



2000年12月25日

スプライトを用いたMPEG-4超高圧縮映像符号化方式を開発 ～インターネットやモバイル環境でも 高画質かつなめらかな動きの映像を実現～

NTTは、光ソフトサービスにおける高品質な映像配信アプリケーションの構築に向けて、MPEG-4超高圧縮映像符号化方式(VideoESPER^{*1})を開発し、インターネットやモバイル環境でも高画質かつなめらかな動きの映像を実現しました。映像符号化では、従来、MC+DCT^{*2}をベースにした符号化方式(ノーマルモード)を用い、各動画像フレームから動きの特徴(動きベクトル)を求めて予測画像を生成し、その画像と実際の画像の差分を符号化する方式が多く利用されてきました。一方、複数の動画像フレームの背景から「スプライト」と呼ばれる1枚の静止画像を生成し、このスプライト画像を利用することで大きな圧縮効果を得られるスプライト符号化方式^{*3}も既にMPEG-4の映像符号化方式として標準化されていますが、スプライトの生成が手動や半自動であるなどの難点がありました。今回開発した方式では、簡易なビデオオブジェクトモデルの考案などにより、圧縮符号化の対象となる映像の特徴に基づいて、スプライトを自動生成し符号化するアルゴリズム(スプライトモード)を開発しました。また、どのような映像に対しても効果的な符号化を行うために2種類のMPEG-4符号化方式(ノーマルモードとスプライトモード)を自動的に切り替えるアルゴリズムも考案し、符号化の完全自動処理を実現しました(図1)。これにより、数百kbit/sというビットレートでも、従来方式比で2～3倍のフレームレートという動きの滑らかな映像作成が可能となり、MPEG-4の映像コンテンツ、およびそれを利用したサービスの実現が促進されます。特に、陸上競技のようなカメラの動きの激しい映像に対する圧縮効果が大きく、将来的には、中高速光IPネットワークやIMT-2000などを通じて高品質なスポーツ映像観戦が可能となります。

〈開発の背景〉

インターネット、そして携帯電話の爆発的な普及に伴い、映像配信サービスが増加しつつあります。このようなサービスは数百kbit/s というビットレ

ートで映像を伝送するために、映像信号を符号化してデータ量を圧縮する必要があります。MPEG-4はこのような用途に向けた符号化方式の国際標準で、これに準拠した製品も市場に登場してきています。そして、これらの製品に使われているのは、MC+DCTをベースにした符号化方式（ノーマルモード）です。

このMC+DCT符号化方式は、高速で動く被写体を追尾したり、ズームやパン、チルトなどのカメラの操作⁴がある映像に対しては、低いビットレートでは十分なフレームレートや画質が実現できないという問題点がありました。

〈技術のポイント〉

1. スプライトモードのポイント

カメラ操作に伴う背景の激しい動きがある映像に適用する「スプライトモード」は、背景部分をスプライト符号化、前景部分をオブジェクト符号化⁵というMPEG-4に備えられている符号化ツールを組み合わせることで符号化します。また、「スプライトモード」では自動的に前景・背景に切り分けて符号化します。この「スプライトモード」を適用するショットにおいて、画像を前景オブジェクトと背景スプライトに効率的に分離するために、「簡易ビデオオブジェクトモデル」を新たに考案しました。このモデルでは、カメラの動きを反映する領域を背景、それ以外の動物体すべてを1つの前景として、自動的に分離します。これにより、レイヤは常に背景と前景の2層に固定されるため、従来問題となっていたフレームにまたがる複数のビデオオブジェクトの対応問題を回避でき背景・前景の分離を良好に行うことが可能となりました。

「スプライトモード」が効果的な映像において、「スプライトモード」は「ノーマルモード」で符号化した場合に比べ、同程度の画質、同程度のフレームレートを1/2～1/4の符号量で達成することが可能です（[図2](#)）。

2. 符号化モード判定の自動化

切り替える2種類のモードは、MPEG-4シンプルプロファイル符号化方式による「ノーマルモード」とNTT新開発の「スプライトモード」です。前者は、従来から用いられてきた'MC+DCT符号化方式'を踏襲する符号化方式で、静止カメラで撮影した映像のような、背景の動きが少ない映像の符号量低減に適した方式です。それに対し後者は、前者が苦手としていた、カメラ操作による激しい動きを伴う映像に対して符号量の大幅な低減を可能とする方式です。映像の各ショットがどちらの符号化モードに適しているかを自動的に判断・切り替えて符号化しますので、どのような映像に対しても全体の符号量の低減が可能となります。

「スプライトモード」が効果的なショットは、カメラの動きがあること、前景がない、もしくはある程度小さな領域であること、そしてある程度の長さをもつショットであることです。これらの知見を用いて1)ショットにおけ

るフレームの数、2)カメラモーションの有無、3)前景率算出、の順番で符号化モードを決定して行きます。カメラモーションの算出と前景率の算出には、ビデオオブジェクトの自動生成の技術を利用しています。「ノーマルモード」となるショットの決定を早い段階で行い、効率的な判定を実現しています。

以上の技術でモードの判別から切り替えまでの処理を自動化できるため、これまでのように映像の性質によって符号化方式を使い分ける必要もなくなるので、編集後の映像に対してもモード判定や画像処理の中身を意識せずに符号化を行うことができます。また、従来方式に比べて、同一符号量でも高品質、高フレームレートな映像の提供が可能になります。

〈今後の展開〉

NTTでは今後、エンコーダ（符号化装置）の高速化や機能向上と同時に、音声と多重化し、ストリーミング配信を念頭においた開発を進め、NTTが展開する光ソフトサービスにおける映像配信サービスのキープロダクトとして展開を図って行く予定です。また、以下のURLにてデモンストレーション映像とデコーダをダウンロードすることができます。今回提供するデコーダは純粹にビットストリームをデコードして表示する機能のみを実現し、システムレイヤは組み込んでおりません。

<http://www.netviker.com/sprite>

○ 用語解説

*1 VideoESPER

Video Efficient SPrte-aided Encoder

*2 MC (Motion compensated coding) + DCT(discrete cosine transform)

本稿では、ノーマルモードと呼んでいる符号化方式です。MPEGの符号化アルゴリズムは、動き補償フレーム間予測符号化（MC）と離散コサイン変換（DCT）を結合したハイブリッド方式という大枠で共通です。MCは、画像フレーム間の時間軸方向の冗長度を取り除き、DCTは2次元画像平面上の空間方向の冗長度を取り除く技術です。したがって、前景・背景を分離するといった考え方はありません。画面上の各マクロブロック（16×16画素）の動きベクトルを一様に求め、そのフレームから予測画像を生成し、その予測画像と実際のフレームの画像との差分を符号化する方式です。

*3 スプライト符号化

MPEG-4符号化ツールの一つ。カメラ操作を反映した映像中の背景領域はパノラマ画像にすることが可能です。沢山のフレームから成る背景部分の動画像を一枚のパノラマ静止画にすることで、この領域を劇的に圧縮することが可能です。

*4 カメラ操作

ズーム、ロテート、パン、チルト等。ズームはカメラレンズの焦点距離を変化させる操作、ロテートはレンズ軸を中心にカメラを回転させる操作、パンはカメラを左右に振る操作、チルトはカメラを上下に振る操作をいいます。

*5 オブジェクト符号化

MPEG-4符号化ツールの一つ。従来符号化法は矩形の画像しか扱えませんでした。任意の形状の画像（ビデオオブジェクト）を符号化することが可能です。ビデオオブジェクトはテクスチャ情報と形状情報からなります。

別紙

- ・ [図1 MPEG-4超高圧縮映像符号化方式の概要](#)
- ・ [図2 スプライトモードの概要](#)

＜本件問い合わせ先＞

NTTサイバーコミュニケーション総合研究所
広報・情報戦略担当 坂本、萩野
TEL：0468-59-2032
e-mail：ckoho@tamail.rdc.ntt.co.jp



[NTT NEWS RELEASE](#)