本研究の成果に基づき、コンピュータビジョン手法の汎用性の向上に貢献できるよう、課題の解決に向け、精力的に研究開発を推進していきたい。

付録A 光感応蛋白質を利用する インテリジェントイメージセンサ

A.1 概要

VLSI技術の進歩と、脊椎動物の網膜をモデルにしたビジョンチップの研究開発を契機として、生態系をモデルとしたアルゴリズムばかりでなく、画像認識アルゴリズムをアナログVLSI上に実現するアプローチ [197] が、平滑化による雑音除去、エッジ検出といった初期視覚問題への適用 [198]~[201] を中心に進められてきた [202].

多くの物体追跡処理では、撮像・蓄積された各画像フレームデータ間で対応点探索などの処理を行う。その際のフロー算出処理は、画像中の対象物体領域のみならず全画素に対して行われることがほとんどである。しかしながら、処理の効率化の観点からは、処理対象とすべきは該当領域の画素データのみで十分なはずである。更には、撮像デバイスレベルで必要な情報（例えば輪郭情報）のみを直接抽出できるならば、物体追跡処理の格段の効率化が期待できる。近年、当該処理の高効率化を目的とし、撮像と特徴抽出処理の機能を併せ持たせたビジョンチップ（または、スマートセンサ）の研究開発が行われている [203]~[206]. これらは、いずれもCMOS半導体プロセス技術を基本に受光画素部であるフォトダイオード（PD）から出力された信号を、同一チップ回路上で信号処理してしまうために処理効率が良い。ただ、受光部分と信号処理部分は別々にならざるを得ない。

蛋白質は生体内において、分子、光、電気などの入出力を精密に分子レベルで高効率に行なう分子機能素子である。もし、機能性蛋白質分子を信号（情報）処理素子として利用するための技術が確立できれば、超高効率・超微小サイズのインテリジェントセンサ、更にはインテリジェントデバイスの実現が期待できる。蛋白質バクテリオロドプシン（Bacteriorhodopsin: BR）は、人の眼の受光蛋白質ロドプシンと類似した光応答特性を持つ。光照射（動画像投射）により微分型の光応答（動き特徴検出）を行い、電気または水素イオンにより信号出力する。これまでに、蛋白質BRの機能を用いた移動物体の輪郭や移動方向の検出が報告されている [207].

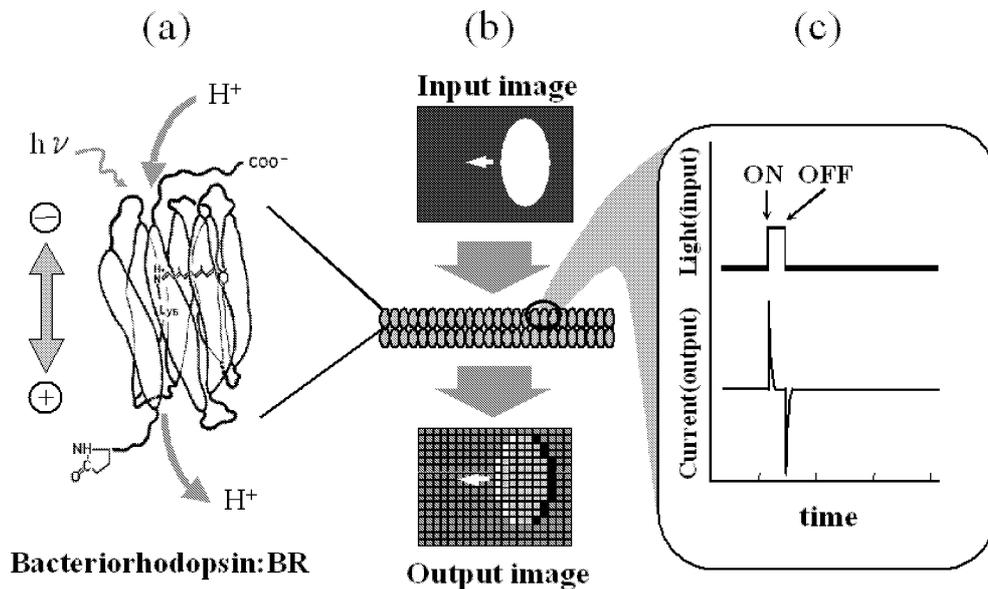
本章では、受光機能と輪郭抽出などの特徴抽出処理機能を兼ね備えた、BR分子の高効率なインテリジェント（スマート）機能を利用したインテリジェントイメージセンサの構築を念頭に、蛋白質BRを用いた移動物体特徴抽出として、BR分子の光電変換時に発生する電気分極を検出信号として利用する手法について述べる [208]. 現

在は、センサデバイス構築技術の基礎的検討段階ではあるが、具体的な処理過程の検討を行った結果、蛋白質機能を利用するセンサの有効性が明らかになった。まず最初に、BRによる特徴抽出原理と当該センサ実験システムについて述べる。次に、得られた蛋白質センサ出力の基本応答特性、移動物体の特徴抽出結果、および計算処理による特徴量抽出処理について紹介した後、本蛋白質の機能を利用したセンサの有効性と今後の課題について述べる。

A.2 光感応蛋白質による移動物体の特徴抽出

A.2.1 特徴抽出原理

BR分子の由来は、高度好塩菌（*Halobacterium Salinarum*）という細菌の細胞膜を貫通する膜蛋白質である。光エネルギー（最大吸収波長 568 nm の可視光）を吸収することにより光化学反応サイクルを開始する。その過程で分子構造を変化させつつ、図 A.1(a) のように細胞膜を貫通する方向にプロトン（水素イオン）を運ぶ（プロトンポンプ機能）とともに、微分応答型の電気分極を発生する。本来は、この機能により細菌の活動エネルギーが産生される。



(a)BR分子と機能, (b)BRフィルムによる特徴抽出, (c)BRの微分型光電変換応答

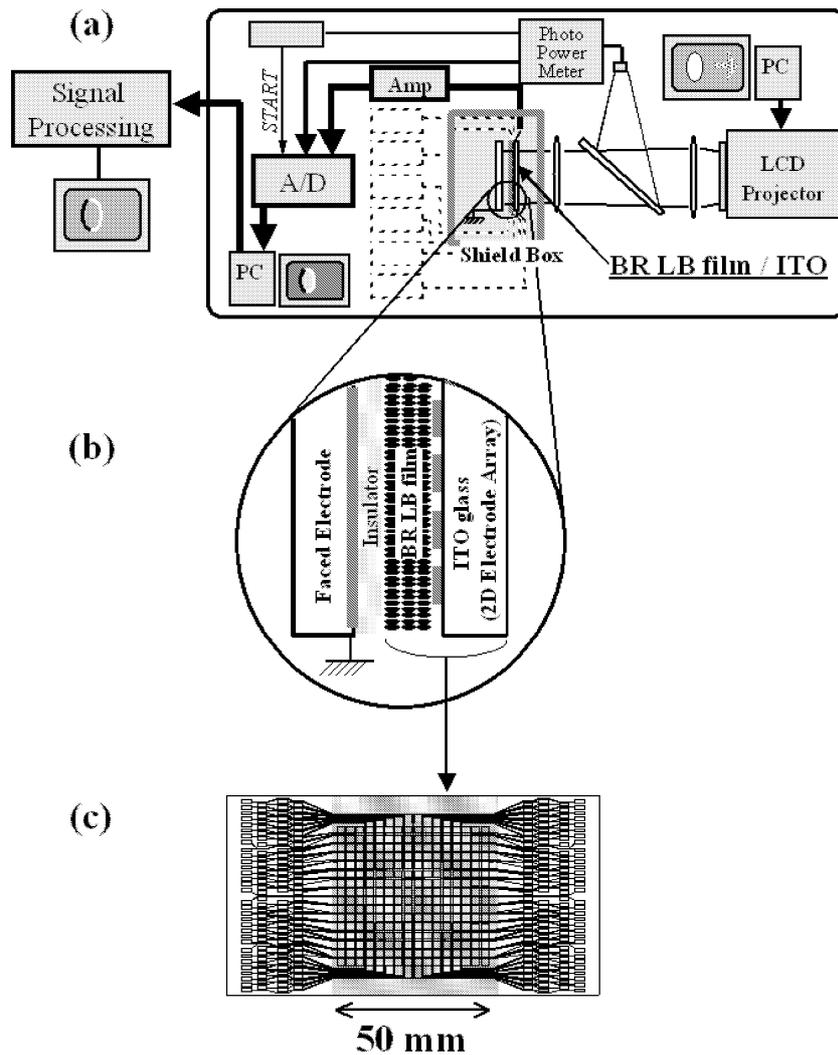
図 A.1: 特徴抽出原理

微分型応答は、BRが照射光の強度変化を捉えた際に出力される（同図(c)）。従って、BR分子の配向を揃えて2次元状に配列したBRフィルムを作製し、そこに動画

像を照射，その応答を2次元出力として検出すれば，同図 (b) のごとく移動物体輪郭等を抽出できることとなる。

A.2.2 実験システム

図 A.2(a) に実験に用いたシステムを示す。



(a) システム構成, (b) BR 受光部構造, (c) 電極アレイ基板

図 A.2: 実験システム

このセンサ系では，BR が光電変換時に発生する分子の電気分極を，マクスウェル変位電流検出法を応用した構成 [209] によって電流として検出している．同図 (b) の配置の各画素電極近傍の BR 分子が光電変換により発生した電気分極は，その極性と

逆極性の電荷を最近傍の画素電極に誘起する。その誘導電荷は対向電極と各画素電極間を短絡した回路を流れ、誘導電流として検出することができる。画素の信号検出に電流を用いたのは、画素間の信号のクロストークを抑制できると考えるからである。

蛋白質 BR は高度好塩菌 *H.Salinarum* の細胞膜から分離・精製した [210]。BR 分子の配向配列フィルムを作製するには、単分子膜累積法である電場印加 Langmuir-Blodgett (LB) 法 [211] を用いた。16 × 16 の画素電極 (2.5 × 2.5 mm²/画素) パターンを形成した ITO 透明ガラス電極基板 (図 A.2(c)) 上に、BR 単分子膜を 100 層程度累積 (膜厚約 0.5 μm) した。更に絶縁層として、不活性化 (界面変性) BR 単分子膜を数層程度累積 (膜厚約 0.01 μm) した。画素電極アレイの対向電極は、目的によって ITO 透明ガラス電極基板や銅基板を用いた。

本センサの基本的応答特性を検討することを目的とし、投影する動画像は簡単なものを作成・使用した。PC で再生した動画像 (AVI ファイル形式) は液晶プロジェクタから 30frames/sec で出力され、光学系を通して集光後、図 A.2(a) 中央部にある 50 × 50mm² の受光部の BRLB フィルム面に投影した。256 個 (16 画素 × 16 画素) ある各画素電極からの誘導電流出力 (256 の時系列データ) は、16 ビット、サンプル周波数 50kHz で 1 画素ずつ AD 変換し、データストレージ (PC) に蓄積して 1 次データとした。この 1 次データを更に計算処理する際のプログラムは C 言語で記述し、計算はワークステーション (EWS) で実行した。

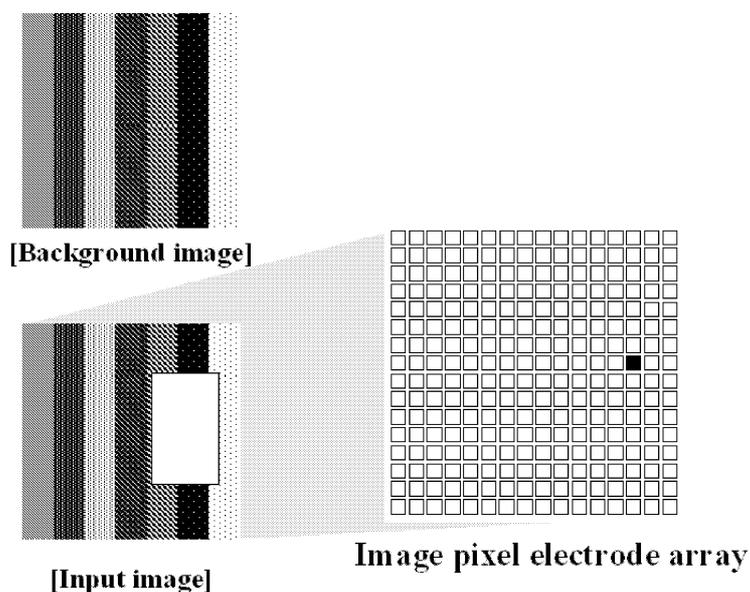
A.3 光感応蛋白質イメージセンサの出力特性

A.3.1 光感応蛋白質イメージセンサ受光部

対向電極に ITO 透明ガラスを用いた際には、作製した蛋白質 BR 受光部を光透過型にすることができる。これは、使用した両 ITO ガラス電極基板の光透過率が 80% 程度であり、かつ累積されている BRLB フィルム自体の光透過率が、BR の最大吸収波長である 568nm でも 90 数% あったからである。また、作製された BRLB フィルムは、機能的にも機械的にも非常に安定していた。BR 面を実験室内で空気に曝した状態のままにしておいても、作製直後と同程度の光電変換出力を数年間維持し続けている。更に、フィルム最上面に絶縁層として累積されている界面変性 BR 単分子膜は、その機械的強度が非常に強いいため結果的にフィルムの保護膜の役割も果たしている。

A.3.2 画素電極アレイにおける BR 応答出力

任意の画素電極において、動画像入力による光強度変化が BR の誘導電流の時系列応答として検出できる。図 A.3 は入力動画像パターンと背景画像、及び図 A.4 の電流



動画像中の白い資格は移動物体、
電極アレイ中の黒い画素は図 A.4 の電流時系列データの測定画素位置

図 A.3: 入力動画像，背景画像および 2D 画素電極アレイ

時系列データを測定する電極アレイ上の観測画素位置（黒い画素）を示している．背景画像には，移動物体の移動方向と垂直なストライプ状の濃淡模様がある．

図 A.4 は、図 A.3 の画素電極アレイ中に示した測定画素位置を，移動物体が一度通過した場合 (a) と，二度通過（往復）した場合 (b) の入射光強度に対する出力電流の各時系列データである．画素電極面上の BR 分子群が光強度の変化（ON や OFF）に反応して電気分極を生じ，その結果電流が出力される．同図では光 OFF 時の電流の立下りが，ON 時の微分応答の緩和に重畳されていて見えにくい．しかし，後述の信号処理により変曲点として容易に検出できる．更に OFF の応答により，ベース電流以下の負のピークが現れる．正の応答ピークは，動画像の輪郭通過などの光強度変化が強くなる場合であり，負の応答ピークは弱くなる場合に対応している．光強度変化の原因が図 A.3 のような背景よりも明るい移動物体であれば，正負の応答ピークはそれぞれ移動物体の前端と後端に対応する．従って，画素の位置，応答ピークの正負，および出現時間差から，対象物体の移動方向と移動速度を容易に計算することができる．

入力動画像の背景パターンは静止画像であるため，BR が動きだけに反応する微分応答特性と，図 A.2(b) のキャパシタ型の電極構造の効果により，背景画像の照射による出力が抑制される．背景画像が照射された直後は微分応答により各画素電極からノイズ様の出力が検出されるが，1 秒以内には初期レベル（電流ゼロまたはベース電流

値)に落ちている。即ち、動画像中の動きだけが検出される。図 A.5 に、図 A.3 の移動物体が一定速度で右方向（または、手前から奥の方向）に移動した際の、画素電極アレイの中央横 1 列に並ぶ各画素電極における応答を示す。

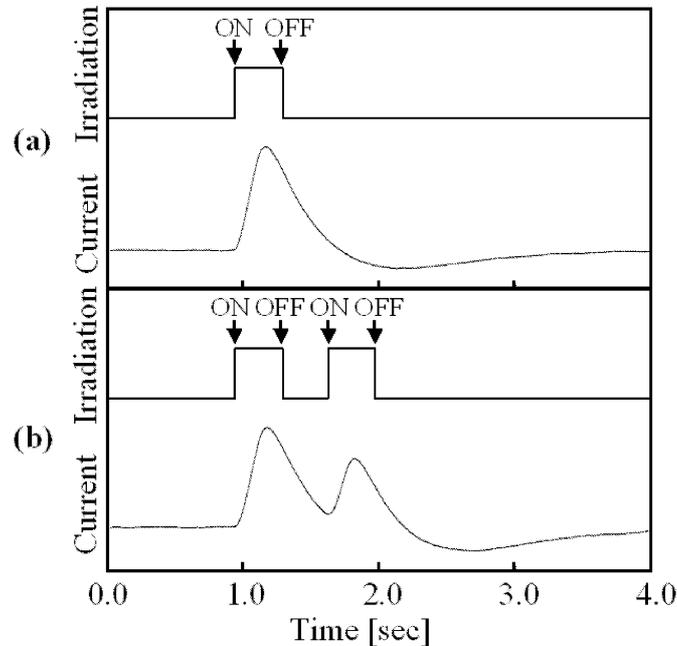
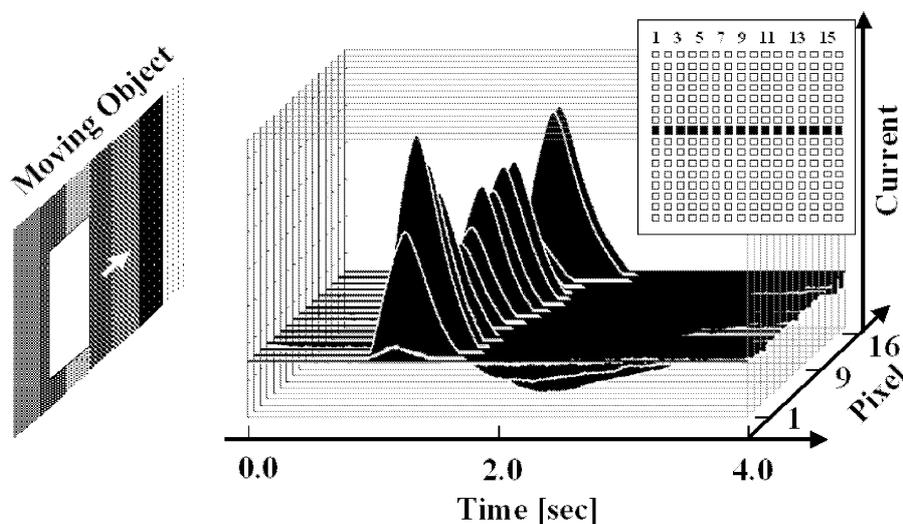


図 A.4: 電流時系列データ

移動物体像がそれぞれの電極に差しかかる前は、背景画像による応答（変化）は一切出力されていない。各時系列データのピークの高さが異なるのは、BR の応答出力が入力光の輝度変化量にほぼ比例しているからである。本データでは、背景と移動物体との輝度変化量が、背景パターンの位置によって異なることに起因する。また、移動物体の通過に伴う各時系列における電流応答の立上り時刻は、手前から奥の画素電極の順に後になっている。本実験で用いた画素電極中央の間隔は 3.1mm であり、隣接画素電極間の微分型応答のピーク時刻の間隔は 35msec であった。この結果を用いれば、移動物体の移動速度を求めるのは容易である。ただ、今回用いた再生用動画像や照射用液晶プロジェクタの動作速度の制約 (30frames/sec) から、本実験では 33 msec 以下の応答時間分解能に関しては議論できない。本蛋白質センサ自体の時間分解能の詳細な議論については、別途検討が必要である。

A.3.3 移動物体の特徴量分離抽出処理

本 BR イメージセンサでは、BR 分子が受光と特徴抽出を同時に行った後に AD 変換されるので非常に効率が良い。そのために、通常の撮像により蓄積されるフレーム



挿入図の電極アレイ上の中央横一列の電極で検出

図 A.5: 画素電極（横一列分）における電流応答

画像に相当する中間データが無い．ところで，本センサは各画素出力がそれぞれ独立（並列）に処理されるが，その全ての AD 変換は同期を受ける．そこで，本センサではこの同期タイミング毎に得られるデータを仮に“フレーム”と呼ぶこととする．今回の実験システムでは，50kHz で AD 変換しているので，本センサのフレームデータ取り込み速度は，50000frames/sec という非常に大きな値になる．実際，この速度でフレームデータは取り込まれてはいるが，実効的な処理速度は種々の要因によって決まる．今回の実験では，前節で述べたように動画像入力系の制約から，時間分解能は実時間（30frames/sec）である．

画素電極で移動物体の輪郭を捕えた時刻に，BR の微分型応答が検出される．ところで，図 A.4 に見られるように，本実験における BR 微分型応答のプロファイルは非常に緩やかであり，初期立上り後の緩和時の時定数は数 100msec にもなる．もし，この波形プロファイルのピーク時刻を利用して移動物体の輪郭検出を行おうとするならば，高速処理はもとより実時間処理さえ困難である．そもそも自然環境と同じ溶液中（湿度 100%）では，BR の光電変換応答の緩和の時定数は数 msec と非常に速いが，本実験環境では，BR が湿度 40～60% の乾燥状態のため，応答速度が遅くなっているものと考えられる [212]．ただ，受光部のキャパシタ型電極構造など，他の要因も原因の一つと考えられる．ところで，溶液中の BR のような速い応答速度であれば，ピーク時刻を利用しても実時間処理は可能であろう．しかし，物体追跡処理の目的に照らした場合，実時間での特徴抽出処理は決して十分な処理速度ではないと考えている．今回は，処理速度向上の観点から後段の信号処理による特徴抽出の高速化を検討した．

図 A.4(b) から分かるように，入射光強度の変化を捉えた瞬間に生じる微分応答の初

期立上りは比較的急峻である。一方，移動物体の輪郭検出において最も必要なのは，輪郭を捉えた瞬間の微分応答の立上り（立下り）時刻である。図A.6に，微分型応答の電流時系列データ（1次データ）から閾値付きの微分処理を用いて，輪郭・移動方向・移動物体の通過時間領域などの特徴量を分離抽出する処理過程を示す。

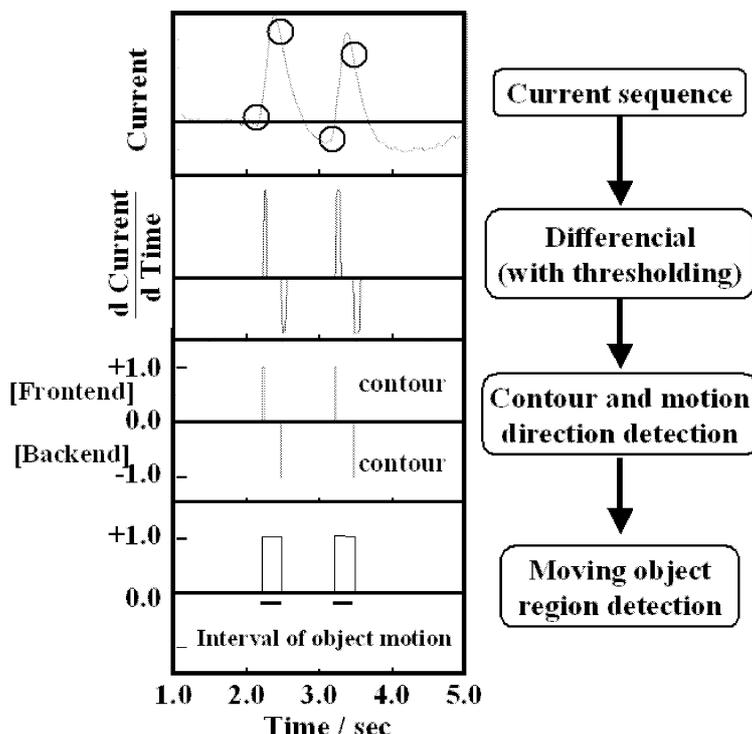


図 A.6: 特徴量分離抽出処理

この閾値付きの微分は，ノイズなどによる微分成分を排除し，急峻な初期立上り（初期立下り）成分のみを抽出するために必須である。閾値には正負の適当な値を設定した。今回の実験では，入力動画像中の移動物体の輝度が背景よりも明るい画像を使用している。従って，閾値付き微分の結果から抽出された輪郭の極性は，移動物体の前端では正值になり後端では負値になる。前端の輪郭検出の初期時刻に“+1.0”を付与し、後端に“-1.0”を付与した。この極性情報は，対象物体の移動方向を判断する際の指標となる。更に，これら分離抽出された情報から，移動物体が画素電極位置を通過している時間領域を抽出することができる。例えば，ある画素で+1.0の輪郭が検出されれば，その時刻以後この画素は移動物体像が通過している状態になる。次に-1.0の輪郭が検出されれば，その時刻に移動物体の後端が通過したことになる。この時間領域情報は，各フレームにおける移動物体画素領域抽出の際の情報として使用できる。各時刻のフレームで，移動物体の通過中か否かを画素位置毎に抽出すれば，移動物体領域の時系列を表示することができる。このように，特徴量分離抽出処理は

非常に簡単であり計算負荷は軽微である。これは、BR 自体が特徴抽出処理の殆どを行っているため、結果として後段の特徴量分離抽出処理の負担を軽くできているためといえる。

図 A.7 は、センサへの動画像入力から特徴量分離抽出処理までの過程を示した例である。

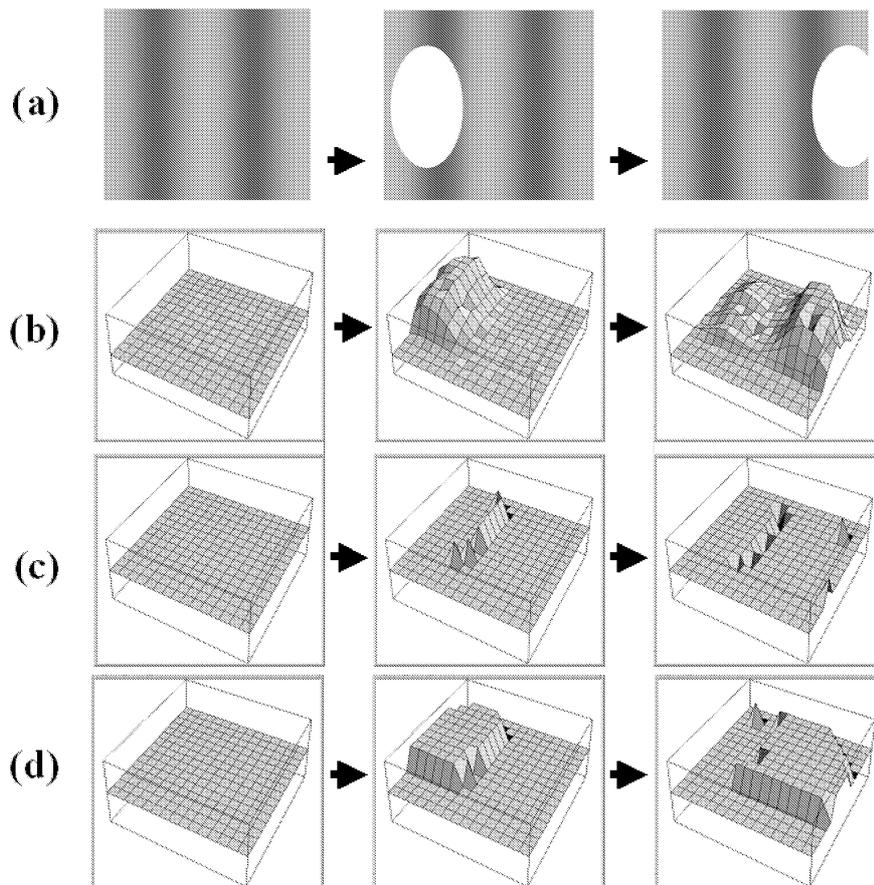


図 A.7: (a) 入力動画像, (b) 1 次データ, (c) 輪郭・移動方向抽出結果, (d) 物体領域抽出結果

同図 (a) には、パターンのある背景の前を右方向に一定速度で移動する楕円状物体を入力動画像として照射したパターンについて、時間順の適当な 3 フレームについて離散的にサンプリングして表示している。同図 (b) から分かるように、BR 受光部から出力される 1 次データでは移動物体に対応した応答出力がなされているが、BR 微分型応答の緩やかな緩和の影響が大きい。各画素の時系列データ毎に特徴量の分離抽出処理を実行した結果が同図 (c),(d) である。輪郭・移動方向を示す極性や、移動物体領域が明瞭に抽出されているのが分かる。特徴量分離抽出処理において移動物体の輪郭検出に要する時間は、微分型応答の立上りや立下りの瞬間を検出する時間のみである。

それゆえ、処理の全体を通して、動画像の受光から特徴量分離抽出処理までの実時間処理は十分可能であると考えられる。処理速度の見積りについては現在検討中である。

しかしながら、分離抽出された輪郭位置（移動物体領域）を元の入力動画像（同図(a)）と比較すると、輪郭の後端が後方に引きづられているのが分かる。この原因は、緩い緩和を示す1次データに主に起因すると考えられる。ただ、本分離抽出処理が画素間相互に関する処理を一切行っていない結果であることを考慮すると、本結果は満足すべきものであると考えている。以上、物体追跡処理（特徴抽出処理）に蛋白質BRを利用する有効性は明らかである。この課題の解決には、特徴量分離抽出処理の更なる検討が必要であるが、同時にセンサ受光部の検討も併せて行う必要があると考えている。

A.4 まとめ

物体の運動追跡といった実時間動画像処理において、各画像フレームにおける特徴抽出や、フレーム間での対応探索等に関する処理の効率化・高速化を念頭に、物体追跡処理に光電変換機能を持つ蛋白質BRを利用することにより、高効率の動画像特徴抽出処理を行うインテリジェントイメージセンサの構築が可能であることを示した。

提案した動画像特徴抽出系は、BR分子を高度に配向・配列したフィルムを含む受光部と、特徴量分離抽出処理部の二つの機能部位から構成されている。受光部は、蛋白質であるBRを使用しているにも関わらず、機能的および機械的に非常に安定であることが確認できた。

また、BRの微分型応答特性により、動画像中の背景における模様の有無に関係なく、移動物体のみを抽出できる機能を明らかにした。

センサからの一次出力データより、前記データ中に十畳されていた移動物体の輪郭・移動方向・領域といった情報が分離抽出可能であることを示した。ここで、センサー部では各画素レベルで処理対象物体の動きを捕えることができ、また特徴量分離抽出処理部では該当時間領域のデータにまで出力を削減できていることから、蛋白質BRの機能の利用は一種のデータ圧縮ということもでき、動画像特徴抽出処理の高効率化への有効性が確認できた。

今回、特徴量分離抽出処理はEWS上のプログラムにて行った。しかしながら、上述の特徴量の分離抽出処理は非常に軽く、回路規模にして数千〜一万ゲート程度で実装可能なレベルと見積ることができるため、処理速度の更なる向上を目指し、実回路による処理を計画している。

今後は、生体機能分子レベルにおける信号検出についての更なる検討による受光部の機能向上も必要かつ重要と考えている。

付録B 平面拘束定理

[平面拘束定理]

3次元空間中の任意の平面図形について、標準カメラモデルにおけるステレオ画像上の一方への投影像と、他方のそれとは互いにアフィン変換可能である。

この定理は、透視投影により得られる画像上のセグメントであっても、同一平面上にあるセグメントの集合は、ステレオ画像間でアフィン変換可能であることを表す[94]。これにより、セグメントを直接的に3次元空間に逆投影することなく画像上で検証することができる。

標準カメラモデルにおいて、3次元平面 $ax + by + cz = 1$ 上の一点 $P(x, y, z)$ の一方のカメラ C の画像平面への投影点の座標を (X, Y) とする。 $x = zX$, $y = zY$ であるから、これを平面方程式に代入し、 z について解くと、式 (B.1) を得る。

$$\begin{aligned} x &= \frac{X}{aX+bY+c} \\ y &= \frac{Y}{aX+bY+c} \\ z &= \frac{1}{aX+bY+c} \end{aligned} \quad (\text{B.1})$$

他方のカメラ C' は $(u, v, 0)$ だけ平行移動したところにあるとすると、カメラ C' から見た点 P の位置 (x', y', z') は、式 (B.2) より得られる。

$$\begin{aligned} x' &= x - u \frac{X}{aX+bY+c} - u \\ y' &= y - v \frac{Y}{aX+bY+c} - v \\ z' &= z \frac{1}{aX+bY+c} \end{aligned} \quad (\text{B.2})$$

その結果、点 P のカメラ C' の画像平面上への投影点の座標 (X', Y') は、式 (B.3) のように (X, Y) のアフィン変換で与えられる。

$$\begin{aligned} \begin{pmatrix} X' \\ Y' \end{pmatrix} &= \begin{pmatrix} \frac{x'}{z'} \\ \frac{y'}{z'} \end{pmatrix} \\ &= \begin{pmatrix} X - u(aX + bY + c) \\ Y - v(aX + bY + c) \end{pmatrix} \\ &= \begin{pmatrix} 1 - ua & -ub \\ -va & 1 - vb \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} + \begin{pmatrix} -uc \\ -vc \end{pmatrix} \end{aligned} \quad (\text{B.3})$$

更に、通常のカメラ構成では $\mathbf{v} = \mathbf{0}$ で \mathbf{y} 座標は変化しないと仮定すれば、次式 (B.4) のように簡単な変換となる。

$$\begin{pmatrix} X' \\ Y' \end{pmatrix} = \begin{pmatrix} 1 - ua & -ub \\ 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} + \begin{pmatrix} -uc \\ 0 \end{pmatrix} \quad (\text{B.4})$$

従って、逆に三点の対応が得られれば、アフィン変換のパラメータが求まり、更にカメラ間距離 u により、平面の方程式を求めることも可能となる。

謝辞

本論文の作成にあたり，親切なる御指導，御鞭撻を賜りました，本学システム情報工学研究科 大田友一教授に心より感謝いたします。また，研究内容について多くの助言を頂きました本学連携大学院客員教授，産業技術総合研究所 情報技術研究部門 喜多泰代様，及び本学システム情報工学研究科 亀田能成准教授に厚く御礼申し上げます。

本論文の執筆に際して，丁寧な御教示，御助言を頂きました，本学システム情報工学研究科 葛岡英明教授，本学システム情報工学研究科 坪内孝司教授に心より御礼申し上げます。

ゼミを通じ，貴重な御意見ならびに御助言を頂きました，本学システム情報工学研究科 北原格講師，本学システム情報工学研究科 掛谷英紀准教授に深く感謝いたします。

数々の御助言ならびに励ましの御言葉を頂きました，産業技術総合研究所 知能システム研究部門 富田文明様にも厚く御礼を申し上げます。

本学への入学以前より，多大なる御支援を頂くと共に，入学後も数々の御助言や暖かな御支援を頂きました，本学学務部学務課教務御担当 関瑞穂様に深く感謝いたします。

また，本研究を進めるにあたり数々の助言や議論を頂きました，産業技術総合研究所 情報技術研究部門 植芝俊夫様，産業技術総合研究所 知能システム研究部門 角保志様，産業技術総合研究所 イノベーション推進室 蔵田武志様に深く感謝いたします。

そして，本研究の御許可ならびに御鞭撻を賜りました，三洋電機株式会社 研究開発本部 デジタルシステム研究所 富田義数所長，片山立部長に御礼申し上げます。

さらに，本研究に御協力頂きました，株式会社 生体分子計測研究所 杉山幸宏様，三洋電機株式会社 研究開発本部 デジタルシステム研究所 大上靖弘様に深く感謝いたします。

最後に，本研究の推進にあたり，励ましの言葉と温かいサポートにより，心の支えとなってくれた，父，母，ならびに妻と娘に，心より感謝いたします。

参考文献

- [1] Y.Shirai and M.Suwa, "Recognition of polyhedrons with a rangefinder;" *Proc. 1st International Joint Conference on Artificial Intelligence (IJCAI-71)*, pp.80-87, 1971.
- [2] 井口 征士, 佐藤 宏介, "三次元画像計測", 昭晃堂, 1990.
- [3] K.Sugimoto and F.Tomita, "Boundary segmentation by detection of corner, inflection and transition points," *Proc. IEEE Workshop on Visualization and Machine Vision (WVMV'94)*, pp.13-17, 1994.
- [4] 杉本和英, 富田文明, "ステレオ画像の境界表現に基づくシーンの解釈", 信学論 (D-II), Vol.J76-D-II No.3, pp.524-533, 1993.
- [5] 杉本和英, 富田文明, "ステレオによる幾何モデリング", 日本ロボット学会誌, Vol.15, No.3, pp.431-438, 1997.
- [6] 大上靖弘, 杉本和英, 北村徹, 角保志, 富田文明, "ネットワーク型並列計算環境における物体認識", 信学論 (D-II), Vol.J82-D-II No.12, pp.2307-2315, 1999.
- [7] K.Sugimoto and F.Tomita, "Detection of skewed-symmetrical shape," *Proc. International Conference on Image Processing (ICIP'94)*, pp.696-700, 1994.
- [8] K.Sugimoto, H.Terasaki, and F.Tomita, "Detection of SkewedSurfaceSymmetrical Primitives for Grasping Unknown Objects," *Proc. Asian Conference on Computer Vision (ACCV'95)*, pp.II-757-761, 1995.
- [9] K.Sugimoto and F.Tomita, "Rigid Hand Motion Tracking based on a Generic Model for Manipulating 3-D Models of Objects," *Proc. IEEE Pacific-Rim Conference on Multimedia (PCM2000)*, pp.202-205, 2000.
- [10] C.Harris and M.Stephens, "A combined corner and edge detector," *Proc. 4th Alvey Vision Conference*, pp.147-151, 1988.
- [11] J.Shi and C.Tomasi, "Good features to track," *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'94)*, pp.593-600, 1994.

- [12] F.Chabat, G.Z.Yang, and D.M.Hansell, "A corner orientation detector," *Image and Vision Computing*, Vol.17, No.10, pp.761-769, 1999.
- [13] K.Rangarajan, M.Shah, and D.V.Brackle, "Optimal corner detector," *Computer Vision Graphics and Image Processing*, Vol.48, No.2, pp.230-245, 1989.
- [14] K.Rohr, "Modeling and identification of characteristic intensity variations," *Image and Vision Computing*, Vol.10, No.2, pp.66-76, 1992.
- [15] A.Singh and M.Shneier, "Gray level corner detection: A generalization and a robust real time implementation," *Computer Vision Graphics and Image Processing*, Vol.51, No.1, pp.54-69, 1990.
- [16] S.Baker, S.K.Nayar, and H.Murase, "Parametric feature detection," *International Journal of Computer Vision*, Vol.27, No.1, pp.27-50, 1998.
- [17] P.Perona, "Deformable kernels for early vision," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.17, No.5, pp.488-499, 1995.
- [18] H.P.Moravec, "Towards Automatic Visual Obstacle Avoidance," *Proc. 5th International Joint Conference on Artificial Intelligence (IJCAI-77)*, pp.584, 1977.
- [19] S.M.Smith and M.Brady, "SUSAN - A New Approach to Low Level Image Processing," *International Journal of Computer Vision*, Vol.23, No.1, pp.45-78, 1997.
- [20] L.T.Watson, K.Arvind, R.W.Ehrich, and R.M.Haralick, "Extraction of lines and regions from grey tone line drawing images," *Pattern Recognition*, Vol.17, pp.493-507, 1984.
- [21] R.Kasturi and S.Siva, "Techniques for line drawing interpretation:An overview," *Proc. IAPR Workshop on Machine Vision Applications*, pp.151-160, 1990.
- [22] P.Gerken, "Object-Based analysis-synthesis coding of image sequences at very low bit rates," *IEEE Trans. Circuits & Systems Video Technology*, Vol.4, No.3, pp.228-235, 1994.
- [23] K.Sugimoto and F.Tomita, "Shape Coding by Detecting Dominant Points on Digital Curves," *Proc. International Conference on Virtual Systems and Multi Media (VSMM'96)*, pp.153-158, 1996.
- [24] 名倉正計, "手書き線図形の直線と円弧による近似", *信学論 (D)*, Vol.J64-D, No.9, pp.839-845, 1981.

- [25] J.L.Dunham, "Optimum uniform piecewise linear approximation of planar curves," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.8, No.1, pp.66-75, 1986.
- [26] 堀 修, 谷川 智秀, "輪郭線と細線を用いた直線・円弧当てはめによる図面のベクトル化", *信学論 (D-II)*, Vol.J77-D-2, No.1, pp.91-100, 1994.
- [27] T.Pavlidis, "Algorithms for shape analysis and waveforms," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.2, No.7, pp.301-312, 1980.
- [28] M.A.Fischler and H.C.Wolf, "Locating perceptually salient points on planar curves," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.16, No.2, pp.113-129, 1994.
- [29] F.Tomita and M.Koizumi, "A step toward generic object recognition," *Proc. International Conference on Pattern Recognition (ICPR'92)*, pp.632-636, 1992.
- [30] A.Rosenfeld and A.C.Kak, "Digital picture processing," *Academic Press*, New York, 1976.
- [31] J.Q.Fang and T.S.Huang, "A corner finding algorithm for image analysis and registration," *Proc. AAAI Conf.*, pp.46-49, 1982.
- [32] R.Mehrotra and S.Nichani, "Corner Detection," *Pattern Recognition*, Vol.23, No.11, pp.1223-1233, 1990.
- [33] A.Rattarangsi and R.T.Chin, "Scale-based detection of corners of planar curves," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.14, No.4, pp.430-449, 1992.
- [34] H.Asada and M.Brady, "The curvature primal sketch," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.8, No.1, pp.2-14, 1986.
- [35] A.P.Witkin, "Scale space filtering," *Proc. International Joint Conference on Artificial Intelligence (IJCAI-83)*, Karlsruhe, West Germany, pp.1019-1022, 1983.
- [36] A.L.Yuille and T.Poggio, "Scaling theorems for zero crossings," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.8, No.1, pp.15-25, 1986.
- [37] 守田了, 川嶋稔夫, 青木由直, "輪郭線形状の階層的記述によるマッチング", *信学論 (D-II)*, Vol.J73-D-II, No.5, pp.717-727, 1990.

- [38] 出口光一郎, "マルチスケール曲率による図形輪郭線の特徴抽出", 計測自動制御学論, Vol.24, No.10, pp.75-81, 1988.
- [39] O.A.Zuniga and R.Haralick, "Corner detection using the facet model," *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'83)*, pp.30-37, 1983.
- [40] C.H.Teh and R.T.Chin, "On the detection of dominant points on digital curves," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.11, No.8, pp.859-872, 1990.
- [41] K.Sugimoto and F.Tomita, "How to detect dominant points on 3-D Curves," *Proc. European signal processing conference (EUSIPCO'98)*, pp.2177-2179, 1998.
- [42] H. P. Moravec, "Visual Mapping by a Robot Rover," *Proc. 6th International Joint Conference on Artificial Intelligence (IJCAI-79)*, pp.598-600, 1979.
- [43] S.T.Barnard and M.A.Fischler, "Computational stereo," *Computing Survey*, Vol.14, No.4, pp.553-572, 1982.
- [44] W.E.L.Grimson, "Computational experiments with a feature based stereo algorithm," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.7, No.1, pp.17-34, 1985.
- [45] 大田友一, 正井康之, 池田克夫, "動的計画法によるステレオ画像の区間対応法", 信学論 (D), Vol.J68-D, No.4, pp.554-561, 1985.
- [46] G.Medioni and R.Nevatia, "Segment-based stereo matching," *Computer Vision Graphics and Image Processing*, Vol.31, pp.2-18, 1985.
- [47] K.Homma and K.S.Fu, "A stereo vision method on region segmentation," *IEEE Workshop on Visual Languages*, pp.14-19, 1984.
- [48] D.Marr and T.Poggio, "A theory of human stereo vision," *Proc. the Royal Society of London*, Vol.B204, pp 301-328, 1979.
- [49] S.T.Barnard and W.B.Thompson, "Disparity Analysis of Images," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.2, No.4, pp.333-340, 1980.
- [50] H.H.Baker, "Edge based stereo correlation," *Proc. Image Understanding Workshop (IUW-88)*, pp.168-175, 1988.

- [51] B.D.Lucas and T.Kanade, "An iterative image registration technique with an application to stereo vision," *Proc. International Joint Conference on Artificial Intelligence (IJCAI-81)*, pp.674-679, 1981.
- [52] J.Weng, N.Ahuja, and T.S.Huang, "Two-View Matching," *Proc. 2nd International Conference on Computer Vision (ICCV'88)*, pp.64-73, 1988.
- [53] H.H.Baker and T.O.Binford, "Depth from Edge and Intensity Based Stereo," *Proc. International Joint Conference on Artificial Intelligence (IJCAI-81)*, pp. 631-636, 1981.
- [54] N.Ayache and F.Lustman, "Fast and Reliable Passive Trinocular Stereovision," *Proc. 1st International Conference on Computer Vision (ICCV'87)*, pp.422-427, 1987.
- [55] H.S.Lim and T.O.Binford, "Structural correspondence in stereo vision," *Proc. Image Understanding Workshop (IUW-88)*, Vol.2, pp.794-808, 1988.
- [56] 佐藤 淳, 富田文明, "テクスチャー領域の両眼立体視", 情処研報, CV-54-5, 1988.
- [57] Y.Ohta and T.Kanade, "Stereo by intra-and-inter-scanline search using dynamic programming," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.7, No.2, pp.139-154, 1985.
- [58] 渡辺正規, 大田友一, 池田克夫, "3眼ステレオ視による奥行き決定精度の改善", 情処論, Vol.27, No.11, pp.1086-1094, 1986.
- [59] M.Okutomi and T.Kanade, "A multiple-baseline stereo," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.15, No.4, pp.353-363, 1993.
- [60] 松浦友彦, 佐藤清秀, 中村裕一, 大田友一, "隠れ検出可能な多眼ステレオ法", 信学論 (D-II), J80-D-II, No.6, pp.1432-1440, 1997.
- [61] 渡辺正規, 大田友一, 池田克夫, "画像理解のための並列処理機構", 信学論 (D-II), Vol.J73-D-II, No.6, pp.820-829, 1990.
- [62] 渡辺正規, 大田友一, "複数ステレオ対応探索アルゴリズムの協調的統合", 信学論 (D-II), Vol.J73-D-II, No.10, pp.1666-1676, 1990.
- [63] 富田文明, 高橋裕信, "ステレオ画像の境界線表現に基づく対応", 信学論 (D), Vol.J71-D, No.6, pp.1074-1082, 1988.
- [64] 石山豊, 富田文明, "等輝度線のステレオ視による曲面の復元", 信学論 (D-II), Vol.J77-D-II, No.9, pp.1673-1679, 1994.

- [65] D.Walts, "Generating semantic descriptions from drawings of scenes with shadows, in *The Psychology of Computer Vision*," ed. P.Winston, pp.19-92, McGraw Hill, New York, 1975.
- [66] M.Herman and T.Kanade, "Incremental reconstruction of 3D scenes from multiple complex images," *Artificial Intelligence*, Vol.30, No.3, pp.289-341, 1986.
- [67] Le Bras-Mehlman E. L., M.Schmitt, O.D.Faugeras, and J.D.Boissonnat, "How the Delaunay triangulation can be used for representing stereo data," *Proc. 2nd International Conference on Computer Vision (ICCV'88)*, pp.54-63, 1988.
- [68] D.D.Morris and T.Kanade, "Image-Consistent Surface Triangulation," *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR2000)*, Vol.1, pp.332-338, 2000.
- [69] 富田文明, "3次元物体の境界線に基づく記述の生成と認識", *信学論 (D)*, Vol.J71-D, No.2, pp.343-349, 1988.
- [70] Y.Yakimovsky, "A System for Extracting Three-Dimensional Measurements from a Stereo Pair of TV Cameras," *Computer Graphics and Image Processing*, Vol.7, No.2, pp.195-210, 1978.
- [71] D.B.Gennery, "Stereo-Camera Calibration," *Proc. Image Understanding Workshop (IUW-79)*, pp.101-107, 1979.
- [72] S.Ganapathy, "Decomposition of Transformation Matrices for Robot Vision," *Proc. International Conference on Robotics and Automation (ICRA'84)*, pp.130-139, 1984.
- [73] R.Y.Tsai, "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses," *IEEE Journal of Robotics and Automation*, Vol.RA-3, No.4, pp.323-344, 1987.
- [74] R.Hartley, "In defense of the eight-point algorithm," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.19, No.6, pp.580-593, 1997.
- [75] L.B.Wolff, "Accurate Measurement of Orientation from Stereo Using Line Correspondence," *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'89)*, pp.410-415, 1989.
- [76] 高橋裕信, 富田文明, "ステレオカメラのセルフキャリブレーション", *情処論*, Vol.31, No.5, pp.650-659, 1990.

- [77] 蚊野 浩, 金出 武雄, ”任意のカメラ配置におけるステレオ視とステレオカメラ校正”, 信学論 (D-II), Vol.J79-D-II, No.11, pp.1810-1818, 1996.
- [78] 高橋裕信, 富田文明, ”ロボットビジョンのためのカメラキャリブレーション”, 日本ロボット学会誌, Vol.10, No.2, pp.177-184, 1992.
- [79] T.Ueshiba and F.Tomita, ”Plane-based Calibration Algorithm for Multi-camera Systems via Factorization of Homography Matrices,” *Proc. 9th International Conference on Computer Vision (ICCV'03)*, Vol.II, pp.966-973, 2003.
- [80] 植芝俊夫, 富田文明, ”平面パターンを用いた複数カメラシステムのキャリブレーション”, 情処論: コンピュータビジョンとイメージメディア, Vol.43, No.SIG11(CVIM5), pp.89-99, 2003.
- [81] P.F.Sturm and S.J.Maybank, ”On Plane-Based Camera Calibration: A General Algorithm, Singularities, Applications,” *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'99)*, Vol.1, pp.432-437, 1999.
- [82] Z.Zhang, ”Flexible Camera Calibration by Viewing a Plane from Unknown Orientations,” *Proc. 7th International Conference on Computer Vision (ICCV'99)*, Vol.1, pp.666-673, 1999.
- [83] Z.Zhang, ”A Flexible New Technique for Camera Calibration,” *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.22, No.11, pp.1330-1334, 2000.
- [84] H.Malm and A.Heyden, ”Stereo Head Calibration from a Planar Object,” *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'01)*, Vol.2, pp.657-662, 2001.
- [85] O.D.Faugeras, ”What can be seen in three dimensions with an uncalibrated stereo rig?,” *Proc. European Conference on Computer Vision (ECCV'92)*, Vol.588, pp.563-578, 1992.
- [86] Z.Zhang, R.Deriche, O.D.Faugeras, and Q.T.Luong, ”A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry,” INRIA technical report, No.2273, 1992.
- [87] L.Robert, M.Buffa, and M.Hebert, ”Weakly-Calibrated Stereo Perception for Rover Navigation,” *Proc. International Conference on Computer Vision (ICCV'95)*, pp.46-51, 1995.

- [88] J.Weng, P.Cohen and M.Hemiou, "Camera Calibration with Distortion Models and Accuracy Evaluation," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.14, No.10, pp.965-980, 1992.
- [89] G.P.Stein, "Lens Distortion Calibration Using Point Correspondences," *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'97)*, pp.602-608, 1997.
- [90] M.T.Ahmed and A.A.Farag, "Differential Methods for Nonmetric Calibration of Camera Lens Distortion," *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR2000)*, Vol.2, pp.477-482, 2000.
- [91] R.Swaminathan and S.K.Nayar, "Calibration of Wide-Angle Lenses and Polycameras," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.22, No.10, pp.1172-1178, 2000.
- [92] F.Devemay and O.D.Faugeras, "Straight lines have to be straight - Automatic calibration and removal of distortion from scenes of structured environments," *Proc. Machine Vision and Applications*, Vol.13, pp.14-24, 2001.
- [93] 石山豊, 角保志, 河井良浩, 植芝俊夫, 富田文明, "セグメントベーストステレオにおける対応候補探索", *映像情報メディア学会誌*, Vol.52, No.5, pp.723-728, 1998.
- [94] H.Takahashi and F.Tomita, "Planarity Constraint in Stereo Matching," *Proc. International Conference on Pattern Recognition (ICPR'88)*, pp.446-449, 1988.
- [95] M. Yachida, Y. Kitamura, and M. Kimachi, "Trinocular vision: New approach for correspondence problem," *Proc. International Conference on Pattern Recognition (ICPR'86)*, pp.1041-1044, 1986.
- [96] N.Ayache and F.Lustman, "Trinocular stereo vision for robotics," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.13, No.1, pp.73-85, 1991.
- [97] S.C.Park, M.K.Park, and M.G.Kang, "Super-resolution image reconstruction: a technical overview," *IEEE Signal Processing Magazine*. Vol.12, No.3, pp.21-36, 2003.
- [98] J.D.van Ouwerkerk, "Image Super-resolution survey," *Image and Vision Computing*, Vol.24, pp.1039-1052, 2006.
- [99] S.Borman, and R.L.Stevenson, "Super-Resolution from Image Sequences-A Review," *Proc. the 1998 Midwest Symposium on Circuits and Systems*, 1998.

- [100] B.Zitova and J.Flusser, "Image registration methods: a survey," *Image and Vision Computing*, Vol.21, No.11, pp.977-1000, 2003.
- [101] C.A.Glasbey and K.V.Mardia, "A review of image warping methods," *Journal of Applied Statistics*, Vol.25, No.2, pp.155-171, 1998.
- [102] S.Baker and T.Kanade, "Limits on super-resolution and how to break them," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.24, No.9, pp.1167-1183, 2002.
- [103] Z.Lin and H.Y.Shum, "Fundamental limits of reconstruction-based superresolution algorithms under local translation," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.26, No.1, pp.83-97, 2004.
- [104] T.Gotoh and M.Okutomi, "Direct Super-Resolution and Registration Using Raw CFA Images," *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'04)*, Vol.2, pp.II-600-607, 2004.
- [105] S.Farsiu, M.D.Robinson, M.Elad, and P.Milanfar, "Fast and robust multiframe super resolution," *IEEE Trans. on Image Processing*, Vol.13, No.10, pp.1327-1344, 2004.
- [106] S.Farsiu, D.Robinson, M.Elad, and P.Milanfar, "Robust shift and add approach to super-resolution," *Proc. SPIE Conf. Applications of Digital Signal and Image Processing*, pp.121-130, 2003.
- [107] S.Farsiu, D.Robinson, M.Elad, and P.Milanfar, "Advances and challenges in super-resolution," *Int. J. of Imaging Systems and Technology*. Vol.14, No.2, pp.47-57, 2004.
- [108] M.Elad and Y.Hel-Or, "A fast super-resolution reconstruction algorithm for pure translational motion and common space-invariant blur," *IEEE Trans. on Image Processing*, Vol.10, No.8, pp.1187-1193, 2001.
- [109] N.K.Bose and N.A.Ahuja, "Superresolution and Noise Filtering Using Moving Least Squares," *IEEE Trans. on Image Processing*. Vol.15, No.10, pp.2239-2248, 2006.
- [110] 金出 武雄, コンラッド ポールマン, 森田 俊彦, "因子分解法による物体形状とカメラ運動の復元", *信学論 (D-II)*, Vol.J76-D-II, No.8, pp.1497-1505, 1993.
- [111] S.Ullman, "The Interpretation of Structure from Motion," *Proc. the Royal Society of London*, pp.405-426, 1979.

- [112] W.N.Martin and J.K.Aggarwal, "Volumetric Descriptions of Objects from Multiple Views," IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI), Vol.5, No. 2, pp.150-158, 1983.
- [113] T.Kanade, P.Rander, and P.J.Narayanan, "Virtualized Reality: Constructing Virtual Worlds from Real Scenes," IEEE MultiMedia, Vol.4, No.1, pp.34-47, 1997.
- [114] 山崎公俊, 友納正裕, 坪内孝司, 油田信一, "カメラを搭載した移動ロボットによる物体のオンライン3次元モデリング", 映像情報インダストリアル, December, pp.61-68, 2004.
- [115] M.A.Fischler, and R.C.Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography," Communication of the ACM, Vol.24, pp.381-395, 1981.
- [116] 富田文明, "3次元ビジョンの実現に向けて", 日本ロボット学会誌, Vol.12, No.8, pp.1124-1127, 1994.
- [117] 加藤顕剛, "1秒で物体の位置と姿勢が認識できる眼がほしい", 日本ロボット学会誌, Vol.12, No.1, pp.15-16, 1994.
- [118] 松山隆司, 浅田尚紀, 青山正人, 浅津英樹, "再帰トラス結合アーキテクチャにおける並列対象認識のためのデータレベル並列プロセスの構成", 情処論, Vol.36, No.10, pp.2310-2320, 1995.
- [119] 許昭倫, 佐藤完, 岡崎信一郎, 藤田善弘, "一次元プロセッサアレイに基づくリアルタイム画像処理システムの開発環境", 情処学研報, 97-CVIM-106, pp.195-202, 1997.
- [120] M.Rygor, S.Pollard, and C.Brown, "Multiprocessor 3D vision System for pick and place," Image and Vision Computing, Vol.9, No.1, pp.33-38, 1991.
- [121] B.R.Modayur and L.G.Shapiro, "Fast parallel object recognition," *Proc. International Conference on Pattern Recognition (ICPR'94)*, Vol.III, pp.284-289, 1994.
- [122] <http://www.csm.ornl.gov/pvm/>
- [123] <http://www.planetary.or.jp/setiathome/>
- [124] D.G.Lowe, "Fitting parameterized three-dimensional models to images," IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI), Vol.13, No.5, pp.441-450, 1991.

- [125] D.Koller, K.Daniilidis, and H.H.Nagel, "Model-based object tracking in monocular image sequences of road traffic scenes," *International Journal of Computer Vision*, Vol.10, No.3, pp.257-281, 1993.
- [126] 野村 由司彦, セーハンドゥスイト, 藤井 省三, "2次元投影像からの3次元物体の位置・姿勢推定方法", *信学論 (D-II)*, Vol.J77-D-II, No.1, pp.101-107, 1994.
- [127] 山根 定章, 泉 正夫, 福永 邦雄, "モデルベースに基づく物体の位置・姿勢推定", *信学論 (D-II)*, Vol.J79-D-II, No.2, pp.165-173, 1996.
- [128] 角 保志, 石山 豊, 富田 文明, "ステレオビジョンシステムのためのモデルベーストアプローチによる自由曲面体の実時間運動追跡", *信学論 (D-II)*, Vol.J84-D-II, No.8, pp.1693-1700, 2001.
- [129] I.Biederman, "Recognition-by-Components: A Theory of Human Image Understanding," *Psychological Review*, Vol.94, No.2, pp.115-147, 1987.
- [130] 柳井啓司, "一般物体認識の現状と今後", *情処論*, Vol.48, No.SIG16(CVIM19), pp.1-24, 2007.
- [131] M.B.Clowes, "On Seeing things," *Artificial Intelligence*, Vol.2, No.1, pp.79-116, 1971.
- [132] J.M.Tanenbaum and H.G.Barrow, "Experiments in Interpretation Guided Segmentation," *Artificial Intelligence*, Vol.8, pp.241-274, 1977.
- [133] Y.Ohta, "Knowledge-based Interpretation of Outdoor Natural Color Scenes," Pitman Advanced Publishing Program, Boston, 1985.
- [134] B.Draper, R.Collins, J.Brolio, A.Hanson, and E.Riseman, "The Schema System," *International Journal of Computer Vision*, Vol.3, No.2, pp.209-250, 1989.
- [135] T.Matsuyama and V.S.Hwang, "SIGMA:A knowledge-based aerial image understanding system," Plenum Press, New York, 1990.
- [136] D.Marr, "Vision," W.H.Freeman and Co., 1982.
- [137] R.A.Brooks, "Model-Based Three-Dimensional Interpretations of Two-Dimensional Image," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.5, No.2, pp.140-150, 1983.

- [138] A.R.Pope, "Model-Based Object Recognition: A Survey of Recent Research," Technical Report TR-94-04, Univ. of British Columbia, Computer Science Department, 1994.
- [139] L.Stark and K.Bowyer, "Achieving Generalized Object Recognition through Reasoning about Association of Function to Structure," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.13, No.10, pp.1097-1104, 1991.
- [140] T.M.Strat and M.A.Fischler, "Context-Based Vision: Recognizing Objects Using Information from Both 2-D and 3-D Imagery," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.13, No.10, pp.1050-1065, 1991.
- [141] 松山隆司, 尾崎正治, "LIVE : トップダウンセグメンテーションのための画像エキスパートシステム", *情処論*, Vol.27, No.2, pp.191-204, 1986.
- [142] 長谷川純一, 久保田浩明, 鳥脇純一郎, "LIVE : サンプル図形提示方式による画像処理エキスパートシステム IMPRESS", *信学論 (D)*, Vol.J70-D, No.11, pp.2147-2153, 1987.
- [143] V.Clement and M.Thonnat, "A Knowledge-Based Approach to Integration of Image Processing Procedures," *Computer Vision Graphics and Image Processing*, Vol.57, No.2, pp.166-184, 1993.
- [144] M.J.Swain and D.H.Ballard, "Color Indexing," *International Journal of Computer Vision*, Vol.7, No.1, pp.11-32, 1991.
- [145] 村瀬洋, V.V.Vinod, "ヒストグラム特徴を用いた高速物体探索法", *信学論 (D-II)*, Vol.J70-D-II, No.9, pp.2035-2042, 1998.
- [146] K.Kashino, T.Koizumi, and H.Murase, "A Quick Search Method for Audio and Video Signals Based on Histogram Pruning," *IEEE Trans. Multimedia*, Vol.5, No.3, pp.348-357, 2003.
- [147] H.Murase and S.K.Nayar, "Visual Learning and Recognition of 3-D Objects from Appearance," *International Journal of Computer Vision*, Vol.14, No.9, pp.5-24, 1995.
- [148] S.Ullman, "High-level Vision," The MIT Press, 1996.
- [149] R.Fergus, P.Perona, and A.Zisserman, "A Object Class Recognition by Unsupervised Scale-Invariant Learning," *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'03)*, pp.264-271, 2003.

- [150] R.Fergus, P.Perona, and A.Zisserman, "A Sparse Object Category Model for Efficient Learning and Exhaustive Recognition," *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'04)*, pp.380-387, 2004.
- [151] L.Fei-Fei, R.Fergus, and P.Perona, "A Bayesian Approach to Unsupervised One-Shot Learning of Object Categories," *Proc. International Conference on Computer Vision (ICCV'03)*, pp.1134-1141, 2003.
- [152] B.Leibe, A.Leonardis, and B.Schiele, "Combined object categorization and segmentation with an implicit shape model," *Proc. ECCV Workshop on Statistical Learning in Computer Vision*, 2004.
- [153] D.Crandall and D.Huttenlocher, "Weakly supervised learning of part-based spatial models for visual object recognition," *Proc. European Conference on Statistical Learning in Computer Vision*, pp.I:16-29, 2006.
- [154] C.Schmid and R.Mohr, "Local Greyvalue Invariants for Image Retrieval," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol,19, No.5, pp.530-535, 1997.
- [155] D.G.Lowe, "Object recognition from local scale-invariant features," *Proc. International Conference on Computer Vision (ICCV'99)*, pp.1150-1157, 1999.
- [156] K.Mikolajczyk and C.Schmid, "A performance evaluation of local descriptors," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol,27, No.10, pp.1615-1630, 2005.
- [157] G.Csurka, C.Bray, C.Dance and L.Fan, "Visual categorization with bags of keypoints," *Proc. ECCV Workshop on Statistical Learning in Computer Vision*, pp.1-22, 2004.
- [158] N.Otsu and T.Kurita, "A new scheme for practical flexible and intelligent vision systems," *Proc. IAPR Workshop on Computer Vision*, pp.431-435, 1988.
- [159] 大津展之, 栗田多喜夫, 関多巖, "パターン認識—理論と応用", 朝倉書店, 1996.
- [160] T.Hofmann, "Unsupervised Learning by Probabilistic Latent Semantic Analysis," *Machine Learning*. Vol.43, pp.177-196, 2001.
- [161] D.Blei, A.Ng, and M.Jordan, "Latent Dirichlet Allocation," *Journal of Machine Learning Research*. Vol.3, pp.993-1022, 2003.

- [162] L.Fei-Fei and P.Perona, "A Bayesian Hierarchical Model for Learning Natural Scene Categories," *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'05)*, pp.524-531, 2005.
- [163] D.G.Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, Vol.60, No.2, pp.91-110, 2004.
- [164] F.Jurie and B.Triggs, "Creating Efficient Codebooks for Visual Recognition," *Proc. International Conference on Computer Vision (ICCV'05)*, pp.I:604-610, 2005.
- [165] D.Comaniciu and P.Meer, "Mean Shift: A Robust Approach toward Feature Space Analysis," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.25, No.5, pp.603-619, 2002.
- [166] J.Zhang, M.Marszalek, S.Lazebnik, and C.Schmid, "Local Features and Kernels for Classification of Texture and Object Categories: A Comprehensive Study," *International Journal of Computer Vision*, Vol.73, No.2, pp.213-238, 2007.
- [167] A.Holub, M.Welling, and P.Perona, "Combining Generative Models and Fisher Kernels for Object Recognition," *Proc. International Conference on Computer Vision (ICCV'05)*, pp.136-143, 2005.
- [168] S.E.Palmer, E.Rosch, and P.Chase, "Canonical Perspective and the perception of objects," *Attention and Performance*, Vol.9, pp.135-151, 1981.
- [169] L.Roberts, "Machine perception of three-dimensional information, in Tippett," J. (Ed.), *Optical and Electro-Optical Information Processing*, 159-197, MIT Press, 1965.
- [170] D.G.Lowe, "Perceptual Organization and Visual Recognition," Kluwer Academic Publishers, Boston, 1985.
- [171] R.Mohan and R.Nevatia, "Perceptual Organization for Scene Segmentation and Description," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.14, No.6, pp.616-635, 1992.
- [172] T.Kanade, "Recovery of the three-dimensional shape of an object from a single view," *Artificial Intelligence*, Vol.17, pp.409-460, 1981.
- [173] 吳, 浅田, 辻, "図形中に存在する斜め対称軸の抽出", *信学論 (D)*, Vol.J70-D, No.3, pp.553-559, 1987.
- [174] S.A.Friedberg, "Finding axes of skewed-symmetry," *Computer Vision Graphics and Image Processing*, Vol.34, No.2, pp.138-155, 1986.

- [175] K.Sugimoto, H.Terasaki, and F.Tomita, "Detection of Skewed-Surface-Symmetrical Primitives for Grasping Unknown Objects," *Proc. Asian Conference on Computer Vision (ACCV'95)*, pp.II-757-761, 1995.
- [176] H.Terasaki and T.Hasegawa, "Motion planning for intelligent manipulations by sliding and rotating operations with parallel two-fingered grippers," *Proc. International Conference on Intelligent Robots and Systems (IROS'94)*, pp.119-126, 1994.
- [177] Y. Matsumoto, H.Terasaki, K.Sugimoto, and T.Arakawa, "A portable three-dimensional digitizer," *Proc. International Conference on Recent Advances in 3-D Digital Imaging and Modeling (3DIM'97)*, pp.197-204, 1997.
- [178] Y.Matsumoto, K.Fujimura, and T.Kitamura, "CyberModeler: A Compact 3D Scanner Based on Monoscopic Camera," *Proc. SPIE Electronic Imaging (EI 1999)*, Vol.3640, pp.2-10, 1999.
- [179] T.Terauchi, Y.Oue, and K.Fujimura, "A Flexible 3D Modeling System Based on Combining Shape-from-Silhouette with Light-Sectioning Algorithm," *Proc. International Conference on 3-D Digital Imaging and Modeling (3DIM 2005)*, pp.196-203, 2005.
- [180] <http://www.web3d.org/>
- [181] 原川健一, "コンピュータを指の動きで操作するシステム", 画像ラボ, Vol.11, No.4, pp.29-33, 2000.
- [182] 金谷健一, "画像理解-3次元認識の数理-", 森北出版, 1990.
- [183] Y. Kameda, M.Minoh, and K.Ikeda, "Three Dimensional Pose Estimation of an Articulated Object from its Silhouette Image," *Proc. Asian Conference on Computer Vision (ACCV'93)*, pp.612-615, 1993.
- [184] N. Shimada and Y.Shirai, "3-D Hand Pose Estimation and Shape Model Refinement from a Monocular Image Sequence," *Proc. International Conference on Virtual Systems and Multi Media (VSMM'96)*, pp.423-428, 1996.
- [185] I.P.Vladimir, S.Rajeev, and T.S.Huang, "Visual Interpretation of Hand Gestures for Human-Computer Interaction : A Review," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.19, No.7, pp.677-695, 1997.
- [186] C.Maggioni and B.Kammerer, "Gesture Computer-History, Design and Applications In Computer Vision for Human-Machine Interaction," Cambridge University Press, 1998.

- [187] Q.Delamarre and O.D.Faugeras, "Finding pose of hand in video images : a stereo-based approach," *Proc. of The 3rd Conf.on Face and Gesture Recognition*, pp. 585-590, 1998.
- [188] A.Utsumi, J.Ohya, and R.Nakatsu, "Multiple-Hand-Gesture Tracking using Multiple Cameras," *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'99)*, pp. 473-478, 1999.
- [189] P.A.Walcott and T.J.Ellis, "The Localization of Objects in Real World Scenes Using Color," *Proc. Asian Conference on Computer Vision (ACCV'95)*, pp.III-243-247, 1995.
- [190] A.Utsumi, J.Ohya, and R.Nakatsu, "Real-time Hand Gesture Recognition System," *Proc. Asian Conference on Computer Vision (ACCV'95)*, pp.II-249-253, 1995.
- [191] D. G. Lowe, "Robust Model-based Motion Tracking Through the Integration of Search and Estimation," *International Journal of Computer Vision*, Vol.8, No.2, pp.113-122, 1992.
- [192] Z.Zhang and O.D.Faugeras, "Three-dimensional Motion Computation and Object Segmentation in a Long Sequence of Stereo Frames," *International Journal of Computer Vision*, Vol.7, No.3, pp.211-241, 1992.
- [193] M.Kass, A.Witkin, and D.Terzopoulos, "Snakes: active contour models," *International Journal of Computer Vision*, Vol.1, No.4, pp.321-331, 1988.
- [194] K. Fujimura, N. Yokoya, and K. Yamamoto, "Motion Analysis of Nonrigid Objects by Active Contour Models Using Multiscale Images," *Systems and Computers in Japan*, Vol.25, No.2, pp.81-91, 1994.
- [195] P.J. Besl and N.D.Mckay, "A Method for Registration of 3-D Shapes," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.14, No.2, pp.13-17, 1994.
- [196] M.Hebert and T.Kanade, "Outdoor scene analysis using range data," *Proc. IEEE International Conference on Robotics and Automation (ICRA'86)*, pp.1426-1532, 1986.
- [197] 安田稔, 山口幸也, 福島邦彦, 長田昌次郎, "視覚系受容野の電子回路モデル", *信学論 (54-C)*, No.6, pp.514-521, 1971.
- [198] 石川正俊, "マトリクス状センサからの出力分布の中心の位置と総和の検出方法", *計測自動制御学会論文集*, 19-5, pp.381-386, 1983.

- [199] S.C.Liu and J.Harris, "Generalized Smoothing Networks in Early Vision," *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'89)*, pp.184-191, 1989.
- [200] J.L.Wyatt, D.L.Standley, and W.Yang, "The MIT Vision Chip Project : Analog VLSI Systems for Fast Image Acquisition and Early Vision Processing," *Proc. IEEE International Conference on Robotics and Automation (ICRA'91)*, pp.1330-1335, 1991.
- [201] A.M.Chiang and M.L.Chuang, "A CCD Programmable Image Processor and its Neural Network Applications," *IEEE J. Solid-State Circuits*, Vol.25, No.6, pp.1891-1894, 1991.
- [202] M.Mathur, S.C.Liu, and H.T.Wang, "Analog Neural Networks for Focal-Plane Image Processing," *SPIE 1242*, pp.141-151, 1990.
- [203] C.Mead and M.Mahowald, "A Silicon Model of Early Visual Processing," *Neural Network*, vol.1, pp.91-97, 1988.
- [204] 船津英一, 倉持宗一郎, 長渕洋二, 鹿毛裕史, 坂下徳美, 村尾文秀, 久間和夫, "射影機能を持つ人口網膜 LSP", *映像情報メディア学会誌*, Vol.55, No.10, pp.1329-1336, 2001.
- [205] 小室孝, 並木明夫, 石川正俊, "小特集 次世代イメージセンサ技術, イメージセンサ技術 対機能な目", *映像情報メディア学会誌*, Vol.56, No.3, pp.356-359, 2002.
- [206] 小室孝, 鏡慎吾, 石川正俊, "ビジョンチップのための動的再構成可能な SIMD プロセッサ", *信学論 (D-II)*, Vol.J86-D-II, No.11, pp.1575-1585, 2003.
- [207] T.Miyasaka and K.Koyama, "Image sensing and processing by a bacteriorhodopsin-based artificial photoreceptor," *Appl. Opt.*, Vol.32, No.31, pp.6371-6379, 1993.
- [208] 杉山幸宏, 杉本和英, "移動物体像の特徴抽出処理に光感応蛋白質を利用するインテリジェントイメージセンサ", *画像電子学会誌*, 第31巻, 第5号, pp.762-767, 2002.
- [209] 杉山幸宏, 田岡峰樹, 杉本和英, "光感応蛋白質の配向フィルムを用いた動物体像抽出", *電子情報通信学会*, D-12-76, pp.264, 2000.
- [210] D. Oesterhelt and W. Stoecknius, "Isolation of the Cell Membrane of Halobacterium halobium and Its Fractionation into Red and Purple Membrane," in *Methods in Enzymology*, eds. S. Fleischer and L. Packer (Academic Press, New York) Vol.31, Pt. A, pp.667-678, 1974.

- [211] Y.Sugiyama, T.Inoue, M.Ikematsu, M.Iseki, and T.Sekiguchi, "Controlling the Orientation of Purple Membrane Fragments on an Air/Water Interface by a New Method of Direct Electric Field Application during Purple Membrane Spreading," *Japanese Journal of Applied Physics*, 36, pp.5674-5679, 1997.
- [212] R. Korenstein and B. Hess, "Analysis of Photocycle and Orientation in Thin Layers," in *Methods in Enzymology*, ed. L. Packer (Academic Press, New York) Vol.88, Pt. I, pp.180-193, 1982.
- [213] M.Dissanayake, P.Newman, S.Clark, H.Durrant-Whyte, and M.Csorba, "A solution to the simultaneous localization and map building (slam) problem," *IEEE Trans. on Robotics and Automation*, Vol.17, No.3, pp.229-241, 2001.
- [214] A.J.Davison, "Real-time simultaneous localisation and mapping with a single camera," *Proc. International Conference on Computer Vision (ICCV'03)*, Vol.2, pp.1403-1410, 2003.
- [215] T.Bailey and H.Durrant-Whyte, "Simultaneous localisation and mapping (SLAM): Part I - The essential algorithms," *IEEE Robotics and Automation Magazine*, Vol.13, No.2, 2006.
- [216] T.Bailey and H.Durrant-Whyte, "Simultaneous localisation and mapping (SLAM): Part II - State of the art," *IEEE Robotics and Automation Magazine*, Vol.13, No.3, 2006.

公表論文リスト

国内論文誌

1. 杉本和英, 富田文明, ”ステレオ画像の境界表現に基づくシーンの解釈”, 信学論 (D-II), vol.J76-D-II No.3, pp.524-533, Mar. 1993.
2. 杉本和英, 富田文明, ”ステレオによる幾何モデリング”, 日本ロボット学会誌, Vol.15, No.3, pp.431-438. Apr. 1997.
3. 大上靖弘, 杉本和英, 北村徹, 角保志, 富田文明, ”ネットワーク型並列計算環境における物体認識”, 信学論 (D-II), vol.J82-D-II No.12, pp.2307-2315, Dec. 1999.
4. 杉山幸宏, 杉本和英, ”移動物体像の特徴抽出処理に光感応蛋白質を利用するインテリジェントイメージセンサ”, 画像電子学会誌, 第31巻, 第5号, pp.762-767. Sep. 2002.

査読付論文 (国際学会)

1. K.Sugimoto, H.Takahashi, and F.Tomita, ”Scene Interpretation Based on Boundary Representations of Stereo Images,” *Proc. IEEE International Conference on Pattern Recognition (ICPR'88)*, pp.155-159, Nov, 1988.
2. K.Sugimoto, H.Takahashi, and F.Tomita, ”Integration and Interpretation of Incomplete Stereo Scene Data,” *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'91)*, pp.683-685, Jun. 1991.
3. K.Sugimoto and F.Tomita, ”Boundary Segmentation by Detection of Corner, Inflection and Transition Points,” *Proc. IEEE Workshop on Visualization and Machine Vision (WVMV'94)*, pp.13-17, Jun. 1994.
4. K.Sugimoto and F.Tomita, ”Detection of Skewed-Symmetrical Shape,” *Proc. International Conference on Image Processing (ICIP'94)*, pp.696-700, Nov. 1994.
5. K.Sugimoto, H.Terasaki, and F.Tomita, ”Detection of Skewed-Surface-Symmetrical Primitives for Grasping Unknown Objects,” *Proc. Asian Conference on Computer Vision (ACCV'95)*, pp.II-757-761, Dec. 1995.

6. K.Sugimoto and F.Tomita, "Shape Coding by Detecting Dominant Points on Digital Curves," *Proc. International Conference on Virtual Systems and Multi Media (VSMM'96)*, pp.153-158, Sep. 1996.
7. K.Sugimoto and F.Tomita, "How to detect dominant points on 3-D Curves," *Proc. European signal processing conference (EUSIPCO'98)*, pp.2177-2179, Oct. 1998.
8. K.Sugimoto and F.Tomita, "Rigid Hand Motion Tracking based on a Generic Model for Manipulating 3-D Models of Objects," *Proc. IEEE Pacific-Rim Conference on Multimedia (PCM2000)*, pp.202-205, Dec. 2000.

その他の公表論文

1. 杉本和英, 高橋裕信, 富田文明, "水平3眼ステレオ画像のB-r e pに基づく対応", 第36回情報処理全国大会, 5V-7, Mar. 1988.
2. 杉本和英, 高橋裕信, 富田文明, "ステレオによるオクルージョンの検出", 第36回情報処理全国大会, 5V-9, Mar. 1988.
3. 杉本和英, 富田文明, "ステレオによるシーンの解釈", 情報処理学会コンピュータビジョン研究会, Jul. 1988.
4. 杉本和英, 富田文明, "ステレオによるシーンの解釈", 第37回情報処理全国大会, Sep. 1988.
5. 杉本和英, 富田文明, "ステレオの移動による世界モデルの構築と更新", 第39回情報処理全国大会, Nov. 1989.
6. 杉本和英, 富田文明, "モデルベースステレオビジョンによるロボットハンドの認識", 第39回情報処理全国大会, Nov. 1989.
7. 杉本和英, 富田文明, "不完全なステレオデータの統合と解釈", 電子情報通信学会春季全国大会, Mar. 1990.
8. 杉本和英, 富田文明, "ハンドアイロボットのためのステレオビジョンと物体認識", 第41回情報処理全国大会, Sep. 1990.
9. 杉本和英, 富田文明, "移動ロボットのためのステレオビジョンによる環境の幾何モデリング", 情報処理学会コンピュータビジョン研究会, Sep. 1990.
10. 杉本和英, 富田文明, "移動ロボットのためのステレオビジョンによる環境の幾何モデリングにおける問題点", 電子情報通信学会秋季全国大会, Oct. 1990.

11. 杉本和英, 高橋裕信, 富田文明, ”ハンドアイロボットのためのステレオビジョン”, 情報処理学会コンピュータビジョン研究会, Jan. 1991.
12. 杉本和英, 富田文明, ”曲面物体の形状プリミティブの発見”, 第43回情報処理学会全国大会, Oct. 1991.
13. 杉本和英, 富田文明, ”図形中の斜め対称軸の発見 (斜め対称図形の発見)”, 電子情報通信学会秋季大会, Sep. 1992.
14. 杉本和英, 富田文明, ”境界線のセグメンテーション”, 第47回情報処理学会全国大会, Oct. 1993.
15. 杉本和英, 富田文明, ”Boundary Segmentation by Detecting Corner, Inflection and-Transition Points”, RWC SYMPOSIUM, Jun. 1994.
16. 杉本和英, 富田文明, ”輪郭線の屈曲点, 変曲点, 遷移点の検出”, 画像の認識・理解シンポジウム, MIRU '94, Jul. 1994.
17. 角保志, 石山豊, 植芝俊夫, 河井良浩, 杉本和英, 富田文明, ”画像の境界表現のデータ構造とインタフェース”, 第49回情報処理学会全国大会, 第2分冊, No.3F-7, pp.123-123, 1994.
18. 杉本和英, 富田文明, ”斜め対称プリミティブの統合と解釈”, 第49回情報処理学会全国大会, Sep. 1994.
19. 杉本和英, 寺崎肇, 富田文明, ”Detection of skewed-surface-symmetrical primitives for grasping unknown objects,” RWC Symposium, Jun. 1995.
20. 杉本和英, 富田文明, ”斜め面对称プリミティブの未知対象物体把持への利用”, 第51回情報処理学会全国大会, Sep. 1995.
21. 杉本和英, 富田文明, ”特徴点検出に基づく形状記述手法”, 電子情報通信学会総合大会, Mar. 1997.
22. 杉本和英, 富田文明, 荒川勉, ”3次元点列上の特徴点検出”, 第55回情報処理学会全国大会, 5AB-10, Sep. 1997.
23. 杉本和英, 富田文明, ”平面パターンの対応に基づく画像間のエピポーラ幾何と3次元形状復元”, 電子情報通信学会総合大会, Mar. 1998.
24. 杉本和英, 荒川勉, 富田文明, ”3次元点列上の特徴点検出”, 画像の認識・理解シンポジウム, MIRU '98, Jul. 1998.

25. 杉本和英, 富田文明, "3D Curve Partitioning for Modeling of Objects with Smooth Surfaces," RWC Symposium, Jun. 1998.
26. 松本 幸則, 藤村 恒太, 杉本 和英, 北村 徹, "動的参照点生成によるカメラ位置計算手法とその三次元モデリングへの応用", 情報処理学会第 57 回全国大会講演論文集, No.5, 5P-12, Sep, 1998.
27. 杉本和英, 大上靖弘, 北村徹, "手指動作による立体モデルのハンドリング", 人工知能学会全国大会 (第 13 回) , Jun. 1999.
28. 杉本和英, 松本幸則, 北村徹, "Hand Motion Estimation based on a Generic Model," RWC Symposium, Jan. 2000.
29. 杉山幸宏, 田岡峰樹, 杉本和英, "光感応蛋白質の配向フィルムを用いた動物体像抽出", 電子情報通信学会, D-12-76, pp.264, 2000.
30. 杉本和英, 大上靖弘, 富田文明, "モノクロ単眼動画像からの手指の検出と運動追跡", 画像の認識・理解シンポジウム, MIRU2000, Jul. 2000.
31. J. Grant, 寺内 智哉, 藤村 恒太, 杉本 和英, "An Advanced 3D Voxel Data Decompression and Voxel-packing Algorithm Implemented in RealSculptorGL," 2002 年電子情報通信学会総合大会講演論文集, No.D-11-149, 2002.