

スプライトを適応的に利用した MPEG-4 映像符号化システム

日本電信電話株式会社 NTT サイバースペース研究所
 東日本電信電話株式会社 法人営業本部
 日本電信電話株式会社 NTT サイバースペース研究所
 早稲田大学国際情報通信研究センター

秦泉寺 久 美
 岡田 重 樹
 八島 由 幸
 正会員 渡 辺 裕

〈あらまし〉 著者らは、MPEG-4 Main Profile におけるスプライトという新しい符号化ツールに着目し、動画像におけるカメラ操作のみを反映した背景領域にスプライトを適用する方法を検討してきた。スプライトが効果的な場合は、従来符号化法に比べて劇的な符号量の削減をもたらすことを明らかにしてきた。一方で、スプライトがどのような画像にも効果があるとは限らないということも知られており、適応的な利用が望まれていた。本稿では、スプライトを適応的に符号化へとりこみ、MPEG-4 Main Profile 準拠の符号化システムとして開発した。本符号化システムは、スプライトが効果があると推定されるショットを自動的に抽出するスプライトショット抽出法を取り入れ、符号化品質の向上を図った。また、ストリーミング配信への適用も考慮し、低遅延スプライトに適用する2層ビデオオブジェクトレート制御法を新たに考案した。これにより、コンスタントビットレートの符号化データの生成を可能とした。また、カメラの動きのあるスポーツ映像を用いて、本提案符号化システムの評価を行い、その優位性と適応条件について考察した。

キーワード：MPEG-4, 超高圧縮, 映像符号化, スプライト, オブジェクト符号化, システム開発

〈Summary〉 MPEG-4, a new audio-visual coding standard has strong compression tools, such as sprite coding. We have focused on sprite coding and proposed "Sprite Mode". In sprite mode, video sequence is divided into foreground object and background object. Foreground object is compressed using MPEG-4 object coding, while background object is compressed using sprite coding. Sprite mode does not always adopt any video sequence such that images having still camera motion can be compressed in enough quality in low bit-rate. So, in this paper, we developed new MPEG-4 video encoder using sprite coding. In the proposed MPEG-4 encoder, multi-mode coding is newly proposed. Video shot suitable for conventional coding method is coded by normal mode (conventional method), while video shot suitable for sprite coding is coded by sprite mode. These two modes are automatic selected in any video sequences. We improved the processing flow using stuck system in order to consequently process video such as AVI format. We improved global motion detection algorithm, foreground object and background sprite generation algorithm. Then we proposed the simple rate-control for sprite mode. Then, we have some experiment on MPEG-4 video coding. The result shows effectiveness of proposed encoder in low bit-rate coding.

Key words: MPEG-4, Video compression, Sprite, Object coding, System development

1. はじめに

筆者らは、MPEG-4 Main Profile¹⁾における「スプライト」という新しい符号化ツールに着目し、動画像におけるカメラ操作^(注1)のみを反映した背景領域にスプライトを適用する方法を検討してきた。スプライトとは、パラメトリックな動きをする動画像の一連の領域を一枚の静止画像(パノラマ画像、モザイク画像)に射影したもので、元の領域はスプライトを幾何変換することで得られる^{2),3)}。著者らは、スプライトが効果的に働く場合は、従来符号化法に比べて劇的な符号量の削減をもたらすことを明らかにしてきた⁴⁾。同時に、スプライトがあらゆる動画像に効果があるとは限らないことも知られており、適応的な利用が望まれていた。スプライトの生成および符号化の研究の歴史は長く、1990年代以降、Irani²⁾やWang³⁾は、フレーム間の大域的な動き(グローバルモーション)の抽出とスプライト(モザイク画像)生成について検討を行っている。また、Lee⁵⁾やSmolic⁶⁾らはスプライトを符号化に適用する研究を行っている。しかし、これらの研究はあらかじめビデオオブジェクトとして抽出された背景スプライトを符号化方式に適用するのに留まり、汎用的な符号化システムにスプライトを導入するための検討はなされていなかった。

本稿では、動画像の圧縮率向上を目的とし、スプライトを適応的に利用することで、MPEG-4 Main Profile準拠の汎用的な高圧縮映像符号化システムを開発した。本符号化システムを構築するにあたって、以下の2点において新たな提案を行う。

一つは、スプライトが劇的な圧縮効果を生み出す場合もあれば、逆に符号化効率を悪化させる場合もあるという特性⁴⁾をかんがみ、動画像からスプライトが効果を発揮する動画像の一連のフレーム群(便宜上、これをスプライトショット^(注2)と呼ぶ)を自動的に抽出する方法を提案する。符号化ツールであるスプライトを適用するには、フレーム単位ではなく、ショット単位に適用性を判断する必要がある。符号化ツールの適応的選択の例は、古くは、マクロブロックをフレーム内予測とフレーム間予測を切り替えて符号化する方法や、ピクチャタイプ^(注3)を適応的に切り替える方法⁷⁾の提案がなされている。しかし、これらの方法は動き補償(MC: Motion Compensation)を基本ツールとし、基本的にフレーム毎においても判定が可能である。一方で、スプライトショットの抽出に関しては、著者ら⁸⁾が先んじて提案して

きたが、スプライトの効果のある要因の追求およびその組み合わせによる検討に留まり、実際に符号化システムに導入するためのアルゴリズムの検討はなされていなかった。本稿では、フレーム単位で適応性を判断する従来技術とは異なり、ショット単位にスプライトの適応性を判定し、汎用符号化システムへの導入を試みる。

二つ目は、2層ビデオオブジェクトレート制御法の提案である。MPEG-4 Main profileにおいては、複数のビデオオブジェクトを個別に符号化することが可能である。スプライトも一種のビデオオブジェクトである。複数の任意形状のビデオオブジェクトを同時にレート制御する方法として、Vetroら⁹⁾がマルチビデオオブジェクトレート制御法の提案を行っている。しかし、任意形状ビデオオブジェクトとは異なる性質を持つスプライトには適用できなかった。本稿では、ストリーミング配信への適用も考慮し、低遅延スプライトと任意形状ビデオオブジェクトからなる2層ビデオオブジェクトのレート制御法を提案する。

以降、2章では本論文における符号化システムの概要を述べ、3章および4章では、本システムを構成するにあたってスプライトショット抽出法および2層ビデオオブジェクトレート制御法をそれぞれ新規に提案する。5章では符号化実験による提案システムの評価を行い、6章で結論を述べる。

2. スプライトを適応的に用いた符号化システムの提案

本章では、スプライトを適応的に用いた映像符号化システムの提案を行う。以下に、スプライトショット符号化の概念、スプライトが効果的な条件、および、本提案符号化システムの概略について述べる。

2.1 スプライトショット符号化

スプライトを使った符号化をスプライト符号化という。著者らは文献4)にて、画像を前景オブジェクトと背景スプライトに自動的に分離し、背景スプライトにスプライト符号化を、前景オブジェクトに任意形状ビデオオブジェクト符号化を適用する、スプライトショット符号化^(注4)を提案している。図1にスプライトショット符号化の概念を示す。生成された二つのビデオオブジェクトのビットストリームは多重化されて送信される。デコーダ側では多重化されたビットストリームを多重分離し、各々のビットストリームを復号し、合成して表示する。一方、これに対して、スプライトを使用しない、MC+DCTを基本とする符号化をノーマルショット符

(注1) パン、チルト、ズームといったカメラの動き。

(注2) 一連のフレーム群をショットと言う。

(注3) I, P, Bピクチャなど。

(注4) 文献4)ではスプライトモードと呼んでいる。

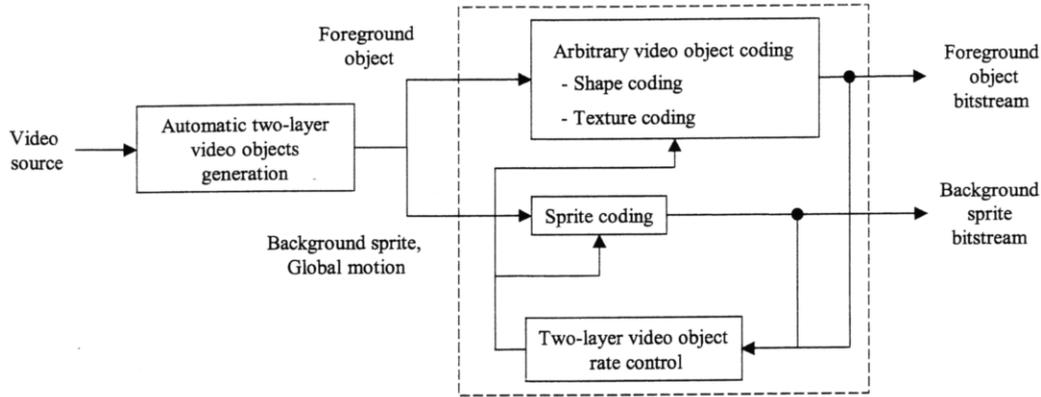


図 1 スプライトショットの概念
Fig. 1 The concept of sprite shot coding.

号化と呼ぶことにする。

スプライト符号化のビットストリーム(符号化データ)は、一枚の背景スプライトを静止画符号化したデータと、フレームごとのグローバルモーションデータからなる。また、オブジェクト符号化のビットストリームは、形状符号化データとテクスチャ符号化データからなる。スプライトは連続するフレームの一連の領域が一枚の静止画であらわされるので、その領域の占める割合が画面に対して大きいほど劇的な圧縮効果をもたらす。しかし、高い主観画質を従来符号化法に比べて小さいデータ量で表現できるが、元の領域のデータを忠実に再現するわけではないので、元データと復号データの PSNR は劣化するという特徴をもつ。

2.2 スプライトショット符号化に適したショットの条件

以下に、スプライトショット符号化が効果的に働く動画画像の条件として、以下の項目を挙げる^{8),4),10)}。

- (1) ショットチェンジを含まないこと
- (2) グローバルモーションが十分大きいこと
- (3) 前景領域が画面全体に対して十分小さい領域であること

グローバルモーションとは、画像の大局的な動きである。グローバルモーションがほとんどない場合は、前のフレームの同じ位置にあるマクロブロックの値をそのまま利用する“Not Coded”マクロブロックが多数出現する⁸⁾。よって、従来方式による符号化でも十分な画質を得ることが可能である。また、前景オブジェクトの画面全体に対する割合が 30%を超えると、スプライトショット符号化の効率が MC+DCT を基本とする従来符号化方式より劣化する⁴⁾。

2.3 ショット単位での処理の流れ

スプライトの適用性を判定するにはショット単位での処理が必須である。よって、画像データは T 秒単位の

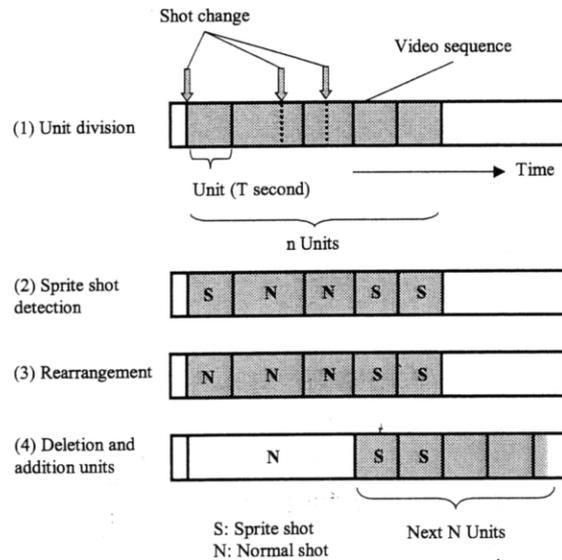


図 2 本提案符号化システムにおけるユニット処理の概念
Fig. 2 The concept of unit processing.

フレーム群(これをユニットと呼ぶ)ごとに処理される。本符号化システムは、ユニットごとにスプライトショットを逐次検出しながら、同時に符号化処理も行う。図 2 にユニット処理の概念を示す。

- (1) ユニット分割処理部において、 T 秒ごとのユニットに分割する。次の t 秒 ($t < T$) 以内にショットチェンジがある場合は、ユニットの長さを $t + T$ 秒に延長する。前ユニットがショットチェンジで終わっていて、かつ、 t ($t < T$) 秒以内にショットチェンジがある場合は、ショットチェンジをまたぐ T 秒を強制的にノーマルショットとする(図 2-(1))。
- (2) ユニット数が n 個になったところでいったん画像の読み込み処理を停止する(図 2-(2))。
- (3) n 個のユニットにおいて、スプライトショット抽出を行う。

- (4) 孤立するノーマルショットおよびスプライトショットを除去するために、判定結果に対し、ノーマルショットで囲まれた単独で存在するスプライトショットをノーマルショットへ変更する(図2-(3))。
- (5) 1個以上 $n-1$ 個以下のユニットの判定を確定し、スタックを空ける。空いたスタックにユニット数が n 個になるまで映像データを読み込む(図2-(4))。
- (6) 以上を繰り返す。

このように、スタックには常に n 個のユニットが格納され、1個以上 $n-1$ 個以下のユニットから成るショットの判定が随時行われる。判定後のユニットはスタックから削除され、削除された数のユニットが新たに挿入される。

2.4 提案符号化システムの概略

図3に、提案符号化システムのブロックダイアグラムを示す。スプライトショット抽出部と2層ビデオオブジェクトレート制御部が本符号化システムの主な新規な部分である。以下に提案符号化システムのブロックダイアグラムの説明を記す。

入力された画像に対して、まずショットチェンジを検出する。それと同時にユニット分割を行い、「2.2 スプライトショット符号化に適したショットの条件」で述べた三つの条件の一つである、(1)ショットチェンジの有無による一次判定を行う。

一次判定にてノーマルショットと判断されたユニット

は、以降の判定処理を省略し、ノーマルショット符号化する。スプライトショット候補と判定されたユニットに対して、局所的な動きベクトルを算出し(LME)、それを基に順方向のグローバルモーションを算出する(GME)。

ここで、(2)グローバルモーションの有無による二次判定を行い、ノーマルショットと判定されたユニットは以降の判定処理を省略し、ノーマルショット符号化する。スプライトショット候補と判定されたユニットに関しては、仮スプライト算出し、それを利用して前景オブジェクトを算出する。

ここで、(3)前景比率の大きさによる三次判定によってスプライトに適したユニットであるかを判定する。前景比率とは、画面全体の面積に対する前景オブジェクトの面積比率である。ノーマルショットと判定されたユニットは、ノーマルショット符号化される。ここでスプライトショットとして判定されたユニットは、前述のショット内処理で示したように、2, 3の連続するスプライトショットは一つのスプライトショットとして統合する(Unit integration)。また、孤立したスプライトショットはノーマルショットとして再判定され、ノーマルショット符号化をされる。

複数のユニットを統合したスプライトショットは、スプライト生成に必要な逆方向グローバルモーションの算出(Inverse GME)と背景スプライトの生成を行い、スプライトショット符号化を行う。

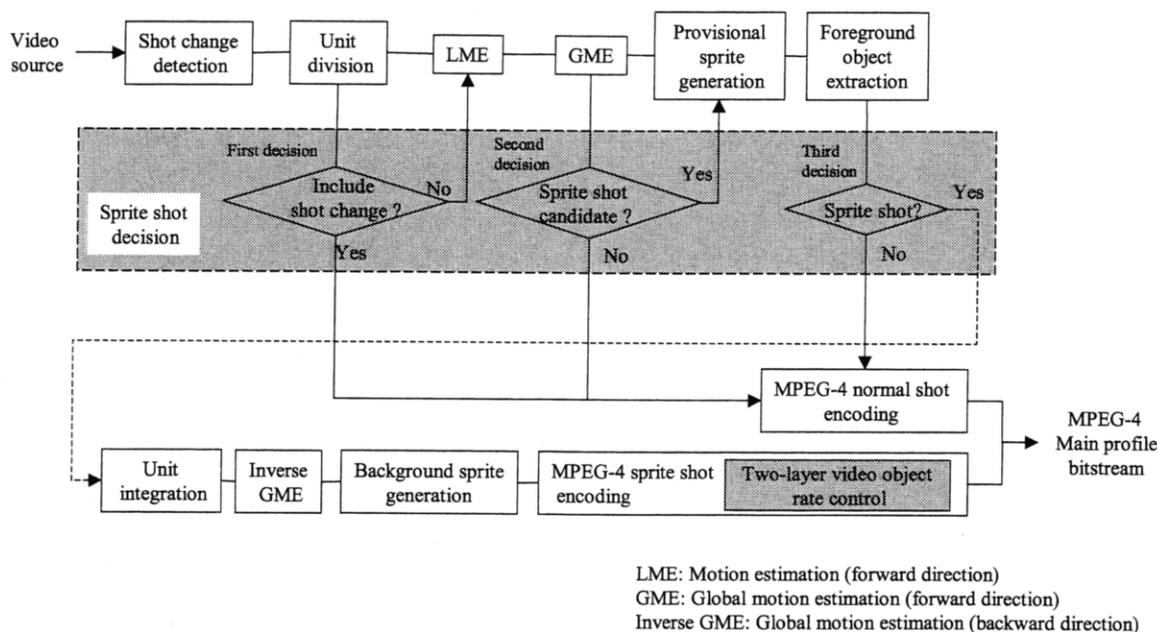


図3 本提案符号化システムのブロックダイアグラム
 Fig. 3 The block diagram of proposed encoding system.

3. スプライトショット抽出アルゴリズム

本章では、本符号化システムの、スプライトショット抽出アルゴリズムの詳細を述べる。図3からスプライトショット抽出処理を列挙すると、以下の手順になる。

- (1) ショットチェンジの抽出
- (2) ショットチェンジの有無による一次判定
- (3) 順方向グローバルモーション算出
- (4) グローバルモーションの大きさによる二次判定
- (5) 仮スプライト生成
- (6) 前景オブジェクト抽出
- (7) 前景比率による三次判定

本提案符号化システムにおけるグローバルモーション算出、前景オブジェクト抽出法、仮スプライトおよび背景スプライト抽出は文献4)11)の方法を踏襲する。

以下に、グローバルモーションの大きさによる二次判定(手順4)および、前景比率による三次判定(手順7)の詳細を示す。

3.1 グローバルモーションによる二次判定

本稿では、パン、チルトの組み合わせから二次判定を行う。ユニット内の先頭フレームを基準フレームとして、手順3により順方向にグローバルモーションを算出する。ここで、基準フレームの座標系を基本座標系とする。各フレームにおける各座標はグローバルモーションによって基準座標系の座標にマッピングすることができる。ここで、各フレームの中心座標(0, 0)を基準座標系へのマッピングする。マッピング後の座標は、各フレームの x 軸および y 軸方向への変移、すなわち、パンとチルトの値に相当する。このパン、チルトの基準座標系における時間的な変移をトラジェクトリと呼ぶ。ここで、ユニット内のトラジェクトリの大きさを算出し、この大きさを用いて二次判定を行う。判定方法の詳細を以下に述べる。

- (1) ショット内の各フレームの中心座標において、グローバルモーションによるトラジェクトリを算出する。
- (2) 基本座標系における、トラジェクトリの x 軸および y 軸方向の最大値、最小値 SX_{max} , SY_{max} , SX_{min} , SY_{min} をそれぞれ算出し、その差分絶対値 SX , SY を算出する。

$$SX = abs(SX_{max} - SX_{min}) \quad (1)$$

$$SY = abs(SY_{max} - SY_{min}) \quad (2)$$

- (3) $SX \geq Th_{tri}$ もしくは、 $SY \geq Th_{tri}$ を満たす場合はスプライトショット候補、満たさない場合はノーマルショットと判定する。ここで、ユニット内のフレーム数を $frames$ として、 $Th_{tri} =$

$Th_{tri} = frames * 2.0$ とした。これは、フレーム当たりの x 方向および y 方向の動きが平均して2画素以下の場合をグローバルモーションが十分に小さいとして、ノーマルショットと判断することを意味する。

3.2 前景比率による三次モード判定

本小節では、前景比率による三次モード判定を提案する。前景オブジェクトの抽出は文献4)10)を踏襲する。文献4)にて、画面に対する前景の面積(前景比率)が10-15%程度のところでスプライトが効果的に作用し、30%を超えると、逆にスプライトの効率が通常の符号化に比べて悪くなることが示されている。そこで、前景比率によって、スプライトが適用できるか否かを判定する三次判定を行う。ユニット内の前景比率の平均値が十分小さくても、ショット内には前景の面積が急激に増えている場合がある。これは、例えば、カメラの前を突然人が横切るなどの場合に起きる。この現象を回避するために、前景比率が極端に大きくなっているフレーム数を算出し、これが全フレーム数のある割合以上を占めるショットをノーマルショットと判定する。以下に、三次判定の流れを記す。

最初に、ユニットごとに前景オブジェクトおよび前景比率を算出する。前景の大きさを fg_{size} (単位:画素)、フレームの大きさを $frame_{size}$ (単位:画素) とすると、前景比率 FG_{ratio} は以下の式で表される。

$$FG_{ratio} = fg_{size} / frame_{size} \quad (3)$$

更に、ユニット内の平均前景率を $AVEfg_{ratio}$ ユニット内のフレーム数を $frames$ 、スプライトショットとして許容できる前景比率の最大値のしきい値を Th_{max} 、前景比率が極端に大きくなっているフレームの割合のしきい値を Th_{large} として、以下のように判定する。

- (1) ユニット内の平均前景率 $AVEfg_{ratio}$ を算出。
- (2) ユニット内の各々のフレームの前景比率 FG_{ratio} が Th_{max} を超えるフレーム数(count)を算出する。
- (3) 以下の判定法で判定する。“S”はスプライトショット、“N”はノーマルショットである。

if $AVEfg_{ratio} \geq Th_{ratio\ mode} = N$;

else if $(count \geq frames * Th_{large}) \ mode = N$;

else $mode = S$;

後述の評価実験では、 $Th_{max} = 0.2$ (20%), $Th_{large} = 0.2$ (20%), $Th_{ratio} = 0.17$ (17%) を用いた。

4. 2層ビデオオブジェクトレート制御法

スプライトショットは、前景オブジェクトと背景スプライトの2層のビデオオブジェクトを符号化する。よっ

て, Verification Model Ver. 17.0¹²⁾等のレート制御法をそのまま適用することはできない。また, Vetroら⁹⁾が提案したマルチビデオオブジェクトレート制御法はスプライトには適用できない。また, スプライトをストリーミング配信するためには, 低遅延スプライト符号化を行う必要がある。よって, 本稿では, 低遅延スプライトに対応した2層ビデオオブジェクトレート制御法¹³⁾を提案する。

前景オブジェクトと背景オブジェクトの符号量比率に基づいて符号量を割り当て, 低遅延スプライトを用いて符号化を行う。具体的には以下の手順で処理を行う。

- (1) スプライトショットにおいて, 量子化値 $QP=15$ (QP : Quantization Parameter)にて前景オブジェクトおよび背景スプライトを符号化し, その配分を前景背景比とする。
- (2) ショット全体のトータルビット T [bit]と前景背景比から, 背景スプライトに割り当てられる符号量 T_{b0} [bit]を算出する。
- (3) 背景スプライトのQP値を1から31まで変化させてQP固定(1フレームに一つのQP値)で符号化し, 前項で求めた背景スプライトに配分される符号量 T_{b0} [bit]に最も近くなるQP値 QP_b およびその符号量 T_b [bit]を求める。
- (4) 残りのビット数を計算し, 前景オブジェクトの総ビット数 $T_f = T - T_b$ [bit]とする。
- (5) T_f [bit]を秒数で除した値を目標ビットレートとして, 前景オブジェクトをVM 17.0¹²⁾で符号化する。これにより, 前景のフレームレート f_f [fps]が算出される。
- (6) スプライトを前景のフレームレート f_f [fps]に合わせて, QP_b (固定)にて低遅延モードでスプライト符号化する。

5. 提案符号化システムの評価

提案符号化システムを表1の画像を用いて, PSNR, 達成フレームレート, 主観画像によって評価した。

5.1 評価画像

画像はカメラ操作を含む10-30秒程度のスポーツ映像(競馬, スケートボード, サッカー, テニス)である。以下に画像の具体的内容を示す。

5.1.1 horserace(a), (b)

“horserace(a)”は競馬のスタートから手前の直線コース, 第一コーナーを回るまで, “horserace(b)”はその続きの第二コーナーから奥の直線コース, 第三コーナーに至るの競走馬をフォローショットしたものである。

表1 実験画像

Table 1 Video Sequences.

Image	Frames	size	contents
horserace(a)	999	SIF	pan, zoom
horserace(b)	999	SIF	pan, zoom
skateboard	390	SIF	pan, tilt
stefan	300	SIF	pan, zoom

5.1.2 skateboard

カメラがほぼ静止した状態で, 画面右端からスケートボードが登場, 画面の左に向けて移動する様子をカメラが追従したものである。最後にスケートボードが向き逆方向に変えて画面から消える。カメラはほぼ静止する。

5.1.3 stefan

標準画像である。左右にボールを拾うテニスプレーヤをフォローショットし, 最後は画面の右端から左にかけて移動する様子をカメラが高速に追従している。

5.2 符号化実験

前述の五つの画像で符号化実験を行った。比較対象は以下の二つである。

- MC+DCTを基本とする従来符号化法(MPEG-4 Main profile 準拠 VM 17.0¹²⁾)
- スプライトを用いた本提案符号化法(MPEG-4 Main profile 準拠)

本提案符号化システムにおけるノーマルショットはBピクチャを含まない。ユニット分割処理におけるショットの区切りは $T=1$ 秒, またユニット数 $n=5$ とした。また, 目標ビットレートおよび目標フレームレートを以下に示す。

- 128 kbps, 15 fps
- 384 kbps, 30 fps

表2にスプライトショット抽出結果の例(“skateboard”)を示す。“second decision”はグローバルモーションの大きさによる判定, “Third decision”は前景比率による判定である。フォローショットの部分のスプライトショットとして抽出されている。また, 表3にスプライトショット符号化の選択率と平均前景率を示す。“stefan”は標準画像として知られるが, その平均前景比率は約10%, すべてのショットにてスプライトショット符号化が選択されている。“horserace(a)”と“horserace(b)”に着目すると, “horserace(a)”の方が, 静止している時間が長く(スタート時), またカメラに近いところを競走馬が走ることから, 前景比率が高く, “horserace(b)”に比較しスプライトショットになる比率が低い。また, スプライトショットにおける平均

表 2 スプライトショット抽出結果の一例“skateboard”
Table 2 Example of sprite shot extraction “skateboard”

Start frame	End frame	Second decision	Third decision	FG ratio	Final decision	Final shot
1	30	N	N		N	N
31	60	N	N		N	
61	90	N	N		N	
91	120	S	N	0.214	N	
121	150	S	S	0.092	S	S
151	180	S	S	0.101	S	
181	210	S	S	0.120	S	
211	240	S	S	0.063	S	S
241	270	S	S	0.097	S	
271	300	S	S	0.106	S	
301	330	S	S	0.093	S	N
331	360	N	N		N	
361	390	N	N		N	

表 3 平均前景率, スプライトモード選択率
Table 3 Average foreground ratio and sprite mode decision ratio

Image	Sprite mode ratio	Foreground ratio
horserace(a)	0.478	0.140
horserace(b)	0.775	0.083
skateboard	0.538	0.096
stefan	1.000	0.094

前景比率も高い結果となっている。以上、表 2, 3 から、本スプライトショット抽出が効果的に作用していることが分かる。また、図 4 に “skateboard” におけるフレームにおける平均 QP 値の推移を示す。表 2 から、従来符号化法にて QP 値が 30 前後に張り付いている場合でも、提案手法では、スプライトショット符号化が採用され、前景・背景とも 10~15 程度の QP 値が与えられており、より細かい量子化が実現されている。すなわち、従来手法では、スケートボードをフォローショットする部分では十分な符号量が与えられず、量子化値が大きくなったのに対し、提案手法ではこの部分にスプライトショット符号化が用いられることによって、より高い画質(細かな量子化)を実現した。また、従来方式では、ショットの最初、中盤、後半では、あきらかに中盤に対して粗い量子化が行われていたのに対し、本提案手法では、安定した画質(量子化値)を実現していることがわかる。

図 5 にスプライトショットの合成画像の例を示す。表

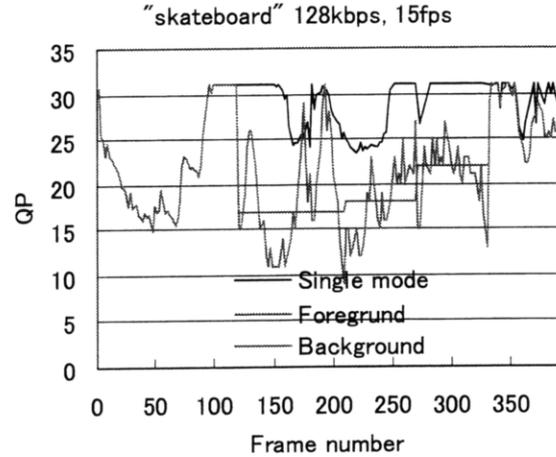


図 4 フレームにおける平均 QP の変化
Fig. 4 Average QP by frame.

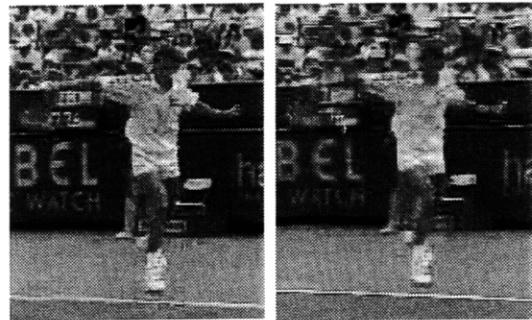


図 5 符号化画像の例。左、提案方式(本提案符号化システム)、右、従来方式(VM 17.0)
Fig. 5 Examples of encoded images. Left: proposed method, right: conventional method (VM 17.0).

4 に符号化効率, 平均 QP (Quantization Parameter: 量子化パラメータ), 平均 PSNR (信号対雑音比), 達成フレームレートの比較の例を示す。“Mode” における “Proposed” は提案符号化法, “Conventional” は従来符号化法である。表 4 において, 128 kbps, 15 fps の符号化条件の達成フレームレートに着目すると, “horserace(a)”, “horserace(b)”, “skateboard”, “stefan” では, 従来符号化では各々 8.33 fps, 7.39 fps, 12.85 fps, 8.00 fps のフレームレートであるのに対し, 提案符号化法では, 12.88 fps 14.89 fps, 14.92 fps, 15.00 fps の高フレームレートが達成されている。この符号化条件では, 平均 QP 値に関しても, 提案符号化法が従来符号化法より小さい値が割り当てられている。“horserace(b)”, “skateboard” の二つの画像では, 提案符号化法におけるノーマルショットのフレームレートと従来符号化法のフレームレートはほぼ同程度であることから, 従来符号化法がフレーム落ちを起こす部

表4 符号化結果
Table 4 Coding results

Image	Mode	Object type	128 kbps, 15 fps			384 kbps, 30 fps		
			Ave. QP	Ave. SNR [dB]	frame rate [fps]	Ave. QP	Ave. SNR [dB]	frame rate [fps]
horserace (a)	Proposed	FG	19.99	25.35	15.06	10.51	29.45	30.00
		BG	20.42	19.76	15.06	10.31	19.82	30.00
		Norm.	30.10	24.80	10.88	23.77	25.90	30.00
		Total	—	23.24	12.88	—	24.21	30.00
	Conventional	—	30.37	24.97	8.83	25.12	20.20	30.00
horserace (b)	Proposed	FG	14.41	28.06	15.00	6.91	33.43	29.65
		BG	14.27	22.74	15.00	7.07	22.99	29.65
		Norm.	17.36	29.31	14.53	10.90	32.32	30.00
		Total	—	25.39	14.89	—	26.50	29.73
	Conventional	—	27.00	26.94	7.39	14.12	29.40	30.00
skateboard	Proposed	FG	18.97	24.29	15.00	9.36	30.07	27.57
		BG	18.71	20.29	15.00	9.62	20.25	27.57
		Norm.	24.07	14.46	14.83	10.95	26.83	30.00
		Total	—	23.16	14.92	—	24.46	28.69
	Conventional	—	26.40	24.46	12.85	18.15	26.38	30.00
stefan	Proposed	FG	17.07	6.44	15.00	8.34	31.48	30.00
		BG	17.00	18.19	15.00	8.30	18.23	30.00
		Norm.	—	—	—	—	—	—
		Total	—	19.38	15.00	—	19.52	30.00
	Conventional	—	30.84	24.15	8.90	23.67	25.37	30.00

分が提案符号化法ではスプライトショット符号化され、目標フレームレートに近いフレームレートを実現していることが分かる。

384 kbps, 30 fps の符号化条件では、従来符号化法と提案符号化法の両者ともおよそ 30 fps のフレームレートを達成している。しかし、平均 QP 値に着目すると、圧倒的に、提案符号化法が小さい値が割り当てられており、より細かい量子化が行われている。“horserace (b)”, “skateboard” の二つの画像では、提案符号化法におけるノーマルショットの平均 QP と従来符号化法の平均 QP の値がほぼ同じであることから、従来符号化法が十分な品質を与えられない部分が提案符号化法ではスプライトショット符号化され、平均的に高い品質を実現していることが分かる。

一方、提案符号化法におけるトータルの PSNR は従来符号化法の PSNR よりも低い値を示している。これは、背景スプライトが各々のフレームを貼り合わせて生成されたものであるため、そこから切り出す背景画像は

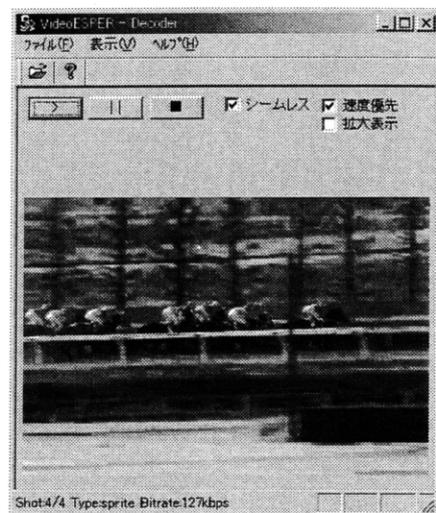


図6 提案符号化システム専用デコーダ
Fig. 6 Original decoder for proposed encoder system.

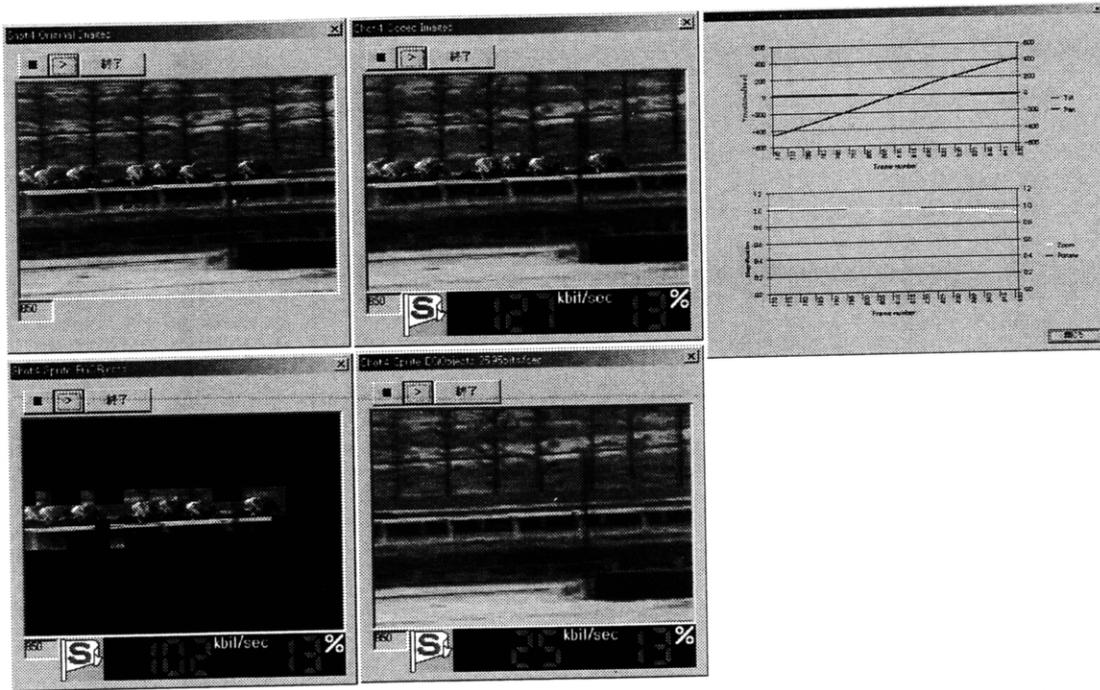


図 7 スプライトビューワによる生成オブジェクトおよびグローバルモーションの例。上段左より、原画像、符号化画像、グローバルモーション。下段左より、前景オブジェクト、スプライトから切り出された背景画像

Fig. 7 Sprite Viewer. From upper left: Original image, coded image, global motion. From lower left: foreground object image, background image extracted from sprite.

元の信号を忠実に再現していないというスプライト符号化特有の現象である(本稿第 1 章を参照)。しかし、前景オブジェクト (FG) に着目すれば、トータルの PSNR よりも 1~2 dB 高い PSNR を達成している。前景オブジェクトに割り当てられる QP 値が従来符号化法よりも小さい値が割り与えられおり、より細かい量子化が行われている。図 5 より、同じビットレートにて、提案符号化の方が従来符号化法よりも高い主観画質を実現していることがわかる。

以上のように、提案符号化法は、従来符号化法が不得意とするフォローショット画像において、高いフレームレートを達成し、細かい量子化による主観品質の向上をもたらす。よって、安定した品質を提供することが可能である。

図 6 に専用デコーダによってデコード表示されている様子を示す。また、図 7 にスプライトビューワによる原画像、本提案符号化システムによる符号化画像、前景画像、背景画像、グローバルモーションを表示した様子を示す。

本提案符号化システムは 1.5 GHz Pentium IV プロセッサを有する汎用的な PC にて、実時間比にて 10 倍程度の処理時間でエンコードを行うことが可能であった。また、400 MHz Pentium III 以上、256 Mbyte 以上のメモリを有する汎用的な PC にて実時間デコードが可

能である。

6. おわりに

著者らは、スプライト符号化を用いた汎用的な MPEG-4 動画像符号化システムを開発した。スプライトが効果的なショットを自動的に抽出する、スプライトイショット抽出アルゴリズムを提案した。また、前景オブジェクトと背景スプライトに自動的に符号量を配分するための 2 層ビデオオブジェクトレート制御法を提案した。また、実際に符号化実験を行って、低レートにおける従来符号化方式に対する本提案符号化システムの優位性を示した。

謝辞 本研究を遂行するに当たって有意義なディスカッションをいただいた NTT サイバースペース研究所画像メディア通信プロジェクト映像符号化技術グループの皆様にご感謝いたします。

参考文献

- 1) Information technology-Coding of audio-visual objects-Part 2: Visual Amendment 1: Visual extensions ISO/IEC 14496-2
- 2) M. Irani, P. Anandan, J. Bergen, R. Kumar, and S. Hsu: "Efficient Representations of Video Sequences and Their Applications", *Signal Processing: Image Communication*, Vol. 8 pp. 327-351 (1996).
- 3) J. Y. A. Wang, and Edward H. Adelson: "Representing

Moving Images with Layers”, IEEE Trans. on Circuits and Systems for Video Technology, Vol. 3, No. 5, September (1994).

- 4) 秦泉寺久美, 渡辺 裕, 岡田重樹, 小林直樹: “スプライト符号化を用いた MPEG-4 動画像超高圧縮”, 電子情報通信学会和文論文誌 D-II, Vol. J 84-D-II No. 5 pp. 758-768 May (2001).
- 5) M.-C. Lee, W. Chen, C.-I. Bruce Lin, C. Gu, T. Markoc, S. I. Zabinsky, and R. Szeliski: “A Layerd Video Object Coding System Using Sprite and Affine Motion Model”, IEEE Trans. on Circuits and Systems for Video Technology, Vol. 7, No. 1, pp. 130-145 February (1997).
- 6) Aljoscha Smolic, T. Sikora, J.-R. Ohm: “Long-Term Global Motion Estimation and Its Application for Sprite Coding, Content Description, and Segmentation”, IEEE Trans. on Circuit And Systems for Video Technology, Vol. 9, No. 8, December (1999).
- 7) 米山暁夫, 柳原広昌, 笹野義二, 中島康之: “画像特徴に応じた MPEG-4 ビデオ符号化方式”, 情報処理学会第 37 回 AVIM 研究会, pp. 9-14 July (2002).
- 8) 岡田重樹, 秦泉寺久美, 渡辺 裕, 小林直樹: “スプライト符号化を利用した MPEG-4 マルチモード符号化方式の研究”, PCSJ 2000, pp. 91-92 (2000-11).
- 9) A. Vetro, H. Sun, Y. Wang: “MPEG-4 Rate Control for Multiple Video Objects”, IEEE Trans. on Circuits And Systems for Video Technology, Vol. 9, No. 1, pp. 186-199 February (1999).
- 10) K. Jinzenji, S. Okada, N. Kobayashi, H. Watanabe: “MPEG-4 Very Low Bitrate Video Coding by Means of Sprite Coding”, ICME (2001).
- 11) 秦泉寺久美, 渡辺 裕, 小林直樹: “スプライト生成のためのグローバルモーション算出法と符号化への適用”, 電子情報通信学会和文論文誌 D-II, Vol. J 83-D-II, No. 2, pp. 535-544 (2000).
- 12) “MPEG-4 Video Verification Model version 17.0”, ISO/IEC JTC 1/SC 29/WG 11/N 3515
- 13) K. Jinzenji, S. Okada, N. Kobayashi, H. Watanabe: “MPEG-4 Very Low Bitrate Video Coding by Adaptively Utilizing Sprite”, ICME (2002).

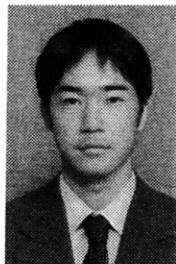
(2003年10月17日受付)

秦泉寺 久美



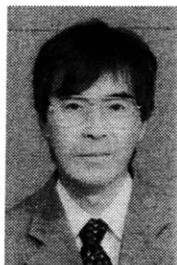
平1, 上智大・理工・電気電子工卒。平3早大大学院修士課程修了。同年NTTに入社。平13年2月～平14年3月まで、早稲田大学国際情報通信研究センター客員研究員。高機能高エネルギー通信のための画像処理、符号化の研究に従事。現在、日本電信電話(株)NTTサイバースペース研究所画像メディア通信プロジェクト研究主任。平10学術奨励賞受賞。平13 The 6th SCI “One of the Best Papers”。電子情報通信学会、情報処理学会、映像情報メディア学会会員。

岡田 重樹



平9, 慶應大・理工・計測工卒。平11, 同大学大学院計算機科学修士課程了。同年NTTに入社。高エネルギー通信のための画像符号化技術の研究開発に従事。現在、東日本電信電話株式会社法人営業本部ブロードバンドビジネス部勤務。

八島 由幸



昭56, 名大・工・電子工学卒。昭58同大学院工学研究科電子工学専攻修士課程了。同年日本電信電話公社(現NTT)入社。以来、主として画像信号の高エネルギー符号化、HDTV信号符号化、動画像信号処理、MPEG関連システムの研究開発に従事。現在、NTTサイバースペース研究所画像メディア通信プロジェクト映像符号化技術グループリーダー、主幹研究員。工博、IEEE、情報処理学会、映像情報メディア学会各会員。

渡辺 裕 (正会員)



昭56, 北大・工・電子卒。昭60, 同大学院博士課程修了。同年NTT入社。NTT研究所において画像の高エネルギー符号化に関する研究および動画像通信システムの開発に従事。平12, 早稲田大学国際情報通信研究センター教授。平14, 早稲田大学大学院国際情報通信研究科教授。マルチメディアの情報生成および分配システムに関する研究に興味をもつ。平11年よりISO/IEC JTC 1/SC 29議長。工博、IEEE、電子情報通信学会、情報処理学会、映像情報メディア学会、日本音響学会各会員。