

論文 / 著書情報
Article / Book Information

題目(和文)	デジタル放送の画質評価に関する研究
Title(English)	
著者(和文)	中須英輔
Author(English)	Eisuke Nakasu
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第7306号, 授与年月日:2008年3月26日, 学位の種別:課程博士, 審査員:羽鳥 好律
Citation(English)	Degree:Doctor of Engineering, Conferring organization: Tokyo Institute of Technology, Report number:甲第7306号, Conferred date:2008/3/26, Degree Type:Course doctor, Examiner:
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

東京工業大学

学位論文

デジタル放送の画質評価に関する研究

中須 英輔

目 次

第 1 章	序論	1
1.1	本研究の背景と目的	1
1.2	本論文の構成と概要	2
	参考文献	3
第 2 章	デジタル放送における画質とその評価	4
2.1	はじめに	4
2.2	デジタル放送における画質の特徴	4
2.2.1	デジタル放送の映像フォーマット	4
2.2.2	映像符号化方式の概要	6
2.2.3	デジタル放送の画質の特徴	9
2.3	画質の評価技術	11
2.3.1	画像の主観評価技術	11
2.3.2	放送応用における画質への要求条件	14
2.3.3	客観画質評価技術	15
2.3.4	クリティシティ	20
2.4	まとめ	21
	参考文献	22
第 3 章	クリティシティ（符号化難易度）	24
3.1	はじめに	24
3.2	デジタル符号化画像の画質の絵柄依存性	24
3.2.1	MPEG-2 符号化における画質とビットレートの関係	24
3.2.2	デジタル放送と NTSC との画質比較	30
3.3	クリティシティの提案	33
3.4	クリティシティと画質の関係	38
3.4.1	クリティシティ測定装置	38
3.4.2	評価画像のクリティシティ測定	38
3.4.3	クリティシティと画質の関係	40
3.5	まとめ	46
	参考文献	46

第 4 章	デジタル放送の画質の統計的解析	47
4.1	はじめに	47
4.2	クリティカリティに基づく放送番組の画質の統計的解析手法	47
4.3	放送番組のクリティカリティ	48
4.3.1	放送番組のクリティカリティの測定	48
4.3.2	番組ジャンルによるクリティカリティの違い	50
4.4	時間率から見た放送番組の画質	51
4.4.1	シンプルプロファイル(SP@ML)とメインプロファイル(MP@ML)の画質比較	52
4.4.2	番組ジャンルによる画質差	54
4.5	まとめ	55
	参考文献.....	56
第 5 章	異種符号化映像フォーマットの総合画質評価	57
5.1	はじめに	57
5.2	デジタル放送における総合画質の評価	57
5.2.1	デジタル放送の画質と評価の課題	57
5.2.2	評価シーケンスとクリティカリティ	58
5.2.3	解像度の異なる基準画像を用いた符号化画像の評価	60
5.3	異種映像フォーマットの主観画質評価	64
5.3.1	評価実験の概要と使用機器	64
5.3.2	評価結果および映像フォーマット間の画質比較	68
5.3.3	放送における映像フォーマット選択に関する考察	72
5.4	まとめ	73
	参考文献.....	73
第 6 章	結論	75
	謝辞.....	76
	研究業績.....	77

第1章 序論

1.1 本研究の背景と目的

放送のデジタル化は、有限な資源である周波数の有効利用、放送の高品質・高機能化、放送・通信の融合・連携サービス等の面から国際的に目覚しく進展している。日本におけるデジタル放送は、表 1.1 に示すように、BS デジタル放送が 2000 年 12 月に開始され、地上デジタル放送が 2003 年 12 月に始まるなど、高画質・高音質のハイビジョン放送、視聴者の多様なニーズに応える多チャンネル放送、データ放送や双方向サービス等の便利で高機能なサービス等の魅力から順調に普及が進んでいる。

表 1.1 日本におけるデジタル放送の進展

狭帯域CS	1996年10月開始
BS	2000年12月開始
CATV	2000年以降本格化
広帯域CS	2002年4月開始
地上	2003年12月：東名阪広域圏、2006年：その他の地域
ワンセグ	2006年4月開始
2007年10月	アナログハイビジョン放送終了
2011年7月	アナログテレビ放送終了

デジタル放送は、図 1.1 に示すように、時間帯によってハイビジョン放送と多チャンネル放送を切替えたり、野球の試合が延伸した場合に野球中継を継続しつつ、定時番組を放送するなど、柔軟な番組編成が可能である。このようなサービスの多様性とシステムの柔軟性がデジタル放送の大きな特徴であるが、その特徴を十分活かすためには伝送ビットレートや映像フォーマットなどを適切に選択することが必要であり、その選択のための品質の評価法と基準が重要となる。

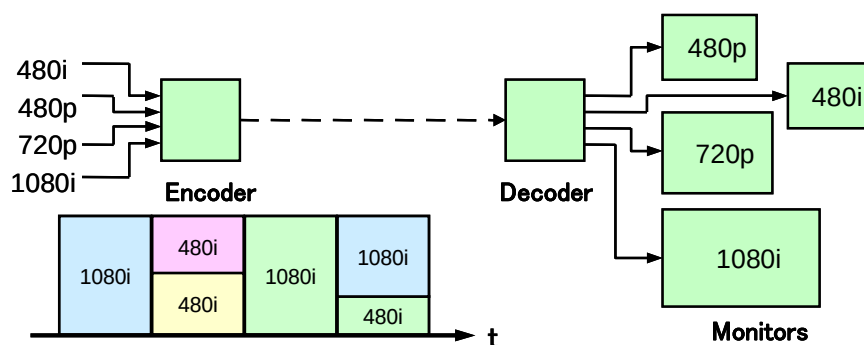


図 1.1 デジタル放送における多様な映像フォーマットと柔軟な番組編成の例

デジタル放送では、もとの画像のデータ量をおよそ 1/30～1/70 程度まで圧縮して伝送しているが、圧縮率が一定の場合でも、符号化対象映像のテクスチャやカメラ・被写体の動きなどに応じて画像品質(画質)が変動する性質がある[1]。このため、従来のアナログ放送の画質評価のように数種類程度の画像を用いた評価では、放送サービスの画質を適切に判断することができない。放送技術の国際標準化を所掌する ITU-R(国際電気通信連合無線通信部門)では、このようなデジタル符号化における画質の絵柄依存性を考慮し、圧縮の難易度を表現する「クリティカルティ」の考え方を勧告 BT.500-7[2]の中で示した。クリティカルティは、評価画像の客観的な選択や長時間映像の画質解析を行ううえで重要な概念であるが、汎用で具体的な定義は行われていない。一方、衛星放送サービスでは、降雨減衰等により受信不良の状態が生じるため、サービス時間率という概念で品質の評価が行われてきた。デジタル放送サービスでは、画質変動を考慮した長時間にわたる画質確認が必要であり、衛星放送における降雨減衰の場合と同様に時間率でサービス品質を評価し、大量映像データを対象として統計的に画質を解析する手法が求められる。

また、デジタル放送の映像フォーマットは複数の形式が規定されており[3]、前述のように同一伝送チャンネル内で放送番組の数を変更したり、データ放送サービス等とのビットレート配分比を変えるなど、柔軟な運用が可能である。したがって、伝送ビットレートや番組数に応じて最も良好な画質を与える映像フォーマットを選択することが重要となる。このため、従来の画質評価のように符号化劣化だけに注目するのではなく、画像の解像感なども含めた総合的な画質の評価が必要であり、それに基づく最適な映像フォーマットの選択基準の確立が求められる。

本研究は、デジタル符号化画質の絵柄依存性を考慮した評価基準を明確にするため、画像の絵柄と符号化時の画質の関係を示すクリティカルティを具体的に定義し、それに基づいてデジタル放送のサービス品質を時間率で評価する手法を提案する。さらに、伝送ビットレートや映像フォーマットを比較的自由に選択できるデジタル放送の利点を活かすため、異なるテレビ方式(映像フォーマット)の符号化画像において、画像の解像感を含んだ総合画質を評価する手法を示し、最適な放送映像フォーマット選択の指針について考察する。

1.2 本論文の構成と概要

本論文は6章から構成される。

第1章では研究の背景と目的を述べ、本研究の位置付けを示す。第2章でデジタル放送の画質の特徴を示し、これまでの画質評価技術の研究を整理するとともに、放送における画質評価の課題を明らかにする。第3章でデジタル映像符号化における画質の絵柄依存性を定量的に表現するクリティカルティ(符号化難易度)を具体的に定義し、符号化画像の主観評価値との関係を調べ、その有効性を示す。次に、第4章で様々な絵柄の映像素材を扱うデジタル放送における画質の変動要素を考慮して、放送の画質分布を統計的に

解析し、サービス品質を時間率で評価する手法を提案する。すなわち、第3章で求めたクリティカリティと画質の関係に基づき、番組のクリティカリティ測定結果から放送番組の画質分布を求め、その解析結果からデジタル放送における画質を時間率で評価する。さらに、第5章では、柔軟性の高いデジタル放送システムの特徴を考慮して、異なる映像フォーマットの符号化画像を統一的に評価し、符号化劣化だけでなく、画像の解像感等も含めた総合画質の評価方法を提案する。また、異種フォーマット映像の総合画質を画像のクリティカリティから評価し、各ビットレートにおける最適な映像フォーマットの選択基準について考察する。第6章では、本研究を総括し、得られた成果をまとめるとともに、今後の研究課題、展望について述べる。

参考文献

- [1] 中須英輔, “画質とユーザ要求 (MPEG 特集) 5-1 画質とユーザ要求,” テレビ誌, 49, 4, pp. 497-501 (1995)
- [2] Rec. ITU-R BT.500-7, “Methodology for the Subjective Assessment of the Quality of Television Pictures” (1995)
- [3] ARIB STD-B32, “デジタル放送における映像符号化、音声符号化及び多重化方式,” 1.4 版 (2003)

第2章 デジタル放送における画質とその評価

2.1 はじめに

デジタル符号化・伝送技術の進展により、従来のアナログ NTSC 放送と同じ伝送帯域内でハイビジョン放送や多チャンネル放送が可能となり、高画質・高音質のデジタル放送の普及が進んでいる。デジタル放送の画質は、伝送エラーによるノイズや符号化劣化が時空間的に局所に出現し、シーンや映像内容によって劣化程度が変化するなど、従来のアナログ放送と性質が大きく異なる[1]。

デジタル放送の品質を考えるうえで、このような映像符号化特有の画質劣化の性質を考慮する必要があり、従来の波形測定を中心としたアナログ放送の品質管理手法をそのままデジタル放送に適用することは難しく、デジタル符号化の性質を前提とした新たな品質管理、品質評価手法の開発が望まれている。このため、デジタル放送の画質評価においては、画質の絵柄依存性などの劣化の性質を考慮して、どのような絵柄において顕著な画質劣化が生じるか、また、実際の放送番組では各種ビットレートでどの程度の頻度で画質の低下が生じるのかという観点で評価することが求められる。

本章では、デジタル放送で用いられる映像符号化と画質の特徴を示し、デジタル放送における画質評価の課題をまとめる。さらに、デジタル放送の画質の評価に主に用いられてきた主観評価技術と、近年盛んに研究が進められている客観画質評価の現状を整理して述べる。

2.2 デジタル放送における画質の特徴

2.2.1 デジタル放送の映像フォーマット

アナログ放送は、番組制作時のスタジオ信号と同一形式の映像信号で伝送、表示されることが基本であった。一方、デジタル放送は複数形式の映像信号を選択して同一チャンネルで伝送可能であり[2]、時間帯によってハイビジョン放送と標準テレビ放送を切替え、番組数を変えて放送することが可能である。また、伝送ビットレートなどに応じてエンコード前に符号化劣化が目立たないようなフォーマットに変換して符号化、伝送することもできる。一般にデジタル放送受信機は、放送番組映像のフォーマットをモニタの種類や解像度に

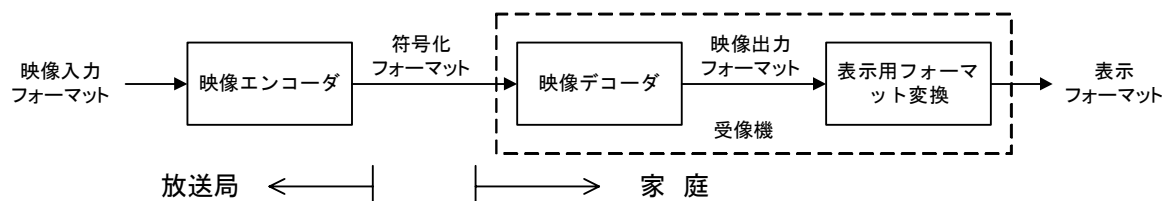


図 2.1 デジタル放送における映像フォーマット

応じて受信機内で変換し、表示する。このように、デジタル放送では、図 2.1 に示すように、それぞれの段階で必要に応じてフォーマット変換処理が行われる。

映像入力フォーマットは、放送局の映像エンコーダの入力信号に相当し、通常、番組制作に用いられるスタジオ信号と同じ形式である。符号化フォーマットは、デジタル圧縮処理を行う映像フォーマットで、ビットレート等に応じて適切なフォーマットを選択することが可能である。受像機内の映像デコーダの出力である映像出力フォーマットは、一般に、エンコーダに入る映像入力フォーマットと同一である。デジタル放送で用いる映像入力フォーマットを表 2.1 に、符号化フォーマットを表 2.2 に示す。表 2.2 に示すように幅広いビットレートでの運用が可能であり、1080i, 480i においては伝送ビットレートと画質を考慮して、水平方向にサンプル数を落として符号化することができる。

表 2.1 デジタル放送の入力映像フォーマット

パラメータ		480i	480p	1080i	720p
走査線数		525 本	525 本	1125 本	750 本
有効走査線数		483 本	483 本	1080 本	720 本
走査方式		1本おき飛越し	順次	1本おき飛越し	順次
フレーム周波数		30/1.001Hz	60/1.001Hz	30/1.001Hz	60/1.001Hz
フィールド周波数		60/1.001Hz	-	60/1.001Hz	-
画面の横と縦の比		16:9 または 4:3	16:9	16:9	16:9
水平走査の繰り返し周波数		15.750/1.001Hz	31.500/1.001Hz	33.750/1.001Hz	45.000/1.001Hz
標本化周波数	輝度信号	13.5MHz	27MHz	74.25/1.001 MHz	74.25/1.001 MHz
	色差信号	6.75MHz	13.5MHz	37.125/1.001 MHz	37.125/1.001 MHz
1走査線当たりの標本化数	輝度信号	858	858	1650	2200
	色差信号	429	429	825	1100
1走査線当たりの有効標本化数	輝度信号	720	720	1280	1920
	色差信号	360	360	640	960

表 2.2 符号化フォーマットとビットレート

方式名	水平 サンプル数	垂直 ライン数	画面アスペクト 比(横:縦)	フレーム 周波数	走査 方式	ビットレート [Mbps]
1080i	1920, 1440	1080	16:9	30/1.001Hz	飛越し	BS: 12~24 地上: 8~20
720p	1280	720	16:9	60/1.001Hz	順次	BS: 12~24 地上: 8~20
480p	720	480	16:9	60/1.001Hz	順次	BS: 4~24 地上: 4~20
480i	720, 544, 480	480	4:3 あるいは 16:9	30/1.001Hz	飛越し	BS: 1.5~15 地上: 1.5~15

2.2.2 映像符号化方式の概要

デジタルテレビ放送の映像符号化方式は、MPEG-2 規格 (ISO/IEC 13818-2 | ITU-T H.262)[3]が採用されている。MPEG-2は、放送、通信、DVDなど様々な応用分野に用いられ、表 2.3 に示す7つのプロファイルが規格化されている。プロファイルは色解像度、予測方式、スケーラブル(階層)機能などの面から分類されている。

デジタル放送には、このうちメインプロファイルが採用されている。ただし、メインプロファイルのデコーダはシンプルプロファイルのビットストリームを復号できるので、遅延時間を少なくしたい場合には、シンプルプロファイルの仕様で放送することもできる。各プロファイルは、扱う画像の解像度によって 4 つのレベルが定義され、それぞれ符号化解像度の最大値が規定されている。480i はメインレベル、480p は H14 レベル、1080i および 720p はハイレベルにそれぞれ相当する。

表 2.3 プロファイルと符号化機能の関係

プロファイル名 符号化機能	シンプル	メイン	SNR スケ ーラブル	スペーシャル スケーラブル	ハイ	4:2:2	マルチ ビュー
双方向予測	なし	あり					
SNR スケーラブル	なし		あり			なし	
スペーシャル スケーラブル	なし			あり		なし	
テンポラル スケーラブル	なし						あり
色差信号 フォーマット	4:2:0				4:2:2		4:2:0

MPEG-2 は、図 2.2 に示すように「動き補償予測」、「離散コサイン変換(DCT)」、「量子化」、「可変長符号化」を組み合わせることで映像信号の圧縮を行っている。映像信号は膨大な情報量を持つ一方で、様々な種類の削減可能な情報を含んでおり、その冗長な部分を効果的に除去することで、画質を損なうことなく、大幅な情報量の圧縮を可能としている。MPEG-2 の後に標準化された MPEG-4[4]および AVC/H.264[5]もほぼ同様なアルゴリズムである。

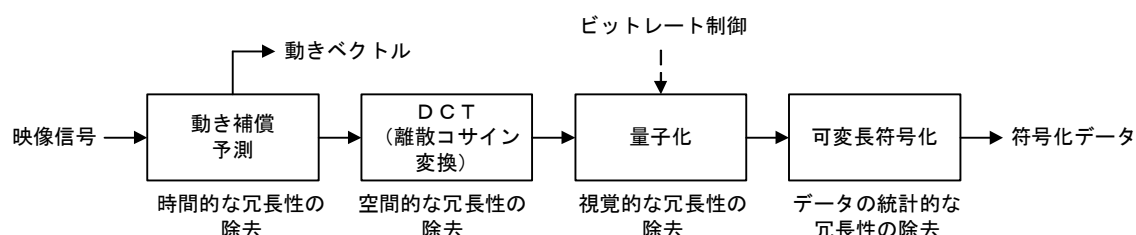


図 2.2 ハイブリッド DCT 符号化方式の圧縮原理

「動き補償予測」は、時間的に前後するフレームの画像からの動き量を検出し、その動き量に応じて補正した信号と原信号との差分をとって時間的な冗長性を除去する。次に、直交変換の一種である「DCT」により、原画像あるいは原画像と予測信号の差信号を 8×8 画素のブロック単位で空間周波数成分に変換し、低周波成分にエネルギーを集中させる。「量子化」は、DCT 変換係数を視覚特性を反映して適当な間隔の代表値に丸め込み(量子化)を行うことによりデータ量を低減する。「可変長符号化」は、量子化後の DCT 係数データは統計的な偏りがあるという性質を利用して、異なるビット数の符号語を割当てトータルの情報量を減らす。

MPEG-2 エンコーダのブロック図を図 2.3 に示す。MPEG-2 映像符号化では、図 2.4 に示すように、予測方式の異なる3種類のピクチャ(フレームあるいはフィールド)を周期的に選択して符号化を行う。ここに、 N は I ピクチャ(イントラ符号化)の間隔、 M は P ピクチャ(順方向予測)の感覚を示す。図 2.4(a)はシンプルプロファイル、(b)はメインプロファイルによる符号化に相当する。B ピクチャは前後の I あるいは P ピクチャを参照して双方向予測を行い符号化する。P ピクチャと B ピクチャでは、絵柄に応じて画面内でマクロブロック(16 画素 $\times$ 16 ライン)毎にイントラと動き補償予測を選択している。さらに、B ピクチャのマクロブロックでは、順方向／逆方向／双方向予測の切替えを行う。

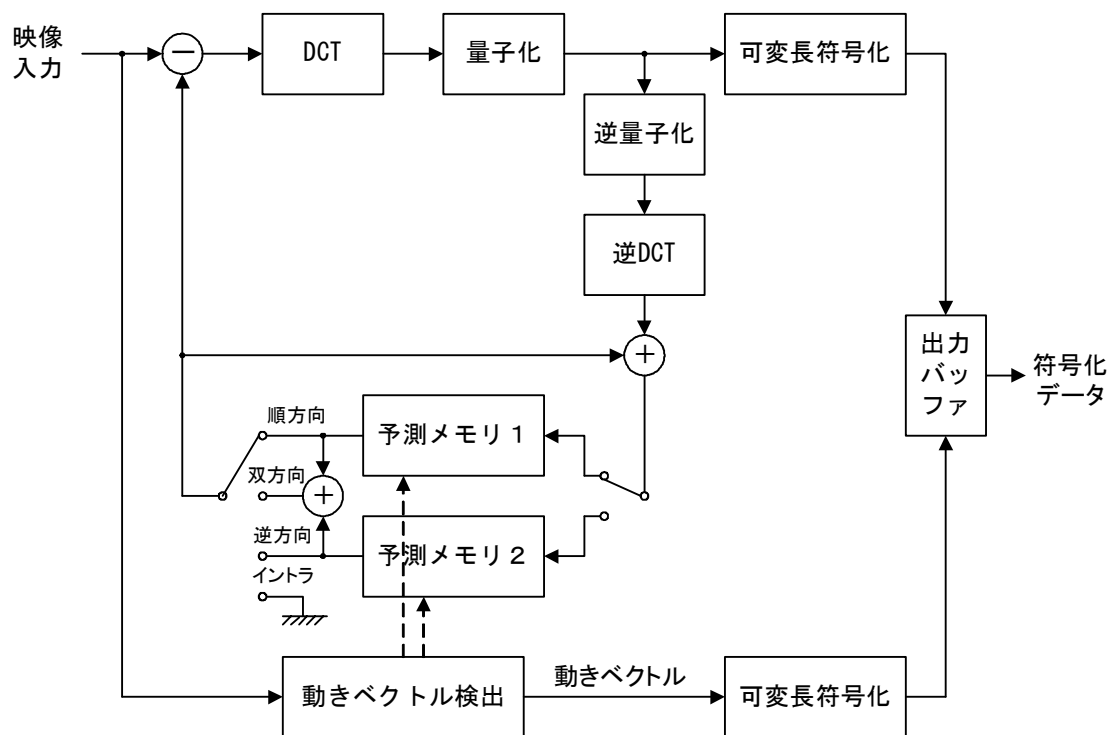
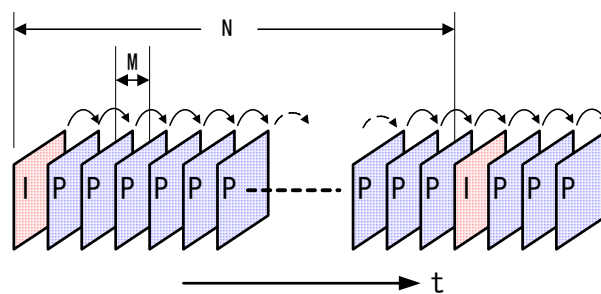
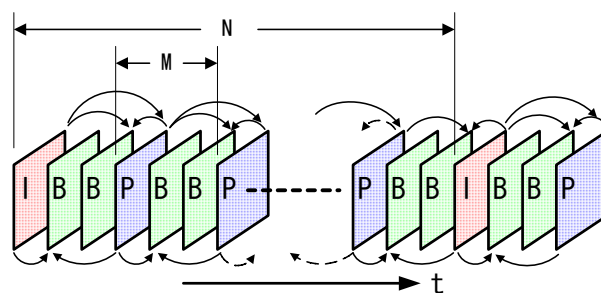


図 2.3 エンコーダのブロック図



(a) 順方向予測 (M=1, N=15)



(b) 双方向予測 (M=3, N=15)

図 2.4 順方向予測と双方向予測

2.2.3 デジタル放送の画質の特徴

デジタル符号化における主な劣化の例を表 2.4 に示す。

表 2.4 デジタル符号化における主な劣化の例

歪みの種類		劣化の見え方
静的劣化	ボケ	エッジや細かい模様がぼけて、明瞭でなくなる劣化
	偽輪郭	輝度や色の緩やかな変化部分で等高線状に生じる偽の輪郭
	グラニュラノイズ	平坦部で生じるランダムな粒子状の雑音
	ブロック歪み	もともとの絵柄に無い幾何学パターン(ブロック状)の歪み
動的劣化	ジャーキ・ムーブメント	動きがなめらかでなく、ギクシャクする劣化
	ダーティウィンドウ	動物体の後にノイズがあるように見える劣化
	フリッカ	輝度レベルが変動し、ちらついて見える劣化
	動きボケ	動領域がぼけて見える劣化
	モスキートノイズ	エッジの周辺に生じる、蚊が飛びまわるように見える雑音
	エッジビジネス	エッジ部分がザラザラして、ちらついて見える劣化

MPEG-2 映像符号化での劣化は、主として DCT 係数の量子化に起因して生じ、ボケ、ブロック歪み、モスキートノイズなどが代表的である。一般に、符号化処理等による歪量(誤差量)は、(2.1)式の PSNR (Peak Signal to Noise Ratio) で示す。

$$PSNR[dB] = 10 \log_{10} \left(\frac{255^2}{MSE} \right) \quad (2.1)$$

ここに、MSE(平均二乗誤差)は原画像と符号化画像間の誤差電力である。同程度の歪量であっても、劣化の種類によって見え方が違うため、画質への影響も異なることがデジタル符号化劣化の特徴の一つである。

発生する歪の形態は、符号化対象の波形とともに予測方式とも深く関係する。また、符号化劣化の程度はピクチャタイプによって異なり、図 2.5 に示すように時間的に劣化程度が変化してフリッカとして検知される場合がある。特に、画面全体が細かい模様の絵柄では、大画面フリッカが生じやすい。

一般に、デジタル符号化方式では、絵柄や想定ビットレート等に応じて様々なパラメータや符号化モードを適応的に切替えて処理を行うため、単純な絵柄を符号化する場合と複雑な絵柄の画像を符号化する場合とではエンコーダの動作が異なる。このため、デジタル映像符号化においてはテスト信号を用いた特性確認だけでは十分でなく、様々な要素を含んだ自然画像を用いて総合的に画質を評価する必要がある。

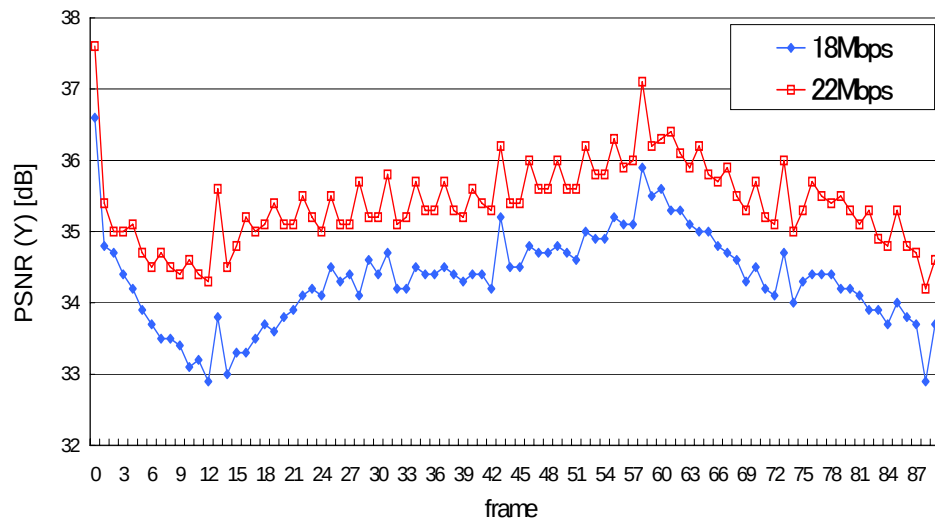


図 2.5 MPEG-2 における符号化 PSNR の変動例

符号化画像の画質は、圧縮率(ビットレート)によって支配的に決まるが、符号化対象映像の絵柄や動きなどの画像内容にも依存する。定性的には図 2.6 に示すように、圧縮が容易な画像(a)は低いビットレートでも高い画質が維持される一方で、画像(c)のように圧縮が難しい画像では急激に画質劣化が増大する。

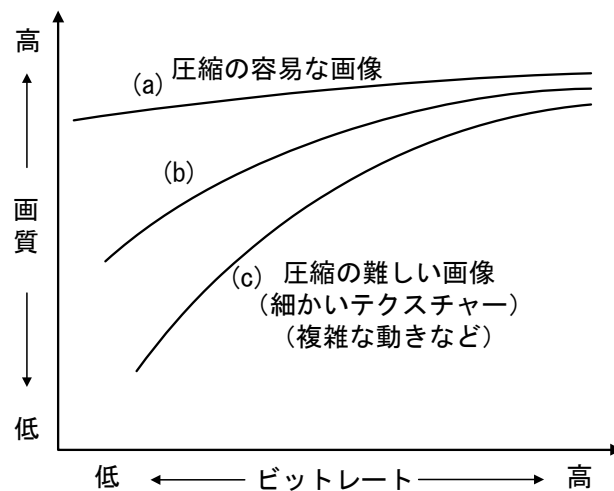


図 2.6 各種画像におけるビットレートと画質の関係

一般に、細かいテクスチャを含んだ画像や複雑な動きを含んだ画像、フェードイン／アウトやストロボ照明のように輝度変化の激しい画像は、圧縮の難しい画像である。回転や拡大・縮小、複雑な動きや速い動き、シーンチェンジの前後などでは予測誤差が大きくなり顕著な劣化が生じやすい。

放送番組のように多様な画像を扱う応用においては、絵柄による画質変動を十分考慮しなければならない。図 2.7 に示すように、比較的高いビットレートが確保できる場合には、絵柄による変動も小さく、一定の高い画質を維持することができるが、ビットレートが低くなると、画像による画質差が大きくなり、非常に圧縮の難しい画像では顕著な劣化が生じる危険性が有る。

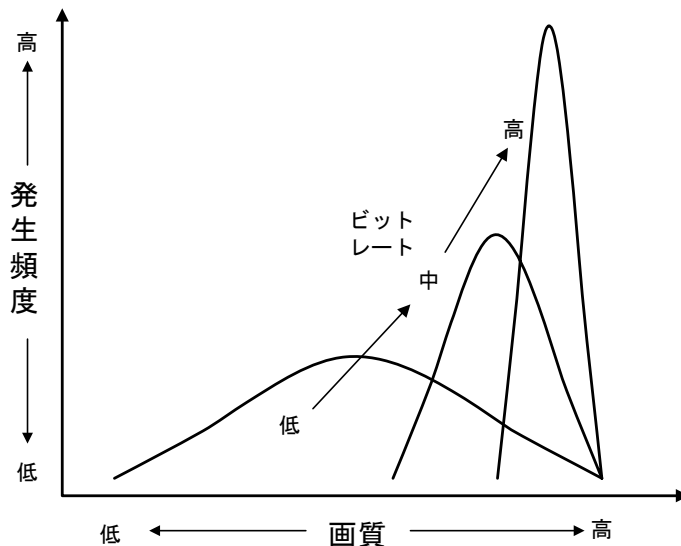


図 2.7 放送番組における画質分布の概念

ビットレートと画質の関係において、ビットレートの違いによる画質の差よりも画像内容の違いによる画質差のほうが大きい場合がある。このため、限られた性質の少数の画像を用いて評価した結果からビットレートと画質の関係を論じて、正しくシステムを評価することはできない。放送のような応用では、画質を長時間にわたってモニタし、所望の画質を確保するのに必要なビットレートを慎重に調べる必要がある。

2.3 画質の評価技術

2.3.1 画像の主観評価技術

映像メディアの評価は、人間が映像を見て、その品質を判断する主観評価により通常行われる。主観評価は、人間が映像や音などの刺激を受けてその品質等を評価するものであり、評価は個々人の主観に基づき下される。主観評価法には、目的や用途に応じて多くの方法がある[6]が、画質の評価には、物理的刺激に対する人間の主観的感覚を定量的尺度にあてはめる評定尺度法[7]が広く用いられる。評定尺度法は、「良い」「悪い」などの順序付けしたカテゴリーに数値を与え、対象に対する評価の程度との対応を求める方法である。この尺度は、厳密には各対象の順序を与える「順序尺度」であるが、通常、加減算、平均値算出などが可能な「距離尺度」とみなして扱うことが多い。その尺度には

品質の良し悪しを表現する「品質尺度」と劣化程度を表現する「劣化尺度」がある。評価に用いるカテゴリーの数は通常、5 ないし 7 が良く用いられる。表 2.5 に、5 段階の品質尺度および劣化尺度を示す。

表 2.5 劣化尺度および評価尺度 (5 段階)

評点	品質尺度	劣化尺度
5	非常に良い	わからない
4	良い	わかるが、気にならない
3	普通	気になるが、邪魔にならない
2	悪い	邪魔になる
1	非常に悪い	非常に邪魔になる

評価法には、基準となる刺激を与えたうえで評価対象を評価する方法と、評価対象だけを示して評価する方法がある。前者はより安定した結果が得られる反面、評価に要する時間が長くなり、評価項目を限定する必要が生じる場合もある。また、評定者が刺激を調整したり、自由に刺激を切替えながら評価を行う方法もあるが、評定者が納得するまで時間をかけることで安定した結果が得られる反面、評価時間は通常長くなる。

人間の判断に基づく主観評価では個人差による評点のバラツキが伴うため、一定数以上の評定者により評価テストを実施する。通常、評価データに基づく評定者のスクリーニングを行い、統計的にデータ処理して評価結果が得られる。また、他の評価実験結果と比較できる客観的で普遍性のある結果を得るため、評価法の適切な選定に加え、使用するモニタやその周囲環境、評定者、視距離などの観視条件、評価シーケンス等の共通化が必要である。このため、ITU-R では、評価法、標準観視条件、評価用標準動画像などを標準規格で細部にわたり規定している。

デジタル符号化画像の評価には、表 2.6 に示す DSIS (Double-Stimulus Impairment Scale) 法と DSCQS (Double Stimulus Continuous Quality Scale)法が通常用いられる[8]。DSIS 法は EBU 法とも呼ばれ、隣接チャンネルの電波からの妨害程度など、基準画像に対する劣化の評価に用いられる。DSCQS 法はデジタル符号化画像の評価に一般的に用いられ、基準画像と比べた評価画像の品質差と評価画像の絶対品質を同時に評価することができる。

表 2.6 DSIS 法と DSCQS 法の比較

	DSIS法(EBU法) 二重刺激劣化尺度法	DSCQS法 二重刺激連続品質尺度法
目的	伝送エラーなどによる劣化程度の評価	システムや伝送路の品質評価
提示方法	参照画像と評価画像をペアで表示 (1 回あるいは 2 回提示)	参照画像と評価画像をペアで表示 (2 回提示)
参照画像	明示(先に提示)	明示しない
尺度	5 段階劣化尺度	連続品質尺度 (5 段階品質尺度と対応)
評点	評価画像のみを評価	参照画像と評価画像の両方を評価

画質評価において、観視条件を最適に設定することは評価結果に普遍性を与え、また信頼性を高めるうえで重要である。ITU-R ではモニタの多様化や観視環境の違いによる画質への影響を考慮して、表 2.7 に示すように観視条件が厳密に設定された実験室環境と、家庭でテレビを視聴する環境に近い家庭環境それぞれを規定している[8][9][10]。モニタもスタジオ用モニタだけでなく、民生用受像機による観視条件も規定されている。

表 2.7 標準観視条件

項 目	パラメータ値		
	SDTV(勧告 BT.1129-2, BT.500-9)		HDTV(勧告 BT.710-4)
	実験室環境	家庭環境	
視距離 (H:画面の高さ)	4H と 6H	6H	3H
画面のピーク輝度	70cd/m ²	200cd/m ²	150～250cd/m ² (当面 70 cd/m ² 以上)
非発光画面輝度のピーク輝度に対する比	≤0.02	≤0.02	≤0.02
黒レベル輝度のピーク輝度に対する比 (於:暗室)	≐0.01	—	≐0.01
背景輝度のピーク輝度に対する比	≐0.15	—	≐0.15
室内照明(周囲照度)	低いこと	—	低いこと
モニター設置場所の背景色	D65	—	D65
モニター及び背景を観視する角度	43°H×57°W	53°H×83°W	53°H×83°W
画面サイズ	20in 以上	(4:3)25～29in、 (16:9)32～36in	55in. 少なくとも 30in 以上
最大観視角度(垂線に対する角度)CRT	30°	30°	30°
周囲から画面への入射照度	—	200lux	—

デジタル符号化画像の品質は、同一ビットレートであっても絵柄によって大きく左右されるため、テスト画像の選定方法は大変重要である。通常、評価画像システムに対し、少なくとも4種の「クリティカルであるが極端にクリティカルではない」絵柄を選ぶ。これは、例えば、4種のテスト画像のうち2つは評価システムにとって明らかに厳しい絵柄とし、残りの2つは適度に厳しい絵柄を使用することを意味する。デジタル符号化画像の評価では、絵柄による影響をできるだけ少なくするため、できるだけ多くの画像を用いて評価することが望ましいと規定されている。

2.3.2 放送応用における画質への要求条件

放送および二次分配(ケーブルによる家庭への伝送)システムにおける画質に対するユーザ要求は、ITU-R 勧告 BT.1122[11]で規定されている。その内容を表 2.8 に示す。符号化画像の画質評価には ITU-R 勧告 BT.500 に記載される DSCQS 法が通常用いられるため、画質のユーザ要求条件も DSCQS 法による評価結果に基づいている。ただし、BT.1122 には、どのような評価画像を用いるか明確には規定されておらず、使用する画像によって評価結果が左右される危険性を含んでいる。

表 2.8 放送と二次分配システムに対する要求画質 (ITU-R 勧告 BT.1122)

基本画質 (良好な受信状態)	・DSCQS で 12%以内の劣化 ・ITU 勧告 BT.710, BT.802[12], BT.1128[13], BT.1210[14]から最低 4つの評価画像を選ぶこと。 これらのうち最低 75%の画像で上記条件を満たすし、残りの画像についても 30%以内の劣化であること。
-------------------	--

ARIB では、デジタル放送方式の検討にあたり、各種 MPEG-2 符号化映像を DSCQS 法により評価し、上記要求画質に基づいて所要ビットレートを検討した[15][16]。その結果を表 2.9 に示す。1080i および 480i の評価は 1997 年に、720p については 1999 年に実施されたが、480p

表 2.9 ITU-R の放送要求条件を満足する所要ビットレート

(ARIB での評価実験[15][16]に基づいた結果)

テレビ方式	評価画像	モニタサイズ	視距離 (H:画面高)	所要ビットレート
1080i	10 種	41 インチ	3H	22Mbps
720p	4 種+4 種*	32 インチ	3H	18Mbps
480i	6 種+6 種*	29 インチ	4H	6Mbps

*1080i からの変換画像

の評価は行われていない。また、1080iは水平1920画素で符号化した映像を用いて評価が行われた。720pおよび480iは、当該テレビ方式のカメラで撮像した映像に加え、1080iから変換した映像についても評価された。その結果では、絵柄による画質の違いが、撮像フォーマットの違いや方式変換による影響よりも大きく、720pあるいは480iのカメラ映像と1080i原画像からの変換画像との間で傾向の違いは認められなかった。また、DSCQS法では符号化による画質劣化のみを評価することになるため、表2.9に示すように解像度の低い画像ほど所要ビットレートが低い結果となっている。

デジタル符号化による画質劣化程度は、映像内容に大きく依存するため、放送サービスのように多種多様な画像を扱う応用分野では画質にばらつきが生じる。画質のばらつきはビットレートによって変わり、低いレートほど画像による画質差は大きくなる。したがって、放送応用では、以下のような点に注目して、システム設計することが重要となる。

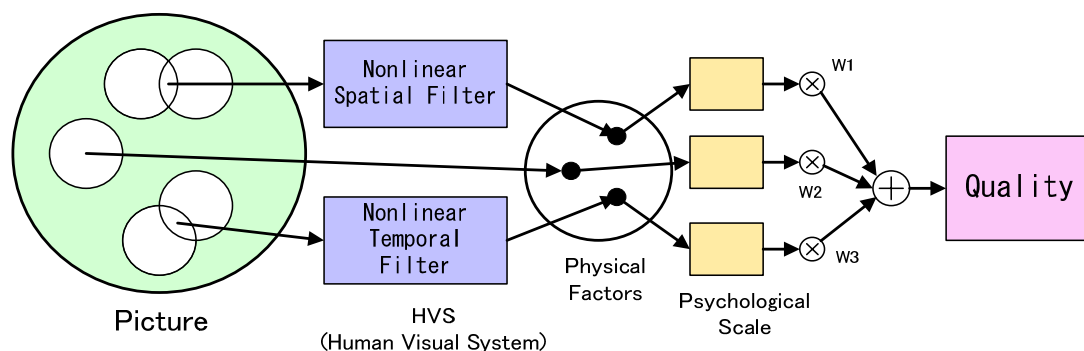
- ・ 符号化に厳しい画像でどの程度の画質を確保するか
- ・ 顕著な画質劣化が生じる時間率をどの程度に抑えるか

このような視点でシステム設計を行ううえで、画像内容による画質変動を考慮し、統計的な処理から画質を時間率で評価する手法の確立が望まれる。

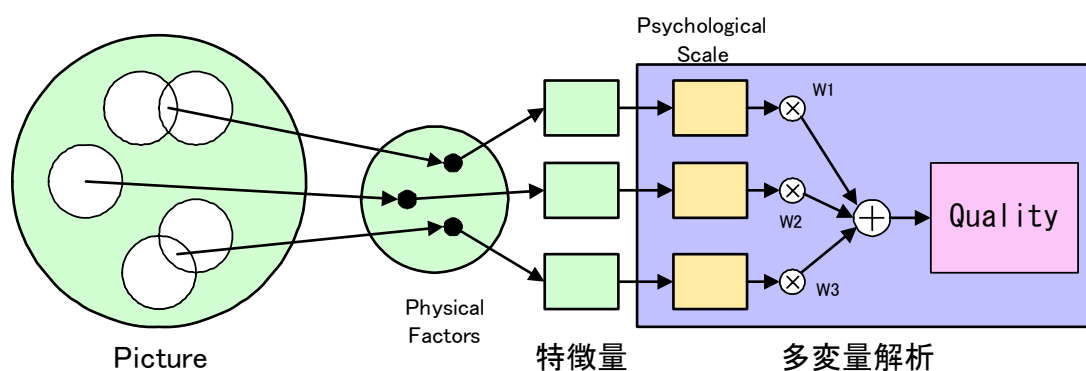
2.3.3 客観画質評価技術

映像や音など刺激の物理量と人間の視聴覚特性、心理量との対応付けで主観評価の代わりに評価結果を導き出す客観評価が近年注目されている。従来、画質の評価で主に用いられてきた主観評価は、適正な評価法の選定、観視条件の調整など、実験計画の立案および評価実験の実施に際して評価に関する専門知識と経験が必要となる。また、一定数以上の評定者を集め、評価結果を統計的にデータ処理するなど、実施には手間と時間もかかる。主観評価では一般に長時間の測定が難しく、リアルタイムのオンライン測定ができないなどの問題もある。

客観画質評価技術は、このような主観評価の欠点を補い、可能ならば主観評価に置き換わるものを目指して研究が進められている。客観画質評価は、符号化画像および符号化雑音・歪を物理的に測定し、解析する手法をいうが、図2.8(a)に示す、視覚・知覚モデルを忠実に再現して画質を推定する方式と、図2.8(b)のように、対象とする画像および符号化劣化の特徴量を抽出して画質と関係付ける方式がある。



(a) 視覚・知覚モデルを用いる方法



(b) 画像・符号化劣化の特徴量から解析する方法

図 2.8 客観評価モデルの分類

(a) 視覚・知覚モデルに基づく客観評価モデル

視覚・知覚モデルを用いた画質推定法は、符号化誤差信号やデコード画像に人間の視覚・知覚特性を適用するものである。図 2.9 に示すように、フィルタリングや非線型処理した結果を加重平均等して評価値へ変換する方法が代表的である[17]。この手法は、劣化要因や劣化の種類に依存しない汎用的なアプローチである。

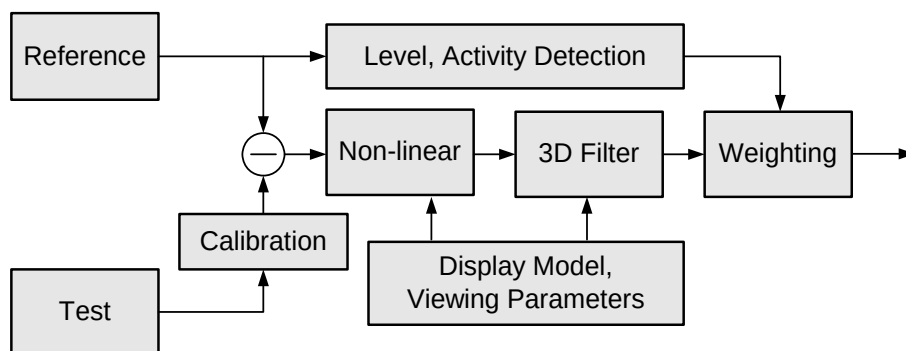


図 2.9 客観評価アルゴリズムの例

原画および符号化劣化の周波数成分の解析手法としては、時空間フィルタを用いるほか、FFTを用いる方法[18]、あるいはガウシアン・ピラミッドで周波数バンドに分割する方法[19]などがある。さらに、フリッカ、輝度信号、カラー信号のコントラストを計算し、検知限を考慮した閾値処理を行う方式[20]もある。被写体に対する知識などが影響する認知・認識機能については高次処理のため、学習や知識ベースの処理が必要となる。ここでは、画面全体を数フレームにわたり解析し、オブジェクト抽出や注視点解析など、複雑な解析が必要となる[21]。

(b) 画像の特徴量に基づく客観評価モデル

画像の特徴量を用いる方法は、画像の特徴量の変化や、符号化歪み・雑音の特徴量を計算し、その結果から画質を推定する手法である。空間周波数、エッジ電力、動き電力等の特徴量[22]と主観画質との関係を多変量解析等の手法で予め解析し、そこで求めた重み付けに基づいて総合的な画質を推定する方法が代表的である。ITU-T[23]で各種特徴量を規定しており、代表的なものに SI (Spatial Information)と TI (Temporal Information)がある。SI はエッジ検出オペレータの1つである Sobel フィルタ(図 2.10)の出力結果を用いる。

$$SI = \max_{time} \{ \text{std}_{space} [Sobel(F_n)] \} \quad (2.2)$$

-1	0	1
-2	0	2
-1	0	1

-1	-2	-1
0	0	0
1	2	1

(a)水平エッジ検出用 (b)垂直エッジ検出用

図 2.10 Sobel オペレータ

一般に、TI は次式に示すフレーム間差から求める。

$$TI = \max_{time} \{ \text{std}_{space} [F_n(i, j) - F_{n-1}(i, j)] \} \quad (2.3)$$

ここに、 $F_n(I, j)$ は n フレーム目の (I, j) 番目の画素値である。

図 2.11 に SI と TI の例を示す。



(a) 原 画



(b) SI (Sobel オペレータ出力)

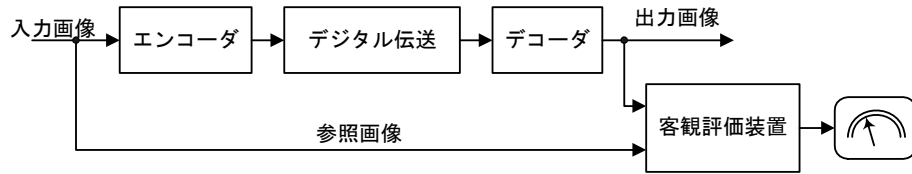


(c) TI (フレーム間差)

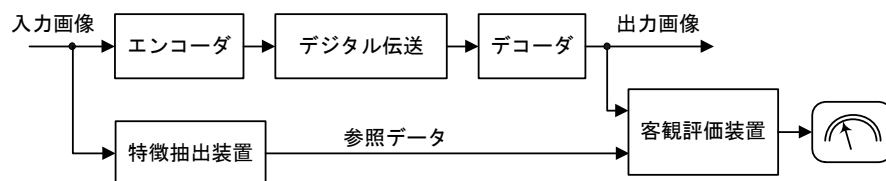
図 2.11 画像特徴量の例

(c) 客観評価の標準化動向

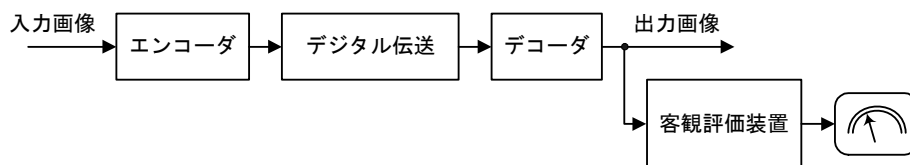
客観評価方式の標準化を進めている専門家グループ VQEG (Video Quality Experts Group) [24]では、図 2.12 に示すように、使用する画像データの種類によって3種に分類して標準化を進めている。



(a) 基準画像を使用する方法;FR(Full Reference)方式



(b) 基準画像の一部の情報を使用する方法;RR(Reduced Reference)方式



(c) 基準画像を使用しない方法;NR(No Reference)方式

図 2.12 客観評価法の分類

図 2.12(a)の FR 方式は、デコード画像と原画像(基準画像)の両方を用いて解析する方法であり、両者の差を求めて符号化歪みを解析する。入出力の比較に際しては空間/時間アライメント、レベルオフセット、ゲイン調整等が必要となる。評価に用いる画像が既知の場合には基準画像を伝送する必要は無く、受信側で予め蓄積しておいた基準画像を用いることができるが、一般に FR 方式では基準画像を忠実に伝送するために高レートの伝送路が要求され、エンコーダとデコーダが遠隔地にある場合には実現性が低い。このため、基準画像のフレームや画素を間引いたり、受端で解析するのに必要な基準画像の特徴量のみを送る図 2.12(b)の RR 方式が考えられる。画像の特徴量としてはエッジ情報や画面フリーズが検知できるフレーム間差情報等が挙げられる。図 2.12(c)の NR 方式は、符号化映像のみを手がかりに画質評価を行う方法である。基準画像を必要とせず、復号画像のみから画質を推定することができれば、画質のモニタなどでは大変有用である。デコーダと客観評価装置が近傍に設置される場合には、符号化ビットストリームを解析に利用することも考えられる。

FR 方式は、ITU においても標準化が進められ、勧告 ITU-R BT.1683[25]および ITU-T 勧告 J.144[26]で4方式が勧告されている。それぞれの方式の概要を以下に示す。

(1) BT 方式

処理画の水平勾配の符号反転回数、3 階層ピラミッド変換画像上での PSNR、原画と処理画のエッジ度の比較、原画と処理画をブロックマッチングした結果のばらつき指標、およびその差分信号の PSNR などを検出し、時間平均と各パラメータへの重み付けを行って最終評価値を算出。

(2) Yonsei(延世)大方式

原画に Sobel 演算を施して生成したエッジ部の PSNR を非線形変換、スケーリングにより最終評価値を導出。

(3) CPqD 方式

原画を平坦部、境界部、複雑な絵柄領域の 3 領域に分割し、それぞれに劣化度を計算。劣化度は、原画、処理画に Sobel 演算を施して得たエッジ画像の差分絶対値平均から算出。さらに、固定量子化器でフレーム内符号化した結果についても同じ処理を行い、それらの値を非線形変換、加重加算、メディアンフィルタリング処理して最終的な評価値を算出。

(4) NTIA 方式

輝度信号のエッジの増減、フレーム内標準偏差と単純フレーム間差分標準偏差、色差信号の差異に基づいて生成した指標を用いて、評価値を算出。

これらの方式の主観評価値との相関係数は、0.835～0.937 (525/60 映像)であり、PSNR(相関係数 0.804)と比較して統計的に優位性が認められている。一方、RR 型、NR 型については、精度および安定性で FR 型に匹敵するものは未だなく、今後の研究が期待されている。

2.3.4 クリティシティ

クリティシティは、符号化による画質劣化と関係付けられる、画像の特徴量の一つである。ITU-R 勧告 BT.500 では、Annex1 で符号化画像の絵柄依存性を概念的に示し、符号化の難しさを示す指標として「クリティシティ」の考え方を示した。クリティシティは、符号化に対する画像の固有の性質である。しかし、勧告では具体的なクリティシティの定義は示していない。

Lodge らは、対象とするアルゴリズムとビットレートを想定し、固定レート符号化動作におけるバッファ占有量をクリティシティと定義した[27]。そのブロック図を図 2.13 に示す。符号化装置において、符号化時の発生情報量によりバッファ占有量が変化し、その量に応じて量子化器がフィードバック制御される。クリティカルな画像では、発生情報量が増える傾向にあるため、バッファ占有量が増大する。したがって、このバッファ占有量をクリティシティと定義を行った。

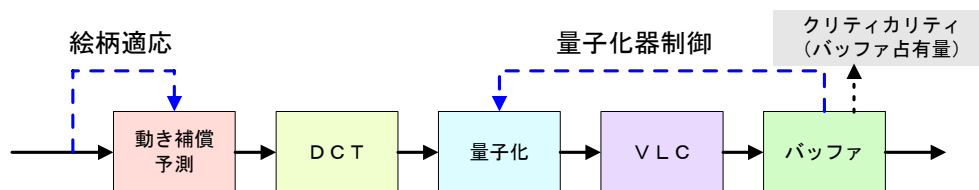


図 2.13 クリティシティの定義

しかし、彼らの定義では、想定するビットレート以外のレートでの符号化劣化の推定は行えず、また、フィードバック処理による時間遅延もあるため、時間制度良く動画シーケンスのクリティシティの測定が行えないなどの課題があった。また、具体的に、各種符号化条件における画質との関係付けは明らかにされておらず、その有効性や実用性についてはこれまで示されてはいない。

2.4 まとめ

デジタル映像符号化画像の画質は、絵柄によって変動し、局所的に劣化が生じるなど固有の特性を示す。デジタル放送のサービス品質を考えるうえで、画質変動を考慮して時間率でサービス品質を評価する手法の開発が求められる。

一般に、符号化画像の画質評価は、主観評価を用いて行われる。ただし、信頼性の高い結果を得るには、評価法や実施方法に対する高い専門性が必要である。また、主観評価テストは、多くの人を集め、ディスプレイや観視条件の調整、データの統計解析など、大変手間のかかる作業である。客観評価法は、様々な手法が提案されているが、未だ主観評価に代わる高い信頼性は得られていない。主観評価では難しい、多量の画像データの連続測定が可能であるが、復号後の符号化画像の画質を評価する技術であり、異なる符号化条件での画質は推定できない。クリティシティは、符号化劣化と関係付けられる画像の特徴量の一種である。これまでに提案された方式では、特定の符号化条件について他のシーケンスの画質は推定できるが、ビットレートや符号化条件の異なる場合の画質推定は行えなかった。

参考文献

- [1] 中須英輔, “テレビ画像のデータ圧縮とその画質評価,” ディスプレイアンドイメージング, 5, 3 (1997)
- [2] ARIB STD-B32, “デジタル放送における映像符号化、音声符号化及び多重化方式,” 1.4 版 (2003)
- [3] ISO/IEC 13818-2 | ITU-T Rec. H.262, “Information Technology - Generic Coding of Moving Pictures and Associated Audio: Video” (2000)
- [4] ISO/IEC 14496-2, “Information Technology - Coding of audio-visual objects part 2: Visual” (2004)
- [5] ITU-T Rec. H.264 (2003), “Advanced Video Coding for generic audiovisual services” | ISO/IEC 14496-10:2003, “Information Technology, Coding of Audio-visual Objects – Part 10: Advanced video coding”
- [6] 大串健吾, 中山剛, 福田忠彦著, テレビジョン学会編, “画質と音質の評価技術,” 昭晃堂(1991)
- [7] 宮川洋監修, テレビジョン学会編, “テレビジョン画像の評価技術,” p.91-92, コロナ社 (1986)
- [8] Rec. ITU-R BT.500-9, “Methodology for the subjective assessment of the quality of television pictures” (1998)
- [9] Rec. ITU-R BT.1129-2, “Subjective Assessment of Standard Definition Digital Television (SDTV) Systems” (1998)
- [10] Rec. ITU-R BT.710-4, “Subjective Assessment Method for Image Quality in High-Definition Television” (1998)
- [11] Rec. ITU-R BT.1122-1, “User Requirements for Emission and Secondary Distribution Systems for SDTV, HDTV and Hierarchical Coding Schemes” (1995).
- [12] Rec. ITU-R BT.802-1, “Test picture and sequences for subjective assessments of digital codecs conveying signals produced according to Recommendation ITU-R BT.601” (1994)
- [13] Rec. ITU-R BT.1128-2, “Subjective Assessment of Conventional Television Systems” (1997)
- [14] Rec. ITU-R BT.1210-1, “Test materials to be used in subjective assessment” (1997)
- [15] 中須英輔, 中田安優, 浦野丈治, 清水勉, “デジタル放送の符号化画質評価実験,” 映像学技報, 22, 19, pp13-18 BCS'98-3 (Mar. 1998)
- [16] 小澤啓爾, 羽鳥好律, 小熊透, 川手雅之, “720P 実証実験について”, 信学技報, 99,

213, pp. 83-90 SAT99 60-73 (1999)

- [17] B. Watson, "Digital Images and Human Vision," MIT Press, London (1993)
- [18] 浜田高宏, 松本修一, 村上仁己, "量子化器最適化に基づく動き補償 DCT 方式の特性限界," 信学論(B-I), J78-B-I, 11, pp.689-699 (Nov. 1995)
- [19] 宮原誠, 堀田裕弘, 豊芦穰, 片山昌和, ムハマド ユスフ エフエンディ, "動画像の客観評価尺度," テレビ学技報, BCS95-80, pp.1-6 (1995)
- [20] 浜田浩行, 斎藤恭一, 難波誠一, "標準 TV 動画像品質の客観評価尺度の検討," 信学技報, IE-93-94, pp.39-45 (1994)
- [21] 浜田高宏, 宮地悟史, 松本修一, "デジタル圧縮画質評価システムの開発," テレビ学技報, BCS96-28, pp.13-18 (June 1996)
- [22] ANSI Standard T1. 801, "Digital Trans- port of One-way Video Signals - Parameters for Objective Performance Assessment" (Mar. 1996)
- [23] ITU-T Rec. P. 910, "Subjective video quality assessment methods for multimedia applications" (1996)
- [24] P. Corriveau, A. Webster, "VQEG Evaluation of Objective Methods of Video Quality Assessment," SMPTE J., 108, 9, pp.645- 648 (1999)
- [25] Rec. ITU-R BT.1683, "Objective perceptual video quality measurement techniques for standard definition digital broadcast television in the presence of a full reference" (2004)
- [26] ITU-T Rec. J.144, "Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference" (2001)
- [27] N.K.Lodge, D.Wood, "Subjective Optimising Low Bit-rate Television," Proc. on IBC 94, Amsterdam, The Netherlands, pp.333-339 (Sept. 1994)

第3章 クリティシティ(符号化難易度)

3.1 はじめに

第2章で述べたように、デジタル符号化を行った画像の画質は、映像内容やカメラの動き、照明の変化などで大きく影響を受ける。このため、絵柄による画質変動を考慮した評価基準の明確化が求められる。本章では、評価画像の絵柄の性質と符号化時の画質の関係を示す、画像のクリティシティ(符号化難易度)を具体的に提案する。まず、各種映像を MPEG-2 符号化で圧縮したときの画質を主観画質評価実験により評価し、その結果を示す。次に想定する符号化アルゴリズムをもとにクリティシティを定義し、主観評価結果との関係を調べ、その有効性について考察する。

3.2 デジタル符号化画像の画質の絵柄依存性

3.2.1 MPEG-2 符号化における画質とビットレートの関係

デジタル放送で用いられる MPEG-2 映像符号化[1]を用いて各種標準テレビ映像を符号化し、その画質を DSCQS 法により評価した。符号化条件および評価条件を表 3.1 に示す。評価実験は、MPEG-2 メインプロファイル(MP@ML)およびシンプルプロファイル(SP@ML)に準拠した符号化により 4, 6, 9, 15Mbps で符号化した結果について行った。表示には 29 インチスタジオモニタを使用し、4H および 6H(H: 画面高)の視距離で評価実験を行った。また、アナログ放送との画質比較のため、NTSC 画像についてもあわせて評価を行った。NTSC 画像は、コンポーネントの原画像を NTSC 符号化した後、NTSC 復号して、MPEG-2 符号化画像と同様にコンポーネント信号で表示を行った。その際、クロスカラー妨害による劣化が極度に生じないように、予めカラーサブキャリア付近の高周波数成分をフィルタ処理により低減した後に、NTSC 符号化を行った。また、使用した評価シーケンスを表 3.2 に示す。ITU-R および ITE 標準画像に加え、NHK の放送番組から比較的クリティカルな画像と電子的に生成した文字スーパーとフェード画像など、計15種類について評価を実施した。

表 3.1 評価テスト条件

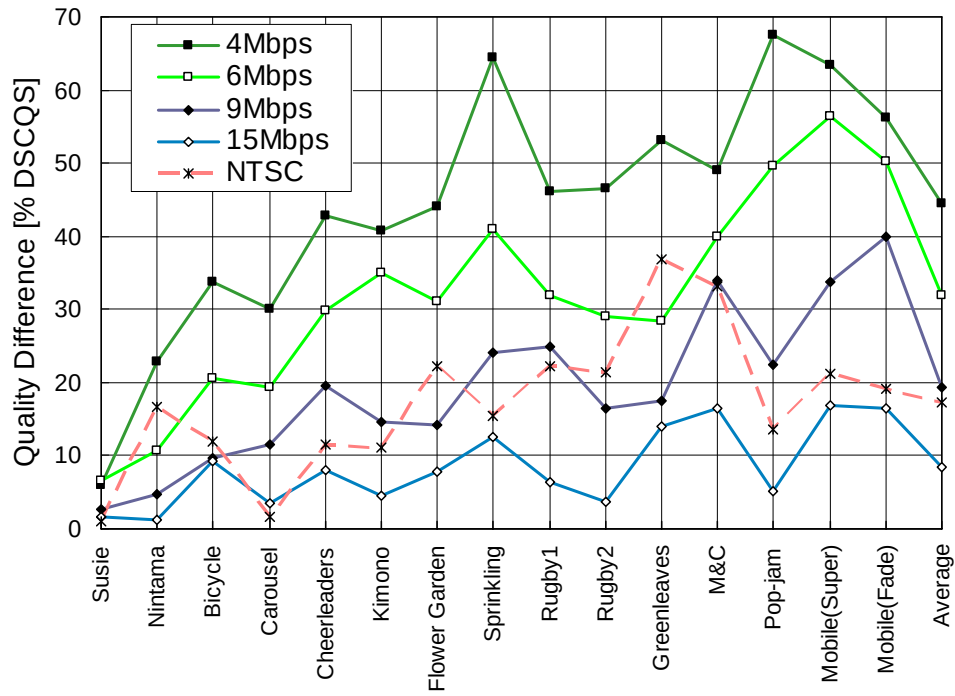
符号化方式	MPEG-2 SP@ML, MP@ML
ビットレート	4, 6, 9, 15 Mbps
テスト画像	現行テレビ(コンポーネント信号) 15 種類、各 5 秒
評価方法	二重刺激連続品質尺度(DSCQS)法 ITU-R 勧告 BT.500-7
モニタ	29 インチ(シバソク製スタジオモニタ)
評定者	映像に関する非専門家 20 名
視距離	4H, 6H(H: 画面高)

表 3.2 実験に用いたテスト画像

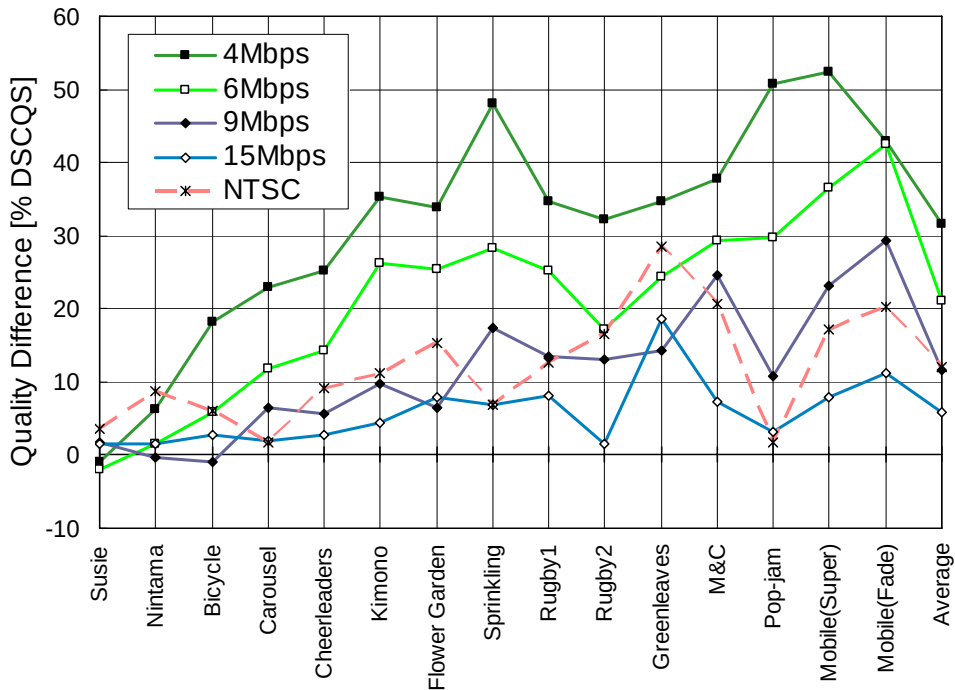
シーケンス名	ソース	シーケンスの特徴
Susie	a	電話中の女性のアップ
Flower Garden	a	花畑のある家のトラックショット。大きな木が視界を横切る。
Mobile&Calendar	a	汽車の模型と細かな模様のカレンダーが壁紙の前を動く。パンニング。
Cheerleaders	a	チアリーダーの速い複雑な動き。アンカバードバックグラウンドあり。
Bicycle	a	林の中で自転車をこぐ2人の女性。回転する車輪へズームアップ。
Carousel	b	画面を素早く横切るメリーゴーランド。
Sprinkling	c	女性のバストショットとシャワー状の水をクロマキー合成。
Green Leaves	c	並木道をゆっくりズームアップ。木漏れ日によりコントラストの高い テクスチャを多く含んだ画像。
Kimono	d	彩度の高い細かい絵柄の着物のパンニング。
Rugby1	d	スタジアムのルーズショットからフィールド内のアップへカット切替。
Rugby2	d	観客席を背景としたフィールド内の選手の激しい動き。
Pop Jam	d	音楽ステージ。ストロボ状にライトの点滅有り。カット切替。
Nintama	d	アニメーション。カット切替有り。
Mobile (Super)	e	“Mobile&Calendar”に文字スーパー縦スクロール。
Mobile (Fade)	e	黒画面から“Mobile&Calendar”へフェードイン。再び黒画面へフェードアウト。

a: ITU-R 標準動画像、 b: MPEG 評価用シーケンス、
c: BTA ハイビジョン標準動画像のダウンコンバート画像、
d: NHK ハイビジョン番組のダウンコンバート画像、
e: コンピュータソフトウェアによる合成画像

SP@ML および MP@ML の評価結果を図 3.1 に示す。ITU-R 勧告 BT.1122[2]で規定する DSCQS12%以下の劣化は、MP@ML では 15Mbps で全ての画像で満足した。一方、SP@ML では“Green Leaves (GL)”は 15Mbps でも 12%以上の劣化であった。9Mbps では、幾つかのシーケンスで 12%以上の劣化となったが、平均的には 12%以内であった。“Susie” と“Nintama”は符号化が容易な画像であり、どちらのプロファイルにおいても、どのビットレートでも良好が画質となっている。“Mobile & Calendar (M&C)”, “Sprinkling”, “Mobile (Fade)”は SP@ML での劣化が大きく、“Rugby2”, “Cheerleaders” は MP@ML で劣化が大きい。“Mobile (Super)”と“Pop Jam”は、どちらのプロファイルに対しても大変クリティカルな画像であった。

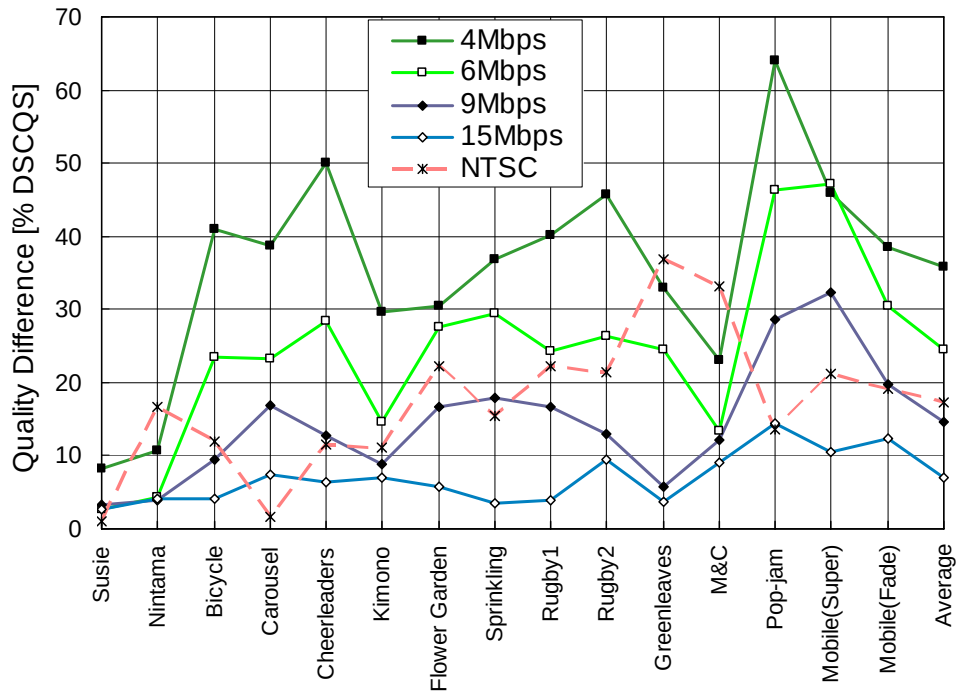


(a) SP@ML at 4H

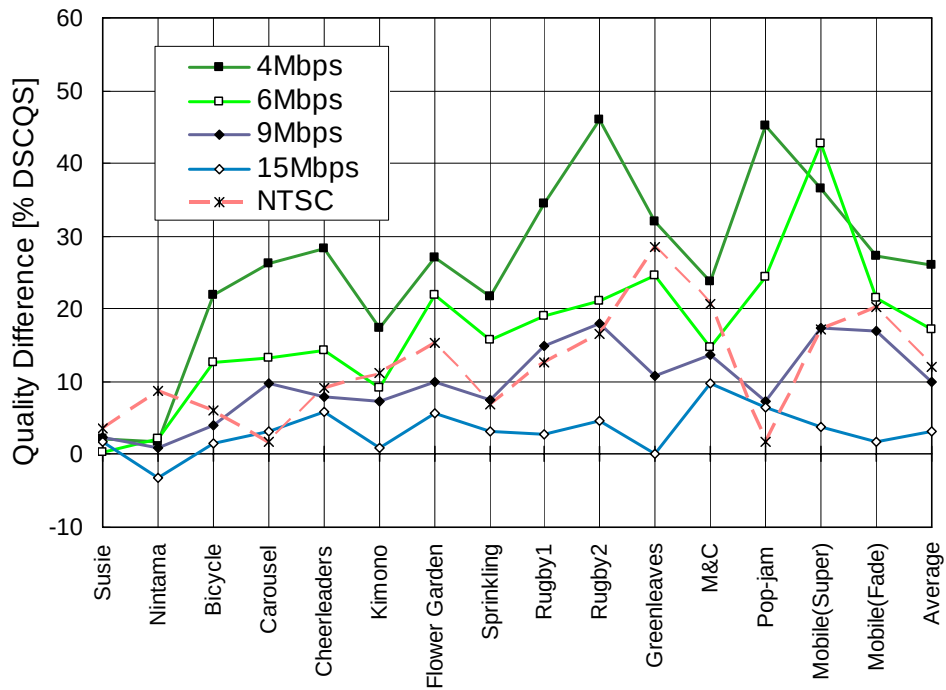


(b) SP@ML at 6H

図 3.1 SP@ML および MP@ML で符号化した各種画像の評価結果



(c) MP@ML at 4H



(d) MP@ML at 6H

図 3.1 SP@ML および MP@ML で符号化した各種画像の評価結果(続き)

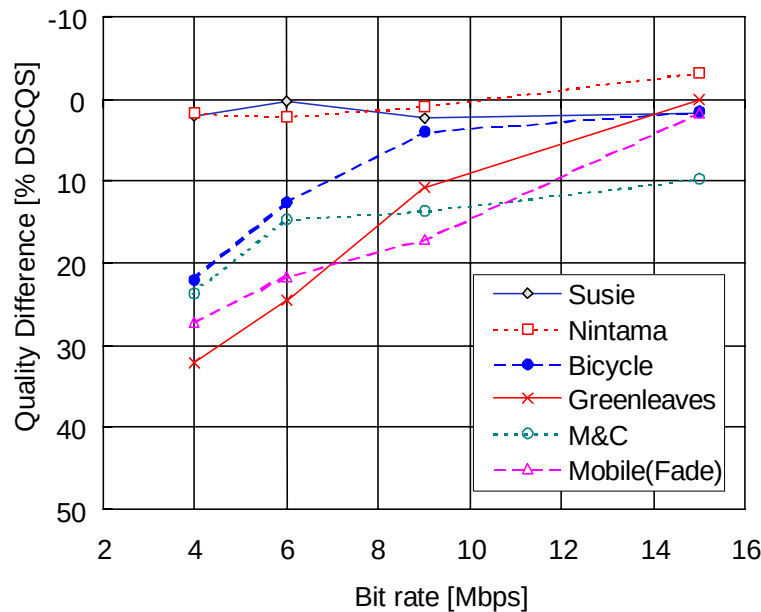


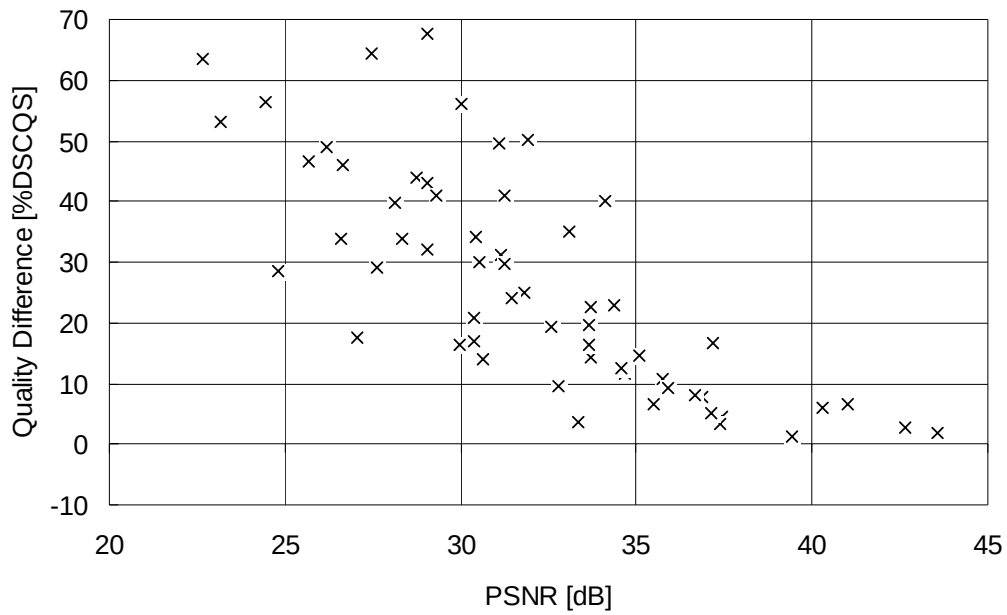
図 3.2 ビットレートと画質の関係 (MP@ML, 6H)

図 3.2 に、各画像について、ビットレートと画質劣化の関係を示す。動きの少ない“Susie”やアニメーションの“Nintama”では、ビットレートが低くなってもほとんど画質は低下しない。これに対して、細かいテクスチャを含む絵柄の“Green Leaves”やフェード画像“Mobile (Fade)”では、ビットレートが低くなると画質劣化が顕著となる。“Sprinkling”, “Rugby2”, “Mobile (Fade)”など、画像によって SP@ML と MP@ML で画質の違いが見られるものの、定性的には前章の図 2.6 で示したように、低いビットレートにおける画質は画像に大きく依存することが確認された。

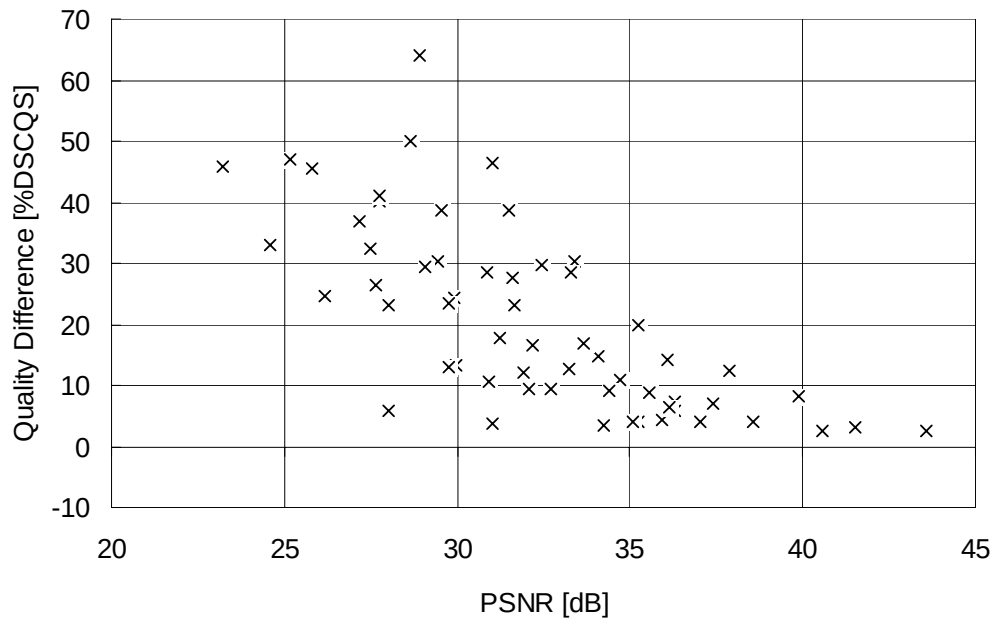
図 3.3 に、DSCQS 法による主観画質評価値と PSNR の関係を示す。SP@ML と MP@ML、あるいは視距離によってばらつき方は多少異なるが、表 3.3 に示すように相関係数はいずれも 0.7～0.8 となっている。この相関係数の値は、ITU で報告されている PSNR の相関係数 0.804 より多少低い値となっているが、本章の実験ではより絵柄的に特徴的なシーケンスを多く用いているためと考えられる。符号化劣化は、絵柄や符号化方式によって歪み方が異なるため、人間の視覚特性や認知特性などの要因が入ってくる主観画質では、PSNR の間で必ずしも相関が高くない場合がある。例えば、“Sprinkling”, “Mobile (Fade)”, “Pop Jam”は、他の同程度の PSNR の画像と比較して主観画質が低い結果となっているが、これらの画像では画面の一部に目立った局所的劣化が生じ、全体的な画質低下を招いたと考えられる。また、“Green Leaves”は、同程度の PSNR の画像よりも良い主観画質となっているが、これは、細かい木の葉のテクスチャに符号化劣化がマスキングされたために劣化が目立たなかったと考察される。

表 3.3 PSNR と画質(%DSCQS)の相関係数

SP@ML (M=1)		MP@ML (M=3)	
4H	6H	4H	6H
0.78	0.75	0.72	0.78

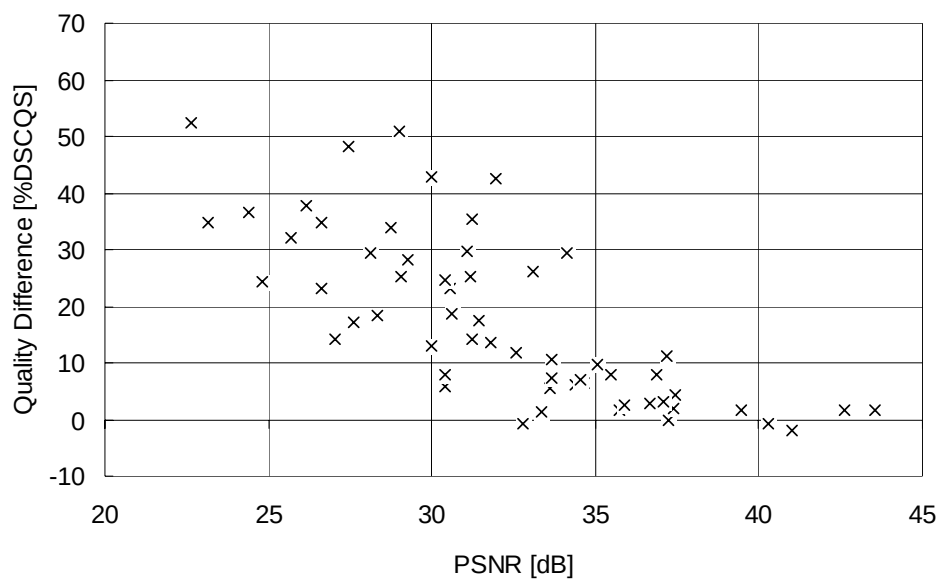


(a) SP@ML at 4H

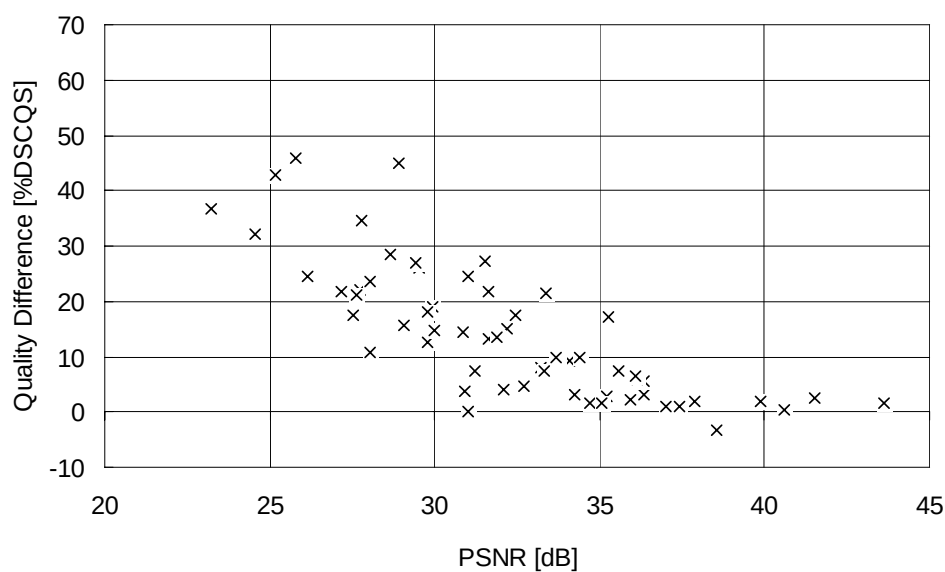


(b) MP@ML at 4H

図 3.3 PSNR と画質(%DSCQS)の関係



(c) SP@ML at 6H



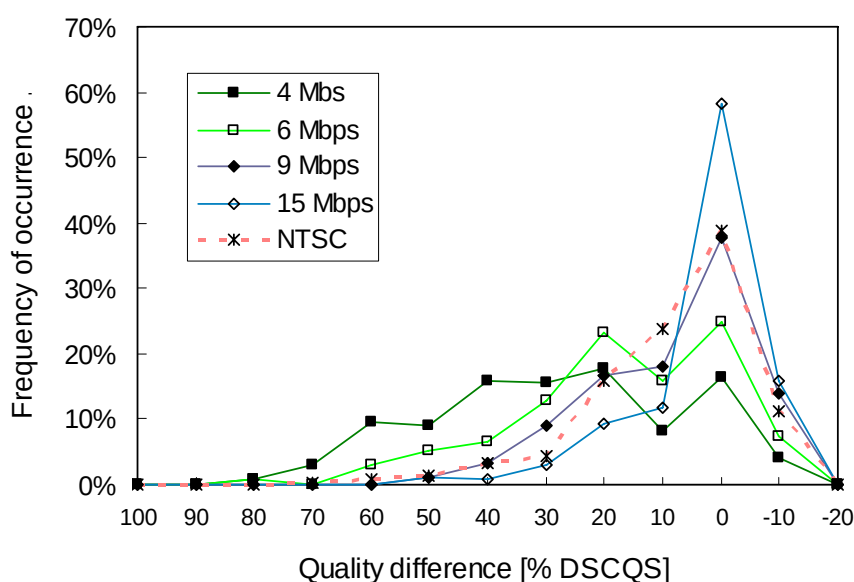
(d) MP@ML at 6H

図 3.3 PSNR と画質(%DSCQS)の関係(続き)

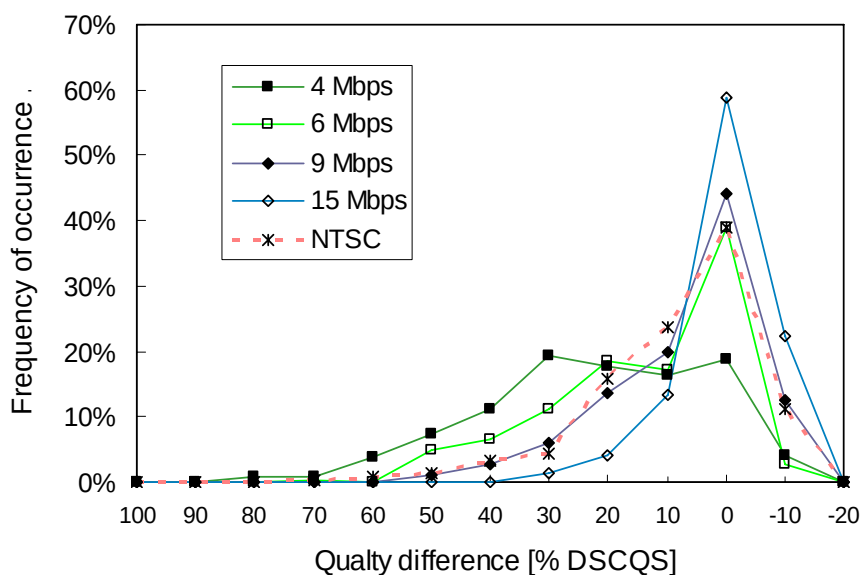
3.2.2 デジタル放送と NTSC との画質比較

図 3.4 は、6H での全ての評定者および評価シーケンスの評価データの頻度分布を示したものである。SP@ML、MP@ML とともに約 9Mbps でスタジオ NTSC と同等の分布となっている。SP@ML、MP@ML の平均値と標準偏差を図 3.5 に示す。ビットレートが低いと、平均的に劣化が大きくなるとともに、画像によるばらつきを示す標準偏差も増大する。ただし、MP@ML の標

準偏差は、SP@ML よりも小さく、メインプロファイルで符号化することで画像による画質のばらつきは小さくなることが示された。表 3.4 は、平均と標準偏差がスタジオ NTSC と同等となるビットレートを示す。これより、NTSC と同等の画質を得るには、メインプロファイルで約 8Mbps が必要ということがわかる。

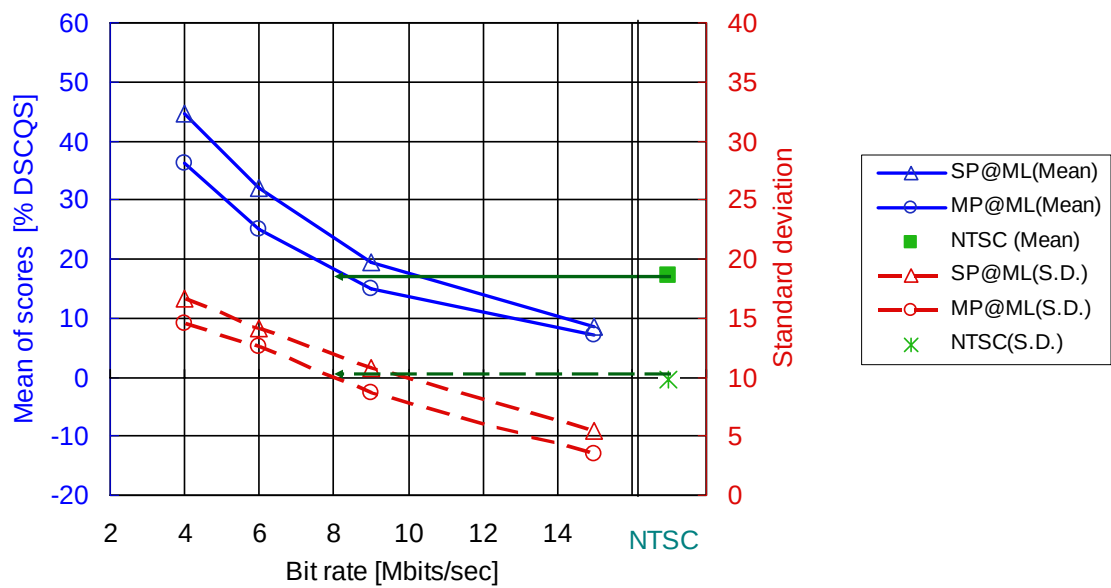


(a) SP@ML

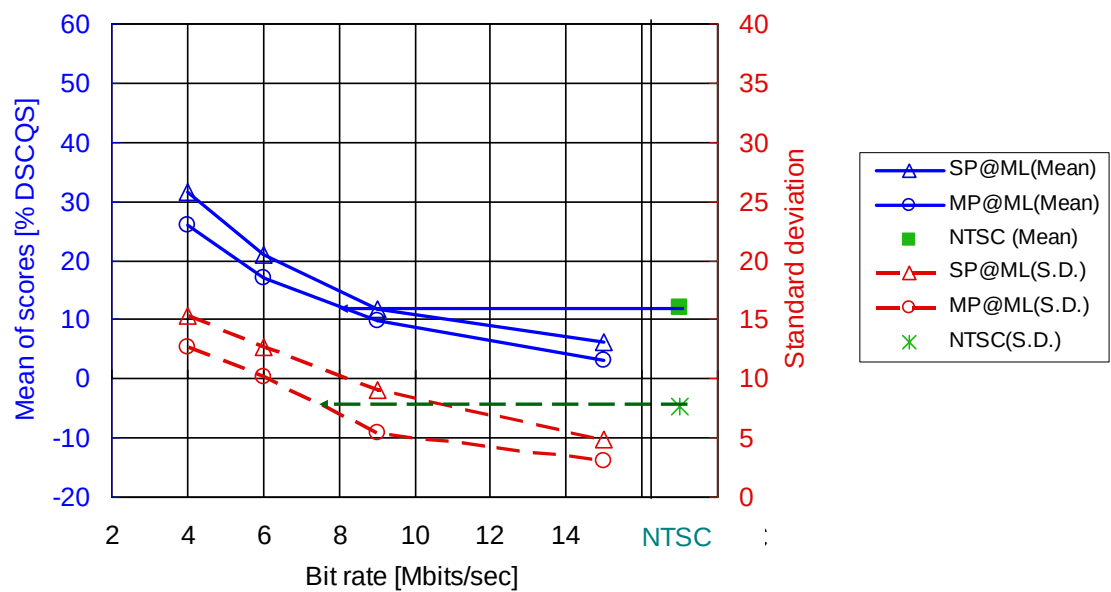


(b) MP@ML

図 3.4 視距離6H における評点分布 (全評価シーケンス)



(a) 視距離 4H



(b) 視距離 6H

図 3.5 評価結果の平均値と標準偏差

表 3.4 スタジオ NTSC と同等の品質を与えるビットレート

SP@ML		MP@ML	
4H	6H	4H	6H
10 Mbps	9 Mbps	8 Mbps	8 Mbps

3.3 クリティカリティの提案

クリティカリティを定義する際、リアルタイムの測定が可能で、ビットレートの違いや若干の符号化パラメータの違いにも適用可能な実用的定義が望まれる。これらの点を考慮し、本章ではクリティカリティの定義を新たに提案する。

Rate Distortion 理論より、入力信号の分散を σ_x^2 とした場合、ビットレート R と量子化誤差分散 σ_q^2 の関係は(3.1)式で表現される[3]。

$$R = \frac{1}{2} \log_2 \left(\frac{\sigma_x^2}{\varepsilon_q^2 \sigma_q^2} \right) \text{ [bit/sample]} \quad (3.1)$$

入力信号 x は、イントラ符号化時には映像信号、インター符号化時には予測誤差信号に相当する。 ε_q^2 は入力信号の統計的性質に応じたパラメータで、均一分布の時は $\varepsilon_q^2 = 1.00$ 、ガウス分布の場合 $\varepsilon_q^2 = 2.71$ 、ラプラス分布の場合 $\varepsilon_q^2 = 4.50$ である。2次元 DCT の AC 係数はラプラス分布で良く近似され、DC 係数はイントラ符号化時にガウス分布で、インター符号化時にはラプラス分布で近似されることが知られている[4]。

ここで、量子化特性を線形とし、その量子化ステップ幅 Δ が十分小さく入力信号が区間内で均一に分布すると仮定すると、 σ_q^2 と Δ との関係は次式で示される。

$$\sigma_q^2 = \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} (x^2 \times \frac{1}{\Delta}) dx = \frac{\Delta^2}{12} \quad (3.2)$$

2次元 DCT 符号化における各係数(m, n)の量子化誤差分散を $\sigma_{qc}^2(m, n)$ としたとき、ブロック内の平均誤差分散 σ_q^2 は次式で与えられる。

$$\sigma_q^2 = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \sigma_{qc}^2(m, n) \quad (3.3)$$

各係数に対する量子化ステップサイズ $\Delta_c^2(m, n)$ と量子化制御パラメータ Q (q_scale)、量子化マトリックス $w^2(m, n)$ との間に $\Delta_c(m, n) = Q \cdot w(m, n)$ の関係があるとき、 σ_q^2 は(3.2)、(3.3)式より次式となる。

$$\begin{aligned} \sigma_q^2 &= \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \frac{\Delta_c^2(m, n)}{12} \\ &= \frac{Q^2}{12MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} w^2(m, n) \end{aligned} \quad (3.4)$$

したがって、(3.1)式は(3.5)式で表現される。

$$R = -a_0 \log_2 Q + b_0 \quad (3.5)$$

ここで、

$$a_0 = 1 \quad (3.6)$$

$$b_0 = \frac{1}{2} \log_2 \left[12MN \cdot \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \frac{\sigma_{xc}^2(m, n, p)}{\varepsilon_{qc}^2(m, n, p) \cdot w^2(m, n, p)} \right] \quad (3.7)$$

である。すなわち、係数 a_0 は一定値(=1)となり、切片 b_0 は映像信号の性質と予測方式 (p=I,P,B のいずれか)に依存した量となる。

ここでは、実際の MPEG-2 符号化において、量子化ステップ幅と符号化誤差電力 MSE および発生ビット数 R_c の関係をシミュレーションにより確認する。エンコード法は TM-5[5]を用い、 $M=1$, $N=15$ で量子化器を固定して符号化した。

図 3.6 に量子化器固定($q_scale=8$)で符号化した時の符号化誤差 MSE と発生ビット数 R_c を示す。実験に用いた“M&C”と“GL”は絵柄的に大変異なった画像であり、それぞれの画像におけるカメラの動きも“M&C”はパンニング、“GL”はズームングと異なる。シミュレーションの結果から、P ピクチャ(Frame #2~15、#17~20)の MSE は双方の画像シーケンスではほぼ同じ値となっているが、 R_c はシーケンスに大きく依存し、“M&C”では絵柄の変化に伴って変化していることが認められた。

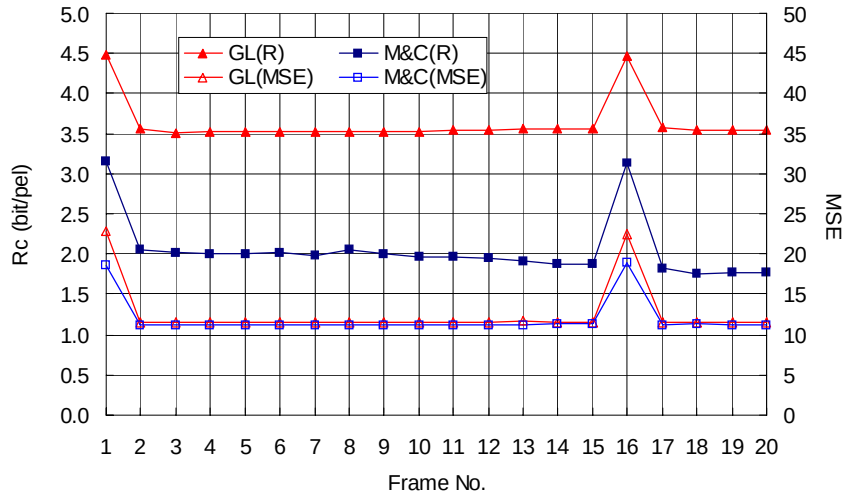
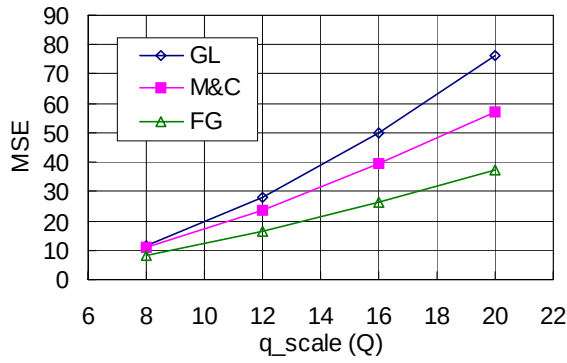
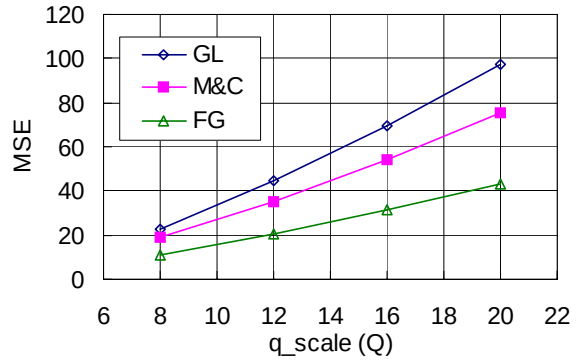


図 3.6 量子化器固定($q_scale=8$)時の発生情報量 R と符号化誤差 MSE の関係

量子化ステップ幅がさらに大きくなると、量子化区間内の信号分布は映像信号の性質に依存するため、(3.5)式の関係が成り立たず、符号化誤差量は絵柄によって変わってくると考えられる。量子化ステップ幅を変えたときの符号化誤差電力 MSE を図 3.7 に示す。量子化ステップ幅が大きくなるにしたがって画像間の誤差量の差が大きくなっている。図 3.8 に示すように、発生ビット数との関係においても、量子ステップ幅が大きくなると(3.5)式における傾き a_0 の値が画像によって異なる結果となっている。

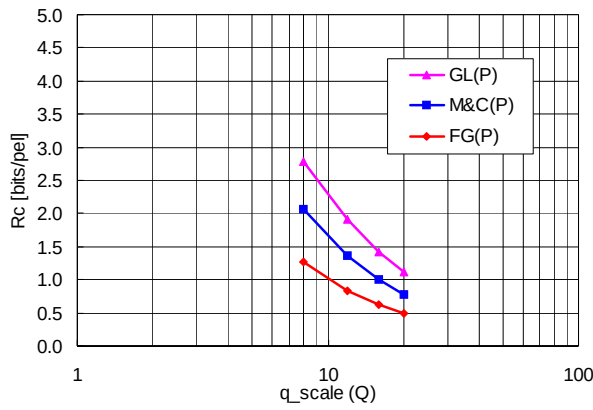


(a) P ピクチャ

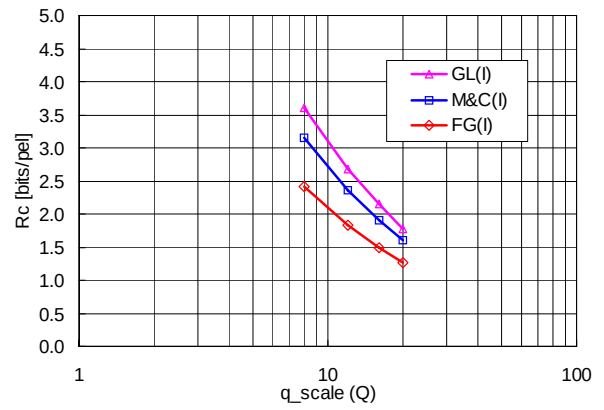


(b) I ピクチャ

図 3.7 量子化ステップ幅と MSE の関係



(a) P ピクチャ



(b) I ピクチャ

図 3.8 量子化ステップ幅と発生ビット数の関係

一方、ステップ幅が比較的大きい場合も含め、(3.8)式の関係が一般に成り立つことが報告されている[6][7]。

$$\log_2 R_c = -a \log_2 Q + b \quad (3.8)$$

(3.8)式は、次式の形式でも表現される。

$$R_c \cdot Q^a = 2^b \quad (3.9)$$

異なる量子化ステップ幅 Q (q_scale)を与えたときの I および P ピクチャでの発生ビット数 R_c [bit/pel]の関係を図 3.9 に示す。

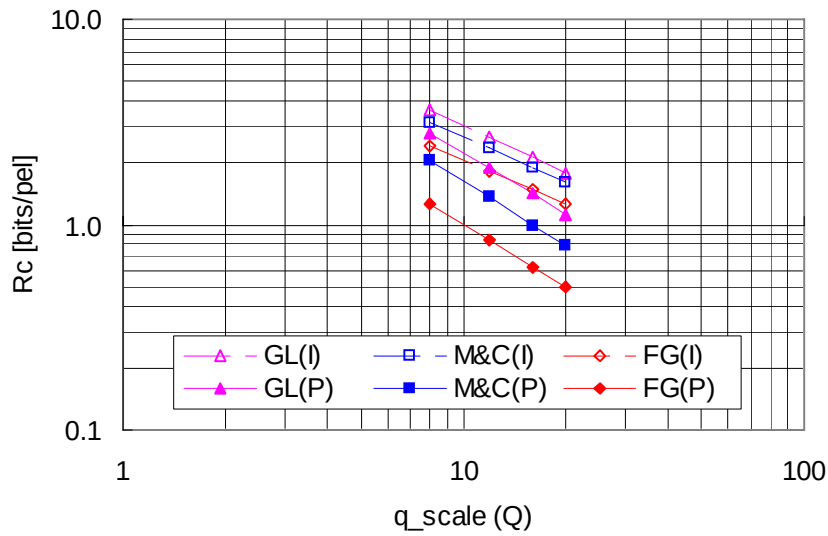


図 3.9 量子化ステップ幅 Q (q_scale) と発生ビット数 R の関係

この結果より、傾き a は各画像ではほぼ一致し、Pピクチャで $a=1$ となることが確認された。一方、切片 b の大きさは画像の種類と予測方式に依存している。この性質は、式(3.6), (3.7)の関係と良く一致する。図 3.10 に Q と $(Rc \cdot Q)$ の関係を示す。

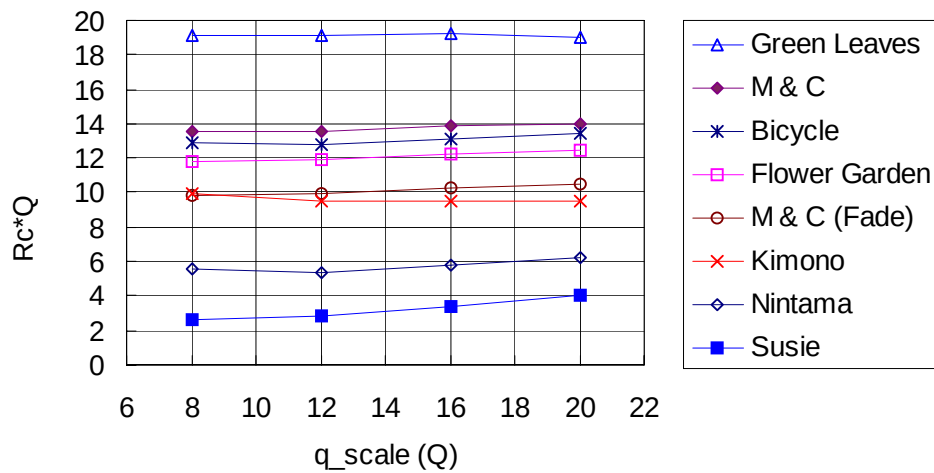


図 3.10 量子化ステップ幅 Q と $(Rc \cdot Q)$ の関係

図からわかるように、Pピクチャの $(Rc \cdot Q)$ は量子化精度(ビットレート)に依存しない画像固有な値となる。そこで、画像の符号化難易度を示すクリティカルティを次式で定義することとする。

$$Criticality = X = R_c \cdot Q \quad (3.10)$$

ここに、 R_c は P ピクチャにおける発生ビット数[bit/pel]、 Q は MPEG-2 の量子化制御パラメータ q_scale である。(3.10)式の定義は TM-5 の符号量制御で用いるパラメータ“global complexity measure”と等価である。

(3.10)式より、クリティカリティ値 X の画像を固定レート R_0 で符号化したときの平均 Q は次式となる。

$$Q = X/R_0 \quad (3.11)$$

一般に、動き補償予測誤差信号の DCT 係数をステップ幅 Q で量子化したときの全係数平均誤差電力(MSE)は Q に比例する[6]ことから、(3.11)式より固定レート符号化時の誤差電力は画像のクリティカリティ値と比例関係にある。

$$MSE = \frac{\alpha}{R_0} \cdot X \quad (\alpha: \text{定数}) \quad (3.12)$$

符号画像の画質と符号化誤差電力との関係では、人間の視覚感度特性の他、歪みの局所的な性質や絵柄によるマスキング効果等を考慮する必要がある[7][8]。(3.10)式で定義したクリティカリティは、図 3.11 に示すような量子化マトリックス $w(m, n)$ で視覚的な重み付けを行っているものの、入力画像の大局的な性質を考慮した量であり、局所的な歪みの影響や視覚的マスキング効果の要素は含んでいない。このため、符号化結果に基づいて計算される客観評価法の結果と比較すると、個々の画像における相関は必ずしも十分高くはないと予想される。但し、クリティカリティは、多くの画像サンプルを対象にした統計的な解析への適用を目的としており、この目的から十分な精度を示すか、3.4 章で検証する。

		→ u															
↓ v		8	16	19	22	26	27	29	34	16	16	16	16	16	16	16	16
		16	16	22	24	27	29	34	37	16	16	16	16	16	16	16	16
		19	22	26	27	29	34	34	38	16	16	16	16	16	16	16	16
		22	22	26	27	29	34	37	40	16	16	16	16	16	16	16	16
		22	26	27	29	32	35	40	48	16	16	16	16	16	16	16	16
		26	27	29	32	35	40	48	58	16	16	16	16	16	16	16	16
		26	27	29	34	38	46	56	69	16	16	16	16	16	16	16	16
		27	29	35	38	46	56	69	83	16	16	16	16	16	16	16	16
W[0][v][u]: intra		W[1][v][u]: non intra															

図 3.11 量子化マトリックスの値

3.4 クリティカリティと画質の関係

3.4.1 クリティカリティ測定装置

(3.10)式で定義したクリティカリティを実時間で測定するため、クリティカリティ測定装置を製作した。クリティカリティ測定装置のブロック図を図 3.12 に示す。クリティカリティ値は映像信号のフレーム毎に計算され、ディスクに蓄積される。本装置の開発により、放送番組のクリティカリティを長時間にわたり測定することが可能となった。

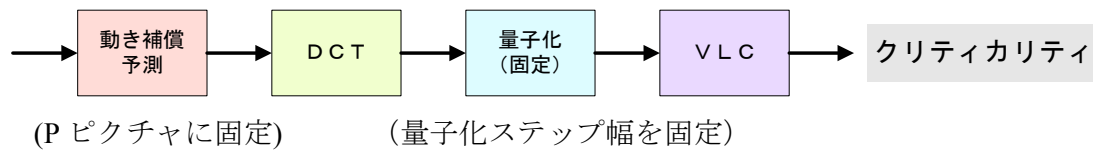


図 3.12 クリティカリティ測定装置のブロック図

クリティカリティ測定装置では、全て P ピクチャ(フィールド単位の順方向予測のみ)として符号化し、リフレッシュはイントラ・マクロブロックで行っている。従来のクリティカリティの定義では、実際の固定ビットレート符号化と同様な動作で測定を行い、その符号化条件およびビットレートにおける符号化劣化の推定しかできなかった。本提案方式では、動き予測および量子化に関する適応処理を固定して測定することで、他の符号化条件およびビットレートでの画質劣化の推定を可能とした。

量子化特性、量子化マトリックスなどは MPEG-2 エンコード・モデル TM-5 と同様であり、量子化器はイントラおよびインターとも線形特性である。イントラ/インター判定、フレーム/フィールド DCT の適応処理も TM-5 と同様に行っている。量子化精度を示す q_scale は 2 から 62 の値のうち 2 刻みで任意に設定することができる。

3.4.2 評価画像のクリティカリティ測定

各評価画像のクリティカリティを測定した結果(5 秒間の平均値と標準偏差)を図 3.13 に示す。クリティカリティはフレーム毎に測定し、測定にあたっては量子化精度を予備実験の結果から $q_scale=12$ に設定した。多くの画像は 5 秒間ほぼ一定のクリティカリティ値を示すが、“Mobile (Fade)”, “Rugby1”, “Pop Jam”, “Sprinkling”などではクリティカリティ値が大きく変動する。図 3.14 に“Mobile(Fade)”, “Sprinkling”, “Nintama”の各シーケンスにおけるクリティカリティの時間変化を示す。“Mobile (Fade)” は、はじめの 1.5 秒 (1~45 Frame)で黒から“Mobile&Calendar”へフェードインし、終わりの 1.5 秒 (105~150 Frame)で再び黒にフェードアウトするシーケンスである。フェードの程度に応じてクリティカリティ値が変化している。“Pop Jam”は照明がストロボ効果を狙って細かく変化するため、クリティカリティ値も細かく変動している。

“Nintama”はアニメーション画像でクリティカルティは低いが、カット切替時にクリティカルティが急激に増大する。また、“Sprinkling”は水しぶきの背景と女性をクロマキー合成した画像で、水しぶきの量に応じてクリティカルティ値も変動している。

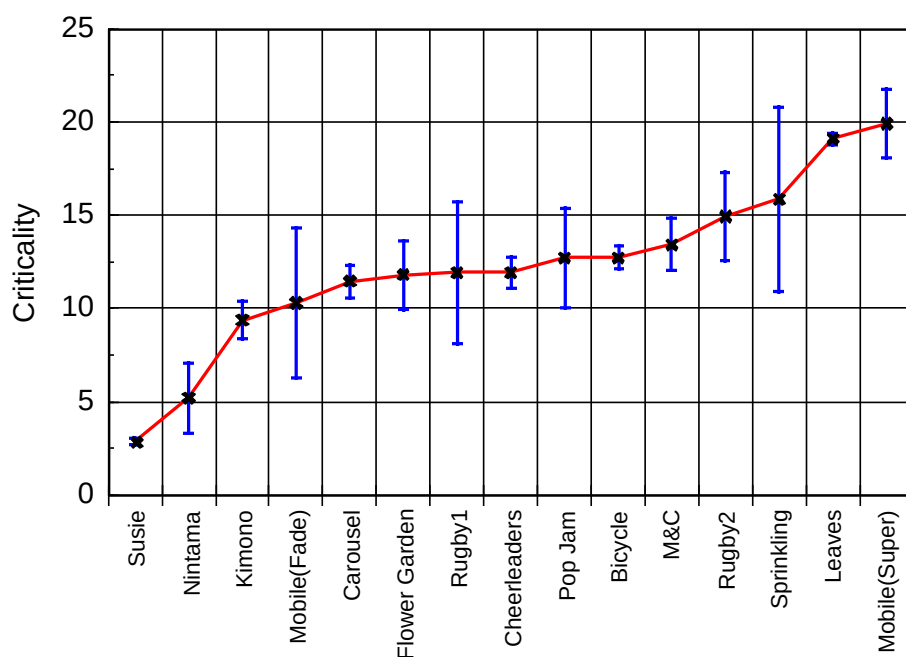


図 3.13 テスト画像のクリティカルティ値 (5 秒間の平均と標準偏差)

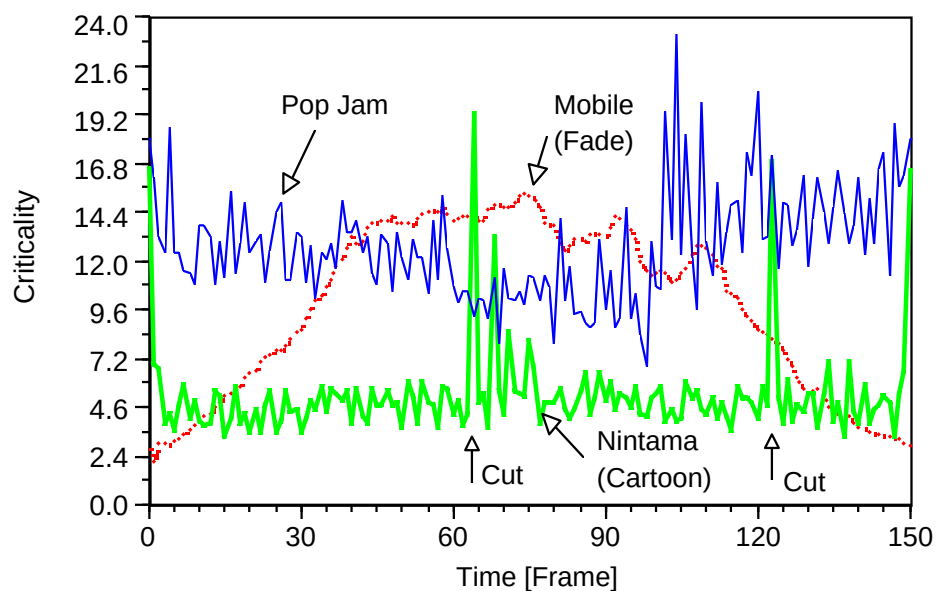


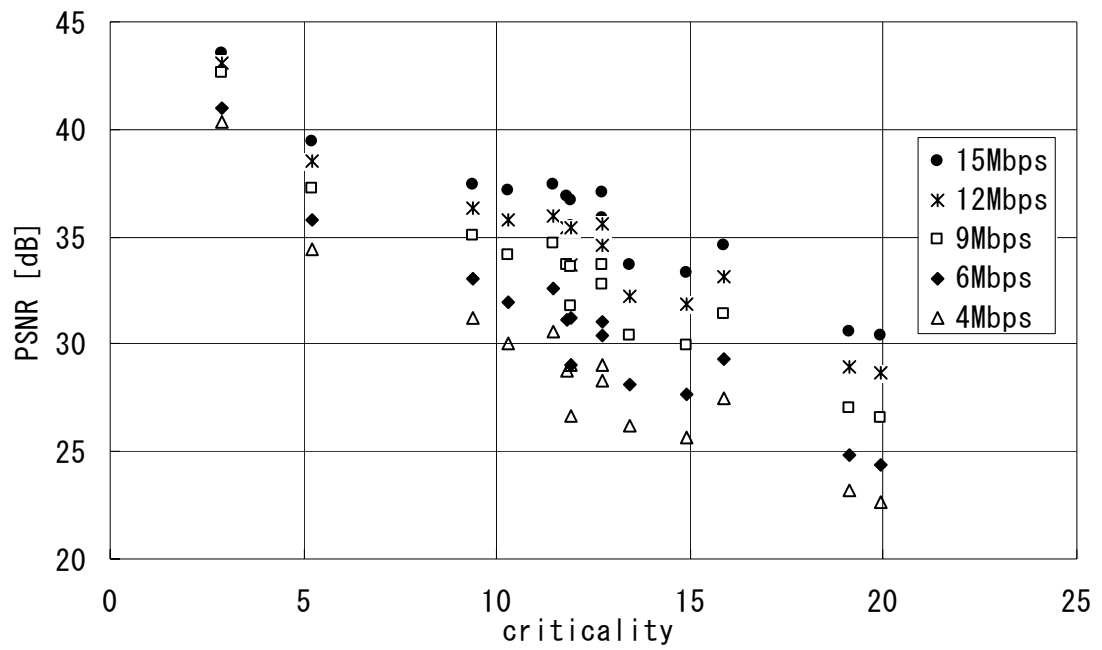
図 3.14 クリティカルティの時間変化

3.4.3 クリティシティと画質の関係

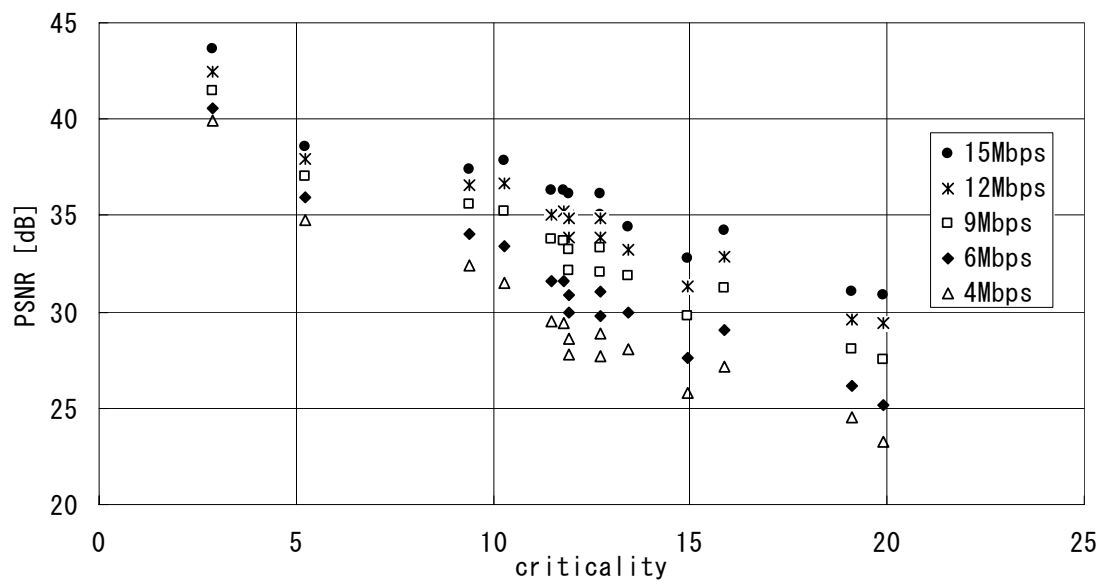
ここでは、各画像のクリティシティと画質の関係を調べる。図 3.15 にクリティシティ(シーケンスの平均値)と符号化誤差(PSNR)の関係を示す。ここに、符号化は、シンプルプロファイル(SP@ML, M=1)およびメインプロファイル(MP@ML, M=3)を用い、それぞれ固定レート符号化でビットレートを変えて測定した結果である。画像によって、予測方式の異なる SP@ML と MP@ML で符号化誤差は多少異なるが、ほぼクリティシティの大きさに応じた誤差量となっている。表 3.5 にクリティシティ値と PSNR の相関係数を示すが、いずれも-0.97~-0.95 と高い相関となっている。同程度のクリティシティの画像に対して、“Sprinkling”(クリティシティ:15.9)の PSNR は高く、“Rugby2”(14.9), “M&C”(13.4), “Rugby”(11.9), “Nintama”(5.2)の PSNR は低い結果となっている。SP@ML と MP@ML の PSNR を比較した場合、動きが速く複雑な“Carousel”, “Bicycle”は SP@ML が良く、“Kimono”, “M&C (Fade)”, “Rugby1”, “M&C”, “Green Leaves”は MP@ML が良いなど、符号化方式による差も見られる。

表 3.5 クリティシティと各種ビットレートにおける PSNR の相関係数

	15Mbps	12Mbps	9Mbps	6Mbps	4Mbps
SP@ML	-0.96	-0.96	-0.96	-0.96	-0.95
MP@ML	-0.96	-0.97	-0.97	-0.97	-0.96



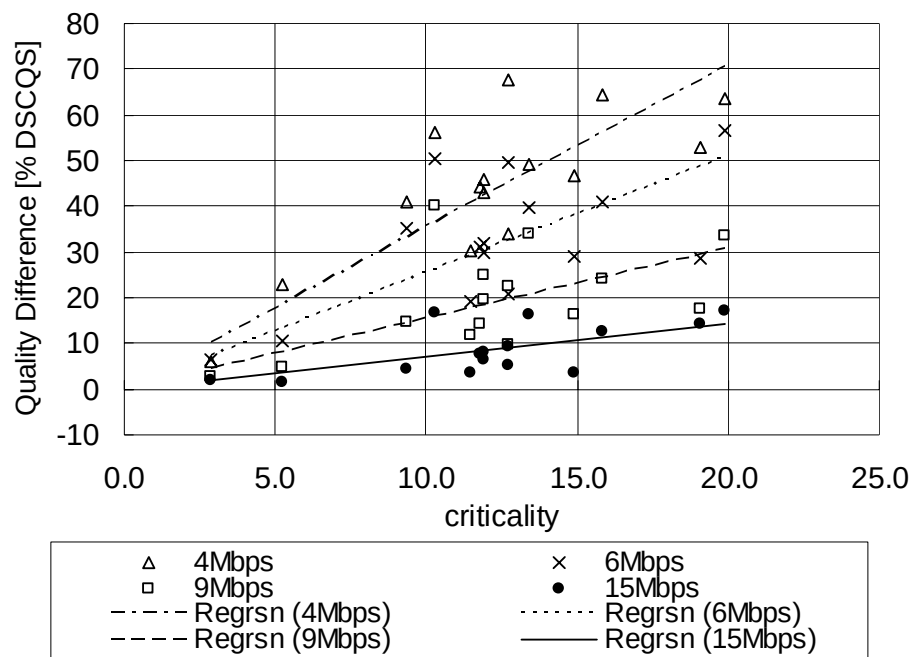
(a) SP@ML



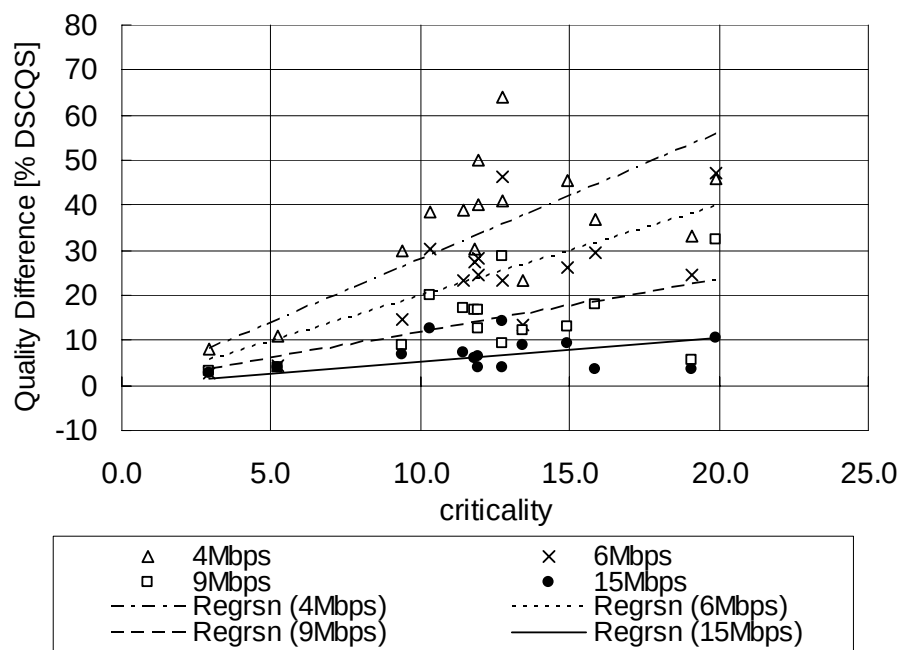
(b) MP@ML

図 3.15 クリティカリティと符号化誤差(PSNR)との関係

図 3.16 に、3.2 章で示した DSCQS 法による主観評価結果とクリティシティの関係を示す。

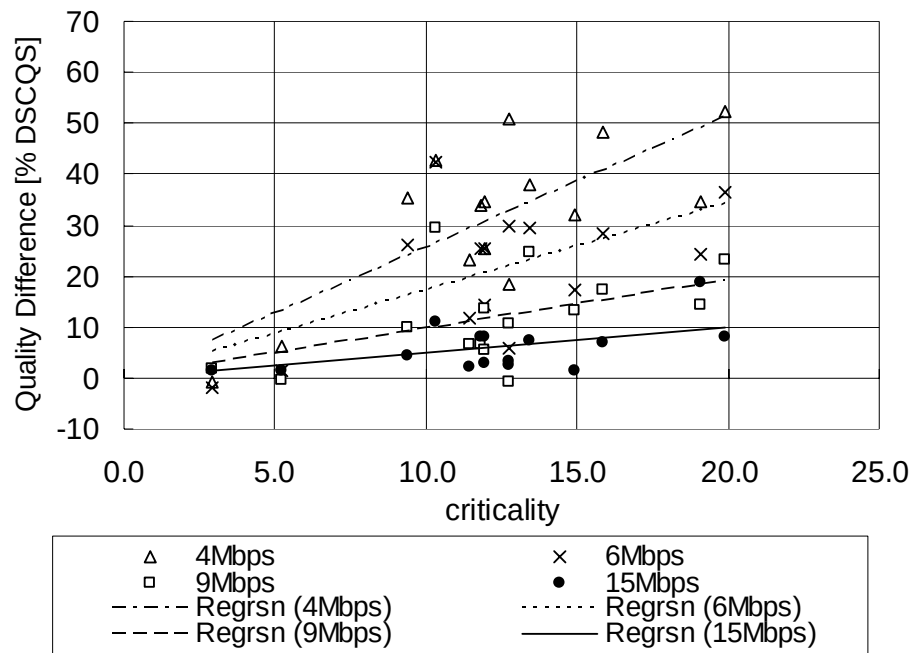


(a) SP@ML at 4H

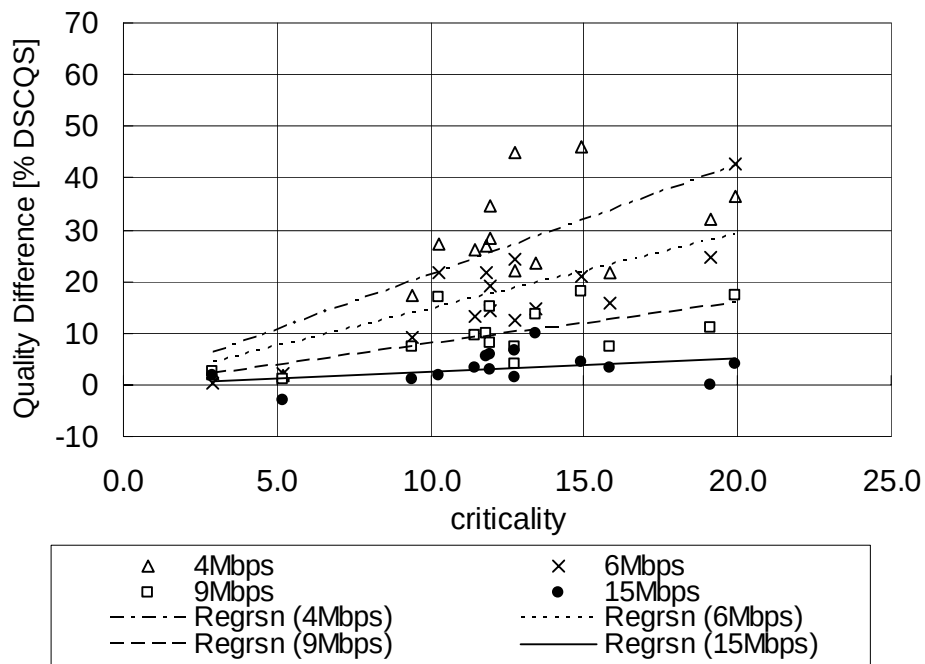


(b) MP@ML at 4H

図 3.16 画質とクリティシティの関係



(c) SP@ML at 6H



(d) MP@ML at 6H

図 3.16 画質とクリティカリティの関係(続き)

デジタル符号化による劣化は局所性を持つため、空間あるいは時間的に顕著な劣化が局所的に生じる画像に対しては、シーケンスの平均値を用いたクリティカリティでは画質推定が難しい。例えば、“Pop Jam”や”Mobile(Fade)”のように時間的に絵柄が変動し、局所的に劣化が発生するシーケンスでは、実際の主観画質は平均的クリティカリティで推定した画質より悪くなる傾向がある。また、符号化方式によって画質の差が大きいシーケンスに対しても、画像特徴量として定義するクリティカリティを用いた画質推定は誤差が生じやすい。

前章の考察より、画質評価値とクリティカリティ値は比例関係にあると仮定し、最小自乗法により直線回帰を行った。表 3.6 に回帰分析の結果を示す。回帰係数を t 検定した結果、いずれ

表 3.6 各種条件における回帰係数と検定結果

(a) 視距離 6H

Profile Level	bit rate	決定係数 R^2	回帰係数	標準誤差	t	回帰係数の有意性 (1%有意水準)	95%信頼区間
SP@ML	4Mbps	0.551	2.58	0.20	12.66	有意	0.44
	6Mbps	0.346	1.72	0.20	8.42	有意	0.44
	9Mbps	0.271	0.96	0.15	6.27	有意	0.33
	15Mbps	0.305	0.49	0.08	6.23	有意	0.17
MP@ML	4Mbps	0.531	2.13	0.17	12.25	有意	0.37
	6Mbps	0.697	1.46	0.12	12.26	有意	0.26
	9Mbps	0.363	0.80	0.09	9.32	有意	0.18
	15Mbps	0.103	0.26	0.06	4.46	有意	0.12

(b) 視距離 4H

Profile Level	bit rate	決定係数 R^2	回帰係数	標準誤差	t	回帰係数の有意性 (1%有意水準)	95%信頼区間
SP@ML	4Mbps	0.609	3.55	0.21	16.57	有意	0.46
	6Mbps	0.394	2.55	0.23	11.30	有意	0.48
	9Mbps	0.286	1.55	0.18	8.50	有意	0.39
	15Mbps	0.437	0.71	0.08	8.55	有意	0.18
MP@ML	4Mbps	0.351	2.81	0.24	11.47	有意	0.53
	6Mbps	0.507	1.99	0.17	11.43	有意	0.37
	9Mbps	0.273	1.16	0.14	8.22	有意	0.30
	15Mbps	0.062	0.53	0.08	7.02	有意	0.16

の場合も 1%有意水準で回帰が有意であることが確かめられた。また、クリティカリティは P ピクチャの発生ビット数で定義しているが、双方向予測(B ピクチャ)を用いる MP@ML についても SP@ML と同等の相関が得られた。表 3.6 の結果から、画質とビットレートの関係調べる目的において、プロファイルおよび視距離が異なる場合についてもクリティカリティの有効性が確認された。すなわち、個々のシーケンスに対してはクリティカリティと画質の関係にばらつきはあるものの、長時間の映像データをもとにシステムパラメータや運用条件を検討する目的において、提案したクリティカリティによる画質推定法は有用な情報を提供するものと認められる。

これは、実用化レベルのエンコーダでは、局所的な画像特徴やシーンチェンジ、視覚的な特性などを考慮して量子化や符号量制御等の最適化がなされており、局所的な性質よりも大局的な画像の性質(エントロピー)が符号化画像の画質の支配要因となっているためと考えられる。したがって、大局的な画像性質としてクリティカリティを定義した場合でも、多くのシーケンスを対象にした画質推定で十分な精度を示したものと考ええる。さらに、シーケンス内でのクリティカリティの時間変動や、画面内の領域によるクリティカリティの分散などを考慮することで、個々のシーケンスについての画質推定精度を高めることができると考える。

次に、(3.12)式の関係が主観評価値とクリティカリティ値の間で成立するかを確認する。一般的には、画質と符号化 MSE とは必ずしも高い相関を示さないが、ここでは同一の符号化アルゴリズムを用い、多くの画像サンプルを対象とすることを前提としている。各種条件における回帰係数の大きさとビットレート R の関係を図 3.17 に示す。この結果より、回帰係数はビットレートの逆数(1/R)とほぼ比例関係にあると認められ、画質評価値(Q.D.)とクリティカリティとの間にも(3.12)式と同様な関係が成立することが確認された。

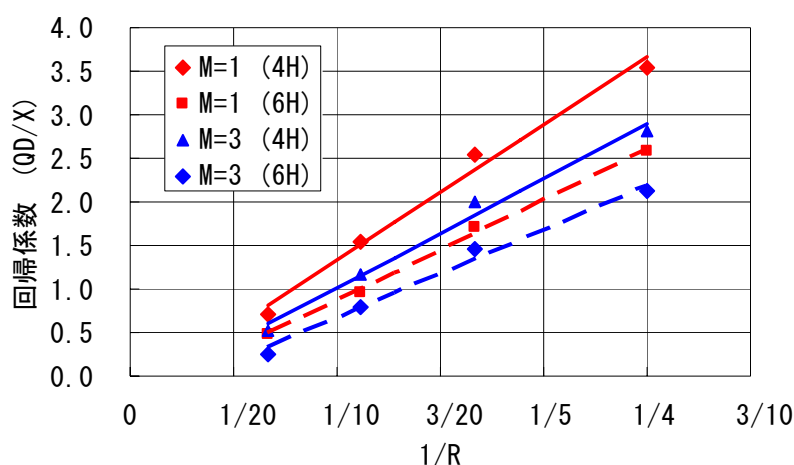


図 3.17 ビットレートと回帰係数の関係

$$Q.D. = \frac{\beta}{R} \cdot X \quad (\beta: \text{定数}) \quad (3.13)$$

ここで、 β はプロファイル(予測方式)あるいは視距離等で決まる定数である。

以上の結果から、本提案のクリティカルティは符号化劣化の絵柄依存性を適切に表現し、(3.13)式の関係でビットレートと画質の関係を表現する有効な定義となっている。

3.5 まとめ

デジタル映像符号化では、映像内容に応じて画質変動が生じやすいことが知られている。デジタル圧縮の厳しさの指針となりうる画像のクリティカルティを「固定量子化したときのPピクチャの発生情報量と量子化幅 q_scale の積」と定義し、各種画像のクリティカルティ値を測定した。クリティカルティは量子化幅によらず、画像に固有な値であることを各種画像の符号化実験で確認した。さらに、クリティカルティと各種符号化条件における画質評価値との関係を調べた結果、ある程度多くの画像を対象とした場合、直線回帰により画質を有効に推定できることが明らかとなった。

これにより、放送番組のクリティカルティを連続的に測定することで、デジタル放送の画質分布を各種ビットレートに対して求め、サービス品質を時間率で評価できることが示された。

参考文献

- [1] ISO/IEC 13818-2 | ITU-T Rec. H.262, “Information Technology - Generic Coding of Moving Pictures and Associated Audio: Video” (1995)
- [2] Rec. ITU-R BT.1122-1, “User Requirements for Emission and Secondary Distribution Systems for SDTV, HDTV and Hierarchical Coding Schemes” (1995)
- [3] N.S.Jayant, P.Noll: “Digital coding of waveforms”, Englewood Cliffs NJ, Prentice-Hall (1984)
- [4] R.C.Reininger, J.D.Gibson, “Distributions of the two-dimensional DCT coefficients for images,” IEEE Trans. Comm., COM-31, 6, pp.835-839 (1983)
- [5] Test Model Editing Committee, “Test Model 5,” ISO/IEC JTC1/SC29/WG11/N0400 (1993)
- [6] 甲藤二郎, 太田睦, “MPEG 圧縮効率の理論解析とその符号量制御への応用,” 信学技報, IE-95-10, pp.71-78 (1995)
- [7] 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おける動画像符号化制御方式,” テレビ誌, 50, 7, pp958-964 (1996)
- [8] 中須英輔, 金次保明, 矢島亮一, 青木勝典, 久保田啓一, “MPEG-2 映像符号化における画質の絵柄依存性の検討～画像のクリティカルティ～,” 映メ学誌, 51, 9, pp.1568-1576 (1997)

第4章 デジタル放送の画質の統計的解析

4.1 はじめに

放送システムへの応用においては、画質、チャンネル切替時間、符号化遅延時間、伝送エラー耐性など、様々なユーザ要求を考慮して、符号化パラメータを十分吟味することが望まれる。なかでもビットレートと画質の関係は、サービスの質を考える上で最も重要な要因である。

- 画質は映像の内容に依存して変わる。
- 各種符号化パラメータによってビットレートと画質の関係は変わる。
- 画質は主にエンコーダの性能によって決まる。このため、エンコード方法の改善によって画質とビットレートの関係は変わる。

これらのことを考慮すると、パラメータの選択やエンコード方法の改善にも柔軟に対応して、それぞれの場合の放送サービスにおける画質を推定できる手法の開発が望まれる。

第3章において、画像の性質を物理量として表現する「クリティカリティ」を定義し、測定装置を開発した。ここでは、この装置を使って長期間、放送番組のクリティカリティを測定した結果について述べる。この結果に基づき、先に求めた主観的な画質評価の結果とクリティカリティとの関係を用いて、放送番組における画質を時間率で評価する手法を提案する。これにより、各種ビットレートにおける放送番組の画質を時間率で評価し、番組のジャンルによる画質の違いについて議論する。さらに、MPEG-2 規格のシンプルプロファイルとメインプロファイルの画質差、視距離による画質差についても議論する。

4.2 クリティカリティに基づく放送番組の画質の統計的解析手法

放送番組のクリティカリティに基づき、デジタル放送の画質を時間率の面から統計的に評価する手法について述べる。図 4.1 に提案する解析手順を示す。

デジタル符号化の画質は主にエンコーダの性能によって決まり、エンコード方法の改善によって画質とビットレートの関係は変わる。このため、パラメータの変更やエンコード方法の改善にも柔軟に対応できる画質解析法が望ましい。本手法では、異なる符号化条件における画質やアルゴリズムの改善効果を確認する場合にも、評価対象システムの主観評価 (Step 3) を従来通り実施すれば良い。最も測定に時間がかかり、同一映像による測定が極めて困難である放送番組の測定(Step 1)結果は、他のシステムの解析にも用いることができる。

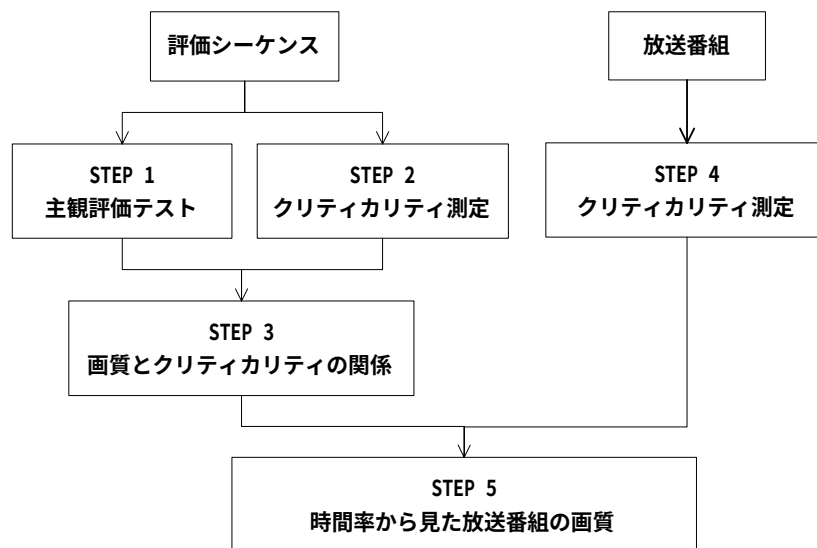


図 4.1 放送番組の画質(時間率)を求める手順

- step 1: テスト画像を各種ビットレートで符号化し、主観評価テストにより、その画質を求める。
- step 2: 主観評価テストに用いたテスト画像のクリティカリティを測定する。
- step 3: 画質(step 1 の結果)とクリティカリティ(step 2 の結果)の関係を導く。
- step 4: 放送番組のクリティカリティを測定し、クリティカリティの発生頻度(時間率)を求める。
- step 5: 画質とクリティカリティの関係(step 3 の結果)とクリティカリティの頻度分布から、放送番組の画質の頻度分布(時間率)を求める。

4.3 放送番組のクリティカリティ

4.3.1 放送番組のクリティカリティの測定

NHK 総合テレビのクリティカリティを1週間にわたり、放送直前の放送局内の信号を用いて測定した。測定は 1995 年 2 月 15 日 17 時から 22 日 17 時まで行い、測定した時間は深夜の放送していない時間帯を除く約 129.5 時間(約 1400 万フレーム)である。クリティカリティ測定装置の入力はコンポーネント信号であるため、放送用の NTSC 信号をデコードした後に測定装置に入力した。測定は、 $q_scale=12$ ($quantizer_scale_code=6$)で行った。

1週間のクリティカリティの度数分布を図 4.2 に示す。平均値は 5.28、標準偏差が 2.46、モード(最大の度数を示す値)が 3.36 となる分布が得られた。また、累積頻度で示した結果を図 4.3 に示す。

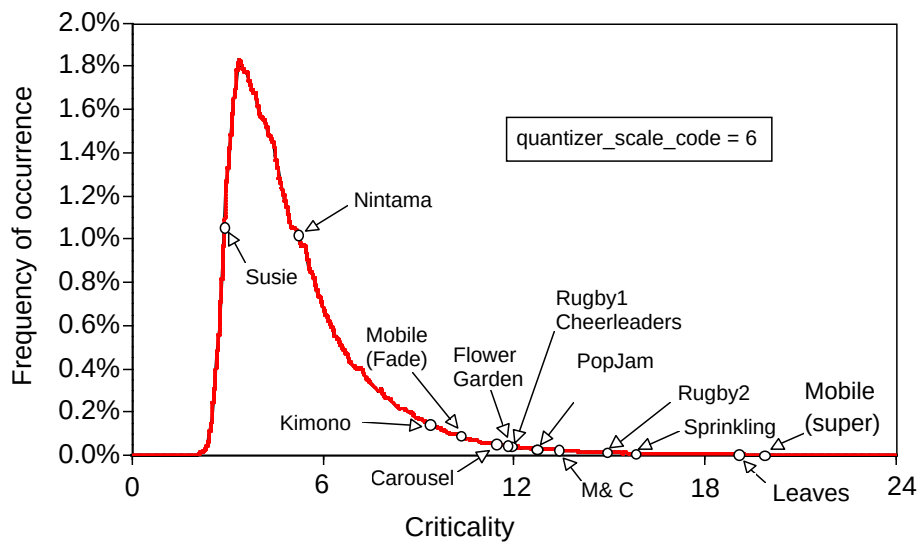


図 4.2 総合テレビ1週間のクリティカルティ分布

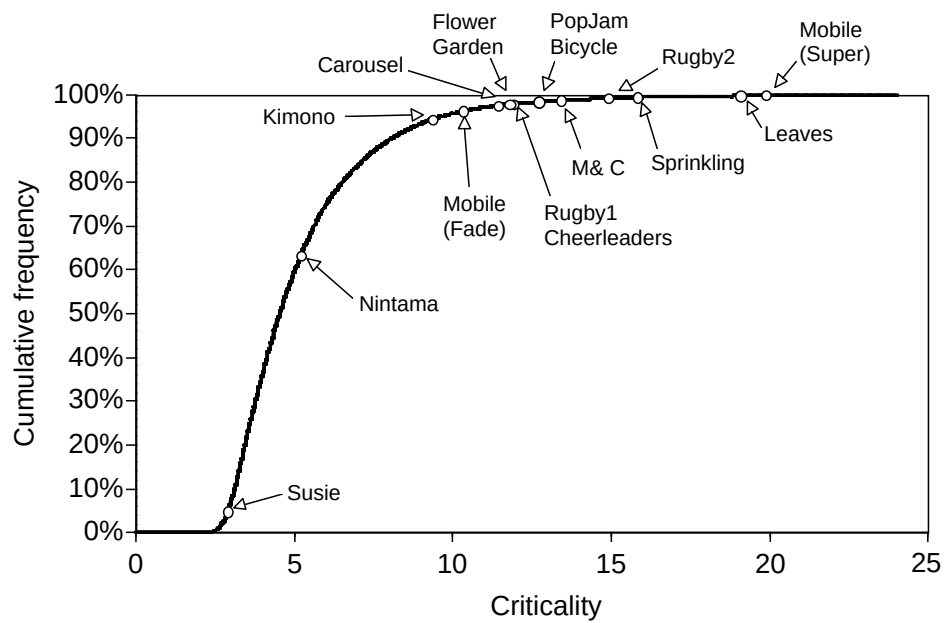


図 4.3 放送番組のクリティカルティ分布と評価画像のクリティカルティ値

4.3.2 番組ジャンルによるクリティカリティの違い

図 4.4 に1日の中のクリティカリティの変動を示す。番組によってクリティカリティの大きさが変わる傾向が認められる。

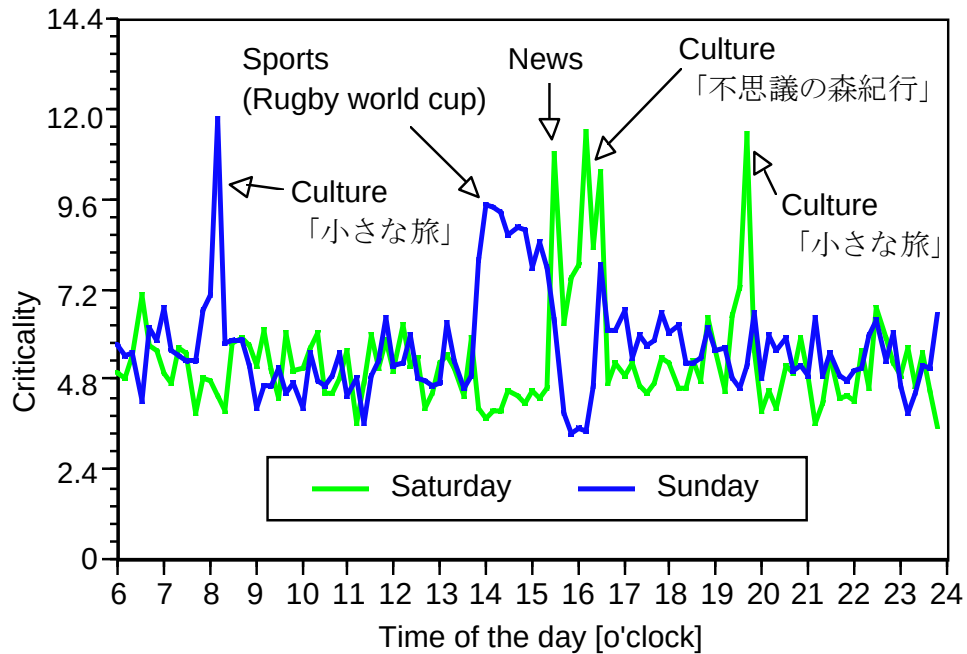


図 4.4 一日の放送におけるクリティカリティの変動

1週間の番組を、ドラマ、教養(含ドキュメンタリ)、ニュース、スポーツ、バラエティの5種類に分類して解析を行った。ジャンル別に分類した番組のクリティカリティの度数分布を図 4.5 に示す。このようにジャンルにより、クリティカリティの分布に違いがあることがわかる。

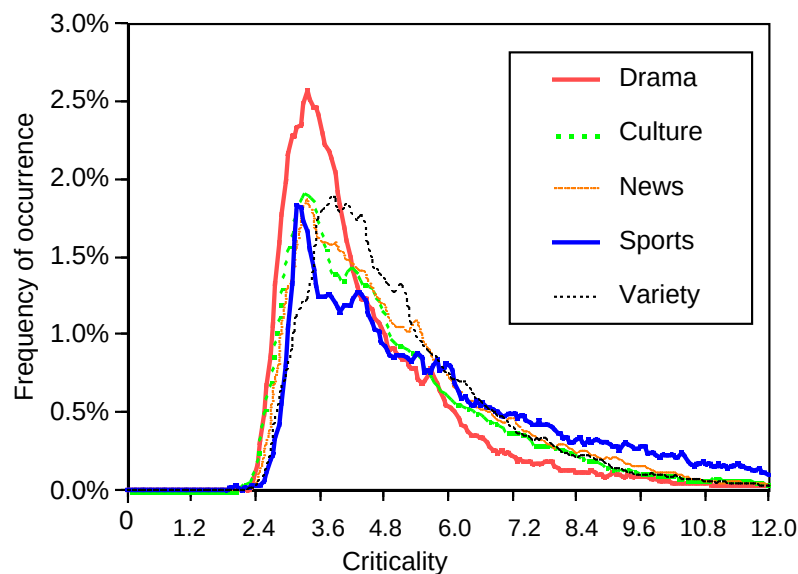


図 4.5 番組のジャンルによる分布の違い

各ジャンルにおけるクリティシティの平均値と標準偏差を表 4.1 に示す。

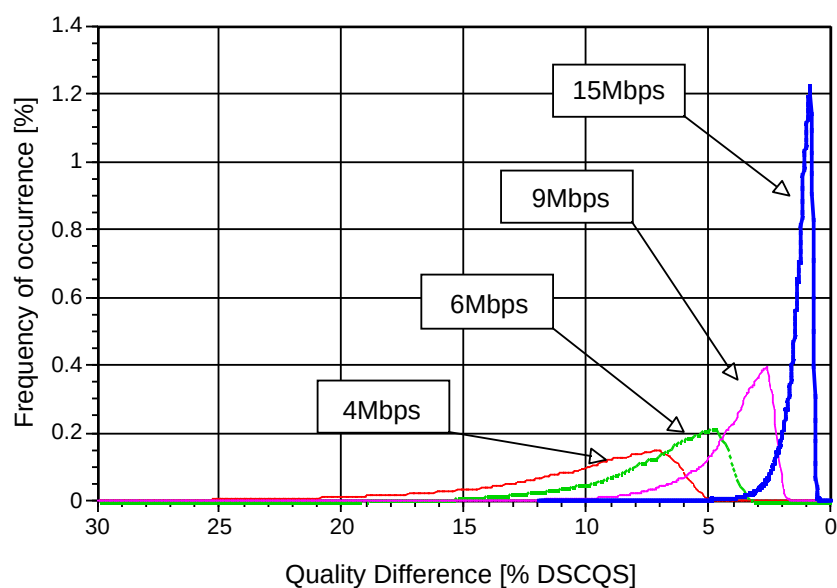
表 4.1 番組ジャンル毎の平均値および標準偏差

Category	Mean	Standard deviation
Drama	4.62	2.02
News	5.29	2.29
Variety	5.33	2.18
Culture	5.4	3.07
Sports	6.16	3.05
Overall	5.28	2.46

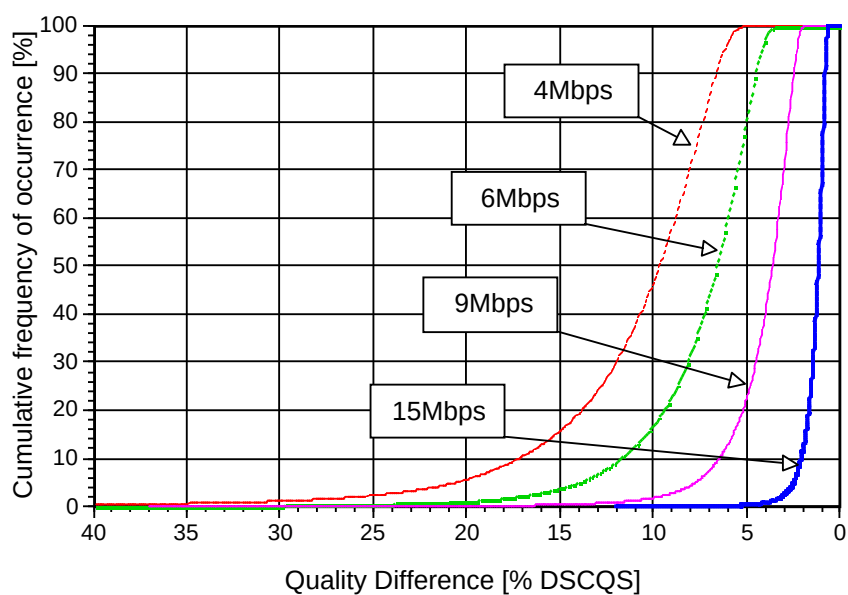
クリティシティの平均値はドラマでは低く、スポーツで高いことがわかる。これは、一般的にドラマでは被写体やカメラの動きが比較的少ないのに対して、スポーツでは激しい動きのシーンが多いためと考えられる。標準偏差はスポーツと教養番組(ドキュメンタリ番組を含む)で大きく、放送番組の中に多様な映像が含まれていると考えられる。また、ニュースや教養番組では、クリティシティの値が 10 以上で符号化が難しく、劣化が顕著となりやすいシーンが存在することが確かめられた。このように符号化の難しい絵柄は発生頻度こそ全体から見ると少ないが、ビットレートに強く依存して画質が変わるため、放送組の画質を考えるうえで、重要な要素となるであろう。

4.4 時間率から見た放送番組の画質

第3章で求めたクリティシティと画質の関係と、4.3 で示した放送番組のクリティシティの頻度分布から、図 4.1 の手順に従って放送番組における画質劣化を時間率で表現した結果を図 4.6 に示す。ここに、頻度計算は 0.01% DSCQS 毎とした。本手法により、各種ビットレートにおける品質分布を、実際の放送番組データを用いて定量的に示した。



(a) 発生頻度



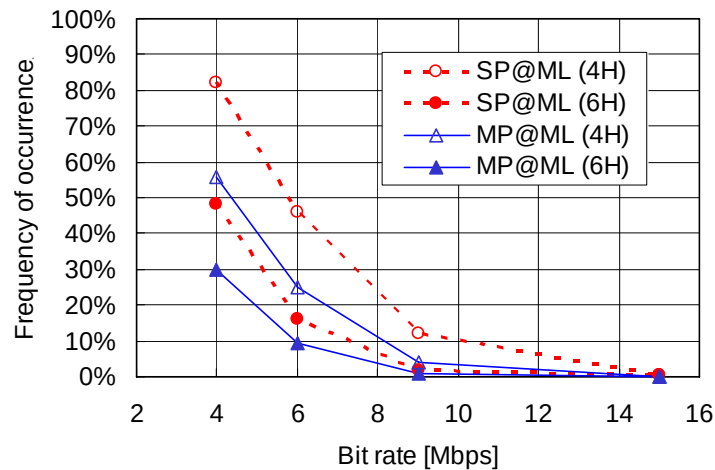
(c) 累積頻度

図 4.6 各種ビットレートにおける画質の分布 (MP@ML at 6H)

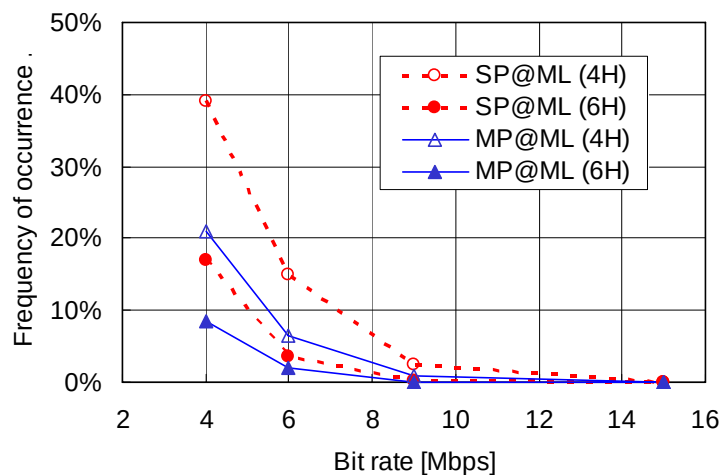
4.4.1 シンプルプロファイル(SP@ML)とメインプロファイル(MP@ML)の画質比較

SP@ML、MP@ML について、DSCQS 12%以内および 18%以内となる時間率を求めた結果を図 4.7 に示す。画質は視距離によって異なるが、SP@ML は MP@ML よりも視距離の影響を受けやすい。メインプロファイルは時間率からも良好な画質を示しているが、プロファイルの違いによる画質差以上に、視距離の違いによる差が大きいことが確認された。

視距離 6H の場合の要求条件 (DSCQS 12%以内)を満たす時間率を表 4.2 に示す。このように、ビットレートが低くなると、要求条件を満たす時間率は急激に減少する。99%以上の時間率で要求画質を維持するには、メインプロファイルで 9Mbps 以上のビットレートが必要であることがわかる。



(a) 12%劣化の時間率



(b) 18%劣化の時間率

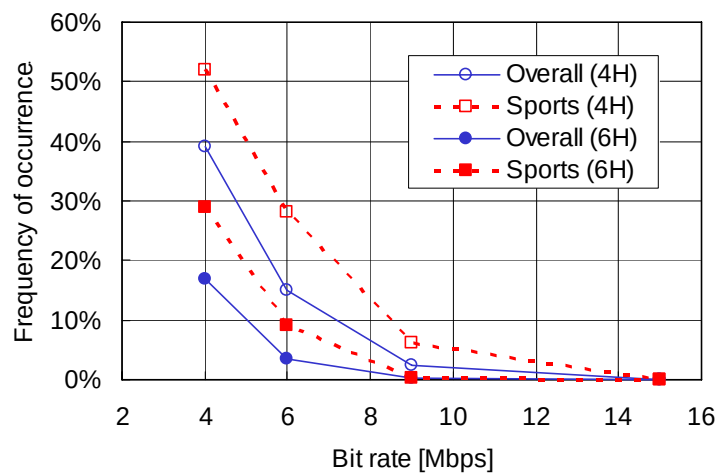
図 4.7 シンプルプロファイルとメインプロファイルの画質比較

表 4.2 画質劣化 12%以内の時間率、視距離 6H (95%信頼区間)

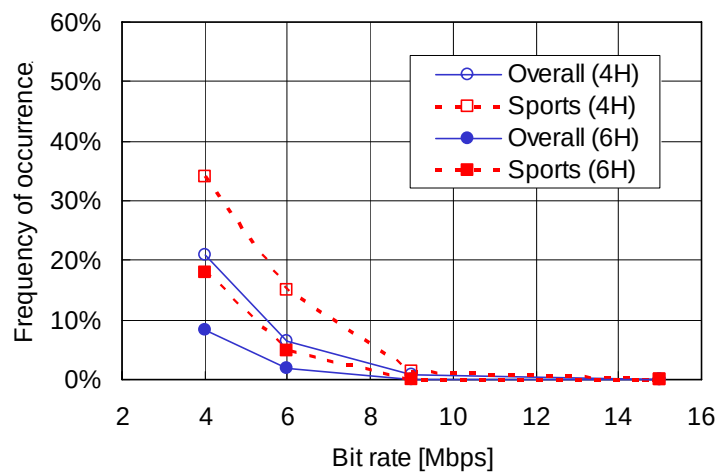
	4Mbps	6Mbps	9Mbps	15Mbps
MP@ML	70% (56~83%)	91% (84~96%)	99% (98~99.7%)	99.99%以上
SP@ML	52% (36~70%)	84% (70~94%)	98% (94~99.7%)	99.5% (99~99.9%)

4.4.2 番組ジャンルによる画質差

以上の結果は放送番組全体の平均値であるが、表 4.2 に示すようにスポーツ番組は符号化の難しい画像が出現する頻度が高い。スポーツ番組だけを取り出し、番組全体と比較した結果を図 4.8 に示す。また、視距離 6H における DSCQS 12%以上の劣化時間率の比較結果を表 4.4 に示す。



(a) SP@ML



(b) MP@ML

図 4.8 放送番組における DSCQS18%以内の画質劣化時間率

表 4.3 スポーツ番組における劣化 12%以内の時間率

(a) MP@ML、視距離 6H (95%信頼区間)

	4Mbps	6Mbps	9Mbps	15Mbps
Sports	56% (44～70%)	80% (71～89%)	98% (95～99.7%)	99.99%以上
Overall	70% (56～83%)	91% (84～96%)	99% (98～99.7%)	99.99%以上

(b) SP@ML、視距離 6H (95%信頼区間)

	4Mbps	6Mbps	9Mbps	15Mbps
Sports	42% (28～56%)	71% (55～86%)	95% (86～99.6%)	99.97% (99.5～99.99%以上)
Overall	52% (36～70%)	84% (70～94%)	98% (94～99.7%)	99.5% (98～99.9%)

ビットレートが低くなると番組ジャンルによる劣化時間率の相違は拡大する。表 4.3 の結果より、スポーツ番組において劣化が生じる時間率は、4～6Mbps で番組全体の平均よりも 10%程度高いことがわかる。すなわち、6Mbps 以下のビットレートでは、画質は放送番組の内容に比較的依存するが、9Mbps 以上ではジャンルによらずほぼ一定した品質を維持できる。

4.5 まとめ

デジタル放送システムを設計する上で重要なビットレートと画質の関係について、放送画質を時間率から評価する方法について提案した。デジタル放送の画質は、従来行われてきたような、数秒程度の極限られた数の画像を用いた主観評価テストの結果だけでは、十分に評価することはできない。ここでは、クリティカリティの概念を導入し、放送番組を長時間にわたり測定した結果から、デジタル放送の画質を時間率で解析した。その結果、番組のジャンルによる画質差などを定量的に評価することができ、その有効性が示された。

NHK 総合テレビのクリティカリティを1週間測定し、スポーツ番組では平均的にクリティカリティが高く、放送番組のジャンルによってクリティカリティ分布が異なることを確認した。すなわち、番組のジャンルによって画質劣化の程度は異なり、スポーツ番組では通常番組より検知限以下の良好な画質の時間率が減少する。特に低いビットレートでは、時間率の低下が顕著であった。

参考文献

- [1] 青木, 水野, 神田, 中須, 久保田, “MPEG-2 符号化におけるビットレートと画質に関する検討,” TV 学技報, BCS95-25 (May, 1995)
- [2] 中須, 青木, 西田, 久保田, “デジタル放送における時間率から見た画質の検討,” TV 学技報, BCS95-26 (May, 1995)
- [3] E. Nakasu et. al.: “Statistical Analysis of Picture Quality for Digital Broadcasting for Digital TV Employing MPEG-2 Video Standard,” 137th SMPTE Technical Conference, New Orleans (6-9 Sep. 1995)
- [4] E. Nakasu, K. Aoki, R. Yajima, Y. Kanatsugu, K. Kubota, “A Statistical Analysis of MPEG-2 Picture Quality for Television Broadcasting,” SMPTE Journal, 105, 11, pp.702-711 (Nov.1996)
- [5] I. Yuyama, Y. Nishida, E. Nakasu, “Objective Measurement of Compressed Digital Television Picture Quality,” SMPTE Journal, pp.348-352 (June 1998)

第5章 異種符号化映像フォーマットの総合画質評価

5.1 はじめに

デジタル放送は、時間帯によって HDTV と標準テレビを切り替えて放送するなど、柔軟な番組編成が可能で、放送事業者は映像素材や番組編成などに応じて送出する映像フォーマットを変更することができる。したがって、想定するビットレートで最良の画質が得られる映像フォーマットを選択する基準は大変重要である。また、平面ディスプレイの普及に伴い、デジタル放送の映像をテレビ受像機のディスプレイに応じて解像度やアスペクト比、走査方式などを変換して表示することが通常となっている[1][2][3]。このように、デジタル放送は同一のディスプレイで異なるフォーマットの映像を視聴する形態であり、その画質は符号化劣化だけでなく、解像感や走査方式等の違いによる品質差も含めて総合的に評価する必要がある。そこで本章では、放送局のエンコーダおよび送出フォーマットの指針を得るため、異なる映像フォーマットの符号化画像の品質を単一ディスプレイにより同一条件で評価を行い、映像フォーマットと総合画質の関係について考察する。

5.2 デジタル放送における総合画質の評価

5.2.1 デジタル放送の画質と評価の課題

2章で述べたように、デジタル放送映像の評価は DSCQS 法あるいは DSIS 法が主に用いられるが、それぞれ映像フォーマットごとに符号化劣化について評価し、異なるフォーマットの符号化画像の画質を直接比較してはいない。

これまで、異種映像フォーマットの画像を主観的に評価する方法として、各フォーマットのディスプレイをそれぞれ使用して基準画像と評価画像を同時に表示して評価する方法が提案されている[4]。この方法は、画質評価に二台のディスプレイを使用するため、各ディスプレイの特性を完全に合わせることが難しく、また、ディスプレイを並べて配置するため同じ視距離で同時に複数の評定者で評価することが難しいという課題があった。この他、1080p, 1080i, 720p の符号化画像の画質を相対的に比較した例があるが[5]、ここでは主観評価による直接的な画質比較は行われず、映像フォーマットごとの主観評価結果と符号化時の発生ビットレートを介して異種フォーマットの画質比較を試みている。このように、複数の異なる映像フォーマットの符号化画像を同一ディスプレイで同時に評価し、総合画質を論じた例はない。

本章では動画像の評価において信頼性の高い DSIS 法 Variant II[6]を異種フォーマットの画像の評価に適用し、マルチフォーマット対応のディスプレイとコーデックを使用してデジタル放送映像の総合画質の評価を行った。ここでは、同解像度の基準画像を用いた従来の評価と、より高解像度の基準画像を用いた評価の結果を比較することで、異種映像フォーマットの符号

化画像の品質評価の有効性を検証し、画質劣化要因を定量的に分析した結果を示す。また、DSIS 法を異種フォーマットに適用して、1080i、720p、および 480p 符号化画像を評価した結果を示すとともに、放送における時間率の視点から映像フォーマットと総合画質の関係について考察する。

制作映像フォーマットの高精細度化が進み、放送映像フォーマットの選択に自由度が高まることにより、ディスプレイや視聴環境を考慮したフォーマット選択がさらに重要となっている。デジタル放送では、サービスや伝送容量などに応じて、最も良い品質の得られる映像フォーマットに変換して符号化・伝送することが求められる。一般に、画面の画素数や毎秒フレーム数が少ないほど低ビットレート時の符号化劣化は抑えられるが、解像感の低下や動きの不連続性を伴うため、必ずしも総合的な画質向上が得られるとは限らない。したがって、放送サービスを行ううえで想定する伝送路において最も良好な画質となる符号化映像フォーマットを選択する指針が求められる。

一方、テレビディスプレイは、従来の CRT に代わり FPD (Flat Panel Display) が急速に普及し、画面の大型化、高精細化が進み、放送画質に対する要求もより厳しくなっている。FPD は液晶や PDP、有機 EL、SED など多様な方式があり、それぞれ特有の駆動方法と信号処理が用いられている。このため、CRT と比較して個々の方式や製品間における表示特性の違いが大きく、表示モードやパラメータの設定により画質も大きく影響を受ける。このような問題に対し、ITU-R など FPD を用いた画質評価法についての検討が開始されているが、表示デバイスや様々な信号処理による画質への影響はまだ十分に調べられておらず、現状では CRT を用いた評価結果を基準とせざるを得ない。

5.2.2 評価シーケンスとクリティシティ

本章の評価実験で用いた動画像を図 5.1 に示す。いずれも ITU-R 勧告 BT.1210-2[7]に記載されている HDTV 標準動画像[8]である。図 5.1 に示す評価シーケンスのクリティシティの測定結果を表 5.1 に示す。クリティシティの大きい画像ほど符号化が難しいことを示す。“Cognac and Fruit” (以下、“Cognac”)は、カメラや被写体の動きが小さく、比較的符号化が容易な画像である。一方、“Whale Show”, “Green Leaves”, “Sprinkling”は、水面や観客席、木立、水しぶきなど細かい絵柄が画面内の多くの面積を占め、物体の動きが複雑なため、符号化が難しいクリティカルな画像である。特に、“Sprinkling”は背景と人物をクロマキー合成した画像で、背景の水しぶきがランダムな動きで細かいテクスチャであり、高いエントロピーを有する影響で前景の女性の顔や着衣が著しく劣化しやすい、クリティシティの顕著に高い画像である。

図 5.2 は、1999 年 12 月 9 日から 2000 年 1 月 26 日までの NHK の HDTV 番組(トータル 550 時間)のクリティカリティ測定値から算出した時間率および累積時間率である。q_scale は 16 で測定を行った。測定した放送番組全体のクリティカリティの平均値は 3.84、最頻値は 2.72 であり、放送番組の 99%はクリティカリティ 10.2 以下の画像であった。各評価シーケンスのクリティカリティに相当する画像の累積時間率を表 5.1 に示す。ほとんどの番組素材は“Cognac”のように比較的符号化のやさしい画像で、今回の評価実験では一般の放送番組と比較してクリティカリティの高い画像を主に使用している。特に、“Sprinkling”は、クロマキー合成して生成した特殊なシーケンスであり、放送番組全体のなかでも非常に稀な画像といえることができる。

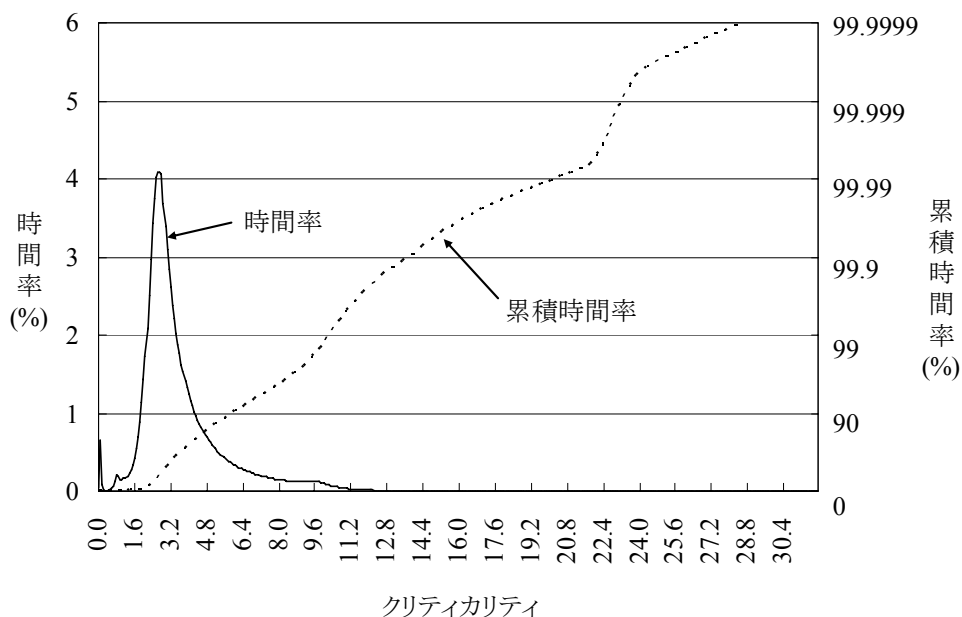


図 5.2 HDTV 放送番組のクリティカリティ

5.2.3 解像度の異なる基準画像を用いた符号化画像の評価

ここでは、異種フォーマットの符号化画像の評価に先立ち、従来のように同解像度の基準画像を用いた評価と、より高解像度の基準画像を用いた評価の結果を予備実験により比較した。さらに、両者の評価結果から、符号化による劣化とフォーマット変換による解像感の低下と画質の関係を考察した。画質評価法は DSIS 法と同じく二重刺激法を用い、基準画像の提示後、評価画像を提示して評価した。DSIS 法では非専門家を評定者とするが、本予備実験では専門家 4 名を用いた。また、専門家による評価であることを考慮して、シーケンスごとに符号化ビットレートを順に下げて提示し、各提示時間は評定者が採点を終わるまでとした。評価尺度は、5 段階劣化尺度(「5:わからない」「4:わかるが気にならない」「3: 気になるが邪魔にならない」「2:邪魔になる」「1:非常に邪魔になる」)を用い、採点は 0.1 精度で行った。

本評価実験では、図 5.1 の画像のうち、“Cognac”, “Whale Show”, “Green Leaves”, “Sprinkling”の 4 種各 10 秒を用いた。これらの HDTV 標準動画を 480p に方式変換後、MPEG-2 符号化装置で符号化したものを評価画像とし、基準画像は 1080i 原画像および 480p 原画像(1080i から変換)とした。表示は 1080i および 480p の信号をそのまま表示可能な 32 インチのマルチスキャンスタジオモニタを使用し、画面高の 3 倍の視距離で評価を行った。実験に使用した符号化装置、方式変換器およびモニタの詳細は 5.3 章で述べる。図 5.3 に 480p および 1080i の原画像を基準画像として評価した結果を示す。評点は評定者 4 人の平均値である。評価画像と同じフォーマットである 480p 原画像を基準画像とする場合、符号化による劣化程度を示し、クリティカリティの大きさに応じた評価結果となっている。

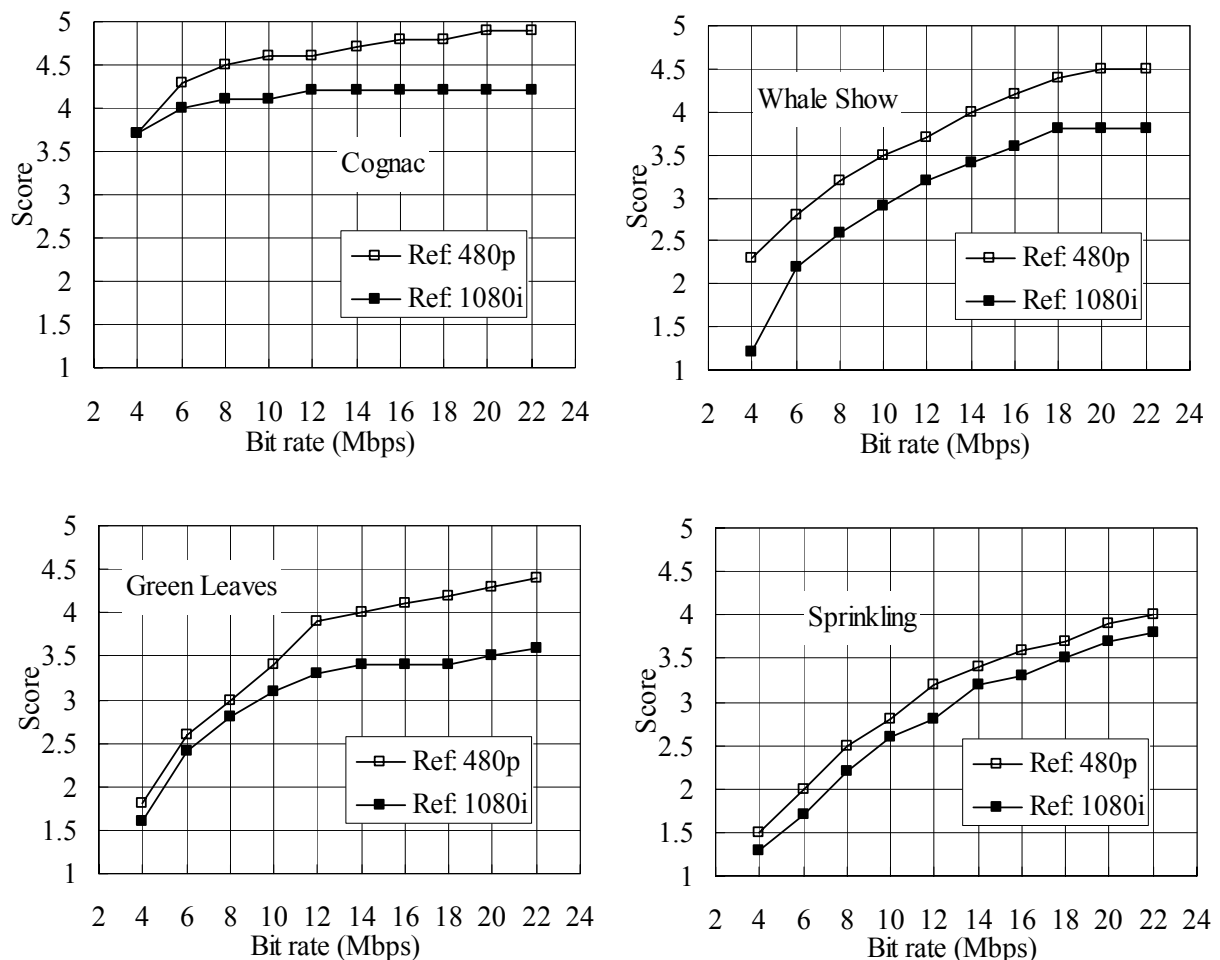


図5.3 主観評価結果(480p)

一方、1080i 原画像を基準とする評価結果は符号化劣化に加え、1080i との解像感の違いも含んだものとなり、評価シーケンス間の画質差は 480p を基準とした場合より小さい。また、1080i 原画像を基準とした場合、評価値が低下し始めるビットレートは絵柄によって異なる。“Cognac”、“Green Leaves”では、画質が低下し始めるビットレートは他より低く、ビットレートが低くなると基準画像の違いによる評点の差は少なくなっている。一方、“Whale Show”と“Sprinkling”では、ビットレートによらず、1080i と 480p の基準画像に対する画質差はほぼ一定となっている。

図 5.4 および図 5.5 は、図 5.3 の評価結果から 480p 原画像を基準としたときの原画像との評点との差を「符号化劣化」、1080i 原画像を基準としたときの評点と 480p 原画像を基準としたときの評点の差を「解像度低下」とみなしてグラフ化した結果である。

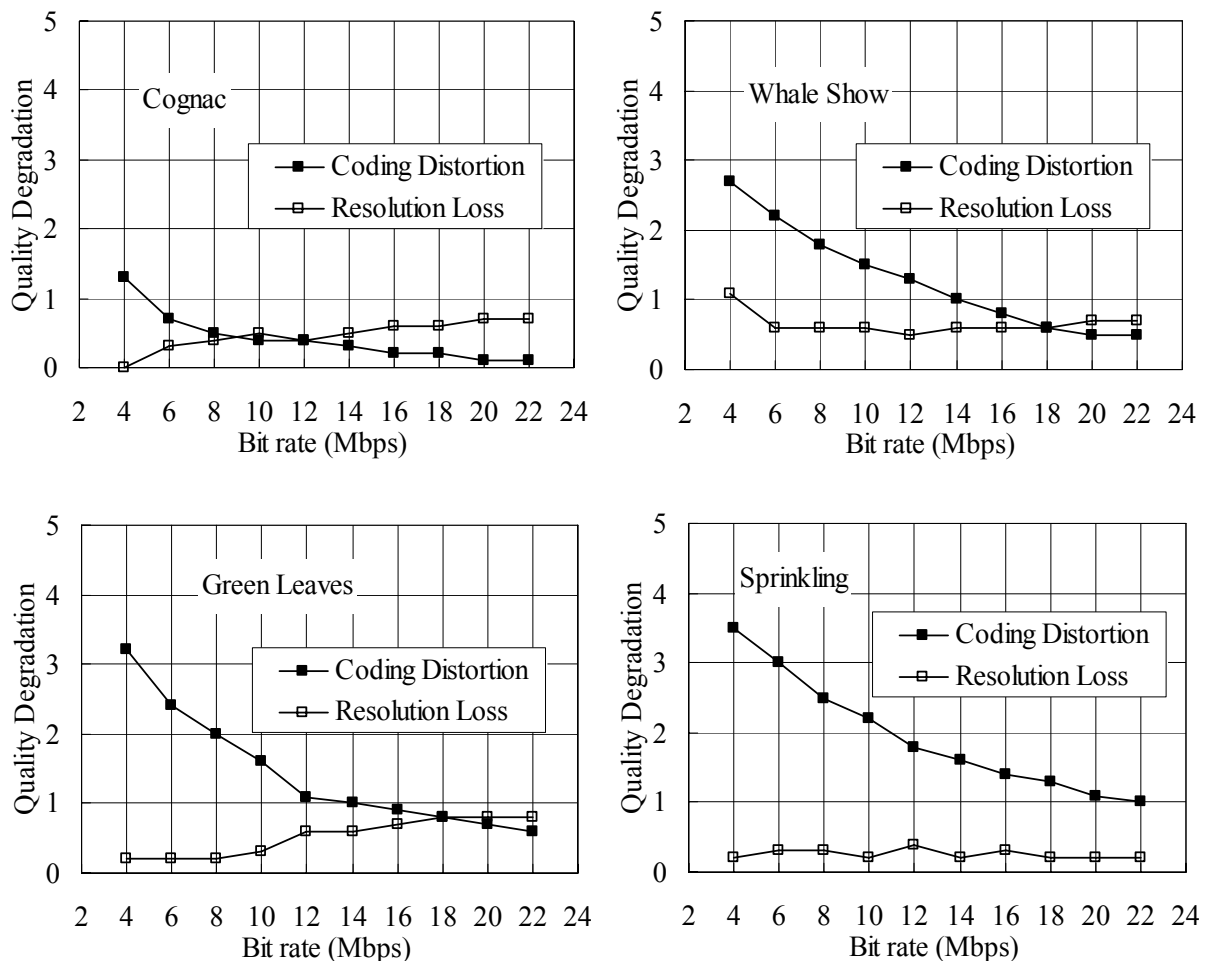


図 5.4 符号化劣化と解像度低下による画質劣化への寄与程度の比較

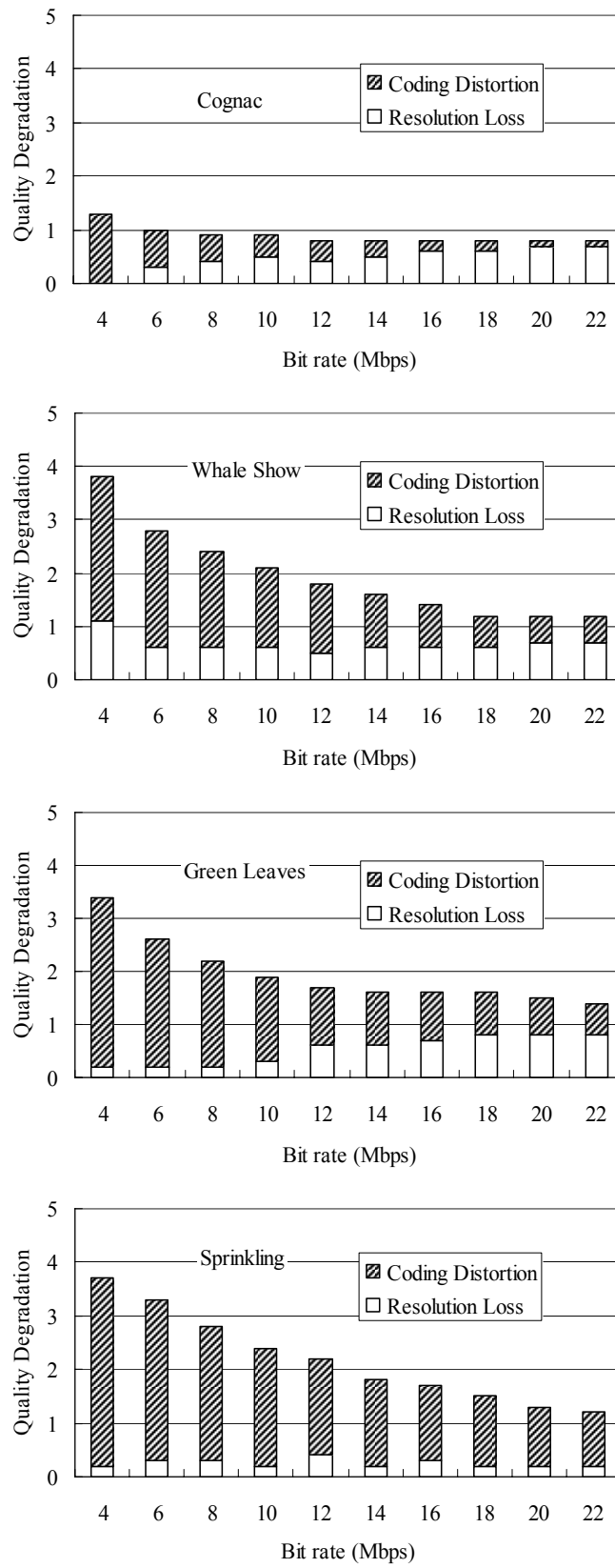


図 5.5 画質劣化の要因

最もクリティカルな“Sprinkling”は、解像感よりも符号化劣化が画質劣化の主要因であり、高精細で比較的動きの少ない“Cognac”では 14Mbps 以上で解像感の寄与が大きいことがわかる。“Whale Show”と“Green Leaves”はその中間的な性質を示し、16Mbps 以下で符号化劣化が主な画質劣化要因となっている。“Whale Show”と“Sprinkling”はビットレートが低くなっても解像感の低下があまり見られないのに対し、“Green Leaves”と“Cognac”は解像感が低下する。これは、高周波成分が量子化により低減するとともに、解像感に関する映像成分が符号化劣化にマスクされたためと考えられる。“Whale Show”は一定の解像感を保ったまま符号化劣化が増大しているが、これは動きに伴う劣化が支配的であったためとみなせる。

このように、より高解像度の基準画像を用いた評価結果を解像感の違いと符号化による劣化を含んだ総合画質とみなし、同解像度の基準画像を用いた評価結果と比較することで、それぞれの要素の総合画質への影響を定量的に分析する方法を示した。本手法は、より詳細なエンコーダの特性評価あるいは評価シーケンスの性質の解析に有効であると考えられる。

5.3 異種映像フォーマットの主観画質評価

本章では、評価画像の中で最も解像度の高いフォーマットを基準とし、DSIS 法 Variant II を異種映像フォーマットの評価に適用して解像度の異なる符号化画像の総合画質を評価した。この評価結果に基づき、各種ビットレートにおいて最適な符号化映像フォーマットを考察した結果について述べる。

5.3.1 評価実験の概要と使用機器

成田らは、DSCQS 法と DSIS 法の比較を行い、2 回提示による DSIS 法 Variant II が動画や極軽微な劣化に対し、評定者の負担が少なく信頼性が高いことを示した[9]。DSCQS 法では、基準画像と評価画像の提示順をランダムに変え、基準画像を明示しないため、特に非専門家による評価では、基準画像と評価画像の区別が付きにくい極軽微な劣化に対して両者の差をあまり付けられない傾向が見られる。一方、DSIS 法は、最初に基準画像、続いて評価画像を提示することで評定者に基準画像を明示するため、劣化程度によらず安定した評価結果を得ることができる。今回のように基準画像と評価画像の映像フォーマットが異なる場合を含んだ評価では、フォーマットの違いによる評点への影響をできるだけ少なくするため、基準画像を明示する DSIS 法 Variant II を採用することとした。

表 5.2 に評価実験の概要を示す。評価画像のフォーマットは 1080i, 720p, 480p であり、基準画像はこのなかで最も空間解像度の高い 1080i 原画像を使用し、基準画像に対する劣化程度を 5 段階劣化尺度で評価した。評価画像は、絵柄と映像フォーマット、符号化ビットレートがそれぞれランダムになるように提示した。提示時間は図 5.6 に示すように、基準画像と評価画

像を交互に2回ずつ各10秒で提示する。ここで、ITU-R 勧告 BT.500 で規定する DSIS 法は、提示する画像と画像の間はグレイ信号を挿入するが、今回の評価実験では映像フォーマットの切替時にショックが入って評定者にフォーマットの違いを気付かせることがないように、画像間はブラック信号を挿入した。また、各セッションの最初3画像はダミーとし、評価結果の解析から除外した。クリティカルな画像の画質劣化程度を考慮し、アンカー画像は用いていない。

表 5.2 画質評価実験の概要

項 目	内 容
評価方法	DSIS 法 Variant II (5 段階劣化尺度)を異種映像フォーマットの評価に適用
視距離	3H (H:画面高)
評価シーケンス	HDTV 標準動画像 (ITU-R 勧告 BT.1210-2) 6 種: “Cognac and Fruit”, “Walk through the Square”, “Bronze with credits”, “Whale Show”, “Green Leaves”, “Sprinkling” 各 10 秒
基準画像	1080i 原画像
評価画像	1080i, 720p, 480p 符号化画像
評定者	非専門家 24 名
観視条件	ITU-R 勧告 BT.500-10 および BT.710-4 準拠(実験室環境および家庭環境)

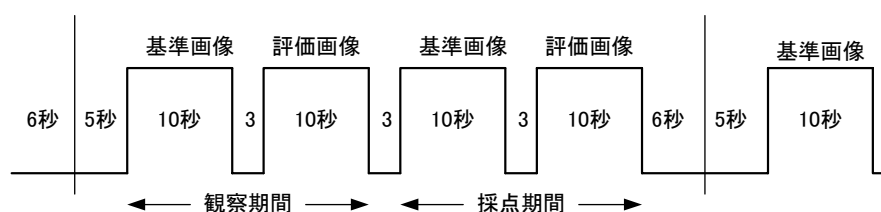


図 5.6 本評価実験における提示時間

評価シーケンスは、図 5.1 に示す 6 種の画像を用いた。評定者は非専門家 24 名で、6 名ずつ 4 グループに分かれて評価した。評価実験では 2 台のモニタを使用し、1 台のモニタに各 3 名の評定者が同時に評価を行った。観視条件は ITU-R 勧告 BT.710[10]準拠とし、視距離は評価映像フォーマットに関わらず 3H (H:画面高)とした。ITU-R 勧告で規定する視距離はフォーマットによって異なるが、実際のテレビ視聴では異なる解像度の放送番組であっても視距離を変えることはないため、本実験ではいずれのフォーマットも同じ視距離とした。

今回の評価実験ではスタジオモニタに加え、マルチフォーマット対応の民生用受像機についても評価を行った。民生用受像機の観視条件は、ITU-R 勧告 BT.500 で規定する家庭での視聴環境相当とし、室内の蛍光灯を点灯して背景輝度を約 30～50cd/m²とした。

図 5.7 に評価実験の系統図を示す。ただし、基準画像および各映像フォーマットの復号画像は予め VTR に収録しておき、評価時に各 VTR を同期運転して再生し、評価者が提示映像の切替を行った。実験に用いた主な機器の仕様を表 5.3 に示す。

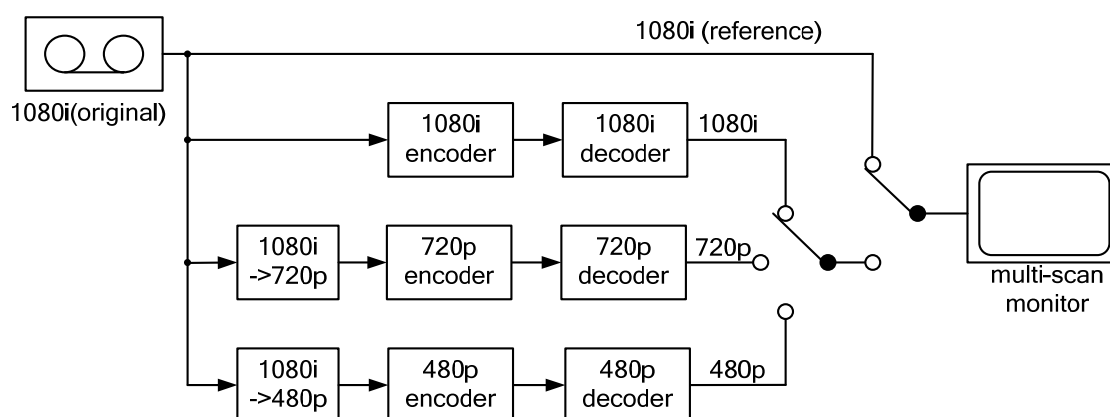


図5.7 評価実験系統図

表 5.3 実験に使用した主な機器の仕様

項 目	内 容
符号化映像フォーマット	1080i (水平 1920 画素および 1440 画素), 720p, 480p
方式変換器	AJ-UFC1800 (松下電器製) 1080i から 720p, 480p への変換
符号化装置	エンコーダ: MV-400 (JVC 製) MPEG-2 Main Profile 準拠 1080i: M=3, N=15, 720p, 480p: M=3, N=30 デコーダ: DM-D4600 (JVC 製) 1080i, 720p, 480p に対応
記録・再生用 VTR	1080i, 480p: D6VTR GBR1000 (東芝製) 非圧縮 720p: D5VTR AJ-HD3000 (松下電器製) 1/4 フレーム内圧縮 (映像 230Mbps)
ディスプレイ	32 インチマルチスキャン CRT ・スタジオモニタ: BVME32D1WU (ソニー製) ・民生用受像機: KV-32DZ950 (ソニー製)

映像素材は、1080i から各フォーマットへ同一の方式変換装置を用いて変換し、機器間の入出力はデジタル I/F で行った。方式変換後の 720p および 480p の解像度（モノスコチャートを用いた目視による実測値）は表 5.4 の通りである。符号化は、実際のデジタル放送で使用されている機器と同等の性能を有するマルチフォーマット対応の MPEG-2 ハードウェアエンコーダを用いた。評価を行ったビットレートを表 5.5 に示す。ハードウェアの制約で必ずしも同一ビットレートとはなっていないが、各フォーマットの画質が同程度となると予想される 12～16Mbps を中心に評価した。デコーダは、全フォーマットに対応する機器を使用した。

表 5.4 方式変換後の解像度（実測値）

	720p	480p
水平解像度	720 TV 本	400 TV 本
垂直解像度	600 TV 本	480 TV 本

表 5.5 評価ビットレート

映像フォーマット	ビットレート
1080i (1920)	12, 14, 16, 20 Mbps
1080i (1440)	10, 12, 14, 16 Mbps
720p	9, 12, 14, 16 Mbps
480p	5, 7, 9, 12, 14 Mbps

ディスプレイは、32 インチ CRT のスタジオモニタと民生用受像機を各 2 台ずつ用いた。画素ピッチ（中央）は、スタジオモニタは 0.32mm、民生用受像機は 0.42mm である。スタジオモニタのアパーチャ補正は off とし、民生用受像機の表示モードは、工場出荷時の設定「スタンダード」とした。また、白ピーク輝度（測定値）はスタジオモニタで 70cd/m²、民生用受像機 2 台はそれぞれ 373, 375cd/m² である。スタジオモニタの色温度は D65 (x:0.3127, y:0.3290)、民生用受像機 2 台の(x,y)色度値は(0.270, 0.273)、(0.272, 0.272)であった。図 5.8 に使用したディスプレイの MTF (Modulation Transfer Function)（実測値）を示す。民生用受像機では、エンハンサのため、約 5～10MHz の範囲で強調された特性を示している。

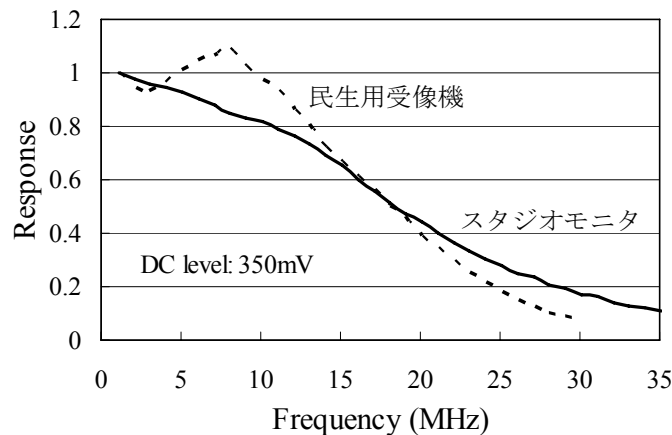


図 5.8 使用したディスプレイの MTF 実測値

5.3.2 評価結果および映像フォーマット間の画質比較

スタジオモニタでの評価結果を図 5.9 に、民生用受像機での評価結果を図 5.10 に示す。それぞれ評定者の平均評点と 95%信頼区間を示す。ITU-R 勧告 BT.500 に基づき評定者の検定を行ったが、今回の実験で除外される評定者はなかった。符号化を行っていない原画像の画質は、ほぼ 1080i, 720p, 480p の順であるが、スタジオモニタでは“Green Leaves”を除き、有意な差は認められなかった。

12~16Mbps のレートにおける画質をフォーマット間で比較した場合、“Walk through the Square” (以下、“Square”) と“Green Leaves”では、ほとんど有意な差はなかった。“Cognac”は、5.2 章に示すように符号化による劣化が少なく、スタジオモニタ、民生用受像機とも原画像の解像度の差が 12~16Mbps での画質差に反映されている。文字スーパーを含む“Bronze with credits” (以下、“Bronze”) では、1080i の画質と 720p、480p の画質との間で有意な差が認められた。一般に文字スーパーは輪郭部に高い周波数成分を有するため符号化が難しく、文字の周りにモスキートノイズが生じやすい。同一サイズのディスプレイで異なる解像度の映像を表示する条件では、解像度の低いフォーマットのほうがモニタ上の表示ブロックサイズが相対的に大きくなり、モスキートノイズが広がって目立つ傾向がある。このため、解像度の高い 1080i が画質的に有利であったと考えられる。一方、“Whale Show”では主にカメラの水平パンニングのシーンで 1080i(1920)の符号化劣化が大きい。クリティシティが非常に高く、5.2 章の結果で示されるように符号化歪による画質劣化が顕著な“Sprinkling”では、解像度の高いフォーマットほど低ビットレートにおける劣化程度が大きい。このシーケンスでは 480p の評価を行った 14Mbps 以下で 480p の画質が最も良く、クリティシティが非常に高い画像に対して 480p 符号化が有効であることが示された。

民生用受像機においては、図 5.8 に示すように表示可能な最高映像周波数が低いにもかかわらず、スタジオモニタの評点とほぼ同様な傾向を示している。ただし、民生用受像機は輝度が高く、エンハンサにより符号化歪を強調して表示するため、およそ 0.5 ポイント程度スタジオモニタより低い評点となっている。特に“Whale Show”では、1080i(1920)がスタジオモニタの場合と比較して 1 ポイント程度画質劣化が大きくなっている。このシーケンスでは、パンニングのシーンで 1080i 符号化の色劣化が顕著であり、特に民生用受像機では色彩が強調されて表示されるため、他の映像フォーマットよりも評点が低かったと考えられる。しかし、水平 1440 画素で符号化することにより 1080i は他フォーマットとほぼ同程度の画質となっている。このように、12～16Mbps のビットレートではスタジオモニタ、民生用受像機とも 1080i を水平画素 1440 で符号化する効果が確認された。

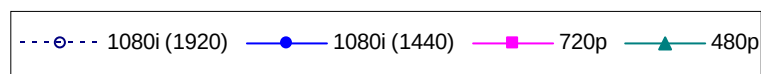
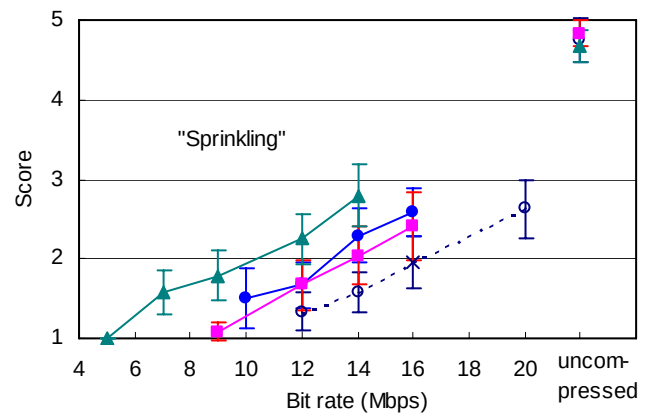
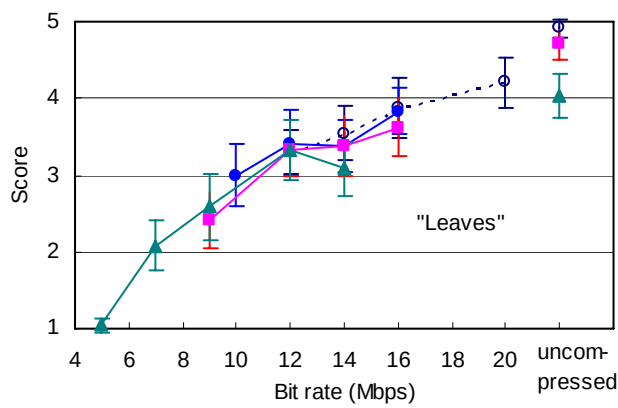
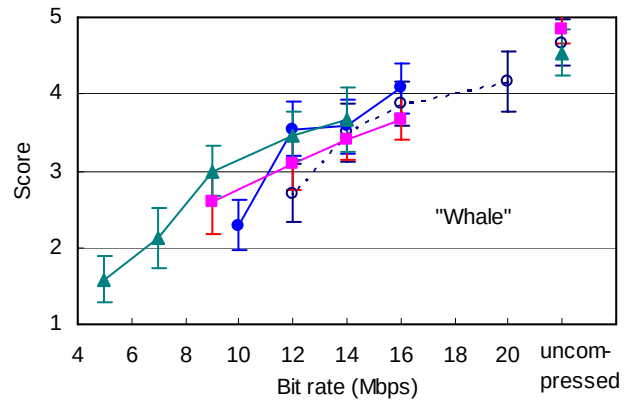
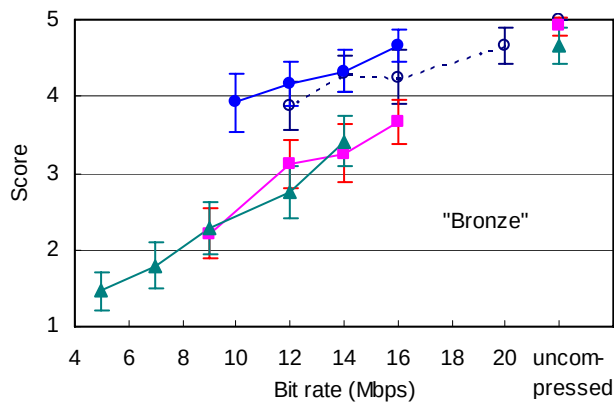
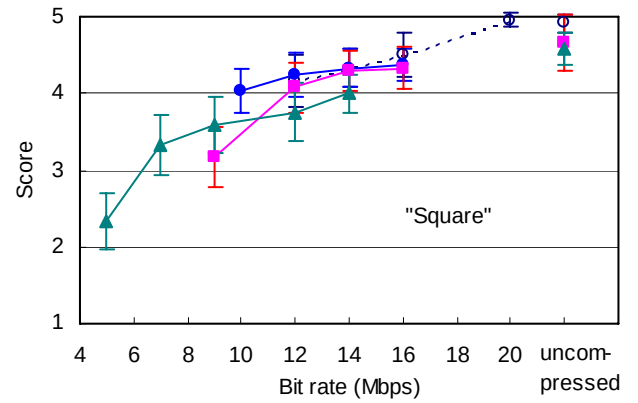
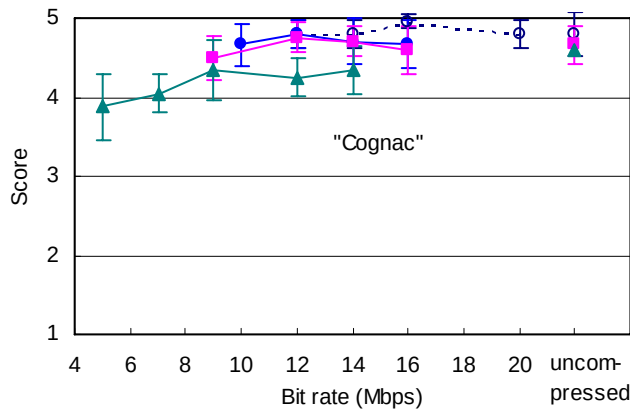


図 5.9 ビットレートと画質の関係 (スタジオモニタ)

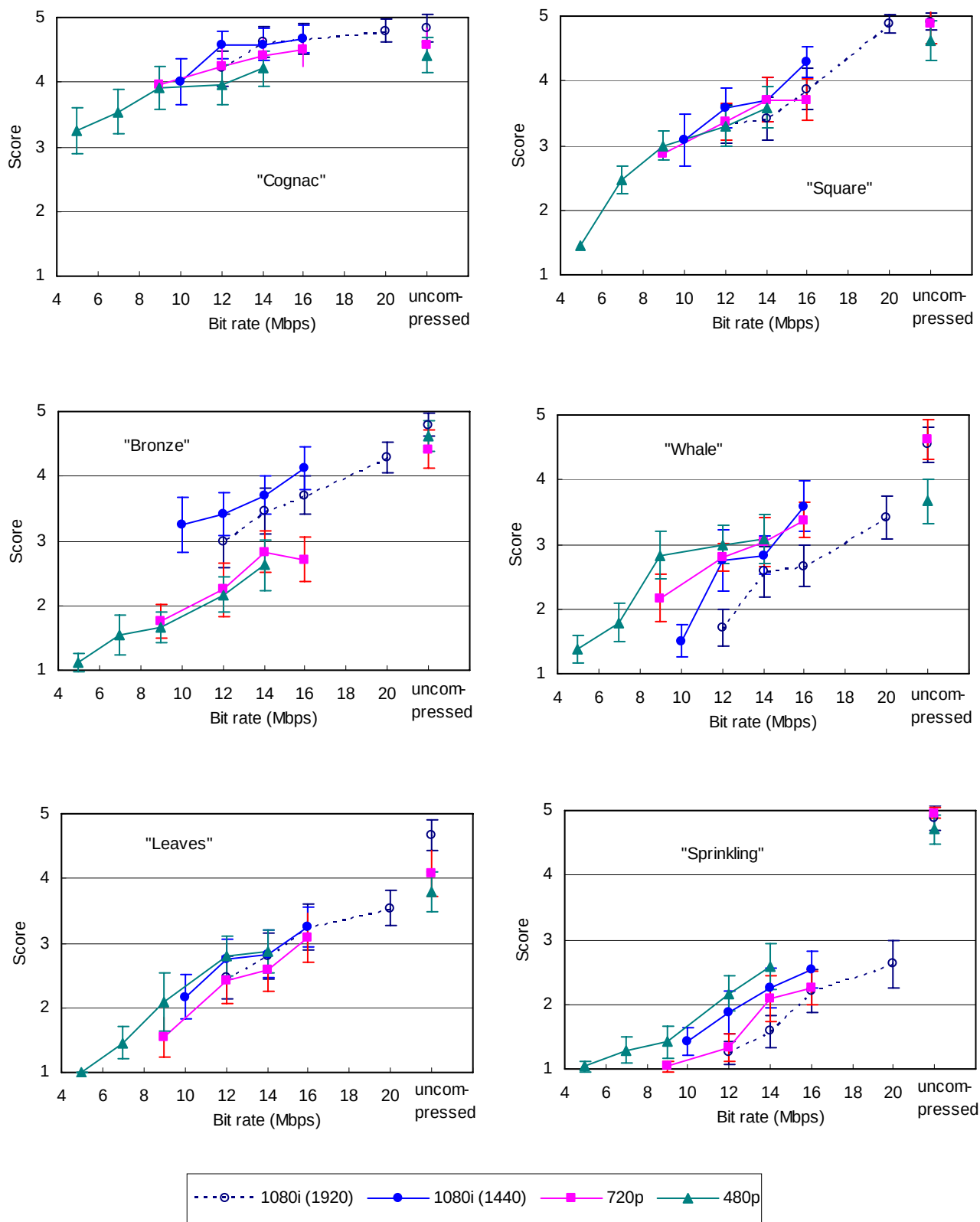


図 5.10 ビットレートと画質の関係(民生用受像機)

5.3.3 放送における映像フォーマット選択に関する考察

ここでは、評価画像のクリティカリティの値に基づいて、放送番組における時間率の視点から各ビットレートに適した映像フォーマットを考察する。12Mbps および 14Mbps において、スタジオモニターで平均評点が最も高い映像フォーマットを表 5.6 に示す。12Mbps では、1080i(1440) 符号化により放送番組の 99.99%の時間率で最も良い画質が得られた。14Mbps では、放送番組の 90.5%に相当するシーケンスで 1080i(1920)が良く、1440 画素に変換して符号化することにより 98%の時間率で他より良い画質となる。16Mbps についても同様の結果であり、非常にクリティカルな画像については 480p の評価結果がないため優劣は判断できないが、99.99%の時間率で 1080i の画質が良いと認められる。9Mbps においては 720p と 480p の評価を行ったが、“Cognac”を除き 480p が良い。1080i(1440)は 10Mbps で符号化を行ったため、同一ビットレートで比較はできないが、図 5.9 の結果から 480p および 720p の 10Mbps における画質を 9Mbps および 12Mbps の評点から線形内挿により推定して比較を行うと、放送番組の 98%の時間率で 1080i(1440)の画質が良いと考察される。480p はクリティカリティが 1 以上のシーケンスに対して幅広いビットレートで相対的に良い画質を示しており、今回の評価結果では 720p の優位性は認められなかった。民生用モニターについても同様な傾向であるが、“Green Leaves”については 14Mbps で 480p の優位性が高まっている。

表 5.6 各評価画像において最も高い評点を与える映像フォーマット(スタジオモニター)

評価画像	クリティカリティ	放送番組での 累積時間率 (HDTV 550時間)	ビットレート			
			10Mbps 注 1)	12 Mbps	14 Mbps	16 Mbps 注 2)
Cognac	3.52	65.1%	1080i(1440)	1080i(1920/1440)	1080i(1920)	1080i(1920)
Square	6.08	90.5%	1080i(1440)	1080i(1440)	1080i(1920/1440)	1080i(1920)
Bronze	9.44	98.0%	1080i(1440)	1080i(1440)	1080i(1440)	1080i(1440)
Whale Show	16.0	99.96%	480p	1080i(1440)	480p	1080i(1440)
Green Leaves	19.8	99.99%	1080i(1440)	1080i(1440)	1080i(1920)	1080i(1920)
Sprinkling	26.7	99.9998%	480p	480p	480p	1080i(1440)

注1) 10Mbps においては 1080i(1920)の評価は行っていない。480p, 720p については
9Mbps と 12Mbps の評価値の値から線形内挿した 10Mbps の値で比較。

注2) 16Mbps については 480p の評価は行っていない。

5.4 まとめ

異種フォーマットの符号化画像品質を、単一ディスプレイを用いて同一条件で主観評価する手法を示し、デジタル放送の各種映像フォーマットの画質比較を行った。まず、より高解像度な原画像と同一解像度の原画像の2種類の基準画像を用いた二重刺激法による評価実験から、解像感の違いと符号化による劣化の総合画質への寄与程度を定量的に検証した。次に、DSIS 法 Variant II を異種映像フォーマットの評価に適用し、1080i 原画像を基準として 1080i, 720p, 480p 符号化画像の画質比較を行った。本評価手法では、符号化劣化に加えて解像感の違いを含む総合的な画質の比較が可能である。

本評価実験は、1080i を原画像として各種映像フォーマットに変換後、符号化画像の評価を行ったが、14Mbps 以上のビットレートで 1080i(水平 1920 画素)が放送番組の約 90%の時間率で他フォーマットより良い画質となることが認められた。また、水平画素数を 1440 にダウンサンプリングして符号化することで、12~16Mbps においてクリティシティが高い動画像に対して 1080i の画質改善が図れることが示された。この評価実験の結果から、水平画素の切替により 1080i が 12Mbps 以上のビットレートで他のフォーマットと比較し、良好な画質が得られるという指針が得られた。

参考文献

- [1] ARIB STD-B32, “デジタル放送における映像符号化, 音声符号化及び多重化方式,” 1.4 版 (2003)
- [2] R. Salmon, “The changing world of TV displays – CRTs challenged by flat-panel displays,” EBU Technical Review (April 2004)
- [3] 中須英輔, “デジタル放送とフラットパネルディスプレイ,” 2006 年映情学年次大, S3-6 (2006)
- [4] 成田長人, 杉浦幸雄, “異種放送システムの主観評価に関する考察,” 信学論 A, J81-A, 2, pp.269-279 (1998)
- [5] H. Hoffmann, T. Itagaki, D. Wood, A. Block, “Studies on the bit rate requirements for a HDTV format with 1920x1080 pixel resolution, progressive scanning at 50 Hz frame rate targeting large flat panel displays,” IEEE Trans. Broadcasting, 52, 4, pp. 420-434 (Dec. 2006)
- [6] Rec. ITU-R BT.500-11, “Methodology for the Subjective Assessment of the Quality of Television Pictures” (2002)
- [7] Rec. ITU-R BT. 1210-3, “Test materials to be used in subjective assessment” (2004)
- [8] 映メ学会監修, “ハイビジョン・システム評価用標準動画像 解説書” (1998)

- [9] 成田長人, “画像符号化の主観評価法に関する考察,” 信学論 B-I, J77-B-I, 2, pp.102-111 (1994)
- [10] Rec. ITU-R BT. 710-4, “Subjective Assessment Methods for Image Quality in High-Definition Television” (1998)

第6章 結論

本論文では、デジタル放送におけるサービス品質の基準と運用指針を明確にすることを目的に、画像のクリティカリティに基づいて放送番組の画質を統計的に解析してサービス品質を評価する手法を提案するとともに、各種映像フォーマットの総合画質を時間率から評価し、最適なビットレートや映像フォーマットを選択する指針を考察した。

本研究の概要を以下にまとめる。

第2章では、デジタル放送の画質評価における課題を明らかにし、画質評価に関するこれまでの研究を整理し、本研究の方向性を示した。第3章では、デジタル映像符号化における画質の絵柄依存性を定量的に表現するクリティカリティ(符号化難易度)の定義を提案し、実験により符号化画像の主観評価値との関係を確認し、有効性を検証した。第4章では、多様な絵柄の映像を扱うデジタル放送での画質変動の性質を考慮して、サービス品質を時間率で評価する手法を検討した。実際の放送番組のクリティカリティを長時間測定し、放送番組の画質分布を統計的に解析して所要ビットレートを検討するなど、本手法の有用性を明らかにした。第5章では、各種映像フォーマットを選択できるデジタル放送の特徴を考慮して、異なるテレビ方式(映像フォーマット)の符号化画像を統一的に評価し、画像の解像感等を含めた総合画質の評価により、最適な映像フォーマットの選択基準について考察した。

以上のように、本論文では、多様な映像を扱うデジタル放送における画質評価の課題を解決するため、絵柄依存性を考慮した統計的な画質解析手法、および異なる映像フォーマットの符号化画像に対する総合画質評価法を提案し、各種画像の評価実験と放送番組のクリティカリティ測定の結果に基づいてデジタル放送の運用に関する指針を示した。

本研究は、BS デジタル放送および地上デジタル放送で使用されている MPEG-2 映像符号化を対象として画質評価に関する検討を行った。H.264 | MPEG-4 AVC のデジタル放送への導入も進められており、これらの新たな符号化方式に対する適用性の確認も求められる。また、多様化するディスプレイや受信機における信号処理、各種視聴条件を考慮した画質評価法の確立も望まれる。今後、コンテンツや符号化・伝送技術、ディスプレイ・端末の多様化がさらに進むと予想されるが、そのようなデジタル映像メディアのサービス品質を統一的に評価する技術の開発が期待される。

謝辞

本論文の作成にあたり、ご指導、ご鞭撻を賜りました東京工業大学大学院総合理工学研究科 羽鳥好律教授に深く感謝いたします。また、東京工業大学大学院総合理工学研究科 内川恵二教授、小林隆夫教授、西田眞也教授、理工学研究科 八島由幸教授、金子寛彦准教授からは、数々の有益なご助言ならびにご示唆をいただきました。ここに、心から感謝します。

本論文は、著者が日本放送協会 放送技術研究所において、1995 年から 2006 年の間にを行った研究をまとめたものです。本論文をまとめる機会を与えて頂き、多大なるご支援を賜った放送技術研究所 谷岡健吉所長に深謝します。また、本研究を進めるにあたり、研究の機会を与えて頂いた泉武博元所長、西澤台次元所長、吉野武彦元所長（元技師長）、山田宰元所長（現パイオニア(株)）、三宅誠元所長（元理事、現東京情報大学）、榎並和雅前所長（現情報通信研究機構）に、深く感謝します。

久保田啓一元次長（現情報システム局局長）には、本研究をはじめめるにあたり、貴重なご助言をいただくとともに、その後も引き続き、ご指導、ご教示いただきました。心より感謝いたします。湯山一郎元次世代テレビ方式研究部部長（現宇都宮大学）、熊田純二元衛星デジタルシステム部長（現(株)NHK アイテック）、福井一夫元マルチメディアサービス部長（現神奈川工科大学）、下田茂元映像情報部長（現(株)NHK アイテック）、伊藤崇之（人間・情報）部長には、研究の機会を与えていただくとともに、ご指導いただきました。厚く感謝いたします。

また、クリティカリティの研究および画質評価実験の実施にあたり、貴重なご助言をいただいた野尻裕司研究主幹、金澤勝研究主幹、合志清一主任研究員、大塚吉道元主任研究員（現パイオニア(株)）、栗田泰市郎（材料・デバイス）部長、金次保明システム担当部長に深く感謝します。

本研究を進めるにあたり、西田幸博主任研究員、青木勝典主任研究員、神田菊文元研究員（現放送技術局）、水野修元研究員（現放送技術局）、黒住正顕元研究員（現放送技術局）、市ヶ谷敦郎研究員をはじめ、次世代テレビ方式研究部（当時）およびマルチメディアサービス（当時）のメンバーには、多大なる助言と協力をいただきました。心より感謝します。

放送番組のクリティカリティ測定に協力いただいた NHK 技術局技術開発センター（当時）、放送技術局番組送出センター（当時）の関係各位、ならびに主観評価テストを実施し、共に検討を行って頂いた多くの方々に感謝致します。

最後に、長年の研究生活において、常に支援・協力してくれた家族に心から感謝します。

研究業績

論文

- (1) E.Nakasu, K.Aoki, R. Yajima, Y.Kanatsugu, K.Kubota “A Statistical Analysis of MPEG-2 Picture Quality for Television Broadcasting”, SMPTE Journal, vol.105, no.11, pp.702-711 (Nov.1996) *SMPTE Journal Certificate 受賞 (1996)*
- (2) 中須英輔、金次保明、矢島亮一、青木勝典、久保田啓一「MPEG-2 映像符号化における画質の絵柄依存性の検討～画像のクリティカルティ～」映像情報メディア学会誌、vol.51, no.9, p.1568-1576 (1997)
- (3) I.Yuyama, Y.Nishida, E.Nakasu: “Objective Measurement of Compressed Digital Television Picture Quality”, SMPTE Journal, vol.107, no.6, pp.348-352 (June 1998)
- (4) 中須英輔、神田菊文、市ヶ谷敦郎、黒住正顕、西田幸博「異種符号化フォーマット画像の総合品質評価」映像情報メディア学会誌、Vol.62, No.2, pp.262-270 (2008)

関連論文

- (1) E. Nakasu, Y. Nishida, M. Maeda, M. Kanazawa, S. Yano, M. Sugawara, K. Mitani, K. Hamasaki, Y. Nojiri “Technical Development Toward Implementation of Ultra High-Definition TV System”, SMPTE Motion Imaging Journal, pp.279-286 (July/August 2007)

国際会議

- (1) E. Nakasu, K. Aoki, R. Yajima, Y. Kanatsugu, K. Kubota: “Statistical Analysis of Picture Quality for Digital TV Employing MPEG-2 Video Standard”, p.163-177, 137th SMPTE Technical Conference (1995)
- (2) E. Nakasu, Y. Nishida, M. Maeda, M. Kanazawa, S. Yano, M. Sugawara, K. Mitani, K. Hamasaki, and Y. Nojiri “Technical development towards implementation of extremely high resolution imagery system with more than 4000 scanning lines”, IBC2006, pp.345-352, Amsterdam (Sept. 2006)

解説論文

- (1) 中須英輔「画質評価 画質とユーザ要求」特集 MPEG、テレビ誌 vol.49, no.4, pp.497-501 (1995)
- (2) 中須英輔「テレビ画像のデータ圧縮とその画質評価」ディスプレイアンドイメージング、vol.5, no.3 (1997)

- (3) 中須英輔「デジタル映像の評価(2)」講座マルチメディアのための品質評価、映像メ
誌 vol.54, no.3,p.357-363 (2000)
- (4) 中須英輔「画質評価の今後の展開」小特集 デジタル動画の画質評価、映像メ
誌 vol.59, no.8,p.1154-1155 (2005)

国内学会・研究会

- (1) 中須英輔、青木勝典、西田幸博、久保田啓一「デジタル放送における MPEG2 映像符
号化方式の検討～ビットレートと画質の関係～」テレビジョン学会技術報告(放送方式研
究会, 放送現業研究会)、vol.19, no.26, BCS95-26, BFO95-38, p.19-24 (1995)
- (2) 中須英輔、小澤啓爾、日高恒義、浦野丈治、坂中靖志「MPEG-2 4:2:2 プロファイルの
概要とベリフィケーションテスト結果」テレビジョン学会技術報告(放送方式研究会, 無
線・光伝送研究会)、vol.20, no.35, BCS96-26, ROFT96-60, p.1-6 (1996)
- (3) 中須英輔、中田安優、浦野丈治、清水勉「デジタル放送の符号化画質評価実験」映像
情報メディア学会技術報告(放送方式研究会)、vol.22, no.19, BCS98-3, p.13-18 (1998)
- (4) 中須英輔「MPEG-2 映像符号化」映像情報メディア学会技術報告, vol.22, no.52,
pp.19-24, BCS98-41 (Sep.1998)
- (5) 中須英輔「符号化画像の評価技術」電子情報通信学会技術報告 画像工学会
pp.25-32, IE2000-12 (June 2000)
- (6) 中須英輔「ビデオ系の品質評価」画像電子学会 第 5 回 VMA 研究会 (Sept. 2000)
- (7) 中須英輔「デジタル符号化画像の画質評価技術」電子情報通信学会技術報告 pp.7-12,
EID2000-46 (Sept. 2000)

著書

- (1) 「ハイビジョンのすべて～ソフトからハードまで～」デジタル帯域圧縮技術、p.288-295、
NHK 出版(1996)
- (2) NHK 放送技術研究所編「マルチメディア時代のデジタル放送技術事典」第 2 章
p.154-157、丸善(1994)
- (3) 「C 言語で学ぶ実践画像処理」第 9 章 p.127-137、12 章 p.187-206、オーム社(1992)
- (4) テレビジョン学会編「MPEG」8 章 8-1 p.179-188、オーム社(1996)
- (5) 映像情報メディア学会編「放送システム」第 2 部 7 章 7.1 p.128-149、コロナ社(2003)
映像情報メディア学会 丹羽高柳賞著述賞 受賞
- (6) 映像情報メディア学会編「デジタル放送ハンドブック」第 2 編 1 章 p.26-41、オーム社
(2003)