



# 高信頼度高能率映像配信サービス実現のためのコンテンツハンドリング技術に関する研究

酒澤, 茂之

---

(Degree)

博士 (工学)

(Date of Degree)

2005-03-11

(Date of Publication)

2013-02-15

(Resource Type)

doctoral thesis

(Report Number)

乙2794

(URL)

<https://hdl.handle.net/20.500.14094/D2002794>

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



神戸大学博士論文

高信頼度高能率映像配信サービス実現のための  
コンテンツハンドリング技術に関する研究

平成 17年 3月

酒澤 茂之

## 内容梗概

インターネットの高速化・低コスト化を受けて多種多様な映像コンテンツの配信サービスが注目されている。本論文では、高信頼度で高能率な映像配信サービスの実現を目指して、コンテンツの制作、素材の収集、最終視聴者への配信の各過程におけるコンテンツハンドリング技術に関する研究を行う。コンテンツの制作においては、非常に関心が高まっている著作権保護の観点から、著作権者の情報をコンテンツ内に埋め込む電子透かし技術を提案する。素材の収集においては、ライブ中継等の実時間伝送への要求に対応して、高能率に多数のコンテンツを伝送する技術を提案する。また、コンテンツ制作のための編集素材などの実時間性が要求されないコンテンツファイル伝送について、様々なネットワーク環境に対応して高信頼度で高速にファイルを伝送する技術を提案する。そして、多様なネットワーク環境にある多数の受信者向けに高能率に映像を配信する技術を提案する。

まず、電子透かし技術に関して、映像コンテンツの流通が MPEG で符号化されたファイルにより行われることに注目し、電子透かしを MPEG ビットストリームから直接検出することのできる方式を提案する。提案方式では、電子透かしは原画像の局所ブロックの DCT 係数を操作することにより埋め込まれる。このとき、単一の局所ブロックに電子透かしを埋め込むのではなく、複数の画像フレーム上の様々な位置に属する局所ブロックを選択し、その集合に対して同一の DCT 係数操作を加えることにより電子透かしの埋め込みを行う。検出にあたっては、MPEG ビットストリームに含まれている DCT 係数の中で、上記の局所ブロック位置の DCT 係数を取り出し、その集合において DCT 係数に加えられた操作を統計的に検出する。この方式は、MPEG 復号が不要であるため極めて高速に埋め込まれている電子透かしを検出することができる。提案方式に対して MPEG 再符号化や電子透かし攻撃ツール StirMark に含まれるツールにより擬似的に攻撃を加えたときの電子透かしの耐性を評価し、再符号化や周波数領域における攻撃には強い耐性があることを示した。さらに、検出された電子透かしの信頼性に関する検討も行っており、統計的検定を応用することにより、本来存在しないはずの情報を読み出してしまう “false positive” 問題を避ける方法も示した。

次に、複数の映像コンテンツを高能率に実時間伝送する技術について、複数の映像エンコーダを統一的制御の下で動作させるマルチエンコーダシステムを提案する。このシステムでは、各エンコーダの出力を多重して固定速度の通信路に入力するため、エンコーダの符号化速度の総和を一定に保つ制約の下で、各エンコーダへの符号化速度配分を決定する必要がある。予備検討として MPEG 符号化で用いられる MC+DCT 符号化方式における符号化速度-歪み関係を解析し、単純なベキ乗関数で近似できることを示した。これに基づいて、歪みの総和を最小化する意味で最適な符号化速度配分方式を導いた。また、符号化速度配分を定期的に再計算す

ることで、映像シーンの切り替わりに追従して最適配分が更新される方式も提案した。計算機シミュレーションの結果、3台のエンコーダの多重のとき、符号化速度配分を行わない従来方式と比較して35%もの歪みの削減が達成されることを示した。また、従来方式では情報量の特に多い映像シーンで大きな画質劣化が発生していた問題点に関して、提案方式では当該シーンに高い符号化ビットレートを配分することにより画質劣化を回避することができることも示した。

一方、非実時間での映像伝送においては、伝送された映像をコンテンツ制作のための編集用素材として用いるために、符号化された映像ファイルを伝送エラーなしで高速に伝送することが求められる。また、受信途中の映像データへアクセスし、映像内容の確認を行うなど、映像コンテンツ特有の機能要求がある。それに対して、従来の伝送プロトコルはファイル伝送を目的として設計されていたため、上記の機能要求を満たすことができなかった。そこで、到達確認型プロトコルによる複数のコネクションを同時に用いて伝送速度を高速にするとともに、受信中のデータの整列済み領域と不連続領域を管理することで、整列領域からの映像再生を実現する手法を提案した。また、ネットワークの輻輳を発生させないために伝送ビットレートシェーピングが必要とされているが、複数のコネクションを利用する際に有効な送信側でのパケット送出制御手法を考案した。これらの提案手法に関して実験室内でのネットワークシミュレータ環境、ならびに日米間の実際のネットワークにおいてもその有効性を確認した。

最終視聴者へのネットワーク経由での映像配信方式に関しては、複数の受信者に対して一斉同報的にパケットを届けることのできるマルチキャスト技術が注目されている。マルチキャストでは、特別なマルチキャスト用のIPアドレスへ送出されたパケットは、それを受信するための登録をネットワークに対して行った受信端末にのみ配送される。このため、不必要なパケット配送が発生しない利点がある。しかし、単純なマルチキャスト技術では、受信者ごとに異なる利用可能帯域に適応させることができず、所定の固定速度での画一的な映像配信となってしまう、正常に受信できない受信者が発生する問題点があった。そこで、MPEG符号化された映像データを階層分割し、その階層別にマルチキャストする手法を考案した。これにより、受信者の利用可能帯域に応じて階層データのマルチキャストチャンネルを追加／削除することにより、マルチキャストの映像配信能率の高さを保ちつつ、各受信者に最適な映像配信を実現することが可能となる。特に、階層データ間の映像復号における依存関係に基づいて、受信側におけるバッファによる階層データ間の時間調整過程を詳細に理論解析し、具体的な初期バッファ値ならびに最大バッファ量の算出法を与えている。

以上のように本論文が映像配信サービスに果たす貢献は、まず、映像符号化方式と親和性の高い電子透かし技術によって映像コンテンツの著作権保護を行うことにある。ついで、複数の映像コンテンツの実時間伝送において、それぞれの映像シーンの情報量に応じて最適に符号化

ビットレートを配分することで実時間伝送を高能率に実現する点にある。また、映像コンテンツの非実時間伝送についても到達確認型プロトコルを複数コネクション同時に用いることで高信頼、高能率かつ高機能な非実時間伝送を実現することにある。最後に、最終視聴者向けに通信ネットワークを用いて受信環境に適応的に映像配信する手法を示した点である。これらのコンテンツハンドリング技術により、コンテンツの制作、伝送、配信という映像配信サービスの一連の過程を高信頼かつ高能率に実現することが可能となる。

# 目次

<b>第1章 緒論</b>	<b>1</b>
<b>第2章 電子透かし</b>	<b>5</b>
2.1 緒言	5
2.2 提案方式の概要	6
2.3 DCT係数の操作	6
2.4 2ステージ電子透かし	8
2.4.1 テンプレート	9
2.4.2 第一ステージ電子透かし	9
2.4.3 第二ステージ電子透かし	9
2.5 攻撃耐性の評価	11
2.5.1 計算機シミュレーション条件	11
2.5.2 検出結果	14
2.5.3 埋め込み情報量拡大時の評価	16
2.6 考察	18
2.6.1 透かし検出計算量	18
2.6.2 検出結果の信頼度	19
2.7 結言	21
<b>第3章 実時間映像伝送方式</b>	<b>24</b>
3.1 緒言	24
3.2 マルチエンコーダシステム	24
3.3 MC+DCTエンコーダのレート歪み特性	25
3.3.1 ステップサイズ対MSE関係	27
3.3.2 ステップサイズ対ビットレート関係	28
3.3.3 レート歪み特性	29

3.4	ビットレート配分制御方式 . . . . .	29
3.4.1	ビットレート配分式 . . . . .	29
3.4.2	ビットレート配分の更新 . . . . .	30
3.4.3	マルチエンコーダ利得 . . . . .	32
3.5	計算機シミュレーションによる評価 . . . . .	33
3.6	結言 . . . . .	36
<b>第4章</b>	<b>非実時間映像伝送方式</b>	<b>38</b>
4.1	緒言 . . . . .	38
4.2	映像伝送に対する要求条件 . . . . .	39
4.3	データ伝送の原理 . . . . .	39
4.4	プロトコル . . . . .	41
4.4.1	通信手順 . . . . .	41
4.4.2	アプリケーションレイヤ向け通信サービス . . . . .	42
4.4.3	伝送レートシェーピング . . . . .	43
4.5	アプリケーションの構成法 . . . . .	44
4.6	伝送特性 . . . . .	44
4.6.1	スループット特性の理論解析 . . . . .	44
4.6.2	実験結果 . . . . .	45
4.7	システム実装例 . . . . .	51
4.7.1	システム構成 . . . . .	51
4.7.2	フィールド実験結果 . . . . .	52
4.8	結言 . . . . .	53
<b>第5章</b>	<b>高能率映像配信方式</b>	<b>55</b>
5.1	緒言 . . . . .	55
5.2	マルチチャンネル伝送モデル . . . . .	55
5.3	ビデオ階層符号化方式 . . . . .	56

5.3.1	各種階層符号化方式の比較 . . . . .	56
5.3.2	本章で採用する階層符号化方式 . . . . .	57
5.3.3	階層データの伝送チャンネルへのマッピング . . . . .	58
5.4	システムの安定動作条件 . . . . .	59
5.4.1	解析の前提条件 . . . . .	59
5.4.2	各チャンネルの解析 . . . . .	59
5.4.3	伝送制御方式および遅延の解析 . . . . .	62
5.4.4	フレーム発生情報量とバッファサイズの理論解析 . . . . .	65
5.4.5	計算機シミュレーションによる評価 . . . . .	66
5.5	階層間レート配分 . . . . .	67
5.5.1	最適化の対象 . . . . .	67
5.5.2	階層間レート配分の実験的検討 . . . . .	67
5.5.3	受信画質 . . . . .	70
5.6	結言 . . . . .	70
<b>第6章</b>	<b>結 論</b>	<b>73</b>
	<b>謝 辞</b>	<b>75</b>
	<b>参考文献</b>	<b>77</b>
	<b>関連発表</b>	<b>82</b>

## 目 次

図 2.1	電子透かし埋め込み、圧縮、検出の流れ . . . . .	7
図 2.2	マクロブロックへの埋め込み (フレーム構造での表現) . . . . .	7
図 2.3	2 ステージ電子透かし . . . . .	8
図 2.4	第一ステージ電子透かしのためのテンプレートの例 . . . . .	10
図 2.5	第二ステージ電子透かしのテンプレートの例 . . . . .	10
図 2.6	第二ステージにおける 1 ビットの検出の過程 . . . . .	12
図 2.7	攻撃耐性の評価手順の流れ . . . . .	13
図 2.8	透かし埋め込み済み画像 “Susie” の第 10 フレーム . . . . .	14
図 2.9	第一ステージのグループ分類誤りが第二ステージ検出に与える影響 . . . . .	17
図 3.1	マルチエンコーダシステムの構成 . . . . .	25
図 3.2	量子化ステップサイズ対 MSE 関係 . . . . .	26
図 3.3	予測誤差 DCT 信号の性質 . . . . .	27
図 3.4	量子化ステップサイズ対ビットレート関係 . . . . .	30
図 3.5	ビットレート対 MSE 関係 . . . . .	31
図 3.6	マルチエンコーダ利得特性 . . . . .	33
図 3.7	ビットレート対 MSE 関係 . . . . .	34
図 3.8	ビットレート配分の時間変化 . . . . .	35
図 3.9	符号化歪み特性 . . . . .	35
図 4.1	マルチコネクションデータ伝送のブロック図 . . . . .	40
図 4.2	データパックフォーマット . . . . .	41

図 4.3	ログイン手順	41
図 4.4	プロトコルスタック	42
図 4.5	SVFTP における TCP コネクション	42
図 4.6	スループット特性	47
図 4.7	データ到着ジッタ	49
図 4.8	SVFTP の他トラヒックへの影響	50
図 4.9	蓄積伝送システム	53
図 4.10	無線通信路におけるスループット特性	53
図 4.11	国際回線上でのスループット特性	54
図 5.1	複数のチャンネルを用いたマルチキャストビデオ伝送モデル	56
図 5.2	Case 1、P チャンネルに対するバッファ遷移	60
図 5.3	Case 1、B チャンネルに対するバッファ遷移	61
図 5.4	Case 2、I+P チャンネルに対するバッファ遷移	62
図 5.5	各チャンネルでの復号タイミング	63
図 5.6	チャンネル間の時間関係	64
図 5.7	フレーム発生情報量累計の時間経過	66
図 5.8	階層間のレート配分実験結果	69
図 5.9	復号画像	70



## 表 目 次

表 2.1	埋め込み強度	13
表 2.2	適用される攻撃	13
表 2.3	画質	15
表 2.4	第一ステージ電子透かしによるグループ分類結果 (分類の正解率)	16
表 2.5	第二ステージの 4bit 透かし検出結果 (ビット判定の正解率)	17
表 2.6	電子透かし埋め込み後の PSNR(dB) (大容量情報埋め込みの場合)	18
表 2.7	第二ステージ電子透かしの検出結果 (80bit 埋め込み中の正解率)	18
表 2.8	t 検定結果	20
表 5.1	チャンネルへの階層データの割り当て	58
表 5.2	必要バッファ量と遅延	67

## 第1章 緒 論

映像メディアは音声やテキストなどの他のメディアと比較して情報量が非常に多く、人間にとって最も受容性の高いメディアであり、ニュース等の情報伝達や教育から娯楽に至るまで幅広い目的で用いられている。逆に、その情報量の多さのために伝送には広帯域、蓄積には大容量が必要とされる。それに対して、近年の映像高能率符号化技術やデジタル伝送技術の進歩により、映像の伝送に必要な帯域の削減と伝送路の速度の増大が進んでいる。これを受けて、放送の分野においては、デジタル衛星放送やデジタルCATVにおいて100チャンネル以上の多チャンネル放送が実用化されている。通信の分野においても、インターネットの基幹ネットワークおよびアクセスネットワークの高速化・低コスト化が急速に進んだことを背景として、多種多様な映像コンテンツがインターネットを用いて配信されている。

このような多チャンネル化の状況の下で、映像配信サービスに用いられる映像コンテンツへの需要が高まっている。ただし、今後の映像配信は従来の大人数の視聴者への画一的放送ではなく、比較的少人数の視聴者に対する多種多様な内容の配信になると想定される。従って、個々の映像配信に要するコストの低廉化が必要であり、映像コンテンツの取得から配信に至るシステム全体の能率の高さが要求されている。また、近年、コンテンツのデジタル化とインターネットの高速化を背景として、違法コンテンツ流通が大きな問題となっており、それを防止するための映像配信の高信頼化技術も求められている。

高信頼度映像配信に関しては、違法コンテンツ流通を抑制するために、コンテンツの著作権保護技術への要求が高まっている。その一つの技術として、映像コンテンツ中に情報を密かに埋め込む電子透かし技術が注目されている。これは、違法なコンテンツが流通した場合に、埋め込まれた正当な著作権者の情報に基づいて権利の保護手段を講じることを可能にすることや、購入者情報をコンテンツに埋め込んでおき、海賊版から流出経路を特定する等を狙いとしたものである。

高能率映像配信に関しては、まず映像コンテンツ制作過程の能率化のために、映像コンテンツ制作に必要な映像素材を高画質かつ低コストで収集する技術が求められている。映像コンテンツの収集には、コンテンツの特性によって異なる要求条件が存在する。一つは実時間性を要求するもの、もう一つは非実時間でよいものである。前者に関しては、例えばオリンピック中継などにおいて固定速度の通信路を用いて複数の映像素材を同時に高能率に高画質で伝送する要求がある。後者に関しては、伝送した素材を後刻、編集に用いることが想定されるため、完全に誤りなしで伝送することが求められる。高能率映像配信のもう一つの側面である視聴者への配信過程では、比較的少数の複数の視聴者に映像コンテンツを高能率に配信する際に、必要な視聴者にだけ配信を実施することが可能な通信技術による配信方式が適している。

そこで、本論文では、高信頼かつ高能率な映像配信サービスに向けて、映像コンテンツ制作、伝送、視聴者への配信に関する一連のプロセスを包含するコンテンツハンドリング技術に関する研究成果を報告する。

本論文は大きく分けて次の4つの研究テーマからなる。

- (1) 映像電子透かし方式
- (2) 実時間映像伝送方式
- (3) 非実時間映像伝送方式
- (4) 高能率映像配信方式

研究テーマ(1)から(3)は、コンテンツの制作側に属するものであり、テーマ(4)は視聴者への映像配信時の技術を取り上げている。

まず、コンテンツ制作者を取り巻く状況として、今日のインターネットの発達により大容量コンテンツの正規の配信が容易になるだけでなく、違法な海賊版コンテンツの流通も容易になっていることが挙げられる。その結果、コンテンツの著作権者の権利侵害が発生し、コンテンツの提供自体が縮減する恐れが出ている。この問題を解決するために、電子透かし技術が注目されている。電子透かしは、メディア情報の内部に人間には知覚できないようにメッセージを埋め込む技術であり、例えばコンテンツの著作権者に関する情報をメッセージとして埋め込み、それが検出された海賊版コンテンツの流通の差し止めを要求するなど著作権の主張を容易にするものである。従来から様々な電子透かし技術が提案されているが、そのほとんどは静止画像を対象としたものである。これには二つの理由が考えられる。第一に埋め込む対象の情報源の冗長度が高いほどメッセージの埋め込みが容易であるため、テキストやオーディオメディアよりも画像メディアが埋め込みに適していること。第二としては、静止画像向け電子透かし方式はそのまま映像メディアに適用可能で、それゆえ商業価値の高いコンテンツの多くを対象とできるためである。ところで、静止画像向け方式を映像に適用する場合、映像の各画像フレームを静止画像とみなして埋め込みを行うが、映像メディアのコンテンツは必ず圧縮形式で流通することを考慮すると、静止画像向け方式の適用は能率が良くない。すなわち、同方式で透かしを検出するためには、いったん圧縮映像を復号し、非圧縮映像に戻した上でフレームごとの処理を行って透かしの検出を行う手順となり、本来の透かし検出以外に要する復号処理が能率低下要素となっている。そこで、研究テーマ(1)では、圧縮映像コンテンツを復号することなく、その中に含まれる電子透かし情報を検出する方式を取り扱う。特に、現在、最も一般的な映像圧縮方式であるMPEGに対応し、その中に含まれているDCT係数値を意図的に変化させ、微小な変化量を多数集めることによって、埋め込まれた情報を統計的に検出

している。これにより、映像コンテンツの検査を効率よく実施し、映像コンテンツの正当な制作者の認証ならびに権利保護を実現することにより高信頼映像配信を可能とする。

次に、実時間映像伝送にあたっては、通常は固定速度の映像伝送専用ネットワークを用いて映像コンテンツの伝送を行う。このとき、通信路に合わせて映像符号化も固定速度で行われるため、画質がそのシーンの内容によって変化するという問題点があった。例えば、ニュース映像で人間のバストショットのみが含まれている場合には高画質だが、動きの激しいスポーツ映像に切り替わるとともに画質が低下する現象はよく見られる。一つの解決方法は映像のシーンに合わせて可変速度で映像符号化することであるが、ネットワークをそれに対応した可変速度で利用することは一般的に困難である。そこで、研究テーマ(2)では、固定速度の1本の伝送路を用いて複数の映像コンテンツを伝送する際に、符号化が易しいシーンの映像コンテンツの符号化速度を減じ、難しいシーンの映像コンテンツの符号化速度を増大させることにより、伝送される映像コンテンツ全体としての画質を最適化する方式を提案している。最適化にあたって、まず、映像シーンのレート歪み特性を直前の符号化結果に基づいて簡便に測定する方法を提案している。ついで、固定速度の伝送路を用いる前提のもとで映像コンテンツ全体の歪みの総和を最小化するためのレート配分方式を導いている。

一方、映像の非実時間伝送に関しては、従来は映像素材を収録したVTRテープを物理的に輸送する方法をとっていた。しかし、非実時間伝送とはいえ、即時性が大きく損なわれる問題や、自由な取材が認められていない国や地域からの映像コンテンツの送付ができないという問題があった。その問題の解決手段として、映像コンテンツをファイル化し、ファイル伝送プロトコルによって通信路を用いて伝送する方法が用いられている。しかしながら、通常のファイル伝送プロトコルでは、その到達確認型プロトコルの本質的な特性のために、衛星や国際間のネットワーク等の伝送遅延が長い経路で伝送する場合、伝送速度が低下するという問題がある。そこで、研究テーマ(3)では、伝送誤りを防止するために代表的な到達確認型プロトコルであるTCPを複数同時に利用することで、全体としての伝送速度を高く保つ非実時間映像伝送方式を取り扱う。このとき、ネットワークの資源を過度に占有しないよう、トラフィックシェーピング方式を新たに取り入れ、他のネットワーク上のトラフィックと共存可能な伝送方式としている。また、この伝送プロトコルを利用したコンテンツ伝送システムにより、フィールド試験を行った結果についても述べる。

最後に、制作されたコンテンツを視聴者に配信する技術が必要となる。比較的少数の視聴者に映像コンテンツ配信を行うには、放送よりも通信が適しているが、受信者の数だけ映像配信サーバから配信するための通信路ならびにサーバ資源を必要とする。この問題に対して、インターネット上で一対多の放送型配信を行うマルチキャスト配信が注目されている。この配信技術では、一台の映像配信サーバから複数台の受信端末へ映像データが伝送されることになる

が、配信に先立って受信端末からネットワークに対して自分が受信する映像コンテンツの配信を申請し、ネットワークは配信申請のあった受信端末へのみデータ伝送を行う。しかし、現実のネットワークでは全体としては均質ではなく、高速な箇所と低速な箇所が存在している。そうした環境で、ある単一の符号化ビットレートで映像を符号化・配信した場合、高速通信路上の受信端末にはその通信路速度を生かせない低画質での配信にもかかわらず、低速通信路上の受信端末には受信できない速度での配信となる問題がある。そこで、研究テーマ(4)では、伝送する映像データを階層化し、階層化映像ごとに個別にマルチキャスト配信する方式を提案している。同方式では、利用可能な受信速度に応じて階層化映像を受信することが可能となり、非均質なネットワークにおいても受信者ごとに適切な映像データを提供できる。また研究テーマ(4)では、代表的な映像圧縮方式である MPEG を対象として、特に階層化表現はされていない通常の MPEG データを圧縮データ形式上で階層データに高速に分離する方式を提案している。さらに、階層データをパケット化して送出する際に、復号開始までの遅延を最も短くする最適な送出方式を提示している。

本論文は以下のように構成されている。まず第1章において、高信頼度高能率映像配信サービスを実現するためのコンテンツハンドリング技術に関する4つの研究テーマが定義される。それを受けて、第2章において、研究テーマ(1)に関してコンテンツの著作権保護を実現する電子透かし技術について述べる。ついで、研究テーマ(2)のコンテンツ素材をリアルタイムで高能率に収集するための映像伝送技術について第3章で示す。引き続き第4章では、研究テーマ(3)のコンテンツ素材を非実時間ながら確実に伝送する映像ファイル伝送技術を示す。そして、第5章では、制作されたコンテンツを放送的にインターネット上で同報配信する階層型マルチキャスト技術を研究テーマ(4)にて紹介する。最後に第6章において、結論として研究成果をまとめる。

## 第2章 電子透かし

### 2.1 緒言

映像コンテンツ配信においては、映像の高能率符号化は必須であり、とりわけDVDやデジタル放送に採用されているMPEG[1]が重要な位置を占めている。そこで、本章では高信頼映像配信に必要な電子透かし技術として、MPEGに対応した方式を検討する。

電子透かしの実用化を考えたとき、コンテンツに対する電子透かしの埋め込みは一度であるのに対して、検出は多数回実施されることから電子透かし検出に要する演算量は低く抑える必要がある。映像コンテンツに関しては、MPEG圧縮形式で配信されること、MPEG復号にも相当程度の演算量が必要であることから、MPEG圧縮形式からMPEG復号することなく直接、電子透かしを検出できることが望ましい。

本章では、特にMPEG-2を対象として、その中に含まれているDCT係数を利用した電子透かし検出方式と、それに対応する電子透かし埋め込み方式を提案する。その基本原理は、MPEG-2の動き補償付きフレーム間予測DCT映像符号化によっており、MPEG-1およびMPEG-4とも共通であるため、提案方式は容易に他のMPEG符号化方式に適用できる。

MPEG電子透かしに関しては、いくつか先行研究がある。Linnartzら[2]は、MPEG符号化で映像フレームの符号化タイプを意図的に操作することで情報を埋め込む方式を提案している。しかし、この方式はMPEG再符号化によって電子透かしが容易に破壊されるという欠点がある。Dittmanら[3]は、DCTブロックベースの方式を提案している。これはDCTブロックを用いて、埋め込み位置をランダム化するという点で本章の方式と共通の要素があるが、彼らの方式では検出にあたってMPEG復号を必要としている。Hartungら[4]は、MPEG符号化されたビットストリームに電子透かしを直接埋め込む方式を提案しているが、やはり検出にはMPEG復号を必要としている。

MPEG復号することなく直接電子透かしを検出するのは、利用できる情報がフレーム間予測残差のDCT係数だけに限られる点に困難さがある。また、DCT係数の意図的な操作は画質劣化を避けるために微小なものでなくてはならないが、あまりに微小であると元々の映像信号が持っている乱雑さに紛れ、検出精度が低下する。そこで、本章で提案する方式では、多数のDCT係数の集合を操作の対象とし、それらを複数フレームに分散させることで検出精度の向上と攻撃に対する頑強さを実現する。

## 2.2 提案方式の概要

提案方式では、MPEG-2 符号化データ中に含まれる DCT 係数を利用する。DCT 係数は透かし埋め込みによって人為的な操作が加えられており、それを透かし検出側で検出する。しかし、画像情報の DCT 係数にはある程度の乱雑さがあるため、単一の DCT 係数の観測だけでそうした操作を判定することは困難である。そこで、透かし埋め込み時に埋め込み情報 1 ビットにつき、複数の DCT 係数が同様の性質を持つように一斉に操作を加え、検出側ではそれらの DCT 係数の集合から統計的に検出を行う。また、DCT 係数は、画像フレームの様々な位置から選択しており、その位置を“テンプレート”として透かし埋め込み側と検出側で予め共有しておく。

透かし埋め込み側では、テンプレートに従って原画像の様々な位置から 8×8 画素ブロックが選択され、DCT 変換の後にそれらの DCT 係数が操作され、再び DCT 逆変換により空間領域に戻される。そして、その透かし入りのステゴ画像が MPEG-2 により符号化されて、利用者へ配信される。一方、透かしの検出側では、MPEG-2 ビットストリーム中の DCT 係数がテンプレートに従って抽出され、電子透かしが統計的に検出される。テンプレートは予め埋め込み側と検出側で共有され、共通鍵暗号通信システムにおける秘密鍵と同様の役割を果たす。すなわち、テンプレートを持たない第三者が MPEG-2 ビットストリームを観測しても、その中に隠されている情報を取り出すことはできない。一連の手順を図 2.2 に示す。

提案方式は、1 ビットの情報の埋め込みに対して、複数の映像フレームに渡って DCT 係数を操作するため、フレーム間分散型電子透かし (Inter-frame diffusion watermark: IFD watermark) と呼ぶ。IFD 方式において、複数の DCT 係数のグループ化は、検出精度および攻撃に対する耐性に対して大きな影響を持つ。このグループ化方式については、2.4 節において述べる。

## 2.3 DCT 係数の操作

映像信号の輝度値からなるフレームを 16×16 画素にタイル状に分割し、それぞれをマクロブロック (MB) とする。2.4 節において述べる規則に従って、複数の MB が選択され、その集合に対して状態“0”または“1”が埋め込み情報として割り当てられる。MPEG-2 の場合、インタレース画像を対象としており、符号化効率を高めるために符号化時に MB ごとにフレーム画像とフィールド画像を適応的に切り替えて取り扱われる。従って、本章では図 2.2 に示すように、フレーム原画像とフィールド原画像の双方に電子透かしの埋め込みをすることで、いずれのモードが選択された場合でも検出が可能な方式を提案する。

MB は 8×8 画素からなる 4 つのブロック B0, B1, B2, B3 から構成される。状態“1”を埋め込むときは、フレーム画像での B0 の DCT 係数およびフィールド画像での B1 と B3 のトップ

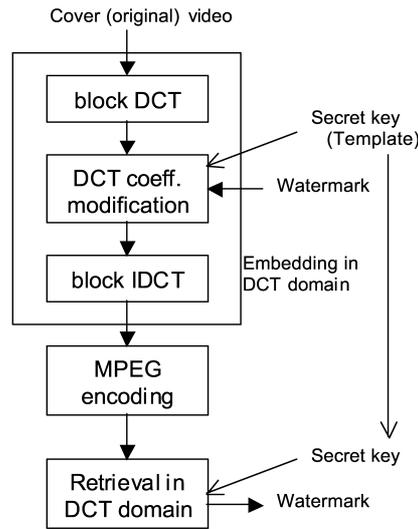


図 2.1 電子透かし埋め込み、圧縮、検出の流れ

Fig. 2.1 Flow of embedding, compression and retrieval.

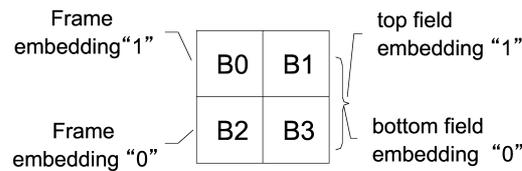


図 2.2 マクロブロックへの埋め込み (フレーム構造での表現)

Fig. 2.2 Embedding for Macroblock (Expression in frame structure).

フィールド側の DCT 係数が操作される。逆に、状態 “0” を埋め込むときは、フレーム画像での B2 の DCT 係数およびフィールド画像での B1 と B3 のボトムフィールド側の DCT 係数が操作される。上記の DCT 係数の操作は、DCT(1,1) 係数の絶対値を増大させるように行われる。

$$x' = \text{sign}(x) \times (|x| + \gamma) \quad (2.1)$$

ここで、 $x$  は元の DCT 係数値であり、 $\gamma$  は操作強度、 $\text{sign}(x)$  は  $x$  の符号 (+/-) を与え、 $x'$  は操作後の DCT 係数値である。 $x = 0$  の場合は、 $\gamma = 0$ 、すなわち埋め込みなしとする。なぜなら  $x = 0$  はその画像ブロックが極めて平坦な領域であることを示しており、そこに埋め込みを行うことは視認可能な歪みを生じるためである。操作強度は人間の視覚特性である周波数感度とマスキング特性 [5] [6] を考慮し、埋め込みが視認できないように決定される。ここで、利用する DCT 係数に関して DCT 高周波係数を用いると、画像の主観画質に与える影響が小さいため容易に圧縮等の信号処理によって取り除かれてしまう。逆に低周波係数への操作は圧

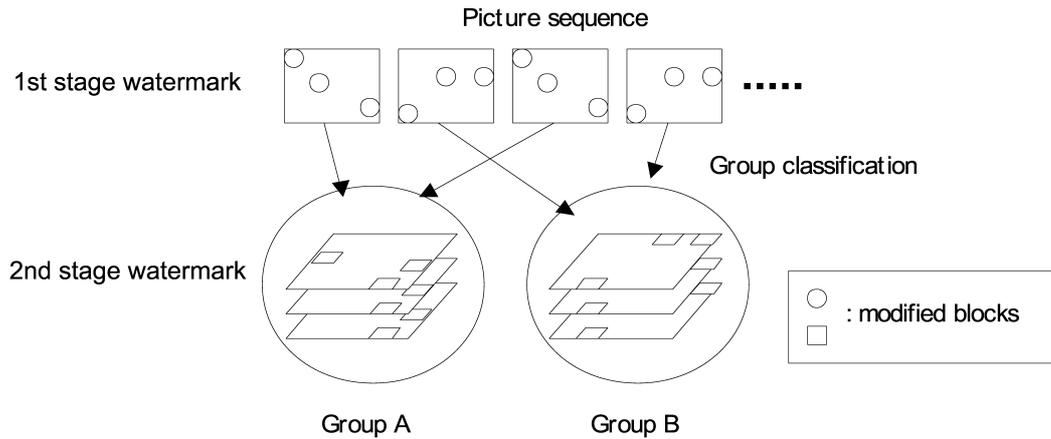


図 2.3 2ステージ電子透かし

Fig. 2.3 2-stages watermarking.

縮後も残るが、より視認されやすい問題がある。低周波係数の中では斜め方向の係数が比較的目に付きにくいDCT(1,1)係数を選択している。

検出側では、Iピクチャだけでなく、PおよびBピクチャのDCT係数も用いて検出を行う。一般に、Iピクチャでは埋め込みによって操作されたDCT係数そのものが観測されるのに対して、P, BピクチャではMPEG-2符号化過程でフレーム間予測符号化を行っているため、観測できるDCT係数は埋め込みの結果そのものではなく、予測残差のDCT係数となり、検出対象にできないと考えられている。しかし、埋め込みによるDCT係数操作の結果、フレーム間予測時にそのDCT係数に予測残差が発生し、MPEG-2符号化後も非零のDCT係数を持つため、間接的ながら埋め込みの影響をP, Bピクチャにおいても観測可能である。

## 2.4 2ステージ電子透かし

IFD電子透かしでは、複数の画像フレームに渡ってMBを選択して構成された集合が電子透かしの埋め込みおよび検出時に用いられる。このとき、すべての画像フレームで同じ位置のMBが同じように操作されていると、そのDCT係数の性質の規則性から第三者に容易に電子透かしの埋め込まれているMB位置が露呈する。そこで、本章では、連続する画像フレームにおける操作対象のMB位置をランダム化するために、映像フレームを複数のグループに分け、図2.3に示すように、グループごとに異なる位置のMBを選択する方式を提案する。そして、映像フレームをグループ分類するための電子透かしの第一ステージ電子透かし、そしてグループ分類後に複数のフレームにまたがって情報を埋め込む電子透かしの第二ステージ電子透かしと呼ぶ。

### 2.4.1 テンプレート

テンプレートは、電子透かし埋め込みと検出に用いられる MB の座標のリストである。その座標は、ステージごとに、またグループごとに異なるよう構成される。

### 2.4.2 第一ステージ電子透かし

まず、各映像フレームをグループに割り当て、グループを識別するための電子透かしの埋め込み。第一ステージ電子透かしのテンプレートの例を図 2.4 に示す。

グループ識別電子透かしは、テンプレートによって選択されるすべての MB に対して、状態“1”を埋め込む。図 2.4 において、 $MB_{A,0}$ ,  $MB_{A,1}$ ,  $MB_{A,2}$ ,  $MB_{B,0}$ ,  $MB_{B,1}$ ,  $MB_{B,2}$  に対して状態“1”を埋め込む。ここで、 $MB_{A,i}$  はグループ A における  $i$  番目の MB を意味し、 $c_{A,i}$  は  $MB_{A,i}$  において操作対象となっている DCT(1,1) 係数を指す。

検出側では、画像フレームごとに第一ステージ用のテンプレートとのマッチングを行い、所属グループを決定する。マッチングにおいては、各グループのテンプレートが指示する MB を選択し、その DCT(1,1) 係数の絶対値和をグループごとに計算し、最大値を与えるグループがその画像フレームの属するグループであると判定する。計算式は以下のとおり。

$$S_p = \sum_i |c_{p,i}|. \quad (2.2)$$

ここで、 $p$  はグループのインデックス、 $i$  はグループ  $p$  内の MB のインデックスを示す。 $c_{p,i}$  は DCT(1,1) 係数、 $S_p$  はグループ  $p$  の第一ステージ用 DCT 係数値の絶対値和を示す。そして、判定結果のグループ  $p^*$  は以下のように定められる。

$$S_{p^*} = \max\{S_p\}.$$

正しいグループのテンプレートを当てはめた場合、その中のすべての MB が状態“1”を持つものに対して、それ以外のグループでは、状態“1”と“0”が入り交じって観測されることになり、両者には統計的に有意な差が生じると考えられる。

### 2.4.3 第二ステージ電子透かし

第二ステージでは、第一ステージとは異なる位置の MB を用いる。それを表現するテンプレートは、グループごとに画像フレーム中の異なる位置の MB を指示したものとなる。そして、1 ビットの情報を埋め込むために、各グループから一つずつ MB を選択した集合に対して、同一の操作を行う。従って、透かし埋め込みに用いられる MB は一つの画像フレームの中限定されず、複数の画像フレーム上の異なる位置から選択されたものとなる。図 2.5 を用いて

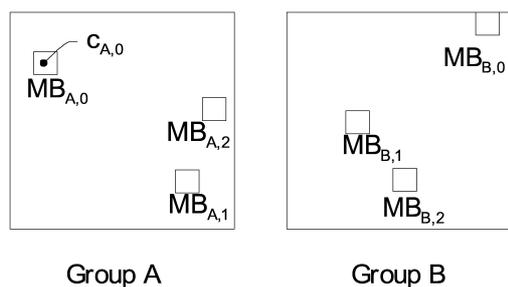


図 2.4 第一ステージ電子透かしのためのテンプレートの例

Fig. 2.4 Template for 1st stage watermark.

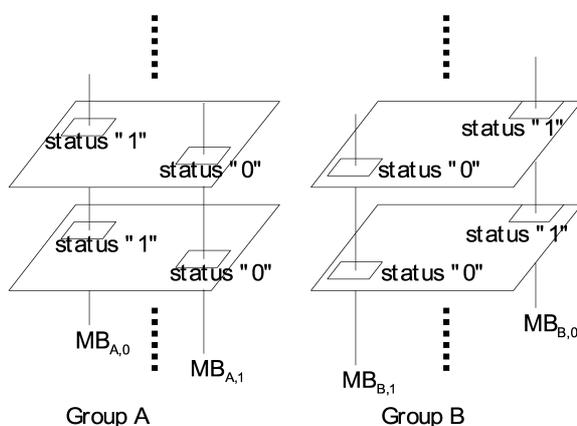


図 2.5 第二ステージ電子透かしのテンプレートの例

Fig. 2.5 Template for 2nd stage watermark.

例示する。メッセージの第一ビット目の情報“1”を埋め込むとき、グループ A の位置  $MB_{A,0}$  およびグループ B の位置  $MB_{B,0}$  にあるすべての MB において状態“1”を埋め込む。同様にメッセージの第二ビット目の情報“0”を埋め込む場合は、 $MB_{A,1}$  および  $MB_{B,1}$  にある MB に状態“0”を埋め込む。このとき、 $MB_{A,0}$  と  $MB_{B,0}$  は画像フレーム上で異なる位置にあるため、透かしの検出能力を高める効果がある。例えば、 $MB_{A,0}$  のみを情報埋め込みに利用するとき、当該領域が静止領域でフレーム間符号化において予測残差が生じないような場合は、MPEG ビットストリーム中に DCT 係数が現れず、透かし検出も不可能となる危険性がある。それに対して、提案方式のように、画像フレーム中の異なる位置に分散して埋め込みを行うことにより、DCT 係数が全く観測されなくなる危険性を抑制し、検出能力を高めることができる。このことから、本章の方式における埋め込み情報容量は、一つのグループが 1 画像フレームあたり利用する MB の個数となる。

次に、図 2.6 を用いて、透かし検出過程について述べる。テンプレート  $MB_{A,0}$  と  $MB_{B,0}$  に位

置するすべての MB を集め、その MB 集合が状態 “0” であるか “1” であるかを判定し、1 ビットの情報を検出する。 $f_{p,i}^k$  と  $d_{p,i}^k$  をそれぞれ状態 “0” に対応する DCT (1,1) 係数および状態 “1” に対応する DCT (1,1) 係数とする。それぞれの DCT 係数値の絶対値和を求め、値の大きくなる方に相当するビット情報として判定する。

$$S_i^0 = \sum_k \sum_p f_{p,i}^k \quad (2.3)$$

$$S_i^1 = \sum_k \sum_p d_{p,i}^k \quad (2.4)$$

$$b_i = \begin{cases} 0 & (S_i^0 > S_i^1) \\ 1 & \text{otherwise} \end{cases} \quad (2.5)$$

ここで、 $p$  はグループインデックス、 $i$  はメッセージビットの項番、 $k$  は画像フレームのインデックスである。なお、DCT 係数  $f_{p,i}^k$  と  $d_{p,i}^k$  は同じ MB に属しているため、量子化ステップサイズやフィールド/フレームモードなど符号化に関するパラメータは同一である。従って、 $f_{p,i}^k$  と  $d_{p,i}^k$  の相違は透かし埋め込みによって生じていると考えられる。

## 2.5 攻撃耐性の評価

### 2.5.1 計算機シミュレーション条件

IFD 電子透かし方式を “StirMark” [7] を含む様々な攻撃のもとで評価する。

透かし埋め込みパラメータとして、グループ数は 8 (A~H)、第一ステージ透かし埋め込みに用いる MB 数は 4、第二ステージ透かし埋め込みに用いる MB 数は 4 とし、DCT 係数の操作強度  $\gamma$  は表 2.1 に従って DCT ブロックの AC エネルギーに応じて適応的に変更する。ここで、AC エネルギーは DCT の AC 係数の二乗和である。

試験に用いた映像は、Football (F)、Mobile&Calendar (M)、および Susie (S) であり、それぞれ 150 フレーム、画像解像度は 720 画素  $\times$  240 ライン、59.94 フィールド/秒のインタレース映像である。8 種類のグループを画像フレームに対して順番に周期的に割り当てた。すなわち、グループ A, B, ..., H, A, B, ..., である。全部で 150 フレームであるため、一つのグループに割り当てられた画像フレーム数は 18 もしくは 19 となる。

まず、電子透かしを原画像フレームに対して埋め込み、ついで MPEG-2 エンコーダ [8] を符号化ビットレート 15Mbit/s、フレームストラクチャ、フィールド/フレーム適応 DCT により動作させた。攻撃者はこのビットストリームに対して、いったん非圧縮画像に復号し、攻撃を施した後に MPEG-2 エンコーダで 15Mbit/s に再圧縮して配布する。電子透かしの検出試験は、この再圧縮後の MPEG-2 ビットストリームに対して実施する。一連の手順を図 2.7 に示す。

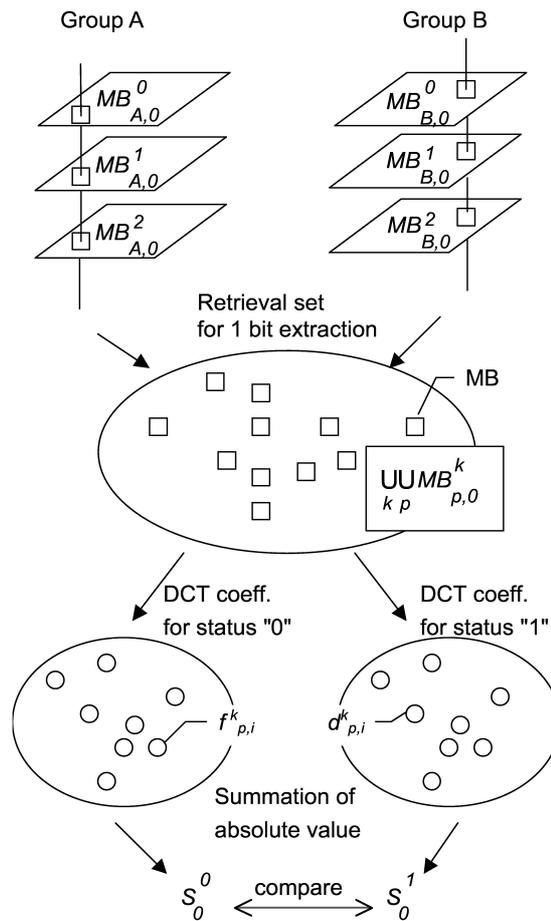


図 2.6 第二ステージにおける 1 ビットの検出の過程

Fig. 2.6 Retrieval process of 1 bit extraction in the 2nd stage.

使用した攻撃方式を表 2.2 に示す。その中で、no attack に関しても MPEG-2 15Mbit/s での符号化が実施されていることに留意する。攻撃 4) から 8) までは StirMark に含まれるツールであり、それぞれデフォルトパラメータを用いている。FMLR[9] はラプラシアン演算子に基づく一種の尖鋭化フィルタによる攻撃である。Default StirMark 攻撃は幾何的歪みと JPEG 圧縮を組み合わせた方法である。幾何的歪みに関しては、カメラぶれ状の画質劣化 [7] を避けるために全画像フレームで同一の幾何的歪み特性を適用している。攻撃 9) は、画像解像度を水平垂直ともに半分ずつにいったん縮小し、その後に再び元の画像解像度に戻す処理である。攻撃 10) は映像信号を業務用 VTR に録画し、再生した映像信号を改めて PC にキャプチャ・デジタル化したものである。

表 2.3 (a) は、透かし埋め込み前の原画像と電子透かし埋め込み直後の画像の間の平均 PSNR である。表から PSNR の値は十分に大きく、透かし埋め込みによる画質劣化は極めて小さいことが分かる。これは、1 画像フレームあたり、少数の MB のみが操作されていることによる。

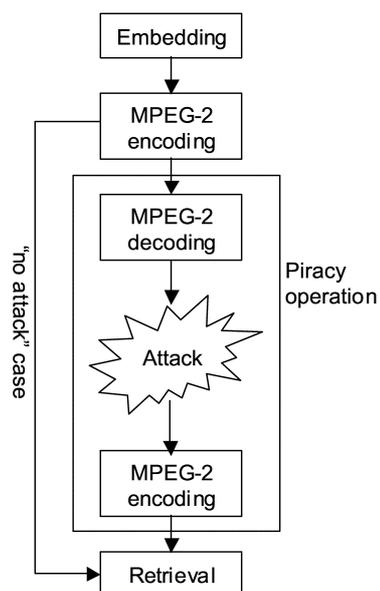


図 2.7 攻撃耐性の評価手順の流れ

Fig. 2.7 Flow diagram for evaluation of attack resilience.

表 2.1 埋め込み強度

Table 2.1 Embedding intensity.

AC energy	0~5	5~20	20~40	40~100	100~
Intensity	15	20	25	30	35

表 2.2 適用される攻撃

Table 2.2 Applied attacks.

No.	Attack
1	no attack (including 15Mbit/s MPEG-2 compression)
2	2Mbit/s MPEG-2 compression
3	3Mbit/s MPEG-2 compression
4	FMLR attack
5	Gaussian filter
6	Median filter
7	Sharpening
8	Default StirMark
9	size conversion/restore
10	DA/AD conversion



図 2.8 透かし埋め込み済み画像 “Susie” の第 10 フレーム

Fig. 2.8 10th frame of the embedded video “Susie”.

また、透かし埋め込み後の画像である図 2.8 から、主観的にも視認可能な画質劣化が生じていないことも分かる。この画像は平坦な領域が多いため、特に画質劣化を識別しやすいことから、他の画像ではさらに透かしによる主観画質劣化は識別されないと言える。表 2.3 (b) は透かし埋め込み前の原画像と攻撃後の画像の間の平均 PSNR である。表から、攻撃の種類によってはその副作用として大きく画質劣化が生じていることがわかる。

## 2.5.2 検出結果

透かし検出結果を表 2.4 と表 2.5 に示す。表 2.4 は第一ステージの電子透かしによるグループ分類の正解率を示す。表からグループ分類は必ずしも 100% 確実ではないことが分かる。攻撃 2), 3), 5), 9) と 10) においては、埋め込みによって生じた状態 “0” と “1” の間の相対的な DCT 係数の違いが弱められ、画像の元来の局所的非常さに紛れて誤検出を引き起こしている。攻撃 4) と 7) においては、DCT 高周波係数は強調されるが、透かし埋め込みに用いている低周波係数への影響は限定的なため、誤検出は抑制されている。攻撃 8) に関しては、幾何学的歪みによって画像フレーム上の画素位置が変化し、DCT 係数の値が大幅に変化するため、グループ分類の失敗につながっている。なお、映像 Mobile&Calendar における正解率が低い理由は、MPEG-2 符号化において映像の高周波エネルギーの高さと複雑な動きのために予測残差が大きく、それに伴って量子化ステップサイズも大きくなり、埋め込んだ DCT 係数自体が消失することが多いためである。

表 2.5 は第二ステージの電子透かし検出結果を示している。ここで、第二ステージの検出にあたっては、第一ステージで誤ったグループ分類がなされた画像フレームも式 (2.3) と (2.4) における総和  $S_i^0$  と  $S_i^1$  の計算に含まれている。表から第一ステージでの誤分類にもかかわらず、ほとんどの場合で正しく埋め込み情報が検出されていることが分かる。これは、誤分類された

表 2.3 画質

Table 2.3 Picture quality.

(a) PSNR (dB) after watermarking

Video scene		
F	M	S
62.37	63.65	62.47

(b) PSNR (dB) after attacks

	Video scene		
Attack	F	M	S
1	38.79	33.59	43.31
2	27.41	22.57	36.29
3	28.25	24.63	38.43
4	35.57	30.82	37.85
5	27.24	23.86	35.29
6	26.59	23.92	34.35
7	17.89	16.55	23.52
8	25.18	14.22	28.72
9	27.28	22.71	35.80
10	35.23	28.40	40.47

表 2.4 第一ステージ電子透かしによるグループ分類結果 (分類の正解率)

Table 2.4 Classification result by 1st stage watermark (percent of correct group classification).

Attack	Video scene		
	F	M	S
1	89	97	100
2	47	33	92
3	65	71	100
4	78	94	100
5	86	97	100
6	76	80	96
7	77	96	100
8	34	8	37
9	87	90	100
10	94	92	100

画像フレームの DCT 係数値が総和  $S_i^0$  と  $S_i^1$  の計算に与える影響が小さいことを示している。図 2.9 を用いて説明する。誤分類された画像フレームの MB では、意図的な操作が行われておらず、状態 “0” および状態 “1” のいずれでもない。そのため、その DCT 係数値は統計的には有意な差は生じず、総和  $S_i^0$  と  $S_i^1$  への影響もない。また、MPEG-2 符号化時に予測残差が小さく、DCT 係数を含まない (skipped) MB の場合もまた、総和への影響を与えない。従って、第一ステージの正解率が 100% でなくても第二ステージでの埋め込み情報の検出は可能である。

表 2.5 の攻撃 2) と 8) においては、埋め込み情報の検出に失敗している。攻撃 2) に関して、映像 Mobile&Calendar を 2Mbit/s で符号化するとき、この攻撃によって画質が大幅に劣化 (PSNR = 22.6 dB) していることから、この場合は映像コンテンツの価値自体が大きく損なわれており、電子透かしによる保護の対象外であると見なせる。他方、StirMark のデフォルト攻撃では幾何学的歪みを含んでいるため、画素位置のずれが発生し、正しい DCT 係数が観測されず、情報検出に失敗していると考えられる。

### 2.5.3 埋め込み情報量拡大時の評価

第二ステージにおける情報埋め込み量は、1 画像フレーム中に存在する MB 数により制限される。本章で試験に用いているテスト画像の場合、画像フレームあたり 1350 個の MB が存在しているが、8 グループに分類し、グループで重畳しないように MB を選択することを考慮すると一つのグループあたり 168 個の MB が利用可能である。従って、理論的には 168bit が最大埋め込み容量となる。本節では、埋め込み情報量を増加させたときの画質並びに検出精度に

表 2.5 第二ステージの 4bit 透かし検出結果 (ビット判定の正解率)

Table 2.5 Retrieval of 2nd stage watermark (4bit) (percent of correct bit retrieval).

Attack	Video scene		
	F	M	S
1	100	100	100
2	100	75	100
3	100	100	100
4	100	100	100
5	100	100	100
6	100	100	100
7	100	100	100
8	75	50	100
9	100	100	100
10	100	100	100

与える影響を評価するために 16, 48, 80bit の埋め込みを実施する。80bit は、コンテンツの権利保有者情報を埋め込むのに必要な情報量である [10]。

表 2.6 は埋め込み前の原画像と埋め込み後のステゴ画像との間の平均 PSNR を示している。埋め込み情報量の増大に伴って、操作される DCT 係数の個数が増加するため PSNR が減少している。ただし、80bit 埋め込みの場合でも PSNR の値は十分に高く、画質も高く保たれていることが分かる。

表 2.7 は 80bit を埋め込んだ場合の第二ステージの検出結果を示している。検出精度は、4bit の埋め込みの場合と同様に、デフォルト StirMark による幾何学的歪み攻撃で、特に映像 Mobile&Calendar においてビット誤り率が高い。映像 Mobile&Calendar では線やエッジが多い映

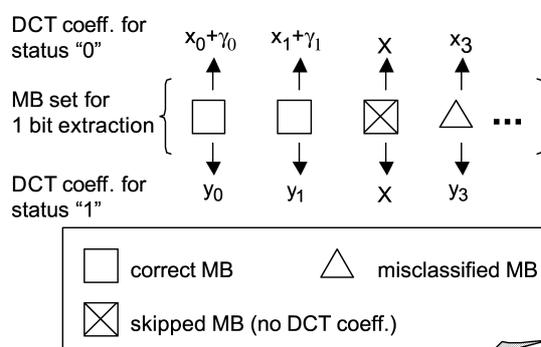


図 2.9 第一ステージのグループ分類誤りが第二ステージ検出に与える影響  
Fig. 2.9 Effect of misclassification in the 1st stage on the 2nd stage retrieval.

表 2.6 電子透かし埋め込み後の PSNR(dB) (大容量情報埋め込みの場合)

Table 2.6 PSNR (dB) after watermarking (large bits embedding).

embedded bits	Video scene		
	F	M	S
16	58.94	59.78	59.72
48	54.74	56.21	55.76
80	52.68	54.18	53.72

表 2.7 第二ステージ電子透かしの検出結果 (80bit 埋め込み中の正解率)

Table 2.7 Retrieval of 2nd stage watermark (80bit) (percent of correct bit retrieval).

Attack	Video scene		
	F	M	S
1	100	100	100
2	100	85	100
3	100	100	100
4	100	100	100
5	100	100	100
6	100	100	100
7	100	100	100
8	94	71	98
9	100	100	100
10	100	100	100

像であるため、画素位置のずれが DCT 係数に大きく影響を与える。他方、映像 Football や Susie の場合、画面中には平坦な領域と複雑なテクスチャの領域が入り交じっていることから、第二ステージで用いている画面内の様々な場所の MB を用いる戦略による効果が大きく、検出性能の向上に寄与している。

## 2.6 考察

### 2.6.1 透かし検出計算量

IFD 電子透かし方式の利点は、他の DCT に基づく電子透かし方式 [3][4] と比較して、演算量が少ない点にある。第一に、電子透かし検出にあたって MPEG-2 復号が不要である。MPEG 復号にはテーブル参照による可変長符号の復号および逆量子化、数値演算を含む逆 DCT、そして画像フレームメモリ参照と数値演算を伴う動き補償処理が必要である。IFD 電子透かし

方式では、このうち可変長復号および逆量子化処理しか必要としないため、演算量を1/3に削減することができる。

第二に、電子透かし検出処理自体の演算も少ないという特長がある。画像フレーム中の画素の数を  $N$ 、埋め込みビット数を  $k$ 、画像フレーム数を  $F$  とするとき、Dittman らの方式 [3] では、MPEG 復号後の映像データにおいて、埋め込み情報の 1bit あたり 3つの DCT 係数の比較が必要であることから、 $3k$  回の DCT 演算と  $3k$  回の比較演算が行われる。Hartung らの方式 [4] では、画素値の相関演算が必要であり、 $N$  回の積算、 $N-1$  回の加算、 $k$  回の比較が行われる。それに対して、IFD 電子透かし方式では、第一ステージで  $32F$  回の加算と  $8F$  回の比較、第二ステージで  $2kF$  回の加算および比較が行われる。 $k$  と  $F$  が  $10^2$  の桁数、 $N$  が  $10^5$  の桁数、DCT 変換演算が  $10^2$  オーダの積算と加算であることを考慮すると、IFD 電子透かし方式では Dittman [3] と比較して一桁、Hartung [4] と比較して 3桁も演算量が少ないことが分かる。従って、IFD 電子透かし方式は MPEG 復号と透かし検出処理の両面において、極めて演算量の少ない方式であることが示された。

## 2.6.2 検出結果の信頼度

IFD 電子透かし方式では、DCT 係数値が操作されているか否かを多数の DCT 係数値から判定する。本節では特に第二ステージにおける判定結果の妥当性に関して、DCT 係数値の観測値に関する統計的解析により議論する。

第二ステージでは、1bit の情報の検出に複数の DCT 係数からなる集合を用いる。検出過程において、状態 “0” に相当する DCT 係数と “1” に相当する DCT 係数の絶対値和を求め、それらの相対比較によって状態を決定している。このとき、両者の差が透かし埋め込み時の意図的な DCT 係数値操作によるものと画像信号の元来の性質によるものの双方の影響を受けていると考えられるため、統計的に解析する必要がある。

本節では、統計的検定により評価を行う。このとき、観測値 (DCT 係数) は以下の性質を持つ。1) 観測値はつねに状態 “0” と “1” に対応するペアとして得られる。2) 透かしが埋め込まれている場合には、ペアのうち、いずれか一方が常に大きな絶対値を持つ。これらから、この統計的検定の問題は以下のように定式化される [11]。

$x_i$  と  $y_i$  を状態 “0” と “1” に対応する DCT 係数とすると、観測値ペアの系列  $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$  から求められる平均  $\mu_x$  と  $\mu_y$  が異なるか否かを検定する。ここで、平均  $\mu_x$  と  $\mu_y$  を比較することと、検出過程における  $S_0$  と  $S_1$  の比較は等価であることに留意する。このとき、帰無仮説は  $H_0: \mu_x = \mu_y$  とする。そしてそれが棄却された場合、 $\mu_x \neq \mu_y$  を意味し、より大きな平均値を持つ側が埋め込みが行われている状態に対応すると解釈される。一方、棄却されなかったときは、 $\mu_x$  と  $\mu_y$  に有意な差がないことを示し、透かし埋め込みの存在

表 2.8 t 検定結果

Table 2.8 Result of t-test.

(a) 3Mbps compression

n	$t_0$	$t(\phi, \alpha)$	$H_0$ rejection
10	1.212872	2.364623	no
20	1.210878	2.093025	no
50	2.679433	2.009574	yes
100	3.788409	1.984217	yes

(b) Gaussian filtering

n	$t_0$	$t(\phi, \alpha)$	$H_0$ rejection
10	1.939702	2.262159	no
20	0.924734	2.093025	no
50	2.290927	2.009574	yes
100	2.859046	1.984217	yes

(c) no attack

n	$t_0$	$t(\phi, \alpha)$	$H_0$ rejection
10	4.810927	2.262159	yes
20	3.845278	2.093025	yes
50	5.749094	2.009574	yes
100	8.352571	1.984217	yes

(d) no watermark

n	$t_0$	$t(\phi, \alpha)$	$H_0$ rejection
10	0	2.262159	no
20	0.555158	2.093025	no
50	0.176436	2.009574	no
100	0.307429	1.984217	no

は不確実であることを意味する。計算手順を以下に示す。

$$t_0 = \frac{\bar{d}}{\frac{s}{\sqrt{n}}} \quad (2.6)$$

$$t_0 \geq t(\phi, \alpha) \text{ のとき } H_0 \text{ は棄却.} \quad (2.7)$$

ここで、

$$\begin{aligned} d_i &= x_i - y_i \\ \bar{d} &= \frac{1}{n} \sum d_i \\ s &= \sqrt{\frac{\sum (d_i - \bar{d})^2}{n-1}} \\ \phi &= n-1 \end{aligned}$$

$t(\cdot)$  は t-分布、 $\phi$  は自由度、そして  $\alpha (= 0.05)$  は危険率を示す。

映像 Football に 4bit 埋め込みを行い、攻撃として (a) MPEG-2 3 Mbit/s compression および (b) Gaussian filtering を行ったときの  $t_0$ 、 $t(\phi, \alpha)$  および  $H_0$  の検定結果を表 2.8 に示す。さらに比較のために、(c) 攻撃なし、および (d) 透かしの埋め込みなし、についても示す。また、それぞれの表において、観測するフレーム数を変化させたときの、それぞれの評価値を示す。

(a) と (b) から  $H_0$  の棄却には 20 画像フレームでは不足で、情報検出結果が信頼できないが、50 フレームあれば十分であることがわかる。(c) の場合は、10 画像フレームの観測で信頼できる結果が得られることが分かる。(d) については、100 フレームを観測しても帰無仮説が棄却されず、状態 “0” と “1” の間には有意な差がないことを示している。この性質を利用することにより、電子透かしの検出者は電子透かしが埋め込まれていないことを判定し、“false positive” [12] を避けることができる。

## 2.7 結言

IFD 電子透かし方式では、MPEG-2 ビットストリームを復号することなく直接電子透かしの検出を行うことが可能であり、周波数領域の攻撃に対して耐性があることを示した。IFD 方式では、埋め込み時の画像フレームのグループ指定は任意であり、擬似乱数を用いてランダム化することにより、画像フレームを時間方向に連続的に観測したときの秘匿性を高める効果がある。また、検出側ではこの擬似乱数に関する知識は不要であり、テンプレート情報さえ保持していれば検出が可能である。従って、IFD 方式はブラインド電子透かし方式 [12] の一つになる。

以上により、本章で提案した IFD 電子透かし方式は、低演算量で著作権者の権利を保護するためのコンテンツハンドリング技術を与え、高信頼かつ高能率な映像配信を実現することができる。



## 第3章 実時間映像伝送方式

### 3.1 緒言

実時間映像伝送に対する需要は、オリンピック等の大規模イベントにおける会場からの各種競技映像の伝送や、日常的な放送局間の提携等によるニュース映像、スポーツ映像、経済情報映像レポートなど、即時性を求められるものに対して存在している。特に衛星放送やCATVなどにおける映像の多チャンネル化の進展を背景として、放送局間の大容量多チャンネル映像伝送への要求が高まっている。また、一般に大容量の回線ほど情報量あたりの通信コストが安くなることから、複数の映像を一括伝送する方式が有利である。

複数の映像の一括伝送には、多重化効果により伝送効率を向上させられるという特長が知られている。一般に、映像の持つ情報量は時々刻々と変化しており、符号化により発生する情報量も時間により変動する。このとき、複数のエンコーダ出力を観察すると、映像情報源が異なる場合にはそれぞれの発生情報量が同時に大きな値になる確率は小さい。そこで、複数のエンコーダ出力を多重することにより情報量の総和が平滑化されるという大群化効果を得ることができる。

通常、この大群化効果は独立に動作するエンコーダの場合、最低でも数十台以上のエンコーダ出力の多重化が必要である [13] が、各エンコーダを外部からの統一的制御の下に動作させることにより少数のエンコーダ多重化でも効果が得られることが古閑らにより示されている [14]。古閑らにより提案されている方式は、各エンコーダの出力バッファの読出しクロックを操作するものであるが、エンコーダの出力バッファの大きさやエンコーダごとのバッファ蓄積状態の相違によって、意図した制御が実際にエンコーダの符号化制御に反映されるまでに不定な時間差が生じる問題がある。

本章では、複数のエンコーダ出力を多重化し、複数の映像プログラムを一括伝送するマルチエンコーダシステムにおいて、エンコーダにおける符号化映像品質を客観的に把握しつつ制御を行うことを目的として、各エンコーダでのレート歪み特性に基づいてエンコーダ間でのビットレート配分を行う方式を提案する。そして、本方式による符号化利得の解析から、マルチエンコーダシステムの特性を示す。

### 3.2 マルチエンコーダシステム

マルチエンコーダシステム [15] は複数のエンコーダを統一的制御の下で動作させるシステムである。本論文においては、その制御目標を総ビットレート一定の下での歪み総和の最小化とする。したがって、マルチエンコーダシステムの制御における操作対象は、各エンコーダの

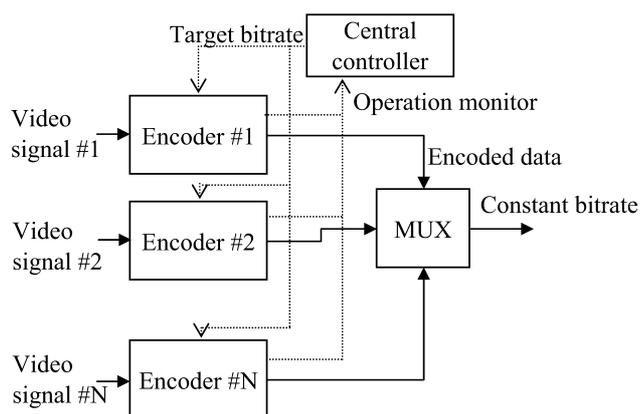


図 3.1 マルチエンコーダシステムの構成

Fig. 3.1 Configuration of a multi-encoder system.

符号化ビットレートとなる。図 3.1 にマルチエンコーダシステムの構成図を示す。構成要素は  $N$  台のエンコーダ、集中制御装置、およびエンコーダ出力の多重化装置である。

以下に各構成要素の機能について述べる。

集中制御装置は各エンコーダの動作状態を示すパラメータを入力信号として、制御アルゴリズムに基づいてビットレート配分を計算し、各エンコーダへ出力する。

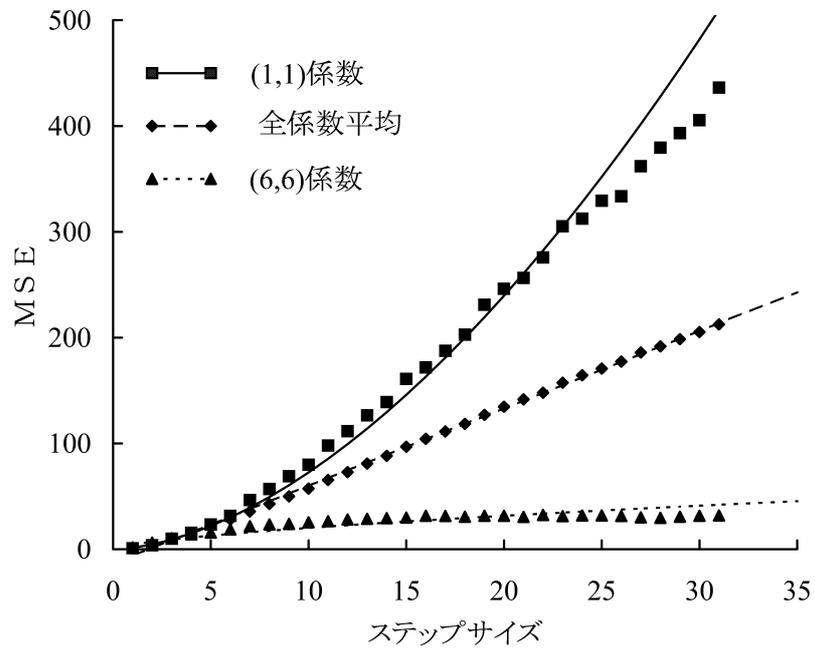
各エンコーダは集中制御装置から指定されたビットレートで動作するためのレートコントロール機構を備えており、画像信号を入力し、符号化ビットストリームと動作状態パラメータを出力する。本論文にて取り扱う符号化アルゴリズムは、映像に対して最も標準的な MPEG で利用されている MC+DCT 方式とする。

多重化装置は各エンコーダの出力を多重化し、固定速度の通信路に送り出す。本章においてはエンコーダの出力ビットレートの総和のみに着目し、各符号化ビットストリームの多重・分離に関しては考察の対象としない。

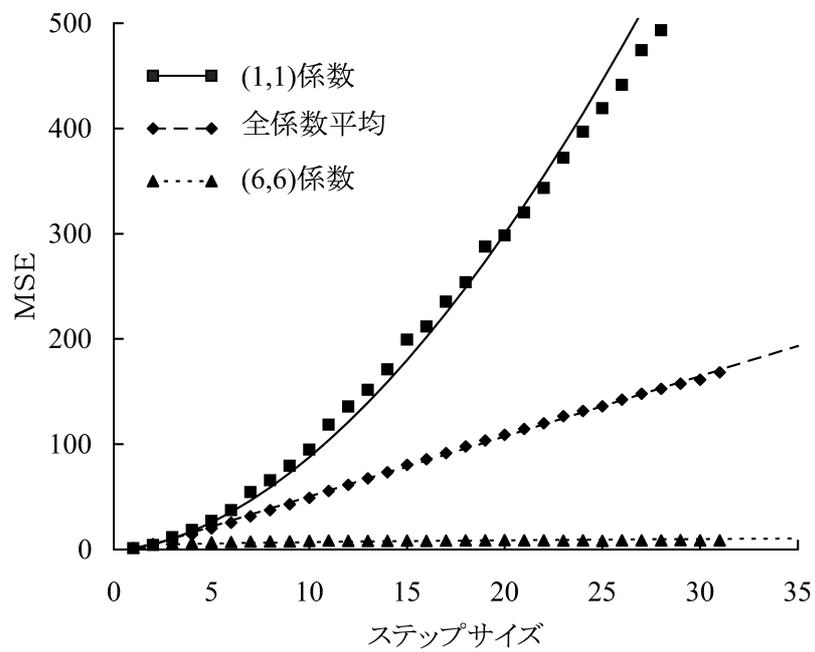
### 3.3 MC+DCT エンコーダのレート歪み特性

本節では各エンコーダのレート歪み特性の定式化を目的として、MC+DCT 符号化器のビットレート対平均二乗誤差 (MSE) 特性を表現する関数を求める [16]。

ここではレート歪み特性を求めるにあたり、量子化ステップサイズを媒介変数として解析を行う。まず、量子化ステップサイズ対 MSE の関係を調べ、ついで量子化ステップサイズ対ビットレートの関係を調べる。最後にその 2 つの結果を併せることで目的のレート歪み特性を得る。



(a) FLOWER GARDEN



(b) CHEER LEADER

図 3.2 量子化ステップサイズ対 MSE 関係

Fig. 3.2 Relation of quantization step-size and MSE.

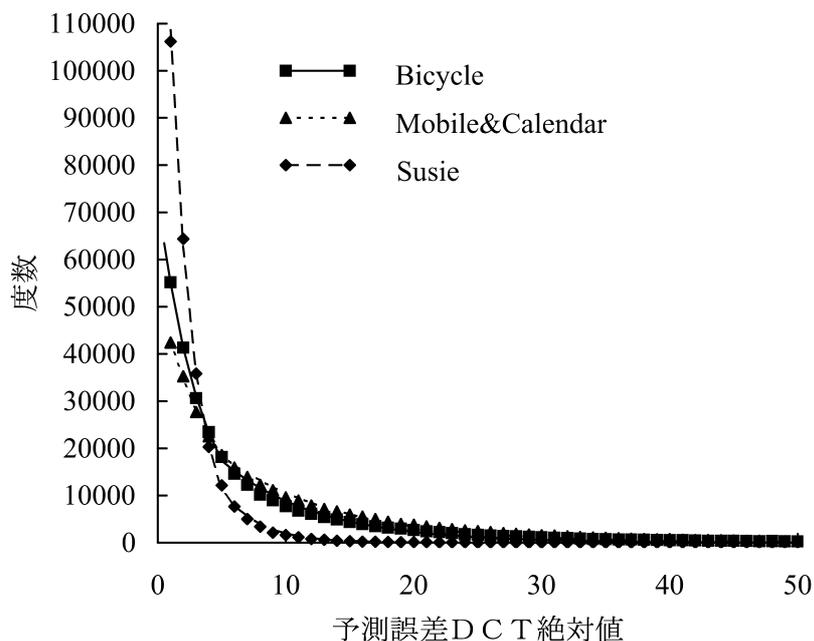


図 3.3 予測誤差 DCT 信号の性質

Fig. 3.3 Characteristics of prediction error DCT signal.

### 3.3.1 ステップサイズ対 MSE 関係

エンコーダとして MPEG-2 の TM5 [17] を用いて、計算機シミュレーションによりステップサイズ  $\Delta$  対 MSE の関係を調べた。映像 FLOWER GARDEN と CHEER LEADER の第 1 予測符号化フレーム（最初の I フレーム直後の P フレーム）において、動き補償誤差信号の DCT 係数に対してステップサイズを変えながら DCT (1,1) 係数、(6,6) 係数および全係数の平均の MSE を求めた結果を図 3.2 に示す（ここで、DCT 係数のインデックスは 0 から始まる）。また、図 3.2 において、(1,1)、(6,6) 係数についてはべき乗関数で、全係数平均については直線でそれぞれ近似したところ、いずれも良好な結果を得た。ここで、それぞれの近似が成り立つ理由について考察する。

まず、低次係数については、一般的ガウス分布 [18] に従うことが報告されており、この線形量子化特性は MSE を  $D_l$ 、 $c_l$  を比例定数とすると

$$D_l = c_l \Delta^\lambda, \quad 1 < \lambda < 2 \quad (3.1)$$

となる。

一方、(6,6) 係数に代表される高次係数は電力が小さいため、ステップサイズが大きくなる

と切り捨てられるので、 $\Delta - \text{MSE}(D_u)$  特性としては図 3.2 のように頭打ちになる。よって

$$D_u = c_u \Delta^\mu, \quad 0 < \mu < 1 \quad (3.2)$$

ここで  $\Delta^\lambda$  を  $\Delta = a$  で Taylor 展開すると

$$\begin{aligned} \Delta^\lambda &= \left\{ (1 - \lambda) + \frac{\lambda(\lambda - 1)}{2} - \frac{\lambda(\lambda - 1)(\lambda - 2)}{6} + \dots \right\} \alpha^\lambda \\ &+ \left\{ \lambda - \lambda(\lambda - 1) + \frac{\lambda(\lambda - 1)(\lambda - 2)}{2} - \dots \right\} \alpha^{\lambda-1} \Delta \\ &+ \left\{ \frac{\lambda(\lambda - 1)}{2} - \frac{\lambda(\lambda - 1)(\lambda - 2)}{2} + \dots \right\} \alpha^{\lambda-2} \Delta^2 + \dots \end{aligned} \quad (3.3)$$

となる。 $\Delta^\mu$  についても同様の結果が得られる。このとき、 $\Delta^\lambda + \Delta^\mu$  を考えると  $1 < \lambda < 2$ 、 $0 < \mu < 1$  の条件から  $\Delta$  の 1 次以外の項が  $\Delta^\lambda$  と  $\Delta^\mu$  で互いに逆符号となり、打ち消し合う傾向となる。よって、粗い近似として次式を得る。

$$D_l + D_\mu \sim 2 - (\lambda + \mu) + (\lambda + \mu)\Delta. \quad (3.4)$$

すなわち、全係数平均歪みはステップサイズに対して線形特性に近くなる。

### 3.3.2 ステップサイズ対ビットレート関係

ビットレートは主として動きベクトルなどのサイド情報に割り当てられる符号量と DCT 係数に割り当てられる符号量から決まる。サイド情報符号量はステップサイズにはあまり依存せず、数 Mb/s 以上の符号化レートではその割合も小さいので、ここでは DCT 係数の符号量について考える。ここで、DCT 係数の符号量は量子化代表値を表現するためのイベント数と 1 イベントあたりの平均符号長から決まる。そこで、図 3.3 に種々の画像における最初の I フレーム直後の P フレームにおける予測誤差信号の DCT 係数絶対値の度数分布を示す。図 3.3 より、DCT 係数絶対値はべき乗的に減少する確率分布となることがわかる。このことから、ステップサイズが大きくなるにつれて、量子化代表値が 0 ではない有意係数の個数がべき乗的に減少することがわかる。ところで、有意係数の個数は量子化後のジグザグ走査を考えると、(ゼロラン, 量子化代表値) からなるイベントの個数に等しい。したがって、ステップサイズによらず 1 イベントあたりの平均符号長があまり変わらなければ、発生情報量はイベント数、すなわち有意係数個数に比例する。以上の考察により、ステップサイズ対ビットレート関係はほぼべき乗関係となるものと予想される。

図 3.4 に実験により得た、最初の I フレーム直後の P フレームにおけるステップサイズ対ビットレート関係を示す。グラフ中に示したべき乗近似曲線は良好な近似を与えており、予想したべき乗関係が実際に成り立つことがわかる。

### 3.3.3 レート歪み特性

ステップサイズ対歪み関係が線形であること、ステップサイズ対ビットレート関係がべき乗に従うことから、ビットレート ( $R$ ) 対 MSE( $D$ ) 関係は  $a, b$  をパラメータとして次式にて表現できる。

$$D = 10^b R^a \quad (3.5)$$

$$\log_{10} D = b + a \log_{10} R. \quad (3.6)$$

図 3.5 にビットレート対 MSE 関係の実験値によるグラフを示す。図 3.5(a) は線形軸によるグラフ、図 3.5(b) は両対数軸によるグラフであり、いずれも最初の 15 フレームにわたって平均した MSE の値を用いている。図 3.5(b) における実験値を表すドットが直線で結ばれることから、(3.5),(3.6) 式による関係式が成り立つことがわかる。

ここで、図 3.5(b) における各直線の傾きが画像によらず、ほぼ等しいことが観察される。したがって直線の傾きである  $a$  は符号化アルゴリズムに依存するパラメータであると考えられる。そこで、(3.5),(3.6) 式における  $a$  を一定値と近似し、 $b$  の値を画像ごとに図 3.5(b) から求めた。この新しい近似式に基づいて描いた曲線を図 3.5(a) に併せて示す。図 3.5(a) から、この  $a$  を一定とする近似が精度良くレート歪み関係を表現していることがわかる。したがって、以下の節では (3.5),(3.6) 式における  $a$  の値は一定として取り扱うこととする。

## 3.4 ビットレート配分制御方式

### 3.4.1 ビットレート配分式

マルチエンコーダシステムの制御に関して、制御の目標は各エンコーダで発生する歪み量 (MSE) の総和を最小にすることで、制御の対象となるのは各エンコーダに割り当てるビットレートであり、制約条件はビットレートの総和を一定に保つことである。これらを数学的に表現すれば

$$\text{制約条件} : R_{total} = \sum_{i=1}^N R_i \quad (3.7)$$

$$\text{最小化式} : D_{total} = \sum_{i=1}^N D_i = \sum_{i=1}^N 10^{b_i} R_i^a \quad (3.8)$$

となる。ここで、 $R_i$  は各エンコーダの割り当てビットレート、 $R_{total}$  は総ビットレート、 $D_{total}$  は総歪み、 $N$  はエンコーダ台数である。この問題は条件付き最小化問題であるので、Lagrange の未定乗数法 [20] により、 $R_i$  を求めると

$$R_i = \frac{R_{total}}{10^{b_i/(a-1)} \sum_{i=1}^N 10^{-b_i/(a-1)}} \quad (3.9)$$

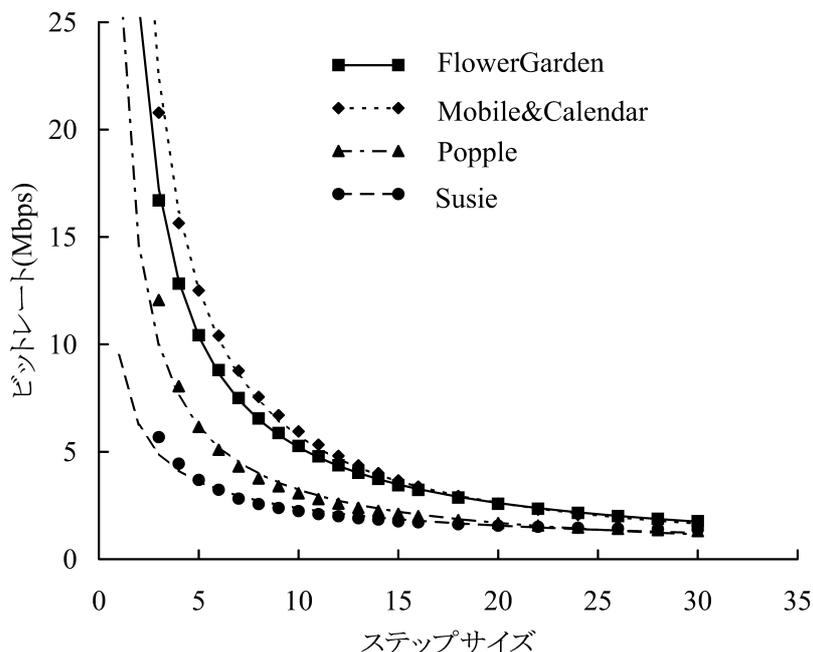


図 3.4 量子化ステップサイズ対ビットレート関係  
 Fig. 3.4 Relation of quantization step-size and bitrate.

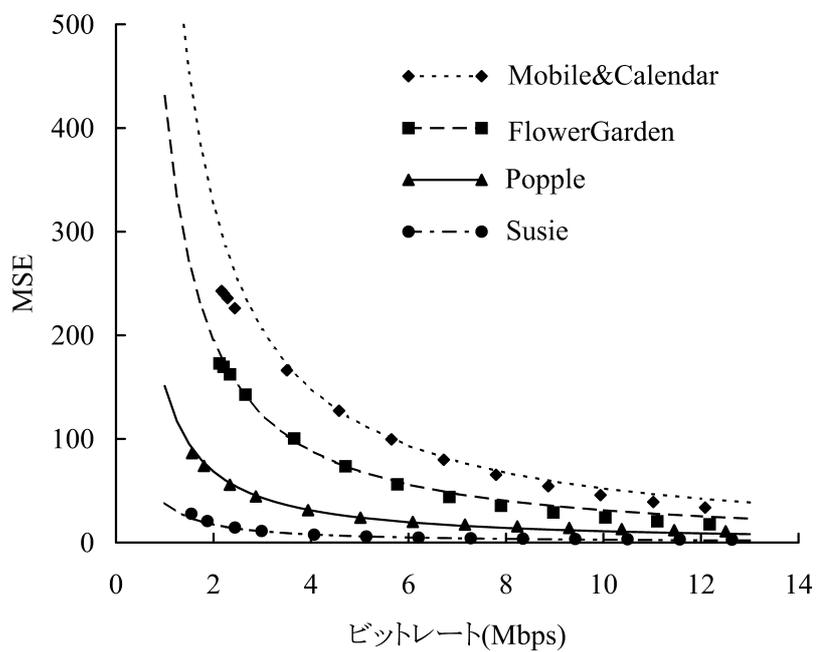
となる。

### 3.4.2 ビットレート配分の更新

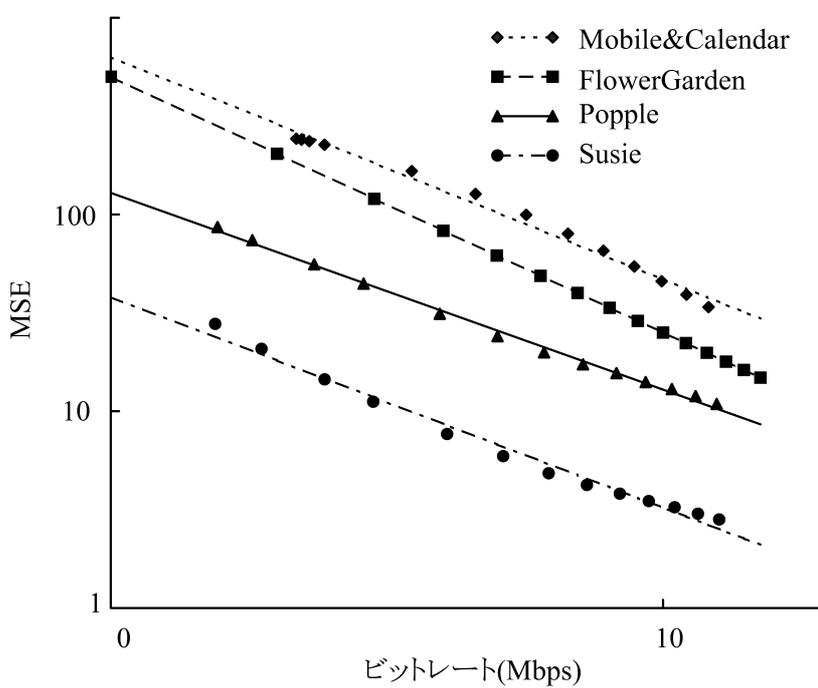
画像のレート歪み特性は画像のシーンによって変化する。一般に映像は時間方向に非定常であるから、時間が経過するとシーンが移り変わって (3.9) 式のパラメータ  $b_i$  も変化する。そこで、時間経過につれて各エンコーダへの割り当てビットレートを更新しなければならない。

本章ではパラメータを1フレームごとに更新することとする。比較的短い間隔で更新する理由は、更新間隔が長すぎると映像シーンの変化に追従できず、特に符号化の易しいシーンから難しいシーンへ移ったときに重大な画質劣化を引き起こすためである。

更新の方法は、まず直前に符号化したフレームのレート  $R_i$  と歪み  $D_i$  の測定値を (3.5) 式へ代入し、固定値  $a$  を用いてパラメータ  $b_i$  を決定する。ついで、このようにして各エンコーダで求めたパラメータ値  $b_i$  を (3.9) 式に代入することで、各々のエンコーダに対する割り当てビットレートを求めることができる。



(a) Linear scale



(b) Bi-logarithmic scale

図 3.5 ビットレート対 MSE 関係  
 Fig. 3.5 Relation of bitrate and MSE.

### 3.4.3 マルチエンコーダ利得

本節では、エンコーダ間でビットレート配分を行うことによる利得をマルチエンコーダ利得と定義し、マルチエンコーダシステムの性能の特性について検討する。

マルチエンコーダ利得は、ビットレート配分制御を行う場合のMSEの総和に対する、行わない場合のMSEの総和の比と定義される。制御を行わない場合、各エンコーダに指定されるビットレートは、総ビットレート  $R_{total}$  をエンコーダ台数  $N$  で割った値 ( $R_{total}/N$ ) とする。したがって、制御を行わない場合のMSE総和  $D_{total}^*$  は(3.8)式から直ちに得られる。一方、制御を行う場合のMSE総和  $D_{total}$  は、(3.9)式を(3.8)式に代入して得られる。以上より、マルチエンコーダ利得  $G$  は  $D_{total}$  と  $D_{total}^*$  の比であるから(3.10)式となる。

$$G = \frac{\sum_{i=1}^N \frac{10^{b_i}}{N^a}}{\sum_{j=1}^N \frac{10^{b_j}}{10^{b_j/(a-1)} \sum_{k=1}^N 10^{-b_k/(a-1)}}}. \quad (3.10)$$

まず、マルチエンコーダ利得の最大値について、エンコーダ2台の場合について解析する。 $N$ が2の場合の(3.10)式におけるパラメータ  $b_1$ 、 $b_2$  を変化させてマルチエンコーダ利得  $G$  を数値的に計算したところ、 $b_1$  を固定して  $b_2$  を大きくすると  $G$  の最大値は2.02に収束した。ここで  $a$  の値は図3.5(b)より  $a = -1.015$  である。一方  $G$  は  $b_1 = b_2$  の時に最小となり、最小値は1となった。すなわち、(3.9)式を用いたマルチエンコーダシステムは、エンコーダを制御しない場合と比較して性能が劣ることはない。マルチエンコーダシステムは割り当てビットレートを各画像に対して等しくすることができるので、ビットレート割り当て制御のない場合を含むためである。逆に利得が最大になるのは、 $b_1$  と  $b_2$  がかなり異なるとき、すなわち符号化の難易が異なる画像がペアになっているときである。一般の場合を考えてみると、エンコーダの台数が多くなるほど利得の最大値は増加する。このとき最大利得を得るための条件は、ひとつの画像のみが難しく、他の画像すべてが易しいことである。なぜなら、易しい画像に割り当てるビットレートを増やす際に、易しい画像の数が多いほど難しい画像に割り当てられるビットレートが増えるためである。

ところで、利得が大きくなるのは画像の符号化難易度が極度に異なる場合で、現実の利得の値は適用する画像に依存する。そこで、図3.6に画像”FLOWER GARDEN”(F)、“MOBILE & CALENDAR”(M)、“SUSIE”(S)を2つずつ組合せた場合(2チャンネル多重)の座標点を示す。ここで  $b_i$  の値は各画像の最初の15フレームを用いて求めた値である。図3.6において、利得の最大値は2.02であるのに対して、最大利得となる組合せ(S,M)でも1.40である。

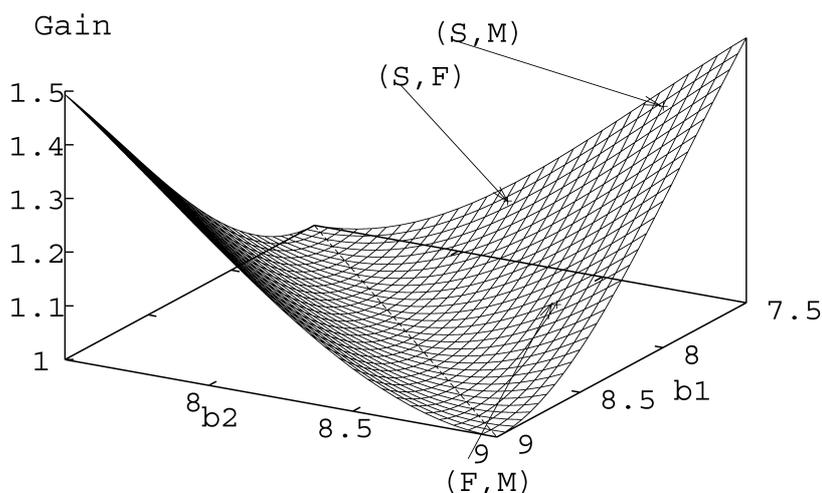


図 3.6 マルチエンコーダ利得特性

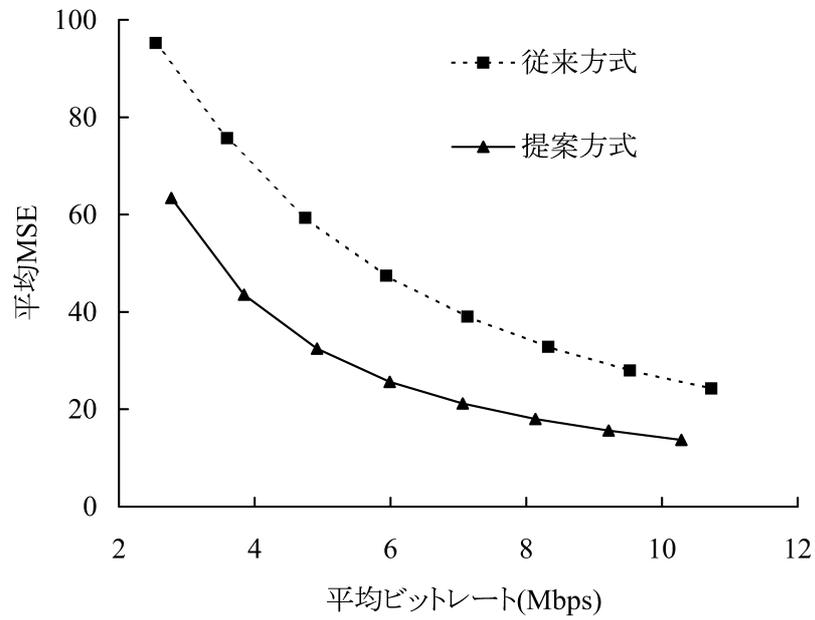
Fig. 3.6 Characteristics of multi-encoder gain.

### 3.5 計算機シミュレーションによる評価

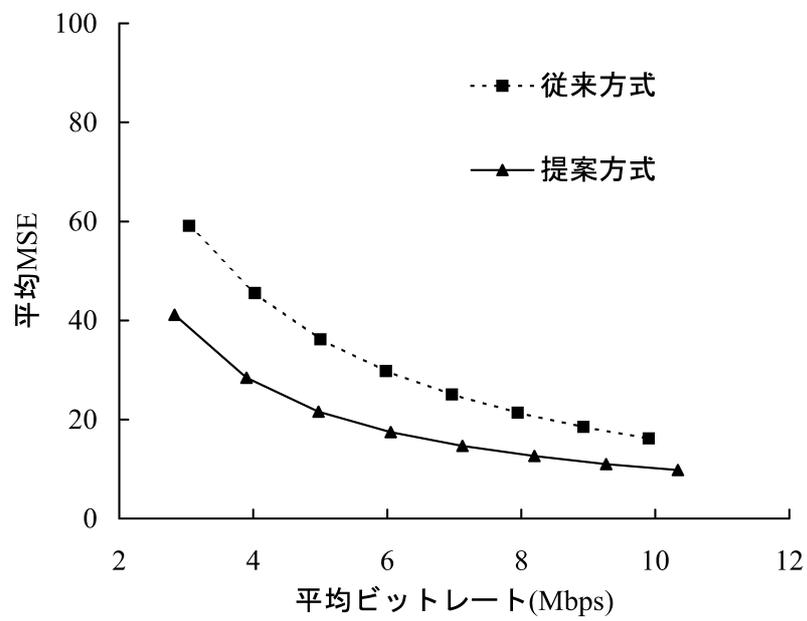
ビットレート配分を行う提案方式と、すべてのエンコーダが同じビットレートで動作する従来方式を比較する目的で、平均ビットレート対平均歪み特性、割り当てビットレートの時間変化、そして符号化画像の局所的な符号化歪みの分布について実画像を用いた計算機シミュレーションを行った。シミュレーションに際して、MPEG-2 TM5をエンコーダとして用いた。TM5のパラメータは、MP@ML, N=15, M=3, Dual Prime 予測有りである。

図 3.7 に符号化特性を示す。これは各画像の最初の 15 フレームに対する平均ビットレートと平均 MSE の関係について、提案方式と従来方式におけるシミュレーション結果を示している。図 3.7(a) はエンコーダが 3 台の場合、(b) はエンコーダが 4 台の場合をそれぞれ示す。3 台の場合に用いた画像は”POPPLE”, ”MOBILE & CALENDER”, ”SUSIE” で、4 台の場合はそれに”SUSIE”の第 100 フレーム以降のフレームを加えている。図 3.7 から、提案方式が常に従来方式よりも少ない平均二乗誤差で符号化を行っていることがわかる。平均ビットレート 5 Mb/s のときで、エンコーダ 3 台の場合、従来方式に対して MSE が約 35% 減少、SNR でみれば約 2dB 向上している。したがって、エンコーダ間でのビットレート配分がシステム全体としての性能向上につながったことがわかる。一方、エンコーダ 4 台の場合では、MSE が約 45% 減少、SNR で約 2.7dB の向上となり、エンコーダ 3 台の場合に対してかなり優れた結果となった。この結果から符号化の易しい画像の数が増えるほどマルチエンコーダシステムの性能が向上することがわかる。さらに、理論値とシミュレーション結果がほぼ一致していることから、前節で行った考察が成り立っていることを示している。

次に、各エンコーダを割り当てられたビットレートの時間変化を図 3.8 に示す。図 3.8 では、



(a) エンコーダ 3 台の場合



(b) エンコーダ 4 台の場合

図 3.7 ビットレート対 MSE 関係  
 Fig. 3.7 Relation of bitrate and MSE.

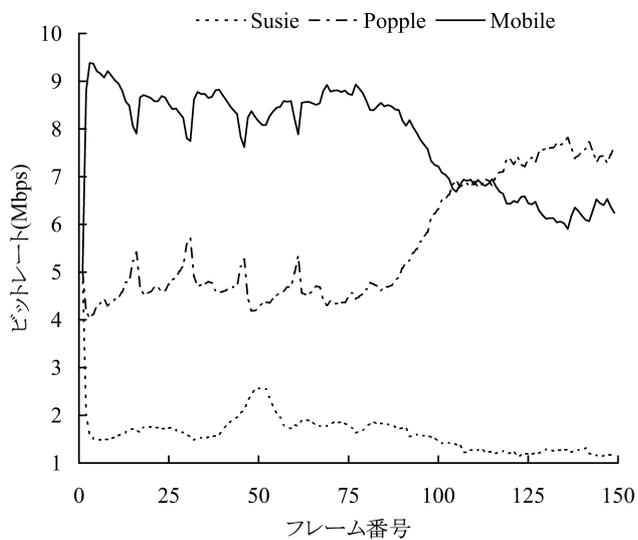


図 3.8 ビットレート配分の時間変化

Fig. 3.8 Characteristics of assigned bitrate for each encoder.

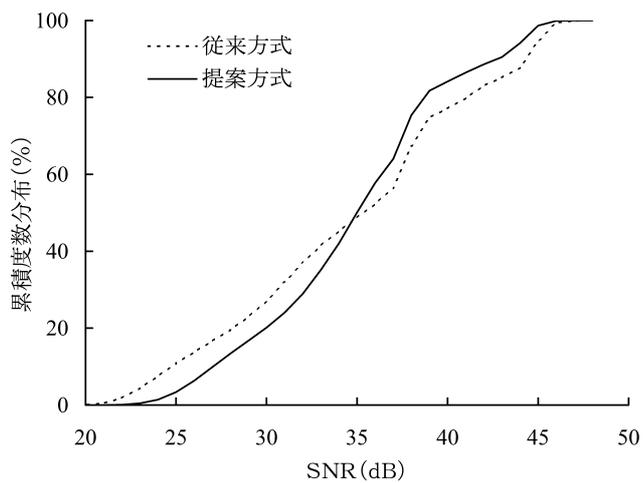


図 3.9 符号化歪み特性

Fig. 3.9 Characteristics of coding distortion.

平均ビットレートを 5Mb/s としている。それゆえ、各エンコーダの第 1 フレームの割り当てビットレートが 5Mb/s となっている。図 3.8 で最も特徴的なのは、画像”POPPLE”に対して割り当てられたビットレートの変化である。”POPPLE”は 75 フレームあたりまでは平坦な背景の中で鳥かごが回っており、それ以降のフレームでは回る鳥かごへズームアップする画像である。したがって、75 フレームまでは比較的符号化の易しい画像であるのに対して、75 フレーム以降は難しいシーンとなる。図 3.8 から、このシーンの移り変わりを反映して、割り当てられるビットレートが 75 フレームあたりから増加していることがわかる。

このような画像の性質に応じたビットレート配分を行うことで、画質の極度な劣化を防ぐ効果が現れる。それを示したのが図 3.9 である。図 3.9 では図 3.8 と同じ条件でシミュレーションを行い、画像のマクロブロック (16 × 16 画素) ごとに SNR を計算し、ある SNR 以下のマクロブロック数が全体の何%存在するかをグラフに示したものである。図 3.9 より、提案方式では 35dB 以下の劣化を従来方式 (すべて 5Mb/s で符号化) に比べて大幅に抑えていることがわかる。

### 3.6 結言

複数の映像プログラムを効率良く一括して伝送するマルチエンコーダシステムに関して検討を行った。

まず、比較的少数のエンコーダ数においても多重化効果を得るために、各エンコーダを統一的制御の下に動作させるシステム構成を提案した。次に、映像符号化に広く用いられている MC + DCT 符号化方式において、ビットレート対 MSE の関係がべき乗の関係で表現されることを示し、各エンコーダで発生する MSE の緩和を最小にするためのエンコーダ間ビットレート配分方式を導いた。また、時間的に変化する映像符号化特性に適応させるために、ビットレート配分を更新する方式を提案した。この提案方式を実映像に適用した計算機シミュレーション検証の結果、提案方式はシステム全体の符号化効率を向上させるだけでなく、符号化画像の局所的な画質の極度な劣化も抑制する効果のあることがわかった。本章では符号化歪みの尺度として MSE を用いたが、すべての議論は視覚特性など各種の重み付けを行った MSE[21] に対しても全く同様に適用できる。

本章の方式により、映像配信に必要な映像コンテンツを実時間で高能率かつ高画質で伝送することが可能となる。



## 第4章 非実時間映像伝送方式

### 4.1 緒言

ドラマや映画などのエンターテインメント映像コンテンツや、編集作業や吹き替え等を行って映像コンテンツを制作するのに必要な映像素材のように必ずしも即時性は要求されない非実時間映像伝送の需要がある。従来、こうした映像伝送はテープなどの物理的メディアの配送によって実施されてきたが、近年のインターネットの高速化と利用可能地域の拡大により、ネットワーク経由での伝送が求められている。

インターネットにおけるデータ伝送では、UDP と TCP が代表的なプロトコルとして用いられる。UDP は誤ったデータの再送メカニズムを持たないため、低遅延での伝送が可能という特徴があり、ビデオストリーミング等のリアルタイム伝送に用いられる。しかし、それゆえに受信側へのデータの到着が保証されず、受信した映像コンテンツの品質も損なわれることがある。一方、TCP は完全なデータの到着を保証するために、再送メカニズムを持っており、実時間性に関する保証がない代わりに、受信した映像コンテンツの品質は送信側と同一に保たれる。これらの理由から、本章で提案する非実時間映像伝送技術に TCP を用いることとする。

TCP には国際間の長距離伝送や無線通信路の利用など伝送路遅延が大きく、かつ、パケットロスが発生する場合にスループットが低下する問題がある。一つの解決方法として、TCP の機能拡張であるウィンドウスケールオプションや選択 ACK 法が提案されている [22]。また、別の解決方法として、複数の TCP コネクションを同時に利用することでそれらを合計したトータルのスループットを改善する方式が提案されている [23][24]。

Allman らの複数の TCP コネクションを用いる方式は、特に大きな遅延を含む衛星通信を対象として提案され、TCP の機能拡張を用いるのと比較して優れたスループット性能を達成することが示されている [23]。この方式は広く用いられているファイル転送プロトコル ftp の内部の TCP の置き換えとして提案されたものであり、ファイル伝送を高速化することが目的とされている。しかし、映像伝送の観点で必要とされる、受信途中での再生開始などの機能面での多様な要求に応えることができない。さらに、複数の TCP コネクションを用いる方式は、その生成するトラフィックがネットワークに対して急激な負荷変化をかける危険性が指摘されている [25]。本章では複数の TCP コネクションを用いて、映像伝送の機能要求に応え、さらにトラフィックのシェーピング機能を有する新しい非実時間映像伝送プロトコルを提案する。

## 4.2 映像伝送に対する要求条件

メディア伝送において、特に番組制作等の受信後の二次利用・加工に用いるための素材伝送を対象として、機能要件をまとめる。

- 1) **高速性** 利用できるネットワークの送信端・受信端間でのベアラ実効帯域を有効に利用し、メディア伝送を行う必要がある。高遅延環境等においてプロトコルの制約を原因として、実行帯域に比してスループットが低下してはならない。
- 2) **高信頼性** ベストエフォートネットワークでは、伝送誤りや実効帯域の変動が発生することがある。これらの問題を解決し、送信側のメディアデータが完全に受信側に到着する必要がある。
- 3) **高機能性** メディアデータの伝送を意図的に中断し、再開できることが必要である。これにより、複数のメディアデータがあるときに、伝送途中での伝送順序の変更など柔軟な伝送のスケジューリングが可能となる。
- 4) **実時間性** ネットワークの実効帯域が転送するメディアデータの符号化速度に比べて十分に大きいとき、ライブ伝送・再生が可能であることが望ましい。一方、ネットワークの実効帯域が十分ではない場合、メディアデータの既受信部分の再生が可能であることが望ましい。
- 5) **ユーザ認証機能** メディアデータの伝送開始前に、正規利用者によるアクセスか否かの判定のために端末の認証を行う必要がある。

## 4.3 データ伝送の原理

上記の要件を満たすために、本論文ではマルチ TCP コネクション型データ伝送モデルを用いたプロトコル SVFTP (Swift Very large File Transfer Protocol) を提案する。SVFTP の構成を図 4.1 に示す。

送信側ではハードディスクから所定のサイズのデータを読み出し、データパックを構成してキューに蓄積する。このとき、実際にはディスク I/O 速度が低速であるため、ディスク上のセクタサイズの整数倍の単位でデータをメモリ上へ読み出し、データパックを構成する。

次に、このデータパックを各 TCP/IP コネクションに割り当てる。このとき、割り当てられたデータパックの送信を終了したコネクションに対して次のデータパックを割り当てることで、コネクションがアイドル状態となることによるスループット低下を回避できる。

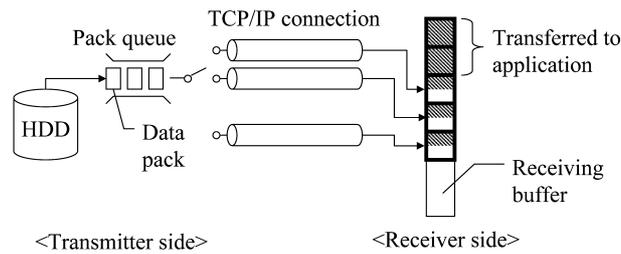


図 4.1 マルチコネクションデータ伝送のブロック図

Fig. 4.1 Block diagram of multi-connection data transmission.

また、あるコネクションが異常終了した場合には、そのコネクションに割り当てられていたデータパックを再度キューの先頭に配置することにより、次にアイドルとなったコネクションでそのデータパックが送信される。なお、伝送対象のファイルがエンコード途中であり、読み出し可能データが存在していない場合には、データパックに相当するデータの蓄積を待つものとする。

受信側では、各コネクションから到着するデータはデータパック単位で受信バッファへ入力され、整列して格納される。このため、データパックは図 4.2 に示すヘッダ情報を持つ。“type” フィールドは、このデータパックが伝送ファイル中の最後のデータであるか否かを示すフラグである。“sequence number” は、データパックを受信側で整列するための情報である。また、size により、データパックサイズを示すことができるため、データパックを可変長とすることも可能である。

ここで、受信バッファ上の仮想ファイルでは、コネクションによって伝送速度が異なりうるため、図 4.1 に示すように到着データが非連続状態となる。そのため、SVFTP 内部では連続データ領域の終点を示すポインタを持ち、連続部分のデータを上位プログラムからのリクエストに応じて書き出す構成としている。したがって、上位プログラムでは、SVFTP 内部のデータ整列やデータパック構造とは独立に構成可能である。

さらに、図 4.1 の構成をとることにより、伝送途中での TCP コネクション数の増減も容易に実現可能である。新たに追加されたコネクションは、送信側においてデータパックの送信の終了したコネクションとして扱われ、新たなデータパックが割り当てられ、データ伝送に供される。受信側ではデータパックのヘッダ情報を参照して仮想ファイルに書き込みを行うため、特にコネクションの区別をする必要はない。逆に、コネクション数を減じる場合は、割り当てられたデータパックの伝送を終了した TCP コネクションを切断するだけでよい。

なお、ワイヤレス向けに推奨されているラージウィンドウオプションについては、TCP コネクションを確立する際のオプション指定により SVFTP でも実装している。また、選択肯定応答については、オペレーティングシステムによる TCP の実装の中で実現されている。

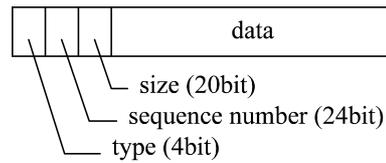


図 4.2 データパックフォーマット

Fig. 4.2 Data pack format.

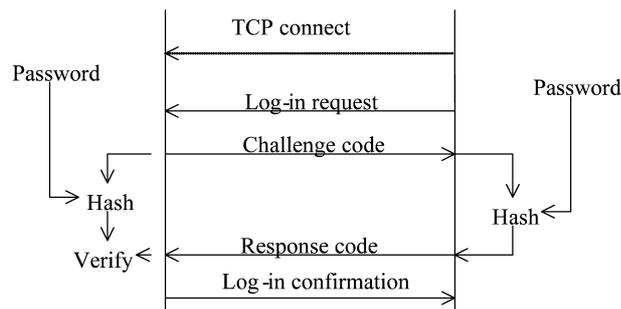


図 4.3 ログイン手順

Fig. 4.3 Log-in procedure.

## 4.4 プロトコル

### 4.4.1 通信手順

SVFTP は、重要なメディアデータ伝送に適用されることを考慮して、通信開始時に通信相手の認証サービスを提供する。図 4.3 に手順を示す。これは、一般的なチャレンジ・レスポンス手順に従ったもので、あらかじめ送受信端間でパスワードを共有しておくことが前提である。まず、ログイン要求を受けた側（サーバ）が乱数値を発生し、ログイン要求を発行した側（クライアント）に対して伝送する。両端において、その乱数とパスワードからハッシュ値を計算し、クライアントはその結果をサーバへ伝送し、サーバは自身の計算結果と比較して一致すればアクセスを認める。

SVFTP では、上記のログインからログアウトまでの処理をセッションと呼称する。次に送受信端間セッション管理のための TCP コネクションを確立する。このセッション管理コネクションを用いて、高速データ伝送のためのマルチ TCP コネクション開設・終結の制御を行う。このデータ伝送の動作のことをトランザクションと呼称する。従って、一つのセッションの中で複数のトランザクションが存在することがありえる。

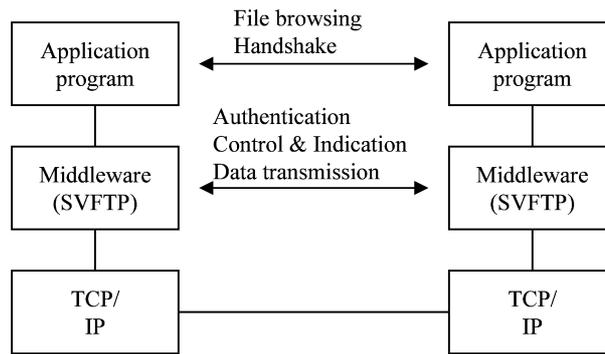


図 4.4 プロトコルスタック

Fig. 4.4 Protocol stack.

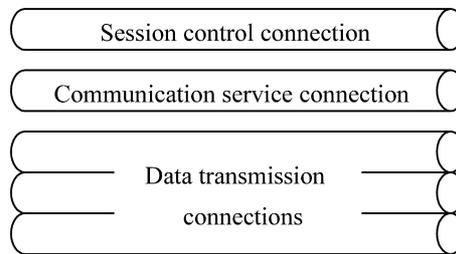


図 4.5 SVFTP における TCP コネクション

Fig. 4.5 TCP connections for SVFTP.

#### 4.4.2 アプリケーションレイヤ向け通信サービス

SVFTP のデータ伝送機能は、上位アプリケーションの観点からは、送信側の SVFTP バッファに書き込んだデータが、受信側の SVFTP バッファから出力されるものである。このとき、送信側のバッファ書き込みと受信側バッファ書き込みは非同期であり、送信側の上位アプリケーションは送信したデータが実際に受信側の上位アプリに到着したことを直接確認することはできない。また、ファイル転送用のアプリケーションを考えると、上位アプリケーションレベルでファイルの一覧確認を行う必要がある。そこで、図 4.4 に示すように、アプリケーション間での通信をサポートする必要がある。SVFTP 自体はミドルウェアとして動作し、アプリケーションからの指示に基づいて、相手方 SVFTP の認証、データ伝送、データ伝送に伴う制御と通知を行う。

以上の機能をまとめると、セッションを制御する機能、アプリケーションに対して通信サービスを提供する機能、そしてデータ伝送を行う機能が必要となる。SVFTP は、それぞれの機能のために個別に TCP コネクションを確立して使用している。図 4.5 に、SVFTP に用いられる TCP コネクションの総体を示す。

### 4.4.3 伝送レートシェーピング

マルチ TCP コネクションによりネットワークに送出されるトラフィックは、公平性の観点と輻輳の起因となる点から問題が指摘されている [25]。まず、公平性に関しては、TCP によるトラフィックの特性として理想的にはネットワーク中で利用可能な帯域を各 TCP コネクションで均等に分配するようにフロー制御が働くが、マルチ TCP コネクションでは  $N$  本のコネクションを同時に用いることで、他の TCP トラフィックの  $N$  倍の帯域を占有しようとするように働く。一方、輻輳に関しては、TCP における輻輳ウィンドウ拡大過程と関係がある。複数の TCP コネクションで同時に輻輳ウィンドウが拡大された場合、急激にネットワークに送出されるパケットが増加し、ネットワークに輻輳を引き起こす恐れがあるというものである。

これらに対処するために、SVFTP では伝送レートシェーピング機能を持つ。SVFTP の送信側システムでは、図 1 に示すデータパックキューは TCP のソケットバッファに空きがあると通知してきた TCP コネクションに対してデータパックを割り当てる。そこで、レートシェーピング機構をこの過程に導入し、下記のアлゴリズムに基づいてデータパックを実際に割り当てるまでに遅延を挿入する。

$T$  を SVFTP の送信開始時点からの経過時間、 $K$  を現時点までの総伝送データ量、 $t_x$  を前回のデータパック割り当て時刻から次のデータパック割り当てまでの時間、 $N$  をデータパックのサイズ (データ量)、そして  $v_{target}$  をシェーピングビットレートとする。このとき、ネットワークへの送信ビットレートをシェーピングビットレートに保つために、下記の等式が満たされなければならない。

$$v_{target} = \frac{K + N}{T + t_x} \quad (4.1)$$

実際に送信側システムで制御に用いるパラメータ  $t_x$  は、上式を解くことにより求められる。 $t_x$  の値が正の場合、送信ビットレートがレートシェーピング速度を超える状態にあることを示しているため、データパックは  $t_x$  秒遅延させてから TCP コネクションに割り当てられる。その間に新たに別のソケットバッファの空き通知があった場合は、その次のデータパックに関して改めて  $t_x$  を計算し、その結果に従ってデータパック割り当てを遅延させる。一方、 $t_x$  の値が負の場合、伝送ビットレートがレートシェーピング速度を下回っていることを示しているため、直ちにデータパックを TCP コネクションに割り当てる。

式 (4.1) は、一定時刻ごとに評価するのではなく、TCP コネクションからのソケットバッファ空き通知ごとに評価する。これは、一定時刻ごとに評価すると、データパックを送出すべき時刻に対して実際にデータパックを割り当てる時刻が遅れる現象が発生し、スループット低下要因となるためである。

## 4.5 アプリケーションの構成法

4.2節の要件を満たす上で、1)の高速性はマルチTCPコネクションの利用、2)の高信頼性はTCP内部での誤データ再送、5)ユーザ認証機能についてもSVFTPにて機能が提供されているので、これらの要件を満たすアプリケーションは容易に実現できる。一方、3)の高機能性(伝送の途中再開)と4)の実時間性(追いかけて再生)については、転送プロトコルと上位アプリケーションとの連携が不可欠である。そこで、これらを実現するための上位アプリケーションの構成法について述べる。

3)を実現するためには、上位アプリケーションは、再送開始個所を指示するポインタ(再送ポインタ)を記録し、それに基づいて送受両側が同じ位置からデータ伝送を再開する必要がある。受信側では、SVFTPから上位アプリケーションへ渡されるデータは整列済みであるため、上位アプリケーション側において受信済みのデータ位置を記憶すればよい。そして、伝送を再開する時点で、受信側の再送ポインタを送信側へ通知し、送信側のファイル読み出しをその箇所から行うことで途中再開が可能となる。

一方、4)の実現は、受信側のデータはSVFTP内部では完全には連続的なデータとはなっていないが、上位アプリケーションに渡されるのは連続となっているデータであるため、上位アプリケーションは渡されたデータの範囲でデコーダへデータを入力し、メディア再生を行うことができる。また、送信側においてもエンコード中のデータを順次SVFTPバッファに書き込むことで、通信帯域さえ十分にあれば、受信側の読み出し操作と合わせて、ライブ伝送によるストリーミングが実現可能となる。

## 4.6 伝送特性

本節では、伝送特性の理論特性と、実験結果の対比・評価を行う。

### 4.6.1 スループット特性の理論解析

TCP/IPのスループット特性に関する理論検討の中で、本論文ではPadhye[26]らによるモデル式を用いる。

$$B(p) = \min \left\{ \frac{W_{max}}{RTT}, \frac{L}{RTT \sqrt{\frac{2bp}{3} + T_o \min\{1, 3\sqrt{\frac{3bp}{8}}\}} p(1 + 32p^2)} \right\} \quad (4.2)$$

このモデル式では、パケットロス発生環境におけるウィンドウサイズの変化をモデル化し、その定常状態に基づいてスループットの理論値を求めている。TCPの輻輳制御メカニズムはRenoアルゴリズムを前提としており、高速再送ならびにタイムアウトをモデル式に組み入れている。

る。 $W_{max}$  は、最大ウィンドウサイズ、RTTは往復遅延時間、 $b$ は重複 ACK 法における重複数(一般的には $b=2$ )、 $p$ はパケットロス率、 $T_o$ はタイムアウト時間、 $L$ はパケットサイズである。式(4.2)右辺の最小値演算の左項は、パケットロス率が0または極めて小さい場合に有効となる。従って、以下の実験において、パケットロスが0の場合は理論値として左項を用いている。逆に、パケットロスを発生させた実験では、理論値として右項を用いている。

次に、マルチ TCP コネクションの場合、ネットワーク上で輻輳が発生しない限り、各 TCP コネクションは独立して働くとみなすことができる。そこで、マルチ TCP コネクションのスループットモデル式は、 $N$  コネクションの場合、

$$B_N(p) = B(p) * N \quad (4.3)$$

と書くことができる。ただし、エンコード中の映像ストリームのライブ伝送を行う場合、映像のエンコード速度が伝送スループットの上限となる。従って、(4.3)式は、エンコード速度を  $B_{Live}$  とすると、

$$B_N(p) = \min\{B_{Live}, B(p) * N\} \quad (4.4)$$

となる。

#### 4.6.2 実験結果

実験では、ネットワークエミュレータを用いて、遅延やパケットロスなどの環境を模擬している。実験に用いた機器は、以下のとおり。

- 送信側 PC: CPU Pentium 3 1.2GHz、メモリ 256MByte、Windows 2000
- 受信側 PC: CPU Pentium 3 1.2GHz、メモリ 256MByte、Windows 2000
- ネットワークエミュレータ: NTT エレクトロニクス社 NE-2000

##### (a) ファイル転送速度

まず、基本的な特性として、ファイルの転送開始から終了までに要した時間から求めたファイル転送スループットを測定する。ネットワーク環境としては、

- a) 往復遅延 150msec、パケットロス率 0%
- b) 往復遅延 150msec、パケットロス率 1%

とした。この往復遅延は、ほぼ日米間での伝送遅延に相当する。このとき、ウィンドウサイズとコネクション数を変化させて、スループットを測定した結果を図 4.6(a) (b) に示す。

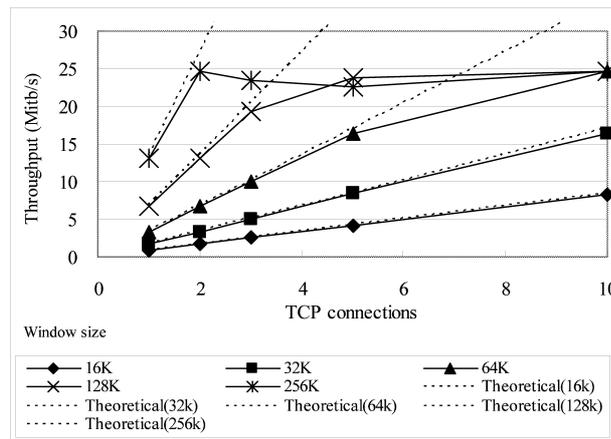
図 4.6(a) より、コネクション数に比例してスループットが増大することが分かる。また、ウィンドウサイズを大きくするほどスループットが高くなることも観測される。ただし、ネットワークエミュレータの処理能力が 25Mbps であるため、それがスループットの上限值となる。また、TCP のラージスケールオプションは、このグラフではコネクション数が 1 の時に、ウィンドウサイズによるスループットの違いを見ることで、効果が確認できる。ただし、128KByte のラージスケールの効果は、16KByte のウィンドウサイズで 8 本のコネクション、32KByte で 4 本、64KByte で 2 本の場合と同等である。また、理論スループットはウィンドウサイズと往復遅延から (4.3) 式によって求められるが、ネットワークエミュレータの処理能力の範囲内では、実験結果とよく一致していることがわかる。

一方、図 4.6(b) より、パケットロスの発生する環境では、ウィンドウサイズによるスループットの違いはほとんどないことが分かる。これは、パケットロスの発生により、TCP の輻輳制御メカニズムが働き、輻輳ウィンドウが最低限まで縮小されることによる。その場合でも、コネクション数を増やすことで、輻輳ウィンドウを拡大することと等価な働きとなり、スループットが改善されている。また、理論スループットにおいては、パケットロス率が高いため (4.2) 式の  $\min$  演算の右項が支配的となり、最大ウィンドウサイズとは無関係な特性となるため、実験結果と符合する。ただし、実験結果では最大ウィンドウサイズが小さい場合のスループットが低くなっていることが観察される。これは、輻輳ウィンドウが通信中に拡大・縮小の遷移を繰り返す際に、最大ウィンドウサイズが上限となり、理論式での輻輳ウィンドウサイズの平均値より小さくなるためと考えられる。以上の実験結果より、基本的なファイル伝送においては、コネクション数と最大ウィンドウサイズをともに増加させることが効果的であるといえる。

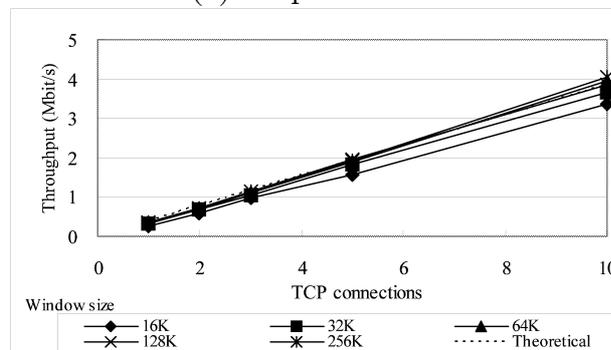
## (b) 追いかけて再生特性

次に、メディア伝送において重要な追いかけて再生機能について評価する。ここでは、送信側から 1.8Mbps の映像ストリームが SVFTP によって伝送され、受信側において、それを順次再生する動作を模擬している。送信側では映像データを 1.8Mbps の速度で SVFTP のバッファに書き込み、受信側では SVFTP のバッファを 10msec 間隔でポーリングし、到着済みデータを読み出す操作を行った。もし、SVFTP による伝送スループットが 1.8Mbps よりも高ければ、ライブ伝送が可能なシステムとなっている。

SVFTP パラメータは、ウィンドウサイズを 32KByte、コネクション数を 1, 2, 10 本とした。また、受信側上位アプリケーションヘデータを渡す単位は 64KByte 単位となっている。この



(a) No packet loss



(b) With packet loss (PLR: 1%)

図 4.6 スループット特性

Fig. 4.6 Throughput characteristics.

とき、受信側の上位アプリケーションに到着したデータについて、時間経過による総データ量の推移を図 4.7 に示す。

図 4.7(a) では、パケットロスが発生していないので、スループット特性は (4.2) 式の第一項に基づく (4.4) 式に従う。情報源速度が 1.8Mbps であるため、スループットの上限值も 1.8Mbps となるが、伝送開始直後には送信側においてデータがバッファリングされているため、1.8Mbps を超えた伝送速度が一時的に出る可能性がある。コネクションが 1 本の時のスループットは 1.7Mbps で、1.8Mbps の速度を示す理論直線よりもわずかに直線の傾きが小さくなっている。ただし、データの到着は 64KByte 単位に極めて安定していることが分かる。一方、コネクション数が 2 本の場合は、スループットは十分であるがデータの到着間隔が不均一であることが観察される。コネクション数が 10 本になると、スループット自体は送信側でのデータ供給が 1.8Mbps で抑えられることから、それがスループット上限値となる一方で、データの到着間隔がさらに不均一となることが分かる。

次に、図 4.7(b) ではネットワークで意図的にパケットロスが発生させており、スループット

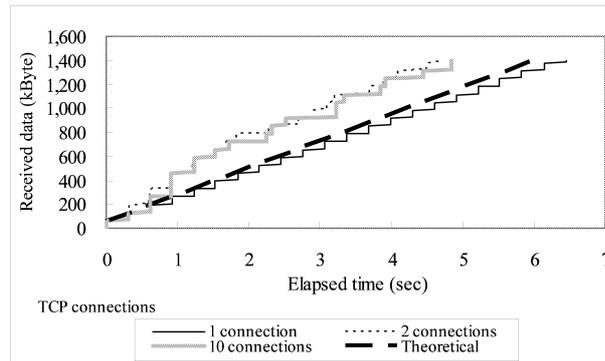
特性は(4.2)式の第二項に基づく(4.4)式に従う。このとき、コネクション数が1,2本では情報源速度に足りずリアルタイム伝送が不可能であるが、コネクション数を10本にすることにより、データ到着間隔が不均一であるもののリアルタイム伝送が可能となっている。従って、伝送路特性等の理由によりネットワークにおいてパケットロスが発生する場合には、コネクション本数を増加させてスループットを向上させる方法が有効であることが分かる。

複数コネクションの場合の不均一なデータ到着は、図4.1で説明したように、SVFTPの受信バッファにおいて受信データの整列を行った後に上位アプリケーションに引き渡すことによっている。すなわち、あるコネクションのデータ到着が遅れたとき、それ以外のコネクションのデータが到着済みであっても受信バッファに留め置かれ、遅れたデータが到着した段階で一時的に出力される。データ到着の遅れの原因としては、コネクション数が多くなるほど送信側におけるTCP伝送スレッド間の干渉が大きくなるためと考えられる。TCP伝送処理は、それぞれ独立なスレッドによって処理されているが、各スレッドは最大でその輻輳ウィンドウサイズ(ここでは32kByte)相当のデータの送信終了まではCPUを占有しており、その占有時間はコネクション1本あたりのスループット1.70Mbpsから約150msecである。あるスレッドの待ち時間は、コネクション数10本の場合、最大150msecの待ち時間×0~9(本)の間に分布することとなり、数100msecオーダーとなる。このことから、コネクション数は必要なスループットを達成できる最低限の本数にすることが最適であるといえる。

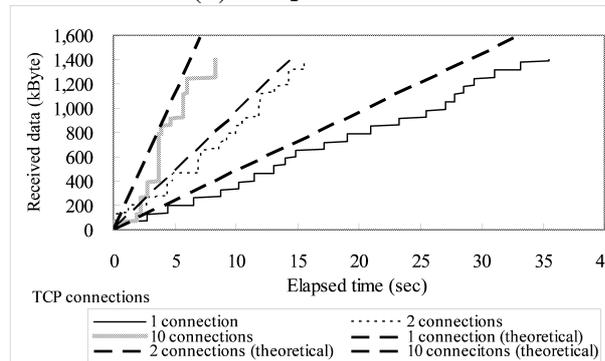
### (c) レートシェーピングの有効性評価

マルチTCPコネクション伝送における最大の懸念は、その帯域占有の貪欲さと生成されるトラヒックのバースト性に向けられている。そこで、マルチTCPコネクションSVFTPと他のトラヒックが伝送帯域を共有する場合に、SVFTPが他トラヒックに及ぼす影響を評価する。そのため、2対向のPCが10MbpsのEthernetハブ(リピータ)を共有したときのスループット特性を測定した。1対向はSVFTPによるファイル伝送を行い、もう1対向は通常のftpもしくは1Mbpsの固定速度UDPトラヒックによる伝送を行った。このとき、ネットワーク環境としては、4.4.3節のa)と同様に往復遅延150msec、パケットロス率0%としている。SVFTPのパラメータとしては、ウィンドウサイズを32KByteとし、コネクション数を変化させている。さらに、SVFTPにおいて3MbpsのレートシェーピングをON/OFFした場合の結果も合わせて示す。

まず、ftpトラヒックへ与える影響を図4.8(a)に示す。レートシェーピングを適用しない場合、コネクション数の増加に伴いスループットも増大するが、レートシェーピングを適用した場合には設定どおりの3Mbpsに収束していることが分かる。一方、ftpトラヒックに関しては、SVFTPトラヒックのない場合(コネクション数0)のスループットは、ネットワーク遅延



(a) No packet loss



(b) With packet loss (PLR: 1%)

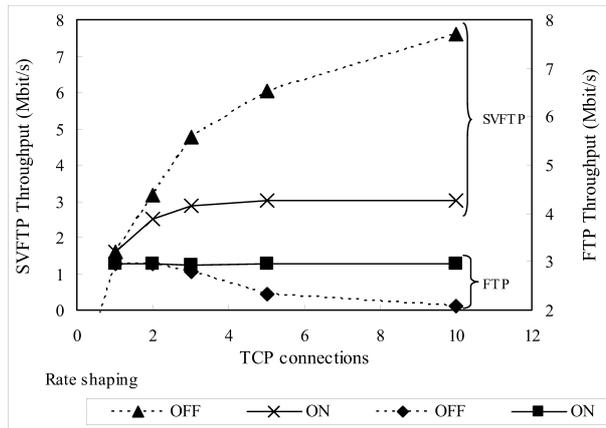
図 4.7 データ到着ジッタ

Fig. 4.7 Data arrival jitter.

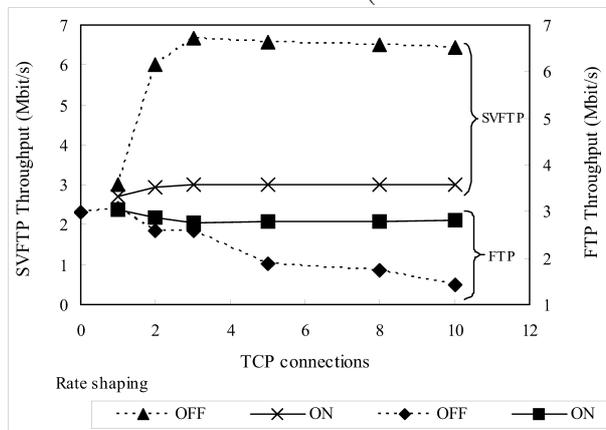
のために 3.0Mbit/s である。コネクション数が 10 本の SVFTP トラヒックと混在したときは、レートシェーピングを適用しないとき 2.1Mbit/s であるのに対して、レートシェーピングを適用した場合には 3.0Mbit/s となり、ftp スループットが大幅に改善された。

次に、SVFTP で用いる TCP ソケットバッファサイズを 64kByte へと拡大した場合のレートシェーピングの効果を評価する。図 4.8 (b) から分かるように、レートシェーピングがない場合はコネクション数の増加に伴い SVFTP のスループットが急速に 7Mbit/s まで上がっており、ネットワーク上で輻輳が発生し、ソケットバッファが小さい場合と比較して、ftp トラヒックがより大きな影響を受けている。レートシェーピングはこの場合でも有効であることが図 4.8(b) から読み取れる。SVFTP のスループットはコネクション数を増加させても正確に 3.0Mbit/s に固定されており、SVFTP コネクション数が 10 本の時の ftp トラヒックのスループットは 1.4Mbit/s から 2.8Mbit/s に回復しており、大きな改善が見られた。

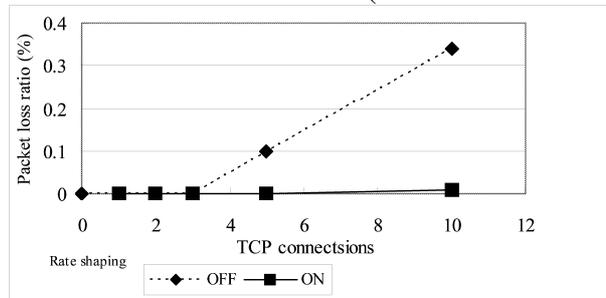
次に、UDP トラヒックへ与える影響を図 4.8(c) に示す。UDP トラヒックが他のトラヒックと混在したときの影響は、パケットロス率となって現れるため、図 4.8 (c) では SVFTP のコネクション数の増加に伴うパケットロス率の変化に注目する。その結果、レートシェーピング



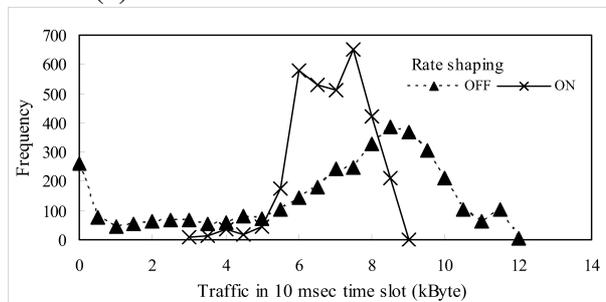
(a) Interaction with FTP traffic (Window size: 32kBytes)



(b) Interaction with FTP traffic (Window size: 64kBytes)



(c) Interaction with UDP traffic



(d) Data burst frequency distribution

図 4.8 SVFTP の他トラフィックへの影響

Fig. 4.8 Interaction between SVFTP and other traffic.

が非適用の場合にはコネクション数が増加するにつれてパケットロス率が上昇するのに対して、シェーピング適用時には非常に低く抑えられていることがわかる。これは、レートシェーピング非適用時にはSVFTPがより多くのネットワーク帯域を占有することが主因と考えられる。さらに副次的な要因として、SVFTPが出力するトラヒックのバースト性が緩和されることも一因である。これは、シェーピング非適用時には(輻輳ウィンドウサイズ)×(コネクション数)のデータが一斉に送出される可能性があるが、シェーピング適用時にはTCPソケットバッファのレベルで送信トラヒックが平滑化されるためである。

SVFTPのトラヒックのバースト性に関して評価を行う。2台のPCが150msecの往復遅延が挿入された10Mbit/sのリピータで接続されており、人為的なパケットロスを発生させていない環境で、10msecごとにネットワーク上のパケットをキャプチャし、そのタイムスロット単位でのデータ転送量を測定している。そして、そのデータ量の分布を階級幅を0.5kByteとしたヒストグラムで図fig:Interaction between SVFTP and other traffic(d)に示す。SVFTPのパラメータは、コネクション数が10、ソケットバッファサイズが32kByte、データパックサイズが1024bitである。レートシェーピング有無でのSVFTPの特性を比較するために、まずレートシェーピング無しでスループットを測定したところ5.5Mbit/sであった。そこで、レートシェーピング有りでのスループットも5.5Mbit/sに設定した。従って、レートシェーピング有りの場合の10msecタイムスロットにおけるデータ転送量の理論値は6.8kByte(=5.5Mbit/s×10msec)である。図4.8(d)から、レートシェーピング非適用の場合には小さなデータ転送量から大きなデータ転送量まで広く分布し、トラヒックのバースト性が高いのに対して、レートシェーピングを適用すると6から8kByteの間に分布のピークがあり、ほぼ一定間隔でデータが転送されており、トラヒックのバースト性が大きく緩和されていることが分かる。

以上の実験結果および議論から、マルチTCPコネクション伝送における問題はSVFTPの伝送レートシェーピング機能により解決されていることが示された。

## 4.7 システム実装例

### 4.7.1 システム構成

非実時間伝送を許容し、高品質な映像伝送を行うシステムとしては、蓄積伝送システムが最適である。蓄積伝送では、送信側は映像データを一旦ファイルに蓄積し、回線速度に合わせて伝送する。受信側では到着したデータをファイルに蓄積後、再生する。これにより、回線速度によらない高画質な映像伝送が可能となる[27]。そこで、蓄積伝送の概念に基づいて、MPEG-2ビデオコーデックとSVFTPによる伝送を組み合わせたシステムを開発した[28]。システム概要を図4.9に示す。本システムでは、入力されたAV信号をMPEG-2で圧縮後、ハードディスクに蓄積し、SVFTPを用いて伝送する。また、マルチTCPコネクションによる伝送

に加えて、伝送の途中再開機能や追いかけ再生機能を実現している。さらに、圧縮中のデータをファイル蓄積と平行してSVFTPによって送信し、受信側では受信中のデータをハードディスク蓄積と平行して復号・再生するパイプライン的処理によって、SVFTPスループットがビデオ符号化速度以上であれば、ライブ伝送も可能である。また、SVFTPのサービスチャネルを利用して、相手側のファイル一覧をブラウズし、必要な映像ファイルのみを伝送可能する機能も実装している。

#### 4.7.2 フィールド実験結果

##### 奄美大島無線ネットワーク

奄美大島における離島遠隔医療実験のための無線アクセスネットワークシステム [29][30]の上で、上記システムを用いた映像伝送実験を行った。同ネットワークは、奄美大島と周辺の3つの島をKDDI研究所のCFO-SS10Aによる無線リンクで結んだものである。奄美大島の古仁屋診療所と、与路島の僻地診療所との間で伝送実験を行った。CFO-SS10Aは2.4GHz帯の周波数を用いて、物理ベアラ速度10Mbpsで約10km離れた地点との無線リンクを構築することが可能である。実験では、5リンクを経由し、うち1リンクは11km隔たった海上を接続している。スループットの測定結果を図4.10に示す。図から伝送路の実効帯域は4Mbps程度であったことが分かる。また、パケットロス率は0.5%程度であり、伝送路の往復遅延は30msecであった。このとき、伝送路の実効帯域を利用するためには、TCPコネクションが10本必要であった。逆に、単純なTCP転送(コネクション1本)では、伝送誤りと伝送路の遅延のためにネットワークの実効帯域を使い切ることができないことがわかる。この結果から、ウィンドウサイズよりもコネクション数を増大させる効果が明らかであり、これは図4.6(b)の結果とも合致する。

##### 日米間IPv6ネットワーク

IPv6上で動作するSVFTPを用いて、WIDEを経由した日米間IPv6ネットワーク上で伝送実験を実施した。WIDEの日米間回線は155Mbpsの帯域であるが、実験で用いたアクセス回線が1.5Mbpsであったため、エンド・エンドでの利用可能帯域は1.5Mbpsである。このとき、SVFTPの最大輻輳ウィンドウサイズ設定は32kByte、ネットワークの往復遅延は143msecであり、またネットワークモニタリングソフトウェアiperfにより1.3MbpsのUDP負荷をかけた際のパケットロス率は0%であった。従って、理論的には(4.2)式の第二項に基づいた(4.3)式に従うことになるが、コネクション数が1本の時点で理論的には1.79Mbpsとなり利用可能帯域を超えるため、それを下回る範囲でのスループットとなっている。

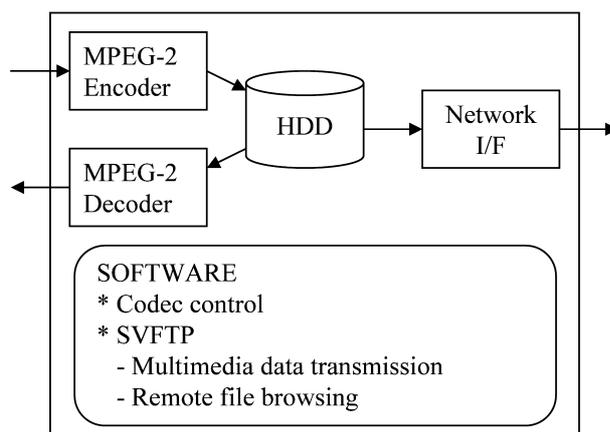


図 4.9 蓄積伝送システム

Fig. 4.9 Store-and-forward System.

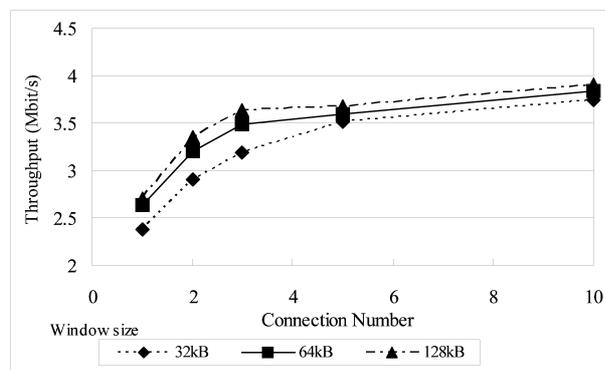


図 4.10 無線通信路におけるスループット特性

Fig. 4.10 Throughput characteristics over wireless network.

## 4.8 結言

IP ネットワークで高速かつ高信頼度なメディアデータの伝送を行うプロトコル SVFTP を提案した。SVFTP の特徴は、大容量メディアデータ伝送において必須となる、コンテンツの柔軟な配信スケジューリングを可能とする伝送の中断・再開機能を実現し、さらに受信途中のメディアデータにアクセスして「追いかけて再生」を可能とした点にある。また、マルチ TCP コネクション伝送の問題点であった過大なネットワーク占有に関しても、新たにレートシェーピング機能を実現することにより解決している。SVFTP の理論的なスループット特性について検討、室内実験による検証の結果、マルチ TCP コネクション伝送の効果を明らかにした。また、ストリーミング伝送を実現する際にはコネクション数の最適値が存在することも示した。SVFTP の映像伝送システムへの応用例も示し、実環境での試験により、その有効性も明らか

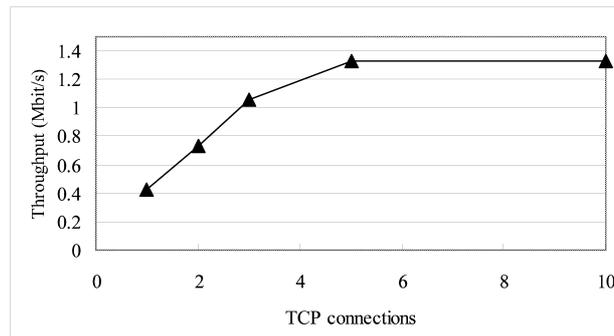


図 4.11 国際回線上でのスループット特性

Fig. 4.11 Throughput characteristics over international link.

となった。

## 第5章 高能率映像配信方式

### 5.1 緒言

インターネットの普及に伴い、その上での映像放送が注目されている [31][32]。放送型サービスを実現する通信技術としては、放送受信を要求している受信者に対してのみデータパケットを能率良く同報的に配信可能なマルチキャストが適している [33][34]。

しかしながら、マルチキャストにおいては同じパケットを通信参加者全員が受け取るようになるため、受信者毎に異なる要求に個別に対応することはできない。例えば、太いアクセス回線を持つ受信者向けに高ビットレートの映像を送信した場合、細いアクセス回線の受信者はパケットが途中で失われてしまうため正常に受信できない。そこで、複数のマルチキャストチャンネルを用いてデータ伝送を行い、受信側において必要なチャンネルを選択・受信する方式が提案されている [35][36]。

本章では、このような複数のマルチキャストチャンネルを用いる映像配信に適した映像の階層符号化ならびに伝送制御方式を提案する。映像の階層符号化方式として、時間方向の冗長除去を充分に行える MPEG 型の I、P、B フレーム間予測符号化と、復号処理の容易なデータパーティショニングを組合せて利用する階層符号化方式を提案する [37]。また、受信側における初期バッファリングを最小化して遅延を削減する一方で、安定した映像の受信・復号・表示を可能とするための伝送制御に関して検討する。さらに、画質を最適化するための階層化の制御方法 [38] についても検討を行っている。

### 5.2 マルチチャンネル伝送モデル

マルチキャスト [39] を用いることにより、ネットワーク帯域の有効利用が図られる反面、受信端末毎の利用可能帯域や受信・復号能力に関係なく一律に同じパケットが配送されてしまう。そこで、この問題を解決するため、複数の固定速度チャンネルを用いて、チャンネル毎に階層符号化したビデオデータをマルチキャストすることにより、受信端末毎に適合したビデオデータを受信する方式が提案されている [36]。図 5.1 に概念図を示す。図 5.1 においてビデオデータは5層に分割されており、5つのマルチキャストチャンネルにより伝送されている。このとき、利用可能帯域の小さいクライアント3は1つの階層しか受信できないため、それに対応するチャンネルを要求するメッセージ [40] をルータ2に送り、その結果ルータ1からルータ2へルーティングが行われている。また、クライアント1は利用可能帯域が大きいので全階層を受信できる一方で、クライアント2は復号能力が低く2つの階層しか受信できないため、それらの要求に応じたルーティングが行われている。

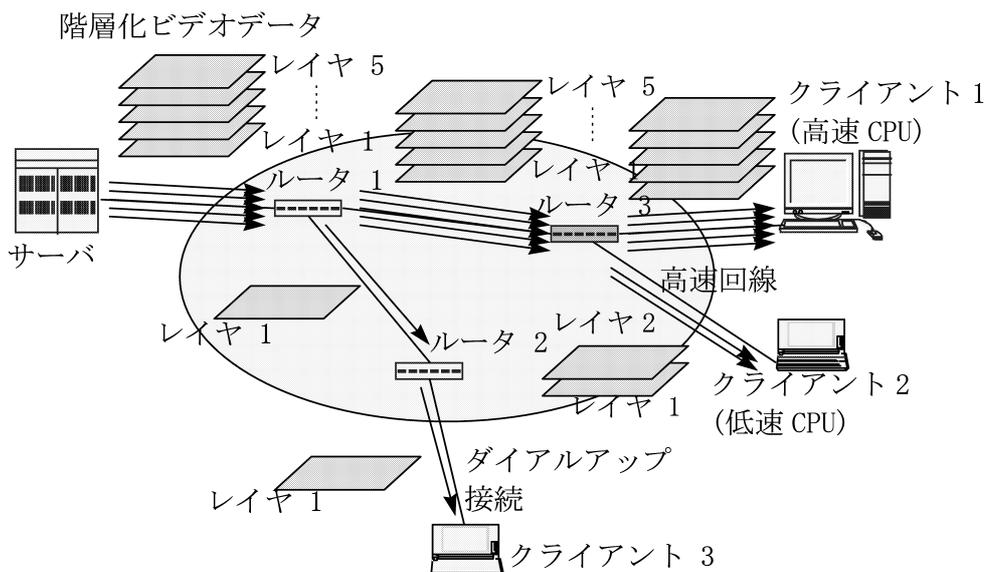


図 5.1 複数のチャンネルを用いたマルチキャストビデオ伝送モデル

Fig. 5.1 Video multicast model using multiple channels.

伝送システムの制御という観点からは、この方式の利点は、受信側からの選択によって伝送帯域および画像品質が制御され、送信側はそれを意識することなく単純にマルチキャストを行えばよいことである。すなわち、受信端末の数が増減する場合でも送信側に変更は必要なく、参加・離脱する受信端末がメッセージをルータに送るだけでよい。したがって、この方式は参加者数に関わりなく動作することの可能なスケラブルなシステムとなる。

### 5.3 ビデオ階層符号化方式

本節ではビデオマルチキャストに適した階層符号化方式について検討し、その階層データのマルチチャンネル伝送モデルへの適用方式について述べる。

#### 5.3.1 各種階層符号化方式の比較

静止画用の階層符号化方式であるラプラシアンピラミッドを適用した方式 [41] や、符号化対象フレームが前フレームと変化している度合いをブロックごとに判定し、変化が大きい場合にそのブロックをイントラ符号化する条件付きブロック更新方式 [42] が提案されている。これらの方式は、誤りの時間波及がないため誤り耐性に優れている。しかし、いずれもフレーム間の予測誤差を符号化するのではなく、直接イントラ符号化しているため符号化効率は H.261 [43] 等の方式に劣る。

一方、H.261 では時間方向の冗長を十分に除去して符号化を行う動き補償予測 DCT 符号化を採用しているので符号化効率は高い。しかし、符号化データの階層的な表現ができないことや、いったん符号化データが失われると、その影響が後続のフレームに波及するという大きな欠点がある。

MPEG-1[44] では、時間方向の冗長除去として H.261 と同じ動き補償予測 DCT 符号化を採用しているが、比較的短い周期で I フレーム（フレーム内符号化）を挿入しているため、誤りの波及を防止することができる。また、予測構造に P フレーム（片方向予測）および B フレーム（両方向予測）を利用しているため、それぞれの従属関係によって階層的に表現することが可能となっている。さらに、MPEG-2[45] においては空間解像度に関する階層化方式として、データパーティショニングや各種スケーラビリティが規定されている。これらの空間解像度階層化は、時間解像度の階層化と独立に適用することができる。ただし、スケーラビリティ方式は符号化・復号ともに DCT や量子化操作を 2 段構成で適用するため、それに伴う演算量が 2 倍となってしまい、パーソナルコンピュータでのソフトウェアによる実現が困難である。

### 5.3.2 本章で採用する階層符号化方式

符号化効率と誤り耐性の両方を満足する符号化方式としては、上記の考察から、データ損失の影響の時間的波及を防止し、かつ時間方向の階層化を同時に実現できる MPEG 型の I, P, B フレーム間予測 DCT 符号化が最も優れている。このとき、I フレームは単独で復号できるが、P フレームの復号には I フレームの復号が事前に必要で、B フレームの復号には P フレームの復号が事前に必要である。したがって、I フレームを基本階層、P フレームをその高位階層、そして B フレームをさらにその高位階層とすることで時間解像度の階層化を実現することができる。

次に、時間方向に階層化された各フレームに対して空間解像度に関する階層化を行う。本章では、空間解像度の階層化に復号処理の単純なデータパーティショニングを用いる。データパーティショニングとは、符号化される DCT 係数を低周波数部と高周波数部に分離する方式である。低周波数部はそれ自身で復号できるが、高周波数部は低周波数部と組合せなければ復号できない。このとき、 $8 \times 8$  マトリックスで表現される DCT 係数はジグザグスキャンされているので、そのスキャンの途中で分離・合成を行うことになる。したがって、スケーラビリティとは異なって DCT・量子化演算は 1 段でよい。なお、P および B フレームでは動き補償のための動ベクトルデータを含んでおり、これは復号のために必須のデータであるから低周波数部に含まれる [37]。

表 5.1 チャンネルへの階層データの割り当て

Table 5.1 Allocation of layered video data to each channel.

チャンネル	Case 1	Case 2
1	$I_L$	$I_L + P_L$
2	$I_H$	$I_H + P_H$
3	$P_L$	$B_L$
4	$P_H$	$B_H$
5	$B_L$	—
6	$B_H$	—

### 5.3.3 階層データの伝送チャンネルへのマッピング

上記により、ビデオデータは時間解像度に関して I, P, B フレーム種別による分割、空間解像度に関して動ベクトルと DCT 低周波係数の組および、DCT 高周波係数への分割が行われることとなった。次に、これらを完全に個別のチャンネルで伝送する方式、あるいはそれらを組合せて伝送する方式が考えられる。表 5.1 にその例を示す。ここで、 $I_H$  とは I フレームの DCT 高周波係数、 $P_L$  とは P フレームの動ベクトルと DCT 低周波係数を意味する。その他の記号についても同様である。

表 5.1 の Case 1 は、分離可能な階層ごとにチャンネルを割り当てているため、受信者側の細かな要求に対応することが可能である。ただし、Case 1 ではチャンネル 1, 2 は I フレーム専用のチャンネルであるため、チャンネルの空き時間を作らないようにするには、低い伝送速度で I フレーム挿入間隔に相当する時間をかけて 1 フレームの I フレームを伝送する必要がある。このため、Case 1 では遅延が大きくなる欠点がある。一方、Case 2 では  $I_L$  と  $P_L$ 、 $I_H$  と  $P_H$  というように周波数部ごとに組合せることで遅延を削減することが可能となっている。これは、チャンネル 1, 2 で I と P フレームを伝送するため、Case 1 よりもチャンネル 1, 2 の伝送速度が速くなっており、I フレームの伝送に要する時間が短縮されることによる。この点については、5.4.5 節でも述べる。

なお、受信・復号するときには必ずしもチャンネル番号の順序で受信するチャンネル数を増やす必要はない。例えば、Case 1 では、チャンネル 1, 3, 5 の組合せによる復号も可能である。ただし、P フレームを復号せずに B フレームを復号することができないなど、階層間の従属関係に基づく制約があることに注意が必要である。

ところで、ここで述べた階層符号化・伝送方式では、階層データが個別のチャンネルで伝送される環境下で I, P, B 構造から要求される復号時刻の相互関係を満たさなくてはならず、符号化伝送・復号システムとして安定に動作させるための条件は複雑である。従来、このような

問題を解析した報告はないため、次節において復号器バッファに注目したシステム安定動作条件について解析し、伝送制御方式を提案する。

## 5.4 システムの安定動作条件

本節では、5.3節で述べたビデオ階層符号化方式をマルチチャンネル伝送モデルに適用した際に、システムが安定動作するための条件を受信バッファの解析を通じて導出する。具体的には、受信バッファがアンダーフローもしくはオーバーフローすることによる復号エラーの発生を避けることを目的として、符号化時の各階層画像のフレームあたり発生情報量と受信バッファ容量との関係を定式化する [38]。これにより、リアルタイム符号化・伝送時には、伝送開始時に定まっているバッファの制約に対応して符号化制御を行うことが可能となる。あるいは、符号化済みのビデオデータをリアルタイム伝送する場合には、既知の発生情報量から受信側に必要なバッファ量を定めることが可能となる。

### 5.4.1 解析の前提条件

受信側バッファの解析を行うにあたり、ネットワークにおける遅延および遅延変動はないものとする。これは、ビデオ階層符号化とバッファの関係に絞ることで解析を明確化するためである。

次にビデオ符号化に関して、GOP パラメータ  $N = 15$ ,  $M = 3$  である MPEG-1 を用いることとし、受信側でビデオ復号および表示は瞬時に行われるとする。このとき、空間解像度の階層化については、例えば  $P_L$  と  $P_H$  のように同一フレーム種別であれば復号時刻は同じであるので、一般式の導出過程では考慮しない。したがって、時間解像度の階層化のみ考慮すればよいので、表 5.1 の Case 1, 2 にしたがって、I, P, B を独立チャンネルあるいは I と P を同一チャンネルで B を別チャンネルとする場合についてバッファ解析を行うこととする。

### 5.4.2 各チャンネルの解析

Case 1

#### (a) Iチャンネル

Iフレームの場合は、15フレームごとにデータがバッファからデコーダへ移されるので、バッファ蓄積量の変化は以下の漸化式により表現できる。

$$d_j^{[1]} = d_{j-1}^{[1]} + R_1 T - q_j \quad (5.1)$$

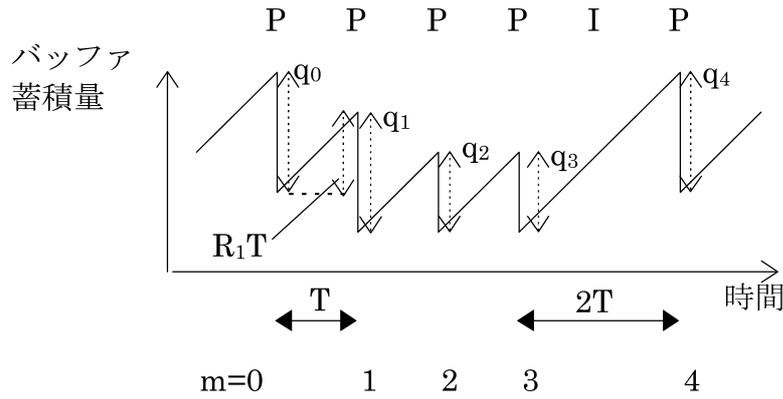


図 5.2 Case 1、P チャンネルに対するバッファ遷移

Fig. 5.2 Buffer state transition for P-channel of Case 1.

ここで、 $j$  は I フレームごとに増加するインデックス、 $d_j^{[1]}$  は  $j$  におけるデータ移送直後のバッファ蓄積量、 $q_j$  は移送されるフレーム情報量、 $R_1$  はチャンネル 1 の伝送速度、 $T$  は I フレーム間隔 ( $T = 15\tau$ )、 $\tau$  はフレーム間隔 (1/30 秒) である。また、 $D_0^{[1]}$  を Channel 1 の移送される前のデコード開始に必要なバッファ初期蓄積量とすると、以下の一般式で表される。

$$d_j^{[1]} = D_0^{[1]} + jR_1T - \sum_{i=0}^j q_i \quad (5.2)$$

一方、バッファ蓄積量に対して、1 つ前の I フレーム復号でバッファがアンダーフローしない条件  $d_{j-1}^{[1]} \geq 0$ 、かつ、後続の I フレームでバッファがオーバーフローしない条件  $d_j^{[1]} + R_1T \leq D_{max}^{[1]}$  が課せられる。ただし、 $D_{max}^{[1]}$  はバッファ容量である。これらの条件に (5.1) 式を組合せることにより、以下の条件式を得る。

$$R_1T - q_j \leq d_j^{[1]} \leq D_{max}^{[1]} - R_1T \quad (5.3)$$

### (b) P チャンネル

受信バッファの遷移を図 5.2 に示す。P フレームの復号タイミングは一定間隔ではなく、I フレームの復号が途中にあるため 4 つの P フレーム毎にバッファデータ移送間隔が広がる。そこで、P フレームのバッファ蓄積量は以下の漸化式で表される。

$$\begin{cases} d_m^{[2]} = d_{m-1}^{[2]} + 2R_2T - p_m & (m = 4i) \\ d_m^{[2]} = d_{m-1}^{[2]} + R_2T - p_m & (\text{otherwise}) \end{cases} \quad (5.4)$$

ここで、 $m$  は P フレームインデックス、 $p_m$  は P フレーム情報量、 $T$  は P フレーム間隔 ( $T = 3\tau$ )、そして  $i$  は非負整数である。これより、一般式は以下で表される。ただし、 $[x]$  は  $x$  を超

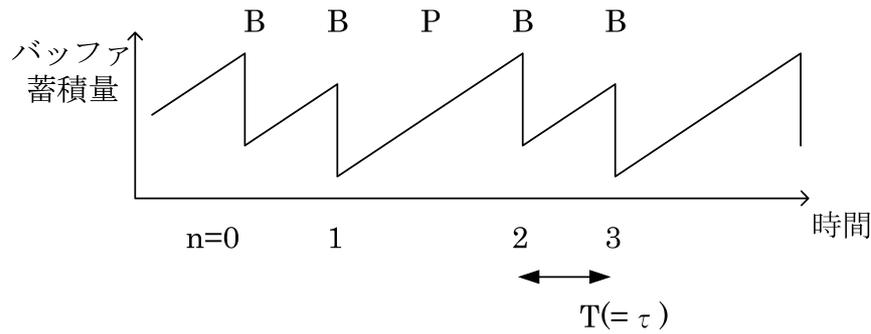


図 5.3 Case 1、B チャンネルに対するバッファ遷移

Fig. 5.3 Buffer state transition for B-channel of Case 1.

えない最大整数である。

$$d_m^{[2]} = D_0^{[2]} + mR_2T + \lfloor m/4 \rfloor R_2T - \sum_{i=0}^m p_i \quad (5.5)$$

また、バッファ蓄積量は I フレーム前後の P フレームに対する場合分けによって、次の制約を受ける。

$$\begin{cases} R_2T - p_m \leq d_m^{[2]} \leq D_{max}^{[2]} - 2R_2T & (m = 4i + 3) \\ 2R_2T - p_m \leq d_m^{[2]} \leq D_{max}^{[2]} - R_2T & (m = 4i) \\ R_2T - p_m \leq d_m^{[2]} \leq D_{max}^{[2]} - R_2T & (\text{otherwise}) \end{cases} \quad (5.6)$$

### (c) B チャンネル

B フレームの場合は、図 5.3 に示すように 2 つの B フレームごとにバッファデータ移送タイミングが変更されることに注意して、P フレームの場合と同様に以下の一般式と制約式を得る。ここで、 $n$  は B フレームインデックス、 $r_n$  は B フレーム情報量、 $T$  は B フレーム間隔 ( $T = \tau$ )、そして  $i$  は非負整数である。

$$d_n^{[3]} = D_0^{[3]} + nR_3T + \lfloor n/2 \rfloor R_3T - \sum_{i=0}^n r_i \quad (5.7)$$

$$\begin{cases} R_3T - r_n \leq d_n^{[3]} \leq D_{max}^{[3]} - 2R_3T & (n = 2i + 1) \\ 2R_3T - r_n \leq d_n^{[3]} \leq D_{max}^{[3]} - R_3T & (n = 2i) \end{cases} \quad (5.8)$$

## Case 2

### (a) I+P チャンネル

この場合は、1 つのチャンネルで I フレームと P フレームを伝送するので、図 5.4 に示すように 1GOP 時間中に 1 つの I フレームと 4 つの P フレームを伝送することになる。したがって、バッファ蓄積量の一般式は

$$d_j^{[1]} = D_0^{[1]} + jR_1T - \sum_{i=0}^{\lfloor j/5 \rfloor} q_i - \sum_{i=0}^{j-\lfloor j/5 \rfloor-1} p_i \quad (5.9)$$

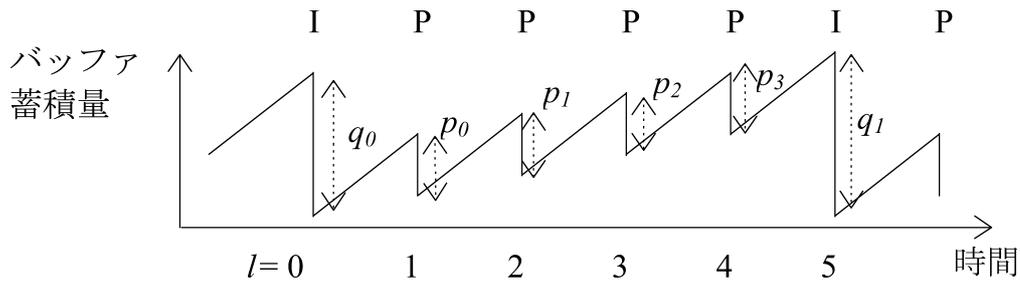


図 5.4 Case 2、I+P チャンネルに対するバッファ遷移

Fig. 5.4 Buffer state transition for I+P-channel of Case 2.

となる。ここで、 $j$  は I もしくは P フレームごとに増加するインデックス、 $d_j^{[1]}$  は  $j$  におけるデータ移送直後のバッファ蓄積量、 $q_i$  は I フレーム情報量、 $p_i$  は P フレーム情報量、 $R_1$  は I+P チャンネルの伝送速度、 $T$  は P フレーム間隔 ( $T = 3\tau$ )、 $\tau$  はフレーム間隔 (1/30 秒)、 $D_0^{[1]}$  は I+P チャンネルのバッファ初期蓄積量である。また、バッファの制約条件は、I フレームと P フレームの復号タイミングでの場合分けから、以下のようになる。

$$\begin{cases} R_1T - q_j \leq d_j^{[1]} \leq D_{max}^{[1]} - R_1T & (j = 5i) \\ R_1T - p_j \leq d_j^{[1]} \leq D_{max}^{[1]} - R_1T & (\text{otherwise}) \end{cases} \quad (5.10)$$

#### (b) B チャンネル

Case 1 の B チャンネルと同じ。

### 5.4.3 伝送制御方式および遅延の解析

各フレームタイプごとの復号処理のタイミングを図 5.5 に示す。図 5.5 において、インデックス  $j, m, n$  が値を持つ時、そのインデックスに該当するフレームデータがバッファから取り出され、その時刻  $k$  において復号される。図 5.5 は、チャンネル内での復号時刻を示しているのに加えて、各チャンネル間の復号時間関係をも示しており、時間インデックス  $k$  を用いて各チャンネル間の時間関係を以下のように表すことができる。

Case 1 の場合

$$\begin{cases} k = 15j \\ k = 3(m+1) + 3\lfloor m/4 \rfloor \\ k = n + 4 + \lfloor n/2 \rfloor \end{cases} \quad (5.11)$$

Case 2 の場合

$$\begin{cases} k = 3j \\ k = n + 4 + \lfloor n/2 \rfloor \end{cases} \quad (5.12)$$

時間インデックス :k	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
表示画像フレーム	-	-	-	I <sub>0</sub>	B <sub>0</sub>	B <sub>1</sub>	P <sub>0</sub>	B <sub>2</sub>	B <sub>3</sub>	P <sub>1</sub>	B <sub>4</sub>	B <sub>5</sub>	P <sub>2</sub>	B <sub>6</sub>	B <sub>7</sub>	P <sub>3</sub>	B <sub>8</sub>	B <sub>9</sub>	I <sub>1</sub>	
復号画像フレーム	I <sub>0</sub>	B*	B*	P <sub>0</sub>	B <sub>0</sub>	B <sub>1</sub>	P <sub>1</sub>	B <sub>2</sub>	B <sub>3</sub>	P <sub>2</sub>	B <sub>4</sub>	B <sub>5</sub>	P <sub>3</sub>	B <sub>6</sub>	B <sub>7</sub>	I <sub>1</sub>	B <sub>8</sub>	B <sub>9</sub>	P <sub>4</sub>	
case	チャンネル	インデックス																		
1	I	j	0													1				
	P	m			0		1			2			3						4	
	B	n				0	1		2	3		4	5		6	7		8	9	
2	I+P	j	0			1		2			3			4			5			6
	B	n					0	1		2	3		4	5		6	7		8	9

図 5.5 各チャンネルでの復号タイミング

Fig. 5.5 Decoding timing for each channel.

ここで、それぞれのチャンネルにおける復号開始時刻は  $j, m, n = 0$  とし得られる  $k$  の値である。なお、図 5.5 から B フレームは復号と同時に表示されるが、I, P フレームは復号から表示までに  $3\tau$  時間かかることに注意する。

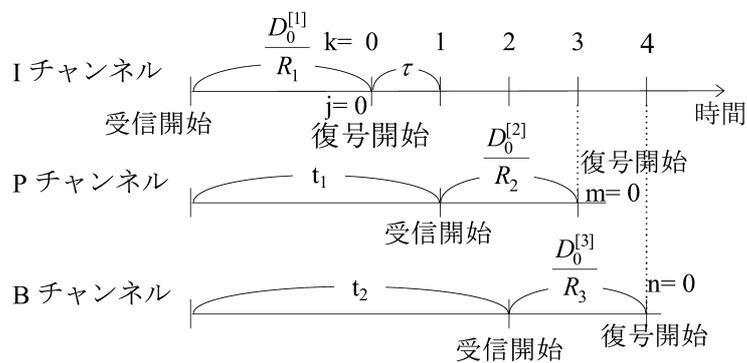
次に、各チャンネルの受信開始時刻を導くため、復号開始時刻と初期バッファ蓄積時間についてチャンネル間の相互関係を正確に記述したものを図 5.6 に示す。図 5.6(a) において、まず復号開始時刻については、図 5.5 で  $I_0$  と  $P_0$  の時間差が  $3\tau$  であることから、P チャンネルは I チャンネルの復号開始から  $3\tau$  秒だけ後から復号を開始しなければならないことがわかる。次に、P チャンネルの受信を開始すべき時刻は、バッファに  $D_0^{[2]}$  蓄積されるのに要する時間だけ復号開始よりも前の時刻となる。以下、B チャンネルや、図 5.6(b) についても同様に時間関係が規定されている。

さて、図 5.6 から、受信開始（すなわち送信開始）時刻は Case 1 のチャンネル I-P 間で  $t_1$ 、チャンネル I-B 間で  $t_2$ 、Case 2 のチャンネル (IP)-B 間で  $t_3$  だけ遅らせる必要があることがわかる。 $t_1, t_2, t_3$  はそれぞれ以下の式で表される。

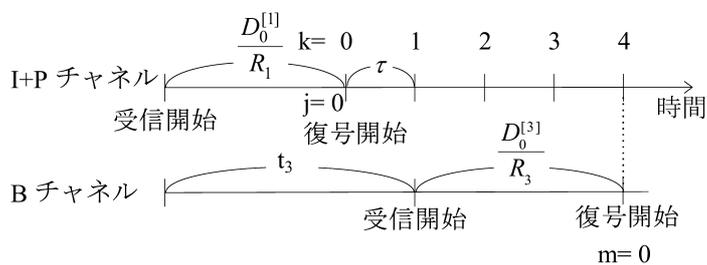
$$t_1 = \frac{D_0^{[1]}}{R_1} + 3\tau - \frac{D_0^{[2]}}{R_2}, \quad t_2 = \frac{D_0^{[1]}}{R_1} + 4\tau - \frac{D_0^{[3]}}{R_3} \quad (5.13)$$

$$t_3 = \frac{D_0^{[1]}}{R_1} + 4\tau - \frac{D_0^{[3]}}{R_3}$$

一方、マルチチャンネル伝送に伴う遅延は、伝送遅延や遅延変動吸収のためのバッファを無視すると、I フレームの受信開始から、それを表示するまでに要する時間である。いま、復号が瞬時に行われると仮定しているので、復号してから表示するまでの時間が  $3\tau$  であるから、



(a) Case 1



(b) Case 2

図 5.6 チャンネル間の時間関係

Fig. 5.6 Time relation between channels in the decoding process.

遅延は次式となる。

$$\frac{D_0^{[1]}}{R_1} + 3\tau \quad (5.14)$$

#### 5.4.4 フレーム発生情報量とバッファサイズの理論解析

チャンネル間に必要な時間差や遅延の定性的な振る舞いを解析するため、フレーム発生情報量と  $D_0$  や  $D_{max}$  のバッファサイズとの関係について検討する。ここでは、最も単純な系である Case 1、Iチャンネルについて解析する。(5.2),(5.3) 式より、以下の不等式が成り立つ。

$$\begin{aligned} D_0 &\geq \sum_{i=0}^{j-1} q_i - (j-1)R_1T \\ D_{max} &\geq D_0 + (j+1)R_1T - \sum_{i=0}^j q_i \end{aligned} \quad (5.15)$$

また、 $R_1$  に関して、長期的にはネットワークで伝送されるデータ量とフレーム発生情報量の総和が等しくならなければならないので、次の関係式が成り立つ。

$$R_1 = \frac{\sum_{i=0}^{N-1} q_i}{NT} \quad (5.16)$$

このとき、時間経過に対するフレーム発生情報量の累計は図 5.7 のようになる。(5.15) 式を図 5.7 を用いて解釈すると、発生情報量累計のグラフが  $R_1t$  による直線に対して上方に位置するほど  $D_0$  の値が大きくなり、逆に発生情報量累計のグラフが  $R_1t$  による直線に対して下方に位置するほど  $D_{max}$  の値が大きくなる。なお、図 5.7 における  $R_1t$  による直線は、ネットワークから届くデータ量の累計である。ここで、直線よりもグラフが上方に位置する状況はネットワークから届くよりも多くのデータがデコードに必要なことを意味し、 $D_0$  はバッファのアンダーフローを防ぐためのバッファ量であることを考えると、 $D_0$  を大きくすることにより、あらかじめそのデータを蓄積していると思なすことができる。逆に、直線よりもグラフが下方に位置する状況は、ネットワークから届いたデータがデコーダへ移送されずにバッファに残っていることを意味し、その分だけ  $D_{max}$  を大きくしなければならないことになる。

上記の考察により、 $D_0$  と  $D_{max}$  を共に小さくするためには、 $q_i = R_1T$  とすればよい。このとき

$$D_0 = R_1T \quad (5.17)$$

$$D_{max} = D_0 = R_1T \quad (5.18)$$

となる。

以上では、Case 1、Iチャンネルの場合について考察したが、定性的傾向については他の場合にも当てはまること (5.5)(5.6),(5.7)(5.8),(5.9)(5.10) 式のそれぞれの組合せからわかる。

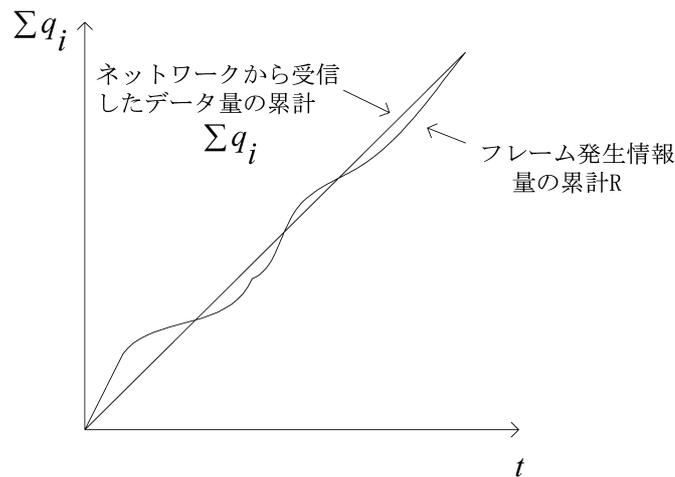


図 5.7 フレーム発生情報量累計の時間経過

Fig. 5.7 Cumulative amount of frame data versus time elapse.

#### 5.4.5 計算機シミュレーションによる評価

次に、計算機シミュレーションにより遅延値等の具体的数値例を挙げる。ここでは、MPEG-1 ( $N = 15, M = 3$ ) を TM5[46] 型のレートコントロールにより動作させ、1.2 Mbit/s で CIF 画像 Salesman 420 フレームを符号化して得られたフレーム別発生情報量  $q_j, p_m, r_n$  に基づいて、必要バッファ量や遅延を求める。まず、フレーム別発生情報量から各チャンネルの総発生情報量を求め、それを伝送に要する時間 (420 フレームに相当する時間) で除することにより、伝送レートを算出する。このようにして得られた各チャンネルの伝送レート  $R_1, R_2, R_3$  を用いて、初期バッファ蓄積量  $D_0$ 、バッファ容量  $D_{max}$ 、チャンネル間の時間差  $t_1, t_2, t_3$ 、および遅延を求めた結果を表 5.2 に示す。

表 5.2 から、Case 2 の方が遅延を大幅に小さくできることがわかる。これは、Case 1 では 1 GOP 時間で 1 枚の I フレームを伝送するため、必ず 500 ms 以上の遅延を生じる一方、Case 2 では 1 GOP 時間で 1 枚の I フレームと 4 枚の P フレームを伝送するため、発生する遅延を低く抑えられるためである。その代償として、Case 2 ではチャンネル選択の自由度が Case 1 よりも小さくなっているが、遅延を減少させる観点からは Case 2 が有利である。

また、バッファ量については Case 1 のチャンネル I, P よりも Case 2 の I+P チャンネルの方が、 $D_0, D_{max}$  共に大きくなっている。これは、5.4.4 節での考察を当てはめると、Case 2 の場合には I と P フレームを合わせることでフレーム発生情報量のばらつきが増えたため、それを吸収するためにバッファサイズが大きくなったと考えられる。特に  $D_0$  に対して  $D_{max}$  の変化が大きいのは、P フレームの復号中に送られてくる次の I フレームデータを長時間バッファに保持しておかなくてはならないためである。

表 5.2 必要バッファ量と遅延

Table 5.2 Amount of required buffer and delay.

(a) Case 1

チャンネル	伝送レート (kbit/s)	$D_0$ (kByte)	$D_{max}$ (kByte)
I	$R_1 = 270$	16.8	16.9
P	$R_2 = 558$	14.6	19.2
B	$R_3 = 372$	3.3	7.1

$t_1 = 391\text{ms}$ ,  $t_2 = 563\text{ms}$ , 遅延 600 ms

(b) Case 2

チャンネル	伝送レート (kbit/s)	$D_0$ (kByte)	$D_{max}$ (kByte)
I+P	$R_1=828$	17.5	25.8
B	$R_3=372$	3.3	7.1

$t_3 = 232\text{ms}$ , 遅延 269ms

## 5.5 階層間レート配分

本節では、提案方式の Case 2 について、画質最適化の観点から階層間のビットレート配分の検討を行う。

### 5.5.1 最適化の対象

いま、MPEG-1 であらかじめ符号化されたビットストリームから、データパーティショニングにより階層データを作成することを想定する。このとき、I, P, B ピクチャそれぞれの符号化ビットレートは所与であるから、階層分割の際のパラメータは DCT 係数のジグザグスキャンにおける低次・高次の分離次数となる。特に、チャンネル 1 は I ピクチャの低周波成分と P ピクチャの低周波成分から構成されるため、チャンネル 1 で利用できるビットレートを両者の間で配分する必要がある。チャンネル 2 に関しては、I, P ピクチャの高周波成分であるから、チャンネル 1 のビットレート配分から自動的に決まる。チャンネル 3 については、B ピクチャの低周波成分だけから構成されるので、使用できる伝送チャンネル速度から決定される。したがって、階層間のビットレート配分に関して最適化が必要なのは、チャンネル 1 の I ピクチャと P ピクチャ間の関係である。

### 5.5.2 階層間レート配分の実験的検討

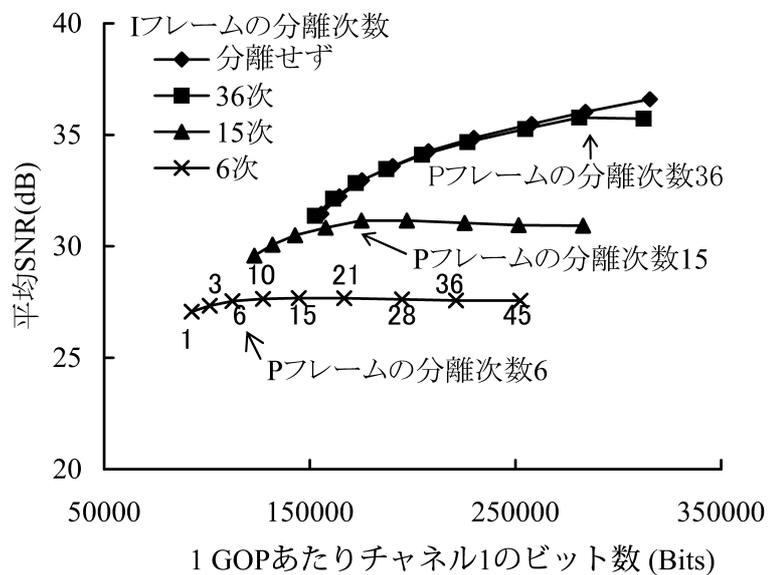
一般に、データパーティショニングにより分離した DCT 高次係数を伝送途中で廃棄すると、廃棄それ自体による SNR 劣化と、送信側と受信側で動き補償に用いるフレームバッファの内

容がミスマッチすることによる SNR 劣化が生じる。いま、I フレームと P フレームの集合から廃棄する高次係数ビット数を一定に保つとすると、I フレームから多く廃棄する場合にはミスマッチによる SNR 劣化が大きくなり、P フレームから多く廃棄する場合には廃棄自体による劣化が大きくなると予想される。

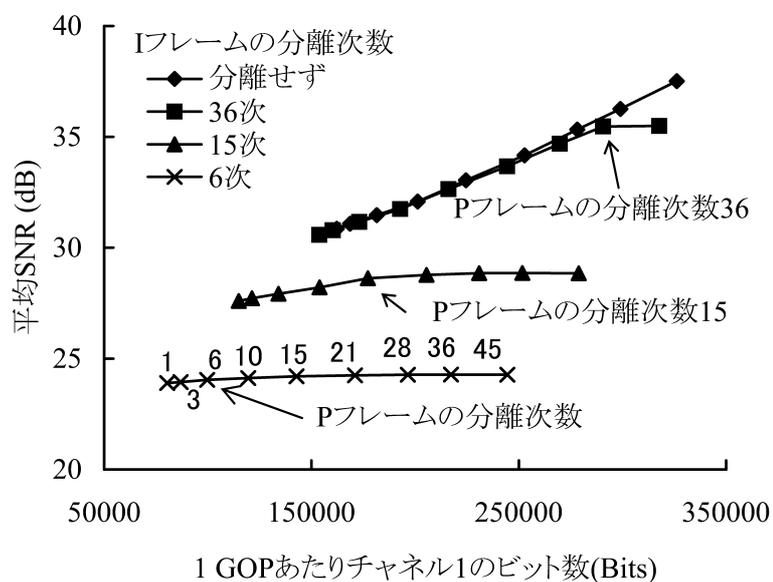
そこで、MPEG-1 ( $N = 15, M = 3$ ) で符号化したピクチャについて、I ピクチャのスキヤン分離次数と P ピクチャのスキヤン分離次数を独立に変動させたときの P ピクチャの SNR について計算機シミュレーションによって調べた。このとき、MPEG-1 の符号化ビットレートは全体として 1.2 Mbit/s、I フレームのビット数は 150 kbit/frame、P フレームのビット数は 70 kbit/frame、B フレームのビット数は 17 kbit/frame である。図 5.8 に CIF 画像 Salesman と Container について、4 通りの I ピクチャのスキヤン分離次数について、P ピクチャのスキヤン分離次数を変化させてプロットしたグラフを示す。このグラフでは、横軸には I ピクチャと P ピクチャの低周波成分を合わせた 1 GOP あたりビット数、縦軸には P ピクチャの SNR を 1 GOP で平均した値をとっている。したがって、ある横軸上の一点において、最も SNR の高くなるような I, P ピクチャのスキヤン分離次数の組合せのとき、最適なビットレート配分となる。ここで分離次数は、DCT 係数をジグザグスキヤンする順番に番号を付与し、指定した番号以上の高次係数をチャンネル 2 へ移すことを意味するので、分離次数が高次であるほどチャンネル 1 に残るビット数が増える。

図 5.8 で、例えば I 分離次数が 15 のときには P 分離次数が 15 より高次側である間は SNR に変化がなく、P 分離次数をそれよりも低くすると SNR が低下することがわかる。また、I 分離次数が 36 のときには、P 分離次数を 36 よりも小さくしてビット数を減じると SNR が低下する一方、それよりも分離次数を高次へ変更し、ビット数を増加させても SNR が変化していない。以上の観察は図 5.8(a), (b) 双方で成り立ち、画像によらないことから、I ピクチャで廃棄される係数の次数よりも高次の P ピクチャ係数は P ピクチャの SNR 向上には寄与しないことがわかる。また、どちらの画像についても、I ピクチャの係数分離を行わない場合が、チャンネル 1 のビット数によらず SNR が最大になっている。したがって、最適なレート配分は P ピクチャのみデータパーティショニングを行って低次係数をチャンネル 1 へ、高次係数をチャンネル 2 へ割り当て、I ピクチャについてはデータパーティショニングを行わず、すべての I ピクチャデータをチャンネル 1 へ割り当てることであると結論できる。

上記の結果は、フレーム間予測符号化において、参照画像に送受でのミスマッチによる雑音が含まれているときに、その雑音よりも高い周波数成分の雑音はマスクされてしまい影響を及ぼさなくなること、そして、最初の参照画像に与えられたミスマッチ雑音は、後続の P ピクチャによる予測差分値の加算では回復不能であることを示している。



(a) Salesman



(b) Container

図 5.8 階層間のレート配分実験結果

Fig. 5.8 Experimental result of bitrate allocation between layers.



図 5.9 復号画像

Fig. 5.9 Decoded picture.

### 5.5.3 受信画質

提案方式の Case 2 の場合、例えば MPEG パラメータ  $N = 15, M = 3$  では、チャンネル 1 から復号すると 10 frame/s で高周波成分のないぼやけた画像、チャンネル 2 を加えると 10 frame/s で精細な画像となる。さらに、チャンネル 3 を加えると 30 frame/s で、ぼやけた画像となり、チャンネル 4 の追加により精細な画像となる。

ここで、主観的な画質評価を行うことを目的として、チャンネル 1 だけから復号した P フレーム画像を図 5.9 に示す。この画像は、GOP 中の最後の P フレームであるため、最もデータパーティショニングによるミスマッチの影響を受けている画像である。符号化パラメータとしては、MPEG-1 ( $N = 15, M = 3$ ) 1.2 Mbit/s で符号化しており、I フレームは 150 kbit/frame、P フレームは 70 kbit/frame、B フレームは 17 kbit/frame でそれぞれ符号化されている。階層化パラメータは、P フレームの係数分離次数 6 で、 $P_L$  画像は 25 kbit/frame である。したがって、チャンネル 1 は 1 秒あたり 2 枚の I フレームと 8 枚の  $P_L$  フレームから構成されるので 500 kbit/s である。また、チャンネル 2 は 8 枚の  $P_H$  フレームなので 360 kbit/s となる。さて、図 5.9 からは、若干のブロック状の雑音が見られるほかは、全体に低域通過フィルタが適用されたような画質となっており、ミスマッチによる極度の画質劣化は生じていないことが確認できる。

## 5.6 結言

インターネット上で効率良く、受信者の資源に応じてフレキシブルに映像放送を行うための映像階層符号化方式を検討した。その結果、MPEG に見られるような I, P, B フレーム構造を

利用した時間解像度の階層化と、データパーティショニングによる空間解像度の階層化を組み合わせる方式が適していることを示した。そして、この方式をマルチチャンネル伝送に適用した場合のシステムの動作解析により、符号化制御および受信・復号制御が満たすべき条件を定式化した。また、画質最適化の観点から階層間のビットレート配分方法について実験的に検討を行った。

これにより、通信網を用いて複数のユーザに対して高能率な映像配信サービスを提供することが可能となる。



## 第6章 結 論

本論文では、高信頼かつ高能率な映像配信サービスにおいて、コンテンツの制作から配信までの各過程におけるコンテンツハンドリング技術の確立を目的として、四つのテーマについて研究を行った。それらは、MPEG 符号化に対応した電子透かし(第2章)、実時間でのコンテンツの高能率な伝送(第3章)、非実時間でのコンテンツの高能率な伝送(第4章)、通信網を利用した最終利用者向けのコンテンツの高能率配信(第5章)である。

第2章では、MPEG 符号化データを画像に復号することなく高速に電子透かしを検出するために、MPEG 符号化と共通の DCT ブロックベースの透かし埋め込み・検出方式を提案した。MPEG 符号化データを利用する電子透かしの問題点として、攻撃に対する耐性の低さと検出結果の信頼性の低さがあった。それに対して、提案方式では、時間的空間的に分散された位置にある DCT 係数を多数利用し、その係数集合から統計的に透かしを検出する手法を用いることにより、透かしの検出性能を高めることができた。また、その係数集合における観測値の分布から、透かし検出結果の信頼性を測定する手法も同時に提案しており、透かしの誤検出の抑制も達成した。これにより、映像配信におけるコンテンツの著作権保護を高信頼かつ高能率に実現することが可能となった。

第3章では、ライブ映像中継において一般的な固定速度の通信路を用いて、複数の映像信号を多重化して高能率に伝送する技術を提案した。従来の多重化伝送方式では 25 チャンネルもの多重化が必要とされていたのに対して、提案手法では映像の符号化器を能動的に制御することにより、3チャンネルの多重化から大きな効果を達成している。最適な制御を実現するために、映像符号化におけるレート歪み特性を観測値から同定し、多重化チャンネルのレートの総和が一定の条件下で歪みの総和を最小化する問題へと定式化を行った。これを映像シーンの変化に伴って動的に解くことにより、常に最適な伝送を行うことができる。

第4章では、実時間性は必ずしも要求されない映像番組の高能率な伝送を実現するための映像ファイル高速伝送プロトコルを提案した。この伝送プロトコルの基礎となっているのは一般的に広く使われている到達確認型手順に基づく TCP であるが、その本質的な問題点である大遅延・高エラー環境におけるスループット性能の低下に対して、複数の TCP コネクションの同時利用による解決方法を示している。さらに映像ファイル伝送に用いるために、受信途中での映像データの再生を実現するためのデータ整列手法や、他のトラヒックと共存するためのレートシェーピング機構の提案を行い、シミュレーションによる効果を確認している。日米間の実回線においても、その効果は確認されている。

第5章では、インターネットを用いて効率よく複数の視聴者に映像を配信するための映像データの階層型マルチキャスト方式を提案した。従来の階層符号化伝送では時間方向の冗長

除去が十分でなく、通信帯域に対して画質が低い問題があった。それに対して、時間方向にも高能率な符号化を行う MPEG で圧縮されたコンテンツを階層的に分割し、階層ごとにマルチキャスト配信を行う手法により画質を改善した。ただし、MPEG 階層データでは階層間に時間的な依存関係が生じるため、受信側でのバッファにより時間差の吸収を行う必要がある。そこで、その系が安定して動作するための理論解析を行い、具体的な初期バッファ値ならびに最大バッファ量の算出法を与えている。

以上のように本論文では、高信頼かつ高能率な映像配信サービスを実現するためのコンテンツハンドリング技術を提案した。特に高能率性を実現する上で、コンテンツ素材の収集や蓄積、伝送にあたって、圧縮が広く使われていることを考慮する必要があった。そこで、MPEG 圧縮を例に挙げ、それと親和性の高い電子透かし技術や、符号化制御を組み込んだ高能率多重化伝送手法、そして MPEG ファイルを元にした配信手法を提案した。また、高信頼性に関しては、様々な攻撃にも耐えられる電子透かし技術や、ならびにインターネット経由で確実に映像コンテンツを届ける映像伝送手法により、著作権者の権利保護と映像利用者における画質保証が実現された。以上のことをまとめると、本論文で提案した各研究テーマにおける研究成果は、映像配信サービスの高信頼化および高能率化に貢献し得ると考えられる。

## 謝 辞

本研究をまとめるに際して御指導、御鞭撻、ならびに様々な御配慮を賜った神戸大学工学部電気電子工学科 田中初一教授に謹んで深謝の意を表します。

本研究を論文としてまとめるにあたって貴重な御教示と多くの御助言を頂きました神戸大学工学部情報知能工学科 上原邦昭教授、神戸大学工学部電気電子工学科 沼 昌宏教授に深く感謝致します。

また、本研究の機会を与えて頂き、研究の途上で終始御熱心、御懇切に御指導、御助言を頂いた KDDI 株式会社技術開発本部 村上仁己本部長、株式会社 KDDI 研究所 浅見徹所長、松本修一取締役、中島康之執行役員に深く感謝致します。

KDDI 株式会社開発推進部 和田正裕部長には、本研究全過程に涉り卓越した見通しを持って御指導、御教示を賜った。深く感謝致します。また、KDDI 研究所マルチメディア通信グループ 滝嶋康弘グループリーダーには、日頃より多大の御助言と励ましを賜った。謹んで感謝の意を表します。

最後に、本研究を進めるにあたり、様々なご協力を頂いた同研究所マルチメディア通信グループの皆様には感謝の意を表します。



## 参考文献

- [1] ISO/IEC JTC1/SC29/WG11 13818-2 Information Technology — Generic Coding of Moving Pictures and Associated Audio Information: Video, 1995.
- [2] J. Linnartz and J. Talstra, “MPEG PTY-Marks: Cheap Detection of Embedded Copyright Data in DVD-Video,” Proc. ESORICS’98, pp.221–240, Sept. 1998.
- [3] J. Dittman, M. Stabenau, and R. Steinmetz, “Robust MPEG Video Watermarking Technologies,” Proc. ACM Multimedia’98, pp.71–80, Sept. 1998.
- [4] F. Hartung and B. Girod, “Watermarking of Uncompressed and Compressed Video,” Signal Processing, vol.66, no.3, pp.283–302, May 1998.
- [5] B. Chitprasert and K.R. Rao, “Human Visual Weighted Progressive Image Transmission,” IEEE Trans. Commun., vol.38, no.7, pp.1040–1044, July 1990.
- [6] 数井, 酒澤, 滝嶋, 和田, “DCT 係数操作に基づく電子透かし方式における埋め込みしきい値に関する検討,” 2000 年電子情報通信学会総合大会, D-11-157, p.157, March 2000.
- [7] F. Petitcolas, R. Anderson, and M. Kuhn, “Evaluation of Copyright Marking Systems,” Proc. IEEE Multimedia Systems (ICMCS’99), vol. 1, pp. 574–579, June 1999.
- [8] ISO/IEC JTC1/SC29/WG11, MPEG 93/457, Test Model 5, April 1993.
- [9] R. Barnett, D. E. Pearson, “Frequency Mode LR Attack Operator for Digitally Watermarked Images,” Electronics Letters 19 (1998) pp. 1837-1839, September 1998.
- [10] M. Kutter, F. A. P. Petitcolas, “A fair benchmark for image watermarking systems,” in Proc. SPIE Security and Watermarking of Multimedia Contents, vol. 3657, pp. 226-239, January 1999.
- [11] P.G. Hoel, Elementary Statistics, 2nd Edition, Wiley, 1966.
- [12] F. Petitcolas, R. Anderson, and M. Kuhn, “Information Hiding — A survey,” Proc. IEEE, vol.87, no.7, pp.1062–1078, July 1999.
- [13] 岸野, 真鍋, 安田, “パケット符号化を考慮した映像信号の統計的性質,” Proc. PCSJ’87, 8-5, 1987.
- [14] T. Koga, Y. Iijima, K. Iinuma, and T. Ishiguro, “Statistical Performance Analysis of an Inter-frame Encoder for Broadcast Television Signals,” IEEE Trans., COM-29, 12, pp.1868-1876, Dec. 1981.

- [15] 酒澤, 和田, “マルチエンコーダシステムにおける符号化制御に関する検討,” テレビ学技報 ICS94-47, 18, 39, 1994.
- [16] 酒澤, 和田, “MPEG-2 マルチエンコーダシステムにおけるビットレート配分に関する検討,” 1994 テレビ年次大, 11-11, 1994.
- [17] ISO/IEC JTC1/SC29/WG11: MPEG93/457(Test Model 5)
- [18] N. Farvardin and J. Modestino, “Optimum Quantizer Performance for a Class of Non-Gaussian Memoryless Sources,” IEEE Trans., IT-30, 3, pp.485-497, May 1984.
- [19] 酒澤, 滝嶋, 和田, “MC+DCT 符号化器におけるレート歪み特性とそのマルチエンコーダシステムへの応用,” 信学技法, IE94-127, 1995.
- [20] N.S. Jayant and P. Noll: Digital Coding of Waveforms, Prentice-Hall, 1984.
- [21] 浜田, 松本, “画像の局所変化度による雑音マスキング効果を考慮した直交変換係数の最適量子化法,” 信学論 B-I, J75-B-I, 12, 1992.
- [22] WAP Forum, “Wireless Profiled TCP,” Version 31, March, 2001.
- [23] M. Allman, C. Hayes, H. Kruse, and S. Ostermann, “TCP Performance over Satellite Links,” Proceedings of the 5th International Conference on Telecommunication Systems, March, 1997.
- [24] M. Allman, H. Kruse, and S. Ostermann, “An Application-Level Solution to TCP’s Satellite Inefficiencies,” Proceedings of the First International Workshop on Satellite-Based Information Services (WOSBIS), Rye, NY, November, 1997.
- [25] M. Allman, S. Dawkins, D. Glover, J. Griner, D. Tran, T. Henderson, J. Heidemann, J. Touch, H. Kruse, S. Ostermann, K. Scott, and J. Semke, “Ongoing TCP Research Related to Satellites,” RFC2760, IETF, February 2000.
- [26] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, “Modeling TCP Throughput: A Simple Model and its Empirical Validation,” Proceedings of ACM SIGCOMM’98, September, 1998.
- [27] M. Wada, Y. Takishima, and S. Sakazawa, “High Quality Remote Video Duplication System for Packet Networks,” Proceedings of AVSPN97, pp.49-54, Aberdeen, Scotland, UK, 1997.
- [28] 酒澤, 滝嶋, 山下, 和田, “IP ネットワークによる高画質映像伝送システムの開発,” 映像情報メディア学会技術報告, NIM’99-86, vol.23, no.50, pp.1-6, August 1999.
- [29] 福家, 蕨野, 石川, 杉山, 篠永, “奄美大島の離島遠隔医療実験における CFO-SS 無線ネットワーク,” 2001 年電子情報通信学会ソサイエティ大会, B-5-197, September 2001.

- [30] 北辻, 藤川, 山崎, “奄美大島の離島遠隔医療実験における無線 LAN の TCP 性能の評価について,” 2001 年電子情報通信学会ソサイエティ大会, B-5-199, September 2001.
- [31] インターネットマガジン編集部: “最新リアルタイム・オーディオ&ビデオツール,” インターネットマガジン, pp.282-295, Jan. 1997.
- [32] S. McCanne and V. Jacobson: “vic: A Flexible Framework for Packet Video,” in Proc. ACM Multimedia’95, pp.511-522, 1995.
- [33] M.R. Macedonia and D.P. Brutzman: “Mbone Provides Audio and Video Across the Internet,” IEEE Computer, pp.30-36, Apr. 1994.
- [34] 日経エレクトロニクス編集部: “ILJ, IP マルチキャストの実験サービスを開始,” 日経エレクトロニクス, No.694, 1997.7.14, p.23, 1997.
- [35] S. McCanne, V. Jacobson and M. Vetterli: “Receiver-driven Layered Multicast,” in Proc. SIGCOMM’96, pp.117-130, ACM, 1996.
- [36] N. Shacham and H. Yokota: “Admission Control Algorithms for Multicast Sessions with Multiple Streams,” IEEE J. Sel. Areas Commun., vol.15, no.3, pp.557-566, Apr. 1997.
- [37] S. Sakazawa, Y. Takishima, M. Wada and K. Amano: “Bandwidth-Adaptive Video Transmission Method for Heterogeneous Network Environment,” pp.49-54, Int. Workshop on New Video Media Technology, 1997.
- [38] 酒澤, 滝嶋, 和田: “IP ビデオマルチキャストにおける画質制御方式の検討,” 信学技報, IE97-93, pp.19-24, Nov. 1997.
- [39] マルチメディア通信研究会編: “通信プロトコル辞典,” pp.156-165, アスキー出版局, 1996.
- [40] W. Fenner: “Internet Group Management Protocol Version 2,” Internet Eng. Task Force, Internet Draft, Nov. 1997.
- [41] N. Chaddha: “Software Only Scalable Video Delivery System for Multimedia Applications over Heterogeneous Networks,” in Proc. Int. Conf. Image Processing, Vol.1, pp.25-28, 1995.
- [42] S. McCanne, M. Vetterli and V. Jacobson: “Low-Complexity Video Coding for Receiver-Driven Layered Multicast,” IEEE J. Sel. Areas Commun., **15**, 6, pp.983-1001, Aug. 1997.
- [43] ITU-T: Recommendation H.261 “Codec for Audiovisual Services at px64 kbit/s,” 1990.
- [44] ISO/IEC JTC1/SC29/WG11 11172-2 Information Technology — Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1,5 Mbit/s : Video, 1993.

- [45] ISO/IEC JTC1/SC29/WG11: 13818-2 Information Technology — Generic Coding of Moving Pictures and Associated Audio Information : Video, 1995.
- [46] ISO/IEC JTC1/SC29/WG11: MPEG93/457 Test Model 5, 1993.



## 関連発表

### 学術論文

- [1] S. Sakazawa, Y. Takishima and M. Wada, "A Watermarking Method Retrievable from MPEG Compressed Stream," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol.E85-A, no.11, pp.2489-2497, Nov. 2002. (2003年度電子情報通信学会論文賞受賞)
- [2] 酒澤茂之、滝嶋康弘、和田正裕、羽鳥好律, "マルチエンコーダシステムによる動画像符号化伝送制御方式," *テレビジョン学会誌*, vol.50, no.7, pp.958-964, 1996年7月.
- [3] S. Sakazawa, Y. Takishima, Y. Kitatsuji, Y. Nakajima, M. Wada and K. Hashimoto, "Video Data Transmission Protocol "SVFTP" using Multiple TCP Connections and its Application," *IEICE Trans. Inf. & Syst.* (2004年11月29日採録通知)
- [4] 酒澤茂之、滝嶋康弘、和田正裕, "ビデオマルチキャストにおける符号化伝送制御方式," *映像情報メディア学会誌*, vol.52, no.6, pp.863-870, 1998年8月.

### 国際学会

- [1] S. Sakazawa, Y. Takishima, M. Wada and Y. Hatori, "Coding Control Scheme for a Multi-Encoder System," *Proc. of 7th International Workshop on Packet Video*, pp. 83-88, Mar. 1996.
- [2] S. Sakazawa, Y. Takishima and M. Wada, "A Watermarking Method Retrievable from MPEG-2 Compressed Stream," *Proc. of the Picture Coding Symposium 2001*, pp.77-80, Apr. 2001.
- [3] S. Sakazawa, Y. Takishima, Y. Nakajima, M. Wada and K. Hashimoto, "Multimedia Contents Management and Transmission System "VAST-web" and Its Effective Transport Protocol "SVFTP" ," *Proc. of 2004 International Conference on Multimedia and Expo*, 0-7803-8603-5/04, June 2004.

## 学術講演

- [1] 酒澤茂之, 浜田高宏, 松本修一, “異なる解像度からなる画像信号に対する視覚感度を用いた最適量子化器設計,” 1993年度画像符号化シンポジウム (PCSJ), pp.39-40, 1993年10月.
- [2] 酒澤茂之, 滝嶋康弘, 和田正裕, “IP ビデオマルチキャストにおける画質制御方式の検討,” 電子情報通信学会 画像工学研究会, IE97, pp.19-24, 1997年11月.
- [3] 酒澤茂之, 山下鉄司, 滝嶋康弘, 和田正裕, “MPEG-1 ビットストリームから検出可能な電子透かし方式の検討,” 2000年暗号と情報セキュリティシンポジウム, D26, 2000年1月.
- [4] 酒澤茂之, 渡辺泰利, 滝嶋康弘, 和田正裕, “Web ベース動画像蓄積伝送システム VAST-web の開発,” 映像情報メディア学会 ネットワーク映像メディア研究会, vol.25, pp.13-18, 2001年10月.
- [5] 酒澤茂之, 滝嶋康弘, 和田正裕, 羽鳥好律, “電子透かし攻撃ツール StirMark の動画像品質に及ぼす影響,” 第1回 情報科学技術フォーラム, 情報科学技術レターズ, pp.179-180, 2002年9月.