

WWW サーバ上でのカット点を利用した動画像情報検索・再生システム

川崎 由起子†, 鈴木 麗†, 砂原 秀樹†,

†奈良先端科学技術大学院大学

概要

近年の WWW の普及と次世代ストリーム転送系ネットワークの将来性を考慮し、本研究では WWW ブラウザを利用した大容量の動画像情報の検索・再生システムの構築を行なった。本システムではサーバ側で長時間のストリームからなる動画像情報をコンテンツに基づいたショットに分割し、ショットを代表する静止画像をカット点とし作成する。ユーザはそのカット点を参照することで、目的のショットを効率良く検索・再生できる。WWW ブラウザを利用することによってユーザは動画像情報を時間・場所の制約をうけずに見ることができるだけでなく、目的の動画像情報を取得しそれを編集して、各ユーザのデータベースを作成できる。またユーザ自身の画像からの検索も行なえる。本論文では本学の電子図書館を対象として、提案するシステムのプロトタイプを設計・実装し評価を行なう。

キーワード WWW, 電子図書館, MPEG2, ショット, カット点, 情報検索, 情報編集

The Motion Image Information Retrieval and Browsing System with a Cutting Point Concept for WWW

Yukiko Kawasaki†, Rei Suzuki†, Hideki Sunahara†,

†Nara Institute of Science and Technology

Abstract

In recent years, WWW and next generation stream transmission networks have spread by leaps and bounds. Considering this situation, we suggest to construct a image retrieval and browsing for WWW which can easily and efficiently handle a large size of motion image information using the technique of cutting points. In this system, the server divides any long stream of motion image information into smaller shots by contents, and creates the cutting points at the pictures used as the representation of the shot. Users can efficiency refer and replay the shot of his choice using the cutting points. Moreover user can retrieval his own image. This paper describes a prototype of such a system based on the Digital Library of our institute(NAIST). We will also discuss a evaluation of our system in this paper.

key words WWW, Digital Library, MPEG2, shot, cutting point, information retrieval, information editing

1 はじめに

WWW の普及、次世代ストリーム転送系ネットワークの普及とともにないインターネット上での動画像情報転送の需要が増加すると考えられる。しかし現在の動画像情報転送は、伝送経路媒体の物理的制約によりテキスト情報や静止画像情報に比べ需要及び供給も低い。その上インフラが十分に整っていない環境下で大容量の動画像情報のストリームをネットワーク上に流すことは、ネットワークのトラヒック増加や、ユーザのハードウェア的負荷へつながる。奈良先端科学技術大学院大学 (NAIST) の電子図書館では、ビデオの閲覧機能を提供している。しかし、そこではユーザにとって目的に応じた場面の検索や再生を行なうのが容易ではない。

そこでサーバ側で連続する動画像ストリームをコンテンツに基づいたショットに分割して、ユーザに提供するシステムの構築を行なう。本システムではサーバ側でショットの代表となる静止画像をカット点としてユーザに提示し、ユーザはそのカット点から、目的のショットを効率良く検索・再生できる。さらにユーザ側ではそのショットの動画像情報を簡単に編集でき、独自のデータベースに蓄積ができる。

本稿では、まず現在動画像情報を提供している NAIST の電子図書館の現状と問題点について述べ、その解決法として、コンテンツを元にショットに分割する方式を提案する。次に WWW において統合した支援システムの設計を述べ、一部実装したものを示す。最後に今回の実装を評価し、今後の課題を述べまとめとする。

2 WWW 上での動画像情報

現在インターネット上で個人レベルで大容量の動画情報を維持管理するのは困難であり、多くは企業や各種団体などによるものである。そこで本論文ではこうした大容量の動画情報を取り扱っている一般的な例として、本学の電子図書館をとりあげ、その形態と問題点について述べる。

2.1 NAIST 電子図書館

現在運用中の NAIST の電子図書館システムは資料情報を電子化してデータベースを構築しており、

ユーザはネットワークを経由したアクセスが可能である。検索エンジンは http サーバとして機能しているため、ユーザは web ブラウザによって情報の検索・閲覧が行なえる。またテキスト情報以外にビデオ情報も取り扱っている。ビデオ情報は、そのデータの性質を考慮して独立したシステムとして構成されている。ビデオ情報の取り込みは MPEG2 エンコーダによって行なわれる。現在はビデオ情報を 4Mbps のビットレートで格納している。ビデオサーバのファイルサーバは、180GB の容量を独立して用意されており、これは 4Mbps (NTSC の画質レベル相当) のビットレートの MPEG2 ビデオを 100 時間分格納する容量に相当する。ビデオ再生用のネットワークは 100Mbps の独立した FDDI を使用している。

なお、現在のシステムではビデオ情報を表示するためにハードウェア MPEG2 デコーダを必要とするため、学内に設置された専用の端末でのみ表示が可能である [1]。

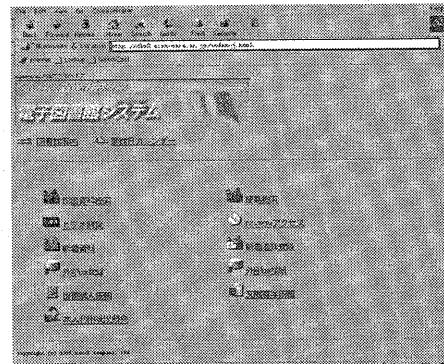


図 1: NAIST の電子図書館

2.2 NAIST 電子図書館での問題点と提案

電子図書館では WWW 上で動画像情報を提供するシステムが増加している [2] [3]。しかし現在の NAIST の電子図書館の動画像情報提供システムは、ユーザにとって検索・再生が容易ではない。

そこで、今回提案するシステムでは以下の 2 点を実現する。

- 検索機能の提供

- ・ブラウズを容易にするユーザインターフェイスの提供

以上の機能を提供するために、ショット分割とカット点作成・提示の概念を導入する。

2.3 「ショット」及び「カット点」

本論文では「コンテンツ」「ショット」及び「カット点」という用語を用いているが、これらの定義を以下に示す。

コンテンツ：単一の話題など、ひとまとまりの情報
 ショット：コンテンツに基づく一連の動画情報単位
 カット点：ショットの内容を表現する代表的静止画

2.3.1 ショット分割

動画像情報は一般に複数のコンテンツの組み合わせで構成されていると言える。例えばニュースは複数の数分のコンテンツで構成されている。また、大学の講義も、各章（あるいは項）といった内容から分割される複数のコンテンツで構成されていると考えられる。^[4] このコンテンツをショットとして分割し、そのショットを提示する。さらに各ショットにインデックスだけでなくそのショットの内容を代表する静止画像であるカット点を付加する。動画像情報を数分のショットに分割した動画情報を提供することで、ユーザは必要とする情報だけを素早く検索できる点で効果的である。また、ジャンプ機能としても使用でき、ビデオディスクよりも早く目的の場面を検索・再生できる。

ショット分割は情報提供者側が多くの情報や知識を持っていることと、ニュースが速報性を重視することを考慮し、手動で行なうものとする。またユーザ側でも任意のショット分割が行なえ、ユーザ自身による独自の編集ができる。

2.3.2 カット点の作成

カット点とは情報提供者側が作成するもので、そのショットのコンテンツを代表している静止画像である。カット点はショットの先頭フレームであると必ずしもいえない。今回対象とするニュースにおいては、ニュースの項目であるテロップが出てきた時点をカット点とする。字幕が出てこないものについ

ては、その内容を表すと考えられる静止画像をカット点とする。ユーザはニュースの代表画像であるカット点をもとに、目的の情報検索を行なう。これによってユーザは効率的にショットを検索・再生できる。

3 設計

3.1 システムの構成図

本のシステムの処理の流れを図2に示す。

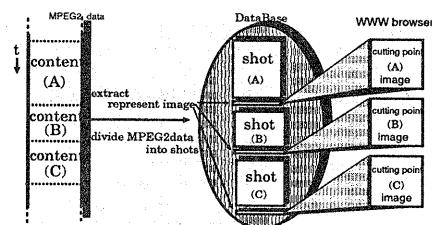


図2: システムの処理の流れ構成図

本システムでは MPEG2 データをコンテンツに基づきショットに分割し、さらにそのショットの代表となる静止画像をカット点とし抽出する。そのショットとカット点をデータベースに蓄積する。WWW ブラウザにはカット点を表示し、カット点をもとにユーザは目的の動画像情報を容易に検索でき、データベースのショットをランダムに再生できる。

今回の設計には、動画像情報として NAIST の電子図書館で取り扱っているニュース情報を取り扱った。

3.1.1 MPEG2 を用いたショット分割

我々の試作システムでは MPEG2 を用いてショット分割を行なう。NAIST の電子図書館では動画像情報伝送方式として MPEG2 を使用している。高压縮の MPEG2 は、蓄積メディア、通信メディア、放送メディアへの適応が考慮され、MPEG1 を越える適用領域の広さと高品質画像を提供しており、次世代通信方式として知られている^[5]。この次世代通信方式である MPEG2 のソフトウェア（エンコーダ／デコーダ）を用いて、動画像情報の圧縮とショット分割を行なう。

MPEG2 は複数枚のフレームを 1 まとまりの単位とした GOP (Group Of Picture) 構造をとっている。

GOP は I,P,B の 3 種類のフレームタイプの集合であり、このうち I フレームはフレーム間予測を使わずに画面の全てを符号化しており、GOP の独立性を保っている。

MPEG2 でショット分割を行なう時、任意のフレームからの分割は MPEG2 のデータと認識できないため再生が行なえない。(図 3(b))

そこで、分割を行なう最初の GOP の先頭に画像フォーマットを指定するシーケンスヘッダを付加し、MPEG2 のデータと認識させ再生を可能とする(図 3(c))。

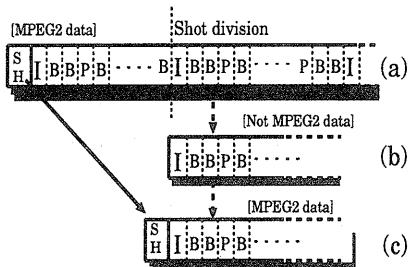


図 3: MPEG2 によるショット分割

3.2 データベース蓄積システム

データベース蓄積において、動画像は NTSC(現行テレビ)の伝送速度 100Mbps を 25 分の 1 に圧縮する MPEG2 を用い、各ショットの検索用として代表静止画であるカット点を蓄積する。

3.3 カット点による検索システム

カット点画像のデータベース検索において本システムでは、パターンマッチングの際 kl 展開に基づく固有空間法を使用している。kl 展開は、画像信号の相関関数(固有空間)を用い、入力画像に対する 2 乗誤差の期待値を最小とする直交固有関数を求めて変換するものである。これによって増大なデータ量をもつ動画像に対し、低損失の高压縮が可能でかつ高速なパターンマッチングが行なえる。一方短所としてデータベースに画像が追加、削除されるごとに固有空間を作成し直さなければならないという点がある [6]。

またテキスト検索と異なる特徴として、静止画像による検索ではユーザは自分のもつ画像からサーバ

側のデータベースのカット点の検索もできることができられる。

4 インターフェース

上記のシステムのインターフェースとして WWW ブラウザを利用した。

4.1 WWW ブラウザの利用

現在 WWW は企業や学術機関で広く使用されており、一般家庭にも普及している。WWW の特徴として目的の情報が得られるというのに加え、時間や場所の制約を受けず情報が得られ、さらに双方発信が可能である。

この特徴から、例えばテキスト情報より情報量の多いニュースなどの動画像情報を、時間や場所の制約なく参照できる。また、その双方向性から情報提供者側から一方向に情報を与えるだけでなく、ユーザからのフィードバックを組み込むことが可能となる。さらにユーザ間でも情報交換ができるという利点がある。

以上の WWW の特徴と電子図書館への適応性を考慮して WWW ブラウザをインターフェースとして採用した。

5 実装

実装は BSD/OS 2.0.1 の環境で行なった。MPEG2 のソフトウェアデコーダである mpeg2play を利用し [7] 必要な機能を追加するためのコーディングは C 言語で行ない、インターフェース部分として Xwindow のライブラリを使用した。画像データの再生・編集は WWW ブラウザ上からデータをローカルディスクに読み込んで行なう。

5.1 データベースの作成

5.1.1 ショット分割とカット点の作成

MPEG2 で圧縮した動画像情報のニュースをショットに分割し、カット点を決める。MPEG2 を再生中に、ショットの先頭・カット点・最後の場面でクリックする。それによってその場面の静止画像とタイム

コード、フレーム情報を切り出す。カット点はこのフレーム情報を使用する。(図 4)



図 4: ショット分割: 再生動画像(左上)をクリックするとその場面の静止画像が作成される(他3枚)

図 4 の情報を元にショットの先頭と最後のタイムコード、入力 MPEG2 動画像ファイル名と出力 MPEG2 動画像ファイル名を指定し、MPEG2 形式のショットを作成する(図 5)。

TIMECODE (Start)	0009c680
TIMECODE (End)	0009e000
File name (Input)	HTS.m2v
File name (Edited mpeg2)	index.m2v
<input type="button" value="OK"/>	

図 5: ショット作成例

またこのシステムをユーザ側でも使用することによって、取得したショットの編集ができる。

5.2 表示と検索機能

情報提供者側は WWW ブラウザ上でニュースの各ショットをカット点を用いて提示する。検索用の静止画像は 150x100 ピクセルの大きさである。(図 6) ユーザが目的のショットからのカット点をクリックすると、ショットが再生される。動画像再生は 324x240 ピクセルの大きさである。

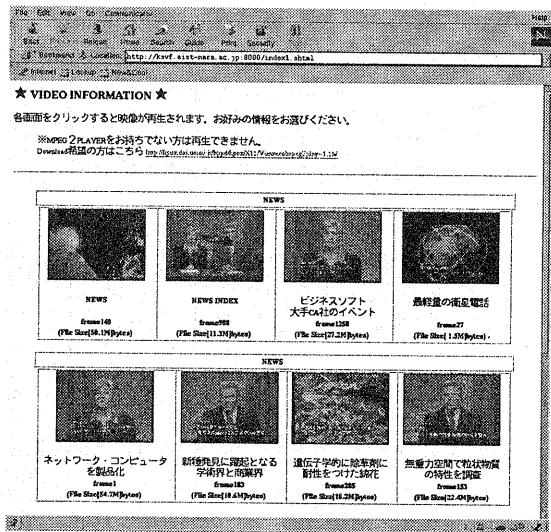


図 6: WWW ブラウザ

5.2.1 カット点間のマッチング

ユーザはカット点間の類似画像のマッチングを行なえ、そのマッチング結果を(3枚)表示する(図 8)。

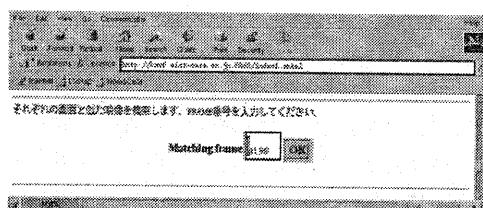


図 7: カット点間のマッチングのための WWW インターフェース

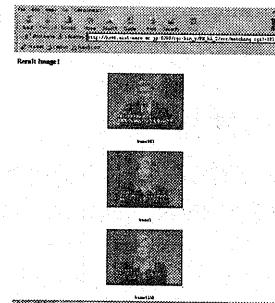


図 8: マッチングの結果

5.2.2 入力画像からのマッチング

静止画像による検索を利用し、ユーザ所有の画像と類似したカット点の検出も可能である。検索のキーにしたいユーザ側の画像を URL を指定することで入力可能とし、その画像からの検索をできようとした(図 9)。

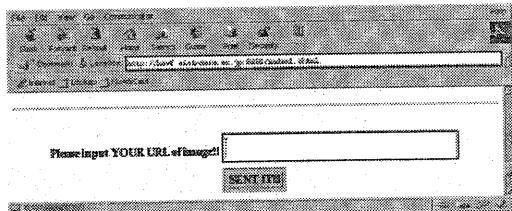


図 9: ユーザ所有画像からのマッチングの WWW インターフェース

6 評価

今回 WWW サーバ上での、ビデオ動画像情報の検索・閲覧のシステムを設計し一部実装した。

動画像情報検索・再生においてはショットを代表する静止画像であるカット点をユーザに提示することで、より取り扱い易く検索・再生が容易に行なえるようになった。

ソフトウェア MPEG2 エンコードは、エンコードと同時にリアルタイム編集ができるハードウェアと比較すると時間がかかる。しかしソフトウェア MPEG2 における、ショット分割とカット点作成は安価である。そのためユーザ側のショット分割といった、編集ツールとしての使用に適すると考えられる。

画像による画像の検索は、ニュースという対象ではカット点が似ている画像のため類似画像が検索できた。

7 今後の課題

今回は動画像情報のみを取り扱ったので、音声情報も扱って、電子図書館に組み込んで行く。

カット点についていくつかのキーワードを付加し、文字列検索や音声による検索といったマルチモーダル的機能を目指す。

類似画像の検索においては、スポーツといった動きの激しいものや特定の顔の検索が困難である。K1 展開法を 1 枚の静止画像全体で行なっているため、類似画像の検索結果を得られないと予想される。ユーザ側の画像からの検索もニュース画像のみを取り扱っているので、同様のことが予想される。

データベース蓄積においては分散型データベースにすることで、データベースへのアクセスによる負荷を軽減できると考えられる。

ニュースだけでなく、講義といった比較的ショット分割の容易なものからショット分割の困難な題材(映画や講演会)にも対応したい。

8まとめ

今回ユーザが WWW で動画像情報を容易に検索・再生・編集できるシステムを設計し、本大学の電子図書館のニュースを題材として、実装した。

検索においてはそれぞれのカット点から見たい項目がすぐに再生できるようになった。編集時にはショットをユーザが編集してデータベースに蓄積できるようにした。

参考文献

- [1] 鈴木麗, 川崎由起子, 砂原秀樹:電子図書館におけるマルチメディアデータの管理モデル, 情報処理学会マルチメディア通信と分散処理研究会 85-33,pp.189-194,1997/11
- [2] 金出武雄, 佐藤真一:Informedia :CMU ディジタルビデオライブラリプロジェクト, 情報処理, Vol.37, No.9/1996, pp.841-847
- [3] 市山俊治, 神谷俊之, 宮本温子:NEC 技報, Vol.49, No.7/1996, pp.33-39
- [4] 八幡孝, 菅野昭博, 下條真司, 西尾章治郎, 宮原秀夫:ニュース・オン・デマンドシステムにおけるニュースデータベースの設計, 電子情報通信学会データ工学研究会, 1995/5
- [5] ISO/IEC DIS 13818-2 Information technology – Generic coding of moving pictures and associated audio information: Video
- [6] 黒田英夫:画像符号化技術, 昭晃堂, 1996
- [7] <http://hpx.dsi.unimi.it/hppd/hpx/X11/Viewers/mpeg2play-1.1b/>