

## 画質の時間的推移による符号化動画像の評価法

稲積 泰宏 堀田 裕弘 村井 忠邦  
富山大学 工学部 電気電子システム工学科

930-8555 富山市五福 3190

TEL:0764-45-6758

E-mail:horita@ecs.toyama-u.ac.jp

あらまし 白黒動画像の各フレームに対して静止画像の画質評価尺度 ( $PQS_{still2}$ ) を適用した動画像の画質評価尺度 ( $PQS_{vs}$ ) の推定方法に対して人の記憶の概念に基づいた、劣化の時間的推移を考慮した新しい評価法を提案する。

これは、人の記憶の系列位置曲線を近似し、その近似曲線による重み付け平均を画質評価値とする方法である。近似曲線のパラメータは GA を用いて算出した。MOS と  $PQS_{vs}$  の相関が一番高くなるパラメータを求めた結果、系列位置曲線に近い特性が得られ、推定精度は先に提案した ( $PQS_{vs2}$ ) より改善された。本手法によって、劣化の時間的推移を評価値に反映させることができた。

キーワード 画質評価, 視覚特性, ロジスティック関数, 人の短期記憶

## Quality estimation method of coded video considering time transition of quality

Yasuhiro Inazumi Yuukou HORITA Tadakuni MURAI

Department of Electric and Electronic Engineering,  
Faculty of Engineering, Toyama University

3190, Gofuku, Toyama-shi, 930-8555 Japan

TEL:0764-45-6758

E-mail:horita@ecs.toyama-u.ac.jp

Abstract We propose an objective picture quality scale ( $PQSvs$ ) considering time transition of quality for monochrome video sequence. Quality of each frame is estimated by using the Picture Quality Scale of Monochrome Still Picture ( $PQS_{still2}$ ).  $PQSvs$  which is considered short-term character of the human memory ( $PQSvs3$ ), are carried out. The optimal parameters of short-term character are estimated using GA. The results show that  $PQSvs3$  is similar to character of the human memory. In addition,  $PQSvs3$  estimate well with the Mean Opinion Score (MOS).

key word picture quality evaluation, human visual system, logistic function, short-term memory

## 1 まえがき

動画像品質を高精度に評価するためには、動画像を1枚の静止画像として見た場合の「静的な劣化」と動画像特有の「動的な劣化」を定量化すること、人間が時間的に変動している動画像を評価している際に「提示時間の中の平均的な画質で評価しているか」、「ある時間において最も劣化した画質で評価しているか」等を見極めることが重要である [1]。これまでに文献 [2, 3] では上記2つのポイントを考慮した検討がなされている。しかしながら、これらの手法は3次元FFTや数十フレームの複雑な3次元信号処理を必要とするため膨大な計算時間を必要とした。そこで、品質を評価するための信号処理が比較的容易で評価精度が良好な画質評価法の開発が必要である。

これまでに、動画像を1フレームの静止画像の時系列データと考えた静止画像の評価方法を動画像に拡張する検討もされている [2]。これは、動画像の時間に対する動的な品質変動を考慮する方法ではあるが、品質評価を行うために必要とされる時空間の複雑な3次元信号処理を必要とするため膨大な計算時間を要した。

これに対して我々は、静止画像の品質評価法を動画像の品質評価に適用した、計算時間が短くかつ精度の良い新しい品質評価法 ( $PQS_{\text{seq}}$ ) を提案し、シーケンスを通して画質評価値の変動の少ない画像に対して、MOSは各フレームの画質評価値の平均値と最低値の線形重回帰モデルで推定可能であることを示した [4]。

しかし、この手法は劣化の最低値がシーケンスのどの時刻に生じているかが考慮されていない。ITU-Rで勧告されている提示シーケンスの時間が定められた主観画質評価法では、画質劣化のひどい部分が提示されたシーケンスのどこにあるかで、被験者の総合画質評価値 (MOS, Imps) に変動が生じる。すなわち、画質評価において劣化の程度と、その現れる時刻が評価値に強く影響すると考えられる。そこで本研究では、従来の画質評価法に、人の記憶に関する概念を適用することによりMOS推定改善の検討を行う。

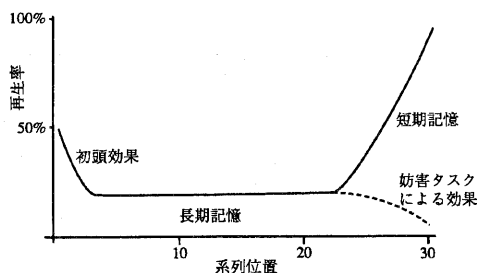


図 1: 系列位置曲線

## 2 人の記憶について

人の記憶には、外部からの干渉などに対して頑健になった長期にわたる記憶 (長期記憶) と話されたことをそのまま復唱できる範囲の記憶 (短期記憶) がある。

30個の無関係な単語を1つずつ1秒間提示し、できるだけ多くの単語を記憶する心理実験において、単語が提示された位置と再生率の関係をグラフにした系列位置曲線を図1に示す [5]。この図から最初の部分の再生率が高く (初頭効果)、最後の7個程度の再生率が極めてよいことがわかる。また、全ての単語を提示した後には妨害タスクを与えるとこの部分が消失する。ここで、消失する成分が短期記憶の成分、消失しない成分が長期記憶の成分と考えられている。

また、短期記憶の一つの感覚情報貯蔵は、感覚器官で受け取られた情報がどれだけ利用できるかという実験で、利用する時間間隔が500msまで急激に低下してそれを越えると約40%で一定になる。

この系列位置曲線は、単語の記憶による心理実験よりもたらされたものではあるが、動画像の画質評価も特徴ある数十フレームが時系列に提示されていると考えることができるので、この曲線が画質評価法に適用可能であると考えた。また、主観評価実験では評価シーケンスを見てすぐに評点をつけるので、妨害タスクに相当する部分がなく図1の実線の特性をとり、最後から500ms程度の劣化が結果に大きな影響を与えていると考えられる。

### 3 PQSvs システムの構築

#### 3.1 静止画像の画質評価尺度

白黒静止画像に対する PQS は、静止画像に妨害を与えている歪みをその特徴に基づいて基礎歪み要因に分類し、各基礎歪み要因の量の線形和で表される [6]。PQS の基礎歪み要因は画像全体にわたる global な妨害だけでなく local feature を考慮したものであり、次の 5 種類が定義されている。

- F1: ランダムノイズ荷重平均 2 乗誤差
- F2: 視覚の弁別閾を考慮した平均 2 乗誤差
- F3: 符号化サブブロック間の誤差変化量
- F4: 誤差の自己相関関数
- F5: 輪郭近傍の visibility を考慮した平均 2 乗誤差

これらを定量的に求め、各歪み要因間の相関性を削除するために主成分分析を行い、得られた主成分  $Z_n$  と主観評価実験で得られた  $MOS_{still}$  との線形重回帰モデルにより PQS を構成することができる。したがって、 $b_0^*, b_n^*$  を重回帰分析より得られる偏回帰係数とすると、 $PQS_{still}$  は次式で求められる。

$$PQS_{still} = b_0^* + \sum_{n=1}^J b_n^* Z_n \quad (1)$$

PQS を線形重回帰モデルのみで計算すると、高品質な画像では評点が 5 をこえたり、劣化のひどい画像では 1 より小さくなるという問題がある。そこで本研究では、ロジスティック関数を導入した非線形重回帰モデルを考えた PQS を用いる ( $PQS_{still2}$ ) [7]。この  $PQS_{still2}$  は次式で定義される。

$$\left. \begin{aligned} PQS_{still2} &= \frac{4}{1 + \exp(-\beta(PQS^* - 3))} + 1 \\ PQS^* &= b_0^* + \sum_{n=1}^J b_n^* Z_n \end{aligned} \right\} \quad (2)$$

ここで、 $\beta$  は  $PQS^* = PQS_{still2} = 3$  の点での傾きを表す。また、 $Z_n$  は基礎歪み要因を

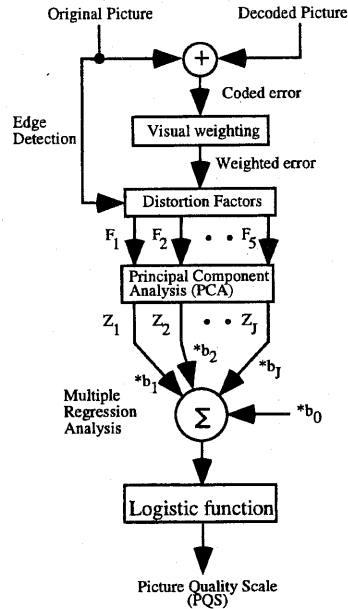


図 2:  $PQS_{still2}$  システム

主成分分析した主成分を表し、 $b_0^*, b_n^*$  はロジスティック関数を考慮した非線形パラメータの最小 2 乗推定より得られる偏回帰係数である。図 2 に  $PQS_{still2}$  システムを示す。

#### 3.2 $PQS_{vs}$ システム

動画における各フレームについて  $PQS_{still2}$  を適用し、動画の PQS を求める  $PQS_{vs}$  システムを構築する (図 3)。このシステムは、原画像と符号化画像の各フレームを静止画像として取り出し、各フレームについて  $PQS_{still2}$  を計算し、これらの PQS データを  $MOS_{video}$  に近似するように処理をし、動画全体での  $MOS_{video}$  の推定値を出力する。

##### 3.2.1 人の記憶の概念を考慮した画質評価法 ( $PQS_{vs3}$ )

動画の画質評価において、シーケンス中のどの部分に強烈な画質劣化が生じるかで被験者の総合画質値が変動すると考えられるので、劣化部分の発生時間に対する画質評価の依存性を

解析し、これらを品質評価に反映させなければならない。

そこで図 1 に示した人の記憶の系列位置曲線の特徴を、近似的に図 4 と仮定して、この近似関数にしたがって  $PQS_{still2}$  の重み付け平均をとり、これを  $PQS_{vs}$  とする。

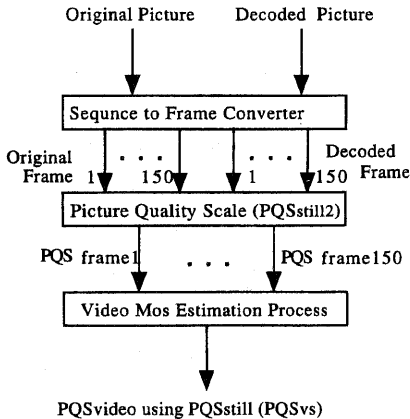


図 3:  $PQS_{vs}$  システム

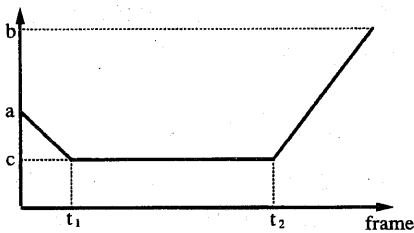


図 4: 系列位置曲線の近似

## 4 主観評価実験

### 4.1 テスト画像の種類

テスト画像は、パン、ズーム、シーンカット、複数の動き物体を含む 5 種類の動画画像シーケンス (FG: Flower Garden, MC: Mobile & Calendar, PP: Popple, BI: Bicycle, TT: Table Tennis) を用いた。符号化レートは 4, 9, 15Mbps

表 1: 静止画像の主観評価実験の条件

視距離	4H(H:画像の高さ)
照明	なし(準暗室)
最大輝度	75(cd/m <sup>2</sup> )
評定者数	33名
評価尺度	5段階妨害尺度
評価方法	2重刺激劣化尺度 (EBU) 法

表 2: 動画画像の主観評価実験の条件

視距離	4H(H:画像の高さ)
照明	なし(準暗室)
最大輝度	70(cd/m <sup>2</sup> )
評定者数	18名
評価尺度	5段階妨害尺度
評価方法	2重刺激劣化尺度 (EBU) 法

の 3 種類を用いた。これらは、150 フレームで 1 シーケンスが構成されている。

### 4.2 フレーム毎の主観評価実験

$PQS_{vs}$  システムで用いる  $PQS_{still2}$  を作成するために主観評価実験を次に示す条件で行った。

動画画像の全 150 フレームについて各シーケンスに対して 2 フレーム (5, 147 フレーム) について、表 1 に示す条件で主観評価実験を行った。

ここで 2 フレームとしたのは、フレーム数を 2, 4, 8 と増やしても MOS 推定は改善されなかったため、2 フレームで十分であると考えられる [4]。

### 4.3 動画画像の主観評価実験

$MOS_{video}$  は文献 [3] で行われた主観評価実験のデータを用いることにした。これは表 2 に示す条件下で行われていた。また、 $PQS_{vs}$  の推定精度については、5 段階評価の EBU 法を用いている関係上、現段階では最大誤差が 0.5 以下と考えているが、実際には文献 [7] で行われたように  $MOS_{video}$  の個人内ばらつきから推定精度の上限を求める必要がある。

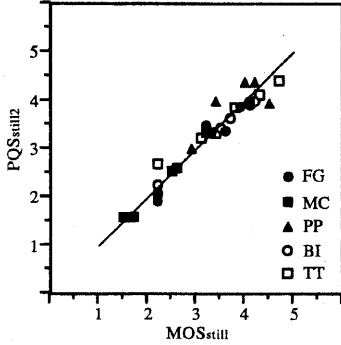


図 5:  $MOS_{still}$  と  $PQS_{still2}$  との関係

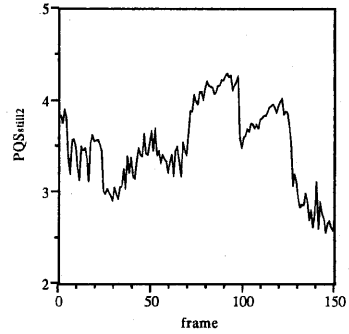


図 6: 150 フレームにおける  $PQS_{still2}$  の結果 (Table Tennis, 4Mbps)

## 5 実験結果

### 5.1 $PQS_{still2}$ の推定精度

$PQS_{still2}$  を構築するために、各シーケンスから前述の2フレームを選択し、この  $MOS_{still}$  を用いて  $PQS_{still2}$  を作成する。主観評価実験で得られた  $MOS_{still}$  と  $PQS_{still2}$  の相関関係を図5に示す。また、このとき相関係数は0.97、推定の最大誤差は0.59であった。

次に動画「Table Tennis」に対して、フレーム毎の品質評価値 ( $PQS_{still2}$ ) を求めた結果を図6に示す。これよりシーケンス中の画質の動的変動がわかり、2つのシーンカット前後での画質評価値が激しく変化していることが示されている。特に、最終シーンの品質評価値は局所的に急激に落ちていることがわかる。

### 5.2 $PQS_{vs3}$ を用いた MOS の推定

図4で示した系列位置曲線の近似曲線のパラメータ ( $a, b, t_1, t_2$ ) を求めるために、とりうる範囲を

$$0 \leq a, b \leq 1000 \quad (3)$$

$$c = 10 \quad (4)$$

$$0 \leq t_1, t_2 < 150 \quad (5)$$

として遺伝的アルゴリズム (GA) を用いて  $MOS_{video}$  と  $PQS_{vs3}$  の相関が最大になるときのパラメータを求めた。このときのパラメータ

表 3: GA による系列位置曲線のパラメータ

	a	b	$t_1$	$t_2$	error	R
ALL	10	1000	0	128	1.02	0.92
noFG	10	1000	0	128	1.02	0.91
noMC	10	1000	0	128	1.02	0.88
noPP	10	1000	0	128	1.02	0.93
noBI	10	1000	0	128	1.02	0.92
noTT	1	250	50	144	0.63	0.97

の値と相関係数、最大誤差を表3に示す。ここで、ALLは5種類のシーケンス全てを用いた場合、noFGはFGを除いた場合の結果であり、noMC、noPP、noBI、noTTも同様である。また、errorは最大誤差、Rは相関係数である。ALLのときの  $MOS_{still}$  と  $PQS_{still2}$  の相関関係を図7に示す。

この表より、系列位置曲線の初頭効果の部分は現れなかったが、後半の急激な傾きはどの場合も現れている。また、「Table Tennis」を除いたときは傾きが比較的緩やかになることから、「Table Tennis」のシーン最後の20フレームの影響が画質評価値に顕著に現れているといえる。これらの結果より、短期記憶の感覚情報貯蔵と類似した特性が得られたと考えられる。

比較として文献[4]の  $PQS_{still2}$  に対する平均値と最低値の重回帰分析により  $MOS_{video}$  を推定する方法 ( $PQS_{vs2}$ ) では、 $MOS_{video}$  と  $PQS_{vs2}$  との関係は図8のようになり、このとき相関係数は0.87で最大誤差は0.94となっていた。

今回提案した  $PQS_{vs3}$  と  $PQS_{vs2}$  を比較する

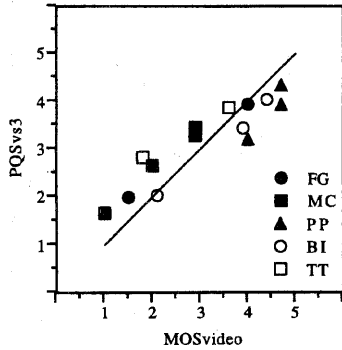


図 7:  $MOS_{video}$  と  $PQS_{vs3}$  との関係

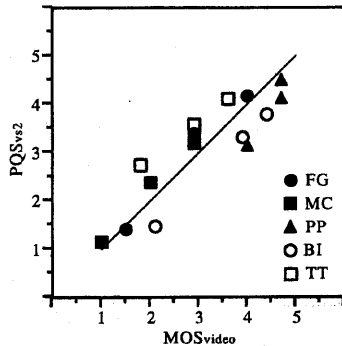


図 8:  $MOS_{video}$  と  $PQS_{vs2}$  との関係

と、最大誤差は若干大きくなったが相関が改善された。

## 6 まとめ

本研究では、白黒動画の各フレームに対して静止画像の画質評価尺度 ( $PQS_{still2}$ ) を適用した動画の画質評価尺度 ( $PQS_{vs}$ ) の推定方法の検討を、人の記憶の観点から劣化の時間的推移を考慮して行った。

まず、人の記憶の系列位置曲線を近似的に表し、その曲線の最適なパラメータを GA を用いて算出した。

次に、算出したパラメータの画像シーケンスに対する依存性を調査した結果、“Table Tennis” の影響が大きいことがわかった。しかし、

系列位置曲線を近似した関数の特性は大きく変化せず、シーン最後の影響が大きく、初頭効果は現れなかった。

GA による最適なパラメータで  $PQS_{vs3}$  を構築すると、 $PQS_{vs2}$  に比べ相関が改善された。 $PQS_{vs3}$  によって、劣化の時間的推移を評価値に反映させることができた。

今後の課題として、実際の画質評価実験における記憶の特性を明らかにすることと、より多くの画像で評価実験を行い、より高精度に推定可能なアルゴリズムを検討する必要がある。

## 参考文献

- [1] 中須 英輔:”画質評価 -画質とユーザ要求-”, テレビ誌, vol.49, No.4, pp.497-501(1993-10)
- [2] Hiroyuki HAMADA, Seichi NAMBA:”A Study on Objective Picture Quality Scales for Pictures Digitally Encoded for Broadcast”, IEICE TRANS. COMMUN. Vol.E77-B, No.12, pp.1480-1488 Dec. 1994
- [3] 堀田 裕弘, 宮原 誠, 村井 忠邦:”白黒動画における客観的画質評価尺度 ( $PQS_{video}$ )”, 信学論, Vol.J80-B-I, No.6, pp.496-504 (1997-06)
- [4] 堀田 裕弘, 稲積 泰宏, 村井 忠邦:”フレーム毎の画質評価を考慮した符号化動画の画質評価法”, 信学技報, IE98-21(1998-06)
- [5] 平井 有三:”視覚と記憶の情報処理”, 培風館, pp.221-247 (1995)
- [6] 宮原 誠, 小谷 一孔, 堀田 裕弘, 藤本 強:”客観的画質評価尺度 (PQS) -local feature の考慮と汎用性-”, 信学論, Vol.J73-B-I, No.3, pp.208-218 (1990-03)
- [7] 堀田 裕弘, 宮原 誠, 村井 忠邦:”白黒静止画像における客観的評価尺度の推定改善の検討”, 信学論, Vol.J80-B-I, No.6, pp.505-514 (1997-06)