

# 動画画像セグメント検索のためのモデル化手法\*

7R-4

牛尼 剛聡† 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

近年、計算機の処理能力の向上と二次記憶媒体の大容量化により、計算機上で動画画像を利用することが容易になった。それにともない、種々の応用分野において大量の動画画像をデータベース化して統一的に管理・利用することが要求されている。このようなデータベース・システムでは動画画像の内容に基づいた検索を実現することが重要である。本稿では、内容に基づいた動画画像のモデル化手法と検索の概要を述べる。

## 2 動画画像の特徴とモデル化

我々は内容に基づいた動画画像の検索のために、以下に挙げる動画画像の特徴を反映したモデル化が必要であると考える。

### ● 動画画像の指示内容の多様性

動画画像はさまざまな内容を指示する。この性質は次の二つの多様性に起因すると考えられる。

#### 1. 対象指定の多様性

動画画像は時間的な広がりや空間的な広がりを持ち、部分によって指示内容が異なるため、動画画像の部分を表現することが必要である。

#### 2. 指定された対象の内容指示の多様性

動画画像中の同一の部分が異なる内容を指示することがあるため、指示する内容によって同一の動画画像の部分を区別することが必要である。

### ● 動画画像の指示内容の共有

異なる動画画像（またはその部分）が同一の内容を提示する場合があるため、内容の一貫性を保証することが必要である。

### 2.1 セグメントと解釈

動画画像は対象指定の多様性があるため、動画画像の時間的 / 空間的な部分を形式的に扱う必要がある。そのためにセグメントの概念を導入する。一つの動画画像がある基準の下に排他的に分割した一つの要素を原子セグメントと呼ぶ。原子セグメントは部分を指定するために利用

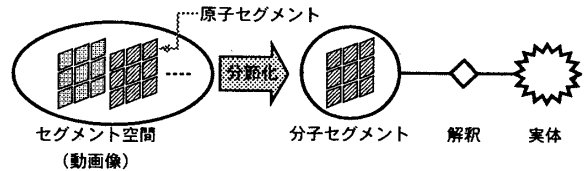


図 1: セグメントと解釈

可能な最小の単位である。一つの動画画像の全ての原子セグメントで構成される集合をセグメント空間と呼ぶ。ある動画画像のセグメント空間は動画画像自身である。セグメント空間内の原子セグメントの集合を分子セグメントと呼び、分子セグメントを生成する操作を分節化と呼ぶ。分子セグメントは、動画画像の時間的 / 空間的な部分を表す。

例えば、動画画像を画素に分割した場合、個々の画素を原子セグメントとすることができる。このときのセグメント空間は動画画像に含まれる全画素から構成される集合である。また、動画画像を構成する一つの静止画像はこのセグメント空間上の分子セグメントの例である。

分子セグメントと実世界の物理的 / 概念的な実体との関係を解釈と呼ぶ。ある分子セグメントと関係する実体を、そのセグメントの内容と呼ぶ。

セグメントと解釈の概念図を図 1 に示す。

### 2.2 モデル化の基本的な枠組

我々の提案する手法では、モデル化の基本単位としてユニットを用いる。全てのユニットは識別子を有し、一意に識別可能である。また、全てのユニットは型のインスタンスとして生成され、型はインスタンスを保持する。型にはユニットに対する操作をメソッドとして定義可能である。ユニットは以下の三種類に分類される。

#### 1. コンセプト

動画画像 (分子セグメント) の内容となる実世界の物理的 / 概念的な実体を表す。コンセプトは属性を有し、その値として他のコンセプトをとる。コンセプトの属性は、動画画像の内容となる実体の普遍的な性質を表すのに用いられる。人物の親子関係、生年月日などは実体の普遍的な性質の例である。

#### 2. サブジェクト

動画画像のセグメントと解釈を表す。サブジェクトはメ

\*A Modeling Method for Retrieving Segments of Movies

†Taketoshi USHIAMA and Toyohide WATANABE

‡Department of Information Engineering, Graduate School of Engineering, Nagoya University  
{ushiama, watanabe}@watanabe.nuie.nagoya-u.ac.jp

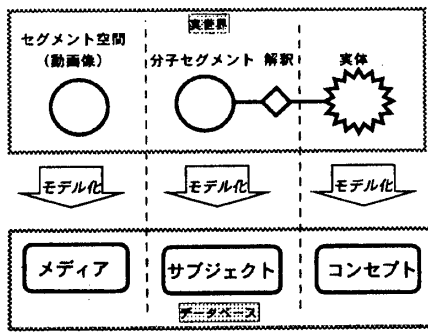


図 2: モデル化の概念図

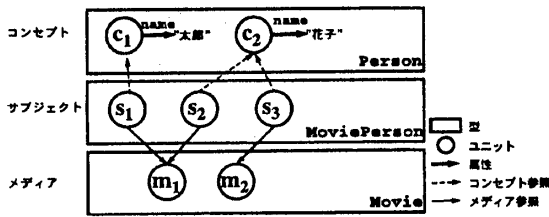


図 3: モデル化の例

ディアに対する参照と分子セグメントに対する指標を有し、参照されるメディアのメソッドを分子セグメントの指標を引数として実行することで分子セグメントを得る。また、サブジェクトはコンセプトに対する参照を有し、分子セグメントの解釈を表す。同一のメディアへの参照を持つ分子セグメントは時間 / 空間軸を共有するため、それらの間に時間的 / 空間的な関係を考えることができる。例えば、分子セグメントを連続するフレーム(カット)としたとき、それを一つの時間区間と考え、Allen[1]が示した時間関係を定義可能である。これらの関係はサブジェクトの型に定義可能とする。

### 3. メディア

個々の動画像を表す。メディアは動画像をデジタル化したビット列を有する。動画像に対する操作はメソッドとして定義される。また、分子セグメント指標から分子セグメントへのマッピングもメソッドとして定義される。

上記のモデル化要素を用いたモデル化の概念を図2に示す。また、この手法に基づいて動画像をモデル化した例を図3に示す。図3において、 $c_1$ は太郎を $c_2$ は花子を表すコンセプトであり、 $m_1, m_2$ はそれぞれ異なる動画像を表すメディアである。 $s_1$ は太郎の $m_1$ 内での表現を、 $s_3$ は花子の $m_1$ 内での表現を、 $s_3$ は花子の $m_2$ 内での表現を表すサブジェクトである。

## 3 ユニット間の関係と検索

### 3.1 ユニット間の関係

二つのユニット  $u_1, u_2$  に対して次の関係を定義する。

- 識別子等価 ( $u_1 =_i u_2$ , または  $u_2 =_i u_1$ )  
 $u_1, u_2$  が同一の識別子を持つ。

- コンセプト等価 ( $u_1 =_c u_2$ , または  $u_2 =_c u_1$ )

- $u_1$ と $u_2$ が共にサブジェクトであり、かつそれぞれが参照するコンセプトが識別子等価である。

- $u_1$ がサブジェクト、 $u_2$ がコンセプトであり、 $u_1$ が参照するコンセプトと $u_2$ が識別子等価である。

- メディア等価 ( $u_1 =_m u_2$ , または  $u_2 =_m u_1$ )

- $u_1$ と $u_2$ が共にサブジェクトであり、かつそれぞれが参照するメディアが識別子等価である。

- $u_1$ がサブジェクト、 $u_2$ がメディアであり、 $u_1$ が参照するメディアと $u_2$ が識別子等価である。

- 属性関係 ( $u_1 \text{ att } u_2$ , または  $u_2 \text{ att}^{-1} u_1$ )

$u_1$ が属性  $\text{att}$  を有し、その値が  $u_2$  と識別子等価である。

### 3.2 検索

二つのユニット集合間の演算である選択演算を用いて検索を行なう。 $A$ と $B$ がユニット集合を表し $\theta$ がユニット間の関係を表すとき、選択演算を $A/[ \theta ] B$ として表す。この演算は集合 $A$ の要素のうちで集合 $B$ の要素と関係 $\theta$ を持つ要素からなる集合を演算結果とする。選択演算の形式的な定義を以下に示す。

$$A/[ \theta ] B = \{ a | a \in A, b \in B, a \theta b \}$$

図3に対して、「太郎が撮影されている動画像の部分を検索せよ」という質問は選択演算を用いて次のように実現される。

$$\text{MoviePerson} / [ =_c ] ( \text{Person} / [ \text{name} ] \{ \text{太郎} \} )$$

ここで、 $\text{MoviePerson}$ と $\text{Person}$ はその型のインスタンスの集合を表す。

上の検索結果を $A$ とすると、「花子が撮影されている動画像の部分のうち、太郎が撮影されている動画像に含まれるものを検索せよ」という質問は次のように実現可能である。

$$( \text{MoviePerson} ) / [ =_c ] ( \text{Person} / [ \text{name} ] \{ \text{花子} \} ) / [ =_m ] A$$

## 4 おわりに

本稿で示した手法は、セグメントを適切に定めることによって静止画像や音などの表現形式にも適用可能である。現在、複数のセグメントを合成し、複合的な動画像を表現可能とするように拡張している。

### 参考文献

[1] J. Allen: "Maintaining Knowledge about Temporal Intervals", *Comm. of ACM*, Vol. 26, No. 11, pp. 832-843 (1983).