

H.264/MPEG-4 AVCにおける符号化モード適応選択手法に関する一検討

吉野 知伸[†] 内藤 整[†] 小池 淳[†]

[†] 株式会社 KDDI 研究所
〒 356-8502 埼玉県ふじみ野市大原 2-1-15
E-mail: †{to-yoshino,sei,koike}@kddilabs.jp

あらまし ブロードバンドサービスの普及に伴い HDTV を IP ネットワーク上で配信するサービスの開発が期待されるが、現在普及している既存専用線を活用することを考慮すると、10Mbps 前後の伝送レートでの HDTV 伝送が求められる。MPEG-2 を用いて HDTV を符号化した場合、これを伝送するためには 20Mbps 程度の伝送レートが必要であるといわれているため、IP ネットワーク上で HDTV を伝送するためには、より高い符号化効率を実現できる符号化方式が必要となる。一般に、動画像において高い符号化効率を実現するためには動き補償と DCT を用いた符号化が有効であるとされており、近年では、H.264/MPEG-4 AVC が、この種の符号化の中で最も高い圧縮率を実現している。H.264/MPEG-4 AVC High Profile では高精細映像に特化して高い符号化効率を実現するための符号化ファンクションが規定されており、特に符号化モードの適応的選択は符号化効率の改善に大きく寄与するといわれている。しかし、一般に、符号化モードの適応選択においては、符号化モードを決定するに至るまで多大な処理が必要とされているため、同ツールの適用にあたっては画質を保ったまま符号化モードの適応選択における処理量を削減する必要がある。本研究では、符号化モード選択について映像の特徴との依存関係を検証し、この性質を利用して符号化モード選択に要する処理量を削減する手法を提案する。また、本手法の評価実験を行い、本手法により画質を保ちつつ符号化モード選択に要する処理量が削減されることを確認する。

キーワード H.264/MPEG-4 AVC, HDTV, 符号化モード適応選択, 低ビットレート符号化

A study on adaptive coding mode decision scheme for H.264/MPEG-4 AVC

Tomonobu YOSHINO[†], Sei NAITO[†], and Atsushi KOIKE[†]

[†] KDDI R&D Laboratories Inc.
2-1-15 Ohara, Fujimino-shi, Saitama, 356-8502 Japan
E-mail: †{to-yoshino,sei,koike}@kddilabs.jp

Abstract It is hoped that the service of HDTV broadcast over IP network will be developed along with the popularization of broadband service, then transport rate of HDTV over IP network is needed about 10Mbps considering expansion of HDTV broadcasting service with use of prevailing leased line. Transport bit-rate about 20Mbps is needed to transport HDTV coded by MPEG-2, therefore more effective video coding standard is needed for HDTV broadcast over IP network. Generally it is said that coding based on MC and DCT is effective to achieve high effective encoding for video, recently H.264/MPEG-4 AVC achieves the most effective video coding among this kind of them. It is recommended that coding functions to achieve high effective coding for HDTV by standard issue of H.264/MPEG-4 AVC High Profile, adaptive coding mode decision make a contribution to improve the coding effectiveness largely. But generally computing cost to decide coding mode is very large, therefore applying it to video encoding, computational cost to decide coding mode should be decreased on the basis that its image quality was kept. In this paper, we evaluated that selected coding mode were dependent on its image features, and we proposed a coding mode decision scheme, by which computational cost could be decreased. Finally we experimented about this scheme, and confirmed that this scheme decreased computational cost to decide coding mode on the basis that its image quality was kept.

Key words H.264/MPEG-4 AVC, HDTV, coding mode decision, low bit-rate coding

1. はじめに

現在のデジタル放送で用いられている HDTV は、MPEG-2 を用いて符号化されており、これを伝送するためには 20Mbps 程度の伝送レートが必要であるといわれている。一方で、ブロードバンドサービスの普及に伴い、HDTV を IP ネットワーク上で配信するサービスの開発が期待されるが、現在普及している ADSL などの既存の専用線設備を活用することを考慮すると、10Mbps 前後の低いビットレートでの HDTV 伝送が求められる。このため、HDTV に対する符号化方式としては、MPEG-2 よりもさらに高い符号化効率を実現できる符号化方式が必要となる。

一般に、動画像において高い符号化効率を実現するためには動き補償と DCT を用いた符号化が有効であるとされており、近年では、H.264/MPEG-4 AVC [1] が、この種の符号化の中で最も高い圧縮率を実現している。

H.264 では、高効率な符号化や符号化データのエラー耐性を高めるなどの目的に応じてプロファイルが定められており、H.264 High Profile では高精細映像に特化して高い符号化効率を実現するための符号化機能が規定されている。これらについての規定はデコードに関する規定であり、エンコーダにおけるチューニングを含めた実装方法についてはエンコーダの開発者に委ねられており、符号化制御の最適化アプローチについて検討が行われている [2]~[4]。

H.264 High Profile における高い符号化効率を目的とした符号化機能の中でも、符号化モードの適応的選択は符号化効率の改善に大きく寄与するといわれており、低いビットレートにおける HDTV の符号化においては同ツールの適用が不可欠である。

しかし、一般に、符号化モードの適応的選択においては、符号化モードを決定するに至るまで多大な処理が必要とされているため、同ツールの適用にあたっては画質を保ったまま符号化モードの適応的選択における処理量を削減する必要がある。

本研究では、H.264 High Profile において HDTV に対して低ビットレートで符号化を行った場合について、符号化モード選択について映像の特徴との依存関係を検証し、その性質を用いて符号化モード選択処理量の削減手法を提案する。最後に、提案手法の評価実験を行い、その効果を検証する。

2. リファレンスエンコーダ

2.1 符号化モード判定手順

H.264 のリファレンスエンコーダである JM では、符号化モードを適応的に選択する単位として、MB 単位とピクチャ単位の 2 種類を選択することができる。

MB 単位での符号化モード適応的選択においては、該当するマクロブロックペア (以下、MB-pair) について、フレーム符号化モードとフィールド符号化モードそれぞれで符号化を行う。符号化の際には、符号化に用いられる全ての MB サイズについて符号化を試行する。なお、インター符号化を行うフレームの MB-pair については、各 MB サイズでの符号化を試行する際

に動きベクトルのサーチを行い、各 MB サイズにおける参照フレームを決定する。各 MB サイズについての符号化を試行した結果として、各 MB サイズにおけるコスト値を算出する。それぞれの符号化モードについて、最もコスト値が小さい MB サイズのコスト値をその符号化モードにおけるコスト値とする。最後に、各符号化モードのコスト値を比較し、値が小さい方を該当 MB-pair の符号化モードとする。MB 単位での符号化モード適応的選択の処理フローを図 1 に示す。

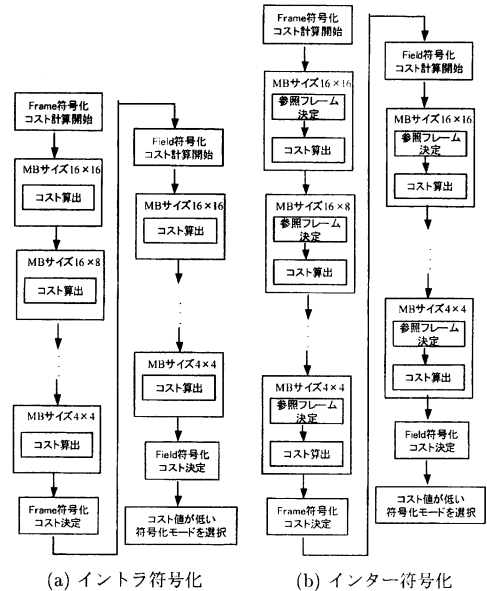


図 1 MB 単位での適応的選択処理フロー

ピクチャ単位での符号化モード適応的選択についても、MB 単位と同様に、該当するピクチャについて、フレーム符号化モードとフィールド符号化モードでそれぞれ符号化を行う。ピクチャ単位における符号化モードのコスト値は、ピクチャを構成する MB のコスト値の和である。符号化モードの決定は、各符号化モードのコスト値を比較してコスト値が小さい方を該当ピクチャの符号化モードとする。ピクチャ単位での符号化モード適応的選択の処理フローを図 2 に示す。

なお、コスト値の算出においては、絶対値誤差和を用いる手法 (以下、SAD) と、誤差にアダマール変換を施した結果の絶対値和を用いる手法 (以下、SATD) を選択することができる。インター符号化を行う場合には、原画像に対して DCT、量子化、逆量子化、逆 DCT を施した結果と原画像の差分を求め、SAD および SATD を算出 (図 3(a)) する。また、インター符号化を行う場合については、参照画像に対して直交変換、量子化、逆量子化、逆直交変換を施した結果と原画像の差分を求め、SAD および SATD を算出 (図 3(b)) する。

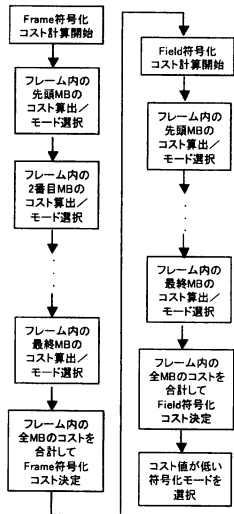


図2 ピクチャ単位での適応選択処理フロー

コスト値は、発生ビット量に Lagrange 乗数を乗じた値に SAD および SATD を加えた値の和が用いられる。発生ビット量を *GenBit*, Lagrange 乗数を λ としたとき, SAD を用いたコスト算出式を式 (1) に示す。なお, SATD を用いたコスト算出式は、式 (1) の SAD を SATD に置き換えた式である。

$$SAD + \lambda \times GenBit \quad (1)$$

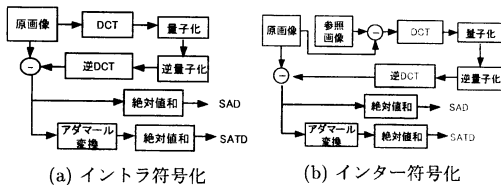


図3 コスト算出フロー

2.2 リファレンスエンコーダの符号化結果

予備実験として、H.264 のリファレンスエンコーダである JM10.1(以下、単に JM と記載)を用いて、ピクチャ単位および MB 単位それぞれで符号化モードを適応的に選択して符号化を行った。また、コスト値の算出については、SAD および SATD を用いた。符号化に用いた素材は ITE HDTV 標準動画 (1920 画素 \times 1080 ライン、8bit) の 10 シーケンス (Sequence No.07, 12, 16, 19, 20, 23, 25, 30, 34, 43) 各 450 フレームである。符号化を行う際の条件を表 1 に示す。符号化結果の SNR 値について、表 2 に示す。なお、表中では、ピクチャ単位で符号化モード選択を行った結果を PIC、MB 単位で符号化モード選択を行った結果を MB と記載する。

表 2 より、符号化モードを適応的に選択する単位をピクチャ単位とした結果に比べて、MB 単位で符号化モードを適応的に選択した結果の方が、より高い符号化効率が得られている。特

表 1 符号化条件

設定項目	設定値
フレームサイズ	1920 画素 \times 1080 ライン
フレーム周波数	30fps
プロファイル	Hight Profile
ビットレート	8Mbps
DCT 変換サイズ	4 \times 4 および 8 \times 8
符号化方式	CABAC
アダマール変換	なし
動き予測	UMHexagon Search
R-D 最適化	あり
GOP 構成	IBBPBBPBBPBBPBB
参照ピクチャ	I および P ピクチャ
参照ピクチャ数	2
動きベクトル探索範囲	$\pm 32 \times \pm 32$

表 2 符号化結果

Sequence No.	SAD		SATD	
	PIC (dB)	MB (dB)	PIC (dB)	MB (dB)
07	34.73	34.74	34.75	34.76
12	32.94	33.03	32.95	33.03
16	28.92	28.66	28.92	28.66
19	28.10	28.12	28.12	28.14
20	28.67	28.64	28.68	28.64
23	24.73	25.22	24.79	25.26
25	41.27	41.24	41.30	41.28
30	36.76	36.92	36.80	36.92
34	38.74	38.76	38.77	38.78
43	30.89	31.01	30.90	31.03

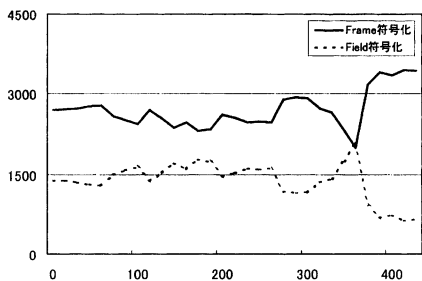
に Sequence No.30 や 43 では MB 単位で符号化モードを適応的に選択することにより、ピクチャ単位での符号化モード選択に対して 0.2dB 程度の SNR 値向上が見られる。一方、Sequence No.07 や 19 では、符号化モード選択の単位が MB 単位とピクチャ単位で同等の結果となっている。MB 単位での符号化モード適応選択において、映像の種類による符号化効率改善効果の違いについては、次節 2.3 で述べる。

一方、コスト算出方法として SAD および SATD を用いる手法については、シーケンスおよび符号化モード選択の単位によらず、符号化効率向上の効果が見られない。したがって、表 1 の符号化条件に示すような低ビットレートにおける HDTV の符号化においては、符号化モードを適応的に選択するための評価方法として SATD を用いる意義は無いと言える。

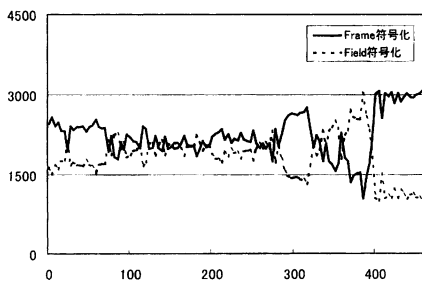
2.3 符号化モード選択結果

2.3.1 映像の種類と符号化モード選択結果

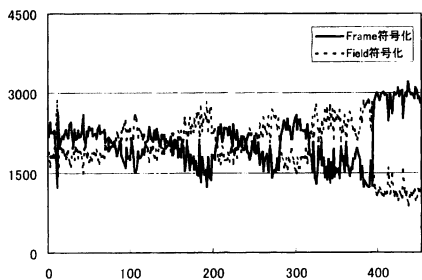
ここで、MB 単位で符号化モードを適応的に選択したとき、映像の種類によって符号化効率の改善効果が異なることについて考察する。



(a) I フレーム



(b) P フレーム

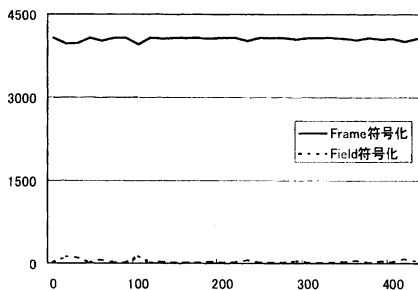


(c) B フレーム

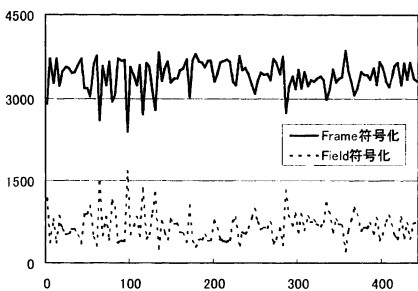
図 4 シーケンス 30

図 4.5 は、Sequence No.30 および Sequence No.19 を、MB 単位での符号化モード適応選択手法を用いて符号化した際に、フレーム符号化モードが選択された MB-pair 数およびフィールド符号化モードが選択された MB-pair 数をフレーム毎に集計したグラフである。横軸はフレーム番号であり、縦軸は各符号化モードが選択された MB-pair の数を表している。なお、フレームサイズは 1920 画素 × 1080 ラインであり、MB-pair の大きさは 16 画素 × 32 ラインであるため、1 フレームあたり 4080 個の MB-pair が存在する。符号化の条件は、表 1 に従うこととした。

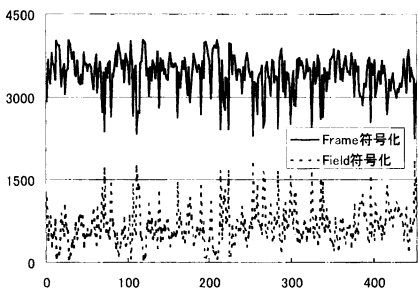
ここで、符号化モード選択がピクチャ単位であるときと MB 単位であるときの、符号化構造の関係について考察する。ピクチャ単位での符号化モード選択においては、各ピクチャ内の全ての MB が同一の符号化モードとなる。すなわち、ピクチャ単位での符号化モード選択は、MB 単位で全て同じ符号化モード



(a) I フレーム



(b) P フレーム



(c) B フレーム

図 5 シーケンス 19

が選択されることと等価である。

図 4 の結果については、フレーム符号化モードが選択された MB-pair とフィールド符号化モードが選択された MB-pair が同一フレーム内において混在している。このため、MB 単位での符号化モード適応選択により、効率的に符号化されたと考えられる。

一方、図 5 の結果については、同一フレーム内において、選択された符号化モードに偏りが生じているために、MB 単位で符号化モードを選択しているにも関わらず、ピクチャ単位での符号化モード選択と同等の性能に止まっていると考えられる。

2.3.2 動きベクトルと符号化モード選択結果

フレーム符号化モードおよびフィールド符号化モードそれぞれの特徴として、一般に、フレーム符号化モードは動きを持たない映像に適しており、フィールド符号化モードは動きが大き

い映像に適していると言われる。本節では、MB 単位での符号化モード選択について、符号化モードの判定結果と映像の動きの大きさの関連について検証する。

ITE HDTV 標準画像 10 シーケンスを表 1 に示す条件の下で符号化し、各シーケンスの B フレームおよび P フレームについて、符号化モード選択結果と動きベクトルの大きさとの相関を調べた。なお、コスト算出においては、SAD を用いた。表 3 は各シーケンスの 61 フレーム目 (B フレーム)、表 4 は各シーケンスの 33 フレーム目 (P フレーム) の結果である。表 3, 4 では、フレーム内の各 MB-pair について、動きベクトルの大きさが 0 (表中 $|MV|=0$ と記載) の MB-pair 数と 0 より大きい動きベクトルを持つ (表中 $|MV|>0$ と記載) MB-pair 数を、符号化モードごとに分けて表している。

表 3 B フレームにおける動きベクトルと符号化モード選択結果

Sequence No.	フレーム符号化		フィールド符号化	
	$ MV =0$	$ MV >0$	$ MV =0$	$ MV >0$
07	2870	0	617	793
12	931	0	33	3116
16	2653	0	157	1270
19	3188	0	752	140
20	1586	0	16	2478
23	1083	0	111	2886
25	1436	0	2124	520
30	2206	0	398	1476
34	2474	0	47	1559
43	889	0	30	3161

表 4 P フレームにおける動きベクトルと符号化モード選択結果

Sequence No.	フレーム符号化		フィールド符号化	
	$ MV =0$	$ MV >0$	$ MV =0$	$ MV >0$
07	2958	0	1122	0
12	2413	0	0	1667
16	2988	0	1092	0
19	3453	0	0	627
20	2205	0	0	1875
23	3174	0	0	906
25	3893	0	187	0
30	2339	0	1741	0
34	2906	0	1174	0
43	1059	0	3021	0

表 3 より、B フレームにおいては、フィールド符号化が選択された MB については動きベクトルが 0 および 0 より大きいものが存在する。一方、フレーム符号化が選択された MB については、動きベクトルが 0 のものしか存在しない。同様に、表 4 より、P フレームについても、フィールド符号化が選択された MB については動きベクトルが 0 および 0 より大きいものが存在する。一方、フレーム符号化が選択された MB については、動きベクトルが 0 のものしか存在しない。なお、他の B フレーム、P フレームについても同様の傾向が見られた。

以上より、SAD を用いたコスト評価に基づいて MB 単位で符号化モードの適応選択を行った場合、最終的にフレーム符号化が選択される MB については、0 より大きい動きベクトルを持たない傾向にあると考えられる。

3. 提案手法

SAD を用いたコスト評価に基づいて MB 単位で符号化モードの適応選択を行う手法について、符号化モード選択の処理量を削減する手法を提案する。提案手法を以下に示す。

イントラ符号化を行うフレームについては、図 1(a) に示す処理に従うものとし、インター符号化を行うフレームについては、以下の手順に示す処理に従うものとする。なお、インター符号化における提案手法の処理フローを図 6 に示す。また、コスト算出の処理については、図 3 に従うものとする。

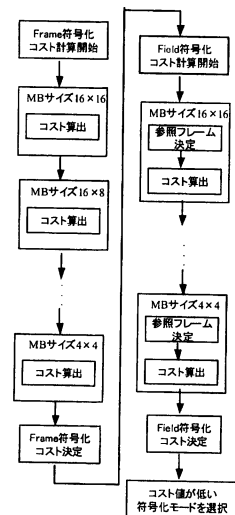


図 6 提案手法処理フロー

- Step.1 該当 MB-pair について、全ての MB サイズで Step.2 から Step.3 の処理を行う。最もコスト値が小さいモードのコスト値を、フレーム符号化モードのコスト値とし、Step.4 へ。
- Step.2 動きベクトルの大きさを 0 とし、参照フレームを決定する。
- Step.3 フレーム符号化を行いコスト値を算出する。
- Step.4 該当 MB-pair について、全ての MB サイズで Step.5 から Step.6 の処理を行う。最もコスト値が小さいモードのコスト値を、フィールド符号化モードのコスト値とし、Step.7 へ。
- Step.5 サーチレンジに基づいた範囲で動き補償を行い、参照フレームを決定する。
- Step.6 フィールド符号化を行いコスト値を算出する。
- Step.7 フレーム符号化モードおよびフィールド符号化モードにおけるコスト値を比較し、コスト値が小さい符号化モードを該当 MB-pair の符号化モードとする。

4. 実験

4.1 実験手順

提案手法の評価を行うために、JM に提案手法の処理を組み込み、符号化実験を行った。

符号化に用いた素材は ITE HDTV 標準動画像 (1920 画素 × 1080 ライン, 8bit) の 10 シーケンス (Sequence No.07, 12, 16, 19, 20, 23, 25, 30, 34, 43) 各 450 フレームである。符号化を行う際の条件を表 5 に示す。コスト値の算出には、SAD を用いた。

比較として、インター符号化の処理として図 1(b) に示す手順を行う処理を従来手法とし、表 5 に示す条件の下で符号化を行った。

4.2 実験結果

手法の評価は、符号化モード選択処理に要する処理時間の短縮率、符号化処理全体に要する処理時間の短縮率、SNR 値の変化量の 3 点について行った。

4.1 の実験についての評価結果を表 6 に示す。

表 5 符号化条件

設定項目	設定値
フレームサイズ	1920 画素 × 1080 ライン
フレーム周波数	30fps
プロファイル	High Profile
ビットレート	8Mbps
DCT 変換サイズ	4 × 4 および 8 × 8
符号化方式	CABAC
アダマール変換	なし
動き予測	フルサーチ
R-D 最適化	あり
GOP 構成	IBBPBBPBBPBBP
参照ピクチャ	I および P ピクチャ
参照ピクチャ数	2
動きベクトル探索範囲	± 32 × ± 32

表 6 符号化モード選択に要する処理量の改善効果

Sequence No.	SNR 変化量 (dB)	符号化モード選択処理短縮率 (%)	符号化処理短縮率 (%)
07	0.10	17.4	17.5
12	0.12	18.9	19.0
16	0.09	21.2	21.5
19	0.07	16.4	16.5
20	0.06	20.2	20.3
23	0.18	20.2	20.3
25	0.03	17.9	18.1
30	0.05	20.6	20.8
34	0.05	20.0	20.1
43	0.20	25.6	25.6

表 6 より、本手法による処理量の削減効果については、シーケンスによって効果に差が見られるものの、符号化モードの適

応選択処理に要する処理量を 20%程度削減されており、符号化処理全体についても処理量の 20%程度が削減された。一方で、本手法を用いることにより SNR 値が 0.1%程度低下した。

5. まとめ

本研究では、HDTV を低ビットレートにおいて符号化する条件の下で、符号化モード適応選択手法の有効性について検証を行い、MB 単位での符号化モード適応選択手法の有効性を確認した。また、この条件の下では、符号化モード適応選択に用いるコスト値の算出方法において、SAD と SATD どちらを用いた場合でも同等の結果が得られることを確認した。さらに、MB 単位の符号化モード適応選択において、符号化モード選択結果の偏りについて検証し、符号化モード選択における処理量の削減に向けた選択手法を提案した。最後に、提案手法の評価実験を行い、低ビットレートにおける HDTV の符号化において、符号化モードの適応選択処理量が 20%程度削減されることを確認した。

本研究では、符号化モードの選択手法として SAD を用いたコスト算出方法を採用したが、符号化モードの選択手法はフレーム符号化およびフィールド符号化の特徴を考慮した手法が望ましい。今後の課題として、SAD を用いた手法の妥当性を検証すると共に、符号化モード適応選択処理の最適化を行い、符号化モード選択処理の短縮および SNR 値の改善を図る。

なお、本研究は独立行政法人情報通信研究機構による委託研究「ソフトウェア符号化技術に関する研究開発」として実施したものである。

文献

- [1] ITU-T Rec. H.264 — ISO/IEC 14496-10, "Advanced video coding for generic audiovisual service", May 2003.
- [2] 谷沢, 古藤, 中條, "H.264 におけるレート-ひずみ最適化モード判定の高速化", 信学論 (D), Vol.J89-D, no.1, pp.27-39, Jan. 2006.
- [3] 常松, 渡辺, "H.264/AVC における Intra 予測モードの決定手法に関する検討", 信学論 (D-II), Vol.J88-D2, No.9, pp.1979-1981, Sep. 2005.
- [4] 高木, 酒澤, 滝嶋, "H.264 符号化における Intra 予測方式の効率化に関する一検討", 映像メディア学会冬季大会 '12-3, Dec. 2004.
- [5] Karsten Suhling, H.264/AVC Software Coordination, <http://ip.hhi.de/index.htm>