

画像符号化における
信号の圧縮と表現に関する研究

浅井光太郎

画像符号化における信号の圧縮と表現に関する研究

浅井光太郎

要旨

本論文は画像符号化における信号の圧縮と表現について、論文提出者自身が行った3つの研究を中心として考察を行った結果をまとめたものである。本論文は全5章から構成される。

第1章は研究の背景と目的について述べる。

画像は訴求力が強く、高度情報社会に必要なメディアであり、画像符号化は画像信号を伝送・蓄積するために必須の技術である。画像符号化において、圧縮は効率よい伝送・蓄積のための機能であり、表現は圧縮状態で何らかの情報の構造を表し、分析などの二次利用を可能とする機能である。本章では準備として画像符号化の基本的な技術について説明し、論文全体の構成と各章の要旨を述べる。

第2章はベクトル量子化を用いた画像符号化について述べる。

ベクトル量子化は複数のサンプルをブロック化して、これを多次元信号空間における入力ベクトルとして、量子化代表値である出力ベクトルに置き換える技術であり、画像では2次元の画素ブロックを入力ベクトルとする。ベクトル量子化は原理的には圧縮限界に漸近する性質を持ち、少ない符号量での画像符号化に適している。画像のベクトル量子化は1980年代に研究が本格化したもので、論文提出者は初期から同技術に取り組み、ベクトルから平均値を分離して正規化することで汎用的なベクトル量子化器を設計できることを示した。また正規化されたベクトルが画像の構造的なパターンを表し、画像を基本構造パターンの配列で表現可能であることを示した。さらに同技術を動画に適用し、64kbit/sの低ビットレートで動画の符

号化伝送が可能であることを実証し、実用化も行った。本章ではこれらを実現した技術と結果について述べる。

第3章は動的適応ウェーブレットパケットによる画像符号化について述べる。

ウェーブレット変換は信号を複数の周波数帯域に分割して扱う技術であり、1980年代終盤から画像符号化への適用が本格化したものである。ウェーブレット変換では基本的に低周波に信号電力が集中することを利用して効率のよい符号化を実現する。しかし現実には入力画像の電力分布によって効率が低下する。ウェーブレットパケットはウェーブレットの一般化であり、入力信号に適応した帯域分割を行うことが可能である。画像の場合にはさらにウェーブレットパケットを動的に適応させ、すなわち空間を分割して分割ごとにウェーブレットパケットを選択することで効率を最適化することができる。論文提出者は画像の帯域-空間分割に対する最適化を行い、最適な分割が良い画像符号化性能を与え、同時に信号電力の分布を可視化することを示した。本章ではこれらを実現した技術と結果について述べる。

第4章は動き補償予測付き DCT を用いる画像符号化標準について述べる。

動き補償予測付き DCT は、現在の国際標準動画符号化方式に採用されている方式である。同方式は基本的枠組みとしては 1990 年発行のテレビ会議向け標準 H. 261 に始まり、世代進化を続けて 2013 年策定の HEVC|H. 265 に到るまで使われている。論文提出者は 1980 年代終盤から標準化活動に参加して提案を行ってきた。例えば符号量の変動に対して安定動作を確保するバッファ制御方式、インターレース構造を持つテレビ信号の処理、低ビットレート向け動き補償ブロックの設定などについて研究を行って、標準方式に貢献してきた。本章では最新の HEVC|H. 265 標準への提案内容を中心にして、動き補

償予測付き DCT の改善技術について述べる。

第 5 章はまとめと今後の課題について述べる。

まず、第 2 章から第 4 章まで述べた技術について、現在の視点から振り返って考察を述べる。特に、産業にとって重要な国際標準について、現在も検討が続けられている符号化方式の現状とこれから進むべき方向について論文提出者の考えを述べる。最後に画像符号化技術の長期ビジョンについて、過去に提案された長期ビジョンや最近の研究動向を参照しつつ考察を述べて論文の結びとする。

目次

第 1 章 研究の背景と目的	1
1.1 はじめに	1
1.2 画像符号化について	3
1.2.1 画像符号化伝送システムのモデル	3
1.2.2 画像信号の情報量と符号化	4
1.2.3 画像符号化の要素技術	5
1.3 本論文の構成と概要	12
第 2 章 ベクトル量子化を用いた画像符号化	13
2.1 はじめに	13
2.2 画像信号のベクトル量子化	15
2.2.1 ベクトル量子化	15
2.2.2 画像信号の統計的性質	16
2.2.3 平均値分離正規化ベクトル量子化	19
2.3 ベクトル量子化器の構成と符号化性能	22
2.3.1 全探索ベクトル量子化器の構成	22
2.3.2 木探索ベクトル量子化器	23
2.3.3 木探索ベクトル量子化器の符号化性能	25
2.3.4 振幅補正と Gain/Shape 型ベクトル量子化	27
2.4 コードテーブルの生成	28
2.4.1 一般化 Lloyd アルゴリズム	28
2.4.2 クラスタリングにおける初期設定の問題	31
2.4.3 トレーニング系列の選別と拡大による汎用化	32
2.4.4 基本パターンによるモデル表現と認識適用の可能性	32
2.4.5 対称性を導入した木探索用コードテーブルの設計	36
2.5 低ビットレート動画像符号化への展開	40
2.5.1 カラー画像のベクトル量子化	40

2.5.2	動画像のダイナミックベクトル量子化	40
2.5.3	ダイナミックベクトル量子化による符号化性能	42
2.6	実時間符号化装置による実証と実用化への展開	43
2.7	この章のまとめ	45
第3章	動的適応ウェーブレットパケットによる画像符号化	46
3.1	はじめに	46
3.2	ウェーブレットによるサブバンド符号化	48
3.2.1	画像のウェーブレット変換	48
3.2.2	ウェーブレットパケット	49
3.2.3	ウェーブレット基底の最適性	50
3.3	適応的ウェーブレットパケット	54
3.3.1	符号量-歪み特性が最適なウェーブレットパケット	54
3.3.2	動的に適応するウェーブレットパケットの設計	56
3.4	ウェーブレットパケットによる画像符号化	58
3.4.1	ゼロツリーの一般化	58
3.4.2	量子化器の決定と符号語の割当	61
3.5	分割結果と符号化性能	62
3.5.1	周波数帯域-空間分割の最適化結果	62
3.5.2	周波数帯域-空間分割およびゼロツリーの最適化結果	63
3.6	周波数帯域と空間の分割による信号表現	67
3.6.1	試験信号の最適分割による検証	67
3.6.2	符号量-歪み特性のトレードオフと最適分割	69
3.7	この章のまとめ	71
第4章	動き補償予測付き DCT を用いる画像符号化標準	73
4.1	はじめに	73
4.2	国際標準化の歴史	75
4.3	動画像符号化国際標準への貢献	77

4.3.1	独自方式から H.261 へ	77
4.3.2	MPEG-1 から MPEG-2 へ	79
4.3.3	MPEG-4 から AVC H.264 へ	81
4.3.4	符号化方式の複雑化と性能の向上	83
4.4	HEVC H.265 への提案	84
4.4.1	方式検討の前提	84
4.4.2	提案方式の概要	85
4.4.3	符号化性能	96
4.5	符号化性能と複雑度の関係	98
4.5.1	符号化性能と複雑度	98
4.5.2	符号化性能と複雑度のトレードオフ改善	100
4.6	この章のまとめ	102
第 5 章 まとめと今後の課題		103
5.1	画像符号化研究の成果とその後	103
5.2	産業技術としての画像符号化	106
5.3	画像符号化のビジョン	108
付録 2 分割フィルタバンクによる局在化利得について		115
謝辞		119
参考文献		123
主要研究業績		140

目次

1-1	シャノンの通信モデルにおける符号化	4
1-2	フレーム内予測に用いる予測器の例	7
1-3	フレーム間予測	7
1-4	8×8ブロックに対する2次元DCTの基底関数	9
1-5	サブバンド符号化の系 (N=2の例)	10
1-6	ベクトル量子化の概念図	11
1-7	パターンマッチング符号化としてのベクトル量子化	11
2-1	歪み $d(\mathbf{x}, \mathbf{0}) = \text{const.}$	16
2-2	平均値分離後の画像信号分布	17
2-3	画像信号のブロック内標準偏差の分布	18
2-4	ブロックサイズと σ_l の画像内標準偏差の関係	18
2-5	平均値分離後の画像信号の相関	19
2-6	シミュレーションに用いた画像	19
2-7	平均値分離正規化ベクトル量子化の概念図	21
2-8	全探索ベクトル量子化符号化器の構成	23
2-9	2進木に対応する信号空間の分割	24
2-10	2進木探索ベクトル量子化器の第n+1段符号化部	25
2-11	AERIALのベクトル量子化符号化性能	26
2-12	AERIALのベクトル量子化符号化結果 (N=2048)	26
2-13	振幅補正の説明図[2-39]	27
2-14	一般化Lloydアルゴリズムのフローチャート	29
2-15	一般化Lloydアルゴリズムによる平均値分離正規化ベクトル量子化コードブック生成のクラスタリングにおける収束曲線	29
2-16	平均値分離正規化されたコードブックの例	30
2-17	基本パターンのセットによる画像信号源のモデル	35
2-18	対称性を持つ2進木に対応する出力ベクトル群	36
2-19	対称折り返しを行う2進木出力ベクトルの生成	38
2-20	正規化ベクトルの対称折り返し	39

2-21	対称性を用いて生成した 2 進木探索用出力ベクトル	39
2-22	ダイナミック多段ベクトル量子化の構成	41
2-23	ダイナミック多段ベクトル量子化の符号化性能	42
2-24	世界初の実時間動画ベクトル量子化符号化装置	
3-1	2 分割フィルバンクによるウェーブレット分解	48
3-2	2 次元のウェーブレット分解 (L: 低周波、H: 高周波)	49
3-3	(a)は Barbara 原画 (512×512×8)、(b)は 2 次元ウェーブレット分解された Barbara を可視化したもの	49
3-4	ウェーブレットパケットの概念図	51
3-5	1 次マルコフ過程 ($\rho = 0.9$) に対する刈込みと局在化利得	52
3-6	相関係数 $\rho = 0.9, 0.1$ の 1 次マルコフ過程を M バンドに分解する時に局在化利得を最大化する部分木 (左が低周波側)	53
3-7	符号量-歪み特性の最適化	55
3-8	刈り込みの決定	56
3-9	2 つの木に対応する時間-周波数平面の分割	57
3-10	4 進木の空間 (領域) 分割	57
3-11	DCT 係数の符号化における EOB	58
3-12	ゼロツリーの一般化	59
3-13	係数を走査する方向のグループ区分	60
3-14	係数の走査順序	60
3-15	(a) Barbara の分割、(b) Barbara の復号画像	62
3-16	ゼロツリーの最適量子化を含むウェーブレットパケットの最適化による符号化系	65
3-17	動的適応ウェーブレットパケットと SFQ による分割結果	67
3-18	CZP と動的適応ウェーブレットパケットによる分割	68
3-19	画像 Barbara に対する動的適応ウェーブレットパケット	70
4-1	動画符号化に関する標準化組織	74

4-2 MPEG 動画符号化標準の進展	75
4-3 動画符号化に関する MPEG 標準と ITU-T 勧告	76
4-4 静止画符号化に関する JPEG 標準ファミリー	77
4-5 DCT とベクトル量子化 (VQ) を包含する枠組み	77
4-6 H.261 符号化器のブロック図	79
4-7 MPEG-1 符号化器のブロック図	80
4-8 MPEG-2 符号化器のブロック図	80
4-9 MPEG-4 符号化器のブロック図	80
4-10 AVC H.264 符号化器のブロック図	82
4-11 動き補償予測の高度化と性能の向上	83
4-12 実験に用いたシーケンス	85
4-13 階層的なブロック分割 (動きベクトルの割当単位)	86
4-14 非矩形を含むブロックの分割	87
4-15 画像 Traffic に対する階層的ブロック分割の導入効果	87
4-16 動きベクトルに対する 2 種類の予測	88
4-17 ダイレクトモードにおける動きベクトルの候補と予測ブロック	89
4-18 MPEG-4 AVC H.264 のイントラ予測	90
4-19 符号化側におけるピラミッド型イントラ予測のブロック図	91
4-20 復号側におけるピラミッド型イントラ予測のブロック図	92
4-21 ピタミッド型イントラ予測 (BPP) による改善効果	94
4-22 提案方式による符号化性能 (1/2)	97
4-23 提案方式による符号化性能 (2/2)	98
4-24 既存国際標準から参照した符号化性能と複雑度の評価	99
5-1 MPEG-4 によるシーンの表現例	110
5-2 学習後のニューロンの可視化例	112
5-3 人工的な 1 次元信号源のモデルと波形	113

表目次

1-1 画像符号化の代表的な要素技術	6
2-1 異なるトレーニング系列を用いて生成されたコードブックによる符号化特性 (N=256 の場合)	21
3-1 符号語の割当	61
3-2 動的適応ウェーブレットパケットによる符号化結果	62
3-3 符号化性能の比較 ([3-21]の TABLE H より)	66
4-1 動画符号化標準の比較 (複雑度の増大)	83
4-2 ピラミッド型イントラ予測の改善効果 (QP=50, 46, 42, 38)	93
4-3 ピラミッド型イントラ予測の改善効果 (QP=34, 30, 26, 22)	93
4-4 提案方式による符号化性能 (AVC H.264 との比較)	97
5-1 画像符号化方式の世代区分 ([5-19]より)	108

第1章 研究の背景と目的

1.1 はじめに

画像情報は受容者に大きな訴求力を持つ情報メディアであり、画像情報の活用によってメッセージを伝え、未来に残すことができる。受容者は画像情報によって学び、楽しむことが可能になる。画像情報は膨大な量であるため、これを圧縮形式で表現して伝送・蓄積する画像符号化技術は豊かな高度情報社会に必須の技術である。

日本の状況を振り返ると、テレビ放送は1996年にCS放送、2000年にBS放送、2003年に地上波放送がデジタル化され、HDTVが日常のものとなった。2016年には8K（水平画素数7680を表す）スーパーハイビジョンの試験放送が開始された。2001年に開始された第3世代携帯電話サービスでは個人ベースのテレビ電話や映像サービスが可能になり、今日スマートフォンでストリーミング動画を視聴することは日常である。多くの会員を擁するSNSサービスでは動画像クリップが日々更新されている。モバイルに限らず高速インターネットが普及した今日、ストリーミング動画サービス業者が会員数を伸ばし、HDTVや4K動画の配信サービスを行なっている。1996年に商用化されたDVDビデオはアナログのVHSビデオソフトを駆逐し、2003年にはブルーレイディスクレコーダが登場し、今日デジタル配信がパッケージ販売を凌駕しつつある。このように動画像を活用するサービスは巨大な産業となり、社会を支える要素となった。この基盤となる画像符号化技術には長い研究開発の歴史がある。以下、筆者の関わりを含めて簡単に振り返る。

1930年代から1940年代にかけて各国でアナログのテレビ放送が開始され、動画像を遠隔地に伝達することが現実になった。1950年代から1960年代にはベル研でテレビ電話を想定してテレビ信号の性質や圧縮技術が研究された[1-7]。1970年代に入るとデジタル化した信号に対する予測符号化など高能率符号化の研究が行われ[1-9]、動き補償予測の概念が提案された[1-10,1-11,1-12]。変換符号化について画像信号に対するDCTの優位性が示された[1-15]。業務

用途のテレビ会議を目的とした開発や製品化も行われた。当時はテレビ会議用途の画像符号化伝送に 6~7Mbit/s の伝送速度を要した。

1980年代に入ると計算機シミュレーションの環境が整い、一般公衆向けデジタル回線サービス ISDN (64kbit/s~1.5Mbit/s、欧州では 2.0Mbit/s まで) の開始が視野に入ってテレビ会議・テレビ電話向け画像符号化の研究が活発化した。1.5Mbit/s のテレビ会議システムが実用化された。筆者は当時、研究が本格化する前夜にあった画像のベクトル量子化に取り組み、静止画像および 64kbit/s 低レート動画像符号化のための方式研究を行い、装置化および製品化まで関わった。この研究について第 2 章で述べる。

1980年代はまた画像符号化に関する国際標準化が始められた時期でもあった。1984年には予測符号化を用いる H.120[4-4] (1次群: 1.5Mbit/s または 2.0Mbit/s) が勧告化され、1990年には動き補償予測と DCT による変換符号化のハイブリッド方式を用いる H.261[4-5] ($p \times 64\text{kbit/s}$: $p=1 \sim 30$) が勧告化された (原型は 1988年に作成)。以降、動画像符号化の国際標準方式は H.261 を原型として改善を重ねていくこととなる。

動画像符号化国際標準方式の開発と並行して別の流れの研究も行われた。その中で重要な一つが 1980年代終盤から 1990年代にかけて盛んに研究されたサブバンド符号化、特にウェーブレット符号化である。ウェーブレット符号化は動画像符号化の国際標準には採用されていないが静止画符号化の標準 JPEG2000[4-13]に採用されている。筆者は当時、ウェーブレットを応用した符号化の最適化に取り組む機会を得て、圧縮符号化の効率を最適化する適応ウェーブレットパケットの研究を行った。この研究について第 3 章で述べる。

H.261 を原型とする動画像符号化国際標準の流れは一般ユーザ向けサービスに適用された規格としては 1993年の MPEG-1[4-6]、1995年の MPEG-2 | H.262[4-7]、1996年の H.263[4-8]、1999年の MPEG-4 Visual[4-9]、2003年の MPEG-4 AVC | H.264[4-10]、2013年の HEVC | H.265[4-11]へと続く。筆者は H.261 標準化の終盤から

参画し、初期にはベクトル量子化による動画コーデックの研究開発で得た知見により、後には既存標準方式の改善という視点で、共同研究者とともに継続的にこれらの標準化に提案を行ってきた。この研究について第4章で述べる。

本論文は上に述べた画像符号化に係る研究開発の歴史の中で筆者が行った研究をまとめるものである。画像符号化の研究において筆者が関心を持つ特性は、実応用に最も重要な圧縮 (compression) を第1として、もう1つは表現 (representation) である。表現とは各符号化方式がいかなるモデルで画像信号を記述しているかという側面である。情報を的確なモデルで構造的に表現できれば、分析・分類や加工等、二次的な機能性が高まる。本章の以下の節では、まず準備として画像符号化伝送システムと画像符号化の要素技術について整理し、次いで本論文の構成について述べる。

1.2 画像符号化について

1.2.1 画像符号化伝送システムのモデル

1.1 節で述べたように、動画像の圧縮を目指す符号化技術の開発は、初期において大容量の動画像データを狭帯域の通信回線で伝送する必要性によって発展した。従って動画像を符号化して伝送するシステムをシャノンの情報通信モデルに対比して考えることで位置付けを明確にすることができる[1-1,1-32,1-35]。

送信側ではまず撮像された画像信号に前処理を施して画像信号系列を形成する。同信号系列に対して予測・変換などの処理によって冗長性を分離した後、量子化レベルの密度を調整して情報圧縮を行う。さらに統計的性質に基づくエントロピー符号化を行い、入力された信号系列を符号化系列に変換する。ここで述べた信号系列から符号化系列への変換が情報源符号化である。符号化系列は暗号化や他の情報との多重化、誤り訂正符号の付与などの通信路符号化を経て伝送路に送出される。受信側では送信側と逆の手順によって動画像データを復号する。

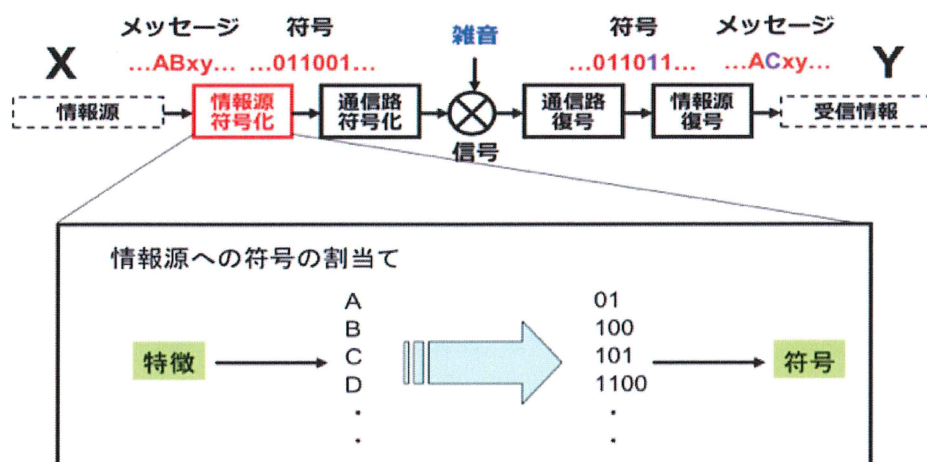


図 1-1 シヤノンの通信モデルにおける符号化

通常、画像符号化という場合には情報源符号化を指す。一方、国際標準 H.261 は情報源符号化だけでなく、BCH(511,493)誤り訂正符号などの通信路符号化を包含している。これは H.261 が ISDN への適用を前提にした規格であり、システム全体の性能が重視されたからである。後年の画像符号化は用途を特定しない方向に向かったため、検討対象は情報源符号化に限定され、国際標準の検討においても通信路符号化とのインターフェースが抽象化される方向となった。本論文では情報源符号化の問題に限定するが、最終的なシステム設計にはシステム全体の合目的性が確認されなければならないことを確認としてここに明記しておく。

1.2.2 画像信号の情報量と符号化

今、デジタル化された画像があり、その画素数が $M \times N$ 、各画素の量子化レベル数が L とすると同画像の表現に必要な情報量は以下の通りである。

$$H = M \times N \times \log_2 L \text{ (bit)}$$

通常、各画素の量子化レベルは一様分布でないため、実効エントロピーを用いた画素あたりの情報量は画素レベル x_i の確率 $P(x_i)$ を用い

て以下で表される。

$$H(x_i) = -\sum_{x_i=1}^L P(x_i) \log_2 P(x_i) \quad (\text{bit/pel})$$

ただし現実の画像は近傍の画素および動画像であれば過去の画像フレームにおける近傍の画素との間に強い相関を持つ。従って画素あたりの実効的な情報量は近傍の状態に依存する条件付きエントロピーである。近傍画素の状態を S 、画素レベル x_i の条件付き確率を $P(x_i|S)$ とすると条件付きエントロピーは以下で表される。

$$H(x_i|S) = -\sum_{x_i=1}^L P(x_i|S) \log_2 P(x_i|S) \quad (\text{bit/pel})$$

常に $H(x_i|S)$ を最小とする状態 S を得ることができれば相関による冗長性を分離した最小の情報量で表すことが可能になる。

許容可能な歪みを前提として圧縮を行う規範となる概念が符号量-歪み理論 (Rate-Distortion theory) である [1-6]。信号系列の個々のシンボルに対して確率分布に基づく最適量子化を行い、実効エントロピーを用いても圧縮の限界には達しない。複数のシンボルをブロック化してブロックサイズを拡大した時に漸近する特性が符号量-歪み関数と呼ばれる圧縮限界である。分散 σ^2 のガウス性無記憶情報源に対して歪みを平均二乗誤差で表した場合の符号量-歪み関数は以下で表される [1-30]。

$$R(d) = \begin{cases} \frac{1}{2} \log_2 \frac{\sigma^2}{d} & (0 < d \leq \sigma^2) \\ 0 & (d > \sigma^2) \end{cases}$$

以上を踏まえ、信号系列間に存在する画像特有の相関を利用して圧縮効率の向上を図る技術が画像符号化、より詳細に言えば高能率符号化である。

画像符号化の符号量を表すには画素 (pel: picture element) あたりのビット量 (bit/pel) を用い、歪みには PSNR (Peak Signal Noise Ratio) を用いる。PSNR は SNR や S_{p-p}/N_{rms} とも書かれる。

$$\text{PSNR} = 20 \log_{10}(S_{p-p}/N_{rms})$$

8 ビット画像の符号化を行う場合、 S_{p-p} として定数 255 を用いる。

1.2.3 画像符号化の要素技術

表 1-1 に代表的な画像符号化の手法を示す。筆者が調べた範囲で現在用いられている画像符号化のほとんどは表 1-1 にある技術を組合せ、またはその展開を用いている。フラクタル符号化[1-22]はベクトル量子化 (VQ: Vector Quantization) との相似性および信号源である画像の自己相似性を記述するという観点で特徴抽出符号化に分類した。近い将来、表 1-1 に追加すべき可能性がある要素技術として、ウィナーフィルタや凸射影[1-23,1-24]など、従来は画像劣化の復元に用いられてきた手法の積極利用が考えられる。以下、代表的な手法について説明する。

表 1-1 画像符号化の代表的な要素技術

予測符号化	フレーム内予測	前値予測、線形予測、適応切替予測
	フレーム間予測	フレーム差分、動き補償、背景予測
変換符号化	直交変換	DFT、WHT、KLT、DCT、LOT
階層/サブバンド符号化	階層符号化	ビットプレーン、解像度ピラミッド
	サブバンド符号化	QMF、SSKF、ウェーブレット
ベクトル符号化	ベクトル量子化	正規化VQ、格子VQ、多段VQ
	特徴抽出符号化	輪郭線、等高線、フラクタル
エントロピー符号化	可変長符号化	ハフマン、ゴロム、ランレングス置換
	ユニバーサル符号化	算術符号、Ziv-Lempel符号、順列符号

DFT: Discrete Fourier Transform, WHT: Walsh-Hadamard Transform, KLT: Karhunen-Loeve Transform, DCT: Discrete Cosine Transform, LOT: Lapped Orthogonal Transform, QMF: Quadrature Mirror Filter, SSKF: Symmetric Short Kernel Filter

(1) 予測符号化

予測符号化は符号化すべき画素を近傍の符号化済み画素から予測し、予測誤差を量子化する方式である。予測精度を高めることによって予測誤差の分散最小化が図られる。予測符号化はDPCM[1-8,1-9]の応用であり、しばしばAR (Auto Regressive) モデルに基づく予測器が設計される[1-30]。ARモデルによる予測符号化は以下の式で表される。

$$x_i = -\sum_{l=1}^m a_l x_{i-l} + e_i$$

ここで x_i は信号系列、 a_i は AR 予測器の係数、 e_i は予測誤差である。信号系列が定常エルゴード過程と近似可能であれば、AR モデルは信号系列の有効なモデルである。画像信号は統計的に高い相関を持つため、総体として AR モデルに基づく予測が有効に機能する。しかし画像信号は本来非定常であるため、予測精度が低下した際には誤差が過大となる勾配過負荷雑音が発生する可能性がある。このため、しばしば予測器の動的適応手法[1-25]や予測誤差の適応量子化などの手法が用いられる。

基本的な画像信号の予測手法として、同一画面の中で近傍の画素を予測のために参照するフレーム内予測、符号化済みの別フレームから予測を行うフレーム間予測がある。

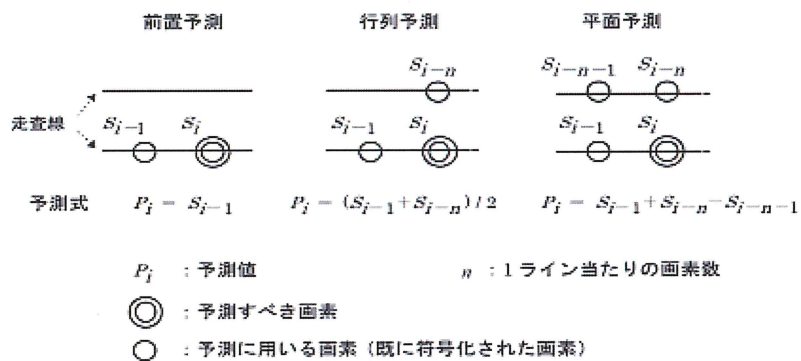
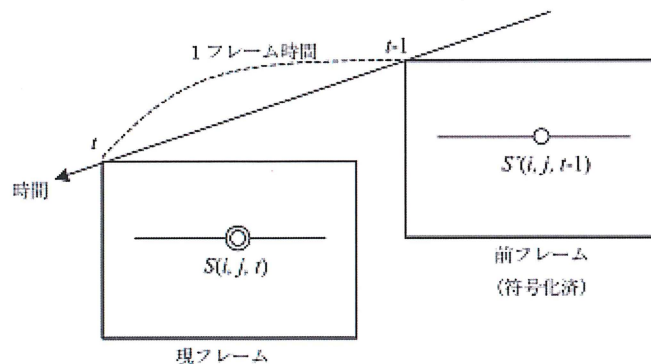


図 1-2 フレーム内予測に用いる予測器の例[1-27]



現フレームの画素 $S(i, j, t)$ の予測値を前フレームの同じ位置の符号化済画素 $S(i, j, t-1)$ とする

図 1-3 フレーム間予測[1-27]

フレーム間予測には動き補償予測が導入され、主流はブロック単位に動きベクトルによる参照フレームからの変位を割り当てる手法である。過去と未来から予測を行う双方向予測や端数精度の動きを想定した画素をフィルタで内挿するなどの改善が行われている。

(2)変換符号化

変換符号化は互いに大きな相関を持つ信号系列をブロック化して無相関な係数に線形変換し、特定の変換係数に電力が集中することを利用して、割り当てる符号量の配分を変えて全体の符号量・歪み特性の向上を図るアプローチである。

変換には離散フーリエ変換 (DFT)、ウォルシュ=アダマール変換 (WHT) などがある。無相関化の点で優れているのは信号に依存して導出されるカルーネン=レーベ変換 (KLT) [1-14] である。最もよく使われる変換は、画素間の相関が高い時の KLT に近い特性を持つ離散コサイン変換 (DCT) である [1-15]。 $N \times N$ の画素ブロックを $S(i,j)$ とすると、DCT 係数 $F(u,v)$ への変換は次の式で与えられる。

$$F(u,v) = \frac{2C(u)C(v)}{N} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} S(i,j) \cos \left[\frac{(2i+1)u\pi}{2N} \right] \cos \left[\frac{(2j+1)v\pi}{2N} \right]$$

$$\text{ただし、 } C(u), C(v) = \begin{cases} \frac{1}{\sqrt{2}} & (u,v=0) \\ 1 & (u,v \neq 0) \end{cases}$$

$F(u,v)$ から画素値 $S(i,j)$ への逆変換 IDCT は次の式で与えられる。

$$S(i,j) = \frac{2}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} C(u)C(v)F(u,v) \cos \left[\frac{(2i+1)u\pi}{2N} \right] \cos \left[\frac{(2j+1)v\pi}{2N} \right]$$

変換及び逆変換の処理は $N \times N$ 行列の乗算で実現される。

上記の変換と逆変換は画素ブロックを $N \times N$ の基底関数の線形和で表現することに相当する。 $N=8$ の場合、基底関数は図 1-4 に示す形となる。

DCT による変換符号化では係数の量子化に十分な符号量を割り当てられない時に発生する画質の劣化がある。ブロック歪みは隣接

画素ブロックとの境界に不連続が目立つ劣化であり、モスキート歪みは画像のテクスチャがエッジやステップを含む場合に電力が係数に分散することから発生する劣化である。こうした劣化を回避するための手法として、LOT (Lapped Orthogonal Transform) などブロックをオーバーラップさせる変換[1-17]が提案されており、静止画符号化に適用されている。(音響信号への適用に前例がある)

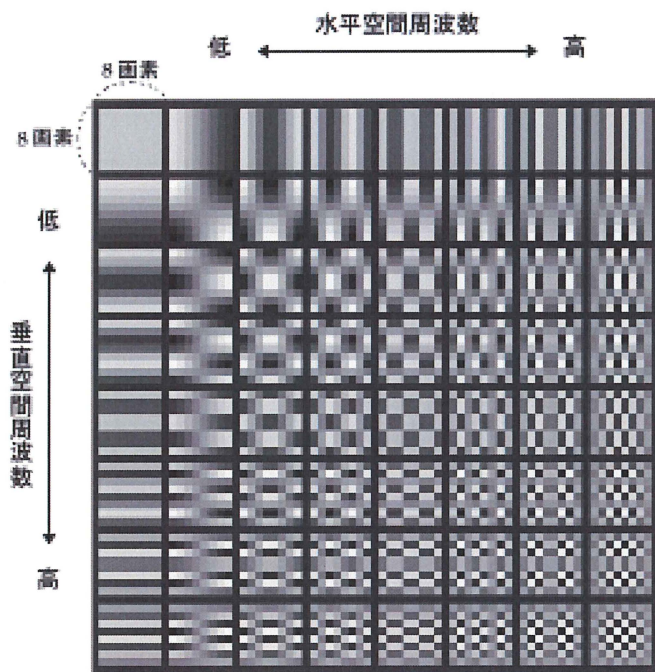


図 1-4 8×8 ブロックに対する 2 次元 DCT の基底関数

(3) サブバンド符号化

サブバンド符号化はフィルタバンクによって信号を N 個の帯域 (サブバンド) に分割し、各帯域の信号を $1/N$ にサブサンプルする。特定の帯域に電力が集中することを利用して、割り当てる符号量の配分を変えて全体の符号量・歪み特性の向上が実現される。復号時は各帯域を復号の上、 N 倍にアップサンプルした後に合成を行う。サブバンド符号化の系を図 1-5 に示す。

サブバンド符号化は基底関数への展開という意味で数学的には変換符号化と等価である[1-18]が、実現手法の相違から異なる符号化

特性を持つ。サブバンド符号化の利点は基底関数がブロックに閉じないためにブロック歪みが生じないこと、2分割フィルタバンクを低周波側に再帰的に適用することで得られるウェーブレット[1-19,1-20]のようにエッジの高周波成分には短い基底を持つフィルタバンクを用いればモスキート歪みにも有効との利点がある。ウェーブレットの活用に関する詳細な議論は第3章で行う。

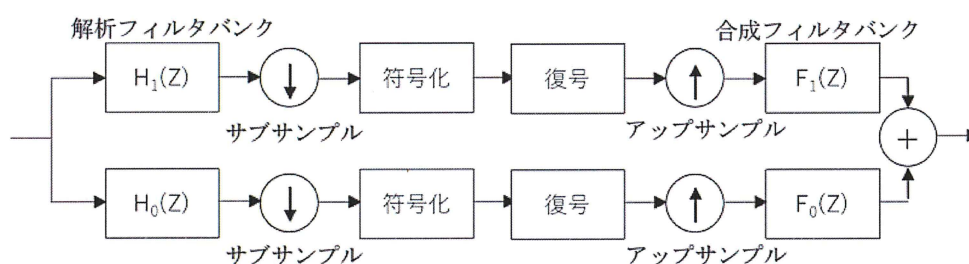


図 1-5 サブバンド符号化の系 (N=2 の例)

(4)ベクトル量子化

ベクトル量子化は複数のシンボルをブロック化し、これを多次元信号空間で量子化してインデックスを符号語とすることにより、系列間の相関を利用するとともにブロックサイズの拡大によって符号量-歪み特性の向上を利用するアプローチである。

ベクトル量子化に関する詳細な記述は第2章で行うこととし、ここでは図 1-6 を用いて概要の説明を行う。符号化すべき K 個のシンボル (x_1, x_2, \dots, x_K) を入力ベクトル \mathbf{x} と定義し、 \mathbf{x} を含む K 次元信号空間 R^K の N 個の分割を (R_1, R_2, \dots, R_N) 、部分空間 R_i の代表点である出力ベクトル $\mathbf{y}_i = [y_{i1}, y_{i2}, \dots, y_{iK}]$ のセットを $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N]$ とすると、ベクトル量子化は K 次元信号空間における入力ベクトル \mathbf{x} を最近傍にある出力ベクトル \mathbf{y}_i へ量子化することに対応する。この時 \mathbf{y}_i の添字 i が符号化出力となる。また図 1-7 のように K 個のシンボルが隣接する2次元の画素ブロックを入力ベクトルとする場合、ベクトル量子化はパターンマッチングによる符号化と直観的に捉えることもできる。

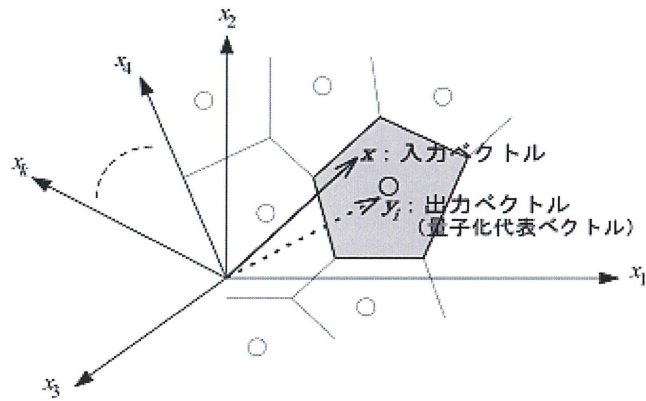


図 1-6 ベクトル量子化の概念図 [1-35]

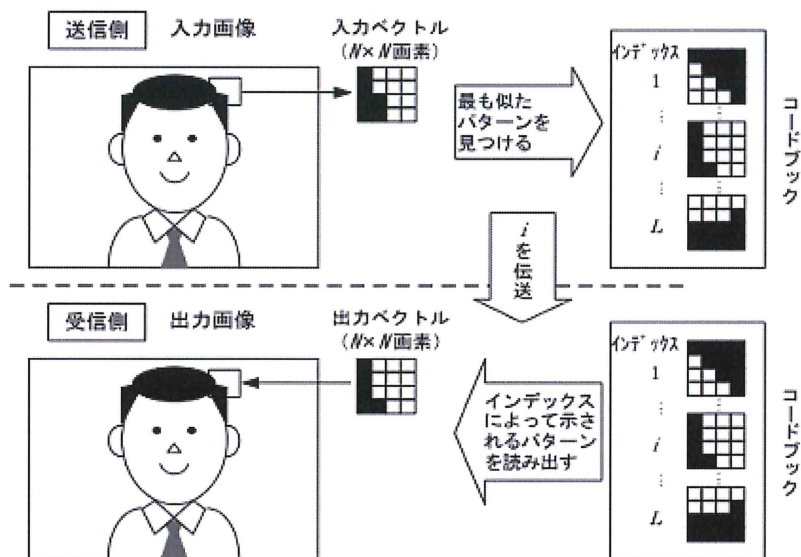


図 1-7 パターンマッチング符号化としてのベクトル量子化 [1-35]

(5) エントロピー符号化

エントロピー符号化ないし可変長符号化は符号化系列の統計的性質に基づいて可逆な符号を与える方式であり、通常は(1)~(3)の方式によって得られた符号化系列に対して最終的な符号語を割り当てて圧縮効率を高める過程で用いられる。基本的には有限個のシンボルを持つ系列に対して生起確率に対応して異なる長さの符号語を与

えて平均符号語長を最小化する。ハフマン符号が古くから使われているが、生起確率を想定した符号語テーブルが必要なことから、生起確率にラプラス分布などを仮定できる場合には規則的な構造を持つ符号も用いられる。また、生起確率を固定的に想定しないユニバーサルな符号化として、算術符号化も国際標準を含めて近年よく用いられる。

1.3 本論文の構成と概要

次章以降の本論文の構成は以下の通りである。第2章では画像信号のベクトル量子化について述べる。同章の基になっているのは1980年代初頭に行った研究であり、隣接する画素のブロックを入力ベクトルとし、これを平均値分離正規化した多次元信号空間することによって汎用性のある出力ベクトルのセットを生成できることを示す。さらに平均値分離正規化された出力ベクトル群が画像の基本構造パターンとも言い得る性質を有することについても触れる。第3章では2分割フィルタを用いて生成されるウェーブレット基底を含む木構造に対応した基底群から符号量-歪み特性的に最適な基底を動的に探索する手法について述べる。同章の基になっているのは1990年代初頭に行った研究であり、最適な基底を選ぶことで符号化性能が向上すること、選択された基底自身が信号源の性質を表現していることを指摘する。第4章では動き補償予測付きDCTを基本とする国際標準符号化方式の開発および改善について述べる。同章の基になっているのは1980年代終盤からほぼ現在に至るまで国際標準において筆者が行ってきた活動である。第5章では本論文のまとめと現時点での整理、今後の展望について述べる。

第 2 章 ベクトル量子化を用いた画像符号化

2.1 はじめに

量子化すべき信号系列が一様な分布を取るならば、固定的なステップサイズで量子化を行うことで最適な結果を得る。分布が一様でない場合、量子化による平均二乗誤差を最小にする最適量子化とは、入力系列に対して最近傍の量子化代表値が出力されるように量子化区間が設計されている最近傍条件と量子化代表値が量子化区間における信号系列の分布に対応した重心である重心条件の 2 つの条件を満たす量子化である。スカラ量子化の場合、これは Lloyd と Max によって定式化されている [2-1,2-2]。

ベクトル量子化は複数のサンプルをまとめてブロックを形成し、多次元空間において最も近傍にある量子化代表値への写像を行う量子化である。画像信号系列において系列間に強い相関があることは、これをブロック化した入力ベクトルの多次元信号空間における分布が一様でなく、これを分布関数で表現したとしても同関数は多次元信号空間の各座標軸に分離可能な形式にはならないことを示唆している。従って多次元信号空間において任意の量子化代表値を設定可能で、その近傍であるポロノイ領域を量子化区間としたベクトル量子化はスカラ量子化に対して優れた特性を持つ。また、仮に信号系列間に相関が全くない場合にも、平均距離を最小化する代表値の配置は座標軸に変数分離可能ではなく、ベクトル量子化はスカラ量子化に対して利得を持つ [2-19]。1.2.2 において複数のシンボルをブロック化したサイズを拡大していくと符号量-歪み関数に漸近するのはこの性質による。

ベクトル量子化の概念は古くから使われてきたブロック量子化 (block quantization) と表されていたが、時間ないし空間的に隣接する信号系列を直接ブロック化して多次元信号空間で量子化することを強調してベクトル量子化と呼ばれるようになった。

画像信号に対するブロック量子化という言葉が使われた初期の文献の一つ [2-4] では、線形変換を用いた画像符号化の符号化復号系に

おける量子化器としてブロック量子化が議論された。主な議論の対象はフーリエ変換 (DFT) やウォルシュ=アダマール変換 (WHT)、カルーネン=レーベ変換 (KLT) の特性であり、変換後の係数をむしろ無相関化して量子化することが主眼のブロック符号化であった。ここでは相関のあるマルコフ情報源に対する符号化効率が高いという特性は重視されていない。1977年の Habibi によるレビュー [2-6] では、適応的な画像符号化の分類として予測符号化、変換符号化、エントロピー符号化と並んでクラスタ符号化 (cluster coding) [2-5] が議論された。クラスタ符号化は、予測符号化や変換符号化が信号源を無相関化するのに対し、相関あるデータをグループ化して量子化する方式である。文献 [2-5] ではクラスタ符号化を LANDSAT マルチスペクトル画像の認識と圧縮 (recognition and compression) に適用する例が示された。注目すべき点は認識と圧縮の双方がクラスタ符号化の特徴とされている点である。このアプローチは本論文のテーマに繋がるベクトル量子化の原型と言える。

実質にベクトル量子化であるブロック量子化の研究は 1970 年代の終盤に主として音声信号を対象に圧縮と認識の両面で行われた [2-7,2-8]。国内では山田、田崎らによって画像への適用を行う先駆的な試みが報告された [2-25]。

1980 年、Linde らはベクトル量子化器の設計方式と音声信号への適用に関する論文を発表した [2-9]。同論文で提案された設計方式は Lloyd [2-1] および Max [2-2] の最適量子化器設計を多次元に拡張したものである。例えば文献 [2-6] において、クラスタ符号化は多次元信号系列をクラスタに分類し、各クラスタに属する信号系列の重心を求め、これを繰り返すことによって収束させると記述されていたように、初めてのことでなかったが、最適条件を含む設計手法として定式化されたのは初めてであった。村上と筆者は文献 [2-25] の著者から説明を受けて画像のベクトル量子化に関する研究を開始、1982 年 7 月に発表した [2-27]。これに前後して 4 月には [2-10]、10 月には [2-11] が発表され、画像のベクトル量子化が注目を集めるよ

うになった。

本章では次節以降、画像信号のベクトル量子化を数学的に記述し、画像信号の統計的性質を調べた上で平均値分離正規化ベクトル量子化を符号化器の構成とともに提案する。平均値分離正規化によって汎用的な出力ベクトルを生成することができること、その生成法についても示す。さらに動画像にベクトル量子化を適用し、高圧縮のビデオコーデックが実現可能であることを示し、実機の開発についても述べる。

2.2 画像信号のベクトル量子化

2.2.1 ベクトル量子化

今、 K 次元信号空間 R^K における入力ベクトルを $\mathbf{x} = (x_1, x_2, \dots, x_K)$ に対し、 R^K を N 個の部分空間 R_1, R_2, \dots, R_N に分割し、各部分空間 R_i の代表点である出力ベクトル $\mathbf{y}_i = [y_{i1}, y_{i2}, \dots, y_{iK}]$ のセットを $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N]$ 、 \mathbf{Y} のインデックスセットを $\mathbf{I} = [1, 2, \dots, N]$ とすると、ベクトル量子化 Q は符号化 C と復号化 D の接続として以下のように表される [2-12, 2-27, 2-35]。

$$Q = D \cdot C$$

$$C: \mathbf{x} \rightarrow i \text{ if } d(\mathbf{x}, \mathbf{y}_i) < d(\mathbf{x}, \mathbf{y}_j) \text{ for all } j \neq i$$

$$D: i \rightarrow \mathbf{y}_i$$

$$R_i \cap R_j = \emptyset \text{ for all } j \neq i$$

$$\cup_{\text{all } j} = R^K$$

C は入力ベクトル \mathbf{x} を出力ベクトル群 \mathbf{Y} の中で、歪み測度に従って最も近距離にある出力ベクトル \mathbf{y}_i のインデックス i への写像であり、 D はインデックス i から出力ベクトル \mathbf{y}_i への逆写像である。出力ベクトルのセット \mathbf{Y} をコードブックと呼ぶ。

この時符号化データの伝送レート T (bit/pel)は以下となる。次元数 K が増大するにつれ符号化特性は向上し圧縮限界に漸近する。

$$T = K^{-1} \log_2 N$$

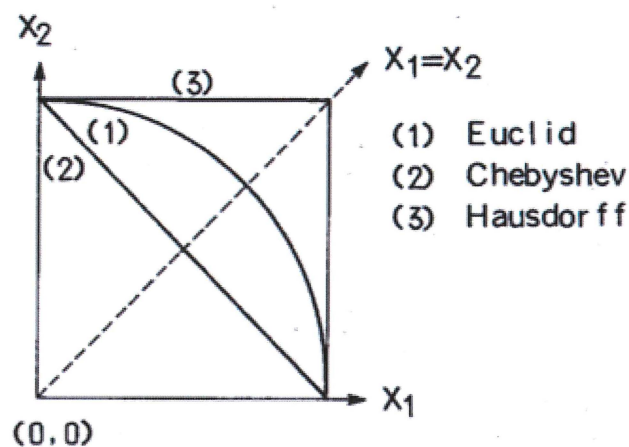


図 2-1 歪み $d(\mathbf{x}, \mathbf{0}) = \text{const.}$ [2-36]

歪み測度 $d(\mathbf{x}, \mathbf{y}_i)$ の例として以下が想定できる。

$$d(\mathbf{x}, \mathbf{y}_i) = \left[\sum_{j=1}^K |x_j - y_{ij}|^2 \right]^{1/2} \quad (\text{Euclid 距離})$$

$$d(\mathbf{x}, \mathbf{y}_i) = \sum_{j=1}^K |x_j - y_{ij}| \quad (\text{Chebychev 距離})$$

$$d(\mathbf{x}, \mathbf{y}_i) = \max_j |x_j - y_{ij}| \quad (\text{Hausdorff 距離})$$

上記の歪み測度を 2次元信号空間の場合の $d(\mathbf{x}, \mathbf{0})$ で表すと図 2-1 のようになる。歪みとしての一般性や数学的扱いの容易さでは Euclid 距離が優れており、後に述べるクラスタリングにおいて収束性が厳密に保証されるのは Euclid 距離の場合である。一方、そのほかの定義には計算時間や装置化上の利点があり、現実的には Chebychev 距離や Hausdorff 距離を用いたクラスタリングの場合も良好な収束性を得られることが確かめられている [2-27]。本論文では Euclid 距離を想定して議論を行うが、過去の研究成果として発表した実験または装置においては Chebychev 距離や Hausdorff 距離を用いた例もある。これらの場合はその都度明記する。

2.2.2 画像信号の統計的性質

画像信号の 2次元配列 $S_{m,n}$ を K 個毎にブロック化した小領域を A とする。画像をブロック分割したシーケンス番号を I として、 A における平均値および標準偏差を次のように定義する。 $E^A(\cdot)$ は A にお

ける期待値を表す。

$$\mu_l = E^A(S_{m,n})$$

$$\sigma_l = \left[E^A(S_{m,n} - \mu_l)^2 \right]^{1/2}$$

図 2-2 は 4×4 のブロックを単位として各画素から μ_l を分離した信号分布 $P(S_{m,n} - \mu_l)$ を画像 GIRL ($512 \times 512 \times 8$) と AERIAL SCENE ($512 \times 512 \times 8$ 、以下 AERIAL) について示している。 μ_l の分離により、両画像ともラプラス分布に近づいている。

図 2-3 は 4×4 のブロックについて σ_l の分布 $P(\sigma_l)$ を示している。2枚の画像は同じ形の分布をとるが、空間周波数成分の相違によって分散は大きく異なる。図 2-4 はブロックサイズをパラメータとして σ_l の分散をプロットしたものである。GIRL ではブロックサイズ 20×20 程度で飽和し、AERIAL ではブロックサイズ 6×6 程度で最大値をとっている。この分散が最大であるときに後述する正規化の効果が最大になると期待される。一方、ブロックサイズを大きくすることはベクトルの次元を拡大することであり、目標のビットレートを実現するために必要な出力ベクトル数が指数的に増大する。実時間の装置化（1980 年代初頭のデバイスのレベルが前提）を想定し、ブロックサイズを 4×4 として検討した。

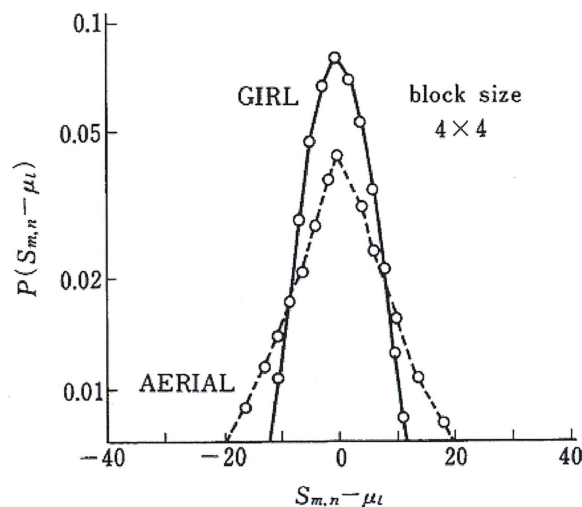


図 2-2 平均値分離後の画像信号分布 [2-35]

ブロックごとに信号から μ_l を分離し、 σ_l で正規化した後のブロック内相関 $R_{g,0}$ を以下に従って測定した値を水平方向について図 2-5 に示す。正規化後もブロック内に指数関数的な相関が残存していることがわかる。垂直方向にも同様なことが観測される。すなわち、 μ_l を分離し、 σ_l で正規化した後の画像信号は異なる画像に対して似通った分布をとる。この性質を利用すれば、出力ベクトルのセットの規模が膨大になることを避けることが可能になり、同セットを汎用化する効果がある。

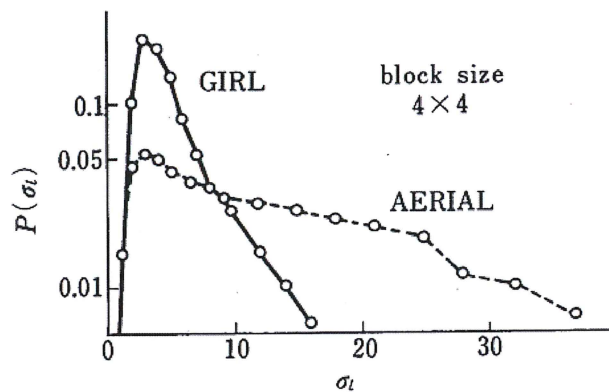


図 2-3 画像信号のブロック内標準偏差の分布 [2-35]

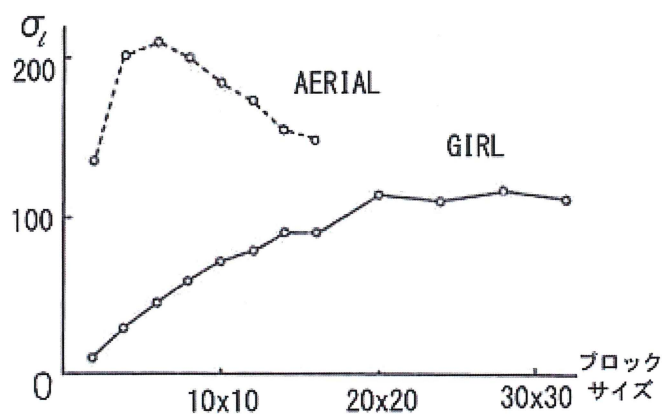


図 2-4 ブロックサイズと σ_l の画像内標準偏差の関係

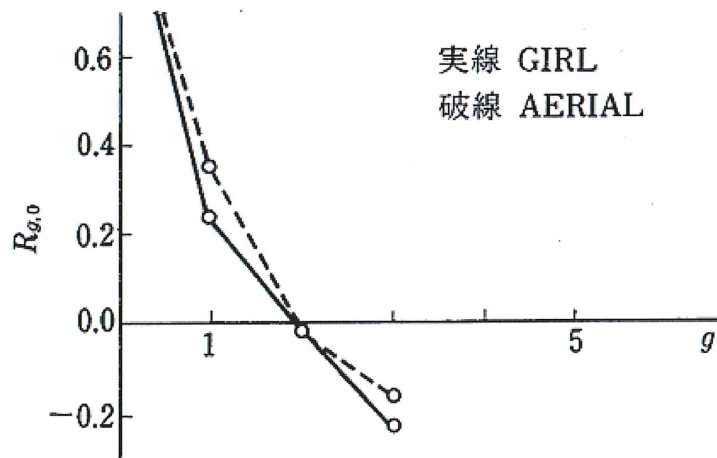


図 2-5 平均値分離正規化後の画像信号の相関[2-35]



図 2-6 シミュレーションに用いた画像 (GIRL と AERIAL SCENE)
(いずれも 512×512 画素 \times 8 ビット)

2.2.3 平均値分離正規化ベクトル量子化

本節ではベクトル量子化の条件である次元数、出力ベクトルの数、歪み測度の定義について以下の条件で画像信号の平均値分離正規化ベクトル量子化を行う。

- (i) 次元数 4×4 の隣接画素をブロック化して得る 16 次元とする
- (ii) 出力ベクトル数 $N = 128, 256, 512$, 後述する 2 進木探索ベクトル量子化には $N = 128 \sim 2048$ の 2 のべき乗とする

(iii) 歪み測度 演算の高速化を図り Chebyshev 距離による絶対値歪み測度を用いる

$$d(\mathbf{x}, \mathbf{y}_i) = \sum_{j=1}^K |x_j - y_{ij}|$$

(iv) 出力ベクトル 平均値分離正規化したトレーニング系列に対して一般化 Lloyd アルゴリズム [2-9] を用いて生成する (詳細は 2.4.1 節で述べる)

平均値分離正規化の処理は次の式によって行う。信号源から入力される原入力ベクトルを $\mathbf{S}_l = [S_1, S_2, \dots, S_{16}]$ 、平均値分離正規化された入力ベクトルを $\mathbf{x} = (x_1, x_2, \dots, x_{16})$ とすると、

$$\mu_l = [\sum S_j] / 16$$

$$\sigma_l = [\sum |S_j - \mu_l|] / 16$$

$$x_j = (S_j - \mu_l) / \sigma_l$$

このとき、 μ_l は画像の局所的な輝度レベル、 σ_l は局所的なコントラスト、平均値分離正規化されたベクトルは正規化された波形パターンに対応する。言うまでもなく、入力ベクトル及び出力ベクトルには以下の拘束条件がある。

$$\sum x_j = 0$$

$$\sum |x_j| / 16 = 1$$

入力ベクトル \mathbf{x} に対して最小歪み $\min d(\mathbf{x}, \mathbf{y}_i)$ を与える出力ベクトル \mathbf{y}_i を探索し、インデックス i をベクトルの波形パターンを表す符号化情報とする。復号の過程では以下の式によって画像信号を復元する。

$$S'_j = \sigma'_i \cdot y_{ij} + \mu'_i$$

$$\mathbf{S}'_i = [S'_1, S'_2, \dots, S'_{16}]$$

図 2-7 は上記の処理を概念図で示したものである。

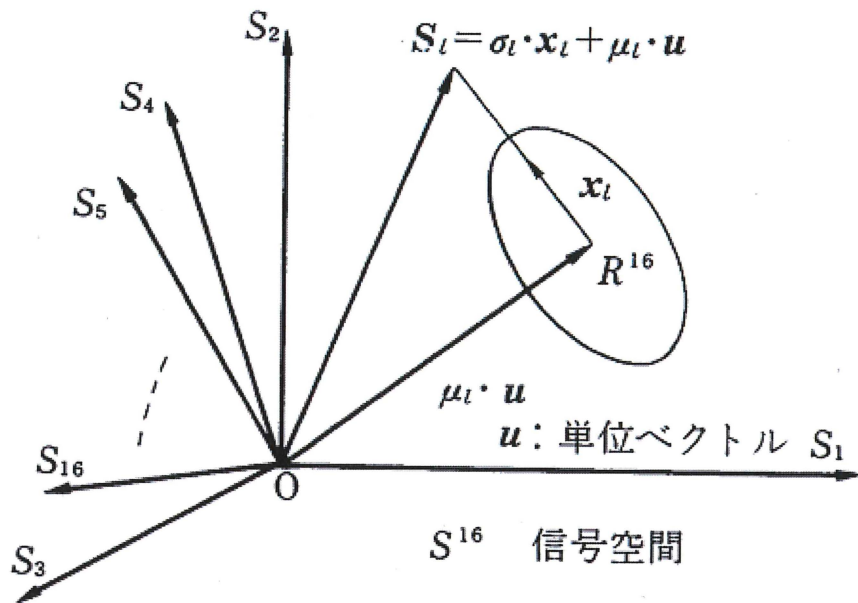


図 2-7 平均値分離正規化ベクトル量子化の概念図

このとき μ_l と σ_l に割り当てるビット数をそれぞれ t_μ, t_σ とすると、伝送すべき符号化レートは以下のようなになる。

$$T = [\log_2 N + t_\mu + t_\sigma] / 16 \quad (\text{bit/pel})$$

以上の条件で、GIRL および AERIAL から出力ベクトルのセット (コードブック) を生成し、両画像を量子化した結果を表 2-1 に示す。

表 2-1 異なるトレーニング系列を用いて生成されたコードブックによる符号化特性 ($N = 256$ の場合) [2-35]

出力ベクトルの セットの生成	GIRL		AERIAL	
	$S_{\rho-\mu}/N_{rms}$ (dB)	ENTROPY (bit/pel)	$S_{\rho-\mu}/N_{rms}$ (dB)	ENTROPY (bit/pel)
GIRL トレーニング	35.4	1.21	24.8	1.29
AERIAL トレーニング	35.2	1.21	26.0	1.31

GIRL に対する符号化結果はトレーニング系列が GIRL であっても AERIAL であってもほとんど同等であり、AERIAL に対する符号化結果はトレーニング系列が AERIAL の場合が GIRL の場合より 1.2dB 高い。これより以下のことがわかる。

- ・ 平均値分離正規化によって汎用性のあるコードブックが生成されたため、全く異なる画像をトレーニング系列として交互に使用しても極端なミスマッチは生じない
- ・ 絵柄の複雑な AERIAL の方が比較的単純な GIRL に比べて豊富なパターンを持つトレーニング系列であり、両者の関係は AERIAL が GIRL をほぼ包含していると思われる

比較のため、現画像を 3×2 画素ごとにブロックの平均値で置き換え、ほぼ同等のビットレート (1.3bit/pel) にした画像の PSNR は GIRL で 32.1dB、AERIAL で 22.9dB であった。単純なサブサンプルに比べて 2~3dB 強の利得が得られていることがわかる。

[2-12,2-27,2-35]

2.3 ベクトル量子化器の構成と符号化性能

2.3.1 全探索ベクトル量子化器の構成

入力ベクトル \mathbf{x} に対して最小歪みを与える出力ベクトル \mathbf{y}_i を求めるには、各出力ベクトルによる歪みの比較を行う必要がある。全ての出力ベクトルとの歪みを比較する全探索ベクトル量子化器の構成図を図 2-8 に示す。

符号化部では入力ベクトル \mathbf{x} とコードブックから順次読み出される出力ベクトル \mathbf{y}_i との歪み $d(\mathbf{x}, \mathbf{y}_i)$ が計算され、最小歪みを与える出力ベクトルが収納されたコードブック上のアドレス i が正規化パラメータの μ_l 、 σ_l と共に出力される。

復号部では、コードブックからインデックス i に対応する出力ベクトル \mathbf{y}_i が読み出され、 μ_l 、 σ_l との演算を経て画像信号が再生される。

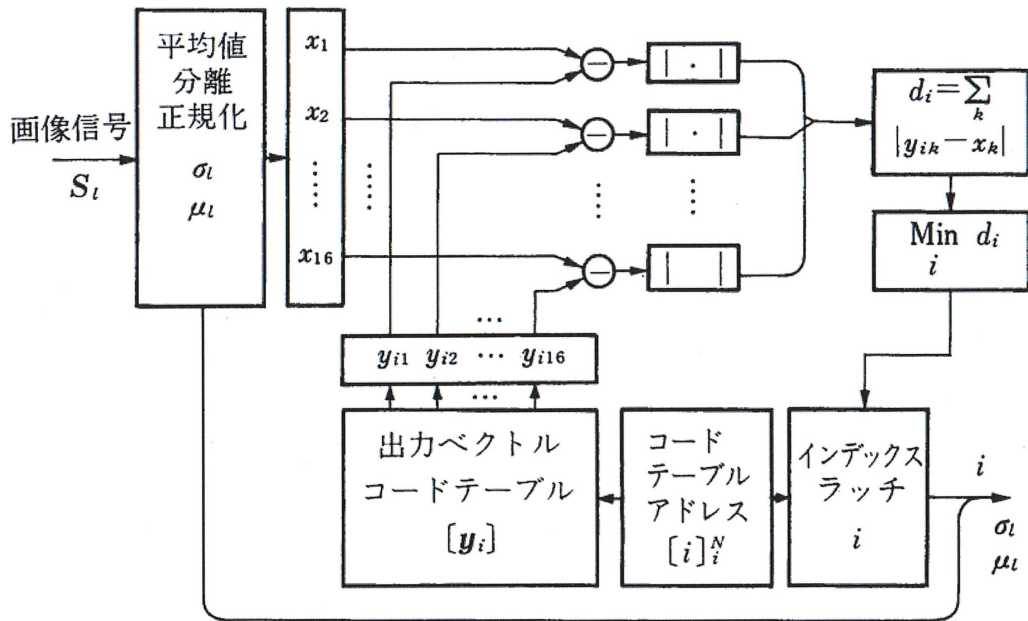


図 2-8 全探索ベクトル量子化符号化器の構成

2.3.2 木探索ベクトル量子化器

木探索による高速ベクトル量子化は画像信号より前に音声信号に対して行われている [2-8]。ここでは画像信号を対象として平均値分離正規化手法を用いた木探索型ベクトル量子化器について検討する。具体的には平均値分離正規化処理を施したトレーニング系列を階層的に分割して木構造を持つ出力ベクトルのセットを生成する [2-27]。\$M\$ 進木を用いて \$N\$ 個の出力ベクトルに対する探索を行う場合、全探索型ベクトル量子化器に比べて歪み計算回数は \$N\$ 回から \$M \log_M N\$ 回に減る。

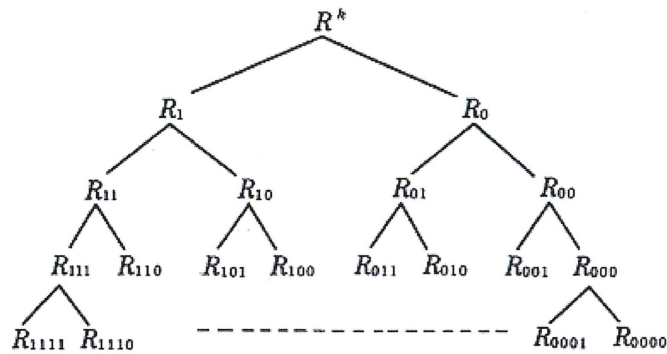


図 2-9 2進木に対応する信号空間の分割

ここでは図 2-9 のような 2 進木を考える。木の根は K 次元信号空間 R^K 、各節点は R^K を段階的に分割した空間に対応する。終端でない各節点には対応する階層的分割を決定する出力ベクトルが設定され、これらとの歪み計算により終端接点の出力ベクトルへの経路が選択される。各接点で 2 つのベクトルからの距離による分割を行うことは多次元信号空間を超平面で分割することに相当し、超平面上ないし超平面近傍に分布するベクトルの分割に対する最適性は保証し得ないので木の終端接点に対応する信号空間の重複は避けられない。損失となる分は終端接点に対応する出力ベクトルの数を増やすことで補う必要がある。

図 2-10 の構成図は各節点の処理をパイプライン化してさらに高速化を図ったものを第 $(n+1)$ 段についてのみ示している。図中、 $b(n)$ は第 n 段までの履歴を表す n 桁の 2 進数列である。各段のコードブックには木構造を持つ出力ベクトルのセットの各段に対応するベクトルが記憶されており、 $b(n)$ をアドレスとして 2 つのベクトルを出力する。

トレーニング系列の分割は 2 つの出力ベクトルに対する一般化 Lloyd アルゴリズムによって行う。各分割が 2 分割であるため初期設定が分割結果に大きく影響することが予想される。ここでは実験的に決定した以下の手順を用いる。

- 分割すべき部分空間に属するトレーニング系列の重心を求め、重心から最も離れた空間内のトレーニング系列のベクトルを第 1 のベクトルとする
- 第 1 の点から最も離れたトレーニング系列のベクトルを第 2 のベクトルとする
- 第 1 のベクトル及び第 2 のベクトルと重心の間にそれぞれ $\sqrt{2}:1$ の重みをかけて 2 つの初期ベクトルとする

このようにして、初期設定の 2 つのベクトルが近距離になることを避け、トレーニング系列に混入した特異なベクトルに影響を受けることを想定した。これらのベクトルを用いて次の分割がトレーニング系列に対して最小歪みとなるように繰り返して収束させる。

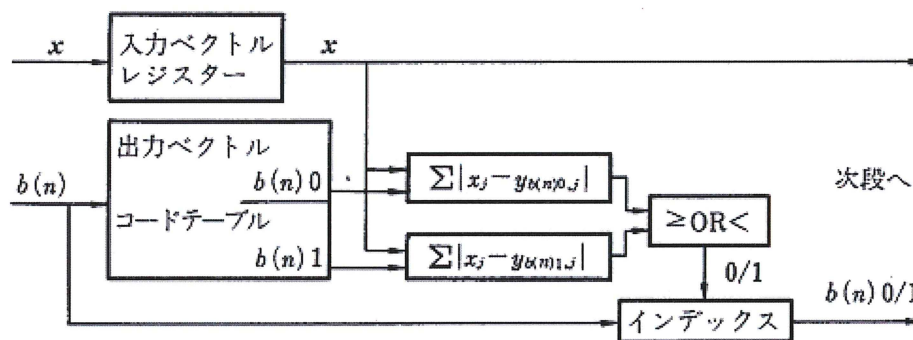


図 2-10 2進木探索ベクトル量子化器の第 n+1 段符号化部

2.3.3 木探索ベクトル量子化器の符号化性能

全探索型、木探索型ともに AERIAL をトレーニング系列として 16 次元平均値分離正規化ベクトル量子化のための出力ベクトルを生成させる。図 2-11 に AERIAL を符号化対象としたベクトル量子化の符号化特性を示す。この結果より以下のことが読み取れる。

- ・ 2進木探索の場合、全探索に比べて 0.1bit/pel 程度のロスがある
- ・ 2進木探索で全探索と同程度の SNR を得るにはおよそ 3 倍程度の出力ベクトルが必要である

また、グラフの範囲では出力ベクトル数を増やしていくにつれ、全探索の特性に近づく傾向が見える。ただしこの場合はトレーニング系列が符号化対象画像であることに注意が必要である。AERIALをトレーニング系列として生成した出力ベクトルを用いてGIRLの2進木探索ベクトル量子化を行った場合は、全探索型の $N=256$ でSNR 34.6dB、木探索型の $N=2048$ でSNR 34.0dBであった。すなわち、2進木探索ベクトル量子化器を2.3.2で述べたように設計した場合、全探索型に比べてトレーニング系列への依存度が高まっていると考えられる。

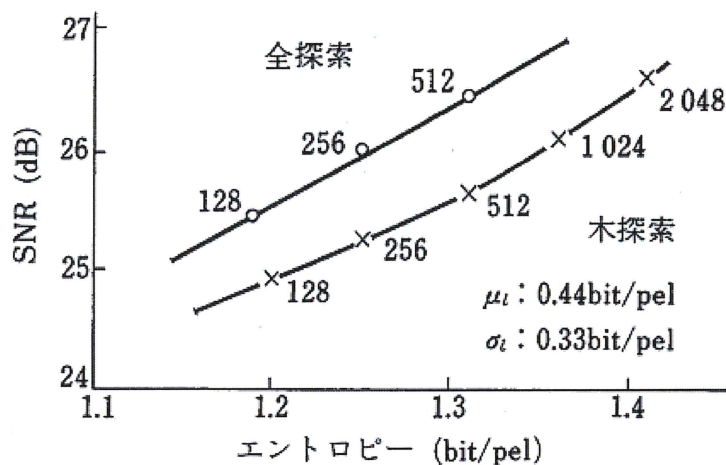


図 2-11 AERIAL のベクトル量子化符号化特性 [2-35]

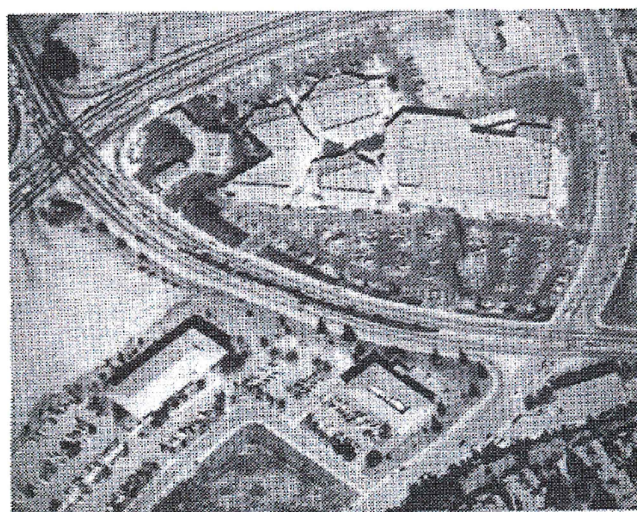


図 2-12 AERIAL のベクトル量子化符号化結果 ($N=2048$)

2.3.4 振幅補正と Gain/Shape 型ベクトル量子化

平均値分離正規化ベクトル量子化では信号再生のために出力ベクトルに振幅を乗じる際、量子化誤差を共に増幅している。これを緩和するため、量子化誤差を用いて振幅値を調整することを考える。

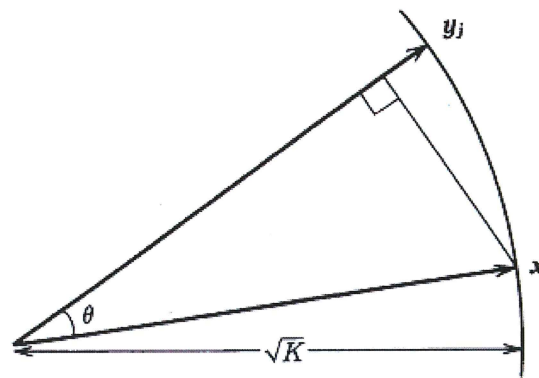


図 2-13 振幅補正の説明図 [2-39]

正規化に用いる σ をベクトルの標準偏差とした場合、入出力ベクトルは多次元信号空間の超球面上にある。図 2-13 に示すように入力ベクトル $\mathbf{x} = (x_1, x_2, \dots, x_K)$ が $\mathbf{y}_i = [y_{i1}, y_{i2}, \dots, y_{iK}]$ に量子化されたとする。ユークリッド距離で評価するなら、 \mathbf{y}_i をスカラ倍して \mathbf{x} の最良近似を得られるのは振幅が $\cos \theta$ の時である。 \mathbf{x} と \mathbf{y}_i の歪みを

$$d(\mathbf{x}, \mathbf{y}_i) = \sum_{j=1}^K (x_j - y_{ij})^2$$

と定義すると、 $\cos \theta$ を

$$\cos \theta = 1 - \frac{d(\mathbf{x}, \mathbf{y}_i)}{2K}$$

と書くことができる。符号化側では量子化の過程で算出された歪みから振幅補正係数 $\cos \theta$ を求め、補正済みの振幅を伝送すれば良い。標準偏差による正規化と二乗歪み測度を用いる場合、入出力ベクトルのノルムは正規化されているので、 $\cos \theta$ は両ベクトルの内積を表

す。この場合の振幅補正は入力ベクトルと平均値分離正規化された出力ベクトルとの内積を歪み測度としてベクトル量子化を行い、最大内積を振幅成分としてスカラ量子化する Gain/Shape 型のベクトル量子化 [2-14,2-38] と等価になる。振幅や歪み測度に近似値を用いた場合、振幅補正の係数を解くことはできない。この場合、入力ベクトルの分布を近似したモデル系列を用意し、出力ベクトルとの歪みを観測して出力ベクトルごとの補正係数を求めておくことが考えられる。実装は平均歪みが大きい出力ベクトルは信号空間の超球面より補正係数に比例して原点寄りに減衰させることに相当する。

2.4 コードブックの生成

2.4.1 一般化 Lloyd アルゴリズム [2-9]

2.2 節で用いたコードブックは GIRL または AERIAL から 4×4 画素ブロックの 16 次元ベクトルを抽出したトレーニング系列に対して一般化 Lloyd アルゴリズムによるクラスタリングを行ったものである。2.1 で述べたように、同アルゴリズムは Lloyd と Max による最適量子化器の設計を多次元に拡張したものである。クラスタリングによる分類手法には先行例があり、同アルゴリズムは k-means 法 [2-3] と本質的に同等である。文献 [2-9] の意義は Lloyd による最近傍条件と重心条件という最適条件を初めて多次元で定式化し、信号の圧縮という応用を含めて議論した点にある。同アルゴリズムは著者名の頭文字をとって LBG アルゴリズムとも呼ばれるが、著者の Gray は一般化 Lloyd アルゴリズムの呼称を提案している [2-24]。

一般化 Lloyd アルゴリズムのフローを図 2-14 に示す。図 2-14 における記号の定義は以下の通りである。

$\|R_i\|$ 部分空間 R_i に含まれるトレーニング系列のベクトルの数

ϵ 歪み収束判定のしきい値

w ワークエリアとして用いるコードブックの添字

c クラスタリングのループカウンタ

N 出力ベクトルの数

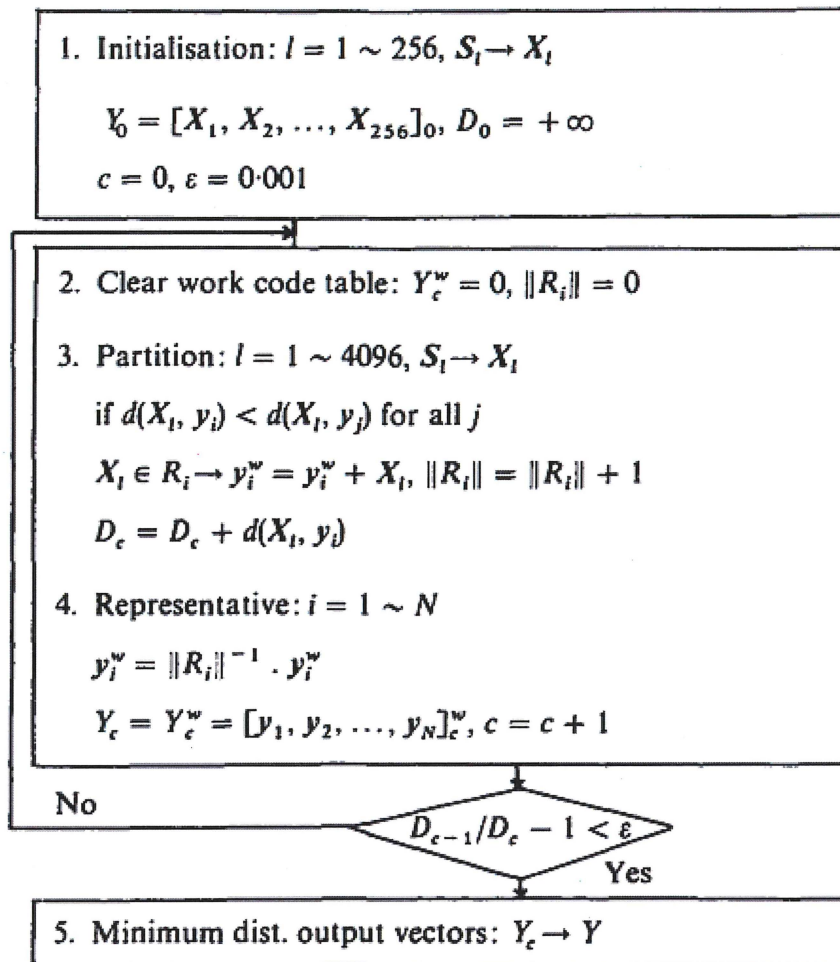


図 2-14 一般化 Lloyd アルゴリズムのフローチャート [2-12]

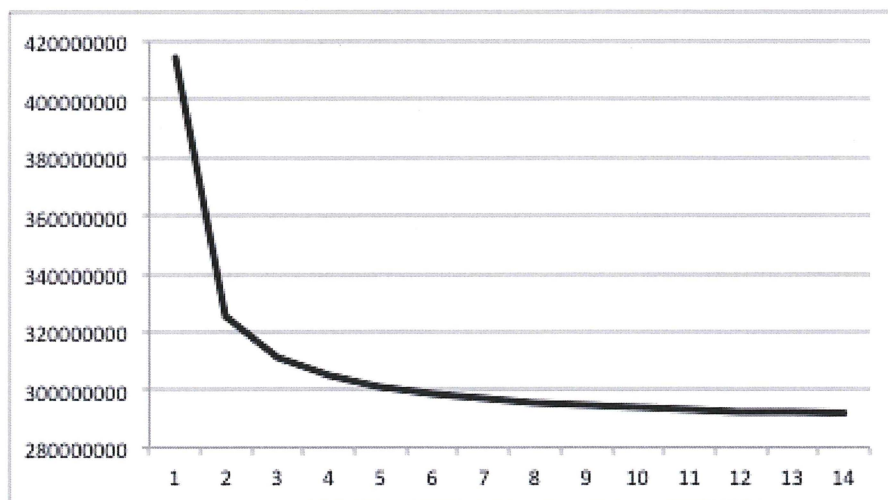


図 2-15 一般化 Lloyd アルゴリズムによる平均値分離正規化ベクトル量子化器コードブック生成のクラスタリングにおける収束曲線

また図 2-15 は一般化 Lloyd アルゴリズムによって平均値分離正規化ベクトル量子化器のコードブック生成のためのクラスタリングを行った際の収束の様子を示している。

以下の図 2-16 は生成された出力ベクトル 256 個を持つコードブックを可視化した例である。各方向の傾斜ないしエッジに対応するパターンが観察できる。

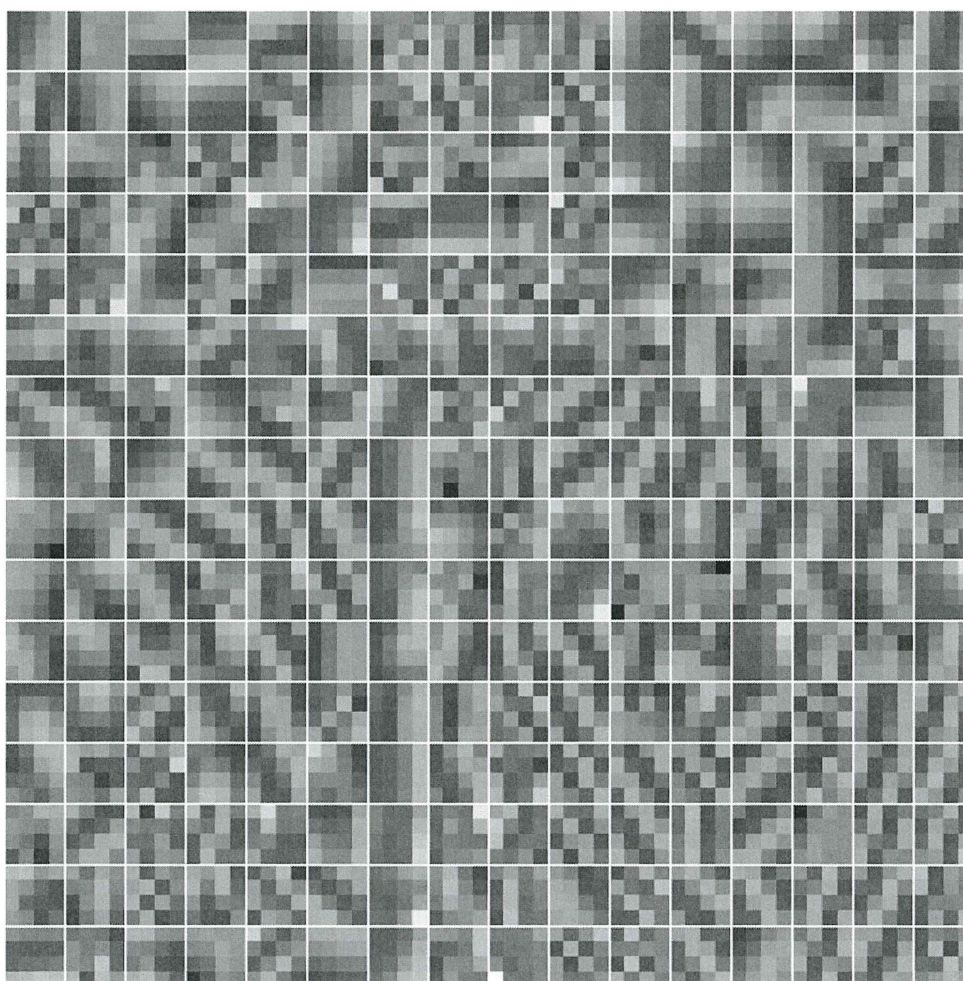


図 2-16 平均値分離正規化されたコードブックの例
(出力ベクトルの数は 256、各ベクトルに対して一律に平均値 128、振幅 64 を与える条件で 8 ビット画像に可視化してある)

2.4.2 クラスタリングにおける初期設定の問題

一般化 Lloyd アルゴリズムにおける出力ベクトルの生成は初期設定に依存する。実際、同アルゴリズムが導く収束解は初期設定に近い局所最適解にすぎない。極端な場合、初期設定の出力ベクトル \mathbf{y}_i^0 に対して写像されるトレーニング系列の入力ベクトル数が 0 であれば、(クラスタリングによって他の出力ベクトルが移動するために最近傍ベクトルが同ベクトルから \mathbf{y}_i^0 に移動する入力ベクトルが現れる偶然を除き) \mathbf{y}_i^0 は更新されることがなく、歪み量の低減にも貢献しないので、別の出力ベクトルに置き換えられなければならない。トレーニング系列の一部を用いて初期のコードブックを設定する場合、上記のようなことは発生しないが、ある出力ベクトル \mathbf{y}_i^0 に対して写像されるトレーニング系列の入力ベクトル数が 1 ないしは他の \mathbf{y}_j^0 ($i \neq j$) に比べて極めて小さいという事象は可能性がある。そこで初期設定を評価するために、貢献度の低い出力ベクトルと貢献を要する出力ベクトルを次のように考える。

① 貢献度の低い出力ベクトル

- ・ 当該ベクトルに量子化される入力ベクトル数が少ない
- ・ 当該ベクトルをコードブックから取り除いた場合にトレーニング系列に対する量子化歪みの総和の変化が小さい

② 貢献を要する出力ベクトル

- ・ 当該ベクトルに量子化される入力ベクトル数が多い
- ・ 当該ベクトルに量子化された入力ベクトルとの平均歪みが大きい

クラスタリングの計算を 1 度実行し、量子化結果を参照して①②にしきい値を導入して極端な例が観察された場合、①の出力ベクトルを棄却し、同ベクトルのインデックスには②の出力ベクトルに対して微小な変動を与えた出力ベクトルを割り当てることで出力ベクトル設定の適正化が行えると期待できる。

筆者が実験を行なった範囲では、平均値分離正規化を行わない直接ベクトル量子化のコードブック生成では①に相当する場合が発生

して設定の適正化を行なったが、平均値分離正規化ベクトルのクラスタリングを行なった場合は量子化結果が分散し、極端な①や②が発生することはなくなった。つまり平均値分離正規化は全てのベクトルを超球面上に配置することによって初期設定への強い依存を緩和する効果を持つと考えられる。ただし、このような調整は初期設定への依存を完全に解決するものではなく、ベクトルの次元数を拡大した場合や次元数に対する出力ベクトルの数が小さい場合などに、上記のような初期設定の調整が有効となる可能性がある。

2.4.3 トレーニング系列の選別と拡大による汎用化

平均値分離正規化ベクトル量子化では正規化により、振幅 σ とベクトルのパターンを分離しているが、両者には下記のように相関関係があると考えられる。

- ・ 画像のエッジ領域では σ の値が大きく、ベクトルのパターンはエッジの形状となる
- ・ 画像の平坦な領域では σ の値が小さく、ベクトルのパターンは雑音の影響を受けてランダムに近づく

復号の過程で出力ベクトルに振幅を乗算する時にはベクトル量子化雑音も増幅される。従って振幅の高いベクトルに対する量子化精度を高めることによって最終的な符号化特性の改善に対して有効と考えられる。2.4.1 に示したコードブックの出力ベクトルには傾斜や境界など画像の基本的な構造に対応するパターンが観察されることもこの予想を裏付ける。一方、符号化雑音がブロック内でランダム化されるために主観的に劣化が目立ち難いこともベクトル量子化の利点の一つなので、ランダムなパターンも出力ベクトルとして必要である。これより以下のようにトレーニング系列を抽出することを考える。

M 枚の画像群 $I = [I_1, I_2, \dots, I_M]$ からトレーニング系列を抽出するものとして、正規化前のベクトルの振幅 σ によってトレーニング系列を例えば以下のようにクラス分けする。

トレーニング系列 T_0 : $\sigma < Th_0$ のベクトル群

トレーニング系列 T_1 : $Th_0 \leq \sigma < Th_1$ のベクトル群

トレーニング系列 T_2 : $Th_1 \leq \sigma$ のベクトル群

トレーニング系列 T_i ($i=0,1,2$) に含まれるベクトルの数を $\|T_i\|$ とし、それぞれの関係が以下のようになるように抽出を行う。

$$\|T_0\| \ll \|T_1\| < \|T_2\|$$

全てのトレーニング系列を統合して、コードブック生成用のトレーニング系列 T とする。

$$T = T_0 \cup T_1 \cup T_2$$

さらにトレーニング系列におけるパターンの偏りを回避するため、正規化後のベクトルに以下の変形を適宜加え、また画像からブロックを切り出す位置を変えてトレーニング系列を拡大した。

$$x_{i,j}^{ne} = -x_{i,j} \quad (\text{反転})$$

$$x_{i,j}^{lr} = x_{i,N-1-j} \quad (\text{上下反転})$$

$$x_{i,j}^{ud} = x_{N-1-i,j} \quad (\text{左右反転})$$

これらは少ないトレーニング系列から汎用性のあるコードブックを生成するための工夫である。

2.4.4 基本パターンによるモデル表現と認識適用への可能性

2.4.1 で生成した平均値分離正規化出力ベクトルを可視化すると、傾斜やステップ状など画像の基本構造に対応するパターンが観察される[2-30]。こうした基本構造型に着目し、トレーニング系列に用いる画像への依存性を排して基本パターンを設計する試みとして、三角関数を用いて人工的に出力ベクトルを設計した[2-29]。ここではパターンを①傾斜、②エッジ、③2次曲線とクラス分けして 4×4 の仮想的画素ブロックを生成した。画素配列を $P(x,y)$ ただし $(x,y = 0,1,2,3)$ とする。

①の傾斜と②のエッジには共通の式(*)を用い、①の波形をピークホールドして②とした。また③の2次曲線には式(**)を用いた。

$$p(x, y) = \sin\left(\frac{2\pi}{T_x}x + \frac{2\pi}{T_y}y + \varphi\right) \quad (*)$$

$$p(x, y) = \cos\left[\frac{\pi}{T}\{(x-a)^2 + (y-b)^2\}\right] \quad (**)$$

各パラメータは以下の通りである。

$$T_x, T_y = 0, 4, 6, 8, 10, 12, 16$$

$$\varphi = \pi \cdot (m/T_x + n/T_y) \quad m, n \in Z$$

$$T = 4, 6, 9, 12, 18$$

$$a, b = 0, 1, 2, 3$$

これらのパターンから重複を除き、256個の平均値分離正規化出力ベクトルのコードブックをクラスタリングにより求めた。同コードブックによって画像 AERIAL と GIRL を符号化すると、AERIAL について S_{p-p}/N_{rms} は 25.3dB、GIRL については同 35.0dB であった。

2.2.3 の符号化結果を参照すると人工的に生成したパターンによるコードブックを用いた符号化結果は AERIAL をトレーニング系列として生成したコードブックの結果にわずかに劣る程度である。このことは傾斜やエッジを主な構造パターンとして人工的に生成したコードブックが有効である可能性を示している。

一方、AERIAL の復号画像を観察すると、地面や植物の部分のように、振幅は高いが傾斜でもエッジでもないテクスチャ部分に明瞭な傾斜パターンが観察される。上記のように生成したパターンにはランダムノイズを含む細かなテクスチャが無いいため、この結果は自明である。完備性を有するパターンの集合を持つには、傾斜やエッジのパターン以外にテクスチャのパターンを持つ必要がある。パターンを出力ベクトルに内包するには画素ブロックを拡大すること、もしくは隣接・近傍の画素ブロックとの間の相関や連続性を何らかの形で表現することが課題となる。

図 2-17 は平均値ゼロで単位振幅を持つ基本パターンをコードブックとして持ち、パターンを拡大・縮小など変形した後に適当な振幅と平均値を与える形式で画像の表現を行う画像信号源のモデルの

概念図である。ここで拡大・縮小・変形としている処理は言わば「幾何的な正規化」の逆関数を意図している。

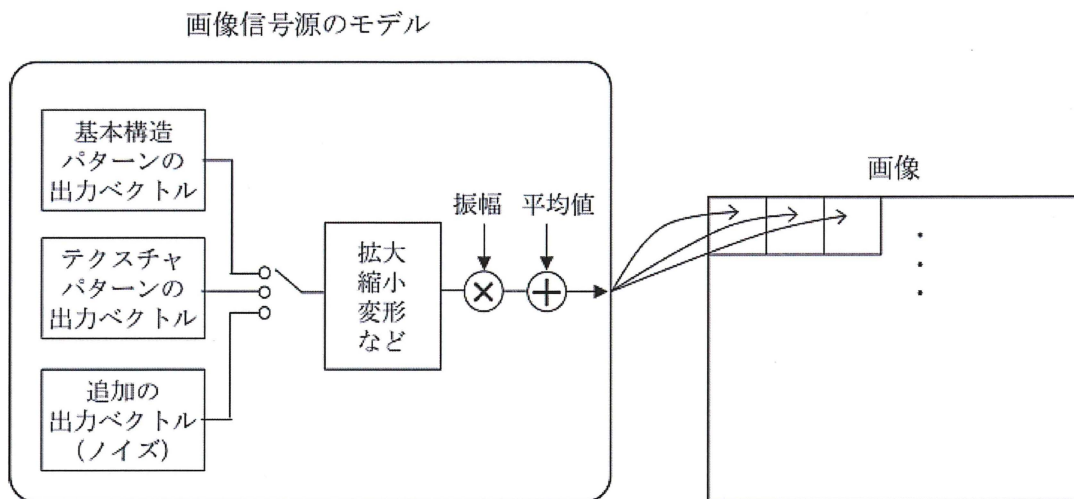


図 2-17 基本パターンのセットによる画像信号源のモデル

図 2-17 のモデルに基づいて符号化を行うには、平均値分離正規化のみならず、幾何的変形による正規化を含めたパターンマッチングの評価が必要になる。このアプローチは画像の特徴解析に基づく符号化とも言える。表 1-1 においてベクトル量子化と特徴抽出符号化を近いクラスに置いた動機はこの視点を考慮したことである。また表 1-1 の特徴抽出符号化に分類したフラクタル符号化は、技法においてベクトル量子化と多くの共通点を持つ。1990 年に濃淡画像への適用が提案されたフラクタル符号化[1-22]は画像の自己相似性に注目し、予め生成したコードブックでなく、画像のある領域から各種の変形を含む縮小写像によって画素ブロックを表現するものである。

コードブックに含まれる人工的なパターンあるいは実信号から抽出されたパターンが、ある用途にとって有意である場合、ベクトル量子化による符号化とパターン認識を同時に実現することができる。先行例として、1970 年代のクラスタ符号化は衛星画像に対して分類と圧縮を同時に行う PR/DC (Pattern Recognition and Data

Compression) 符号化の側面を持っていた [2-5]。筆者らは上記のように人工的に生成した基本パターンと実画像をトレーニング系列とする出力ベクトルの両方を含むコードブックを使用し、PR/DC 符号化の基本的検討を行った [2-34]。PR/DC 符号化では用途に応じた特徴ベクトルの設計が重要となる。

2.4.5 対称性を導入した木探索用コードブックの設計

本節の最後に、トレーニング系列への依存度が大きい木探索ベクトル量子化のコードブックを汎用的に生成するために、ベクトルの対称性を導入する手法について考える [2-15]。

木探索ベクトル量子化のコードブックを設計することはトレーニング系列のベクトルが分布する k 次元信号空間を部分空間へと再帰的に分割していくことに相当する。2 進木探索の場合、親に相当する部分空間（親空間）を子である 2 つの部分空間（子空間）に分割する。ここで考える対称性とは、2 つの子空間の代表ベクトルが親空間の重心（当該空間の代表ベクトル）について対称となる関係である。すなわち、図 2-18 の 2 進木において根から実線で結合されている節点の出力ベクトルを対称性に基づいて展開し、破線で結合されている出力ベクトルを得るコンセプトである。

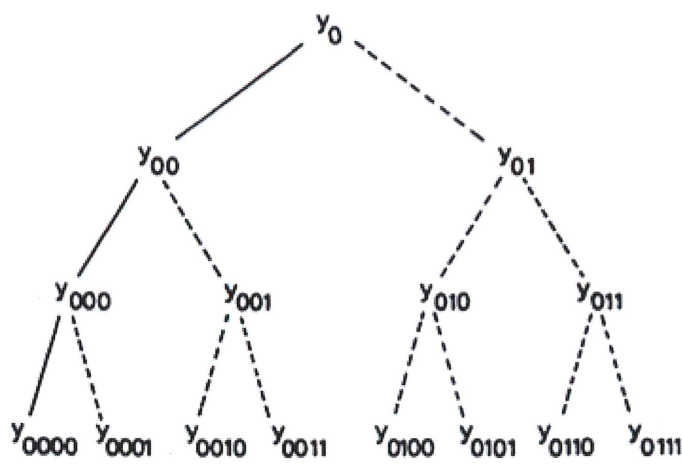


図 2-18 対称性を持つ 2 進木に対応する出力ベクトル群

上記の関係は親空間におけるベクトルの分布が重心に対して対称性を持つ場合に合理的である。2進木の根の段階ではベクトルが信号空間で球対称な分布を持つので上記仮定が成り立つ。この時、2進分割は2.3.2に述べた手法と同じ分割となる。ここでは信号空間を2分割するごとにトレーニング系列に属するベクトルの対称折り返しを行い、出力ベクトル決定段階での対称性を確保する。

図2-18におけるベクトル $y_{0\dots 0}$ を生成する手順を図2-19に示す。ただし、 l はトレーニング系列のシーケンス番号であり、 c はクラスタリングのループ回数、 ϵ は収束判定のしきい値、 n は段数を表す番号であり、最大段数は r である。また出力ベクトルの添字について n 個のゼロ連を $n(0)$ 、ベクトル x_l を $y_{0(n)}$ について対称に折り返したベクトルを $\text{sym}(x_l, y_{0(n)})$ 、トレーニング系列の部分集合 $[x_l]$ の代表ベクトルを $\text{rep}(x_l)$ と表す。

図2-19における $\text{sym}(x_l, y_{0(n)})$ の折り返し操作は図2-20と次の式で表現される。 $\langle x_l, y_{0(n)} \rangle$ はベクトル x_l と $y_{0(n)}$ の内積である。

$$x_l^* = 2\langle x_l, y_{0(n)} \rangle \cdot y_{0(n)} - x_l$$

上記の操作で得た出力ベクトル $y_{0\dots 0}$ を2進木に展開するために以下の手順を用いる。

$$y_{0(p) \cdot 1 \cdot b(q)} = \text{sym}(y_{0(p+1) \cdot b(q)}, y_{0(p)})$$

ただし第 n 段の出力ベクトルの添字を p と q 、長さ p のゼロ $0(p)$ 、長さ q の2進数列 $b(q)$ で表し、 $p+q+1=n$ 、 $q=0$ の時に $b(q)=\phi$ とする。図2-21は求めたコードブックを5段目まで表示した。

2.3.2で触れたように、段数ごとの分岐数 M が小さい M 進木探索ベクトル量子化コードブックでは段数が深くなるに従って信号空間の重複が大きくなり、これが量子化損失となる。この損失に加え、対称折り返しでは、本来はトレーニング系列によって決定される信号空間においてベクトルの分布密度が低い領域にもベクトルが折り返される可能性がある。このことはコードブックの汎用性につながるが、原信号のベクトルが有するベクトル内部の相関による利得が

小さくなる可能性を含んでいる。

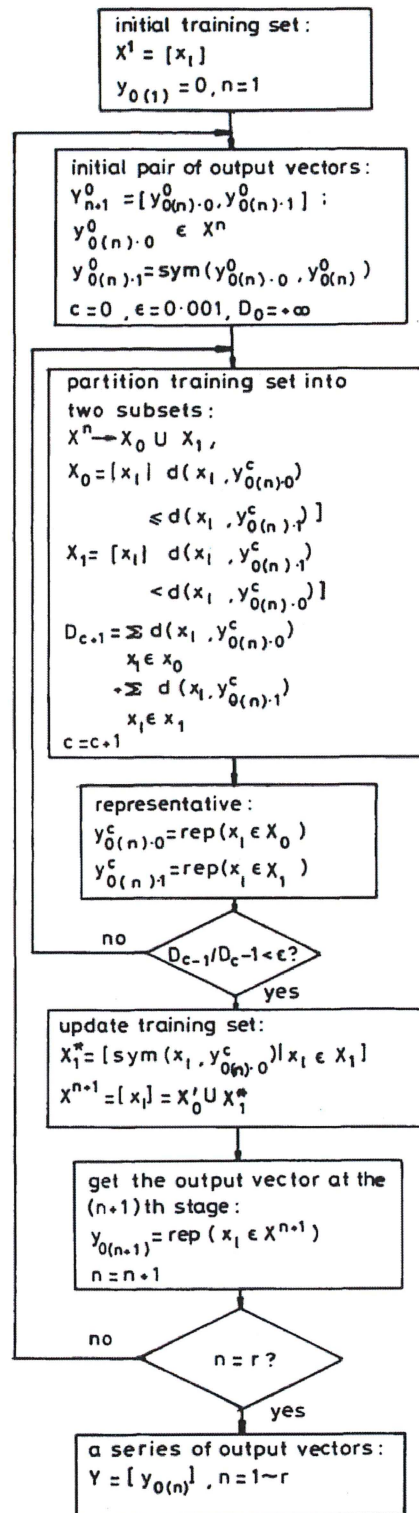


図 2-19 対称折り返しを行う 2 進木出力ベクトルの生成 [2-15]

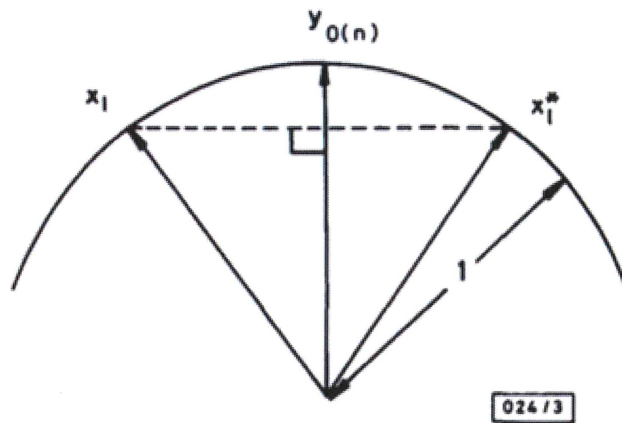


図 2-20 正規化ベクトルの対称折り返し [2-15]

本手法を用いることにより、サイズが大きくユニバーサル性を有する木探索ベクトル量子化のコードブック生成が可能になる。さらに対称性の確保によりトレーニング系列に依存する効果が緩和される。一方、段数を深めていった時に、部分信号空間における分布の対称性が低下することも想定され、折り返されたベクトル間の距離、ベクトルが分布する信号空間の重複性について定量的な評価が必要である。

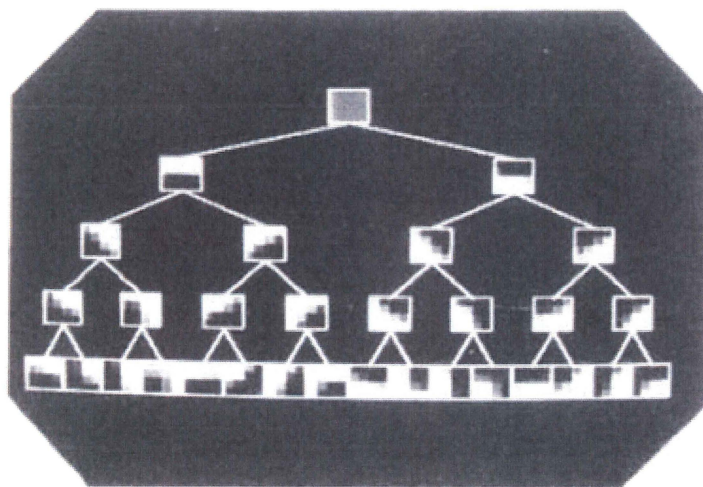


図 2-21 対称性を用いて生成した 2 進木探索用出力ベクトル [2-15]

2.5 低ビットレート動画像符号化への展開

2.5.1 カラー画像のベクトル量子化

カラー画像は RGB またはこれに準ずる 3 成分あるいは輝度と 2 つの色差成分で表される。カラー画像のベクトル量子化では、各成分を独立にブロック化する手法と成分を結合してブロック化する手法が考えられる。筆者らは動画像には各成分を独立にブロック化する手法、高解像度静止画には成分を結合した手法が適していると考えて検討を行った [2-39]。

対象画像を限定した時の符号化性能は輝度・色成分間の相関をコードブック設計に取り込むことのできる成分結合型のベクトル量子化が優れている。一方、輝度・色成分間の相関は両者が同時に変化する変化点でこそ顕著であり、汎用性には課題があり、量子化誤差による色ズレは主観的に顕著に観察される。静止画では対象画像に個別の付加情報を与える可能性があるが、動画像では汎用性が必須になるため、成分独立型のベクトル量子化が合理的である。このため本節で扱う動画像では 3 成分を個別にベクトル量子化するアプローチを用いる。

2.5.2 動画像のダイナミックベクトル量子化 [2-36]

動画像に対する符号化性能を高めて低ビットレート動画像符号化を実現するため、出力ベクトルを動的に生成するダイナミックベクトル量子化 (Dynamic Vector Quantization: DVQ) を前段に置き、これに対する誤差を後段で平均値分離正規化ベクトル量子化する多段構成の符号化を考える。ここでは後段の平均値分離正規化ベクトル量子化を DVQ に対して AVQ (Adaptive VQ) と呼ぶ。

図 2-22 において前段 DVQ の入力ベクトルを \mathbf{s}_r 、これに対する出力ベクトルを $\widehat{\mathbf{s}}_r$ とする。 r は \mathbf{s}_r の画面上の位置を表すものとする。両ベクトルは $M \times N (= K)$ 個の画素ブロックとする。後段 AVQ の入力を ϵ_r 、これに対する局部復号出力を $\hat{\epsilon}_r$ とする。両ベクトルは $m \times n (= k)$ 個の

画素ブロックとする。必ずしも $M = m, N = n$ ではなく、 $M \times N$ の画素ブロックは画面上で整数個の $m \times n$ 画素ブロックを包含するものとする。

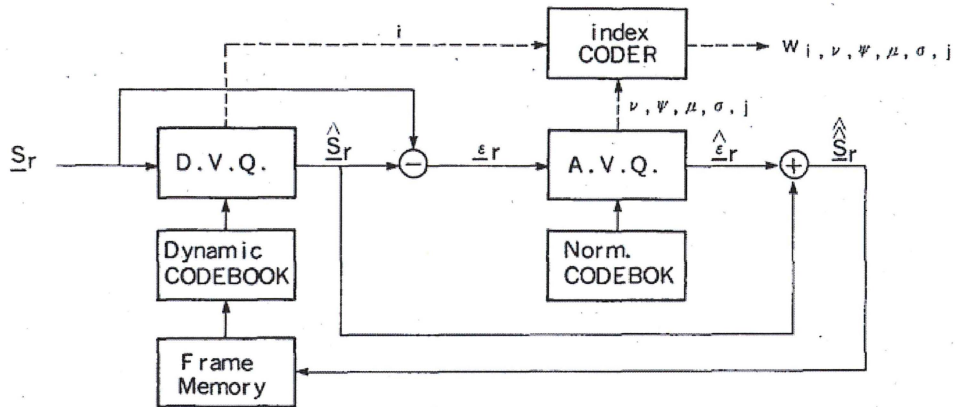


図 2-22 ダイナミック多段ベクトル量子化の構成 [2-36]

前段 DVQ の出力ベクトルのセットは 2 つのグループに分かれる。第 1 のグループは符号化すべき画素ブロックと空間的な同一位置にある符号化済みの過去のフレーム上の画素ブロック \widehat{S}_r^p に位置変位 ν_i ($i=1, 2, \dots, V$) を与えた画素ブロック $\widehat{S}_{r+\nu_i}^p$ によって形成される V 個のベクトル群である。第 2 のグループは平均値のみを成分とする固定的な平均値成分のみからなる L 個のベクトル群である。第 1 のグループだけで位置変位を先験的に与えた場合、DVQ は動き補償予測と等価である。DVQ の設計では位置変位のベクトルのセットを初期設定し、動画シーケンスに対して初段 DVQ が最小の誤差となるようクラスタリングによって最終的な位置変位ベクトルのセットを求めた。第 2 のグループはシーンチェンジなど第 1 のグループの変位で近似しきれないフレーム間の変化があった場合や、フレーム間符号化によるループ内の蓄積があった場合に、これをリセットする効果がある。

後段の AVQ は基本的に前節までの平均値分離正規化ベクトル量子化と同じであるが、符号化対象が実質的にフレーム間の予測誤差信号となるため、トレーニング系列として同等条件のベクトルを用

意すること、またフレーム間予測符号化による積分の収束性を保証するため、原点に対称な量子化器であることが必要となる。このため、フレーム間予測誤差から抽出したトレーニング系列に反転ベクトルを加えて対称なコードブックを生成する。

2.5.3 ダイナミックベクトル量子化による符号化性能

ダイナミックベクトル量子化による符号化性能を確認するためのシミュレーションを行った。目標設定のビットレートは 64kbit/s である。1980 年代初頭、ISDN の導入に向けてテレビ会議・電話への期待があり、ISDN の基本チャンネルである 64kbit/s で動画像の符号化伝送を可能にすることが想定目標の一つであったからである。

テレビ会議用途を想定し、話者が話す動画像（輝度 352×240 画素、色差 88×120 画素）を用いてシミュレーションを行った結果を図 2-23 に示す[2-16]。

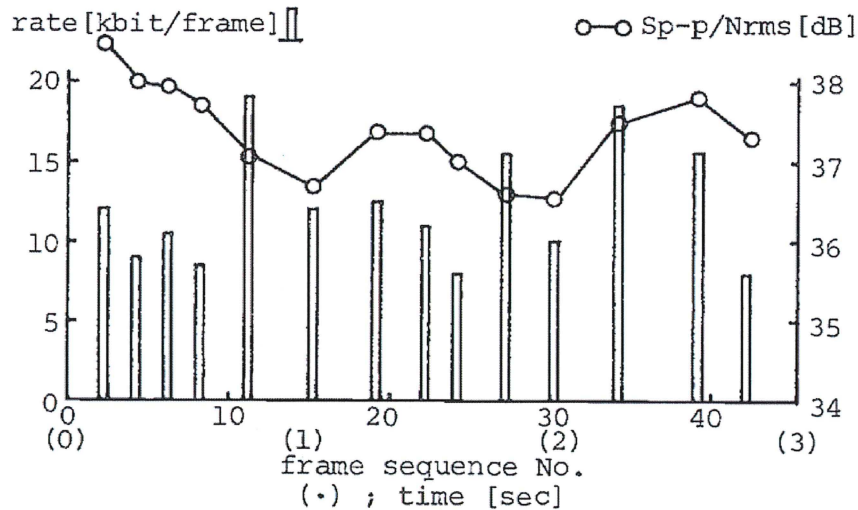


図 2-23 ダイナミック多段ベクトル量子化の符号化性能[2-16]

ここで前段 DVQ において $M=16$, $N=8$ ($K=128$)、 $V=25$ 、 $L=7$ であり、後段 AVQ では $m=n=4$ ($k=16$)、平均値分離正規化出力ベクトル

ルの数は 512 とした。後段 AVQ における平均値、振幅はそれぞれ DPCM 符号化を行った。

画面上において動きのある領域が 10-20% の時、画素あたり 0/1-0.2bit/pel の符号量で平均 S_{p-p}/N_{rms} は 37.4dB であった。これより、伝送速度に 64kbit/s を仮定して秒あたり 4~8 フレームの動画像が伝送可能であることが確認された。

2.6 実時間符号化装置による実証と実用化への展開

筆者は同僚の共同研究者と、まず 1982 年から 1983 年にかけて、NTSC 信号を入力として実時間で動作する動画像の平均値分離正規化ベクトル量子化装置を開発した。同装置ではフレーム間予測による予測誤差をブロック化し、平均値分離正規化ベクトル量子化を実行した。正規化後のベクトルは 8 ビットに整数化して扱った。平均値 μ と振幅 σ をしきい値処理することによって 4×4 ブロックごとに有効/無効ブロックの判定を行い、しきい値以下の誤差がループ内に蓄積するのを避けるため、ループゲインを抑える低域通過型のループフィルタを導入した。図 2-24 に装置写真を示す。

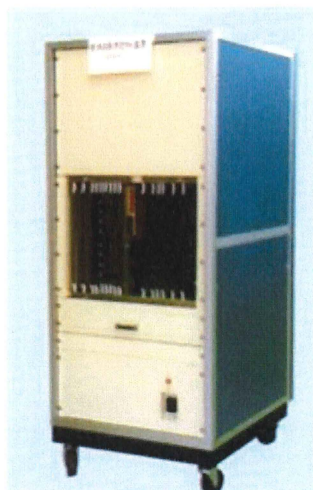


図 2-24 世界初の実時間動画ベクトル量子化符号化装置

装置は基板前縁にロータリースイッチと LED 表示器を備え、ロ

一タリスイッチにより有効/無効ブロック判定のしきい値を設定することが可能であり、LED表示器で有効ブロックの数を読み取ることが可能であった。

2.5.4 記載のように、ISDNの導入に向けて64kbit/sでの符号化伝送が想定目標であった。筆者らは同僚や自らが被写体となり、画面に映るサイズ、動きの程度としきい値設定、有効ブロックの数からビットレートとの関係を実験し、テレビ会議・電話レベルの動きであれば64kbit/s以下で実用的な画質による符号化伝送が可能であろうという感覚を得ることができた。

本装置開発の後、2.5節で述べたダイナミックベクトル量子化および動き適応フィルタの導入、有効/無効ブロックのグループ可変長符号化、情報発生量によるしきい値のフィードバック制御などを加え、テレビ会議・電話システムとして製品化された。[2-42]。

ベクトル量子化による超低レート符号化への期待が高まり、多くのアプローチが生まれ、実用化された。国内だけでなく、米国や英国でも実用化された[2-22]。国内における自社製品以外の実用化事例を2つ挙げる。

一つのアプローチは予測誤差を輝度成分と色差成分の予測誤差を多次元信号空間でベクトル量子化する手法である[2-17,2-40]。本手法は色成分が変化する箇所では輝度成分が同時に変化することが多いというPlateau Coding[2-48]の着目点をベクトル量子化に活かしたものである。同手法は予測誤差を対象とすることで平均値成分をほぼキャンセルし、近隣の符号化済み画素の分散から予測誤差の振幅を調整することで正規化効果を持たせ、コードブック汎用化設計の問題を解決した。同方式はプロトタイプの開発及び発表と実証実験を経て実用化された[2-43]。

もう一つのアプローチは平均値分離正規化された木探索用コードブックに対して木探索の段数を適応化させることによって符号長の制御を可能としたベクトル量子化である。同方式は実機に採用され、ISDNでのテレビ会議システムに活用された[2-45,2-46]。

2.7 この章のまとめ

本章では画像信号のベクトル量子化技術を検討し、同技術が圧縮手法として優れた特性を有することを示し、同時に画像信号をパターンとして表現する可能性を示した。本章で得られた知見は以下の通りである。

- ・ 符号化すべき画像を画素ブロックに分割し、これを平均値分離正規化してベクトル量子化することにより、符号化効率の良いベクトル量子化が実現可能である
- ・ ベクトル量子化のコードブックはトレーニング系列を用いたクラスタリングによって得られるが、平均値分離正規化はトレーニング系列への依存度を軽減し、汎用性を与える効果を持つ
- ・ 平均値分離正規化された出力ベクトルは画像を構成するパターンとしての基本構造型を有する
- ・ コードブックの基本構造型を担保するための手法としてトレーニング系列の拡大や対称性の導入などが有効である
- ・ 動的にコードブックを書き換えるダイナミックベクトル量子化との多段構成により、低ビットレートの動画符号化が実現され、実用に供された

以下に、本研究の時点で追究し尽くしていない課題を挙げる。

- ・ ベクトル量子化の符号化効率を高める次元数およびコードブックの拡大について、演算とメモリの制約から追究不足である
- ・ 高ビットレート化によってベクトル量子化の符号化誤差をゼロに漸近させる高効率またはスマートな手法が追究不足である
- ・ 画像を基本構造パターンでモデル化して表現すること、パターン認識と融合したフレームワークについて追究不足である

最後の課題はベクトル量子化による画像符号化を圧縮の道具以上に活用する視点である。

本章で触れなかった同時代および後年のベクトル量子化に関する研究については第5章(5.1)で述べる。

第3章 動的適応ウェーブレットパケットによる画像符号化

3.1 はじめに

第1章で述べたように、変換符号化はブロック化した画像信号を変換して無相関化し、電力が一部の係数に集中する性質を利用して圧縮を行う。変換のブロックサイズを $N \times N$ とし、入力画素ブロックを \mathbf{X} 、 N 個の変換基底関数を行ベクトルとする変換行列を \mathbf{A} とすると、変換係数の行列 \mathbf{Y} は次の式で求められる。

$$\mathbf{Y} = \mathbf{A} \mathbf{X} \mathbf{A}^T$$

すなわち変換は基底関数への写像である。 \mathbf{X} から \mathbf{Y} への変換によって得られる変換利得 G_{TC} は変換係数の数を $M (=N^2)$ 、各係数の分散を σ_i^2 とすると以下の式で表される [1-4]。

$$G_{TC} = 1/M \left[\sum_{i=0}^{M-1} \sigma_i^2 \right] / \left[\prod_{i=0}^{M-1} \sigma_i^2 \right]^{1/M}$$

ただしこの利得は最適量子化と符号語割当を前提とするものであり、最終的な符号化利得として実現する量子化器や符号語割当を保証するものではない。第1章で触れたように、国際標準方式など画像信号に対して最もよく用いられる変換は DCT [1-15, 1-16] である。DCT による変換係数の符号化に割当られる符号量が十分でない場合、復号画像にブロック歪みやモスキート歪みが現れる。これに対して隣接するブロックをオーバーラップさせることでブロック歪みの発生を抑えるオーバーラップ変換 [1-17] が提案されている。一方、オーバーラップ変換では基底関数が隣接ブロックに重なる分だけ長くなり、劣化の現れる範囲が拡大するため、モスキート歪みの解消にはさほど有効ではない。

サブバンド符号化は、解析フィルタバンクによって画像信号を複数の帯域に分割してサブサンプルした後に、電力が一部の帯域に集中する性質を利用して圧縮を行う。復号側ではアップサンプルした後に折り返し歪みを除去する特性を持つ合成フィルタバンクによって信号を復元する。解析フィルタバンクと合成フィルタバンクによ

って完全再構成が可能なサブバンド符号化は基底関数への写像という点で変換符号化と等価である。

サブバンド符号化は画像信号以前に音声や音響信号の符号化に使われていた [3-1]。完全再構成が可能な各種フィルタバンクについての研究も行われている [3-2,3-7]。画像信号への適用は 1980 年代半ばのことである [3-4,3-5]。また、基底としては冗長であるが、低解像度の画像を階層的に生成して積み上げるピラミッド符号化 [3-3] もサブバンド符号化の流れに連なる。

サブバンド符号化と変換符号化やオーバーラップ変換とは本質的な等価性 [1-18,1-26] を有するが、符号化技法の相違により、性能や画質劣化の現れ方は異なる。サブバンド符号化では基本的にブロック境界が存在せず、ブロック歪みは発生しない。モスキート歪みについては基底関数の長さに依存するため、サブバンド符号化が常に変換符号化より有利とは言えない。ただこの点では基底関数の長さが一定である必要のないサブバンド符号化が有利であり、次の節で述べるウェーブレットは高周波成分で短い基底長を持つため、モスキート雑音に強いという特長を持つ。

上記の範囲ではサブバンド符号化が優れているが、上記以外の相違として、変換符号化がブロック単位の処理であり装置化や並列化に適しているのに比べ、サブバンド符号化は装置化の面で不利になること、動画像符号化でしばしば用いられるブロック単位の動き補償予測との親和性が良くないことが挙げられる。

本章では次節以降、筆者の研究の前提となるウェーブレットとウェーブレットパケットについて説明し、相関の高い信号に対してウェーブレットが電力局在化の意味で効率の良い基底を与えることを示す。その後、与えられた画像に対して符号量-歪み特性の観点で実現可能な最適性を持つウェーブレットパケットを画像の領域ごとに設計して全体最適とする方式と符号化性能について示す。さらにそのようにして得られた周波数と空間の分割が信号源である画像の電力分布の変化を表現していることを指摘する。

3.2 ウェーブレットによるサブバンド符号化

3.2.1 画像のウェーブレット変換

ウェーブレットは”小さな波”であり、周波数的にも時間的にも局在化された関数 $\psi(t)$ を基本として、 $\psi(t)$ を時間方向に伸縮および時間シフトさせることで得られる基底 $\psi_{m,n}(t)$ がウェーブレット基底である。ウェーブレット変換とは信号のウェーブレット基底による分解に他ならない。

$$\psi_{m,n}(t) = 2^{-\frac{m}{2}}\psi(2^{-m}t - n) \quad m, n \in \mathbf{Z}$$

ウェーブレット基底は1980年代初頭にMorletらによって示され、1980年代後半、Mallatによる画像の多重解像度表現への適用[1-19]、Daubechiesによる論文[1-20]などに続き画像符号化への適用に関する研究が盛んに行われた。国際標準JPEG 2000にも採用されている。

信号のウェーブレット基底による分解もしくはウェーブレット変換は、図3-1に示すように2つの帯域に分割を行う解析フィルタバンクを低周波側に再帰的に適用することで実現される。

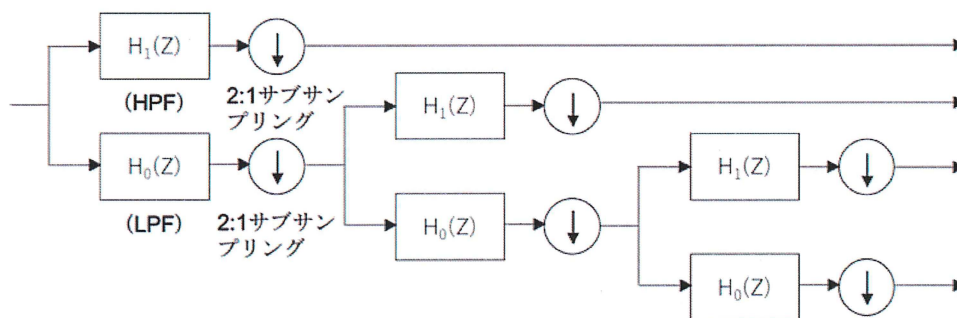


図 3-1 2分割フィルタバンクによるウェーブレット分解

画像信号の場合、フィルタバンクを水平方向と垂直方向の組にして低周波側に適用すればよい。図3-2は分割されたサブバンド、すなわちウェーブレット変換による係数の2次元配置を示している。

図 3-3 は実際の画像に対するウェーブレット変換の例である。

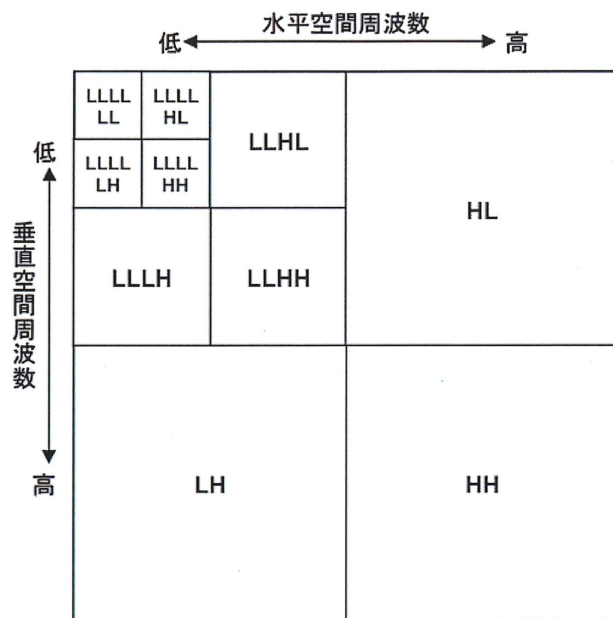


図 3-2 2次元のウェーブレット分解 (L:低周波, H:高周波)

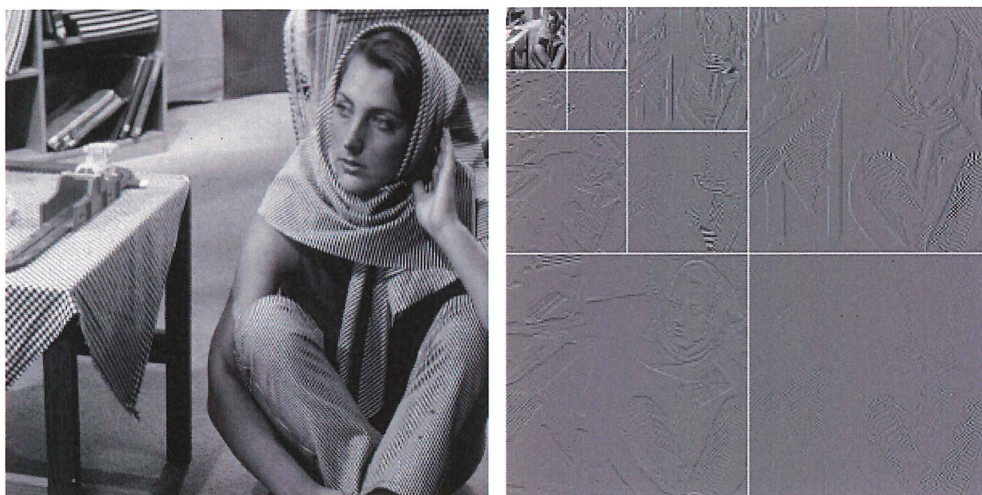


図 3-3 (a)は Barbara 原画 (512×512×8)、(b)は 2次元ウェーブレット分解された Barbara を可視化したもの

3.2.2 ウェーブレットパケット

ウェーブレット変換ではフィルタバンクを低周波側にのみ再帰的

に適用するが、高周波側にも同様に適用することができる。ウェーブレット変換のための2帯域フィルタバンクを適用することを2進分岐の木を成長させる(2分割を繰り返して枝を伸ばす)ことに対応させて考える。完全2進木は短時間フーリエ変換(STFT: Short Term Fourier Transform)に対応する。この時、完全2進木の枝を刈り込んで得られる部分木の集合全体はウェーブレット変換を含む変換の集合すなわち基底の集合に対応する。Coifmanら[3-9,3-10]はこの基底群をウェーブレットパケットと呼び、ウェーブレット基底以外にも有効な基底があることを示した。Ramchandranら[3-15]は符号量-歪み特性の面で最適な符号化を実行するウェーブレットパケットの決定方法を示した(3.3節で述べる)。

図3-4は1次元の信号におけるウェーブレットパケットの概念図である。図3-4上に示すように、2進木における親節点と2つの子節点の組が2帯域の解析フィルタバンクを表すものとする。図3-4の(a)(b)(c)は左側に帯域分割を表す2進木、右側には2進木に対応する時間-周波数平面の分割を示している。(a)(b)(c)は3種類のウェーブレットパケットに対応しており、(a)は短時間フーリエ変換、(b)はウェーブレット変換、(c)は一般の(特段の名前を持たない)ウェーブレットパケットに対応する。ウェーブレット基底では、低周波ほど時間解像度が低く(基底長が長く)、高周波では時間解像度が高い(基底長が短い)。このことからウェーブレット変換ではDCTのようなモスキート歪みが目立ちにくいことが直観的にも理解される。

3.2.3 ウェーブレット基底の最適性

ここではウェーブレットパケットの集合の中でウェーブレットが相関の強い信号に対して電力局在化の意味で優れていることを示す[3-27]。電力局在化の効果を表すパラメータを局在化利得と呼ぶことにし、2つの帯域に分割した後の信号電力の相加平均と相乗平均の比を用いる。これは3.1で変換符号化における最適符号量割当を前提とした変換利得 G_{TC} に相当する。

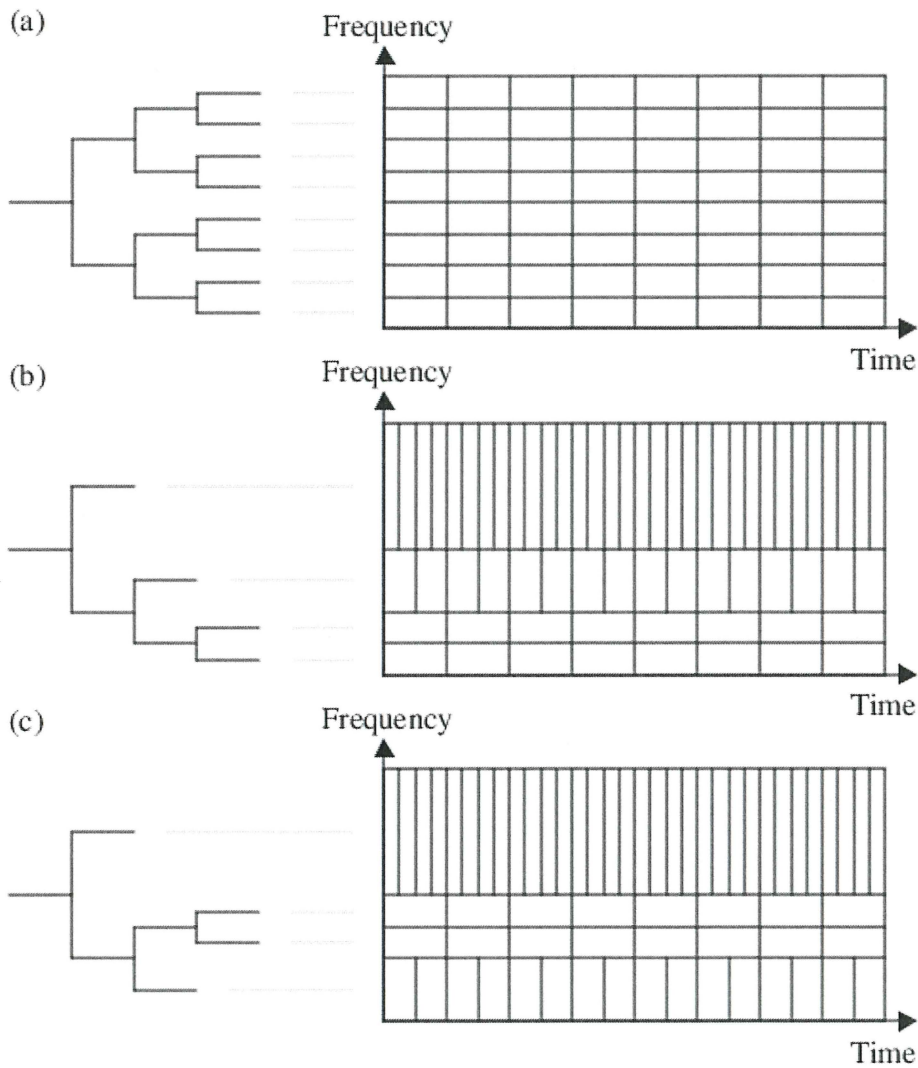
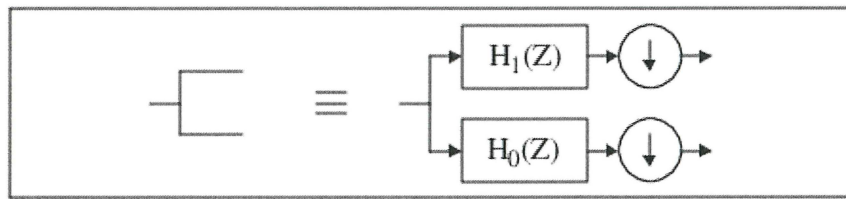


図 3-4 ウェーブレットパケットの概念図 (2 帯域分割フィルタバンクの適用により 2 進木に対応する時間-周波数平面の分割が変わることを示している。(a)は短時間フーリエ変換、(b)はウェーブレット変換、(c)は一般のウェーブレットパケットに対応する)

分割の対象とする信号源を1次元の1次マルコフ過程（相関係数 $\rho = 0.9, 0.1$ ）でモデル化し、Daubechiesの4タップフィルタを適用して2進4段の木に対応する分割を行う。この木の枝を局在化利得が最大となるよう刈り込むために一般化BFOSアルゴリズム[2-21]を用いる。

一般化BFOSアルゴリズム[2-21]はBreimanらによる最適な木の刈り込みアルゴリズム[2-47]を応用したもので、木の節点もしくは終端節点に対応づけられる複数パラメータが加算的で全体のパラメータの値に対応付けられる要件を満たす時、複数のパラメータに最適なトレードオフを与える部分木を連続的に求める手法である。すなわち求められた節点数で希望するパラメータの組合せに対応する部分木を知ることができる。一般化BFOSアルゴリズムの適用要件とここでの問題への適用の詳細については付録に記述する。

今、複数のパラメータを帯域分割数 M および局在化利得 G として相関係数 $\rho = 0.9$ の1次マルコフ信号源に一般化BFOSアルゴリズムを適用した結果を図3-5に示す。

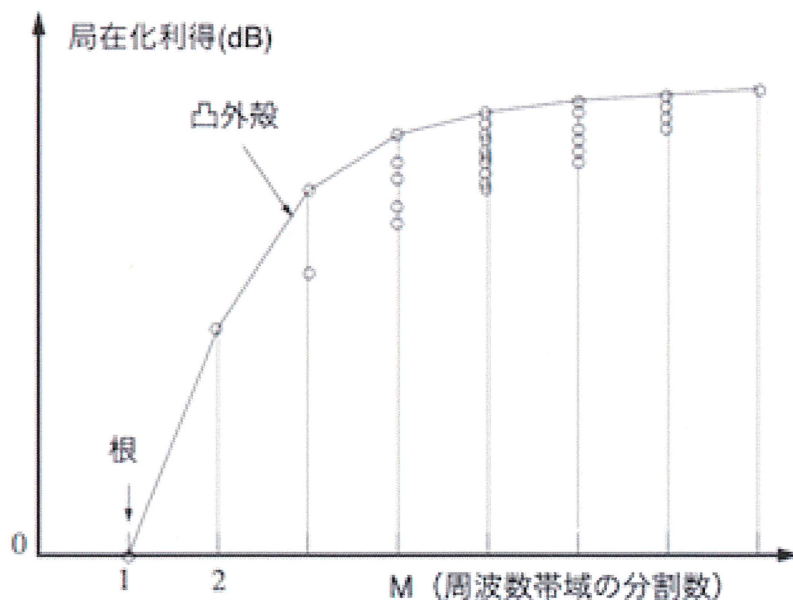


図 3-5 1次マルコフ過程 ($\rho = 0.9$) に対する刈り込みと局在化利得

図 3-5 において縦軸は局在化利得、横軸は分割数であり、○で示されている点が部分木である。各分割数における最適な部分木は実線で示された凸外殻上に存在する。

M	$\rho = 0.9$	$\rho = 0.1$
12		
11		
10		
9		
8		
7		
6		
5		
4		
3		

図 3-6 相関係数 $\rho = 0.9, 0.1$ の 1 次マルコフ過程を M バンドに分割する時に局在化利得を最大化する部分木（左が低周波側）

また図 3-6 は相関係数 $\rho = 0.9, 0.1$ の信号源について、これらを M バンドに分割する時に局在化利得を最大化する部分木の形状を示したものである。左が低周波側であり、 $\rho = 0.9$ の時にウェーブレットに対応する部分木が得られている。この結果からも、相関の高い画像信号に対するウェーブレット基底の優れた特性が読み取れる。ここで最適化された利得は電力局在化の効果に関する評価に基づくものであり、現実の符号化性能を意味するものではない。

3.3 適応的ウェーブレットパケット

3.3.1 符号量-歪み特性が最適なウェーブレットパケット

前節では局在化利得の観点で最適なウェーブレットパケットについて調べた。本節では画像符号化を行う場合に符号量-歪み特性を最適にするウェーブレットパケットについて検討する。

符号量-歪み特性が最適であるとは、使用可能な符号量で歪みを最小にすること、もしくは許容される歪み以下の歪みで符号量を最小にすることである。通常、符号量を大きくすれば歪みは小さくなるので両者はトレードオフの関係にある。最適なトレードオフを与える組み合わせを求めるために Lagrange の未定乗数法を用いる。

今、コスト関数 $J(\lambda)$ を以下のように定義する。

$$J(\lambda) = D + \lambda R$$

ここで、 D は歪み、 R は符号量を表す。 λ は Lagrange の乗数である。最適なトレードオフは $J(\lambda)$ を最小にする D と R の組み合わせで得られる。このことについて図 3-7 を用いて説明する。

図 3-7 は符号量と歪みの関係を示すグラフである。グラフに書かれた黒丸の集合は、我々が最適化しようとする符号化器および符号化結果の（ありうる組み合わせに対応する）母集号である。部分木（完全 2 進木を含む）すなわちウェーブレットパケットが決まれば量子化や符号割当が最適に選択されるものと仮定すれば、各々の黒丸は各々のウェーブレットパケットに対応する。我々が関心を持つ符号量-歪み特性が最適なウェーブレットパケットはグラフ上で可

能な限り原点に近い凸外殻の上にある。

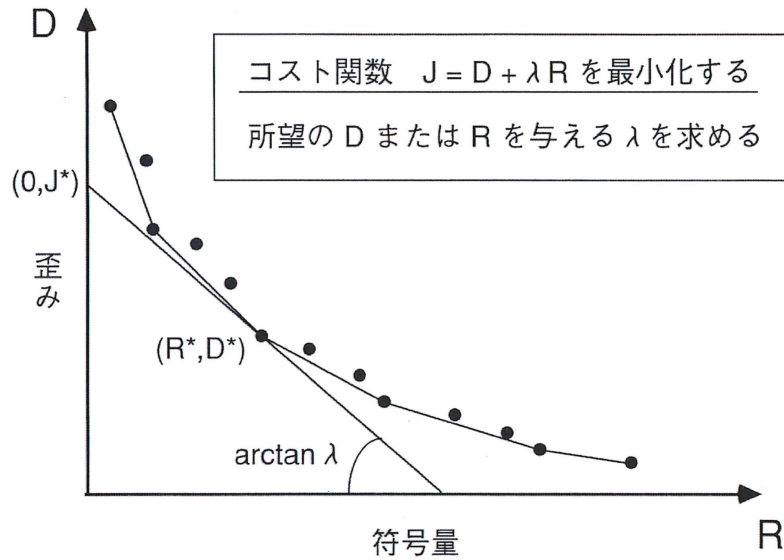


図 3-7 符号量-歪み特性の最適化

グラフの上で傾き $-\lambda$ を持ち、黒丸を通る直線が縦軸（符号量=0）と交わる切片は $J(\lambda)$ に等しい。従って $J(\lambda)$ を最小にする黒丸は凸外殻の上にある。このことから、求める最適なトレードオフは以下の手順で得られる。

- (1) ある λ について、 $J(\lambda)$ を最小にする R と D の組み合わせを求める
- (2) 与えられた符号量 R もしくは許容される歪み D を満たす λ を求める

上記(1)を実行する手法として、母集合である2進木全体（完全2進木を含む部分木）の全ての組み合わせを調べることはせず、完全2進木の終端節点から開始し、以下の条件で判定を行いながら刈込みを行う。

if ($J(\lambda)_{parent} < J(\lambda)_{child0} + J(\lambda)_{child1}$) then 刈込む
 else 刈り込まない

ここで行う2進木の刈り込みは3.2.3の場合と異なり、一般化BFOSアルゴリズムの単調増加（または減少）条件を満たすことが保証さ

れない。そのため凸外殻上にある最適な部分木が入れ子状とは限らず、 λ ごとに完全2進木からの刈り込みが必要となる。

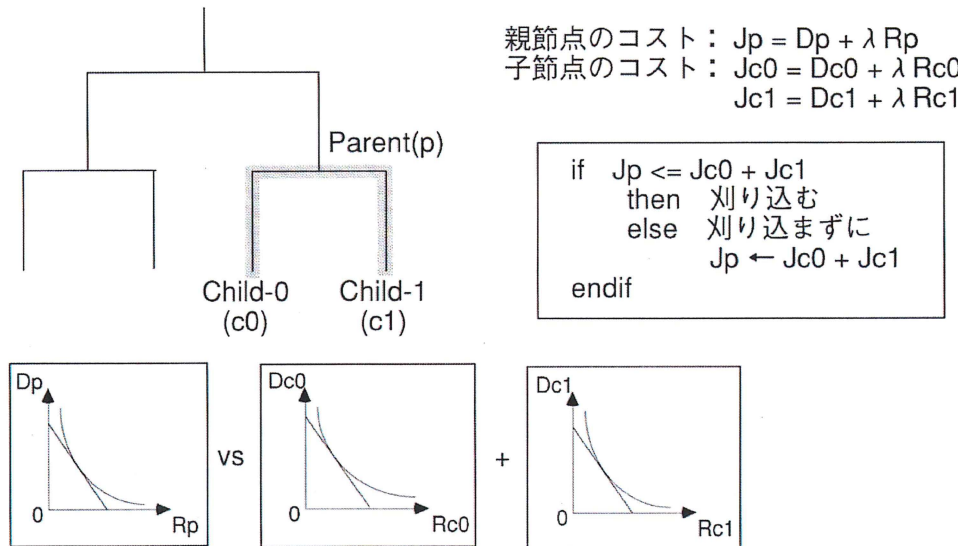


図 3-8 刈り込みの決定

上記の手法により、符号量-歪み特性の点で最適な適応ウェーブレット packets を得ることが可能になる。

3.3.2 動的に適応するウェーブレット packets

3.3.1 で得られる最適なウェーブレット packets は入力信号に依存する。一方、信号の性質は時間によって変動すると想定され、画像信号の場合には一つの画面の中でも非定常な特性の変化がしばしば観察される。そこで 3.3.1 の最適なウェーブレット packets を動的に適応させることを考える。この場合は 2 つの木を用いる。図 3-9 は図 3-4 に倣って周波数の分割と時間の分割とにおける 2 つの木のイメージを示している。 T_{WP} は周波数の分割すなわちウェーブレット packets に対応する木であり、 T_{TV} は時間分割に対応する木である。

今、完全 2 進木と刈り込みを行なった部分木 T_{WP} 全体の集合を **WP** (*Wavelet Packets*)、同様に T_{TV} 全体の集合を **TV** (*Time-varying*)

とすると、動的適応ウェーブレット packets は以下の最小化問題 $J_{min}(\lambda)$ を λ について解き、次いで条件を満たす λ を求めることに相当する [3-17]。

$$J_{min}(\lambda) = \min_{T_{TV} \in TV} \min_{T_{WP} \in WP} [D(\lambda, T_{TV}, T_{WP}) + \lambda R(\lambda, T_{TV}, T_{WP})]$$

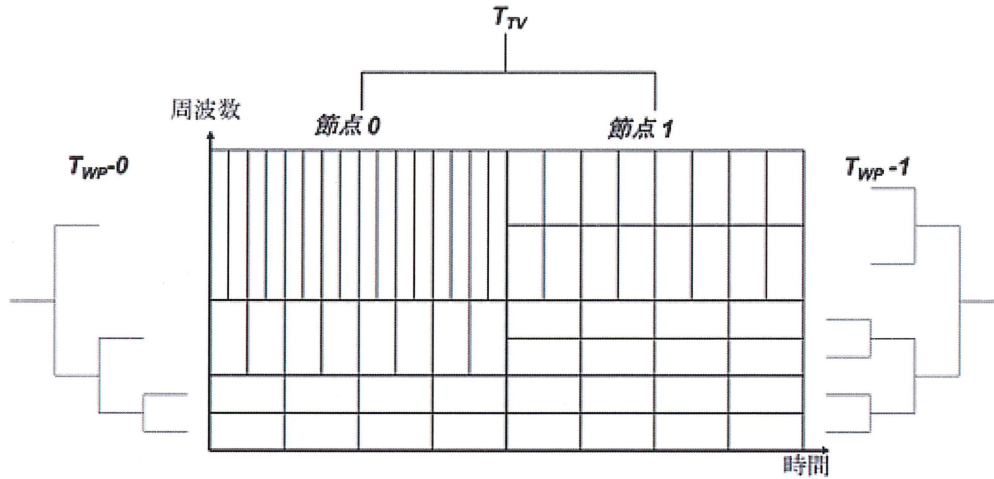


図 3-9 2つの木に対応する時間-周波数平面の分割 [3-17]

最適な T_{TV} および T_{WP} は 3.3.1 と同様に 2 進木の刈り込みによって求める。すなわち T_{WP} を T_{TV} の節点 (ある時間) において最適化し、これを前提として T_{TV} が最適な時間の分割となるように刈り込みを行う。画像に対する時間の扱いとして、4 進木の空間分割を用いる。

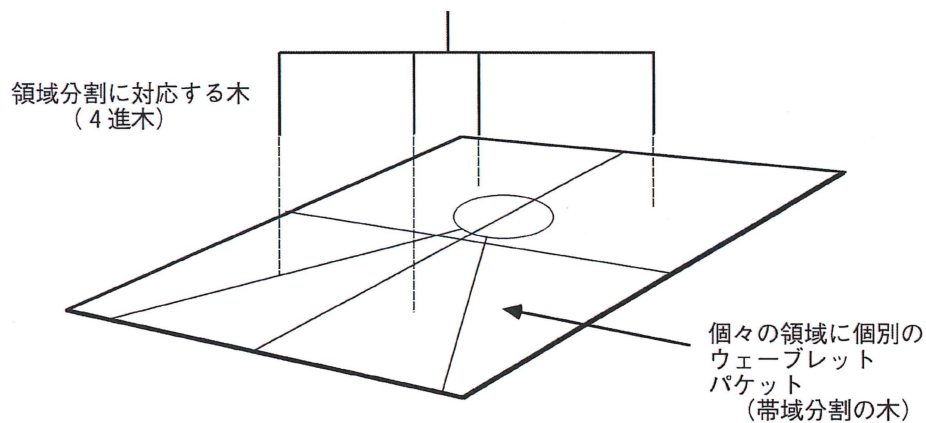


図 3-10 4 進木の空間 (領域) 分割

3.4 ウェーブレット packets による画像符号化

3.4.1 ゼロツリーの一般化

変換符号化やサブバンド符号化の符号語割当においては統計的に小電力となる高域信号の扱いを考慮することが必要である。先行する DCT 符号化ではブロックごとの EOB (End Of Block) が活用されている。EOB の本質は空間的な同一位置に相当する変換係数を低周波側から高周波側に向かって走査 (統計的には電力の大きな係数から小さな係数に向かって走査することに相当) し、ある係数より走査順序が後の係数は全てゼロであるとして走査完了を表すコードである。図 3-11 に示すように、DCT 符号化における EOB は国際標準 H.261 で導入され、その後の標準にも採用されている。低ビットレート向けの H.263 や MPEG-4 Visual では、EOB に個別の符号割当を行うことなく、ノンゼロ係数に最後の係数かどうかの 1 ビット識別情報を追加して符号割当した。これは低ビットレートでは多くの高域係数がゼロになり、ノンゼロ係数が最後のノンゼロである確率が高いからである。つまり EOB は確率に基づくエントロピー符号化の手法でランレングスのバリエーションである。

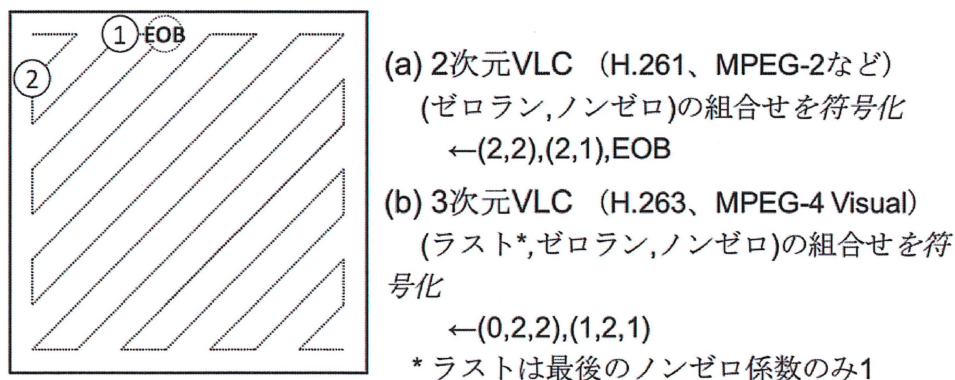


図 3-11 DCT 係数の符号化における EOB

ウェーブレット変換の場合、水平・垂直方向に 2 分割を繰り返すため、低域から高域に向かって同一空間位置に対応する係数は分割

図 3-12(b)はウェーブレット packets におけるゼロツリーの例である。ゼロツリーの走査経路を適切なクラスタ分けと順序付けで定義する。図 3-13 は 3つの方向に対応するクラスタ定義の例であり、図 3-14(a)は完全に周波数帯域の分割を行った状態で番号付けを行った例である。同図(b)のように枝の刈り込みを行ったウェーブレット packets の場合、刈り込み前の枝に割当てられていた番号の中で最小の番号が残るものとする。

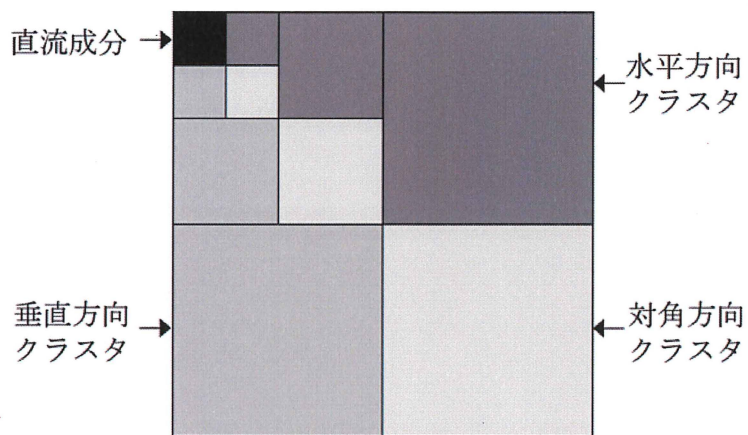


図 3-13 係数を走査する方向のグループ分け [3-17]

1	2	5	6	17	18	25	26
3	4	7	8	19	20	27	28
9	10	13	14	21	22	29	30
11	12	15	16	23	24	31	32
33	34	37	38	49	50	53	54
35	36	39	40	51	52	55	56
41	42	45	46	57	58	61	62
43	44	47	48	59	60	63	64

(a)

1	2	5	6	17	25
3	4	7	8		
9		13		21	29
33		37		49	
41		45			

(b)

図 3-14 係数の走査順序 [3-17]

3.4.2 量子化器の決定と符号語の割当

量子化対象の周波数領域における符号量-歪みの関係を以下の式と仮定する。ただし κ は信号の確率分布に依存する定数である。

$$D \approx 2^{-2R} \sigma^2 \kappa$$

この曲線に傾き $-\lambda$ の直線が接する点の R と D が求める組合せである。上記曲線の傾きは次のように計算される。

$$\frac{\partial D}{\partial R} = -(2 \ln 2) \cdot D$$

傾き λ の条件から目標とする歪み D_{target} を以下とする。

$$D_{target} = \frac{\lambda}{2 \ln 2}$$

一様分布に対する量子化ステップサイズ Δ と歪みの関係 $D = \Delta^2/12$ からステップサイズの初期候補 Δ_0 を以下のように推定する。

$$\Delta_0 = \sqrt{12 D_{target}} = \sqrt{\frac{6\lambda}{\ln 2}}$$

Δ_0 を用いて D_0 と R_0 を計算し、次の候補 Δ_1 を以下のように求める。

$$\Delta_1 = \sqrt{\frac{D_{target}}{D_0}} \cdot \Delta_0 = \sqrt{\frac{\lambda}{2 \ln 2 D_0}} \cdot \Delta_0$$

以降、 Δ_2 も同様に求められる。実験では Δ_0 から Δ_2 までの量子化において最小コスト $J_k = D_k + \lambda R_k$ を与えるステップサイズを使用した。

符号語の割当については量子化値のラプラス分布性を仮定し、[3-13,3-14]を参照して以下のようにクラス分けを行った。

表 3-1 符号語の割当

クラス	インデックス	ビット幅	クラス符号語
0	0, ゼロツリー	1	0
1	-1, 1	1	10
2	-3, -2, 2, 3	2	110
3	-11 ~ -4, 4 ~ 11	4	1110
4	<-11, >11	8	1111

3.5 分割結果と符号化性能

3.5.1 周波数帯域-空間分割の最適化結果

前節のアルゴリズムに従って、周波数帯域-空間分割を符号量-歪み特性の観点で最適化した時の分割結果と符号化性能の例を画像 Barbara(512×512画素×8ビット)について図 3-15 に示す[3-17]。Barbara は画像の領域によって信号電力が集中する周波数帯域が異なると期待できる画像である。またウェーブレットと比較した符号化結果について表 3-2 に示す。

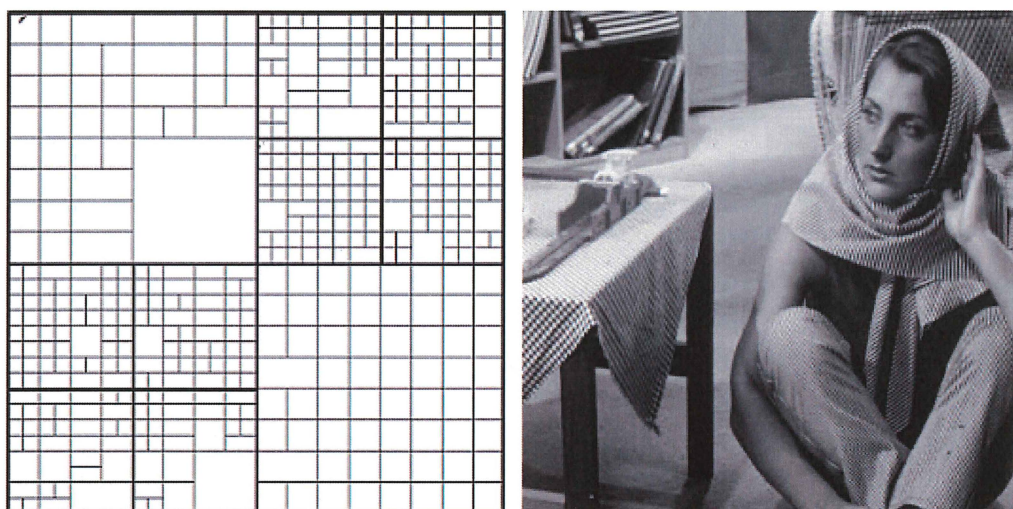


図 3-15 (a) Barbara の分割 (太線は空間の分割、細線は周波数帯域の分割) (b) Barbara の復号画像 (1.02bit/pel, SNR=35.2dB)

表 3-2 動的適応ウェーブレットパッケージによる符号化結果

基底	個数		符号量 (bit/pel)	SNR (dB)
	ノンゼロ係数	ゼロツリー		
ウェーブレット (固定)	50326	18270 (8.7)	1.06	34.15
ウェーブレットパッケージ (空間分割あり)	41596	7133 (18.9)	1.02	35.23

フィルタバンクには Daubechies の 4 タップフィルタを用い、4 進 2 段の T_{WP} 、2 進 4 段の T_{TV} に対して刈り込みを行なった。ターゲットとしたビットレートは 1.0bit/pel である。ゼロツリーの走査には図 3-13 のクラスタと図 3-14(a)の順番付けを用い、最終的な符号割当結果に対して算術符号化を行った。符号量には部分木の形状を表す情報を含んでいる。

表 3-2 より以下のことがわかる。

- ・ 動的適応ウェーブレット packets はウェーブレットとの比較において 1dB 以上優れている
 - ・ 動的適応ウェーブレット packets のノンゼロ係数がウェーブレットより 14%程度少ないことは信号電力がより少ない係数に局在化されたことを示している
 - ・ 動的適応ウェーブレット packets においてゼロツリーの数が半分以下で平均長（ゼロツリーに含まれるゼロ係数の平均個数）が 2 倍以上となっていることも符号化性能に貢献している
- 一方、同じ表から、ゼロツリーに含まれないゼロ係数は以下のように推定される。

- ・ ウェーブレット :

$$512 \times 512 - 50326 - 18270 \times 8.7 = 52869$$

- ・ 動的適応ウェーブレット packets :

$$512 \times 512 - 41596 - 7133 \times 18.9 = 85734.3$$

係数の走査経路上における最後のノンゼロ係数より前に、このように多数のゼロ係数が存在することは、走査の順序、複数係数をベクトル化した量子化ないし符号割当、もしくはゼロツリーの長さを符号量-歪み特性の観点で最適化する余地があることを意味している。3.5.2 ではこれらの観点でさらに最適化を行う手法を示す。

3.5.2 周波数帯域-空間分割およびゼロツリーの最適化結果

Xiong らは符号量-歪み特性の観点によるゼロツリーの最適化を含む量子化最適手法を SFQ (Space-Frequency Quantization) と呼

び、[3-19]においてウェーブレット係数、筆者が共著者の[3-20]においてウェーブレット packets に適用した。さらに Ramchandran らは筆者が共著者の[3-21]において、3.5.1 の動的適応ウェーブレット packets に SFQ を適用した。SFQ は符号量-歪み特性の観点でノンゼロ係数をゼロに切り捨て、量子化ステップサイズを最適に選択することで、3.5.1 の符号化結果に残存した「ゼロツリーに含まれないゼロ係数」の符号量を削減する。SFQ の最適切り捨ては Ramchandran ら [3-18] が DCT 係数について行なった最適切り捨てのウェーブレット係数およびウェーブレット packets 係数への適用に相当する。ノンゼロ係数の切り捨てはすなわちゼロツリーの最適な成長であり、SFQ はゼロツリーの最適化を伴う量子化に相当する。ゼロツリーの最適化は前節の議論と同じように歪み D と符号量 R を Lagrange の乗数 λ で結合したコスト関数 $J = D + \lambda R$ を最小化する原則に基づき、以下の手順で行う。

[ゼロツリーの最適化]

- ・ ウェーブレット変換あるいはウェーブレット packets においてまず係数の初期的な量子化を行い、ノンゼロ係数とゼロ係数を生成する。
- ・ 上記量子化によって形成される 2 つのツリーを観察する。一つはゼロツリーであり、もう一つはノンゼロ係数およびゼロツリーに含まれないゼロ係数からなるツリーである（ゼロツリーに含まれないゼロ係数とは、走査経路において当該係数より高域方向にノンゼロ係数が存在するゼロ係数である）。
- ・ 上記後者のツリーの節点に存在するノンゼロ係数について、当該係数および子孫に相当する節点のノンゼロ係数をゼロと切り捨てることによる符号量-歪み特性の変化 ($\Delta D / \Delta R$) を λ と比較し、コスト関数 $J = D + \lambda R$ を小さくするなら切り捨てる。

[3-20]では周波数分割の木 T_{WP} の各節点においてゼロツリーの最適化を行い、 T_{WP} を最適に決定した後に最終的な量子化ステップサイズをコスト関数 $J = D + \lambda R$ で評価して最適化した。式で表すと以

下となる。

$$J_{min}(\lambda) = \min_{q \in Q} \min_{T_{WP} \in W_P} \min_{ZT} [D(\lambda, q, T_{WP}) + \lambda R(\lambda, q, T_{WP})]$$

ここで $\min_{ZT}(\cdot)$ はゼロツリーの最適化、 $\min_{q \in Q}(\cdot)$ は最適量子化ステップサイズの決定を表す。この段階でも固定的なウェーブレットの場合 [3-19] より SNR で 0.7dB 程度優れた符号化性能が示されている。図 3-16 は [3-20] の方式による符号化器の説明図である。

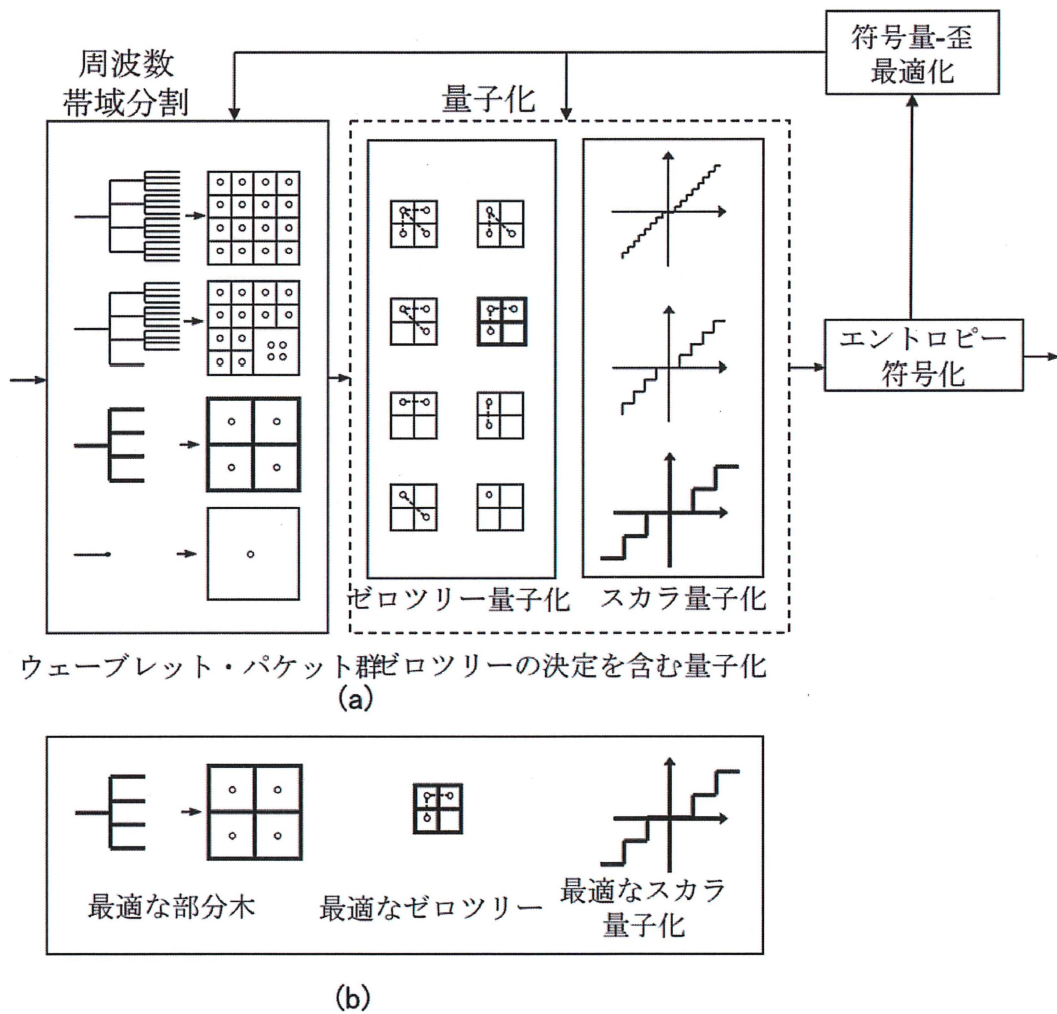


図 3-16 ゼロツリーの最適量子化を含むウェーブレットパケットの最適化による符号化系 [3-19]、(a) はブロック図であり、(b) は最適化された要素の例を表す (ウェーブレットパケット、ゼロツリー、量子化)

[3-21]では動的適応ウェーブレット packets による周波数帯域-空間分割およびゼロツリーの最適化を含む量子化が行われた。ここで行われた最適化は以下に相当する。

$$J_{min}(\lambda) = \min_{q \in Q} \min_{T_{TV} \in TV} \min_{T_{WP} \in WP} \min_{ZT} [D(\lambda, q, T_{TV}, T_{WP}) + \lambda R(\lambda, q, T_{TV}, T_{WP})]$$

[3-21]の符号化結果を表 3-3 に示す。同表には Shapiro による EZW の符号化結果[3-14]も示している。また図 3-17 には画像 House に対する空間分割結果と、絵柄確認のために画像 Lena の原画を示している。これらの結果より以下のことがわかる。

- ・ 入力画像に依存する周波数帯域-空間分割およびゼロツリーの最適化により、ウェーブレット符号化のリファレンスである EZW を 1.0~2.0dBdB 程度上回る性能が得られる
- ・ 画像 Lena ではウェーブレット、ウェーブレット packets、動的適応ウェーブレット packets の性能差がわずかである
- ・ 画像 Barbara ではウェーブレット、ウェーブレット packets、動的適応ウェーブレット packets の性能差が最も大きい

表 3-3 符号化性能の比較 ([3-21]の TABLE II より)
EZW による Lena と Barbara に関する数値は[3-14]より引用

基底	Lena Barbara House				Lena Barbara House				Lena Barbara House			
	符号量 (bit/pel)	SNR (dB)	SNR (dB)	SNR (dB)	符号量 (bit/pel)	SNR (dB)	SNR (dB)	SNR (dB)	符号量 (bit/pel)	SNR (dB)	SNR (dB)	SNR (dB)
EZW	1.0	39.55	35.14	-	0.5	36.28	30.53	-	0.25	33.17	26.77	-
ウェーブレット	1.0	40.51	36.94	37.61	0.5	37.31	31.16	32.80	0.25	34.24	27.12	29.70
ウェーブレット packets	1.0	40.55	37.13	37.83	0.5	37.40	31.53	32.88	0.25	34.27	27.32	29.70
動的適応 ウェーブレット packets	1.0	40.55	37.58	38.01	0.5	37.40	32.13	33.00	0.25	34.27	27.85	29.70

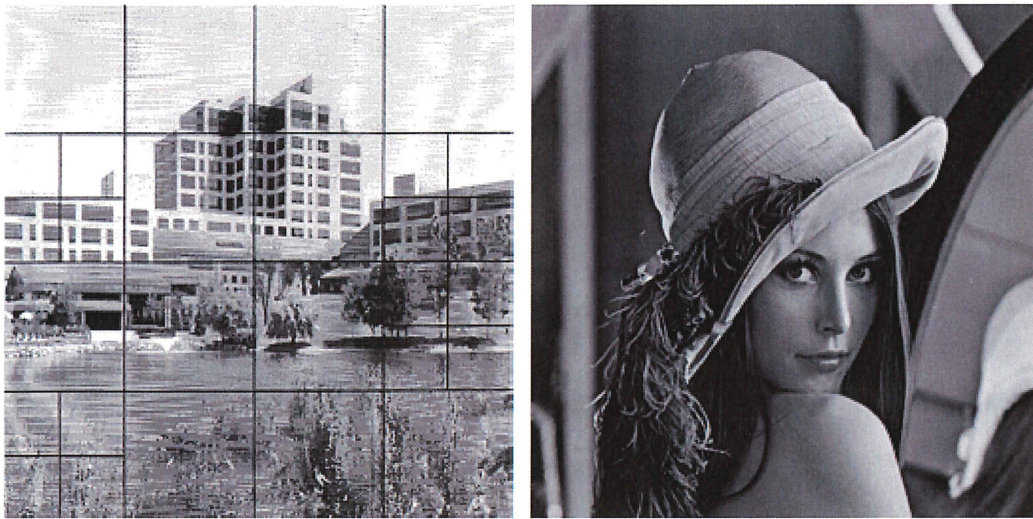


図 3-17 動的適応ウェーブレットパケットと SFQ による分割結果
(a)の分割結果は文献[3-21]による、(b)は Lena 絵柄確認のための原画

Lena における性能改善が小さいことは最適なウェーブレットパケットがウェーブレットとほぼ一致し、動的適応の空間分割においてすべての枝が刈り込まれた結果である。このことは同画像が比較的定常に近い画像であること、ウェーブレット基底が有効であることを示している。また、4進分割木では空間分割に制約が強すぎることも想像される。Barbara における性能改善が大きいことは、同画像が各種の高周波成分を含む標準画像を意図して作られたからであり、図 3-15(a)の最適分割結果からもそのことが了解される。図 3-17(b)の House における空間分割結果は直観的にも了解性の高い結果である。

3.6 周波数帯域と空間の分割による信号表現

3.6.1 試験信号の最適分割による検証

本節では前節までに検討した周波数帯域と空間の分割について、圧縮符号化だけでなく信号の特性を表現する側面について検討する。なお 3.6 節で用いるデータはすべて 3.5.1 における最適化による。すなわち SFQ を用いないデータである。

最適なウェーブレット packets は信号電力が帯域間で大きな差を持つ時に分割を行う。可視化されたウェーブレット packets の分割木は微分で表された電力分布のマップであると言える。最適な空間分割の木は領域間で電力分布が異なる時に分割を行う。空間分割は電力分布の違いを領域で塗り分けるマップであると言える。

上記を直観的に確認するため、CZP (Circular Zone Plate) を入力画像として動的適応ウェーブレット packets の分割を行なった結果を図 3-18 に示す。同図(a)は入力とした CZP であり、左上を原点 (0,0)、水平右方向を x 軸、垂直下方向を y 軸として、以下の式で 512×512 画素の画像を生成した。A=128, N=512 である。

$$P(x,y) = A \cos \left[\left(\frac{x^2}{2N} + \frac{y^2}{2N} + 1 \right) \pi \right] + 128$$

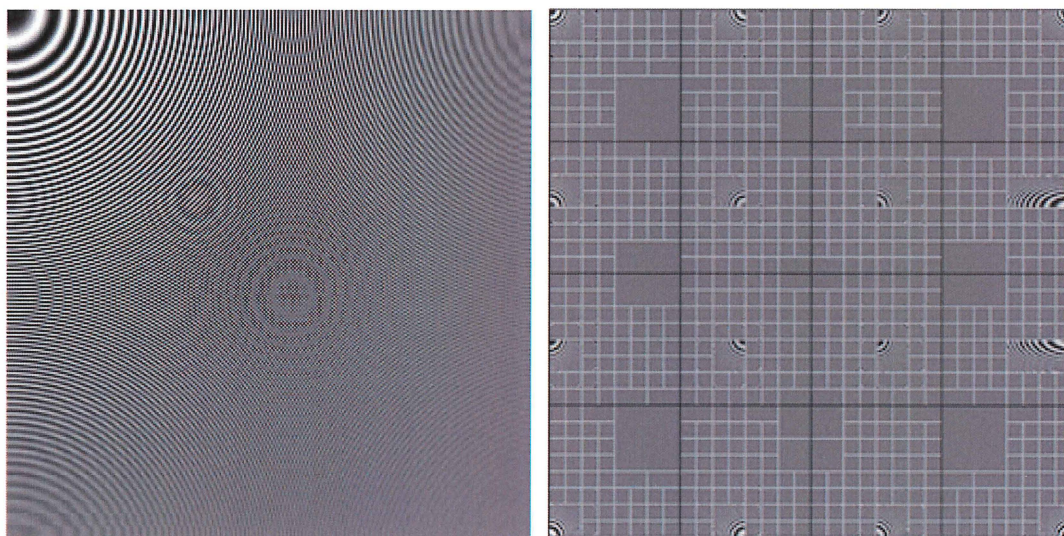


図 3-18 CZP と動的適応ウェーブレット packets による分割

(a) CZP (左上原点の空間周波数を(0,0)とする線形掃引)

(b) 空間-周波数分割された CZP (黒線は空間、白線は周波数分割)

CZP は光学機器の周波数特性を測定するためのテストチャートからテレビジョンシステムの評価にも用いられるようになった空間周波数の 2 次元線形掃引 [3-23] であり、ここではその四半分を用い

ている。2次元空間周波数(ω_h, ω_v)は原点(0,0)で $\omega_h=0, \omega_v=0$ 、原点からの距離に従って変化し、画像の右上(N-1,0)では1画素ごとに余弦関数の位相が反転し、左下(0,N-1)では1ラインごとに位相が反転する。

(b)の分割結果は空間が設定された条件の中で空間分割の木が刈り込まれず、等方な分割となっている。空間上の位置に依存して信号電力が特定の周波数帯域に偏在している本画像に関する合理的な分割結果である。各空間領域では、位置に対応して最大電力を持つ周波数領域がピークの平坦性により若干刈り込まれ、電力が最も小さい(周波数分割において最大周波数領域から離れた)周波数領域が大きく刈り込まれており、それ以外の周波数領域は信号電力の急峻な減衰による微分の大ききで刈り込まれない結果となっている。

動的適応ウェーブレットパッケージによって非定常性を顕在化できるという特長は、圧縮と同時に異常検知などの信号解析を行う可能性を示唆している。例えば菊池ら[3-24]は心電図波形の解析と圧縮にウェーブレットを用い、ガソリンエンジンの異常燃焼振動の判別にウェーブレットパッケージを用いる自らの事例を紹介した。

もう一つの可能性は動的適応ウェーブレットパッケージの分割結果から部分木の構造と適当な付加情報ないし条件だけで信号復元を行う(つまり係数の網羅的符号化を行わない)ことである。こうした方向の先行例として、ウェーブレットの零交差や極値を用いた信号復元のアプローチ[3-26,3-29]がある。

3.6.2 符号量-歪み特性のトレードオフと最適分割

図3-19は画像Barbaraに対して λ を変化させた際の分割の変化を示している。 λ の変化により符号量と歪みのトレードオフは変化する。(d)は絵柄と分割の関係を見比べられるように加えている。これらの観察から最適分割は符号量の条件に対して敏感に反応することがわかる。入力依存ではあるが、入力画像が与えられた時に最適な分割が一意に決まるものではない。従って当面は動的な符号量

制御に適していない。

符号化シミュレーションでは分割木や分割された帯域や空間の符号化パラメータを表す付加情報を符号量に加えている。目標符号量が極めて小さい時や分割後の信号に対して量子化器や符号語割当が適応できなかつた場合、こうした付加情報は相対的に分割を刈り込む方向に働きうる。一方で目標符号量に関わらず、最適分割は固定的な分割、例えばウェーブレットを選択肢に含み、なお符号化利得の得られる条件を選び出しているのである。

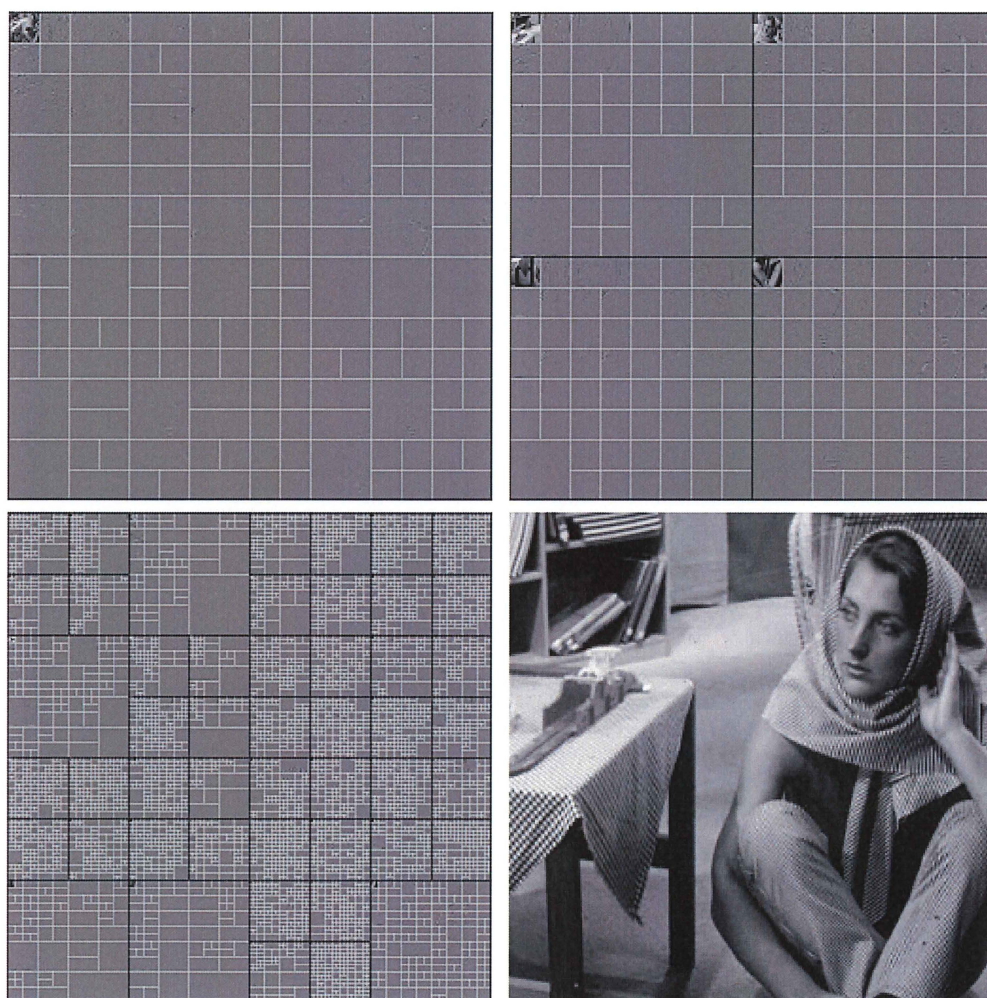


図 3-19 画像 Barbara に対する動的適応ウェーブレット packets (左上、右上、左下、右下の順に(a)(b)(c)(d)) (a)(b)(c)は空間-周波数領域の分割結果、(d)は参照用の原画)

画像 Barbara は女性の衣服やショール、テーブルクロスや網目など、異なるテクスチャを含んでいる。図 3-19 (c)は圧縮符号化の最適性と同時に信号の電力分布と非定常性を図示している。直観的にテクスチャが強いと感じられる箇所では周波数帯域が細かく分割され、テクスチャが異なると思われる箇所では空間も分割されている。このことから、圧縮と同時に、テクスチャの分類や抽出への提要が有効と想像できる。

3.7 この章のまとめ

本章では画像符号化の性能を高めることを第1の目標として動的に適応するウェーブレット packets を画像に適用して優れた圧縮符号化特性が得られることを示した。同時に、符号化結果として得られる周波数帯域と空間の分割が信号電力の分布をよく表現していることから、圧縮と信号解析を同時に行う可能性を指摘した。本章で得られた知見は以下の通りである。

- ・ 現実の量子化器と符号語割当を用い、符号量-歪み特性の観点で動的適応ウェーブレット packets を最適化することにより、効率のよい画像符号化が実現可能である
- ・ 最適化された動的適応ウェーブレット packets は周波数帯域と空間の最適な分割を与える
- ・ 上記最適化の結果として得られる分割は目標符号量の条件によって変動する
- ・ 上記最適化の結果として得られる分割は画像信号の電力の分布や非定常性を表現しており、信号解析として活用の可能性がある最適な動的適応ウェーブレット packets が条件によって異なることから、動画のように実時間で符号量制御を行う必要のある場合には本方式を適用し難いようにも思われる。しかしこの問題は複数のウェーブレット packets を予め定義しておき、適宜切り替えることで複雑度と利得のバランスを取ることが可能だろう。従来、ウェー

ウェットが動画の符号化に不向きとされた理由には、ブロック単位の動き補償予測がブロック境界にエッジの立った予測誤差画像を生成するという問題もあった。この点についても、オーバーラップ動き補償やアフィン変換を含む動き補償予測が現実的になるほどにデバイスが進化したことを考えれば、動画符号化へのウェットおよびウェットパケットの適用検討が行われていい時期である。また、信号電力の分布を顕在化するという動的適応ウェットパケットの特質は、信号解析や信号表現の道具として検討するに値するものである。

第4章 動き補償予測付き DCT を用いる画像符号化標準

4.1 はじめに

画像符号化技術の研究開発に関する歴史は国際標準化の歴史に重なる部分が多い。第1章で述べたように、画像を効率よく伝送したいという要求は古くから存在した。1950年代に米国のベル研究所でテレビジョン信号の圧縮技術が研究されたのは偶然のことではない。音声による電話サービスが実現された後、互いの顔を見て会話するテレビ電話が発想されたのは自然なことであり、電話会社の研究所が先駆的な研究を行うことは必然でもあった。電話に代表される通信サービスは普及が相互接続性に依存することから、標準化は必須であり、通信分野で先行した。

通信分野における画像符号化の国際標準化は通信事業者を中心とする ITU-T (1993年以前は CCITT) が主導している。ITU-T はテレビ会議・電話用の符号化・伝送フォーマット・プロトコル・端末定義など、多くの規格を勧告化してきた。放送分野における標準化は放送事業者を中心とする CCIR およびに CCITT と CCIR の合同委員会である CMTT がかつて主導し、現在画像符号化の関連はほぼ ITU-T に集約されている。蓄積分野、例えば DVD などのパッケージやレコーダーについては機器メーカーからの参加者が大きな役割を果たしている。ISO 傘下で SC 2/WG 8 に属し、CCITT と共同で静止画符号化を担当した JPEG[4-15]の標準はデジタル写真の代名詞的な標準となった。同じく WG 8 に属していた MPEG は動画・音響符号化を担当し、蓄積メディア対象から始まり、通信・放送を含む汎用的な動画符号化について標準化を行う中心的存在となった。1991年には WG 8 が ISO/IEC JTC 1/SC 29 に昇格して今日に至っている。WG 8 および SC 29 の幹事国および議長は日本が務めており、MPEG、JPEG 標準化に大きく貢献している(筆者は 2006 年から SC 29 の議長を拝命している)。動画符号化に関係する組織の協力関係を図 4-1 に示す。

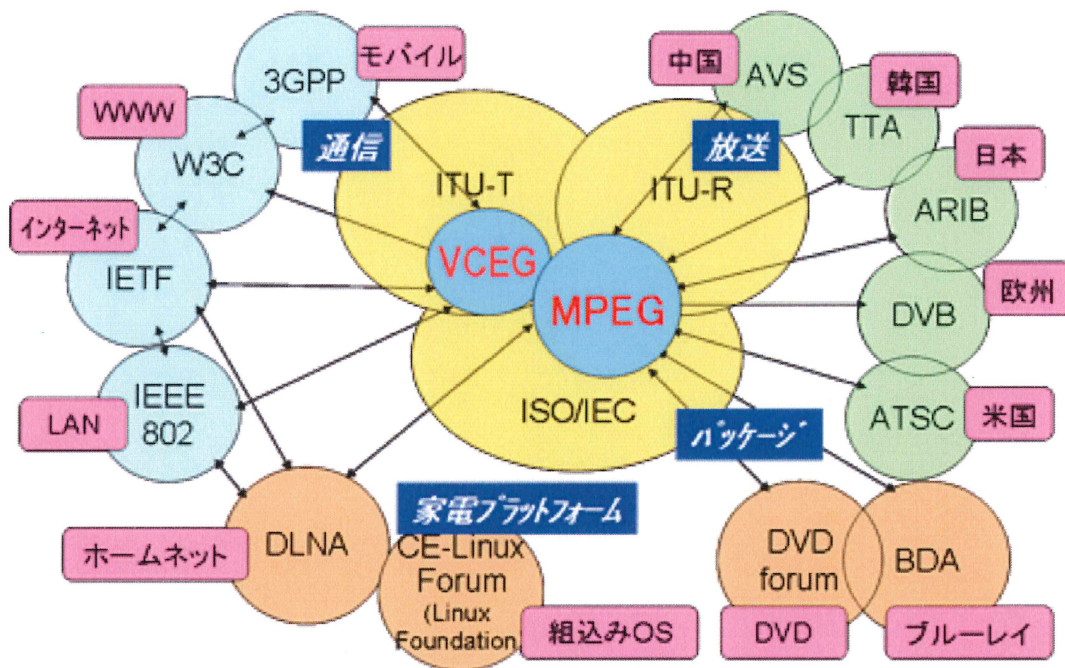


図 4-1 動画符号化に関する標準化組織[1-36]

国際標準化の利点は相互接続性や相互運用性が保証され、機器やコンテンツが広域に渡って流通することである。サービスに必要な技術が国際標準となることで、仕様がメンテナンスされ、周辺技術の進歩とともに更新される可能性が高まる。品質や安全性も考慮される。TBT 協定など国際調達の条件としての役割も大きい。機器・システムの提供者にとって、製品の市場が世界規模で大きいことは魅力であり、量産効果によるコストダウンも可能になる。ユーザにとっては機器が安価に入手できる上に継続的に使えることが期待でき、サービスの品質が高いことは利点となる。

画像符号化の標準策定は既存技術から優れたものを選択するのではなく、多くの参加者による提案を基に開発を行う手法をとっている。こうした開発のプロセスは世界と競争すると同時に協力するオープンイノベーションの効果も有している。

本章では次節以降、国際標準化の歴史を簡潔に述べ、主な国際標準とそれらの社会への貢献について説明し、国際標準が継続的な技

術な進歩を取り込んできたこと、筆者の活動と貢献について述べる。さらに現時点で主要かつ最新の国際標準 HEVC|H.265 に筆者と共同研究者が行なった提案方式とその性能を示し、最後に符号化性能と方式の複雑度について考える。

4.2 国際標準化の歴史

本節では画像符号化に係る標準化の歴史を簡単に振り返る。図 4-2 は MPEG 標準を中心に動画符号化の国際標準を圧縮率やアプリケーションとともに時系列で示している。MPEG 標準は ITU-T 勧告 H.261[4-8]で採用された動き補償予測付き DCT 符号化を共通の骨格とし、CD-ROM への適用を想定した MPEG-1[4-9]、汎用テレビジョンの MPEG-2|H.262[4-10]、モバイルへの適用を想定した MPEG-4[4-12]、ストリーミングの AVC|H.264[4-13]と改善のための追加や改訂を続け、符号化性能を向上させている。

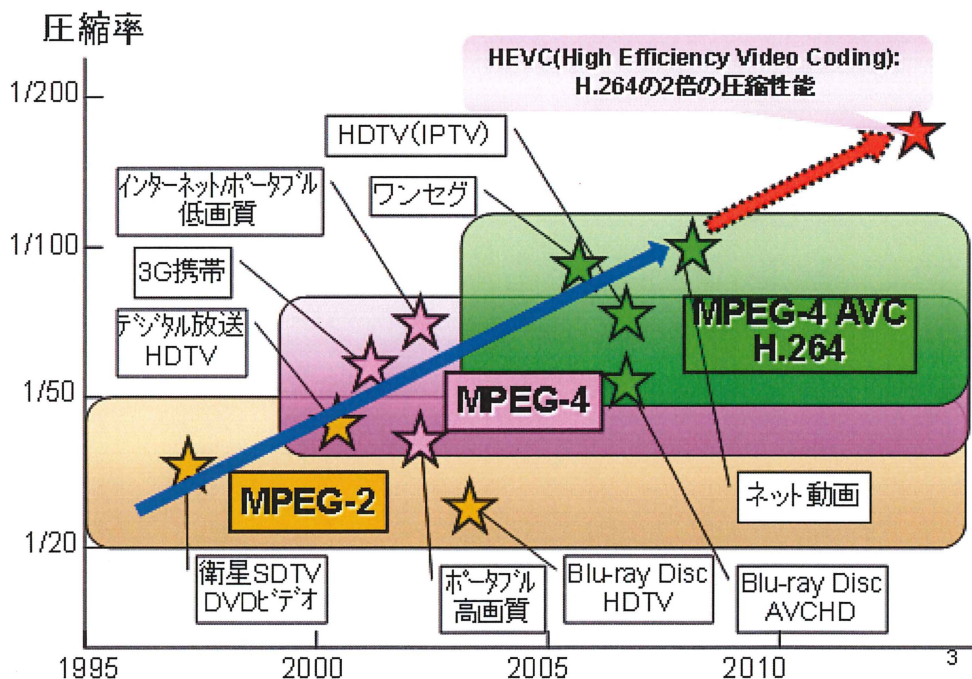


図 4-2 MPEG 動画符号化標準の進展 [1-36]

図 4-3 は MPEG 映像符号化標準と ITU-T 勧告との関係を示している。MPEG-2|H.262 と AVC|H.264 は ISO/IEC と ITU-T との協力の産物である（H.26x は ITU-T 勧告のシリーズ番号）。1994 年に標準化された MPEG-2 | H.262 は SDTV、HDTV とともに現行デジタル放送や DVD ビデオに採用されている。1999 年に標準化された MPEG-4 は低レートかつ低品質の回線に対するロバスト性を重視して策定され、第 3 世代携帯電話の動画やストリーミング、デジカメ動画などに適用された。2003 年に標準化された AVC|H.264 は細かな改善要素を盛り込んで圧縮率を大きく向上し、ブルーレイ、IPTV、スマートフォン動画などに適用され、市場に浸透している。

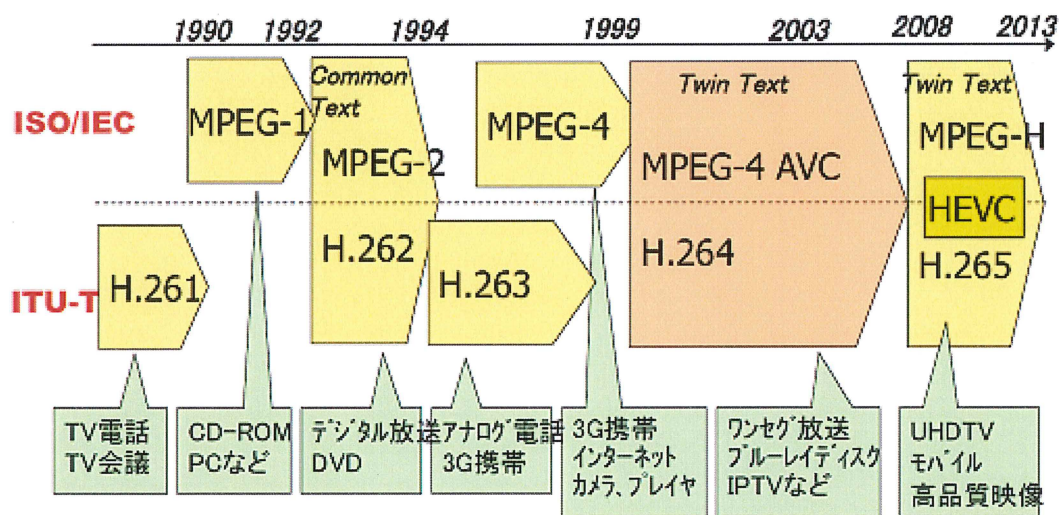


図 4-3 動画符号化に関する MPEG 標準と ITU-T 勧告 [1-36]

動画に比べて演算負荷や実時間伝送への要求度が低い静止画の符号化標準である JPEG は共通の枠組みを使わず、DCT を用いた最初の JPEG 標準が 1992 年、ウェーブレットを用いた JPEG 2000 [4-16] が 2002 年、オーバーラップ変換を用いた JPEG XR [4-17] が 2011 年に、それぞれ国際標準が出版された。これらの性能と複雑度は図 4-4 の関係にある。

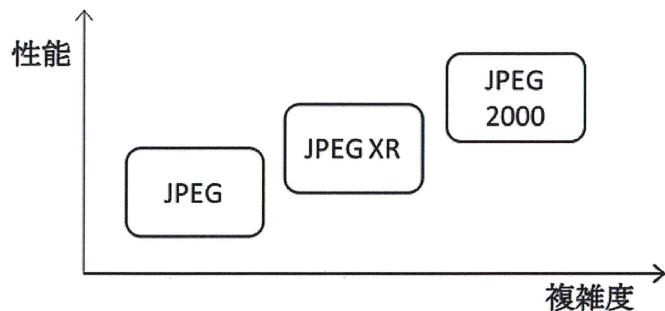


図 4-4 静止画符号化に関する JPEG 標準ファミリー

4.3 動画像符号化国際標準化への貢献

4.3.1 独自方式から H.261 へ

第 2 章で述べたように、筆者らは 1980 年代にベクトル量子化を用いた独自方式でテレビ会議システムを開発した。最初の製品は 1985 年に発売された。国際標準については 1984 年に 1.5-2Mbit/s 向け H.120[4-7]が策定され、 $n \times 384\text{ kbit/s}$ 、 $m \times 64\text{ kbit/s}$ の低ビットレート向け標準化活動に移行しつつあった。筆者らは 80 年代後半に標準化活動に参加した。国際標準は DCT 採用に進行したが、図 4-5 の枠組みで捉えれば、システム要件の多くが共通していた。

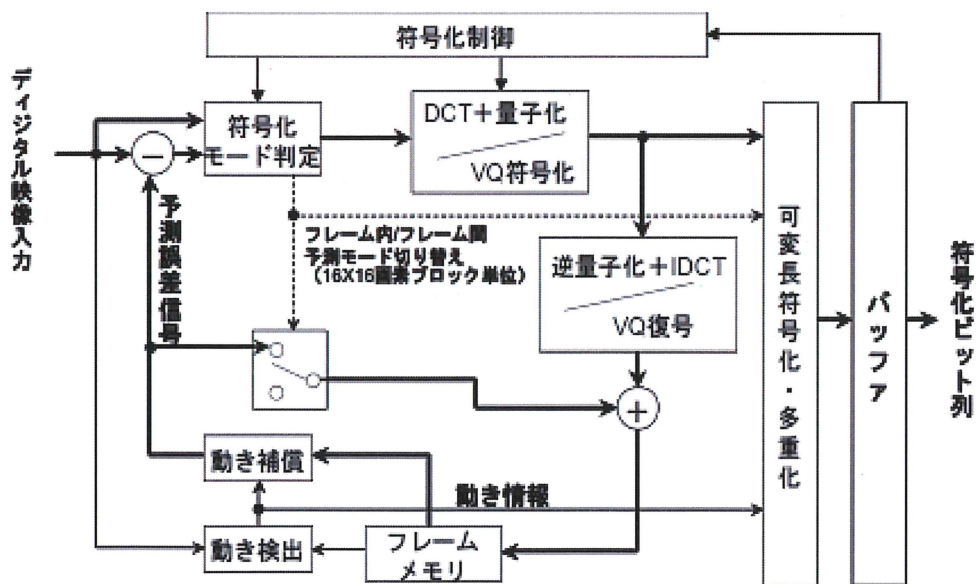


図 4-5 DCT とベクトル量子化 (VQ) を包含する枠組み

1990年に勧告化された標準 H.261 は動き補償予測付き DCT を採用した。同標準における動き補償と次段符号化を階層的ブロックで行うマクロブロック構造、マクロブロック内の有効/無効ブロック識別情報の可変長符号化、マクロブロック単位のイントラ（フレーム間予測を用いない）モード、雑音蓄積を除去する（動き情報で特性を変化させる）ループフィルタ、フィードバック制御、符号量が増大した時のコマおとしなどは細部で異なるにせよ、実質的に筆者らの製品に導入済みであった。

国際標準として直面した相互接続性の課題は IDCT ミスマッチとバッファ制御であった。前者は実数演算である DCT の逆変換において演算精度の相違により誤差が蓄積する問題である。標準では周期的なリフレッシュという対症療法が採られた。後者も DCT に顕在的で、信号がエッジやランダムノイズ状のパターンを含む場合に電力局在効果が低下して係数情報が大きくなる、つまり符号量の変動が大きいため平滑化バッファの制御に制約が必要という課題である。同課題に対して考えた方式[4-18]について説明する。

動画像の各フレームにおいて発生する符号量を B_i 、フレーム周期を T 、ビットレートを R とすると、受信バッファにおいて消化しきれずに蓄積する符号量は以下で表される。

$$\sum_{i=0}^{n-1} (RT - B_i)$$

この値が受信バッファの容量を超えるとオーバーフローとなるため、第 1 のしきい値 $Th_1 > 0$ を設定する。

$$\sum_{i=0}^{n-1} (RT - B_i) > Th_1$$

符号化側では上記不等式が真とならないよう、 B_i を制御しなければならない。逆に符号量が大きいため受信側に 1 フレーム分の符号情報が届けられない場合を次のように考える。以下、 $Th_2 > 0$ である。

$$\sum_{i=0}^{n-1} (RT - B_i) < -Th_2$$

上位不等式が真となれば、 $B_n = 0$ として第 n フレームをコマ落としにして上記不等式が成立しなくなるようタイミングを調整する。

こうしたバッファ制御の制約はバッファオーバーフローやアンダーフローを排した回避して相互接続性の保証に必要であり、標準に盛り込まれた。バッファ制御に関する論文作成はなく議論と特許[4-18]のみである。(同特許は後に MPEG-2 の必須特許と認定)

図 4-6 は H.261 符号化器のブロック図である。以降の標準もほぼ同じブロック図を用い、変化点の技術を吹き出しで示すこととする。

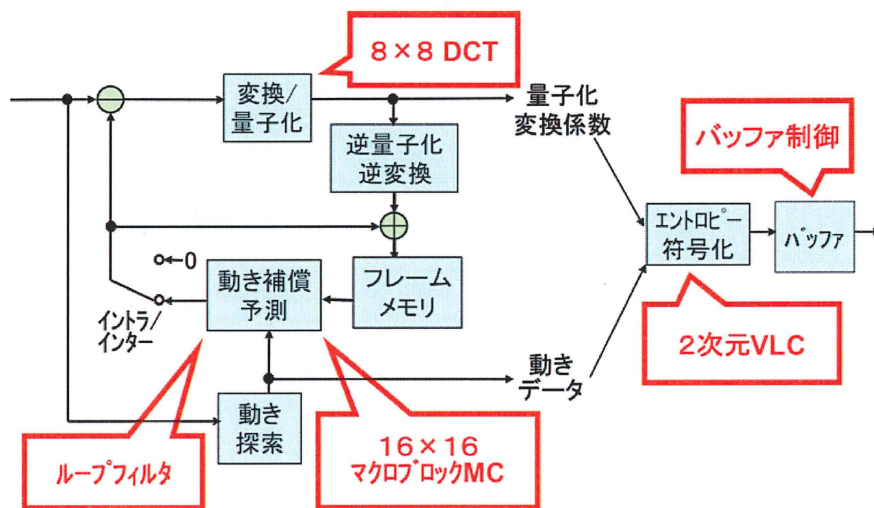


図 4-6 H.261 符号化器のブロック図

4-3-2 MPEG-1 から MPEG-2 へ

MPEG-1 (1992 年最終文書) は実質、H.261 に周期的なイントラフレームを含む双方向予測と半画素精度動き補償を加え、ループフィルタを除いた仕様である。双方向予測は MPEG-1 の想定用途が CD-ROM で遅延が許容されたため、周期的なイントラはランダムアクセスのために必要であった。半画素精度動き補償が LPF の効果を持つため、ループフィルタは割愛された (図 4-7)。

MPEG-2 (1994 年最終文書) はデジタル放送を見据え、汎用的な標準を目指したため、テレビジョン信号のインターレース構造が課題であった。

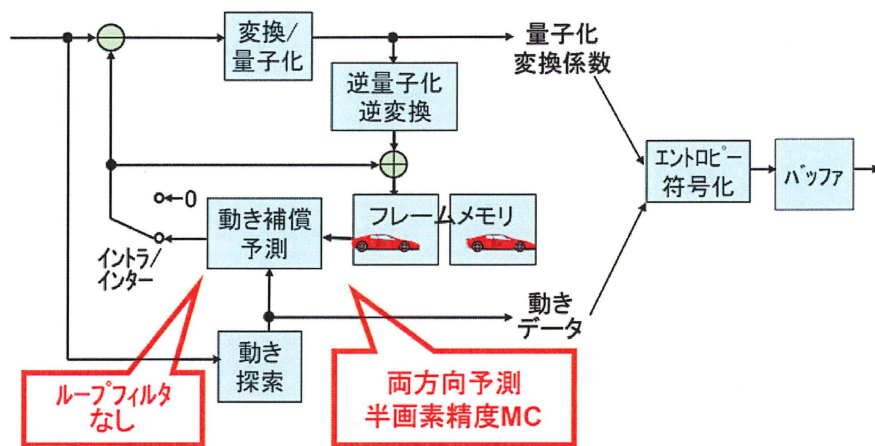


図 4-7 MPEG-1 符号化器のブロック図 (フレーム並べ替えを除く)

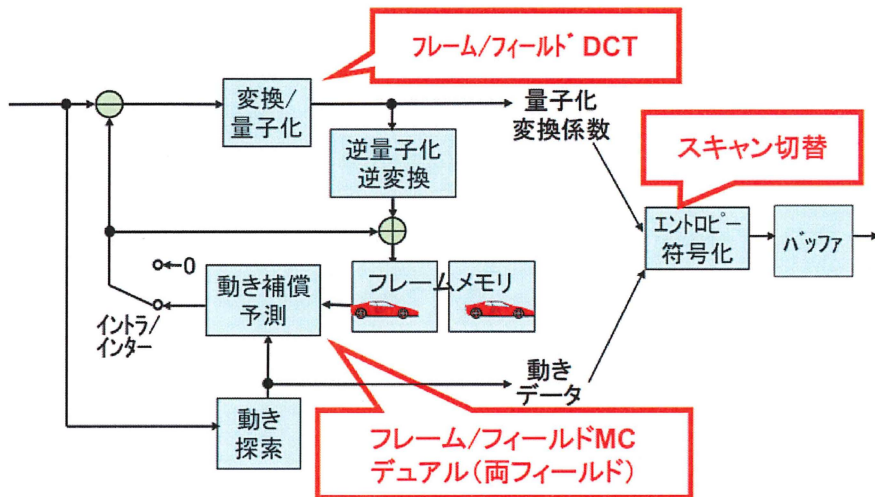


図 4-8 MPEG-2 符号化器のブロック図 (フレーム並べ替えを除く)

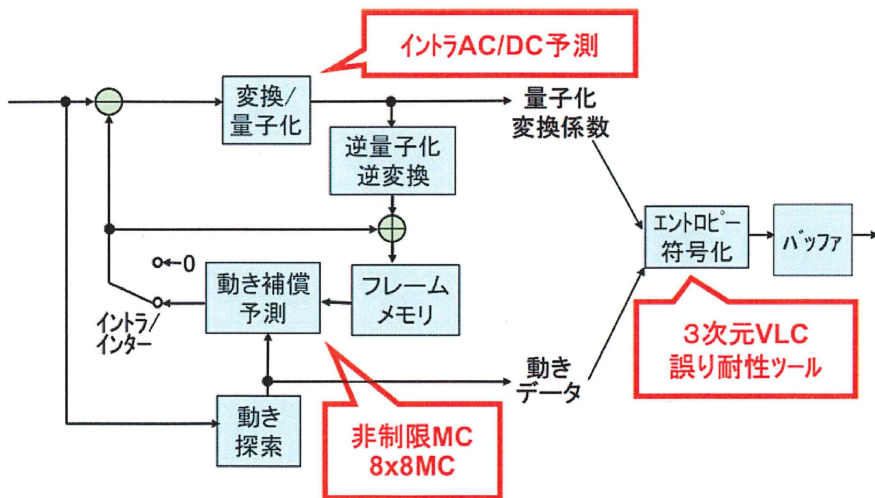


図 4-9 MPEG-4 符号化器のブロック図 (同標準 Visual Simple Profile)

筆者らはインターレースへの対応としてフレームまたはフィールド単位の予測、両フィールドの加算平均を用いる予測、DCTのためのブロックをフレームまたはフィールド単位で構成すること、DCT係数の走査順序切替を含む提案を行い、文献[4-19]に発表した。同じ着目点を持つ他の提案を含めた議論の結果、MPEG-2標準の技術は筆者らの提案の基本部分を取り入れた内容となっている。その後、MPEG-2は今日に至る放送やDVDに広く活用されている。

4-3-3 MPEG-4 から AVC | H.264 へ

MPEG-4（1999年第1版最終文書）はMPEG-2標準化の終盤から検討を開始して長い時間を要した標準である。その理由はやや発散した目標設定にあるが、第1目標は第3世代携帯を想定したモバイルであった。そのため、超低レート、低解像度、符号誤りへの耐性が主な課題となった。

筆者らは超低レート、低解像度の課題に対して動き補償ブロックの分割、短周期と長周期の予測に用いる2つのフレームメモリを含む提案を行い、文献[4-20]に発表した。動き補償ブロックの分割については低解像度画像を対象とする必然からMPEG-4の8×8単位動き補償に反映され、短周期/長周期のフレームメモリは陽に採用されず、基本部分が次のAVC | H.264に反映された。

AVC | H.264（2003年最終文書）はHDTVを含むテレビジョンのMPEG-2、モバイル低解像度のMPEG-4完成後の活動であり、当初はITU-Tの長期視点の活動であった。標準化を後押ししたのはインターネットのポテンシャルである。MPEG-4標準化から時間が短く、大きな改善方式の目算はなかった。符号量を10%削減できる提案が7つあれば圧縮率は2倍以上になる($0.9^7 = 0.478 \dots < 0.5$)とも言われ、小さな改善を捨てずに集積する方針となった。もちろん個々の改善は独立でなく、単純な加算的効果は期待できない。このためAVC | H.264は多くの改善の集積であり複雑度が高い標準となった。個々の改善方式をツールと呼ぶことも定着した。

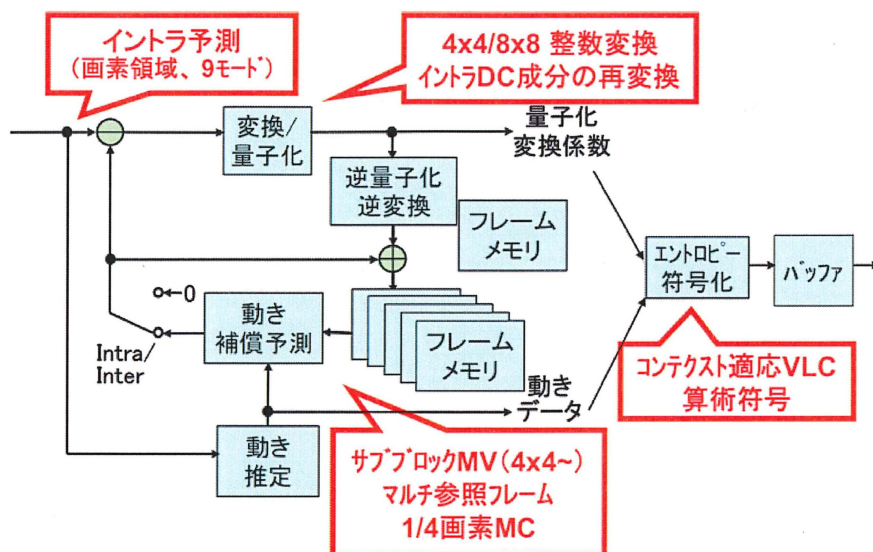


図 4-10 AVC | H.264 符号化器のブロック図

AVC | H.264 の特長の一つに、用途やビットレート条件を限定しないユニバーサル性がある。H.261 や MPEG-2 で採用された DCT 係数の 2 次元 VLC を低レート向けの H.263[4-11]や MPEG-4 で 3 次元 VLC としたのは低レート条件での確率分布を想定したからであり、ユニバーサルな符号化では条件を特定できない。ユニバーサル性を有するエントロピー符号化が必要とされ、適応的な VLC、算術符号が導入されることは必然であった。シンボルに指数的な語長を持つ符号語を割当て算術符号で圧縮するアプローチは、例えば筆者も第 3 章のウェーブレット packets の実験で用いており、1990 年代前半の研究動向 ([3-13,3-14]等) を踏襲したものである。

この点における筆者らの貢献は算術符号の誤り回復力の強化である。算術符号は動的適応性に優れるが誤りに弱い。そこでストリーム途中からのリカバリが可能ないように、適応状況を示す情報を周期的に (同期回復点に) 多重化することを提案[4-21]し、標準に採用されている。

4.3.4 符号化方式の複雑化と性能の向上

表 4-1 は MPEG 動画符号化標準の主な仕様を示している。サイズや精度の数値から、複雑度が向上していることがわかる。図 4-11 は動き補償予測方式に着目して、性能の向上をグラフ化したものである。性能が確実に向上していることがわかる。

表 4-1 動画符号化標準の比較（複雑度の増大）

	予測ブロックサイズ	動きベクトル精度	参照ピクチャ数	イントラ予測	変換ブロックサイズ	エントロピー符号化	ループフィルタ
MPEG-2	16x16	半画素	2以下	DC予測	8x8	2D-VLC	—
MPEG-4	8x8, 16x16	1/4画素 (8tap, mirror)	2以下	AC/DC予測	8x8	3D-VLC	—
AVC/H.264	4x4~16x16 & H/V shape	1/4画素 (6tap)	5(通常)	9モード空間予測	4x4, 8x8 整数変換	CABAC CAVLC	デブロッキングフィルタ
HEVC/H.265	4x4~64x64 & H/V shape & AMP	1/4画素 (8&7tap)	5(通常)	35モード空間予測	4x4~32x32 整数変換	CABAC	デブロッキング、画素適応オフセット

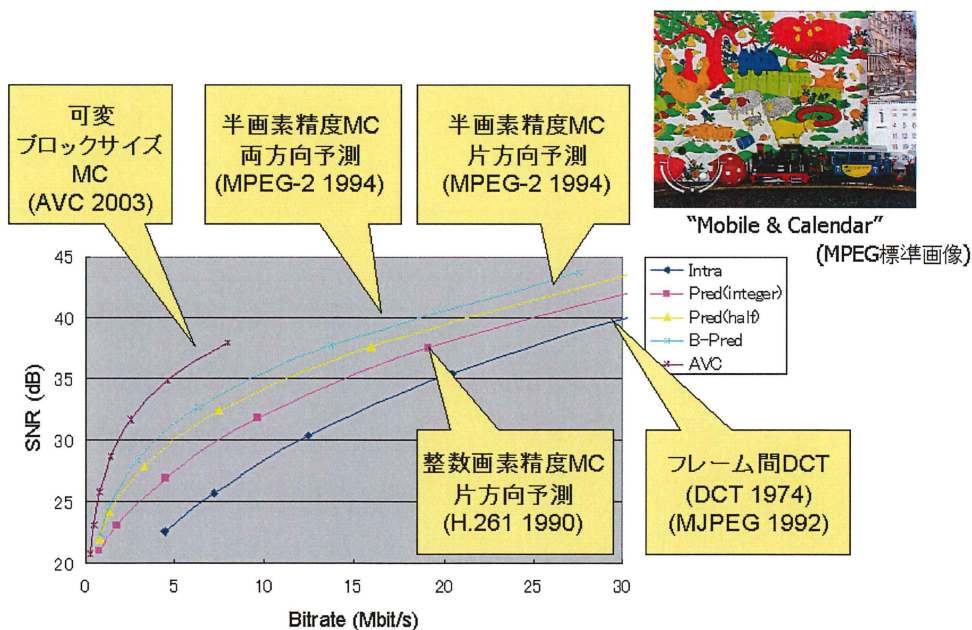


図 4-11 動き補償予測の高度化と性能の向上 [1-35]

4.4 HEVC|H.265 への提案

4.4.1 方式提案の前提

本節では HEVC|H.265 標準化における筆者らの提案方式について、論文化した文献[4-30]を用いて説明する。方式検討にあたり、以下の3つを前提とした。

[1] 既存の動き補償予測付き DCT の骨組みを踏襲する

既存国際標準との親和性や設計上の知見及び流用性を活かして改善を行う。互換性の確保は困難としても、共通要素を多く含む方式はデバイスやソフトウェア開発の視点から有利であり、産業的に期待されている。既存方式には長年の洗練を経て最適化されており、異なるアプローチでこれに対抗することは現実的ではない。

[2] HD/UHD コンテンツを符号化対象として AVC/H.264 の 2 倍の符号化性能を目標に設定する

既に HD クラスのカメラやディスプレイは市場に溢れ、4K に対応した機材が入手可能になっているため、HD/UHD の大容量コンテンツを対象にすることは必然である。

[3] 符号化性能と複雑度のトレードオフを考える。

高解像度映像は文字通り膨大な画素データを持つため、符号化・復号装置には大容量のメモリや高性能のプロセッサが必要となる。符号化方式の複雑度は方式を活用する時点で妥当なコストにより実現可能でなければならない。この点で AVC/H.264 に比べて 2~3 倍の複雑度を目標に設定する。複雑度の評価に符号化の計算機シミュレーションに要する時間を用いる。これについては 4.5 で議論する。

符号化性能を検証するためのテストシーケンスとして、標準化活動で採用されている中で HD (1920×1080) 以上の解像度のシーケンスを用いる。これらを図 4-12 に示す。(a)(b)は 2560×1600 画素×30 フレーム/秒、(c)(d)(e)(f)(g)は 1920×1080 画素で、(c)は 60 フレーム/秒、(d)(e)は 50 フレーム/秒、(f)(g)は 24 フレーム/秒である。

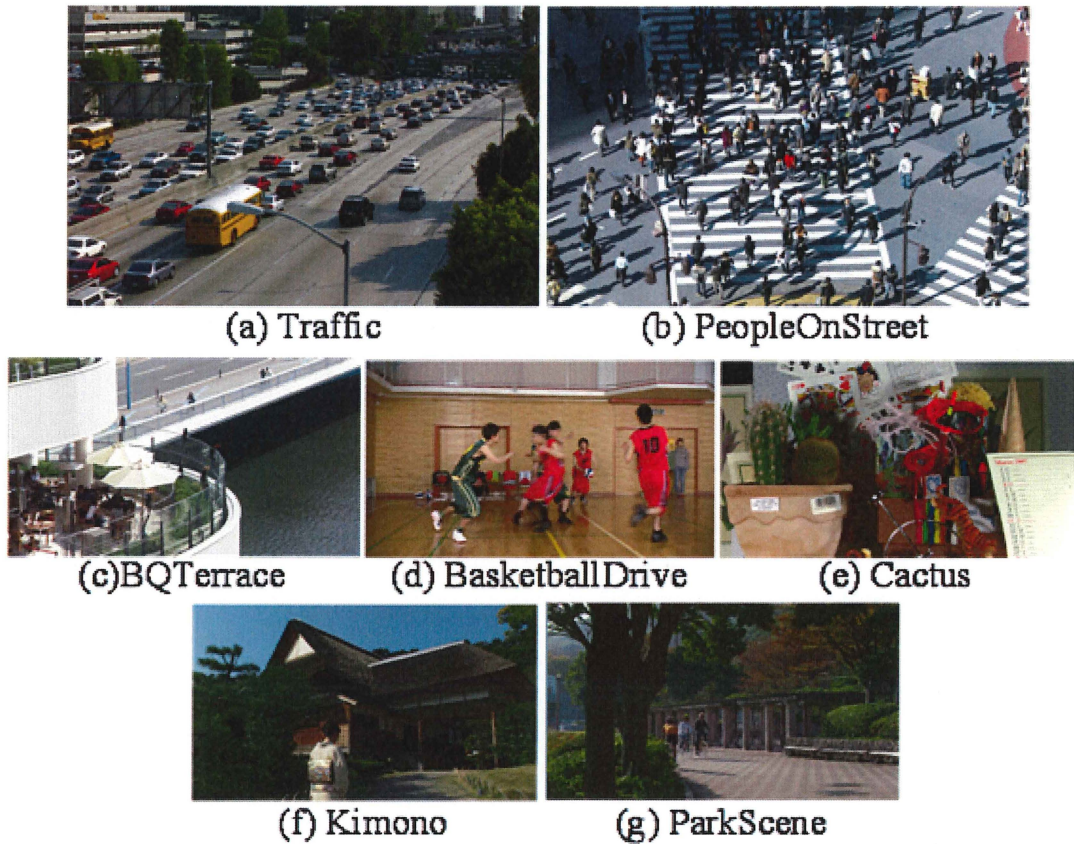


図 4-12 実験に用いたシーケンス

4.4.2 提案方式の概要

基本構造として動き補償予測付き DCT を用い、AVC/H.264 を改善前の土台と見做し、以下の工夫によって性能強化を行う。

- A) ブロックサイズの拡張と非矩形を含む階層的ブロック分割
- B) 動きベクトル予測器の適応的選択
- C) ブロックベースのピラミッド型フレーム内予測
- D) 複数のブロックサイズによる適応的な変換
- E) デブロッキングフィルタとウィナーフィルタの併用

A) ブロックサイズの拡張と非矩形を含む階層的ブロック分割

マクロブロックのサイズは H.261 以来、 16×16 画素である。動きベクトルなど、マクロブロック単位のデータは有意な符号量とな

るため、ブロックサイズの拡大は全体符号量の削減につながる。このためブロックサイズを拡大する試みが報告されている[4-23]。ここではサイズの拡大に加えて非矩形への分割を含む階層的な分割を行い、これらの分割された領域に対して動きベクトルを割り当てる。

階層的な分割の構造を図 4-13 に示す。最大のマクロブロック階層を 64×64 とし、分割による最小サイズの 32×32 を L0 階層としてさらに分割する。以下同様に、L1 階層は 16×16 、L2 階層は 8×8 と定義する。図の範囲で動きベクトルは最大 64×64 のマクロブロックから最小 8×8 のブロックまで割り当てられる。ブロックサイズは被写体の大きさや動きの局所的な相違によって選択する。

さらに図 4-14 に示すように、ブロックを対角方向に分割するモードを設定する。この時、ブロックは矩形と非矩形の「L 型」に分割される。最小のブロックサイズは 8×8 となる。非矩形の分割は動く被写体の境界部分で有効である。

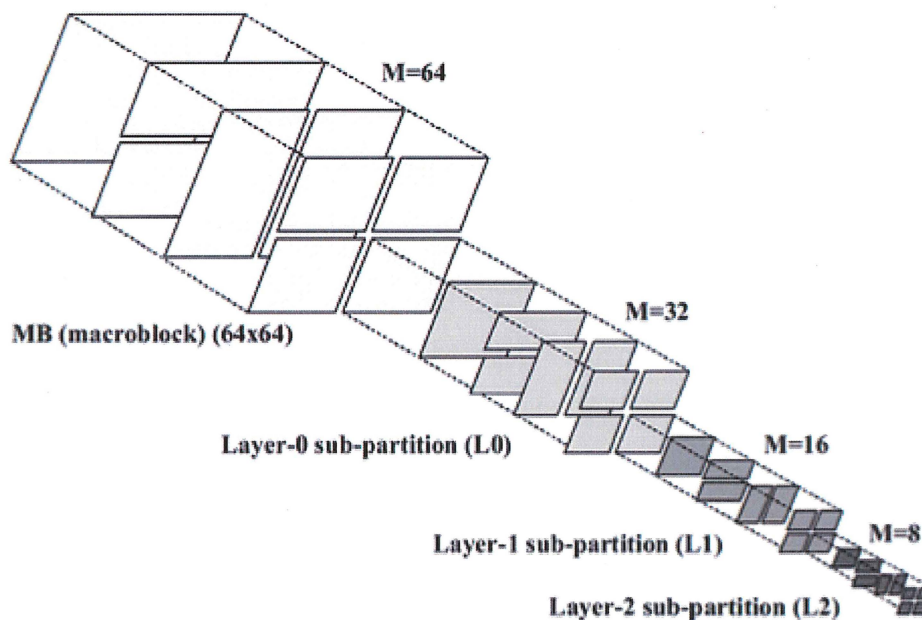


図 4-13 階層的なブロックの分割（動きベクトルの割当単位）

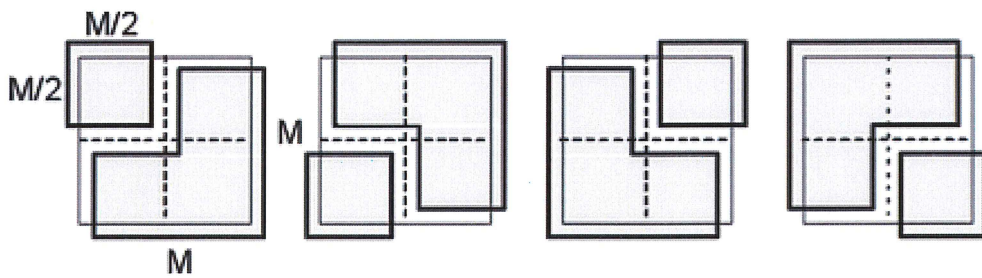


図 4-14 非矩形を含むブロックの分割

2560×1600 画素の画像 Traffic に対する階層的ブロック分割の導入効果を輝度成分の符号量-歪み特性で図 4-15 に示す。ビットレートにして平均 10%以上の符号量削減効果が得られている。

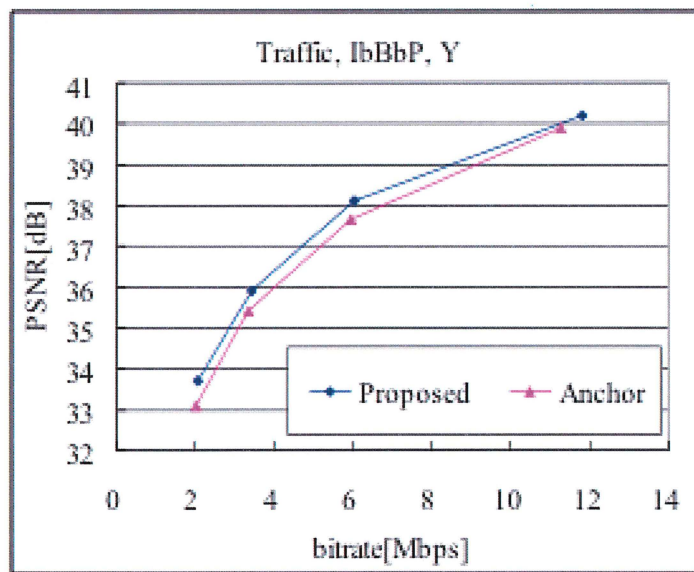


図 4-15 画像 Traffic に対する階層的ブロック分割の導入効果

B) 動きベクトル予測器の適応的選択

動きベクトルを割り当てる領域を適応的に選択することで動きベクトルの数を適応化したことに続けて、動きベクトルの符号量を低減させることを考える。動いている被写体が、動き補償予測を行う時間の中ではほぼ剛体と近似可能なら、動きベクトルは近傍の領域

でもほぼ同じ値を持つことが期待される。また被写体が一定速度で動いていると近似可能なら、動きベクトルは微小時間前の（符号化済みの）近傍領域とほぼ同じ値を持つことが期待される。こうした想定により、AVC/H.264には動きベクトルの適応的な予測器が導入されている。

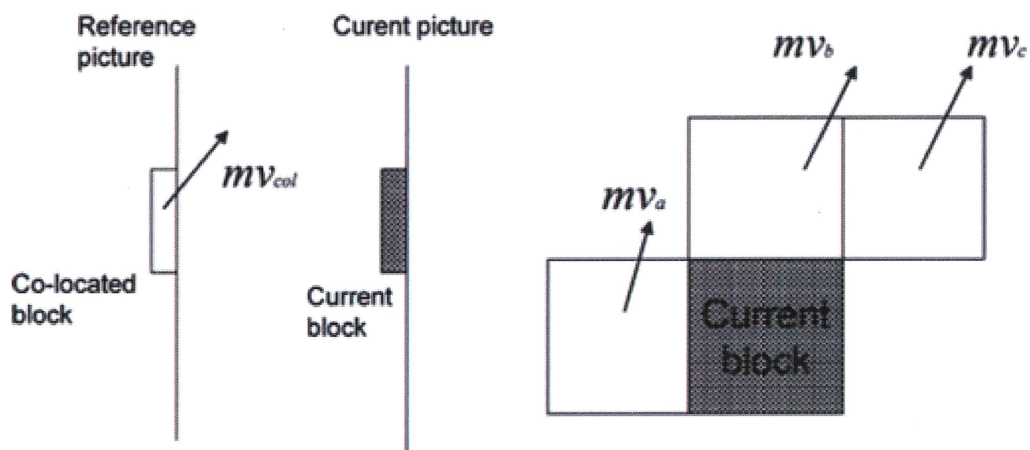


図4-16 動きベクトルに対する2種類の予測

図4-16は動きベクトルに対する2種類の予測を示している。網がけのなされた部分は動きベクトルを予測したい符号化対象のブロックを表す。図の左側は、動き補償予測に参照するピクチャの同一位置における動きベクトルを予測に用いる時間的予測である。また、図の右側は同一ピクチャにおいて既に符号化済みである左側のブロック、上側のブロック、右上のブロックの動きベクトルを参照し、これらのメディアンを予測に用いる空間的予測である。

AVC/H.264では動きベクトルの符号量が特に大きい双方向予測を行うBピクチャにおいてダイレクトモードと呼ばれる動きベクトルの予測にこれら2種類の予測を用いる。符号化の段階で2種類の予測から一方を選択し、選択結果の情報を符号化して復号側に送る。この1ビット情報をさらに削減することを考える。

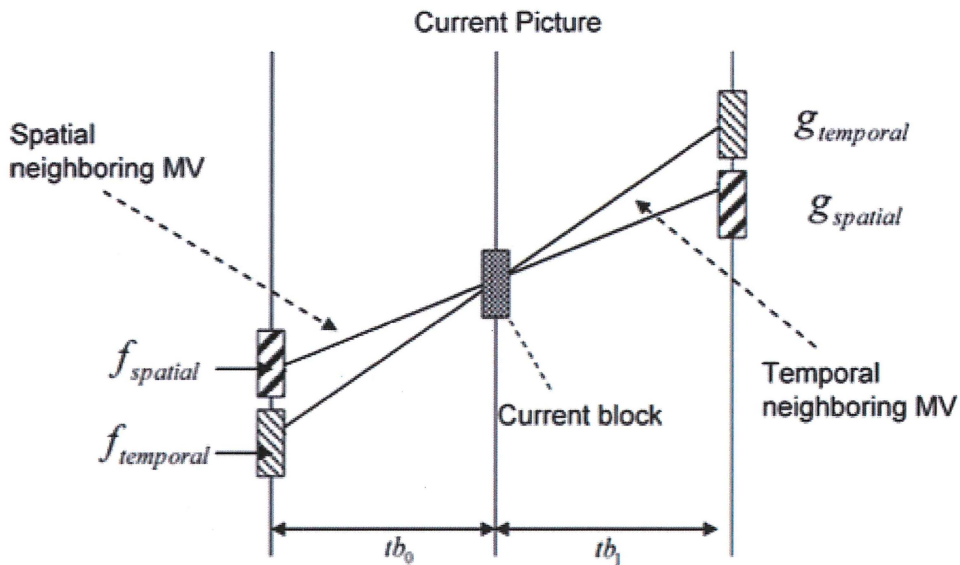


図4-17 ダイレクトモードにおける動きベクトルの候補と予測ブロック

図4-17において、水平方向は時間方向とし、中央の縦線は符号化すべきBピクチャを表すものとする。左右の縦線は既に符号化が終了した第1および第2の参照ピクチャを表す。今、符号化すべきピクチャの符号化すべきブロックを基準にして、時間的予測で与えられる動きベクトルを仮定して第1の参照ピクチャから得られる予測ブロックを $f_{temporal}$ 、第2の参照ピクチャから得られる予測ブロックを $g_{temporal}$ とする。また同様に、空間的予測で与えられる動きベクトルを仮定して第1の参照ピクチャから得られる予測ブロックを $f_{spatial}$ 、第2の参照ピクチャから得られる予測ブロックを $g_{spatial}$ とする。この時、第1および第2の参照ピクチャにおいて時間的予測と空間的予測の精度を評価する2つの値をそれぞれの予測によって得られる予測ブロックの間の差分絶対値和として以下のように求める。

$$SAD_{temporal} = |f_{temporal} - g_{temporal}|$$

$$SAD_{spatial} = |f_{spatial} - g_{spatial}|$$

ここで差分絶対値和の小さい予測は精度の高い予測が可能であると推定する。すなわち、 $SAD_{temporal}$ が小さい場合には時間的予測、 $SAD_{spatial}$ が小さい場合には空間的予測を選択する。 $SAD_{temporal}$ と $SAD_{spatial}$ の計算に必要な情報は符号化側と復号側の両方に共通して得られるため、このような推定を両方で行うことにすれば、選択結果の情報を送らずにすませることが可能である。この効果により、1.5%の符号量削減効果が得られたと報告されている[4-24]。

C) ブロックベースのピラミッド型イントラ予測

国際標準では動き補償を用いたフレーム間予測に改善を加えて符号化効率を向上させているが、時間方向の予測を使わないフレーム内符号化の効率向上が課題である。MPEG-2やMPEG-4 Visualなどの標準では周波数領域すなわちDCT係数の領域で予測を行なっている。一方、AVC/H.264ではフレーム内符号化を行う際に空間領域すなわち画素領域での予測を行って符号量の削減を図っている。これに対する改善を考える。

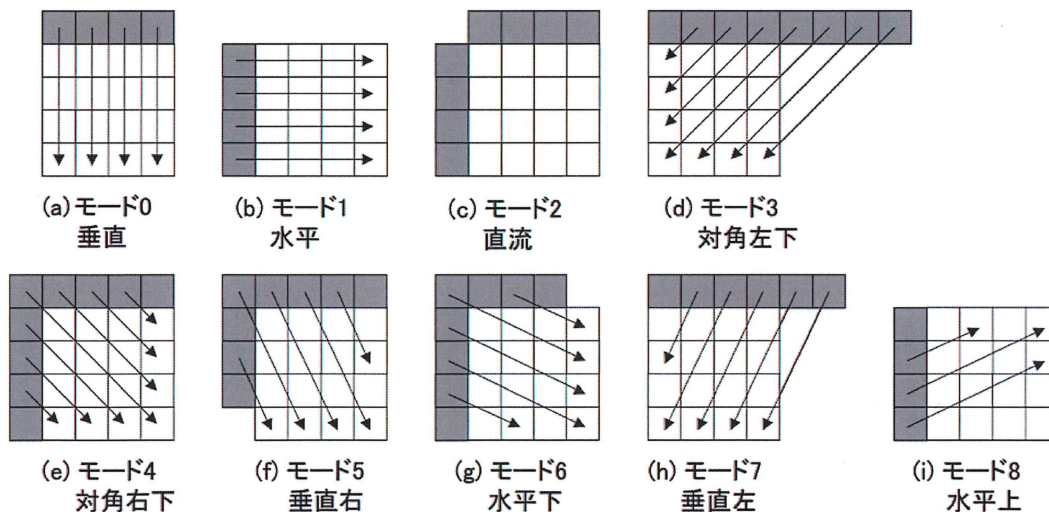


図4-18 MPEG-4 AVC|H.264のイントラ予測

AVC/H.264 の画素領域の予測はブロックサイズによって異なる

が、特徴的なものは 4×4 画素ブロックに適用されるイントラ 4×4 予測である。同方式は図 4-18 に示す 9 つのモードのいずれかをブロック単位に選択するものであり、復号処理を終えた周囲のブロック（左上、上、右上、左）の画素を予測に用いる。

推測されるように、図中の矢印で示される予測の方向に強い相関のあるエッジが存在する場合、予測誤差は小さくなり、同方式は優れた特性を示す。しかしどの方向にも一致しない方向がある場合、同方式の予測は選ばれた予測の方向に予測誤差を生成してしまう効果を持つ。この課題を解決するために、ピラミッド型のイントラ予測を導入する。ピラミッド型の予測とは、画素を上位の階層（画素数が少ない）から下位の階層（画素数が多い）に連なる多重解像度のピラミッドで表現し、上位から下位を予測することである。符号化器側のブロック図を以下に示す。

符号化器側では入力の 16×16 画素ブロック B_0 を 4×4 画素ブロック単位ごとに平均値処理することによって 4×4 にダウンサンプルされたブロック B_s を出力する。一方、平均値予測では既に符号化の終了した隣接ブロックから予測値 d を算出する。減算器ではダウンサンプルされたブロック B_s から d を減算して残差のブロック

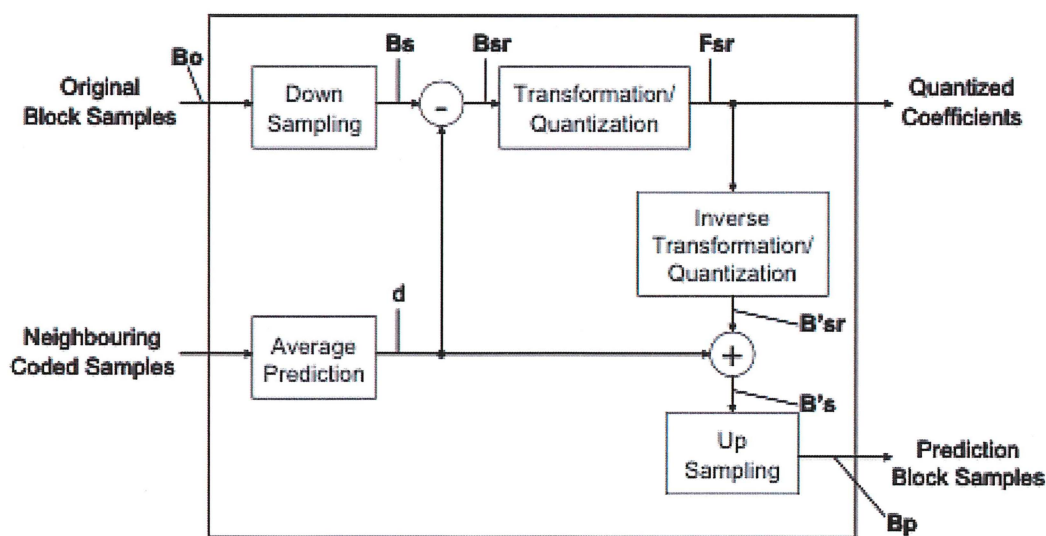


図 4-19 符号化側におけるピラミッド型イントラ予測のブロック図

B_{sr} を得る。

$$B_{sr}(i, j) = B_s(i, j) - d, \quad (i, j) \in B_s$$

残差となったブロック B_{sr} に対して AVC/H.264 の符号化器が備える変換と量子化を施す。量子化された変換係数 F_{sr} は符号化情報として出力されるとともに量子化代表値に対する逆変換処理によって復号された残差ブロック B'_{sr} となる。加算器では予測値 d を復号された残差ブロック B'_{sr} に加算してブロック B'_s を得る。

$$B'_s(i, j) = B'_{sr}(i, j) + d, \quad (i, j) \in B'_{sr}$$

B'_s は B_s に量子化誤差が加算されたブロックであり、これを AVC/H.264 の符号化器が備えている 1/4 画素精度動き補償のサンプルを生成するためのアップサンプルフィルタを用いてアップサンプルし、 16×16 画素の予測ブロック B_p を得る。

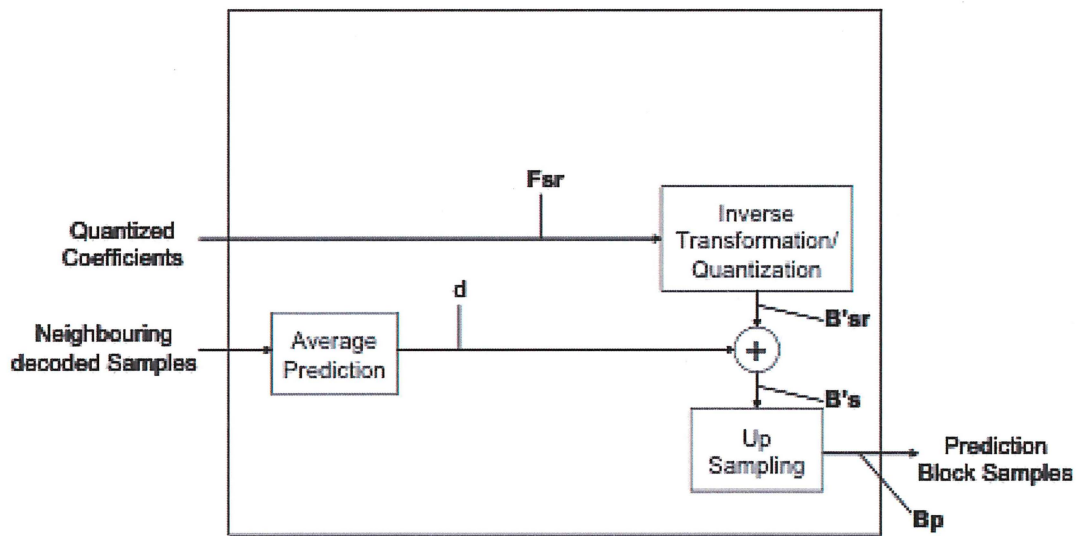


図 4-20 復号側におけるピラミッド型イントラ予測のブロック図

図 4-20 はピラミッド型イントラ予測を実現する復号器側のブロック図である。量子化された変換係数 F_{sr} が与えられ、量子化代表値と逆変換によって B'_{sr} が求められる。一方、既に復号された隣接ブロックから予測値 d を算出する。加算器では予測値 d を復号され

た残差ブロック $B'sr$ に加算してブロック $B's$ を得る。 $B's$ を AVC/H.264 の復号器が備えている 1/4 画素精度動き補償のためのアップサンプルフィルタを用いてアップサンプルし、 16×16 画素の予測ブロック Bp を得る。この予測ブロックは符号化器側の Bp と等しい。

ピラミッド型のイントラ予測は低解像度から高解像度への予測であり、方向性のあるエッジを予測するのに適した AVC/H.264 の予測器に対して補完的に作用し、画像の中の高周波成分が弱い領域、輝度が緩やかに変化する領域に有効である。

AVC/H.264 にピラミッド型イントラ予測を追加し、本予測の効果を検証した。高解像度 (2560×1600 及び 1920×1080) のテストシーケンスに対する結果は表 4-2, 4-3 の通りである。

表 4-2 ピラミッド型イントラ予測の改善効果(QP=50, 46, 42, 38)

Sequence	BD-rate (%)	BD-PSNR (dB)
1920x1080	-5.638	0.255
Kimono	-9.419	0.470
ParkScene	-5.316	0.176
Cactus	-4.022	0.180
BasketballDrive	-6.323	0.323
BQTerrace	-3.109	0.126
2560x1600	-3.032	0.167
Traffic	-4.341	0.229
PeopleOnStreet	-1.722	0.105

表 4-3 ピラミッド型イントラ予測の改善効果 (QP=34, 30, 26, 22)

Sequence	BD-rate (%)	BD-PSNR (dB)
1920x1080	-0.050	0.002
Kimono	-0.169	0.006

ParkScene	-0.032	0.002
Cactus	-0.004	0.000
BasketballDrive	-0.004	0.000
BQTerrace	-0.041	0.002
2560x1600	-0.021	0.001
Traffic	-0.004	0.000
PeopleOnStreet	-0.037	0.002

表はピラミッド型イントラ予測により、QP が高い場合すなわち高圧縮の場合に、2560×1600 画素のシーケンスには平均-3.03%、1920×1080 画素のシーケンスには平均-5.64%の符号量削減効果があることを示している。このとき、符号量削減効果は最小で-1.72%、最大で-9.42%に及び、全てのシーケンスで改善が得られていることがわかる。一方、QP が低い場合すなわち低圧縮の場合には改善効果がわずかである。ただし全てのシーケンスにおいて改善の方向であることがわかる。

図 4-21 はシーケンス Kimono に関する符号量-歪み特性を示している。ビットレートの低い条件で改善効果が高いことが見て取れる。

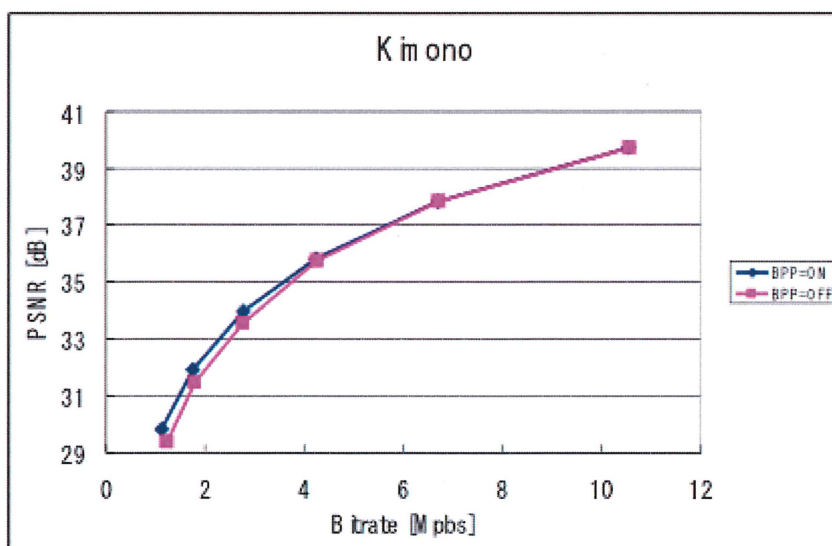


図 4-21 ピラミッド型イントラ予測 (BPP) による改善効果

D) 複数のブロックサイズによる適応的な変換

動き補償予測による予測誤差をサイズが適応的に変化するブロックにより変換する。AVC/H.264 では 4×4 画素または 8×8 画素の変換を用いたが、ここでは高精細画像に対する性能を考慮して 16×16 画素のブロックを加え、3種類の DCT を用いる (4×4 , 8×8 , 16×16 DCT)。

ブロックサイズの選択を示すフラグが必要となるが、オーバーヘッドの増大を避けるため、動き補償予測のために選択したブロック分割のサイズによって以下の制約条件を設け、1 ビットのフラグでサイズを示すようにする。

- ・ ブロック分割が 32×32 より大きい場合、DCT は 16×16 のみ
- ・ ブロック分割が 32×32 以下かつ 16×16 より大きい場合、DCT は 8×8 または 16×16
- ・ ブロック分割が 16×16 以下の場合、DCT は 4×4 または 8×8

これらの制約条件により、AVC/H.264 に対してフラグの符号量を増やすことなく、適応的なブロックサイズを選択することが可能になる。

E) デブロッキングフィルタとウィナーフィルタの併用

i. デブロッキングフィルタ

DCT を用いる符号化では符号量不足によるブロック歪みの発生が不可避である。AVC/H.264 ではブロック境界の不連続を平坦化するデブロッキングフィルタを導入している。ブロック境界の平坦化はテクスチャの連続性を損ねる可能性もあるため、フィルタの強度を適応的に調整する。調整の基準はマクロブロック外縁のブロック境界であるか、マクロブロック内部のブロック境界であるか、また動き補償予測の参照画像は隣接ブロックと同一であるか、動きベクトルは同等か、などの条件であり、有効に機能することが知られている。

ここでは基本的に AVC/H.264 のデブロッキングフィルタを用い、ブロックサイズを適応化して 16×16 画素の DCT を導入したことによる改訂程度とする。

ii. 適応ウィナーフィルタ

適応ウィナーフィルタは文献[4-25]によって提案された手法である。もとよりウィナーフィルタは画像復元の分野において、画質劣化プロセスの逆問題を解くアプローチとして研究されてきた手法であり、画質改善効果が高いと同時に複雑度が高い。ここでは画質改善効果を重視して採用する。符号化器側で局部復号された画像と原画像の情報を参照してウィナーフィルタの係数を設計し、スイッチング情報とともに符号化情報に多重化して送出する。復号器側では同情報を参照してウィナーフィルタを実現し、符号化ループの整合を確保する。

4.4.3 符号化性能

4.4.2 で説明した提案方式について、図 4-12 のシーケンスを用い、計算機シミュレーションによって符号化性能を調べ、AVC|H.264 との比較評価を行なった。比較対象には AVC|H.264 の参照ソフトウェア JM16.2 を用いている。比較結果を表 4-4 に示す。表 4-4 において提案方式は同一の SN 相当で 17%-33%の符号量削減（平均 25.7%削減）を実現している。

性能への貢献度は厳密に分解できないが、個々の要素の検証状況から、動き補償ブロックの拡大、変換ブロックの適応化、ループフィルタがそれぞれ 8%の改善、動きベクトル選択の適応化で 2%の改善と推定される。演算時間は AVC|H.264 の 2 倍程度である。

図 4-22 および図 4-23 に、輝度成分に対する符号量-歪み曲線を示す。提案方式の符号量-歪み曲線はいずれも AVC|H.264 の曲線の上方向にあり、ビットレートの広い範囲で提案方式が AVC|H.264 より優れていることを示している。

表 4-4 提案方式による符号化性能 (AVC|H.264 との比較)

sequence	size	frame rate	no. of frames	BD-rate [%]	BD-PSNR [dB]
Traffic	2560x1600	30	150	-24.8	1.07
PeopleOnStreet	2560x1600	30	150	-16.7	0.94
BQTerrace	1920x1080	60	600	-33.2	0.70
BasketballDrive	1920x1080	50	500	-27.7	1.05
Cactus	1920x1080	50	500	-24.1	0.84
Kimono	1920x1080	24	240	-32.2	1.40
ParkScene	1920x1080	24	240	-21.0	0.85
Average	-	-	-	-25.7	0.98

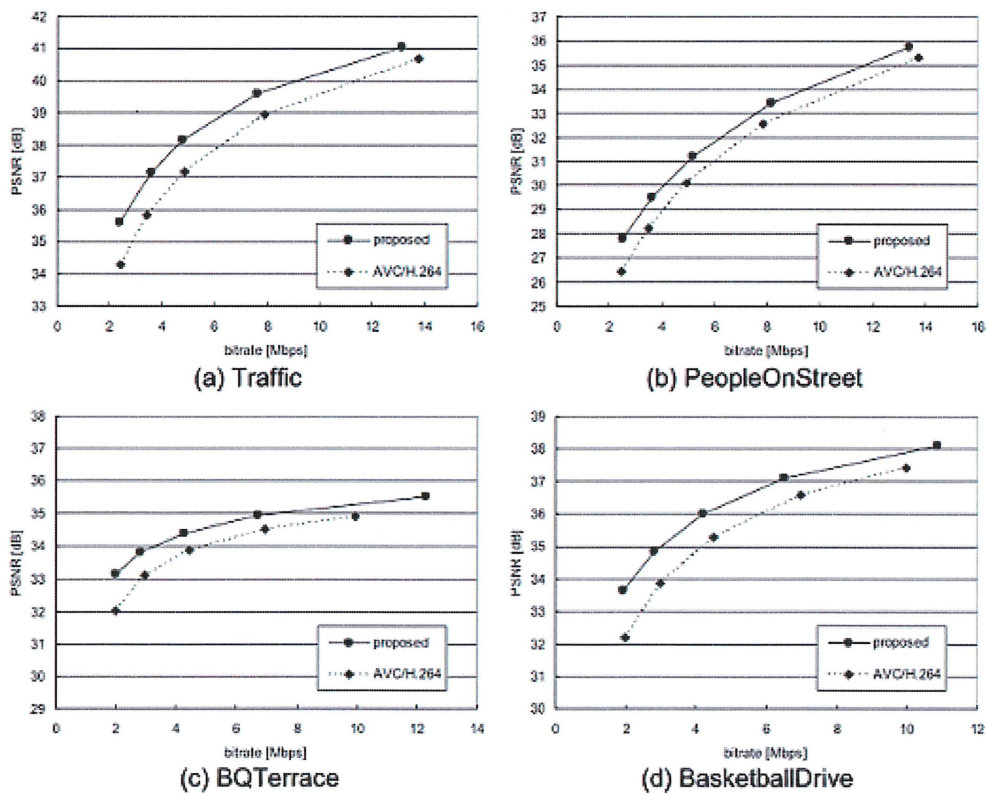


図 4-22 提案方式による符号化特性 (1/2) [4-30]

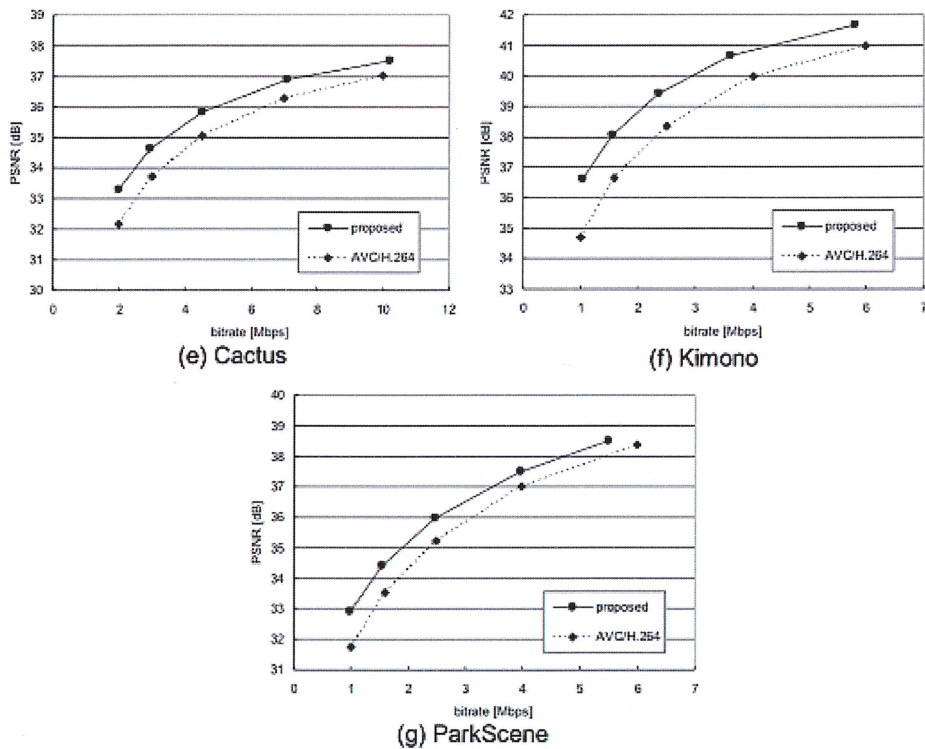


図 4-23 提案方式による符号化特性 (2/2) [4-30]

4.5 符号化性能と複雑度の関係

4.5.1 符号化性能と複雑度

本節では前節で述べた HEVC | H.265 への提案方式を基にして符号化性能と複雑度との関係について考える。まず、同標準の提案募集に応じた提案と筆者らの提案方式を性能-複雑度平面にプロットすることを試みた結果が図 4-24 である。符号化性能については AVC | H.264 と比べた符号量削減効果を用い、複雑度については符号化に要する演算時間を用いた。いずれも提案書類に報告されたものである。性能の数値は 2560×1600 画素のシーケンスに対する値を用いた。

複雑度として符号化の演算時間を用いたのは以下の理由による。

- ・ 符号化処理の演算時間は当該アルゴリズムが演算に要する能力とメモリを近似的に表すと考えられる
- ・ 符号化処理の演算時間は当該アルゴリズムが持つ適応的な符号

化モードなどの選択肢の多さを反映すると考えられる

- ・ 符号化処理の演算は通常、復号に要する演算よりかなり大きい
が、今後ユーザにとって重要度が増す要素である

一方では符号化処理の演算時間はソフトウェア実装や計算機環境、最適化演算の程度に依存する。図にプロットされた数値はそれぞれ異なる環境で得られたものであり、正規化された数値ではない。

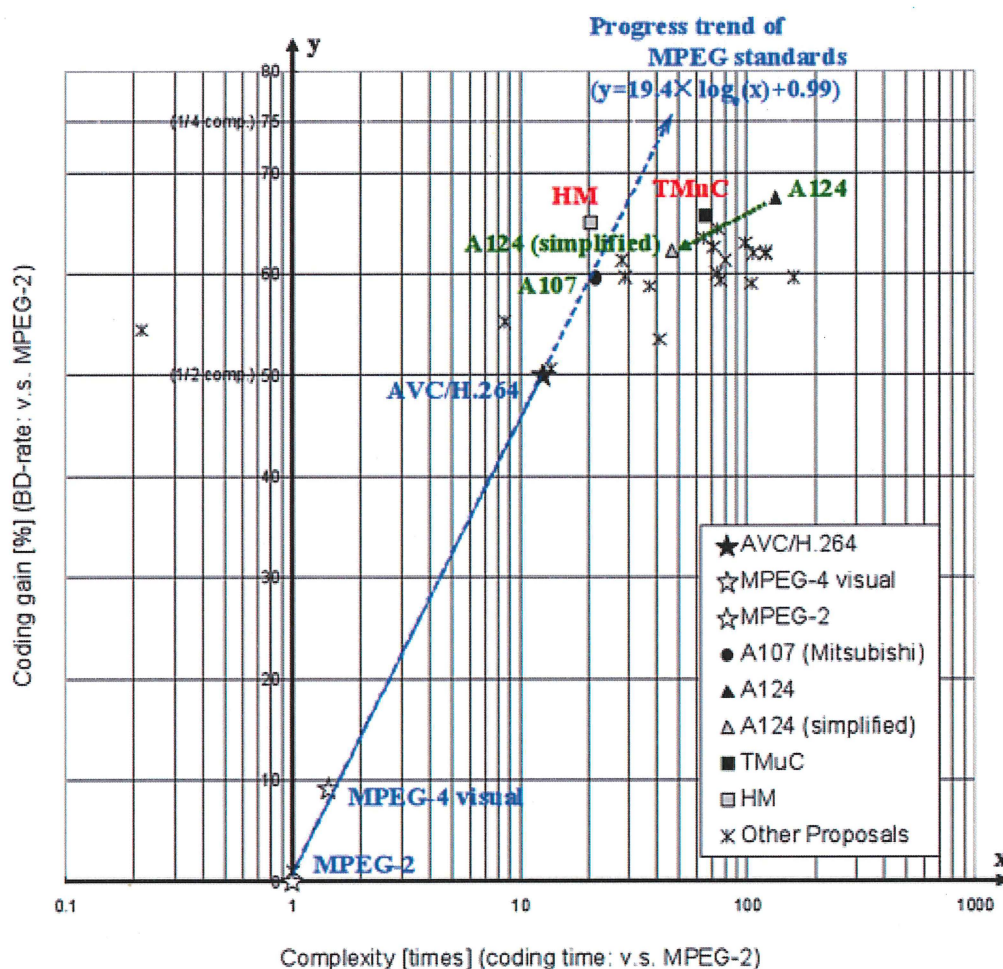


図 4-24 既存国際標準から参照した符号化性能と複雑度の評価

従って同図を持って個々の提案方式を評価することは適当でない。筆者らの目的は複数の提案方式の傾向やばらつきを観察することにある。図 4-24 には参照点として MPEG-2、MPEG-4 Visual、

AVC|H.264 をプロットしている。MPEG-2 と MPEG-4 Visual は AVC|H.264 の参照ソフトウェア JM16.2 の設定によって模擬したものである。表が示すように、後年の標準ほどパラメータの選択肢が増える方向にあり、適応化の範囲はしばしば新しい標準が古い標準を包含するように拡張される。ただし厳密な包含ではないので、あくまで複雑度の参考値として使うためのエミュレーションである。

AVC|H.264 の符号化性能は MPEG-2 に比べて概ね性能が 2 倍、演算時間が 10 倍である。MPEG-4 Visual は両者を結ぶ実線の線分上にある。この線分を AVC|H.264 の先に破線で延長して参照線とする。HEVC 標準の目標は AVC|H.264 の 2 倍の性能であり、複雑度については標準化完了時の市場で現実的であること、というものである。ムーアの法則によるデバイス技術の進歩を考慮しても、HEVC 標準の複雑度は最終的に AVC|H.264 の 4~5 倍とすることが妥当かもしれない。このことは最初からその範囲で提案方式を考えるべきだということの意味しない。ムーアの法則に依存する以外に、アルゴリズムの理解や実装、適応化の洗練が進むにつれて性能は向上し、複雑度は下がるからである。複雑度を恐れるべきではない。必要なのは、複雑度の参照基準と複雑度を推定する手段を確かに持つことである。

4.5.2 符号化性能と複雑度のトレードオフ改善

図 4-24 において、筆者らの提案方式は A107[4-26]としてプロットされている。A107 は AVC|H.264 に比べて 20%の符号化利得と 2.5 倍の符号化演算時間である。A107 のソフトウェア実装は JM を基本とした改造である。

A124[4-27]は最も優れた性能値を示した提案の一つであり、符号化性能の改善は 35%、符号化の演算時間が 11 倍であった。図 4-24 にはもう一つ、[A124(simplified)]の点をプロットしてある。この点は A124 のソフトウェアを用い、同提案の中の新しい要素の中で多くの提案に共通する要素以外を停止して実行した結果である。この

状態でも、A124(simplified)は 25%の性能改善と 37%の複雑度増を示している。つまり A124 提案から一部を除いたサブセットであってもトップクラスの性能を保ち、性能と複雑度のバランスを改善することが可能である。

HEVC の標準化においてまず、提案の中から有効と考えられる要素を集めた参照モデル TMuC (Test Model under Consideration) [4-28]が定義された。図 4-24 において TMuC は 32%の性能改善と 5.2 倍の演算時間としてプロットされている。さらに実験と議論を重ね、正式な標準開発のための共通テストモデル HM1.0[4-29]が定義された。HM1.0 は TMuC に比べてわずか 4%の性能劣化を代償として、35%の演算時間減を実現している。AVC|H.264 に対しては 30%の性能改善と 1.6 倍の演算時間増である。

我々が図 4-24 のような参照基準と参照手段を持つと仮定する。図中の x 軸と y 軸を用い、ある要素を動作停止させることで生じる変化 $\Delta y/\Delta(\log x)$ を評価して傾きの小さい要素を取り除くようにすれば、バランスを改善できる。これは符号量-歪み特性の最適化においてしばしば行われる最適刈込みの考え方と同じである。

図 4-24 は動き補償予測付き DCT という基本的枠組みを前提とした範囲で連続性を仮定できると考えられる。感覚的ではあるが、同図からはもう少し直線の先に標準を策定する可能性もあるように思われ、実際に検討が行われている。同じ条件で進むなら、まず複雑度を恐れることなく性能改善の材料を集めることが合理的であるとも言える。一方では逆方向に、 $\Delta y/\Delta(\log x)$ の小さな提案はもとより採用すべきでない判断できるはずである。どのようなアプローチを採用し、方式開発のためのリソース投入をどこまで合理的に考えるかは提案者と標準化の推進リーダーに委ねられるべきだろう。本研究の主張は、複雑性を恐れず、しかし複雑度を参照する基準とその手段を精度よく共有することが必要だということである。

4.6 この章のまとめ

本章では基本的枠組みに動き補償予測付き DCT を採用した動画像符号化の国際標準方式について、歴史と筆者らの貢献について示した。標準方式はデジタル映像が必要とされた歴史の転換、すなわち放送のデジタル化、第 3 世代以降の携帯電話の世界的普及、インターネット利用の爆発的増大、ストリーミングメディアの勃興などと相互作用を持ち、ムーアの法則に象徴される基盤技術の進展に支えられて複雑度を増し、符号化性能を向上させてきた。筆者と共同研究者はそれぞれの段階で課題とされた点について提案と貢献を行ってきた。本章で示された知見は以下の通りである。

- ・ 動画像の国際標準は H.261 以来、動き補償予測と DCT のハイブリッド符号化を基本としており、予測・変換・エントロピー符号化の適応化を進めて圧縮効果を高めてきた
- ・ H.261 の仕様には 1970～1980 年代の研究成果が盛り込まれており、符号量の変動に対抗して相互接続の安定動作を保証するためにバッファ制御が規定された（筆者らの発明を含む）
- ・ MPEG-2 はテレビジョン信号のデジタル化と圧縮を可能にした標準であり、特にインターレース構造による特性を生かした要素が性能向上に貢献した（筆者らの発明を含む）
- ・ MPEG-4 は低レート低解像度向けに動き補償予測ブロックの細分化が導入され、3 次元 VLC が導入された（ブロック細分化は筆者らの提案と共通、長周期予測の提案は AVC | H.264 に反映）
- ・ AVC | H.264 では数多くのツールが導入され、ユニバーサル性を重視したため、算術符号が本格的に導入された（算術符号の誤り回復力を強化する筆者らの提案が採用された）
- ・ HEVC | H.265 向け提案では 4K/8K 映像を想定したマクロブロックサイズの拡大が効果的であることを確認した（拡大かつ階層的なマクロブロックサイズが標準に採用された）
- ・ 標準方式の策定には、性能と複雑度の関係について基準と評価手段を共有し、複雑度を恐れずに進むべきである

第5章 まとめと今後の課題

5.1 画像符号化研究の成果とその後

本節では第2章から第4章までの研究成果とその後を2017年の視点で振り返る。

第2章のベクトル量子化に関する議論は1980年代初頭から1990年頃までの研究成果である。筆者らは国際標準が $n \times 64\text{ kbit/s}$ ($n=1 \sim 2$) と $m \times 384\text{ kbit/s}$ ($m=1 \sim 5$) とに分かれて議論されていた頃、64kbit/sでの伝送を想定してベクトル量子化方式を開発し、同方式によるテレビ会議システムの製品展開を進めていた。標準化前のDCTを用いたシステムも市場で競合していた。ベクトル量子化には絵柄の変化に対する強さ、誤差がブロック内でランダム化されて目立ち難いことなど低レート符号化での優位性があった。一方で適用ビットレートの拡大や圧縮率をさらに向上させるために必要な次元数やコードブックの拡大に対し、当時の演算規模やメモリ制約からすぐに対応することが難しかった。ベクトル量子化の画像符号化に関する研究が当時はほぼ1980年代の歴史しか持っていなかったためでもあるだろう。

一方、1970年代から研究が続けられていたDCTの活用は技術的洗練を進めており、同じ枠組みで対応可能なビットレートの範囲も広がった。H.261は最終的に適用範囲を拡大して $p \times 64\text{ kbit/s}$ ($p=1 \sim 30$) 向けとなった。

製品展開や標準化以外のベクトル量子化研究については、汎用性と符号化性能の向上、コードブックの構造化等が研究された。例えば信号空間の変換が性能向上のアプローチの一つであり(平均値分離正規化もその一つ)、DCT係数やウェーブレット変換のベクトル量子化も検討された[2-41,5-1]。近傍ブロック間の相関を利用するアプローチとして(筆者らの平均値や振幅のDPCM符号化もその一つ)符号化済みブロックの画素を参照するサイドマッチ型の有限状態ベクトル量子化、また可変長符号化との組合せでコードブックの自己組織化を行う手法も検討された[5-2,5-3,5-4]。符号量-歪み特性の最

適化ではエントロピー拘束型ベクトル量子化や木探索コードブックの最適刈り込みが報告された[2-20,2-21]。木探索ベクトル量子化のコードブック設計に主成分分析の活用も提案された[2-44]。

コードブックの設計と蓄積における負荷を削減するために構造化を持つコードブックの検討も行われた。代表的なアプローチは格子(lattice)型のベクトル量子化である。第2章の対称性を持つ木探索コードブックとの関係で興味深いのは鏡映群による超球面一様分布向けのベクトル量子化である[5-5]。平均値分離正規化されたベクトルは超球面上に対称な分布を持つが、一様分布ではない。一様性を前提とすればコードブックは格子型となる。[5-5]は超平面に対する鏡像によって対称な部分空間とするため、これを万華鏡ベクトル量子化と表現した。これに対して第2章で議論した対称性は親節点の出力ベクトルに対する点対称である。

今日、画像に関するベクトル量子化は、画素をブロック化したパターンに対するストレートな圧縮手段に限らず、複数シンボルの非線形な相関を活用する要素技術として、他の手法と組合せて活用されている[例えば5-6]。またテクスチャの分類など信号解析あるいは合成にも活用されている[5-7]。かつて制約だった演算規模やメモリ容量の制約はかなり緩和された。現在の条件で2.4.4に述べたPR/DCの活用を再訪する価値があるのではないか、新たな角度で認識との接点がありうる(5.3参照)と筆者は考えている。

第3章の動的適応ウェーブレットパケットによる適応変換の議論は1992-1996年の研究成果である。当時の成果は筆者の知る限り、静止画圧縮について2017年の現在もState of the artである。固定的なウェーブレット変換は静止画符号化の国際標準JPEG 2000に採用され、業務用途を中心に活用されているが、普及は旧世代のJPEG標準に遠く及ばない。JPEG 2000の複雑度低減については、現在でも簡略化や簡略版の標準策定について要望があり、普及に至らない一つの要因は複雑性と考えられる。

ウェーブレットの動画符号化への提案は何度か試みられているが、

標準への採用に至っていない。複雑性以外の課題の一つは動き補償予測との親和性である。本課題について例えば、オーバーラップ動き補償やアフィン変換を含む動き補償予測によって動き補償予測誤差信号のブロック境界における不連続性を改善するなど、利用可能な演算規模が拡大した今でこそ可能な手段の活用が期待される。

ウェーブレットは信号解析のツールとして圧縮以外の用途にも盛んに研究が行われている。圧縮が最優先の目標でない時には基底として過完備つまり変換係数の数が入力画像の画素数より多い変換も使われる。この時は基底でなくフレームと呼ばれる。画像解析におけるアプローチの一つはテクスチャの分類である。研究例として、ウェーブレットによるフレームを用いたテクスチャの分類・分割[5-8]や周波数帯域内あるいは帯域間に渡る相関を用いた分類の提案[5-9]などがある。第3章と関連する方向ではウェーブレットパケットと混合ガウスモデルをパターン認識志向で用いたテクスチャ分類・分割の研究[5-10]がある。

またウェーブレットの定義からは派生となるが、「2次元画像におけるエッジの方向」の弁別性を重視する変換が検討されている。ラドン変換を用いて2次元画像をエッジの角度と周波数の平面で表すリッジレット変換やカーブレット変換、コンターレット変換などのアプローチである[1-34,5-11,5-12,5-13]。

第4章の国際標準化に関する議論は1988年からほぼ現在に至る成果である。1990年のH.261以来、動き補償予測付きDCT符号化という枠組みに時代のニーズと同時代的な実現性を考慮した複雑度を取り入れて進化してきた。現在もなお、同じ枠組みで次の進化のサイクルを回す議論を行なっている[5-2,5-3]。これについては5.2節で述べる。

以下、本章では5.2節で産業のための画像符号化について現状とこれからに対する筆者の考えを述べ、最終5.3節では画像符号化が長期的視点で目指すものについて考えて結びとする。

5.2 産業技術としての画像符号化

産業技術としての画像符号化技術は国際標準を抜きにして考えられない。圧縮率を高めることについて今も要求は止まない。かつて MPEG-2 や MPEG-4 標準が策定され、通信速度や蓄積デバイス容量がムーアの法則に従って増大することから新しい圧縮の必要性は低いと言われた。MPEG 標準化が検索の MPEG-7 やコンテンツ流通の MPEG-21 に進んだ背景にはこの認識があった。しかし今日、高解像度化、ストリーミングによる映像トラフィックの増加、SNS を含めたモバイル環境での動画視聴の拡大が、圧縮性能の向上を求めている。符号量を 30%削減できれば、帯域あたりのストリーム本数を 40%増やせる。それがサービス業者の収入増につながるかもしれない。今日、産業上の動機に支えられ、国際標準の対抗軸としてボランティアな共同開発や IT 企業による動画像符号化技術の開発が行われている。しかし国際標準はなお大きな存在であり続けると筆者は考えている。

現在の国際標準化は動画にせよ、静止画にせよ、さらに圧縮率を高めた符号化方式、光線空間や点群など 3 次元画像の新たな形式、VR や AR との親和性を拡大する方向で進んでいる。例えば MPEG における Future Video Coding の検討では、HEVC|H.265 にツールを追加することによってほぼ 30%の符号量削減が可能、ただし演算量は 10 倍以上という報告がなされている [5-15]。提案されているツールは様々であるが、マクロブロックサイズをさらに拡大する提案は有力である。いずれ図 4-24 の範疇に収まるものだろう。

現状が示すのは、動き補償予測付き DCT 符号化という枠組みの強靱さである。この枠組みは四半世紀に渡る洗練や改善、性能強化の工夫を経ており、標準化に携わるコミュニティに絶大な信頼を得ている。4.5 節で述べたように図 4-24 が象徴する最適化がなされている。この環境下で、全く別の視点による試みが既定の枠組みを崩すことは難しい。例えばウェーブレットやフラクタル、マッチングパーシュートなども過去に提案されたが、動画像符号化の標準には採

用されていない。現状の枠組みで考えるか、または長期的視点ないし圧縮以外の機能性で訴求するかを選択しなければならない。

本節は前者の視点で考えるものとして、筆者は2つの点を挙げたい。一つは温故知新、もう一つはリポジトリの整理活用である。

温故知新とは、これまで検討・提案されたが本格的に採用されていないツールの見直しである。かつて性能改善と複雑度のバランスから棄却された手法、洗練不足であった手法が現在の環境では有効かもしれない。例えば前節で触れたように、オーバーラップ動き補償（H.263 に採用実績はあるが活用は不十分）やアフィン変換（MPEG-4の一部に実績あり）とウェーブレットの組合せなど、色々な可能性があるだろう。

リポジトリの整理活用とは、既に標準に採用されたツール、これから採用されるツールを含めて再評価の上、リポジトリ化して動的に組合せ可能な仕組みを作ることである。それぞれのツールは万能でなく、特に効果を発揮する入力画像や動き特性への得意不得意を持つ。それを見極めて、ツールの新たな組合せを何種類か定義する。これを入力画像に適応して切り替えれば、全てのツールを包含するよりも効率が良いはずである。今後追加されるツールも前提とする特性を見極めて整理すべきだろう。

後者に通じる考え方は2004年頃からMPEGの中にもあった。この活動はRVC（Reconfigurable Video Coding）と呼ばれた。当時、筆者はRVCに関するアドホックグループの議長を務め、RVCの啓発とツール整理の準備を行っていた[5-16]。動的再構成可能なデバイスへの期待があった時期でもあり、RVCはデバイス側面から見られることも多かったが、本命は動的なソフトウェア実装であった。現在、RVCはツールの組合せを記述するISO/IEC 23001-4: Codec Configuration representation[5-17]とツールを集積したリポジトリであるISO/IEC 23002-4: Video Tool Library[5-18]に分かれて標準化され、今もツールが追加されている。惜しむらくは各ツールの特性整理と活用が伴っていないことである。

5.3 画像符号化のビジョン

画像符号化の目標はまず圧縮によって通信・放送や蓄積の問題を解決することであった。この問題はある程度解決したと言える。その上で高解像度化や3D化に拡張していく動きがある。それらはそれぞれの方向で前節の範疇である。今後の長期的視点や圧縮以外の機能についてどう考えるか。本節では3つの視点から考える。

第1の視点は知的符号化である。原島は1988年に「知的画像符号化」の概念について述べた[5-19]。情報のいかなるレベルまで立ち入るかによる世代区分が提示された。表5-1は同文献からの引用である（原典は[5-20]に遡る）。

表 5-1 画像符号化の世代区分（[5-19]より）

世代区分	符号化方式	符号化に用いる知識	符号化レート (bit/s)	符号化方式の例
第0世代	直接符号化	振幅分布	$10^7 \sim 10^8$	標本値ごとのハフマン符号化など
第1世代	統計的冗長圧縮符号化	画素間相関などの統計的性質	$10^5 \sim 10^7$	予測符号化 変換符号化 ベクトル量子化など
第2世代	構造抽出符号化 特徴抽出符号化	動き情報・輪郭情報などの特徴情報	$10^4 \sim 10^5$	輪郭符号化 構造・特徴情報に基づく適応符号化
第3世代	分析合成符号化 (モデルによる符号化) (パラメータ符号化)	画像のモデル 図形分離 特徴抽出	$10^3 \sim 10^4$	構造モデルに基づく顔画像の分析符号化
第4世代	認識合成符号化 (コマンド符号化)	画像の認識と生成に必要な知識と規則	$10^2 \sim 10^3$	
第5世代	知的符号化	画像の概念モデル 推論・思考 異種情報の統合	$10 \sim 10^2$	

さらに知的符号化の応用として、超圧縮、情報の構造化、知的インターフェース、知的情報通信ネットワークが挙げられた。

原島ら自身は第3世代分析合成符号化の例として人間の顔の構造を3次元ワイヤフレームモデルで表し、顔画像のテクスチャを張

り付ける例を示した。国際標準を含めて予測・変換符号化など多くの符号化技術は第 1 世代に分類される。

上記知的画像符号化は、画像符号化をキーとして、画像処理、パターン認識、コンピュータービジョン、グラフィックス、人工知能、知識情報処理を統合しようとする視点を持っている。圧縮と伸長だけではなく、信号の構造的表現を経て認識と合成への接点を意識する方向性である。この視点は現在も有効で、進むべき方向に示唆を与えると筆者は考えている。

第 2 の視点には 1994 年に Wang と Adelson が示した画像符号化方式の分類、ローレベル、ミッドレベル、ハイレベル符号化 [5-21] を引用した斎藤らの議論 [5-22] を用いる。ローレベルは波形近似、ミッドレベルは領域やレイヤーを分離した構造的表現、ハイレベルは被写体に対応した特徴量で被写体を表現する符号化である。

Wang らの関心は自らの提案方式をミッドレベル符号化と位置付けることであり、斎藤らの関心は学会誌の MPEG 特集で読者の関心を慮りつつ未来志向を担うことにあったと筆者は解釈している。それぞれの関心が文献名の "Representing moving images ~ [5-21]" と 「ビデオデータ圧縮 ~ [5-22]」 に現れている。

斎藤らの文献ではミッドレベル符号化の紹介に力点が置かれた。ミッドレベルでは 2,3 次元のオブジェクト（被写体）モデル [5-23] やセル画のアニメーションのように複数のレイヤー（前景、背景など）で画像を表現するモデル [5-21] を用い、画像提示の側面を主眼としつつ、自然画像をレイヤー表現する大域的セグメンテーションを通して符号化に取り組む技術と事例が紹介された。画像を提示する技術が先行したのは研究者にコンピュータービジョンのチームが多かったことが要因だろう。またハイレベル符号化では原島らのように顔や人物肩上の 3 次元構造を活用した研究が紹介された。（筆者らのチームも行って [5-24]。）

これらの研究は MPEG-4 標準化に大きな影響を与えた [5-25]。第

4章では矩形画像の符号化に関する議論に集中したが、個々の被写体を形状情報とテクスチャ情報で表し、画面の上にオブジェクトとして配置する符号化、人間の身体をモデル化してアニメーション表現を可能にする符号化も MPEG-4 標準の一部として策定された(図 5-1 参照)。しかし汎用的な入力画像をオブジェクトやレイヤーに分離するには当時(2000年頃)の技術は未熟であり、分離を伴う符号化標準としての本格的な普及には至っていない。

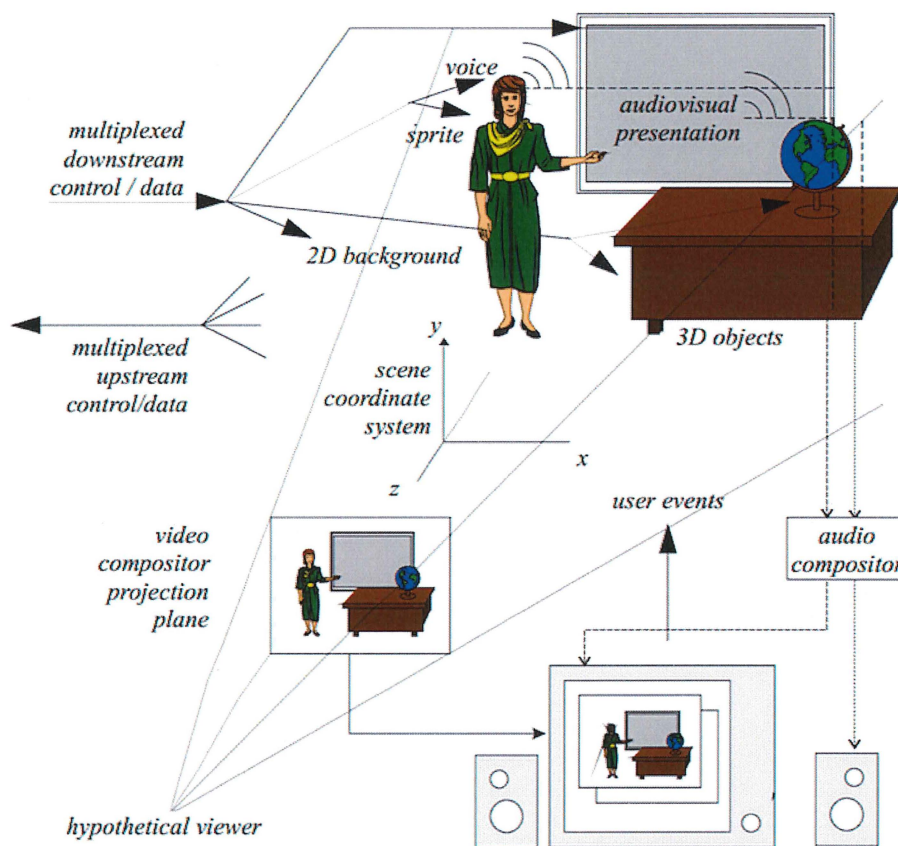


図 5-1 MPEG-4 によるシーンの表現例 ([5-25]より)

今日、オブジェクトやレイヤーを分離する要素技術の成熟は進んでおり、例えばスマートフォンが備えるカメラとセンサ情報を活用した分析で 3D 表現や VR/AR コンテンツ生成を行うアプリケーションが登場している。また顔画像のハイレベル符号化を想起させる事例として、俳優が(モーションキャプチャ用のマークをつけて)演

じた映像のワイヤフレームモデルに、故人となった俳優のテクスチャをマッピングすることでリアルな映像が製作されるようになった[5-26]。つまり制約付きで画像以外の情報も利用してはいるが、グラフィクス先導でオブジェクト符号化の環境条件は確実に進展しているのである。

図 5-1 にある MPEG-4 の概念は既に VR や AR を包含している。過去の研究、MPEG-4 の議論を再訪し、現時点の産業活用を含めて再検討することが有効だと筆者は考えている。

第 3 の視点は画像認識との関係である。この分野では近年、深層学習 (Deep Learning) という進展がある。2012 年、同技術を採用したチームが画像認識のコンペティションで飛び抜けた成績を示した[5-27,5-29]ことから火がついた状況は今も続いている。深層化されたネットワークは学習の過程で与えられた膨大な画像データを基に高精度の「分類器」として機能した。そのメカニズムが完全に解明されたわけではないが、ヒントはある。CNN (Convolutional Neural Network) についてニューロンの学習結果を可視化した際、浅い Convolution 層で各パターンのエッジが観察されることが報告されている[5-28,5-29]。画像を入力すれば絵柄に近いエッジのパターンがフィルタとして強い出力を持ち(すなわちニューロンが発火)、Pooling 層で縮小や強いエッジが選択的に生き残る処理が行われる。つまりスケールや位置の情報が正規化される。続く階層にテクスチャを含むパターンが観察される。ここまでは認識の準備として納得感がある。段数が進むと画像の「部分」に相当する高次元のパターンが現れるとの報告があるが、過学習の可能性もあり、一般化した説明は困難なようである。第 2 章のベクトル量子化コードブック、特に木探索コードブックの浅い段に出力ベクトルの構造的なパターンが現れることを観察してきた筆者は CNN の学習結果に対して既視感めいた印象を持たずにいられない。

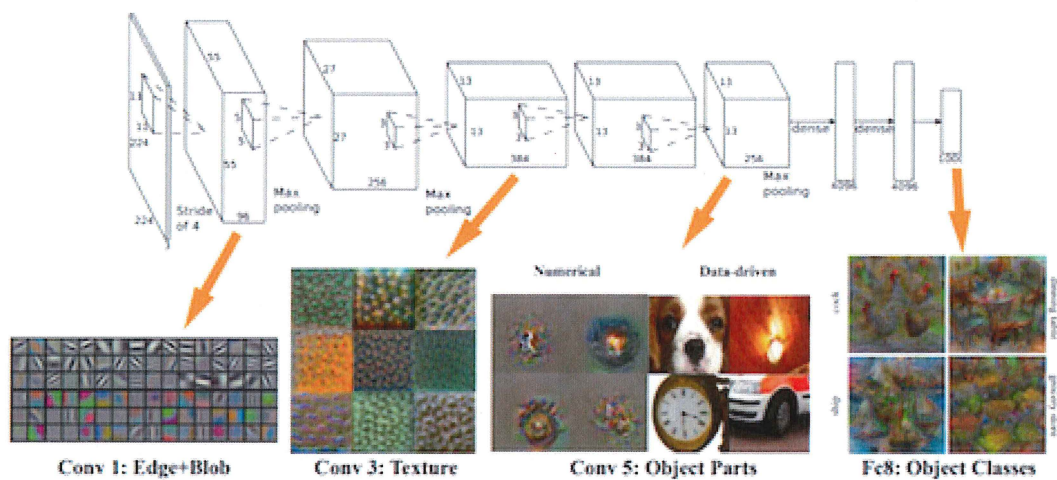


図 5-2 学習後のニューロンの可視化例 ([5-29]より)

深層学習の意義を、従来は人間が設計していた特徴量を計算機が表現学習する点だとすれば、深層学習は人間が見つけられなかった特徴量を見つけてくれるかもしれない。学習結果を分析することで、新たな特徴表現 (representation) の方法を見つけることができるかもしれない。気づかなかった基本パターンを示してくれるかもしれない。

例えば図 5-3 に示す人工的な信号源のモデルを考える。図 5-3 の左に書かれた基本波形がこの信号源の 2 つの本質的な特徴ベクトルである。信号源に適応するカルーネン=レーベ変換は直交条件を持つためにこの 2 つの波形を見つけてくれない。この例は単純だが、もっと複雑な場合に基本波形がどれだけあり、それらがどのような波形であるかという問題をどのようにして解くか。

図 5-3 は図 2-17 の人工的なアナロジーに過ぎない。図 2-17 のコードブックに新たなパターンあるいはテクスチャの表現手法を追加して、画像を汎用的にパターンの組合せで表現できるのでないか、基本パターンの集合はいかにして見出せばいいか、基本パターンは画像を構成する以外の認識的な用途に使えるのではないか、超圧縮と認識を統合したフレームワークを考えられるのではないか。基本

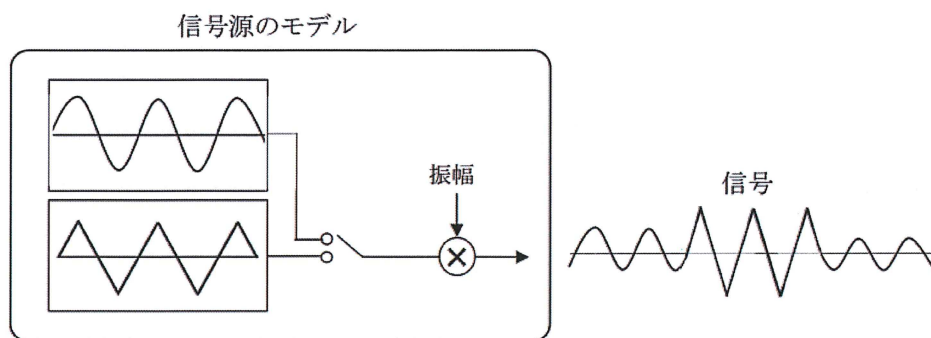


図 5-3 人工的な 1 次元信号源のモデルと波形

は単純すぎるほど単純な発想ではあるが、フレームワークとしては有効であると筆者は考えている。深層学習に過剰な期待はしないとしても、基本パターンの表現や生成を試行する一つの手段になるかもしれない。MPEG 標準化においても MPEG-7 の流れを組む特徴記述の標準化に深層学習を特徴量抽出の道具として使う検討が行われている [5-14]。

以上、画像符号化の今後について 3 つの視点で述べた。

第 1 の知的符号化の視点は、画像情報には低次の物理的な信号値から意味・概念に至る階層性がありうると語っている。高次符号化の階層に至るには課題が多く残っているが、依然として今後の研究の方向性に示唆を与えてくれる。

第 2 の視点は、画像が 3 次元の被写体や背景を投影した構造的な情報であると考えた場合、構造的な情報の抽出と提示を組みにして符号化を考える必要があると語っている。特に提示のためのグラフィックス技術は大きく進展している。また従来は困難であった構造の抽出についても、演算性能やセンサ情報の活用によって実現性が高まっている。

第 3 の視点は、情報の高次階層に到達可能と言える段階ではないとしても、分類・認識に向けた新しい道具が利用可能になりつつあることを告げている。

本研究のまとめとして、いささか抽象的ではあるが、今後の方向性を以下のように書いてみたい。

「(画像) 情報が持つ信号としての構造および意味・概念としての構造をいずれかのレベルで捕捉する特徴量の組を見出してパラメータ化することにより、高効率の圧縮と信号表現、さらには表現に基づく分析によって工学的価値を提供すること」

画像符号化における信号圧縮の機能は今日、筆者の研究成果を含み、既に実用化されている。今後の研究開発の目標について、現在以上の圧縮という要求もあるが、むしろ新たな価値やサービスを実現することを優先した目標設定が望まれる。遠方に高次の目標を想定するにせよ、途中の段階においても価値を創出する成果を生み続ける計画が必要である。我々は今、どの段階にいるのか、どんな価値を提供できるのかを考え、可能な範囲で貢献していきたいと考えている。

付録

2 分割フィルタバンクによる局在化利得について

2 分割の正規直交フィルタバンクを用いるウェーブレットパケットは 2 進木に対応づけて表すことが可能である。この時、電力局在性の観点で最適なウェーブレットパケットを求めるために、帯域分割の数と電力の局在化利得に対して一般化 BFOS アルゴリズム [2-21]を用いる。ここでは本文中に詳しく書かなかった同アルゴリズムの適用要件について説明する。

一般化 BFOS アルゴリズムとは、Breiman らによるアルゴリズム [2-48]を Chou らが一般化し、符号量-歪み特性的に最適な木探索ベクトル量子化器の設計をはじめとして、木構造に対応づけられる多くの問題に適用可能としたアルゴリズムである。同アルゴリズムを適用するための条件は、節点に対応付けられた関数が linear もしくは affine の意味で加算的であること、それらの関数が木の段数に従って単調（単調増加の場合は増加ないし同値、単調減少の場合は減少ないし同値）であることである。

Linear であるとは終端節点の関数値の和が木全体の関数値となることを意味し、affine であるとは全節点の関数値の和が木全体の関数値となることを意味する。

フィルタバンクによって分割される帯域の数は終端節点の数に対応するので明らかに linear で上記の条件を満たす。次に、帯域分割による利得が affine として上記の条件を満たすことを確認する。帯域は 2 分割される度に 2:1 サブサンプリングされるものとし、信号源は十分大きなサンプル数と信号電力を持つものとする。すなわち帯域分割されたサンプルの分散は各分割において 0 より大きな値を持つものと仮定する。以下、最初に完全 2 進木の場合について考える。

帯域を 2 分割することによる利得 g を以下の式、すなわち分散の相加平均と相乗平均の比で表す。局在化利得は木の段数に従って増

加する（または同値の）単調増加関数である。

$$g_j^{i-1} = 1/2 (\sigma_{2j}^{i-2} + \sigma_{2j+1}^{i-2}) / \sqrt{\sigma_{2j}^{i-2} \cdot \sigma_{2j+1}^{i-2}}$$

2進木の何段目であるかを i と表し、第 i 段における接点の番号を j ($j = 0, 1, \dots, 2^i - 1$) と表すものとする。上の式は $i > 0$ を前提としており、 $i = 0$ は 2 進木の根に対応するものとする。また、各節点に対応付けられた帯域に属するサンプルの分散を σ_j^{i2} と表すものとする。

今、完全 2 進木が第 i 段までであるとすると、第 i 段に属する終端節点の数は 2^i であり、この時の局在化利得 G^i は [1-4] から引用して、

$$G^i = 1/2^i \sum_{j=0}^{2^i-1} \sigma_j^{i2} / \left[\prod_{j=0}^{2^i-1} \sigma_j^{i2} \right]^{1/2^i}$$

である。これを第 $i-1$ 段以前にさかのぼって表現することを考える。完全 2 進木が第 $i-1$ 段までであれば、局在化利得は以下で表される。

$$G^{i-1} = 1/2^{i-1} \sum_{j=0}^{2^{i-1}-1} \sigma_j^{i-12} / \left[\prod_{j=0}^{2^{i-1}-1} \sigma_j^{i-12} \right]^{1/2^{i-1}}$$

2 分割フィルタバンクの正規直交条件より

$$\sigma_{2j}^{i-2} + \sigma_{2j+1}^{i-2} = 2\sigma_j^{i-12}$$

これを G^i の式の分子に代入する。分子は G^{i-1} と同じ形となる。

$$G^i = 1/2^{i-1} \sum_{j=0}^{2^{i-1}-1} \sigma_j^{i-12} / \left[\prod_{j=0}^{2^i-1} \sigma_j^{i2} \right]^{1/2^i}$$

ところで、第 $i-1$ 段から第 i 段への分割による局在化利得を第 $i-1$ 段の親節点について考えれば、最初に定義したように、

$$g_j^{i-1} = 1/2 (\sigma_{2j}^{i-2} + \sigma_{2j+1}^{i-2}) / \sqrt{\sigma_{2j}^{i-2} \cdot \sigma_{2j+1}^{i-2}}$$

これを以下のように変形する。

$$g_j^{i-1} = \sigma_j^{i-12} / \sigma_{2j}^{i-2} \cdot \sigma_{2j+1}^{i-2}$$

さらに第 i 段の要素を第 $i-1$ 段の要素で表すために、以下のように変形する。

$$\sigma_{2j}^i \cdot \sigma_{2j+1}^i = \sigma_j^{i-1^2} / g_j^{i-1}$$

次に G^i の式に戻り、 G^{i-1} と g_j^{i-1} を用いて G^i を書き換える。

$$\begin{aligned} G^i &= G^{i-1} \cdot \left[\prod_{j=0}^{2^{i-1}-1} \sigma_j^{i-1^2} \right]^{1/2^{i-1}} / \left[\prod_{j=0}^{2^{i-1}-1} \sigma_j^{i^2} \right]^{1/2^i} \\ &= G^{i-1} \cdot \left[\prod_{j=0}^{2^{i-1}-1} \sigma_j^{i-1^2} \right]^{1/2^{i-1}} / \left[\prod_{j=0}^{2^{i-1}-1} (\sigma_{2j}^i \cdot \sigma_{2j+1}^i)^2 \right]^{1/2^i} \\ &= G^{i-1} \cdot \left[\prod_{j=0}^{2^{i-1}-1} \sigma_j^{i-1^2} \right]^{1/2^{i-1}} / \left[\prod_{j=0}^{2^{i-1}-1} (\sigma_j^{i-1^2} / g_j^{i-1})^2 \right]^{1/2^i} \\ &= G^{i-1} \cdot \left[\prod_{j=0}^{2^{i-1}-1} \sigma_j^{i-1^2} \right]^{1/2^{i-1}} / \left[\prod_{j=0}^{2^{i-1}-1} (\sigma_j^{i-1^2})^2 \prod_{j=0}^{2^{i-1}-1} g_j^{i-1-2} \right]^{1/2^i} \\ &= G^{i-1} \cdot \left[\prod_{j=0}^{2^{i-1}-1} \sigma_j^{i-1^2} \right]^{1/2^{i-1}} / \left[\prod_{j=0}^{2^{i-1}-1} \sigma_j^{i-1^2} \right]^{1/2^{i-1}} \cdot \left[\prod_{j=0}^{2^{i-1}-1} g_j^{i-1-1} \right]^{1/2^{i-1}} \end{aligned}$$

分母と分子から共通因子をくくり出して約分する。

$$G^i = G^{i-1} \cdot \left[\prod_{j=0}^{2^{i-1}-1} g_j^{i-1} \right]^{1/2^{i-1}}$$

すなわち、第 i 段の分割による利得は第 $i-1$ 段の分割による利得に対して、第 $i-1$ 段から第 i 段への 1 段分割によって生じる利得を（段に応じた節点数すなわち信号源を分割するサイズの比率に応じたべき乗根を含めて）乗算することにより漸化的に表される。従って帰納的にさかのぼり、 $G_0 = 1$ として第 0 段である根からの局在化利得をすべて乗算した形式で表現することができる。

$$G^h = \prod_{i=1}^h \left[\prod_{j=0}^{2^{i-1}-1} g_j^{i-1} \right]^{1/2^{i-1}}$$

両辺の対数を取り、利得の乗算を対数の加算で表す。

$$\log G^h = \sum_{i=1}^h \left[\frac{1}{2^{i-1}} \sum_{j=0}^{2^{i-1}-1} \log g_j^{i-1} \right]$$

対数で表現した局在化利得 (dB) は各分割による利得に対して加算的である。上の式において最初の Σ は段に関する和、次の Σ は各段の接点に関する和となる。

第 $i-1$ 段の第 j 節点における利得 $\log g_j^{i-1}$ を $1/2^{i-1}$ 倍した上で、第 i 段の子節点、第 $2j$ 節点と第 $2j-1$ 節点に半分ずつ配分し、これを局在化利得に相当する関数とすれば、同関数は **affine** となる。

上記は完全 2 進木として考えたが、刈り込みのある 2 進木は刈り込まれた枝 (分岐) における局在化利得 g_j^{i-1} を 1 (=0dB) と考えれば上記の場合と同等である。

以上より、節点に対応づける関数を帯域分割数と局在化利得とした場合、これらは一般化 BFOS アルゴリズムの適用要件を満たす。

謝辞

本論文の基になった研究は筆者が勤務する三菱電機株式会社、留学の機会を得たコロンビア大学、また一部は東京大学において、長期間に渡り行われたものである。多くの人の助力があって可能になった研究でもある。これらの組織だけでなく、学会や国際標準化活動、国家プロジェクト、数々の共同開発において出会った大学や研究機関、企業の人々にも多くを負っている。続いて御礼を申し上げる方々のご所属は特に注記のない限り、当時のものである。

画像符号化の研究を始めたのは猪瀬博・齊藤忠夫両先生（東京大学）の研究室で画像符号化を卒業論文のテーマにした時からである。当時博士課程におられた谷中一寿氏（現神奈川工科大学教授）に直接の指導を受けた。指導は今や古典に属するレビュー文献や画像符号化シミュレーションに必要なソフトウェアライブラリ的设计など、大変に有益な内容を含んでいた。三菱電機に入社してから短期間の内に実験を始めることができたのは氏の指導があったからである。

三菱電機入社以降、最もお世話になったのは村上篤道氏である。村上氏は直接の指導者であり共同研究者であった。社会人としての常識すら理解していなかった筆者にとってはメンターでもあった。氏による指導は信号処理・符号化の基礎から研究の姿勢、特許の書き方、開発手法、さらに礼儀や振る舞いにまで及んだ。その後の研究開発に筆者が機会を得たことについても、氏に多くを負っている。大変残念なことに村上氏は2015年10月に急逝された。氏の生前に本論文をまとめられなかったことは筆者の悔いの一つである。氏には何度感謝を申し上げても過ぎることはない。

筆者が新人の頃に同僚としてご指導を頂いた杉山康夫博士、井上徹氏、河原林成行氏、また音声のグループを率いておられた中島邦男氏に感謝する。彼らを含む通信研究部諸氏の多くが研究者として極めて真摯でありながら強い個性の持ち主でもあったことは入社後何年か実感することがなかった。社会人生活の最初にそうした環境にいられたことは筆者にとって幸運であった。

共同研究者である多くの同僚とかつての同僚に感謝する。全ての人の名前を挙げることはできないが、特に本論文につながる共著の機会を得た伊藤敦氏、山岸篤弘氏、上澤功氏、西田正実氏、加藤嘉明氏、大平英雄氏、長谷川由里氏、松崎一博氏、福原隆浩氏、西川博文氏、山田悦久氏、関口俊一氏、井須芳美氏、杉本和夫氏、黒田慎一氏、山岸秀一氏、共著に加えて共同開発で濃密な時間を共にした浅野研一氏、丹野興一氏、鈴木光義氏、鈴木隆太氏、小川文伸氏に深く感謝する。INS 大船工場、郡山製作所、コミュニケーション・ネットワーク製作所等における開発で多くの人々と共有した時間が筆者の活力になったことも書いておきたい。海外研究所勤務の同僚 Huifang Sun 氏、Anthony Vetro 氏、Miroslaw Bober 氏には国際標準化活動に関連した共同研究と国際協力に感謝する。筆者が研究者として最も実りの多い時期に研究所長として高所からご指導を頂いた片木孝至氏、野間口有氏に感謝を申し上げたい。上に挙げた三菱電機諸氏の助力なしに本論文が書かれることはなかった。

1992 年から 93 年にかけての 1 年間、会社派遣によってコロンビア大学に滞在する機会を得た。共同研究者であった Martin Vetterli 教授、Kannan Ramchandran 氏、Antonio Ortega 氏、Cormac Haley 氏、Truong-Thao Nguyen 氏に感謝する。この時の研究成果が第 3 章に現れている。この時期の出会いはその後の国際標準化における再会や交流につながっている。その意味で第 4 章の一部にも繋がる有益な滞在であった。

筆者が画像符号化技術の研究に従事した時期は同技術が学会、国際標準化および産業界において大きく進展した時期でもある。結果的には企業内研究者として自然なことであるが、大変幸福なことでもあった。学会では原島博教授（東京大学）、谷本正幸教授（名古屋大学）に現在に至る長い間、ご指導を頂いている。斎藤隆弘教授（神奈川大学）、相澤清晴教授（東京大学）にもご指導や多くの刺激を受けた。その影響は第 1 章から第 5 章までの全てに現れている。博士論文のご相談にのって頂いた青山友紀教授、博士課程でお世話にな

った森川博之教授と研究室諸氏（東京大学）に感謝する。同研究室では社会人博士課程に在籍しつつ、筆者の事情で学位取得には至らなかったが、多くの刺激と交流の機会を得た。第4章の一部には研究室での議論も反映されている。

筆者にとって第4章の国際標準化活動は大きな存在である。H.261標準化については終盤のみ参加の機会を持ったが、大久保栄議長（NTT）の着実な進行と今日に到る映像符号化標準の基盤構築に居合せることができたのは幸運だった。MPEG標準化については創設者であり親委員会の議長、安田浩氏（NTT）とMPEG議長のLeonardo Chiariglione氏（CSELT）に深く感謝する。安田氏（現東京電機大学学長）の後継議長であり筆者の前任でもある渡辺裕氏（NTT、現早稲田大学）とは学会や実開発で何度も意見を交わす機会があった。滝沢正明氏（日立製作所）、小杉康弘氏（富士通研究所）、山口博久氏、和田正裕氏（以上KDD）、大塚吉道氏（NHK）には多くの示唆や刺激を受けた。太田睦氏（NEC）、中島康之氏（KDD）、酒井潔氏（富士通研究所）、高橋俊也氏（松下電器）、行武剛氏（松下通信工業）、米満潤氏、矢ヶ崎陽一氏（以上ソニー）、杉山賢二氏（JVC）、片山康男氏（GCT）、尾高敏則氏（東芝）、鹿喰善明氏（NHK）との議論や競争は研究開発の活力になった。MPEG-4標準化では栄藤稔氏（松下電器）、中屋雄一郎氏（日立製作所）の提案に刺激を受けた。MPEGのサブグループの議長も務められた小暮拓世氏（松下電器）には標準化を含めた多くを教わった。標準化で出会った海外の人は数多く、名前の列挙すらできないが、Barry G. Haskell氏（AT&T）、Geoffrey Morrison氏（BT）の両氏はそれぞれ米国と欧州を代表する参加者として多くの意見交換の機会があった。第2章のベクトル量子化研究の頃、また第3章に対応するコロンビア大学留学時期に接点があり、第4章の国際標準化MPEG-4 AVC|H.264以降の議長を務めるGary Sullivan氏とも長い交流がある。全ての方々に感謝申し上げる。

学位論文の執筆を促してくれた多くの方々に感謝したい。その度

に口を濁してきた筆者の怠慢をお詫び申し上げます。ごく最近まで励まし続けてくれた西村隆司氏、開発も共にした本多孝司氏（三菱電機）の言葉がなかったら断念していたかもしれない。

本論文執筆の直接の契機になったのは水野忠則教授（愛知工業大学）の強い励ましであった。水野教授のご指導を受け、さらに同学の中條直也教授、森本正志教授、田村隆善教授にご指導とお世話を頂くことができた。いざ書き始めてみれば執筆の時間は非常に楽しい時間でもあった。深く感謝申し上げます。

本論文の執筆やその基になった研究開発の時間は充実していたが、長い時間でもあった。家族にとって筆者は会社や出張から帰らない放蕩息子で、戻ってきても研究開発を引きずっている厄介者であったと思う。長い時間を辛抱強く見守ってくれた家族に感謝する。

参考文献

第 1 章の参考文献

- [1-1] Claude E. Shannon, "A Mathematical Theory of Communications," Bell System Technical Journal, Vol. 27, 1948.
- [1-2] Arun Netravali, John O. Limb, "Picture Coding: A Review," Proc. of the IEEE, Vol. 68, No. 3, pp.366-406, March 1980.
- [1-3] Anil K. Jain, "Image Data Compression: A Review," Proc. of the IEEE, Vol. 69, No. 3, pp.349-389, March 1981.
- [1-4] N. S. Jayant and Peter Noll, "Digital Coding of Waveforms," Prentice-Hall, Englewood-Cliffs, NJ, 1984.
- [1-5] Anil K. Jain. "Fundamentals of Digital Image Processing," Prentice-Hall, Englewood Cliffs, NJ, 1989.
- [1-6] Toby Berger, "Rate Distortion Theory," Prentice Hall, 1971.
- [1-7] W. M. Goodall, "Television by Pulse Code Modulation," Bell System Technical Journal, Vol. 30, pp. 33-49, January 1951.
- [1-8] P. Elias, "Predictive coding," IRE Trans. Information Theory, IT-1, pp.16-33, March 1955.
- [1-9] J. B. O'Neal, "Predictive Quantizing Systems (Differential Pulse Code Modulation) for the Transmission of Television Signals," Bell System Technical Journal, Vol. 45, pp. 689-721, May 1966.
- [1-10] 瀧保夫, 羽鳥光俊, 田中茂, "動きに追随するフレーム間符号化", 信学全大, pp. 1263, 1974年7月.
- [1-11] Barry G. Haskell, "Entropy Measurements for Non-Adaptive Frame-to-Frame, Linear Predictive Coding of Video Telephone Signals," Bell System Technical Journal, Vol. 54, No. 6, pp. 1155-1174, August 1975.

- [1-12] Arun N. Netravali and J. D. Robbins, "Motion Compensated Television coding – part I," Bell System, Technical Journal, vol. 58, pp. 631-670, March 1979.
- [1-13] Yuichi Ninomiya and Yoshimichi Ohtsuka, "A motion-compensated interframe coding scheme for television pictures," IEEE Trans. Comm., COM-30, pp. 201-211, January 1982.
- [1-14] M. Loeve, "Probability Theory," 2nd Edition, pp. 478, Van Nostrand, Princeton, 1960.
- [1-15] N. Ahmed, T. Natarajan and K. R. Rao, "Discrete Cosine Transform," IEEE Trans. Comp., C-23, pp. 90-93, 1974.
- [1-16] W. H. Chen and W. K. Pratt, "Scene Adaptive Coder," IEEE Trans. Comm., Vol.32, No. 3, pp. 225-232, March 1984.
- [1-17] H. S. Malvar, "The LOT: Transform Coding without Blocking Effects," IEEE Trans. ASSP, Vol. 37, No. 4, pp. 553-559, April 1989.
- [1-18] J. H. Princen, "Analysis/Synthesis Filter Bank Design based on Time Domain Aliasing Cancellation," IEEE Trans. ASSP, Vol. 34, No. 5, pp. 1153-1161, October 1986.
- [1-19] Stephane G. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," IEEE Trans. PAMI, Vol. 11, No. 7, July 1989.
- [1-20] Ingrid Daubechies, "Orthonormal Bases of Compactly Supported Wavelets," Communications on Pure and Applied Mathematics, Vol. XLI 909~996, 1988.
- [1-21] Ingrid Daubechies "Ten Lectures on Wavelets," Society for Industrial and Applied Mathematics, 1992.
- [1-22] Arnaude E. Jacquin, "Image Coding Based on a Fractal Theory of Iterated Contractive Image Transformations," IEEE Trans. Image Processing, Vol. 1, No. 1, January 1992.

- [1-23] Henry Stark, "Image Recovery: Theory and Application,"
1st Edition, Academic Press, Elsevier, March 1987.
- [1-24] Kohtaro Asai, Nguyen T. Thao, Martin Vetterli, "A Study
of Convex Coders with an Application to Image Coding,"
IEEE ICASSP 94, 1994.
- [1-25] 浅井光太郎, 村上篤道, 太田堯久, "適応型 AR 予測器による
DPCM 符号化特性," 信学全大, 688, 1982 年.
- [1-26] 太田睦, "変換・予測ハイブリッド動画像符号化方式の研究,"
博士論文, 2000 年 5 月.
- [1-27] 村上篤道, "高品位映像の高能率符号化と蓄積・伝送技術へ
の応用に関する研究," 博士論文, 2002 年 3 月.
- [1-28] 映像情報メディア学会編, "映像情報メディアハンドブック,"
オーム社, 2000 年.
- [1-29] 原島博監修, 映像情報メディア学会編 "画像情報圧縮," オ
ーム社, 1991 年.
- [1-30] 有本卓, "信号・画像のデジタル処理," 産業図書, 1980 年
- [1-31] 吹抜敬彦, "画像のデジタル信号処理," 日刊工業新聞社,
1981 年 5 月.
- [1-32] 村上篤道, "高能率符号化技術," テレビジョン学会誌, Vol. 42,
No. 11, pp. 1198-1204, 1988 年 11 月.
- [1-33] 酒井善則, 吉田俊之, "映像情報符号化," オーム社, 2001 年.
- [1-34] 電子情報通信学会, "知識ベース 2 群 5 編 (画像符号化)".
<http://member.ieice-hbkb.org/portal/>
- [1-35] 小野定康, 村上篤道, 浅井光太郎編著, "ユビキタス技術
動画像の高能率符号化," オーム社, 2005 年.
- [1-36] 村上篤道, 浅井光太郎, 関口俊一編, "高効率映像符号化技術
HEVC/H.265 とその応用," オーム社, 2013 年.

第 2 章の参考文献

- [2-1] Stuart P. Lloyd, "Least squares quantization in PCM," Unpublished Bell Laboratories Technical Note. Portion presented at the Institute of Mathematical Statistics Meeting, September 1957 (published in March 1982).
- [2-2] J. Max, "Quantizing for minimum distortion," IRE Trans. Information Theory, Vol. 6, No. 1, pp. 7-12, March 1960.
- [2-3] J. MacQueen, "Some methods for classification and analysis of multivariate observation," Proc. of the Fifth Berkeley Symposium on Math. Stat. and Prob., Vol. 1, pp. 281-296 1967.
- [2-4] Ali Habibi and Paul A. Wintz, "Image Coding by Linear Transform and Block Quantization," IEEE Trans. Communication, COM-19, No. 1, February 1971.
- [2-5] E. E. Hilbert, "Joint Pattern Recognition Data Compression Concept for ERTS Multispectral Imaging," Proceedings of SPIE, Vol. 66, pp. 122-137, August 1975.
- [2-6] Ali Habibi, "Survey of Adaptive Image Coding Techniques," IEEE Trans. Communications, COM-25, No. 11, November 1977.
- [2-7] Robert M. Gray, Andrés Buzo and Y. Matsuyama, A. H. Gray Jr., and J. D. Markel, "Source coding and speech compression," Proc. of the International Telemetry Conference, Vol. XIV, pp. 871-878, November 1978.
- [2-8] Andrés Buzo, A. H. Gray Jr., R. M. Gray and J. D. Markel, "Speech coding based upon vector quantization," IEEE ICASSP, April 1980.
- [2-9] Yoseph Linde, Andrés Buzo, Robert M. Gray, "An algorithm for vector quantizer design", IEEE Trans. COM, Vol. 28, pp. 84-95, 1980.

- [2-10] Allen Gersho and B. Ramamurthi, "Image coding using Vector Quantization," IEEE ICASSP, Vol. 1, pp. 428-431, April 1982.
- [2-11] Richard L. Baker and Robert M. Gray, "Image compression using non-adaptive spatial vector quantization," Asilomar Conference on Circuits Systems and Computers, October 1982.
- [2-12] Tokumichi Murakami, Kohtaro Asai, Eizo Yamazaki, "Vector Quantiser of Video Signals", Electronics Letters, pp. 1005-1005, November 1982.
- [2-13] Allen Gersho, "On the structure of vector quantizer", IEEE Trans. IT, Vol. 28, pp.157-166, 1982.
- [2-14] Robert M. Gray, "Vector quantization," IEEE ASSP Magazine, 1:4-29, April 1984.
- [2-15] Tokumichi Murakami, Kohtaro Asai, Atsushi Itoh, " Design Method of Symmetric Vector Quantiser", Electronics Letters, Vol.22, No.16, pp.848-849, 1986.
- [2-16] Tokumichi Murakami, Kohtaro Asai, Atsushi Itoh, "Vector Quantization of Color Images", IEEE ICASSP 86, 4.2.1, 1986.
- [2-17] Masahiro Wada, Hirohisa Yamaguchi and Hideo Yamamoto, "64kb/s Integrated Visual Telecommunication System," 2nd International Conference on Image Processing and its Application, 1986.
- [2-18] N. M. Nasrabadi and R. A. King, "Image Coding using vector Quantization: A Review," IEEE Trans. Comm., Vol. COM-36, pp. 957-971, August 1988.
- [2-19] J. H. Conway and N. J. A. Sloan, "Sphere Packings, Lattices and Groups," Springer-Verlag, New York 1988.

- [2-20] Philip A. Chou, Tom Lookabaugh, Robert M. Gray,
“Entropy-constrained vector quantization”, IEEE Trans.
ASSP, Vol. 37, pp.31-42, January 1989.
- [2-21] Philip A. Chou, Tom Lookabaugh, Robert M. Gray,
“Optimal Pruning with Applications to Tree-Structured
Source Coding and Modeling”, IEEE Trans. IT, Vol. 35, No. 2,
pp.299-315, March 1989.
- [2-22] Jim Brown, “PictureTel, VideoTelecom bridges out,”
Network World, Vol. 6, Np. 19, p. 25, May 1989.
- [2-23] William H. Equitz, “A new vector quantization clustering
algorithm”, IEEE Trans. ASSP, pp.1568-1575, October 1989.
- [2-24] Allen Gersho and Robert M. Gray, “Vector Quantization
and Signal Compression,” Kluwer Academic Publishers, 1992.
- [2-25] 山田芳郎, 藤田, 田崎三郎, “白黒静止画のブロック量子化に
対する一考察,” 信学技報 CS79-235, pp. 9-12, 1980年2月.
- [2-26] 山田芳郎, 正木一雄, 田崎三郎 “分散の異なる幾つかの情報
源群に対するベクトル量子化方式”, テレビジョン学会技術報
告, 5(32), pp.97-102, 1982年1月.
- [2-27] 村上篤道, 浅井光太郎, 太田堯久, “画像信号のベクトル量子
化器”, 信学技報, IE82-39, pp.25-32, 1982年7月.
- [2-28] 村上篤道, 浅井光太郎, 山崎英蔵, “高速ベクトル量子化器の
一構成法と符号化性能”, 信学技報, CS82-108, pp.17-24,
1983年.
- [2-29] 浅井光太郎, 村上篤道, 山崎英蔵, “画像ベクトル量子化コー
ドブックの一生成法”, 信学全大, 277, 1983年.
- [2-30] 村上篤道, 浅井光太郎, 山崎英蔵, “画像の基本構造パターン
の抽出と高能率符号化”, テレビ全大, 1983年.
- [2-31] 山田芳郎, 田崎三郎, “画像信号を対象としたベクトル量子
化器設計”, 信学論, J-B66, No. 8, pp.965-972, 1983年8月.

- [2-32] 田崎三郎, 山田芳郎, “ベクトル量子化”, 電子情報通信学会誌, Vol. 67, No. 5, pp.532-536, 1984年5月.
- [2-33] 村上篤道, 浅井光太郎, 山崎英蔵, “適応型ベクトル量子化方式フレーム間符号化”, 信学技報, IE84-1, pp.1-8, 1984年.
- [2-34] 村上篤道, 浅井光太郎, 山崎英蔵, “PR/DCパターンマッチング符号化”, 信学全大, 1984年.
- [2-35] 村上篤道, 浅井光太郎, “画像信号のベクトル量子化器,” テレビジョン学会誌, Vol. 38, No. 5, pp. 452-457, 1984年5月.
- [2-36] 村上篤道, 伊藤敦, 浅井光太郎, “画像のダイナミック多段ベクトル量子化”, 信学論, J-B68, No.1, pp.69-76, 1985年1月.
- [2-37] 山田芳郎, 田崎三郎, “逐次近似ベクトル量子化”, 信学論, Vol. J-A68, No. 5, pp.466-473, 1985年5月.
- [2-38] 斉藤隆弘, 武尾英哉, 相澤清晴, 原島博, 宮川洋, “画像信号の多段 Gain/Shape ベクトル量子化: 直交変換領域における設計とその性能評価”, テレビジョン学会技術報告, Vol. 9, No. 3, pp.57-64, 1985年5月.
- [2-39] 村上篤道, 浅井光太郎, 伊藤敦, “カラー画像のベクトル量子化手法”, テレビジョン学会誌, Vol.39, No.10, 1985年10月.
- [2-40] 山口博久, 和田正裕, 64kbit/s ビデオ通信のためのカラー画像符号化方式, テレビジョン学会誌, VOL.39, No.10, 1985年
- [2-41] 相澤清晴, 原島博, 宮川洋, “離散コサイン変換ベクトル量子化”, テレビジョン学会誌, Vol. 39, No. 10, pp.920-925, 1985年10月.
- [2-42] 浅井光太郎, 村上篤道, 丹野興一, 高野広志, “64~384Kbps カラー動画コーデック MVC-3000 の開発”, 信学全大, 1305, 1987年3月.
- [2-43] 和田正裕, 山口博久, “64kbps 統合ビデオ通信におけるマルチメディア伝送方式,”信学論, J70-B, No.4, 1987年4月.
- [2-44] 渡辺裕, 八島由幸, “主成分分析を用いた2進木探索コードブックの設計法”, 信学技報, IE87-109, 1988年.

- [2-45] 渡辺裕, 星野, 黒田英夫, 橋本秀雄, “適応的木探索ベクトル量子化を用いた 64kbit/s ビデオ符号化装置”, 信学技報, IE87-120, 1988年2月.
- [2-46] http://www.hct.ecl.ntt.co.jp/panel/pdf/3_Hd_3_2-4.pdf
- [2-47] Leo Breiman, Jerome H. Friedman, Richard A. Olshen, and Charles J. Stone, “Classification and Regression Trees,” Wadsworth Statistics/Probability Series. Belmont, 1984.
- [2-48] John O. Limb, Charles B. Rubinstein, "Plateau Coding of the Chrominance Component of Color Picture," Proc. IEEE, 60, pp. 809-820, June 1972.

第3章の参考文献

- [3-1] R. E. Chrocherie, S. A. Webber and J. L. Flanagan, “Digital coding of speech in sub-bands,” Bell System Technical Journal, Vol. 55, pp. 1069-1085, October 1976.
- [3-2] J. D. Johnston, “A filter family designed for use in quadrature mirror filter banks,” Proc. ICASSP 1980, pp. 291-294, April 1980.
- [3-3] P. J. Burt and E. H. Adelson, “The Laplacian pyramid as a compact image code,” IEEE Trans. Communication, Vol. COM-31, pp. 552-540, April 1983.
- [3-4] Martin Vetterli, “Filter Banks Allowing Perfect Reconstruction,” Signal Processing, Vol. 10, No. 3, pp. 219-244, April 1986.
- [3-5] J. W. Woods and S. O’Neil, “Sub-band Coding of Digital Images using Two-dimensional Quadrature Mirror Filtering,” Proc. ICASSP 86, pp. 1005-1008, April 1986.
- [3-6] E. H. Adelson, E. Simoncelli, and R. Hingorani. Orthogonal pyramid transforms for image coding. In Proc. SPIE, volume 845, pages 50–58, 1987.

- [3-7] Didier Le Gall and Ali Tabatabai, "Sub-band Coding of Digital Images," Proc. ICASSP, pp. 761-764, April 1988.
- [3-8] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers," IEEE Trans. Acoust., Speech, and Signal Proc., 36(9):1445-1453, September 1988.
- [3-9] R. Coifman, Y. Meyer, S. Quake and V. Wickerhauser, "Signal processing and compression with wavelet packets," Technical Report of Numerical Algorithms Research Group, Yale University, 1990.
- [3-10] R. Coifman and V. Wickerhauser, "Entropy-based algorithms for best basis selection," IEEE Trans. Inform. Theory, vol. 38, pp. 713-718, March 1992.
- [3-11] J. M. Shapiro, "An embedded wavelet hierarchical image coder," In Proc. IEEE ICASSP, pp. 657-660, March 1992.
- [3-12] J. W. Woods and T. Naveen, "A filter based bit allocation scheme for subband compression of HDTV," IEEE Trans. on IP, 1:436-440, July 1992.
- [3-13] Y. Huang, H. M. Dreizen and N. P. Galatsanos, "Prioritized DCT for Compression and Progressive Transmission of Images," IEEE Trans. IP, Vol. 1, pp. 477-487, October 1992.
- [3-14] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," IEEE Trans. on Signal Proc., 41(12): pp. 3445- 3462, December 1993.
- [3-15] Kannan Ramchandran and Martin Vetterli, "Best wavelet packet bases in a rate-distortion sense," IEEE Trans. Image Proc., 2(2):160-175, April 1993.
- [3-16] Cormac Herley, Jelena Kovacevic, Kannan Ramchandran and Martin Vetterli, "Tilings of the Time-frequency Plane:

- construction of arbitrary bases and fast tiling algorithms,”
IEEE Trans. SP, Vol. 41, No. 12, pp. 3341-3359, 1993.
- [3-17] Kohtaro Asai, Kannan Ramchandran, and Martin Vetterli,
“Image representation using time-varying wavelet packets,
spatial segmentation and quantization,” Proc. of Conf. on Inf.
Science and Systems, March 1993.
- [3-18] Kannan Ramchandran and Martin Vetterli,
“Rate-distortion optimal fast thresholding with complete
JPEG/MPEG decoder compatibility,” IEEE Trans. IP, Vol. 3,
No. 5, pp. 700-704, 1994.
- [3-19] Zixiang Xiong, Kannan Ramchandran, and Michael
Orchard, “Joint optimization of scalar and tree-structured
quantization of wavelet image decomposition,” in *Proc.
Asilomar Conf*, Pacific Grove, CA, vol. 2, pp.891-895,
November 1993.
- [3-20] Zixiang Xiong, Kannan Ramchandran, Michael Orchard
and Kohtaro Asai, “Wavelet Packets-based Image Coding
using Joint Space-Frequency Quantization,” in *Proc. ICIP*,
vol.III, pp.324-328, November 1994.
- [3-21] Kannan Ramchandran, Zixiang Xiong, Kohtaro Asai, and
Martin Vetterli, “Adaptive Transforms for Image Coding
Using Spatially Varying Wavelet Packets”, IEEE Trans. IP,
Vol. 5, No. 7, July 1996.
- [3-22] Martin Vetterli, Jelena Kovacevic, “Wavelets and
Subband Coding”, Prentice Hall, 1995.
- [3-23] John O. Drewery, “The Zone Plate as a Television Test
Pattern”, SMPTE Journal, Vol. 88, No. 11, November 1979.
- [3-24] 菊池久和, 中静真, “ウェーブレット変換による異常検知～
広がるウェーブレット変換の世界～,” シミュレーション第 16
巻第 4 号, 1997 年 12 月.

- [3-25] 太田睦, 西谷隆夫, “ウェーブレット変換について,” 信学技報 IE92-7, 1992年4月.
- [3-26] 渡邊慎治, 小松隆, 斎藤隆弘, “ウェーブレット変換領域における零交差表現と信号再構成” 信学論, J79-A(12), 2014-2022, 1996年12月.
- [3-27] 浅井光太郎, 村上篤道, “領域分割とウェーブレットパケットの適応化による画像符号化”, 信学全大, D-315, 1994年.
- [3-28] Kohtaro Asai, “Best Wavelet Packet Basis in Energy Localization Property,” Progress Report in Image Lab., Columbia University, 2013 (研究室内報告で未発表、参考のみ).
- [3-29] Stephane Mallat, “Zero-crossings of a wavelet transform,” IEEE Trans. IT, Vol. 37, No. 4, pp. 1019-1033, July 1991.

第4章の参考文献

- [4-1] 安田浩 “マルチメディア符号化の国際標準,” 丸善, 1991年.
- [4-2] 大久保榮, “テレビ会議教科書,” VTV ジャパン, 2017年4月.
<https://vcbook.vtv.co.jp/pages/viewpage.action?pageId=720963>
- [4-3] 小野文孝, 渡辺裕, “国際標準画像符号化の基礎技術,” コロナ社, 1998年2月.
- [4-4] 浅井光太郎, 守谷健弘, “映像音声符号化技術と国際標準化,” 情報処理学会デジタルプラクティス 1(2), pp.77-84, 2010年4月.
- [4-5] Sakae Okubo, “Reference Model methodology – A tool for the collaborative creation of video coding standards”, Proc. IEEE, Vol. 83, No. 2, 1995.
- [4-6] Kohtaro Asai, “Trends in Standardization of Multimedia Coding,” Keynote Speech, IEEE GCCE, October 2016.
- [4-7] ITU-T Recommendation H.120 “Codecs for videoconferencing using primary digital group transmission”, 1984.

- [4-8] ITU-T Recommendation H.261 “Video codec for audiovisual services at px64kbits”, 1990.
- [4-9] ISO/IEC 11172-2 (MPEG-1 Video) “Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s -- Part 2: Video,” 1993.
- [4-10] ISO/IEC 13818-2 (MPEG-2 Video) | ITU-T Recommendation H.262 “Generic coding of moving pictures and associated audio information -- Part 2: Video,” 1994.
- [4-11] ITU-T Recommendation H.263 “Video coding for low bit rate communication,” 1996.
- [4-12] ISO/IEC 14496-2 (MPEG-4 Visual) “Coding of audio-visual objects -- Part 2: Visual,” 1999.
- [4-13] ISO/IEC 14496-10 (AVC) | ITU-T Recommendation H.264, “Coding of audio-visual objects -- Part 10: Advanced Video Coding,” 2004.
- [4-14] ISO/IEC 23008-2 (HEVC) | ITU-T Recommendation H.265, “High efficiency coding and media delivery in heterogeneous environments -- Part 2: High efficiency video coding,” 2013.
- [4-15] ISO/IEC 10918-1 (JPEG: Digital compression and coding of continuous-tone still images -- Part 1: Requirements and guidelines) | ITU-T Recommendation T.81, 1992.
- [4-16] ISO/IEC 15444-1 (JPEG 2000 image coding system -- Part 1: Core coding system) | ITU-T Recommendation T.800, 2002.
- [4-17] ISO/IEC 29199-2 (JPEG XR image coding system -- Part 2: Image coding specification), 2011.
- [4-18] 浅井光太郎, 村上篤道, “画像の符号化伝送装置,” 特許公報, 特公平 8-28872, 1996年 (出願 1989年 2月) .
- [4-19] Kohtaro Asai, Hirofumi Nishikawa, Yoshihisa Yamada, Shin-ichi Hattori, Tokumichi Murakami, "Rate and Quality

- Control for Full-field Video Compression around 10Mbps",
IEEE GLOBECOM, pp.1081-1085, December 1992.
- [4-20] Takahiro Fukuhara, Kohtaro Asai, Tokumichi Murakami,
"Very Low Bitrate Coding with Block-Partitioning and
Adaptive Selection of Two-Differential Frame Memories",
IEEE Transactions on Circuits & Systems for Video
Technologies, Vol. 7, No. 1, pp. 212-220, 1997.
- [4-21] Shun-ichi Sekiguchi, Yoshihisa Yamada, Kohtaro Asai,
"A New Syntax for Error Resilient Transmission of CABAC
Stream," JVT-C098, May 2002.
- [4-22] "Joint Call for Proposals on Video Compression
technology," ISO/IEC JTC 1/SC 29/WG 11 N11113, January
2010.
- [4-23] Sei Naito and Atsushi Koike, "Efficient coding scheme for
super high definition video based on extending H.264 high
profile," Proc. VCIP 2006, January 2006.
- [4-24] Yusuke Itani, Shun-ichi Sekiguchi and Yoshihisa Yamada,
"Adaptive direct vector derivation for video coding," Proc.
PCS 2010, December 2010.
- [4-25] T. Watanabe, N. Wada, G. Yasuda, A. Tanizawa, T. Chujoh,
and T. Yamakage, "In-loop filter using block-based filter
control for video coding," in Proc. ICIP'09, Nov. 2009.
- [4-26] Kazuo Sugimoto, Yusuke Itani, Yoshimi Isu, Norimichi
Hiwasa and Shun-ichi Sekiguchi, "Description of video coding
technology proposal by Mitsubishi Electric," JCTVC-A107,
April 2010.
- [4-27] Ken McCann, W.-J. Han and I.- K. Kim, "Samsung's
response to the call for proposals on video compression
technoplogy," JCTVC-A124, April 2010.
- [4-28] "Test model under consideration," JCTVC-B205, July

2010.

- [4-29] Thomas Wiegand, W.-J. Han, Jens Reiner Ohm and Gary J. Sullivan, "High Efficiency Video Coding text specification working draft 1," JCTVC-C122, October 2010.
- [4-30] Kohtaro Asai, Tokumichi Murakami, Shuichi Yamagishi, Akira Minezawa, Yusuke Itani, Kazuo Sugimoto, Shun-ichi Sekiguchi, Yamada, Yoshiaki Kato, "New Video Coding Scheme Optimized for High-Resolution Video Sources", IEEE Journal of Selected Topics in Signal Processing, Vol. 5, No. 7, pp. 1290-1297, November 2011.

第 5 章の参考文献

- [5-1] Marc Antonini, Michael Barlaud, Pierre Mathieu, Ingrid Daubechies, "Image Coding Using Wavelet Transform," IEEE Trans. IP, Vol. 1, No. 2, pp. 205-220, April 1992.
- [5-2] Taejeong Kim, "Side Match and Overlap Match Vector Quantizers for Images," IEEE Trans. IP, Vol. 1, No. 2, pp. 170-185, April 1992.
- [5-3] 斎藤隆弘, 小松隆, 原島博, "画像信号の自己組織型リストを用いるベクトル量子化法," テレビジョン学会誌, Vol. 43, No. 10, pp.1020-1027, 1989年10月.
- [5-4] 長谷川由里, 浅井光太郎, 村上篤道, "木探索ベクトル量子化の自己組織化に関する一検討," PCSJ 89, 7-6, 1989年10月.
- [5-5] 山根延元, 森川良孝, 林一久, 藤原友紀, 松本和洋, "超球面上一様分布のための鏡映群に基づくベクトル量子化器," 信学論, J88-A, No. 11, pp. 1334-1342, 2005年11月.
- [5-6] Yushin Cho, David M. Barr, Jean-Marc Valin and Timothy B. Terriberry, "Applying Perceptual Vector Quantization outside of DAALA," Xiph.Org Foundation.
<https://people.xiph.org/~tterribe/daala/daala-icip2017.pdf>

- [5-7] Li-Yi Wei, marc Levoy, “Fast Texture Synthesis using Tree-structured Vector Quantization,” Proc. of SIGGRAPH 00, pp. 479-488, July 2000.
- [5-8] Michael Unser, “Texture Classification and Segmentation using Wavelet Frames,” IEEE Trans. IP, Vol. 4, No. 11, pp. 1549-1560, November 1995.
- [5-9] Eero P. Simoncelli, Javier Portilla, “Texture Characterization via Joint Statistics of Wavelet Coefficient Magnitudes,” IEEE ICIP 98, Vol. I, pp. 62-66, October 1998.
- [5-10] Soo Chang Kim, Tae Jin Kang, “Texture Classification and Segmentation using Wavelet Packet Frame and Gaussian Mixture Model,” The Journal of Pattern Recognition, Vol. 40, pp. 1207-1221, September 2006.
- [5-11] Minh N. Do, Martin Vetterli, “The Finite Ridgelet Transform for Image Representation,” IEEE Trans. IP, Vol. 12, No. 1, pp. 16-28, January 2003.
- [5-12] J. L. Starck. E. J. Candès, D. L. Donoho, “The Curvelet Transform for Image Denoising,” IEEE Trans. IP, Vol. 11, No. 6, pp. 670-684, June 2002.
- [5-13] Minh N. Do, Martin Vetterli, “The Contourlet Transform: An Efficient Directional Multiresolution Image Representation,” IEEE Trans. IP, Vol. 14, No. 12, December 2005.
- [5-14] 浅井光太郎, “マルチメディア符号化技術の標準化動向 ～ ISO/IEC JTC 1/SC 29 の活動から ～”, 電子情報通信学会通信方式研究会特別招待講演, 2016年12月.
- [5-15] 河村圭, “HEVCを超える次世代動画像符号化方式の標準化動向”, 情報処理学会短期集中セミナー「画像・音声符号化伝送技術～標準化動向最前線および開発事例～」, 2016年11月.

- [5-16] 浅井光太郎, “リコンフィギュラブル・ビデオ・コーディング”, 半導体理工学研究センター, 2007年11月.
- [5-17] ISO/IEC 23001-4: Codec Configuration representation, August 2014.
- [5-18] ISO/IEC 23002-4: Video tool library, April 2014
- [5-19] 原島博, “知的画像符号化と知的通信”, テレビジョン学会誌 42-6, pp.519-525, 1988年6月.
- [5-20] 原島博, 相澤清晴, 斎藤隆弘 “次世代画像符号化の構想-分析符号化から知的符号化へ-”, 信学技報, IE87-1, 1987年4月
- [5-21] J. W. A. Wang and E. H. Adelson, “Representing Moving Images with Layers,” IEEE Trans. Image Process., 3, 5, pp.625-638, Sept. 1994.
- [5-22] 斎藤隆弘, 相澤清晴, “ビデオデータ圧縮の新しい展開”, テレビジョン学会誌 49-4, pp.523-534, 1995年4月.
- [5-23] H. G. Musmann, M. Hotter and Joern Ostermann, “Object Oriented Analysis-Synthesis Coding of Moving Images,” Signal Process.: Image Communication, Vol. 1, No. 2, pp. 117-138, October 1989.
- [5-24] Takahiro Fukuhara, Kohtaro Asai, Tokumichi Murakami, "Model-based Image Coding using Stereoscopic Images and Hierarchical Structuring of New 3-D Wire-Frame Model", Picture Coding Symposium, 1991.
- [5-25] Rob Koenen, “MPEG-4 Overview (Tokyo version)”, ISO/IEC JTC 1/SC 29/WG 11/N2196, March 1998.
- [5-26] Samit Sarkar, “Rogue One filmmakers explain how they digitally recreated two characters,” Polygon.com, December 2016.
- [5-27] A. Krizhevsky, I. Sutskever, G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Neural Information Processing Systems 25*, 2012.

- [5-28] Mathew D. Zeiler, Rob Fergus, “Visualizing and Understanding Convolutional Networks,” arXiv:1312.2901v3, November 2013.
- [5-29] Donglai Wei, Bolei Zhou, Antonio Torralba, William T. Freeman, “mNeuron: A Matlab Plugin to Visualize Neurons from Deep Models,” MIT.
http://vision03.csail.mit.edu/cnn_art/index.html#v_single

主要研究業績

学術論文（海外誌掲載）

- (1) Tokumichi Murakami, Kohtaro Asai, Eizo Yamazaki, "Vector Quantiser of Video Signals," *Electronics Letters*, Vol.18, No.23, pp.1005-1006, September 1982.
- (2) Tokumichi Murakami, Kohtaro Asai, Atsushi Itoh, "Design Method of Symmetric Vector Quantiser," *Electronics Letters*, Vol.22, No.16, pp.848-849, July 1986.
- (3) Kannan Ramchandran, Zixiang Xiong, Kohtaro Asai, Martin Vetterli, "Adaptive transforms for image coding using spatially varying wavelet packets," *IEEE Transactions on Image Processing*, Vol. 5, No. 7, pp. 1197-1204, July 1996.
- (4) Takahiro Fukuhara, Kohtaro Asai, Tokumichi Murakami, "Very Low Bitrate Coding with Block-Partitioning and Adaptive Selection of Two-Differential Frame Memories," *IEEE Transactions on Circuits & Systems for Video Technologies*, Vol. 7, No. 1, pp. 212-220, February 1997.
- (5) Ajay Divakaran, Anthony Vetro, Kohtaro Asai, Hirofumi Nishikawa, "Video browsing system based on compressed domain feature extraction," *IEEE Transactions on Consumer Electronics*, Vol. 46, No. 3, pp. 637-644, August 2000.
- (6) Kohtaro Asai, Tokumichi Murakami, Shuichi Yamagishi, Akira Minezawa, Yusuke Itani, Kazuo Sugimoto, Shun-ichi Sekiguchi, Yoshihisa Yamada, Yoshiaki Kato, "New Video Coding Scheme Optimized for High-Resolution Video Sources," *IEEE Journal of Selected Topics in Signal Processing*, Vol. 5, No. 7, pp. 1290-1297, November 2011.

学術論文（国内誌掲載）

- (1) 村上篤道, 浅井光太郎, "画像信号のベクトル量子化器," *テレビ*

ジョン学会誌, Vol.38, No.5, 1984年4月.

- (2) 村上篤道, 伊藤敦, 浅井光太郎, "画像のダイナミック多段ベクトル量子化," 信学論誌, J68-B, No.1, pp.69-76, 1985年1月.
- (3) 村上篤道, 浅井光太郎, 伊藤敦 "カラー画像のベクトル量子化手法," テレビジョン学会誌, Vol.39, No.10, 1985年10月.
- (4) 浅井光太郎, 守谷健弘, "映像音声符号化技術と国際標準化," 情報処理学会デジタルプラクティス, 1(2), pp.77-84, 2010年.
- (5) 泉井良夫, 渋谷昭宏, 浅井光太郎, "スマートグリッドとセンサネットワーク," 信学論誌, J95-B, No.11, pp.1378-1387, 2012年11月.

国際会議

- (1) Tokumichi Murakami, Kohtaro Asai, Eizo Yamazaki, "Interframe coding using Adaptive Vector Quantizer," Picture Coding Symposium, Mar. 1983.
- (2) Tokumichi Murakami, Kohtaro Asai, Atsushi Itoh, "Vector Quantization of Color Images," IEEE ICASSP 86, 4.2.1, April 1986.
- (3) Tokumichi Murakami, Kohtaro Asai, Atsushi Itoh, "64Kbps Motion Video Codec using Vector Quantizer," Picture Coding Symposium, June 1987.
- (4) Tokumichi Murakami, Kohtaro Asai, Atsushi Itoh, "Progressive Vector Quantization of Color Images," IEEE GLOBECOM, 35-8, November 1987.
- (5) Tokumichi Murakami, Kohtaro Asai, Masami Nishida "64Kbps Scene Adaptive Vector Coding of Motion Picture," Int'l workshop on 64kbps coding of moving video, June 1988.
- (6) Tokumichi Murakami, Kohtaro Asai, Koh Kamizawa, "1x8kbps Color Motion Picture Transmission using Vector Quantization," IEEE ISIT 88, pp.204-205, June 1988.

- (7) Tokumichi Murakami, Kohtaro Asai, Koh Kamizawa,
"A Feasibility Study of 1x8kbps Videophone Codec," Picture
Coding Symposium, September 1988.
- (8) Tokumichi Murakami, Kohtaro Asai, Yoshiaki Kato, Kazuhiro
Matsuzaki, "Scene Adaptive Vector Quantization for Image
Coding," IEEE GLOBECOM, 33-2, December 1988.
- (9) Takahiro Fukuhara, Kohtaro Asai, Tokumichi Murakami,
"Hierarchical Division of 3-D Wire-Frame Model and Vector
Quantization in a Model-Based Coding of Facial Image,"
Picture Coding Symposium, March 1990.
- (10) Takahiro Fukuhara, Kohtaro Asai, Tokumichi Murakami,
"Model-based Image Coding using Stereoscopic Images and
Hierarchical Structuring of New 3-D Wire-Frame Model,"
Picture Coding Symposium, September 1991.
- (11) Kohtaro Asai, Hirofumi Nishikawa, Yoshihisa Yamada,
Shin-ichi Hattori, Tokumichi Murakami, "Rate and Quality
Control for Full-field Video Compression around 10Mbps,"
IEEE GLOBECOM, pp.1081-1085, December 1992.
- (12) Kohtaro Asai, Kannan Ramchandran, Martin Vetterli,
"Image Representation using Time-varying Wavelet Packets,
Spatial Segmentation and Quantization," Conference on
Information Sciences and Systems, pp.761-766, March 1993.
- (13) Kohtaro Asai, Nguyen T. Thao, Martin Vetterli, "A Study
of Convex Coders with an Application to Image Coding," IEEE
ICASSP 94, V581-584, April 1984.
- (14) Nguyen T. Thao, Kohtaro Asai, Martin Vetterli, "Set
Theoretic Compression with an Application to Image Coding,"
IEEE ICIP 94, pp.336-340, November 1994.
- (15) Zixiang Xiong, Kannan Ramchandran, Michael T. Orchard,
Kohtaro Asai, "Wavelet Packets-based Image Coding using

- Joint Space-Frequency Quantization,” IEEE ICIP 94, pp.324-328, November 1994.
- (16) Kohtaro Asai, Yoshihisa Yamada, Tokumichi Murakami, “Video Sequence Coding Based on Segment-Model and Priority Control,” Picture Coding Symposium, September 1994.
- (17) Takahiro Fukuhara, Kohtaro Asai, Tokumichi Murakami, “A Study of Segment-based MC as a Tool of Object-based Image Coding,” International Workshop on Coding Technologies for Very Low Bitrate Video”, November 1995.
- (18) Shun-ichi Sekiguchi, Takahiro Fukuhara, Kohtaro Asai, Tokumichi Murakami, “A Study on Very Low BitRate Video Coding with Segment-based Motion Compensation,” Picture Coding Symposium, March 1996.
- (19) Shun-ichi Sekiguchi, Takahiro Fukuhara, Kohtaro Asai, Tokumichi Murakami, “Region Adaptive Video Coding with Variable-sized Blocks and VPDL,” International Symposium on Information Theory and its Application, September 1996.
- (20) Kohtaro Asai, Shun-ichi Sekiguchi, Yoshimi Isu, Hirofumi Nishikawa, Shin-ichi Kuroda, “A Study on Region-based Video Coding with Optimal Segmentation,” Picture Coding Symposium, September 1997.
- (21) Tokumichi Murakami, Kohtaro Asai, Hideo Ohira, “Compression and Packetization for MPEG-2/4 Video and Audio,” EUSIPCO 98, September 1998.
- (22) Ajay Divakaran, Kohtaro Asai, Hirofumi Nishikawa, “A description scheme for video based on feature extraction in the compressed domain,” IEEE ICCE, pp. 278-279, June 2000.
- (23) Miroslaw Bober, Kohtaro Asai, Ajay Divakaran, “MPEG-4/7 Based Internet Video and Still Image Browsing

System,” SPIE Conference on Multimedia Systems and Applications, Vol. 4209, December 2000.

- (24) Kohtaro Asai, “Status and trends in MPEG multimedia technologies,” Japan-America Frontiers of Engineering Symposium, August 2003.
- (25) Shun-ichi Sekiguchi, Yuichi Idehara, Kazuo Sugimoto, Kohtaro Asai, “A low-cost video frame-rate up conversion using compressed-domain information,” IEEE ICIP, Vol.2, pp.974-7, September 2005.
- (26) Shun-ichi Sekiguchi, Yoshimi Isu, Kazuo Sugimoto, Yoshihisa Yamada, Kohtaro Asai, Tokumichi Murakami, “On Separate Color-Plane Prediction for Direct 4:4:4 Video Coding,” IEEE ICIP, pp.1717-1720, October 2006.
- (27) Kohtaro Asai, “Beyond H.264/MPEG-4 AVC,” Chair of Panel, PCS, December 2010.
- (28) Kohtaro Asai, “Trends in Standardization of Multimedia Coding,” Keynote Speech, IEEE Global Conference on Consumer Electronics, October 2016.

研究会・シンポジウム

- (1)谷中一寿, 浅井光太郎, 齐藤忠夫, 猪瀬博, “2値・濃淡混合画像の帯域圧縮符号化,” テレビ学技報, Vol. 4, No. 44, pp.49-54, 1981年3月.
- (2)浅井光太郎, 村上篤道, 杉山康夫, “サブナイキストサンプリングにおける予測符号化方式,” 信学技報, IE82-4, pp.25-30, 1982年
- (3)村上篤道, 浅井光太郎, “最大傾斜法による画像の適応型予測符号化,” 情報理論とその応用研究会, 1982年.
- (4)村上篤道, 浅井光太郎, 太田堯久, “画像信号のベクトル量子化器,” 信学技報, IE82-39, pp.25-32, 1982年7月.
- (5)村上篤道, 浅井光太郎, 山崎英蔵, “高速ベクトル量子化器の一構

- 成法と符号化性能,” 信学技報, CS82-108, pp.17-24, 1983 年.
- (6) 村上篤道, 浅井光太郎, 山崎英蔵, ”適応型ベクトル量子化方式フレーム間符号化,” 信学技報, IE84-1, pp.1-8, 1984 年.
- (7) 村上篤道, 伊藤敦, 浅井光太郎 “画像のダイナミック多段ベクトル量子化,” 信学技報, IE84-18, pp.9-16, 1984 年.
- (8) 村上篤道, 浅井光太郎, 山崎英蔵, “ベクトル量子化による画像高能率符号化,” 信学技報, IT85-61, pp.77-82, 1985 年.
- (9) 浅井光太郎, 上澤功, 村上篤道, “ビデオコーデックの LSI 化, DSP 化に関する一検討,” PCSJ 87, 3-1, 1987 年 10 月.
- (10) 松崎一博, 浅井光太郎, 西田正実, 村上篤道, “シーン・アダプティブ・ベクトル量子化を用いた静止画符号化,” PCSJ 88, 2-5, 1988 年 10 月.
- (11) 長谷川由里, 浅井光太郎, 村上篤道, “木探索ベクトル量子化の自己組織化に関する一検討,” PCSJ 89, 7-6, 1989 年 10 月.
- (12) 福原隆浩, 浅井光太郎, 村上篤道, “顔画像の輝度情報を用いた 3 角形パッチの階層的分割とパッチのベクトル量子化についての検討,” PCSJ 89, 8-4, 1989 年 10 月.
- (13) 西田正実, 加藤嘉明, 浅井光太郎, 山岸秀一, 村上篤道, “テレビ会議・テレビ電話用映像伝送方式の検討,” 情報理論とその応用シンポジウム, 1989 年.
- (14) 西川博文, 浅井光太郎, 村上篤道, “サブバンドフィルタ特性にタップ数/演算精度が与える影響と直交変換との比較検討”, PCSJ 90, 1-5, 1990 年 10 月.
- (15) 福原隆浩, 浅井光太郎, 村上篤道, “分析合成符号化方式における 3 次元構造モデルの自動生成とその階層構造化についての検討,” PCSJ 90, 3-3, 1990 年 10 月.
- (16) 西川博文, 浅井光太郎, 村上篤道, “Quadrature Mirror Filter の特性と直交変換との比較について,” 情報理論とその応用シンポジウム, 1990 年.
- (17) 村上篤道, 浅井光太郎, 西川博文, 山田悦久, “適応的ブロッ

キングによるインターレース画像の高能率符号化", 情報理論とその応用研究会, 1991年.

- (18) 西川博文, 加藤嘉明, 西田正実, 浅井光太郎, "サブバンド分割を用いた階層的動画符号化方式の検討," PCSJ 91, 7-3, 1991年10月.
- (19) 坂戸美朝, 佐々木源, 浅野研一, 浅井光太郎, 海老沢秀明, 吉田英夫, "ディジタル SNG システム用ビデオコーデック," テレビ学技報, Vol. 17, No. 13, 1993年2月.
- (20) 福原隆浩, 浅井光太郎, "オブジェクト指向を導入した動画符号化方式の一検討," PCSJ, 7-7, 1995年10月.
- (21) 黒田慎一, 西川博文, 浅井光太郎, "色の再現性を考慮した超高精細静止画像の MPEG-2 イントラ符号化," PCSJ, P-3.7, 1996年10月.
- (22) 関口俊一, 井須芳美, 浅井光太郎, "レート歪み最小化規範に基づく動画の領域分割符号化に関する一検討," PCSJ, P-1.7, 1997年10月.
- (23) 関口俊一, 西川博文, 井須芳美, 浅井光太郎, "低レートビデオ符号化における最適符号化に関する一検討," PCSJ, P-2.5, 1998年10月.
- (24) 山田悦久, 浅井光太郎, "GOP 構造とフレーム/フィールド構造による MPEG-2 符号化効率の改善," PCSJ, P-5.3, 1998年10月.
- (25) 山田悦久, 日和佐憲道, 浅井光太郎, "相関性の低い動画信号に対する符号化効率の検証," PCSJ, P-5.3, 1999年9月.
- (26) 山田悦久, 杉本和夫, 浅井光太郎, "復号画像から推定されたパラメータを利用する再符号化手法," PCSJ, P-P1.22, 2000年11月.
- (27) 杉本和夫, 井須芳美, 西川博文, 浅井光太郎, "画像特徴を利用した監視画像の蓄積再生に関する一検討," PCSJ, P-2.1, 2001年11月.

- (28) 山田悦久, 関口俊一, 浅井光太郎, “適応ブロックサイズを用いた動き補償予測の符号化効率検証,” PCSJ, P-4.1, 2002年11月.
- (29) 山岸秀一, 井須芳美, 関口俊一, 山田悦久, 浅井光太郎, 村上篤道, “4:4:4 映像信号に対する色成分独立符号化方式の検討,” 情報理論とその応用シンポジウム, Vol. 29, No. 2, pp.541-544, 2006年11月.
- (30) 浅井光太郎, “映像符号化技術の国際標準化,” 情報処理学会 DICOMO, 2010年7月.

全国大会

- (1) 浅井光太郎, 村上篤道, 太田堯久, “適応型 AR 予測器による DPCM 符号化特性,” 信学全大, 688, 1982年.
- (2) 浅井光太郎, 村上篤道, 山崎英蔵, “画像信号の高速ベクトル量子化,” 信学全大, S-18, 1983年.
- (3) 浅井光太郎, 村上篤道, 山崎英蔵, “画像ベクトル量子化コードブックの一生成法,” 信学全大, 277, 1983年.
- (4) 浅井光太郎, 村上篤道, 山崎英蔵, “カラー画像信号のベクトル量子化,” 信学全大, 1274, 1984年.
- (5) 浅井光太郎, 村上篤道, “平均値分離正規化ベクトル量子化における量子化歪の低減,” 信学全大, 1220, 1985年.
- (6) 浅井光太郎, 西田正実, 村上篤道, “低ビットレート動画コーデックの構成に関する一検討,” 信学全大, 1314, 1986年.
- (7) 浅井光太郎, 村上篤道, 丹野興一, 高野広志, “64~384Kbps カラー動画コーデック MVC-3000 の開発,” 信学全大, 1305, 1987年.
- (8) 浅井光太郎, 西田正実, 村上篤道, “9600bps ベクトル量子化方式カラー動画像伝送,” 信学全大, 383, 1987年.
- (9) 浅井光太郎, 松崎一博, 西田正実, 村上篤道, “シーン・アダプティブ・ベクトル量子化による低ビットレート動画像符号化方式,” 信学全大, D-93, 1989年.

- (10) 浅井光太郎, 西川博文, 村上篤道, “サブバンドベクトル量子化を用いた画像符号化方式,” 信学全大, D-276, 1990年.
- (11) 浅井光太郎, 村上篤道, “領域分割とウェーブレット・パケットの適応化による画像符号化,” 信学全大, D-315, 1994年.
- (12) 浅井光太郎, “MPEG-4 ビデオ検証モデル,” 映情冬季大, pp.33-38, 1996年12月.

解説論文・解説記事

- (1)浅井光太郎, “次世代放送のための符号化技術,” 映情学誌, Vol.51, No.9, pp.1381-1384, 1997年9月.
- (2)浅井光太郎, “マルチメディアを目指す MPEG-4 標準化動向,” 映情学誌, Vol.51, No.12, pp.1999-2003, 1997年12月.
- (3)浅井光太郎, 山田悦久, 小川文伸, “新技術の動向 MPEG 符号化技術,” 映像情報, Vol.30, No. 21, pp.14-21, 1998年11月.
- (4)浅井光太郎, “MPEG とは何か?,” 映情学誌, Vol.54, No.6, pp.814-815, 2000年6月.
- (5)浅井光太郎, 加藤嘉明, 岡進, “離陸する MPEG-4～映像検索・配信,” 映情学誌, Vol.55, No.4, pp.501-502, 2001年.
- (6)浅井光太郎, “放送・通信融合時代の MPEG,” ARIB 機関誌, No.32, 2003年.
- (7)浅井光太郎, 岡進, “ユビキタス社会におけるデジタル放送の役割,” 電気評論誌, 2003年6月.
- (8)浅井光太郎, “画像符号化技術の発達と活用,” 電気学会誌, Vol.125, No. 5, pp.296-299, 2004年4月.
- (9)浅井光太郎, “符号化圧縮技術の現状,” 映情学誌, Vol.63, No.12, pp.1757-1760, 2009年12月.
- (10) 浅井光太郎, “JTC 1/SC 29 における国際標準化,” 標準化と品質管理, Vol. 63, No.5, pp.84-88, 2010年6月.
- (11) 関口俊一, 浅井光太郎, “次世代動画像符号化 (HEVC) の標準化動向,” 映情学誌, Vol.65, No.3, pp.302-308, 2011年3月.

- (12) 小島洋之, 神田英伸, 浅井光太郎, 西川博文, “映像セキュリティシステムの動向,” 画像ラボ, 2017年4月.

講師・その他

- (1)浅井光太郎, “MPEG4 ビデオ検証モデル,” テレビジョン学会映像メディア部門招待講演, 1996年.
- (2)浅井光太郎, “次世代画像サービスのための符号化技術,” 信学会関西支部専門講習会, 1997年.
- (3)浅井光太郎, “見えてきた MPEG-4 の全貌～5 今後の課題,” 映像情報メディア学会講習会, 1997年.
- (4)浅井光太郎, “MPEG-4 の概要と要求条件,” 学術振興第165委員会, 1998年.
- (5)浅井光太郎, “MPEG-4 の概要と動向,” 情報処理学会連続セミナー, 1999年.
- (6)浅井光太郎, “情報符号化技術～映像信号の符号化～,” 信学会関西支部情報通信技術基礎講座, 2001年.
- (7)浅井光太郎, “MPEG-4 の実用化と標準化活動の状況,” 情報処理学会～標準化セッション, 2001年.
- (8)浅井光太郎, “画像符号化とネットワーク映像サービスへの期待,” 第14回情報伝送と信号処理ワークショップ, 2001年.
- (9)浅井光太郎, “MPEG 国際標準化の近況,” 映情学会, 2006年2月.
- (10) 浅井光太郎, 伊藤浩, 鈴木光義, 加藤嘉明, 村上篤道, “デジタル映像の電子透かし埋め込み技術,” 文部科学省科学技術振興事業デジタルシネマフェスティバル公開講座, 2006年11月.
- (11) 浅井光太郎, “リコンフィギュラブル・ビデオ・コーディング,” 半導体理工学研究センター, 2007年11月.
- (12) 浅井光太郎, “高品質映像サービスの可能性と新たな符号化技術への期待,” FIT, 2009年9月.

編著書

- (1) “ユビキタス技術 動画像の高能率符号化,” 共著（小野定康, 村上篤道, 浅井光太郎）, オーム社, 2005年.
- (2) “『知識の森』5群（通信・放送）8編 放送・CATV,” 編幹事, 電子情報通信学会, 2012年.
- (3) “高効率映像符号化技術 HEVC/H.265 とその応用,” 共編著（村上篤道, 浅井光太郎, 関口俊一） オーム社, 2013年.

表彰

テレビジョン学会丹羽高柳賞論文賞受賞, 1986年
電子情報通信学会篠原記念学術奨励賞受賞, 1989年
映像情報メディア学会技術振興賞進歩賞, 1997年
関東地方発明奨励賞受賞, 2001年（2005年, 2011年）
電子情報通信学会業績賞受賞, 2002年
情報処理学会情報規格調査会標準化貢献賞受賞, 2002年
映像情報メディア学会ハイビジョン技術賞受賞, 2002年
映像情報メディア学会著述賞受賞, 2006年
国際標準化貢献者表彰 産業技術環境局長表彰受賞, 2008年
電子情報通信学会フェロー称号, 2014年
工業標準化事業表彰・経済産業大臣表彰受賞, 2015年