

Title	画質クラスを考慮したカラー符号化画像の画質評価モデル
Author(s)	古性, 淑子
Citation	
Issue Date	1998-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1173
Rights	
Description	Supervisor:小谷 一孔, 情報科学研究科, 修士

修 士 論 文

画質クラスを考慮した
カラー符号化画像の画質評価モデル

指導教官 小谷一孔 助教授

北陸先端科学技術大学院大学
情報科学研究科情報処理学専攻

古性淑子

1998年2月13日

要旨

符号化によって劣化した画像の画質評価には専ら主観評価が用いられているが評価条件等により結果が変動する等の問題があり、人間の視覚特性を考慮した客観的な評価方法が必要とされている。本研究ではカラー画像の JPEG 符号化誤差を原画像と符号化再生画像の差である色差として計算し、色差から画質劣化の状態を表す画質劣化要因を定義し、多変量解析の手法を用いて画質評価モデルを構築する。また、画質クラスを考慮する事により評価精度の向上をめざす。本方式により得られた評価モデルは人間による主観評価値を約 87%の精度で近似できた。更に評価モデルの構築に視点停留特徴を測定した結果を適用する事を試みた。

目次

1	序論	1
1.1	研究の背景と目的	1
1.2	研究の概要	2
1.3	本論文の構成	3
2	画像符号化歪みと画質評価において考慮すべき視知覚特性	4
2.1	JPEG 符号化方式	4
2.1.1	標準方式のカテゴリ分け	5
2.1.2	ベースライン・プロセスのアルゴリズム	6
2.1.3	データ構造	15
2.1.4	cjpeg における quality	16
2.2	JPEG 符号化によって生じる歪み	16
2.3	画質評価において考慮すべき視知覚特性	17
2.3.1	画像全体に関する視覚特性	17
2.3.2	画像の局所領域における視覚特性	19
2.3.3	空間相関の知覚	20
3	主観評価実験	21
3.1	テスト画像	21
3.1.1	原画像	21
3.1.2	符号化の種類	25
3.2	観視条件	29
3.3	評価者	30

3.3.1	評価者のトレーニング	30
3.3.2	トレーニングの結果	33
3.4	主観評価実験結果	38
4	種々の色空間における画質評価モデル	40
4.1	評価モデルの定義	41
4.1.1	システムの構成	41
4.1.2	劣化の種類と画質劣化要因の関数近似	42
4.1.3	画質劣化要因 F_i の主成分分析	44
4.1.4	評価値の計算	45
4.2	RGB 色空間における画質評価	45
4.2.1	画質評価結果	45
4.2.2	必須画質劣化要因	48
4.3	YCbCr 色空間における画質評価	49
4.3.1	画質評価結果	50
4.3.2	必須画質劣化要因	52
4.4	CE L*a*b*色空間における画質評価	53
4.4.1	画質評価結果	53
4.4.2	汎用性	55
4.4.3	必須画質劣化要因	63
4.5	CE L*u*v*色空間における画質評価	64
4.5.1	画質評価結果	65
4.5.2	汎用性	67
4.5.3	必須画質劣化要因	69
4.6	むすび	70
5	画質クラスを考慮した画質評価モデル	73
5.1	画質クラスを考慮した評価モデルの定義	73
5.1.1	システムの構成	73
5.1.2	基礎画質劣化要因と主成分分析	74
5.1.3	評価値の計算	74

5.2	評価モデルによる画質評価	74
5.2.1	画質評価結果	74
5.2.2	汎用性	77
5.2.3	必須画質劣化要因	81
5.3	むすび	82
6	画質評価時における視点停留領域	83
6.1	視点停留領域測定実験の条件	83
6.1.1	測定システム構成	83
6.1.2	テスト画像と符号化の種類および評価者	84
6.1.3	観視条件	88
6.2	視点停留領域の定義	89
6.3	視点停留領域の測定結果	90
6.3.1	原画像に対する視点停留領域	90
6.3.2	符号化再生画像に対する視点停留領域	91
6.4	むすび	96
7	画質評価時における視点停留領域の評価モデルへの適用	97
7.1	視点停留領域を考慮した画質評価モデルの定義	97
7.1.1	システムの構成	97
7.1.2	視点停留情報の関数近似と画質劣化要因の主成分分析	98
7.1.3	画質劣化要因 F_i の主成分分析と評価値の計算	99
7.2	評価モデルにおける画質評価結果	99
7.3	むすび	101
8	結言	102

第 1 章

序論

1.1 研究の背景と目的

画像データを高速かつ高効率に伝送するには符号化技術が欠かせない。JPEG 等の圧縮率の高い非可逆符号化では、符号化誤差によるブロック歪みや輪郭の不連続などの劣化を生じる。符号化誤差により劣化した画像の評価には主観評価が多く用いられているが、主観評価は人間が画像を見て評価を行なう為、労力（‘1 回のセッションで約 30 分、最高で 40 枚の画像を提示’の繰り返し）を必要とし、信頼性を満たす為には多くの評価者による膨大な回数の実験が必要となる。

評価結果の変動を軽減する為、主観評価に関しては ITU によって勧告 [1] がされているが、この手法を用いても十分な信頼性を確保するのが困難であるという報告 [2] もある。また時間と労力の削減の為に客観的な画質評価方法がいくつか提案されている。画質評価方法を実現する際に、それが人間の視知覚特性をどれだけ忠実に反映しているかという事が重要である。従来から用いられている荷重平均 2 乗誤差 (WMSE) は、ランダムノイズによる劣化は評価できるが、これより 10 倍以上大きな妨害と人間が感じる局所的な歪みによる劣化は評価できない。人間の視覚特性を考慮し、局所的な特徴の損傷まで捕らえられる高精度な画質評価方法が実現されれば、単に劣化した画像を評価するだけでなく、その過程および結果を符号化方法に反映できる為、より優れた符号化技術の開発が効率良く行なえる [3]。

人間の視覚特性を考慮し、局所的な特徴の損傷をも捕らえる事の出来る客観的な画質評価モデルが構築されれば、上記の問題は解決出来る。現在、修正 Munsell 色空間にける色

差を計算し、評価モデルを構築する事で、良好な評価結果を得られることが報告されているが [4][5]、修正 Munsell 色空間が最適であるかどうかは確認されていない。

本研究では、原画像を基準として符号化による画質劣化の程度を評価する為に、誤差を原画像と符号化画像の差である色差で定義した評価モデルを構築する。評価モデルを用いて更に、画質クラスと視点停留を考慮した評価モデルを構築し、評価精度の向上を試みる。

1.2 研究の概要

評価モデルの構築にあたって JPEG 符号化による画質の劣化の要因を色差の関数として与える。画質の劣化の要因を計量する為、発生している種々の歪みをグローバルな歪み、テクスチャ状のパターンを作る歪み、画像内容に関わるローカルな歪みに分類し、近似的に定式化した次の 4 つの基礎画質劣化要因を用いて計算を行なう。

F_1 : 一画素あたりの色差の量 (ランダムノイズに相当)

F_2 : 符号化サブブロック間の誤差変化量

F_3 : 誤差の自己相関係数

F_4 : 輪郭部分における誤差の重みづけ (画質評価時の視点停留領域)

評価値はこれらの基礎画質劣化要因 F_i 間の相関を取り除く為に主成分分析を施し基本画質劣化要因 Z_i を求め、この Z_i と主観評価値とで重回帰分析を行ない、 Z_i の線形和として得る。また 符号化した際の画質によって知覚される歪みの特徴および量が異なるという考えから、画質クラスを考慮して評価モデルを構築する。

評価用画像には ITE のデジタル標準画像の他、人物, CG, 風景, 建築構造物, 静物等の多種の原画像 15 枚を PDS として広く用いられている `cjpeg` で画質をコントロールするパラメータを 8 種類に変えて符号化して得た符号化再生画像を用いる。

構築した評価モデルの妥当性を確認する為、また評価モデルを構築する際に主観評価値を用いて重回帰分析を行なう理由から、ITU の勧告 [1] に基づき主観評価実験を行なう。評価者に評価に対する熟練度を求めるという観点から主観評価実験のトレーニングを行なう。

人間が画質評価を行なう際の視点停留領域を評価モデルに適用することで評価精度が向上すると考えられる。そこで評価者にアイマークレコーダーを装着させ、視点停留情報を得て、これを評価モデルに適用する。

1.3 本論文の構成

本論文は8章で構成される。第1章は序論である。第2章では、JPEG符号化方式と、それによって生じる歪みおよび画質評価において考慮すべき視知覚特性について述べる。第3章では、人間が実際に画像を見て評価する主観評価実験の実験条件および結果を示す。第4章においては、本研究で構築する評価モデルの定義および、RGB, YCbCr, CIE $L^*a^*b^*$, CIE u^*v^* 色空間における画質評価結果を示す。第5章では、画質クラスを考慮した評価モデルについて示す。第6章では、人間が画質評価を行なう際の視点停留領域の測定法およびその結果についてまとめる。第7章において、視点停留領域の測定結果を基礎画質劣化要因として評価モデルに適応した結果をまとめる。第8章で結論を示す。付録として、ITU-R BT Rec.500-7 で勧告されている主観評価について記載する。

第 2 章

画像符号化歪みと画質評価において考慮すべき視知覚特性

客観的画質評価モデルを構築する際、その評価モデルが人間の視覚特性を考慮しているという事が重要である。ここでは、人間の視覚特性を画像全体に関する視覚特性、局所領域における視覚特性、および空間相関の知覚の 3 つに分けて説明する [3]。

2.1 JPEG 符号化方式

カラー静止画像の符号化方式の標準化は、1982 年より検討が開始された [7]。カラー静止画像の符号化の検討は当初ビデオテックスでの利用を想定して、ISO/IEC/JTC1/SQ2/WG8 で進められた。一方 1986 年 3 月に CCITT (Common Component for Image Communication) SG VIII NIC (New Image Communication) グループにプログレッシブ・ビルドアップ機能を持つ符号化方式が KDD より提案された事を機会に、CCITT SG VII でもテレマティックス・サービス全般に適用できるカラー静止画像符号化方式の標準化の必要性が認識された。個々で独自のカラー静止画像符号化方式を用いる事は相互接続の観点から望ましい事とは言えない為に、テレマティックス・サービスに幅広く適用できるカラー静止画像の研究が始められ、CCITTとISOとが同様の研究を行なう事となった。この為に共同作業を行なう標準化組織として JPEG (Joint Photographic Expert Group) が設立された。JPEG アルゴリズムは CCITT SG VII ラポータ・グループと CCITT と ISO/JTC1/CS29/WG10 のジョイント・グループ JPEG で進められた為、このような名前で

呼ばれている。

JPEG 設立後、符号化方式選定の為のコンテストが2度行なわれ、最終的に DCT (Discrete Cosine Transform 離散コサイン変換) をベースとした符号化方式が基本標準案として決定され、各種作業の後に標準化がなされた [8]。

2.1.1 標準方式のカテゴリ分け

JPEG アルゴリズムには図 2.1に示すように大きく2つの圧縮方式に分ける事ができる。第1の方式は DCT を基本とした非可逆方式であり第2の方式は可逆方式である。DCT 方式はベースライン・プロセス (必須機能) と拡張 DCT プロセス (オプション機能) の二つに分類される。

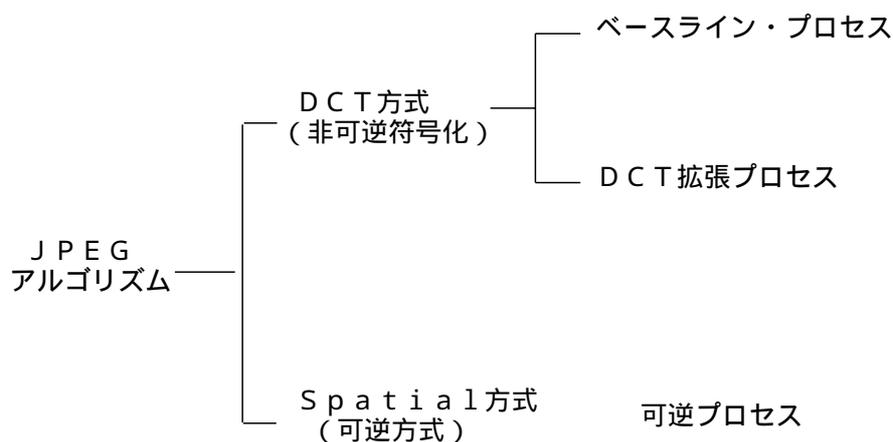


図 2.1: JPEG アルゴリズムのカテゴリ

各プロセスの必須機能をまとめると以下のようなになる。

—ベースライン・プロセス—

ベースライン・プロセスは、DCT 方式をサポートする全ての復号器が持たなくてはならない最小限の機能であり、具体的には

1. 8 ビット精度の画像を取り扱えなくてはならない
2. シーケンシャル・ビルドアップによる画像再生が出来る

3. ハフマン符号化をサポートする。符号化テーブルとして2つのACおよびDC符号化テーブルをサポートする

という機能が必要とされる。

—拡張 DCT プロセス—

拡張 DCT プロセスは、より広い領域のアプリケーションを満足させる為に用意されたベースライン・プロセスのオプション機能で、具体的には

1. 8 ビットまたは 12 ビット精度の画像を取り扱えなくてはならない
2. シーケンシャル・ビルドアップによる画像再生または、プログレッシブ・ビルドアップによる画像再生が出来る
3. ハフマン符号化または算術符号化をサポートする。符号化テーブルとして4つのACまたは4つのDC符号化テーブルをサポートする

という機能が必要とされる。

—可逆プロセス—

可逆プロセスにおいては

1. 予測処理を行なう
2. N ビット精度 ($2 \leq N \leq 16$) の画像を取り扱えなくてはならない
3. シーケンシャル・ビルドアップによる画像再生が出来る
4. ハフマン符号化または算術符号化をサポートし、符号化テーブルとして4つの符号化テーブルをサポートする

という機能が必要とされる。

2.1.2 ベースライン・プロセスのアルゴリズム

JPEG アルゴリズムの基本となるベースライン・プロセスのアルゴリズムを以下に説明する。

—基本構成—

図 2.2に基本構成のブロック図を示す。

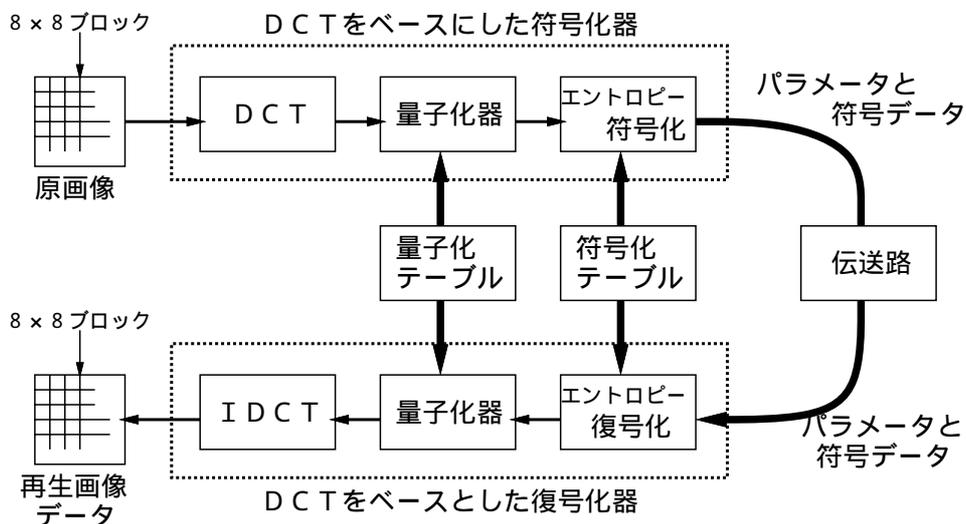


図 2. 2: 基本構成

入力画像は符号化器において、DCT、量子化、エントロピー符号化が行なわれ、圧縮される。復号器では、圧縮されたデータに対して、エントロピー復号化、逆量子化、逆DCTが行なわれ、復号画像が出力される。

—DCT—

図 2. は示すように、入力画像は 8×8 画素のブロックに分割される。そのブロック内の画素の値を $P_{xy}(x, y = 0, \dots, 7)$ と示す。

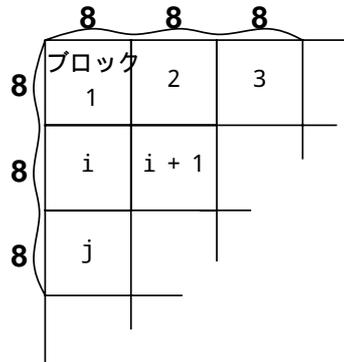


図 2.3: 画素ブロック

分割された 8×8 画素ブロックに対して式 (2.1) による 2 次元 DCT 変換を行なう。

$$S_{uv} = \frac{1}{4} C_u C_v \sum_{x=0}^7 \sum_{y=0}^7 (P_{xy} - L_s) \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2y+1)v\pi}{16} \quad (2.1)$$

ただし、

$$\begin{aligned} x, y &= \text{ブロック内の画素の位置} \\ u, v &= \text{DCT 係数の位置} \\ C_u, C_v &= \begin{cases} \frac{1}{\sqrt{2}} & u, v = 0 \\ 1 & \text{otherwise} \end{cases} \\ L_s &= \begin{cases} 128 & P_{xy} = 8 \text{ ビット精度} \\ 2048 & P_{xy} = 12 \text{ ビット精度} \end{cases} \end{aligned}$$

である。

変換の結果、64 個の係数 S_{uv} が得られる。係数 S_{00} は DC 係数と呼ばれ、残りの 63 個の係数は AC 係数と呼ばれる。DC 係数は 8×8 画素の平均値 (直流成分) を示している。各画素値から L_s を引く事により、DC 係数の期待値を 0 にレベルシフトしている。すなわち 8×8 画素の平均値が 128 の場合は、DC 係数の値は 0 となる。

変換係数 S_{uv} に対する基底は左上と右下がそれぞれ変換係数 S_{00} 、 S_{77} に対応する。左から右に進むにつれて高周波の水平周波数成分を、上から下に進むにつれ高周波の垂直成分を多く含む事になる。任意の 8×8 画素はこの 64 個の基底ベクトルの線形結合で表現する。

原画像を DCT 変換した結果の一部を拡大すると任意の 8×8 画素は 64 個の基底ベクトルの線形結合で表現されている事が確認できる。

また、式 (2.2) に示す逆変換で変換係数 S_{uv} から、 8×8 画素のブロック画像 $P_{xy}(x, y = 0, \dots, 7)$ を得る事が出来る。

$$P_{xy} = \frac{1}{4} \sum_{x=0}^7 \sum_{y=0}^7 C_u C_v S_{uv} \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2y+1)v\pi}{16} + L_s \quad (2.2)$$

ただし、

$$\begin{aligned} x, y &= \text{ブロック内の画素の位置} \\ u, v &= \text{DCT 係数の位置} \\ C_u, C_v &= \begin{cases} \frac{1}{\sqrt{2}} & u, v = 0 \\ 1 & \text{o therwize} \end{cases} \\ L_s &= \begin{cases} 128 & P_{xy} = 8 \text{ ビット精度} \\ 2048 & P_{xy} = 12 \text{ ビット精度} \end{cases} \end{aligned}$$

である。

各画素値に L_s を加える事により、レベルシフト分を元に戻す。逆変換された P_{xy} 値が指定されたビット精度を越える場合は次に示すクリッピング処理を行なう必要がある。

・ P_{xy} のビット精度が 8 ビットの場合

$$P_{xy} = \begin{cases} 255 & P_{xy} > 255 \\ 0 & P_{xy} < 0 \end{cases}$$

ブロックは 64 個の画素を含んでおり、その数は DCT 変換を行なっても変わらない。画像を DCT 変換した時の電力分布を見ると、輝度が高いほど電力が大きくなり、多くの情報を含んでいる事が確認できる。DCT 変換を行なうと、画像情報が一部の係数に集中する。(DC 成分が特に大きく、AC 成分においても低周波成分に電力が集中する。) したがって、少ないビット数で効率的に符号化を行なう為には、低周波成分の係数に多くのビット数をもたせて符号化を行なう。

—量子化—

64 個の DCT 係数 S_{uv} をそのまま符号化して復号器に伝送すれば 8×8 画素のブロック画像を得る事が出来るが、高い符号化効率を得る事は出来ない。JPEG のアルゴリズムでは

再生画像の画質を選択する事が出来、そのコントロールは量子化処理の時に行なわれる。

DCT 係数 S_{uv} は符号化量子化テーブル Q_{uv} を用いて係数位置ごとに異なるステップサイズで線形量子化される。量子化された係数 r_{uv} は式 (2.3) で決定される。

$$r_{uv} = \text{round}(S_{uv}/Q_{uv}) \quad (2.3)$$

round とは最も近い整数への整数化を意味する。

Q_{uv} の値を小さい値に設定すると、画質の良い符号化画像を生成する事が出来る。逆に Q_{uv} の値を大きくすると S_{uv} の値が小さくなり、符号化情報量が減少するが画質は劣化する。このように量子化テーブル Q_{uv} を変える事により、画質と符号化情報量を自由にコントロールする事が出来る。量子化テーブルの値にデフォルトはなくアプリケーションや画像特性毎に出力装置に対して最適化する事が出来る。

一般的に人間の視覚特性はローパスフィルタの特性を持っているといわれている為、高周波成分については多少粗い量子化を行なってもその影響は目立たない。この為量子化テーブルの高周波成分には大きな値が割り当てられている。

—エントロピー符号化—

量子化された r_{uv} 係数はエントロピー符号化されて出力される。エントロピー符号化はハフマン符号化方式を使用する。DC 係数 (r_{00}) と AC 係数 (r_{00} 以外) とでは符号化方法は異なる。

[DC 係数の符号化]

DC 係数のハフマン符号化のブロックダイアグラムを図 2.4に示す。

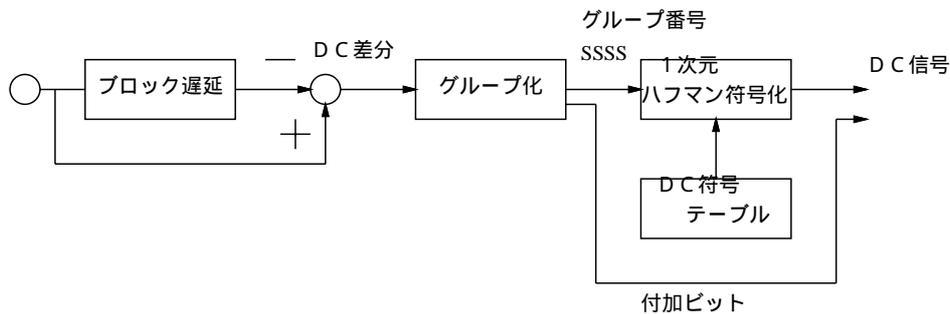


図 2.4: DC 係数のハフマン符号化

DC 係数は図 2. に示すように 1 つ前に符号化した同一色成分のブロックの DC 係数との差分を符号化する。

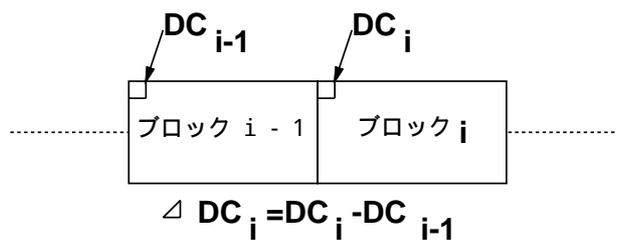


図 2.5: 差分

DC 係数は 8×8 画素ブロックの平均値を示している。一般的な画像においては隣のブロックとの平均値が大きく変化する事はあまりないので、1 つ前の DC 係数との差分値は 0 の近傍に集中する。差分を符号化する事により高能率な符号化を期待する事が出来る。DC 係数との差分値を基準にしたがってグループ化する。DC 差分値は、このグループ番号 (SSSS と付加ビットとで表現される。

[AC 係数の符号化]

AC 係数のハフマン符号化のブロックダイヤグラムを図 2.6 に示す。

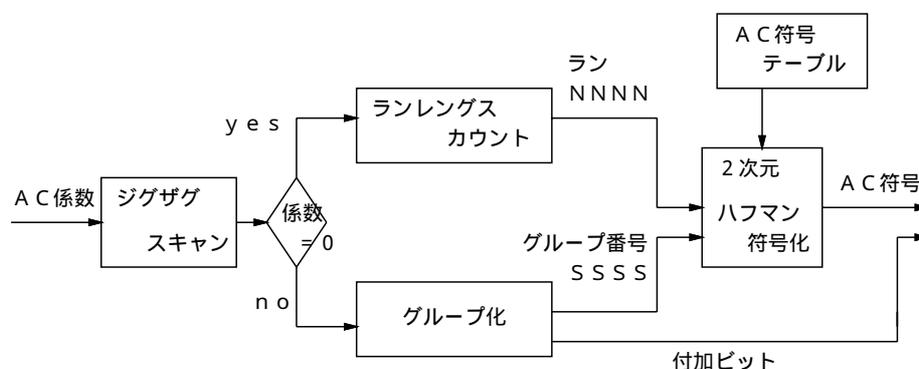


図 2. 6: AC 係数のハフマン符号化

AC 係数は図 2. 7 に示すようなジグザグスキャンによって 1 次元に並びなおされる。

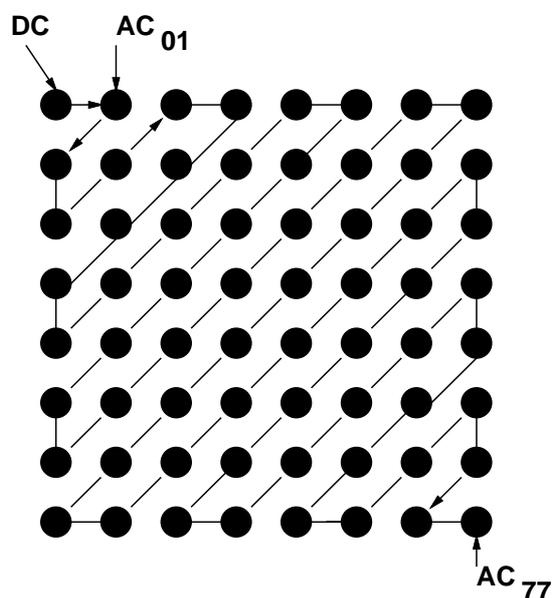


図 2. 7: ジグザグスキャン

この 1 次元に並びなおしをした AC 係数は、連続する 0 の係数 (無効係数) の長さを表すラン長と 0 以外の係数 (有効係数) の値を用いて符号化される。すなわち無効係数のラ

ン長 (NNNN) と有効係数のグループ番号 (SSSS) を用いてハフマン符号化する。NNNN と SSSS の組合せによってハフマン符号が割り当てられている。SSSS は有効係数が属するグループを限定するだけで、その中の特定の値を示す為に付加ビットが用いられる。

図 2.8 に 2 次元ハフマン符号の構成を示す。

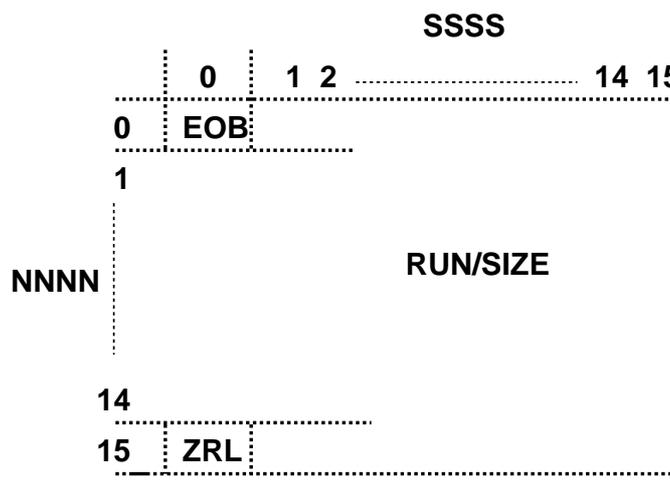


図 2. 8: 2 次元ハフマン符号の構成

ブロック内の最後の AC 係数が 0 の時は、最後の有効係数に対する符号の直後に EOB (End of Block) をつけ、ブロックの符号化を終了する。ただしブロック内の最後の AC 係数 (AC_{77}) が 0 以外の場合は EOB はつけない。また、無効係数のラン長が 15 を越える場合は 16 の無効係数のラン長を示す ZRL を残りのラン長が 15 以下になるまで出力した後に、残りのラン長を NNNN として符号化する。

—ハフマン符号化テーブル—

ハフマン符号化テーブルはベースライン・プロセスで 2 個設定でき、成分毎に切替えて用いる事が出来る。このハフマン符号テーブルにもデフォルトはなく、アプリケーション毎にデフォルトハフマン符号テーブルを定める事が出来る。しかし異なるアプリケーション間では復号化の為にハフマン符号化テーブルを送る必要がある。ただしこの場合、符号化テーブルを直接送るのではなく符号テーブルを作成する為の情報として以下の 2 つを送れば良い。

- 何ビットの符号が何個あるかというテーブル。ただしハフマン符号の最長ビットは16である。
- 発生頻度順に並べた符号化要素 (DC 係数なら SSSS、AC 係数なら NNNSSSS)

受信側でこれらの情報をもとにハフマン符号テーブルを作成する。

—複数色成分の転送順序—

画像が複数の色成分で出来ている場合、エンコーダーは規定された色成分の順に符号化し、色成分に応じてテーブルも切替える。テーブルとは量子化テーブルと、ハフマン符号テーブルの事である。符号化の順序は以下に示す通り、大きく分けて ノンインタリーブとインタリーブがある。

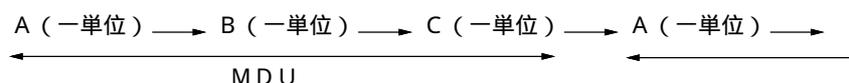
[ノンインタリーブ]

1つの色成分に対し、画像全体の符号データを出力し終えた後に、次の色成分を符号化する。3つの色成分 (A, B, C) の場合は以下のような順番で符号化と出力が行なわれる、

A(符号化、出力) → B(符号化、出力) → C(符号化、出力) → 終了

[インタリーブ]

1つの色成分に対し一単位の画素領域を符号化した後に、次の色成分の一単位を符号化する。繰り返される単位のセットを MDU (Mini Data Unit) という、



MDU には、画素とブロックが定められており、それぞれ画素インタリーブ、ブロックインタリーブ、という。

DCT方式のシーケンシャルでは、ノンインタリーブか8×8画素のブロックを一単位としたブロックインタリーブを選択する事が出来る。DCT方式のプログレッシブではDC係数はノンインタリーブか8×8画素のブロックを一単位としたブロックインタリーブを選択する事が可能であるが、AC係数はノンインタリーブのみである。インタリーブでは符号化成分が順次スイッチすると、量子化テーブルなどのパラメータも同期してスイッチする。

—逆量子化—

復号側では逆量子化を行なう。ハフマン符号化で得られた係数を r_{uv} とすると、式 (2.4) で逆量子化を行なう事が出来る。

$$S_{uv} = r_{uv} \times Q_{uv} \quad (2.4)$$

2.1.3 データ構造

図 2. 9に符号化データのデータ構造を示す。

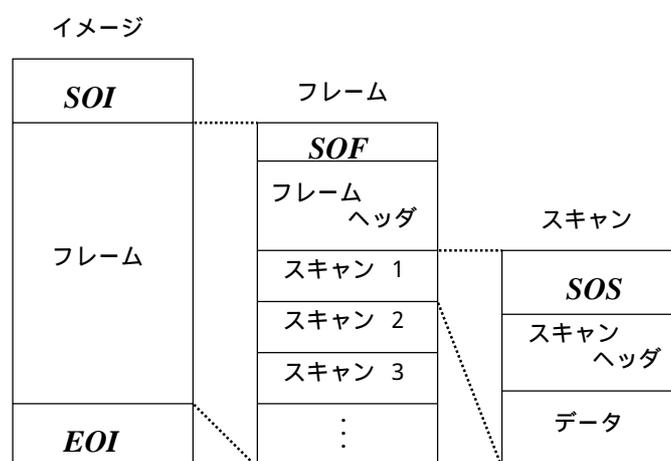


図 2. 9: 符号化データのデータ構造

符号化データの構造は図 2. 9に示したように、イメージ/フレーム/スキャンという大きな3種類の構造から構成されている。1枚の画像は、画像データの始まりを示す SOI(Start of Image) マーカーと画像データの終了を示す EOI (End of Image) マーカーで、挟まれたフレームで構成される。

フレームの先頭には、そのフレームに関連する諸パラメータを指定したフレームヘッダがつき、そのあとに1つ以上のスキャンが続く。スキャンの先頭にはそのスキャンに関連する諸パラメータを指定したスキャンヘッダがつき、その後には符号化データが続く。

全ての色成分がインタリーブされている場合は、全ての色成分の情報を一回の符号化処理で符号化されている為、フレーム内には1つのスキャンしか含まれない。しかし、ノン

インタリーブの場合は、色コンポーネント毎に符号化処理が行なわれる為、色コンポーネントの数だけスキャンが存在する。

2.1.4 cjpeg における quality

本研究では JPEG 符号化に広く用いられている PDS である cjpeg を使用した。この cjpeg ではオプションの quality パラメータを変化させる事により画質をコントロールする事が出来る。quality の値を変化させると、量子化のステップ数が変化する。

画質を決定する為のパラメータの値は、1 から 100 までの間で設定する事ができ、この値から scale 値を式 (2.5) を用いて決定している。

$$\text{scale} = \begin{cases} 5000 / \text{quality} & \text{quality} < 50 \\ 200 - \text{quality} \times 2 & \text{otherwise} \end{cases} \quad (2.5)$$

ここで得られた scale の値と基本量子化テーブルの値 64 個と、式 (2.6) を用いて、新しい量子化テーブルの値 64 個を作成している。

$$\text{new quant table}_i = (\text{basic quant table}_i * \text{scale} / 50) / 100 \quad (2.6)$$

ここで i の範囲は ($1 \leq i \leq 64$) であり、計算された値が 0 以下の場合には値を 1 とし、255 以上だった場合は 255 に値を変換してリミッターをかけている。

quality=100 の時は scale は 0 となり、この場合は new quant table の要素は全て 1 とする。DCT 係数 S_i は量子化テーブルの要素 Q_i のステップサイズで線形量子化される。量子化された係数 r_i は式 (2.7) で決定される。

$$r_i = \text{round}(S_i / Q_i) \quad (2.7)$$

round とは最も近い整数への整数化を意味する。

quality の値が小さいほど量子化を行なう際のステップサイズが大きくなり粗く量子化される為、画質が劣化する。逆に quality の値が大きいほど得られる量子化テーブルの値は小さくなり細かく量子化される為、画質が向上する。

2.2 JPEG 符号化によって生じる歪み

JPEG 符号化方式によって生じる歪みは以下のように大別される。

1. グローバルな歪み
2. テクスチャパターンを作る歪み
3. 輪郭損傷のような画像内容に関わる局所的な歪み

これらの歪みが生じる原因を以下に示す。

(1) 輝度信号と色信号のランダムノイズ

ランダム雑音に近い性質を持つ符号化雑音は画像全体にばらまかれる。視覚の感度が鋭敏になる画像の平坦部において、量子化雑音として検知され画質劣化の原因となる。

(2) ブロック歪みや輪郭の不連続

画像を 8×8 ブロックに分割してブロック毎に符号化を行なう事により、ブロック境界で不連続な歪みを生じる。その結果、サブブロック構造が知覚されたり、連続した輪郭がジグザグ状の歪みとなり、大きな妨害となる。

(3) モスキート雑音

輪郭などのエッジ周りには、ざらざらした雑音や色のにじみが発生する。これをモスキート雑音といい、高周波成分を粗く量子化した為に発生する。

2.3 画質評価において考慮すべき視知覚特性

2.3.1 画像全体に関する視覚特性

—Weber-Fechner の法則—

一般に刺激 I が与えられている時、やっと知覚できる刺激の増分 (Just Perceptible Difference) を ΔI とすると $\Delta I / I$ に対する人間の感覚はある I の範囲において一定となる。これを Weber の法則といい、比例定数を k として

$$\Delta E = k \cdot \Delta I / I \quad (2.8)$$

で表される。

ただし、 ΔE は刺激 I が存在する時の刺激の増分 ΔI に対する感覚の増分である。 ΔI は充分小さくなると仮定して ΔI を dI と書けば式 (2.8) を積分して

$$E = k \cdot \log I \quad (2.9)$$

と書く事が出来、式 (2.9) を Weber-Fechner の法則という。

広いレンジの刺激に対しては k は一定として扱えない為、Stevens が実験的に式 (2.10) で感覚の強さを示した。この法則を Stevens の法則という。

$$E = k(R - R_0)^\beta \quad (2.10)$$

ここで R_0 は刺激閾である。明るさの感覚の例では、見込み視覚 5° の指標に対して $\beta = 0.33$ となる。人間の視覚特性は線形的ではなく、式 (2.9) のような非線形変換で表した時にも k が刺激に対して一定とならず、扱いが難しいが、テレビジョンの CRT を観察する範囲においてはほぼ $k = 0.02$ で一定である。

被写体の明るさと信号レベルの関係が先に示したような非直線性をもつように補正してから信号を均等量子化、符号化等の処理をする事によって、符号化誤差を視覚的妨害量として結びつけやすい。一方これとは無関係にテレビの伝送系では受像管の発光輝度が被写体の輝度変化と直線的に変化するように予め送り側で受像管の入力信号-光発光強度の非直線補正を行なっている。この補正を γ 補正という。好運にもこの γ 補正が、式 (2.9) に基づく非直線変換の特性と比較的良く一致しており、 γ 補正されている輝度信号を用いる事により、補正を行なわなくても近似的に視覚的に比例する量として取り扱う事が出来る。 $\gamma = 3$ の場合が最も近似精度が良い。

— 視覚の空間周波数特性 —

視覚の空間周波数特性が、画像符号化や処理において検討すべき重要な特性であるのは以下に示す理由からである。

1. 信号処理は周波数領域で行なう事が多く、その効果を視覚の空間周波数特性を利用して判断・説明できる。
2. 画像の処理や符号化によって生じた誤差が画質に及ぼす影響を周波数領域で視覚の空間周波数特性を乗じる事により判断できる。

(1) の一例として、画像の輪郭がシャープではなく視覚上のマッハ効果が現れない時に、マッハ効果に関連付けられる空間周波数特性で荷重する事によって画像の鮮鋭度を上げる事が出来る。この例は視覚の空間領域と空間周波数領域が、フーリエ変換で結びつけられている例である。

(2) の手法として、視覚系を線形であると仮定し、原画像と処理画像の差分誤差をフーリエ変換し、周波数領域において、視覚の空間周波数特性を乗じ、再び逆フーリエ変換して視覚的妨害量を近似する。

—色覚の空間周波数特性—

無彩色を中心とする補色間の正弦波パターンの色差弁別により視覚の色度空間周波数特性を測定した結果より、周波数特性はバンドパスフィルター型で色度パターンの空間周波数特性は色相によって大きく変化することが明らかになっている。色相を変えて色度空間周波数特性を求めると、R-G の方向が最も帯域が広く、Y-B の方向が最も帯域が狭い [10]。

評価モデルを構築するに当たって、色度の空間周波数特性を荷重するが、現在色度の空間周波数特性は R-G, Y-B 特性のみが明らかになっている為、RGB 信号を変換して特性の荷重を行なった。

2.3.2 画像の局所領域における視覚特性

—マッハ効果—

視覚の MTF(帯域通過特性) と線形性により解釈できる輪郭の見え方として、マッハ効果がある。この現象は、輝度のステップパターン入力に対して明るさのアンダーシュート、オーバーシュートを生じる。したがってレベルの低い境界はより低く感じられ、レベルの高い境界はより高く感じられ、結果として白黒のエッジが強調されて見える [7]。

—マスキング効果—

大きな刺激が存在すると、同時あるいは近傍に存在する小さな刺激は知覚され難くなる事をマスキング効果という。視覚系が線形システムではない事を示す一例である。画像符号化において“強い輪郭部分では、少々誤差を生じてもその誤差は見えにくい”性質として

有名である。画像信号の符号化では、この効果は量子化ステップ数の削減の根拠とされている。しかし、この性質は輪郭に直交する方向に対しては正しいが、輪郭に沿う方向には必ずしも正しいとは言えない [3]。

2.3.3 空間相関の知覚

人間の視知覚系には、隣接する画素間の相関から画像内容を認識する性質がある。すなわち、空間的には周囲とのつながりから認識する。この為、小面積内の誤差は、個々では検知限以下であっても空間的に相互に関連を持っていると、人間の目に大きな妨害として映る。

この妨害には次のようなものがある。

1. 小さな誤差であるが、空間的なつながりや、まとまりを持ってテクスチャ状のパターンを知覚させる妨害
2. 滑らかであった輪郭が符号化誤差によって蛇行したり、ジグザグ状に見える妨害

(1) の例としてブロック歪みが挙げられる。また、(2) の妨害は、輪郭の連続性を損ね、副尺視力によって高感度に知覚される。

これらの妨害は、画像の輪郭などの局所的特徴を壊してしまう。人間は局所特徴や、まとまり、つながりに注目して画像を認識しているから、これらの特性は、画質評価尺度の検討において、非常に重要な知識となる [1] 。

第 3 章

主観評価実験

偏回帰係数を計算する際に主観評価値を用いる為、また得られた評価尺度の妥当性を確認する為に予め主観評価実験を行ない、評価値を求めておく。

ここでは、劣化特性を測定する為、主観評価実験に関する勧告に従い、2重刺激劣化尺度法 (EBU method) を用いる [1]。

以下に評価実験の条件および、結果を示す。

3.1 テスト 画像

3.1.1 原画像

画像符号化によって生じる誤差の特徴は、画像の平坦部に生じるグローバルな歪み、画像内容にかかわる局所的な歪み、テクスチャパターンを生じる歪み等が典型的であることを 2.2 節で示した。これらの歪みが生じるのは画像の平坦部と輪郭領域である事から、符号化誤差の量は画像の平坦部と輪郭部分の量等の画像内容の影響を受ける。画像の内容が大きく異なる画像を使用して定義された評価モデルが主観評価値を良好に近似できれば他の画像もカバーできると考え、ITE デジタル標準画像の他、CG, 風景, 静物, 建築構造物, 人物等、表 3.1 および図 3.1 から図 3.15 に示す 15 種類の画像を用いた。これらの画像は画像サイズ 256×256 [pixel]、RGB 各 8 [bit/pl] の階調をもつカラー画像である。

表 3.1: 評価に用いた原画像

カテゴリ	原画像名
人物	barbar a
	boy childr en hai r band weather
風景	bar a wi ne
静物	br ush s ake

カテゴリ	原画像名
建築構造物	bui l chur ch ggbr dg jai st
CG	cg1 cg-j ai st



图 3.1: barbara



图 3.2: boy



图 3.3: childr en



图 3.4: hai r band



图 3.5: weather



图 3.6: bara



图 3.7: wi ne



图 3.8: br ush



图 3.9: sake



☒ 3.10: buil



☒ 3.11: church



☒ 3.12: ggbr dg



☒ 3.13: jai st



☒ 3.14: cg1



☒ 3.15: cg-j ai st

3.1.2 符号化の種類

各原画像とも符号化は PDS として広く用いられている cjpeg を使用し、伝送レート 0.40 ~ 4.36[bit/pixel]の間でパラメータを 8 種類に変化させ JPEG 符号化を行なう。基本的な歪み形状を網羅するように F_i を定義している為、符号化で発生する他の歪みにも充分対応できると考えられる。原画像および cjpeg のよる符号化再生画像の例を図 3.16 から図 3.24 に示す。



图 3.16: barbara
原画像



图 3.17: barbara
符号化再生画像
(quality=20)



图 3.18: barbara
符号化再生画像
(quality=30)



图 3.19: barbara
符号化再生画像
(quality=40)



图 3.20: barbara
符号化再生画像
(quality=50)



图 3.21: barbara
符号化再生画像
(quality=60)



图 3.22: barbara
符号化再生画像
(quality=70)



图 3.23: barbara
符号化再生画像
(quality=80)



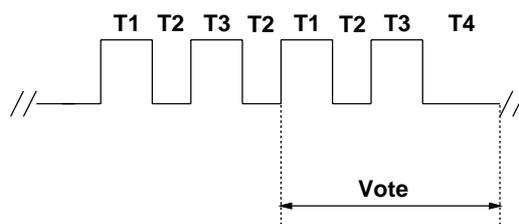
图 3.24: barbara
符号化再生画像
(quality=90)

3.2 観視条件

観視条件は表 3.2に示す通りである。

表 3. 2: 観視条件

視距離	4H(H：画像高)
室内照明	暗室
ピーク輝度	68[cd/m ²]
提示方法	連続提示(図 3.25 参照)
観視時間	(原画像 10 秒、符号化画像 10 秒)× 2
評価尺度	5 段階劣化尺度



- T1: original image 10sec
- T2: mid grey 3sec
- T3: codeing image 10sec
- T4: mid grey 10sec

図 3. 25: 画像提示方法

—評価語—

評価語は表 3. 3に示す 5 段階劣化尺度を使用する。この評価語は評価語間の重みのばらつきが多くならないように心理実験で決定されている。

表 3.3: 5 段階評価における評点と評価語

評点	評価語
5	(劣化が) 分からない
4	(劣化が) 分かるが気にならない
3	(劣化が) 気になるが邪魔にならない
2	(劣化が) 邪魔になる
1	(劣化が) 非常に邪魔になる

3.3 評価者

本研究において、評価者には評価に対する熟練度、すなわち安定性を求めている。小林の研究 [6] によって合計 10 時間程度の評価経験を積ませる事により、平均評点が収れんと報告されている。

本研究でも安定した評価結果を得る為に画質評価試験の経験のない非専門家 22 名 (女性 1 名を含む, 全ての評価者の色覚は正常) に合計 10 時間のトレーニングを行な確認した。以下にトレーニングの方法と結果を示す。

3.3.1 評価者のトレーニング

トレーニングには画質評価モデルを構築する際に使用する画像と異なる画像、図 3. 26 から図 3. 33 を用いている。画像サイズ 256×256 [pixel]、RGB 各 8 [bit/pl] の階調をもつカラー画像、“人物 (man, child)”, “建築構造物 (bui12, bui13)”, “CG (cg3, cg4)”, “風景 (ni wa, ni wa2)” の 4 種類の画像各 2 枚を jpeg を使用して quality=20, 40, 60, 80 の 4 種類に変化させ使用した。

1 回の評価実験は 30 分であるから 20 回の評価実験で 10 時間のトレーニングを行なった事になる。

観視条件は評価実験と同じ表 3. 2 に示す条件で行ない、評価者には勧告に定められた教示以外は一切行っていない。



图 3.26: man



图 3.27: child



图 3.28: bui l 2



图 3.29: bui l 3



图 3.30: cg3



图 3.31: cg4



☒ 3.32: niwa



☒ 3.33: ni wa2

3.3.2 トレーニングの結果

22名の評価者に対し、トレーニングを行ない、これによって結果が安定した例と、安定しなかった例を図3.34から図3.41に示す。図の縦軸には評点を横軸には実験回数を示している。

評価用画像は quality=20, 40, 60, 80の4種類の符号化を行っており、どの画像も合計8時間程度の評価経験を積む事で結果が安定している事が分かる。なお、トレーニングによって結果が安定しなかった評価者は22名中2名であった。

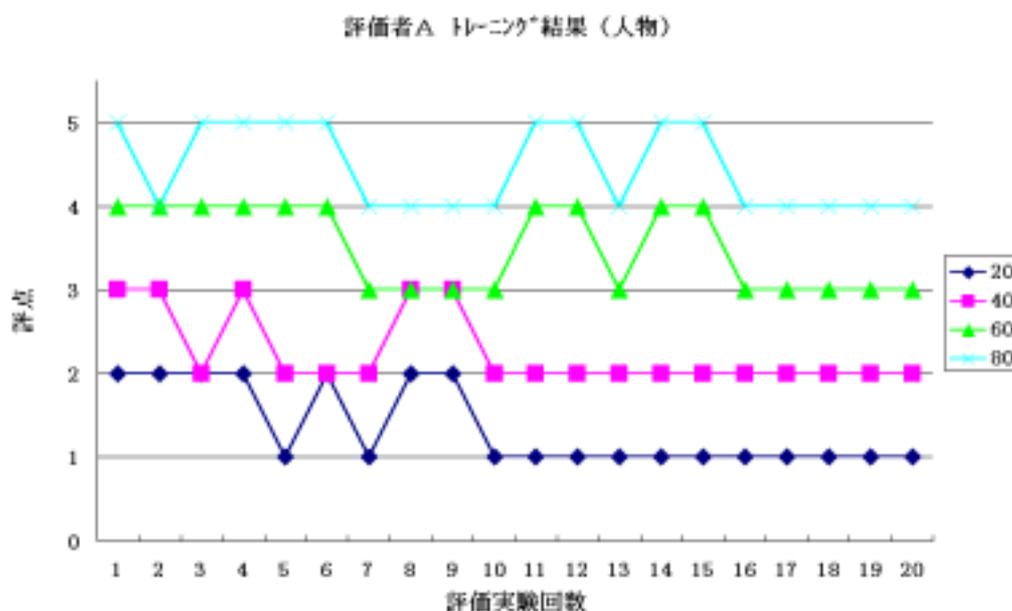


図 3.34: 評価者 A の人物画像に対するトレーニング結果 (安定した例)

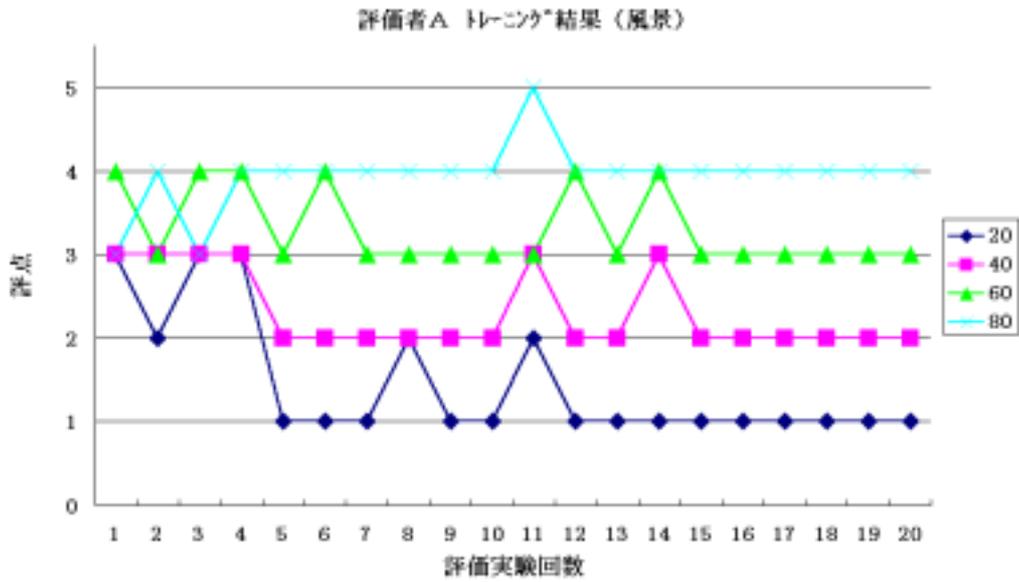


図 3.35: 評価者 A の風景画像に対するトレーニング結果 (安定した例)

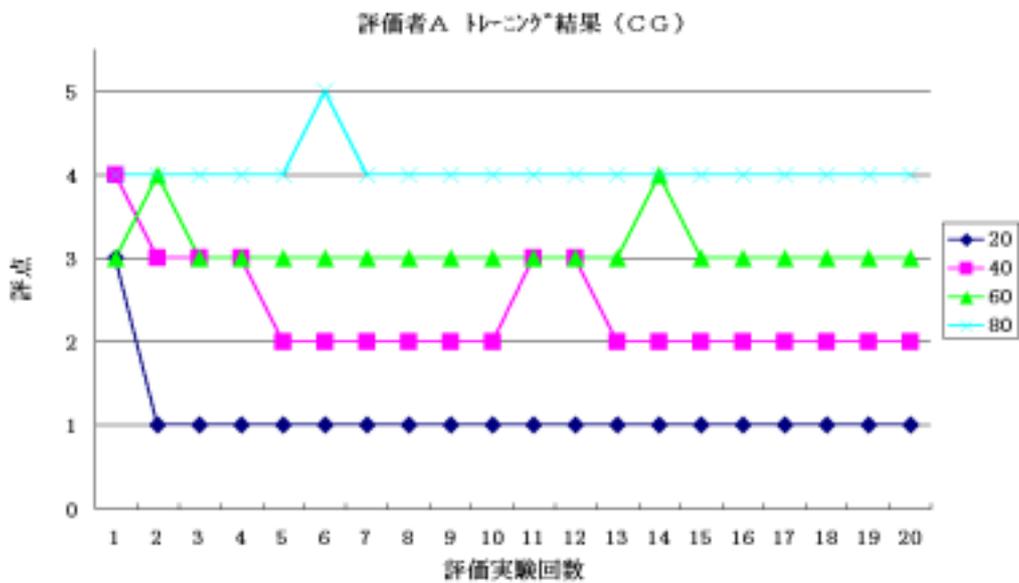


図 3.36: 評価者 A の CG に対するトレーニング結果 (安定した例)

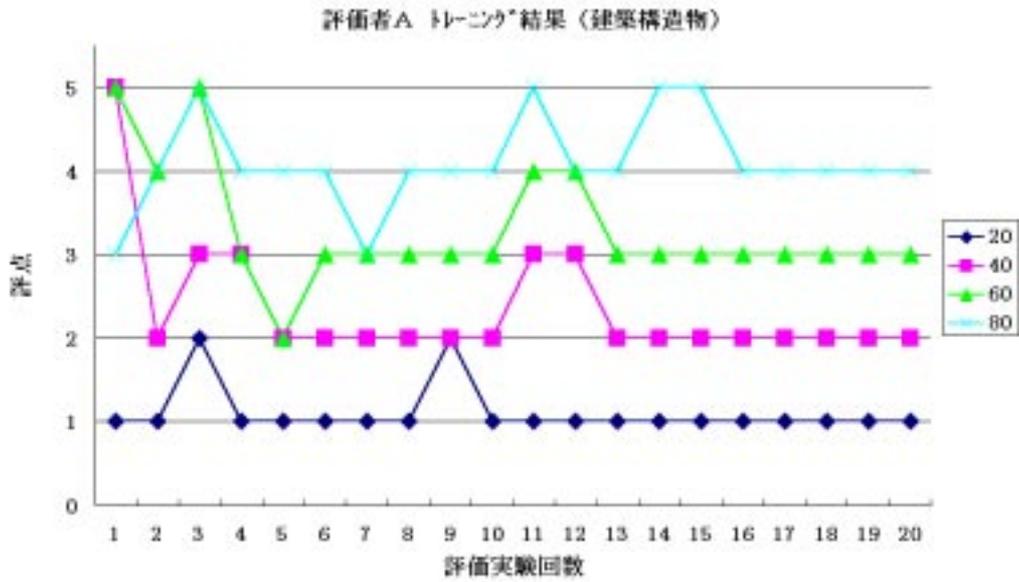


図 3.37: 評価者 A の静物画像に対するトレーニング結果 (安定した例)

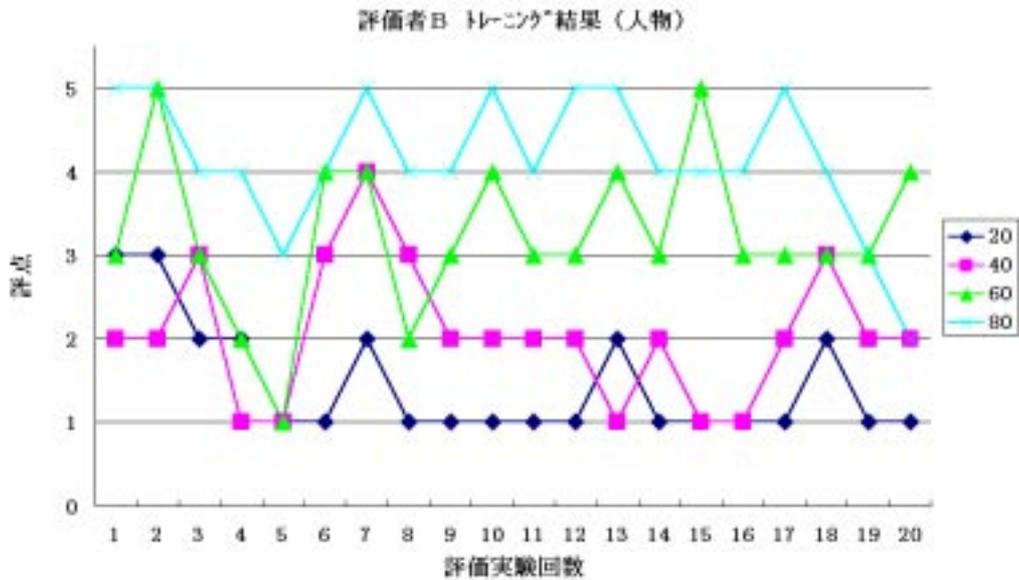


図 3.38: 評価者 B の人物画像に対するトレーニング結果 (安定しなかった例)

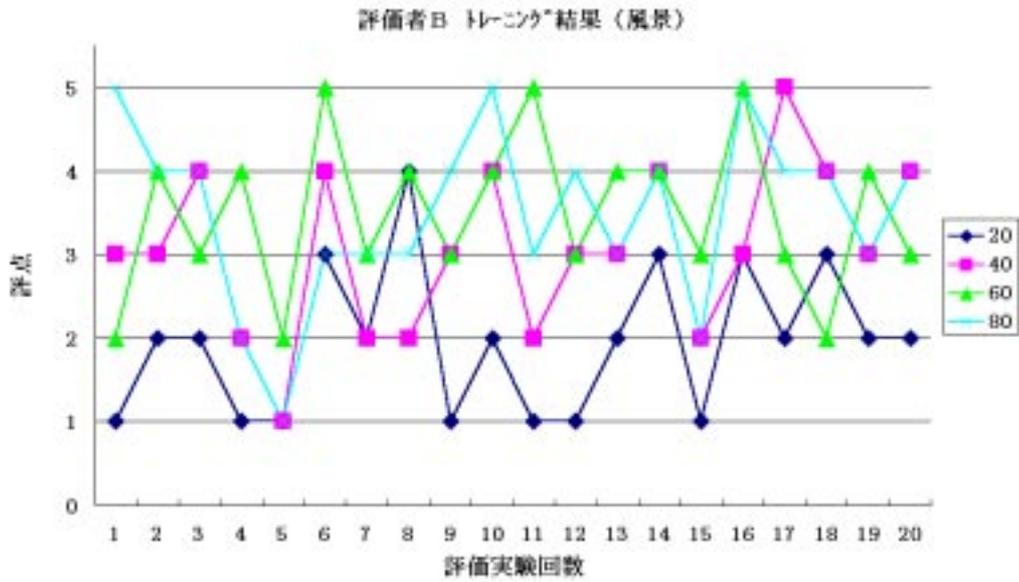


図 3.39: 評価者 B の風景画像に対するトレーニング結果 (安定しなかった例)

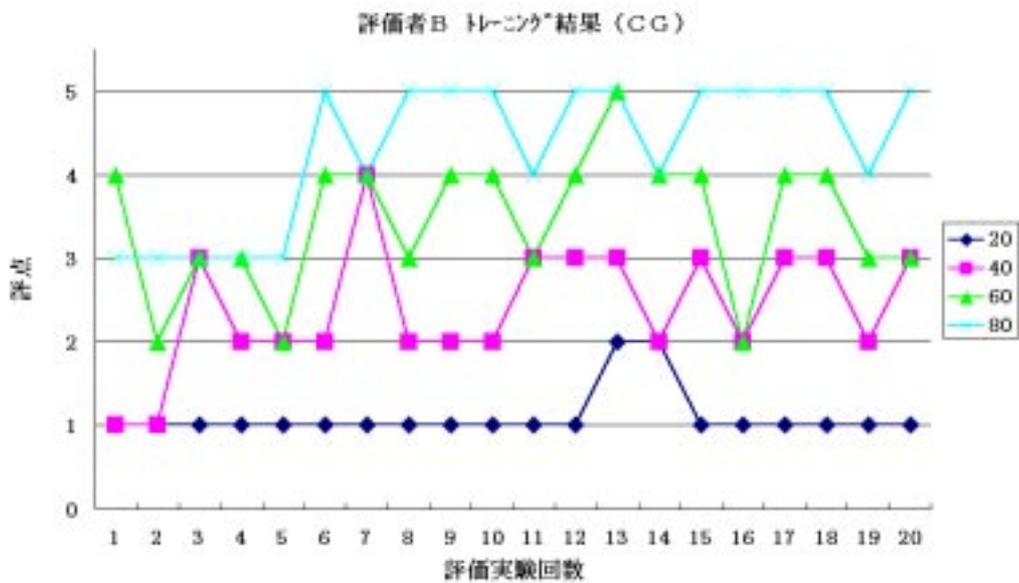


図 3.40: 評価者 B の CG に対するトレーニング結果 (安定しなかった例)

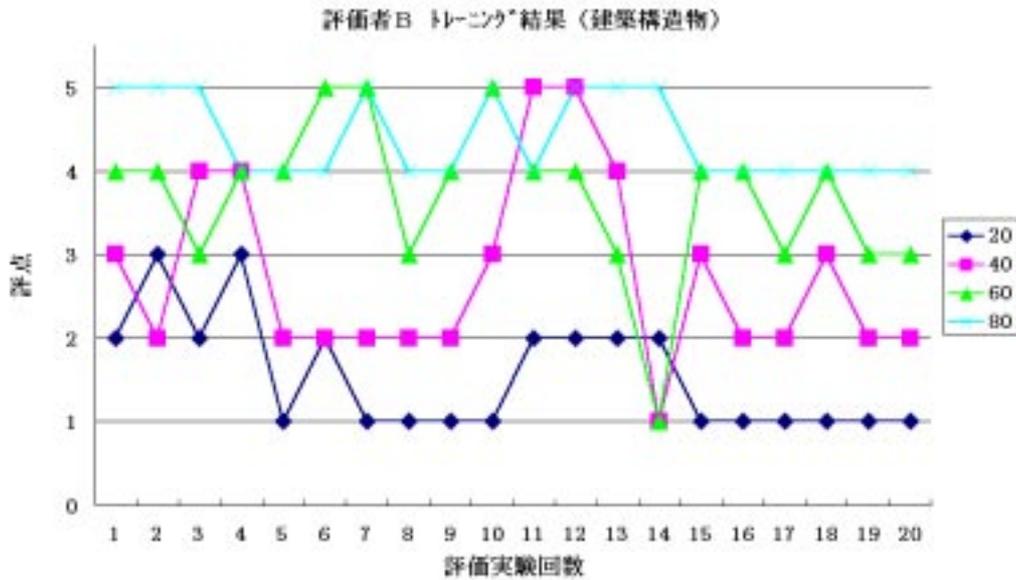


図 3.41: 評価者 B の静物画像に対するトレーニング結果 (安定しなかった例)

—全評価者のトレーニング結果の検定—

合計 10 時間のトレーニングを行なう事により平均評点は安定するという報告 [6] を基に、合計 10 時間のトレーニングが修了した後、全評価者の平均評点 (MOS) を式 (3.3) より求め、表 3.4 の条件で平均値の t 検定を行なった。

t 検定とは、仮説 $H_0: \bar{x}_1 = \bar{x}_2$ の基で統計量の実現値

$$T = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(n_1-1)u_1^2 + (n_2-1)u_2^2}{n_1+n_2-2} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}} \quad (3.1)$$

は自由度 n_1+n_2-2 の t 分布*に従う性質を利用して、実現値 t_0 が

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(n_1-1)u_1^2 + (n_2-1)u_2^2}{n_1+n_2-2} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}} > t_{n_1+n_2-2}(\alpha) \quad (3.2)$$

ならば仮説 H_0 を有意水準 α で棄却する手法である。

表 3.4: 平均値の検定条件

仮説 H_0	$MOS_{20} = MOS_{other}$
対立仮説 H_1	$MOS_{20} \neq MOS_{other}$
有意水準	5 %
検定手法	t 検定 (両側検定)
棄却域	$t_{40}(0.05) = \pm 2.02$ より $ t \geq 2.02$

ここで MOS_{20} は 20 回目の実験の平均評点を、 MOS_{other} は検定を行なう回の平均評点を指す。この結果、どの画像も 15 回目から 16 回目までは、仮説が棄却されたが、それ以後は有意差が認められず、結果が安定したと見る事が出来る。

図 3.34 のように評価者個別に評点の推移をプロットした場合も 16 回 (合計 8 時間の評価経験) 程度で結果が安定しており、文献 [6] の裏付けも行なう事が出来た。

3.4 主観評価実験結果

評価者が安定して評価を行なえる事を確認した上で、評価モデルを構築する際の主観評価実験を行なった結果を示す。

評価者数を n 、各評価者の評価値を A_i として主観評価値 (MOS) を式 (3.3) により計算する。

$$MOS = \frac{1}{n} \sum_{i=1}^n A_i \quad (3.3)$$

主観評価実験の結果を図 3.42 に示す。図より、画像内容に関わらず画質に応じた主観評価結果が得られている事が分かる。

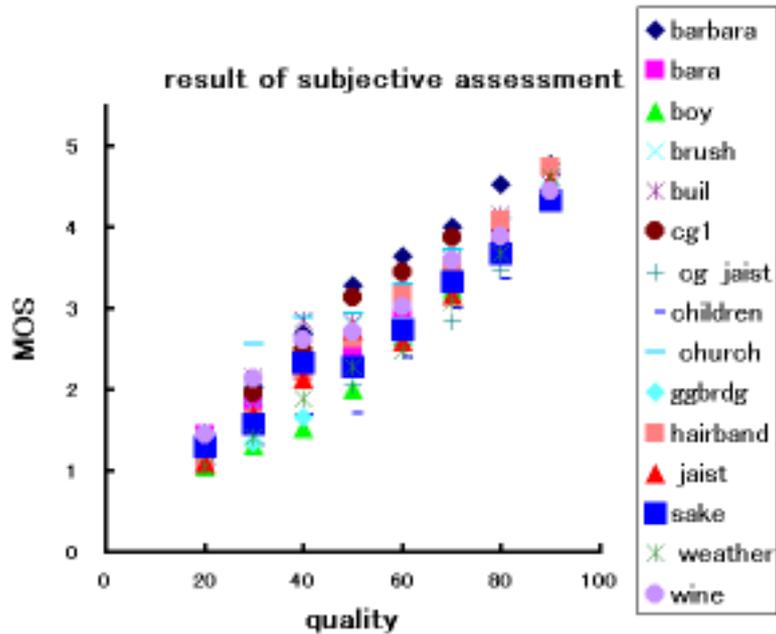


図 3.42: 主観評価実験結果

— 評点の検定 —

主観評価値の信頼性を増す為に、以下の条件に該当した評点は MOS から除外した。

1. 1 回の評価セッションの中で提示した全く同じ画像の評点の差が 2 以上あった場合
2. 1 の条件をクリアした上で画像毎の標準偏差を求め、評点が標準偏差の 3 倍を超える場合

1 の条件に関しては評価値の検定方法として勧告 [1] に記載されている方法である。また 2 の条件において標準偏差の 3 倍とした理由は、一般に品質管理において標準偏差の 3 倍を超えるデータを除外する点および、チェビシェフの定理¹ を用いて、勧告で必要とされている 15 名の評価者数を下まわらず、かつ符号化による評点の差が保たれる最低の倍数とした為である。

¹[チェビシェフの定理]

確率変数 X の値がその平均値 m から、標準偏差 σ の k 倍より小さい範囲に入る確率は $1 - \frac{1}{k^2}$ より小さくならない。 $\sigma \neq 0, k$ は任意の整数。

$$P\{|X - m| < k\sigma\} \geq 1 - \frac{1}{k^2}$$

第 4 章

種々の色空間における画質評価モデル

JPEG 符号化による誤差は、平坦部に生じるランダムなもの、輪郭周辺に生じる誤差、テクスチャパターンを形成するものに大別される。これらは、ランダムノイズや、偽輪郭の発生、輪郭の不連続や、不鮮明化色にじみやムラなどを生じる。輪郭の不連続やテクスチャパターンを作る誤差のような空間的つながりのある誤差はランダムな誤差より 10 倍以上知覚されやすい事からこの誤差を計量して評価モデルを構築する必要がある。

本研究における画質評価は、画像自身の画質を絶対評価するものではなく、原画像を基準として符号化による劣化の程度を客観的に評価する事を目的としている為、原画像と符号化再生画像の差である色差を用いて画質劣化を関数近似し評価モデルを構築する。

評価モデルを構築する対象の色空間とその関係を図 4.1 に示す。

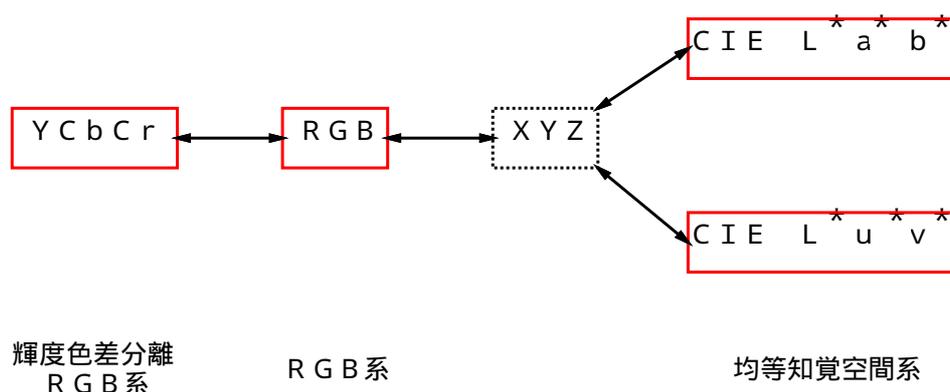


図 4.1: 評価モデルを構築する色空間とその関係

4.1 評価モデルの定義

4.1.1 システムの構成

本モデルにおいて評価値を得るプロセスを図 4.2 に示す。

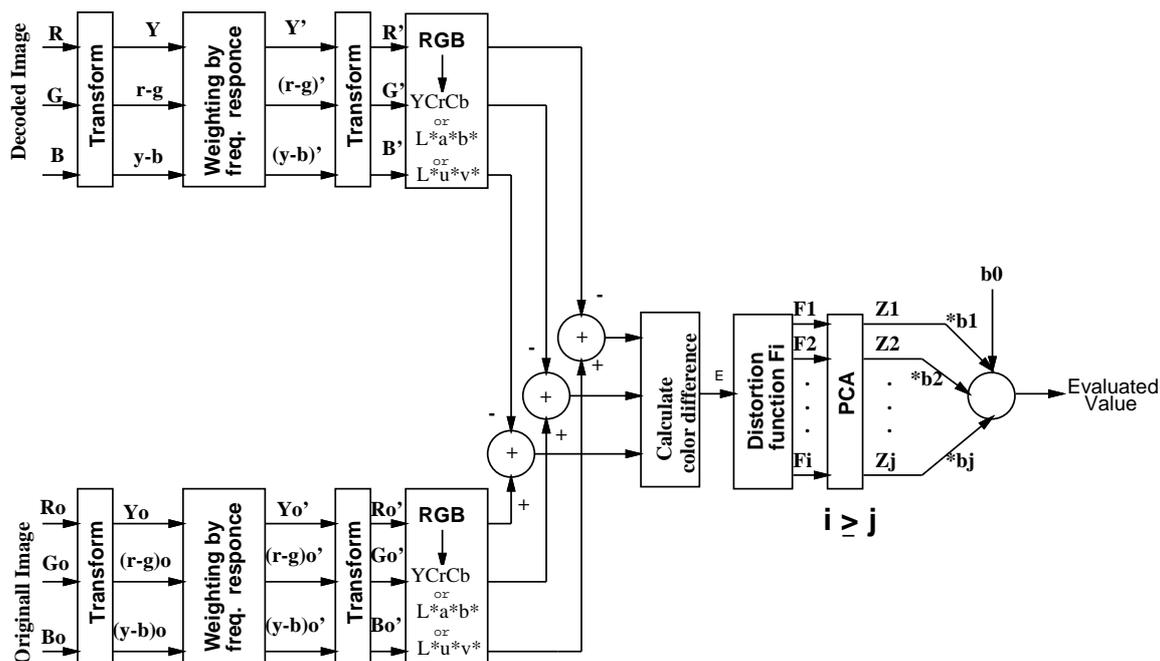


図 4.2: 評価値を得るプロセス

原画像と符号化再生画像は、最初に視覚の空間周波数特性を荷重し、評価を行なう表色系に変換する。原画像と符号化再生画像の差信号から基礎画質劣化要因を定義する。

各基礎画質劣化要因間に相関がある可能性があり、これを取り除く為に主成分分析を行ない、得られた主成分の累積寄与率を求め充分大きな累積寄与率が得られる主成分を基本画質劣化要因として抽出する。

得られた基本画質劣化要因を説明変数として主観評価値との重回帰分析を行ない偏回帰係数を求め、その線形和として式 (4.1) を用いて評価値を得る。

$$\text{Evaluated Value} = b_0 + \sum_{j=1}^J b_j Z_j \quad (4.1)$$

4.1.2 劣化の種類と画質劣化要因の関数近似

画像符号化による画質劣化の原因となる画像歪みの代表的なものを挙げると次の様になる。

1. グローバルな歪み
2. テクスチャパターンを作る歪み
3. 輪郭損傷のような画像内容に関わる局所的な歪み

これらの歪みによって生じる画質劣化の程度を関数近似する各画質劣化要因 F_i は、色差 $e(m, n)$ の関数として与える。

$$F_i = f_i\{e(m, n)\} \quad (4.2)$$

$$e(m, n) = \begin{cases} \sqrt{(\Delta R)^2 + (\Delta G)^2 + (\Delta B)^2} & \text{RGB} \\ \sqrt{(\Delta Y)^2 + (\Delta Cr)^2 + (\Delta Cb)^2} & \text{YCbCr} \\ \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2} & \text{CIE L*a*b*} \\ \sqrt{(\Delta L^*)^2 + (\Delta u^*)^2 + (\Delta v^*)^2} & \text{CIE L*u*v*} \end{cases}$$

i は画質劣化要因の種類を表し、関数 $f_i\{\cdot\}$ は F_i が画質劣化の特徴を精度良く近似すると共に、 F_i 相互の相関ができるだけ小さくなるように定義する。使用する画質劣化要因が、基本的で抜けが無ければ、実現した評価方法は広範囲な種類の画像や符号化方法に対しても適用可能な汎用性の高いものになる。本研究では、これまでの研究において良好な評価結果を得ている画質劣化要因 [4][5] を基本として、色差から 4 つの基礎的な画質劣化要因を定義した。

- F_1 : 色差の平均値
- F_2 : 符号化サブブロック間の誤差変化量
- F_3 : 誤差の自己相関係数
- F_4 : 輪郭近傍に生じる誤差

— F_1 : 色差の平均値—

従来のランダムノイズに相当する誤差を計量するものであり、この値はこれまでも符号化画像の評価に用いられていることから、研究の継続性を配慮して評価に取り入れる。

$$F_1 = E\{e(m, n)\} \quad (4.3)$$

$e(m, n)$ は色差、 (m, n) は画素の座標、 $E\{\cdot\}$ は期待値を表す。

— F_2 : 符号化サブブロック間の誤差変化量—

JPEG 符号化方式では、画像を数画素毎のサブブロックに分割し、各サブブロック毎に独立に符号化する為、サブブロックの境界が現れたり、輪郭部分が不連続になるような劣化 (ブロック歪み) が生じる。この歪みはランダムノイズより知覚されやすく、大きな妨害と感じる。このブロック間の誤差の不連続性を式 (4.4) で定式化し、 F_2 とする。

$$F_2 = \sqrt{[E\{\Delta e(m, kN)\}]^2 + [E\{\Delta e(lM, n)\}]^2} \quad (4.4)$$

$$\Delta e(m, kN) = \{e(m, kN) - e(m, kN + 1)\}^2$$

$$\Delta e(lM, n) = \{e(lM, n) - e(lM + 1, n)\}^2$$

M, N は垂直、水平のサブブロックの大きさを表し、 k, l は垂直、水平のサブブロックの位置を示す。一般的に $M = N$ とする事が多い。

— F_3 : 誤差の自己相関係数—

強い相関を持ったテクスチャはランダムなパターンより、約 10 倍知覚されやすい。誤差についても同様で、ランダムノイズより、ブロック歪みの方が知覚されやすい。この歪みに対して垂直方向、水平方向の誤差の自己相関係数を曲面とみなして、その体積に相当するものを計算し、 F_3 とする。

$$F_3 = \sqrt{R_x^2 + R_y^2} \quad (4.5)$$

$$R_x = \sum_{m=1}^M \sum_{dn=1}^{N-1} \left[\frac{R\{e(m, dn)\}}{e_{0x}} \right]^2$$
$$R_y = \sum_{dm=1}^{M-1} \sum_{n=1}^N \left[\frac{R\{e(dm, n)\}}{e_{0y}} \right]^2$$

$$R\{e(m, dn)\} = E\{e(m, n + dn) \times e(m, n)\}$$

$$R\{e(dm, n)\} = E\{e(m + dm, n) \times e(m, n)\}$$

$$e_{0x} = \text{Max}[R\{e(m, 0)\}], \quad e_{0y} = \text{Max}[R\{e(0, n)\}]$$

e_{0x}, e_{0y} はサブブロック内の各位置における誤差のパワーの最大値 (距離 0 の自己相関係数) を示す。

— F_4 : 輪郭近傍に生じる誤差—

マスキング効果により輝度変化の大きな輪郭近傍の誤差は知覚されにくい、輪郭の様な画像の構造を壊す誤りは大きく知覚される。マスキング効果を考慮した上で輪郭近傍の誤差の総和を F_4 と定義する。

$$F_4 = \sqrt{D_x^2 + D_y^2} \quad (4.6)$$

$$D_x = \sum_p^P \sum_q^Q \sum_{j=-l}^l |e(p, q - j) \times M_x(p, q - j)|$$

$$D_y = \sum_p^P \sum_q^Q \sum_{j=-l}^l |e(p - j, q) \times M_y(p - j, q)|$$

$$M_x(r, s) = \exp(-0.04V_x), \quad M_y = \exp(-0.04V_y)$$

$$V_x = \frac{1}{2} \{ |x(r, s + 1) - x(r, s - 1)| \}$$

$$V_y = \frac{1}{2} \{ |x(r + 1, s) - x(r - 1, s)| \}$$

(p, q) は輪郭点の座標位置、 P, Q は $(2l + 1)^2$ 領域内の画像の輪郭画素数の総和を示す。

$M_x(r, s), M_y(r, s)$ は点 (r, s) におけるマスキング量を表す。

4.1.3 画質劣化要因 F_i の主成分分析

各画質劣化要因 F_i は、基本的な歪みを代表するように与えられているが、互いに独立である保証はなく、相関を持っている可能性がある。この為、主成分分析を行ない冗長性を除去して基本的な画質劣化要因を抽出する。

I 個 の F_i から I 個 の主成分 Z_i が計算されるが、評価値への寄与率が低いものは近似的に省略できる。各主成分 Z_i の全体に対する累積寄与率を式 (4.7) で計算し、十分大きな (0.90 程度) の累積寄与率を有する J 個まで ($I \geq J$) の主成分 Z_i を基本劣化要因とし、これを基に評価値を計算する。

$$\text{累積寄与率} = \frac{\sum_j^J \lambda_j}{\sum_i^I \lambda_i} \quad (4.7)$$

ここで λ_j は j 番目の主成分に対する固有値を表す。

4.1.4 評価値の計算

式 (4.8) で示すように基本劣化要因 Z_j の荷重和で評価値を与える。ここで b_0, b_j は線形回帰式における偏回帰係数で、主観評価実験で既に得られている平均評点と式 (4.8) との重回帰分析によって与えられる。

$$\text{Evaluated Value} = b_0 + \sum_{j=1}^J b_j Z_j \quad (4.8)$$

4.2 RGB 色空間における画質評価

ディスプレイに表示される信号は RGB 各信号である。このため、RGB 色空間において評価モデルを構築し、画質評価を行なった。

4.2.1 画質評価結果

RGB 色空間において評価モデルを構築し、画質評価を行なった結果を示す。

—主成分—

F_i を正規化した後、共分散行列を作り主成分分析を行なう。図 4.3 に累積寄与率を示す。図より、主成分 $Z_1 \sim Z_3$ までの累積寄与率が 0.99 を越えている事から、この 3 つの主成分のみでも十分な評価精度が得られる事が分かる。

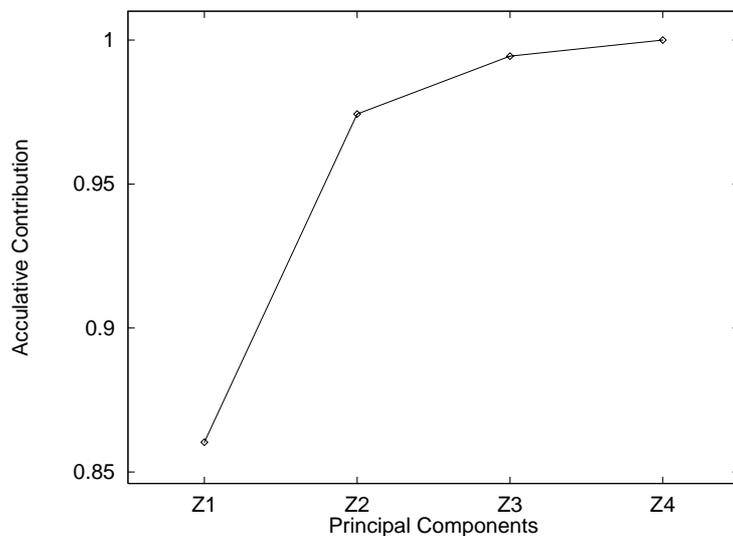


図 4.3: 各主成分 Z_i の累積寄与率

— 偏回帰係数 —

基本画質劣化要因 $Z_1 \sim Z_3$ に対する偏回帰係数を式 (4.8) と MOS との重回帰分析により計算し、結果を表 4.1 に示す。

表 4.1: 偏回帰係数 b_j (RGB)

b_0	b_1	b_2	b_3
4.936	0.165	0.0064	2.99^{-9}

— 評価精度 —

基本画質劣化要因 $Z_1 \sim Z_3$ を用いて式 (4.8) により計算した評価値と MOS との関係を求める。ここでは open data に対する評価を行なう為、全評価用画像セットから 1 枚の画像を除外して評価モデルを作成した後、除外した画像をこのモデルによって評価する。除外する画像を逐次変えて評価を行ない、その値をプロットし図 4.4 に示す。図より評価値と MOS を近似している事が確認できるが、最大誤差が 2.70 と高く、その精度はあまり高くない。

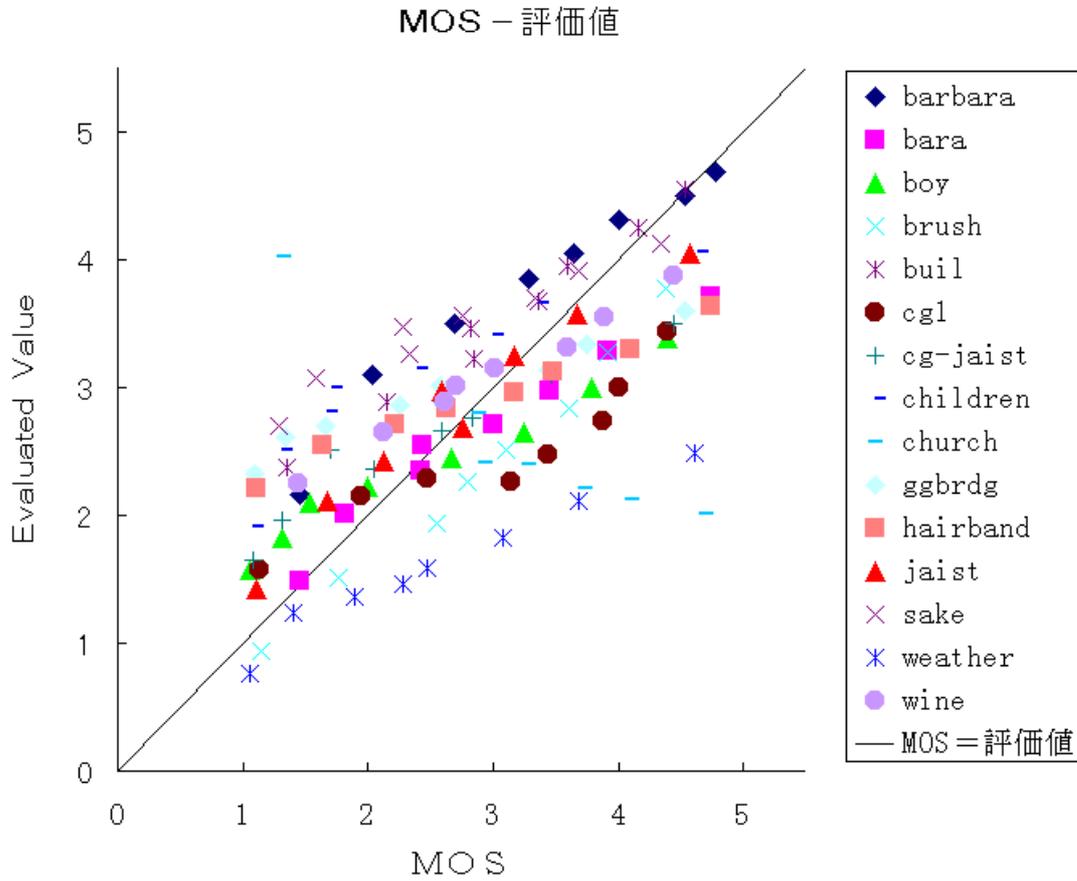


図 4.4: 評価値と MOS との関係

更に、評価値と MOS の近似度を定量的に表す為に評価値と MOS の重相関係数 R を式 (4.9) により計算する。

$$R = \frac{\delta_{\text{evaluated value}}}{\delta_{\text{mos}}} \quad (4.9)$$

ここで、 $\delta_{\text{evaluated value}}$ と δ_{mos} は各々の評価値と MOS の標準偏差を表わす。

本手法による R は、0.67 で、高い精度を得る事が出来ない。RGB 色空間における評価モデルの構築は、あまり有効ではない事が確かめられた。

4.2.2 必須画質劣化要因

本手法では画質の劣化を基本画質劣化要因の線形和でモデル化しており、画質劣化要因 F_i は基本的で抜けがない事が望ましく、かつ不要な F_i は除去すべきである。 $F_1 \sim F_4$ の中で必要不可欠な画質劣化要因を検討する。

各 F_i の組合せを変えて評価値を計算し、 F_i の個数に対する補正を行なった自由度調整済み重相関係数 R^{**} を計算する。重相関係数 R と R^{**} の間には式 (4.10) の関係がある。

$$R^{**} = \sqrt{\frac{R^2(n-1)(n+p+1) - 2np}{(n+1)(n-1-p)}} \quad (4.10)$$

ここで n はデータの種類、 p は F_i の個数である。 R^{**} が大きいほど MOS との近似度が良い。式 (4.9) では p が増加すれば R は単調に増加するが、 R^{**} は単調に増加しない。無意味な F_i を追加して R が増加しないと R^{**} は減少する特性がある為、 R^{**} を使用する事により F_i の有意性を評価する事ができる。

表 4.2 に各組合せを総当たりで計算した結果を示す。できるだけ少ない F_i で R^{**} が高い方が望ましく、RGB 色空間における評価モデルでは、全ての要因を用いた場合が評価精度が高くなる。しかし、 F_4 を除外する事で評価精度が極端に減少する事から、 F_4 が必須画質劣化要因である事が言える。

表 4.2: 評価値-MOS 間の自由度調整済み重相関係数 (RGB)

組合せ i (F_i)	R^*
1	0.302
2	0.209
3	0.007
4	0.592
1 2	0.319
1 3	0.403
1 4	0.653
2 3	0.407
2 4	0.629
3 4	0.657
1 2 3	0.421
1 2 4	0.648
1 3 4	0.656
2 3 4	0.663
1 2 3 4	0.701 *

4.3 YCbCr 色空間における画質評価

RGB 系で表現されたデジタル画像データを輝度色差分離系の YCbCr 色空間に変換した場合、色差軸 (C_b, C_r) ではヒストグラムが狭い分布特性となり、効率的な情報圧縮の可能性があるとされている [16]。また、jpeg においても RGB 色空間から YCbCr 色空間に変換して符号化を行なっていることから、YCbCr 色空間における画質評価モデルを構築し画質評価を行なった。なお RGB 色空間から YCbCr 色空間への変換は式 (4.11) を用いた。

$$\begin{aligned}
 Y &= 0.29900 \times R + 0.58700 \times G + 0.11400 \times B \\
 C_b &= -0.16874 \times R - 0.33126 \times G + 0.50000 \times B \\
 C_r &= 0.50000 \times R - 0.41869 \times G - 0.08131 \times B
 \end{aligned}
 \tag{4.11}$$

4.3.1 画質評価結果

YCbCr 色空間における画質評価結果を以下に示す。

—主成分—

E_i を正規化した後、共分散行列を作り主成分分析を行なう。図 4. 5に累積寄与率を示す。図より、主成分 $Z_1 \sim Z_3$ までの累積寄与率が 0.99 を越えている事から、この3つの主成分のみでも十分な評価精度が得られる事が分かる。

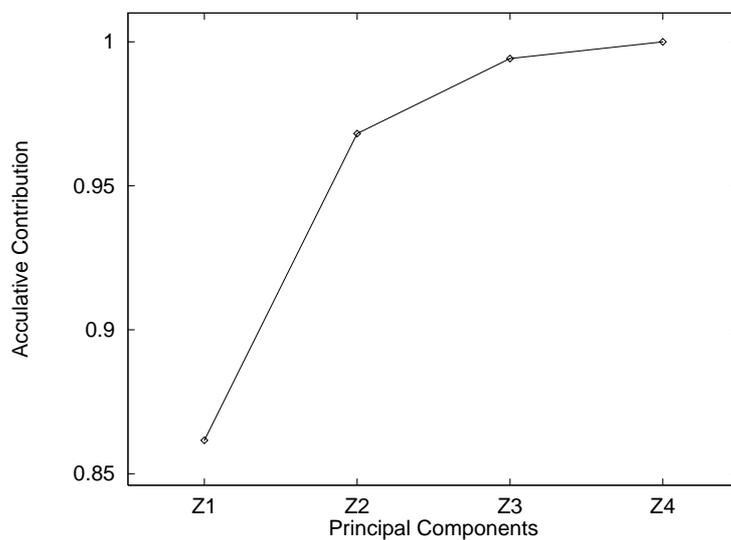


図 4. 5: 各主成分 Z_i の累積寄与率

—偏回帰係数—

基本画質劣化要因 $Z_1 \sim Z_3$ に対する偏回帰係数を式 (4.8) と MOS との重回帰分析により計算し、結果を表 4. 3に示す。

表 4.3: 偏回帰係数 b_j (YCbCr)

b_0	b_1	b_2	b_3
5.15 7	0.3 5 7	-0.0 1 0	$4. 1^{-2}$

—評価精度—

基本画質劣化要因 $Z_1 \sim Z_3$ を用いて式 (4.8) により計算した評価値と MOS との関係を求める。ここでは open data に対する評価を行なう為、全評価用画像セットから 1 枚の画像を除外して評価モデルを作成した後、除外した画像をこのモデルによって評価する。除外する画像を逐次変えて評価を行ない、その値をプロットし図 4.6 に示す。図より、評価値が MOS を良好に近似している事が確認できる。

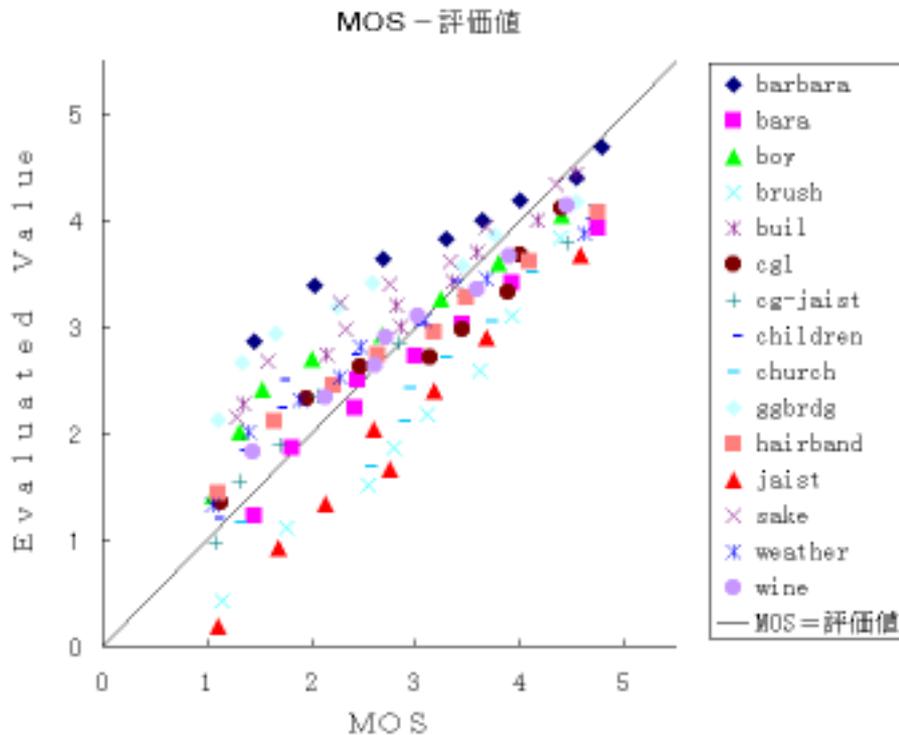


図 4.6: 評価値と MOS との関係

更に、評価値と MOS の近似度を定量的に表す為に評価値と MOS の重相関係数 R を式 (4.9) により計算する。

本手法による R は、0.844 となった。比較を行なう為、修正 Munsell 色空間において符号化誤差を Godlove 色差として計算し、画質の劣化要因を色差の関数として与え、多変量解析を行ない画質モデルを構築した。このモデルの R は 0.848 となり、YCbCr 色空間における画質評価モデルの構築が有効である事が確かめられた。

4.3.2 必須画質劣化要因

4.2. 節と同様に、表 4.4 に各画質劣化要因の組合せを変えて R^{**} を計算した結果を示す。 F_1 と F_4 の組合せが最も良く、YCbCr 色空間において評価モデルを構築する際の必須画質劣化要因も RGB 色空間の時と同様 F_4 である事が分かる。

表 4.4: 評価値-MOS間の自由度調整済み重相関係数 (YCbCr)

組合せ $i (F_i)$	R^{**}
1	0.522
2	0.428
3	0.314
4	0.797
1 2	0.530
1 3	0.529
1 4	0.848 *
2 3	0.463
2 4	0.831
3 4	0.819
1 2 3	0.520
1 2 4	0.845
3 4	0.845
2 3 4	0.829
1 2 3 4	0.843

4.4 CIE L*a*b*色空間における画質評価

CIE L*a*b*色空間は、輝度 L* と色相と彩度からなる色知覚の属性を示す a*, b* で表現される色空間で、Munsell 色空間を近似する均等色空間である [13][14]。CIE L*a*b* 色空間への変換は式 4.13 を用いた。

$$\begin{aligned} L^* &= \begin{cases} 116(Y/Y_n)^{\frac{1}{3}} - 16 \\ 116\{7.767(Y/Y_n)^{\frac{1}{3}}\} - 16 \end{cases} & (Y/Y_n \leq 0.008856) \\ a^* &= 500\{(X/X_n)^{\frac{1}{3}} - (Y/Y_n)^{\frac{1}{3}}\} \\ b^* &= 500\{(Y/Y_n)^{\frac{1}{3}} - (Z/Z_n)^{\frac{1}{3}}\} \end{aligned} \quad (4.13)$$

ここで

$$\begin{aligned} X &= (0.608 \times R + 0.174 \times G + 0.200 \times B)/255.0 \\ Y &= (0.299 \times R + 0.587 \times G + 0.114 \times B)/255.0 \\ Z &= (0.066 \times G + 1.112 \times B)/255.0 \\ X_n &= X/1.020, \quad Y_n = Y/1.00, \quad Z_n = Z/0.847 \end{aligned} \quad (4.14)$$

4.4.1 画質評価結果

構築した評価モデルを用い、画質評価実験を行なった。評価値と MOS の一致度と主成分の解析結果を示す。

—主成分—

F_i を正規化した後、共分散行列を作り主成分分析を行なう。図 4.7 に累積寄与率を示す。図より、主成分 $Z_1 \sim Z_3$ までの累積寄与率が 0.99 を越えている事から、この 3 つの主成分のみでも十分な評価精度が得られる事が分かる。

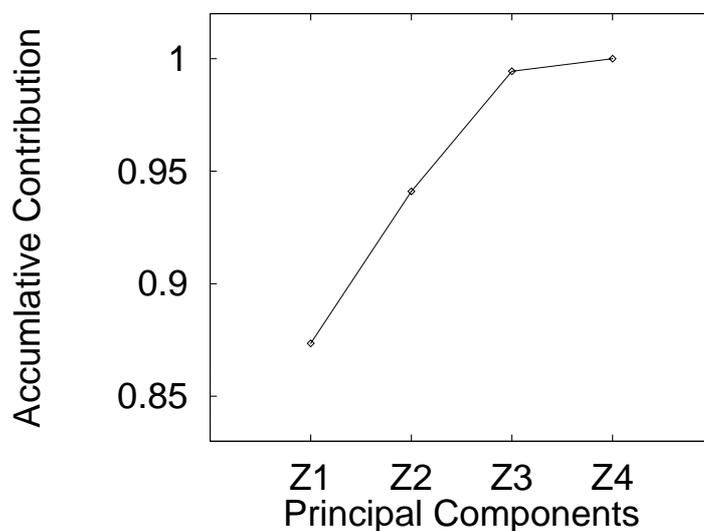


図 4.7: 各主成分 Z_i の累積寄与率

— 偏回帰係数 —

基本画質劣化要因 $Z_1 \sim Z_3$ に対する偏回帰係数を式 (4.8) と MOS との重回帰分析により計算し、結果を表 4.5 に示す。

表 4.5: 偏回帰係数 b_j (CIE $L^*a^*b^*$)

b_0	b_1	b_2	b_3
6.021	-0.136	0.019	$1.43 \cdot 10^{-7}$

— 評価精度 —

基本画質劣化要因 $Z_1 \sim Z_3$ を用いて式 (4.8) により計算した評価値と MOS との関係を求める。ここでは open data に対する評価を行なう為、全評価用画像セットから 1 枚の画像を除外して評価モデルを作成した後、除外した画像をこのモデルによって評価する。除外する画像を逐次変えて評価を行ない、その値をプロットし図 4.8 に示す。図より、評価値が MOS を良好に近似している事が確認できる。

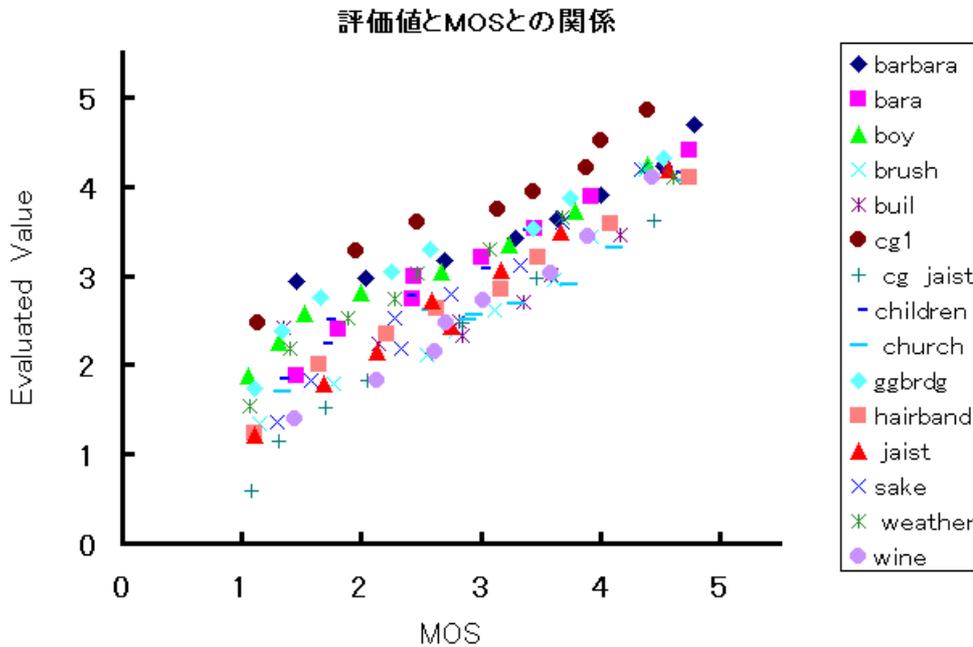


図 4.8: 評価値と MOS との関係

更に、評価値と MOS の近似度を定量的に表す為に評価値と MOS の重相関係数 R を式 (4.9) により計算する。

本手法による R は、0.859 となった。WMS E と同等の F_1 のみを使用した場合の R は 0.62、修正 Munsell 色空間において符号化誤差を Godlove 色差として計算し、画質の劣化要因を色差の関数として与え、多変量解析を行なって作成した線形回帰モデルの R は 0.848 となり、CIE $L^*a^*b^*$ 色空間における画質評価モデルの構築が有効である事が確かめられた。

4.4.2 汎用性

図 4.8 においては、全評価用画像セットから 1 枚の画像を除外してモデルを作成し、除外した画像の評価を行なっているが、この手法では汎用性の確認はできない。そこで、モデルを作成する際に使用する画像種類を限定 (15 種類の中から 1 種類を除外) したモデル

または符号化効率を限定 (8 種類の中から 1 種類を除外) したモデルを作成し、モデルを作成する際に除外した画像を評価する事によって汎用性の確認を行なう。

—画像内容に関する汎用性—

全評価画像種から 1 種類を除外したモデルを 15 種類作成し、得られた偏回帰係数を使用して除外した画像の評価を行なう。その結果を図 4.9 に示す。この結果、低画質時の誤差が最大 1.7 と大きいものの、平均誤差約 0.5 と全体的に主観評価値を良好に近似している事が分かり、画像内容に関しての汎用性が確保されていると言える。

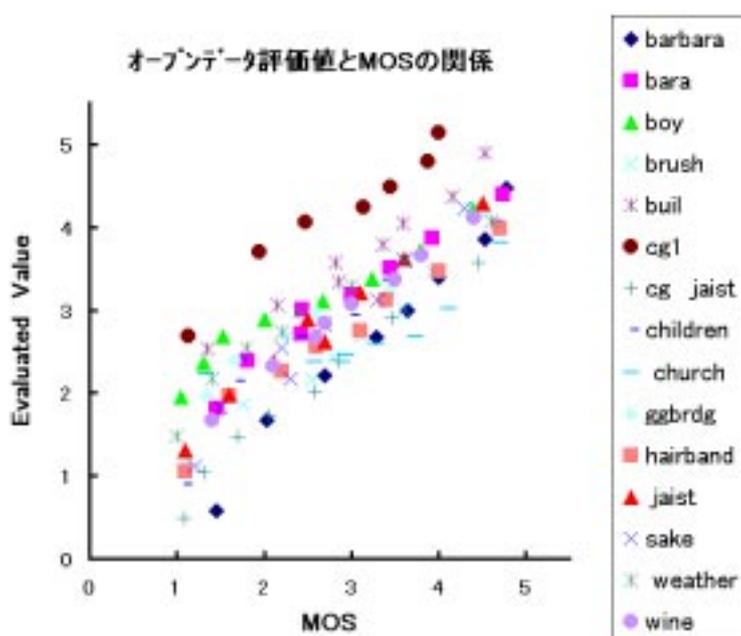


図 4.9: オープンデータ評価値と MOS との関係 (画像内容)

—符号化効率に関する汎用性—

15 枚の原画像に対して 8 種類の符号化効率で符号化した画像を用いてモデルを作成し評価を行なった。符号化効率に関する汎用性を確認する為に特定の伝送レート 1 種を除いたモデル 5 種類を作成し、得られた偏回帰係数を使用して除外した伝送レートの画像の評価を行なう。その結果を図 4.10 に示す。図より、低画質時の評価精度が悪く一つのモデルで全ての画像範囲の評価を行なう事が困難である事が確認できる。この理由としては、

低画質時に大きな誤差として生じる偽輪郭のような歪みを現在の F_4 では計量できないのだと考える。この問題を解決する為に低画質用、高画質用と画質クラスを考慮した評価モデルの作成の必要である。

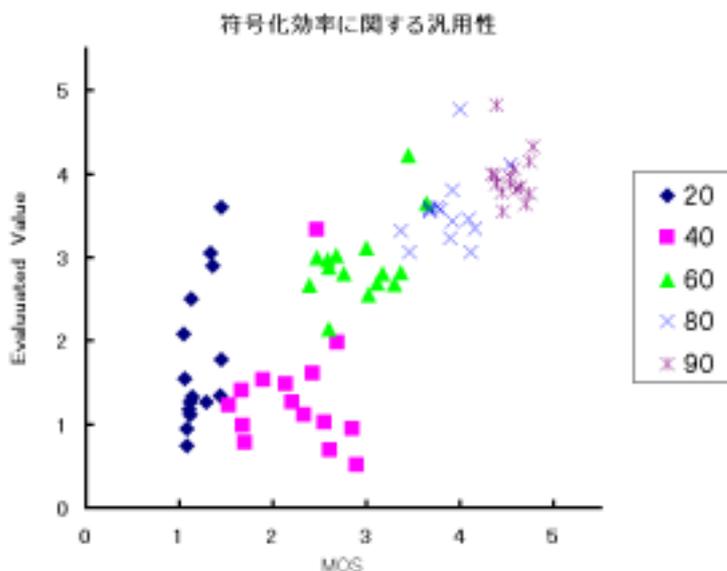


図 4.10: オープンデータセット評価値と MOS との関係 (符号化効率)

—モノクロ画像に対する汎用性—

構築した評価モデルはカラー画像を対象としていたがモノクロ画像をカラー画像の一部として考えると、モノクロ画像も同じモデルで評価できる事が望ましい。そこで、作成した評価モデルを用いてモノクロ画像の評価を行ない、モノクロ画像に対する汎用性の確認を行なう。モノクロ画像に対する主観評価はカラー画像の評価の時と同様で、トレーニングを積んだ 20 名の評価者に対して行なう。評価に使用したモノクロ画像を表 4. 6に、画像を図 4. 11から 4. 15に示す。

これらの原画像をカラー画像の時と同様に jpeg で 8 種類に符号化し、符号化再生画像として用いた。

表 4.6: モノクロ評価用画像

カテゴリ	原画像名
人物	barbar a hair band weather
建築構造物	bui l chur ch



図 4. 11: bar bar a-mono



図 4. 12: hai r band- mono



図 4. 13: weat her- mono



図 4. 14: bui l- mono



図 4. 15: chur ch- mono

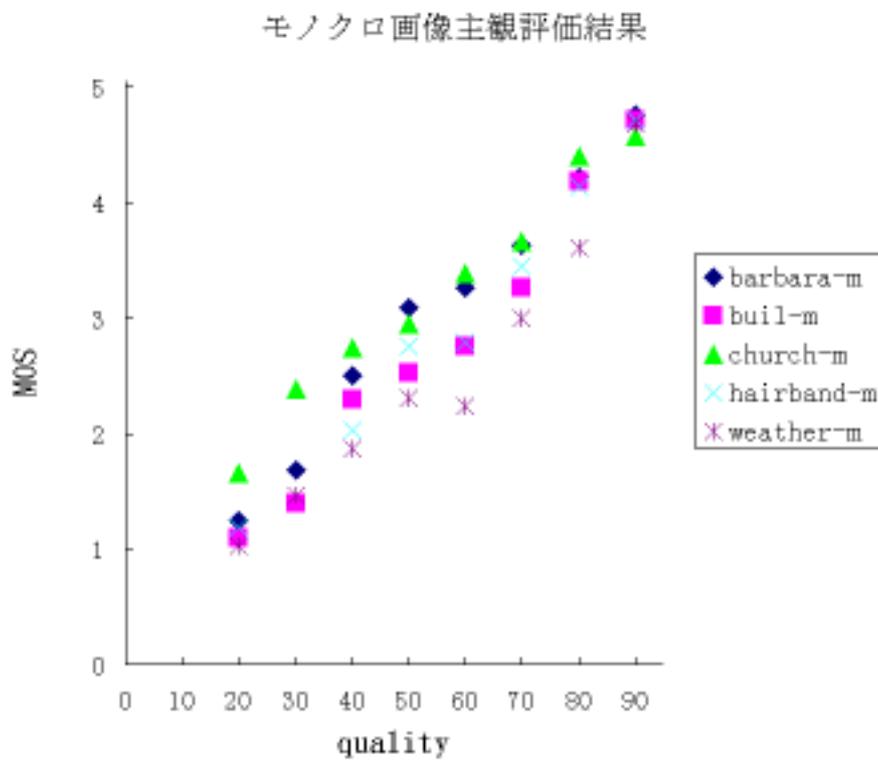


図 4.16: モノクロ画像主観評価結果

主観評価実験結果を図 4.16に示す。ここで得られた主観評価値を用いて、以下の1から3の手法でモノクロ画像の汎用性を確認した。

[手法 1: カラー画像対象の偏回帰係数を使用して評価]

カラー画像を対象とした評価モデルで得られている偏回帰係数(表 4.5)を使用してモノクロ画像の評価を行なう。その結果を図 4.17に示す。

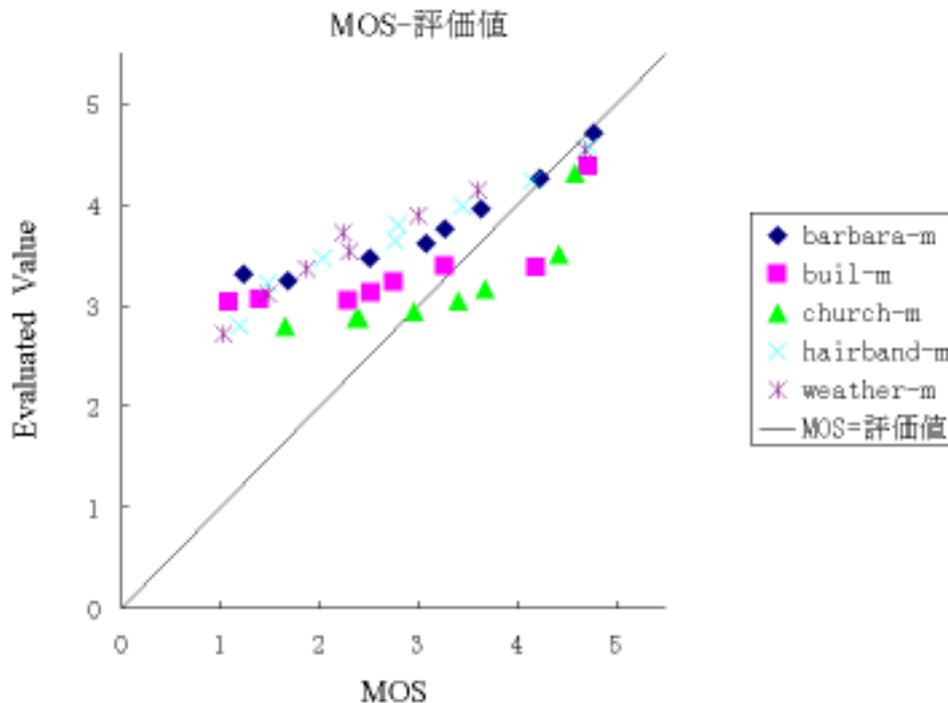


図 4.17: MOS-評価値 (手法 1)

この図から、カラー画像を対象とした評価モデルで得られた偏回帰係数を使用し、モノクロ画像の評価を行なえない事が分かる。この時点ではモノクロ画像の評価はモノクロ画像評価用のモデルを使用する必要がある。

[手法 2:モノクロ画像を $L^*a^*b^*$ 色空間に変換して評価]

手法 1 からモノクロ画像の評価にはカラー画像の評価モデルと別に扱う必要があると考察されたが、色に関する成分がない場合はどのように評価されるかを確認する為、モノクロ画像のみを用いて、 $L^*a^*b^*$ 色空間における評価モデルを作成し直し、画質評価を行なう。その結果を図 4. 18に示す。

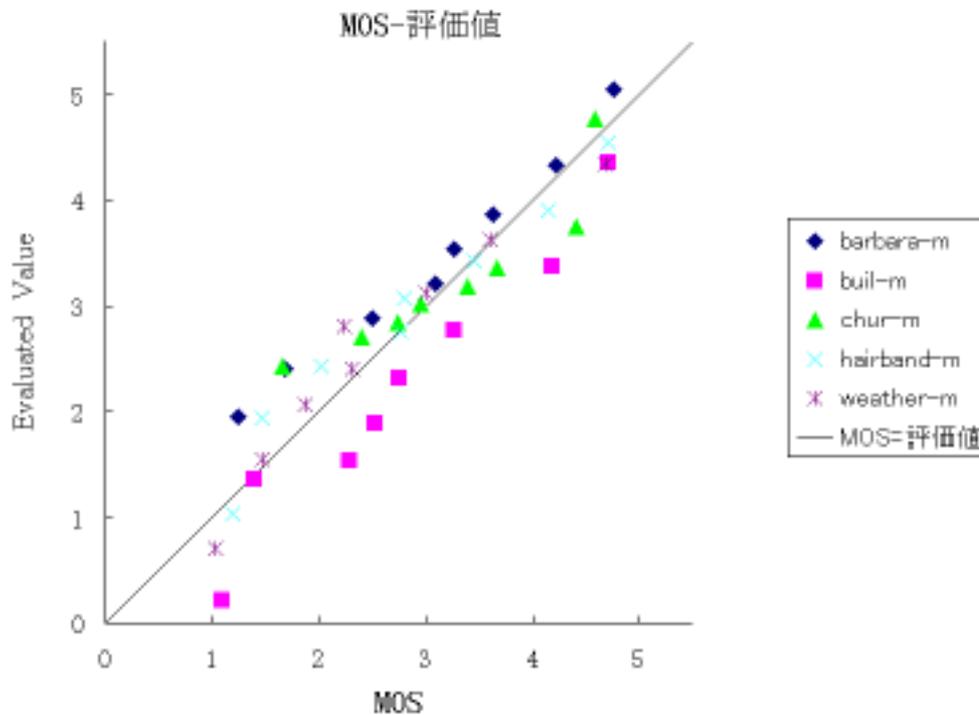


図 4.18: MOS-評価値 (手法 2) オープンデータ

モノクロ画像のみを用いて評価モデルを作成した際の重相関係数は 0.92 と高い評価精度を得ていることから、このモデルをモノクロ画像に適用する出来る可能性があることが分かる。

[手法 3:モノクロ画像とカラー画像を用いて評価モデルを構築]

モノクロ画像を用いて、 $L^*a^*b^*$ 色空間における評価モデルを作成した結果より、評価モデルをモノクロ画像に適用できる事が確認できた。この為、モノクロ画像 5 種類とカラー画像 15 種類の合計 20 種類の画像を使用してモデルを作成し画質評価を行なう。その結果を図 4. 19に示す。図 4. 20にカラー画像のみの結果、図 4. 21にモノクロ画像の結果のみをプロットした。

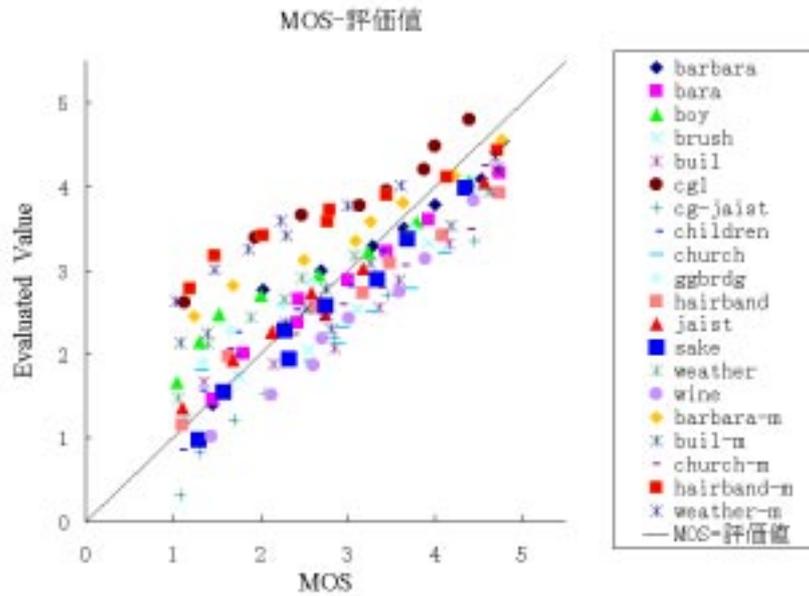


図 4.19: MOS-評価値 (手法 3) オープンデータ

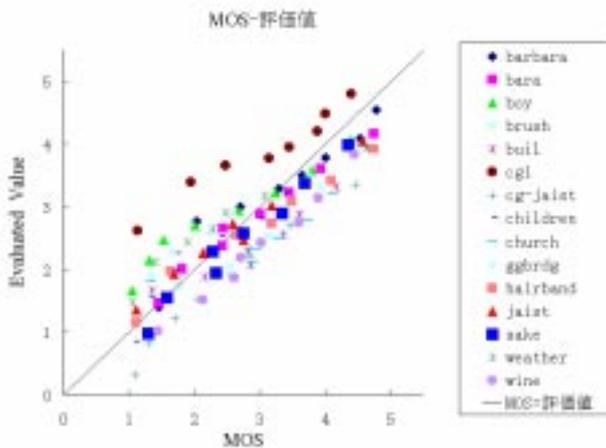


図 4.20: MOS- 評価値 (手法 3)(mono)

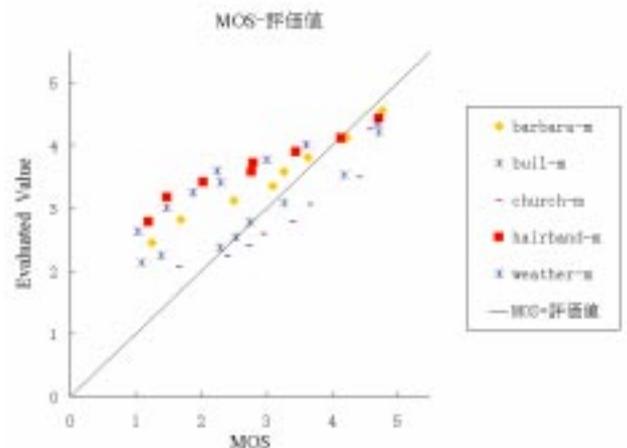


図 4.21: MOS- 評価値 (手法 3)(color)

カラー画像とモノクロ画像を用いて評価モデルを作成した際の重相関係数は0.81とカラー画像のみで評価モデルを作成した際と同等の精度(カラー画像で作成したモデルの重相関係数0.86)を得ている。しかし、図4.21からも分かるようにモノクロ画像の低画質時のばらつきが大きくなる傾向が見られる。

cjpeg において、通常使用される quality は 50 以上 (デフォルト 75) である事から主観評価値 3 以上を良好に評価できる事が望ましい。また、モノクロ画像を含めた画像を用いて評価モデルを作成する事により、モノクロ画像も同一のモデルで評価が可能となると考えられるが、全ての画質を評価できるモデルの構築を考えると、モノクロ画像の評価はカラーの画質評価とは別に扱う方が望ましい。

4.4.3 必須画質劣化要因

本手法では画質の劣化を基本画質劣化要因の線形和でモデル化しており、画質劣化要因 F_i は基本的で抜けがない事が望ましく、かつ不要な F_i は除去すべきである。 $F_1 \sim F_4$ の中で必要不可欠な画質劣化要因を検討する。表 4.7 に各組合せを総当たりで計算した結果を示す。できるだけ少ない F_i で R^{**} が高い方が望ましい為、*マークをした F_3 と F_4 の組合せが最も良い。この事から、主観評価においては、ランダムノイズよりも局所的な歪みの方が評点の決定に重要な要因となっている事が考えられる。

表 4.7: 評価値-MOS 間の自由度調整済み重相関係数 (CIE L*a*b*)

組合せ i (F_i)	R**
1	0.621
2	0.477
3	0.519
4	0.863
1 2	0.681
1 3	0.613
1 4	0.863
2 3	0.509
2 4	0.869
3 4	0.872 *
1 2 3	0.707
1 2 4	0.869
1 3 4	0.870
2 3 4	0.870
1 2 3 4	0.869

4.5 CIE L*u*v*色空間における画質評価

CIE L*a*b*色空間も、輝度 L^* と色相と彩度からなる色知覚の属性を示す u^*, v^* で表現される色空間であり、明るさ一定の条件のもとで jnd(丁度可知差異) の均等性について検討されている [13][14]。CIE L*u*v*色空間への変換は式 (4.15) を用いた。

$$\begin{aligned}
 L^* &= \begin{cases} 116(Y/Y_n)^{\frac{1}{3}} - 16 & (Y/Y_n > 0.008856) \\ 116\{7.767(Y/Y_n)^{\frac{1}{3}}\} - 16 & (Y/Y_n \leq 0.008856) \end{cases} \quad (4.15) \\
 u^* &= 13L^*(u' - u'_n), \quad u' = 4X/(X + 15Y + 3Z), \quad u'_n = 4X_n/(X_n + 15Y_n + 3Z_n) \\
 v^* &= 13L^*(v' - v'_n), \quad v' = 9Y/(X + 15Y + 3Z), \quad v'_n = 9Y_n/(X_n + 15Y_n + 3Z_n)
 \end{aligned}$$

ここで X, Y, Z, X_n, Y_n, Z_n は式 (4.14) を参照。

4.5.1 画質評価結果

構築した評価モデルを用い、画質評価実験を行なった。評価値と MOS の一致度と主成分の解析結果を示す。

—主成分—

F_i を正規化した後、共分散行列を作り主成分分析を行なう。図 4.22に累積寄与率を示す。図より、主成分 $Z_1 \sim Z_3$ までの累積寄与率が 0.99 を越えている事から、この 3 つの主成分のみでも十分な評価精度が得られる事が分かる。

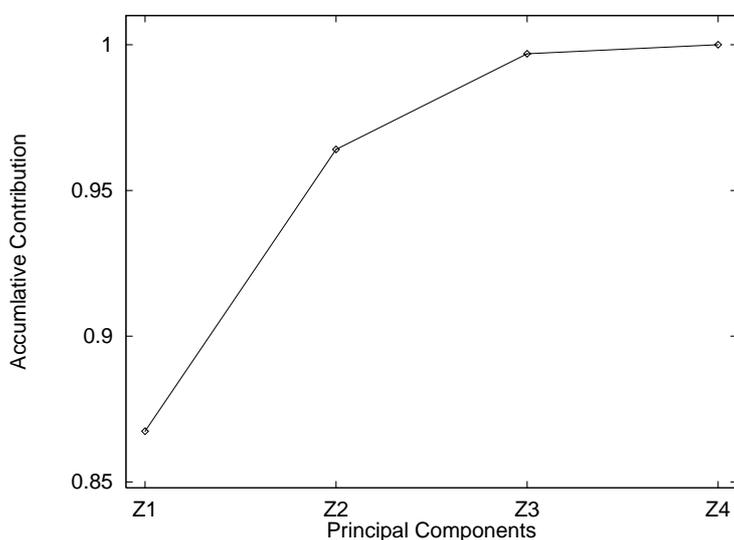


図 4. 22: 各主成分 Z_i の累積寄与率

—偏回帰係数—

基本画質劣化要因 $Z_1 \sim Z_3$ に対する偏回帰係数を式 (4. 8)と MDS との重回帰分析により計算した、表 4. 8に示す。

表 4. 8: 偏回帰係数 b_j (CIE $L^*u^*v^*$)

b_0	b_1	b_2	b_3
5.491	0.527	-0.054	$3.93 \cdot 10^{-6}$

基本画質劣化要因 $Z_1 \sim Z_3$ を用いて式 (4.8) により計算した評価値と MOS との関係を求める。ここでは open data に対する評価を行なう為、全評価用画像セットから 1 枚の画像を除外して評価モデルを作成した後、このモデルによって除外した画像を評価する。除外する画像を逐次変えて評価を行ない、その値をプロットし図 4.23 に示す。図より、評価値が MOS を良好に近似している事が確認できる。

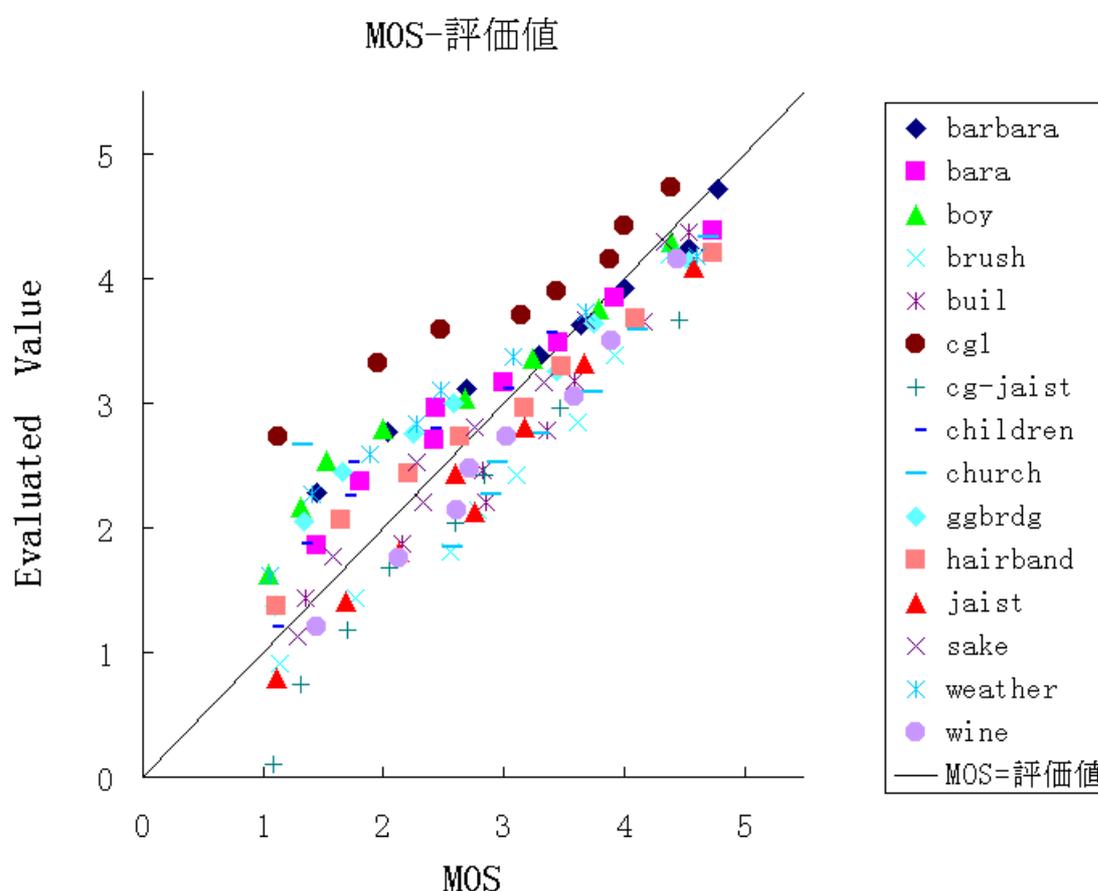


図 4.23: 評価値と MOS との関係

更に、評価値と MOS の近似度を定量的に表す為には評価値と MOS の重相関係数 R を式 (4.9) により計算する。

本手法による R は 0.886 となった。WMS E と同等の F_1 のみを使用した場合の R は 0.537

修正 Munsell 色空間において符号化誤差を Godlove 色差として計算し、画質の劣化要因を色差の関数として与え、多変量解析を行なって作成した線形回帰モデルの R は 0.848 となり、CIE $L^*u^*v^*$ 色空間における評価モデルの構築が有効である事が確かめられた。

4.5.2 汎用性

図 4.23 においては、全評価用画像セットから 1 枚の画像を除外してモデルを作成し、このモデルを用いて除外した画像の評価を行なっているが、この手法では汎用性の確認はできない。そこで、モデルを作成する際に使用する画像種類を限定 (15 種類の中から 1 種類を除外) したモデルまたは符号化効率を限定 (8 種類の中から 1 種類を除外) したモデルを作成し、モデルを作成する際に除外した画像を評価する事によって汎用性の確認を行なう。

—画像内容に関する汎用性—

全評価画像種から 1 種類を除外したモデルを 15 種類作成し、得られた偏回帰係数を使用して除外した画像の評価を行なった結果を図 4.24 に示す。この結果、低画質時の誤差が最大 1.60 と大きいものの、平均誤差約 0.5 と全体的に主観評価値を良好に近似している事が分かり、画像内容に関しての汎用性が確保されていると言える。

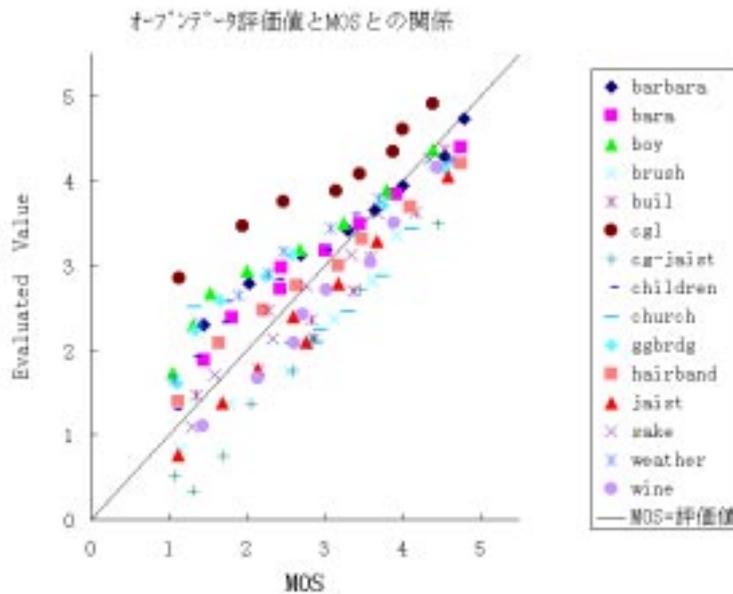


図 4.24: オープンデータ評価値と MOS との関係 (画像内容)

—符号化効率に関する汎用性—

15 枚の原画像に対して 8 種類の符号化効率で符号化した画像を用いてモデルを作成し評価を行なった。符号化効率に関する汎用性を確認する為に特定の伝送レート 1 種を除いたモデル 5 種類を作成し、得られた偏回帰係数を使用して除外した伝送レートの画像の評価を行なう。その結果を図 4.25 に示す。図より、低画質時の評価精度が悪く一つのモデルで全ての画像範囲の評価を行なう事が困難である事が確認できる。この理由としては、低画質時に大きな誤差として生じる偽輪郭のような歪みを現在の F_4 では計量できない為だと考えられる。この問題を解決する為に低画質用、高画質用と画質クラスを考慮した評価モデルの作成が必要である。

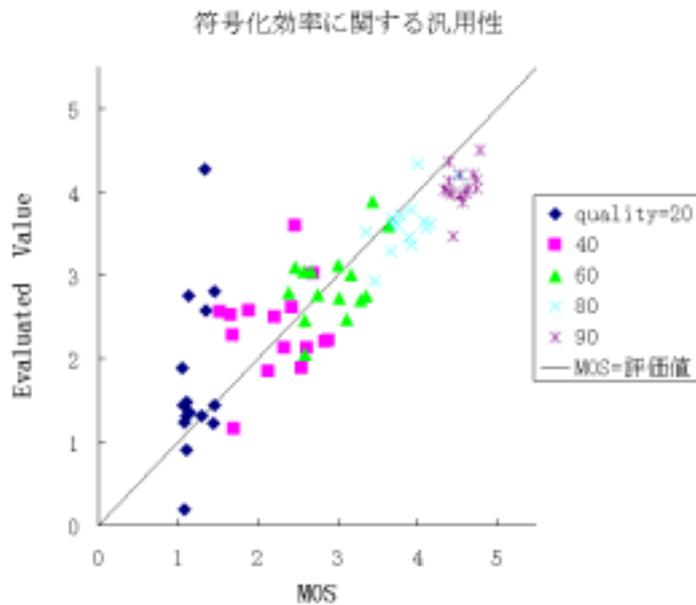


図 4.25: オープンデータセット評価値と MOS との関係 (符号化効率)

4.5.3 必須画質劣化要因

本手法では画質の劣化を基本画質劣化要因の線形和でモデル化しており、画質劣化要因 F_i は基本的で抜けがない事が望ましく、かつ不要な F_i は除去すべきである。 $F_1 \sim F_4$ の中で必要不可欠な画質劣化要因を検討する。

表 4.9 に各組合せを総当たりで計算した結果を示す。できるだけ少ない F_i で R^{**} が高い方が望ましい為、*マークをした F_3 と F_4 の組合せが最も良い。この事から、主観評価においては、ランダムノイズよりも局所的な歪みの方が評点の決定に重要な要因となっている事が考えられる。

表 4.9: 評価値-MOS 間の自由度調整済み重相関係数 (CIE L*u*v*)

組合せ i (F_i)	R^{**}
1	0.534
2	0.524
3	0.377
4	0.860
1 2	0.528
1 3	0.537
1 4	0.886
2 3	0.602
2 4	0.884
3 4	0.880
1 2 3	0.601
1 2 4	0.884
1 3 4	0.886 *
2 3 4	0.882
1 2 3 4	0.885

4.6 むすび

原画像と符号化再生画像の差である色差で画質劣化要因を定義し、多変量解析の手法を用いて線形回帰評価モデルを構築した。

表 4.10に、評価モデルを構築した色空間における平均誤差、最大誤差、重相関係数 R を示す。この結果、輝度色差分離系の色空間において評価値が良好に主観評価結果を近似できる事を示した。修正 Munsell色空間を使用した場合とほぼ同じ精度が得られ、輝度色差分離系の色空間における評価モデルの構築が有効である事が分かった。これらの色空間においては、どの色空間もほぼ同じ評価精度が得られるが、わずかではあるが、均等色空間で画質評価モデルを構築する方が精度が高い。

表 4.10: 構築した各色空間における評価精度

色空間の系統	色空間	平均誤差	最大誤差	重相関係数 R
輝度色差非分離	RGB	0.64	2.70	0.669
輝度色差分離	YCbCr	0.48	1.14	0.845
	CIE L*a*b*	0.44	1.72	0.859
	CI EL*u*v*	0.43	1.60	0.871

均等色空間で構築した評価精度を比較する為、修正 Munsell 色空間での結果を加え表 4.11に示す。表より CIE L*u*v* 色空間でのモデルの評価精度がわずかに高い事が示された。しかし、CI EL*a*b* 色空間と CI EL*u*v* 色空間の評価精度の差はほとんど無い為今後は色に関する研究で多く用いられている CI EL*a*b* 色空間において検討を進める。

表 4.11: 均等空間における評価精度の比較

色空間	平均誤差	最大誤差	重相関係数 R
修正 M u n s e l l	10.45	1.71	0.848
CI EL*a*b*	0.44	1.72	0.859
CI EL*u*v*	0.43	1.60	0.871

表 4.11に2 評価モデルを構築した色空間において最小の画質劣化要因で、良い結果を得られる組合せと必須画質劣化要因を示す。この結果、どの評価モデルにおいても輪郭近傍に生じる誤差を計量する F_4 が必須画質劣化要因となっており、画像の局所特徴に着目して評価モデルを構築する重要性が明らかになった。

表 4.12: 各色空間における必須画質劣化要因

色空間の系統	色空間	最小组合せ	必須画質劣化要因
輝度色差非分離	RGB	F_1, F_2, F_3, F_4	F_4
輝度色差分離	YCbCr	F_1, F_4	F_4
	CIE L*a*b*	F_3, F_4	F_4
	CI EL* u* v*	F_1, F_3, F_4	F_4

また、CI EL* a* b* および CI EL* u* v* 色空間において構築したモデルの汎用性を確認した結果、画像内容に対して良好な汎用性を示している事が確認できたが、符号化効率の汎用性において低画質時の近似精度が悪く、全ての画質範囲を網羅する単一モデルを作成するのが困難である可能性が見い出された。低画質時の評価精度が低い理由としては、低画質時に生じる輪郭が不連続となる誤差を必須画質劣化要因として抽出された F_4 では計量できない為と考える。この為全ての画像範囲を網羅する単一モデルではなく、画質クラスを考慮し低画質用、高画質用に分けて処理するモデルを構築する事によって評価精度が向上すると考えられる。

CI EL* a* b* 色空間においてモノクロ画像に対する汎用性を検討した結果、モノクロ画像とカラー画像の両方を用いて評価モデルを構築した場合、高画質時の評価精度は高いが、低画質時の評価精度は低い事が確認できた。この事からモノクロ画像の評価はカラー画像の評価とは別に扱う事が望ましいといえる。

第 5 章

画質クラスを考慮した画質評価モデル

先に構築した評価モデルの符号化効率に関する汎用性を確認した際に、低画質時の評価精度が悪く、全ての画像範囲を網羅する単一モデルではなく、低画質用、高画質用と画質クラスを考慮してモデルを作成する事によって、評価精度が向上する可能性を示した。

本章では主観評価値を参考に cjpeg における quality60 未満を低画質、60 以上を高画質とクラス分けし画質クラスを考慮した評価モデルを作成する。

このモデルにおいて、色差の計算には式 (4.3) に示す CIE $L^*a^*b^*$ 色空間における色差式を用いる。

5.1 画質クラスを考慮した評価モデルの定義

本モデルは評価対象を低画質、高画質にクラス分けし、それぞれの画質クラスに応じた評価モデルを作成する。

5.1.1 システムの構成

評価値を得る基本的なシステムは図 4.2 と同様である。評価モデルを構築する際に 低画質用の評価モデルでは quality60 未満の符号化再生画像と主観評価値を、高画質用の評価モデルには quality60 以上の符号化再生画像と主観評価値を使用し、低画質用と高画質用の偏回帰係数と基本画質劣化要因を得た。評価値は基本画質劣化要因の線形和として式 (4.1) を用いて計算される。

5.1.2 基礎画質劣化要因と主成分分析

本評価モデルにおける基礎画質劣化要因は前の章で示したものと同様である。低画質用と高画質用それぞれ別に求めた画質劣化要因も相関を取り除く為に主成分分析を行ない冗長性を除去して低画質評価用、高画質評価用の基本的な画質劣化要因を抽出する。ここでも I 個の F_i から I 個の主成分 Z_i が計算されるが、評価値への寄与率が低いものは近似的に省略できる。各主成分 Z_i の全体に対する累積寄与率を式 (4.7) で計算し、十分大きな (0.90 程度) の累積寄与率を有する J 個まで ($I \geq J$) の主成分 Z_i を基本劣化要因とし、これをもとに評価値を計算する。

5.1.3 評価値の計算

評価値は得られた 各々の偏回帰係数と基本画質劣化要因の線形和として式 (4.8) を用いて得られる。

5.2 評価モデルによる画質評価

5.2.1 画質評価結果

対象となる個々の原画像を 8 種類の符号化を行ない符号化再生画像を得た。以下に評価値と MOS の一致度と主成分の解析結果を示す。

—主成分—

画質クラスに応じて、それぞれ F_i を正規化した後、共分散行列を作り主成分分析を行なう。図 5.1、図 5.2 にそれぞれ低画質クラス、高画質クラスの累積寄与率を示す。図より、 $Z_1 \sim Z_3$ までの累積寄与率が 0.99 を越えているので、この 3 つの主成分のみでも十分な評価精度が得られる事が分かる。

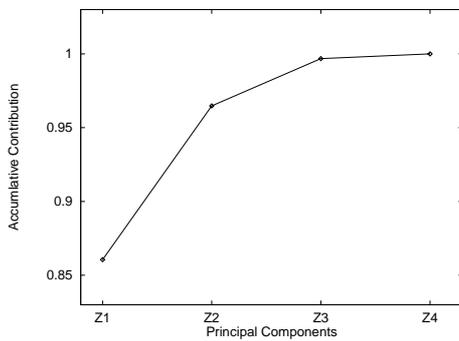


図 5.1: 各主成分 Z_i の累積寄与率
(低画質クラス)

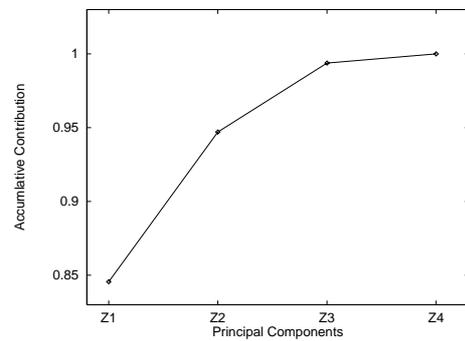


図 5.2: 各主成分 Z_i の累積寄与率
(高画質クラス)

— 偏回帰係数 —

$Z_1 \sim Z_3$ に対する偏回帰係数を式 (4.8) と MOS との重回帰分析により計算した結果を低画質クラスについて表 5.1、高画質クラスについて表 5.2 に示す。

表 5.1: 偏回帰係数 b_j
(低画質クラス)

b_0	b_1	b_2	b_3
2.583	-0.278	0.038	-2.62^{-7}

表 5.2: 偏回帰係数 b_j
(高画質クラス)

b_0	b_1	b_2	b_3
5.576	0.213	-0.017	5.83^{-7}

— 評価精度 —

$Z_1 \sim Z_3$ を用いて式 (4.8) により計算した評価値と MOS との関係を求める。ここでは open data に対する評価を行なう為、全評価用画像セットから 1 枚の画像を除外して評価モデルを作成した後、除外した画像をこのモデルによって評価する。除外する画像を逐次変えて評価を行ない、その値をプロットし図 5.3 および図 5.4 に示す。これらの図より、評価値が MOS を良好に近似している事が確認できる。

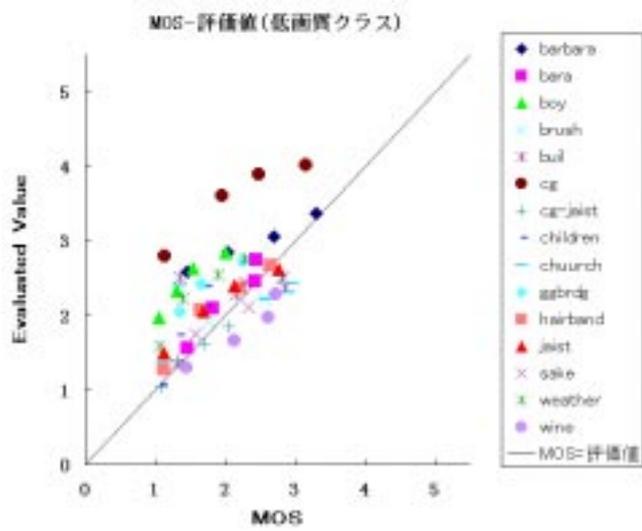


図 5.3: 評価値と MOS との関係
(低画質クラス)

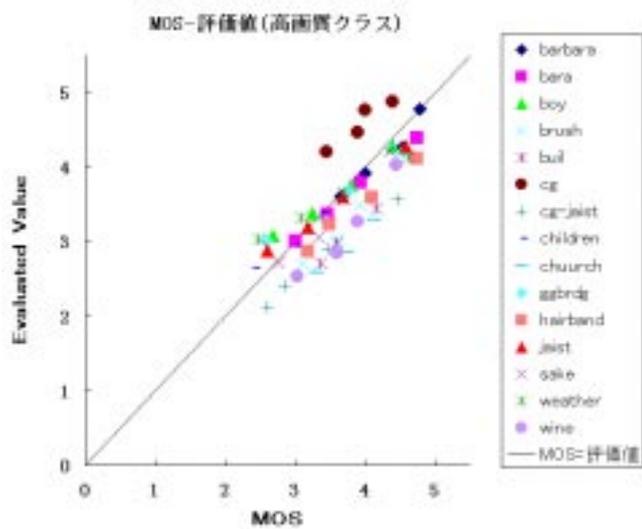


図 5.4: 評価値と MOS との関係
(高画質クラス)

更に、評価値と MOS の近似度を定量的に表す為に評価値と MOS の重相関係数 R を式 (4.9) により計算する。低画質クラスの R は 0.73、高画質クラスの R は 0.81 となり、画質クラスを考慮しない場合と比較して低画質時の評価精度が向上し、画質クラスを考慮したモデルが有効である事が確かめられた。

5.2.2 汎用性

図 5.3 および図 5.4 においては、それぞれのクラスとも全評価用画像セットから 1 枚の画像を除外してモデルを作成し、除外した画像の評価をこのモデルを用いて行なっているが、この手法では汎用性の確認はできない。この為、モデルを作成する際に使用する画像種類、符号化効率を限定してモデルを作成し、モデルを作成する際に使用しなかった画像を評価する事によって汎用性の確認を行なう。

—画像内容に関する汎用性—

低画質クラス、高画質クラスの両方について、全評価画像から 1 セットを除いた 15 種類のモデルを作成し、得られた偏回帰係数を使用して除外した画像の評価を行なう。低画質クラスにおける結果を図 5.5 に、高画質クラスの結果を図 5.6 に示す。この結果、平均誤差約 0.5 と全体的に主観評価値を良好に近似している事が分かり、画像内容に関しての汎用性が確保されていると言える。

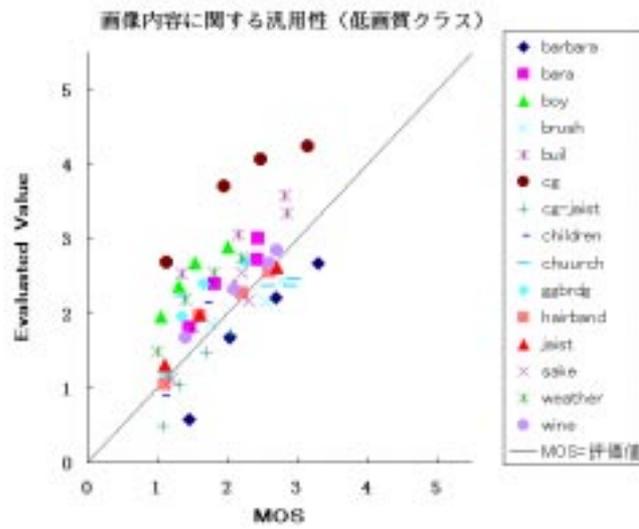


図 5.5: オープンデータ評価値と MOS との関係 (画像内容)
(低画質クラス)

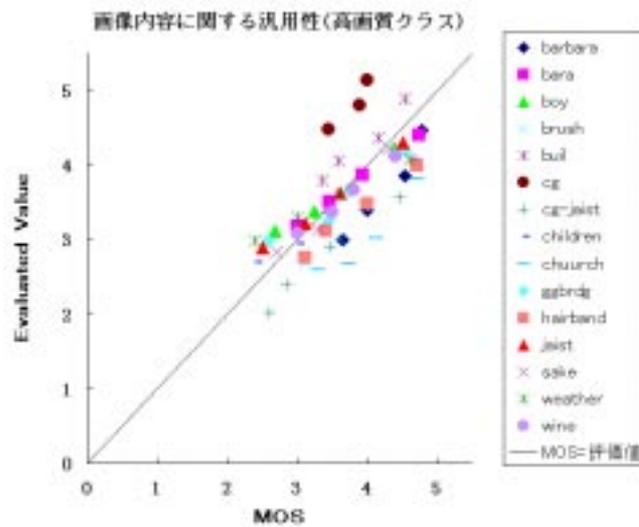


図 5.6: オープンデータ評価値と MOS との関係 (画像内容)
(高画質クラス)

—符号化効率に関する汎用性—

15 枚の原画像に対してクラスに応じた符号化効率で符号化した画像を用いてモデルを作成し評価を行なう。更に、符号化効率に関する汎用性を確認する為に特定の符号化効率 1 種を除いたモデル作成し、得られた偏回帰係数を使用して除外した画像の評価を行なう。低画質クラスにおける結果を図 5.7 に、高画質クラスにおける結果を図 5.8 に示す。これらの結果より、全ての画像範囲を網羅するモデルでは低画質時の評価精度が悪いという問題があったが、画質クラスを考慮する事によって、低画質時の評価精度が向上されている事が分かる。

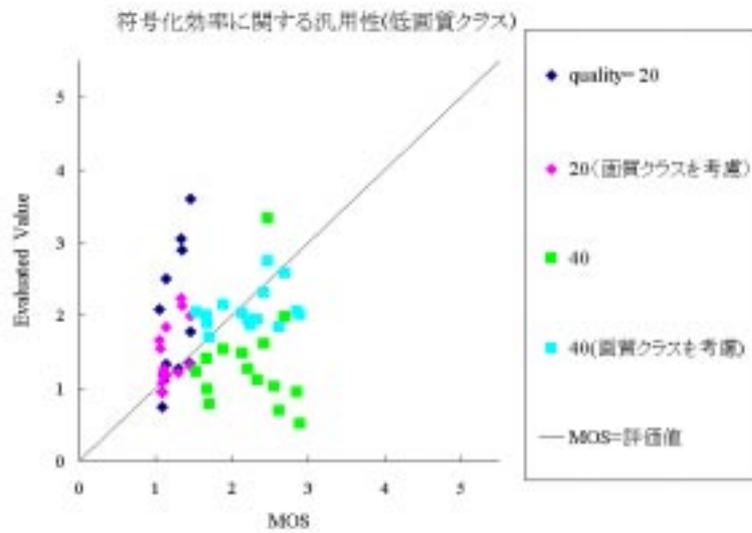


図 5.7: オープンデータセット評価値と MOS との関係 (符号化効率)
(低画質クラス)

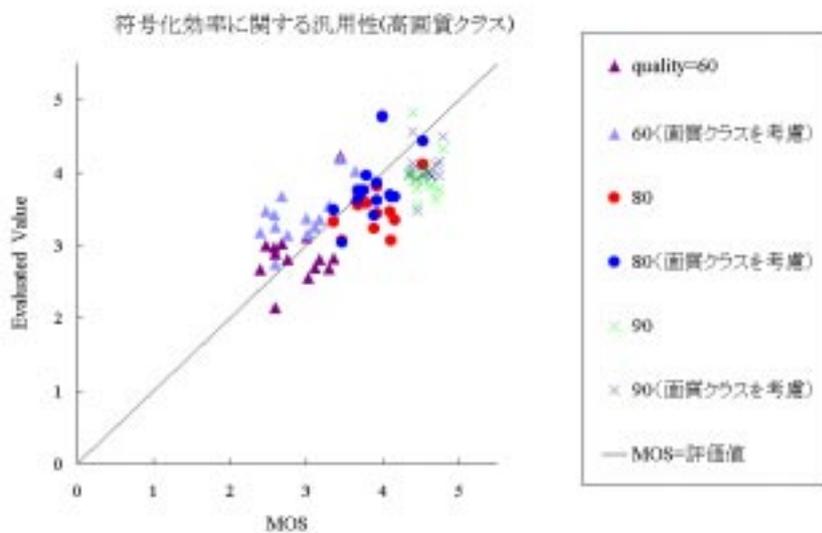


図 5.8: オープンデータセット評価値と MOS との関係 (符号化効率)
(高画質クラス)

5.2.3 必須画質劣化要因

画質クラスを考慮したモデルも、画質の劣化を基本画質劣化要因の線形和でモデル化しており、画質劣化要因 F_i は基本的で抜けがない事が望ましく、また不要な F_i は除去すべきである。それぞれのモデルにおいて、 $F_1 \sim F_4$ の中で必要不可欠な画質劣化要因を式 (4.10) で求められる自由度調整済重相関係数を用いて検討する。各モデルとも基本画質劣化要因の組合せを変えて R^{**} を計算した結果を表 5.3、表 5.4 に示す。

表 5.3 によると、低画質クラスの評価モデルにおいては、*マークをした F_1, F_2, F_3 の組合せが最も良く画像の輪郭部分の損傷よりも、ランダムノイズや誤差の自己相関係数を計量する画質劣化要因が重要となっている。これは、低画質クラスにのみ発生するブロック歪みのようなテクスチャパターンを作る歪みや相関を持った誤差を計量する為と思われる。

高画質クラスになると、ブロック歪みや偽輪郭は知覚されない。しかし、画像の輪郭損傷のような画像内容に関わる歪みが大きな妨害と感じる為に、画像の局所的特徴を計量する画質劣化要因 F_4 が必要不可欠となってくる。

表 5.3: 評価値-MOS 間の自由度調整済み重相関係数

(低画質クラス)	
組合せ $i (F_i)$	R^{**}
1	0.539
2	0.292
3	0.665
4	0.576
1 2	0.667
1 3	0.519
1 4	0.588
2 3	0.625
2 4	0.702
3 4	0.666
1 2 3	0.731 *
1 2 4	0.692
3 4	0.676
2 3 4	0.716
1 2 3 4	0.726

表 5.4: 評価値-MOS 間の自由度調整済み重相関係数

(高画質クラス)	
組合せ $i (F_i)$	R^{**}
1	0.355
2	0.376
3	0.465
4	0.797
1 2	0.320
1 3	0.424
1 4	0.809 *
2 3	0.428
2 4	0.808
3 4	0.808
1 2 3	0.409
1 2 4	0.800
3 4	0.806
2 3 4	0.802
1 2 3 4	0.798

5.3 むすび

全ての画質範囲を網羅する単一の評価モデルにおいて、低画質時の評価精度が悪い問題を、画質クラスを考慮する事で解決した。

また、必須画質劣化要因はその画質クラスによって異なる事を示し、その要因の妥当性を確認する為に主観評価時における視点停留領域の測定を行なう必要性が生じた。

第 6 章

画質評価時における視点停留領域

画質クラスを考慮したモデルを構築した結果、画質クラスによって必須画質劣化要因が異なる事が示された。低画質クラスにおける評価精度の低さは、現在の画質劣化要因では偽輪郭のような誤差を計量できない為と考えられる。

JPEG 符号化による誤差は、平坦部でのランダムなもの、輪郭周辺に生じるもの、テクスチャパターンを形成するものに大別される。これらはランダムノイズや、偽輪郭の発生、輪郭の不連続や不鮮明化、色にじみやむらを生じる。

人間は輪郭の不連続やテクスチャパターンを作る誤差のような空間的つながりのある誤差はランダムな誤差より 10 倍以上知覚しやすいという知見がある。

この知見を裏付け、評価モデルに視点停留情報を適応する事を目的とし、主観評価時の視点停留領域を測定する事によって評点を決定する際に人間が画像のどの部分に注目したか、どのような歪みが評価値に大きく寄与したかを解明する。

6.1 視点停留領域測定実験の条件

6.1.1 測定システム構成

測定にはアイマークレコーダー (竹井機器製 EYE MOVEMENT MONITOR) を使用し、評価者の視点の追跡および、視点の停留点を解析する。

測定システム構成を図 6.1 に示す。

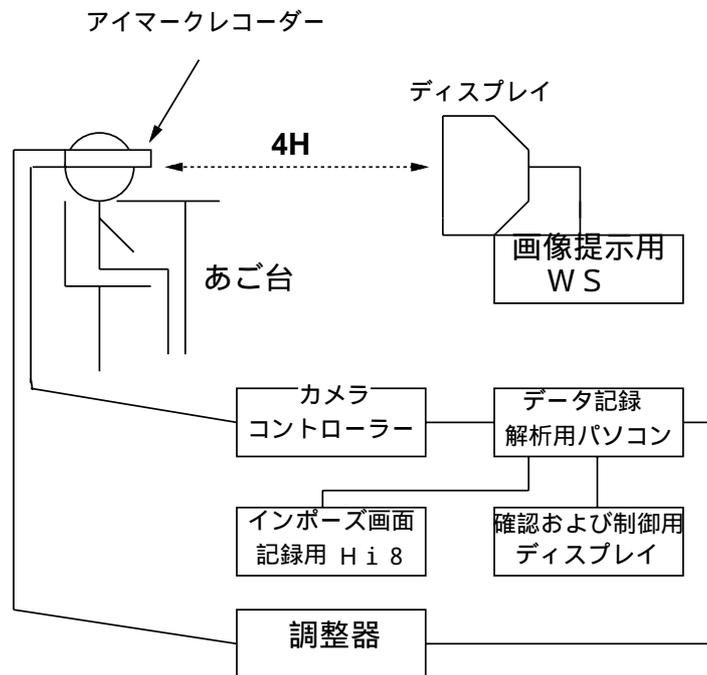


図 6.1: 視点停留領域測定システム

評価者は主観評価実験の際と同様の条件下で画質評価および視点停留領域の測定を行った。アイマークレコーダーより得られたデータは、パソコンおよび接続された Hi8 のテープに記録される。

6.1.2 テスト画像と符号化の種類および評価者

—テスト画像と符号化の種類—

評価実験経験者と未経験者との特性を比較する為に、表 6.1 に示す人物、静物、風景、建築構造物、CG の 5 種の画像を、表 6.2 に示す quality で低画質、中画質、高画質に符号化した。

表 6.1: 視点測定用原画像

カテゴリ	原画像名
人物	hairband
静物	sake
風景	wi ne
建築構造物	bui l
CG	cg1

表 6.2: 符号化時の quality

画質	quality
低画質	20
中画質	50
高画質	80

図 6.2 から図 6.15 に、視点測定に用いた原画像およびそれぞれの符号化再生画像と、符号化誤差を可視化したものの一例を示す。



図 6.2: 原画像
(hairband)



図 6.3: 再生画像
(hairband 低画質)



図 6.4: 符号化誤差 ($E_{L^*a^*b^*}$)
(hairband 低画質)



図 6.5: 再生画像
(hairband 中画質)



図 6.6: 符号化誤差 ($E_{L^*a^*b^*}$)
(hairband 中画質)



図 6.7: 再生画像
(hairband 高画質)



図 6.8: 符号化誤差 ($E_{L^*a^*b^*}$)
(hairband 高画質)



图 6.9: 原画像
(sake)



图 6.10: 再生画像
(sake 低画質)



图 6.11: 符号化誤差 ($E_{L^*a^*b^*}$)
(sake 低画質)



图 6.12: 再生画像
(sake 中画質)

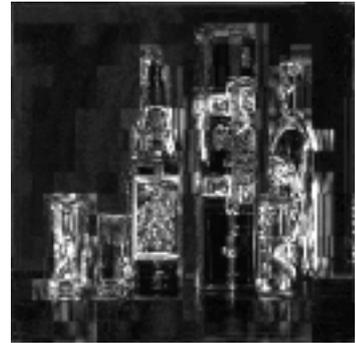


图 6.13: 符号化誤差 ($E_{L^*a^*b^*}$)
(sake 中画質)



图 6.14: 再生画像
(sake 高画質)



图 6.15: 符号化誤差 ($E_{L^*a^*b^*}$)
(sake 高画質)

—評価者—

評価者には画質評価実験の経験者(ここでは10時間以上評価経験を積んだ者)と未経験者の画質評価時の特性の違いについて考察する為

- 画質評価実験未経験者 4名
- 画質評価実験経験者 5名(女性1名を含む)

について測定を行なった。なお評価者の色覚は全て正常である。

6.1.3 観視条件

視点測定は通常の画質評価時と同様の観視条件および評価語を用いたが、測定中の頭部のずれを補正する為、コンスタントにキャリブレーションを行なう必要がある。また画像の提示方法も若干異なっている。竹井機器製 EYE MOVEMENT MONITOR を用い、左目の視点を測定した。その他の実験条件を表 6.3に示す。キャリブレーションは測定器固有の方法の他、ディスプレイ上に図 6.16のような上下左右、中央にマーカーをつけた 256×256 画素のキャリブレーション用画像を提示し、指定するマーカーを視点停留させる事で行なった。

表 6.3: 視点測定実験条件

視距離	4H(H: 画像高)
室内照明	暗室
ピーク輝度	68[cd/m ²]
提示方法	連続提示(図 6.1参照)
観視時間	キャリブレーション 45 秒 + (原画像 10 秒、符号化画像 10 秒)× 2
評価者数	各グループ合計 9 名
評価尺度	5 段階劣化尺度

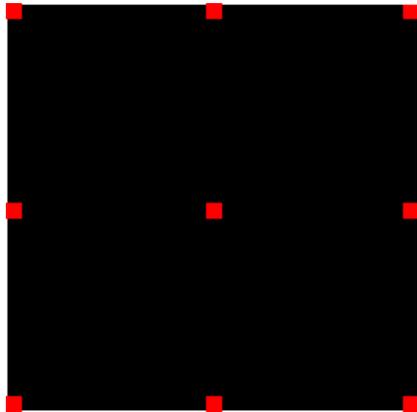


図 6.16: キャリブレーション画像

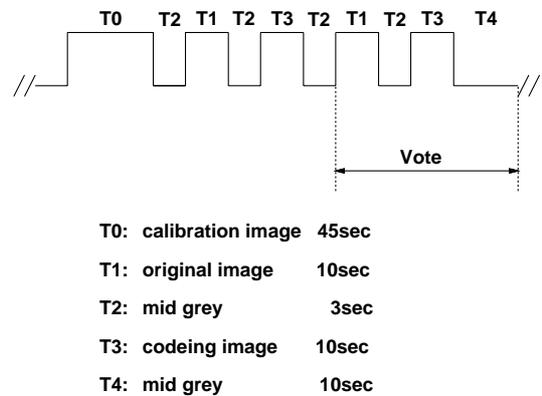


図 6.17: 画像提示方法

6.2 視点停留領域の定義

本研究では視点が視覚 1° の範囲に $300[\text{ms}]$ 以上停留した際に、その領域を視点停留領域とする。文献 [9] によれば視点停留領域を定義する際の視点の停留時間を $150[\text{ms}]$ から $300[\text{ms}]$ 以上としているが、本研究において $200[\text{ms}]$ 以上視点が停留した場合と、 $300[\text{ms}]$ 以上視点が停留した場合とを比較した結果、視点停留領域の違いが現れた。そこで、より厳しく定義する為に $300[\text{ms}]$ 以上視点が停留した領域を視点停留領域とした。画質評価経験者の同種かつ同画質における視点領域測定結果を図 6.18および 6.19に示す。ここで図 6.18は視点停留時間を 200ms とした場合、図 6.19は視点停留時間を 300ms とした場合である。なお、視点停留領域は画像中に白抜きで示す。

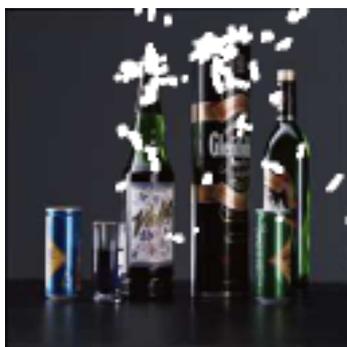


図 6.18: 停留時間 200ms 以上



図 6.19: 停留時間 300ms 以上

6.3 視点停留領域の測定結果

視点停留領域の測定から評価経験の有無による特性の違いを考察する。

6.3.1 原画像に対する視点停留領域

符号化再生画像の視点停留領域を解析する前に、まず原画像の視点停留領域測定結果を述べる。

—画質評価未経験者の視点停留領域—

原画像に対する画質評価未経験者の視点停留領域の例を図 6.20と図 6.21に示す。

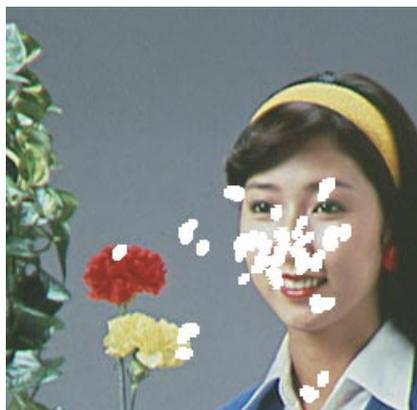


図 6.20: 画質評価未経験者の視点停留領域 (hairband) 図 6.21: 画質評価未経験者の視点停留領域 (sake)

両図から、画質評価未経験者は、画像特徴の特定の領域に視点が停留する傾向が強い事が分かる。

—画質評価経験者の視点停留領域—

原画像に対する画質評価経験者の視点停留領域測定の例を図 6.22と図 6.23に示す。



図 6.22: 画質評価経験者の視点停留領域
(hairband)



図 6.23: 画質評価経験者の視点停留領域
(sake)

画質評価経験者は評価未経験者と異なり、広範囲な画像輪郭に視点が停留する傾向が強い。

6.3.2 符号化再生画像に対する視点停留領域

—画質評価未経験者の視点停留領域—

図 6.24から図 6.41に符号化再生画像に対する画質評価未経験者の視点停留領域と符号化誤差を示す。



図 6.24: 符号化再生画像
(hairband 低画質)



図 6.25: 視点停留領域
(hairband 低画質)



図 6.26: 符号化誤差
(hairband 低画質)



図 6.27: 符号化再生画像
(hairband 中画質)



図 6.28: 視点停留領域
(hairband 中画質)



図 6.29: 符号化誤差
(hairband 中画質)



図 6.30: 符号化再生画像
(hairband 高画質)



図 6.31: 視点停留領域
(hairband 高画質)



図 6.32: 符号化誤差
(hairband 高画質)



図 6.33: 符号化再生画像
(sake 低画質)



図 6.34: 視点停留領域
(sake 低画質)



図 6.35: 符号化誤差
(sake 低画質)



図 6.36: 符号化再生画像
(sake 中画質)



図 6.37: 視点停留領域
(sake 中画質)



図 6.38: 符号化誤差
(sake 中画質)



図 6.39: 符号化再生画像
(sake 高画質)



図 6.40: 視点停留領域
(sake 高画質)



図 6.41: 符号化誤差
(sake 高画質)

[低画質時～高画質時]

画質評価未経験者の場合、画質の程度に関係なく、低画質時から高画質時にかけてどの画質においても、画像輪郭の一部に視点が停留しており、画像全体に生じている誤差全てを評価するのではなく、あくまで一部を見て、画質評価を行なう傾向が強い。

—画質評価経験者の視点停留領域—

図 6.42から 6.59に符号化再生画像に対する画質評価経験者の視点停留領域と符号化誤差を示す。

[低画質時]

低画質では、視点は画像の平坦部に生じた偽輪郭周辺、各画像輪郭に生じた大きな誤差や、ブロック歪みのようなテクスチャパターン部分に停留する傾向がある。また、色のにじみを生じている部分にも、視点の停留が見られる。



図 6.42: 符号化再生画像
(hairband 低画質)



図 6.43: 視点停留領域
(hairband 低画質)



図 6.44: 符号化誤差
(hairband 低画質)



図 6.45: 符号化再生画像
(sake 低画質)



図 6.46: 視点停留領域
(sake 低画質)



図 6.47: 符号化誤差
(sake 低画質)

[中画質時]

中画質になると平坦部の偽輪郭はほとんどなくなり、視点は各画像の輪郭部分に生じた大きな誤差と、テクスチャパターンを作る誤差の部分に視点が停留する傾向が見られる。



図 6.48: 符号化再生画像
(hairband 中画質)



図 6.49: 視点停留領域
(hairband 中画質)



図 6.50: 符号化誤差
(hairband 中画質)



図 6.51: 符号化再生画像
(sake 中画質)

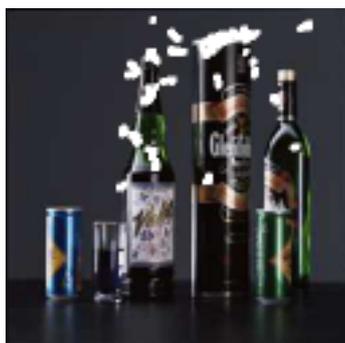


図 6.52: 視点停留領域
(sake 中画質)



図 6.53: 符号化誤差
(sake 中画質)

[高画質時]

符号化によって生じる誤差の知覚される領域も、誤差の量も減少する為、原画像を見ている時と、同様な視点停留が見られる。また、高画質時においても輪郭近傍には誤差が知覚されやすい為、輪郭近傍に視点が停留する傾向が強い。



図 6.54: 符号化再生画像
(hairband 高画質)



図 6.55: 視点停留領域
(hairband 高画質)



図 6.56: 符号化誤差
(hairband 高画質)



図 6.57: 符号化再生画像
(sake 高画質)



図 6.58: 視点停留領域
(sake 高画質)



図 6.59: 符号化誤差
(sake 高画質)

6.4 むすび

画質評価時における視点停留を測定した結果、画質評価経験者は大きな誤差を生じる輪郭近傍や、偽輪郭を生じる平坦部およびブロック歪みのようなテクスチャパターンを生じている部分を広範囲に評価しているが、画質評価未経験者は画像の限られた部分の誤差に着目し、その部分の誤差の量で画質評価を行なう傾向が見られる。

第 7 章

画質評価時における視点停留領域の評価モデルへの適用

7.1 視点停留領域を考慮した画質評価モデルの定義

7.1.1 システムの構成

本モデルにおいて評価値を得るプロセスを図 7.1 に示す。図 4.2 との相違は視点停留領域情報を新たに加えた事である。

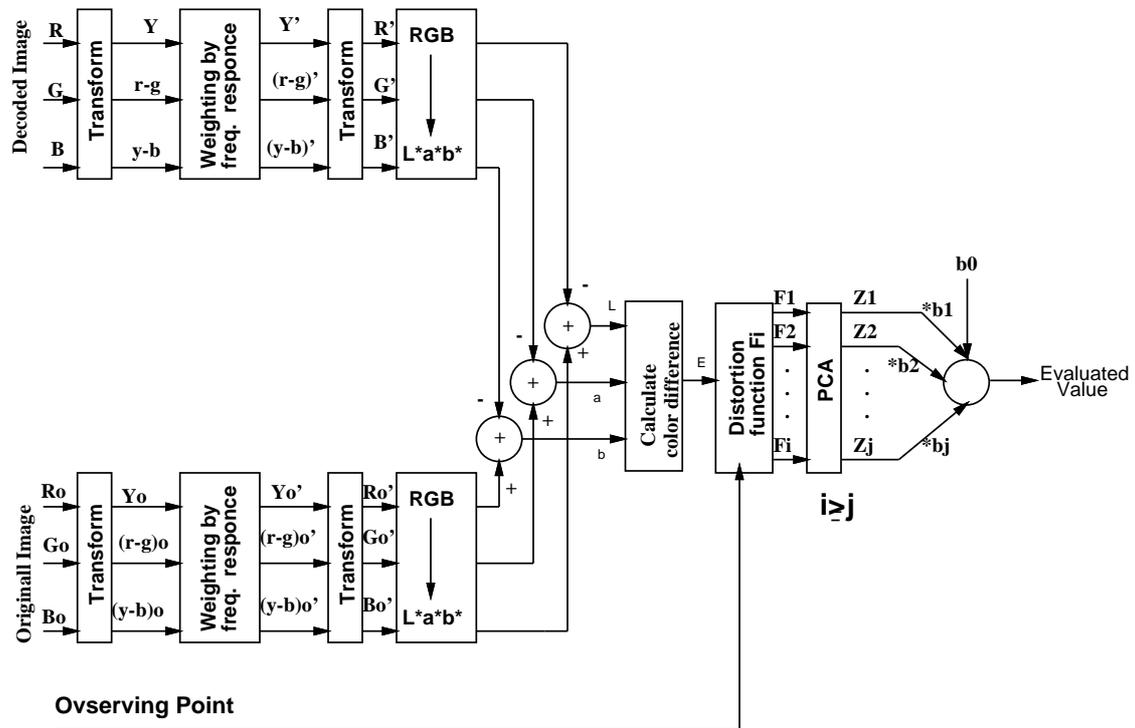


図 7.1: 評価値を得るプロセス

7.1.2 視点停留情報の関数近似と画質劣化要因の主成分分析

— F_4 : 画質評価時の視点停留情報—

符号化再生画像中の誤差は評価者がそれを視覚で捕らえてのみ知覚される。測定より得られた視点停留領域から、評価者は画像の全ての画像領域をまんべんなく注視しているのではなく、特徴的な歪みが生じている領域を注視して評価する傾向がある事が分かる。視点停留領域をアイマークレコーダーで捕らえ、その領域の誤差を計量して F_4 とする。

$$F_4 = \sum_{t=0}^T \sum_{a=-A}^A \sum_{b=-B}^B \{e(p+a, q+b)\} \quad (7.1)$$

p, q は評価時刻 t における視点停留点座標を示し、 T は総評価時間、 A, B は視点停留点を中心とする誤差計量範囲 (視点停留領域) を示す。本来は円形の範囲とするべきであるが、計算の容易さから、正方領域とした。先に定義したように視点停留領域は視覚 1° の領域内に視点が $300[\text{ms}]$ 以上停留する点とし、同じ領域に複数回視点が停留する場合はそれぞれ

を合計する。評価実験経験者は画像の一部ではなく広い範囲に視点を移して評価している事から、同一領域に 300[ms] より長い時間視点が停留しても計量は一度しか行わない。

7.1.3 画質劣化要因 F_i の主成分分析と評価値の計算

ここで定義した画質劣化要因 F_4 も、他の画質劣化要因と独立である保証がない為、相関を持っている可能性がある。この為、他の画質劣化要因と主成分分析を行ない冗長性を除去して基本的な画質劣化要因を抽出する。評価値は式 (4.8) で示すように基本劣化要因 Z_j の荷重和で与える。

7.2 評価モデルにおける画質評価結果

対象となる個々の原画像を 8 種類の効率で符号化を行ない符号化再生画像を得た。評価値と MOS の一致度と主成分の解析結果を示す。

—主成分—

F_i を正規化した後、共分散行列を作り主成分分析を行なう。図 7.2 に累積寄与率を示す。図より、主成分 $Z_1 \sim Z_3$ までの累積寄与率が 0.99 を越えているので、この 3 つの主成分だけでも十分な評価精度が得られる事が分かる。

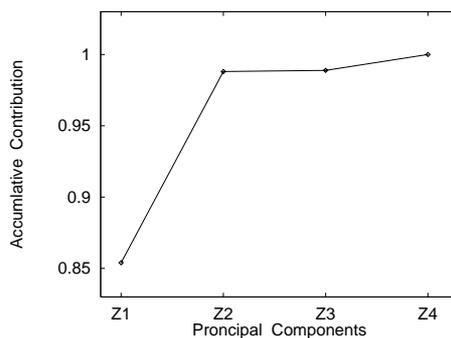


図 7.2: 各主成分 Z_i の累積寄与率

— 偏回帰係数 —

基本画質劣化要因 $Z_1 \sim Z_3$ に対する偏回帰係数を式 (4.8) と MOS との重回帰分析により計算し、表 7.1 に示す。

表 7.1: 偏回帰係数 b_j

b_0	b_1	b_2	b_3
5.607	-1.022	0.009	-2.03×10^{-7}

— 評価精度 —

$Z_1 \sim Z_3$ を用いて式 (4.8) により計算した評価値と MOS との関係を求める。ここでは open data に対する評価を行なう為、全評価用画像セットから 1 枚の画像を除外して評価モデルを作成した後、除外した画像をこのモデルによって評価する。除外する画像を逐次変えて評価を行ない、その値をプロットし図 7.3 に示す。図より、評価値と MOS の近似度があまり高くない事が分かる。

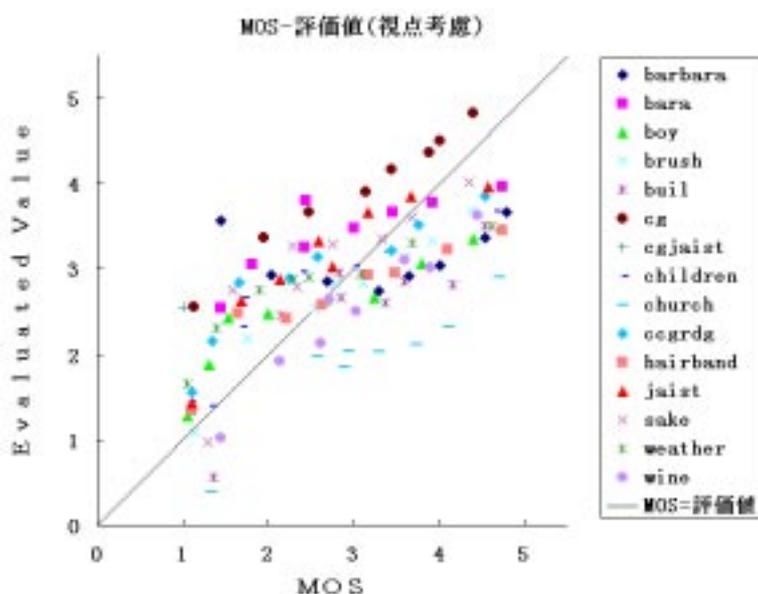


図 7.3: 評価値と MOS との関係

更に、評価値と MOS の近似度を定量的に表す為に評価値と MOS の重相関係数 R を式 (4.9) により計算する。

本手法による R は 0.71 となった。修正 Munsell 色空間において符号化誤差を Godlove 色差として計算し、画質の劣化要因を色差の関数として与え、多変量解析を行なって作成した線形回帰モデルの R は 0.72 となり、どちらの色空間においても高い評価精度が得られない。この原因として、注視領域測定の評価者数が不十分である為に評価時の個人性を排除できなかった為と考える。しかし、評価者数を増やし、一般的な注視領域情報を得る事により評価精度が向上すると考えられる。

7.3 むすび

本章では原画像を基準として JPEG 符号化による画質劣化の程度を評価する為、原画像と符号化再生画像の差を均等色空間である CIE $L^*a^*b^*$ 色空間における色差から定義した基礎的な画質劣化要因に画質評価時の視点停留情報を加え、多変量解析の手法を用いて線形回帰評価モデルを構築し、評価値が主観評価結果を近似できる事を示した。

修正 Munsell 色空間を使用した視点停留情報を加えたモデルと比較を行なったところ、ほぼ同じ精度が得られ CIE $L^*a^*b^*$ 色空間での評価モデルの構築が有効である事が分かった。

視点停留測定の評価者数が不十分であった為に、個人性を排除できず主観評価値を良好に近似していない画像もあるが、これは評価者数を増やす事により改善され評価精度が向上すると考えられる。

第 8 章

結言

JPEG 符号化によって生じる主な歪みは以下に示す通りである。

1. グローバルな歪み
2. テクスチャパターンを作る歪み
3. 輪郭損傷のような画像内容に関わる局所的な歪み

本研究では JPEG 符号化による画質劣化を客観的に評価する為に、劣化を原画像と符号化再生画像の差である色差を用いて関数近似した以下に示す基礎画質劣化要因 F_1 から F_4 を用いて画質評価モデルを構築した。

F_1 : 一画素あたりの色差の量 (ランダムノイズに相当)

F_2 : 符号化サブブロック間の誤差変化量

F_3 : 誤差の自己相関係数

F_4 : 輪郭部分における誤差の重みづけ (画質評価時の視点停留領域)

評価モデルは定式化した基礎画質劣化要因を基に多変量解析の手法を用いて構築し、評価値が輝度色差分離系の色空間において良好に主観評価値を近似する事を示した。これらの評価モデルは主観評価値を約 87%の精度で近似できる。輝度色差分離系の色空間の中でも、均等色空間の評価精度がわずかに高く、その中でも CIEL*u*v*色空間が評価に適していると考えられる。

しかし、この評価モデルにおいては低画質時の評価精度が低い為に、評価対象を低画質クラス、高画質クラスのような画質クラスを考慮したモデルを作成し、低画質時の評価精度を向上させた。このモデルの個々の画質劣化要因の組合せによって自由度調整済重相関係数を求める事により必須画質劣化要因を抽出した。この結果、低画質クラスにおいては誤差の自己相関係数を計量する画質劣化要因が重要である事を示した。高画質クラスにおいては、テクスチャパターンのような画像歪みが生じない為、より顕著に知覚される画像輪郭部分の誤差を計量する画質劣化要因が重要である事が示された。

画質クラスによって必須画質劣化要因が異なった事から、主観評価時の視点停留領域を測定する事によって評価精度の向上を試みた。視点停留領域の測定により、画質評価経験者と未経験者では視点停留領域が異なる事を示し、また画質クラスによって、特徴的な視点停留が起こる事を示した。低画質クラスにおいては、輪郭の不連続な部分はもちろんのこと、低画質クラスに特徴的に現れるブロック歪みのような大きな画像歪みに視点停留領域が集中する傾向を得た。高画質クラスになると、ブロック歪みのような大きな画像歪みは生じないが、輪郭部分に歪みが大きな妨害となり、輪郭領域に視点停留領域が存在する傾向を得た。

これらの視点停留情報を評価モデルの基礎画質劣化要因に取り入れ評価モデルを構築したが、時間の関係で評価者数が充分でなかった為に視点停留領域の個人性を排除する事が出来ず、高い評価精度を得る事が出来なかった。しかし、視点停留領域の傾向から分かるように個人性を排除した視点停留領域を求められれば、評価精度は向上すると期待できる。

本研究において画質劣化要因を式(4.3)の色差式より求めた色差を用いて画質劣化要因を定式化した。式(4.3)から得られるCIE L*a*b*色空間における誤差を可視化した図8.2と各成分の誤差を同じ条件で可視化した図8.3、図8.4、図8.5との比較比較により、符号化再生画像上では各成分により誤差の特徴が異なる。

今後、各成分を個別に扱って画質劣化要因を計算する事によって評価精度が向上すると考えられる。これは、各成分を個別に扱うと以下に示す各成分の特徴を、より計量できるからである。

1. 明度成分は平坦部でのランダム誤差に加え、輪郭周辺に大きな誤差を生じる。
2. 色成分は輪郭部分での大きな誤差に加えて、テクスチャパターンをもつ誤差が輪郭周辺に生じる。



図 8.1: 符号化再生画像
(quality=20)



図 8.2: 符号化誤差
($E_{L^*a^*b^*}$)



図 8.3: 符号化誤差
(E_{L^*})



図 8.4: 符号化誤差
(E_{a^*})



図 8.5: 符号化誤差
(E_{b^*})

また、視点停留領域測定の評価者を増し評価精度を向上させる事、および JPEG 符号化以外の符号化方式に評価モデルを拡張する事などが今後の課題として挙げられる。

なお、本研究は 1997 年 11 月 11 日 ~ 13 日に工学院大学で行なわれたカラーフォーラム JAPAN'97 に於いて優秀ポスタ賞を授賞致しました。

謝辞

本論文をまとめるにあたり、熱心に御指導下さいました本学 小谷 一孔助教授に深く感謝致します。また終始有益な御助言と御指導を頂きました本学 宮原 誠教授に感謝致します。

さらに、主観評価実験および視点停留領域測定実験にお忙しい中でも御協力を頂きました関係者各位と小谷研究室および宮原研究室ならびに堀口・阿部研究室の皆様にご心からお礼を申し上げます。

また評価用画像の入手に関し、御協力を頂いた大和様、三瓶様、貞弘様に感謝致します。

参考文献

- [1] Recommendation ITU-R BT.500 – 7 “Method for the subjective assessment of the quality of television pictures” 1995
- [2] 成田 長人 “画像符号化の主観評価法に関する考察” 電子情報通信学会論文誌, E J77-B1, No.2 pp.102-111 1994
- [3] 宮原誠 “統計的画像符号化” アイピーシー 1990
- [4] K. Kotani, Q. Gan, M. Myahara and V. R. Al gazi “Objective Picture Quality Scale for Color Image Coding” ICP’95 Washington, D. C, Vol. 3 pp.133-136, 23-26. Oct. 1995
- [5] K. Kotani, Q. Gan and M. Myahara “Objective Picture Quality Scale for Color Image Coding” PCS’94 Sacramento, pp.100-103. Sept. 1994
- [6] テレビジョン学会編 “テレビジョン画像の評価技術” pp. 85-86, 1986
- [7] 原島博 “画像情報圧縮” オーム社, 1991
- [8] インターフェース “カラー静止画像の国際標準符号化方式” pp. 160-182, Dec. 1991
- [9] 樋渡編著 “視聴覚情報概論” pp. 41-45, 昭晃堂, 1987
- [10] 坂田・磯野 “視覚における色度の空間周波数 (色差弁別閾)” テレビジョン学会, Vol. 31, No. 1, pp. 29-35, 1977
- [11] 宮原・小谷・堀田・藤本 “客観的画質評価尺度 (PQ) – local feature の考慮と汎用性 –” 電子情報通信学会, Vd. J73-B1 No.3 pp.218-218 1990
- [12] 日本色彩学会編 “新編 色彩科学ハンドブック” 東京大学出版会

- [13] 太田 登 “色彩工学” 東京電機大学出版局
- [14] 池田 光男 “色彩工学の基礎” 朝倉書店
- [15] テレビジョン学会編 “画質と音質の評価技術” pp.19-20,1991
- [16] 石川・花村・会津・池上 “画像通信に用いる色空間の評価—符号化効率による比較—” 画像電子学会誌, Vol.23, No1, pp10-23, 1986

付録 1 主観評価法

[評価尺度の水準]

主観評価では評価結果に客観的普遍性を持たせる事が重要である。画質の評価はまず言葉によって表現されるが、言葉による表現は曖昧であり、極めて定量性を欠く事が多い。これを防ぐ為に物理的な刺激に対する人間の主観的な感覚を定量的に一つの尺度上に当てはめる事が必要となる。

測定とは、ある規則にしたがって対象に数字を割り当てる事であるが、この規則の違いにより 4 つの尺度が作り出される。この 4 つの尺度は対象に加える経験的操作と数字の系列との間に対応関係が存在する為に定義する事が出来る。ここでは Stevens による 4 つの評価尺度の水準について述べる。[15]

—名義尺度—

名義尺度とは数字の大小が問題になるのではなく、単なる符号、レットルとしてたまたま数字を用いたものである。名義尺度も 2 つに分類されその数字が 1 つのものを同定する場合に用いられる場合と、形式番号のような同じ番号を持つものが複数存在する場合とがある。

—順序尺度または序数尺度—

順序尺度とは大小関係 (順位付け) という操作によって決定される。この場合与えられる数値はあくまで順序を示すものであるからその尺度値では数学的処理を加える事は出来ない。

—間隔尺度または距離尺度—

間隔尺度は対象に割り当てられた数字の順序だけではなく、相互の間隔(距離)が意味を持ってくる。しかしこの尺度上には絶対的な零点がない。この尺度には加算、減算、平均値や相関係数を求める操作など尺度値の比を求める以外の全ての操作が可能になる。

—比率尺度または比例尺度—

比率尺度では間隔尺度と事なり絶対零点が存在し、この尺度上では尺度値の比を求める操作を含めた全ての数学的操作が可能になる。明るさなどの感覚には零点が存在するので比較尺度上で表現する事が出来る。

[2 重刺激劣化尺度法]

EBU法(The double-stimulus impairment scale method)とも言われシステムの劣化特性の測定に用いられる手法である [1]。

—評価実験系統—

評価実験の系統図を図.6に示す。評価者には、まず信号源から直接ディスプレイに画像が提示され、次に同一の画像が評価対象システムを経由して提示される。

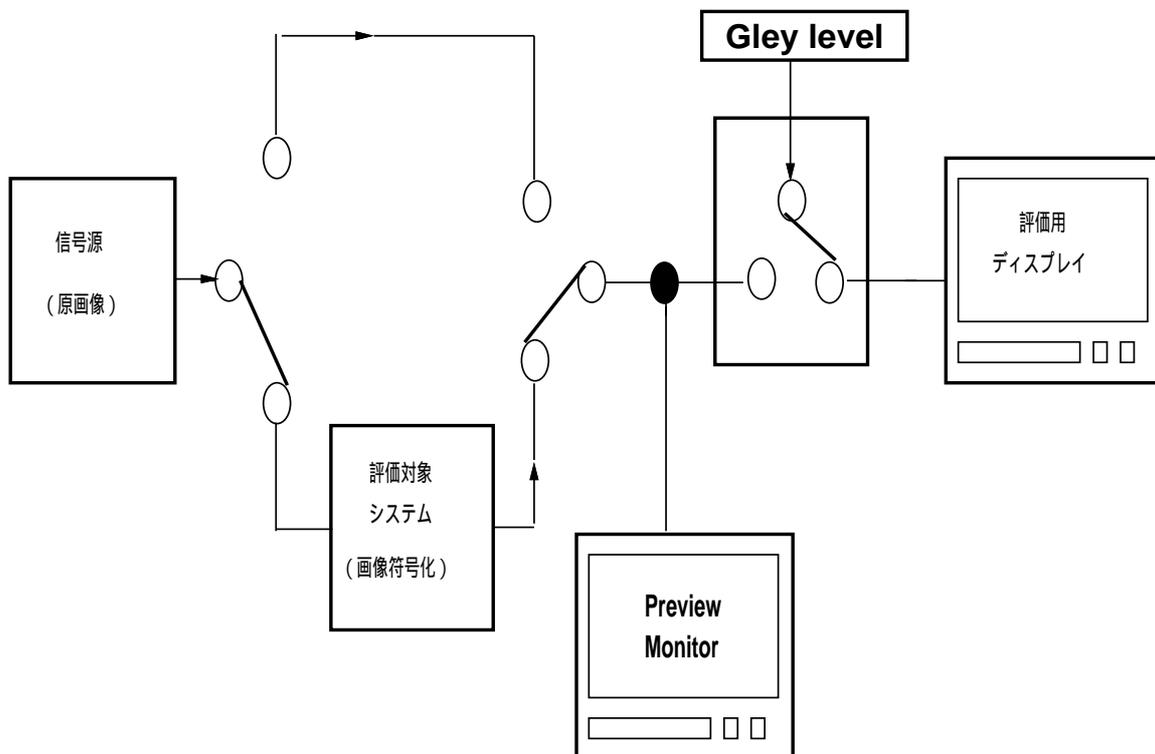


図 .6: EBU 法の実験系統図

—信号源—

安定な結果を得る為には、劣化の無い原画像を用いる事が重要である。この事から信号源にはデジタル信号形式を用いるのが望ましい。

—観視条件—

表.1の観視条件の他、ディスプレイは22~26インチの高品質のものを使用する事が望ましい。

表 .1: 観視条件

視距離	4H(H : 画像高)
非発光画面のピーク輝度に対する比	≤ 0.02
暗室で黒レベル輝度のピーク輝度に対する比	$\simeq 0.01$
ディスプレイの brightness,contrast	PLUGE 経由でセットアップ*
観視角度の上限	30°
モニター設置場所の背景と画像輝度ピーク比	$\simeq 0.15$
室内照明	低い事
モニター設置場所の背景色	D ₆₅ (色温度 6500 度) のカーテン

* : Recommendations ITUR BT814,815 参照

—テストセッション—

テストセッションにおいて考慮しなければならない事は以下の通りである。

- 1つのセッションは説明と予備試験の時間を含めて 30 分以上続けない(これは長時間のセッションで評価者が疲労し、正しい評価が行なえなくなることを防ぐ為)。
- 画像の提示数は 40 以内に収める。
- 同一条件の状態をセッション内で 2 回提示する。
- 絵柄と劣化の大きさをランダムに組み合わせて提示する。ただし、劣化の大きさのいかに問わず、同じ絵柄の画像を続けて提示してはいけない。
- 大多数の評価者が全ての種類の評点を使い、評価試験における全ての評点の平均値が 3 になるような劣化範囲を選ぶ。

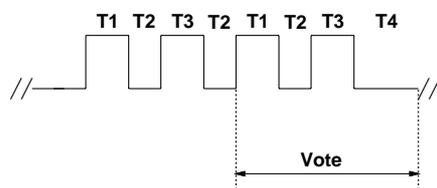
テストセッションの流れとしては 最初 4 回の提示トレーニングセッションを行ない、質疑応答、そしてメインのテストセッションとなる。

—評価者—

通常は評価試験の経験の少ない非専門家を 15 名用いる。またセッションの開始前に視力と色覚が正常であることを調べる。

—提示方法—

EBU 法における画像の提示方法を図.7に示す。



- T1: original image 10sec
- T2: mid grey 3sec
- T3: codeing image 10sec
- T4: mid grey 10sec

表 .2: 5 段階劣化尺度

評点	評価用語
5	(劣化が) 分からない
4	(劣化が) 分かるが気にならない
3	(劣化が) 気になるが邪魔にならない
2	(劣化が) 邪魔になる
1	(劣化が) 非常に邪魔になる

図 .7 EBU 法の提示方法

—評価尺度—

表.2に示す 5 段階劣化尺度を使用する。この評価語は評価語間の重みのばらつきが多くなるように心理実験で決定された。

—教示—

評価を行なう際に評価方法 (評価尺度、提示方法、判定、記入など) と生じる劣化の種類について評価者に説明を行なう。

[2 重刺激連続品質尺度法]

DSCQS(The double-stimulus continuous quality scale method) と略されることもある方法で、基準に対する評価対象システムの品質の測定に用いられる手法である [1]。

—評価実験系統—

評価実験の系統図を図.8に示す。評価者は、信号源から直接ディスプレイに入る画像と評価対象システムを経由して提示される画像を対で評価する。この時「原画像対原画像」という組合せも評価対象に加えておく必要がある。

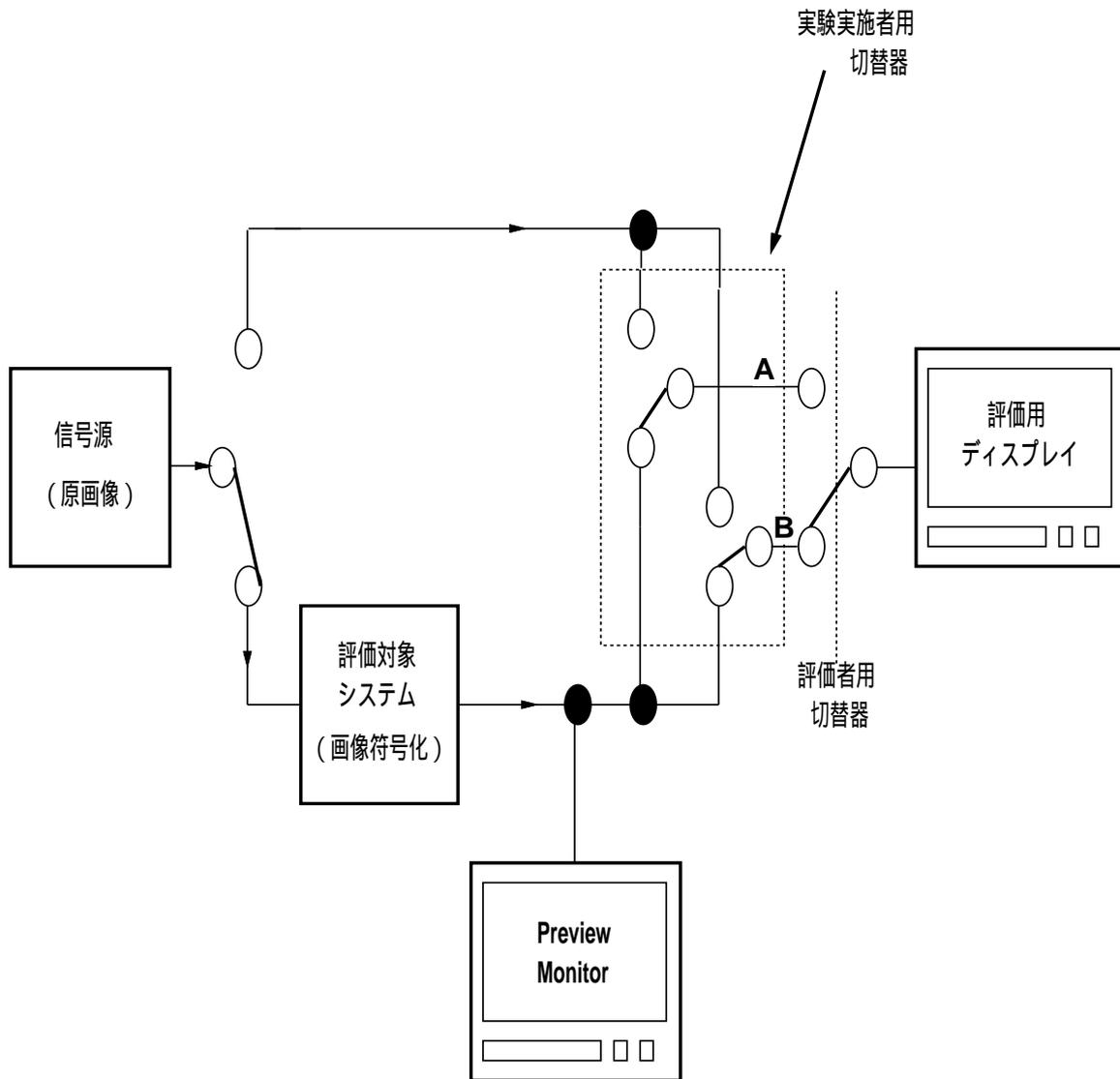


図 .8: DSCQS 法の実験系統図

—提示方法—

提示方法には次の 2 通りの方法がある。

1. 評価者に画像対を自由に (通常 10 秒間で 2~3 回) 切替えさせて評価をさせる。
2. 一つの状態を 3~4 秒、対で 5 回 (30~40 秒) 繰り返し提示し、最後の対を提示している最中に判定・記入をさせる (ただしこの提示法の結果の安定性はまだ検討中で

あると報告されている)。

—評価尺度—

対で提示される画像の個々の総合画質を、表.3に示す5段階品質尺度と照らし合わせ評価試験用紙のグラフ上に印をつけさせる。この評価語も評価語間の重みのばらつきが多くならないように心理実験で決定された。図.9に2重刺激連続品質尺度法の評価試験用紙を示す。

表 .3 5 段階品質尺度

評点	評価用語
5	非常に良い
4	良い
3	普通
2	悪い
1	非常に悪い

	No.1		No.2		No.3	
	A	B	A	B	A	B
非常に良い						
良い						
普通						
悪い						
非常に悪い						

図 .9 評価試験用紙

—信号源, 観視条件, テストセッション, 評価者, 教示—

これらの項目については EBU 法と同じである為ここでは省略する。