

内容の関連性にもとづいたビデオ映像の記述

3 Q-9

是津耕司† 上原邦昭†† 田中克己††† 木邑信夫†

†通信・放送機構 神戸リサーチセンター

††神戸大学工学部情報知能工学科

†††神戸大学大学院自然科学研究科知能科学専攻

1. はじめに

データベース化することを目的としてビデオ映像に対しインデックスづけを行なう際、これまでは主にシーンの情景やストーリーにもとづいた映像記述を利用していた。しかし、シーンがめまぐるしく変化してストーリー性が乏しい映像、例えば予告編映像などは、こうした従来の方法による映像の記述が難しくデータベース化することは困難である。そこで本研究では、これまでのアプローチとは異なり、映像の内容に関する断片的な記述によりビデオ映像をデータベース化すること目指し、これが予告変映像のようなこれまでは扱うことが難しかった映像もうまく扱えることを示す。

2. 内容の関連性にもとづいた記述

これまでに提案されたビデオ映像記述方式は、キーワード [1] やレコード [2] あるいはグラフ構造 [3][4] により特定の時間区間に上映されるシーンの情景やストーリーを記述し、これを時間軸上に並べることで映像全体を記述していた。これらの方法では、求めるシーンを検索するために各時間区間における映像シーンを正確に記述することが要求されるが、これは非常に困難な作業である。

困難と感じるのは、記述者である我々自身が同じ映像シーンに対してさまざまな印象や感想を持ち、それらを拠り所に映像に対する説明を与えようとするからである。我々が映像を理解する際には、その前に流れた映像との因果関係や、論評や宣伝など映像以外から得た情報を使いながら理解することが多い。これまでの記述方法のように、時間で区切られた映像部分の記述を時間軸上に並べるだけでは、こうした我々の映像理解にもとづいた記述を行なうことは難しい。

A video description method based on contents association
Kohji Zettsu†, Kuniaki Uehara††, Katsumi Tanaka†††, and
Nobuo Kimura†

†Kobe Research Center, TAO

††Department of Computer and Systems Engineering, Faculty of Engineering, Kobe University

†††Division of Intelligence Science, Graduate School of Science and Technology, Kobe University

そこで、厳密なシーンの説明によって映像を記述するのではなく、内容の意味的なまとまりやつながりをもとに映像を記述することを考える。すなわち、映像の内容に関するいくつかの断片的な表現とこれらの間にある関連性に着目して映像を記述する。このような内容の関連性にもとづいた映像記述は、我々人間の映像理解により近い記述方法と言える。

3. 時刻印付ノードリンクグラフによるビデオ映像の記述と検索

本研究では、内容の関連性にもとづいたビデオ映像の記述として「時刻印付ノードリンクグラフ」[5] による映像記述を行なっている。時刻印付ノードリンクグラフでは、記述者が映像を見て思いついた印象や感想から得られるキーワードをノードとし、キーワード間の意味的な関連性の有無をリンクで表したグラフ構造でビデオ映像を記述する。キーワードにはそれが指し示す映像部分に対応した時刻（タイムコード）が付けられる。映像の検索では、中に含まれるキーワードとリンクを使って検索要求として与えられた文（検索文）を表現できるような最小の部分グラフを見つけ出し、この部分グラフに含まれるノードに対応する映像（カット）を選び出す。時刻印付ノードリンクグラフによるビデオ映像の記述と検索の例を図1に示す。この例から、内容の関連性にもとづいた映像の記述および検索が可能であることが示される。

時刻印付ノードリンクグラフにより、時間的に離れていても意味的にはつながっている映像間の関係をリンクによって表現することができる。従来の記述方法のように各シーンに対する記述を時間軸に沿って並べて記述するだけでは、このような関係を表現することはできない。また、グラフの連結性に着目してグラフをまとめると、映像の意味的なまとまりを見い出すことができる。

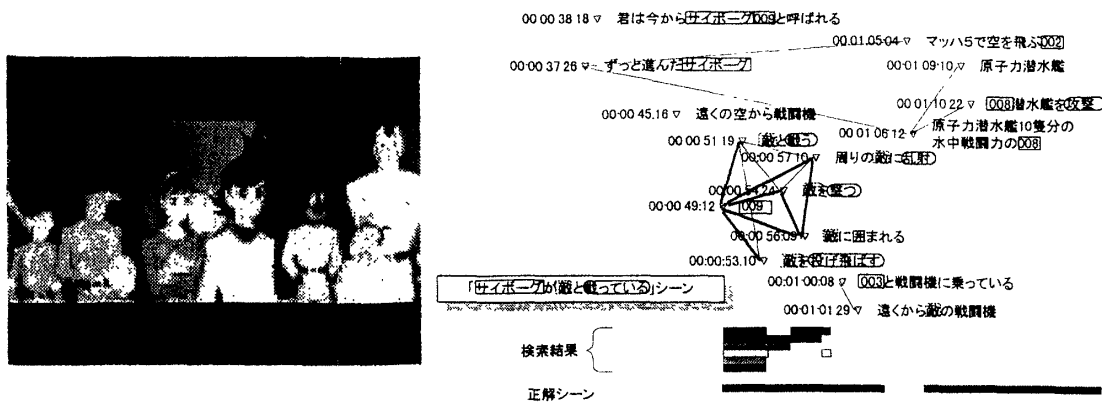


図1: 時刻印付ノードリンクグラフによるビデオ映像の記述と検索

東映アニメ予告編映像「サイボーグ009」(1966年東映動画スタジオ製作、3分18秒)の一部を記述し、この記述をもとに「サイボーグが敵と戦っている」シーンを検索した結果を表す。検索の結果選出されたカットは「検索結果」の欄に対応する時間区間として示す。また検索要求に対し正解と思われるシーンを「正解」欄に併記しておく。

4. 今後の課題

内容の断片的な表現とそれらの間の関連性にもとづいた映像記述では、記述者の映像理解に合わせた柔軟な記述が可能であるが、その反面、記述に許した曖昧さのために生じる問題もある。

図1に示した時刻印付ノードリンクグラフによる検索の例では、検索結果としてとびとびの映像が選ばれてしまったものがあるが、これはキーワードと映像の対応づけが曖昧であるために、記述グラフに対応する適切な映像を選択することができなかったためである。キーワードに付けられた時刻はそのキーワードを「思いついた」時刻であり、キーワードにはこの時刻より前に上映された映像に対する内容が含まれている可能性がある。そのため、従来の記述方法のようにキーワードと特定映像部分を単純に対応づけることができない。したがって、記述に対応する映像を選び出す際に適切な映像を使って補間し、無意味な映像が選出されることを防ぐ必要がある。

また、記述者の自由な表現で書き出したキーワードの中には、主語や目的語、動詞などが抜けているために内容的に不完全なものもある。こうしたキーワードが多く含まれていると、記述内容が全般的に隙間の多い(sparseな)ものになってしまう、適切な関連を表現できないために検索に悪影響を及ぼす恐れがある。そこで、キーワード間の関連をたどって欠けた言葉を相互に補間しながらキーワードを集約していくことで、全体の記述内容をより密度の濃いものにしていくことが考えられる。

5. おわりに

今回は、映像の内容に関する断片的な記述とこれらの関連にもとづいて映像を記述および検索することについて述べ、この記述方法により予告編映像のようなストーリー性の薄いビデオ映像もうまく扱うことができることを示した。今後はビデオ映像のデータベース化を目指し、本記述方法をより有用なものへと発展させていく。

謝辞

本研究のために貴重な映像資料を提供していただいた、東映株式会社様に感謝致します。本研究は、一部、文部省科学研究費重点領域研究(課題番号08244103)による。

参考文献

- [1] 柴田 正啓: 映像の内容記述モデルとその映像構造化への応用, 電子情報通信学会論文誌 D-II, Vol.J78-D-II, No.5, pp754-764 (1995)
- [2] E.Oomoto, K.Tanaka: OVID: Design and Implementation of a Video-Object Database System, IEEE Transactions on Knowledge and Data Engineering, Vol.5 No.4, pages 629-643 (1993)
- [3] 前原 恵太, 福田 慶朗, 上原 邦昭: グラフ表現による動画像内容記述からの階層的場面構造の構築, 情報処理学会研究会報告, 96-DBS-109, Vol.96 No.68 (1996)
- [4] 柴田 正啓: デスクトップ映像製作, ADBS '96 (1996)
- [5] 是津耕司, 上原邦昭, 田中克己, 木邑信夫: 時刻印付ノードリンクグラフによるビデオ映像のデータベース化, 情報処理学会研究会報告, 97-DBS-111 (1997)