

特徴量抽出方式による IP ベース映像配信サービスの自動画質監視方式

Automatic video quality monitoring system for IP-based television transmission service based on reduced reference method

杉本 修† Osamu Sugimoto
川田 亮一† Ryoichi Kawada小池 淳† Atsushi Koike
松本 修一† Shuichi Matsumoto

1. まえがき

近年の FTTH に代表される高速アクセス回線の普及により、IP ネットワークを利用した放送サービスが今後本格的に普及していくと予想されている。これらの IP ベース映像配信サービスは、アクセス回線の高速性・低廉性を活用したサービスが提供できる一方、パケットロスなどの伝送路の品質劣化により受信画質が変動するという特性を合わせ持っている。しかし、放送サービスにおいては常時劣化のない映像を提供することがサービスプロバイダには求められるため、この画質変動を正確に監視し、映像の劣化が激しい場合には対策を講ずることがサービス品質の管理上非常に重要となる。従来、映像の監視は、ネットワーク運用者が中継点や受信点の映像を目視により見続ける人的監視により行われてきた。しかし、この方法は運用者に相当の負担を強い、複数の地点・チャンネルの監視には相応の人数が必要となる。今後、ネットワークのさらなる大容量化による多チャンネル化が進んだ場合、人的監視は限界を迎えると考えられている。それゆえ、これを自動化する技術への要求が高まってきており、ITU などでも標準化のための検討が行われている[1]。

そこで本稿では、IP ベース放送サービスにおける受信映像品質の自動監視を目的として、特徴量抽出方式を利用した画質自動監視方式を提案する。提案方式は、映像配信ネットワーク上の任意の 2 地点の受信映像から画像特徴量を自動抽出し、これを比較することにより、ネットワーク上の任意のリンク間の映像品質を測定する。この際、映像配信ネットワークの上り帯域を利用し、画像特徴量をネットワーク運用者に伝送することにより、遠隔自動監視を可能としている。提案方式により、数十 kbps の上り回線帯域を使用するだけで、パケットロスなど伝送路の品質変動による受信画質の劣化を高精度に検出可能であることを示す。

2. IP 放送サービスにおける映像品質評価

2.1. 映像配信ネットワークの構成

本稿で想定するコンテンツ配信ネットワーク(CDN; Contents Delivery Network)の構成を図 1 に示す。図 1 は映像の送出側、受信側に分けられる。送出側では、映像コンテンツを MPEG-2 圧縮して得られた TS パケットを RTP[2]パケットにマッピングし、エラー訂正情報を付加した後、CDN にマルチキャストで送出する。送出されたマルチキャストパケットは CDN 内で中継され、最終的にユーザ宅のレジデンシャルゲートウェイ(R-GW)を経由してセットトップボックス(STB)がこれを受信する。STB は RTP パケットから TS のストリームを再構成し、MPEG-2 デコードを行い、再生画像を TV に表示する。なお、パケットロスが発生した場合には、STB はエラー訂正情報に基づき損失パケットの回復を試みる。

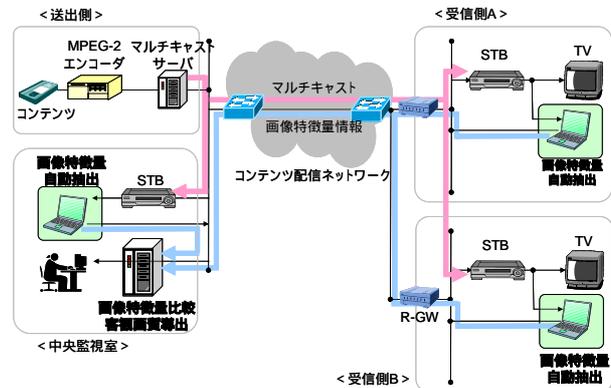


図 1 コンテンツ配信ネットワークの構成

2.2. 映像劣化の発生要因

マルチキャストはコネクションレスな通信方式であり、伝送中にエラーが発生しても再送を行わない。そのため、パケットロスが発生した場合、ビットストリーム情報が欠落してしまい、復号画像に劣化が生じる。通常、映像をマルチキャストで配信する場合には、経路上のルータにおいて映像トラフィックを優先させたり、RTP パケットを構成する際にエラー訂正情報を付加させたりするなどの対策を行うが、これらの手段でもパケットロスおよびそれに伴う映像劣化を完全に排除することはできない。そこで、次節に示すような自動監視システムを提案する。

2.3. 映像監視システムの構成

提案方式では、CDN 内の任意の 2 地点の STB が出力する映像信号の各フレームから画像特徴量を抽出し、これを中央監視室で比較することにより 2 点間の受信画質 PSNR を測定する。測定対象の 2 地点の組み合わせとしては、図 1 に示されるように、中央監視室の STB 出力と受信側の STB 出力の比較が効果的である。これは、中央監視室側では CDN の直前で折り返されたエラーフリーのストリームを受信することが可能であるため、エラーを含んでいる可能性のある受信側の映像出力と比較することにより、受信画像の劣化の検出が可能となるからである。

中央監視室と測定対象の 2 地点はそれぞれ離れた場所に位置しているため、画像特徴量は CDN の上り帯域を用いて中央監視室まで伝送される。映像配信ネットワークでは、上り回線(受信側 中央監視室の方向)はマルチキャストが配信される下り回線に比べると帯域には余裕があるため、画像特徴量にはある程度の情報量を持たせることも可能である。しかし、中央監視室には複数の受信地点からの画像特徴量情報が集中するため、やはり 1 地点あたりの特徴量情報の帯域は低く抑える必要がある。そこで、提案方式では、筆者らが基幹映像伝送ネットワークの監視のために開発した低ビットレート監視回線用の特徴量抽出方式[3]を適用し、上り回線の使用量を抑制する。

†株式会社 KDDI 研究所, KDDI R&D Laboratories Inc.

3. 特徴量抽出に基づく PSNR 推定方式の概要

3.1. 特徴量抽出方式

提案方式による特徴量抽出方式の概要を示す。提案方式では、受信画像の各フレームの映像信号を 8×8 画素ブロックに分割し、1 ブロックから 1 つの特徴量係数を抽出する。特徴量は、画素ブロックの信号 $x_i(n)$ に対してスペクトル拡散およびウォルシュ - アダマル変換(WHT)を反復実行して得られた係数のうち、成分 s_0 の係数を抽出し、さらにこれをパラメータ M で量子化した代表値のパリティにより与えられる。すなわち、式 3 の $F[i]$ のとおり定義される。

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

$$Y_i[s] = \text{WHT}\{X_i[k] \times s_{iPN}[k]\} \quad (2)$$

$$F[i] = \{\text{round}(X_i[s_0] / M)\} \bmod 2 \quad (3)$$

ここで、 $s_{iPN}(n)$ 、 $s_{iPN}[k]$ はスペクトル拡散に使用される PN 系列、 $\text{WHT}\{\}$ はウォルシュアダマル変換のオペレータである。式 3 より、画像特徴量は (0,1) のいずれかの値を持つ 1 ビットの情報となる。このように、画像特徴量を最小限の情報量に抑制することで、特徴量伝送のための帯域を大幅に減少させることが可能となる。

3.2. PSNR 推定方式

中央監視室では、送信側、受信側双方から抽出された画像特徴量を受信し、両者を比較することにより受信側画像の PSNR を推定する。画像特徴量は受信画像の周波数成分の値に基づくため、2 地点間の画像信号の差が大きい場合には、特徴量の差分も大きくなると予想される。そのため、まずフレーム内の画像特徴量の差分平均を求める。なお、前述のとおり画像特徴量は 1 ビット情報であるため、実際にはフレーム内のビット相違率を求めることとなる。送信側、受信側の特徴量をそれぞれ $F_T[i]$ 、 $F_R[i]$ とすると、ビット相違率 FDR は、

$$\text{FDR} = \sum_{b=0}^{N_B-1} |F_T[i] - F_R[i]| / N_B \quad (4)$$

により与えられる。ここで、 N_B はフレーム内のブロックの総数であるとする。FDR と受信画像の PSNR の間には、絵柄やビットレートによらない相関があることが知られているため [3]、FDR をもとに受信画質を推定することが可能となる。

3.3. 特徴量抽出機能の Windows PC への実装

前述の画像特徴量の抽出および送出機能について、Windows PC へのソフトウェア実装をおこなった。なお、STB の出力映像を入力するため、Windows PC にはビデオキャプチャデバイスを追加している。なお、STB の映像信号はアナログコンポジット信号であるため、画像の取り込みの際に A/D 変換や解像度変換などによる劣化が生じるが、本稿で想定するシステム構成では、同一仕様の STB の出力を比較することを前提としているため、この劣化は映像品質測定に対して大きな影響を与えないと考える。

4. シミュレーション実験

4.1. 実験ネットワーク構成

実験ネットワークの構成を図 2 に示す。本実験では、subnet1, subnet2 という 2 つのサブネットを用意し、subnet1 に

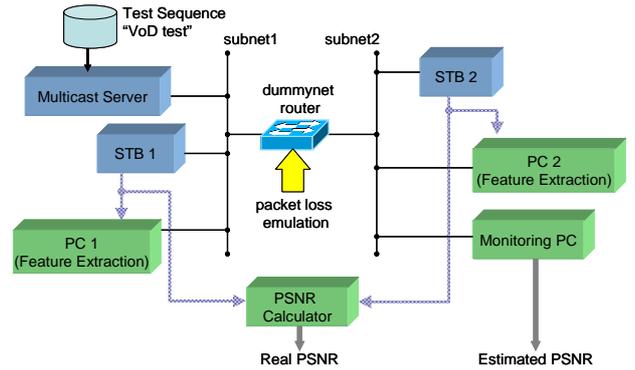


図 2 実験用ネットワークの構成

表 1 シミュレーション条件

テスト画像	vod_drama, vod_festa
フォーマット	720 × 480, YUV 4:2:0, 30fps
コンテンツ長	160秒
エンコーダ	商用MPEG-2エンコーダ
符号化方式	Video: 4Mbps(MP@ML), Audio: なし
復号画像のキャプチャサイズ	320 × 240, YUV 4:2:2, 30fps
特徴量ビットレート	1200bit/frame (36kbps)

映像配信サーバを接続する。配信コンテンツの符号化パラメータ等の条件は表 1 のとおりである。本実験では、4Mbps で符号化された映像のみを送送するものとし、RTP パケットにエラー訂正などの情報は付加していない。STB は各サブネットに 1 台ずつ接続され、サブネット間は dummynet ルータにより接続される。dummynet ルータは映像トラフィックを中継する際に所定の確率でパケットロスを発生させる。よって、STB1 は図 1 の中央監視室の STB に相当し、STB2 は、受信側の STB に相当する。

PC1,2 は STB の映像出力をキャプチャし、画像特徴量を Monitoring PC に送出する。ここで、STB が受信する MPEG-2 ストリーム自体は 720 × 480 の解像度を持っているが、PC1,2 でキャプチャする際にはこれを 320 × 240 にダウンコンバートしている。これは、STB のアナログ出力の解像度がそれほど高くないため、解像度を下げても PSNR 測定精度が大きく落ちないと考えられることと、キャプチャ解像度を下げることにより、特徴量の情報量を 36kbps に抑制でき、さらに、処理ブロック数の減少により、PC の負荷を低減できるという理由に基づく。なお、特徴量を受信する Monitoring PC は subnet2 に接続されるため、PC1 からの特徴量情報はルータを経由して送出されるが、ルータでは映像トラフィックに対してのみロスを発生させるため、PC1 からの特徴量情報についてはエラーなしで Monitoring PC まで伝送されることに注意されたい。Monitoring PC では、受信した特徴量から式 4 に基づきビット相違率 FDR を求める。さらに PSNR の真値を求めるため、PSNR Calculator においても 2 台の STB の映像出力をキャプチャし、画素単位の比較により真の PSNR を求め、Monitoring PC における FDR との相関を求められるように構成する。

4.2. 実験結果

まず、PSNR Calculator で得られる真の PSNR とパケットロス率の関係を図 3 に示す。パケットの欠落は、subnet2 側の RTP パケットのシーケンス番号の連続性を見て判定しており、コンテンツの先頭から 10 秒単位の区間での欠落パケット数をもとにロス率を求めている。図 3 より、パケットロスが発生していない区間では PSNR が 40dB 以上の値を維持していることがわかる。本来、subnet2 においてパケットロスが発生していなければ、STB1,2 から完全に同一の復号画が得られるため、真の MSE 値は 0 となり、PSNR= となるべきであるが、本実験では、STB のアナログ出力を評価しているため、実際にキャプチャされる画像は多少異なっており、PSNR はエラーなしであっても 40~80dB 程度の有限値となっている。しかし、パケットロスが発生している区間では PSNR は 40dB よりさらに低下しているため、パケットロスなしで MSE=0 とならなくても PSNR を監視することで、十分に画質劣化を判定することが可能であるといえる。

以上は画素単位で PSNR を測定した場合の結果であるが、特徴量の比較に基づき PSNR を精度よく推定することが可能であれば、提案方式も画質監視に有効であることが確認できる。そこで、次に提案方式の PSNR 測定精度について調べる。図 4 は式 4 により求めた各フレームの特徴量ビット相違率 (FDR) と PSNR と真の PSNR の関係を示したものである。同図より、両者は $PSNR=SNR=8.662\ln(FDR)+9.10$ という関数で近似できることがわかる。筆者らは、画像劣化が主に圧縮符号化により発生すると前提するとき、FDR と PSNR は $PSNR=A \ln(-\ln(FDR))+B$ (A, B は定数) なる関係で近似できることを以前に示しているが[4]、本実験ではこれとは異なる関係が得られたこととなる。これは、本実験では、比較の対象となっているのが符号化の前後の画像ではなく、エラーの発生の有無を除いてはほぼ同一のストリームを復号した場合の出力であるための違いと考えられる。具体的には、符号化劣化を推定する場合には、復号画像の周波数成分の劣化モデルとしてラプラス分布を前提としていたが、今回はこの劣化モデルが一様分布に近いために相関の関係がわずかに変わっていると考えられる。

最後に近似の精度について言及すると、図 4 における FDR-PSNR 間の相関係数は 0.945 と非常に高い。よって、提案方式の PSNR 近似の精度は非常に高く、それゆえパケットロスによる劣化検出の精度も非常に高いということが確認された。

5. むすび

画像特徴量の自動抽出方式に基づく IP ベース映像配信サービスの自動画質監視方式について検討した。提案方式は、スペクトル拡散と WHT を反復実行して得られた直交変換係数の 1 ビットパリティ情報を特徴量とすることで、伝送すべき特徴量の帯域を大幅に低減させることが可能である。また、映像ストリームの送信側折り返し画像と受信側画像から提案方式により画像特徴量を抽出し、その PSNR を測定することにより、パケットロスによる復号画像の劣化を正確に検出することが可能であることを示した。また、WindowsPC ベースの測定システムにより、特徴量の伝送レート 36kbps の条件のもとで受信画像の PSNR を推定した場合、相関係数 0.945 という高い精度で測定が可能であることを示した。

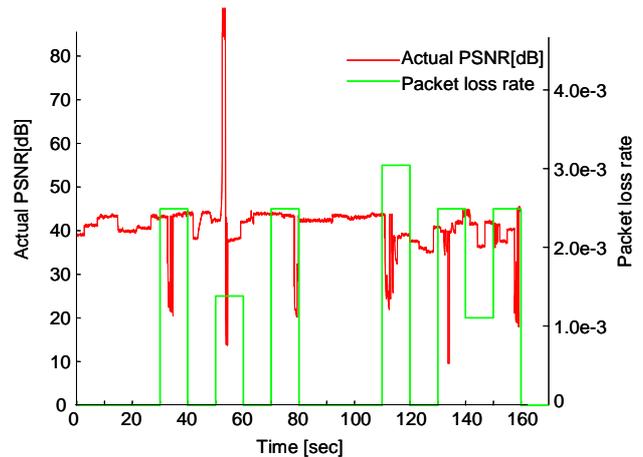


図 3 PSNR とパケットロス率の関係

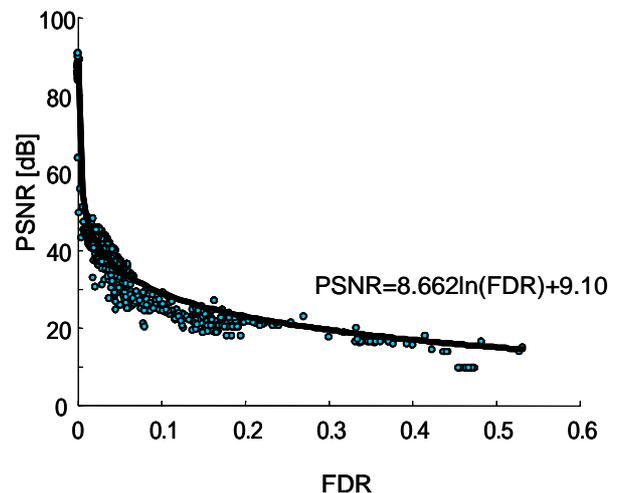


図 4 特徴量ビット相違率 FDR と PSNR の関係

(相関係数 0.945)

今後の課題としては、FEC やインターリーブなどのエラー耐性ツールを実装したストリームに対しての性能測定があげられる。

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] RFC 3550, "RTP: A Transport Protocol for Real-Time Applications" (July 2003)
- [3] 杉本, 川田, 小池, 和田, "低速監視回線に適用可能な特徴量抽出型画質監視方式", FIT2003 LJ-11 (Sep. 2003)
- [4] 杉本, 川田, 小池, 松本, "画像特徴量の多値表現に基づく遠隔自動画質監視方式の高精度化の検討", 映像学 2003 冬季大会, 9-4 (Dec. 2003)