

LJ-010

低速監視回線に適用可能な特徴量抽出型画質監視方式

Study on a method for monitoring video quality applicable to low bitrate reference path based on extraction of image features

†杉本 修
Osamu Sugimoto

†川田 亮一
Ryoichi Kawada

†小池 淳
Atsushi Koike

†和田 正裕
Masahiro Wada

1. はじめに

テレビ伝送のデジタル化の普及とともに、デジタル映像伝送における品質管理技術への要求が高まってきている。特に、放送局間を結ぶ素材伝送系では、送受信される映像は番組素材として利用されるため、画像符号化や伝送障害による受信画の劣化を監視することが強く求められている。従来、これらの伝送監視はネットワーク層や MPEG-2 TS(Transport Stream)層におけるエラー率測定により行われていたが、近年では、ベースバンド映像を自動監視することにより、より正確な映像品質測定を可能とする自動画質評価方式が注目されている。

中でも、テレビ伝送網向けの画質監視方式として有力視されているのが、特徴量抽出方式 (RR 方式; Reduced Reference method) である。RR 方式は、画像特徴量を自動的に抽出し、遠隔の監視地点まで映像回線とは別のデータ回線により伝送することで、遠隔自動監視を実現するフレームワークであり、現在 ITU でも標準化の対象として注目されている[1]。筆者らは先にこの RR 方式による自動画質評価方式について提案を行った[2]。しかし、提案方式の実装には、画像特徴量の伝送のために高速なデータ回線が必要となるため、従来から利用されてきた機器の遠隔制御のための低速回線を監視回線として利用することができないという問題があった。そこで、本稿では、これら低速回線において RR 方式を適用するための特徴量抽出方式について検討する。本検討では、従来方式において送受信画像の各画素ブロックから数ビットのデータ幅で抽出していた画像特徴量を、1 ビットの情報に置換することにより、情報量の削減を実現する。計算機シミュレーションにより、従来方式の 1/8 の監視回線容量で従来方式とほぼ同等の精度での画質監視が可能であることを示す。

2. 従来方式のフレームワークと問題点

2.1. RR 方式のフレームワーク

RR 方式の概要を図 1 に示す。図 1 は、送受信局間の映像伝送を示しており、伝送の過程で圧縮符号化が適用され、受信画像には符号化劣化が印加される。送受信側それぞれの映像信号は A, B 点で映像伝送の本線から分離され、画像特徴量 $F_A[i]$, $F_B[i]$ が抽出される。画像特徴量は映像回線とは別に用意されたデータ回線である監視回線を経由して中央監視室に送出される。中央監視室では、A, B 点の画像特徴量を比較することにより、客観画質を求める。RR フレームワークでは、限られた監視回線速度のなかでいかに効率よく画像特徴量を抽出することが重要となる。

2.2. 従来方式による特徴量抽出方式

従来方式による画像特徴量の抽出手順を図 2 に示す。画素

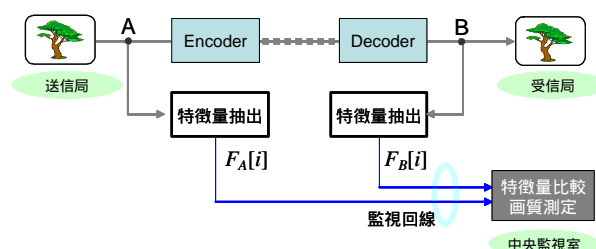


図1 特徴量抽出方式のフレームワーク

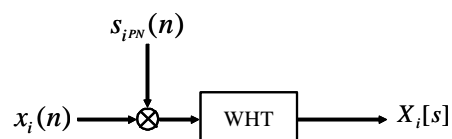


図2 特徴量抽出方式

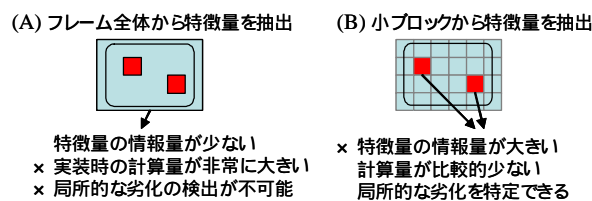


図3 ブロックサイズの設定

ブロック毎に分割された画像信号とPN系列との乗算によるスペクトル拡散とWHTによる直交変換により得られた係数のうち、特定の成分を抽出することにより画像特徴量は求められる。任意のサイズの画素ブロックに分割された入力信号を $x_i(n)$ とすると、ブロック i における特徴量係数 $F[i]$ は以下により求められる。

$$X_i[s] = \text{WHT}\{x_i(n) \times s_{iPN}(n)\} \quad (1)$$

$$F[i] = \text{round}(X_i[s_0] / Q) \quad (2)$$

ここで、 s_0 は抽出対象のシーケンス成分、 Q は $F[i]$ を有限長のデータで表すために行う丸めのためのパラメータである。 Q の値は、ブロックサイズおよび特徴量を何ビットで表現するかによって変化する。受信画像の各フレームのMSEは、 $F_A[i]$, $F_B[i]$ を図1のA, B点における画像特徴量とすると、以下により与えられる。

$$MSE = \sum_{i \in \text{frame}} \{F_A[i] - F_B[i]\}^2 / N_B / \text{scale} \quad (3)$$

ここで、 N_B は特徴量を抽出したブロックの総数、 scale は前述の丸めパラメータ Q に応じたスケール定数である。

画像信号は、一般にスペクトルの偏りを有しているため、特定の直交変換係数成分のみを抽出してもMSEを精度よく推定することはできない。そのため、多数の成分を抽出しなければなら

ず、それゆえ特徴量伝送のための監視回線容量が増大してしまふ問題があった。しかし、本方式においては、スペクトル拡散の効果により、画像信号はランダム化され、スペクトルの偏りが解消されるため、特徴量として一つまたは少数の成分のみを抽出するだけで高精度な画質測定が可能となる。

2.3. 従来方式の問題点

前述の通り、従来方式はスペクトル拡散による画像スペクトルのランダム化を利用して、少量の特徴量から高精度に客観画質を推定している。拡散の効果、すなわちランダム化の度合いは、ブロックサイズが大きいくほど高く、特徴量の抽出が少数で済むため、ブロックサイズの選択としては、図 3(A)のようにフレーム全体を 1 ブロックとみなしてそこからいくつかの係数を抽出することが理想的である。しかし、このようにブロックサイズを拡大させることは、同時に局所的な劣化の検出を困難にしたり、特徴量抽出のための計算量を増加させることにつながる。

一例として、ブロックサイズによる計算量の差異について考察する。 n 点の WHT に要する計算量のオーダが $O(n \log n)$ であるとすると、フレーム全体を 1 ブロックとするときのブロックサイズを N_{B1} 、小ブロックごとの抽出のときの 1 ブロックの画素数を N_{B2} とすると、それぞれのオーダは、

$$\text{フレーム全体の場合} : O_1 = N_{B1} \log N_{B1}$$

$$\text{小ブロックごとの場合} : O_2 = L \times N_{B2} \log N_{B2}$$

となる。なお、 L は、フレーム内の小ブロックの総数である。文献 [2] に示されたパラメータ $N_{B1}=1024 \times 256$ 、 $N_{B2}=16 \times 16$ 、 $L=660$ を用いて比較すると、 O_1 と O_2 の比はおおよそ 3.5 倍となる。計算量の増加は、実装時にハードウェア規模や処理負荷の増加を招くため、現実的には図 3 (B) のように、ブロックサイズをある程度小さくして処理を行う必要がある。しかし、ブロックサイズを小さくすると、抽出すべき特徴量の数が増えてしまい、結果として特徴量伝送のための監視回線の帯域が増大してしまうというトレードオフが従来方式では生じてしまう。このため、小ブロックサイズからの抽出が可能でかつ低速監視回線に対応する特徴量抽出方式が求められる。

3. 提案方式

提案方式は、1 ブロックあたりの特徴量の情報量を従来方式に比べて大きく減らすことにより、低速監視回線に対応する。従来方式は、式 2 に示されるように、式 1 で得られた WHT 係数成分を有限長に丸めることにより特徴量を求めていた。しかし、特徴量の情報量を減らすために丸め幅を大きくすることは、PSNR の推定精度に影響する。そのため、式 2 におけるパラメータ Q を大きくするには限界があると考えられる。

そこで、提案方式では、不可視マーカによるリファレンス画像なしの画質評価 [3] で採用されたリファレンス情報の付与方式を活用する。すなわち、次式のとおりに、各ブロックの特徴量を 1 ビットであらわす。

$$F[i] = \{\text{round}(X_i[s_0] / M)\} \bmod 2 \quad (4)$$

ここで $X_i[s]$ は式 1 で定義されるスペクトル拡散後の画像の WHT 成分、 M は任意の量子化パラメータである。不可視マーカ方式では、このリファレンス情報を画像信号中に埋め込んでいたが、提案方式ではこの 1 ビット情報を送受信画像から抽出し、監視回線により中央監視室まで伝送する。監視室では、送受信間の特徴量の比較により当該ブロックの劣化度を判定する。

● $X_i[s_0]=800$, $M=200$ の場合

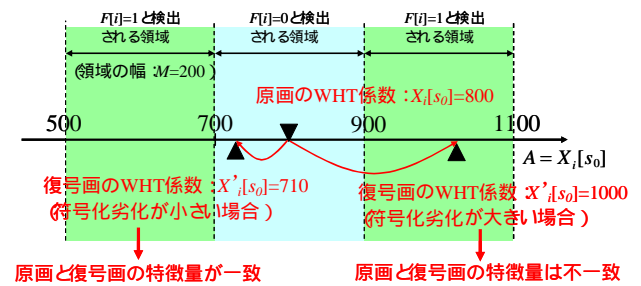


図 4 特徴量比較による劣化度の推定

図 4 に特徴量の比較と劣化度の関係を示す。同図では、原画の WHT 係数 $X_i[s_0] = 800$ 、特徴量抽出のパラメータ $M=200$ の場合を表している。1 ビットの特徴量は式 4 に基づき、 $M=200$ の周期で $F[i]=0$ または $F[i]=1$ が与えられる。同図の例では、原画の特徴量は、 $X_i[s_0] = 800$ であるため、 $F[i]=0$ と認識される。一方、復号画像における WHT 係数成分は、符号化劣化などの伝送劣化が加えられるため、一般に原画の WHT 係数とは異なる。ただし、図 4 (A) のように符号化劣化が小さい場合には、WHT 係数 $X_i[s_0]$ の変動が小さく、原画と同じく $F[i]=0$ の領域の中に収まると考えられる。一方、図 4 (B) のように符号化劣化が大きい場合には、WHT 係数 $X_i[s_0]$ の変動が大きくなり、原画とは異なる特徴量 $F[i]=1$ を示すことが考えられる。よって、原画と復号画の特徴量が一致するかどうかを各ブロックごとに検査することにより復号画像の伝送劣化の程度を検出可能であることがわかる。なお、パラメータ M により与えられる $F[i]=0,1$ の周期 (領域の幅) が小さければ、劣化に対する感度が高くなり、逆に M の値が大きければ、感度が下がることになる。そのため、パラメータ M は適切に選ぶ必要がある。

提案方式では、フレーム内の劣化度の指標として、特徴量ビットの相違率を用いる。ビット相違率以下により定義される。

$$FDR = \sum_{b=0}^{N_B-1} |F_A[b] - F_B[b]| / N_B \quad (5)$$

このビット相違率と MSE の間には以下の関係が存在することが知られている [3]。

$$\frac{1}{MSE} \propto \frac{\ln(FDR)}{M} \quad (6)$$

式 6 は、 $\ln(FDR)$ と MSE の間に相関があることを示す。よって、予備実験により予め FDR-MSE 特性を求めておけば、以降は特徴量の比較により FDR が得られさえすれば PSNR の推定が可能となる。

4. 計算機シミュレーション

4.1. 実験条件

6 種類のテスト画像を MPEG-2 TM5 で符号化し、提案方式により復号画像の PSNR を推定する実験を行った。実験条件を表 1 に示す。本実験では、特徴量抽出のブロックサイズを変化させることにより、監視回線速度を 9.9kbps ~ 158kbps の幅で変化させ、輝度成分の 15 フレームの PSNR 平均値を推定する。また、提案方式においては、1 フレーム単位での PSNR 推定値を求めることが可能であるが、本実験では VQEG による RR 方式評価テスト RRNR-TV での要求 Q サンプル / 秒 [4] に基づき、15 フレーム平均の PSNR を求めることとする。また、特徴量

抽出のためのパラメータ M は、予備実験によりブロックサイズ 8×8 において PSNR 推定誤差を最小とする最適値を求め、他のブロックサイズでも共通の値を用いた。また、従来方式として、2.2 節の WHT 係数成分の丸めによる特徴量抽出方式による PSNR 推定もあわせて行った。従来方式による 1 ブロックあたりの特徴量の情報は 8 ビットとし、提案方式のブロックサイズと同じブロックサイズを適用した。すなわち、従来方式は提案方式の 8 倍の監視回線帯域を使用することとなる。

4.2. 実験結果

各ブロックサイズの PSNR 推定誤差平均値を表 2 に示す。ここで、推定誤差とは、各シーケンス、各ビットレートにおける推定誤差の絶対値の平均を意味する。図 5 は、ブロックサイズ 8×8 の場合の FDR-PSNR 特性である。縦軸の PSNR とは原画と復号画の輝度値から求めた真の PSNR を示し、横軸の $-\ln(\text{FDR})$ は、式 5 を用いて得られたビット相違率の対数を示している。実際の推定においては、真の PSNR を知ることはできないため、FDR-PSNR 特性から得られた近似曲線を用いて復号画像の PSNR を推定する。図 5 の例でいうと、近似曲線は

$$\text{PSNR} = 14.19 \ln(-\ln(\text{FDR})) + 25.32 \quad (7)$$

となる。推定誤差とは、これらプロット点と近似曲線との距離の平均によって定義される。なお、式 7 の近似式は、ブロックサイズが変わった場合には、新たに別の近似式を求めることとする。

表 2 の提案方式および従来方式における PSNR 推定誤差について考察する。従来方式では、ブロックサイズの拡大に伴い推定誤差が拡大しているが、提案方式では逆にブロックサイズが大きいほど推定誤差が改善するという傾向が見られる。一般に、リファレンスとなる特徴量の情報は多いほど精度が高くなると考えられる。しかし、提案方式による画質推定は、スペクトル拡散による周波数成分のランダム化の効果に依存しており、拡散の効果はブロックサイズが大きいほど高いため、情報量の増加よりスペクトル拡散によるランダム化の効果のほうが推定精度の改善に寄与しているものと考えられる。

従来方式と提案方式の推定精度を比較すると、情報量が高い分だけ従来方式の推定誤差がわずかに提案方式を下回っているが、ブロックサイズ 64×16 では、監視回線速度が $1/8$ となっても推定誤差の差は 0.076dB とごく僅かであり、実用上ほとんど差がないといえるため、提案方式による情報量削減の効果を確認できる。ただし、これは 15 フレームという比較的長い期間の平均 PSNR を推定しているからであり、1 フレーム単位の PSNR 推定精度を比較した場合には、やはり監視回線速度の高い従来方式が有利であると考えられる。よって、提案方式の適用には若干の制限 (サンプル / 秒程度の間隔での評価値の出力のときに適用可能) があるといえるが、監視回線速度に制限がある場合には有効な方式である。

5. おわりに

特徴量抽出型のフレームワークに基づく自動画質監視方式を低速監視回線環境において適用するための特徴量抽出方式について検討した。送受信画像の各画素ブロックから 1 ビットの特徴量を抽出し、これを比較することによって得られるビット相違率をもとに、復号画像の PSNR を高精度で推定可能であることを示した。提案方式は、 64kbps 以下といったごく低速な監視回線環境下においても、15 フレーム平均の PSNR 推定誤差が約 0.34dB という実用上十分な精度で画質監視を行うことが可能であるため、機器制御用に設けられた低速監視回線のみを

表 1 実験条件

テスト画像	Cheerleaders, Flamingoes, Green Leaves Marching in, Mobile and Calendar, Soccer Action (30frames)
画像サイズ	704(H)×480(V) 4:2:0 30fps
コーデック	MPEG-2 Test Model 5
ビットレート	6.9,12,15 Mbps (MP@ML)
画質評価尺度	PSNR(15フレーム平均)
ブロックサイズ	8×8 , 16×8 , 16×16 , 32×16 , 64×16
特徴量抽出 パラメータ M	$M=200$ (ブロックサイズ 8×8 , スケーリングなし WHT係数に対して)

表 2 各ブロックサイズにおける PSNR 推定精度

ブロックサイズ	従来方式		提案方式	
	監視回線速度 [kbps]	推定誤差 [dB]	監視回線速度 [kbps]	推定誤差 [dB]
8×8	1276.2	0.199	158.4	0.361
16×8	633.6	0.221	79.2	0.388
16×16	316.8	0.223	39.6	0.418
32×16	158.4	0.222	19.8	0.339
64×16	79.2	0.252	9.9	0.338

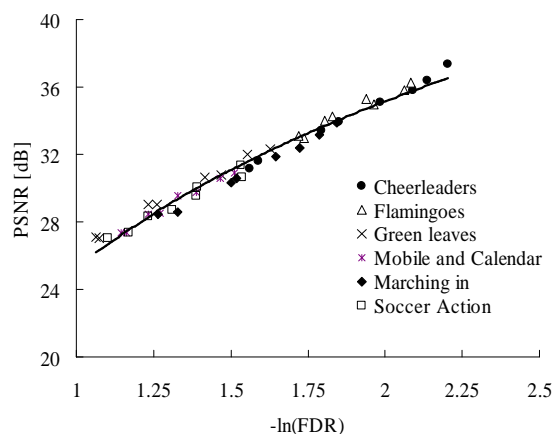


図 5 FDR-PSNR 特性 (ブロックサイズ 8×8)

有するような既存のテレビ伝送系へも適用可能であることを示すことができた。

今後の課題としては、本検討では伝送劣化として、MPEG 符号化による劣化のみを前提としていたが、それだけでなく、伝送障害による映像キズによる劣化の検出を可能とする監視方式の確立が挙げられる。

参考文献

- [1] Recommendation of ITU-T J.143, "User requirements for objective perceptual video quality in digital cable television," (May 2000)
- [2] R. Kawada, O. Sugimoto, M. Wada, S. Matsumoto, "Precise monitoring of remote picture PSNR by extracting reference data," IEEE ISPCS 2001 (Nov.2001)
- [3] 杉本, 川田, 和田, 松本, "WHT を用いた不可視マージ力埋め込みによる画質評価方式," 情処研報 2001-AVM33-26 (June 2001)
- [4] Video Quality Expert Group, "RRNR-TV Test Plan version 1.4," <http://www.vqeg.org/>