

H. 265/HEVC 動きベクトルの深層学習に基づく シーン認識性能の評価

宮地 美海矢[†] 神保 悟[†] 王 冀[†] 八島 由幸[†]

[†] 千葉工業大学大学院情報科学研究科

1. はじめに

近年、4K・8K といった超高精細映像が普及しつつあるが、その情報量の膨大さゆえ、映像コンテンツは H.265/HEVC (以下 HEVC) に代表される映像圧縮が施された状態で取り扱われることが少なくない。一方、動画像の認識は動画のカテゴリ分類や人物の行動認識など幅広く応用されており、近年では2つの CNN を並列で利用する Two-Stream CNN[1]が高いパフォーマンスを示している。動画像のシーン認識の課題の一つにリアルタイム性の確保がある。本研究では HEVC 符号化データを部分デコードすることによって得られる情報のみから動画像のシーン認識を高速に行う手法を検討する。

2. 提案手法

文献[2]では映像の符号化時に得られる動きベクトルを利用することによってこの問題の解決を図っているが、本研究ではこれをさらに拡張し、Two-Stream CNN の枠組みを応用して認識を行う。認識の流れを図 1 に示す。Two-Stream CNN は入力に RGB 画像とオプティカルフローを利用するが、本研究では H.265/HEVC の符号化データから部分デコードで得られる情報を入力データとして扱う。符号化ストリームから抽出される情報の例として DCT 係数、動きベクトル、イントラインターフラグ、ブロック分割構造などが挙げられる。本稿では初期検討として HEVC の動きベクトルを利用した認識性能を評価する。

3. 実験結果

実験には UCF101 データセット[3]から 5 つのクラスを利用する。解像度は 320×240 である。HEVC の動きベクトル検出はブロックサイズ 8×8 固定、探索範囲±16、1/4 画素精度、全探索で行った。イントラブロックについては動きベクトルは伝送されない。CNN の学習には動きベクトルの水平成分と垂直成分それぞれの絶対値を画像化して用いる。イントラ予測ブロックについては動きに寄与しないと考える。今回、比較のためにイントラ予測ブロックでも動きベクトルを求めて CNN 設計に利用する手法も試した。動きベクトル画像は 40×30 の解像度となるが、N フレーム分の動きベクトルを水平成分、垂直成分の順に交互に重ね、合計で 2N フレーム分を CNN 入力とした。

実験結果を表1に示す。表 1 より、フレーム数 N を増やすほど認識率は高くなるが、動きベクトルのみでは十分な認識性能は得られない。また、イントラ予測ブロックの動き

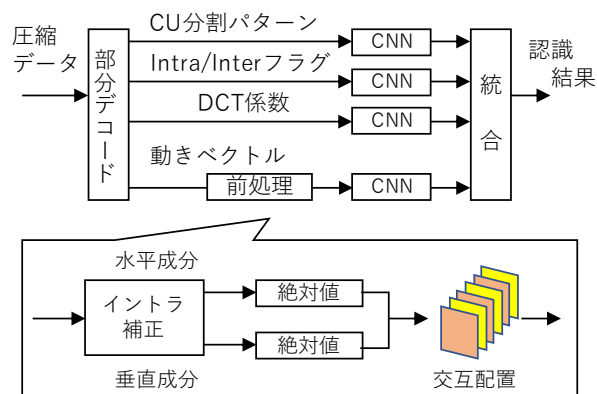


図1. 符号化データによる動画像シーン認識

表1. 実験結果(10回の実験の平均値, 単位%)

| フレーム数(N) | 25 | 50 | 75 |
|----------------|------|------|------|
| イントラベクトル無し(=0) | 56.3 | 62.2 | 62.6 |
| イントラベクトル有り | 59.6 | 61.9 | 64.6 |

ベクトルは認識率を向上させる傾向にあるが、「0」と取り扱うことの是非は結論付けできない。送信されてこないイントラ予測部分の動きベクトルを何らかの形で補完する実験も必要であり、さらに符号化データに含まれる動きベクトルが実際の動きを反映しているとは限らないため、信頼性の高いベクトルに整形するような前処理も必要と考える。

4. まとめ

本稿では、符号化データを用いた動画像のシーン認識の着手段階として、HEVC の動きベクトルに基づいた動画像のシーン認識性能を検討した。今後は様々な符号化データを用い、どのように統合すれば高い認識率が得られるかを検討していく予定である。

参考文献

- [1] K. Simonyan, A. Zisserman, “Two-stream convolutional networks for action recognition in videos,” NIPS’14, pp. 568-576, 2014.
- [2] B. Zhang, L. Wang, Z. Wang, Y. Qiao, H. Wang, “Real-time action recognition with enhanced motion vector CNNs,” CVPR’16, pp.2718-2726, 2016.
- [3] K. Soomro, A.R. Zamir, M. Shah, “UCF101: A dataset of 101 human actions classes from videos in the wild,” CoRR, abs/1212.0402, 2012.