

ハイパービデオを用いたビデオ・データベースの 効率的なブラウジング手法*

3K-8

牛尾 剛聡† 広部 一弥† 渡邊 豊英†

名古屋大学大学院工学研究科情報工学専攻†

1 はじめに

ビデオは時系列データであり、利用者が内容を観測するためには記録に必要とした時間と同程度の時間が必要である。したがって、大量のビデオを扱うビデオ・ライブラリ、対話型テレビ、ビデオ・オン・デマンド等の応用のために、ビデオ・データベース上でのビデオの効率的なブラウジング機構の開発が期待されている [1]。

現在、我々はビデオ・データベース・システム STRIKE (Stream data Retrieval system based on Indexing with Key Events) を開発中であり、本稿では、STRIKE 上でビデオの内容に基づいた効率的なブラウジングを実現する手法について述べる。本手法では、データベース内に保持されたビデオを内容に基づいて自動的にハイパービデオ化することにより、利用者はビデオ中の興味のあるシーンを文脈に応じて非線形的にアクセス可能である。

2 ビデオ・データのブラウジング

ブラウジングとは、利用者がデータベース内に蓄えられたデータの内容を把握する操作である。これまでに提案されてきたブラウジング手法の多くは、ビデオの概要を提示するアプローチを採用している。代表的な概要の形態としては、ショット遷移グラフ [1] や、r-frame と呼ばれるショットのパノラマ画像 [2] を提示するものがあり、ビデオの時間的な広がりや空間的な広がりとして表現することで内容の一覧性を高めている。これらの概要は、多種類のショットが存在するドラマやニュース番組等には有効であるが、スポーツ中継のように類似するショットが繰返し出現する場合には、ビデオの内容把握が困難である。また、時間的な広がりや空間的な広がりとして表現するため、ビデオという表現形式が本質的に有する直観的なわかり易さという特長が失われる場合がある。

我々は、上記の問題点を解決するために、データベース内のビデオをハイパービデオ形式でブラウジング可能とする手法を採用する。ハイパービデオによるブラウジングでは、シーン間に設定されたリンクにより、誘導的に意味内容に基づいたビデオの非線形的なアクセスが可能である。この手法では、利用者に対する直観的なわか

り易さを失わずに、簡単な操作でビデオの文脈に応じて興味のあるシーンをアクセス可能である。

3 ハイパービデオ・モデル

我々は、ビデオの内容に基づいて多種多様なシーンを検索可能なデータモデルであるイベント・アクション・モデルを提案している [3]。本モデルでは、ビデオ中に記録された継続時間を持たない出来事をイベントとして表現する。例えば、野球中継においては「打者がボールを打つ」、「野手が補球する」等のイベントを考えることができる。ビデオの内容はイベントの系列として表現される(これをコンテキストと呼ぶ)。このとき、ビデオに表現された継続した動作は、コンテキストの部分系列として表現可能であり、アクションと呼ばれるイベントの系列パターンによって指定可能である。本モデルでは、コンテキストを構成するそれぞれのイベントをフレームに対応付け、アクションを指定することでシーンを検索可能である。

本節では、内容に基づいたハイパービデオを自動生成するために、イベント・アクション・モデルに基づくハイパービデオ・モデルを導入する。

3.1 プレゼンテーション・コンテキスト

ビデオ・データベースでは2種類のビデオを考える必要がある。一方はデータベース内に蓄えられているビデオであり、他方は利用者が試聴するビデオである。前者をソース・ビデオと呼び、後者をプレゼンテーション・ビデオと呼ぶ。プレゼンテーション・ビデオはソース・ビデオを編集することにより作成される。なお、簡単化のために、本稿では以下の条件を満足するプレゼンテーション・ビデオを対象とする。

1. プレゼンテーション・ビデオは同一のソース・ビデオに含まれるシーンの並びである。
2. プレゼンテーション・ビデオを構成するシーンは順方向にのみ再生される。

ソース・ビデオの内容をイベント・アクション・モデルに基づいて表現した際のコンテキストをソース・コンテキストと呼ぶ。ソース・コンテキストは2項 (E, \prec) として表現される。ここで、 E はイベントの集合であり、 \prec はイベント間の全順序関係である。順序関係 \prec は、対象世界におけるイベントの発生順序を表す。

ソース・ビデオと同様に、プレゼンテーション・ビデオの内容をイベント系列として表現することができる。これをプレゼンテーション・コンテキストと呼ぶ。プレ

*An Efficient Browsing Method with Hyper-Video on Video Databases

†Taketoshi USHIAMA, Kazuya HIROBE and Toyohide WATANABE

‡Department of Information Engineering, Graduate School of Engineering, Nagoya University

ゼンテーション・コンテキストにおけるイベントの並びは利用者がプレゼンテーション・ビデオを介して観測するイベントの順序である。プレゼンテーション・ビデオにおけるシーンの境界では、ソース・ビデオ中では観測されないイベントが発生する場合がある。これをカット・イベントと呼ぶ。図1にカット・イベントの例を示す。

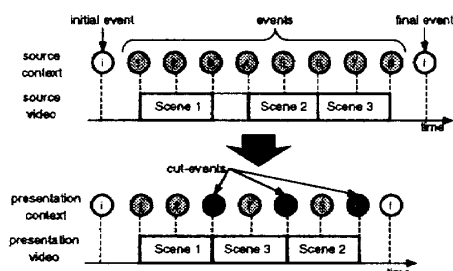


図1: プレゼンテーション・ビデオとカット・イベント

カット・イベントは、連続する2つのシーンにおいて、先行シーンの最終イベントと連続シーンの終了イベントの2項関連として表現可能である。このとき、プレゼンテーション・コンテキストは2項(sc, CE)として表現できる。ここで、 sc はソース・コンテキストであり、 CE はソース・コンテキストを構成するイベントの2項関係(e_i, e_j) (カット・イベント)の集合である。ソース・コンテキスト上でカット・イベントをイベント間のリンクとして表現した例を図2に示す。図中では(e_3, e_6), (e_8, e_4), (e_6, e_7)という3個のカット・イベントが表現されている。

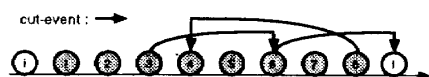


図2: ソース・コンテキスト上のカット・イベント

3.2 プレゼンテーション・スキーマ

同一のソース・ビデオに対して複数のプレゼンテーション・ビデオを考えることができる。同一のソース・ビデオから生成可能な複数のプレゼンテーション・ビデオを表現するために、プレゼンテーション・スキーマを導入する。プレゼンテーション・スキーマはソース・コンテキスト上にカット・イベントを指定することにより表現する。あるプレゼンテーション・ビデオにおけるカット・イベントが他のプレゼンテーション・ビデオにおいて発生しない場合がある。そこで、カット・イベントを選択カット・イベントと単純カット・イベントに分類する。単純カット・イベントは全てのプレゼンテーション・ビデオにおいて発生し、選択カット・イベントは選択的に発生する。

プレゼンテーション・スキーマは5項($sc, CE_{sim}, CE_{sel}, L, \phi$)として表現できる。ここで、 sc はソース・コンテキストであり、 CE_{sim} は単純カット・イベントを表すイベント間の2項関係であり、 CE_{sel} は選択カット・イベントを表すイベント間の2項

関係であり、 L はラベルの集合であり、 ϕ はマッピング $\phi: L \rightarrow CE_{sel}$ である。図3にプレゼンテーション・スキーマの例を示す。プレゼンテーション・スキーマ上のカット・イベントをリンクとみなすことにより、ソース・ビデオをハイパービデオ化可能である。

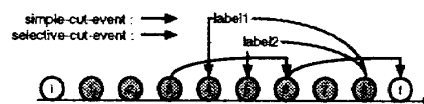


図3: プレゼンテーション・スキーマの例

3.3 プレゼンテーション・ポリシー

プレゼンテーション・スキーマは、個々のソース・コンテキストに対して指定されるため、大量のビデオ・データを扱うビデオ・データベースでは、個々のプレゼンテーション・スキーマの人手による指定は非現実的である。そこで、カット・イベントをプレゼンテーション・ポリシーと呼ばれる一般的な形式で表現する。プレゼンテーション・ポリシーは(a_1, a_2, l)として表現する。ここで、 a_1 と a_2 はアクションであり、 l はラベルである。プレゼンテーション・ポリシーは、 a_1 に対応するシーン s_1 と a_2 に対応するシーン s_2 がプレゼンテーション・ビデオにおいて連続することを表す。すなわち、 a_1 にマッチするソース・コンテキスト上の部分系列の最終イベントと、 a_2 にマッチする部分系列の先頭イベントの2項関係をラベル l を持つカット・イベントとすることで、プレゼンテーション・スキーマを生成可能である。なお、複数のプレゼンテーション・ポリシーが指定されているとき、次の場合には選択的カット・イベントとなる。1) カット・イベントの第一要素が他のシーンに含まれる場合、2) 同一のイベントを第一要素とする他のカット・イベントが存在する場合。

4 おわりに

本ハイパービデオ・モデルでは、それぞれのシーンを最後まで観測しなければ、次のシーンへ移動できない。しかし、利用者はシーンの途中から他のシーンへ移動したい場合があるため、シーン中の任意のフレームから、関係するシーンへ移動可能とするように本モデルを拡張する必要がある。

本稿では、ビデオをハイパービデオ化するための基本的なモデルを示したが、具体的なモデル化の方針について意味的な視点から考察していない。今後、効率的なブラウジングを実現するために、設定すべきカット・イベントの意味的な基準を明確化しなければならない。

参考文献

- [1] B.L. Reo and M.M. Yeung: "Retrieving and Visualizing Video", *Comm. of ACM*, Vol. 40, No. 12, pp. 43-52 (1997).
- [2] M. Flickner, et al.: "Query by Image and Video Content: The QBIC System", *IEEE Computer*, Vol. 28, No. 9, pp. 23-32 (1995).
- [3] 牛尾剛聡, 広部一弥, 渡邊豊英: "イベント・アクション・モデルに基づく動画検索システム STRIKE", *情処研報*, 98-DBS-115, pp. 79-86 (1998).