

LJ_003

符号化アーチファクトを抑制する H.264 符号化モード選択方式

An Enhanced Mode Decision to Suppress Coding Artifacts in H.264 Encoding

蝶野 慶一 宮本 義弘
Kei-ichi Chono Yoshihiro Miyamoto

NEC メディア情報研究所
Media and Information Research Laboratories, NEC Corporation

1. はじめに

H.264[1]では、イントラ予測モードおよび動き補償予測モードの各々について複数の符号化モードが規定されている。このため、エンコードにおいては、画像の特性に応じて、複数の符号化モードの中から符号化効率の高い符号化モードを選択する必要がある。比較的低演算量で、なお且つ、高性能な符号化モードの選択方式の一つとして、画像の予測誤差だけでなく、モードのヘッダ符号量を考慮した符号化モード選択方式(以後、これを従来方式と呼ぶ)が、H.264 リファレンスソフト[2]に実装されている。

ところで、H.264 においては、符号化済みの周辺ブロックの符号化データから予測したパラメータ(イントラ予測モードあるいは動きベクトル)を、符号化対象ブロックの符号化パラメータとしてそのまま利用できる符号化モードがある。このような符号化モードは、他の符号化モードよりも伝送するヘッダ符号量が相対的に少ないため、低ビットレートで選択されやすくなる。

しかしながら、低ビットレートの符号化で、ヘッダ符号量の少ない符号化モードが連続して選択されると、入力画像に存在しないテクスチャパターンが発生したり(イントラアーチファクト)、さらには、入力画像の動きとは異なる動きベクトル場が形成されて(モーションアーチファクト)、主観的な画質劣化が生じる問題がある。本稿では、前記の主観画質劣化の発生原理を説明し、その抑制を可能とする符号化モード選択方式を提案する。

2. 従来方式と符号化アーチファクトの発生原理

従来方式においては、式(1)のコスト関数を利用して符号化モードを選択する。

$$Cost(mode, QP) = SATD(mode) + \lambda(QP) \times R(mode) \quad (1)$$

ここで、QP は量子化パラメータ、SATD(x) はモード x の予測誤差の Hadamard 変換係数の絶対値総和、R(x) はモード x のヘッダ符号量、 $\lambda(Q) = 2^{\lceil (Q-12)/6 \rceil}$ である。式(1)から、低ビットレートになるほどコスト関数におけるヘッダ符号量の比重が高くなるのが分かる。このため、SATD が小さい平坦領域においては、低ビットレートになるほどヘッダ符号量の少ないモードが優先的に選択されるようになる。ヘッダ符号量の少ないモードが連続して選択されると、H.264 規格の構造によって、I ピクチャにおいてはイントラアーチファクト、P ピクチャにおいてはモーションアーチファクトが発生する。各アーチファクトの発生原理を以下で説明する。

2.1 イントラアーチファクトの発生原理

H.264 の Intra_4x4 あるいは Intra_8x8 の符号化モードにおいては、符号化対象ブロックのイントラ予測モード mode とその左もしくは上に隣接する符号化済みブロックが

ら予測したイントラ予測モード pmode が一致する場合(mode==pmode)、当該ブロックのヘッダ符号量を 3bit 削減できる[1]。従来方式は、この削減効果を考慮し、イントラ予測モードのコスト関数の R(mode) を式(2)で計算する。このため、低ビットレートでは、pmode が優先的に選択されるようになる。

$$R(mode) = \begin{cases} 0 & \dots \text{ if } (mode == pmode) \\ 4 & \dots \text{ Otherwise} \end{cases} \quad (2)$$

ここで、H.264 規格においては、前記符号化済みブロックの中でモード番号が小さい方のイントラ予測モードが pmode に設定されること、モード番号が表 1 のように定義されていることに着目する。すると、この規格の構造に起因して、平坦な画像の一部でモード番号の小さい V および H が選択すると、選択した V あるいは H のイントラ予測モードが後続する平坦領域に伝播することが分かる(図 1)。V や H のイントラ予測モードが平坦領域に連続して伝播すると、主観的に目立つテクスチャパターン(文献[3]の線状ノイズとは性質の異なる特有のノイズ)が発生する。これが、イントラアーチファクトの発生メカニズムである。

表 1: モード番号とイントラ予測方向

mode	Spatial direction
0	Vertical
1	Horizontal
2	DC
3	Diagonal down left
4	Diagonal down right
5	Vertical right
6	Horizontal down
7	Vertical left
8	Horizontal up

2.2 モーションアーチファクトの発生原理

H.264 においても、P ピクチャのスキップ MB モード(以後、単純にスキップモードと呼ぶ)では、動きベクトル、および DCT 係数などの情報を伝送しない。このことを考慮して、従来方式は、スキップモードのコスト skip_cost を式(3)で計算する。

$$SkipCost(sMV, QP) = SATD(sMV) - 16 \times \lambda(QP) \quad (3)$$

ここで、sMV はスキップモードの動きベクトルである。該ブロックの左もしくは上に隣接するブロックの動きベクトルが零である場合の sMV は零、そうでない場合の sMV は隣接 3 ブロックの動きベクトルの中央値(pMV)となる。なお、pMV は 16x16 ブロックサイズの動き補償予測モードにおいて動きベクトル MV16 の予測値となる。

式(3)から明らかなように、従来方式を利用すると、低ビットレートでスキップモードが優先的に選択されるようになる。しかしながら、パンションなどで sMV=0 のスキップモードが選択されると、上記 sMV の計算方法とメディア

ン予測による動きベクトルの伝播特性から(図2)、カメラの動きとは異なる零の動きベクトル場が平坦領域に形成される(図3)。この結果、パンニングの動きと符号化された画像の動きが大きく異なり、主観的に不自然な動きやテクスチャの引きずりが見える。これが、モーションアーチファクトの発生メカニズムである。

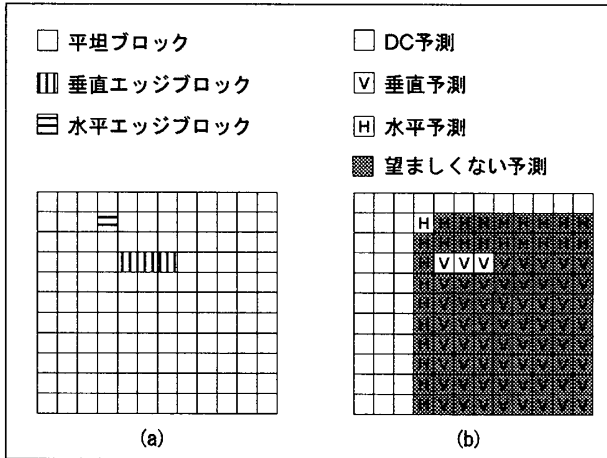


図1: (a):原画像の特性、(b):原画像に対して従来方式が選択するイントラ予測モード

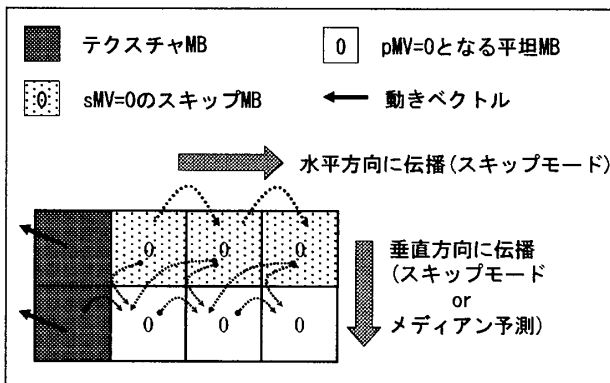


図2: 動きベクトルの伝播特性

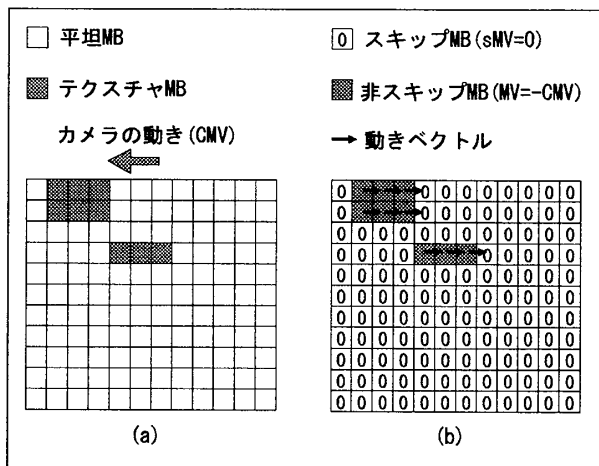


図3: (a):原画像の特性、(b):原画像に対して従来方式が形成する動きベクトル場

3. 提案方式

提案方式では、上述したイントラ予測モードあるいは動きベクトルの画面内伝播特性を考慮して、モード選択を行う。提案方式のイントラ予測モード選択、スキップモード選択をそれぞれ以下で説明する。

3.1 イントラ予測モードの選択

表1の構造に着目し、pmodeがVあるいはHの場合、DCモードのコスト関数のヘッダ符号量を0として計算する。すなわち、イントラ予測モードのコスト関数の $R(mode)$ を式(4)で計算する。

$$R(mode) = \begin{cases} 0 & \dots \text{if}(mode = pmode) \\ 0 & \dots \text{else if}(pmode = V \ \& \ mode = DC) \\ 0 & \dots \text{else if}(pmode = H \ \& \ mode = DC) \\ 4 & \dots \text{Otherwise} \end{cases} \quad (4)$$

式(4)のヘッダ符号量を用いて当該ブロックのイントラ予測モードを選択した後、後段のMBモード判定のために、式(5)でコスト関数の値を補正する。

$$Cost(mode, QP) = \begin{cases} Cost + 4 \times \lambda(QP) & \dots \text{if}(pmode = V \ \& \ mode = DC) \\ Cost + 4 \times \lambda(QP) & \dots \text{else if}(pmode = H \ \& \ mode = DC) \\ Cost & \dots \text{Otherwise} \end{cases} \quad (5)$$

このようにすることで、VあるいはHのイントラ予測モードが選択されたブロックに後続する平坦領域に対するイントラ予測モードの伝播を防ぐことができる(図4(a))。また、提案方式は、すべてのブロックで3ビットのヘッダ符号量を伝送するのではなく、VあるいはHに隣接する平坦ブロックだけで伝送するので、極端にヘッダ符号量が増えることはない。

3.2 スキップモードの選択

2.2で述べたスキップモードとメディアン予測による動きベクトルの伝播特性を考慮し、提案方式は、スキップモードのコスト $skip_cost$ を式(6)で計算する。

$$SkipCost(sMV, QP) = \quad (6)$$

$$SATD(sMV) + \lambda(QP) \times RDMV(sMV - MV16)$$

ここで、 $RDMV(x)$ は差分ベクトル x の符号量を計算する関数である。式(6)の第2項によって、動き推定が検出した $MV16$ と sMV のずれが大きい場合、スキップモードが選ばれ難くなる。なお、式(6)の代わりに式(7)を利用しても、同様の効果を期待できる。

$$SkipCost(sMV, QP) = \quad (7)$$

$$SATD(sMV) + \lambda(QP) \times RDMV(sMV - pMV)$$

このようにすることで、パンニングなどで $sMV=0$ のスキップモードが発生し、カメラの動きとは異なる動きベクトル場が形成されることを防ぐことができる(図4(b))。また、提案方式は、パンニングなどにおいて、すべてのMBではなく、実際の動きと異なる sMV の平坦MBだけでスキップモードを選ばれ難くする。このため、提案方式によって、極端にヘッダ符号量が増えることはない。

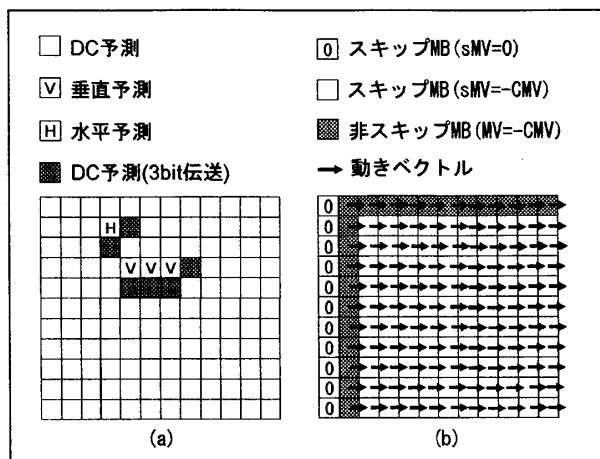


図 4: (a): 図 1(a) の原画像に対して提案方式が選択するイントラ予測モード、(b): 図 3(a) の原画像に対して提案方式が形成する動きベクトル場

4. 実験

HDTV(1920x1080i)の ITE 標準テスト画像(ITE12, ITE22, ITE23)に対して符号化実験を行った。符号化条件は、RD 最適化 OFF、MBAFF、N=15、M=3、CABAC、動き探索範囲±128、デフォルト量子化マトリクス、固定量子化(QP=28/32/36/40、Bピクチャでは+3)である。エンコードソフトには、RD 最適化 OFF が正常に動作するようにバグフィックスをした JM[2] を利用した。JM に 3.1 のイントラ予測モード選択を実装した方式を Method1、Method1 に 3.2 のスキップモード選択を実装した方式を Method2 として、各方式の符号化結果をグラフ 1-3 に示す。グラフ 1-3 から明らかなように、提案方式(Method1 および Method2)の RD 特性が、JM と殆ど変わらないことを確認できる。これは、3 節で述べたように、提案方式においてヘッダ符号量を多く伝送するのは、特定の条件を満たすブロックあるいは MB だけになるからである。

続いて、提案方式による主観的な画質改善の効果を確認するために、原画像と対応する符号化画像の一部を図 5-10 に示す。JM の I ピクチャ(図 6)では、イントラアーチファクト(垂直方向にテクスチャパターンが伝播)が発生しているのに対して、提案方式(図 7)では発生していない。また、JM の P ピクチャ(図 9)では、モーションアーチファクト(斜めのラインの周りに過去フレームのラインが残存している)が発生しているのに対して、提案方式(図 10)では発生していない。

以上の結果から、提案方式によって、符号化アーチファクトを効果的に抑制できていることが確認できる。

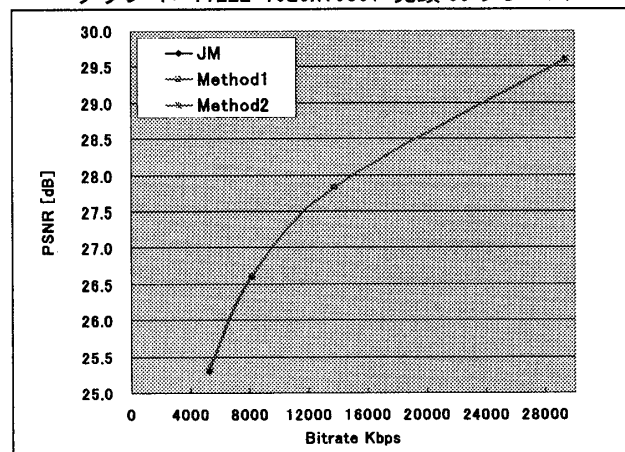
5. まとめ

本稿では、H. 264 による低ビットレートの符号化において、ヘッダ符号量の少ない符号化モードが連続して選択されることによって生じる主観画質劣化とその発生原理を述べ、その抑制を可能とする符号化モード選択方式を提案した。実験結果により、提案方式によって、レート歪み特性を損なうことなく符号化アーチファクトを効果的に抑制し、主観画質を改善できることを確認した。

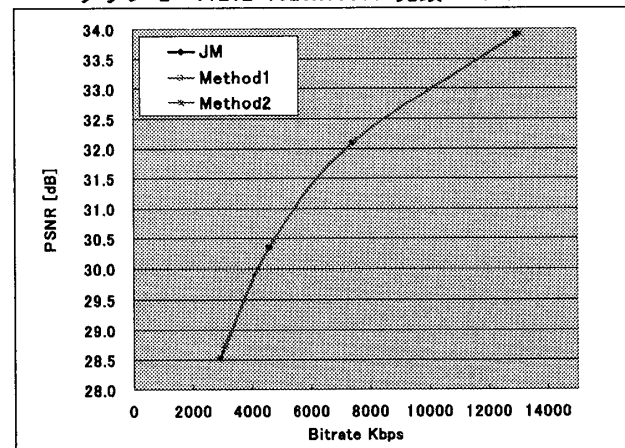
参考文献

- [1] ISO/IEC FDIS 14496-10 "Advanced Video Codec" (AVC) & ITU-T Rec. H. 264, 2003.
- [2] JVT Reference Software: <http://iphome.hhi.de/suehring/tml/>
- [3] 谷田他, "H. 264/MPEG-4 AVC における線状ノイズ抑制アルゴリズムの検証", 電子情報通信学会総合大会、D-11-38、2004.

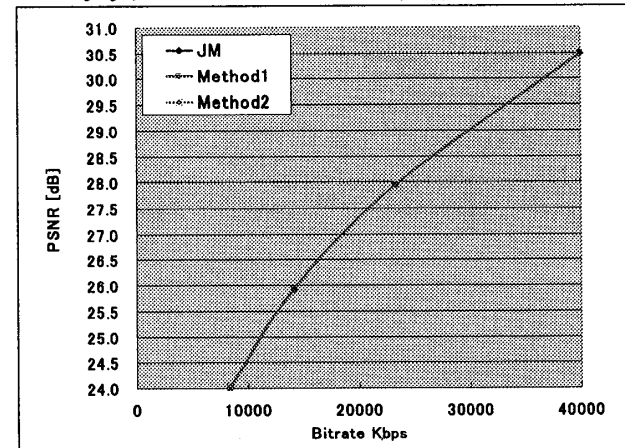
グラフ 1: ITE22 1920x1080i 先頭 60 フレーム



グラフ 2: ITE12 1920x1080i 先頭 60 フレーム



グラフ 3: ITE23 1920x1080i 先頭 60 フレーム



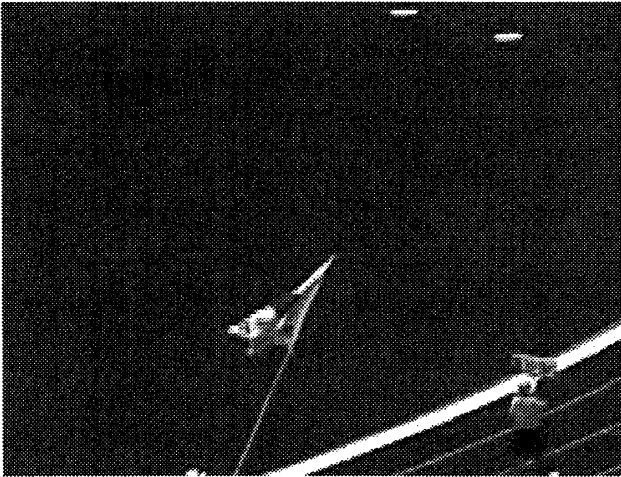


図 5: 原画像 (ITE22) の部分領域 (先頭フレーム)



図 8: 原画像 (ITE22) の部分領域 (第 9 フレーム)

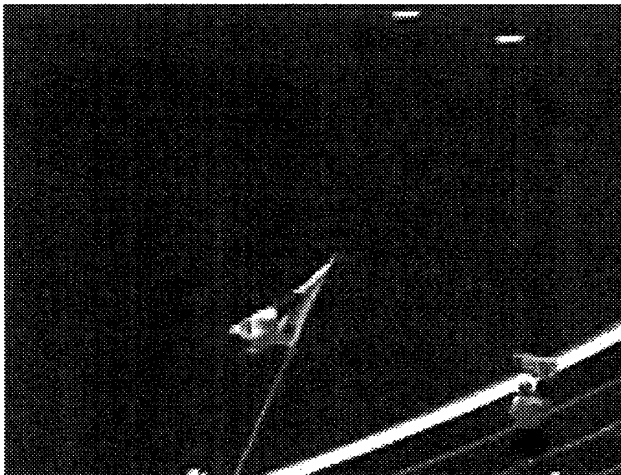


図 6: JM I ピクチャの部分画像
(QP=36, 8100.53 Kbps, 26.60[dB])



図 9: JM P ピクチャの部分画像
(QP=36, 8100.53kps, 26.60[dB])

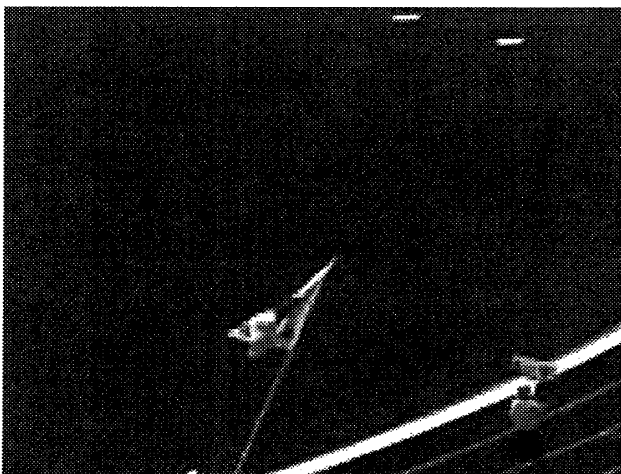


図 7: Method1 I ピクチャの部分画像
(QP=36, 8103.53 Kbps, 26.60[dB])



図 10: Method2 P ピクチャの部分画像
(QP=36, 8113.70 Kbps, 26.62 [dB])