

リーク予測による異種DCTフレーム間 符号化方式の整合化

黒田 英夫・藤村 誠
下田 裕・マワンサ デニス シレッシユ

Trying to match up interframe coding and decoding using different DCT algorithms by leaky prediction

by

Hideo KURODA, Makoto FUJIMURA
Yutaka SHIMODA and Mwansa Dennis Chileshe

In the international standard video coding specification H. 261 for videophone systems, DCT (Discrete Cosine Transform) is used for the interframe difference signals. If the used DCT algorithms are different each other, FDCT-IDCT mismatch may occur and cause the introduction of mismatch noise in the frame memories of sender and receiver. The leaky prediction is useful for removing the mismatch noise although it causes decreasing of insignificant blocks or deteriorating of picture quality on static area.

In this paper, the leaky interframe prediction coding system is proposed. The system however prevents insignificant blocks decreasing by raising up the threshold value to decide between insignificant block and significant block. Moreover, by dividing the decoded signals by leaky prediction coefficient, the decoded picture quality is improved for displaying.

1. まえがき

ISDNによるテレビ会議/テレビ電話システムの実現を可能とするキーテクノロジーとして、動画像用フレーム間符号化方式の研究が行われている¹⁾。その国際標準符号化方式であるH. 261では、NTSC, PAL, SECAMなど異なったテレビ信号方式間の相互接続が可能ないように共通ビデオフォーマットを決め、符号化効率の観点からフレーム間差分信号に対して離散コサイン変換を行う方式を採用している。このような国際標準符号化方式としては、特定の企業のみにも有利になることなく、各開発機関が自社のデジタルシグナルプロセッサ(DSP)を用いて独自に符号化装置を開発できるように、公平な仕様を決定する必要がある。

離散コサイン変換の出力は、DSPの演算精度や変換

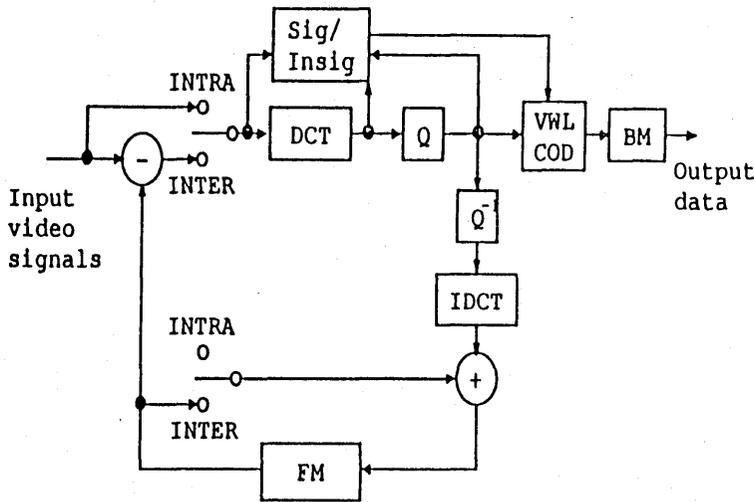
アルゴリズムによって、必ずしも一致はしない。従って、異種DCT方式を用いた符号化装置間で通信を行った場合、送受信部の間で不整合が生じることになる。これはFDCT-IDCTミスマッチ(FDCTはDCT変換、IDCTはDCT逆変換)と呼ばれている。このような異種DCT方式間でも相互接続が可能ないように、H. 261では、周期的にフレーム内符号化データを伝送することにより送受間の整合化を図る方式を採用している。

しかし、将来低ビットレートでより高画品質を要求するサービスが開発されるようになった場合には、この送受整合化用フレーム内符号化データの伝送が重荷になることが考えられる。

ここでは、リーク予測²⁾⁻⁴⁾を導入することにより送受の整合を可能とするフレーム間符号化方式について

平成4年9月30日受理

電気情報工学科 (Department of Electrical Engineering and Computer Science)



Sig/Insig: Decision circuit of significant block/insignificant block
 DCT: Discrete cosine transformer
 Q: Quantizer
 Q^{-1} : Inverse quantizer
 IDCT: Inverse DCT
 VWL COD: Variable word length coder
 BM: Buffer memory
 FM: Frame memory

Fig. 1 International standard coding system H. 261

検討する。まず、2においてH. 261国際標準フレーム間符号化方式の構成、3で予測係数と不整合雑音の減衰特性、4で符号化効率、5で無効ブロックにおけるリーク予測、6でリーク予測フレーム間符号化方式の構成、7でシミュレーション実験について述べる。

2. H. 261国際標準フレーム間符号化方式の構成

図1に国際標準フレーム間符号化方式H. 261の基本構成を示す。画像入力信号はフレーム間予測信号を引かれ、得られたフレーム間差分信号が8ライン×8画素からなるブロック毎にDCT処理を施される。DCT係数は量子化された後エントロピー符号化され、符号化データが受信側に伝送される。また、量子化出力が逆量子化され、IDCTされた後、予測信号と加算され局部復号信号が作成される。この局部復号信号が次のフレームの予測信号として使用される。また、勧告案で決められている標準のFDCT-IDCT方式で再生される各画素値に対して、各機関で開発する装置の不整合誤差が $1/256$ （以下1と記す）を越えないことと規定とすることにより各社独自のFDCT-IDCT方式を許容している。その代わり、132回符号化する度に1度はフレーム内符号化を挿入することが決められている。いま、64kbit/s符号化において、1秒間に10フレーム伝送した場合を想定すると、約13秒に1度フレーム内符

号化を行うことになる。

3. 予測係数と不整合雑音の減衰特性

本章では、FDCT-IDCTのミスマッチを防ぐ方法としてリーク予測について検討する。

リーク予測を用いるフレーム間符号化方式において、予測係数を P とすると、ある時点で発生したIDCTの不整合雑音 N は、1フレーム毎に P 倍された値に減衰し、 n フレーム後には P の n 乗倍まで減衰する。図2に不整合雑音のリーク状態を示す。 $P=0.9$ とした場合、例えばH. 261の国際標準勧告で規定されているフレーム内符号化の挿入周期と同じ132フレーム後の不整合雑音の量は $9 \times 10^{-7}N$ となる。国際標準勧告で N が1を越えないこととあることから考えると、 9×10^{-7} は画品質上十分小さい値であり、フレーム内符号化の挿入の必要はなくなる。画品質上からは N は $1/2$ に減衰すれば十分であり、このことを達成する時間としては、 $P=0.9$ の場合7フレームで達成可能である。

また、H. 261では、最大不整合雑音を1とすることの他、10,000個のブロック内の全画素点における平均二乗誤差を0.06以下とする等の規定もあり、FDCT-IDCTミスマッチの規定は更に厳しいものになっている。しかし、低廉な符号化装置を開発したい場合にはこの規定を満足しない符号化装置が接続され、かつ

それなりに使用できる通信状態が起きることも十分考えられる。

そこで、本稿では、最悪の場合をも想定して、前述したNが毎フレーム発生した場合についても検討した。この場合、受信側におけるNの累積雑音は式(1)のように表される。

$$\begin{aligned}
 N_s &= \lim_{n \rightarrow \infty} N \sum_{k=0}^n p^k \\
 &= \lim_{n \rightarrow \infty} N \cdot \frac{(1-p^{n+1})}{1-p} \quad [p < 1] \\
 &= \frac{N}{1-p} \tag{1}
 \end{aligned}$$

図3にnフレーム後の累積雑音を示す。図3の各カーブは $p < 1$ の時、式(1)の値 $N/(1-p)$ に収束する。即ち、リーク予測とすることにより、毎フレーム不整合雑音Nが発生し続けたとしても無限大時間後に $N/(1-p)$ の雑音が残るだけである。現象としては、早い時期に発生した雑音は図2で示した特性で減衰するため無限大時間後には殆ど影響しない。一方、観測直前に発生した雑音は影響が大きく現れる。

以上のことは、Nの発生条件としては最悪ケースの場合であり、現実的にはNの値の発生確率及びNの発

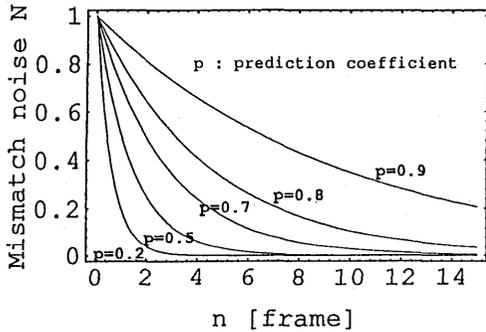


Fig. 2 Leaky's condition for mismatch noise

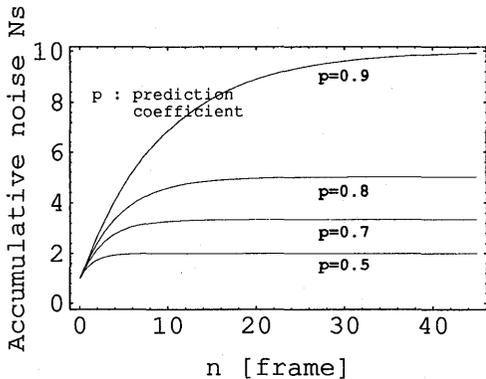


Fig. 3 Accumulative noise for n frames

生タイミングはランダムであると思われる。ここではNの平均値を用いて定量的に検討する。即ち、平均値 N_m の雑音が毎フレーム発生した場合を想定して検討する。

図4に無限大時間後の N_m の累積不整合雑音を示す。 $p=0.9$, $N_m=0.1$ の場合累積雑音は1であり、画品質上は問題ないものと思われる。

4. 符号化効率

リーク予測を採用することにより、FDCT-IDCT mismatchに対して、フレーム内符号化の挿入を必要としない見通しが得られた。そこで、次に符号化効率について検討する。

入力画素値を x 、1フレーム前の値を x_F 、予測係数を p 、平均信号電力を σ^2 、 x と x_F の相関係数を ρ 、予測誤差を e とおく。このとき、予測誤差電力は次式のように表される。

$$\begin{aligned}
 E = e^2 &= \overline{(x - px_F)^2} \\
 &= \overline{x^2} - 2p\overline{xx_F} + p^2\overline{x_F^2}
 \end{aligned}$$

ここで、定義により、

$$\overline{x^2} = \overline{x_F^2} = \sigma^2 \quad (\text{平均信号電力})$$

$$\rho = \overline{xx_F} / \sigma^2$$

とおくと、

$$E = \sigma^2 - 2p\rho\sigma^2 + p^2\sigma^2 \tag{2}$$

この予測誤差電力 E を予測係数 p で微分して式(3)を求める。

$$\partial E / \partial p = 2\sigma^2 (p - \rho) = 0 \tag{3}$$

式(3)を満たす p が予測誤差電力を最小とするものであり、 $p = \rho$ である。即ち、 $p = \rho$ の時最適予測符号化が行われることになる。

式(2)、(3)で表される予測係数 p と予測誤差電力 E の関係を図5に示す。図より相関係数の ρ の大きさによ

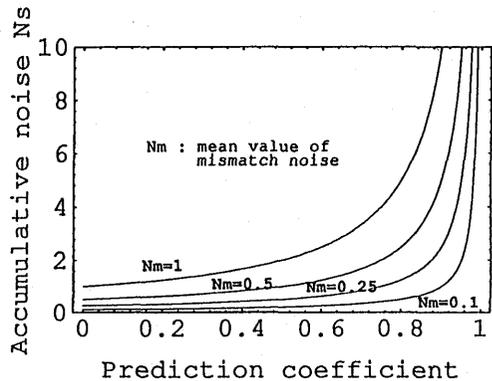


Fig. 4 Accumulative noise as time approaches infinity

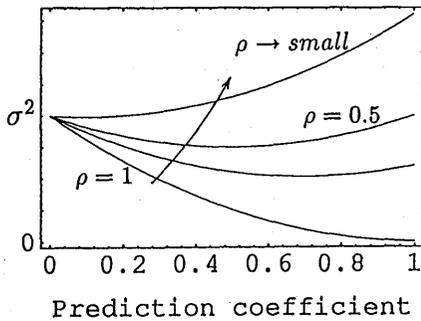
Prediction error power E 

Fig. 5 Relationship between prediction coefficient and prediction error power

り、即ち、画像の動きにより E と p の関係が異なることがわかる。いずれにしても、 $p = \rho$ の時最適予測となるため、相関係数の値を知って予測係数を制御するのが望ましい。予測係数 p が 1 以下の場合、これはとりもなおさず、リーク積分形の予測回路になっていることを表している。また、 $\rho = 1$ 即ち静止画像に対しては、当然のことながら、 $p = 1$ の完全積分形が望ましい。

5. 無効ブロックにおけるリーク予測

無効ブロックでは、DCT係数を伝送せず、フレーム間差分を零として処理するため、IDCTに起因する送受の不整合が生じることはない。従って、無効ブロックについて完全積分とすることは一見妥当なことのように思える。しかし、既に累積されている不整合雑音がフレームメモリに残ったまま、このブロックが無効ブロックになった場合、このブロックが無効ブロックになっている間リークされないため次に有効ブロックとなるまで保持されることになる。保持される不整合雑音は、あるブロックが無限遠の過去からずっと、有効ブロックとなる毎に毎フレーム最大の不整合雑音が発生し続けた場合に最大値となる。一般的には有効ブロックになる毎に最大不整合を生じることはなくほとんど問題になるような状況は発生しないものと思われるが、極端な場合には上述したような大きな累積雑音が画面上に現れる可能性が考えられる。

そこで、無効ブロックにおいても一定のリークをかけることが望ましい。しかし、一方で無効ブロックにおいて、リークをかけるということは、無効ブロックを有効ブロック化するということになり、フレーム間符号化方式の符号化効率で重要な役割を果たす無効ブロック化の増加を損なうことになる。従って、本稿で

は、有効・無効ブロックの判定閾値を高くすることにより無効ブロック数の減少を防ぐ方法を提案する。この場合、無効ブロックではフレーム間差分信号を強制的に零にするため、局部復号画像の振幅はリーク係数分だけ次第に減衰し、再生画像が暗くなったり、あるいはコントラストが弱くなることになる。更に、画像の静止領域が長時間続く時、無効ブロックがフレーム方向に連続して発生し、時々有効ブロックが発生することになるため、再生画像にフリッカが生じることがある。このような事態に対して、ここでは、以下のような方法を提案する。

①有効・無効ブロック判定閾値を高め設定し、無効ブロックを確保する。

②リーク予測により減衰された局部復号画像に対して、 $1/P$ 倍することにより、表示画像の画素値を補償する。

③予測画像用フレームメモリの他に表示用フレームメモリを設置し、無効ブロックが連続している間、このフレームメモリの出力を表示することにより、表示画像のフリッカの発生を防止する。

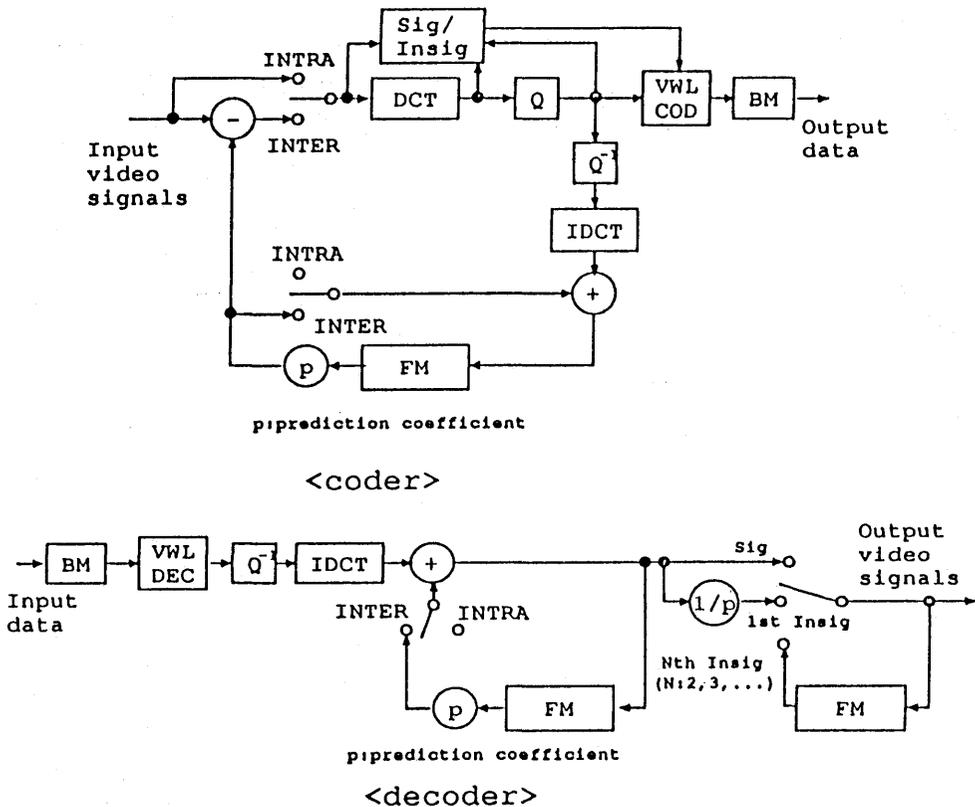
④入力信号が徐々に暗くなって、丁度リーク予測信号と同じ大きさになった場合は無効ブロックとなることがあり得る。この時②、③の処理を行うと、画像振幅が逆に拡大されることになる。従って、このような場合には有効ブロックとみなして処理する。有効ブロックとしても、もともと入力信号が変化しているブロックであるので特に問題はない。

6. リーク予測フレーム間符号化方式の構成

図6に本方式のソース・復号化部の基本構成を示す。画像入力信号はリーク係数分をリークした予測信号、即ち予測係数 P 倍分の信号を引かれ、得られたフレーム間差分信号が 8 ライン \times 8 画素からなるブロック毎にDCTされる。DCT係数は量子化された後エントロピー符号化され、符号化データが受信側に伝送される。送信側では更に量子化出力が逆量子化され、IDCTされた後、予測信号と加算され局部復号信号が作成される。この局部復号信号が次のフレームの予測信号として使用される。

受信側ではエントロピー復号化された量子化データがソース復号化部に入力される。量子化データは先ず逆量子化され、更にIDCTされ、フレーム間差分信号となる。このフレーム間差分信号がフレーム加算回路に入力され、送信側と同様にリーク積分される。

また、フレーム加算結果は表示画像作成回路に入力され、ここで表示画像用の信号が作られる。即ち、フ



Sig/Insig: Decision circuit of significant block/insignificant block
 DCT: Discrete cosine transformer
 Q: Quantizer
 Q⁻¹: Inverse quantizer
 IDCT: Inverse DCT
 VWL COD: Variable word length coder
 VWL DEC: Variable word length decoder
 BM: Buffer memory
 FM: Frame memory

Fig. 6 Leaky interframe prediction coding and decoding system

フレーム間差分が送られてくる有効ブロックについてはそのままスイッチを介して出力され、無効ブロックとなった最初のフレームにおいてはそのブロックの信号を $1/p$ 倍し、画素値を補償して出力する。2フレーム目以降の無効ブロックについては、表示画像用フレームメモリの内容を繰り返し読み出して、出力する。

7. シミュレーション実験

無効ブロックにおける復号画像信号及び表示画像信号の状況を確認するため、シミュレーション実験を行った。1画面全体が無効ブロックとなる静止画像を例にした。図7に復号画像信号及び表示画像信号のSNRを示す。リーク予測係数が小さいほど復号画像信号のSNRが低くなっているのに対し、表示画像信号のSNRはほぼ一定になることが分かる。これは復号画像

信号の振幅がリーク係数だけ次第に減衰するのに対し、表示画像信号の振幅が常に原画像信号の振幅と同じ値に補償されているからである。

これら処理画像の例を図8に示す。(A)は原画像、(B)はフレームメモリに蓄積される復号画像、(C)は振幅を補償された表示画像である。

8. むすび

国際標準符号化方式H. 261において、フレーム内符号化を用いることなく、異種DCT方式間で生じるFDCT-IDCTミスマッチを除去する方式として、リーク予測を用いたフレーム間符号化方式を提案した。

先ず、リーク予測を導入することにより、H. 261で規定されている最大の不整合雑音が毎フレーム発生し続けた場合でも累積不整合雑音は一定値以上にはなら

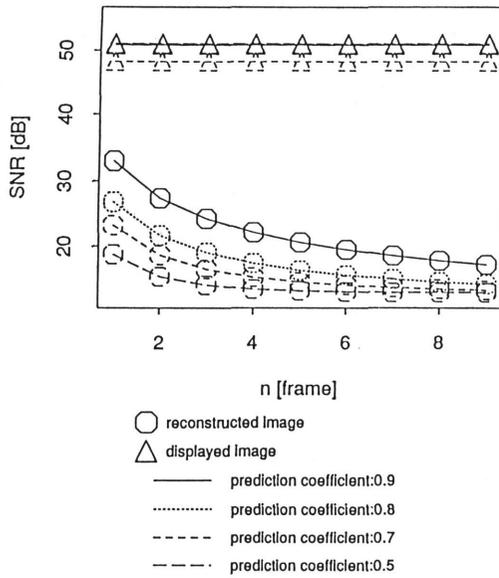


Fig. 7 Comparison of reconstructed and displayed image

ず、不整合雑音の抑圧に有効であることを明らかにした。しかし、静止領域にリークをかけた場合には無効ブロックが減少し、符号化効率が大幅に損なわれる。このため、特に無効ブロックを確保し、かつリークも掛ける方法として、以下の方法を提案した。

- (1) 無効ブロックにおけるリーク予測係数を高くする、あるいは有効・無効ブロック判定閾値を高めにして無効ブロックを確保する。
- (2) 受信側で $1/P$ をかけることにより、再生表示画像が暗くなることを防ぐ。
- (3) 受信側に表示画像用フレームメモリを付加し、無効ブロックが継続する間は、2フレーム目以降の無効ブロックに対してこのフレームメモリの出力を表示することにより、再生画像にフリッカが生じるのを防ぐ。

提案した方式についてシミュレーション実験を行い、無効ブロックを確保し、かつ安定した再生表示画像を得られることを示した。この結果、フレーム内符号化を用いることなく、リーク予測により異種DCT方式間の不整合雑音を削減できることが明かとなった。

今後、FDCT-IDCT ミスマッチのみでなく、伝送路エラーや信号源の切り替えに基づく送受の不整合に対するリーク予測係数について検討を行う予定である。

参考文献

- 1) 大久保：「テレビ会議／電話用符号化の標準化動



(A) Original image



$\rho=0.5$ $n=2$
(B) Reconstructed image



$\rho=0.5$ $n=2$
(C) Displayed image
Fig. 8 Examples of image

向”，PCSJ89画像符号化講演会，pp43-48，1989，11月

- 2) 甲藤，安田：「階層的符号化を用いた映像パケット通信におけるセル廃棄対策」，電子情報通信学会論文誌，Vol. J72-B-I，No.11，pp1094-1102，Nov. 1989

- 3) 滝沢, 吹抜: “複合差分予測に適した伝送誤り伝搬の減衰方法”, PCSJ86, 6. 2, pp73-74, 1986
- 4) T. Fukinuki: “Optimization of D-PCM for TV Signals with Consideration of Visual Property”, IEEE Trans. on Com., Vol. COM-22, No. 6, pp. 821-826, (June 1974)