

論文 / 著書情報
Article / Book Information

題目(和文)	映像メディアからの道路標識抽出法に関する研究
Title(English)	
著者(和文)	三輪道雄
Author(English)	
出典(和文)	学位:不明, 学位授与機関:東京工業大学, 報告番号:甲第4467号, 授与年月日:2000年3月26日, 学位の種別:課程博士, 審査員:佐藤 誠
Citation(English)	Degree:, Conferring organization: Tokyo Institute of Technology, Report number:甲第4467号, Conferred date:2000/3/26, Degree Type:Course doctor, Examiner:
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

映像メディアからの
道路標識抽出法に関する研究

平成11年度

知能システム科学専攻

三輪道雄

目 次

1 序論	8
1.1 研究の背景と目的	8
1.2 脳における映像処理と映像メディア解析	9
1.3 本研究の産業分野への応用	10
1.4 本論文の構成	13
2 従来の研究と解決すべき研究課題	17
2.1 まえがき	17
2.2 映像メディア解析に関する従来の研究	18
2.2.1 動画像の構造解析	18
2.2.2 色情報の解析	18
2.2.3 奥行き情報の解析	20
2.2.3.1 ステレオ立体視	20
2.2.3.2 光切断法	22
2.2.3.3 合焦法による奥行き情報の解析	22
2.2.4 テクスチャ解析	26
2.2.5 道路標識の認識	26
2.3 本研究で解決すべき課題	27
2.4 まとめ	27
3 合焦法を用いた遠距離物体の検出	30
3.1 まえがき	30
3.2 合焦法の課題	31
3.2.1 カメラモデル	31
3.2.2 多重フォーカス画像	32
3.2.3 距離計測法	33
3.2.4 物体深度	33
3.2.5 遠距離計測における問題点	34
3.3 スクリーン投影形撮像系	35
3.3.1 スクリーン投影形撮像系の提案	36
3.3.2 提案観測系と従来方式の比較	38
3.4 一定距離範囲内に存在する物体検出	39

3.4.1	焦点を変えて撮影した3枚の画像による物体検出	39
3.4.2	スクリーン投影形撮像系による物体検出	41
3.5	DCTを用いた合焦判定法	43
3.5.1	DCT係数と焦点ぼけの関係	44
3.5.2	合焦判定アルゴリズム	46
3.6	試作実験	49
3.6.1	スクリーン投影形撮像系の試作	49
3.6.2	遠距離計測実験	50
3.6.3	DCTを用いた物体認識実験	54
3.6.4	離散フーリエ変換との比較	56
3.7	まとめ	59
4	多元情報を用いた映像メディアからの道路標識の抽出	63
4.1	まえがき	63
4.2	多元情報による標識抽出法	64
4.2.1	標識の特徴	64
4.2.2	標識抽出の課題とシステム構成	65
4.3	色フィルタの設計	65
4.4	距離情報による背景画像との識別	68
4.5	検出ウインドウの設計	72
4.6	標識抽出に適したテクスチャ解析法の選択	77
4.7	道路標識の抽出実験	79
4.7.1	実験の概要	79
4.7.2	標識領域抽出アルゴリズム	80
4.7.3	実験結果および考察	82
4.7.4	距離情報の道路標識抽出への適用	84
4.7.5	他の手法との比較	85
4.7.5.1	単独の手法との比較	85
4.7.5.2	シーン内の文字列抽出法との比較	87
4.8	道路管理用画像データベースへの適用	88
4.9	まとめ	89
5	結論	92
5.1	本研究で得られた成果	92

5.2 今後の研究課題	93
謝辞	94
参考文献	95
発表論文	101

図 目 次

1	脳の受容野の位置	9
2	視覚大脳皮質の局所並列階層モデル	10
3	従来の道路管理方法	11
4	将来の道路管理方法	12
5	将来の映像メディアの利用方法	14
6	論文の流れ	16
7	本研究における映像メディアの取得の様子	17
8	ビデオの構造	19
9	時空間画像の切断面	20
10	円筒物体表面上での色分布	21
11	R G B空間でのカラーヒストグラム	21
12	ステレオ視	22
13	光切断法	23
14	D F Dによる距離測定	24
15	多元情報を用いた映像のコンテンツ解析	28
16	カメラモデル	32
17	物体深度	33
18	受光面サイズと画角	34
19	C C Dカメラで撮影する場合のFナンバーと物体深度の関係	35
20	焦点距離と受光面サイズの変化	36
21	スクリーン投影形撮像系	37
22	提案観測系における受光面サイズの変化	38
23	カメラモデル	40
24	スクリーン投影形撮像系	43
25	スクリーン投影形撮像系の性能例 (D の単位は mm)	44
26	物体検出のための光学系	45
27	ぼけの影響による波形の変化	47
28	D C T係数のぼけによる減衰	48
29	D C T係数配列	49
30	試作したスクリーン投影形撮像系の概観	50
31	撮影対象の概略図	51

32	多重フォーカス画像	52
33	距離計測結果	53
34	目標物	54
35	バンド幅の変化による正解率の変化 (DCT, 目標物 (a))	55
36	領域抽出結果 (左:手前, 中:正解, 右:遠く)	55
37	領域抽出結果その2 (左:手前, 中:正解, 右:遠く)	56
38	バンド幅の変化による正解率の変化 (DFT, 目標物 (a))	57
39	バンド幅の変化による正解率の変化 (DCT, 目標物 (b))	57
40	バンド幅の変化による正解率の変化 (DFT, 目標物 (b))	58
41	廊下の風景	60
42	廊下のシーンからの検出結果 (閾値なし)	60
43	廊下のシーンからの検出結果パターン (閾値なし)	61
44	廊下のシーンからの検出結果 (閾値あり)	61
45	廊下のシーンからの検出結果パターン (閾値あり)	62
46	典型的な道路標識	64
47	標識抽出の手順	66
48	緑色フィルタの範囲	67
49	距離情報を用いた標識の認識	69
50	実験で用いた目標物	70
51	抽出された物体	71
52	領域抽出時のパターン	71
53	目的距離と判別されたブロックの比	72
54	標識とビデオカメラで撮影した画像の関係	73
55	ウインドウを通過する標識の面積	75
56	標識が検出ウインドウからはみ出す場合	75
57	標識と検出ウインドウが重なる場合	76
58	標識が検出ウインドウに含まれる場合	76
59	Clumpiness と Missibility の例	77
60	形状作用素を適用した例 (a) 原画 (b) $t=0$ (c) $t=9$	78
61	湾岸線の動画像解析結果のヒストグラム (全体)	80
62	湾岸線の動画像解析結果のヒストグラム (部分)	81
63	選択されたフレームと標識	81
64	前後の動画フレーム	82

65	標識以外	82
66	誤認識例	83
67	背景に標識と似通った色の物体がある場合の画像	85
68	標識部分の画像	86
69	標識の検出結果	86
70	路面上の文字と標識のある画像	87
71	文字領域抽出結果	88
72	道路管理用画像データ管理システム	89
73	抽出した標識を地図上に表示した様子	90

表 目 次

1	試作したスクリーン投影形撮像系の仕様	50
2	道路交通標識に用いる表面色	64
3	標準照明光	68
4	マルチスケールによる標識の分類	78
5	認識結果	84
6	テクスチャ解析の結果	84

1 序論

1.1 研究の背景と目的

テレビ番組や映画、あるいはこれらを記録したビデオなどの映像メディアを記録する媒体には、今日まで磁気テープが用いられてきたため、記録再生は時間に沿って連続的に行われていた。しかし、光磁気ディスクを始めとするランダムアクセス可能な記録媒体に大量のデータを記録できるようになり、映像メディアを時間にかかわらず記録再生できるようになった。撮影した位置や撮影されている物体の情報を映像メディアに添付しておけば、位置や物体を元にして映像メディアを参照することができる。このような記録媒体に、我々の住む環境を記録すれば、時間を遡って参照したり、これから訪れようとする場所をあらかじめ映像によって確認したり、目的とする物体がどの場所にあるかを調べたりすることができる。

こうした映像メディアの産業的な応用として期待されているもののひとつに、道路標識などの道路設備管理業務がある。道路設備は、延長1000km以上にもおよぶ長い区間に設置されている。そして、道路設備の点検を行ったり、新たに道路設備を設置する場所を事前に調査する必要がある。撮影場所の位置情報は、衛星を用いた位置測定システム（Global Positioning System）によって取得することができる。しかし、撮影されている道路設備を映像メディアから抽出することができなくては、目的とする道路設備の検索をすることができず、映像メディアを有効に利用することができない。

本研究の目的は、道路を走行する車に搭載されたカメラによって撮影された映像メディアの中から道路標識を抽出するための手法の提案とその実証を行うことである。本研究では、新たに2つの方法を提案している。第1の提案は色、形、奥行きなどの視覚に関する多元的な情報を組み合わせた映像メディアからの道路標識抽出法であり、第2の提案は奥行き情報を取得するための撮像系の提案である。道路の周辺には、道路標識と似通った色や形をもつ物体が数多く存在する。このため、色や形を識別する単独の解析方法では、道路標識を識別することは困難である。そこで、第1の提案では、物体を識別する際に、物体の特徴や奥行き情報を識別する機能を有機的に組み合わせている。そして、視覚に関する人の脳の働きを参考にしてシステムを構成し、道路標識の識別を行っている。また、第2の提案では、視覚情報の一つである奥行き情報を合焦法を利用して取得している。合焦法は遠距離の物体が識別できないため、道路標識を識別することは困難であったが、本提案では、主レンズで得られ像を一旦スクリーンに投影し、これをカメラで撮影する方法によりこの課題を解決している。また、焦点ぼけの大きさを測る方法として、離散コサイン変換（DCT）が利用できる事を示している。DCTは画像の符号化に用いられており、

DCTを行う高速なアルゴリズムやハードウェアを、物体までの距離情報を取得するためにも用いることができることを示している。

本研究の提案により、長い経路に沿って存在する物体を、物体の特徴に応じたフィルタの組み合わせや、物体までの距離を測定する光学系の工夫によって抽出可能である。この方法は、道路のみならず鉄道等にも応用が可能である。

1.2 脳における映像処理と映像メディア解析

対象物の持つ多角的な情報を利用して、映像メディアの中から対象物を抽出するために、脳の働きを参考にする。図1は、人の脳を示している。脳のうち、物体の認識などの

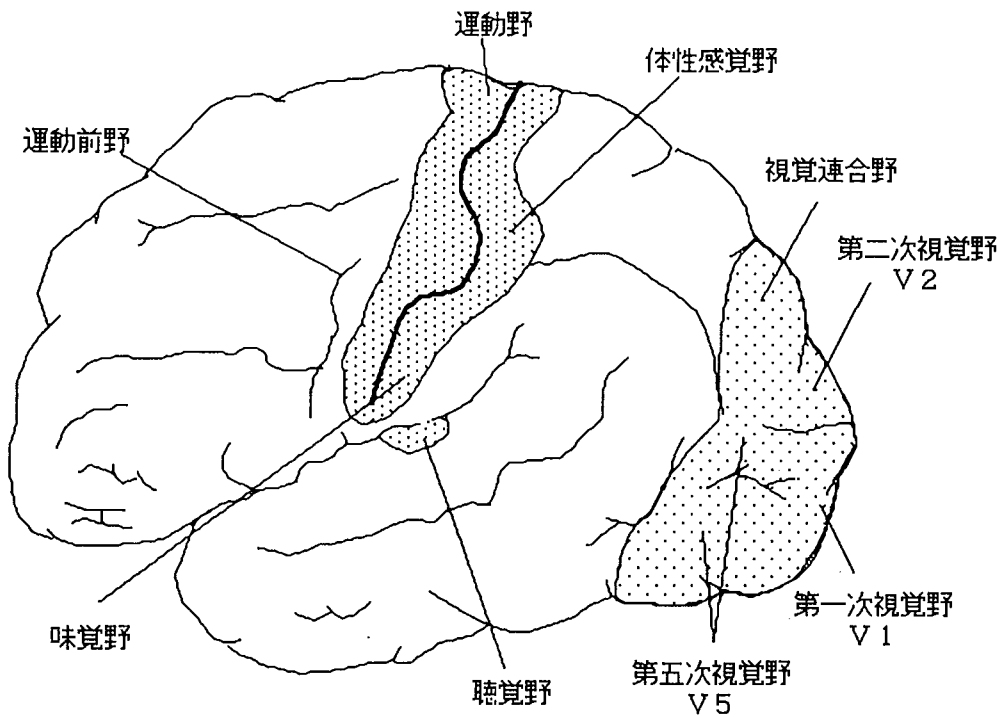


図 1: 脳の受容野の位置

高度な作業は大脳皮質で行われている。大脳皮質は視覚野聴覚野などの受容野と呼ばれる部分に別れており [1], それぞれの受容野毎に異なった処理を行っている [7]. こうした受容野毎に行われた処理結果が統合されてさらに高度な処理が行われる。視覚に関しては、網膜で捕らえられた光は視覚野で処理される。これらは第1次視覚野 (V1) から、2次

(V2), 3次(V3), 4次(V4), 5次(V5)によって構成されている。

図2に、これらの視覚野の働きの関連を示す[5]。この関連図は、視覚に関する大脳皮質の局所並列階層モデルと呼ばれている。第一次視覚野では、網膜によって得られた光に対して画像処理を行い、色や形、テクスチャを得るため基本的な画像処理を行う。第二次視覚野ではこの情報を元に、観察者から見た物体の奥行きや方向の表現である $2\frac{1}{2}$ 次元スケッチを構成する。さらに、第三次、第四次視覚野では、物体の立体的情報や色、テクスチャが認識され、MT (middle temporal area), MST (middle superior temporal area), IT (inferotemporal area) によって、個々の物体の動き、観察者の動きが検出され、3次元物体の記憶像と照らし合わせながら物体が認識される。このように脳内では、視覚に関して異なった機能を持つ視覚野が存在し、これらが協調して働きながら物体の認識を行っている。この中で、特に注目したいのは、脳は視覚に関しては、基本的な画像処理で得られた情報を基に、奥行き、色、テクスチャ、物体の動きを求めていること、最終的な物体の認識には、脳内にある3次元物体の記憶像をもとにしていることである。本研究では、このような人の脳の働きにならない、物体の持つ特徴を捉える手法やカメラとの位置関係を利用して映像メディアの解析を行っていく。

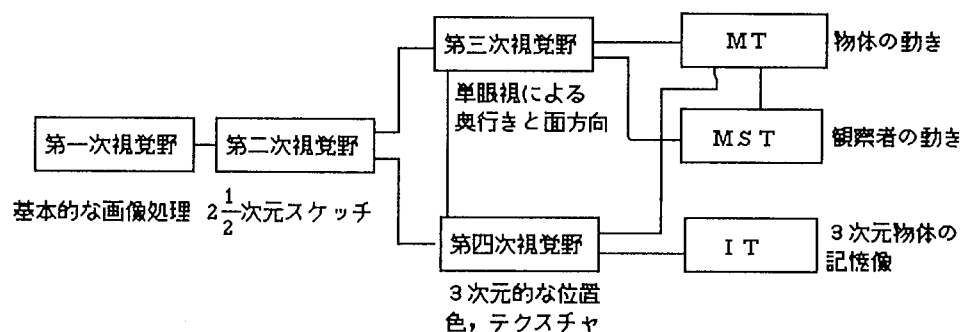


図 2: 視覚大脳皮質の局所並列階層モデル

1.3 本研究の産業分野への応用

本研究は、標識などの道路関連設備の保全管理用の画像データベースへの応用を目指している。道路は、遠隔地に人や物資を運ぶために重要な役割を果たしており、交通の流れを円滑に保つことが必要であり、このためには、道路管理業務が不可欠である。道路管理業務には、道路の破損状況を把握し修繕を行うことや、標識を始めとする新たな設備を設

置ることがある。道路の現状を把握する際に重要となるものが、映像情報である。従来の道路管理は、図3のように、現場に出かけて行き、写真を取り、これを台帳に張り付けて管理していた。しかし、この方法では、交通を止める必要がある、車から降りて撮影する必要があり危険かつ時間がかかる、台帳の作成や検索に時間がかかる、といった問題点があった。

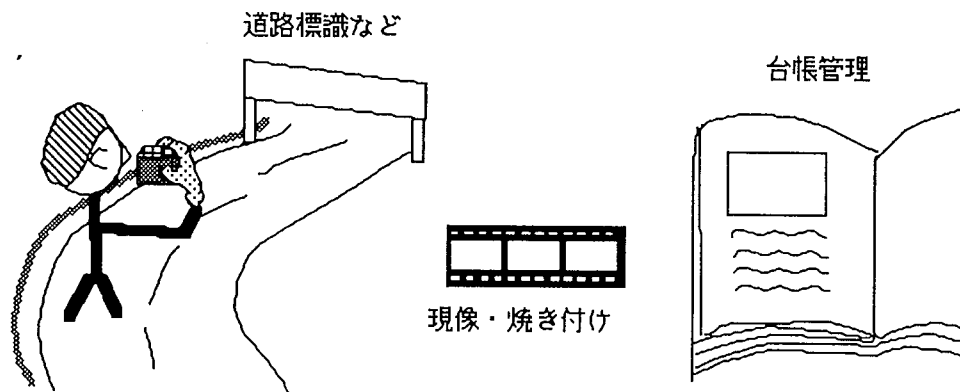


図 3: 従来の道路管理方法

この問題を解決するシステムが、図4である。このシステムでは、車にビデオカメラとカーナビを搭載し、道路を走行しながら撮影して行く。そして、動画フレーム毎に位置情報を添付して管理する。映像メディアの蓄積媒体として、大容量のハードディスクや書き替え可能な光ディスク装置などを用いれば、位置情報を元に画像を検索することができる。

しかしながらこのシステムには、撮影した物体にもとづいて検索することができないという課題がある。検索する物体の位置情報をあらかじめ登録しておくという方法もあり、たとえばD-GPS¹を用いれば、数メートルの誤差で位置を求めることができ、物体の近辺で撮影した映像を見ることはできる。しかし、撮影位置の数メートルの誤差やカメラの撮影方向のわずかなずれにより、対象物の一部が欠けたり、対象物に対して遠すぎる位置から撮影したものが選択されてしまう。このため、映像メディアの中から対象物が最適な位置で撮影されたものを抽出する技術が不可欠なものとなる。さらに、映像メディアの中から撮影対象物を抽出することができれば、前回撮影した道路管理設備と比較してその汚れや、破損状況を自動検出することが可能になる。

¹ あらかじめ位置の分かっている場所で衛星からの電波を受け補正情報を作成し、この補正情報をFM波を用いて送信する。衛星からの電波によって位置を知る時は、この情報に基づいて補正を行う。

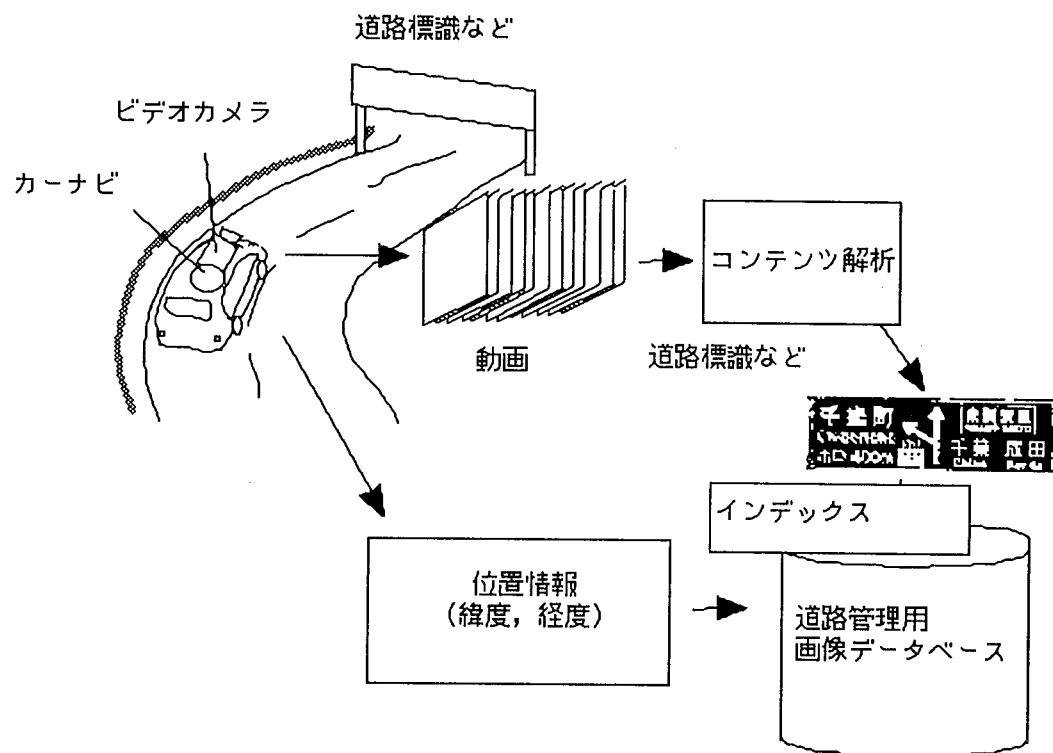


図 4: 将来の道路管理方法

ところで、映像メディアに我々の身の回りの環境を記録でき、内容や位置に基づいて検索できるようになると、我々の生活はどのように変わるのであろうか。図5は将来の映像メディアの利用方法を示したものである。大量の映像メディアを安価に蓄積することができるようになり、ビデオカメラの性能が向上すれば、高画質の動画情報を取得し蓄積することができる。撮影方法を考慮する必要があるが、例えば人の一生の間に眼にする映像をすべて蓄積することも考えられ、ある人が立ち寄ったすべての場所や、すべての人物を記録できる²。また、大学や企業の敷地や建物の内部を時間を追ってすべて記録しておくこともできる。このような情報は、ネットワークを通じて世界中の人々と共有することもできる。例えば、これから訪れてみたい場所の情報を様々な角度からあらかじめ調べることができる。また、その場所の季節的な変化を見ることもできる。そのためには、こういった世界においては、映像メディアに何が撮影されているかを抽出することが重要である。本研究は、そういった将来への第一歩をめざしているのである。

1.4 本論文の構成

本論文は5つの章から構成されている。第1章「序論」では、本研究の背景として、時間軸によらずランダムな読み書きが可能で大容量の情報を記録できる蓄積媒体の登場がある事を述べ、本研究の目的はこの蓄積媒体に記録された映像メディアの中から物体を抽出することであることを述べた。また、映像メディア解析の手法は人の脳の働きに学び、撮影対象物の色、対象物までの奥行き、対象物の立体的な位置、対象物の持つテクスチャといった多元的な情報を組み合わせる事の重要性、および物体までの奥行き情報を取得する事の重要性を述べた。さらに、本研究の応用目的が道路管理業務であることを述べ、道路管理業務における映像情報の重要性を述べ、映像情報を有効に利用するためには、標識を始めとする道路関連施設を映像メディアから抽出することが重要であると述べた。

第2章「映像メディアからの情報抽出」では、映像メディアに関する研究の現状と課題について述べた後、本研究における映像メディアの解析法について述べる。この中で、従来手法の問題点を指摘すると共に、本研究の課題は、映像メディアに関する様々な情報を統合して利用する方法、および映像メディア解析のうち物体までの奥行き情報を取得する方法である事を述べる。

第3章「合焦法を用いた遠距離物体の検出」では、物体の奥行き情報を取得するための

² 例えば1フレーム1Mバイトで蓄積するとすれば約75ペタバイト $1(Mbyte/frame) \times 30(frame/sec) \times 60(sec/minute) \times 60(minutes/hour) \times 24(hours/day) \times 365(days/year) \times 80(years/life) = 75,686,400,000 Mbytes \doteq 75 petabytes$ のデータ量であり、近い将来実現されると思われる

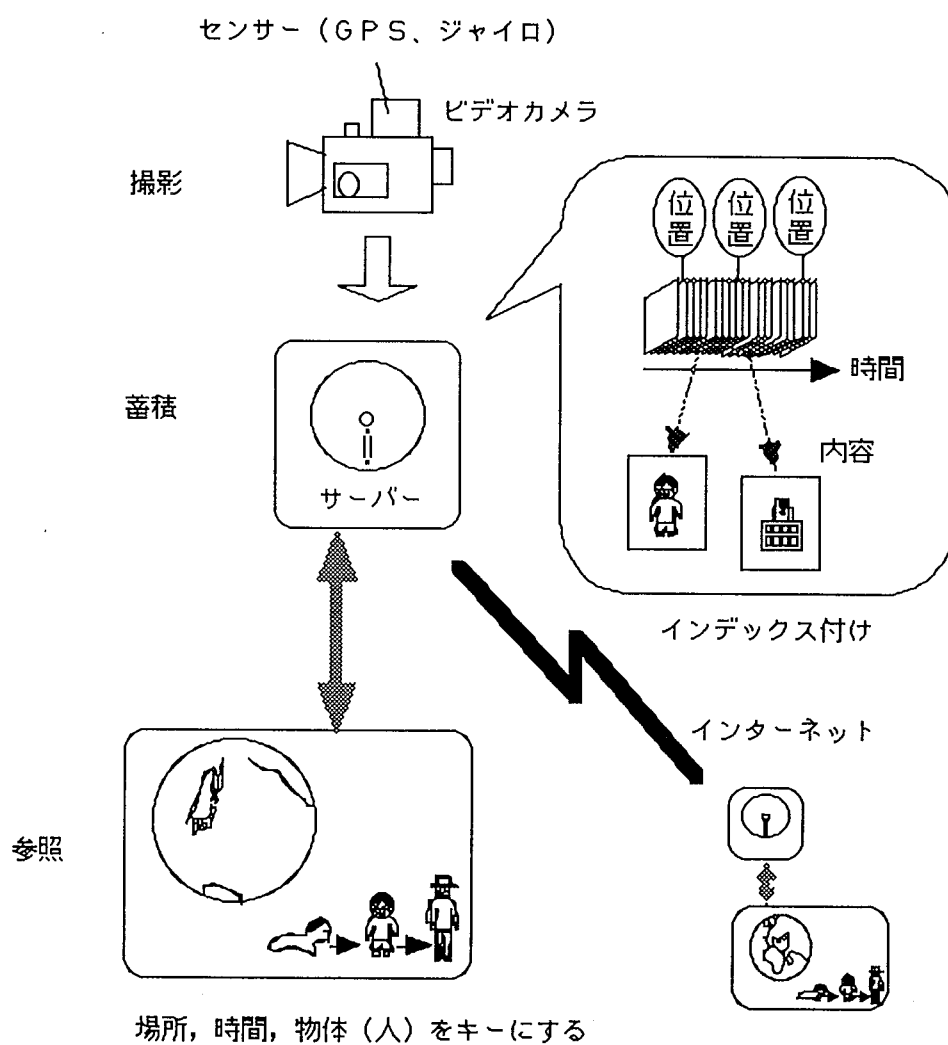


図 5: 将来の映像メディアの利用方法

新たな方法として、合焦法を利用した撮像系の提案をする。最初に、レンズの焦点合わせの位置を変えて、焦点の異なる複数の画像を撮影し、焦点ぼけの大きさを比較することによって物体までの距離を求める合焦法は、遠距離になればなるほど焦点ぼけの差が生じにくく、物体までの距離を求めることが困難であることを述べる。

提案する手法では、口径の大きなレンズで得られる遠距離物体に対しても焦点ぼけの差の大きな像を一旦スクリーンに投影し、この像を内視鏡などに用いられる市販のテレビカメラで撮影する。これにより、市販のテレビカメラを用いながらも10 m以上離れた物体を検出することが可能になることについて述べる。また、焦点ぼけの大きさを測る方法として、動画像の圧縮に用いられる離散コサイン変換(DCT)を利用可能なことを理論的かつ実験的に実証する。また、従来用いられてきた離散フーリエ変換と性能的に変わらないことを実験により確かめる。これにより、映像メディアの圧縮に使われる符号化方式が焦点ぼけの大きさを測る際にも利用可能な事について述べる。また、焦点の異なった3枚の画像を撮影することにより、撮像系から10 m離れた物体を検出できることについて理論的かつ実験的に実証する。

第4章「多元情報を用いた映像メディアからの道路標識の抽出」では、多元情報を用いて車載カメラから撮影して得られた映像メディアの中から、道路標識を抽出する方法について述べる。この方式は、道路標識が、周囲と区別しやすい色に着色されている、路面からの高さや大きさなど空間的な位置が決められている、単色の背景に文字が規則正しく並んでいる、といった特色を利用するものであり、色フィルタ、検出ウインドウ、テクスチャ解析を組み合わせることにより、道路を走行する車から撮影したビデオ画像から標識を抽出できることを述べる。また、標識の背景に標識と同じ色を持つ木や山などの物体がある場合は、抽出できない場合がある。そこで、距離を測定することにより、背景に同じ色の物体がある場合も識別できるようする。さらに、画像の中から文字領域を切り出す従来の単独の手法と比較して、本提案方式の有効構成を実験により実証する。また、実際に道路管理用画像データベースシステムに、標識の認識システムを実装した事について述べる。

第5章「結論」では、本研究を総括し、本研究で得られた研究成果をまとめた上で、今後の課題、将来の展望について述べる。

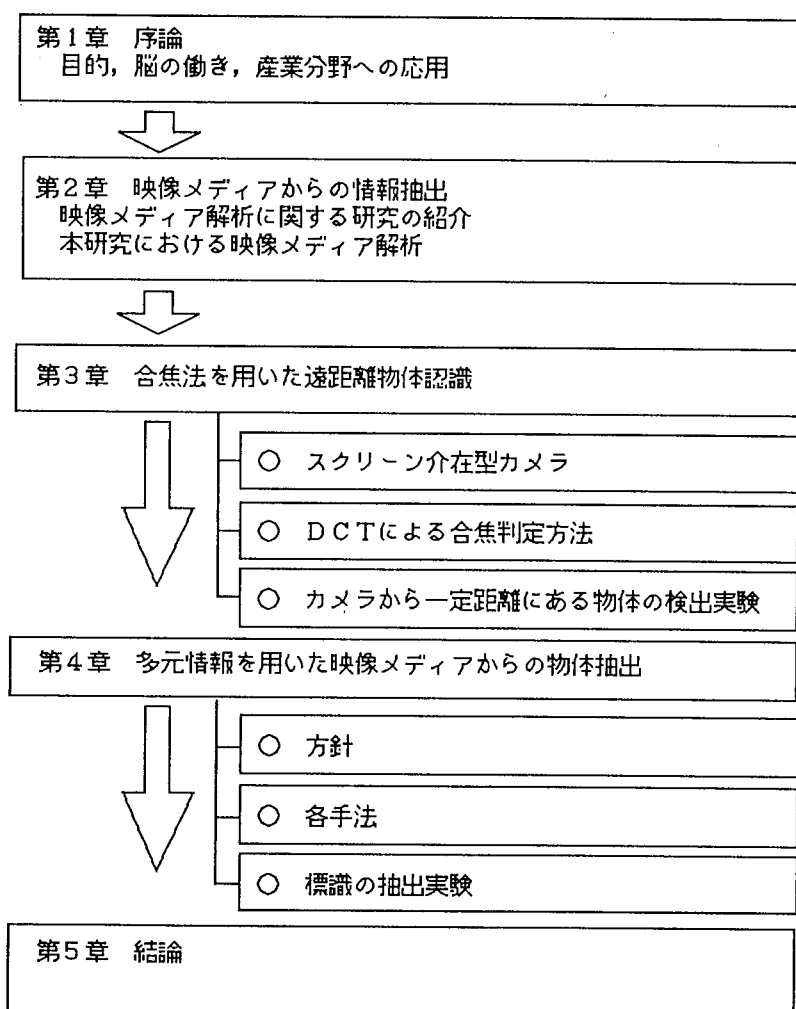


図 6: 論文の流れ

2 従来の研究と解決すべき研究課題

2.1 まえがき

道路を撮影した映像メディアから道路標識を抽出するためには、大量の映像データから効率良く目的の画像を抽出でき、標識の背景に撮影されている様々な物体から標識のみを区別する手法が必要である。そのためには、第1章で述べたように、人間の脳の働きに学び、物体の持つ色、テクスチャといった画像的な特徴を捕らえる手法と、物体の動きや奥行きといった空間的な特徴を捕らえる手法を組み合わせることとする。

ところで、一般的に映像メディアは、画像・音・言語・図形等のマルチモーダル情報の同期的なデータストリームを指しており、具体的にはテレビ番組、映画、そしてこれらを記録したビデオである。また、焦点距離などのカメラのキャリブレーション情報は含まない[6]とされている。しかし、本研究で扱う映像メディアは道路を撮影したものであり、映像メディアから抽出したい対象物も道路標識を始めとして、道路に設置された施設である。このため、映像メディアの中で特に画像情報を扱う対象とし、これに、撮影時に取得する物体までの距離情報によって解析を行う。図7は、本研究における映像メディア

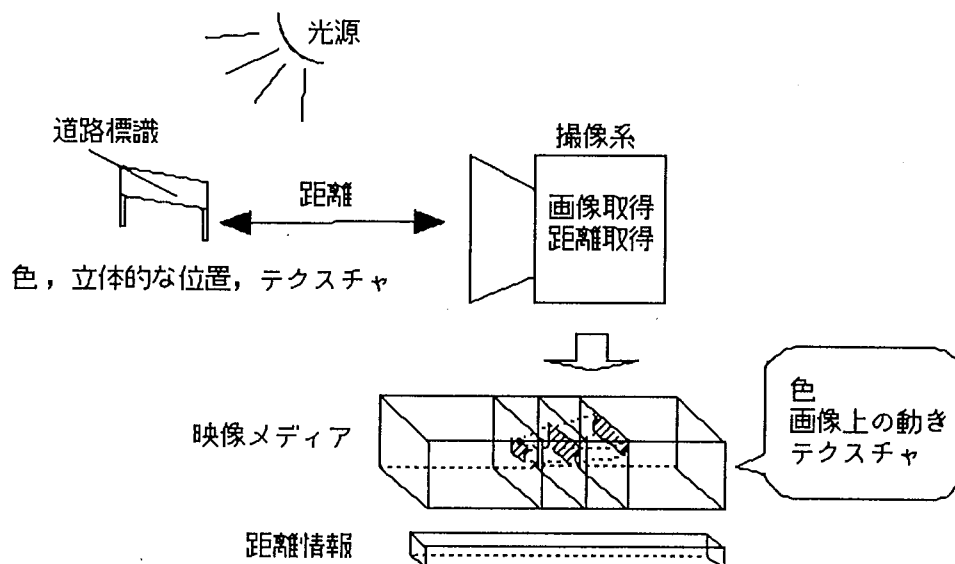


図 7: 本研究における映像メディアの取得の様子

ア取得の様子を示している。対象物である道路標識は、色、立体的な位置、テクスチャにおいて他の物体と異なった特徴を持っている。具体的な特徴については、4.2 節に詳しく述べる。撮像系によって道路標識を撮影する場合の色は、標識の本来持つ色と光源の色によって決定される。また、道路標識の画像上の位置は道路標識の立体的な位置と、撮像系の位置関係によって決定され、撮像系を移動しながら撮影することから、その位置が移動する。映像メディアには、道路標識の情報が色、物体の動き、テクスチャとして記録される。さらに、撮像系と物体までの距離情報も利用する。以下の節では、映像メディアに関する主な研究について述べた後に、本研究で提案する方法について述べる。

2.2 映像メディア解析に関する従来の研究

2.2.1 動画像の構造解析

ビデオは動画フレームと呼ばれる画像の連続によって構成されており、各動画フレームは時間によって管理されている (図 8)。ビデオを構造化する研究としては、動きや色、テクスチャを解析することによりビデオ構造を視覚化したり ([21])、ビデオに撮影されているスキーヤーや野球選手を識別する研究 [22] がある。阿久津らはビデオ画像に対して時空間投影画像を求め (図 9)、この画像からカメラの動きを求めている [23]。また、ステレオ画像に用いるエピポーラー線を利用したもの [26]、大きさの変化する物体の領域を追跡しながら解析する研究 [27] などがある。シーンの切れ目を検出するために、フレーム毎の色平均を求めてフレームの特徴量とし、この変化を調べてシーンを分類する研究 [24] や、フレームの DCT 係数の変化からニュース番組を録画したビデオの解析を行う研究 [25] がある。また、画像中に双曲線型のスリットを設け、このスリットを通過する画像を解析することにより空間中の直線情報を抽出する研究 [31] がある。

2.2.2 色情報の解析

観測される色は、物体の持つ色と光源の色によって決定される [9]。西辻、富永らは反射モデルのパラメータ推定法について述べている [8]。画素 $\mathbf{x}$ で観測された色ベクトルを $\mathbf{c}(\mathbf{x})$ とし、光源の正規化ベクトルを $\mathbf{e}$ 、拡散反射成分による物体色ベクトルを $\mathbf{c}_d$ とおくと、色ベクトル $\mathbf{c}(\mathbf{x})$ は式 (1) によって表される。右辺第 1 項は拡散反射成分で、第 2 項は鏡面反射成分であり、 $w_d(\mathbf{x})$ と $w_s(\mathbf{x})$ は拡散反射と鏡面反射の重み係数である。

$$\mathbf{c}(\mathbf{x}) = w_d(\mathbf{x})\mathbf{c}_d + w_s(\mathbf{x})\mathbf{e} \quad (1)$$

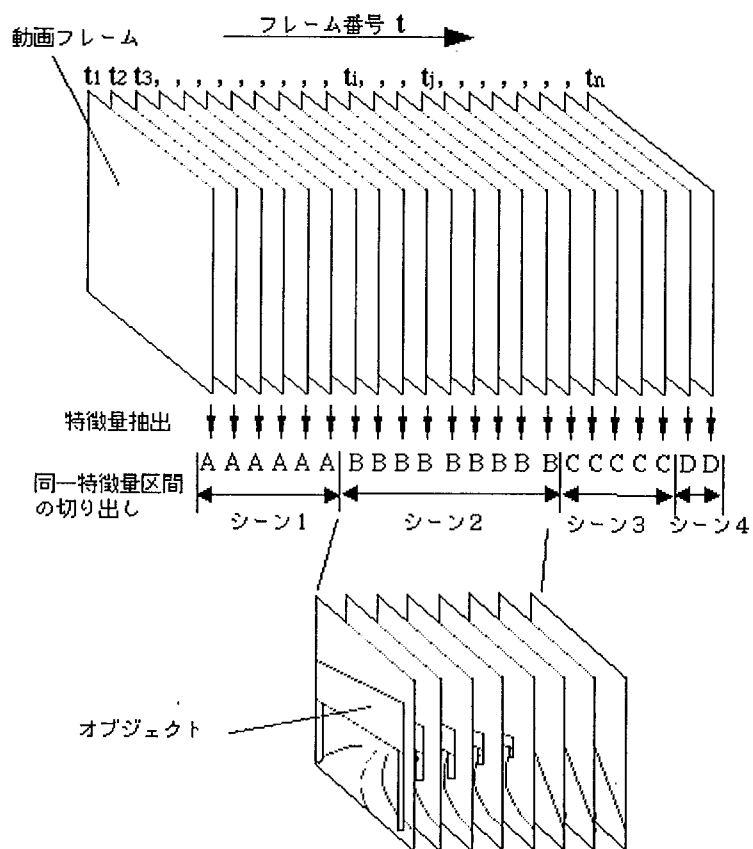


図 8: ビデオの構造

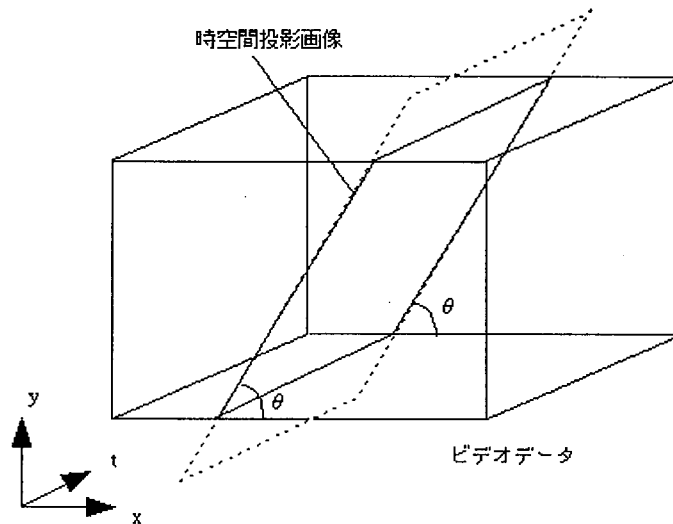


図 9: 時空間画像の切断面

図 10 のような円筒物体上の表面色を観測すれば，鏡面反射によるハイライトクラスタと拡散反射による物体色クラスタを観測することができる（図 11）．

こうした，観測される色と，光源および物体色の関係を元にして，物体間の相互反射がある場合の画像の解析法 [33] や，不均質物体が存在する場合 [34]，カラー画像のセグメンテーション [10] の研究が行われている．

2.2.3 奥行き情報の解析

奥行き情報を取得する方法としては，物体の位置を計測する方法として，2 眼式ステレオ法があり，1 台のカメラを移動させることにより 2 眼式と同じ効果をうるモーションステレオ法もある [48, 49]．また，光切断法や，レンズによる焦点合わせの位置を利用して物体の位置を求める合焦法がある．以下，奥行き情報の取得の代表的な方法として，ステレオ視，光切断法，合焦法について紹介する．

2.2.3.1 ステレオ立体視 ステレオ立体視は，図 12 のように，人間が左右の 2 つの目で物体までの距離を知るのと同じ原理によって物体までの距離を知る方法である．左右 2 台のカメラとお互いの位置関係を求めておき，物体上の点 P の左右のカメラで撮影した画像上の位置， P_L ， P_R を求める事により，三角測量の原理で物体までの距離を知る事ができる．

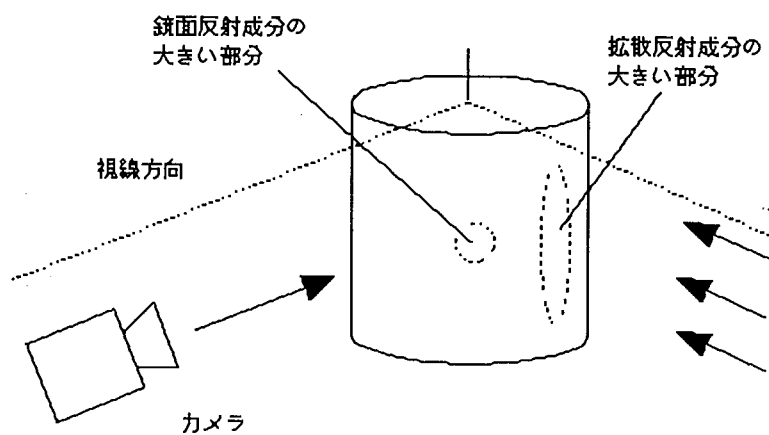


図 10: 円筒物体表面上での色分布

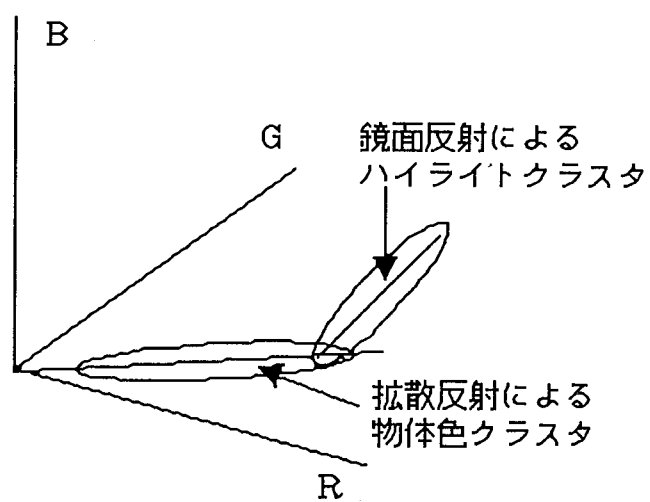


図 11: RGB空間でのカラーヒストグラム

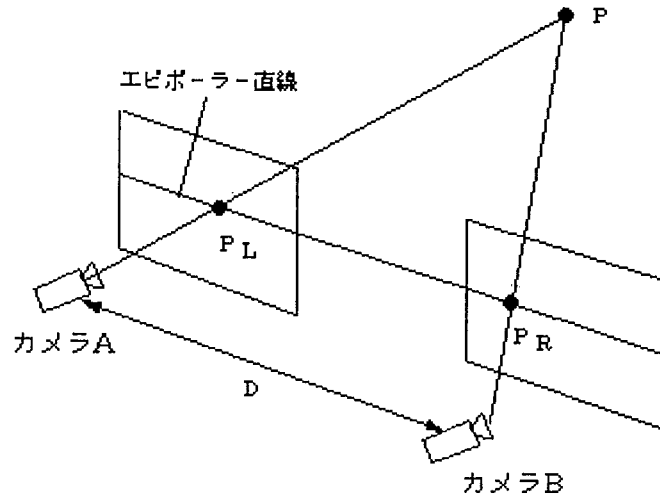


図 12: ステレオ視

カメラの位置関係を知らなくても，2台のカメラで撮影した画像上の8点以上の対応を取る事により，カメラの位置と物体の位置を同時に求める方法もある[13]．しかし，いずれにしても左右のカメラで撮影した画像上で同じ物体の点の対応を求める，いわゆる対応点問題を解決する必要がある．

2.2.3.2 光切断法 光切断法は，物体にスリット光などをあて，物体に投影されたスリットの形状から物体の形を求めるものである．スリット光としてレーザー光線を利用した研究がある[11]（図13）．スリット光を当てた時，物体の表面が均一でない場合正確な測定を行うことができない．そこで，カメラのダイナミックレンジを広げ，照射する光の量も可変にする研究がある[12]．スリット光を複数本にしてビデオレートで物体形上の計測を可能にする方法などが研究されている[15]．

2.2.3.3 合焦法による奥行き情報の解析 合焦法による距離測定は，レンズの焦点距離 f とレンズから物体までの距離 u およびこの物体の像が合焦するレンズからの距離 v の間に，式（2）の関係があることを利用したもので，比較的古くから知られている．

$$\frac{1}{f} = \frac{1}{u} + \frac{1}{v} \quad (2)$$

距離を求める方法として，焦点を変えた多数の画像を撮影し，この中でもっとも焦点の合っている画像を求め，その画像を撮影した際の投影面の位置から距離を求める Depth

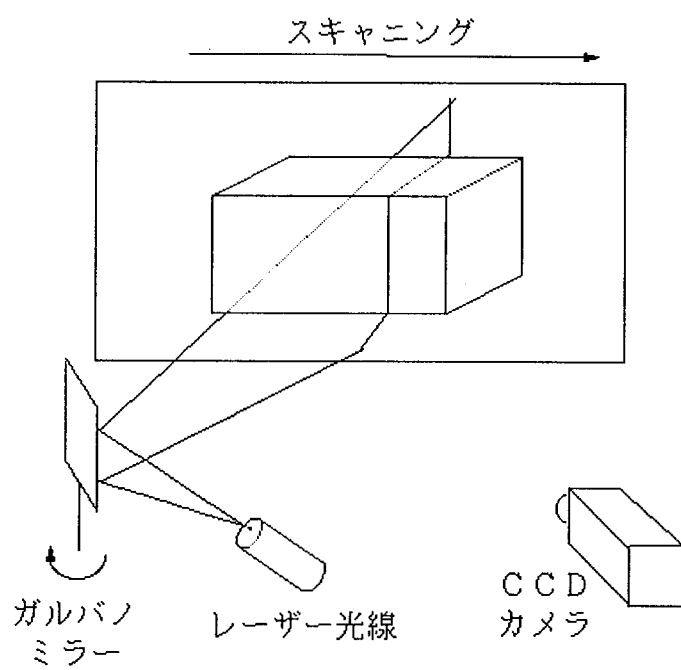


図 13: 光切断法

From Focus (DFF)[58, 16] や、ぼけモデルを用いて数枚の画像から物体の位置を求める Depth From Defocus (DFD)[51, 63, 64] がある。

本研究では、DFF によって物体までの距離を求めるが、ここでは関連の深い DFD について詳しく述べる。図 (14) は DFD の原理を説明した説明図である。

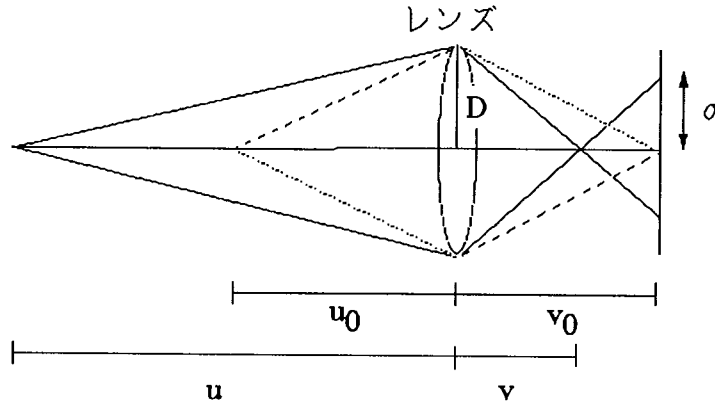


図 14: DFD による距離測定

式 (2) より,

$$u = \frac{fv}{v - f} \quad (3)$$

$v = v_0$ の時,

$$u_0 = \frac{fv_0}{v_0 - f} \quad (4)$$

図 (14) から,

$$\tan \theta = \frac{D}{v} = \frac{\sigma}{v_0 - v} \quad (5)$$

式 (2) および式 (5) より、物体までの距離 u は式 (6) で与えられる。

$$u = \frac{frv_0}{Dv_0 - (D + \sigma)} \quad (6)$$

ぼけの量はフーリエパワーで求めることができる。ぼけ円の中心を (x_0, y_0) としぼけ関数を $f_1(r, \theta)$ とすると、ぼけ関数は

$$f_1 = (r, \theta) = I_1(x_0 + r \cos \theta, y_0 + r \sin \theta) \quad (7)$$

で与えられる. f_1 のフーリエ変換を $F_1(t, \theta)$ とする. 同じように $f_2(r, \theta)$ と $F_2(t, \theta)$ を考える.

ぼけのない時の f_1 および f_2 を $f_0(r, \theta)$ とし, f_1 および f_2 がガウス関数 $G(r, \sigma)$ との畳込み演算で得られるとすれば, 次の式がなりたつ. なお, $\otimes$ は畳み込み積分を表している.

$$\frac{f_1(r, \theta)}{f_2(r, \theta)} = \frac{f_0(r, \theta) \otimes G(r, \sigma_1)}{f_0(r, \theta) \otimes G(r, \sigma_2)} \quad (8)$$

$f(r, \theta) = e^{-\pi r^2}$ であり, $F(\lambda, \theta) = e^{-\pi \lambda^2}$ であるので, F_1, F_2 および F_0 の間には, 以下の関係が成り立つ.

$$F_1(\lambda, \theta) = \frac{F_0(\lambda, \theta) G\left(\lambda, \frac{1}{\sqrt{2\pi}\sigma_1}\right)}{\sqrt{2\pi}\sigma_1} \quad (9)$$

$$F_2(\lambda, \theta) = \frac{F(\lambda, \theta) G\left(\lambda, \frac{1}{\sqrt{2\pi}\sigma_2}\right)}{\sqrt{2\pi}\sigma_2} \quad (10)$$

結局

$$\frac{F_1(\lambda)}{F_2(\lambda)} = \frac{G(\lambda, \sigma_1)\sigma_2}{G(\lambda, \sigma_1)\sigma_1} = \frac{\sigma_2^2}{\sigma_1^2} \exp(\lambda^2 2\pi^2 (\sigma_2^2 - \sigma_1^2)) \quad (11)$$

両辺に対して $\log$ をとる事により,

$$\ln \frac{\sigma_2^2}{\sigma_1^2} + \lambda^2 2\pi^2 (\sigma_2^2 - \sigma_1^2) = \ln F_1(\lambda) - \ln F_2(\lambda) \quad (12)$$

となる, これは, ぼけの量の差がフーリエパワーの差によって, 表されることを示しており, フーリエパワーの差を求めることによって, ぼけの変化を調べることができる.

物体にテクスチャがない場合, ぼけの変化を調べることができない. そこで, カメラ側からパターン光を照射して精度をあげる. 例としては, フーリエ変換を用いた解析に適したパターン光を照射する研究 [17] がある. また, 多重フォーカス画像からエッジを検出し, ぼけによるエッジの濃淡の変化の程度を利用する方法 [53], レンズと物体の間に設置する絞りに工夫を加えフーリエ変換によって物体の正確な位置を求めることができる様にした研究 [18] などがある. 距離を求めるのではないが, 焦点の異なる複数の画像から全焦点画像を求める研究 [19] もある.

焦点ぼけを利用した距離測定の問題点は, 遠距離の物体の距離が測定できないことである.

2.2.4 テクスチャ解析

テクスチャには統計的なテクスチャと構造的なテクスチャの2つのレベルがある[20]. 統計的なテクスチャ特徴は、画像に対して統計的な処理を行うことによって得られる. 統計的なテクスチャ特徴の計算法として、濃度ヒストグラム、同時生起行列、差分統計量、ランレングス行列、パワースペクトルを用いるものなどがある. 構造的なテクスチャ特徴は、画像を構成する要素が一定の配列規則に従っている場合のことで、繰り返しパターンを検出することができる. 繰り返しパターンの代表的な抽出方法として、要素間の相対的位置ベクトルを用いるもの、フーリエ変換を用いるもの、形状作用素を用いるもの等がある[37].

以上のように、テクスチャ解析は解析する領域を特定する必要がある、映像メディアの解析においても、領域を切り分けることが重要である.

2.2.5 道路標識の認識

道路標識は、ドライバーが道路を安全かつ短時間で目的地に到着するために重要な役割を果たす. 道路標識には、目的地や通りの場所などを示す案内標識、踏み切りや行く手の交差点の形などを示す警戒標識、速度制限や一方通行等を示す規制標識、駐車場や横断歩道などを示す指示標識およびこれらの説明を行なう補助標識がある. 近年、自動走行車両の研究がすすめられており、道路標識認識の研究も行われている.

道路標識は、周囲の風景と区別しやすくするために、色や形に特徴がある. 移動するカメラで撮影した動画像から物体を抽出する場合、色情報は形の情報と比較して、安定に取りだせる[39]ため、色情報によって標識をおおまかに切り出した後、形状や絵柄を用いてさらに詳しい解析を行う場合が多い[43, 44, 45]. 色の解析方法としては、あらかじめ多数のサンプル画像によって標識の色による特徴空間を求め、観測された画像の解析を行う方法[41], また、標識の縁取りの色に注目して標識の領域を求る方法[42], RGB空間や色彩を用いる方法があり[4, 28], 色を解析する専用ハードに関する研究[46]もある. 形状に関しては、パターンマッチングによって標識を識別する研究[47]や、円形度を求めて行う研究がある[29].

これらの方式は、画像を1枚毎に解析するものであり、映像メディア解析のように一旦撮影された映像から物体を抽出する場合は、似通った画像に対して同じ処理をくり返す必要があり効率が悪い. また、背景に似通った色を持つ物体を持つ場合の解析方法については未解決である.

2.3 本研究で解決すべき課題

以上述べてきたように、従来の映像メディア解析では、画像の持つ特徴に関して個別の検討がなされてきた。しかし、映像メディアからの道路標識抽出、という観点に立つと、例えば色情報のみでは周囲の似通った色と判別することが難しい、ステレオ視では、対応点問題を解決する必要がある、という課題がある。また、道路標識の認識に関する研究では、色やテクスチャの情報が使われているが、背景に同じ様な特徴を持つ物体がある時の識別が困難である。

以上の事から、本研究では、図 15 のような方法で道路標識の認識を行う。すなわち、距離情報の取得可能な撮像系により、対象物の撮影と距離測定を同時に行う。この時、道路を走行しながら映像を取得するため、大量の画像データが得られる。次に、色フィルタや距離に基づく解析手法を用いて、道路標識の色や道路脇にある物体が撮影された映像のみを抽出する。こうして、大量の画像情報から道路標識の撮影されている一連の映像のみを抽出する。一連の画像情報には、同じ物体が少しずつ位置を変えながら撮影されている。そこで、画像上の物体の位置の変化を求めることにより、一連の画像の中から道路標識が最も大きく撮影されているものを抽出する。この時、道路標識の大きさと空間的な位置関係から、道路標識の領域を同時に切り出すことができる。最後にテクスチャ解析によって、道路標識以外のものを除外する。この方法の中で、物体までの距離情報を取得する方法に関しては、新たな撮像系を提案する。これは、距離を測定する場合、ステレオ視では対応点問題を解決する必要がある、合焦法では遠距離の測定ができない、といった課題に対応するためである。

以上をまとめると、本研究で解決しようとする課題は以下の 2 つである

映像に関する多元的な情報解析が標識の認識に有効なことを示す 色、立体的な位置、テクスチャといった標識の持つ特徴に適合した解析手法を組み合わせることにより、映像メディアから標識を効率良く抽出できることを実証する。

合焦法により遠距離の物体が識別可能なことを示す ステレオ視や従来の合焦法の課題を解決するために、合焦法を改良したスクリーン投影形撮像系を提案し、その効果を実証する。

2.4 まとめ

本章では、映像メディア解析に関する現在の技術について述べた。ビデオ構造を解析してシーンの切り分けを行う研究やカメラワークを求める研究があり、内容に基づく検索の

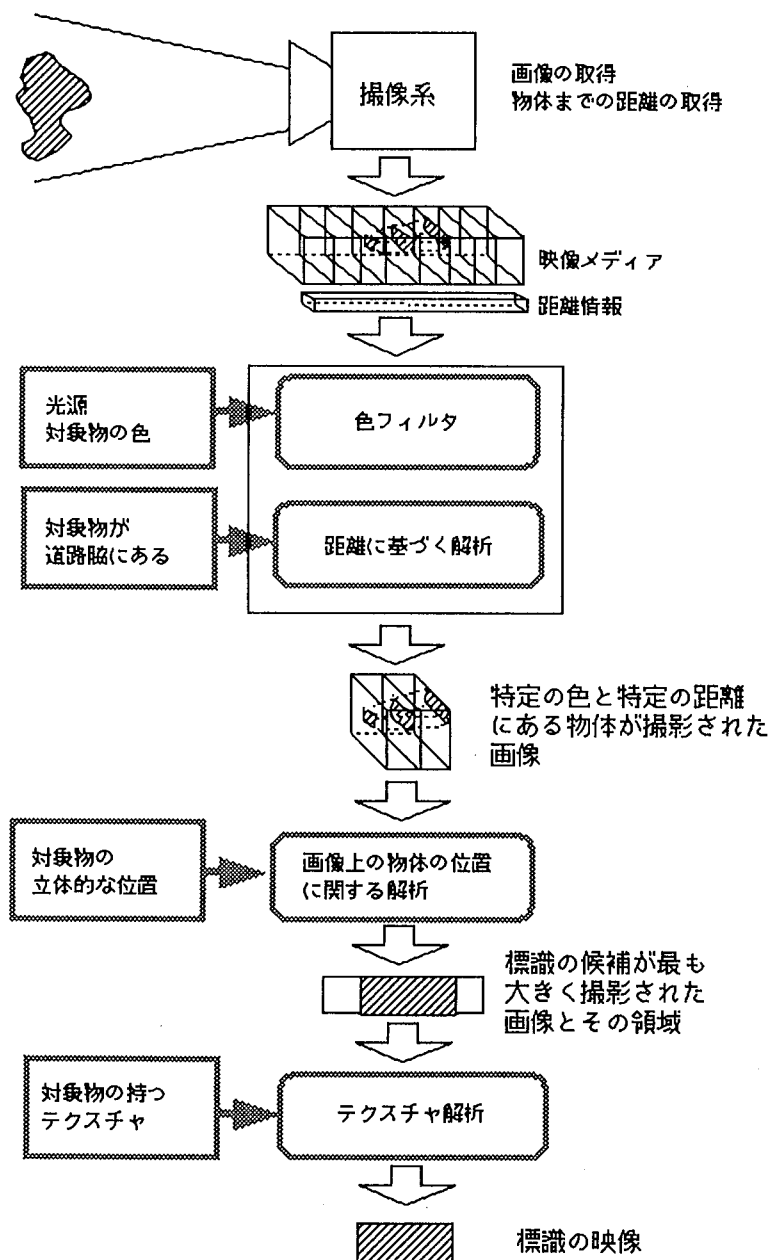


図 15: 多元情報を用いた映像のコンテンツ解析

研究も行われている。色や奥行き情報，テクスチャ解析などの研究も行われている。標識の認識は，色や形をもとにして行われている。しかし，解析しようとする物体の持つ個別の情報のみを利用したり，画像1枚毎に同じ処理をくり返す必要があった。また，背景に同じ色を持つ物体がある場合の識別も困難であった。本研究では，物体の持つ様々な情報を多元的に用いる方法と，物体までの距離情報を取得し利用する方法を提案とその実証を行う。

3 合焦法を用いた遠距離物体の検出

3.1 まえがき

抽出する物体の背景に同じ色を持つ物体がある場合、色情報のみでは物体を識別することが困難である。このため、物体までの距離情報を利用する。画像からのレンジデータの取得は、撮影装置や画像処理が単純で、任意の対象に対して行えるような汎用性の高い手法が望ましい。一般に用いられているステレオマッチング[60]では、三角測量の原理を用いてレンジデータを取得しているが、画像間の特徴点の対応づけは一般に困難であるうえ、遠距離では視差がでにくいという欠点がある。また、レーザーのスリット光を物体に向けて投射し、物体に投影されたパターンの形状から物体までの距離を測る光切断法[59]があるが、10 mを超える遠方では用いるのが困難である。

一方、画像処理や精密工学などの分野で最近用いられるようになってきた手法に、焦点ぼけを利用してレンジデータを取得する合焦法がある[50, 51, 52, 53]。合焦法では、焦点を変化させながら撮影した複数枚の画像(多重フォーカス画像)から対象に焦点が合う位置を検出することでレンジデータを取得する。この手法では撮影装置が単純で画像間の対応づけ問題もほぼ回避でき、これまで数多くの研究がなされているが、いずれも対象は部屋の中などの近距離にあるものであり、CGやVRで扱うような風景や建築物などの遠距離にある対象は扱っていない。これは、対象が遠距離にある場合には物体深度が深くなってぼけの量を認識しづらくなり、この手法の適用が困難なためである。

合焦法により遠距離計測を行うためには、物体深度を浅くしなければならない。そのため画像観測系には焦点距離が長くFナンバーが小さいレンズが必要になる。しかしそのようなレンズを用いて広画角で撮影するためには、CCDカメラでは実現が困難なほどに大きなサイズの受光面が必要になる。そこで本研究では、合焦法で遠距離を計測するために、スクリーン投影形撮像系を新たに提案する。スクリーン投影形撮像系では、CCDカメラとは別にレンズとスクリーンを用い、まずレンズを通った光をスクリーン上に結像させる。そしてその投影画像を背面からCCDカメラによって撮影する。この観測系の受光面はスクリーンであり、CCDカメラで撮影した範囲がそのサイズとなるから、CCDカメラをスクリーンから離すことで受光面サイズを容易に大きくできる。そのため、焦点距離の長いレンズを用いて画角の広い多重フォーカス画像を撮影し、遠距離計測を行うことができる。

さらに、受光面サイズを変えながら多重フォーカス画像を撮影することにより、撮影と同時に各画像間の画素の対応付けを行うことも可能である。従来の、テレセントリック光学系を用いて撮影時に画素の対応付けを行う手法[54]では画角が狭くなってしまうという

問題があったが、提案観測系による手法を用いた場合、撮影する画角は任意に決定することができる。

焦点ぼけの大きさを測定する方法として、従来は画素間の分散を用いたり [58] フーリエ変換が用いられてきたが [64]、離散コサイン変換 (DCT) を用いることにより、画像の圧縮方法と共通に用いることができる。DCT を用いて合焦法により物体の位置を求める場合、画像に焦点ぼけが加わった際のぼけの大きさとこの画像に DCT を施して得られる DCT 係数の関係を調べ、ぼけの大きさ変化を調べるための DCT 係数の用い方を決定する必要がある。また、合焦法は画像のぼけの大きさの変化によって物体までの距離を測定するものであるが、光学的な特性からカメラから物体までの距離が遠くなるにつれて、焦点の合う範囲が広がってしまい、遠距離になる程物体の位置を求めることが困難になってしまう。

次章以降では、まず合焦法の基礎と、それを遠距離に適用した場合の問題点を挙げる。次にその解決策として提案するスクリーン投影形撮像系について説明し、実験によって提案観測系が有効であることを確認する。また、合焦法を用いて特定距離にある物体の検出を行う際の光学系について述べる。次に、筆者らの開発したスクリーン投影形撮像系 [66, 67, 68] を用いて、カメラから 10 m 程度離れた遠距離でも合焦法により物体位置の測定が可能なことを示す。さらに、画像のぼけの大きさが画像の持つ周波数にどのように影響をおよぼす影響を調べる。そして、DCT 係数を用いることによって、物体を検出することが可能なことを実験的に示す。

3.2 合焦法の課題

3.2.1 カメラモデル

一般のカメラは複数枚のレンズから構成されているが、合焦法では図 16 に示すモデルのような焦点距離 f の単一薄凸レンズとして扱う。

図 16 に示すモデルでは光軸はレンズ中心を通り、受光面に対し垂直に交わっている。また、受光面が光軸と平行に移動することで焦点が変化し、レンズから受光面までの距離 (受光面位置) v と、そのときに合焦となるレンズからの距離 (焦点位置) u との間には式 (13) に示すレンズの公式が成り立つ。

$$\frac{1}{f} = \frac{1}{u} + \frac{1}{v} \quad (13)$$

式 (13) より、受光面位置が $v = f$ のときに焦点位置が無限遠になるため、受光面が $v = f$ の位置よりもレンズに近づくことは考えない。

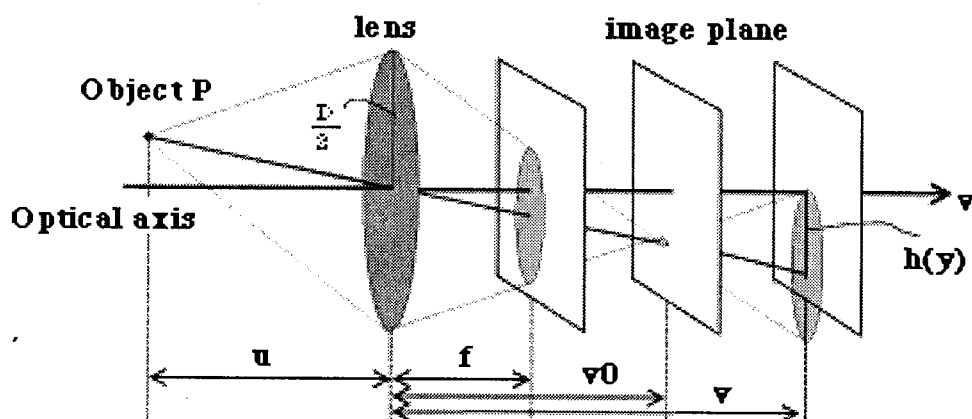


図 16: カメラモデル

3.2.2 多重フォーカス画像

図 16 における受光面位置 v を変化させながら撮影された複数枚の画像をまとめて多重フォーカス画像と呼ぶ。受光面が $v = f$ から $v = v_{max}$ まで動くとき、点 P の像はぼけの量が次第に小さくなり、ある点 $v = v_0$ で一点に結像し、再びぼけの量が次第に大きくなる。同時に、ぼけ円の中心は点 P とレンズ中心を結ぶ直線上を移動するため、受光面上の像は次第に拡大されていくことになる。受光面位置が v のときの、ぼけ円の中心と光軸との距離を $h(v)$ とすると、受光面位置が f のときの像と任意の距離 v のときの像との拡大率 $m(v)$ は式 (14) で表され、合焦法における画素の対応付けは式 (14) を用いて多重フォーカス画像を拡大・縮小することで行う。

$$m(v) = \frac{h(v)}{h(f)} = \frac{v}{f} \quad (14)$$

また、合焦法における画素の対応付けを不要にする手法として、テレセントリック光学系による多重フォーカス画像を撮影する研究がされている。テレセントリック光学系とは焦点距離の位置に絞りが入っているような観測系で、焦点を通過して入ってくる光はレンズと垂直に出ていくという性質を利用している。そのため $m(v) = 1$ となり、受光面位置が変わっても受光面上の像は拡大されない。しかしその性質上、結像面はレンズと同程度までしか大きくすることはできないため、画角が狭くなるという問題がある。

3.2.3 距離計測法

合焦法では、合焦となる受光面位置を決定することでレンジデータを取得する。そのため、焦点が合っている度合いを表す量(合焦測度)を定め、画像処理によりこれを測定する必要がある。合焦測度は、受光面が合焦となる位置に近づくにつれて単調に増加または減少するものが望ましく、画像間での合焦測度の比較によって各画素ごとの合焦となる受光面位置を決定する。

3.2.4 物体深度

ある受光面位置 v に対する焦点位置 u が決まるとき、図 17 のように u の前後にある物体の像には焦点ぼけが生ずる。しかし、このぼけ円が受光素子の大きさ ϵ よりも小さいときは、そのぼけの量を認識することはできない。

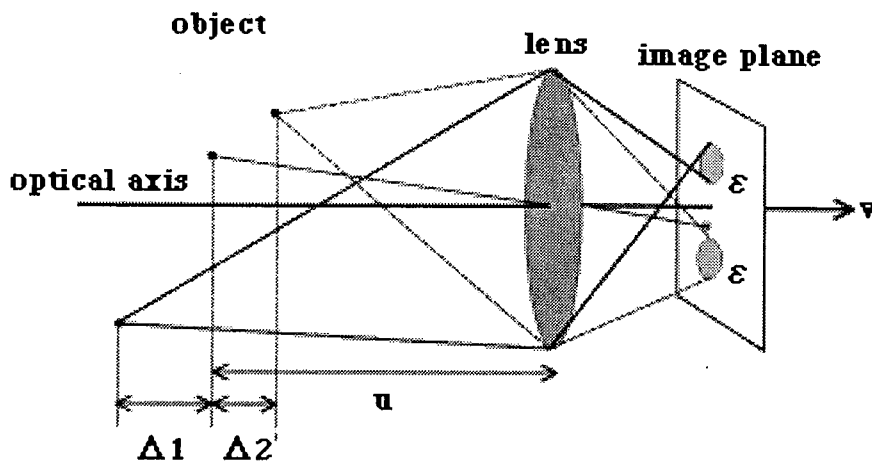


図 17: 物体深度

ぼけ円の直径が ϵ に等しくなる Δ_1, Δ_2 は、それぞれ後方、前方物体深度と呼ばれ、本研究では物体深度 Δ を式 (15) のように与える。ただし、 D はレンズ直径である。

$$\Delta = \frac{1}{2}(\Delta_1 + \Delta_2) = \frac{\epsilon u f D(u - f)}{f^2 D^2 - \epsilon^2(u - f)^2} \quad (15)$$

3.2.5 遠距離計測における問題点

合焦法では、焦点が対象とずれているときに生ずるぼけの量から合焦となる焦点を決定するため、被写界深度 Δ は浅くすることが望ましい。また、撮影対象を限定しないためには撮影範囲を広くし、遠距離でも計測できる必要がある。

ここで、CCDカメラを用いる場合の物体深度 Δ の変化について考える。撮影範囲を広くするためには図 18 で示される画角 2ω を広くしなければならず、必要な画角 2ω を決めると、受光面の対角の長さ d を用いて、そのときのレンズの焦点距離 f を式 (16) により求めることができる。

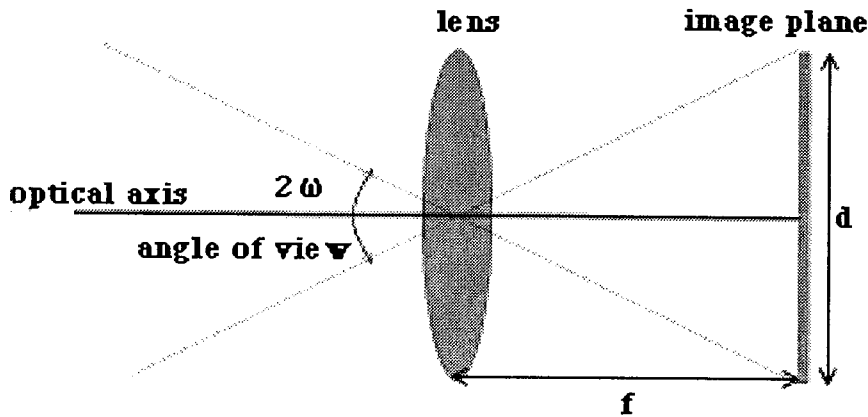


図 18: 受光面サイズと画角

$$f = \frac{d}{2 \tan \omega} \quad (16)$$

また、式 (41) により分解能 ϵ を求めることができる。

$$\epsilon = f \tan \frac{2\omega}{N} \quad (17)$$

ただし、 N は d に対する画素数である。これを用いると式 (15) は式 (18) のようになる。

$$\Delta = \frac{uD(u-f) \tan \frac{2\omega}{N}}{D^2 - (u-f)^2 \tan^2 \frac{2\omega}{N}} \quad (18)$$

例として撮影する画角を $2\omega = 24$ 度、CCDカメラとして $\frac{1}{2}$ インチ CCDカメラ (512 × 480 画素) を用いると、受光面の対角に対する画素数が $N = 702$ 画素となり、必要なレン

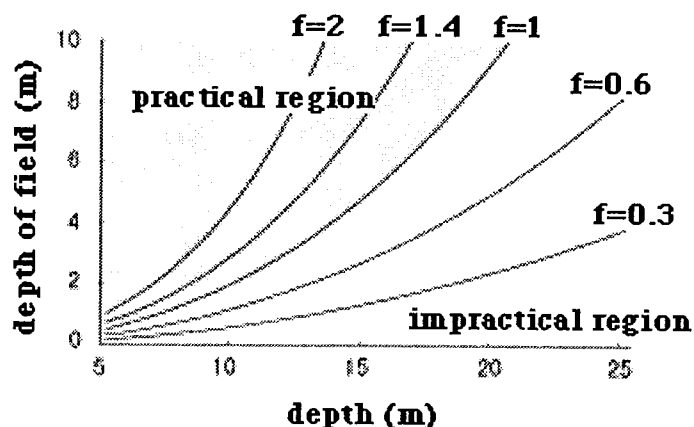


図 19: C C Dカメラで撮影する場合のFナンバーと物体深度の関係

ズの焦点距離は $f = 30mm$ となる. このときレンズのFナンバーを $F = 0.3, 0.6, 1, 1.4, 2$ とした場合の, 対象までの距離 u に対する物体深度 Δ の変化を図 19 に示す. ただし, レンズ直径を $D = \frac{f}{F}$ とした. 図 19 の実線に示すように, C C Dカメラで撮影した多重フォーカス画像では, 精密な遠距離推定をするためにFナンバーが1より小さいレンズが必要になる. しかしそのようなレンズの設計・製造は容易ではないため, 通常のC C Dカメラで多重フォーカス画像を撮影して遠距離計測を行うのは困難である.

3.3 スクリーン投影形撮像系

最初に合焦法で遠距離計測をするために必要な画像観測系の性能について考える. 計算機に入力する画像の対角に対する画素数を N , 撮影する画角を 2ω とすると, 距離 u とその物体深度 Δ から, 式(18)よりレンズの焦点距離 f と直径 D の関係が決まる. ここで, レンズのFナンバーを F にすると, レンズの焦点距離 f が式(19)より一意に決まり, 同時に, 受光面サイズ d も式(20)により求まる.

$$f = \frac{Fu(u^2 + \Delta^2) \tan \frac{2\omega}{N}}{u\Delta + F(u^2 + \Delta^2) \tan \frac{2\omega}{N}} \quad (19)$$

$$d = 2f \tan \omega \quad (20)$$

例として, $F = 1.4, 2.8, 5.6, 8, 11$ の5種類のレンズについて, 対象までの距離 u に対して, 物体深度を $\Delta = 2m$ とした場合の焦点距離 f , 受光面サイズ d の変化を図 20 に

示す。ただし、 $2\omega = 24$ 度、 $N = 702$ 画素とする。

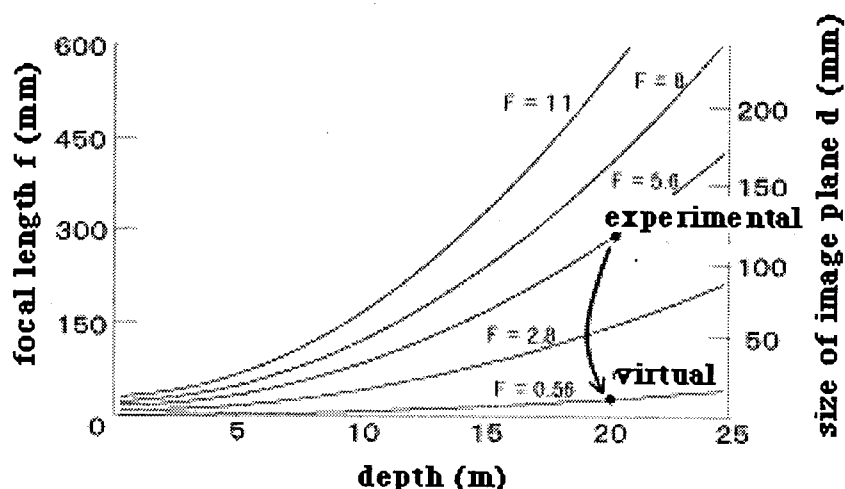


図 20: 焦点距離と受光面サイズの変化

図 20 から、合焦法で遠距離計測をするために用いるレンズと受光面の関係がわかる。受光面サイズを小さくするためには、焦点距離が長く F ナンバーが小さいレンズを用いなければならない。通常の CCD カメラの受光面サイズでは F ナンバーが極端に小さなレンズが必要になる。しかし 3.2.5 節で述べたとおり、そのようなレンズの製造・設計は容易ではない。逆にできるだけ焦点距離が長く F ナンバーが小さいレンズを選んだ場合でも、CCD カメラよりかなり大きな受光面サイズが必要になる。つまり合焦法で遠距離計測をするためには、CCD カメラよりも大きな受光面サイズが容易に得られるような観測系を用いなければならない。

3.3.1 スクリーン投影形撮像系の提案

本研究では、合焦法で遠距離計測をするための画像観測系として、図 21 のようなスクリーン投影形撮像系を提案する。スクリーン投影形撮像系では、レンズを CCD カメラに依らずに選択できる。また、レンズを通してスクリーン上で結ばれた像を CCD カメラによって撮影するため、CCD カメラで撮影されるスクリーンの範囲が受光面となる。これにより、CCD カメラで撮影できる範囲内であれば受光面サイズを大きくすることができる。

図 21 のように焦点距離 f_o のレンズを持つ CCD カメラにおいて、レンズと受光面と

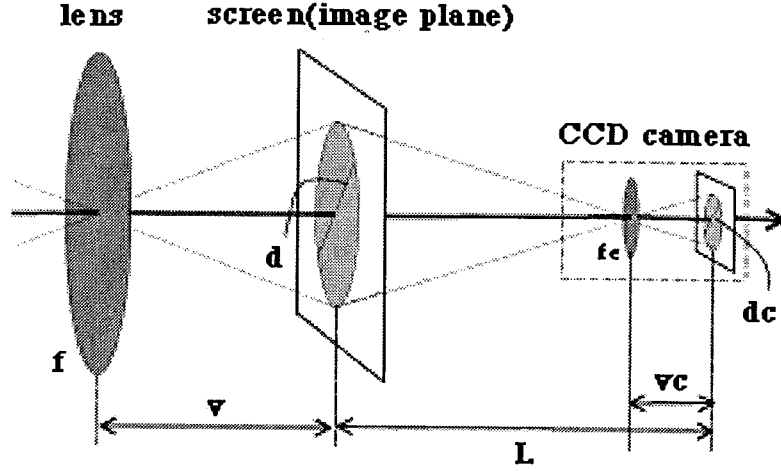


図 21: スクリーン投影形撮像系

の間の距離を v_c 、受光面サイズを d_c とすると、スクリーンと CCD カメラの受光面との間の距離 L を用いて、レンズの公式として式 (21) が成り立つ。

$$\frac{1}{f_c} = \frac{1}{L - v_c} + \frac{1}{v_c} \quad (21)$$

また、図 21 から $d : (L - v_c) = d_c : v_c$ が成り立つので、必要な受光面サイズ d が決定したとき、式 (22) より CCD カメラの受光面位置 v_c が、式 (23) よりスクリーンと CCD カメラの受光面との間の距離 L がそれぞれ求まる。

$$v_c = f_c \left(1 + \frac{d_c}{d} \right) \quad (22)$$

$$L = \frac{\left(\frac{f_c d}{d_c} + 1 \right)^2}{\frac{f_c d}{d_c} + 1 - f_c} \quad (23)$$

逆に L が決定したとき、式 (24) から v_c を、式 (25) から d を、それぞれ求めることができる。ただし $4f_c < L$ とする。

$$v_c = \frac{2f_c}{1 + \sqrt{1 - \frac{4f_c}{L}}} \quad (24)$$

$$d = \frac{L d_c}{4f_c} \left(1 + \sqrt{1 - \frac{4f_c}{L}} \right)^2 \quad (25)$$

$\frac{1}{2}$ インチの CCD カメラを用いた場合 ($d_c = 12.7\text{mm}$) の、スクリーン投影形撮像系の受光面サイズの変化を図 22 に示す。図 20 に対して十分な受光面サイズが容易に得られており、提案観測系では広い画角と浅い物体深度を得ることが十分可能であるということがわかる。そのため、提案観測系を用いれば合焦法で遠距離計測を行うことができる。しかも、提案観測系は物体深度を浅くするためのものであるから、従来の近距離計測においてはさらに精度を高めることが可能である。

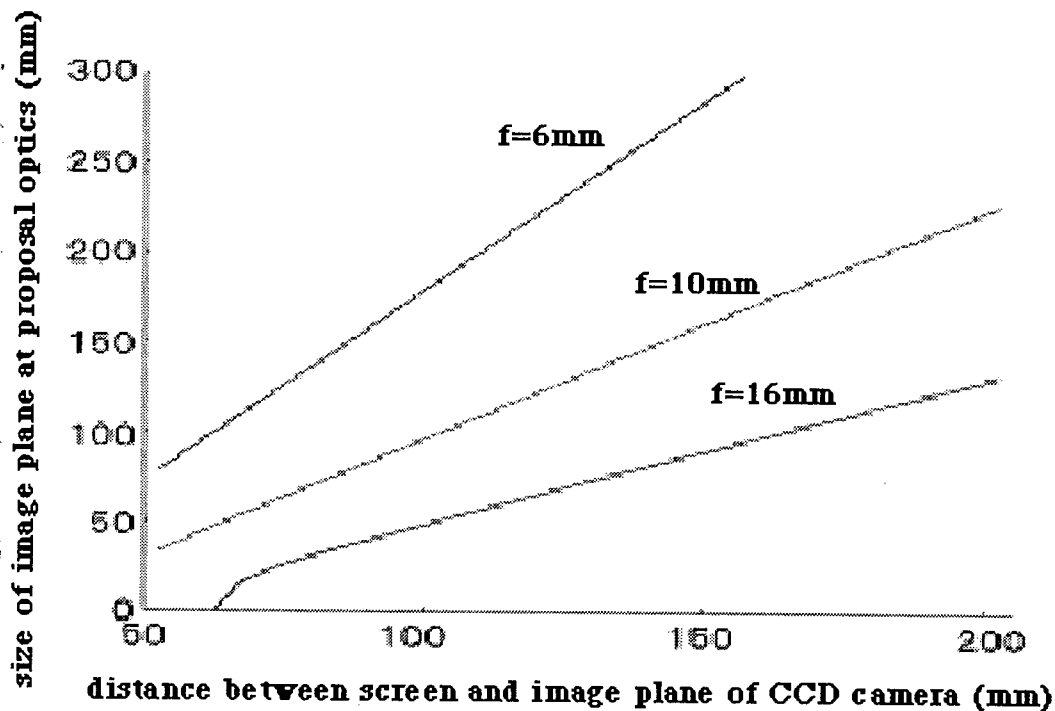


図 22: 提案観測系における受光面サイズの変化

3.3.2 提案観測系と従来方式の比較

提案観測系と同じ性能の画像観測系を、仮想的に CCD カメラで実現した場合のレンズの F ナンバー (仮想 F ナンバー) F_c について考える。まず提案観測系の受光面サイズを d 、CCD カメラの受光面サイズを d_c とし、その縮小率を $\kappa = \frac{d_c}{d}$ とする。撮影する画角を $2w$ とすると、式 (16) より提案観測系のレンズの焦点距離 f と仮想レンズの焦点距離 f_c はそれぞれ式 26、式 (27) のように表される。

$$f = \frac{d}{2 \tan \omega} \quad (26)$$

$$f_c = \frac{d_c}{2 \tan \omega} = \kappa f \quad (27)$$

また式 (19) により撮影する対象までの距離 u とその物体深度 Δ から、レンズのFナンバー F と焦点距離 f との関係は式 (28) のように表される。ただし N はCCDカメラの対角に対する画素数であり、対象は十分遠距離にあるとする。

$$\frac{f}{F} = \frac{u^2 \tan \frac{2\omega}{N}}{\Delta} \quad (28)$$

ここで u, ω, N, Δ を定数とみなすと、 $\frac{f}{F}$ も定数である。つまり、提案観測系に用いる実レンズのFナンバー(実Fナンバー)を F とすると、 $\frac{f}{F} = \frac{f_c}{F_c}$ となり、仮想Fナンバー F_c と実Fナンバー F との間に式 (29) のような関係が成立する。

$$F_c = \kappa F \quad (29)$$

3.4 一定距離範囲内に存在する物体検出

3.4.1 焦点を変えて撮影した3枚の画像による物体検出

焦点を変えた3枚の画像を撮影することにより、カメラから一定範囲内にある物体を検出する際の光学系を図23に示す。レンズから物体までの距離を u 、この物体の合焦位置にある投影面のレンズからの距離を v 、レンズの焦点距離を f とするとレンズの公式 (13) が成り立つ。図23でレンズの左側を物体側、右側を投影面側と呼ぶことにする。投影面側に投影面l、投影面mおよび投影面nを配置して画像を撮影する。この時、レンズから各投影面までの距離をそれぞれ、 v_l, v_m, v_n とする。また、これらの投影面に合焦する物体までの距離をそれぞれ、 u_l, u_m, u_n とする。

u_l と u_m の間に物体があり、レンズからこの物体までの距離を u_a とし、この物体の合焦位置のレンズからの距離を v_a とする。 u_a にある物体の像は、投影面lおよび投影面m上ではぼけている。投影面におけるぼけの大きさの広がりをぼけ円と呼びぼけ円の投影面lおよび投影面m上の直径 r_{al} および r_{am} は、レンズの直径を D とすると、式 (31) および式 (32) によって表すことができる。

$$r_{am} = \frac{|v_m - v_a|}{v_a} D \quad (30)$$

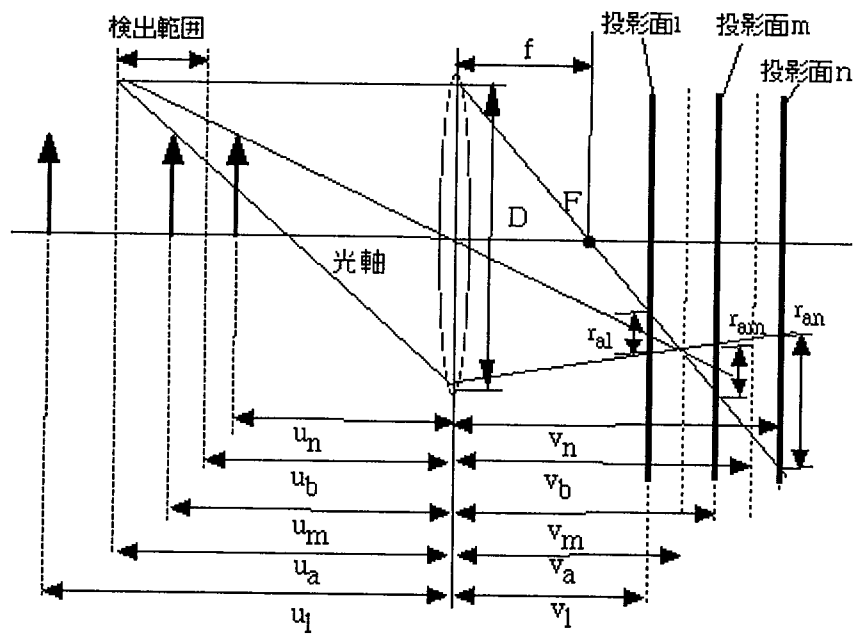


図 23: カメラモデル

$$r_{al} = \frac{|v_l - v_a|}{v_a} D \quad (31)$$

$$(32)$$

$r_{am} = r_{al}$ となる時、 u_a にある物体のぼけの大きさは投影面 f および投影面 l 上で等しくなる。この時、 u_a は式 (33) によって求めることができる。

$$u_a = \frac{f(v_m + v_l)}{v_m + v_l - 2f} \quad (33)$$

レンズから物体までの距離が u_a より大きくなると、投影面 m におけるぼけの大きさの方が投影面 l におけるぼけの大きさよりも大きくなる。一方、投影面 m と投影面 n に関しても同じ議論が成り立ち、投影面 n 上での物体のぼけの大きさを v_n とし、投影面 m および投影面 n 上でぼけの大きさが同じになる物体までの距離を u_b とすると、 u_b は式 (34) によって求めることができる。

$$u_b = \frac{f(v_n + v_m)}{v_n + v_m - 2f} \quad (34)$$

物体までの距離が u_b より小さくなると、投影面 m におけるぼけの大きさの方が投影面 n におけるぼけの大きさよりも大きくなる。この結果、物体からレンズまでの距離 u が、 $u_a < u < u_b$ の時、投影面 m におけるぼけの大きさが投影面 l や投影面 n のぼけより小さくなり、物体がレンズから $u_a < u < u_b$ の距離にあることを検出することができる。

一方、 u_a 、 u_b が最初に与えられている場合、投影面を等間隔に配置するとすれば、式 (33) および式 (34) より、投影面の位置は式 (35) ～ (37) となる。

$$v_f = -\frac{1}{2} \frac{fu_b}{u_b - f} + \frac{3}{2} \frac{fu_a}{u_a - f} \quad (35)$$

$$v_n = -\frac{1}{2} \frac{fu_a}{u_a - f} + \frac{3}{2} \frac{fu_b}{u_b - f} \quad (36)$$

$$v_m = \frac{v_f + v_n}{2} \quad (37)$$

なお、投影面上の像の大きさはレンズの焦点 F から投影面までの距離に比例して大きくなる。ぼけの大きさを比較する際には、画像の拡大縮小が必要である。

3.4.2 スクリーン投影形撮像系による物体検出

カメラから u の距離に焦点を合わせた時、 u の前後にある物体のぼけの大きさはぼけ半径 r として表現される。カメラの解像度を ϵ とした時、

$$r < \epsilon \quad (38)$$

のとき、物体の像のぼけは観測されない。物体が式 (38) を満たしながら u の前後 Δ の範囲を移動した時 Δ を物体深度と呼ぶ。物体深度は、レンズの焦点を f 、レンズの口径を D とした時、式 (39) で表すことができる。

$$\Delta = \frac{\epsilon u f D (u - f)}{f^2 D^2 - \epsilon^2 (u - f)^2} \quad (39)$$

物体深度はレンズから物体までの距離が長くなればなるほど大きくなり、物体の位置検出が困難になる。物体深度を小さくするためには、レンズの口径 D を大きくするか、焦点距離 f を長くする必要がある。しかし、焦点距離を長くすると、投影面のサイズが同じ場合、視野角が狭くなる。投影面を大きくすることは CCD の面積を大きくすることに相当し半導体の製造上、容易ではない。また、焦点距離を同じにしたままでレンズの口径を大きくすることは、レンズを構成するガラスなどの屈折率を大きくする必要があり、光学的な歪みが大きくなりやはり困難である。

そこで、筆者らはスクリーン投影形撮像系 [66] を考案した。このカメラは主レンズによって得られる像を一旦スクリーンに投影し、この像を CCD カメラによって撮影するものである。スクリーン投影形撮像系の撮像系を、図 24 に示す。

スクリーン投影形撮像系の解像度 ϵ_s は、CCD の対角線の長さを d 、スクリーン上の CCD の撮影している範囲の対角線の長さを d_s とすれば式 (41) によって表される。また、スクリーン投影形撮像系の物体深度 Δ_s は、主レンズの口径を D_s 、主レンズの焦点距離を f_s とすれば式 (41) によって表される。

$$\epsilon_s = \frac{d_s}{d} \epsilon \quad (40)$$

$$\Delta_s = \frac{\epsilon_s u_s f_s D_s (u_s - f_s)}{f_s^2 D_s^2 - \epsilon_s^2 (u_s - f_s)^2} \quad (41)$$

この撮像系では、市販の CCD カメラを用いながら、主レンズとして口径が大きく焦点距離の長いレンズを用いることができ、投影面も大きくすることが可能ため、物体深度を浅くしかも視野角も広くすることが可能である。

図 25 は CCD カメラに直接レンズを取り付けた場合とスクリーン投影形撮像系の比較である。CCD カメラに直接レンズを取り付けた場合の物体深度は式 (39) によって表される。この際想定しているのは、 $\epsilon = 0.018\text{mm}$ 、 $f = 30\text{mm}$ であり、レンズの性能を表す F 値 ($= \frac{\text{焦点距離}(f)}{\text{レンズの口径}(D)}$) を用いて表現している。

一方スクリーン投影形撮像系で想定しているのは、CCD カメラは同じものを使用し、 $d_s = 126\text{mm}$ 、 $f_s = 299.2\text{mm}$ 、口径 $D_s = 53.4\text{mm}$ である。

15m 近辺で、物体深度が 2m 程度になっており、F 値として 0.6 以下のレンズを実現出来たことになる。

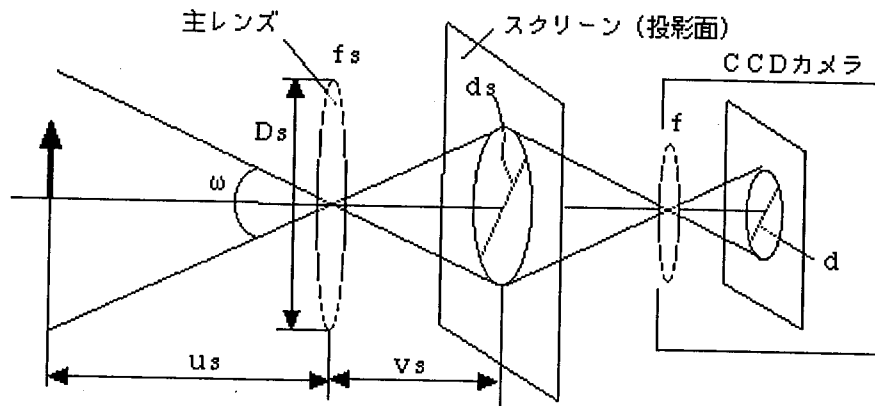


図 24: スクリーン投影形撮像系

参考のために、スクリーン投影形撮像系を用いて3枚の画像を同時に撮影する撮像系を図26に示す。主レンズを通した像をビームスプリッターによって分光し投影面に投影する。投影された像をCCDカメラで撮影する。主レンズから、スクリーンまでの距離は v_l , v_m , v_n に設定する。レンズから投影面までの距離をかえると投影面上の画像の大きさは変化するが、 $C1$, $C2$, $C3$ を調整することにより、CCDカメラで撮影する像の大きさを同じにすることが可能である[66]。

3.5 DCTを用いた合焦判定法

監視カメラやロボット用のカメラにおいて、画像を撮影すると同時に撮影対象までの距離を知ることができれば、対象物の検出を容易に行うことができる。物体までの距離を知る方法の一つとして、前述したように焦点を変えた複数の画像を撮影し、画像のぼけの大きさの差を利用する合焦法[63, 65]がある。この際、画像のぼけの大きさを測定するために、画素間の分散値を比較したり[58]、画像の持つ周波数成分を比較する。画像の持つ周波数成分を調べる方法としては、フーリエ変換を利用する方法[64]があるが、離散コサイン変換(DCT)[70]を利用することができれば、画像の圧縮方法と共通に使用することができる。これは、動画像を記録するフォーマットとしてMPEG[61, 62]等が利用され、画像の圧縮にはDCTが用いられるからである。DCTにおいては、高速に行うことのできるアルゴリズムやハードウェアが利用可能であるため距離計測を高速に行ったり、監視カメラなどのシステムを構築する際、システムの作成に必要な資源を少なくすることができる。以下、本章では画像にDCTを施し、その周波数係数を比較することにより、

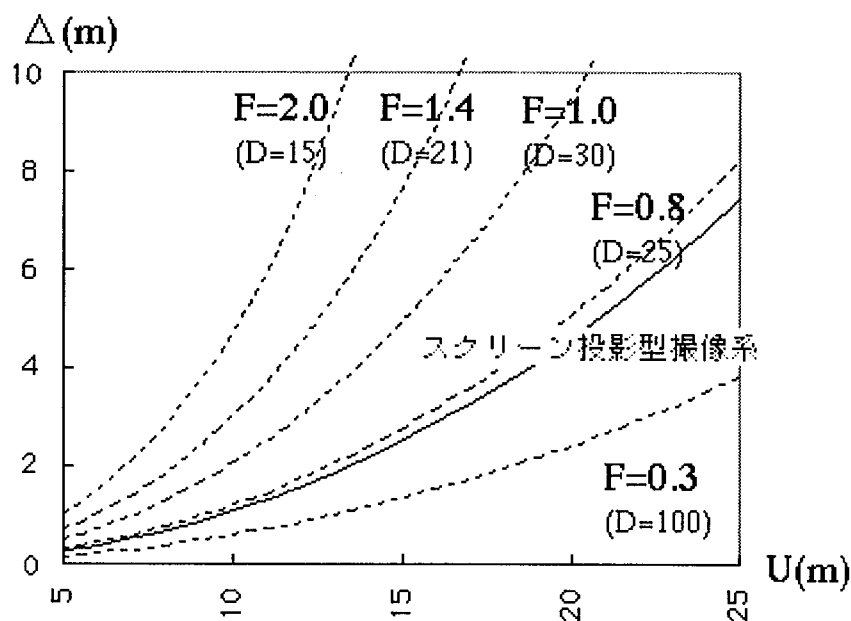


図 25: スクリーン投影形撮像系の性能例 (D の単位は mm)

画像の相対的なぼけの大きさの比較を行う。最初に、焦点ぼけによる画像の周波数成分毎の減衰の様子を、1次元関数を用いて検討する。次に、実際の2次元画像の合焦判定法について述べる。

3.5.1 DCT係数と焦点ぼけの関係

2次元画像にぼけが加わった時のぼけの大きさとDCTとの関係を1次元波形について明らかにする。関数 $f(x)$ を $0 \leq x \leq 1$ で定義される積分可能な関数とし、コサイン変換形式で表す [70] (式(43))。

$$f(x) = \sum_{n=0}^{\infty} a_n \cos(n\pi x) \quad (42)$$

$$\text{ただし } a_n = \int_0^1 f(x) \cos(n\pi x) dx \quad (43)$$

関数 $f(x)$ を周期2で半周期毎に対称な $-\infty \leq x \leq \infty$ で定義される連続関数とし、 $f(x)$ は式(45)および式(46)を満足するように拡張する。

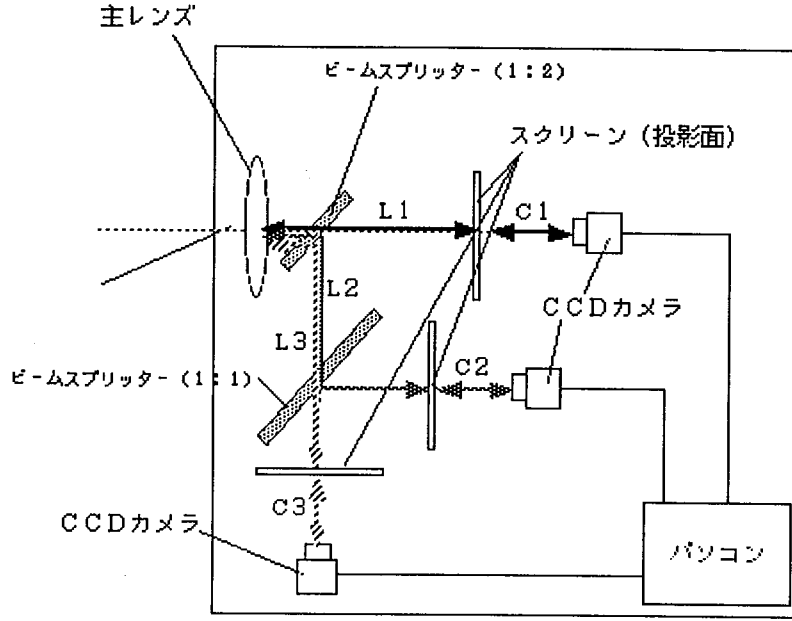


図 26: 物体検出のための光学系

$$f(x + 2L) = f(x) \quad (44)$$

$$f(x + 2L - 1) = f(1 - x) \quad (45)$$

$$(46)$$

画像のぼけの大きさは、ガウス関数 (式 (47)) との畳み込み積分で近似できる [69].

$$g(x, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \quad (47)$$

$f(x)$ と $g(x, \sigma)$ の畳み込み積分によって得られる関数を $h(x, \sigma)$ とすれば、 $h(x, \sigma)$ は式 (49) になる.

$$\begin{aligned} h(x, \sigma) &= \int_{-\infty}^{\infty} g(x', \sigma) f(x - x') dx' \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x'^2}{2\sigma^2}} \sum_{n=0}^{\infty} a_n (\cos(n\pi x) \cos(n\pi x') \\ &\quad - \sin(n\pi x) \sin(n\pi x')) dx' \end{aligned} \quad (48)$$

$$\begin{aligned}
&= \frac{1}{\sqrt{2\pi}\sigma} \sum_{n=0}^{\infty} a_n \left(\int_{-\infty}^{\infty} e^{-\frac{x'^2}{2\sigma^2}} \cos(n\pi x') dx' \right) \\
&\quad \times \cos(n\pi x) \\
&= \sum_{n=0}^{\infty} a_n e^{-\frac{(n\pi)^2 \sigma^2}{2}} \cos(n\pi x)
\end{aligned}$$

この結果、 $f(x)$ のコサイン変換係数 a_n は $a_n e^{-\frac{(n\pi)^2 \sigma^2}{2}}$ となり、 n が大きいすなわち周波数が高い程減衰が大きいことを示している。

$f(x)$ を区間 $0 \leq x \leq 1$ で $x_i = \frac{2i+1}{16}$ (ただし、 $i = 0, 1, \dots, 7$) で離散コサイン変換すれば、式(49)で $n > 7$ は意味を持たない。この結果、式(49)は式(49)となる

$$h(x_i, \sigma) = \sum_{n=0}^7 a_n e^{-\frac{(n\pi)^2 \sigma^2}{2}} \cos(n\pi x_i) \quad (49)$$

図 27 は式(49)で表される a_n をすべて等しくし、 $\sigma = 0.0$ 、 $\sigma = 0.5$ および $\sigma = 1.0$ にした結果を示している。また、図 28 はこれらの信号に DCT を施した後の、周波数成分を示している。周波数が高い程、 σ を大きくした時の減衰が大きいことが分かる。

以上は 1 次元の場合であるが、 8×8 画素で構成される 2 次元画像の場合は、式(43)、式(43)、式(49)は式(50)、式(52)、式(52)となる。

$$f(x, y) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} a_{nm} \cos(n\pi x) \cos(m\pi y) \quad (50)$$

ただし、

$$a_{nm} = \int_0^1 \int_0^1 f(x, y) \cos(n\pi x) \cos(m\pi y) dx dy \quad (51)$$

$$h(x, y, \sigma) = \sum_{n=0}^7 \sum_{m=0}^7 a_{nm} e^{-\frac{(n^2+m^2)\pi^2 \sigma^2}{2}} \cos(n\pi x) \cos(m\pi y) \quad (52)$$

3.5.2 合焦判定アルゴリズム

3.1 に示すように、周波数成分が高い程ぼけによる減衰が大きい。しかし、撮影された画像の持つ周波数成分は対象に依存するので、合焦判定に適した帯域を実験で求めることにする。ここでは、合焦判定に用いる帯域の指定方法と、対象物を検出するためのアルゴリズムについて述べる。

8×8 画素で構成される焦点を変えて撮影した 3 枚の 2 次元画像を $f^{(i)}(x, y)$ ($i = 1, 2, 3$) とし、 $f^{(i)}(x, y)$ に DCT を施して得られる係数を $a^{(i)}(n, m)$ とする。 i が大きくなるに従って焦点を合わせるカメラから物体までの焦点距離が遠いとする。そして、(53) 式で与えられる $v^{(i)}$ を計算する。

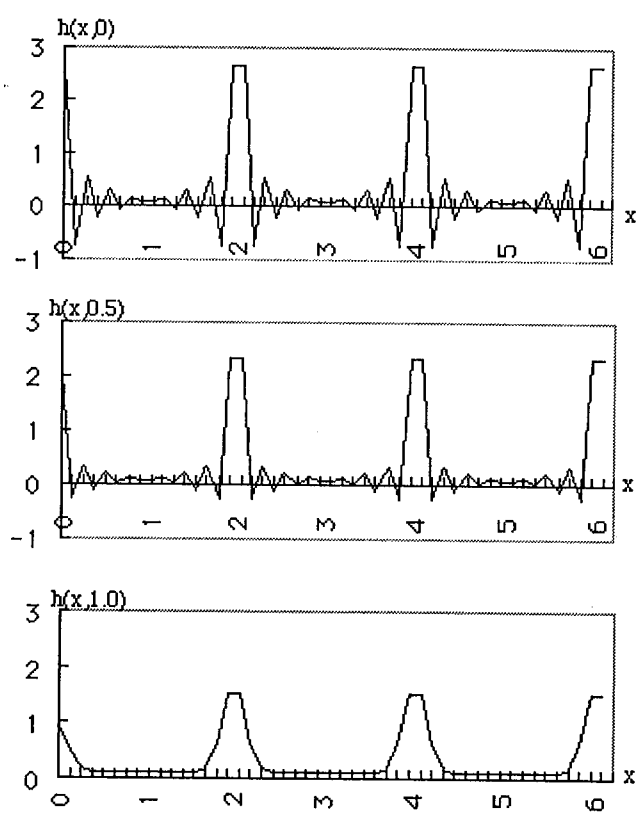


図 27: ぼけの影響による波形の変化

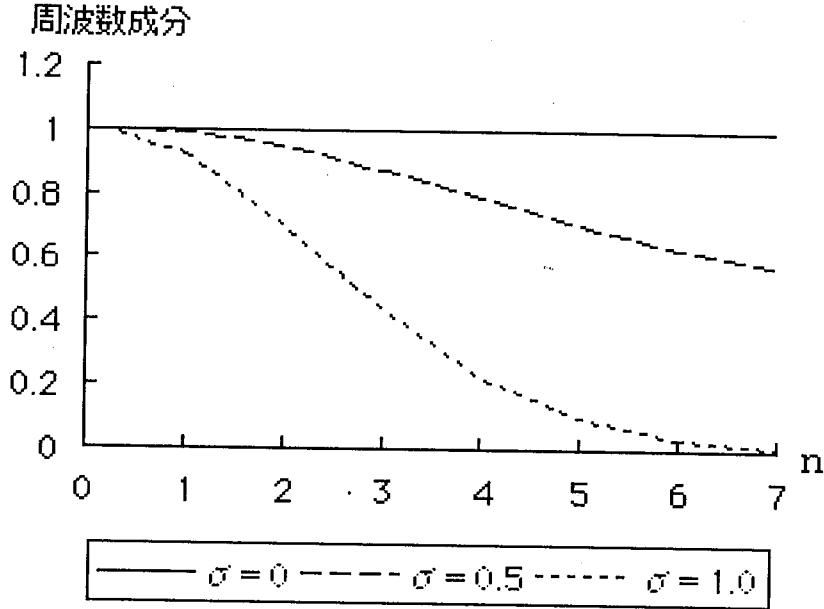


図 28: D C T 係数のぼけによる減衰

$$v^{(i)} = \sum_{n=0}^7 \sum_{m=0}^7 o(n, m) (a^{(i)}(n, m))^2 \quad (53)$$

ただし $o(n, m) = \begin{cases} o(n, m) = 1 & \text{if } p_l \leq m + n \leq p_h \\ o(n, m) = 0 & \text{other} \end{cases}$

ここで D C T 係数のうち合焦判定に用いる p_l と p_h によって指定されるバンド幅の指定方法について述べる. 図 29 は 6 4 個の D C T 係数の配列を表しており, 左上が直流成分, 右下が高周波成分である. この時, 右上から左下に向かう直線上の周波数成分 $a(n, m)$ の n, m の和は一定であり, これを p とする ($p = n + m$). p_l は直流成分に近い直線上, p_h は高周波成分に近い直線に対応している. p_l と p_h は実験によって決定する.

最後に, 3 つの $v^{(i)}$ に関して (54) 式が成り立った時, $f^{(i)}(x, y)$ として撮影された対象物が目的距離にあると判断する.

$$v^{(2)} > v^{(1)} \quad \text{かつ} \quad v^{(3)} \quad (54)$$

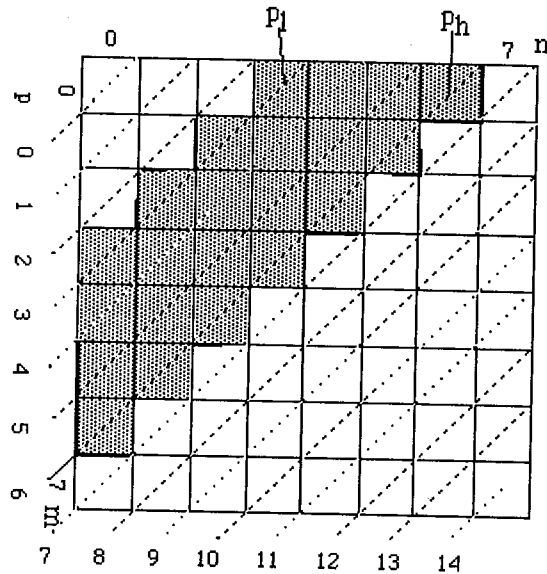


図 29: DCT 係数配列

3.6 試作実験

多重フォーカス画像からの遠距離計測が可能であることを示すために、スクリーン投影形撮像系を試作し、遠距離にある複数の物体に対して距離計測実験を行った。

3.6.1 スクリーン投影形撮像系の試作

表 1 に示す各機材を用いて実際にスクリーン投影形撮像系を試作した。その概観を図 30 に示す。試作観測系ではスクリーンと CCD カメラとの間の距離を固定しており、それらは単軸ロボット (YAMAHA エコ単シリーズ PS) によってレンズとの距離を変えることができる。また、CCD カメラの焦点位置をスクリーンに合わせてある。

今回の実験ではスクリーンと CCD カメラとの間の距離 $L = 178\text{mm}$ 、受光面サイズ $d = 126\text{mm}$ 、画角 $2\omega = 23.8$ 度とした。このとき試作観測系と $\frac{1}{2}$ インチ CCD カメラとの受光面との比の値は $K = 9.92$ となり、図 20 に示すように、式 (29) より仮想的に F ナンバーが 0.56 のレンズを用いた CCD カメラを実現したことになる。そのようなレンズを実現するのは製造上困難であり、スクリーン投影形撮像系は遠距離測定のための撮像系として適していることがわかる。

表 1: 試作したスクリーン投影形撮像系の仕様

レンズ (Nikon W300mmF5.6)	
F ナンバー	5.6
焦点距離	299.2mm
スクリーン (タチカワ フレネルレンズ)	
サイズ	202.4mm × 253mm × 2.1mm
CCDカメラ (Panasonic GP-US522HP)	
焦点距離	16mm
受光面サイズ	12.7mm

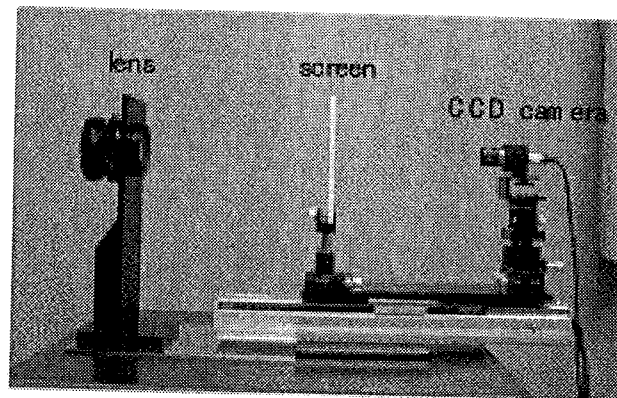


図 30: 試作したスクリーン投影形撮像系の概観

3.6.2 遠距離計測実験

10m, 15m, 20m, 25m の距離にパネルを置いた廊下を対象として、焦点位置を 10m から 25m まで動かしながら多重フォーカス画像を撮影して遠距離計測を行った。撮影した画像の概略図を図 31 に、撮影した多重フォーカス画像のうちの 1 枚を図 32 にそれぞれ示す。撮影した多重フォーカス画像はすべてグレイスケール 256 階調, 512 × 480 画素である。

多重フォーカス画像から各画素の合焦となる焦点位置を計測した結果を図 33 に示す。図 33 ではそれぞれのパネルに焦点を合わせて 4 枚の画像を撮影している。また、距離計測の際には光軸に相当する点を中心として画像を縮小することにより画素の対応づけを

行い[56], 注目画素の 5×5 近傍領域における分散値を合焦測度として用いている. ただし, 分散値の最大値と最小値との差がしきい値以下のときは計測不能とみなす. 図 33 では各画像において合焦となる画素が白色で表されており, 計測不能とみなした画素はいずれの画像においても黒色にしている.

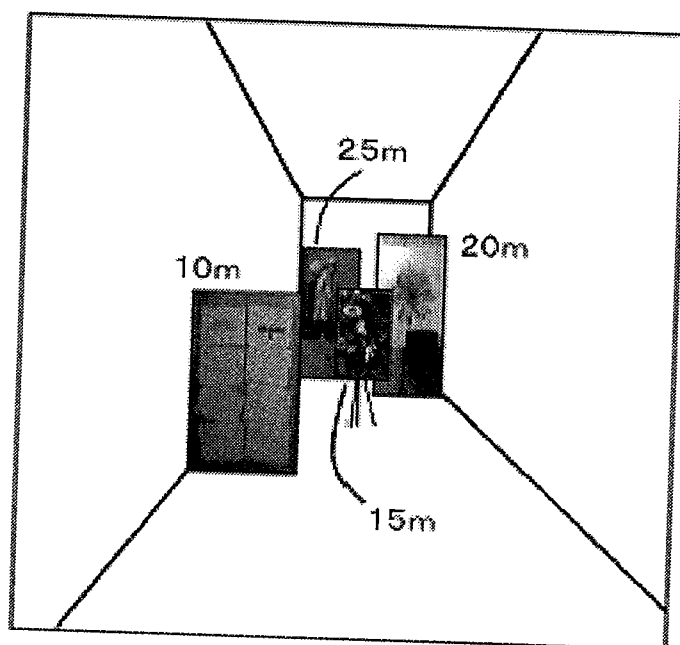


図 31: 撮影対象の概略図

図 33 のようにそれぞれのポスターが物体深度よりも離れていれば区別することができ, 合焦法で遠距離計測を行うためには物体深度を浅くすることが必要であることがわかる. そして, そのための画像観測系としてスクリーン投影形撮像系は有効であると思われる. まだ一部の直線エッジの距離測定がうまくいかず距離測定結果にノイズが残っているが, これは試作観測系のキャリブレーションがまだ不十分なためであり, 正確なキャリブレーションを行うことでこのようなノイズは除去することができる.

また, 合焦法そのものの問題ではあるが, 床・壁・天井などテクスチャのない部分では焦点ぼけが生じないために距離計測が全くできなかった. そして, 蛍光灯のような発光体や, ガラスなどによって光が反射している部分では距離計測がうまくできなかった. これは, レンズに入ってくる光が強すぎるためで, 一様に明るい場合には CCD カメラのアイリスの調整をすることにより回避できるが, 今回のように撮影対象のコントラストが大き

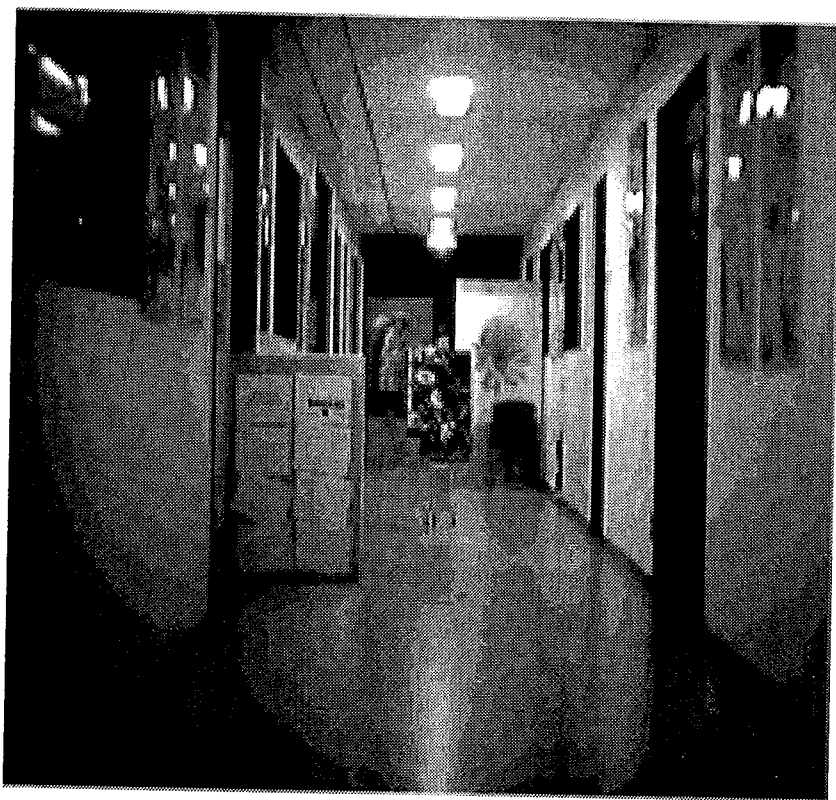


図 32: 多重フォーカス画像

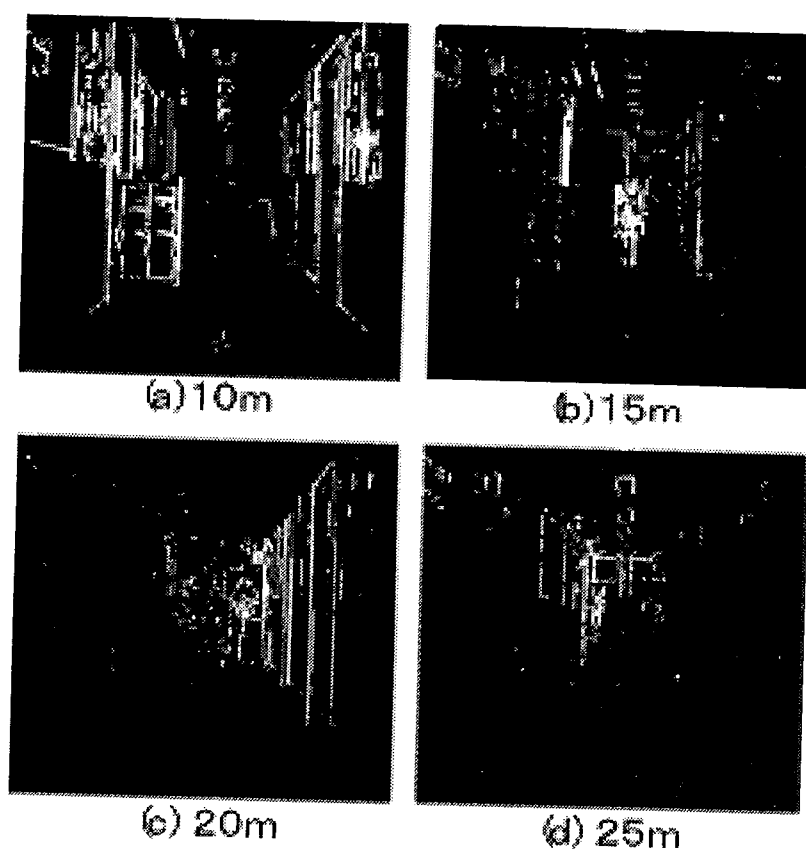


図 33: 距離計測結果

いとうまく距離計測ができない場合がある。

3.6.3 DCTを用いた物体認識実験

合焦判定に適したバンド幅を決定するために、図 30 に示したカメラを用いて図 34 に示す 2 つの目標物を目的距離において画像を撮影した。それぞれ 105 個、96 個の小領域で構成されている。

図 35 は目標物 (a) について p_l と p_h を変えながら小領域の判定を行い、小領域が目標距離範囲内と判別された時の正解率を示している。 $p_l = 3$, $p_h = 4$ の時最も正解率が高く 80% 以上である。直流成分に近い係数を用いない方が良い結果が得られているのは、ぼけによって周囲のブロックの影響が及んだ場合、直流成分に近い係数に影響があるためと考えられる。また、高周波の係数を除いた方が良い結果が得られるのは、今回撮影した対象物および撮像系では、得られた画像に高周波が含まれておらず、高周波の係数のみを使用すると、ノイズの影響などで判別に誤りが発生するためと考えられる。

同じ評価を目標物 (b) についても行い、これらを総合して $p_l = 3$, $p_h = 6$ を採用することにした。図 36, 図 37 は $p_l = 3$, $p_h = 6$ のとき、目標物の小領域が目標距離範囲内、手前、遠くのどこに判定されたかを示している。どちらの画像でも、模様の顕著でないところで判別エラーが発生している。

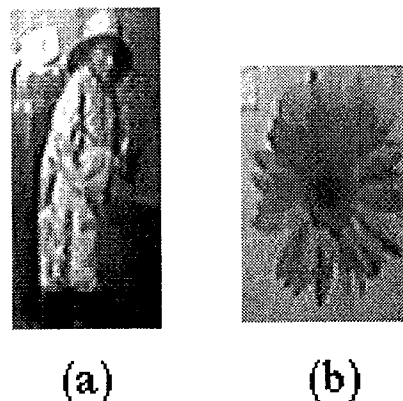


図 34: 目標物

次に廊下に物体を配置し評価した。図 41 は廊下の風景である。枠の中は目標距離範囲を示している。目標物を含む風景の中から目標距離範囲にあるものを検出したものが図 42 である。図 43 は判定結果をブロック単位に色分けしたものである。黒いブロックが目

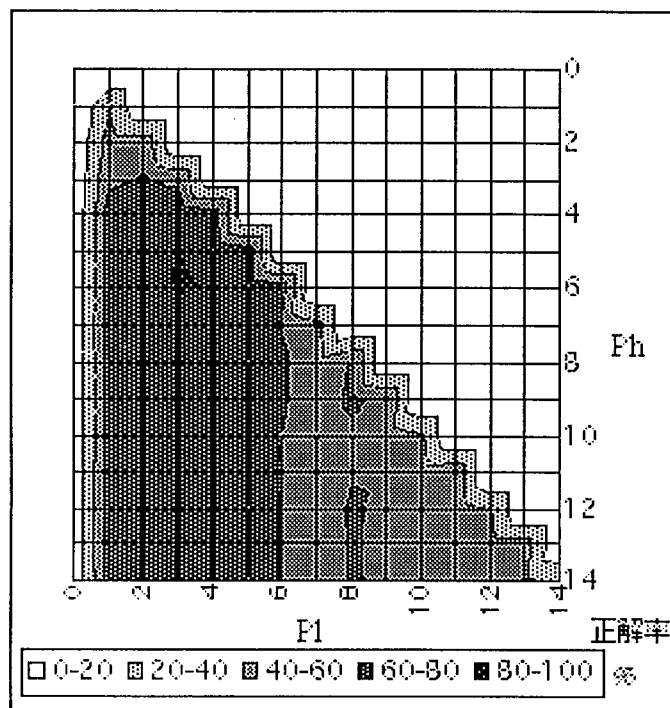


図 35: バンド幅の変化による正解率の変化 (DCT, 目標物 (a))

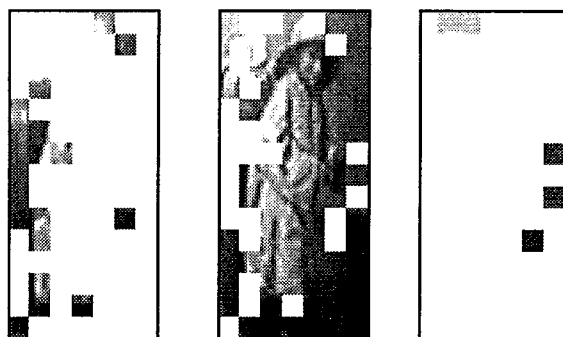


図 36: 領域抽出結果 (左: 手前, 中: 正解, 右: 遠く)

目標距離範囲にあると判定されたブロックで、白いブロックが目標距離範囲外にあると判定されたブロックである。

図 42 では、多数の目標距離範囲外のブロックが目標距離範囲と判別されている。これは、対象物がテクスチャを多く持たない時、DCT 係数の値が小さくなって誤った検出が行われていると考えられる。

そこで、 V に閾値を設け、各ブロックの 3 つの V のどれかがこの閾値を超えない時は判定から除外する。図 44、図 45 は、全てのブロックの V の中の最大値の 50% を閾値として、判定を行ったものである。図 44 では、目標距離内にある人の画像が検出されていることが分かる。図 45 で黒ブロック白ブロック以外は閾値より小さく判定されなかった部分である。判別対象となったブロックは全体の 19% であり、目標距離内の正解率は 73%、目標距離外の正解率は 76% であった。廊下の壁などで、同じ小領域の中に、距離の異なったものが含まれている場合誤認識が多くなっている。

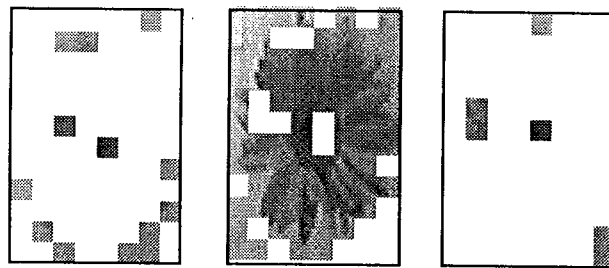


図 37: 領域抽出結果その 2 (左：手前，中：正解，右：遠く)

3.6.4 離散フーリエ変換との比較

DCT によって合焦判定が可能なことを示したが、同じ処理を離散フーリエ変換 (DFT) を用いて行い、比較した。図 38 が、目標物 (a) に対して DFT 係数を用いて合焦判定を行った結果、また、図 40 が、目標物 (b) に対して判定を行った結果である。それぞれ最終的な正解率はおなじあるが、DFT 係数を用いた場合では、高周波成分を用いても良い結果が得られている。これは、画像を周波数に分解した場合、DCT の方が情報が直流成分に近い方に片寄る [70] と考えられるが、さらに詳しい解析が必要である。

DFT との性能の比較をするために、目標物 (a) および目標物 (b) について、バンド幅を変えながら DFT 係数の実部と虚部の 2 乗和を比較して正解率を求めた。目標物 (a) の場合、最大の正解率は DCT の場合 81%、DFT の場合 80% であった。目標物 (b) では、

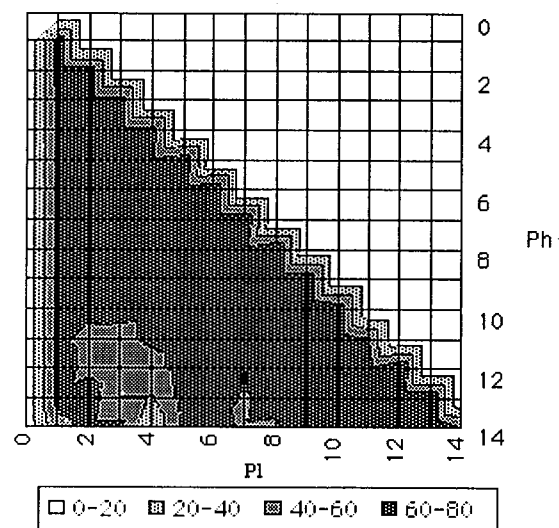


図 38: バンド幅の変化による正解率の変化 (DFT, 目標物 (a))

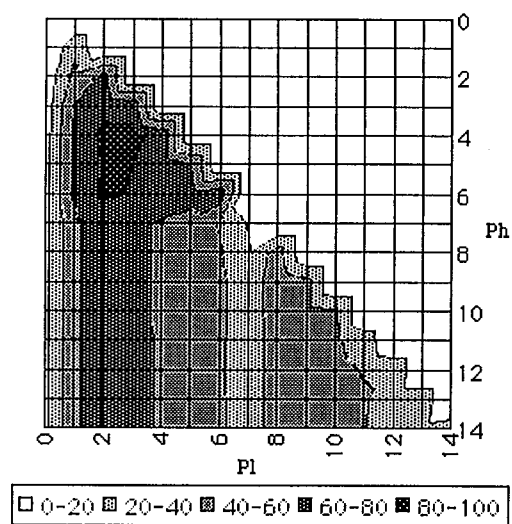


図 39: バンド幅の変化による正解率の変化 (DCT, 目標物 (b))

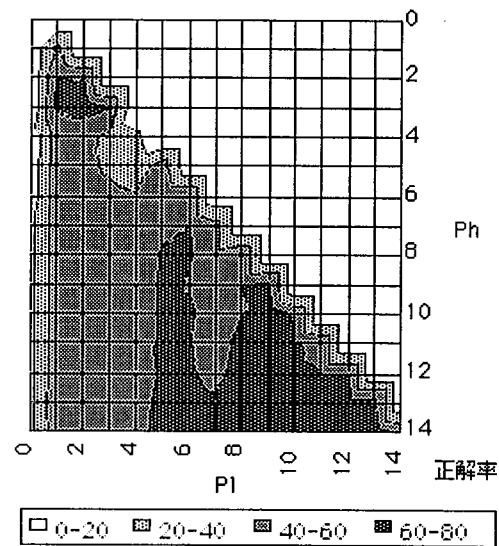


図 40: バンド幅の変化による正解率の変化 (DFT, 目標物 (b))

DCT, DFT共に77%で同じであった。

3.7 まとめ

本章では、まず、合焦法を用いて遠距離の計測を行うための観測系として、スクリーン投影形撮像系を提案した。合焦法で遠距離を計測するためには、広い画角と浅い物体深度を得ることができる観測系が必要となる。提案する観測系では、レンズで屈折した光を一度スクリーン上で結像させて背面からその像を撮影することにより、通常のCCDカメラよりも大きな受光面が得られる。そのために、焦点距離の長いレンズを用いて物体深度を浅くし、受光面を大きくして画角を広くすることができる。

そして実際に提案する観測系を試作して実験を行い、その有効性を示した。試作した観測系のキャリブレーションは不十分であったが、それでも大まかな遠距離計測ができており、スクリーン投影形撮像系のキャリブレーションの精度が十分に得られればより正確な距離の計測が期待できる。

DCT係数を用いたぼけの大きさの測定を使用して、合焦法により物体位置の検出が可能であることを示した。DCTによってぼけの大きさを測定することができれば、画像の圧縮と同時に物体までの距離推定を行うことが可能である。ぼけの大きさの測定に利用するDCTの係数を制限することにより、対象物が平面の場合は80%の精度で対象物が目的距離範囲内にあるか否かを識別できた。廊下を対象とした実験では、周波数成分を多く含まないブロックを判定から除外するために、閾値を設けることにより73%のブロックが目標距離内にあると識別出来た。また、スクリーン投影形撮像系を用いることにより、従来の光学系では不可能であった10 m以上の距離における物体検出が可能となった。

課題としては、今回の実験では、ぼけの大きさの評価に用いるDCT係数のバンド幅は撮影対象物2つの目標物についての実験で決定しているが、適切なバンド幅は撮影対象物によって異なると考えられる。そこで、撮影対象物によって評価に用いるバンド幅を可変にする方法を検討する必要がある。また、ブロック内に距離の異なったものが撮影されている場合判定の誤りが発生するため、隣接ブロックとの関連を考慮した判定が必要である。今後はリアルタイムの物体検出をするために、3枚の画像を同時に撮影する撮像系の実現をしていきたい。

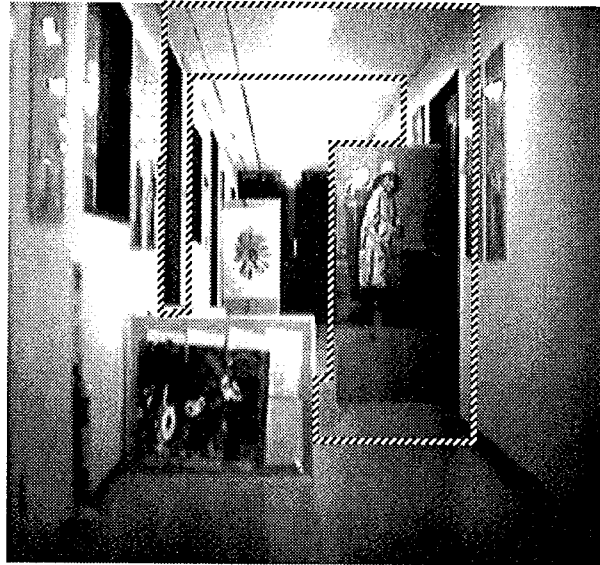


図 41: 廊下の風景

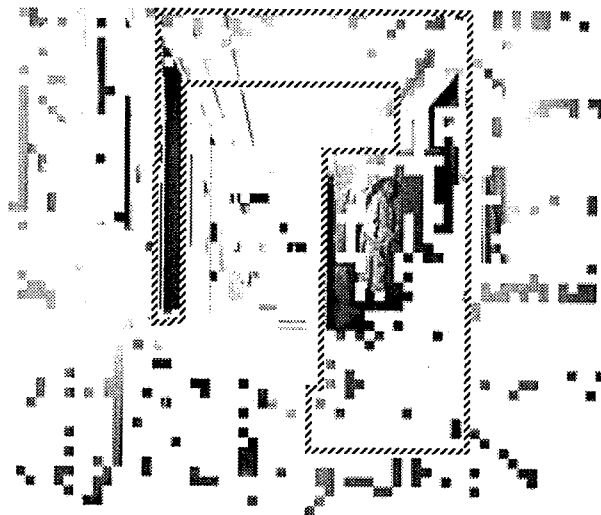


図 42: 廊下のシーンからの検出結果（閾値なし）

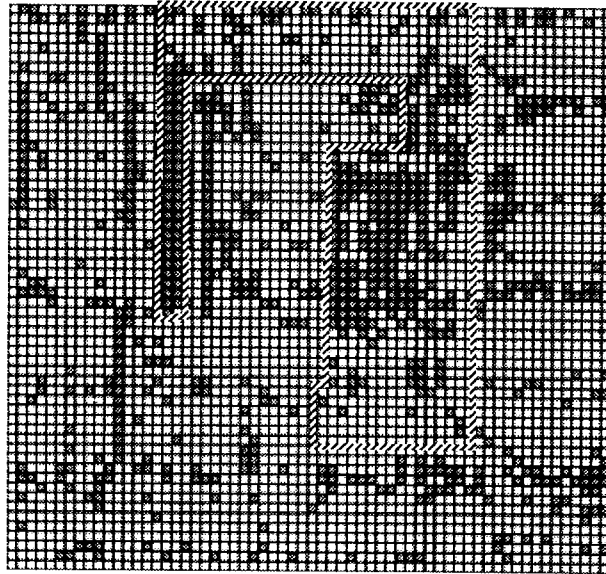


図 43: 廊下のシーンからの検出結果パターン（閾値なし）

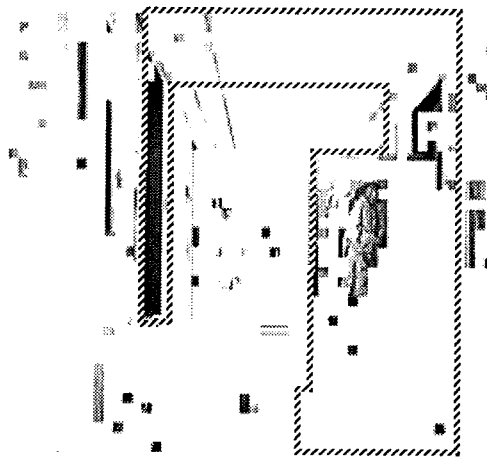


図 44: 廊下のシーンからの検出結果（閾値あり）

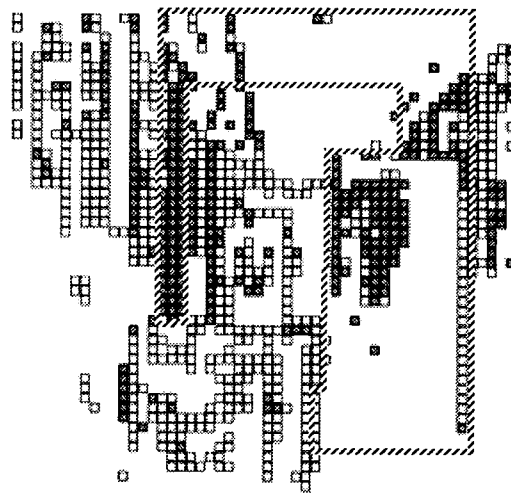


図 45: 廊下のシーンからの検出結果パターン (閾値あり)

4 多元情報を用いた映像メディアからの道路標識の抽出

4.1 まえがき

道路を撮影した動画像から標識を抽出することができれば、道路管理用の画像データベースを作成したり、カーナビゲーション用の画像を作成したりすることができる。この際、大量の動画フレームの中から標識が撮影されているものを精度良く抽出する必要がある。道路標識を色をもとに抽出する研究については、小節 2.2.5 で述べた研究があり、情景中の文字領域を抽出する研究 [30] もある。複数の動画フレームを利用する例としては、時空間画像に対してスリットを設定することにより空間中の直線成分のみを抽出する研究 [31] があるが、陸橋やビルなどの直線成分と標識の直線成分を区別することは困難である。標識は色や形等の画像としての特徴のみならず、大きさや位置が決められており空間的な特徴を持っている。このため、画像的な特徴を捉える手法と、空間的な特徴を捉える手法を組み合わせる事により、大量の動画フレームから精度良く標識の抽出を行うことができると考えられる。この際、標識の背景に木などの同じ色持つ物体があると、標識の抽出が出来ないことがある。

本章では、道路標識の持つ特徴である色とテクスチャ、および空間的な特徴に着目し、色フィルタ、検出ウインドウ、テクスチャ解析の 3 つの方法を組合せることにより大量の動画フレームの中から標識の撮影されているものを抽出する方法について述べる。また、背景に標識と同じ色を持つ物体がある場合は、距離情報を用いて標識の抽出を行う方法について述べる。

以下、4.2 節では、本研究で対象とする標識の特徴と標識抽出の方針について述べる。4.3 節では色フィルタの設計、4.4 節では距離情報の利用、4.5 節では検出ウインドウ、4.6 節では、テクスチャ解析について述べる。4.7 節の実験では、実際に道路標識を撮影した実験を行った。この際、撮像系の関係から、最初に背景に標識と同じ色を持たない路線で撮影した動画像を用い、色フィルタ、検出ウインドウ、テクスチャ解析を用いた標識抽出実験を行った。距離情報の利用に関しては、別に実験を行い、同じ色を持つ背景からの分離が可能なことを実証した。

4.2 多元情報による標識抽出法

4.2.1 標識の特徴

案内標識は、走行車線上に設置されておりF形と門形[32]である(図46)。路面からの高さは5.0mである。また、標識の縦の長さは2m前後、横幅は3mから6m程度である。

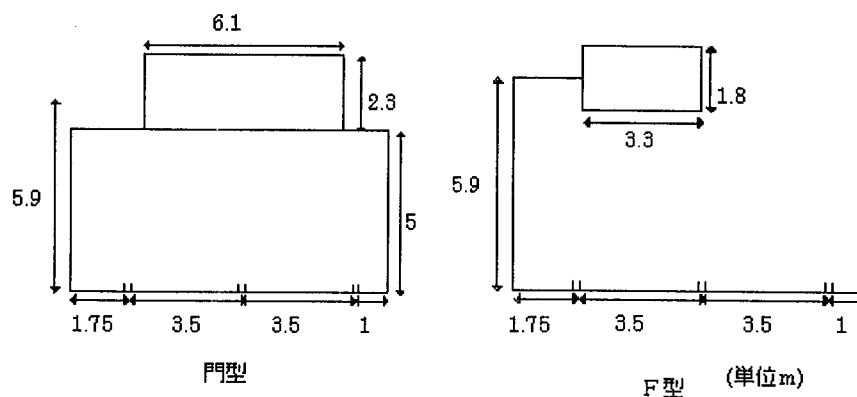


図 46: 典型的な道路標識

色や文字については表1のように決められている。案内標識では緑色が使われている。文字の大きさは文字の種類によって決められており、文字から標識の端までの間隔は、文字の大きさを基準として文字の大きさの0.5から1倍の余白が設けられている。以上のように標識は、色、空間的な位置、テクスチャという3つの特徴を持っている。

表 2: 道路交通標識に用いる表面色

標識の色	赤	(0.56, 0.39)
	黄	(0.47, 0.45)
	緑	(0.24, 0.43)
文字	漢字	30cm
	アルファベット	漢字の 1/2 から 3/4
	数字	漢字の 1 倍から 0.5 倍

4.2.2 標識抽出の課題とシステム構成

道路を撮影した動画像から道路標識を抽出する際の課題は、(1) 大量の動画フレームから標識を探し出す必要がある、(2) 同じ標識を撮影した動画フレームが多数ある、(3) 周囲の風景に標識と似通った色や形を持つ物体が存在することである。また、撮影条件によって得られる物体の色が異なるという問題がある。

標識の設置されている道路を走行すると、標識は遠くからしだいに近づき、ついには視野から消えて行く。車の進行方向に向けてカメラを搭載し道路を撮影すれば、道路上に設置されている案内標識は、画面の上部から消えて行く。標識は特徴のある色に着色されているので、画像上部を通過する画素を調べれば標識の通過を知ることができる。色情報は物体の動きを捉えやすい[39]ことに着目し、標識の色を捉えやすい色フィルタを設け、この色フィルタによって画像上部に設置した検出ウインドウ内の画素を調べる。この結果得られるヒストグラムによって、動画フレームの中から標識の位置と撮影されている大きさを推定することができる。色フィルタのみでは、標識か否かを判別することができないので、テクスチャ解析によって標識以外のものを除外する。このように、検出ウインドウ、色フィルタ、テクスチャ解析といった3つの手法を組合せれば、動画像中の標識を検出することができる(図47)。

ところが、標識と同じ色を持つ物体が背景にある場合は、この方法では標識を抽出することが出来ない。そこで、物体までの距離情報を用いて背景との分離を行う。以下、各々の解析方法について述べる。

4.3 色フィルタの設計

色の度合いは x y 色度図上で表すことができる。カメラによって観測される色の度合いは、物体の持つ色度と光源の色度の合成によって求めることができる[33, 34]。標識の色度は表1のように決められている[32]。また、野外自然物体の持つ色度分布は、図3のように分布していることが報告されている[36]。空などの野外自然物は黒体軌跡周辺に分布していることがわかる。一方、光源の色は光源の持つ温度によって決まり、色度図上では黒体軌跡と呼ばれる(1)式に沿って移動することが知られている[35, 36]。例えば、光源の特性を決めるものとして標準照明光A, B, C, D_{65} が設定されており、これらの色は黒体軌跡上にある(表3)。

$$B(x) = -3.000x^2 + 2.870x - 0.275 \quad (55)$$

この結果、カメラによって観測される標識の色は、黒体軌跡と標識や各物体の持つ色の

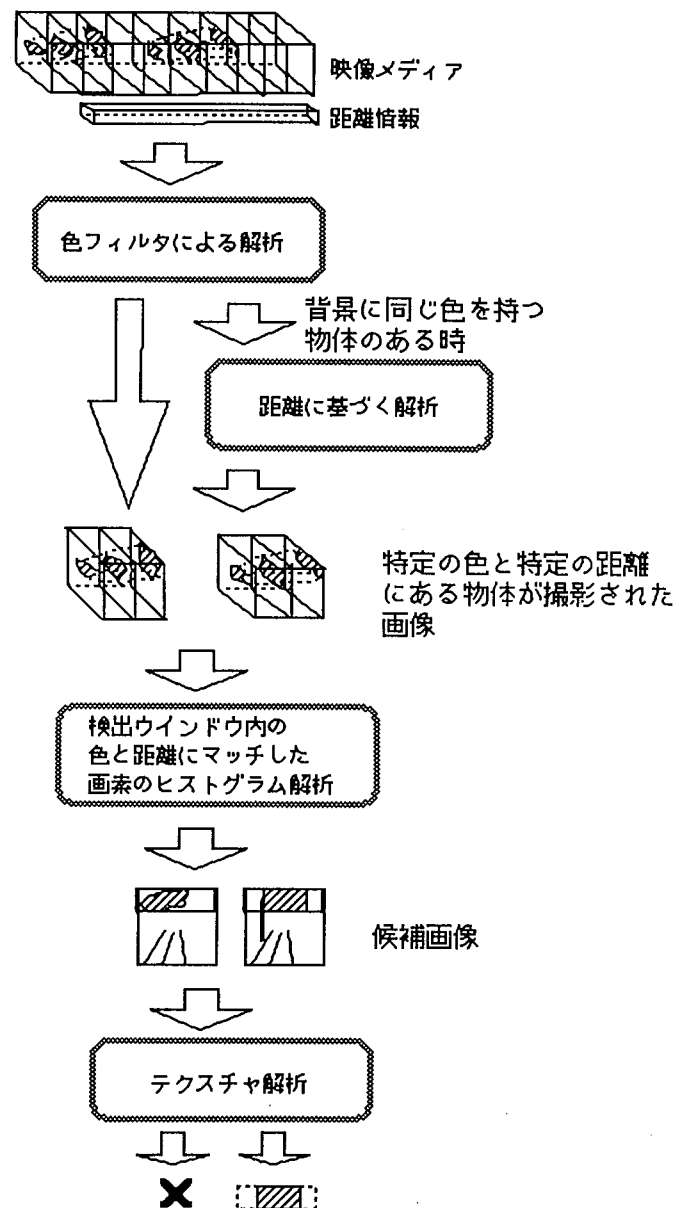


図 47: 標識抽出の手順

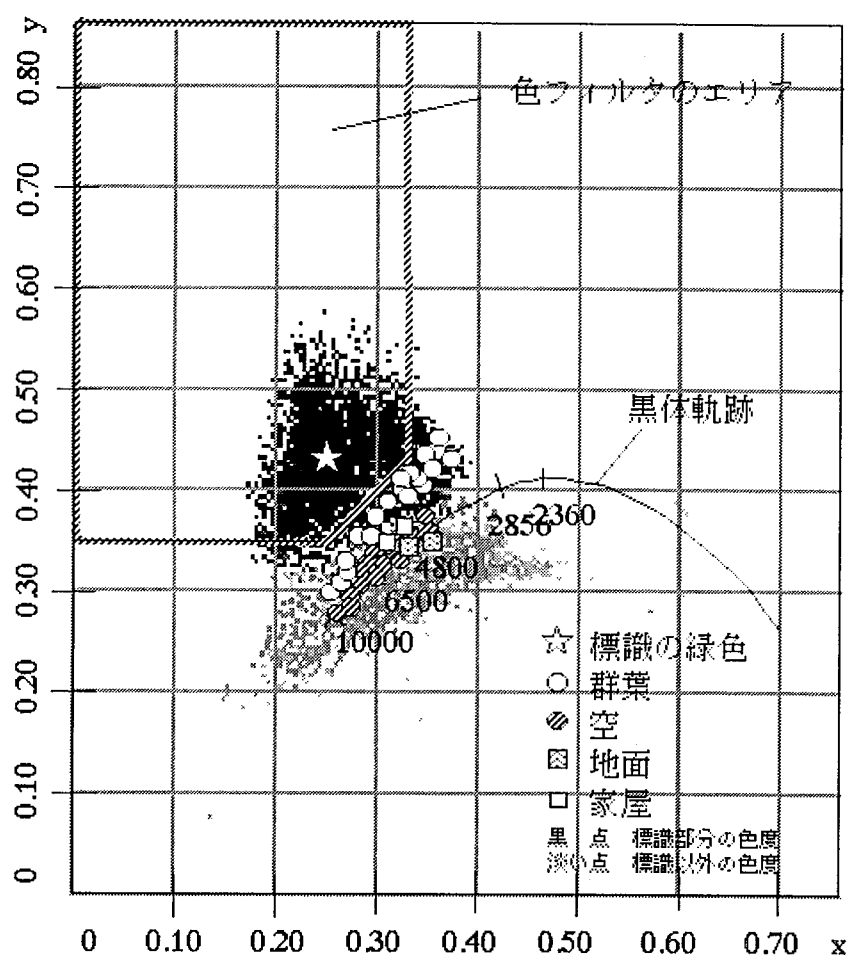


図 48: 緑色フィルタの範囲

関係で決定される。

表 3: 標準照明光

標準照明光	種別	黒体温度	x	y
A	黒体	2856K	0.4476	0.4074
B	直射日光	4870K	0.3484	0.3516
C	昼間の光の平均	6770K	0.3101	0.3163
D65	全天空からの光の平均	6500K	0.3127	0.3290

図 48 に、実際に晴れた日の昼間に実際に道路で撮影した標識が含まれる 14 枚の画像について、標識の部分と標識以外の部分の xy 色度分布を求め、その和集合を取ったものを示す。標識の領域には白い文字も含まれている。図 48 で濃淡の濃いものが標識の部分、淡いものが標識以外の部分の色度分布である。実際に観測される物体の色度は色度図上で広く分布している。標識と標識でない部分の色度は明瞭にわけることができないが、色フィルタの範囲の設定に際しては、野外自然物体の色を含まないことと、領域を単純な形にするという方針にした。以上の結果から、フィルタの領域 $F(x)$ は

$$F(x) > B(x) + 0.05$$

$$x < 0.33$$

$$y > 0.25$$

とした。

なお、曇りの日や逆光の場合は、画像の色度分布が図 3 とは異なるので、色フィルタの範囲を変えたり文献 [40] のような方法を用いる必要がある。

4.4 距離情報による背景画像との識別

図 49 に、3 章で述べた光学系と画像処理により、カメラから特定の距離にある物体のみを抽出するための装置の一つを示す。主レンズを通した光は、ビームスプリッターによって光路を分けられ、3つのスクリーンに像を結ぶ。主レンズからスクリーンまでの距離、 $L1$, $L2$, $L3$ を小節 3.4.1 のようにセットすれば、一定距離範囲にある物体の映像のみを切り出すことができる。

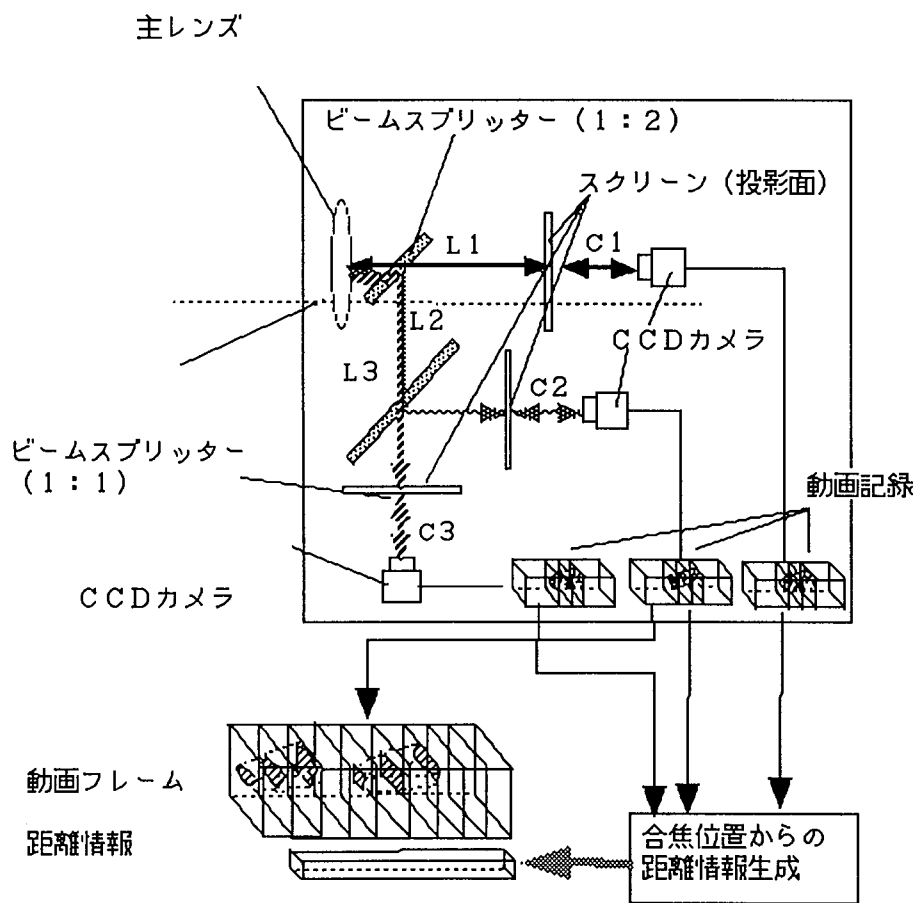


図 49: 距離情報を用いた標識の認識

小節 3.4.1 では、一定距離範囲にある物体の検出について述べたが、カメラが移動した時の物体検出方法について示したものが、図 50 から、図 53 である。図 50 に示すような目標物をセットし、カメラと目標物の距離を変え、焦点の異なる画像を取得して行く。

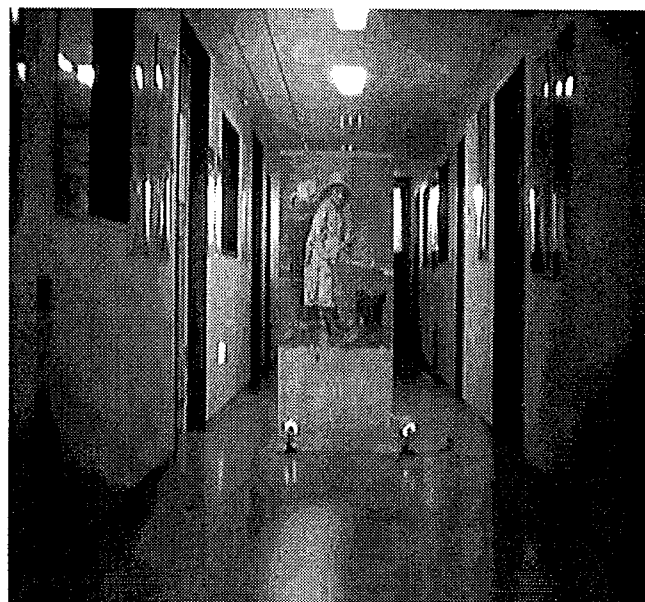


図 50: 実験で用いた目標物

カメラと物体の距離を 20 m から 6 m まで 2 m 間隔で変えながら撮影した。この時、3 枚の画像の焦点距離はカメラから 10 m から 16 m の距離の物体が検出できるようにセットしてある。図 51 はこうして得られた 3 枚の画像から、物体を抽出した結果を示している。また、図 52 はこのときの領域を示している。黒く塗られているのが目的距離物体があると判別されたブロックで、白はそれ以外の距離にあると判別されたものである。また、ブロックのない場所は 3 枚の画像の対応するブロック間で周波数成分の差が少ない部分で、判別の対象から除外してある。

カメラと目標物の距離を変えながら、目標物が撮影されている範囲で、目的距離と判別されたブロックの割り合いを調べたものが図 53 である。このことから、物体がカメラに近づいていった時、物体が目的距離に入ったことを検出できることがわかる。また、標識と同じ色をした物体が遠方にあったとしても、距離情報を用いることにより、これを除外できることがわかる。



図 51: 抽出された物体

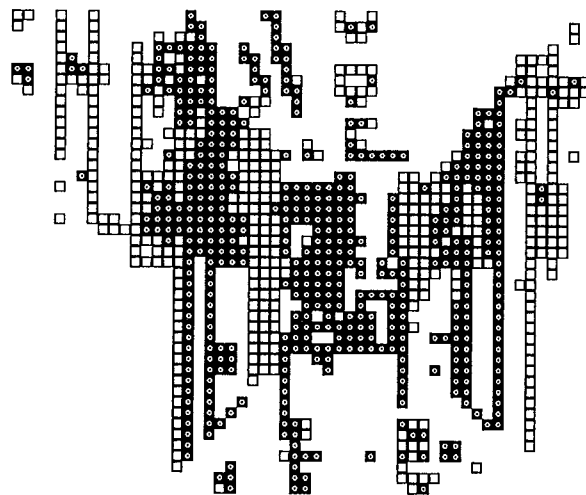


図 52: 領域抽出時のパターン

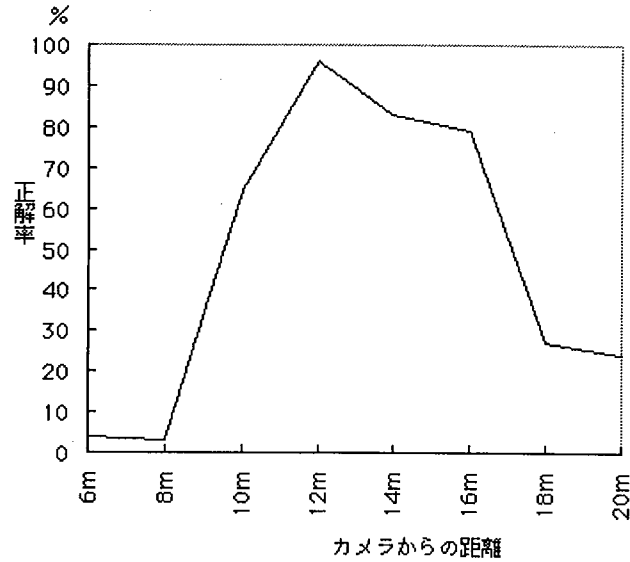


図 53: 目的距離と判別されたブロックの比

4.5 検出ウインドウの設計

標識はF形で道路の左車線の上部中央にあるものとし、車は左車線を走行しているとす
る。標識中央真下の路面に世界座標系 $X Y Z$ の原点を置き、垂直方向に Y 軸、道路と平行
でカメラの方向に Z 軸、 $Y Z$ 平面に垂直に X 軸を取る。これに対してカメラの座標系 $x y$
 z は視点に原点を取り、 z 軸は世界座標の Z 軸と等しく、 x 軸 y 軸は各々世界座標系の X
軸 Y 軸と平行に取る。視点の位置は世界座標系で $(0, 0, L)$ 、カメラの投影面は視点
から f の位置にあるとする。標識とビデオカメラで撮影した画像の関係を図 54 に示す。

この時、カメラ座標系の物体の点 $P(x, y, z)$ に対応する画像上の点 $P(u, v)$ は

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} f \frac{x}{z} \\ f \frac{y}{z} \end{pmatrix} \quad (56)$$

によって表すことができる。標識の上端までの高さを h_1 、下端までの高さを h_2 とす
る。検出ウインドウの上端と下端を画像の中心から H_1, H_2 とする (図 54)。標識の左右
の幅を u_0 とすると画像上における標識の上端、下端、および幅は $f \frac{h_1}{L}$, $f \frac{h_2}{L}$, $f \frac{u_0}{L}$ によっ
て表される。カメラが標識に近づく時、検出ウインドウに含まれる標識の面積を S とする
と、 S には 4 つの場合がある。これを S_1, S_2, S_3, S_4 で表す (図 5)。

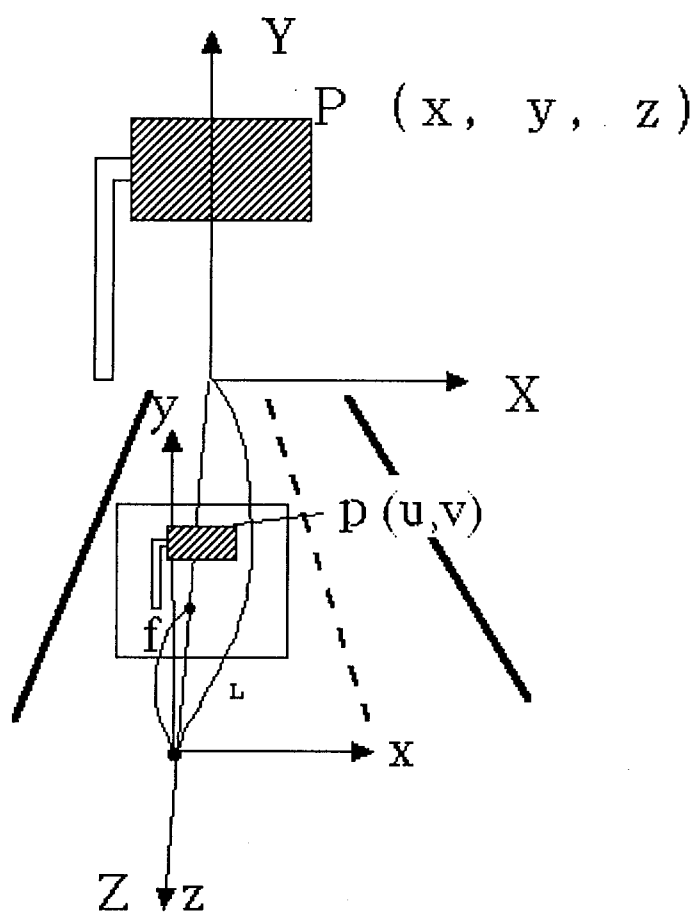


図 54: 標識とビデオカメラで撮影した画像の関係

$$S_1 = \frac{f}{Z}u_0\left(\frac{f}{Z}h_1 - H_2\right) \quad (L_2 < Z < L_1) \quad (57)$$

$$S_2 = \frac{f}{Z}u_0(H_1 - H_2) \quad (L_3 < Z < L_2) \quad (58)$$

$$S_3 = \frac{f}{Z}u_0\left(\frac{f}{Z}h_1 - \frac{f}{Z}h_2\right) \quad (L_2 < Z < L_3) \quad (59)$$

$$S_4 = \frac{f}{Z}u_0\left(H_1 - \frac{f}{Z}h_2\right) \quad (L_4 < Z < L_3) \quad (60)$$

検出ウインドウ内の標識の面積を示すヒストグラム曲線は、カメラが移動するに従って、 $S_1 \rightarrow S_2 \rightarrow S_4$ （標識がウインドウからはみだす場合、図 56）、 $S_1 \rightarrow S_4$ （標識とウインドウが同じ幅になる場合、図 57）、または $S_1 \rightarrow S_3 \rightarrow S_4$ （標識がウインドウにはいつてしまう場合、図 58）と変化する。この時 L_1 は標識がウインドウにかかり始めた時、 L_2 は標識の上端がウインドウの上端に重なった時、 L_3 はウインドウの下端と標識の下端が重なった時、 L_4 はウインドウの上端に標識の下端が重なった時のカメラから標識までの距離を示している。検出ウインドウの上下の幅は、標識の領域切り出しを考慮し検出ウインドウの幅と画像内の標識の幅が等しくなる、図 57 の場合に設定する。この時、 $S_2 = S_3$ となり、(4) および (5) 式から

$$Z(H_1 - H_2) = f(h_1 - h_2) \quad (61)$$

となる。この時のカメラの位置は標識からの距離が (8) 式で与えられる。

$$Z = \frac{fh_1}{H_1} \quad (62)$$

(7) 式において f 、 h_1 、 h_2 は定数である。このため、 Z が小さく、すなわちカメラが標識に近づけば近づくほど $H_1 - H_2$ の値すなわち検出ウインドウの幅を大きくすることができる。これは検出ウインドウの幅を広くすることに相当している。 H_1 の限界は H_0 （画像の上端と等しい）の時であり、この時 (8) 式より、 $Z = \frac{fh_1}{H_0}$ となる。検出ウインドウの上下の幅は以下の時が最適となる。

$$H_1 = H_0 \quad (63)$$

$$H_2 = H_1 - \frac{f}{Z}(h_1 - h_2) = H_0 \frac{h_2}{h_1} \quad (64)$$

検出ウインドウの横幅は、認識対象を走行車線上に設置されている案内標識としているので、画面の横幅いっぱいにする。

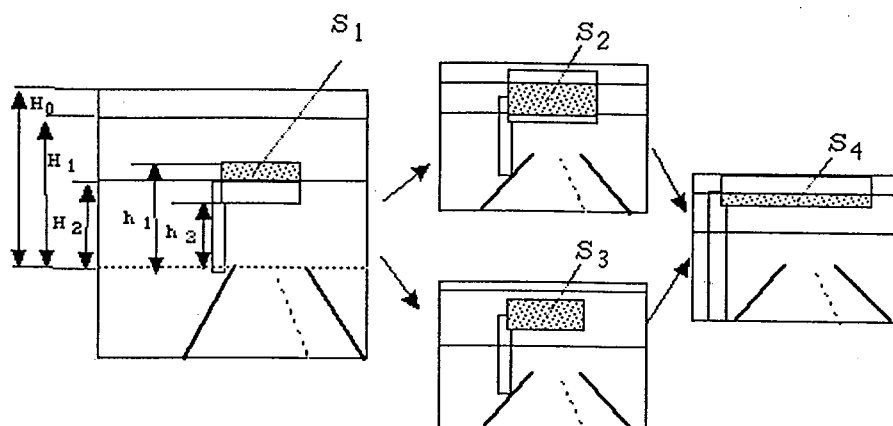


図 55: ウィンドウを通過する標識の面積

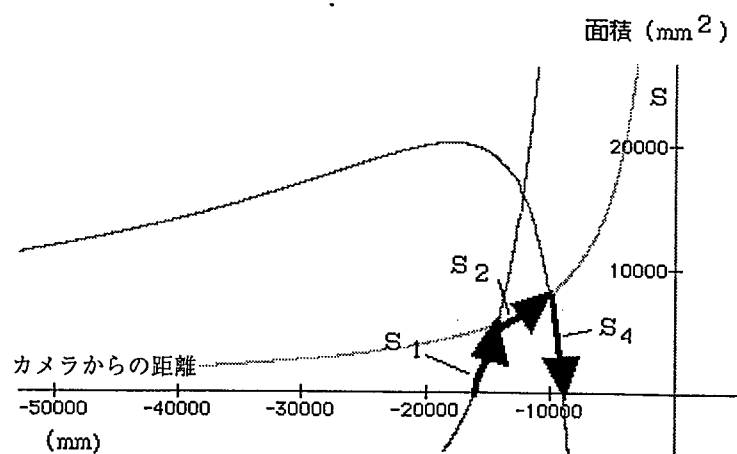


図 56: 標識が検出ウィンドウからはみ出す場合

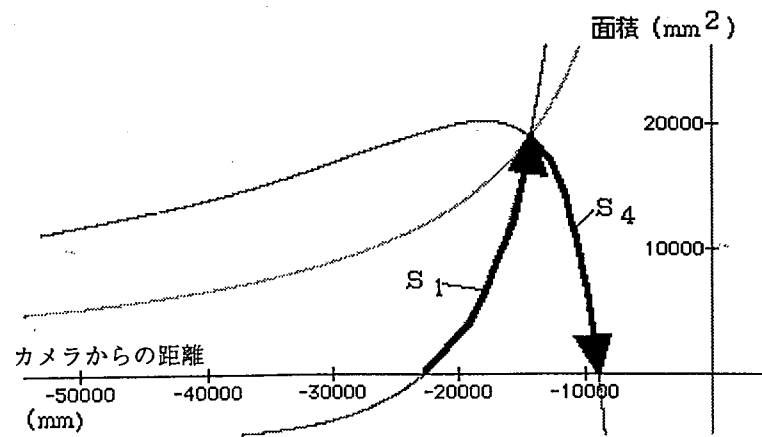


図 57: 標識と検出ウィンドウが重なる場合

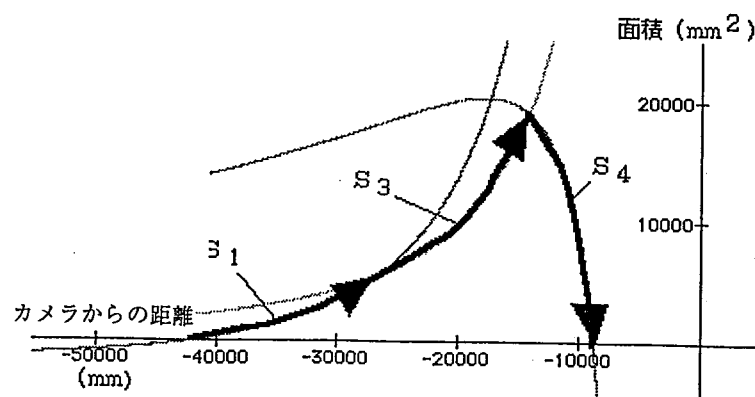


図 58: 標識が検出ウィンドウに含まれる場合

4.6 標識抽出に適したテクスチャ解析法の選択

テクスチャ解析に求められる条件としては、標識の特徴である単色の背景に文字が規則正しく並んでいるものと、そうでないものを識別できるものが望まれる。ここでは、スカンヤが開発した Surface-Shape オペレータ (SS オペレータ) を用いた解析手法を用いる [37, 38].

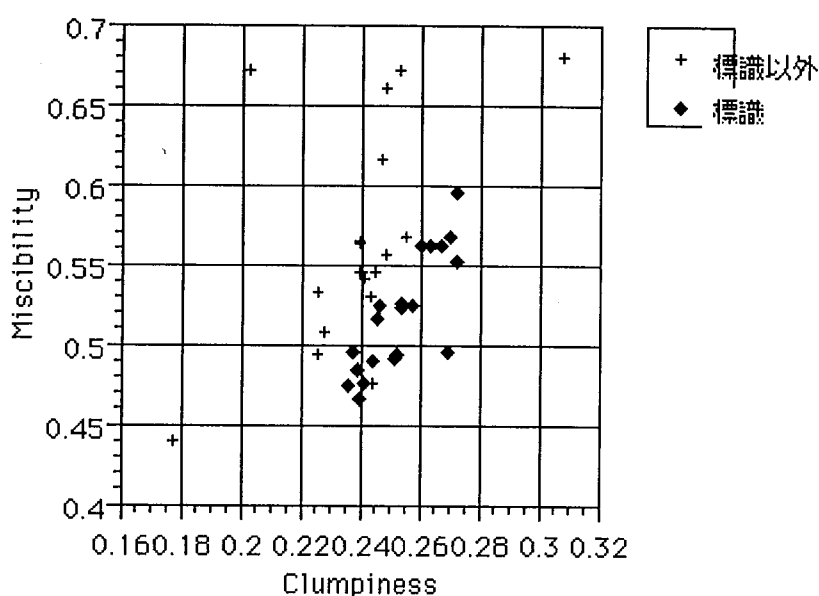


図 59: Clumpiness と Miscibility の例

この手法によれば、画像を水平、垂直、対角、反対角の4方向のそれぞれについて Clumpiness と Miscibility と呼ぶ合計8次元のベクトルによって特徴付けることができる。図59に水平方向の Clumpiness と Miscibility を求めた例を示す。マルチスケール画像を用いことによって、分類の精度を高めることができる。表73に、シングルスケールと、2階層のマルチスケールによる結果を示す。サンプルはいずれも標識および標識以外それぞれ20枚ずつである。表73で t_1, t_2 はガウス関数の分散値である。

標識の識別は、教示画像として、標識および標識以外の画像それぞれ20枚ずつについて、SSオペレータによるそれぞれの特徴量の平均値を求めておく。判別に際しては、評価すべき画像の特徴量とこの平均値とのユークリッド距離を求め、どちらのグループに近いかに基づいて行った。なお、本実験では原画像とガウシアン分散値が9のマルチスケール画像を用いた (図60)。

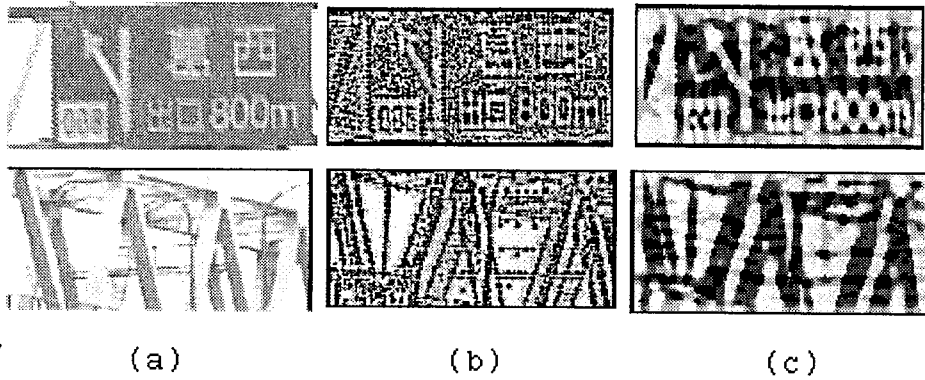


図 60: 形状作用素を適用した例 (a) 原画 (b)t=0(c)t=9

表 4: マルチスケールによる標識の分類

シングルスケール (t_1)	分類ミス	マルチスケール (t_1, t_2)	分類ミス
0	9	0, 1	6
1	7	0, 4	5
4	9	0, 9	4
9	8	0, 16	6
16	10	1, 4	8
		1, 9	7
		1, 16	6
		4, 9	8
		4, 16	8

ここでSSオペレータについて簡単に説明する．2次元の濃淡画像パターンを2変数関数 $f(x,y)$ によって構成される曲面とした時，ヘシアン（式（66））は曲面の凹凸を記述する要素として用いられる．SSオペレータは，式（66）として定義され， θ_f の大きさによって曲面の構造を知ることができる．

$$H_f = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix} \quad (65)$$

$$\theta_f = \arctan \left(\frac{\sqrt{(f_{xx} - f_{yy})^2 + 4f_{xy}^2}}{f_{xx} + f_{yy}} \right) \quad (66)$$

SSオペレータを画像全体に施したのちに，式（68）で表されるauraマトリックスを定義する． a_{ij} は画素 s の近傍を N_s ($N_s = s \pm d$)としたとき，例えば垂直方向に関しては，

式 (69) によって定義される。画像全体の特徴を表す量として、Clumpiness (式 (70)) と Miscibility (式 (70)) を定義する。

$$A = [a_{ij}] \quad (67)$$

$$a_{ij} = \text{Prob}(\theta_f(x, y) = i \wedge \theta_f(x, y \pm d) = j) \quad (68)$$

$$C = \frac{\sum_i \sum_j a_{ij}}{((i - j)^2 + 1)} \quad (69)$$

$$M = \frac{\sum_i \sum_j a_{ij}}{((i + j - n)^2 + 1)} \quad (70)$$

clumpiness と Miscibility は水平、垂直、対角、反対角の 4 方向について定義されたそれぞれ 4 次元のベクトルである。SS オペレータはマルチスケール画像に対しても施すことができる。マルチスケール画像は、Gaussian (分散値を t とする) を式 (72) とした時、式 (73) で表される。この時、マルチスケールに対応する形状作用素は式 (74) で表される。

$$G(x, y; t) = \frac{1}{2\pi t} \exp\left(-\frac{x^2 + y^2}{2t}\right) \quad (71)$$

$$L(x, y; t) = \iint f(u, v) G(x - u, y - v; t) du dv \quad (72)$$

$$\theta_L = \arctan\left(\frac{\sqrt{(L_{xx} - L_{yy})^2 + 4L_{xy}^2}}{L_{xx} + L_{yy}}\right) \quad (73)$$

$$(74)$$

4.7 道路標識の抽出実験

4.7.1 実験の概要

以上の考察に基づき実験システムの構築を行った。実験システムに用いたカメラは、1/2 インチ CCD ビデオカメラで、RGB 各 40 万画素、画像サイズ 720 k a × 480 画素の画像が得られる。用いたレンズは焦点距離 16 mm のものである。このカメラを吸盤式のカメラ台によって車のフロントガラスの内側にとりつけ、ビデオデッキと接続して撮影した。カメラの座標系が路面より約 1 m 高くなるので、その分標識の高さを差し引いて考える。この時、カメラは車の進行方向に向けて設置し、水平線が画像の中央になるようにする。撮影時のシャッタ速度は 1/1000 秒で固定しており、絞りも固定している。トンネル内は照明条件が異なるので対象外とした。

撮影は、晴れた日の昼間の午前10時から午後3時までに首都高速湾岸線および環状線の左または中央車線を走行して行った。動画フレームとしては合計61500枚ある。色フィルタは第2章で述べたものを使用した。この中には走行車線上に設置されている目視で確認できる案内標識が、合計27個ある。

最後に、距離情報を用いて背景に標識と似通った色の物体があった時の標識の抽出実験について述べる。

4.7.2 標識領域抽出アルゴリズム

入力した動画像に対して動画フレーム毎に検出ウインドウを当てはめる。検出ウインドウの幅は(10)式を元に決定するが、実験では画像上端から110画素の幅とした。検出ウインドウ内の各画素に対して色フィルタを施す。ノイズ対策のため、色フィルタに該当する画素が上下左右のいずれか2画素以上連続しているものをカウントし、ヒストグラムを作成する。ヒストグラムのピーク毎の最大値に相当する動画フレームを選択する。

標識領域の切り出しは、検出ウインドウ内で行っている。ヒストグラムを求める時カウントした画素を囲む矩形を作成する。矩形の各辺は検出ウインドウの辺と平行にしている。1つの検出ウインドウ内で複数の矩形が得られることがあるが、面積が最大のものを選ぶ。最後にこの矩形毎にテクスチャ解析を行う。以上の方法では、道路と標識が斜めの場合や、標識が2枚以上隣り合っている場合、標識が斜めに切り出されたり、1枚の標識しか認識対象とならないが、本研究では標識の撮影されている動画フレームを検出することが目的であるので、支障はない。

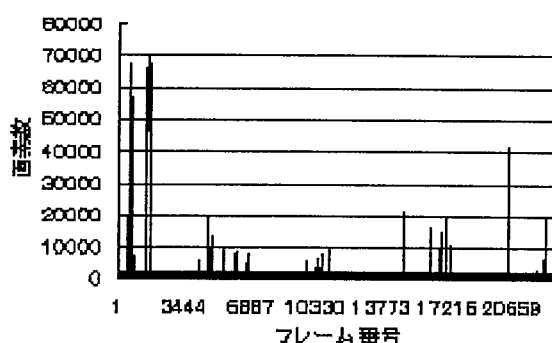


図 61: 湾岸線の動画像解析結果のヒストグラム (全体)

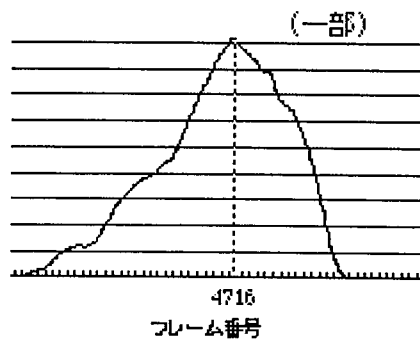


図 62: 湾岸線の動画像解析結果のヒストグラム (部分)

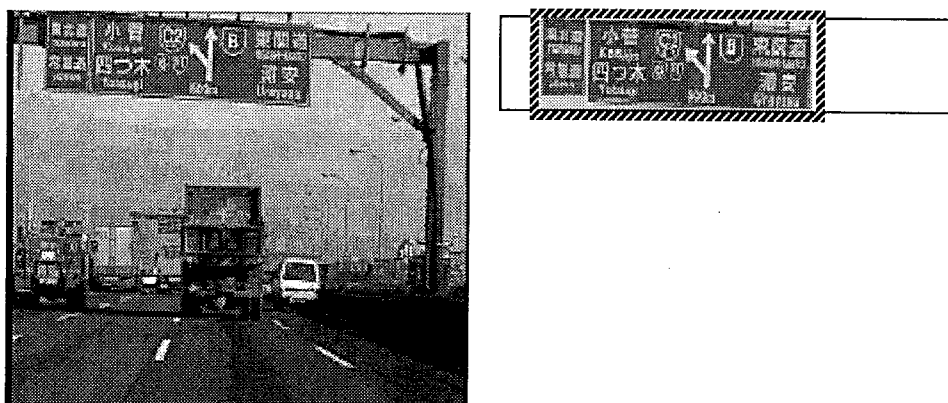


図 63: 選択されたフレームと標識

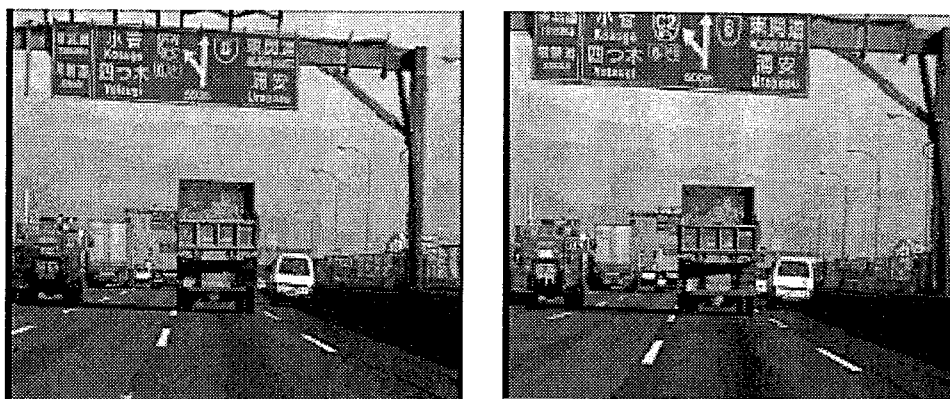


図 64: 前後の動画フレーム

4.7.3 実験結果および考察

首都高湾岸線で撮影した画像のヒストグラムを表したものを図 61（全体）に示す．このうち道路標識が撮影されている附近のヒストグラムを拡大したものを図 62 に示す．図 63 はこのヒストグラムの最大値に対応する動画フレームと切り出された標識であり，ヒストグラム解析の結果得られた画像に対して，色情報を用いた切り出し処理を行うことによって，標識の領域のみが切り出されている．図 64 に 4 フレーム前と後の動画フレームを示す．

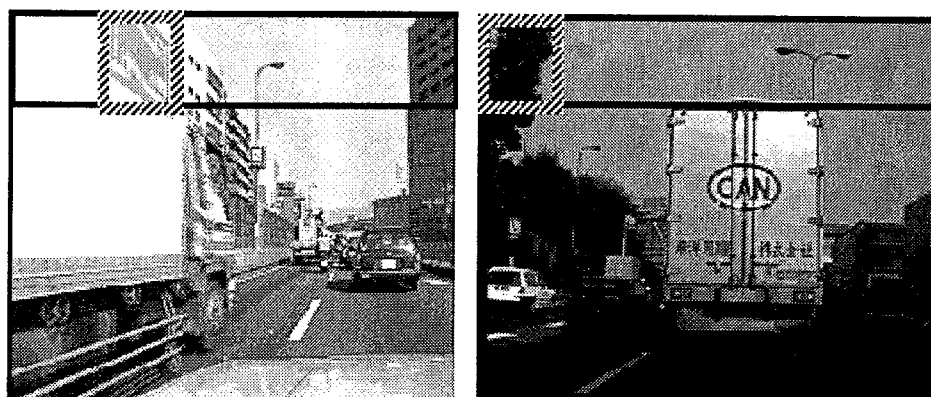


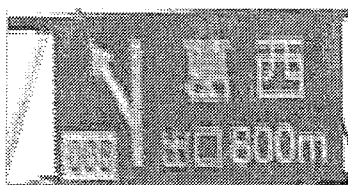
図 65: 標識以外

図 65 に標識以外の例を示す。図で太い斜線内部が切り出された部分、細い線の上部が検出ウインドウである。標識以外のものには、隣を走るトラックに書かれた広告、ビル、木などがある。

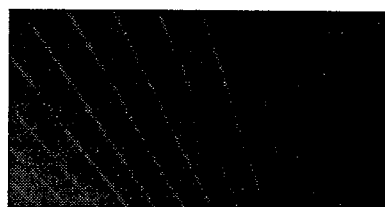
表 4 にヒストグラム解析の結果を示す。対象となる 27 枚の標識はすべて含まれている。標識の左右が欠ける場合もあったが、テクスチャ解析に影響をおよぼすほどではなかった。標識以外のものも多数含まれている。表 5 にヒストグラム解析によって抽出された画像に対してテクスチャ解析を行った結果を示す。標識と標識以外のものが精度良く分類されていることがわかる。

異なった手法を組合せる事により、大量の動画フレームの中から、精度良く標識の撮影されている動画フレームの抽出ができた。これは、色フィルタと検出ウインドウにより、標識が撮影されている動画フレームおよび領域が抽出できており、この結果、テクスチャ解析による判別結果も良くなっているためである。単独の方法、例えば、色フィルタのみでは、表 4 の結果より精度を向上させるのは難しい。また、テクスチャ解析のみでは、大量の動画フレームの中から対象となる動画フレームと領域を決定することが困難である。

本来標識であるのに、標識と認識されなかったもの（図 66 左）、標識以外なのに、標識と認識されたもの（図 66 右）（立体交差している陸橋の下側）がそれぞれ 1 枚ずつあった。図 66 左については、標識領域を切り出す際に、標識以外の部分も含めて切り出してしまったためと考えられ、切り出しのアルゴリズムを改善する必要がある。また、図 66 右のパターンは、単色の背景に直線が規則正しく並んでいるものであり、標識と同じ特徴を備えている。今回の実験では、標識以外のものを 1 つのカテゴリとして扱っているが、このような特別なパターンに対しては別のカテゴリを新たに設けることにより精度が改善されると考えられる。



(左)



(右)

図 66: 誤認識例

表 5: 認識結果

路線	含まれる標識の数	全フレーム数	抽出された フレーム数	内訳	
				標識	標識以外
湾岸線	16	23000	30	16	14
環状線 1	7	20700	21	7	14
環状線 2	4	17800	10	4	6
合計	27	61500	61	27	34

表 6: テクスチャ解析の結果

路線	抽出フ レーム数	上：標識と認識 下：標識以外と認識	判別結果		上：正当率 下：誤認識率
			標識	標識以外	
湾岸線	30	16	15	1	93%
		14	0	14	0%
環状線 1	21	7	7	0	100%
		14	1	13	7%
環状線 2	10	4	4	0	100%
		6	0	6	0%
合計	61	27	26	1	96%
		34	1	33	3%

4.7.4 距離情報の道路標識抽出への適用

スクリーン介在型カメラを路上にセットし、標識と背景の分離ができるか否かの実験を行った。図 67 は、道路脇に標識と似通った色の街路樹があるため、標識の識別が困難な場合の画像を示している。色フィルタと検出ウインドウでは、このような場合、標識の画像を切り出すことが難しい。図 68 はこの撮影位置にスクリーン介在型カメラを設置し、物体までの距離により物体の判別を行ったものであり、図 69 はこの時のブロックパターンを示している。背景との分離ができていることがわかる。なお、この実験で用いたスクリーン介在型カメラで得られる画像はグレースケールであるが、カラーカメラを用いても同様の結果が得られることは明らかである。この場面のように、背景に標識と似通った色の物体がある場合は、色フィルタと共に距離情報も用いることにより標識を抽出すること

が可能である。



図 67: 背景に標識と似通った色の物体がある場合の画像

4.7.5 他の手法との比較

以上述べてきた多元情報の組み合わせによる物体抽出方法を、単独の手法のみに頼る場合、および他の方法を用いる場合との比較を行う。

4.7.5.1 単独の手法との比較 小節 4.7.3 の表 5 で述べたように、色フィルタと検出ウインドウによっては、前フレーム数 61,500 枚の中から 61 枚の動画フレームが抽出できたが、この中には標識を含むものと含まないものが同数程度ずつ含まれている。標識と同じような立体的な位置にあり、色によって区別できない物体は、色フィルタと検出ウインドウによってはこれ以上の判別は困難である。色フィルタと検出ウインドウの組み合わせによって得られた結果に対してさらにテクスチャ解析を施したものが表 6 であった。この結果、標識以外の物はほぼ排除出来ている。

一方、色フィルタと検出ウインドウを用いなくて、テクスチャ解析のみを行う場合を考える。本研究で用いた標識の抽出を例に取れば、6 万枚以上の動画フレームに対して、標識が画像上のどの領域に存在するのを判断せねばならず、テクスチャ解析のみでは困難である事が分かる。この結果からも、本研究の提案方法である、異なった手法の組み合わせ

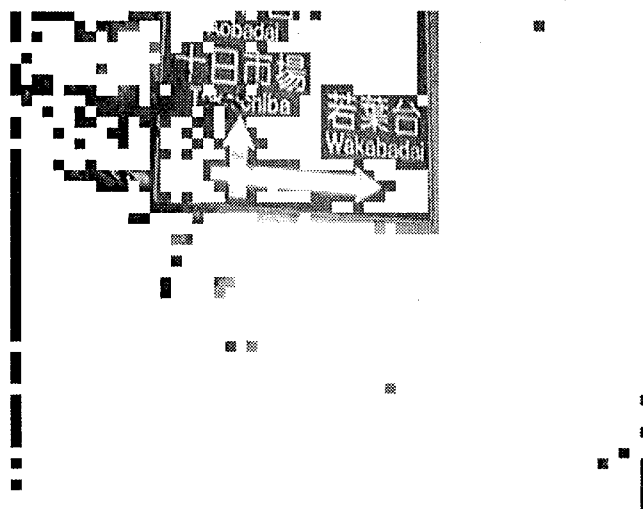


図 68: 標識部分の画像



図 69: 標識の検出結果

による物体抽出が有効に機能していることがわかる。

4.7.5.2 シーン内の文字列抽出法との比較 劉 [30] らは、シーン内の文字列を空間周波数解析などを用いて行っている。空間周波数解析による方法との比較実験を行った。図 70 および図 71 は、道路標識の撮影されている動画フレームに対して、空間周波数解析を行ったものである。



図 70: 路面上の文字と標識のある画像

この空間周波数解析では、最初に部分領域 (i, j) での離散的なフーリエ変換を式 (75) で求める。

$$F(k, l) = \sum_{i=0}^{P-1} \sum_{j=0}^{Q-1} f(i, j) \exp\left\{-j2\pi\left(\frac{k(i-1)}{P} + \frac{l(j-1)}{Q}\right)\right\} \quad (75)$$

ただし、

$$k = 0, 1, \dots, P-1; l = 0, 1, \dots, Q-1; J = \sqrt{-1}$$

この $F(k, l)$ から、式 () によって振幅で重み付けされた、部分領域の平均値を求める。

$$\omega_{sub} = 2\pi \sum_{k=0}^{\frac{P}{2}} \sum_{l=0}^{\frac{Q}{2}} \|F(k, l)\| \sqrt{k^2 + l^2} \quad (76)$$

ただし、

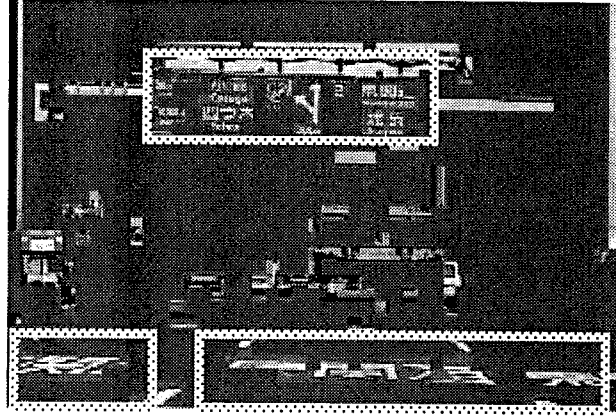


図 71: 文字領域抽出結果

$$\|F(k, l)\| = \sqrt{\{Re F(k, l)\}^2 + \{Im F(k, l)\}^2}$$

式 () の ω_{sub} はその部分領域が高周波成分をどの程度持っているかを反映している。そこで、画像全体に対して、式 (77) と式 () により ω_{sub} の平均 $\bar{\omega}$ と標準偏差 σ^2 を求める。

$$\bar{\omega} = (MN)^{-1} \sum_{m=1}^M \sum_{n=1}^N \omega_{sub}(m, n) \quad (77)$$

$$\sigma^2 = (MN)^{-1} \sum_{m=1}^M \sum_{n=1}^N (\omega_{sub}(m, n) - \bar{\omega})^2 \quad (78)$$

ただし、画像全体を $M \times N$ の部分領域に分けているものとする。

そして、以下の式によって文字を含む部分領域か否かを判別する。

$$\omega_{sub} \begin{cases} \geq \bar{\omega} - \frac{\sigma}{8} \longrightarrow \text{文字候補} \\ < \bar{\omega} - \frac{\sigma}{8} \longrightarrow \text{非文字候補} \end{cases}$$

しかし、図 71 の結果が示すように、路面上の文字まで切り出されてしまう。路面上の文字と、標識を区別することはやはり、空間周波数の解析のみでは不可能である。

4.8 道路管理用画像データベースへの適用

本研究で得られた標識認識技術を道路管理用画像データベースに適用したので、その概要について述べる。

図 72 は、画像データ管理システムを示している。撮影時には、デジタルビデオカメラを車に装着し、道路を走行しながら画像を撮影する。これとともに、カーナビの位置情報を取得し、ノートパソコンによってビデオカメラの動画フレーム番号と位置情報の対応

テーブルを作成していく。こうして得られた動画像および位置とフレーム番号の対応テーブルは、撮影終了後画像データ管理用のパソコンに読み込まれる。こうして得られた画像データベースから標識を抽出し、地図上にマッピングした様子を図 73 に示す。

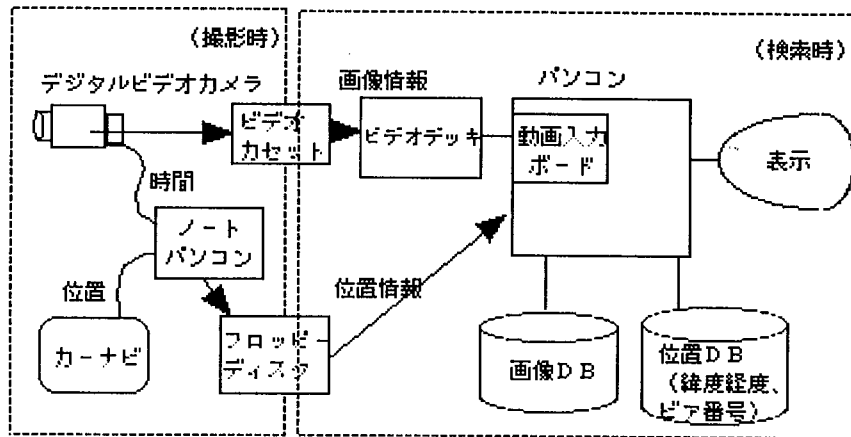


図 72: 道路管理用画像データ管理システム

道路管理においては、安全かつ迅速な対応が求められる。このシステムは、道路の現状を把握する事を目的としているが、物体の検出精度が高まれば、標識などの設備の破損状況なども自動的に検出できるようになるであろう。また、道路を定期的に走行するバスなどにカメラを設置し、画像データの更新をすみやかに行うことができれば、一般車両への情報提供も可能になる。

4.9 まとめ

緑色案内標識の持つ色、空間的な位置、テクスチャの3つの特徴に着目し、色フィルタ、検出ウィンドウおよびテクスチャ解析を組合せた緑色道路標識の認識方法について検討した。撮影条件の良い時間帯であれば、精度良く標識を抽出することができた。色フィルタと検出ウィンドウの組み合わせにより、標識の撮影されている動画フレームを検出でき、検出ウィンドウを元に標識領域の切り出しも効果的に行うことができた。テクスチャ解析により、標識と標識以外の判別も行うことができた。結果的に大量の動画フレームの中から精度良く標識を抽出することができた。

今後の課題としては、標識と似通った規則正しいパターンを排除するためにテクスチャ解析のカテゴリを増やすこと、天候や太陽光の向きにより標識の色が異なって観測される

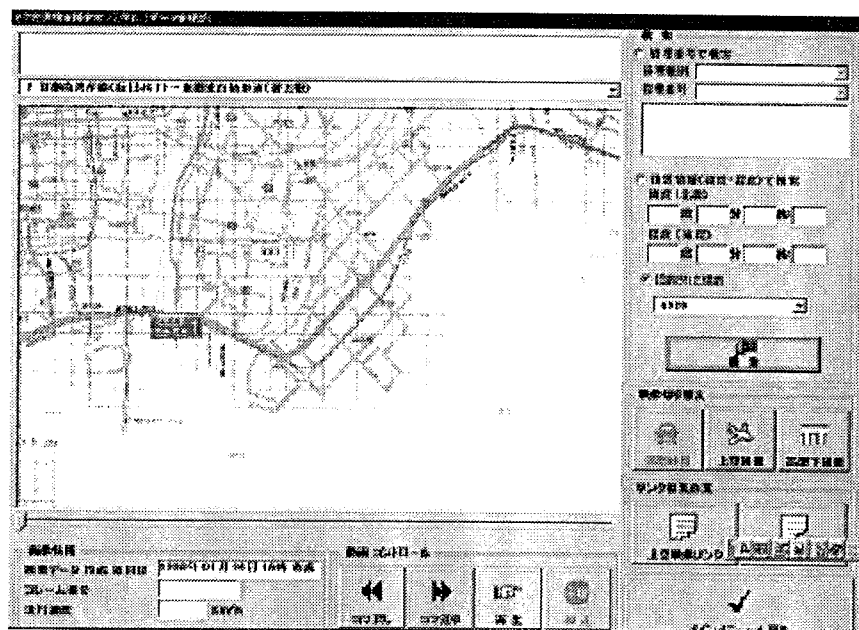


図 73: 抽出した標識を地図上に表示した様子

問題に対する対策，緑色以外の標識に対する対策，標識の領域の切り出しを正確に行うための対策がある。また，道路の凹凸が激しい場合や道路の曲がり急な場合について検討が必要である。さらに，今回は文字の書かれた案内標識を対象としているが，図柄の描かれたものについては別に検討する必要がある。

5 結論

5.1 本研究で得られた成果

本論文では、映像メディアからの道路標識抽出法について述べた。

第1章「序論」では、本研究の目的が、車に搭載したビデオカメラによって取得した映像メディアから、道路標識を抽出することであると述べた。そのためのシステムは、人の視覚に関する脳の働きを参考にして構成し、道路標識の画像的な特徴や奥行き情報、空間的な位置情報を多元的に利用すると述べた。研究の背景には、大容量の映像記憶媒体によって我々を取り巻く環境全体を記録することが可能になり、記録された物体を抽出することが不可欠な技術になることがある。

第2章「従来の研究と解決すべき研究課題」では、物体識別に関する従来手法を概観し、道路標識の識別に応用する際の問題点を指摘した。そして、映像メディアに記録された道路標識を周囲の物体から識別するために、多元的な情報を組み合わせた手法を提案することが第1の研究課題であると述べた。また、多元的な情報の一つである物体までの奥行き情報を利用するために、カメラから道路標識までの奥行き情報を取得する手法を提案することが第2の研究課題であると述べた。

第3章「合焦法を用いた遠距離物体認識の提案」では、道路標識までの奥行き情報を取得するための合焦法を利用した撮像系の提案をした。合焦法はアルゴリズムが単純である、ステレオ視などに発生する対応点問題がないという特長があるが、従来の合焦法では、道路標識を識別するために十分な視野角を保ちながら遠距離の奥行き情報を取得するのが困難であった。これは、合焦法は焦点を変えた多数の画像を撮影し、画素毎の焦点ぼけの大きさを調べ、焦点ぼけの大きさが最も小さな画像を撮影した焦点位置から物体の位置を求めるものであるが、遠距離の物体ほど焦点ぼけの変化が小さくなるため奥行き情報を得にくい。一方、焦点ぼけの変化を大きくするためには口径の大きなレンズを用いればよいが、口径の大きなレンズはレンズ素材の屈折率の限界から焦点距離が長くなり、視野角を広く保つために必要となる十分な広さの撮像面を半導体で構成することが困難だからである。提案手法は、口径の大きな主レンズによって得られる像をスクリーンに投影し、主レンズとスクリーンの距離を変えて焦点ぼけの大きさを変えながらスクリーン上の像をカメラによって撮影するもので、焦点ぼけの変化の大きな像を、広いスクリーンで得る事が可能である。この結果、標識を捕らえるための十分な奥行き情報と視野角を同時に得ることができる。そして、理論的な解析を行い実験によりその性能を実証した。さらに、焦点ぼけの大きさを測る方法として、映像メディアを蓄積する際の圧縮方法として利用される離散コサイン変換(DCT)を利用可能なことを理論的かつ実験的に実証した。これ

により、DCTを高速に実行できるアルゴリズムやハードウェアが距離測定にも利用できることを示した。

第4章「多元情報を用いた映像メディアからの道路標識の抽出」では、映像メディアから道路標識を抽出するための多元的な情報の組み合わせ手法を提案した。道路標識の周囲には、道路標識と似通った色や形をしたものがあり、道路標識と同時に撮影されてしまう。このため、色や形をもとにした解析手法のみによっては、映像メディアから道路標識を抽出することは困難である。提案手法は、物体の特徴や奥行き情報を識別する個別の機能を多元的に組み合わせて、物体の認識を行う人の脳の働きを参考にしたものであり、道路標識の特徴をとらえるための、色フィルタ、奥行き判定法、画像上の検出ウインドウ、テクスチャ解析によって構成されている。そして、まず最初に色フィルタと奥行き判定法により、道路標識と同じ色を持ち道路沿いにある物体のみを識別する。次に、検出ウインドウにより、道路標識と同じ空間的な位置にある物体のみを識別し、同時に道路標識の候補となる領域を切り出す。最後に、テクスチャ解析により道路標識のみを選択する。提案手法の有効性を実証するために実験システムを構築し、実際に道路を撮影した結果では、適合率再現率とも96%以上の精度で道路標識を抽出できた。

第5章「結論」では、本研究を総括し、本研究で得られた研究成果をまとめた上で、今後の課題、将来の展望について述べた。

5.2 今後の研究課題

映像に関する多元的な情報を用いた解析手法は、視覚に関する人の脳の働きに学んでシステムを構成したが、機能を多元的に組み合わせた段階にとどまっている。人の脳はあらかじめ知識として保持している物体に関する3次元的なイメージ像と実際に目から入った情報を比較しながら、組み合わせの方法や解析のしくみ自体を自在に変えていることが知られている。このような仕組みを組み入れることにより、光源状態が変わったり、緑地に白い文字の書かれたもの以外の道路標識や他の道路設備にも対応させていきたい。

スクリーン介在形撮像系については、性能評価の段階までを行ったが、小型化を図り操作性の向上を検討する必要がある。DCT係数を利用したボケの評価については、DFT係数を用いる場合と同程度の性能が得られる事を確認した。しかし、良い結果の得られる周波数帯域を比較した実験結果を見ると、DCTの方が狭い周波数成分で同じ性能を得られているように思われる。これは、実画像に関してはDCTの方が情報の圧縮効率が良く[70]、少ない係数成分で元の画像を復元可能な事と関連していると思われる。そこで、DCTを用いる場合の有効性をさらに検証していきたい。

謝辞

佐藤誠教授には博士課程入学以前から今日に至るまで、研究の機会を与えていただき、研究に対する考え方や進め方についても熱心に指導していただき、幾たびか困難な状況に陥った時も暖かい励ましを頂きました。心より感謝いたします。

本研究をすすめるにあたり、有益な御助言と意見をいただいた、精密工学研究所教授河原田弘博士（現桐蔭横浜大学教授）、同助教授小池康晴博士、同助手石井雅博博士に感謝します。また、像情報工学研究施設教授長橋宏博士、大学院総合理工学研究科教授廣田薫博士、像情報工学研究施設助教授長尾智晴博士、大学院総合理工学研究科助教授山田誠二博士には論文をまとめるにあたり、適切な御助言を頂きました。感謝いたします。

本研究を共同で進めていただいた、スカンヤ・ポンスーパツ博士（現タイ国マヒドル大学講師）、大原知之氏（現日本電気（株）勤務）、フュンフ・クァン・フイ・ベト氏（現修士1年）に感謝します。また、日頃から議論の相手となって下さり、研究の支援をして下さった精密工学研究所技官丸田英徳氏、本研究室OBおよび学生の皆様に感謝します。

私の所属する松下電器産業株式会社マルチメディアシステム研究所の方々には、多大な支援を頂きました。心より感謝いたします。特に日頃から励ましていただきました安川秀樹氏（同チームリーダー）、研究の機会を与えていただいた元同所長櫛木好明氏（現松下電器産業（株）取締役）、前同所長菅野昌志氏（現同経営企画室長）、同所長戸倉毅氏、同グループリーダー菊地忠一氏、同グループリーダー西川宏博士に感謝します。

また、本研究の産業分野への応用に関して研究の機会を与えていただいた鈴木正輝氏（松下通信工業（株）勤務）、道路標識の抽出に当たり議論をしていただき開発に協力していただいた矢野公嗣氏（同）、深澤宗昭氏（同）、青山智子氏（松下情報システム勤務）に深く感謝します。

最後に、常に暖かく励ましてくれた妻美智子、千尋、七海に感謝します。

参考文献

- [1] S.Zeki,J.D.G.Watson,C.J.Lueck,K.J.Friston,C.Kennard,R.S.J.Frackowiak,"A Direct Demonstration of Functional Specilalization in Human Visual Cortex", Journal of Neuroscience , 11(3),pp.641-649, Mar. 1991.
- [2] 永野俊, 梶真寿, 森晃徳, " 視覚系の情報処理", 啓学出版, 1993.
- [3] 坂口豊, " 触感覚における感覚統合と能動的認知", 電子情報通信学会誌, 76, 11, pp.1222-1227, 1993.
- [4] 中村雅司, 若松英樹, 新保孝, 梅野正義, " 車載用視覚補助システムの研究", 信学技報, PRMU97-36, pp.89-94, 1997.
- [5] " 視聴覚情報科学", A T R先端テクノロジーシリーズ, オーム社, 1994.
- [6] 馬場口登, " メディア理解による映像メディアの構造化", 信学技法, MVE-99-38, pp.39-46, July 1999.
- [7] Marr,D."Vision-A computational Investigation into the Human Representation and Processing of Visual Information", Freeman, 1982
乾敏郎, 安藤広志訳, " ビジョンー視覚の計算理論と脳内表現", 産業図書, 1987.
- [8] 西辻順一, 富永昌治, " 反射モデルのパラメータ推定法", 信学技法, PRU95-207, pp.69-74, Jan.1996.
- [9] Gudrun J. Kkinker,Steven A.Shafer,Takeo Kanade,"A Physical Approach to Color Image Understanding", Internationa Journal of Computer Vision, 4, pp.7-38, 1990.
- [10] Michael H.Brill,"Image segmentation by object color, a unifying framework and connection to color constancy", Optical Society of Ameria, vol.7, no.10, pp.2041-2047, Oct. 1990.
- [11] 服部数幸, 佐藤幸男, " スキャン式符号化法による小形高速レンジファインダ", 電子情報通信学会論文誌 vol.J76-D-II, no.8, pp1528-1535, Aug.1993.

- [12] 樋口和則, 塚田敏彦, 山本新, "表面反射が部位によって異なる物体の3次元形状計測", 映像情報メディア学会誌, vol.52, no.3, pp316-321, 1998.
- [13] Faugeras, O., "What Can Be Seen in Three Dimensions with an Uncalibrated Stereo Rig?," Proc. European Conference on Computer Vision'92 (G. Sandini, ed., Lecture Notes in Computer Science, vol.588), Springer-Verlag, pp.563-578, 1992.
- [14] Brown, L.G., "A Survey of Image Registration Techniques", ACM Computing Surveys, vol.24, no.4, pp.325-376, 1992.
- [15] 吾妻健夫, 登一生, 魚森謙也, 森村淳, 樋口和則, "マルチスリット方式ビデオレートレンジファインダの開発", 第5回画像センシングシンポジウム講演論文集, pp209-214, Jun. 1999.
- [16] Murali Subbarao and Jenn-Kwei tyan, "Selecting the Optical Focus Measure for Autofocusing and Depth-From-Focus", IEEE Transactions On Pattern Analysis and Machine Intelligence, vol.20, no.8, pp.864-870, Aug. 1998.
- [17] Shree K. Nayer, Masahiro Watanabe, Minori Noguchi, "Real-Time Focus Range Sensor", IEEE Trans. On Pattern Analysis and Machine Intelligence, vol.18, no.12, pp.1186-1198, Dec. 1996.
- [18] 日浦慎作, 松山隆司, "構造化瞳を持つ多重フォーカス距離画像センサ", 電子情報通信学会論文誌 (D-II), vol-J82-D-II, no.11, pp.1912-1920, Nov. 1999.
- [19] 内藤将彦, 児玉和也, 相澤清晴, 羽鳥光俊, "複数の異なる焦点画像からの焦点外れ画像の生成を利用した全焦点画像の強調的取得", 電子情報通信学会論文誌 (D-II), vol.J97-D-II, no.6, pp.1046-1053, Jun. 1996.
- [20] 高木幹雄, 下田陽久, "画像解析ハンドブック", 東京大学出版会, 1991.
- [21] 上田博唯, 宮武孝文, 炭野重雄, 長坂晃朗, "動画像解析に基づくビデオ構造の視覚化とその応用", 電子情報通信学会論文誌 (D-II), vol.J76-D-II, no.8, pp.1572-1580, Aug. 1993.
- [22] Shi-Fu Chang, William Chen, Horace J. Meng, Hari Sundaram, Di Zhong, "VideoQ: An Automated Content Base Video Search System Using Visual Cues", ACM Multimedia 97, Seattle Washington, pp.313-324, 1997.

- [23] 阿久津明人, 外村佳伸, ” 投影法を用いた映像の解析方法と映像ハンドリングへの応用”, 電子情報通信学会論文誌 D-II, no.5, pp.675-686, May 1996.
- [24] 長坂晃朗, 宮武孝文, ” 時系列フレーム特徴の圧縮符号化に基づく映像シーンの高速分類手法”, 電子情報通信学会論文誌 D-II, vol.J81-D-11, no.8, pp.1831-1837, Aug.1998.
- [25] 有木康雄, ” DCT特徴のクラスタリングに基づくニュース映像のカット検出と記事切り出し”, 電子情報通信学会論文誌 D-II, no.9, pp.2421-2427, Sep. 1997.
- [26] Robert C. Bolles, H. Harlyn Baker, David H. Marimont, ”Epipolar-Plane Image Analysis: An Approach to Determining Structure from Motion”, International Journal of Computer Vision, 1, pp.7-55, 1987.
- [27] Mark Tabb, Narendra Ahuja, ”A Multiscale Region-Based Approach to Image Matching”, IEEE, pp.415-419, 1994.
- [28] 日比哲也, 鎌田正樹, ” 色相彩度変換と座標変換を用いたカラー自然画像からの交通標識領域の抽出および認識法”, 電気学会論文誌D, vol.115, no.12, pp.1484-1490, 1995.
- [29] 薮木登, 三木成彦, ” 道路情景中の最高速度標識の検出および認識”, 電子情報通信学会論文誌 (D-II), vol.J77-DII, no.7, pp.1393-1394, July 1994.
- [30] 劉詠梅, 山村毅, 大西昇, 杉江昇, ” シーン内の文字列領域の抽出について”, 電子情報通信学会論文誌 (D-II), vol.J81-D-II, no.4, pp.641-650, Apr. 1998.
- [31] 李春曉, 全 火丙 東, 坂内正夫, ” 双曲線スリットによる平面時空間画像の生成と3次元情報の獲得”, 電子情報通信学会論文誌 (D-II), vol.J81-D-II, no.6, June 1998.
- [32] (社) 日本道路協会, ” 道路標識設置基準同解説書”, Jan. 1987.
- [33] 富永昌治, 岡山敏之, ” ハイライトや相互反射を含むカラー画像の解析法”, 電子情報通信学会論文誌 (D-II), vol.J80-D-II, no.6, Jun. 1997.
- [34] 富永昌治 今村淳志, ” 複数の不均質物体に対するカラー画像の解析”, 情報処理学会論文誌, vo.35, no.1, pp.79-91, Jan. 1994.

- [35] 池田光男, "色彩工学の基礎", 朝倉書店, 1980.
- [36] 日本色彩学会編, "色彩科学ハンドブック", Feb. 1975.
- [37] スカンヤポンスーパツ, 高松亮, 佐藤誠, "形状作用素とそのテクスチャ解析への応用", 映像情報メディア学会誌, vol.51, no.10, pp.1730-1737, Oct.1977.
- [38] P. Sukanya, R. Takamatsu, M. Sato, "The Surface-Shape Operator and Multiscale Approach for Image Classification," IEICE Trans.Fundamentals, vol.E81-A,no.8,pp.1683-1689,Aug. 1998.
- [39] T.D.Grove, K.D.Baker, T.N.Tan,"Color Based Object Tracking", ICPR'98 Proceedings, 14th international Conference on Pattern Recognition August 16-22, pp.1442-1444,1998.
- [40] 大矢百合, 小尾高史, 山口雅浩, 大山永昭, 小宮康宏, "正確な色再現手法と遠隔医療への応用", 映像情報メディア学会技術報告, May 1998.
- [41] Yong-Jian Zheng, Werner Ritter, Reinhard Janssen,"An Adaptive System for Traffic sign Recognition", Proceedings of Intelligent Vehicle'94 Symposium IEEE Ind. Electro. Sci. pp.165-170, 1994.
- [42] Jung Hak An, Tae Young Choi,"Dominant Color Transform and Circular Pattern Vector for Traffic Sign Detection and Recognition", IEICE Trans. Fundamentals, vol.E81-A, no.6, pp.1128-1135, Jun. 1998.
- [43] W.Ritter, F.Stein and R.Janssen,"Traffic Sign Recognition Using Colour Information", Mathematical and Computer Modelling, vol.22, no.4-7, pp.149-161, 1995.
- [44] Arturo de la Escalera, Luis E. Moreno, Miguel Angel Salichs, Jose Maria Armingol, "Road Traffic Sign Detection and Classification", IEEE Trans. on Industrial Electronics, vol.44, no.6, pp.848-859, Dec. 1997.
- [45] Sang-Kyun Kim,"An Efficient Roadsign Detection and Recognition Algorithm", Part of the SPIE Conference on Enhanced and Synthetic Vision, pp.336-347, SPIE vol.3364, 1998.
- [46] Lluís Pacheco, Joan Batlle, Xevi Cufí,"A New Approach to Real Time Traffic Sign Recognition Based on Colour Information", proceedings of the intelligent vehicle '94 symposium . IEEE Ind.Electron.Soc, pp.339-344, 1994.

- [47] D.M. Gavrilu,"Multi-feature Hierarchical Template Matching Using Distance Transforms", Proceedings Fourteenth International conference on Pattern Recognition, IEEE Comput. Soc., pp.439-444, 1998.
- [48] T. Kanade, P. J. Narayanan, P. W. Rander, "Virtualized Reality : Concepts and Early Results, " IEEE Workshop on the Representation of Visual Scenes, pp.69-76, Jul. 1995.
- [49] 中嶋正之, 小西宏, 高橋裕樹, "連続画像を用いた風景画像からの3次元モデルの生成, " 信学技報, IE97, no.IE97-23, pp.33-40, Jun. 1997.
- [50] S. K. Nayer, "Shape from Focus System, " Proc. of IEEE Conf. Computer Vision and Pattern Recognition, pp.302-308, 1992.
- [51] M. Subbarao, G. Surya, "Depth from Defocus : A Spatial Domain Approach, " International Journal of Computer Vision, vol.13, no.3, pp.271-294, 1994.
- [52] 石原満宏, 佐々木博美, "合焦法による高速三次元形状計測", 精密工学会誌, vol.63, no.1, pp.124-128, Jan. 1997.
- [53] 竹村岳, 松山隆司, "多重フォーカス画像を用いた実時間3次元距離計測", 信学技報, 97-CVIM-103, no.CVIM103-5, pp.35-42, Jan. 1997.
- [54] M. Watanabe, S. K. Nayer, "Telecentric optics for focus analysis," IEEE Trans. Pattern Analysis and Machine Intelligence, vol.19, no.12, pp.1360-1365, Dec. 1997.
- [55] 三宅和夫, "幾何光学", 共立出版, 東京, 1979.
- [56] 児玉和也, 大西隆之, 相澤清晴, 羽鳥光俊, "反復法に基づく複数画像からの任意焦点画像の生成", 映像情報メディア学会誌, vol.51, no.12, pp.2072-2081, Dec. 1997.
- [57] 沼尾利夫, 奥富正敏, "ズームレンズのキャリブレーションにおけるターゲットと精度に関する検討", 映像情報メディア学会誌, vol.51, no.12, pp.2126-2132, Dec. 1997.
- [58] 工藤朋之, 三池秀敏, "カメラフォーカスを連続変化させた動画画像を用いた奥行き分布検出", 情処 CV98-2, pp7-12, Jan 1996

- [59] 井口, 佐藤, " 三次元画像計測", 昭光堂, 1990.
- [60] 吾妻健夫, 魚森謙也, 森村淳, " ステレオ画像の中間視点画像生成のためのエッジ情報を用いた視差推定", 映像情報メディア学会誌, vol.52, no.3, pp322-330, 1998.
- [61] ISO,"Coding of moving pictures and associated audio for digital storage media at up to about 1.5Mbits/s Parts Video", ISO/IEC 11172-2.
- [62] ISO,"Coding of moving pictures and associated audio Part2 Video", ISO/IEC 13818-2 (Committe Draft).
- [63] A.Pentland,"A New Sense for Depth of Field",IEEE Transaction on Pattern Analysis and Machine Inteligence, vol.PAMI-9, no.4, July 1987
- [64] A.Pentland,T.Darrell,"A Simple, Real-time Range Camera", Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, California, June 1989..
- [65] Masahiro Watanabe, Shree K.Nayer,"Minimal Operator Set for Passive Depth from Defocus", Proc.of IEEE Conf.Computer Vision and Pattern Recognition.pp.431-438,1996.
- [66] 大原知之, 三輪道雄, 石井雅博, 小池康晴, 佐藤誠, " 遠距離測定のための多重フォーカス画像観測系の提案", 信学技報, MVE98-80, pp.17-24, Dec. 1998.
- [67] 三輪道雄, 大原知之, 石井雅博, 小池康晴, 佐藤誠, " 焦点ボケを用いた遠距離物体領域の抽出法", 第5回画像センシングシンポジウム論文集, pp.221-226, Jun. 1999.
- [68] 三輪道雄, 大原知之, 石井雅博, 小池康晴, 佐藤誠, " 焦点ボケを用いた遠距離物体の推定", 3次元画像コンファレンス'99, pp.305-307, June 1999.
- [69] Murali Subbarao, Gopal Surya,"Depth from Defocus: A Spatial Domain Approach", , International Journal of Computer Vision, vo.13, no.3, pp.271-294, 1994.
- [70] K.R.Rao,P.Yip 著, 安田, 藤原訳, " 画像符号化技術", オーム社, 1992.

本論文に関する研究業績

学術論文

1. 三輪道雄, 大原知之, スカンヤポーンスパツ, 石井雅博, 小池康晴, 佐藤誠, ”色フィルタとテクスチャ解析を用いた緑色道路標識の抽出”, 映像情報メディア学会誌, vol.53, no.5, pp.722-729, May 1999.
2. 大原知之, 三輪道雄, 石井雅博, 小池康晴, 佐藤誠, ”合焦法遠距離計測のためのスクリーン投影型画像観測系の提案”, 電子情報通信学会論文誌 (D-II), Vol.J82-D-II, no.10, pp.1750-1757, Oct. 1999.
3. 三輪道雄, フンフ・クァン・フイ・ベト, 石井雅博, 小池康晴, 佐藤誠, ”DCTを用いた合焦判定による物体検出法の提案”, 映像情報メディア学会誌, vol.54, No.2, pp252-259, Feb. 2000.

国際会議

1. M.Miwa,T.Oohara,P.Sukanya,M.Ishii,Y.Koike,M.Sato,”Extracting a Traffic Sign From the Video Stream Using Color Filtering, Histogram Analysis, And Texture Analysis”,11th Scandinavian Conference on Image Analysis(SCIA99), Knagerlussuaq, Greenland, pp.243-250,June 1999.
2. Hunph Quang Huy Viet, Michio Miwa,Yasuharu Koike, Makoto Sato,”Video stream analysis for a traffic sign extraction using color filtering and histogram analyssis, texture analysis”, Asian Conferences on Computer Vision(ACCV2000),pp.668-674,Taipei,Ta Jan. 2000.
3. Michio Miwa,Yasuharu Koike, Makoto Sato,”A study of screen intermediate camera for ranging far-away objects”, International Workshop on Advanced Image Technology(IV 225,Kanagawa, Japan, Jan. 2000.

雑誌掲載

1. 三輪道雄, 佐藤誠, ”異なった手法の組み合わせによる緑色道路標識の抽出”, 画像ラボ, Vol.10, No.11, pp.1-5, Nov. 1999.

口頭発表（全国大会，研究会，等）

1. 三輪道雄，大原知之，スカンヤポーンスーパツ，石井雅博，小池康晴，佐藤誠，” 動画像からの道路標識の抽出法”，映像情報メディア学会冬期大会，pp.102-102，Dec. 1998.
2. 大原知之，三輪道雄，石井雅博，小池康晴，佐藤誠，” 遠距離計測のための多重フォーカス画像観測系の提案”，映像情報メディア学会冬期大会，pp.80-80，Dec. 1998.
3. 大原知之，三輪道雄，石井雅博，小池康晴，佐藤誠，” 遠距離測定のための多重フォーカス画像観測系の提案”，信学技法 MVE98-80，pp.17-24，Dec. 1998.
4. 三輪道雄，大原知之，スカンヤポーンスーパツ，石井雅博，小池康晴，佐藤誠，” 色フィルタとテクスチャ解析を用いた動画像からの道路標識抽出”，情処研資，CVIM114-13，pp.97-103，Jan. 1999.
5. 三輪道雄，大原知之，石井雅博，小池康晴，佐藤誠，” 焦点ボケを用いた遠距離物体領域の抽出法”，第5回画像センシングシンポジウム講演論文集，pp.221-226，June 1999.
6. 三輪道雄，大原知之，石井雅博，小池康晴，佐藤誠，” 焦点ボケを用いた遠距離物体の推定”，3次元画像コンファレンス，pp.305-307，June 1999.
7. 三輪道雄，石井雅博，小池康晴，佐藤誠，” 焦点ボケ画像によるDC Tを用いた物体位置推定”，映像情報メディア学会年次大会，pp.302-303，Aug.1999.

出願特許

1. 三輪道雄，佐藤誠，” 仮想空間生成装置”，特願平 8-342882.
2. 佐藤誠，三輪道雄，高松亮，” 位置検出装置”，特願平 9-132984.
3. 三輪道雄，佐藤誠，石井雅博，大原知之，” 遠距離画像撮影装置”，特願平 10-233940.
4. 三輪道雄，佐藤誠，大原知之，” 標識認識装置”，特願平 10-307673.
5. 三輪道雄，佐藤誠，” 物体検出装置”，特願 2000-021627