

動画データモデルの提案

7C-6

池谷勇一・飯田卓郎・石丸知之**・中森眞理雄*・植村俊亮***
 *東京農工大学 ** (株)東芝 ***奈良先端科学技術大学院大学
 工学部情報工学大講座 青梅工場 情報科学研究科

1. はじめに

動画データをデータベースで扱う場合は、従来の文字や数値のデータとは異なり、時間と空間の概念を必要とする。ここでは動画の一つとしてビデオ映像を取り上げ、ビデオ映像を検索、分析、編集するためのデータモデルについて提案する。

2. データ構造

ビデオ映像のデータ構造^[1]は、物理構造と論理構造に分けられる。

2.1 物理構造

物理構造は、メディア上におけるビデオ映像の位置を表すための構造である。デバイスに依存する内容を分離し、ビデオ映像を物理的に管理、操作するためのデータ構造である。この構造は使用するメディアに依存し、ビデオ映像の内容には依存しない。今回は、ビデオデッキ PV-S98 (NEC 社製) を使用した (図1)。PV-S98 は、ビデオテープ識別子 Header と1秒単位の位置情報 Address をビデオテープに直接録画することができ、1/30 秒単位の位置情報 Frame を使ってビデオ操作が可能となっている。したがって、図1の物理構造を計算機側で管理することにより、ビデオ映像の管理および操作を行うことができる。

物理構造の外部インターフェイスは、時間を基本単位とする。

2.2 論理構造

論理構造は、ビデオ映像の意味を表現する構造である。ビデオ映像の内容に注目すると、ビデオ映像はシーン、およびシーン間の関連によって表現できる。次にシーンの定義および各種の関連について述べる。

(1) シーン

シーンには、基本シーンと複合シーンがある。基本

シーンは、ある一つの対象・事柄などに注目して取り出した一区切りの映像のことである。複合シーンは、基本シーンの集まりまたは、基本シーンと複合シーンの集まりである (図2)。このように、シーンを階層構造にすることによって、内容をより自然に表現できるようにした^[2]。

基本シーンは、オブジェクト識別子、複数の属性、複数の条件付属性そしてビデオ映像識別子からなる。

基本シーン = {オブジェクト識別子; 属性; 条件付属性; ビデオ映像識別子;}
 複合シーン = {オブジェクト識別子; 属性; 条件付属性; 構成シーン;}
 構成シーン = {}

オブジェクト識別子は、シーンを一意に識別するための情報である。属性は、属性名と属性値の組の集合であり、シーンの詳細を表現する。条件付属性は、時間によって変化する属性を表現する。つまり、その属性がどの時間内で有効であることを示す条件式と、その条件式にあった場合の属性名および属性値を用いて、条件付属性が定義される。ビデオ映像識別子は、シーンの格納場所情報である物理構造を指している。つまり、一区切りの Frame 区間の情報である。構成シーンは、複合シーンを構成する各々のシーンのオブジェクト識別子からなる。

(2) 時間関連

シーンとシーンの間には時間関連が存在する。時間関連とは時間の前後関係を示す。ビデオ映像における

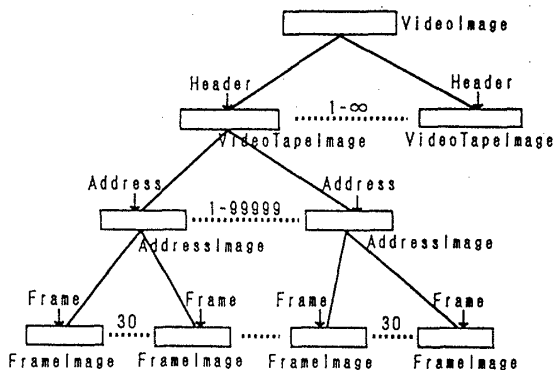


図1. ビデオ映像の物理構造

A Proposal for Modelling Video Images
 Yuuichi IKEYA*, TIIDA*, TISHIMARU**, M.NAKAMORI*, S.UEMURA***
 *Tokyo University of Agriculture and Technology
 **Ome Works, TOSHIBA Corporation
 ***Nara Institute of Science and Technology

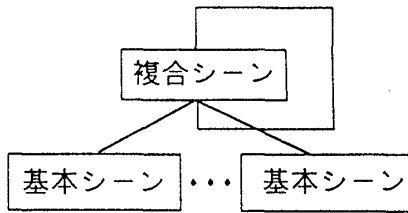


図2. 論理構造

時間には物理時間と論理時間の二つの種類があるため、時間関連においても物理時間関連と論理時間関連が存在する。ここで述べた物理時間とはシーンの再生順序に沿った時間であり、論理時間とはシーン中の世界における時間である。

例えば、ビデオテープのはじめから順番にシーンA、B、C、Dが録画されていて、シーンCは回想シーンであり、シーンAより前に起こった事柄とする。その場合、物理時間は A→B→C→D の順であり、論理時間は C→A→B→D の順となる。このようにビデオ映像のシーン間には、物理時間関連と論理時間関連が存在する。

今二つのシーンA、Bを仮定すると、時間関連は次の四つに分類できる。

- (1) AとBが重ならない
- (2) AとBが一部重なる
- (3) AはBと同じ
- (4) AがBに含まれる

この四つの関連は論理時間関連においても物理時間関連においても存在する関連である。このように、シーンは時間軸上において順番を表現できるべきである。

(3) 画面位置

画面位置は、物体が画面上のどの位置に存在するかを示す。この画面位置は複数の画面上の点で囲まれた範囲で表現する。

物体と物体の二次元的重なりは、画面位置の重なりを調べることによって分かる。

(4) 位置関連

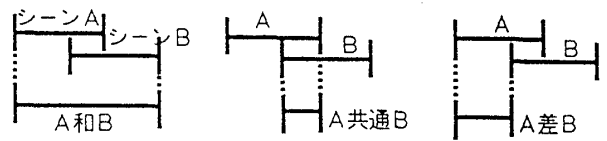
ビデオ映像において、同一フレームに登場する物体と物体の間には位置関連が存在する。ビデオ映像は三次元の世界を描き出している場合が多いので、位置関連も三次元を考慮して表現する必要がある。したがって位置関連を、左右関係、上下関係、前後関係さら

に物体同士の包含関係の四つの関係を組み合わせて表現する。

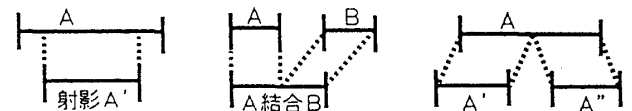
3. データ操作

ビデオ映像はシーンの集まりである。このシーンに対するデータ操作は次のようなものが考えられる。これはビデオ映像の編集のとき必要になる。

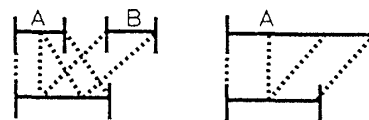
- (1) 和
- (2) 共通
- (3) 差



- (4) 射影
- (5) 結合
- (6) 分離



- (8) 挿入
- (9) 削除



以上のような操作をシーン操作として用意すべきである。他にも便利な操作が考えられるが、これらを組み合わせれば足りる。ここで重要なことは、和、共通、差、射影などの操作をした結果のシーンが、はたして何の属性を受け継ぐかということである。

4. おわりに

ビデオ映像の物理、論理構造およびシーン操作について述べた。特にシーン間に必要と思われる相互の関連を整理した。他にも、シーンの動作関連やビデオ映像に特有な時間、空間の概念を特に表現できるよう検討していきたい。実装の手法も今後の課題である。

5. 参考文献

[1]池谷, 飯田, ビデオ映像の構造について, 電子情報通信学会, 第4回データベースワークショップ, 1993
 [2]Junichi Takahashi, Temporal Query Processing for Scene Retrieval in Motion Image Databases, Third INTERNATIONAL SYMPOSIUM ON DATABASE SYSTEMS FOR ADVANCED APPLICATIONS, 1993