

空間・時間属性に基づくビデオ映像のビジュアル検索

黎 亜 和[†] 田 中 譲[†]

従来のビデオ映像検索システムは、各フレームの映像オブジェクト間の空間関係記述を用いる検索法を、数多く提案している。本論文では、映像フレームシーケンス中における、これらの空間関係の変化を記述できるビデオ映像内容記述法を提案する。さらに、映像オブジェクト間の2項時空間関係系列とこれらの系列の正規表現を定義し、映像フレームシーケンス内の2つのオブジェクト間の時空間関係をビジュアルに指定するビジュアル検索要求言語を提案する。本論文は、2項時空間関係系列の正規表現の演算子ごとにSQL表現記述への展開法とを与え、任意のビジュアル検索要求をSQL言語の間合せに変換するアルゴリズムを与えている。最後に、プロトタイプシステムの実現についても述べる。

Visual Query Processing of Video Databases Based on Spatio-temporal Quantification

YAHE LI[†] and YUZURU TANAKA[†]

Most existing systems for searching video sequences for specified ones only handle descriptions of the spatial relationships between video objects within a single frame of the video. This paper introduces a method for expressing the temporal changes of these spatial relationships in a video clip. By defining a sequence of binary spatial relationship types to express the spatio-temporal relationship between two video objects in a video clip, a query representing a desired sequence of spatial relationships can be specified as a regular expression over a set of different binary spatial relationship types. Queries are visually specified as a regular expression, then translated into SQL. We describe both an algorithm for this translation and a prototype system implementation.

1. はじめに

マルチメディア・データベース中に格納した情報に本来備わっている可視的な属性値に対して、ビジュアルに検索条件を指定することは、内容に基づく検索要求を表現する効率かつ自然な方法である。このようなビジュアルな検索を実現するためには、ビジュアルに指定される検索条件を、データベースの内容とのマッチングがとれて検索処理ができる内容記述形式に変換することが必要である。画像あるいはビデオ映像の元データを直接処理したりアクセスすることを避けるために、システムは画像やビデオ映像の内容を記述するための適当な内容記述法を提供しなければならない。

画像に関するビジュアルな検索に関しては、いろいろな研究がなされてきた。Swainら¹⁾、Binaghiら²⁾、Niblackら³⁾は、メニューから検索対象オブジェクトのカラーとテキストチャーを選択すると、データベース

中に格納した画像中のオブジェクトのカラーヒストグラムやテキストチャーの種類との比較を行って、検索を行う方法を提案した。Hirataら³⁾、Niblackら⁴⁾が提案した略図を用いるビジュアルな検索方法は、検索対象オブジェクトの形状を簡略図で描くと、この略図の形状属性によって、データベース中の画像データとのマッチングがとられる。また、Changら^{5),6)}とMohanら⁷⁾などの研究者たちは画像オブジェクト間の空間関係を符号射影⁶⁾という手法で記述表現する方法を提案した。

単一画像に対する内容検索法とそのビジュアルな指定法に関しては、数多くの提案があるが、画像が連続したビデオ映像フレームシーケンスに対する内容検索法とそのビジュアルな指定法についての提案はまだ少ない。Nagasakaら⁸⁾や、Hampapurら⁹⁾たちは連続するフレームのカラーヒストグラムの分析によって、ビデオクリップのカットの変わり目を探す方法を提案している。カラーヒストグラムの連続性が保たれているフレームシーケンスの初めのフレーム画像を索引情報として、ビデオ映像に対して索引を付加することがで

[†] 北海道大学知識メディア・ラボラトリー

Meme Media Laboratory, Hokkaido University

きる。Walter ら¹⁰⁾ は意味ネットワークの構造を用いて、短い映像シーケンスの中の複数映像オブジェクトの間に成立する空間・時間関係を記述することにより、映像シーケンスの内容記述を行っている。Davis¹¹⁾、Amdt ら¹²⁾ は符号射影を用い、単一フレーム中の複数映像オブジェクト間の空間関係を 2D スtring で表現し、シーケンス中の空間関係の変化を 2D String に時間タグを付加することにより、記述する手法を提案した。Del Bimbo¹³⁾ は時間論理と符号射影の拡張により、映像シーケンスの内容を記述する手法を提案した。これらのアプローチはあらかじめ決められたビデオ映像シーケンスの単位ごとに内容記述と検索処理を行うため、記述範囲と検索対象に関する柔軟性は十分といえない。

本論文では、著者ら¹⁵⁾ がすでに提案している、映像フレームごとの映像オブジェクト間の空間関係記述を用い、映像フレームシーケンス中における、これらの空間関係記述の時間変化の系列を記述する方法を提案する。また、2 項空間関係記述系列の正規表現を定義し、この正規表現の演算子ごとに SQL 表現記述への展開法を与える。映像フレームシーケンス内の映像オブジェクトの時空間関係をビジュアルに指定するビジュアル検索要求言語を提案し、空間・時間属性に基づくビデオ映像内容検索システムの実現について述べる。

提案したビデオ映像内容記述法¹⁵⁾ では、映像中の特定の映像オブジェクトを覆う最小の矩形をリファレンスフレーム（以下では RF と略記する）と呼び、各映像オブジェクトに対して、その RF をマニュアル操作で指定することにより分節化^{*}を行い、個々の映像オブジェクトを含む映像シーンの ID、これらの映像オブジェクトの名称、フレーム内における位置と大きさ、現れる時刻（フレームナンバ）などの情報を RF の属性として定義する。

マルチメディアオブジェクトの内容検索は種々のものが考えられるが、本論文では、2 つの映像オブジェクト間に成立する 2 項空間関係の時間的な変化の様子を検索条件として指定する内容検索を取り扱う。より多くの映像オブジェクト間の空間時間関係を検索条件とするには、本論文で扱う 2 項空間時間関係を組み合わせ

せて用いることができる。これらの 2 項空間関係を少数の種類に分類し、2 項空間関係の型を A1, A2 などの記号を用いて記述すると、これらの記号をフレームごとに 1 個ずつ逐次並べた記号列を用いて、空間関係の型の時間変化を記述することができる。この記号列の正規表現を検索条件の記述として用い、この記述を満たす映像フレームシーケンスを検索することができると思われる。本論文は、このような検索記述法を提案し、その検索処理法を明らかにする。

本論文の構成は以下のとおりである。2 章では、空間・時間属性に基づきビデオ映像内容を記述する方法について述べる。3 章では時空間関係の記号系列の正規表現を定義し、この正規表現の演算子の順に従って SQL を用いる検索手法を明らかにする。4 章ではビデオ映像の空間・時間属性に関する検索条件をビジュアルに指定するビジュアル検索要求言語について述べる。5 章はシステムの実装とその評価を報告する。最後に 6 章において、まとめを述べる。

2. ビデオ映像全フレームの内容記述

ビデオ映像は時間的にも空間的にも連続した対象を表現したものである。ビデオ映像の空間属性は個々のフレーム中の映像オブジェクト間の相対的な空間関係により表現できる。また、フレームシーケンスにおいて、フレームごとの空間関係の変化を記述することにより、ビデオ映像の時空間属性を記述することが可能である。さらに、映像オブジェクトの名称などの属性とともに、ビデオ映像の内容を記述することができる。

本論文では、RF を用いてビデオ映像の全フレーム内容記述法を提案する。この方法では、基盤として、先に著者ら¹⁵⁾ が開発したビデオ映像の内容記述法を用い、映像フレームごとに現れる映像オブジェクトを分節化し、名称などの内容や位置、サイズなどの空間に関する RF の属性値を求め、データベースの関係表中にこれらの情報をあらかじめ格納しておく。さらに、映像オブジェクトが現れる時刻（フレームナンバ）も RF の属性として追加することにより、ビデオ映像シーケンス中における、映像オブジェクト間の空間関係の時間変化を記述可能にする。

図 1 と表 1 に、RF を用いて、ビデオ映像オブジェクトに関する空間・時間属性を記述した例を示す。図 1 に示すようにビデオ映像シーンの中の各映像オブジェクトに対して、フレームごとに映像オブジェクトの名称、それを覆う最小の矩形の位置とサイズ、このフレームを含むシーンの ID、映像フレームナンバなどの属性値を求め、映像オブジェクト（すなわち RF）

^{*}「分節化」とは、ビデオ中に現れる映像オブジェクトをその内容が記述される記述単位に分割することである。映像オブジェクトの抽出について、既存の画像処理技術ではそれを自動的に抽出することはそもそも困難である。開発したシステムでは、各映像オブジェクトに対する RF をユーザが直接指定することにより、映像オブジェクトの分節化が行われる。

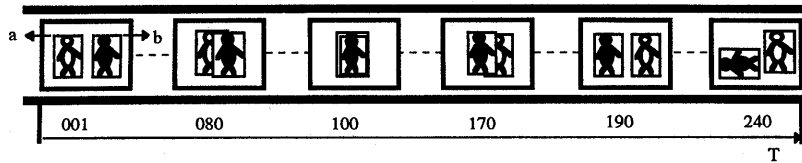


図 1 ビデオ映像オブジェクト間の空間・時間関係

Fig. 1 The spatio-temporal relationship among objects on video.

表 1 ビデオ映像の空間・時間属性と値を記述するレコード構造

Table 1 The record structures to store spatio-temporal properties of each video object.

Reference								Scene			
id	Object Name	LocationX	LocationY	SizeX	SizeY	SceneId	FrameNo	id	Event	Start Frame	Duration
001	a	xa1	ya1	Sxa1	Sya1	scene1	001	scene1	e1	001	240
002	b	xb1	yb1	Sxb1	Syb1	scene1	001				
160	a	xa80	ya80	Sxa80	Sya80	scene1	080				
161	b	xb80	yb80	Sxb80	Syb80	scene1	080				
200	a	xa100	ya100	Sxa100	Sya100	scene1	100				
201	b	xb100	yb100	Sxb100	Syb100	scene1	100				
340	a	xa170	ya170	Sxa170	Sya170	scene1	170				
341	b	xb170	yb170	Sxb170	Syb170	scene1	170				

ごとに1つのレコードを生成し、データベースに格納する。この結果を表1のReferenceとSceneの2つの関係表に示す。属性LocationX, LocationY, SizeX, SizeYにより、映像オブジェクトの位置とサイズに関する空間属性を記述し、属性FrameNoにより、映像オブジェクトの時間属性を記述している。以上述べたビデオ映像の内容記述法は、個々の映像フレーム内の個々の映像オブジェクトごとの記述を独立して取り扱うので、ビデオデータベースの構築に際して、検索条件の指定法や検索対象ビデオ映像シーケンスに関して、制約は何も課していない。

3. 空間・時間属性を用いるビデオ映像の内容検索

3.1 空間関係記号系列の正規表現

個々の映像フレームにおける2つの映像オブジェクト間の相対的な2項空間関係は、図2の第1行に示す左右関係のみの指定と、第1列に示す上下関係のみの指定と、その他の部分に示す左右上下関係を指定する48種類の型に分類することができる。以下では、これらの2項空間関係の型をA1, A2などの記号を用いて表現する。これらの記号を順次並べた記号列を用いて、2項空間関係の型の時間変化を記述できる。これらの記号列の正規表現を2項空間関係記号系列の

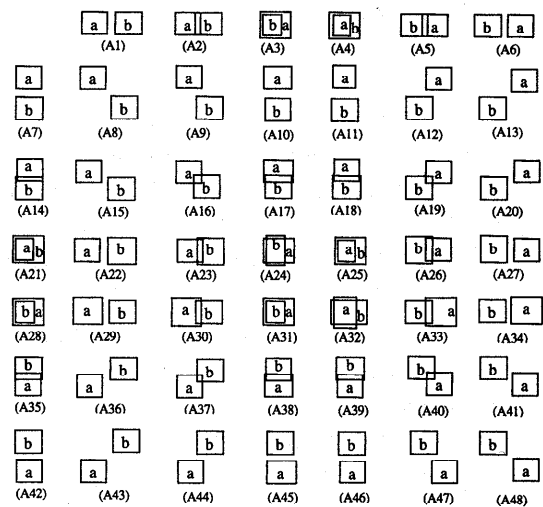


図 2 2項空間関係の型

Fig. 2 Different types of binary spatial relationships.

正規表現と定義する。

定義：2項空間関係記号系列の正規表現と、それとマッチする映像フレームシーケンスを以下のように再帰的に定義する。

- (1) +, ·, † は演算子である。
- (2) 演算子ではない長さ1の2項空間関係記号系列の正規表現 A は記号 A が表記している2項

時空間関係を満たすフレームとマッチする

- (3) α と β が 2 項時空間関係記号系列の正規表現である場合、 $\alpha + \beta$ も 2 項時空間関係記号系列の正規表現であり、かつ α あるいは β が表記している 2 項時空間関係のいずれかを満たすフレームシーケンスとマッチする。
- (4) α と β が 2 項時空間関係記号系列の正規表現である場合、 $\alpha \cdot \beta$ も 2 項時空間関係記号系列の正規表現であり、かつ α が表記した 2 項時空間関係を満たす映像フレームシーケンスの後に β が表記した 2 項時空間関係を満たす映像フレームシーケンスを結合した映像フレームシーケンスとマッチする。
- (5) α が 2 項時空間関係記号系列の正規表現である場合、 α^+ は 2 項時空間関係記号系列の正規表現であり、かつ α が表記した 2 項時空間関係を満たす映像フレームシーケンスを 1 つ以上並べて結合した映像フレームシーケンスとマッチする。

以上の定義によって、各 2 項空間関係の型を満たすビデオ映像フレームとその変化系列を満たす映像フレームシーケンスを規定することができるようになる。たとえば、A1, A2, A5, A6 が異なった 2 項空間関係型とすると、2 項時空間関係記号系列の正規表現 $A1^+ \cdot (A2 + A5)^+ \cdot A6^+$ では A1 が規定する 2 項空間関係を満たす映像フレームが続くシーケンスの後に、A2 あるいは A5 が規定する 2 項空間関係を満たす映像フレームが続くシーケンスが続き、その後に A6 が規定する 2 項空間関係を満たす映像フレームが続くシーケンスが結合されたビデオ映像クリップを規定している。

以下の節で、正規表現の演算子ごとに SQL 表現記述への展開法を与える。検索の途中結果はすべて以下に示す型を持った検索結果表に格納されるものと仮定する。この表の属性 Sid, FMin, FMax は検索条件を満たす映像クリップを含む映像シーンの ID と、その映像クリップの開始フレームナンバと終了フレームナンバである。この型の表をフレームシーケンス表と呼ぶことにする。

R (Sid, FMin, FMax)

3.2 A の映像フレームシーケンスの検索

以下の式 (1) に示した SQL 文は、A が規定する 2 項空間関係を満たすフレームを求める検索要求を記述する。この中の a と b は映像オブジェクトの名称である。Condition(A) は 2 項空間関係型 A に関する条件式である。異なる 2 項空間関係の型を指定するには、

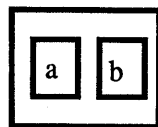


図 3 2 つの映像オブジェクトが離れている空間関係

Fig. 3 The spatial relationship between two objects that are apart from each other.

この 2 項空間関係の条件式の記述を変更すればよい。
SQL(A)

```
⇒ CREATE TABLE R(Sid, FMin, FMax) (1)
AS SELECT R1.SceneId AS Sid,
          R1.FrameNo AS FMin,
          R1.FrameNo AS FMax
FROM Reference R1, Reference R2
WHERE R1.ObjectName = a
AND R2.ObjectName = b
AND R1.SceneId = R2.SceneId
AND R1.FrameNo = R2.FrameNo
AND Condition (A)
```

例として図 3 に示したような 2 つの映像オブジェクトが左右に離れていることを表す 2 項空間関係に対しては、以下の式 (2) に示した 2 項空間関係の条件式を式 (1) に代入すればよい。

Codition(A) :

```
R1.LocationX + R1.SizeX < R2.LocationX (2)
```

3.3 $\alpha \cdot \beta$ の映像フレームシーケンスの検索

$\alpha \cdot \beta$ が規定する検索条件を指定する場合、親文の FROM 句で演算対象の α と β を満たすフレームシーケンス表 R_1 と R_2 を指定し、これらの間の結合条件を親文の WHERE 句で規定することができる。したがって、この SQL 文は以下の式 (3) のようになる。
SQL($\alpha \cdot \beta$) (3)

```
⇒ CREATE TABLE R(Sid, FMin, FMax)
AS SELECT R1.Sid, R1.FMin, R2.FMax
FROM R1, R2
WHERE R1.FMax + 1 = R2.FMin
AND R1.Sid = R2.Sid
```

例として、式 (3) の SQL 文中の R_1 と R_2 を表 2 の (a), (b) に示す。これらの表は式 (1) のような SQL 文を実行することにより得られる検索結果で、図 4 に示した映像オブジェクトの 2 項空間関係の型 A1 と A2 を満たすフレームシーケンスのみを集めた 2 つのフレームシーケンス表である。 R_1 と R_2 を結合する条件は、2 つの 2 項空間関係を満たす映像フレームが連続していること、すなわち $R1.FMax + 1 = R2.FMin$ を満たすことである。この検索結果のフレームシーケ

表2 α を満たす R_1 と β を満たす R_2 に対し, $\alpha \cdot \beta$ を満たす R
 Table 2 Two relations R_1 and R_2 respectively satisfying α and β , and the relation R that satisfies $\alpha \cdot \beta$.

Sid	FMin	FMax	Sid	FMin	FMax	Sid	FMin	FMax
011	030	030	011	031	031	011	030	031
011	032	032	011	033	033	011	032	033
011	035	035	011	036	036	011	035	036
(a) R_1			(b) R_2			(c) R		

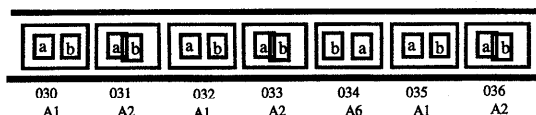


図4 2項空間関係とその系列

Fig. 4 Binary spatial relationships and a sequence of them.

ンス表は表2(c)のようになる。

3.4 $\alpha + \beta$ の映像フレームシーケンスの検索

式(4)に示したようなSQL文で, $\alpha + \beta$ が規定する2つの2項時空間関係の和演算を指定することができる。

SQL($\alpha + \beta$) (4)

```
⇒ CREATE TABLE R(Sid, FMin, FMax)
AS SELECT Sid, FMin, FMax
FROM R1
UNION
SELECT Sid, FMin, FMax
FROM R2
```

このSQL文は, 演算対象の α と β を満たすフレームシーケンス表 R_1 と R_2 を指定し, UNION演算子によってこの2つフレームシーケンス表の集合和を求めている。

3.5 $\alpha^\dagger$ の映像フレームシーケンスの検索

推移的な閉包を探す機能を付加したOracle SQL¹⁶⁾のSTART WITHとCONNECT BYという階層問合せ句を用いて, $\alpha^\dagger$ という連続したフレームシーケンスを検索することが可能である。以下の式(5)のSQL文で, α が規定する2項時空間関係を満たすすべての連続したフレームシーケンスの開始と終了フレームナンバを求めることができる。

SQL($\alpha^\dagger$) (5)

```
⇒ CREATE TABLE R(Sid, FMin, FMax)
AS SELECT MIN(Sid) AS Sid,
MIN(FMin) AS FMin, FMax
FROM (SELECT MIN(Sid) AS Sid,
MIN(FMin) AS FMin
MAX(FMax) AS FMax
```

```
FROM (SELECT Sid, FMin, FMax,
(ROWNUM-LEVEL) AS X
FROM R1
CONNECT BY PRIOR
FMax + 1 = FMin
AND PRIOR Sid = Sid)
GROUP BY X)
```

GROUP BY FMax

たとえば, α が図4に示した映像オブジェクトの2項空間関係の型A1とA2の積演算 $A1 \cdot A2$ である場合, $\alpha^\dagger$ とは, 式(3)のSQL文で得られるフレームシーケンス表から, 接続しているフレームシーケンスをすべて1つにまとめたシーケンスを求め, 得られるフレームシーケンスの開始と終了フレームナンバを求めることに相当する。式(5)の検索要求文のFROM句で指定している表 R_1 は, α が規定する2項時空間関係を満たすすべてのフレームシーケンスをタプルとして持つフレームシーケンス表である。表はシステムによって自動的に作成されるので, タプルが時間順に並ぶことになるが, 本節で述べた検索処理自体はこれを仮定する必要がある。

階層問合せを実行するSELECT文では, 階層のレベルを示すLEVEL疑似列と呼ばれるOracle SQLの機能を用いている。LEVEL疑似列は仮想的に定義された属性と見なすことができる。他のDBMSでも推移閉包が計算できるならば同様の処理を行うことが可能である。式(5)のSQL文は, フレームシーケンス表の行番号ROWNUMとLEVELの差によって連続するフレームシーケンスをグループ化し, 表中の連続したフレーム階層の開始と終了フレームナンバを探し出している。式(5)のSQL文の処理の途中経過を, 表3の(a), (b), (c)に示す。(a)はフレームシーケンス表 R_1 を示す。(b)と(c)はROWNUMとLEVELの差の値と, それによってグループ化を行った結果を示す。さらに, これらのフレーム階層の終了フレームナンバFMaxの値でグループ化し, α で規定した2項時空間関係を満たすすべての連続するフレームシーケンスの開始と終了フレームナンバを(d)のように求めている。

表 3 検索例の処理における途中の関係表と結果の関係表

Table 3 The intermediate relation and the result relation for the example query processing.

Sid	FMin	FMax
011	030	031
011	032	033
011	035	036

(a) フレームシーケンス表 R_1

Sid	FMin	FMax
011	030	033
011	032	033
011	035	036

(c) X をグループ化した結果

Sid	FMin	FMax	X
011	030	031	000
011	032	033	000
011	032	033	002
011	035	036	003

(b) ROWNUM と LEVEL の差

Sid	FMin	FMax
011	030	033
011	035	036

(d) FMax をグループ化した結果

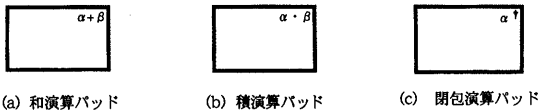


図 5 正規表現の和, 積, 閉包演算パッド

Fig. 5 Regular expression pads for the union, the concatenation, and the transitive closure operations.

4. ビジュアルな検索言語の実現

我々¹⁵⁾ はすでに, IntelligentPad システムを用いて, ビジュアル検索要求インタフェースの構築, と映像オブジェクト間の空間関係をビジュアルに指定する方法を提案した. 本章ではこれを拡張し, 2 項時空間関係の記号系列正規表現を用いて, 2 項時空間関係をビジュアルに指定する方法と, これに基づくデータベースの検索処理法を与える.

4.1 ビジュアルな検索要求の指定

2 項空間関係記述の時間変化の系列をビジュアルに指定するために, 2 項時空間関係記号系列の正規表現の定義に従って, 図 5 に示す 2 項時空間関係記号系列の正規表現の和, 積, 閉包演算パッドを用意する. 演算パッドの上に貼り付けたパッドは演算対象に対応する. 映像オブジェクトとそれらの間の 2 項空間関係を示すために, 映像オブジェクトを示す映像オブジェクトパッドと, 2 項空間関係を示す空間関係パッドを用意する. これらのパッドを用いることにより, 図 6 に示す検索要求インタフェース上に 2 つの異なる映像オブジェクト間の空間・時間関係をビジュアルに規定して, 検索要求を指定することができる.

4.2 検索要求を処理するアルゴリズム

2 項時空間関係記号系列の正規表現を定義したビジュアル検索要求の指定方法とその演算の SQL 文表現によって, ビジュアル検索要求の SQL 文の自動生成と

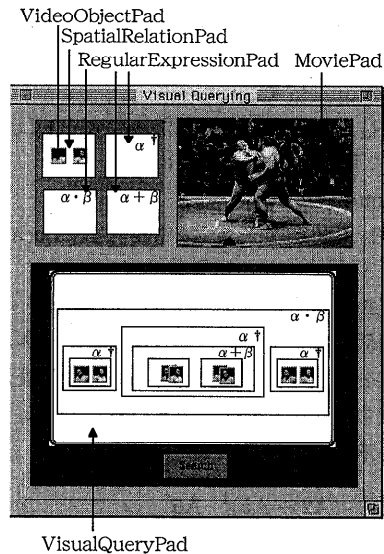


図 6 ビジュアルな検索要求インタフェース

Fig. 6 A visual query interface.

それを用いるビデオ映像の検索を行うことができる. 以下にこの処理のアルゴリズムについて述べる.

- (1) 図 6 に示すビジュアル検索要求上の個々の空間関係パッドに対し, 2 項空間関係の型とその条件式を求める. 対象となる空間関係パッド上の映像オブジェクトパッド間の 2 項空間関係は, 映像オブジェクトパッドの位置とサイズによって, 我々が先に示したアルゴリズム¹⁵⁾ により, 映像オブジェクトの名称とその間の 2 項空間関係の型を確定し, この 2 項空間関係の条件式 Condition (A) を生成することができる. たとえば, 図 6 に示すビジュアルな検索要求に対して, 以下の 4 つの型の 2 項空間関係とその条件式が得られる.

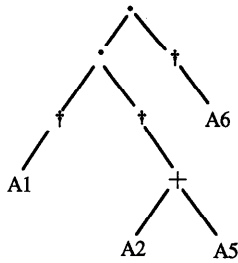


図 7 演算順序木
Fig. 7 An operation tree.

Condition(A1):

$$R_1.LocationX + R_1.Size < R_2.LocationX$$

Condition(A2):

$$R_1.LocationX + R_1.Size/2 \leq R_2.LocationX \\ \leq R_1.LocationX + R_1.Size$$

Condition(A5):

$$R_2.LocationX + R_2.Size/2 \leq R_1.LocationX \\ \leq R_2.LocationX + R_2.Size$$

Condition(A6):

$$R_2.LocationX + R_2.Size \leq R_1.LocationX$$

- (2) 正規表現演算パッドの貼り付け構造と 2 項空間関係の型を用いて定義されたビジュアル検索要求を 2 項時空間関係記号系列の正規表現式に変換する。たとえば、図 6 に示すビジュアル検索要求は、以下の式 (6) になる。

$$A1^+ \cdot (A2 + A5)^+ \cdot A6^+ \quad (6)$$

- (3) 手順 (2) で求められた 2 項時空間関係記号系列の正規表現式に対して、以下の式 (7) に示した演算順序によって、演算順序木を求める。例として、式 (6) の演算順序木は、図 7 に示す。
- $$() \rightarrow + \rightarrow \cdot \rightarrow + \quad (7)$$
- (4) 演算順序木の各リーフノードに対し、3.2 節の SQL 文を用いて、フレームシーケンス表を求める。
- (5) 子ノードのフレームシーケンス表がすべて求まっているノードに対し、演算の種類に応じて、3.2~3.5 節の SQL 文を用いて、フレームシーケンス表を求める。
- (6) (5) をルートノードに至るまで繰り返す。

5. システムの実装と評価

5.1 システムの実現環境

Oracle7 RDBMS と IntelligentPad システムを用い、すでに開発したメタデータの管理に基づくビデオ

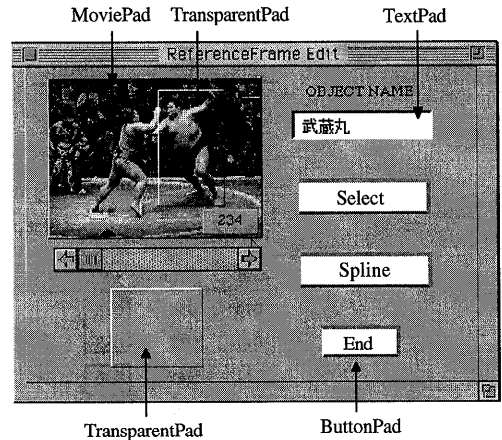


図 8 マニュアル操作で RF を規定するためのサポートツール
Fig. 8 The support tool to specify reference frames.

データベースアーキテクチャのシステム¹⁵⁾を拡張して、システムの実装を行った。

5.2 RF の属性の抽出

映像フレームごとに RF の属性値を求めるために、実装システムでは我々が開発した支援ツール EditProcessPad¹⁵⁾の機能を拡張し、RF の属性の抽出処理を行う。図 8 に示したように、選択したビデオ映像を再生しながら、参照したい映像オブジェクトの位置とサイズの変化の状況によって、サンプル映像フレームを選ぶ。この映像フレームに対して、ユーザはマニュアル操作で映像オブジェクトの名称を入力し、図 8 の右上の MoviePad 上に示した映像オブジェクトの位置とサイズに合わせて、図 8 の左下に用意した透明なパッドを用いて貼り付けると、映像オブジェクトの位置、サイズが自動的にシステムに入力される。さらに、システムは補間計算によって、2つのサンプル映像フレーム間の各映像フレームごとに映像オブジェクトの位置、サイズを求め、これらの属性値の他に映像オブジェクトの名称や映像フレームナンバなどの属性値を含むレコードを生成し、データベース中の関係表 Reference (表 1 参照) に登録する。

5.3 ビジュアルな検索要求のインタフェース

ビジュアルな検索要求インタフェースは、すでに我々が IntelligentPad システムを用いて開発したビジュアル検索要求インタフェース¹⁵⁾に新たなパッドと処理機能を加えて実現した。図 6 に示したように、インタ

☆ RDBMS と IntelligentPad システムを用いたシステム実現の詳細については、文献 15) を参照。本論文で提案するように検索記述能力と検索処理法を拡張しても、RDBMS と IntelligentPad の間の統合のフレームワークは同一のものを使用することができる。

フェースの左上の部分には、VideoObjectPad, SpatialRelationPad, RegularExpressionPad が用意されている。ユーザは、これらのパッドを複製し、VisualQueryPad 上に貼り合わせる。プルダウンメニューで映像オブジェクトを選び、図 6 下の部分に示すように VisualQueryPad 上にこれらを並べて、映像オブジェクト間の 2 項空間関係の時間変化を正規表現でビジュアルに指定できる。図 6 の例では、「力士 1 が力士 2 の右側にいる場面から始まり、力士 1 が力士 2 の左側にいる場面で終わり、これらの間には力士 2 が力士 1 の右側に重なったり、力士 1 が力士 2 の右側に重なったりする場面がいくつか続く」という力士 1 対力士 2 の対戦シーンのビデオ映像クリップを探し出すことが要求されている。検索結果のビデオ映像クリップは図 6 の右上に示した MoviePad により再生される。

5.4 評価

システムの実装例では、相撲の 2 日間の対戦の映像を用い、QuickTime ムービを用いてデジタル化した。2 人の力士が立ち上がった後勝負がつくまでのシーンのみを対象として前処理を行ったところ、平均の長さが 16 秒であるシーンが 40 シーンあった。

ムービのレートが 15 fps であるため、映像フレームの総数は 9600 になった。力士を対象として RF の属性の抽出処理をした結果、各フレームには 2 力士が存在するので、RF の総数は 19200 個になった。これは Reference 表のレコード総数でもある。相撲の一場所における Reference 表のレコード総数は約 144,000 となる。1 年間の場合、Reference 表のレコード総数は約 864,000 程度になると推定される。

相撲の 2 日間の対戦の 40 シーンに対して、ユーザとシステムの対話でフレームごとに RF の属性値を求める前処理には、約 2 時間を要した。この時間の長さはサンプル映像フレームの数に依存する。データベースに格納する取り組み数を増やすと、前処理の時間は長くなる。Nagasaka ら⁸⁾や、Hampapur ら⁹⁾たちは連続するフレームのカラーヒストグラム、あるいは累積したピクセルの相違点の分析によって、ビデオクリップのカットの変わり目を探す方法を提案しているが、既存の画像処理技術では映像オブジェクトを自動的に分節化することはそもそも困難である。実装システムでは、補間計算により、2 つのサンプル映像フレーム間の各映像フレームを自動的に処理することによって、RF の属性値を求める前処理に要する時間を大幅に減少することができた。

6. おわりに

本論文では、ビデオ映像データの空間・時間属性を記述するために、RF を用いるビデオ映像の全フレームの内容記述法を提案した。この方法では、映像フレームごとに、そこに現れる映像オブジェクトの名称などの内容や位置、サイズなどが RF の属性として求められ、これらがレコードとしてまとめられ、あらかじめデータベース中の関係表に格納される。

映像フレームシーケンス中の映像オブジェクト間の時空間関係のビジュアルな指定とその検索処理を実現するために、本論文では、2 項時空間関係系列の正規表現記述を定義し、この正規表現の演算子ごとに SQL 表現記述への展開法を明らかにした。これによって、映像オブジェクト間の時空間関係をビジュアルに指定する方法と、空間・時間属性に基づきビデオ映像を検索する手法を示した。

提案された手法をもとに、Oracle DBMS と IntelligentPad システムを用いて、空間・時間属性に基づくビデオ映像のビジュアル検索が実現できることを示した。

本論文で提案された手法では、既存のデータベース技術をもとに、映像フレームシーケンス上のオブジェクト間の空間・時間関係のような動的かつ連続的な検索条件を記述し、これらの検索条件を用いた内容検索を可能にしている。映像オブジェクトの名称、これらの間の時空間関係の変化など意味的な情報が検索条件として指定でき、検索対象のビデオ映像シーケンスに関し何らの制約も仮定していないので、柔軟性の高い検索が実現できている。

今後の課題として、3 つ以上の映像オブジェクトの間の時空間関係を検索条件とするために、2 項時空間関係を組み合わせる用いる対処方法に関する研究を進めていきたい。

参考文献

- 1) Swain, M.J. and Gallard, D.H.: Color Indexing, *Int'l J. Computer Vision*, Vol.7, No.1, pp.11-32 (1991).
- 2) Binaghi, E., Gagliardi, I. and Schettini, R.: Indexing and Fuzzy Logic based Retrieval of Color Images, *IFIP Trans. A-7, Visual Database Systems II*, Knuth, E. and Wegner, L.M. (Eds.), pp.79-992 (1992).
- 3) Niblack, W., et al.: The QBIC Project: Querying Images by Content Using Color, Texture, and Shape, Res. Report 9203, IBM Res. Div.

- Almaden Res. Center (1993).
- 4) Hirata, K. and Kato, T.: Query by Visual Example: Content-based Image Retrieval, *Advances in Database Technology-EDBT '92*, Pirotte, A., Delobel, C. and Gottlob, G. (Eds.), Lecture Notes on Computer Science, Vol.580, pp.56-71, Springer Verlag, Berlin (1992).
 - 5) Chang, S.K. and Liu, S.H.: Picture Indexing and Abstraction Techniques for Pictorial Databases, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.6, No.4, pp.475-484 (1984).
 - 6) Chang, S.K., Shi, Q.Y., Dimitroff, C.W. and Arndt, T.: Iconic Indexing by 2D String, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.9, No.3, pp.413-427 (1987).
 - 7) Mohan, L. and Kashyap, R.L.: An Object-Oriented Knowledge Representation for Spatial Information, *IEEE Trans. Softw. Eng.*, Vol.14, No.5, pp.675-681 (1988).
 - 8) Nagasaka, A. and Tanaka, Y.: Automatic Video Indexing and Full Video Search for Object Appearances, *IFIP Trans. Visual Database System II*, Knuth, E. and Wegner, L.M. (Eds.), pp.113-127 (1992).
 - 9) Hampapur, A., Jain, R. and Weymouth, T.: Digital Video Indexing in Multimedia Systems, *Proc. AAAI '94 Workshop Indexing and Reuse in Multimedia*, pp.187-198 (1994).
 - 10) Walter, I.M., Sturm, R. and Lockemann, P.C.: A Semantic Network Based Deductive Database System for Image Sequence Evaluation, *IFIP Trans. Visual Database System II*, Knuth, E. and Wegner, L.M. (Eds.), pp.251-276 (1992).
 - 11) Davis, M.: Media Dstreams, an Iconic Visual Language for Video Annotation, *Teletronik*, No.4, pp.59-71 (1993).
 - 12) Amdt, T. and Chang, S.K.: Image Sequence Compression by Iconic Indexing, *IEEE VL '89 Workshop on Visual Languages*, Roma, Italy, pp.177-182 (1989).
 - 13) Del Bimbo, A.: Symbolic Description and Visual Querying of Image Sequences Using Spatio-Temporal Logic, *IEEE Transaction on Knowledge and Data Engineering*, Vol.7, No.4 (1995).
 - 14) Akur, R. and Henzinger, T.A.: Logics and Models of Real Time: A Survey, Technical Report, No.92-1262, Dept. of Computer Science, Cornell Univ., Ithaca, N.Y. (1992).
 - 15) 黎 亜和, 田中 譲: メタデータの管理に基づくビデオデータベースの構成, 情報処理学会論文誌, Vol.39, No.4, pp.1137-1145 (1998).
 - 16) Oracle7 Server SQL 言語リファレンス, リリース 7.2, pp.4398-4402, Oracle Corporation (1995).

(平成 10 年 1 月 20 日受付)

(平成 10 年 11 月 9 日採録)

黎 亜和 (学生会員)



1982 年中国中南工業大学計算機科学科卒業。同年中国非鉄金属鉱産地質研究院に勤務。主任技師。1993～1994 北海道大学客員研究員。現在、同大学院工学研究科博士後期課程在学中。データベース、マルチメディアデータベース等に興味を持つ。中国計算機学会会員。

田中 譲



1972 年京都大学電気工学科卒業, 1974 年同大学院電子工学専攻修士課程修了。工学博士。1974 年北海道大学工学部助手。1977 年同講師, 1985 年同助教授を経て, 1990 年同教授, 現在に至る。1996 年北海道大学知識メディアラボラトリー長。この間, 1985 年 10 月より 1 年間, IBM 社 T.J. ワトソン研究所客員研究員。ソフトウェア学会, 人工知能学会, 米国 IEEE 各会員。データベース理論, データベース・マシン, 並列処理アーキテクチャ, メディア・アーキテクチャ等の研究に従事。「コンピュータ・アーキテクチャ」(雨宮氏との共著, オーム社)等の著書あり。1994 年に IntelligentPad の開発に関して日経 BP 技術賞大賞受賞。