

時空間概念データモデルの実装

天 笠 俊 之^{†,☆} 有 次 正 義[†]
 田 中 貴 之[†] 金 森 吉 成[†]

本稿では我々が提案した時空間概念データモデルをオブジェクトデータベース上のクラスライブラリとして実装する手法について述べる。このクラスライブラリは時間クラス、空間クラス、時空間クラスの三つのクラスから構成される。時間クラスには我々が提案している時区概念データモデルに基づくクラスライブラリを用いる。このためにまず空間データを表現するクラスを構築し、これを時間クラスと組み合わせることによって時空間クラスを構築する。また、このクラスライブラリを用いた時空間の問合せの例を示す。更に、R*-tree に基づいたインデックス構造を提案し、時空間の問合せが効率良く処理できることを示す。

An Implementation of Spatio-Temporal Conceptual Data Model

TOSHIYUKI AMAGASA,[†] MASAYOSHI ARITSUGI,[†] TAKAYUKI TANAKA[†]
 and YOSHINARI KANAMORI[†]

This paper describes a way of implementing a spatio-temporal conceptual data model which we have proposed as a class library on an object database system. The class library is composed of three classes: temporal, spatial, and spatio-temporal classes. As the temporal part of our class library, we exploit time-interval class based on time-interval conceptual model which we have been proposed. We first build a class for spatial concepts, and then integrate the class with the time-interval class. We also show some examples of spatio-temporal queries using the class library. Moreover, we investigate an indexing structure based on R*-tree for the class library so that any spatio-temporal query can be efficiently processed with it.

1. はじめに

静止画像、動画像などの内容検索は重要な技術課題であり、近年活発に研究が行われている^{6),8),11),12),14),15),17)~21)}。その手法は大きく、(1) 画像データから特徴量を抽出し、その類似度によって所望の画像を獲得する方法、(2) 画像中に出現する検索対象をあらかじめオブジェクトとして認識しておき、そのオブジェクトに関する検索を行う方法の二種類に分けられる。前者の手法では、検索対象のオブジェクトの認識が必要でない、またはそれが不可能な場合に用いられる。例えば、画像から受ける印象や色の配置といった特徴量から絵画や写真を検索することができる。一方、後者の手法ではオブジェクトの画像内における位置や動きなどから検索を行うことができるので、医療画像や動画

像の検索などに用いられる。我々の研究グループでは、後者の手法を用いて時系列画像の検索を可能にする時空間概念データモデルを提案している^{5),6),21)}。

画像中に現れるオブジェクトをモデル化するのについていくつかの手法が提案されているが、多くの研究において外接長方形 (Minimum Bounding Rectangle: MBR) が用いられている。外接長方形とは物体の外形がちょうど収まるような長方形のことである。例えば図 1 に示す CT による患者の断面図を考える。この画像には肝臓、大動脈および腫瘍が含まれ、これらが外接長方形によって表現されている。一般に、このような画像は時間の経過に伴い複数枚が撮影され、それぞれの画像にはそれが撮影された日付が時間情報として付与される。外接長方形を用いたアプローチはモデルが簡潔であり、必要なデータ量および計算量が少なく済む。その反面、得られる解に計算誤差が含まれるため、必ずしも正確な解を得られるとは限らないという欠点を有する。

本モデルでは画像内に現われる検索対象を外接長方形によって表現する。また外接長方形に時区間を組み合わせることによって時空間直方体 (Minimum Bounding

[†] 群馬大学工学部情報工学科

Department of Computer Science, Gunma University

[☆] 現在、奈良先端科学技術大学院大学情報科学研究科

Presently with Graduate School of Information and Science, Nara Institute of Science and Technology

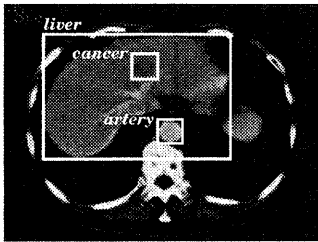


図 1 CT 画像の例

Fig. 1 An example of CT images

Cuboid: MBC) を構成し, その物体が画像内で有効であった時間を表現する. 画像内を時間と共に移動する物体は, 時間軸に沿った時空間直方体の系列によって表現することができる. これを複合時空間直方体 (Composite MBC: CMBC) と呼ぶ.

本稿では, 時空間概念データモデルをオブジェクトデータベース ObjectStore¹⁶⁾ 上のクラスライブラリとして実装する手法について論じる. このクラスライブラリは以下に示す三つの構成要素から成る. すなわち空間クラス (Interval, MBR), 時間クラス (TMInterval, CollectionTM, CompositeTM), 時空間クラス (MBC, CMBC) である. 時間クラスに関しては, 我々がすでに提案している時区間クラスライブラリ^{2)~4)} を利用し, 空間クラスおよび時空間クラスを構築する. 後で述べるように, 空間クラスを実装すれば, これを時間クラスと組み合わせることによって時空間クラスを容易に構築することができる. これらのクラスには, モデルで定義されている全ての演算, 時間的關係, 空間的關係がメソッドおよび演算として実装されるので, モデルで可能な表現は全てこのクラスライブラリで表現可能である.

これに加えて本稿では, 上記クラスライブラリのためのインデックス構造について考察する. これは R*-tree⁷⁾ に基づいており, その特徴として, 物体の座標値だけでなく, その面積や持続期間などを予め計算しておき, 座標値と一緒にインデックスする点が挙げられる. この手法はマルチメディアデータに関するメタデータとして数年来研究が行われている¹⁰⁾. 本稿は静止画像の検索に対してメタデータの手法を用いた場合の定量的な性能評価を与えるものである.

本稿の構成は以下の通りである. 第二章は時空間概念データモデルの概要を述べる. 第三章は本モデルを実装したクラスライブラリについて述べる. 第四章は本クラスライブラリのためのインデックス構造について述べる. 第五章は関連研究について述べる. 第六章はまとめである.

表 1 区間上の演算

Table 1 Operations on intervals

演 算	返 り 値
$intersection(t, u)$	$(\max(t_s, u_s), \min(t_e, u_e)),$ if $t \cap u \neq \emptyset$ 0, otherwise
$cover(t, u)$	$(\min(t_s, u_s), \max(t_e, u_e)),$ if $t \cap u \neq \emptyset$ 0, otherwise
$Re(t, u)$	t と u の間の区間的関係

2. 時空間概念データモデル

本章では時空間概念データモデルの概要を以降の議論に必要な部分に絞って述べる. 詳細な記述に関しては文献^{6), 21)} を参照されたい.

2.1 時空間直方体

本モデルでは画像に現れる対象物を外接長方形 (MBR) によって表現する. さらに, この対象物は次の画像が作成される時間までその形を保っていると仮定すると, その物体は外接長方形に時区間を加えた三次元空間中の直方体と捉えることができる. これを時空間直方体 (MBC) と呼ぶ.

定義 1 (区間) i_s, i_e ($i_s < i_e$) を座標軸上の点とする. 区間 i をこの二点を両端点とする閉区間 (i_s, i_e) と定義する. □

以下では, i の開始点, 終了点および長さをそれぞれ $i.sp = i_s, i.ep = i_e, i.dur = t.ep - t.sp + 1^*$ として参照する. また, $t = (t_s, t_e), u = (u_s, u_e)$ をそれぞれ区間とする場合の区間上の演算を表 1 に示す.

定義 2 (外接長方形) 画像中の物体 O を X 軸および Y 軸に射影してできる区間をそれぞれ $O = (Ox.sp, Ox.ep), Oy = (Oy.sp, Oy.ep)$ とする. O の外接長方形を以下の領域 (x, y) と定義する.

$$Ox.sp \leq x \leq Ox.ep \wedge Oy.sp \leq y \leq Oy.ep$$

以下では, これを $(Ox, Oy; O)$ と表す. □

以下では, 外接長方形 O_{mbr} の保持する区間を $O_{mbr}.x = Ox, O_{mbr}.y = Oy$ として参照する. また, $O_{mbr} = (Ox, Oy; O), P_{mbr} = (Px, Py; P)$ をそれぞれ物体 O, P を表す外接長方形とする場合の外接長方形の演算を表 2 に示す.

定義 3 (時空間直方体) 物体 $(Ox, Oy; O)$ が有効な時間を区間 $Ot = (Ot.sp, Ot.ep)$ で表す. O の時空間直方体を以下の領域 (x, y, t) と定義する.

$$Ox.sp \leq x \leq Ox.ep \wedge Oy.sp \leq y \leq Oy.ep \\ \wedge Ot.sp \leq t \leq Ot.ep$$

* 本モデルでは, 座標軸に離散モデルを採用しているため.

表 2 外接長方形上の演算
Table 2 Operations on MBRs

演算	返り値
$O_{mbr}.area$	$O_x.dur \times O_y.dur$
$O_{mbr}.gravity$	$(\frac{O_x.sp+O_x.ep}{2}, \frac{O_y.sp+O_y.ep}{2})$
$SR(O_{mbr}, P_{mbr})$	$(Re(O_x, P_x), Re(O_y, P_y); O, P)$
$TR(O_{mbr}, P_{mbr})$	O_{mbr} と P_{mbr} の間の位相的關係
$DR(O_{mbr}, P_{mbr})$	O_{mbr} と P_{mbr} の間の方角的關係
$Dist(O_{mbr}, P_{mbr})$	O_{mbr} と P_{mbr} の重心間の距離
$SDist(O_{mbr}, P_{mbr})$	O_{mbr} と P_{mbr} の最短距離

以下では、これを $(Ox, Oy, Ot; O)$ と表す。□

以下では、時空間直方体 O_{mbc} の保持する区間、および外接長方形をそれぞれ $O_{mbc}.x = Ox$, $O_{mbc}.y = Oy$, $O_{mbc}.t = Ot$, $O_{mbc}.mbr = (Ox, Oy; O)$ として参照する。

本モデルの特徴の一つに、物体が現れていた時間だけでなく消滅していた時間も表現できる点が挙げられる。前者は実時間区間 (real-time interval), 後者は空時間区間 (null-time interval) によって表現することができる⁴⁾。具体的には、物体の消滅を表す空外接長方形および空時空間直方体を導入し、それぞれ $(\varepsilon, \varepsilon; O)$, $(\varepsilon, \varepsilon, Ot; O)$ のように表現する。

画像中の物体は時間の経過と共に変化する。これには例えば、移動、拡大、縮小、出現、消滅などがある。これらを表現するために複合時空間直方体 (CMBC) を導入する。

定義 4 (複合時空間直方体) n 枚の画像を含む時系列画像を考え、それぞれが i ($1 \leq i \leq n$) によって順序付けされているとする。

物体 O が与えられ、その時空間直方体を $O_{mbc_i} = (Ox_i, Oy_i, Ot_i; O)$ とすると、 O の複合時空間直方体は以下のように定義される。

$\{O_{mbc_1}, O_{mbc_2}, \dots, O_{mbc_n} |$
 $(\dots((Ot_1 \text{ meets } Ot_2) \text{ meets } Ot_3) \dots \text{ meets } Ot_n)\}$
 ここで $Ot_i \text{ meets } Ot_{i+1}$ は $Ot_i.ep + 1 = Ot_{i+1}.sp$ を表す。□

図 2 に複合時空間直方体の例を示す。図では物体 A が外接長方形によってモデル化されており、画像の左上から時間の経過に伴い下に移動している。途中、時刻 10 から 20 の間 A は画像から消失しているが、その様子は空時間区間によって明示的に表現されている。

物体 O を表す複合時空間直方体を $O_{cmbc} = \{O_{mbc_1}, O_{mbc_2}, \dots, O_{mbc_n}\}$ とする場合の複合時空間直方体の演算を表 3 に示す。

2.2 時空間的關係

本モデルでは物体の間の関係を表現するために Allen による区間的關係¹⁾を用いる。このため物体の形状を

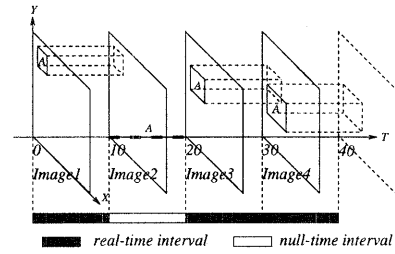


図 2 複合時空間直方体 (CMBC)

Fig. 2 Composite minimum bounding cuboids (CMBC)

表 3 複合時空間直方体上の演算
Table 3 Operations on CMBCs

演算	返り値
$O_{cmbc}.mbcs$	$\{O_{mbc_i} 1 \leq i \leq n\}$
$O_{cmbc}.r$	$\{O_{mbc_i} O_{mbc_i}.x \neq \varepsilon \wedge O_{mbc_i}.y \neq \varepsilon \wedge 1 \leq i \leq n\}$
$O_{cmbc}.n$	$\{O_{mbc_i} O_{mbc_i}.x = O_{mbc_i}.y = \varepsilon \wedge 1 \leq i \leq n\}$

X 軸上および Y 軸上に射影することによって二つの区間を導く。すると物体間の関係をそれぞれの軸に並行な区間の関係の組み合わせによって表現することができる。本モデルではさらに区間関係を拡張することによって、物体間の距離や重なり具合の表現が可能になっている²¹⁾。

二つの外接長方形の間には 8 つの位相的關係 (Disjoint, Meets, Insides, Contains, Overlaps, Equals, Covered_by, Covers) が存在する¹⁵⁾。また、物体間の方角としては上 (above), 左下 (left-below) などが定義されている。これらは X 軸および Y 軸に並行な区間関係を組み合わせることによって表現することができる。この定義の特徴は、物体が互いに重なっているような場合でも方角関係を記述することができる点である²¹⁾。さらに距離的關係として重心間の距離と最短距離の二種類を定義している。

2.3 問い合わせ

ここでは図 3 から図 10 に示した肝臓の CT 画像を例にとって、時空間概念データモデルにおける問い合わせの記述例を示す。これらの図において肝臓には腫瘍が見られるが、治療によってだんだんと縮小して行き、図 10 では消滅している。

これを複合時空間直方体を用いてモデル化すると以下のようなになる。ここで L_{cmbc} と C_{cmbc} はそれぞれ肝臓と腫瘍を表す識別子であるとする。また、それぞれのオブジェクトの座標値は、患者の体を 100×100 の正方形に正規化した上で計算している。

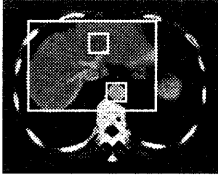


図3 93/11/22
Fig. 3 93/11/22

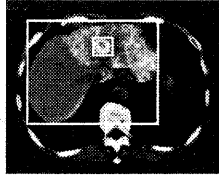


図4 93/12/03
Fig. 4 93/12/03

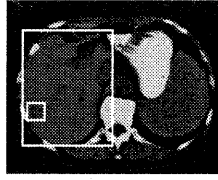


図5 94/07/22
Fig. 5 94/07/22

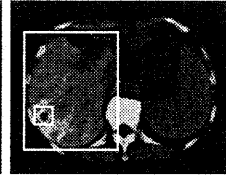


図6 94/08/05
Fig. 6 94/08/05

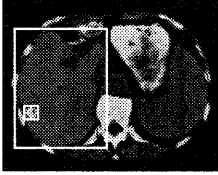


図7 94/09/05
Fig. 7 94/09/05

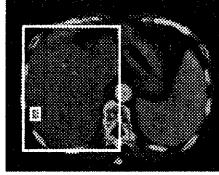


図8 94/10/06
Fig. 8 94/10/06

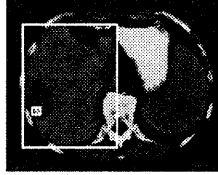


図9 95/01/31
Fig. 9 95/01/31

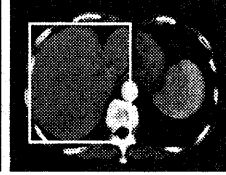


図10 95/04/13
Fig. 10 95/04/13

$$\begin{aligned}
 L = & \{((11, 95), (8, 64), (93/11/22, 94/07/21); L), \\
 & ((4, 62), (11, 83), (94/07/22, 94/08/04); L), \\
 & ((6, 65), (12, 85), (94/08/05, 94/09/04); L), \\
 & ((4, 63), (12, 85), (94/09/05, 94/10/05); L), \\
 & ((3, 70), (9, 88), (94/10/06, 95/01/30); L), \\
 & ((6, 66), (10, 85), (95/01/31, 95/04/12); L), \\
 & ((5, 71), (10, 82), (95/04/13, now); L)\} \\
 C = & \{((41, 64), (15, 29), (93/11/22, 93/12/02); C), \\
 & ((40, 64), (15, 30), (93/12/03, 94/07/21); C), \\
 & ((9, 22), (44, 68), (94/07/22, 94/08/04); C), \\
 & ((10, 23), (45, 70), (94/08/05, 94/09/04); C), \\
 & ((9, 20), (46, 69), (94/09/05, 94/10/05); C), \\
 & ((10, 20), (46, 69), (94/10/06, 95/01/30); C), \\
 & ((7, 14), (45, 67), (95/01/31, 95/04/12); C), \\
 & ((\varepsilon, \varepsilon), (\varepsilon, \varepsilon), (95/04/13, now); C)\}
 \end{aligned}$$

このようにモデル化されたデータに対して問い合わせを行うには、上述の演算を複合時空間直方体に対して適用すれば良い。この例では代数的表記によって問い合わせを行っているが、3章で述べるクラスライブラリでは、同様の演算を用いて OQL 風の記述が可能である。

例1 肝臓の右側に現れた腫瘍を検索せよ。

$$\begin{aligned}
 \{C \mid DR(C_{mbc_i}.mbr, L_{mbc_i}.mbr) = (<_x) \wedge \\
 TR(C_{mbc_i}.mbr, L_{mbc_i}.mbr) = \\
 \text{Insides} \vee \text{Covered_by})\}
 \end{aligned}$$

$= C$

ここで TR は引数として与えられた二つの外接長方形の位相的關係を返す二項演算である。

例2 肝臓の左から右側に広がった腫瘍を検索せよ。

$$\begin{aligned}
 \{C \mid DR(C_{mbc_i}.mbr, L_{mbc_i}.mbr) = (>_x) \wedge \\
 TR(C_{mbc_i}.mbr, L_{mbc_i}.mbr) = \text{Inside} \wedge \\
 DR(C_{mbc_j}.mbr, L_{mbc_j}.mbr) = (<_x) \wedge \\
 TR(C_{mbc_j}.mbr, L_{mbc_j}.mbr) = \text{Inside} \wedge \\
 Re(C_{mbc_i}.t, C_{mbc_j}.t) = (\text{before} \vee \text{meets})\} = C
 \end{aligned}$$

例3 肝臓にある腫瘍が消滅した時間を求めよ。

$$\begin{aligned}
 \{C_{mbc_i}.t \mid C_{mbc_i}.mbr.x \neq \varepsilon \wedge C_{mbc_i}.mbr.y \neq \varepsilon \wedge \\
 C_{mbc_j}.mbr.x = \varepsilon \wedge C_{mbc_j}.mbr.y = \varepsilon \wedge \\
 Re(C_{mbc_i}.t, C_{mbc_j}.t) = \text{meets}\} \\
 = (95/01/31, 95/04/12)
 \end{aligned}$$

この場合、腫瘍を表す MBR が空 (ε) に変化した時区間を探せば良い。

以上の議論をまとめると、本モデルは (1) 時空間の間合せを区間により統一的に記述することができる、(2) 位相、方角、距離などのさまざまな空間的關係が定義され、これらを用いた間合せ表現が可能である、(3) ある特定の応用に特化していない、などの特徴を持つ。ただし、このモデルでは物体の形状を外接長方形により近似しているために、その記述には潜在的な誤差が含まれている。すなわち本モデルから得られる解は候補に過ぎず、正確な解を得るためには、さらに物体形状の詳細な検討が必要である。

3. 時空間クラスライブラリ

3.1 クラスライブラリの構成

本節では時空間オブジェクトを管理するための時空間クラスライブラリについて論じる。このクラスライブラリでは時空間概念データモデルに基づき、区間、外接長方形、時空間直方体、複合時空間直方体の四つの概念をそれぞれクラスとして実装し、時系列画像のモデル化を可能にしている。

本クラスライブラリは、空間クラス (Interval, MBR), 時間クラス (TimeInterval, CollectionTM, CompositeTM), 時空間クラス (MBC, CMBC) から構成される (図 11)。時間クラスに関しては、我々がすでに実装している時区間クラスライブラリ^{2)~4)}を

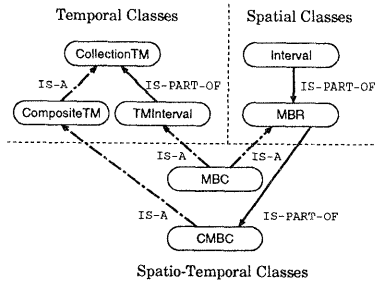


図 11 クラス構成

Fig. 11 Class composition

利用するため、実際、時空間クラスは空間クラスと時間クラスを組み合わせることによって容易に実現することができる。時間クラスの定義については付録を参照されたい。なお、実装には `ObjectStore`¹⁶⁾ を用いている。

3.2 クラス定義

3.2.1 区間クラス (Interval Class)

区間クラスは属性としてその開始点及び終了点の座標値を保持する。Interval クラスで定義されているメソッドを表 4 に示す。

`IntervalRelation`, `PIntervalRelation` は Allen による区間的関係を表現するためのクラスである。特に後者は我々のモデルで拡張されたパラメータ付きの区間的関係²¹⁾ を表現するためのクラスであり、これにより区間の間の距離や重なり具合を表現することが可能になっている。

3.2.2 外接長方形クラス (MBR Class)

外接長方形は X 軸および Y 軸に並行な二つの区間から構成されるので、外接長方形クラスでは二つの区間オブジェクトを保持させる。また、これに加えて外接長方形がモデル化するオブジェクトの `OID` を保持させる。MBR クラスで定義されているメソッドを表 5 に示す。

3.2.3 時空間直方体クラス (MBC Class)

時空間直方体は空間的には外接長方形とみなすことができるのと同時に、時間的には時区間とみなすこともできる。このことから時空間直方体は外接長方形と時区間の双方を継承することによって実現できる。従って MBC クラスは固有のメソッドを持つ必要がない。

3.2.4 複合時空間直方体クラス (CMBC Class)

定義 4 で述べたように、複合時空間直方体とは時間順に整列し、隣接する要素が時間的に接合している時空間直方体の系列である。すなわち、複合時空間直方体は空間的側面からは外接長方形の系列と見なすことができる一方、時間的側面からは隣接する要素が互いに接合している時区間の集合 (複合時区間⁴⁾) と見ることができる。これに従い本実装では、複合時空間直方体を

(1) 空間座標を保持する MBR オブジェクトの系列、および (2) 時間に関する値を保持する `CompositeTM` オブジェクトとの組合せによって実現した。ここで `CompositeTM` クラスとは時区間クラスライブラリ^{2),3)} における複合時区間の実装である。なお、順序付きの集合は `ObjectStore` の `os_List` クラスによって実現している。CMBC クラスで定義されているメソッドを表 6 に示す。

3.2.5 時空間クラスライブラリを用いた問合せの例

本クラスライブラリでは、このようにモデル化された CT 画像に対して OQL⁹⁾ 風のような問合せ表現が可能である。ここで `Cancer` および `Liver` は、それぞれ腫瘍オブジェクトと肝臓オブジェクトを表す識別子とする。

例 4 肝臓の右側に現れた腫瘍を検索せよ。

```

select c
from Cancer c, Liver l
where exists i in [1 ... c.cardinality()]:
  c.select(i).directionRelation(l.select(i)).
    x_relation() > 0 and
  (c.select(i).topologicalRelation(l.select(i))
    == insides ||
  c.select(i).topologicalRelation(l.select(i))
    == coveredby)
  
```

例 5 肝臓の左から右側に広がった腫瘍を検索せよ。

```

select c
from Cancer c, Liver l
where exists i in [1 ... c.cardinality()]:
  exists j in [1 ... c.cardinality()]:
    (c.select(i).directionRelation(l.select(j)).
      x_relation() > 0 and
    c.select(i).topologicalRelation(l.select(j))
      == inside and
    c.select(j).directionRelation(l.select(i)).
      x_relation() < 0 and
    c.select(j).topologicalRelation(l.select(i))
      == inside) and
  (c.select(i).relation(c.select(j)) == before or
  c.select(i).relation(c.select(j)) == meets)
  
```

例 6 肝臓にある腫瘍が消滅した時間を求めよ。

```

select c1
from (select c.mbc() from Cancer c) c1, c2
where c1.x_interval() != 0 and c1.y_interval()
  != 0 and
  c2.x_interval() == 0 and c2.y_interval()
  == 0 and
  c1.relation(c2) == meets
  
```

我々のクラスライブラリでは、時空間概念データモデルにおける全ての演算をメソッドおよび演算子として実装している。すなわち、モデルで可能な全ての問合せをクラスライブラリ上でも表現することができる。従って、時間、区間、空間、位相、方角、距離などのさまざまなオブジェクト間の関係を利用した問合せも可能である。

表 4 Interval クラスのメソッド
Table 4 Methods of Interval class

メソッド名	返り値の型	処理
startPoint	int	開始点を返す
endPoint	int	終了点を返す
duration	int	幅を返す
intersection	Interval*	引数の Interval との共通部分を返す
cover	Interval*	引数の Interval と被覆する Interval を返す
relation	IntervalRelation*	引数の Interval との区間的関係を返す
relationParam	PIntervalRelation*	引数の Interval とのパラメータつき区間的関係を返す

表 5 MBR クラスのメソッド
Table 5 Methods of MBR class

メソッド名	返り値の型	処理
x_interval	Interval*	X 軸の区間を返す
y_interval	Interval*	Y 軸の区間を返す
objectID	int	OID を返す
area	int	面積を返す
spatialRelation	SpatialRelation*	引数の MBR との空間的關係 (X 軸および Y 軸の区間的関係の組み合わせ) を返す
topologicalRelation	TopologicalRelation*	引数の MBR との位相的關係を返す
directionRelation	DirectionRelation*	引数の MBR との方角的關係を返す
distance	double	引数の MBR との重心間の距離を返す
shortestDistance	double	引数の MBR との最短距離を返す

表 6 CMBC クラスのメソッド
Table 6 Methods of CMBC class

メソッド名	返り値の型	処理
mbcs	os_List<MBC*>	MBC の系列を返す
real	os_List<MBC*>	実時空間直方体の系列を返す
null	os_List<MBC*>	空時空間直方体の系列を返す
objectID	int	OID を返す
cardinality	int	要素の数を返す
select	MBC*	順番を指定し, 該当する MBC を返す
append	void	新たな MBC を加える

4. 時空間クラスライブラリのためのインデックス構造

4.1 基本構成

時空間データに対する検索処理の効率化のためには多次元インデックスが有効である。中でも R-tree の改良版である R*-tree⁷⁾ は最も効率の良いインデックスの一つである。R*-tree はその効率のみならず, 多次元データを外接長方形によって扱うため外接長方形に基づいたモデルと親和性が高いことなどから, 多くの研究で利用されている。本研究でも R*-tree を利用し, 時空間クラスライブラリのための効果的なインデックスについて考察する。

多次元インデックスで時空間データを管理する際には, 時間データと空間データをいかに管理するかが問題となる。Vazirgiannis 等は巨大なマルチメディアデー

タのための R-tree に基づいた効果的なインデックス構成法を提案している²²⁾。彼等の研究では以下の手法が比較検討された。

(1) 時空間オブジェクトを以下の二つのインデックスで管理する単純な方法

- 空間に関する属性 (id, x_1 , x_2 , y_1 , y_2) を格納する空間 (2D) インデックス。ここで id はオブジェクト識別子である
- 時間に関する属性 (id, t_1 , t_2) を格納する時間インデックス

(2) 単一の空間インデックス (3D) で時空間に関する全ての情報を取り扱う方法

この結果, 後者の手法がより効果的であることが示されている。

そこで本研究でも同様の手法を用いることにする。すなわち, オブジェクト O の時空間直方体を $O_{mbc} =$

($Ox, Oy, Ot; O$) とするとき, X 軸 ($Ox.sp, Ox.ep$), Y 軸 ($Oy.sp, Oy.ep$), 時間軸 ($Ot.stp, Ot.etp$), オブジェクト ID (O) から成る 4 次元の R*-tree に格納する.

4.2 付加情報のインデックシング

多くのアプリケーションにおいて, オブジェクトの属性に関する問合せに加えて, それから導出される様々な値をインデックシングしておくことは, 問合せ処理の効率化に有効である. これはマルチメディアデータに関するメタデータとしてかねてから研究がなされている¹⁰⁾. 例えば, 本論文で扱っている腫瘍を含む CT 画像に対しては, 腫瘍の面積を予め計算しておくことによって, 「腫瘍の面積が 93/11/22 の時点と比べて半分になったのはいつか」といった問い合わせを効率よく処理できるようになる.

本研究で用いるインデックスには, オブジェクト $O_{mbc} = (Ox, Oy, Ot; O)$ を格納するために X 軸 ($Ox.sp, Ox.ep$), Y 軸 ($Oy.sp, Oy.ep$), 時間軸 ($Ot.stp, Ot.etp$), オブジェクト ID (O), 面積 ($Ox.dur \times Oy.dur$), 持続期間 ($Ot.dur$) の 6 つの軸を用意した. このため, オブジェクトをインデックスに挿入する際には, その面積および持続期間を計算する余分な計算を行う必要がある. 一方, これらの値に対して検索を行う場合には, 通常のインデックスとは違い, 対応する軸に関して範囲検索を行うだけでよい.

4.3 性能評価

付加情報のインデックシングによる検索効率と空間効率の得失について実験による評価を行った. 実験では, 本研究で提案した付加情報を加えたインデックス (6D R*-tree) と通常の時空間インデックス (4D R*-tree) について, 両者のサイズ, およびデータの挿入, 期間による検索, 時空間の範囲検索における両者の性能を比較した (図 12 から 図 15). 実験に用いた環境は Sun Ultra 30 (UltraSPARC-II 296 MHz CPU, 128MB memory), データベース ObjectStore 5.0.0.0 である.

実験に用いたデータでは, 200×200 の画像中にオブジェクトをランダムに生成した. 各オブジェクトは複数の時空間直方体からなり, 各時空間直方体は X 軸, Y 軸, 時間軸に関して一様に分布した区間から構成される. 図 12 はインデックスのサイズである. 横軸は画像に現れるオブジェクトの数を表している. 図から 6D R*-tree の方がおよそ 20% ほどこインデックスのサイズが大きいたことが分かる. これは 4D R*-tree に比べて次元数が大きく, 格納されるデータの量も多いためである. これに従い, データの挿入にかかる時間 (図 13)

表 7 平均ヒット数
Table 7 Avg. # of hits

# of objs	duration	range
10	9.8	0.0
50	45.8	15.2
100	91.3	35.1
150	145.3	52.5
200	194.5	75.0
250	242.7	89.9
300	301.8	108.8

も 6D R*-tree の方が約二倍となっている. この場合, インデックスのサイズだけでなく, 分割軸の増加によるノード分割のコスト増やデータ挿入の際に要する計算の手間などが要因となり, 両者の差が顕著になっている.

実験では, 上記データベースに対して, 以下に示す二つの問合せを行った.

- (1) 期間の検索

```
select c_mbc
from (select c.mbc() from Cancer c) c_mbc
where dur_min <= c_mbc.dur() and
      c_mbc.dur() <= dur_max
```
- (2) 時空間の範囲検索

```
select c_mbc
from (select c.mbc() from Cancer c) c_mbc
where (c_mbc.topologicalRelation(q_mbc)
      != disjoint) and
      (c_mbc.topologicalRelation(q_mbc)
      != meets) and
      (c_mbc.intersection(q_mbc) != NULL)
```

表 7 は, (1) に関しては dur_min と dur_max , (2) に関しては q_mbc をランダムに与えながら十回検索を行い, その結果得られた時空間直方体の平均数である. たとえば表 7 で, オブジェクト数 100 における duration の平均ヒット数 91.3 とは, オブジェクト 100 個のうちの 91.3 個のオブジェクトを検索したのではなく, 範囲検索により平均 91.3 個の時空間直方体が検索されたことを示す.

図 14 は (1) の問合せの処理時間である. 6D R*-tree では, 期間を検索するために, インデックスの対応する軸で単純な範囲検索を行えばよいのに対し, 4D R*-tree ではそれぞれのオブジェクトに対して持続期間を計算しなくてはいけないので, およそ倍の処理時間が必要になっている.

図 15 は q_mbc と重なりを持つような全ての時空間直方体を検索する処理である. 結果より, 6D R*-tree の方がインデックスのサイズが大きくなっている分性能の悪化が見られるが, その差は少なく, 軸を増やしたことによる性能の悪化はほとんど無視できると考えられる.

4.4 議論

付加情報をインデックシングする手法では, インデックス構築に要するコストを犠牲にする代わりに検索性能

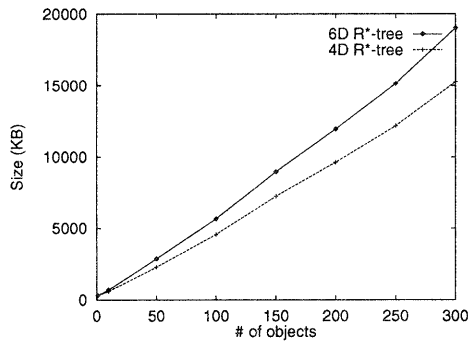


図 12 インデックスのサイズ
Fig. 12 Sizes of indexes

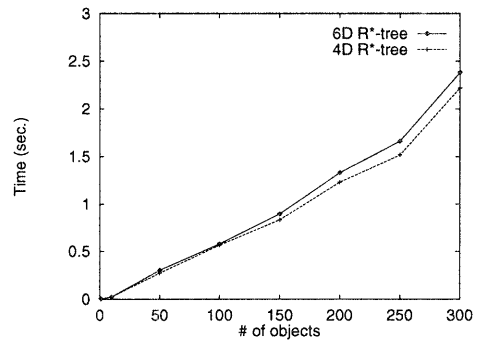


図 15 時空間の問合せ
Fig. 15 Spatio-temporal queries

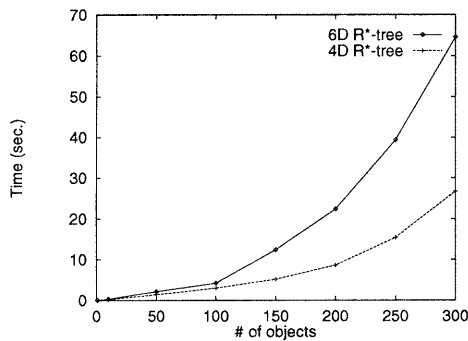


図 13 挿入
Fig. 13 Insertion

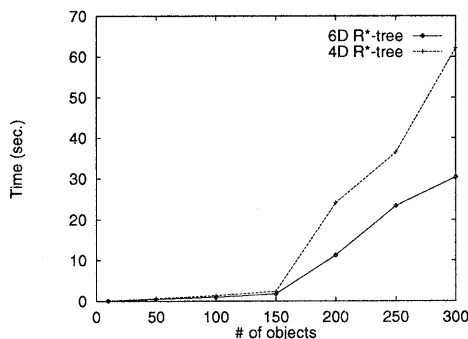


図 14 期間の問合せ
Fig. 14 Queries on duration

の効率を向上させていると考えられる。一般に、静止画像や動画画像などはデータに対する更新の頻度は少ないので、付加情報に関わる計算コストはほぼデータの登録時のみに発生すると考えてよい。従って、この手法は検索の効率を上げるために有効な手段であると言える。

検索を効率化する方法としては、まず、予め計算可能なメソッドの値を計算してインデックスしておく方法が考えられる。例えば、上の実験の *area* や *duration* などのように、メソッドの返り値が属性値から導出できる場合がこれにあたる。この場合、問い合わせの中で対

応するメソッドの呼び出しが行われた場合、単純にメソッドを起動する代わりにインデックスを検索するだけでよい。

さらに、今回行った実験のように、検索処理にかかる時間を定量的に評価しておき、その結果を基に適用するオペレーションの順序を決定する方法も考えられる。この場合、上記手法と組み合わせて更なる効率化を図ることができる。これは今後の課題である。

5. 関連研究

Vazirgiannis 等は時空間オブジェクトを空間および時間の独立したインデックスで管理するよりも、単一の時空間インデックスで管理した方が効率的であることを示した²²⁾。

堀之口等は時空間データモデル Universe のための R*-tree に基づいたインデックスを提案している¹³⁾。この研究において、彼等は R*-tree に正規化の手法を取り入れることによって検索処理の効率化を図っている。この手法は、R*-tree の各軸が表す単位や意味が異なる場合にも空間分割を悪化させないという特徴を持つ。同様の手法は我々の提案するインデックスでも有効であると考えられる。

我々の研究とこれらの研究との相違は、我々のインデックスでは外接長方形の面積やオブジェクトの持続期間などの付加的な情報をもインデックスに取り込んでいる点である。これらの値は頻繁に検索に利用されるので、効率的な検索を期待できる。また、付加情報によりインデックスのサイズが増大するが、それは通常の検索性能にはほとんど影響を及ぼさないことを実験によって示している。

6. おわりに

本稿では、時系列画像をモデル化するためのクラスライブラリについて論じた。このライブラリは時空間概念

データモデルに基づいており、さまざまな時間的、空間的關係を用いたオブジェクトの内容検索を可能にする。内容検索の例として CT 画像による医療画像を示し、これがどのようにモデル化され、検索処理がどのように行われるのかを示した。これに加えて、検索処理を効率的に行うためのインデックス構造を提案した。このインデックスでは、オブジェクトの座標値に加えて外接長方形の面積やオブジェクトの持続期間など、よく問合せられる値をもあらかじめインデックシングしておく。これにより効率的な検索処理を期待することができ、データの更新が少ない静的なデータに関して有効である。

謝辞

本研究を進めるにあたり実験に御協力頂いた金森研究室大学院生櫻井紀美子氏、医療画像を提供して頂いた群馬大学医学部井上登美夫助教授に深く感謝する。また、本研究の一部は文部省科学研究費特定領域研究「高度データベース」(課題番号 08244101)、文部省科学研究費基盤研究(C)(2)(課題番号 10680333)によるものである。

参 考 文 献

- 1) Allen, J. F.: Maintaining Knowledge about Temporal Intervals, *Communications of the ACM*, Vol. 26, pp. 832-843 (1983).
- 2) Amagasa, T., Aritsugi, M. and Kanamori, Y.: Implementing Time-Interval Class for Managing Temporal Data, *9th International Workshop on Database and Expert Systems Applications (DEXA'98)*, IEEE Computer Society, pp. 843-849 (1998).
- 3) Amagasa, T., Aritsugi, M. and Kanamori, Y.: An Implementation of Interval Based Conceptual Model for Temporal Data, *IEICE Trans. Info. & Syst.*, Vol. E82-D, No. 1, pp. 136-146 (1999).
- 4) Amagasa, T., Aritsugi, M., Kanamori, Y. and Masunaga, Y.: Interval-Based Modeling for Temporal Representation and Operations, *IEICE Trans. Info. & Syst.*, Vol. E81-D, No. 1, pp. 47-55 (1998).
- 5) Amagasa, T., Aritsugi, M., Tanaka, T. and Kanamori, Y.: Implementing Class Library and Index for Managing Spatio-Temporal Data, *Int. Workshop on New Database Technologies for Collaborative Work Support and Spatio-Temporal Data Management (NewDB'98)*, LNCS, No. 1552, pp. 532-541 (1998).
- 6) Aritsugi, M., Tagashira, T., Amagasa, T. and Y. Kanamori: An Approach to Spatio-Temporal Queries —Interval-Based Contents Representation of Images—, *8th International Conference on Database and Expert Systems Applications (DEXA 97)*, LNCS, No. 1308, pp. 202-213 (1997).
- 7) Beckmann, N., Kriegel, H.-P., Schneider, R. and Seeger, B.: The R*-tree: An Efficient and Robust Access Method for Points and Rectangles, *Proc. ACM SIGMOD Conf.*, pp. 322-331 (1990).
- 8) Bimbo, A. D., Vicario, E. and Zingoni, D.: Symbolic Description and Visual Querying of Image Sequences Using Spatio-Temporal Logic, *IEEE Trans. Knowledge and Data Engineering*, Vol. 7, No. 4, pp. 609-621 (1995).
- 9) Cattell, R. G.G., Bartels, D., Berler, M., Eastman, J., Gamerman, S., Jordan, D., Springer, A., Strickland, H. and Wade, D.: *The Object Database Standard: ODMG 2.0*, Morgan Kaufmann Publishers (1997).
- 10) Chu, W.W., Hsu, C.-C., Jeong, I.T. and Taira, R. K.: Content-Based Image Retrieval Using Metadata and Relaxation Techniques, *Multimedia Data Management Using Metadata to Integrate and Apply Digital Media*, McGraw-Hill, chapter 6 (1998).
- 11) Day, Y. F., Dagtas, S., Iino, M., Khokhar, A. and Ghafoor, A.: Object-Oriented Conceptual Modeling of Video Data, *Proc. ICDE*, pp. 401-408 (1995).
- 12) Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., Gorkani, M., Hafner, J., Lee, D., Petkovic, D., Steele, D. and Yanker, P.: Query by image and video content: The QBIC system, *IEEE Computer*, Vol. 28, No. 9, pp. 23-32 (1995).
- 13) 堀之口浩征, 黒木進, 牧之内顕文: 時空間データベースインデックス正規化 R*-tree の実装と性能テスト, 情報処理学会論文誌, Vol. 40, No. 3, pp. 1225-1235 (1999).
- 14) Masunaga, Y.: The Block-World Data Model for a Collaborative Virtual Environment, *Proc. WWCA'98*, pp. 309-324 (1998).
- 15) Nabil, M., Shephred, J. and Ngu, A. H.H.: 2D-Projection Interval Relationships: A Symbolic Representation of Spatial Relationships, *Proc. 4th Int. Symp. on Large Spatial Databases*, pp. 292-309 (1995).
- 16) Object Design, I.: *ObjectStore C++ API Reference Release 5.0* (1995).
- 17) Papadias, D. and Sellis, T.: Qualitative Representation of Spatial Knowledge in Two-Dimensional Space, *VLDB Journal*, Vol. 4, No. 3, pp. 479-516 (1994).
- 18) Papadias, D., Theodoridis, Y., Sellis, T. and

- Egenhofer, M. J.: Topological Relations in the World of Minimum Bounding Rectangles: A Study with R-trees, *Proc. SIGMOD Conf.*, pp. 92-103 (1995).
- 19) Sharma, J. and Flewelling, D. M.: Inferences from Combined Knowledge about Topology and Directions,, *Proc. 4th Int. Symp. on Large Spatial Databases*, pp. 279-291 (1995).
- 20) Smith, J. R. and Chang, S.-F.: VisualSEEK: A fully automated content-based image query system, *Proc. ACM International Conference on Multimedia*, pp. 87-93 (1996).
- 21) Tagashira, T., Amagasa, T., Aritsugi, M. and Kanamori, Y.: Interval-Based Representation of Spatio-Temporal Concepts, *Proc. 9th Conf. on Advanced Information Systems Engineering (CAiSE*97)*, LNCS, No. 1250, pp. 231-244 (1997).
- 22) Vazirgiannis, M., Theodoridis, Y. and Sellis, T.: Spatio-temporal Composition and Indexing for Large Multimedia Applications,, *Multimedia Systems*, Vol. 6, pp. 284-298 (1998).

付録 時区間クラスライブラリの概要

時区間クラスライブラリ³⁾は時区間概念モデル⁴⁾に基づいており, `TMInterval`, `CollectionTM`, `CompositeTM`の三つのクラスから構成される. それぞれのクラスには, モデルで定義される演算が全てメソッドおよび演算子として実装されるので, モデルで表現可能な問合せは全てこのクラスライブラリでも表現可能である.

A.1 `TMInterval` クラス

`TMInterval` クラスは時区間概念モデルの時区間に対応するクラスである. `TMInterval` は属性として開始時刻, 終了時刻, およびその時間に有効であった事象のIDの集合を持つ. 図16に `TMInterval` クラスのクラス定義を示す.

A.2 `CollectionTM` Class

`CollectionTM` クラスは時区間概念モデルの時区間の収集に対応するクラスであり, 時間順に整列した時区間の系列から構成される. 図17に `CollectionTM` クラスのクラス定義を示す.

A.3 `CompositeTM` Class

`CompositeTM` クラスは時区間概念モデルの複合時区間に対応するクラスである. 複合時区間とは時区間の収集の特殊な場合であり, ある特定の事象に関する履歴全体を表現することができる. 図18に `CompositeTM` クラスのクラス定義を示す.

```
class TMInterval {
private:
    int start; // start time
    int end;   // end time
public:
    TMInterval(); // constructor
    ~TMInterval(); // destructor
    virtual os_Set<int> event(); // get set of event
    int startTime(); // get start time
    int endTime(); // get end time
    int duration(); // get duration
    TMInterval* real(); // return this if real-time itvl
    TMInterval* null(); // return this if null-time itvl
    TMInterval* include(os_Set<int>); // return this if it contains
                                // given set of events
    TMInterval* intersection(TMInterval*);
    TMCollection* cover(TMInterval*);
    IntervalRelation relation(TMInterval*);
                                // Allen's interval relation
};
```

図16 `TMInterval` クラス

Fig. 16 `TMInterval` class

```
class CollectionTM {
private:
    os_Set<TMInterval*> collection;
public:
    CollectionTM(); // constructor
    ~CollectionTM(); // destructor
    int cardinality(); // # of intervals
    TMInterval* select(int n); // select n-th interval
    os_Set<int> startTime(); // apply TMInterval::startTime() to each
                            // element and return the results as a
                            // collection
    // All methods defined in Interval class are
    // re-defined like startTime method.
    ...
};
```

図17 `CollectionTM` クラス

Fig. 17 `CollectionTM` class

```
class CompositeTM: public CollectionTM {
public:
    CompositeTM() {} // constructor
    ~CompositeTM() {} // destructor
    void append(TMInterval&); // append a time interval
};
```

図18 `CompositeTM` クラス

Fig. 18 `CompositeTM` class

(平成11年3月20日受付)

(平成11年6月27日採録)

(担当編集委員 宝珍 輝尚)

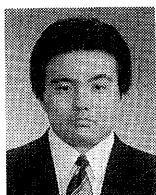


天笠 俊之 (正会員)

1994年群馬大・工・情報工卒.

1999年同大大学院博士後期課程修了. 博士(工学). 1999年から奈良先端大・情報科学研究科助手. 時間データベースに興味を持つ. 電子情報通信学会会員.

報通信学会会員.

**有次 正義 (正会員)**

1991年九大・工・情報工卒。1996年同大大学院博士後期課程了。1994-1995文部省派遣留学生としてMcGill大(院)在籍。博士(工学)。1996年から群馬大・工・情報工助手。並列分散データベース、マルチメディアデータベース等に興味を持つ。電子情報通信学会会員。

**金森 吉成 (正会員)**

1969年東北大大学院博士課程(電子工学専攻)修了。工学博士。東北大電気通信研究所助手、同助教授、仙台電波高専教授を経て、1990年群馬大・工・情報工教授。オブジェクト指向データベース、マルチメディアデータベースに関する研究に従事。ACM, IEEE-CS, 電子情報通信学会各会員。

**田中 貴之 (学生会員)**

1998年群馬大・工・情報工卒。同大大学院博士前期課程(情報工学専攻)在学中。
