

7R-3

動画像中のオブジェクトに注目した データモデルと問合せ処理

石川正敏 高倉弘喜 植村俊亮

奈良先端科学技術大学院大学 情報科学研究科

1. はじめに

動画像データの増加に伴い、利用者の様々な要求に対応できる動画像データベースへの要求が高まっている。このため、Allenの時区間代数¹⁾などに基づいて動画像の時間関連を記述する動画像データベースが提案されている^{2), 3), 4)}。また、動画像中のオブジェクトの位置などに注目した動画像データベースも提案されている^{5), 6)}。しかし、これらは、動画像を2次元平面の静止画の列と考えている。このため、動画像の奥行き情報を保持できず、空間関連を十分に表現できない場合がある。

本稿では、予め3次元の位置情報が得られる3次元CGアニメーションなどを対象にし、空間関連記述法およびデータモデルを提案する。また、空間関連を用いた問合せについて考察する。

2. 空間関連記述法と動画像データモデル

2.1 オブジェクトの空間表現

まず、2次元のオブジェクトの空間表現について述べる。一般に、2次元のオブジェクトに対し空間関連の判定を容易にするためにオブジェクトをそれに外接する最小の面積の長方形を用いて近似する。この近似手法を最小領域矩形 (Minimum Bounding Rectangles) 法という。この手法は、動画像のオブジェクトの空間表現として幾つかの研究で用いられている^{5), 6)}。

本稿では、2次元の空間表現を3次元のオブジェクトの空間表現に拡張するため、オブジェクトに外接する最小の体積の直方体を用いて3次元のオブジェクトを近似する。

さらに、空間表現では、直方体の他に、重心と向きの属性を用いて表現する。オブジェクトの重心は、オブジェクトを近似する直方体の重心とする。重心は、大まかな空間関連の判定やオブジェクト間の距離の計算に利用する⁶⁾。向きは一意に定義する。例えば、人の場合、人が立っている時は、体の前面の方向を向きとして定義する。向きは、オブジェクトから見た空間関連の定義に利用できると考える。

オブジェクトの近似モデルの例を図1に示す。図1の点A, Bは、オブジェクトを近似する直方体を表す代表点であり、点Cは、重心、ベクトルDは向きベクトルを表す。

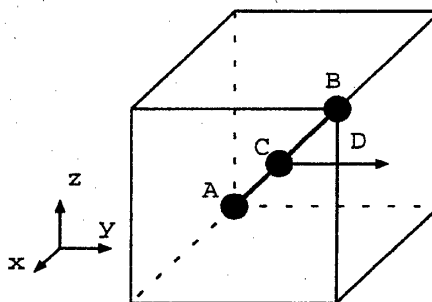


図1 オブジェクトの近似モデルの例

2.2 空間関連

3次元空間は、 x 軸、 y 軸、 z 軸の3つの軸で表現できる。

まず、軸が一つの場合についての関連を考える。

Allenは、一つの時間軸上の2つの時区間の関連を *equal*, *meets*, *before* などの13の関連で表現できることを示している¹⁾。1次元空間上の2つの区間についても、時間軸と形式が同じであるのでAllenの時区間代数を用いることができる。

Liらは、1次元の空間関連を x 軸、 y 軸のそれぞれに適用し、その組合せによって2次元の空間関連を定義している⁶⁾。Liらは次のような関連を定義した。

- ・方向関連 *north*, *southeast* などの8関連
- ・位置関連 *above*, *below*, *left*, *right* の4関連
- ・位相関連 *inside*, *cover* などの6関連

本稿では、Liらの手法を3次元に拡張する。つまり、1次元の空間関連を x 軸、 y 軸、 z 軸それぞれに適用して、その組合せによって3次元の空間関連を定義する。Liらの定義した位置関連の拡張を行ない、次のように定義する(A, Bは、オブジェクト)。

- ・*above* A *above* B (Aの上にBが存在する。)
- ・*below* A *below* B (Aの下にBが存在する。)
- ・*right* A *right* B (Aの右にBが存在する。)
- ・*left* A *left* B (Aの左にBが存在する。)
- ・*front* A *front* B (Aの前にBが存在する。)
- ・*back* A *back* B (Aの後ろにBが存在する。)

さらに、オブジェクトの向きの利用により、周囲の環境の影響を受けずに空間関連を効率的に表現できる。

例えば、オブジェクトA, Bが、並んで写っている動画を考える。A, Bをカメラで前から写すのと後ろから写すのでは空間関連が入れ替わる。しかし、オブジェクトの向きを導入することにより、オブジェクトに基づいて空間関連を記述できるので、カメラの向き

による影響を除去できる。

2.3 動画像データモデル

Liらは、動画像データモデルとして共通動画オブジェクト木モデル (Common Video Object Tree model) を提案している⁶⁾。このモデルは、クリップ間のすべてに共通するオブジェクトを発見することや、オブジェクトに従ってクリップをまとめることを目的としたものである。木構造は、クリップまとまりを表現するのに用いられる。

動画像データモデルで木構造を採用することにより、動画像内のオブジェクトを効率的に検索できる。

Liらのデータモデルでは、クリップをいくつかのオブジェクトをまとめた構成要素として定義している。それに対し、本稿の動画像データモデルのクリップは、対象とするオブジェクトを一つに限定した構成要素である。これにより、動画像中のオブジェクト間の時間関連を効率的に検索できる。

データモデルの構造は、次のような構成要素をもつ木構造を考える (図2)。

クリップ 対象とするオブジェクトが登場するフレームの集まり
 ノード 関連するクリップもしくはノードを集めたもの (再帰表現が可能)
 ルート クリップもしくはノードをすべて集めたもの

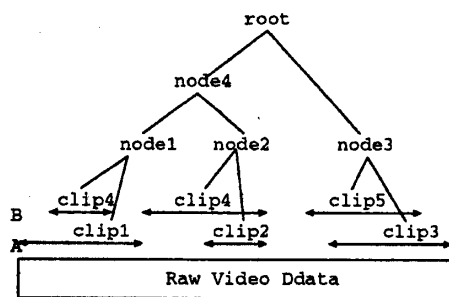


図2 動画像モデル

このモデルの木構造は、動画像からオブジェクト毎にクリップを抽出し、関連するクリップをノードにまとめ、最終的にルートに集めるようにした構造である。

3. 問合せ

本稿で提案する空間関連記述を用いた問合せについて例を用いて考察する。

例1 “Aの右にオブジェクトBが並んでいる映像を表示する。”

従来の空間関連は、オブジェクトの向きを考慮しないので、カメラの向きによって、同じ意味の映像であっても異なる関連になる。従って、目的の映像を取り出すためには、問合せに合致すると考えられるすべての関連を記さなければならない。これに対し、本稿で提案するモデルは、オブジェクトの向きも扱うので、カメラの向きに関係なく同じ意味の映像は同じ関連で表現でき、問合せにおいても容易に関連を記述できる。

例2 “オブジェクトAが振り向く映像を表示する”

従来の空間表現は、オブジェクトの向きを考慮しないので、向きの変化を伴う動作を表現できない。本稿の空間表現は、向きを代数的に表現し、向きの変化を伴う動作は、向きの代数的な時間変化をAllenの時区間代数を用いて表現できる。従って、向きの変化に関する問合せも効率的に処理できる。

4. 今後の課題

極端に大きさの異なるオブジェクトの空間関連の表現は、利用者の認識と大きくかけ離れたものなる。例えば、人とビルの空間関連を記述する場合、利用者は「ビルの前に人が立っている」と認識するが、ビルの重心が人よりはるか上方にあるため、空間関連の記述では「人はビルの下前の方に立っている」になる。従って、今回提案したオブジェクトの空間表現法の改良を考えなければならない。また、最後に動画像への問合せをさらに効率良く記述するため、動画像の時間関連も統合して扱える手法が必要である。

謝辞: 植村研究室の皆様には、本論文を作成するにあたり数々の御助言を頂き、感謝いたします。

参考文献

- 1) J. F. Allen: "Maintaining Knowledge about Temporal Intervals", *Comm. of the ACM*, Nov. 1983, pp832-843
- 2) 堀内 優希 他: “映像データベースのための論理データモデルとその実装”, 電子情報通信学会 第6回 DEWS'95 論文集, Mar. 1995, pp79-86
- 3) 小川 政行 他: “圧縮映像データベースシステムにおける映像演算と実現手法”, 情報処理学会第106回データベースシステム研究会研究報告, Jan. 1996, pp1-8
- 4) 増水 良文 他: “マルチメディアオブジェクト間の時間的関連記述の一フレームワーク”, 電子情報通信学会論文誌 D-II, Apr. 1996, pp492-501
- 5) Y.F.Day, et al.: "Spatio-Temporal Modeling of Video Data for On-Line Object-Oriented Query Processing", *Proceeding of the International Conference on Multimedia Computing and Systems*, May 1995, pp98-105
- 6) J.Z.Li, et al.: "Modeling Video Spatial Relationships in an Object Model", *University of Alberta Technical Report TR 96-06*, Mar. 1996