

関数型データモデルに基づく 映像データベースの構成とその実現

富井尚志 由井仁 石川英彦 有澤博

横浜国立大学 工学部 電子情報工学科

画像処理技術やネットワーク技術の発展により、映像データベースの普及が現実味をおびてきた。その基礎となる理論として著者らのグループでは、映像の中に映し込まれている主題とフレーム列の関係などを自然に表すことのできる関数型データモデルに基づいて、映像をモデル化した。そして、映像の検索や、検索結果に強調処理を施すこと、並列処理などを見越した高速で最適な検索操作が可能であることなどの特徴を持つデータベース操作系を導入した。また、これらの妥当性を検証するために、リレーショナルデータベースを利用した映像データベースのプロトタイプを開発した。

Design and Prototyping of Image Sequence Database based on Functional Data Model

Takashi Tomii, Hitoshi Yui, Hidehiko Ishikawa, Hiroshi Arisawa

Division of Electrical and Computer Engineering,
Faculty of Engineering,
Yokohama National University

In various fields of engineering, the technology of Image Sequence (or "movie shower") Database is growing to be quite significant and practical. This paper presents a new functional datamodel and its functional operators as the theoretical bases of the ISD. The operator set is based on Backus's FP, but introducing new atomic functions, it enables us to realize to highlight a specific object in a logical unit of image sequence.

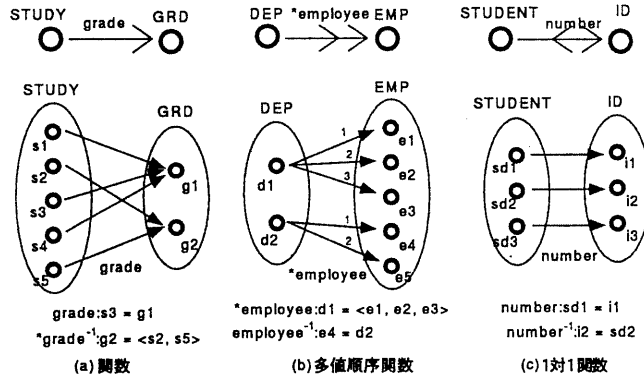


図 1: AIS モデルの関数

1 はじめに

プロダクション・エンジニアリング、建設物設計、医療情報処理など、多くのコンピュータ応用分野で、図形や映像などいわゆるマルチメディア情報を取り扱う要望が高まっている [1]。特に近年のワークステーションの CPU の高性能化や、画像処理技術、ネットワーク技術の発展により、映像情報を蓄積、検索、および配送する映像データベースの普及がにわかに現実味をおびてきた。

主として文字・数値などの二次情報を扱うための従来のデータベースに対して、フルテキスト、音声、図形、動画、画像、映像などの一次情報を統合して扱うデータベースをマルチメディアデータベースと呼ぶ。しかし、マルチメディアデータベースは、一次情報、二次情報など、異なる意味レベル間でのデータの扱いや、異なるメディア間での同一物の認識などの点で、データモデル、データ操作系の根本的な性格が従来型のデータベースとまったく異なる [2]。

近年、映像メディアを計算機上で扱うことはめづらしいことではなくなってきたが、これらのほとんどは映像に映し込まれている「もの」のレベルまで踏み込んでおらず、映像をデータベース化しているとはいえない。映像データベースは、利用者（エンドユーザ）側からの期待や要望が強いわりには、それを扱うデータモデルやデータ操作系などの基礎となる理論が確立されていない。

そこで著者らのグループでは、多様なマルチメディア情報を単純な要素 (primitives) の組合せだけで直観的かつ自然に表現することができる意味データモデルとして AIS モデル [3][4] を導入し、マルチメディアデータベースのカーネルとする研究を進めている。

このデータモデルは、関数型データモデルを基礎としており、現実世界の事象はデータベース上で主体 (entity) と主体間の「つながり」である関数によって表現し、イメージ情報とそれに伴う意味情報を統合化してモデル化している。さらに操作系として、FP を基とした関数型操作系を導入し、この操作系のもとで、映像データベースにとって必要な多種多様な検索操作を記述することを可能にする。

本稿では、AIS モデルと、操作系の枠組を導入し、その枠組の中で、映像を論理カットと呼ばれる単位に分割して蓄積する映像データベースの設計を行なう [5][6]。そして、導入した操作系によって、強調処理などを含めたデータベース操作を実現する手法を提案する。さらに、以上の考えのもとに作成した、試作映像データベースシステムによって、導入した映像データベースや操作系の妥当性と、映像データベースシステムの有効性を検証した。

2 データモデル

映像を、連続する画像のフレーム群であるとしてとらえたとき、個々のフレームデータは膨大であり、それらが長大な列をなしていると考えられる。また映

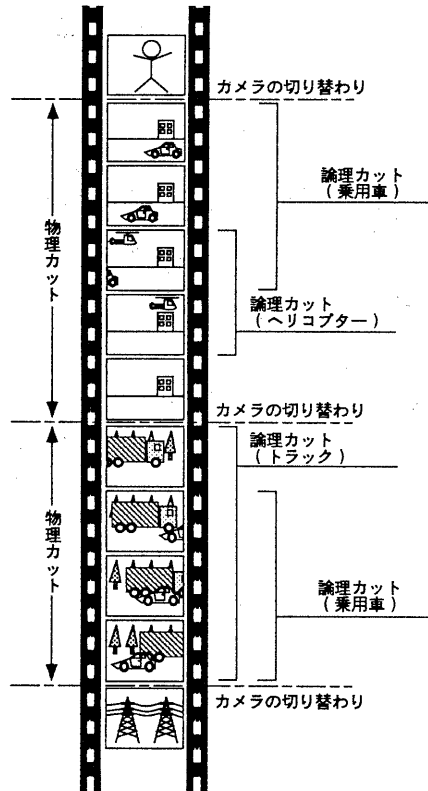


図 2: 論理カットと物理カット

像には、被写体の時間的・空間的位置などの情報を含むが、それらは明示的には示されていない。そして、そのような映像情報をデータベース化し検索する際に、被写体の強調などの画像処理特有のオペレータを埋め込んで検索したい。従来のデータモデルでは以上のような問題に対して、映像データを取り扱うことは困難である。

本稿では、それらの問題を解決するために新しいデータモデルの導入が必要であると考え、そのようなマルチメディア情報を表現するために関数型データモデルを基礎とする AIS モデルを導入する。

このデータモデルでは、現実世界の事象を、それらを表す主体と、主体同士を結ぶ関数による 2 項関係で表す。また、関数によってデータベースを操作する操作系として 2 項関係とマッチングのよい関数型操作系を導入する。以下の方法では、関数型操作

系によって多様な操作が可能となる。

ある共通の性質 (学生であるなど) を持つ主体の集合を主体型と呼ぶ。関数は、主体間の「つながり」を表し、主体型から主体型への写像によって表される。写像の元となる主体型を関数の定義域、写像の像となる主体型を関数の値域と呼ぶ。特に関数の返り値が順序を持った集合 (列) となるものを多値順序関数と呼び、図 1 に示すように関数名に * をつけることで、返り値が単値 (atomic) である関数と区別する。関数の逆関数は、多値順序関数である。これらの特殊な場合として、関数の対応が 1 対 1 となるものも存在する。

主体には 2 種類あり、現実世界の事物・事象を表す、内部構造を持たない点としてデータベース中で扱われるものと、値を示す一例えば数値や、文字列など一主体に分類される。後者を特に値 (value) と

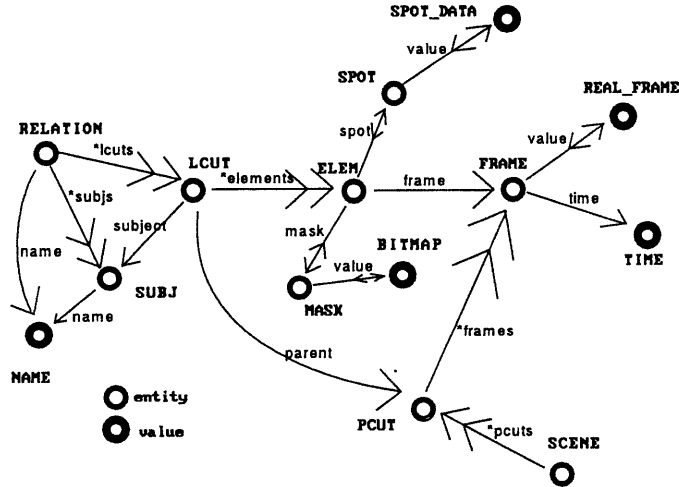


図 3: 映像データベーススキーマ

呼ぶ。動画像の 1 フレームを表すビットマップも 1 つの値である。

3 映像のモデリング

前章で述べた AIS モデルに基づき、映像データのモデリングを行なう。本節では、AIS モデルの特質や映像データの持つ特徴、そして操作系やインターフェイスなど、映像データベースシステムに対する要求を考慮したスキーマの設計を行なう。

3.1 映像データベースの構成要素

映像をモデリングするのに際し、映像を単位化する必要がある。映像の単位としては、放送などの分野で用いられている、カメラやフィルムの変り目を利用した「カット」が一般的である。しかし、ユーザがカットを示す際にもっともよく用いるのは「ある物体が映っている場面」等、被写体などをもとにした映像単位である [7]。そこで本稿では、この被写体が映っている区間を特に論理カットと呼び、前者のカットと区別する。前者をカメラの切り替わりなどによって、自動的に検出される物理的なカット—物理カット—と呼ぶ。

ユーザが必要とするであろう情報は、多くは論

理カットに含まれる。したがって、映像をモデリングするにあたって、この論理カットを中心に考えることとする。ある論理カットには、映し込まれている主題や、主題同士の関連などの情報も付加される。さらに、論理カットの各フレームに、被写体の位置情報があれば、「被写体にスポットライトをあてる」などの強調処理を行なうことができるが、この強調処理のためのパラメタも論理カットの付随情報として付加される。また、物理的な変り目をうまく利用するため、物理カットをまたぐ論理カットは存在しないものとする。物理カットと論理カットの例を図 2 に示す。

3.2 映像データベーススキーマ

前節で述べた映像データベースの構成要素に、意味情報などを付加して AIS モデルによってスキーマの定義をする。

現時点では、映像という一次情報の持つ「意味」を自動的に抽出することは困難であるから、「人間が被写体の位置を指し示す」などの半自動的手段によって抽出 [8] された情報のみを検索の対象とする。これらの情報を表すためには、以下のような記述が考えられる:

あるフレーム (FRAME) のどの部分が被写体なのか、どの位置に被写体があるのか (SPOT)、論理

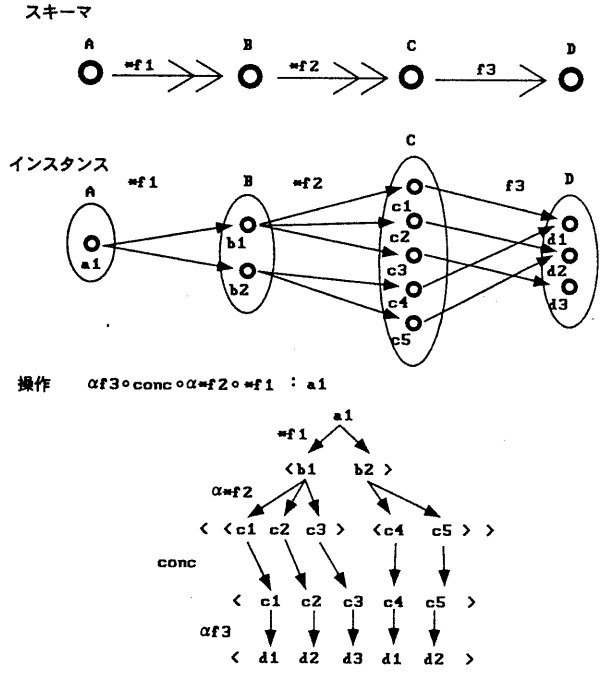


図 4: 操作例

カット (LCUT) はどのような順序でコマ (ELEM) が並んだものなのか、また、何を被写体としているのか (SUBJ)、被写体間に何らかの関係 (加工や衝突など) があるのか (RELATION)、関係や被写体の名称 (NAME)、フレームはどのくらいの時間間隔でサンプル (例えば 1/30 秒) されたか (TIME)、など。

これらを AIS モデルの要素を用いて表現したスキーマを図 3 に示す。

4 関数型操作系

AIS モデルの単純なプリミティブは、関数型操作系としての FP[9] と良くマッチする。AIS モデルによって表されたデータベースの関数は、1つの主体に適用すると、0,1 ないし複数の主体を返すが、FP の関数は1つの対象 (Object) に適用すると対象または対象の列 (Sequence) を返す。この類似性は、FP の操作系の上でデータベースの情報を扱うことを可能としている [10]。本稿では、この FP を基盤として、関数型操作言語を導入する。

4.1 関数形式と原始関数

データベース操作の際、例えば「ある被写体が写っているフレームすべてにある操作をせよ」というような、並列処理抽出の容易な操作が多くなる。本稿では、データベース操作がそのような並列処理・最適化処理を意識した操作系によって実現するように、再帰的記述を排除した宣言的な操作系を提案する。

したがって、FP に、データベース操作系としてのある拡張を行なう。宣言的データベース操作のために、以下のような関数形式、原始関数を定義する。

- 関数形式 Generate ! [10]
 指定された主体型に属するすべての主体を列にして返す。主体型 A に、 n 個の主体 a_1, a_2, \dots, a_n が属するとき、
 $!A : x \equiv A$ が主体型ではない $\rightarrow x$;
 $\langle a_i | a_i \text{ は主体型 } A \text{ に属する主体} \rangle$

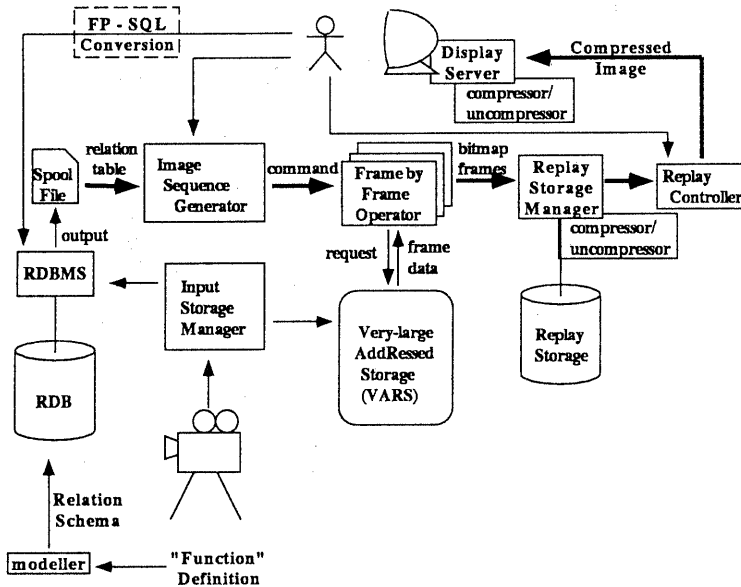


図 5: 試作映像データベースシステム

- 関数形式 Restrict [10]
Apply to All と同様に、すべての対象に関数を適用して、真となる対象だけを返す。

$$\begin{aligned}
 |f : x &\equiv x = \phi \rightarrow \phi; \\
 &x = \langle x_1, x_2, \dots, x_n \rangle \rightarrow \\
 &\langle x_i | f : x_i = T \rangle; \perp
 \end{aligned}$$

関数型操作系による操作例を図 4 に示す。

4.2 関数型操作系による映像データベースの操作

- 原始関数 Group by
列中の各要素の、先頭要素について、distr の逆演算を行なう。

$$\begin{aligned}
 grpby : x &\equiv x = \langle \langle a, y \rangle \dots \langle a, z \rangle, \dots \rangle \\
 &\rightarrow \langle \langle a, \langle y, z \rangle \rangle, \dots \rangle; \perp
 \end{aligned}$$

関数型操作系による映像データベースの操作は、操作関数と、操作対象の 2 つを作成することによって可能である。前者は、検索操作の宣言であり、後者はそのパラメタである。例えば、被写体の名前からフレームを生成する関数 f を定義したとき、対象である被写体の名前 A に関数 f を適用することによって、データベース操作が行なわれる。

- 原始関数 Concatenate
リストに含まれるリスト群のブラケット $\langle \rangle$ をはずして、それらの内容をつなげる。

$$\begin{aligned}
 conc : x &\equiv x = \langle \langle y_1 \rangle, \langle y_2 \rangle, \dots, \langle y_n \rangle \rangle \\
 &\rightarrow \langle y_1, y_2, \dots, y_n \rangle; \perp
 \end{aligned}$$

ところで、関数型操作系は、作成した関数に新しい名前をつけること (関数定義) によって、複雑な操作を一つの関数として取り扱うことが可能である。データベースに対する複雑な操作は、自由度の高い関数型操作言語によるプログラミングであると考えてもよいが、頻繁に使われる操作については、あらかじめ定義されている (名前を付けられている) のが望ましい。そこで本節では、データベース操作に関

して「便利な」関数をいくつか定義し、それらを組み立てることによってデータベース操作を実現することにする。

ところで、マルチメディア情報を扱う際には、「画像処理を行なう」などの、扱うメディアに依存した処理を行なう必要がある。

例えば強調効果などの操作を行なう場合には、画像処理モジュールが必要となる。さらにメディア依存処理の性質は、検索処理の性質とは(並列性・処理量などの点で)全く異なる。このような多数のフレーム列に対して、画像処理演算を(並列性を意識して)可能とする要素を導入することが望まれる。

そこで、このような画像メディアに依存した画像処理モジュールをFFO(Frame by Frame Operator)と呼ぶ。この関数は、画像メディア(動画画像メディア)の場合には、画像処理の要求列から実際に画像処理済みのデータを返す。このような関数は、別の研究テーマである音声メディアにも専用のものが必要である。これらのメディア依存処理は、専用のモジュール(DSPや画像処理チップなど)を用いてシステム構成すれば、高速な処理を実現することが可能である。

FFOには、「#1のフレームに、#5なるパラメタを用いてフレーム合成せよ。」というような指示を与える必要があるが、これをフレームスクリプトと呼ぶ。FFOは関数型操作系からは一つの関数のようにみえるように定義する。

FFOの持つ機能は、必要最低限の画像処理であり、処理のために多量の計算機資源を必要としないように設計すれば、FFO実行の際のフレーム単位の並列化によって、極めて高速に操作することが可能となる。ユーザーは検索操作として、1.フレームスクリプトを生成して、2.FFOに入力することにより、画像処理済みの映像を得ることができる。

以下のように、よく用いられる操作関数を定義する。

- 主題の名前から、その主題を示す論理カットの列を返す関数

```
Def nm_lc ≡ α1 ◦
  |(eq ◦ [name ◦ subject ◦ 1, 2])
  ◦ distr ◦ [!LCUT, id]
```

- あるコマ(ELEM)の主体と、強調加工の種類からなる列からフレームスクリプトを生成する関数

```
Def gen_fs ≡ eq ◦ [2, "Mask"] →
  α[frame, mask] ◦ 1;
  eq ◦ [2, "Spot"] →
  α[frame, mask] ◦ 1;
  α[frame, φ] ◦ 1
```

- 論理カットの主体と、強調加工の種類からなる列からフレームスクリプトを生成する関数

```
Def ap_op ≡ gen_fs ◦ [*elements ◦ 1, 2]
```

- 主題の名前と、その主題を示す論理カットへの強調加工の種類からなる列からフレームスクリプトを生成する関数

```
Def f_script ≡ αap_op ◦ distr
  ◦ [nm_lc ◦ 1, 2]
```

上記の関数 f_script を用いれば、主題名とそれに対する強調操作によって構成される列群に対して以下の操作によってその映像を得ることができる。

```
αFFO ◦ grpby ◦ conc ◦ αf_script
```

すなわち、主題名と強調加工の種類からなる列それぞれに対して、フレームスクリプトを生成し、同じフレームによって構成されるフレームスクリプトは $grpby$ によってひとまとめにされる。この操作によって検索された列は、同じフレームが重複しないフレームスクリプト列である。したがって、それらのフレームスクリプトそれぞれに対して、FFOによって強調処理画像が生成されれば、検索結果の動画映像を得ることができる。

5 試作映像データベースシステム

前節までに述べた考えを利用して、映像データベースシステムを試作した。

この試作映像データベースシステムのシステム構成図を図5に示す。

我々は、本稿で述べた映像のモデリング手法に基づいて、リレーショナルモデルによって映像を表現し、SQLによって操作できる映像データベースシステムを実現した。

ユーザーが発行した、FPによって記述された問い合わせを、SQLに変換し、SQLによってリレーショナルデータベースを検索する。

これによって得られる表は、検索結果として得られる動画像を生成する生成プログラムへのパラメータ群であり、ファイルSpool Fileへと出力される。システム構成要素Image Sequence Generatorは、Spool Fileを解釈してFrame by Frame Operatorによって、画像列を生成する。生成された画像列は、Replay Controllerによって動画像として出力される。

このシステムによって、映像をその持つ意味によって検索し、検索結果に強調処理などの加工をして出力する映像データベースシステムを実現した。その結果、映像データベースシステムの有効性や、映像のモデリング手法と操作システムの妥当性が検証された。

6 まとめ

本稿では、映像データベースに求められる要求から、従来とは異なるデータモデルであるAISモデルを導入して、それによって映像データベースを構築した。また、AISモデルとマッチングのよいFPから、FPをもとにした関数型操作系を導入し、それによって映像データベースを操作することが可能であることを示した。

さらに、それらの考えをもとにした映像データベースシステムを試作し、映像データベースシステムの有効性や操作システムの妥当性を検証した。

参考文献

[1] 坂内正夫, 佐藤幸男, 全炳東, 井上誠喜: “画像応用”, テレビジョン学会誌, Vol.46, No.7, pp.888-

894 (1992)

- [2] 岡崎彰夫: “画像データベース概論”, 情報処理, Vol.33, No.5, pp.448-456 (1992)
- [3] Hiroshi Arisawa, Hisayoshi Nagae, Yasuko Mochizuki: “Representation of Complex Objects in Semantic Data Model “AIS” and Implementation of Set Operators”, IEICE Transactions, Vol.E 74, No.1, pp.191-203 (1991)
- [4] 永江尚義: “意味データベースシステムの操作系と実現手法に関する研究”, 横浜国立大学工学部1993年度博士論文 (1994)
- [5] 有澤博, 由井仁, 富井尚志: “映像データベースシステムの構成の方式”, 1993年アドバンスデータベースシンポジウム論文集, 情報処理学会, pp.181-190 (1993)
- [6] 有澤博, 由井仁, 富井尚志: “知的作業支援基盤としての映像データベース”, 1994年情報学シンポジウム講演論文集, 日本学術会議他, pp.247-256 (1994)
- [7] 長坂晃朗, 田中謙: “カラービデオ映像における自動索引付け法と物体探索法”, 情報処理学会論文誌, Vol.33, No.4, pp.543-550 (1992)
- [8] 井上誠喜: “画像合成のための対象物抽出法”, 電子情報通信学会論文誌 D-II, Vol.J74-D-II, No.10, pp.1411-1418 (1991)
- [9] John Backus: “Can Programming Be Liberated from the von Neumann Style? A Functionnal Style and Its Algebra of Programs”, CACM, Vol.21, No.8, pp.613 (1978)
- [10] Peter Buneman, Robert E. Frankel, Rishiyur Nikhil: “An Implementation Tequique for Database Query Languages”, ACM Transactions on Database Systems, Vol.7, No.2, pp.164-186 (1982.6)