

## VVC/H.266におけるアフィン予測効率化の一考察

## A study on the efficiency of affine prediction in VVC/H.266

川口 昇太<sup>†</sup> 丹羽 直也<sup>†</sup> 岩崎 裕江<sup>†</sup>  
 Shota Kawaguchi Naoya Niwa Hiroe Iwasaki

## 1. はじめに

映像トラヒックの増大に伴い、IP トラヒックが爆発的に増加したことから、更なる符号化効率向上に向けて、2020年にH.266/VVC(Versatile Video Coding)が標準化された。VVCは、新たに360度映像に対する効率的な符号化も可能である。VVCは前世代のHEVCと比較すると、主観評価において、約50%符号化効率を改善しており、優れた性能を実現している一方で、符号化時間がHEVCの10倍に増加し、リアルタイムでの符号化を困難にしている[2]。そのため、VVCエンコーダの様々な高速化手法が研究されている[3]。

VVCでは、より高い符号化効率を実現するために、新たに様々な符号化ツールが導入され、動き補償には、より複雑な動きに対応したアフィンモードが追加されている。

本研究では、VVCのリファレンスソフトウェアVTM(VVC Test Model)にて、各テストシーケンスのエンコード結果に与えるアフィンモードの影響を分析し、アフィンモードによる予測の効率化につなげる考察を行った。

## 2. 動き補償及びアフィンモードについて

## 2.1 動き補償

動き補償では、現在のフレームの各領域が、過去のフレームのどこから移動してきたピクセル群なのかを動きベクトル(Motion Vector: MV)として算出し、MVを用いて移動後のピクセル位置を予測した画像を生成する。この予測画像と実際の画像との差分を符号化する事で、より少ないビットでの符号化を実現している。映像符号化は $64 \times 64 \sim 4 \times 4$ 画素の符号化単位(Coding Unit: CU)ごとに行われ、MVはCUにつき一本のみ算出される。このように一本のみのMVで予測画像を生成する動き補償は並進動き補償(Translational Motion Compensation: TMC)という。

## 2.2 アフィンモード

アフィンモードは、CUを $4 \times 4$ 画素のサブブロックに分割し、それぞれで並進動き補償を行う事で回転や拡大・縮小等の複雑な動きに対応する。アフィンモードでは図1に示すようにCUの隅にControl Point Motion Vector (CPMV)というベクトルを設ける。CPMVの配置方法には、2隅に配置する4パラメータモデルと3隅に配置する6パラメータモデルがある。これらのCPMVを用いて、線形補間の計算を行い、サブブロック単位のMVを決定する。

## 3. 実験

本研究では、様々なテストシーケンスにおいて、アフィンモードによる符号化効率と符号化時間の変化を評価することで、アフィンモードが有効な映像や効率化の余地がある

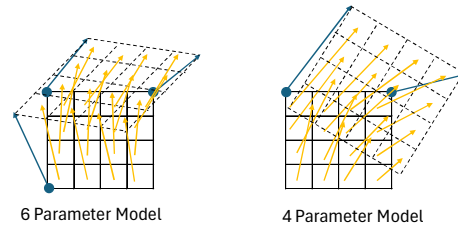


図1 サブブロックと CPMV

映像の傾向を調査する。

## 3.1 実験設定

VTMバージョン20.2を用いて映像のエンコードを行った。エンコード対象のテストシーケンスは、超高精細・広色域標準動画像Bシリーズ[4]とし、QPを22, 27, 32, 37の場合で実施した。それぞれ冒頭から数えて360から392の33フレームを使用した。アフィンモードが有効な時と無効な時でのエンコード結果を取得した。

## 3.2 評価指標

評価指標として、符号化時間増加率とBD-rateを用いた。符号化時間増加率は、アフィンモードが無効な時に対して、アフィンモードが有効な時の符号化時間の変化率を

$$R_{EncT} = \frac{EncT_{Affine} - EncT_{no\_Affine}}{EncT_{no\_Affine}} \quad (1)$$

とする。ただし、 $R_{EncT}$ は符号化時間増加率、 $EncT_{Affine}$ はアフィンモードが有効の時の符号化時間、 $EncT_{no\_Affine}$ はアフィンモードが無効の時の符号化時間を表す。

BD-rateは基準となるエンコード結果に対して、同等の画質を得る時に必要なビットレートの相対的な量を示す指標である。本実験では、アフィンモードが有効な時を基準として、アフィンモードが無効な時のビットレートの相対的な量をそれぞれ算出している。正の方向に大きい値をとるほど、アフィンモードの符号化効率への寄与が大きい事を意味する。

## 3.3 実験結果

実験結果は、図2に符号化時間増加率、図3にBD-rateを各シーケンス別に並べて、左から降順で示している。

結果を見ると、HorseRaceシリーズは符号化時間も長くなるものの、アフィンモードによって符号化効率が改善している傾向がある。一方で、WaterPoloシリーズ、MarathonシリーズはBD-rateがほとんど変化しておらず、アフィンモードの効果が無いといえる。これらのケースでは、符号化時間のばらつきが大きい。

<sup>†</sup>東京農工大学 Tokyo University of Agriculture and Technology

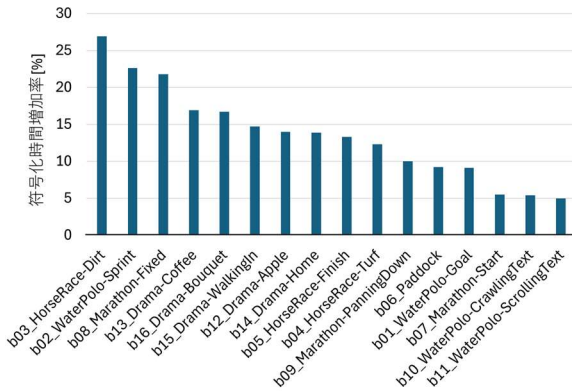


図2 アフィンモードによる符号化時間の変化

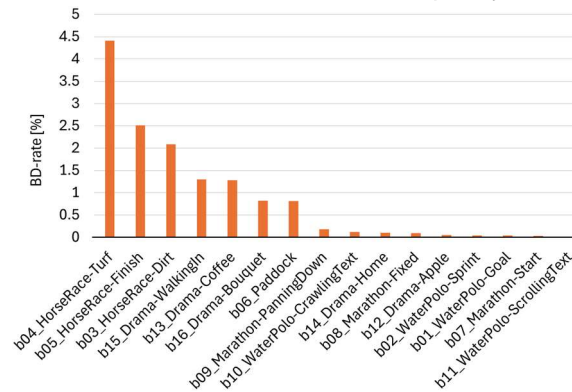


図3 アフィンモードによる符号化効率の変化

## 4. 考察

符号化効率と符号化処理のオーバーヘッドの観点で考察を述べる。

### 4.1 符号化効率

HorseRace シリーズはいずれも、アフィンモードによる改善度が高い。これらの映像は図4のように、カメラがズームインまたはパンフォローされる映像になっており、アフィンモードによる予測と映像の特徴が合致しているためと考えられる。

図5に示した、大量のマラソンランナーが道路を埋め尽くしている様子を上から撮影している Marathon シリーズ、水球の模様を写した WaterPolo シリーズは符号化効率の傾向が近い。Marathon シリーズではランナー、WaterPolo シリーズでは水面の動きに対して予測を試みたが、動きが細かく、入り乱れている箇所もあるので効率的な予測に繋がらなかったのではないかと考えられる。

### 4.2 符号化処理のオーバーヘッド

Marathon シリーズ、WaterPolo シリーズでは、BD-rate の変化がほとんど無いのに対して符号化時間増加率は 20% 台のケースも存在し、オーバーヘッドがアフィンモードによる利得を上回っているといえる。ランナーや水面の激しく細かい動きが占める画面上の割合が大きいケースほどオーバーヘッドが大きい傾向がある。

細かく複雑で動きを予測する CPMV は各 CU で向いている方向が乱れると予想でき、他のケースと比べて近傍 CU を参照して収集した CPMV 候補が別の方向を向いている事

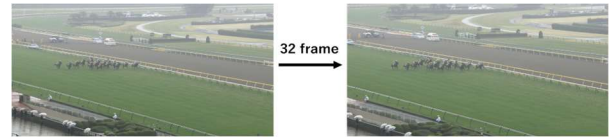


図4 b04\_HorseRace-Turf の映像



図5 アフィンモードによって符号化効率が改善しない映像が多いのではないかと考えられる。このような場合には、事前に映像の特徴量で判断することによる高速化が効果的であると考えられる。

## 5. 終わりに

本研究では、VVC の符号化時間を削減するため、特に複雑な動きの予測に用いられるアフィンモードが符号化時間に与える影響を分析した。エンコード結果から、映像の内容によってはアフィンモードが符号化効率にほとんど寄与しないにもかかわらず、符号化時間を増加させるケースがあることを示した。

今後の研究では、より効率的な符号化を実現するため、アフィンモードで生成される CPMV 候補が示す方向における統計を分析し、適応的にアフィンモードを無効化する実装の導入を検討している。

### 謝辞

本研究の一部は JSPS 科研費 23K24827, 24K20751 及び一般財団法人テレコム先端技術研究支援センター(SCAT)、公益財団法人放送文化基金、公益財団法人電気普及財団、2024 年度国立情報学研究所公募型共同研究 (24S0103, 24S108) の支援による。

### 参考文献

- [1] 一般社団法人 電波産業会 デジタル放送システム開発部会, “地上デジタル放送方式高度化に関わる 適用技術検討作業 中間報告”, 一般社団法人 電波産業会, (2021).
- [2] Qinghua Sheng, Hongzhao Chen, Changcai Lai, Xiaofeng Huang, Xiaofang Huang, Yuanyuan Liu, and Haibing Yin, “Fast Linear Equation Solving Algorithm and its Pipelined Hardware Architecture Design for VVC Affine Motion Estimation”, IEEE Transactions on Circuits and Systems for Video Technology, Vol.34, Issue.11 (2024).
- [3] Rikita Uchiyama, Karin Onouchi, Naoya Niwa, Masayuki Sato, Hiroaki Kobayashi, and Hiroe Iwasaki, “A Fast Block Partitioning Decision Method Using Luminance Textures for VVC Encoders”, IEEE International Conference on Consumer Electronics (2025).
- [4] 一般社団法人 映像情報メディア学会, 一般社団法人 電波産業会, “超高精細・広色域標準動画画 - B シリーズ 解説書”, 一般社団法人 映像情報メディア学会, 一般社団法人 電波産業会, (2017).