

サブサンプリングを用いたマルチストリーム動画像伝送方式の検討

村松宏基[†] 藤井俊章[†] 谷本正幸[†]

[†] 名古屋大学大学院工学研究科電子情報学専攻 〒464-8603 名古屋市千種区不老町

E-mail: [†]{hmurama,fujii}@tanimoto.nuee.nagoya-u.ac.jp, ^{††}tanimoto@nuee.nagoya-u.ac.jp

あらまし 本研究では、画像をサブサンプリングすることにより相関を持った複数のストリームに分割し、動画像のエラー耐性能力を向上させる方式を提案した。提案方式は、あるストリームにエラーが生じた場合、分割した画像間の相関を利用して失われた画素を予測することができる。通信路で生じるエラーを想定した実験の結果、提案方式は従来方式に比べてエラーの影響を受けにくく、劣悪な通信環境下において有効であることを示した。

キーワード 動画像通信, Multiple Description Coding, サブサンプリング

Multi-stream Video Transmission System Using Subsampling for Error Resilience

Hiroki MURAMATSU[†], Toshiaki FUJII[†], and Masayuki TANIMOTO[†]

[†] Department of Information Electronics, Graduate School of Engineering, Nagoya

University Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8603 Japan

E-mail: [†]{hmurama,fujii}@tanimoto.nuee.nagoya-u.ac.jp, ^{††}tanimoto@nuee.nagoya-u.ac.jp

Abstract We propose a new video transmission method for error resilience. The proposed method divides the image into some correlative streams by using subsampling. If the error occurs, the proposed method restores lost pixels by using the correlation between the divided images. The experiment with the transmission error shows that the proposed method is more robust than the conventional method and effective in the ignoble transmission environment.

Key words Video Transmission, Multiple Description Coding, Subsampling

1. はじめに

移動体通信やインターネット上において、マルチメディアデータをリアルタイムに通信することへの要求が高まっている。しかし、移動体通信でのフェージングの影響やインターネット上での混雑時のパケットの破棄などにより、伝送路では高い比率でエラーが生じるため、動画像通信での品質の劣化が大きくあらわれる。これは、動画像データは一般に符号化効率をあげるためにフレーム間差分符号化を用いているため、エラーが伝搬しやすいためである。このため、動画像通信におけるエラー耐性のための通信方式、符号化方式などの検討が重要となる。

エラー耐性を向上させる方法の1つに、Multiple Description Coding [1] がある。これは、データを Description と呼ばれる相関を持った複数のストリームに分割してそれぞれを独立に伝送するものである。このように、相関を残すようにデータを分割することは、情報圧縮の観点からすると反していることになるため符号化効率は低下するが、ある Description にエラーが生じた場合には他の Description との相関を利用して失われたデータを修復することができるため、データに誤り耐性を持たせることができる。本稿ではこの Multiple Description Coding に注目し、その動画像への実現方法として、サブサンプリングを用いることにより画像を複数の相関を持ったストリームに分割し、動画像データのエラー耐性能力を向上させる方式を提案する。

2. 動画像符号化の概要

ここでは、現在の動画像圧縮の主流である DCT(離散コサイン変換)+動き補償を用いた動画像の符号化について述べる。符号化の基本的な考えは、主に次の4つのものからなる。

- フレーム間符号化による、時間方向の冗長性の削減
- DCT による空間的情報圧縮
- DCT 係数の量子化ステップ制御による、発生符号量の制御
- DCT 係数、動きベクトルの発生確率の偏りを利用したエントロピー符号化

動画像の符号化には、動き補償を行い時間方向の予測を行うフレーム間符号化と、時間方向の予測を行わないフレーム内符号化がある。フレーム内符号化はフレーム間符号化に比べて符号量が多いため通常はフレーム間符号化を行うが、エラーが生じた際にその伝搬を防ぐ等の理由で定期的にフレーム内符号化を行う必要がある。

図1に動画像符号化器の基本構成を示す。まず、動き補償を行い他フレームとのフレーム間予測を行うことより時間方向の冗長性を削減する。フレーム間予測によって得られた差分画像は離散コサイン変換され、量子化される。量子化された DCT 係数、動きベクトル等の情報は、可変長符号を用いてエントロピー符号化される。

2.1 フレーム構造

入力された画像は、 16×16 画素のマクロブロックと呼ばれるものに分割される。マクロブロックは 8×8 画素の4つの輝

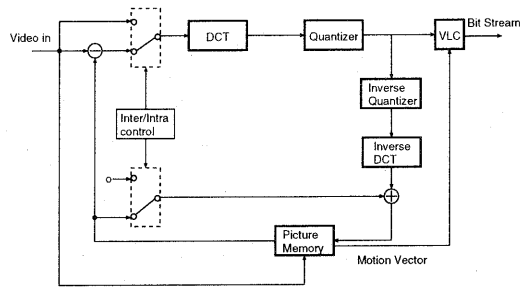


図1 動画像符号化器の基本構成

Fig. 1 Conventional Video Encoder

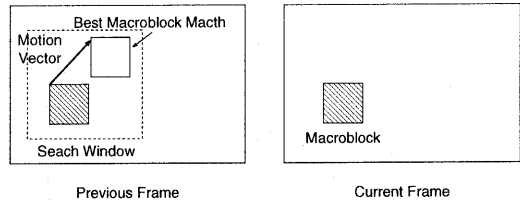


図2 動き補償

Fig. 2 Motion Compensate

度信号ブロックと縦、横方向にダウンサンプリングされた2つの色差信号ブロックから構成される。画像は、GOB(Group of Picture) または slice と呼ばれる複数のマクロブロック群に分割され、GOB(or slice) 単位で独立に復号できるように符号化される。

2.2 符号化のための技術

2.2.1 動き補償

フレーム間符号化では現在のフレームと前フレームとの差分がとられるが、フレーム間の動きを検出し、水平、垂直方向にずらして差分をとることを動き補償という。通常、動き補償はマクロブロック単位で行われる。

動き補償の様子を図2に示す。動き情報は水平成分、垂直成分の二次元で表現され、これを動きベクトルと呼ぶ。動きベクトルはブロックマッチングを用いて推定され、現在のフレームのマクロブロックと前のフレームの候補となるマクロブロック間の mismatches が最小となる位置を求めることによって得られる。 mismatches を測る尺度としてはいくつかのものがあるが、多くの場合二乗誤差 (MSE: Mean Square Error) が用いられ、次のように定義される。

$$MSE = \sum_{k=1}^{16} \sum_{l=1}^{16} \{B_{i,j}(k,l) - B_{i-u,j-v}(k,l)\}^2 \quad (1)$$

ここで、 $B_{i,j}(k,l)$ は現在のフレームの (i,j) にあるマクロブロック内の (k,l) 画素、 $B_{i-u,j-v}(k,l)$ は参照フレームの $(i-u,j-v)$ にあるマクロブロック内の (k,l) 画素を示している。

通常、 ± 15 画素の探索範囲を設けて二乗誤差を計算し、探索範囲内で二乗誤差が最小となる (u,v) を動きベクトルとする。

2.2.2 離散コサイン変換 (DCT)

原画像、または動き補償によって得られた差分画像は、 8×8 画素のブロックに分割され DCT がされる。これにより、低周波成分に信号が偏ることになる。 8×8 DCT は次のように定義される。

$$C_{m,n} = \alpha(m)\beta(n) \sum_{i=1}^8 \sum_{j=1}^8 B_{i,j} \cos\left\{\frac{\pi(2i+1)m}{16}\right\} \cdot \cos\left\{\frac{\pi(2j+1)n}{16}\right\}, \quad 0 \leq m, n \leq 7 \quad (2)$$

ただし、 $\alpha(0) = \beta(0) = \sqrt{1/8}$, $\alpha(m) = \beta(n) = \sqrt{1/4}$ である。ここで、 $B_{i,j}$ は 8×8 ブロックの (i, j) 画素、 $C_{m,n}$ は DCT 後の 8×8 ブロックの係数を示す。また、逆 DCT (IDCT) は次のように与えられる。

$$B_{i,j} = \sum_{m=1}^8 \sum_{n=1}^8 C_{m,n} \alpha(m) \cos\left\{\frac{\pi(2m+1)i}{16}\right\} \cdot \beta(n) \cos\left\{\frac{\pi(2n+1)j}{16}\right\}, \quad 0 \leq i, j \leq 7 \quad (3)$$

2.2.3 量子化

MPEG [2] では、量子化マトリクスを用いて 8×8 DCT のそれぞれの要素に対して対応する量子化値が割り当てられ、次のように量子化される。

$$C_{m,n}^q = \frac{C_{m,n}}{Q_{m,n}}, \quad 0 \leq i, j \leq 7 \quad (4)$$

ここで、 $C_{m,n}$ は (m, n) の DCT 係数、 $Q_{m,n}$ は (m, n) の量子化値、 $C_{m,n}^q$ は量子後の DCT 変換係数である。

また、H.26x [3] では、フレーム内符号化の直流成分に対してステップサイズ 8 の直線量子化、その他の係数については量子化特性に応じたデッドゾーン付きの量子化が行われる。

2.2.4 エントロピー符号化

量子化後の DCT 変換係数、動きベクトルは、符号化情報の発生確率の偏りを利用して、ハフマン符号に基づくエントロピー符号化が行われる。

量子化後の変換係数は、図 3 に示すように低周波から高周波に向かって二次元的にジグザグスキャンされ、一次元に並べられる。その後、ゼロ係数の継続個数 RUN とそれに続く非ゼロ係数の LEVEL の組合せに対して、発生頻度に応じて二次元可変長符号化される。

また、動きベクトルは水平成分、垂直成分の二つで表され、一つ前のマクロブロックの動きベクトルとの差分が可変長符号化される。

3. 提案方式

3.1 サブサンプリングを用いた動画のマルチストリーム化

提案方式の符号化器のブロック図を図 4 に示す。フレーム内符号化における原画像、フレーム間符号化における差分画像は、図 5 のようにサブサンプリングを用いて 2 つの画像に分割さ

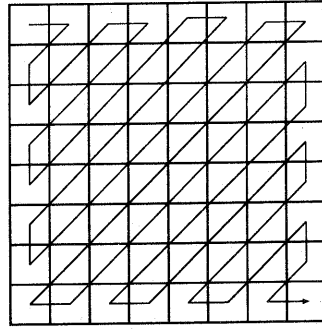


図 3 ジグザグスキャン

Fig. 3 Zigzag Scan

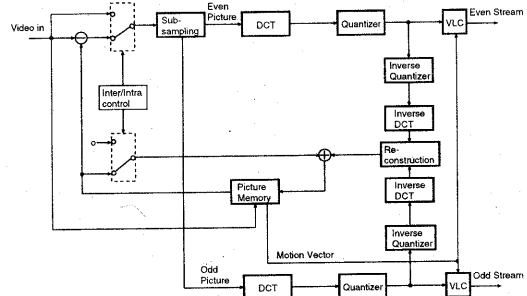


図 4 提案方式の符号化器のブロック図

Fig. 4 Proposed Encoder

れ、分割後は水平方向につめるような形でサイズが半分になる。分割された画像はそれぞれ DCT、量子化され、別のストリームとして伝送される。

復号側 (図 6) では、2 つのストリームが正しく受信された場合にはそのまま復号が行われる。仮にどちらか一方のストリームにエラーが生じ場合には、正しく受信できたストリームの画像から失われた部分を補間し復号される。補間は次式のように行われる。

$$B(i, j) = \{B(i-1, j) + B(i+1, j) + B(i, j-1) + B(i, j+1)\} / 4 \quad (5)$$

ここで、 $B(i, j)$ は画像内の (i, j) の画素値とする。odd picture にエラーがおこった場合の補間の例を図 7 に示す。

3.2 マルチストリーム化の意味

サブサンプリング後の分割画像は、隣の画素間の相関が低くなるため、DCT による圧縮効率の低下につながる。また、それぞれのストリームに共通の動きベクトル情報を付加するので、動きベクトルの情報量は 2 倍となる。このため、マルチストリーム化することで符号化効率の低下がおこる。しかし、分割した画像間には強い相関があり、どちらかのストリームにエラーが発生した場合にもう一方を補間することができるため、動画データに誤り耐性を持たせることができる。

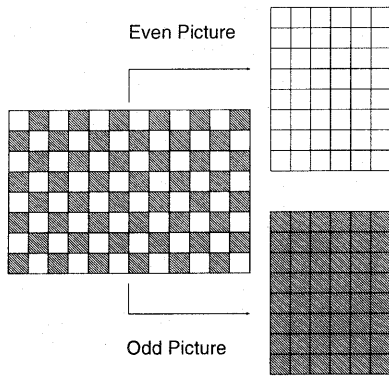


図 5 サブサンプリングによるマルチストリーム化

Fig. 5 Multistreaming Method Using Subsampling

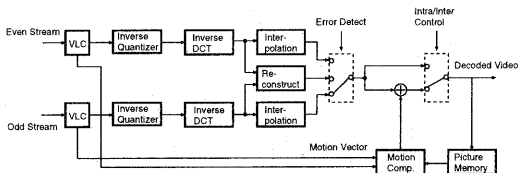


図 6 提案方式の復号器のブロック図

Fig. 6 Proposed Decoder

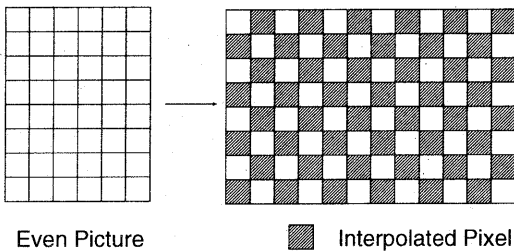


図 7 補間の例

Fig. 7 Example of Interpolation

4. 実験

4.1 実験方法

ここでは、伝送路においてエラーが発生することを想定し実験を行う。従来方式、または提案方式で符号化したデータに対して誤りを与え、それを復号した際の画質を評価する。このとき、量子化特性値を変化させることで符号化後の符号量を制御し、ビットレートと画質との関係を見る。実験では画質の客観的な評価尺度として Peak Signal-to-Noise Ratio (PSNR) を用いる。PSNR は次式で与えられる。

$$\text{PSNR} = 10 \log \frac{1}{M} \sum_{n=1}^M \frac{255^2}{(o_n - r_n)^2} \quad (6)$$

ここで、 M は画素数、 o_n, r_n は原画像および復号画像の輝度値である。また PSNR による客観評価の他に、MOS (Mean

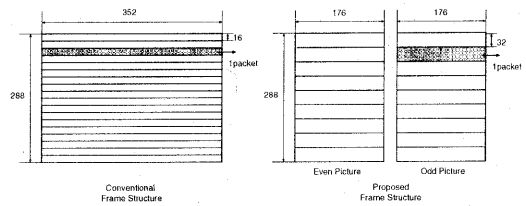


図 8 実験に用いたフレーム構造

Fig. 8 Frame Structuer using Experiment

表 1 実験条件

Table 1 Experimental Condition

使用した動画像	Flower Garden
画像サイズ	CIF 解像度 (352 × 288 画素)
フレームレート	10fps
符号化方式	H.261 に準拠
フレーム構成	1 フレーム 18GOB (図 8)
パケット	1GOB につき 1 パケット (18 packet/frame)
伝送エラー	パケット単位

Opinion Score) による主観評価も行う。

主な実験条件を表 1 に示す。使用する動画像は、画像サイズ CIF 解像度 (352 × 288 画素) の Flower Garden を用い、フレームレートは 10fps とする。動画像符号化方式は H.261 に準拠したものとし、各フレームは図 8 のように 18 個の GOB で構成されるものとする。パケットは 1 つの GOB データから構成され (18packet/frame)、伝送路ではエラーがパケット単位でランダムで起こるものとする。また、誤り伝搬を防ぐために 10 フレーム毎にフレーム内符号化を行い、エラーが生じた際のエラーコンシールメント [4] は前のフレームをコピーすることで行うものとする。

4.2 実験結果

4.2.1 客観評価

図 9 にビットレートと画質の関係を示す。エラーがないとき、従来方式は提案方式と同等の画質を得るために約 1.4 倍のビットレートが必要であることが分かる。これは、提案方式ではサブサンプリングを行って画像を分割するため、分割後の画像は隣の画素間との相関が少なくなり、DCT による圧縮効率が減少するためである。次に、パケットエラーを加えたときの特性をみると、従来方式はエラーが生じると急激に画質の劣化が起こるが、提案方式はエラーの影響を受けにくく、特にパケット誤り率 10% といった劣悪な通信環境下においても、良い特性を示していることが分かる。

次に、図 10 に従来方式のビットレートを 745Kbps、提案方式のビットレートを 744Kbps とほぼ同じにしたときのパケット誤り率と画質の関係を示す。従来方式は誤りが増加すると急

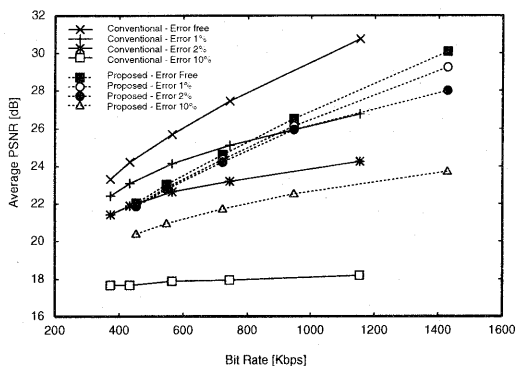


図9 ビットレートと Average PSNR の関係

Fig.9 Bitrate vs Average PSNR

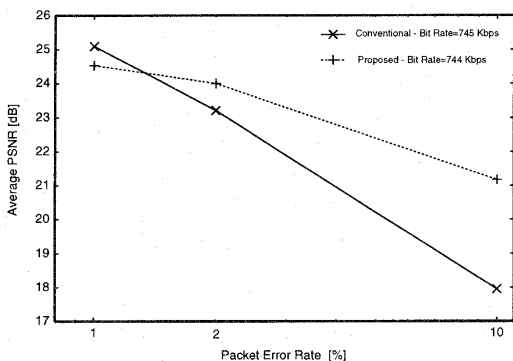


図10 パケット誤り率と Average PSNR の関係

Fig.10 Packet Error Rate vs Average PSNR

激に画質の劣化がおこるが、提案方式は誤りの増加による画質の劣化が緩やかであることが分かる。

図11は図10と同じビットレート、パケット誤り率10%における各フレームのPSNRを示したものである。ただし、図11では従来方式、提案方式それぞれに同じフレームにエラーが発生するようにしている。従来方式は、エラーがないときは提案方式に比べてよい画質を得られるが、少しでもエラーが生じると急激に画質が劣化することが分かる。一方、提案方式は、多少のエラーが生じてても急激には画質が劣化せず、エラーに対する耐性があることが分かる。しかし、図11の18から30フレームあたりにエラーが多く生じると画質の劣化につながる可能性がある。これは、両方のストリームに同時にエラーが発生したために、画像を補間により復元できなかったことによるものである。

4.2.2 主観評価

主観評価は、14人の評価者が原画像と復号画像を比較し、表2に示す5段階の評価基準を用いて最適な数値を選び出し、その評価値を平均して決定した。

表3にMOSの5段階評価による主観評価の結果を示す。従来方式はパケット誤り率が大きくなるとPSNRの低下とともに

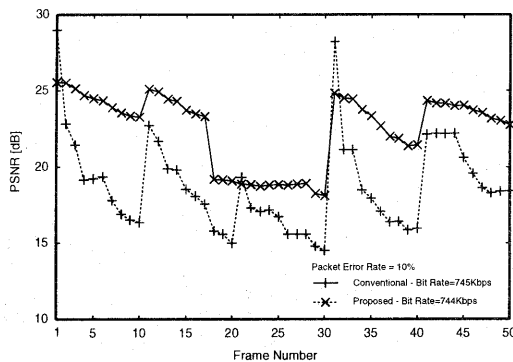


図11 各フレームのPSNR

Fig.11 PSNR of Each Frame

表2 主観評価基準

Table 2 Measure of Subjective Quality

MOS 値	評価基準
5	劣化がわからない
4	劣化がわ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

表3 主観評価結果 (従来方式 745kbps, 提案方式 744kbps)

Table 3 Subjective Quality (Conventional Method 745kbps, Proposed Method 744kbps)

パケット誤り率 [%]	従来方式		提案方式	
	PSNR[dB]	MOS 値	PSNR[dB]	MOS 値
0	27.5	4.6	24.8	4.1
1	25.1	2.6	24.6	4.2
2	23.2	2.4	24.4	4.5
10	18.0	1.8	21.2	4.0

にMOS値が低下するのに対し、提案方式はPSNRは低下するがMOS値はあまり変化しないことが分かる。パケット誤り率10%における21フレーム目の復号画像を図12,13に示す。それぞれのPSNRは従来方式が19.3dB、提案方式が18.9dBである。提案方式はエラーの様子ははっきりと分かるのに対して、提案方式はPSNRは18.9dBと低いにもかかわらず、あまりエラーが目立たないことが分かる。これは、提案方式がサブサンプリングを利用して画像を分割しているため、エラーが生じた場合に失われた画素を補間できるからである。

5. むすび

本稿では、Multiple Description Codingに注目し、その動画像への実現方法として、サブサンプリングを用いることにより画像を複数の相関を持ったストリームに分割し、動画像デー



図 12 復号画像の例：従来方式，パケット誤り率 10%，745Kbps，
PSNR=19.3dB

Fig. 12 Example of Decoded image :Conventional Method, Paket
Error Rate =10%, 745kbps, PSNR=19.3dB



図 13 復号画像の例：提案方式，パケット誤り率 10%，744Kbps，
PSNR=18.9dB

Fig. 13 Example of Decoded image :Proposed Method, Paket Er-
ror Rate =10%, 744kbps, PSNR=18.9dB

文 献

- [1] Vivek K Goyal, "Multiple Description Coding: Compression Meets the Network", IEEE Signal Processing Magazine, pp. 74-93, Sept. 2001
- [2] 藤原 洋, "最新 MPEG 教科書", マルチメディア通信研究会, 1994
- [3] Guy Cote, Berna Erol, Micheal Gallant, Faouzi Kossentini, "H.263+: Video Coding at Low Bit Rate", IEEE Trans. Circuits and System for Video Technology, vol 8, No.7, Nov. 1998
- [4] Yao Wang, Stephan Wenger, Jiangtao Wen, Aggelos K. Katsaggelos, "Error Resilient Video Coding Techniques", IEEE signal processing Magazine, pp. 61-82, JULY 2000

タのエラー耐性能力を向上させる方式を提案した。そして、伝送路でエラーが発生することを想定した実験を行い、PSNRを用いた客観評価、MOSの5段階評価基準を用いた主観評価において提案方式の特性を評価した。提案方式は従来方式に比べ符号化効率が悪く、エラーがない場合は同等の画質を得るために約1.4倍の符号量が必要である。しかし、提案方式は相関を持った2つの画像を別々に伝送するためエラーに対する画質の劣化が緩やかであることを示した。また、提案方式はエラーが多い環境では2つのストリームに同時にエラー生じて画質の劣化につながる可能性があるが、主観評価実験より提案方式ではその際にあらわれる画質の劣化が目立たず、劣悪な通信環境下においても有効であることを示した。