

博士論文

3次元統合画像符号化の基礎検討

平成6年12月20日

指導教官 原島 博

東京大学工学系研究科 電子工学専攻

学籍番号 27121

藤井俊彰

この論文は、Copyを原島研秘書さんから
もらったものである。(原本は、図書にとどけておく)

目次

1 序論	6
1.1 本研究の背景と目的	7
1.2 本論文の構成	7
2 3次元統合画像通信の構想	9
2.1 はじめに	10
2.1.1 3次元画像入力	10
2.1.2 3次元画像の表示	11
2.1.3 3次元画像の符号化・圧縮方式	15
2.2 「3次元統合画像通信」の提案	17
2.3 空間符号化へ	20
参考文献	22
3 3次元画像の取得と表示方式	24
3.1 はじめに	25
3.2 視野と視域	25
3.3 3次元画像入力方式の分類	25
3.4 カメラ入力法の解析	26
3.4.1 正投影と中心投影	26
3.4.2 カメラ配置による分類	29
3.4.3 さまざまなカメラ配置の例	33
3.4.4 フレネルレンズを用いて撮影した場合	35
3.5 幾何学的補正	37
3.5.1 キーストーン歪み補正	37
3.6 視差画像の位置合わせ	38
3.6.1 定式化	38
3.6.2 例1：一般的な線形関数	39
3.6.3 例2：非線形関数	39
3.7 画像の幾何学的変換	40
3.7.1 アフィン変換	41
3.7.2 回転, 平行移動, スケーリング	41
3.8 コンピュータによる入力	43
3.9 ホログラム入力	44
3.10 表示方式	45

3.10.1	2眼ステレオ方式	45
3.10.2	レンティキュラー方式	45
3.10.3	ホログラフィ方式	45
参考文献		46
4	光線群による3次元画像の統一的表現	47
4.1	はじめに	48
4.2	光線の物理的性質	48
4.3	光線群による被写体空間の記述	49
4.4	空間中の一点を通る光線群	51
4.5	撮像による光線群のサンプリング	53
4.6	光線空間のもつ冗長性	54
4.7	3次元画像符号化の定式化	55
4.8	誤差評価法	56
4.9	視差画像の補間合成の問題	57
4.10	$f(x, u)$ の帯域	58
4.11	その他の応用例	60
4.12	まとめ	61
参考文献		62
5	視差補償予測に基づく3次元画像符号化	64
5.1	はじめに	65
5.2	従来の研究	65
5.2.1	2次元画像の圧縮・符号化	65
5.2.2	ステレオ画像の相関	65
5.2.3	3次元画像符号化の研究動向	65
5.3	動画像標準符号化方式の適用	66
5.3.1	標準符号化方式 H.261	66
5.3.2	H.261 を用いた多眼画像の符号化	71
5.3.3	符号化特性の評価	71
5.4	幾何学的制約を用いた視差補償(1)	76
5.4.1	H.261 における視差ベクトルの内挿	76
5.4.2	視差内挿予測符号化シミュレーション	79
5.5	幾何学的制約を用いた視差補償(2)	83
5.5.1	mn 画像の導入と多眼画像間の幾何学的関係	83
5.5.2	mn 画像からの視差ベクトルの抽出	86
5.5.3	アフィン変換に基づく視差ベクトルの補間	89
5.5.4	非撮影方向からの画像の合成, 補間	93
5.5.5	背景と物体の分離に基づく予測特性の改善	96
5.6	まとめ	109
参考文献		110

6	3次元構造推定に基づく3次元画像符号化	112
6.1	はじめに	113
6.2	構造抽出に関する従来の研究	113
6.3	被写体空間と視差画像の幾何学的な関係	114
6.3.1	投影モデル	114
6.3.2	視差画像空間と規格化空間	114
6.4	3次元構造モデルを用いた視差画像空間の符号化	116
6.4.1	問題の定式化	116
6.4.2	3次元構造モデル	117
6.4.3	評価値の設定	118
6.4.4	モデルパラメータの最適化	119
6.5	3次元構造抽出符号化の実際	120
6.5.1	モデルの制約条件	120
6.5.2	評価値空間の導入	120
6.5.3	動的計画法の利用	121
6.5.4	アルゴリズム	121
6.6	実験1	122
6.6.1	画像の取得	122
6.6.2	アルゴリズム	122
6.6.3	処理例1	123
6.6.4	処理例2	123
6.7	実験2	127
6.7.1	モデル記述に要する付加情報量と予測特性の検討	127
6.7.2	評価値の設定を変化させた場合	128
6.8	評価値空間のセグメンテーションによる予測特性の向上	132
6.9	テクスチャと予測誤差の符号化	132
6.10	まとめ	135
	参考文献	138
7	ホログラムパターンの圧縮, 符号化	141
7.1	はじめに	142
7.2	光の回折の基礎理論	142
7.2.1	回折の近似式	142
7.2.2	回折の計算法	143
7.2.3	ホログラフィの原理	145
7.3	ホログラフィテレビジョン技術の歴史と現状	146
7.3.1	テレビジョン伝送の試み	147
7.3.2	情報量の計算	147
7.3.3	ホログラフィ研究の停滞期	148
7.3.4	ホログラフィテレビジョン研究の最近の動向	148
7.3.5	今後の展望	152
7.4	ホログラムパターン圧縮の研究動向	154
7.5	ホログラムパターンと光線空間との相互変換	154

7.5.1 離散フレネル変換の導入	154
7.5.2 ホログラムから光線空間への変換実験	155
7.6 まとめ	160
参考文献	161
8 結論	170
8.1 本論文の主たる成果	170
8.2 今後の課題と展望	171
発表文献	173

第 1 章

序論

立体テレビジョンは長い間人類の夢のテーマであった。近年、レンズの微細加工技術の発展、レンチキュラシートとの位置合わせの容易なディスプレイの開発、計算機の高性能化、デジタル伝送技術、CG 技術の発展など周辺技術が整備されてきたこと、ホログラフィの動画化の試み、また VR(Virtual Reality) 技術や臨場感通信への期待などがあいまって、再び 3 次元画像のブームがおとずれているように思われる。このような背景の中、3 次元画像の蓄積、伝送を効率よく行なうため、必要とされる要素技術、特に圧縮符号化技術を確立することが重要なテーマとなってきている。

1.1 本研究の背景と目的

3次元画像技術の歴史は古く、例えば1600年頃に2枚の手書き絵による両眼視の試みがあることが報告されている。その後、19世紀には眼鏡を用いたアナグリフ方式、偏光方式が考案され、20世紀初頭には眼鏡なしのレンチキュラ方式が提案されている。さらに、20世紀後半にはホログラフィによる3次元空間像形成技術が開発されている。

考えてみれば我々が生活している空間は本来3次元であり、それと同じ仮想環境を人工的に実現することは、人類の夢であった。この立場から多くの3次元画像技術が開発され、最近では、バーチャルリアリティ（仮想現実感）、臨場感通信が注目を浴びてきている。

しかし、これまでの3次元画像に関する研究開発状況を振り返ってみると、その大部分が3次元ディスプレイを中心とするハードウェア技術であり、体系的な検討はほとんど行われていないことに気づく。関連して奥行き知覚を中心とする人間の視覚特性も調べられているが、それも心理学的あるいは生理学的な立場からの検討が中心である。

このような3次元画像技術は、今後どのようにして体系的に統合されていくのだろうか。3次元画像技術は、将来の立体テレビジョン放送や3次元画像通信の実現にどのように結びついていくのであろうか。近年のマルチメディア化、大規模ネットワーク化にともなって、将来の立体テレビジョン放送や3次元画像通信が現実のものとなろうとしている今、3次元画像の効率的な符号化手法の確立は、重要な課題として残されている。

さらには、ネットワーク化されたバーチャルリアリティ環境下においては、ネットワークを通じて環境そのものがやりとりされるという状況が想定される。このような状況における符号化とは、もはや「画像」の伝送ではなく、「空間そのもの」をいかに伝送するという考えを考へるべきであろう。画像を符号化するのではなく、空間そのものをある中間的な記述に変換し符号化すること、すなわち、画像符号化 (picture coding) ではなく、「空間符号化 (space coding)」がこれからの符号化の中心的な課題となっていくと思われる。

このような背景から、本論文では、従来のデバイス中心の符号化法ではなく、さまざまな方式の3次元ディスプレイが共存することを前提とした「3次元統合画像通信」(Integrated 3-D Visual Communication) を提案する。これは3次元画像の中間的な記述フォーマットを定め、入出力デバイスを独立に規定しようという試みである。

従来、研究されてきている3次元画像に関する研究をこの観点からとらえ直し、統一的に論ずることが、本論文の一つの目的である。

1.2 本論文の構成

本論文の具体的な構成は以下の通りである。

第1章. 序論

第2章. 3次元統合画像通信の構想

第3章. 3次元画像の取得と表示方式

第4章. 光線群による3次元画像の統一的表現

第5章. 視差補償予測に基づく3次元画像符号化

第6章. 3次元構造推定に基づく3次元画像符号化

第7章. ホログラムパターンの圧縮、符号化

第8章. 結論

これらの章の間を模式的に表すと、図1.1のようになる。

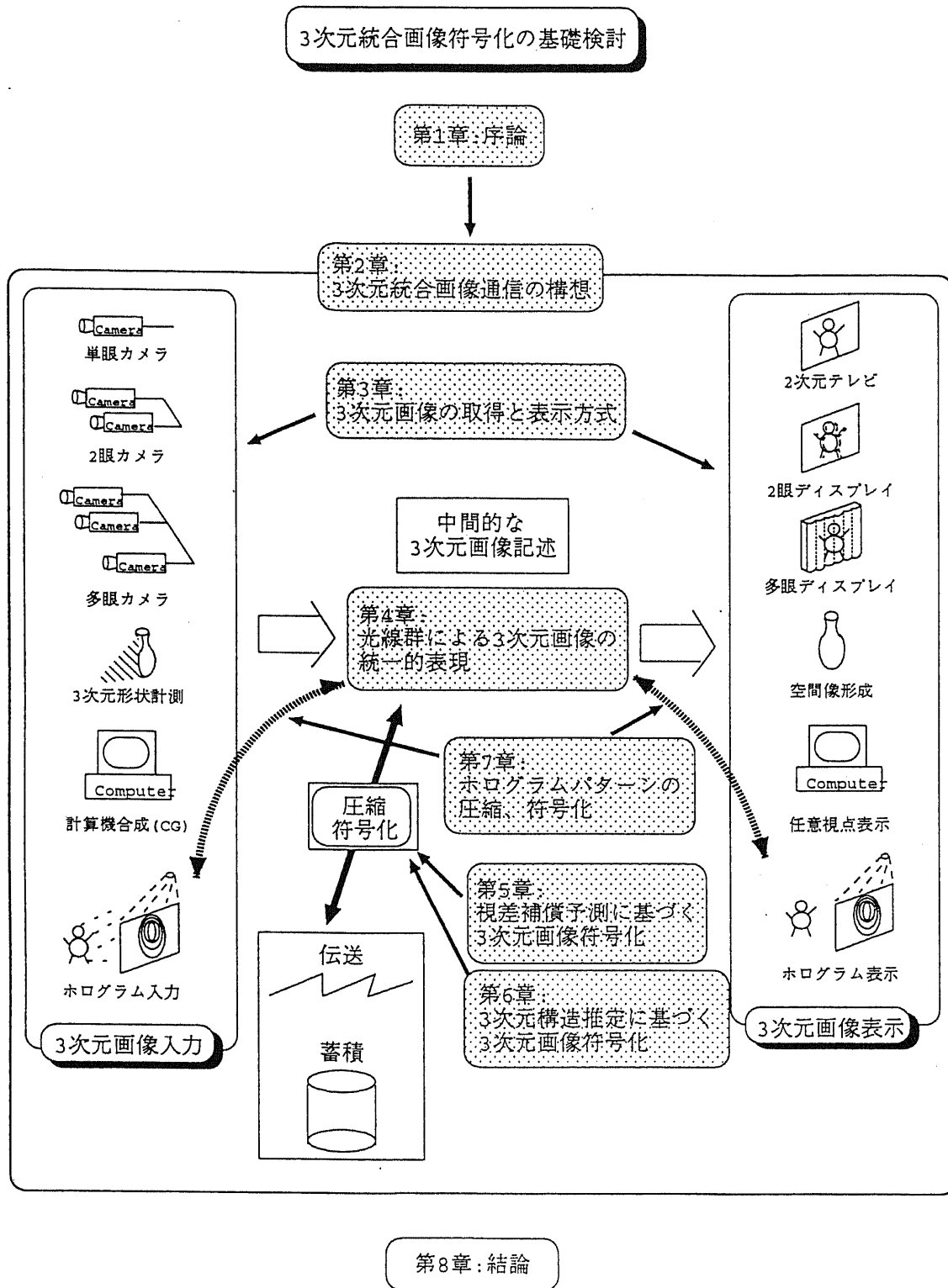


図 1.1: 本論文の構成

第 2 章

3 次元統合画像通信の構想

一口に 3 次元画像といってもさまざまな種類があり, 用途に応じてそれぞれの分野で使い分けられている. 本章ではまず, 現在知られている 3 次元画像入出力方式について概観し, 将来における「3 次元統合画像通信 (Integrated 3-D Visual Communication)」の構想について述べる. また, 3 次元統合画像通信の基盤技術となる 3 次元画像符号化に関して, その技術課題をまとめるとともに, 今後の展望について述べる.

2.1 はじめに

本章では、将来における「3次元統合画像通信 (Integrated 3-D Visual Communication)」の実現に向けて、その基本構想を明らかにする。また、3次元統合画像通信の基盤技術となる3次元画像符号化に関して述べる。

以下に紹介するように、3次元映像表示技術にはさまざまなものがあり、それぞれ一長一短がある。現在の技術でフルカラーの3次元画像の実時間伝送、表示が可能なものとしては、2眼式が最も有望である。レンティキュラーやバララクスバリアを用いた方式も、研究段階から商品化へと進んできている。ホログラフィは完全な3次元画像として期待されているが、リアルタイム表示にはまだ克服すべき課題が多い。そこで、より自然な3次元像をめざした研究をおこなっていくことはもちろんのこと、現在までに開発されている3次元表示装置がそれぞれの特徴を生かせる場所で用いられることを前提に、体系的な検討を行うことも重要であると考えられる。ここでは、3次元画像通信の形態を図2.1のような形で考えてみたい。入力、符号化・伝送（蓄積）、表示とに分け、以下、入出力方式について概観した後、符号化方式を見ていくことにする。入出力方式についての詳細な解析は3章で述べる。

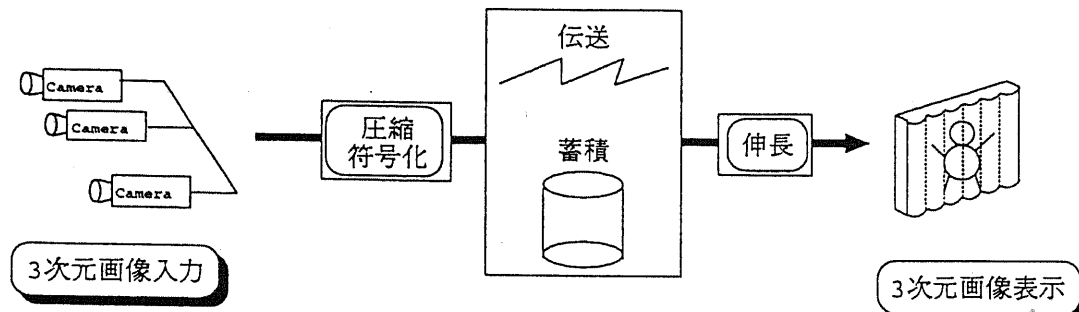


図 2.1: 3次元画像通信の形態

2.1.1 3次元画像入力

2次元画像の場合、入力装置は単眼のテレビカメラである。放送品質が要求される場合には撮像管が多く用いられており、また家庭用としてはCCDカメラを用いたものが広く用いられている。また最近ではCG技術が発達し、コンピュータで合成した画像を実時間で表示することも可能となってきており、この場合には2次元画像の入力はコンピュータであると考えられる。

さて、3次元の場合、入力装置としてはどのようなものがあるのだろうか。

現在のところ、入力装置は出力装置に対応して決まってしまうというのが現状で、3次元画像の出力装置に特化したデータ取得装置という位置付けとなっている。例えば、

- 2眼式表示装置の場合、2眼カメラが入力
- 多眼式表示装置の場合、多眼カメラが入力
- ホログラフィの場合、ホログラム撮影が入力

という入力形態が一般的である。

また、3次元画像においても2次元画像におけるCGと同様にコンピュータを入力とすることが可能である。CG手法を用いて疑似的に2眼カメラや多眼カメラを入力としたものと等価な画像を合成して表示する方法がとられる。またホログラフィにおいても、コンピュータ内部に仮想的な2次元または3次元物体を仮定し、光の伝搬を計算することによってコンピュータ内部で干渉縞を計算するという、計算機合成ホログラム (Computer Generated Hologram, CGH) の手法がある。

コンピュータ内に物体の3次元構造情報を構築するためには、CAD/CAMで用いられている手法、すなわち、

- 形状のプリミティブ (球, 平面, 円柱など) を組み合わせてコンピュータ内部で仮想3次元物体を作り上げていく

といった手法や、実物体を3次元計測する、すなわち、

- 可視光, レーザ光, 超音波などを使った能動的3次元計測

を行ない、物体の3次元構造をコンピュータ内に取り込むといった手法が行なわれる。これらも広い意味での3次元画像の入力方式の一つととらえることができよう。

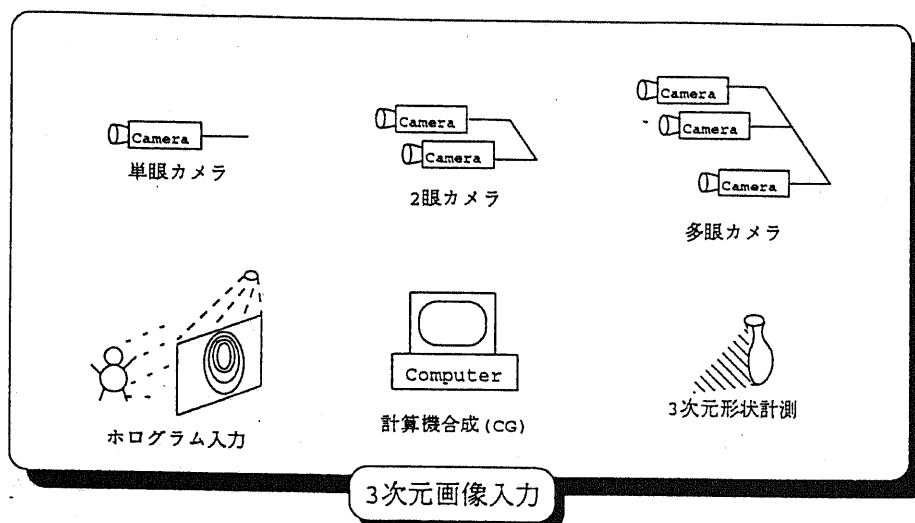


図 2.2: さまざまな3次元画像入力方式

2.1.2 3次元画像の表示

現在、3次元画像表示技術は、図 2.4に示すように多くの方式が提案されている。このうち、代表的なものについて説明を加える。3次元画像表示方式の詳細については、例えば文献 [4, 1, 2, 4] を参照されたい。また、() 内には奥行き知覚の要因を付しておく。

アナグリフ方式：

白地に赤と青で描かれた左右画像を、赤-青眼鏡を通して観察することにより左右画像を分離する方式である。簡単な方法 (印刷物など) で立体映像が得られるという利点があ

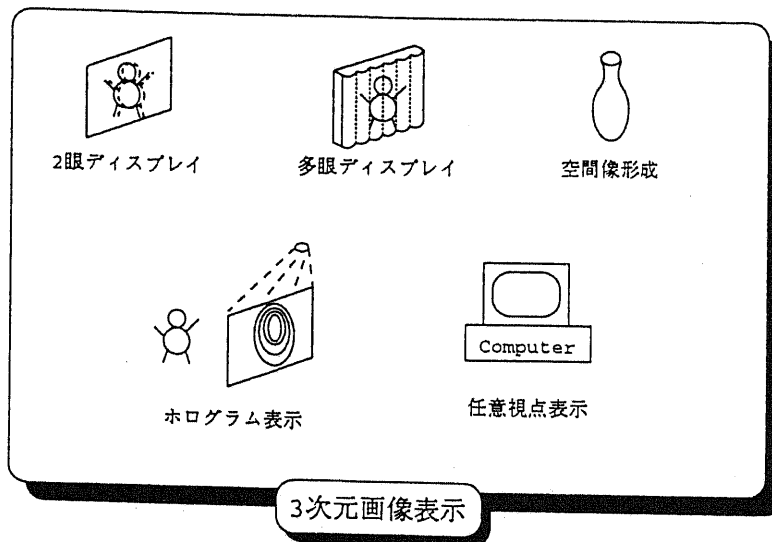


図 2.3: さまざまな3次元画像出力方式

る。しかし、本質的にフルカラー化が不可能であること、明るさやコントラストが十分でないなどの問題がある。(両眼視差, 輻輳)

偏光方式 :

偏光眼鏡方式は直線偏光の振動方向の違い, あるいは円偏光の回転方向の違いをそれぞれ利用して左右眼像を分離する方法である。この方式は、高解像度のカラー動画像表示が可能で、同時に多数の人に対して立体画像を見せるのに適している。(両眼視差, 輻輳)

時分割方式 :

両眼に対応した画像をテレビのフィールド毎に交互に切り換えて提示し、提示する画像に同期させて開閉するシャッター眼鏡を用いて立体視する方式である。例として1985年にビクターが技術発表したレーザーディスクがある [12]。フレーム周波数が左右画像各々30Hzであるため画像が暗い、フリッカが生じるという問題があった。また、ソフトが続かず廃れてしまった感がある。その後、フリッカレス時分割立体テレビ方式が提案され、この原理に基づいた立体テレビ装置が開発されている [13]。この方式は現行テレビと両立性が良い。(両眼視差, 輻輳)

濃度差式 :

ブルフリットヒ効果と呼ばれる錯視現象を利用した方式である。片方の眼に減光フィルターを付けて左右の明るさを変えることによって網膜から脳への情報伝達速度が異なることを利用する。垂直面内で運動している振り子を減光フィルター付きの眼鏡をかけて観察すると、奥行き方向に楕円起動を描いて運動しているように知覚する。(両眼視差, 輻輳)

レンチキュラ方式 :

多眼式レンズ板3次元像は、レンチキュラシート (Lenticular Sheet, LS) と呼ばれる、かまぼこ状の形をした微小レンズの配列からなる板状光学素子を用いて、適当な視差を持った像を見せ、立体感を出す方式である。これにより眼鏡なしで立体像を観察することが可能となる。この多眼式レンズ板3次元像は、視点が断続的に変化するような

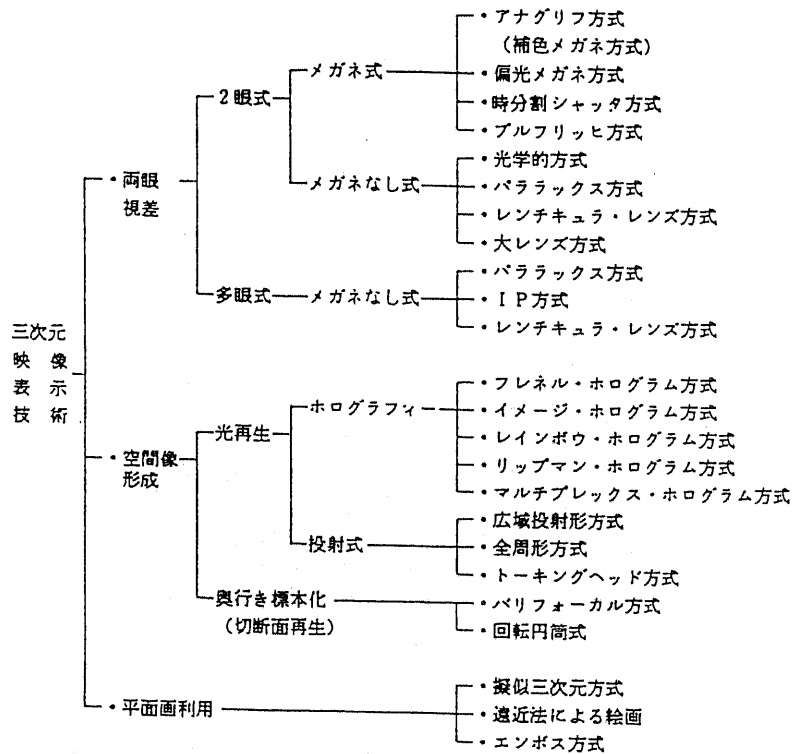


図 2.4: 3次元画像表示技術の分類

像ではなく、連続的に変化するような自然な3次元像を知覚することができ、連続視域性 (autostereoscopicity) が保たれている。実時間表示が可能な程度に表示装置で取り扱うべき情報量が少ない。そのため、像を得る時間や経費の面で有利である、伝送、記録が容易である、といった特長を持つ。現在では立体絵はがき、3次元テレビジョン、立体カメラなどへ広く応用されている。ディスプレイとしては、レンチキュラスクリーンに投影する方式と、レンチキュラレンズをディスプレイ (CRT, プラズマディスプレイ, 液晶テレビなど) にとりつけた直接表示方式とがある。(両眼視差, 輻輳, 運動視差)

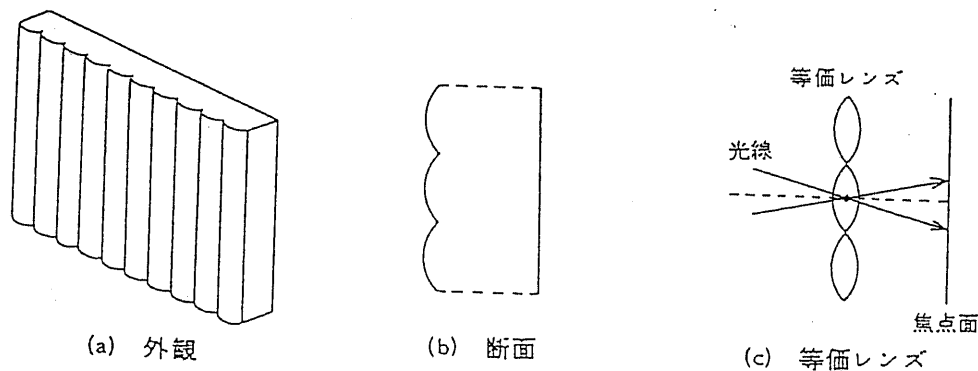


図 2.5: レンティキュラ板の概観図とその原理

空間像形成:

バリフォーカルミラーは、スピーカの前面に振動可能な薄膜ミラーを取りつけ、薄膜ミ

ラーが凸面ミラーや凹面ミラーに変わるのに同期させて、ブラウン管上に表示された物体の各切断面像をミラーに反射させる方式である。この他、ファイバーシートや、LEDアレイを回転させて3次元像を表示するもの、可動スクリーンに断層像を投射するものがある。これらの空間像形成方式は内部が透けて見えるファントムイメージング現象が特徴になっているので、相互の位置関係を正しく判断しなければならない航空管制や医療分野での活用、コンピュータ・グラフィックス出力画像の表示用に適している。図2.6にLEDアレイを用いたディスプレイの例を挙げる [36]。(両眼視差, 輻輳, 運動視差)

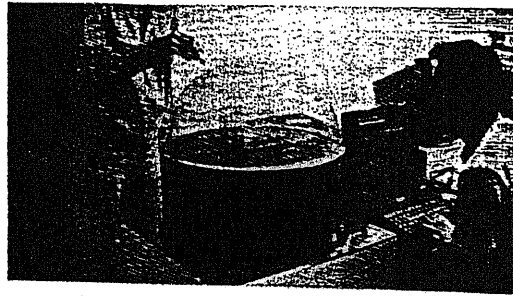


図 2.6: 空間像形成方式

ホログラフィ:

ホログラフィは物体の無数の点からの光の振幅と位相の両方の情報を含んだ干渉パターン(ホログラム)を記録・再生する技術である。ホログラフィによる3次元画像表示は、任意の視点から3次元映像を見ることができ、立体視の要因となる両眼視差, 運動視差, 焦点調節, 輻輳など, すべてを備えているため, 最も理想的な3次元表示技術といわれている。ホログラフィの原理, 種類などは後の章に譲る。(両眼視差, 輻輳, 運動視差, 水晶体の調節)

これらさまざまな3次元画像表示技術の中で, 比較的将来性が高いと考えられているのは次の3方式であろう。

(1) 両眼ステレオ方式

両眼視差を利用するステレオ方式は, 眼鏡を必要とし, アナグリフ方式(赤青方式), 時分割シャッター方式, 偏光方式など多くの種類がある。簡単で情報量が少なく, 大画面, フルカラー動画化も可能である。しかし, 視線方向はももとの撮影方向に限定され, 疲れやすく, 長時間使用には適さない。

(2) 多眼レンチキュラ方式

蒲鉾状のレンズを用いる多眼レンチキュラ方式は眼鏡なしで立体視できることが特徴である。液晶ディスプレイと組み合わせて, フルカラー動画化も可能である。しかし, 視域は限定され, 多人数での鑑賞には適さない。

(3) ホログラフィ方式

眼鏡なしで真の3D画像が可能なホログラフィは, その記録方式によってフレネルホログラム, リップマンホログラム, レインボーホログラムなどに分類できる。印刷技術と結びついてペンダントやカードの装飾として用いられているが, 動画化への試み(ホロビデオ)は始まったばかりである [11]。

両眼ステレオ方式は, 眼鏡を必要とするがカラー化, 動画化, 大画面化が可能であり, イベント等での多人数鑑賞用としてはほとんど唯一の実用方式である。

眼鏡なしでは、多眼レンティキュラ方式が完成度が高い。カラー化、動画化も可能である。しかし、視域が限定されるので、個人使用（例えば、テレビ電話やコンピュータ端末）に限られよう。

ホログラフィ方式は、眼鏡なしで任意視点の立体像が得られると言う意味で本命視されることが多い。しかし、技術的には未熟で、当面はデザインやアートなどに用途が限られるであろう。

少なくとも、現状を見る限り、3次元画像表示方式には、表2.1に示すように一長一短があり、1つの方式のみが主流になるようには思われない。むしろ応用あるいは目的に応じて、技術の棲み分けができつつあるように見える。

Table 2.1: 3次元画像表示方式の比較

	両眼ステレオ方式	多眼レンティキュラ方式	ホログラフィ方式
メガネなしで立体視が可能か	× 眼鏡が必要	○ 眼鏡なし	○ 眼鏡なし
任意視点から観察可能か	× 撮影方向に限定される	△ 視域が制限される	○ 視域が広い
フルカラー化が可能か	○ 偏光方式ならば可能	○ 可能	△ 可能になりつつある
動画化が可能か	○ 容易に可能	○ 平面ディスプレイとの組合せ	× 研究が始まったばかり(△)
大画面化が可能か	○ 投射型とすれば容易に可能	△ 大型パネルの製造技術次第	× 画面の大きさには限界がある
多人数で同時に観賞可能か	○ 劇場でも可能	× 視域が制限され個人用	△ 視域制限はないが、少人数
特徴のまとめ	眼鏡を必要とするが、大画面かつ多人数観賞可能	眼鏡なしでよいが視域が限定されるので個人用	真の3次元空間像が合成できるが、動画化など技術的課題が多い
とりあえずの用途は？	映画・イベントなどでの多人数鑑賞向け	コンピュータ端末やテレビ電話端末等の個人使用向け	デザインや芸術分野向け

2.1.3 3次元画像の符号化・圧縮方式

2次元画像における符号化・圧縮方式については、すでに数多くの研究がなされており、現在標準化作業が進められている。静止画像符号化においてはJPEG、動画像符号化においてはDCT (Discrete Cosine Transform, 離散コサイン変換) とMC (Motion Compensation, 動き補償) を基幹とした符号化方式 (通信用符号化方式としては H.261, MPEG2, 蓄積用符号化

方式として MPEG1 など) が提案されている。JPEG, H.261, MPEG1, MPEG2 についてはすでに勧告が出されている。これらの画像標準符号化方式は、MC により時間軸方向の冗長性を削減し、DCT により空間軸方向の冗長性を削減している。

それでは、3次元画像の符号化方式としてどのような研究がなされているのだろうか。

従来行なわれてきた研究を眺めると、2眼画像あるいは多眼画像についての符号化方式の提案と、ホログラムパターンの圧縮の提案がなされている。以下、それぞれについて簡単に述べる。

(1) 2眼画像または多眼画像の圧縮、符号化

2眼画像または多眼画像においては、各画像間に強い相関がある。これは、ある対象を少しずつ視点を変えて撮影した画像であることからくる冗長性である。これを、適当な方法で取り除けば大幅な情報圧縮が期待できる。一方、各画像は普通の2次元画像であるから、空間的な冗長性があると考えられる。したがって、視差軸方向にはフレーム間予測の手法を用い、空間軸方向には変換符号化を用いるという動画像符号化方式と同様な手法を適用することができる。

この立場から、それぞれの画像間の視差ベクトルを検出し、その視差を補償して予測符号化をおこなう「視差補償予測符号化」が提案されている。これと時間方向の動き補償と組み合わせる試みもある(図2.7) [19][20]。現在のステレオ画像符号化方式の提案は、何らかの形で視差補償を用いているものが多い。視差補償予測符号化の詳細や従来の研究については、章を改めて5章で述べることにする。

この考え方は、レンティキュラ方式などの多眼画像へも自然に拡張される。図2.8はその一例である [21]。この視差補償符号化では、中央の2次元画像を基準として、これに視差画像を付け加えることによって3次元化を図っている。その意味では、この符号化は2次元画像システムと両立性のある方式であると言える。

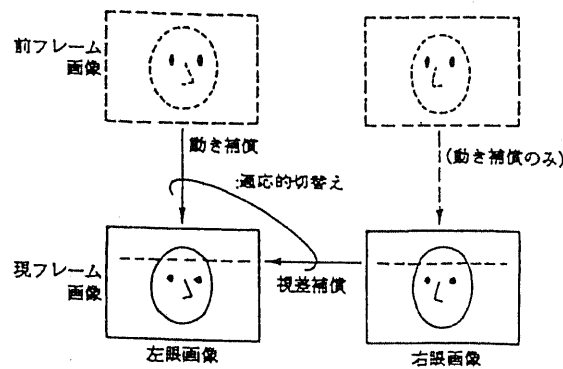


図 2.7: ステレオ動画像符号化の例 [40]

(2) ホログラムパターンの圧縮、符号化

ホログラフィの符号化は、直接的には干渉縞の符号化になる。これまで光学系による情報圧縮の手法の他には、2次元静止画像標準符号化方式の JPEG を用いた方式が提案されている。ホログラムパターンは多くの高周波成分を含むことからホログラムパターンに直接 JPEG を適用してもあまり圧縮率は期待できない。そこで、JPEG を施す前にホログラムパターンをブロック分割した後に DCT を施し、その結果に対して JPEG を施すことで品質の劣化の少な

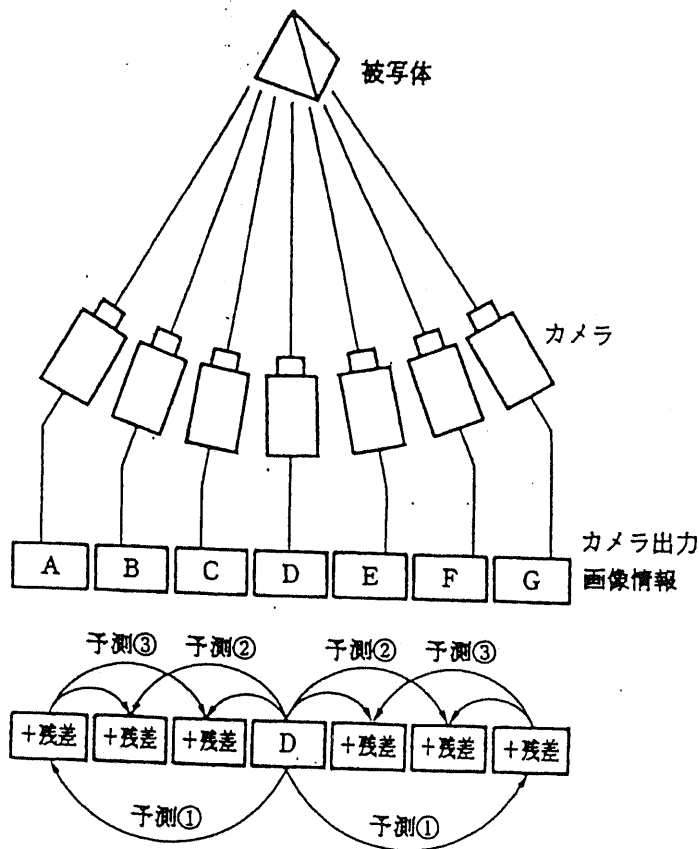


図 2.8: 多眼式 3 次元画像の視差補償符号化 [21]

い圧縮が可能であるとしている [37]. ホログラムパターンの圧縮, 符号化に関しては章を改めて7章で述べることにする.

2.2 「3次元統合画像通信」の提案

現在利用できる3次元画像表示方式はそれぞれ一長一短があり, 少なくとも近未来において一つの方式が主流になるとは思われない. これらは用途に応じて使い分けられていくであろう. したがって, 3次元画像通信システムの検討においては複数の方式の共存を前提すべきであろう. また, 将来3次元画像の技術が進歩したとしても, 画像通信の基本は2次元動画であると考えられ, 現在標準化作業が進められている2次元動画画像通信方式との両立性(コンパチビリティ)も考慮すべきである. このような技術の現状を眺めてみると, 少なくとも近未来においては, 1つの方式に統一されることはありえず, 3通りあるいはそれ以上の方式が目的に応じて使い分けられると考える方が自然である. システム的な検討も, 複数の方式の共存を前提としておこなうべきであろう.

それでは, 共存を前提としたときの画像通信システムのイメージはどのようなものになるのでしょうか.

符号化の研究もそれぞれの表示方式に特化して断片的になされてきたという印象を拭えない. このような背景から, 先に我々は将来の3次元画像通信の形態として「3次元統合画像通信

(Integrated 3-D Visual Communication)」を提案した。これは、すべての3次元画像表示方式に共通の中間的な画像記述を規定することにより、入出力を独立に規定できることを目指すものである。

その一つは、通常の2次元画像システムとのコンパチビリティ（両立性）である。将来3次元画像技術が進歩したとしても、画像通信の基本は2次元動画であろう。3次元画像システムは単独で存在するのではなく、必要に応じて2次元に一次元追加する形で実現されることが望ましい。

さらに、3次元画像も複数の方式の共存を前提とすれば、それぞれ全く別々のシステムとして存在するのではなく、当然それぞれの間のデータの交換、共有、さらにはネットワーク化が必要になってこよう。

この観点から、ここでは将来のより柔軟な3次元情報環境へ向けて、「3次元統合画像通信」(Integrated 3-D Visual Communication)を提案したい。

1988年にサービスが開始されたサービス総合デジタル通信網（ISDN）では、音声・データ・映像情報をすべて同じ情報記述すなわちデジタル化することで、伝送、交換、サービスの統合を可能にした。

これと同じように「3次元統合画像通信」では、入力方式、出力（表示）方式に依存しない中立的な情報記述を採用することによって、統合的な3次元情報環境を実現することを目指している（図2.9）。その特徴は次の通りである。

(1) 入力方式と出力方式を独立に規定できる

従来の3次元技術では、画像の入力方式と出力方式が、原則として対応していた。すなわち、

ホログラフィ入力	—	ホログラフィ表示
多眼カメラ入力	—	レンティキュラ表示
2眼カメラ入力	—	両眼ステレオ表示

これに対して、3次元統合画像通信において方式間の自由なデータの交換が可能になれば、3次元画像の入力方式と出力方式（表示方式）を独立に規定できる。

従って、入力方式は、画像表示方式とは無関係に、状況に応じて最も入力が容易な方式（例えば多眼カメラ入力）を採用することができる [15]。

一方、画像出力すなわち表示方式は、その利用法に応じて最適なものを選べばよい。すなわち、

- シアター等における多人数による鑑賞では両眼ステレオ方式
- コンピュータ端末、TV電話端末等における個人使用では多眼レンティキュラ方式
- 真の立体像が必要となるデザイン・芸術分野ではホログラフィ方式

その他、画面は通常の2次元であるが、任意方向からの像をCG技術によって自由に合成表示できる「任意視点合成表示方式」も、疑似的な3次元表示方式とみなすことができる。考えてみれば、3次元画像において立体視は必要不可欠ではない。例えばデザインや宣伝（カタログプレゼンテーションなど）分野では、任意方向からの画像が合成できれば、立体視は必ずしも要請されない。

(2) すべての方式に共通の中間的な画像記述を規定

3次元画像方式の共存を前提として、それぞれのデータ交換を個別に行うと、組合せの数だけの交換方法が必要になる。

これに対して、ネットワークにおいてすべての方式に共通の中間的な画像記述法を規定しておけば、方式相互のデータ交換方法をすべて考える必要はなく、中間記述との間のデータ交換方法すなわち符号化方法だけを考えるだけでよい。

これは、次に述べるように「3次元画像情報の統一符号化」へとつながる発想である。

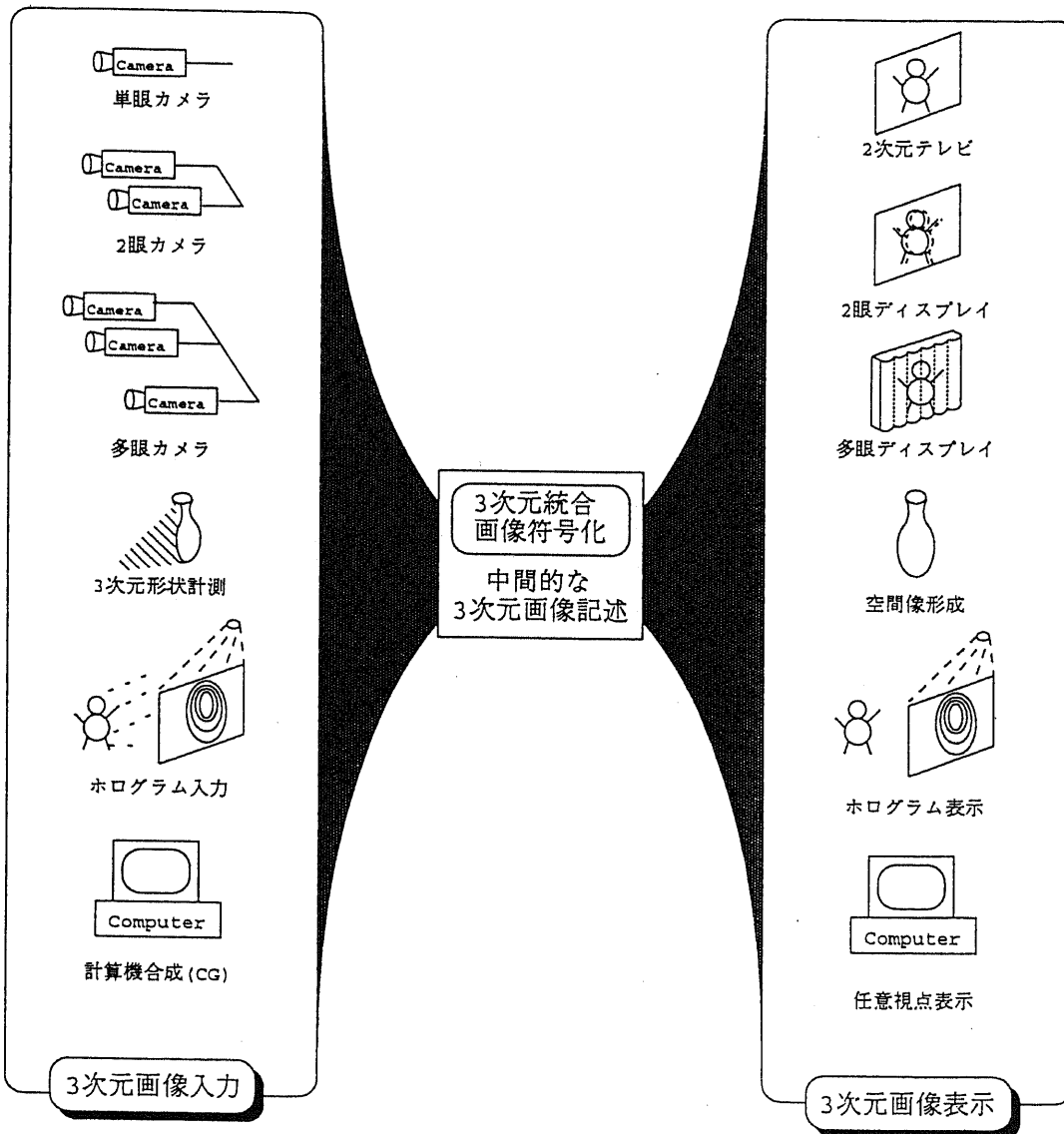


図 2.9: 3次元統合画像通信

3次元統合画像通信の実現には「3次元画像表示方式に共通の中間的な画像記述」を見出すことが必要であるが、これが一体何であるかはすぐには分からない。3次元統合画像通信は入出力を独立に規定できることを目指すものであるから、とりあえずは各入出力方式の間の相互変換を当面の目標として考えていくことにする。

現在、3次元画像の入力方式としてもっとも有望視されているものは「水平方向に並んだ多眼カメラによる入力」である。この理由としては、多眼カメラ入力は2眼カメラ、単眼カメラ入力を含んでおり、両眼ステレオ表示、任意視点合成表示が容易であること、また風景などにレーザ光を照射することは不可能でありホログラム入力が困難であること、十分に密な間隔で撮影した多眼画像からホログラムの合成が可能である（ホログラフィックステレオグラムという）ことなどである。また水平方向に並んだカメラを用いる理由は、多くの3次元ディスプレイが、縦方向の視差を犠牲にして水平方向の視差のみを与えていることによるものである。

多眼カメラによる入力を考えた場合、カメラの間隔、台数が重要なパラメータとなる。ホログラムの合成を考えた場合なるべく密な間隔で撮影しておくことが必要であるが、一方カメラの台数が増加することにより伝送すべき情報量が増えるのみならず入力装置のコストが増加するという問題も生じる。したがって、入力側ではなるべく少ないカメラ数で撮影し、その中間の視点からの画像は出力側で補間合成できることが望ましい。

これらの点を鑑みて、3次元画像符号化に向けた当面の目標として次の3つを掲げたい。

- 高能率圧縮
- 任意方向からの画像の合成
- 2次元画像符号化方式との両立性

これらを満たす符号化方式を開発することが3次元統合画像通信への足掛かりを与えてくれることになろう。

2.3 空間符号化へ

さらに将来に目を向けた場合、3次元統合画像通信の向うには何が見えるであろうか。

ネットワーク化されたバーチャルリアリティ環境下においては、ネットワークを通じて環境そのものがやりとりされるという状況が想定される。このような状況における符号化とは、もはや「画像」の伝送ではなく、「空間そのもの」を効率良く伝送することが次なる目標となってくるであろう。画像の符号化といった観点ではなく、空間そのものがある中間的な記述に変換し符号化するという観点が重要になってくる。

先に見たように3次元画像の取得にはさまざまな方法があるが、3次元統合画像通信の立場から見れば、空間そのものが持っている情報の一部を異なった形態で観測しているにすぎない、ととらえられるのではないだろうか。

3次元統合画像通信が必要としている「中間的な記述」とは、空間そのものが持っている情報であるに違いない。この観点からは、画像符号化 (picture coding) ではなく、「空間符号化 (space coding)」とも呼ぶべき概念へと発展していくのではなかろうか。その意味で、3次元統合画像通信とは、「空間符号化」へと続く道の第一歩となるであろう。

また、CGと実写が融合した仮想空間や、仮想空間内でのインタラクティブ操作といった技術は、今後ますます発展していくことが予想される。この点を考慮すると、

- 情報量を減らし、かつ編集操作などの処理が行ないやすい形に空間がもつ情報を「構造化」しておくこと、

が究極の目標となるであろう。

これら問題はさまざまなテーマへと発展していく可能性を持っており、今後研究が活発化する分野となるであろう。

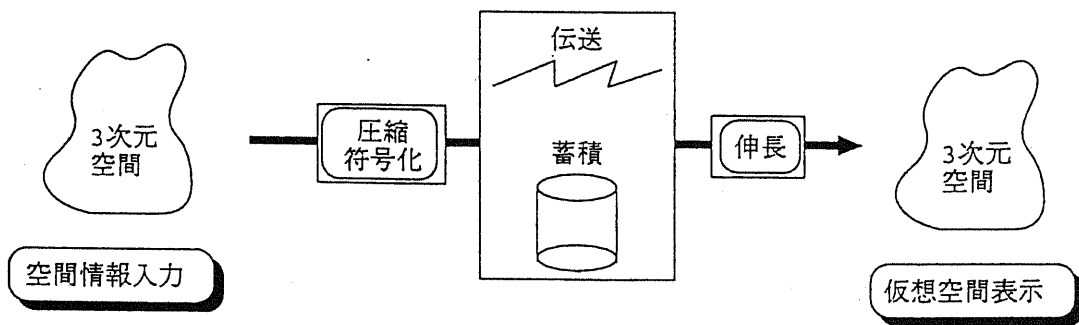


図 2.10: 空間符号化の概念

参考文献

- [1] 原島, 藤井, 苗村: “3次元統合画像通信へむけて”, PCSJ '92(Oct. 1992).
- [2] 原島, 藤井, 苗村: “三次元統合画像通信へ向けて”, ディスプレイアンド イメージング, Vol. 1, pp. 121-131 (Dec. 1992).
- [3] 原島, 藤井, 苗村: “3次元統合画像通信の構想”, 信学技報, HC92-58 (Jan. 1993).
- [4] 大越孝敬: “三次元画像工学”, 朝倉書店 (1991)
- [5] 増田千尋: “3次元ディスプレイ”, 産業図書 (1990)
- [6] 稲田修一 (編著): “三次元映像”, 昭晃堂 (1991)
- [7] 小特集 “3次元ディスプレイの応用”, テレビ誌, 41, 7, pp.604-627 (1987)
- [8] 小特集 “立体視・三次元表示”, テレビ誌, 43, 8, pp.754-789 (1989)
- [9] 特集 “立体感と3次元情報”, テレビ誌, 45, 4, pp.422-519 (1991)
- [10] 濱崎: “3次元情報技術の現状と将来”, テレビ誌, 45, 4, pp.423-430 (1991)
- [11] 佐藤: “動画ホログラフィの研究動向”, テレビ誌, 46, 5, pp.586-593 (1992)
- [12] 首藤: “偏光式立体表示, 液晶シャッタを用いたフィールドシーケンシャルステレオ表示装置”, テレビ誌, Vol. 43, No. 8, pp. 763-767 (Aug. 1989)
- [13] 日置, 外山, 原田, 近藤, 磯野, 安田, 山田: “フリッカレス光学式立体ビデオディスク装置”, 1987 テレビ全大, No. 7-8, pp. 165 (July 1987)
- [14] 安田 (浩), 一之瀬, 渡辺: “3次元画像通信”, テレビ誌, 43, 8, pp.786-789 (1989)
- [15] 井口, 佐藤: “三次元画像計測”, 昭晃堂 (1990)
- [16] 山口, 立平, 秋山, 小林: “2眼式立体画像の情報圧縮と奥行き形状の再現性”, 信学論 D-II, J73-D-II, 3, pp.317-325 (1990)
- [17] M.G.Perkins: “Data compression of stereopairs”, IEEE Trans. Commun., 40, 4, pp.684-696 (1992)
- [18] M.E.Lukacs: “Predictive coding of multi-viewpoint image sets”, ICASSP '86, 1, pp.521-524 (1986)
- [19] W.A.Schupp, 安田 (靖): “視差補償および動き補償を用いたステレオ動画のデータ圧縮”, 1988年画像符号化シンポジウム (PCSJ88), pp.63-64 (1988)

- [20] 泉岡, 渡辺: “視差補償を用いたステレオ動画像符号化”, 1989 信学会春季全大, D-273 (1989)
- [21] 鈴木, 平田: “記録方式”, テレビ誌, 45, 4, pp.467-474 (1991)
- [22] 吉川: “ホログラフィック 3次元画像と符号化”, 平成 3 年電気・情報関連学会連合大会, P2-7 (1991)
- [23] 吉川, 佐々木: “電子ホログラムディスプレイにおける情報圧縮の検討”, 信学技報, IE92-12 (1992)
- [24] 原島, 森川, 青木: “三次元構造モデルを用いる画像の符号化・処理・表示”, 1989 年信学会春季全大, SD-3-12 (1989)
- [25] 原島: “知的画像符号化と知的通信”, テレビ誌, 42, 6, pp.519-525 (1988)
- [26] 伊藤, 藤井, 中川, 三宅: “知的画像入力技術”, テレビ誌, 44, 11, pp.1537-1542 (1990)
- [27] Benton, S.A. (吉川訳): “Experiments in Holographic Video Imaging”, 3D映像, 5, 2, pp.36-68 (1991)
- [28] 宮川: “多次元空間内の定常的確率変数の標本化定理”, 信学誌, 42, 4, p.421 (1959)
- [29] T.Berger: “Rate distortion theory”, Prentice-Hall (1971)
- [30] D.Slepian and J.K.Wolf: “Noiseless coding of correlated information sources”, IEEE Trans. Inf. Th. 19, pp.471-480 (1973)
- [31] 近藤, 森川, 原島: “動画像符号化における Shape from X 手法の検討”, 1990 年画像符号化シンポジウム (PCSJ90), pp.121-124 (1990)
- [32] Morikawa H. and Harashima H.: “3D structure extraction coding of image sequences”, Journal of Visual Communication and Image Representation, 2, 4, pp.332-344 (1991)
- [33] 近藤, 浅井, 原島: “情報の統合による画像からの三次元構造抽出”, 信学技報, IE90-34 (1990)
- [34] 森川, 青木, 原島: “ステレオ動画像からの 3次元構造復元”. 信学技報, PRU89-57 (1989)
- [35] 藤井: “多眼式レンズ板 3次元像のデータ圧縮と拡張”, 東京大学修士論文 (1992)
- [36] 亀山, 大富: “並進表示面を持つ体積走査型ディスプレイの開発とその応用” テレビ誌, Vol. 48, No. 10, pp. 1253-1260 (1994 年 10 月).
- [37] 佐々木, 丹治, 吉川: “ホログラフィック 3次元画像の情報圧縮” テレビ誌, Vol. 48, No. 10, pp. 1238-1244 (1994 年 10 月).
- [38] 本田: 「ホログラフィのはなし」, 日刊工業新聞社.
- [39] 辻内順平 編著: 「ホログラフィックディスプレイ」, 産業図書.
- [40] 原島 博 監修: 「画像情報圧縮」, オーム社 (1991).
- [41] 原島: “知的符号化と知的通信”, テレビ誌, 42, 6, pp. 519-525 (Jun. 1988).
- [42] 画像電子学会「先端の 3次元画像技術」(1992 年 4 月).

第 3 章

3 次元画像の取得と表示方式

3次元画像情報の取得は、主にロボットビジョン、計測の分野でとり扱われてきた内容である。一方、3次元画像の表示に関してはもっぱら表示装置中心の議論がなされてきた。3次元画像の入出力方式を画像通信の観点から体系的に述べたものは少ない。

3次元画像通信の観点から入出力方式の詳細な解析を行なうこと、および画像の位置合わせ、歪み補正、ホログラムの高速計算法など、実際に入出力を行なう際に重要となることについて、体系的にまとめることが本章の目的である。

3.1 はじめに

現状ではさまざまな形態の3次元ディスプレイが提案、実験されている段階であり、入力方式も出力方式に依存して1対1に対応している場合がほとんどである。3次元統合画像通信においては、入力出力方式に依存しない符号化方式の確立が目的であるから、各入力方式ではどのようなデータが得られるのか、また各3次元ディスプレイ上に表示するためにはどのようなデータが必要とされるかについて、詳しい検討が必要となる。

本章では、3次元画像の入出力を考える上で極めて重要な概念である視野と視域について定義した後、現在利用できる3次元画像代表的な入出力方式について概観し、後の符号化方式の提案の際に必要な解析をおこなう。また、実際に実写画像を入力する際に不可欠となる画像の位置合わせ、歪み補正、また入力にコンピュータを用いた場合の高速計算法についても言及する。

3.2 視野と視域

図3.1に示すように、視野 (visual field) とは物体空間の幅 ω を、また視域 (viewing zone) とは視点の動き得る範囲 D を表す。また ω が視点に対して張る角度 ψ を視野角、 D が再生像に対して張る角度 Ω を視域角と呼ぶこともある [1]。

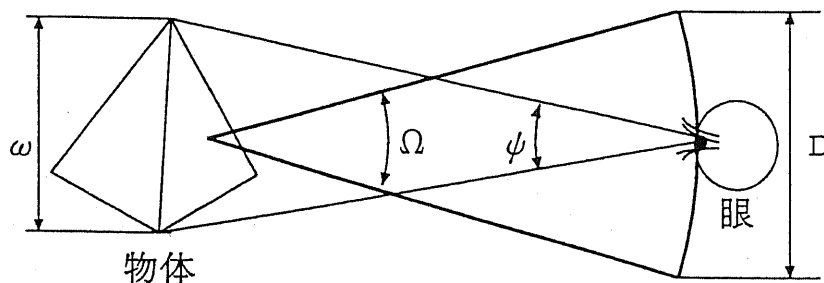


図 3.1: 視野 ω , 視域 D , 視野角 ψ , および視域角 Ω の定義

3.3 3次元画像入力方式の分類

3次元画像の入力方式として、以下のものを考える。

1. 単眼カメラ入力
2. 2眼カメラ入力
3. 多眼カメラ入力
4. コンピュータによる入力 (CGによる合成画像、3次元モデルなど)
5. ホログラム入力
6. キノフォーム入力

上記1から3までは入力としてカメラを用いるものであり、カメラの台数、配置をどのようにするかという点が異なるだけである。したがっていずれの場合でも幾何学的な射影関係を用いて解析できる。したがって、これらを一つのグループとして考える。

4は、CAD/CAMの手法を用いてコンピュータ内部に3次元モデルを構築していくことや、レンジファインダを用いて3次元形状を計測しコンピュータ内部に物体の3次元情報を取り込む場合をも含んだものとして考える。

また、5と6は光の伝搬理論を用いて同じように解析できるものであるから、これらも一つのグループとする。計算機合成ホログラム(CGH)もこの範疇に含めるものとする。

以上のように、入力を3つのカテゴリーに分け、それぞれ詳細な解析を行なう。

3.4 カメラ入力法の解析

文献[1] 4.2節によれば、LS画像(レンティキュラーシートを用いた3次元画像)の撮影法として13項目挙げている。このうち、3次元画像通信の観点からは「連続法¹」は現実的でないと思われ、また多眼カメラの間隔を密に配列したものと等価であると考えられることから、本論文では割愛する。

ここでは、撮影方法を次のように分類してみた。

投影法：

正投影と中心投影がある。通常のカメラによる入力は中心投影を仮定せざるをえない場合が多いが、コンピュータによる3次元画像入力を考えるならば、正投影と等価な画像を得ることができる。

カメラ配置：

視点位置の配置の仕方、およびカメラレンズの光軸の向きについての分類である。前者は、平行配列法、円弧配列法などがあり、後者は平行法、輻輳法などがある。

レンズの光軸と投影面との関係：

レンズの光軸と投影面の原点とを一致させるかどうか、あるいは、レンズの光軸と投影面とが垂直であるかどうかといった分類である。一般に光軸と投影面は垂直であるが、3次元画像取得の場合には両者が垂直ではない「あおり撮影」法が必要な場合もある。

その他の撮影方法：

大口径レンズ(フレネルレンズなど)を通して撮影する方法などがある。

以下の解析では、主にカメラ配列について分類し、それぞれについて2つの投影法(正投影、中心投影)を考える。また、特に必要があると思われる部分でレンズの光軸と投影面との関係を考えていくことにする。これらの組合せを表3.1に示す。3.4.1節で2つの投影方法について見た後に、それぞれにカメラ配置法についての検討を行なう。

3.4.1 正投影と中心投影

本節では、正投影と中心投影について、物体空間中の点から投影面への幾何学的な写像関係を求めることが目的である。

¹カメラを連続的に動かしながら撮影する方法

Table 3.1: カメラによる入力の分類

	直線配列		円弧配列	
平行撮影	正投影 (a)	中心投影 (b)	正投影 (c)	中心投影 (d)
輻射撮影	正投影 (e)	中心投影 (f)	正投影 (g)	中心投影 (h)

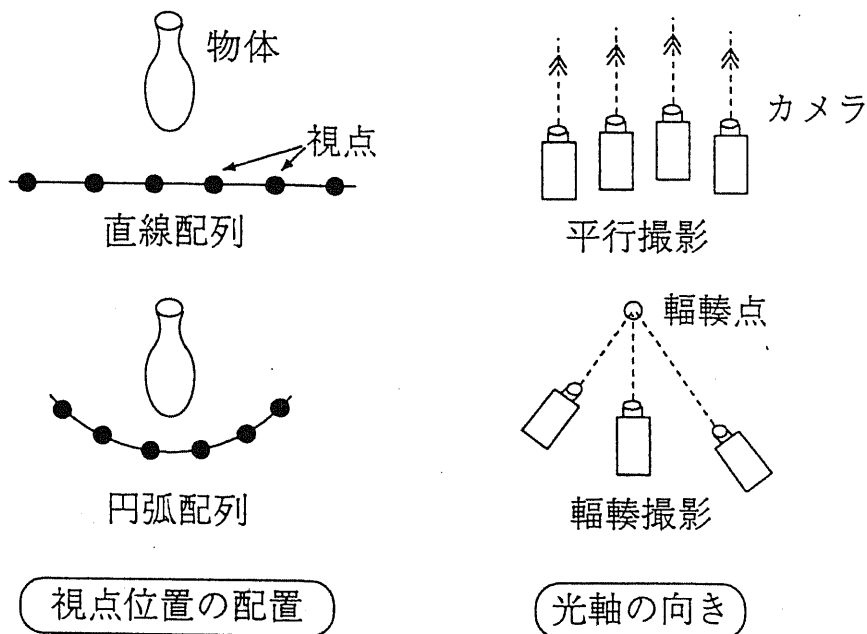


図 3.2: カメラ配置の分類

(1) 正投影 (orthogonal projection)

図 3.3 のように物体空間座標 $O-X-Y-Z$ をとる. 今, 投影面を原点 $O(0,0,0)$ を通る平面とする. 投影面座標を $o-x-y$ とし, x 軸, y 軸は X, Y と同じ方向にとるものとする. この場合, 物体空間座標 (X, Y, Z) から投影面座標 (x, y) への投影は,

$$x = X \tag{3.1}$$

$$y = Y \tag{3.2}$$

と与えられる. 視点の平行移動を考える際に便利のいいように, 同次座標系 (homogeneous coordinate) で表現しておく. 同次座標を (X_h, Y_h, Z_h, W_h) とすると,

$$x = \frac{X_h}{W_h}, \quad y = \frac{Y_h}{W_h} \tag{3.3}$$

ただし, $X_h = X, Y_h = Y, Z_h = 0, W_h = 1$

となる。これを行列表現すると、

$$\begin{bmatrix} X_h \\ Y_h \\ Z_h \\ W_h \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (3.4)$$

となる。(3.4)式の変換行列を P_{ortho} とし、正投影行列と呼ぶことにする。(3.3), (3.4)式により、物体空間から投影面への正投影変換が定義される。

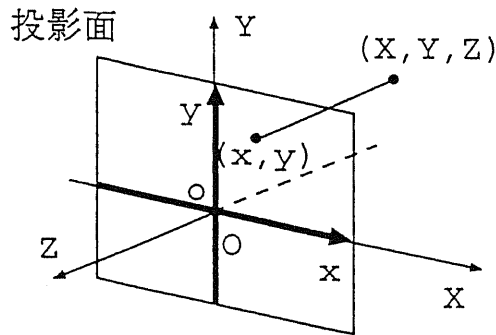


図 3.3: 正投影 (orthogonal projection)

(2) 中心投影 (perspective projection)

図 3.4のように物体空間における座標 $O-X-Y-Z$ をとる。今、投影中心を物体空間の原点 $(X, Y, Z) = (0, 0, 0)$ とし、投影面を $Z = -f$ の平面とする。 f は焦点距離である。また、 $Z = -f$ の平面内での座標系を $o-x-y$ とし、それぞれ X, Y と同じ方向にとる。

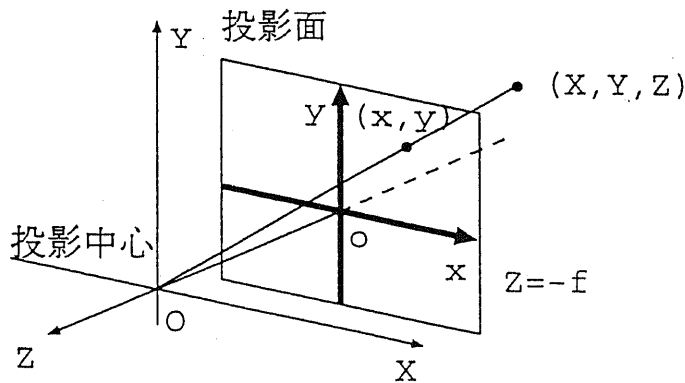


図 3.4: 中心投影 (perspective projection)

さて、物体空間座標 (X, Y, Z) と投影面座標 (x, y) との関係は次式で与えられる。

$$x = -\frac{f}{Z}X \quad (3.5)$$

$$y = -\frac{f}{Z}Y \quad (3.6)$$

これを同次座標 (X_h, Y_h, Z_h, W_h) を用いて表すと,

$$x = \frac{X_h}{W_h}, \quad y = \frac{Y_h}{W_h} \quad (3.7)$$

$$\text{ただし, } X_h = X, \quad Y_h = Y, \quad Z_h = 0, \quad W_h = -\frac{Z}{f}$$

この時、物体空間座標 (X, Y, Z) から同次座標 (X_h, Y_h, Z_h, W_h) への変換は次式で与えられる。

$$\begin{bmatrix} X_h \\ Y_h \\ Z_h \\ W_h \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{f} & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (3.8)$$

(3.8) 式の変換行列を P_{pers} とし、中心投影行列と呼ぶことにする。(3.7), (3.8) 式により、物体空間から投影面への変換が定義される。

3.4.2 カメラ配置による分類

本節では、3.4節で定義した8つの分類に従い、それぞれのカメラ入力方式についての解析を行なうことが目的である。解析に先立ち、以下に平行移動、回転を表す行列を定義しておく。

まず、 (X_0, Y_0, Z_0) 平行移動を表す行列を $T_{(X_0, Y_0, Z_0)}$ とおくならば、

$$T_{(X_0, Y_0, Z_0)} = \begin{bmatrix} 1 & 0 & 0 & X_0 \\ 0 & 1 & 0 & Y_0 \\ 0 & 0 & 1 & Z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.9)$$

と与えられる。また、 Y 軸回りの θ 回転を表す行列を $R_Y(\theta)$ 、 X 軸回りの θ 回転を表す行列を $R_X(\theta)$ 、とすると、それぞれの変換行列は次のようになる。

$$R_Y(\theta) = \begin{bmatrix} \cos \theta & 0 & \sin \theta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \theta & 0 & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad R_X(\theta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta & 0 \\ 0 & \sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (3.10)$$

これらを用いて、先に挙げた8つの分類に従って解析をしていく。

(a),(b),(c),(d) : 平行撮影の場合

正投影の場合、基本的には3.4.1節で述べた投影法をカメラ配列に従って平行移動するだけでよい。

視点²を $V(X_0, Y_0, Z_0)$ におくとする。この場合、視点が原点となるように平行移動した後で投影を行えば良い。すなわち、空間中の点を $(-X_0, -Y_0, -Z_0)$ 平行移動したあと投影すれ

²中心投影法の場合には投影中心である(言葉に注意)。平行投影法の場合の視点の定義はあいまいであるが、ここでは便宜上3.4.1節における原点 $O(0, 0, 0)$ を視点と呼ぶことにする。

ばよく、正投影の場合、

$$\begin{bmatrix} X_h \\ Y_h \\ Z_h \\ W_h \end{bmatrix} = P_{ortho} T_{(-X_0, -Y_0, -Z_0)} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (3.11)$$

となり、中心投影の場合、

$$\begin{bmatrix} X_h \\ Y_h \\ Z_h \\ W_h \end{bmatrix} = P_{pers} T_{(-X_0, -Y_0, -Z_0)} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (3.12)$$

となる。

後は、カメラ配列にしたがって $V(X_0, Y_0, Z_0)$ を定めれば良い。直線配列の場合、

$$X_0 = m\delta X \quad (3.13)$$

$$Y_0 = 0 \quad (3.14)$$

$$Z_0 = F \quad (3.15)$$

となり、円弧配列の場合、

$$X_0 = F \sin m\delta\theta \quad (3.16)$$

$$Y_0 = 0 \quad (3.17)$$

$$Z_0 = F \cos m\delta\theta \quad (3.18)$$

となる。ここで、 m は視点の指標を表す整数変数であり、 F は視点から物体までの距離である。直線配列では δX の等間隔でカメラを配列した場合について、また円弧配列では $\delta\theta$ の間隔でカメラを配列した場合についての例を挙げた。この他にも、 $V(X_0, Y_0, Z_0)$ の定義の仕方により、平面配列³、球面配列などの場合についても、写像関係を記述することができる。

(e),(f),(g),(h) : 輻輳撮影の場合

輻輳撮影の場合、視点と輻輳点との関係に応じてレンズの光軸の角度を変える必要がある。カメラを回転させる場合、視線⁴回りの回転の自由度があるため、カメラ回転の順番には注意が必要である。

ここでは、視点を $V(X_0, Y_0, Z_0)$ とし、カメラの光軸を常に $O(0, 0, 0)$ に輻輳させて撮影させた場合について考える。ここでは、計算上便利なように極座標表現を用いることにし、投影中心の物体空間座標を図のように $V(r, \alpha, \beta)$ とする。 $P(r, \alpha, \beta)$ と (X_0, Y_0, Z_0) との関係は次のようになる。

$$d = \sqrt{X_0^2 + Y_0^2 + Z_0^2}$$

³カメラアレイによる撮影はこれに当たる

⁴視点と注視点（ここでは輻輳点）とを結ぶ直線として定義する。

$$\begin{aligned} \alpha &= \tan^{-1} \frac{X_0}{Z_0} \\ \beta &= \sin^{-1} \frac{Y_0}{d} \end{aligned} \tag{3.19}$$

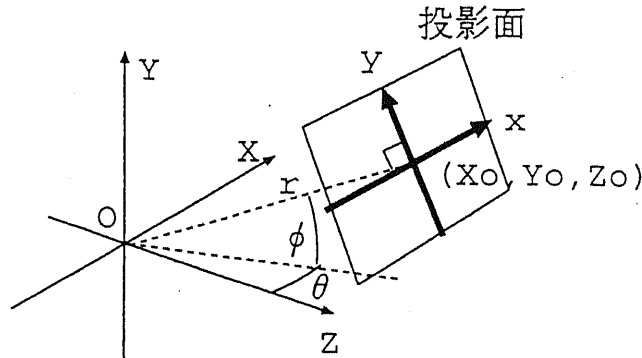


図 3.5: 輻輳撮影

輻輳点 $(0, 0, 0)$ の輻輳撮影の場合の投影変換は、視点 $V(r, \alpha, \beta)$ を回転と平行移動によって原点 $O(0, 0, 0)$ に変換した後、投影を行なうことにより求められる。視点の変換操作は、次の3つのステップにより求められる。

1. Y 軸回りに $-\alpha$ 回転させる (変換行列 $R_Y(-\alpha)$)
2. X 軸回りに $+\beta$ 回転させる (変換行列 $R_X(\beta)$)
3. Z 軸方向に $-d$ 平行移動させる (変換行列 $T_{(0,0,-d)}$)

3つの合成変換を表す行列はつぎのようになる。

$$T_{(0,0,-d)}R_X(\beta)R_Y(-\alpha) = \begin{bmatrix} \cos \alpha & 0 & -\sin \alpha & 0 \\ -\sin \alpha \sin \beta & \cos \beta & -\cos \alpha \sin \beta & 0 \\ \sin \alpha \cos \beta & \sin \beta & \cos \alpha \cos \beta & -d \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.20}$$

$(-X_0, -Y_0, -Z_0)$ 平行移動したあとと投影すればよく、正投影の場合、

$$\begin{bmatrix} X_h \\ Y_h \\ Z_h \\ W_h \end{bmatrix} = P_{ortho} T_{(0,0,-d)} R_X(\beta) R_Y(-\alpha) \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{3.21}$$

となり、中心投影の場合、

$$\begin{bmatrix} X_h \\ Y_h \\ Z_h \\ W_h \end{bmatrix} = P_{pers} T_{(0,0,-d)} R_X(\beta) R_Y(-\alpha) \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{3.22}$$

となる。後は、先ほどの例と同様にカメラ配列にしたがって $V(X_0, Y_0, Z_0)$ を定められることにより、任意のカメラ配列の場合で輻射撮影した場合の投影変換が求められる。

レンズの光軸と投影面との関係

レンズの光軸と投影面の原点とを一致させるかどうか、あるいは、レンズの光軸と投影面とが垂直であるかどうかといった分類である。一般に光軸と投影面は垂直であるが、3次元画像取得の場合には両者が垂直ではない「あおり撮影」法が必要な場合もある。

このような場合の解析として、投影面に並進、回転を施した撮影方法について考える。光軸を Z 軸とし、投影面を X - Y 面に平行な平面であると考ええる。

まず、並進については X - Y 平面内での並進のみ考えればよいので、 $T_{(X_d, Y_d, 0)}$ を考えればよく、

$$T_{(X_d, Y_d, 0)} = \begin{bmatrix} 1 & 0 & 0 & X_d \\ 0 & 1 & 0 & Y_d \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.23)$$

と与えられる。

次に、投影面の回転について考える。ここでは、投影面が投影面内座標原点を中心に回転する場合を考える。したがって、正投影の場合には $(0, 0, 0)$ を中心、中心投影の場合には $(0, 0, -f)$ を中心とした回転である。

今、考えやすいように $(0, 0, 0)$ を中心とした回転を仮定する。投影面の法線ベクトルを n とし、 n と Z 軸がなす角度を α 、 X - Z 平面がなす角度を β とする。この場合、

$$R_Y(\alpha) = \begin{bmatrix} \cos \alpha & 0 & \sin \alpha & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \alpha & 0 & \cos \alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, R_X(\beta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \beta & -\sin \beta & 0 \\ 0 & \sin \beta & \cos \beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.24)$$

となる。

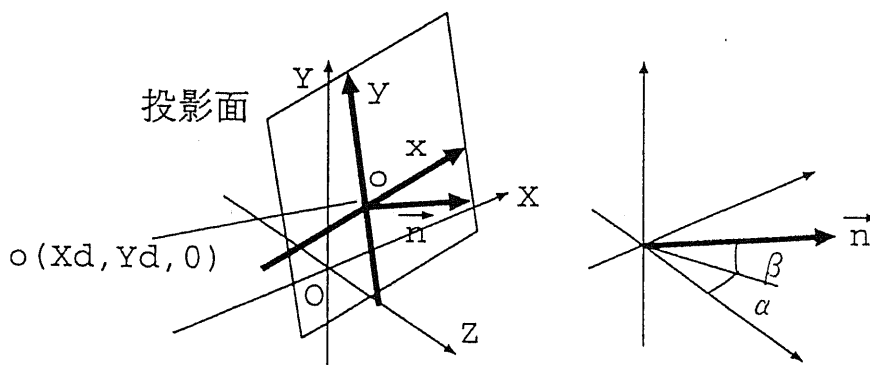


図 3.6: レンズの光軸と投影面との関係 (Z 軸が光軸)

中心投影での平行撮影（輻射、あおりなし）に加えて、物体空間の原点 $O(0, 0, 0)$ が投影面の原点に投影されるように投影面を Z 一定の平面内で平行移動させた場合の投影変換を求め

る。これは、物体空間の原点を基準点とし、各画像を位置合わせをする場合に相当する。(3.28), (3.29)式において、物体空間の原点 $O(0,0,0)$ は投影面上の $(-f\frac{X_0}{Z_0}, -f\frac{Y_0}{Z_0})$ に投影されるので、この点が投影面上の原点となるように移動する。この結果、

$$x = -f \frac{X - \frac{X_0}{Z_0} Z}{Z - Z_0} \quad (3.25)$$

$$y = -f \frac{Y - \frac{Y_0}{Z_0} Z}{Z - Z_0} \quad (3.26)$$

となり、確かに物体空間の原点が投影面上の原点に写像される。ただしこの場合に、投影中心から投影面原点までの距離(=焦点距離)が長くなっていることに注意する必要がある。この場合の焦点距離を F とすると、前章で定義した角度 α を用いて $F \cos \alpha = f$ と表せる。

3.4.3 さまざまなカメラ配置の例

前節までの解析を元に、実際よく行なわれる撮像系についての例をみていくことにする。

直線(平面)配列, 平行撮影

ここでは、光軸を Z と平行に保ったまま視点(投影中心)を平行移動させた場合について述べる。投影中心の物体空間座標を $P(X_0, Y_0, Z_0)$ とする。この時、投影面は $(X_0, Y_0, Z_0 - f)$ を原点とし、 X - Y 平面と平行な面である。この場合の投影変換は、物体空間を平行移動させ、投影中心 $P(X_0, Y_0, Z_0)$ が原点 O と一致するように平行移動させてから中心投影を行なうことにより求めることができる。

であるから、

$$\begin{aligned} \text{CT} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{f} & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & -X_0 \\ 0 & 1 & 0 & -Y_0 \\ 0 & 0 & 1 & -Z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 & -X_0 \\ 0 & 1 & 0 & -Y_0 \\ 0 & 0 & 0 & -Z_0 \\ 0 & 0 & -\frac{1}{f} & \frac{Z_0}{f} \end{bmatrix} \end{aligned} \quad (3.27)$$

従って、

$$x = \frac{X_h}{W_h} = -f \frac{X - X_0}{Z - Z_0} \quad (3.28)$$

$$y = \frac{Y_h}{W_h} = -f \frac{Y - Y_0}{Z - Z_0} \quad (3.29)$$

を得る。

直線（平面）配列, 輻輳撮影

さらに, 原点 $O(0, 0, 0)$ を投影中心, 投影面を $Z = -f$ とした中心投影を行なうことにより, 輻輳撮影の場合の投影変換を求めることができる. 変換行列は次のようになる.

$$CA_3^{-1}A_2^{-1}A_1^{-1} = \begin{bmatrix} \cos \alpha & 0 & -\sin \alpha & 0 \\ -\sin \alpha \sin \beta & \cos \beta & -\cos \alpha \sin \beta & 0 \\ 0 & 0 & 0 & 0 \\ -\frac{1}{f} \sin \alpha \cos \beta & -\frac{1}{f} \sin \beta & -\frac{1}{f} \cos \alpha \cos \beta & \frac{d}{f} \end{bmatrix} \quad (3.30)$$

同次座標系での成分は,

$$\begin{aligned} X_h &= X \cos \alpha - Z \sin \alpha \\ Y_h &= -X \sin \alpha \sin \beta + Y \cos \beta - Z \cos \alpha \sin \beta \\ Z_h &= 0 \\ W_h &= -\frac{1}{f}(X \sin \alpha \cos \beta + Y \sin \beta + Z \cos \alpha \cos \beta - d) \end{aligned} \quad (3.31)$$

であり, これにより投影面座標に変換される. 以上により, 輻輳撮影による投影変換は次のようになる.

$$x = \frac{X_h}{W_h} = -f \frac{X \cos \alpha - Z \sin \alpha}{X \sin \alpha \cos \beta + Y \sin \beta + Z \cos \alpha \cos \beta - d} \quad (3.32)$$

$$y = \frac{Y_h}{W_h} = -f \frac{-X \sin \alpha \sin \beta + Y \cos \beta - Z \cos \alpha \sin \beta}{X \sin \alpha \cos \beta + Y \sin \beta + Z \cos \alpha \cos \beta - d} \quad (3.33)$$

また, 特に実用上重要な水平面内での回転 ($\beta = 0$) について解を求めておくと, $\sin \beta = 0, \cos \beta = 1$ より,

$$x = -f \frac{X \cos \alpha - Z \sin \alpha}{X \sin \alpha + Z \cos \alpha - d} \quad (3.34)$$

$$y = -f \frac{Y}{X \sin \alpha + Z \cos \alpha - d} \quad (3.35)$$

となる.

円弧配列, 輻輳撮影

物体を全周から撮影した場合に相当し, 「全周画像」とも呼ばれる. この場合には,

$$X_0 = F \sin m\delta\theta \quad (3.36)$$

$$Y_0 = 0 \quad (3.37)$$

$$Z_0 = F \cos m\delta\theta \quad (3.38)$$

とおけばよい.

その他の応用例

今後バーチャルリアリティの発展にともない, 空間がもつ情報をそのまま伝送するというニーズが高まってくるであろう. その際に, 重要となるであろう撮像方式について見ておく. これらの幾何学的な投影関係は前節までの式の組合せにより容易に求めることができる.

まず一つ目は円柱配列である。この場合、

$$X_0 = r \sin m\delta\theta \quad (3.39)$$

$$Y_0 = n\delta Y' \quad (3.40)$$

$$Z_0 = r \cos m\delta\theta \quad (3.41)$$

とおけばよい。

これをさらに発展させ、球面配列というのも考えられるであろう。

$$X_0 = r \cos n\delta\phi \sin m\delta\theta \quad (3.42)$$

$$Y_0 = r \sin n\delta\phi \quad (3.43)$$

$$Z_0 = r \cos n\delta\phi \cos m\delta\theta \quad (3.44)$$

とおけばよい。

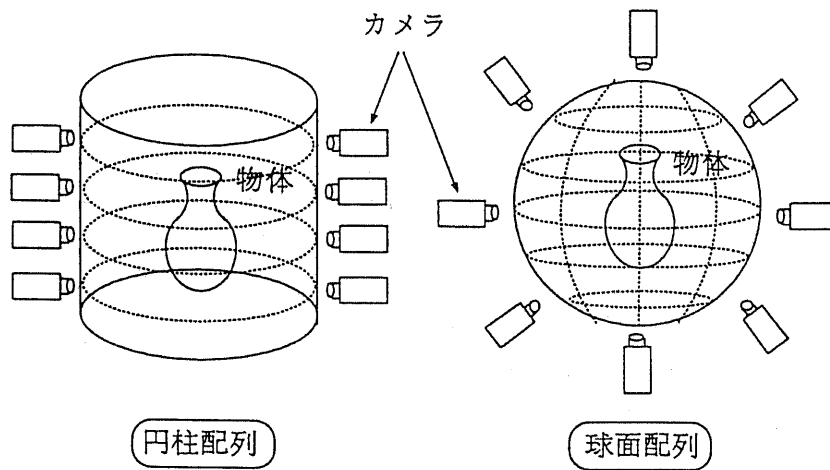


図 3.7: 円柱配列, 球面配列 (輻輳撮影)

3.4.4 フレネルレンズを用いて撮影した場合

図 3.9 に大口径レンズを用いた撮影法の概念図を示す。カメラ配列は直線配列であり、大口径レンズの焦点位置に被写体を配置する。カメラのピントは無限大に設定している。

解析に先立ち、幾何光学的な光の伝搬を記述するのに便利な光線行列について述べる。光線の位置と向きを成分とする 1×2 の行列として表す。光線の伝搬は、行列をかけることにより表される。代表的なものについて述べる。

- 何もない空間中の距離 d の伝搬

$$\begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix} \quad (3.45)$$

- 焦点距離 F の凸レンズ

$$\begin{pmatrix} 1 & 0 \\ -\frac{1}{F} & 1 \end{pmatrix} \quad (3.46)$$

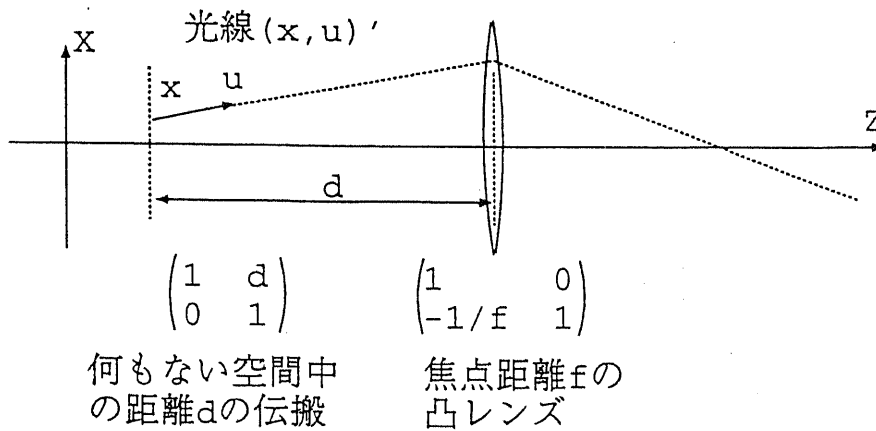


図 3.8: 光線行列

さて、この光線行列を用いて、フレネルレンズを用いて撮影した場合についての解析を行なう。図のように座標系をとる。Z = F の位置に焦点距離 F のフレネルレンズをおく。投影中心 P から投影面までの距離を f₁、P からフレネルレンズまでの距離を f₂ とする。f₂/f₁ = f と定義する。投影面と投影中心の位置の座標は、下向きを正にとることにする。これは、実際の撮影条件との整合を考慮したものである。Z 軸から投影中心までの距離を X₀、投影面上での座標を x とする。さて、投影面上の点 x には物体空間のどの点が投影されるかについて考察する。x と投影中心を結ぶ光線の、フレネルレンズ上での光線行列は、(X₀ - fx, x/f)' である。この光線を凸レンズおよび距離 F の空間を伝搬させる。Z = 0 における平面上では、

$$\begin{pmatrix} 1 & F \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1/F & 1 \end{pmatrix} \begin{pmatrix} X_0 - fx \\ x/f \end{pmatrix} = \begin{pmatrix} Fx/f \\ -1/F(X_0 - fx) + x/f \end{pmatrix} \quad (3.47)$$

となる。この結果は次の重要な点を意味している。すなわち、このような撮像系を用いた場合、

x に対応する Z = 0 の平面上の点 Fx/f は、投影中心位置 X₀ および f₂ に依存しない

のである。これをくだけて言うならば、

Z = 0 にある正方形は、どの視点から撮影しても等しい大きさの正方形として投影される

とも言えるだろう。これが3次元画像の情報処理を行なう上でしばしば取り扱い易い幾何学的な性質を与えている。

また、レンズ上から Z = 0 の平面上への像の拡大率は、

$$\frac{Fx/f}{fx} = \frac{F}{f_1 f} = \frac{F}{f_2} \quad (3.48)$$

と与えられる。特に、拡大率が1(f₂ = F)となる時、任意の視点からの像は、物体空間中の平行光線を記録することになる。このような撮像系でサンプリングされた物体空間は規格化空間と呼ばれ、3次元画像の解析を行なう際には非常に重要である [7]-[12].

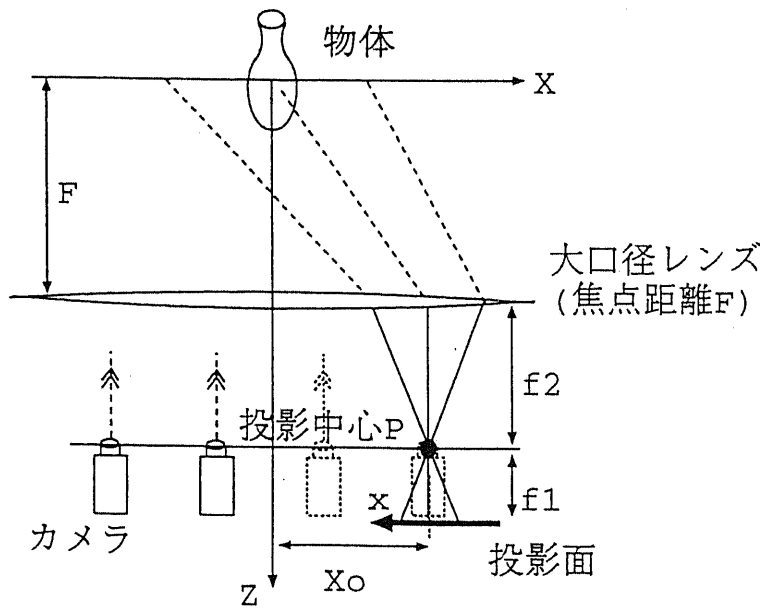


図 3.9: 大口径レンズを用いた撮影法

3.5 幾何学的補正

3.5.1 キーストーン歪み補正

先に導入した「平行撮影+原点位置合わせ」は、後の処理を簡単にするために極めて重要なものである。しかしながら、実際の撮影においては被写体が画角からはみだすことを防ぐため、「輻射撮影」を行なうことが必要な場合が多い。本章では「輻射撮影」により得られた画像から「平行撮影+原点位置合わせ」による画像への補正を行なう幾何学的変換手法について述べる。文献 ([12]) でおこなったキーストーン補正もこの一種である。

投影中心を極座標表示で $P(r, \alpha, \beta)$ とし、 P から投影面の原点 O' までの距離を F とおく。「輻射撮影」における投影面 S 上の座標系を $O'xy$ とし、「平行撮影+原点位置合わせ」における投影面 S' 上の座標系を $O'x'y'$ とおく。2つの座標系の原点 O' は一致している。「輻射撮影」から「平行撮影+原点位置合わせ」への変換は、 S 上の点 $Q(x, y)$ から直線 PQ と S' 面との交点 $Q'(x', y')$ への変換として与えられる。

まず、 S 上の点 $Q(x, y)$ の物体空間座標を求める。直線 PQ を求める際便利なように直交座標系で求めることにする。 S 上の点 $Q(x, y)$ は、点 $(x, y, d - F)$ を X 軸回りに $-\beta$ 、 Y 軸回りに α 回転させた点であるから、点 Q の物体空間座標 $Q(X_q, Y_q, Z_q)$ は先に定義した変換行列を用いて、

$$\begin{bmatrix} X_q \\ Y_q \\ Z_q \\ 1 \end{bmatrix} = A_2 A_1 \begin{bmatrix} x \\ y \\ d - F \\ 1 \end{bmatrix} = \begin{bmatrix} \cos \alpha & -\sin \alpha \sin \beta & \sin \alpha \cos \beta & 0 \\ 0 & \cos \beta & \sin \beta & 0 \\ -\sin \alpha & -\cos \alpha \sin \beta & \cos \alpha \cos \beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ d - F \\ 1 \end{bmatrix} \quad (3.49)$$

と表せる。

一方、投影中心座標を $P(X_p, Y_p, Z_p) = (r \sin \alpha \cos \beta, d \sin \beta, d \cos \alpha \cos \beta)$ とすると、直線 PQ は、

$$\frac{X - X_p}{X_q - X_p} = \frac{Y - Y_p}{Y_q - Y_p} = \frac{Z - Z_p}{Z_q - Z_p} \quad (3.50)$$

である。直線 PQ と S' 面との交点 $Q'(X'_q, Y'_q, Z'_q)$ は、 S 面の方程式 $Z = (r - F) \cos \alpha \cos \beta$ を (3.49) 式に代入することにより得られ、 $Q'(X'_q, Y'_q, Z'_q)$ から S' 面内の座標系 $Ox'y'$ への変換は、

$$(x', y') = (X_q - (r - F) \sin \alpha \cos \beta, Y_q - (r - F) \sin \beta) \quad (3.51)$$

で与えられる。以上より、「輻輳撮影」から「平行撮影+原点位置合わせ」への変換は、

$$x' = \frac{Fx \cos \beta}{x \sin \alpha + y \cos \alpha \sin \beta + F \cos \alpha \cos \beta} \quad (3.52)$$

$$y' = \frac{F(x \sin \alpha \sin \beta + y \cos \alpha)}{x \sin \alpha + y \cos \alpha \sin \beta + F \cos \alpha \cos \beta} \quad (3.53)$$

となる。特にカメラが水平面内で移動する場合 ($\beta = 0$) には、

$$x' = \frac{Fx}{x \sin \alpha + F \cos \alpha} = \frac{1}{\frac{\sin \alpha}{F} + \frac{\cos \alpha}{x}} \quad (3.54)$$

$$y' = \frac{Fy \cos \alpha}{x \sin \alpha + F \cos \alpha} = \frac{y}{1 + \frac{x}{F} \tan \alpha} \quad (3.55)$$

となる。

実際に、(3.32),(3.33) 式を (3.52),(3.53) 式に代入することで (3.26),(3.26) 式と一致することを確認することができる。

3.6 視差画像の位置合わせ

精密な撮像装置を使うことが困難な場合や、スキャナを用いて視差画像を入力下場合には、被写体に位置合わせ用の印をつけておき、撮影した画像の印を元に画像の位置合わせを行なうことが必要となる。ここでは2つの例を挙げ、視差画像の位置合わせの手法について述べる。

図は視差画像の例である。原画像はCGにより作成されたものであるが、スライドフィルムに焼付け後、スライドとして提供されたものである。これを元にして視差画像を得るためには、まずこれをフィルムスキャナでデジタル化し、マークにしたがって位置合わせをする必要がある。

本稿ではまず、位置合わせのための理論的考察を行なう。線形、非線形変換についてパラメータを最小2乗法を用いて求める一般的な手法について概観する。次に、画像の幾何学的変換についての線形、非線形の例を挙げ各々の場合について変換パラメータ算出の方法について述べる。

3.6.1 定式化

最小2乗法を次のように定式化する。

関数 f は、 θ をパラメータとし、 x を変数とする関数とする。ただし、

$$y = f(\theta, x) \quad (3.56)$$

$$\theta = (\theta_1, \theta_2, \dots, \theta_p)' \quad (3.57)$$

$$x = (x_1, x_2, \dots, x_m)' \quad (3.58)$$

$$y = (y_1, y_2, \dots, y_n)' \quad (3.59)$$

l 個の測定データ $(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)$ をもとに θ を求める問題として定式化する。 ($l > p$)

3.6.2 例1: 一般的な線形関数

$n = 1, p = m = k$ として, 線形関数とする.

$$\begin{cases} y_1 = x_{11}\theta_1 + x_{12}\theta_2 + \cdots + x_{1k}\theta_k \\ y_1 = x_{21}\theta_1 + x_{22}\theta_2 + \cdots + x_{2k}\theta_k \\ \vdots \\ y_l = x_{l1}\theta_1 + x_{l2}\theta_2 + \cdots + x_{lk}\theta_k \end{cases} \quad (3.60)$$

$B = (y_1, y_2, \dots, y_l)'$, $\theta = (\theta_1, \theta_2, \dots, \theta_k)'$, $A = (x_{ij}) (1 \leq i \leq l, 1 \leq j \leq k)$ とおくと, 上式は $B = A\theta$ と表すことができ,

$$J = \|B - A\theta\|^2 = (B - A\theta)'(B - A\theta) \quad (3.61)$$

を最小にする θ を求めることが目的である. $\frac{\partial J}{\partial \theta} = 0$ より, 次の正規方程式を得る.

$$A'A\theta = A'B \quad (3.62)$$

さて, $A'A$ が正則ならば逆行列が存在して,

$$\theta = (A'A)^{-1}A'B \quad (3.63)$$

となる.

3.6.3 例2: 非線形関数

評価関数 J に関して,

$$\frac{\partial J}{\partial \theta_1} = \frac{\partial J}{\partial \theta_2} = \cdots = \frac{\partial J}{\partial \theta_p} = 0 \quad (3.64)$$

となるところは, 前章までと同様である. ただ, (3.64) 式の連立方程式が非線形であるために, 一回の行列計算で解が求まらない. ここでは Newton 法により近似解を求めることとする.

いま, $\theta = (\theta_1, \theta_2, \dots, \theta_p)'$, $F(\theta) = (F_1(\theta), F_2(\theta), \dots, F_p(\theta)) = (\frac{\partial J}{\partial \theta_1}(\theta), \frac{\partial J}{\partial \theta_2}(\theta), \dots, \frac{\partial J}{\partial \theta_p}(\theta))$ と定義すると, 解くべき連立方程式は,

$$F(\theta) = 0 \quad (3.65)$$

とかける. 解に十分近い初期近似値 $\theta^0 = (\theta_1^0, \theta_2^0, \dots, \theta_p^0)'$ から始めて,

$$\theta^{k+1} = \theta^k + \Delta\theta^k \quad (3.66)$$

として, パラメータ θ を更新していく. ここで, $\Delta\theta^k$ は k 回目の繰り返しにおけるパラメータの更新値であり,

$$J(\theta^k)\Delta\theta^k = -F^k \quad (3.67)$$

の解である. $J(\theta^k)$ は $F(\theta)$ の点 θ^k における Jacobi 行列である. (3.67) 式を行列表現で書き下すと次のようになる.

$$\begin{pmatrix} \frac{\partial F_1^k}{\partial \theta_1} & \frac{\partial F_1^k}{\partial \theta_2} & \cdots & \frac{\partial F_1^k}{\partial \theta_p} \\ \frac{\partial F_2^k}{\partial \theta_1} & \cdots & \cdots & \frac{\partial F_2^k}{\partial \theta_p} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial F_p^k}{\partial \theta_1} & \cdots & \cdots & \frac{\partial F_p^k}{\partial \theta_p} \end{pmatrix} \begin{pmatrix} \Delta\theta_1^k \\ \Delta\theta_2^k \\ \vdots \\ \Delta\theta_p^k \end{pmatrix} = - \begin{pmatrix} F_1^k \\ F_2^k \\ \vdots \\ F_p^k \end{pmatrix} \quad (3.68)$$

3.7 画像の幾何学的変換

本章では、2次元画像の幾何学的変換について、その種類と最小2乗法によるパラメータの算出方法について述べる。

今、画像中の n 個の点をベクトル $A_1 = (x_1, y_1), A_2 = (x_2, y_2), \dots, A_n = (x_n, y_n)$ と表し、これに幾何学的な変換を施して、 $B_1 = (X_1, Y_1), B_2 = (X_2, Y_2), \dots, B_n = (X_n, Y_n)$ と一致させるようにする問題を考える。 θ をパラメータとした幾何学的な変換を f_θ とし、 A_i が A'_i に写されるとすると、

$$A'_i = f_\theta(A_i) \quad (3.69)$$

とかける。ただし、 $A'_i = (x'_i, y'_i)$ である。ここで、評価関数 J を、

$$\begin{aligned} J &= \sum_{i=1}^n \|B_i - f_\theta(A_i)\|^2 \\ &= \sum_{i=1}^n \|B_i - A'_i\|^2 \\ &= \sum_{i=1}^n \{(X_i - x'_i)^2 + (Y_i - y'_i)^2\} \end{aligned} \quad (3.70)$$

と定め、 J を最小にするパラメータ θ を求めることが目的である。 θ_j で偏微分することにより、

$$\frac{\partial J}{\partial \theta_j} = 2 \sum_{i=1}^n \left\{ -X_i - x'_i \frac{\partial x'_i}{\partial \theta_j} - (Y_i - y'_i) \frac{\partial y'_i}{\partial \theta_j} \right\} = 0 \quad (3.71)$$

を得る。ここで、 $F_j = -\frac{1}{2} \frac{\partial J}{\partial \theta_j}$ と定義すると、

$$F_j = \sum_{i=1}^n \left\{ (X_i - x'_i) \frac{\partial x'_i}{\partial \theta_j} + (Y_i - y'_i) \frac{\partial y'_i}{\partial \theta_j} \right\} = 0 \quad (3.72)$$

となり、解くべき方程式は、

$$F_j = 0 \quad (1 \leq j \leq p) \quad (3.73)$$

となる。この解を Newton 法により求める。 F_j を次のように x と y の2つに分解する。

$$F_j = F_{xj} + F_{yj} \quad (3.74)$$

ただし、

$$\begin{cases} F_{xj} = \sum_{i=1}^n (X_i - x'_i) \frac{\partial x'_i}{\partial \theta_j} \\ F_{yj} = \sum_{i=1}^n (Y_i - y'_i) \frac{\partial y'_i}{\partial \theta_j} \end{cases} \quad (3.75)$$

(3.67) 式によれば、

$$\sum_k \left(\frac{\partial F_{xj}}{\partial \theta_k} + \frac{\partial F_{yj}}{\partial \theta_k} \right) \Delta \theta_k = -(F_{xj} + F_{yj}) \quad (3.76)$$

$$\begin{cases} \frac{\partial F_{xj}}{\partial \theta_k} = \sum_{i=1}^n \left(-\frac{\partial x'_i}{\partial \theta_k} \frac{\partial x'_i}{\partial \theta_j} + (X_i - x'_i) \frac{\partial^2 x'_i}{\partial \theta_k \partial \theta_j} \right) \\ \frac{\partial F_{yj}}{\partial \theta_k} = \sum_{i=1}^n \left(-\frac{\partial y'_i}{\partial \theta_k} \frac{\partial y'_i}{\partial \theta_j} + (Y_i - y'_i) \frac{\partial^2 y'_i}{\partial \theta_k \partial \theta_j} \right) \end{cases} \quad (3.77)$$

以上より、(3.76) 式を解くためには $\frac{\partial x'_i}{\partial \theta_j}, \frac{\partial y'_i}{\partial \theta_j}$ (1次微分)、 $\frac{\partial^2 x'_i}{\partial \theta_k \partial \theta_j}, \frac{\partial^2 y'_i}{\partial \theta_k \partial \theta_j}$ (2次微分) が必要となる。以下に、種々の f_θ による変換の例を挙げ、各々についてのパラメータ θ の算出方法について述べる。非線形変換の場合、微分表をつける。

3.7.1 アフィン変換

アフィン変換とは、次の式による変換である。

$$\begin{pmatrix} x'_i \\ y'_i \end{pmatrix} = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \begin{pmatrix} x_i \\ y_i \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \quad (3.78)$$

この場合、 $\theta = (a_1, b_1, c_1, a_2, b_2, c_2)$ となり6つのパラメータとなる。したがって、3点以上のデータが必要となる。この変換は線形変換であるから、1章で述べたパラメータ算出手法がそのまま利用できる。

3.7.2 回転、平行移動、スケーリング

アフィン変換の一部であるが、パラメータ意味が陽に表れる変換を考える。具体的には次のような変換を考える。

$$\begin{pmatrix} x'_i \\ y'_i \end{pmatrix} = \begin{pmatrix} s_x & 0 \\ 0 & s_y \end{pmatrix} \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} x_i \\ y_i \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix} \quad (3.79)$$

ただし、

$$\begin{cases} \text{回転角} & : \phi \\ \text{平行移動パラメータ} & : t_x, t_y \\ \text{スケーリングパラメータ} & : s_x, s_y \end{cases} \quad (3.80)$$

である。この場合、 $\theta = (\phi, t_x, t_y, s_x, s_y)$ となり5つのパラメータとなる。したがって、3点以上のデータが必要となる。前節で考察したアフィン変換に比べてパラメータの数が1つ減っていることに注意すべきである。これにより例えば、回転、スケーリング、回転といった変換はできず、したがって斜め方向に縮小するといった変換は、この式では表現できない。また、この変換は非線形変換であるから、逐次近似によりパラメータを収束させる必要がある。

上記5つのパラメータ、 $\theta = (\phi, t_x, t_y, s_x, s_y)$ をすべて変数としてもよいが、場合によってはそのうち幾つかを固定し、パラメータの数を減らすことも考えられる。

スケーリングパラメータを次の3つの場合に分類し、各々の場合について(1)回転のみ、(2)平行移動のみ、(3)回転+平行移動を考える。

$$\begin{cases} s_x = s_y = 1 (\text{スケーリングなし}) \\ s_x = s_y (\text{縦横とも同じスケーリング}) \\ s_x, s_y (\text{縦横でスケーリングが独立}) \end{cases} \quad (3.81)$$

$s_x = s_y = 1$ (スケーリングなし) の場合

$$\begin{pmatrix} x'_i \\ y'_i \end{pmatrix} = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} x_i \\ y_i \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix} \quad (3.82)$$

パラメータは3つであり、 $\theta = (\theta_1, \theta_2, \theta_3) = (\phi, t_x, t_y)$ とする。

したがって2点(すなわち4つの座標)のデータがあれば未知パラメータを定めることができる。1次および2次微分表は次のようになる。

	x'_i	y'_i
$\frac{\partial}{\partial \theta_1}$	$-x_i \sin \phi - y_i \cos \phi$	$x_i \cos \phi - y_i \sin \phi$
$\frac{\partial}{\partial \theta_2}$	1	0
$\frac{\partial}{\partial \theta_3}$	0	1

	$\frac{\partial x'_i}{\partial \theta_1}$	$\frac{\partial x'_i}{\partial \theta_2}$	$\frac{\partial x'_i}{\partial \theta_3}$
$\frac{\partial}{\partial \theta_1}$	$-x_i \cos \phi + y_i \sin \phi$	0	0
$\frac{\partial}{\partial \theta_2}$	0	0	0
$\frac{\partial}{\partial \theta_3}$	0	0	0

	$\frac{\partial y'_i}{\partial \theta_1}$	$\frac{\partial y'_i}{\partial \theta_2}$	$\frac{\partial y'_i}{\partial \theta_3}$
$\frac{\partial}{\partial \theta_1}$	$-x_i \sin \phi - y_i \cos \phi$	0	0
$\frac{\partial}{\partial \theta_2}$	0	0	0
$\frac{\partial}{\partial \theta_3}$	0	0	0

$s_x = s_y$ (縦横とも同じスケール) の場合

$$\begin{pmatrix} x'_i \\ y'_i \end{pmatrix} = \begin{pmatrix} s & 0 \\ 0 & s \end{pmatrix} \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} x_i \\ y_i \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix} \quad (3.83)$$

パラメータを, $\theta = (\theta_1, \theta_2, \theta_3, \theta_4) = (\phi, t_x, t_y, s)$ とする. 微分表は次のようになる.

	x'_i	y'_i
$\frac{\partial}{\partial \theta_1}$	$s(-x_i \sin \phi - y_i \cos \phi)$	$s(x_i \cos \phi - y_i \sin \phi)$
$\frac{\partial}{\partial \theta_2}$	1	0
$\frac{\partial}{\partial \theta_3}$	0	1
$\frac{\partial}{\partial \theta_4}$	$x_i \cos \phi - y_i \sin \phi$	$x_i \sin \phi + y_i \cos \phi$

	$\frac{\partial x'_i}{\partial \theta_1}$	$\frac{\partial x'_i}{\partial \theta_2}$	$\frac{\partial x'_i}{\partial \theta_3}$	$\frac{\partial x'_i}{\partial \theta_4}$
$\frac{\partial}{\partial \theta_1}$	$s(-x_i \cos \phi + y_i \sin \phi)$	0	0	$-x_i \sin \phi - y_i \cos \phi$
$\frac{\partial}{\partial \theta_2}$	0	0	0	0
$\frac{\partial}{\partial \theta_3}$	0	0	0	0
$\frac{\partial}{\partial \theta_4}$	$-x_i \sin \phi - y_i \cos \phi$	0	0	0

	$\frac{\partial y'_i}{\partial \theta_1}$	$\frac{\partial y'_i}{\partial \theta_2}$	$\frac{\partial y'_i}{\partial \theta_3}$	$\frac{\partial y'_i}{\partial \theta_4}$
$\frac{\partial}{\partial \theta_1}$	$s(-x_i \sin \phi - y_i \cos \phi)$	0	0	$x_i \cos \phi - y_i \sin \phi$
$\frac{\partial}{\partial \theta_2}$	0	0	0	0
$\frac{\partial}{\partial \theta_3}$	0	0	0	0
$\frac{\partial}{\partial \theta_4}$	$x_i \cos \phi - y_i \sin \phi$	0	0	0

s_x, s_y (縦横でスケールが独立) の場合

$$\begin{pmatrix} x'_i \\ y'_i \end{pmatrix} = \begin{pmatrix} s_x & 0 \\ 0 & s_y \end{pmatrix} \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} x_i \\ y_i \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix} \quad (3.84)$$

パラメータを, $\theta = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5) = (\phi, t_x, t_y, s_x, s_y)$ とする. 微分表は次のようになる.

	x'_i	y'_i
$\frac{\partial}{\partial \theta_1}$	$s_x(-x_i \sin \phi - y_i \cos \phi)$	$s_y(x_i \cos \phi - y_i \sin \phi)$
$\frac{\partial}{\partial \theta_2}$	1	0
$\frac{\partial}{\partial \theta_3}$	0	1
$\frac{\partial}{\partial \theta_4}$	$x_i \cos \phi - y_i \sin \phi$	0
$\frac{\partial}{\partial \theta_5}$	0	$x_i \sin \phi + y_i \cos \phi$

	$\frac{\partial x'_i}{\partial \theta_1}$	$\frac{\partial x'_i}{\partial \theta_2}$	$\frac{\partial x'_i}{\partial \theta_3}$	$\frac{\partial x'_i}{\partial \theta_4}$	$\frac{\partial x'_i}{\partial \theta_5}$
$\frac{\partial}{\partial \theta_1}$	$s_x(-x_i \cos \phi + y_i \sin \phi)$	0	0	$-x_i \sin \phi - y_i \cos \phi$	0
$\frac{\partial}{\partial \theta_2}$	0	0	0	0	0
$\frac{\partial}{\partial \theta_3}$	0	0	0	0	0
$\frac{\partial}{\partial \theta_4}$	$-x_i \sin \phi - y_i \cos \phi$	0	0	0	0
$\frac{\partial}{\partial \theta_5}$	0	0	0	0	0

	$\frac{\partial y'_i}{\partial \theta_1}$	$\frac{\partial y'_i}{\partial \theta_2}$	$\frac{\partial y'_i}{\partial \theta_3}$	$\frac{\partial y'_i}{\partial \theta_4}$	$\frac{\partial y'_i}{\partial \theta_5}$
$\frac{\partial}{\partial \theta_1}$	$s_y(-x_i \sin \phi - y_i \cos \phi)$	0	0	0	$x_i \cos \phi - y_i \sin \phi$
$\frac{\partial}{\partial \theta_2}$	0	0	0	0	0
$\frac{\partial}{\partial \theta_3}$	0	0	0	0	0
$\frac{\partial}{\partial \theta_4}$	0	0	0	0	0
$\frac{\partial}{\partial \theta_5}$	$x_i \cos \phi - y_i \sin \phi$	0	0	0	0

標準動画画像からの変換（フラワーガーデンの例）

等速度で移動しているカメラから本画像を撮影したと仮定した場合、空間中の各点は画像中で一様に消失点に向かって近付いていくが、近づく速度は消失点に近づくにつれて落ちてくる（分数関数的に；別紙参照）。しかし、注目している物体空間中の点が画像中での消失点から遠い場合には、画像中では等速度で移動していると近似することが出来る。本画像の場合においてそれを確かめるため、注目点の21フレームにわたってのx座標の変化を調べてみた（グラフ：X points）。各点はほぼ直線的に移動していると見なすことができ、上述の近似をおこなっても差し支えないと判断した。

このような考えのもとに、画像を次のように補正するものとする。まず、実験によりオプティカルフローの消失点を求めたところ、 $P(-753.0, 5.6)$ であることが分かった。画像の右上をA(351, 239)点、右下をB(351, 0)点とし、三角形PABの範囲に入っている画像を有効な画像とする。

xを0から351まで走査する。各xについてy軸に平行な直線を考え、これがPA, PBによって切りとられる範囲の画像を、補正画像の高さと一致するように拡大（線形補間）する。以上のようにして補正画像を得る。

3.8 コンピュータによる入力

コンピュータを用いた3次元画像の入力としては、次のようなものが考えられる。

1. CADの手法により3次元モデルを作る
2. レンジファインダにより被写体の3次元形状を直接測定して3次元モデルを作る
3. 多眼画像をCGで作成する
4. キノフォーム、ホログラムパターンをコンピュータにより合成する

1, 2については得られるものが3次元モデルおよび表面テクスチャである。3次元形状を伝送する際になるべく少ない情報量で伝送するための検討が必要であるが、この問題についてはここでは割愛する。

4については第7章で考察することとし、本節では3について考察する。

コンピュータグラフィックスにおいては、米国 Silicon Graphics 社のグラフィックライブラリ GL が事実上の標準となろうとしている。ここでは、GL における投影変換パラメータとの関係を考えることにする。

GL においては投影変換、およびビューイング変換により投影方法を決定している。以下に GL における関数の例を挙げる。

GL における投影変換

```
perspective()
window()
orth()
```

GL におけるビューイング変換

```
lookat()
polarview()
```

(1) --- 正投影

これが一番の基本であるが、x 軸の長さが変化する。

回り込んで見ると x 軸が縮んで見えるということである。

```
{set w, c, f;
loop n:
  ortho(-w/2.0, w/2.0, -w/2.0, w/2.0, 1.0, 1000.0);
  lookat(n*c, 0., f, 0., 0., 0., 0);
}
```

(2) --- 正投影 (x 軸の長さ不変の場合)

(x, z) の位置から見ると、画角は x 軸方向に $\cos(\theta)$ ($\theta = \text{atan}(x/z)$) 倍されて見える。これに対処するため、

「ortho の x 方向を $\cos(\theta)$ 倍し、 x 方向を $1/\cos(\theta)$ 倍に拡大して見えるようにする」の方法をとった。

```
{set w, c, f;
loop n:
  ortho(-w/2.0*cos(th), w/2.0*cos(th), -w/2.0, w/2.0, 1.0, 1000.0);
  lookat(n*c, 0., f, 0., 0., 0., 0);
}
```

3.9 ホログラム入力

ホログラムの干渉縞を直接入力する方法と、コンピュータにより合成する方法とがある。入出力の実際については後の章でまとめて述べることにする。

3.10 表示方式

3.10.1 2眼ステレオ方式

観察者が眼鏡を装着し、眼鏡により左右像の分離を行なう方式である。主なものを以下に挙げる。なお、かっこ内は奥行き知覚の要因である。

- アナグリフ方式（両眼視差, 輻輳）
- 偏光方式（両眼視差, 輻輳）
- 時分割方式（両眼視差, 輻輳）
- 濃度差式（両眼視差, 輻輳）

3.10.2 レンティキュラー方式

多眼式レンズ板3次元像は、レンティキュラシート（Lenticular sheet, LS）と呼ばれる、かまぼこ状の形をした微小レンズの配列からなる板状光学素子を用いて、適当な視差を持った像を見せ、立体感を出す方式である。眼鏡なしで立体像を観察することが可能である。（両眼視差, 輻輳, 運動視差）

3.10.3 ホログラフィ方式

ホログラフィは物体の無数の点からの光の振幅と位相の両方の情報を含んだ干渉パターン（ホログラム）を記録・再生する技術である。ホログラフィによる3次元画像表示は、任意の視点から3次元映像を見ることができ、立体視の要因となる両眼視差, 運動視差, 焦点調節, 輻輳など、すべてを備えているため、最も理想的な3次元表示技術といわれている。（両眼視差, 輻輳, 運動視差, 水晶体の調節）これについては章を改めて述べる。

参考文献

- [1] 大越孝敬:「三次元画像工学」, 朝倉書店, 1991.
- [2] 稲田修一 編著:「三次元映像」, 昭晃堂, 1991.
- [3] 本田:「ホログラフィのはなし」, 日刊工業新聞社.
- [4] 増田千尋:「3次元ディスプレイ」, 産業図書.
- [5] 辻内順平 編著:「ホログラフィックディスプレイ」, 産業図書.
- [6] 井口征士, 佐藤宏介 共著:「三次元画像計測」, 昭晃堂, 1990.
- [7] 濱崎, 深澤:“連続視域レンズ板三次元像の幾何学的性質”, 生産研究, Vol.41, No.11, pp.815-823.(1989)
- [8] J.Hamasaki,M.Fukazawa and R.Ishima, "Data compression using geometrical properties of multi-view lens-plate 3D images", Appendix of "Sampling errors and data compression of multi-view lens-plate 3D images", SPIE vol.1319, pp.350-351 (1990), Journal of 3D Images, 4, 4, pp. 40-41 (1990).
- [9] 深澤, 濱崎:“多眼式レンズ板三次元像の標本化誤差”, テレビ学会誌, Vol.44, No.5, pp.598-607.
- [10] 石間, 濱崎:“多眼式レンズ板三次元像のデータ圧縮”, テレビ学技報, 15, 118, pp.67-72 (1991).
- [11] 石間:“多眼式レンズ板三次元像のデータ圧縮”, 東京大学修士論文 (1991)
- [12] 藤井, 浜崎:“多眼式レンズ板三次元像のデータ圧縮と拡張”, 東京大学修士論文, (1992).

第 4 章

光線群による 3 次元画像の統一的表現

我々の目が観測しているのは物体から反射してくる光線であって、この観点から光線そのものを記述するという方法が考えられる。このように 3 次元空間の情報を「光線群」により記述することによって、通常の 2 次元画像から自然な次元の拡張によりホログラフィを含めたさまざまな 3 次元画像の記述、符号化方式を統一的に論じることが可能となる。

4.1 はじめに

前章でさまざまな撮像方式、表示方式を見てきた。3次元統合画像通信においてはこれらの入出力を独立に規定することを目指している。しかしながら、撮影においては被写体までの距離、カメラ間隔、カメラ数などのパラメータが、また再生においては視野、視域、眼数などのパラメータが異なり、入出力方式の相互変換には複雑な処理が必要となる。そこで本章では、3次元画像を光線群の集まりであるとして、撮像を光線群のサンプリング、表示を光線群の再生であるという見方を導入する。これにより3次元画像の統一的表現が可能となり、3次元画像の伝送とは「像再生空間の伝達」であることが見えてくる。また、光線群のサンプリングという観点から、視差画像の補間合成の問題についても言及する。

4.2 光線の物理的性質

光線の物理的性質について、主にCGで用いられている技術を以下にまとめておく。

観測される光の強さ¹を I とする。物体そのものが光を発していると考えた場合（光源を考慮しない場合がこれに相当する）には、

$$I = k_i \quad (4.1)$$

となる。また、物体そのものが光を発しているのではなく、環境光が物体にあたって反射するというモデルを考えるならば、

$$I = I_a k_a \quad (4.2)$$

となる。ここで、 I_a は環境光であり、すべての物体に共通である。 k_a は反射係数であり、0から1の値をとる。

次に、点光源が存在する場合を考える。視点の移動に伴って表面の光具合が変化するモデルである。次に、ランバート反射と点光源減衰モデルについて考える。

1. ランバート反射

$$I = I_d k_d \cos\theta \quad (4.3)$$

内積記号・を使うならば、

$$I = I_d k_d (N \cdot L) \quad (4.4)$$

と表すことができる。

$$I = I_a k_a + I_d k_d (N \cdot L) \quad (4.5)$$

2. 点光源減衰 (Light-source attenuation)

光源からの距離に従った減衰を考慮した場合、

$$\begin{aligned} I &= I_a k_a + f_{att} I_d k_d (N \cdot L) \\ f_{att} &= \frac{1}{d^2} \end{aligned} \quad (4.6)$$

となる。 f_{att} が減衰項を表す。

¹画像にした場合の輝度値と考えれば良い

また奥行き感を出す技術として、デプスキューが用いられることもある。これは、物体が観測者から遠くなるにしたがって、暗い輝度値で描くことで奥行き感を出すものである。

$$I'_\lambda = s_0 I_\lambda + (1 - s_0) I_{dx\lambda} \quad (4.7)$$

ここで、

$$s_0 = s_b + \frac{(z_0 - z_b)(s_f - s_b)}{z_f - z_b} \quad (4.8)$$

と定義する。

鏡面反射としては、Phong のモデルが知られている。視線の方向が反射方向から α の角度を持っているとする。 α が 0 の時に反射は最大となり、 α が増加するにしたがって反射光は少なくなる。この少なくなり方を $\cos^n \alpha$ で近似するものが Phong モデルと呼ばれるものである。 Phong のモデルは次のように表される。

$$I_\lambda = I_{a\lambda} k_a O_{d\lambda} + f_{att} I_{p\lambda} [k_d O_{d\lambda} \cos \theta + W(\theta) \cos^n \alpha] \quad (4.9)$$

ここで、反射方向を R 、視線方向を V と表すならば、 $\cos \alpha = R \cdot V$ となる。 n は、反射係数の指数係数であって、通常 1 から数百の値をとる。

光の屈折についてはスネルの法則が知られている。スネルの法則 (Snell's law) は、

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{n_2}{n_1} \quad (4.10)$$

というものである。ここで、 θ_1, θ_2 はそれぞれ光線の入射角度、出射角度、 n_1, n_2 は媒質の屈折率である。

4.3 光線群による被写体空間の記述

図 4.1 を用いて実空間の情報を光線群により記述する考え方を述べる。被写体空間の座標を $O-X-Y-Z$ とする。 Z 軸に垂直な面 $Z = z$ を考え、この面を横切る光線を考える。今、光線が $Z = z$ の平面を横切る位置を (x, y) とし、光線の方向を経度・緯度に相当する 2 つの変数 θ, ϕ により表すことにすると、光線は x, y, z, θ, ϕ の 5 つの変数により一意に定められる。この光線の光強度を表す関数を f と定義すると、この空間の光線群データは $f(x, y, z, \theta, \phi)$ と表すことができる。このように光線に着目した場合、3次元空間に加えて光線の方向を表す次元が 2 つ増えるため 5次元の空間となる。さらに、光線群データの時間変化まで考慮するならば、 $f(x, y, z, \theta, \phi; t)$ となり、6次元の空間となる。従来、3次元空間+時間の4次元空間で物体の記述を行えば良いと考えられていたが、実は光線に着目するならば5次元+時間の6次元空間まで考慮しなくてはならないのである。これを「光線空間」と呼ぶことにする。

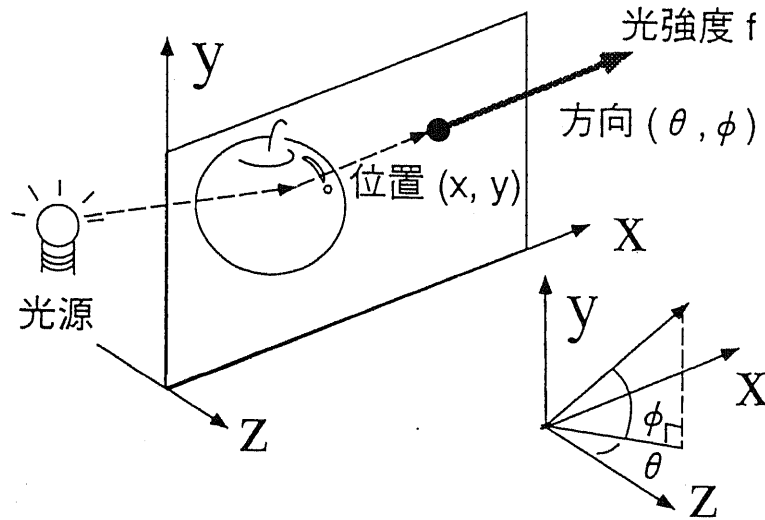
次に、光線空間のデータの観測をおこない、画像、あるいはホログラムとして記録する。まず通常の画像とは、「一点に集まる光線群を記録したもの」と考えることができる。したがって、方向に関する 2 つのパラメータをもつ「2次元」の画像となる。例えば (x_0, y_0, z_0) に視点をおいた場合、 $f(x, y, z, \theta, \phi; t)$ において、

$$f(\theta, \phi)|_{x=x_0, y=y_0, z=z_0, t=t_0} \quad (4.11)$$

という 2次元のデータを記録したものが従来の画像である。一方ホログラフィは、ある断面を通過する光線群をそのまま記録するものである。例えば、 $Z = z_0$ にホログラム面があるとするならば、

$$f(x, y, \theta, \phi)|_{z=z_0, t=t_0} \quad (4.12)$$

ある面を横切る光線を (x, y, θ, ϕ) で表し、
その光強度を表す関数 $f(x, y, \theta, \phi)$ により
3次元画像を記述する



3次元画像は、 $f(x, y, \theta, \phi)$ の4次元空間の
データで記述される

(縦方向の視差を考慮しない場合は $f(x, y, \theta)$)

$x_{min} < x < x_{max}, y_{min} < y < y_{max}$: 視野幅を規定

$\theta_{min} < \theta < \theta_{max}, \phi_{min} < \phi < \phi_{max}$: 視域角を規定

図 4.1: 光線群による空間情報の記述

という4次元空間のデータを記録しているといえる。この4次元光線群データを参照光と干渉させることにより2次元の濃淡値として記録し、光の回折作用によって4次元空間の再生をしているのである。ホログラフィは莫大な情報量が必要と言われるが、これは4次元のデータを2次元面上に記録しようとしているからに他ならない。このように観測、記録とは、6次元の光線空間から部分空間を抽出して記録しているということが出来る。このような表記により、どのような複雑な形状を持った物体であっても、またどのような鏡面反射がおこっている被写体であっても記述ができる。

さて、このように得られた3次元画像情報を3次元ディスプレイ上で表示を行なう。ここで、理想的な3次元ディスプレイを考えてみる。図4.2に3次元ディスプレイの水平断面図を示す。 $z=0$ の面にディスプレイの面があるとし、物体が3次元表示されているとする。この場合の3次元画像は連続変数 x, θ を用いて、 $f(x, \theta)$ と表すことができる。すなわち、 x の位置を通り、 θ の方向への光線の輝度値を $f(x, \theta)$ と定義するのである。ここで、 x は視野内の値をとり、 θ は z が正の方向を0として $-\pi < \theta < \pi$ の値をとる連続変数である。この $f(x, \theta)$ により、どのような複雑な形状を持った物体であっても、またどのような鏡面反射がおこっている被写体で

あっても記述ができる。逆に、この表記の範囲では従来の概念での「3次元画像」と呼ぶことができないものまで含まれていることに注意が必要である。極端な例では視点を变えるごとに全く異なった画像を提示することも可能である。「視点が変わると全く違った絵が見える画像は3次元画像といえるのか?」という疑問もあろうかと思うが、この点に関しては、2次元画像が空間軸方向、時間軸方向に全く相関を持たない画像であってもディスプレイは表現できる能力をもっていることと事情は同じである。

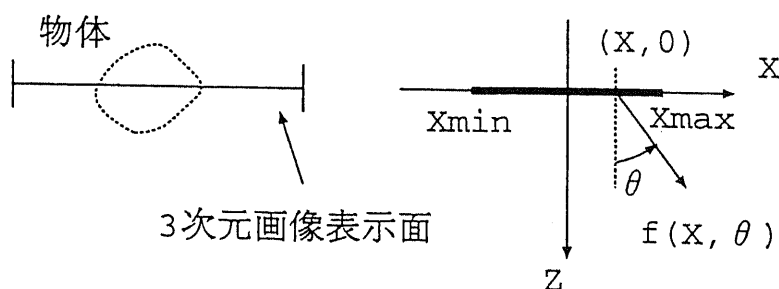


図 4.2: 3次元画像の連続系表現

2次元画像はすべての方向 θ に対して同じ輝度値の光線を発すれば良いので $f(x, y)$ と表すことができる。縦方向の視差をもたない理想3次元ディスプレイは、 $f(x, y, \theta)$ と表現することができ、縦方向の視差を有する理想3次元ディスプレイは $f(x, y, \theta, \phi)$ と表現できる。

すなわち3次元ディスプレイの表現能力としてはこれだけ要求されるということである。このように、

- 2次元画像は「2次元」
- 縦方向の視差をもたない3次元画像は「3次元」
- 縦方向の視差を有する3次元画像は「4次元」

という点が興味深い。

4.4 空間中の一点を通る光線群

さて、光線群の性質をもう少し詳しく見ていくことにしよう。

例として、3次元ディスプレイなどのデバイスにより3次元空間を再現する場合について考察してみよう。この場合、3次元ディスプレイが発することのできる光線群はディスプレイ面上を通過する光線群だけであり、ディスプレイ面より観察者側においては反射や屈折などがなく自由空間を光線が伝搬すると仮定できる。

このように、ある断面を通過する光線群のみを考えることとし、また縦方向の視差は考慮しないことにする。これは f が ϕ に対して一定であることを意味する。さらに、具体的なイメージをつかみやすくするため、ある水平走査線に着目して被写体空間と光線空間との関係を調べてみる。

例として、図4.2に縦方向の視差を考慮しない3次元ディスプレイの水平断面図を示す。 $Z = 0$ 上の点 x を通る θ 方向の光線によりこのディスプレイの光線群表示は、

$$f(x, \theta) \cdot (-x_{min} \leq x \leq x_{max}, -\theta_{min} \leq \theta \leq \theta_{max}) \quad (4.13)$$

となる。ここで、 θ は z 軸正の方向を0とし、反時計回り方向を正とする。また $-x_{min}, x_{max}$ は視野幅を規定し、 $-\theta_{min}, \theta_{max}$ は視域角を規定する。すなわち、被写体空間中の $Z = 0$ 上の一点を通る一本の直線が $f(x, \theta)$ における一点に対応することになる。

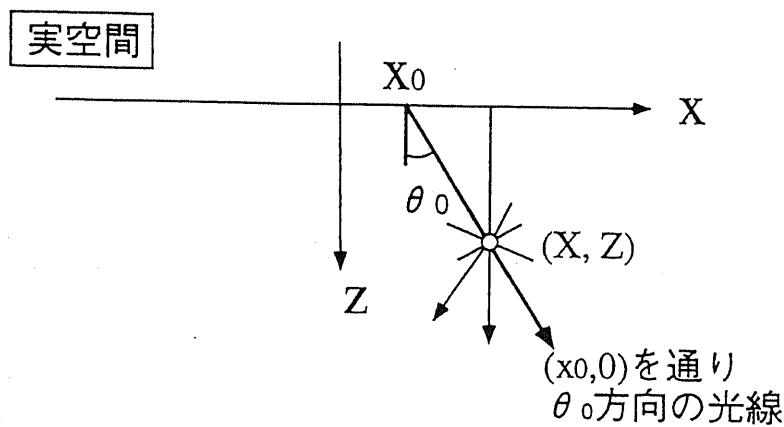
さて、幾何学的な関係により、空間内の一点 (X, Z) を通る光線 (x, θ) は、

$$X = x + Z \tan \theta \tag{4.14}$$

を満たすことが分かる。ここで、便宜のため光線の方向の変数 u を $u = \tan \theta$ と定義し、 (x, u) 空間に変換すると、

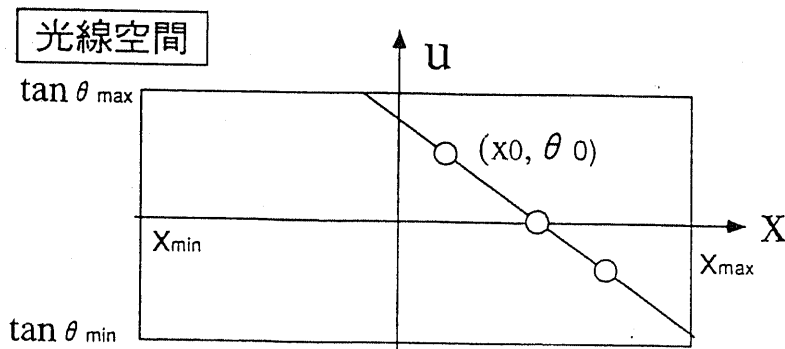
$$X = x + uZ \tag{4.15}$$

となり、被写体空間中の一点 (X, Z) を通る光線群は (x, u) 空間内においては直線上にあることが分かる。



(X, Z) を通る光線は、幾何学的な関係により、

$$X = x + uZ \quad (\text{ただし、} u = \tan \theta)$$



(x, u) 空間では直線となる

図 4.3: 被写体空間 (x, z) と $f(x, u)$ 空間との関係

4.5 撮像による光線群のサンプリング

4.15が意味することは、カメラを (X, Z) において撮影した場合に (x, u) 空間中の4.15式の直線に沿ったデータが記録されるということである。すなわち、1台のカメラによる撮影とは、 $f(x, u)$ のデータの内、一つの直線に沿ったデータのみを記録していることになる。多眼カメラによる撮影とは、 $f(x, u)$ のデータのうち、多数の直線に沿ったデータをサンプリングして記録していることに相当する。カメラを等間隔で配置した場合には、図3のような多数の直線上のデータをサンプリングして記録することとなる。

4.15式のもう一つの特徴は、 (X, Z) に視点のある観察者に対しディスプレイが提示すべきデータが4.15式に沿ったデータであるということである。これは、 $f(x, u)$ のデータがすべて利用できるならば視域内の任意の点からの視差画像が合成できることを意味する。このデータを元にしてディスプレイまでの視距離、眼数の違いなどに応じて4.15式を参照して $f(x, u)$ から必要なデータを取り出して提示すればよい。ある視域、視野に限った場合の任意の点からの視差画像の合成を行なうためには、 $f(x, u)$ のデータをすべて再生できればよい。ホログラフィにおいては視域内の任意視点からの画像が観察できるが、 $f(x, u)$ のデータをすべて記録しているからに他ならない。

「カメラによる撮影」を $f(x, \theta)$ との関連から論じる。

まず簡単のため正投影 (orthographic projection) を仮定する。すなわち、ある視点からの画像は、ある $\theta = T$ からの画像であるとする。さて、 x 方向のサンプリング間隔は画像の場合一定としてよく、 Δx の間隔で M 画素の画像を撮影するものとする。この撮影により得られた1次元の画像は、

$$F(m)_T = f(m\Delta x, T) \quad (4.16)$$

と表すことができる。

また、 K 個の視点から撮影したものとして、視点に相当する θ を $\theta_1, \theta_2, \dots, \theta_k$ とおく。このように定義することにより多眼画像を画素のインデックス m と視点のインデックス n で表すことができ、撮影により得られた多眼画像の離散表現を $F(m, n)$ とおくと、

$$F(m, n) = f(m\Delta x, \theta_n) \quad (4.17)$$

と表すことができる。

多眼3次元画像を得るためには、このように連続な光線群からなる被写体空間を、「ある解像度、ある眼数」で、撮影をおこなう。すなわち、多眼3次元画像とは、連続信号 $f(x, u)$ を x, u それぞれに対し、ある間隔でサンプリングしたものということができる。

次に実際のカメラモデルである中心投影 (perspective projection) を仮定する。まず被写体空間内の一点と $f(x, \theta)$ との関係を述べる。点 $P(X, Z)$ にある点光源からの光は、次の式に従って $f(x, \theta)$ 空間に投影される。

$$X = x + \tan \theta Z \quad (4.18)$$

便宜のため、光線の方向の変数 t を

$$u = \tan \theta \quad (4.19)$$

と定義し、 $f(x, u)$ 空間で考えると、

$$X = x + uZ \quad (4.20)$$

となり、被写体空間内の点は、 $f(x, u)$ 空間には直線として投影されることが分かる。

点光源からの光と考えたが、カメラのレンズ中心の一点を通る光線群が4.15式により与えられると考えることができる。

以上まとめると、

カメラによる撮影とは、 $f(x, u)$ 内の (4.15) 式に沿ったデータを、あるサンプリング間隔でサンプリングしたもの

ということができる。一本の直線がカメラの位置と1対1に対応している。先に考えた正投影とは、 $t = \text{const.}$ の直線に沿ったサンプリングである。多眼カメラによる撮影とは、 $f(x, u)$ 内の多数の直線に沿ったデータを記録していることになる。

このように $f(x, u)$ は本質的に直線群から構成される。すなわち、

カメラを直線上に配置することにより (m, n) 画像は直線群から構成される

のではなく、

(m, n) 画像は撮像系によらず本質的に直線群から構成されるものである

ということである。

「多眼画像を並べて水平断面をとったもの」という表現はエビポーラ平面画像を表しているが、特にカメラを直線上に配列し、しかも正撮影（光軸が常に平行）を行った場合に「エビポーラ平面画像は直線群から構成される」のである。

4.6 光線空間のもつ冗長性

光線空間による表現は3次元統合画像通信における中間的記述としての可能性を持つことが示された。次の問題は、光線空間をいかに記述し、符号化するかという点である。それにはまず光線空間のもつ冗長性についてみておく必要がある。

1. (x, y) 断面において：普通の画像と同様な統計的冗長性をもつ。
2. (x, θ) 断面において：物体上の一点から発生する光線の方向による変化を示すもの。光源の位置、物体の傾き、表面の材質に依存する冗長性がある。

1の性質はDCTなどの従来の2次元画像符号化方式を適用することで冗長性の削減が可能である。一方、2の性質は、物体上の点は鏡面反射などの特殊な場合を除いて視線方向により大きな明るさの変化があることは少ないことを意味する。この性質を利用したものが「視差補償予測符号化」である。

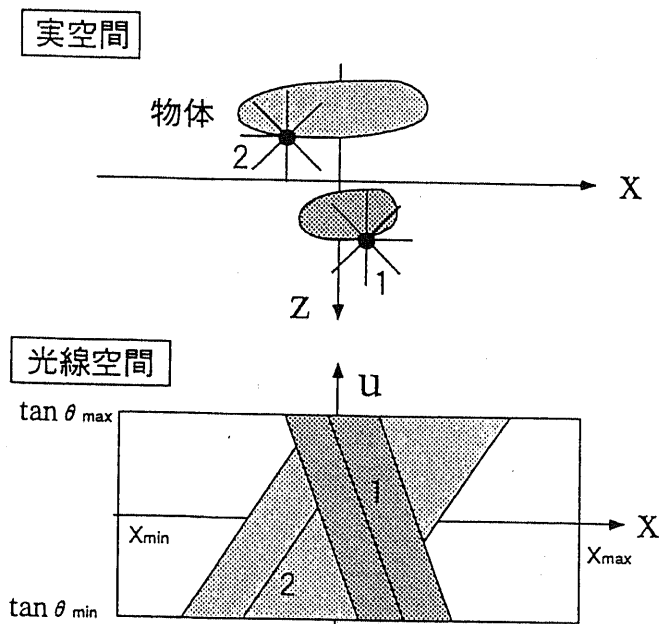
また、具体的な符号化方法としては、

1. レンジ画像+テクスチャとして記述
2. 3次元構造モデルとして記述

といった方法を思いつくことができる。これらの手法を光線空間の記述という観点から眺めると、直観的な従来の手法がこの枠組みの中に含まれていることが分かる。実際、3次元構造モデルとして記述する手法は $u = 0$ の軸上で光線空間を構成する直線の傾きと輝度値とに符号化したと考えることができる。また、物体の表面反射モデルは、直線に沿った輝度値の変化の仕方をモデリングしたものにとらえることができる。

さて、「光線空間の記述」という観点から視差補償予測符号化、構造抽出符号化、モデルベース符号化を同じ土俵で考えることができる。

二つ目は、(1)式における直線の傾きが物体の奥行き情報に対応していること、および $f(x, \theta)$ の $\theta = 0$ の輝度情報が正面からの正投影像に相当していることに着目し、光線空間を(2次元画像) + (奥行き情報)としてモデリングし、SNRが最も良くなるように奥行き情報の最適化をおこなったものである。 $f(x, \theta)$ のデータを被写体空間 (X, Z) 空間における折れ線形状 + 表面テクスチャで近似するものであり、 y 方向も考慮するならば三角形パッチ + 表面テクスチャにより $f(x, y, \theta)$ のデータを近似していることに相当する。この手法による処理例を図4に示す。(2次元画像) + (付加情報)により3次元画像の記述ができるため、2次元画像符号化方式との両立性(コンパチビリティ)がある。



物体上の一点から発せられる光線データ

視線方向の変化に伴って
明るさが大きく変化することは少ない



(x, u)空間の冗長性

図 4.4: (x, u) 空間の冗長性

4.7 3次元画像符号化の定式化

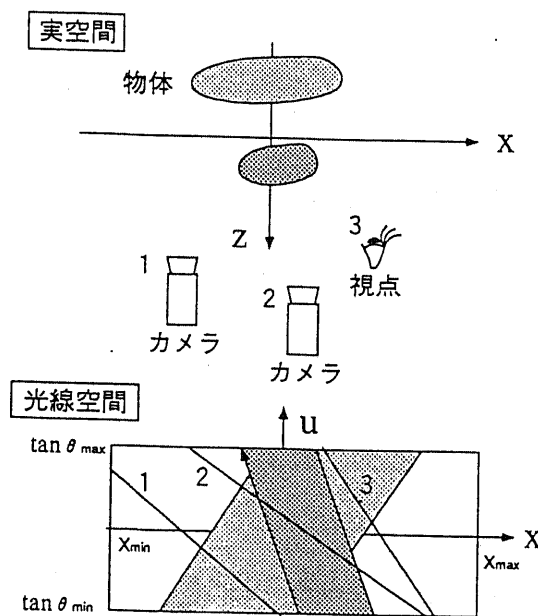
多眼画像の圧縮、補間という問題は、

伝送すべき情報量、復号画像の SNR などを総合的に考えた上で、いかに (m, n) を圧縮、符号化するかという問題

であり、かつ、

サンプリングされた u_i 以外の u' における画像を補間合成する問題

としてとらえることができる。このように考えていくと、多眼画像の符号化を $F(m, n)$ の圧縮符号化という様に狭くとらえるべきではなく、究極的な目的は、送信側における $f(x, u)$ を復号側において再構成することであるということが見えてくる。このように $f(x, u)$ 表現は被写体空間の情報を全て持っているものと考えることができ、これを符号化、伝送することは被写体空間を符号化、伝送するものと考えられる。これを特に「像空間の符号化」と呼ぶことにする。図4.7に像空間符号化の全体像を示す [9].



- 1, 2. カメラにより記録される光線データ (x, u) 空間からのサンプリング
3. 視点位置に対応して提示すべき光線データ (x, u) 空間内のデータの取り出し

図 4.5: 3次元画像の連続系表現

4.8 誤差評価法

3次元画像の評価尺度としてどのようなものを用いればよいのだろうか。究極的には人間が見て自然な3次元画像と見えるかどうかで評価すべきであるが、人間の立体視時の視覚特性についてはまだ研究段階であり、はっきりとした指標がないのが現状である。ここでは、画像符号化の考え方を導入し、SNRを用いた画質評価という考え方に立つことにする。前節の考えからすれば、究極的には送信側における $f(x, \theta)$ が受信側においてどれだけ忠実に再生されるかにより評価すべきであろう。

そこで以下のように x, θ に関する誤差関数を定義する。

$$e(x, \theta) = f(x, \theta) - f'(x, \theta) \quad (4.21)$$

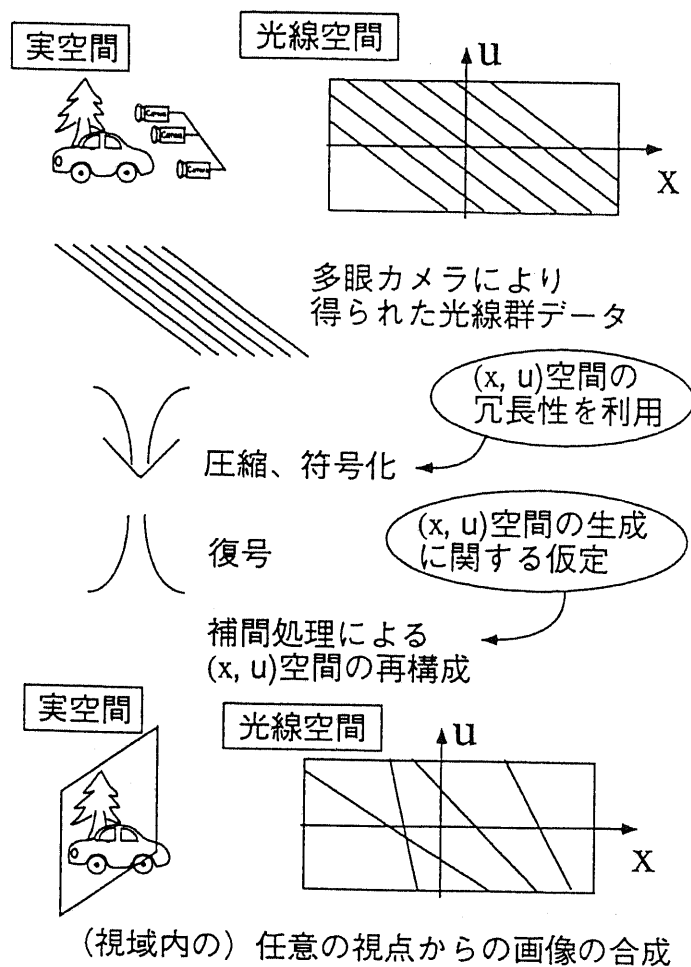


図 4.6: 光線空間伝送による3次元画像の伝送

これを用いて次の誤差評価尺度を定義することができる。

$$\text{Error} = \int_{\Theta} \int_X |e(x, \theta)|^2 dx d\theta \quad (4.22)$$

ここで X は視野を表し、 Θ は視域を表す。また、ある視点からの画像の誤差評価は

$$\text{MSE}(n) = \int_X |e(x, \theta_n)|^2 dx \quad (4.23)$$

と定義できる。2次元画像符号化方式との両立性を考慮し、中央画像の品質を高めたい場合には、 $\text{MSE}(0)$ に重みをつけて Error を評価すれば良い。

4.9 視差画像の補間合成の問題

全節において補間合成の問題を、「サンプリングされた u_i 以外の u' における画像を補間合成する問題」とした。本節ではこの点について考察する。

$f(x, \theta)$ は全く制約条件がなければ単なる2次元信号とみなせるので、 θ 方向に対してサンプリング定理を適用することにより、

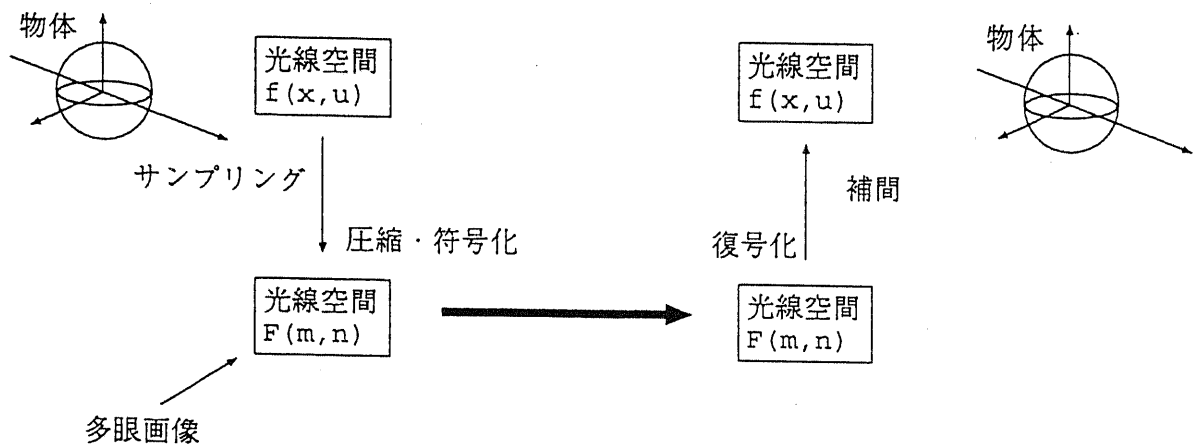


図 4.7: 像空間符号化の全体像

$f(x, \theta)$ において θ 方向の最高周波数の 2 倍の周波数でサンプリングする

ことが必要であり、これが満たされない場合には折り返し歪みを生じることになる。実際には $f(x, \theta)$ は現実世界の制約にしたがって生成された信号であるが、対象が複雑な形状をしていたり、複雑な反射率、透過率を持ったものである場合、現実世界の制約をすべて考慮することは不可能である。このような最悪の場合には、 $f(x, \theta)$ を制約条件のない 2 次元信号とみなすこととなり、上記の条件が解を与えている。

一方、純粋な数学的な興味からは、ある制約条件のもとで $f(x, \theta)$ が生成されている場合いくつかの視点からの、またどの視点からのデータがあれば任意視点からの画像の合成ができるかについての回答が欲しいところである。例えば、

- 視点の方向により輝度データは変わらない (ランベルト則)
- 視点の移動による見え隠れ (オクルージョン) がない

という制約条件がある場合には、2 つの視点からのデータがあり、点の対応関係が完全に分かっているならば (対応点問題が完全に解かれていけば) 任意視点からの画像の合成ができることが分かる。現実的には、「ランベルト則」を仮定できる対象は多いが、「オクルージョンがない」と仮定することは現実的とはいえない。そこで、

- 視点の方向により輝度データは変わらない (ランベルト則)
- 物体の自己隠れ (オクルージョン) はなく、前後 2 つの物体だけからなる

などと制約条件を課した場合に、どの視点からのデータがあれば補間合成ができるかという問題に対する回答はないものだろうか。これらは未解決の問題として残されており、3 次元画像情報理論へと発展していく可能性をもっている。

4.10 $f(x, u)$ の帯域

このようにして導入した $f(x, u)$ であるが、視差画像の補間合成という問題を、この観点から見てみたい。

$f(x, u)$ を純粋な2次元信号として見た場合には、

$f(x, u)$ がもつ帯域の2倍以上の周波数でサンプリングする必要がある

ということになる。例えて言えば、

「穴の中にある宝石を補間合成により再生することは不可能である」

というのはサンプリング周波数が低いことから必然的に生じる結果なのである。サンプリングというとらえ方をした場合、必要な眼数、補間ができる、できないという問題は、 $f(x, u)$ がどれだけの帯域をもつかという点と関連する。次に $f(x, u)$ の帯域という観点から考察してみたい。

まず第1に、ディスプレイの表示能力の問題が挙げられる。レンティキュラーの場合、クロストークといって、ある視点への輝度情報が隣の視点へ「しみだして」しまうという現象がある。これは、ディスプレイの t 方向への再生周波数に限界があると見ることができるだろう。

第2に、撮影によって帯域が制限されてしまうという点が挙げられる。先の議論ではピンホールカメラモデルを使用し、カメラレンズは無限小の開口をもつとしたが、実際の撮影においてはカメラレンズはある有限の開口をもっている。このような場合に、撮影によって t 方向の帯域制限が起きる。これについて以下に述べる。

まず有限の開口をもったレンズを用いた場合、ピントをどこに合わせるかという問題があるが、これは「ディスプレイ面」に合わせるものとする。すなわち、

$$\frac{1}{s} + \frac{1}{Z} = \frac{1}{f} \quad (4.24)$$

が成立しているものと仮定する。ここで s はレンズとフィルムとの距離、 f はレンズの焦点距離である。このようにレンズを用いた撮影の場合には、一点としてフィルムに記録される情報は (x, u) を中心として t 方向に幅を持つことになる。この幅は、レンズの直径を D とすると、

$$t = \frac{D}{Z} \quad (4.25)$$

となる。したがってこのレンズを通して撮影した場合には、 t 方向に dt の大きさで平滑化した結果を記録していることになる。結局、レンズの大きさが平滑化する幅を決定し、レンズが大きくなれば、それだけ低い周波数でのフィルタリングがなされるということになる。

このような帯域制限が意味をもってくる例を挙げよう。ホログラフィックステレオグラムでは100枚以上の枚数の視差画像が必要であると言われるが、このようにカメラ間隔（基線長）が短くなってくるとカメラ間隔が絞りの大きさ以下になることも有り得る。例として、焦点距離60mmのレンズを用いて、絞りF5.6で撮影したとする。この場合レンズの開口は概算で約10mmとなり、これは40cmの視域に100枚の視差画像を撮影するカメラ間隔(4mm)の倍以上の数値となっている。

この数値は、すでにサンプリング定理を満足していることを意味し、たとえ200枚の視差画像が欲しいからといってこの撮像系を用いて200枚の画像を撮影することは無意味であり、また適当なアルゴリズムを用いて200枚の画像を補間合成することも無意味である。信号として t 軸方向への厳密な再生が可能なのである。

これと関連して人間の目のレンズについても同様のことが言える。以下に「レンティキュラーは目の焦点調節と矛盾し、ホログラフィは矛盾しない」という点を考察してみたい。

先に、レンズの大きさに対応して t 方向の平滑化が起こるという話をしたが、平滑化をした結果、ぼける点、ぼけない点が出てくるならば、これはピント調節に対応しているといえる。す

なわち、 t 方向の平滑化の幅に対して提示される $f(x, u)$ の解像度が高ければよいということである。

目の焦点調節機構にあう、あわないという話は、目の水晶体レンズの大きさに比べ、レンティキュラーなどの表示系における $f(x, u)$ の t 方向の解像度が劣っているからに他ならない。極端な話、人間の目のレンズが数 cm の大きさをもつのであれば、レンティキュラーであっても焦点調節に対応するとも言えるのである。

ホログラフィは暗い場所で展示してあることが多く、目の瞳孔を開かせて目の焦点に整合している効果をいっそう引き立てている役割をはたしている。

4.11 その他の応用例

光線空間の概念は入出力デバイスに依存しない空間の中間的なフォーマットを目指すものであるから、多眼画像やCGについてもまったく同じ土俵で考えることができる。例えば、多眼画像は2次元画像の集まりであるが、CGは形状モデル、表面テクスチャモデル、光源モデル、表面反射モデルなどにより記述されている。しかし、これらを光線群表現という観点からは両者の違いはない。したがって、両者を光線群表現に変換し前後の重なりに注意して重ねることによりCGと実写画像が自然に合成される。これにより、例えば多眼カメラにより実写画像を取得し、CGにより合成されたモデルと融合させて任意視点からの画像の再生を行なうことができる。これはCGと実写画像の融合による仮想空間の構築などに応用することが可能である。

また、仮想空間内で人間が奥にあるものを手前にもってくるといったインタラクティブ操作は、物体中心の記述の観点からは物体の奥行きを変化させることになるが、光線群の観点からは光線空間の幾何学的変換とみることができる。先に見たように、 $f(x, \theta)$ 空間内においては直線の傾きが奥行きに対応している。したがって光線群情報を奥行き情報に従ってあらかじめ分割しておき、光線空間内のデータに回転を施すことにより等価的に物体の奥行きを操作することができる。

このように、CGと実写が融合した仮想空間内でのインタラクティブ操作は、光線群の観点から見通し良く取り扱えるようになる。具体的な手法、光線群を容易に取り扱えるようなアーキテクチャをもったコンピュータの研究などがこれからの研究課題として期待される分野である。

4.12 まとめ

3次元統合画像通信の枠組で、空間の情報を光線群により記述する概念の提案をおこなった。本章では、3次元画像のもつ情報は光線群データであるとの観点から撮像系、表示系について考察し、また、3次元画像の伝送を光線群の伝送であると位置付けた。またこの観点から3次元画像の符号化、誤差の定式化をし、また視差画像の補間合成についても言及した。光線に着目すると、従来4次元と思われていた空間は、実は6次元のデータであることが示され、(ホログラフィを含む)画像による記録とは6次元空間からの部分空間の抽出であると定式化した。この観点から3次元画像の符号化を「光線群の記述、符号化」として統一的に論ずることが可能となる。したがって、3次元統合画像通信の章で述べた「入出力に依らない中間的な3次元画像記述」としての有力な候補として光線群記述を考えることができるのではないかと思われる。

参考文献

- [1] 大越孝敬:「三次元画像工学」, 朝倉書店, 1991.
- [2] 稲田修一 編著:「三次元映像」, 昭晃堂, 1991.
- [3] 本田:「ホログラフィのはなし」, 日刊工業新聞社.
- [4] 増田千尋:「3次元ディスプレイ」, 産業図書.
- [5] 辻内順平 編著:「ホログラフィックディスプレイ」, 産業図書.
- [6] 原島 博 監修:「画像情報圧縮」, オーム社 (1991).
- [7] 画像電子学会「先端の3次元画像技術」(1992年4月).
- [8] 安居院猛, 長尾智晴, 共著:「画像の処理と認識」, 昭晃堂, 1992.
- [9] 原島:“知的符号化と知的通信”, テレビ誌, 42, 6, pp. 519-525 (Jun. 1988).
- [10] 日下:“3次元情報の基礎” テレビ誌, Vol. 41, No. 7 pp. 604-609 (1987).
- [11] 安田, 一之瀬, 渡辺:“3次元画像通信”, テレビ誌, Vol. 43, No. 8, pp. 786-789 (Aug. 1989).
- [12] 原島, 藤井, 苗村:“3次元統合画像通信へむけて”, PCSJ '92(Oct. 1992).
- [13] 原島, 藤井, 苗村:“三次元統合画像通信へ向けて”, ディスプレイアンド イメージング, Vol. 1, pp. 121-131 (Dec. 1992).
- [14] 原島, 藤井, 苗村:“3次元統合画像通信の構想”, 信学技報, HC92-58 (Jan. 1993).
- [15] Yamamoto, “Motion Analysis by Visualized Locus Method”, *Transactions of Information Processing Society of Japan*, vol. 22, no. 5, pp. 442-449 (Sep. 1981).
- [16] 山本:“連続ステレオ画像からの3次元情報の抽出”. 信学論. Vol. J69-D No. 11, pp. 1631-1638 (1986).
- [17] R. C. Bolles, H. H. Baker, and D. H. Marimont. “Epipolar-Plane Image Analysis: An Approach to Determining Structure from Motion”. *International Journal of Computer Vision*, 1, pp. 7-55 (1987).
- [18] H.H. Baker, R.C. Bolles, “Generalizing Epipolar-Plane Image Analysis on the Spatiotemporal Surface”, *International Journal of Computer Vision*, 3, pp. 33-49 (1989).

- [19] 苗村, 藤井, 中屋, 原島: “三次元画像符号化基礎検討(1)-誤差評価法-”, 1993年信学会春季全大, D-326 (Mar. 1993).
- [20] 藤井, 原島: “構造抽出に基づく三次元画像符号化” *PCSJ94* (1994-10).
- [21] T. Fujii, H. Harashima, “Coding of autostereoscopic 3D image sequence”, *VCIP '94*, 2308-91 (Chicago, Sep. 1994).

第 5 章

視差補償予測に基づく 3 次元画像符号化

多眼画像とは、同じシーンを多数の方向から撮影したものであるから、視差画像間には強い空間的な相関がある。本章で述べる「視差補償予測符号化」とは、多眼画像データを信号とみなし、信号の統計的冗長性を除くという手法により圧縮を実現するものである。2次元画像符号化方式においては、動き補償予測符号化と離散コサイン変換を基幹とした符号化方式はすでに標準化されていることから、本手法は2次元画像符号化方式との整合性がよいという特徴をもつ。本章では、元となる考え方について述べ、さらにこの手法を多眼画像に適用した実験結果について述べる。また、多眼画像の幾何学的な特徴を用いた効率的な予測符号化について述べ、実験結果を述べる。

5.1 はじめに

多眼画像は水平方向に並んだカメラから撮影した画像であり、これは水平方向に移動するカメラにより撮影した動画画像と等価である。従って、2次元動画画像標準符号化方式における動き補償予測符号化の考え方を応用することが可能であり、これを視差軸方向に適用したものが「視差補償予測符号化」である。本章では、現在標準化がなされている動画画像符号化方式を多眼画像に対して適用した例について述べる。

5.2 従来の研究

5.2.1 2次元画像の圧縮・符号化

2次元画像における符号化・圧縮方式については、すでに数多くの研究がなされており、現在標準化作業が進められている。静止画像符号化においてはJPEG、動画画像符号化においてはDCT (Discrete Cosine Transform, 離散コサイン変換) とMC (Motion Compensation, 動き補償) を基幹とした符号化方式 (通信用符号化方式としてはH.261, 蓄積用符号化方式としてMPEG1, MPEG2など) が提案されている。JPEG, H.261, MPEG1についてはすでにCCITTの勧告が出されており、MPEG2についても93年に勧告が出されている。動画画像符号化方式は、MCにより時間軸方向の冗長性を削減し、DCTにより空間軸方向の冗長性を削減している。

5.2.2 ステレオ画像の相関

2眼画像、あるいは多眼画像は、同一のシーンを視点を少しずつ変えて撮影した画像であるから、各画像間には強い相関があることは容易に想像がつく。実際にステレオペア画像がどの程度の相関があるかを調べた報告がある。近景、中景、遠景についてそれぞれ相関係数を図5.1に示す [2]。

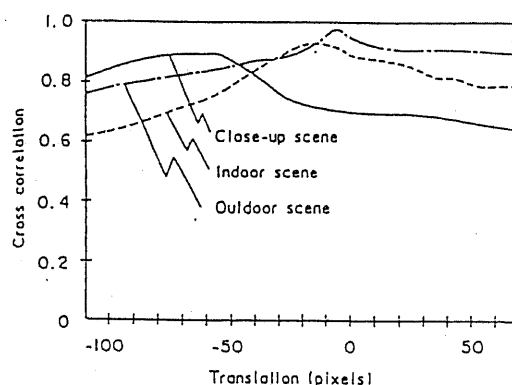


図 5.1: 左右画像の相互相関

5.2.3 3次元画像符号化の研究動向

前節で見たように、ステレオ画像は左右画像の相関が非常に強い。これは、動画画像における隣接フレームと同様の性質をもっており、動画画像符号化において採用されている動き補償を、左右2枚のステレオペア画像に適用する方法が検討されている。空間軸方向に対して、ある視

差をもったペア画像における対応する領域の移動量を、動きベクトル推定と同様な手法により視差推定を行ない、補償するもので、視差補償と言われている。現在のステレオ画像符号化方式の提案は、何らかの形で視差補償を用いているものが多い。最近の研究の中からいくつかを紹介する。

従来、ステレオ画像については符号化の検討がいくつかなされている。ステレオ画像対は水平方向の相関が非常に高く [2]、この性質を利用した符号化方法として「視差補償予測符号化」が提案されている。これは、2次元動画画像符号化標準方式で採用されている「動き補償予測符号化」を視差方向に適用したものである。ステレオ動画を対象として、視差補償予測符号化と時間軸方向の動き補償予測符号化とを組み合わせ、視差補償と動き補償を適応的に切り替える方式 [6] や、左右の動き補償フレーム間差分画像から安定に視差ベクトルを検出する方式の提案 [4] などがある。また、実際にステレオ動画を 9Mbps で符号化した実験の報告がなされている [9]。いくつかの代表的な例を次にまとめる。

- 静止画に対して、片方の画像をブロックに分割し、これを平行移動させて他方との差分をとる方法 [2]。
- ステレオペア画像のうち、一方の画像はDCTを施してそのまま伝送し、もう一方の画像を伝送する代わりに視差情報のみを伝送し、復号側でペア画像を再構成する方法 [16]。
- 動画の場合、視差補償をおこなうより、動き補償を用いる方が有効な場合があるため、この2つを適応的に切替える方式が提案されている。ブロックサイズを 8×8 としたステレオ動画の実験の例で、動き補償フレーム間予測符号化方式に比べ、約 10 - 20 % 程度削減できることを示している [6]。
- 上記の両眼視差による補償を、左右の動き補償フレーム間予測差分画像に対して適用する方式を提案して、さらに、10 - 20 % 程度削減できることを示している。 [4]
- 多眼式3次元像は水平方向に一定の間隔をおいて撮影した視差像の集まりであるから、隣あった2つの視差像をステレオペア画像とみなし、上記の圧縮の手法を適用することが可能である [5]。

この考え方は、レンティキュラ方式などの多眼画像へも自然に拡張される。この視差補償符号化では、中央の2次元画像を基準として、これに視差画像を付け加えることによって3次元化を図っている。その意味では、この符号化は2次元画像システムと両立性のある方式であると言える。

5.3 動画標準符号化方式の適用

本章では、視差補償予測符号化の考え方に基づいて、多眼画像を一種の動画像とみなし、動画像の標準符号化方式である H.261 を用いて符号化する実験を行う。このように、視差像の集まりである多眼画像を一種の動画像とみなした場合、多眼画像の視差補償予測符号化には、盛んに研究が進められているテレビ電話・会議用符号化技術が応用できると考えられる。

5.3.1 標準符号化方式 H.261

映像コーデック勧告 H.261 は「px64kb/s オーディオビジュアルサービス用ビデオ符号化方式」と題され、テレビ会議もしくはテレビ電話を主要な用途と想定し、64kb/s から 2Mb/s まで

の動画通信用の映像符号化として、標準化されたものである。以下に標準符号化方式 H.261 の概要を記す。

1. 基本構成

符号器の構成をあらわす機能ブロック図を図 5.2 に示す。情報源符号器では入力信号に含まれる冗長度を除き、残りの信号を一定の誤差の範囲内で量子化する。ビデオマルチプレックス符号器では、画像信号のほか各種ブロック属性信号を可変長符号化した後、定められたデータ構造の符号列に多重化し、送信バッファで符号化レートの変動を吸収して、伝送路レートに整合させる。また、伝送路符号器は伝送路誤りの影響を防ぐ誤り訂正符号化を行い、符号化制御は送信バッファの残留量によって量子化ステップなどの符号化パラメータを変化させて、発生情報量が一定となるように制御する。

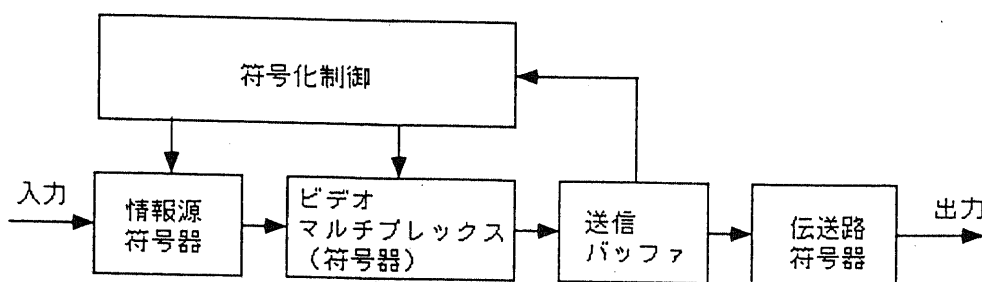


図 5.2: H.261 の構成

2. 画像フォーマット

情報源符号器への入力信号は、テレビ方式の異なる地域間の通信を可能とするため、表 5.1 に示す共通中間フォーマット (CIF: Common Intermediate Format) を規定している。また低いビットレートでの動作のために、解像度を水平・垂直とも半分にした QCIF (Quarter CIF) を定義し、両者が対向した場合は小さい方のフォーマットにより通信することになっている。

パラメータ		CIF	QCIF
1 ラインの画素数 (Y:輝度, C:色差)	Y	360	180
	C _b	180	90
	C _r	180	90
1 フレームのライン数 (Y:輝度, C:色差)	Y	288	144
	C _b	144	72
	C _r	144	72
毎秒フレーム数		29.97	
インタレース		なし	

Table 5.1: 共通中間フォーマット (CIF) の規格

3. 情報源符号化

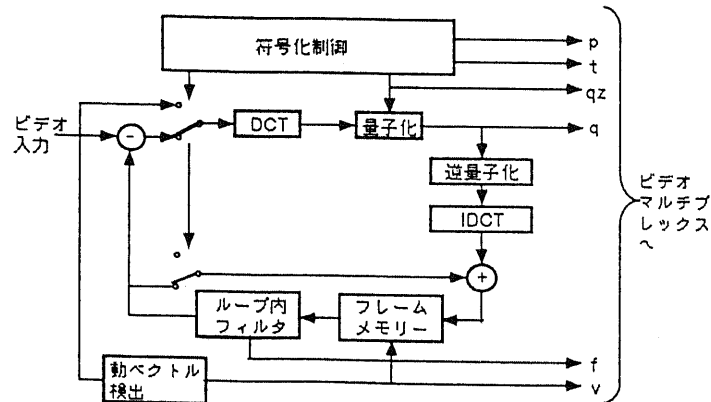


図 5.3: 情報源符号器

図 5.3 に H.261 の情報源符号化方式を示す。基本的には、動き補償フレーム間予測された誤差信号を DCT 符号化し、これを可変長符号化するものである。

符号化モードの単位は、 8×8 画素の輝度信号ブロック 4 個と、同じく 8×8 画素の 2 個の色差信号ブロックであり、これをマクロブロック (MB) 呼ぶ。単位ブロックが比較的大きいのは、低レート符号化における画素あたりのサイド情報 (符号化モード、ブロックアドレス、動きベクトルなど) 量を減らすためである。1 画像フレームは、12 の GOB (Group of Block) から構成され、1GOB は 33 の MB から構成される。

動き補償はこの MB を単位として行われ、その範囲は縦横共 ± 15 画素/フレームで、精度は整数画素である。動き補償は ON/OFF 可能であり、動きベクトル伝送のオーバーヘッドを考慮して、予測誤差電力の減少率により選択される。また動きベクトルは直前のマクロブロックの動きベクトル値との差分が符号化される。差分を取ることでダイナミックレンジは 2 倍に広がるが、表 5.2 に示す折り返し符号を用いることにより符号数を削減している。例えば直前のベクトル値が $+10$ であった場合、差分ベクトル値は $-25 \sim +5$ の範囲にしかありえないので、“ $-8 \& 24$ ” が送られてくれば、差分値は -8 と定めることができる。

動き補償予測は、一般に画像の低域成分に関して有効であるから、予測値に図 5.4 のようなローパス特性を有するループ内フィルタを適用することにより、予測効率を向上させることができる。ただし、比較的高い符号化レートにおけるカメラのパニングのような単純な動きに対しては、動き補償予測の精度が高いため、フィルタをかけない方が効率が高い。このためループ内フィルタも ON/OFF 可能となっている。

またシーンチェンジなどの場合に有効なフレーム内符号化モード (入力信号がそのまま DCT 符号化される) も用意されている。

DCT は 8×8 画素のブロックに対して適用される。一般にフレーム間予測を用いた符号器では符号器の局部復号器と復号器のそれが全く同じ動作をする必要がある。たとえわずかな誤差であっても、符号化のたびに発生する誤差は、フレームメモリに蓄積され、や

動きベクトル	符号語	符号長
-16 & 16	0000 0011 001	11
-15 & 17	0000 0011 011	11
-14 & 18	0000 0011 101	11
-13 & 19	0000 0011 111	11
-12 & 20	0000 0100 001	11
-11 & 21	0000 0100 011	11
-10 & 22	0000 0100 11	10
-9 & 23	0000 0101 01	10
-8 & 24	0000 0101 11	10
-7 & 25	0000 0111	8
-6 & 26	0000 1001	8
-5 & 27	0000 1011	8
-4 & 28	0000 111	7
-3 & 29	0001 1	5
-2 & 30	0011	4
-1	011	3
0	1	1
1	010	3
2 & -30	0010	4
3 & -29	0001 0	5
4 & -28	0000 110	7
5 & -27	0000 1010	8
6 & -26	0000 1000	8
7 & -25	0000 0110	8
8 & -24	0000 0101 10	10
9 & -23	0000 0101 00	10
10 & -22	0000 0100 10	10
11 & -21	0000 0100 010	11
12 & -20	0000 0100 000	11
13 & -19	0000 0011 110	11
14 & -18	0000 0011 100	11
15 & -17	0000 0011 010	11

Table 5.2: 動きベクトルの可変長符号化

がては目につく雑音になってしまう。したがって H.261 では演算の精度を規定し、予想される送受のミスマッチに対しては、一定の間隔 (132 回符号化する間に 1 度) でフレーム内符号化を行うことにより誤差の蓄積を防いでいる。

DCT 係数は図 5.5 に示すようにジグザグにスキャンされて量子化された後、連続したゼロの数 (ラン) とそれに続くゼロ以外の値 (レベル) の組み合わせとして符号化される。

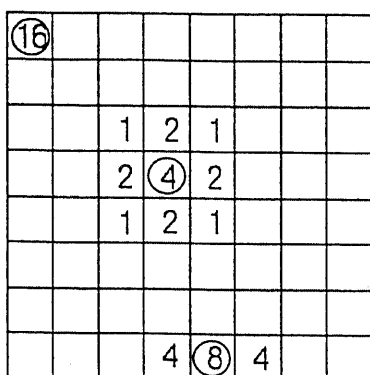
量子化は基本的に線形であり、フレーム内符号化の直流成分を除くすべての DCT 係数に同一の量子化器が適用される。符号化レートを制御するため、量子化器は 31 種類用意されている。

また、量子化代表値はすべて奇数となっている。これは、符号器と復号器のミスマッチを減少させるため、IDCT を行ったときに、 n (整数) + 0.5 となるような値を避けるためである。

量子化の結果、有意係数を持たないブロックは無効ブロックとなる。6 ブロックすべてが無効の MB の情報は伝送せず、次の有意な MB の先頭に 1 つ前の有効 MB からの距離を示すアドレスを付加して表現する。

4. ビデオマルチプレックス

情報源符号器で発生する各種の符号 (変換係数の量子化インデックス、動きベクトル情報、マクロブロックタイプ情報等) は、可変長符号化され、ビット量を削減された後、ビデ



数字は○の画素に対する重みを表す

図 5.4: ループ内フィルタの特性

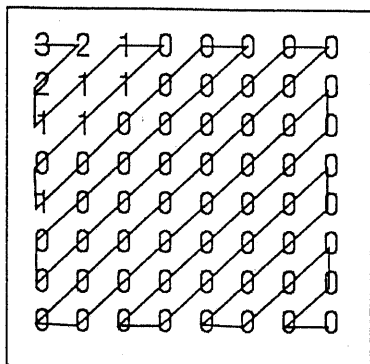


図 5.5: DCT 係数のジグザグスキャン

オマルチプレックスで 1 つのビット列に多重化される。これは、画像データの階層構造に対応した構造を持つ。

5. 伝送路符号化

符号化された画像信号は高度に圧縮されているため、伝送符号誤りがあるとその影響は広範囲に波及する。ここでは、BCH(511,493) 誤り訂正符号が用いられ、511 ビットブロック中の 2 個までのランダム誤り訂正あるいは 6 個までのバースト誤り訂正が可能である。

6. 符号化制御

ビデオマルチプレックスの出力データ量は、入力画像信号に含まれるシーンチェンジや動きの大小により変動する。したがって送信バッファをオーバーフローさせないように、発生情報量を制御する必要がある。制御方式には、発生した情報量と送信バッファの残留量から符号化パラメータを制御するフィードバック型と、入力画像信号の統計的性質を測定して符号化パラメータを制御するフィードフォワード型がある。また、発生情報量の制御には、量子化ステップサイズを変化させるものが一般的である。

5.3.2 H.261 を用いた多眼画像の符号化

5.3.1節で示した標準符号化方式 H.261 を用いて画像データの符号化実験を行った。実験の具体的な方法は次のように定めた。

- アルゴリズムは RM8(Reference Model 8)[1] の方式に従う
- ビットレートは約 6.4kbit/frame(H.261 の 64kbit/s, 10frame/s に相当) と約 3.2kbit/frame の場合について調べる
- 最初に 19 枚の画像の中央の画像 (10 番画像) をフレーム内符号化し, 両端 (1 番画像, 19 番画像) に向かってフレーム間符号化する
- 多眼画像の場合, 隣接した画像間では原理的に水平方向の視差のみが存在する。そこで動きベクトル (視差ベクトル) の探索範囲は
 1. H.261 の規格である ± 15 画素 (水平), ± 15 画素 (垂直)
 2. ± 7 画素 (水平), ± 7 画素 (垂直)
 3. 視差ベクトルの特徴を考慮した ± 7 画素 (水平), ± 1 画素 (垂直)

の 3通りとする。

実験に用いた画像は, カメラのレンズの光軸が基準点を通るように輻射撮影した後, キーストーン補正 (3章参照) + 位置合わせを施しているが, 厳密な位置合わせは困難であることから, \pm の両方向に対して視差ベクトル探索を行う必要がある。

前述のように, H.261 は動きベクトルの符号化において, 直前のマクロブロックの動きベクトル値との差分を, 可変長符号を用い符号化している。(2) と (3) で視差ベクトルの探索範囲を狭くした場合には, 表 5.3 で示すように, 可変長符号の折り返し位置を変更することによって符号語の割当を変え, 符号量を減少させることができる。

このようにして, 視差ベクトルの情報量が減少すれば, その分だけ量子化ステップサイズを細かくすることができ, SNR が向上すると考えられる。

5.3.3 符号化特性の評価

以上のように視差ベクトルの探索範囲を変えて符号化シミュレーションを行ったときの復号画像の一部を図 5.10 に, 視差ベクトルの平均符号量を表 5.4 に, 輝度値の SNR の平均を表 5.5 に示す。また, フレームごとの輝度値の SNR と視差ベクトルのビット数を図 5.6, 5.7, 5.8, 5.9 に示す。

ビットレート約 6.4kbit/frame では全体のビット数に占める視差ベクトルの割合が約 16~26% と少ないため, 探索範囲を狭くして視差ベクトルの符号量を減らしても, SNR にその効果はほとんど見られない。

一方, 約 3.2kbit/frame のときは視差ベクトルの占める割合が約 33~50% と大きいため, 符号量減少の効果が大きく, SNR は改善されている。

探索範囲 ±7			探索範囲 ±1		
視差ベクトル	符号語	符号長	視差ベクトル	符号語	符号長
-8 & 8	0000 0101 11	10	-2	0011	4
-7 & 9	0000 0111	8	-1	011	3
-6 & 10	0000 1001	8	0	1	1
-5 & 11	0000 1011	8	1	010	3
-4 & 12	0000 111	7	2	0010	4
-3 & 13	0001 1	5			
-2 & 14	0011	4			
-1	011	3			
0	1	1			
1	010	3			
2 & -14	0010	4			
3 & -13	0001 0	5			
4 & -12	0000 110	7			
5 & -11	0000 1010	8			
6 & -10	0000 1000	8			
7 & -9	0000 0110	8			

Table 5.3: 探索範囲変更後の視差ベクトルの可変長符号化

視差ベクトル 探索範囲	視差ベクトル平均符号量 (bit)	
	6.4kbit/frame	3.2kbit/frame
15 × 15	1679(26.54%)	1679(49.85%)
7 × 7	1339(21.20%)	1403(41.89%)
7 × 1	1027(16.26%)	1081(32.59%)

Table 5.4: 視差ベクトルの平均符号量 (全符号量に占める割合)

視差ベクトル 探索範囲	6.4kbit/frame		3.2kbit/frame	
	平均 SNR(dB)	最低 SNR(dB)	平均 SNR(dB)	最低 SNR(dB)
15 × 15	37.84	34.88	34.91	28.00
7 × 7	37.86	34.93	35.31	28.37
7 × 1	37.84	34.81	35.56	29.03

Table 5.5: 輝度値の平均 SNR と最低 SNR

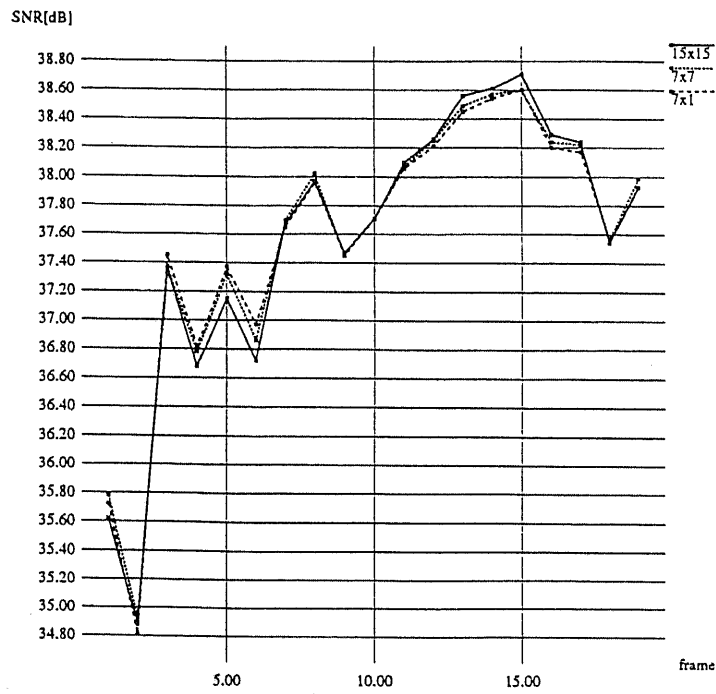


図 5.6: 輝度値の SNR(6.4kb/frame)

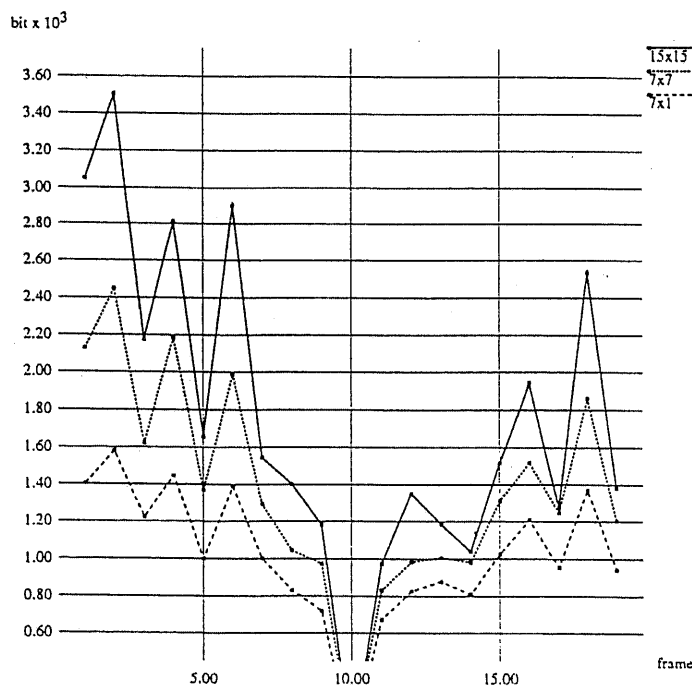


図 5.7: 視差ベクトルのビット数(6.4kb/frame)

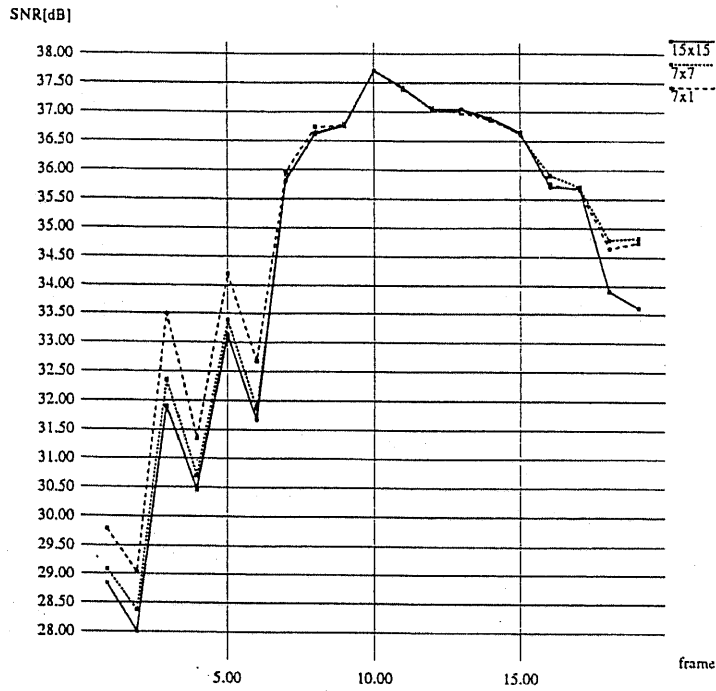


図 5.8: 輝度値の SNR(3.2kb/frame)

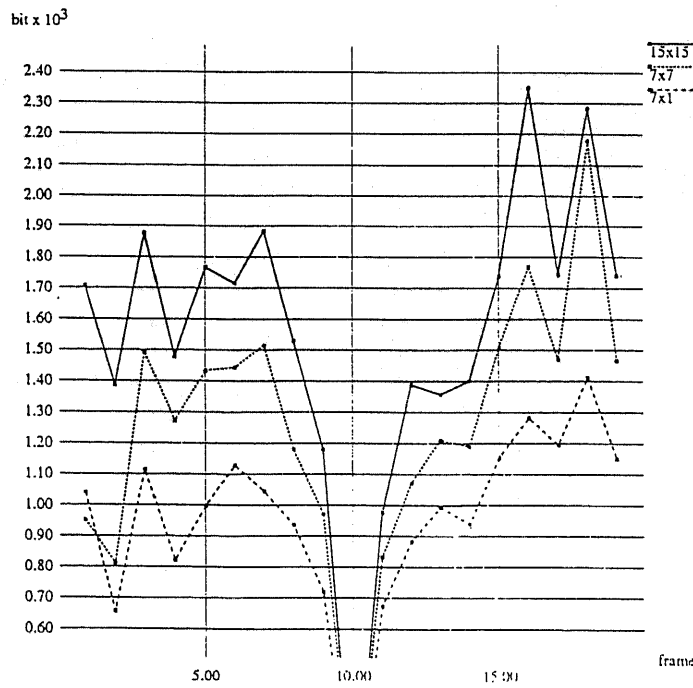


図 5.9: 視差ベクトルのビット数(3.2kb/frame)



7 × 1



15 × 15

6.4kbit/frame



7 × 1



15 × 15

3.2kbit/frame

図 5.10: 復号画像の一部 (18 番画像)

5.4 幾何学的制約を用いた視差補償(1)

多眼画像における隣接した画像は、視差が異なるだけで水平方向の相関は非常に高い。しかも、多眼画像は水平方向に一定の間隔をおいて撮影した視差画像の集まりであるから、すべての隣接した画像間には一定の視差が存在し、それらは幾何学的な制約により関係付けられる。この性質を積極的に利用して、より効率的な予測符号化ができないか、というのが本章の骨子である。

5.4.1 H.261 における視差ベクトルの内挿

本章では H.261 をもとにした符号化法の検討を行なう。具体的には、多眼画像を1フレーム飛ばして視差補償予測符号化し、その間の視差画像は視差ベクトルの内挿により予測しようという考え方である。

ここでは、以下のような符号化方法を考える。

1. 中央の5番画像(Iフレーム)を基準画像としてDCTでフレーム内符号化する。
2. 3番画像, 7番画像(Pフレーム)に対してH.261と同様の方法で視差ベクトルを検出して、フレーム間符号化を行う。
3. 5番画像から3番画像, 7番画像への視差ベクトルを等分して、4番画像, 6番画像(Dフレーム)を予測し、予測誤差を符号化する。
4. 3番画像, 7番画像から1番画像, 9番画像(Pフレーム)へフレーム間予測し視差ベクトルを求め、後は同様に2番画像, 8番画像(Dフレーム)を予測符号化する。

ここでは1枚飛ばして符号化する方法を考えたが、2, 3枚飛ばしても同様の方法で符号化が可能である。

この符号化法の利点としては、

- Dフレームの符号化の際、視差ベクトル情報を符号化しないので情報量が少なくてすむ。
- Dフレームの符号化の際、視差ベクトルの検出をしないので計算量が大幅に減少する。

などがあげられる。

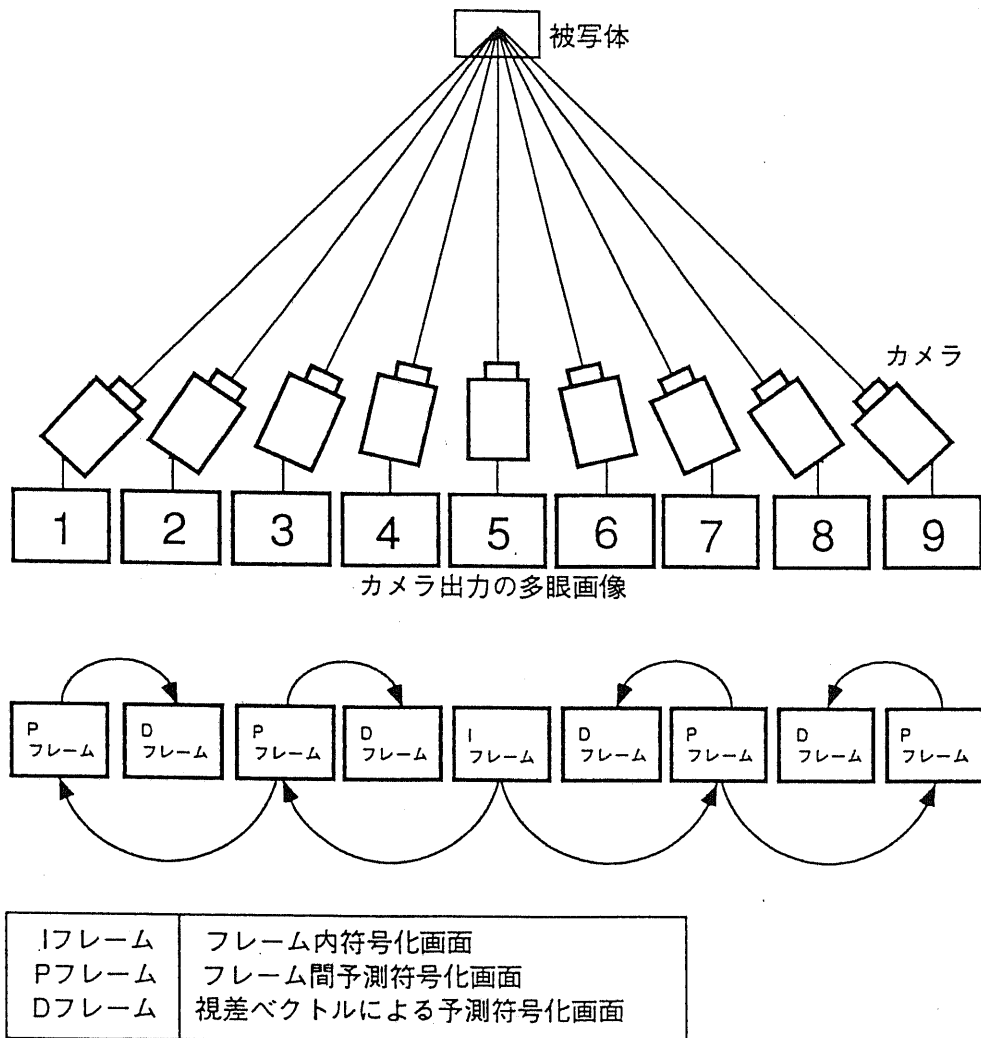


図 5.11: 視差補償予測符号化

以下で本方式の具体的な手法について述べる。

1. 符号化アルゴリズム

Iフレームのフレーム内符号化とPフレームのフレーム間符号化の基本構成はH.261(RM8)と同様とする。

しかし、多眼画像であるため水平方向の視差が大部分であること、フレーム飛ばしのため視差ベクトルが大きいことを考慮して、Pフレームの視差ベクトル探索範囲は、±15画素(水平)、±1画素(垂直)とする。そしてDフレームの予測画像には、H.261と同様のループ内フィルタを常にONとする。

2. Dフレーム画像の予測法

Pフレーム符号化の際に、16×16画素のマクロブロックのブロックマッチングで検出された視差ベクトル(dvx, dvy)は、Dフレームをつくる時1/2にして用いられる。その様子を図5.12に示す。

視差ベクトルの精度は整数画素であるため、垂直方向の視差ベクトルは1/2にすると0になる。よって、以下に述べる予測画像に発生する隙間の問題は、水平方向だけ考えればよい。

図5.12のように隣あう2個のマクロブロックの視差ベクトルが

$$\frac{dvx_n}{2} - \frac{dvx_{n+1}}{2} > 0$$

のような場合には、予測画像に隙間ができてしまう。

そこで

$$\frac{dvx_n}{2} - \frac{dvx_{n+1}}{2} = a > 0$$

なら、n+1番のマクロブロックを左側にa画素広げ、隙間をつくらないようにする。また、マクロブロックが重なる場合は、右側のマクロブロックの画素値を利用する。

いずれの場合も予測に用いるマクロブロックの画素値には、Dフレームの両隣のフレームの、対象マクロブロックの平均値を用いる。

3. 情報量の制御

H.261は低ビットレートの通信用符号化方式であるため、あらかじめ固定した伝送レートに合わせて、フレームごとにビット数が一定になるような制御を行っている。

しかし今回の実験では、静止画の多眼画像を扱っているため、フレームごとのビット数よりも、全フレームのビット数の和の方が重要な意味を持つ。

つまりH.261のようなバッファ制御ではなく、本方式に適した符号量制御を行う必要があるということである。

本符号化方式の場合、Dフレームの符号化では視差ベクトル情報を符号化する必要がないため、その分だけデータ量が少なくなる。また、Dフレームの予測画像はPフレームをもとにして作られるのであるから、Pフレームの品質をできるだけ高くする必要がある。

そこで、Dフレームの符号化に割り当てるビット数を少なく、Pフレームの符号化に割り当てるビット数を多くしてやる。そうすれば、総ビット数を同じにした場合、H.261の方式より画質は改善されるはずである。

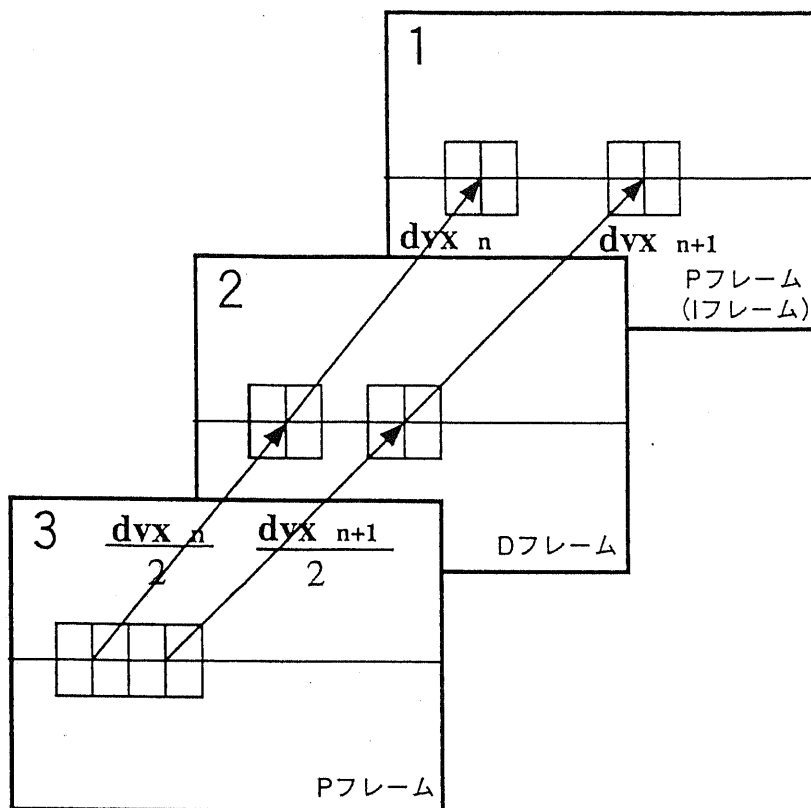


図 5.12: D フレーム画像の予測法

5.4.2 視差内挿予測符号化シミュレーション

視差補償予測符号化方式を用いて、画像データの符号化実験を行った。実験の流れは次のようになる。

1. 10 番画像を I フレームとしてフレーム内符号化する。
2. 12 番画像を P フレームとして符号化する。
3. 11 番画像を D フレームとして符号化する。
4. 14 番画像を P フレームとして符号化する。
5. 13 番画像を D フレームとして符号化する。
6. 後は交互に P フレーム、D フレームとして 18 番画像まで符号化する。

比較のため H.261 を持ちいて、探索範囲 ± 15 (水平), ± 1 (垂直) の場合の実験を行なった、また、I フレームの量子化ステップサイズは 32 とし、P フレーム、D フレームの量子化ステップサイズは、P フレーム、D フレームのビット配分が P フレーム重視で、しかも 1 フレームあたりの平均ビット数が約 6.4 kbit になるように定める。

実際には、P フレーム、D フレーム共にステップサイズ 21 のときが最もよい SNR の値を示したので、この値を用いることにする。

以上のように実験を行ったときの復号画像の一部を図 5.15 に、P フレームと D フレームの平均ビット数を表 5.6 に、H.261 との輝度値の SNR の比較を表 5.7 に示す。また、フレームごとの輝度値の SNR を図 5.13 に、フレームごとの全ビット数を図 5.14 に示す。

	P フレーム	D フレーム	平均
ビット数 (bit)	8832	3893	6362

Table 5.6: P フレームと D フレームの平均ビット数

	H.261	本方式
平均 SNR (dB)	38.22	38.52
最低 SNR (dB)	37.49	38.19

Table 5.7: SNR の比較

平均 SNR は H.261 方式で得られた 38.22dB から 38.52dB に 0.3dB 改善され、最低 SNR は 0.7dB 改善された。

以上の結果より、フレームごとに符号化のビット配分を変更することによって、視差ベクトル情報量の減少を利用することができたといえる。

より効率の良い符号化のためには、以下のような点を改善すべきであると考えられる。例えば、本方式では D フレームの符号化で予測画像を作るときに、視差ベクトルの精度を整数画素単位としている。予測の際、P フレームの符号化で求められる視差ベクトルを 1/2 にして用いていることを考えると、半画素単位での予測が必要であり、その結果、予測精度の改善が期待できる。また、視差ベクトル探索時のブロックマッチングに用いるブロックを小さくしたり、3 枚以上の多眼画像を用いて視差ベクトルを検索することによって、より正確な視差ベクトルが求められれば、補間するフレームを増やすなどして、さらに情報量を削減することが可能であると考えられる。

また別の見方をすれば、この結果は、ブロックごとに視差ベクトルを求めるという、従来の波形符号化の方式の限界とみることもできる。

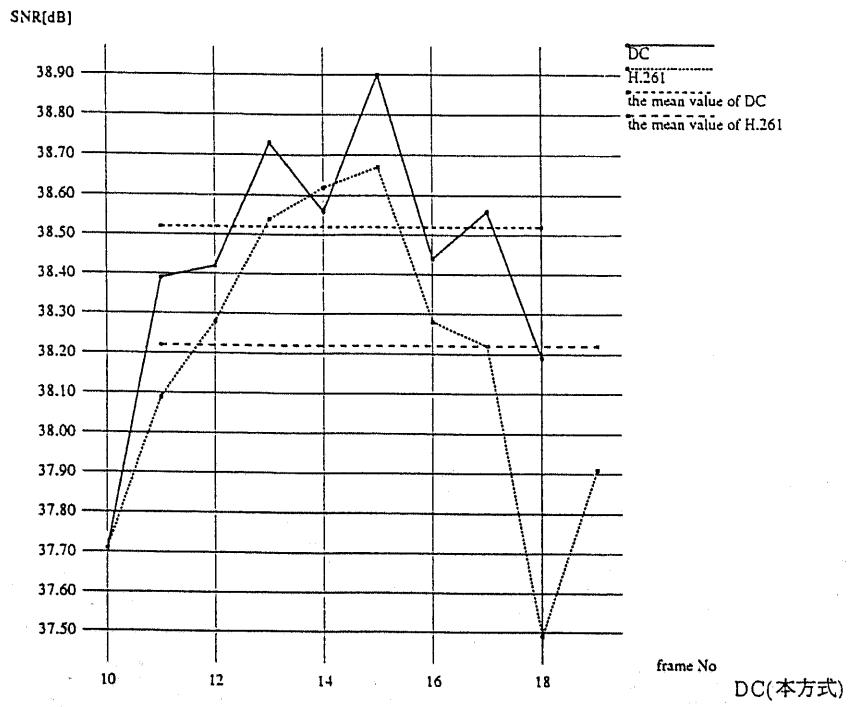


図 5.13: 輝度値の SNR

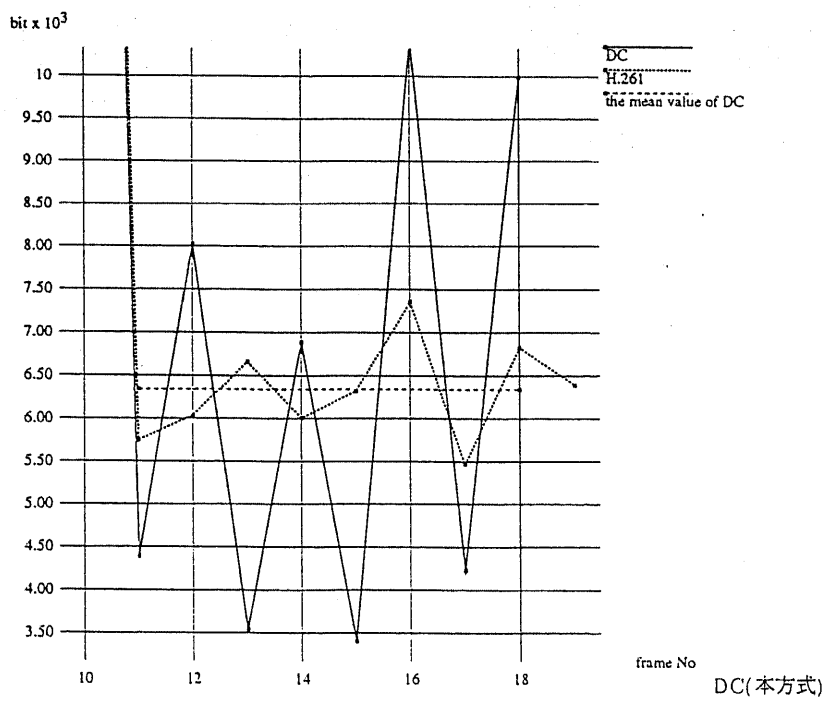


図 5.14: 総ビット数



H.261



本方式

図 5.15: 復号画像の一部 (17 番画像)

5.5 幾何学的制約を用いた視差補償 (2)

前章では、数枚の視差画像の間にある視差ベクトルの幾何学的関係を用いた符号化法について検討した。この考え方は動画像においてもあまり動きが激しくない画像に対しては有効な手法であり、実際、蓄積用動画像標準符号化方式である MPEG1 においては双方向予測として実現されている。

さて、多眼画像の場合には、数枚だけではなく全ての視差画像に渡って視差ベクトルの幾何学的な関係が存在する。本章ではこの幾何学的関係を考慮し、多眼画像の効果的な圧縮方式を検討する。

5.5.1 mn 画像の導入と多眼画像間の幾何学的関係

各フレームの水平 1 ラインに注目し (y を固定), (X, Z) 水平面内の二次元座標系を考える。図 5.16 の左に被写体空間の水平断面 (X, Z) 面を示す。

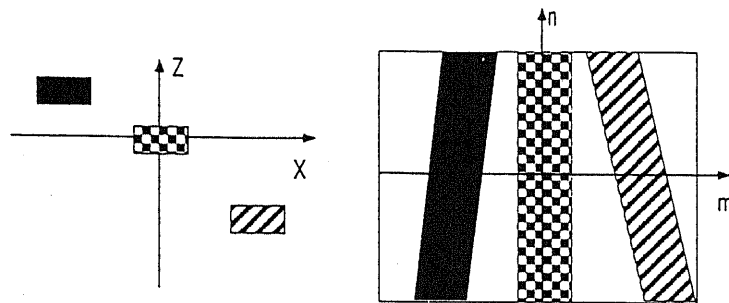


図 5.16: レンズ板によって再生される空間と (m, n) 平面

平行投影を仮定し、指標 n で表される視点からみた画像を $I_n(m)$ とする。すなわち、画素に対応する指標を X 座標と同じ値を持つ整数指標 m で表し、光線の方向 X/Z を指標 n で表すことにする。光線は n が整数である方向に量子化されているものとする。このような空間は、「規格化空間」と呼ばれており、実際の 3 次元像が表示される空間に容易に変換することができる。一組のパラメータ (m, n) により定められる光線を「特徴光線」と呼ぶ。特徴光線 (m, n) とは、 $X = m, Z = 0$ の位置を通り、 n の方向への光線である。3 次元像はこのような特徴光線群によって再生される。

この特徴光線 (m, n) は、次のように考えることもできる。

各フレームから同じ y 座標値の水平 1 ラインを取り出して並べる、すなわち、多眼画像を並べて水平断面をとると、 (m, n) 平面が出来上がる。この (m, n) 平面を mn 画像と呼ぶことにする。図 5.17 に mn 画像を示す。

具体的に多眼画像の幾何学的な関係を考えてみる。

図 5.18 より次のことが明らかである。基準点より奥にある被写体は、多眼画像の中心画像から右に 1 枚目の画像にかけて、ブロックが水平右方向に動く。同様に考えると、基準点より手前にあるものは、逆に水平左方向に動くはずである。また基準点からの奥行きは多眼画像の幾何学的関係により、ブロックの水平方向への移動量に比例するはずなので、2 枚目の画像では先ほどの 1 枚目の画像の 2 倍だけ必ず平行移動しているはずである。

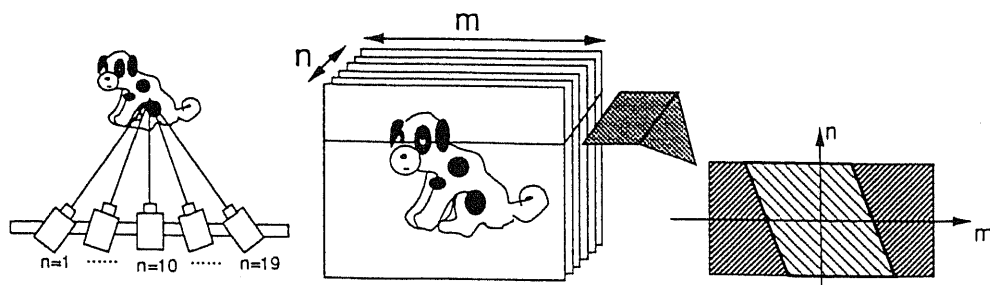


図 5.17: mn 画像

以上により, 多眼画像の幾何学的な関係から, 規格化空間内の点 (X, Z) を通る特徴光線群 (m, n) は, 次の式を満たす.

$$m = X + nZ \quad (5.1)$$

一点 (X, Z) は, 式 5.1 に従って複数の mn 画像上の点に記録される. これを利用すれば, 式 5.1 の関係を用いて, 多眼画像の視差補償を行なうことができる.

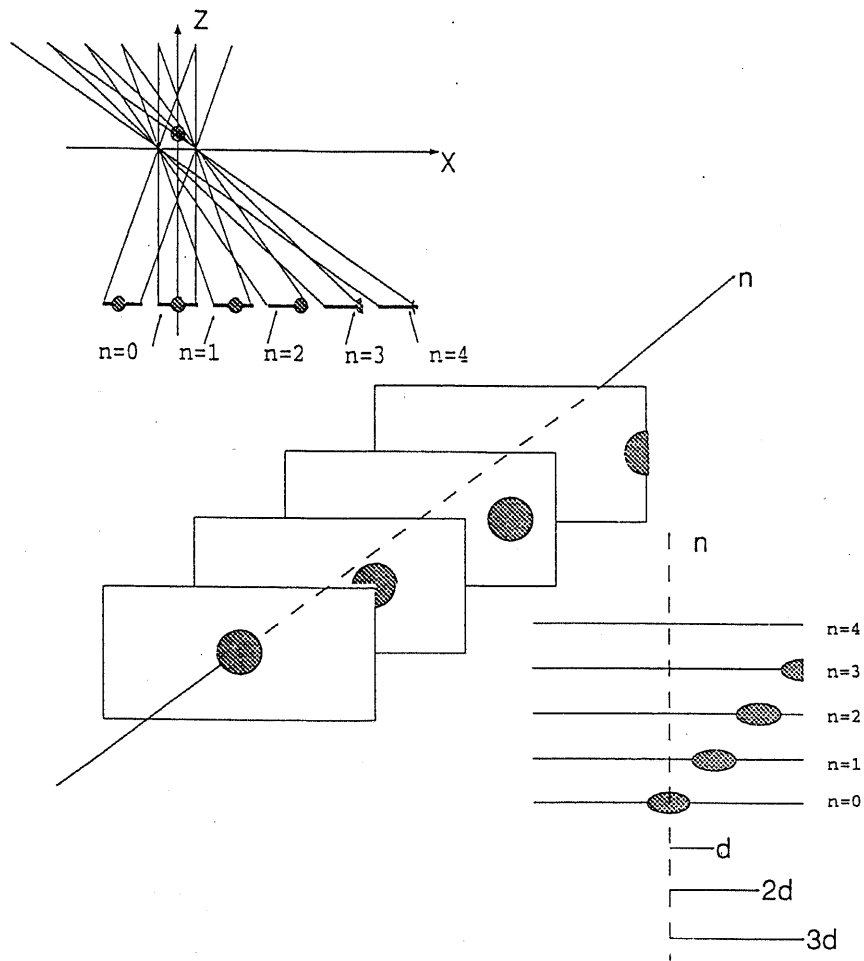


図 5.18: 奥行きと水平方向移動量

5.5.2 mn 画像からの視差ベクトルの抽出

被写体空間内の点は、mn 画像内では基準面からの距離に比例した傾きの直線にそって多数のフレームに記録されている。このためデータの冗長性が生じている。この冗長性を除くため、mn 画像のテクスチャを構成する直線群に着目する。この直線群の傾きと、この直線に沿って各視差像に記録されている輝度データの平均値をもって代表値とすることにより、データの圧縮が可能である。また、この圧縮過程はある物体の規格化空間座標における位置 (X, Z) 、及びその物体のテクスチャの輝度データを抽出していることに等しい。

19 枚の画像の中心画像を基準に考える。中心画像から左右に1枚目の画像に対しては、中心画像中のあるブロックからの移動量が水平方向に d であるブロックを、左右に2枚目の画像には今度は(中心画像の元のブロックから)水平方向に $2d$ 移動した位置にあるブロックを、同様に左右に3枚目では $3d$ 、4枚目には $4d$ 、9枚目には $9d$ だけ移動した位置でそれぞれのブロックを対応するブロックとし、対応するブロック間で、各ピクセル単位で分散をもとめる。そして、最適な水平方向の移動量を求める評価関数は、これらの分散の和が最小になるときとする。次に、また(中心画像内の)別のブロックにおいて同様に最適水平方向移動量 d' を求める。そしてすべてのブロックに対して適当な水平移動量を求める。

視差ベクトル抽出実験の具体的な方法は次のように行なった。

- ブロックサイズは8ピクセルとする。
- 視差ベクトル(d)の細かさは、サブピクセルまで用い、隣のフレームに対しては1/3ピクセル、中心画像(10番画像)から端の画像(1,19番画像)だと3ピクセルの細かさとする。
- 動きベクトルの検索範囲は、隣のフレームに対して ± 6 画素。
- 中心画像を基準として考えた時カメラの視野から外れる部分において復号画像に穴が空くのを防ぐために M_{max} に加えて両襟につけるブロックは、左右に48ピクセルずつとする。

上記「カメラの視野から外れる部分」に対しては特別な処理が必要となる。それについて以下に述べる。

カメラの視野から外れる部分の処理

本来 mn 画像の m は、 (X, Z) 平面における X と同じであるので、犬の画像の場合 256 ピクセルでよい。しかし例えば1フレームの右や19フレームの左など、中心画像を基準として考えた時、画角から外れる部分は中心画像に写っておらず、復号画像に穴が空く。これを防ぐために以下の処理を行なう。

例えば図5.19の A は中心画像では見えていないが、中心画像以降で見えてくる領域である。すなわち、領域 A の中心画像での対応領域は、図5.19の右図に見られるように中心画像に写っている部分より左に位置する領域である。従って $n=10$ を基準と考えた背景画像は、 M_{max} では不足し、 M_{max} に加えて両襟に余分にブロックをつけなくてはならない。この考えに基づき、 M_{max} 及び、その両襟のブロックの輝度と傾きを伝送する。

視差ベクトル抽出結果

いくつかの mn 原画像および、5.5.1節で示した方法により復号された mn 画像を以下に示す。なお、mn 画像は m が 256 ピクセルに対し n が 19 ピクセルであるので、縦に比べて横に長く

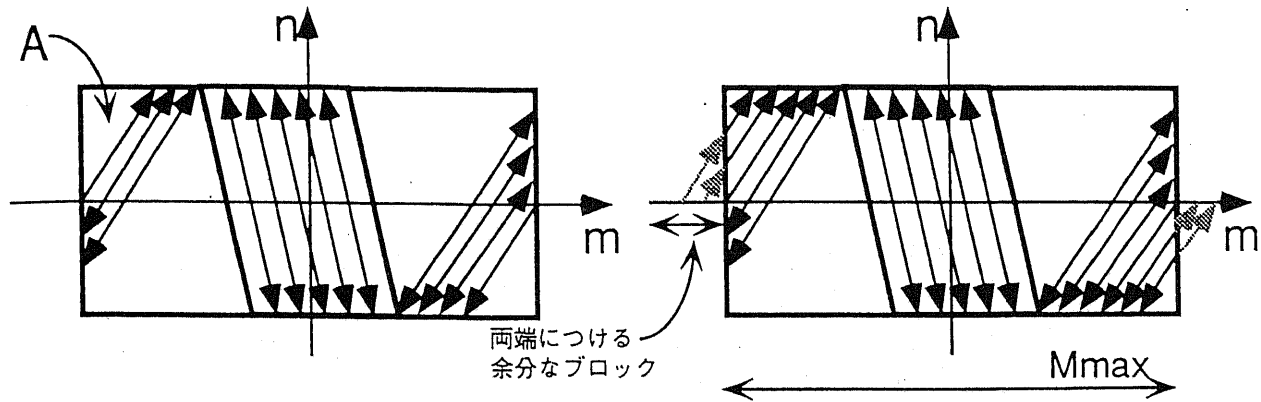


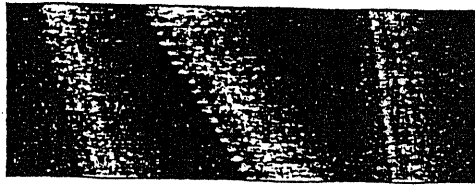
図 5.19: 画角から外れる部分の処理

見にくいので便宜上 n 方向 (縦方向) に 5 倍に拡大している。

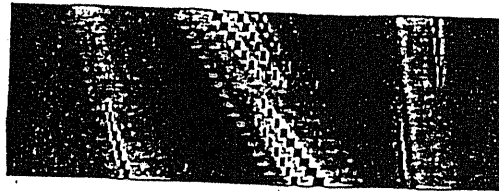
以上の方法により圧縮データは 8 ピクセル単位の視差ベクトルと 1 枚分のテクスチャ情報だけとなり、多眼画像を効率良く圧縮することができた。その際の符号量は初めの 1 枚 ($256 \times 256 \times 3 \times 8(\text{bit})$) を送った後は 67584bit の視差ベクトル情報で残りの 18 枚を復号できることになる。

しかしながら、本手法では中心画像に写っていない部分は、復号することはできない。従って中心以外の方向からの復号画像に穴が空いてしまう。以下に穴が生じる原因として考えられる理由をあげる。

- 物体など比較的奥行きの変化がある物体など視点位置により見かけの大きさが変化するもの ($Z=0$ の平面に平行でない平面など) に対しては穴が空いてしまう。
- 物体と背景の境界など、奥行きが急激に変化するところではオクルージョン (手前の物体により、背後の物体が隠されること) が発生してしまい、直線的に視差ベクトル抽出をすることができず、犬のように物体に比して背景が単調な画像の場合だと物体と背景の境界の背景側に穴が空いてしまい、予測特性が低下してしまう。



原画像

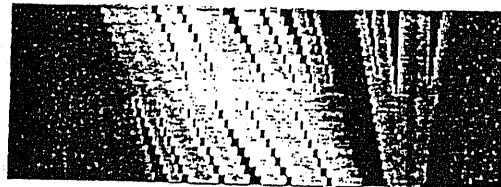


復号画像

$y = 50$

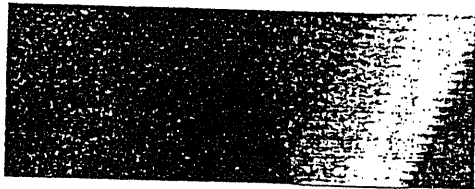


原画像

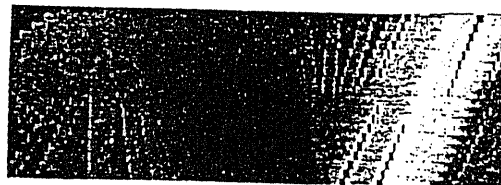


復号画像

$y = 100$

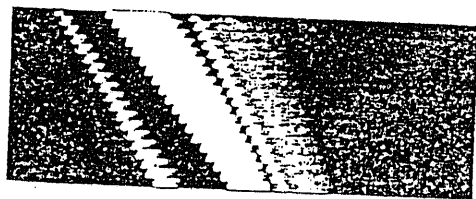


原画像

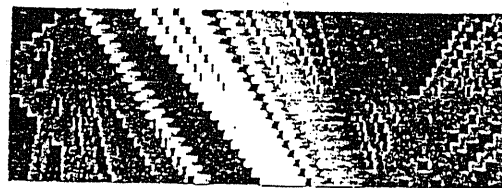


復号画像

$y = 150$



原画像



復号画像

$y = 200$

5.5.3 アフィン変換に基づく視差ベクトルの補間

視点位置により見かけの大きさが変化するもの ($Z=0$ の平面に平行でない平面など) に対しては中心画像の復号画像以外には穴が空いてしまうことが実験から明らかになった。この節では、視点位置により見かけの大きさが変化する物体により、復号画像に穴が空いてしまう理由を定性的に考えてみる。

前節の方式では、8ピクセル単位で視差ベクトルを抽出した後、受信側でそのまま8ピクセルとも同じ視差ベクトル(傾き)で復号画像に書き込んでいる。すなわち図5.20のように第1フレームから第19フレームまで大きさが変わることがなく、同じ8ピクセルとして書き込んでいるのである。

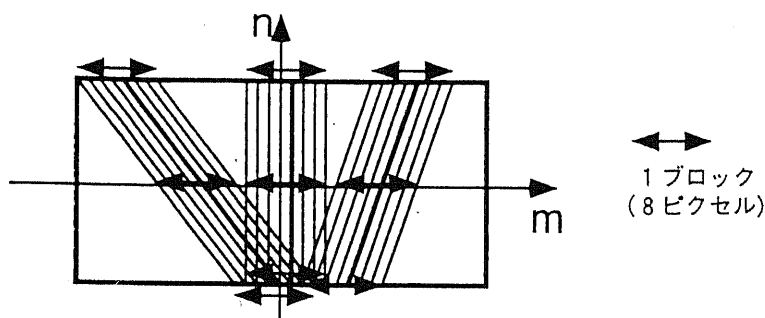


図 5.20: 前節の方式による復号画像

8ピクセルごとに同じ傾きのまま復号画像を作るということはすなわち、あるブロックに含まれる8ピクセルに対応する空間上の物体はすべて同じ奥行きを持っているとみなすことになる。それを図に表すと図5.21の左図のようになる。

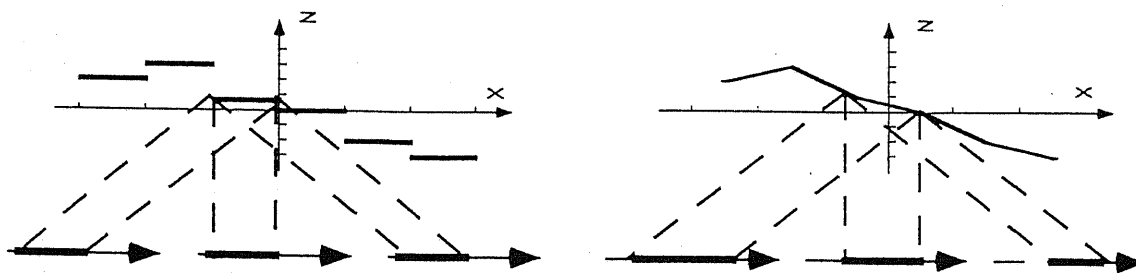


図 5.21: 傾きの内挿

従って中央画像以外には穴が空くのは図5.21の左図から明らかである。

ここでは前述の復号画像の穴を埋めるためにアフィン変換を導入し、視差ベクトルを内挿することを試みる。視差ベクトルを内挿する手順は以下の通りである。

- 8ピクセルのうち中央のピクセルの視差ベクトル(傾き)を、前節の方式によってそれぞれのブロックごとに求められた傾きとする。
- 8ピクセルのうち中央より左のピクセルの視差ベクトル(傾き)は、そのブロックより一つ左のブロックの視差ベクトル(傾き)との間で傾きを内挿する。

- 同様に8ピクセルのうち中央より右のピクセルの視差ベクトル(傾き)は, そのブロックより一つ右のブロックの視差ベクトル(傾き)との間で傾きを内挿する.

これを図に表すと図5.22のようになる.

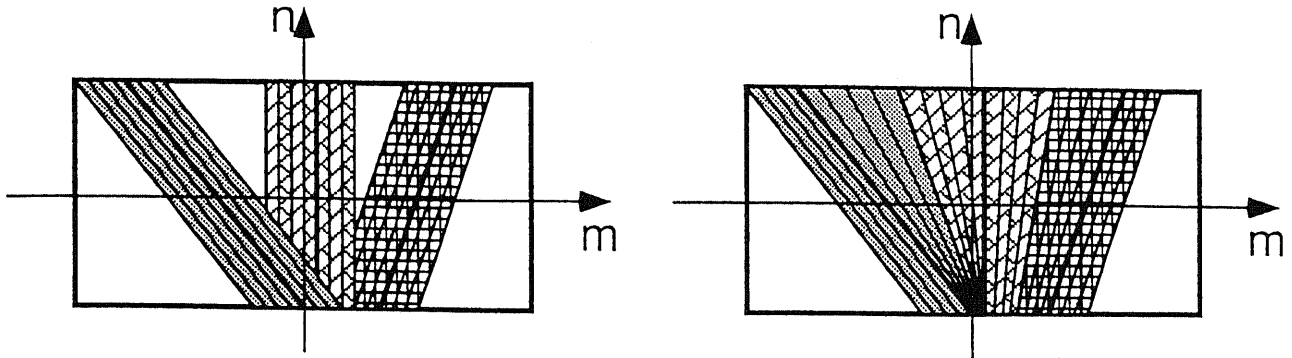


図 5.22: アフィン変換による傾きの内挿

以上のような方法で, 隣あったブロック間で視差ベクトルを内挿することはすなわち, 各ブロックの中央の奥行きを, 前節の方式によって求められた奥行きに固定し, それ以外のブロック内のピクセルの奥行きを隣のブロックの奥行きとの間で内挿することである. このことにより, 急激に奥行きが変わらないようにするのである. それを図に表すと図5.21の右のようになる. 隣あったブロック間で傾きを内挿することにより中央画像以外には穴が空かなくなるのは図5.21の右図から明らかである.

色の内挿

傾きの内挿により隙間を埋めることが必要とされる領域は, (X, Z) 平面において Z 方向の変化があり, 直線的に視差補償したのではあまり都合の良い領域である. 従って, Z 方向の変化の少ない領域に比べてブロックマッチングした結果の分散は比較的大きくなっているはずである. そのような色を補間せずにそのまま隣のピクセルまでベタ塗りしたのでは, SNR の値が下がってしまう.

この考えから, 傾きを内挿した後の隙間を, 隣のピクセルの色との間で内挿する.

今 mn 画像中の8ピクセルからなるあるブロックを考える. 求められた1ブロック(8ピクセル)の色を左から順に $(r_0, g_0, b_0), (r_1, g_1, b_1) \dots, (r_{\bar{1}}, g_{\bar{1}}, b_{\bar{1}})$ とする. また, ある $n (= N)$ においてその色が対応すると認識された点を順に $m_0, m_1 \dots, m_{\bar{1}}$ とする. この時, mn 画像上の点の色 $(C_r(m, n), C_g(m, n), C_b(m, n))$ ($m_1 \leq m < m_2, n = N$) は式5.2のように表せる.

$$\begin{pmatrix} C_r(m, n) \\ C_g(m, n) \\ C_b(m, n) \end{pmatrix} = \begin{pmatrix} r_1 + \frac{r_{\bar{1}} - r_1}{m_2 - m_1} (m - m_1) \\ g_1 + \frac{g_{\bar{1}} - g_1}{m_2 - m_1} (m - m_1) \\ b_1 + \frac{b_{\bar{1}} - b_1}{m_2 - m_1} (m - m_1) \end{pmatrix} \quad (5.2)$$

視差ベクトルの補間実験

視差ベクトルを補間することにより, 5.5.2節において復号画像に穴が空いていたのを埋めた. アルゴリズムは以下の通りである.

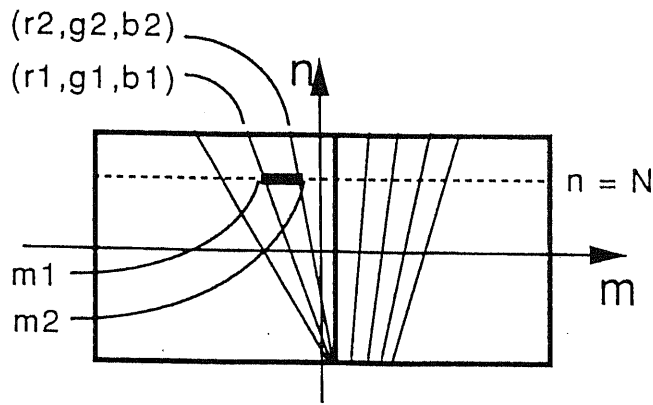


図 5.23: 色の内挿

- エンコーダは前節と全く同じものを用い、ブロックサイズ、サブピクセルの細かさ、動きベクトルの探索範囲、両襟につけるブロックの大きさなどは先の方式に従う。
- デコーダは前節の方式に5.5.3節で示した視差ベクトルを内挿するアルゴリズム及び5.5.3節で示した色を内挿するアルゴリズムを加えたものとする。
- 8ピクセルのうち、奥行きを変化させない中心ピクセルは左から4番目とする。

これにより前章と同じ符号量で、かつSNRは平均33.4dBとなった。画像のうち犬の部分ではディスプレイ上で見る限りあまり差はわからない程度になった。

図5.24にいくつかの復号mn画像を示す。また、フレームごとのSNRを図5.25に示す。

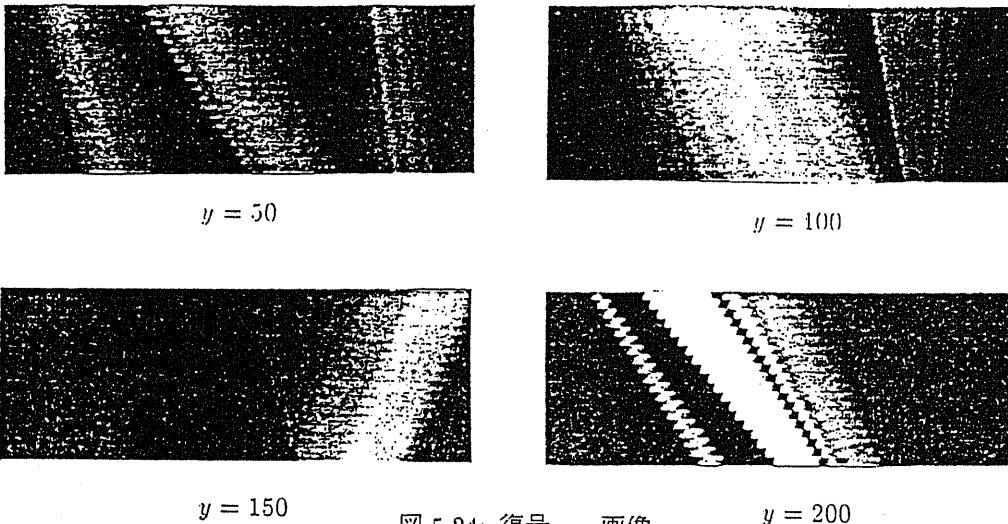


図 5.24: 復号 mn 画像

しかしながら、オクルージョンが激しい領域では予測特性が低下するという問題がある。とりわけ物体と背景の境界付近では、視差ベクトルが正しく抽出できていないだけに、傾きの内挿をしても復号画像はあまり綺麗にはならない。ただ、この犬の画像に限って言えば、背景が比較的単調なので、間違っても求められた(背景部分の)視差ベクトルを内挿してもさほどSNRの値は落ちていないようである。

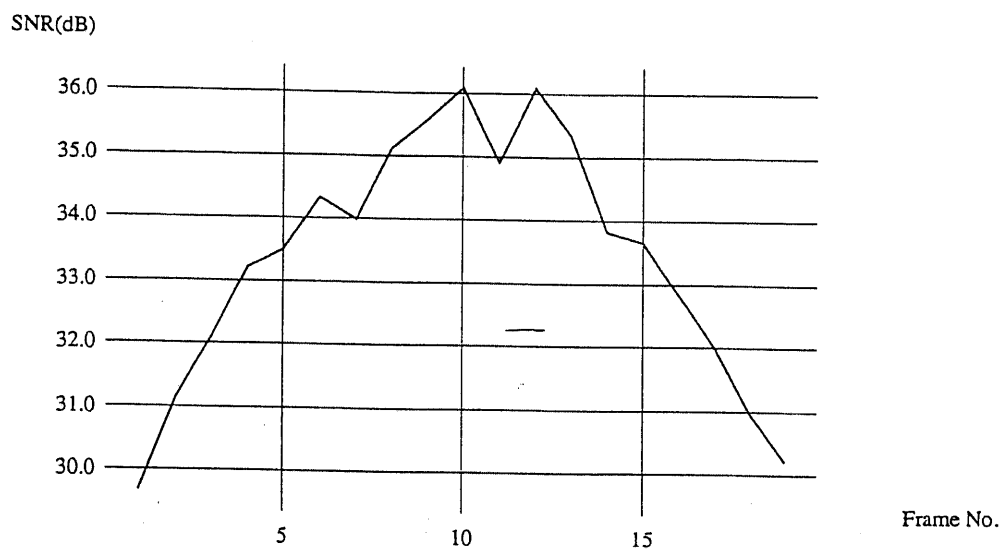


図 5.25: 復号画像の SNR

5.5.4 非撮影方向からの画像の合成, 補間

多眼画像においては, その補間 (非撮影方向から見た像の合成) 手法の確立が, 情報圧縮及び撮像系の簡略化のために重要な技術課題と言える。

今回の実験で用いた画像は従来の二次元カメラを用いて, 多眼カメラ入力を行なったものだが, 一般に入力系が複雑になるのを避けるためには, 入力側のカメラ数をあまり大きくとるべきでない。一方, 自然な立体視を得るためには3次元ディスプレイの眼数がある程度大きくする必要がある。このような相反する要求から, 少ない眼数を出力側に合わせるために, 非撮影方向から見た像を合成, 補間する必要が出てくると考えられる。また, 受信した情報に合わせて出力側で補間していたのでは, 情報量削減の意味からも効率的ではない。

本節では, 以上のような問題点を踏まえ, 前節及び5.5.3節で述べた方式と同様の方式で, 特別な操作を必要とせず自由に画像を合成・補間できることを示す。この目的が達成されれば, 送信側で得られた情報から, 各受信側でそれぞれのディスプレイにあわせた眼数の画像を自由に合成できるようになる。また, 入力側のカメラの数を減らし, 入力系の簡略化にも貢献し得る。

非撮影方向からの画像の合成

偶数フレームから奇数フレームを補間した際の mn 画像を図 5.26 に, 1,4,7,10,13,16,19 番画像から他を補間した際の mn 画像を図 5.27 に示す。また, フレームごとの SNR を図 5.28 に示す。

- エンコーダ, デコーダともに視差ベクトルの補間実験と同じものを用いる。
- 次の2通りの内挿実験を行なった。
 - 1~19 フレームのうち偶数フレームから奇数フレームを内挿 (ただし, 第1,19 フレームは外挿となる) した場合。
 - 1,4,7,10,13,16,19 番フレームからその他を内挿した場合。

非撮影方向からの画像の合成実験に関しては, SNR から見る限り次のことがいえる。

- 偶数画像から奇数画像を合成

まず, 従来の補間を用いない場合と, 偶数画像から奇数画像を補間した場合を比べてみると, 補間された奇数フレームではもちろん補間していない画像の方が SNR の値は良くなる。しかしながら 3, 5, 7, ..., 15, 17 番画像についていえば補間を用いない場合とそれほど大差はない。これは, これらのフレームが内挿により合成されたためで, 今回のように1フレーム飛ばしの内挿程度であれば全く問題がないといえる。しかし, 1, 19 フレームに関していえば, 外挿により合成されているので, SNR は極めて悪い。外挿されたフレームは他のフレームで見られない部分が見えているので, 視差ベクトルが正しく抽出されていないのである。

また, 偶数フレームでは, 偶数画像から奇数画像を補間したものの方が一般的に SNR が良くなっている。これは標本となる元の画像数が少なくなったので, 標本としたフレームの SNR の値が従来の補間しない方式より良くなったのだと思われる。
- 1,4...,16,19 番画像から他を合成

理論的には標本となる元の画像数を減らせば減らすほど, 標本となったフレームの SNR

は上がり、補間されたフレームの SNR は下がるはずなのであるが、フレーム数が少なくなると正しい視差ベクトルを求めることが困難になる。そのため、標本となるフレーム数を減らしたにもかかわらず、全般的に SNR が下がってしまった。

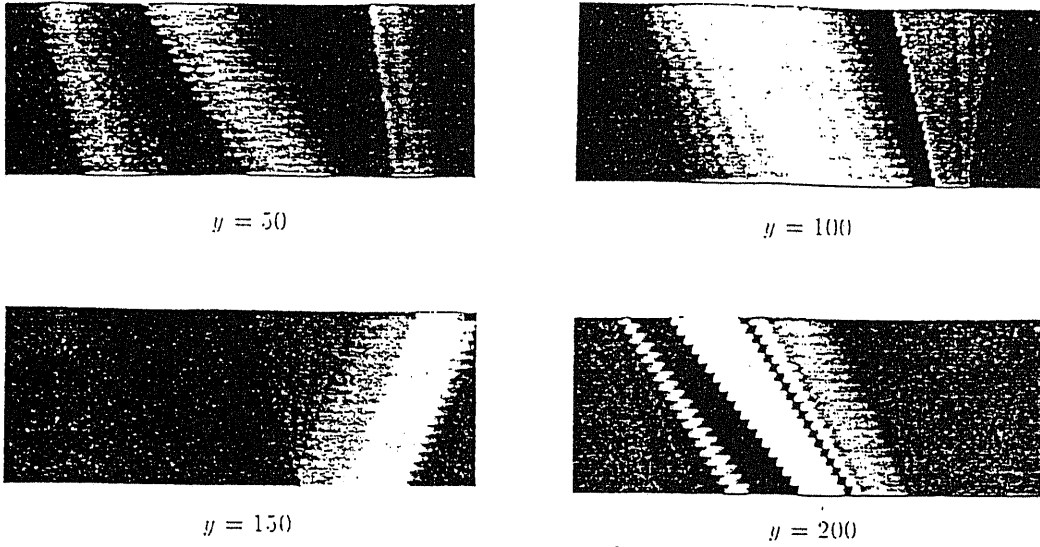


図 5.26: 偶数フレームから奇数フレームを合成した際の mn 画像

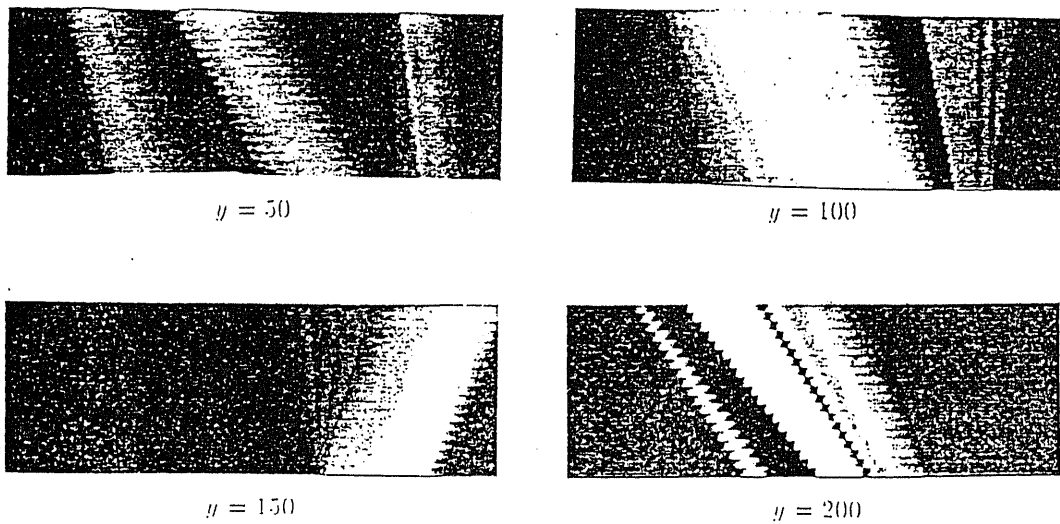


図 5.27: 1,4,7,10,13,16,19 番画像から他を合成した際の mn 画像

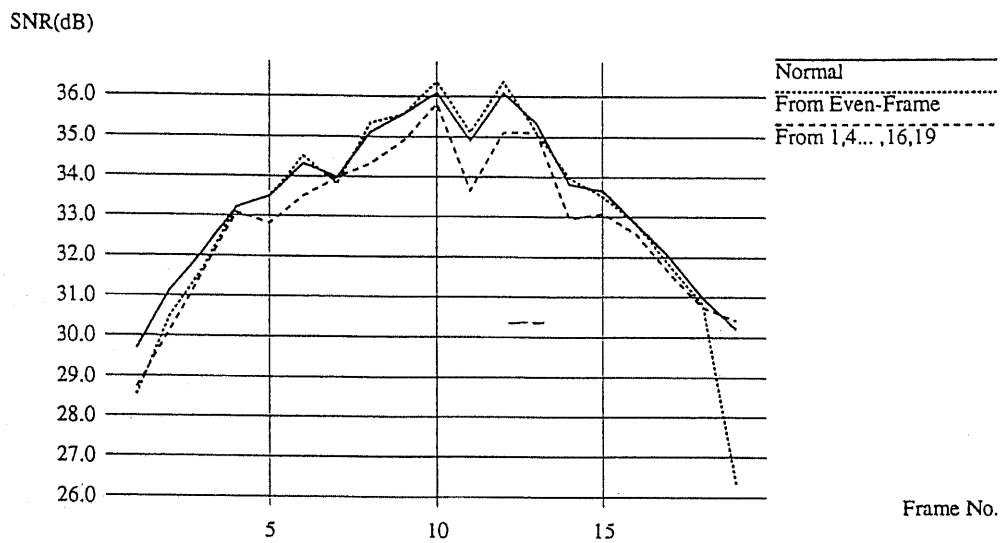


図 5.28: 補間画像の SNR

5.5.5 背景と物体の分離に基づく予測特性の改善

前節までの実験では、SNRの値で見ると比較的高い画質を達成しているものの、アフィン変換を用いて視差ベクトルを内挿している関係上、オクルージョンが激しいシーンに対しては予測特性の劣化が予測される。

全ての視差画像に渡る視差ベクトルを求めているので、オクルージョンが生じた画像に対しては正しい視差ベクトルが求められることは期待されない。正しい視差ベクトルを求めなければ、犬のように背景のテクスチャが比較的単調な画像ならば良いが、背景のテクスチャが複雑な画像では、たとえ視差ベクトルの補間により復号画像の隙間を埋めたとしても、SNRの向上は望めない。

これらの問題を検討するため、オクルージョンの激しい多眼画像を用いて背景と物体の分離に基づく予測特性の改善を試みた。

オクルージョンの激しい多眼画像の例として、背景のテクスチャの複雑な動画標準画像である「フラワーガーデン」を多眼画像としてみなせるように補正した。オクルージョンが生じている場合について検討を行なうため、具体的に用いたフレームはflower525～flower545の21枚の画像である。この範囲の画像は中心に木があり、オクルージョンの問題を考えるのには都合が良い。

mn 画像上での背景と物体の分離方式

物体と背景の差は何であろうか。(X, Z) 平面で考えると、それは明らかに奥行きである。すなわち基準点(0, 0)からの距離Zである。5.5.1節での考察により、視差ベクトルの傾きは基準点からの奥行きZに比例するので、mn 画像中のブロックの傾きに閾値を設けて、傾きが閾値より小さい(右上がり)場合は背景で、閾値より大きい(右下がり)場合は物体であると判断することにより物体と背景の分離ができるものと思われる。

しかし、前述の視差ベクトル抽出法では、図5.29のように、mn 画像において物体近辺の背景は物体の傾きにひきずられ物体と同じ程度の視差ベクトルを算出してしまふ。これでは物体のそばの背景が物体と認識されてしまふ、その部分のためにSNRの値が落ちてしまふ。

mn 画像を構成する要素はすべて直線であることに注目すると、物体と背景のエッジもまた、直線になっているはずである。この仮定のもとで以下の分離方式を提案する。

今、物体と背景のエッジの背景側の1ブロック(図5.29左図)に注目する。先ほど述べたように、このブロックにおいて普通に1～19フレームを用いてブロックマッチングを行ない視差ベクトルを求めた場合、図5.29の左図に示すように物体にひきずられて物体のエッジとほぼ同じ傾きを算出してしまふ。

しかしながら、このブロックの本来の視差ベクトルの方向は図5.29の右図に示す方向である。この際、前述の通りに1～19フレームを用いて直線的にブロックマッチングをしていたのでは、この直線に沿った分散が大きくなってしまふこの方向は算出されない。これを避けるため、従来の1～19フレームを用いてブロックマッチングをする方式に加え、1～10フレーム及び10～19フレームを用いてブロックマッチングをする方式を加えることにする。この方式を適用すると、例えば5.29の右図の左のブロックの傾きは1～10フレームでブロックマッチングをすることにより求められ、右のブロックは10～19フレームでブロックマッチングをすることにより求めることが可能になる。

そして、1～19, 1～10, 10～19フレームの中から、もっとも分散が少ないものをそのブロックの正しい視差ベクトルとみなし、先ほど述べたように、そのブロックの傾きが閾値より小さい(右上がり)場合は背景で、閾値より大きい(右下がり)場合は物体であると判断する。また、閾

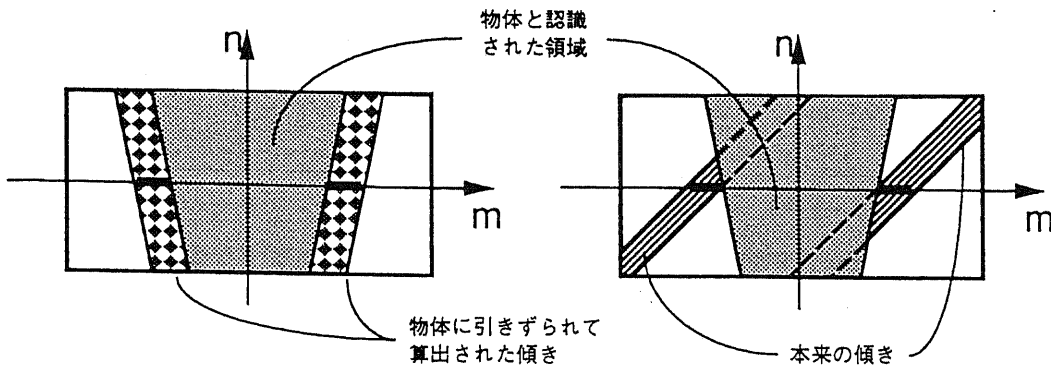


図 5.29: 従来の方法で算出される傾きと, 本来の傾き

値より大きく物体と認識された領域は 1~19 フレームを用いてブロックマッチングした結果の視差ベクトルと平均の色をそのまま用いる。

背景に関しては, 図 5.30 に示すように, 先ほどの分離方式により背景と認識された領域のみで再び 1~19 フレームでブロックマッチングを行なう。もちろんこの際, 検索する傾きは閾値より小さいもののみとする。

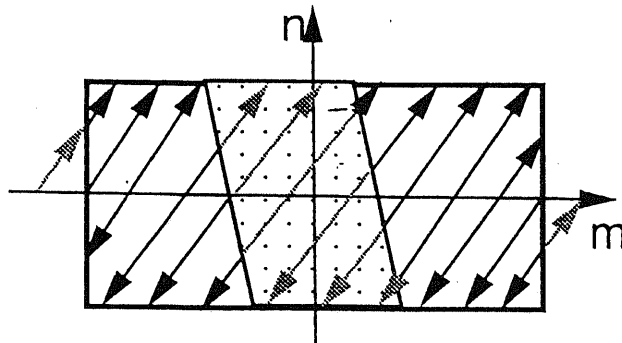


図 5.30: 背景と物体を分離した後の背景の視差ベクトルの抽出法

実験方法

5.5.5 節で示した方式を用いて物体と背景を分離することで背景上のブロックの視差ベクトルを求めてみる。

• 犬の画像

- 物体と背景を分離する際の視差ベクトルの閾値は, この犬の画像の場合+3 ピクセルとする。
- 背景の横幅は従来通り $256 + 48 \times 2$ ピクセルとする。一方多眼画像が主に 3 次元画像として用いられることを考えると, 物体が画角からはみ出ることがないと考えられるので, 物体の横幅は 256 ピクセルで十分だと考える。

- 非撮影方向からの画像の合成は5.5.4節と全く同じように行なう。
- 非撮影方向からの画像として、19眼画像の各フレーム間に2つずつフレームを内挿することで、55眼画像を作成した。
- フラワーガーデン
 - ブロックサイズは同様に8ピクセルとする。
 - 犬の画像とは違いフラワーガーデンは非常に背景のテクスチャが複雑で、かつ色の濃淡がはっきりしている。そのため視差ベクトルに関しては極めてシビアであり、犬の画像の場合より細かくとらなければならない。今回は隣のフレームに対して1/7ピクセル、中心画像(532番画像)から端の画像(525,539番画像)だと1ピクセルの細かさとした。
 - 視差ベクトルの検索範囲は隣のフレームに対して同様に±6画素。
 - 物体と背景を分離する際の視差ベクトルの閾値は、フラワーガーデンの場合-3ピクセルとする。
 - 両襟につけるブロックは従来通り48ピクセルとする。
 - 背景の横幅は $352 + 48 \times 2$ ピクセルとし、物体の横幅は352ピクセルとする。

実験結果

- 犬の画像の場合

5.5.5節で示した方法を用いて背景と物体を分離することで予測特性の向上を試みた。例として $y = 100$ のmn画像をとり、以下にその結果を示す。



図 5.31: $y = 100$, 物体と認識された領域

この領域のみの視差ベクトルを図 5.32 に示す. また, それに加え, 背景と認識された領域のみで視差ベクトルを抽出したのが図 5.33 である.

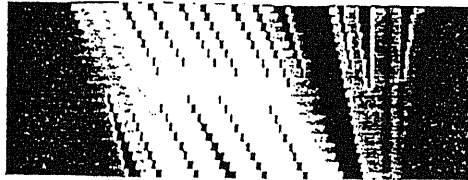


図 5.32: $y = 100$, 物体と認識された領域の視差ベクトル

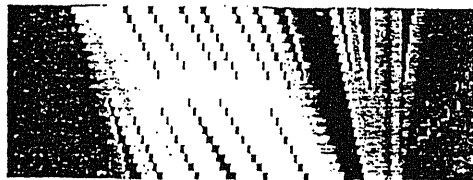


図 5.33: $y = 100$, 背景と物体を分離して求めた視差ベクトル

図 5.33 で求められた視差ベクトルを 5.5.3 節で示した方式により内挿したのが図 5.34 である.



図 5.34: $y = 100$, 視差ベクトルを内挿した結果

また、その際の SNR を以下に示す。図 5.35 が mn 画像ごとの SNR、図 5.36 がフレームごとの SNR である。

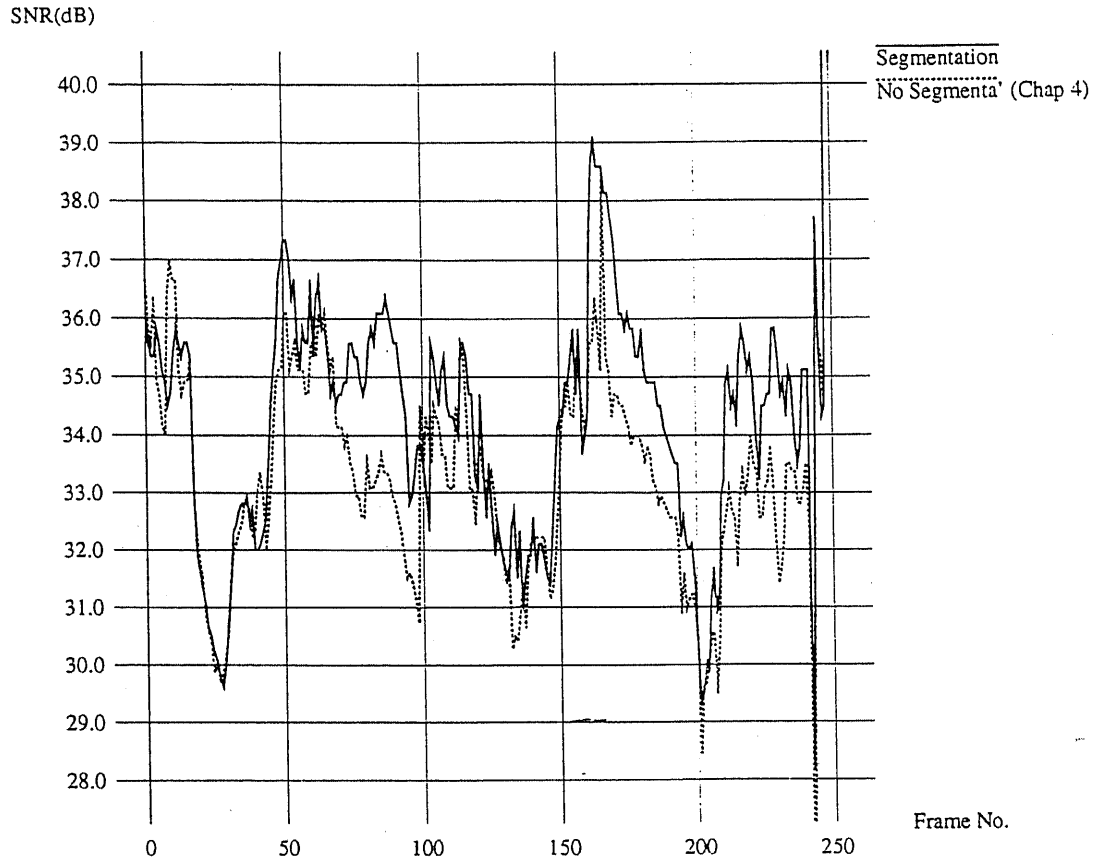


図 5.35: mn 画像の SNR

また、以下に物体と背景の分離を行なって、補間した結果の SNR を示す。

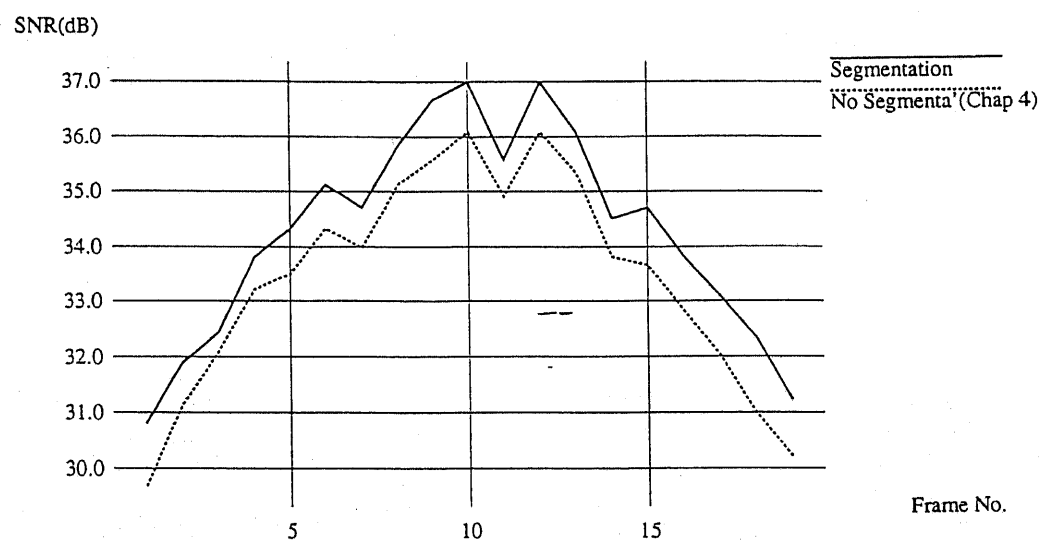
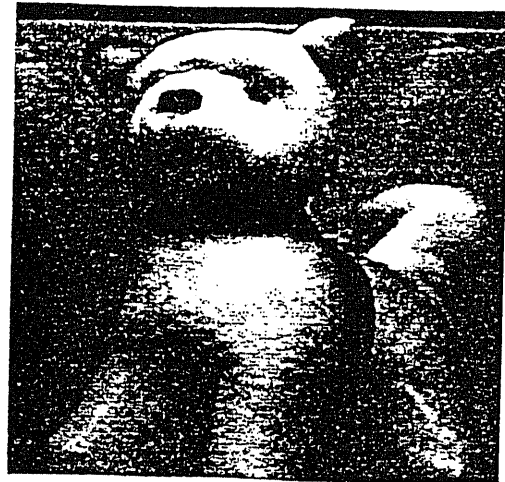


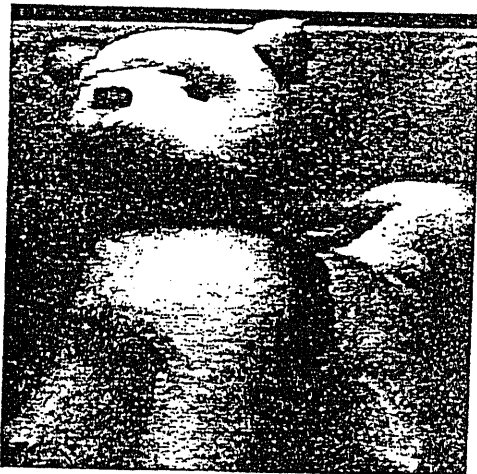
図 5.36: 背景抽出した時としなかった時の SNR



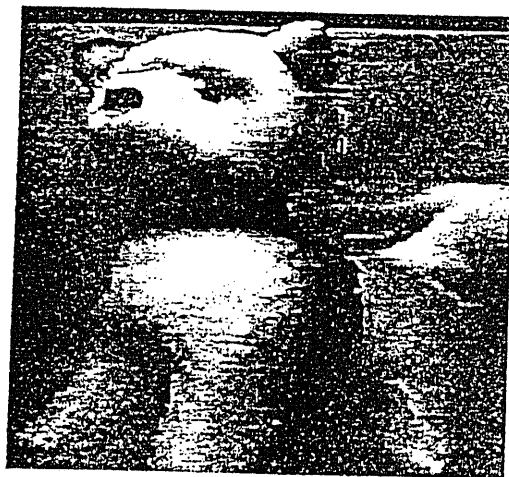
FrameNo.5



FrameNo.10



FrameNo.15



FrameNo.19

図 5.37: 復号画像

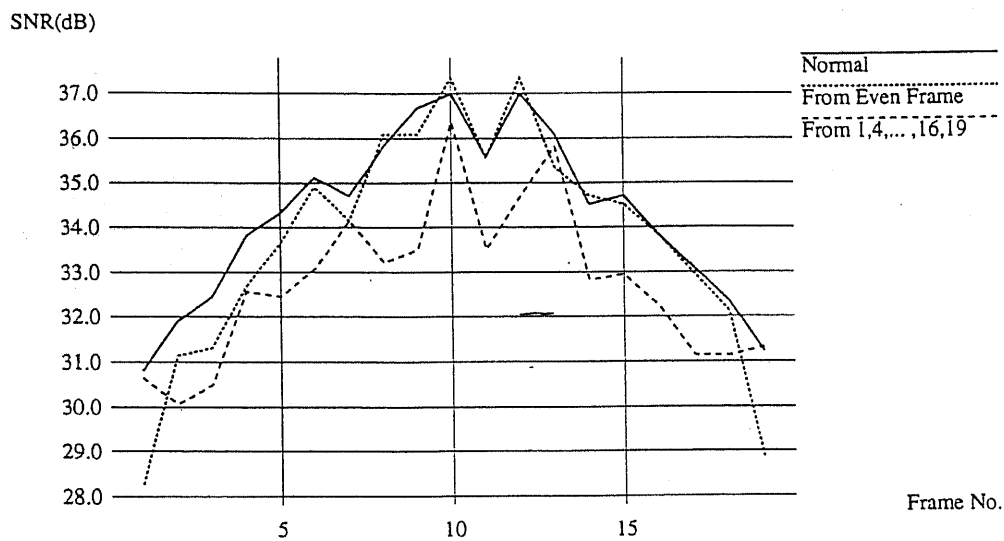


図 5.38: 補間画像の SNR

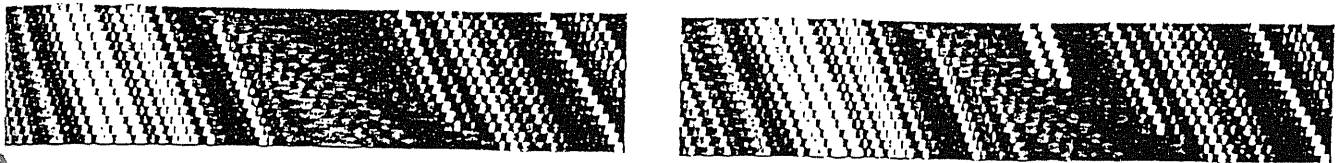
● フラワーガーデンの場合

同様にフラワーガーデンにも適応してみた。いくつかの復号 mn 画像を図 5.39 に、復号画像を 5.40 に示す。左が背景抽出をしていない画像で、右が背景抽出をした画像である。

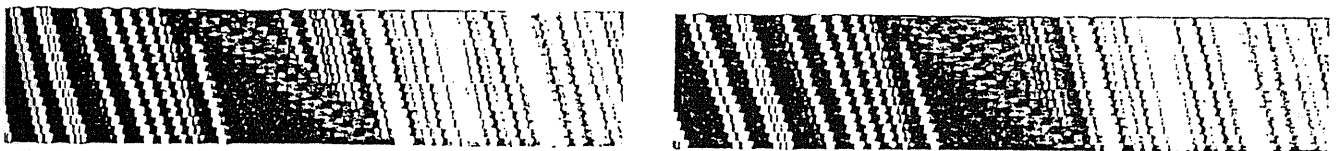
また、その際の SNR を図 5.41 に示す。



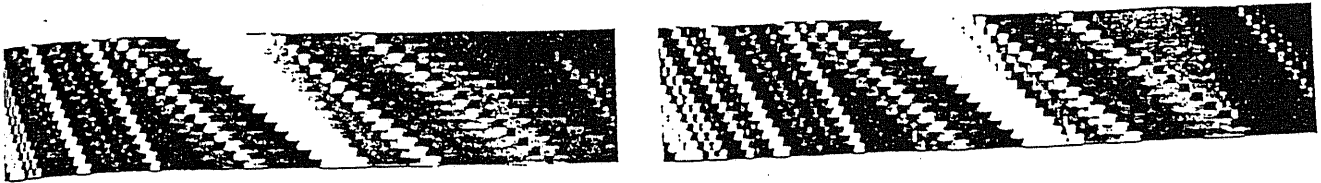
$y = 0$



$y = 50$

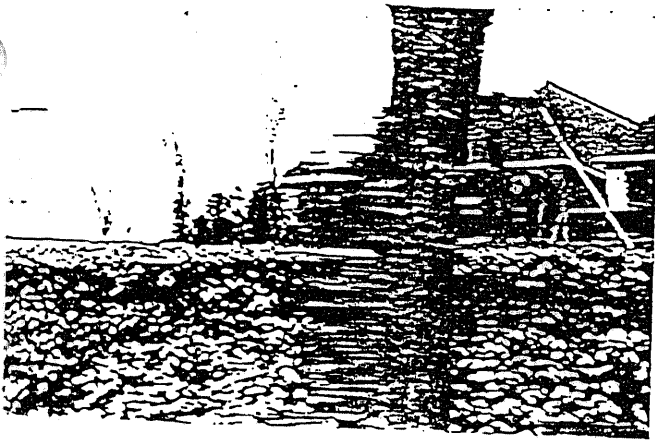


$y = 100$

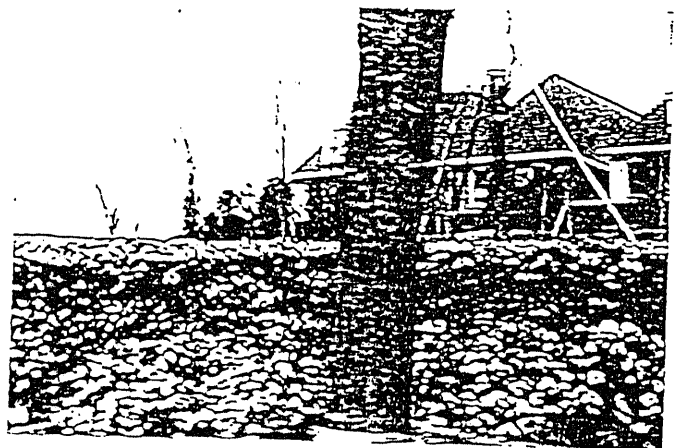
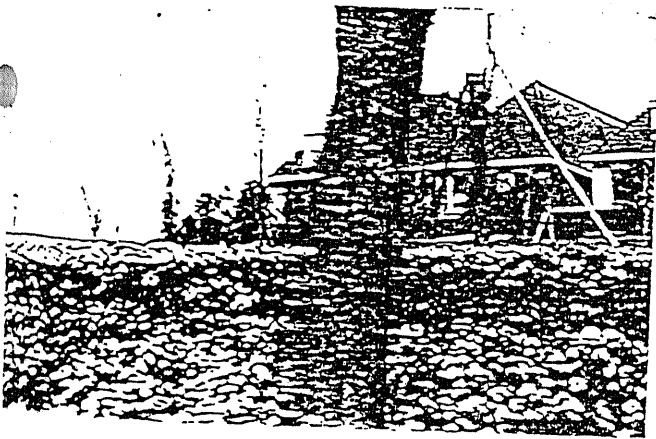


$y = 150$

図 5.39: 背景抽出した時としなかった時の復号 mn 画像



FrameNo.525

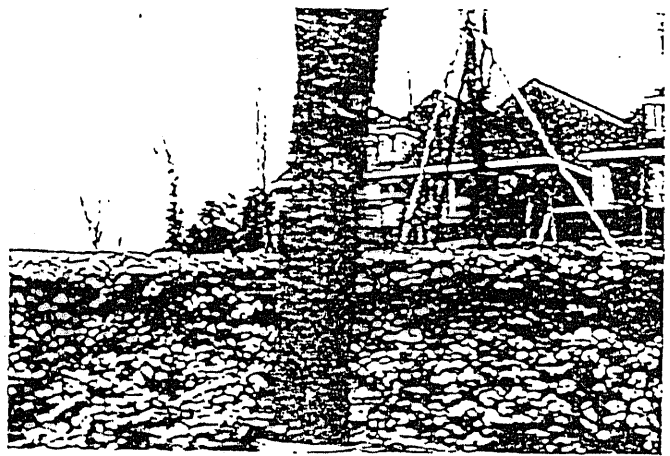
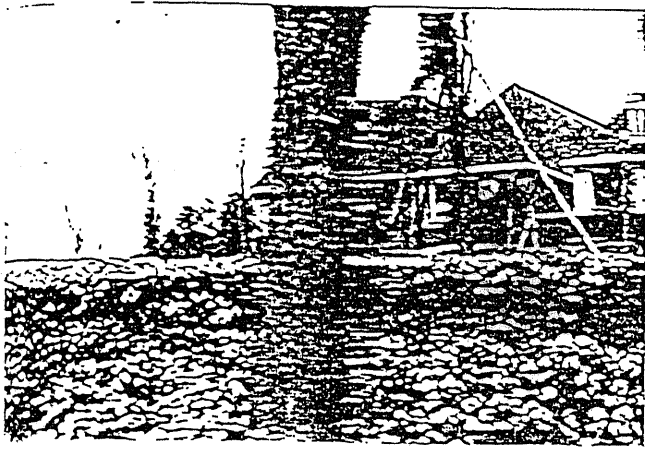


FrameNo.530

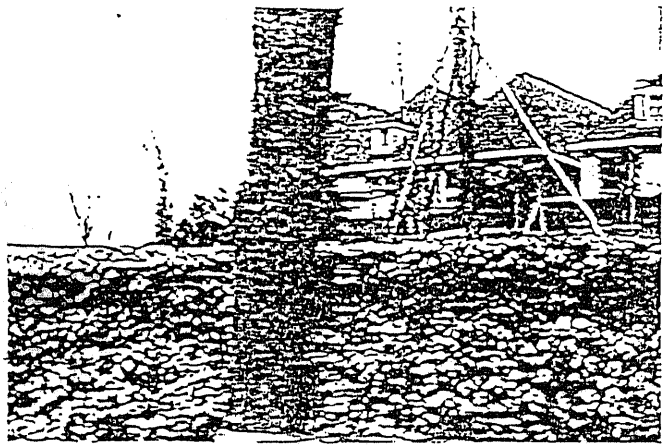
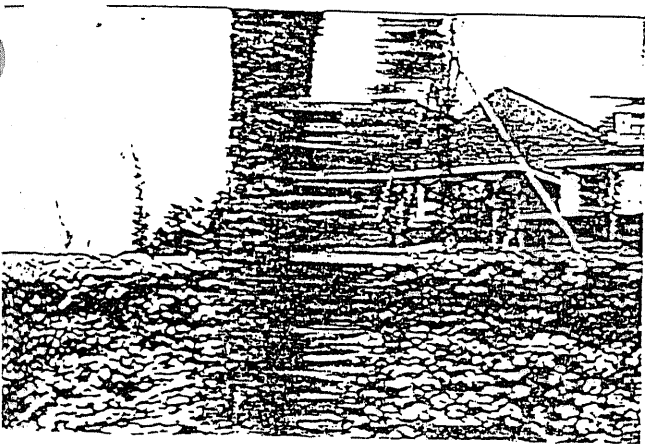
多眼画像の幾何学的拘束を利用した視差補償において、物体と背景の分離を導入してみた。その結果、物体と背景の初めの画像を別々に送った後の符号量は、従来通りの 67584bit (背景) に加えて、49152bit (物体) 増加するものの、これだけの情報量の視差ベクトルを送ることにより、SNR は 34.3dB となり、背景分離しない時に比べ 1.1dB 上がった。フレーム単位で見た場合、1.1dB しか上がっていないが、図 5.35 に示したように、mn 画像単位で見ると $y = 80 \sim 100$ 近辺では SNR は 2.5dB 程度改善しているのがわかる。

また、犬の画像のように背景の単調な画像では 1.1dB 程度しか符号量が増えないが、フラワーガーデンのように背景のテクスチャが複雑な画像では、背景分離をしないときで平均 21.6dB、したときで平均 23.6dB となり 1.8dB の向上が見られた。なお、フラワーガーデンの復号画像の SNR が、中心画像において下がってしまうのは、空間上の各点が消失点に近づく速度が直線的でないため mn 画像を構成する直線群が曲がっているためである。

しかしながら本方式によって物体と背景を分離した場合、次のような問題が生じる。まず、犬の $y = 175$ の mn 原画像、及び本方式による物体と背景の分離を行なった結果を図 5.42 に示す。このように本方式で物体と背景の分離を行なうと、図 5.43 に示すような中心画像写っていない、両端のフレームに行くにつれて現れてくる領域は剥し残ってしまい、5.44 に示すように復号画像の背景に物体の側面の色がついてしまう。



FrameNo.535



FrameNo.539

図 5.40: 背景抽出した時としなかった時の復号画像

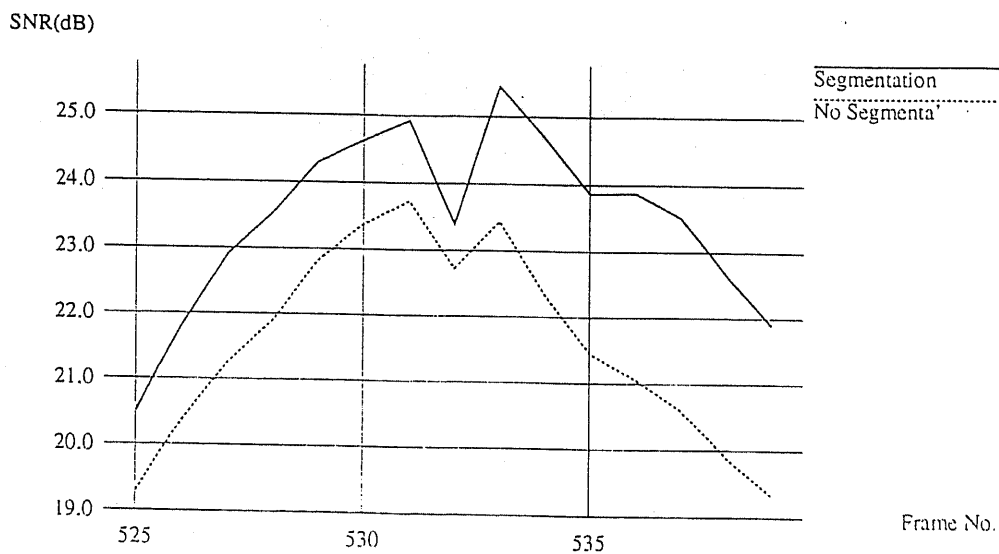


図 5.41: 背景抽出した時としなかった時の SNR

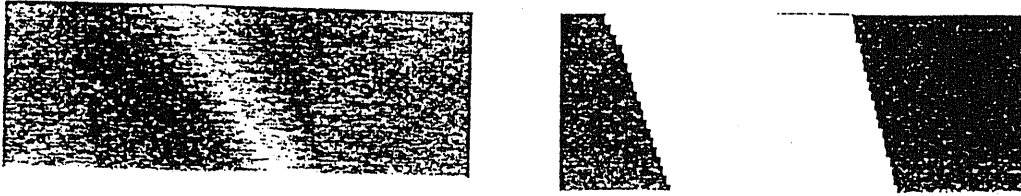


図 5.42: $y = 175$, 原画像及び物体と認識された領域

この際見た目には分かりにくいですが、物体側面に図 5.43 に示すような剥し残りの領域が出てしまう。

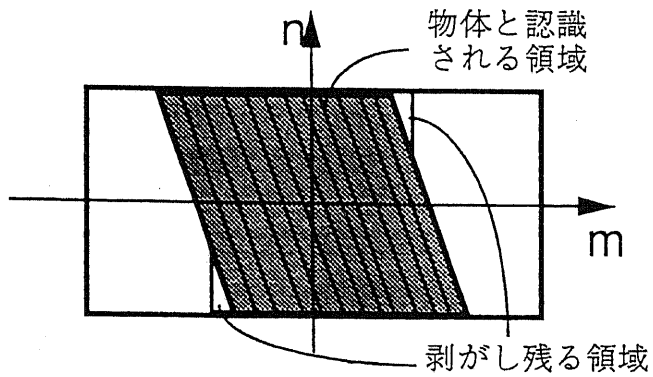


図 5.43: 剥し残る領域

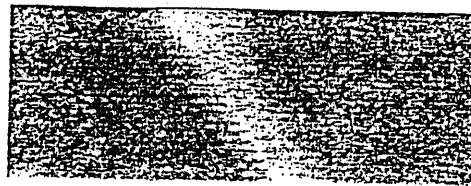


図 5.44: $y = 175$, はがし残りによる結果

5.6 まとめ

3次元画像の波形符号化のアプローチとして「視差補償予測符号化」の検討をおこなった。

まず第1に、ステレオ画像、多眼画像は従来の動画像と同様の性質をもつことから、動画像標準符号化方式を適用することができる。本章では H.261 をベースとして多眼画像への適用例について述べた。

次に、多眼画像の間にある幾何学的な関係を積極的に利用した符号化法についての検討をおこなった。全ての視差画像を参照して視差ベクトルを抽出する手法、アフィン変換により視差ベクトルを補間して表示する手法、予測特性向上のために背景と物体を分離する手法について検討し、シミュレーションをおこなった。

本章で得られた主な結果は以下の通りである。

1. 標準符号化方式 H.261 による多眼画像の符号化において、視差ベクトルの探索範囲を限定して、視差ベクトル符号量を減少させた結果、32kbit/frame という低レートのときだけ SNR が改善された。
2. 1 フレームおきにフレーム間符号化を行ない、そこで求めた視差ベクトルを2等分することで、間のフレームの予測画像を作る方式の符号化では、間のフレームの視差ベクトル符号量をゼロにすることが可能であり、減少した視差ベクトルの符号量分だけフレーム間符号化に割り当てる符号量を増加させ、量子化ステップサイズを小さくしたところ、全般的に画質の改善がみられたが、期待されるほどではなかった。
3. 多眼画像の幾何学的な拘束を用いて視差ベクトルの抽出した。その結果、初めの1枚の画像を送った後は67584bitの視差ベクトルを送ることにより残りの18枚の画像を復号できる。また、それを補間することにより、SNR は平均33.4dB になった。
4. 情報量は増えるものの物体と背景の近傍など、オクルージョンが激しい領域での予測特性を改善するため、物体と背景を分離した後に、背景画像の視差ベクトルを抽出した。その結果、物体と背景の分離をしない時に比べ SNR が平均で1.1dB 改善した。
5. フラワーガーデンに適用することにより、背景のテクスチャが複雑な画像では物体と背景の分離が有効であることがわかった。
6. 情報量の圧縮及び撮像系の簡略化のため、非撮影方向からの像の補間・合成実験を行なった。その結果本方式を用いれば内挿して合成する画像であれば、補間のための特別な処理を施すことなく、容易に非撮影方向からの像を合成することができた。

以上をまとめると、予測画像の作成に多眼画像の特徴である視差を利用することによって、符号化の品質が向上することを確認できたといえる。しかし、非常に冗長度の高い多眼画像の符号化実験であることを考慮すると、期待されたほどではなかったといえる。

本論文で提案した圧縮法は多眼画像の幾何学的拘束を用いた、第一段階の圧縮法である。近くの標本点は輝度データが非常に似通っており、従来の2次元画像における帯域圧縮の手法を適応することにより、さらなるデータ圧縮が可能である。さらにmn画像でデータ圧縮を行なったため縦方向の相関を用いておらず、縦方向の視差ベクトルの相関を用いることで符号量が削減され、復号画像の予測特性が改善すると思われる。また、従来の動画像圧縮技術に見られるように、予測誤差の符号化も必要であろう。今後の研究が期待される。

参考文献

- [1] CCITT SGXV: "Description of RMS", doc. no. 525 (1989-06).
- [2] 山口, 立平, 秋山, 小林: "2眼式立体画像の情報圧縮と奥行き形状の再現性", 信学論, D-2, Vol. J73-D-2, No. 3, pp. 317-325 (Mar. 1993).
- [3] W.A. Schupp, 安田(靖): "視差補償および動き補償を用いたステレオ動画像のデータ圧縮", *PCSJ88*, pp. 63-64 (1988).
- [4] 泉岡, 渡辺: "視差補償を用いたステレオ動画像符号化", 1989 信学会春季全大, D-273 (1989).
- [5] M.E. Lukacs, "Predictive Coding of Multi-Viewpoint Image Sets", *ICASSP 86*, pp. 521-524 (1986).
- [6] W.A.Schupp, Y.Yasuda, "Efficient coding of 3-D moving pictures with adaptive motion/disparity compensation", *Journal of three dimensional images*, Vol.3, No.1, pp.47-52(1989)
- [7] M.G. Perkins: "Data Compression of Stereopairs", *IEEE Trans. Commun.*, Vol. 40, No. 4, pp. 684-696 (Apr. 1992).
- [8] M.G. Perkins: "Data Compression of Stereopairs", *IEEE Trans. Commun.*, Vol. 40, No. 4, pp. 684-696 (Apr. 1992).
- [9] B. Choquet, et al., "3D TV studies at CCETT", *Proc. of TAO First International Symposium on Three Dimensional Image Communication Technologies*, S-1-2 (Dec. 1993).
- [10] Y. Nakaya and H. Harashima, "An iterative motion estimation method using triangular patches for motion compensation," *Proc. of SPIE Visual Communications and Image Processing '91: Visual Communications*, vol. 1605, pp. 546-557, Nov. 1991.
- [11] S. Nagata: "Visual Effects in Multidirectional Steroscopic Images", *The international symposium on three dimensional image technology and arts, Seiken symposium* Vol. 8 (Tokyo, February 1992).
- [12] A. Shertz, "Source coding of stereoscopic television picture," *Proc. of 1st International Symposium on 3D Images*, pp. 165-171, Sep. 1991.
- [13] A.Tamtauoi, C.Labit, "Joint estimation of motion and disparity for stereoscopic sequence coding", *First Int'l Symp. on 3D Images*(Paris,1991)

第 6 章

3次元構造推定に基づく3次元画像符号化

通常の撮影条件においては、光線空間には大きな冗長性がある。これは、光線空間が3次元空間内の物体の投影像の集まりであって、かつ、投影像は視点の位置変化に対して大きく変化しないという性質に由来するものである。

本章では、「光線空間のモデル化」という観点から、3次元構造モデルを用いた3次元画像符号化について検討する。具体的には、視差補償予測符号化を幾何学的な制約を用いて多眼画像へと拡張することにより、対応点問題を避けながら構造抽出符号化手法の利点と直接高能率圧縮の利点を組み合わせた方式を提案する。具体的には、復号画像のSNRを評価基準として設定し、物体の形状と表面テクスチャの決定を最適化問題として解くという手法である。

- [14] F.Chassaing, D.Pele, "3DTV compatible transmission of stereoscopic television signal", First Int'l Symp. on 3D Images(Paris,1991)
- [15] J.L.Dugelay, B.Choquet, "A 3D image analysis algorithm and stereoscopic television", First Int'l Symp. on 3D Images(Paris,1991)
- [16] Ziegler, M. and Tengler, W., "Stereo Compensated Image Coding", *Proc. of The International Symposium on Three dimensional Image Technology and Arts, Seiken symposium*, vol. 8, pp. 163-170, Feb. 1992.
- [17] M. Ziegler and F. Seytter, "Coding of Stereoscopic Sequences using Disparity and Motion Estimation", *PCS '93*, 12.4, Mar. 1993.
- [18] A.Shertz., "Source coding of stereoscopic television picture", First Int'l Symp. on 3D Images (Paris, 1991)
- [19] A.Tamtaoui, C.Labit, "Joint estimation of motion and disparity for stereoscopic sequence coding", First Int'l Symp. on 3D Images (Paris, 1991)
- [20] F.Chassaing, D.Pele, "3DTV compatible transmission of stereoscopic television signal", First Int'l Symp. on 3D Images (Paris, 1991)
- [21] J.L.Dugelay, B.Choquet, "A 3D image analysis algorithm and stereoscopic television"
- [22] A. Kopernik and D. Pele, "Disparity Estimation for Stereo Compensated 3DTV Coding", *PCS '93*, 16.5, Mar. 1993.
- [23] 吉岡, 苗村, 藤井, 中屋, 原島: "三次元画像符号化基礎検討(2) -多眼画像の視差補償予測符号化-", 1993年信学会春季全大, D-327 (Mar. 1993).
- [24] 苗村, 吉岡, 藤井, 中屋, 原島: "多眼画像の視差補償予測符号化", 1993 TV 学会年次大会, 24-4 (1993-7).
- [25] 藤井, 原島: "アフィン変換による多眼画像の視差補償予測符号化", 1993年信学会秋季全大, D-200 (Sep. 1993).
- [26] 藤井, 原島: "多眼式三次元画像符号化 -視差補償予測符号化から構造抽出符号化へ-", *PCSJ '93*, 7-3 (Oct. 1993).
- [27] 藤井, 原島: "アフィン変換を用いた多眼式三次元画像符号化", *SITA '93*, T11-1 (Oct. 1993).

6.1 はじめに

3次元構造抽出符号化の実現のためには、コンピュータビジョンの分野で研究されてきた成果を利用し、それらを組合せ、最も良い結果を出す手法を流用すればよいかのようにも思われる。コンピュータビジョンの分野において2次元画像からの3次元構造の復元の問題は、“Shape from X”として知られている。Xに入るものとしては、輪郭(contour)、陰影(shading)、動き(motion)、テクスチャ(texture)、ステレオ(stereo)などが挙げられる。特に、多眼画像を入力とした場合にはモーションステレオ(motion stereo)の手法を利用することができる。

しかしながら、本論文においては、

物体の正確な3次元構造の推定が目的ではない、

ことに注意する必要がある。本論文における目的は、3次元画像の圧縮、符号化であって、3次元画像の統一的な符号化方式を模索することである。先に挙げた「光線空間の伝達」という観点からは、送信側における光線空間を受信側でいかに再生するかということを考える必要がある。すなわち、

3次元構造の推定が究極の目的ではなく、光線空間の再生が究極の目的

であり、3次元構造推定はその手段に過ぎないという点を念頭においておく必要がある。

6.2 構造抽出に関する従来の研究

物体の3次元構造にまで立ち入って符号化する方法は、2次元動画の符号化のみならず、3次元画像の圧縮、補間に非常に有望な方法であると思われる[11]。対象に対する知識をあらかじめ有している場合には、3次元構造モデルを当てはめ、形状と動きパラメータを符号化、伝送すればよく、また対象に対する知識を持たない場合には、まず物体の3次元構造を推定し、形状情報と表面テクスチャ情報とを求めた後にそのパラメータを符号化することになる。

画像からの物体の3次元構造の復元問題に対しては、従来からコンピュータビジョンの分野において多くの研究の蓄積がある。特にステレオ画像が利用できる場合には、単眼画像からの3次元構造復元の手がかりに加えて両眼の視差が利用でき、視差の検出すなわち対応点問題が本質的な課題となる[15, 11]。さらに多眼画像が利用できるならば、2枚ごとの対応づけの結果を統合することで対応のあいまい性を減少できることに加えて、時空間画像解析の手法を用いることができる。等速運動するカメラで撮影した時空間画像の軌跡が直線となることを利用して物体の3次元構造を求める手法[5]や、エピポーラ平面画像(EPI)における軌跡を解析することにより物体の3次元構造を求める手法[6]、因子分解法を用いて物体の3次元構造とカメラの回転パラメータを安定に求める手法[19]、多眼の動画からの物体の構造およびカメラの動きを推定する手法[16]などが提案されている。

任意視点からの画像の合成に関しては、物体の3次元構造、表面テクスチャ、光源情報が完全に復元されていればCGの手法を用いて補間画像を合成することができる。その他の手法として、視差ベクトルの内挿予測をすることにより中間像の合成をおこなったもの[11, 25]や、EPIのもつ幾何学的特徴を用いるもの[17]、EPIの周波数解析により視差画像補間を行なう手法などが報告されている[18]。

さて、本研究においては、3次元ディスプレイに表示する際に必要なデータの圧縮、補間画像の合成が目的であり、それを実現する符号化手法の構築が目的である。構造抽出的な手法は3次元画像の圧縮、補間画像合成に非常に有望な方法であると思われるが、対応点問題を安定に

解くのは難しく、3次元構造の完全な復元が困難であるという問題点がある。しかし、本研究の目的は3次元構造の正確な復元ではない。ここでもう一度原点に立ち返り、3次元画像の伝送という観点から視野、視域を限った場合の3次元画像のもつ本質的な情報について検討を加え、それをふまえた上で圧縮、符号化手法を構築する必要がある。

6.3 被写体空間と視差画像の幾何学的な関係

光線空間と被写体空間との関係については先に述べた。本章では、3次元構造モデルを用いた予測符号化を行なう際に便利のように、被写体空間と視差画像の幾何学的な関係についても一度整理しておく。ここで導入する視差画像空間は光線空間と比べて幾何学的にゆがんだ構造となっているが、視差補償予測との整合が良く、実装を行い易くなっている。

6.3.1 投影モデル

被写体空間の座標系（物体座標系）を $O-X-Y-Z$ とし、投影面における座標系（カメラ座標系）を $o-x-y$ とする。像の逆転を防ぐため、図 6.1 に示すように仮想的に投影面をレンズと物体の間にもってくるものとする。視点位置を物体座標系で (X_v, Y_v, Z_v) とし、被写体空間中の点 $P(X_p, Y_p, Z_p)$ から視差画像上の画素 (i, j) への投影の関係を導出する。

レンティキュラーシートなどを用いた多眼3次元画像においては、ディスプレイ面上に提示する点を基準点として各視差画像の位置合わせを行なう。ここでは、物体座標系の原点を基準点にとることとし、物体座標系の原点がカメラ座標系の原点に写されるとする。投影モデルとしては被写体空間の $X-Y$ 面と投影面 $x-y$ とが平行な中心投影を仮定する。このような投影法を仮定した場合、物体上の点 $P(X_p, Y_p, Z_p)$ から投影面への投影は、

$$x = -f \frac{X_p - \frac{X_v}{Z_v} Z_p}{Z_p - Z_v} \quad (6.1)$$

$$y = -f \frac{Y_p - \frac{Y_v}{Z_v} Z_p}{Z_p - Z_v} \quad (6.2)$$

と与えられる。ここで、視点（カメラのレンズ中心）と投影面との距離を f とした。撮影された視差画像の画素の指標を i, j とすると、

$$\begin{aligned} i &= Cx \\ j &= Cy \end{aligned} \quad (6.3)$$

と表すことができる。ここで C はカメラ座標から画素への対応を表す定数であり、例えばデジタイズの際のスケーリングパラメータなどを含む。

6.3.2 視差画像空間と規格化空間

多眼3次元画像の撮影の際の視点の典型的な配置法として (X_v, Y_v, Z_v) を規則正しく配置した場合について考察する。例として $X-Y$ 平面に平行な平面内で視点を格子状に配置した場合、視点 $V(X_v, Y_v, Z_v)$ は次のようになる。

$$\begin{aligned} X_v &= m\Delta L \\ Y_v &= n\Delta L \\ Z_v &= F \end{aligned} \quad (6.4)$$

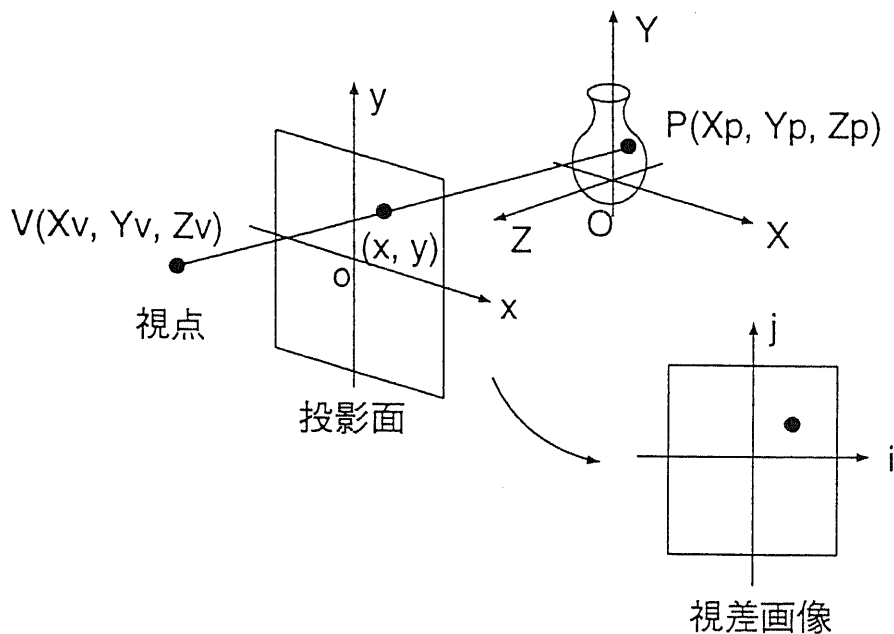


図 6.1: 投影モデル

ここで、 m, n は視点の指標を表す整数値、 F は視点から X - Y 平面までの距離、 ΔL は視点の間隔である。ここで視点の間隔は X 方向、 Y 方向とも同じであるとした。これにより多眼画像は、 m, n の位置から撮影した視差画像として $I_{m,n}(i, j)$ と表すことができる。

さて、視点の指標 m, n 、および画素の指標 i, j を実数変数に拡張し、

$$f(m, n, i, j) = I_{m,n}(i, j) \quad (6.5)$$

と定義すると、4次元空間 (m, n, i, j) 中の格子点上でのみ値をもつ関数 $f(m, n, i, j)$ を考えることができる。この空間のことを本論文では「視差画像空間」と呼ぶことにする。

さらに、視差画像空間と被写体空間の関係を見通し良くするため、被写体空間を規格化する。 X, Y 座標を画像内の座標 i, j で規格化したものを X', Y' とし、また奥行き座標 Z' を、隣あった視差画像に Z' pixel 分の視差を与える量として規格化する。すなわち、 (X', Y', Z') を

$$\begin{aligned} X' &= \frac{f}{C} \cdot \frac{X}{Z_v - Z} \\ Y' &= \frac{f}{C} \cdot \frac{Y}{Z_v - Z} \\ Z' &= \frac{f}{C} \cdot \frac{\Delta L}{Z_v} \cdot \frac{Z}{Z_v - Z} \end{aligned} \quad (6.6)$$

と定義する。この時、点 (X', Y', Z') から (m, n, i, j) への投影は、

$$\begin{aligned} i &= X' - mZ' \\ j &= Y' - nZ' \end{aligned} \quad (6.7)$$

と与えられることになる。このような仮想的な被写体空間を「規格化空間」 (X', Y', Z') と呼ぶことにする。

(6.7)式によれば、規格化空間内の一点 (X', Y', Z') は視差画像空間 (m, n, i, j) 中において平面を構成することが分かる。特に、視点を水平面に等間隔で配置した場合を考えるならば、 $n=0$ とおけばよく、この場合、 $m-i$ 平面に平行な断面がエピポーラ平面画像(Epipolar Plane Image; EPI)と一致する。したがって、このようなカメラ配置の場合のエピポーラ平面画像は直線から構成されることが分かる。

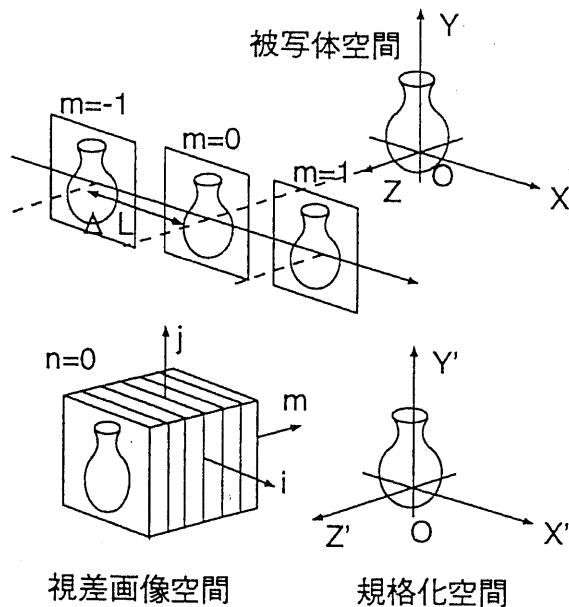


図 6.2: 視差画像間の幾何学的な関係と規格化空間

6.4 3次元構造モデルを用いた視差画像空間の符号化

6.4.1 問題の定式化

多眼画像の圧縮の問題とは、視差画像空間をいかに効率的に記述し、符号化するかという問題となる。例えば、動画画像の標準符号化方式である H.261 や MPEG を多眼画像に対して適用した符号化手法は、 $f(m, n, i, j)$ のデータのうち、 i, j 方向の冗長性を離散コサイン変換 (Discrete Cosine Transform; DCT) により削減し、 m, n 方向の冗長性を近接した視差画像 $m \pm \alpha, n \pm \alpha$ からの予測により削減するという手法であると考えることができる。

また、中間像の補間合成の問題とは、視差画像空間のうち m, n が整数でない部分のデータをいかにして補間合成するかという問題となる。視差ベクトルの内挿により中間像の合成を行なう手法は、隣あった m, n との間で視差を求め、線形補間することにより中間像を合成する手法である。また、被写体の3次元構造を正確に復元できるならば $f(m, n, i, j)$ の全てのデータを再構成できるが、 (m, n, i, j) の格子点上のデータのみから被写体の3次元構造を完全に復元することは一般には困難である。

以上の観点から、中間像の合成ができるような形で情報圧縮を実現するためには、 m, n が整数でない部分のデータの合成ができる形でいかにして $f(m, n, i, j)$ のデータを効率良く記述するかという問題となる。したがって、これら2つの条件を満足するような符号化法であるな

らば、被写体の3次元構造が正確に復元されることは必ずしも必要ではない。

3次元構造抽出符号化においては、得られた多眼画像を、あるモデルの形状と表面テクスチャとに圧縮、符号化することになる。この際、

伝送すべき情報量、復号画像のSNRなどを総合的に考えた上で、

モデルの決定、モデルパラメータの推定、表面テクスチャの決定を行なう必要がある。これらは相互に依存関係にあり、上記の意味での最適な解を決定するのは非常に困難である。

本章では、視差画像空間 (m, n, i, j) と規格化空間 (X', Y', Z') との写像関係を積極的に利用した予測符号化について検討する。一般に、フレーム間予測符号化においては、

1. ブロックや三角形パッチなど、予測のもととなるモデルを仮定し、
2. 2乗誤差などの評価値を設定し、
3. モデルの移動、変形パラメータを最適化する、

という段階から成り立っている。これを3次元構造モデルを用いた予測符号化に拡張すると、

1. (X', Y', Z') 内に3次元構造モデルを仮定する。この際、3次元形状モデルの種類、表面の性質などの制約条件を決定する。
2. 評価値を設定する。例えば復号視差画像の平均的なSNRなど。
3. 上で定めた評価値を規範としてモデルパラメータを最適化する。

という手法となる。以下、それぞれについて述べる。

6.4.2 3次元構造モデル

符号化を目的とした3次元構造モデルとしては、まずデータ量が小さいこと、そして形状の表現能力に汎用性があり、変形に対する柔軟性を有していることが求められる。また、用いた3次元構造モデルの表現能力と実際の情景の複雑さとの関係により予測特性が異なるので、モデルを適切に選ぶ必要がある。

以下に、3次元構造モデルの簡単な例を挙げ、従来の視差補償予測符号化との関連について述べる。直観的な理解を助けるために、規格化空間の $X'-Z'$ 断面と視差画像空間の $m-i$ 断面との関係で示すことにする。

平面モデル：

1枚の平面で表現するものであり、モデルとしては最も簡単なものである。モデル上の各点は同じ奥行き z にあり、 (m, i) 空間内では同じ傾きの直線群となる。 Z' 軸に垂直な平面を $Z' = 0$ 上に固定し、中央から見た画像を表面テクスチャとして貼りつけた場合、中央の画像 $I_{0,0}(i, j)$ をそのまま各画像の予測画像として用いる場合に相当し、動画像のフレーム間予測における前値予測と等価な手法となる。またパラメータを一つ増やし、 Z' 軸方向に自由度を持たせるならば、視点移動量に比例させて画像全体を平行移動して予測する方法となり、この方法だけでもステレオ画像におけるかなりの冗長性を除くことができること報告されている [2]。

ブロックモデル :

対象をZ'軸に垂直なブロックの集まりで表現し、各ブロックに対してZ'軸方向に自由度を持たせたものである。ブロックマッチングに基づくステレオ画像の視差補償は、一方の画像を中央の画像とみなし、他方の画像から中央画像を予測する方式とみなすことができる。1つのブロック内での各点は同じ奥行きにあるが、隣あったブロック同士では奥行きが不連続に変化する。したがって、 (m, i) 空間内では傾きが不連続に変化する直線群となる。図から分かるように、 $u = 0$ 以外では重なる部分や穴があいてしまう部分がある。穴が空いた部分は復号画像において欠損部分となってしまうので、復号時に何らかの方法で埋める必要がある。

平面パッチモデル :

対象を平面からなる連続なパッチの集まりとして表現するものである。 xz 平面での断面をみると折れ線となる。この場合、奥行きが直線的に連続に変化するので、復号画像に穴が空くことはない。CGモデルにおいては三角形パッチモデル、四角形パッチモデルがよく用いられる。

その他、スプライン曲面、円筒モデル、球面モデル、超二次曲面体などさまざまなモデルを考えることができる。これらは非線形な予測符号化となる。さらには、このようなモデルを情景内に2つ、3つと増やしていてもよい。このようなモデルを情景内に2つ、3つと増やしていくことによりより複雑な (m, i) 空間の表現が可能である。モデルの違いによって (m, i) 空間の表現能力、モデル記述のための情報量が異なる。また同じモデルであってもブロックサイズや奥行き方向の量子化数などの違いによっても、モデル記述のための情報量とともに予測特性が異なる。すなわち、より複雑なモデルを用いることで (m, i) 空間の表現能力を上げることができるが、その反面、モデル記述のための情報量が増加することに注意が必要である [31]。

さらに、光源、物体表面の反射率などをモデル化することにより、同じ点であっても視点の移動にともなって輝度値が変化するような記述も可能となる。この場合も、モデル記述のための情報量および処理量の増加と復号画像のSNRの向上とのトレードオフを考慮する必要がある。

6.4.3 評価値の設定

復号視差画像の評価値として、ここでは画像符号化において標準的に用いられている信号対ノイズ比 (signal-to-noise ratio; SNR) を用いることにする。定義は次の通りである。

$$\begin{aligned} \text{SNR}(m, n) &= 20 \log_{10} \frac{255}{\sqrt{\text{MSE}(m, n)}} \\ \text{MSE}(m, n) &= \frac{1}{S} \sum_{i, j} (I_{m, n}(i, j) - I'_{m, n}(i, j))^2 \end{aligned} \quad (6.8)$$

ここで、 $I_{m, n}(i, j)$ は m, n の指標で表される視点からの視差画像であり $I'_{m, n}(i, j)$ は復号視差画像である。SNR(m, n) は、指標 m, n の復号視差画像のSNRを表し、MSE(m, n) は指標 m, n の復号視差画像の平均2乗誤差を表す。また S は視差画像1枚あたりの画素数である。

さて、本手法は視差画像空間の符号化であって、この立場からは、各視差画像のSNR(m, n) に対して重みをつけて全体の評価値とするなど、より柔軟な評価値の設定の仕方が可能である。

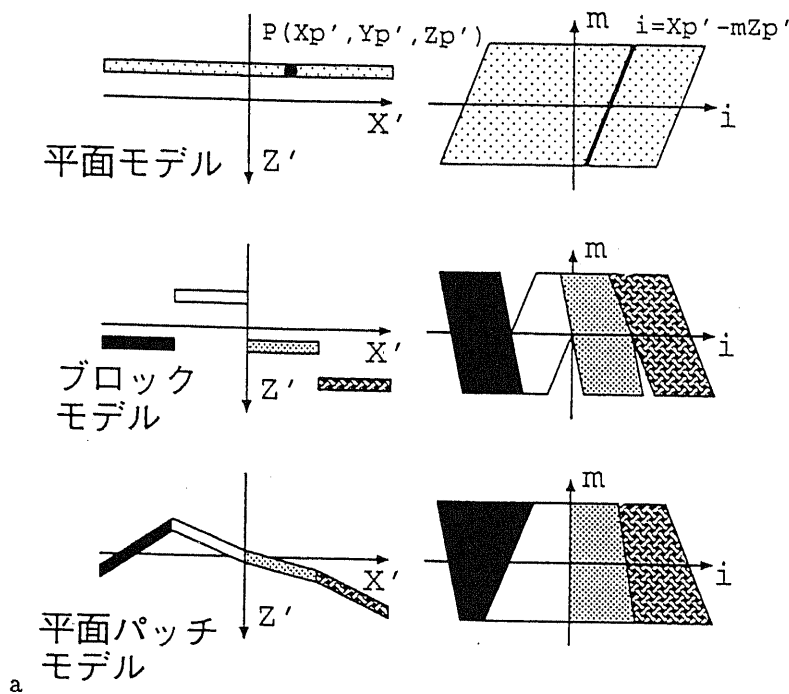


図 6.3: 3次元構造モデルと視差画像空間の記述能力の関係

例として、視差画像の SNR 評価の重み係数 $a_{m,n}$ を定義する。 $a_{m,n}$ はどの視点からの復号視差画像の SNR を重視するかということを表す係数である。 $a_{m,n}$ は 0 から 1 までの値をとり、

$$\sum_{m,n} a_{m,n} = 1 \quad (6.9)$$

であるとする。その重み付け平均を評価値とするならば、

$$\text{total SNR} = \sum_{m,n} a_{m,n} \text{SNR}(m, n) \quad (6.10)$$

と評価値を設定できる。例えば、視差画像全体の平均の SNR を上げることを目的とするならば、

$$a_{m,n} = \frac{1}{MN} \quad (6.11)$$

ただし M, N は m, n 方向の視差画像の枚数

と設定すればよく、また中央画像の SNR を上げることを目的とするならば、

$$a_{m,n} = \begin{cases} 1 & m = n = 0 \\ 0 & \text{otherwise} \end{cases} \quad (6.12)$$

と設定すればよい。この場合、従来の 2次元画像と同じ評価値となり、3次元画像符号化の特別な場合として 2次元画像符号化を位置付けることができる。

6.4.4 モデルパラメータの最適化

先に定義した評価値のもとで、3次元構造モデルの形状、表面テクスチャ、光源の有無、表面の性質などのパラメータを最適化していく。一般にはパラメータの数は非常に多く、最適化に

際しては莫大な計算コストを必要とする。本論文においては、次章で述べるようにモデルに対して制約条件を設けることにより、最適化を実行し易くしている。

6.5 3次元構造抽出符号化の実際

6.5.1 モデルの制約条件

本章では、モデルが持つ制約条件として、

- 物体表面はランベルト表面であるとする、
- 視域内の任意の視点においてオクルージョンは生じない、

という2つの制約条件を仮定する。ランベルト表面を仮定するということは視線の移動に対して表面の輝度値の変化がないということであり、オクルージョンがないと仮定するということは、視域内の任意の視点からモデルの表面が隠れることはないということである。

6.5.2 評価値空間の導入

モデルに対してこのように制約条件を定めることにより、モデルパラメータの最適化が見通し良く行なえるようになる。次の段階でのモデルパラメータの最適化処理の際に便利のように、以下に評価値の空間を定義しておく。(6.7)式で与えられる視差画像空間と規格化空間の関係を利用すると、評価値の空間は、

$$E(X', Y', Z') = \sum_{m,n} a_{m,n} \{I_{m,n}(i, j) - T(X', Y', Z')\}^2 \quad (6.13)$$

$$\begin{aligned} \text{ただし, } i &= X' - mZ' \\ j &= Y' - nZ' \end{aligned} \quad (6.14)$$

と与えられる。これを最小にする $T(X', Y', Z')$ は、

$$T(X', Y', Z') = \sum_{m,n} a_{m,n} f_{m,n}(x, y, z) \quad (6.15)$$

で与えられる(付録参照)。これは、 (X', Y', Z') の位置に $T(X', Y', Z')$ のテクスチャを持った仮想物体により予測画像を合成した場合の、各視差画像における予測誤差の重み付け平均を $E(X', Y', Z')$ と定義していることに相当する。したがって、 $E(X', Y', Z')$ の値が小さい領域を通るように物体の形状を定め、その表面に $T(X', Y', Z')$ の値を貼り付ければよい。

前節のように評価値の空間を定義することにより、最適な3次元構造パラメータは評価値の空間内でモデルの表面に沿った評価値の和の最小値を与えるものとなり、直観的に理解しやすくなる。例えば、平面モデルや平面ブロックモデルの場合には、ブロック内での評価値の和が最小となるように、各々のブロックの奥行きを定めればよい。これは従来のブロックマッチングと等価な操作である。三角形パッチモデルなどの場合には、ノード位置の各々と全体の評価値が独立ではないので、動的計画法などを用いて効率的に最適解を求めていく必要がある。

6.5.3 動的計画法の利用

三角形パッチモデルの場合には表面に沿った評価値の和を最小にする最適化問題の場合には、以下のように動的計画法が利用できる。

ここで考える問題は、分散空間中で分散の小さい領域を通るように折れ線状の形状を決定することである。この際、オクルージョンが起らないという制約条件をつけることが必要である。すなわち折れ線上のあらゆる点がすべての視点から見えなくてはならないということである。この制約は視域に依存する。極端な場合、例えば視域が0、すなわち一つの視点だけから撮影したもの（2次元画像）であれば、どんなに激しい凹凸があってもよいことになる。視域が大きくなるにしたがい、平面に近づいていくような制約条件となる。

ここでは簡単のためあらかじめ均等に配置した格子点に対して奥行き方向のみの移動を許すものとする。

この場合、問題は

奥行き方向に自由度を持った格子点を結んでできる経路で、隣あった格子点の奥行き座標の差はある一定値以下であるという制約を持った経路全体のうち、経路上の分散値の和が最小となるような経路を求める問題

として定式化できる。動的計画法 DP(Dynamic Programming) とは、 n 段階の決定を必要とする問題を、 n 個の1段階決定問題に分解して解くことにより、計算量を対数のオーダーに低減させることができるものである。これにより比較的短時間で最適な形状関数を求めることができるようになる。

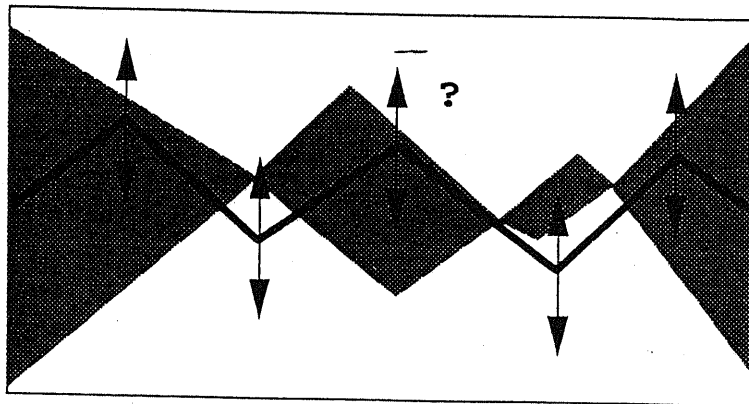


図 6.4: DP による最適経路探索

6.5.4 アルゴリズム

以上より、符号化のアルゴリズムは以下の通りとなる。

1. モデルの種類を決定する。
2. 評価値の空間 $E(X', Y', Z')$ を計算する。
3. 規格化空間 (X', Y', Z') 内にモデルを配置し、

4. モデルの表面に沿った $T(X', Y', Z')$ の値の和を計算する.
5. 4. の評価値が最小となるようにモデルパラメータを最適化する.

6.6 実験1

6.6.1 画像の取得

実験に用いた画像は4種類である。70cm離れた犬の玩具を実際に写真撮影した19枚の多眼画像、標準動画像「フラワーガーデン」に補正を施して21枚ずつの多眼画像としたもの、およびテクスチャの貼られた球の多眼画像をCGにより合成したもの2種類である。

犬の玩具はレンズの光軸が基準点を通るような輻輳撮影をおこなっている。輻輳撮影を行なうと、例えば長方形の被写体を斜めから撮影した場合に、投影面上では台形に見える（キーストーン歪み）という問題があり、視差が水平方向のみとなるような補正をかけることが必要となる。

また、標準動画像「フラワーガーデン」は並進するカメラから撮影された画像である。この場合カメラの移動方向が光軸と垂直ではないため、画像中に「消失点」が存在する。この場合にも画像を解析して消失点を求め、視差が水平方向のみとなるように補正を施す必要がある。

また、半径10cmと5cmの球を100cm離れた位置から3cmおきに19枚撮影した画像をCGにより合成した。投影方法は中心投影 (perspective projection) であり、光源を設定した画像を合成した。

また自然画像を実験に用いる場合、輝度値の分散を評価として実験を行なう都合上、画像の平均的な明るさが画像間で異なると不都合である。したがって、各画像の平均的な明るさが同じになるように明るさの補正を行なっておくことが必要となる。

6.6.2 アルゴリズム

ここでは先に挙げた3つのモデルについて実験を行なっている。3次元構造モデルパラメータの自由度は、3つのモデルとも奥行きZ方向1つのみに限定している。したがって、平面モデルや平面ブロックモデルの場合には、各ブロックの奥行きを変化させながらブロック内の評価値の和を計算し、それが最小となるように各ブロックの奥行きを決定すればよい。これに対して三角形パッチモデルの場合にはやや複雑な処理が必要となる。横一列に並んでいるノードを一組とし、その中で各ノードの奥行きを組合せすべてについて評価値を求め、最小の評価値を与えるノードの奥行きを組合せを求めることになる。この際、全探索を行なう必要はなく、動的計画法を利用して効率良く最適解を求めることができる。

以下に三角形パッチモデルを用いた場合の実験のアルゴリズムを挙げる。

1. 評価値の空間 $Var(X, Y, Z)$ を計算する.
2. 格子点の (X, Y) 座標を定め、被写体空間内にメッシュを張る (底辺、高さがともに Meshsize pixel の二等辺三角形)
3. 格子点の存在する水平ラインについて、DP を実行し最適な格子点の奥行き配置を求める。格子点間のコストとしては、格子点を結ぶ直線と y 方向に上下 Meshsize/2 pixel 以内の点の分散値の和を用いる.
4. 格子点の存在するすべての水平ラインについて上記の処理を繰り返す.

このようにして求めた三角形パッチの形状は、物体の形状に近いものとなる。

また、貼りつけるテクスチャとしては、全ての視点からのSNRを評価値として用い、定められたモデルの表面に沿った $Avg(X, Y, Z)$ の値 (平均のテクスチャ) を用いる。テクスチャはいずれのモデルでも視差画像約1枚分の情報量である。

6.6.3 処理例1

原画像の大きさは、犬の画像が 256×256 pixel, フラワーガーデンが 352×240 pixel である。評価値空間の大きさは、縦方向は原画像と同じ、横方向は原画像に左右40 pixel ずつ拡大したものとし、奥行きZ方向については $Z = \pm 20$ 内の空間を80段階に量子化している。

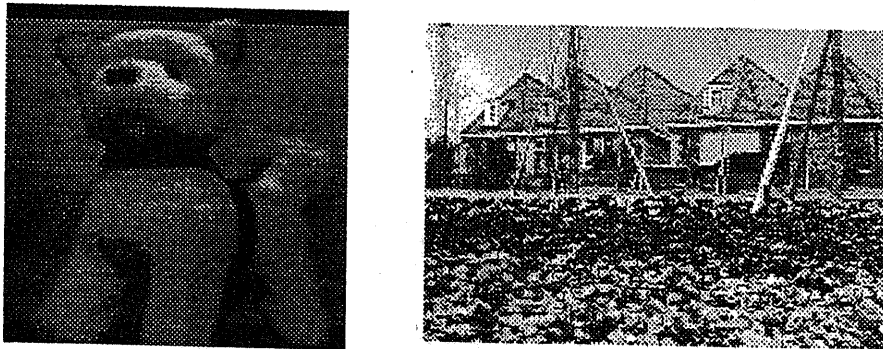


図 6.5: 原画像 (中央画像: $m=10$). Toy dog(左), Flower Garden (右).

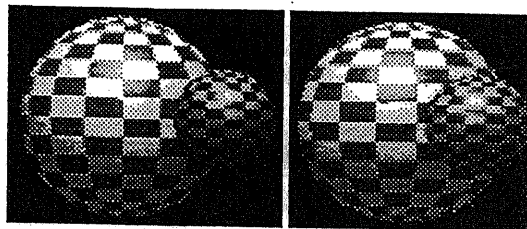


図 6.6: 原画像 "eluminated spheres" (中央画像: $m=9$, 右からの画像: $m=18$)

6.6.4 処理例2

ここで興味深いのは、光源を設定した球の場合である。ここではランベルト表面を仮定して分析を行なっているのであるが、その制約の範囲内で光の当たった部分を再生するような形状とテクスチャが解として出ている。

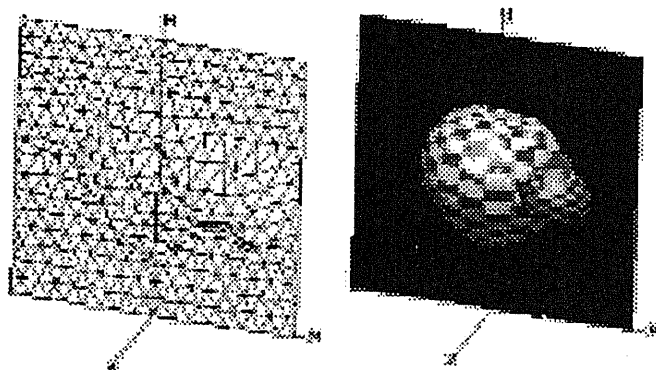


図 6.7: “eluminated spheres” の実験結果, 圧縮データの構造, テクスチャを貼り付けたもの.

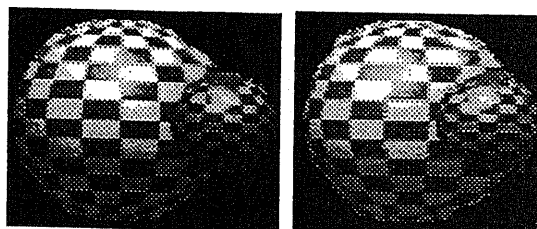


図 6.8: 復号画像 “eluminated spheres”(中央画像: $m=9$, 右からの画像: $m=18$)

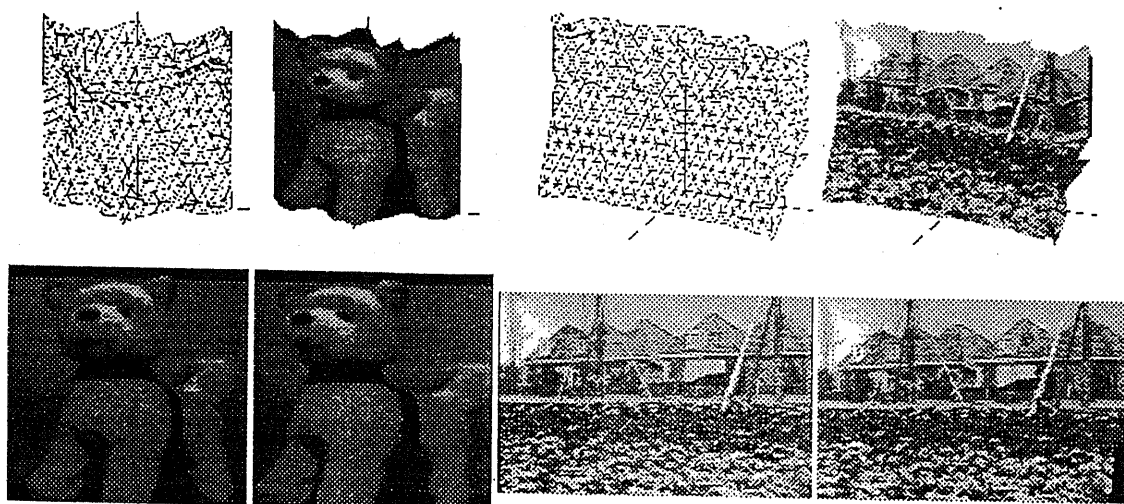


図 6.9: “Toy Dog” と “Flower Garden” の実験結果, 圧縮データの構造, テクスチャを貼り付けたもの, および復元画像 ($m=10, 19$).

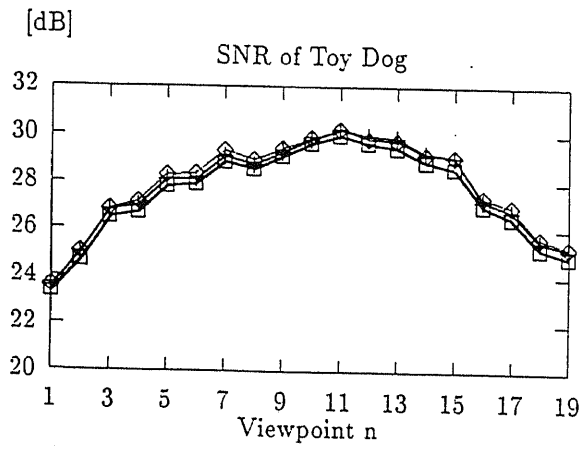


図 6.10: 復号画像の SNR, “Toy Dog”. メッシュサイズは上から 4, 8, 16 pixel.

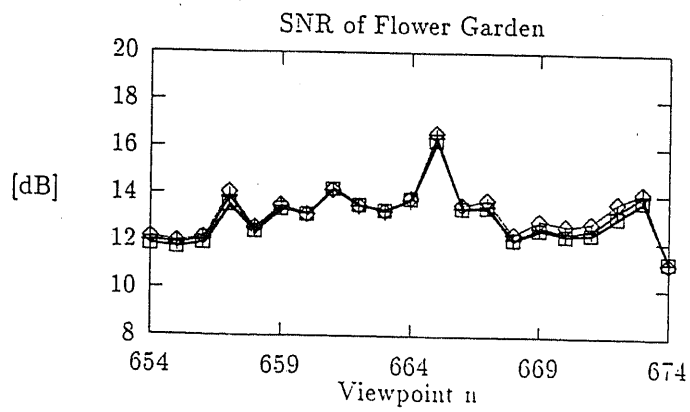


図 6.11: 復号画像の SNR “Flower Garden”

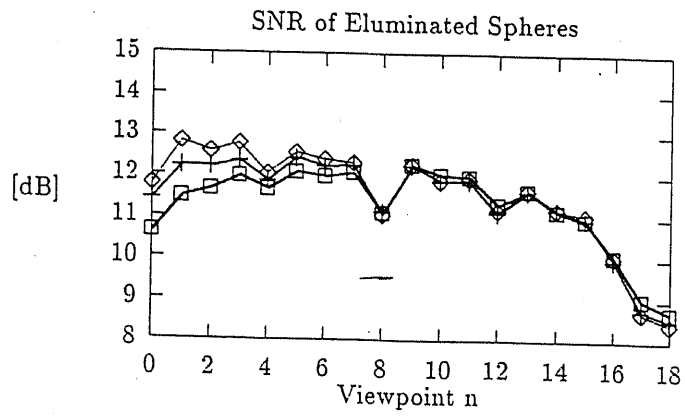


図 6.12: 復号画像の SNR “Eluminated spheres”

6.7 実験2

符号量と予測特性との関係や、評価値の設定の仕方による違いについて詳細に検討するため、さらにCGモデルを用いて実験を行なった。

実験に用いた画像は、半径10cmの球を100cm離れた位置から3cm間隔で縦横方向に5枚ずつ、合計25枚撮影したことに相当する画像(256×256 pixel)をCGにより合成したものである。 m, n の範囲は-2から2までであり、 $M = N = 5$ である。投影法としては、前章で述べたように物体座標系の原点がカメラ座標系の原点に写されるとし、被写体空間のX-Y面と投影面とが平行な中心投影を仮定した。合成した視差画像の例を図6.13に示す。

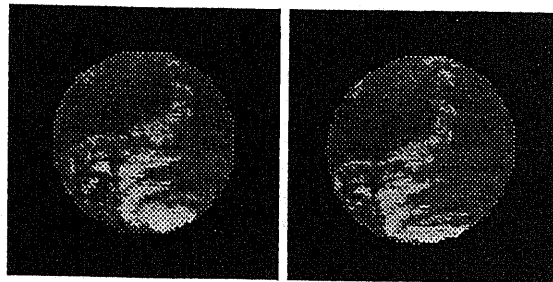


図 6.13: 視差画像合成に用いたモデルと視差画像の例 ($(m, n) = (0, 0)$ と $(2, 2)$)

このようにして得られた25枚の視差画像を、前章で述べたアルゴリズムにしたがってモデルの形状と表面テクスチャとに圧縮する。評価値空間の大きさは、縦方向、横方向ともに原画像と同じとし、奥行き Z' 方向については $Z' = \pm 20$ 内の空間を80段階に量子化している。

6.7.1 モデル記述に要する付加情報量と予測特性の検討

3次元構造モデルとして、次の2種類について実験を行なった。

1. 平面モデル
2. 三角形パッチモデル

1番目として Z' 軸に垂直な平面モデルで実験を行なった。パラメータの最適化は、平面モデルの Z' の値を全探索により求めた。規格化空間は Z' 方向に80段階に量子化しており、 Z' 軸方向の情報量として1 byte 割り当てている。

次に、形状表現に汎用性があるものとして、三角形パッチモデルを用いた実験を行なった。三角形パッチモデルの形状は二等辺三角形とし、大きさによって付加情報量が決まる。三角形パッチモデルの形状の変更は、ノードの奥行き Z' 方向のみを移動させることにした。視域内の任意の視点から見た場合にオクルージョンはないという制約条件を設けているので、三角形パッチの面の向きに対して視域に依存した制約条件を課すことが必要である。三角形パッチの面の法線ベクトルを (x_n, y_n, z_n) とおくと、制約条件は、

$$1/m_{min} < z_n/x_n < 1/m_{max} \quad (6.16)$$

$$1/n_{min} < z_n/y_n < 1/n_{max} \quad (6.17)$$

$$\text{ただし, } m_{min}, n_{min} < 0 \quad m_{max}, n_{max} > 0$$

となる。このような制約条件のもとでノードの位置を最適化する。付加情報は各ノードごとに1 byte 割り当てるものとする。また、この実験における評価値としては、すべての視差画像の平均のSNRとし、

$$a_{m,n} = \frac{1}{MN} \quad (6.18)$$

と設定した。

図6.14に付加情報量1 byte, 1105 byte の場合について、求められた形状とテクスチャを示す。また、図6.15に付加情報量と復号視差画像の平均のSNR とのグラフを示す。

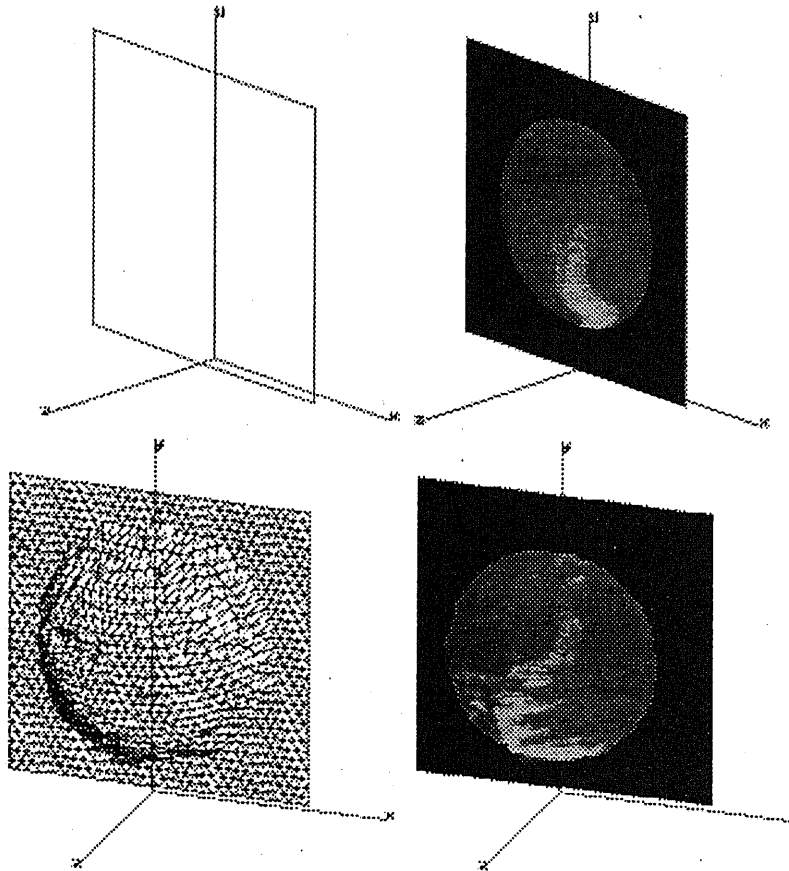


図 6.14: 求められた形状とテクスチャ。付加情報量 1byte(上), 1105 byte(下)。

6.7.2 評価値の設定を変化させた場合

異なった評価値のもとでは、解として求められるモデルの形状、表面テクスチャが違ったものとなることを確かめるため、評価値を変えた実験をおこなった。3次元構造モデルは評価値によらず同じとし、底辺、高さがともに8 pixelの三角形パッチとした。したがって付加情報量は1105 byteとなる。評価値として以下の2種類について実験を行なった。

- すべての視差画像の平均のSNRを評価値とした場合
- 中央画像のSNRだけを評価値とした場合

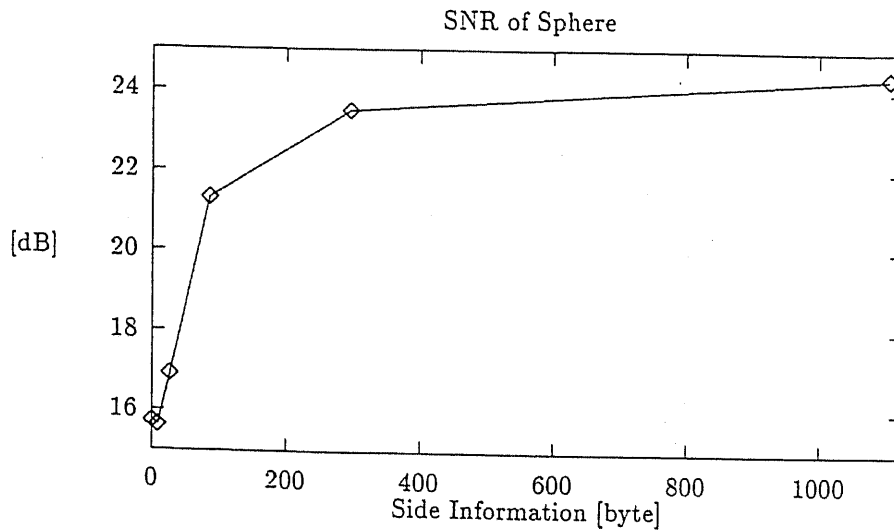


図 6.15: モデルパラメータのデータ量に対する復号視差画像の平均 SNR

それぞれの評価値のもとでの重み係数 $a_{m,n}$ は次の通りである。

Criterion 1:

$$a_{m,n} = \frac{1}{MN} \quad (6.19)$$

Criterion 2:

$$a_{m,n} = \begin{cases} 1 & m = n = 0 \\ 0 & \text{otherwise} \end{cases} \quad (6.20)$$

形状記述のための付加情報はいずれの評価値の場合でも同じであり、1105 byte とした。それぞれの評価値のもとで得られたモデルの形状を図 6.14, 6.16 に示す。図 6.14 の下の図が本実験においてすべての視差画像の平均の SNR を評価値とした場合に相当し、図 6.16 が中央画像の SNR だけを評価値とした場合に得られた形状とテクスチャである。また、復号視差画像の SNR のグラフを図 6.17, 6.18 に示す。

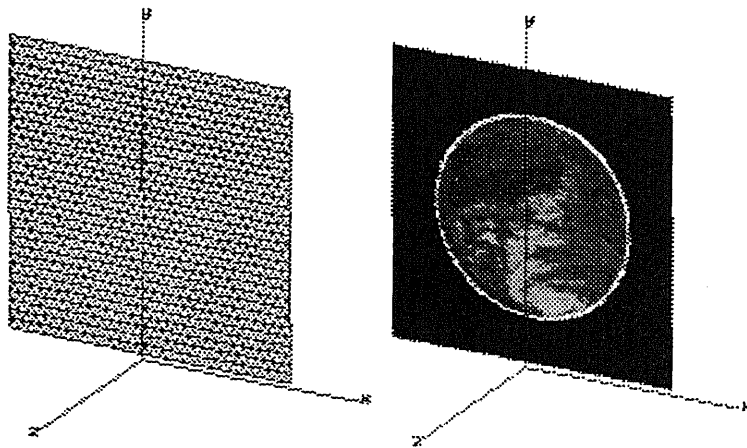


図 6.16: 中央の視点からの視差画像の SNR を評価値とした場合に得られた形状とテクスチャ

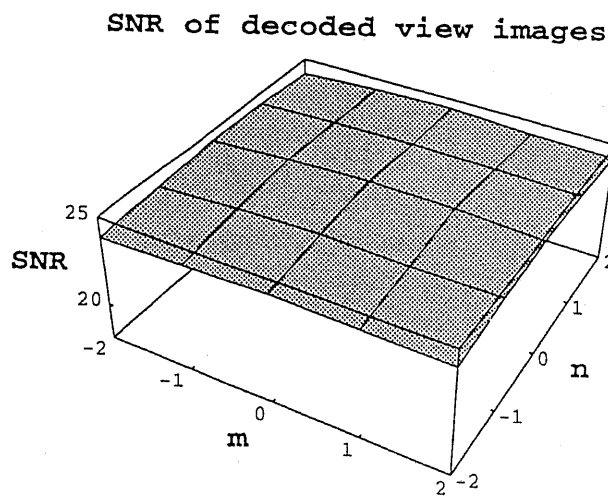


図 6.17: すべての視差画像の平均の SNR を評価値とした場合の復号視差画像の SNR

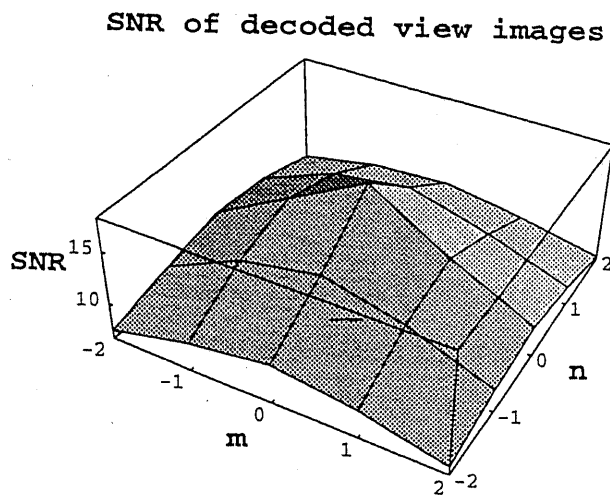


図 6.18: 中央の視点からの視差画像の SNR を評価値とした場合の復号視差画像の SNR

6.8 評価値空間のセグメンテーションによる予測特性の向上

この場合の解決法として、まずすべての視点から見えている領域を抽出、符号化し、その領域を原画像から取り除いた残りの画像に対して同様の手法を繰り返すことにより、手前にある物体から符号化していくこととなる。この例を図6.19に示す。この手法の場合、本稿におけるモデルを2層、3層にしていく場合に相当する。モデルを複雑にしていくことでモデルパラメータ伝送に要する情報量が増加し、2次元画像符号化方式との両立性に乏しくなる半面、予測特性が向上し、かつ任意視点からの合成画像の品質が向上する [6, 23]。

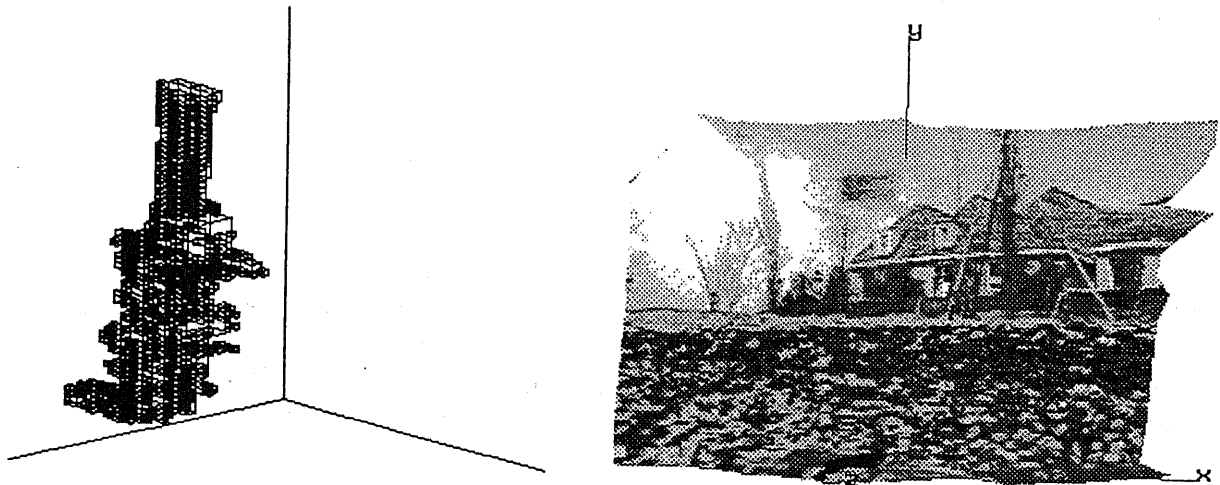


図 6.19: 評価値空間の領域分割による前景と背景の分離

6.9 テクスチャと予測誤差の符号化

レンズ素子は、 θ 方向の輝度値の分散の小さい領域に集めているので θ 方向に対しては低周波数成分に片寄った信号となっているはずである。これを2次元画像符号化で用いられているDCT（離散コサイン変換）により圧縮・符号化を符号化することを検討していく。

図6.21は前節までのようにして形状関数を定めることにより、表示すべきデータ $((m, n)$ 画像) がどのように変換されるかを表した図である。特に形状関数が直線 $Z = 0$ の時にはもとの (m, n) 画像と一致する。この変換された (m, n) 画像に対して2次元DCTを施して圧縮する。変換された (m, n) 画像は、 m 方向の性質は通常の2次元画像と同じであるが、 n 方向の相関が非常に高く、ほとんど直流成分に電力が集中している。DCT係数の選び方として、 n' 方向に対しては低周波数成分へ電力が集中していることを積極的に利用する。この様子を分かりやすく説明するため、図6.22に水平断面で説明する。左上から変換前の (m, n) 画像、変換後の (m', n') 画像、 n' 方向に対してDCTの直流成分のみを用いて復号したもの、逆変換して復号したものを表している。3次元ブロックに対し3次元DCTを施す処理を画像全体に対し実行すると、計算時間は数十分程かかる [13]。

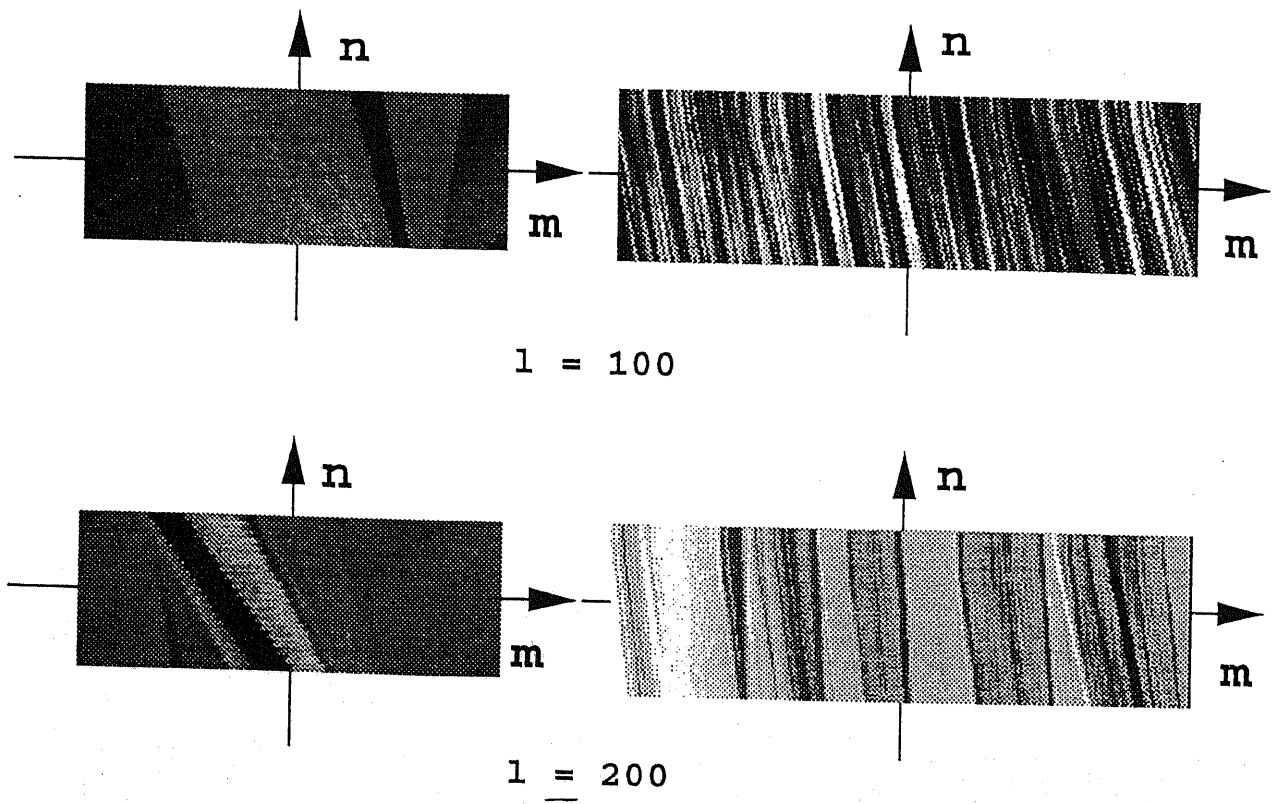


図 6.20: mn 画像の例 “Toy Dog” (左), and “Flower Garden”(右).

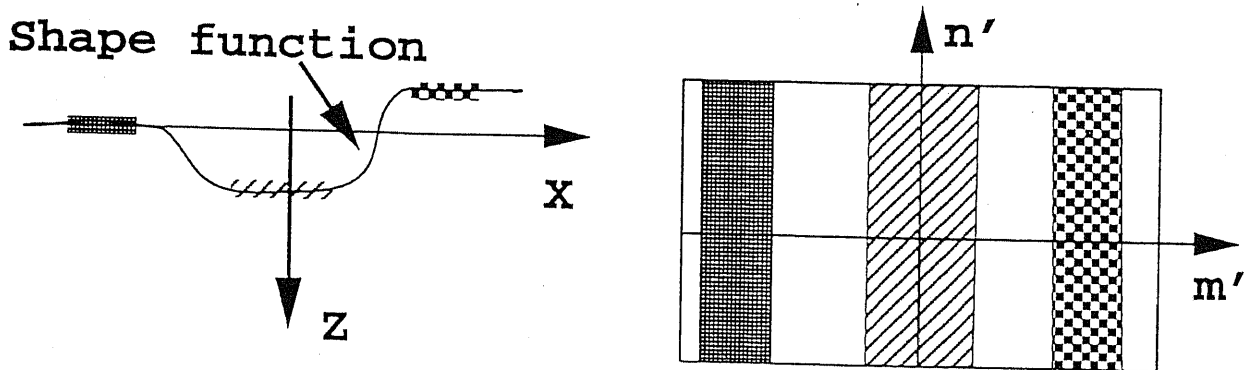


図 6.21: レンズ列の再配置による新しい (m', n') 画像の生成

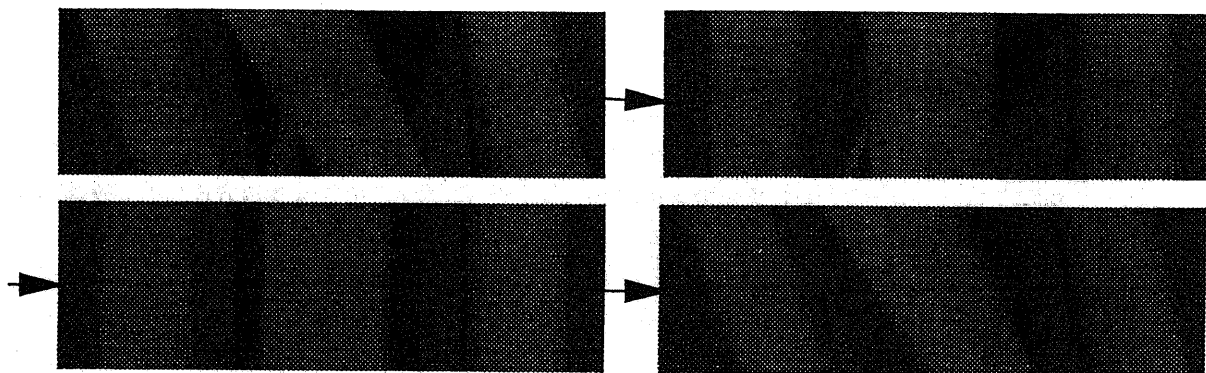


図 6.22: 変換前後の (m,n) 画像の DCT 符号化結果

6.10 まとめ

多眼画像の情報圧縮, 視域内の任意視点からの画像の合成を目指し, 3次元構造モデルによる3次元予測符号化を提案, 実験的検討を行なった. 本手法は, 被写体の正確な3次元構造を求めることを目的とするのではなく, なるべく少ない情報量で高品質に多眼画像の情報を記述しようとする試みである. 4次元の空間で表される「視差画像空間」と被写体空間を規格化した「規格化空間」との写像関係を利用し, 復号視差画像のSNRを評価値とした最適化問題として3次元構造モデルの形状, および表面テクスチャを決定した.

本論文においては, 被写体の正確な構造を求めることが目的ではなく, 3次元構造モデルを用いて多眼画像を圧縮符号化することが目的である. したがって, Shape from Xに代表されるビジョン的アプローチでは考慮されていない, 以下のような点に留意する必要がある.

第1に, 全体としての情報量が最も小さくなるよう, 情報量と符号化特性との関係の中で3次元構造を求めていくことが必要となる. 実験1では, 三角形パッチモデルを用いた場合について予測特性と付加情報量の関係を調べた. 実験の結果, 付加情報量を増やしていくに従って予測特性は向上するが, ある一定値以上に付加情報量を増やしても予測特性の向上は飽和する傾向が確認された. 付加情報量を増やしていくにしたがって被写体の細かな構造を表現することができるようになり, 予測特性が向上するものと考えられる. しかしながら, 3次元構造モデルとして, ランベルト表面およびオクルージョンがないと仮定したことにより, このモデルの制約の中で付加情報量を増やしていても, あるところで飽和すると考えられる. これに対処する一つの方法としては, 今回の実験で用いたモデルを情景内に複数個用意し, より複雑なモデルを用いることにより予測特性を上げるというアプローチが考えられる. 一方, 2次元動画像標準符号化方式の考え方からすれば, 「フレーム内符号化」を行なうか, もしくは予測誤差を伝送することによりSNRの向上をはかることになる. いずれの場合でも付加情報の記述に要するデータ量の増加と復号画像の品質の向上とのトレードオフのもとで決定する必要がある. また画像を構造化しオクルージョンを抽出することによる予測特性の向上を確かめる実験例について述べた.

次に, 符号化という観点からは復号画像の品質が評価値であるという点が重要である. ここでは復号画像のSNRを評価規範と定め, これを最大にするような形状, およびテクスチャを決定する手法を提案した. ここでは多眼画像が符号化対象であるため, 非常に柔軟な評価値の設定の仕方が可能である. 実験2では, 視差画像ごとに評価値の重みが異なる場合についての実験を行なった. この結果は, 3次元ディスプレイの眼数などに応じた柔軟な符号化方式が可能であることを示しており, さらに, 3次元符号化方式の1つの特殊例として2次元画像符号化方式を位置付けることが可能であることを示している.

また, 視域内の任意視点からの画像の合成を行なうためには, 被写体の構造が正確に復元されている必要がある. 今回の実験で用いたような制約を設けたモデルを用いた場合には, 任意視点からの画像合成においては正確さが失われてしまうという問題がある. 何枚の視差画像があれば被写体の構造が正確に復元できるかについては被写体の複雑さに依存するため, 一般的な数値を述べることは困難であり, 今後の課題として残されている.

今後の課題としては, まず第1番目に, ささまざまなモデルに対して, モデルの記述に割り当てる情報量と予測特性との関係をさらに詳細に検討することである. 本章における実験結果をみると, モデルの記述に割り当てる情報量と予測特性とは互いにトレードオフの関係になっていることが分かる. これは直観的にもうなずける結果となっている. この関係を模式的に示したものが図6.23である. ユーザはこれらの中からビットレート, 品質が平衡点となるように選択できるような符号化方式を模索していくことが今後の重要な課題であろう.

2番目として、時間軸方向への統合をおこなうことである。時間軸方向への統合として、現在、考えているのは次の2つの方法である。

- 伝送した形状情報+テクスチャ情報を用いて形状の変形情報を伝送する。適宜テクスチャ更新情報を伝送する。
- テクスチャ情報は2次元動画画像符号化方式により伝送する。付加情報としてノードの奥行きのみを伝送する。

モデルベース符号化のセンスからは1番目の方式が望まれる。しかし、本方式は次節で述べるように2次元動画画像符号化方式との両立性を備えることが可能であるので、情報量は多くなっても安定な2番目の方式も有望であると考えられる。

本章で述べた3次元構造モデルによる「光線空間のモデル化」は非常に有望な手法であって、今後十分な検討を重ねていくことにより「3次元統合画像符号化」における中心的な手法となり得ると思われる。今後のさらなる研究の発展が期待される。

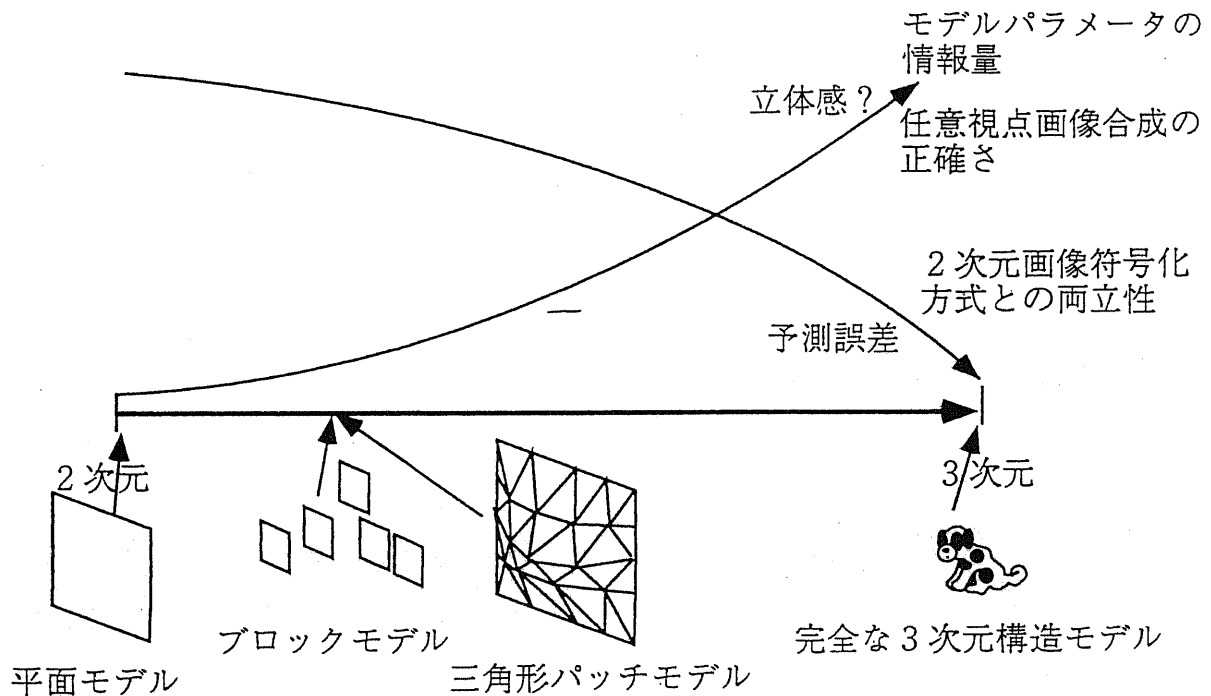


図 6.23: 構造モデルと情報量, 予測特性との関係の模式図

付録

評価値の空間は,

$$E(x, y, z) = \sum_{i=-N}^N a_i \{f_i(x, y, z) - T(x, y, z)\}^2 \quad (6.21)$$

と定義された。これを最小にするような $T(x, y, z)$ を求める。 E を T で偏微分すると、

$$\frac{\partial E(x, y, z)}{\partial T(x, y, z)} = \sum_{i=-N}^N 2a_i \{f_i(x, y, z) - T(x, y, z)\}(-1)$$

となり、これを 0 とおくことにより、

$$\begin{aligned} \sum_{i=-N}^N a_i f_i(x, y, z) &= T(x, y, z) \sum_{i=-N}^N f_i(x, y, z) \\ T(x, y, z) &= \frac{\sum_{i=-N}^N a_i f_i(x, y, z)}{\sum_{i=-N}^N f_i(x, y, z)} \\ T(x, y, z) &= \sum_{i=-N}^N a_i f_i(x, y, z) \end{aligned} \quad (6.22)$$

を得る。

参考文献

- [1] 原島, 藤井, 苗村: “3次元統合画像通信の構想”, 信学技報, HC92-58 (Jan. 1993).
大田, 正井, 池田: “動的計画法によるステレオ画像の区間対応法”, 信学論, Vol. J68-D, No. 4, pp. 554-561 (Apr. 1985).
- [2] Y. Ohta, T. Kanade, “Stereo by Intra- and Inter-Scanline Search Using Dynamic Programming”, *IEEE Trans on Pattern Anal. and Machine Intell.*, Vol. PAMI-7, No. 2, Mar. 1985.
大田, 山田: “動的計画法によるパターンマッチング”, 情報処理, Vol. 30, No. 9, pp. 1058-1066 (Sep. 1989).
渡辺, 大田: “複数のステレオ対応探索アルゴリズムの協調的統合”, 信学論 D- , Vol. J73-D- , No. 10, pp. 1666-1676 (Oct. 1990).
- [3] K.E. Price, “Relaxation Matching Technique - A Comparison”, *IEEE Trans. on Pattern Anal. Machine Intell.*, Vol. PAMI-7, No. 5, Sep. 1985.
- [4] Yamamoto, “Motion Analysis by Visualized Locus Method”, *Transactions of Information Processing Society of Japan*, vol. 22, no. 5, pp. 442-449 (Sep. 1981).
- [5] 山本: “連続ステレオ画像からの3次元情報の抽出”, 信学論, Vol. J69-D No. 11, pp. 1631-1638 (1986).
- [6] R. C. Bolles, H. H. Baker, and D. H. Marimont, “Epipolar-Plane Image Analysis: An Approach to Determining Structure from Motion”, *International Journal of Computer Vision*, 1, pp. 7-55 (1987).
- [7] H.H. Baker, R.C. Bolles, “Generalizing Epipolar-Plane Image Analysis on the Spatiotemporal Surface”, *International Journal of Computer Vision*, 3, pp. 33-49 (1989).
- [8] 蓼沼, 湯山: “立体画像の左右対応点検索法の最適化”, 信学技報 IE93-69 (1993-11).
- [9] W.Richards: “Structure from stereo and motion”, *J. Opt. Soc. Amer. A.*, Vol. 2, No. 2, February 1985, pp.343-349.
- [10] U.R. Dhond, J.K. Aggarwal: “Structure from Stereo - A Review”, *IEEE Transactions on Systems, Man, Cybernetics*, Vol. 19, No. 6, pp. 1489-1510 (Nov./Dec. 1989).
- [11] S.T. Barnard and W.B. Thompson: “Disparity analysis of images,” *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. PAMI-2, pp. 333-340 (July 1980).

- J. Liu and R. Skerjanc, "Stereo and motion correspondence in a sequence of stereo images", *Signal Processing:Image Communication* 5, pp. 305-318 (1993).
- R. Skerjanc and J. Liu, "A three camera approach for calculating disparity and synthesizing intermediate pictures", *Signal Processing:Image Communication* 4, No. 1, pp. 55-64 (1991).
- [12] R. Thoma and M. Bierling: "Motion compensating interpolation considering covered and uncovered background", *Signal Procossing: Image Communication*, Vol. 1, No. 2, pp. 191-212(Oct. 1989).
- [13] 藤井: "3次元画像からの視差検出 - 3次元形状の推定と帯域圧縮-", 3次元映像, Vol. 8, No. 1 (1994).
- [14] K. Satoh and Y. Ohta, "Passive Depth Acquisition for 3D Image Displays", *IEICE Trans. Inf. and Syst.*, Vol. E77-D, No. 9, pp. 949-957 (Sep. 1994).
- [15] U.R. Dhond, J.K. Aggarwal: "Structure from Stereo - A Review", *IEEE Transactions on Systems, Man, Cybernetics*, Vol. 19, No. 6, pp. 1489-1510 (Nov./Dec. 1989).
- [16] Jin Liu, Robert Skerjanc, "Stereo and motion correspondence in a sequence of stereo images", *Signal Processing: Image Communication*, vol. 5, no. 4, pp. 305-318 (Oct. 1993).
- [17] 田村, 北村: "3次元仮想空間生成・表示へのモデル系&非モデル系アプローチ", 信学技報HC研究会/SICE HI 研究会資料 (May 1995).
- [18] 高橋, 本田他: "ホログラフィックステレオグラムにおける視差画像の補間", 光学, Vol. 22, No. 11(Nov. 1993).
- [19] 金出, コンラッド, ポールマン, 森田: "因子分解法による物体形状とカメラ運動の復元", 信学論 D-II, Vol. J76-D-II, No. 8, pp. 1497 - 1505 (Aug. 1993).
- [20] 緒方, 佐藤: "運動視のモデル-両眼立体視モデルとの統合をめざして-", テレビ誌, Vol. 45, No. 3, pp. 299-305 (1991).
- [21] 藤井, 三宅: "両眼視差抽出と3次元表面の復元", テレビ誌 Vol. 45, No. 3, pp. 315-322 (1991).
- [22] 原島, 森川, 青木: "三次元構造モデルを用いる画像の符号化・処理・表示", 1989 信学会春季全大, SD-3-12 (1989).
- [23] 森川, 原島: "3次元構造・運動情報に基づく動画像符号化". 信学論, D-II, Vol. J73-D-II, No. 7, pp. 982-991 (1990年7月).
- [24] 藤井, 原島, 濱崎: "三次元画像符号化の基礎検討(3) - 多眼画像からの構造抽出-", 1993年信学会春季全大, D-328 (1993-3).
- [25] 苗村, 吉岡, 藤井, 中屋, 原島: "多眼画像の視差補償予測符号化". 1993 TV 学会年次大会, 24-4 (1993-7).

- [26] 藤井, 原島: “アフィン変換による多眼画像の視差補償予測符号化”, 1993 年信学会秋季全大, D-200 (1993-9).
- [27] 藤井, 原島: “多眼式三次元画像符号化 - 視差補償予測符号化から構造抽出符号化へ -”, *PCSJ '93*, 7-3 (1993-10).
- [28] 藤井, 原島: “アフィン変換を用いた多眼式三次元画像符号化”, *SITA '93*, T11-1 (1993-10).
- [29] 嵯峨田, 藤井, 原島: “多眼画像の視差補償予測符号化-物体と背景の分離に基づく視差ベクトルの抽出と補間-”, 1994 年信学会秋期全大, D-325 (1994-3).
- [30] 藤井, 原島: “多眼画像の視差補償予測符号化-セグメンテーションに基づく予測特性の改善-”, 1994 年信学会秋期全大, D-326 (1994-3).
- [31] 藤井, 原島: “3次元構造推定に基づく多眼3次元画像の圧縮と補間”, 3次元画像コンファレンス'94, pp. 25-31 (Jul. 1994).
- [32] 藤井, 原島: “構造抽出に基づく三次元画像符号化” *PCSJ94* (1994-10).
- [33] T.Fujii, H.Harashima, J.Hamasaki: “Data Compression for an Autostereoscopic 3-D Image”, *PCS '93*, 13.21 (Mar. 1993).
- [34] T. Fujii, H. Harashima, “Data compression of an autostereoscopic 3-D image”, *SPIE Proc. Vol. 2177 Stereoscopic Displays and Applications V* (San Jose, CA, Feb. 1994).
- [35] T. Fujii, H. Harashima, “3-D Image Coding Based on Affine Transform”, *ICASSP '94*, 81.8, Adelaide (1994-4).
- [36] T. Fujii, H. Harashima, “Data Compression and Interpolation of Multi-View Image Set”, *IEICE Trans. on Inf. and Sys.*, Vol. E77-D, No. 9, pp. 987-995 (Sep. 1994).
- [37] T. Fujii, H. Harashima, “Data Compression and Interpolation for an Autostereoscopic 3-D Image”, *PCS '94* (Sep. 1994).
- [38] T. Fujii, H. Harashima, “Coding of autostereoscopic 3D image sequence”, *VCIP '94*, 2308-91 (Chicago, Sep. 1994).
- [39] 安居院猛, 長尾智晴, 共著: 「画像の処理と認識」, 昭晃堂, 1992.
- [40] R. Nevatia 著, 南敏 監訳: 「画像認識と画像理解」, 啓学出版, 1986.
- [41] 原島 博 監修: 「画像情報圧縮」, オーム社 (1991).
- [42] 谷内田正彦: 「ロボットビジョン」, 昭晃堂.

第7章

ホログラムパターンの圧縮, 符号化

ホログラフィの伝送の試みは, 1965年に, ベル研究所の Enloe らによりなされたのが最初であろう [5]. 当時, ホログラムのもつ情報量の多さからホログラム伝送の実現は当分見込みがないとされた. 近年, コンピュータの高性能化, メモリの大容量化, デジタル信号処理技術の発展に支えられ, 動画ホログラフィの実験例が報告されるようになってきており, 再びホログラムパターンの圧縮の必要性が高まってきたと言えるだろう.

本章では, 光線空間の観点からのホログラムパターンの圧縮, 符号化方式について提案を行ない, 基礎的な実験について述べる. 光線空間の概念をホログラムパターンに適用することで, 3次元統合画像通信の枠組でホログラムパターン圧縮を論ずることを目的とする.

7.1 はじめに

ホログラフィは理想的な3次元ディスプレイとして注目されており, ホログラフィテレビジョンはTV技術者の間のみならず人類の夢の技術であると言えよう。

ホログラフィの伝送の試みは, 1965年に, ベル研究所の Enloe らによりなされたのが最初であろう [5]. 当時, ホログラムのもつ情報量の多さからホログラム伝送の実現は当分見込みがないとされた. その後, 光学系における圧縮の試みがなされてきたものの, 決定的な解決をみるには至っていない。

近年, コンピュータの高性能化, メモリの大容量化, デジタル信号処理技術の発展に支えられ, 動画ホログラフィの実験例が報告されるようになってきており, 再びホログラムパターンの圧縮の必要性が高まってきたと同時にホログラムパターンの圧縮の試みが行なえる状況になってきた。

本章ではホログラムパターンの圧縮, 符号化についての検討を行なう。まず, 光の伝搬理論について述べ, ホログラムの生成, 再生の原理について述べる。また実際に離散系でシミュレーションを行なうときに必要となるサンプリング間隔などのパラメータについて考察し, 高速計算方法などについても言及する。

ホログラフィ技術の歴史と現状を概観した後に, ホログラムパターンの圧縮, 符号化の研究の現状を述べる。最後に, 本論文における提案である, 光線空間とホログラムパターンとの相互変換に基づくホログラムパターンの圧縮, 符号化の提案を行ない, 基礎的なシミュレーション結果について述べる。

7.2 光の回折の基礎理論

7.2.1 回折の近似式

ここでは, 以下の式から出発することにする。

$$\frac{1}{j\lambda} \int \int_{-\infty}^{\infty} g(x_o, y_o) \frac{\exp(jkr)}{r} dx_o dy_o \quad (7.1)$$

さて, 直交座標系において, 開口面上 ($z=0$) の入力光分布 $g(x_o, y_o)$ が衝立面上 ($z=z_i$) に作る回折像 $u(x_i, y_i)$ を求めるとすると,

$$r = \sqrt{z^2 + (x_o - x_i)^2 + (y_o - y_i)^2} \quad (7.2)$$

となる。これをそのまま積分できるケースは非常に少ないので, 以下のように近似をする。

$$\sqrt{z^2 + (x_o - x_i)^2 + (y_o - y_i)^2} = z \left(1 + \frac{(x_o - x_i)^2 + (y_o - y_i)^2}{2z^2} \right) \quad (7.3)$$

また, 分母の値も $r = z_i$ と近似する。これにより, 次のフレネル近似式を得る。

$$u(x_i, y_i) = \frac{1}{j\lambda z_i} \exp(jkz_i) \int \int_{-\infty}^{\infty} g(x_o, y_o) \exp(jk \frac{(x_o - x_i)^2 + (y_o - y_i)^2}{2z_i}) dx_o dy_o \quad (7.4)$$

また z_i が十分大きい場合には, さらに近似ができて,

$$u(x_i, y_i) = \frac{1}{j\lambda z_i} \exp[jk\{z_i + \frac{(x_i^2 + y_i^2)}{2z_i}\}] \int \int_{-\infty}^{\infty} g(x_o, y_o) \exp\left(-j2\pi\left(\frac{x_i x_o}{\lambda z_i} + \frac{y_i y_o}{\lambda z_i}\right)\right) dx_o dy_o \quad (7.5)$$

となる。これをフラウンホーファ回折という。フレネル近似を使うかフラウンホーファ近似を使うかを決定するのは、一般に $k(x_o^2 + y_o^2)/2z_i$ の値が $\pi/2$ より大きい小さいかによって決める。

7.2.2 回折の計算法

次に、フレネル回折の計算法について述べる。以下の3つを考える。

1. 数値計算
2. フーリエ変換と畳み込みとの関係を利用
3. フーリエ変換の利用

数値計算

数値積分をする際に問題となるのは x_o の刻み幅である。被積分関数中の \exp による位相変化分が $\pi/2$ を越えないように刻み幅を設定する必要がある。 $x_o + \Delta x_o$ とした場合、 \exp の項の位相の変化分が $\pi/2$ を越えないためには

$$k \frac{2(x_o - x_i)\Delta x_o + \Delta x_o^2}{2z_i} < \frac{\pi}{2} \quad (7.6)$$

でなくてはならない。この不等式と $\Delta x_o > 0$ とから、

$$0 < \Delta x_o < -(x_o - x_i) + \sqrt{(x_o - x_i)^2 + \frac{\pi z_i}{k}} \quad (7.7)$$

を得る。右辺の値は $(x_o - x_i)$ に対して単調減少であるから、 $(x_o - x_i)$ が最大となった時が Δx_o の最大値を規定する。数値例として、

$$\begin{aligned} (x_o - x_i) \text{ の最大値} &= 200\text{mm} \\ \text{波長} \lambda &= 632\mu\text{m} \\ z_i &= 500\text{mm} \end{aligned}$$

とすると、

$$0 < \Delta x_o < 395\mu\text{m} \quad (7.8)$$

となる。

フーリエ変換と畳み込みの関係を利用

(7.4) 式は畳み込みの記号を用いて次のように表すことができる。

$$u(x_i, y_i) = g(x_i, y_i) f_{z_i}(x_i, y_i) \quad (7.9)$$

ただし、

$$f_{z_i}(x_i, y_i) = \frac{1}{j\lambda z_i} \exp(jk(z_i + (x_i^2 + y_i^2)/2z_i))$$

フーリエ変換の性質により、

$$F[u(x_i, y_i)] = F[g(x_i, y_i)] \cdot F[f_{z_i}] \quad (7.10)$$

となる。ここで $F[\cdot]$ はフーリエ変換を表す。 f_{z_i} のフーリエ変換は分かっており、

$$F[f_{z_i}] = \exp(jkz_i) \exp(-j\pi\lambda z_i(f_x^2 + f_y^2)) \quad (7.11)$$

と与えられる（付録参照）。したがって $g(x_i, y_i)$ をフーリエ変換した後 $f[f_{z_i}]$ をかけ、それに逆フーリエ変換を施すことにより $u(x_i, y_i)$ が得られる。なお、

$$\begin{aligned} f_x &= \frac{x_i}{\lambda z_i} \\ f_y &= \frac{y_i}{\lambda z_i} \end{aligned} \quad (7.12)$$

である。

フーリエ変換の利用

(7.4) 式を変形し、

$$u(x_i, y_i) = \frac{1}{j\lambda z_i} \exp[jk\{z_i + \frac{(x_i^2 + y_i^2)}{2z_i}\}] \int \int_{-\infty}^{\infty} g(x_o, y_o) \exp(jk \frac{x_o^2 + y_o^2}{2z_i}) \exp\left(-j2\pi\left(\frac{x_i x_o}{\lambda z_i} + \frac{y_i y_o}{\lambda z_i}\right)\right) dx_o dy_o \quad (7.13)$$

とすると、これはフーリエ変換の式になっていることに気づく。すなわち、

$$g(x_o, y_o) \exp(jk \frac{x_o^2 + y_o^2}{2z_i}) \quad (7.14)$$

を計算し、これをフーリエ変換することにより $u(x_i, y_i)$ が得られる。

FFT による光伝搬の計算

7.2.2節で述べたフーリエ変換の利用についての離散的な計算方法について述べる。まず、

$$g(x_o, y_o) \exp(jk \frac{x_o^2 + y_o^2}{2z_i}) \quad (7.15)$$

を、 Δx_o の間隔で N 点サンプリングをする。この際、 x_o の変域としては、

$$-\frac{N\Delta x_o}{2} \leq x_o \leq \frac{N\Delta x_o}{2} \quad (7.16)$$

となる。サンプリング間隔が Δx_o であることから、 ω_i の変域としては、

$$-\frac{2\pi}{2\Delta x_o} \leq \omega_i \leq \frac{2\pi}{2\Delta x_o} \quad (7.17)$$

であり、 x_i で考えるならば、

$$-\frac{\lambda z_i}{2\Delta x_o} \leq x_i \leq \frac{\lambda z_i}{2\Delta x_o} \quad (7.18)$$

となる。この範囲を N 点でサンプリングしていることから、像面におけるサンプリング間隔は、

$$\frac{\lambda z_i}{N\Delta x_o} \quad (7.19)$$

と与えられる。

数値例：開口面におけるサンプリング間隔を 250nm , 物体面での開口の大きさを 10mm , 物体面 - 像面の距離を 50cm とする. $N = 10\text{mm}/250\text{nm} = 40000$, $\Delta x_o = 2.5 \times 10^{-7}$, $z_i = 0.5$, $\lambda = 632 \times 10^{-9}$ を代入することにより, 像面におけるサンプリング間隔は,

$$\frac{\lambda z_i}{N \Delta x_o} = 31.6 \mu\text{m}$$

となり, 像面の大きさは,

$$\frac{\lambda z_i}{\Delta x_o} = 63.2\text{cm}$$

となる.

7.2.3 ホログラフィの原理

ホログラフィの基本的な発明は1948年にガボール (Gabor) によりなされた. 電子顕微鏡の改良中に着想した方法であり, ガボールはこの功績により後にノーベル物理学賞を授賞したが, 3次元ディスプレイとしてとらえた場合, 物体が透明でなければならないこと, 真の回折像と共役像が重なり2重像になってしまうという欠点があった. 1962年には米国ミシガン大学のリース (Leith) とウパトニクス (Upatnieks) により2光束干渉法が考案され [2], 1960年のレーザの発明とあいまって, 2重像の分離に成功, 格段に良質な画像の記録, 再生に成功した. これ以降ホログラフィは, 目の焦点調節にも矛盾しない極めて自然な立体像を記録, 再生する技術として注目を浴びることとなる. 以下, この2光束法について簡単に説明する.

図7.1に示すようにレーザから出た光をハーフミラーで2つに分け, 一方を参照光とし, 他方を物体に照射し, 参照光と物体光の2つの光によって生ずる干渉縞を感光材料に記録する. このようにして記録, 現像したホログラムにもとの参照光の方向から再生光を照射することにより, 撮影時と同じ位置に虚像が再生される.

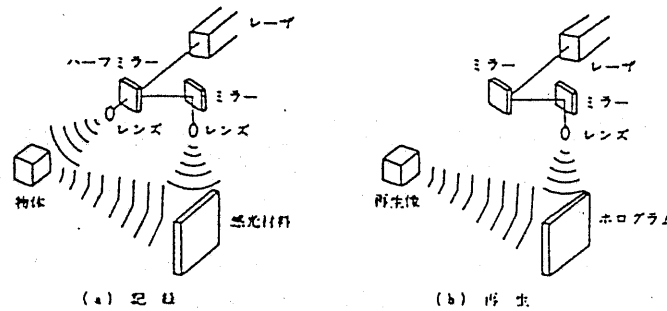


図 7.1: ホログラムの記録, 再生

さて, ホログラフィテレビジョンを実現するためにはこの干渉縞が表示可能な素子があれば良い. 干渉縞の分布は, 記録, 再現したい物体光の光波面と, 斜め方向より入射させた参照光の光波面の形により一義的に決定される. もっとも簡単な場合として, 図7.2のように物体光と参照光がともに平面波の場合を考えよう. この場合, 記録面上での干渉縞パターンは直線平行状になり, その縞の間隔 p は次式で与えられる.

$$p = \frac{\lambda}{\sin \theta} \quad (7.20)$$

ここで λ はレーザ光の波長であり, θ は斜め方向より入射させる平面波の入射方向と記録平面の法線とがなす角度である. 例として He-Ne レーザ (波長 $\lambda = 0.633\mu\text{m}$) を用いた場合, $\theta = 30^\circ$ ならば $p = 1.2\mu\text{m}$ となる. これらの考察から, off-axis 型のホログラムパターンの解像度としては, 最低でも 1000 lp(line pair)/mm 必要であるということになる. ホログラフィに用いられるフィルムは通常 1500 lp/mm 以上の解像度をもつものが使われる. したがって, ホログラフィテレビジョンの実現のためには 1000 lp/mm 以上の解像度で振幅変調, もしくは位相変調が可能であり, かつ実時間で書き込み, 消去が可能なデバイスが必要であるということになる. 現在の商用の液晶パネルのピクセルサイズが $100\mu\text{m}^2$ 弱であることを考えると, この数字がいかに小さいものかが分かるであろう.

また, (7.20) 式から分かるように, 参照光の入射角 θ が干渉縞の間隔に大きく影響する. 参照光の入射角を小さくすることで解像度の低減が可能である. 例えば $\theta = 1^\circ$ とすれば干渉縞の間隔は $p = 36.3\mu\text{m}$ となる. このように参照光の進む方向が物体光の進む方向とほぼ一致するホログラムをオンアキス (on-axis) ホログラムと呼び, 参照光が物体光とは異なる方向のものをオフアキス (off-axis) ホログラムと呼ぶ. on-axis 型のもはキャリア周波数を含まないので, 解像度が小さくてすむ反面, 透過光や共役像との分離が困難となる. ガボール型のホログラムは本質的に on-axis 型である. 理想的には off-axis 型のホログラムが構成できればよいのだが, ホログラフィテレビジョンの研究は表示素子の解像度との戦いでもあり, しばしば on-axis 型のものも利用される.

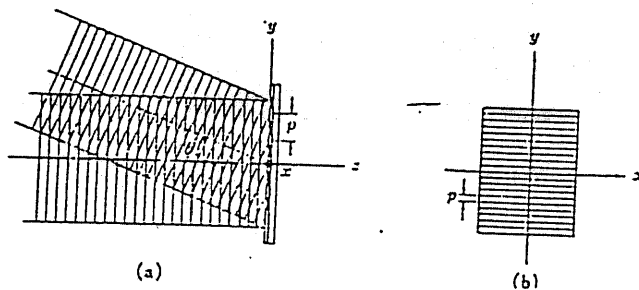


図 7.2: 平面波同士の干渉で生じる干渉縞パターン

7.3 ホログラフィテレビジョン技術の歴史と現状

ホログラフィの基本的な発明は, 1948年にガボール (Gabor) によってなされた. その後, 1962年に米国のリース (Leith) とウパトニクス (Upatnieks) (ミシガン大学) により 2 光束干渉法が考案され, ホログラフィの研究が活発化した. 3次元ディスプレイとしてのホログラフィーが注目を浴びるとすぐに動画化への検討が始められたものの, 通常の TV に比べて 10^4 - 10^5 倍程度の帯域が必要であるとする研究結果が報告され, その情報量の膨大さからホログラフィテレビジョンに対して悲観的とならざるを得なかった. その後しばらくはホログラフィテレビジョンに関する本格的な研究は現れていない.

このような経緯から, つい 4, 5 年前まではホログラフィテレビジョンは不可能と考えられていた. ところが, 状況は, ここへ来て急激に動きは始めている. MIT メディアラボでは AOM を用いたホログラフィビデオが提案され, また近年, 高精細化されてきた液晶を空間光変調器

として用いたホログラフィの実験が行われるようになってきており、夢が現実のものとなろうとしている。

本節では、初期のテレビジョン伝送実験について述べた後、現在、活発に研究が行われているいくつかの方式について述べ、最後にこのような技術が可能になった背景、今後の見通しについて述べる。

7.3.1 テレビジョン伝送の試み

TVにおけるホログラフィの応用は、2光束干渉法が発明されて間もない1965年に、ベル研究所のEnloeらがホログラフィの伝送を試みたのが最初であろう[5]。Enloeらは「BELL」の4文字を記録したスライドを物体とし、Mach-Zehnder干渉計で参照光の入射角を約18分と小さく選び、干渉縞をテレビカメラ(ビジコン)の光電面で直接受けてTV信号を作り、これを受像器のCRTに表示し、いったんフィルムに記録してホログラムをつくり、He-Neレーザーで再生して像を得ている。再生像を図7.3に示す。

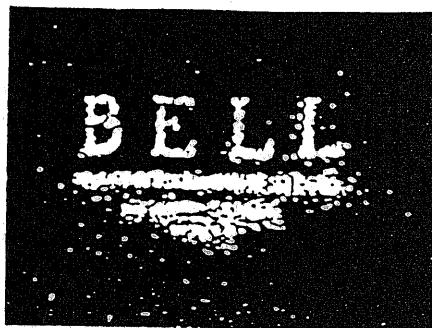


図 7.3: テレビ伝送されたホログラムからの再生像

7.3.2 情報量の計算

これと前後して、ミシガン大学のリースとウパトニクスらによりホログラム伝送に必要な情報量の計算がなされている[3]。この論文では、ホログラムの撮影装置、ディスプレイ装置といったハードウェアのことは棚上げにして、ホログラムをTV信号として伝送したときどれだけの帯域幅が必要かを検討したものである。彼らは具体的な数値を挙げ、9m離れて存在する $6 \times 6\text{m}^2$ の物体(2次元)を $\lambda = 500\text{nm}$ の光で $10 \times 10\text{cm}^2$ のホログラムに記録したとき、Gabor型のホログラムとしても全体で 4.9×10^9 個の標本点となるので、これを1秒30フレームのレートで伝送するとすれば、1秒間に 1.5×10^{11} の標本点を伝送しなければならず、通常のTVの能力 $8.4 \times 10^6 / \text{s}$ とは大きな隔たりがあることを指摘し、これを解決するには搬送波の周波数を光の周波数程度に引き上げなければならない。将来、光通信でも実現すれば伝送も可能になるであろうと結んでいる。

このような経緯から、当時のホログラフィ研究者たちは、光学系においてなんとか情報圧縮を試みようとする方向と、当分実現の見込みはないと悲観的な見方をする方向とに別れたという。ホログラムのテレビジョン伝送に関する研究は、以後20年間、目立った進歩は見られなかった。

7.3.3 ホログラフィ研究の停滞期

ホログラフィテレビジョンの研究が停滞している間、現在におけるホログラフィテレビジョンの研究へとつながる新しいホログラフィの方式が提案され、実験が行なわれていた。

1つめはホログラフィックステレオグラム¹の発明である [80]。ホログラフィックステレオグラムとは、ある間隔で観察点が次々と異なった多数枚の画像列を原画として使い、1枚のホログラムに合成して立体像を表示する方式である。ホログラフィックステレオグラムは1968年にカリフォルニア工科大学の McCrickerd と George により提案された。ホログラフィックステレオグラムの合成は、光学的な方法でももちろん可能であるが、異なる観察点より見た画像列が得られればよいことから、可視光以外の像の表示や計算機で合成した架空物体の表示も可能である。

2つめにはキノフォームの実験が挙げられる。キノフォームは厳密にはホログラムとは異なり、物体光波面の位相分布そのものを記録面上で実現し、光波面を実現するものである（ホログラムは参照光波面との干渉による光強度（または位相）分布の記録および変調であった。）。レンズと同様な結像作用により入力パターンを再生するものと考えてもよい。キノフォームを作成するためには、計算機において物体からの光を記録面上での波面に変換し、得られた波面の振幅は一定と近似して位相の分布を求め、記録媒体の屈折分布などによって位相の分布を記録すればよい。このように計算機内にある物体を表示させることが可能であり、またキノフォームは on-axis 型の光学素子であることから解像度の点でも有利である。しかしながら、実物体の表示を行うことは記録面上での位相分布の検出が行なえない限り無理である。したがって現時点では必然的に CGH (Computer Generated Hologram) となる。

ホログラフィックステレオグラムの発明により物体にレーザー光を当てなくてもホログラムが合成できるようになり、また、キノフォームにより on-axis 型で（すなわち低解像度の素子でも）波面の制御が可能となった。しかしながら、CGH において波面の計算に莫大な計算コストがかかること、また肝心の表示デバイスがなかったことからこれらの技術をすぐにホログラフィテレビジョンに応用することはできなかったのである。

7.3.4 ホログラフィテレビジョン研究の最近の動向

計算機の高速化、メモリの大容量化、液晶デバイスの高精細化、CG 技術の進歩などに支えられ、1990年前後、ホログラフィテレビジョンに関する研究は再び活発化する。近年盛んに研究されている方式は次の3つである。

1. 直接伝送方式
2. 超音波光変調器 (AOM) 方式
3. キノフォームによる方法

それぞれの方式について、最近の研究を紹介していく。

直接伝送方式

直接伝送方式とは、物体にレーザーを照射し光学的に干渉縞を作成し、干渉縞を高画素数の CCD カメラで直接撮影、電気信号に変換して伝送し、そのビデオ信号で高画素数の液晶ディスプレイパネルを駆動して干渉縞を表示し、そのパネルにレーザー光を当てて、元の像を再生する

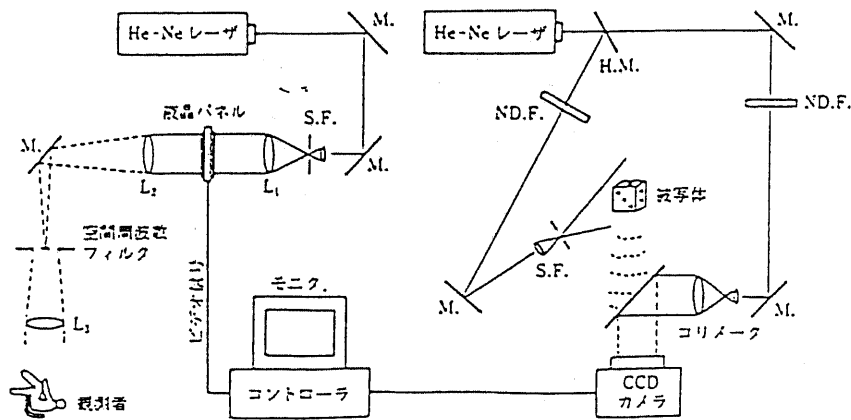
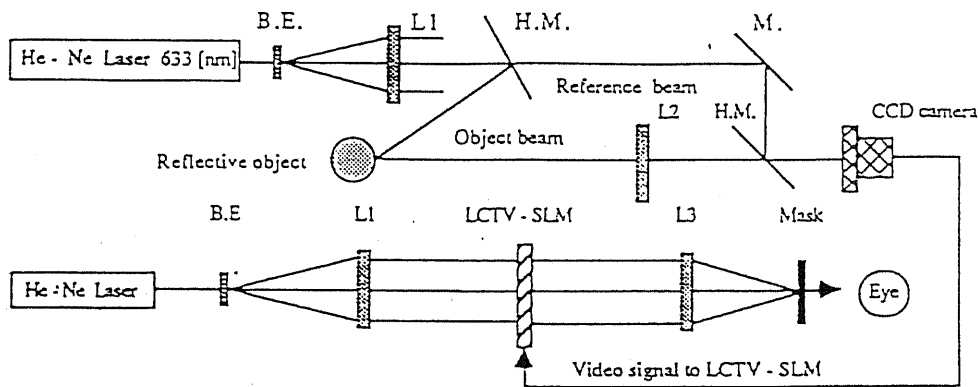


図 7.4: ホログラフィテレビジョン実験光学系



B.E. Beam expander L1. Collimate lens L2. Imaging lens
L3. Convergent lens H.M. Half mirror M. Mirror

図 7.5: ワンステップイメージホログラムの光学系

方式である。アイデアとしては先に紹介した初のテレビジョン伝送と似ているが、高解像度の CCD カメラおよび高画素数の液晶パネルが手に入るようになったことから現実的となった。

図 7.4 に湘南工科大学でおこなわれたホログラフィテレビジョンの実験の光学系を示す [109, 112]。まず被写体にコヒーレント光を照射し、これと参照光をテレビカメラの受光面において干渉させて光学的ホログラムを発生させる。Mach-Zehnder 干渉計を用いて、入力パターンとして反射物体を一方の光路に入れて光学的ホログラムを作成している。次に、これをテレビカメラで直接撮影し、干渉縞を電気信号に変換した後に、テレビ信号の伝送系を用いて伝送する。さらにコントローラを介して液晶パネルにホログラムとして表示したものに、再生用のレーザを照射して実時間で像を再生する。得られたホログラムは on-axis 型ホログラムとなっている。そのため、レンズ+ピンホールにより空間的フィルタリングをおこない、0 次透過光および回折光を除去する。

この実験で用いている液晶はセイコーエプソンの市販の液晶プロジェクタであり、液晶セルは約 $80 \times 90 \mu\text{m}$ である。解像度は 10 数本/mm で、通常では立体像を再生するためには不十分であるが、本実験では on-axis 型のホログラムを用いているためにキャリア成分を含まず、立体像の表示が可能である。

一方、シチズンでは同様に光学的な方法でさらに解像度の高い液晶パネルを用いて、off-axis 型ホログラムの作成を行っている [134]。液晶セルは $30 \times 60 \mu\text{m}$ である。作成されたシステムの概要を図 7.5 に、結果の再生像を図 7.6 に示す。また、本手法ではイメージ型ホログラムを用

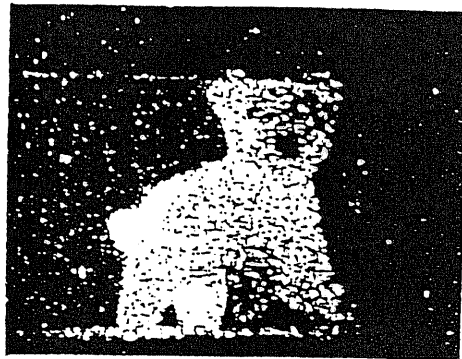


図 7.6: ワンステップイメージホログラムの再生像

いているために、像がホログラムの近傍にできるので白色光再生が可能である。

超音波光変調器 (AOM) 方式

1990 年はじめに MIT メディアラボのベントン教授らは、計算機によって合成されたホログラムを準実時間的に表示するシステム「ホログラフィビデオ」を提案した [43, 44, 45]. このシステムの概略を図 7.7 に示す. このシステムでは計算機内で発生させたパターンをもとに干渉縞を超高速パラレル・コンピュータ (コネクションマシン) により作成し、ホログラムの 1 フレームを 2×10^6 サンプルまで圧縮しメモリに収める. メモリからの信号により AOM (acousto-optical modulator; 音響光学変調器) を駆動してホログラム信号を超音波に変換し、レーザ光の回折方向と振幅を変化させ空間像を再現する. 水平方向には長い AOM (ドライブ周波数は 50-100MHz) でレーザ光を (時) 空間変調しており、垂直方向に対しては、ガルバノミラーで走査している. また AOM で光波面が流れる (約 700m/sec) のを止めるため、ポリゴンミラーを用いて機械的走査をおこなっている (Mark I system). また、同様の方式で AO 素子の各チャンネルに RGB の信号を入力することによりカラーホログラムが得られている [46]. 再生像の一例を図 7.8 に示す. このシステムで用いられているホログラム仕様は、読みだし速度 40Hz, 解像度 192 本/mm, ホログラムの大きさは 10cm^2 であり、また再生像の大きさ

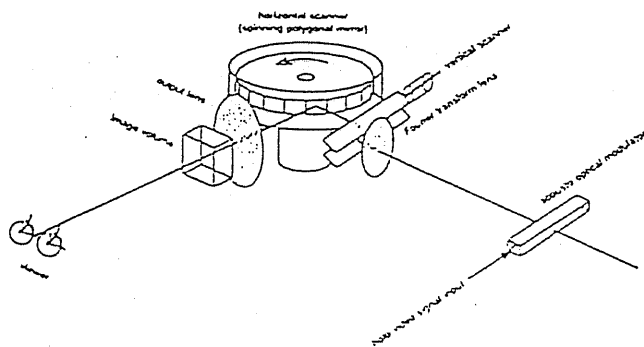


図 7.7: ホロビデオの概念図 (Mark I system)

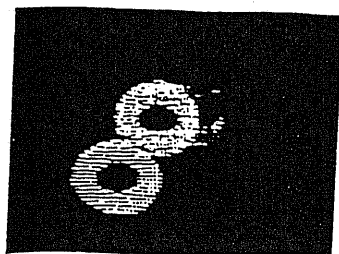


図 7.8: ホロビデオの再生像 (Mark I system)

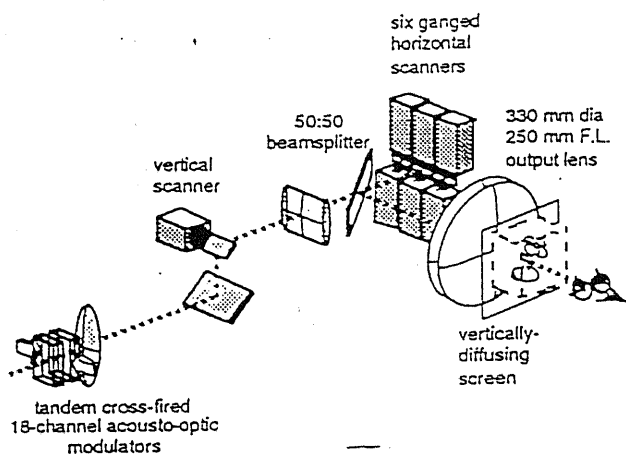


図 7.9: ホロビデオの概念図 (Mark II system)

4cm², 視域角は水平方向 12°, 垂直方向 23°程度である。

また最近では, 18 チャンネルの AOM を用い, ポリゴンミラーの代わりに6つのガルバノミラーを用いて水平走査を行うシステムが開発され (図 7.9), 再生像の大きさ 135 × 70 mm, 視域角 30°が得られている (Mark II system)[47].

キノフォーム方式

物体光波面の位相分布そのものをビデオ的に実現できる可能性のあるデバイスとして, 液晶パネルが注目された. 先に挙げた直接方式においては, Twisted Nematic(TN) 液晶を偏光板ではさんでレーザー光強度空間変調デバイスとして用いているが, ここでは光波位相空間変調素子として用いるのである.

液晶を用いたキノフォームの最初の実験は 1986 年に位相変化を 2 段階のみとした基礎的な実験が初めてなされ, また 1989 年には TN(Twisted Nematic) 液晶を用いた本格的なキノフォームの実験がなされた [123]. その後, TN 液晶より位相変調特性の優れたホモジニアス液晶を用いたキノフォームの作成が, 1990 年, 初めてセイコーエプソン社によって行われた [119]. これは, ホモジニアス液晶を用い, アクティブマトリックス方式で駆動している. 位相変調を用いるために偏光板ははずしてある. 液晶パネルの仕様は, 液晶サイズ 25.6 × 19.8mm, ピクセルサイズ 80 × 90μm, 厚さ 6.0μm である. 図 7.10にキノフォームパターンと再生像を示す.

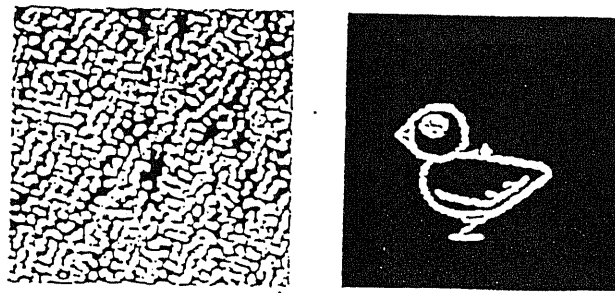


図 7.10: キノフォームパターンと再生像

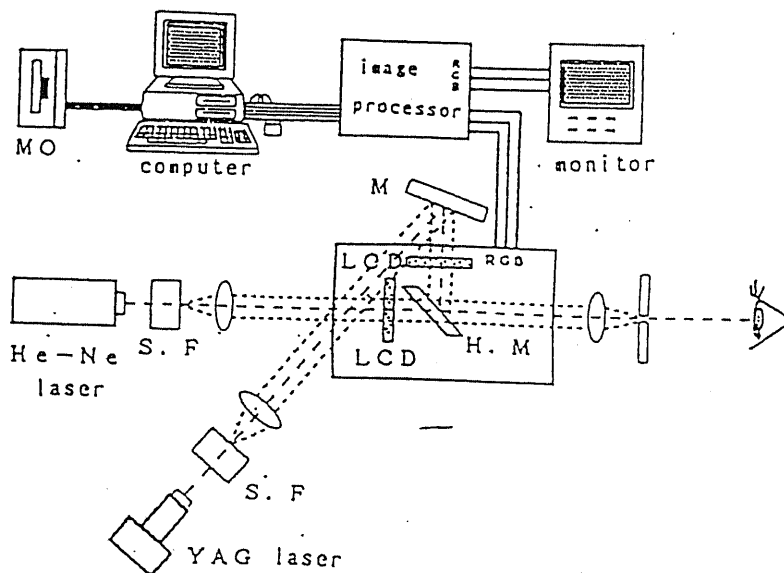


図 7.11: カラー表示可能なキノフォーム実験光学系

標本点数は 128×128 で量子化数は 16 である。

また湘南工科大学においても同様の実験がなされている [114]. 液晶などの各数値は上と同じであるが, He-Ne レーザ (赤) に加え YAG レーザ (緑) を使用してこれらをハーフミラーで合成することによりカラー化を目指している. 図 7.11 に実験システムを示す.

7.3.5 今後の展望

以上, ホログラフテレビジョンについて歴史的な経緯, 現在の研究の動向についてみてきた. 近年においてホログラフテレビジョンの研究が再び活発になってきたわけであるが, その背景としては次のものが考えられる.

1. 計算機の高性能化 (高速演算, 大容量メモリ)
2. 液晶技術の進歩 (高精細化, コントラスト, 位相変調特性の向上)
3. 情報圧縮技術の進歩 (縦方向の視差の放棄, ホログラムのサンプリング)

まず, 計算機の高性能化がキーテクノロジーとなったのは MIT 方式のホログラフィビデオであろう。MIT 方式のホログラフィビデオは, 高速なコンピュータを用いた CGH の方式によりインタラクティブ操作が可能なレベルにまで来ている。今後イメージサイズを拡大していくにつれ必要な情報量は増加していくのでさらに高速な計算機, 大容量メモリが必要となる。現在のところ機械的な走査を伴うのが難点だが, 現行のテレビジョンも最初は機械的な走査方法で始まったことを考えると, これが将来どのように発展していくか非常に興味深い。

液晶技術の進歩に支えられているものは, 直接伝送方式, キノフォーム方式である。液晶方式においては, 今後さらに高精細な液晶パネルを作成することが課題である。キノフォームによる方法では液晶パネルはそれほど大きな解像度を必要としないことから, 比較的早期に実現されると思われる。

情報圧縮技術として用いられているものは, 現在のところ縦方向の視差の制限とホログラムのサンプリングであり, MIT 方式に採用されている。今後はホログラムパターンを対象とした情報圧縮, 補間, 拡大・縮小などの信号処理技術も必要とされるだろう [104]。これらについてはまだ研究段階であり, 発展途上にある。

これらの技術は今後ますます発展していくと考えられ, ホログラフィテレビジョンの実現も夢物語ではなくなりつつある。本章で紹介した方式のいずれか, または全く新しい方式のホログラフィテレビジョンが, 将来の3次元ディスプレイとして実用化される日がくることであろう。

7.4 ホログラムパターン圧縮の研究動向

ホログラムパターンそのものを圧縮, 符号化しようとする研究はあまり多くは報告されていない。代表的なものとして, 文献 [76] において, ホログラムパターンをまびくことにより圧縮を実現し, 失われた部分を転写により補うという手法が提案されている。

近年になり, ホログラムパターンの直接高能率符号化の試みも報告されている [104, 105]。ホログラムパターンをブロック分割した後離散コサイン変換 (Discrete Cosine Transform; DCT) により変換し, 静止画標準符号化方式である JPEG を適用することで圧縮を実現するものである。この手法は, ホログラムパターンが高い空間周波数まで一様に含むことを利用するものである。DCT を用いてホログラムを周波数領域に変換すると, すべての周波数の係数がほぼ同じような値を持つ。したがって, これらの係数の間の相関をとり除くことで圧縮が期待できる。この手法は, DCT および JPEG といった, 現在画像符号化で利用されている手法を積極的に応用することにより, 高速に実現できることを目指すものである。

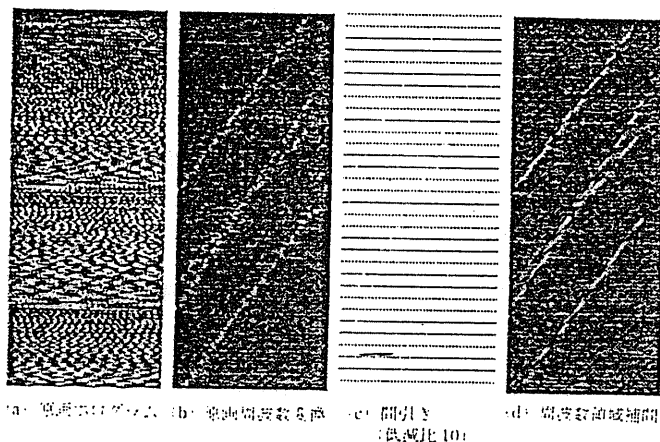


図 7.12: 計算機ホログラムの周波数変換による情報圧縮

7.5 ホログラムパターンと光線空間との相互変換

これまで, 3次元画像の統一的な記述法として光線空間を考えてきた。ホログラムパターンを光線空間で記述することや, またはその逆ができるならば, 3次元統合画像符号化への道を開くことになるだろう。本節では「ホログラムパターンと光線空間との相互変換」に焦点を当て, 離散フレネル変換を利用した手法について検討する。

7.5.1 離散フレネル変換の導入

フレネルーキルヒホッフの式より, フレネル領域での回折は次のように表される。

$$G(u) = \frac{1}{j\lambda z} \exp(j\frac{2\pi}{\lambda}z) \int_{-\infty}^{\infty} g(x) \exp(j\frac{\pi}{\lambda z}(u-x)^2) dx \quad (7.21)$$

このフレネル変換を計算機を利用して離散系で計算するために, 文献 [151] においてフレネル変換, 逆変換が導入されている。フレネル変換, 逆変換の定義を以下に示す。

$$\text{フレネル変換: } G(u) = \int_{-\infty}^{\infty} g(x) \exp(j \frac{\pi}{\lambda z} (u-x)^2) dx \quad (7.22)$$

$$\text{フレネル逆変換: } g(x) = \frac{1}{\lambda z} \int_{-\infty}^{\infty} G(u) \exp(-j \frac{\pi}{\lambda z} (u-x)^2) du \quad (7.23)$$

これを離散化することにより, 以下のように離散フレネル変換, 離散フレネル逆変換が定義できる.

$$\text{離散フレネル変換: } G_s(kT_u) = \sum_{n=0}^{N-1} g(nT_x) \exp(j \frac{\pi}{\lambda z} (kT_u - nT_x)^2) dx \quad (7.24)$$

$$\text{離散フレネル逆変換: } g_x(nT_x) = \frac{1}{N} \sum_{k=0}^{N-1} G(kT_u) \exp(-j \frac{\pi}{\lambda z} (kT_u - nT_x)^2) du \quad (7.25)$$

このとき, 次の条件を満たす必要がある.

$$\frac{T_x T_u}{\lambda z} = \frac{1}{N} \quad (7.26)$$

また, 回折面での値を持つ範囲を B とすると,

$$\frac{\lambda z}{T_x} \geq B \quad (7.27)$$

また, 物体面での値を持つ範囲を A とすると,

$$\frac{\lambda z}{T_u} \geq A \quad (7.28)$$

を満たさなくてはならない.

7.5.2 ホログラムから光線空間への変換実験

前節で導入したフレネル変換, 逆変換を用いて, ホログラムパターンから光線空間への変換の検討を行なう. その手順は以下の通りである.

まず, 仮想物体をもとにコンピュータにより基準面における振幅位相分布を計算する. この振幅位相分布に参照光を当てて干渉させ, 振幅分布のみを記録した場合には, 狭い意味でのホログラムとなる. また, 振幅位相分布そのものをデータとした場合にはキノフォームとなる. ここでは, 振幅位相分布が記録されているものと仮定して以下のシミュレーションを行なう. 狭い意味でのホログラムは, 位相の項を無視した特殊な場合として位置付けることができるからである.

次に, 基準面における振幅位相分布に対して前節で導入した離散フレネル変換を施す. これを光線空間において考えた, 「光線の位置と方向」という記述になるように得られたデータを並び替える.

以下, それぞれについて詳細に説明していく.

基準面における振幅位相分布の計算

コンピュータで作成した仮想物体からホログラムを計算する。そのときに、計算時間、情報量を低減するために縦方向の視差は無視するものとする。また、波長 $\lambda = 1\mu\text{m}$ 、物体は一辺が 2cm の正立方体とし、ホログラムは縦横 $5 \times 5\text{cm}$ とする。横方向のサンプリング間隔は、 $0.25\mu\text{m}$ とし、縦方向のサンプリング間隔は、 $125\mu\text{m}$ とする。なお、仮想物体とホログラム面との距離は 3cm 、 6cm の2種類について実験を行なった。

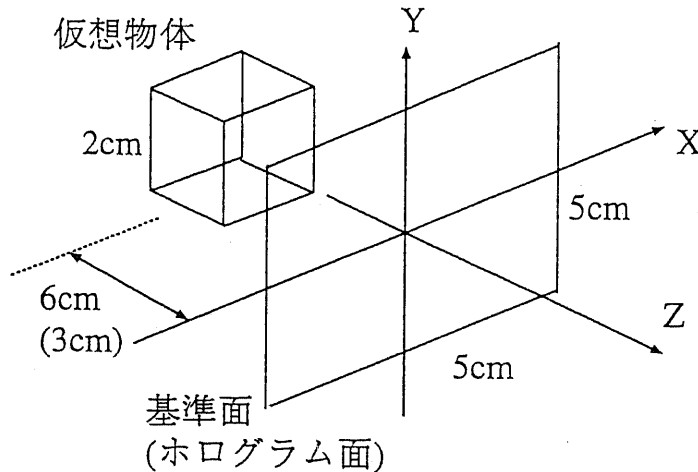


図 7.13: ホログラムの計算

離散フレネル変換の適用による光線空間への変換

図 7.14 はホログラム面、光線空間の任意の y 座標での水平断面図を示している。1次元のホログラムを等しい幅で分割し、それぞれを離散フレネル変換を用いて変換像を作る。ホログラム面から離れたイメージ面上での各々の変換像を図のように重ねると光線空間ができる。また、光線空間からホログラムへ変換するには、変換像を離散逆フレネル変換する。これにより、ホログラムと光線空間の相互変換が可能となる。

縦横 $5 \times 5\text{cm}$ のホログラムを横方向に 196 分割し、 5.12cm 離れたイメージ面上に各々の変換像を作る。次に変換像の $y = 0$ の水平断面から光線空間の断面図を作成した。イメージ面上の横方向のサンプリング間隔を $200\mu\text{m}$ とする。

図 7.15 に各々の変換像を示す。ここでは 3 種類の変換像を示しており、 x が最小の部分 (図 7.14 における A の部分)、中央の部分 (Z 軸上)、および x が最大の部分 (図 7.14 における C の部分) である。また、図 7.16 にホログラムパターンから計算された光線空間の水平断面図を示す。

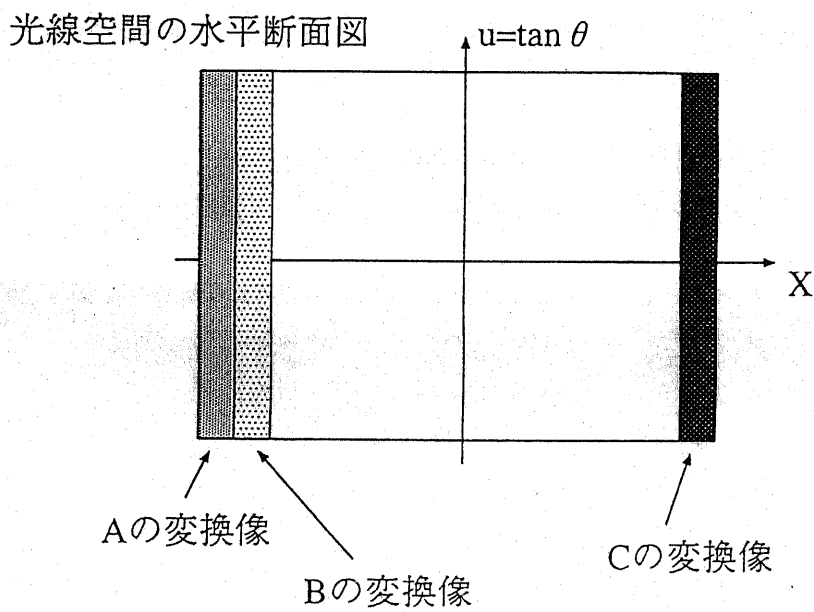
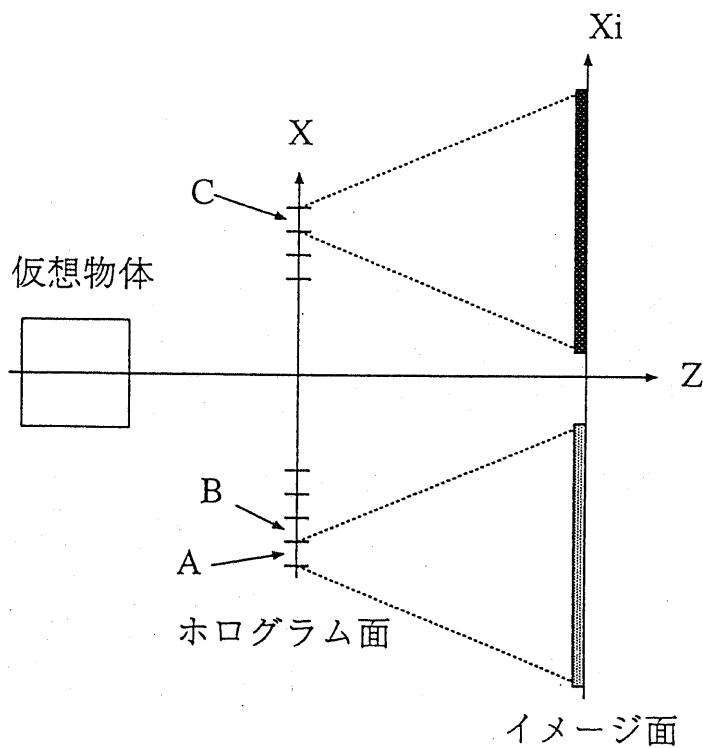


図 7.14: ホログラムと光線空間の水平断面図

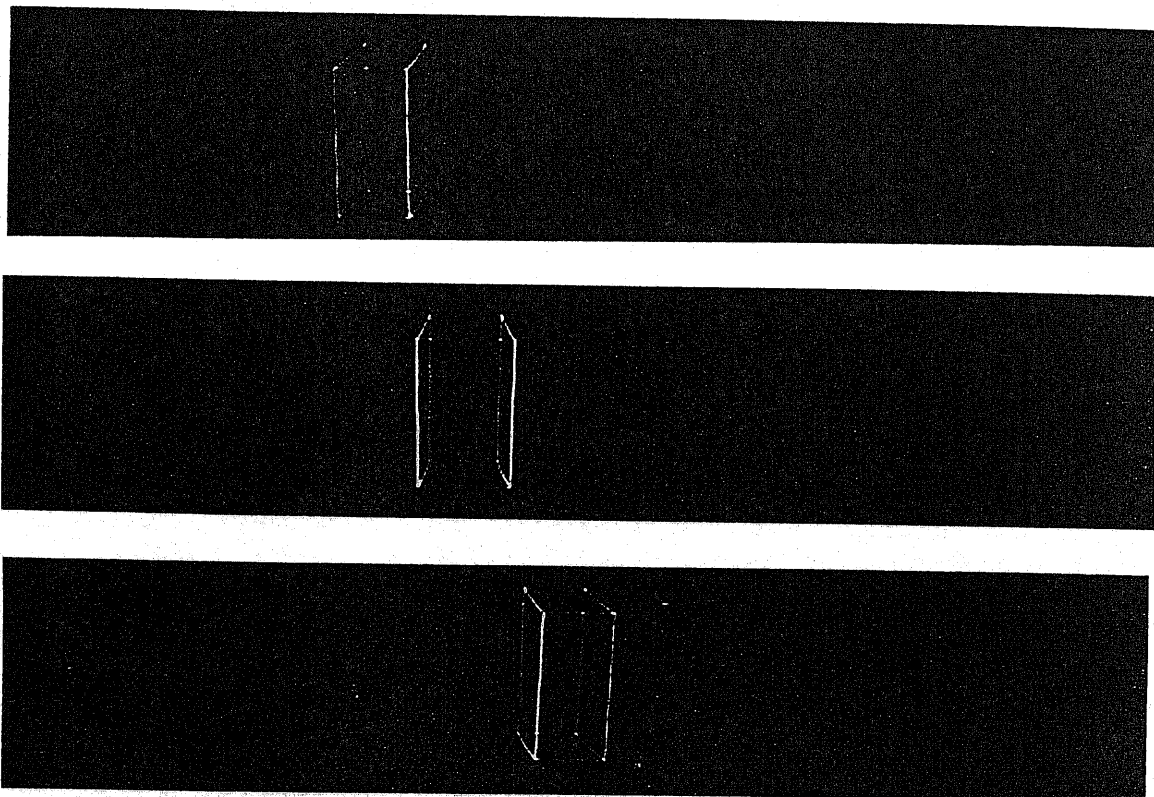


図 7.15: 要素ホログラムの離散フレネル変換像. 被写体とホログラム面の距離が 6cm の場合.



図 7.16: 求められた光線空間の水平断面図. 被写体とホログラム面の距離が 6cm の場合.

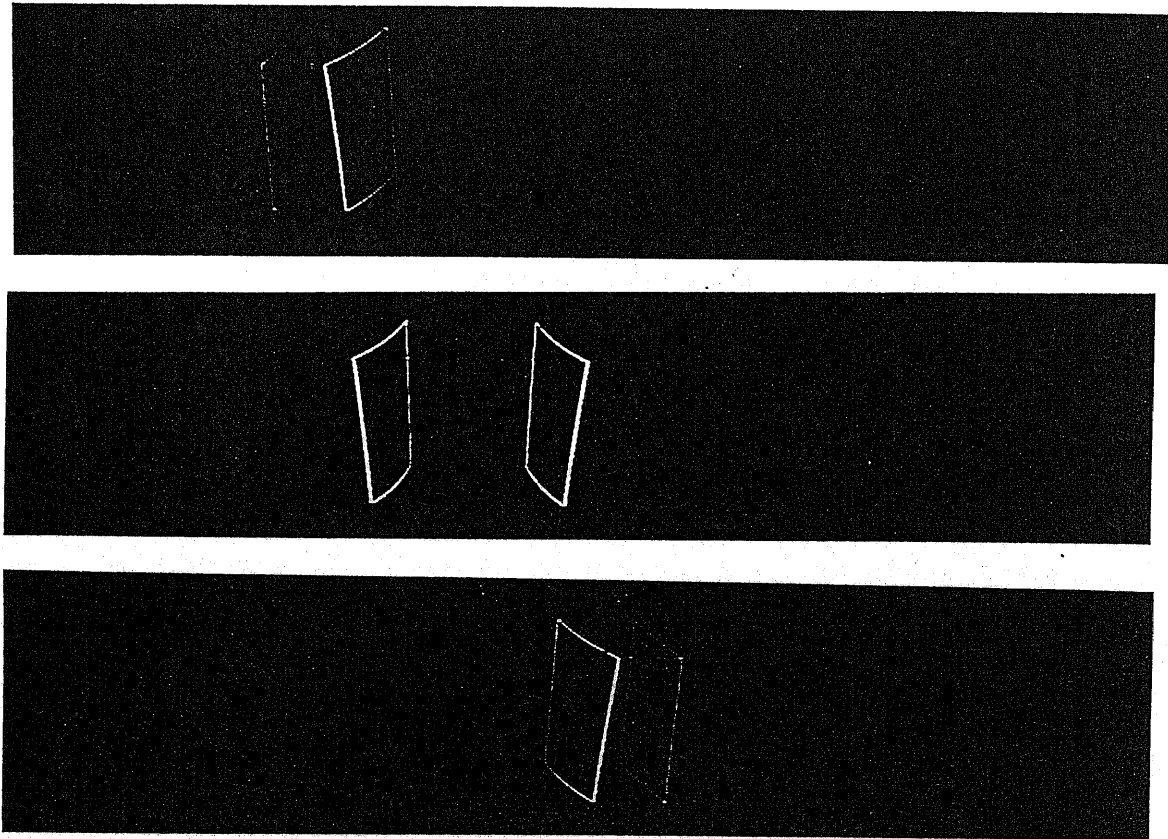


図 7.17: 要素ホログラムの離散フレネル変換像. 被写体とホログラム面の距離が 3cm の場合.



図 7.18: 求められた光線空間の水平断面図. 被写体とホログラム面の距離が 3cm の場合.

7.6 まとめ

近年, 動画ホログラフィの実験例が数多く報告されるようになってきており, 再びホログラムパターンの圧縮の必要性が高まってきたといえるだろう. このような背景のもと, 本章では, 3次元統合画像符号化に向けたホログラムパターンの情報処理について基礎的な検討を行なった.

「光線空間の符号化」という枠組へホログラムパターンを乗せることができるならば2次元画像符号化方式をも含めた圧縮, 符号化手法をホログラムへと適用することが可能であると思われる. したがって, 光線空間とホログラムパターンとの相互変換の研究を進めていくことで, 3次元統合画像符号化への道が開けてくるのではないかと思われる.

付録

$$F[e^{j\pi x^2/\lambda z_i}] = \sqrt{z_i \lambda} e^{j\frac{\pi}{4}} e^{j\pi \lambda z_i f_x^2} \quad (7.29)$$

であって, y_0 方向に対しても同じ式を適用することにより答を得る.

参考文献

- [1] D. Gabour, "A new microscopic principle," *Nature*, No. 161, pp. 777-779 (May 1948).
- [2] E.N. Leith and J. Upatnieks, "Reconstructed wavefronts and communication theory", *J. Opt. Soc. Am.*, 52, 10, pp. 1123-1130 (Oct. 1962).
- [3] E. N. Leith, J. Upatnieks, B.P. Hildebrand, and K. Haines, "Requirements for a Wavefront Reconstruction Television Facsimile System", *J. SMPTE*, 74, 10, pp. 893-896 (Oct. 1965).
- [4] E. N. Leith, J. Upatnieks, A. Kozma, and N. Massey, "Hologram Visual Displays" , *J. SMPTE*, 75, 4, pp. 323-326 (Apr. 1966).
- [5] L. H. Enloe, J. A. Murphy, C. B. Rubinstein, "Hologram Transmission Via Television", *Bell System Technical Journal*, 45, 2, pp. 335-339 (1966).
- [6] C. B. Burckhardt and L. H. Enloe, "Television Transmission of Holograms With Reduced Resolution Requirements on the Camera Tube", *Bell System Technical Journal*, 48, 5, pp. 1529-1535 (1969).
- [7] J. E. Berrang, "Television Transmission of Holograms Using a Narrow-Band Video Signal", *Bell System Technical Journal*, 49, 5, pp. 879-887 (1970).
- [8] 本田: "ホログラフィック・ディスプレイ技術の展望", *O plus E*, No. 146, pp.87-92 (1992年1月).
- [9] 動画ホログラフィ研究会会報.
- [10] 佐藤: "動画ホログラフィの研究動向", *テレビ誌*, 46, 5, pp. 586-593 (May 1992).
- [11] 辻内: "最近のホログラフィック・ディスプレイ", *テレビ誌*. 44, 1 (1990).
- [12] ベントン: "CGとホログラフィの結合", *サイエンス*. 1. pp. 38-39 (1989).
- [13] 樋口: "動画ホログラフィへの挑戦". *3D映像*. 5. 1. pp. 3-11 (1991).
- [14] 樋口: "活発化する動画ホログラフィ研究". *画像ラボ*. 9 (Sep. 1991).
- [15] 本田: "ホログラフィック・ディスプレイの現状と将来展望". *オプトロニクス*, 10 (Oct. 1991).
- [16] Y. N. Denisyuk, "Three-dimensional and pseudodeep holograms", *J. Opt. Soc. Am. A*, 9, 7, pp. 1141-1147 (July 1992).

- [17] H. Farhoosh, Y. Fainman, K. Urquhart, S. H. Lee, "Real-time Display of 3-D Computer Data Using Computer Generated Hologram", SPIE Proc. 1052, pp. 172-176 (1989).
- [18] G. Tricoles, "Computer generated holograms: an historical review", Appl. Opt., 26, 20, pp. 4351-4360 (1987).
- [19] 津田, 黒川, 森中: "フォトクロミック媒体を用いた書換え可能ホログラム", 1989年秋応用物理学会, 27p-8F-8, 第3分冊, p. 747 (1989).
- [20] 山口, 岡田: "バイオフィルム使いホログラフィ撮影", 日本経済新聞 (1992.1.27).
- [21] Y. Tsunoda, Y. Takeda. "Three-Dimensional Color Display by Projection-Type Composite Holography", IEEE Trans., ED-22, 9, pp. 784-788 (1975).
- [22] 杉浦, 山口, 本田, 大山: "1ステップ・リップマン・ホログラフィックステレオグラムによる3-Dアニメーション", 第21回画像工学コンファレンス, No. 12-7, pp. 243-246 (1990).
- [23] 樋口, 山, 石川: "ホログラフィ映画技術の実験的検討", 第22回画像工学コンファレンス, No. 9-10, pp. 231-234 (1991).
- [24] D. J. De Bitetto, "A Holographic motion Picture Film With Constant Velocity Transport", Appl. Phys. Lett., 12, 9, pp. 295-?? (May 1968).
- [25] Underkoffler, J. S., "Toward Accurate Computation of Optically Reconstructed Holograms", S. M. Thesis, MIT (1991).
- [26] 五十嵐, 坂本, 青木: "ホログラム面における平面からの複素振幅分布の計算法", 1992年信学秋季全大, D-227, p. 6-229 (Sep. 1992).
- [27] 五十嵐, 坂本, 青木: "3次元物体の計算機合成ホログラム作成法" 1992年信学春季全大, D-463, p. 7-205 (Mar. 1992).
- [28] 丸山, 中村, 高橋, 志水: "線光源の計算機合成ホログラムを用いた3次元画像の構成の方式(2)", 1992年信学秋季全大, D-230, p. 6-232 (Sep. 1992).
- [29] 中村, 高橋, 志水: "計算機合成ホログラムによるカラー画像再生に関する一手法", 1992年信学秋季全大, D-231, p. 6-233 (Sep. 1992).
- [30] E. Leith, C. Chen, Y. Chen, D. Dilworth, J. Lopez, J. Rudd, P. C. Sun, J. Valdmanis, and G. Vossler, "Imaging through scattering media with holography", J. Opt. Soc. Am. A, 9, 7, pp. 1148-1153 (July 1992).
- [31] 伊藤, 山内: "電子ビーム露光装置による高精細ホログラム作成条件の検討", 1992年信学秋季全大, D-226 (1992).
- [32] 丸山, 中村, 高橋, 志水: "線光源の計算機合成ホログラムを用いた3次元画像の構成の方式(2)", 1992年信学秋季全大, D-230 (1992).
- [33] 中村, 高橋, 志水: "計算機合成ホログラムによるカラー画像再生に関する一手法", 1992年信学秋季全大, D-231 (1992).

- [34] 五十嵐, 坂本, 青木: "ホログラム面における平面からの複素振幅分布の計算法", 1992年信学秋季全大, D-227 (1992).
- [35] 五十嵐, 坂本, 青木: "3次元物体の計算機合成ホログラム作成法", 1992年信学秋季全大, D-463 (1992).
- [36] J. D. Stack and M. R. Feldman, "Recursive mean-squared-error algorithm for iterative discrete on-axis encoded holograms", *Appl. Opt.*, 31, 23, pp. 4839-4846 (Aug. 1992).
- [37] 谷口, 高木: "デジタル画像処理技術によるホログラフィック振動パターンより干渉縞野抽出に関する一考察", 1991年テレビ年次大, 12-1, pp. 237-238 (1991).
- [38] 谷口, 他: "ホログラフィックパターン計測法の振動解析への応用", *信学論 C*, J71-C, 1, pp. 74-82 (昭 63-1).
- [39] S. Ellingsrud and G. O. Rosvold, "Analysis of a data-based TV-holography system used to measure small vibration amplitudes", *J. Opt. Soc. Am. A*, 9, 2, pp. 237-251 (Feb. 1992).
- [40] L. Onural and M. T. Oezgen, "Extraction of three-dimensional object-location information directly from in-line holograms using Wigner analysis", *J. Opt. Soc. Am. A*, 9, 2, pp. 252-260 (Feb. 1992).
- [41] 市野, 藤川, 高橋, 志水: "ホログラフィック手法を用いたニューラルネット型3次元情報検出", 1992年信学秋季全大, D-232 (1992).
- [42] 本田: "主要文献題目集, 光学, 2, 1, pp. 57-63 (Feb. 1973).
- [43] P. St. Hilaire, S. A. Benton, M. Lucente, M. L. Jepsen, J. Killin, H. Yoshikawa, and J. Underkoffler, "Electronic Display System for Computational Holography" *Proc. SPIE*, 1212 (Practical Holography IV), pp. 174-182 (Jan. 1990).
- [44] S. A. Benton, "Experiments in Holographic Video Imaging", *Proc. SPIE*, IS#08, 3D 映像, 5, 2, pp. 36-68 (Apr. 1991).
- [45] P. St. Hilaire, S. A. Benton, M. Lucente, J. Underkoffler, H. Yoshikawa, "Real-time Holographic Display: Improvements using a multichannel acousto-optic modulator and holographic optical elements", *Proc. SPIE*, No. 1461, 37, pp. 254-261 (Feb. 1991).
- [46] P. St. Hilaire, S. A. Benton, M. Lucente, P. M. Hubel. "Color Images with the MIT Holographic Video Display", *Proc. SPIE*, 1667-33 (Feb. 1992).
- [47] S. A. Benton, "The Second Generation of the MIT Holographic Video System", *Proc. TAO First International Symposium on Three-Dimensional Image Communication Technologies*, S-3-1 (Dec. 1993).
- [48] P. St. Hilaire, S. A. Benton, M. Lucente, M. L. Jepsen, J. Killin, H. Yoshikawa, and J. Underkoffler, "Electronic Display System for Computational Holography," *Proc. SPIE*, 1212 (Practical Holography IV), pp. 174-182 (Jan. 1990).
- [49] S. A. Benton, "Experiments in Holographic Video Imaging." *Proc. SPIE*. IS#08.

- [50] S. A. Benton, "Experiments in Holographic Video Imaging," Proc. SPIE, IS#08, 3 D 映像, 5, 2, pp. 36-68 (Apr. 1991).
- [51] P. St. Hilaire, S. A. Benton, M. Lucente, J. Underkoffler, H. Yoshikawa, "Real-time Holographic Display: Improvements using a multichannel acousto-optic modulator and holographic optical elements", SPIE Proc., No. 1461, 37, pp. 254-261 (Feb. 1991).
- [52] S. A. Benton, "Elements of Holographic Video Imaging", Proc. 4th Int. Symposium on Display Holography, pp. ?? (July 1991).
- [53] M. Lucente, "Optimization of Hologram Computation for Real-Time Display", Proc. SPIE, 1667, pp. ?? (Feb. 1992).
- [54] P. St. Hilaire, S. A. Benton, M. Lucente, P. M. Hubel, "Color Images with the MIT Holographic Video Display", Proc. SPIE, 1667-33 (Feb. 1992).
- [55] S. A. Benton and M. Lucente, "Interactive Computation of Display Holograms", Proc. Comput. Graphics Int. '92, (June 1992).
- [56] M. Lucente, P. St. Hilaire, S. A. Benton, "New Approaches To Holographic Video", Proc. SPIE, 1732, pp. ? (?).
- [57] P. St. Hilaire, S. A. Benton, and M. Lucente, "Synthetic aperture holography: a novel approach to three-dimensional displays", J. Opt. Soc. Am. A, 9, 11, pp. 1969-1977 (Nov. 1992).
- [58] M. Lucente, "Interactive computation of holograms, using a look-up table", J. Electronic Imaging, 2, 1, pp. 28-34 (Jan. 1993).
- [59] P. St. Hilaire, S. A. Benton, M. Lucente, J. D. Sutter, W. J. Plesniak, "Advances in Holographic Video", Proc. SPIE, 1914, pp. ?? (?).
- [60] S. A. Benton, P. St. Hilaire, M. Lucente, J. D. Sutter, W. J. Plesniak, "Real-time computer-generated 3-D holograms", Proc. SPIE, 1983, pp. 536-543 (Aug. 1993).
- [61] S. A. Benton, "The Second Generation of the MIT Holographic Video System", Proc. TAO First International Symposium on Three-Dimensional Image Communication Technologies, S-3-1 (Dec. 1993).
- [62] M. W. Halle, "Holographic stereograms as discrete imaging systems", Proc. SPIE, 2176, pp. ?? (Feb. 1994).
- [63] S. A. Benton, S. M. Birner, and Akira Shirakura, "Edge-Lit Rainbow Holograms", Proc. SPIE, 1212, pp. 149-157 (1990).
- [64] W. J. Farmer, S. A. Benton, M. A. Klug, "The Application of the Edge-Lit Format to Holographic Stereograms", Proc. SPIE, 1461 pp. ?? (Feb. 1991).
- [65] M. W. Halle, S. A. Benton, M. A. Klug, J. S. Underkoffler, "The Ultragram: A Generalized Holographic Stereogram", Proc. SPIE, 1461, pp. ?? (Feb. 1991).

- [66] P. M. Hubel, "Recent advances in color reflection holography", Proc. SPIE, 1461, pp. ??(1991).
- [67] "Rainbow holograms", Optics and Photonics News (July 1991).
- [68] S. A. Benton, "Holographic imaging research at the MIT media lab", Proc. Third International Conference on "Holographic Systems, Components and Applications" (Sep. 1991).
- [69] P. M. Hubel and L. Solymar, "Color-reflection holography: theory and experiment", Appl. Opt., 30, 29 pp. 4190-?? (Oct. 1991).
- [70] P. M. Hubel, and M. A. Klug, "Color holography using multiple layers of Du Pont photopolymer", Proc. SPIE, 1667, pp. ?? (1992).
- [71] M. A. Klug, M. W. Halle, P. M. Hubel, "Full Color Ultragrams", Proc. SPIE, 1667, pp. ?? (Feb. 1992).
- [72] P. M. Hubel, "True-Color Holography", LEONARDO, 25, 5, pp. 507-515 (1992).
- [73] M. A. Klug, M. W. Halle, M. Lucente, W. J. Plesniak, "A Compact Prototype one-step Ultragram printer", Proc. SPIE, 1914, pp. ?? (?).
- [74] ベントン, 吉川:" ホログラフィックビデオ画像に関する実験", 3D映像, 5, 2, pp. 57-68 (1991).
- [75] ベントン, 吉川:" ホログラフィックビデオ画像に関する最近の実験", 3D映像, 5, 4, pp. 34-39 (1991).
- [76] C. B. Burckhardt, "Information Reduction in Holograms for Visual Display", J. Opt. Soc. Am., 58, 2, pp. 241-248 (Feb. 1968).
- [77] K.A. Haines and D.B. Brumm:"Holographic data reduction", *Appl. Opt.*, 7, pp.1185 (June 1968)
- [78] 永田, 浜崎:" レンズ列によるホログラム情報の削減法", 信学論, 56-C, 8, pp. 488-495 (1973).
- [79] 大越孝敬:「三次元画像工学」(朝倉書店) 第8章 "三次元画像の情報量とその削減"
- [80] J. T. McCrickerd and N. George, "Holographic Stereogram from Sequential Component Photographs", *Appl. Phys. Lett.* 12, 1, pp. 10-12 (Jan. 1968).
- [81] 高橋, 本田他: "ホログラフィックステレオグラムにおける視差画像の補間", 光学, Vol. 22, No. 11 (1993年11月).
- [82] 山口, 本田, 大山:" ホログラフィック・ステレオグラムによる像再生特性に関する考察", 光学, 22, 1, pp. 714-720 (Nov. 1993).
- [83] 高橋, 本田他: "少ない視差画像からのホログラフィック・ステレオグラムの原画作成" 3次元画像コンファレンス'93, 1-1 (1993-7).

- [84] 吉川：“ホログラフィックステレオグラム方式による電子ホログラムディスプレイ”，1991年第22回画像工学コンファレンス，No. 4-4，pp. 107-110 (1991).
- [85] S. A. Benton (吉川浩 訳)：“Experiments in Holographic Video Imaging”，3D映像，5，2，pp. 36-68 (Apr. 1991).
- [86] 吉川：“ホログラフィック3次元画像と符号化”，平3連大，P2-7 (Sep 1991).
- [87] 吉川，佐々木：“電子ホログラムディスプレイにおける情報圧縮の検討”，信学技報，IE92-12 (May 1992).
- [88] 吉川，佐々木：“電子ホログラムディスプレイにおける情報圧縮の検討”，1992年信学会秋季全大，D-228 (Sep. 1992).
- [89] 吉川，松丸：“電子ホログラムディスプレイのための隠面およびシェーディング処理”，信学秋季全大，D-229 (Sep. 1992).
- [90] 吉川：“ホログラフィック3次元画像の圧縮技術”，TV誌，Vol.47，No.2，pp.177-182 (1993年2月).
- [91] H. Yoshikawa, K. Sasaki, "Information reduction by limited resolution for electro-holographic display", SPIE Proc. 1914, 30 (Feb. 1993)
- [92] 松丸，吉川：“電子ホログラフィックディスプレイのシェーディングにおける半つや消し表現の一手法”，動画ホログラフィ研究会会報，7，pp. 11-14 (Mar. 1993).
- [93] 佐々木，吉川：“電子ホログラフィックディスプレイにおける再生像の拡大・縮小の検討”，動画ホログラフィ研究会会報，7，pp. 15-18 (Mar. 1993).
- [94] 佐々木，吉川：“電子ホログラフィックディスプレイにおける再生像の拡大・縮小の検討”，3次元画像コンファレンス'93，1，pp.77-80 (July 1993).
- [95] 松丸，吉川：“電子ホログラフィックディスプレイのためのシェーディングにおける影処理”，信学技報，IE93-71 (Nov. 1993).
- [96] 佐々木，吉川：“電子ホログラフィックディスプレイにおける再生像の拡大・縮小の検討”，3次元画像コンファレンス'93，No. 1，pp. 77-80 (July 1993).
- [97] 丹治，吉川：“ホログラフィック3次元画像の高効率符号化に関する基礎的検討”，テレビ年次大，19-2 (1993).
- [98] 吉川，佐々木，亀山：“3次元ディスプレイのための変換を利用したホログラムの計算法”，1993年テレビ年次大，19-1 (July 1993).
- [99] 吉川，佐々木：“動画ホログラフィの情報低減”，画像電子学会，Vol. 22，No. 4 (Sep. 1993).
- [100] 亀山，吉川：“奥行き情報を考慮したホログラフィックステレオグラムの計算法”，1993年信学秋季全大，SD-7-9 (Sep. 1993).
- [101] 吉川，丹治：“ホログラフィック3次元画像の高効率符号化”，テレビ学技法，ICS'93-37 (1993).

- [102] 佐々木, 吉川: “電子ホログラフィックディスプレイにおける情報低減の検討”, 信学論, J76-D-II, 9, pp.2151-2152 (Sep. 1993).
- [103] 佐々木, 吉川: “電子ホログラフィックディスプレイにおける画像情報補間の検討”, 信学技報, IE93-72 (Nov. 1993).
- [104] H. Yoshikawa, “Digital holographic signal processing”, *Proc. TAO First International Symposium on Three-Dimensional Image Communication Technologies*, S-4-2 (Dec. 1993).
- [105] 佐々木, 丹治, 吉川: “ホログラフィック3次元画像の情報圧縮” テレビ誌, Vol. 48, No. 10, pp. 1238-1244 (1994年10月).
- [106] T. H. Barnes et al., “Phase-only Modulation Using a Twisted Nematic Liquid Crystal Television”, *Appl. Opt.*, 28, 22, pp. 4845-4852 (1989).
- [107] 佐藤, 道下, 小林, 樋口, 勝間: “ホログラフィ・テレビジョン (ホロビジョン) へ向けての基礎実験”, 1991年春応用物理学会, 28P-A-14, 第3分冊, p. 792 (1991).
- [108] 佐藤, 道下, 小林, 樋口, 勝間: “液晶表示デバイスを用いたホログラフィ・テレビジョンの基礎実験 (IV)”, 1991年秋応用物理学会, 28P-A-14, 第3分冊, p. 792 (1991).
- [109] 佐藤, 樋口, 勝間: “液晶表示デバイスを用いたホログラフィ・テレビジョンの基礎実験”, テレビ誌, 45, 7, pp. 873-875 (Jul. 1991).
- [110] 佐藤, 道下, 小林, 樋口, 勝間: “ホログラフィ・テレビジョンへ向けての基礎実験 (II)”, 1991年テレビ年次大, 11-6, pp. 229-230 (1991).
- [111] 佐藤, ?:"?", 1991年信学秋季全大, D-185, p. 6-187 (Sep. 1991).
- [112] K. Sato, K. Higuchi, H. Katsuma, “Holographic Television by Liquid Crystal Device”, *Proc. SPIE*, No. 1667-02 (1992).
- [113] 佐藤, 樋口, 勝間: “液晶表示デバイスを用いたホログラフィテレビジョンの基礎実験 (V)”, 1992年信学秋季全大, D-462 (Sep. 1992).
- [114] K. Sato, “Characteristics of CGH by LCD and its application to display the animated color 3D image”, *Proc. TAO First International Symposium on Three-Dimensional Image Communication Technologies*, S-3-3 (Dec. 1993).
- [115] 佐藤: “液晶を用いたホログラフィックディスプレイ”. 写真工業, 8 (Aug. 1991).
- [116] 尼子, 曾根原: “液晶空間光変調器を用いたキノフォーム”. 1990年秋応用物理学会, 26a-H-10, 第3分冊, p. 754 (1990).
- [117] 尼子, 三浦, 曾根原: “液晶空間光変調器を用いたキノフォーム (II)”. 1991年春応用物理学会, 29a-SF-13, 第3分冊, p. 975 (1991).
- [118] 尼子, 三浦, 曾根原: “液晶空間光変調素子を用いた光の振幅位相変調” 1991年秋応用物理学会, 10a-ZK-2, 第3分冊, p. 861 (1991).

- [119] J. Amako, T. Sonehara, "Kinoform Using Electrically Controlled Birefringence Liquid Crystal Spatial Light Modulator", *Appl. Opt.*, **30**, 32, pp. 4622-4628 (Nov. 1991).
- [120] 当山, 石井:" 液晶空間光変調器による計算機ホログラムの記録", 1989年秋応用物理学会, 27p-ZF-3, 第3分冊, p. 745 (1989).
- [121] 田, 谷田貝:" 液晶テレビによる計算機ホログラムの記録とその応用", 1991年春応用物理学会, 30a-PA-8, 第3分冊, p. 844 (1991).
- [122] F. Mok, J. Diep, H. K. Liu and D. Psaltis, "Real-time Computer Generated Hologram by Means of Liquid-crystal Television Spatial Light Modulator", *Opt. Lett.*, **11**, pp. 748-750 (1986).
- [123] T. H. Barnes et al., "Phase-only Modulation Using a Twisted Nematic Liquid Crystal Television", *Appl. Opt.*, **28**, 22, pp. 4845-4852 (1989).
- [124] 阿須間, 廬, 本田, 大山:" 液晶パネルの位相変調特性", *光学*, **20**, 2, pp. 98-102 (1991).
- [125] 岡田:" 液晶を用いたホログラム", *ホログラフィックディスプレイ研究会会報*, No. 1, pp. 20-25 (1991).
- [126] 木下:" 液晶ディスプレイとサーモプラスチック・フィルムを用いたホログラム", 1990年春応用物理学会, 29a-D-7, 第3分冊, p. 766 (1990).
- [127] 黒川, 福島:" 空間光変調器を用いた実時間ホログラム記録・再生", *ホログラムディスプレイ研究会会報*, No.2, pp. 15-18 (1991).
- [128] U. Mahlab, J. Rosen, and J. Shamir, "Iterative generation of holograms on spatial light modulators", *Opt. Lett.* **15**, pp. 556-558 (1990).
- [129] J. Rosen, L. Shiv, J. Stein, and J. Shamir, "Electro-optic hologram generation on spatial light modulators", *J. Opt. Soc. Am. A*, Vol. 9, No. 7, pp. 1159-1166 (July 1992).
- [130] M. Yonemura, "Holographic Contour Generation by Spatial Frequency Modulation", *Appl. Opt.*, **21**, 20, pp. 3652-3658 (Oct. 1982).
- [131] 橋本, 諸川:" 液晶空間光変調器を用いたホログラフイテレビジョン", *テレビ学技法*, Vol. 15, pp. 25-30 (1991).
- [132] N. Hashimoto, S. Morokawa, K. Kitamura, "Real-time Holography using the High-resolution LCTV-SLM", *SPIE Proc.*, No. 1461 (Practical Holography V), **44**, pp. 291-300 (1991).
- [133] 橋本:" 新技術を用いた動画ホログラフイ", *画像ラボ*, No. 12, pp. 58-62 (1991).
- [134] 橋本, 諸川:" 液晶空間光変調器を用いた電子動画ホログラフイ", *3D映像*, Vol. 6, No. 1, pp. 4-10 (Mar. 1992)
- [135] 橋本, 諸川:" 液晶TVを用いた実時間ホログラフイ", *光学*, **3**, 21, (1992).
- [136] N. Hashimoto, S. Morokawa, K. Hoshino, "Improved real-time Holography System with LCDs", *SPIE Proc.*, No. 1667, **1** (1992).

- [137] R. J. Doyle et al., "Remote real-time reconstruction of holograms using the Lumatron", Appl. Opt. pp. 1261-1264 (1971).
- [138] 勝間:"私はホログラフィが好きです", 写真工業, 1989年3月, 7~11月号.
- [139] 奥川:"リアルタイムホログラフィ", 1990.11.30. 大学院論文輪講資料.
- [140] Holography, 1992年目白祭資料(日本女子大).
- [141] Intelligent Visual Illusion, 工学部計数工学科, 1993年五月祭資料.
- [142] 藤井:"動画ホログラフィの研究動向", 保立研演習資料 (July 1993).
- [143] 大野:"光情報処理の最近の動向", 1993.11.26. 大学院論文輪講資料.
- [144] 電子ホログラフィシステム, (財)イメージ情報科学研究所 (Nicograph 出展).
- [145] S. A. Benton and M. Lucente, "Interactive Computation of Display Holograms"
- [146] S. A. Benton, "Computational Holography and 3-D Visualization"
- [147] "Summary of Medical Imaging-Related Research at the Media Laboratory" (Jan. 1992)
- [148] "Overview of the Holographic Terrain Map Project"
- [149] 辻内:"[医用画像のマルチプレックス・ホログラムによる立体表示], 3D映像, Vol.3, no.4, pp.56~57 (1989)
- [150] 志村, 山口, 本田, 大山:"[マルチプレックス・ホログラムの大型化], 1989 第20回画像工学コンファレンス, 10-4, pp.315-318 (1989).
- [151] 小林, 本田, 塩見, 岡田:"[フレネル回折像からの計算機による像再生], 第25回画像工学コンファレンス, 6-6, pp. 141-144 (1994年12月).

第 8 章

結論

8.1 本論文の主たる成果

本論文では、従来のデバイス中心の符号化法ではなく、さまざまな方式の3次元ディスプレイが共存することを前提とした「3次元統合画像通信」(Integrated 3-D Visual Communication)の構想を提案し、入出力方式の解析、光線群による3次元画像記述の提案、視差補償予測符号化、3次元構造モデルを用いた符号化、およびホログラムパターンの圧縮、符号化について述べてきた。

従来の3次元画像の研究でおこなわれてきたデバイス中心の符号化法ではなく、入出力方式に依存しない「3次元統合画像通信」の構想は2次元画像符号化をも含めた統一的な符号化の枠組へと発展していく可能性を秘めている。

本論文では、まず現在利用できる3次元画像代表的な入出力方式について概観し、後の符号化方式の提案の際に必要な解析をおこなった。また、実際に実写画像を入力する際に不可欠となる画像の位置合わせ、歪み補正、また入力にコンピュータを用いた場合の高速計算法についても言及した。

次に、3次元統合画像通信においては「3次元画像の中間的な記述方式」の確立が重要な要素技術である。本論文においては、3次元画像を光線群の集まりであるにとらえ、撮像を光線群のサンプリング、表示を光線群の再生であるという見方を導入した。これにより3次元画像の統一的表現が可能となり、3次元画像の伝送とは「像再生空間の伝達」であることが見えてくる。また、光線群のサンプリングという観点から、視差画像の補間合成の問題についても言

及した。

「光線空間」の冗長性を利用した圧縮、符号化の技術の確立が次なる課題となる。

本論文においては、まず第1に、波形符号化的な考え方に基づく「視差補償予測符号化」について、2次元動画画像符号化方式の適用、および多眼画像の幾何学的な性質を用いた符号化法の提案を行ない、実験により特性評価を行なった。

第2には、3次元構造モデルを用いた光線空間の構造化という観点から、「3次元構造抽出符号化」について実験的検討を行なった。従来の特徴点のマッチングに基づく3次元構造の推定という枠組ではなく、符号化が目的であるという立場をはっきりと打ち出し、復号画像のSNRを評価基準とした構造推定手法の検討を行なった。また、符号化という観点から、モデルの記述に要するデータ量、予測特性について詳細に検討し、さらには、より構造化を進めることによる予測特性の評価、3次元DCTを用いた予測誤差およびテクスチャデータの符号化について実験的検討を行なった。

また、ホログラムパターンの圧縮、符号化も3次元統合画像通信を実現する上での重要な課題である。本論文においては、ホログラムパターンから「光線空間」へのデータ変換、およびその逆について検討し、ホログラムパターンの圧縮、符号化の問題を多眼画像など他の3次元画像の圧縮、符号化の問題と同じ土俵の上で考えることができる基礎を作った。

従来、研究されてきている3次元画像に関する研究を「3次元統合画像通信」の枠組からとらえ直し、統一的に論ずることが目標であるとはじめに掲げたが、本論文がそのひとつの足掛かりとなれば幸いである。

8.2 今後の課題と展望

さらに将来に目を向けた場合、3次元統合画像通信の向うには何が見えるであろうか。

ネットワーク化されたバーチャルリアリティ環境下においては、ネットワークを通じて環境そのものがやりとりされるという状況が想定される。このような状況における符号化とは、もはや「画像」の伝送ではなく、「空間そのもの」を効率良く伝送することが次なる目標となってくるであろう。画像の符号化といった観点ではなく、空間そのものをある中間的な記述に変換し符号化するという観点が重要になってくる。

先に見たように3次元画像の取得にはさまざまな方法があるが、3次元統合画像通信の立場から見れば、空間そのものが持っている情報の一部を異なった形態で観測しているにすぎない、ととらえられるのではないだろうか。

3次元統合画像通信が必要としている「中間的な記述」とは、空間そのものが持っている情報であるに違いない。この観点からは、画像符号化 (picture coding) ではなく、「空間符号化 (space coding)」とでも呼ぶべき概念へと発展していくのではなかろうか。その意味で、3次元統合画像通信とは、「空間符号化」へと続く道の第一歩となるであろう。

また、CGと実写が融合した仮想空間や、仮想空間内でのインタラクティブ操作といった技術は、今後ますます発展していくことが予想される。この点を考慮すると、

- 情報量を減らし、かつ編集操作などの処理が行ないやすい形に空間がもつ情報を「構造化」しておくこと、

が究極の目標となるであろう。

これら問題はさまざまなテーマへと発展していく可能性を持っており、今後研究が活発化する分野となることが期待される。

謝辞

本研究を通じて終始御指導下さいました原島博教授,そして数々の助言をいただき熱心に討論して下さいました金子正秀助教授に心より感謝致します。また,私が修士課程在籍の折りに御指導下さいました,現,東亜大学の濱崎襄二名誉教授には,本研究全般に渡って有益な御助言をいただきました。心より感謝の意を表します。

また,研究室において数々の便宜を図って下さった松尾達郎 技官をはじめとする職員の方々,秘書の方々,諸先輩,後輩の方々に深く感謝します。特に,斉藤隆弘 博士,森島繁生 博士,相澤清晴 博士,森川博之 博士,中屋雄一郎博士,鄭文濤氏,許博程 氏,永田 明德氏,木藤 英二氏,苗村健 氏には,本研究全般にわたって有益な御助言をいただくとともに熱心な討論をして下さいました。また,本論文の5章に述べた研究は吉岡克真氏,差峨田淳氏の協力によりなされたものであり,深く感謝致します。最後に,生活面においてもこれらの方々を始めとして,数多くの方々のお世話になりました。ここに深く感謝の意を表します。

発表文献

学会誌論文

- [1] T.Fujii and H. Harashima, "Data Compression and Interpolation of Multi-View Image Set", IEICE Special Issue on 3D Image Processing(1994-9).
- [2] 藤井, 原島: "3次元構造モデルに基づく多眼3次元画像の圧縮と補間", テレビジョン学会誌(投稿中).

国際会議

- [3] Toshiaki Fujii, Joji Hamasaki, Reiji Ishima "Experiments on data compression for an autostereoscopic lenticular 3D TV", First International Symposium on 3D Images, Paris (1991-1).
- [4] Toshiaki Fujii, Joji Hamasaki, and Mathias Pusch, "Data compression of an autostereoscopic 3D image", The international symposium on three dimensional image technology and arts, Seiken symposium, Tokyo (1992-2).
- [5] T.Fujii, H.Harashima, J.Hamasaki: "Data Compression for an Autostereoscopic 3-D Image", *PCS '93*, 13.21 (1993-3).
- [6] T. Fujii, H. Harashima, "Data compression of an autostereoscopic 3-D image", SPIE Proc. Vol. 2177 *Stereoscopic Displays and Applications V*, San Jose, CA (1994-2).
- [7] T. Fujii, H. Harashima, "3-D Image Coding Based on Affine Transform", *ICASSP '94*, 81.8, Adelaide (1994-4).
- [8] T. Fujii, H. Harashima, "Data Compression and Interpolation for an Autostereoscopic 3-D Image", *PCS '94* (1994-9).
- [9] T. Fujii, H. Harashima, "Coding of an Autostereoscopic 3-D Image Sequence" *VCIP '94* (1994-9).

学会研究会

- [10] 原島, 藤井, 苗村: "3次元統合画像通信へむけて". *PCSJ '92*(1992-10).
- [11] 原島, 藤井, 苗村: "3次元統合画像通信の構想". 信学技報. HC92-58 (1993-01).

- [12] 藤井, 原島: “多眼式三次元画像符号化-視差補償予測符号化から構造抽出符号化へ”, *PCSJ '93*, 7-3 (1993-10).
- [13] 藤井, 原島: “アフィン変換を用いた多眼式三次元画像符号化”, *SITA '93*, T11-1 (1993-10).
- [14] 藤井, 原島: “3次元構造推定に基づく多眼3次元画像の圧縮と補間”, 3次元画像コンファレンス'94, (1994-7).
- [15] 藤井, 原島: “構造抽出に基づく三次元画像符号化” *PCSJ94* (1994-10).

国内大会

- [16] 藤井俊彰, 石間礼次, 濱崎襄二, “多眼式レンズ板三次元像のデータ圧縮の実験(テクスチャー解析とガウスフィルターの応用)”, 電子情報通信学会秋季全国大会 (1991-9).
- [17] 苗村, 藤井, 中屋, 原島: “三次元画像符号化基礎検討(1)-誤差評 Wed Dec 7 11:11:22 JST 1994 価法”, 1993年信学会春季全大, D-326 (1993-3).
- [18] 吉岡, 苗村, 藤井, 中屋, 原島: “三次元画像符号化基礎検討(2)-多眼画像の視差補償予測符号化-”, 1993年信学会春季全大, D-327 (1993-3).
- [19] 藤井, 原島, 濱崎: “三次元画像符号化の基礎検討(3)-多眼画像からの構造抽出-”, 1993年信学会春季全大, D-328 (1993-3).
- [20] 苗村, 吉岡, 藤井, 中屋, 原島: “多眼画像の視差補償予測符号化”, 1993 TV 学会年次大会, 24-4 (1993-7).
- [21] 藤井, 原島: “アフィン変換による多眼画像の視差補償予測符号化”, 1993年信学会秋季全大, D-200 (1993-9).
- [22] 嵯峨田, 藤井, 原島: “多眼画像の視差補償予測符号化-物体と背景の分離に基づく視差ベクトルの抽出と補間-”, 1994年信学会秋期全大, D-325 (1994-3).
- [23] 藤井, 原島: “多眼画像の視差補償予測符号化-セグメンテーションに基づく予測特性の改善-”, 1994年信学会秋期全大, D-326 (1994-3).

その他

- [24] 原島, 藤井, 苗村: “三次元統合画像通信へ向けて”, *ディスプレイアンド イメージング*, Vol. 1, pp. 121-131 (1992-12).
- [25] 藤井: “3次元画像からの視差検出-3次元形状の推定と帯域圧縮-”, 3次元映像, Vol. 8, No. 1 (1994).