

ビデオ管理システムVOMのビデオに対する索引付け

2Q-8

佐藤通夫 伊藤秀昭 伊藤誠 福村晃夫

中京大学大学院情報科学研究科

1. はじめに

近年、ビデオデータに対する索引付け、検索および編集を支援するために供するマルチメディアデータベースシステムに対する要求が高まっている。この要求に応えるために、我々はドラマを対象として、ビデオ管理システム VOM (Video Object Management System) の開発を進めている [1]。

本システムでは、1つのビデオを複数のオブジェクトに分割して、オブジェクトから成る階層によりビデオを表現する。また、オブジェクトに対してキーワードを付与することによって索引付けが行われる。このとき、階層の上位オブジェクトのキーワードが下位オブジェクトのキーワードから求められ、下位オブジェクトに付与されたキーワードの集まりを代表するキーワードが上位オブジェクトに付与される必要がある。

本論文では、ビデオ管理システム VOM のビデオに対する索引付けについて述べる。

なお、本システムは、シリコングラフィックス社の O₂ により稼働し、C を用いて記述されている。

2. VOM の概要

VOM の主たる機能は構造化機能、編集機能および問い合わせ機能である。

構造化機能はオブジェクトによって1つのビデオを階層構造により記述することを支援する機能である。編集機能は構造化されたビデオを合成したり、合成されたビデオを編集する機能である。問い合わせ機能はオブジェクトの集まりを検索する機能である。

ビデオを表現するためのオブジェクトには、物理オブジェクト、シーンおよび論理オブジェクトの3種がある。物理オブジェクトは、動画を管理するための最小単位となるオブジェクトである。VOM はビデオを構成する2つのフレーム間の色に関する相関を用いて物理オブジェクトの切り出しを支援しており、利用者によって物理オブジェクトにはキーワードが付与される。シーンは下位オブジェクトとなる連続する物理オブジェクトを管理するオブジェクトである。論理オブジェクトは、下位オブジェクトとなるシーン、または他の論理オブジェクトを管理するオブジェクトである。

3. キーワードの受け渡し

キーワードの受け渡しは、下位オブジェクトのキーワードの集まりより半自動的に上位オブジェクトのためのキーワードを選択することである。受け渡しの方法として、和集合、積集合および特異値分解 [2] による3つの方法を備えている [1]。

特異値分解は、行に用語、列に文献、その要素が出現頻度である出現頻度行列を、用語行列、特異値行列および文献行列の3つに分解する演算である。VOM では用語がキーワード、文献がオブジェクトに相当する。

下位オブジェクトのキーワードは、上位オブジェクトであるシーンまたは論理オブジェクトを生成するときに、上位オブジェクトに受け渡される。利用者は和集合、積集合、特異値分解による方法の中からキーワードの受け渡し方法を選択する。和集合、積集合による方法が選択されれば、下位オブジェクトのキーワードの和集合、または積集合が上位オブジェクトに受け渡される。一方、特異値分解による方法が選択されれば、特異値分解によりキーワード間、またはオブジェクト間の距離が求められる。利用者はどちらの距離を利用してキーワードを受け渡すのかを選択する。選択された距離を利用してキーワードの受け渡しが支援される。

3.1 キーワード間の関係

キーワード間の関係によるキーワードの受け渡しでは、まず、下位に位置するオブジェクト群のキーワードを特異値分解によってキーワードや、オブジェクトから成る空間におけるキーワード間の距離を求める。次に、キーワード間の距離が近いキーワードの集まりであるクラスタを構成する。さらに、求めたクラスタに対して1つまたは複数の代表的なキーワードを付与する。このキーワードを上位オブジェクトのキーワードとして下位オブジェクトから上位オブジェクトへ受け渡す。

特異値分解により求められたキーワード間の距離に基づいて得られる空間の一部を表示するための画面を図1に示す。この図は、約1分間の物理オブジェクト(10個)のキーワード(22個)を特異値分解によって求められたキーワード間の距離に基づいた2次元上でのキーワードの位置を表現している。図中の四角の点と丸の点はキーワードを表しており、楕円中の4つの四角の点は1つのクラスタの構成要素である。四角の点と丸の点は、それぞれクラスタ1に属するキーワード、およびクラスタ1に属さない

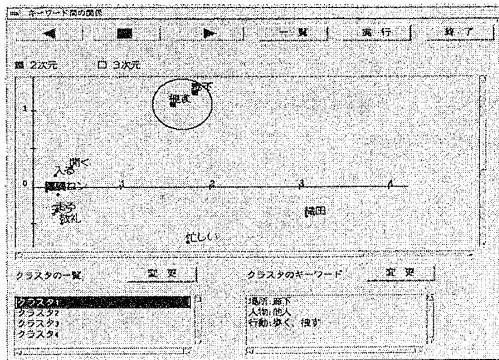


図1 特異値分解によるキーワード間の関係

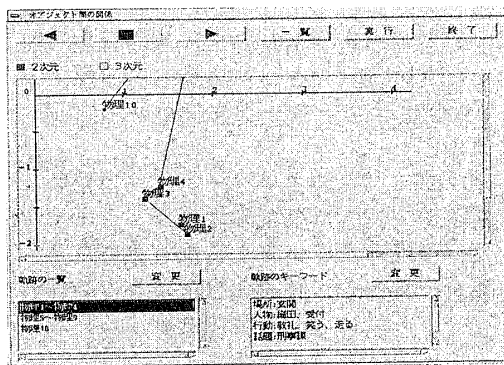


図2 特異値分解によるオブジェクト間の関係

キーワードを表している。

3.2 オブジェクト間の関係

オブジェクト間の関係によるキーワードの受け渡しでは、まず、下位に位置するオブジェクトを特異値分解に基づいて空間上におけるオブジェクトの位置を計算する。次に、各々のオブジェクトをオブジェクトが再生される時間順序に従いリンクを設定する。このリンクとリンクにより結合されたオブジェクトから成るグラフを軌跡と呼ぶ。2つのオブジェクトを結合するリンクの長さが長い場合には、リンクを切断して、得られている軌跡をいくつかの軌跡に分割する。さらに、分割された軌跡に対して1つまたは複数のキーワードを付与する。このキーワードを上位オブジェクトのキーワードとして下位オブジェクトから上位オブジェクトに受け渡す。

特異値分解によって得られた軌跡の一部を表示するための画面を図2に示す。この図は、約1分間の物理オブジェクト(10個)を特異値分解によって得られた2次元上での軌跡を表現している。図中の四角の点と丸の点は、オブジェクトを表している。点を結ぶ線は、オブジェクトの時間順序を表すリンクである。四角の点を線で結んだオブジェクト群は、1つの軌跡を表している。

4. 実験

連続する物理オブジェクトから1つのシーンを生

表1 精度と再現率の平均値

	平均精度	平均再現率
和集合	0.21	0.57
積集合	0.38	0.16
キーワード間の関係	0.40	0.61
オブジェクト間の関係	0.64	0.79

成するときのキーワードの受け渡し例について実験した。キーワードの受け渡し方法として、和集合による方法、積集合による方法、特異分解を用いたキーワード間の関係による方法、特異値分解を用いたオブジェクト間の関係による方法の4つの方法を用いている。これらの4つの方法を用いて、20件のビデオデータに対してキーワードの受け渡しを試みた。キーワードの適切さを計るために、シーンに付与されたキーワードの精度と再現率を以下に示す式により求める。

$$\text{精度} = \frac{\text{システムが得たキーワードのうち適切であるキーワード数}}{\text{システムが得たキーワード数}}$$

$$\text{再現率} = \frac{\text{システムが得たキーワードのうち適切であるキーワード数}}{\text{利用者により付与されているキーワード数}}$$

20件の生成されたシーンのキーワードに対する精度と再現率の平均値を表1に示す。特異値分解を用いたキーワード間の関係とオブジェクト間の関係により求められたキーワードの精度と再現率は、和集合、積集合により得られたキーワードと比較して高い。これは、和集合、積集合による方法で得られなかったキーワードが、特異値分解を用いることによってクラスタや軌跡に新たなキーワードが付与されるためである。

5. まとめ

本論文では、特異値分解によりキーワードやオブジェクト間の関係を用いてキーワードの受け渡し方法について述べた。今後の課題は、利用者によるキーワード付与負担の削減のためのキーワードの取捨選択のメカニズムを開発することである。

参考文献

- [1] 佐藤, 伊藤, 伊藤, 福村: ビデオ管理システム VOM におけるキーワードの構造化, 人工知能学会全国大会論文集, pp.199-200, 1999.
- [2] M.W.Berry, et al; Computational Methods for Intelligent Information Access, Proceedings of SuperComputing's 95, San Diego, CA, 1995.
- [3] V.Kobla, et al; Video Trails: Representing and Visualizing Structure in Video Sequences, Proc. of ACM Multimedia, pp.335-356, 1997.