蓄積メディア用動画像符号化における 逆方向復号化について

正会員 上 倉 一 人[†], 正会員 東 野 豪[†]
正会員 渡 辺 裕[†],

あらまし 動き補償フレーム間予測を用いた蓄積メディア用動画像符号化方式において、逆方向復号化を可能とする符号化手法を提案した。また、提案手法による符号化特性を計算機シミュレーションにより求め、平行移動の多い画像に対して有効であることを確認した。

1. ま え が き

蓄積メディアの発達に伴って、蓄積メディア用動画像符号化方式が検討され始めている¹⁾²⁾。蓄積メディアでは、その利用形態を考慮すると、逆方向再生、高速再生、ランダムアクセスといった特殊機能を実現できることが望ましい。一方、動画像符号化では、動き補償フレーム間予測を導入することによって、その符号化効率を大幅に向上させることが可能である。しかし、従来の動き補償フレーム間予測を蓄積メディア用動画像符号化に導入した場合、画像を時間的に逆方向に復号化していくことは不可能であった。

本論文では、動き補償フレーム間予測を用いた符号化方式において、逆方向復号化を可能とする符号化手法を提案する。さらに、提案した手法による符号化特性を計算シミュレーションにより示す。

2. 逆方向復号化手法

第 n フレームにおける画素の復号値を $x_n(i, j)$, $x_{n-1}(i, j)$ に対応するフレーム間予測誤差信号の量子化後の値を $\varepsilon(i, j)$, 第 $(n-1)$ フレームにおける画素の復号値を $x_{n-1}(i, j)$ とする。

フレーム間予測に動き補償が含まれていない場合には、 $x_n(i, j)$ は次式により算出される。

$$x_n(i, j) = x_{n-1}(i, j) + \varepsilon(i, j) \quad (1)$$

したがって、逆方向復号化時には次式により $x_{n-1}(i, j)$ を算出することによって第 $(n-1)$ フレームのすべての画素を復号化できる²⁾。

$$x_{n-1}(i, j) = x_n(i, j) - \varepsilon(i, j) \quad (2)$$

次に、フレーム間予測に動き補償が含まれている場合を考える。 $x_n(i, j)$ に対応する動きベクトルを $v(u, v)$ とする。これは第 $(n-1)$ フレームから n フレームへの動きベクトルである。ただし u, v はそれぞれ i, j 方向の画素数を表す。逆方向復号化時には次式により第 $(n-1)$ フレームの画素の一部を復号化できる。

$$x_{n-1}(i-u, j-v) = x_n(i, j) - \varepsilon(i, j) \quad (3)$$

写真 1 動き補償フレーム間予測値として用いられなかった画素の空間分布

Spatial distribution of pels not used as motion compensated interframe prediction values

[†] NTT ヒューマンインタフェース研究所

“On Reversal Decoding Methods of Moving Picture Coding for Digital Storage Media” by Kazuto Kamikura, Suguru Higashino and Hiroshi Watanabe (NTT Human Interface Laboratories, Yokosuka)

しかし、(3)式において $x_{n-1}(i-u, j-v)$ と $x_n(i, j)$ とは必ずしも1対1に対応しているとは限らず、その場合には第 $(n-1)$ フレームに復号化できない画素が生じる。このような例を写真1に示す。写真1において黒く表されている領域は、次フレームの符号化時に動き補償フレーム間予測値として用いられなかった画素の領域である。これらの画素では、逆方向復号化時には(3)式による復号化が不可能である。そこで、第 n フレームの符号化時に、第 $(n-1)$ フレームのこれらの画素に対してのみ、逆方向復号化時に必要な情報を以下の手法により追加する。この追加情報により逆方向復号化が可能となる。ただし、符号化アルゴリズムにおいてループ内フィルタは使用していないことを前提としている。

手法(a)

第 n フレームの符号化時に動き補償フレーム間予測値として用いられなかった第 $(n-1)$ フレームの画素の復号値を $x_{n-1}(i_b, j_b)$ とする。本手法では、 $x_{n-1}(i_b, j_b)$ と第 n フレームにおける画素の復号値 $x_n(i_b, j_b)$ とのフレーム間差分を追加情報とする。すなわち、追加情報は次式の $\xi_1(i_b, j_b)$ で定義される。

$$\xi_1(i_b, j_b) = x_{n-1}(i_b, j_b) - x_n(i_b, j_b) \quad (4)$$

逆方向復号化を連続して行うためには、 $\xi_1(i_b, j_b)$ は量子化できない。したがって、追加情報は直接エントロピー符号化される。

手法(b)

動き補償フレーム間予測値として用いられなかった画素(復号値 $x_{n-1}(i_b, j_b)$) を、その画素に隣接する8画素(復号値 $x_{n-1}(i_b+p, j_b+q)$ $p, q = -1, 0, 1$ ただし、 $p=q=0$ を除く)のうち、動き補償フレーム間予測値として用いられた画素、および、すでに内挿ずみの画素を用いて内挿し、その内挿誤差を追加情報とする。内挿誤差 $\xi_2(i_b, j_b)$ は次式で定義される。

$$\xi_2(i_b, j_b) = x_{n-1}(i_b, j_b) - \hat{x}_{n-1}(i_b, j_b) \quad (5)$$

ただし、

$$\begin{aligned} \hat{x}_{n-1}(i_b, j_b) &= \left(\sum_{p,q} \frac{1}{\sqrt{p^2+q^2}} x_{n-1}(i_b+p, j_b+q) \right) \\ &\quad \cdot f(i_b+p, j_b+q) \\ &\quad / \left(\sum_{p,q} \frac{1}{\sqrt{p^2+q^2}} f(i_b+p, j_b+q) \right) \end{aligned} \quad (6)$$

ここで、 $f(i_b+p, j_b+q)$ は、 $x_{n-1}(i_b+p, j_b+q)$ が動き補償フレーム間予測値として用いられた画素、または、すでに内挿ずみの画素の復号値である場合に $f(i_b+p, j_b+q)=1$ 、それ以外の場合に $f(i_b+p, j_b+q)=0$ である。

したがって、内挿値 $\hat{x}_{n-1}(i_b, j_b)$ は、画素の配置が等方的であると仮定して、内挿に用いられる画素の復号値の、内挿しようとする画素からの距離で重み付けした平均値である。内挿は上のライン(走査線)から下に向かって順に、各ライン内では左端から右に向かって順に行う。したがって、(6)式の $f(i_b+p, j_b+q)$ がすべて0となる可能性があるのは、 $i_b=j_b=0$ の場合だけである。 $f(0+p, 0+q)$ がすべて0の場合には、 $\hat{x}_{n-1}(0, 0)$ を定数とする。

3. 符号化特性

逆方向復号化を可能にする本符号化手法による符号化特性を、計算機シミュレーションにより求めた。符号化方式は、動き補償フレーム間予測と離散コサイン

図1 符号化特性(画像1)
Coding performance. (image 1)

図2 符号化特性(画像2)
Coding performance. (image 2)

変換 (DCT) ³⁾⁴⁾を組合せた方式を基本としている。シミュレーションに用いた画像フォーマットは CIF フォーマット⁵⁾であり、画像の輝度信号 (360 画素×288 ライン) のみを用いた。画像信号は [0,255] の整数値 8 ビットのデータ精度である。したがって、(6)式の内挿値、(4)、(5)式の追加情報は整数値に量子化された値をとるものとする。また手法(b)において、(6)式の $f(i_b+p, j_b+q)$ が $i_b=j_b=0$ においてすべて 0 の場合、内挿値は $\hat{x}_{n-1}(0,0)=128$ としている。

シミュレーションは 5 フレームずつ 2 種類の画像に対して行った。画像 1 はカメラのパンニングにより、主として画面全体が平行移動している画像である。画像 2 は回転する物体を含んだ画像である。画像 1 および画像 2 に対する提案手法による符号化特性をそれぞれ図 1 および図 2 に示す。図 1 および図 2 には比較のために、追加情報がない場合の符号化特性 (すなわち動き補償フレーム間予測と DCT を組合せた手法) および動き補償を含まないフレーム間予測と DCT を組合せた手法 (以下、単にフレーム間予測 DCT と略す) による符号化特性も併せて示す。なお、符号量は Huffman 符号により設計された符号のほか、符号化テーブル識別のための符号、動きベクトル識別のための符号等、復号化に必要な符号をすべて含めた場合の符号量である。また S/N は次式で定義する。

$$S/N(\text{dB}) = 10 \log_{10} \frac{255^2}{D^2} \quad (7)$$

(D^2 : 符号化誤差電力)

図 1 および図 2 より、提案手法(b)は手法(a)に比較して効率が良いことがわかる。図 1 より、主として同一方向に動いている画像 1 では、逆方向復号化のための追加情報が比較的少なく、動き補償の効果によって符号量が大きく減少する。符号量が 0.5~1.3 b/pel の範囲で、手法(b)ではフレーム間予測 DCT に比べ S/N が 2~2.5 dB 向上している。一方、図 2 より、様々な方向への動きを含んでいる画像 2 では、逆方向復号化のための追加情報が比較的多い。そのため手法(b)は、フレーム間予測 DCT より特性が悪くなっている。このことは、手法(b)の有効性が画像の動きに依存することを意味する。しかし、再生専用型の蓄積メディアへの符号化の場合には処理のリアルタイム性を除外することができる。したがって、フレーム毎に両者を適応的に選択することにより、画像の種類に依

存することなく優れた符号化効率を得られると考えられる。

4. む す び

動き補償フレーム間予測を用いた蓄積メディア用動画像符号化において、逆方向復号化を可能とする符号化手法を提案し、提案手法による符号化特性を計算機シミュレーションにより示した。その結果、本手法は、主として同一方向に動いている画像に対して、動き補償を含まないフレーム間予測符号化に比べ優れた特性を示すことがわかった。

今後、提案手法と動き補償を含まないフレーム間予測符号化とを適応的に切り換える手法について検討する予定である。また、画像の動きと本手法の符号化効率との関係について定量的な検討を加えていく予定である。

(1989 年 4 月 17 日受付)

〔参 考 文 献〕

- 1) 茨木 久: “蓄積系動画像符号化方式の一検討”, 信学春季全大, D-111 (1989)
- 2) 花村, 富永: “ディジタル映像記録のための符号化方式に関する一検討”, 信学春季全大, D-99 (1989)
- 3) N. Ahmed, T. Natarajan, and K. Rao: “Discrete Cosine Transform”, IEEE Trans., C-23, 1, pp.90-93 (Jan., 1974)
- 4) W. H. Chen, C. H. Smith, and S. Fralick: “A Fast Computational Algorithm for the Discrete Cosine Transform”, IEEE Trans., COM-25, 9, pp.1004-1009 (Sep., 1977)
- 5) R. C. Nicol and N. Mukawa: “Motion Video Coding in CCITT SG XV—The Coded Picture Format”, Proc. of 1988 IEEE GLOBECOM, 2, 31.1 (Nov., 1988)

かみくら かずと
上倉 一人 昭和 61 年, 東京理科大学大学院修士課程修了。同年, NTT に入社。現在, 同社ヒューマンインタフェース研究所に勤務。画像の高効率符号化方式の研究に従事。正会員。

ひがしの たかし
東野 豪 昭和 63 年, 東京工業大学大学院修士課程修了。同年, NTT に入社。現在, 同社ヒューマンインタフェース研究所に勤務。画像の高効率符号化方式の研究に従事。正会員。

わたなべ ひろし
渡辺 裕 昭和 60 年, 北海道大学大学院博士課程修了。同年, NTT に入社。画像信号の高効率符号化方式の研究に従事。現在, NTT ヒューマンインタフェース研究所主任研究員, 工学博士。正会員。