

画像解析を用いた H.265/HEVC 符号化における 画面間予測の切替に関する一検討

坂本悠輔^{†1} 雑賀新太郎^{†1} 竹内健^{†1} 松尾康孝^{†2} 甲藤二郎^{†1}

概要 : H.265/HEVC では、従来符号化法と比較して、符号化効率の向上と引き換えに処理コストが増大している。特に、高解像度画像や高フレームレート画像においては、処理コストの増大が顕著である。このため、なるべく符号化効率を落とさずに、処理コストを減少させる手法が求められる。本研究では、周波数解析、フレーム間特徴点マッチングといった映像特徴量解析の結果を基に、事前の画像解析を用いることで、画面間予測を用いて符号化するかを推定し、高効率、低負荷な符号化を行うための手法について提案する。

キーワード : H.265/HEVC, 画面内予測, 画面間予測, DFT, SURF

1. はじめに

H.265/HEVC では、符号化効率の向上とともに処理コストも増大している。特に画面間予測を用いる場合、動き補償に多くの処理時間を要する[1]。しかし、動画像の中には画面内予測のみを用いた方が画面間予測を用いるよりも符号化効率が高くなるものも存在する。そこで本研究では、事前の画像解析の結果を用いて画面間予測を用いて符号化するかを決定し、符号化にかかる処理コスト減少させる手法を検討する。

2. 符号化実験

高解像度画像における H.265/HEVC の符号化効率と処理時間を調査するために、予備実験として HM16.15 による符号化実験を行った。符号化方式には intra(画面内予測), と randomaccess(画面間予測)の二つを用いた。使用動画像は図 1 に示す 3 種類の画像と、これらの画像にフィルタを施した 3 種類の合計 6 種類である。いずれも 8K(7680 x 4620)解像度, YUV4:2:0 形式, フレームレートは 59.94Hz, ビット深度は 8bit である。フィルタは図 2 のような周波数特性を持つ 13x13 タップの 2 次元 FIR フィルタを用い, SteelPlant は高周波成分を減衰させ, TrainsA, Inaho は高周波成分を強調させた。以降, フィルタを施した動画像については名前の末尾に”F”をつける。

図 3 に符号化の結果を示す。一定方向への動きを持つ TrainsA, および TrainsAF は, 画面間予測を用いることで符号化効率が向上した。一方, 火花が激しく飛び散る複雑な

動きを持つ SteelPlant は画面内予測のみを用いた方が符号化効率は高くなった。ただし, SteelPlantF は TrainsA と同様画面間予測を用いることで符号化効率は向上した。画面全体にわたって稲穂が揺れる動画像である Inaho については, TrainsA と同様の傾向であったが, InahoF では SteelPlant に近い傾向となった。

処理時間については, どの動画像についても画面間予測を用いると, 画面内予測のみを用いた場合と比べて 2.5 倍から 3 倍程度の増加がみられた。これらの結果から, 事前の解析によって適切な符号化方式を選択することができれば, 符号化効率を保ったまま処理時間を抑えることができると推測される。

3. 画像解析

2 章で示したような, 画面内予測のみを用いた方が符号化効率が高くなる動画像にみられる特徴を調べるため, 2 章で用いた 6 種類の動画像それぞれに対して空間周波数解析と, 動きの解析を行った。空間周波数解析では, 動画像中の 1 フレームを抜き出し, 水平方向, および垂直方向それぞれに 1 次元 DFT を施す。得られたパワースペクトルについて, 全帯域のうち, 直流成分を含む低周波側の 25%にあたる帯域幅がもつエネルギーが全エネルギーの何%を占めるかを測定した。測定結果を表 1 に示す。SteelPlant, InahoF, TrainsAF は他と比べて低い値となっている。画面内予測のみを用いた方が符号化効率が高くなる動画像は高周波成分を多く含むことが分かるが, SteelPlant, InahoF のみにみられる特徴ではない。

動きの解析では, [2]で利用されている SURF 特徴量によるフレーム間特徴量マッチングを行った。マッチング点間



図 1 使用動画像

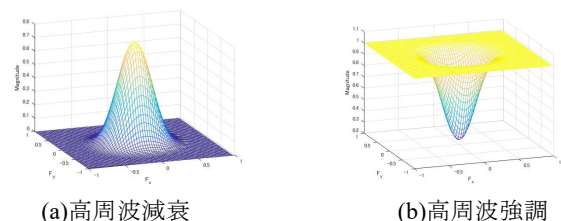


図 2 フィルタの周波数特性

^{†1} 早稲田大学
Waseda University
^{†2} NHK 放送技術研究所
Science and Technology Research Laboratories, NHK

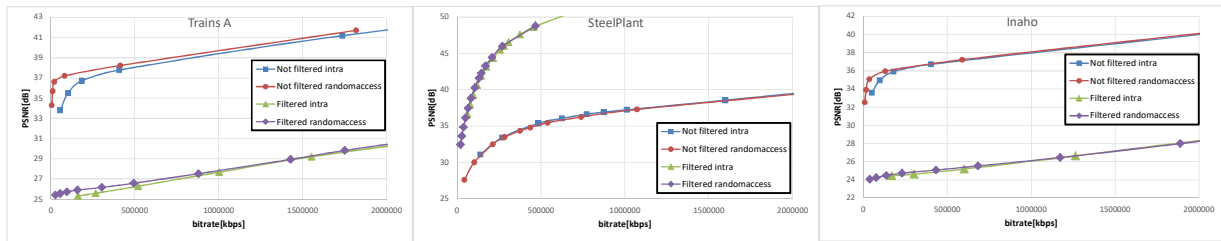


図 3 符号化実験結果 (RD 曲線)

表 1 空間周波数解析結果

sequence	horizontal	vertical
TrainsA	80.6154	84.71474
SteelPlant	62.59716	64.61094
Inaho	76.49665	79.82578
TrainsAF	56.9478	62.60855
SteelPlantF	92.06148	93.12108
InahoF	50.54624	54.34577

表 2 特徴点マッチングの結果

sequence	keypoints1	keypoints2	matched points	matched ratio[%]	Variance dist	Variance angle
TrainsA	11402	12664	5035	39.75837	3765.773	78.2675
SteelPlant	19576	34709	220	0.633841	2294711	8846.175
Inaho	7305	6811	10	0.136893	11.24202	4457.417
TrainsAF	24661	25919	3680	14.19808	2599.744	161.2923
SteelPlantF	12971	23200	216	0.931034	2088043	8158.992
InahoF	17188	16698	0	0	0	0

の距離および角度の分散に加えて、マッチング点の数、および検出点数に対するマッチング点数の割合も計測した。表 2 に計測結果を示す。表 2 のうち、keypoints1 は基準フレームで検出された特徴点の数、keypoints2 は基準フレームから 16 フレーム先のフレームで検出された特徴点の数、matched points はマッチングが取れた特徴点の数、matched ratio は keypoints1 または 2 のいずれか大きい方の値に対する matched points の値の割合である。SteelPlant は、距離と角度の分散値は他と比べて極端に大きな値となっている。しかし、SteelPlantF も同様の結果である。また、Inaho, InahoF は特徴点は検出できているものの、マッチング点数は他と比べて極端に少なく、マッチングが取れていない。

以上の結果から、画面内予測のみを用いた方が符号化効率が高くなるかどうかは、空間周波数と動きの二つの要素を合わせて判断する必要があるといえる。

4. 提案手法

3 章での実験結果を踏まえて、画像解析の結果を用いて符号化方式を決定する手法を提案する。フローチャートを図 4 に示す。提案手法では、初めに動画の 1 フレームを対象に空間周波数解析を行う。得られたパワースペクトルについて、低周波成分へのエネルギーの集中度合いが閾値を下回った場合は SURF 特徴点マッチングを行う。マッチング点間の距離および角度の分散値が閾値を超えた場合、または閾値を下回った場合で、検出点数に対するマッチング点数の割合が閾値を下回った場合は画面内予測のみを用いる。そうでない場合は画面間予測を用いる。画面内予測のみを用いたほうが符号化効率が高くなるような動画の場合、提案手法を用いることで処理コストを減少させることができる。

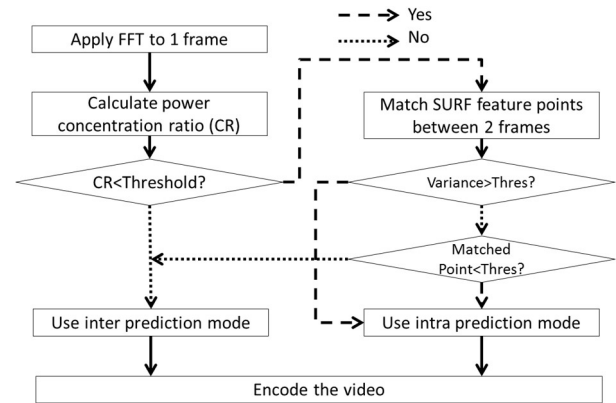


図 4 提案手法(フローチャート)

5. おわりに

本稿では、初めに符号化実験により画面間予測を用いても符号化効率が向上しない動画があることを示した。そして、空間周波数解析および SURF 特徴量による動きの解析を行い、それらの動画像にみられる特徴を明らかにした。最後に、符号化実験と画像解析の結果を踏まえて、事前の解析の結果を用いて符号化方式を決定するための手法を提案した。今後は、より多くの動画像に対しても解析を行うこと、提案手法を実装して符号化実験を行い、提案手法にかかる処理コストを評価することが課題である。

参考文献

- [1] F. Bossen, B. Bross, K. Suhring, D. Flynn, HEVC complexity and implementation analysis, IEEE Trans. Circuits Syst. Video Technol. 2012, vol. 22 p. 1685-1696.
- [2] S. Orihashi, R. Harada, Y. Matsuo, J. Katto, "Improvement of H.265/HEVC encoding for 8K UHD TV by detecting motion complexity," Consumer Electronics (ICCE), 2016 IEEE International Conference on. IEEE, 2016.