

動画圧縮符号内の動きベクトルの応用

高寺 達也[†] 久保田 光一[‡]中央大学大学院 理工学研究科 情報工学専攻^{†‡}

要約: 動画圧縮符号化方式の 1 つである H.264/MPEG-4 AVC を取り上げ、符号化される情報を整理し、その内の 1 つである動きベクトルの画像解析への応用を検討した。

キーワード: 動画圧縮符号, H.264/MPEG-4 AVC, 動きベクトル

1 背景

デジタル映像機器の普及に伴い、映像コンテンツ (以下ではすべてデジタルコンテンツを対象とする) の制作は容易になり、世界中にはすでに相当量の映像コンテンツが蓄積されていると思われる。また映像機器の進歩に伴い、映像コンテンツ自体も高精細になり大容量化している。

それらの映像コンテンツは通常、動画圧縮符号化技術を用い圧縮され、伝送・蓄積されるが、映像コンテンツの増大、高精細・大容量化に伴い、圧縮された動画データ処理・解析する手間も増す一方である。

2 目的

本研究は、圧縮された動画データの加工・解析を動画圧縮符号化技術に着目し高速に処理することを目的としている。

通常、動画データの加工・解析を行うためには、符号化された情報を復号してから解析を行う必要がある (加工の場合は再符号化も必要となる)。高度に圧縮された動画データほど、その復号や符号化に手間がかかるが、ある種の加工・解析では、符号化情報を利用することでその手間を省略できるものと考えられる。

本稿では、近年普及が進みつつある動画圧縮符号化方式 H.264/MPEG-4 AVC を取り上げ、その符号化される情報の整理を行うとともに、その中の 1 つである動きベクトルについて、画像解析への応用を検討する。

3 動画圧縮符号化方式 [2]

映像コンテンツは、一定時間間隔の静止画像 (ピクチャ) の連なりからなるが、単純にそれら 1 枚 1 枚の画像を保存し動画データとすると、容量面で非効率的である。そこで、何らかの動画圧縮符号化方式を用い、連続する画像を圧縮符号化して動画データとする方法が一般的にとられる。

近年の動画圧縮符号化方式は、予測、変換、可変長符号化という主に 3 つの技術を中心に構成されている。

予測とは、符号化対象ピクチャの前後のピクチャや、ピクチャ内の符号化対象領域の周辺領域などから符号化対象の画素値を予測し、その予測と実際の画素値との差分を符号化対象とする技術である。

また変換とは、画素値や予測によって生成された差分値に偏りがあることを利用し、それらの値をより圧縮効率を高められる表現形式へ変換する技術である。

そして可変長符号化とは、値の出現頻度に応じて長さの異なる符号を割り当てることで、固定長符号化に比べ、情報を符号

化した際の符号長の総計を減らす技術である。

動画圧縮符号化方式には、代表的なものに H.262/MPEG-2 があり、DVD への映像記録や BS/CS・地上デジタル放送に利用されている。

本研究では、次節で説明する H.264/MPEG-4 AVC を研究対象として採用した。

4 H.264/MPEG-4 AVC [1]

4.1 概要

H.264/MPEG-4 AVC は、ISO/IEC JTC-1 と ITU-T が合同チーム JVT (Joint Video Team) を結成し策定した動画圧縮符号化標準である。

Blu-ray Disc や HD-DVD の映像記録方式として採用され、またワンセグ放送にも利用されるなど、近年急速に普及が進んでいる。

圧縮符号化の基本的な考え方は、現在広く利用されている H.262/MPEG-2 と同様のものであり、また競合技術である VC-1 と類似点が多い。

最大の特徴は圧縮効率の高さであり、H.262/MPEG-2 と比較して同程度の画質で 2 倍程度の圧縮率があるが、符号化・復号処理にはより手間を要する。

4.2 主要技術

H.264/MPEG-4 AVC では、ピクチャをスライスと呼ばれる領域に分ける。スライスはマクロブロックと呼ばれる 16×16 画素の小領域からなり、このスライスとマクロブロックは符号化の際に 1 つの重要な単位となる。それらを踏まえて、H.264/MPEG-4 AVC における予測、変換、可変長符号化の主要技術をまとめる。

まず予測には、大きく 2 つ、画面内予測と動き補償予測が用いられる。画面内予測は、符号化対象マクロブロックに対し、その近傍の符号化済みマクロブロック (左上, 上, 右上, 左のマクロブロック) を用いて画素値を予測する手法である。一方、動き補償予測は、時間的に近傍の符号化済みピクチャの空間的な近傍領域を予測値とする手法である。

H.264/MPEG-4 AVC の動き補償予測では、マクロブロックを 4 分割したサブマクロブロックを含む可変ブロック・サイズでの予測ができる。さらに、 $1/4$ 画素単位での参照位置指定を行うことができ、参照ピクチャに対する時間的な制約も少なく、また重み付き予測なども可能である。

次に変換には、整数精度 DCT と離散アダマール変換 (DHT) が用いられる。マクロブロックを 4×4 または 8×8 画素単位に分けて、整数精度 DCT で変換し、直流成分にはさらにそれらのみを集めて DHT を施す。従来は実数 DCT が用いられていたが、H.264/MPEG-4 AVC では整数精度 DCT になり、デコーダの復号結果が一致するようになった。

最後に可変長符号化には、指数ゴロム符号 (Exp-Golomb codes), CAVLC (Context-based Adaptive Variable Length Coding), CABAC (Context-based Adaptive Binary Arithmetic Coding) が中心に用いられる。変換後の DCT 係数値などには CAVLC や CABAC が使われ、その他のシンタックス要素には指数ゴロム符号などが使われる。

Application of Motion Vector in Video Compression Code

[†] Tatsuya TAKADERA, Information and System Engineering Course, Graduate School of Science and Engineering, CHUO University

[‡] Koichi KUBOTA, Information and System Engineering Course, Graduate School of Science and Engineering, CHUO University

表1 NAL ユニットの種類

識別子	内容
1...5	各種ピクチャのスライス
6	SEI (Supplemental Enhancement Information)
7	SPS (Sequence Parameter Set)
8	PPS (Picture Parameter Set)

4.3 符号化情報

H.264/MPEG-4 AVC で符号化された情報は、NAL (Network Abstraction Layer) ユニットと呼ばれる単位でまとめられ、これは伝送・蓄積するための下位システムにマッピングする単位となる。

NAL ユニットは、ピクチャのスライス情報を持つ VCL (Video Coding Layer) NAL ユニットと、それ以外の符号化にかかわる情報を持つ非 VCL NAL ユニットとに大別できる。

NAL ユニットの種類には主要なものに表1にあげるものがあり、NAL ユニットのヘッダには種類を表す識別子が収められている。

このうち、SPS と PPS には、一連のピクチャをまとめたシーケンスレベル、あるいはピクチャレベルでの符号化にかかわる情報が含まれている。ゆえに、復号に際しても重要な役割を果たす。

一方、VCL NAL ユニットに相当する識別子1...5のNAL ユニットには、画素値や差分値を変換・符号化したものと、ブロックサイズや予測に用いられる各種パラメータが含まれている。この中には、次節で説明する動きベクトルも含まれる。

5 動きベクトルの応用

5.1 動きベクトル

動き補償予測では、マクロブロックを 16×16 から 4×4 までの可変サイズで分けて、各ブロックごとに相対的な参照位置を示す動きベクトルを符号化する。参照するピクチャはマクロブロックごとあるいはサブマクロブロックごとに符号化される。

H.264/MPEG-4 AVC における動き補償予測は $1/4$ 画素精度で行われる。動きベクトルの水平成分が取得可能な範囲は $[-2048, 2047.75]$ (単位は画素) であり、動きベクトルの垂直成分が取得可能な範囲は、デコーダの性能を規定する各レベルごとに限界が定められており、最大で $[-512, 511.75]$ (同上) である。

動きベクトルの符号化には、動きベクトルに対する予測も用いられる。隣接する符号化済みブロックの中央値を予測値として、求めた動きベクトルとの差分を符号化する。

また、予測値をそのまま動きベクトルとして扱うダイレクト・モードと呼ばれる予測方法もある。この場合、動きベクトルはいっさい符号化されない。ダイレクトモードには、時間ダイレクト・モードと空間ダイレクト・モードがあり、いずれの場合でも、表示順序で後方にある最も近い参照ピクチャをアンカー・ピクチャとして、符号化対象ブロックと同位置にあるアンカー・ピクチャのアンカー・ブロックの動きベクトルを利用する。

5.2 画像解析への応用

符号化時に動きベクトルを決定するときは、一般的に符号化対象マクロブロックの近傍から探索を行い決定する手法が用いられる。そのため、動きベクトルはおおよそ映像の動きを表し

ていると考えられる。この仮定の下に、動きベクトルの画像解析への応用を検討する。

特に、映像自身の動きが緩やかである場合、動きベクトルはオプティカルフローの代替として使える可能性がある。動画画像圧縮符号化に用いられる動きベクトル検出法は、オプティカルフローのブロックマッチングに近く、符号内の動きベクトル情報を代わりに利用できれば、復号・解析の手間を省くことができると考えられる。

また、固定カメラから取られた映像コンテンツの場合、動きベクトルを移動体検出に利用できると考えられる。カメラ側が映像をあらかじめ符号化して伝送する場合には、その復号・解析の手間を省くことができる。他にも、固定カメラで長時間録画された中から、映像の変化が大きい部分をピックアップすることにも応用できると考えられる。

5.3 問題点

動画画像圧縮符号内の動きベクトルを画像解析に応用するにはいくつか問題点がある。

まず、動きベクトルはすべてのピクチャに必ず存在するわけではない。動きベクトルは動き補償予測が用いられるマクロブロックにのみ存在する。H.264/MPEG-4 AVC では、ある程度の間隔ごとに IDR (Instantaneous Decoding Refresh) ピクチャがあり、IDR ピクチャは画面内符号化のみで構成されるため、ある程度の間隔ごとに動きベクトルがまったく検出されないピクチャが存在することになる。

また、H.264/MPEG-4 AVC では、動き補償予測を基本とするスライスに画面内予測を取り入れることができ、マクロブロックによっては動きベクトルが存在しないものもある。

さらに、根本的な問題になるが、動画画像圧縮符号化方式は符号に対する制約を定めているだけで、どのような手法で符号化するかは完全にエンコーダ依存であり、動きベクトルが目的に則したものであるという保証はない。

しかしながら、動きベクトルの応用が実用的に有効かどうかは、さまざまなソースで実験し、その相関を調べる必要がある。

6 実装

C++ 言語で H.264/MPEG-4 AVC デコーダを実装し、その復号過程において特定の情報のみを取得するプログラムを作成した。このプログラムを用いることで、例えば符号中の動きベクトルのみを解析することが可能になる。

7 まとめ

近年の実用的な動画画像圧縮符号化方式、特に H.264/MPEG-4 AVC について符号化情報をまとめ、動きベクトルの画像解析への応用を検討し、実験のためのプログラムを作成した。今後は、実際の動画画像データを数多く解析しどの程度の有用性があるのかを検証していく必要がある。

謝辞

本研究を進めるにあたり研究室の同輩、後輩にはお世話になった。ここに記して感謝の意を表す。

参考文献

- [1] ITU-T, "ITU-T Recommendation H.264 (03/2005) Advanced video coding for generic audiovisual services", ITU-T, 2005.
- [2] 大久保 榮 監修, 角野 真也, 菊池 義浩, 鈴木 輝彦 共編, "インプレス標準教科書シリーズ 改訂版 H.264/AVC 教科書", インプレス, 2006.