

HFR対応映像エンコーダに適した 予測処理軽量化手法に関する一考察

大森 優也^{1,a)} 大西 隆之¹ 岩崎 裕江¹ 清水 淳¹

概要: 近年、放送や映像配信サービスにおいて、4K や 8K に代表される高解像度化の流れが急速に進みつつある。高臨場感映像においては、このような高解像度化に加え、被写体の動きの滑らかさを表現するために、120P など高フレームレート (HFR) もあわせて重要となる。本稿では、120P 映像サービスに向けたリアルタイム HEVC エンコーダ実現に向け、大きな演算量を必要とする動き探索処理の削減を想定したうえで、それにより生じる符号化効率低下を抑制する動き探索手法について提案する。また、提案手法を適用した際の符号化性能の評価結果について示す。

Study of Efficient Motion Estimation Methods for High Frame Rate Video Encoders

YUYA OMORI^{1,a)} TAKAYUKI ONISHI¹ HIROE IWASAKI¹ ATSUSHI SHIMIZU¹

1. はじめに

近年、4K 映像を利用した配信サービスが次々と開始され、さらに 2016 年には 8K 映像の試験放送開始 [1] が予定されており、高解像度映像通信サービスへの期待が高まりつつある。また、高品質映像サービスの追求は、解像度の高度化のみで達成されるものではなく、映像の滑らかさを表現するために、フレームレートの高度化も不可欠である。ITU (国際電気通信連合) により勧告されている UHDTV (Ultra High Definition Television) [2] では、現状考えうる最高品質の映像規格として、図 1 に示す通り、解像度 8K、フレームレート 120P、画素値深度 12bit、色表現 4:4:4、かつ広色域、といった規格が策定されている。

フレームレートは、映像の本質ともいえる動きの滑らかさに直結する要素であり、主観評価実験によって、120P では 60P に比べ主観画質が倍増し、240P 付近で主観画質が飽和することが示されているが [3][4]、現行の映像サービスは主に 30P/60I, 60 である。120P の高フレームレート

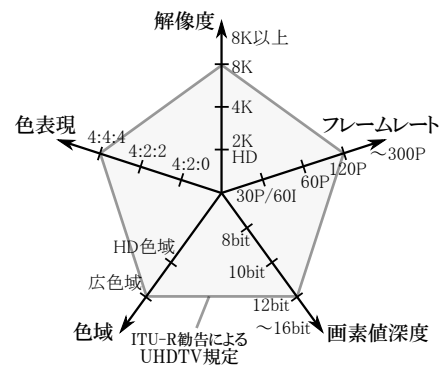


図 1 UHDTV による映像規格
Fig. 1 UHDTV video specification.

(HFR) 対応機器は、撮像系・表示系については登場しつつあるものの [5][6]、リアルタイムエンコーダについては、今後の開発が期待されている。そこで本稿では、先行して普及する 60P 対応の受信機およびデコーダでもデコード可能な "時間方向スケーラビリティ" への対応を前提とし [7]、60P 基本レイヤおよび 120P 拡張レイヤに分割されたストリーム構成に準拠する、120P HFR 対応リアルタイムエンコーダの実現に向けた符号化方式の検討を行う。

120P HFR 対応リアルタイムエンコーダ構成としては、60P 基本レイヤについては、既存の HEVC[8][9] エンコー

¹ 日本電信電話株式会社, NTTメディアインテリジェンス研究所
Nippon Telegraph and Telephone Corporation, NTT Media Intelligence Laboratories

^{a)} omori.yuya@lab.ntt.co.jp

ダを活用し、追加の 120P 拡張レイヤについては、FPGA (Field Programmable Gate Array) を用いた構成を想定する。FPGA は、拡張レイヤの要求条件に応じたハードウェア設計が可能であるが、一方で、ASIC タイプの LSI と比較して実装可能な回路規模は限定される。このため、拡張レイヤのエンコーダ実装時には演算量の大きい処理、代表的には動き探索処理について探索範囲を小さく限定することで演算量を減らす必要があるが、この場合、画面内の動きの速い物体に対して十分な動き追従が困難となる。特に拡張レイヤ特有の課題として、動き探索処理削減時に

- 同一ピクチャ内における遠方方向の動き探索精度低下
 - レイヤ内における特定のピクチャの動き探索精度低下が生じ、これらの要因が拡張レイヤの符号量増加を招く。
- これらの課題を解決するため、120P/60P 時間方向スケラビリティの参照構造の特性に着目し、符号化効率低下を抑制する以下の 2 つの動き探索手法を提案する。

- 近方方向動き探索結果を用いた遠方方向探索中心決定
- 拡張レイヤ間事前探索による探索中心決定

本提案手法で計算機シミュレーションを行った結果、120P 全体の符号化効率の低下を 10% 未満に抑制可能となる。

以下に本稿での構成を説明する。2 章では、120P/60P 時間方向スケラビリティ方式について説明し、時間スケラビリティに対応したエンコーダ構成方式の想定について述べる。3 章では、符号化効率劣化を抑制するための提案手法について説明する。4 章では、評価結果と考察について述べる。5 章では、本稿をまとめる。

2. 前提条件と課題

本章ではまず、120P エンコーダ実現時の前提条件として、時間方向スケラビリティと想定するエンコーダ構成について述べ、続いて、予測処理制限に起因する符号化効率低下の課題について示す。

2.1 HEVC 120P/60P 時間方向スケラビリティ方式

H.265/HEVC に準拠した時間方向スケラビリティ対応符号化方式として、ARIB (一般社団法人 電波産業会) STD-B32 3.3 版では 2160/120/P 映像及び 4320/120/P 映像の時間方向スケラビリティ方式が規定されている [7]。ARIB 規定では、図 2 に示すように、120P ビットストリームは 60P 基本レイヤと 120P 拡張レイヤとに分離されて伝送される。また、60Hz 対応受信機でのデコード時間の偏りを防ぐ目的で、基本レイヤと拡張レイヤの復号タイミングが交互に発生するよう、インター予測における特殊なピクチャ参照構造が規定されている。当該参照構造の例を図 3 に示す。TemporalID= 0, 1, 2, 3 に属するピクチャが 60P 基本レイヤ、TemporalID= 4 が 120P 拡張レイヤである。120P 拡張レイヤ内のピクチャは、60P 基本レイヤ内のピクチャに比べて近距離の参照が多い一方で、前方向動

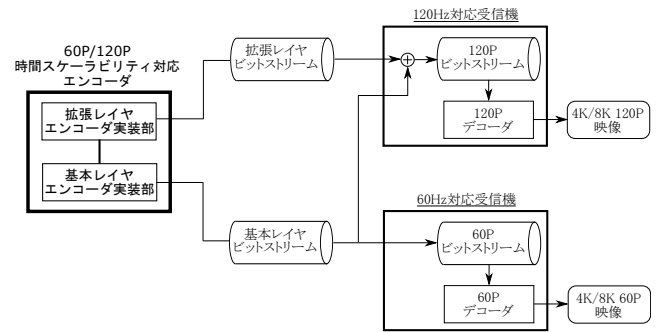


図 2 120P/60P 時間スケラビリティ対応映像配信
Fig. 2 120P/60P temporal scalable video distribution.

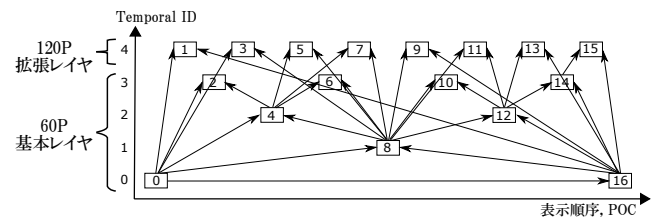


図 3 120P/60P 時間方向スケラビリティ符号化における参照構造 (ARIB STD-B32 規定)
Fig. 3 Reference image structure of 120P/60P temporal scalable encoding (ARIB STD-B32).

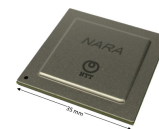


図 4 HEVC リアルタイムエンコーダ LSI “NARA”
Fig. 4 HEVC real-time encoder LSI “NARA”.

き予測と後方向動き予測で参照距離が異なる非対称性を有し、参照距離も大きく変動する点に特徴がある。

2.2 120P エンコーダ構成

本稿では、前節で述べた ARIB 規定のレイヤ構成および参照構造を遵守し、時間方向スケラビリティに対応した 120P リアルタイムエンコーダを想定する。60P 基本レイヤについては、既存の HEVC リアルタイムエンコーダ LSI (開発コード名: NARA[10]) を使用して符号化を行う (図 4)。これに対し、120P 拡張レイヤは新規構成が必要となり、拡張レイヤの要求条件に応じたハードウェア設計を可能とするため、書き換え可能なハードウェアである FPGA (Field Programmable Gate Array) による実装を想定する。

2.3 課題

映像符号化における動き探索処理では、一般的に、縮小画像による広範囲の事前動き探索処理や、近傍ブロックの動きベクトルを用いた予測ベクトルを中心とする近傍探索が行われることが多い。60P 基本レイヤにおいては、NARA LSI を用いて、図 5 に示すように、(1) 左近傍ブ

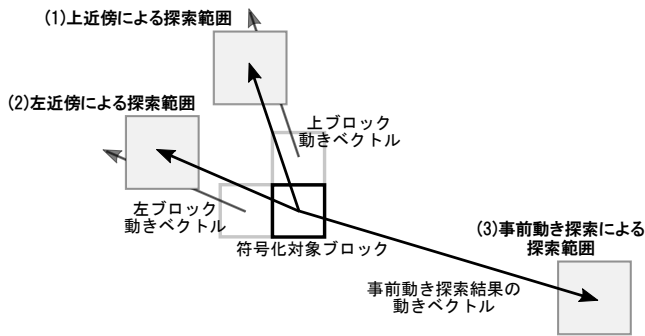


図 5 NARA LSI における動き探索中心
Fig. 5 Motion search center of NARA LSI.

ロック動きベクトル結果, (2) 上近傍ブロック動きベクトル結果, (3) 事前動き探索結果, の 3 点を動き探索中心とし, 3 つの探索範囲の中で符号化コストが最小となる点を探索結果ベクトルとして出力する。

一方で, FPGA は LSI と比較して実装可能な回路規模が限られるため, FPGA を用いた 120P 拡張レイヤのエンコード実装時には, 特に膨大な演算量を要する動き探索処理を大幅に削減しなければならない。具体的には, 探索範囲の削減, 及び事前動き探索処理の省略, が必要となる。しかしながら, 上記の動き探索処理削減による符号化効率の低下は不可避であり, 特に画面内に動きの速い物体が存在する場合は, 動きベクトルが探索範囲からはみ出し十分な動き追従を行うことができず, 符号化効率が大幅に低下する傾向にある。特に, 120P 拡張レイヤで特有の問題として, 動き探索処理の削減時には

- 同一ピクチャにおける動き予測参照距離の非対称性に起因する, 遠方方向の動き探索精度低下
- 同一 Temporal ID における動き予測参照距離の変動に起因する, 特定のピクチャの動き探索精度低下

が生じ, これらの要因が拡張レイヤの符号量増加を招く。

3. 提案手法

本章では, 前章で述べた課題を解決するため, 120P/60P 時間方向スケーラビリティの参照構造の非対称性に着目し, 以下の動き探索手法を提案する。

- 近方方向探索結果を用いた遠方方向探索中心の決定
- 拡張レイヤ間事前探索による探索中心の決定

3.1 近方方向探索結果を用いた遠方方向探索中心の決定

図 3 で示したように, 基本レイヤでは全てのピクチャで前方向と後方向の動き予測参照距離が等しいのに対し, 拡張レイヤではピクチャ 15 を除く各ピクチャにおいて, 前後方向の参照距離が異なる。ここで参照距離の近い方向を近方方向, 遠い方向を遠方方向と定義すると, 特にピクチャ 3 以外の拡張レイヤでは, 近方方向は隣接フレーム参照となっている。このような近方方向では小さな探索範囲でも

高い予測精度が期待できる一方, 遠方方向では動きが探索範囲からはみ出しやすく, 予測精度の低下が予想される。

以上のような拡張レイヤの特徴に注目し, 予測精度の高い近方方向動き探索を利用した, 遠方方向動き探索の探索中心決定手法について述べる。以下の処理は CTU^{*1} 単位で実行し, 本手法では, 近方方向動き探索を遠方方向動き探索よりも 1CTU 前に実行する。図 6 に, 処理の流れをフローチャートで示す。近方方向の参照フレーム間隔を d_1 , 遠方方向の参照フレーム間隔を $d_2 (\geq d_1)$ とする。近方方向動き探索, および遠方方向動き探索における探索中心を, それぞれ図 7, 図 8 に示す。

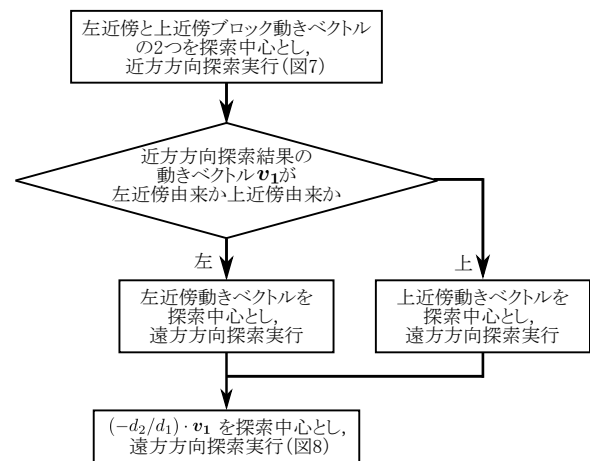


図 6 提案手法 1: フローチャート
Fig. 6 Proposed method 1: flowchart.

以上の処理により, 遠方方向探索においても, 予測精度の高い近方方向の動きベクトルを利用することで, 映像の大きな動きに追従することが容易となる。本手法は, 特に線形な動きを有する映像において大きな効果が期待できる。

3.2 拡張レイヤ間事前探索による探索中心の決定

図 3 で示すように, 60P 基本レイヤにおいては同一 Temporal ID 内の全ピクチャで参照距離が等しい一方, 120P 拡張レイヤ (Temporal ID=4) ではピクチャごとの参照距離が大きく変動する特徴がある。このため, 120P 拡張レイヤでは各ピクチャに動き予測精度の差が生じる。

特にピクチャ 3 では, 図 9 に示すように近方方向でも 3 フレーム離れた参照となるため, 前後方向ともに隣接フレーム参照となるピクチャ 15 と比べて数倍の符号量が発生する。このため, ピクチャ 3 の動き予測精度の向上が, 拡張レイヤ全体の符号量削減に対して有効と考えられる。

ここでピクチャ 1 に着目する。ピクチャ 1 は図 9 の通り近方方向が隣接フレーム, 遠方方向が 15 フレーム先の参照となっており, 非対称性が最も強く, 最終的な符号化結果も近方方向のみの片方向予測となるケースが多い。そこ

^{*1} Coding Tree Unit. H.264/HEVC におけるブロック分割処理の起点

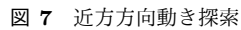


図 8 遠方方向動き探索

Fig. 9 Reference image structure of picture 1 and 3.

3 フレーム離れた近方方向探索を、より短いフレーム間隔による動きベクトル結果 $v_{3 \rightarrow 1}$ と $v_{1 \rightarrow 0}$ の加算として表現することで、動きへの追従性を高めることを目的とする。符号化効率への影響が少ないピクチャ 1 の遠方方向探索処理を、ピクチャ 3 の拡張レイヤ間事前探索処理に割り当てることで、ピクチャ 1 の符号量増加を抑えながらピクチャ 3 の符号量を削減し、拡張レイヤの符号化効率を改善することが期待できる。^{*2}

```

graph TD
    A[ピクチャ1の近方方向動き探索を  
左近傍と上近傍ブロック動きベクトル  
を探索中心として実行(図8)] --> B{ピクチャ1の  
近方方向動きベクトルが  
左近傍由来か上近傍由来か}
    B -- 上 --> C[上近傍動きベクトルを  
探索中心とし、  
ピクチャ1遠方方向探索実行]
    B --> D[左近傍動きベクトルを  
探索中心とし、  
ピクチャ1遠方方向探索実行]
    C --> E[ピクチャ3から1への  
拡張レイヤ間事前動き探索を実行  
(左近傍と上近傍のいずれか一方)]
    D --> E
    E --> F[左近傍と上近傍のいずれか一方  
を探索中心として  
ピクチャ3の近方方向動き探索を実行]
    F --> G[ピクチャ3から1への動きベクトルに  
ピクチャ3の近方方向動きベクトルを合計した  
 $v_{3 \rightarrow 1}(x, y) + v_{1 \rightarrow 0}((x, y) + v_{3 \rightarrow 1}(x, y))$   
を探索中心として(図11)  
ピクチャ3の近方方向動き探索を実行]

```

Fig. 10 Proposed method 2 : flowchart.

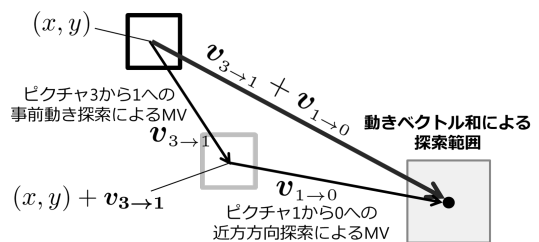


図 11 拡張レイヤ間事前動き探索を利用した探索中心決定
Fig. 11 Search center decision using pre-motion search in the enhancement layer.

本実験で用いる画像シーケンスは、図 12 に示す 4 つの 2K/120P 評価映像である。評価映像パラメータは表 1 の通りである。全て画面内の動きが大きい映像であり、評価映像 1 および 2 は、静止背景画面内に動きの速い物体が存在し、評価映像 3 および 4 では、カメラのパンにより画面全体が大きく動く。符号化パラメータを表 2 に示す。それぞれの映像において、TemporalID = 0 のピクチャを QP = 22, 27, 32, 37 の 4 種類で符号化を行い、TemporalID が 1 上がるごとに QP を 1 上げて符号化シミュレーションを実行している。

表 3 に実験結果を示す。拡張レイヤの動き探索処理削減

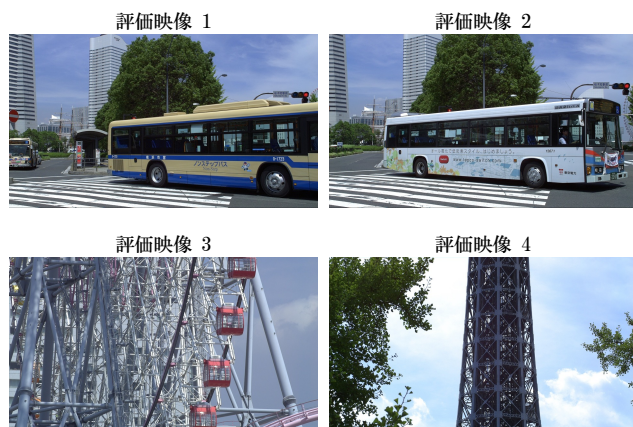


図 12 120P 評価映像

Fig. 12 120P video sequence.

表 1 評価映像パラメータ

Table 1 Video sequence parameters.

項目	条件
解像度	1920 × 1080
フレームレート	120P
画素値深度	8bit
色フォーマット	4:2:0
フレーム数	65 フレーム

表 2 符号化パラメータ

Table 2 Encoding parameters.

項目	条件
GOP size	16
Intra period	64
CTU size	64
QP	22, 27, 32, 37

として、FPGA 実装を見込み、事前動き探索処理の削除、及び探索範囲の削減 (± 8 点 $\rightarrow \pm 3$ 点) を行っている、動き探索処理を削減せずに符号化を行った場合の結果をアンカーとして、符号化効率の劣化比率 (BD-bitrate) を示している。動き探索処理削減時における、提案手法適用前の符号化効率劣化を示し、さらに、2つの提案手法をそれぞれ個別に適用した結果、および両提案手法を同時に適用した結果、を示している。表 3 には、拡張レイヤのみに注目した符号化効率劣化と、基本レイヤを含めた 120P 全体の符号化効率劣化の、両方を示す。

図 13 に、拡張レイヤにおける R-D 曲線を示す。動き探索処理削減前、動き探索処理削減後、動き探索処理削減かつ両提案手法適用後、の 3つの R-D 曲線をそれぞれ示す。

4.3 考察

提案手法 1 は、映像における動きの線形性を利用した手法であるため、特に、動きが線形かつ大きい映像に対して遠方方向の動き探索精度低下が抑制される。このため、評

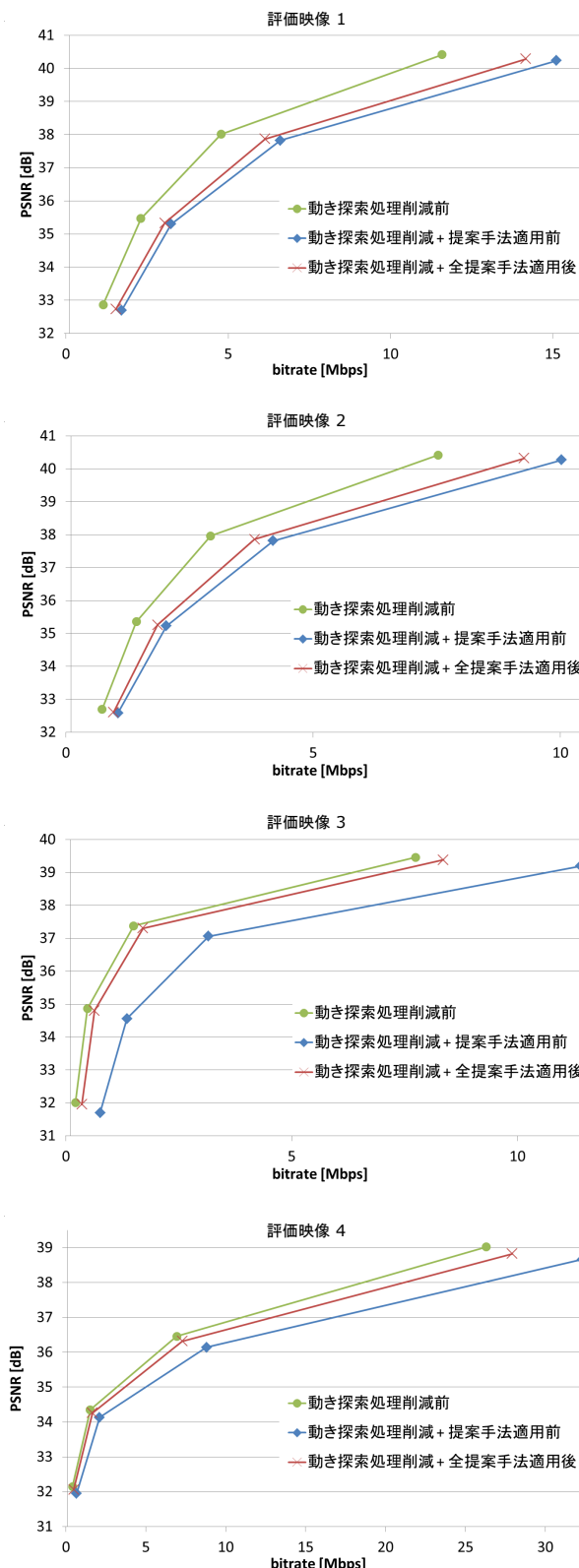


図 13 拡張レイヤにおける R-D 曲線

Fig. 13 Coding efficiency of the enhancement layer

価映像 3 や評価映像 4 のように画面全体が大きくパンをする映像において、提案手法を適用せずに動き探索処理削減のみを行った場合に比べ、拡張レイヤにおいて 35% 以上の劣化抑制効果が示されている。

提案手法 2 は、各ピクチャにおける予測精度及び符号量

表 3 実験結果：動き探索処理削減時の符号化効率劣化

Table 3 Experiment results : deterioration of encoding efficiency

		提案手法適用前	提案手法 1 適用後	提案手法 2 適用後	両提案手法適用後
評価映像 1	全体	11.7%	10.7%	10.2%	8.9%
	拡張レイヤのみ	46.9%	42.9%	40.3%	35.1%
評価映像 2	全体	10.4%	9.5%	8.7%	7.4%
	拡張レイヤのみ	48.0%	44.5%	39.7%	33.6%
評価映像 3	全体	16.3%	6.7%	14.1%	2.9%
	拡張レイヤのみ	181.9%	81.3%	150.7%	31.5%
評価映像 4	全体	12.4%	4.8%	11.0%	3.9%
	拡張レイヤのみ	57.9%	22.5%	49.2%	16.3%

のばらつきに着目し、特に符号量の大きいピクチャの効率改善を行う手法であるため、符号量のばらつきが大きい映像、符号量が大きい映像に対し、効果を発する。このため、評価映像 3 のように動き探索処理削減による符号量増加が大きい映像において、拡張レイヤにおいて 30% の劣化抑制効果が示される。

両提案手法適用時は、提案手法を適用しない場合に比べ、拡張レイヤのみにおいて 11.8% 以上、基本レイヤを含めた 120P 全体において 2.8% 以上の符号化効率劣化抑制が見られる。120P 全体においても、符号化効率劣化が最大 8.9% 以内に抑えられている。

評価映像 3 では、提案手法 1, 2 を個別に適用した場合以上の符号化効率劣化抑制が、両提案手法適用後に確認できる。これは、提案手法 2 によってピクチャ 3 の近方向予測の精度が向上した結果、提案手法 1 によってピクチャ 3 の遠方向の予測精度も向上し、劣化抑制において相乗効果が生じたためだと考えられる。一方で、評価映像 4 では、両提案手法適用後の符号化効率劣化抑制の効果は、提案手法 1, 2 を個別に適用した場合よりも低下している。提案手法 1 が提案手法 2 と併用される場合、ピクチャ 1 の遠方向予測は近傍ブロック探索中心のみとなり提案手法 1 を適用できないため、提案手法 1 によるピクチャ 1 の符号量抑制効果が大きい評価映像 3 では、両提案手法適用後に個別適用時よりも若干の効果低下が生じたと考えられる。

図 13 では、全ての評価映像において提案手法適用前より提案手法後の R-D 曲線が優位であり、特に評価映像 3 および 4 のように、画面全体が大きく動く映像において提案手法は大きな効果を示している。

5. おわりに

本稿では、被写体の動きの滑らかさを表現するためのフレームレート向上を主たる目的と定め、現状の 60P を超える 120P HFR 対応リアルタイム映像エンコーダ実現に向けた検討を行った。120P/60P 時間方向スケーラビリティ対応のエンコーダとし、大きな演算量を必要とする動き探索処理の大幅な削減を想定したうえで、それにより生じる符

号化効率低下を抑制するため、時間方向スケーラビリティの参照構造の特性を利用した動き探索手法について提案した。(1) 近方向動き探索結果を用いた遠方向探索中心の決定 (2) 拡張レイヤ間事前探索による探索中心の決定の 2 つの提案手法を合わせて適用することで、探索処理を大幅に削減しつつも全体符号量の増加を 10% 未満に抑制可能であることを示した。今後は、基本レイヤを含め全体の回路構成について検討を深め、4K/8K 120P HFR 対応リアルタイム映像エンコーダ実現に向けた検討を進める。

参考文献

- [1] 総務省, 4K・8K ロードマップに関するフォローアップ会合 中間報告参考資料, http://www.soumu.go.jp/main_sosiki/kenkyu/4k8kroadmap/, Sep. 2014.
- [2] ITU-R BT.2020, "Parameter values for ultra-high definition television systems for production and international programme exchange", Aug. 2012.
- [3] Y. Kuroki, et al., "Improvement of motion image quality by high frame rate," SID '06. vol. 37. no. 1, pp. 14-17, 2006
- [4] 江本 正喜, et al., "高フレームレート撮像テレビジョンシステムの動画質改善効果," 映像情報メディア学会誌, vol. 65. no. 8, pp. 1208-1214, 2011
- [5] H. Shimamoto, et al., "120Hz Frame-Rate Super Hi-Vision Capture and Display Devices," SMPTE Mot. Imag. J., vol. 122, no. 2, pp. 55-61, 2013
- [6] T. Kumakura, "8K LCD: Technologies and challenges toward the realization of SUPER Hi-VISION TV," The 20th Asia and South Pacific Design Automation Conference, pp. 680-683, Jan 2015.
- [7] 電波産業会, "デジタル放送における映像符号化、音声符号化及び多重化方式 ARIB STD-B32 3.2 版," 2015.
- [8] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, no. 7, pp. 560-576, December 2012.
- [9] D. Flynn, et al., "High Efficiency Video Coding (HEVC) Range Extensions text specification: Draft 7," ITU-T SG16 WP3 and ISO/IEC JCT1/SC29/WG11 Doc. JCTVC-Q1005, June 2014.
- [10] Takayuki Onishi, et al, "Single-chip 4K 60fps 4:2:2 HEVC Video Encoder LSI with 8K Scalability," Dig. Symp. VLSI Circuits, pp.C54-CS55, June 2015.