

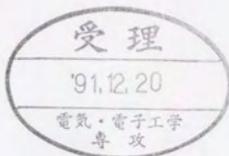
学 位 論 文

サブバンド符号化に基づく画像符号化の統一表現と
その特性評価に関する研究

平成3年 12月

甲 藤 二 郎

①



目次

サブバンド符号化に基づく画像符号化の統一表現と
その特性評価に関する研究

東京大学大学院工学系研究科 電子工学専攻 97082

甲藤 二郎

指導教官 安田 靖彦 教授

目次

1 序論	4
1.1 序	5
1.2 画像符号化方式の概略	11
1.2.1 予測符号化	11
1.2.2 変換符号化	13
1.2.3 サブバンド符号化	14
1.2.4 階層的符号化	22
1.3 数学的基盤	25
1.3.1 データ圧縮特性の理論評価尺度	25
1.3.2 サブバンド符号化における完全再構成条件	29
2 画像符号化の統一表現	33
2.1 行列の利用	34
2.1.1 サブバンド符号化の行列表現	34
2.1.2 直交性の検証	39
2.1.3 基底ベクトルの相互比較	44
2.2 MFBの利用	48
2.2.1 MFBによる統一表現	48
2.2.2 予測符号化の変形(内挿予測の概念の導入)	52
2.3 考察・検討	55
3 サブバンド符号化の特性評価	56

3.1	符号化ゲインに基づく特性評価	57
3.1.1	Coding gain の利用	57
3.1.2	AR(1) モデルに対する特性評価	57
3.2	UCG に基づく特性評価	63
3.2.1	MFB におけるビット割当ての最適化 (Unified coding gain の導出)	63
3.2.2	AR(1) モデルに対する特性評価	68
3.3	2次元のサブバンド符号化の特性評価	77
3.3.1	UCG の2次元への拡張	77
3.3.2	2次元入力モデルに対する特性評価	79
3.3.3	シミュレーション	85
3.4	考察・検討	94
4	完全再構成フィルタの設計	96
4.1	ハイパスフィルタの簡略化	97
4.1.1	1次元の場合	97
4.1.2	2次元の場合	98
4.1.3	具体的なフィルタ設計	100
4.1.4	考察・検討	101
4.2	数論変換に基づく完全再構成フィルタ	103
4.2.1	サブバンド符号化の一般化	103
4.2.2	数論変換に基づく完全再構成フィルタ	104
4.2.3	考察・検討	105
4.3	圧縮効率の最適化	106
4.3.1	問題の定式化	106
4.3.2	直線位相フィルタバンクの最適化	107
4.3.3	直交フィルタバンクの最適化	123
4.3.4	考察・検討	131

5	サブバンド符号化の ATM 用画像符号化への応用	132
5.1	ATM 用画像符号化について	133
5.2	サブバンド符号化の有効性	137
5.2.1	高能率符号化として	137
5.2.2	セル廃棄対策として	137
5.2.3	ユニバーサル符号化として	143
5.2.4	画質一定可変レート符号化として	143
5.3	シミュレーション	145
5.3.1	符号器の基本構成	145
5.3.2	リーク予測の効果	145
5.3.3	動き補償の導入	149
5.4	考察・検討	151
6	結論	152
	謝辞	155
	参考文献	156
	発表文献	163
	付録	166
A.1	DPCM の符号化ゲインの厳密な評価	166
A.2	パラメータ B_k の導出	167
A.3	多段接続の場合のパラメータ B_k の導出	167

Chapter 1

序論

映像情報といえば、テレビや映画に代表されるように、我々はそれを受け取るものと考えてきた。これは単に映像信号の情報量が莫大であり、それを通信システムに組み込むには技術がまだ未熟だったことに起因する。しかし、最近の技術の進歩は、ようやく我々に映像情報の自由な送受信を可能にしようとしている。この基盤となるのが、インフラストラクチャの整備、および画像信号の高エネルギー符号化技術である。

この高エネルギー符号化方式として近年注目を集めているのがサブバンド符号化である。このサブバンド符号化は、従来方式に対するさまざまな利点を有するだけでなく、広帯域 ISDN の情報転送方式である ATM に適した画像符号化方式としても有効である。しかし、その理論的な背景にはいくつか不明確な点があり、その普遍性に疑問の感じられる論旨もまま見受けられる。

本論文は、サブバンド符号化に対する圧縮効率の理論の構築を主要な骨子とする。画像符号化の統一表現はこの問題を解決するための道具として考えてもよいが、さまざまな符号化手法の相違点・相互関係を理解するのに有効である。そして本論文で与えるフィルタ構成は、フィルタ長を制約とした条件下において、圧縮効率としての観点からのサブバンド符号化の最適化を意味することになる。

1.1 序

画像信号の有する情報量はきわめて莫大である。例えばテレビ信号は1チャンネルで数MHzもの広帯域を占有するし、それをデジタル化すれば数十~数百Mb/sにも達することになる。一方、ISDN、さらには広帯域ISDNに向かうインフラストラクチャの整備は、画像メディアを通信システムに組み込むことを可能にする。しかし、音声と同様のユーザ間の自由な送受信を実現するには、その情報量はあまりにも大きすぎる。ただし、画像信号の相関の高さは古くから知られており、その統計的な性質を利用した帯域圧縮の可能性は今から40年以上も前に指摘されている。それ以来、画像信号の帯域圧縮、あるいはデジタル入力を前提としたデータ圧縮に関する研究が盛んに進められ、すでにFAXのように大規模な成功を収めている例もある。現在では、JPEG(カラー静止画符号化)、H.261(通信用動画像符号化)、MPEG(蓄積用動画像符号化)などの高能率符号化方式の国際標準化も終了しようとしており、いよいよ画像メディアの本格的な展開が実現されようとしている[8]。

画像信号の高能率符号化を支えてきたのは

- 予測符号化
- 変換符号化
- ベクトル量子化
- 階層的符号化 [22]-[25]

などのアルゴリズムである [1]-[9]。国際標準化も、これらの手法をベースとして進められてきた。しかし、これらの手法自体はどれも10年以上も前に提案されたものであり、研究対象としてはもはや成熟しきった感がある。何か、新しい符号化方法はないか？

現在、研究レベルで注目を集めているのが

- サブバンド符号化 [27]-[60]
- オーバーラップ変換 [61,62]
- wavelet 変換 [63]-[67]

などのアルゴリズムである [10]-[17]。それぞれが画像符号化方式として初めて論文誌に掲載された時期を見ると、サブバンド符号化が1986年 [44]、オーバーラップ変換が1989年 [61]、wavelet 変換が1988年 [63]となる。これらの方式はすべて優れた圧縮効率を実現するだけでなく、変換符号化の弱点であったブロックひずみ(blocking distortion)の問題を解決している。ブロックひずみの問題は特に低レート的环境下において顕著となり、ブロックひずみの排除は主観的な画質の向上に帰着することになる。

さらにサブバンド符号化や wavelet 変換では、コンピュータビジョンへの応用 [64,65] や ATM (asynchronous transfer mode) 用画像符号化への適用 [80,83] などが検討されている。まず、コンピュータビジョンへの応用の背景としては、サブバンド符号化の周波数分割の様子が、網膜応答の線形モデルである多チャンネルモデルに類似したものであることが指摘されている。あるいは、その解像度変換手法としての特徴を活かして、大まかな様子を表した低解像度画像からボトムアップ的に画像解析を進めていく階層処理手法の数学的な基盤となっている。このような手法は、多解像度解析 (multiresolution analysis) と呼ばれている。

一方、ATM 用画像符号化としては、ATM 環境において生じる諸問題の解決に、サブバンド符号化の有する周波数領域における階層構造が有効となることが指摘されている。この ATM は次世代の通信網である広帯域 ISDN の情報伝送手段として活発な検討が進められているものであり、従来の回線交換とは異なり、高速パケット交換を基盤としている。これによって可変レート符号化の採用などの利点が得られる一方、網のふくそう時のセルの廃棄などの新たな検討課題が生じることになる [73,74]。これに対して、サブバンド符号化の階層構造を利用してやれば、これらの課題をうまく解決することが可能である。

ただし、サブバンド符号化、オーバーラップ変換、wavelet 変換を分類すること自体は便宜的なものであり、決して本質的なものではない。言い替えば、これらの方式の間にはきわめて密接な関係があり、オーバーラップ変換と wavelet 変換は共にサブバンド符号化の一実現例として考えてもよい。さらにはまた、前述の変換符号化もサブバンド符号化に包含される概念であることが知られている [16,17,26]。

考えてみれば、画像符号化の多くは線形変換に基づいている。それに関わらず、このような相互関係が話題となるのは、予測符号化には線形予測モデル、変換符号化には直交行列、サブバンド符号化には z 変換などと、それぞれの方式ごとにそれぞれの理論展開に適した相異なる表現手法が用いられてきたことに起因する。これに対して画像符号化の統一表現を図ってやれば、各種方式間の相互関係・相違点が明確になるのみならず、それぞれに対して構築された理論体系の相互利用が可能となる。

画像符号化の目的は、言うまでもなく入力信号の冗長度の削減によるデータ圧縮の実現である。歴史の古い予測符号化や変換符号化では、これに対応するデータ圧縮の理論が確立されている [1]。最適予測係数や KLT (Karhunen Loeve transform) は、これらの理論に基づく最適解として与えられるものである。また DCT (discrete cosine transform) が国際標準化の主流として君臨していただけるのも、それに対する KLT の漸近性が理論的な背景として与えられていることが非常に影響している。

しかし、サブバンド符号化には、未だにこのようなデータ圧縮の理論が確立されていない。Jayant と Noll は理想的な周波数特性を持つフィルタバンクに対する圧縮効率の評価尺度を与えているが [1]。現実には理想フィルタを有限長のタップ数で実現することは不可能であり、あるいはまたその周波数特性に類似性が見られないさまざまなフィルタ構成の提案が行われている。このために、実際のフィルタ構成の圧縮効率を評価するには、シミュレーションに頼るしかないのが現状である。ただし、シミュレーションの結果は画像に依存したものであり、そこに普遍性を見出すことは難しい。

さまざまなフィルタ構成に対して、その圧縮効率に普遍性を見出すためには、予測符号化や変換符号化と同様のデータ圧縮の理論をサブバンド符号化に対しても構築することが必須である。そして前述のように、サブバンド符号化が従来の変換符号化、オーバーラップ変換、wavelet変換を包含したものであることから、この理論はこれらの符号化方式に対しても適用することが可能である。さらにはまた、サブバンド符号化と従来の予測符号化や変換符号化との組合せも、理論として評価を行うことが可能となるかもしれない。

一方、サブバンド符号化に対する圧縮効率の理論の応用として、フィルタ設計の問題が考えられる。これまでのサブバンド符号化のフィルタ設計の歴史を見てみると、そこでは周波数特性の最適化が設計基準として用いられているものが多い。ただし、DCTとDFT(discrete Fourier transform)の例を見るように、必ずしも急峻な周波数特性が圧縮効率の最適化を意味しているとは考え難い。これに対して、理論として与えられる圧縮効率の評価尺度を最大にするようなフィルタバンクを設計できれば、これはすなわちサブバンド符号化の圧縮効率の最適化を意味することになる。

本論文は、サブバンド符号化に対する圧縮効率の理論の構築を骨子とするものである。これに先立ち、各種画像符号化方式の統一表現を試みるが、これによってそれらの相互関係・相違点が明確になる。そして、その統一表現に対する圧縮効率の理論構築を図り、その結果として、すべての画像符号化に対して適用可能な圧縮効率の評価尺度を導出する。

続いてこの評価尺度に基づき、既存のフィルタ構成の特性比較と圧縮効率の最適化を目的とするフィルタ設計に関する検討を行う。前者の結果は、これまで曖昧なままに残されてきたサブバンド符号化の圧縮効率に対して定量的な説明を与えるものであり、また後者の結果は、従来の周波数特性とは異なり、圧縮効率を設計基準として導入した新しいフィルタ設計手法の提案となる。

本論文ではまた、サブバンド符号化の応用としてATM用画像符号化に関する検討も行う。ここではまずサブバンド符号化の有効性に対して定性的な説明を行い、さらにシミュレーションに基づく評価によって符号器の具体的な構成について言及する。

本論文の構成

本論文の構成を示したのが図1.1である。第1章は、序論として本論文の研究背景を述べている。次節である1.2節では、これまでに報告されている画像符号化方式の総括を行う。そして1.3節では、第2章以降の理論展開の数学的な基礎として、JayantとNollによって与えられている各種符号化方式の圧縮効率の理論的評価尺度である符号化ゲイン(coding gain)の説明、およびサブバンド符号化におけるフィルタ設計の基礎となる完全再構成条件(perfect reconstruction condition)の説明を行う。

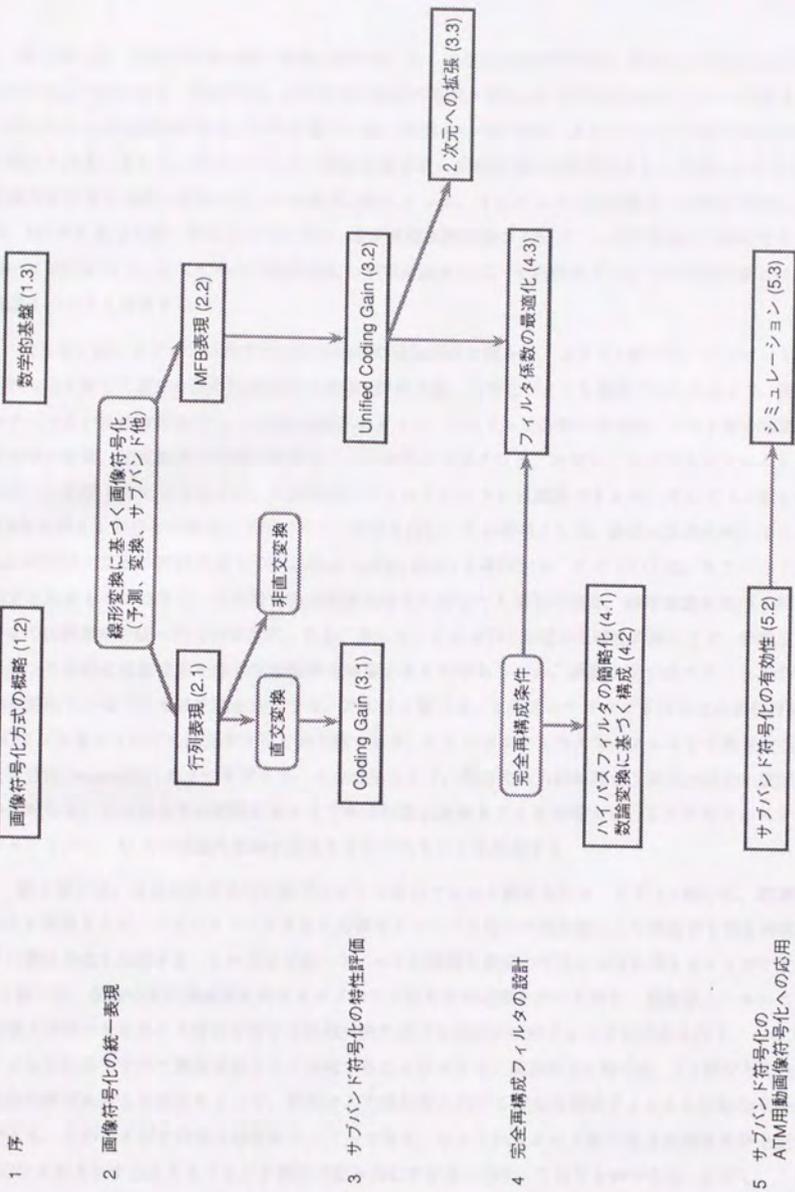


図1.1: 本論文の構成

1 序

2 画像符号化の統一表現

3 サブバンド符号化の特性評価

4 完全再構成フィルタの設計

5 サブバンド符号化の
ATM用動画像符号化への応用

第2章では、画像符号化の統一表現を試みる。ここで用いる表現手段は、行列とMFB(multirate filter bank)の2手法である。歴史的に、行列は主に変換符号化に対して、MFBは主にサブバンド符号化に対して用いられてきた経緯がある。行列に基づく統一表現(2.1節)では、まずサブバンド符号化の扱いについて検討を行う。そして、従来のフィルタ構成が直交系と非直交系に分類されることを明らかにする。また、各種画像符号化方式の基底ベクトルの相互比較によって、それぞれの方式の長所・短所を明確にする。一方、MFBに基づく統一表現(2.2節)では、まず多段接続に基づくサブバンド符号化と予測符号化の扱いについて検討を行う。そしてその予測符号化の解釈を基にして、並列型サブバンド符号化の新しいフィルタ構成についても言及する。

第3章では、サブバンド符号化の圧縮効率の理論解析を試みる。まず3.1節では、サブバンド符号化の行列表現に対して直交変換の圧縮効率の理論的評価尺度、符号化ゲインを適用することにより、直交系フィルタバンクの特性評価を行う。これらの結果によって、ピラミッド分割の妥当性、バンド数の増加に伴う圧縮効率の改善、直交変換の併用の効果などが定量的に実証される。ただし、この手法はフィルタの直交性の成立を前提としているために、非直交系のフィルタバンクには適用できない。そこで3.2節では、まずMFBのビット割当ての最適化問題について検討を行い、その結果として、非直交変換に対しても適用可能な圧縮効率の理論的評価尺度UCG(unified coding gain)を導出する。このUCGは、サブバンド符号化に限定されるものではなく、予測符号化や変換符号化に対しても適用可能な、線形変換に基づく画像符号化方式の圧縮効率の統一的な評価尺度となる。そして、このUCGを用いた特性評価により、非直交系フィルタバンクの中には直交系を凌ぐ圧縮効率を実現するものがあること、非直交系ではピラミッド分割が絶対的に優れていることなどを明らかにする。次に3.3節では、2次元のサブバンド符号化の特性評価を行う。UCGは容易に2次元に拡張することが可能であり、ここでは2次元の入力モデルとして等方性(isotropic)と可分性(separable)の2つを考える。その結果として、等方性入力の結果が1次元の場合に類似しているのに対して、可分性入力の結果にはいくつかの相違点があることを示唆する。ここではさらにシミュレーションを行い、以上の理論的推測が妥当なものであることを確認する。

第4章では、完全再構成条件に基づくいくつかのフィルタ設計を行う。まず4.1節では、階層的符号化からの類推として、ハイパスフィルタを入力信号とローパス出力の差分器として構成する完全再構成フィルタの構成方式を提案する。この方式では、フィルタの総数を従来の手法の半分に抑えることができる。次に4.2節では、任意の直交関数系に対するサブバンド符号化の定義を行った後に、整数環上においてフーリエ変換と同様のたたみこみ特性を有する数論変換に基づく完全再構成フィルタの提案を行う。この方式では、フィルタ処理をすべて整数演算として実現することができる。最後に4.3節では、3.2節で与えたUCGと完全再構成条件との結合によって、特定の入力情報源に対する完全再構成フィルタの圧縮効率の最適化を試みる。そのアイデア自体は非常にシンプルであり、与えられたタップ数と完全再構成条件を制約としてUCGを最大にするようなフィルタ係数の組み合わせを見い出す、というものである。ただし、これらの操作は予測符号化における最適予測係数、変換符号化におけるKLTを求める操作に等価であり、“サブバンド符号化の圧縮効率の理論的な最適化”を意味している。

第5章では、サブバンド符号化の応用としての ATM 用動画符号化に関する検討を行う。広帯域 ISDN の情報伝送手段である ATM(asynchronous transfer mode) は高速パケット交換を基盤としたものであり、従来の回線交換とは異なる点が多い。このために、符号化方式も従来とは異なった構成を採ることになる。これに対して、サブバンド符号化は ATM との親和性が高く、セル廃棄対策、ユニバーサル符号化など、長短問わず注目を集めている。ここではまず ATM 用画像符号化としてのサブバンド符号化の有効性に関する定性的な説明を行い、さらにシミュレーションを併用しつつ、リーク予測の導入や動き補償の導入方法などについて検討を進めていく。本章は基本的に筆者の修士論文 [94] の延長線上に位置づけられるものである。

最後に第6章では、本論文の結論を述べる。

1.2 画像符号化方式の概略

画像信号の高効率符号化に関する研究の歴史は長く、古くから予測符号化、変換符号化、ベクトル量子化などの方式が検討対象とされてきた。そして最近では、その優れた圧縮特性に加えて、ブロックひずみの排除による主観的な画質の向上効果が期待されるサブバンド符号化やオーバーラップ変換 (LOT, MLT) などの方式が注目を集めている。

図 1.2 は画像符号化の変遷の歴史をまとめたものである。ただし、ベクトル量子化は、その本質が信号の変換ではなく量子化に置かれたものであるから、ここでは省略している。この図が示すように、80 年代後半になってからさまざまな新方式の提案が行われている。

ここでは画像符号化を、変換符号化、サブバンド符号化 (並列、多段接続)、階層的符号化、予測符号化の 4 系統に分類したが、これらが全く異なる手法であることを意味するわけではない。図中に矢印で示したように、LOT と MLT、および CQF と wavelet は、それぞれの具体的なフィルタ係数は異なるものの、ある程度の制約条件下では等価なものと考えてよい。すなわち、LOT と MLT は $2M \times M$ ブロックの正規直交基底として、CQF と wavelet は 2 バンド分割に帰着する正規直交基底として、それぞれ位置づけることが可能である。ただし、それぞれの方式を理解する上でこのような分類が有効であることは間違いがなく、以下これらの符号化方式の概括を行っていく。

1.2.1 予測符号化

静止画像は統計的に、空間的に隣接する画素間 (フレーム内) の相関が強く、“空間的冗長度が大きい”ことが知られている。動画画像信号も、画面の動きがあまり激しくない場合には、時間的に隣接する画素間 (フレーム間) の相関が強く、“時間的冗長度が大きい”ことが知られている。これらの冗長性を利用して、すでに符号化された画素の値から次に符号化すべき画素の値を予測し、その予測誤差を符号化することによって高効率符号化を実現することができる。このような考えに基づく方式は予測符号化 (predictive coding) と呼ばれる。

予測には一般的に

$$\hat{x}(n) = \sum_{j=1}^N h_j x(n-j) \quad (1.1)$$

として線形予測が利用されることが多い。ここで h_j は線形予測係数であり、見方を変えれば FIR フィルタのインパルス応答にはかならない。そして、予測誤差 $d(n) (= x(n) - \hat{x}(n))$ の分散を最小にする h_j を求めることにより、データ圧縮としての観点からの予測符号化の最適化が図られる。

ただし、このままでは復元信号に量子化誤差の蓄積が生じてしまうために、圧縮の効果がほとんど得られない。そこで符号器にフィードバックループとして局部復号器を挿入し、量子化済みの画素値を用いて予測するのが普通である。局部復号器を挿入しない方式は開ループ (open-loop) DPCM と呼ばれ、挿入す

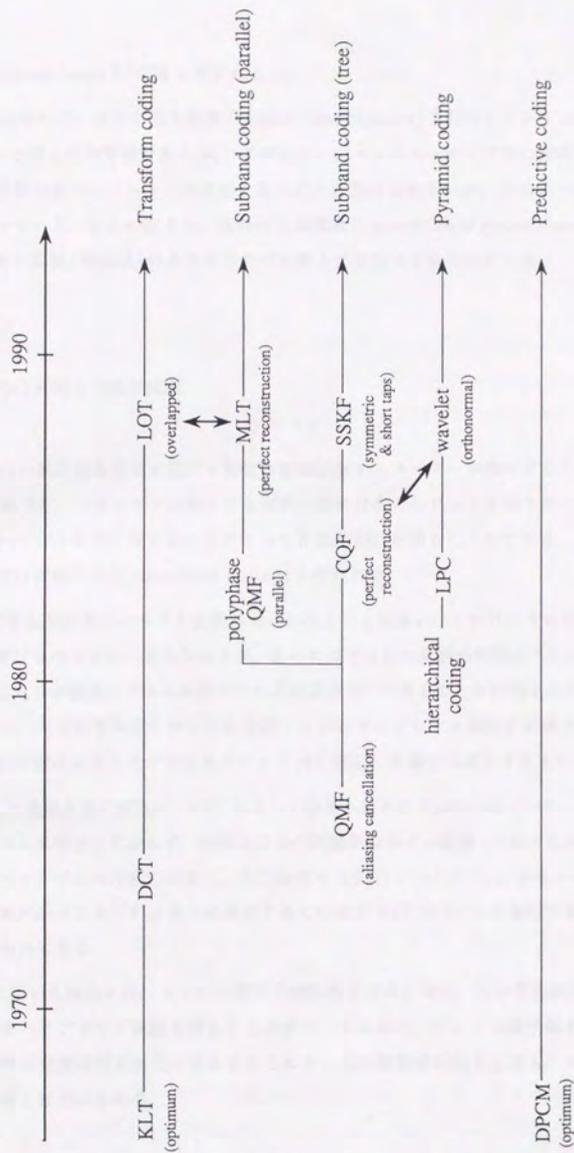


図1.2: 画像符号化の変遷

る方式は閉ループ (closed-loop) DPCM と呼ばれる [1].

対象が動画像の場合には、さらに動き補償 (motion compensation) を行ってフレーム間の相関を増加させ、圧縮効率の向上を図るのが普通である [6]. この場合、ブロックマッチング等の方式を用いる場合には動きベクトルの送信分のオーバーヘッドが必要となることは避けられないが、そのオーバーヘッドを上回る圧縮効率を得られている。あるいはまた、条件付き画素補充 (conditional replenishment) のように、予測が十分に行われない領域 (動領域) のみを符号化の対象とする方式も知られている。

1.2.2 変換符号化

ある信号系列 $x(n)$ に対して線形変換

$$y = Hx \quad (1.2)$$

を施した場合、 $x(n)$ の統計的な性質に応じて特定の変換係数にエネルギーが集中することがある。このような線形変換 H に基づき、エネルギーが集中する変換係数には多くのビットを割り当て、逆に集中しない変換係数には少ないビットを割り当てることによって冗長度削減を図ることができる。このような考えに基づく高能率符号化は変換符号化 (transform coding) と呼ばれる。

変換手法としては直交変換とユニタリ変換が用いられることが多い。これは、その数学的な扱い易さが一番の理由として挙げられるものと考えられるが、さらにはそれらの変換の特徴が“入力信号を互いに無相関な成分に分解するための基底ベクトルに求められる必要条件”であることが指摘されている [1]. ただし、計算時間の問題から、入力信号を小ブロックに分割してからそのブロック単位に変換を施しているのが現状であり、上記の無相関化はあくまで対象をブロック内に限定した場合に成立するものである。

データ圧縮として最適な直交変換は、KLT によって与えられることが知られている。ただし、この KLT には高速アルゴリズムが存在しておらず、実用上は DCT (離散コサイン変換) が用いられることが多い。これは、その高速アルゴリズムの存在に加えて、入力信号を AR(1) プロセス (1 次ガウスマルコフ情報源) と仮定した場合に、モデルパラメータ ρ が 1 に漸近するにつれて KLT が DCT に漸近することが理論的に明らかにされているためである。

動画像符号化に用いる場合には、フレーム間の予測誤差を求めた後に、その予測誤差に対して DCT 等の変換符号化を施すハイブリッド構成を採ることが多い。このほか、フレーム間予測を十分に行えない動領域ではフレーム内の変換符号化に切り替えることによって圧縮効率の向上を図る、フレーム間/フレーム内適応予測もしばしば用いられる。

オーバーラップ変換

KLT や DCT では、画像全体を小ブロック (8×8、あるいは 16×16 程度) に分割した後、そのブロックごとに変換を行う。これは計算量の増大を抑えるためであるが、このような“ブロック変換”ではブロックと

しては最適な符号化が施されたとしても、画像全体としては、特に低レート環境において、そのブロックの境界部が目につくことがある。このようなブロック変換特有のひずみのことを、“ブロックひずみ” (blocking distortion) と呼ぶ。

ブロックひずみを低減する方式としては、復元画像にローパスフィルタを掛けたり、あるいは境界部をオーバーラップさせた上でブロック変換を施す、などの方式が提案されてきた。しかし、前者の場合にはブロックの境界部ではない部分まで歪んでしまい、また後者の場合には符号化の対象となる画素数が増加するという問題があった。

オーバーラップ変換は、これらの問題を起こすことなくブロックひずみを解決する変換手法として、近年注目を集めているものである。この場合、ブロック境界部をオーバーラップさせた上で変換を施すのは上述の後者の方式と同じであるが、その変換を工夫することによって変換出力の総数が原画像のそれに一致するようにしている。具体例としては LOT (lapped orthogonal transform) [61], MLT (modulated lapped transform) [62] などが知られており、これらは共にウィンドウ (FIR フィルタ) の長さを $2N$ とすると、その変換出力の数が N であるような直交変換となっている。そして受信側では、 N 個の入力に対して出力数が $2N$ であるような合成フィルタが導入される。

1.2.3 サブバンド符号化

入力信号をバンドパスフィルタによって複数の周波数領域に分割した場合、変換符号化の場合と同様に、入力信号の統計的な性質に応じて特定の周波数帯域にエネルギーが集中する。よって、このエネルギーが集中する周波数帯域に多くのビットを割り当て、逆に集中しない周波数帯域には少ないビットを割り当てることにより、変換符号化と同様の冗長度削減を図ることができる。このような考えに基づく高能率符号化をサブバンド符号化 (subband coding) と呼ぶ。

サブバンド符号化の最も基本的なシステム構成は、図 1.3 のような 1 次元 2 バンドの“分析合成” (analysis/synthesis) システムとして表される。ここでそれぞれ、 $x(n)$ は入力、 $h_i(n)$ は送信側の分析フィルタ、下向きの矢印は 2:1 のサブサンプリング (decimation)、上向きの矢印は 1:2 の 0 値補間 (interpolation)、 $g_i(n)$ は受信側の合成フィルタ、 $\hat{x}(n)$ は最終的な復元信号を表している ($i=0,1$)。

一般的に $h_0(n)$ はローパスフィルタ、 $h_1(n)$ はハイパスフィルタを表しており、これらのフィルタ処理によって入力信号は低周波成分と高周波成分とに分離される。ここで理想フィルタを用いた場合には、それぞれの信号系列の周波数帯域は入力信号の周波数帯域の半分になるから、標準化定理によってそれぞれの信号を 2:1 にサブサンプリングすることが許容される。このサブサンプリングの導入によって、符号化の対象となる信号の数が原信号のそれよりも多くなることはない。一方、受信側ではまず 0 値補間操作を行うが、この操作は周波数領域では帯域の拡大を意味する (中間周波数を軸として、周波数帯域が折り返す)。そこでローパスフィルタ $g_0(n)$ とハイパスフィルタ $g_1(n)$ によって、それぞれのチャンネルのサブサンプリング前の周波数帯域の回復を図り、最後にこれらの信号系列の和を取ることによって (原信号の周波数帯域の

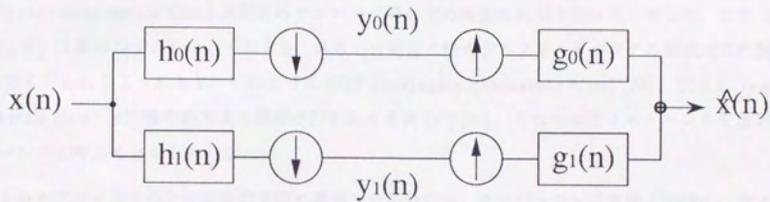


図 1.3: 1次元 2 バンド帯域分割システム

復元), 原信号の再生を実現する.

以上の操作で問題となるのは, 有限タップ数のフィルタでは理想的な周波数特性を実現できないことにある. このために, 復元信号 $\hat{x}(n)$ には一般的に

- 折り返し (aliasing)
- 振幅 (amplitude)
- 位相 (phase)

の3種類のひずみが発生することになる.

ただし, フィルタ間にある関係式が成立すれば, 折り返しひずみは完全にキャンセルすることが可能である (1.3.2 節参照). これを最初に理論として明らかにしたのが Esteban らであり [27], ここで提案された QMF (quadrature mirror filter) が現在のサブバンド符号化の隆盛の発端となった. そして, この QMF に端を発して, 上記のひずみを小さく抑える, あるいは完全に除去するフィルタバンクの構成方式に関する検討が活発に行われるようになり, これまでに CQF (conjugate quadrature filter) [33], SSKF (symmetric short kernel filter) [47] 等の新方式の提案が行われてきた [27]-[36]. これらのフィルタバンクを総称として QMF バンクと呼ぶこともある [13].

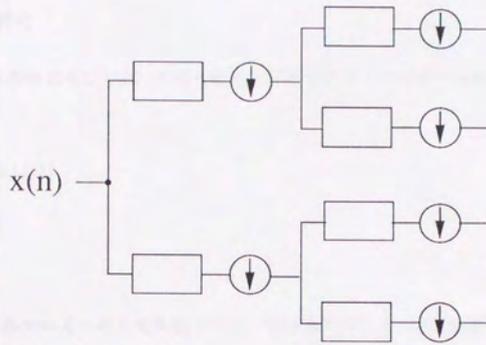
以上のサブバンド符号化を画像符号化に適用した場合には, オーバーラップ変換と同様に, ウィンドウ (フィルタ) が互いにオーバーラップしているために, ブロックひずみが生じないと言う利点が得られる. あるいはその 2:1 のサブサンプリング操作を利用して, 解像度変換に利用することも可能である. これは, コンピュータビジョンにおける多解像度解析や映像サービスに依存しない符号化, すなわちユニバーサル符号化の実現に向けて非常に有効な特性である. あるいはまた, その周波数領域における階層構造を利用して, ATM におけるセル廃棄対策としての応用も盛んに検討されている.

Tree structure

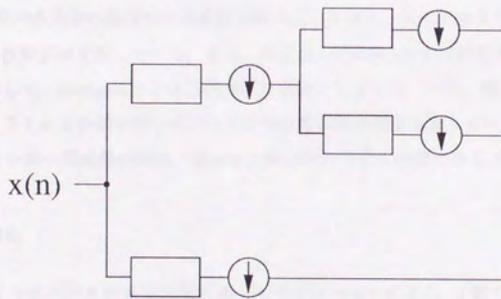
サブバンド符号化を用いる場合に, 図 1.3 のような 2 バンド分割だけで用いられることはほとんどない. 一般的には, 予測符号化や直交変換と組み合わせられるか, あるいは図 1.4 のような多段接続によって複数のバンド分割を図る. ここで図 1.4 のような構造は “木構造” (tree structure) と呼ばれる.

木構造周波数分割にはさまざまな形態が考えられるが, 図 1.4(a) のように分割できる信号系列はすべて再分割する方式を, 本論文では “フル分割” と呼ぶ. また, 図 1.4(b) のように低周波成分だけを再分割する方式を, 本論文では “ピラミッド分割” と呼ぶ (Akansu らは前者を規則的 (regular), 後者を不規則的 (irregular) と呼んでいる [55]).

後者のピラミッド分割は, その周波数分割の様子から後述する階層的符号化と等価な概念であると考え



(a) Full decomposition (regular)



(b) Pyramid decomposition (irregular)

图 1.4: 木構造周波数分割

られる。そこで、その周波数領域における階層構造を利用して、セル廃棄対策、あるいはユニバーサル符号化等の ATM 用動画符号化としての応用も盛んに検討されている。

2次元のサブバンド符号化

2次元のサブバンド符号化としては、その2次元サブサンプリングパターンの相違に基づいて、これまで

- 可分型 (separable) [43]
- quincunx [43,45]
- hexagonal [45]

の3種類の方法が知られている。ここで後者2つは、可分型に対して“非可分型” (non-separable) と総称されることがある。これは、可分型の場合には1次元フィルタ $\{h(n)\}$ から

$$h(m, n) = h(m) \cdot h(n) \quad (1.3)$$

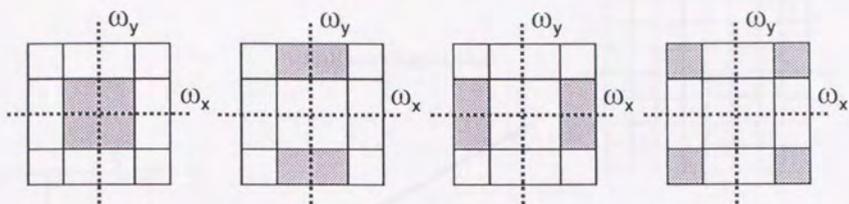
として2次元のフィルタ係数が決定されるのに対して、非可分型の場合にはこのような関係が成立しないことによる。

2次元のサブバンド符号化の周波数分割の様子を、理想フィルタを前提とした場合について示したのが図1.5である。最も外側の多角形が原信号の周波数帯域を表しており、灰色に塗りつぶした領域が1回の周波数分割で得られる周波数領域を表している。なお、可分型と quincunx では原信号の標本化に正方向格子を前提としているのに対して、hexagonal では三角格子を前提としている。一方、図1.6には、可分型でフル分割を行った場合とピラミッド分割を行った場合との周波数分割の相違を表している (縦列段数=2)。前述のように、ピラミッド分割の周波数分割は、後述する階層的符号化に等価である (図1.7参照)。

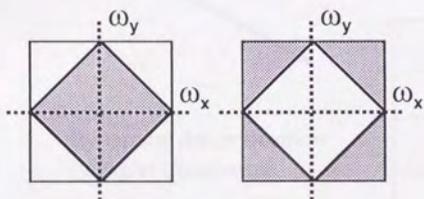
並列型サブバンド符号化

図1.4のようなフィルタバンクの多段接続に基づく周波数分割ではなく、1段で並列的に $M(> 2)$ 個のバンド分割を行うサブバンド符号化に関する検討も盛んに行われている [37]-[42]。このような並列型サブバンド符号化の場合には、図1.3の一般化として、図1.8によってシステム表現が図られることが多い。この図1.8はマルチレートフィルタバンク (MFB) と呼ばれており、 $h_i(n)$ 、 $g_i(n)$ は共にバンドパスフィルタを表すものと考えてよい。ただし、それが図のように K チャンネル構造を採っていても、サブサンプリングと0値補間がそれぞれ $K:1$ 、および $1:K$ となるとは限らない。

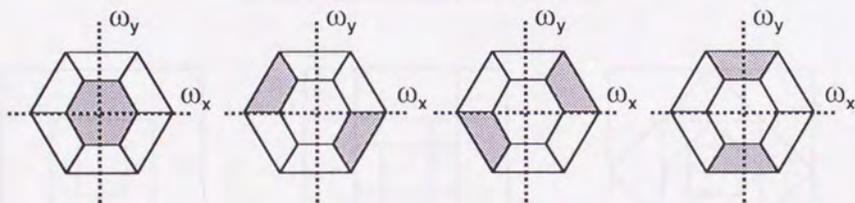
オーバーラップ変換のところで述べた MLT は、その発端を辿れば、並列型サブバンド符号化として検討が開始された変換方式である。言い替えれば、並列型サブバンド符号化とオーバーラップ変換は等価な



(a) Separable



(b) Quincunx



(c) Hexagonal

図 1.5: 2次元のサブバンド符号化

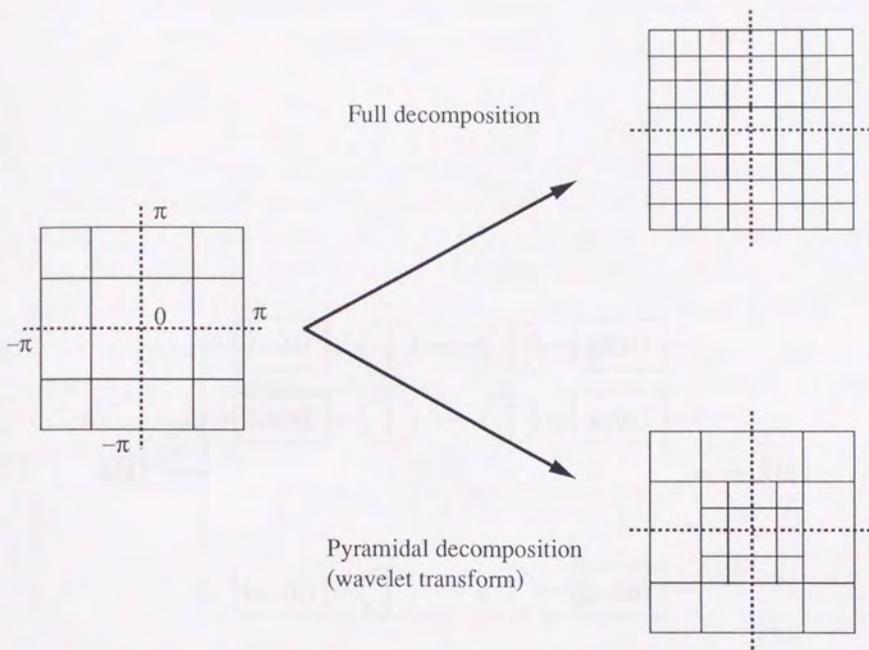
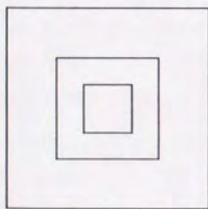
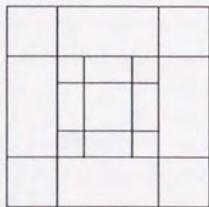


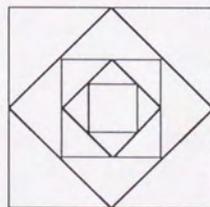
図 1.6: 2次元のフル分割とピラミッド分割



(a) Laplacian pyramid



(b) Separable



(c) Quincunx

図 1.7: 階層的符号化とピラミッド分割

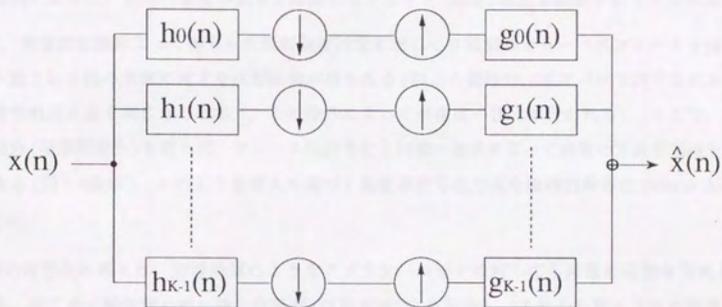


図 1.8: マルチレートフィルタバンク

ものであり、その区別は余り意味をなさない。あるいはまた、従来のブロック直交変換も MFB に包含されるものであり、最近ではこれらすべてをサブバンド符号化と呼ぶ人々もいる。

1.2.4 階層的符号化

入力画像に対してローパスフィルタとサブサンプリング操作を繰り返し施してやれば、サブサンプリングの効果によって、大幅に情報量の削減された低解像度画像が得られる。これに対して、画像信号のエネルギーは一般的に低周波成分に集中しており、上記の操作によって得られた低解像度画像からも、原画像の大体の様子は把握することができる。以上のような操作は、“解像度変換”と総称される。そして、この解像度変換によって、空間解像度の異なる画像のピラミッド(階層)構造が構成されることになる。

一方、解像度変換によって得られた低解像度画像に対して0値補間とローパスフィルタを施すと、解像度変換が施される前の画像に対する近似画像が得られる(以上の操作は、サブバンド符号化における1チャネルの信号処理と全く同じのものであり、この操作によって解像度の復元が行われる)。そこで、これらの画像間の差分(階層間差分)を取れば、フレーム間符号化と同様の効果によって画像の冗長度削減を実現することができる(図1.9参照)。このような考えに基づく高能率符号化方式を階層的符号化(hierarchical coding)と呼ぶ[22]。

階層的符号化の考えは、画像検索のようなアプリケーションに対しても非常に有効な方式となりうる。すなわち、はじめに解像度の粗い最上位階層(低周波成分)を送信し、それから順に下位の階層を送信することによって原画像の復元を図る。受信側では、最初の低解像度画像から徐々に解像度が改善されて行くことになり、情報量の少ない伝送の初期段階において、その画像の取捨選択を決定することができる。このような画像伝送方式は、“段階的伝送”(progressive transmission)と呼ばれる。

階層的符号化の意味を周波数領域で考えると、図1.7(a)のようになる。最も内側の正方形の領域が最上位階層の周波数帯域を表しており、階層間の差分情報の周波数帯域は正方形間の差分領域として表される。一方、図1.7(a)、(b)はそれぞれ、可分型と quincunx においてピラミッド分割を行った場合の周波数分割の様子を表している。これらの図より、階層的符号化における低周波成分の処理はサブバンド符号化のそれと全く等価であることがわかる。

ただし、サブバンド符号化とは異なり、階層的符号化では符号化の対象となる画素の数が原信号のそれよりも増加するという問題がある。これは高周波成分(階層間差分情報)のオーバーサンプリングに起因しており、原画像の画素数に対する各階層間差分情報の符号化の対象となる画素数の割合が、全周波数帯域に対して階層間差分情報の占める面積の割合よりも大きくなる問題として理解される。

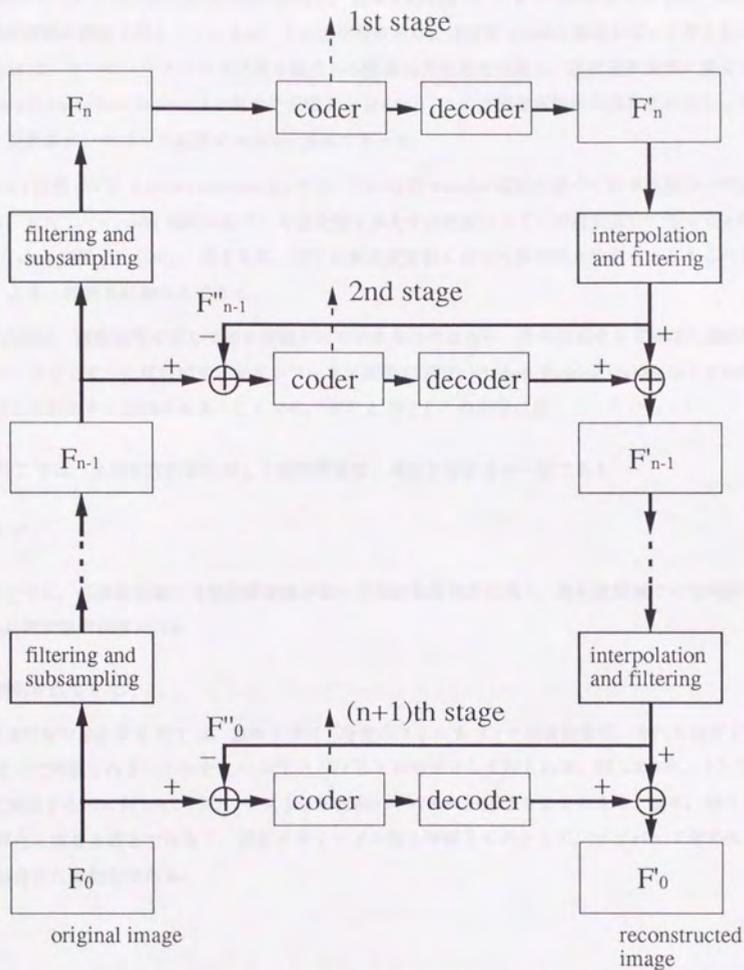


图 1.9: 階層的符号化

Wavelet transform

階層的符号化における画素数の増加の問題は、近年 Daubechies によって解決された [63]。ここで彼女は直交関数展開の理論を導入しているが、その直交関数系は正規直交 wavelet 基底によって与えられる。一方、Mallat は、コンピュータビジョン的な観点から階層的符号化を見直し、直交関数展開に基づく多解像度解析 (multiresolution analysis) の概念を提唱した [64,65]。この多解像度解析の具体化に当たって用いられた直交関数系が、やはり正規直交 wavelet 基底であった。

Wavelet 変換 (WT: wavelet transform) とは、これらの wavelet 基底に基づく線形変換の一手法を意味している。ただし、wavelet 基底は必ずしも直交性を満たす必然性はなく、非直交基底であっても構わない (例えば Gabor 変換 [57]-[60])。要するに、WT は解像度変換に対して数学的な枠組みを与えるのに非常に有効な、より一般的な理論体系である。

WT 自体は、離散信号に対してのみ定義されているものではない。その起源をたどれば、連続信号の変換手法の一つとして、しばしばウィンドウフーリエ変換 (WFT: window Fourier transform) を比較の対象として論じられてきた経緯がある。ここでは、WT と WFT との相違点は

- WFT では、全周波数領域に対して空間解像度と周波数解像度が一定である

のに対して

- WT では、低周波領域では空間解像度が低い周波数解像度は高く、高周波領域では空間解像度は高い周波数解像度は低い

として説明されてきた。

離散信号を対象とする WT は、常に 2 バンド分割のフィルタバンクの多段接続、それも必ずピラミッド分割によって実現される。これを上の WT と WFT との対比として捉えれば、図 1.6 の右上 (フル分割) が WFT に対応するのに対して、右下 (ピラミッド分割) が WT に対応することになる。故に、WT はサブバンド符号化と異なる概念ではなく、逆にピラミッド分割と等価なものとして、サブバンド符号化の枠組みの中に包含される概念である。

1.3 数学的基盤

1.3.1 データ圧縮特性の理論評価尺度

これまで各種画像符号化方式の概括を行ってきたが、それではこれらの符号化方式によって、どれほどのデータ圧縮が図れるのであろうか？ シミュレーションを用いた検証も一つの結果ではあるが、その結果は画像に依存したものであり、決して普遍的なものではない。

Jayant と Noll が示した符号化ゲイン (coding gain) G は、ある信号系列をある一定の伝送レートで符号化する場合に、変換 TR を用いた時の再生誤差の分散 $\sigma_{r,TR}^2$ と PCM を用いた時の再生誤差の分散 $\sigma_{r,PCM}^2$ の比

$$G = \frac{\sigma_{r,PCM}^2}{\sigma_{r,TR}^2} \quad (1.4)$$

として定義される [1]。この値が 1 よりも大きければ、同じレートで符号化したときに PCM よりも小さな再生誤差を実現できることになり、変換 TR がデータ圧縮手法として有効な方式であることを意味する。逆に、 G が 1 よりも小さければ、変換 TR にはデータ圧縮としての効果が望めないことを意味する。

よって、入力情報源のモデルを規定し、そのモデルに対する符号化ゲインの値を算出することにより、変換 TR のデータ圧縮の効果を理論的、かつ定量的に評価することができる。Jayant らは予測符号化、変換符号化、およびサブバンド符号化の場合について、より具体的な形で以下のような符号化ゲインを定義している。

予測符号化の場合

入力信号を $x(n)$ 、予測誤差を $d(n)$ 、予測誤差の量子化誤差を $q(n)$ 、再生誤差を $r(n)$ とする。ここで閉ループ DPCM の場合には量子化誤差の蓄積が起こらず

$$\sigma_{r,DPCM}^2 = \sigma_q^2 \quad (1.5)$$

が成立する。

次に、1 次の線形予測

$$d(n) = x(n) - h \cdot x(n-1) \quad (1.6)$$

を考える。ここで h は予測係数を表す。このとき、予測誤差の分散 σ_d^2 は、入力 $x(n)$ が安定した相関特性を示す場合には

$$\sigma_d^2 = (1 + h^2 - 2\rho \cdot h)\sigma_x^2 \quad (1.7)$$

として与えられる。ここで ρ は

$$\rho = \frac{E[x(n)x(n+1)]}{\sigma_x^2} \quad (1.8)$$

によって与えられ、 $x(n)$ の1次の自己相関係数を表している。

次に伝送レートを R とし、この条件下で予測誤差を量子化することを考える。このとき、Jayant らは、予測誤差とその量子化誤差の間に

$$\sigma_q^2 \simeq \epsilon^2 2^{-2R} \sigma_d^2 \quad (1.9)$$

の関係式が成立することを明らかにした。ただし ϵ は量子化器入力の特徴とその量子化方法に応じて決定される定数である。この関係は、入力信号がそのまま量子化される PCM の場合も同様に成立し、

$$\sigma_{r,PCM}^2 \simeq \epsilon^2 2^{-2R} \sigma_x^2 \quad (1.10)$$

の関係式が成立する。

よって(1.4)式に基づき、1次の閉ループ DPCM の符号化ゲイン G_{DPCM} が

$$G_{DPCM} = (1 + h^2 - 2\rho \cdot h)^{-1} \quad (1.11)$$

によって与えられることになる。これが $N(>1)$ 次の場合についても同様に求められることは自明である。

- 1次の閉ループ DPCM に対する G_{DPCM} は、 $h = \rho$ のときに最大値

$$G_{DPCM,max} = (1 - \rho^2)^{-1} \quad (1.12)$$

を取る。この値は、入力情報源を AR(1) プロセスでモデル化したときの理論的最適値を与えるものである。

- 厳密に言えば、(1.12) 式に示した理論的最適値は、符号化レートが十分に大きくなければ実現できない。これは、(1.5) 式が閉ループ DPCM に対して定義されたものであるのに対して、(1.7) 式が開ループ DPCM に対して定義されたものであることに起因する。このために、閉ループ DPCM は高レートでは理論的最適値に近い特性を示すが、低レートでは特性が大きく劣化する現象が見られる (Appendix A.1)。

直交変換の場合

直交変換の場合、入力信号の分散 σ_x^2 と変換係数の分散 σ_k^2 の和は等しく、

$$\sigma_x^2 = \frac{1}{N} \sum_{k=0}^{N-1} \sigma_k^2 \quad (1.13)$$

が成立する。ここで N は行列のサイズを表している。同様に、再生誤差の分散 $\sigma_{r,TC}^2$ と各変換係数の量子化誤差の分散 σ_{qk}^2 の和も等しく、

$$\sigma_{r,TC}^2 = \frac{1}{N} \sum_{k=0}^{N-1} \sigma_{qk}^2 \quad (1.14)$$

が成立する。非直交変換の場合には、これらの関係式は成立しない。

次に、各変換係数に対して割当てられたビット数を R_k とし、次のようなビット割当ての最適化問題を考える。

$$\sum_{k=0}^{N-1} R_k = R(\text{const.}) \quad (1.15)$$

を制約条件とした

$$\sigma_{r,TC}^2 = \frac{1}{N} \sum_{k=0}^{N-1} \sigma_{q_k}^2 \quad (1.16)$$

の最小化、

ここで量子化器入力 (各変換係数) とその量子化誤差の関係式

$$\sigma_{q_k}^2 \simeq \epsilon^2 2^{-2R_k} \sigma_k^2 \quad (1.17)$$

を考え、さらにラグランジュの未定乗数法を用いてこの最適化問題を解くと、再生誤差分散の最小値が次式によって与えられる。

$$\min\{\sigma_r^2\} = \epsilon^2 2^{-2R} \cdot \left[\prod_{k=0}^{N-1} \sigma_k^2 \right] \quad (1.18)$$

よって、直交変換の符号化ゲイン G_{TC} が次式によって定義される。

$$G_{TC} = \frac{\frac{1}{N} \sum_{k=0}^{N-1} \sigma_k^2}{\left(\prod_{k=0}^{N-1} \sigma_k^2 \right)^{\frac{1}{N}}} \quad (1.19)$$

一方、変換係数の分散 σ_k^2 は、

$$R_{\theta\theta} = H R_{xx} H^t \quad (1.20)$$

の対角成分として求めることができる。ただし、 H は変換行列としてサイズ $N \times N$ の直交 (or ユニタリ) 行列を表しており、 R_{xx} は入力信号の自己相関行列、 $R_{\theta\theta}$ は変換係数の自己相関行列をそれぞれ表している。そこで、入力信号のモデルを自己相関行列 R_{xx} として与えることにより、その入力モデルに対する直交変換 H の圧縮効率を (1.19) 式によって定量的に評価できる。

- (1.19) 式を最大にする直交変換は、KLT であることが知られている。さらには、入力信号を AR(1) プロセスによってモデル化する場合、 $N \rightarrow \infty$ とするにつれて、 G_{KLT} が理論的最適値 (1.12) 式に漸近することが明らかにされている [1]。

サブバンド符号化の場合

入力信号を M 個のサブバンドに分割する場合を考える。ここでそれぞれのバンドパスフィルタが理想的な周波数特性を実現すると仮定すると、バンド分割後のエネルギー保存式

$$\sigma_x^2 = \frac{1}{M} \sum_{k=0}^{M-1} \sigma_k^2 \quad (1.21)$$

、および量子化誤差分散のエネルギー保存式

$$\sigma_{r,SBC}^2 = \frac{1}{M} \sum_{k=0}^{M-1} \sigma_{qk}^2 \quad (1.22)$$

等の関係式が成立するものと考えてよい。以下は直交変換の場合と同様の計算によって、サブバンド符号化の符号化ゲイン G_{SBC} が

$$G_{SBC} = \frac{\frac{1}{M} \sum_{k=0}^{M-1} \sigma_k^2}{\left(\prod_{k=0}^{M-1} \sigma_k^2 \right)^{\frac{1}{M}}} \quad (1.23)$$

によって与えられる。ここに σ_k^2 は、各サブバンドの分散を表している。

- 入力信号を AR(1) プロセスによってモデル化する場合、直交変換の場合と同様に $M \rightarrow \infty$ とすれば、 G_{SBC} は理論的最適値 (1.12) 式に漸近する [44]。
- G_{SBC} の定義では、すべてのフィルタが理想的な周波数特性を有することを前提としている。しかし、現実には有限タップ数の制約のため、理想的な周波数特性を実現することは不可能である。あるいはまた、具体的なフィルタ構成としても、その周波数特性に類似性が見られない多種多様な方式が提案されている。故に、 G_{SBC} に代わるサブバンド符号化の圧縮効率の評価尺度の見直しが急務となる。

1.3.2 サブバンド符号化における完全再構成条件

図 1.3 に示した 1 次元 2 バンドのサブバンドシステムを考える。前述のように、理想的な周波数特性を呈するフィルタを有限のタップ数で実現することは不可能であり、一般的には復元信号に 3 種類のひずみ、折返し、振幅、位相ひずみが発生する。ただし、フィルタ群 $\{h_0(n), h_1(n), g_0(n), g_1(n)\}$ の間に以下に示す関係式 (完全再構成条件) が成立する場合には、復元信号を原信号に完全に一致させることが可能となる。

まずサブサンプリングパターンが上下チャンネルで一致する場合、そのサブサンプリング関数が上下チャンネルで等しく

$$f_0(n) = f_1(n) = \frac{1 + (-1)^n}{2} \quad (1.24)$$

によって与えられることから、復元信号 $\hat{x}(n)$ は z 変換領域において

$$\begin{aligned} \hat{X}(z) &= \frac{1}{2}[H_0(z)G_0(z) + H_1(z)G_1(z)]X(z) \\ &+ \frac{1}{2}[H_0(-z)G_0(z) + H_1(-z)G_1(z)]X(-z) \end{aligned} \quad (1.25)$$

と表される。ここで上式右辺の第 2 項が折返し成分を表しており、フィルタ間に

$$\begin{bmatrix} H_0(z) & H_1(z) \\ H_0(-z) & H_1(-z) \end{bmatrix} \begin{bmatrix} G_0(z) \\ G_1(z) \end{bmatrix} = \begin{bmatrix} 2 \cdot z^{-N} \\ 0 \end{bmatrix} \quad (1.26)$$

の関係が成立するならば完全再構成が実現される。ただし N は任意の整数値であり、画像においては N 画素分の遅延に対応する。

一方、サブサンプリングパターンが上下チャンネルで反転する場合には、サブサンプリング関数がそれぞれ

$$f_0(n) = \frac{1 + (-1)^n}{2} \quad (1.27)$$

$$f_1(n) = \frac{1 - (-1)^n}{2} \quad (1.28)$$

として与えられる。故に、復元信号 $\hat{x}(n)$ は z 領域において

$$\begin{aligned} \hat{X}(z) &= \frac{1}{2}[H_0(z)G_0(z) + H_1(z)G_1(z)]X(z) \\ &+ \frac{1}{2}[H_0(-z)G_0(z) - H_1(-z)G_1(z)]X(-z) \end{aligned} \quad (1.29)$$

と表される。故に、完全再構成条件は、次式によって与えられる。

$$\begin{bmatrix} H_0(z) & H_1(z) \\ H_0(-z) & -H_1(-z) \end{bmatrix} \begin{bmatrix} G_0(z) \\ G_1(z) \end{bmatrix} = \begin{bmatrix} 2 \cdot z^{-N} \\ 0 \end{bmatrix} \quad (1.30)$$

表 1.1: 完全再構成フィルタの構成例

	$H_1(z)$	$G_0(z)$	$G_1(z)$	サンプリング	完全再構成条件
Esteban (QMF)	$H_0(-z)$	$2 \cdot G_0(z)$	$-2 \cdot G_0(-z)$	一致	$H_0(z)^2 - H_0(-z)^2 = z^{-k}$
Galand (odd-QMF)	$H_0(-z)$	$2 \cdot H_0(z)$	$2 \cdot H_0(-z)$	反転	$H_0(z)^2 + H_0(-z)^2 = z^{-k}$
Smith (CQF)	$-H_0(-z^{-1})z^{-N}$	$H_1(-z)$	$-H_0(-z)$	一致	$H_0(z)H_0(z^{-1}) + H_0(-z)H_0(-z^{-1}) = 2$
Le Gall (SSKF)	\dots	$H_1(-z)$	$-H_0(-z)$	一致	$H_0(z)H_1(-z) - H_1(z)H_0(-z) = 2 \cdot z^{-k}$

具体例

これまでに報告されているフィルタの構成例を表 1.1 に示す。これらの方式はすべて、

$$\begin{cases} G_0(z) = m \cdot H_1(-z) \\ G_1(z) = \pm m \cdot H_0(-z) \end{cases} \quad (1.31)$$

とすることによって折り返しひずみの完全な除去 (相殺) を実現している。ここで \pm は、サンプリングパターンが一致する場合にはマイナスになり、反転する場合にはプラスになる。これは、空間領域で見れば

$$\begin{cases} g_0(n) = m \cdot (-1)^n h_1(n) \\ g_1(n) = \pm m \cdot (-1)^n h_0(n) \end{cases} \quad (1.32)$$

となることを意味している。

Esteban らが 1977 年に示した QMF [27] では、まず $\{h_0(n)\}$ が係数対称の直線位相 FIR フィルタであることを前提としている。この場合、タップ数は必ず偶数でなければならない。そして、さらにフィルタ $\{h_1(n)\}$ を

$$h_1(n) = (-1)^n h_0(n) \quad (1.33)$$

すなわち

$$H_1(z) = H_0(-z) \quad (1.34)$$

とすることによって、完全再構成の実現を試みている。ただし、この QMF では折返しひずみと位相ひずみは完全に除去されるが、振幅ひずみはフィルタのタップ数が 2 と無限大の場合を除いて完全に除去することはできない (準完全再構成)。一方、Galand らは上下チャネルのサンプリングパターンを反転させることにより、奇数タップの QMF の構成方式を明らかにしたが、この方式も準完全再構成である [30]。

次に、Smith らが 1986 年に示した CQF [33] では、

$$h_1(n) = (-1)^n h_0(L-1-n) \quad (1.35)$$

すなわち

$$H_1(z) = -H_0(-z^{-1})z^{-L+1} \quad (1.36)$$

とした条件下において $\{h_0(n)\}$ の導出を図り、完全再構成を実現することに成功した。ここで L はフィルタ長を表すが、これは偶数でなければならない。この方式ではフィルタに直線位相性を持たせることはできないが、3 種類のひずみは完全に除去されている。

最後に、Le Gall らが 1988 年に示した SSKF [47] では、(1.30)(1.31) 式に基づいて完全再構成の問題を

$$F(z) - F(-z) = \frac{2}{m} \cdot z^{-k} \quad (1.37)$$

を満たす z の多項式 $F(z)$ を $H_0(z)$ と $H_1(-z)$ に因数分解する問題に帰着させ、この条件下でいくつかの完全再構成フィルタの提案を行っている。ただし、 $F(z)$ はその奇数次の項で係数が 0 でない項がただ 1

つしかなない多項式でなければならない。このフィルタバンクは完全再構成を実現するのみならず、すべてのフィルタが直線位相性を実現している。

- Daubechies が 1988 年に示した正規直交 wavelet [63] では、その理論的背景が階層的符号化に置かれていたにも関わらず、結果としてサブバンド符号化の完全再構成フィルタが与えられている。ただし、そのフィルタ間の関係は (1.35) 式によって与えられ、ここにサブバンド符号化の流れから生まれた CQF と、階層的符号化の流れから生まれた正規直交 wavelet とが完全に一致するという興味深いエピソードが報告されている。これらの相違点を強いて挙げれば、具体的なフィルタ係数の決定に当たり、前者が周波数特性を基準としたのに対して、後者が調和条件 (regularity condition) を基準とした点にある。

Chapter 2

画像符号化の統一表現

これまでの画像符号化の研究の流れを見ると、予測符号化は線形予測モデルとして、変換符号化は直交行列に基づく線形変換として、サブバンド符号化は z 変換を用いたフィルタ処理、あるいは MFB として、それぞれが最も取り扱い易い表現手法に基づいて理論展開が図られてきた。しかし、これらがすべて線形変換であることは明白であり、それらを統一的に表現することによって相互関係が明らかになると共に、それぞれに対して構築された理論体系の相互利用が可能となる。

そこで本章では、行列、および MFB に基づく各種画像符号化方式の統一表現を試みる。

2.1 行列の利用

2.1.1 サブバンド符号化の行列表現

サブバンド符号化は、これまでは z 変換や MFB を用いて表されるのが普通であった。これに対して、変換符号化と同様の行列表現を用いてサブバンド符号化を表現しようとした試みもいくつか報告されている [26,45]。本節では、これらの検討結果を基礎として、まず 2 バンド分割の場合のサブバンド符号化の行列表現に関する検討を行い、次にそれを多段接続した場合について検討を行う。あるいはまた、サブバンド符号化にブロック変換を組み合わせた場合についても同様の行列表現が可能であることを明らかにする。

2 バンド分割の場合

図 1.3 に示した 1 次元 2 バンドの帯域分割システムを考える。このとき、入力を $x(n)$ 、 $h_i(n)$ によるフィルタ処理の後にサブサンプリングされた出力を $y_i(n)$ とすると ($i=0,1$)、 $x(n)$ と $y_i(n)$ の関係は、サブサンプリングを考慮に入れた線形たたみこみ操作として

$$y_i(n) = \sum_{k=-\infty}^{\infty} h_i(k)x(2n-k) \quad (2.1)$$

と表される。一方、 $y_i(n)$ を 0 値補間した後フィルタ $g_i(n)$ を施し、さらにそれらを加算したものをシステム出力 $\hat{x}(n)$ とするとき、 $\hat{x}(n)$ もまたたたみこみ操作として

$$\hat{x}(n) = \sum_{k=-\infty}^{\infty} g_0(n-2k)y_0(k) + \sum_{k=-\infty}^{\infty} g_1(n-2k)y_1(k) \quad (2.2)$$

と表される。

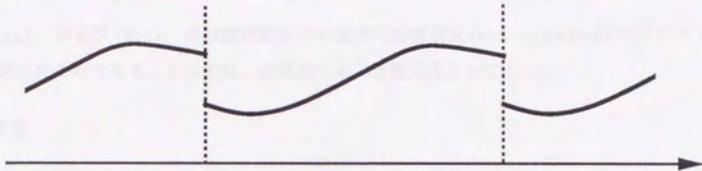
ここで上記の操作の時間領域 (画像の場合には空間領域) における行列表現を試みる。まず、以下に示す 3 つのベクトルを規定する (t は転置を表す)

$$\begin{aligned} \mathbf{x} &= (\cdots, x(0), x(1), x(2), \cdots)^t \\ \mathbf{y} &= (\cdots, y_0(0), y_0(1), \cdots, y_1(0), y_1(1), \cdots)^t \\ \hat{\mathbf{x}} &= (\cdots, \hat{x}(0), \hat{x}(1), \hat{x}(2), \cdots)^t \end{aligned} \quad (2.3)$$

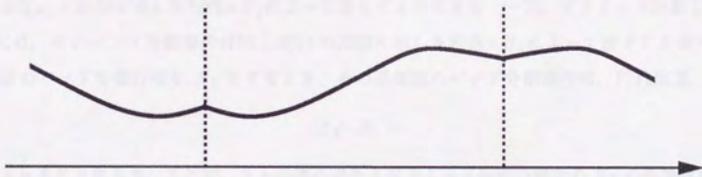
すると (2.1) 式、および (2.2) 式の操作は、それぞれフィルタ係数から構成される行列 H 、 G を用いて

$$\mathbf{y} = H\mathbf{x} \quad (2.4)$$

$$\hat{\mathbf{x}} = G\mathbf{y} \quad (2.5)$$



(a) Circular convolution



(b) Symmetric expansion

図 2.1: 有限入力数に対する境界部の処理

多段接続の場合

サブバンド符号化が図 1.3 のような 2 バンド分割だけで用いられることは希であり、一般的には多段(縦列)接続に基づく $M (> 2)$ 個以上のバンド分割, あるいは DCT 等の直交変換と組み合わせて利用されることが多い。これらの“階層処理”は、以下のような行列の積として表現することができる。

図 2.2 の (a), および (b) は、多段接続数が 2 の場合の木構造型 (tree-structured) サブバンド符号化の行列表現を試みたものである。ここでは、木構造バンド分割方式として

- フル分割
- ピラミッド分割

の 2 種類の分割方法を考えている。前者は分割可能なすべての信号系列のバンド分割を行う方式であるが、後者は低周波成分のみの再分割を行う方式であり、これは階層的符号化方式と同様の周波数分割にはかならない。

図 2.2 に示した右側の行列は、(2.6) 式のフィルタ係数行列、すなわち第 1 段目のバンド分割行列を概念的に表したものである。このとき、それぞれの短冊がフィルタ係数群に対応しているが、行列のサイズが $N \times N$ のとき (これは入力数が N であることを意味する)、上下それぞれの短冊の数は $N/2$ に等しくなる。以下、この行列を H_1 と呼ぶことにする。次に、2 段のフル分割 (4 バンド分割) の場合には、そのバンド分割操作は図 2.2 (a) の左側に示した行列 $\times H_1$ によって表すことができる。一方、ピラミッド分割 (3 バンド分割) の場合には、そのバンド分割操作は図 2.2 (b) の左側に示した行列 $\times H_1$ によって表すことができる。結局、第 2 段目のバンド分割行列を H_2 とするとき、その送信側のバンド分割操作は、行列演算

$$H_2 \cdot H_1 \quad (2.8)$$

によって表されることになる。ここで、フル分割の場合とピラミッド分割の場合の H_2 の相違点は、フル分割の場合の第 1 段目の高周波出力の再分割を表す行列の部分が、ピラミッド分割では単位行列によって置き換えられていることに表れている。

結局、 K 段縦列のサブバンドシステム全体の送信側の処理を (2.4) 式に対応づけるとき、その全体の処理行列 H は各段のフィルタ処理行列 H_k ($k = 1, 2, \dots, K$) の積

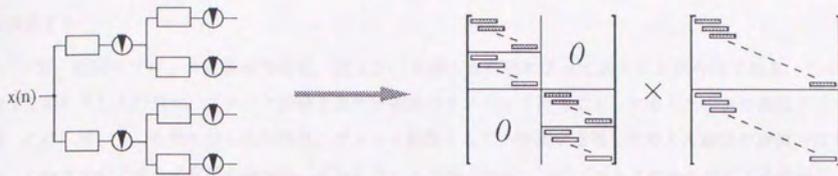
$$H = H_K \cdots H_2 H_1 \quad (2.9)$$

として表されることになる。一方、受信側の処理行列 G は、送信側と同様にして

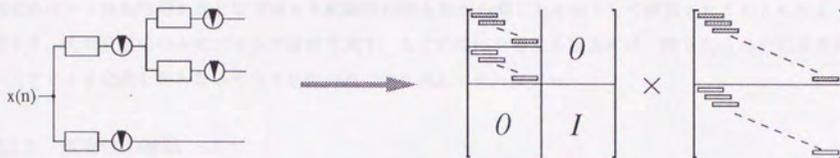
$$G = G_1 G_2 \cdots G_K \quad (2.10)$$

と表される。ここで完全再構成が実現される場合には、

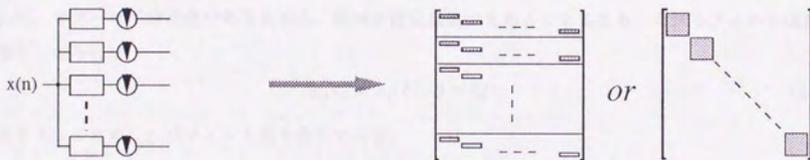
$$G_k \cdot H_k = I \quad (2.11)$$



(a) Full decomposition (regular)



(b) Pyramid decomposition (irregular)



(c) Block transform

図 2.2: サブバンド符号化の多段接続

が成立する。

一方、従来のブロック変換符号化は、図 2.2(c) に示した行列によって表すことが可能である。ここで、まず左側に示した行列は、ブロック変換を並列分割型のサブバンド符号化として考えた場合の表現方法である。これに対して、右側に示した行列は、ブロック変換としての特徴を重視して考えた場合の表現方法である。これまでのブロック変換の表現は、右側に示した表現方法の一つのブロックのみに着目する傾向があったが、このように入力全体を意識した表現を行うことにより、ブロック変換がサブバンド符号化の中に包含される概念であることが理解される。すなわち、ブロックサイズを M とするとき、これは並列 $M(> 2)$ バンド分割のサブバンドシステムの一実例として考えることができるわけである。これは、LOT、MLT 等のオーバーラップ変換についても同様のことが言える。

サブバンド符号化にブロック変換を組合せることはしばしば行われるが、これらの処理はサブバンド符号化のバンド分割行列に対してブロック変換の行列を左から乗じたものとして表現されることになる。このとき、低周波成分のみにブロック変換を施す、などの操作を考える場合には、図 2.2(c) の行列表現においてブロック変換を行わない部分を単位行列で置き換えてやればよい。

2.1.2 直交性の検証

行列 H が“直交条件”

$$H \cdot H^t = H^t \cdot H = I \quad (2.12)$$

を満足する場合、 H は“直交行列”と呼ばれ、また H に基づく線形変換は“直交変換”と呼ばれる(ユニタリ行列、ユニタリ変換もまた複素領域において同様に定義される)。一方、サブバンド符号化の行列表現においてバンド分割行列 H が直交条件を満足する場合、受信側の合成行列 G は H^t に等しくなる。言い替えれば、サブバンド符号化でありながら、同時に直交変換でもあることになる。これをフィルタ係数間の関係としてみれば、

$$g_i(k) = h_i(L-1-k) \quad (2.13)$$

が成立する。ここに、 L はフィルタ長を表している。

これまでに報告されている著名なブロック変換(KLT, DCT, DWHT等)やオーバーラップ変換(LOT, MLT等)はすべて直交変換であり、前節のプロセスに従ってバンド分割行列 H を構成すると、それらが上記の直交条件を満足するは明らかである。それでは、サブバンド符号化における各種のフィルタ構成方式は、やはりこの直交条件を満足するのであろうか？ここでは、その典型例として、QMF, CQF, SSKF に対する直交性の検証を行う。

まず図 2.3は、16 タップの QMF から構成されるバンド分割行列 H に対する $H \cdot H^t$ の様子を表したものである。この図から明らかのように H は直交条件を満足しており、QMF を用いたサブバンド符号化は直交変換にほかならない。従来の KLT や DCT などの直交変換とは異なり、ブロックの概念がないために

その直交性が直感的には分かりにくい、要するに QMF を用いたサブバンド符号化は“ブロックサイズが入力数に一致する直交変換”にはかならない。前述のように、その受信側の合成行列 G は H^t に等しい。

- 厳密には QMF は準直交行列であり、QMF の準完全再構成の問題は、タップ数の増加に伴う直交行列の近似の度合の向上、として考えることができる。

CQF から得られるバンド分割行列も直交条件を満足する。この CQF は、Daubechies らによって指摘されているように、正規直交 wavelet 基底の一実現例として考えることができる。言い替えれば、正規直交 wavelet 基底を用いたサブバンド符号化はすべて直交変換である。

一方、図 2.4 は、 5×3 の SSKF に対する $H \cdot H^t$ の様子を表したものである。ここで、5 はローパスフィルタのタップ数、3 はハイパスフィルタのタップ数をそれぞれ表している。この図より明らかなように、SSKF ではもはや直交性は成立していない。

これらの結果を踏まえて、サブバンド符号化における直交性の問題を、その完全再構成の問題と絡めてまとめてみると

- CQF : 直交行列, 完全再構成
- QMF : 準直交行列, 準完全再構成
- SSKF : 非直交行列, 完全再構成

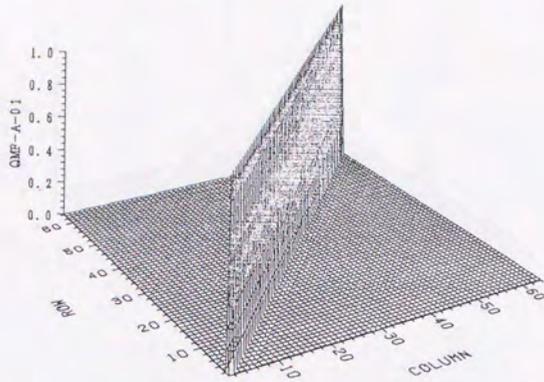
となる。また、図 2.5 には、既存の画像符号化方式の直交性の成立をまとめてみた。この図からわかるように、ブロック変換やオーバーラップ変換、あるいは正規直交 wavelet 基底等を含めて、画像符号化方式の多くは直交変換である。

- 以下の表記において、タップ数が n の QMF や CQF を、 $QMF(n)$ 、 $CQF(n)$ のように表す。この場合、ローパスフィルタとハイパスフィルタのタップ数は等しい。一方、ローパスフィルタとハイパスフィルタのタップ数が異なる SSKF の場合には、 m をローパスフィルタのタップ数、 n をハイパスフィルタのタップ数として、 $SSKF(m \times n)$ のように表記する。
- サブバンド符号化における完全再構成の問題は、(2.4)(2.5) 式から明らかなように

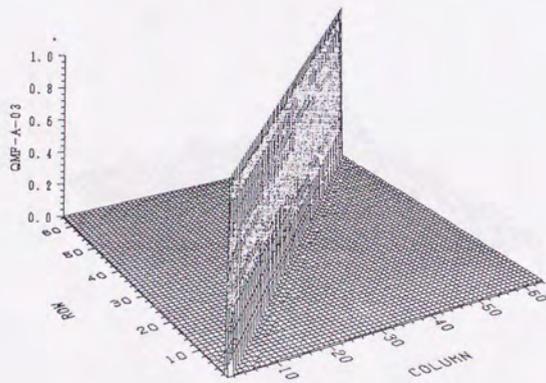
$$G \cdot H = I \quad (2.14)$$

の成立を意味しているに過ぎない。言い替えれば、この条件は必ずしも直交条件の成立を意味するものではない。

- 予測符号化に導入される理論として、入力信号と予測誤差の直交性の実現を目的とする“直交化原理”(orthogonality principle) が知られている [1]。ただし、この直交化原理は、予測符号化自体の直交性を意味するものではない。

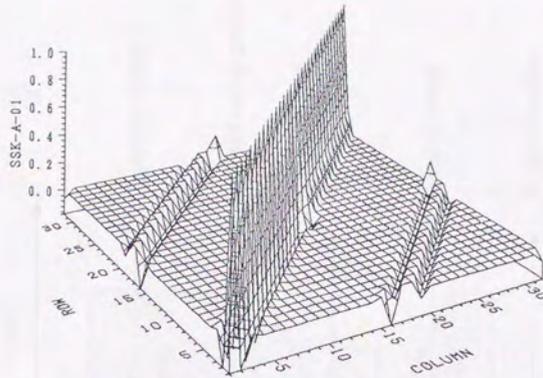


(a) 階層処理 1 回の場合

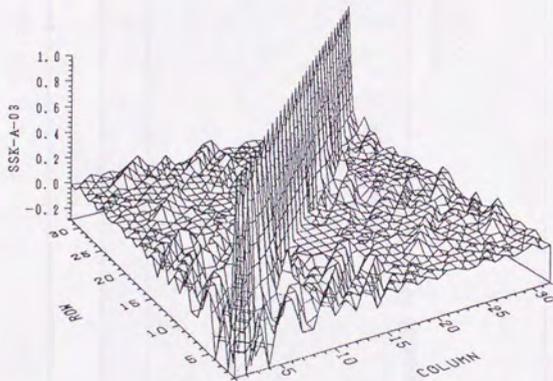


(b) 階層処理 3 回の場合

図 2.3: QMF の直交性



(a) 階層処理 1 回の場合



(b) 階層処理 3 回の場合

図 2.4: SSKF の非直交性

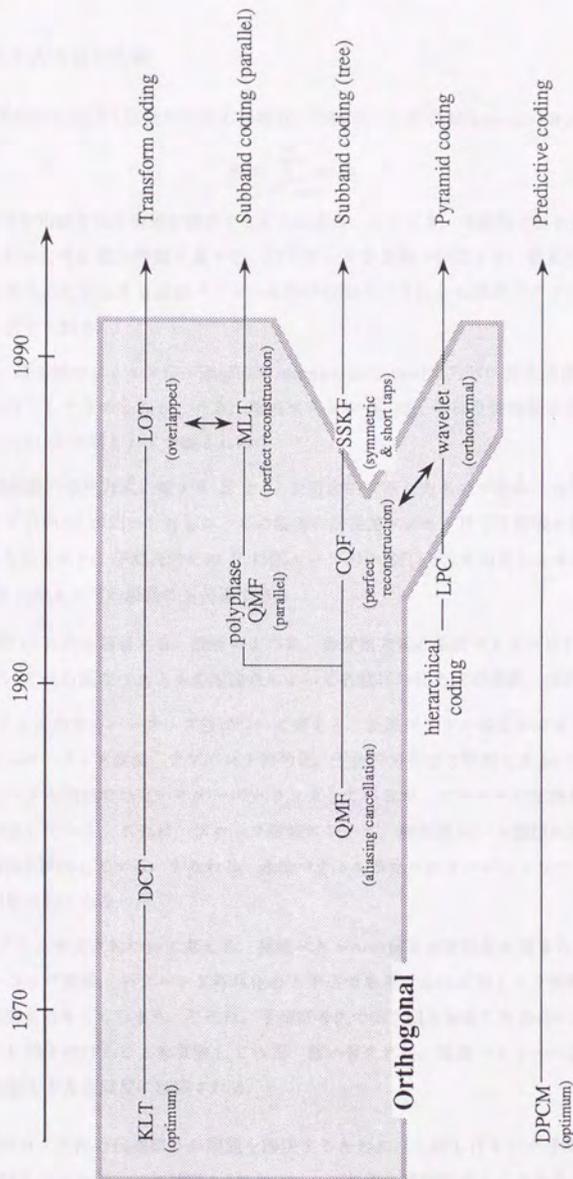


図2.5: 直交変換に基づく画像符号化

2.1.3 基底ベクトルの相互比較

一般的に、線形変換に基づく画像符号化の目的は、基底ベクトル (basis vector) b_j の重み付け加算和

$$\hat{x} = \sum_{j=-\infty}^{\infty} w_j b_j \quad (2.15)$$

によって、入力信号を可能な限り忠実に復元することにある。このとき、送信側ではある規則に従って重み係数 w_j の算出を行い、その重み係数を量子化、符号化して受信側へ伝送する。次に受信側では、伝送されてきた重み係数をそれに対応する基底ベクトルに掛け合わせ、それらの基底ベクトルの和を取ることで入力信号の復元を図る。

これに対して、送信側のフィルタが FIR (finite impulse response) ならば、重み係数 w_j は、行列 H による線形変換の出力として求められる。一方、基底ベクトル b_j は、やはり受信側の合成フィルタが FIR である場合には、行列 G の列として定義される。

図 2.6 は、各種画像符号化方式に対する H と G を概念的に示したものである。ただし、一般的に予測符号化では閉ループ DPCM が用いられるが、この場合の送信側フィルタは IIR 構成を採るために行列表現を適用できない。このために、予測符号化の H は開ループ DPCM (1 次) に対応したものである。なお、受信側の合成行列 G は閉ループの場合にも共通である。

以後、合成行列 G へのみ着目する。前述のように、各変換方式の基底ベクトルは行列 G の列として与えられる。そこで、これら基底ベクトルの相違点によって各種符号化方式の長所、短所を明らかにする。

まず、基底ベクトルのオーバーラップ性について考える。基底ベクトルが互いにオーバーラップしている変換方式は、オーバーラップ変換、サブバンド符号化、予測符号化の 3 手法である。これに対してブロック変換では、ブロックの内部では互いにオーバーラップしているが、ブロックの境界部において基底ベクトルの不連続性が生じている。これは、ブロック変換において、特に低レート環境において顕著となるブロックひずみの問題を反映している。すなわち、基底ベクトルが互いにオーバーラップしているならば、ブロックひずみの問題は起こらない。

次に、基底ベクトルの長さについて考える。基底ベクトルの長さが有限長に収まる変換方式は、ブロック変換、オーバーラップ変換、サブバンド符号化の 3 手法である。これに対して予測符号化では、基底ベクトルの長さが無限長になってしまう。これは、予測符号化で伝送誤りが生じた場合に、その影響が後続する信号系列に永久に残り続けることを反映している。言い替えれば、基底ベクトルの長さが有限長であれば、伝送誤りの影響もまた有限長に制限される。

- 予測符号化では、上記の伝送誤りの問題を解決するために、しばしばリーク予測 (leaky prediction) の導入が検討されている。これは図 2.6 において、 ρ の値の絶対値が 1 より小さければ予測符号化の基底ベクトルは徐々に 0 に漸近する形状となり、伝送誤りの影響もまた次第に減衰して行くことに対応している。

広義のサブバンド符号化

サブバンド符号化の利点としては

1. 優れた圧縮効率を実現できること
2. ブロックひずみがないこと
3. 伝送誤りの影響が有限長に制限されること

等が挙げられる。このうち、2番目と3番目の利点は、上述のように基底ベクトルの相互比較によって説明することができる。一方、オーバーラップ変換は並列型サブバンド符号化の一形態にはかならず、結果としてサブバンド符号化は、ブロック変換と予測符号化の双方の利点を兼ね備えた符号化方式であると考えることができる。そこで本論文では

- 基底ベクトルが互いにオーバーラップしており、かつその長さが有限であるものを、“広義の”サブバンド符号化と呼ぶ

ことにする(図2.7)。従来のサブバンド符号化では、概してそのフィルタの周波数特性が重要視されてきたが、本論文ではあえてその問題を意識しない。

- サブバンド符号化の利点として、視覚特性のビット割当てへの反映、が指摘されることがあるが、これは必ずしもサブバンド符号化特有の利点とはならない。予測符号化ではノイズフィードバック符号化(NFC: noise feedback coding)が知られており、ブロック変換では変換係数ごとに重み係数を考慮するなどして、視覚特性の反映が図られている。

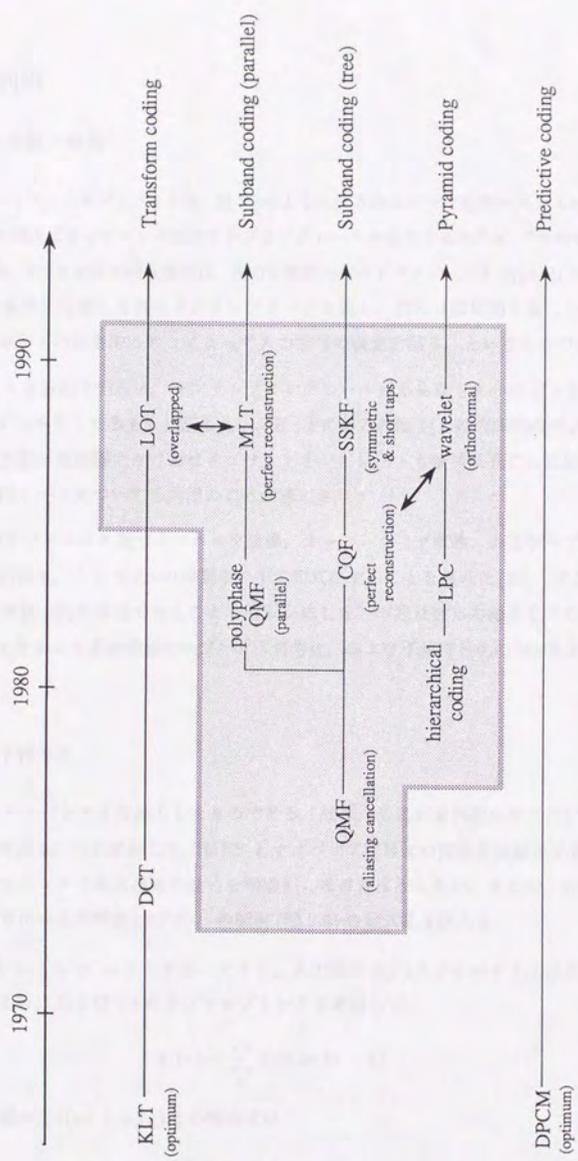


図2.7: 広義のサブバンド符号化

2.2 MFBの利用

2.2.1 MFBによる統一表現

MFB (マルチレートフィルタバンク) とは、図 1.8 のように表されるバンド分割システムの総称であり、サブサンプリング操作の導入によってバンド毎にサンプリングレートが変化するために、“マルチレート” という言葉が用いられている。MFBの基本的な機能は、入力を複数のバンドパスフィルタ $h_i(n)$ ($i = 0, 1, \dots, K-1$) によって K 個の信号系列に分割した後サブサンプリングを施し、次に 0 値補間を施した後バンドパスフィルタ $g_i(n)$ を掛け、その和を取ることにによって入力信号の復元を図る、というものである。

ここで図 1.8 のような表記を行うと、サブサンプリングレートがあたかもすべて $K:1$ であるかのような錯覚に陥るが、必ずしもそうであるとは限らない。 $K:1$ になるのは $1/K$ の等間隔のバンド分割を行った場合であり、バンド分割が等間隔でなければサブサンプリングレートも信号系列ごとに異なるのは明らかである。これは、補間レートについても同様のことが言える。

Vetterli らは、FIR フィルタに基づくブロック変換、オーバーラップ変換、およびサブバンド符号化における完全再構成の問題を、この MFB の枠組みの中で定式化することを試みた [26]。すなわち、MFB によってこれら 3 手法が統一的に表現できることを明らかにした。本節は彼らの結果をそのまま踏襲するものとし、彼らが言及しなかった多段接続のサブバンド符号化、および予測符号化の MFB としての位置づけを明らかにする。

多段接続のサブバンド符号化

MFB は並列フィルタバンクを前提としたものである。故に、これが並列型のサブバンド符号化を包含していることは自明である。これに対して、MFB とサブバンド符号化の関係を完結させるためには、2 バンド分割等のフィルタバンクの多段接続の扱いを明確にしなければならない。そこで、ここでは多段接続構成のサブバンド符号化の並列構成 (MFB) への転換 (図 2.8) の定式化を試みる。

図 2.8 において $K = L = M = 2$ とする。すると、入力信号 $x(n)$ とフィルタ $h_1(n)$ の出力 $y_1(n)$ との関係は、線形たたみこみ、および 2:1 のサブサンプリングを考慮して、

$$y_1(n) = \sum_k h_1(k)x(2n-k) \quad (2.16)$$

として表される。同様に、 $x(n)$ と $y_2(n)$ との関係式は

$$\begin{aligned} y_2(n) &= \sum_{k'} h_2(k')y_1(2n-k') \\ &= \sum_{k'} \sum_l h_2(k')h_1(l)x(4n-2k'-l) \end{aligned}$$

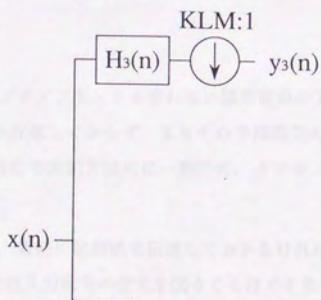
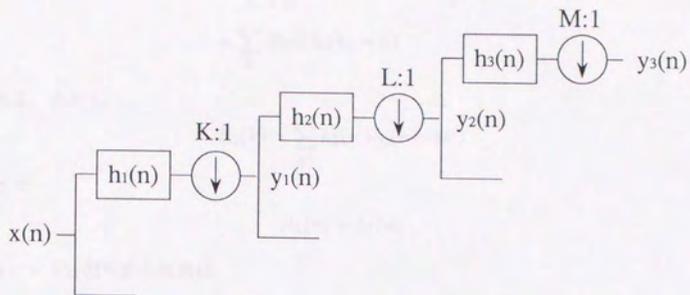


図 2.8: 多段接続構成の並列構成への転換

$$\begin{aligned}
&= \sum_k \left[\sum_{k'} h_2(k') h_1(k - 2k') \right] x(4n - k) \\
&= \sum_k H_2(k) x(4n - k) \tag{2.17}
\end{aligned}$$

として表される。ただし、

$$H_2(k) = \sum_{k'} h_2(k') h_1(k - 2k') \tag{2.18}$$

である。そこで

$$H_1(n) = h_1(n) \tag{2.19}$$

とすると、2バンド分割の多段接続は

$$\begin{cases} y_i(n) = \sum_k H_i(k) x(2^i \cdot n - k) & (i \geq 1) \\ H_i(n) = \sum_k h_i(k) H_{i-1}(n - 2^{i-1} \cdot k) & (i \geq 2) \end{cases} \tag{2.20}$$

として並列化することが可能である。このとき、バンドごとのサブサンプリングレートは(2.20)式の 2^i によって規定されることになり、これらがバンドごとに異なることは明らかであろう。

受信側、および2バンド分割以外の場合については省略するが、以上の操作がこれらに容易に適用できることは明らかであろう。これによって、多段接続のサブバンド符号化のMFBへの転換が完了する。

予測符号化

予測符号化は、直感的にはサブサンプリングを伴わない線形変換のように見える。これは、入力信号に施される予測フィルタが一つしか存在しておらず、またその予測誤差のみに関心が向けられてきたことに起因する。このために、予測符号化の表記方法には一般的に、サブサンプリング操作の存在しない構成が用いられてきた。

しかし、より厳密に考えると、最初に初期値を伝送しておかなければ予測符号化はまったく意味をなさない。すなわち、予測誤差だけでは入力信号の復元を図ることはできない。一方、MFBにおいてフィルタ $h_1(n)$ に着目した場合、それらがすべて異なるフィルタでなければならないという必然性はない。

そこで、この初期値の扱いを分離して、予測符号化に対してMFBの考え方を導入すると、開ループDPCMは図2.9のように表すことができる。このとき、 $h_0(n)$ が初期値通過フィルタであり、そのほかの $\{h_1(n), h_2(n), \dots\}$ は予測フィルタとして、遅延分を除いてすべて同じ形状となる。一方、受信側の合成フィルタ $g_i(n)$ はすべて同じ形状となり、1次予測の場合には、前節の行列表現から $\{\dots, \rho^2, \rho^1, 1\}$ によって与えられることになる。さらにはまた、入力信号の総数が N であれば、そのサンプリング間隔は $N:1$ によって規定されることになる。

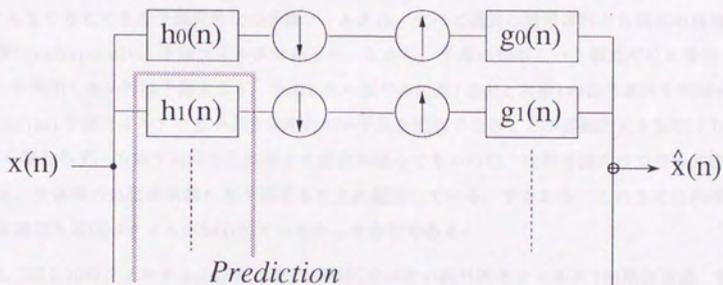


図 2.9: 予測符号化の MFB としての位置づけ

2.2.2 予測符号化の変形 (内挿予測の概念の導入)

前節の検討により、各種画像符号化方式の MFB としての解釈が完了する。MFB はサブサンプリング操作を伴っているところから、標本化定理を背景としたバンドパスフィルタ群として見られがちであるが、予測符号化の例を見るように、必ずしもそれらに拘束される概念ではない。さらにはまた、以下のように、フィルタバンクの新しい構成を考えることが可能である。

まず、図 2.9 において、予測フィルタを有限個で打ち切ることを考える (図 2.10)。この構成は、言うまでもなく、周期的リフレッシュを伴う予測符号化にはかならない。この場合、サブサンプリング操作によるレート変換はすべて $K:1$ に等しく、見方を変えれば並列型 K バンド分割のサブバンド符号化の一実現例として考えることもできる。

一方、これまで考えてきた予測符号化の予測フィルタは、すべて過去の信号系列から現在の信号の予測を行う“外挿”(extrapolative)予測フィルタであった。しかし、予測の効率という観点だけに着目すれば、過去の情報しか利用しない外挿予測よりも、予測したい信号の両側(過去と未来)の信号系列を利用する“内挿”(interpolative)予測フィルタの方がより効率の高い予測を実現できることが理論的にも証明されている [1]。それに関わらず、内挿予測符号化に関する研究が盛んでないのは、内挿予測だけで符号器構成を行おうとすると、受信側の処理が複雑になり過ぎることに起因している。すなわち、これまでは内挿予測符号化に対する適切な基底ベクトルが知られていなかったわけである。

それでは、図 2.10 のフィルタ $h_0(n)$ だけを内挿補間ではない例外的なフィルタ (初期値通過、等)として制約条件を緩和すれば、そのほかのフィルタ $h_i(n)$ ($i \geq 1$) はすべて内挿補間フィルタとして機能するような完全再構成のフィルタバンクを構成することができるのではないだろうか?

- 予測の効果のないフィルタの挿入による圧縮効率の低減が危惧されるが、これは木構造バンド分割としてのフィルタバンクの多段接続に基づく入力信号の相関の段階的な除去によって、ある程度解決できる。この点から、本節で検討を行うフィルタバンクは、予測符号化とサブバンド符号化の中間的な存在として位置づけることが可能である。

Le Gall らによって提案された SSKF を考える。SSKF の中で、SSKF(1×3) のハイパスフィルタは $\{-1/2, 1, -1/2\}$ として与えられるが、これはまさに内挿補間フィルタにかならない。この場合ローパスフィルタは $\{1\}$ で与えられるが、これは前述の初期値通過フィルタと考えることができる。すなわち、SSKF(1×3)こそ、上の問いに対する解となるフィルタバンクであることになる。

- SSKF(1×3) は図 2.10 の $K=2$ の場合に対応している。さらに $K \geq 3$ の場合についても、同様の完全再構成フィルタを構成することは可能である [42]。
- 上のようなフィルタバンクは、動画画像符号化において時間軸方向に適用すれば、フレーム内挿的な効果を実現する。このために、MPEG のような蓄積系動画画像符号化に対する応用も考えられる。

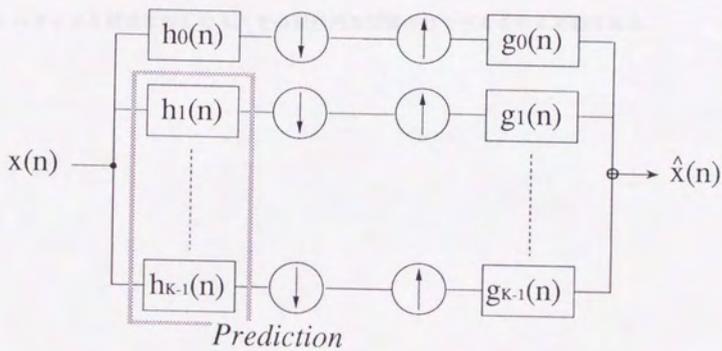


図 2.10: 予測符号化の変形

次に SSKF(5×3) を考える。この場合のハイパスフィルタは SSKF(1×3) のそれに一致するが、ローパスフィルタは $\{-1/8, 1/4, 3/4, 1/4, -1/8\}$ として与えられる。これはすなわち、 $h_i(n)$ ($n \geq 1$) が内挿補間フィルタである特性は維持しつつ、さらに $h_0(n)$ が単純な初期値通過フィルタではないフィルタ構成が可能であることを意味している。ここでフィルタバンクの多段接続を前提として、フィルタ $h_0(n)$ の周波数特性を考えた場合、全域通過特性の初期値通過フィルタよりも、よりローパス特性を呈する SSKF(5×3) の $h_0(n)$ の方が圧縮効率も優れていることは容易に想像が行く。小松らは、以上のような観点から新しい並列型完全再構成フィルタの構成方法に関する検討を行い、具体的なフィルタ構成を提案している。そして、1 次元の場合に、DCT(8) 並の圧縮特性を実現できる $K=3$ のフィルタバンクを明らかにした [42]。

本論文では、以降これらの並列フィルタバンクについては言及しないが、本論文で述べる理論展開は必然的にこれらの並列構成に対しても適用可能である。いずれにせよ、以上の検討によって画像符号化の統一表現が完了する。この画像符号化の統一表現は次章以降の理論展開の基礎となるだけではなく、本節のような新しいフィルタ構成に対しても、その定性的な位置づけを与えるのに有効である。

2.3 考察・検討

本節では、線形変換に基づく画像符号化方式の行列と MFB を用いた統一表現を試みた。線形変換が最大の枠組みであることは自明であるが、この中で各種画像符号化方式は直交系と非直交系、基底ベクトルが互いにオーバーラップしているものとしていないもの、その長さが有限であるものとそうでないもの、などに分類されることになる。いずれにせよ、このような統一表現により、それぞれの符号化方式の相互関係・相違点を明らかにすると共に、サブバンド符号化の有効性を定性的な観点から示すことができた。

一方、本節の統一表現では、各種符号化方式において同じ基底ベクトルが周期的に現れることを仮定している。これは MFB では、同じ合成フィルタが周期的に使用されることに対応している。入力情報源が均一であればこれで何等問題はないが、時間的にその特性が変化するような入力であれば、基底ベクトルを時間的に変化させることによって特性の改善を図るのが普通である。これは予測符号化では適応的に予測係数を修正する操作、変換符号化では適応的に変換基底を変化させる操作に対応する。

サブバンド符号化でこのような適応操作を導入する場合には、まず多段接続においてバンドの分割方法を適応的に切り替える、などの方式が考えられる。あるいはまた、KLT のようにフィルタバンクを適応的に変化させる方法も考えられるが、変換がブロック単位でないために、問題が複雑になることが予想される。ただし、適応処理として自然な検討課題ではある。

Chapter 3

サブバンド符号化の特性評価

予測符号化や変換符号化では、任意のフィルタ構成に対してその圧縮効率を評価する理論解析手法が確立されている。これに対して、サブバンド符号化では、理想フィルタを用いた場合の理論解析手法が存在するだけで、実際に用いられるフィルタ構成の圧縮効率の評価はシミュレーションに頼らざるを得ない。実際、サブバンド符号化の研究動向を見ると

1. 使用するフィルタ (QMF, CQF, wavelet, SSKF etc.)
2. 分割バンド数
3. バンド分割方法 (フル分割, ピラミッド分割, 可分型, 非可分型 etc.)
4. 直交変換や予測符号化との組合せ方法

などについての共通した認識の得られないままに、多種多様な符号化方法の提案が行われているのが現状である。このような混沌を招いた最大の要因は、上記のように、その圧縮効率の理論解析手法が未だに確立されていないことにあるものと考えられる。

そこで本章では、サブバンド符号化に関する多くのシミュレーション結果に理論的な裏付けを与え、さらには効率的な符号器設計への指針を与えることを目的として、サブバンド符号化の圧縮効率の理論解析手法の構築を試みる。

3.1 符号化ゲインに基づく特性評価

QMF や CQF を用いたサブバンド符号化は、入力信号の数がブロックサイズに等しい直交変換として考えることができる。そこで、これらのサブバンド符号化の圧縮効率、直交変換と同様の理論解析を図ることが可能である。

3.1.1 Coding gain の利用

Jayant と Noll によって与えられた直交変換の圧縮効率の評価尺度、符号化ゲイン (coding gain) G_{TC} は、以下のように定義される (1.3.1 節参照)。

H を、所望の直交変換に対応するサイズ $N \times N$ の直交 (or ユニタリ) 行列とする。ここで入力信号の自己相関行列を R_{xx} とすると、行列 H を用いた線形変換の変換係数の自己相関行列 $R_{\theta\theta}$ は

$$R_{\theta\theta} = H R_{xx} H^t \quad (3.1)$$

によって与えられる。そして、この $R_{\theta\theta}$ の対角要素 σ_k^2 ($k = 0, 1, \dots, N-1$) を用いて、符号化ゲイン G_{TC} が次式によって定義される。

$$G_{TC} = \frac{\frac{1}{N} \sum_{k=0}^{N-1} \sigma_k^2}{\left(\prod_{k=0}^{N-1} \sigma_k^2 \right)^{\frac{1}{N}}} \quad (3.2)$$

上式はスカラー量子化として変換係数ごとのビット割当ての最適化を行った場合の PCM との再生誤差の分散の比を表しており、この G_{TC} の値によって PCM に対する圧縮効率の向上効果を定量的に評価できる。

これに対して、2.1 節のサブバンド符号化の行列表現では、QMF、CQF に対応する行列 H が直交行列であることを実証した。そこでこの行列 H を上記のプロセスにそのまま当てはめてやれば、QMF と CQF に対応する符号化ゲインが求められることになる。すなわち、直交性を持つフィルタバンクに限定されているが、ここに“理想的な周波数特性を持たないフィルタバンクを用いたサブバンド符号化の圧縮効率の理論解析”が初めて可能となる。

3.1.2 AR(1) モデルに対する特性評価

入力信号を AR(1) プロセスによってモデル化する。この AR(1) プロセスは画像信号のモデルとして古くから用いられており、

$$x(n) = \rho \cdot x(n-1) + z(n) \quad (3.3)$$

によって規定されるものである。ここで $x(n)$ が画像信号に対応しており、 $z(n)$ はホワイトノイズ (ガウス雑音) を表している。モデルパラメータ ρ は結果として $x(n)$ と $x(n-1)$ との相関係数を規定する役割を果

たすことになり、画像のような相関の高い入力に対しては 0.95 や 0.90 程度の値が適用されることが多い。このとき、(3.1) 式の自己相関行列 R_{xx} は

$$R_{xx} = \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{N-1} \\ \rho & 1 & \rho & & \\ \rho^2 & \rho & 1 & & \\ \vdots & & & \ddots & \\ \rho^{N-1} & & & & 1 \end{bmatrix} \quad (3.4)$$

によって表されることになる。そしてこの R_{xx} とサブバンド符号化の行列表現 H を用いて、(3.1)(3.2) 式から直交系フィルタバンクの符号化ゲインが求められる。

- DCT の有効性を論じる場合、 $\rho \rightarrow 1$ としたときの KLT の DCT への漸近性がしばしばその理論的な根拠として引用される。この場合の入力モデルは、言うまでもなく AR(1) プロセスである。
- 画像信号のモデルとして AR(1) では不十分ではないか、という疑問が生じるのは当然である。しかし、経験的には AR(1) によってほぼ十分な結果が得られており、より高次の AR プロセスやこれらとは異なるモデルを用いてもあまり大きなブレイクスルーは得られていないのが現状である。そこで本論文でも AR(1) のみを検討対象とする。

まず図 3.1 は、QMF(16) に対するフル分割とピラミッド分割の符号化ゲインの比較を表したものである。ここで横軸は多段接続(木構造)に基づく周波数分割を前提とした場合のフィルタバンクの多段接続数を表しており、結果としてバンドの分割数に対応することになる。まずフル分割の場合には、ステージ数 1 は 2 バンド、ステージ数 2 は 4 バンド、ステージ数 3 は 8 バンドをそれぞれ表している。一方、ピラミッド分割の場合には、ステージ数 1 はフル分割と同じになるが、ステージ数 2 は 3 バンド、ステージ数 3 は 4 バンドを表すことになる。縦軸は符号化ゲインを表しており、図中の点線は理論上の最適値 $(1-\rho^2)^{-1}$ を表している。また、行列のサイズは $N = 256$ としている。

ここではモデルパラメータ ρ の値として 0.95 と 0.90 を考えたが、この図からは

- ピラミッド分割によって、フル分割とほとんど変わらない十分なデータ圧縮が実現できる

ことが示唆される。この結果は階層的符号化や wavelet 変換のアイデアに理論的な妥当性を与えることになるが、裏返せば高相関入力はいかに低周波成分にエネルギーが集中しているかということを示すものである。図 3.2 には $\rho = 0.95$ の場合の AR(1) モデルのパワースペクトル密度を示すが、この図からも低周波成分へのエネルギーの集中が理解できる。なお、上記の結果には QMF の準完全再構成の影響が反映されていないが、16 タップもの長タップになると、準完全再構成の影響よりも量子化誤差の影響の方が大きくなり、これは本質的な問題とはならないものと考えられる。

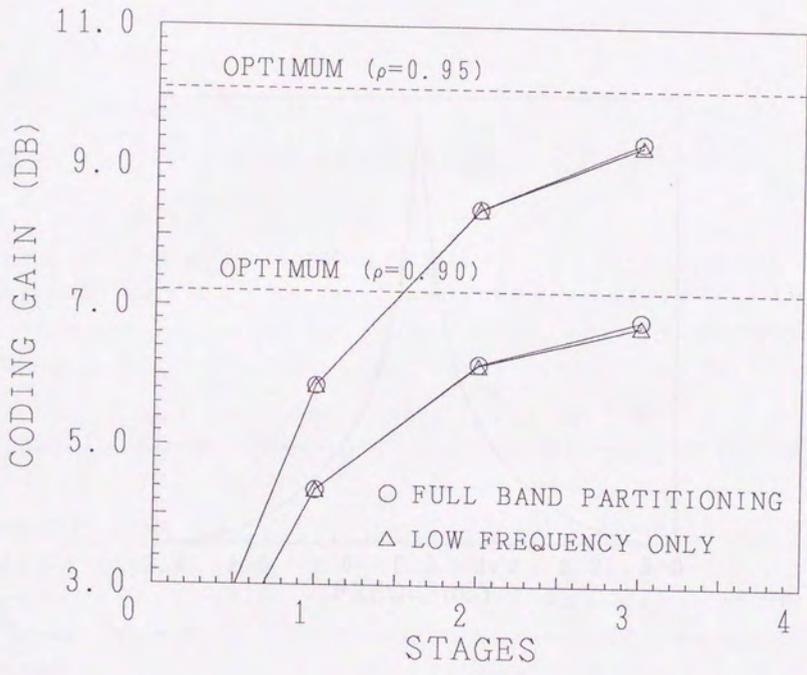


図 3.1: サブバンド符号化の圧縮効率(1): フル分割 vs. ピラミッド分割

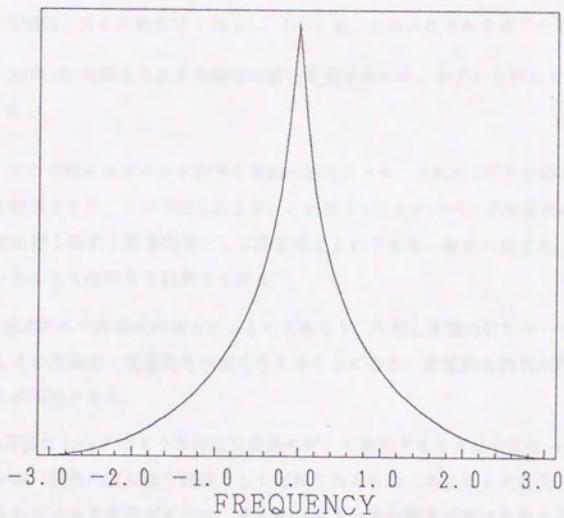


図 3.2: AR(1) モデルのパワースペクトル密度 ($\rho = 0.95$)

次に図 3.3は、QMF(16)とCQF(16)をそれぞれ単独にピラミッド分割として用いた場合の符号化ゲインと、そのピラミッド分割によって得られた低周波成分に対してさらにDCT(8)を適用した場合の符号化ゲインを並記したものである($\rho = 0.95$)。ここで後者はサブバンド符号化に直交変換を組み合わせた符号化方式の圧縮効率の一例であり、これらがサブバンド符号化単独の場合と同様の行列表現が可能なのは既に2.1節で明らかにしている。ここではまた比較として、DCT(2^n)単独の場合の符号化ゲインも付記している。

この図からは

- サブバンド数の増加と共に圧縮効率も向上し、DCT並、あるいはそれを凌ぐ圧縮特性を実現できる
- QMF(16)とCQF(16)はあまり大きな特性の違いを見せないが、わずかながらCQF(16)の方が良好な特性を呈する

ことなどがわかる。ここで特にサブバンド符号化単独の場合よりも、それにDCTを組み合わせた場合の方がより優れた圧縮を実現できることが示唆されるが、これはDCTをサブバンド符号化の一種と考えてやれば、バンド数の増加に伴う特性の改善効果として考えることができる。あるいはまた、低周波成分には相関がかなり残っていることを証明する結果でもある。

以上の結果は、直感的には容易に想像されるものであろう。ただし本節の目的の一つは、このような直感的な推測に対してその理論的・定量的な根拠を与えることにある。定量的な評価が行えてこそ、初めて最適性を論じることが可能となる。

しかし、本節の方法はSSKFのような非直交変換に対して適用することはできない。これは(3.2)式に示した符号化ゲインが、変換の直交性を前提として定義されたものであることに起因する。一方、予測符号化に対して与えられている符号化ゲインは、そこにバンド分割の概念がないためにそのまま利用するのは無理であろう。そこで非直交変換の特性評価のためには、これまでにはない新しい評価尺度を定義してやらねばならない。

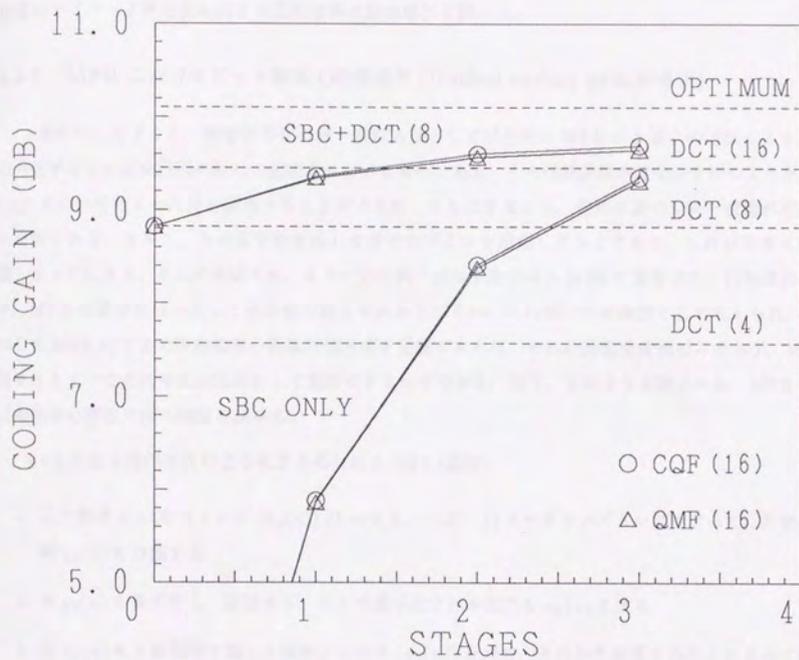


図 3.3: サブバンド符号化の圧縮効率 (2): 直交系フィルタバンク

3.2 UCG に基づく特性評価

SSKF のような非直交変換の圧縮効率の理論解析方法は、これまで非直交変換自体が軽視されてきたために前例がない。ここでは MFB に対するビット割当ての最適化問題に関する検討を行い、その結果として非直交変換にも適用可能な評価尺度 UCG (unified coding gain) を導出する。そしてその UCG を用いて、任意のサブバンド符号化に対する圧縮効率の理論解析を試みる。

3.2.1 MFB におけるビット割当ての最適化 (Unified coding gain の導出)

2 章に示したように、画像符号化の統一表現方法としては行列と MFB の 2 通りの方法がある。前節の符号化ゲインは直交行列に対して定義されているものであり、その行列表現が直交性を持つような QMF と CQF についてはこの尺度を適用することができた。これはすなわち、行列に基づく統一表現が有効に機能した例である。ただし、その直交性を外して符号化ゲインを定義しようとする、これが非常に難しい問題となってしまう。そこで本節では、もう一方の統一表現手法である MFB に着目する。行列表現とは異なり、MFB の場合には一見して直交性の概念がわかりにくい。これは一つの長所として考えられ、図 3.4 に示した MFB のままで圧縮効率の理論評価尺度を定義できれば、それは非直交変換のみならず、MFB で表現されるすべての符号化方式に対して適用できるはずである。以下、このような観点から、MFB に対する圧縮効率の評価尺度の構築を試みる。

MFB の基本操作は次のようにまとめられる (図 3.4 参照)。

1. 入力信号 $x(n)$ をフィルタ $\{h_k(n)\}$ ($k = 0, 1, \dots, K-1$) とサブサンプリングによって、 K 個の信号系列 $y_k(n)$ に分割する。
2. 各 $y_k(n)$ を量子化し、伝送する。ここで量子化された出力を $u_k(n)$ とする。
3. 各 $u_k(n)$ に 0 値補間を施した後フィルタ $\{g_k(n)\}$ を掛け、それらを加算することによって復元信号 $\hat{x}(n)$ を得る。

このとき、再生誤差 $r(n)$ 、および量子化誤差 $q_k(n)$ は、それぞれ

$$r(n) = x(n) - \hat{x}(n) \quad (3.5)$$

$$q_k(n) = y_k(n) - u_k(n) \quad (3.6)$$

によって与えられる。

次に $x(n)$ の総数を N 、 $y_k(n)$ の総数を N_k とし、パラメータ $\alpha_k = N_k/N$ を定義する。このとき伝送情報の総数が入力数に一致する、いわゆるクリティカルサンプリング (critical sampling) であれば、

$$\sum_{k=0}^{K-1} \alpha_k = 1 \quad (3.7)$$

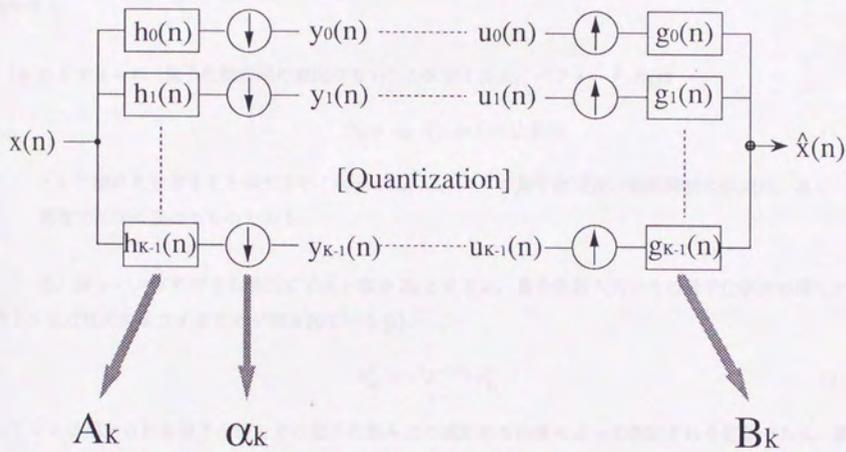


図 3.4: 量子化操作を含む MFB

が成立する。この条件は、言い替えれば、符号化の対象となる信号の数が原信号の数に一致することを表している。

次に分散を σ で表し、

$$\sigma_{y_k}^2 = A_k \cdot \sigma_x^2 \quad (3.8)$$

$$\sigma_r^2 = \sum_{k=0}^{K-1} B_k \sigma_{q_k}^2 \quad (3.9)$$

を満たすパラメータ A_k , および B_k を考える。ここで A_k はフィルタ係数 $h_k(n)$ と入力信号の相関係数から、 B_k はフィルタ係数 $g_k(n)$ と量子化誤差の相関係数、および上記のパラメータ α_k からそれぞれ求めることができる。

- ここでさらに“量子化誤差間の相関はない”と仮定すると、パラメータ B_k は

$$B_k = \alpha_k \cdot \{g_k(n)\} \text{の自乗和} \quad (3.10)$$

として簡単に求めることができる (Appendix A.2)。この量子化誤差の無相関性の仮定は、高レート環境では特に妥当なものである。

一方、第 k -バンドに対する割当てビット数を R_k とすると、量子化器入力とその量子化誤差の間には次のような近似式が成立することが知られている [1]。

$$\sigma_{q_k}^2 \simeq \epsilon^2 2^{-2R_k} \sigma_{y_k}^2 \quad (3.11)$$

ここで ϵ は用いられる量子化器とその量子化器入力の統計的な性質によって決定される定数であり、量子化パフォーマンスファクタ (quantization performance factor) と呼ばれる。

以上の表記に基づき、次のようなビット割当ての最適化問題を考える：

$$\sum_{k=0}^{K-1} \alpha_k R_k = R(\text{const.}) \quad (3.12)$$

を制約条件とした

$$\sigma_r^2 = \sum_{k=0}^{K-1} B_k \sigma_{q_k}^2 \quad (3.13)$$

の最小化。

ここで各信号系列に対する量子化パフォーマンスファクタが等しいものと仮定し、ラグランジェの未定乗数法

$$\frac{\delta}{\delta R_k} \left\{ \sigma_r^2 - \lambda \left(R - \sum_{k=0}^{K-1} \alpha_k R_k \right) \right\} = 0 \quad (3.14)$$

を適用する。するとその結果として、第 k -バンドに対する最適ビット割当て $R_{k,opt}$ およびそれに対応する量子化誤差分散 $\sigma_{qk,opt}^2$ がそれぞれ

$$R_{k,opt} = R + \frac{1}{2} \log_2 \frac{A_k B_k}{\alpha_k \prod_{k=0}^{K-1} \left(\frac{A_k B_k}{\alpha_k} \right)^{\alpha_k}} \quad (3.15)$$

$$\sigma_{qk,opt}^2 = \frac{\alpha_k}{B_k} \cdot \prod_{k=0}^{K-1} \left(\frac{A_k B_k}{\alpha_k} \right)^{\alpha_k} \cdot \epsilon^2 2^{-2R} \sigma_x^2 \quad (3.16)$$

として求められる。故に、再生誤差分散の最小値は

$$\min\{\sigma_r^2\} = \prod_{k=0}^{K-1} \left(\frac{A_k B_k}{\alpha_k} \right)^{\alpha_k} \cdot \epsilon^2 2^{-2R} \sigma_x^2 \quad (3.17)$$

によって与えられる。

- 量子化パフォーマンスファクタ ϵ がすべての信号系列に対して等しいという仮定は厳密性に欠けるが、理論解析としてはある程度やむを得ないものである。直交変換の符号化ゲインの導出に当たり、Jayant と Noll も同様の仮定を行っている

符号化ゲインは一般的に、“等レート環境下における各変換手法の再生誤差分散 $\min\{\sigma_{r,TR}^2\}$ と PCM の場合の再生誤差分散 $\sigma_{r,PCM}^2$ の比”として定義される。故に PCM では

$$\sigma_{r,PCM}^2 = \epsilon^2 2^{-2R} \sigma_x^2 \quad (3.18)$$

が成立するから [1]、結局 MFB に対する符号化ゲイン G_{MFB} が

$$G_{MFB} = \frac{1}{\prod_{k=0}^{K-1} \left(\frac{A_k B_k}{\alpha_k} \right)^{\alpha_k}} \quad (3.19)$$

によって与えられることになる。

(3.19) 式に示した評価尺度の意義は、それが MFB として表現されるすべての変換方式に対して適用可能なことにある (図 3.5)。すなわち、本節の課題であった非直交変換の特性評価のみならず、この評価尺度は従来の予測符号化や直交変換に対しても適用可能である。そこでこの評価尺度を以下 “unified coding gain” (UCG) と呼ぶ。

予測符号化への適用

1 次の閉ループ DPCM を考える。このとき、 $K = 1$ 、 $\alpha_0 = 1$ 、 $A_0 = 1 - \rho^2$ 、 $B_0 = 1$ であり、 $G_{DPCM} = (1 - \rho^2)^{-1}$ が得られる。

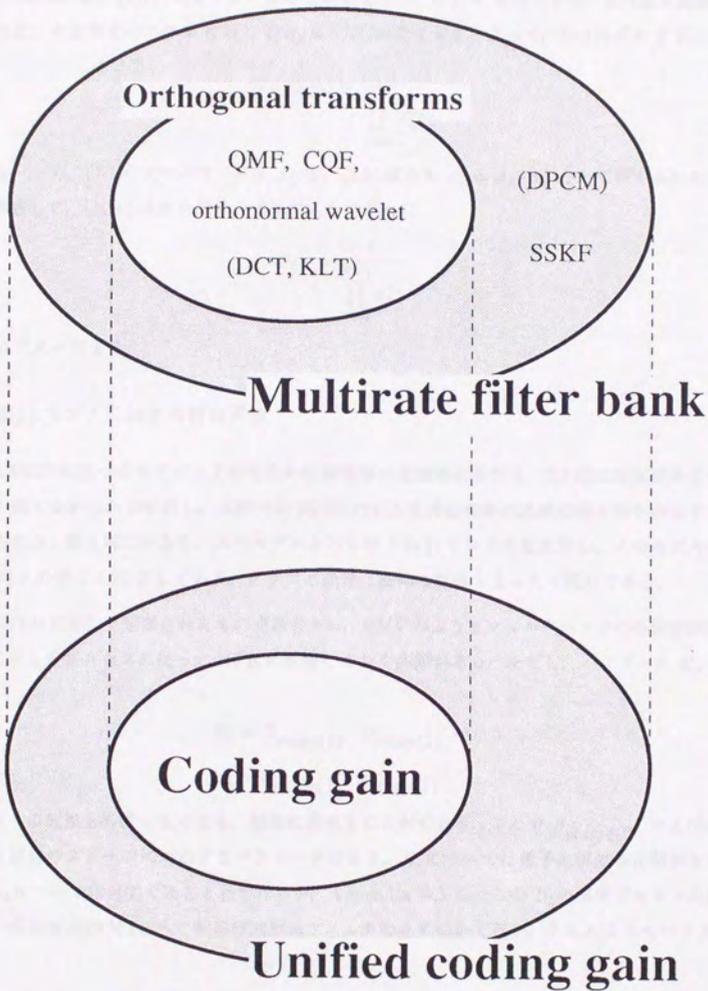


図 3.5: UCG の適用範囲

直交変換への適用

直交変換の場合、まず(1.14)式よりすべての k に対して $B_k = 1/N$ が成立する。また直交変換を MFB として考えれば、やはりすべての k に対して $\alpha_k = 1/N$ が成立する。よって、(3.19)式により、

$$G_{MFB,TC} = \frac{1}{\prod_{k=0}^{K-1} A_k^{\alpha_k}} \quad (3.20)$$

が与えられる。一方、(1.19)式のパラメータ σ_k^2 は、(3.8)式から $\sigma_k^2 = A_k \cdot \sigma_x^2$ として与えられる。ここで(1.13)式を考慮して、(1.19)式の分母に上式を代入してやれば

$$G_{TC} = \frac{1}{\prod_{k=0}^{K-1} A_k^{\alpha_k}} \quad (3.21)$$

となり(3.20)式に一致する。

3.2.2 AR(1) モデルに対する特性評価

本節では UCG に基づくサブバンド符号化の圧縮効率の理論解析を行う。3.1 節では直交系のサブバンド符号化しか扱えなかったのに対し、本節では SSKF のような非直交系の圧縮効率も明らかにする。比較条件は基本的に 3.1 節と同じである。入力モデルとしては AR(1) プロセスを仮定し、そのモデルパラメータ(相関係数) ρ の値は 0.95 としている。グラフの横軸、縦軸の意味もまったく同じである。

UCG は MFB に対して定義されたものであるから、QMF のようなフィルタバンクの多段接続はあらかじめ 2.2 節に示したプロセスに従って MFB に転換しておく必要がある。ただし、パラメータ B_k と α_k については

$$B_k = B_{\text{stage}(1)} \cdot B_{\text{stage}(2)} \cdots \quad (3.22)$$

$$\alpha_k = \alpha_{\text{stage}(1)} \cdot \alpha_{\text{stage}(2)} \cdots \quad (3.23)$$

として MFB への転換も考慮しなくとも、簡単に求めることができる。ここで $B_{\text{stage}(i)}$ および $\alpha_{\text{stage}(i)}$ はそれぞれ i 段目のステージに対応するパラメータであり、 B_k については量子化誤差の相関がないと仮定している。 α_k については自明であると思われるが、Appendix A.3 にはこの B_k の導出プロセスの証明を与える。故に、多段接続の MFB への転換は送信側フィルタのみについて行い、これによってパラメータ A_k を算出する。

- ブロック直交変換に関しては G_{MFB} と G_{TC} は完全に一致するが、直交系サブバンド符号化に関しては G_{MFB} の方が若干高い値を示す。これは 3.1 節の解析結果には有限入力数の問題に起因する環状たみこみ(2.1 節参照)の影響が反映されているのに対して、本節の結果にはこの問題が反映されていないためである。この問題については“有限入力 vs. 無限入力”として後述する。

直交 vs. 非直交

まず図 3.6は、直交系フィルタバンクとしての QMF(16), CQF(16), 正規直交 wavelet(16)に加えて、非直交系として SSKF(5×3), および SSKF(3×5)の UCG の比較を行ったものである。ここではすべてピラミッド分割を仮定し、参考として DCT(2ⁿ)の符号化ゲインの値も付加している。前述のように、DCTの符号化ゲインと UCG は完全に一致するが、直交系フィルタバンクの場合には UCGの方が若干高い値を示しており、図 3.3とはわずかに異なった結果となっている。

この図からは

- SSKF(3×5)を除き、各フィルタバンクの圧縮効率はバンド数が増加するにつれて DCT を凌ぐようになる
- SSKF(5×3)は、タップ数が少なく、かつ非直交であるにも関わらず、多段接続数が少ない、すなわちバンド数が少ない限りにおいては直交系を凌ぐ圧縮効率を実現する

ことなどがわかる。

バンド数が少ないという制約があるにせよ、少タップ、非直交、かつ周波数特性の優れないフィルタバンクが多タップ、直交、かつシャープな周波数特性を有するフィルタバンクを凌ぐという結果は、いくらか奇妙な感じがするものと思われる。しかし、以下のようないくつかの例を見れば、これも妥当な結果と言わざるを得ない。

• タップ数の反例

- 閉ループ DPCM は、高レート環境では KLT や DCT よりも優れた、ほぼ最適な圧縮効率を実現できる (1.3 節参照)。この場合、タップ数はわずかに 2 に過ぎない。

• 直交変換の反例

- 単位行列も直交行列であるが、データ圧縮に使えばそれは PCM にほかならず、符号化ゲインはない。

• 周波数特性の反例

- DFT と DCT を比べた場合、周波数特性は明らかに DFT の方が優れているが、圧縮効率は明らかに DCT の方が優れている。

従来の線形変換の中で、圧縮効率の最適化を目的としてフィルタ係数が決定されているのは DPCM と KLT だけである。DCT が広く用いられている理由は、KLT が相関の高い入力 ($\rho \rightarrow 1$) に対して DCT への漸近性を示す、という理論的な基盤があるからであり、周波数特性はそこではあまり大きな意味を持たな

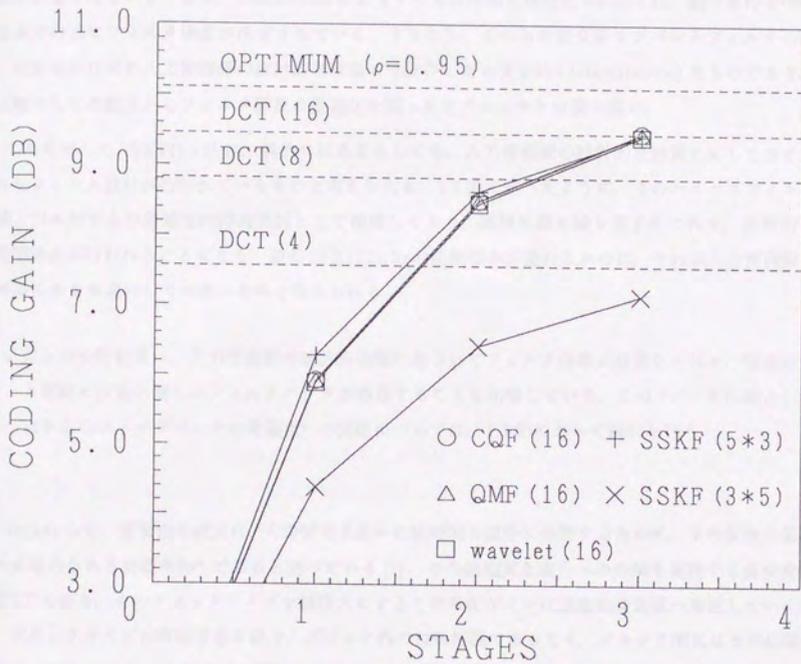


図 3.6: サブバンド符号化の特性評価 (3): 直交 + 非直交フィルタバンク

い。あるいはまた、直交変換ではあっても、アダマール変換やハール変換などの圧縮効率はそれほど優れたものではない。すなわち、データ圧縮を目的とする変換手法の規定において大切なことは“入力情報源の統計的な性質に基づくフィルタ係数の最適化”であり、直交性の成立やシャープな周波数特性の実現はデータ圧縮として本質的な問題解決には帰着しない。

一方、QMFとCQFは共に、(フーリエ領域における)周波数特性の最適化を設計基準としてフィルタ係数が決定されている。また、Daubechiesによって与えられた正規直交 wavelet は、調和条件と呼ばれる拘束条件の基にフィルタ係数が決定されている。すなわち、これらの直交系サブバンドフィルタの設計基準、拘束条件は共に入力情報源の統計的な特徴には依存しない決定的 (deterministic) なものであり、データ圧縮としての観点からフィルタ係数の最適化を図ったアプローチとは言い難い。

これに対してSSKF(5×3)は、偶然ではあるとしても、入力情報源の統計的な性質に対してほぼ最適化されたフィルタ設計が行われているものと考えられる。2.2節に述べたように、そのハイパスフィルタは高相関入力に対する準最適内挿補間器として機能しており、階層処理を繰り返すにつれて、段階的に入力の相関除去が行われることになる。逆にSSKF(3×5)の圧縮効率が優れないのは、それが入力情報源の統計的特徴にあまり適合していないためと考えられる。

- 以上の検討結果は、入力情報源の統計的な性質に基づいてフィルタ係数が最適化された、従来のフィルタ構成にはない新しいフィルタバンクが存在することを示唆している。この“データ圧縮としての観点からのフィルタバンクの最適化”の問題については、4.3節において検討を行う。

Jayantらは、直交性の成立は“入力信号を互いに無相関な成分に分解するために、その変換の基底ベクトルに求められる必要条件”であると述べている [1]。この無相関な成分への分解を実現する直交変換こそがKLTであり、そのブロックサイズを無限大にすると符号化ゲインは理論的最適値へ漸近していく。しかし、ブロックサイズが有限である以上、ブロック内では無相関であっても、ブロック間にはまだ相関が残っている。故に、ブロックサイズの増加と共にブロック間の相関が小さくなり、圧縮効率が向上することになる。

一方、サブバンド符号化では、たとえ直交系フィルタバンクを用いたとしても、バンド分割数が少なければその低周波成分間に強い相関が残っていることは容易に想像できる。その相関を定量的に評価したわけではないが、バンド分割数が少なればかなり高い値を示すであろう。さらにその高周波成分に着目した場合、その相関はもともと小さいであろうが、予測誤差的なSSKF(5×3)の出力の方がより小さな相関値を示すであろう。故に、サブバンド符号化では、バンド数が少なければ変換出力の無相関化は原理的に困難であり、直交系が非直交系に劣る特性を呈したとしても何等不思議ではない。

フル分割 vs. ピラミッド分割

3.1節では、直交系のフィルタバンクを用いた場合にピラミッド分割はフル分割と同等の圧縮効率を実現できることを理論的に明らかにした。ただし、直交系の場合にはピラミッド分割がフル分割を凌ぐことは絶対にありえない。これは非直交系のフィルタバンクでも成立するか？

図 3.7は、CQF(16) と SSKF(5×3) についてフル分割とピラミッド分割の特性比較を行ったものである。ここでは直交フィルタの代表として CQF(16) を使ったが、これは CQF(16)、QMF(16)、正規直交 wavelet(16) の中で、CQF(16) がわずかではあるが理論値として最大値を示したからである。

この図より

- CQF の場合はピラミッド分割よりもフル分割の方が良好な特性を呈する

のに対して、

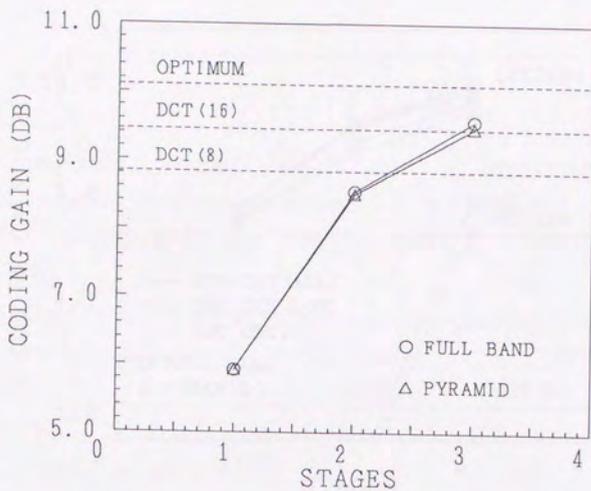
- SSKF の場合はフル分割よりもピラミッド分割の方が良好な特性を呈する

ことがわかる。理想フィルタの場合には、バンド分割を繰り返せば繰り返すほど圧縮効率が向上するから、周波数特性の良好な CQF の結果はこれを反映したものとなっている。しかし、SSKF の結果は、この理論的な予測とは矛盾する結果となっている。この結果は定量的には以下のように説明できる：前述のように、SSKF のハイパスフィルタは高相関入力に対する“内挿予測フィルタ”として考えることができる。よってその出力は予測誤差と見なすことができ、出力信号間の相関は非常に低くなることが予想される。一方、無相関入力 ($\rho=0$) のバンド分割を試みた場合、直交フィルタであれば必ず $G_{MFB}=0$ (dB) となって符号化利得は良くも悪くもならないが、非直交フィルタである SSKF の場合には $G_{SBC}=-0.65$ (dB) となり、符号化利得の値が負になってしまう。すなわち、SSKF のハイパス出力をさらに SSKF で分割すると、結果として符号化利得の低下をもたらすことになる。

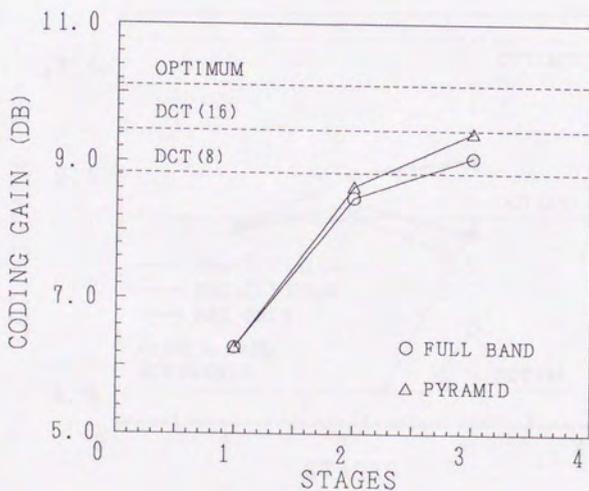
続いて図 3.8は、CQF(16)、SSKF(5×3) それぞれに対してさらに DCT(8) を施した場合の特性比較を行ったものである。ここで ALL はすべてのバンドに対して DCT を施した場合であり、LOW は低周波成分のみ DCT を施した場合である。この結果からは、

- CQF、SSKF ともに図 3.7と同様の特性を呈する
- CQF には理論的最適値への漸近性が見られるが、SSKF は 9.7dB 付近で頭打ちになる

ことなどがわかる、

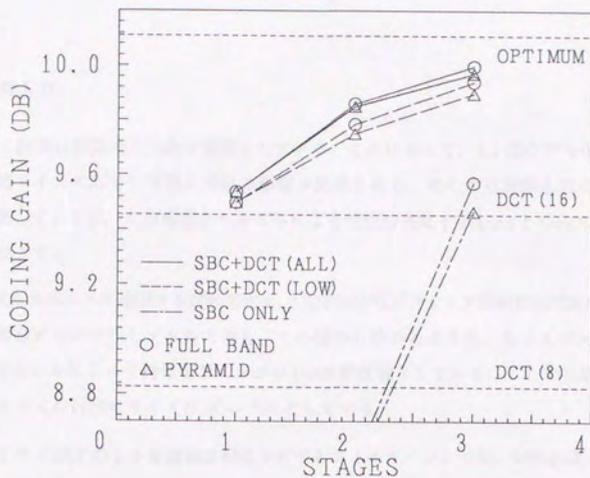


(a) CQF(16) の場合

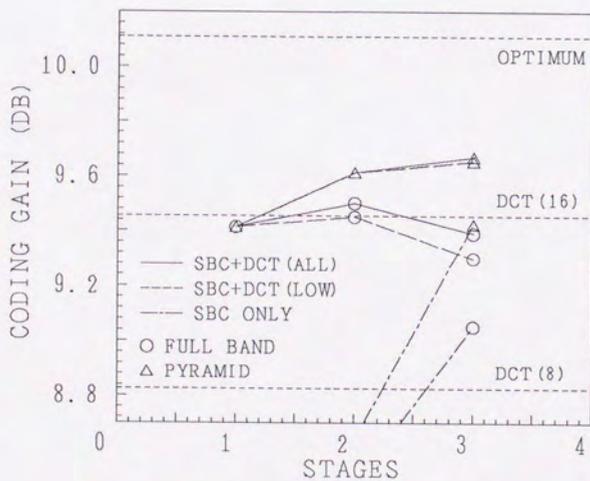


(b) SSKF(5x3) の場合

図 3.7: サブバンド符号化の特性評価 (4): フル分割 vs. ピラミッド分割



(a) CQF(16) の場合



(b) SSKF(5x3) の場合

図 3.8: サブバンド符号化の特性評価 (5): フル分割 + DCT vs. ピラミッド分割 + DCT

有限入力 vs. 無限入力

UCGに基づく評価は無限の入力数を前提にしている。これに対して、3.1節の符号化ゲインに基づく評価では、その行列サイズに応じて有限入力数の影響が反映される。ここでは有限入力の問題を環状たたみこみによって解決しているが、この環状たたみこみによる特性の劣化を G_{MFB} と G_{TC} の値の差として定量的に示すことができる。

図3.9は、環状たたみこみに起因する特性劣化を、CQF(16)のピラミッド分割に低周波成分のみのDCT(8)を組み合わせた場合について示したものである。この図からわかるように、たとえばステージ数が3の場合には、環状たたみこみによって符号化ゲインが0.15dB程度低下している。ここで有限入力数、すなわち G_{TC} の算出に当たっての行列のサイズは $N = 256$ としている。

ただし、QMFやSSKFのような直線位相性を有するフィルタバンクでは、対称拡張(2.1節参照)によって境界部の問題を解決できる。この場合には境界部で信号系列を鏡像的に折り返すために、相関係数が低下することはない。このために、UCGを用いた結果がそのまま利用できるものと考えられる。

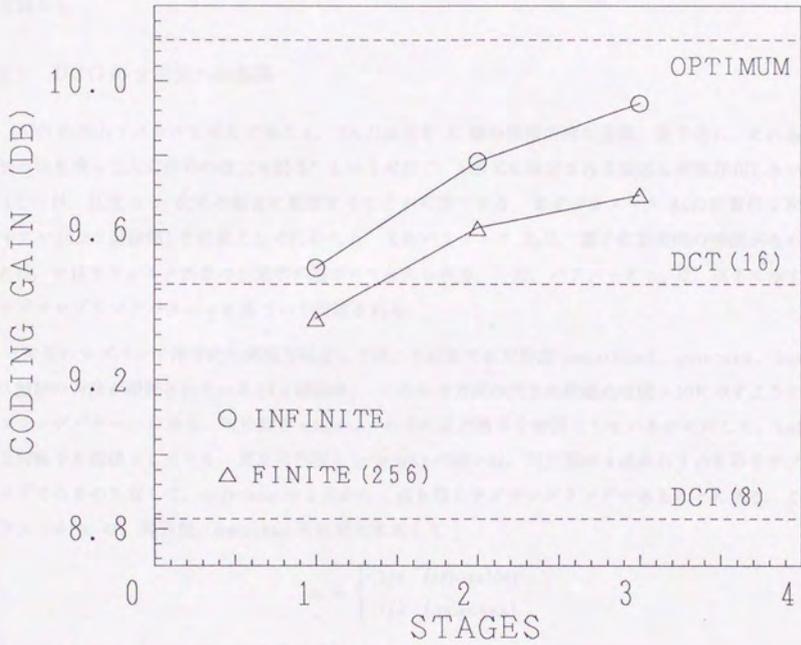


図 3.9: サブバンド符号化の特性評価 (6): 有限入力 vs. 無限入力

3.3 2次元のサブバンド符号化の特性評価

画像信号のサブバンド符号化は、一般的に2次元入力を前提としている。また quincunx サブサンプリングに基づく非可分型フィルタのように、2次元固有のフィルタ構成が用いられることもある。これに対して、前節までの1次元を対象とした結果が2次元にもそのまま適用可能であるとは考えにくく、あるいはまた非可分型の特性評価は不可能である。そこで本節では、2次元のサブバンド符号化の圧縮効率の理論解析を試みる。

3.3.1 UCG の2次元への拡張

UCG の導出プロセスを考えてみると、“入力信号を K 個の信号系列に分離、量子化し、それらの信号系列の和を取って入力信号の復元を図る”というだけで、1次元に限定される要因は何等存在しない。そこで UCG は、任意の m 次元の場合に拡張することが可能である。まずパラメータ A_k の計算は2次元の入力モデル(3.3.2節参照)を対象として行われる。またパラメータ B_k は、量子化誤差間の相関がないと仮定すれば、やはりフィルタ係数の自乗和に基づいて与えられる。一方、パラメータ α_k は、以下に示す2次元のサブサンプリングパターンに基づいて決定される。

2次元のサブバンド符号化の実現方法としては、これまでに可分型(separable), quincunx, hexagonal の3種類の方法が報告されている(1.2節参照)。これらの方式の大きな相違点は図3.10に示すようにサブサンプリングパターンにあり、可分型と quincunx が共に正格子を前提としているのに対して、hexagonal は三角格子を前提としている。また可分型と quincunx の違いは、可分型が4点から1点を取るサブサンプリングであるのに対して、quincunx は2点から1点を取るサブサンプリングであることにある。このときパラメータ α_i は、可分型, quincunx それぞれに対して

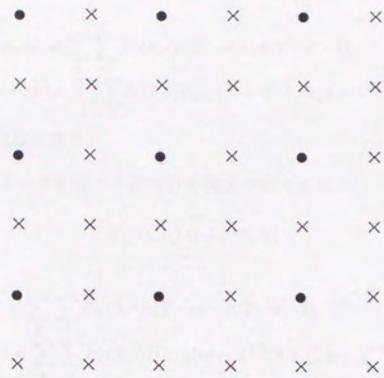
$$\alpha_i = \begin{cases} 1/4 & (\text{separable}) \\ 1/2 & (\text{quincunx}) \end{cases} \quad (3.24)$$

として与えられる。本節では対象を可分型と quincunx に限定し、三角格子を前提とする hexagonal については検討を行わない。

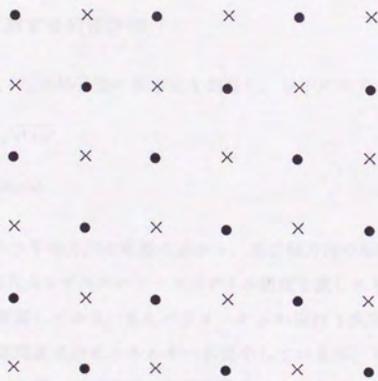
一方、2次元のサブバンド符号化も、1次元と同様にフィルタバンクの多段接続によって周波数分割を行うことが多い。故に、このような多段接続構成に対する UCG の適用にあたり、その並列構成への転換手法を定式化しなければならない。ただし、 B_k と α_k の扱いは3.2節の1次元の場合と同じであり、並列構成への転換の目的はパラメータ A_k の算出にある。

まず、可分型構成の場合には、

$$H_1(m, n) = h_1(m, n) \quad (3.25)$$



(a) Separable



(b) Quincunx

図 3.10: 2次元のサブバンド符号化のサブサンプリングパターン

として

$$\begin{cases} y_i(m, n) = \sum_k \sum_l H_i(k, l) x(2^i \cdot m - k, 2^i \cdot n - l) \\ H_i(m, n) = \sum_k \sum_l h_i(k, l) H_{i-1}(m - 2^{i-1} \cdot k, n - 2^{i-1} \cdot l) \end{cases} \quad (3.26)$$

によって並列構成への転換が可能となる。

また quincunx サブサンプリングに基づく非可分形構成の場合には、

$$H_1(m, n) = h_1(m, n) \quad (3.27)$$

として

$$\begin{cases} y_{2i}(m, n) = \sum_k \sum_l H_{2i}(k, l) x(2^i \cdot m - k, 2^i \cdot n - l) \\ H_{2i}(m, n) = \sum_k \sum_l h_{2i}(k, l) H_{2i-1}(m - 2^{i-1}(k+l), n - 2^{i-1}(-k+l)) \end{cases} \quad (3.28)$$

および

$$\begin{cases} y_{2i+1}(m, n) = \sum_k \sum_l H_{2i+1}(k, l) x(2^i(m+n) - k, 2^i(-m+n) - l) \\ H_{2i+1}(m, n) = \sum_k \sum_l h_{2i+1}(k, l) H_{2i}(m - 2^i \cdot k, n - 2^i \cdot l) \end{cases} \quad (3.29)$$

によって並列表現が可能となる。ただし、 $y_{2i+1}(m, n)$ の座標軸は、 $y_{2i}(m, n)$ のそれに対して 45° 回転したものとなる。

3.3.2 2次元入力モデルに対する特性評価

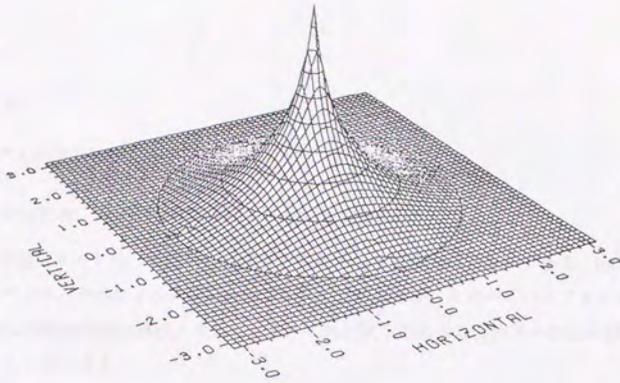
2次元入力モデルとしては、正方格子型の標本化を前提に、以下に示す2つのモデルを考える。

1. 等方性モデル $\rho_{x,y} = \rho^{\sqrt{x^2+y^2}}$
2. 可分性モデル $\rho_{x,y} = \rho^{|x|+|y|}$

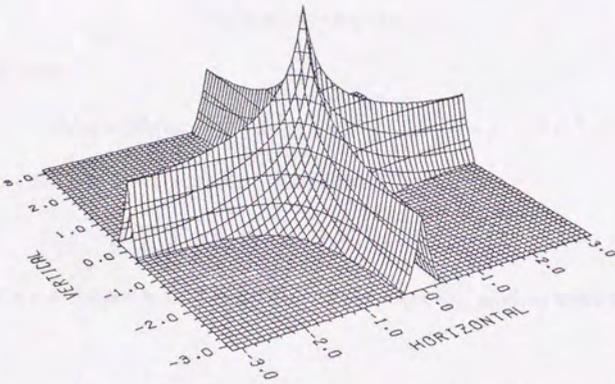
ここで $\rho_{x,y}$ は、2つの画素間の水平軸方向の距離の差が x 、垂直軸方向の距離の差が y の場合の相関係数を表している。図 3.11 はこれらの入力モデルのパワースペクトル密度を表したものである ($\omega_x, \omega_y = -\pi \sim \pi$)。ここで z 軸はログスケールで表現しており、またパラメータ ρ の値は 1 次元の場合と同様に 0.95 としている。等方性モデルの場合には低周波成分にエネルギーが集中しているが、可分性モデルの場合にはその相関の方向性が強く現れており、水平、垂直方向の高周波成分のエネルギーが DC 成分に対して相対的に大きくなっている。

画像データの多くは以上の2つのモデルのどちらかに従うことが、実験結果として明らかにされている。一般的に、自然物体は等方性モデルに、人工物は可分性モデルに適合するものと考えられている。また、それらのモデルに対するデータ圧縮の効果は、2次元の予測符号化を用いた場合の結果が明らかにされており、高周波成分が相対的に大きいにも関わらず、可分性モデルの圧縮効率のはるかに高くなる [1]。

次に評価の対象としては、以下の3つのフィルタ構成を考える。



(a) 等方性モデル



(b) 可分性モデル

図 3.11: 2次元入力モデル

1. CQF(16)
2. SSKF(5×3)
3. 非可分型フィルタ [50,53]

ここで CQF と SSKF は、前述の可分型構成によって 2 次元に拡張する。

一方、非可分型フィルタは、当初から 2 次元を前提として構成されたものであり、quincunx サンプリングを前提としている。そのフィルタ係数の配置を図 3.12 に示すが、そのハイパスフィルタは SSKF(5×3) と同様に 2 次元の内挿補間的な構成となっている。これは以下のような 2 次元の完全再構成条件の一つの解として求めることができる。

図 1.3 において、信号系列を 2 次元に、サブサンプリングを quincunx に置換した構成を考える。このとき、完全再構成条件は、折り返しひずみの除去条件

$$G_0(z_0, z_1) = H_1(-z_0, -z_1) \quad (3.30)$$

$$G_1(z_0, z_1) = -H_0(-z_0, -z_1) \quad (3.31)$$

を満足する上での

$$H_0(z_0, z_1)G_0(z_0, z_1) - H_0(-z_0, -z_1)G_0(-z_0, -z_1) = 2 \cdot z_0^{-M} z_1^{-N} \quad (3.32)$$

の成立として表される。この (3.32) 式は、空間領域では

$$\sum_k \sum_l h_0(k, l) g_0(2i - k, 2j - 1 - l) = \delta_i \delta_j \quad (3.33)$$

と書き換えることが可能であり、この条件下でフィルタ $h_0(m, n)$ 、 $g_0(m, n)$ を決定してやればよい。

まず図 3.13、および図 3.14 は、等方性入力、可分性入力それぞれに対する評価結果を表している ($\rho = 0.95$)。ここでは、上記の 3 つのフィルタ構成に対してフル分割とピラミッド分割を行った場合の UCG を求めている。ただし、横軸は可分型フィルタに対するステージ数を表しており、これは分割バンド数に対応することになる。このとき特に可分型と非可分型のマッチングを図るために、それぞれの理想フィルタを用いた場合にはその低周波成分が一致するように、非可分型のステージ数を考慮した上で重ね合わせを行っている。例えばピラミッド分割の場合、ステージ数が 1 の可分型は 4 バンド分割となるが、これに対応する非可分型は 3 バンド分割になる。あるいはまた、ピラミッド分割でステージ数が 2 の可分型は 7 バンド分割になるが、これに対応する非可分型は 5 バンド分割になる (図 1.7 参照)。

これらの結果から

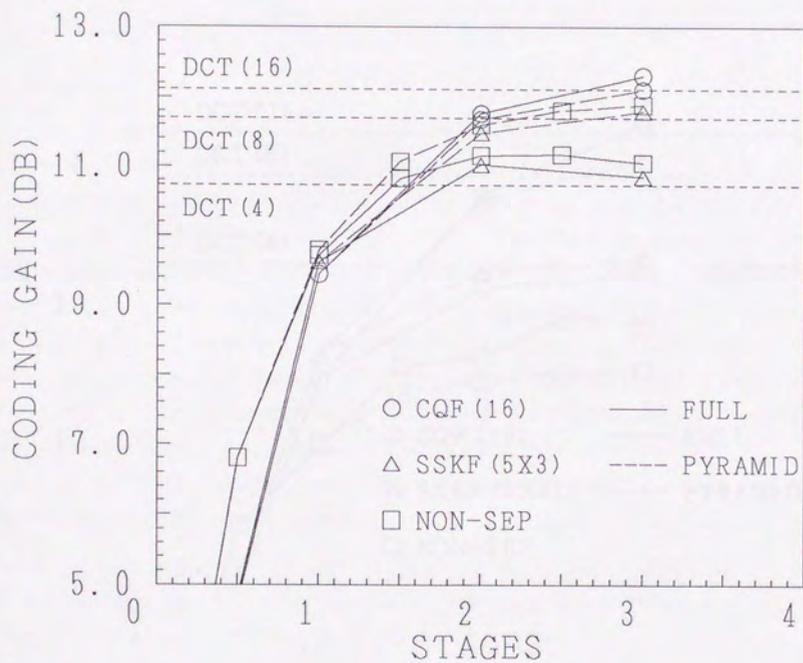


図 3.13: サブバンド符号化の特性評価 (7) : 2 次元等方性入力

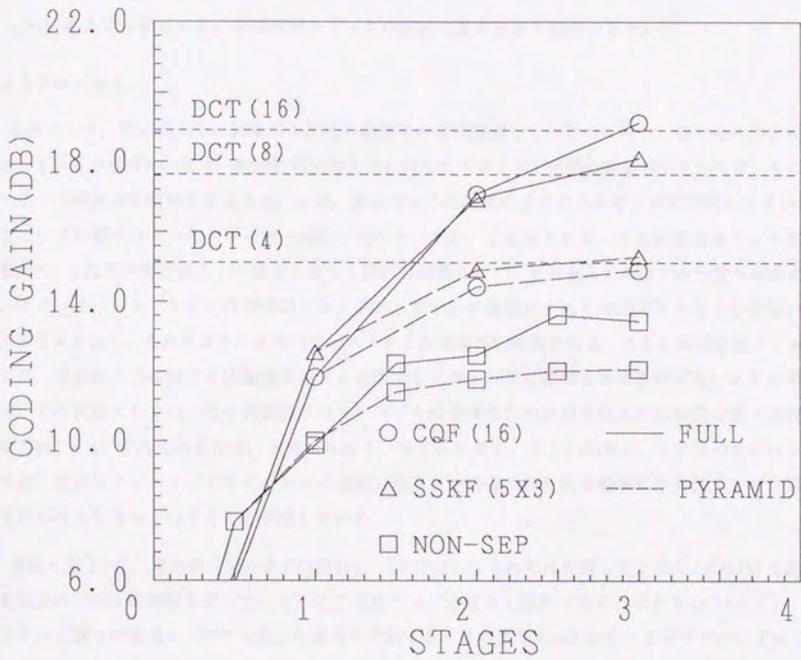


図 3.14: サブバンド符号化の特性評価 (8): 2 次元可分性入力

- 直交系フィルタバンク (CQF) の場合には、等方性入力、可分性入力それぞれに対してフル分割が有効である。
- 非直交系フィルタバンク (SSKF, 非可分型) の場合には、等方性入力に対してはピラミッド分割が有効であるが、可分性入力に対してはフル分割が有効である。
- 可分性入力の場合には、ピラミッド分割の効果があまり期待できない。
- 可分性入力の場合には、非可分型フィルタの圧縮効果があまり期待できない。

ことなどがわかる。

総括として、等方性入力の結果が1次元の結果に非常に類似しているのに対し、可分性入力の場合には異なっている点が多い。まず一般的に可分性入力に対してピラミッド分割の効果が小さいのは、そのパワースペクトル密度から明らかのように、水平、垂直方向の高周波成分のエネルギーが相対的に大きいために、ピラミッド分割では十分なエネルギー集約が図れないため、と推測される。また非直交系フィルタバンクの場合に、1次元や等方性入力の場合とは全く反対の結果として、可分性入力にはフル分割が有効になるのは、そのハイパスフィルタが内挿補間フィルタとして十分に機能せずにその出力にかなりの相関が残るため、と考えられる。これもまた、そのパワースペクトル密度から推測される。さらに非可分型フィルタの場合には、可分性入力に対する圧縮効果がフル分割でもピラミッド分割でもあまり伸びないことが示唆されるが、この理由としては、その偶数段目のステージでは予測の方向が可分性入力の相関の強い方向(水平、垂直方向)と45°ずれているため、と考えられる。いずれにせよ、以上の結果は、2次元のサブバンド符号化では、用いるフィルタバンクやバンドの分割方法などを画像の統計的な性質に十分注意して決定して行かなければならないということを示唆している。

最後に表 3.1 は、可分型フィルタ CQF(16)、SSKF(5×3) それぞれに対して2次元 DCT(8) を組み合わせた場合の UCG の評価を表している。ここではバンド分割は1回行っただけであり(4バンド)、図 3.13 のステージ数1の場合に DCT を施した場合に対応する。この表中の all はすべてのサブバンドに DCT を施した場合であり、また low は低周波成分のみに DCT を施した場合である。その結果自体は、図 3.13 の結果と同様に評価することができる。

3.3.3 シミュレーション

これまででは理論解析のみを行ってきたが、これだけでは心許ない。そこで本節では、いくつかのシミュレーションの結果を提示する。

図 3.15-図 3.18 は、256×256 のモノクローム画像 GIRL, MOON, COUPLE, AERIAL に対してサブバンド符号化を施した場合のレートひずみ特性を表したものである。それぞれの結果は、左上から順に

1. ステージ数1

表 3.1: サブバンド符号化の特性評価 (9): 可分型フィルタ + DCT

	CQF+DCT(8)		SSKF+DCT(8)		DCT(8)	DCT(16)
	all	low	all	low		
isotropic	12.70	12.11	11.86	11.83	11.70	12.12
separable	19.57	15.42	18.83	15.69	17.65	18.91

2. ステージ数 2 (ピラミッド分割)
3. ステージ数 2 (フル分割)
4. ステージ数 3 (ピラミッド分割)
5. DCT のみ
6. ステージ数 1 + DCT (すべてのバンド)
7. ステージ数 1 + DCT (低周波のみ)

を表している。ここでステージ数は図 3.13 等と同様に、可分型 2 次元フィルタに対して定義したものである。また右下には、それぞれの画像のパワースペクトル密度を示したが、これらの結果はほぼ 3.3.2 節の理論解析の結果に一致するものとなっている。

次に図 3.19, 図 3.20 は, GIRL, MOON に対する符号化処理画像の例を示したものである。これらの結果より, サブバンド符号化ではブロックひずみが生じないために, 主観的に良好な画像が得られていることがわかる。

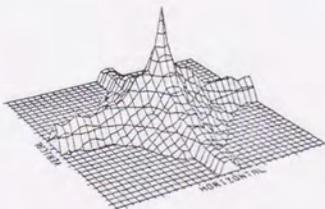
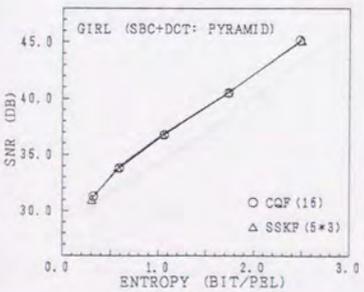
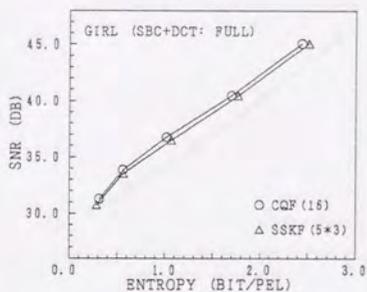
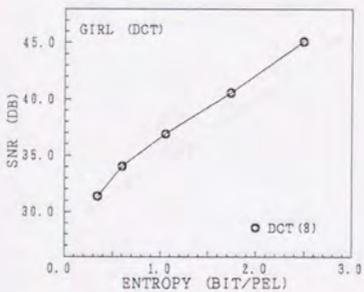
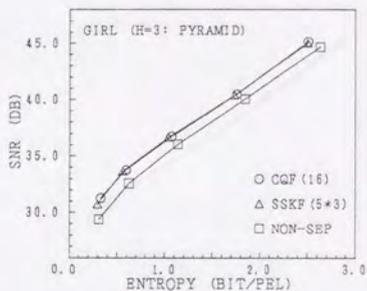
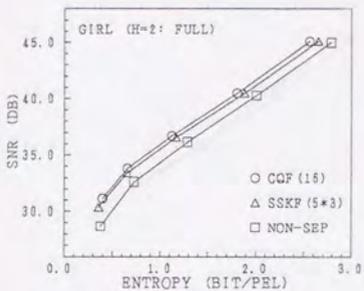
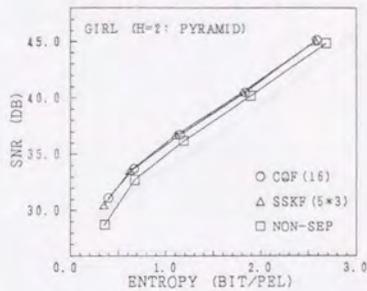
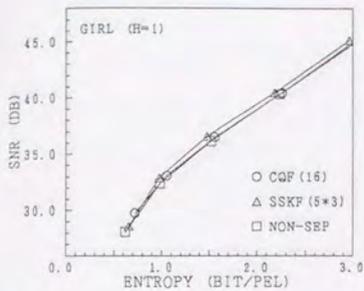


図3.15: シミュレーション結果(GIRL)

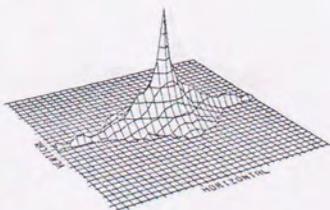
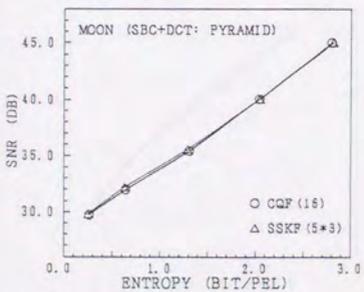
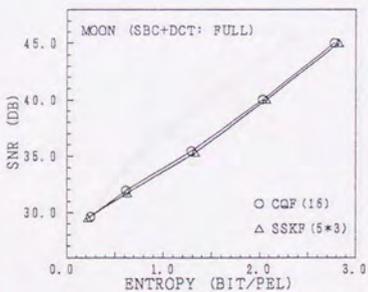
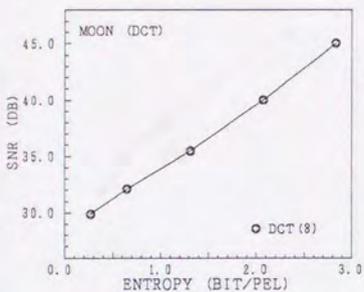
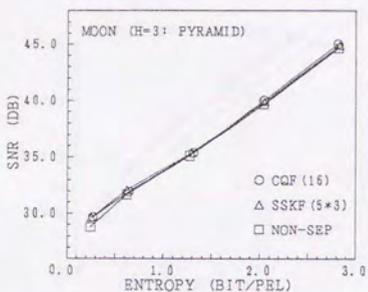
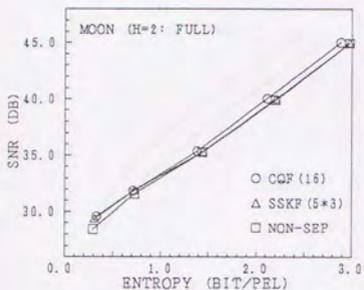
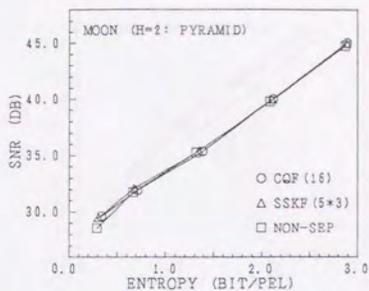
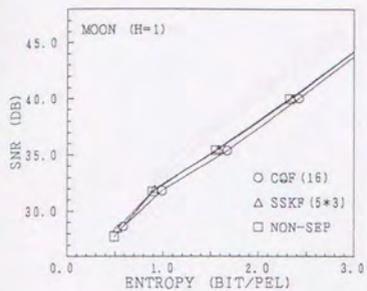


図3.16: シミュレーション結果(MOON)

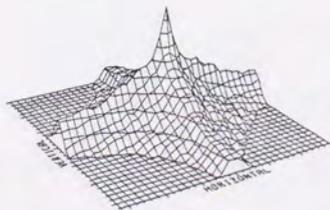
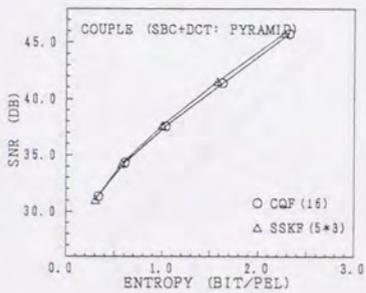
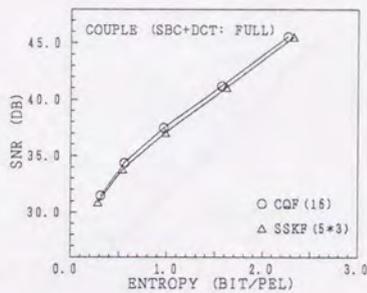
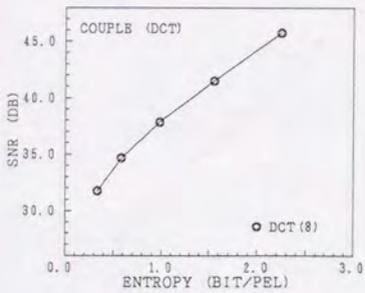
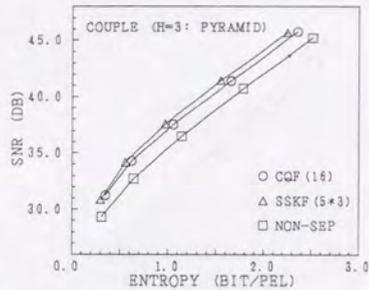
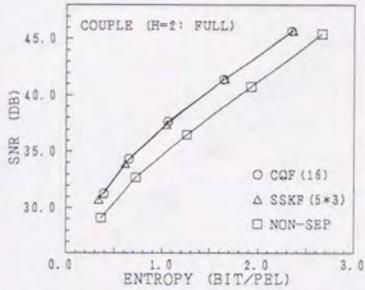
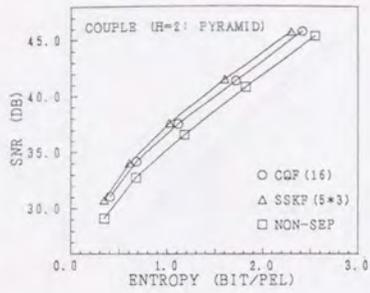
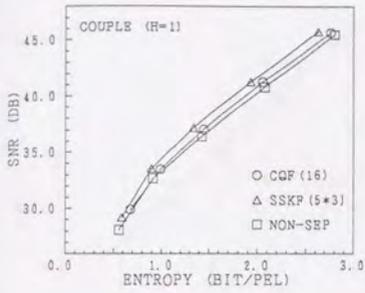


図3.17: シミュレーション結果(COUPLE)

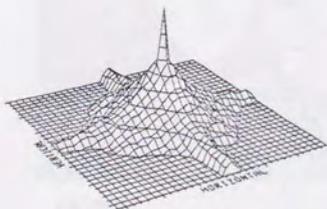
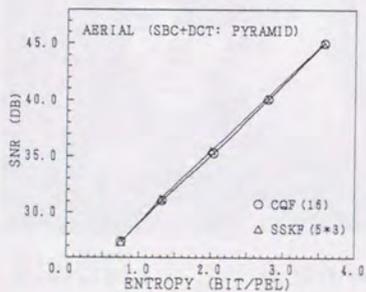
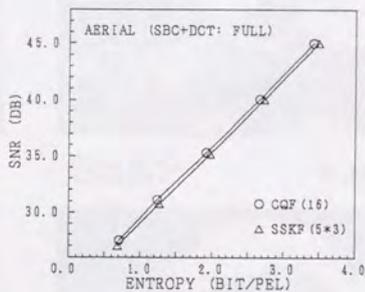
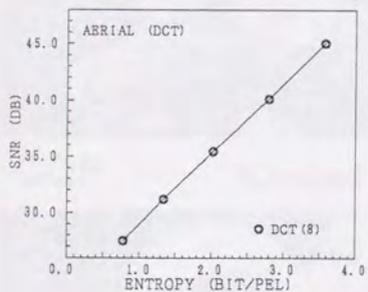
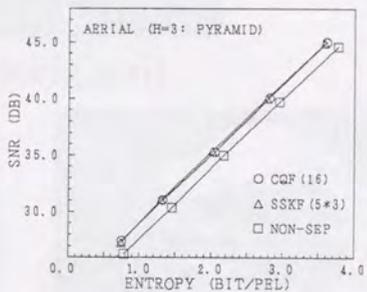
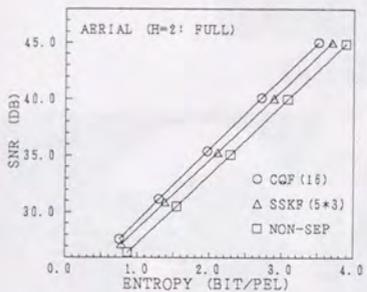
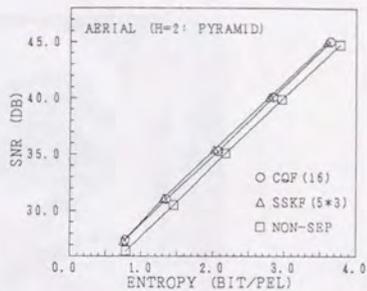
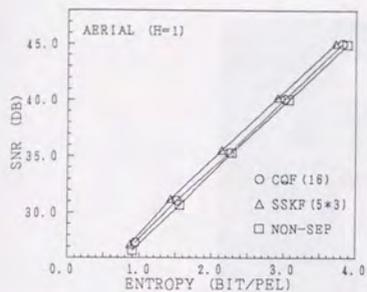


図3.18: シミュレーション結果(AERIAL)



GIRL (original)



CQF

0.3bit/pel



SSKF

0.3bit/pel



NSEP

0.3bit/pel



DCT

0.3bit/pel



MOON (original)



CQF

0.3bit/pel



SSKF

0.3bit/pel



NSEP

0.3bit/pel



DCT

0.3bit/pel

3.4 考察・検討

本章の結果として、まずサブバンド符号化では

- 多段接続数の増加、あるいは DCT との組み合わせにより、DCT 単独の場合を凌ぐ圧縮効率を実現できる

ことが理論として立証された。Wood らによって初めてサブバンド符号化が画像に適用されて以来さまざまな提案が行われてきたが、上記の結果はこれらに対して理論的な裏付けを与えるものとなる。サブバンド符号化では、そのセールスポイントとしてしばしばブロックひずみからの解放が引用されてきたが、それだけではなく、レートひずみ特性としても非常に優れた符号化方式であることがわかる。

ただし、そのバンド分割方法を工夫しないと、バンド分割してもほとんどゲインが得られなかったり(1次元直交系, 2次元直交系等方性モデル)、逆に特性が劣化することすらある(1次元非直交系, 2次元非直交系等方性モデル)。このために、入力画像の統計的な特徴に合わせて適応的にバンド分割方法を決定するなどの工夫を施す必要があるものと考えられる。

あるいはまた、本章では

- SSKF(5×3)のような非直交変換でも、分割バンド数が少ない限りにおいて、直交変換を凌ぐ圧縮効率を実現できる

ことを明らかにした。これは直交性や周波数特性が必ずしも圧縮効率の最適化に帰着する要因ではないことを示唆するものであり、非直交変換でも、使い方によっては非常に優れた符号化方式として機能する可能性があることになる。その計算量の極端な少なさ、直線位相性もまた非常に魅力的である。

もっとも、SSKF を発表した Le Gall らの主張は少タップ構成による完全再構成の実現とその直線位相性に重点が置かれており、圧縮効率はあまり重要視していなかった風がある。実際、本章では省略したが、SSKF(5×3)に並記されて提案されている SSKF(4×4)の圧縮効率はそれほど高い値は示さず、サブバンド符号化のリーダー的な存在である Vetterli にして、効率の悪い SSKF(3×5)をシミュレーションに用いているほどである [80]。あるいはまた、直線位相フィルタを用いた完全再構成の実現は、現在では双直交(biorthogonal)フィルタバンクとして活発な検討が進められているテーマである。しかし本章の結果は、これまでに提案された方式の中では SSKF(5×3)だけが優れた圧縮効率を実現するフィルタであることを示唆している。

いずれにせよ、直交性、周波数特性、フィルタ長の3つの要因は、これまではサブバンド符号化で圧縮効率を上げるための必須条件として考えられてきた風潮があった。しかし、本章の結果は、これらの先入観を理論として否定するものである。

Jayantらが述べているように、直交性の成立は“入力信号を互いに無相関な成分に分解するために、その変換の基底ベクトルに求められる必要条件”であるとされる [1]。また図 3.8が示すように、理論として明らかにした訳ではないので適当なことは言えないが、SSKF(5×3)には理論的最適値への漸近性に疑問が残る。しかしながら、たとえ直交性が理論的最適値への漸近性を保証する必要条件であるとしても、非直交変換によってそれに非常に近いところまで達することが可能であることも事実である。故に、直交・非直交に拘ることなく、入力情報源の統計的な特徴に適合したフィルタバンクの利用方法について検討を進めていくことが重要であると考えられる。

Chapter 4

完全再構成フィルタの設計

サブバンド符号化におけるフィルタ設計の問題は、そのトリガとなった QMF の提案以来十数年が経過した今日においても、デジタル信号処理の分野における大きな研究テーマとなっている。完全再構成条件 (1.3 節参照) を満足するフィルタは無限に存在しており、3 章で取り挙げたフィルタ構成はほんの一例に過ぎない。本節でもまた、この完全再構成フィルタに関する検討を行い、いくつかの新しい方式提案を行う。

4.1 ハイパスフィルタの簡略化

サブバンド符号化と階層的符号化の概念としての等価性は1.2節に述べた通りである。ただし、サブバンド符号化では複数のバンドパスフィルタによって並列的に周波数分割を図るのに対して、階層的符号化ではその高周波成分を原信号とローパス出力との差を取ることによって求めている。そこで本節では、この階層的符号化からの類推として、ハイパスフィルタが原信号とそのローパス信号との差分器として構成される完全再構成フィルタについて検討を行う。

4.1.1 1次元の場合

サブサンプリングパターンが一致している場合

上下チャネルのサブサンプリングパターンが一致する場合、行列式

$$\Delta(z) = H_0(z)H_1(-z) - H_0(-z)H_1(z) \quad (4.1)$$

を用いて、完全再構成条件(1.26)式を書き換えると

$$\begin{bmatrix} G_0(z) \\ G_1(z) \end{bmatrix} = \frac{1}{\Delta(z)} \begin{bmatrix} H_1(-z) & -H_1(z) \\ -H_0(-z) & H_0(z) \end{bmatrix} \begin{bmatrix} 2 \cdot z^{-N} \\ 0 \end{bmatrix} \quad (4.2)$$

となる。故に、 $\Delta(z) = 2\alpha \cdot z^{-N}$ が成立するならば、

$$\begin{aligned} G_0(z) &= \frac{1}{\alpha} H_1(-z) \\ G_1(z) &= -\frac{1}{\alpha} H_0(-z) \end{aligned} \quad (4.3)$$

によって完全再構成を実現できる。

一方、ハイパスフィルタ $H_1(z)$ を、原信号とローパス出力の差分器

$$H_1(z) = \beta - H_0(z) \quad (4.4)$$

として構成することを考える。ただし、 β はハイパス出力から原信号の DC 成分を除去するために原信号に乗ずるパラメータであるが、 $H_0(z)$ の係数選択は任意であり、 $H_0(\omega = 0) = \alpha$ となるようにフィルタ係数を定めても何等問題はない。そこで以下 $\beta = \alpha$ として検討を進めていくことにする。また煩雑化を避けるために $H_0(z) = H(z)$ とする。

次に、これらの関係式を(4.1)式に代入すると、完全再構成条件は

$$H(z) - H(-z) = 2 \cdot z^{-N} \quad (4.5)$$

にまで簡略化される。この(4.5)式は z 変換の定義から

$$\sum_{n=-\infty}^{\infty} \{1 - (-1)^n\} h(n) z^{-n} = 2 \cdot z^{-N} \quad (4.6)$$

に等価であるが、

$$1 - (-1)^n = \begin{cases} 0 & (n: \text{even}) \\ 2 & (n: \text{odd}) \end{cases} \quad (4.7)$$

であるから、フィルタ係数 $h(n)$ が

$$h(n) = \begin{cases} \text{arbitrary} & (n: \text{even}) \\ 1 & (n = 2K + 1 \text{ only}) \\ 0 & (\text{otherwise}) \end{cases} \quad (4.8)$$

を満足する限り、完全再構成が実現される。

一方、受信側では $G_1(z) = G(z)$ とすれば、フィルタ係数 $g(n)$ は

$$g(n) = \frac{1}{\alpha} (-1)^n h(n) \quad (4.9)$$

によって規定され、具体的なシステム構成は図 4.1 のようになる。

サブサンプリングパターンが反転している場合

上下チャネルのサブサンプリングパターンが反転している場合には、行列式 $\Delta(z)$ は

$$\Delta(z) = -H_0(z)H_1(-z) - H_0(-z)H_1(z) \quad (4.10)$$

となる。この場合には $\Delta(z) = -2 \cdot z^{-N}$ が成立すれば、 $G_0(z) = H_1(-z)$ 、 $G_1(z) = H_0(-z)$ によって完全再構成 FIR フィルタを実現できる。

この $\Delta(z)$ に対し、前節と同様に $H_0(z) = H(z)$ 、および $H_1(z) = \alpha - H(z)$ を代入すると、完全再構成条件は

$$\alpha [H(z) + H(-z)] - 2H(z)H(-z) = 2 \cdot z^{-N} \quad (4.11)$$

として与えられる。この場合には、前節とは異なり一般解を求めることは困難であるが、 $\alpha = 1$ とした場合の特殊解として次のような例が考えられる。

$$\begin{cases} H_0(z) = z^{-1} \\ H_1(z) = 1 - z^{-1} \end{cases} \quad (4.12)$$

以上の操作は、単純なサブサンプリングと前値予測の組合せとして考えることができる。

4.1.2 2次元の場合

ここではサブサンプリングパターンが一致する場合の2次元フィルタの構成方式について検討を行う。ただし、可分型 (separable) 構成が可能であることは明白であり、ここでは quincunx サブサンプリングを前提とする非可分型構成に関する検討を行う。

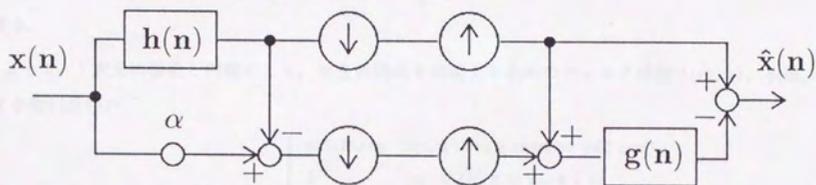


図 4.1: ハイパスフィルタを簡略化した完全再構成フィルタ

図 4.1における表記を、それぞれ単純に $x(m, n)$, $h(m, n)$, $g(m, n)$, $\hat{x}(m, n)$ によって 2 次元に拡張するものとする。一方, quincunx サブサンプリングは, 2 次元のサブサンプリング関数

$$f(m, n) = \frac{1 + (-1)^{m+n}}{2} \quad (4.13)$$

によって規定される。この場合の完全再構成条件は, 2 次元の z 変換領域において次式によって与えられる:

$$H(z_x, z_y) - H(-z_x, -z_y) = 2 \cdot z_x^{-M} z_y^{-N} \quad (4.14)$$

ただし,

$$\begin{cases} H_0(z_x, z_y) = H(z_x, z_y) \\ H_1(z_x, z_y) = \alpha - H(z_x, z_y) \\ G_0(z_x, z_y) = 1 - \frac{1}{\alpha} H(-z_x, -z_y) \\ G_1(z_x, z_y) = -\frac{1}{\alpha} H(-z_x, -z_y) \end{cases} \quad (4.15)$$

である。

よって, 1 次元の場合と同様にして, 完全再構成を実現するためのフィルタ係数 $h(m, n)$, $g(m, n)$ に対する制約条件が

$$h(m, n) = \begin{cases} \text{arbitrary} & (m, n : \text{even even or odd odd}) \\ 1 & (m = 2K, n = 2L + 1 \text{ or} \\ & m = 2K + 1, n = 2L) \\ 0 & (\text{otherwise}) \end{cases} \quad (4.16)$$

$$g(m, n) = \frac{1}{\alpha} (-1)^m (-1)^n h(m, n) \quad (4.17)$$

として与えられる。

4.1.3 具体的なフィルタ設計

構成例 1

(4.8)式に基づくもっとも簡単な例としては, $\alpha = 2$ とした場合の

$$\begin{cases} H_1(z) = 1 + z^{-1} \\ H_2(z) = 1 - z^{-1} \end{cases} \quad (4.18)$$

が考えられる。この操作は, 従来の 2×2 のブロック直交変換にはかならない。

構成例 2(理想ハーフバンドフィルタ)

階層的符号化と同様の画像信号の段階的表示への応用を考え, $H(z)$ を直線位相のローパスフィルタとして構成することを考える。

1次元の場合、フィルタ係数に対して $h(0)$ の回りの対称性条件 $h(-n) = h(n)$ を課すと (簡略化のために、(4.8)式における奇数次の唯一の非0のフィルタ係数 $h(2K+1)$ を $h(0)$ で置き換えている)、 $H(z)$ の振幅特性 $H(\omega T) (= |H(e^{j\omega T})|)$ は

$$H(\omega T) = h(0) + 2 \sum_{n=0}^{\infty} h(2n+1) \cos(2n+1)\omega T \quad (4.19)$$

によって与えられる。このとき、フィルタ係数の値とは無関係に

$$H(\omega T) + H(\omega T + \pi) = 2 \cdot h(0) \quad (4.20)$$

が成立する (自由度の劣化)。

ただし、この(4.20)式の制約の基でもさまざまなローパスフィルタの構成が可能である。1次元の場合の具体例としては、理想ハーフバンド特性

$$H'(\omega T) = \begin{cases} 1 & (|\omega T| \leq \frac{\pi}{2}) \\ 0 & (|\omega T| \geq \frac{\pi}{2}) \end{cases} \quad (4.21)$$

の逆フーリエ変換から求まる

$$h'(n) = w(n) \cdot \frac{\sin \frac{n\pi}{2}}{\frac{n\pi}{2}} \quad (4.22)$$

が考えられる ($w(n)$ は任意の窓関数)。

また、非可分型構成の場合も同様に、理想ダイヤモンド分割を実現する

$$h'(m, n) = \begin{cases} 1 & (m = n = 0) \\ 0 & (|m| = |n| \neq 0) \\ -\frac{2}{\pi^2} \cdot \frac{(-1)^m - (-1)^n}{m^2 - n^2} & (\text{otherwise}) \end{cases} \quad (4.23)$$

を具体例として挙げることができる。

一方、ハイパス信号において、入力信号のDC成分を除去するためには、

$$h(0) = 2 \sum_{n=0}^{\infty} h(2n+1) \quad (4.24)$$

を満足するように ($H(\pi) = 0$) フィルタ係数を設定した環境下において $\alpha = 2$ とすればよい。これは非可分形の場合についても同様である

4.1.4 考察・検討

本方式の特徴としては

- フィルタ数が半減できること

- 少タップのフィルタで完全再構成を実現できること
- 送信側ローパスフィルタの周波数特性を理想フィルタ（ハーフバンドフィルタ）のそれにいくらかでも近づけられること

などが挙げられる。一方、本方式の欠点としては、ローパスフィルタの周波数特性を急峻にしてもハイパスフィルタの周波数特性はなだらかなこと、ローパスフィルタとハイパスフィルタを同時に直線位相にできないこと、などが挙げられる。また本方式をデータ圧縮に用いた場合には、ピラミッド分割によってDCT(4)程度の圧縮効率を実現することは可能であるが、DCT(8)を凌ぐような高い圧縮効率の実現は困難である。

ただし、近年 Ansari らによって、本方式の改良が図られ、上記の周波数特性の問題を大幅に改善できることが明らかにされている [36]。その場合の圧縮効率はまだ明らかではないが、このような工夫によってまだ改善の余地は残されているものと考えられる。

4.2 数論変換に基づく完全再構成フィルタ

サブバンド符号化では、たたみこみに基づくフィルタ処理によって入力信号を複数の信号系列に分割し、それぞれの信号系列をサブサンプリングした後伝送を行う。これらの処理は、以下に述べるように、フーリエ直交関数系のたたみこみ特性の成立を基礎としていると考えることが可能である。

一方、たたみこみ特性が成立する直交変換は複素数体上ではフーリエ変換しかないが、対象を整数環とすれば、数論変換もまた同様の特徴を有する直交変換であることが知られている [21]。このことはすなわち、数論変換に基づくサブバンド符号化の理論展開が可能であることを示唆している。

そこで本節では、サブバンド符号化の一般化に続き、数論変換に基づくサブバンド符号化、およびその場合の完全再構成フィルタの構成方式に関する検討を行う。

4.2.1 サブバンド符号化の一般化

サブバンド符号化の一般化として、有限長の離散信号系列 $\{x(n)\} (n=0, 1, \dots, N-1)$ に対する直交関数系列 $\{T(n, k)\}$

$$\begin{cases} X(k) = \sum_{n=0}^{N-1} x(n)T(n, k) \\ x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k)T^*(n, k) \end{cases} \quad (4.25)$$

を考え、さらに $\{h(n)\}$ によるフィルタリング操作を

$$Y(k) = H(k)X(k) \quad (4.26)$$

によって定義する (図 1.3 参照)。このとき、入力 $x(n)$ と出力 $\hat{x}(n)$ の関係は

$$\begin{aligned} \hat{X}(k) = & \frac{1}{2} [G_0(k)H_0(k) + G_1(k)H_1(k)] X(k) \\ & + \frac{1}{2} \left[G_0(k) \sum_{n=0}^{N-1} (-1)^n y_0(n)T(n, k) \right. \\ & \left. + G_1(k) \sum_{n=0}^{N-1} (-1)^n y_1(n)T(n, k) \right] \end{aligned} \quad (4.27)$$

によって与えられる (ただし $y_0(n)$, $y_1(n)$ はそれぞれ、フィルタ $h_0(n)$, $h_1(n)$ の出力である)。

このとき、上式第 2 項が一般的な意味としての折り返し成分 (aliasing) に対応しており、直交関数系が

$$(-1)^n T(n, k) = T(n, f(k)) \quad (f(\cdot) : \text{任意の関数}) \quad (4.28)$$

の性質を満足するならば、(たたみこみ特性の成立とは無関係に)

$$\begin{cases} g_0(n) = (-1)^n h_1(n) \\ g_1(n) = -(-1)^n h_0(n) \end{cases} \quad (4.29)$$

とすることによって、この折り返し成分は完全にキャンセルすることが可能になる。このような特性を持つ直交関数系としては、フーリエ変換以外にも、アダマール変換や数論変換などが挙げられる。

- フーリエ直交関数系に基づく従来のサブバンド符号化は

$$\begin{cases} T(n, k) = e^{jn\pi k} \\ f(k) = k + \pi \end{cases} \quad (4.30)$$

から定義される。これは一般的に無限長の離散信号系列に対して定義されるが、DFTのような有限長を対象とする直交関数系に対してもサブバンド符号化を定義できることは明らかである。

4.2.2 数論変換に基づく完全再構成フィルタ

数論変換に基づく直交展開は

$$\begin{cases} X(k) = \sum_{n=0}^{N-1} x(n)\alpha^{nk} \pmod{M} \\ x(n) = N^{-1} \sum_{k=0}^{N-1} X(k)\alpha^{-nk} \pmod{M} \end{cases} \quad (4.31)$$

によって定義される。ただし、 $\alpha^N = 1 \pmod{M}$ 、 $\alpha^i \neq 1 (i = 1, 2, \dots, N-1)$ 、 $M = N+1$ 、 $N \cdot N^{-1} = 1 \pmod{M}$ である。

この数論変換は、整数環上において(4.26)式のフィルタリングをたみこみ操作で実現でき、さらには(4.28)式の折り返し除去条件も満足する… $(-1)^n T(n, k) = T(n, k + \frac{N}{2})$ 。また、 $\hat{x}(n) = x(n-L)$ であるときには、 $\hat{X}(k) = \alpha^{Lk} X(k)$ が成立する。よって

$$H_0(k)H_1(k + \frac{N}{2}) - H_0(k + \frac{N}{2})H_1(k) = 2 \cdot \alpha^{Lk} \pmod{M} \quad (4.32)$$

を満たす $h_0(n)$ 、 $h_1(n)$ によって、数論変換に基づく完全再構成フィルタが構成できる。

(4.32)式からフィルタ係数を求める方法としてはさまざまなアプローチが考えられるが、ここではLe Gallらの方式[47]を基礎として $N = 256$ 、 $M = 257$ 、 $\alpha = 3$ の場合について求めた完全再構成フィルタの係数の例を以下に挙げる($L = 3$)。

$$\begin{aligned} h_0(n) &= \{1, 2, 1\} \\ h_1(n) &= \{16, 32, 161, 32, 16\} \\ g_0(n) &= \{16, 225, 161, 225, 16\} \\ g_1(n) &= \{256, 2, 256\} \end{aligned} \quad (4.33)$$

4.2.3 考察・検討

本方式の特徴は、言うまでもなくそれが整数演算しか必要としないために、きわめて高速のフィルタ処理が実現できることにある。ただし、これまでの数論変換の歴史がそうであったように、フーリエ変換とは異なりその物理的なイメージがはっきりしない。このために、数学的にはすっきりとまとめることができたが、その応用が明確ではない。データ圧縮への応用はかなり厳しそうな感があり、強いて挙げれば、音声のスクランブルのような応用が考えられるであろう。

4.3 圧縮効率の最適化

各種符号化方式の圧縮効率を考えた場合、予測符号化には最適予測係数、変換符号化には KLT と、それぞれ理論的な最適解が与えられている。しかし、サブバンド符号化の最適解は未だに明らかにされておらず、フィルタ設計の混沌とした状況が続いている。

一方、本論文では、3.2 節においてサブバンド符号化の圧縮効率の理論評価尺度 UCG を与えた。そこで完全再構成フィルタの設計に当たり、この UCG を最大にするフィルタ係数を求めてやれば、圧縮効率として最適化されたフィルタバンクを与えることが可能となる。

4.3.1 問題の定式化

1 次元 2 バンドのサブバンド符号化 (図 1.3) の完全再構成条件は、 z 変換領域において、折り返しひずみの除去条件

$$G_0(z) = H_1(-z) \quad (4.34)$$

$$G_1(z) = -H_0(-z) \quad (4.35)$$

を満足する上での

$$H_0(z)G_0(z) - H_0(-z)G_0(-z) = 2 \cdot z^{-m} \quad (4.36)$$

の成立の問題、として表現される。これらの関係式を空間領域で表現すれば

$$h_1(n) = (-1)^n g_0(n) \quad (4.37)$$

$$g_1(n) = -(-1)^n h_0(n) \quad (4.38)$$

の下における

$$\sum_n h_0(n)g_0(2k-1-n) = \delta_k \quad (4.39)$$

の成立に等価である。

そこで完全再構成条件 (4.39) 式の制約の基に、UCG

$$G_{MFB} = \frac{1}{\prod_{k=0}^{K-1} \left(\frac{A_k B_k}{\alpha_k} \right)^{\alpha_k}} \quad (4.40)$$

を最大にするようなフィルタ $\{h_0(n)\}$, $\{g_0(n)\}$ の組合せを見つけられれば、それはすなわち“データ圧縮としての観点からの完全再構成フィルタの最適化”を意味している。

一方、(4.39) 式に基づく完全再構成フィルタは“biorthogonal”フィルタバンクと総称される。ただし (4.39) 式を満足するフィルタ係数の組合せは無限に存在しており、一般的にはある拘束条件を付加することによって具体的なフィルタ係数の導出を行っている。これまでに与えられている拘束条件の例としては

1. 直交性 (orthogonality)
2. 直線位相性 (linear phase)
3. 調和性 (regularity)

の3つが挙げられる。この中で“調和性”は、フィルタの z 変換を因数分解したときにそれが $(1+z^{-1})^m$ の項を含むという拘束条件であり、 m の値が大きいかほど高次の“調和条件”(regularity condition)を満足していると言われる。これは、周波数領域で見れば $H(\pi) = 0$ の成立に他ならない。

Biorthogonalの枠の中で、既存のフィルタ構成 QMF, CQF, wavelet, SSKF は、それぞれ図 4.2 のように位置付けられる。この中で、QMF は上記の3つの拘束条件をすべて満足するフィルタ構成として考えることができるが、1.3 節に述べたように、完全再構成を実現するのはタップ数が2と無限大の場合に限られている。また、直線位相性と直交性を共に満足する完全再構成フィルタは、これまでのところタップ数が2の場合しか知られていない。

本節では、これらの biorthogonal フィルタバンクの中で、特に直線位相性を有するフィルタバンクと直交性を有するフィルタバンクの最適化を試みる。上に述べた3つの拘束条件が、データ圧縮としてどのような意味を持つかは明らかにされていないが、直線位相性の成立は(画像には好ましくないと考えられている)位相ひずみの除去に役立ち、また直交性の成立は QMF, CQF, 正規直交 wavelet 等の直交系完全再構成フィルタに対して理論的な最適解を与えるのに有効である。

4.3.2 直線位相フィルタバンクの最適化

フィルタ係数(インパルス応答)の配置を対称、または反対称とすることによって、直線位相の FIR フィルタを構成できる。一般的に、タップ数が偶数の場合には、係数配置が対称のフィルタはローパスフィルタとして($H(\pi) = 0$)、反対称のフィルタはハイパスフィルタとして機能する($H(0) = 0$)。一方、タップ数が奇数の場合には、反対称のフィルタはハイパスフィルタとして機能するが($H(0) = 0$)、対称のフィルタがローパスとなるかハイパスとなるかはその係数に依存している [20]。

ただし本節では、フィルタ $\{h_0(n)\}$ はローパスフィルタとして機能することを前提とする。故に、そのタップ数に関わらず、 $\{h_0(n)\}$ の係数配置は対称になる。一方、サブバンド符号化の概念に基づき、 $\{h_0(n)\}$ がローパスフィルタであるならば $\{g_0(n)\}$ もローパスフィルタでなければならない。よって、 $\{g_0(n)\}$ の係数配置もまた対称となる。

このように、係数配置の対称性を前提として、以下の4通りが最適化の対象となる。

1. $\{h_0(n)\}$ のタップ数が奇数、 $\{g_0(n)\}$ のタップ数が奇数
2. $\{h_0(n)\}$ のタップ数が偶数、 $\{g_0(n)\}$ のタップ数が偶数

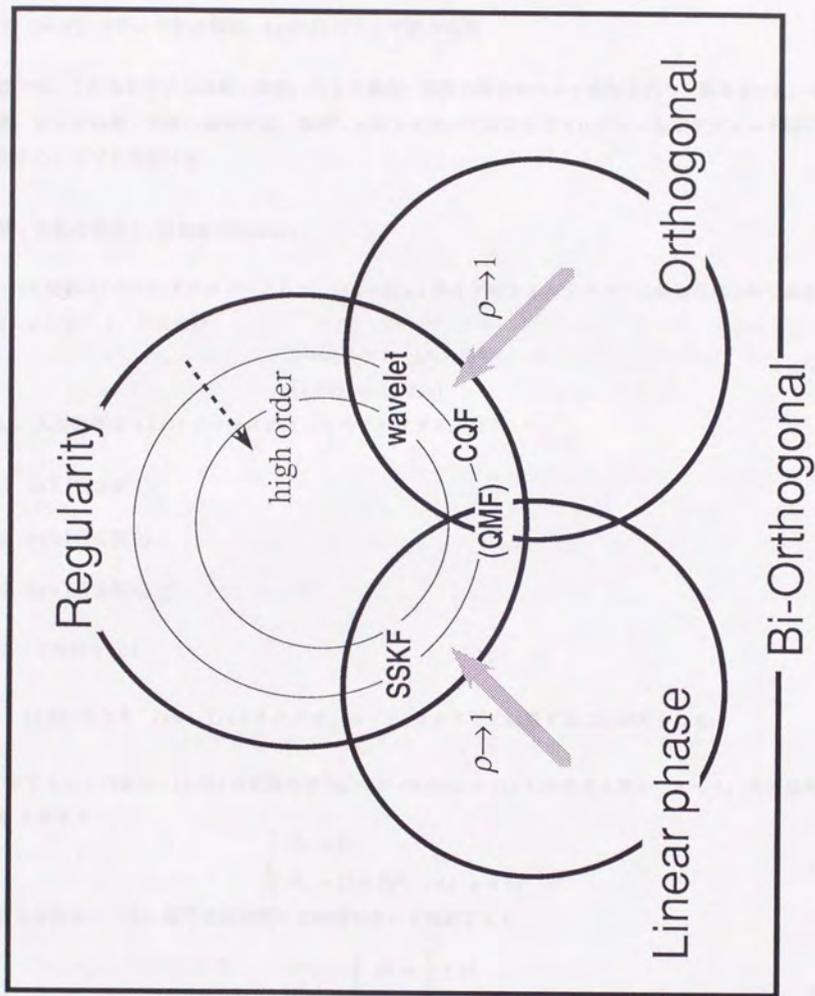


図 4.2: Biorthogonal フィルタバンク

3. $\{h_0(n)\}$ のタップ数が奇数, $\{g_0(n)\}$ のタップ数が偶数

4. $\{h_0(n)\}$ のタップ数が偶数, $\{g_0(n)\}$ のタップ数が奇数

本節では、これらの中から奇数・奇数、および偶数・偶数の場合について検討を行う。参考までに、奇数・偶数、および偶数・奇数の場合には、調査した限りにおいて良好な符号化ゲインを示すフィルタ構成を見出すことができなかった。

奇数, 奇数の場合 1 (直線位相性のみ)

SSKF(5×3) からのアナロジーとして、 $(4n-3) \times 3$ タイプのフィルタバンクの最適化について検討を行う ($n=1, 2, \dots$)。ここでは

$$\begin{aligned} \{h_0(n)\} &= \{\dots, c, b, a, b, c, \dots\} \\ \{g_0(n)\} &= \{q, 1, q\} \end{aligned} \quad (4.41)$$

とし、入力信号は AR(1) プロセスによってモデル化する。以下、

1. 1×3 (自由度 1)

2. 5×3 (自由度 2)

3. 9×3 (自由度 3)

について検討を行う。

- (4.39) 式より、 $(4n-1) \times 3$ タイプは $(4n-3) \times 3$ タイプに縮退することが示される。

まず $n=1$ の場合、(4.39) 式を満たす $h_0(n)$ は $\{h_0(n)\} = \{1\}$ しか存在しない。よって、入力信号の相関を ρ とすると、

$$\begin{cases} A_0 = 1 \\ A_1 = (1 + 2q^2) - 4q \cdot \rho + 2q^2 \cdot \rho^2 \end{cases} \quad (4.42)$$

が求められる。一方、量子化誤差間には相関がないと仮定すると、

$$\begin{cases} B_0 = \frac{1}{2} + q^2 \\ B_1 = \frac{1}{2} \end{cases} \quad (4.43)$$

が求められる。よって、予め設定された ρ に対して、UCG はパラメータ q の関数として表されることになる。そこで UCG を最大にするような q の値を探索してやれば、目的とするフィルタの最適化が実現されることになる。一例として、 $\rho = 0.95$ の場合に q の値を 0.4 から 0.6 の間に 0.002 刻みで変化させてやると、最適値として $q = 0.490$ が求まる。このとき、 $G_{1 \times 3, \max} = 5.586(\text{dB})$ となる。

次に $n = 2$ の場合、完全再構成条件は

$$\begin{cases} a + 2 \cdot q \cdot b = 1 \\ q \cdot b + c = 0 \end{cases} \quad (4.44)$$

と表される。そこで UCG のパラメータは

$$\begin{cases} A_0 = (a^2 + 2b^2 + 2c^2) + 4(ab + bc)\rho \\ \quad + 2(b^2 + 2ac)\rho^2 + 4bc \cdot \rho^3 + 2c^2 \cdot \rho^4 \\ B_1 = \frac{1}{2}a^2 + b^2 + c^2 \end{cases} \quad (4.45)$$

によって与えられる。ただし、 A_1 、 B_0 は $n = 1$ の場合とまったく同じである。一方、(4.44)式から、 (a, c) は (q, b) の関数として表されるから、UCG は 2 変数 (q, b) の関数として表されることになる。そこで $n = 1$ の場合と同様に $\rho = 0.95$ とし、 (q, b) の値を変化させることによって $G_{5 \times 3}$ の変化を調べた結果を図 4.3 に示す。この結果より、 $q = 0.51$ 、 $b = 0.29$ のときに最大値 $G_{5 \times 3, \max} = 6.307(\text{dB})$ が得られることがわかる。

なお、SSKF(5×3) は $q = 0.50$ 、 $b = 0.25$ の場合に対応しており、この場合には $G_{SSKF(5 \times 3)} = 6.277(\text{dB})$ が得られる。

一方、 (q, b) ではなく、 (a, c) を変化させた場合の符号化ゲインの推移を 3 次元表示した結果を図 4.4 に示す ($\rho = 0.95$)。この図が示すように、5×3 フィルタで優れた圧縮効率を実現するためには、 a の値にはかなりの許容範囲が見られるが、 c の値は -0.125 周辺の非常に限られた範囲に限定されることがわかる。

$n = 3$ の場合には、完全再構成条件が

$$\begin{cases} a = 1 - 2 \cdot q \cdot b \\ c = -q(b + d) \\ e = -q \cdot d \end{cases} \quad (4.46)$$

によって与えられる。この制約条件下では、 $\rho = 0.95$ に対して、 $q = 0.51$ 、 $b = 0.29$ 、 $d = -0.04$ のときに、最大値 $G_{9 \times 3, \max} = 6.322(\text{dB})$ が得られる。

以上のまとめとして、 $n = 1, 2, 3$ それぞれの場合について符号化ゲインの最大値をプロットしたのが図 4.5 である。これから

- n を大きくしても特性はほぼ飽和している
- SSKF(5×3) の特性が最適値に非常に近い

ことがわかる。

ただし、現実的な問題としてハイパスフィルタの周波数特性は入力 DC 成分を完全に除去できることが望ましい ($G_0(\pi) = 0$)。これは後述する段階的表示への応用と共に、実際の画像が統計的にはほぼ AR(1) でモデル化できるとしても、よりミクロに見れば、画像はある物体の表面上の滑らかな部分と物体の境界部のエッジとしての急峻な部分から構成されているからである。本節の最適解は直線位相性は満足するが、

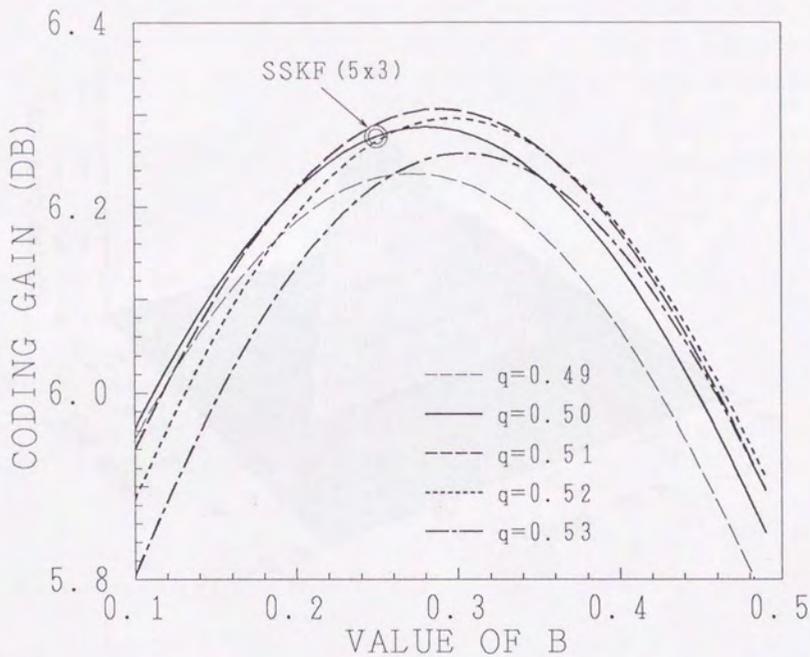


図 4.3: 5×3 直線位相フィルタバンクの最適化 (1)

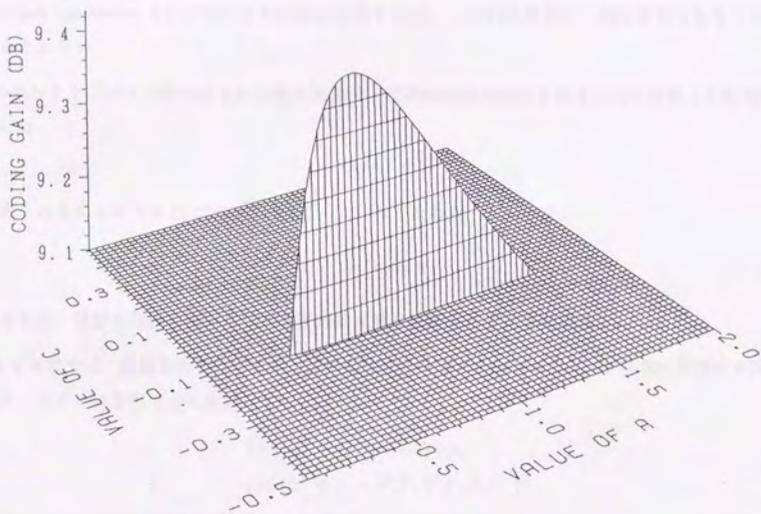


図 4.4: 5×3 直線位相フィルタバンクの最適化 (2)

ハイパスフィルタは入力 DC 成分を完全に除去するものではない。そこで次に、新たな制約条件として $G_0(\pi) = 0$ を課した場合の最適化について検討を行う。

奇数，奇数の場合 2 (直線位相性 + 1 次の調和条件)

階層的符号化のような画像の段階的表示等の応用を考えた場合，低解像度のサブサンプリング画像 (低周波成分) に補間処理を施して解像度の回復を図ることが起こりうる。この場合，入力画像の直流成分だけはそのまま直流成分として復元できることがフィルタ構成に望まれる。これは Burt らによって equivalent contribution constraint として導入された概念であり [23]，この拘束条件の一般化を図ったものが調和条件にほかならない。

この条件をサブバンド符号化として考えた場合，直線位相性の成立とはまったく独立した問題として，数式的には

$$G_0(\pi) = 0 \quad (4.47)$$

として表されることになる (1 次の調和条件)。これはすなわち

$$H_1(\pi) = 0 \quad (4.48)$$

に等価であり，前節の検討結果からリーク予測的な解が排除されることになる。

そこで本節では，直線位相特性にこの 1 次の調和条件を加えた場合のフィルタ係数の最適化に関する検討を行う。ここでは表記を少し変更して

$$\begin{aligned} \{h_0(n)\} &= \{\dots, c, b, a, b, c, \dots\} \\ \{g_0(n)\} &= \{\dots, c', b', a', b', c', \dots\} \end{aligned} \quad (4.49)$$

とし，特に $b = 0.25$ に固定する。そして調査を行ったフィルタ構成は

1. 1×3 (自由度 0)
2. 5×3 (自由度 1)
3. 9×3 (自由度 2)
4. 5×7 (自由度 2)
5. 7×5 (自由度 2)

の 5 つである。このように新たな拘束条件の付加により，フィルタ係数の自由度は前節の検討よりも 1 だけ減少することになる。

上記のフィルタごとの完全再構成条件は，

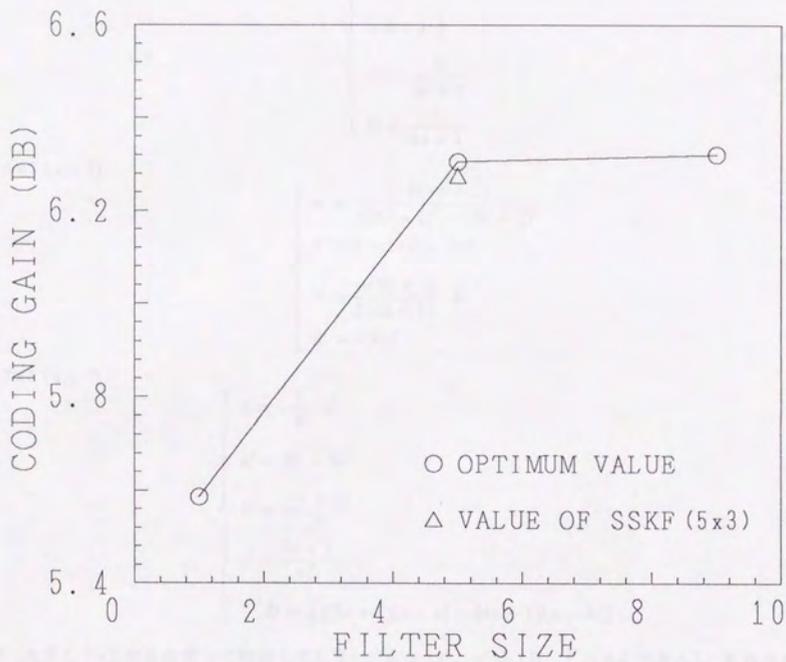


図 4.5: $(4n - 3) \times 3$ 直線位相フィルタバンクの符号化ゲインの推移

1. 5×3 (a ?)

$$\begin{cases} c = -0.125 \\ a' = \frac{4}{4a+1} \\ b' = \frac{2}{4a+1} \end{cases} \quad (4.50)$$

2. 9×3 (a, c ?)

$$\begin{cases} d = -2(c + \frac{1}{8}) \\ e = c + \frac{1}{8} \\ a' = \frac{4}{4a+1} \\ b' = \frac{2}{4a+1} \end{cases} \quad (4.51)$$

3. 5×7 (a, c ?)

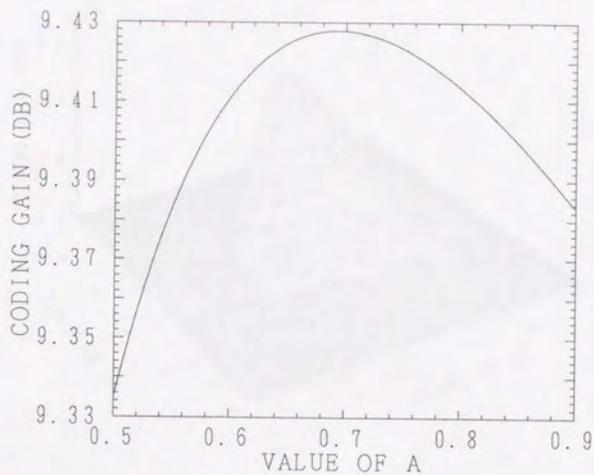
$$\begin{cases} a' = \frac{8(4a+1)}{2(4a+1)^2 - (8c+1)^2} \\ b' = 2 - 2aa' - 4cc' \\ c' = \frac{-(8c+1)}{2(4a+1)} \cdot a' \\ d' = -4cc' \end{cases} \quad (4.52)$$

4. 7×5 (a, c ?)

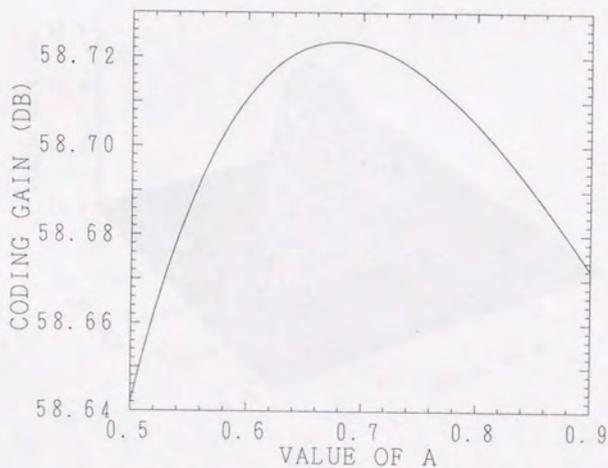
$$\begin{cases} d = -\frac{c}{b'} \cdot c' \\ a' = 2b' - 2c' \\ b' = \frac{-a+3c}{D} \\ c' = \frac{8c+1}{4D} \\ D = \frac{1}{2}[(8c+1)(c-a) - (4a+1)(a-3c)] \end{cases} \quad (4.53)$$

となる (ただし 1×3 は自由度 0 で確定してしまい, $a = 1.0$, $a' = 1.0$, $b' = 0.5$ である). これらの完全再構成条件の基でフィルタ係数の探索を行い, その符号化ゲインの推移の様子を表したのが図 4.6-図 4.9 である. ここでは 3 階層のピラミッド分割を前提としており, それぞれ $\rho = 0.95$, および $\rho \rightarrow 1$ の場合の評価結果である. また, 表 4.1 には, これらのフィルタ係数の最適解を示す. 参考までに $\rho = 0.95$ に対する符号化ゲインの理論的最適値は 10.11dB であり, CQF(16), SSKF(5×3) の 3 段ピラミッド分割, および DCT(8), DCT(16) の符号化ゲインはそれぞれ 9.47dB, 9.42dB, 8.83dB, 9.46dB となる (括弧内の数値はフィルタ長を表す).

次に図 4.10 は, 以上の検討結果を基に $\rho = 0.95$, 3 階層のピラミッド分割に対して, それぞれのフィルタバンクの符号化ゲインの最大値をプロットしたものである. ここで横軸はフィルタ $\{h_0(n)\}$ のタップ数, 縦軸はフィルタ $\{g_0(n)\}$ のタップ数を表している. そしてこの図に示すように

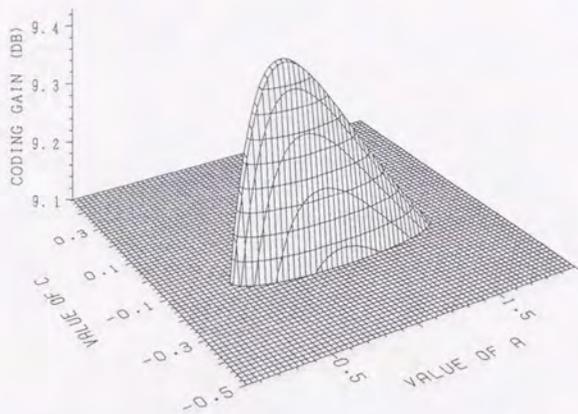


(a) $\rho = 0.95$ の場合

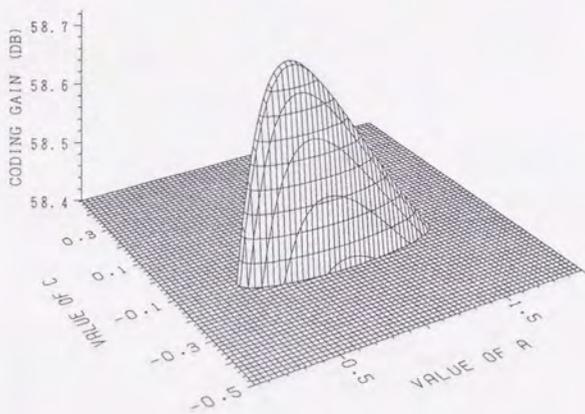


(b) $\rho \rightarrow 1$ の場合

図 4.6: 5×3 直線位相フィルタバンクの最適化

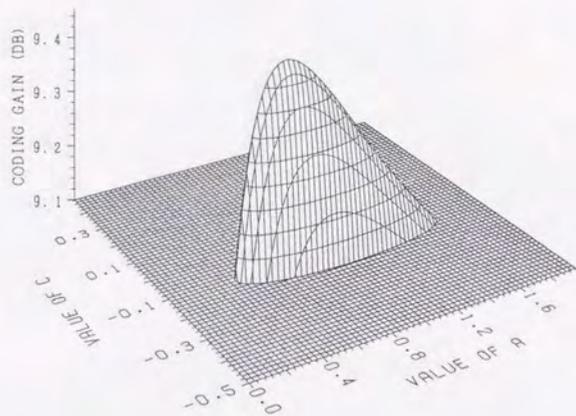


(a) $\rho = 0.95$ の場合

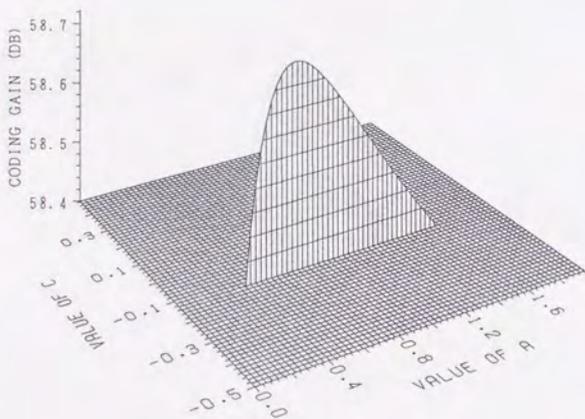


(b) $\rho \rightarrow 1$ の場合

図 4.7: 9×3 直線位相フィルタバンクの最適化

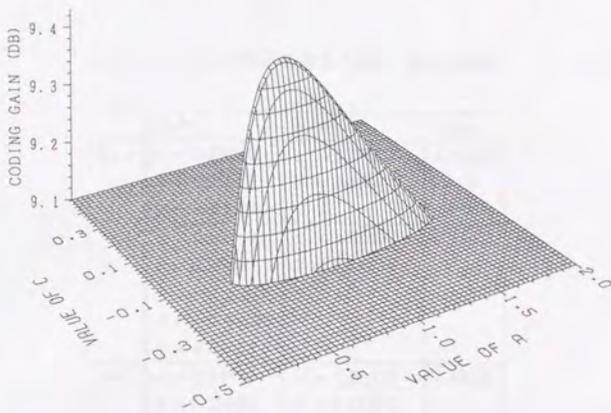


(a) $\rho = 0.95$ の場合

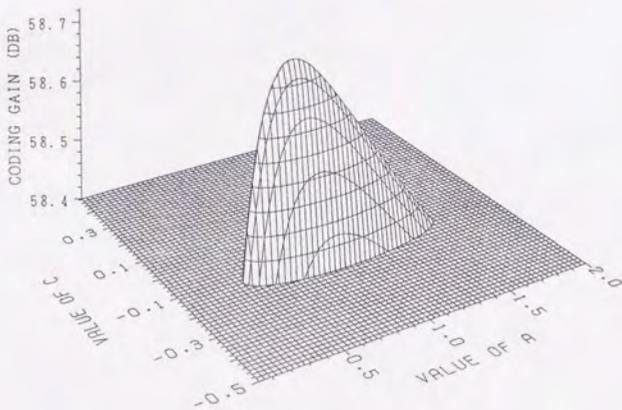


(b) $\rho \rightarrow 1$ の場合

図 4.8: 5×7 直線位相フィルタバンクの最適化



(a) $\rho = 0.95$ の場合



(b) $\rho \rightarrow 1$ の場合

図 4.9: 7×5 直線位相フィルタバンクの最適化

表 4.1: フィルタ係数の最適化 (奇数, 奇数の場合)

	$h_0(n)$	$g_0(n)$	G_{SBC}
5×3	$a = 0.69280$ $b = 0.25000$ $c = -0.12500$	$a' = 1.06067$ $b' = 0.53034$	9.43dB
9×3	$a = 0.68750$ $b = 0.25000$ $c = -0.10938$ $d = -0.03125$ $e = 0.01563$	$a' = 1.06667$ $b' = 0.53333$	9.43dB
5×7	$a = 0.68750$ $b = 0.25000$ $c = -0.10938$	$a' = 1.06726$ $b' = 0.52474$ $c' = -0.01779$ $d' = -0.00778$	9.45dB
7×5	$a = 0.68750$ $b = 0.25000$ $c = -0.10938$ $d = -0.00337$	$a' = 1.07146$ $b' = 0.51974$ $c' = -0.01599$	9.43dB

- 5×7 と 7×5 の最適解は, ρ が 1 に近づくに連れて 5×3 に漸近して行く ($c = -0.125$)

ことが明らかになった.

偶数, 偶数の場合

本節ではタップ数が偶数・偶数である場合の検討を行う. ここでは,

1. 2×2 (自由度 0)
2. 4×4 (自由度 1)
3. 6×6 (自由度 2)
4. 8×8 (自由度 3)

の 4 通りについて検討を行い, フィルタ係数を

$$\begin{aligned} \{h_0(n)\} &= \{\dots, c, b, a, a, b, c, \dots\} \\ \{g_0(n)\} &= \{\dots, c', b', a', a', b', c', \dots\} \end{aligned} \quad (4.54)$$

と表す. そして今回は, $a = 1.00$ に固定する.

奇数・奇数の場合との大きな相違点は, 偶数・偶数の場合には 1 次の調和条件

$$G_0(\pi) = 0 \quad (4.55)$$

が自動的に満足されることである. このために, 直線位相性だけを拘束条件として課してやればよい.

フィルタごとの完全再構成条件は,

1. 4×4 (b ?)

$$\begin{cases} a' = \frac{1}{2(1-b^2)} \\ b' = -\frac{b}{2(1-b^2)} \end{cases} \quad (4.56)$$

2. 6×6 (b, c ?)

$$\begin{cases} a' = -\frac{b(b-c)}{2D} \\ b' = \frac{b^2(b+c)}{2D} \\ c' = -\frac{bc(b+c)}{2D} \\ D = b(b-c)(b+c+1)(b+c-1) \end{cases} \quad (4.57)$$

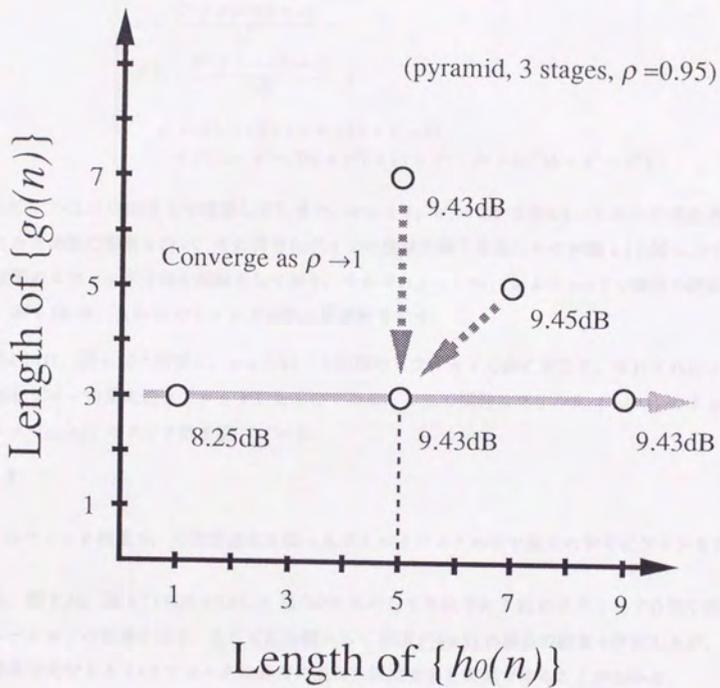


図 4.10: フィルタサイズと符号化ゲイン (奇数, 奇数の場合)

3. 8×8 (b, c, d ?)

$$\left\{ \begin{array}{l} a' = \frac{d^2 + (1 - 2bc)d - c(b - c)}{2D} \\ b' = \frac{bd^2 - (b + 2c)d + bc(b + c)}{2D} \\ c' = \frac{d^2 + d - c(b + c)}{2D} \cdot c \\ d' = -\frac{d^2 + d - c(b + c)}{2D} \cdot d \\ \\ D = c(b - c)(b + c + 1)(b + c - 1) \\ \quad + d\{(1 - b^2 - 4bc + c^2) + (1 + b^2 + bc + 2c^2)d - d^2 - d^3\} \end{array} \right. \quad (4.58)$$

となる (ただし 2×2 は自由度 0 で確定してしまい, $a = 1.0$, $a' = 0.5$ である). これらの完全再構成条件の基でフィルタ係数の探索を行い, その符号化ゲインの推移の様子を表したのが図 4.11-図 4.12 である. ここでは 3 階層のピラミッド分割を前提としており, それぞれ $\rho = 0.95$, および $\rho \rightarrow 1$ の場合の評価結果である. また, 表 4.1 には, これらのフィルタ係数の最適解を示す.

次に図 4.13 は, 図 4.10 と同様, $\rho = 0.95$, 3 階層のピラミッド分割に対して, それぞれのフィルタバンクの符号化ゲインの最大値をプロットしたものである. ここで横軸はフィルタ $\{h_0(n)\}$ のタップ数, 縦軸はフィルタ $\{g_0(n)\}$ のタップ数を表している.

このとき

- 8×8 のフィルタ構成が, 今回最適化を図ったフィルタバンクの中で最大の符号化ゲインを実現する

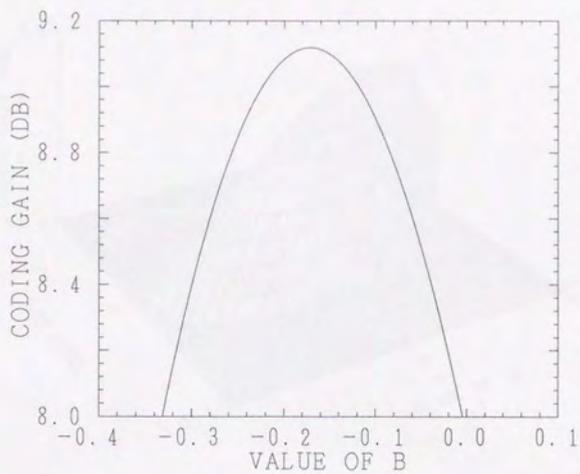
ものであり, 図 4.14, 図 4.15 には GIRL と MOON に対してそれぞれ 3 段のピラミッド分割を図った場合のシミュレーションの結果を示す. ここでは比較として SSKF(5×3) の場合の結果も併記したが, これらの図より, 提案方式である 8×8 フィルタは非常に優れた圧縮効率を実現できることがわかる.

一方, 図 3.15, 図 3.16 の結果と比較した場合, 8×8 フィルタは CQF(16) に対しても優れた特性を示していることが解る. これは理論とは矛盾する結果であるが, 今回のシミュレーションでは, 画像の境界部の処理に当たり, 8×8 フィルタには対称拡張を適用している. しかし, CQF は非対称フィルタであるためにそれが適用できず, 今回のシミュレーションのような結果が得られたことになる.

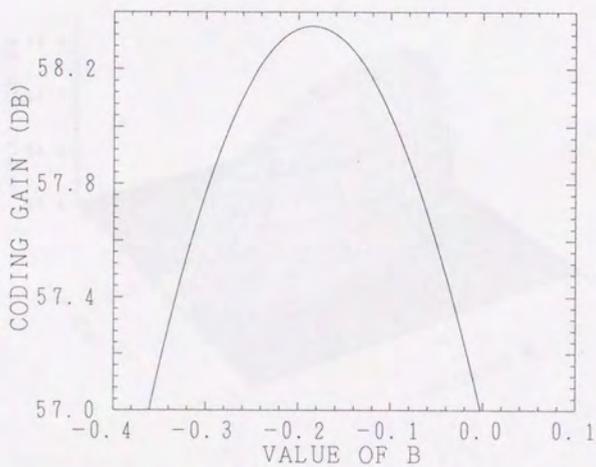
4.3.3 直交フィルタバンクの最適化

(4.34)-(4.36) 式に対して, さらに制約条件として

$$H_1(z) = -H_0(-z^{-1})z^{-L} \quad (4.59)$$

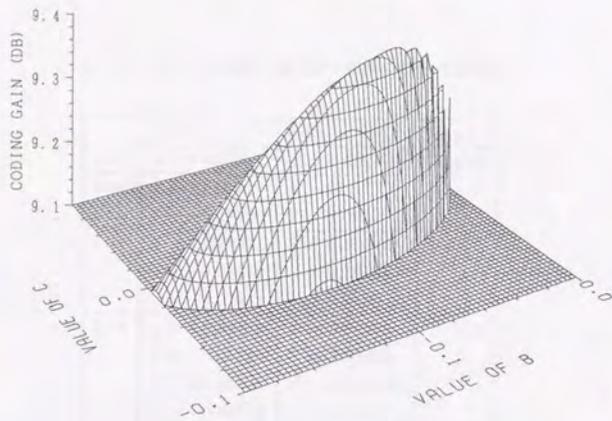


(a) $\rho = 0.95$ の場合

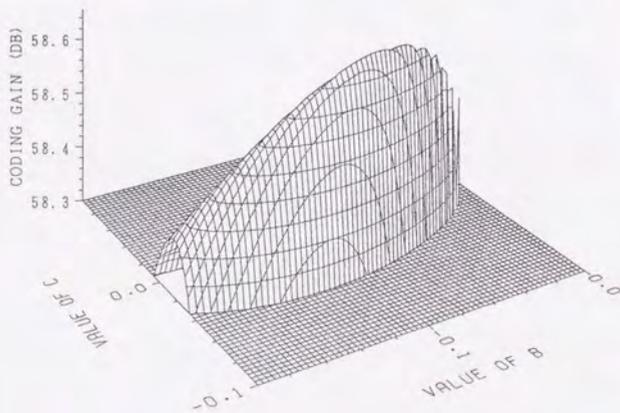


(b) $\rho \rightarrow 1$ の場合

図 4.11: 4×4 直線位相フィルタバンクの最適化



(a) $\rho = 0.95$ の場合



(b) $\rho \rightarrow 1$ の場合

図 4.12: 6×6 直線位相フィルタバンクの最適化

表 4.2: フィルタ係数の最適化 (偶数, 偶数の場合)

	$h_0(n)$	$g_0(n)$	G_{SBC}
2×2	$a = 1.00000$	$a' = 0.50000$	7.94dB
4×4	$a = 1.00000$ $b = -0.18500$	$a' = 0.51772$ $b' = 0.09578$	9.12dB
6×6	$a = 1.00000$ $b = -0.08437$ $c = -0.03438$	$a' = 0.50715$ $b' = 0.10163$ $c' = -0.04140$	9.40dB
8×8	$a = 1.00000$ $b = -0.05125$ $c = -0.05500$ $d = -0.01625$	$a' = 0.50289$ $b' = 0.09403$ $c' = -0.03849$ $d' = 0.01137$	9.45dB

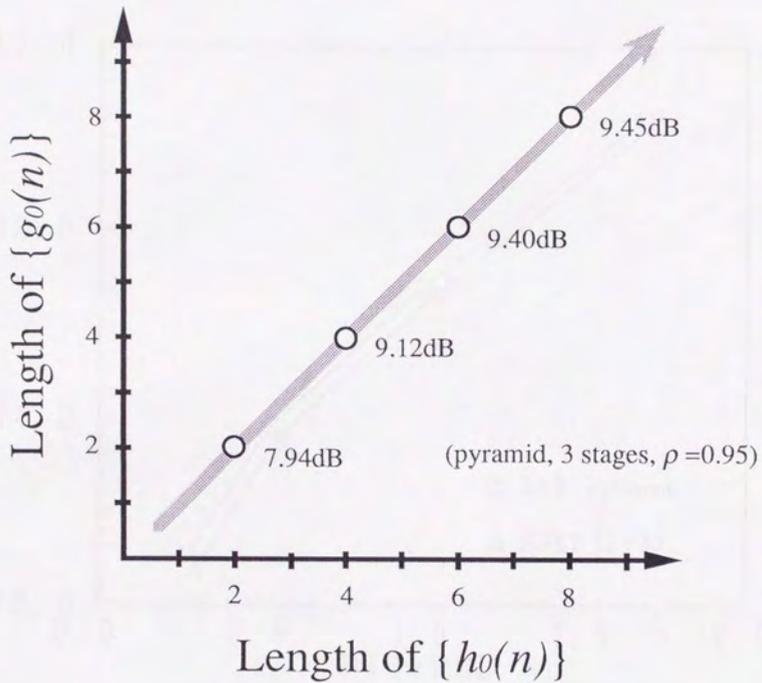


図 4.13: フィルタサイズと符号化ゲイン (偶数, 偶数の場合)

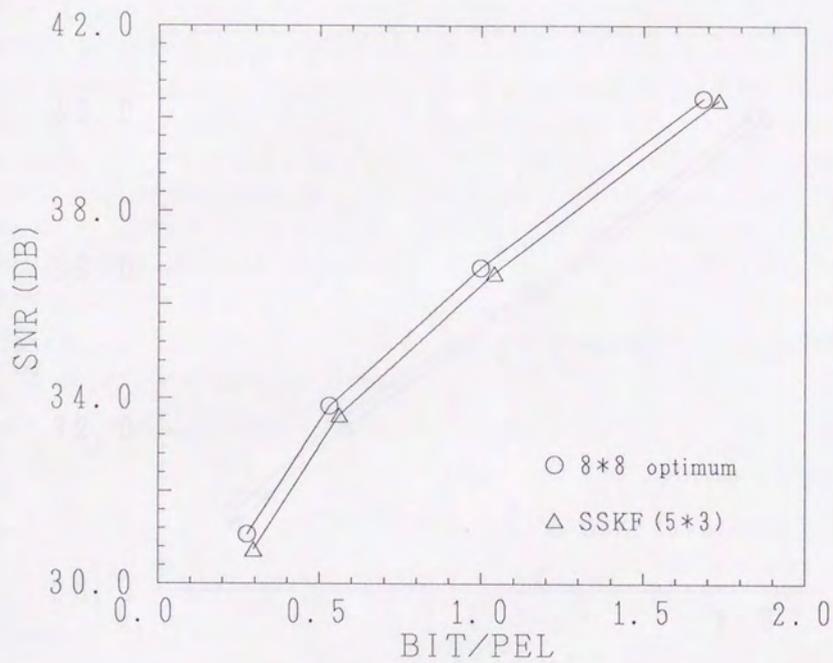


図 4.14: シミュレーション結果 (GIRL)

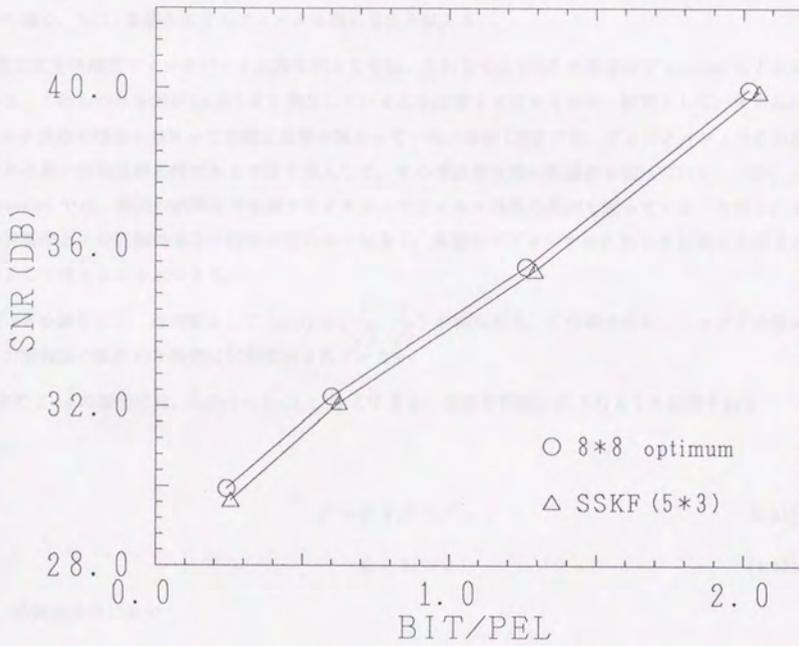


図 4.15: シミュレーション結果 (MOON)

を課すと(ただし L はフィルタ長を表す), 完全再構成フィルタが直交性を持つようになる. この場合, (4.36) 式は空間領域において

$$\sum_n h_0(n)h_0(n+2k) = \delta_k \quad (4.60)$$

と書き換えられ, $\{h_0(n)\}$ から残りのフィルタ $\{h_1(n)\}$, $\{g_0(n)\}$, $\{g_1(n)\}$ が一意に決定されることになる. ただし, (4.60) 式を満足する $\{h_0(n)\}$ が無数に存在することは明らかである. また, フィルタ長 L は偶数に限られ, $L=2$ の場合を除いて直線位相特性を実現することはできない. 本節では, (4.60) 式の拘束条件の基で, UCG を最大にするフィルタ係数の導出を試みる.

直交完全再構成フィルタバンクの具体例としては, これまでに CQF や正規直交 wavelet などが知られている. これらの具体例が (4.60) 式を満足していることは言うまでもないが, 結果として, その具体的なフィルタ係数の導出に当たっての設計基準が異なっていた. まず CQF では, デジタルフィルタの設計理論である最小位相分解と呼ばれる手法を導入して, その周波数特性の最適化を図っている. 一方, 正規直交 wavelet では, 前述の調和条件を課すことによってフィルタ係数の導出を図っている. ただし, これらの設計基準とその圧縮効率との関係は明らかではなく, 本節のアプローチはこれらとは異なる第 3 の設計基準として考えることができる.

$L=2$ の場合には, 最適解として $h_0(n) = \{\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\}$ が得られる. この場合には, フィルタ係数に対して入力情報源の統計的な特徴は何等反映されていない.

次に $L=4$ の場合には, $h_0(n) = \{a, b, c, d\}$ とすると, 最適化問題は以下のように説明される.

$$a^2 + b^2 + c^2 + d^2 = 1 \quad (4.61)$$

$$ac + bd = 0 \quad (4.62)$$

の拘束条件の基で

$$(ab + bc + cd) + ad \cdot \rho^2 \quad (4.63)$$

を最大にする $\{a, b, c, d\}$ の組合せを見出すこと

上式に対してさらに, 制約条件として

$$a + c = b + d \quad (4.64)$$

を課すと(1 次の調和条件), 最適解として

$$a = \frac{\sqrt{2} + \sqrt{6}}{8}, \quad b = \frac{3\sqrt{2} + \sqrt{6}}{8}, \quad c = \frac{3\sqrt{2} - \sqrt{6}}{8}, \quad d = \frac{\sqrt{2} - \sqrt{6}}{8} \quad (4.65)$$

が得られる. この場合にも入力情報源の統計的な特徴は反映されていないが, これらの係数は $L=4$ の場合の正規直交 wavelet に一致する.

4.3.4 考察・検討

本節では、前節で求めたサブバンド符号化の圧縮効率の理論的評価尺度 UCG と、1.3 節に示した完全再構成条件とを組み合わせることにより、完全再構成フィルタの圧縮効率の最適化を図った。本節で求めた具体例は、与えられたフィルタ長の制約の基において最適な圧縮効率を実現するものであり、SSKF(5×3)以外にも、タップ数の短い非直交系のフィルタバンクとして、従来の直交系を凌ぐ圧縮効率を実現できるものが存在することを示している。あるいはまた、biorthogonal の枠の中で非常に自由な提案の行われてきた完全再構成フィルタに対して、理論的な最適解を与える結果となるものと考えられる。

ただし、今回のフィルタ係数の導出には、計算機に頼った全探索的な手法を用いている。これではフィルタ長が長くなった場合に問題が非常に複雑になってしまい、問題の定式化としては明確であっても、直交系の場合のようにその具体例の導出が非常に困難である。そこで、今後の課題としては、何等かの洗練された最適化アルゴリズムを用いた、より系統的な理論解析手法の構築が必要となってくるものと考えられる。

Chapter 5

サブバンド符号化の ATM 用画像符号化への 応用

広帯域 ISDN の情報転送方式である ATM 環境下の画像通信である“映像パケット伝送”では、セル廃棄対策やユニバーサル符号化などの問題が、従来にはない新たな検討課題として持ち上がってくる。一方、サブバンド符号化は、その周波数領域における階層構造の利用によって、これらの諸問題を解決する符号化方式として現在注目を集めている。そこで本章では、このサブバンド符号化の ATM 用画像符号化への応用を目的として、定性的、定量的な観点からその有効性を明らかにする。

5.1 ATM 用画像符号化について

ISDN から広帯域 ISDN への移行にあたり、その情報伝送方式として ATM(Asynchronous Transfer Mode) が採用されていることは周知の通りである。ここでは、従来の 1 対 1 の通信ばかりではなく、1 対 N(放送型)、N 対 1(情報収集型)、ならびに N 対 N(会議型)の多様な通信接続形態が実現され、また統計的に独立な多数の情報源から発生するバースト情報を網内で多重化することにより、網資源の有効利用が図られる。ユーザにとっては、ネットワークに対するレートフリーなアクセスが可能となり、これまでの回線交換をベースとした ISDN において課せられていた多くの伝送上の制約から解放されることになる。その ATM の恩恵を被る通信サービスの典型的な例が、画像、あるいは映像の伝送である [73,74]。

動画をデジタル伝送する場合、その伝送レートは数十 Mb/s から数 Gb/s にも及ぶ膨大なものとなる。しかし、一般的には画像情報の冗長度は大きく、限られた網資源の有効利用、あるいは LSI 技術の発展に伴う端末コストの低減、などの観点から見ても、動画通信に対するデータ圧縮技法の導入は必須である。

一方、動画の高効率符号化方式は、空間方向の相関を利用するフレーム内符号化と、時間軸方向の相関を利用するフレーム間符号化とに大別されるが、一般的には後者の方がより高い圧縮効率を実現できる。また、このような高効率符号化を施した場合、その発生情報量は動きの程度、シーンチェンジなどによって時間的に変動することになるが、フレーム間符号化の場合にはこの変動がより顕著となる。

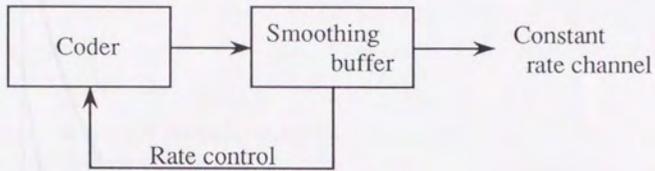
従来の回線交換網ではピークレートよりも小さなレートを割り当てることによって伝送路の使用効率の向上を図っていたが、このためには固定レート制御として量子化ステップサイズの変更、コマ落とし、などの操作を施す必要が生じる。故に、画質の時間的な変動は、不可避なものとなっていた。画質評価実験の結果としても、部分的、瞬間的な画質劣化によって総合評価が決定されるとする報告もあり、この画質変動の問題は動画通信における大きな問題となっていた。

しかし、通信網が伝送レートの時間変動に対応できれば、もはや上記の問題は解決され、可変レート符号化の採用による画質の安定した動画通信が実現される。ATM は、まさにこの要求を満足する情報伝送方式である。

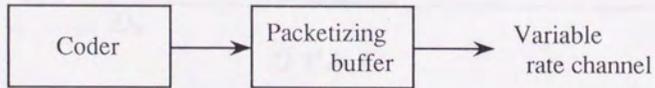
ATM における動画通信、すなわち映像パケット伝送は、以上のような背景の基に、通信網の研究者と画像符号化の研究者の双方を巻き込んだ形で、活発な検討が行われているテーマである。すでに、IEEE COMSOC 主催の基に 4 回の国際ワークショップが開催されており [89,90,91]、1993 年には第 5 回のワークショップが開催される予定である。

図 5.1 は、固定レートの伝送路を対象にした従来の画像伝送の送信側 (a) と、レートフリー(可変レート)の伝送路を対象にした画像伝送の送信側 (b) における符号器の、基本的な構成方式の相違を示している。

一般的に、高効率符号化によって発生する瞬時情報量は個々の画像や画像の局所的な特徴に応じて大きく変化するものであり、その出力は時間的に大きく変動することになる。このために、固定レートチャネル

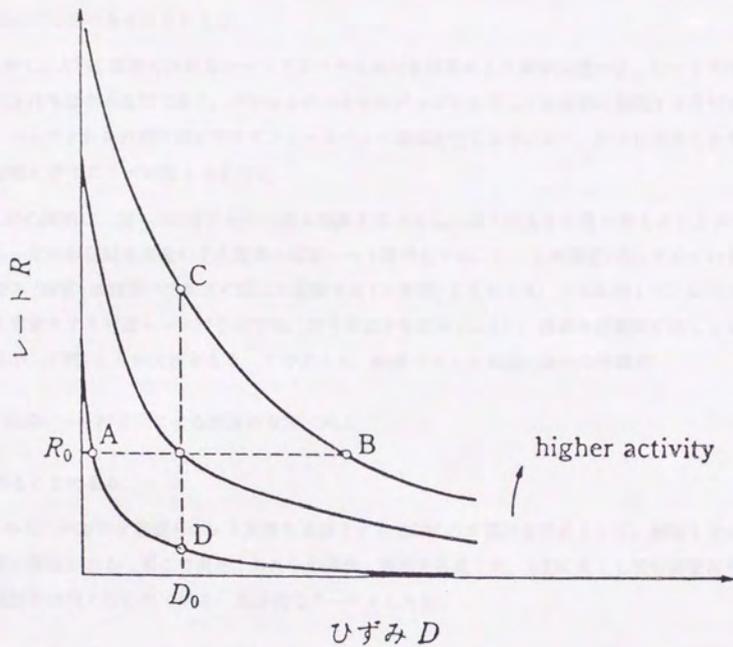


(a) Constant rate coding



(b) Variable rate coding

図 5.1: 画像符号化方式の比較



M.Maglaris et al., "Performance Analysis of Statistical Multiplexing for Packet Video Sources," IEEE Trans. on Commun., July, 1988.

図 5.2: レートひずみ関数

を対象にした画像伝送では、符号化出力をバッファメモリで平滑化した後にチャンネルに送出する必要がある。ただし、バッファの最大容量はハードウェアコストや許容遅延時間によって制限され、伝送路の利用効率の問題もある。そこで、バッファから符号器に対するフィードバック制御(固定レート制御)が行われ、オーバーフローやアンダフローの発生が規制される。すなわち、固定レート環境では符号化品質の時間的な変動が不可避であることになる。

しかし、ATM 環境におけるレートフリーチャンネルを対象にした画像伝送では、レート平滑化のためのバッファはもはや不必要であり、パケット化のためのバッファを介して伝送路に送出するだけでよい。すなわち、バッファからの符号器に対するフィードバック制御を行う必要はなく、符号化品質をあらかじめ定めたい定値に保つことが可能となる [75]。

上述の説明は、図 5.2 に示すレート歪み関数を用いると、以下のように言い替えることができる [76]。レート一定の伝送路を対象にする従来の固定レート符号化では、レートが固定 (R_0) されているために、符号化歪み(画質)は画像の活動度に応じて変動する(A-B間)ことになる。これに対して、レートフリーな伝送路を対象にする可変レート符号化では、符号化歪みを固定 (D_0) し、画面の活動度に応じてレートを変化させる(C-D間)ことが可能となる。このように、映像パケット伝送の最大の特徴は

- 可変レート符号化による画質の安定、向上

が図れることにある。

しかし、一方では高速パケット交換を基盤とする ATM の本質的な欠点として、輻輳などに起因するセル廃棄の問題がある。そこで現在、これらの長所、短所を考慮した、ATM 用としての画像符号化方式に関する検討が活発に行われている。具体的なテーマとしては、

- セル廃棄に強い画像符号化方式 [77]-[81]
- 多様な画像サービスに共通した画像符号化方式 (Universal coding) [82]
- CCITT H.261 との互換性を有する画像符号化方式
- 視覚特性を反映した画質一定符号化方式 [83]

などが挙げられる。

5.2 サブバンド符号化の有効性

5.2.1 高能率符号化として

ATM 環境における画像符号化の最大の利点は、可変レート符号化の導入による画質の安定性の実現、そしてその結果としての主観的な画質の向上効果にある。しかし、前述したセル廃棄などの問題のために、従来の固定レートを対象とした符号化方式 (CCITT H.261 等) から固定レート制御を外すだけでは十分な符号器構成が図れない。

ただし、ATM 用であるとはいえ本質的には高能率符号化であり、セル廃棄対策などの工夫を凝らすことによって圧縮効率の大幅な低減を招いてはならない。すなわち、特殊な細工によって圧縮効率が大きく低減するのであれば、新たに ATM 用画像符号化について検討を行う必要はなく、周期的リフレッシュを伴う従来の固定レート方式で十分であるということになる。

結局、ATM 用動画像符号化は、セル廃棄などの問題を解決するだけでなく、従来の固定レート方式に同等の圧縮効率を実現するものでなければならない。

一方、本論文第3章で明らかにしたように、サブバンド符号化は DCT と同等、あるいはそれを凌ぐ圧縮効率を実現できる。この意味においてサブバンド符号化は、ATM 用画像符号化方式の有力な候補の一つとして挙げられることになる。

5.2.2 セル廃棄対策として

動画像の高能率符号化として、フレーム間符号化は、一般的にフレーム内符号化よりも優れた圧縮効率を実現できる。しかし、それが時間軸方向の相関を利用した方式であるために、セル廃棄の影響が後続フレームにまで伝搬し、結果として大きな画質劣化が生じることになる。このために、セル廃棄に耐性を有するフレーム間符号化方式の構築が大きな検討課題となる。

これまでに報告されているフレーム間符号化に対するセル廃棄対策の例としては、

- 階層符号化 [77]-[80],[94]
- セルのインターリーブ構成 [81]

などが挙げられる。このとき、前者は情報源符号化的なアプローチ、後者は通信路符号化的なアプローチであり、互いに相入れない概念ではない。そこで、両者の混合によって廃棄対策を実現することは可能である。

階層符号化を用いたセル廃棄対策の基本的な概念は、“伝送情報を画質に大きな影響を及ぼすもの (一般的に低周波成分) と及ぼさないもの (一般的に高周波成分) とに分離し、画質への影響の大きな情報につい

ては廃棄に対する特殊な保護(プライオリティ)を施す”というものである。その具体例としては、それぞれ周波数領域における符号化として考えられる

- 直交変換(DCT, LOT等)
- 階層的符号化
- サブバンド符号化

などを用いた方式が報告されている。

DCTに基づく階層符号化方式の分類

ここでは特に、DCTに基づく階層符号化方式について考えてみる。これまでの報告例を見ると、これらは図5.3のように分類することが可能である、

まず図5.3(a)は、最初に所望の画質を保証する符号化(保証情報)を行った後にさらにその復号画像と原画像との差分情報(強調情報)を伝送する方式であり、“再量子化”(re-quantization)と呼ばれている[77]。この場合、保証情報が廃棄保護の対象となり、強調情報が廃棄されたとしても保証情報分の画質は保たれることになる。あるいはまた、強調情報は高周波成分として考えることができるが、そのフレーム間の相関が低いことは直感的にも明らかであり、その符号化には廃棄の影響が伝搬するフレーム間符号化を用いずにフレーム内符号化を用いても、伝送情報量が大幅に増加することはない。前述のように、フレーム内符号化であればセル廃棄の影響は1フレームだけに限定され、それ以降のフレームに廃棄の影響が残ることはない。

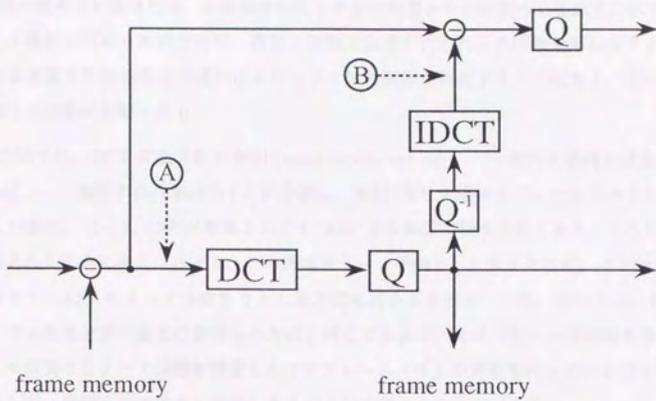
一方、図5.3(b)は、廃棄発生後の画質の回復を主な目的として“リーク予測”(leaky prediction)を導入する方式である。このリーク予測の概念を簡単に説明すれば、以下のようになる：

リーク予測の場合、フレーム間差分 $d(n)$ (n はフレーム番号) は

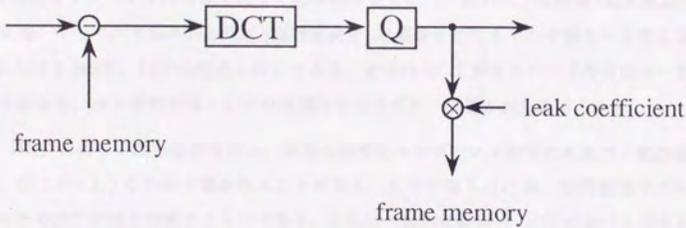
$$d(n) = x(n) - a \cdot x(n-1) \quad (5.1)$$

として求められる。この式において a がリーク係数を表しており、 $a = 1$ の場合がフレーム間符号化、 $a = 0$ の場合がフレーム内符号化に対応することになる。ここで a の絶対値の値が0に近ければ(フレーム内符号化に近ければ)、廃棄発生後の画質の回復速度が速くなる。

このリーク予測をDCT等の直交変換に組み合わせて使う場合には、フレーム間差分信号の変換係数に対してリーク係数を乗ずることになる。この場合、リーク係数の値が変換係数ごとに異なってもよい。そしてそのリーク係数で重みづけられた変換係数を用いて、フレームメモリの更新が行われる。



(a) Re-quantization



(b) Leaky prediction

図 5.3: DCT に基づく階層符号化方式の分類

リーク係数の値が1に近ければ、圧縮効率は向上するが廃棄からの画質の回復速度は低下する。これに対して、リーク係数の値が0に近ければ、画質の回復は急速に行われるが圧縮効率は低下する。このために、画質の回復速度と圧縮効率との間にはトレードオフの関係が存在することになり、リーク係数の設定に際しては細心の注意が必要である。

岸野らの方式では、DCT変換係数をMSP(most significant part…一般的に低周波成分)とLSP(least significant part…一般的に高周波成分)とに分類し、MSPだけを用いてフレームメモリの更新を行っている[78]。この場合、たとえLSPが廃棄されてもMSP分の画質は保証されており、これによってセル廃棄対策が実現されることになる。ただし、この方式をリーク予測として見てやれば、それはすなわちMSPのリーク係数を1、LSPのリーク係数を0とした方式にほかならない。一方、野村らは、MSP、LSPの分類に基づくセル廃棄対策の概念は岸野らの方式と同じであるが、そのフレーム間相関を考慮して、DCT変換係数ごとに相異なるリーク係数を設定した上でフレームメモリの更新を行っている[79]。この方式は、リーク予測として、岸野らの方式を一般化したものとして考えることができる。

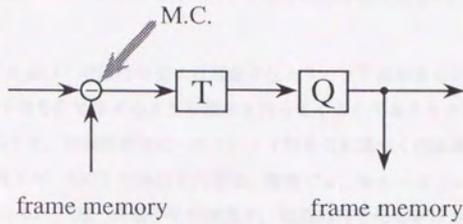
もっとも、再量子化とリーク予測を併用することは可能である。再量子化では強調情報の符号化にフレーム内符号化を仮定することが多いが、これは単にリーク係数を0とした場合に対応しており、より一般的にリーク予測を導入することによって圧縮効率の改善を図ることが可能である。

階層的符号化・サブバンド符号化に基づく階層符号化

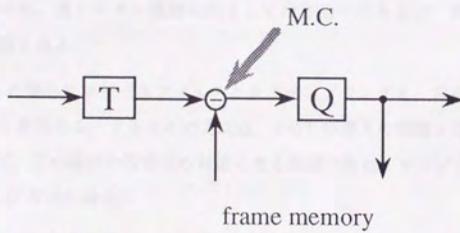
階層的符号化やサブバンド符号化に基づく階層符号化では、一般的に上位階層(低周波成分)が廃棄保護の対象となる。そして、そのほかの階層(高周波成分)は廃棄されてもやむを得ないと考える。これは上記のDCTにおけるMSP、LSPの概念と同じであり、従来のDCTがサブバンド符号化の一形態として考えられることから、セル廃棄対策としての階層符号化方式の一貫性が見られる[80,94]。

ただし、DCTに基づく動画像符号化と、階層的符号化・サブバンド符号化に基づく動画像符号化の表記方法には、図5.4のような区別が置かれることがある。ここで図5.4(a)は、空間領域でフレーム間差分を取った後にその差分情報を変換するものであり、これは一般的に従来のDCTに基づく方式に対応している。一方、図5.4(b)は、空間領域で変換(周波数分割)を行った後にその変換成分ごとにフレーム間差分を取る方式であり、これは一般的に階層的符号化・サブバンド符号化に基づく方式に対応している。もっとも、これらは直感的な分類であり、数式的には両者が等価であることは自明であろう。すなわち、前者をサブバンド符号化で実現することも、後者をDCTで実現することも可能である。

ただし、図5.4は、動き補償の導入形態を考えるのに非常に都合のよい表現方法となる。すなわち、図5.4(a)は原画像の動きベクトルを求める方式に対応しており、図5.4(b)は変換成分ごとの動きベクトルを求める方式に対応している。DCTに対する動き補償の導入形態とえば、図5.4(a)ばかりが検討対象であったことは言うまでもない。8×8のDCTと言えども、サブバンド符号化としては64バンド分割にもな



(a) Interframe difference in space domain



(b) Interframe difference in transform domain

図 5.4: 動画像符号化の表記法・動き補償の導入形態

るのであるから、DCT に対しては今後も図 5.4(a) が支配的であろう。しかし、サブバンド符号化では、図 5.4(a) も図 5.4(b) も共に十分実用的な導入形態となり得る。あるいはまた、サブバンド符号化では、低周波成分に対して求めた動きベクトルをそのほかの高周波成分に対して利用することも十分考えられる。このように、階層的符号化・サブバンド符号化に対する動き補償の導入方法は、一つの大きな研究テーマとなる。

一方、すでに DCT に基づく階層符号化には再量子化とリーク予測があることを示したが、これらは階層的符号化・サブバンド符号化ではどのような意味を持ってくるのであろうか？ まずリーク予測の概念は、図 5.4 からも明らかなように、階層的符号化・サブバンド符号化に基づく動画像符号化に対してもそのまま適用可能である。このために、DCT の場合と同様に、階層ごと、あるいはバンドごとの適切なリーク係数の決定が一つの課題となる。一方、再量子化の概念は、階層的符号化における階層間差分を求める操作に類似していることは直感的にも明らかである。そこで次に、再量子化と階層的符号化・サブバンド符号化との関係について検討を行う。

再量子化・階層的符号化・サブバンド符号化の対比

DCT に基づくセル廃棄対策の中で、特に再量子化について考えてみる。この方式の符号化の対象となる画素の数は、そこにサブサンプリング操作が何等含まれていないことから、原信号の 2 倍になることは明らかである。このために、確かにセル廃棄対策としては有効ではあるが、データ圧縮としての効果はあまり望めないことが予想される。

次に、図 5.3 中の A の部分にローパスフィルタとサブサンプリングを、B の部分に 0 値補間とローパスフィルタを挿入することを考える。するとこの方式は、DCT の導入の問題とは無関係に、階層的符号化方式に等価になる。よって、この場合の符号化の対象となる画素の数は、サンプリングの効果によって原信号の 5/4 倍に抑えられる (2 次元の場合)。

一方、階層的符号化に対するサブバンド符号化の利点として、しばしば後者では符号化の対象となる画素の数が原画像のそれに一致することが指摘されている。よって、以上の検討をまとめると、セル廃棄対策として再量子化、階層的符号化、サブバンド符号化の 3 種類の方式を考える場合、その符号化の対象となる画素数の問題から直感的に

1. サブバンド符号化
2. 階層的符号化
3. 再量子化

の順に良好なデータ圧縮特性が実現されるものと考えられる。

5.2.3 ユニバーサル符号化として

一口に画像サービスと言っても、テレビ電話、テレビ会議から HDTV、さらには将来の超高精細画像と、その画質要求、伝送レートの大幅に異なる多種多様なものが存在する。このために、サービスに限定されない画像符号化方式(ユニバーサル符号化)の必要性が各所で論じられている [82]。このユニバーサル符号化を実現するためには、各映像サービス間のフォーマット変換を規定しなければならないが、サブバンド符号化が必然的に有するサンプリング構造が、そのままこのフォーマット変換にも応用できる可能性がある。

空間軸方向では、これはすなわち 2:1 の解像度変換としての応用となる。たとえば原画像が 1024×1024 のサイズであれば、サブバンド符号化による解像度変換によって順に 512×512 , 256×256 , 128×128 のサイズの画像が得られることになる。一方、空間-時間軸方向では、quincunx サンプリングの応用として、ノンインターレース-インターレース変換が可能となる [50,53]。

5.2.4 画質一定可変レート符号化として

ATM における画像伝送の最大の利点は、可変レート符号化による“画質”の安定にあるとされる。そこで、画質が一定になるような可変レート符号化器の設計が一つのテーマとなる [83]。その一方では、画質の客観的評価尺度として広く用いられている SNR と主観評価結果が必ずしも一致しない場合があることから、工学的により汎用性のある画質評価尺度の構築が古くから指摘されている。このために、SNR を設計基準とする可変レート符号器では、必ずしも最適とはいえないことになる。

SNR に変わる画質の客観的評価尺度としてしばしば用いられるのが、視覚の空間周波数特性を反映した WSNR(weighted SNR)である [84,85]。この他にも、いくつかの視覚特性を考慮した評価尺度の提案が行われている [86]-[88]。そこでこれらの評価尺度を制御対象として、可変レート符号器を設計することが考えられる。

ただし、制御対象があまり複雑すぎても、SNR によってかなりの画質の反映が図られることを考慮すると、実用的には問題が大きいのと言わざるを得ない。そこで視覚の空間周波数特性を線形システムとして近似した場合の WSNR が一つの妥協案となる。ここで、たとえば直交変換を前提とした場合に、以下のような最適化問題を考える。

$$\sigma_r^2 = \frac{1}{K} \sum_{k=0}^{K-1} w_k \sigma_{q_k}^2 = \text{const.} \quad (5.2)$$

を制約条件とした

$$\frac{1}{K} \sum_{k=0}^{K-1} R_k = R \quad (5.3)$$

の最小化。

これはすなわち，“ひずみを一定とする条件下において，伝送レートを最小とすること”を意味しており，従来の固定レートを対象としたビット割り当ての問題においてレートとひずみを入れ換えた形となっている．そして，上式における w_k が視覚の空間周波数特性を反映した重み係数となっており，これによって WSNR が規定される．ただし， w_k がすべて 1 であれば，ひずみ尺度は SNR に一致することは自明である．

上記の問題の解は

$$\sigma_{q_k}^2 = \frac{\sigma_r^2}{w_k} \quad (5.4)$$

として与えられる．これはすなわち，可変レート符号化における量子化器の設計問題が，各変換係数ごとに定められる許容歪み $\sigma_{q_k}^2$ を実現する量子化器を最小ビット数で実現する問題に等価になる，ということの意味している．そしてこの制約条件の基に，画質一定の可変レート符号器の設計が図られることになる．

サブバンド符号化の場合には，その符号化の対象がすでに周波数成分であるために，上記の手法をそのまま踏襲することが可能である．

5.3 シミュレーション

5.3.1 符号器の基本構成

サブバンド符号化に基づく ATM 用動画符号化方式として、図 5.5 に示す符号器構成について検討を行う。ここでは、まずフレーム内においてサブバンド分割を行い、各バンドごとにフレーム間符号化を行うことにする。そしてセル廃棄対策としては、低周波成分(最上位階層)にプライオリティを付与することを前提とし、ネットワークのふくそう時などにはそのほかの周波数成分が廃棄されても構わないものとする。ただし、廃棄発生後の画質の回復を目的として、最上位階層以外の周波数帯域のフレーム間符号化に対してはリーク予測を導入する。

5.3.2 リーク予測の効果

セル廃棄対策としては、その廃棄からの画質回復の速度を考慮して、リーク係数の絶対値の値が小さいことが望まれる。一方、データ圧縮としての観点からは、フレーム間の最適予測としての観点から、リーク係数の値がフレーム間の相関係数に一致することが望まれる。

ここで、第 k バンドのリーク係数を a_k 、フレーム間の相関係数を ρ_k とする。経験的に、低周波成分の場合には ρ_k は 1 に近い値を取るが、高周波成分の場合には ρ_k は比較的小さい値になる。故に、高周波成分に対するリーク予測の導入は、セル廃棄対策としても、データ圧縮としても、きわめて自然なものである。図 5.6 は、モノクローム画像 Split/Trevor(10frames/sec) に対して SSKF(5×3) を階層的に施した場合の、バンドごとのエネルギーの平均値、およびそのフレーム間の相関係数を表しているが、この結果から高周波成分に対するリーク予測の導入の必然性が認識される。ただし、最上位階層のフレーム間相関は非常に高く、さらに動き補償の導入によって相関が高められることを考慮すると、最上位階層に対するリーク予測の導入はあまり得策ではないものと考えられる。このような結果は、非可分型の場合についても同様に認められる。

次に図 5.7 は、最上位階層以外のリーク係数を共通にした条件下で、リーク係数を変化(0.7, 0.8, 0.9, 1.0)させた場合の符号化特性の比較を行ったものである。ここでは SSKF(5×3) を用いて、フレーム周期 10frames/sec の Salesman の Y 信号に対するシミュレーションを行った。具体的には、低周波成分の再分割を 2 回施し(ピラミッド分割)、各バンドのフレーム間差分に対して DCT を施した後に線形量子化し、その 20 フレームに渡る統計量から平均の SN、およびエントロピーを求めている。なお、低周波成分(最上位階層)に対しては、そのフレーム間相関の高さを考慮して、動き補償を施している。ここで DCT は 8×8 単位、動きベクトルの探索範囲は 12×12 の領域である。

この図より、0.8、あるいは 0.9 程度のリーク予測であれば、圧縮効率の大きな低下を招かないことがわかる。これは階層的符号化を用いた場合について行った検討結果 [94] に一致する結果である。これは、非

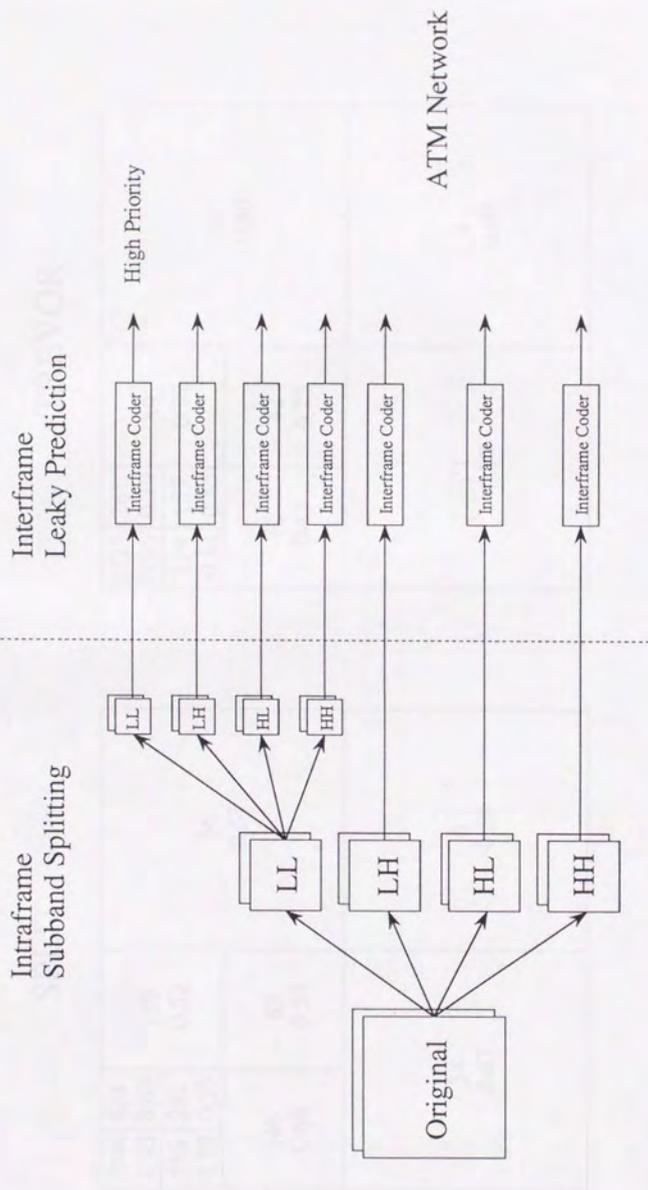


図5.5: ATM用動画画像符号化の基本構成

SPLIT

1946	424	159 0.52	34 0.51
0.93	0.65		
353	292	82 0.51	6 0.39
0.70	0.55		
146	0.68	34 0.67	6 0.39
0.70	0.55		

TREVOR

1524	261	89 0.72	36 0.80
0.95	0.79		
139	155	42 0.23	4 0.09
0.61	0.45		
72	0.43	21 0.37	4 0.09
0.43	0.23		

図5.6: バンドごとの平均パワーとフレーム間相関係数

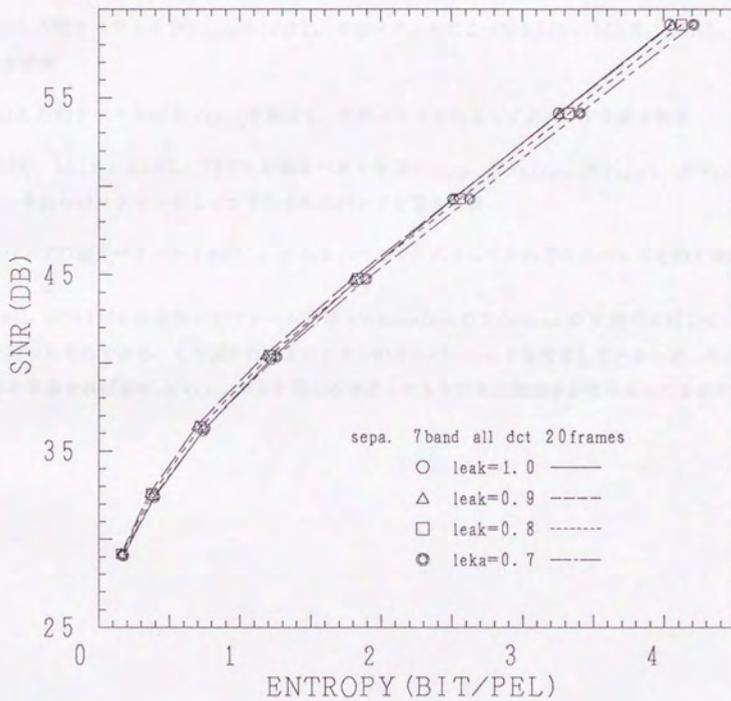


図 5.7: リーク係数と符号化効率

可分型フィルタについても同様に成立する。

5.3.3 動き補償の導入

図 5.5 において、次のような 5 通りの動き補償の導入方法について検討を行った。

1. 最上位階層 (LLLL) の動きベクトル MV_{LLLL} を検出し、そのベクトルによって LLLL のみ動き補償
2. LLLL の動きベクトル MV_{LLLL} を検出し、そのベクトルによって LLLL, LLLH, LLHL, LLHH を動き補償
3. LLLL の動きベクトル MV_{LLLL} を検出し、そのベクトルによって全バンドを動き補償
4. LLLL, LLLH, LLHL, LLHH の動きベクトル MV_{LLLL} , MV_{LLLH} , MV_{LLHL} , MV_{LLHH} を検出し、それらのベクトルによってそれぞれのバンドを動き補償
5. 全バンドの動きベクトルを検出し、それらのベクトルによってそれぞれのバンドを動き補償

図 5.8 は、SSKF(5×3) を用いてフレーム周期 10frames/sec の Salesman の Y 信号に対してシミュレーションを行ったものである。この図では動きベクトルのオーバーヘッドを考慮していないが、そのオーバーヘッド分を考慮すれば逆に MV_{LLLL} のみを用いる方式 1 でも十分な圧縮効率を得られることが示唆される。

5.4 動き補償

動き補償は、動きのある信号を伝送する際に、動きの少ない信号と比べて、より多くのビットを必要とする。動き補償の効率を評価するために、動き補償と符号化効率の関係を示すグラフが示されている。

動き補償の効率を評価するために、動き補償と符号化効率の関係を示すグラフが示されている。動き補償の効率を評価するために、動き補償と符号化効率の関係を示すグラフが示されている。

動き補償の効率を評価するために、動き補償と符号化効率の関係を示すグラフが示されている。動き補償の効率を評価するために、動き補償と符号化効率の関係を示すグラフが示されている。

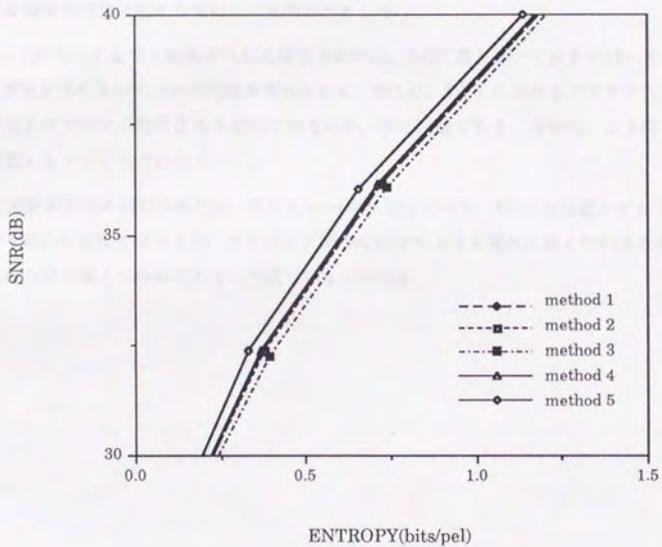


図 5.8: 動き補償と符号化効率

5.4 考察・検討

本章ではサブバンド符号化に基づく ATM 用画像符号化に関する検討を行い、その有効性を明らかにした。ここではまずサブバンド符号化の、高能率符号化、セル廃棄対策、ユニバーサル符号化、画質一定可変レート符号化としての長所を、おもに定性的な観点から明らかにした。

そして次に、最上位階層 (低周波成分) にはプライオリティを付与することを前提とし、リーク予測の導入によって、圧縮効率の大きな劣化を招くことなくセル廃棄対策が実現できることを明らかにした。また、サブバンド符号化に対する動き補償の導入方法として、その最上位階層の動きベクトルのみを用いた場合でも、大きな特性の劣化は起こらないことを明らかにした。

サブバンド符号化に基づく画像符号化に関する検討は、DCT 等と比べて歴史が浅いため、本章で検討を行った課題以外にもいくつかの問題が考えられる。例えば、DCT におけるジクザクスキャン [70] などの実用化手法をサブバンド符号化にも適用できないか、等の問題である。今後は、より実用的な観点からの検討が重要になってくるであろう。

ATM 用画像符号化の国際標準化は、早くも 1994 年に迫っている。DCT を基盤とする H.261 や MPEG との互換性も確かに重要ではあるが、サブバンド符号化がそれよりも優れた多くの利点を有している以上、新しい符号化方式の導入に積極的になって欲しいものである。

Chapter 6

結論

本論文では、サブバンド符号化に基づく画像信号の高効率符号化に関する検討を軸として、以下のようないくつかの検討を行った。

まず第2章では、行列とマルチレートフィルタバンクを用いて、線形変換に基づく各種画像符号化方式の統一表現に関する検討を行った。このような統一表現により、各種方式間の相互関係、相違点を明確にすると共に、予測符号化や変換符号化に対して構築された理論をサブバンド符号化にも直接的に適用することが可能となった。

次に第3章では、サブバンド符号化に対するデータ圧縮の理論の構築を試みた。ここで導かれた圧縮効率の理論的評価尺度 UCG は、任意のサブバンド符号化に適用可能であるだけでなく、結果として予測符号化や変換符号化に対して与えられている尺度を包含する統一的な評価尺度となる。そして、この UCG を用いた特性評価により、これまで不明確なままに検討の進められてきたサブバンド符号化の圧縮効率を理論として明らかにしただけではなく、非直交系フィルタバンクの中には分割バンド数が少ない限りにおいて直交系を凌ぐ圧縮効率を実現するものが存在することなどを示した。

第4章では、サブバンド符号化のための完全再構成フィルタの設計に関する検討を行ったが、ここでは特に UCG と完全再構成条件との結合により、完全再構成フィルタの圧縮効率の最適化を試みた。このアプローチは予測符号化における最適予測係数、あるいは変換符号化における KLT を求める操作に等価であり、結果としてサブバンド符号化の圧縮効率の理論的な最適化を意味している。

最後に第5章では、サブバンド符号化の応用としての ATM 用動画像符号化に関する検討を行った。ここでは定性的、かつ定量的な観点から、従来方式に対するサブバンド符号化の ATM 用動画像符号化としての有効性を明らかにした。

画像信号の高効率符号化の国際標準化も、筆者が大学院生活を送っている間に、ほとんどがまとめられてしまった。残された作業は ATM 用画像符号化の標準化ぐらいのものであろう。ただし、これらの標準化の基礎となっているのは、原理的には遥か昔から知られている予測符号化と DCT である。

本論文の結果は、サブバンド符号化が DCT のブロックひずみの問題を解決するのみならず、レートひずみ特性としても DCT を凌ぐパフォーマンスを実現できることを理論的に明らかにしている。それが革新的に優れた効率を示すわけではないために、しばらくは歴史的に熟成の進められた DCT と予測符号化の支配が続くであろう。しかし、今回の標準化が永遠に残り続けるとは考えにくく、いずれはサブバンド符号化が主流となる時代が来ることを期待している。

そしてサブバンド符号化のもう一つの特徴として、それが網膜の視覚特性を線形フィルタとして模倣したものであることが挙げられる [64,65]。これはすなわち、波形符号化の次世代の符号化として、網膜以降の視覚情報処理を導入した画像符号化の基盤としてもサブバンド符号化が君臨することを期待させるものである。言い替えれば、波形符号化から知的符号化への橋渡しとしてサブバンド符号化が活用される可能

性があるわけである。

いずれにせよ、サブバンド符号化を研究テーマとして選べたことは、非常に有意義な博士時代を過ごす上で好運でもあった。

大学4年の時、ある企業の見学に行き、そこである画像処理のデモを見せてもらった。今になれば、それが非常に簡単なアルゴリズム(平均値処理による動物体の消去)に過ぎなかったことがわかるが、当時は非常におもしろくて画像に興味を持ったのが始まりであった。“これからは画像の時代だ”というデモの発表者の言葉も、強く印象に残った。

それが、修士になって学会に参加するようになると、“画像符号化にはもうやることがない”という言葉をししばしば耳にするようになった。大学4年の時に聞いた上の言葉とは矛盾する、悲しい言葉である。確かに、その基本的なアルゴリズムである予測、変換、ベクトル量子化の3手法はもはや成熟段階にあり、そこから新しいアイデアを出すことは難しいように思えた。また、一方では知的符号化が脚光を浴びていたが、他人の後を追いかけるのはつまらないし、またその理論的背景の不透明さが肌に合わなかった。

サブバンド符号化を用いた画像符号化方式が注目を集め始めたのも、筆者が修士の頃である。幸運なことに、サブバンド符号化に密接な関係のある階層的符号化が、筆者の修士論文の核であった。また、このサブバンド符号化に対する関心度は、欧米での活況とは裏腹に、国内ではあまり高くはなかった。それならば、という訳で、修士の中頃からサブバンド符号化を用いた画像符号化の検討を始めた。

ペーパーを読んで行くと、最初は欧米のオリジナリティを求めるパワーに圧倒されるばかりであったが、やがて多くの疑問が湧いてくるようになった。シミュレーションの結果はいくつか与えられてはいるが、サブバンド符号化は本当にDCTを超えることができるのか?さまざまなフィルタ構成が提案されているが、どのフィルタ構成が最も優れた圧縮効率を実現するのか?最適予測やKLTのように、データ圧縮として最適な完全再構成フィルタはどのように設計したらよいのか?

本論文は、筆者がサブバンド符号化に対して感じたこれらの疑問の解決を最大の骨子としている。結果として、これらの疑問はすべて解決できたから、個人的には非常に満足している。線形変換の枠を越えられなかったことには若干の悔いがあるが、本論文がサブバンド符号化に関与するすべての研究者にとって何等かの一助となればと期待している。

謝辞

5年にも渡る大学院生活において、未熟な小生を終始励まされると共に、研究者としての考え方を説かれ、親身の御指導、御鞭撻を賜りました安田靖彦教授に心から感謝の意を表します。

また、筆者の研究を進めるに当たり、貴重な御討論、御助言を頂いた瀬崎薫講師、加藤茂夫助手（現・宇都宮大）、木本伊彦助手、小松邦紀技官に心より御礼申し上げます。

また、日頃御討論いただき、研究室生活でお世話になった横澤一彦助教授、学生の陳艶萍さん、梶谷昭彦氏、黄文翔氏、木村俊一氏、浅井基博氏、大道文雄氏、渥美栄司氏、黄英傑氏、研究生の大澤秀史氏、片山昭宏氏、田中孝一氏、大山昌一氏、受託研究員の恩田勝政氏、釣部智行氏、樋口学氏、井上直幸氏、博士研究員の季華妹さん、留学生の周君に感謝致します。

また、秘書の森真理子さん、小林美香さん、松浦双葉さん、富田喜久子さん、吉田治美さん、いろいろとお世話になりました。

最後に、私をここまで育ててくれた父と母に、感謝の意を表したいと思います。どうも有難うございました。

参考文献

- [1] N.S.Jayant and P.Noll: "Digital coding of waveforms", *Englewood Cliffs NJ*, Prentice-Hall (1984)
- [2] W.K.Pratt: "Digital image processing", John Wiley & Sons (1978)
- [3] A.N.Netravali and J.O.Limb: "Picture coding : a review", *Proc. IEEE*, Vol.68, No.3, pp.366-406 (March.1980)
- [4] 吹抜: "画像のデジタル信号処理", 日刊工業新聞社 (1981)
- [5] 釜江, 吹抜: "デジタル画像通信", 産業図書 (1985)
- [6] H.G.Musmann, P.Pirsch and H.J.Grallert: "Advances in picture coding", *Proc. IEEE*, Vol.73, No.4, pp.523-548 (April.1985)
- [7] M.Kunt, A.Ikonomopoulos and M.Kocher: "Second-generation image-coding techniques", *Proc. IEEE*, Vol.73, No.4, pp.549-574 (April.1985)
- [8] 安田浩綱: "マルチメディア符号化の国際標準", 丸善 (1991)
- [9] 原島博綱: "画像情報圧縮", オーム社 (1991)
- [10] R.E.Crochiere and L.R.Rabiner: "Interpolation and decimation of digital signals - a tutorial review", *Proc. IEEE*, pp.300-331 (March.1981)
- [11] R.E.Crochiere and L.R.Rabiner: "Multirate digital signal processing", Prentice Hall (1983)
- [12] E.Dubois: "The sampling and reconstruction of time-varying imagery with application in video systems", *Proc. IEEE*, pp.502-522 (April.1985)
- [13] P.P.Vaidyanathan: "Quadrature mirror filter banks, M-band extensions and perfect-reconstruction techniques", *IEEE ASSP Magazine*, pp.4-20 (July.1987)
- [14] P.A.Regalia, S.K.Mitra and P.P.Vaidyanathan: "The digital all-pass filter: a versatile signal processing building block", *Proc. IEEE*, pp.19-37 (Jan.1988)

- [15] P.P.Vaidyanathan: "Multirate digital filters, filter banks, polyphase networks, and applications: a tutorial", *Proc. IEEE*, pp.56-93 (Jan.1990)
- [16] O.Rioul and M.Vetterli: "Wavelets and signal processing", *IEEE Signal Processing Magazine*, pp.14-38 (Oct.1991)
- [17] A.N.Akansu and R.A.Haddad: "Signal decomposition techniques: transforms, subbands and wavelets", *Proc. SPIE Visual Communication and Image Processing*, Short Course Notes (Nov.1991)
- [18] 小柳: "フーリエ解析", 培風館 (1979)
- [19] 前田: "デジタル信号処理の基礎", オーム社 (1980)
- [20] 三谷: "デジタルフィルタデザイン", 昭晃堂 (1987)
- [21] 辻井: "デジタル信号処理の基礎", 電子情報通信学会 (1988)
- [22] 安田, 高木, 加藤, 栗野: "階層的符号化法による静止画像の段階的の伝送および表示", 信学論 (B), J63-B, 4 (April.1980)
- [23] P.J.Burt and E.H.Adelson: "The Laplacian pyramid as a compact image code", *IEEE Trans. Commun.*, COM-31, 4, pp.532-540 (April.1983)
- [24] 谷本, 山田: "相補サブサンプリング方式における画像復元", 信学論 (B), J71-B, 12, pp.1511-1516 (Dec.1988)
- [25] 齊藤, 秋吉, 原島: "ラブラシアン・ピラミッド符号化法の改善 - ベクトル量子化法の導入 -", 信学論 (B), J71-B, 12, pp.1517-1527 (Dec.1988)
- [26] M.Vetterli and D.L.Gall: "Perfect reconstruction FIR filter banks: some properties and factorizations", *IEEE Trans. Acoust., Speech & Signal Process.*, ASSP-37, 7, pp.1057-1071 (July.1989)
- [27] D.Esteban and C.Galand: "Application of quadrature mirror filters to split band voice coding schemes", *Proc. ICASSP'77*, pp.191-195 (May.1977)
- [28] J.D.Johnston: "A filter family designed for use in quadrature mirror filter banks", *Proc. ICASSP'80*, pp.291-194 (April.1980)
- [29] R.E.Crochiere: "Subband coding", *Bell Syst. Tech. J.*, 60, pp.1633-1653 (Sep.1981)
- [30] C.R.Galand and H.J.Nussbaumer: "New quadrature mirror filter structures", *IEEE Trans. Acoust., Speech & Signal Process.*, ASSP-32, 3, pp.522-531 (June.1984)

- [31] P.C.Millar: "Recursive quadrature mirror filters - criteria specification and design method", *IEEE Trans. Acoust., Speech & Signal Process.*, **ASSP-33**, 2, pp.413-420 (April.1985)
- [32] M.Vetterli: "Filter banks allowing perfect reconstruction", *Signal Processing*, 10, pp.219-244 (April.1986)
- [33] M.Smith and T.P.Barnwell III: "Exact reconstruction techniques for tree-structured subband coders", *IEEE Trans. Acoust., Speech & Signal Process.*, **ASSP-34**, 3, pp.434-441 (June.1986)
- [34] F.Grenez: "Chebyshev design of filters for subband coders", *IEEE Trans. Acoust., Speech & Signal Processing*, **ASSP-36**, 2, pp.182-185 (Feb.1988)
- [35] T.Kronander: "A new approach to recursive mirror filters with a special application in subband coding of images", *IEEE Trans. Acoust., Speech & Signal Processing*, **ASSP-36**, 9, pp.1496-1500 (Sep.1988)
- [36] C.W.Kim and R.Ansari: "FIR/IIR exact reconstruction filter banks with application to subband coding of images", *1991 Midwest CAS Symposium* (May.1991)
- [37] J.H.Rothweiler: "Polyphase quadrature filters - a new subband coding technique", *Proc. ICASSP'83*, pp.1280-1283 (March.1983)
- [38] P.L.Chu: "Quadrature mirror filter design for an arbitrary number of equal bandwidth channels", *IEEE Trans. Acoust., Speech & Signal Process.*, **ASSP-33**, 2, pp.203-218 (Feb.1985)
- [39] J.P.Princen and A.B.Bradley: "Analysis/synthesis filter bank design based on time domain aliasing cancellation", *IEEE Trans. Acoust., Speech & Signal Process.*, **ASSP-34**, 5, pp.1153-1161 (Oct.1986)
- [40] P.P.Vaidyanathan et al.: "Improved technique for design of perfect reconstruction FIR QMF banks with lossless polyphase matrices", *IEEE Trans. Acoust., Speech & Signal Process.*, **ASSP-37**, 7, pp.1042-1056 (July.1989)
- [41] 黒沢他: "不等間隔完全 QMF システムの構成", *信学技報*, **DSP89-11**.
- [42] 小松, 甲藤, 安田: "並列型完全再構成フィルタの一構成方式", *PCSJ'91* (Oct.1991)
- [43] M.Vetterli: "Multi-dimensional subband coding: some theory and algorithms", *Signal Processing*, 6, pp.97-112 (April.1984)
- [44] J.W.Woods and S.D.O'Neil: "Subband coding of images", *IEEE Trans. Acoust., Speech & Signal Process.*, **ASSP-34**, 5, pp.1278-1288 (Oct.1986)

- [45] E.H.Adelson, E.Simoncelli and R.Hingorani : "Orthogonal pyramid transforms for image coding", *Proc. SPIE Conf. Visual Commun. and Image Processing II*, pp.50-58 (Oct.1987)
- [46] H.Gharavi and A.Tabatabai: "Subband coding of monochrome and color images", *IEEE Trans. Circuit and Systems, CAS-35*, 2, pp.207-214 (Feb.1988)
- [47] D.L.Gall and A.Tabatabai : "Subband coding of digital images using symmetric short kernel filters and arithmetic coding techniques", *Proc. IEEE ICASSP'88*, pp.761-764 (June.1988)
- [48] P.H.Westerink et al., "Subband coding of images using vector quantization", *IEEE Trans. Commun.*, COM-36, 6, pp.713-719 (June.1988)
- [49] 尾高: "完全復元可能なフィルタバンクを用いた画像信号サブバンド符号化方式における基礎特性", *信学技報, IE89-77*.
- [50] M.Vetterli, J.Kovacevic and D.L.Gall: "Perfect reconstruction filter banks for HDTV representation and coding", *Proc. 3rd Int. Workshop on HDTV* (Aug.1989)
- [51] E.P.Simoncelli and E.H.Adelson: "Non-separable extensions of quadrature mirror filters to multiple dimensions", *Proc. IEEE* (May.1990)
- [52] M.J.T.Smith and S.L.Eddins: "Analysis/synthesis techniques for subband image coding", *IEEE Trans. Acoust., Speech & Signal Processing, ASSP-38*, 8, pp.1446-1456 (Aug.1990)
- [53] M.Vetterli, J.Kovacevic and D.L.Gall: "Perfect reconstruction filter banks for HDTV representation and coding", *Signal Processing: Image Communication*, pp.349-363 (Oct.1990)
- [54] A.N.Akansu, R.A.Haddad and H.Caglar: "Perfect reconstruction binomial QMF-wavelet transform", *Proc. SPIE Visual Communication and Image Processing*, pp.609-617 (Nov.1990)
- [55] A.N.Akansu and Y.Liu: "On-signal decomposition techniques", *Optical Engineering*, vol.30, pp.912-920 (July.1991)
- [56] 本間他: "画像のサブバンド符号化のための周期入力生成法", 1991 秋季信学全大, D-128.
- [57] M.Porat and Y.Y.Zeevi: "The generalized Gabor scheme of image representation in biological and machine vision", *IEEE Trans. Pattern Anal. Machine Intell.*, PAMI-10, 4, pp.452-468 (July.1988)
- [58] M.Porat and Y.Y.Zeevi: "Gram-Gabor approach to optimal image representation", *Proc. SPIE Visual Communication and Image Processing*, pp.1474-1478 (Nov.1990)
- [59] T.Ebrahimi, T.R.Reed and M.Kunt: "Video coding using a pyramidal Gabor expansion", *Proc. SPIE Visual Communication and Image Processing*, pp.489-502 (Nov.1990)

- [60] T.Ebrahimi and M.Kunt: "Image compression by Gabor expansion", *Optical Engineering*, vol.30, pp.873-880 (July.1991)
- [61] H.S.Malvar and D.H.Staelin: "The LOT: transform coding without blocking effects", *IEEE Trans. Acoust., Speech & Signal Processing*, ASSP-37, 4, pp.553-559 (April.1989)
- [62] H.S.Malvar: "Lapped transforms for efficient transform/subband coding", *IEEE Trans. Acoust., Speech & Signal Processing*, ASSP-38, 6, pp.969-978 (June.1990)
- [63] I.Daubechies: "Orthogonal bases of compactly supported wavelets", *Comm. Pure Appl. Math.*, pp.909-996 (Nov.1988)
- [64] S.G.Mallat: "A theory for multiresolution signal decomposition: the wavelet representation", *IEEE Trans. Pattern Anal. Machine Intell.*, PAMI-11, 11, pp.674-693 (July.1989)
- [65] S.G.Mallat: "Multifrequency channel decompositions of images and wavelet models", *IEEE Trans. Acoust., Speech & Signal Processing*, ASSP-37, 12, pp.2091-2110 (Dec.1989)
- [66] M.Vetterli and C.Herley: "Wavelets and filter banks: relationships and new results", *Proc. ICASSP'90*, D12.1.
- [67] M.Antonini et al.: "Image coding using vector quantization in the wavelet transform domain", *Proc. ICASSP'90*, M9.10.
- [68] N.Ahmed, T.Natarajan and K.R.Rao: "Discrete cosine transform", *IEEE Trans. Computer*, pp.90-93 (Jan.1974)
- [69] W.H.Chen and S.C.Fralick: "Image enhancement using cosine transform filtering", *Image Science Math. Symp.*, Monterey, CA, pp.186-192 (Nov.1976)
- [70] W.H.Chen and W.K.Pratt: "Scene adaptive coder", *IEEE Trans. Commun*, COM-32, 3, pp.225-232 (March.1984)
- [71] K.N.Ngan: "Two-dimensional transform domain decimation techniques", *Proc. ICASSP'86*, 20.5.1.
- [72] B.Chitprasert and K.R.Rao: "Discrete cosine transform filtering", *Proc. ICASSP'90* D3.1.
- [73] 安田靖彦: "映像バケット伝送技術", *PCSJ'87* (Sep.1987)
- [74] 安田靖彦: "映像バケット伝送", *PCSJ'89 画像符号化講演会* (Nov.1989)
- [75] J.J.Dubnowski and R.E.Crochiere: "Variable rate coding of speech", *Bell Syst. Tech. Journal*, vol.58, 3, pp.577-600 (March.1979)

- [76] B.Maglaris et al.: "Performance analysis of statistical multiplexing for packet video sources", *IEEE Trans. on Commun.*, pp.834-843 (July.1988)
- [77] M.Ghanbari: "Two-layer coding of video signals for VBR Networks", *IEEE Journal on Selected Areas in Commun.*, pp.771-781 (June.1989)
- [78] F.Kishino et al.: "Variable bit rate coding of video signals for ATM networks", *IEEE Journal on Selected Areas in Commun.*, pp.801-806 (June.1989)
- [79] M.Nomura et al.: "Basic characteristics of variable rate coding in ATM environment", *IEEE Journal on Selected Areas in Commun.*, pp.752-760 (June.1989)
- [80] G.Karlsson and M.Vetterli: "Subband coding of video signals for packet-switched networks", *Proc. SPIE Conf. Visual Commun. and Image Processing II*, pp.446-456 (Oct.1987)
- [81] 岸本: "映像信号のバケット廃棄に対する補償法の一提案", 信学技報, IN87-12.
- [82] W.Verbiest: "Video coding in an ATD environment", *Proc. of the 3rd Int. Conf. on New Systems and Services in Telecommun.* (Dec.1986)
- [83] J.C.Darragh and R.L.Baker: "Fixed distortion subband coding of images for packet switched networks", *Journal on Selected Areas in Commun.*, vol.7, 5, 789-800 (June.1989)
- [84] Y.Yamaguchi: "Weighting function for evaluation of random television interferences with different standards", *Journal of the SMPTE*, vol.76, 3, pp.176-179 (March.1967)
- [85] J.B.Nill: "A visual model weighted cosine transform for image compression and quality assessment", *IEEE Trans. on Commun.*, COM-33, 6, pp.551-557 (June.1985)
- [86] D.J.Granrath: "The role of human visual models in image processing", *Proc. of IEEE*, vol.69, 5, pp.552-561 (May.1981)
- [87] F.X.J.Lukas and Z.L.Budrikis: "Picture quality prediction based on a visual model", *IEEE Trans. on Commun.*, COM-30, 7, pp.1679-1692 (July.1982)
- [88] M.Miyahara: "Quality assessment for visual service", *IEEE Commun. Magazine*, pp.51-60 (Oct.1988)
- [89] *Proc. of the 2nd Int. Workshop on Packet Video*, (Sep.1988)
- [90] *Proc. of the 3rd Int. Workshop on Packet Video*, (March.1990)
- [91] *Proc. of the 4th Int. Workshop on Packet Video*, (Aug.1991)

発表文献

• 学会誌論文

1. 甲藤, 安田: “階層的符号化を用いた映像パケット通信におけるパケット廃棄対策”, 電子情報通信学会論文誌 B-I, pp.1094-1102 (Nov.1989)
2. J.Katto and Y.Yasuda: “A New Structure of the Perfect Reconstruction Filter Banks for Subband Coding”, *Trans. IEICE E*, pp.1616-1624 (Oct.1990)
3. J.Katto, K.Onda and Y.Yasuda: “Variable Bit Rate Coding Based on Human Visual System”, *Signal Processing: Image Communication*, pp.313-320 (Sep.1991)
4. J.Katto and Y.Yasuda: “Performance Evaluation of Subband Coding and Optimization of Its Filter Coefficients”, *Journal of Visual Communication and Image Representation*, Academic Press, pp.303-313 (Dec.1991)

• 国際会議

1. J.Katto, S.Kato and Y.Yasuda: “Hierarchical Coding for Packet Loss Protection”, *The 2nd International Workshop on Packet Video* (Sep.1988)
2. J.Katto and Y.Yasuda: “Variable Bit Rate Coding Based on Human Visual Model”, *The 3rd International Workshop on Packet Video* (March.1990)
3. J.Katto and Y.Yasuda: “Variable Bit Rate Coding of Video Signals with Leaky Prediction”, *The 4th International Workshop on Packet Video* (Aug.1991)
4. J.Katto and Y.Yasuda: “Performance Evaluation of Subband Coding”, *Picture Coding Symposium '91* (Sep.1991)
5. J.Katto and Y.Yasuda: “Performance Evaluation of Subband Coding and Optimization of Its Filter Coefficients”, *SPIE Visual Communications and Image Processing '91* (Nov.1991)

• 国内大会

1. 甲藤, 加藤, 安田: “階層的符号化を導入した映像バケット通信におけるロス対策”, 昭 63 秋季・電子情報通信学会全国大会 (Sep.1988)
2. 甲藤, 安田: “帯域分割符号化を用いたセル廃棄対策”, 平 1 春季・電子情報通信学会全国大会 (March.1989)
3. 甲藤, 安田: “視覚特性を考慮した可変レート符号化における品質制御”, 平 1 秋季・電子情報通信学会全国大会 (Sep.1989)
4. 甲藤, 安田: “サブバンド分割における完全再構成フィルタに関する一検討”, 平 2 春季・電子情報通信学会全国大会 (March.1990)
5. 恩田, 甲藤, 安田: “視覚特性を考慮した可変レート符号化に関する一検討”, 平 2 春季・電子情報通信学会全国大会 (March.1990)
6. 甲藤, 安田: “数論変換に基づく帯域分割符号化”, 平 2 秋季・電子情報通信学会全国大会 (Oct.1990)
7. 甲藤, 安田: “サブバンド符号化の特性評価と, フィルタ係数の最適化問題について”, 平 3 春季・電子情報通信学会全国大会 (March.1991)
8. 釣部, 甲藤, 安田: “非可分型フィルタを用いたサブバンド符号化”, 平 3 春季・電子情報通信学会全国大会 (March.1991)
9. 小松, 甲藤, 安田: “計算量をパラメータとしたサブバンド符号化の特性評価”, 平 3 春季・電子情報通信学会全国大会 (March.1991)
10. 甲藤, 安田: “サブバンド符号化の特性評価 (2次元の場合)”, 平 3 秋季・電子情報通信学会全国大会 (Sep.1991)
11. 小松, 甲藤, 安田: “非直交系帯域分割符号化の最適ビット割当て”, 平 3 秋季・電子情報通信学会全国大会 (Sep.1991)
12. 甲藤, 安田: “Biorthogonal フィルタバンクの圧縮効率の最適化”, 平 4 春期・電子情報通信学会全国大会 (March.1992)
13. 小松, 甲藤, 安田: “並列型完全再構成フィルタの圧縮効率の最適化”, 平 4 春期・電子情報通信学会全国大会 (March.1992)
14. 樋口, 甲藤, 安田: “サブバンドコーディングにおける動き補償予測方式の比較検討”, 平 4 春期・電子情報通信学会全国大会 (March.1992)

• 学会研究会

1. 甲藤, 加藤, 安田: “階層的符号化を導入した映像バケット伝送におけるロス対策”, 画像符号化シンポジウム (Oct.1988)

2. 甲藤, 加藤, 安田: “映像バケット伝送におけるセル廃棄対策”, 情報理論とその応用シンポジウム (Nov.1988)
3. 甲藤, 安田: “視覚特性を考慮した可変レート符号化に関する一検討”, 画像符号化シンポジウム (Oct.1989)
4. 甲藤, 安田: “サブバンド分割における無ひずみフィルタの新たな構成法”, 電子情報通信学会技術報告, IE89-98 (Feb.1990)
5. 甲藤, 安田: “帯域分割符号化における完全再構成フィルタの一構成とその画像への応用”, テレビジョン学会技術報告, ICS90-38 (May.1990)
6. 甲藤, 安田: “サブバンド符号化に関する一検討”, 画像符号化シンポジウム (Oct.1990)
7. 甲藤, 安田: “サブバンド符号化の特性評価とそのフィルタ係数の最適化について”, 電子情報通信学会技術報告, IE91-10 (May.1991)
8. 甲藤, 釣部, 安田: “サブバンド符号化に基づく ATM 用動画画像符号化に関する諸検討”, 電子情報通信学会技術報告, IE91-39 (July.1991)
9. 甲藤, 安田: “サブバンド符号化の特性評価 (2次元の場合)”, 画像符号化シンポジウム (Oct.1991)
10. 小松, 甲藤, 安田: “並列型完全再構成フィルタの一構成”, 画像符号化シンポジウム (Oct.1991)

付録

A.1 DPCMの符号化ゲインの厳密な評価

A.1.1 開ループの場合

$h = \rho$ の場合, (1.7)式より

$$A_0 = 1 - \rho^2 \quad (\text{A.1})$$

となる. これに対して, 図2.6より

$$\begin{aligned} B_0 &= 1 + \rho^2 + \rho^4 + \cdots \\ &= \frac{1}{1 - \rho^2} \end{aligned} \quad (\text{A.2})$$

となる. よって, (3.19)式より

$$G_{h=\rho, \text{open}} = 1 \quad (\text{A.3})$$

となり, PCMに対するゲインは全くない. ただし, この開ループDPCMの場合, $h_{\text{opt}} = (1 - \sqrt{1 - \rho^2})/\rho$ であることが知られている [1].

A.1.2 閉ループの場合

局部復号器の挿入によって, (1.5)式より $B_0 = 1$ であることは自明である. これに対して, $h = \rho$ とした場合, (1.7)式は厳密には

$$\sigma_d^2 = (1 - \rho^2)\sigma_x^2 + \rho^2 \cdot \sigma_q^2 \quad (\text{A.4})$$

として表される. よって, (1.9)を用いて

$$G_{h=\rho, \text{closed}} = \frac{1 - \epsilon^2 \cdot 2^{-2R}}{1 - \rho^2} \quad (\text{A.5})$$

が与えられる. (A.5)式は, レートが十分に大きければ理論的最適値 (1.12)式に漸近するが, レートが小さくなるとそれにつれて小さくなっていく. すなわち, 閉ループDPCMは, 高レート環境では優れた圧縮効率を実現できるが, 逆に低レート環境では圧縮効率が大きく低下してしまうことになる.

A.2 パラメータ B_k の導出

簡単のために、2バンド分割の場合を考える ($\alpha_0 = \alpha_1 = 1/2$)。ここでまず (2.2) 式の変形として、量子化操作がない場合には

$$\begin{aligned} x(2n) &= \sum_k g_0(2k)y_0(n-k) + \sum_k g_1(2k)y_1(n-k) \\ x(2n+1) &= \sum_k g_0(2k+1)y_0(n-k) + \sum_k g_1(2k+1)y_1(n-k) \end{aligned} \quad (\text{A.6})$$

が成立する。ただし、ここでは完全再構成 ($\hat{x}(n) = x(n)$) が実現されることを前提としている。一方、量子化操作が入ってくると、上式と同様にして

$$\begin{aligned} \hat{x}(2n) &= \sum_k g_0(2k)u_0(n-k) + \sum_k g_1(2k)u_1(n-k) \\ \hat{x}(2n+1) &= \sum_k g_0(2k+1)u_0(n-k) + \sum_k g_1(2k+1)u_1(n-k) \end{aligned} \quad (\text{A.7})$$

が成立する。ここで $\hat{x}(n)$ は量子化誤差を含んだ形となっている。そして (A.6) 式と (A.7) 式の差を取ることによって、再生誤差が

$$\begin{aligned} r(2n) &= \sum_k g_0(2k)q_0(n-k) + \sum_k g_1(2k)q_1(n-k) \\ r(2n+1) &= \sum_k g_0(2k+1)q_0(n-k) + \sum_k g_1(2k+1)q_1(n-k) \end{aligned} \quad (\text{A.8})$$

として表されることになる。

ここで量子化誤差間の相関はない、と仮定する。そして再生誤差の平均値は 0 であるものとし、(A.8) 式の自乗平均を取ってやれば

$$\sigma_r^2 = \frac{1}{2} \sum_k g_0(k)^2 \cdot \sigma_{q_0}^2 + \frac{1}{2} \sum_k g_1(k)^2 \cdot \sigma_{q_1}^2 \quad (\text{A.9})$$

として再生誤差分散と量子化誤差分散の関係式が求められる。故に、

$$\begin{aligned} B_0 &= \frac{1}{2} \sum_k g_0(k)^2 \\ B_1 &= \frac{1}{2} \sum_k g_1(k)^2 \end{aligned} \quad (\text{A.10})$$

となる。ここで $1/2$ は α_0 および α_1 に対応しており、 $K > 2$ の場合には $1/2$ の部分を α_k で置き換えてやればよい。

A.3 多段接続の場合のパラメータ B_k の導出

簡単のために、2バンド分割の場合を考える。まず多段接続を行わない (1段のみ) 場合には、定義から

$$\sigma_r^2 = B_0 \sigma_{q_0}^2 + B_1 \sigma_{q_1}^2 \quad (\text{A.11})$$

となるが、ここで

$$q_0(n) = y_0(n) - u_0(n) \quad (\text{A.12})$$

であり、 $y_0(n)$ は入力信号 $x(n)$ のフィルタ出力 (量子化入力)、 $u_0(n)$ はその量子化出力である。ここでこの $y_0(n)$ を再分割することを考える。

$y_0(n)$ が $y_{00}(n)$ と $y_{01}(n)$ に分割され、それらを量子化した $u_{00}(n)$ と $u_{01}(n)$ から再構成した出力が $u_0(n)$ であるとする、その定義から

$$\sigma_{q_0}^2 = B_0 \sigma_{q_{00}}^2 + B_1 \sigma_{q_{01}}^2 \quad (\text{A.13})$$

が成立する。故に (A.11)(A.13) 式から、再生誤差分散は

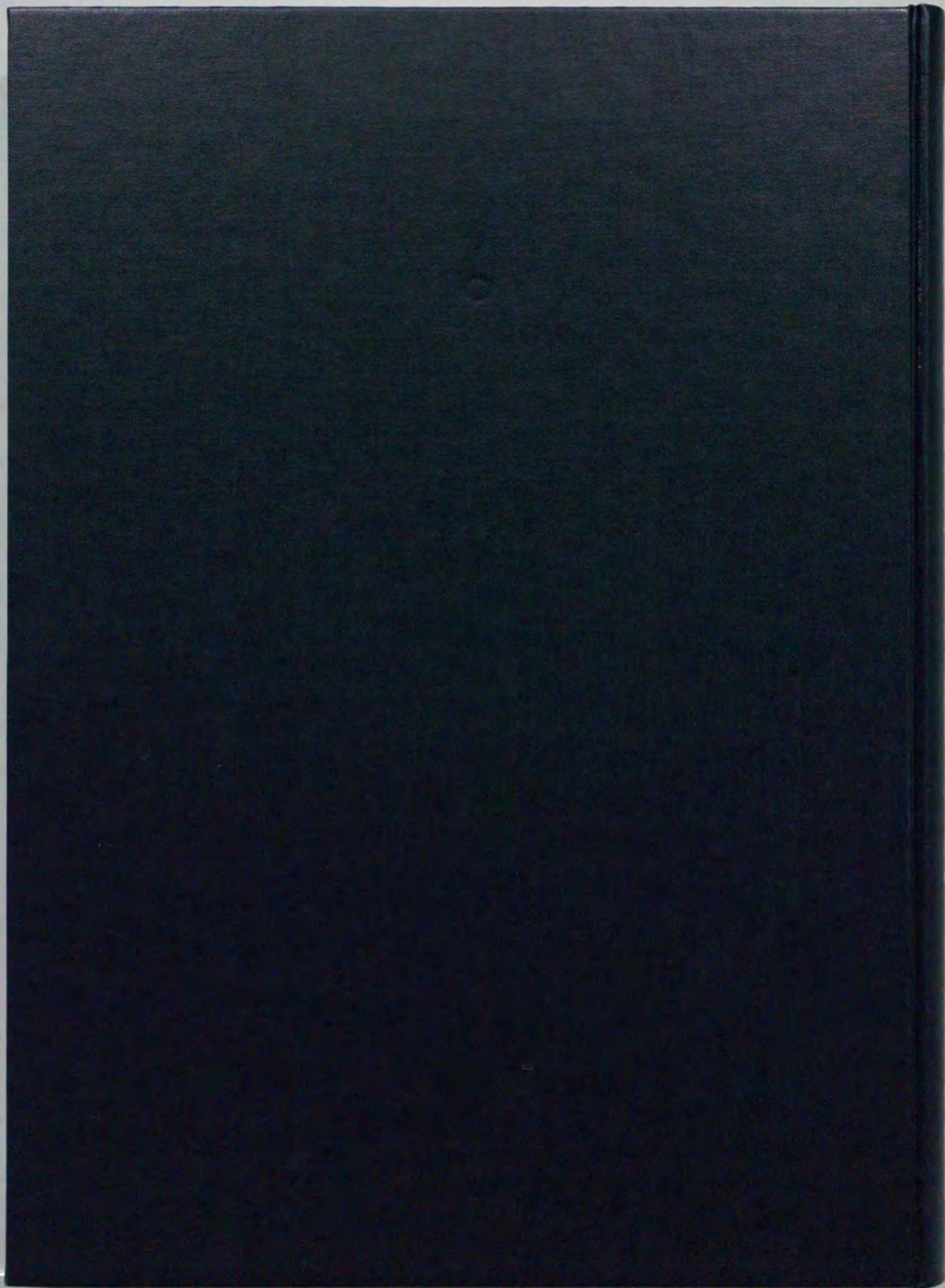
$$\sigma_r^2 = B_0 B_0 \sigma_{q_{00}}^2 + B_0 B_1 \sigma_{q_{01}}^2 + B_1 \sigma_{q_1}^2 \quad (\text{A.14})$$

によって表される。

さらに再分割を進めた場合、あるいは $K > 2$ バンド分割の場合も同様であり、結局

$$B_k = B_{\text{stage}(1)} \cdot B_{\text{stage}(2)} \cdots \quad (\text{A.15})$$

が成立することが証明される。



inches 1 2 3 4 5 6 7 8
cm 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19

Kodak Color Control Patches

© Kodak, 2007 TM Kodak

Blue	Cyan	Green	Yellow	Red	Magenta	White	3/Color	Black

Kodak Gray Scale



© Kodak, 2007 TM Kodak

A 1 2 3 4 5 6 M 8 9 10 11 12 13 14 15 B 17 18 19

