

## Video Browser 技術を用いた動画像検索システムの試作

4 Z - 6

本田文隆\*<sup>1</sup> 谷本正幸\*<sup>1\*2</sup> 田坂修二\*<sup>1\*3</sup> 近藤康二\*<sup>1\*4</sup>

\*<sup>1</sup>通信・放送機構 \*<sup>2</sup>名古屋大学大学院 \*<sup>3</sup>名古屋工業大学 \*<sup>4</sup>ソニー株式会社

### 1 はじめに

文字を対象とした情報検索技術はすでに実用されているが、映像を対象とした検索技術は実用例が少なく、現在もさまざまな研究が行われている[1]。例えば、動画サーバとデータベースサーバが組み合わされた VOD (Video On Demand) システムでは、タイトル名などを対象としたキーワード検索により、コンテンツの検索が可能である。しかし、必ずしもタイトル名がコンテンツ全体の内容を適切に表現しているとは限らず、実際に再生してみないと内容を確認できないのが現状である。

そこで筆者らは、動画像の概略を静止画像として一覧表示する Video Browser 技術[2]を利用した検索システムについて検討を行った。本稿は、その試作システムの報告である。

なお、本研究は、通信・放送機構岡崎リサーチセンターで行われている「異種ネットワーク間高速接続技術研究開発」の一環であり、異種動画サーバへの透過的アクセス機能[3]を実現した後の課題として、動画像検索手法の研究に取り組んでいるものである。

### 2 動画像検索手法

動画像の検索手法に関しては、画像認識技術やシーン抽出技術を用いたものなど、さまざまな観点から研究が行われている。本システムでは、動画像の一覧性に観点をおき、検索手法として Video Browser 技術を利用した。Video Browser 技術は、移動スリット窓によるサンプリングを用いて動画像情報を圧縮表示することにより、動画像の概略を正確な時間軸のもとに静止画像として一覧表示する技術である(図 1)。

本システムは、動画像の概略を示すインデックスの

生成技術と表示技術により実現される。

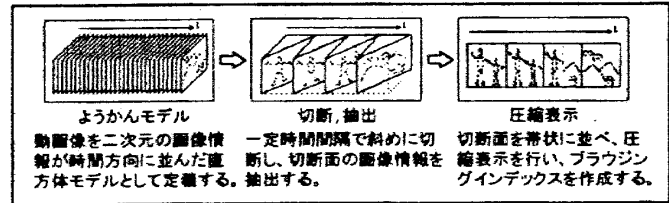


図 1 Video Browser 技術概念図

#### 2.1 インデックス生成技術

Video Browser 技術を利用して生成するインデックス(以下、VD インデックス (Visual Digest Index) と呼ぶ)は、以下の手順で生成する。

- (1) 動画像から静止画像への変換。MPEG ファイルを 1/30 秒ごとに静止画像として抽出する。(以下、この静止画像をフレームと呼ぶ)
- (2) フレームからのライン抽出。各フレームからラインを抽出する。(抽出方法は、縦方向、横方向、部分的、フレーム指定の 4 種類を試作し、対象フレームとライン数は可変とした)
- (3) 抽出した画像の連結。抽出した画像をつなぎあわせ、規定のサイズになったら 1 つのファイル (VD インデックスファイル) として生成する。
- (4) 静止画像の抽出。VD インデックスファイルの中央部分のフレームを抽出する(静止画ファイル)。

生成過程においては、後に VD インデックスを表示することを考慮して、ファイル形式、サイズを適宜変更している。

#### 2.2 インデックス表示技術

VD インデックスは、動画像の長さに比例して、生成される量が増加する。このとき、検索結果としてすべてのインデックスを表示すれば、各動画像の内容も把握しやすいが、全体の検索結果が膨大になり、かえって検索が困難になる。また、インデックス情報は文字情報に比べて情報量が多いため、表示時間も考慮する必要がある。

そこで、VD インデックスを表示する方法として、目的に合わせて以下の2種類の方法を試作した。

- (1) スクロールを用いて1行で表示する方法。
- (2) コンテンツの長さに応じて2行で表示する方法。

(1)の方法は、すべてのVDインデックスを表示するために、両端にスクロールボタンを設置し、必要に応じて表示する方法である。これは、特定のシーンを検索する場合に有効である。(2)の方法は、一覧性を重視し、一目見て動画の内容を把握できるように表示する方法である。この方法では、すべてのVDインデックスの中から表示するインデックスを取捨選択するアルゴリズムを取り入れた。

また、動きの早い映像の場合はVDインデックスでは内容を認識するのが困難であるため、VDインデックスにマウスフォーカスを当てると、その中心部分の静止画像を拡大表示するようにした。さらに、クリックするとそこから動画を再生する機能も実現した。

### 3 システムの概要

本システムは、動画サーバ、WWWサーバ、データベースサーバとクライアントから構成される(図2)。

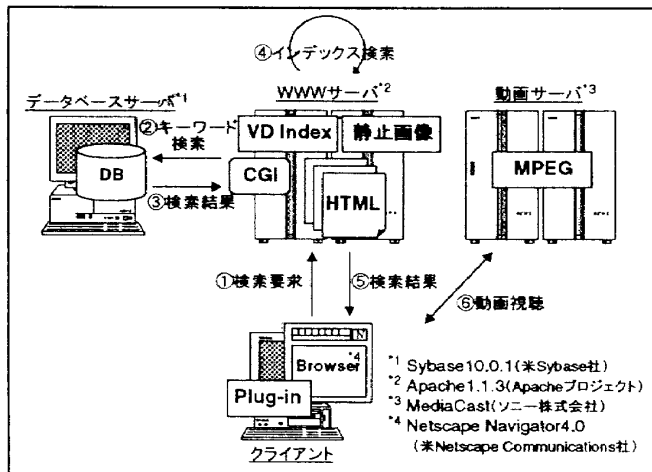


図2 システム構成と処理の流れ

利用者は、WWWブラウザを通じてキーワードを入力すると、検索結果としてタイトル名と、対応するVDインデックスが表示される仕組みになっている。表示された候補をクリックすると、WWWブラウザ内で動画が再生される。なお、本システムで利用した動画サーバには、2分程度の動画を中心に約1,000タイトルが蓄積されている。

### 4 評価

インデックスの生成においては、可視性と生成されるインデックス数を考慮して、1フレームから2ラインずつ縦方向に抽出する方法を採用した。生成には、203MIPSのワークステーションを利用して、動画再生時間の約10倍の処理時間を必要とした。

検索においては、従来の文字情報だけの検索に比べ、VDインデックスのロードのために多少の時間がかかるが、検索結果としてコンテンツの概略が一覧できるため、検索が容易であった(図3)。また、VDインデックスが有効でない部分においても、静止画像と組み合わせることで、シーンの確認が可能であった。

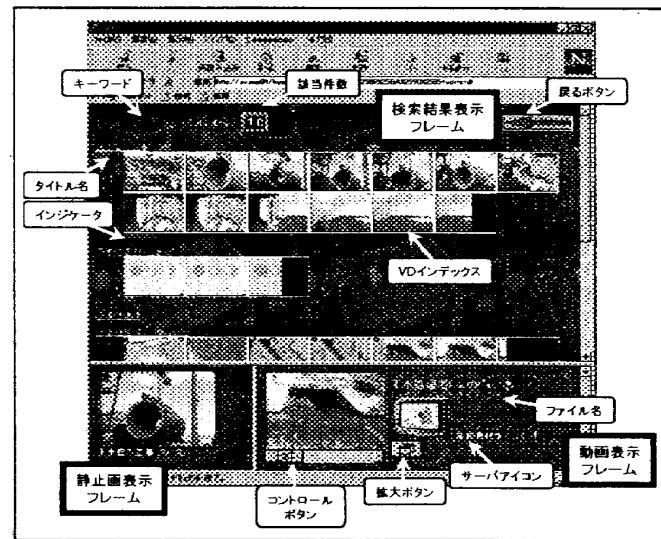


図3 アプリケーション画面

### 5 まとめ

本稿では、Video Browser技術を用いた動画検索システムの試作について述べた。今後は、抽出方法や動画のの違いによる生成されるインデックスの変化と、効果的な表示方法について研究を進める予定である。

### 参考文献

- [1] 美濃: 知的映像メディア検索技術の動向, 人工知能学会誌, Vol11.No1, 1996
- [2] 大場他: 動画一覧表示技術: ビデオブラウザ, テレビジョン学会年次大会, 1991
- [3] 本田他: 異種マルチメディアサーバへの透過的アクセスおよび検索方法の構築, 情報処理学会第55回全国大会論文集, 6G-01, 1997