

視覚障害者のための視覚情報提示システム に関する研究

河井 良浩

目次

1. 序章.....	1
1.1. 研究の背景.....	1
1.1.1. コミュニケーション支援.....	2
1.1.2. 行動支援.....	5
1.2. 研究の目的.....	7
1.3. 本論文の構成.....	8
2. 触覚・聴覚を利用した視覚情報提示システム.....	11
2.1. はじめに.....	11
2.2. システム概要.....	12
2.3. 図面処理.....	13
2.4. 情報提示.....	17
2.4.1. 触覚ディスプレイ.....	17
2.4.2. 音声出力.....	21
2.5. 評価実験.....	22
2.5.1. 方法.....	22
2.5.2. 結果.....	24
2.5.3. 検討.....	25
2.6. まとめ.....	27
3. セグメントベースドステレオ法による3次元情報の取得.....	28
3.1. はじめに.....	28
3.2. 3次元情報取得処理.....	29
3.2.1. カメラキャリブレーション.....	29
3.2.2. 対象物計測手法.....	30
3.2.2.1. セグメントベースドステレオ法.....	31
3.2.2.2. 相関法ステレオ.....	31
3.2.3. 対象物検出.....	32
3.3. セグメントベースドステレオ法.....	34
3.3.1. 対応探索の特徴.....	34
3.3.2. 境界線表現.....	35
3.3.3. 2D B-Rep から 3D B-Rep へ.....	36
3.3.3.1. 対応候補探索.....	37
3.3.3.2. ペアの連結.....	38
3.3.3.3. 対応パス探索.....	39

3.3.3.4.	対応評価	40
3.3.3.5.	対応評価の統合	43
3.3.3.6.	3次元復元	44
3.3.4.	実験結果	45
3.3.5.	セグメントベースドステレオのまとめ	49
3.4.	まとめ	50
4.	3次元点群情報の統合	51
4.1.	はじめに	51
4.2.	レンジデータ	52
4.2.1.	レンジファインダ	52
4.2.2.	データ形式	53
4.3.	3次元データ統合方法	54
4.3.1.	データ取得	55
4.3.2.	前処理	55
4.3.3.	領域対応	56
4.3.4.	初期移動パラメータ計算	57
4.3.5.	移動パラメータ修正	57
4.3.6.	データ統合	59
4.4.	実験	61
4.4.1.	結果	61
4.4.2.	統合実験結果の考察	68
4.5.	まとめ	69
5.	対話型触覚ディスプレイによる3次元情報提示システム	71
5.1.	はじめに	71
5.2.	システムの概要と特徴	72
5.3.	触覚ディスプレイシステム	72
5.3.1.	ステレオカメラシステム(入力部)	73
5.3.2.	触覚ディスプレイ(出力部)	74
5.3.3.	合成音声出力装置(出力部)	75
5.4.	3次元情報の取得	76
5.5.	表示モード	76
5.6.	評価実験	77
5.6.1.	実験1:3次元触覚ディスプレイ	77
5.6.2.	実験2:物体形状表現モード	77
5.6.3.	実験3:全体像モード	78
5.6.4.	参考実験:触地図認識	79

5.7. 考察.....	79
5.8. まとめ.....	80
6. 音響表現を利用した3次元情報提示システム.....	82
6.1. はじめに.....	82
6.2. システム概要.....	83
6.2.1. ステレオカメラシステム.....	84
6.2.2. 3次元音響システム.....	85
6.2.3. 3次元視覚情報処理.....	85
6.2.4. システム制御.....	86
6.2.5. 視覚情報の音響表現.....	86
6.3. 音像定位実験.....	87
6.3.1. 予備実験.....	87
6.3.2. 実験1:水平面内の音像定位.....	88
6.3.3. 実験2:全方位の音像定位.....	89
6.3.4. 実験3:環境音の影響.....	91
6.4. 考察.....	92
6.5. まとめ.....	94
7. 結論.....	95
7.1. まとめ.....	95
7.2. 今後の展開.....	96
謝辞.....	99
参考文献.....	100
付録.....	106
A1. カメラキャリブレーション.....	106
A1.1.はじめに.....	106
A1.2.参照平面を用いた単体カメラの較正.....	106
A1.2.1. Zhang の手法.....	106
A1.2.2. カメラ画像間の対応点.....	107
A1.3.複数カメラの較正.....	108
A1.3.1. 参照平面の観測.....	109
A1.3.2. 幾何学的拘束からの初期値推定と最適化.....	109
A1.3.3. アルゴリズムのまとめ.....	110
A1.4.実データを用いた較正実験.....	111
A1.4.1. 簡易3次元計測システムの構築.....	111
A1.4.2. 共通視野を持たないカメラの較正.....	112
A1.5.まとめ.....	114

A2. 輝度補正	115
A2.1. はじめに	115
A2.2. 処理手順	115
A2.3. 補正結果	117
A2.4. まとめ	119
A3. ステレオ視における3次元復元の高精度化	120
A3.1. はじめに	120
A3.2. 計測誤差	121
A3.3. 平面拘束による誤差の軽減	122
A3.3.1. 標準カメラモデル座標系への変換	122
A3.3.2. 対応点計算	123
A3.3.3. 3次元復元時の問題点	123
A3.3.4. 平面拘束に基づく対応点補正	124
A3.4. 実験	125
A3.4.1. 実験 1: 近距離にある単純図形	126
A3.4.2. 実験 2: 一般的なシーン	127
A3.5. まとめ	129
発表論文	130
論文誌	130
国際会議	132
著書・刊行物	136
講演会・研究会	137
特許	146
プレス発表	147

1. 序章

序章では、視覚障害者のための支援システムに関する研究開発の発展の歴史的背景の説明を通じて、コミュニケーションや行動を支援するために重要な要素技術、および、視覚情報提示システムの研究開発の必要性を言及し、本研究の目的と特徴を述べる。

1.1. 研究の背景

日本における視覚障害者の数は約31万人(成人,平成18年厚労省調べ)と言われている。人間が外界から得る情報の多くが視覚によるものであることを考えると、この機能を失っている場合、日常生活・社会生活において多くの不便と不利益を被ることになる。したがって、視覚障害者のこのような不自由な状況を少しでも解消し、社会復帰を促進するためには、視覚情報を残存感覚である触覚と聴覚を利用して提示する支援技術、システム化に関する研究開発が必要である。そのため、視覚支援システムの研究、開発は古くからいろいろなされている [1] [2] [3] [4] [5]。

視覚障害者といっても個人によりその障害度は異なる。大まかな分類をすると、まったく視覚機能を失っている全盲と、程度の差はあるが視覚機能を有している弱視に分けられる。そして視覚機能を失った時期により、先天盲と後天盲に分類される。ゆえに、視覚障害者といっても千差万別であり、個人一人一人に合わせた視覚機能支援が必要である。

視覚の代わりとなる代行感覚として利用できるものは、触覚と聴覚である。触覚は、皮膚が網膜と同様な2次元的な広がりを持っているため、空間配置、物体形状を把握する場合に適しているが、対象が近くに存在するもので、触るのに適していなければならず、また、接触部分が限られるため、一度には局所情報しか得ることができない。一方、聴覚は、空間状況把握に適しており、リアルタイムで全方向からの情報を受け取ることができる。しかし、そのすべての情報を処理しているわけではなく、意識的に注意を向けることで情報を選択する必要がある。また、音声などによる概念の提示にも優れた感覚である。

失われた視覚情報を、触覚・聴覚によって情報提示する手法の確立、システム化は、健常者と遜色ないレベルでの、視覚障害者の社会復帰を促進し、自立した生活への手助けに必要不可欠である。視覚障害者向けの視覚支援システムの研究開発は古くからいろいろとなされており [1] [2]、最初に、視覚障害者のための支援システムに関し、これまでの研究開発・製品化されたものの経緯を簡単に述べる。何を支援するためのシステムと考えた場合、主に「コミュニケーション」支援と「行動」支援に分類でき、この区分で支援システムを述べる。

読書器としては、光学的に読み取った文字パターンを指先で読みとるオプタコンが有名であるが、近年、一方、文字以外の2次元的情報、すなわち地図や回路図などの図については、わずかに末田らの図面認識補助装置 [6]があるだけである。これはカメラで撮影された図がハードコピーの形で触覚ディスプレイ上に呈示されるもので、その認識はすべて使用者にゆだねられている。

1.1.1. コミュニケーション支援

コミュニケーションの手段としては、主に文字、図、映像などの視覚情報と音声、音響などの聴覚情報を介して行うものがある。聴覚情報に関しては、ラジオや音楽鑑賞など、視覚障害者にとっても特別な変換を要することなく利用可能であるが、視覚情報に関しては何らかの変換を施さなければならない。例えば、本などは人に読んでもらうか、点字本を利用することが可能ではあるが、他の人の手を介する必要がある。また、点字本はその数が絶対的に足りておらず、最新の本などは読むことができない。そしてその大多数はボランティアによる点訳である。障害者も人の手を煩わすことなく自分一人で行いたいという自立願望があり、この自立を促進するシステムが必要である。情報の種類別に支援技術の経緯を記す。

(1) 文字情報

点字は1825年にBrailleによって発明された。文字によるコミュニケーションが可能なものであり、図1-1に示すように2×3点の突起で表わされる。これを指先で走査することで、その突起を触覚し、文字情報を読み取るものである。点字は覚えるのにかなり労力が必要であり、視覚障害者のすべてが必ずしも点字を読み書きできるわけではない。

そこで、文字を点字としてではなく、形として触覚で捉えるシステムとしてオプタコン(キャノン)が開発された(図1-2)。オプタコンは、小型カメラで撮影した画像(文字)を指先大の触覚ディスプレイで表示するシステムで、文字の形が触覚ディスプレイのピンの振動に変換され、指先で読み取るものである。後天盲の人で一般の文字(墨字)の方が慣れている人には適しており、また、墨字を点字に変換する作業を伴わない利点がある。

さらに、発展型として画像処理技術を利用し、スキャナなどで読み取った文字情報を認識して、合成音声で出力する装置(Kurzweil Reading Machine [7])も製品化されている。

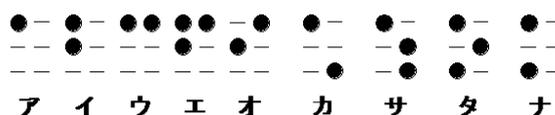


図 1-1 点字例(凸面)



図 1-2 オプタコン(キャノン株式会社)

また、近年の電子書籍の進展には目を見張るものがあり、データが電子データ化されているため、読み上げソフト・サービスとの組合せで、これまでになく手軽に書籍情報にアクセスしやすくなってきている。ただし、著作権保護のために読み上げができない形で提供されていたり、対応する端末が少なかったり、一部ルビの読み上げがおかしいなどの問題もあつたりする状況であり、まだまだ電子書籍が視覚障害者に身近なものとなるには、多くの越えなければならない問題がある。

(2) 図面・画像情報

図面で視覚障害者が利用できるものとして触図がある。触図は点字の図面版とも言えるもので、点や線が凸に浮き上がって図を形成している。これはレイズライタ(ペンで描いた線が浮き上がる特殊な用紙を使用)や立体コピー(特殊なトナーで、熱を加えると発泡して盛り上がる)などで作成できるが、一度描いてしまうと訂正ができないなどの問題がある。

触覚ディスプレイはこの電子版であり、2次元上にピンを多数配置し、その凹凸や振動で図面情報を表示する触覚デバイスである。触覚ディスプレイは、構造が複雑になりがちで、解像度があまり高くないのが欠点ではあるが、図面情報との親和性は高い。末田らは、図面認識補助装置 [6]として、カメラで撮影された図がハードコピーの形で触覚ディスプレイ上に呈示されるものであるが、その認識はすべて使用者にゆだねられている。この他にも図面情報の触覚による提示システムの市販品もある。しかしながら、その図面情報の処理はせいぜい画像の二値化や細線化レベルであり、形そのものの提示しかできず、複雑な形状の表示は難しい。

触覚と同様に使用できる感覚である聴覚は、視覚の次に受容情報量が多いため、画像の表現を音響で行う研究も多く行われている [8]。Meijer は画像の濃淡情報を音の高低、上下方向を周波数の高低、左右方向を時間走査で表現して視覚障害者にそのシーンを認識させる研究を行った(図 1-3) [9]。小林らは、普及のためには安価なシステムでなければならないという観点から、特別な装置を用いることなく簡便にシステムを実現するために、中央の走査線のみを画像処理してエッジ部分を検出し、その位置をステレオサウンドで表現している [10]。上下の情報は被験者自身が頭を振ることにより得るようにした。

現段階では、映像情報に関しては高度な処理はされていないが、今後はシーンの解釈を行い、その情報を3次元の音響で表現する方向で研究が進んでいる。

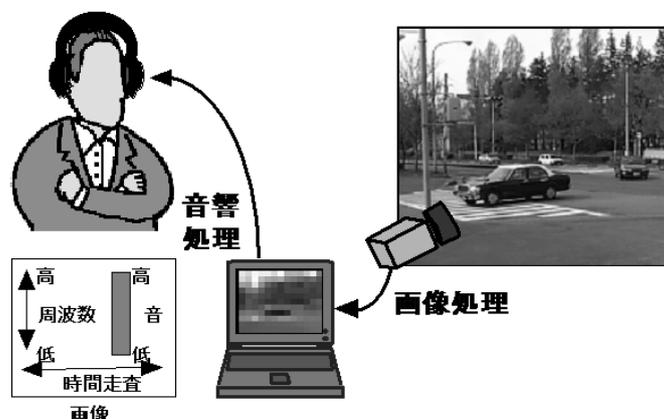


図 1-3 音響による視覚情報表示システム

(3) 映像情報

一方、動画などの映像情報においては、NHK などのテレビ放送の一部が、放映シーンの状況などを説明する、音声放送には対応しているが、現状では視覚障害者に十分な情報を提供しているとは言いがたい状況である。映像情報を提示する実用システムに関しては現状ではほとんど存在しないと言える。

(4) 立体情報

触覚で3次元の形状、配置情報を得るデバイスとして、医療福祉機器研究所において、「盲人用3次元情報表示装置」が開発された [8]。ピンは64×64の千鳥格子配列で、ピン直径2.5 mm、間隔3 mm、高さ0～10 mm(可変)であり、かなり複雑な物体の3次元物体の形状表示が可能なディスプレイであったが、システムの大きさやメンテナンス性に難があり、実験装置の域を出ていないものであった。

(5) 電子機器操作情報

1980年代から急速に普及し始めたコンピュータへのアクセス情報に関しては、80年代のテキストベースのMS-DOSでは、その文字情報を合成音声で出力する方法で、視覚障害者はコンピュータを使用することができた。しかし、アイコンをマウスでクリックする操作が基本のGraphic User Interface(GUI)が主流になり、逆に視覚障害者はコンピュータを使いづらい状況になった。そこで、GUIのシステムに音声や音響を使用してアクセスできるシステムの研究開発が行われてきた(図1-4) [3] [4] [5] [11] [12]。Windows95の時代から、一部のソフトではあるが、音声で読み上げるソフトウェア95Readerが開発・販売され、現バージョンではWindows XPまで対応している。また、NEDOの委託により日立が3次元音響生成技術を用いて、視覚によって認識しているWindows画面上のカーソル位置やアイコン、ウィンドウなどを音響、および、音声で表現するシステムが開発された [13]。しかし、これらのシステムにおいてはOSやアプリケーションから得られる情報を音声・音響化しているだけであり、画像そのものに対する処理は行われていない。



図 1-4 GUI アクセスシステム

近年、液晶画面とタッチパネルで操作をするタブレット・スマートホンが小型情報端末として急速に普及しているが、触覚感のない操作法になったため、視覚障害者にとっては(5)のGUIの普及時と同様にこれらの機器を扱いづらくなった問題が生じている。しかしながら、iPad、iPhoneなどのiOS上で動作するSiriなど、音声認識を用いた機器操作、検索なども行えるようになっており、必ず

しも使えない機器ではなく、視覚障害者においても使える機器であり、これらの機器を所有・使用したいと要望もある。いろいろな操作を行うためには、触覚フィードバック機能を組み込んだインタフェースへの改良が必要であり、その研究開発が望まれている [14].

1.1.2. 行動支援

一方、行動支援に関してもいろいろな手段があるが、ここでは視覚障害者が街中を自立して歩行することができるための支援システムについて述べる。介護者とともに行動をするのが安全ではあるが、前述のように障害者の自立を望む声を満たしてはいないし、いつも介護者の手当てがつくわけではない。自立行動を補助する支援システム別に記す。

(1) 白杖, 盲導犬

視覚障害者の行動を支援する機器としては、白杖と盲導犬が広く普及している。白杖はその手軽さから最も普及しており、視覚障害者の行動範囲を広げている。盲導犬はさらに優れた能力を持っている。しかし、白杖は認識できる範囲が限られており、盲導犬はその数や生き物であるが故の問題もある。そこで、工学的な支援機器が開発されてきた。

(2) 電子白杖, ソニックガイド

電子白杖は、白杖の発展型であり、杖に超音波センサや近赤外線センサを組み合わせることで、杖が届く少し先の範囲の障害物を検知し、杖の握り部分が振動して情報を提示するシステムである。一方、歩行ガイドロボット [68]や、歩行支援システム [69] [70]など、ロボット技術を応用した研究開発もいろいろなされているが、システムの複雑さ、大きさ、価格などの面で実用化には至っていないのが現状である。

ソニックガイドは、超音波計測により得た障害物までの距離情報を音の高さで表現するシステムであり、図 1-5 のような眼鏡の形状をしている。白杖より大きな観測エリアとなるが、この機器を使いこなすにはかなりの訓練を要する。

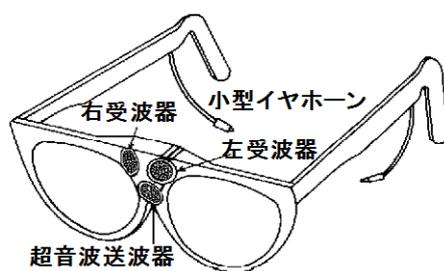


図 1-5 ソニックガイド

(3) GPS システム

福祉分野においても、カーナビゲーションで利用されているGPS(全地球測位システム)を応用する研究が行われている [3] [4] [5] [15] [16] [17]. 視覚障害者の歩行誘導だけでなく、徘徊癖のある高齢者の捜索にも利用できる。同様な簡易システムとして、携帯電話を利用した位置検索はすでに幾つかの自治体で導入されている。図 1-6 にシステムの概要を示す。GPS は现阶段では軍事上の理由で測位精度が悪く、そのままでは歩行補助に使える精度とは言えない。そこで、基準

局を設け、同時に受信した測位データから精度を高める DGPS を用いている。これにより測位誤差は約 1~2m になり、地図データを併用することで、ほぼ正確に位置を特定できる。また、電波を受信できない場所もあるため、加速度センサ、地磁気ジャイロ、振動ジャイロなどを利用して移動の推定を行う。視覚障害者が介護者を伴わないで知らない場所に行く場合、あらかじめ出かける場所までの道順を触地図などで覚えておくのが普通であるが、このシステムにより一人で気軽に外出できることを目標としている。現段階では、位置は特定できるが、使用者の周りの局所的な情報は得ることができない。そこで、カメラで得たその場の状況を携帯電話などで遠方にいるオペレータに送り、指示を仰ぐことなどが必要である。完全自立のためには、見ている方向の障害物検知、看板などの文字情報の読み取り、出入り口の検出などの局所的な情報を映像情報技術の利用により得る必要がある。今後はこのような画像処理・認識モジュールの研究開発が望まれる。

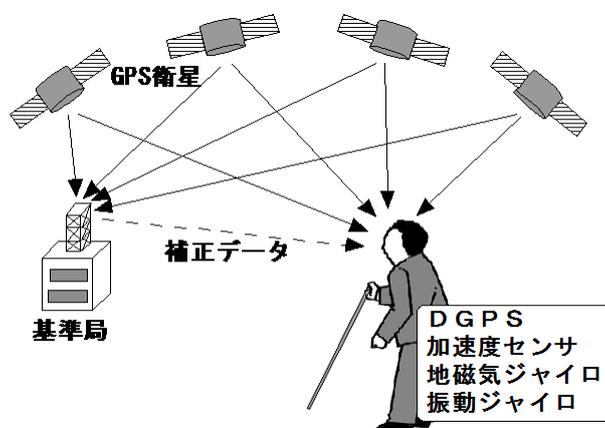


図 1-6 DGPS を利用した歩行補助システム

(4) インフラの整備

視覚障害者が周りの状況を把握しやすいようにインフラを整備することも重要な歩行補助手段である [18]。インフラ整備で実用化されているものとして、点字ブロック、点字パネル、視覚障害者用信号機、音響案内などが上げられる。点字ブロックはかなり普及しており、新築の際にはほぼ設置されている。しかし、狭い路地などは対応が遅れているし、デザインを優先して機能が生かされていない場合もある。視覚障害者用信号機は普及しているが、まだすべてのものが対応しているわけではなく、横断途中で赤信号に変わると音情報が途絶えてしまうため、危険が伴う問題も残されている。地下鉄の出入り口などに設置されている音響案内の普及率もまだまだ低い。

インフラ整備として、街中の主な物体に電子的なタグを取り付け、視覚障害者が携帯する装置と連動して、その情報を音声で提示するシステムがいろいろ研究開発されている [5] [19] [20]。

駅構内での情報(券売機、改札口の位置、階段やホームの端などの情報)を単方向ではなく双方向で案内情報のやりとりができる誘導案内システムが鉄道総研で研究されている [20]。

この分野でも映像情報は基本的には使われていない。現実問題として、歩行時の大きな障害物となる他の歩行者などすべてのものにタグを付けることはできないので、より高度な歩行補助を行う場合には、映像情報取得による局所的な情報を提供する必要がある。

視覚障害者に対する支援システムの現状、研究状況を簡潔に説明した。この分野における映像情報技術は現状ではまだ十分に利活用されていると言えない。特に画像解析、理解など高度な処理技術を利用した実用システムはほとんどない。現在実用化が進んでいるシステムは、画像処理を用いて視覚情報を得るものより、インフラ整備による情報発信を多くし、その情報を利用して視覚障害者に提示するものが多いのが現実である。これは、映像情報技術の未熟さとこの分野への適用の遅れが一因であると思われる。しかし逆を言えば、視覚障害者の日常生活・社会生活の向上のためには、映像情報技術が必要不可欠であると言える。今後は、福祉分野に最新の实用技術が導入され、視覚障害者のニーズに答えられる機器の研究開発が促進される必要である。

1.2. 研究の目的

前節で述べたように、視覚障害者が日常生活をより便利に過ごすことができ、積極的に社会へ参加できるようにするために、本研究では、視覚障害者のコミュニケーション支援、行動支援を行うために必要な、視覚情報取得・処理技術、触覚・聴覚を用いたインタフェース技術の確立、および、これらを利用した支援システムの構築・提供を目的とする。視覚障害者には、様々な障害レベルの方がおられるが、本研究では、主になんらかの原因で視力を失った全盲の人(後天盲)を対象とする。

この目的を達成するためには、

- ・膨大な視覚情報から効率よく必要な情報を取得する技術
- ・残された感覚である触覚と聴覚を最大限活用し、視覚情報を提示する技術

を確立し、視覚障害者の失われた視覚情報を補う支援システムの構築を行う必要がある。以下に、必要不可欠な技術を具体的に挙げる。

前者の視覚情報取得技術として、近年急速に進歩しているコンピュータビジョン技術を積極的に活用する。その中でも特に、使用者が欲する視覚情報をいかに効率よく取得するかがキーポイントである。そのためには、環境や対象物の2次元・3次元情報を取得する方法の確立が重要である。情報取得と言っても、単に形状情報を得るだけでなく、対象とする情景にユーザが必要とする対象の有無の判別、有る場合はどこにあり、どのような状況であるかまでを得る必要がある。本研究では、2次元の図面情報理解支援、3次元の対象物の形状・配置理解支援、3次元のリアルタイムで変化する環境の理解支援を具体的な情報支援対象と定め、2次元の図面の解釈技術、対象シーンの3次元形状情報取得技術、対象目的の3次元シーンからの検出技術に関して、技術確立のために研究開発を行う。

後者の触覚と聴覚を活用した提示技術としては、残存感覚をいかにうまく利用して、視覚情報を伝えるかということが重要であり、各感覚に適した情報への変換、提示の仕方を確立する必要がある。各感覚への情報変換に関しては、触覚・聴覚の特徴を踏まえたものでなければならない。触覚の長所は、空間配置、物体形状を把握する場合に適している点であり、聴覚の長所は空間状況把握に適しており、音声などによる概念の提示にも優れている点である。逆に、触覚は、一度には局所情報しか得ることができない点、聴覚は、形状情報を伝えるには適しておらず、全方位すべての

情報を処理できるわけではない点が短所である。このような長短所を踏まえ、触覚情報としては、対象物の形状や、対象シーンでの配置関係の表示、聴覚情報としては、全体・個別の概要情報、リアルタイムでの状況表示を主目的とした情報変換、情報提示機能、および、これらを具現化できるハードウェアシステムが必要かつ重要である。また、これらの情報を一方的に視覚障害者に提示するだけでなく、使用者が知りたい情報を取得できるようにすることも大切であり、対話的なやり取りが可能なシステムであることも、理解度を深めるためには重要な機能である。

そこで、本研究では、上記した2つの技術の確立とともに、それら技術を適用した情報提示支援システムを構築し、視覚障害者に実際に評価してもらうことで、システムの妥当性、有効性を確認する。また、評価を通じて得られた知見に基づき、支援システムとして必要な技術を洗い出すことも行う。このように、本研究は、視覚障害者が必要とする視覚情報を、コンピュータビジョン技術を活用して効率よく取得できるようにするとともに、残存感覚である触覚・聴覚情報へ的確に変換し、膨大な情報を生で受けるのではなく、必要に応じて視覚情報を取得できるような支援システムを目指している点が特徴である。

本研究で開発する視覚情報支援システムを利用することにより、視覚障害者が普段の生活で取得するのに困っている視覚情報に対して、より簡便にアクセスできるようになることで、コミュニケーションを円滑に行えるようになったり、自立して行動できたりすることの手助けになることが期待できる。

1.3. 本論文の構成

本論文においては、序章に続いて、視覚障害者用の視覚情報提示支援システムを実現するために必要とされる、視覚情報処理機能の開発、視覚情報を残存感覚である触覚や聴覚を通じて提示する機能の開発、および、これらを実装した視覚情報提示支援システムについて述べる。具体的な本論文の構成は以下に示すとおりである。

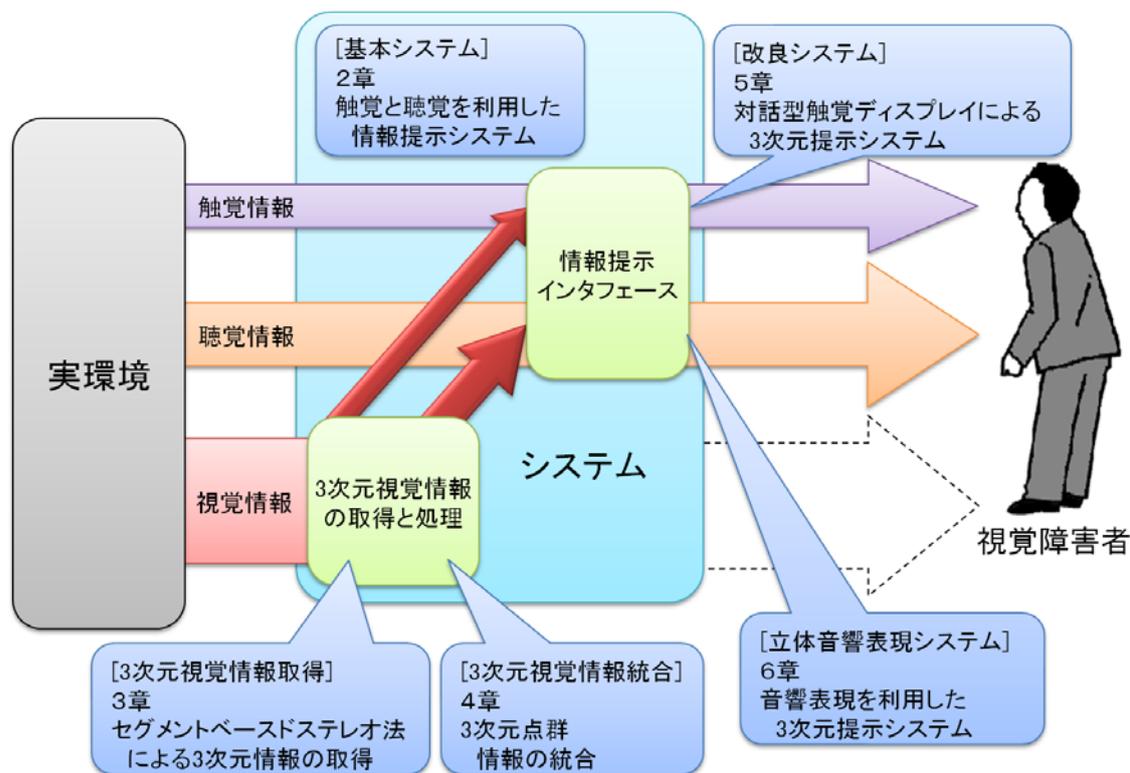


図 1-7 本論文の全体構成図

本章では、視覚障害者用の支援システムの研究の背景、目的、および、構成を説明した。

第 2 章では、基本的な支援システムとして、視覚情報を取得し、変換処理を行い、触覚と聴覚で情報を伝える機能を有する、視覚情報提示支援システムの開発について述べる。2 次元の視覚情報(図面)を、触覚と音声によって提示するシステムである。視覚障害者が触覚のみで図面を理解するのは困難であるため、図面そのものの触覚表示ではなく、パソコンに図面を認識させ、図面の全体情報を触覚で、詳細情報を音声で提示するシステムを提案している。対話的な操作が可能のようにタッチセンサの機能も有する触覚ディスプレイを開発し、対話的操作により視覚障害者が図面を理解する際にかかる負担を軽減できることを示している。

第 3 章では、前章のシステムの研究開発を通じて、視覚障害者が生活していく上で重要な 3 次元空間情報を扱う必要性を感じたため、実環境の 3 次元の視覚情報を得るための方法について述べる。3 次元の視覚情報を得るための一手法として、ステレオ画像解析による、3 次元復元法を開発した。境界線表現で記述されたステレオ画像間の対応探索において、領域の境界を構成する境界線セグメントの連結性に基づいて対応を評価する方法であり、オクルージョンに対して比較的頑健な方法である。また、境界線の類似性に基づく大局的な画像間の対応評価を行うことで、安定した復元結果を得られる。複数の物体が存在する様々なシーンを対象にして 3 次元情報取得実験を行い、手法の有効性を示している。復元された幾何データは単なる 3 次元形状データではなく、3 次元物体の認識や追跡に利用可能な構造を有しているため、応用範囲は広く、視覚障害者が認

識したい物体を検出する際などに利用できるものである。

第 4 章では、第 3 章に引き続き、視覚障害者が必要な 3 次元空間情報取得の際に問題となる、一方向からの観測した視覚情報の不十分さを補うために、複数の観測方向からの情報を統合する手法に関して述べる。取得すべきシーン情報が複雑な場合、一方向からの計測では十分なデータを取得できず、視覚障害者に必要十分な視覚情報を提示することができない場合があるためである。視点方向の位置関係が未知である場合でも、計測データから、オクルージョンの影響が少ない領域を基に初期照合を行い、共通部分全体のずれが最小となるように繰り返し処理をすることで、より正確な位置関係を計算することが可能な方法である。開発した手法を、複雑な形状を有する花などの対象に対して適用し、多視点から計測したデータを統合し、全体の 3 次元形状を正確に得られることを実験により示している。この手法により、視覚障害者が必要とする対象の十分な 3 次元情報を得ることができる。

第 5 章では、第 2 章で行った基本システムの研究開発の経験を基に、第 3 章、第 4 章で開発した 3 次元視覚情報取得・処理技術を適用し、視覚障害者への提示情報量を増やして、必要な 3 次元情報をより得やすいような、対話型触覚ディスプレイによる 3 次元情報提示システムについて述べる。視覚情報として 3 次元の形状を有する物体を対象とし、その視覚情報を 3 次元の触覚情報と音声で提示するシステムである。16×16 のマトリクス状に配置され、6 段階の高さ表現ができる触覚ピンを有し、かつ、タッチセンサの機能を有している 3 次元触覚ディスプレイを開発した。3 次元視覚情報処理技術により取得した物体の立体形状、配置情報を 3 次元触覚ディスプレイに表示し、音声で属性情報を付加することで、視覚障害者がシステムと対話的なやりとりをすることが可能となり、3 次元情報を効率よく取得できる点、空間認知をしやすくなる点を、実験を通じて示している。

第 6 章では、第 5 章で開発したシステムでは伝えにくい、状況が刻々と変化し、空間的な拮りを持つ 3 次元の視覚情報を提示できるように、即応性のある聴覚を利用した 3 次元仮想音響による新たな 3 次元視覚情報提示システムについて述べる。3 次元センサにより取得した 3 次元視覚情報に音響を割り当て、仮想 3 次元空間でその位置・動きを表現する方法を用い、インフラが未整備の場所や情報提供できない対象に対して有用であり、ユーザの周りの状況を提示できる。このうち、3 次元仮想空間での音像定位の特性を評価するために、各種被験者実験を行い、音像定位の問題点、認識率を向上させる方法など、音響による適切な情報提示に関する指針が得られ、視覚障害者は使用に関して多大な学習をすることなく使用できる可能性も示している。

第 7 章では、本研究によって得られた成果を総括し、視覚障害者向けの視覚情報提示支援システムとしての今後の展開を記している。

また、付録資料として、第 3 章、第 4 章での 3 次元視覚情報取得、処理技術を補完する技術として開発を行ったツールの技術に関して説明をしている。付録 A1 では、ステレオカメラの位置姿勢を簡便に計算するための平面パターンを用いたカメラキャリブレーション法、付録 A2 で、ステレオ画像間の輝度調整を自動的に行う方法、付録 A3 でステレオ視においてセグメント間の連結性に基づく平面拘束を利用して、高精度な 3 次元情報を復元する方法に関して述べる。

2. 触覚・聴覚を利用した視覚情報提示システム

視覚障害者支援システムの実現を目指し、視覚情報を取得し、変換処理を行い、触覚と聴覚で情報を伝えるという、情報提示支援の基本となるシステムとして、触覚と音声によって失われた視覚情報を視覚障害者に提示する支援システムについて述べる。視覚障害者が必要とする情報のうち、図面情報を対象とし、触覚情報のみでは図面を理解するのが難しいという問題点を解決するため、図面そのものを触覚表示するのではなく、図面の全体情報を触覚で、詳細情報を音声で提示するシステムを開発した。本システムとの対話的操作により、図面を理解する際に視覚障害者にかかる負担を軽減できることを示す。

2.1. はじめに

第1章の研究の背景で説明したように、人間が外界から得る情報の大部分は視覚によるものであり、文字や図・絵の認識において、視覚は重要な役割を果たしているにも関わらず、その視覚機能を失った人々は、日常生活において多くの不便と不利益を被っている。その情報取得の不十分さを補うための、視覚支援システムの研究、開発が多くなされてきている。

読書器としては、光学的に読み取った文字パターンを指先で読みとるオプタコンなど、製品レベルのものがあるが、文字以外の2次元的情報、すなわち地図や図面などについては、点字と同じように紙に突起で線画表現されている触図が一般的であるが、対応しているものは少なく、また、そのかさばる点も問題となっている。コンテンツが少ない問題を解決するものとして、読み取りたい対象図をカメラで撮影し、点字表示のように多くの触知覚できるピンを有する触覚ディスプレイ上にハードコピーの形で提示されるシステムの研究もあるが、図の認識はすべて使用者にゆだねられており、複雑な内容を有する図の場合、視覚障害者がなかなか理解しづらい問題がある。

今一度視覚障害者に残された感覚である、触覚、聴覚の特性を考えると、皮膚は網膜と同様に2次元の広がりを持つので、触覚は位置関係の提示に適している。しかし、触覚は視覚と比較すると情報の提示量は少なく、空間分解能も劣っている。このため、触覚だけで事物を理解することは、アルファベット文字やカタカナなどの簡単なものの場合には良いが、図面などのような複雑なものになると非常に難しい。一方、音声は概念の提示に優れている。したがって、触覚と聴覚の長所を生かすように、これらの残された感覚をうまく組み合わせ、使用者に視覚情報を効率よく提示することが可能になると考えられる。

そこで、視覚障害者が図面を容易に理解できる情報提示支援システムを提案する。単純に読み取った図面情報を、単純にそのまま視覚障害者に提示するのではなく、画像処理技術により図面の認識を行い、認識結果から触覚に適した情報(図全体の構成)を触覚ディスプレイに表示し、聴覚に適した情報(図を構成する個々の要素名など)は合成音声で与えるシステムである。このシステムにより、使用者は図面そのものの触覚表示ではなく、全体的な情報を触覚で、詳細な情報は聴覚で得られるため図面理解の負担が軽減されると期待できる。つまり、本システムの特徴は、(1)コンピュータによる図面認識というシステムの知能化と、(2)触覚・聴覚という複数のモダリティの最適

利用, 更に, (3)システムとの対話的な操作が可能なことである. このシステムの利用により, 視覚障害者は図面が必要不可欠な分野を学習できるようになる.

本章では, 本システムの概要, 図面処理, 開発した触覚ディスプレイ, 触覚・音声情報の提示方その結果について述べる.

2.2. システム概要

開発したシステム(図 2-1)は①図面入力, ②画像処理, ③情報提示の三つの部分からなり, パーソナルコンピュータ(以下, パソコンと略記)(NEC PC9801 VX21)を中心に構成されている(図 2-2).

- ① 図面入力: 対象図面を CCD カメラ(NEC TL-22P2)で撮影し, 画像をメモリボード(ADS EIP-98)に取り込む.
- ② 画像処理: パソコンで画像を処理し, 図面を認識する.
- ③ 情報提示: 触覚ディスプレイに全体情報を表示し, タッチセンサの機能により指で指定した位置での図面の詳しい情報を音声出力ボード(IO データ PIO-9061A)から出力する.

図面が出力されると, 使用者はまず触覚ディスプレイを指で走査することで図面全体の情報を得る. 指で押している部分の詳しい情報を知りたい場合は, 情報取得キーを押すことでパソコンに指の位置の信号が送られ, 認識結果との比較により音声でその部分の情報が示される. つまり, システムとの対話的な操作を行うことで, 触覚・音声情報が得られる. それらの情報を統合することで図面を理解することができる.

本システムの特徴は次の 3 点である.

- (1) 知能化: 図面そのものを触覚表示することは, 触覚ディスプレイのハード的問題や, 使用者の認識時の負担増などにより効率的でないため, コンピュータに図面を認識させ, 使用者の図面を理解する負担を軽減する.
- (2) 触覚と聴覚の利用: 認識結果を触覚と音声で表示するため, 図面情報が効率良く伝わる.
- (3) 対話的な操作: 触覚ディスプレイはタッチセンサの機能もあり, 情報を提示するだけでなく, 使用者がさわっている位置を知ることができる. ゆえに, システムからの一方的情報の受身的な取得ではなく, 対話的な操作が可能となり, 図面の理解が容易となる.



図 2-1 システム(概観)

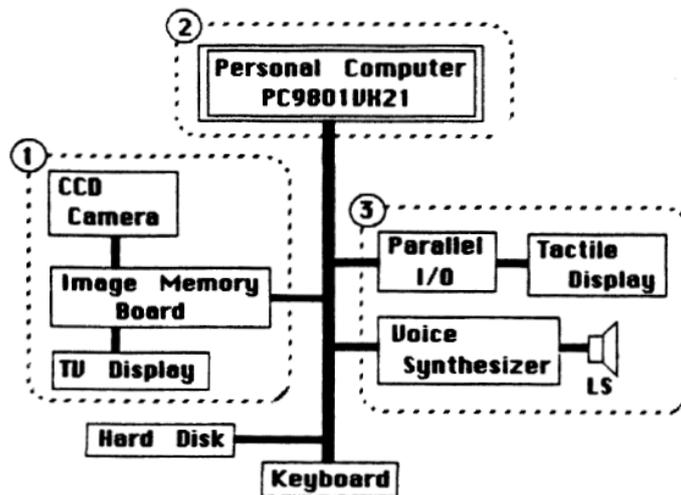


図 2-2 ブロックダイアグラム

2.3. 図面処理

図面には、地図、回路図、プラント図など数多くのものであり、任意の図面を対象とし、認識することは困難であるため、本システムで扱う対象図面を限定する。本研究では、二値の線画で表記されるが、解像度の関係で、そのままの形での触図表現が難しい図面の一例として、電気回路図を対象図面に選んだ。対象は一般の本に掲載されている回路図として、今回は「交流電源、直流電源、コイル、コンデンサ、抵抗、端子」で構成されていて、図面中に文字がないものとする。

図面の処理手順は図 2-3 に示すとおりであり、(a)の電気回路図を入力すると、線図形(b)に変換して図面認識(c)を行い、(d)のように触覚ディスプレイに適した形に変換される。(d)の黒丸●は素子(分岐点を含む)で、配線部分が○(斜線)で表されている。以下に処理の流れを示す。

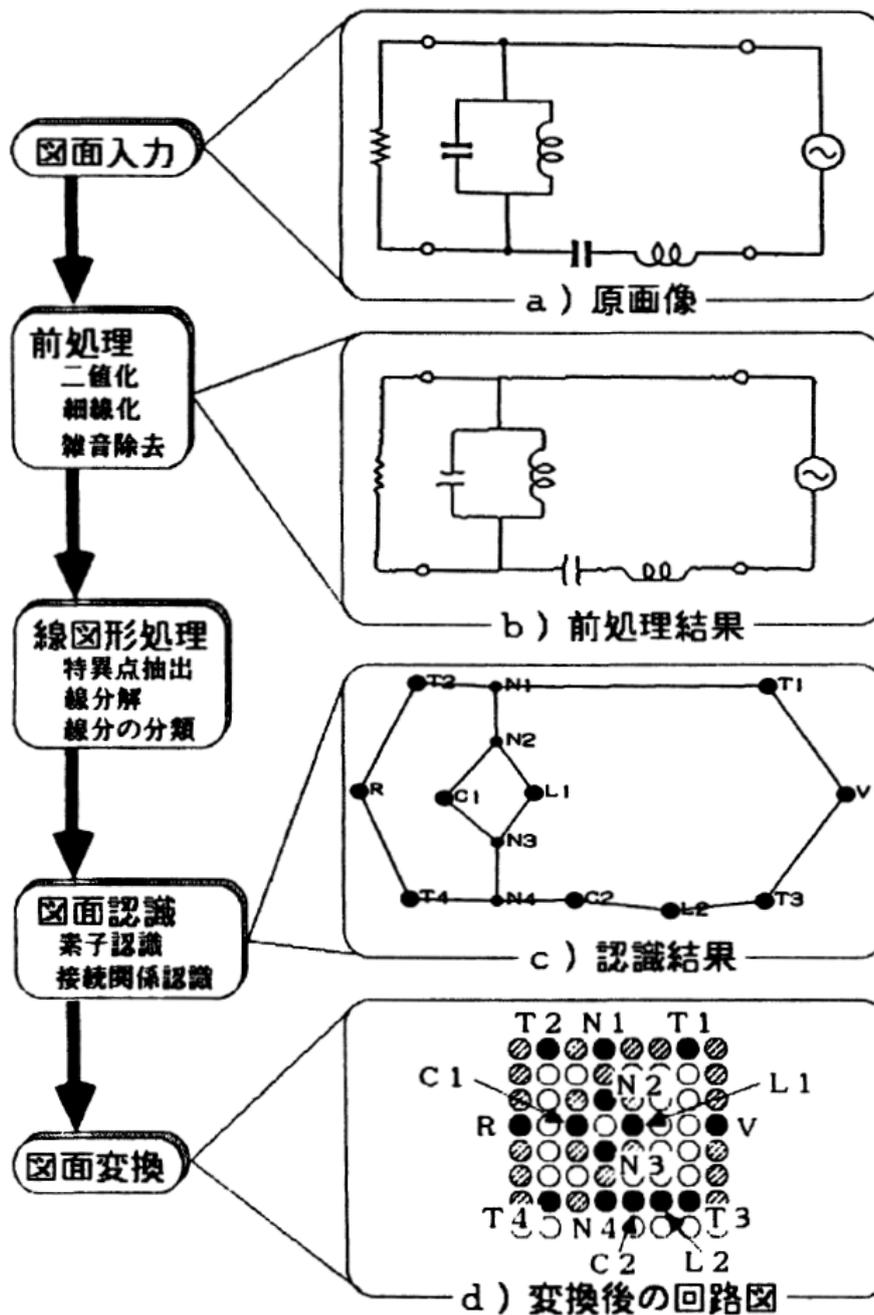


図 2-3 図面の処理手順

【図面入力】

図面は CCD カメラにより 256×240 画素, 64 濃度レベルでパソコンに取り込まれる(図 2-3 a)). 照明条件, 絞り設定, 焦点設定などを適切に行い, 鮮明な画像を撮影する.

【前処理】

対象図面が白黒画像であり, そのまま処理するのではなく, 前処理として二値化, 細線化, 雑音

除去を行い、画像データをきれいな線図形にして処理を簡単にする。以下、順番に処理内容を示す(図 2-3 b))。

【線図形処理】

前処理により得られた線図形を線分に分解して、それがどのような種類の線分であるかを求める。まず、線図形を線分に分解し、特徴量を求める。例えば、図 2-3 b)の線図形は図 2-4 のように、34 個の特異点と 40 の線分に分解することができる。

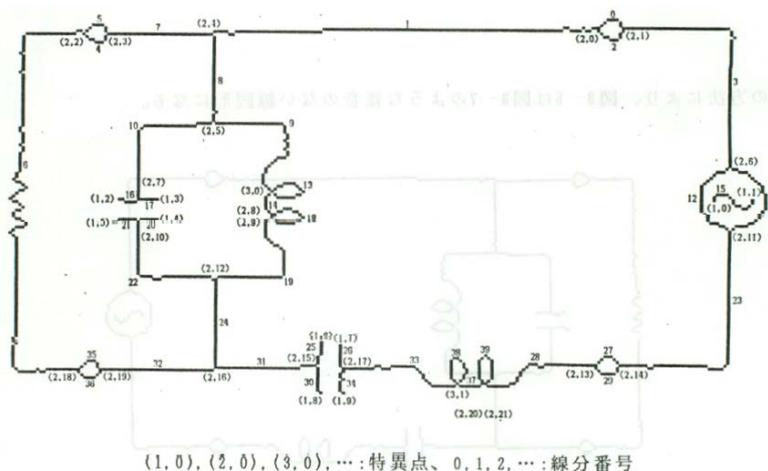


図 2-4 線分解

その後、分解された線分をその特徴量によって直線、半円など 9 種類に分類する(表 2-1)。

表 2-1 分類された線分リスト

番号	種類	番号	種類	番号	種類	番号	種類
0	半円	10	直線の直角	20	直線	30	直線
1	直線	11	半円	21	直線	31	直線
2	半円	12	半円	22	直線の直角	32	直線
3	直線の直角	13	楕円	23	直線の直角	33	配線
4	半円	14	弧	24	直線	34	直線
5	半円	15	波線	25	直線	35	半円
6	抵抗	16	直線	26	直線	36	半円
7	直線	17	直線	27	半円	37	弧
8	直線	18	楕円	28	配線	38	楕円
9	配線	19	配線	29	半円	39	楕円

【図面認識】

線分の種類が認識されたので、それらを組み合わせてどのような素子があり、その接続関係がどうなっているかを調べ、図面の構成を求める。具体的には、線分は表 2-2 に示すようにそれぞれ素子、配線の一部であり、これらを組み合わせることで素子、配線が形成される。

表 2-2 線分の種類とその対応部分

線分の種類	対応部分
直線部分	配線, コンデンサー, 直流電源
直線の直角部分	配線
波線部分	交流電源
半円部分	交流電源, 端子
円部分	端子
楕円部分	コイル
弧部分	コイル
抵抗部分	抵抗
配線部分	配線

表 2-2 の関係と線分の接続関係より, 線分にラベル付けを行い, 回路図を再構築する(図 2-3 c)).

【図面変換】

これまでの処理で, 回路図の構成は認識されたが, 16×16 (もしくは 8×8) 個の刺激ピンをもつ触覚ディスプレイに表示できるように変換しなければならない. 処理として, 図面を 16×16 の大きさに割り当て, 素子を1つのピンで表わし, 素子の接続関係より, 素子間に配線を引き, そして, 無駄な配線部分を削除し, ピン数を減らす. この手順により変換された図面は, 図 2-3 d) のようになる. 8×8 のディスプレイに表示できるように変換したものであり, 8×7 の大きさに 14 の素子 (分岐点を含む) で表現されている. 素子部分が黒丸● (高さ2) で, 配線部分が網目のかかった丸○ (高さ1) で表わされている. 連結方向は上下左右の4方向であり, 元の回路図の形を忠実に表わしている.

【処理結果・考察】

上記手順により, 図 2-5 に示す 26 枚の電気回路図面について実験したところ, すべての図面で回路図は認識でき, 約半数の 12 図面は 8×8 表示への変換ができた. 図面入力から図面表示までの所用時間 (CPU: 80286, 10MHz) は平均約 50 秒であり, その半分以上が前処理の細線化に費やされている (現在入手できるコンピュータでは, 1 秒未満で処理できる内容である).

この図面処理における問題点を挙げると, 以下のようになる.

- 1) 図面入力時における照明状況, 図面の傾きなどに対処できる処理系でなければならない. 今回は全体が均一な状況であるという仮定のもとで二値化しており, その仮定が崩れる (照明が均一でない) と雑音が生じ, 処理ができなくなる. また, 傾きに関しても計算機はマス目の処理系なため, 水平・垂直な線の処理は簡単であるが, 斜めの線に対しては誤差が大きくなり, 処理が難しくなる. ゆえに現段階では, 図面はまっすぐな状態で入力されるという前提に基づいている. 現時点では, 解像度の高い入力系をもちい, 傾き補正するなどの方策が有効である.
- 2) 線分の分類は特徴量を用いて行っており, ベクトル化に比べて処理が楽で, 処理時間も短縮され有効である. しかし, 対象が小さくなると分類が困難になるため, 誤分類が多くなり, 以後の処理が不可能となる. ゆえに, 特徴量の利用法について改善しなければならない. これも高解

像度化により改善されると思われる。

3) 今回扱った電気回路図以外の他図面について扱う場合には、図 2-3 の前処理における雑音除去、線図形処理の線分の分類以降を修正することで適応できるが、その修正が簡単に行えるような処理系への変換が必要である。

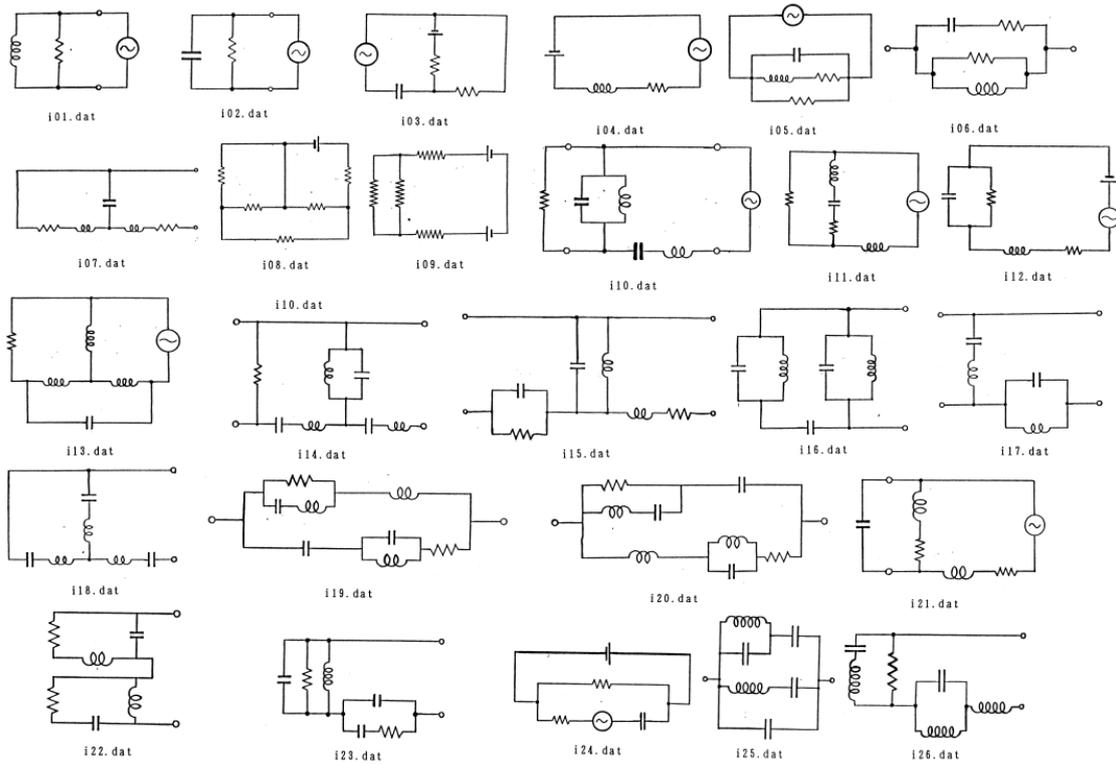


図 2-5 評価用電気回路図

2.4. 情報提示

2.4.1. 触覚ディスプレイ

触覚知覚には能動的触覚知覚と受動的触覚知覚がある。前者は点字を読む場合のように触覚刺激物が固定されていて、使用者が手や指を動かして知覚する方法で、後者はディスプレイが背中や腹部に固定され、触覚刺激物が振動することで知覚する方法である。Gibson(1962)や Cronin(1977), Heller(1980, 1983)らは、手のひらや指先で2次元パターンを触覚知覚する場合、能動的触覚知覚が受動的触覚知覚よりも一般的に優れている結果を報告している [21] [22]。

ゆえに本研究では、凸状の触覚刺激物のパターンを使用者が指先で自由に走査できるような触覚ディスプレイ(図 2-6)を開発した。また、触覚刺激物を駆動するものには、バイモルフ振動子や小形ソレノイドがあるが [23] [24] [25], 開発したディスプレイは小形ソレノイドで駆動される非振動型ピンのアレーである。ピンの突出状態の固定は機械的なものと電気的なものとが考えられるが、表 2-3 の特色を考慮して、制御回路の負担が大きく消費電力も多いが、構造が簡単でメンテナンスがしやすい電気的なピン位置固定法を採用した。

このディスプレイの特色は次の2点である。

- ・触覚による情報が多く提示できるように、オン・オフの情報でなく、高さ 0, 1, 2 の三つの段階がある。

- ・ディスプレイは単に情報を表示するだけでなく、指で触れている位置を計測するタッチセンサの機能もある。

そして、このディスプレイは本体とその制御部、指でさわっている場所の詳しい情報を知りたいときなどに押す情報取得キーで構成されている(図 2-7)。

図 2-8に触覚ディスプレイの構造を示す。ピンは8×8に配列され、小形ソレノイドにより押し上げられる。ピン間隔を狭めるために図に示すような2層構造とした。これにより、1層のものに比べてピン間隔は $1/\sqrt{2}$ 倍となっている。高さ1, 2に対するピンの高さはそれぞれ2mm, 4mmであり、高さ1はソレノイド1に通電することにより、高さ2はソレノイド1, 2をともに通電することで実現される。押し上げられたピンはソレノイドの電磁力により固定され、指先の圧力にも耐え得る。指で押さえているピンの位置は上端部の加圧導電性ゴム(PCR)を用いたスイッチにより計測する。このゴムは軽い圧力に対してもよく反応し、また小さいためこのようなスイッチに適している。ピンのリセットはソレノイドに流している電流を切断することで行われ、1層目のピンはピンの重さによる自然落下、2層目のピンはピンが長いことひっかかり抵抗が大きいので、パネによる強制落下である。また、ソレノイドに電流を流し続けると熱が発生するため、電動ファンを設け、強制冷却ができるようになっている。

図 2-9にディスプレイの制御部の構造を示す。(a)はピン出力部で、(b)は位置入力部である。パラレル I/O を介してパソコンから xy 信号(X0, X1, X2, Y0, Y1, Y2), 高さ信号(H1, H2), リセット信号(RESET), 選択信号(ENABLE)が送られ、それらをもとにピンを出力したり、さわっているピンの位置を検出する。また、情報取得キーは三つのキーからなっており、その押し方により音声情報の選択(全体情報, 個別情報), 内容変更ができる。

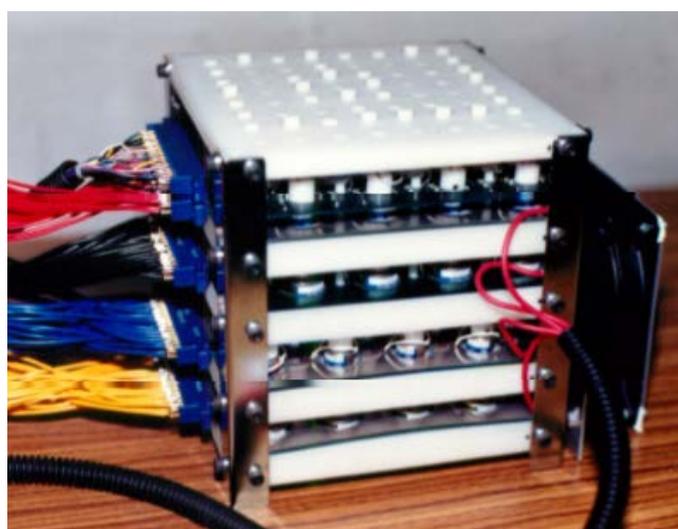


図 2-6 触覚ディスプレイ

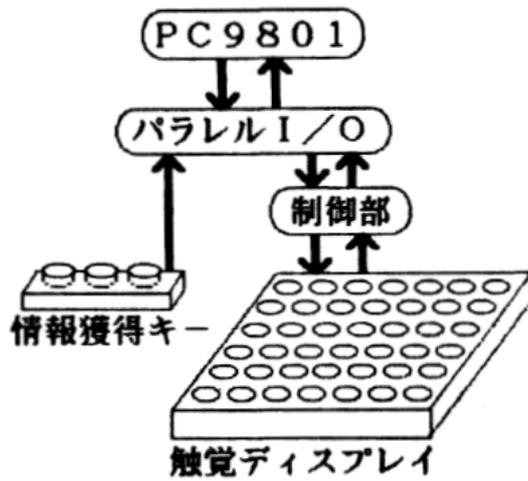


図 2-7 触覚ディスプレイの構成

表 2-3 ピン位置固定の方式の特色

比較項目	機 械 的	電 気 的
構 造	×	○
制 御 回 路	○	×
消 費 電 力	○	×
メ ン テ ナ ンス	×	○

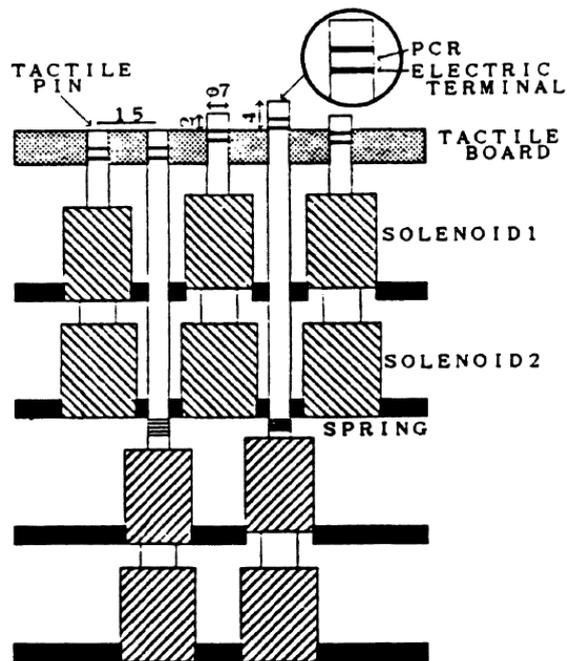
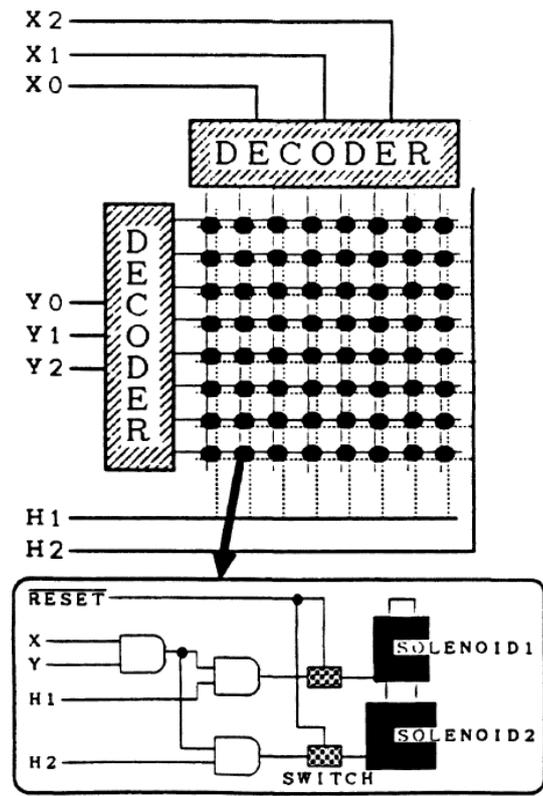
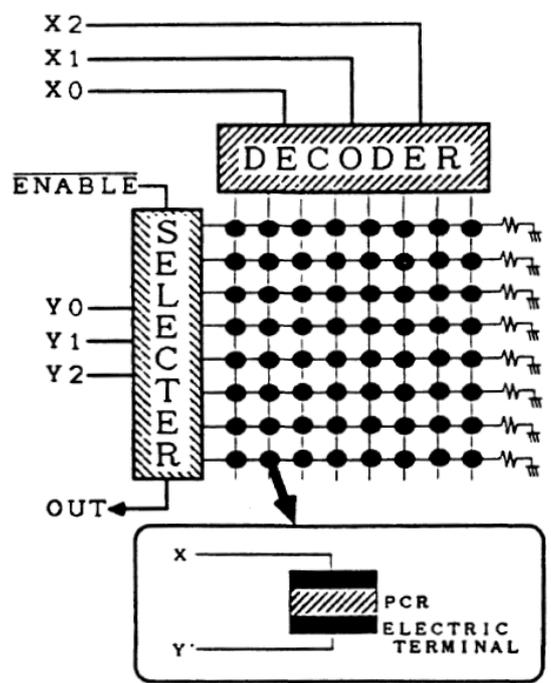


図 2-8 触覚ディスプレイの構造



(a) Circuit for output



(b) Circuit for input

図 2-9 触覚ディスプレイの制御回路

2.4.2. 音声出力

音声出力の方法としては、合成音声出力と録音音声出力の 2 通りがある。前者は任意の文を再生できるが、規則が複雑で辞書が大きくなる。後者はその逆で、データをあらかじめ用意しておかなければならないが、システム自体は小さい。コスト面と上記のことより、録音再生出力のうち、メモリ効率が良く音質も必要十分な ADPCM 方式(8kHz, 67 秒録再生可能)の音声出力ボードを採用した。

音声出力は図 2-10 の手順で行われる。まず、

- ① ファイルの音声データを PC9801 のメモリ上に読み込み、
- ② 必要なデータを取り出して組み合わせ、
- ③ 音声ボードのメモリへ転送し、
- ④ 再生する。

図 2-3 の図面の音声出力例を以下に示す。用意するデータは表 2-4 のように回路情報、素子名、数字など 30 個であり、これらの組合せで出力音声形成される。例えば、全体情報として「閉回路」、「素子の数は」、「個」、「10」、「4」、無音という 6 個のデータを組み合わせて、「閉回路素子の数は 14 個」が出力される。また、左上のコンデンサを押さえた場合は「コンデンサ 1」(同じ種類の素子が図中に複数個あるときは区別するために番号を付ける)と出力される。また、接続関係情報ありの場合は、「コンデンサ 1 分岐点 2 と分岐点 3 と接続」と出力される。

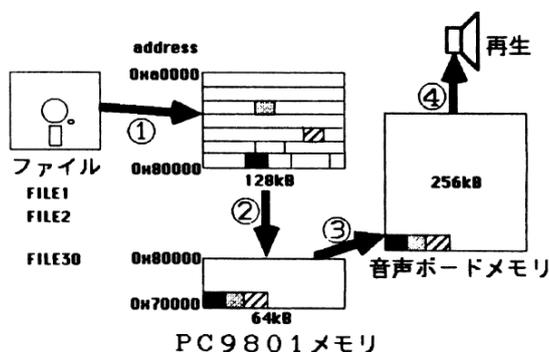


図 2-10 音声出力手順

表 2-4 音声データ

回路情報	閉回路, 2 端子回路網, 4 端子回路網
素子名	交流電源, 直流電源, コイル, 抵抗, コンデンサ, 端子, 分岐点
番号	1, 2, 3, 4, 5, 6, 7, 8, 9, 10
その他	が並列なもの, が直列なもの, と, と接続, 素子の数は, 個, あり, なし, 接続情報, 無音

2.5. 評価実験

本システムを使って、電気回路図を触覚・音声情報で表示し、どのような内容の情報を認識する場合に適しているかを調べる。

2.5.1. 方法

被験者は 20 代前半の 9 人の男子大学生である。全員が晴眼者であり、電気回路図に関する知識を有している(図 2-11)。

実験では、電気回路図(図 2-12 の閉回路, 2 端子網回路, 4 端子網回路)が 8×8 の触覚ディスプレイに表示される。目隠しをした被験者に突出したピンを指先で探索して構成を把握してもらい、素子(高さ 2)のピンを押さえながら情報取得キーを押すと、素子名(例:コイル)が音声で出力される。また、回路情報(例:閉回路素子の数は 7 個)はいつでも聞くことができる。そして、視覚によるフィードバックの影響をなくすため、頭の中に回路図が完全に思い浮かべることができるまでこれらのやりとりを繰り返し行い、その回路図を紙に描いてもらう。

実験の詳細な手順は、以下のとおりである。

[実験1]

図 2-12(a), (b)の回路を図 2-13 のような「標準, 拡大, 縮小, 合成」の 4 パターンで表示する。そして、わかりやすい表示の順位付けをしてもらう。

表示法の特徴はそれぞれ以下のとおりである。

標準: 素子と素子の間に配線部分が少なくとも一つある。

拡大: 標準より素子間隔が空いているもの。

縮小: 素子と素子の間がつまっているもの。

合成: 省略できるものは除き、まとめることができるものを合成する(例:コイルと抵抗が並列なとき、それを一つの素子とみなして表示し、音声でそれらが並列であることを知らせる)。つまり、触覚情報を減らし、音声情報を増やす表示法である。

また、音声情報の内容について、接続関係があるもの(例:抵抗コンデンサと分岐点 2 と接続)とならないもの(例:抵抗)、のどちらの情報の方が触覚情報と統合しやすいのかを調べるため、「標準」で接続情報がある場合とない場合の比較を行う。

[実験2]

被験者が実験1で選んだ表示方法で図 2-12(c)-(e)を表示し、認識してもらう。

[実験3]

(e)と同じ回路図を音声だけで表現し、触覚&音声表示の場合と比較する。音声は「コイルとコンデンサ 2 が直列なものと、コンデンサ 1 が並列に接続していて、それが抵抗と直列に接続している。そして、その両端が交流電源に接続されていて、閉回路を構成している」である。これは実験 2 の (e)の回路の触知覚を行ったすぐ後に行い、前もっては(e)と同じ回路を表現しているとは知らせない。

[アンケート]

実験1の後に, 表 2-5 の(ア), (イ)の質問に, 実験 3 の後に(ウ), すべての実験後に(エ), (オ)の質問に答えてもらう. (ア)は 4 段階評価, それ以外に二者択一の回答を求めている.



図 2-11 評価実験風景

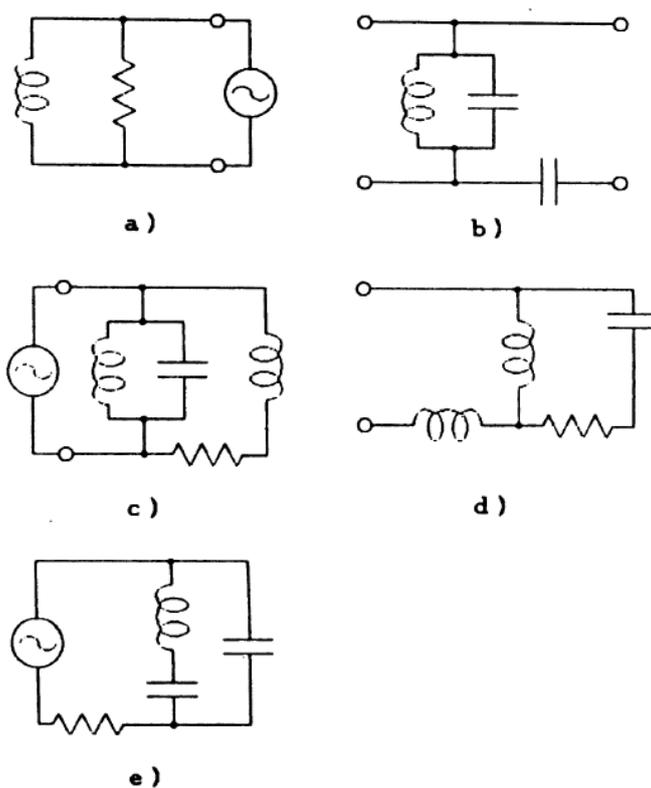


図 2-12 電気回路図

標準	拡大
○ ○ N1 ○ T1 ○ . . . ○ . ○ . . ○ . . . L0 . R0 . . V0 . . . ○ . ○ . . ○ . . . ○ ○ N2 ○ T2 ○ .	○ ○ ○ N1 ○ ○ T1 ○ ○ . . ○ . . . ○ ○ . . ○ . . . ○ L0 . . R0 . . . V0 ○ . . ○ . . . ○ ○ . . ○ . . . ○ ○ ○ ○ N2 ○ ○ T2 ○
縮小	合成
○ ○ N1 ○ T1 . . . L0 . R0 . V0 . . . ○ ○ N2 ○ T2 .	○ ○ ○ U0 . V0 ○ ○ ○ . U0 : L0 と R0 が 並列

図 2-13 回路の 4 種類の表示法(電気回路図:図 2-12 (a)の回路の場合)

表 2-5 アンケート内容

(ア) わかりやすい表示法の順位付け	実験 1
(イ) 音声情報に接続情報は必要か?	実験 1
(ウ) 音声のみの表示と触覚&音声の表示方法ではどちらが理解しやすいか?	実験 3
(エ) 回路情報は認識の手助けとなるか?	実験 1 & 2
(オ) 認識手順は、部分を把握してから全体を組み上げるのか、全体を把握してから部分を調べるのか?	実験 1 & 2

2.5.2. 結果

図面認識は 1 人につき 1 箇所程度の間違い(素子名, 位置関係など)があるが, ほぼ正しく行われた。全素子数に対する誤り率は 4%であった。実験結果をまとめたものが表 2-6 である。

アンケートの(ア)の結果をまとめたものが表 2-6 の(a)であり, 個人によっていろいろな反応が見られる。また, (b)には(イ)-(オ)の質問に対する結果がまとめられており, 9 人の被験者全員が同じ意見であった。まず, 音声情報は素子名だけでよいとする意見のみであり, 図 2-12 (e)の回路で行った音声のみで表現した場合と, 触覚&音声で提示した場合のどちらがわかりやすいのかの比較は, 触覚&音声の二つのモダリティを利用した方が優れているという結果になった。回路情報の必要性の有無に関しては, ある方が便利であり, 認識する手順は全体像を把握してから個々の部分を理解するというものであった。

また、ピン間隔、高さ(0, 2, 4 mm)は図 2-7 に示したようなものであるが、高さの差をもう少しつけた方がわかりやすいとする人が7人いた。

表 2-6 実験結果

(a) (ア)表示法のわかりやすさの結果

被 験 者	表 示 法			
	標 準	拡 大	縮 小	合 成
A	1	1	1	4
B	1	1	3	4
C	1	2	2	4
D	2	1	4	3
E	2	1	4	3
F	1	4	2	3
G	3	3	2	1
H	3	4	2	1
I	3	4	2	1

*わかりやすいものから1, 2, 3, 4

(b) その他の結果

評 価 項 目	結 果
(イ) 接続関係あり/なし	なし
(ウ) 音声のみ/触覚&音声	触覚&音声
(エ) 回路情報あり/なし	あり
(オ) 全体→部分/部分→全体	全体→部分

2.5.3. 検討

各アンケート結果に関して検討を行う。

【アンケート項目(ア)】

表示方法は表 2-6(a)に示されるように、いろいろな結果があり、どの表示法が適しているかを一概に決定することはできない。しかし、Fを除く8人を、A~EとG~Iの2グループに分けることができる。

前者のグループは、標準、拡大の場合の方が縮小、合成よりもわかりやすいとする人で、その理由は以下のとおりである。縮小では素子が連続的に並ぶため、ピンの高さの差がなくなり、どの位置のピンをさわっているかがわかりづらい。また、合成に関しては、ピン数は少なく触知覚が容易

であるが、音声で二つの素子の直・並列が示されるため、全体像に形として組み込みにくいためである。標準がわかりやすい理由は、隣のピンとの高さに差があり、位置関係が把握しやすいためである。拡大は標準より更にピン間隔が空いているため、位置関係がわかりやすい。このグループは、更にA～CとD, Eの二つに分けることができ、A～Cは合成が一番わかりにくいとする人であり、D, Eはピンがつまった縮小が一番わかりづらい人である。

後者のグループでは、前者とは逆に合成、縮小の方がわかりやすいとしている。これらの人には、標準、拡大において表示範囲が大きくなり全体の把握がしにくいことの方が、縮小、合成においてピン間隔がつまっていることにより支障が大きいことを示している。ゆえに、範囲が小さい方が認識しやすくなると言える。

また、被験者 F は上記の二つのグループの中間的な位置であると言える。

以上の結果を考慮すると、触覚ディスプレイへの図面表示はどれが最適であるかは決定できない。ゆえに、使用者が自分に適したものを選択できるようにすべきである。

【アンケート項目(イ)】

音声情報の内容については、素子名だけの表示が良いという結果であったが、これは、接続関係までを表示に入れると音声の情報量が多過ぎて対処できず、そのために以前に覚えていたことまで忘れてしまうからである。接続関係は触覚によって得ることができるため、音声情報は必要不可欠なものだけにすべきであろう。但し、確認のために利用することができるとする人もいるので、現在のシステムでは、ボタンを押すことでその内容を選択できるようにしている。

【アンケート項目(ウ)】

回路表現法で、音声表現のみと触覚&音声では、明らかに後者の方が優れている結果が出た。これは、一度に覚えられる情報量に限りがある点や、音声で位置関係を知らせるのが難しいのに対し、触覚は位置関係の把握では優れているためである。また、音声情報はどちらかという受身的な情報取得であるのに対して、触覚&音声表示では、自分の意思で指を動かすことができるという能動的なもののため、理解しやすいついと言える。ゆえに、図面を扱う場合においては、位置、構成の把握がしやすい本システムの考え方が良いと言える。

【アンケート項目(エ)】

回路情報は、最初におおよその形を頭の中に描くためや、最後に自分の考えているものと一致しているかを確認するために利用され、回路図の把握の負担軽減に役立っている。回路の形を聞いておおよその形を思い浮かべることが、電気回路図についての知識がなければできないが、この学習はシステムを使えば簡単に行える。ゆえに、今回の実験ではすべての人が回路知識を有していたが、素人に対しても有効な結論であるといえる。そして、回路情報は現在、回路形態と素子数が出力されているが、前知識のある人に対して更に有効な情報(「共振回路である」など)も出力できるようにすべきである。

【アンケート項目(オ)】

認識手順に関しては、触覚に慣れないうちは個々の部分を調べながら全体の形を把握しようとする場合もあったが、最終的には、全体の形を把握してから、個々の部分を調べていくようなトップ

ダウン処理になる。回路情報を聞くことで回路のおおよそのイメージがわくということもあるかもしれないが、これは、人間が無意識に行っている処理に基づいているためであり、そのため部分から全体よりも把握しやすくなっていると言える。

以上の晴眼者に対する実験結果が、視覚障害者、特に先元性視覚障害者でも成り立つかどうかは、今後に残された問題である。

2.6. まとめ

本章では、視覚障害者が図面を容易に理解できる図面認識支援システムを説明した。視覚障害者が対話的な操作で図面を触覚と音声情報から理解するシステムであり、開発した触覚ディスプレイと、そのディスプレイを用いた触覚・音声情報の提示方法、評価実験の結果について述べた。2.2 で述べた本システムの各特徴については、

- (1) システムの知能化: 画像処理技術を活用し、図面そのものの情報を触覚で提示するのではなく、図面認識をさせることで提示すべき情報量を低減させ、受動情報量が視覚に比べて少ない触覚や聴覚でも、十分に視覚情報を伝えられる点、
- (2) 触覚と聴覚の利用: 触覚と音声の利点を活かし、全体の配置情報は触覚で、個々の対象の情報は音声で表現することにより、図面情報が効率良く伝わる点、
- (3) 対話的な操作: 各触覚ピンへのタッチセンサの導入により、システムからユーザへの一方的な情報提示でなく、ユーザ側が積極的に情報を取りにいくことができ、不明な部分は何度も聞き直すことができるといった対話的な操作が可能となる点、

を各種被験者実験により示すことができ、本システムを使用することで、図面の理解が容易となる可能性を見出すことができた。

また、被験者実験からは、「触覚の最適な表示方法は被験者によって異なり、すべての人に共通な方法はない」、「音声情報は簡潔なものほど良い」、「事前に図面情報を与えることで認識しやすくなる」などの知見も得られ、これらを考慮したシステムとすることで、視覚障害者は図面情報を容易に取得できると考えられる。

本システムは、情報提示支援システムで必要とされる、情報の取得、処理、提示機能で構成された基本システムであり、今後更新、拡張すべき残された数多くの課題がある。それらの中の重要なものを列挙すると、

- ① いろいろな視覚情報を取り扱える手法の開発。図面情報だけでなく、生活空間は3次元であり、3次元の視覚情報を扱えるべきである。
- ② 触覚ディスプレイの仕様向上。触覚により得られる情報を多くするため、より提示情報量が多い触覚ディスプレイの開発が必要である。
- ③ 各種ヒューマンインタフェースの確立。各種視覚情報を最適に提示するためには、触覚、音声だけでなく、音響などを利用した提示方法、システムの開発が必要である。

などが挙げられ、次章以降でこれら課題解決に関して取り組んだ結果を説明する。

3. セグメントベースドステレオ法による 3 次元情報の取得

第 2 章では、図面情報を視覚障害者に効率よく提示するシステムに関して述べたが、そのまとめで記したように、対象を視覚障害者が日常生活を送る上で重要だと考えられる 3 次元空間情報に適用対象を拡げる必要性を感じた。そこで、ユーザ周辺の 3 次元の視覚情報を取得する手段の確立を目指し、ステレオ画像解析によって 3 次元情報を取得する方法として開発した、セグメントベースドステレオ法について述べる。境界線表現で記述されたステレオ画像間の対応探索において、領域の境界線を構成するセグメントの連結性に基づいた、大局的な対応評価手法を開発した。得られる 3 次元幾何データは、単なる 3 次元形状データではなく、3 次元物体の認識や追跡に利用可能な構造を有しており、視覚障害者が認識したい物体の検出や追跡する際に利用できるものである。

3.1. はじめに

人間が外界から得る情報の大部分は視覚によるものであるため、視覚障害者が生活する上においても大変重要な情報であり、その欠落を補う必要性は高い。特に、生活空間は 3 次元空間であり、3 次元の視覚情報をいかにして得るかが、視覚障害者の日常生活の質を向上させる重要な要因であると言える。

そこで本章では、視覚障害者に 3 次元の視覚情報を提示できるシステムの構築に必要な、周辺環境の 3 次元視覚情報を取得する手法の提案を行う。3 次元の情報を得ることは、視覚障害者への情報提示という目的のみならず、システムの自動化や、ロボットの知能化などでも重要であるため、昔から活発に研究が行なわれてきている [26] [27] [28] [29] [30]。3 次元情報取得する方法として一般的には、計測対象に対してパターン光などを投射するアクティブセンサ系と、複数台のカメラを利用するパッシブセンサ系がある(図 3-1)。前者は、マイクロソフト社から発売されている Kinect センサが代表例である。計測対象にパターンを投影し、その投影デバイスと観測カメラの関係から 3 次元形状を計測するシステムであり、投影パターンと観測画像間での対応付けが一意に決まりやすいため、誤計測が少なく、対象物が均一色のように模様のないものであっても計測可能である。しかしながら、パターンが写らない対象の計測が不可能である点、投影機とカメラ間の位置関係のキャリブレーションが難しいといった問題点がある。後者は前者に比べ後に述べるカメラキャリブレーション手法などで自由な設計ができるなど制約条件が少ないが、画像間の対応探索問題が最大の課題であり、特徴のない均一色のように模様のない対象物に対しては計測ができない欠点がある。

本章では、これらの長所・欠点を鑑み、日常生活上の様々な状況下でも比較的安定して使用できる、後者のパッシブセンサ系のステレオカメラによる 3 次元情報の取得手法に関しての研究開発を行った。上記のように、ステレオ画像間の対応探索問題の解決を行うことを目的とし、この課題解決により、身の回りの 3 次元の空間情報を安定して得られることにつながる。まず 3.2 において、ステレオカメラ系での 3 次元情報の計測、および、計測結果から対象物を検出手法に関して、簡単

に一般的な処理の流れを説明する。これらの処理のうち、3.3 において、ステレオ画像間の対応探索問題を解決し、安定した 3 次元情報を取得することができる、セグメントベースドステレオ法に関して詳細に述べる。

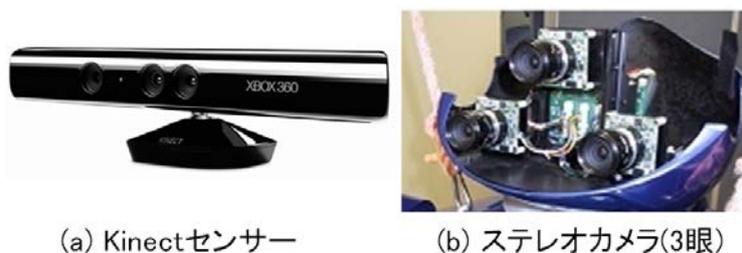


図 3-1 3D センサ

3.2. 3 次元情報取得処理

3 次元情報取得の基本的な処理の流れを、パッシブセンサ系のステレオカメラを対象としたものを示す(図 3-2)。処理は、計測系のカメラのキャリブレーション、計測処理、対象物検出処理に分けることができる。

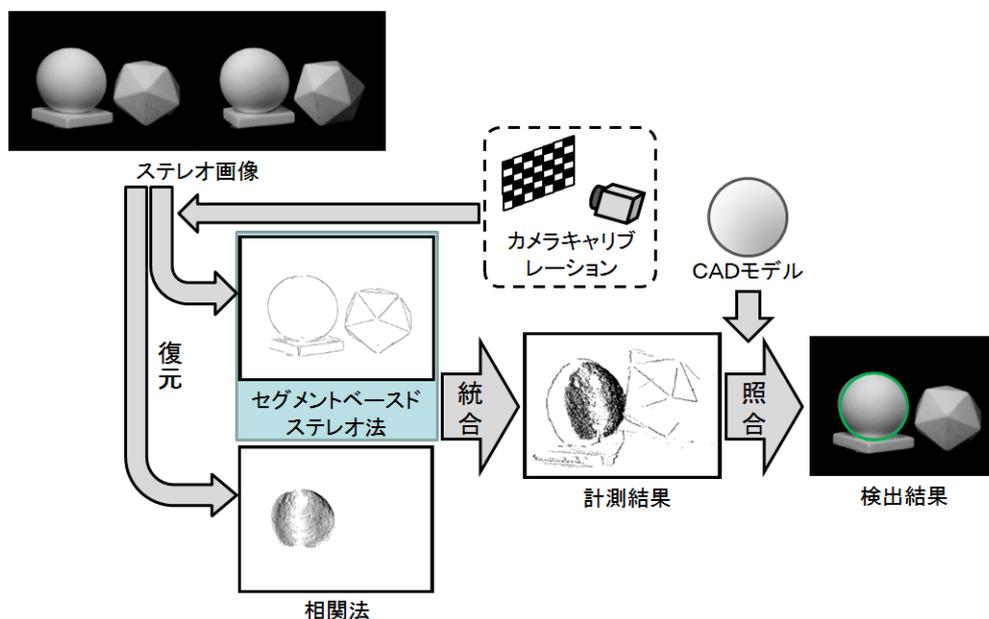
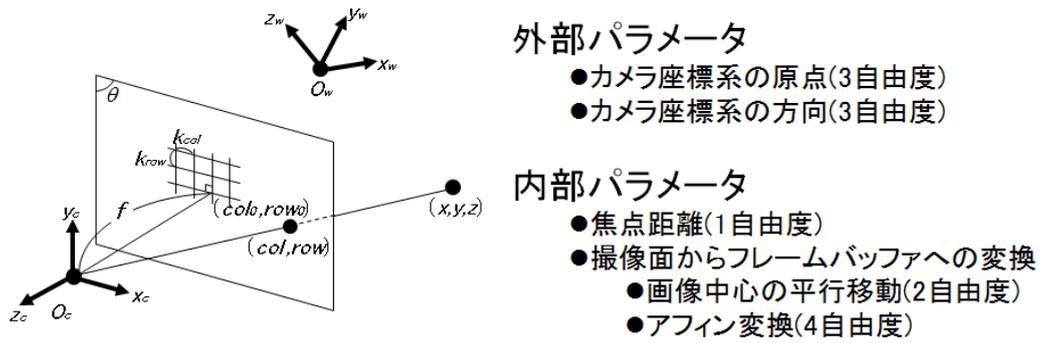


図 3-2 ステレオビジョンシステム

3.2.1. カメラキャリブレーション

計測前に既知パターンなどを用いて観測系のキャリブレーションを行う。つまり、それぞれのカメラの外部パラメータ(位置、姿勢)、および内部パラメータ(焦点距離など)、レンズの歪みパラメータをあらかじめ計算しておく(図 3-3)。ステレオカメラのカメラキャリブレーションに関しては付録 A1 に具体的な一手法を示している。



外部パラメータ

- カメラ座標系の原点(3自由度)
- カメラ座標系の方向(3自由度)

内部パラメータ

- 焦点距離(1自由度)
- 撮像面からフレームバッファへの変換
 - 画像中心の平行移動(2自由度)
 - アフィン変換(4自由度)

図 3-3 カメラキャリブレーション

また、ステレオ画像間の対応関係を求める際、基本的には視点が変わっても注目箇所の輝度・特徴が変わらないという条件で対応を求めており、画像は同一条件で撮影されているという前提に基づいている。しかしながら、シャッタースピードや絞りなど同一値にしてもカメラの個体差などの調整しきれない場合もあり、画像間の輝度調整法に関しても付録 A2 で修正手法の一例を示している。

3.2.2. 対象物計測手法

次に、ステレオ画像間での対応関係を求める。ここでは対象シーンのデータ特徴に応じた代表的な 2 つの形状復元手法、セグメントベースドステレオ法と相関法を説明する。対象画像において、コントラストがはっきりしているものに対してはセグメントベースドステレオ法を、模様があったり、輝度変化が滑らかなシェーディング領域 [31]に対しては相関法を用いる。つまり、図 3-4 に示すように画像データの質に合わせて、画像間に対応させる特徴量を変えて対応関係を求める。

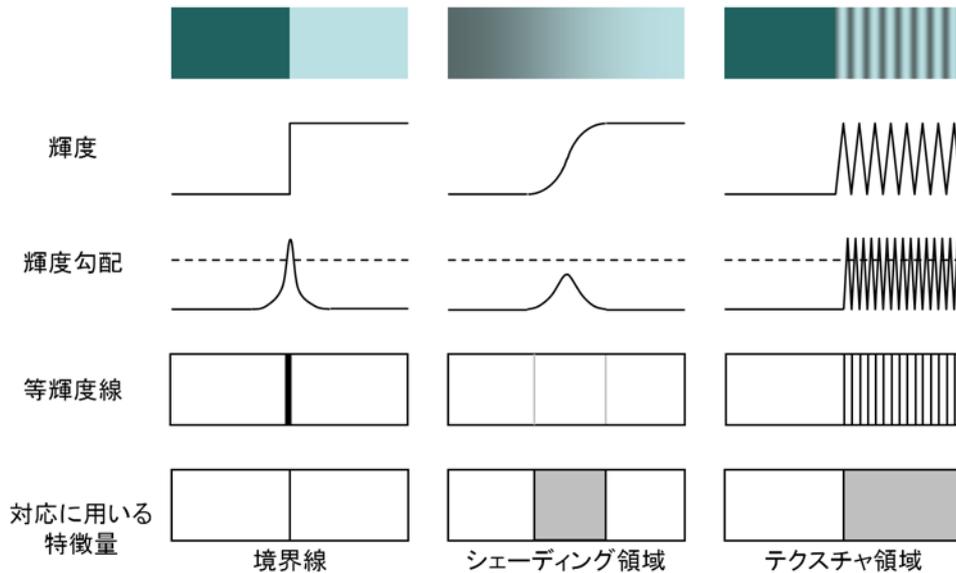


図 3-4 データの状況に応じた特徴量算出

特徴量の対応関係と先に求めたカメラの位置・姿勢情報から、三角測量の原理で観測点の 3 次元の位置情報が計算でき、観測点すべての 3 次元情報を得ることが可能となる。

3.2.2.1. セグメントベースドステレオ法

セグメントベースド法は物体境界線の対応を求め、3次元ワイヤーフレームを復元する手法である [32] [33]。図 3-5 の処理手順に沿って説明する。

左右それぞれの画像で、エッジ検出、細線化、延長処理、小領域除去を行うことで境界線を抽出し、特徴点(線が枝分かれする分岐点や、折れ曲がった頂点、端点など)でセグメントに分割し、境界線表現である B-Rep(boundary representation) [34]で表す。ステレオ対応においては、この B-Rep のセグメントを対応単位として処理を行う。対応探索は局所的なものではなく、領域の境界を構成する境界線セグメントの連結性に基づいて大局的に対応を評価する。つまり、基準となる画像の各境界線と類似の形状・特徴を持つセグメント列を、もう一方の画像中から見つけ出すことで、局所的なセグメント間の対応ではなく、境界線というより大局的な基準を用いた対応を求めることができる。

まず前処理として、エピソード条件、セグメントの属性(輝度、方向性など)で、左右のセグメント対である対応候補を見つける。各候補には左右のセグメントの対応区間の長さ、輝度差を反映した類似度を与える。この処理により対応探索範囲を小さくできる [32]。

次に画像の各境界線において、対応候補のセグメント間の連結を調べる。連結性はセグメント間の近距離性、同輝度性、同角度性の3要素で評価する。これにより、左右セグメント対を節点、連結を辺とする有向グラフができる。このグラフの中から各候補の類似度の和が最大となるパスを探索し、そのパスに属するセグメント対を対応候補として残し、パスを反映した大局的な類似度を各セグメント対の類似度に代入する [33]。

ここまでの段階ではまだ信頼性の低い対応や、左右のセグメント間で一対一の対応関係ではなく、重複、つまり多重対応関係にあるものもある。そこで、類似度を基にその値が低い対応や、どのセグメントにも連結していない単独対応、対応区間の重複する対応を除く処理を行い、左右の一意的対応関係を求める。

以上の処理手順により境界線の視差が求まり、3次元情報を復元できる。本手法の詳細は 3.3 で記載する。

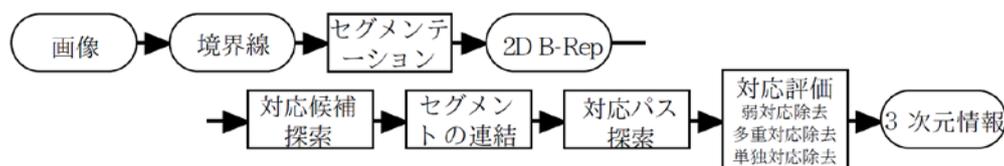


図 3-5 セグメントベースドステレオ法の処理手順

3.2.2.2. 相関法ステレオ

相関法は表面形状データを得るために使われるが、すべての画素に対しては行わない。前処理

として相関法を適用するのに適した領域を抽出し、その領域に対してのみ処理を行う。これにより時間、誤対応問題を軽減できる。本手法では図 3-4 に示すようなテクスチャ領域 [35]とシェーディング領域 [31]を図 3-6 の処理で求め、相関を計算し 3 次元復元を行う。

ここでいうテクスチャ領域は多数の小領域が隣接関係にある集合を意味する。抽出はまず明るさの一様性に基づいてエッジを検出し、画像の分割を行う。小領域の隣接関係を調べ、小領域の集合領域を一つの領域、テクスチャ領域とする [35]。

シェーディング領域はエッジ強度が一定値以上で、方向に連続性があるとし、この領域を求める。この処理でおおよそのシェーディング領域は求まるが、ステップ状エッジを含む領域幅が狭い領域も含まれる。この領域を収縮・膨張処理で削除し、シェーディング領域 [31]を求める。

3 次元復元においては、セグメントベースドステレオで求めた対応する境界線の視差を利用してテクスチャ領域同志、シェーディング領域同志において、画素単位で相関を求める。領域は境界線の内側に存在するので、境界線の視差を相関探索範囲の境界条件とすることで、エポポーラ線上における探索範囲を小さくすることができる。具体的には、画像を順方向、逆方向ラスタ走査を行い、境界線の視差を対応探索範囲の初期値として近傍で相関を計算し、最良の相関が得られる時の視差を選ぶ。この処理で対象領域全体の視差を計算できる。これにより、セグメントベースドステレオ、相関法ステレオのそれぞれの手法で復元された 3 次元データを、お互いに矛盾が生じることなく統合でき、物体の形状データを得られる。

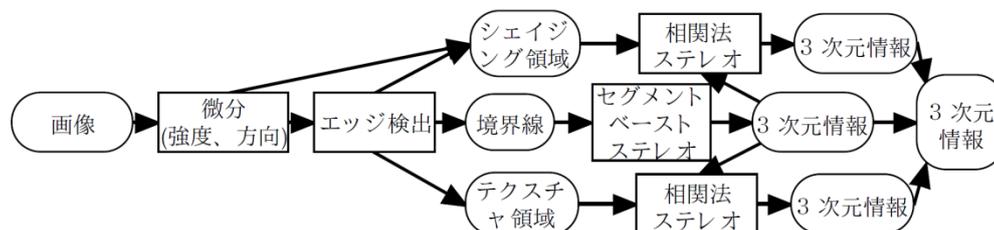


図 3-6 相関法ステレオの処理手順

3.2.3. 対象物検出

最後に必要に応じて、3 次元データ取得後、物体検出処理を行う。対象物体の幾何モデルをあらかじめデータベースに用意しておき、計測データと照合することでその位置・姿勢を検出する、幾何モデルに基づいた物体検出手法である[22] [23] [24] [25] [26] [27] [28] [29]。以下にそれぞれの手法について簡単に説明する。

3 次元空間において、回転行列 R 、平行移動ベクトル t とすると、座標変換行列 T は

$$T = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \quad (3-1)$$

で表される 4×4 の行列式で表現できる。すなわち検出は物体モデルと観測データを照合し、 T を計算する処理である。処理は初期照合(局所的な幾何特徴を用いる)と微調整(全体的な形状を用いる)の二段階で行う(図 3-7)。以下に処理の概要を示す。

初期照合では、基本的には物体モデルと観測データ間で可能性のある全ての組み合わせを考

える。例えば、幾何特徴量を物体の頂点とした場合、モデルの各頂点に着目し、頂点を構成する2つのベクトルの位置、姿勢がほぼ一致する頂点をデータから検索し、その組み合わせを対応候補とする。これから座標変換行列の初期値 T_0 が求まる。これにより求めた組み合わせを対応候補とする。

微調整においては、まずモデルを T_0 で座標変換する。モデルのサンプル点(形状を表すセグメント上に離散的に設定された照合に用いる特徴点)に対応する観測データ点を探索し、その3次元距離が最小となるように T' を計算する。新たに求めた T' で同様な処理を行い、誤差が閾値以下になるまで更新する($T_n = T' T_{n-1}$, n は微調整繰り返し回数)。これらを組み合わせ候補すべてに対して行い、一致度(可視状態のモデル点が計測データとどれくらい一致しているかの評価値)でソートし、一致しきい値以上を満たすものを検出結果とする。

これらの処理により、モデルとデータ間で座標変換 $T (= T_n)$ が求まり、対象シーンにおける物体の位置、姿勢を知ることができる。

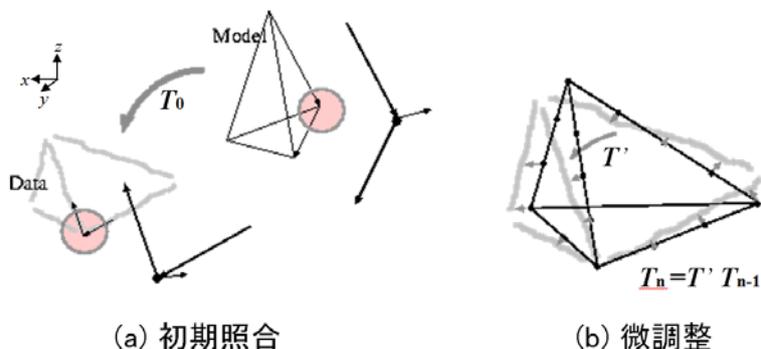


図 3-7 対象物検出処理

3.3. セグメントベースドステレオ法

本項では、視覚障害者が日常生活に必要な 3 次元空間の情報把握するために必要な、3 次元情報取得手段の一手法としての、3 次元情報セグメントベースドステレオ法に関して、その手法を詳しく述べる。

3.3.1. 対応探索の特徴

ステレオ画像における対応探索に関する研究は数多くなされているが [36]、輝度値の局所的な類似性を相関値で評価する相関法がよく用いられている。しかし、画像間の局所的な類似性だけでは探索空間が大きく、正しい対応関係を求めることは難しい。そのため、大局的な類似性も考慮の対象とした粗密解析法 [37] [38]、動的計画法 [39] [40]、構造解析法 [41] [42] [43] [44] [45] などの研究がなされている。

このうち構造解析法は、他の多くの方法が基本的にはエッジ点を対応単位としているのに対し、より高次のエッジセグメントなどを対応単位としている。この方法は処理が複雑になる反面、単位数が激減し、対応単位ごとに形状、輝度などの多くの特徴を持つので、探索空間をかなり狭めることができる。そして、カメラの特性などによる画像間での輝度対応の多少のずれに関して比較的頑健であり、復元結果が 3 次元構造記述そのものになり得るため、モデリングや物体認識などに直接利用できる点が特徴である。しかしながら、対応候補が多くなればなるほど、その対応探索が難しくなったり、対応特徴をどのように選ぶかなどの難しさがある。また、2 眼視ステレオの場合、エッジの対応・復元に関してエピポーラ線に垂直な方向のエッジ部分の対応探索のみで、復元(特に水平方向のセグメント)に関しては不完全な手法や、直線セグメントのみを対象にした手法が多い。

開発した手法は、構造解析法に属する方法であり、対応単位をセグメントとして、画像間で同じ形状の輪郭線の対応関係を求める方法である。従来の方法に比べ、セグメント単位の局所的な対応関係探索ではなく、輪郭線というより大局的な特徴量を基に対応評価を行うため、オクルージョンが発生した場合、その近辺のセグメントの対応関係は一般に不安定になりがちであるが、連結しているセグメントから得られる情報を利用することで、より安定した対応を求めることができる。また、他の方法では対応セグメントを直線近似して 3 次元復元を行うものが多かったが、本手法では対応単位となるセグメントを直線と曲線に分類しているため、任意の形状を扱うことが可能であり、曲線部分の復元精度も良くなる。復元結果はセグメントベースの 3 次元物体認識や運動追跡などに直接利用できる構造となっている点も本手法の特徴である。そして、得られる境界線部分の対応結果は、輝度値の局所的な相関により対応を求める相関法に適用もできる。すなわち、境界線部分の対応関係を相関値計算の境界条件(初期値)とすることで、効率よく、かつ、正しい対応関係を求めることができ、局所的な相関法と大局的な構造解析法を統合した方法に拡張可能な点も特徴である。

本手法の具体的な処理の流れは以下のとおりである。まず画像を境界線表現に変換し、エピポーラ条件、輝度などからセグメント間の対応候補を求め、局所的な類似度を計算する。セグメント同志の連結評価を近距離性、同輝度性、同角度性によって行い、この連結性に基づいて基準となる画像の境界線と類似したセグメント列をもう一方の画像中から求め、先に求めた局所類似度からよ

り大局的な類似度を計算する. この類似度をセグメント間の評価基準として, 誤対応除去, 多重対応除去を行い, 画像間のセグメントの正しい対応を求め, 3次元復元を行う. セグメントの連結を考慮しているため, エピポーラ線に水平になるセグメントの復元も可能である.

本項ではこの提案手法のうち, セグメントベースドステレオにおけるセグメント間の連結性を求める方法, 誤対応・多重対応除去の対応評価法, および, 3次元復元法の処理を詳細に述べ, 実験結果を基に本手法の有効性を示す. なお, 全体の処理過程において, 初期対応候補探索と対応パス探索の処理に関しては文献 [32] [33]に詳細に記載してあるので, 簡潔な説明にとどめる.

3.3.2. 境界線表現

対象画像に対してエッジ検出, 細線化, 延長処理, 小領域除去を行うことにより境界線を抽出する. この境界線を特徴点で分割して各区間をセグメントと定義する [46]. 特徴点は図 3-8 に示すように, 分岐点(境界線が枝分かれする点), 屈曲点(法線方向が急激に変化する点), 変曲点(凹凸属性の異なる曲線がなめらかに接続する点), 遷移点(曲線と直線がなめらかに接続する点), 端点(線の端の点)である.

境界線は2つの異なる領域の境界であり, 両方の領域で意味のある線なので, セグメントの方向を, 領域を右に見込むように表裏両側に定義する. 各セグメントは前後, および, 裏のセグメントへのポイントがある winged-edge 構造になっており, セグメントは点列で構成される. このような領域, 境界線, セグメント, 点の四階層で表現されたデータ構造を 2D B-Rep (Boundary Representation) と呼ぶ(図 3-9) [34]. ステレオ対応においては, この 2D B-Rep のセグメントを単位にして処理を行う.

本ステレオでは図 3-10 に示すように, 2台のカメラの光軸方向が平行である標準ステレオカメラモデルに変換する [32]. これにより, セグメントに属する画像上の各点 (col, row) は以下の式で標準ステレオカメラ系の座標 (X, Y) に変換され, エピポーラ条件は $Y_L = Y_R$ となる.

ステレオにおいてはエピポーラ線と平行なセグメントは対応関係が不定になりやすいため, 図 3-11 に示すように, 切り分けて扱う必要がある. 曲線セグメントにおいて, 構成する点列を Y 値について見た場合に, (a)極大/極小になる点について分割を行う. また, (b)エピポーラ線と平行な部分を含む場合は Y 値を基に水平部分とそれ以外のセグメントに分割する. これらの処理により, 非水平セグメントにおいては, 構成する点列は Y 値について見ると単調増加(減少)が保証される.

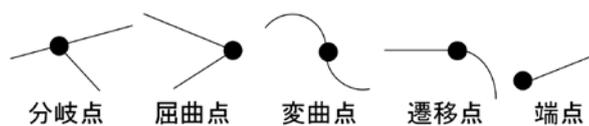


図 3-8 特徴点の種類

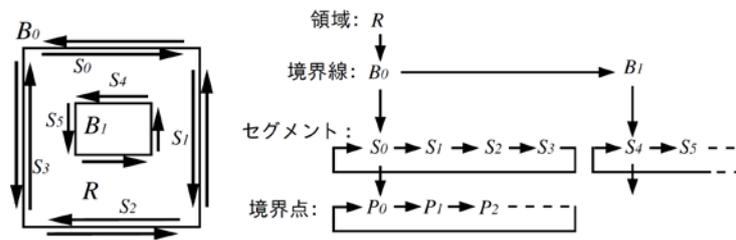


図 3-9 画像の境界線表現(2D B-Rep)

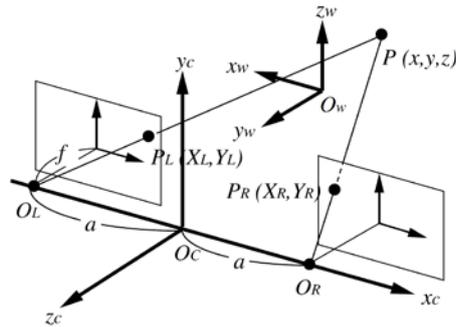


図 3-10 標準ステレオカメラモデル

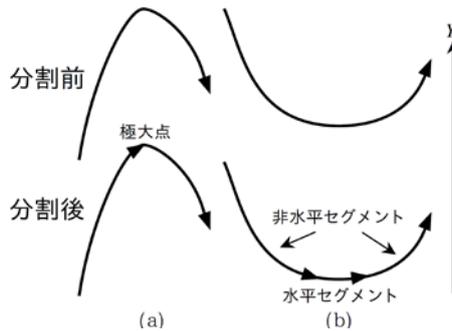


図 3-11 曲線セグメントの分割

3.3.3. 2D B-Rep から 3D B-Rep へ

3.3.2 の処理で 2D B-Rep 化されたステレオ画像に対して、基準画像(以後、実際の処理に合わせて左画像とする)の境界線に基づいて図 3-12 の手順で処理を行う。つまり、図 3-13 の B-Rep の例では、形状、輝度などの特徴を利用して、左画像のセグメント $l_1 \sim l_6$ で構成される境界線と類似したものを、もう一方の画像(右画像とする)のセグメント群 $r_1 \sim r_6$ から検索し、左右のセグメント間の対応を求め、3次元復元を行う。

以下に 2D B-Rep から 3D B-Rep への変換処理の各過程について説明するが、ここでは、図 3-12 の太枠部分の「ペアの連結」、「対応評価」、「3次元復元」処理を中心に述べる。「対応探索」、および、「対応パス探索」に関しては、それぞれ文献 [32] [33] に詳細に記載されている。

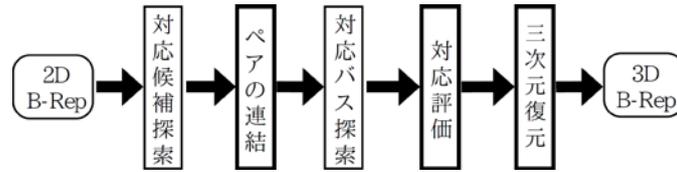


図 3-12 処理手順

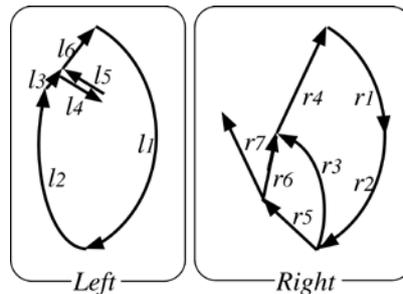


図 3-13 B-Rep 例

3.3.3.1. 対応候補探索

前処理として、セグメントの表裏を考慮せず、エッジポラ条件、セグメント間の輝度、方向(弦)、形状(直線、凹曲線、凸曲線)から対応候補を選ぶ。この処理により得られた対応候補を以下のように定義する。

ペア $[l, r]$: l とその対応候補である r の組み合わせ

各ペアは左右の画像のセグメント l, r の対応区間と類似度 S_{pair} を有する。対応区間は Y 値が共通な区間を意味し、 S_{pair} は、対応区間の長さ、および、輝度の相違度を反映した、つまり、対応区間が長いほど、左右の輝度差が小さいほど類似性があるとする、以下の評価式で求まる局所的な類似度を与える。

$$S_{pair} = \begin{cases} d & : I_{dif} < 0 \\ 0 & : I_{dif} > I_{thr} \\ \frac{I_{thr}}{I_{dif} + I_{thr}} d & : \text{otherwise} \end{cases} \quad (3-2)$$

$$I_{dif} = -|I_l - I_r| + w(|\Delta I_l| + |\Delta I_r|) \quad (3-3)$$

I_{dif} : 輝度の相違度

I_{thr} : 対応候補とみなす輝度差のしきい値,

I_s : セグメント s の輝度値

ΔI_s : セグメント s の微分値

w : 微分値と輝度差の重み

d : 対応区間の長さ

ただし、曲面の境界線である遮蔽輪郭線部分と考えた場合、その部分の輝度は左右で多少異なる可能性が大きいため、輝度の相違度 I_{dif} は式(3-3)のように、局所的な輝度変化を考慮してい

る。

また、入力データが多眼視データである時、対応検証用として他の視点からのデータを利用できる。得られたペアの対応点を 3 次元復元して検証用の画像面上に逆投影し、その近傍に同輝度、同方向の境界線が無い場合は誤対応候補として取り除く。

なお、この対応候補探索については文献 [32] に詳しく述べられている。

3.3.3.2. ペアの連結

ペアの連結性のチェックにおいては、左画像の境界線に基づいた処理を行う。同じ l を持つペア同志 (Step1)、連続する l を持つペア同志 (Step2)、そして「ひげ」を介して連続する l を持つペア同志 (Step3) の 3 つの組み合わせを考える。ペアの連結性は、上記 3 条件下では l に関して連結性は保証されているので、被対応探索画像である右画像におけるセグメント r 同志の連結を意味する。これらの連結性を近距離性、同輝度性、同角度性の 3 要素で評価する。

【Step1】

図 3-13 は B-Rep の一例で、 $[l_1, r_1]$ と $[l_1, r_2]$ 、 $[l_2, r_5]$ と $[l_2, r_6]$ などがこの処理の対象となるペアの組み合わせであり、ペアの l が同一の場合の処理である。連結性評価の条件は以下のとおりである。

[距離] ペア 1 の対応区間の終点とペア 2 の始点間距離がしきい値以下である。

[輝度] r_1 と r_2 の輝度の差がしきい値以下であり、 l と $r_{1,2}$ との輝度の差もしきい値以下である。

[角度] 左セグメントはなめらかに連続する点列なので、右セグメントも同様であるべきことから、

図 3-14 の l の 4 点の○と r の 4 点の●との間でアフィン変換が成立する。

上記の 3 条件をすべて満たす場合、ペア 1 とペア 2 は連結していると見なす。

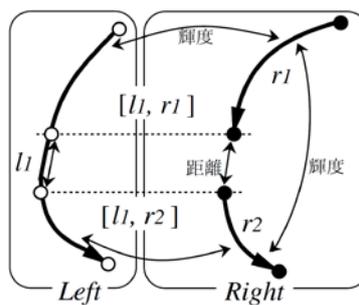


図 3-14 連結性チェック (Step1)

【Step2】

Step2 では図 3-13 の $[l_1, r_2]$ と $[l_2, r_5]$ 、 $[l_2, r_4]$ と $[l_3, r_4]$ などにあたり、連結性評価の条件は以下のとおりである。

[距離] ペア 1 の対応区間の終点とペア 2 の始点間距離がしきい値以下である。(Step1 と同じ)。

[輝度] l_1 と r_1 との輝度の差, l_2 と r_2 との輝度の差がしきい値以下である.

[角度] 図 3-15 の l_1 の終点●と l_2 の始点○での法線方向の差, r_1 の終点■と r_2 の始点□での法線方向の差をそれぞれ求め, その差がしきい値以下である.

上記の 3 条件をすべて満たす場合, ペア 1 とペア 2 は連結していると見なす.

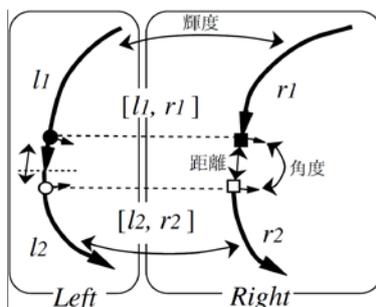


図 3-15 連結性チェック (Step2)

【Step3】

図 3-13 の $[l_3, r_4]$ と $[l_6, r_4]$ が Step3 で連結性評価される対象である.

「ひげ」は, 図 3-13 における l_4, l_5 にあたるもので, 同じ境界線上に属し, かつ表裏の関係にあるセグメントを意味する. エッジ検出処理でこのような「ひげ」が生じた場合, Step2 のみの評価では連結性が途切れる. そこで, 「ひげ」の部分飛ばして連続する l を持つペア同志の連結性を Step2 と同条件で調べる.

図 3-16 の例においては, $[l_1, r_1] \rightarrow [l_4, r_2]$ がこのチェックで追加される.

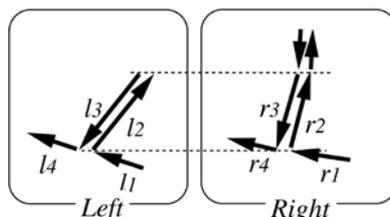


図 3-16 連結性チェック (Step3)

3.3.3.3. 対応パス探索

以上の処理で, 左右セグメント間の輝度・形状の類似度を反映した S_{pair} を節点とし, ペア間の連結を辺とする有向グラフ(図 3-17 (a))が得られる.

この段階での S_{pair} はセグメント間のみの局所的な類似度を表すものであり, より信頼性の高い対応を得るためには, さらに大局的な評価が必要である. そこで, 左画像中の各境界線を基にした処理を行い, S_{pair} を再計算する.

左画像中の個々の境界線ごとに図 3-17 (a)のような有向グラフの中から, 以下の式で表されるパスの類似度 S_{path} を最大にするパスを探索する(図 3-17 (b)). この例では(b)の太線が選ばれたパスである.

$$S_{path} = \sum_{i \in path} S_{pair}(i) \quad (3-4)$$

S_{path} を最大にするパス上のペアを正しい対応として、このパスを構成する各ペアの S_{pair} に S_{path} の値を代入する。

$$S_{path}(i) = S_{path} \quad \text{for } i \in path \quad (3-5)$$

最適候補ペアは、この探索により左画像中の境界線を基にした大局的な評価を反映した類似度を持つ。

なお、この対応パス探索については文献 [33] に詳しく述べられている。

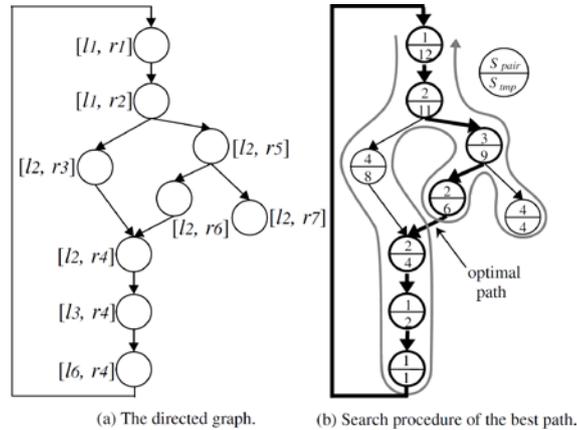


図 3-17 対応パス探索

3.3.3.4. 対応評価

これまでの処理ではペア $[l, r]$ の類似度を求めただけであり、まだ誤対応が存在し、また各 l に対して対応区間が重複した r が対応している状態、つまり多重対応も存在する。そこで、類似度を利用して誤対応や多重対応などの除去をすることで、左右のセグメントの対応が 1:1 になる処理を行う必要がある(図 3-18)。

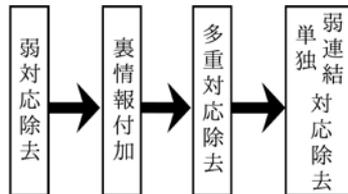


図 3-18 対応評価処理手順

まず、ペアの対応が弱いものは弱対応と見なし、対応リストから外す。

$$[l, r] = \begin{cases} \text{採用} : S_{pair} > S_{T_w}, S_{T_w} : \text{しきい値} \\ \text{除外} : \text{otherwise} \end{cases} \quad (3-6)$$

次に、この段階で初めて、セグメントの表裏の類似度を考慮する。裏セグメントで良い対応が得られている場合、すなわち、ペア $[l, r]$ の l, r の裏セグメント l', r' で構成される裏の対応 $[l', r']$ が存在し、かつ、その裏セグメントの類似度 S'_{pair} が対象ペアの類似度より大きい場合は裏の値

を代入する.

$$S_{pair} = \begin{cases} S_{pair} : S_{pair} > S'_{pair} \\ S'_{pair} : \text{otherwise} \end{cases} \quad (3-7)$$

この前処理を行った後, S_{pair} を基に, 左右のセグメントで区間の重なりが生じないように多重対応除去を行う.

$$[l, r] = \begin{cases} \text{採用} : \text{対応区間の重ならないペア} \\ \text{採用} : \text{対応区間が重なるが, } S_{pair} \text{ が最大値のペア} \\ \text{除外} : \text{otherwise} \end{cases} \quad (3-8)$$

多重対応除去処理を図 3-19 の具体例を用いて説明する.

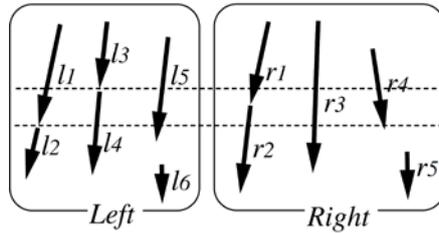


図 3-19 対応評価例

(a) 図 3-19 の例で得られたペアリスト.

$$l_1 : r_1, r_2, r_3$$

$$l_2 : r_2, r_3$$

$$l_3 : r_1, r_3$$

$$l_4 : r_1, r_2, r_3$$

$$l_5 : r_4$$

$$l_6 : r_2, r_5$$

(b) 各 l において, そのセグメントを含むペア $[l, r]$ を類似度 S_{pair} の値を基にソートする.

$$l_1 : r_1 = r_2 > r_3$$

$$l_2 : r_2 > r_3$$

$$l_3 : r_3 > r_1$$

$$l_4 : r_3 > r_1 = r_2$$

$$l_5 : r_4$$

$$l_6 : r_5 > r_2$$

$$l : r_i = r_j \rightarrow S_{[l, r_i]} = S_{[l, r_j]}$$

$$l : r_i > r_j \rightarrow S_{[l, r_i]} > S_{[l, r_j]}$$

(c) 昇順で区間の重なりが生じないペアを選ぶ.

$$l_1 : r_1, r_2$$

$$l_2 : r_2$$

$$l_3 : r_3$$

$$l_4 : r_3$$

$l_5 : r_4$

$l_6 : r_5$

(d) r に関する逆引きリストを作り, S_{pair} を基にソートする.

$r_1 : l_1 > l_3 = l_4$

$r_2 : l_1 = l_2 > l_4 > l_6$

$r_3 : l_3 = l_4 > l_1 = l_2$

$r_4 : l_5$

$r_5 : l_6$

$$S_{[l,r]} = S_{[r,l]}$$

(e) (c)と同様に区間重複除去を行う.

$r_1 : l_1$

$r_2 : l_1, l_2, l_6$

$r_3 : l_3, l_4$

$r_4 : l_5$

$r_5 : l_6$

(f) (e)を l に関するリストに戻す.

$l_1 : r_1, r_2$

$l_2 : r_2$

$l_3 : r_3$

$l_4 : r_3$

$l_5 : r_4$

$l_6 : r_2, r_5$

(g) (c)と(f)で求めたリストを比べ, 対応区間重複が生じるペアがあれば, S_{pair} の値を基に削除する.

$l_1 : r_1, r_2$

$l_2 : r_2$

$l_3 : r_3$

$l_4 : r_3$

$l_5 : r_4$

$l_6 : r_5$

(この例では l_6 に関して不整合が起きているので, $S_{[l_6,r_2]} < S_{[l_5,r_2]}$ であることから,

$[l_6, r_2]$ が削除される.)

以上の処理で, 左右両セグメントに関して多重対応区間を除去することができる. また, この処理によって選ばれなかったペアでリストを作り, 上記の処理を繰り返し行うことにより, S_{pair} は第 2 位以下であったが, 実際は正しい対応のペアを選ぶ処理も行う.

この段階で対応区間の重複問題は解消されるが, 対応の妥当性をセグメント形状のなめらかさで

判断する処理を行う。各左セグメントに対応する右セグメント数が2つ以上の場合、図 3-14 における連結性チェックの角度条件と同じようにしてなめらかさを調べる。つまり、左セグメントはなめらかな形状のセグメントであり、対応する右セグメント群も同様になめらかでなければならない。この条件を満たすならば、左右セグメントにおける対応区間の始点、終点で構成される点群間でアフィン変換が成立する。

$$[l,r] = \begin{cases} \text{採用 : アフィン変換成立} \\ \text{除外 : otherwise} \\ \quad (S_{pair} \text{ が低い方}) \end{cases} \quad (3-9)$$

そして、最後の処理として連結性のチェックを行う。残っているペア間で、もう一度連結性を調べ、連結性がない、もしくは連結性が薄いペアを、偶然に条件を満たし、信頼性が低い対応と見なして対応リストから除く処理を行う。

$$[l,r] = \begin{cases} \text{除外 : 単独対応} \\ \text{除外 : 弱連結(他の 1 つだけのペアと連結} \\ \quad \text{かつ } S_{pair} < S_{T_c}, S_{T_c} : \text{しきい値} \\ \text{採用 : otherwise} \end{cases} \quad (3-10)$$

3.3.3.5. 対応評価の統合

以上は左画像を基準画像としてその境界線を基に処理を行ったが、オクルージョンやエッジ検出失敗などによって境界線の形状が意味のない形になることもある。そこで、図 3-20 に示すように、図 3-12 のうち連結性チェック、対応パス探索、対応評価の処理を右画像の境界線に基づく処理も行い、その結果を統合することでより正確な対応評価をする。

統合においては、それぞれの処理で得られた対応関係が同一でない場合、 S_{pair} の値が大きいものを優先し、小さいものは誤対応として除外する。

$$[l,r] = \begin{cases} \text{採用 } [l,r_1] : S_{[l,r_1]} > S_{[r_2,l]} \\ \text{採用 } [r_2,l] : S_{[l,r_1]} < S_{[r_2,l]} \\ \text{除外 : otherwise} \end{cases} \quad (3-11)$$

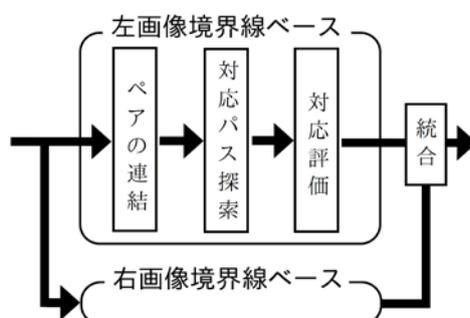


図 3-20 対応評価の統合

3.3.3.6. 3次元復元

求められた対応関係から 3 次元復元を行い, 3D B-Rep を求める. 3D B-Rep は左画像の 2D B-Rep に (x, y, z) の 3 次元座標値のスポットを付け加えたもので, 基本的には対応関係が求まった左画像の各セグメントについて点単位で復元を行うが, エピポーラ線と平行な水平セグメントに関しては特別な処理を行う.

復元精度を向上させるためには, 点単位の復元ではなく, より上位の特徴であるセグメント単位, 面単位などで行う必要があり, その一手法である平面拘束による 3 次元復元高精度化に関しては, 付録 A3 で詳細を説明している.

【水平セグメントでない場合】

対応点の標準カメラモデル座標 (図 3-10) を $(X_L, Y_L), (X_R, Y_R)$ とすれば,

$$\begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} = \frac{2a}{X_L - X_R} \begin{pmatrix} \frac{X_L + X_R}{2} \\ Y_L \\ 1 \end{pmatrix} \quad (3-12)$$

$2a$: 左右カメラの焦点中心間距離

(x_c, y_c, z_c) : カメラ座標系での 3 次元座標

となり, 3 次元座標を計算できる. しかし, セグメント l における各点と, セグメント r の対応する点とにおける Y 値が完全に一致することはほぼない. そこで, セグメントにおける Y 値は単調増加 (減少) が保証されていることを利用し, セグメント l の各点に対応する, セグメント r の対応点を求める. 図 3-21 に示すように, セグメント l のある点に対し, その Y 値である Y_L を挟み込む, 連続する 2 点●をセグメント r において探し, その 2 点を結んだ直線と, エピポーラ線 ($Y = Y_L$) の交点×を (X_R, Y_R) として, (x, y, z) を計算する.

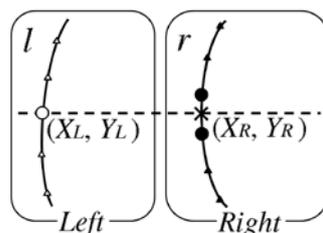


図 3-21 3次元復元

【水平セグメントの場合】

エピポーラ線と平行な水平セグメントのペアに関しては, 対応するセグメントの点は $Y_L \cong Y_R$ であるから, 相手側のどの点にも対応するといえるので, 以下の処理を行う.

2眼の場合:

水平部分の補間については, 情報が不十分であるとして補間しない. または, 両側にそれぞれ連結するペアがある場合のみ, 端点同志を直線で結び補間する.

3 眼以上の場合:

水平セグメントにならない他の視点からの画像と組み合わせ、図 3-12 と同様な処理で 3 次元復元を行い、その結果を使用する。

3.3.4. 実験結果

3.3.3 で示した処理手順により、3 次元構造を復元した実験結果を以下に示す。3 台のカメラで構成されるステレオカメラを用い、中間のカメラを左右の 2 台のカメラ間の上方に設置した。カメラの解像度が 640×480 pixels で、256 階調のグレイ画像が取得できる CCD カメラと、焦点距離が 32 mm のレンズを使用した。計測時の左右のカメラ間距離は約 600 mm、対象物までの距離は約 1.6 m の実験環境である。なお、3.3.3 で用いた各種しきい値は、画像のサイズ、輝度レベルなどを考慮して決定し、今回の実験においては以下のように定めた。

$I_{thr} : 10.0, w : 0.2$

距離 : 4.0, 輝度 : 20.0, 角度 : 20.0

$S_{T_w} : 5.0, S_{T_c} : 10.0$

実験 1 として、本手法がオクルージョン問題に対して頑健性を有することを簡単な例を用いて示す(図 3-22)。(a)のように同色の対象物が配置された実シーンについて実験を行った。同じ大きさの長方形の白紙を垂直に立てて観測した 3 眼のデータが(b)である。各画像に対して Sobel オペレータで微分し、細線化、延長処理、エッジを抽出、小領域除去を行って境界線を抽出し、2D B-Rep 構造に変換したものが(c)である。このようにオクルージョンが生じるシーンで得られるステレオ画像は、対象物が同色であるため境界線の一部が検出されない可能性があり、(c)のような場合が多い。この場合における対応関係を示した図が(d)である。(d)の状況では、局所的な特徴量のみで対応関係を求めた場合、エピソード線 $e_1 \sim e_3$ において、 e_1, e_3 ではセグメントの並びに問題はないが、 e_2 においてはオクルージョンが生じているため、垂直方向のセグメントの並び順は $(l_1, l_2, l_3, l_6) : (r_1, r_4, r_5, r_6)$ となり、 $l_2 : r_4, l_3 : r_5$ と間違った対応関係が求まる可能性がある。

しかしながら、本手法では左画像の境界線 b_1, b_2 それぞれに対して連結性を考慮しているため、対応する境界線部分を正しく求めることができる((d)左画像の境界線の黒線:対応あり, 点線:対応なし, 右画像の境界線の黒線:対応部)。本手法でも $l_2 : r_4$ も対応候補としては残るが、対応評価におけるセグメントのなめらかさのチェックで、 l_2 に対して (r_2, r_4) の組み合わせは成立せず、類似度 S_{pair} の比較により $l_2 : r_4$ は誤対応として除かれる。 $l_3 : r_5$ も同様にして誤対応として削除される。本手法により復元された結果が(e)であり、それぞれ左上部、上面、正面から見た 3D 画像である。このようにオクルージョンにより生じる並び順の変化に対して頑健で正しい対応結果を求めることができる。

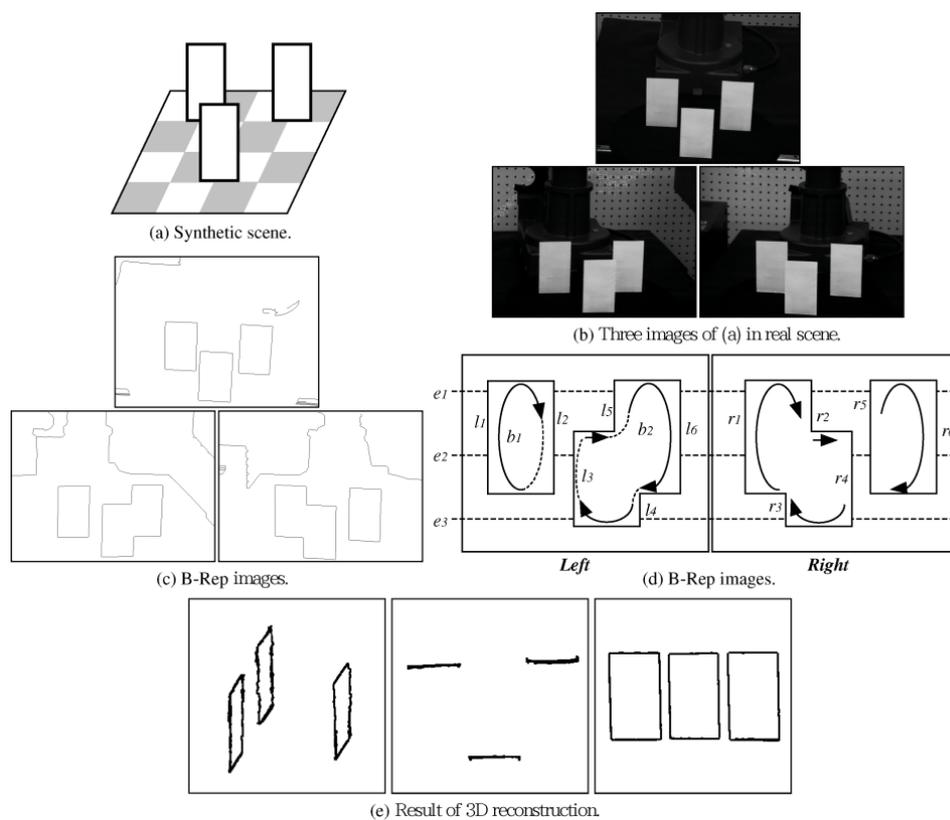


図 3-22 実験結果1

実験2では、コップや空き缶、パイプなどが置かれたシーンを撮影した(図 3-23)。(a)は取得したカメラ画像である。B-Rep 画像は(b)であり、セグメントのペアを求め、左画像の境界線を基に類似度を計算し、多重対応除去前までの処理で3次元を復元した結果が(c)である。この段階では誤対応が多く、十分な復元であるとは言えない。(d)は多重対応のみを除去したものである。大きな誤対応部分は減少しているが、まだ空き缶やパイプの一部分の対応関係に誤りが見られる。そして(e)は(d)に対して単独対応除去などの対応評価処理を行った結果と、(b)の右画像の境界線に基づいて同様な対応評価を行った結果を統合したものを異なる4つの視点から見た図であり、(c),(d)で見られた誤対応がほぼなくなっている。

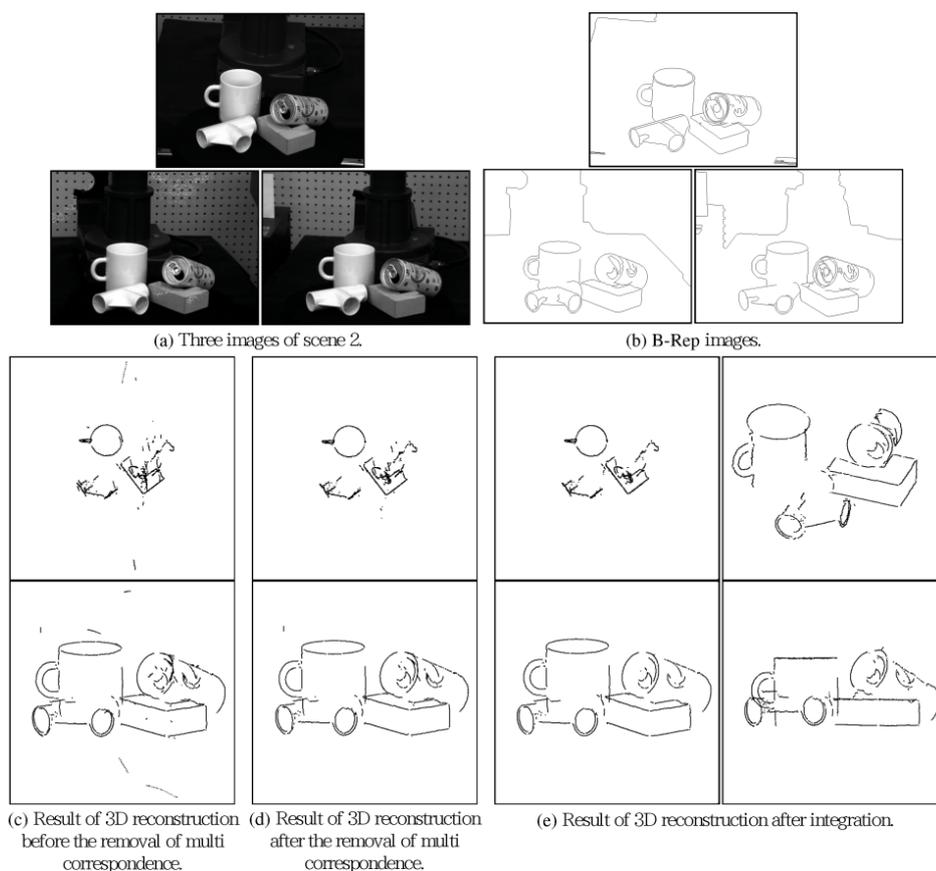


図 3-23 実験結果2

実験3では、図 3-24は水平に置かれた3台のカメラでおもちゃのガソリンスタンドを撮影した。(a)が左画像で、(b)が左右画像の2D B-Repであり、セグメント数はそれぞれ939,750である。左右の2眼のみでの3次元復元結果をそれぞれ斜め上、上面、正面からの視点で射影したものが(c)、中間の画像を対応候補探索時に検証用として用いた結果が(d)である。(c)と(d)の3次元復元を比較すると、当然ではあるが、木の部分の誤対応などがかなり減少している。(d)における対応ペア数は408個であり、うち誤対応は11ペア、誤対応率は2.7%であった。誤対応の大半は木の部分などの微小セグメントの連続している部分であり、主要な構造はこのように複雑なシーンでも十分に復元されている。処理時間(Sun Ultra1 Model 170: UltraSPARC-I,167 MHz, SPECint [252.0], SPECfp[351.0]使用)に関しては、(c)16.5秒(2D B-Rep生成:6.8秒)、(d)12.6秒(2D B-Rep生成:8.8秒)であり、対応探索時に検証する処理などが加わるにも関わらず、対応探索区間が小さくなるので、全体の処理時間は(d)の方が少なくなっており、多眼視の有効性が示されている。

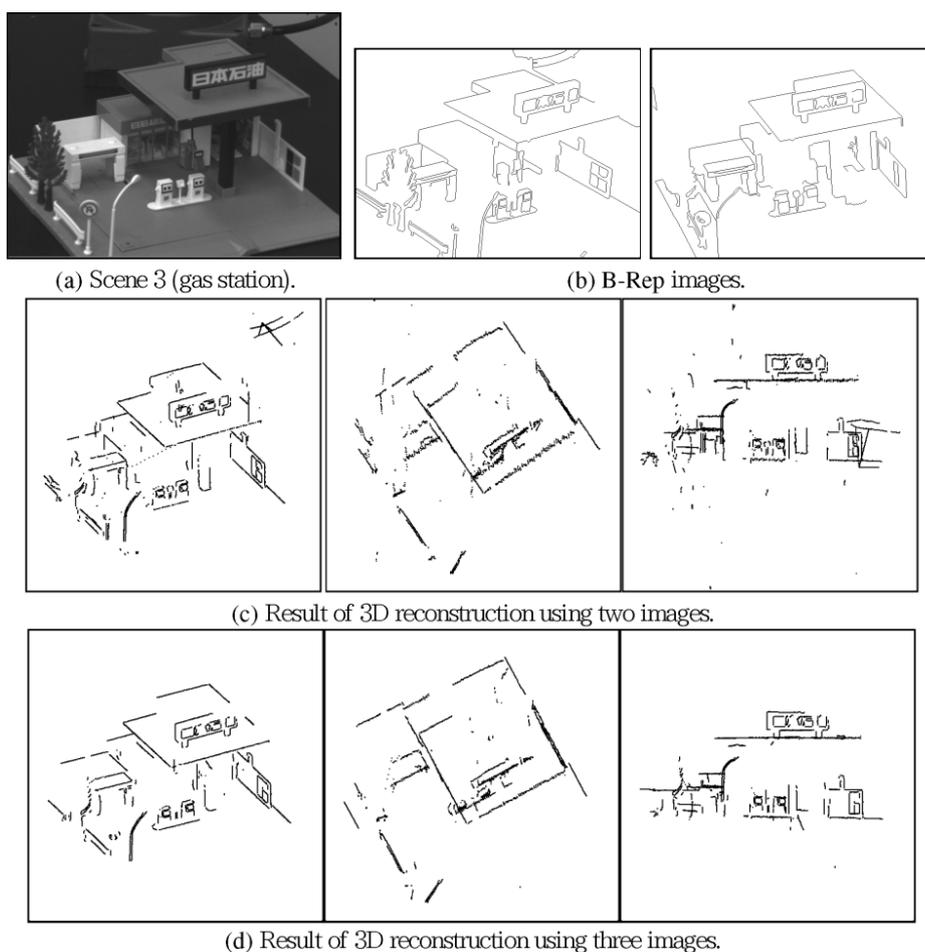


図 3-24 実験結果3

実験 4 では、人間の腕の 3 次元モデルを生成するために計測した(図 3-25)．画像間で対応が取り易いように、腕に円形のシールを貼り撮影した(a)．復元結果を 3 視点(正面、左斜下、右側面)から見たものが(b)である．このような同じ形が連続する場合には、エピソード条件、輝度の判断だけでは難しいが、本手法の境界線の形状という制約条件を付加することにより、安定に対応関係を求めることができる．

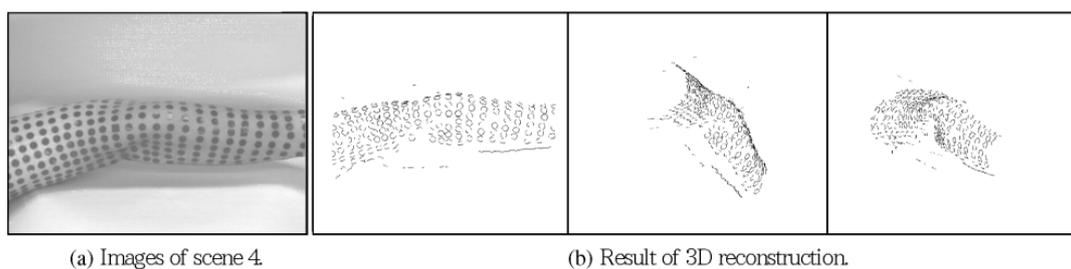


図 3-25 実験結果4

復元精度に関しては、実寸の分っている物体を、ステレオカメラシステムを用いて復元し、比較を

行い、3次元の復元精度は1 mm前後であった。この環境下では、1 pixelのずれは約1 mmの誤差に相当し、計測結果で得られた結果は理論値にほぼ一致するものである。

これらの条件を多少変更しても、復元結果に大幅な影響を与えることはないが、距離の条件を大幅に緩めると、誤対応が多くなり、全体の復元率は悪くなる傾向がある。

3.3.5. セグメントベースドステレオのまとめ

本項では、3次元情報を取得する手法として開発したセグメントベースドステレオ法に関して、特にステレオ画像間の対応探索において、領域の境界を構成する境界線セグメントの連結性に基づいて対応を評価し、3次元復元を行う方法について述べた。2D B-Rep化されたデータをエッジポラ条件、輝度などから対応候補を求め、左右のセグメントからなるペアを定義し、ペア同志の連結性を近距離性、同輝度性、同角度性によって評価した。境界線に基づいて連結するペアによって構成される最適パスを求め、その類似度を評価基準として誤対応・多重対応除去を行い、左右のセグメント間の対応関係を求め、3次元復元を行った。本手法は対応基準として境界線の類似性に基づく探索であるため、大局的な評価が可能となり、安定した復元を行うことができた。

実験として4つの実シーンを3台のカメラで撮影した画像を用い、本手法で3次元復元を行った。復元結果は十分な精度を持つ幾何情報を有しているため、物体認識 [47] [48] [49]、トラッキング [50]などに直接利用できる。また、多眼視画像では、対応候補を絞り込む前処理を行うことで対応探索空間を狭められるので、結果的には全体の処理時間の短縮につながる。また、2眼画像では難しい水平セグメント部の復元が容易になる。しかし、実験結果3で示したように2眼画像の場合でも、多少誤対応のノイズが存在するが、主要な構造に関しては十分な復元結果を得ることができる。

連結性チェックに関しては、ペア間が近距離である場合のみに限って行った。そのため、実際は1つのセグメントだが、オクルージョンなどによりそのセグメントが分割されている場合には、対応関係の不一致が起こる可能性がある。この解決策として、対応関係のチェックを強化し、距離が離れたペアでも滑らかにつながると見なせるセグメント対ならば、連結していると見なして類似性を評価する必要がある。また、対応評価においては、誤対応がまだ若干残っている。この誤対応の主な原因は短いセグメントの連結部分であり、類似度の近似した最終候補が多数残っているためである。この原因を詳しく解析して、改善しなければならない。そして、本手法は境界線に基づく手法であるため、線分が密集するようなテクスチャ部分に対しては誤対応を起こす可能性が高い。しかし、テクスチャ部分を1つの領域として扱い、領域の境界線のみを本手法の対象とすることで対応可能である。そして、テクスチャ内部の対応に関しては、得られた境界線の対応情報を用いて相関法で求める。今後は境界線部分の3次元復元だけではなく、領域内の特徴を用いて面復元などを行うモジュールを追加し、領域情報も有する拡張3D B-Repの生成を行う。また、現段階での対応関係探索処理においては、しきい値の多くは経験則に基づいて決められているが、自動設定、またはその数を減らす方向で改良を加える予定である。

3.4. まとめ

本章では、視覚障害者が必要としている情報の中でも生活上重要だと考えられる 3 次元空間情報の取得を行う方法として開発した、セグメントベースドステレオ法に関して説明した。ステレオ画像間の対応探索において、領域の境界を構成する境界線セグメントの連結性に基づいて対応を評価する方法であり、オクルージョンに対して比較的頑健な方法である。また、境界線の類似性に基づく大局的な画像間の対応評価を行うことで、安定した復元結果を得られる。得られる 3 次元幾何情報は単なる 3 次元形状データではなく、3 次元物体の認識や追跡に利用可能な構造を有しているため、視覚障害者が見つけたい対象の検出や追跡の際の基盤となりうる。この成果を、第 5 章で述べる、第 2 章の情報提示支援システムの発展系であるシステムへ適用する。

しかしながら、取得したい 3 次元の視覚情報が複雑な場合、つまり、対象となるシーンの情報が複雑な場合、ある方向から観測しただけでは十分な視覚情報を得ることができず、視覚障害者に必要十分な視覚情報を提示することができない場合も考えられる。そこで、色々な方向から得られた視覚情報を統合することにより、全体の 3 次元形状を得るための処理手法に関して検討する必要がある。次章でこの問題解決への取り組みを示す。

4. 3次元点群情報の統合

第3章では、視覚障害者が日常生活上重要な3次元空間情報を得る方法に関して述べたが、必要とするシーン状況が複雑な場合、一度の観測では十分なデータを取得できず、視覚障害者に必要十分な視覚情報を提示することができない時がある。この問題点を解決するため、多視点から得られた3次元視覚情報を統合し、視覚障害者に提示するのに必要十分な3次元情報とするための手法について述べる。パッシブセンサ系のステレオ画像の相関法や、アクティブセンサ系のレンジファインダから得られる3次元点群情報の統合方法であり、複数方向から観測した情報を統合する際、データ計測の位置関係が未知であっても、信頼性の高い領域を利用することで対応領域を求め、その関係から位置関係を計算できる統合手法を開発した。植物のように形状が複雑な対象物に対して適用し、多視点から観測した形状を正確に再現でき、視覚障害者が必要な3次元情報を得られることを示す。

4.1. はじめに

視覚障害者が生活する上においても大変重要で、必要な3次元の視覚情報をいかにして得るかが大きな課題であり、第3章では、開発した3次元空間情報の取得を行う方法について述べた。しかしながら、取得すべきシーン情報が複雑な場合、一方向からの計測では十分なデータを取得できず、視覚障害者に必要十分な視覚情報を提示することができない場合がある。提示する3次元の視覚情報を正確に、より十分なものにするためには、その対象を異なる視点から観測して、見えない部分、隠れている部分の情報を別途取得し、合わせる必要がある。

そこで本章では、視覚障害者に正確で、十分な3次元の視覚情報を提示するために必要な、3次元視覚情報を統合する手法の提案を行う。ここでは、対象とする物体の面情報などを得ることができる、パッシブセンサ系のステレオ画像の相関法や、アクティブセンサ系のレンジファインダから得られる3次元の点群データを扱う対象とする。このように、見えない部分や裏側の部分などを計測するために、複数方向から計測し、それらを統合して対象物の完全な3次元形状を生成するための研究開発が行われてきた [51] [52] [53] [54] [55] [56]。対象物が工業部品などのように多面体状のものである時、モデル化が容易なため対応がとりやすく、統合は比較的簡単であるが、複雑なものである場合は難しい。この問題を解決するために、光投影法のシステムに背面鏡を用いることで、死角部分を減らすシステムが開発されている [53]。このシステムでは1方向から計測できる部分は増えているが、対象物全体の形状を得るには至らない。また、ターンテーブル上に対象物体を置き、回転させながら物体の全周囲形状を測定するシステムもある [52] [54] [55]。これらでは、2つのデータを統合する際、移動パラメータが既知のため簡単に合成できる。しかし、上下方向からの計測ができないため、対象物の置き方を変えて測定を行ない、それらのデータをまとめる必要がある。対象物が複雑なものである場合、グラフィックディスプレイを使うなどして、人手を介して統合している [54]。多視点のデータ統合に関しては、自律走行車の環境地図作成でも行なわれている [57]。これらは移動しながら得た地形データを高さ地図に変換し、随時まとめていくことで環

境地図を作成するものである。データ同士のマッチングは特徴のある場所などを用いて行っており、データ間の移動パラメータが車輪の回転などのセンサから高精度ではないが測定できるため、対応問題が軽減されており、また、高さ地図という3次元物体の全体の形状を表現するには適していないデータ構造のため、そのまま本研究のねらいとする状況に適用できない。

本章では、上記のような問題点の解決を目指し、複雑でオクルージョンが生じやすく、不連続領域の多い3次元物体の全体の形状を得るためのデータ統合手法について述べる。まず、未測定の部分をなくすため複数方向からレンジデータを計測する。そしてデータを領域分割して、他の視点からでも見える確率の高い領域を複数選択する。それらの特徴量を計算し、各データ間の対応する領域を検索して領域の対応を定める。その領域の重心同士が一致するようにデータ間の移動パラメータの概算値を求める。精度を上げるため、データを小領域に分割して、比較するデータ間で完全に一致すると考えられる小領域を選ぶ。そして逐次的に小領域に当てはめた面のずれを最小とする処理を繰り返す。この手法により、複雑な形状の物体でも高精度に移動パラメータを計算することができ、各データを統合していくことで物体全体の形状データを取得することができる。本手法の特徴はデータ間の位置関係が未知であっても、視点方向が大幅に変わらないものであればその関係を求めることができ、それらを統合できる点である。この手法により得られる3次元の視覚情報は、面情報も含むため表面形状を表現したり、認識したい対象を検出するのに使用したりし、視覚障害者に提示するための3次元情報として利用できることが期待できる。

4.2. レンジデータ

センサなどにより計測された3次元点群情報をレンジデータと呼び、ここでは、レンジデータを取得するためのデバイスである、アクティブセンサ系のレーザレンジファインダを一例として、それから得られるレンジデータの形式に関して説明する。

4.2.1. レンジファインダ

レンジファインダの基本は、1台のパターン投影プロジェクタと1台のカメラの組合せが基本で、対応関係を対象に照射した特徴を利用し、対応精度を高める手法である。例えば、Kinectのように、近赤外でパターン投影しカメラで観測し、対応関係から距離を計測する方法や、Time-of-Flight (TOF) 法のように、光源から出た光が、対象物で反射してセンサに届くまでの光の飛行時間から、計測対象までの距離を得る方法などもある。ここでは、図4-1に示す電総研で開発された複数光源レンジファインダを説明する [58] [59]。これは光切断法に基づくレンジファインダであり、複数の独立した光源を用いることで高速に計測することができる。具体的には15個の半導体レーザと円筒レンズにより作られる15本の線状光を2進コード状にオン・オフ制御し、4回の画像入力で15本の線状画像を得て、その光点の3次元座標値を計算するものである。光源としてレーザを用いるため、線状光の幅を細くでき、またサブピクセル単位で線状光の位置を計算しているため [58]、高精度な3次元座標データが得られる。測定精度はカメラからの距離が70~110cmの距離範囲では相対誤差が約0.8mm程度である。図4-2は大きさが25cmほどのあじさいを対象物としたレンジデー

タで、カメラから約 80cm の所に置かれている。このデータは線状光間隔が約 0.12 度で計測されたものであり、3 次元データ(測定点数: 約 4 万点)を得るまでの所要時間は約 15 秒である。



図 4-1 レーザレンジファインダ



図 4-2 レンジデータ

4.2.2. データ形式

4.2.1 の装置により計測されたレンジデータは 3 次元距離データである。このデータ形式は図 4-3 に示すような 2 次元のテーブルを介して 3 次元座標値を持つものである(以後このデータ形式を RDS(Raw Data Structure)形式と呼ぶ)。2 次元のテーブルの大きさは任意の設定が可能であるが、本研究においては 512×480 というカメラの分解能と同じものとした。このテーブルで、計測点がある部分はその点の 3 次元座標値を示すテーブルのポインタを持ち、ない点は NULL で埋められる。3 次元的に距離が近いものは、投影した面上で 2 次元的に距離が近いということから、この RDS 形式はある点の近傍の点を探す際に有効なデータ形式である。

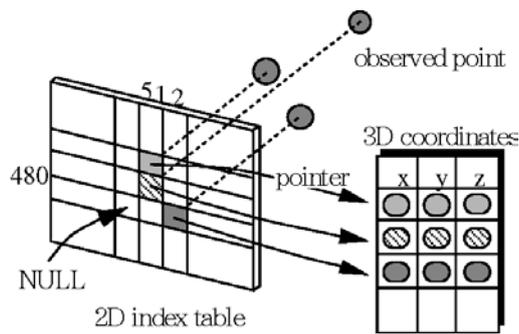


図 4-3 データ形式(RDS 形式)

4.3.3 3次元データ統合方法

レンジデータの取得からデータを統合するまでの処理手順は図 4-4 に示すものであり、大きく分けると対応領域計算，初期移動パラメータ計算，そしてパラメータ修正から成っている。まず位置関係が未知であるデータ間の関係を求めるために、仮の領域を設定し、その特徴量から対応する領域を求める。この対応領域からおおよその位置関係を求めることができ、それを移動パラメータとする。しかしこの値はおおむね適正な値であるが、正確な値とは言えない。そこでこの値を初期値として片方のデータを変換して重ね合わせ、一致している部分のずれを最小とする処理をし、逐次的にパラメータの修正を行なうものである。データ間の位置関係の情報がない場合、データ自身からその関係を求めなければならない。パラメータ値を徐々に変化させてデータ全体の点を変換、比較し、最適なものを捜す全探索的手法があるが、このように点が多い場合は効率が悪い。また比較領域を設定して、移動パラメータの概算値を求めず、共通部分のずれを最初から逐次法で最小にする処理の場合、ローカルミニマムに陥ってしまう可能性が大きく、正しい値を求めることができない。本手法のように、大雑把な初期照合を行ない、それに基づいて精度を上げる方法の方が有効である。



図 4-4 3次元データ統合処理手順

4.3.1. データ取得

複雑な形状で、オクルージョンが生じやすい物体は、複数方向からデータを計測する必要がある。データ間の位置関係は任意で良いが、見え方が大幅に変化してしまうと対応がとれなくなるので、共通部分が存在するように視線方向が大幅に変わらない範囲から計測するものとする。そのような制約下で複数のレンジデータを計測する。

4.3.2. 前処理

位置関係が未知であるデータ同士の対応関係を求めるには、比較する何らかの対象が必要である。点同士を比較するには膨大な計算量を必要とし、実用的でない。また、エッジ部分や曲率が大きく変化する部分を比較の対象とする featurematching があるが、複雑な形状の場合やデータがスパースな場合、計算量が多くなりあまり適していない。そこで、領域を適当に設定し、その特徴量を基に対応を求める手法を用いている。この方法の特徴は、対象が複雑な形状になっても、比較する特徴量を計算する時間があまり変化せず、安定していることである。

上記の考えに基づき、データ同士の対応関係を求めるための前処理として、各データに対し、領域分割、領域選択、そして領域の特徴量計算を施す。

a) 領域分割

ある注目点の近傍に他の点があれば、その点と近傍点に同一ラベルをつけるという処理により、簡単な領域分割を行なう。領域数が少なく、後の処理が実行不可能になる場合は近傍点と見なす距離のしきい値を小さくして再計算する。この分割法は複雑な形状のためオクルージョンが生じやすく、不連続領域が多く発生する対象に適用可能な方法である。この際に、孤立点、面積の小さい領域はノイズ領域として削除する。

b) 領域選択

対応関係を求める際に分割された領域をすべて使用せず、領域対応に関して信頼性が高いと考えられる領域のみを選んで使用する。この信頼性の高い領域として今回、他の視点からのデータでも見える確率の高い領域を選択する。このような領域は対応関係がとりやすい。本研究ではこの条件を満たすものとして、カメラ方向に近いもの(前面にある領域)を N 個選択する。前面にあるものほど、他の領域によるオクルージョンの影響を受けづらく、計測方向が変化しても領域の形状が大幅に変化する可能性が低いためである。また、このように領域をすべて使用せず、選択したものだけにすることで、計算時間の節約にもなる。

c) 特徴量計算

先の処理で選ばれた N 個の領域に対し、対応関係を調べるための特徴量として図 4-5 に示す面積、重心、法線、隣接関係を求める。これらの特徴は簡単に求めることができ、計算時間はあまり必要としない。

(1) 面積: 点の数 = A_i

(2) 重心: $\sum_k x_k / A_i = G_i$

- (3) 法線: 重心での法線ベクトル = NV_i
 (4) 隣接関係: 重心と他の領域の重心間ベクトル = GV_{ij}

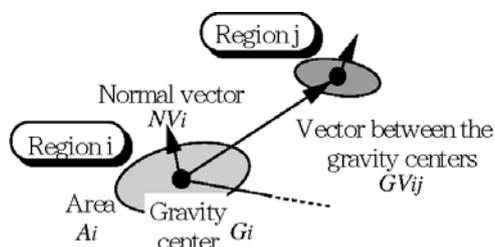


図 4-5 特徴量計算

4.3.3. 領域対応

選択された領域の特徴量を用いて各データの領域がどのように対応しているかを調べる. 仮に設定した領域の対応を求めることで, 比較するデータ同士の, 未知であった位置関係のおおよそを知ることができるためである. 今回用いた方法は, 領域間の位置関係を比較することが特徴であり, 個々の特徴がさほど正確でなくても, 他領域との位置関係はあまり変化しない点に着目したものである. なお, この手法は十分な領域数が得られていることを前提としている.

まず, N 個の選択された領域の中から m 個の領域を選択する ($2 \leq m < N - 1$). 比較するもう片方のデータでも同様に選択する. そして, 選択された m 個の領域により構成される図形が相似であるものを検索する. 相似であるとは, 構成される図形の

- (1) 重心間ベクトルにより構成される図形が相似
- (2) 対応する領域の m 個の面積比が一致
- (3) 対応する領域の m 個の法線ベクトルが一致

の条件をすべて満たすものである. この条件を満たす領域を対応する領域の候補として, 領域対応表(0 で初期化しておく)の該当欄に度数+1 を加算し, この操作をすべての組合せに対して行なう. m が ($2 \leq m < [N/2]$ の場合) 大きくなると, 検索数が多くなり計算量も膨大なものとなる. 一方, m が小さいと相似な図形が多くなり, 誤対応が多くなる. この相反する条件下で, 本研究では $m = 3$ とし, 計算コストと結果の信頼性に共に配慮するようにした. $m = 3$ で比較する図形は図 4-6 に示す 3 角形であり, これに相似な 3 角形を見つけることになる.

そして, 完成した表より領域の対応関係を決定する. 対応関係の決定は, 度数が多いものを選ぶことである. あいまい性のある対応の場合は除いて, 信頼できる対応の組だけを求めるために, 度数が多くても, 本来1つしかあるべきでない対応領域が2つ以上あり, その数にあまり差がなく, 優劣を決定できない場合は除去する. 以上の処理で対応関係が3組以上見つからない場合は, 次の 4.3.4 以降の処理を継続できない. その際は, 4.3.1 のデータ取得に戻り, 別の視点からデータを取得し, 再計算する.

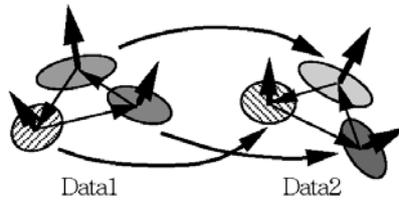


図 4-6 領域対応

4.3.4. 初期移動パラメータ計算

4.3.3 で求めた対応領域を用いて、データ間の移動パラメータ($\mathbf{x}_2 = R_0 \mathbf{x}_1 + \mathbf{t}_0$), すなわち

回転移動成分 $R_0 : (\theta_x, \theta_y, \theta_z)$

並進移動成分 $\mathbf{t}_0 : (t_x, t_y, t_z)$

の初期値を求める。まず対応する各領域の重心を一致させることで、移動パラメータを計算する。視点方向に多少のずれがあっても、対応する重心の位置ずれは少ないと考えられるため、概算値を計算するには適している。重心間ベクトルのずれが最小となるような R_0 を

$$\sum_{i \neq j, i' \neq j'} |R_0 \mathbf{G}V_{1ij} - \mathbf{G}V_{2i'j'}|^2 \quad (4-1)$$

(i と i' , j と j' が対応する領域)

($\mathbf{G}V_{rij}$: データ r の領域 i と領域 j の重心間ベクトル)

の評価関数を用いて最小二乗法の繰り返し処理で求める。次に \mathbf{t}_0 を同様に

$$\sum_{i, i'} |(R_0 \mathbf{G}_{1i} + \mathbf{t}_0) - \mathbf{G}_{2i'}|^2 \quad (4-2)$$

(\mathbf{G}_{ri} : データ r の領域 i の重心ベクトル)

の評価関数で計算する。

4.3.5. 移動パラメータ修正

4.3.4 で求めた移動パラメータは重心を一致させることで計算したものであるが、他視点から見たデータ同士の重心が完全に一致することはない。そこで、より正確なパラメータ値を求めるために、値の修正をしなければならない。その手順は次に示す(a)~(h)であるが、ここでは(b)に示すように新たに小領域を設定している。4.3.3 で得られた領域は大雑把な領域分割によるものであり、正確に一致するものではない。また、対応する領域が前面部分に限られているため、他の一致部分をパラメータ計算に考慮していないため、領域対応がとれた所から離れるに従って誤差が大きくなる傾向がある。そこで両データを小領域に分割し、その中で完全マッチしていると考えられる部分をデータ全体で求める(図 4-7)。この小領域分割により、データ同士では部分マッチングであった問題を、小領域ごとの完全マッチング問題に単純化できる利点がある。そして各小領域のずれを最小とする繰り返し処理を行なうことで移動パラメータを修正する。

- (a) 移動パラメータを $R = R_0$, $t = t_0$, 繰り返し回数 $i = 0$ と設定する.
- (b) 2次元テーブル 512×480 を $sw \times sh$ の小領域に分割する.
- (c) 比較するデータにおいて, 以下の条件をすべて満足する小領域を求める.
- (1) 領域内の点の数が共に N_t 以上
 - (2) その数の差が ND_t 以下
 - (3) 領域の重心同士の距離が GD_t 以下
 - (4) 領域に対して2次曲面を当てはめた際のずれの分散が共に σ_t 以下
($N_t, ND_t, GD_t, \sigma_t$ はそれぞれのしきい値)
- (d) 選択された小領域数が十分でない場合はステップ(c)に戻り, しきい値を緩和して, 条件を満たす小領域を求め直す.
- (e) 選択された各領域の重心近傍の点を用いて, 2次曲面 $z = f(x, y)$ (z は手前方向)を計算する.
- データ1: R, t で変換した点で f_1 を計算
データ2: f_2 を計算(このデータは固定)
- (f) ずれの評価関数を

$$E = \sum_L w_r \sum |f_1 - f_2|^2 \quad (4-3)$$

(w_r は領域 r への重み, L はステップ(c)で選ばれた小領域)とする.

- (g) E の最小の条件は $\partial E / \partial \eta = 0$ ($\eta = \theta_x, \theta_y, \theta_z, t_x, t_y, t_z$)であるが, 評価関数式が非線形のため解析的には解くことができない. そこで,

$$\eta_{i+1} = \eta_i - \kappa \frac{\partial E}{\partial \eta} \eta_i \quad (\kappa (> 0) \text{ は可変数}) \quad (4-4)$$

$$\frac{\partial E}{\partial \eta} = 2 \sum (f_1 - f_2) \frac{\partial f_1}{\partial \eta} \quad (4-5)$$

により η の値を更新していき, E を徐々に小さくしていく.

- (h) E が発散したときは, κ を半分の値とし, ステップ(a)に戻る. また, E が振動したときには, κ の値を徐々に小さくしていき, 振動を抑制するように設定する.
- (i) $E \cong 0$, $\Delta E \cong 0$, 繰り返し回数の設定値に達する, のいずれかを満足した時, 計算を終了する. その他の場合は, 新しく求めた R, t でデータ1を変換し, ステップ(b)に戻る.

この繰り返し処理によりずれを小さくし, 移動パラメータの精度を向上させることができる. このような逐次的処理ではローカルミニマムに陥る可能性が考えられるが, 4.3.4 の処理により真の極値の近傍に初期値を設定しており, その可能性は十分低くなっている.

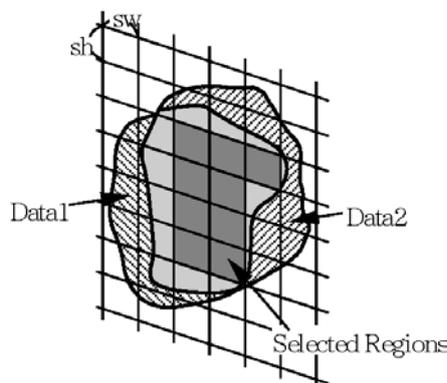


図 4-7 小領域の選択

4.3.6. データ統合

4.3.5 により求めた移動パラメータ R, t によりデータの統合が可能となる. このようにして隣接した各データについて, 上記のような手法で移動パラメータを求めていく. 対応がとれないデータ間の移動パラメータは, 既知のデータを経由していくことで, その間の移動パラメータ R, t を計算する(図 4-8).

そして, ある視点方向からのデータを基準フレームとし, その座標系を基準座標系として選ぶ. 他の視点からのデータはこの基準座標系に変換され, 統合される.

各データは図 4-3 の RDS 形式であるが, この形式は一方向からの形状を表現するのには適しているが, 全体の形状を表すのに適していない. そこで統合データ形式として, 図 4-9(a)に示される 5つのパラメータ (x, y, z, α, β) を持つデータ形式を用いた(以後, SDS(Standard Data Structure)形式と呼ぶ). これは基準座標系における 3次元座標 (x, y, z) と, その点の法線方向を表す角度 (α, β) で表したデータである. この形式の利点は視線方向に左右されず, 点の法線方向によりある視線方向からでは見えない点を判別できることである. 点の法線方向は近傍点を用いて面を当てはめて求めた面の法線方向である.

統合手順を図 4-9(b)に示す. 最初に基準フレームを SDS 形式のデータに変換する. 次に更新する SDS 形式のデータを座標変換して, RDS 形式のデータと同一座標系にし, 図 4-3 の 2次元テーブルのセルに比較するため変換された点を書き込む. このとき (α, β) と視線方向から点が可視かどうか判別し, 見える点だけを書き込む. そしてテーブルの注目するセルにすでに点が存在するならば, すでにある点より視線方向(RDS 形式のデータ座標系の $-z$ 方向)に近い場合だけセルを更新する. 同様にして SDS 形式のデータもテーブル上に書き込む. この処理を繰り返し, 新たに得られた SDS 形式のデータの点を基準座標系に逆変換し, SDS 形式のデータに追加する. 以上の操作を繰り返して, データを統合していく.

この具体例を図 4-10 で説明する. 視点1(VP1)からは面(1),(2),(3),(4)が計測できる. このデータを視点2(VP2)から計測した座標系に変換する. VP2 では法線との関係から VP1 でのデータでは(1),(3)の部分しか見えないことがわかり, 2次元テーブル上にそのデータを書き込む. また, VP2 で

計測できる面は(1),(5),(6),(7)であり, この内(6),(7)部分はテーブル上に何もないのでテーブルに追加される. (1)はVP1での(1)と位置が一致するので追加はしない. (5)の一部が(3)とテーブル上で重なるが, (5)の方が(3)より $-z'$ 方向に近いので, (5)の面をテーブルに追加する. そしてVP2で新たに追加された面をテーブル上から求め, 逆変換してVP1での座標系のデータに直し, VP1でのデータに追加する. この処理により, 面(1)~(7)を統合できる.

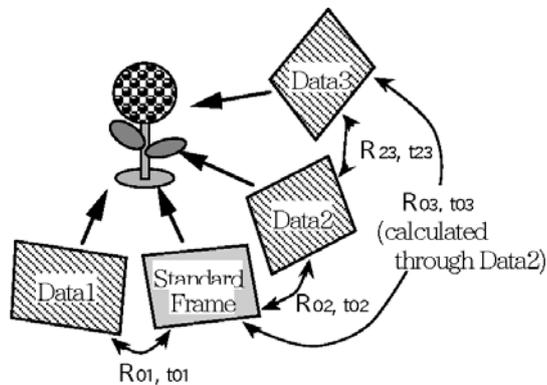
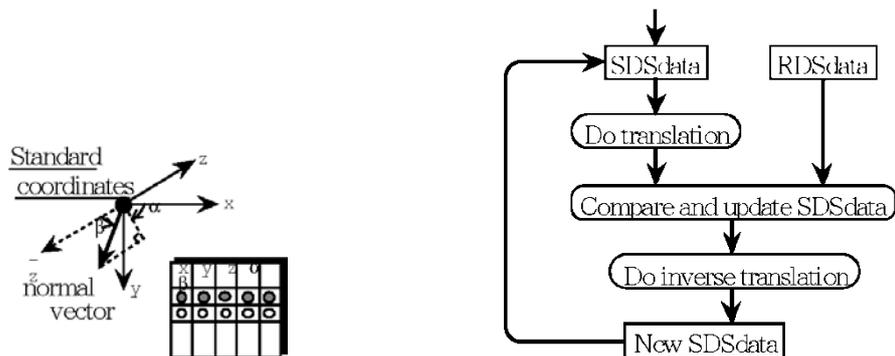


図 4-8 データ間関係



(a) Standard Data Structure (SDS) for integration (b) The flow of integration

図 4-9 データ統合

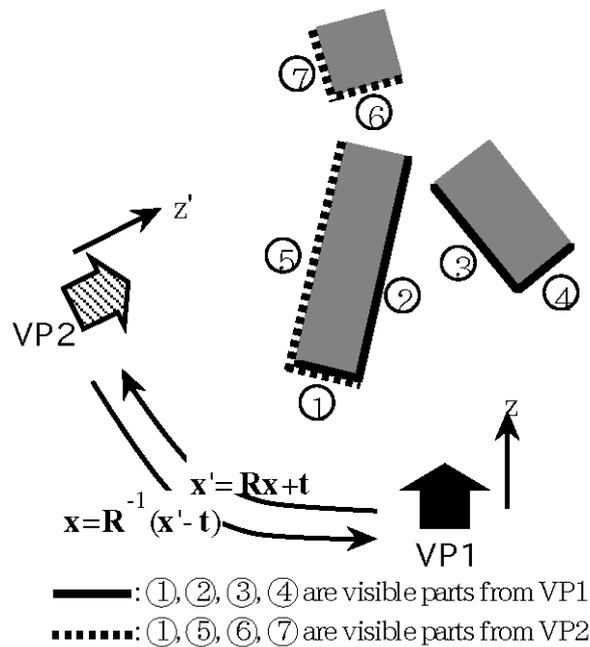


図 4-10 データ統合例

4.4. 実験

4.4.1. 結果

実験は複雑な形状を持ち、従来の方法では対応関係を得るのが難しく、扱われていなかった、図 4-11 に示す花(あじさい)で実験を行なった。あじさいの大きさは約 25cm で、レンジファインダから約 80cm の所に置かれている。これを左から右へ 7° 程度移動しながら9つのデータを取り、処理を行なった($\theta_y = 7.0^\circ$, $\theta_x = \theta_z = 0.0^\circ$)。 7.0° という角度は物体の見え方が大幅に変わらない程度の、目安の角度であり、求まったパラメータとの比較のためにこのように設定した。

各レンジデータに対して、4.3.3 に示した方法で距離による領域分割が行ない(約 50 領域)、信頼性が高いと考えられる領域である、カメラ方向に近い($N =$)20 領域を選択した。この値は、 R, t を計算する際、最低 3 つの対応が必要であることを考慮して実験的に定めた。そして各領域の特徴量である面積、重心、法線、隣接関係が計算される。図 4-12 に選択された領域とその重心、法線が示されている。



図 4-11 計測対象物体(あじさい)

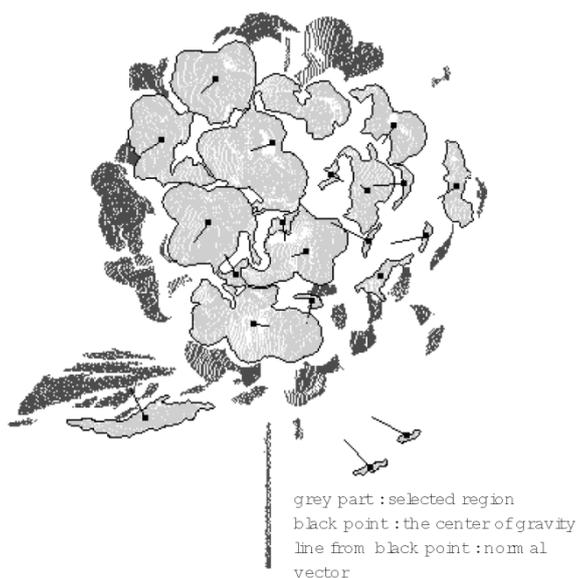


図 4-12 領域分割

次に, 4.3.3 で説明した方法で相似な3角形が検索され, 領域対応表が得られる(図 4-13). 本実験では 4.3.3 で示した条件としてそれぞれ,

- (1) 重心間の距離が 5 mm 以下.
- (2) 面積比の差が $80\sim 120\%$.
- (3) 法線ベクトルのずれが 20° 以下.

とした. これらの設定は実験的に定めた. そして得られた表から, 対応している領域が求められる. 例えば図 4-13 では, データ1の領域1とデータ2の領域3は1:1対応で十分な度数が得られているので領域同士が対応しているとみなせる. 一方, データ1の領域16に対して, データ2の領域4と

領域15の2つの領域が対応している。対応度数はそれぞれ 1, 15 であり、この場合度数差と大きさから領域16(データ1)と領域4(データ2)の組合せは誤対応となる。ゆえに領域16(データ1)と領域15(データ2)の対応が正しい。この例では 20 領域の内 7 領域の対応が求まったことを図 4-13 は示している。9つのデータによる領域対応数は平均7.1であり、領域分割がある程度ラフなものであっても、最低 3 つ以上の対応関係が求めれば移動パラメータを計算することができるので、この実験例の対応数は、計算要件を十分満たしていると言える。

		Data2																					
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20		
Data 1	1	0	0	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	5	0	0	0	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	8	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0	0	0	0	0	0	0	
	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	12	0	0	0	0	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0	
	13	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0	0	0	0	0	0	
	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	15	
	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	16	0	0	0	1	0	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	
	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

*** Result ***

7 regions out of 20 ones were selected.
 (1-3), (5-4), (8-8), (12-13), (13-9), (14-19), (16-15)

図 4-13 領域対応表

比較するデータの対応領域から移動パラメータの概算値、つまり初期値が求まる。重心は視線方向の多少の変化があったとしても大幅に移動するものではないので、これを一致させるようにすることで初期値が計算される。上記の例では、求まった7つの対応する領域の重心が一致するように、重心間ベクトルのずれを利用して回転移動成分 R_0 がまず求められ、それから並進移動成分 t_0 が計算される。この結果は、

$$(\theta_x, \theta_y, \theta_z, t_x, t_y, t_z) = (-0.38^\circ, 7.19^\circ, -1.18^\circ, -29.39mm, 0.36mm, 12.39mm)$$

であった。実験時の機器の設定では、おおよそ $\theta_y = 7.0^\circ$, $\theta_x = \theta_z = 0.0^\circ$ (t_x, t_y, t_z は任意の値) であったので、ほぼ正しい値である。この初期値を利用してデータが同じ座標系に設定される。

そして、移動パラメータの精度を向上させるため、データの比較が 4.3.5 に示した手順で繰り返して行なわれる。まず、2次元テーブル 512×480 が小領域 16×16 ($sw, sh = 16$, 約 1.5 cm^2) に分割される。そして、4.3.5(c)の条件(1)~(4)を満たす領域が求められる。これらの条件のしきい値を本実験では、

$$N_t = 128, \quad ND_t = 32, \quad GD_t = 2.5mm, \quad \sigma_t = 1.0mm$$

とした。これらの値の設定は、データ精度と実験から求めたものであるが、対象が変わったり、値が

変化したりすることで選ばれる領域が大きく変化することはない。図 4-14 に分割された小領域の内、条件を満足する領域が示されている。これらの領域のずれが最小になるように、2次曲面が当てはめられ、最小二乗法の繰り返し処理が200回施された結果、 R, t のパラメータは、

$$(\theta_x, \theta_y, \theta_z, t_x, t_y, t_z) = (-0.13^\circ, 6.82^\circ, -0.02^\circ, -27.59\text{mm}, -0.40\text{mm}, 11.16\text{mm})$$

となった。この処理により、例えば θ_z の値が大幅に改善されていることがわかる。この処理による評価関数 E と回転成分 θ_{zy} の値の推移を図 4-15 に示す。このグラフはずれが徐々に減少し、パラメータ θ_{zy} が一定値に収束しているのを表している。またこの実験での可変数 κ については、

$$\kappa_\theta = 0.05, \quad \kappa_t = 0.001$$

と設定した。初期値が真の極値付近に設定されるので、その極値に落ち着くように実験的に定めた。もし κ が大きいならば発散する可能性があるが、収束速度が速いという利点がある。また、発散した場合は κ の値を小さくして再計算し直すので、問題はない。これらの値は対象に左右されない。表 4-1 に回転パラメータの値を示す。表中の数値は、9つのデータ(左方向から右方向へ順にデータ1, 2, ..., 9)の隣同士の R_0 と R の値の表である。この表によっても初期のパラメータ値より、繰り返し処理後の値が改善されていることが示されている。

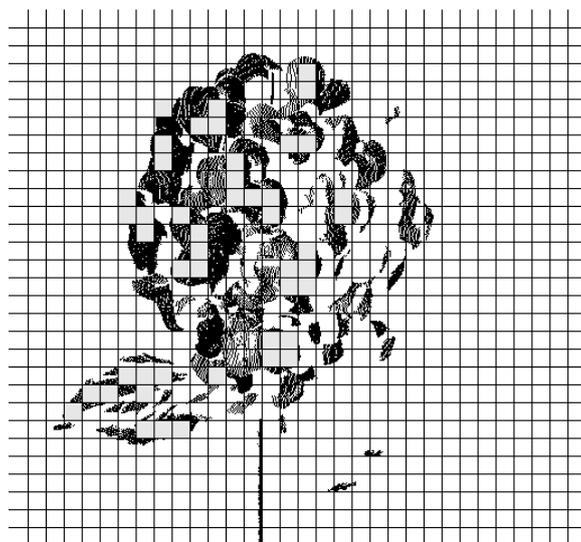


図 4-14 小領域分割

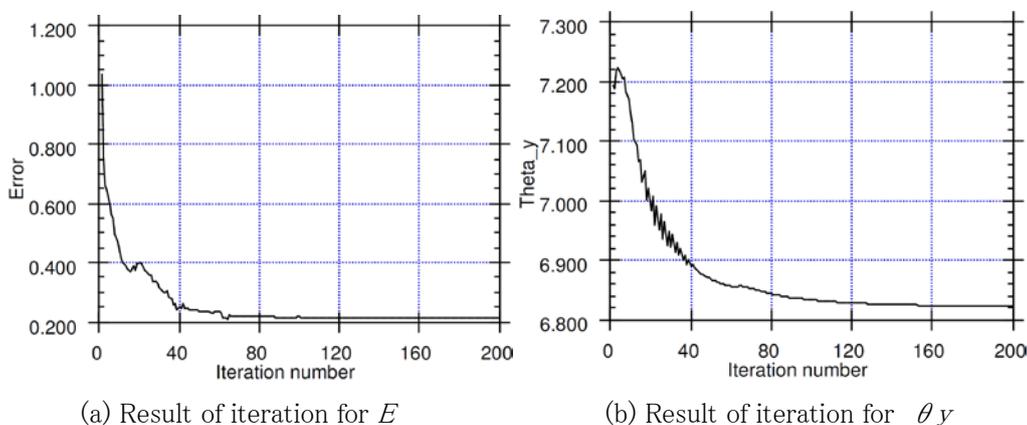
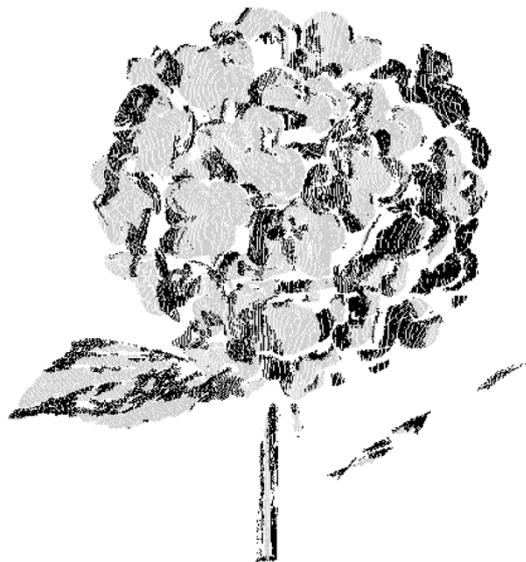


図 4-15 繰り返し処理によるパラメータの修正例

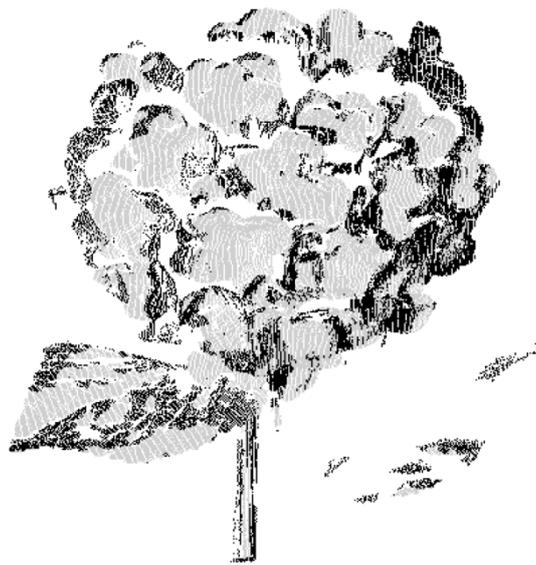
表 4-1 回転パラメータの値

DATA	θ_{x0}	θ_{y0}	θ_{z0}	θ_x	θ_y	θ_z
DATA1:DATA2	-0.329	7.236	-0.204	-0.036	7.110	-0.043
DATA2:DATA3	-0.158	7.291	-0.348	-0.039	7.031	-0.037
DATA3:DATA4	-0.623	6.994	-1.188	-0.097	6.864	-0.150
DATA4:DATA5	-0.380	7.194	-1.177	-0.131	6.822	-0.022
DATA6:DATA5	0.700	-7.794	0.104	0.152	-7.019	0.002
DATA7:DATA6	0.520	-7.864	0.410	0.174	-7.092	0.167
DATA8:DATA7	0.650	-7.572	0.133	0.046	-7.077	0.155
DATA9:DATA8	0.915	-7.360	0.687	0.147	-6.933	0.177

求まった移動パラメータにより、データ5(正面から計測したデータ)を基準フレームとして、9つのフレームを統合した結果を、図 4-16 に示す。濃い部分が基準データでは計測できず、他の視線方向から計測して得られた部分である。1方向からのデータ(図 4-2)に比べ、計測した左右の部分が良く計測されているのがわかる。(a)は統合したデータを正面から見たもので、(b)は左上方から見たものである。この実験では上方からデータを計測しなかったため、見えない部分が生じているが、計測することでこれらを解消することができる。



(a) Front view



(b) Left-upper view

図 4-16 統合結果(あじさい)

図 4-17 は, かきつばたを対象物体とした結果である. (a)のかきつばたを同様に左から右へ約 5° 程度移動しながら, 8 視点で計測し, 統合した. (b)は 8 つのデータを統合した結果を右上方から見たものである. 濃い部分が1方向だけの計測で得られなかった部分であり, オクルージョンにより見えていなかった葉の部分が計測されているのがわかる.



(a) An iris



(b) Result

図 4-17 統合結果2(かきつばた)

同様にチューリップの実験例を図 4-18 に示す. (a)のチューリップを 5° ずつ移動しながら計測した9つのデータを統合し, ボクセル形式(1.5 mm立方体)で表現したものが(b)である. b)の左側が1方向からだけのデータで, 右側が統合したものである. この例でも統合により, 花や葉部分がより多く計測されていることがわかる.



(a) A tulip



(b) Voxel representation (Left side : data from one viewpoint, Right side : integrated data)

図 4-18 統合結果3(チューリップ)

4.4.2. 統合実験結果の考察

あじさい、かきつばたなどを対象物体として行なった実験結果を示した。現在のところ、測定器具の制約により、設定は θ_y 方向(左右方向の回転)しかできない。そのため、上方・下方部からのデータはこの実験では示されていないが、測定に工夫を凝らした実験で、左上方部からのデータでも正確に統合できることがわかっている。つまりこの処理によるデータ統合が測定方向に依存しないことが判明している。

回転移動パラメータ R の統合精度評価は、単純な回転台に対象物を置き、 $\theta_y = 7.0^\circ$, $\theta_x = \theta_z = 0.0^\circ$ 設定でデータを測定して行った。結果は表 4-1 回転パラメータの値に示されているように若干の誤差があり、誤差範囲は $\pm 0.2^\circ$ である。この誤差の原因は2つあり、角度設定が目盛 1° の回転台を手動によるものであるという測定時の誤差と、今回提案した処理により生じた誤差である。しかしながら、 $\pm 0.2^\circ$ という値は十分な結果であると言える。

一方、並進移動パラメータ t の統合精度を評価するために、大きさ約 3 cm の円形や4角形で構成された物体に対して面の一致具合の実験を行なった結果、対応する面の法線同士のずれは約 0.3° 、上下方向は約 0.3 mm 、左右方向は約 0.7 mm であった。測定の相対誤差が 0.8 mm であるから、並進移動パラメータに関しても統合精度は満足できる結果であると考えられる。

統合データ形式を、今回 SDS 形式という3次元座標値と法線方向を表す角度により構成されている値の集まりとしたが、この形式の利点は、視線方向を与えた場合に、可視の点だけを求めることやデータの更新が簡単にできる点である。しかしながらある注目点の近傍点を探索する際、図 4-3 のように参照するテーブルがないため、全データをスキャンしなければならず、時間がかかる欠点がある。3次元データを表す形式として図 4-18 で示したようなボクセル形式による方法もあり、近傍探索が容易という特徴があるが、精度を高くするとデータ量は莫大になってしまい、現時点で

は実用的ではない。

また、領域対応表の上で偽の対応が現われることがあっても、対応度数の違いから正しいものと十分区別できる。一例として、実験で用いたあじさいにおいては、最終的な対応決定の正当率は約 97%であり、十分信頼できる領域対応付けの手段であることが言える。この誤対応が生じる原因を解析してみると、小領域に分割された領域が多い場合、偶然に対応してしまう組み合わせが生じることがあるためである。なお、対象形状がシンプルなものに対する実験では、対応する領域が 3 つ以上求まらない場合もあり、この場合は初期移動パラメータを計算できず、データ間の対応が求まらなかった。この対策としては、改良によるより安定した領域分割法が必要である。

4.5. まとめ

本章では、第 3 章に引き続いて、視覚障害者が生活する上においても大変重要で、必要な 3 次元の視覚情報の取得のうち、多視点からの情報を統合する方法について述べた。観測対象が複雑な物体の場合、1 方向からの計測では十分なデータを取得できないため複数方向から計測する必要があるが、その際にデータ計測の位置関係が未知であっても、信頼性の高い領域を利用することで対応領域を求め、その関係からデータ間の移動パラメータを計算できる。そして、面当てはめによる繰り返し処理を行ないその精度を向上できる。複数のデータはこれらのパラメータを基に統合可能である。花などのように、形状が複雑な物体を対象物として実験を行ない、提案する手法の有効性を示すことができた。統合された 3 次元の点群情報は、第 5 章で述べる 3 次元触覚ディスプレイにより対象の表面形状を表現したり、シーン中の対象物を検出したりする際に利用できる。

統合手法に関する今後の課題としては、

(1) 領域分割法の改良

領域分割は現在のところ、点間の距離による単純な方法で十分な結果が得られているが、エッジ、曲率などの情報を用いたセグメンテーションにより、より正確な領域設定を行い、初期移動パラメータ値の精度向上をはかりたい。

(2) 計算時間の短縮

移動パラメータ精度向上の際、パラメータ値を更新するごとにデータ全体の座標変換を行っているため、精度は良くなるが、時間を要するという欠点があるため、計算時間を短縮するようにアルゴリズムの改良をはかりたい。

(3) データ形式の確立

3 次元物体を正確に表現し、処理しやすく、かつ、サイズが小さいデータ形式を検討したい。が挙げられる。

第 3 章と本章において、視覚障害者が注目する対象や環境の情報を得ようとした場合に必要とされる 3 次元の視覚情報の取得、処理手法に関して述べたが、このような 3 次元視覚情報処理技術であるコンピュータビジョン技術は大変重要である。適用先は障害者支援のみならず、様々な支援システムや、作業の自動化などに不可欠な要素技術であり、まだまだ未解決な技術課題が数多くあり、今後も研究開発に取り組んでいかなければいけない。このようにして得られた 3 次元視覚情

報は、続く第 5 章、第 6 章において、視覚障害者のための支援システムにおける情報取得部に活用されるものである。

5. 対話型触覚ディスプレイによる 3 次元情報提示システム

本章では、今一度視覚障害者への視覚情報提示システムに立ち返り、第 2 章での基本システムの開発の経験を基に、第 3 章、第 4 章での 3 次元視覚情報取得、処理技術を適用し、新たに開発した対話型 3 次元触覚ディスプレイを中心とした、3 次元視覚情報提示システムについて述べる。情報取得部、提示部の改良により、視覚障害者への提示情報量を増やことができ、必要な 3 次元情報をより得やすくなるようにした。多様な 3 次元視覚情報を、音声情報とともに提示でき、立体表面形状も表現できる触覚ディスプレイによる支援システムであり、視覚障害者は、システムと対話的なやりとりで 3 次元情報を効率よく取得できることを、被験者実験の結果を通して示す。

5.1. はじめに

視覚情報者用の支援システムに関する製品は、第 1 章で示したように、OPTACON、Web ページ読み上げソフトなどを始めとしてその大半は文字拡大、読み取り、読み上げなどであり、対象は「文字」が中心である。しかし文字以上の情報量を持ったもの、つまり図面(2 次元)や物体、空間(3 次元)に対するアクセス手段は限られているのが現状であり、人に説明してもらうか、実際に触ってみなければならぬことが多い。このような 2・3 次元に対する視覚支援の研究開発 [7] [60] [61] [62] [63] [64] [65] [66] [67] [68] [69]が望まれている。この視覚障害者の要望を満たすよう、図面認識支援システムを開発した(第 2 章)が、対応すべき課題として、日常生活を送る上で重要だと考えられる 3 次元空間情報を扱えること、触知覚により得られる情報を多くするため、より提示情報量が多い触覚ディスプレイへの性能向上、の必要性を感じた。

3 次元触覚ディスプレイに関しては、第 2 章で開発した触覚ディスプレイは、8×8 のピンで構成され、高さ段階が 2 段階であり、その解像度の低さが問題であった。より高解像度のディスプレイとしては、ドイツ metec 社が製品化しており、表示部は 372×186 mm で 120×60 ピンで構成されている。これは空間解像度も高く、図形をそれなりに表示できるが、このシステムのピンの高さはオン・オフの 2 段階(0, 0.7 mm)であるため、3 次元表示には適していない。より解像度が高い 3 次元触覚ディスプレイに関しては、医療福祉機器研究所において「視覚障害者用三次元情報表示装置」が開発された [67] [68]。ピンは 64×64 の千鳥格子配列で、ピン直径 2.5 mm、間隔 3 mm、高さ 0～10 mm の可変であり、3 次元物体の形状表示が可能なディスプレイである。ただし、システム全体の大きさが机 2 個分あり、配線も複雑で装置のメンテナンスが事実上不可能な装置である。

本章では、第 2 章で開発した情報提示の基本システムで明らかになった課題解決を目指し、日常生活を送る上で重要だと考えられる 3 次元空間情報の提示に適用対象を拡げ、3 次元の情報をより伝えやすい触覚ディスプレイの開発、システム化について述べる。第 3 章、第 4 章において行ったコンピュータビジョン技術を本システムに適用し、タッチセンサの機能を有している 3 次元触覚ディスプレイを用い、視覚障害者がシステムと対話的なやりとりをすることで、必要な 3 次元情報を効率よく取得できる点、空間認知をしやすくなる点を、実験を通じて示す。

5.2. システムの概要と特徴

開発したシステムは、第2章での基本システムで課題となった、日常生活を送る上で必要な3次元の視覚情報を扱えることができ、触知覚による提示情報量をより多くできる触覚ディスプレイの性能向上を目標にした、新たな視覚情報提示支援システムである。視野が広く可変のステレオカメラから非接触に得られる多様な3次元視覚情報を、対話型3次元触覚ディスプレイと音声によって表現する支援システムである(図5-1)。システムは入力、処理、出力部からなり、入力部は2台のカメラシステムを用いている。処理部ではステレオ画像を解析し、シーン中の物体、環境の3次元構造復元を行う。得られた視覚情報を単純に触覚・聴覚にメディア変換するだけでなく、物体認識を計算機側で行うことにより、情報量の圧縮を行い、必要な情報のみを提示する。そして、これらの情報を対話的なインタフェースで触覚・聴覚を通して使用者に伝える。

開発した対話型3次元触覚ディスプレイの特徴は、単に3次元情報を表示できるだけでなく、視覚障害者がどの部分を触っているかが分かる、タッチセンサの機能を有している点である。この機能により視覚障害者は一方的に情報を与えられるだけでなく、システムとのやりとりをしながら対象を理解していくことが可能となる。また、過去の研究例 [64] [65] [66]でその有効性が示されているように、触覚と聴覚の両感覚を用いており、得られた視覚情報を理解しやすいシステムとしている。

本章では、触覚を通して情報提示をする装置である対話型触覚ディスプレイを中心とした3次元物体認識支援システム(図5-1)について説明する。

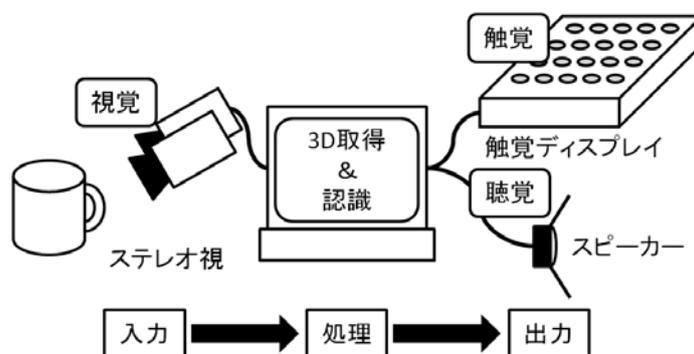


図 5-1 支援システム概念図

5.3. 触覚ディスプレイシステム

対話型触覚ディスプレイシステムのハードウェア部分の詳細を説明する。システム全体は図5-2であり、左から触覚ディスプレイ、カメラ、カメラ制御部、ワークステーション、そして合成音声出力装置が並んでいる。システム構成は図5-3であり、システムの入力・出力部であるステレオカメラシステム、触覚ディスプレイ、合成音声出力装置について以下に説明する。

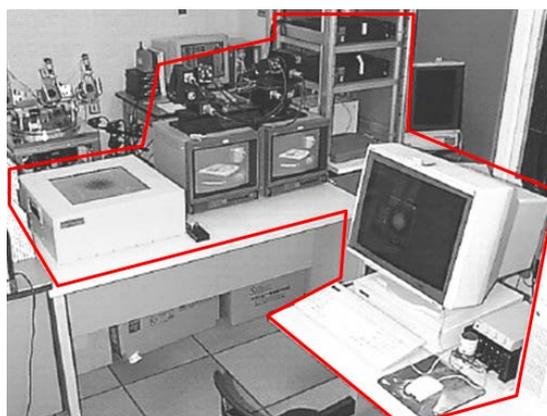


図 5-2 システム全景

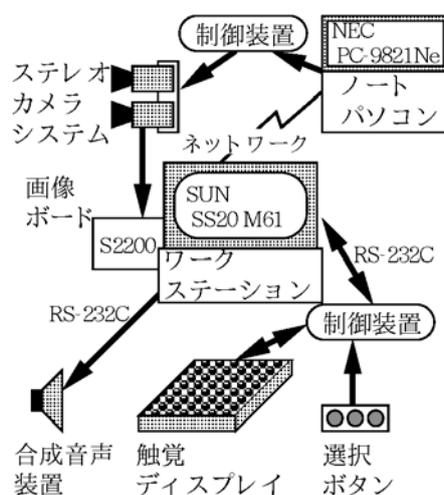


図 5-3 システム構成図

5.3.1. ステレオカメラシステム(入力部)

入力部であるステレオカメラシステム(図 5-4) [70]はパーソナルコンピュータ NEC PC-9801Ne (OS:Lnyx マルチプロセス対応リアルタイムOS)により制御される。各カメラはパルスモータ駆動によるパン、チルト制御が可能で、最大 50° /秒の回転速度であり、移動物体の追従には十分な性能である。カメラは Sony の 3CCD カメラ XC-007 で、画像入力ボードは SUN SPARCstation の s-bus 用 Data Cell Limited 社の S2200 ($640 \times 480 \text{ pixel}$, RGB 画像)を使用している。

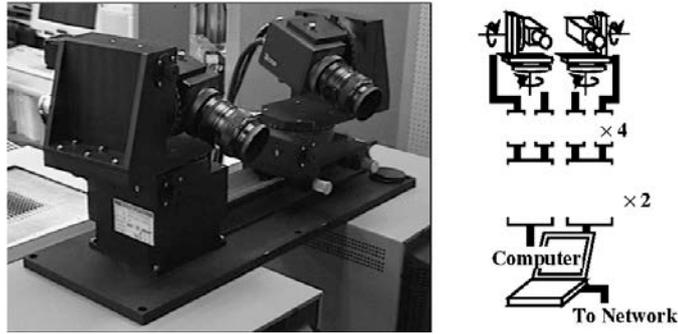


図 5-4 ステレオカメラ

5.3.2. 触覚ディスプレイ(出力部)

触覚を通して情報を提示する出力装置として3次元触覚ディスプレイを開発した。この触覚ディスプレイ(図 5-5)は、2次元の格子状に配置された16×16の凸状のピンで、パターンを表示するものである。ピンの高さは触覚情報量を増加させるためと、3次元形状を表示するために、数段階に制御可能となっている。他の触覚ディスプレイとの大きな相違点は、タッチセンサの機能を有している点である。つまり、出力装置としてだけでなく、入力装置としても利用できる。また、音声情報を聞くことや表示モードを切替えることに使用する選択キーも用意されている。

触覚ディスプレイの仕様は表 5-1 の通りである。

表 5-1 触覚ディスプレイ仕様

ピン配列	16×16 本
表示部寸法	175×175 mm
ピン直径	5 mm
ピン間隔	10 mm
ピン高さ	0~6 mm (1 mm 単位)
ピン形状	
ピン駆動	ステッピングモータ
駆動速度	3 mm/sec.
スイッチ	タクトスイッチ
外寸	550(W)×530(L)×195(H) mm
重量	28 kg

ピンは格子状に配置され、その内部構造を図 5-6 に示す。ステッピングモータを駆動するためのボードが外周部に配置されている。ピンが指で押された場合でも、同じ高さを維持できる力を持った、直径 10 mm のモータを採用した。ピン間隔を狭めるために2層構造にし、ピン間隔を 10 mm にした。ピン上部のスクリュー構造がモータの回転力を垂直移動に変換している。各ピンの最下部にカード式電卓で使用されているタクトスイッチを設け、どのピンが押されているかを知ることができ

る。

また選択キーは計算機のマウスとほぼ同じ大きさで、3つのボタンで構成されている。これは使用者が音声情報を聞く場合や、表示モードを変更する時に利用する。

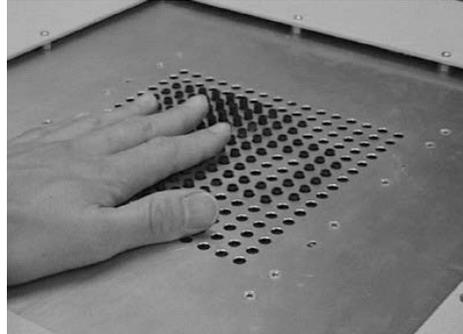


図 5-5 触覚ディスプレイ

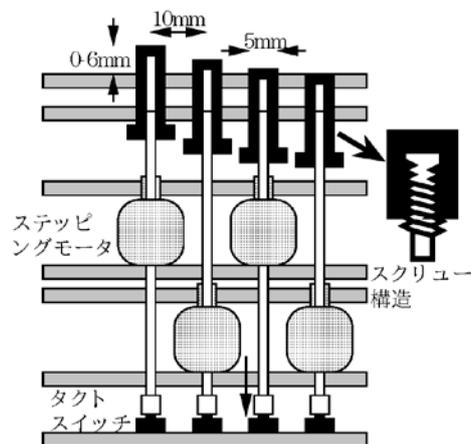
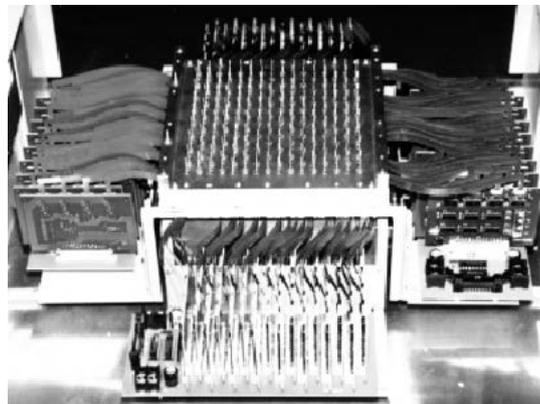


図 5-6 触覚ディスプレイ内部構造

5.3.3. 合成音声出力装置(出力部)

聴覚への情報提示として、本システムでは合成音声出力装置を用い、触覚では伝えられない情報や、触覚より簡潔に提示できる情報のために使われる。RS-232C を通して制御可能で、日本語音声出力可能なリコーの VC-1 を使用している。

5.4. 3次元情報の取得

入力部で得られたステレオ画像は、第3章で説明した手法などで処理され、対象部のワイヤーフレームのように輪郭部分の3次元情報、および、面部分の3次元情報を得ることができる。

5.5. 表示モード

システムと視覚障害者との対話システムに関しては、ディスプレイの解像度は現在16×16とまだまだ十分でないため、一度にたくさんの情報を表示できない。これを補う方法として、多段階表示を採用している(図5-7)。表示モードには以下の3つがある。

- (1) 全体像モード(環境)
- (2) 輪郭像モード(物体)
- (3) 凹凸像モード(物体)

認識された環境は、まず(1)全体像モードでそれぞれの物体の空間的配置関係が把握できるように、上方から見た配置図が表示される。1つのピンの突起が1つの物体を表し、使用者はディスプレイを指で走査することで物体の位置関係を理解する。このモードでは、音声情報キーを押すことで、「全部で4つの物体があります。」といった全体情報を聞くことができる。また、タッチセンサ機能により各ピンを押すと合成音声で「コップです。」といった物体名も聞くことができる。次にその物体の形状を知るために表示モード変更キーを押し、(2)輪郭像モード、または、(3)凹凸像モードに移る。これらのモードでは物体の輪郭線、3次元である凹凸が表され、物体の形状を知覚することができる。第2章での基本システム開発時に、情報表示はユーザによって好ましいものが異なり、選択できるようにすべきであるという知見を得られており、物体表現もこのように2種類用意した。音声情報として、図5-7の例のように「把手のあるコップです。大きさは約10cmで、色は白です。」といった、物体の大きさ、色なども得られる。これら3つの表示モードを行き来し、3次元視覚情報を触覚と聴覚で理解することができる。

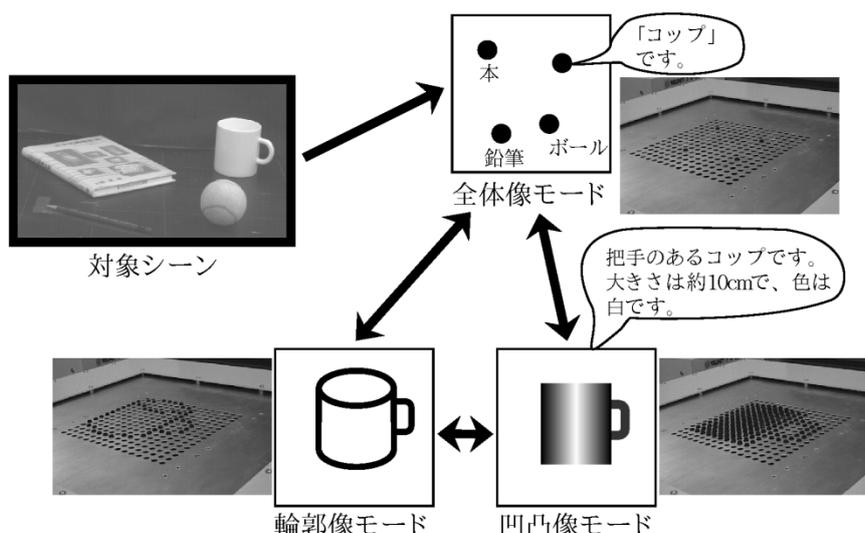


図 5-7 三種の表示モード

5.6. 評価実験

開発した 3 次元触覚ディスプレイを中心とした、情報提示システムに関して、その仕様やインタフェースの使い勝手などを視覚障害者に評価してもらった。被験者は、普段触図を利用している2名の視覚障害者(全盲[後天盲], 弱視[指数:20cm])と使用していない3名(弱視)の、計5名の大学生(20代前半の男子)である。

評価実験を始めるにあたり、まず被験者に対して、システムの趣旨、概要(特に、ディスプレイの制約(縦横、高さの解像度が低いこと)を口頭で説明した。システムに不慣れな状況では、正当な評価が望めず、否定的な見解が多くなることもあるため、約 10 分間の触知覚の練習をしてもらい、システムに慣れ親しんでもらった。

評価実験としては、最初に、3次元触覚ディスプレイの仕様に関する評価実験(実験 1)、物体の形状表現モードに関する評価実験(実験 2)、物の配置を表示する全体像表示モードに関する評価実験を行った。そして、最後に参考実験として、触図利用者の2名に関しては、触覚ディスプレイを触覚地図表示装置として用い、地図表示システムとしての妥当性評価実験を行った。なお、被験者が視覚障害者であることを考慮して、各評価実験に関しては、評価してもらいたい項目を口頭で質問し、口頭返答するものとした。

5.6.1. 実験1:3次元触覚ディスプレイ

実験 1 では、まず 3 次元触覚ディスプレイのハードウェア仕様(表 5-1)について評価をした。特に表示される3次元物体の指先での感覚やピンの配置などについて調べた。被験者は少し練習をした後、音声ガイドなしの状態、ディスプレイに表示された様々な 3 次元形状を手のひらで走査し、表示サイズ、ピンの大きさ、間隔、形状、解像度などの評価を行ってもらった。

この実験結果は、指先で走査するには適切な大きさ、ピンサイズであり、手のひら全体で形状を感じられて良く、特に問題はないとの意見を各被験者から得た。解像度に関しては、より詳細な形状表現をできるものを望む声が多かったが、単純な対象であれば、この解像度でも十分知覚できるとの意見であった。また、ピン駆動をステッピングモータで行っているため、その応答速度に少々難ありとの指摘もあった。

5.6.2. 実験2:物体形状表現モード

実験 2 では、図 5-7 のうち、物体形状表示モード(輪郭像モード、凹凸像モード)における表示方法の評価を行った。第 2 章の基本システム開発での知見から、好みの表示モードはユーザによって異なったためであり、その知見の有効性確認を行うためでもある。評価実験としては、2つの物体形状表示モードでカップ、ボールを表示し、その形状把握度を評価した(図 5-8)。例えばカップに関しては、(a)横から見た輪郭像、(b)上方からの輪郭像、(c)斜め上方からの輪郭像、(d)横からの凹凸像の4パターンで表示し、触知覚だけではどれが一番理解しやすいか、どのような組み合わせが良いかを調べた。

結果として、「(a)がベスト」、「(c)の方がよくわかるが、(d)と組み合わせるとより良い」との回答を得た。(a)が良いとする被験者は、触覚情報が複雑になると分かりづらくなるという観点から、シンプルな表現方法を望んだ。(c)を選んだ被験者は、(a)では他の物体も想像できるが、(c)は自分の持っているコップの形の知識と一番一致するためとのことであった。(d)は触覚ディスプレイの仕様の制約から、高さ方向と縦横方向のアスペクト比が約 1:6 と、かなりつぶれた形となる点や滑らかな変化はエッジ部分を把握しづらい点などから、ベストな選択肢にはならなかった。

一方、ボールに関しては、「輪郭像モードの表示法の(e)ではただの円となり、たとえ音声情報の補助があっても、球という認識はしづらい」との理由で、この物体形状においては凹凸像モードの表示法の(f)が適していた。

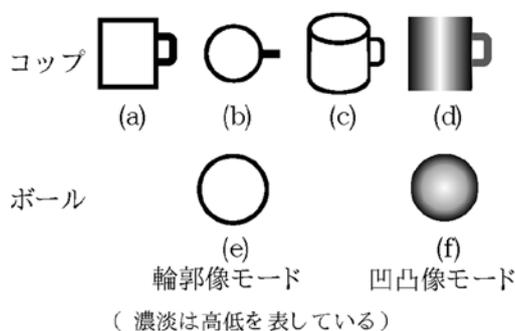


図 5-8 実験2

5.6.3. 実験3:全体像モード

実験 3 では、全体像モードでの表示の仕方に関して評価を行った。まず、いろいろな視点からの表示について調べるため、物体の配置を上方から見た形と、カーナビゲーションで採用されているような上方の視点から斜めに見下ろすような形の表示法の比較を行った。その結果、後者は、ディスプレイの低解像度が影響はしているが、触覚情報が複雑になりすぎる点でわかりづらく、一般的な地図と同様な前者の表示法が一番であるとの結果が得られた。

また、この結果を受け、上方から見下ろした形で表現した場合の、個々の物体を表わす方法として、物体の大きさを反映させるかどうかを検討するため、1つの物体を(a)複数のピン、(b)1つのピンで表示する方法の2通りを比較した(図 5-9)。結果としては、(b)が良く、(a)の表示では物体がいくつあるのかが分りにくくなる理由で評価は低かった。

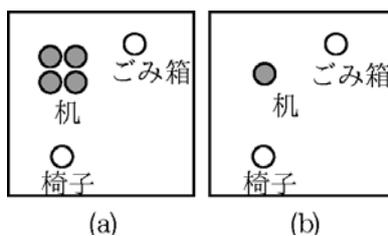


図 5-9 実験3

5.6.4. 参考実験: 触地図認識

最後に参考実験として、触地図使用者に、この対話型触覚ディスプレイを用いた電子触地図の理解のしやすさに関する評価実験を実施した。図 5-10 に示す、被験者が住んでいる地区の概略地図を触覚ディスプレイ上に表現した。地図では、主なランドマーク、道路、交差点、鉄道が異なる高さのピンで表現されており、各ピンを押すと剛性音声でその情報を伝えるようになっている。

実験は、地図が表示されていることを被験者に伝え、利用してもらい、使用後に、地図全体が把握しやすいか、ランドマークの位置が分かりやすいかなどの質問を行い、地図表示システムとしても有効であるかの評価を行った。

結果は、音声アナウンスもあるため地図を把握しやすいとの回答が多く、地図表示システムとしての評価はかなり高かった。基本コンセプトはこれで十分であるが、要望として、地図の倍率を簡単に換えられる機能が必要という意見があった。

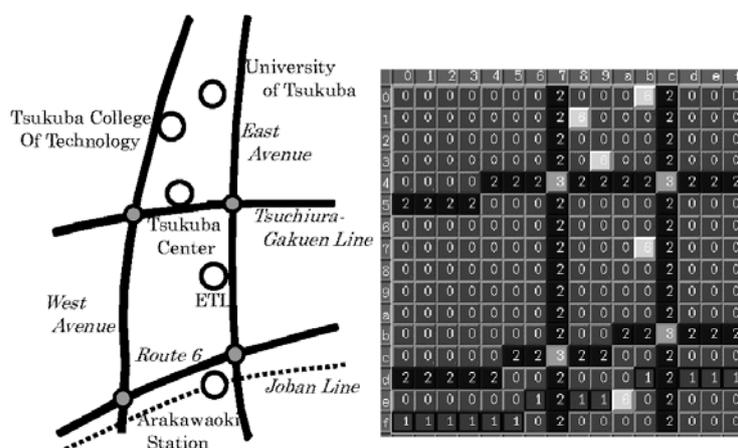


図 5-10 参考実験

5.7. 考察

今回開発した触覚ディスプレイは、ピンの高さが可変である点が特徴であり、3次元情報を表示でき、実験1の結果から、対象、目的によってはこの仕様でも十分であるとの知見が得られた。しかし、まだプロトタイプレベルであり、解像度も16×16と低いため、細かい表現ができない問題点がある。ただし、より高解像度なディスプレイはメンテナンス性が悪く、高価であるといった点で非現実的である。現在のところ触覚ディスプレイ自体の開発よりも、総合的な視覚支援システムにおける触覚ディスプレイの効果的な使用方法としての、聴覚情報をも含めた対話的なインタフェースの研究・開発に重きを置いている状況である。

実験 2, 3 の被験者実験の結果から、表示モードに関しては、設定した3つのモードで十分である点、物体の形状は輪郭像、凹凸像を使い分けられる点、そして全体像では物体の大きさに関係なく1物体は1ピン表示が良い点などが実験を通して得られ、図 5-7 のインタフェースの妥当性が確かめられた。第 2 章で得られた知見がここでも証明されたと言える。

触覚情報の表示方法については、触覚ディスプレイのハード自身の制約が大きく影響するが、複

雑な視覚情報は把握しづらいため、極力シンプルな表示が望ましいと言える。凹凸像表示モードでは立体形状を表現することができるが、縦横の長さ比べ高さ方向の変化が少ないため、正確に3次元形状を表示できるとは言えない。例えば、ボールを表示した場合凸状に盛り上がった円形になり、半球を表示できない。しかし、定性的な立体形状を把握することはできる。またエッジ部分がはっきりしていない場合、つまり、高さがなだらかに変化している場合はその変化を知覚しづらいという傾向がある。凹凸像モードのこれらの欠点を補う意味で輪郭像モードがあり、これら2つの表示モードで形状を認識するインターフェースになっている。また、問題点として応答速度の向上、Undo ボタンの必要性、音声として提供している情報の点字表示などがあげられた。しかし、その際には触覚の切り分けや、手の移動量の増加による混乱が生じないための設計が必要である。

音声に関しては、全体像モードでは情報ボタンを押すと、物体数、各ピンを押すと物体名が出力され、形状表示モードでは情報ボタンを押すと、さらに詳細な物体の特徴を得られる。また、表示モード移動ボタンを押すと、移動先のモード名を合成音声で出力する。この音声表示方法に対する被験者の意見は、全体像モードでも物体名だけでなく、より詳細な情報が必要とのコメントもあったが、触覚情報との組み合わせを考えると、あまりに音声での情報量が多いと忘れてしまうため、一度に出力する量としては、この程度が適量との意見もあった。また、音声は一過性なので、触覚ディスプレイの下部に点字ディスプレイを組み込み、常時触覚で情報を得られる方がうれしい、との意見が点字を利用している被験者2名から出た。

また、触覚ディスプレイを電子触地図へ適用できるかどうかの参考実験でも、触覚と音声の統合による認識のしやすさが認められた。対話的なインターフェースを有する本システムの有効性を別の角度からも示すことができた。

5.8. まとめ

本章では、3次元の視覚情報を取得、提示できる視覚障害者用支援システムとして、触覚と音声を用いた対話型3次元触覚ディスプレイシステムについて説明した。触覚ディスプレイは触覚による情報量を増加するため、および、3次元形状を表現するため、16×16のマトリクス状に触覚ピンが配置され、6段階の高さ表現ができる。また、各ピンにスイッチを設けることでタッチセンサの機能も搭載している。この機構により、使用者は本システムと対話的なやりとりが可能となり、ステレオビジョンにより得られた3次元の視覚情報を効率よく取得できる点、空間認知をしやすくなる点を、被験者実験を通して示した。今後、更なる被験者実験を通してシステムの問題点、改良点を見出すとともに、コンパクト化、実時間化、柔軟性のある対話システムの開発などを行う必要がある。

このシステムは、触覚を主に用いて、視覚障害者に物の形状、配置を伝えることを目的にしたものであるが、視覚障害者が必要としている視覚情報をすべてまかなえるものではない。1.1.2 行動支援で示したように、自立行動を補助する支援システムも必要とされている。残存感覚で利用できる触覚と聴覚のそれぞれの特性も第1章で記載したが、静的状態の対象の状況を情報提示するシステムとしては、このような触覚を中心としたシステムが望ましく、情報もうまく提示することができるが、歩行時などに必要となる、状況が刻々と変化する状況に対応できる情報提示システムに関して

も検討しなければいけない。この問題に関しては、次章で述べる。

6. 音響表現を利用した 3 次元情報提示システム

第 5 章では、3 次元の視覚情報を主に触覚で視覚障害者に情報を提示する支援システムを説明したが、状況が刻々と変化する情報を提示するには、触覚ではなく聴覚を利用すべきである。立体音響表現によって 3 次元の視覚情報を提示支援システムについて述べる。視覚障害者が知りたいと思う 3 次元状況を計測、解析し、ユーザが設定したタスクに応じて、その状況を 3 次元音響で提示するプロトタイプシステムを開発した。3 次元音響提示で重要な、3 次元音像定位に関する被験者実験を行い、音像定位は上下方向、前後方向で誤認識を起こしやすい点、頭を動かすなどの能動的な動作が有効である点などが判明した。また、仮想音を付け足す形で音響情報を提供する方法のため、視覚障害者は使用に関して多大な学習をすることなく、今まで養ってきた聴感覚を損なわずに使用できる可能性も示した。

6.1. はじめに

視覚障害者用の 3 次元の視覚情報提示支援システムとして、第 5 章において、物体の 3 次元形状や物の配置を触覚ディスプレイにより提示するシステムを述べたが、状況が刻々と変化し、空間的な広がりを持つ 3 次元の視覚情報に関しては、触覚では十分に伝えることが難しく、第 1 章で説明したように、触覚ではなく聴覚を用いて情報を提示するシステムを検討する必要がある。

視覚情報を音響表現で提示する支援システムに関する従来の研究としては、画像を音の高低や周波数を変換してユーザに提示する音響表現システム [8] [9]、スピーカーアレイを利用したシステム [71]、ステレオ効果を利用したもの [72] などがある。これらは対象シーンが 2 次元であったり、音像定位によって正面の 2 次元平面における位置認識を試みたりするものであったりする。一方、近年もバーチャルリアリティ技術の発達により、頭部音響伝達関数に基づいて仮想空間における聴かせる技術の進歩も著しく、3 次元音響を利用し、コンピュータの GUI 画面情報を提示する研究 [13] があるが、2 次元視覚情報をの 3 次元音響提示であり、また 3 次元音響情報が十分に生かされているとは言えない。

そこで本章では、空間把握に適していて即応性のある聴覚を利用し、3 次元の視覚情報を提示する支援システムとして、3 次元の仮想音響技術による情報提示支援システムについて述べる。開発したシステムは、3 次元環境を画像解析により自動計測、認識し、ユーザが設定したタスクに応じてその状況を 3 次元音響で提示することを目標としており、従来の研究とは異なる。頭部に装着した、小型のステレオカメラから得られたステレオ画像を解析して空間計測、対象認識を行い、その結果を 3 次元音響出力でユーザに伝える。インフラが未整備の場所や情報提供できない対象に対して有用で、ユーザの周りの状況を提示できるシステムである。また、周辺環境からの音を塞ぐことなく仮想音を付け足す形で音響情報を提供する形を取るため、ユーザは使用に関して多大な学習をすることなく、今まで養ってきた聴感覚を損なわずに使用できる可能性がある。このような特性を活かすことができる適用事例としては、図 6-1 に示すように、視覚障害者の要望アンケートで上位にあげられる、歩行支援やスポーツなどでの利用が考えられる。例えば、歩行支援では、転落の危

険性があるが、インフラが整備されていない下り階段の場所の提示など、スポーツでは、晴眼者が視覚障害者にレベルを合わせるのではなく、普通の道具を使ってキャッチボールを楽しむなど、視覚障害者の要望を満たすためには必要な支援システムの一つになりうる。以下では、構築したプロトタイプシステムの詳細と、3次元仮想空間での音像定位の特性を評価するために行った被験者実験結果について述べる。

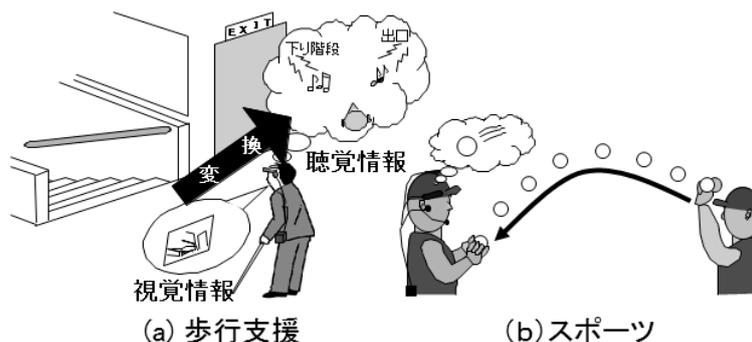


図 6-1 システム実現例

6.2. システム概要

支援システムを開発するための視覚情報処理、システム制御、音響表現に関する実験を行うにあたり、図 6-2(a)に示すプロトタイプシステムを構築した。(b)は視覚センサとしてのステレオカメラシステムであり、(c)は音響出力部と音声入力部を備えたヘッドセットである。現段階では被験者実験を主眼としたプロトタイプのため、システムはウェアラブルな大きさではない。本システム構成は図 6-3 であり、ステレオカメラシステム、3次元音響システム、3次元視覚情報処理、システム制御、および、視覚情報の音響表現に関して以下に説明する。



(a) システム全景



(b) ステレオカメラ



(c) ヘッドセット

図 6-2 支援システム概観

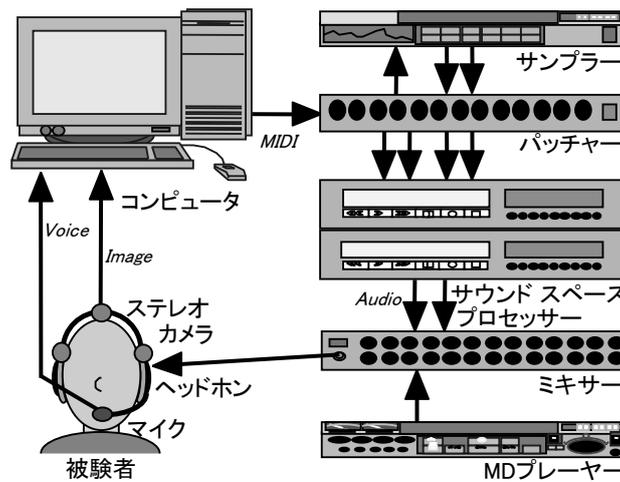


図 6-3 システム構成

6.2.1. ステレオカメラシステム

視覚センサとして超音波センサなどいろいろなデバイスがあるが、本システムでは第 3 章で説明した 3 次元センサのうちのステレオカメラを用い、得られる画像の解析により計測・認識を行う。この方法の利点は遠距離物体の距離計測、認識が可能で(例えば、遠方の信号機の有無や赤／青の識別)、状況が刻々と変化する場面での使用に適していることである。画像解析法は一般に処理が複雑になりがちであるが、第 3 章に一部示したように、多分野・多目的に利用できる汎用な 3 次元視覚システム [73]を開発している。複数台のカメラで得られるステレオ画像を解析することで、対

象シーン中の物体の 3 次元構造の復元, モデルとの照合による認識, そして, 運動追跡などが可能となっている。

視覚情報入力部はユーザが頭部に装着するデバイスであるため, 小型軽量であることが望ましいが, 計測精度などの問題があり, 頑丈であることも必要である。今回は, ヘルメットに 3 台の小型カメラをアルミフレームで固定し, 装着したものを作成した(図 6-2(b))。カメラは 1/4 インチ CCD のカラーカメラで, 直径 7mm, 重さ 68g (3.5m 長のケーブルも含む)と小型軽量のものを使用し, ヘルメット全体の重量は約 650g である。レンズのフォーカスは白杖が届かない 3m 以上に合うような設定とした。カメラを 3 台用いる理由は, ステレオ画像解析処理における水平線の対応問題を軽減するためである。

6.2.2. 3次元音響システム

近年バーチャルリアリティ技術の発達により, 仮想空間における聴かせる技術の進歩も著しい。3次元音響機器もいくつか市販されており, 手軽に仮想空間における立体音響を生成できる。今回は Roland 製のサウンドスペースプロセッサ RSS-10 を中心としたシステムを構築した(図 6-2(a)左)。

RSS-10 は, 無響音室で測定された頭部音響伝達関数 (HRTFs)により, 左右の耳に到達する音の特性をシミュレートすることで音の方向感, 反射特性や減衰などをパラメータとした1次反射音をシミュレートすることで距離感, 残響音で臨場感を, ドップラー効果により移動感を表現できる装置である。与えたオーディオデータを加工することにより, 3次元空間上で1チャンネルの音像を自由に動かすことができる。本システムでは, この装置を最大 4 台用いることで, 3次元仮想音響空間に同時に4つの音源を生成させることができる。ただし, HRTFs は個人により異なるため, この装置で使用されている関数と近い特性を持った人にはかなり正確に仮想音響空間を認識できるが, 大きく外れている人は, 空間認識率が悪くなりやすい。

出力デバイスにはヘッドホンを利用している。ヘッドホンには密閉型, 開放型, イヤホン型, 骨伝導型など様々なものがある。視覚障害者は, 周りの状況把握の際に大きく依存している聴覚の入力部である耳を塞ぐものを使用することには大きな抵抗がある。耳を塞がれることで, 音の微妙な変化を捕らえられないなど, 身につけている聴覚が損なわれ, 不安を感じるからである。しかしながら, 若い世代の視覚障害者は, 日常生活でヘッドホンステレオの使用に慣れているため, このような不安感を生じない場合も多いが, 基本的には図 6-2(c)のような骨伝導タイプのヘッドホンを使用する。これは外耳を塞ぐことがないため, 上記の欠点を和らげることができる。

6.2.3. 3次元視覚情報処理

カメラから得られたステレオ画像の解析により対象シーンの計測・検出を行う。詳しくは第 3 章で説明したとおりで, 処理手順は図 6-4 に示すように相関法とセグメントベースドステレオ法を組み合わせで行っている。これらの2つの手法の特徴の組み合わせにより面情報を含む構造化された 3次元データが得られる。

3次元視覚情報が得られた後、物体の幾何モデルをあらかじめデータベースに用意しておく幾何モデルベースな物体認識処理が行われる [48]。ユーザが設定したタスクに必要な対象物のモデルと計測データ間で座標変換を求めることで、対象シーンにおける物体の有無、位置、姿勢、形状を知ることができる。また、段差や障害物に関しては、距離情報からその位置と状況を検出することが可能である。これらの処理により得られた結果が3次元音響システムに送信される。

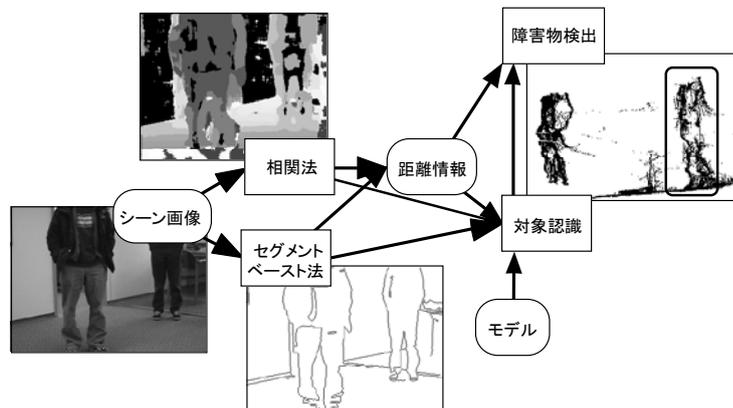


図 6-4 視覚情報処理手順

6.2.4. システム制御

システム全体の制御を図 6-3 を用いて説明する。3次元環境入力部であるステレオカメラから得られた画像はコンピュータに送られ、上記の処理を行うことで対象シーンの3次元構造を復元し、ユーザがあらかじめ与えたタスクで必要な対象を認識、追跡する。これにより得られる対象物体の3次元位置、姿勢と、対象に割り振った音をサウンドスペースプロセッサに入力することで、対象に対する音像を3次元仮想空間にマッピングできる。サンプラーとサウンドスペースプロセッサはコンピュータからのMIDI信号で制御している。複数台の仮想音や環境音(実験時にはMDプレーヤーから出力)はミキサーで合成され、ユーザのヘッドホンに出力される。現段階では、システムに組み込まれてはいないが、ヘッドセットにつけられたマイクで、ユーザは音声指示によりタスクの設定を行う。

6.2.5. 視覚情報の音響表現

ビジョンシステムにより得られた3次元の視覚情報を3次元仮想音響として出力する。ただし、得られる情報すべてを聴覚情報に変換せず、ユーザが行いたいタスクに必要な対象に関してのみ音響表現を行う。1つのタスクに関しては検出・認識対象はせいぜい2,3個にとどめる。これはビジョンシステムの処理能力の問題と、多くの仮想音を加えることで得られる情報量は増大するが、逆にユーザに混乱を招き、認識率が低下するためである。しかしながら、階段、壁、車などの障害物、危険物に対してはタスクとして設定されていなくても、危険な対象であるとして、注意を喚起する音、もしくは音声を出力する。

本システムで使用している6.2.2のサウンドスペースプロセッサは、基本的には認識した物体の位置、動き、および、音源を与えることで、3次元空間上での仮想音源を生成できる。対象が現実世

界で音を発するものである場合は同じ音を用い、それ以外の対象に関しては、ユーザが認識しやすい音を割り当てる。例えば、対象を信号機とした場合、視覚障害者用信号機で使われている音楽を用い、転がるボールなどを対象にした場合は鈴の音などを利用し、できる限り実環境と同じ音響空間を作る。

また、物体の認識には、位置情報だけでなく、姿勢、大きさ、形状、色、表面模様などの属性も有用である。これらを一度に音として表現するのは難しいため、認識対象を定めた後にこれらの属性を個別に聞くモードを設定する。それぞれの属性に関して出力する音は、ユーザの感性に左右されるものであり、一概に決めることは難しいが、例えば、対象の大きさに関しては、大きさに比例するように再生する時間を変化させる方法で表現する。これらの設定は、最初のチューニング段階でユーザの好みに合ったものに変更できる。

6.3. 音像定位実験

3次元音響インタフェースの開発にあたり、基礎実験として3次元音響システムを用いた仮想空間における3次元音響の音像定位実験を行った。予備実験として、視覚障害者に音像定位理解度、ヘッドホンの種類(図6-5(a))による認識への影響に関する実験を行い、この結果を踏まえ、実験1として、音像の位置を被験者の左右の耳を含む水平面に限定した前後左右の定位実験(図6-5(b))を、実験2では、仮想空間の全方位における音像定位実験(図6-5(c))を行った。そして、実験3では、骨伝導タイプのヘッドホン(図6-2(c))を用いて、実験2と同じ実験を環境音のあるなしで行った。今回の実験ではジョイスティックを用いて、ユーザの頭の動きをシミュレートした。

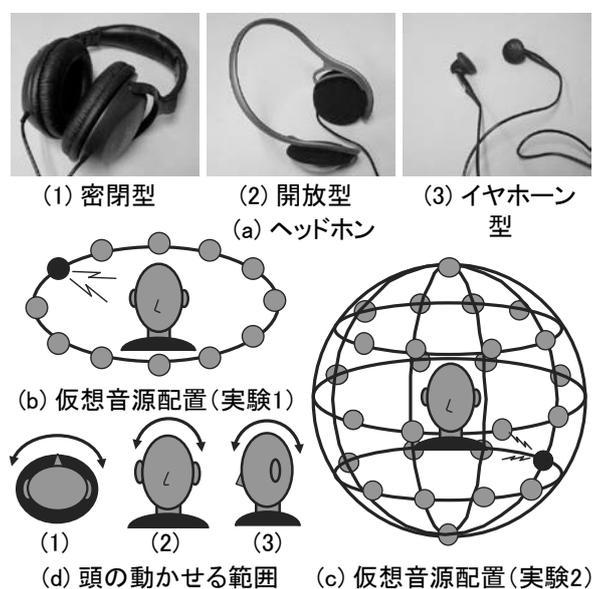


図 6-5 実験

6.3.1. 予備実験

予備実験として、3次元音響システムで生成された、いくつかの仮想音響空間の認識度、また、

使用するヘッドホンの種類による認識度の違いを調べた。被験者は後天盲の全盲者、21歳の男子大学生である。最初に仮想空間を移動する音源をいくつか聴き、その軌跡を口頭で答えてもらった。同様の実験をヘッドホンの種類を替えて行い、種類による認識への影響を調べた。

軌跡認識に関しては、音源としてシンセサイザーで作成したジェット機の音と、ヘリコプターの音を用い、頭上を前方から後方へ通過する音源、上空を旋回する音源、前方を上下する音源を作成し、実験を行った。実験後の被験者の意見として、左右感は全く問題ないが、前後感、特に上下感にはわかりづらい点があげられた。たとえジェット機の通過音であっても、つまり、飛行機が自分の下を飛ぶという状況がないことが分かっているにもかかわらず、上下の判断に戸惑いが見られた。

また、使用するヘッドホンを図 6-5 (a)のような(1)密閉型、(2)開放型、(3)イヤホン型で行った認識具合の差は、(3)、(2)、(1)の順で分かりやすかったという意見であった。(1)は音の聞えは一番良いが、やはり環境音を遮蔽してしまうことが最大の欠点だった。(2),(3)はさほど変わらない使用感とのものであったが、(3)が普段使い慣れているものと同一タイプであるため、一番認識しやすいものとして選ばれた。今回の一連の実験では、準備の都合で骨伝導タイプのヘッドホン(図 6-2(c))に関する実験は行っていない。

6.3.2. 実験1:水平面内の音像定位

予備実験から、上下方向の音像定位は経験に依存しやすい傾向であることが判明していたので、実験1ではまず音像の位置を被験者の左右の耳を含む水平面に限定し、前後左右の定位実験を行った。被験者は視覚障害をもつ3名の男性で、うち2名は日常的に点字を使用している全盲者である。

実験方法は、図 6-5 (b)に示すように、被験者を中心とする半径 1.5m の同一円周上に 30 度間隔で 12 点の音源を配置し、ランダムな順番で計 24 回音を提示した。被験者にはあらかじめ 12 方向からのランダムな音像提示実験であることを伝えておき、認識した音像の方向を時計の針の方向で回答してもらった。再生音として音像定位がしやすいピンクノイズを用いた。また、音響出力装置として開放型のヘッドホンを使用した。

実験内容は、前後左右の音像の受動的／能動的な方法による位置認識に関して行った。受動的な認識実験として、単純に音を鳴らすもの(実験 a)、音源を左右に $\pm 10^\circ$ ずつ 2 秒周期で揺らすもの(実験 b)を、能動的な実験として、図 6-5(d)-(1)のように頭を左右に ± 10 度まで動かせるもの(実験 c)を行った。

結果を表 6-1、図 6-6 に示す。各実験における正答率、認識時間は表 6-1 に示されている。図 6-6 は各実験における音像提示方向と回答方向の関係をプロットしたものである。点が実線の円周上に存在すれば正解、内(外)側にずれるほど左(右)方向へのずれが大きいことを示している。点線上ではそれぞれ左(右)に6時分(180°)ずれていることになる。正答率は平均すると実験 c, b, a の順で良くなっており、認識時間は実験 a が最も短く、実験 b, c はほぼ同じで実験 a の約 1.8 倍である。実験 a, b では前後方向の誤認識が見られる。

表 6-1 実験結果1:正答率と認識時間

	正答率(%)			認識時間(秒)		
	実験a	実験b	実験c	実験a	実験b	実験c
MA	58.3	83.3	95.8	3.1	6.5	5.6
TM	54.2	37.5	62.5	3.6	5.9	7.6
EN	66.7	79.2	75.0	2.6	4.5	4.1
平均	59.7	66.7	77.8	3.1	5.6	5.8

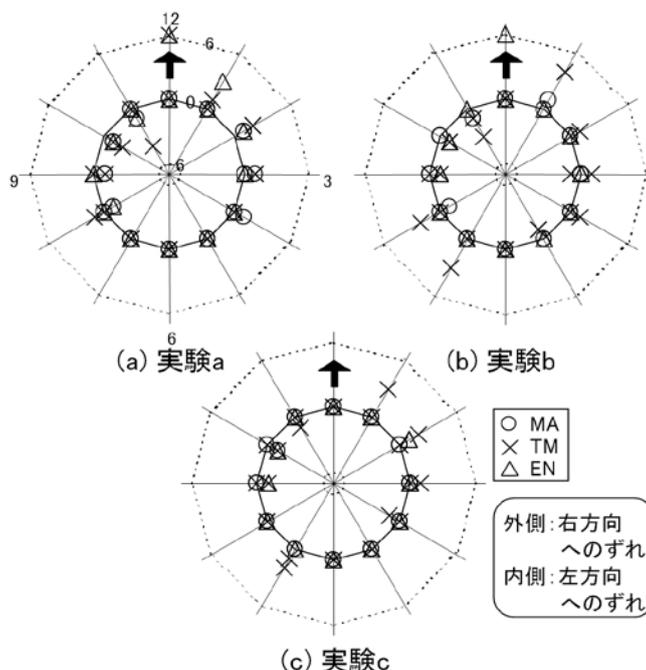


図 6-6 実験結果1:回答方向

6.3.3. 実験2:全方位の音像定位

本実験では全方位の音像定位に関して行った。被験者は大学院生6名で、目隠しをした晴眼者である。

この実験では、図 6-5(c)のように、被験者を中心とする半径 3.0 m の同一球上(床面は中心から 3.2 m に設定)に上下各 1 点、水平面、上下 45 度の水平断面それぞれに 45 度間隔で各 8 点の計 26 点配置し、実験 1 と同様にランダムな順番で計 52 回音を提示した。各点に番号を割り当て、被験者が事前に方向と番号の対応を学習した後に実験を行い、定位した音像の番号を口頭で回答してもらった。再生音、音響出力装置は実験 1 と同じものを使用した。

実験内容は、音像の受動的/能動的な方法による音像定位であり、受動的な実験 a として単純に音を鳴らすものと、能動的な実験 b として頭を $\pm 10^\circ$ まで動かせるもの(図 6-5(d)-(1,2,3))を行った。

実験 a, b での正答率、認識時間の結果を、それぞれ表 6-2, 表 6-3 に示す。この表の正答率 1~4 の意味は以下のとおりである。

正答率 1: 回答位置が完全に一致した場合の正答率

正答率 2: 回答が左右に一つずれていても正解とした場合

正答率 3: 回答が上下に一つずれていても正解とした場合

正答率 4: 正答率 2 と 3 の基準をともに認める場合

一つの音源を定位する認識時間に関しては、実験 a が直感的に回答するため、平均 9.6 秒であったが、実験 b では音源を探すという動作が入るため約 2.5 倍の 25.3 秒平均でかかっている。図 6-7 は、正答の判断基準を変えたときの実験 a, b における平均正答率をプロットしたものである。完全一致の正答率 1 で見てみると、実験 a では 26.0%と低い正解率であった。これに対し、能動的な行動ができる実験 b では 52.6%となり、約 2 倍の正答率となっている。しかしながら、それでも 2 回に 1 回という正解率である。これを隣の音源を答えても正解とする正答率 4 で見てみると、それぞれ、64.4%, 82.7%となり、能動的な音像定位では大雑把な方向の認識はできていたと言える。

表 6-2 実験結果 2a:受動的な認識実験の正答率と認識時間

被験者	正答率1	正答率2	正答率3	正答率4	認識時間
KD	21.2%	32.7%	42.3%	53.9%	10.5秒
KM	23.1%	28.9%	57.7%	63.5%	3.3秒
MY	30.8%	42.3%	63.5%	75.0%	12.4秒
TN	25.0%	42.3%	50.0%	67.3%	10.0秒
WN	17.3%	42.3%	30.8%	55.8%	9.3秒
WA	38.5%	44.2%	65.4%	71.2%	12.1秒
平均	26.0%	38.8%	51.6%	64.4%	9.6秒

表 6-3 実験結果 2b:能動的な認識実験の正答率と認識時間

被験者	正答率1	正答率2	正答率3	正答率4	認識時間
KD	46.2%	53.9%	69.2%	76.9%	48.0秒
KM	38.5%	50.0%	67.3%	78.9%	19.5秒
MY	63.5%	67.3%	88.5%	92.3%	16.2秒
TN	57.7%	69.2%	76.9%	88.5%	19.7秒
WN	44.2%	50.0%	61.5%	67.3%	24.3秒
WA	65.4%	73.1%	84.6%	92.3%	24.1秒
平均	52.6%	60.6%	74.7%	82.7%	25.3秒

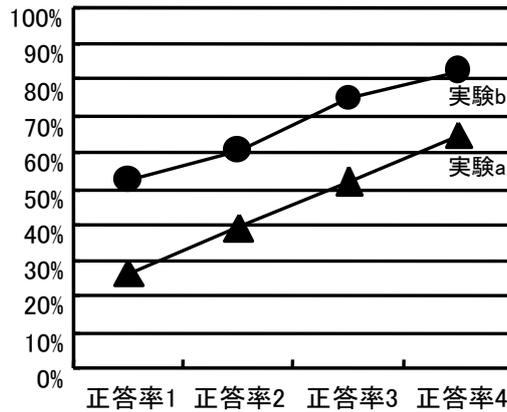


図 6-7 実験結果 2:平均正答率

6.3.4. 実験3:環境音の影響

実験 3 では、より実環境下での使用を考慮して、骨伝導型のヘッドホンを使用して、環境音が全方位の音像定位に与える影響を調べた。環境音として、図 6-8 の交通の激しい交差点付近でバイノーラル録音したものをを用い、定位する音としては、視覚障害者用信号機から発せられる「ピッポ」という音に似せたものを使用した。

実験の基本設定は実験 2b の受動的な認識実験と同じであり、被験者は晴眼の男性 2 名と、女性 1 名である。実験 a としては環境音がない場合、実験 b では環境音がある場合を行い、正答率と認識率を調べた。環境音はイヤホンタイプのヘッドホンを使用して呈示した。表 6-4、表 6-5 に実験 a, b の結果を示す。

完全一致の正答率 1 に関して見れば、ともかなり低い値である。実験 2 より低い値になった理由としては、骨伝導ヘッドホンを使用したことが挙げられる。80dB 以上となる有効周波数帯域が 30 ~ 3000Hz であり、開放型のヘッドホンよりも聞き取りにくいためである。また、別の原因としては、音源から発せられる音が視覚障害者用信号機の音に似せた単純な音を用いたためである。実験 1, 2 で使用したピンクノイズに比べ、周波数の広がりがなく、音像定位には適していなかったかもしれない。実験 a, b の比較で、環境音が音像定位に与える影響は、正答率が低下し、認識時間が増加することから、少なからずあると言える。



図 6-8 交差点での環境音

表 6-4 実験結果 3a:環境音がない場合の正答率と認識時間

被験者	正答率1	正答率4	認識時間
YK	11.5%	48.1%	2.75 秒
TY	19.2%	53.8%	3.67 秒
MT	11.5%	38.5%	8.43 秒
平均	14.1%	46.8%	4.95 秒

表 6-5 実験結果 3b:環境音がある場合の正答率と認識時間

被験者	正答率1	正答率4	認識時間
YK	13.5%	44.2%	2.9 秒
TY	5.8%	34.6%	4.6 秒
MT	9.6%	36.5%	9.5 秒
平均	9.6%	38.4%	5.7 秒

6.4. 考察

ここでは、開発した音響提示システムを利用した音像定位の被験者実験に関する考察を行う。予備実験では、1名の全盲者に対して3次元仮想空間での音像定位基礎実験を行ったが、上下、および、前後の認識は芳しくなかった。この理由としては、被験者とシステムの HRTFs があまり一致していなかったのではないかと考えられる。また、ヘッドホンの種類に関しては、日常生活でイヤホン型のものを使用しているため、これがベストなものに選ばれた。ヘッドホンをして仮想音源を鳴しながら廊下などを歩行する実験も行ったが、この被験者はあまり違和感なく行動できた。大半の視覚障害者が環境音を直接聴けない状況は受け入れがたいとしているが、ヘッドホンを日常生活で良く使っている若い世代の視覚障害者(被験者の実験時の年齢は 21 歳、18 歳の段階で失明。)については、必ずしもこれが正しいとは言い切れないかもしれない。

実験 1 の正答率については、被験者 TM を除き、完全受動的提示の実験 a より受動的提示の実験 b、能動的提示の実験 c の方が良いが、個人差や実験順序による学習などを考えると提示方法による有意な差は見られない。認識時間に関しては、実験 a がどの被験者も短い、この実験においては被験者が直感的に回答しやすいためだと考えられる。実験 b では、静止音源より、動いている音源の方が方向を同定しやすいため、音源を揺らすことで実験 a に比べ認識率がどう変化するかを調べた。結果として、2名の被験者で認識率の向上が見られたが、認識時間は実験 c とあまり変わらず、直感的に回答はできなかったと言え、あまり適した提示方法ではなかった。図 6-6 にある提示方向に対する回答方向の誤りを詳しく見ると、正面提示に対し背面方向と回答するなど、前後方向で誤認識する場合は実験 a、b である。これは、受動的提示は能動的提示よりも音像定位がしづらいことを示している。実験 c の図 6-5(d)-(1)のように自分の首を動かす動作の場合、音像を

能動的に動かせるため、正面と背面の誤認識はなくなっている。より正確に音像定位するためには能動的な動作が必要であると言える。

実験 2 の全方位音像定位認識に関する実験でも受動的な音像定位より、能動的な行動を伴う実験 b の方の結果が良かったが、完全正答(正答率 1)は平均 52.6%と約半分であり、かなり低い認識率である。しかしながら、現実の世界を考えると、実験で設定したような分解能は、必ずしも必要ではない時が多いので、隣の位置を答えても正解とする正答率 4 の基準で見ると 82.7%となり、かなりの高い認識率であった。一方、実験 b では認識時間は平均で 25.3 秒とかなりかかっており、状況を直感的に判断しなければならない場合には問題があると言え、実験 a における正答率をより高める必要がある。正答率 4 の基準でも 64.4%であり、音の出し方や音そのものに何らかの工夫を加える必要があると言える。

また、実験 2 において、6 名の被験者のうち平均的な結果である TN に関して正答率 4 の基準での誤回答を分析してみると、表 6-6 のようになった。実験 a では前後、上下の誤認識とも同じ程度であるが、能動的な探索が行われる実験 b では、上下の誤認識はなくなるが、前後の誤認識は多い。上下・前後方向の音像定位に関しては、上下・前後に非対称な形である頭部を回折して耳に届くことによる、音の周波数の微妙な違いで判別しており [74]、使用しているサウンドスペースシステムの HRTFs でもパラメータに反映されているが、より強調するように音源の周波数を変える必要がある。

同様に、実験 3 における被験者 YK の誤回答の内容分析を表 6-7 に示す。環境音あるなしにかかわらず、ほぼ同じ正答率であったが、実験 b の環境音がある場合は前後の誤認識率がかなり高くなっている。前後の認識に関しては周波数に依存することが知られており [75]、環境音の周波数が前後の認識に影響していると仮定できる。前後の認識率を改善するには、周波数の違いを強調した音を音源から出力すべきであると言える。

また、実験全体を通して見ると、システムとユーザの HRTFs の同一性の問題があるが、最初のうちは音像定位があまりできない人でも、使い慣れてくると同じ内容の実験において正答率がかなり向上する。ただし、被験者によってその慣れるまでの時間は一定ではなく、システムを使いこなすための効率的な学習方法を考える必要がある。

表 6-6 実験 2 における被験者 TN の誤回答の内訳

被験者 TN	誤回答率	内訳		
		前後	上下	その他
実験a	32.7%	41.2%	35.3%	23.5%
実験b	11.5%	50.0%	0.0%	50.0%

表 6-7 実験 3 における被験者 YK の誤回答の内訳

被験者 YK	誤回答率	内訳		
		前後	上下	その他
実験a	51.9%	14.8%	59.3%	33.3%
実験b	55.8%	41.1%	51.7%	20.7%

6.5. まとめ

本章では、視覚障害者に、触覚では伝えにくい、状況が刻々と変化し、空間的な広がりを持つ 3 次元の視覚情報を、3 次元音響による 3 次元空間情報提示支援システムの実現を目指して行った、支援システムの開発と、3 次元音像定位に関する被験者実験の結果を述べた。構築したシステムを使用して行った音像定位の基礎的な実験結果から、音像定位は上下方向、前後方向で誤認識を起こしやすい点、頭を動かすなどの能動的な動作が有効である点などが判明し、音響による適切な情報提示に関する指針が得られた。また、周辺環境からの音を塞ぐことなく、仮想音を付け足す形で音響情報を提供する方法のため、ユーザは使用に関して多大な学習をすることなく、今まで養ってきた聴感覚を損なわずに使用できる可能性も示した。

今後はこれらの結果を踏まえ、単に音と位置情報を入力するだけではなく、誤認識しやすい特定の位置(方向)、移動方向に対して音源に変化を加えることで、音像定位率を向上させる表現方法の開発や、音の種類による認識率の変化、複数音像の定位、仮想音の環境音把握への影響に関する実験、位置情報だけではなく、物体の属性を音響で表示し、認識する実験、そして、音声指示によるタスク設定などのユーザインタフェースの開発を行い、視覚障害者の方々の利用を通して、問題点を一つずつ解決していく必要がある。

7. 結論

最後に、本研究で行った視覚障害者向けの視覚情報提示支援システムの研究開発における要点を述べるとともに、視覚障害者用支援システムに関する今後の研究の展開と課題を述べる。

7.1. まとめ

著者は、視覚障害者の人とのコミュニケーションや行動を支援できるように、視覚情報処理技術を活用した、視覚障害者のための視覚情報提示支援システムに関する研究を行った。工学的アプローチで、失われた視覚情報を他の感覚を通じて視覚障害者に提示するために、対象・環境の視覚情報を得る手段・手法の研究開発、得られた情報を効率的に他感覚情報に変換して視覚障害者に伝える手段・手法の研究開発を行った。以下に、本研究の要点をまとめる。

第1章では、これまでの視覚障害者向けの情報取得・理解支援システムに関する背景、および、これまでの機器開発の推移に関して説明し、本研究開発で行った、効率的な情報提示システムの実現を目指し、視覚情報処理技術と触覚・聴覚表現技術を組合せた新しい情報提示システムなど、本研究の新規性、特徴を示した。

第2章では、視覚障害者が図面情報を読み取る支援システムに関して述べた。図面情報をそのまま触図として提示するのではなく、コンピュータで図面を認識させて情報を変換し、触覚と音声を用いて少ない情報量で、2次元視覚情報を的確に提示する、基本となる提示支援システムを提案した。タッチセンサの機能も有する触覚ディスプレイを開発し、システムとの対話的な操作により視覚障害者が図面情報を理解する際にかかる負担を軽減できることを示した。また実験により、触知覚の最適な表示方法は個人差があり、ユーザごとのカスタマイズが可能なシステムでなければならないこと、事前に図面情報を与えることで認識しやすくなること、触覚で伝える情報は元図から簡略化しすぎないこと、音声情報は簡潔なものほど良いこと、などが判明した。

第3章では、この基本システムの研究開発を通じ、視覚障害者が必要としている情報の中でも生活上重要だと考えられる3次元空間情報の取得を行う必要性が判明したため、実環境の3次元の視覚情報を得るための方法として、ステレオ画像解析による3次元形状復元法を開発した。ステレオ画像間の対応探索において、領域の境界を構成する境界線セグメントの連結性に基づいて対応を評価する方法であり、オクルージョンに対して比較的頑健な方法である。また、境界線の類似性に基づく大局的な画像間の対応評価を行うことで、安定した復元結果を得られる。復元された幾何データは単なる3次元形状データではなく、3次元物体の認識や追跡に利用可能な構造を有しているため、応用範囲は広く、物体検出、追跡の際の基盤となる。3台のカメラで構成されるステレオカメラを用いて、複数の物体が存在する様々なシーンを対象にして実験を行い、本手法により、精度よく対象の3次元形状が得られることを示した。この手法により得られる3次元形状データは、3次元物体の認識や追跡に利用可能な構造を有しているため、視覚障害者が認識したい物体を検出する際などに利用できる。

第4章では、第3章に続いて3次元視覚情報処理技術として、取得すべきシーン情報が複雑な

場合、一方向からの計測では十分なデータを取得できず、視覚障害者に必要十分な視覚情報を提示することができない場合があるため、多視点から得られた3次元視覚情報を統合する方法も開発した。視点方向の位置関係が未知である場合でも、計測データから、オクルージョンの影響が少ない領域を基に初期照合を行い、共通部分全体のずれが最小となるように繰り返し処理をすることで、より正確な位置関係を計算することが可能な方法である。本手法を、複雑な形状を有する花などの対象に対して適用し、多視点から計測したデータを統合し、全体の3次元形状を正確に再現できることを実験により示した。この手法により、視覚障害者が必要とする対象の十分な3次元形状情報を得ることができる。

第5章では、第2章の基本システムの研究を発展させ、第3章、第4章で開発した3次元視覚情報取得・処理技術を適用し、視覚障害者への提示情報量を増やして、必要な3次元情報をより得やすいように、対話型触覚ディスプレイによる3次元情報提示システムを開発した。前述の基本システムの研究を発展させ、視覚情報として3次元の形状を有する物体を対象とし、その視覚情報を3次元の触覚情報と音声で提示するシステムである。16×16のマトリクス状に配置され、6段階の高さ表現ができる触覚ピンを有し、かつ、タッチセンサの機能を有している3次元触覚ディスプレイを開発した。3次元視覚情報処理技術により取得した物体の立体形状、配置情報を3次元触覚ディスプレイに表示し、音声で属性情報を付加することで、視覚障害者がシステムと対話的なやりとりをすることが可能となり、3次元情報を効率よく取得できる点、空間認知をしやすくなる点を、実験を通じて示した。

第6章では、触覚では伝えにくい、状況が刻々と変化し、空間的な広がりを持つ3次元の視覚情報を、即応性のある聴覚を利用して提示する支援システムとして、3次元の仮想音響技術を利用した視覚情報提示システムを構築した。3次元センサにより取得した3次元視覚情報に音響を割り当て、仮想3次元空間でその位置・動きを表現する方法で、インフラが未整備の場所や情報提供できない対象に対して有用であり、ユーザの周りの状況を提示できる。このうち、3次元仮想空間での音像定位の特性を評価するために、各種被験者実験を行った。その結果、音像定位は上下方向、前後方向で誤認識を起こしやすい点、頭を動かすなどの能動的な動作が有効である点などが判明し、音響による適切な情報提示に関する指針が得られた。また、周辺環境からの音を塞ぐことなく、仮想音を付け足す形で音響情報を提供する方法のため、ユーザは使用に関して多大な学習をすることなく、今まで養ってきた聴覚を損なわずに使用できる可能性も示した。

以上のように、視覚情報を取得する技術、視覚情報を提示する技術を確立し、触覚、聴覚の特性を有効に活用して、視覚障害者に視覚情報を提示する支援システムを構築した。

7.2. 今後の展開

本研究で行った視覚障害者用の支援システムは、デバイスの小型化、高機能化、低価格化などで研究開発当時に比べより良いものが開発され、その一部は製品として入手できるようになっている。しかし、ユーザの求めるものは千差万別であり、ユーザへの適用性、つまり、ユーザに応じたカスタマイズが容易に行えることが必要である。これまでに数多くの視覚障害者の方と接してきたが、

この機能があれば、すべての方に満足していただけるというものは難しい。そのため、支援システムの開発方針としては、過去の研究、ユーザの生の声を最大限反映した最大公約数の機能、ユーザインタフェースを有するシステムを作成し、それを基にユーザの希望する機能を追加できるような構成にすることが、今後の支援機器開発において最重要視されるべきことであると思われる。

また、支援システムという観点では、視覚障害者向けに限らず、お年寄りや怪我をした人、肢体不自由者を支援するシステムの一つとして、ロボットを活用した生活支援ロボットという新たな枠組みができてきている。ロボット技術の進展は目覚ましく、また、実空間に作用を及ぼすシステムであり、ユーザが必要な物を取ってきてくれるシステムと考えれば、非常に役立つ支援システムの一つであると言える。このように、支援システムの一つとして有望なものとして捉え、開発した視覚情報取得技術をロボット、例えば、川田工業製 ヒューマノイドロボット HRP-2 や、安川電機製 サービスロボット SmartPal V などに搭載し、頼まれた作業をこなす支援システムの実現に向けた研究開発も行った。

具体的には、産総研の所内プロジェクト、ヒューマノイドロボット型知能ブースタープラットフォーム開発(2003-2005 年度)(図 7-1)、NEDO の次世代ロボット知能化技術開発プロジェクト(2008-2011 年度、2007 年度は経産省プロジェクト)(図 7-2)において、生活環境内での人とロボットが共存し、人に指示された簡易な作業(物を取ってくる)をロボットに代行してもらうために必要なセンサ・認識技術、動作技術、操作技術、システム化技術に関する研究開発を行い、それらを組み合わせることでロボットによる支援作業を実現した。

現段階では、まだまだロボットができる作業が限られており、また、専門家でないと操作できない、価格が極めて高い、安全性の問題など、一般に普及するにはまだ多大な課題を有しているが、先の支援機器と同じように実際のシステムで試用し、問題点を洗い出す必要がある。今後、各機能の研究開発、ユーザ評価に基づくシステムの進展が望まれる。しかしながら、製品として販売する際に必要となる安全性保証に関しては、NEDO の生活支援プロジェクト(2009-2013 年度)において、生活支援ロボットの実用化に必要な安全性を考慮したロボットの設計、評価、認証を行うための技術として、高信頼ソフトウェアツール、リスクアセスメント手法などの開発と国際標準化が行われ、これらの成果に基づいて、ISO/DIS 13482 に基づく世界初の製品認証が実現しており、一歩ずつではあるが、支援システムとしての一翼を担えるようになりつつある。

また、視覚障害者に欠落している視覚情報を取得する方法として、コンピュータビジョン技術による3次元視覚情報処理技術の研究開発を第3章、第4章で行ったが、コンピュータやセンサ類が進歩した現時点においても、まだまだ視覚障害者が日常生活を安心してこなせるだけの情報を取得、処理できる状態までには至っていない。ハードウェアの進歩は必要不可欠であるが、システムの知能化、つまり、ソフトウェア側の進歩も不可欠であり、今後もシーン理解などに関する研究開発を行っていきたい。

以上のように、本研究で行った視覚障害者用の、3次元視覚情報を取得し、提示する支援システムに関する技術は、視覚障害者にとどまらず、身体障害者やお年寄りなどを支援するためにも必

要不可欠で、共通な技術であり、対象を拓げることで更なる技術進歩が進み、そして、様々な分野への展開が期待される。人を支援システムというものには、支援する内容が千差万別であり、明確なゴールがあるわけではなく、各要望を満たすような技術開発を常に行っていく必要がある。今後も、工学的な技術アプローチで、人々の役に立つ技術の研究開発を行うとともに、その技術を社会に還元する努力を続けていきたい。



図 7-1 ヒューマノイドによる作業代行



図 7-2 SmartPal による作業代行

謝辞

本研究は、名古屋大学工学部・工学研究科、通商産業省工業技術院電子技術総合研究所(電総研)、および、組織改編後の独立行政法人産業技術総合研究所(産総研)において行った、視覚障害者に対する情報支援システム、3次元情報取得・処理に関する研究をまとめたものである。

本研究をスタートした時から、本論文をまとめるまで長年に渡り、丁寧かつ熱心にご助言、ご指導して頂いた名古屋大学大学院情報科学科 大西昇教授に深く感謝の意を表します。

本論文の内容をご検討くださり、多くのご助言を頂きました名古屋大学大学院情報科学科 村瀬教授と間瀬教授に深く感謝いたします。

この分野の研究開発に係わるきっかけを作っていただき、電総研への入所に際しても大変お世話になった、高校・大学・電総研の大先輩でもある、杉江昇 名古屋大学名誉教授には、言葉では言い表せないほど深く感謝いたします。

名古屋大学時代において、多くのコメント、アドバイスを頂いた杉江・大西研究室の黄捷様(現在、会津大学准教授)、加藤克己様、箱守聡様、本田英明様、森本康嗣様、吉武敏幸様、荒木均様、高木和久様、山村剛様(現在、愛知県立大学准教授)、彦坂達也様、荻久保友史、長谷部雅彦様、間瀬久雄様、および、松崎規子様に感謝いたします。

電総研時代においては、研究する機会を与えていただき、ご指導頂いた、高瀬國克 知能システム部部长、諏訪基情報科学部部长、大島正毅室長、富田文明室長、吉見隆主任研究官、植芝俊夫主任研究官、角保志主任研究官をはじめとする視覚情報研究室の皆様には、研究開発に関する日々の助言、共同作業を行えたことに深く感謝します。また、共同研究を通じて一緒に研究開発を行い、多くのアドバイスを頂いた元三洋電気の高橋裕信様(現在、アプライド・ビジョン・システムズ代表取締役)、杉本和英様(現在、沖縄高専教授)、スタンレー電気の石山豊様、筑波技術短期大学(現在、筑波技術大学)の宮川弘之教授、小林真博士(現在、同大准教授)、皆川洋樹博士(現在、同大教授)、および、被験者実験を快く快諾し、貴重な意見をいただけた筑波技術短期大学視覚部の学生の皆様に深く感謝します。

産総研においては、研究する機会を与えていただき、ご指導頂いた、知能システム研究部門の谷江和雄元研究部門長(故人)、平井成興元研究部門長(現在、千葉工業大学未来ロボット技術研究センター 副所長)、電総研時代から引き続いて一緒に研究開発を行っていただいた上記の皆様、丸山健一元主任研究員(現在、セイコーエプソン株式会社)、高瀬竜一様、川端聡元特別研究員に感謝いたします。特に、本論文をまとめるに当たり、色々考慮して下さった、比留川博久研究部門長、横井一久研究副部門長、大場光太郎研究副部門長、部門事務室の皆様には感謝いたします。

このように、多くの皆様の協力の下で、研究開発を行うことができたことに改めて感謝致します。

最後に、論文としてまとめることができたのは、家族の日々の協力と支援があった他ならならず、妻と子供たち、両親に感謝の気持ちを伝えたい。

参考文献

- [1] 市川宏, 大頭仁, 鳥居修晃, 和氣典二, 視覚障害とその代行技術, 名古屋大学出版会, 1984.
- [2] D. H. Warren, E. R. Strelow, “Electronic Spatial Sensing for the Blind,” Martinus Nijhoff Publishers, 1985, pp. 35-61.
- [3] “第 21 回感覚代行シンポジウム予稿集,” 感覚代行研究会, 1995.
- [4] “第 22 回感覚代行シンポジウム予稿集,” 感覚代行研究会, 1996.
- [5] “第 23 回感覚代行シンポジウム予稿集,” 感覚代行研究会, 1997.
- [6] 末田統, “盲人用文字・図面認識補助装置の試作 II,” 信学技報, MBE75-53, 1975.
- [7] L. Kay, “Electronic aids for blind persons interdisciplinary subject,” Proc. of IEE, 131, Pt. A, 7, pp.559-576, 1984.
- [8] 伊福部達, 音の福祉工学 音響テクノロジーシリーズ 3, コロナ社, 1997.
- [9] P. B. L. Meijer, “An Experimental System or Auditory Image Representations,” IEEE Trans. Biomed. Eng., 39, 2, pp.112-121, 1992.
- [10] 小林真, 太田道男, “音像定位を利用した能動的情報取得が可能な視覚代行装置,” バイオメカニズム誌, 21, 1, pp.39-42, 1997.
- [11] M. Shimizu, K. Itoh, Y. Yonezawa, “Operation Helping Function of the GUI for the Visually Disabled Using Virtual Sound Screen,” Proc. of ICCHP'98, pp.387-394, 1998.
- [12] Y. Shimizu, M. Shinohara, T. Sasaki, T. Watanabe, H. Minagawa, H. Nagaoka, “An Assisting System for a GUI-Based PC for the Blind,” Proc. of ICCHP'98, pp.373-378, 1998.
- [13] 在塚俊之, 畑岡信夫, “視覚障害者の GUI アクセスを助ける音響技術,” 音響学会誌, 54,5, pp.393-398, 1998.
- [14] 松阪治男, “視覚障害者のタッチスクリーン端末の利用とユーザインタフェースに関する研究,” 筑波技術大学, 修士学位論文, 2013.
- [15] J. M. Loomis, R. G. Golledge, R. L. Klatzky, J. M. Speige, J. Tietz, “Personal guidance system for the visually impaired,” Proc. of ASSETS'94, pp.85-91, 1994.
- [16] “第 24 回感覚代行シンポジウム予稿集,” 感覚代行研究会, 1998.
- [17] “第 25 回感覚代行シンポジウム予稿集,” 感覚代行研究会, 1999.
- [18] 松原広, 後藤浩一, 明星秀一, “視覚障害者向け誘導案内システムの開発,” 鉄道総研報告, 13, 1, pp.31-36, 1999.

- [19] 日経産業新聞, “視覚障害者向け音声案内システム,” p. 1, 1998.
- [20] 明星秀一, 松原広, 深澤紀子, 後藤浩一, “視覚障害者向け誘導案内システム,” 第57回情処学全大, 4, 1S-7, 1998.
- [21] 和気洋美, 和気典三, “指頭による触文字知覚,” 第6回感覚代行シンポジウム論文集, pp.63-67, 1980.
- [22] 和気洋美, 和気典二, “一筆書きそードによる触図形知覚,” 神奈川大学心理・教育研究論集, 第5号, pp.66-78, 1987.
- [23] 鈴木章二, 石田明允, “視覚障害者用ワードプロセッサシステム,” 東京医科歯科大学・医器材研報, 21, pp.51-57, 1987.
- [24] H. Wake, “On Tactile and Visual System for Visually Handicapped Individuals,” Kanagawa University, The Journal of Psychology & Education, 4, pp.47-60, 1986.
- [25] S. F. Frisker-Gibson, P. B. Y-Rita, W. J. Tompkins, J. G. Webster, “A 64-soenoid four-level fingertip search display for the blind,” IEEE Trans. Biomed. Eng., BME-34, 12, pp.963-965, 1987.
- [26] R. A. Jarvis, “A perspective on range finding techniques for computer vision,” IEEE Trans. Pattern Anal. & Mach. Intell., PAMI-5, 2, pp.122-139, 1983.
- [27] M. J. Magee, J. K. Aggarwal, “Using multisensory images to derive the structure of three-dimensional objects - A review,” CVGIP, 32, 2, pp.145-157, 1985.
- [28] 山本裕之, 田宗道弘, 田村秀行, “距離画像の入力と処理,” 信学技報, PRU86-129, 1987.
- [29] 井口征士, “1986 三次元計測研究最近の動向と展望,” 映像情報, 18, 11, 1986.
- [30] 吉見隆, 富田文明, “3次元センシングの研究動向,” 情処技報, CG-90-46-2, 1990.
- [31] 高城英誌, 河井良浩, 石山豊, 富田文明, “微分ベクトルに基づくシェイピング領域の抽出,” 第53回情処学金大, 2, pp.327-328, 1996.
- [32] 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.
- [33] T. Ueshiba, Y. Kawai, Y. Ishiyama, Y. Sumi, F. Tomita, “An Efficient Matching Algorithm for Segment-Based Stereo Vision Using Dynamic Programming Technique,” Proc. MVA'98, pp.61-64, 1998.
- [34] 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論文集, Vol.2, pp.123-124, 1994.
- [35] 佐藤準, 富田文明, “テクスチャー領域の両限立体視,” 情処学 CV 研資, CV55-5,

1988.

- [36] 松山隆司, 久野義徳, 井宮淳(編), “コンピュータビジョン技術評論と将来展望,” アドコム・メディア, 1998, pp. pp.123-137.
- [37] D. Marr, Vision, Freeman, 1982.
- [38] W. E. L. Grimson, “Computational experiments with a feature based stereo algorithm,” IEEE Trans. PAMI, Vol.7, No.1, pp.17-34, 1985.
- [39] Y. Ohta, T. Kanade, “Stereo by intra-and inter-scanline search using dynamic programming,” IEEE Trns. PAMI, Vol.7, No.2, pp.1074-1082, 1985.
- [40] H. H. Baker, T. O. Binford, “Depth from edge and intensity based stereo,” Proc. IJCAI81, pp.63-636, 1981.
- [41] G. G. Medioni, R. Nevatia, “Segment-based stereo matching,” Computer Vision, Graphics, Image Processing, Vol.31, No.1, pp.2-18, 1985.
- [42] N. Ayache, D. Faucher, “Efficient registration of stereo images by matching graph descriptions,” Int. J. Computer Vision, Vol.1, No.2, pp.107-131, 1987.
- [43] R. Horaud, T. Skordas, “Stereo correspondence through feature grouping and maximal cliques,” IEEE Trans. PAMI, Vol.11, No.11, pp.1168-1180, 1989.
- [44] O. Faugeras, Three-Dimensional Computer Vision: A Geometric Viewpoint, MIT Press, 1993.
- [45] 嶋崎真一, 金子俊一, 本多庸悟, “SimulatedAnnealingによるステレオ対応付けにおける隠れ部分の評価について,” 情報処理学会研究会資料, CV93-6, 1995.
- [46] 杉本和英, 富田文明, “輪郭線の屈曲点, 変曲点, 遷移点の検出,” MIRU'94, Vol.I, pp.83-90, 1994.
- [47] 角保志, 富田文明, “ステレオビジョンによる3次元物体の認識,” 信学論, Vol.J80-D-II, No.5, pp.1105-1112, 1997.
- [48] 角保志, 河井良浩, 吉見隆, 富田文明, “セグメントベースステレオによる自由曲面体の認識,” 信学論, Vol.J81-D-II, No.2, pp.285-292, 1998.
- [49] 角保志, 富田文明, “ステレオビジョンによる円錐複合体の認識,” 情処学 CV 研資, CV98-7, 1996.
- [50] 松下俊夫, 角保志, 富田文明, 石山豊, “ピジュアルトラッキングを用いたハンドリング作業状況の認識,” 第15回日本ロボット学会学術講演会, pp.985-986, 1996.
- [51] M. Potmesil, “Generating Models of Solid Objects by Matching 3D Surface Segments,” IJCAI-83, 2, pp.1089-1093, 1983.
- [52] C. K. Wu, D. Q. Wang, R. K. Bajcsy, “Acquiring 3-D Spatial Data of Real Object,”

- CVGIP, 28, 1, pp.126-133, 1984.
- [53] 安居院猛, 齊藤知人, 中嶋正之, “背面鏡を用いた 3 次元物体データ入力システム,” 信学論(D), J70-D, 5, pp.995-1002, 1985.
- [54] 西野治彦, 秋山健二, 小林幸雄, “光切断法による 3 次元立体形状計測と形状再構成,” 信学論(D-II), J72-D-II, 11, pp.1778-1787, 1989.
- [55] C. H. Liu, W. H. Tsai, “3D Curved Object Recognition from Multiple 2D Camera Views,” CVGIP, 50, 2, pp.177-187, 1990.
- [56] 上杉満昭, 猪股雅一, “イメージエンコーダーを用いた 3 次元曲面形状計測,” 映像情報, 21, 25, pp.43-48, 1989.
- [57] I. S. Kweon, T. Kanade, “High Resolution Terrain Map from Multiple Sensor Data,” IEEE IROS'90, 1, pp.127-134, 1990.
- [58] 吉見隆, 植芝俊夫, 大島正毅, “複数光源によるレンジファインダシステム,” 日本ロボット学会誌, 9, 7, pp.803-812, 1991.
- [59] 植芝俊夫, 吉見隆, 大島正毅, “光学系の 3 次元モデルに基づくレンジファインダの高精度キャリブレーション法,” 信学論(D-II), J74-D-II, 9, pp.1227-1235, 1991.
- [60] K. A. Koczmarec, J. G. Webster, W. J. Tompkins, “Electrotactile and Vibrotactile Displays for Sensory Substitution Systems,” IEEE Trans. Biomed. Eng., 38, pp.1-16, 1991.
- [61] R. Kowalik, I. Postawka, “The Concept of a Full Screen Tactile Display (FSTD) Driven by Electrochemical Reactions,” Proc. of ICCHP '94, pp.455-460, 1994.
- [62] J. Fricke, H. Bihring, “A Graphic Input/Output Table for Blind Computer Users,” Proc. of ICCHP '92, pp.172-179, 1992.
- [63] J. Fricke, H. Bahring, “Displaying Laterally Moving Tactile Information,” Proc. of ICCHP '94, pp.461-468, 1994.
- [64] 牧野秀夫, 石井郁夫, 馬場麻理, 大塚清和, 大和淳二, “盲人用地図作製および音声案内システムの開発,” 信学論, J73A, 3, pp.619-625, 1990.
- [65] 皆川洋樹, 大西昇, 杉江昇, “触覚と聴覚による盲人用図表現システム,” 信学論, J77-D-II, 3, pp.616-624, 1994.
- [66] 皆川洋樹, 大西昇, 杉江昇, “盲人の地図および環境の理解に関する実験と考察,” 信学論, J79D II, 5, pp.989-991, 1996.
- [67] M. Shinohara, et al., “Development of a 3-D Tactile Display for the Blind: System Design,” Proc. of ICCHP '92, pp.422-430, 1992.
- [68] “20 年のあゆみ,” 技術研究組合医療福祉機器研究所, 1996.
- [69] G. Jansson, “3-D Perception from Tactile Computer Displays,” Proc. of ICCHP '92,

pp.233-237, 1992.

- [70] H. Takahashi , T. Suehiro, “Stereo Vision for the Mimic Robot,” Computer Vision and Visual Communication, pp.75-78, 1994.
- [71] 伊東一典, 米沢義道 , 城戸健一, “音像定位制御による聴覚を通しての画像情報伝達,” 音響学誌, 42, 9, pp.708-715, 1986.
- [72] J. Loomis, C. Hebert , J. G. Cicinelli, “Active localization of virtual sounds,” J. Acoust. Sot. Am., 88, 4, pp.1757-1764, 1990.
- [73] F. Tomita, T. Yoshimi, T. Ueshiba, Y. Kawai, Y. Sumi, T. Matsushita, K. S. N. Ichimura , Y. Ishiyama, “R&D Versatile 3D Vision System VVV,” Proc. of SMC '98, pp.4510-4516, 1998.
- [74] 中村健太郎, 図解雑学 音のしくみ, ナツメ社, 1999.
- [75] J. Blauert, Spatial Hearing, MIT Press, 1996.
- [76] Z. Zhang, “Z. Zhang,” IEEE Trans. PAMI, 22, 11, pp.1330-1334, 2000.
- [77] T. Ueshiba , F. Tomita, “Plane-based calibration algorithm for multi-camera systems via factorization of homography matrices,” Proc. ICCV 2003, pp.966-973, 2003.
- [78] B. Caprile , V. Torre, “Using vanishing points for camera calibration,” IJCV, 4, 2, pp.127-139, 1990.
- [79] M. Fiala, “ARTag, a fiducial marker system using digital techniques,” Proc. CVPR 2005, 2, pp.590-596, 2005.
- [80] B. Triggs, P. F. McLauchlan, R. I. Hartley , A. W. Fitzgibbon, “Bundle Adjustment – A Modern Synthesis Vision Algorithms: Theory and Practice,” LNCS, Springer, pp.153-177, 2000.
- [81] 井上博允 , 比留川博久, “人間協調・共存型ロボットシステム研究開発プロジェクト,” 日本ロボット学会誌, 19, 1, pp,2-7, 2001.
- [82] M. Okutomi , T. Kanade, “A multiple-baseline stereo,” IEEE Trans. PAMI, 15, 4, pp.353-363, 1993.
- [83] 服部寛, “物体表面の勾配を考慮したステレオマッチング,” 情報学論, 42, SIG6(CVIM2), pp.22-32, 2001.
- [84] J. Weng, P. Cohen , M. Heriou, “Camera calibration with distortion models and accuracy evaluation,” IEEE Trans. PAMI, 14, 10, pp.965-980, 1992.
- [85] C. Baillard , A. Zisserman, “Automatic reconstruction of piecewise planar models from multiple views,” Proc. of CVPR'99, 2, pp.559-564, 1999.
- [86] R. Collins, C. Jaynes, Y. Q. Cheng, X. Wang, F. Stolle, E. Riseman , A. Hanson, “The

- Ascender System: Automated Site Modeling from Multiple Aerial Images,” CVIU, 72, 2, pp.143-162, 1998.
- [87] 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, 10, 2, pp.177-184, 1992.
- [88] 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, 371-D, 6, pp.1074-1082, 1988.
- [89] 井口征士 , 佐藤宏介, 三次元画像計測, 昭晃堂, 1990.
- [90] 谷内田正彦 , 石黒浩, “ロボットの視覚,” 人工知能学会誌, 5, 6, pp.720-730, 1990.
- [91] 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,” ロボット学誌, Vol.12, No.1, pp.15-16, 1994.
- [92] 富田文明, 吉見隆, 植芝俊夫, 河井良浩, 角保志, 松下俊夫, 市村直幸, 杉本和英 , 石山豊, “3 次元視覚システム vvv 研究開発 – 概要,” 情報処理学会研究会資料, CVIM109-1, 1998.
- [93] 石山豊, 角保志 , 富田文明, “ステレオ視による位置姿勢の変化する 3 次元物体の追跡,” 情報処理学会研究会資料, CV95-3, 1995.
- [94] 小谷信司, 清弘智昭 , 森英雄, “視覚障害者のための歩行ガイドロボットの開発,” 映像情報メディア学会誌, 51, 6, pp.878-885, 1997.
- [95] 篠田豊, K. Sasadara , 田所嘉昭, “携帯型視覚障害者歩行支援システムを用いた歩行訓練の評価,” 第 25 回感覚代行シンポジウム予稿集, pp.11-14, 1999.
- [96] 横井茂樹, 鳥脇純一郎 , 福村晃夫, “標本化された二値図形のトポロジカルな性質について,” 電子情報通信学会誌, 56, D, 11, pp.662-669, 1973.
- [97] 美濃導彦, 荒木誠 , 坂井利之, “論理型図面の清書方式とデータ量圧縮,” 電子通信学会, PRL85-4, pp.29-37, 1985.

付録

本付録においては、第3章で説明した画像処理による3次元情報取得・処理に関連する手法において重要と思われる関連研究

- ・ カメラキャリブレーション手法
- ・ ステレオ画像の輝度補正手法
- ・ ステレオ視における3次元復元の高精度化手法

に関して簡単に述べる。

A1. カメラキャリブレーション

A1.1. はじめに

ステレオカメラは周辺環境の情報を一度に取得できるため、作業対象物の認識・把持など幅広く利用されている。しかし、このような複数のカメラを利用するシステムは、実空間と画像面の関係の他に、カメラ同士の相対位置・姿勢を求める必要があり、単眼のものに比べて較正が難しく、導入に障害がある。本節では、単眼の較正法で広く用いられる平面パターンを観測する Zhang の手法 [76]と同様の手順で複数カメラの較正が実現できる、著者らが考案した手法を述べる。

既知の平面パターンを用いる方法は、基本的には市販のプリンタと丈夫な板のみで容易に作成可能なために扱いやすく、パターン間の距離が既知であることからスケールも決定することができる実用的な手法である。既知の平面パターンを利用した複数のカメラに対する較正法は、参照平面と各カメラ画像平面とのホモグラフィを利用するもの [77]、消失点の幾何学的な関係を利用してカメラの相対位置関係を推定するもの [78]、パターンに座標情報を埋め込むもの [79]などが提案されている。しかし、これらはカメラ画像間で対応点を求める必要があったり、観測点を密に配置できないなどの問題があったりするため、実用には適さない手法である。これらに対し本手法は、単純な繰り返しパターンの平面を数回観測するのみで較正でき、対応点も不要であることからカメラ間に共通視野が一切無い場合でも使用できる、簡便かつ汎用性の高い手法である。

A1.2. 参照平面を用いた単体カメラの較正

A1.2.1. Zhang の手法

ピンホールカメラモデルの下で、3次元空間中の点とその画像上の点は次の関係で対応付けられる。

$$\lambda \tilde{\mathbf{x}} = P \tilde{\mathbf{X}} = A[R|t]\tilde{\mathbf{X}} \quad (\text{A1-1})$$

$$A = \begin{bmatrix} \alpha & \gamma & u \\ 0 & \beta & v \\ 0 & 0 & 1 \end{bmatrix} \quad (\text{A1-2})$$

但し、 $\tilde{\mathbf{x}}$ と $\tilde{\mathbf{X}}$ はそれぞれ $\mathbf{x} \in R^2$ と $\mathbf{X} \in R^3$ の同次座標、 P は 3×4 の射影行列、 A は内部パラメータを表す 3×3 の行列、 $[R|t]$ は回転行列 R と平行移動ベクトル t からなる 3×4 行列である。一方、参照平面と画像平面の関係は次式で表される。

$$\lambda \tilde{\mathbf{x}} = A[\mathbf{r}_1 \ \mathbf{r}_2 \ t][X \ Y \ 1]^T = H[X \ Y \ 1]^T \quad (\text{A1-3})$$

これにより、観測から得たホモグラフィ行列とカメラの内部・外部パラメータが関係付けられるため、以下の線形方程式を解くことで、カメラの内部パラメータと観測した各平面パターンの位置・姿勢を求めることができる。

$$\begin{bmatrix} v_{12}^T \\ v_{11} \ v_{22} \end{bmatrix} \mathbf{b} = \mathbf{0} \quad (\text{A1-4})$$

$$B = \{B_{ij}\} = A^T A^{-T} \quad (\text{A1-5})$$

$$\mathbf{b} = [B_{11}, B_{12}, B_{22}, B_{13}, B_{23}, B_{33}]^T \quad (\text{A1-6})$$

$$\mathbf{v}_{ij} = [h_{i1}h_{j1}, h_{i1}h_{j2} + h_{i2}h_{j1}, h_{i2}h_{j2}, h_{i3}h_{j1} + h_{i1}h_{j3}, h_{i3}h_{j2} + h_{i2}h_{j3}, h_{i3}h_{j3}]^T \quad (\text{A1-7})$$

そして適当なレンズ歪みモデルを仮定し、以下の再投影誤差を最小化するバンドルアジャストメント [80]を行う。

$$\sum_{i,j} \|\mathbf{x}_{ij} - \tilde{\mathbf{x}}(A, \kappa_1, \kappa_2, R_i, \mathbf{t}_i, \mathbf{X}_j)\|^2 \quad (\text{A1-8})$$

但し $\tilde{\mathbf{x}}(\cdot)$ は、時刻 i における参照平面上の、 j 番目の 3 次元点 \mathbf{X}_j の画像平面への投影である。実装では、観測点の信頼度を制御できるようにするため、式(A1-8)の二乗誤差の代わりに重み付き最小二乗誤差を用いた。

A1.2.2. カメラ画像間の対応点

ホモグラフィを利用した単眼カメラの較正法は、参照平面の幾何学的な位置とスケールのみが重要であるため、参照平面上に設定する座標系の原点位置や軸の向きに注意を払う必要は無い。ところが、カメラ間の相対位置・姿勢を求めて再投影誤差を最小化する枠組みの上では、3 次元空間で同一の点に対応する各画像上の点の組が必要になる。

単純な繰り返しパターンを複数のカメラで撮影し、それぞれ任意の座標系を設定した場合を考えると、これらの座標系は平面上の平行移動・回転による 3 自由度の座標変換で結ばれることが分かる(図 A1-1)。

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} X' \\ Y' \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \quad (\text{A1-9})$$

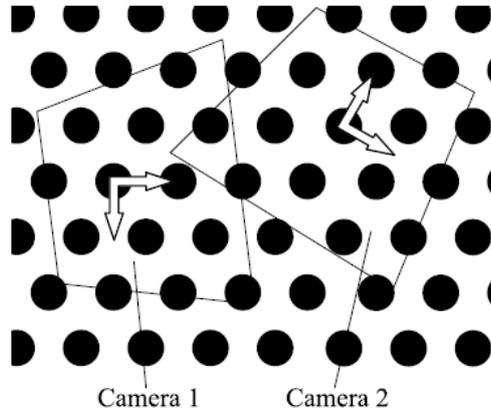


図 A1-1 参照平面上の座標系

さらに繰り返しパターンであることから、このズレはパターン間の距離の整数倍、点対称になる角度の整数倍に制限できることが明らかである。従って何らかの方法で得た座標変換を、このような性質に基づいてパラメータを整数値に丸めるなどの正規化を施すことで、参照平面に対してカメラ毎に設定した座標系を厳密に一致させ、参照平面上の任意の点に対する各カメラ画像上の点、すなわち対応点を求めることができる。換言すると、繰り返しパターンに関して対応点を求めることは上記の座標変換を求めることと等しく、対応点が複数のカメラに同時に撮影されるといった制約が無い場合、参照平面を共通視野内に収める必要がないことが分かる。次に、Zhang の手法によって得られた各時刻・各カメラに対する参照平面の位置・姿勢を用いて、この座標変換とカメラ間の相対位置・姿勢を推定する手法を述べる。

A1.3. 複数カメラの較正

n 台のカメラに対して m 姿勢の参照平面を観測し、各々 Zhang の手法を適用すると、それぞれのカメラに対応する n 組の内部パラメータと、 nm 個の参照平面の位置・姿勢が得られる。

$$\{A_i\}_{i=1}^n, \{R_{ij}, t_{ij}\}_{j=1..m}^{i=1..n} \quad (\text{A1-10})$$

このとき、参照平面の位置を表すベクトル t_{ij} は、設定された原点が観測毎に異なる可能性があるため、一般的には (i, j) のどの組み合わせに対しても参照平面上の同じ位置を指していることは保証されない。ここで、各観測・カメラ毎に割り当てられた参照平面上の座標系を、一時座標系 (temporal coordinate) と呼ぶことにする(図 A1-2)。

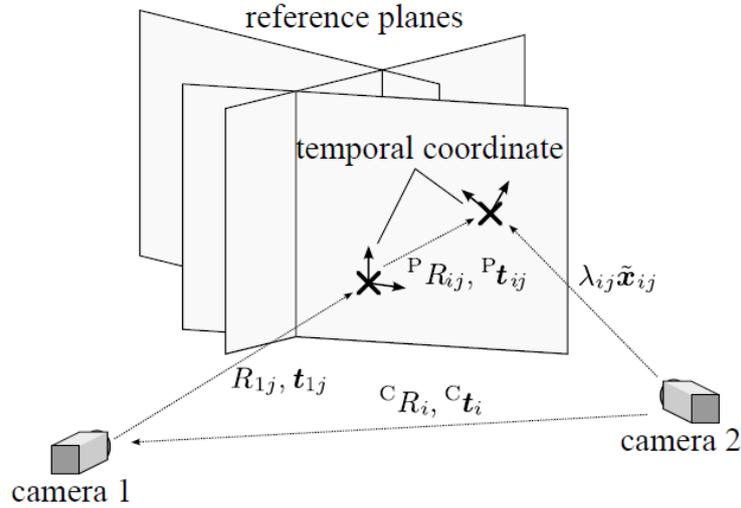


図 A1-2 各座標系の関係

A1.3.1. 参照平面の観測

A1.2.2 で示したように、ある参照平面の姿勢に対する一時座標系は、平面上の平行移動・回転によって表現できる。これにより、Zhang の手法で得た外部パラメータ $\{R_{ij}, t_{ij}\}_{j=1..m}^{i=1..n}$ は、

- 第 i カメラの位置・姿勢 ($[^c R_i \ ^c t_i]$)
- j 番目の参照平面の位置・姿勢 ($[^R R_j \ ^R t_j]$)
- 一時座標系の変換 ($[^P R_{ij} \ ^P t_{ij}]$)

を含んでいると考えられる。つまり、参照平面の観測は次式で表現できる。

$$\lambda_{ij} \tilde{\mathbf{x}}_{ij} = A_i \begin{bmatrix} ^c R_i & ^c t_i \end{bmatrix} \begin{bmatrix} ^R R_j & ^R t_j \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} ^P R_{ij} & ^P t_{ij} \\ \mathbf{0}^T & 1 \end{bmatrix} \tilde{\boldsymbol{\xi}}_j \quad (\text{A1-11})$$

ただし、 $\tilde{\boldsymbol{\xi}}_j$ は統一した参照平面上の座標、 $^P R_{ij}, ^P t_{ij}$ は統一した座標系から各一時座標系への変換である。

ここで、一般性を失うことなく世界座標系を第 1 カメラの座標系と一致させることができる。すると、カメラと参照平面の位置・姿勢は第 1 カメラからの相対位置・姿勢、一時座標系への座標変換は第 1 カメラの一時座標系からの相対位置・姿勢と読み替えられる(図 A1-2)。

A1.3.2. 幾何学的拘束からの初期値推定と最適化

第 1 カメラを原点に置くと、Zhang の手法によって得られた第 1 カメラの外部パラメータが参照平面の相対位置・姿勢に対応するため、初期値を求める必要があるのは、各カメラの相対位置・姿勢と一時座標系間の座標変換である。ここでは、幾何学的な条件を用いて、まず各カメラの相対位置・姿勢を算出する。これは消失点を用いた相対位置・姿勢の幾何学的推定に対応する。

さて、第 i カメラの j 回目の観測に対して得られた外部パラメータを $[\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{t}]$ とすると

次を得る.

$$[(\mathbf{r}_1 \times \mathbf{r}_2)^T, -(\mathbf{r}_1 \times \mathbf{r}_2) \cdot \mathbf{t}] \tilde{\mathbf{X}} = 0 \quad (\text{A1-12})$$

これは参照平面上の点 (X, Y) が満たす 3 次元空間中の平面方程式であり, 一時座標系と無関係な幾何的拘束である. この式は, 平面の法線 \mathbf{n} と原点から平面までの距離 $-c$ に変形でき, j 番目の観測においてシーンに対する参照平面の位置・姿勢は共通であることから, 第 1 カメラと第 i カメラの相対位置・姿勢を求める問題は, それぞれのカメラで観測した参照平面の法線と距離 $\mathbf{n}_{1j}, c_{1j}, \mathbf{n}_{ij}, c_{ij}$ の全ての j に関する次の問題へ変換できる.

$$\text{find } {}^cR_i, {}^c\mathbf{t}_i; \mathbf{X}_i = {}^cR_i \mathbf{X}_1 + {}^c\mathbf{t}_i \quad (\text{A1-13})$$

with

$$\begin{cases} \mathbf{n}_{1j}^T \mathbf{X}_1 + c_{1j} = 0 \\ \mathbf{n}_{ij}^T \mathbf{X}_i + c_{ij} = 0 \end{cases} \quad (\text{A1-14})$$

式(A1-13)を式(A1-14)へ代入し, $\mathbf{X}_1 \leftarrow \mathbf{n}_{1j}$ と置くと次式を得る.

$$\begin{cases} \mu_{ij} = \mathbf{n}_{ij}^T {}^cR_i \mathbf{n}_{1j} \\ \mathbf{n}_{ij}^T {}^c\mathbf{t}_i = \mu_{ij} c_{1j} - c_{ij} \end{cases} \quad (\text{A1-15})$$

ただし, $\mu_{ij} = \pm 1$ は第 i カメラの j 番目の観測において第 1 カメラと同じ面が見えているかどうかを表す係数であり, 参照平面が片面のパターンである場合は常に 1 である. 1 番目の式は回転に関するものであり, この式を最大にする R がカメラの相対姿勢となる. 回転 R が定まれば $\{\mu_{ij}\}$ が既知となるため, 第 2 式を用いて平行移動成分 \mathbf{t} の算出ができる.

次に, 残りの一時座標系間の座標変換の初期値を求める. この時点で各カメラの内部パラメータ A_i , 相対位置・姿勢 $[{}^cR_i \quad {}^c\mathbf{t}_i]$, 参照平面の位置・姿勢 $[{}^R\mathbf{r}_{j1} \quad {}^R\mathbf{r}_{j2} \quad {}^R\mathbf{t}_j]$ が既知であるため, 画像座標 x_{ij} から参照平面上の座標 ξ_j を直接算出できる. そこで元々割り当てられていた一時座標系上の座標を η_j として次を満たす 2 次元回転 ${}^P R_{ij}$, 平行移動 ${}^P \mathbf{t}_{ij}$ を決定する.

$$\eta_j = \frac{1}{\lambda'} H'^{-1} x_{ij} = {}^P R_{ij} \xi_j + {}^P \mathbf{t}_{ij} \quad (\text{A1-16})$$

$$H'_{ij} = A_i [{}^cR_i \quad {}^c\mathbf{t}_i] [{}^R\mathbf{r}_{j1} \quad {}^R\mathbf{r}_{j2} \quad {}^R\mathbf{t}_j] \quad (\text{A1-17})$$

この結果を A1.2.2 に記したように正規化することも可能であるが, ここで以下に示す最適化を行った後 ${}^P R_{ij}$ ${}^P \mathbf{t}_{ij}$ を正規化して固定し, 最終的なバンドルアジャストメントを行う.

$$\sum_{i,j,k} w_{ijk} \|x_{ijk} - \check{x}(A_i, \kappa_{i1}, \kappa_{i2}, {}^cR_i, {}^c\mathbf{t}_i, {}^P R_j, {}^P \mathbf{t}_j, {}^P R_{ij}, {}^P \mathbf{t}_{ij}, \xi_k)\|^2 \quad (\text{A1-18})$$

A1.3.3. アルゴリズムのまとめ

本手法のアルゴリズムを以下にまとめる.

1. Zhang の手法による各カメラの較正
2. 幾何的拘束からのカメラの相対位置・姿勢算出
3. カメラ間の対応点推定
 - (1) 画像座標の逆投影による座標変換の算出

(2) (最適化,) 一時座標系の座標変換の正規化

4. バンドルアジャストメント

A1.4. 実データを用いた較正実験

A1.4.1. 簡易 3 次元計測システムの構築

まず, Point Grey Research 社の IEEE1394b デジタルカメラ, Flea2 (640 × 480, grayscale) を 6 台使用して, 図 A1-3 に示すように配置し, 簡易的な 3 次元計測システムの較正を行った. 提案手法は任意の繰り返しパターンに適用可能であるため, ここではチェスボードパターンを用い, その検出に OpenCV を利用した. 図 A1-4 に示すように参照平面が観測される方向はカメラによって異なり, 平面上に割り当てられた座標系も統一されていないことが確認できる.

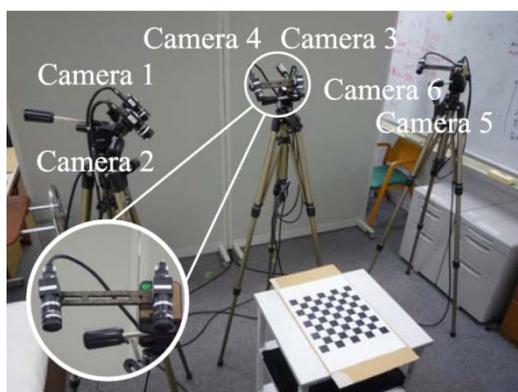


図 A1-3 チェスボードパターンを用いた較正の様子

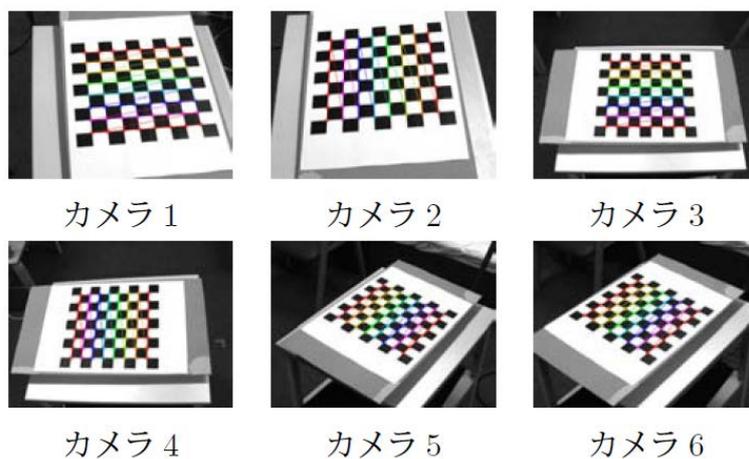


図 A1-4 OpenCV によるチェスボードパターンの検出

本手法によって較正を行った後, カメラ(1, 2), (3,4), (5, 6) の各組み合わせで簡単なテンプレートマッチングによるステレオ計測を行った(図 A1-5). 各組み合わせの基線長は約 17cm, 各カメラから対象物(メロンの模型) までの距離は, 概ね 100cm 前後, 対象物は直径約 13cm である. 図 A1-6(a) に各組み合わせにおける復元結果をテーブル上方からの視点で表示したもの, (b), (c)

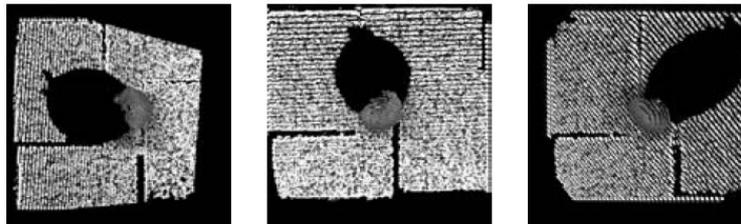
に全ての結果を単純に重ね合わせて表示したものを示す. この図より, 各組み合わせに対して破綻無く 3 次元復元が行えている上, 単純に合成したものを観測時と異なる視点で描画した際においても継ぎ目が目立たないことが確認でき, 提案手法が実用上十分な精度を達成していると考えられる.



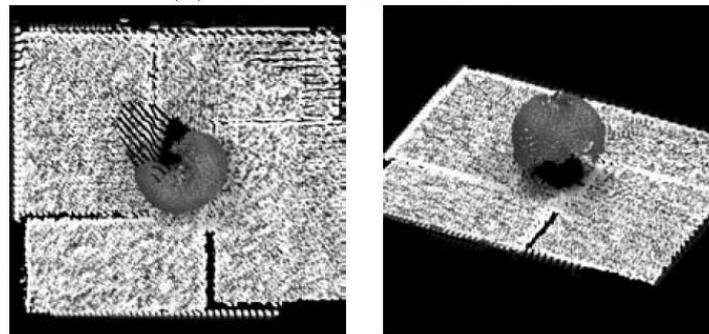
(a) 全景

(b) 計測対象

図 A1-5 計測風景と対象物



(a) 各カメラペアの復元結果



(b) 上方

(c) 斜方

図 A1-6 3次元復元結果

A1.4.2. 共通視野を持たないカメラの較正

ここでは, 図 A1-7に示すように3 台のカメラを用い, 左・右カメラで共通視野を持たないように配置した. そして三脚の自由雲台を操作し, 参照平面を異なる姿勢で計 5 回観測した. これによって得たデータのうち, 左右カメラのみ(共通視野無し), 3 台のカメラ全ての場合で計算を行い, 両者を比較した(図 A1-8, 図 A1-9).

推定された左・右カメラ間の相対位置の差は約 0.7mm , カメラの基線長に対する相対誤差が約 0.26% と良好な結果が得られた. これにより, 提案手法はカメラが共通視野を持たない場合におい

でも, 安定・高精度に計算できることが分かる.



(a) カメラの詳細



(b) 全景

図 A1-7 較正の様子

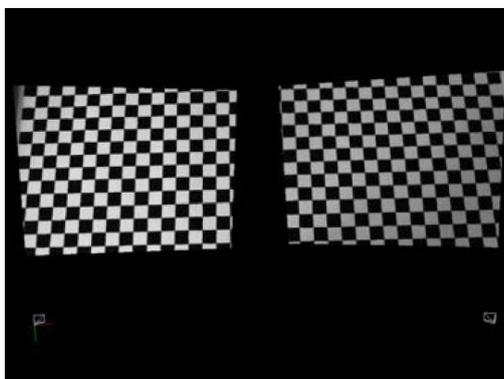


図 A1-8 共通視野のない左右カメラのみを使用した場合のカメラの較正結果

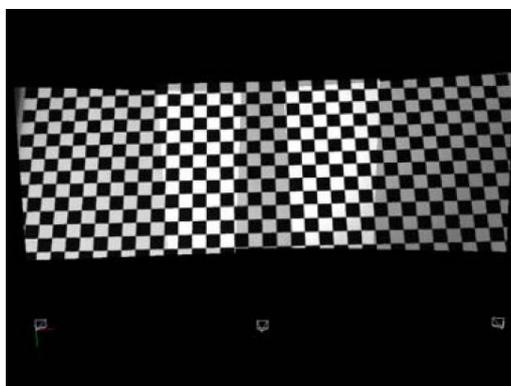


図 A1-9 3台全てを使用した場合のカメラの較正結果

A1.5. まとめ

本節では, 単純な繰り返しパターンを持つ参照平面を用いて, 複数のカメラの内部パラメータと互いの相対位置・姿勢を同時に較正する手法を述べた. 本手法は予め対応点を求める必要がない上, 推定した対応点に対して再投影誤差の最小化を行う, バンドルアジャストメントが実行できる精度の高いものである. 3次元復元システムの較正と, 共通視野が無い場合の較正の実験を行い, 本手法が複数カメラシステムの較正を効果的に実現できることを示した.

A2. 輝度補正

A2.1. はじめに

ステレオ画像において、正確な対応関係を求めるためには画像間の輝度調整が正しく行われている必要がある。特に輝度に基づく相関法の場合は輝度の調整が不可欠である。相関法には大きく二通りあり、正規化された評価式を使う方法[1]と差分評価式を使う方法[2]がある。前者の方法は、図 A2-1 (a)に示すテクスチャなどのような相関窓内で輝度変化が顕著な領域には有効であり、正規化しているため絶対的な輝度一致がなくても良い。しかし、(b)のシェーディング領域のように滑らかに輝度に変化する部分に対しては誤対応を起こす可能性がある。シェーディング領域の正しい視差を計算するためには、後者の方法が有効であり、そのためには画像間の輝度対応を正しくしなければならない。

そこで、輝度の多少の変動には比較的頑健であるセグメントペアズステレオ法[3]を利用して画像間の対応関係を求め、その対応部分の輝度情報を用いて画像の輝度補正を自動的に行う方法を考案し、実験で本手法の有効性を示す。

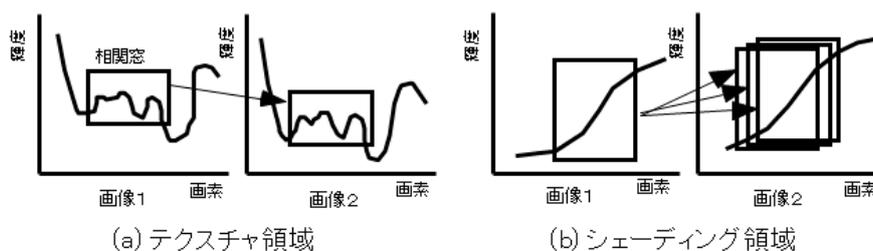


図 A2-1 輝度相関

A2.2. 処理手順

まず、各画像でエッジを検出し、屈折点、変曲点、遷移点などの特徴点に分割し、直線、凸曲線、凹曲線セグメントを求める。セグメントを構成する各点の持つ輝度情報はエッジ部分そのものではなく、各セグメントが属する領域内でその垂直方向の近傍で、微分値が一番小さい点の輝度値と微分値である。

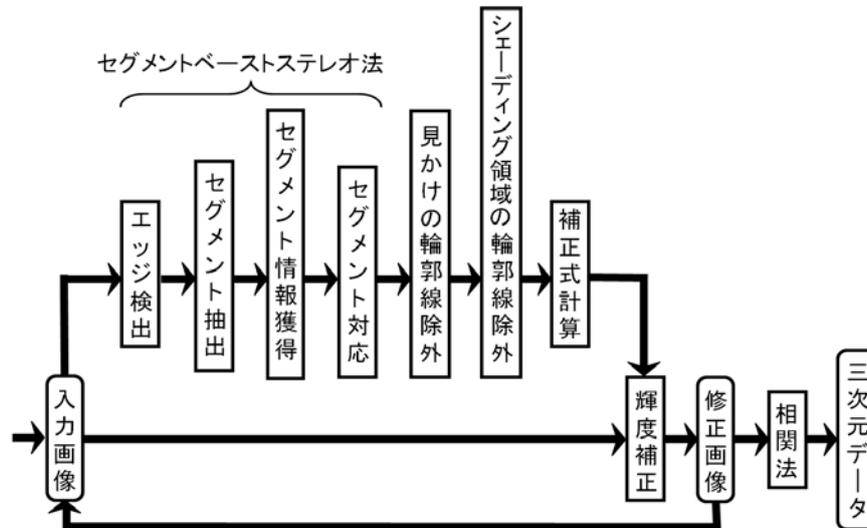


図 A2-2 処理手順

輝度補正を行う際の基準画像(輝度補正をしない画像) I_1 を定め、基準画像と被補正画像 I_2 との間のセグメントの対応を、それらの連続性に基づいたセグメントベースステレオ法で求める。対応を求めるセグメント、つまりエッジ部分は輝度変化の大きい領域間の境界線であり、画像間で絶対的な輝度の一致がなくても、ほぼ同じ場所にエッジは検出される。それゆえに、この特徴量を用いるステレオ法は輝度変化に対して頑健である。

ただし、セグメントベースステレオ法では曲面の輪郭線、つまり、遮蔽輪郭線も対応候補として扱っている。しかし、輝度補正においては、遮蔽始郭線の対応部分は厳密には真の対応ではないため、補正には固定エッジセグメントの対応関係だけを使用する。遮蔽輪郭線部分の輝度変化は大きいため、セグメントを構成する点列の各点における輝度情報の微分値が大きくなる。この性質を利用して、対応セグメントの中から遮蔽輪郭線を除く。また、多眼ステレオ画像の場合には、他視点からの画像を検証用画像として用い、検出位置のずれから遮蔽輪郭線を除くこともできる[4]。

次に、求めたセグメントの対応からステレオ画像間の輝度の関係を求めることができる。輝度値に関して画像間 (I_1, I_2) の対応点の分布から、画像間の輝度補正式を求める。対応点が正しければ、点は直線上に分布するので、直線の補正式を計算する。

$$I_2(0) = aI_1 + b \quad (\text{A2-1})$$

a, b : 係数, $I_i(n)$: n 回補正後の画像

この補正式を基に画像 $I_2(0)$ の輝度補正を行い、補正画像 $I_2(1)$ を求める。基準画像と輝度補正された画像を入力画像として同様な処理を補正式がほぼ $I_2(n) \cong I_1$ になるまで繰り返す(図 A2-2)。

このように補正された画像に対して、SAD(Sum of Absolute Difference)や SSD(Sum of Squared Difference)などの差分評価式による相関法を行うことで正しい対応を求めることができる。

A2.3. 補正結果

提案した手法の実験結果を示す。図 A2-3(a)は、対象とするステレオ画像で、右画像を左画像に対してカメラの絞りを少しクローズにして撮影した。本実験では左画像を基準画像として処理を行った。(b)の細線は検出されたエッジセグメントで、セグメントベースステレオ法により求めた対応セグメント部分が太線である(画像は左画像)。遮蔽輪郭線部分の除外処理後の結果が(c)であり、球や円柱部分の輪郭が除かれている。補正に用いた対応点は 1593 点で、 x 軸を左画像の輝度、 y 軸を右画像の輝度とした輝度分布は(d)となり、ほぼ直線上に分布していて、対応点の輝度は右画像の方が若干低いことがわかる。この分布に直線を当てはめ、輝度補正式を求めると、

$$I_{right}(1) = 0.852347825 I_{left} - 2.554870067$$

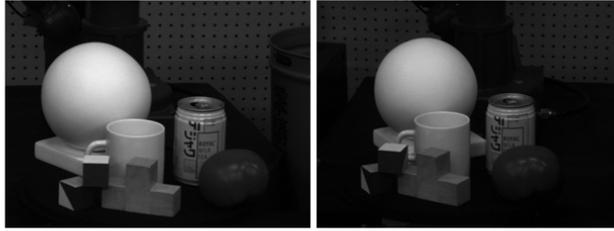
となる。この補正式で右画像の輝度補正を行い、補正画像を入力画像として、 $I_{right}(n) \cong I_{left}$ になるまで同様の処理を繰り返す。この画像では 5 回の繰り返し処理で収束し、最終的な補正式は、

$$I_{right}(5) = 0.831778529 I_{left} - 2.275367400$$

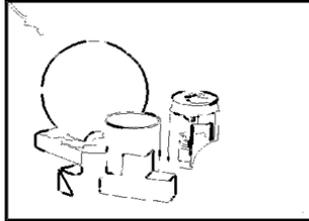
となり、求めた補正式で右画像の輝度を補正した画像が(e)である。補正された(e)の画像に対する、セグメントベースステレオ法による対応点の輝度分布は(f)である。基準画像(左)と補正画像(右)間の輝度関係式は、

$$I'_{right} = 1.019867106 I_{left} - 0.988590436$$

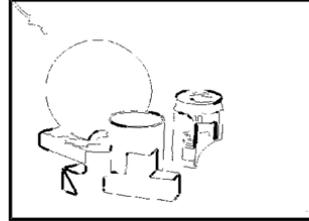
であり、 $I'_{right} \cong I_{left}$ になっている。この(e)の画像を基に、輝度変化のある部分のみの視差をSADによる相関法で計算し、形状復元した結果は(h)であり、補正していない復元結果(g)と比べ、例えば、球面の正しい 3 次元形状が得られた。



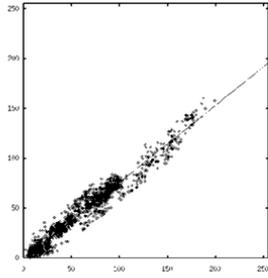
(a) 入力ステレオ画像(左[基準], 右)



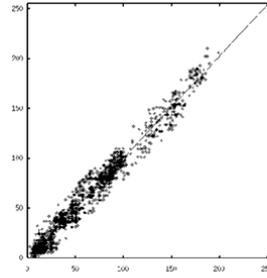
(b) セグメントベーストステレオ法による対応点



(c) 遮蔽輪郭線除去後の対応点



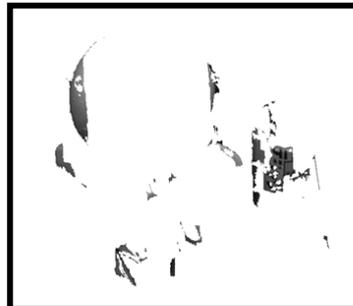
(d) 輝度分布(補正前)



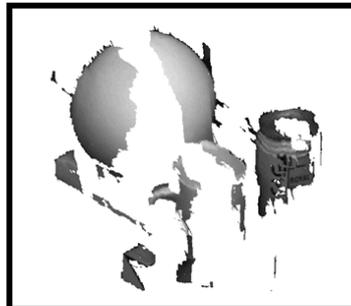
(f) 輝度分布(補正後)



(e) 補正画像



(g) 復元結果(補正前)



(h) 復元結果(補正後)

図 A2-3 補正結果

A2.4. まとめ

輝度の対応関係が完全でないステレオ画像において、輝度変化に強い特徴の対応関係から、輝度の補正を画像から自動的に行うことができる方法を考案し、実験によりカメラの個体差などによる絞りの不調整分を自動で補正できた。

A3. ステレオ視における 3 次元復元の高精度化

ステレオ視において、セグメント間の連結性に基づく平面拘束を利用して、高精度な 3 次元情報を復元する方法について述べる。計測対象までの距離に対して基線長が相対的に短いステレオカメラシステムでは、画像面上での対応点の検出誤差が遠方物体の 3 次元復元に(特に奥行き方向に対して)与える影響が大きい。そのため、対応点間の視差から単純に 3 次元情報を復元した場合、十分な復元結果が得られない場合が多い。提案する手法では、対応するセグメント間において、その連結性から平面拘束を行い、視差を補正する。本手法により、3 次元においてセグメントの連続性を保ちつつ高精度な復元が可能となる。実験結果から本手法の有効性を示す。

A3.1. はじめに

ビジョンセンサとして、ステレオカメラシステムが様々なシステムに組み込まれてきており、例えばヒューマノイドにもステレオカメラシステムが搭載されている [81]。しかしながら、まだ十分な視覚機能がないため、遠隔操作が基本であり、自律行動は限られたものでしかなく、周りの環境、状況を認識できる 3 次元視覚機能の研究・開発が必要とされている。

ステレオ視に関する研究は古くから多くの研究がなされてきているが [36]、最近では輝度値の局所的な相関値の類似性から対応を求める相関法ベースのものが多い [82] [83]。アルゴリズムが単純なためハードウェア化しやすく、リアルタイムで奥行き方向の情報を得られるシステムもある。しかしながら、対象はテクスチャを含むものに限られ、物体認識などを行う場合にはセグメンテーションなどの後処理が必要となる。

一方、画像から高次特徴であるエッジセグメントを抽出し、その対応を求めるセグメントベースドステレオ法は、処理が複雑になる反面、探索空間をかなり狭めることができ、また、カメラの特性などによる画像間の輝度の多少の変化に関して比較的頑健であり、復元結果が 3 次元構造記述そのものになり得るため、モデリングや物体認識 [47] [48]などに直接利用できる利点がある。

前述のヒューマノイドなどでは、顔部分の形状や大きさの制限から、カメラ間距離、つまり、基線長が短くならざるをえない。そのため、画像上におけるデジタル誤差や雑音などが 3 次元の復元精度に大きく影響し、特に遠方物体になるほど十分な精度を得られない。

本節では、計測対象までの距離に対して基線長が相対的に短いステレオカメラシステムにおいて、セグメント間の連結性を考慮した高精度な 3 次元復元法について述べる。3 次元復元精度に影響を与える要因として、レンズの歪曲収差による画像の歪みやエッジ検出誤差などがあり、補正法もいろいろあるが [84]、本節ではこれらを補正した上で、誤差を含む視差情報からの高精度な 3 次元復元法に焦点を絞る。対応点間の視差から単純に 3 次元情報を復元した場合、サブピクセル以下のわずかな誤差でも、奥行き方向に対する誤差に大きく影響する。視差を補正する方法として、画像面上でのセグメント単位で直線、曲線などを当てはめる方法や、頂点間を結ぶ方法などがあるが、それぞれ、3 次元的に連結しているセグメントがねじれの位置の関係になったり、頂点位置検出の不正確さが大きく影響したりするなどの問題がある。そこで、対応するセグメント間において、隣接セグメントとの連結性に基づく平面拘束により視差を補正することで、連結するセグメントとの

関係を保ちつつ 3 次元復元誤差を軽減する方法を開発した。平面拘束を利用した 3 次元復元法には航空写真からの建物形状の復元 [85] [86] などがあるが、直線セグメントの組合せのみを対象としているが、本手法は平面上の曲線セグメントを区別なく扱うことができる点で優れている。

本章では、ステレオ視における計測誤差の問題を説明し、対応するセグメント間の連結性に基づく平面拘束による視差の補正方法を詳細に述べ、実験結果から本手法の有効性を示す。

A3.2. 計測誤差

ステレオカメラシステムにおいて、画像上での点の検出誤差が 3 次元座標計算に与える影響を説明する。画像のデジタル誤差や雑音などの影響を減らすために重み付け平均を行うことなどにより境界線の位置をサブピクセルオーダーで計算しても、その影響を完全に取りきれず、誤差を含んだ検出となる場合が多い。このわずかな誤差が、基線長が短いステレオカメラシステムでは 3 次元復元に大きな影響を与えることを示す。

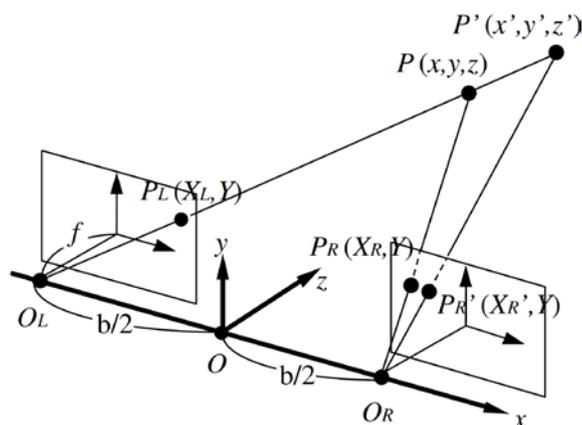


図 A3-1 標準カメラモデル

ステレオカメラとして図 A3-1 に示すように、2 台のカメラの光軸方向が平行である標準ステレオカメラモデルを基に考える。基線長を b 、焦点距離を f とすると、計測点 $P(x, y, z)$ の左右の画像面上への投影点 $P_L(X_L, Y), P_R(X_R, Y)$ は、エピポーラ条件 $Y = \text{一定}$ を満たしており、この場合

$$P_L = \begin{pmatrix} X_L \\ Y \end{pmatrix} = \frac{f}{z} \begin{pmatrix} x + \frac{b}{2} \\ y \end{pmatrix} \quad (\text{A3-1})$$

$$P_R = \begin{pmatrix} X_R \\ Y \end{pmatrix} = \frac{f}{z} \begin{pmatrix} x - \frac{b}{2} \\ y \end{pmatrix} \quad (\text{A3-2})$$

となる。よって、式(A3-1), (A3-2)から計測点 P は

$$P = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \frac{b(X_L + X_R)}{2(X_L - X_R)} \\ \frac{bY}{X_L - X_R} \\ \frac{bf}{X_L - X_R} \end{pmatrix} = z \begin{pmatrix} \frac{X_L + X_R}{2f} \\ \frac{Y}{f} \\ 1 \end{pmatrix} \quad (\text{A3-3})$$

で表される。

画像面上での対応点の計測誤差問題を簡略化するために左画像上での投影点は固定し、右画像上での投影点が X 方向に ΔX_R だけずれた場合を考える。その投影点を $P'_R(X_R + \Delta X_R, Y)$, ずれて計測された点を $P'(x', y', z')$ とする。計測誤差は点 PP' 間の距離であり、式(A3-1)～(A3-3)から

$$\begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} = \begin{pmatrix} x' - x \\ y' - y \\ z' - z \end{pmatrix} = \frac{\Delta X_R z^2}{bf - \Delta X_R z} \begin{pmatrix} \frac{X_L}{f} \\ \frac{Y}{f} \\ 1 \end{pmatrix} \quad (\text{A3-4})$$

と導かれる。一般のカメラを考えた場合、

$$\left| \frac{X_L}{f} \right| < 1, \quad \left| \frac{Y}{f} \right| < 1 \quad (\text{A3-5})$$

が成り立つため、式(A3-4), (A3-5)から、画像面上での X 方向の誤差は、奥行き方向の z に関する誤差に一番大きく影響することがわかる。

ちなみに、式(A3-4)の分母は幾何学的な拘束から、

$$X_L > X'_R = X_R + \Delta X_R \quad (\text{A3-6})$$

であり、これを变形すると、式(A3-3)から

$$bf - \Delta X_R z > 0 \quad (\text{A3-7})$$

が導かれ、常に正となることが言える。よって、式(A3-4)から、基線長が短いほど、または、対象までの距離が遠いほど、画像面上での対応点の計測誤差の影響が大きくなることがわかる。また、実際には左画像の誤差 ΔX_L も加わるため、さらに誤差が大きくなる。つまり、基線長の短いステレオカメラシステムでは、投影面上での対応点の検出誤差が3次元計測誤差に与える影響は大きくなる。

A3.3. 平面拘束による誤差の軽減

ここでは、このような誤差を軽減するための手法を述べる。最初に標準カメラモデル座標系への変換、左右の画像間での対応点の求め方を説明し、一般的な3次元復元法の問題点を挙げる。そして、提案手法であるセグメントの連結性に基づく平面拘束を利用した視差補正法を詳しく説明する。

A3.3.1. 標準カメラモデル座標系への変換

ステレオにおける左右画像での対応計算を簡単にするため、カメラキャリブレーションデータから、

撮影画像を図 A3-1 標準カメラモデルに示すような標準ステレオカメラモデル系に基づく画像に仮想的に変換する [87]. これにより, セグメントに属する画像上の各点 (col, row) は標準ステレオカメラ系の座標 (X, Y) に変換され, エピポーラ条件は $Y = \text{一定}$ となる. また, 焦点距離 f は, 変換後の画像サイズが元の画像のサイズにほぼ同一になるように定める. 以後, 画像の座標に関しては, 標準カメラモデルに変換したものを扱う.

A3.3.2. 対応点計算

セグメントベースドステレオにおける対応セグメントの探索法に関しては, 0に記載されており, ここでは対応セグメント上の対応点の求め方のみを説明する. 対応点はエピポーラ条件を満足するので, 標準カメラモデル(図 A3-1 標準カメラモデル)においては, 左右それぞれ (X_L, Y) , (X_R, Y) となる. しかしながら, 実際の処理においては, セグメンテーションは原画像で行い, 点の座標値のみを標準カメラモデル座標系に変換しているため, 図 A3-2 に示すように, 左画像のセグメントを構成する各点○の Y 値は, 右画像のセグメント上の点の Y 値と通常は一致しない. セグメントにおける Y 値は単調増加(減少)が保証されているので, Y を挟み込む連続する 2 点 ● $(X_{R1}, Y_{R1}), (X_{R2}, Y_{R2})$ を結んだ直線が Y となる点× (X_R, Y) を計算し, この点を (X_L, Y) の対応点とする(図 A3-2).

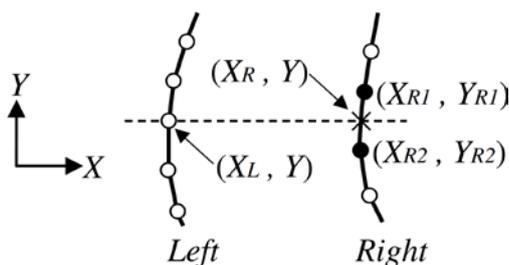


図 A3-2 対応点の決定

A3.3.3.3 次元復元時の問題点

3次元距離は視差から計算できるが, その際の問題点を説明する. まず, 一番単純な3次元復元法は, 上記で求めた対応点 (X_L, Y) , (X_R, Y) を利用して式(A3-3)から計算する方法であるが, エッジ検出時の誤差で生じる視差の誤差が復元精度に大きく影響する. また, 近傍点を利用して平滑化し, 誤差を軽減する方法もありえるが, 大きながたつきがなくなるわけではなく, 本質的な解決にはならない. 次に, 対応するセグメントそれぞれに最小二乗法などで直線や曲線を当てはめ, セグメントを構成する点群のがたつきをなくす方法が挙げられる. これによりセグメント内の点群のぶれがなくなるが, 頂点部分の位置もずれてしまうため, 隣接するセグメントとねじれの位置関係になる場合が多く, 対応点の3次元位置の連続性が保証されない. このようなねじれの位置関係にはならないように, 頂点を連結する方法が挙げられる. この方法はエピポーラ条件を満足するが, エッジ検出過程において頂点自身の位置検出の不安定さがある点や, 中間部分の情報が生かされない点などの問題がある.

A3.3.4. 平面拘束に基づく対応点補正

以上の問題を解決するため、連結するセグメントの情報を利用して対応点の位置補正を行う。セグメント間の対応を求める方法は連結性に基づいた手法であるため、隣接セグメントとの連結性を容易に得ることができる。ゆえに、対象セグメントと連結セグメントの組み合わせで求まる平面に点列が存在する場合は、その平面上に投影することで、法線方向の誤差を軽減できる。対象セグメントと連結セグメントが平面上に存在するかどうかは、以下の平面拘束定理 [88]により調べることができる。この平面拘束により、対応点の補正が可能で、また、前述のねじれ問題も解決できる。

平面拘束定理

標準カメラモデルにおいて、平面上の任意の図形の一方のカメラ画像上の投影像と、他方のカメラの画像上の投影像とはアフィン変換可能である。

対応する左右のセグメント S_{L_i}, S_{R_j} に属する対応点(点数 n)を $(X_{L_{ik}}, Y), (X_{R_{jk}}, Y)$ とし、それらが平面上に存在するならば、平面拘束定理から、

$$P_{R_{jk}} = AP_{L_{ik}} \quad (k = 1 \dots n)$$

$$P_{L_{ik}} = \begin{pmatrix} X_{L_{ik}} \\ Y \\ 1 \end{pmatrix}, \quad P_{R_{jk}} = \begin{pmatrix} X_{R_{jk}} \\ Y \\ 1 \end{pmatrix}, \quad A = \begin{pmatrix} a & b & c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (\text{A3-8})$$

というアフィン変換が成立する。各セグメントに対して、以下のアルゴリズムで連結するセグメントの情報を使用して対応点の補正を行う。

- (1) 対象セグメントの組 (S_{L_i}, S_{R_j}) に対し、連結する前後のセグメントの組 $(S_{L_{i-1}}, S_{R_{j-1}}), (S_{L_{i+1}}, S_{R_{j+1}})$ の存在を調べる。
- (2) 対象セグメントと連結するセグメントが同一直線上に存在する場合は、それらを1つの対象セグメントの組 (S'_{L_i}, S'_{R_j}) と見なして、(1)に戻り連結性を調べ直す。
- (3) 選択したセグメントの組の対応点から、式の ε を最小にする A を求める。

$$\varepsilon = \sum_{l,j}^3 \sum_{k=1}^{n_{l,j}} |AP_{L_{lk}} - P_{R_{jk}}|^2 \quad (\text{A3-9})$$

ただし、連結性がなく (S_{L_i}, S_{R_j}) だけの場合、それが曲線セグメントならば単独で平面を定義できるが、直線セグメントの場合は $a = 1$ として、視差を Y の1次関数で表現、つまり直線フィットを行う。

- (4) 誤差 E を式(A3-10)で定義した場合、 E がしきい値以下であれば、式(A3-9)を満足するとし

て(5)に進む。それ以外は、そのセグメントの組合せでは平面が構成されないとし、連結するセグメントの組で、 E が大きくなる原因の組の方を除いて(3)に戻る。対象セグメントの組だけになった場合は(5)に進む。

$$E = \max \left(\left| AP_{L_i k} - P_{R_j k} \right| \right) \quad (\text{A3-10})$$

(5) 式(A3-8)から、対象セグメントの右画像の対応点の X_R を X'_R に変換する。

$$X'_{R_j k} = aX_{L_i k} + bY_{(i,j)k} + c \quad (\text{A3-11})$$

(6) 対象セグメントの (X_{L_i}, Y) と、補正した対応点 (X'_{R_j}, Y) を用いて式(A3-3)から 3 次元座標を計算する。

これにより求めた点列はセグメントの組み合わせによって決まる平面上に投影されるので、その法線方向に対する誤差が軽減される。

次に、平面上での誤差を軽減する処理を行う。対象が直線セグメントの場合は、各画像で共にエピポーラ線に水平でない ($Y \neq \text{一定}$) 直線に投影されているならば、3 次元的にも直線であることから、その平面上に投影された点列に対して直線を当てはめ、3 次元空間において直線になるように変換を行う。曲線においても同様に平面上で曲線フィットを行う。

A3.4. 実験

本節では、ヒューマノイド用に開発した基線長の短いステレオカメラシステムを使用して行った実験結果を示し、本手法の有効性を示す。

図 A3-3 は小型軽量の 3 眼ステレオカメラシステムで、各カメラには焦点距離 8mm のレンズを搭載したシーアイエス製ボードカメラ BCC-155RE を使用した。なお、使用したレンズの焦点距離が短いため、画像に歪みが生じ、その影響が無視できない。そこで、あらかじめ撮影した平面のキャリブレーションパターンから歪み補正係数を計算しておき、標準カメラモデル座標系に変換する際に歪み補正を行っている。左右、左下の基線長はカメラキャリブレーションの結果、それぞれ 115.6mm , 77.5mm であり、このシステムの左右のカメラにおける式(A3-5)の最大値は、

$$\max \left(\left| \frac{X_L}{f} \right| \right) = 0.30, \quad \max \left(\left| \frac{Y}{f} \right| \right) = 0.22$$

であり、前述のように z 方向への影響が一番大きくなることを示している。式(A3-4)における Δz の変化を基線長 b 、奥行き方向 z を変数としてグラフ化したものが図 A3-4 である ($f = 1082.5\text{ pixel}$, $\Delta X_R = 0.1\text{ pixel}$, $f = 115.6\text{ mm}$, $z = 1642.5\text{ mm}$ を基準とした)。視差の誤差 ΔX_R が 0.1 pixel であっても、誤差 Δz は 7.9 mm とかなり大きくなることがわかる。また、基線長を現在の半分にすると最大計測誤差は約 2 倍になり、対象までの距離が 2 倍になると誤差は約 4 倍になることが読み取れる。

実験 1 では簡単な対象における 3 次元復元精度向上例を示し、実験 2 では遠方の対象物でも精度良く復元でき、位置、姿勢を正しく認識できることを示す。



図 A3-3 ヒューマノイド(HRP-2P)用ステレオカメラシステム

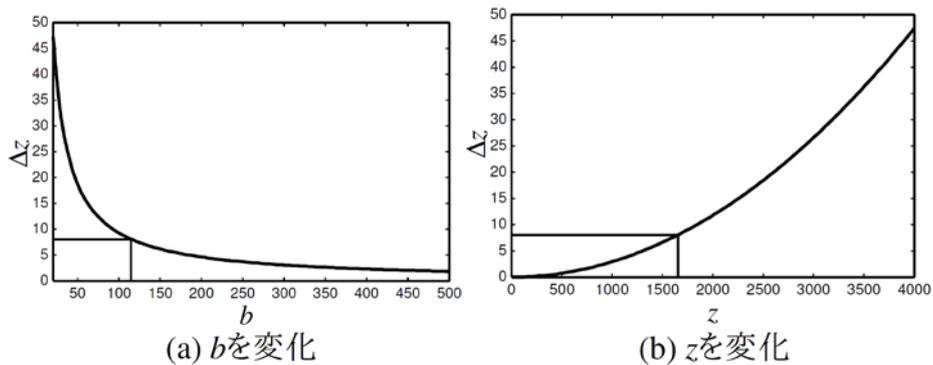


図 A3-4 基線長 b , 奥行き z をそれぞれ変化させた場合の Δz の変化例 (単位は mm)

A3.4.1. 実験 1: 近距離にある単純図形

実験 1 として, 近い距離にある単純図形を対象とし, 提案手法の有効性を示す. 図 A3-5(画像の一部のみ表示)のように, 約 $1.6 m$ 先に設置されたホワイトボード上の 1 辺 $210 mm$ の黒い 4 角形を対象とし, 構成する 4 辺の 3 次元復元について実験した. 各画像においてエッジをサブピクセルオーダーで検出し, 対応点を図 A3-2 の方法で求めた. その対応点間の視差 $X_L - X_R$ と Y の関係をグラフ化したものが図 A3-6 である. 理想的には 4 角形になっているべきであるが, サブピクセルオーダーで見ると, かなりぶれているのがわかる.

この視差をそのまま使用して 3 次元復元を行った結果が図 A3-7(a) であり, 奥行き方向に対して大きく変動している. 図 A3-7(b)は各画像の直線セグメントに最小二乗法で直線を当てはめから視差を計算し, 復元したものである. この場合, 各セグメントの直線性は改善されるが, 隣接するセグメントとはねじれの位置関係になっている. 図 A3-7(c)は本手法による結果で, 連結性を考慮して平面拘束した直線当てはめになるため, より忠実に 4 角形が復元されている.

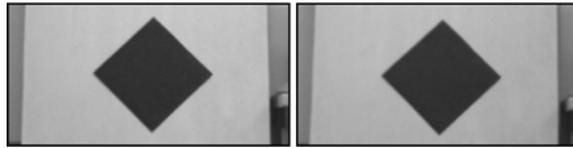


図 A3-5 対象シーン(約 1.6 m 先のホワイトボード上に、
1 辺 210 mm の黒い 4 角形を張り付けたもの)

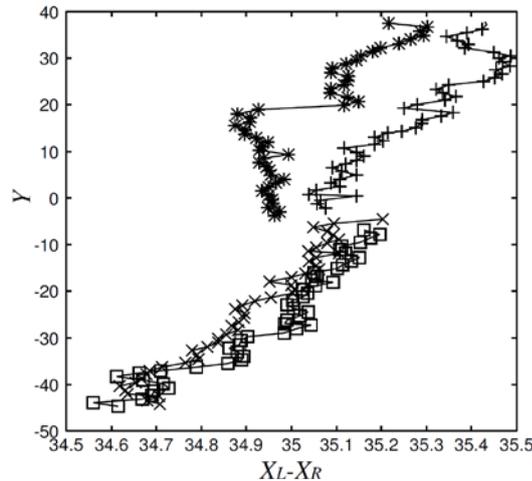


図 A3-6 4 辺の対応点の視差(*, +, x, □は, 各辺の対応点)

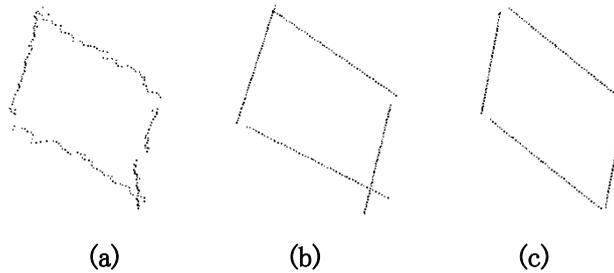


図 A3-7 左斜め上方から見た復元結果

((a)補正なし (b)セグメントに直線当てはめ (c)本手法)

A3.4.2. 実験 2: 一般的なシーン

実験 2 ではより一般的な実験室のシーンを対象とし、遠方の物体でも精度良く復元できることを示す。図 A3-8, 対象シーンであり、約 5000mm 遠方の壁にパネル(600×1800×20 mm)が立てかけられている。図 A3-9(a)が補正しないで 3 次元復元した果であり、図 A3-9(b)が補正後の結果である。図 A3-9(a)では図 A3-9(b)にあるように、カメラに近い対象に比べ、奥にあるパネル部分は奥行き方向に大きく乱れているのが、図 A3-9(b)においては対象の形をより正確に復元できている。

この補正後の復元結果(図 A3-9(b))で遠方のパネルを認識した結果を示す。認識手法は文献 [47]を適用した。図 A3-10(a)は認識対象のパネルのモデルである。このモデルがシーンのどの位

置にどのような姿勢で存在するかを調べ結果が図 A3-10(b)であり, 認識したモデルを重ね合わせて表示している. また, パネルを左画像上に投影したものが図 A3-10(c)である. 対象のパネルの位置姿勢を正確に認識できていることがわかる.

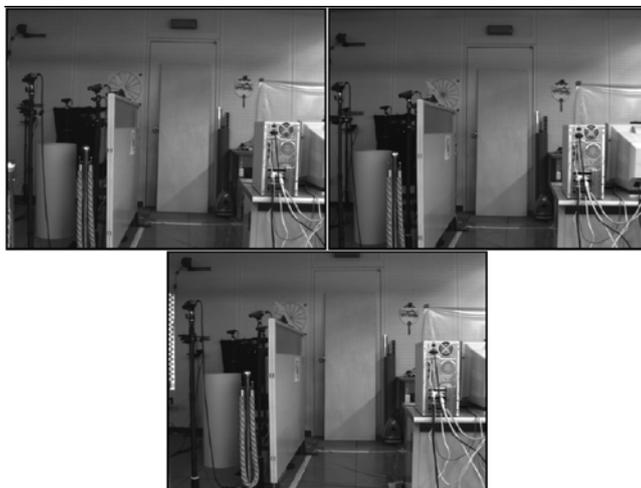


図 A3-8 対象シーン(実験室)

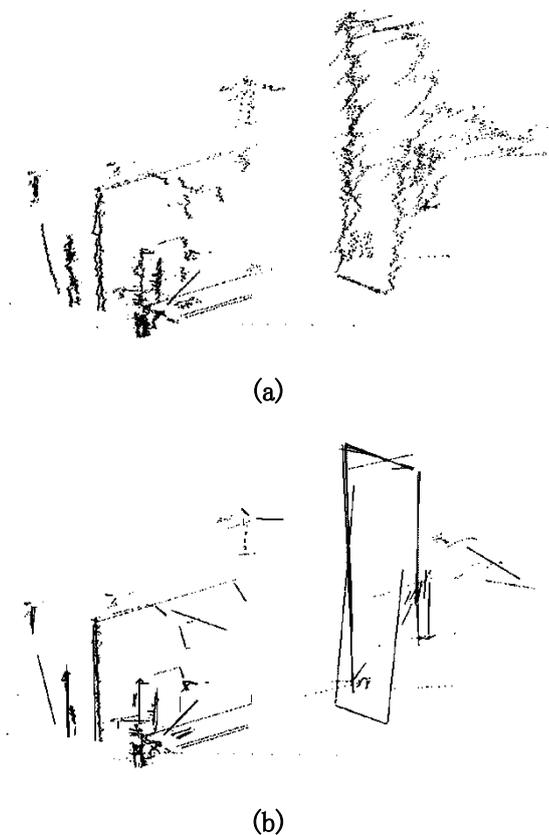


図 A3-9 左斜め上方から見た復元結果 (a)補正なし (b)補正あり

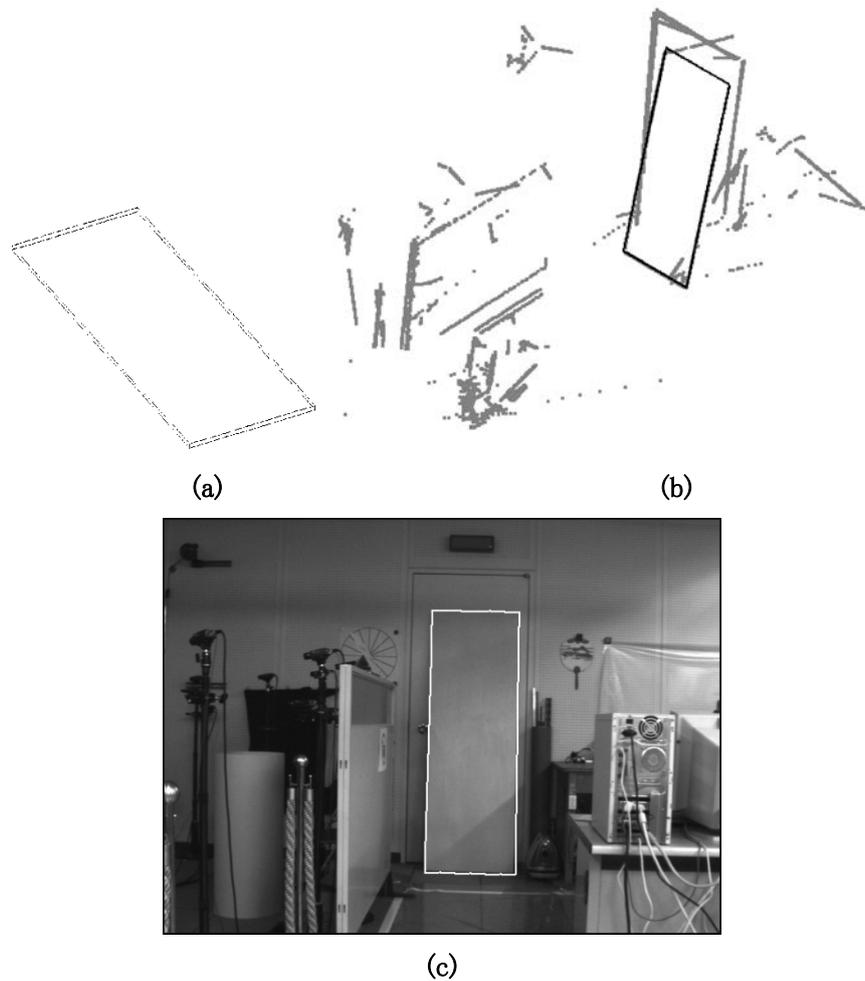


図 A3-10 認識結果((a)対象物体のモデル, (b)認識結果の復元結果への投影, (c)左画像への投影)

A3.5. まとめ

ステレオ視において、計測誤差を軽減するためにセグメントの連結性を利用した平面拘束に基づくアフィン変換による視差補正法を考案し、実験を通してその有効性を確認した。本手法による 3 次元情報の高精度化により、ヒューマノイドロボットなど物理的な制約で基線長を短くせざるをえないステレオカメラシステムの場合でも、精度の良い復元結果を得ることができる。

発表論文

論文誌

1. J. L. Rojas, K. Harada, H. Onda, N. Yamanobe, E. Yoshida, K. Nagata, and Y. Kawai, “Towards Snap Sensing”, *Int. J. Mechatronics and Automation*, Vol. 3, No. 2, pp.69–93, 2013.
2. 永田 和之, 宮坂 崇, 金宮 好和, 山野辺 夏樹, 丸山 健一, 川端 聡, 河井 良浩, “複雑な環境下における指示物体の把持”, *日本機械学会論文集(C 編)*, Vol. 79, No.79, pp.27–42, 2013.
3. 原田 研介, Torea Foissotte, 辻 徳生, 永田 和之, 山野辺 夏樹, 中村 晃, 河井 良浩, “双腕を有するロボットによるピックアンドプレースの動作計画”, *日本機械学会論文集(C 編)*. Vol. 78, No.78, pp.2870–2883, 2012.
4. 原田 研介, 辻 徳生, 永田 和之, 山野辺 夏樹, 丸山 健一, 中村 晃, 河井 良浩, “把持面の柔軟性を考慮したパラレルグリッパの把持計画”, *計測自動制御学会論文集*, Vol. 48, No. 7, pp.264–272, 2012.
5. 渡邊征春, 富田 文明, 丸山 健一, 河井 良浩, 藤村恒太, “ステレオ誤対応に頑健なモデルベース物体位置姿勢計測”, *計測自動制御学会論文集*, Vol. 47, No. 9, pp.356–365, 2011.
6. 川端 聡, 河井 良浩, “参照平面上の局所座標系間の対応推定による複数カメラの較正法”, *電子情報通信学会論文誌*, Vol. J94-D, No. 8, pp.1314–1323, 2011.
7. 吉見 隆, 高瀬 竜一, 河井 良浩, 吉村 剛治, 富田 文明, “主曲率ベクトル分布を用いた中心軸平面を持つ管状曲面物体の位置姿勢推定”, *電子情報通信学会論文誌*, J94-D, No. 5, pp. 893–907, 2011.
8. 丸山 健一, 河井 良浩, 富田 文明, 遮蔽輪郭線を用いた物体モデルに基づく 3 次元物体位置姿勢推定”, *電気学会論文誌*, Vol. 131, No. 4, pp.505–514, 2011.
9. K. Maruyama, R. Takase, Y. Kawai, T. Yoshimi, H. Takahashi, F. Tomita, “Semi-Automated Excavation System for Partially Buried Objects Using Stereo Vision-Based Three-Dimensional Localization”, *ADVANCED ROBOTICS*, Vol. 24, No. 5–6, pp.651–670, 2010.
10. 丸山 健一, 高瀬 竜一, 河井 良浩, 吉見 隆, 高橋裕信, 富田 文明, “非整備環境における部分埋没物体発掘システムのための三次元位置姿勢計測”, *日本ロボット学会誌*, Vol. 28, No. 1, pp.122–132, 2010.
11. 小谷内 範穂, 丸山 健一, 河井 良浩, 富田 文明, “4 輪 4 脚フレキシブルパーソナルロボットの段差跨ぎ超え - 3 眼立体視による段差検出と車輪脚協調歩容 -”, *日本ロボット学会誌*, Vol. 27, No. 8, pp.825–832, 2009.
12. H. Hirukawa, F. Kanehiro, K. Kaneko, S. Kajita, K. Fujiwara, Y. Kawai, F. Tomita, S. Hirai,

- K. Tanie, T. Isozumi, K. Akachi, T. Kawasaki, S. Ota, K. Yokoyama, H. Handa, Y. Fukase, J. Maeda, Y. Nakamura, S. Tachi, and H. Inoue, "Humanoid robotics platforms developed in HRP", *International Journal of Robotics and Autonomous Systems*, Vol. 48, pp.165-175, 2004.
13. 富田 文明, 松下 俊夫, 河井 良浩, "3次元視覚によるプログラミングレスマニピュレーションシステムの開発", *情報処理学会論文誌:コンピュータビジョンとイメージメディア*, Vol. 44, No. SIG 17(CVIM 8), pp.1-12, 2003.
 14. Y. Sumi, Y. Kawai, T. Yoshimi, and F. Tomita, "3D Object Recognition in Cluttered Environments by Segment-based Stereo Vision", *INTERNATIONAL JOURNAL OF COMPUTER VISION*, Vol. 46, No. 1, pp.5-23, 2002.
 15. Y. Kawai and F. Tomita, "A support system for the visually impaired to recognize three-dimensional objects", *IOS Press, Technology and Disability*, Vol. 12, No. 1, pp.13-20, 2000.
 16. 河井 良浩, 小林 真, 皆川 洋喜, 宮川 正弘, 富田 文明, "3次元仮想音響による視覚障害者用支援システム", *電気学会 電子・情報・システム部門誌*, Vol. 120-C, No. 5, pp.648-655, 2000.
 17. 河井 良浩, 植芝 俊夫, 石山 豊, 角 保志, 富田 文明, "セグメントベースドステレオにおける連結性に基づく対応評価", *情報処理学会論文誌*, Vol. 40, No. 8, pp.3219-3229, 1999.
 18. 石山 豊, 角 保志, 河井 良浩, 植芝 俊夫, 富田 文明, "セグメントベースド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. 10, pp.723-728, 1998.
 19. 角 保志, 河井 良浩, 吉見 隆, 富田 文明, "セグメントベースドステレオによる自由曲面体の認識", *電子情報通信学会論文誌*, Vol. J81-D-II, No. 2, pp.285-292, 1998.
 20. 河井 良浩, 富田 文明, "視覚障害者用3次元物体認識支援システム - 対話型3次元触覚ディスプレイシステム -", *映像情報メディア学会誌*, Vol. 51, No. 6, pp.870-877, 1997.
 21. 河井 良浩, 植芝 俊夫, 吉見 隆, 大島 正毅, "多視点レンジデータからの3次元形状復元", *電子情報通信学会論文誌*, Vol. J75-D-II, No. 4, pp.737-748, 1992.
 22. Y. Kawai, N. Onishi, N. Sugie, "A Support System for the Blind to Recognize a Diagram", *Systems and Computers in Japan*, Vol. 21, No. 7, pp.75-85, 1990.
 23. 河井 良浩, 大西 昇, 杉江 昇, "盲人用図面認識支援システム", *電子情報通信学会論文誌*, Vol. J-72-D-II, No. 9, pp.1526-1533, 1989.

国際会議

1. K. Harada, K. Nagata, T. Tsuji, N. Yamanobe, A. Nakamura, and Y.Kawai, “Probabilistic Approach for Object Bin Picking Approximated by Cylinders”, Proc. of ICRA 2013, pp. 3727–3732, 2013.
2. J. L. Rojas, K. Harada, H. Onda, N. Yamanobe, E. Yoshida, K. Nagata, and Y.Kawai, “Gradient Calibration for the RCBHT Snap Verification System”, Proc. of ROBIO 2012, pp. 984–990, 2012.
3. J. L. Rojas, K. Harada, H. Onda, N. Yamanobe, E. Yoshida, K. Nagata, and Y.Kawai, “Probabilistic State Verification for Snap Assemblies using the Relative-Change-Based Hierarchical Taxonomy”, Proc. of Humanoids 2012, pp.96–103, 2012.
4. K. Harada, T. Tsuji, K. Nagata, N. Yamanobe, H. Onda, and Y. Kawai, “Object Placement Planner for Robotic Pick and Place Tasks”, Proc. of IROS 2012, MonCT10.4, 2012.
5. J. L. Rojas, K. Harada, H. Onda, N. Yamanobe, E. Yoshida, K. Nagata, and Y.Kawai, “Relative-Change-Based Hierarchical Taxonomy for Cantilever-Snap Assembly Verification”, Proc. of IROS 2012, MonAT10.2, 2012.
6. J. L. Rojas, K. Harada, H. Onda, N. Yamanobe, E. Yoshida, K. Nagata, and Y.Kawai, “A Constraint-Based Motion Control Strategy for Cantilever Snap Assemblies”, Proc. of ICMA 2012, pp.1815–1821, 2012.
7. K. Harada, T. Foissotte, T. Tsuji, K. Nagata, N. Yamanobe, A. Nakamura, and Y. Kawai, “Pick and Place Planning for Dual-Arm Manipulators”, Proc. of ICRA 2012, pp.2281–2286, 2012.
8. K. Harada, T. Tsuji, K. Nagata, N. Yamanobe, K. Maruyama, A. Nakamura, and Y. Kawai, “Grasp Planning for Parallel Grippers with Flexibility on its Grasping Surface”, Proc. of ROBIO 2011, pp.1540–1546, 2011.
9. A. Nakamura and Y. Kawai, “Recovery Technique from Classified Errors in Adjustment Tasks of Domestic Appliances”, Proc. of AROB 2011, GS18–3, 2011.
10. S. Kawabata and Y. Kawai, “Correspondence-Free Multi Camera Calibration by Observing A Simple Reference Plane”, Proc of. ACCV 2010, Vol. 3, pp.1831–1842, 2010.
11. K. Nagata, T. Miyasaka, D. N. Nenchev, N. Yamanobe, K. Maruyama, S. Kawabata, and Y. Kawai, “Picking up an Indicated Object in a Complex Environment”, Proc. of IROS 2010, pp. 2109–2116, 2010.
12. K. Maruyama, Y. Kawai, and F. Tomita, “3D Contour Model Creation for Stereo-vision Systems”, Proc. of ICPR 2010, pp.1779–1782, 2010.
13. A. Nakamura and Y. Kawai, “Recovery Technique from Classified Errors in Skill-Based Manipulation”, Proc. of AROB 2010, GS9–2, 2010.
14. N. Koyachi, K. Maruyama, Y. Kawai, F. Tomita, and Y. Morikawa, “Step Over Motion of

- Four Wheeled and Four Legged Flexible Personal Robot”, Proc. of ROBIO 2009, pp.616–621, 2009.
15. K. Maruyama, Y. Kawai, and F. Tomita, “Model-based 3D Object Localization Using Occluding Contours”, Proc. of ACCV 2009, MP3-20, 2009.
 16. M. Nakagawa, Y. Kawai, and F. Tomita, “Large-scale stereo for improvement of 3D measurement accuracy in gaze-observation”, Proc. of MVA 2009, pp.58–61, 2009.
 17. M. Nakagawa, F. Tomita, and Y. Kawai, “Development of a Self-adjustment Methodology for a Maintenance-free Stereo Camera”, Proc. of ICIT 2009, pp.353–358, 2009.
 18. M. Nakagawa, F. Tomita, and Y. Kawai, “A Stereo Self-adjustment Methodology for Resuming Active Camera Operation”, Proc. of PSIVT 2009, pp.680–691, 2009.
 19. T. Yoshimi, T. Yoshimura, R. Takase, Y. Kawai, and F. Tomita, “Multi-view Stereo Camera Calibration using Laser Targets for Measurement of Long Objects”, Proc. of IWAIT 2009, 2009.
 20. K. Maruyama, K. Oka, R. Takase, Y. Kawai, T. Yoshimi, H. Takahashi, and F. Tomita, “3D Localization of Partially Buried Object in Unstructured Environment”, Proc. of ICPR 2008, TuBCT8.42, 2008.
 21. K. Maruyama, Y. Kawai, T. Yoshimi, and F. Tomita, “3D Object Localization Based on Occluding Contour Using STL CAD Model”, Proc. of ICPR 2008, TuBCT8.41, 2008.
 22. M. Nakagawa, Y. Kawai, and F. Tomita, “A Stereo Self-adjustment Methodology for Camera Parameter Restoration in Outdoor Environments”, Proc. of DICTA 2008, pp.428–435, 2008.
 23. T. Tsuji, K. Harada, K. Kaneko, F. Kanehiro, and Y. Kawai, “Selecting a Suitable Grasping Motion for Humanoid Robots with Multi-Fingered”, Proc. of Humanoids 2008, pp.54–60, 2008.
 24. N. E. Sian, K. Maruyama, T. Sakaguchi, Y. Kawai, and K. Yokoi, “A Behavior Level Operation System for Humanoid Robots”, Proc. of Humanoids 2006, pp.327–332, 2006.
 25. N. E. Sian, T. Sakaguchi, K. Yokoi, Y. Kawai, and K. Maruyama, “Operating Humanoid Robots in Human Environments”, Proc. of the Robotics, Science and Systems Workshop on for Human Environments, 2006.
 26. K. Yokoi, N. E. Sian, T. Sakaguchi S. Olivier, Y. Kawai, and K. Maruyama, “Humanoid Robot HRP-2 with Human Supervision”, Proc. of ISER 2006, pp.513–522, 2006.
 27. K. Harada, H. Hasunuma, K. Nakashima, Y. Kawai, and H. Hirukawa, “Task Autonomy for a Teleoperated Humanoid Robot”, Proc. of ISER 2006, 2006.
 28. N. E. Sian, S. Olivier, Y. Kawai, T. Sakaguchi, and K. Yokoi, “A Unified On-line Operation Interface for Humanoid Robots in A Partially-Unknown Environment”, Proc. of IEEE Int. Conf. on Intelligent Robots and Systems, pp.4437–4439, 2006.
 29. K. Yokoi, N. E. Sian, T. Sakauchi, H. Arisumi, E. Yoshida, S. Olivier, Y. Kawai, K.

- Maruyama, T. Yoshimi, and S. Kajita, "Humanoid Robot HRP-2 No.10 with Human Supervision", Proc. of IROS 2005, 2005.
30. I. Hara, F. Asano, H. Aso, J. Ogata, N. Ichimura, Y. Kawai, F. Kanehiro, H. Hirukawa, and K. Yamamoto, "Robust Speech Interface Based on Audio and Video Information Fusion for Humanoid HRP-2", Proc. of IROS 2004, pp.2404-2410, 2004.
 31. Y. Fukase, J. Maeda, Y. Kawai, T. Yoshimi, and F. Tomita, "Development of Vision System on Humanoid Robot HRP-2", Proc. 21st International Symposium on Automation and Robotics in Construction, pp.604-609, 2004.
 32. S. Sarata, H. Osumi, Y. Kawai, and F. Tomita, "Trajectory Arrangement based on Resistance Force and Shape of Pile at Scooping Motion", Proc. of ICRA 2004, pp.3488-3493, 2004.
 33. Y. Kawai, F. Tomita, H. Hirukawa, and Y. Fukase, "A Stereo Vision System for the HRP-2 Humanoid Robot to Act Autonomously", Proc. of ACCV 2004, Vol. 2, pp.754-759, 2004.
 34. H. Hirukawa, F. Kanehiro, K. Kaneko, S. Kajita, K. Fujiwara, Y. Kawai, F. Tomita, S. Hirai, and K. Tanie, "Humanoid Robotics Platforms developed in HRP", Proc. of Humanoids 2003, 2003.
 35. H. Hirukawa, K. Kaneko, F. Kanehiro, Y. Kawai, F. Tomita, K. Yokoyama, H. Handa, T. Isozumi, and Y. Fukase, "Cooperative Works by a Human and a Humanoid Robot", Proc. of ICRA 2003, pp.2985-2991, 2003.
 36. K. Yokoi, K. Kaneko, F. Kanehiro, Y. Kawai, F. Tomita, H. Hirukawa, T. Isozumi, K. Yokoyama, H. Handa, and Y. Fukase, "Humanoid Robot HRP-2P", Video Proc. of ICRA 2003, 2003.
 37. T. Yoshimi, Y. Kawai, Y. Fukase, H. Araki, R. Ikeno, Y. Ishiyama, and F. Tomita, "Development of the Vision System for the HRP-2P Humanoid Robot", Proc. of MVA 2002, pp.257-260, 2002.
 38. H. Hirukawa, F. Kanehiro, K. Kaneko, S. Kajita, K. Fujiwara, Y. Kawai, F. Tomita, S. Hirai, K. Tanie, T. Isozumi, K. Akachi, T. Kawasaki, N. Ohta, K. Yokoyama, H. Handa, Y. Fukase, J. Maeda, Y. Nakamura, S. Tachi, and H. Inoue, "Humanoid Robotics Platform", Proc. of 3rd IARP Int. Workshop on Humanoid and Human Friendly Robotics, pp.128-133, 2002.
 39. Y. Kawai and F. Tomita, "A Support System for Visually Impaired Persons to Understand Three-dimensional Visual Information Using Acoustic Interface", Proc. of ICPR 2002, Vol. III, pp.974-977, 2002.
 40. Y. Kawai and F. Tomita, "A Visual Support System for Visually Impaired Persons Using Acoustic Interface - Recognition of 3-D Spatial Information -", Proc. of HCI International 2001, Vol. 1, pp.203-207, 2001.
 41. Y. Kawai and F. Tomita, "A Visual Support System for Visually Impaired Persons Using

- Acoustic Interface”, Proc. of MVA 2000, pp.379–382, 2000.
42. Y. Kawai, M. Kobayashi, H. Minagawa, M. Miyakawa, and F. Tomita, “A Support System for Visually Impaired Persons using Three-Dimensional Virtual Sound”, Proc. of ICCHP 2000, pp.327–334, 2000.
 43. Y. Kawai and F. Tomita, “Intensity calibration for stereo images based on segment correspondence”, Proc. of MVA '98, pp.331–333, 1998.
 44. T. Ueshiba, Y. Kawai, Y. Ishiyama, Y. Sumi, and F. Tomita, “An Efficient Matching Algorithm for Segment-Based Stereo Vision Using Dynamic Programming Technique”, Proc. of MVA '98 , pp.61–64, 1998.
 45. F. Tomita, T. Yoshimi, T. Ueshiba, Y. Kawai, and Y. Sumi, “R&D of Versatile 3D Vision System VVV”, Proc. of SMC '98, pp.4510–4516, 1998.
 46. Y. Kawai and F. Tomita, “EVALUATION OF INTERACTIVE TACTILE DISPLAY SYSTEM”, Proc. of ICCHP '98, pp.29–36, 1998.
 47. Y. Kawai, T. Ueshiba, Y. Ishiyama, Y. Sumi, and F. Tomita, “Stereo Correspondence Using Segment Connectivity”, Proc. of ICPR '98, Vol. 1, pp.648–651, 1998.
 48. Y. Sumi, Y. Kawai, T. Yoshimi, and F. Tomita, “Recognition of 3D Free-Form Objects Using Segment-Based Stereo Vision”, Proc. of ICCV '98, pp.668–674, 1998.
 49. Y. Kawai and F. Tomita, “A Support System for the Visually Disabled to Recognize 3D Objects”, Proc. of MVA '96, pp.325–328, 1996.
 50. T. Yoshimi, Y. Kawai, and F. Tomita, “Range Data Segmentation with Principal Vectors and Surface Types”, Proc. of MVA '96, pp.534–537, 1996.
 51. Y. Kawai and F. Tomita, “Interactive Tactile Display System”, Proc. of ASSETS '96, pp.45–50, 1996.
 52. Y. Kawai, T. Ueshiba, T. Yoshimi, and M. Oshima, “Reconstruction of 3D Objects by Integration of Multiple Range Data”, Proc. of ICPR '92, Vol. I, pp.154–157, 1992.
 53. N. Ohnishi, Y. Kawai, and N. Sugie, “A SUPPORT SYSTEM FOR THE BLIND TO RECOGNIZE A DIAGRAM”, Proc. of the annual Int. Conf. of the IEEE Engineering in Medicine and Biolog Society, pp.1510–1511, 1989.

著書・刊行物

1. 河井 良浩 ほか(共著), “社会まるごとスマートシステム”, カナリヤ書房, 2.2.4, 2013.
2. 河井 良浩, 小笠原 司, “作業知能オープンソースの統合検証”, 日本ロボット学会誌, Vol. 31, No. 1, pp.41-42, 2013.
3. 河井 良浩, “3次元視覚情報処理技術 各種分野への展開”, 電気評論, Vol. 97, No. 10, pp.50-55, 2012.
4. 河井 良浩 ほか(共著), “会議・プレゼンテーションのバリアフリー - だれでも参加”を目指す実践マニュアル -“, 社団法人電子情報通信学会, 4.1.1, 2010.
5. 吉見 隆, 河井 良浩, 富田 文明, “高機能3次元視覚システムVVV”, 映像情報インダストリアル, Vol. 40, No. 11, pp.41-50, 2008.
6. 横井 一仁, 丸山 健一, 河井 良浩, “探査型ヒューマノイドロボットの視覚機能”, O PLUS E, Vol. 29, No. 3, pp.240-243, 2007.
7. 森川 泰, 河井 良浩, 丸山 健一, 富田 文明, “自律型無人ヘリコプタの開発”, 日本ロボット学会誌, Vol.24, No. 8, pp.913, 2006.
8. 河井 良浩, “書評:先天盲開眼者の視覚世界”, 日本ロボット学会誌, Vol. 19, No. 5, pp.48, 2001.
9. 河井 良浩, “視覚支援”, 映像情報メディア学会誌, Vol. 53, No. 1, pp.30-33, 1999.
10. 河井 良浩, 富田 文明, “視覚障害者用 3次元物体認識支援システム -触覚と音声による視覚情報の伝達方法-”, 画像ラボ, Vol. 9, No. 8, 1998.
11. 河井 良浩 ほか(共著), “21世紀における知能システムの姿 -202X年あるテクニカルライターの手記より”, 日本ロボット学会誌, Vol. 12, No. 3, pp.8-11, 1994.

講演会・研究会

1. 金子 健二, 植芝 俊夫, 吉見 隆, 河井 良浩, 森澤 光晴, 金広 文男, 横井 一仁, “ハンド内蔵型レンジセンサの開発”, 第 31 回日本ロボット学会学術講演会, 2013.
2. 西 卓郎, 高瀬 竜一, 吉見 隆, 河井 良浩, 富田 文明, “領域ベースのステレオ対応点探索のための局所的輝度変動に頑強な「標準化係数フィルタ」”, 計測自動制御学会 計測部門 第 90 回パターン計測部会研究会, 2013.
3. 原田 研介, J. Rojas, 音田 弘, 山野辺 夏樹, 吉田 英一, 永田 和之, 河井 良浩, “スナップジョイントを有するパーツの力制御の切り替えに基づく組み立て”, 日本機械学会ロボティクス・メカトロニクス講演会, 2013.
4. 原田 研介, 永田 和之, 辻 徳生, 山野辺 夏樹, 中村 晃, 河井 良浩, “確率的アプローチに基づく円柱近似された対象物のビンピッキング”, 第 18 回ロボティクスシンポジウム, 2013.
5. 河井 良浩, “産総研のロボット技術と今後のものづくりへの取組み”, ロボット技術等活用生産現場完全支援事業講演会, 2013.
6. 河井 良浩, “3次元形状計測・物体検出技術の開発と食品加工機器等への応用”, 北海道産学官プラットフォーム ～HiNT@函館～ 公開セミナー, 2012.
7. 音田 弘, 原田 研介, 吉見 隆, 山野辺 夏樹, 永田 和之, 中村 晃, 北垣 高成, 河井 良浩, “ロボットによる詰込のための仕様・計画・配置技能”, 第 30 回日本ロボット学会学術講演会, 4G3-7, 2012.
8. 阪口 健, 花井 亮, 原田 研介, 中岡 慎一郎, 河井 良浩, 松本 吉央, “移動作業用ロボットプラットフォーム “MobileHIRO” の操作用統合環境の開発”, 第 30 回日本ロボット学会学術講演会, 3J2-3, 2012.
9. 原田 研介, 永田 和之, 辻 徳生, 山野辺 夏樹, 中村 晃, 河井 良浩, “対象物の円筒近似に基づくビンピッキング”, 第 30 回日本ロボット学会学術講演会, 2O3-1, 2012.
10. 原田 研介, 辻 徳生, 永田 和之, 山野辺 夏樹, 音田 弘, 吉見 隆, 河井 良浩, “ロボットハンドによる Object Placement Planner”, 日本機械学会ロボティクス・メカトロニクス講演会, 2012.
11. 原田 研介, T. Foissotte, 辻 徳生, 永田 和之, 山野辺 夏樹, 丸山 健一, 河井 良浩, “双腕ロボットによるピックアンドブレース動作計画”, 第 17 回ロボティクスシンポジウム, 2012.
12. 横山 和彦, 原田 研介, 川端 聡, 丸山 健一, 吉見 隆, 高瀬 竜一, 野崎 俊輔, J. L. Rojas, 河井 良浩, 足立 勝, 大橋 健, 長谷川 勉, 松坂 要佐, 辻 徳生, 浅間 一, 近野 敦, 山口 亨, 小川 秀樹, “作業知能(社会・生活分野)の開発”, 「次世代ロボット知能化技術開発プロジェクト」成果報告会, 2012.
13. 河井 良浩, 原田 研介, 川端 聡, 丸山 健一, 吉見 隆, 高瀬 竜一, 野崎 俊輔, J. L. Rojas, 花井 亮, 三浦 純, 近野 敦, “双腕ロボットプラットフォームを用いた作業知能オープンソースの統合検証 パレタイジング作業”, 「次世代ロボット知能化技術開発プロジェクト」

成果報告会, 2012.

14. 山野辺 夏樹, 永田 和之, 原田 研介, 中村 晃, 河井 良浩, 久場 景太郎, 新井 民夫, “受け手の心地よさを考慮した持ち替えも含む物体の把持動作戦略”, 第 12 回計測自動制御学会・システムインテグレーション部門講演会, 2011.
15. 大原 賢一, 川端 聡, 河井 良浩, “共通カメラインタフェースの提案”, 第 29 回日本ロボット学会学術講演会, 2011.
16. 山野辺 夏樹, 久場 景太郎, 原 辰徳, 浅間 一, 新井 民夫, 永田 和之, 原田 研介, 中村 晃, 河井 良浩, 辻 徳生, “手渡しの際の受け手の心地よさを考慮した日用品の把持”, 第 29 回日本ロボット学会学術講演会, 2011.
17. 原田 研介, 辻 徳生, 永田 和之, 山野辺 夏樹, 中村 晃, 河井 良浩, “双腕ロボットによる持ち替えを伴うピックアンドプレース”, 第 29 回日本ロボット学会学術講演会, 2011.
18. 原田 研介, 音田 弘, 山野辺 夏樹, 吉田 英一, 辻 徳生, 永田 和之, 河井 良浩, “複雑な形状の部品に対する接触状態遷移グラフ”, 日本機械学会ロボティクス・メカトロニクス講演会, 2011.
19. 原田 研介, 辻 徳生, 永田 和之, 山野辺 夏樹, 丸山 健一, 河井 良浩, “把握面の柔軟性を考慮したパラレルグリッパの把握計画と検証”, 第 16 回ロボティクスシンポジウム, 2011.
20. 二宮 恒樹, 小笠原 哲也, 田中 秀幸, 松坂 要佐, 松本 吉央, 原 功, 河井 良浩, 平井 成興, “RT コンポーネント再利用性向上の研究(第 2 報)”, 計測自動制御学会システムインテグレーション部門講演会, 2010.
21. 辻 徳生, 丸山 健一, 河井 良浩, 原田 研介, “2 次元 3 次元再帰計算エッジマッチングによる物体位置・姿勢推定の高速化”, 計測自動制御学会システムインテグレーション部門講演会, 2010.
22. 小谷内 範穂, 丸山 健一, 森川 泰, 河井 良浩, 富田 文明, “4 輪 4 脚フレキシブルパーソナルロボットの段差・階段乗り越え 一段差勾配への適応と摺足遊脚”, 第 28 回日本ロボット学会学術講演会, 2010.
23. 丸山 健一, 川端 聡, 河井 良浩, 富田 文明, “ステレオビジョンシステムを用いた 3 次元物体位置姿勢計測と同一形状物体の計数”, 第 28 回日本ロボット学会学術講演会, 2010.
24. 川端 聡, 丸山 健一, 河井 良浩, “既知の平面パターンを用いたステレオカメラの簡便な較正法”, 第 28 回日本ロボット学会学術講演会, 2010.
25. 原田 研介, 辻 徳生, 永田 和之, 山野辺 夏樹, 中村 晃, 河井 良浩, “把握面に柔軟性を有するパラレルグリッパの把握計画”, 第 28 回日本ロボット学会学術講演会, 2010.
26. 川端 聡, 河井 良浩, “単純な繰り返しパターンの参照平面を用いた複数カメラの較正法”, MIRU 2010, 2010.
27. 丸山 健一, 河井 良浩, 富田 文明, “ステレオビジョンシステムのための 3 次元輪郭モデル生成とその応用”, MIRU 2010, 2010.
28. 永田 和之, 宮坂 崇, 金宮 好和, 山野辺 夏樹, 丸山 健一, 川端 聡, 河井 良浩, “日

常生活環境下での指示物体の把持計画”，日本機械学会ロボティクス・メカトロニクス講演会'10, 2010.

29. 河井 良浩, “高機能 3 次元 3 次元視覚システム VVV の研究開発 — ステレオ画像処理による 3 次元形状計測, 認識とロボットへの応用 —”, 広島県画像処理活用研究会 & 中国地域産総研技術セミナー, 2010.
30. 尹 祐根, 原 功, 中坊 嘉宏, 朴 雲埴, 山野辺 夏樹, 脇田 優仁, 永田 和之, 丸山 健一, 小野 栄一, 神徳 徹雄, 河井 良浩, 末廣 尚士, 山田 陽滋, 北垣 高成, 比留川 博久, “対人サービスロボットの開発”, 計測自動制御学会 システムインテグレーション部門講演会, 2009.
31. 宮坂 崇, 永田 和之, 山野辺 夏樹, 川端 聡, 丸山 健一, 河井 良浩, 金宮 好和, “日常生活環境における人が指示した物体の把持システム”, 計測自動制御学会 システムインテグレーション部門講演会, 2009.
32. 吉見 隆, 高瀬 竜一, 河井 良浩, 富田 文明, “主曲率ベクトル分布を用いた管状曲面物体の位置姿勢推定”, パターン認識・メディア理解研究会, 2009.
33. 丸山 健一, 川端 聡, 河井 良浩, 富田 文明, “ステレオ視による 3 次元物体位置姿勢計測とその RT コンポーネント化”, 第 27 回日本ロボット学会学術講演会, 2009.
34. 川端 聡, 丸山 健一, 河井 良浩, “マーカ 1 点の複数回撮影によるハンドアイシステムの簡便な較正法とその RT コンポーネント化”, 第 27 回日本ロボット学会学術講演会, 2009.
35. 河井 良浩, 丸山 健一, 川端 聡, “施設内生活支援ロボット知能の研究開発 — 作業対象物認識に関する知能モジュール群の開発 (第 2 報) —”, 第 27 回日本ロボット学会学術講演会, 2009.
36. 西村 悠, 吉見 隆, 西 卓郎, 河井 良浩, 富田 文明, “曲面形状表現のための曲率線ネットの生成”, MIRU 2009, 2009.
37. 丸山 健一, 河井 良浩, 富田 文明, “遮蔽輪郭線を用いたモデルベース 3 次元物体位置姿勢計測”, MIRU 2009, 2009.
38. 川端 聡, 永田 和之, 河井 良浩, “環境に固定されたマーカを用いたハンドアイキャリブレーション”, MIRU 2009, 2009.
39. 中川 雅史, 河井 良浩, 富田 文明, “3 次元計測精度向上のための注視観測による Large-scale stereo”, 日本写真測量学会平成 21 年度年次学術講演会, 2009.
40. 森川 泰, 中川 雅史, 河井 良浩, 富田 文明, “ステレオカメラを搭載した無人ヘリコプタによる 3 次元計測”, ロボティクス・メカトロニクス講演会 2009, 2009.
41. 安達 栄輔, 吉見 隆, 河井 良浩, 富田 文明, “境界表現に基づく複数観測点からのステレオデータの統合”, TXテクノロジー・ショーケース in つくば 2009, 2009.
42. 安達 栄輔, 吉見 隆, 河井 良浩, 富田 文明, “境界表現に基づく複数観測点からのステレオデータの統合”, 電子情報通信学会パターン認識・メディア理解研究会, 2009.
43. 原田 研介, 辻 徳生, 金子 健二, 丸山 健一, 河井 良浩, 富田 文明, “視覚情報に基

- づく多指ハンドの把握計画”，第 26 回日本ロボット学会学術講演会，2008.
44. 小谷内 範穂，丸山 健一，河井 良浩，富田 文明，“4 輪 4 脚フレキシブルパーソナルロボットの段差跨ぎ越え —三眼立体視による段差検出と車輪脚協調歩容—”，第 26 回日本ロボット学会学術講演会，2008.
 45. 丸山 健一，岡 浩司，藤原 一正，高瀬 竜一，河井 良浩，吉見 隆，高橋 裕信，富田 文明，“ステレオ視を用いた埋没物体発掘システムの開発”，第 26 回日本ロボット学会学術講演会，2008.
 46. 丸山 健一，富田 文明，河井 良浩，“施設内生活支援ロボット知能の研究開発 —作業対象物認識に関する知能モジュール群の開発—”，第 26 回日本ロボット学会学術講演会，2008.
 47. 丸山 健一，河井 良浩，吉見 隆，富田 文明，“STL CAD モデルを用いた遮蔽輪郭線による任意形状物体認識”，MIRU 2008, pp.1626-1631, 2008.
 48. 中川 雅史，富田 文明，河井 良浩，“相関法に基づくステレオカメラ・セルフアジャストメント”，日本写真測量学会 平成 20 年度年次学術講演会，2008.
 49. 丸山 健一，高瀬 竜一，河井 良浩，吉見 隆，富田 文明，“非整備環境での 3 次元物体認識に関する研究”，TXテクノロジー・ショーケース・イン・ツクバ 2008, 2008.
 50. 松井 公志，中川 雅史，高瀬 竜一，吉見 隆，河井 良浩，富田 文明，“アクティブステレオによる注視制御”，日本写真測量学会 平成 19 年度年次学術講演会，2007.
 51. 本橋 英明，森川 泰，河井 良浩，丸山 健一，富田 文明，柿倉 正義，“小型無人ヘリコプタのホバリング制御”，第 49 回自動制御連合講演会，2006.
 52. 森川 泰，河井 良浩，丸山 健一，富田 文明，本橋 英明，柿倉 正義，“ステレオビジョンを搭載した自律型無人ヘリコプタ”，第 44 回飛行機シンポジウム，2006.
 53. 阪口 健，N. E. Sian，丸山 健一，河井 良浩，横井 一仁，“ステレオ相関法による床上の未知物体の把持点検出と拾い上げ動作”，第 29 回日本ロボット学会学術講演会，2006.
 54. N. E. Sian，丸山 健一，河井 良浩，阪口 健，横井 一仁，“ヒューマノイドロボットの行動レベル操作システム”，第 29 回日本ロボット学会学術講演会，2006.
 55. 丸山 健一，高橋 裕信，高瀬 竜一，河井 良浩，吉見 隆，富田 文明，“不整地環境での 3 次元物体認識システム”，ViEW2005 ビジョン技術の実利用ワークショップ，2005.
 56. 不破 洋平，竹内 義則，大西 昇，角 保志，河井 良浩，富田 文明，“観測シーンに依存しない 3 次元物体認識”，PRMU研究会，2005.
 57. 横井 一仁，N. E. Sian，阪口 健，有隅 仁，吉田 英一，河井 良浩，丸山 健一，吉見 隆，S. Olivier，梶田 秀司，“探査型ヒューマノイドロボット HRP-2 No.10”，第 23 回日本ロボット学会学術講演会，2005.
 58. 丸山 健一，河井 良浩，吉見 隆，高橋 裕信，富田 文明，“非整備環境からの回転体の発見と 3 次元位置計測”，MIRU 2005, 2005.
 59. 横井 一仁，N. E. Sian，有隅 仁，吉田 英一，S. Olivier，河井 良浩，梶田 秀司，金広

- 文男, 原田 研介, “探査型ヒューマノイドロボット HRP-2 No.10”, 日本機械学会ロボティクス・メカトロニクス部門, 2005.
60. 河井 良浩, 坂井 忠裕, “誰でも参加できる学会を目指して – 視覚障害関連情報保障”, 電子情報通信学会 総合大会, HK-1-2, 2005.
 61. N. E. Sian, 横井 一仁, 梶田 秀司, 河井 良浩, 谷江 和雄, “ヒューマノイドロボットの自律・遠隔融合操作”, 第 10 回ロボティクスシンポジウム, 2005.
 62. 不破 洋平, 竹内 義則, 大西 昇, 河井 良浩, 富田 文明, “ステレオ視を使った 3次元環境地図の生成”, 平成 16 年度 電気関係学会東海支部 連合大会, O-453, 2004.
 63. 小谷内 範穂, 富田 文明, 吉見 隆, 河井 良浩, “4 輪 4 脚フレキシブルパーソナルロボットの設計”, 第 22 回日本ロボット学会学術講演会, 3D29, 2004.
 64. 金広 文男, 河井 良浩, 富田 文明, 森澤 光晴, 藤原 清司, 原田 研介, 金子 健二, 梶田 秀司, 比留川 博久, “視覚に基づくヒューマノイドの複雑環境自律移動 – ローカルマップを用いた移動様式の種類 –”, 第 22 回日本ロボット学会学術講演会, 1K18, 2004.
 65. 安念 克洋, 萩川 友宏, 河井 良浩, 岡本 明, 坂根 裕, 竹林 洋一, “点訳機能を有する講演原稿受付システムによる手頃な点字資料作成方式の検討”, FIT 2004, K-083, pp.583-584, 2004.
 66. 深瀬 勇太郎, 前田 純一郎, 河井 良浩, 吉見 隆, 富田 文明, “人間型ロボット視覚の簡易建屋モデル組立て作業への応用”, 第 10 回建築ロボットシンポジウム論文集, pp.63-70, 2004.
 67. 森川 泰, 河井 良浩, 小森谷 清, 富田 文明, “無人ヘリコプタによる三次元画像撮影”, ロボティクス・メカトロニクス講演会 2004, 1P1-H-41, 2004.
 68. 堀口 彰, 坪内 孝司, 皿田 滋, 河井 良浩, 富田 文明, “破碎岩石堆積のステレオ視による 3次元形状計測”, 資源・素材学会 2004 年春季大会, Vol. I, pp.71-72, 2004.
 69. 大上 靖弘, 藤村 恒太, 河井 良浩, 富田 文明, “等輝度線ステレオによる顔の三次元復元”, パターン認識・メディア理解研究会 信学技法, PRMU2003-279, 2004.
 70. 深瀬 勇太郎, 前田 純一郎, 河井 良浩, 吉見 隆, 富田 文明, “ヒューマノイドによる簡易建屋建築作業におけるロボットビジョンの応用”, 第 15 回建築施工ロボットシンポジウム, pp.3-8, 2004.
 71. 横山 和彦, 安田 賢一, 深瀬 勇太郎, 赤地 一彦, 五十棲 隆勝, 富田 文明, 吉見 隆, 河井 良浩, 金広 文男, 金子 健二, 比留川 博久, “ヒューマノイドの屋外共同作業への応用”, 第 21 回日本ロボット学会学術講演会, 3A35, 2003.
 72. 河井 良浩, 富田 文明, “新世代視覚技術の多様な実用化に向けて”, 平成 14 年度 NEDO 先端技術講座, 2003.
 73. 比留川 博久, 河井 良浩, 富田 文明, “ヒューマノイドロボティクスプロジェクト”, 情報処理学会研究会資料, CVIM-135-1, 2002.
 74. 深瀬 勇太郎, 河井 良浩, 池野 良平, 石山 豊, 富田 文明, “HRP-2P 搭載視覚システ

- ムの開発”, 第 20 回日本ロボット学会学術講演会, 1D31, 2002.
75. 横山 和彦, 安田 賢一, 深瀬 勇太郎, 赤地 一彦, 五十棲 隆勝, 富田 文明, 吉見 隆, 河井 良浩, 金広 文男, 金子 健二, 比留川 博久, “ヒューマノイドの屋外共同作業への応用”, 第 20 回日本ロボット学会学術講演会, 1D31, 2002.
 76. 深瀬 勇太郎, 河井 良浩, 富田 文明, 池野 良平, 石山 豊, “HRP-2P 搭載視覚システムの開発”, 第 20 回日本ロボット学会学術講演会, 1D34, 2002.
 77. 河井 良浩, “高機能 3 次元視覚システム VVV について”, 財団法人ソフトピアジャパン, 2002.
 78. 河井 良浩, 深瀬 勇太郎, 池野 良平, 石山 豊, 富田 文明, “自律行動型ヒューマノイド HRP-2P 用視覚システムの開発”, FIT 2002, I-40, 2002.
 79. 佐谷野 健二, 富田 文明, 河井 良浩, 角 保志, 吉見 隆, 植芝 俊夫, 戸田 賢二, “FPGA を利用した 3 次元画像処理の高速化の検討”, 電子情報通信学会 DC 研究会資料, DC-2002-19, 2002.
 80. 河井 良浩, 富田 文明, “ステレオ視における 3 次元復元の高精度化”, MIRU 2002, Vol. 1, pp.159-164, 2002.
 81. 石黒 正揮, 浦井 哲哉, 西山 洋一, 富田 文明, 河井 良浩, 角 保志, 植芝 俊夫, 長島 一雄, 土原 丈友, “3 次元画像認識処理タスクの生成・管理プラットフォームの研究開発”, IPA 2001 年度 先進的情報処理技術開発促進事業 報告書, 2002.
 82. 横山 和彦, 安田 賢一, 深瀬 勇太郎, 河井 良浩, 赤地 一彦, 五十棲 隆勝, 金広 文男, 金子 健二, 比留川 博久, “屋外共同作業への応用分野”, 日本ロボット学会創立 20 周年記念学術講演会, 2002.
 83. 皿田 滋, 河井 良浩, 角 保志, 富田 文明, 片石 一永, 坪内 孝司, “画像による堆積物形状計測について”, 資源・素材 2001, 2001.
 84. 深瀬 勇太郎, 池野 良平, 石山 豊, 河井 良浩, 吉見 隆, 富田 文明, 竹内 啓五, 青木 滋, “ヒューマノイド用ステレオカメラシステムと距離計測性能評価”, 第 19 回日本ロボット学会学術講演会, 2001.
 85. 河井 良浩, 小林 真, 皆川 洋樹, 宮川 正弘, “ICCHP 2000 国際会議報告”, 電子情報通信学会 WIT 研究会資料, WIT00-25, 2000.
 86. 角 保志, 河井 良浩, 吉見 隆, 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物体認識”, 電気学会システム・制御研究会資料, SC00-15, 2000.
 87. 河井 良浩, 小林 真, 皆川 洋樹, 宮川 正弘, 富田 文明, “仮想音響による 3 次元空間認識支援システム”, 電子情報通信学会 WIT 研究会資料, WIT99-26, 2000.
 88. 河井 良浩, 小林 真, 皆川洋樹, 宮川 正弘, “視覚障害者のための 3 次元音響インタフェース”, 平成 11 年度生命工学研究総合推進会議・ニューバイオ技術検討会合同研究発表会, 2000.
 89. 宮川 正弘, 河井 良浩, “仮想音響によるナビゲーションの実現 — 視覚障害者の歩行支

- 援への期待—”, 第17回日本ロボット学会学術講演会, 特別セッション 別冊資料「福祉の現場の声とロボット技術」, pp.57, 1999.
90. 河井 良浩, 小林 真, 皆川 洋樹, 宮川 正弘, “視覚障害者用 3 次元音響システム”, 1999 年電子情報通信学会総合大会, Vol. 1, pp.498-499, 1999.
 91. 小林 真, 皆川 洋樹, 宮川 正弘, 河井 良浩, “3 次元音響による視覚代行システム”, 第 24 回感覚代行シンポジウム, pp.101-104, 1998.
 92. 河井 良浩, 富田 文明, “ステレオ顕微鏡による3次元計測”, 平成 10 年度マイクロマシン総合研究会, 1998.
 93. 河井 良浩, 富田 文明, “セグメントの対応に基づくステレオ画像の輝度補正法”, 第 57 回情報処理学会全国大会, 3E-3, 1998.
 94. 角 保志, 河井 良浩, 石山 豊, 富田 文明, “ステレオビジョンを用いた複数 3 次元データの統合と物体モデルの生成”, 第 57 回情報処理学会全国大会, 3E-8, 1998.
 95. 河井 良浩, 富田 文明, “ステレオ顕微鏡による 3 次元形状計測”, 平成 10 年度機械・航空・宇宙総合推進会議講演会, 1998.
 96. 河井 良浩, “電総研のコンピュータビジョングループにおける研究の紹介”, 理化学研究所 BMC フォーラム, 1998.
 97. 河井 良浩, “対話型触覚ディスプレイの評価実験について”, 平成 9 年度生命工学研究総合推進会議シンポジウム, 1998.
 98. 富田 文明, 吉見 隆, 植芝 俊夫, 河井 良浩, 角 保志, 松下 俊夫, 市村 直樹, 杉本 和英, 石山 豊, “3 次元視覚システム VVV 研究開発 一概要—”, 情報処理学会研究会資料, CVIM-109-1, 1998.
 99. 河井 良浩, 植芝 俊夫, 富田 文明, “ステレオ顕微鏡”, 情報処理学会第 55 回全国大会, 1997.
 100. 河井 良浩, 富田 文明, “触覚と音声による視覚障害者向け三次元認識支援システム”, 電気学会 電子・情報・システム部門大会, 1997.
 101. 河井 良浩, 植芝 俊夫, 石山, 角 保志, 富田 文明, “セグメントベースドステレオにおける連結性と対応評価”, 電子情報通信学会 PRMU 研究会, 1997.
 102. 角 保志, 河井 良浩, 富田 文明, “セグメントベースドステレオによる自由曲面体の認識”, 電子情報通信学会 PRMU 研究会, 1997.
 103. 植芝 俊夫, 河井 良浩, 角 保志, 石山 豊, 富田 文明, “セグメントベースドステレオにおける対応パスの探索”, 電子情報通信学会 PRMU 研究会, 1997.
 104. 石山 豊, 角 保志, 河井 良浩, 植芝 俊夫, 富田 文明, “セグメントベースドステレオにおける対応候補探索”, 電子情報通信学会 PRMU 研究会, 1997.
 105. 松下 俊夫, 角 保志, 石山 豊, 河井 良浩, 植芝 俊夫, 富田 文明, “トラッキング機能を持つステレオビジョンを用いた作業の誘導とモニタ”, 日本ロボット学会学術講演会, 1996.
 106. 河井 良浩, 吉見 隆, 富田 文明, “ステレオレンジファインダ “, 情報処理学会第

- 53回全国大会, 1996.
107. 角 保志, 河井 良浩, 富田 文明, “セグメントベースドステレオによる自由曲面体の認識”, 情報処理学会全国大会, 1996.
 108. 吉見 隆, 河井 良浩, 富田 文明, “主ベクトルと曲面の型によるレンジデータのセグメンテーション”, 情報処理学会全国大会, 1996.
 109. 高城 英誌, 河井 良浩, 石山 豊, 富田 文明, “微分ベクトルに基づくシェイピング領域の抽出”, 情報処理学会全国大会, 1996.
 110. 高城 英誌, 角 保志, 河井 良浩, 石山 豊, 富田 文明, “プリミティブ認識によるカメラキャリブレーション”, 電子情報通信学会全国大会, 1996.
 111. 植芝 俊夫, 石山 豊, 河井 良浩, 角 保志, 富田 文明, “セグメントベースドステレオにおける対応パスの効率的探索”, 電子情報通信学会全国大会, 1996.
 112. 河井 良浩, 植芝 俊夫, 石山 豊, 角 保志, 富田 文明, “セグメントの連結性に基づくステレオ対応探索”, 情報処理学会全国大会, 1995.
 113. 河井 良浩, 富田 文明, “対話型触覚ディスプレイの開発”, 情報処理学会全国大会, 1995.
 114. 河井 良浩, 富田 文明, “対話型触覚ディスプレイの開発”, 生命工学研究総合推進会議合同研究発表会, 1995.
 115. 河井 良浩, 石山 豊, 植芝 俊夫, 富田 文明, “境界線と等輝度線のステレオ視による三次元復元”, 情報処理学会全国大会, 1994.
 116. 角 保志, 石山 豊, 植芝 俊夫, 河井 良浩, 杉本, 富田 文明, “画像の境界表現のデータ構造とインターフェース”, 情報処理学会全国大会, 1994.
 117. 植芝 俊夫, 河井 良浩, 角 保志, 杉本, 富田 文明, “境界線セグメントの拘束伝播によるステレオ対応法”, 情報処理学会全国大会, 1994.
 118. 河井 良浩, 石山 豊, 植芝 俊夫, 角 保志, 高橋 裕信, 富田 文明, “ステレオカメラシステム - パタパタ -”, MIRU '94 , 1994.
 119. 河井 良浩, 富田 文明, “傾斜ループに基づく曲面の分割”, 電子情報通信学会全国大会, 1994.
 120. 河井 良浩, 石山 豊, 植芝 俊夫, 高橋 裕信, 富田 文明, “ステレオカメラシステム - パタパタ - について”, 情報処理学会全国大会, 1994.
 121. 河井 良浩, 石山 豊, 富田 文明, “等輝度線からエッジのステレオ対応”, 電子情報通信学会全国大会, 1992.
 122. 河井 良浩, 植芝 俊夫, 吉見 隆, 大島 正毅, “多視点レンジデータからの3次元形状復元”, 第6回産業における画像センシング技術シンポジウム, 1991.
 123. 河井 良浩, 植芝 俊夫, 吉見 隆, 大島 正毅, “多視点レンジデータからの3次元形状復元”, 電子情報通信学会 PRU 研究会, 1991.
 124. 河井 良浩, 吉見 隆, 大島 正毅, “多視点レンジデータの統合”, 情報処理学会全国大会,

1991.

125. 河井 良浩, 大島 正毅, “レンジデータによる3次元モデル構築”, 電子情報通信学会全国大会, 1990.
126. 河井 良浩, 大西 昇, 杉江 昇, “盲人用図面認識支援システム”, 第4回 生体・生理工学シンポジウム, 1989.
127. 河井 良浩, 大西 昇, 杉江 昇, “盲人用図面認識支援システム”, 電子情報通信学会 ME とバイオサイバネティクス研究会資料, MBE88-118, 1989.
128. 河井 良浩, 大西 昇, 杉江 昇, “視覚障害者用図形読取システム”, 電気学会 計測研究会資料, IM-88-34, 1988.
129. 河井 良浩, 大西 昇, 杉江 昇, “盲人用図面認識システムの開発”, 電気学会 システム・制御研究会資料, SC-88-7, 1989.

特許

1. 吉見 隆, 高瀬 竜一, 河井 良浩, 富田 文明, “ステレオカメラ配置方法とそのシステム”, 特開 2012-211833, 2011/3/31.
2. 高瀬 竜一, 富田 文明, 河井 良浩, 吉見 隆, 西 卓郎, “ステレオ画像処理装置及びステレオ画像処理方法”, 特開 2012-181142, 2011/3/2.
3. 吉見 隆, 河井 良浩, 高瀬 竜一, “キャリブレーション用校正治具, 校正治具を備えた3次元計測システム”, 特開 2012-112911, 2010/11/29.
4. 蔵田 武志, 興梠 正克, 石川 智也, 河井 良浩, 丸山 健一, 多葉井 宏, 染谷 俊介, “検査支援システムおよびプログラム”, 特開 2012-021323, 2010/7/14.
5. 丸山 健一, 河井 良浩, 富田 文明, “3次元物体位置姿勢計測方法”, 特許 5030183, 2009/7/17.
6. 吉見 隆, 高瀬 竜一, 河井 良浩, 富田 文明, 石山 英二, “距離測定装置および方法並びにプログラム”, 特許 5071865, 2008/6/3.
7. 吉見 隆, 高瀬 竜一, 河井 良浩, 富田 文明, 石山 英二, “距離測定装置および方法並びにプログラム”, 特許 5071866, 2008/6/3.
8. 河井 良浩, 富田 文明, “3次元復元方法及び装置”, 特許 3870255, 2002/2/25.
9. 河井 良浩, 富田 文明, “ステレオ画像の輝度修正方法及び装置”, 特許 2958462, 1998/10/1.
10. 角 保志, 河井 良浩, 富田 文明, “3次元物体モデル生成方法及び3次元物体モデル生成プログラムを記録したコンピュータ読み取り可能な記録媒体”, 特許 2961264, 1998/9/21.
11. 角 保志, 河井 良浩, 富田 文明, “3次元物体位置姿勢決定方法”, 特許 2920513, 1997/1/21.
12. 河井 良浩, 富田 文明, 高城 英誌, 石山 豊, “画像解析方法”, 特許 3051936, 1996/9/4.
13. 河井 良浩, 大島 正毅, “多視点距離データの統合方法”, 特許 2138014, 1991/3/11.

プレス発表

1. 知能ロボット開発のための知能ソフトウェアモジュール群, 産総研, 2012/2/23.
2. 生活支援ロボットの実現を目指し共同研究開発, 安川電機, 産総研ほか, 2012/1/27.
3. 日本初のロボット用知能ソフトモジュールを公開, NEDO, 2011/7/27.
4. 農水産物加工向けの全周3次元形状計測システム, 産総研, ニッコー, 2011/6/6.
5. 安全性を考慮した生活支援用ロボットアーム, 産総研, 2009/9/28.
6. 生活活動支援ヒューマノイドロボットプラットフォームを開発, 産総研, 2006/1/23.
7. 人間協調・共存型ロボットシステムプロジェクト最終成果発表, 産総研, 2003/2/26.