



# デジタル画像信号の符号化に関する研究

黒木, 修隆

---

(Degree)

博士 (工学)

(Date of Degree)

1995-03-31

(Date of Publication)

2008-10-31

(Resource Type)

doctoral thesis

(Report Number)

甲1394

(JaLCOI)

<https://doi.org/10.11501/3105417>

(URL)

<https://hdl.handle.net/20.500.14094/D1001394>

※ 当コンテンツは神戸大学の学術成果です。無断複製・不正使用等を禁じます。著作権法で認められている範囲内で、適切にご利用ください。



博士論文

デジタル画像信号の  
符号化に関する研究

平成7年1月

神戸大学大学院自然科学研究科

黒木修隆

# 目次

<b>第 1 章 緒 論</b> .....	<b>1</b>
1.1 画像データ圧縮の背景 .....	1
1.2 本研究の目的 .....	3
<b>第 2 章 モノクロ画像の可逆符号化</b> .....	<b>5</b>
2.1 緒言 .....	5
2.2 JPEGの手法 .....	5
2.2.1 2次元予測 .....	5
2.2.2 JPEG手法の性能 .....	6
2.2.3 問題点 .....	9
2.3 ダイナミック・レンジ変換法 .....	11
2.4 動的予測 .....	12
2.4.1 アルゴリズム .....	12
2.4.2 高速演算手法 .....	14
2.5 計算機シミュレーション .....	17
2.5.1 システム構成 .....	18
2.5.2 実験結果 .....	19
2.6 結言 .....	22
<b>第 3 章 カラー画像の可逆符号化</b> .....	<b>24</b>
3.1 緒言 .....	24
3.2 コンポーネント間予測 .....	24
3.3 3次元予測 .....	26
3.4 勾配予測方式 .....	27
3.5 計算機シミュレーション .....	28
3.5.1 予測係数 .....	28
3.5.2 予測誤差信号のエントロピー .....	31
3.6 結言 .....	33
<b>第 4 章 予測誤差信号のための算術符号化</b> .....	<b>34</b>
4.1 緒言 .....	34
4.2 動的算術符号化 .....	35

4.2.1	確率推定方法	35
4.2.2	動的符号化のアルゴリズム	37
4.3	計算機シミュレーション	39
4.3.1	7種類の符号化手法	39
4.3.2	手法I,II,III,IVおよびVIによる圧縮比	43
4.3.3	手法VおよびVIIによる圧縮比	44
4.3.4	9枚の画像に対する圧縮比	45
4.3.5	予測器との組み合わせに対する圧縮比	46
4.3.6	処理時間	46
4.3.7	リアルタイム処理能力	48
4.4	結言	49
<b>第5章 モノクロ画像の非可逆符号化</b>		<b>50</b>
5.1	緒言	50
5.2	帯域間予測ハール・ウェーブレット変換	51
5.2.1	ハール・ウェーブレット変換	51
5.2.2	帯域間予測	52
5.2.3	2次元への拡張	53
5.3	帯域分割手法としての性能	54
5.3.1	等価フィルタバンク	55
5.3.2	ブロック歪み	57
5.3.3	演算量	60
5.4	計算機シミュレーション	62
5.4.1	変換後の信号のエントロピー	62
5.4.2	圧縮比と再生画像	65
5.4.3	プログレッシブ・ビルドアップ	68
5.5	結言	68
<b>第6章 結 論</b>		<b>71</b>
付 録		73
謝 辞		74
参考文献		75

## 内容梗概

現代の情報化社会において、画像を用いたメディアは大きく注目されつつある。テレビ電話／会議、カラーファクシミリ、ビデオディスク、画像データベース、ナビゲーション・システムなど、画像を用いたメディアは今後大きく発展が期待されている。これらの画像データを効率よく処理、蓄積、そして伝送するために、画像データの圧縮技術は不可欠なものとなっている。本論文ではこの圧縮技術に関する研究が行われている。

第1章では画像データ圧縮の背景、圧縮の概要が述べられた後に、本論文の目的が明らかにされている。これらの目的のもとに、第2章から第5章では四つの研究が行われている。それぞれの研究において、従来まで行われてきた手法の問題点を指摘し、その解決手法を提案している。

第2章ではモノクロ画像の可逆符号化の研究が行われている。従来の予測符号化は冗長抑圧の点では必ずしも最適な手法ではなかった。本章ではこの点が指摘され、冗長抑圧能力を高めることを目的とした研究が行われている。本章では従来の手法が十分な冗長抑圧を行わない原因として、

(1) 予測により信号のエントロピーが増加する可能性があること、  
(2) 画像信号の性質が局所的に変化する可能性があること、  
を指摘している。これに対して本論文は、

(1) ダイナミックレンジ変換法、  
(2) 動的予測法、

による冗長抑圧能力の改善を提案している。実験では提案手法の圧縮能力が評価されている。

第3章ではカラー RGB 画像の可逆符号化の研究が行われている。カラー RGB 信号は YUV 信号に変換すると打ちきり誤差を生じるため、有効な可逆符号化法が行われていなかった。本論文ではこの解決法として、RGB 信号を直接予測符号化する 3次元予測法を提案している。この手法は RGB 信号を 3次元の信号とみなし、空間内の 1点をその周りの値から

予測するという手法である。これによって画素間の冗長性とRGB信号間の冗長性が同時に取り除かれている。実験では提案手法の圧縮能力が評価されている。

第4章では予測誤差信号のための符号化について研究が行われている。本章では符号化における従来の確率推定法の弱点として、

- (1) 雑音によって推定が不安定になる点、
- (2) 局所的な確率の変動に追従しない点、

を挙げている。これらの点を改善するために本論文では新しい確率推定法を提案している。提案手法は

- (1) 手法の特性を調整するパラメータをもたず、
- (2) 経験的に調整されたしきい値や変数テーブルをもたない、

という点で理論的に明快である。実験では圧縮率の改善量が示されている。

第5章ではモノクロ画像の非可逆符号化の研究が行われている。従来のDCTを用いた手法は高い圧縮率を得ることができたが、ブロック歪みの発生を避けることができなかった。また近年、ブロック歪が生じない手法として注目を浴びているウェーブレット変換は、フィルタ長が長いため、高速な処理が行えなかった。これに対して本章では帯域間予測を用いたハール・ウェーブレット変換を提案することにより、ブロック歪が生じない高速変換手法を実現している。さらに本章ではここで用いられている予測処理がSSKFフィルタと等価であることを証明し、提案手法が完全再構成可能なサブバンド符号化の一つであることを述べている。実験では処理速度、ブロック歪み、圧縮率が従来手法と比較され、改善の効果が示されている。

最後に第6章では各章で行われた研究成果がまとめられている。そして将来の展望と今後の研究課題が述べられている。

# 第1章 緒論

## 1.1 画像データ圧縮の背景

現代の情報化社会において、画像を用いたメディアは大きく注目されつつある。テレビ電話／会議、カラー・ファクシミリ、ビデオディスク、画像データベースなど、画像を用いたメディアは今後大きく発展が期待されている。しかし画像のもつデータ量は膨大であり、その取り扱いが容易ではない。テレビ電話／会議、カラーファクシミリ等の通信メディアにおいては、容量の小さな回線で高速な画像伝送を行うための技術が必要である。またビデオディスク、画像データベース等の蓄積メディアにおいては、限られた記憶容量の中に大量の画像データを保存するための技術が求められている。画像データの圧縮技術はこれらのメディアの発展のために不可欠なものとなっている。

画像データの圧縮には様々な技術が応用されている [1]-[4] が、基本的に圧縮技術は信号変換技術と符号化技術の二つの技術から成る。図 1.1 に画像データ圧縮の基本プロセスを示す。前段の信号変換器は画像信号をより冗長性の少ない信号に変換する役目がある。この代表的な手法としては Vector Quantization (VQ) [5]-[7], Discrete Cosine Transformation (DCT) [8], [9], Subband Coding (SBC) [10]-[14], および予測符号化 (DPCM) [15]-[18] などがある。一方、後段の符号化器は変換された信号から符号語を生成する役目がある。符号化の代表的な手法としてはハフマン符号化 [19]-[21], 算術符号化 [22]-[25], および Lempel-Ziv 符号化 [26] などがある。

VQ は高い冗長抑圧が可能であるため、1980年代から注目を浴びてきた。画像信号に対しては予測符号化と組み合わせた Predictive VQ (PVQ) が有効であることが報告されている [7]。しかし VQ は計算量が膨大であるため、高速アルゴリズムの開発が課題とされている。DCT は VQ に比べると簡単な処理で高い冗長抑圧が可能なることから、1991年に Joint Photographic Expert Group (JPEG) と呼ばれる国際機関が標準案として勧

告した[27], [28]. しかしブロック歪みやモスキート雑音が発生することが問題とされている.

これに対して Quadrature Mirror Filter (QMF) [29]-[33] や Symmetric Short Kernel Filter (SSKF) [34] を用いた SBC はブロック歪みの発生を抑えることができるため, DCT に代わる手法として数多くの研究が行われている [10]-[14]. 近年の SBC においては予測符号化と組み合わせた手法 [35]-[37] が注目されている. 同じく 1980 年代から注目を浴びていたウェーブレット解析 [38]-[40] は SBC の一つとして体系づけられ, 近年は SBC と共に画像データ圧縮への応用が盛んに行われている [41]-[43]. 予測符号化[15]-[18] は VQ, DCT および SBC ほど高い圧縮率が期待できないが, 完全に歪みのない画像を再生できるため, 可逆符号化法としての価値が大きい. JPEG は高品質が要求される場合の可逆符号化モードとして予測符号化を提案している [44].

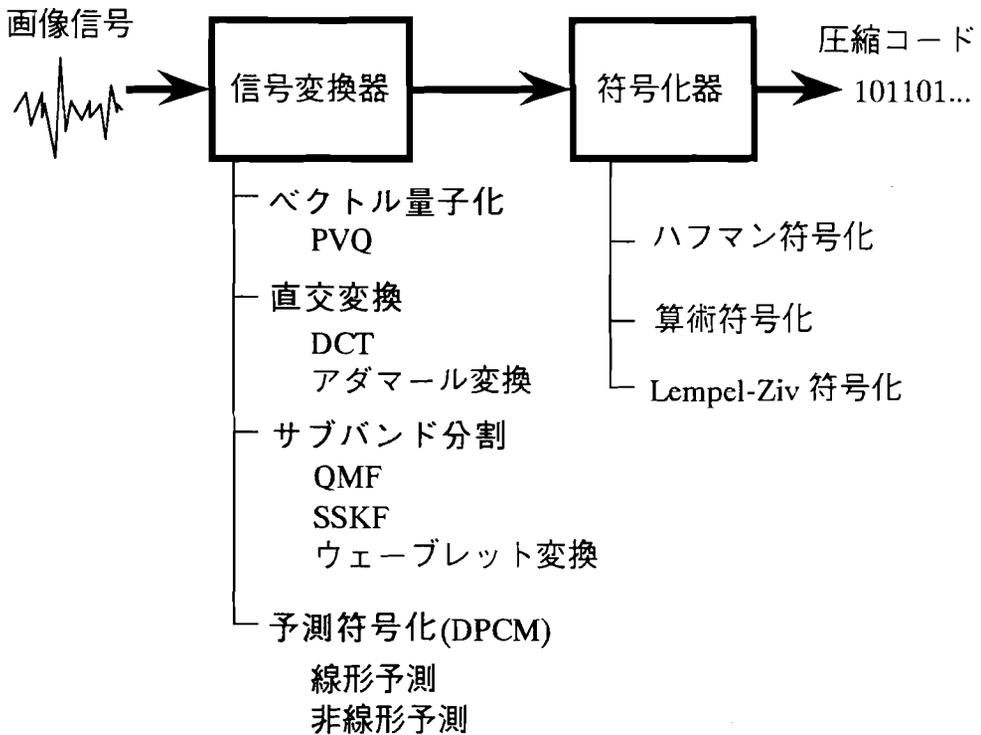


図 1.1 圧縮のプロセス

一方、後段の符号化手法としてはハフマン符号化および算術符号化が主流である。ハフマン符号化は原理が簡単で高速であるため広く用いられている。しかしシンボルの発生確率に合わせて符号化テーブルを構築しなければならないので、動的符号化には適していない。これに対して算術符号化はシンボルの発生確率がわかれば即座に符号語を生成できるので、動的符号化に適している。JPEG は符号化方式としてハフマン符号化と算術符号化の両方のモードを用意している [27], [28]。Lempel-Ziv 符号化は UNIX の "compress" コマンドで広く知られているが、テキスト情報をモデルとして設計されているため、画像信号に対する効果は期待されていない。

近年は B-ISDN の通信体制が整いつつあるため、圧縮率のみならず、処理速度も重視されている。JPEG は標準案の作成において 64 kb/s の ISDN におけるリアルタイム復号化能力を必要条件としている [27]。また復号化の際にプログレッシブビルドアップが可能であるかどうかも重要視されている。このような状況の中で、今後はマルチメディアに柔軟に対応できる符号化設計が重点課題とされている。

## 1.2 本研究の目的

本研究は従来の画像データ圧縮手法において指摘されている様々な問題点を改善し、

- (1) 圧縮の効率を向上させること
- (2) 画像の品質を向上させること

を目的とする。これらの目的は画像蓄積メディアの高密度化、および通信メディアの高速化につながるものである。

本論文ではこれらの目的のもとに、四つの研究を行っている。それぞれの研究は第 2 章から第 5 章に対応している。第 2 章の研究はモノクロ画像の可逆符号化システムにおいて、信号変換器の冗長抑圧能力を高めることを目的としている。第 3 章の研究はカラー画像の可逆符号化システムにおいて、信号変換器の冗長抑圧能力を高めることを目的としてい

#### 4 第1章 緒論

る。第4章の研究はモノクロ画像の非可逆符号化システムにおいて、信号変換器の冗長抑圧能力を高めることを目的としている。第5章の研究は符号化手法の圧縮能力を高めることを目的としている。

# 第1章 緒 論

## 1.1 画像データ圧縮の背景

現代の情報化社会において、画像を用いたメディアは大きく注目されつつある。テレビ電話／会議、カラー・ファクシミリ、ビデオディスク、画像データベースなど、画像を用いたメディアは今後大きく発展が期待されている。しかし画像のもつデータ量は膨大であり、その取り扱いが容易ではない。テレビ電話／会議、カラーファクシミリ等の通信メディアにおいては、容量の小さな回線で高速な画像伝送を行うための技術が必要である。またビデオディスク、画像データベース等の蓄積メディアにおいては、限られた記憶容量の中に大量の画像データを保存するための技術が求められている。画像データの圧縮技術はこれらのメディアの発展のために不可欠なものとなっている。

画像データの圧縮には様々な技術が応用されている [1]-[4] が、基本的に圧縮技術は信号変換技術と符号化技術の二つの技術から成る。図 1.1 に画像データ圧縮の基本プロセスを示す。前段の信号変換器は画像信号をより冗長性の少ない信号に変換する役目がある。この代表的な手法としては Vector Quantization (VQ) [5]-[7], Discrete Cosine Transformation (DCT) [8], [9], Subband Coding (SBC) [10]-[14], および予測符号化 (DPCM) [15]-[18] などがある。一方、後段の符号化器は変換された信号から符号語を生成する役目がある。符号化の代表的な手法としてはハフマン符号化 [19]-[21], 算術符号化 [22]-[25], および Lempel-Ziv 符号化 [26] などがある。

VQ は高い冗長抑圧が可能であるため、1980年代から注目を浴びてきた。画像信号に対しては予測符号化と組み合わせた Predictive VQ (PVQ) が有効であることが報告されている [7]。しかし VQ は計算量が膨大であるため、高速アルゴリズムの開発が課題とされている。DCT は VQ に比べると簡単な処理で高い冗長抑圧が可能なることから、1991年に Joint Photographic Expert Group (JPEG) と呼ばれる国際機関が標準案として勧

告した[27], [28]. しかしブロック歪みやモスキート雑音が発生することが問題とされている.

これに対して Quadrature Mirror Filter (QMF) [29]-[33] や Symmetric Short Kernel Filter (SSKF) [34] を用いた SBC はブロック歪みの発生を抑えることができるため, DCT に代わる手法として数多くの研究が行われている [10]-[14]. 近年の SBC においては予測符号化と組み合わせた手法 [35]-[37] が注目されている. 同じく 1980 年代から注目を浴びていたウェーブレット解析 [38]-[40] は SBC の一つとして体系づけられ, 近年は SBC と共に画像データ圧縮への応用が盛んに行われている [41]-[43]. 予測符号化[15]-[18] は VQ, DCT および SBC ほど高い圧縮率が期待できないが, 完全に歪みのない画像を再生できるため, 可逆符号化法としての価値が大きい. JPEG は高品質が要求される場合の可逆符号化モードとして予測符号化を提案している [44].

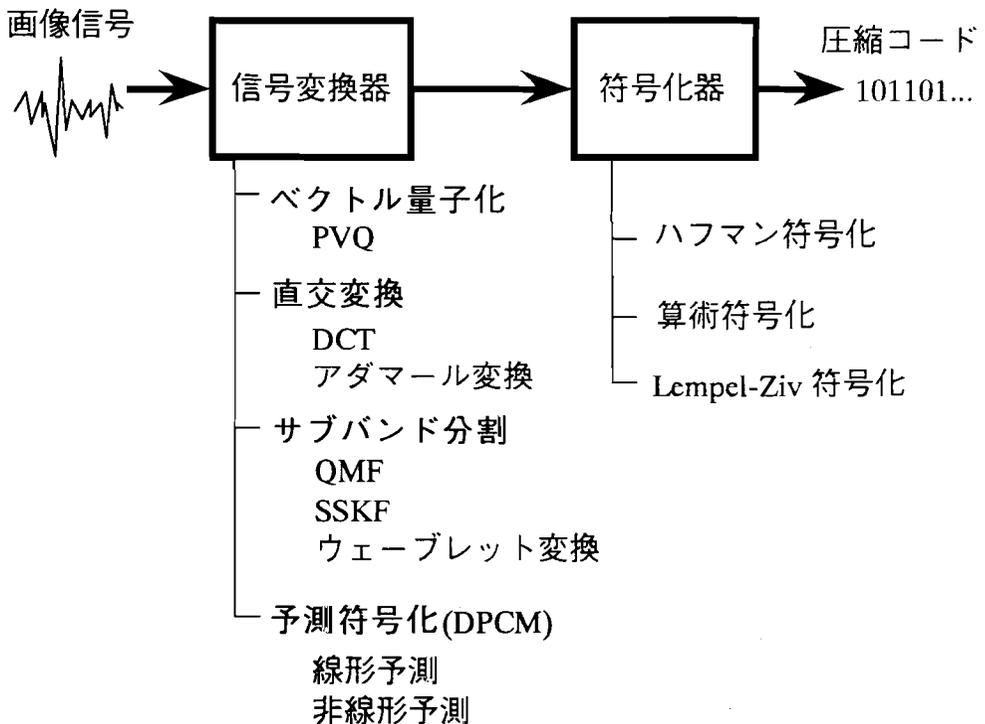


図 1.1 圧縮のプロセス

一方、後段の符号化手法としてはハフマン符号化および算術符号化が主流である。ハフマン符号化は原理が簡単で高速であるため広く用いられている。しかしシンボルの発生確率に合わせて符号化テーブルを構築しなければならないので、動的符号化には適していない。これに対して算術符号化はシンボルの発生確率がわかれば即座に符号語を生成できるので、動的符号化に適している。JPEG は符号化方式としてハフマン符号化と算術符号化の両方のモードを用意している [27], [28]。Lempel-Ziv 符号化は UNIX の "compress" コマンドで広く知られているが、テキスト情報をモデルとして設計されているため、画像信号に対する効果は期待されていない。

近年は B-ISDN の通信体制が整いつつあるため、圧縮率のみならず、処理速度も重視されている。JPEG は標準案の作成において 64 kb/s の ISDN におけるリアルタイム復号化能力を必要条件としている [27]。また復号化の際にプログレッシブビルドアップが可能であるかどうかも重要視されている。このような状況の中で、今後はマルチメディアに柔軟に対応できる符号化設計が重点課題とされている。

## 1.2 本研究の目的

本研究は従来の画像データ圧縮手法において指摘されている様々な問題点を改善し、

- (1) 圧縮の効率を向上させること
- (2) 画像の品質を向上させること

を目的とする。これらの目的は画像蓄積メディアの高密度化、および通信メディアの高速化につながるものである。

本論文ではこれらの目的のもとに、四つの研究を行っている。それぞれの研究は第2章から第5章に対応している。第2章の研究はモノクロ画像の可逆符号化システムにおいて、信号変換器の冗長抑圧能力を高めることを目的としている。第3章の研究はカラー画像の可逆符号化システムにおいて、信号変換器の冗長抑圧能力を高めることを目的としてい

#### 4 第1章 緒 論

る．第4章の研究はモノクロ画像の非可逆符号化システムにおいて，信号変換器の冗長抑圧能力を高めることを目的としている．第5章の研究は符号化手法の圧縮能力を高めることを目的としている．

## 第2章 モノクロ画像の可逆符号化

### 2.1 緒言

可逆符号化とは原データを全く劣化させずに忠実に再構成する符号化法のことであり、高精細画像、出版用画像、医用画像などを扱うアプリケーションにおいて重要である。JPEGはこの可逆符号化法として2次元の予測符号化を提案している。しかしJPEGの手法は冗長抑圧の点では必ずしも最適な手法ではない。本章ではJPEGの手法による冗長抑圧が不十分であることの原因を解明し、これを改善するための手法を提案する。2.2節ではJPEGの予測符号化方式の問題点を明確にする。2.3節および2.4節ではそれらを克服する新しい手法の提案を行う。2.5節では実験を行い、提案手法の有効性を確かめる。

### 2.2 JPEGの手法

本節では始めに2次元予測について説明し、次にJPEG手法の性能を示し、最後にその問題点を指摘する。

#### 2.2.1 2次元予測

予測器の役目は冗長性の多い画像信号を冗長性の少ない誤差信号に変換することである。図2.1に一般的な予測符号化のブロック構成を示す。

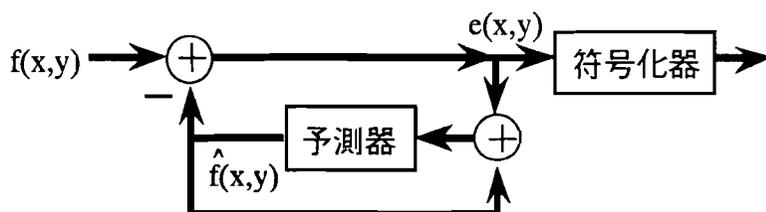


図 2.1 予測符号化システム

本論文では以後、画像の左上を原点としたときの画素の位置を  $(x,y)$  とする。図において  $f(x,y)$  は座標  $(x,y)$  に位置する画素の値（輝度）， $\hat{f}(x,y)$  はその予測値とする。予測器は  $\hat{f}(x,y)$  を算出し，実際に出現した値  $f(x,y)$  との差  $e(x,y)$  を誤差信号として出力する。予測器はできるだけ正確に画像信号を予測し，誤差信号の電力およびエントロピーを下げなければならない。

予測の手法としては2次元線形予測が有効であることが報告されている[15]-[18]。この手法はこれから観測する  $f(x,y)$  の値を，それよりも過去の画素値の線形結合により予測する手法である。過去の画素の中で参照する点の値を  $f_1, f_2, \dots, f_N$  とし，それらに対する予測係数を  $c_1, c_2, \dots, c_N$  とすると予測値  $\hat{f}(x,y)$  は

$$\hat{f}(x,y) = c_0 + \sum_{n=1}^N c_n \cdot f_n \quad (2.1)$$

によって求めることができる。参照点の座標および個数は予測器に依存するが，距離が離れると画素間の相関が急激に低下するので，一般に隣接する4個前後の画素を用いることが多い。

### 2.2.2 JPEG 手法の性能

JPEG は予測手法として，線形予測の一つを用い，Spatial 方式と呼んでいる[27]-[28]。図 2.2 に注目する画素と参照する3画素の配置を示す。この方式は，ある画素の値  $x$  を，それに隣接する3画素の値  $a, b$  および  $c$  の線形結合により予測するものである。予測式としては以下の7種類が用意されている。

$$1: x = a, \quad (2.2)$$

$$2: x = b, \quad (2.3)$$

$$3: x = c, \quad (2.4)$$

$$4: x = a+b-c, \quad (2.5)$$

$$5: x = a+(b-c)/2, \quad (2.6)$$

$$6: x = b+(a-c)/2, \quad (2.7)$$

$$7: x = (a+b)/2. \quad (2.8)$$

予測器はこれらの中の一つを用いて予測値を算出し、実際に出現した値との誤差を出力する。これを繰り返すことにより画像信号を誤差信号に変換する。

この予測手法の性能を評価するために、図 2.3 に示した 4 枚の画像を予測誤差信号に変換した。それぞれの画像は  $256 \times 256$  画素、輝度は 8 ビット、すなわち 256 階調に量子化されている。

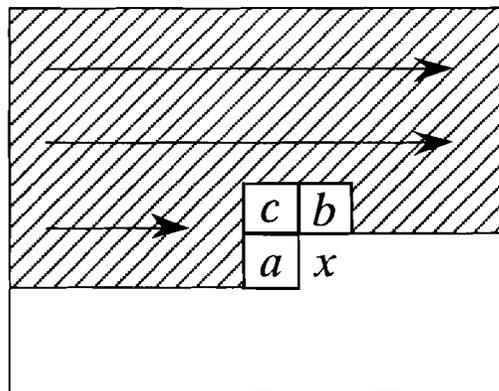


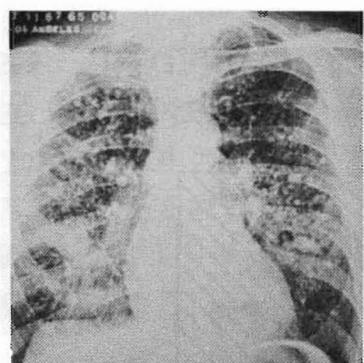
図2.2 注目画素と参照画素の配置



(a) Girl



(b) Couple



(c) Chest



(d) Moon

図 2.3 テスト画像

表2.1 予測誤差信号のエントロピー

予測式 \ 画像	Girl	Couple	Chest	Moon
	$\begin{pmatrix} 256 \\ \times \\ 256 \\ \times \\ 8 \text{ bit} \end{pmatrix}$	$\begin{pmatrix} 256 \\ \times \\ 256 \\ \times \\ 8 \text{ bit} \end{pmatrix}$	$\begin{pmatrix} 256 \\ \times \\ 256 \\ \times \\ 8 \text{ bit} \end{pmatrix}$	$\begin{pmatrix} 256 \\ \times \\ 256 \\ \times \\ 8 \text{ bit} \end{pmatrix}$
原画像のエントロピー	6.41	6.14	5.22	6.71
1: $x = a$	5.06	4.83	5.21	5.54
2: $x = b$	5.11	4.69	5.18	5.26
3: $x = c$	5.42	5.30	5.46	5.58
4: $x = a+b-c$	5.05	4.52	5.87	5.58
5: $x = a+(b-c)/2$	4.89	4.54	6.41	5.38
6: $x = b+(a-c)/2$	4.93	4.46	6.41	5.30
7: $x = (a+b)/2$	4.83	4.60	5.99	5.11

7種類の予測式によって算出された予測誤差信号のエントロピーを表2.1に示す。表2.1におけるエントロピーは、予測誤差の情報源が無記憶情報源であることを仮定した上で算出されたものであり、近傍データとの相関を考慮していない。したがって、この値が必ずしも圧縮率の限界につながるものではないが、少なくともハフマン符号化や算術符号化を用いて、この値に限りなく近い平均符号長で符号化することができる[20]。ゆえに、この値の低さは予測の効果を示す尺度として見ることができる。

表2.1から、平均的には予測誤差のエントロピーが原画像のエントロピーよりも減少していることがわかる。つまり、予測によって原画像はより圧縮され易い形に換えられていることがわかる。特に7番の予測式は効果的であるといえることができる。

### 2.2.3 問題点

上の手法には二つの問題点が見られる。第一に、画像の性質が局所的

に変化する場合、これに予測式を順応させる手段がないという点である。第二の問題点は、表 2.1 の "Chest" に見られるように、予測によって逆にエントロピーが増加する可能性があるということである。

第一の問題を解決するためには、局所的に予測関数を最適化するような手法を開発しなければならない。第二の問題については、まずその原因を説明する。図 2.4 はこれら 4 枚の画像のヒストグラムである。横軸は画素の値であり、縦軸はその値の出現頻度である。"Chest" は他の画像に比べて、かなり不規則な量子化が行われていることがわかる。そのため、予測式が複雑になるほど、それによって得られる予測誤差のシンボル数が増加する可能性がある。エントロピーの増加はこれが原因と考えられる。これを防ぐためには存在しないシンボルを予測の対象外にしなければならない。本論文では以上の二つの問題を改善するための手法を検討する。

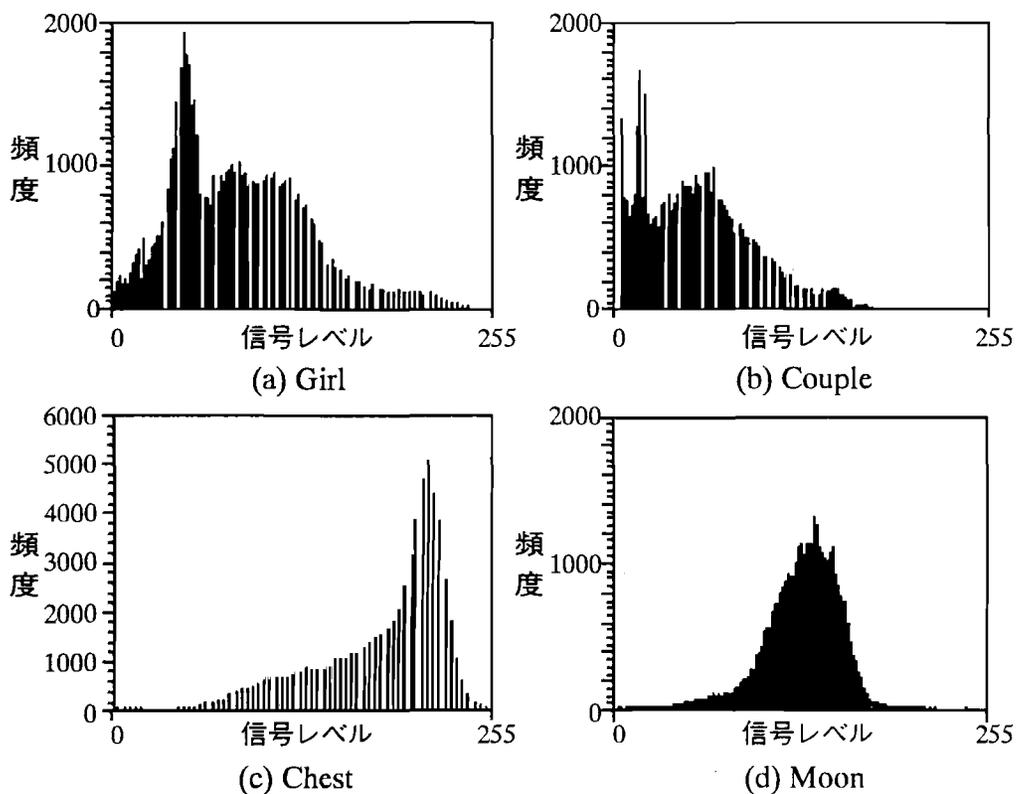


図 2.4 テスト画像のヒストグラム

## 2.3 ダイナミック・レンジ変換法 [45]

本節では予測誤差信号のエントロピーの増加を防ぐ手法として、ダイナミック・レンジ変換法を提案する。

本処理は存在しない信号レベルを削除することにより予測精度を高め、誤差信号のシンボル数を低減するものである。ダイナミック・レンジ変換器は図 2.5(a) に示すように、原画像中に存在する信号レベルに対して、小さいものから順番に番号の割付けを行い、その番号を新たな信号レベルとして出力する。これによって隣り合う信号レベル間の差は 1 に統一されるので、誤差信号も必ず隙間のない整数値をとるようになる。本処理により、誤差信号のシンボル数が無駄に増加するのを防ぐことができる。

変換された信号レベルを元の信号レベルに逆変換するためには、元の信号分布を記憶しておかなければならない。そのため、信号が取り得る 256 のそれぞれの値について、実際に値が存在するか否かを 1 または 0 で記憶する方法をとる。この記憶には 32 バイトがあれば十分である。図

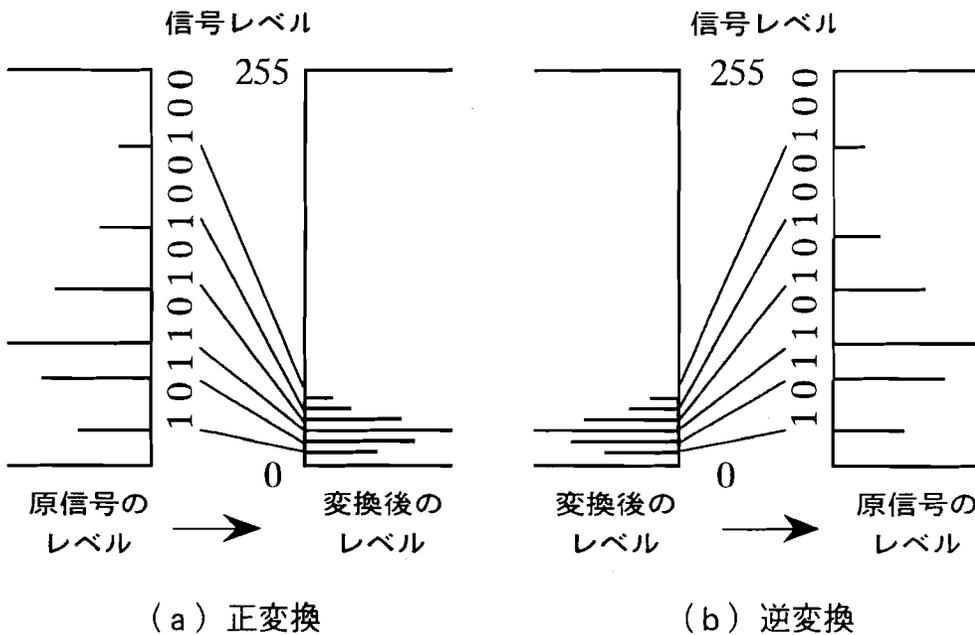


図 2.5 ダイナミック・レンジ変換

2.5(b) に逆変換の様子を示す。逆変換器では 32 バイト (256 ビット) のビット系列を受け取る。逆変換器はこれによって原信号に存在した値を知ることができる。逆変換器はこれを基に信号の変換前の分布範囲と、変換後の分布範囲の対応表を作成するが、このときに信号の大小関係が入れ替わることはないので、逆変換は一意に決めることができる。したがって原画像を完全に復元することができる。

## 2.4 動的予測 [45]

一般的に画像は性質の異なる領域の組み合わせにより表現されているので、一つの予測式では局所的に適合しなくなる可能性がある。領域による画像の性質の変化に対応するためには、予測係数を局所領域の性質に応じて、すなわち、動的に算出しなければならない。そこで本節では予測係数の動的算出を行う予測符号化方式を提案する。

### 2.4.1 アルゴリズム

提案手法では JPEG の Spacial 方式と同様に、 $f(x,y)$  の値を図 2.2 に示す 3 点の線形結合により予測する。ただし三つの画素値に対する係数および直流係数 (DC 係数) は可変であり、予測関数は

$$\hat{f}(x,y) = c_0 + c_1 \cdot f(x-1,y) + c_2 \cdot f(x,y-1) + c_3 \cdot f(x-1,y-1) \quad (2.9)$$

となる。ここで予測係数  $c_k$  ( $k=0, 1, \dots, 3$ ) は 1 画素ごとに更新する。予測係数の算出の際には注目している画素の近傍領域のデータを利用する。この動的予測のアルゴリズムを図 2.6 に示す。以下に図中の各手順について説明する。

手順 1) まず注目している画素の近傍領域を観測する。画素  $(x,y)$  の近傍領域を  $W(x,y)$  と定義する。 $d$  はこの領域の大きさをきめるパラメータである。ただし、 $f(x,y)$  の予測を行う時点では画素  $(x,y)$  よりも先の値は得られていないため、 $W(x,y)$  は画素  $(x,y)$  を含んでいない。

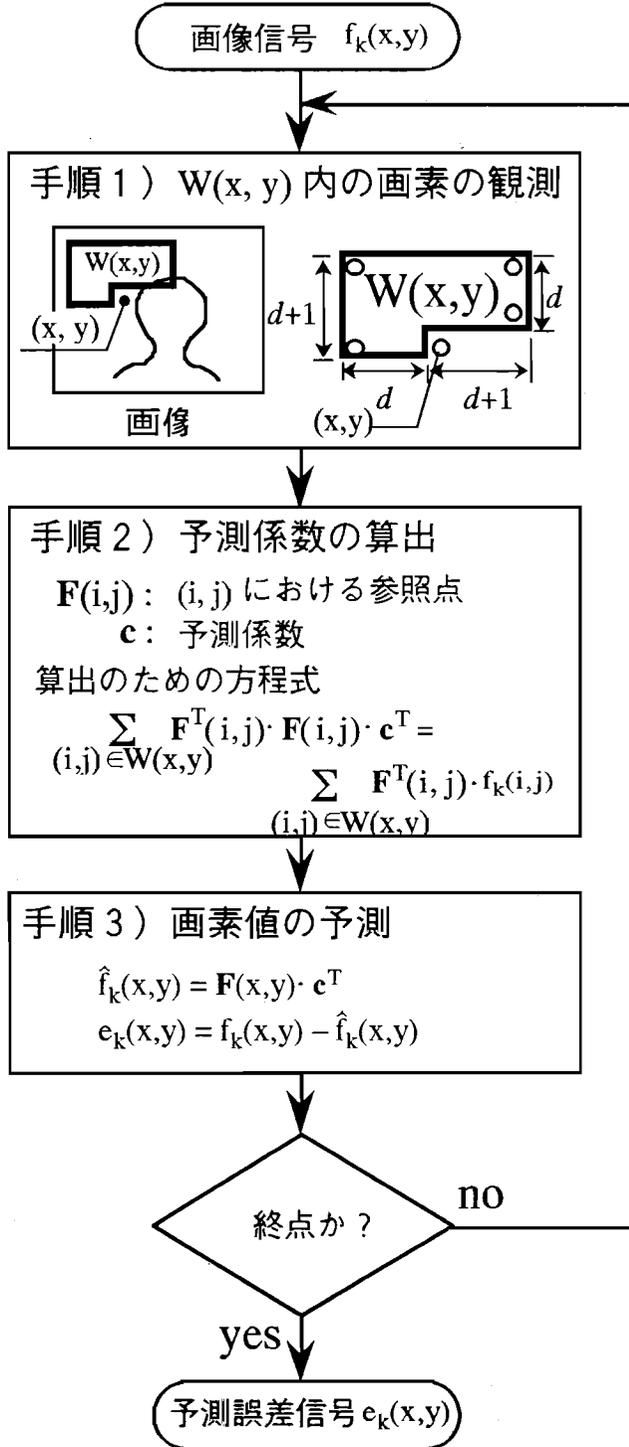


図 2.6 動的予測アルゴリズム

手順2) 領域  $W(x,y)$  内の画素に対して最小二乗法 [46] を適用することにより、この領域内での予測誤差の電力が最小となるような予測係数を求める。厳密には電力を最小にすることとエントロピーを最小にすることは異なるが、最小二乗法はエントロピーをほぼ最小にするための有効な手段である。最小二乗法の演算については付録に示している。

手順3) 手順2) で求められた予測係数を用いて  $f(x,y)$  を予測する。そして予測値と実際の出現値との差を誤差信号として出力する。

以上のステップを 1 画素ごとに繰り返し、全ての画像信号を予測誤差信号に変換する。

局所的に最適な予測係数を算出するためにはパラメータ  $d$  を小さく選ぶ必要がある。しかし領域  $W(x,y)$  内のサンプル数が極端に少ない場合、最小二乗法による解が一意に定まらなくなり、予測が不安定になる可能性がある。そのため  $d$  を極限まで小さくすることはできない。パラメータ  $d$  の大きさについては 2.5 節で議論する。

### 2.4.2 高速演算手法

本手法は、1 画素ごとに予測係数の算出を行うため、付録の式 (A5) を繰り返し解かなければならない。このとき、各項には多くの冗長な演算が含まれている。ここでは演算上の冗長性を取り除くための手法を提案する。

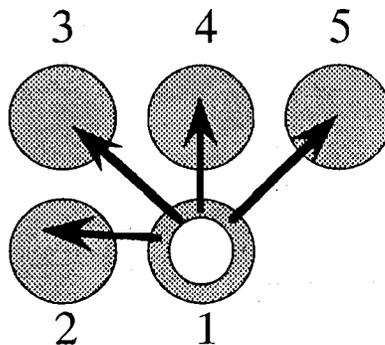


図 2.7 五つの積の方向

まず乗算演算の冗長性を取り除く。参照点が3点のとき、式(A5)の行列  $F(i,j) \cdot F(i,j)$  は、の16個の要素をもち、ベクトル  $F(i,j) \cdot f(i,j)$  は4個の要素をもつ。これらの20個の要素のうち13個は二つの画素値の積である。これらを  $(x,y)$  平面上で見たとき、その積の方向によって分類すると、図2.7に示した5種類に限られる。したがって13個の要素を独立に算出することは冗長である。この冗長性を除去するために、5方向の画素値の積  $R_k(x,y)$  を

$$\begin{aligned}
 R_1(i,j) &= f(i,j) \cdot f(i,j), \\
 R_2(i,j) &= f(i,j) \cdot f(i-1,j), \\
 R_3(i,j) &= f(i,j) \cdot f(i-1,j-1), \\
 R_4(i,j) &= f(i,j) \cdot f(i,j-1), \\
 R_5(i,j) &= f(i,j) \cdot f(i+1,j-1)
 \end{aligned} \tag{2.10}$$

のように定義する。そして  $R_k(i,j)$  を事前に一度だけ算出し、記憶しておく。そして式(A5)の全ての要素を  $R_k(i,j)$  の加算だけで求める。これによって1画素あたりの乗算回数を5回にすることができる。

次に  $R_k(i,j)$  の総和演算の冗長性を除去する。  $R_k(i,j)$  の積算領域  $W(x,y)$  はたかだか1画素ずつ移動するので、その度に総和を求めていたのでは膨大な量の加算演算が重複していることになる。ここでは、2次元平面上の任意の位置にある  $W(x,y)$  内の総和を、わずか6項の加減算により求める手法を提案する。

まず2次元配列上に  $R_k(i,j)$  を準備し、図2.8の手順で2次元累積データ  $S_k(i,j)$  を作成する。このとき2次元配列上に生成されたデータ  $S_k(i,j)$  は  $R_k(i,j)$  を用いて

$$S_k(i,j) = \sum_{n=1}^i \sum_{m=1}^j R_k(n,m) \tag{2.11}$$

のように表すことができる。また逆に、  $R_k(i,j)$  の総和は  $S_k(i,j)$  を用いて

$$\begin{aligned}
 \sum_{i=\text{left}}^{\text{right}} \sum_{j=\text{top}}^{\text{bottom}} R_k(i, j) &= S_k(\text{right}, \text{bottom}) \\
 &\quad - S_k(\text{left}-1, \text{bottom}) \\
 &\quad - S_k(\text{right}, \text{top}-1) \\
 &\quad + S_k(\text{left}-1, \text{top}-1)
 \end{aligned}
 \tag{2.12}$$

のように表すことができる。上の式からわかるように、 $R_k$  平面上の任意の長方形領域内の総和は、 $S_k$  平面上の 4 点の加減算によって求めることができる。  $R_k$  平面上の長方形領域と  $S_k$  平面上の 4 点の関係は図 2.9(a) のようになる。  $S_k$  平面上の 4 点は  $R_k$  平面上の長方形領域の頂点に対応する。積算領域が長方形でない場合は、その領域を長方形に分解して考える。例えば  $W(x,y)$  は図 2.9(b) のように二つの長方形に分けることができるため、 $W(x,y)$  内の積算値は式 (2.11) を二度用いることによって算出できる。この値は図 2.9(b) のように二つの長方形領域の頂点の値の加減算によって表すことができる。八つの頂点の内、二つの頂点は打ち消し合うので、結局、積算値は六つの頂点の加減算により求めることができる。頂点の個数は領域の大きさが変化しても変わらないので、計算量は領域の大きさによらず一定である。例えば  $d=10$  のとき、 $W(x,y)$  内の総和演算を行うと 220 項の加算が必要であるのに対して、本手法を用いれば 6 項の加減算で総和を得ることができる。このように、本手法を用いることによって演算回数を大幅に減らすことができるので、積算処理の速度は飛躍的に向上する。

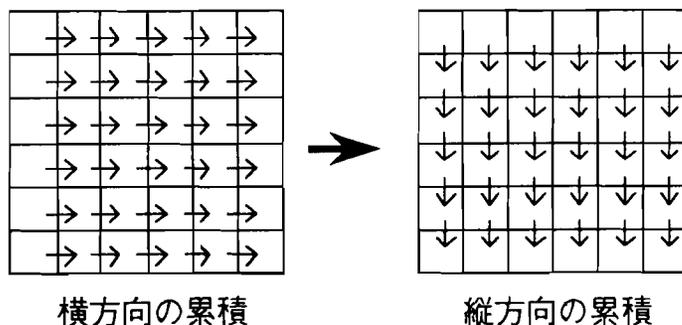


図2.8 2次元配列  $S_k$  の作成

実時間の処理では、一つの画素(x,y)を符号化（複合化）するたびに  $S_k(x,y)$  を算出しなければならないので、図 2.8 に示した累積手順は、すでに符号化（複合化）済みの領域から随時行われ、即座に  $S_k(x,y)$  が算出される。さらに配列  $R_k$  は、 $S_k(x,y)$  を算出するための、中間的に必要な配列にすぎないので、使用済みの領域から解放できる。これをソフトウェアによって実現する場合は、配列を指し示すポインタを巡回させることによって仮想的に画像 1 枚分の記憶領域をとり、実際には上書きによって配列の内容を更新する手段を用いる。

### 2.5 計算機シミュレーション

本章では提案手法を用いた可逆符号化システムを構成し、画像信号のエントロピー低減の効果および圧縮比について考察を行う。

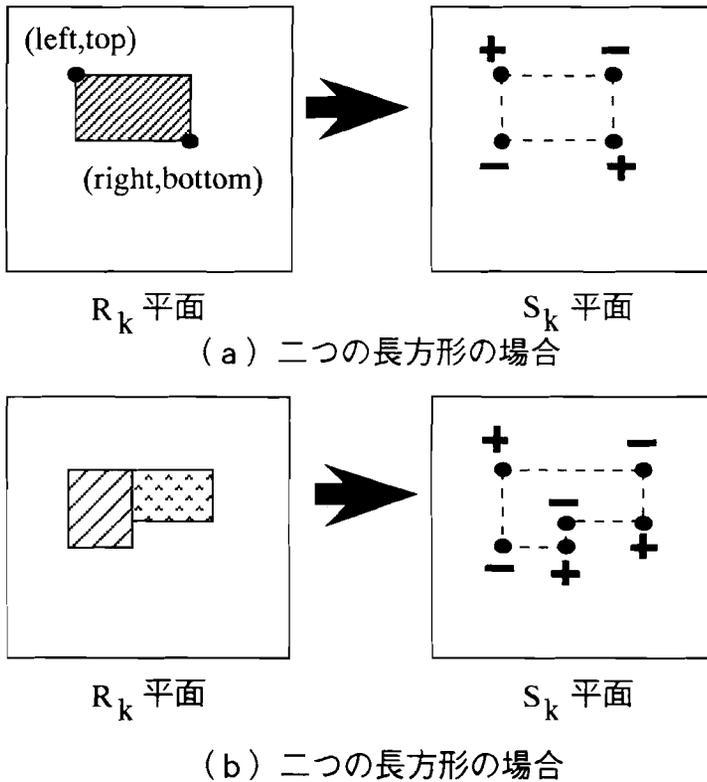
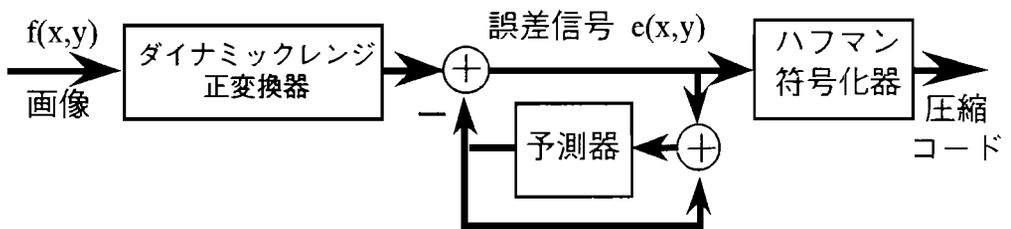


図 2.9  $R_k$  平面上の累積領域と  $S_k$  平面上の加算点との関係

## 2.5.1 システム構成

提案手法を用いた符号化システムのブロック・ダイアグラムを図 2.10(a) に示す。このシステムは、ダイナミック・レンジ変換器、予測器、符号化器の三つの部分から構成されている。第1段階では 2.3 節で提案されたダイナミック・レンジ変換が行われる。変換に関する情報は、ヘッダとして画像信号と共に送信される。第2段階では、2.4 節で提案された動的予測が行われる。画像の外側部分では画素の値が 0 であることを仮定し、注目画素が画像の端の部分であるかどうかにかかわらず、図 2.6 のアルゴリズムを用いる。動的予測では復号化側が符号化側と全く同じアルゴリズムを用いて予測係数を更新するため、予測係数を送信する必要がない。第3段階ではハフマン符号化が行われる。符号化テーブルは圧縮コードと共に受信側である復号化器に送信される。以上の仕様から、圧縮データのフォーマットは図 2.11 のようになる。

次に画像の復号化システムのブロック・ダイアグラムを図 2.10(b) に示す。復号側もまた、符号化側と同じように三つの段階から構成されている。第1段階の復号化器は、まず符号化テーブル情報を受け取る。続いて圧縮コードを予測誤差信号に復元する。第2段階では、符号化システ



(a) 符号化側



(b) 復号化側

図2.10 提案する予測符号化システム

ムで用いた手法と全く同じアルゴリズムを用いて予測係数の更新を行いながら、画素値の予測を行う。そして予測値に予測誤差を加えることによって画素値を復元する。最後の第3段階では、ダイナミック・レンジ逆変換器が変換に関する32バイトのデータを受け取り、圧縮されたダイナミック・レンジを完全に復元する。

### 2.5.2 実験結果

実験には図2.3に示した4枚の画像を用いた。圧縮システムは日本電気株式会社製のパーソナル・コンピュータPC9801上で全てソフトウェアにより実現した。

まず予測係数の局所適応能力を明らかにするために、画像"Couple"に対してJPEGの手法と提案手法を用いた場合の予測誤差の画像を調べ、それぞれ図2.12(a)と(b)に示した。図の画像の明るさが予測誤差の大きさを表している。(b)は(a)に比べ、全体的に予測誤差が少ないことがわかる。特にエッジ部分での予測誤差の減少は顕著である。この様子から予測係数は局所領域に適合していることがわかる。

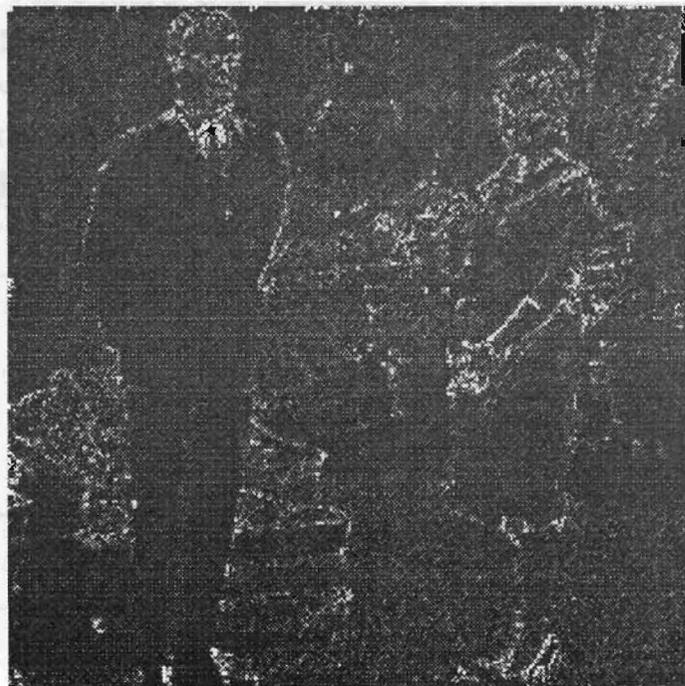
次に予測誤差信号のエントロピーを図2.13に示す。3枚の画像"Girl", "Couple", "Chest"については提案手法のエントロピーがJPEGのどの予測式によるエントロピーよりもかなり小さく、JPEGが示した最も低い値の約80%に抑えられていることがわかる。画像"Moon"の場合はわずか

ダイナミック・レンジ変換情報 (32バイト)
ハフマン符号化テーブル (可変サイズ)
圧縮コード (可変サイズ)

図2.11 保存データのフォーマット



(a) J P E Gの予測式による予測誤差



(b) 提案手法による予測誤差

図2.12 "Couple"の予測誤差信号の比較

な改善しか見られないが、この原因は画像全体が単調で局所的な変化が少ないため、動的手法の効果が殆ど現われなかったためと考えられる。

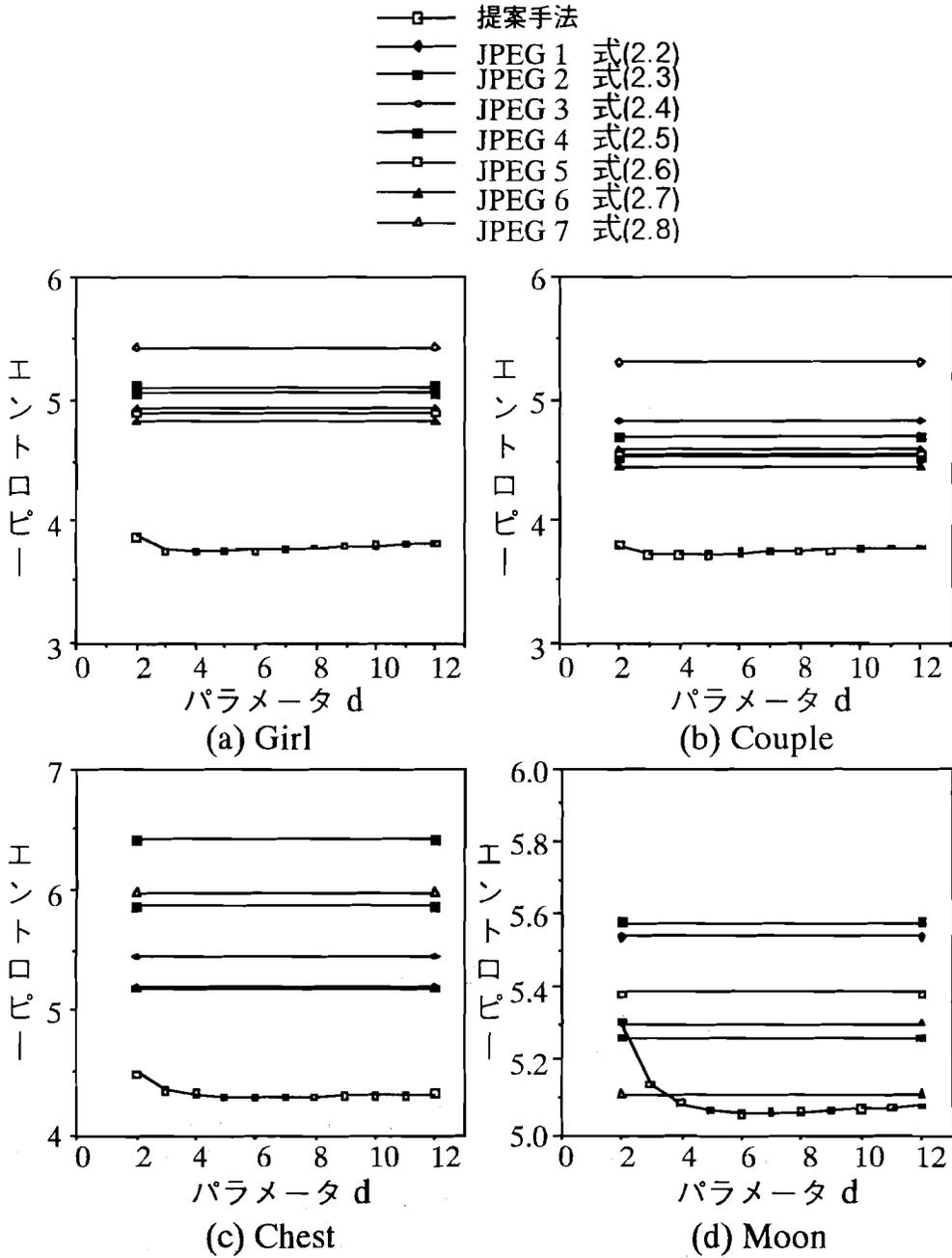


図 2.13 パラメータ d に対する予測誤差信号のエントロピー

表2.2 JPEGの手法による圧縮データサイズ (バイト)

圧縮データ \ 画像	Girl	Couple	Chest	Moon	平均値
符号化テーブル	203	198	223	186	203
圧縮コード	39,550	36,735	42,430	41,570	40,071
合計	39,753	36,933	42,653	41,756	40,274
圧縮比	60.7%	56.4%	65.1%	63.7%	61.5%

表2.3 提案手法 (d=5)による圧縮データサイズ (バイト)

圧縮データ \ 画像	Girl	Couple	Chest	Moon	平均値
ダイナミックレンジ 変換情報	32	32	32	32	32
符号化テーブル	106	114	174	195	147
圧縮コード	30,141	29,898	35,258	41,310	34,152
合計	30,279	30,044	35,464	41,537	34,331
圧縮比	46.2%	45.8%	54.1%	63.4%	52.4%

4枚の画像に共通して見られる提案手法の特長は、 $d=5$ の付近で極小値をもつということである。極小値の右側においてエントロピーが増加する理由は、領域  $W(x,y)$  が過度に大きくなると予測係数が局所的に最適化されないためと考えられる。極小値の左側においてエントロピーが増加する理由は、サンプル点が少なくなりすぎて最小二乗法の解が一意に定まらなくなり、予測が不安定になるためと考えられる。

最後に予測誤差信号をハフマン符号化し、圧縮比を比較する。まずJPEGの7種類の誤差信号の中で最もエントロピーの低い信号を符号化した結果を表2.2に示す。ここで圧縮比とは原画像のデータサイズ 65,536バイトに対する圧縮データの合計サイズの比率とする。4枚の画像の平均では61.5%に達している。提案手法において  $d=5$  に固定した際の圧縮

比を表 2.3 に示す。画像 "Moon" については先に述べたようにエントロピーがあまり低減しなかったため、圧縮比も 0.3 ポイントしか改善されていない。しかし他の 3 枚の画像については 10 ポイント以上圧縮比が改善されていることがわかる。平均の圧縮データサイズは JPEG の結果をさらに 15% 削減する 34,331 バイトに達している。圧縮比は JPEG の結果より 9.1 ポイント低い 52.4% に達している。

## 2.6 結言

本章では、可変係数の 2 次元線形予測法に基づく画像データの可逆符号化を提案し、さらにその高速アルゴリズムを開発した。計算機シミュレーションでは JPEG の手法による圧縮データサイズをさらに 15% 削減することができ、圧縮比は平均で 52.4% に到達した。これらの実験結果から本手法は画像データの可逆符号化法として非常に有効であることが確かめられた。

## 第3章 カラー画像の可逆符号化

### 3.1 緒言

カラー画像の可逆符号化においては、モノクロ画像の場合とことなり、三つのコンポーネント間の冗長性をどのように除去するかが問題となる。本章ではこの冗長性を除去する手法として、3次元予測を用いた予測符号化を提案する。3.2節では従来のコンポーネント間予測について述べ、その問題点を指摘する。3.3節では提案する3次元予測の原理を述べる。さらに3.4節では高速化のための勾配予測を提案する。最後に3.5節では、これらの予測方式を用いて画像の圧縮符号化を行い、その有効性を確認する。

### 3.2 コンポーネント間予測

モノクロ画像を可逆符号化する手法としては、2次元予測符号化が有効であることを、第1章では述べた。この手法をそのままカラー画像に応用することは容易である。JPEGはカラー画像の三つのコンポーネントのそれぞれに対して2次元予測符号化を行っている。しかし三つのコンポーネントが冗長性の多いRGB信号から構成されている場合、2次元予測ではこれらの間の冗長性を除去することができない。しかも可逆符号化の場合は打切り誤差を伴う処理を実行できないので、RGB信号をYUV信号に変換することも不可能である。

Yamaguchi [47] はRGB形式でコンポーネント間の冗長性を取り除くために、コンポーネント間予測を提案した。RGB信号の予測システムを図3.1に示す。図中の $f_R(x,y)$ 、 $f_G(x,y)$ および $f_B(x,y)$ はそれぞれ座標 $(x,y)$ におけるR信号、G信号およびB信号の値とする。それらの予測値は $\hat{f}_k(x,y)$  ( $k=R, G, B$ )で表されている。予測誤差 $e_k(x,y)$  ( $k=R, G, B$ )は

$$e_k(x,y) = f_k(x,y) - \hat{f}_k(x,y) \quad (3.1)$$

のように求める。Yamaguchi のコンポーネント間予測では、まず  $\hat{f}_G(x,y)$  を予測式

$$\hat{f}_G(x,y) = \frac{1}{3} \{f_G(x-1,y) + f_G(x,y-1) + f_G(x+1,y-1)\} \quad (3.2)$$

を用いて求める。その後、復号化した  $f_G(x,y)$  を用いて、 $\hat{f}_k(x,y)$  ( $k = R, B$ ) を予測式

$$\begin{aligned} \hat{f}_k(x,y) = & f_G(x,y) + \frac{1}{3} \{f_k(x-1,y) - f_G(x-1,y)\} \\ & + \frac{1}{3} \{f_k(x,y-1) - f_G(x,y-1)\} \\ & + \frac{1}{3} \{f_k(x+1,y-1) - f_G(x+1,y-1)\} \end{aligned} \quad (3.3)$$

を用いて求める。G信号を最初に予測する理由は、一般的にG-R間、およびG-B間の相互相関の方が、R-B間のそれよりも大きいため、G信号を中心にRおよびB信号を予測する方が誤差が小さくなるからである。

このコンポーネント間予測を用いると2次元予測に比べてよりエントロピーの低い誤差信号を生成することができる。しかし式(3.2)および式(3.3)の予測式は各項に対する係数が1/3に固定されているので、全ての画像に対して有効であるとは限らない。画像信号に適した予測係数につ

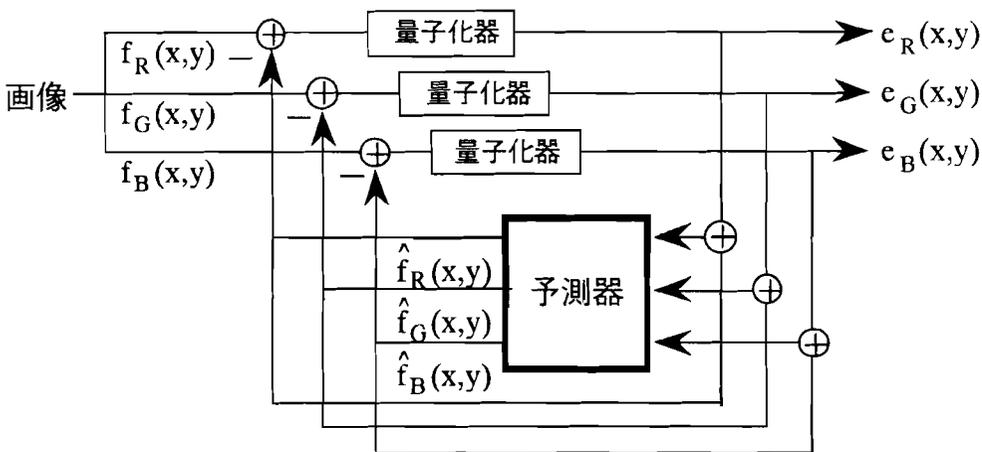


図 3.1 RGB信号の予測システム

いては理論的な根拠に基づいて検討する必要がある。

### 3.3 3次元予測 [48]

RGB形式の画像においては隣接する画素間の冗長性とコンポーネント間の冗長性を同時に除去する必要がある。そこで本章では3次元予測を提案する。

提案する手法は、RGBの各コンポーネントを図3.2のようにz方向に配置し、3次元の信号と見なす。そして注目する点の値をその近傍空間から予測する。本手法も先に述べたコンポーネント間予測と同様に、G・R・Bの順番で予測を行う。G・R・Bの各信号に対する参照点は図3.2のように選ぶ。提案手法は予測式として、

$$\hat{f}_G(x,y) = g_0 + \sum_{n=1}^3 g_n \cdot f_n \quad (3.4)$$

$$\hat{f}_R(x,y) = r_0 + \sum_{n=1}^7 r_n \cdot f_n \quad (3.5)$$

$$\hat{f}_B(x,y) = b_0 + \sum_{n=1}^7 b_n \cdot f_n \quad (3.6)$$

を用いる。この予測式は線形予測の一般形であり予測係数  $r_n$ ,  $g_n$  および  $b_n$  が可変であるため、2次元予測とコンポーネント間予測の両方の性質が実現できる。提案手法ではこれらの予測係数を前章と同様に図2.6の

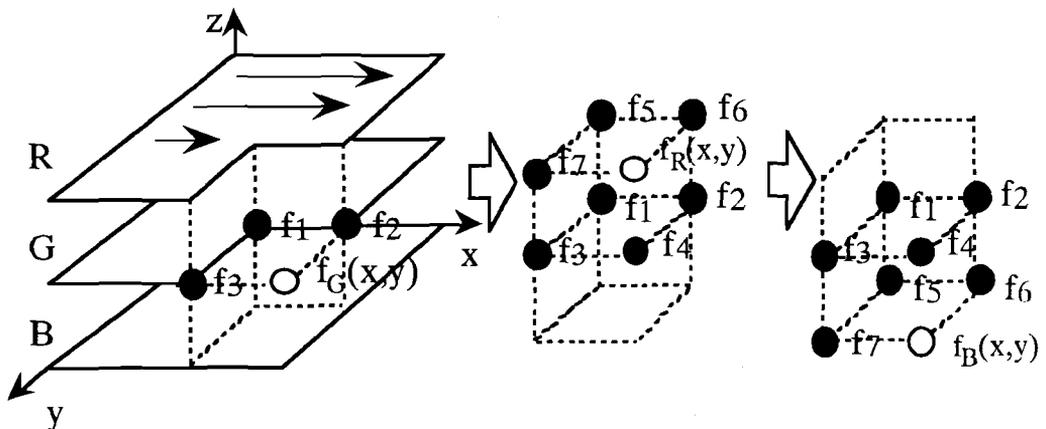


図 3.2 予測順序と参照点

アルゴリズム用いて算出する。また  $d$  も前章と同様にサンプル領域の大きさを表すパラメータとして用いる。

### 3.4 勾配予測方式 [49]

本節では 3次元予測に近い性能をもつもう一つの予測式を提案する。

3.3 節で提案した予測式 (3.5) および式 (3.6) は、第 2 章で提案した予測式 (2.9) に比べて 2 倍の予測係数をもつため、最小二乗法の演算量がほぼ 4 倍になる。また予測係数の数に対するサンプル数の比率が 2 分の 1 に減少するので最小二乗法の解が一意に定まらなくなる可能性がある。そこで本節では予測係数が少ない勾配予測を提案する。

本手法は信号の勾配に注目した予測式を用いる。図 3.3 において矢印はその終点の画素値から始点の画素値を引いた差分、すなわち 2 点間の勾配を表すとする。本手法は予測すべき画素への勾配  $d_G(x,y)$ ,  $d_R(x,y)$ , および  $d_B(x,y)$  を、近傍の勾配の線形結合

$$\hat{d}_G(x,y) = g_1 \cdot d_1 + g_2 \cdot d_2, \tag{3.7}$$

$$\hat{d}_R(x,y) = r_1 \cdot d_1 + r_2 \cdot d_2 + r_3 \cdot d_3 + r_4 \cdot d_4, \tag{3.8}$$

$$\hat{d}_B(x,y) = b_1 \cdot d_1 + b_2 \cdot d_2 + b_3 \cdot d_3 + b_4 \cdot d_4 \tag{3.9}$$

により予測する。 $g_n, r_n$  および  $b_n$  は勾配に対する予測係数である。画素

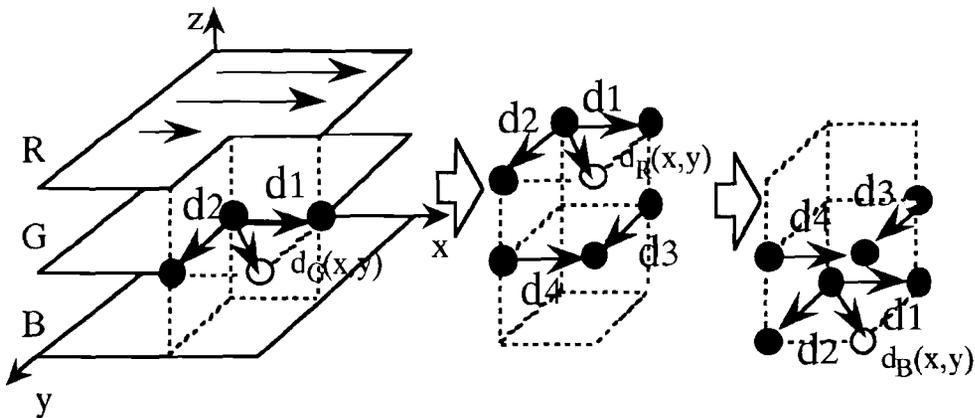


図 3.3 勾配予測の順序と参照点

表3.1 5種類の予測手法の比較

手法	予測係数の個数			画素毎に係数を更新	G信号からRおよびBを予測
	R	G	B		
(I) JPEGの予測	0	0	0	No	No
(II) コンポーネント間予測	0	0	0	No	Yes
(III) 2次元予測	4	4	4	Yes	No
(IV) 3次元予測	8	4	8	Yes	Yes
(V) 勾配予測	4	2	4	Yes	Yes

の予測値  $\hat{f}_k(x,y)$  ( $k=R, G, B$ ) は勾配の予測値  $\hat{d}_k(x,y)$  から間接的に

$$\hat{f}_k(x,y) = \hat{d}_k(x,y) + f_k(x-1,y-1) \quad (3.10)$$

によって求める。本予測方式も基本的に線形予測であるが、各画素に対してではなく各勾配に対して予測係数が決定される点が3次元予測と異なる。

### 3.5 計算機シミュレーション

本節では表3.1に示した5種類の予測器の性能を評価する。テストに用いた画像を図3.4に示す。RGBの各コンポーネントは256×256×8ビットのデータをもっている。

#### 3.5.1 予測係数

まず9枚の画像の合計画素数、約60万サンプルから求めた予測係数の平均値を図3.5に示す。2次元予測および3次元予測を用いた場合の予測係数には次のような性質が見られる。



No.1



No.2



No.3



No.4



No.5



No.6



No.7



No.8



No.9

図3.4 テスト画像

- i) 画像信号が 256 値をもつにもかかわらず，DC 係数  $C_0$  は 2 以下の小さい値をとる。
- ii) 隣接する二つの予測係数の符号が反転している。

このことから，ある信号の値はその近傍の平均値よりも，むしろ近傍の勾配に依存していることがわかる。また R 信号および B 信号の 3 次元予測においては，同一座標の G 信号に対する係数が最も大きいことから，G 信号からの予測が非常に重要であることがわかる。

一方，勾配予測の予測係数は全て正の値であるので，各コンポーネントの勾配が予測値にどのように影響しているか容易に理解できる。これらの係数から， $dR(x,y)$  および  $dB(x,y)$  は同一コンポーネント内の信号勾配からだけでなく，G コンポーネントの信号勾配からも大きく影響を受けていることがわかる。

	G 成分	R 成分	B 成分
2 次元 予測	<p>-0.360 ● --- ● 0.642 0.703 ● ○ <math>C_0 = 1.285</math></p>	<p>-0.392 ● --- ● 0.645 0.733 ● ○ <math>C_0 = 1.364</math></p>	<p>-0.280 ● --- ● 0.615 0.645 ● ○ <math>C_0 = 1.341</math></p>
3 次元 予測	<p>R G -0.392 ● --- ● 0.645 B 0.733 ● ○ <math>C_0 = 1.364</math></p>	<p>-0.246 ● --- ● 0.527 0.655 ● ○ 0.137 ● -0.383 ● -0.440 ● --- ● 0.757 <math>C_0 = 0.286</math></p>	<p>0.014 ● --- ● -0.257 -0.310 ● --- ● 0.647 -0.105 ● --- ● 0.445 0.553 ● ○ <math>C_0 = 1.354</math></p>
勾配 予測	<p>R G 0.658 B 0.753</p>	<p>0.418 0.479 0.407 0.462</p>	<p>0.269 0.566 0.324 0.599</p>

図 3.5 最小二乗法によって算出された予測係数

表 3.2 予測誤差信号のエントロピー

予測法	画像	1	2	3	4	5	6	7	8	9	平均	合計
		JPEG 1 番	R	5.394	5.334	5.887	6.428	6.111	6.699	6.588	5.754	5.063
G	5.082	5.105	5.888	6.372	5.920	6.654	5.585	5.686	5.333	5.736		
B	5.396	5.231	6.018	6.302	5.623	6.486	5.802	5.532	5.025	5.713		
JPEG 2 番	R	5.102	4.747	5.486	5.685	5.814	6.506	6.590	5.463	4.722	5.568	16.29
G	4.748	4.492	5.601	5.386	5.631	6.461	5.480	5.377	4.988	5.352		
B	5.203	4.773	5.738	5.336	5.318	6.290	5.745	5.219	4.727	5.372		
JPEG 3 番	R	5.394	5.334	5.887	6.428	6.111	6.699	6.588	5.754	5.063	5.918	17.37
G	5.082	5.105	5.888	6.372	5.920	6.654	5.585	5.686	5.333	5.736		
B	5.396	5.231	6.018	6.302	5.623	6.486	5.802	5.532	5.025	5.713		
JPEG 4 番	R	5.077	4.577	5.377	5.705	5.552	6.331	7.147	5.125	4.469	5.484	16.26
G	4.825	4.387	5.528	5.475	5.565	6.364	5.785	5.243	4.741	5.324		
B	5.440	4.913	5.640	5.474	5.382	6.258	6.104	5.206	4.645	5.451		
JPEG 5 番	R	4.898	4.577	5.194	5.788	5.384	6.065	6.880	4.931	4.329	5.338	15.82
G	4.630	4.381	5.283	5.744	5.366	6.088	5.595	5.025	4.602	5.190		
B	5.160	4.752	5.422	5.710	5.172	5.984	5.962	4.979	4.453	5.288		
JPEG 6 番	R	4.931	4.504	5.292	5.524	5.644	6.251	6.898	5.129	4.392	5.396	15.88
G	4.619	4.276	5.434	5.294	5.496	6.253	5.603	5.145	4.659	5.198		
B	5.170	4.678	5.540	5.253	5.245	6.131	5.970	5.063	4.497	5.283		
JPEG 7 番	R	4.824	4.626	5.229	5.729	5.602	6.108	6.564	5.077	4.381	5.349	15.72
G	4.515	4.411	5.301	5.698	5.404	6.095	5.395	5.063	4.641	5.169		
B	4.948	4.638	5.413	5.617	5.124	5.972	5.739	4.961	4.416	5.203		
コンポーネント間予測	R	4.731	4.189	5.11	5.752	5.698	4.825	6.101	4.551	3.959	4.991	15.27
G	4.496	4.437	5.334	5.809	5.477	6.175	5.433	5.136	4.711	5.223		
B	4.642	4.169	5.731	4.766	5.062	5.207	5.661	5.376	4.871	5.054		

### 3.5.2 予測誤差信号のエントロピー

JPEG の予測およびコンポーネント間予測を用いた場合の予測誤差信号のエントロピーを表 3.2 に示す。JPEG の 7 種類の式の中では、7 番の式が平均的に低いエントロピーに達することがわかる。JPEG の手法とコンポーネント間予測を比較すると、G 信号のエントロピーはコンポーネント間予測よりも JPEG の 5 番、6 番および 7 番の方が平均的に低い値に達している。しかし R 信号および B 信号のエントロピーはコンポーネント間予測の方が低い値に達しているため、合計ではコンポーネント間予測が最も低い 15.27 に達している。このことから G 信号を用いた予測がかなり効果的であることがわかる。

次に 2 次元予測、3 次元予測および勾配予測による予測誤差信号のエ

ントロピーを図 3.6 に示す。グラフの横軸はパラメータ  $d$ 、縦軸は RGB の各予測誤差信号のエントロピーの合計である。グラフ上の全ての点は 9 枚の画像に対する平均値を用いてプロットされている。グラフにはこれと同時に表 3.2 の JPEG の 7 番の予測式に対する合計エントロピーおよびコンポーネント間予測を用いた場合の合計エントロピーが水平線で示されている。

2次元予測では最小二乗法によって予測係数を求めるため、JPEG の予測に比べて低いエントロピーに達している。しかしその値はコンポーネント間予測によるエントロピーをほとんど下回ることができない。この理由は 2次元予測がコンポーネント間の冗長性を除去しないからである。これに対して 3次元予測では他のあらゆる手法よりも低い 14.25 に達している。この理由は 3次元予測が隣接画素間の冗長性とコンポーネント間の冗長性を同時に除去するからである。

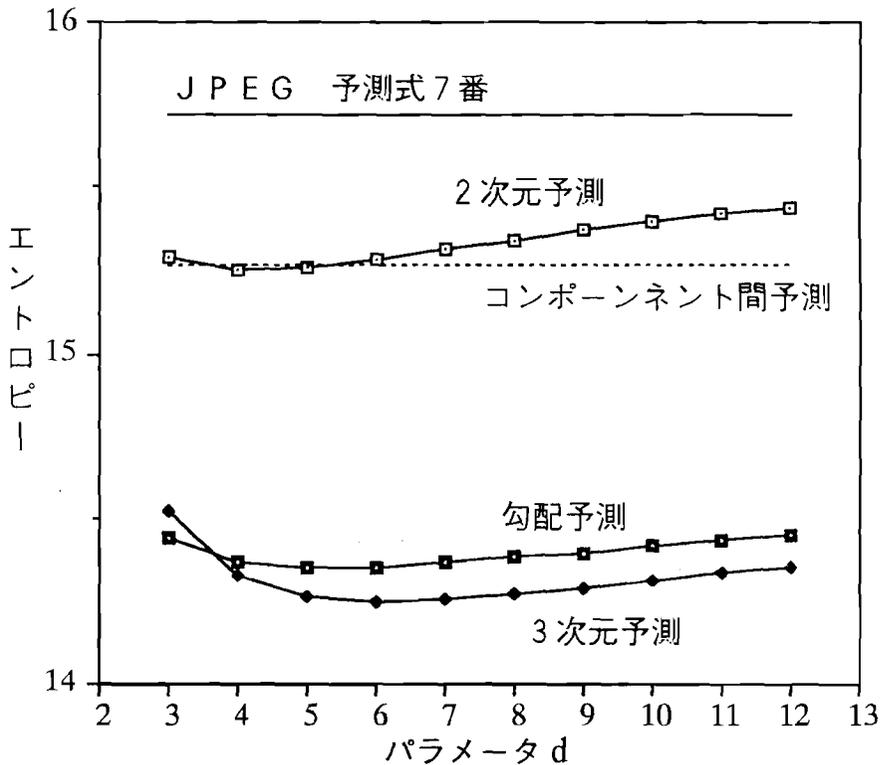


図 3.6 パラメータ  $d$  に対する予測誤差信号のエントロピー

一方、勾配予測は予測係数が少ないにもかかわらず、3次元予測に近いエントロピーに達している。このことから勾配予測は演算上の冗長性が少なく、画像信号の性質に適していると言える。d=2のときに3次元予測よりも低いエントロピーに達する理由は、予測係数が少ないため最小二乗法が安定し易いからである。

### 3.6 結言

本論文ではデジタルカラー画像の可逆符号化法として、3次元予測および勾配予測を用いた予測符号化を提案した。カラー画像において3次元予測を行った結果、JPEGの予測によるエントロピーを10%削減することができた。また勾配予測を用いた結果、演算量を削減することができた。これらのことから提案を行った手法の有効性が確認できた。

## 第4章 予測誤差信号のための算術符号化

### 4.1 緒言

予測符号化システムにおいては、予測器から出力された誤差信号を符号化器が2進コードに変換する。ここで符号化器は入力シンボルの出現確率を推定し、符号長の期待値が最短になるような符号を生成しなければならない。符号化器においてこの確率推定は重要である。第2章の予測符号化システムにおいては議論を簡単にするためにハフマン符号化を行った。そこでは予測誤差信号の出現頻度を全て調べた後に一つの確率分布を想定し、それに基づいて符号化テーブルを作成した。しかし多くの画像では出現確率が局所的に変動するため、一つの確率分布に基づいて符号化を行うことは最適であるとは言えない。理想的には1画素ずつ信号の出現確率を推定し、動的符号化を行うべきである。

動的符号化にはハフマン符号化よりも算術符号化の方が適している。その理由は確率が変化したときにハフマン符号化では符号化テーブルを再構築しなければならないが、算術符号化では確率を更新するだけでよいからである。Witten [25]はこの算術符号化を用いて一つの動的符号化手法を実現している。Wittenの手法では過去のシンボル系列の出現頻度を記録し、それに基づいて次のシンボルの出現確率を推定する。出現したシンボルを符号化すると同時にそのシンボルの頻度を一つ増し、また出現確率を更新する。この手法は文字列の符号化や確率変動の緩やかな信号の符号化に有効であった。

ところが画像のように確率分布が急激に変動するような信号に対してはWittenの手法は適さない。その理由は膨大なサンプルの頻度分布が現在の確率推定に大きく影響しているため、推定機能が局所的な確率の変動に追従しなくなるからである。これに対してJPEGはこの変動に追従するため、注目画素の上と左の2画素から現在の状態を分類し、その状態に応じて確率を選択する手法を提案している。この手法は局所的な確率変動に迅速に応答することができる。しかしわずか2点によって確率

を決定するため、雑音に敏感過ぎるという欠点がある。

そこで本章ではこれらの問題点を改善する新しい確率推定手法を提案する。提案手法は、

- (1) 局所的な確率変動に迅速に応答し、
- (2) 雑音に対して安定である

という特徴がある。4.2節では提案する確率推定手法および動的符号化のアルゴリズムについて詳しく述べる。4.3節では提案手法および従来手法を用いた予測符号化の計算機シミュレーションを行い、圧縮比および処理速度について比較検討を行う。

## 4.2 動的算術符号化

本節では新しい確率推定法を用いた動的算術符号化法を提案する [50]。一般に予測後の誤差信号の頻度分布はラプラス分布を示すことが知られている [51]。そこで本手法はサンプル信号の頻度分布をそのまま確率分布として利用するのではなく、それらの信号から平均値 0、分散  $\sigma^2$  のラプラス分布型の確率分布を推定する。変動するパラメータは分散  $\sigma^2$  のみであると仮定し、これを 1 画素ごとに推定しながら符号化を行う。本手法は少ないサンプル信号で安定した確率推定を行えるため、信号の性質の局所的な変動に追従することが可能である。以下にまず確率の推定方法を述べ、次に動的符号化のアルゴリズムを述べる。

### 4.2.1 確率推定方法

本手法では予測誤差信号  $t$  の確率密度  $p(t)$  は、平均値 0、分散  $\sigma^2$  のラプラス分布

$$p(t) = \frac{1}{\sqrt{2 \cdot \sigma^2}} \cdot \exp\left(\frac{-2 \cdot |t|}{\sqrt{2 \cdot \sigma^2}}\right) \quad (4.1)$$

で表すことができると仮定する。式を簡単にするために、

$$\frac{1}{\sqrt{2} \cdot \sigma^2} = -\log s \quad (0 < s < 1) \quad (4.2)$$

とおくと、式(4.1)は

$$p(t) = -(\log s) \cdot s^{2 \cdot |t|} \quad (4.3)$$

となる。以後、本章では分散  $\sigma^2$  の代わりに  $s$  を用いる。この  $s$  は画素ごとに変動するパラメータとし、その値を以下に述べる方法で推定する。

信号  $t$  は最も近い整数値  $g$  に量子化されて観測されるとすると、整数値  $g$  の出現確率  $P(g)$  は

$$P(g) = \int_{g-1/2}^{g+1/2} p(t) \cdot dt = \begin{cases} 1-s & (g=0) \\ \frac{1}{2}(s^{-1}-s) s^{2|g|} & (g \neq 0) \end{cases} \quad (4.4)$$

となる。このような信号源から  $k$  個の信号 ( $\mathbf{g} = g_1, g_2, \dots, g_k$ ) が観測される確率  $P(\mathbf{g})$  は

$$P(\mathbf{g}) = \prod_{i=1}^k P(g_i) \quad (4.5)$$

となる。ここで確率  $P(\mathbf{g})$  を最大にするような  $s$  を求める。まず式(4.5)の両辺の常用または自然対数をとって

$$\log P(\mathbf{g}) = \sum_{i=1}^k \log P(g_i) \quad (4.6)$$

とする。右辺に式(4.4)を代入すると

$$\log P(\mathbf{g}) = \sum_{\substack{i=1 \\ g_i=0}}^k \log(1-s) + \sum_{\substack{i=1 \\ g_i \neq 0}}^k \left\{ \log\left(\frac{s^{-1}-s}{2}\right) + 2|g_i| \log s \right\} \quad (4.7)$$

となる。  $k$  個の観測値のうち、  $g_i=0$  となる観測値が  $N$  個、それ以外の観

測値が  $M$  個,  $|g_i|$  の総和  $\sum_{i=1}^k |g_i|$  を  $C$  とすると, 式 (4.6) は

$$\log P(\mathbf{g}) = N \cdot \log(1-s) + M \cdot \log \frac{(s^{-1}-s)}{2} + 2 \cdot C \cdot \log s \quad (4.8)$$

となる. 式 (4.8) の両辺を  $s$  で微分すると

$$\frac{d}{ds} \log P(\mathbf{g}) = N \cdot \frac{1}{(s-1)} + M \cdot \frac{(s^2+1)}{s(s^2-1)} + \frac{2 \cdot C}{s} \quad (4.9)$$

となる. 式 (4.9) の右辺を 0 とおくと,

$$(k+2 \cdot C) \cdot s^2 + N \cdot s + (M-2 \cdot C) = 0 \quad (4.10)$$

となる. この 2 次方程式の二つの解のうち,  $0 < s < 1$  となるのは

$$s = \frac{-N + \sqrt{N^2 - 4 \cdot (k+2 \cdot C) \cdot (M-2 \cdot C)}}{2 \cdot (k+2 \cdot C)} \quad (4.11)$$

である. このようにして求めた  $s$  を式 (4.4) に代入することにより信号  $\mathbf{g}$  の出現確率  $P(\mathbf{g})$  を求めることができる. 本手法ではこの出現確率  $P(\mathbf{g})$  を次の信号の確率推定に利用して算術符号化を行う.

#### 4.2.2 動的符号化のアルゴリズム

画像においては, 信号の性質が局所的に変動していると考えられる. したがってある点の信号の出現確率は, その近傍の信号から推定されるべきである. 図 4.1 に提案する動的符号化のアルゴリズムを示す. 以下に各ステップの説明を行う.

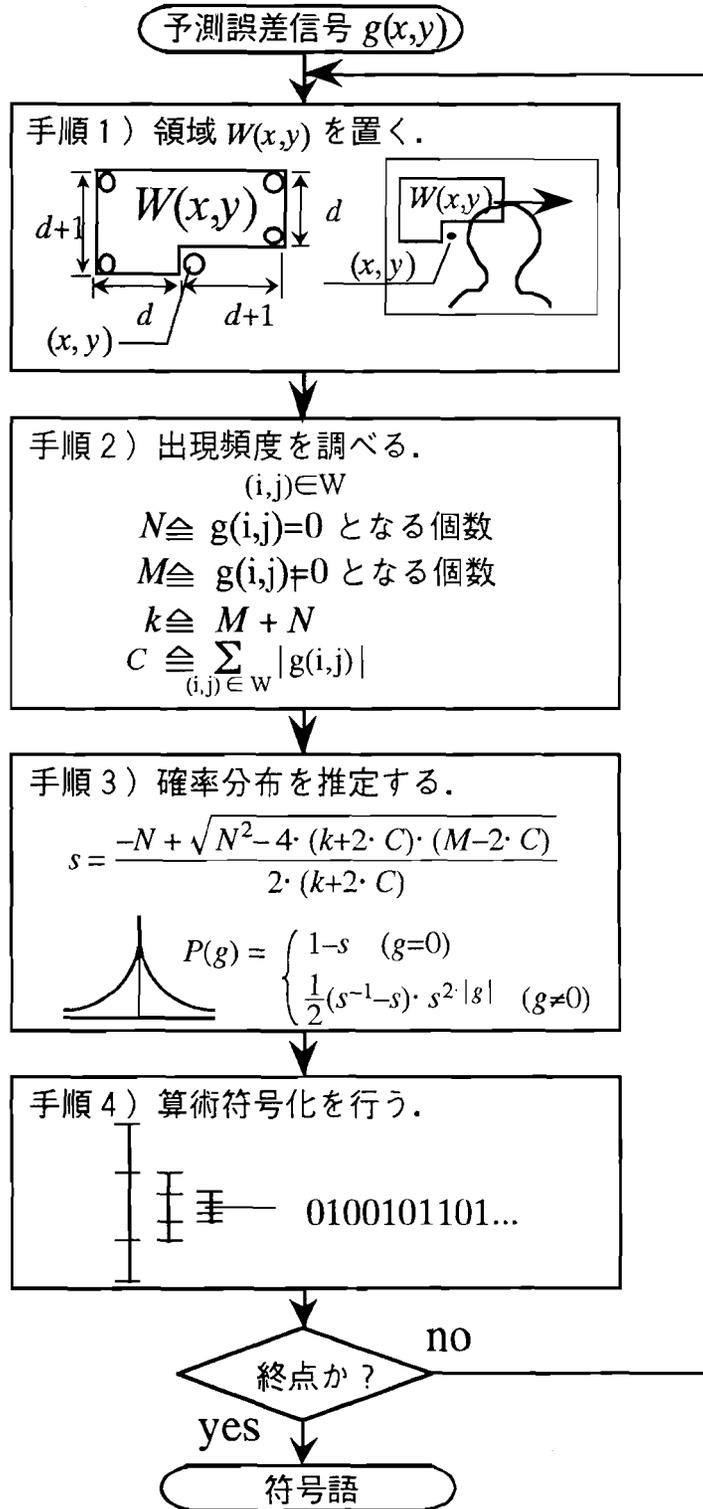


図 4.1 提案する符号化アルゴリズム

手順1) 領域  $W(x,y)$  を置く。

座標  $(x,y)$  の近傍領域を  $W(x,y)$  と定義する。図中の  $d$  は領域  $W(x,y)$  の大きさをきめるパラメータである。

手順2) 出現頻度を調べる。

領域  $W(x,y)$  内において、4.2.1 で述べた  $N$  ( $0$  の出現個数) ,  $M$  ( $0$  以外の出現個数) および  $C$  (絶対値の合計) を計算する。

手順3) 確率分布を推定する。

4.2.1 に述べた手法で確率  $P(g)$  を算出する。

手順4) 算術符号化を行う。

確率  $P(g)$  に基づいて算術符号化を行い、符号を生成する。

一つの信号を符号化するごとに領域  $W(x,y)$  を次の座標に移動し、以上のステップを繰り返す。パラメータ  $d$  の大きさに関しては4.4節で議論する。

### 4.3 計算機シミュレーション

本章では提案手法および従来手法を合わせて7種類の符号化手法によって画像の予測誤差信号を符号化し、圧縮比および処理速度について考察を行う。実験に用いた画像を図4.2に示す。それぞれの画像は65,536バイト ( $256 \times 256 \times 8$  ビット) のデータから成っている。

#### 4.3.1 7種類の符号化手法

符号化手法として、表4.1に示した7種類の手法を用意した。以下に各手法について説明する。

##### (I) Lempel-Ziv 符号化 [26]

Lempel-Ziv 符号化は UNIX 上で広く用いられている手法である。符号化器は信号の中から過去に登録されている系列と同じ系列を探し出し、それをあるコードに置き換える。そして符号化と同時に系列の参照テーブルを作成する。復号化器も復号と同時に参照テーブルを作成すること



No.1



No.2



No.3



No.4



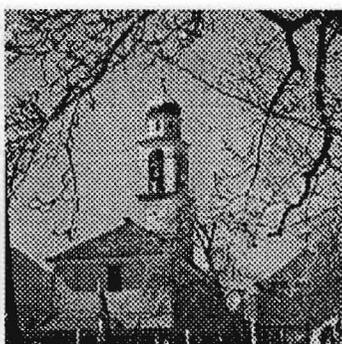
No.5



No.6



No.7



No.8



No.9

図4.2 テスト画像

表 4.1 7種類の符号化手法

手法	静的/ 動的	パス	観測領域	符号化 テーブル
I Lempel-Ziv 符号化	動的	1 パス	処理済み領域	—
II ハフマン符号化	静的	2 パス	画像全体	作成
III 静的算術符号化	静的	2 パス	画像全体	作成
IV 動的算術符号化	動的	1 パス	処理済み領域	—
V 局所動的算術符号化	動的	1 パス	局所近傍	—
VI JPEG 算術符号化	動的	1 パス	上と左のみ	内部定義
VII 提案手法	動的	1 パス	局所近傍	—

ができるので、符号化器は復号化器にテーブルを伝送する必要がない。

#### (II) ハフマン符号化 [19]-[21]

ハフマン符号化は簡単で広く用いられている手法である。先に述べたようにハフマン符号化は動的符号化に適していないので、ここでは静的符号化として用いている。符号化テーブルはあらかじめ信号の頻度分布を調べた後に構築し、1枚の画像を通して固定する。復号側は事前に信号の頻度分布を調べることができないので、符号化器は符号化テーブルを保存し、それを復号側に伝送しなければならない。

#### (III) 静的算術符号化 [22]-[24]

静的算術符号化ではあらかじめ信号の頻度分布を調べ、その分布を正規化したものを出現確率とする。上のハフマン符号化の場合と同様で、復号側は事前に信号の頻度分布を調べることができないので、符号化器はそれを保存し、復号側に伝送しなければならない。

## (IV) 動的算術符号化 [25]

動的算術符号化では一つの信号を符号化するたびに過去の頻度分布を更新し、それを正規化したものを次の信号の出現確率とする手法である。この手法では全てのシンボルが等確率の状態から符号化を始め、最後には静的符号化の際に作成する確率分布と同じものを得て符号化を終了することになる。復号化側でも同じアルゴリズムによって出現確率を算出できるため、符号化側は確率を伝送する必要がない。

## (V) 局所動的算術符号化

局所動的算術符号化は動的算術符号化と似ているが、後者が過去の全ての頻度を調べるのに対して、前者はサンプル領域を  $W(x,y)$  内に限定して頻度を調べる。このように局所的な頻度分布に注目することにより、確率の変動に追従することを狙っている。

## (VI) JPEG 算術符号化 [28]

JPEG の算術符号化の詳細は文献 [28] で述べられている。この手法は注目している画素の上と左の画素値から現在の状態を 25 種類に分類する。その状態に応じて、内部的に定められているテーブルから確率を選択する仕組みである。この手法も他の動的算術符号化と同様に確率を保存する必要がないが、符号化側と復号化側で同じ内部テーブルをもたなければならない。

## (VII) 提案手法

提案手法は 4.2 で述べた。本手法は局所動的算術符号化と同じように領域  $W(x,y)$  内の信号を調べるが、それがシンボルの頻度分布を正規化することにより確率分布を得るのに対して、本手法はシンボルの事後出現確率が最大となるようなラプラス分布を求める。本手法も他の動的算術符号化と同様に出現確率を記録する必要がない。

表4.2 図4.2のNo.4の画像に対する5種類の符号化による圧縮比の比較

手法	コード サイズ	符号化 テーブル	合計	圧縮比
I Lempel-Ziv 符号化	47,557 byte	—	47,557 byte	72.6%
II ハフマン符号化	40,891 byte	266 byte	41,157 byte	62.8%
III 静的算術符号化	40,642 byte	370 byte	41,012 byte	62.6%
IV 動的算術符号化	40,839 byte	—	40,839 byte	62.3%
VI JPEG 算術符号化	38,108 byte	—	38,108 byte	58.1%

#### 4.3.2 手法I,II,III,IVおよびVIによる圧縮比

まず No.4 の画像を式 (2.8), すなわち JPEG の 7 番の予測式で予測誤差に変換し, Lempel-Ziv, ハフマン, 静的算術符号化, 動的算術符号化および JPEG の算術符号化で符号化した. 局所動的算術符号化および提案手法の結果については次の節で述べる. 表 4.2 は生成コードサイズと圧縮比を示している. ここで圧縮比とは生成コードサイズと符号化テーブルまたは確率テーブルの合計サイズの, 原データサイズに対する比率である.

Lempel-Ziv は 5 種類の手法の中で最も大きいコードを生成している. この理由は誤差信号の中に二つの同じ系列が繰り返されている可能性が非常に少ないためである. この手法は誤差信号の符号化には適していないと考えられる. ハフマン符号化と静的算術符号化は Lempel-Ziv の次に大きなコードを生成している. 静的算術符号化はハフマン符号化に比べて少ないコードを生成しているが, 符号化テーブルが大きいため, 両者の圧縮比の差はわずか 0.2 ポイントである. 動的算術符号化は静的算術符号化に比べて少し大きなコードを生成しているが, 前者は符号化テーブルを保存しないため, 後者に比べて 0.3 ポイント小さな圧縮比になっている. JPEG の算術符号化は 5 種類の手法の中で最も小さい 58.1% の圧縮比に達している. この原因は主に 25 種類への分類が有効であるため

と考えられる．表 4.2 の圧縮比からわかるように，JPEG の算術符号化は 5 種類の手法の中で最も有効である．

#### 4.3.3 手法VおよびVIIによる圧縮比

次に上の予測誤差信号を，局所動的算術符号化および提案手法を用いて符号化した．両手法は領域  $W(x,y)$  のサイズを決めるパラメータ  $d$  をもっている．パラメータ  $d$  に対する圧縮比の関係を図 4.3 に示す．図中の水平線は先の表 4.2 に挙げられている圧縮比を示したものである．

局所動的算術符号化は  $d$  が 25 よりも大きい場合に動的算術符号化を下回ることができる．この範囲では  $W(x,y)$  によるサンプル領域の限定が有効であるといえる．しかし  $d$  が 25 よりも小さい場合，圧縮比は急増する．この現象から局所動的算術符号化が安定な確率推定を行うためには  $d > 25$

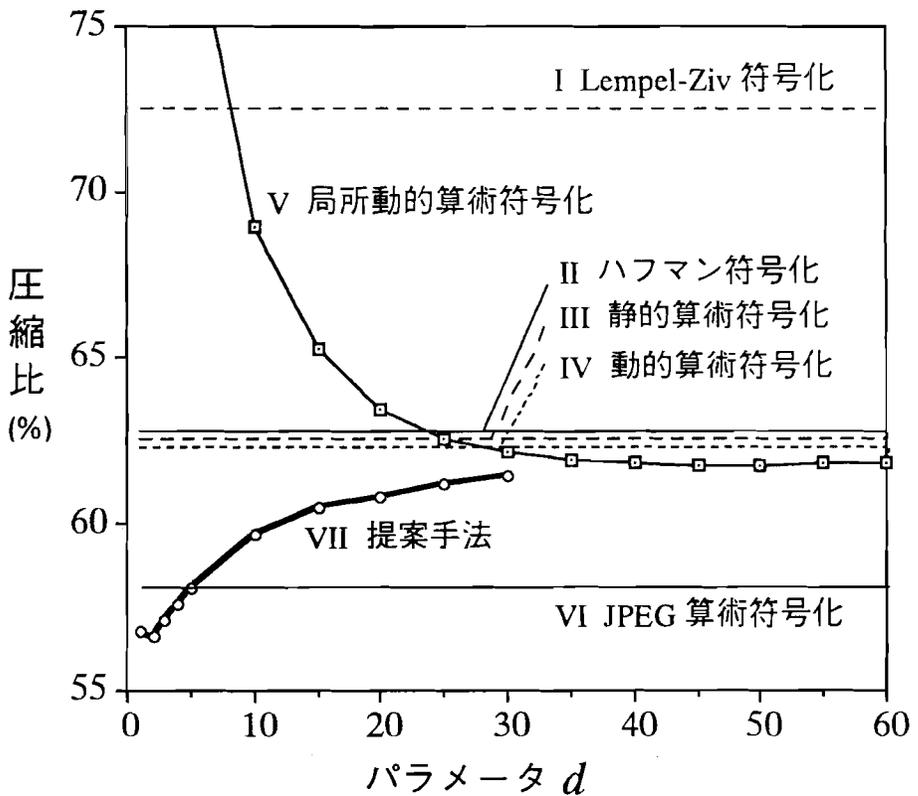


図4.3 パラメータ  $d$  と圧縮比の関係

表4.3 9枚の画像に対する圧縮比

符号化手法 \ 画像	No.1	No.2	No.3	No.4	No.5	No.6	No.7	No.8	No.9	平均
I Lempel-Ziv 符号化	65.0	67.9	69.3	72.6	71.0	76.4	89.5	99.8	99.8	79.0
II ハフマン符号化	57.8	59.5	60.9	62.8	60.7	63.7	74.5	81.5	80.3	66.9
III 静的算術符号化	57.5	59.2	60.6	62.6	60.4	63.6	74.3	81.3	80.2	66.6
IV 動的算術符号化	57.5	59.0	60.4	62.3	60.4	63.6	74.0	80.8	79.8	66.4
V 局所動的算術符号化 ( $d=50$ )	56.9	59.1	60.3	61.8	59.9	64.2	74.5	80.5	79.8	66.3
VI JPEG 算術符号化	54.0	58.8	59.6	58.1	61.0	65.7	75.2	80.9	81.1	66.1
VII 提案手法 ( $d=2$ )	<b>51.7</b>	<b>55.9</b>	<b>56.4</b>	<b>56.6</b>	<b>57.5</b>	<b>63.2</b>	<b>72.0</b>	<b>75.6</b>	<b>77.9</b>	<b>63.0</b>

(1,000 サンプル) 以上の膨大な量の統計をとらなければならないことがわかる。

これに対して提案手法は  $d$  が小さいほど圧縮比が低くなり、 $d=2$  のときに極小値に達している。そしてその値は JPEG の算術符号化を用いた場合よりも小さい。このグラフは提案手法が少ないサンプル数での確な確率推定を行っていることを証明している。また  $d>2$  のときグラフが単調増加していることから、遠くの画素をサンプルすることが逆効果であることがわかる。

#### 4.3.4 9枚の画像に対する圧縮比

9枚の画像に対する圧縮比を表4.3に示す。パラメータ  $d$  は局所動的算術符号化において  $d=50$  に、提案手法において  $d=2$  に固定した。

圧縮比は画像によって異なる。しかしどの画像においても図4.3と同じ傾向が現われている。平均的には Lempel-Ziv が 79% という最も大きな圧縮比を示している。次にハフマン、静的算術符号化、動的算術符号化、局所動的算術符号化および JPEG の算術符号化が 66% から 67% の間の値を示している。この5手法の中では JPEG の算術符号化が最低値を示しているが、その値はハフマン符号化に比べて 0.8 ポイント改善されてい

表 4.4 予測式と符号化器の組合わせに対する圧縮比 (%)

予測式 符号化手法	JPEG1 式(2.2)	JPEG2 式(2.3)	JPEG3 式(2.4)	JPEG4 式(2.5)	JPEG5 式(2.6)	JPEG6 式(2.7)	2次元予測 式(2.9)
I Lempel-Ziv 符号化	83.9	80.1	86.4	81.6	80.7	79.3	70.1
II ハフマン符号化	71.4	68.6	74.4	69.2	68.1	66.9	61.4
III 静的算術符号化	71.3	68.5	74.2	69.0	67.9	66.6	61.2
IV 動的算術符号化	70.9	68.1	73.8	68.8	67.7	66.4	61.2
V 局所動的算術符号化 ( $d=50$ )	70.6	67.9	73.4	68.8	67.6	66.4	61.2
VI JPEG 算術符号化	69.8	67.5	72.4	69.3	68.0	66.8	61.8
VII 提案手法 ( $d=2$ )	<b>67.2</b>	<b>64.6</b>	<b>69.6</b>	<b>66.1</b>	<b>64.6</b>	<b>63.5</b>	<b>58.4</b>

るだけである。これに対して提案手法は 63.0% に達しており、JPEG の算術符号化によるコードサイズをさらに 5% 削減する結果となった。このように圧縮比の点では提案手法は最も優れているといえる。

#### 4.3.5 予測器との組合わせに対する圧縮比

圧縮比は符号化器のみならず、予測器によっても左右される。上の実験では JPEG の 7 番の予測式を用いたが、JPEG の残りの 6 種類の予測式および第 2 章で提案した 2 次元予測式についても同様の実験を行った。表 4.4 に予測式と符号化器の全ての組合わせに対する圧縮比を示す。表中の各値は 9 枚の画像の平均値である。

圧縮比は予測式によって異なるが、どの予測式を用いた場合も提案手法が最も小さな圧縮比に達していることがわかる。平均的に提案手法は JPEG の算術符号化よりも約 3 ポイント小さな圧縮比に達している。このことから提案手法は予測器によらず有効であると言える。

#### 4.3.6 処理時間

次に処理時間について議論する。図 4.4 は SPARCstation ELC (21 MIPS,

16 MB メモリ) 上で 7 種類の符号化手法が 64 k バイトのコンポーネントを符号化するのに要した時間である。

Lempel-Ziv, ハフマン, 静的算術符号化は非常に高速であり, それぞれ 0.40, 1.02, 1.47 秒で符号化を完了している. これに対して他の動的手法は多くの時間を要する. 動的算術符号化は 1 画素ごとに確率を更新するため, 静的算術符号化よりも 7.73 秒多くの時間を要する. JPEG の算術符号化も同様に確率を更新するが, 確率は内部的に定められているテーブルから選択されるだけなので動的算術符号化よりも少し高速である. 局所動的算術符号化および提案手法では領域  $W(x,y)$  のサイズと共に処理時間が増す. 局所動的算術符号化は  $d=50$  のときに 14.5 秒を要する. 提案手法は  $W(x,y)$  内を観測した後に, さらに式 (4.4) および式 (4.11) を算出しなければならないので, これに約 7 秒要する.

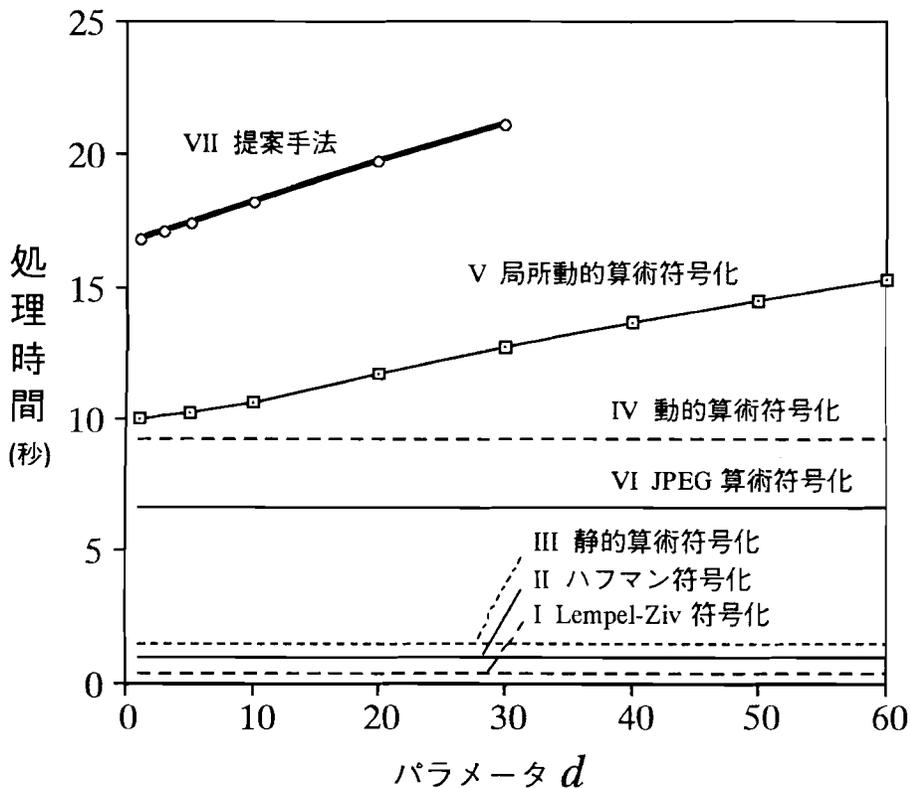


図4.4 処理時間

その結果、 $d=2$  のときには JPEG の算術符号化に比べて約 2.5 倍の時間を要することになる。このように提案手法は圧縮比の代わりに処理時間を犠牲にしている。

#### 4.3.7 リアルタイムの処理能力

7種類の符号化手法 (I-VII) について処理時間と圧縮比の関係を調べ、図 4.5 に示した。提案手法は圧縮比が最も低いが処理時間が最も長いので右下に位置している。図中の破線は

$$y = 100 - 100 / 8 x \quad (4.12)$$

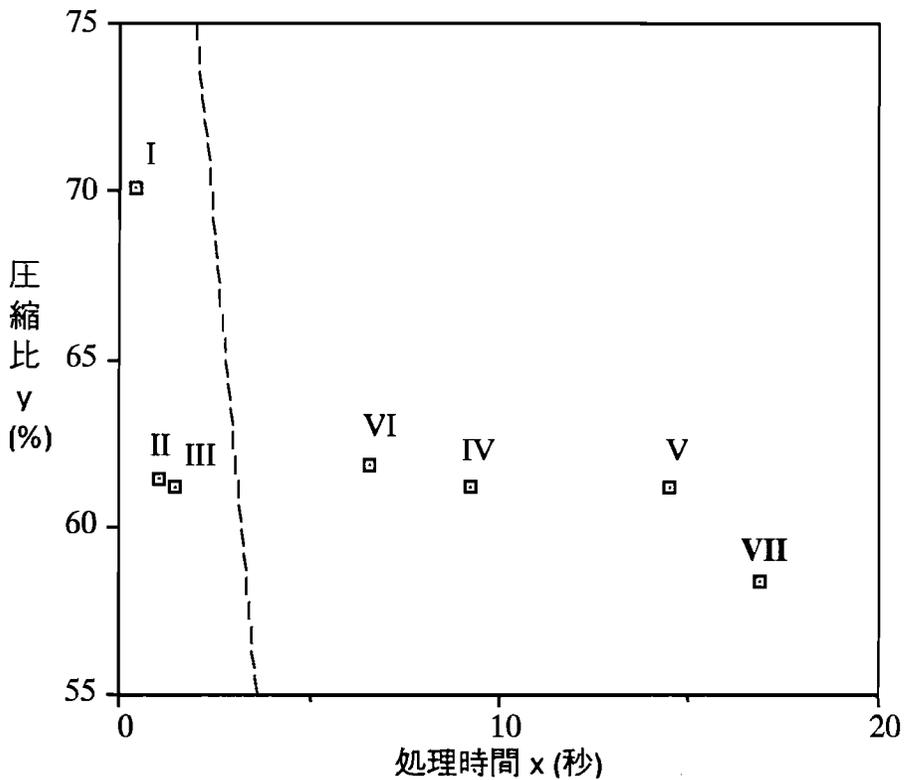


図4.5 処理時間と圧縮比の関係

を表している。式 (4.12) は 64 k バイトのデータを 64 kb/s の通信回線によって伝送すると想定したときの、伝送時間と残りデータの関係を表している。もし 64 k バイトの画像データを圧縮せずに伝送すると、この破線に従って残りデータが減少し、8 秒後に全てのデータを伝送し終えることになる。このデータをリアルタイムに圧縮し、かつ伝送時間を 8 秒よりも短くするためには、プロットが破線よりも左側になければならない。しかし JPEG の算術符号化を含め多くの動的算術符号化は破線の右側に位置する。グラフより、提案手法を破線の左側に位置付けるためには、あと 5 倍から 6 倍の高速化が必要であることがわかる。

圧縮手法に要求される処理速度は用途によって異なるが、リアルタイム処理が要求されるシステムにおいて動的算術符号化を行う場合には、専用ハードウェアによる高速化が必要であると考えられる。

#### 4.4 結言

本章では予測誤差信号のための動的算術符号化を提案した。計算機シミュレーションの結果、提案手法は Lempel-Ziv、ハフマン符号化、静的算術符号化、動的算術符号化、局所動的算術符号化および JPEG の算術符号化によるコードサイズをさらに 5% 以上削減することができた。

提案手法では信号の確率分布がラプラス分布をなすと仮定したが、その他の分布の適合性についても検討する必要がある。また処理の高速化は今後の課題である。

## 第5章 モノクロ画像の非可逆符号化

### 5.1 緒言

非可逆符号化においては、少ないコード量でいかに高品質の画像を再現できるかが重要なポイントとなる。DCT [8] は、高能率な冗長抑圧が可能なることから Joint Photographic coding Experts Group (JPEG) の標準規格として採用され [27], [28], 専用ハードウェアの開発 [9] によって実用化されている。しかし、再生画像にブロック歪みやモスキート雑音が現れ、それらが視覚的な妨害になることが問題点となっている。

この問題点を改善する手法として、近年ウェーブレット変換 [38]-[43] が注目を浴びている。ウェーブレット変換は隣接する基底関数をオーバーラップさせることによってブロックの境界をなくし、視覚的な歪みを低減することができる。しかし、一般にフィルタ長が長いためにその処理は高速ではない。最も短い2タップのフィルタで実現できるハール・ウェーブレット変換 [39] は非常に高速な処理が可能であるが、基底関数がオーバーラップしていないためにブロック歪みが現れる。またハール・ウェーブレット変換は、他の帯域分割手法に比べて変換後の高域信号の電力およびエントロピーが高いことから、画像信号の冗長性の抑圧が不十分であると考えられる。

そこで本章では、高速処理が可能であり、コード量が少なく、かつブロック歪みの少ない再生画像が得られる手法として、帯域間予測を用いたハール・ウェーブレット変換 [52] を提案する。5.2 節では提案手法について詳しく述べる。5.3 節では本変換手法の帯域分割手法としての特性を、これに等価な Symmetric Short Kernel Filter (SSKF) バンク [34] を用いたサブバンド分割の考え方に基づいて解析する。5.4 節では実画像データの帯域分割と符号化、そしてプログレッシブ・ビルドアップによる原画像の再生実験を行う。これらの実験により、圧縮比の向上とブロック歪みの減少を確認する。

### 5.2 帯域間予測ハール・ウェーブレット変換 [52]

本章では高速処理が可能なハール・ウェーブレット変換に帯域間予測を導入することにより、信号電力の低減およびブロック歪の除去を行う。提案するシステムはハール・ウェーブレット変換と帯域間予測の二つの処理から構成されている。以下ではまずハール・ウェーブレット変換を簡潔に述べた後、提案する帯域間予測手法について詳しく述べる。

#### 5.2.1 ハール・ウェーブレット変換 [39]

ハール・ウェーブレット変換は、図 5.1 に示す直交基底関数を用いて多重解像度表現を行うものである。図 5.2 にそのブロック構成を示す。デジタル信号を扱う場合には低域通過フィルタ  $F_0(z)$  と高域通過フィルタ  $F_1(z)$  を、

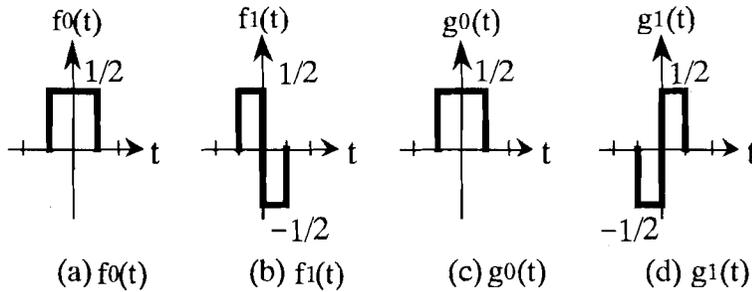


図 5.1 ハール基底

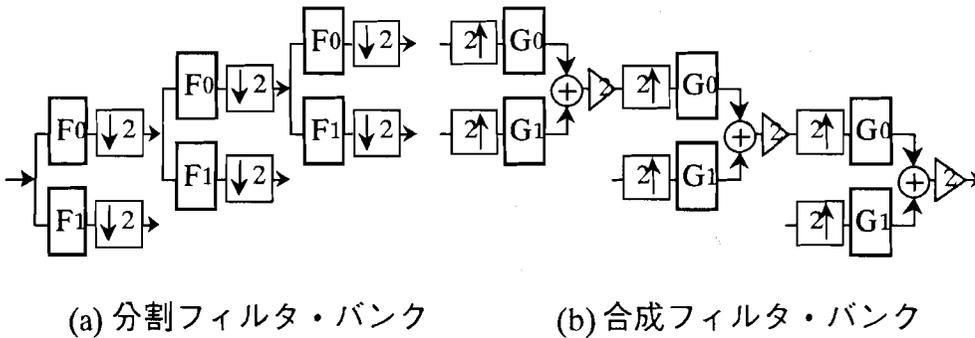


図 5.2 信号のオクターブ分割および再構成のためのブロック構成

$$F_0(z) = (1+z^{-1})/2, \quad (5.1)$$

$$F_1(z) = (1-z^{-1})/2 \quad (5.2)$$

として帯域分割フィルタ・バンクを構成する．分割後の低域信号に対して繰り返しこの帯域分割処理を行うことによって，帯域のオクターブ分割を行う．逆変換の際の合成フィルタ・バンクは分割フィルタ・バンクと

$$G_0(z) = F_1(-z), \quad (5.3)$$

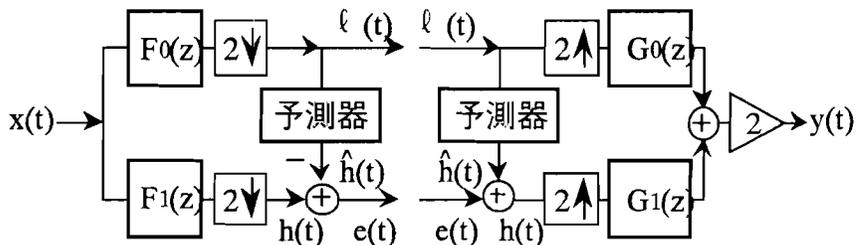
$$G_1(z) = -F_0(-z) \quad (5.4)$$

の関係がある．ハール・ウェーブレット変換を実現するフィルタ・バンクは，直交性と完全再構成性を満足する．

### 5.2.2 帯域間予測

本章で述べる帯域間予測とは，帯域を2分割するたびに低域信号の勾配から高域信号を予測し，それをよりエントロピーの低い補正信号に変換する処理である．2分割フィルタ・バンクへの入力信号を  $x(t)$  とし，分割後の低域側の信号を  $l(t)$ ，高域側の信号を  $h(t)$  とする．ハール・ウェーブレット変換では  $h(t)$  そのものを出力とするが，本変換手法では図5.3(a)のように  $l(t)$  から  $h(t)$  の予測を行い，予測値  $\hat{h}(t)$  と  $h(t)$  との誤差

$$e(t) = h(t) - \hat{h}(t) \quad (5.5)$$



(a) 分割フィルタ・バンク

(b) 合成フィルタ・バンク

図5.3 高域信号の予測

を補正信号として出力する．本変換手法では特に  $l(t)$  の 1 次勾配と  $h(t)$  の間の相関に注目し，予測値  $\hat{h}(t)$  を

$$\hat{h}(t) = \alpha \{ l(t+1) - l(t-1) \} \quad (5.6)$$

によって求める．ここで予測係数  $\alpha$  は信号  $e(t)$  のエントロピーが最小となるように決定すべきであるが，本変換手法では信号  $e(t)$  のエントロピーを最小にする  $\alpha$  と電力を最小にする  $\alpha$  がほぼ一致することを利用し，最小二乗法を用いて  $\alpha$  を決定する．信号が  $0 \leq t \leq N-1$  の範囲で定義されているときの  $\alpha$  の算出式は

$$\alpha = \frac{\sum_{t=0}^{N-1} h(t) \{ l(t+1) - l(t-1) \}}{\sum_{t=0}^{N-1} \{ l(t+1) - l(t-1) \}^2} \quad (5.7)$$

となる．本変換手法では  $h(t)$  の代わりに予測係数  $\alpha$  と補正信号  $e(t)$  を受信側である合成フィルタ・バンクに伝送する．

逆変換の際には図 5.3(b) のように正変換の際に求めた  $\alpha$  を用いて  $l(t)$  から  $h(t)$  の予測値  $\hat{h}(t)$  を求める．そして予測された  $\hat{h}(t)$  に補正信号  $e(t)$  を加算することによって  $h(t)$  を得る． $h(t)$  は予測と補正によって完全に再構成することができるので，本変換手法はハール・ウェーブレット変換と同様に原信号の完全再構成が可能である．

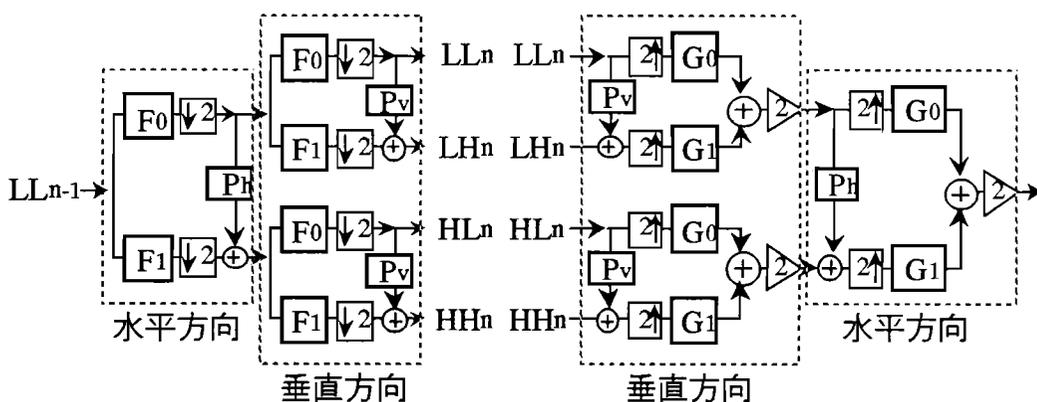
### 5.2.3 2次元への拡張

2次元のオクターブ分割および合成のためのブロック構成を図 5.4 に示す．画像信号においては図 5.4(a) のように水平方向と垂直方向に帯域分割を行い，これを一つの階層 (Layer) における処理とする．第  $n$  階層の分割において得られた四つの帯域を  $LL_n(x,y)$ ,  $LH_n(x,y)$ ,  $HL_n(x,y)$ , および  $HH_n(x,y)$  とする．このうち， $LL_n(x,y)$  は図 5.4(c) のように第  $n+1$  階層でさらに分割を行い，これを再帰的に繰り返すことによって 2次元のオクターブ分割を行う．

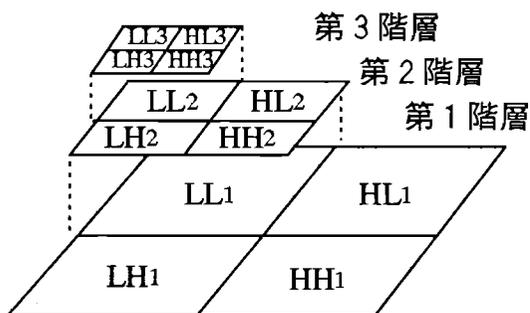
図中の  $P_h$ ,  $P_v$  はそれぞれ水平方向の帯域間予測と垂直方向の帯域間予測を表している. 予測係数は各階層において水平方向と垂直方向の2値を独立に算出する. なおハール・ウェーブレット変換, 帯域間予測とも線形処理であるため, 各階層における予測係数の値は水平方向と垂直方向の処理の順に依存しない.

### 5.3 帯域分割手法としての性能

本章では提案する変換手法の特性を, これに等価なサブバンド分割の考え方に基づいて詳しく述べる.



(a) 2次元分割フィルタ・バンク (b) 2次元合成フィルタ・バンク



(c) 周波数領域におけるオクターブ分割

図 5.4 2次元信号のオクターブ分割

5.3.1 等価フィルタ・バンク

提案する予測手法は線形フィルタとして実現することが可能であるため、図5.3における帯域分割処理と帯域間予測処理は、一つのフィルタ・バンクにまとめることができる。まず図5.3(a)の予測処理の位置を図5.5(a)のようにダウン・サンプリングの直前に移動し、ダウン・サンプリングの前の低域信号を  $l'(t)$ 、高域信号を  $h'(t)$ 、補正信号を  $e'(t)$  とする。このとき式(5.6)は

$$h'(t) = \alpha \{ l'(t+2) - l'(t-2) \} \tag{5.8}$$

と書き直される。式(5.8)と図5.5(a)より、 $x(t)$  と  $e'(t)$  の関係をそれらの  $z$  変換  $X(z)$ 、 $E'(z)$  を用いて表すと、

$$E'(z) / X(z) = F_1(z) - \alpha F_0(z) \{ z^2 - z^{-2} \} \tag{5.9}$$

となる。これに式(5.1)と式(5.2)を代入すると、

$$E'(z) / X(z) = \{ \alpha z^{-3} + \alpha z^{-2} - z^{-1} + 1 - \alpha z^1 - \alpha z^2 \} / 2 \tag{5.10}$$

となるので、 $e'(t)$  は図5.6(b)に示した6タップのフィルタにより  $x(t)$  から直接算出できることがわかる。したがって、提案する分割フィルタ・バンクは図5.7(a)のブロック構成に改めることができる。

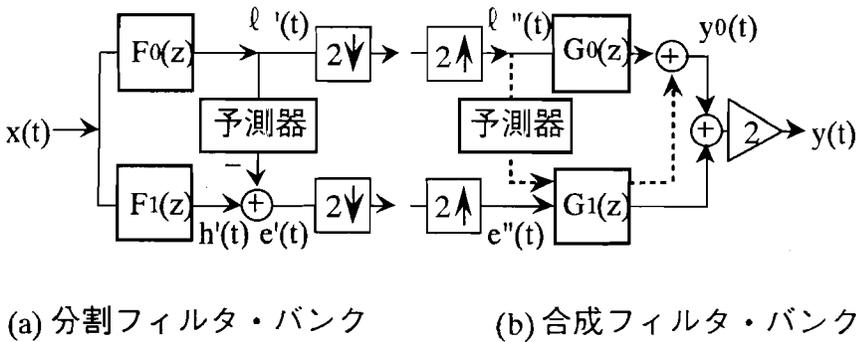


図 5.5 図5.3 の等価構造

次に合成側について述べる．図 5.3(b) の予測処理の位置を図 5.5(b) のようにアップ・サンプリングの直後に移動し，アップ・サンプリング後の低域信号を  $l''(t)$ ，補正信号を  $e''(t)$  とする．ここで予測信号は補正信号と合成せず，独立に  $G_1(z)$  に通した後に  $G_0(z)$  を通過した低域側の信号と合成する．そして得られた低域側の信号を  $y_0(t)$  とする．このとき  $l''(t)$  と  $y_0(t)$  の関係をそれらの  $z$  変換  $L''(z)$ ， $Y_0(z)$  を用いて表すと，

$$Y_0(z) / L''(z) = G_0(z) + \alpha G_1(z) \{z^2 - z^{-2}\} \quad (5.11)$$

となる．これに式 (5.1) と式 (5.2) を代入すると，

$$Y_0(z) / L''(z) = \{-\alpha z^{-2} + \alpha z^{-1} + 1 + z + \alpha z^2 - \alpha z^3\} / 2 \quad (5.12)$$

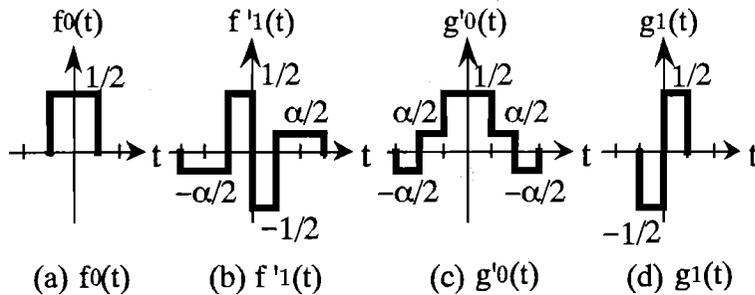


図 5.6 等価フィルタのインパルス応答

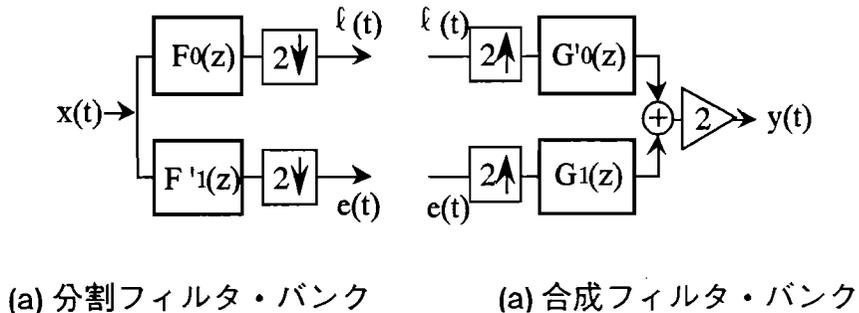


図 5.7 図 5.5 の等価構造

となるので、 $y_0(t)$ は図 5.6(c)に示した6タップのフィルタにより $l''(t)$ から直接算出できることがわかる。したがって、提案する合成フィルタ・バンクは図 5.7(b)のブロック構成に改めることができる。

以上のことから本変換手法は $F_0(z)$ 、 $F'_1(z)$ 、 $G'_0(z)$ 、そして $G_1(z)$ をフィルタ・バンクとするサブバンド分割手法と見なすことができる。さらに図 5.6に示した各フィルタは任意の $\alpha$ について式(5.3)と式(5.4)を満足し、かつ対称形であることから、これらはSSKFバンク[34]と等価であることがわかる。したがって、本変換手法はSSKFバンクに一つの可変パラメータを導入することにより、その周波数特性を可変にしたものと考えることができる。本変換手法は階層ごとに、また水平・垂直の方向ごとに $\alpha$ を決定することにより、画像信号の性質に応じたフィルタ設計を行うという特徴をもつ。また実際には予測を用いることにより、6タップのフィルタ処理をシフト演算と $\alpha$ の乗算のみで行うことができるため、高速な処理が可能である。

### 5.3.2 ブロック歪み

フィルタ通過後の画像に現れるブロック歪みは $G'_0(z)$ の特性に大きく依存している。低域信号はアップサンプラと $G'_0(z)$ を繰り返し通過することによって補間拡張されるので、 $G'_0(z)$ の高周波の遮断特性が良いほ

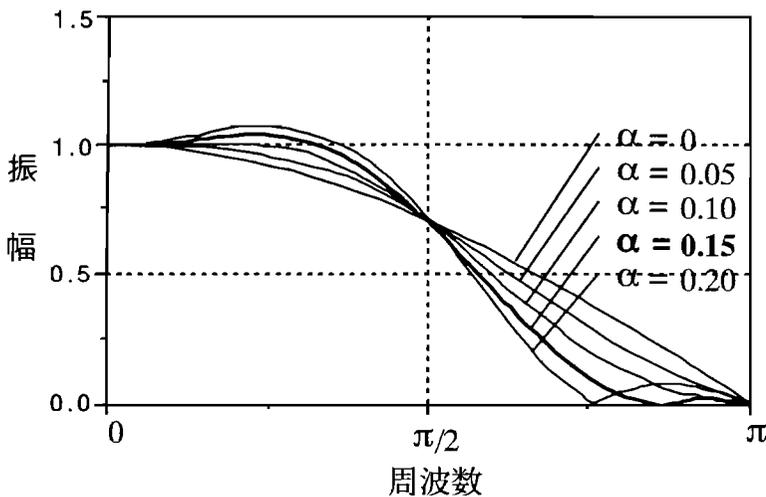


図 5.8  $G'_0(z)$ の周波数特性

どブロック歪みが現れにくい。図 5.8 に  $G'_0(z)$  の周波数特性を示す。 $\alpha=0$  のときはハール・ウェーブレット変換の特性に等しい。 $\alpha$  を増すと高域の遮断特性が向上し、 $\alpha=0.15$  付近で良好な特性が得られることがわかる。図 5.4(b) に示した 2次元の合成フィルタ・バンクにおいて  $\alpha$  を 0 または 0.15 に固定した場合の  $LL_n(x,y)$  ( $n=1\sim 3$ ) に対するインパルス応答を図 5.9 に示す。これらの波形は解像度を  $2^n$  倍 ( $n=1\sim 3$ ) に拡張する際の補間特性として見ることができる。したがってブロック歪みの影響は、これらの応答波形から容易に推測することができる。 $\alpha=0$  の場合（通常のハール・ウェーブレット変換）はインパルス応答が矩形波になっており、このことから量子化誤差などの雑音そのままブロック状の歪みとなって現れることがわかる。これに対して  $\alpha=0.15$  の場合は滑らかな補間特性をもっており、ブロック歪みが生じないことがわかる。ただし、本変換手法は  $e(t)$  の電力が最小となるように  $\alpha$  を決定しているので、この値は画像ごとに異なる。実画像における  $\alpha$  の算出値については 5.4 節で述べる。

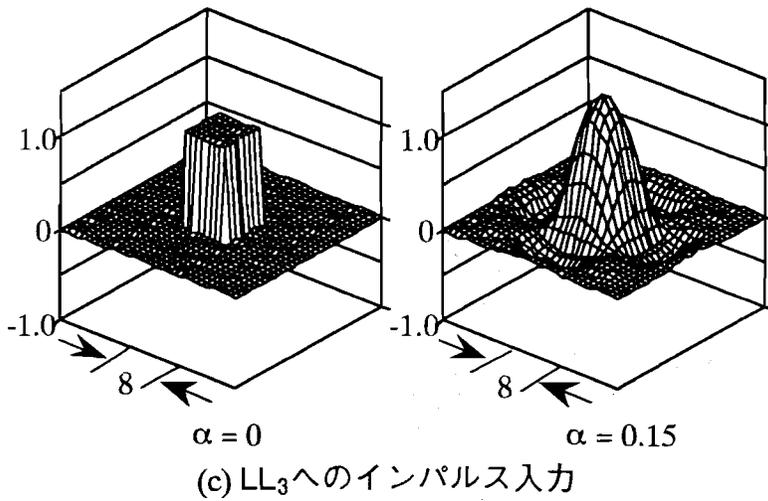
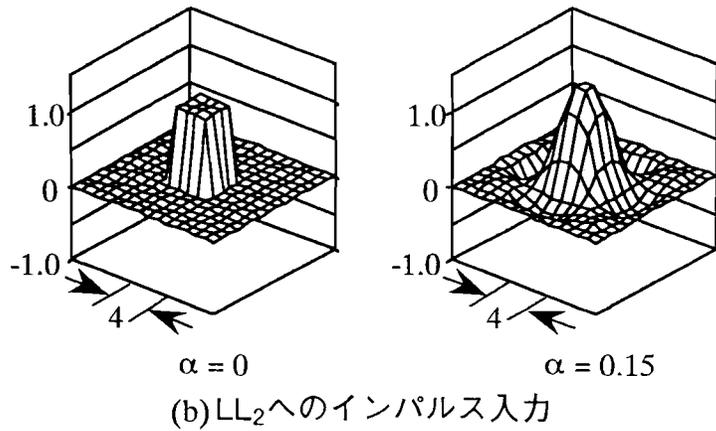
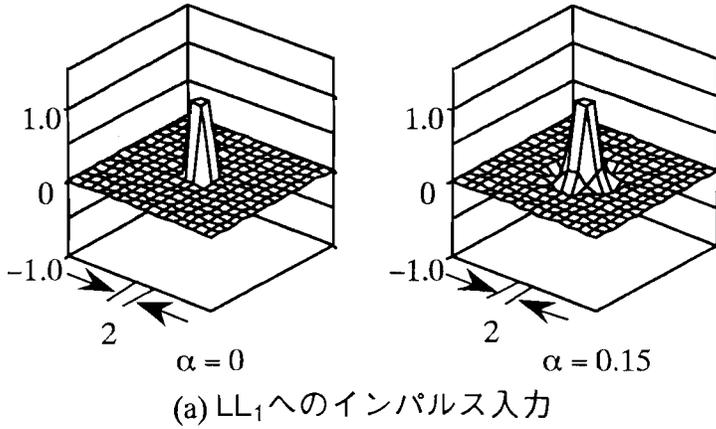


図 5.9 合成フィルタ・バンクのインパルス応答

表 5.1 1画素当たりの乗算回数

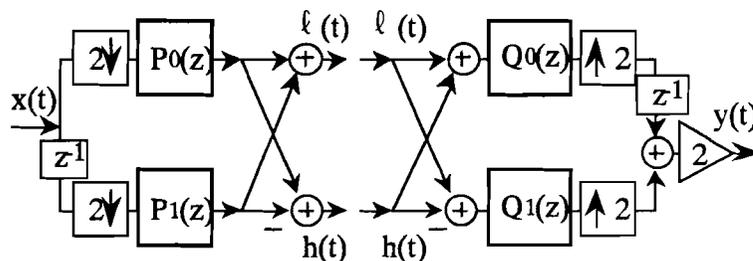
処理内容		ハール	QMF(n タップ)	提案手法
正変換	帯域分割	0	$n/2$	0
	予測係数の算出	—	—	1
	予測	—	—	0.5
	合計	0	$n/2$	1.5
逆変換	帯域合成	0	$n/2$	0
	予測	—	—	0.5
	合計	0	$n/2$	0.5

## 5.3.3 演算量

ハール・ウェーブレット変換, Quadrature Mirror Filter (QMF) [29]-[33]を用いたサブバンド分割, そして本変換手法の処理速度の違いを, 1画素当たりの乗算回数に注目して説明する. それぞれの手法において一つのフィルタ・バンクが1画素当たりが必要とする乗算回数を表 5.1 にまとめた. 以下で表 5.1 に示した乗算回数について述べる.

ハール・ウェーブレット変換においては, 信号を固定小数点型で扱うことにより, 係数 0.5 の乗算をシフト演算で実現することができる. したがって乗算は不要となる.

QMF においては, フィルタ・バンクを図 5.10 のようにバタフライ構造にすることによって乗算回数を削減できる [10]. 前述の  $F_0(z)$ ,  $F_1(z)$ ,



(a) 分割フィルタ・バンク

(a) 合成フィルタ・バンク

図 5.10 バタフライ構造のフィルタ・バンク

$G_0(z)$  および  $G_1(z)$  を  $n$  タップの QMF とすると、図中の  $P_0(z)$ ,  $P_1(z)$ ,  $Q_0(z)$  および  $Q_1(z)$  は

$$P_0(z) = \frac{1}{2} \{F_0(z^{1/2}) + F_1(z^{1/2})\} , \quad (5.13)$$

$$P_1(z) = \frac{z^{1/2}}{2} \{F_0(z^{1/2}) - F_1(z^{1/2})\} , \quad (5.14)$$

$$Q_0(z) = \frac{z^{1/2}}{2} \{G_0(z^{1/2}) + G_1(z^{1/2})\} , \quad (5.15)$$

$$Q_1(z) = \frac{1}{2} \{G_0(z^{1/2}) - G_1(z^{1/2})\} \quad (5.16)$$

と表すことができる。これらのフィルタは全て  $n/2$  タップとなる。さらに図 5.10 ではダウン・サンブラがフィルタの直前に配置されているので、フィルタへの入力画素数が半分になる。したがって QMF バンクの分割側の乗算回数は  $P_0(z)$  と  $P_1(z)$  の合計タップ数の半分、すなわち  $n/2$  回となる。合成側の乗算回数も同様の理由により  $n/2$  回となる。

本変換手法においては、ハール・ウェーブレット変換後に、さらに予測係数の算出と補正信号への変換処理を行う必要がある。式 (5.7) によって予測係数  $\alpha$  を決定するためには 1 画素当たり 2 回の乗算が、さらに  $\alpha$  を決定した後に式 (5.6) によって予測値を求めるには 1 回の乗算が必要である。ただし、このときの画素数はダウン・サンプリングによって半減しているため、入力画像の 1 画素に対する乗算回数は、それぞれ 1 回、0.5 回となる。逆変換時には正変換の際に求めた予測係数を用いるので、 $\alpha$  を算出する必要がなく、乗算回数は予測値を求めるための 0.5 回だけとなる。

以上をまとめると、ハール・ウェーブレット変換、 $n$  タップの QMF、そして本変換手法が 1 画素当たり必要とする乗算回数は、正変換時にはそれぞれ 0 回、 $n/2$  回、1.5 回、逆変換時にはそれぞれ 0 回、 $n/2$  回、0.5 回となる。したがって QMF において一般的なタップ数  $n=8, 12, 16$ 、またはそれ以上を用いた場合に比べると、本変換手法の方が乗算回数が少なく、処理が高速であることがわかる。また本変換手法は逆変換時には特に高速であるため、蓄積メディア用の符号化に有効である。

表5.2 8タップQMFの係数

t	$h_0(t)$	t	$h_0(t)$
0	0.009422118	4	0.4918060
1	-0.07091501	5	0.06968689
2	0.06968689	6	-0.07091501
3	0.4918060	7	0.009422118

表5.3  $\alpha$ の算出結果

階層 \ 方向	水平方向	垂直方向
第1階層	0.163	0.154
第2階層	0.160	0.158
第3階層	0.152	0.152

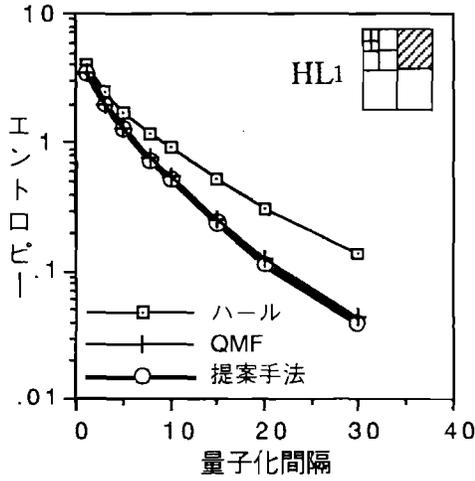
## 5.4 計算機シミュレーション

本変換手法の性能を評価するために、ハール・ウェーブレット変換、QMFを用いたサブバンド分割、そして提案する帯域間予測を用いたハール・ウェーブレット変換の三つの手法を用いて、画像データの帯域分割符号化とプログレッシブ・ビルドアップによる原画像の再生を行った。QMFのタップ数は、本変換手法に最も近いSNR特性をもつ、表5.2に示した8タップのものを用いた。

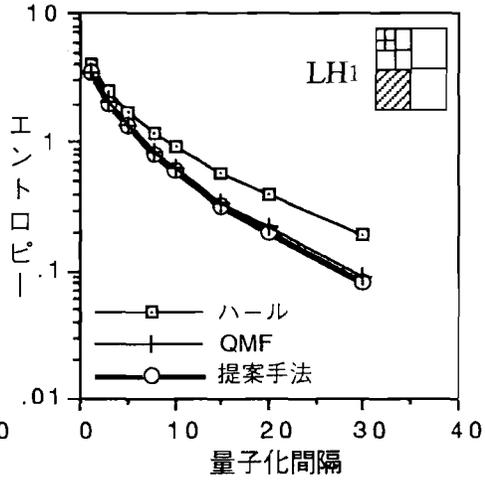
### 5.4.1 変換後の信号のエントロピー

三つの手法を用いて標準画像 "Girl" (256×256×8bit) を第3階層まで分割 (10分割) した。本変換手法において算出された予測係数  $\alpha$  を表5.3に示す。 $\alpha$ は水平・垂直の方向や階層によって異なっており、フィルタ・バンクの特性が変化していることがわかる。しかしながら  $\alpha$ は図5.8に示した  $G_0(z)$ の特性曲線の中で、高域遮断特性に優れた0.15付近に集中しており、このことから予測誤差を小さくすることが高域の遮断につながることをわかる。

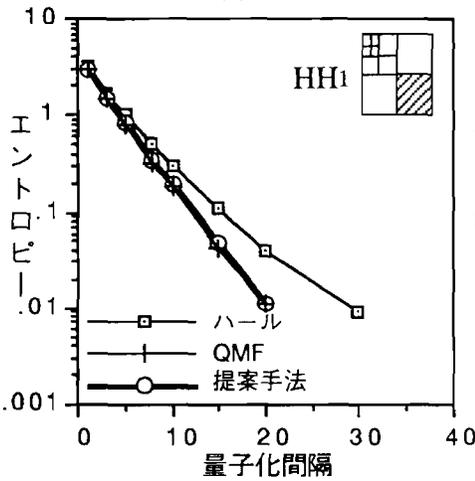
変換後の各帯域信号のエントロピーを図5.11に示す。グラフの横軸は



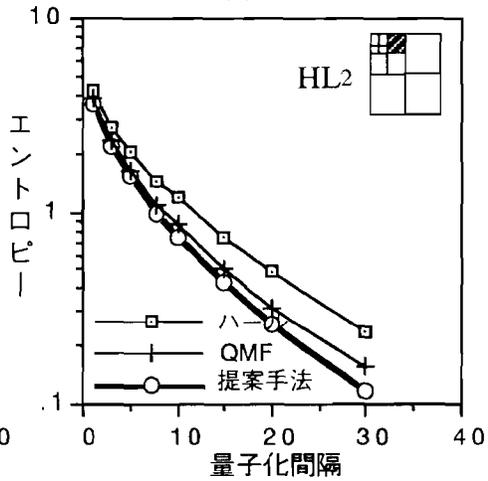
(a) HL1



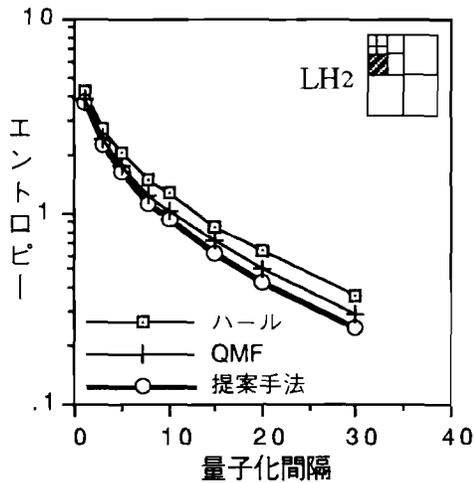
(b) LH1



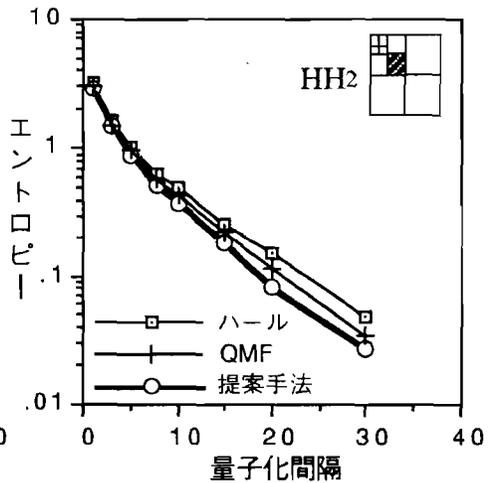
(c) HH1



(d) HL2



(e) LH2



(f) HH2

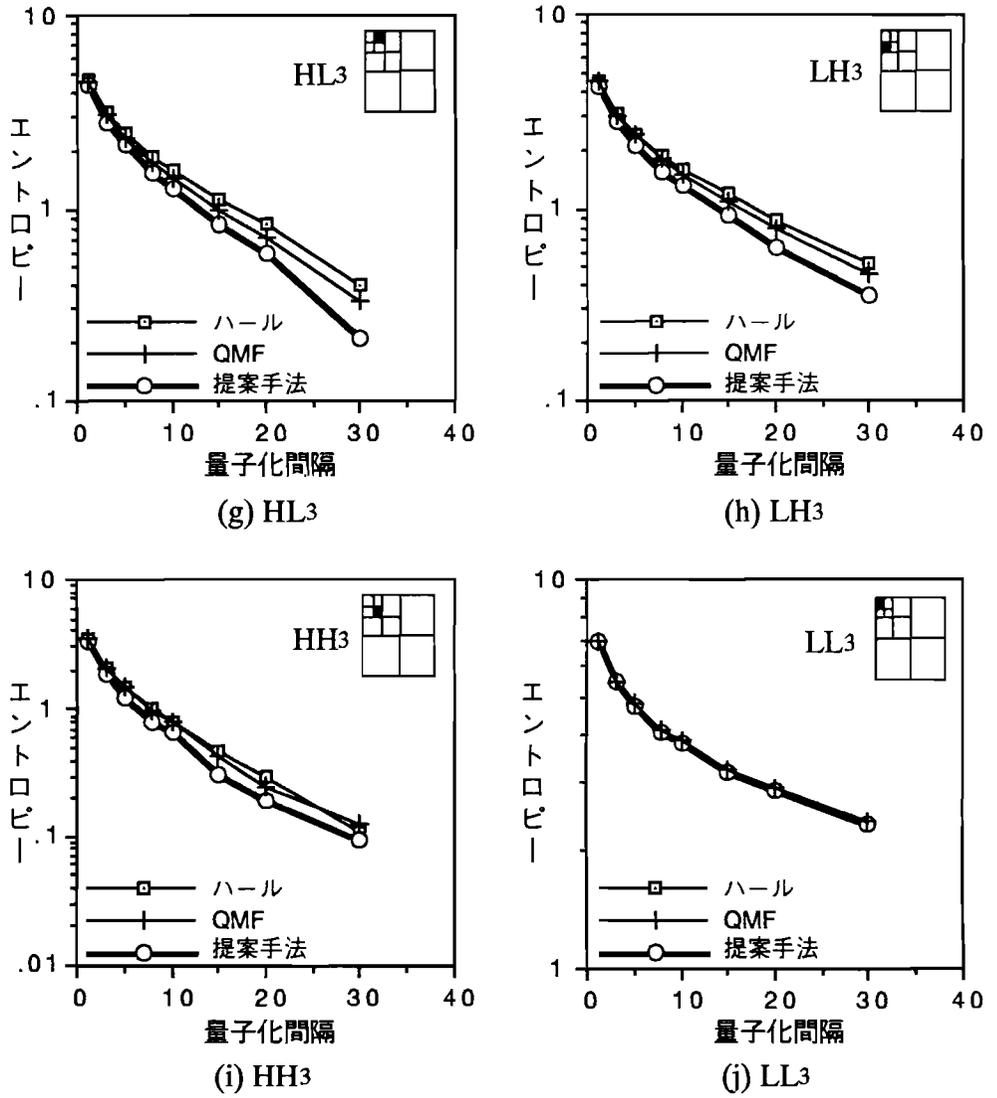


図 5.11 帯域別の信号のエントロピー

量子化間隔，縦軸はエントロピーである．エントロピーは信号の線形量子化を行った後に，信号が無記憶情報源から生じたものと仮定して算出を行った．DCTの直流成分に相当する帯域 LL3 においては，各手法ともほぼ等しい傾向を示している．その他の全ての帯域については，QMF および本変換手法がハール・ウェーブレット変換よりも低い値に達していることがわかる．特に本変換手法は QMF と比較するとフィルタのタップ数が少ないにもかかわらず，第1階層の信号のエントロピーは QMF とほぼ等しい値に，また第2階層，第3階層の信号においては QMF より低い値に達している．このことから本変換手法は少ない演算量で効率良くエントロピーを低減していると言える．

#### 5.4.2 圧縮比と再生画像

三つの手法の性能を比較するために，それぞれの手法による正変換後の信号を以下に述べる手法で符号化し，圧縮比，SNR 特性，そして再生画像の比較を行った．

いずれの手法を用いた場合も，帯域別に線形量子化と符号化を行った．量子化の際には図 5.12 に示したように第  $n$  階層の量子化間隔を  $q/2^n$  とするのが有効であることが報告されている [14]．本実験においてはこの量子化方法を用い， $q$  を可変パラメータとした．符号化手法としては，ゼロの連続個数および非ゼロの値をシンボルとするランレングス・ハフ

q/8	q/8	q/4	q/2
q/8	q/8		
q/4	q/4		
q/2		q/2	

図 5.12 各帯域の量子化間隔

表 5.4 連続個数による分類

グループ	連続個数	付加ビット数
1	1	0
2	2, 3	1
3	4 ~ 7	2
4	8 ~ 15	3
5	16 ~ 31	4
6	32 ~ 63	5
7	64 ~ 127	6
8	128 ~ 255	7
9	256 ~ 511	8
10	512 ~ 1023	9
11	1024 ~ 2047	10
12	2048 ~ 4095	11
13	4096 ~ 8191	12
14	8192 ~ 16383	13
15	16384 ~ 32767	14
16	32768 ~ 65535	15

マン符号化を用いた。ゼロの連続個数は表 5.4 に示したグループに分類し、グループ番号と連続個数を符号化する手法を用いた。

$q = 32, 64, 128$  の場合について再生画像の一部を拡大した結果を図 5.13 に示す。括弧内の数値は圧縮比である。ここで圧縮比とは、生成コードと全ての符号化テーブルの合計ビット数と、原画像データの総ビット数の間の比率である。ハール・ウェーブレット変換は他に比べて圧縮比が低いことがわかる。また量子化係数を大きくするとブロック歪みが目立つようになる。これに対して QMF と本変換手法は圧縮比が高く、ブロック歪みもほとんど見られない。

SNR 特性を図 5.14 に示す。図の横軸は符号化ビットレート、縦軸は SN 比である。同じ圧縮比の下では本変換手法および QMF はハール・ウェーブレット変換に比べて SN 比が 1 dB から 2 dB 程度上回っている。本変換手法は QMF と比較すると視覚的にも、また SNR 特性についてもほぼ等しい特性を示すことがわかる。



ハール (1/11.3)



8タップ QMF(1/15.3)  
(a)  $q = 32$



提案手法 (1/16.4)



ハール (1/23.3)



8タップ QMF(1/30.5)  
(b)  $q = 64$



提案手法 (1/34.0)



ハール (1/51.0)



8タップ QMF(1/62.0)  
(c)  $q = 128$



提案手法 (1/70.2)

図 5.13 量子化間隔と再生画像の関係

### 5.4.3 プログレッシブ・ビルドアップ

静止画像の通信等で用いられているプログレッシブ・ビルドアップ方式は、受け手側になるべく速く良好な画像を伝送するための手法である。三つの各手法において低域信号から順にプログレッシブ・ビルドアップで再生を行った結果を図 5.15 に示す。最も上の画像が LL3(x,y) 信号のみからの再生画像である。続く画像はこれに加えて第3階層、第2階層、そして第1階層の高域信号を加えて再生した結果である。ハール・ウェーブレット変換では再生途中の画像にブロック歪が目立つ。これに対して本変換手法は QMF と同様にブロック歪みが見られない。このように本変換手法は視覚的にも良好な画像が再生でき、しかもすでに述べたように QMF を用いた場合に比べて高速処理が可能であることから、画像通信などに有効に利用できると思われる。

## 5.5 結言

本章では帯域間予測を用いたハール・ウェーブレット変換を提案し、その特性を SSKF バンクを用いたサブバンド分割の考え方に基づいて詳

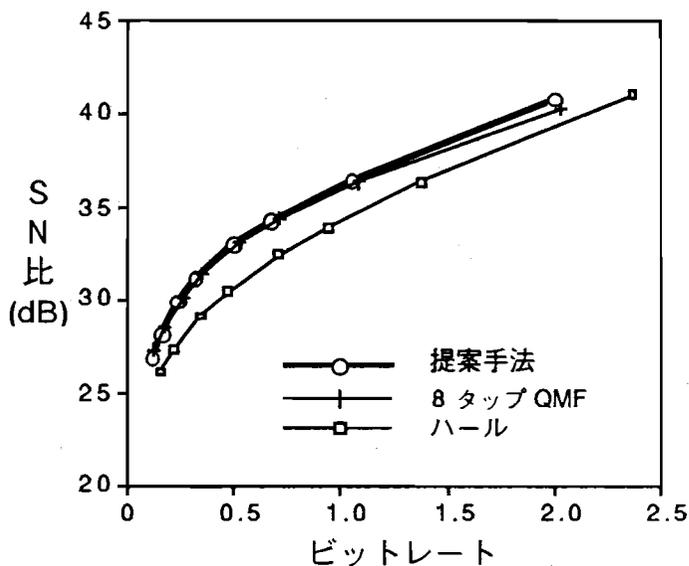


図 5.14 SNR特性



(a) ハール

(b) 8 タップ QMF

(c) 提案手法

図 5.15 プログレッシブ・ビルドアップによる再生

しく述べた。実験では画像信号の帯域分割符号化を行い、本変換手法がハール・ウェーブレット変換に比べて高域信号のエントロピーを低く抑えること、そして高能率な符号化を実現することを示した。また様々な圧縮比に対する再生画像、プログレッシブ・ビルドアップによる再生画像などを示し、本変換手法が QMF に比べて少ない演算量で、ブロック歪みの少ない視覚的に良好な画像を再生できることを示した。これらのことから本変換手法は符号化に有効であり、かつ高速処理が可能な優れた変換手法といえる。

## 第6章 結 論

本論文は画像データの圧縮技術に関して、

- (1) 圧縮の効率を向上させること、
- (2) 画像の品質を向上させること

を目的として研究を行った。これらの目的のもとに、四つの研究を行い、それぞれの研究において従来手法を改善する大きな成果をあげることができた。

第2章ではモノクロ画像の可逆符号化の研究を行った。従来の予測符号化は冗長抑圧の点では必ずしも最適な手法ではなかったが、本章では(1)ダイナミックレンジ変換法、(2)動的予測法、を提案することにより、冗長抑圧能力を改善することができた。その結果、JPEGの手法によって生成されたコード量をさらに15%削減することができた。

第3章ではカラーRGB画像の可逆符号化の研究を行った。カラーRGB信号はYUV信号に変換すると打ちきり誤差を生じるため、有効な可逆符号化法が行われていなかった。本論文ではこの解決法として、RGB信号を直接予測符号化する3次元予測法を提案し、画素間の冗長性とRGB信号間の冗長性を同時に取り除いた。その結果、2次元予測によって得られた誤差信号のエントロピーを、さらに7%削減することができた。

第4章では予測誤差信号のための符号化について研究を行った。本章では符号化における従来の確率推定法の弱点として、(1)雑音によって推定が不安定になる点、(2)局所的な確率の変動に追従しない点、を指摘し、これらの点を改善する新しい確率推定法を提案した。実験の結果、提案手法は従来の動的算術符号化やJPEGの算術符号化に比べて5%以上少ないコードを生成することができた。

第5章ではモノクロ画像の非可逆符号化の研究を行った。従来の手法

はブロック歪みの発生や処理の複雑さの点で問題があった。これに対し本章では帯域間予測を用いたハール・ウェーブレット変換を提案することにより、ブロック歪が生じない高速変換手法を実現した。さらに本予測処理は SSKF フィルタと等価であることを証明し、提案手法が完全再構成可能なサブバンド符号化の一つであることを示した。実験では従来のハール・ウェーブレット変換に伴うブロック歪みを除去し、圧縮コード量を 30% 削減することができた。そして少ない演算量で QMF と同等の特性をもつことを示した。

以上のように本論文は圧縮率の向上と画像品質の向上という当初の目的を達成することができた。これらの成果はテレビ電話／会議、カラーファクシミリなどの通信メディアの高速化に、また画像データベース、ビデオディスクなどの蓄積メディアにおける高密度化に非常に有効であると考えられる。

本論文で提案した全ての手法において、予測係数や分散などのパラメータを数学的に最適化したため、浮動小数点演算に多くの時間を費やしている。しかし実用上は予測係数などのパラメータを必ずしも理論的根拠に基づいて算出する必要はないと思われる。むしろ提案手法は係数テーブルの調整などを行い、高速化を行うべきである。今後は ISDN や光通信などの進展に伴い、大容量通信の時代が到来すると思われる。そのような状況の下では圧縮率のみならず、通信速度に合ったより高速な圧縮処理が要求されるだろう。実用化を目指した高速処理アルゴリズムの研究開発は今後の重要な課題である。

## 付 録

最小二乗法を用いた予測係数の決定法は次の通りである。

座標 $(x,y)$ における画素値を $f(x,y)$ 、その近傍の $N$ 個の参照点を $f_1, f_2, \dots, f_N$ とする。参照点の座標は予測器に依存する。 $\mathbf{F}(x,y)$ は1個の定数と $N$ 個の参照点からなるベクトルであり、

$$\mathbf{F}(x,y) = (1 \ f_1 \ f_2 \ \dots \ f_N) \quad (\text{A1})$$

とする。 $\mathbf{c}$ はそれらに対する予測係数からなるベクトルであり、

$$\mathbf{c} = (c_0 \ c_1 \ c_2 \ \dots \ c_N) \quad (\text{A2})$$

とする。このとき予測値 $\hat{f}(x,y)$ は

$$\hat{f}(x,y) = \mathbf{F}(x,y) \mathbf{c}^T \quad (\text{A3})$$

と表される。ある領域 $A$ 内の予測誤差の二乗和 $E$ は

$$E = \sum_{(i,j) \in A} \{ f(i,j) - \mathbf{F}(i,j) \mathbf{c}^T \}^2 \quad (\text{A4})$$

となる。上の $E$ を最小にするような $\mathbf{c}$ は

$$\sum_{(i,j) \in A} \mathbf{F}^T(i,j) \cdot \mathbf{F}(i,j) \cdot \mathbf{c}^T = \sum_{(i,j) \in A} \mathbf{F}^T(i,j) \cdot f(i,j) \quad (\text{A5})$$

を解くことにより求めることができる。

以上のようにして予測係数を決定する。

## 謝 辞

本論文は、神戸大学大学院自然科学研究科システム科学専攻平野浩太郎教授のもとで、筆者が博士課程在学中に行った研究成果をまとめたものである。研究の始終にわたりご指導とご鞭撻を賜った平野浩太郎教授に感謝の意を表す。

また本研究のための基礎知識をご教示下さり、本論文をまとめるにあたり貴重なご助言を賜った同専攻田中初一教授に深く感謝の意を表す。本研究の応用についてご討論下さり、本論文をまとめるにあたり貴重なご助言を賜った同専攻田中栄一教授に深く感謝の意を表す。

研究の成果をまとめるにあたり、さまざまなご指導とご助言を頂いた神戸大学大学院自然科学研究科システム科学専攻沼昌宏講師に深く感謝の意を表す。

本研究の全般にわたり、熱心なご討論とご助言を頂いた神戸大学工学部電気電子工学科野村孝徳先生、亀田昌志先生に心から感謝の意を表す。

最後に、日頃からさまざまな面でお世話になった神戸大学工学研究科澄川文徳氏、谷篤史氏をはじめとする神戸大学工学部電気電子工学科平野研究室の諸氏に心からお礼の意を表す。

## 参考文献

- [1] A. Habibi, "Survey of adaptive image coding techniques," IEEE Trans. Commun., vol. COM-25, pp. 1275-1284 (1977).
- [2] A. K. Jain, "Advances in mathematical models for image processing," Proc. IEEE, vol. 69, pp. 502-528 (1981).
- [3] 原島 博, 画像情報圧縮, オーム社 (1991).
- [4] "画像データ圧縮の理解と応用," インターフェース, no. 175, pp. 131-231 (1991).
- [5] Y. Linde, A. Buzo and R. M. Gray, "An algorithm for vector quantizer design," IEEE Trans. Commun., vol. COM-28, pp. 84-95 (1980).
- [6] N. M. Nasrabadi and R. A. King, "Image coding using vector quantization: A review," IEEE Trans. Commun., vol. COM-36, pp. 957-971 (1988).
- [7] H. Hang and J. W. Woods, "Predictive vector quantization of images," IEEE Trans. Commun., vol. COM-33, pp. 1208-1219 (1985).
- [8] N. Ahmed, T. Natarajan. and K. R. Rao, "Discrete cosine transform," IEEE Trans. Comput., vol. C-23, pp. 90-93 (1974).
- [9] M. Maruyama, H. Uwabu, I. Iwasaki, H. Fujiwara, T. Sakaguchi, M. T. Sun and M. L. Liou, "VLSI architecture and implementation of a multi-function, forward/inverse discrete cosine transform processor," Proc. Soc. Photo-Opt. Instrum. Eng., vol. 1360, pp. 410-417 (1990).
- [10] M. Bellanger, J. Daguët and G. Lepagnol, "Interpolation extrapolation and reduction of computation speed in digital filters," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-22, pp. 231-235 (1974).
- [11] M. J. T. Smith and T. P. Barnwell. , "Exact reconstruction techniques for tree-structured subband coders," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-34, pp. 434-441 (1986).
- [12] J. W. Woods and S. D. O'neil , "Subband coding of images," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-34, pp. 1278-1288 (1986).

- [13] C. K. Cheong, K. Aizawa, and T. Saito, "Subband image coding with biorthogonal waveletes," IEICE trans. Fundamentals, vol. E75-A, pp. 871-881 (1992).
- [14] J. Chen, S. Itoh, T. Hashimoto, "Scalar quantization noise analysis and optimal bit allocation for wavelet pyramid image coding," Trans. IEICE, vol. E76-A, pp.1502-1514 (1993).
- [15] A. Lempel and J. Ziv, "Compression of two-dimensional data," IEEE Trans. Inf. Theory, vol. IT-32, pp. 2-8 (1986).
- [16] Z. Lin and Y. Attikiouzel, "Two-dimensional linear prediction model-based decorrelation method," IEEE Trans. Pattern Anal. Mach. Intell., vol. PAMI-11, pp. 661-665 (1989).
- [17] S. Ranganath and A. K. Jain, "Two-dimensional linear prediction models-Part I : spectral factorization and realization," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-33, pp. 280-299 (1985).
- [18] P. A. Maragos, R. W. Shafer, R. M. Mersereau, "Two-dimensional linear prediction and its application to adaptive predictive coding of images," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-32, pp. 1213-1229 (1984).
- [19] D. A. Huffman, "A method for the construction of minimum-redundancy codes," Proc. IRE, vol. 40, pp. 1098-1101 (1952).
- [20] N. Abramson, *Information Theory and Coding*, McGraw-Hill (1963).
- [21] "データ圧縮アルゴリズムとその実現," インターフェース, no. 183, pp. 87-169 (1992).
- [22] J. Rissanen, and G. G. Langdon, "Arithmetic coding," IBM J. Res. Dev., vol. 23, pp. 149-162 (1979).
- [23] F. Rubin, "Arithmetic stream coding using fixed precision registers," IEEE Trans. Inf. Theory, vol. IT-25, pp. 672-675 (1979).
- [24] G. G. Langdon, "An introduction to arithmetic coding," IBM J. Res. Dev., vol. 28, pp. 135-149 (1984).
- [25] I. H. Witten, R. M. Neal, and J. G. Cleary, "Arithmetic coding for data compression," Communications of the ACM, vol. 30, pp. 520-540 (1987).
- [26] T. A. Welch, "A technique for high performance data compression," IEEE

- Computer, vol. 17, pp. 8-19 (1984).
- [27] 安田 浩, マルチメディア符号化の国際標準, 丸善 (1991).
- [28] International Organization for Standards / International Electrotechnical Commission (ISO/IEC), Digital Compression and Coding of Continuous-tone Still Image, International Standard 10918-1 (1993).
- [29] Sanjit K. Mitra. and James F. Kaiser , *Handbook for Digital Signal Processing*, John Wiley & Sons, New York (1993).
- [30] H. Gharavi and A. Tabatabai, "Application of quadrature mirror filtering to the coding of monochrome and color images," Proc. of ICASSP'87 pp.2384-2387.
- [31] T. Q. Nguyen and P. P. Vaidyanathan. , "Two-channel perfect-reconstruction FIR QMF structures which yield linear\_phase analysis and synthesis filters," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-37, pp. 676-690 (1989).
- [32] J. D. Johnston, "A filter family designed for use in quadrature mirror filter banks," Proc. ICASSP '80, pp.291-294 (1980).
- [33] D. Esteban and C. Galand, "Application of quadrature mirror filters to split band voice coding schemes," Proc. of 1977 IEEE Int. Conf. Acoust., Speech, Signal Processing, Hartford, CT, may 1977, pp. 191-195 (1977).
- [34] D. Le Gall and A. Tabatabai, "Sub-band coding of digital images using symmetric short kernel filters and arithmetic coding techniques," Proc. ICASSP '88, pp. 761-764 (1988).
- [35] 武田信之, 大竹孝平, " Wavelet変換における帯域間予測の検討" , 電子情報通信学会春季大会, 講演番号 D-330 (1992).
- [36] 渡辺哲史, 森川良孝, 浜田博, " 階層間電力相関を利用したサブバンド-ベクトル量子化" , 電子情報通信学会春季大会, 講演番号 D-320 (1992).
- [37] 酒澤茂之, 浜田高宏, 松本修一, " 2次元サブバンド間予測における符号化利得" , 電子情報通信学会春季大会 , 講演番号 D-248 (1993).
- [38] O. Rioul and M. Vetterli, "Wavelet and signal processing," IEEE Sig. Proc. Magazine, vol. 8, pp. 14-38 (1991).

- [39] 佐藤雅昭, "ウェーブレット理論の数学的基礎," 日本音響学会誌, vol. 47, pp. 405-423 (1991).
- [40] Charles K. Chui, *An Introduction to Wavelets*, Academic Press, INC (1992).
- [41] Stephane G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," IEEE Trans. Pattern Anal. Mach. Intell., vol. PAMI-11, pp. 674-692 (1989).
- [42] Stephane G. Mallat, "Multifrequency channel decompositions of images and wavelet models," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-37, pp. 2091-2109 (1989).
- [43] M. Ohta, M. Yano, and T. Nishitani, "Wavelet picture coding with transform coding approach," IEICE trans. Fundamentals, vol. E75-A, pp. 776-785 (1992).
- [44] R. B. Arps, T. K. Truong, "Comparison of international standards for lossless still image compression," Proc. IEEE, vol. 82, pp. 889-899 (1994).
- [45] N. Kuroki, T. Nomura, M. Tomita, and K. Hirano, "Lossless image compression by two-dimensional linear prediction with variable coefficients," IEICE trans. Fundamentals, vol. E75-A, pp. 882-889 (1992).
- [46] J. R. Treichler, *Theory and Design of Adaptive Filters*, John Wiley & Sons, New York (1987).
- [47] H. Yamaguchi, "Efficient encoding of colored pictures in R, G, B components," IEEE Trans. Commun., vol. COM-32, pp. 1201-1209 (1984).
- [48] N. Kuroki, T. Nomura, M. Kameda, M. Numa, and K. Hirano, "Lossless coding of digital RGB images based on three-dimensional linear prediction," Submitted to IEEE Trans. Circuit and Systems II.
- [49] 黒木修隆, 野村孝徳, 亀田昌志, 平野浩太郎, "勾配予測によるデジタルカラー画像の可逆符号化", 電子情報通信学会秋季大会, 講演番号 D-201 (1993).
- [50] N. Kuroki, T. Nomura, M. Numa, and K. Hirano, "Adaptive arithmetic coding of image prediction errors," Submitted to IEEE Trans. Circuit and Systems II.
- [51] 田中初一, "ラプラス分布をなす信号の最適符号", 電子通信学会論文誌, vol. J65-A, pp. 379-385 (1982).

- [52] 黒木修隆, 野村孝徳, 亀田昌志, 冨田昌宏, 平野浩太郎, ”帯域間予測を用いたハールウェーブレット変換と画像符号化への応用”, 電子情報通信学会論文誌, vol. J77-A (1994). (印刷中)