大規模動画アーカイブシステムにおける視覚的情報検索方式

Visual Data Retrieve Method for the Large Scale Video Archive System

正会員 前原文雄[†],井上由紀子^{††}, 正会員 池田 淳^{††}伊藤 学[†], 長谷川 文雄^{†††}Fumio Maehara[†], Yukiko Inoue^{††}, Jun Ikeda^{†††},Manabu Ito[†] and Fumio Hasegawa^{†††}

あらまし 膨大な映像データベースの中から自分の意図する素材の高速検索や加工編集を可能とし、同時に映像素材が多目的に利用できるペタ(百万 G)バイト級の大容量映像アーカイブシステムと、その視覚的検索方式を開発した。大容量アーカイブでは、アクセス速度の向上から容量、処理速度の異なる複数種類のサーバ、もしくはアーカイブを組合せて、これらを統合管理する必要がある。また、大容量画像データを利用者が内容を確認しつつ効率的に検索する必要がある。このため、マルチ画面同時表示によるブラウジング方式、ならびに画像データのシーン変化点を自動検出し、ダイジェスト映像を作成し検索に供する方式を開発した。ここでは、大規模磁気テープ自動装填ロボット(7200 巻)を中心とした大容量アーカイブシステムと、マルチ画面表示、シーン検出方式によるダイジェスト映像を併用した視覚的情報検索方式について報告する。

キーワード: ペタバイト, アーカイブ, サーバ, シーン検出, 情報検索

1. ま え が き

地上波デジタル放送や、ATM 伝送網の普及に伴ない、そのためのコンテンツの大量供給が大きな課題となっている。一方、放送局等を中心に過去の膨大な映像資産が未整理のまま蓄積され、十分な活用がなされていないという現状がある。この問題の解決には、ペタバイト級の容量を有する蓄積装置と、その中のコンテンツ(放送番組、CM など)の効率的な管理と検索、また、その加工性の向上が求められる。すなわち、膨大な映像の中から自分の意図する素材の高速検索や加工編集を可能とし、同時に映像素材が多目的に利用できる大容量映像アーカイブ技術ならびにその検索技術の確立が必要である¹⁾²⁾。

当、通信・放送機構山形映像アーカイブリサーチセンターでは、7200 巻(約 15000 時間、10 万ファイル相当の映像分)のデジタル VTR カセットを蓄積し自動的に装填する大規模テープロボットを中心に、DVD-RAM をベースとするサーバ、ハードディスクによる高速サーバの 3 種類の異なった蓄積媒体を組合せた大容量アーカイブシステムを開発し、これにより原映像、MPEG 化された 2 次データ、検索用ダイジェストデータ等の配置の最適化、再配置を行い、運用管理するための検討を行っている。

映像アーカイブデータとして想定されているものは、過去の放送番組、VTR ベースの映画、自主撮影・編集された地域映像等で、ここで実験に用いた映像データには、これに準ずるものもある。

また、蓄積された膨大な映像素材の中から自分の意図する素材を探し出し編集加工し、これらを多目的に利用するためには、高速検索が必須である。

従来の映像素材の検索においては、あらかじめ付された属性情報、例えば撮影年月日、撮影場所、撮影内容、キーワードなどを用いて映像素材を絞り込み、ある程度候補が少なくなった状態で、個々の候補テープを再生し、内容を確認することによって所望の映像を入手していた。しかし、このような方法では、数十万タイトル規模での映像アーカイブから所望の映像素材を入手するためには、非常に多くの時間を要する。

1999 年 7 月 30 日受付, 2000 年 2 月 7 日最終受付, 2000 年 2 月 29 日採録

[†] 通信・放送機構 山形映像アーカイブリサーチセンター

(〒 990-2473 山形市松栄 1-3-8, TEL 023-647-8121)

^{††} 松下電器産業株式会社 九州マルチメディアシステム研究所

(〒 820-0067 飯塚市大字川津 693-47, TEL 0948-21-2602)

^{†††} 東北芸術工科大学

(〒 990-9530 山形市上根田 200, TEL 023-627-2181)

[†] Yamagata Video Archive Research Center, Telecommunication Advancement Organization

(1-3-8, Matuei, Yamagata-shi, Yamagata 990-2473, Japan)

^{††} Kyushu Multimedia System Research Laboratory, Matsushita Electric Industrial Co., Ltd.

(693-47, Kawazu, Iizuka-shi, Fukuoka 820-0067, Japan)

^{†††} Tohoku University of Art and Design (200, Kamisakurada, Yamagata-shi, Yamagata 990-9530, Japan)

これを補う方法として、映像そのものを検索の手段として用いるものがある。これには映像の編集カット点を検出し、その先頭フレームを代表的静止画として表示するもの³⁾⁴⁾や、画像の1シーンを形、色、位置関係等の情報としてとらえ、インデックスとするもの等がある^{5)~9)}。

編集作業は、番組作成の本質部分であり、編集カット点は番組編集者の意図が込められており、番組内容のおおまかな把握という意味では、非常に重要な情報部分であると考えられる。

しかし、編集カット点1フレームのみでは動きのある映像の内容を把握するには不十分である。

形、色、位置関係などのデータを用いる方法は、主に静止画データベース等に適用されており、本目的のような大容量動画の検索への適用に対してはまだ十分な研究がなされていない。

ここでは、7200巻(約15000時間相当分)の大規模映像データベースに対する視覚的映像検索方法として、キーワードなど文字による属性検索により得られた複数の候補映像素材を、圧縮した動画ダイジェスト映像とし、さらに16ないしは64というマルチ画面に同時に表示することにより、複数映像を同時に確認しながら映像を絞り込んで行くという検索方式を開発した。

また、新しい編集カット点検出方式(シーンチェンジ検出方式)を開発した。これにより編集カット点を自動検出し、そこから3秒分(任意可変)の映像をつなぎ合わせ、ダイジェスト映像を生成し、これを前記のマルチ画面として表示することができる。このマルチ画面上にある1つのダイジェスト映像をマウスで選択することにより、実映像を検索できる。このような構成の視覚的検索方式において、種々の検討を行った。

本論文では、ペタバイト級の容量を有する大容量映像アーカイブの構成と、この内容を視覚的に検索するためのマルチ画面ダイジェスト映像を用いたブラウジング方式、ま

た、このための短縮されたブラウジング画面を作成するための編集カット点検出方式について述べる。

本論文の2章では本システムの全体構成を説明し、3章においてマルチ画面ブラウジング方式について述べる。4章においてブラウジング画面作成に用いる編集カット点検出のためのシーンチェンジ検出方式について述べ、5章にてまとめとする。

2. システム構成

2.1 全体構成

機能の概略構成を図1に、また、システムの全体構成を図2に示す。

システムは、大きく2つに大別される。ひとつは、7200巻のプロ用VTR(DVCPRO)カセットを蓄積・検索する大容量アーカイブ(アーカイブロボット)とハードディスクベースの高速サーバ(HDDサーバ)からなる、動画像蓄積システム。もうひとつは、これらの蓄積素材をベースにしたマルチメディア番組の作成、およびそれらをネットワークへ配信することなどを目的とした、番組配信制作システム

図1 システム機能構成
Functional structure of the system.

図2 ペタバイト級マルチユースアーカイブシステムの構成
System structure of the peta-byte multi-use archive system.

ムである。動画蓄積システムは放送用画質である DV 圧縮方式(Digital Video Coding Format：4 章参照)を採用しており、番組配信制作システムはこれを MPEG に変換してネットワークによる配信を行う。

上記システムの他に、アーカイブすべき映像あるいはすでにアーカイブされている映像を編集する素材作成・編集システム、システム全体の映像系、VTR 系の制御・切り換え・管理を行う映像配信制御・保管情報管理システム、アーカイブされた映像を多人数で評価するための番組再生・表示システムなど、合計 5 つのサブシステムで構成される。これらは、映像系については SDI(Serial Digital Interface)ルータにて、また情報系については 100 base-T にて接続されている。

システムの概略機能は表 1 に示すように、映像素材(番組)の登録、ダイジェスト映像の作成、マルチ画面の表示、編集カット点検出、登録素材の検索出力、登録素材の削除等である。

文字情報による検索項目の一覧を表 2 に示す。

映像素材を文字で検索する場合は、タイトル名(部分一致)の他に 2 種類の素材分類項目(各 100 個：完全一致)、ならびにキーワード 4 種(部分一致：トータル 1000 個まで)で検索を行う。キーワードは OR 検索であり、タイトルならびに分類項目は AND による検索となる。2 種類の

表 1 機能一覧
Function of the system.

機 能	内 容
番組登録受付	・ 素材サーバ装置の監視 ・ タイトル番組の生成
番組登録	・ 画像タイトルからの入力 ・ タイトル入力による登録
ダイジェスト生成	・ 編集カット点検出 ・ ダイジェスト生成と登録
マルチ画面表示	・ ブラウジング用 64 画面表示 ・ 送出用 16 画面表示
素材の削除	・ 登録された素材の削除

表 2 文字情報による検索項目
Retrieve item by words.

検索情報	内容・項目
テープ数	7200 巻(2 時間/巻)
記録時間	約 15000 時間
検索ファイル数	約 10 万
素材 ID	自動付与
タイトル	部分一致(最大 20 文字)
素材分類 1	完全一致(最大 100 種)
素材分類 2	完全一致(最大 100 種)
キーワード 1	部分一致
キーワード 2	同上
キーワード 3	同上
キーワード 4	同上

AND
AND
OR

素材分類項目とは、登録素材を例えば第 1 素材分類項目で、人物、風景、出来事、スポーツ等のように大分類し(最大 100 項目)、これとは別に、素材をさらに国別、都道府県別など別種に第 2 素材分類項目(同じく最大 100 項目)を設定できる。例えば、風景・アメリカというように項目の AND によって検索項目を絞る目的に使用する。組合せとして 100×100 のマトリクスの中から検索することになる。

システムは素材蓄積系として動画像蓄積には高画質 DV データ(Digital Video Data：ディジタル VTR に用いられる圧縮画像フォーマットの一種、4 章にて概略を説明)を用い、ダイジェスト映像などのブラウジングのための構内あるいは、外部への配信は低ビットレートの MPEG データを用い、動画蓄積システムと、配信システムとを分離している。両者はネットワークにより接続されている。大容量アーカイブが送出するディジタル映像信号は MPEG 信号に変換し、さらにファイル転送プロトコル(ftp)にデータ変換し、配信系システムに転送される。

また、アーカイブに登録された素材は番組としても登録することができる。番組登録できるものは、素材そのものと、素材と静止画やテキストを組合せた映像から作成したものである。登録された番組は、キーボードやダイジェスト映像から容易に検索し、視聴することができる。このため、編集カット点検出を組合せた場合の効果的画像検索とそのデータ管理、マルチ画面ブラウジングのための画像情報提示方式とデータ管理、マルチ画面ブラウジング画像情報提示のためのシステム制御方式が利用可能なソフトウェア構成となっている。

大容量アーカイブ装置に蓄積された映像データは、遠隔地からも参照、活用されるため、遠隔地においても視覚的検索方式を用いて素材検索の効率化を図る必要がある。しかし、ネットワークを介したデータの配信では、一律にダイジェスト映像を送出した場合、再生端末に至るネットワークの容量によって、円滑な再生が行えない場合が発生する。これをネットワークや生成端末に応じてダイジェスト映像の最適な送出を行う能力が求められる。このためにも、編集カット点を検出してダイジェスト映像を生成する方式は、データ圧縮という点からも有効である。

2.2 ペタデータアーカイブの構成

本システムのハードウェアの中心となるのはディジタル VTR カセット 7200 巻にペタデータコンテンツを格納するアーカイブロボットである。

アーカイブロボットの構成図を図 3 に、また外観図を図 4 に示す。アーカイブロボットは、倉庫ユニット、結合形送出ユニットから構成される。倉庫ユニット部は 7200 巻の VTR を格納する部分であり、BIN と呼ばれる複数種類の棚にカセットが収納されている。カセットは、制御装置であるコンピュータからの指示により、所定の格納位置から取り出される。取り出されたカセットは結合形送出カ

図 3 アーカイプロボットの構成
Structure of the Archive Robot.

図 4 アーカイプロボットの外観
Outlook of Archive Robot.

ート部に転送される。結合形送出カートはデジタル VTR が 4 台内蔵されており、必要に応じて入力もしくは出力が 4 チャンネルまで可能である。送出カート部は倉庫カートから送られてきた 10 個のカセットのうち、必要なものを順次 1 つずつ選択し、4 台の VTR のいずれかに装填する。

アーカイプロボットの仕様として表 3 に示すように、データ転送速度は 270 Mbps (SDI: Serial Digital Interface, 回線を用いている)、また内部圧縮データは DV 圧縮によるもので 25 Mbps である。エラーレートは ECC を付加したレベルで $10e^{-12}$ 、検索時間は最小 10 秒、最大 1 分となっている。

DVD チェンジャの仕様を表 4 に示す。システムはアーカイプロボットその他、DVD (Digital Video Disc) チェンジャを核とする中規模サーバ、ならびにハードディスクベースの高速サーバよりなっており、大容量な素材保存には、アーカイプロボットならびに高速サーバを用い、高画質

表 3 ペタデータアーカイプロボット仕様
Specification of the Archive Robot.

項 目	性 能
SDI ルータ, データ転送速度	270 Mbps
内部圧縮データレート	25 Mbps
インタフェース	SDI (Serial Digital Interface) SDTI
エラーレート	$10e^{-12}$ (ECC 付)
検索時間	10 s (min) - 1 min (max)
リwind時間	110 s
データ容量	23 GB/L カセット (ビデオ 2 時間分)
最大カセット収納数	7200 巻
VTR ドライブ数	4 台
全データ容量	1.14 peta-B

表 4 DVD チェンジャ仕様
Specification of DVD changer.

機 能	仕 様
記録方式	DVD-RAM
容量	2.6 GB/片面
収納枚数	100 枚
アクセス時間	17.5 s, 動画表示まで 25 s
ドライブ数	4 台
キャッシュメモリー	512 MB
蓄積可能動画時間	100 分/1 枚 (100 枚) MPEG-2 3.0 Mbps にて

(DV 圧縮)映像データを保管している。番組としてネットワークに配信する場合はこれを MPEG 変換し、DVD サーバならびに番組配信用サーバに蓄積し、外部端末からの要求に応じて MPEG データを配信する。管理ソフトウェアは蓄積系と配信系を分散管理している。

3. マルチ画面による検索

前節で述べたように映像素材の検索は、基本的には文字入力による。タイトル名やシステム利用者が、例えば人物、風景、スポーツなど、任意に 100 種までの項目名が付与できる 2 階層の素材分類項 (各々 100 項目でトータル 10000 の組合せが可能)、4 種のキーワードの AND, OR の組合せにより行われる。

しかし現在、別システムとして稼動している約 600 巻の映像ライブラリーでの実績では、文字属性などによる検索では、画像の内容確認を含めると平均検索時間は約 900 秒であり、論理的には 10 万巻にこの手法を適用すると、検索までに 259 分かかると計算される。

このため、映像素材内容の簡易な確認方法が重要であると考えられる。

本方式は、アーカイブ映像のダイジェスト映像を画像登録時に自動的に作成しておき、これを検索時の内容確認として複数画像を同時に表示し確認することにより、検索の効率化、容易化を図るものである。

実際には、文字による属性検索により先に複数の候補を選択しておき、これらに対し映像素材を同時に再生し、内容の確認を行う。

図5に本方式により64個の動画を同時再生しているマルチ画面ブラウザの例を、図6にマルチ画面動画再生方式の概略を示す。

マルチ画面ブラウザで表示される個々のデータは、あらかじめ合成サーバのハードディスク上にMPEG-1形式の合成用縮小動画ファイルとして蓄積されている。検索端末からの合成要求に応じて、合成サーバは、64個または16個の縮小動画データを用いてMPEG-1形式の合成ストリーム(マルチ画面動画)1本を作成し、検索端末に送出する。検索端末では、合成サーバからのマルチ画面動画ストリームを、自らMPEG-1デコーダにて復号し、ディスプレイに表示する。

合成ストリームの作成では、それぞれの縮小データ(MPEG-1形式)を復号することなく、すなわち圧縮された状態のまま、所定の位置に配置する。位置情報はスライスデータ部に記述されており、本合成処理では、おのおの縮小動画データにおけるスライスデータ部の位置情報を、逐次所定の位置情報に書き換えている。

一般にスライスデータ部の位置情報領域は、ビット単位での可変長データであるため、位置情報の書き換えにおい

図5 64動画同時再生のマルチ画面ブラウザの例
An example of multi display of 64 sub-divided motion picture.

図6 マルチ画面動画再生方式の概略
Algorithm outline of multi-motion picture display.

てはスライスデータ(画像データ)のビットシフト処理が必要になり、高速合成が困難である。そこで、高速合成を可能にするためにMPEG-1規格に準拠した上で、スライスデータ部の位置情報領域を固定長とし、ここにスライスデータのサイズおよび各スライスのバイト位置を書込んでおき、合成時には固定長の本情報をもとに、ブロック内の必要なスライスデータを取り出して処理することにより、位置情報の書き換えに伴うスライスデータ(画像データ)のビットシフト処理が生じないようにした。

次に、縮小動画データについて説明する。縮小動画データは、映像素材(DVデータ)をアーカイブに登録する際に自動的に作られる。作成の手順は以下の通りである。

まず、映像素材(DVデータ)のダイジェスト映像(DVデータ)を作成する。すなわち、映像素材から、後述するシーンチェンジ検出処理により編集カット点を探し出す。その後、編集カット点を先頭に所定の時間長、例えば3秒のDVデータを切り出し、おのおのの編集カット映像の出現順にしたがって切出したDVデータを連結し、時間方向に縮小したダイジェスト映像をつくる。

次に、このダイジェスト映像(DVデータ)の解像度を、縦、横それぞれ8分の1に縮小する。DVデータは離散コサイン変換(DCT)を用いたフレーム内動画圧縮方式であるため、この縮小処理は一般的なパーソナルコンピュータを用いて高速に行うことができる。最後に、縮小ダイジェスト映像をMPEG-1形式に圧縮し、縮小動画データを得る。このとき、先に説明したように、スライスデータ部の位置情報領域に冗長性を持たせる。

このように、映像素材の概略を表す縮小動画データを複数(64個、16個)同時に再生することにより、候補映像素材の内容確認が容易になる。ゆえに、映像素材そのものの再生を多数繰り返す必要がなくなり、映像素材の高速検索が可能となった。

以上の処理は、高速演算アルゴリズムにより、Pentium-II 400 MHzのパーソナルコンピュータによって、8000フレーム/sの合成が可能である。これは本システムのように10台程度の検索端末を想定した用途には十分な性能である。また、MPEGの伸長、再圧縮の必要がないため、画質劣化のないマルチ画面ブラウジングが実現できた。

4. シーンチェンジ検出

4.1 概要

映像素材を効率的に検索するひとつの手がかりとして、映像情報の編集カット点(シーンチェンジ点)を自動検出する方法がある。これにより、カットから始まる映像素材のある部分をインデックスとして検索したり、また、編集のキーデータとすることが可能である。また、シーンチェンジ点の検出結果をもとに原映像のダイジェストを作成することも可能である。本システムでは、シーンチェンジ点検出により検出された編集カット点から数秒の映像をつ

なぎ合わせることによるダイジェスト映像の生成が可能である。さらには、これによりダイジェスト映像をマルチ画面として一覧表示して、視覚的な検索を可能としている。

4.2 DV 圧縮データの構造

本システムにおけるシーンチェンジ検出ではDV圧縮データを用いている。ここでは、シーンチェンジ検出との関連においてDV圧縮の方式を簡単に述べておく。DV圧縮データとは1フレーム、525-60システムで10本、625-50システムで12本のDIFシーケンス(12kB)からなる。

DIFシーケンスは150個のDIFブロック(80B)からなり、各ブロックにはID(ブロックインデックス)とデータが格納されている。DIF(Data Information)ブロックとは、圧縮されたデータの1かたまりをいう。

1フレームの画像は 8×8 (64)ブロック毎に分けられ、DCT(離散コサイン変換)によりDC(直流)成分とAC(交流)成分に分けられる。DCTには、8-8DCTと2-4-8DCTの2種類があり、エンコードの際に、ブロック内の奇数フィールドと偶数フィールドの間のデータ差が小さいときには8-8DCTが、差が大きい時には2-4-8DCTが用いられる。このDCTの種別を識別するために、DCTモード(1ビット)が記録される。

4.3 シーン変化点検出における問題点

従来のシーンチェンジ検出技術では、映像の1フレーム分あるいは1フレーム中の分割した部分の色のヒストグラムと対応する部分における色のヒストグラムとの類似度の比較により、シーンチェンジを検出していた³⁾⁴⁾。しかし、色のヒストグラムを用いる方法では、色調の似たシーンチェンジ検出において検出漏れを起こし易かった。

また、24フレームの映画フィルムから30フレームのテレビ映像等へのテレシネ変換などに際し、フィールド間にシーンチェンジがある場合、隣接する画像との差はフレーム間にシーンチェンジがある場合に比べて隣接する画像の差は小さくなり、検出漏れが起りやすくなる。

4.4 検出方式

ここでは、色調の似たシーンチェンジ検出において検出漏れを起こし易いという問題を解決するための方式(以下、方式1)と、フィールド間にシーンチェンジがある場合の検出漏れを解決するためのDVデータの特徴を用いた方式(以下、方式2)を開発し、この2つの方式を組合せた高速シーンチェンジ検出の新方式を開発した。以下に、方式1について説明する。

シーンチェンジ点では、上記DVデータに対して隣接する2フレームの画像データの差分が大きくなる。そこで、ここでは図7に示すように、2フレーム間の同一位置にあるブロックのDC部分におけるデータの違い(差分)に着目した。この差分を用いることにより、色調の似たシーンチェンジであっても検出漏れが起りにくくなる。しかし、図8(a)に示すように、この差分だけでは、シーンチェンジ(β)のみでなく動きの激しいシーン(α)において

図7 シーンチェンジ検出方式(1)
Scene change detection method (1).

(a)

(b)

図8 変化量の比較
Comparison of the difference parameter.

も、2フレーム間のデータの差分が大きくなることから、過剰検出が多くみられた。ここで図8(a)を見ると、シーンチェンジ(β)と動きの激しいシーン(α)では、異なる点がある。シーンチェンジ(β)では、そのフレームと前のフレームとの差分が大きいが、その前後における差分は小さい。動きの激しいシーン(α)では、激しい部分全体において前のフレームとの差分が大きい。そこで、判定を行うフレームにおける差分と、その前後のフレームにおける差分とを比較し、差分を取ると、図8(b)に示すような結果が得られる。これを変化量1とする。図8(b)に示すように、動きの激しいシーン(α')ではグラフの凹凸は小さいが、本来のシーンチェンジ点(β')ではこれが大きい。ここでプラスの値を出した差分に注目し、この値がある閾値よりも大きくなる所をシーンチェンジ検出点とする。

このように、変化量1を用いれば、色調の似たシーンチェンジでは検出漏れが起りにくくなり、動きの激しいシーンでは過剰検出を抑制することができる。

図9に示すように、瞬間的な変化の多い映像の対策としては、変化の大きいフレームを挟む2つのフレームは変化が少ないことを利用して、過剰検出を削減した。

以上、変化量1を用いる方式が方式1である。

次に方式2について説明する。方式1では、フィールド間にシーンチェンジが存在する場合、そのフレームと隣接する前後のフレームでは画像データの差分が小さくなるので、検出漏れが起こりやすい。そこで、DVデータの圧縮の特徴であるDCTモードを利用する。

DCTモードは、図10に示すように、ブロック内の奇数フィールドと偶数フィールドの間のデータ差が大きいときにモード1(2-4-8 DCT)が記録されている。これは、2-4-8 DCTを用いることで圧縮効果を上げ、画質を保持するためである。フィールド間にシーンチェンジが存在する場合には、シーンチェンジが存在しない場合に比べてモードが1である割合が非常に高い。そこで1フレーム中のDCTモード=1の数をカウントし、これを変化量2: $SM(N)$ とすると、以下の式で表される。

$$SM(N) = \sum DCT\ mode\ (N, I) \quad (1)$$

N : フレーム番号

I : ブロック番号

フレームにおけるDCTモード1の数 $SM(N)$ が、前後のフレームと比較して、閾値よりも大きくなる場合に、フィールド間にシーンチェンジがあると判断する。

このように、DCTモードを用いることにより、フィー

ルド間に存在するシーンチェンジ点についても良好な検出結果が得られた。

なお、この方式2においても、瞬間的な変化の多い映像において過剰検出することがあるが、この対策としても、変化の大きいフレームを挟む2つのフレームは変化が少ないことを利用して、過剰検出を削減した。

以上、DCTモードを用いる方式が方式2である。

最後に、新方式について説明する。新方式は、方式1の判定結果と方式2の判定結果のいずれかでシーンチェンジが検出されれば、これを用いる。ただし、フィールド間にシーンチェンジが存在する場合は、その前後のどちらのフレームとも差が大きくなり得る。そこで、フィールド間にシーンチェンジが存在すると判定された場合には、その次のフレームで方式1の判定結果がシーンチェンジとなっても、その結果は用いない(つまりフィールド間にシーンチェンジが存在した場合には、次のフレームはシーンチェンジとはなり得ない)。

4.5 実験評価

表5に本方式によるシーン検出実験結果を示す。表において動画像の内容について説明すると、AとGは自主撮影でフィールド混合のシーンチェンジの多いもの。Bは映画で動きが激しく過剰検出し易いもの。Cはテレビ放送番組でフィールド混合フレームの中にシーンチェンジがあるので、検出漏れが起こり易く、動きの多いもの。Dもテレビ放送番組だが、動きの少ないもの。Eは自主撮影・編集画像で、50 Hzの蛍光灯の下で30 fpsのDVカメラで撮影した場合、フリッカが生じることで過剰検出し易いもの。Fは自主撮影で逆光、フラッシュの多いものを用いた。

表5 検出結果
Detection result.

データ名 (フレーム枚数)	シーン数 (ノーマル, フィールド混合)	データの説明	検出結果: 検出数 *(過剰, 漏れ)		
			方式1	方式2	方式3
動画像 A (6740)	69 (7, 62)	フィールド混合のシーンチェンジの多い画像	16 (2, 55)	63 (0, 6)	69 (0, 0)
動画像 B (656)	1 (1, 0)	パン、ズームの激しい映像	2 (1, 0)	1 (0, 0)	2 (1, 0)
動画像 C (623)	10 (1, 9)	街中の様子	6 (0, 4)	10 (0, 0)	10 (0, 0)
動画像 D (903)	5 (1, 4)	風景	2 (0, 3)	5 (0, 0)	5 (0, 0)
動画像 E (1174)	5 (5, 0)	蛍光灯によるフリッカの激しい映像	5 (0, 0)	1 (0, 4)	5 (0, 0)
動画像 F (1445)	1 (1, 0)	逆光、フラッシュの多い映像	1 (0, 0)	1 (0, 0)	1 (0, 0)
動画像 G (1824)	21 (3, 18)	フィールド混合のシーンチェンジの多い画像	3 (0, 18)	19 (0, 2)	21 (0, 0)

*誤検出数 (過剰: 過剰検出数
漏れ: 検出漏れが生じた数)

図9 シーンチェンジ検出方式(2)
Scene change detection method (2).

図10 シーンチェンジ検出方式(3)
Scene change detection method (3).

実験としては、方式1と方式2、およびその両方を用いた新方式について行った。

表5からもわかるように、方式1においては、過剰はほとんどないものの、フィールド間に存在するシーンチェンジは8割近くが検出されていないし、フレーム間とフィールド間のどちらにシーンチェンジが存在するのかわからない。

動画像Aにおいては、過剰検出も見られるが、これは、フィールド間にシーンチェンジがあるフレームとその次のフレームとの差が大きいため検出されたものである。

方式2においては、フレーム間に存在するシーンチェンジは全く検出できていないが、過剰検出はなく、フィールド間に存在するシーンチェンジはすべて検出されている。

新方式においては、動画像Bで過剰検出が1カ所みられたのみで、他の被験画像については、検出漏れ、過剰検出ともに0という結果が得られた。

この新方式を用いて10時間程度の映像データ(自主撮影した自然風景等)をアーカイブしダイジェスト映像を作成しているが、現在のところダイジェスト映像の内容判断、検出精度という面では問題ない(ただし、本映像はフラッシュや特殊効果を用いた編集等は比較的少ない)。

大容量データによる詳細な検出誤りの評価は今後の課題であり、現在さらに大容量データによりその有効性を検証中である。

5. む す び

膨大な映像の中から自分の意図する素材の高速検索や加工編集を可能とし、同時に映像素材が多目的利用できる大容量映像アーカイブシステムを構築し、その上で映像素材の視覚的検索方式について検討を行った。映像データの検索手法には、個々の映像データのダイジェスト映像を一覧表示して視覚的に検索する方式が有効である。とくにマルチ画面表示によるダイジェスト映像表示は、その高速性という意味で効率的であり、内容確認という意味では非常に効率化が図れた。

また、シーン変化のあった場面を抽出し、これからダイジェスト映像を生成することにより、マルチ画面に表示する方式をあわせて開発し、その有効性を確認した。シーンチェンジ点検出は現在さらに多くのデータを収集中であるが、10時間程度のアーカイブデータにより検索作業を行った結果、内容判断、検出制度においては問題はなかった。しかし、シーンチェンジ検出率、またパラメータ閾値の最適決定方法は、今後さらに検討が必要である。

マルチ画面表示では、何画面分割が適当か、またキーワード選択とマルチ画面作成画像の対比の検討などが課題で

ある。シーン検出については実用的にも比較的良好な結果が得られているが、今後の映像データベース整備の中で非常に特殊な変化、また非常に古い映像、テレシネデータ等については十分な評価がなされているとはいえない。

また、データの一元管理、再配置と検索方式の関連、シーン検出精度の改良とユーザインタフェースの相関性、マルチブラウジング機能とデータの統合管理等が今後の課題となってくるであろう。

(文 献)

- 1) 藤井実, 浅井清: “階層的ファイル自動管理システムの設計”, 情報学論, 21, 6, pp. 442-453 (Nov. 1980)
- 2) 鈴木偉元, 西村一敏, 坂本秀樹: “階層化蓄積ビデオサーバの性能解析”, 信学論, J80-D-I, 3, pp. 300-308 (Mar. 1997)
- 3) 長坂晃朗, 宮武孝文, 上田博唯: “カットの時系列コーディングに基づく映像シーンの実時間識別法”, 信学論, J79-D-II, 4, pp. 531-537 (Apr. 1996)
- 4) 谷口行信, 外村佳伸, 浜田洋: “映像ショット切換え検出法とその映像アクセスインタフェースへの応用”, 信学論, J79-D-II, 4, pp. 538-547 (Apr. 1996)
- 5) 赤堀裕志, 五島雪絵: “撮影状態推定による動画像自動編集方式の提案”, 信学技報, IE 94-117 (Jan. 1995)
- 6) 藤本真: “画像検索技術の動向”, テレビ誌, 47, 1, pp. 59-66 (Jan. 1993)
- 7) 坂内正夫: “画像検索技術”, 信学論, 7, 9, pp. 911-914 (Sep. 1988)
- 8) 岩成英一, 有木康雄: “DCT成分もちいたシーンのクラスタリングとカット検出”, 信学技報, PRU 93-119, pp. 23-30 (Jan. 1994)
- 9) 上田博唯, 宮武孝文, 古沢聡: “認識技術を用いた対話型映像編集方式の提案”, 信学論, 75-D-II, 2, pp. 216-225 (Feb. 1992)

まえはら ふみお
前原 文雄 1975年、電気通信大学大学院電気通信専攻修士課程修了。同年、松下電器入社。中央研究所にて音声認識の研究に従事。MIT客員研究員、松下電器台北技術研究所長を歴任。現在、同社ビデオシステム事業部参事。通信・放送機構、山形RC、サブリーダー。工学博士。正会員。

いのうえ ゆきこ
井上由紀子 1993年、九州芸術工科大学芸術工学部画像設計学科卒業。同年、松下電器産業(株)入社。同社、九州マルチメディアシステム研究所において、色彩処理技術および、映像編集に関する研究開発に従事。

いけだ じゅん
池田 淳 1989年、九州大学理学部物理学科卒業。1991年、松下電器産業(株)入社。現在、同社九州マルチメディアシステム研究所において、色彩処理および映像編集に関する研究開発に従事。正会員。

いとう まなぶ
伊藤 学 1994年、千葉工業大学工学部電気工学科卒業。同年、山形ナショナル電機(株)に入社。主として音響・映像システムの設計、販売に従事。現在、通信・放送機構山形映像アーカイブリサーチセンターにてペタバイト級マルチユースアーカイブ技術の素材検索技術を担当。

はせがわ ふみお
長谷川文雄 1974年、電気通信大学大学院電気通信学研究科修了。同年、清水建設(株)入社。1986年、MIT都市工学客員研究員、1989年、東京大学先端科学技術研究センター都市開発工学助教授を経て、1998年、東北芸術工科大学副学長。現在、通信・放送機構山形RCにてプロジェクトリーダー。工学博士。