

MPEG 符号化による劣化を検出するリファレンス画像なしの画質客観評価手法

杉本 修 川田 亮一 和田 正裕 松本 修一

KDD 研究所 画像通信グループ

〒356-8502 埼玉県上福岡市大原 2-1-15

TEL:0492-78-7416 FAX:0492-78-7439

sugimoto@kddlabs.co.jp

あらまし 近年のデジタル映像伝送の急速な普及に伴い、その運用監視を支援する技術の開発が強く求められている。特に、画質監視に関しては、運用者への負担が非常に大きく、これを自動化する技術の開発が求められている。そこで、筆者らはリファレンス画像を用いない画質評価方式について提案する。提案方法は、スペクトル拡散によるデータハイディングを用い、原画像に不可視マーカを埋め込む。この不可視マーカの特徴として、画像の幅広い周波数帯に分布するという特性があるため、この劣化を受信側で検出することによりMPEG符号化による画質劣化を推定することが可能となる。また、本方式は絵柄やビットレートを問わず、また、コーデック多段接続にも対応しているため、あらゆる形態の映像伝送に適応可能である。さらに、マーカ付加による画質劣化も極めて小さく視覚的に検知できないレベルに抑えられているため、高品質の映像伝送にも適用できるという特徴を有する。

キーワード 自動画質評価, 客観画質評価, 画質監視, 不可視マーカ, スペクトル拡散

An objective measurement scheme for perceived picture quality degradation caused by MPEG encoding without reference pictures

Osamu Sugimoto Ryoichi Kawada Masahiro Wada Shuichi Matsumoto

KDD R&D Laboratories

2-1-15 Ohara, Kamifukuoka-City, Saitama, 356-8502 Japan

TEL: +81 492 78 7416 FAX: +81 492 78 7439

sugimoto@kddlabs.co.jp

Abstract The rapid progress in digital transmission technology has spurred demand for developing a technology that supports monitoring of video transmissions. Automating the picture quality assessment process is in particular demand because it currently depends on subjective assessments by human operators and places a heavy burden on them. The authors therefore propose an objective picture quality measurement method that works without reference pictures. In this proposed method, invisible markers are embedded into original pictures by use of the spread spectrum data hiding method.

Since the markers are widely spread over the frequency domain, the degradation caused by MPEG compression can be estimated by detecting the extent of marker degradation. Our method is usable regardless of the kind of picture, bitrate, and number of stages of tandem codec connections. The degradation in picture quality caused by the embedded markers is quite small and not perceivable by the human eye. The proposed method is therefore applicable to a wide range of visual transmission services.

key words automatic picture quality assessment, objective picture quality assessment, picture quality monitoring, invisible marker, spread spectrum

1. はじめに

近年、デジタル放送、DVDなどのMPEG-2を利用したデジタル圧縮が急激に普及している。MPEG-2の特徴として、圧縮により画質劣化が生じ、その劣化は絵柄の精細さや動きの速さなどによって異なるという点が挙げられる。これは、画質劣化が絵柄によらない従来のアナログ映像伝送と大きく異なっている。

したがって、圧縮後の画質劣化の度合いを評価する技術が、特に通信放送の領域で要求され、従来からITU-Rで主観評価方式の勧告化が進められてきた^[1]。しかしながら、人間による主観評価は、監視者が常に復号画面を見続けなければならないため、膨大な作業量が必要となり、現実の運用監視には適さない。

この負担を軽減するため、画質劣化の主観評価値を機械により自動的に評価する客観評価技術の検討が同じくITUで始まっている。ただし、ここで検討されている方式はつねに原画をリファレンスとして用いる方式で、圧縮前の画像と圧縮再生された画像の差分から客観評価値を求めている。これでは、原画と復号画が異なる場所に存在する映像伝送や放送品質の運用監視には利用できない。このような背景から、筆者らは復号画像のみから、MPEG符号化による劣化の客観評価値を求めることができる不可視マーカ挿入を用いた画質評価方式を提案する。提案方式では、MPEG-2符号化の前に不可視マーカを原画の中に埋め込み、この埋め込み情報を復号画から検出して誤検出率を計算する。ここで、誤検出されるマーカ信号は符号化雑音により劣化しているため、誤検出率と符号化雑音の間にはある一定の相関があり、それゆえ誤検出率から符号化劣化を推定することが可能となる。

本稿では、放送用映像伝送の実運用上重要となるコーデック多段接続を前提とし、提案方式により多段接続の各段において高精度の客観画質評価が可能であることを示す。また、マーカ付加による画質劣化が極めて小さく、放送局間の素材伝送やHDTVなどの高品質な映像伝送にも適用可能であることを示す。

2. 関連技術の動向

従来、客観画質評価では、リファレンス画像として原画を用いるFR(Full Reference)方式が一般的であった。FR方式に関しては、VQEG等で盛んに議論され、主観評価方式との高相関化を目指し、複数の方式が提案されてきた^[2]。しかし、これらの方式は、復号画と原画が同時に利用できる状態でのみ評価が可能であるという制約があった。そのた

め、映像伝送のように原画と復号画が離れた2地点に存在する形態には適用できなかった。

そこで、映像伝送に適用可能な方式として、リファレンス画像を全く用いずに復号画やビットストリーム情報のみから画質評価を行うNR(No Reference)方式や、原画から特徴量を抽出し、映像回線とは別の低速回線でこの情報伝送した後、復号画の特徴量と比較するRR(Reduced Reference)方式などが提案された。

NR方式の代表例として、ビットストリームから各マクロブロックのMQMANT(quantizer scale)を抽出し、その値から各マクロブロックの量子化誤差を統計的に推定する方式があげられる。この方式により、シングルエンドの画質評価が可能となるが、ビットストリームからMQMANTを抽出するような特殊なデコーダが必要であり、また、コーデックの多段接続に対応していないため、実運用に適しているとは言い難い。RR方式についても、映像回線とは別にリファレンス情報を伝送する回線が必要であるなどの問題点が存在しており、その改善が待たれている。

以上の点から、提案方式では、映像伝送の実運用への適用を目指し、コーデックに特殊な処理を必要とせず、コーデックの多段接続に対応した画質評価方式を目指す。

3. 提案方式の概要

3.1. 提案方式の構成

本稿で検討するコーデックの多段接続を図1に示す。図1は2地点間の映像伝送を想定したもので、コーデックの多段接続に対して、前処理(マーカ埋め込み)と後処理(マーカ検出)が付加された構成になっている。ここで、前処理器は送信側に唯一存在しているのに対し、後処理器は、各段のデコーダ出力の後段に複数存在するという違いがある。これは、提案方式がEnd-Endにおける画質評価だけでなく、多段接続の各段における画質評価への対応も考慮しているためである。ここで、マーカ検出を行う後処理器は、映像伝送の本線から分岐しているため、本線系に影響

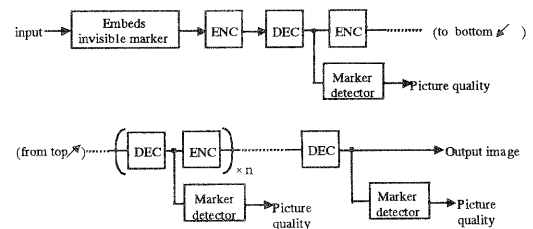


図1 提案方式の構成

を与えることはなく、また、画質評価の必要のない段には機器を配置する必要がないため、効率的な構成が可能となる。

また、同図のようにコーデックに対し前処理・後処理を加えるという構成になっているのは、マーカの埋め込み・検出をコーデックと独立に行えるようにし、汎用のコーデックを使用可能とするためである。

3.2. 不可視マーカ埋め込み・検出方式

マーカ埋め込みと検出の具体的な方式を示す。マーカ信号は、スペクトル拡散を用いたデータハイディング^[14]方式に基づき動画像への2値情報として埋め込まれる。すなわち、マーカ信号には埋め込み時点で(0,1)の情報が与えられる。ここで、マーカ埋め込みにスペクトル拡散を用いるのは、マーカ埋め込みによって生じるノイズを広い帯域に拡散させて不可視化し、さらに、すべての周波数成分をもつマーカ信号の劣化を検出することにより、より高精度な画質推定を実現するという目的による。図2にマーカ埋め込み手順を示す。

3.2.1. 送信側の処理(マーカ埋め込み手順)

1. マーカ埋め込みのための2値情報系列 **bit** を生成する。この2値情報は、受信側で検出されたマーカ2値情報と比較され、誤検出率の計算に用いられる。埋め込む系列としては、任意の系列を使用可能であるが、この系列は送受信間で共有する必要があるため、
 - i.) すべて0または1の系列 **bit**={0000000
...}, **bit**={111111...}を用いる
 - ii.) **bit**={01010101...}のような、単純な繰り返しを持つ系列を用いる
 - iii.) 送受信間で同一の乱数生成アルゴリズムを用い、ランダムな系列を生成する
 などの方法により2値情報の系列を生成する必要がある。
2. 入力信号をマクロブロック単位の信号 $x_b(n)$ に分割したのち直流成分を除去し、交流成分 $x_{ac}(n)$ を求める。提案方式では、この $x_{ac}(n)$ に対してマーカを埋め込んでいる。これは、以下のような理由による。
 - ・ マクロブロックは、MPEG-2 符号化の主要な処理単位のひとつであるため、これに一致させることにより、精度の向上が期待できる。
 - ・ MPEG-2 では、直流成分は他成分より高精度で保存されるため、劣化は主に交流成分に発生すると考え

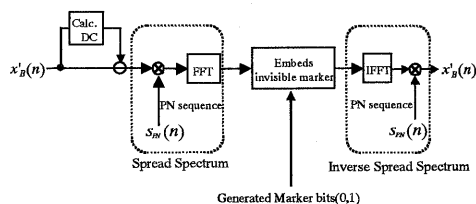


図2 送信側の処理（マーカ埋め込み）

られる。このため、劣化の推定のためには、交流成分を評価すべきである。

3. $x_{ac}(n)$ に対し PN 系列を乗することでスペクトル拡散を施し、拡散スペクトル $X_{ss}[k]$ を得る。

$$X_{ss}[k] = F\{x_{ac}(n) \times s_{pn}(n)\}$$

ここで、 F はフーリエ変換をあらわす。

4. $X_{ss}[k]$ の中から一つの周波数成分 k_0 に注目し、振幅 $A = |X_{ss}[k_0]|$ を操作して2値情報を埋め込む。具体的には、 A を任意のステップサイズ M で量子化した際のパリティが埋め込みビットに一致するように操作することにより、拡散スペクトルから2値情報情報を取り出せるようにする。

いま、埋め込みビットを bit 、 A を量子化した際のパリティが偶数、すなわち $\text{int}(A/M) = 2m$ であるとする、

$$\text{when } bit = 0 \quad A' = M(2m + \frac{1}{2})$$

$$\text{when } bit = 1 \quad \text{if } A < M(2m + \frac{1}{2}) : A' = M(2m - \frac{1}{2})$$

$$\text{else: } A' = M(2m + \frac{3}{2})$$

により A を操作し、マーカが付加されたスペクトル A' を得る。また、パリティが奇数の場合、すなわち、 $\text{int}(A/M) = 2m+1$ の場合も同様に、

$$\text{when } bit = 1 \quad A' = M(2m + \frac{3}{2})$$

$$\text{when } bit = 0 \quad \text{if } A < M(2m + \frac{3}{2}) : A' = M(2m + \frac{3}{2})$$

$$\text{else: } A' = M(2m + \frac{5}{2})$$

によりマーカを付加する。

5. IFFT および逆スペクトル拡散により時間域の信号に変換した後、2で分離した直流成分を足し合わせて、出力信号とする。
6. 以上をフレーム内のすべてのマクロブロックに対して行う。

3.2.2 受信側の処理(マーカ検出)

受信側では、送信側と同様の手順でマーカ信号を検出し、誤検出率を計算する。マーカの検出手順を図 3 に示す。

1. 復号画をマクロブロック単位に分割し、直流成分を分離した後、拡散スペクトル $A = |X_{SS}[k_0]|$ を求める。
2. 得られた拡散スペクトルの値に対して以下の判定式

$$\text{int}\left(\frac{A}{M}\right) = 2m \rightarrow \text{bit} = 0$$

$$\text{int}\left(\frac{A}{M}\right) = 2m + 1 \rightarrow \text{bit} = 1$$

を適用し、マーカを検出する。

3. 以上をすべてのマクロブロックに対して行い、検出されたビット系列と送信側で埋め込んだマーカ信号のビット系列(受信側で既知)を比較し、マーカの誤検出率を求める。
4. マーカの誤検出率と画質評価尺度の間の相関を利用し、マーカの検出率だけから画質評価尺度を導出する。

この両者の間に、絵柄や多段接続の段数によらず成立する関係が存在することを次の計算機シミュレーションで示す。

4. 計算機シミュレーション

提案方式では、マーカの検出率と画質評価尺度の関係が既知であるという前提のもとで、受信側でマーカ誤検出率を求めている。そのため、この関係が成立することを示すとともに、あらかじめこの両者の関係を測定しておくため、計算機シミュレーションによる実験を行った。実験に用いたパラメータを表 1 に示す。

4.1. パラメータ M の決定方法

実験に際して、マーカの埋め込みおよび検出の際に必要なとなるパラメータ M を設定する必要がある。このパラメータは、マーカ信号埋め込みの強度に影響する値であり、これによりマーカの誤検出率や画質推定の精度が変化するため、大変重要なパラメータである。提案方式では、各マクロブロックの交流成分にマーカ信号を埋め込んでいるが、実際の画像では、絵柄ごとに交流成分の電力が異なるため、 M を固定の値として与えた場合、絵柄ごとにマーカ検出率がばらついてしまい、それゆえ、推定精度が悪化してしまうことが考えられる。そこで、画質劣化が同一であるならば、どの絵柄に対しても同程度のマーカ検出率を示すように、 M の値を交流電力に応じて変化させる。具体的には、

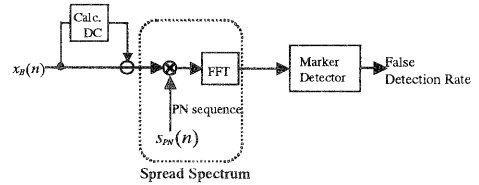


図 3 受信側の処理 (マーカ検出)

表 1 シミュレーションの条件

テスト画像	Cheerleaders, Flamingoes, Green leaves Marching in, Mobile and Calendar, Soccer
画像サイズ/フォーマット	704(H)×480(V), 4:2:2Format
エンコーダ	MPEG-2 TestModel5
ビットレート	10,20,30,40 Mbps
多段接続の段数	3段接続
画質評価尺度	RSNR(30フレーム平均)

表 2 マーカによる画質劣化と RSNR 推定精度

絵柄	マーカによる画質劣化		RSNR推定誤差の 標準偏差[dB]
	平均 MSE	平均 PSNR[dB]	
Cheerleaders	0.94	48.42	1.74 (コーデック1対向)
Flamingoes	0.61	50.27	
Green leaves	1.09	47.76	
Marching in	1.08	47.80	
Mobile and Calendar	1.15	47.52	1.71 (多段接続)
Soccer	1.34	46.86	

交流電力が σ_0^2, σ_1^2 なる 2 つの絵柄が存在する場合、

$$M_1 = \frac{\sigma_1}{\sigma_0} \times M_0$$

により M_1 を与える。ここで、 M_0 については予備実験により最適な値を求めておき、 M_1 を求めるものとする。

4.2. マーカ検出率と RSNR の関係

図4に 6 種類の画像について符号化した場合のマーカ誤検出率と RSNR の関係を示す。同図において、同一の絵柄に対して 4 つの点があるのは、表 1 に示した 4 つのビットレートによる測定結果をあらわしているためであり、誤検出率が低いものほど高いビットレートで符号化した結果となっている。ここで、RSNR および誤検出率は 30 フレームの平均値をあらわしている。なお、RSNR は、復号画の交流成分における劣化を評価するための画質評価尺度であり、

$$RSNR = 20 \log \sqrt{\frac{\sigma_{MB}^2}{MSE}}$$

により定義される。ここで σ_{MB}^2 は、各マクロブロックにおける信号分散のフレーム平均をあらわす。RSNR は、 σ_{MB}^2 の値を用いることで容易に PSNR に変換できるため(劣化が小さい範囲であれば、受信側の σ_{MB}^2 により代用できるの

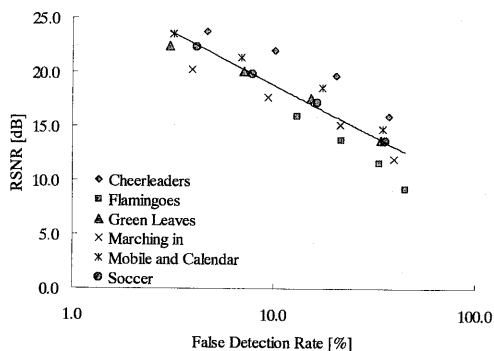


図 4 誤検出率と RSNR の関係 (コーデック 1 対向)

で、一般的な画質評価尺度として使用することができる。

図 4 より、マーカ誤検出率と RSNR の間に大きな相関が存在していることが確認できる。このとき、同図の測定結果から回帰直線を求めると、誤差 1.7dB (誤差の標準偏差) で近似できる。この回帰直線は絵柄やビットレートによらず使用可能のため、一度この直線を求めた後は、誤検出率のみを求めれば RSNR を推定することが可能となる。

また、各絵柄のマーカによる画質劣化を表 2 に示す。ここでは、原画に対する PSNR を求めることにより劣化の程度を求めている。同表より、6 絵柄の平均で 48dB 以上、最も劣化の大きい絵柄でも 46.8dB という PSNR を持つことがわかる。これにより、マーカによる画質劣化は、符号化劣化に比べて極めて小さく、視覚的にも全く検知できないレベルに抑えられていることが確認できる。

4.3. 多段接続における誤検出率と RSNR の関係

図 5 に多段接続の各中継点におけるマーカ誤検出率と RSNR の関係を示す。多段接続においても、前節と同様に絵柄や多段接続の段数によらない相関が存在すれば、誤検出率からの画質推定が可能となる。同図より、多段接続の段数によらない相関が確認でき、また、図 4 と同様の回帰直線が求められることから、提案方式は、コーデック多段接続にも適用可能であることがわかる。

4.4. 推定精度の向上に関する考察

提案方式において、推定精度を向上させる方法について考察する。

本実験では、1 フレームあたり 1320 ビットの情報を埋め込み、その検出率によって画質を推定している。ここで、1 フレームあたりにより多くの情報を埋め込むことができれば、それに伴い誤差分散 (推定誤差の 2 乗平均) が減少するため、推定精度が向上すると予測できる。提案方式において、

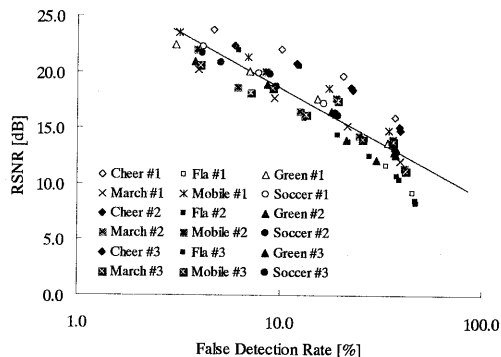


図 5 誤検出率と RSNR の関係 (コーデック多段接続)

1 フレームあたりの埋め込みビット数を増加させる方法としては、

1. マクロブロック当たりの情報埋め込みビット数 (当初は 1 ビット) を増やす
2. マクロブロック (16x16 画素) より小さい単位で情報埋め込みを行う

の 2 つが考えられる。両者の違いは、1. がスペクトル拡散を行う画素ブロックの大きさ自体は変えず、1 ブロックに埋め込む情報量を増やすのに対し、2. はスペクトル拡散を行う画素ブロックを小さくする代わりに、1 ブロックには 1 ビットだけを埋め込むという点である。この両者の効果を確認するために、

- i.) 1 マクロブロックに 2 ビットずつ挿入した場合
- ii.) 16x8 画素ブロックに 1 ビットずつ挿入した場合
- iii.) 8x8 画素ブロックに 1 ビットずつ挿入した場合

の 3 つについて測定し、推定誤差の絶対値平均およびマーカ信号による原画の劣化度を調べた。

表 3 にブロックサイズおよび埋め込みビット数を変更した場合の推定誤差絶対値の平均およびマーカによる画質劣化を示す。また、図 6, 7 に実験 ii.) iii.) における誤検出率 - RSNR の関係を示す。表 3 および図 6, 7 より、ブロックサイズが同一の場合には、1 ブロック当たりの埋め込みビット数を増やしても、推定精度はほとんど向上せず、マーカによる劣化のみが増加し、1 ブロックあたりのビット数が同一の場合には、ブロックサイズを小さくした方が推定精度が向上することが確認できる。すなわち、1 ブロックあたりに複数のマーカを埋め込んだ場合には、各マーカ信号の劣化の相関が強く、ほとんど誤差分散が減少しないということが考えられる。一方、ブロックサイズを小さくした場合には、各ブロックにおけるマーカ信号の相関が小さく、誤差分散

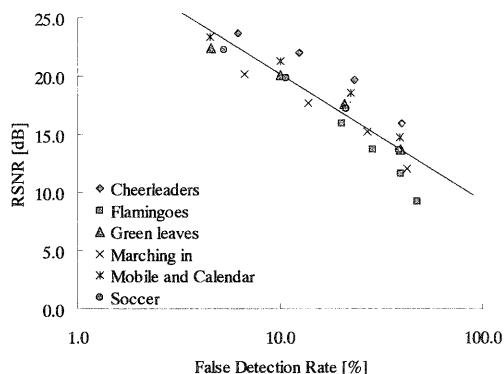


図 6 誤検出率と RSNR の関係 (ブロックサイズ 16×8)

が減少するため、推定精度が向上するといえる。ただし、1 マクロブロック当たり 1 ビットの時と比べて、実験 i.)~iii.)のすべての場合において、マーカによる画質劣化は大きくなっている。すなわち、RSNR 推定精度と画質劣化はトレードオフの関係にある。このため、マーカ埋め込みのブロックサイズは、要求する推定精度と画質劣化の許容度に応じて選択する必要がある。

5. 結論

MPEG 圧縮による画質劣化を検出する不可視マーカ埋め込みによるリファレンス画像なしの画質評価方式について提案した。提案方式では、マーカ検出率と RSNR の相関を利用し、受信側でマーカ検出率を求めるだけで高い精度での画質評価を行うことを可能としている。この相関は、絵柄やビットレート、多段接続の段数によらず成立するため、その適用範囲は非常に広い。また、マーカ信号埋め込みによる画質劣化は極めて小さいという特長を有する。このため、高品質な映像伝送サービスへの適用が可能となる。

今後の課題としては、各段のコーデックの符号化条件 (ビットレート、GOP 構造など) が異なる場合への対応や、アップ/ダウンコンバージョンなどの画像処理に対しても適用可能な方式への拡張があげられる。

参考文献

- [1] ITU-R Recommendation BT.500-8, "Methodology for subjective assessment of the quality of television pictures", Sep. 1998
- [2] Video Quality Experts Group, "Final report from the video

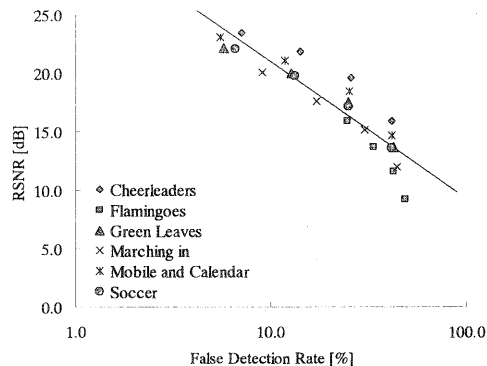


図 7 誤検出率と RSNR の関係 (ブロックサイズ 8×8)

表 3 各ブロックサイズにおける推定精度と画質劣化

ブロックサイズ	1フレーム当たりの埋め込みマーカ数	画質劣化 (PSNR[dB])	推定精度 (標準偏差[dB])
16x16 (1marker)	1320	48.10	1.74
16x16 (2 markers)	2640	46.48	1.74
16x8	2640	47.67	1.46
8x8	5280	46.82	1.30

quality experts group on the validation of objective models of video quality assessment", VQEG meeting number 4, Ottawa, Canada, March, 2000.

- [3] 大西, 岡, 松井, "PN 系列による画像への透かし署名法", 暗号と情報セキュリティシンポジウム SCIS97-26B, Jan.1997
- [4] 松井, "電子透かしの基礎", pp.78-82, 森北出版, 1998
- [5] G Depovere et al., "The VIVA Project: Digital Watermarking for Broadcast Monitoring", ICIP99 26AP1.1, Oct.1999