

映像インデクシング研究のための 大規模放送映像アーカイブシステムの試作

片山紀生 孟 洋 佐藤真一

国立情報学研究所

〒101-8430 東京都千代田区一ツ橋 2-1-2

TEL (03) 4212-2620 E-mail {katayama,mo,satoh}@nii.ac.jp

映像インデクシング研究のさらなる発展のためには、実用規模での映像アーカイブを対象とすることが求められる。そこで、我々のグループでは、テラバイトオーダの放送映像アーカイブシステムを設計、試作したので、その概要について報告する。このシステムの特徴として、複数の市販製品を緩やかに組み合わせて使っている点があり、メインサーバとしての UNIX ワークステーションに加えて、MPEG キャプチャカードと文字放送デコーダカードを装備した PC を複数台使うことでシステムを構築している。また、我々は、保存された映像にアクセスするためのツールとして、Java によるビデオブラウザを試作した。このブラウザは、映像インデクシング研究のためのソフトウェアプラットフォームとして利用することを狙っており、応用プログラムのプロトタイピングや、映像インデクシング、映像解析のツールとして有用であると考えている。本稿では、この映像アーカイブシステムの設計上、実装上の考察事項、ならびに、今後の展望について論じる。

Large-Scale Broadcast Video Archive Testbed for Video Indexing Research

Norio Katayama Hiroshi Mo Shin'ichi Satoh

National Institute of Informatics

2-1-2 Hitotsubashi, Chiyoda-Ku, Tokyo 101-8430, Japan

TEL +81-3-4212-2620 E-mail {katayama,mo,satoh}@nii.ac.jp

It is desired to build the realistic scale video archive system for further advancement of video indexing research. Based on this idea, we designed and are constructing a broadcast video archive system. The system is composed of commodities as key components, such as unix workstations, RAID disk arrays, and MPEG capture cards and closed-caption decoder cards installed in PCs. We also developed the experimental video browser system which is intended to be used as the software platform of the system aiming at rapid prototyping of video applications and video analysis software. This paper discusses designing issues, implementation issues, and future directions of the broadcast video archive system.

1 はじめに

BS デジタル放送の開始などデジタルテレビジョン放送の動向, IMT-2000 や FTTH など個人向け高速ネットワークの実現動向, MPEG-7 や TV Anytime Forum など映像メタデータ記述やサービスに関する動向など, 大量の映像データに対する, 内容に依存した検索などの知的アクセスを各視聴者に提供する環境が整いつつある。われわれの研究グループでは, こうした環境の実現のためのキーテクノロジーとして, 映像解析手法を用いて映像の内容に応じた索引付けを行う手法, 映像アーカイブを効果的に利用するための手法など, 一般に映像インデクシングと呼ばれる技術の重要性に着目し, 研究を行ってきている。今後各視聴者がアクセス可能となるチャンネル数は数千を越えると考えられ, かつ各放送局は連日途切れることなく番組を提供するため, 本質的に極めて大規模で, かつ時々刻々更新されるという動的な性格を持つ「映像空間」が実現されることになる。こうした大規模・動的な映像空間に対応しうる映像インデクシング技術の実現のためには, 相応の規模を持ち, 動的な性格を備えた映像アーカイブの実現と, これを用いた実証的な研究が重要になってくると考えている。

そこで, 我々のグループでは, この要請に応えるため, 将来提供されると考えられる映像空間と同等の規模・性格を持った映像アーカイブを実現するような放送映像アーカイブシステムを設計, 試作すると同時に, これを利用した映像インデクシング研究のソフトウェアプラットフォームとしてビデオブラウザを開発したので, その概要について報告する。このシステムは, 複数チャンネル(現在7チャンネル)の放送映像を終日連続してデジタル記録することが可能であり, 最大 2,300 時間(約2週間)分を MPEG-1 フォーマットで保存することが可能である。また, 特定の番組(ニュース番組)について, これまでに 300 日分を MPEG-2 フォーマットで保存している。また, このシステムの特徴として, 複数の市販製品を緩やかに組み合わせて使っている点があり, メインサーバとしての UNIX ワークステーションならびに RAID ディスクアレイに加えて, MPEG キャプチャカードと文字放送デコーダカードを装備した PC を複

数台使うことでシステムを構築している。システムの運転は自動化されており, 最新映像を逐次保存できる状態になっている。以下, この映像アーカイブシステムの設計上, 実装上の考察事項, ならびに, 今後の展望について論じる。

2 システムに求められる機能

映像インデクシング研究のための放送映像アーカイブシステムに求められるのは, 実用規模に近い研究や, 研究成果の適正な評価などを支援する機能である。理想的には, 映像アーカイブに, より長い期間(～数年), より多くのチャンネルを同時に, かつ日々の放送を順次格納していくことが望まれる。しかしながら, これはシステムの容量などの制約から, 現実的には不可能な要求である。そのため, これらの条件に対して妥協点を見出す必要がある。そこで, 映像アーカイブに対して求められる事項を整理すると, 以下のようにまとめることができる。

(1) 多様性

比較的短い期間に(数日, 数週間など), 番組の種類を問わずできるだけ多くのチャンネルの映像ストリームを並行して記録し, 同時期の多様な映像を格納すること。

(2) 継続性

毎日あるいは毎週放送される特定の番組を長期間(数ヶ月～数年)継続して格納すること。

(3) 自律性

連日放送されている映像を自動的に格納し, アーカイブを動的に更新すること。

多様な番組に対する視聴環境を想定した研究[1]などでは, 一番めの「多様性」を持つ映像アーカイブが必要になる。一方, 特定のタイプの番組に対する高度な内容解析を行う研究, 例えば, ニュース番組[2, 3, 4], スポーツ番組[5, 6], 料理番組[7]を対象とした研究では, むしろ二番めの「継続性」を持つ映像アーカイブが必要になる。従って, 映像インデクシング研究のための映像アーカイブシステムとしては, 多様性と継続性の双方を満足す

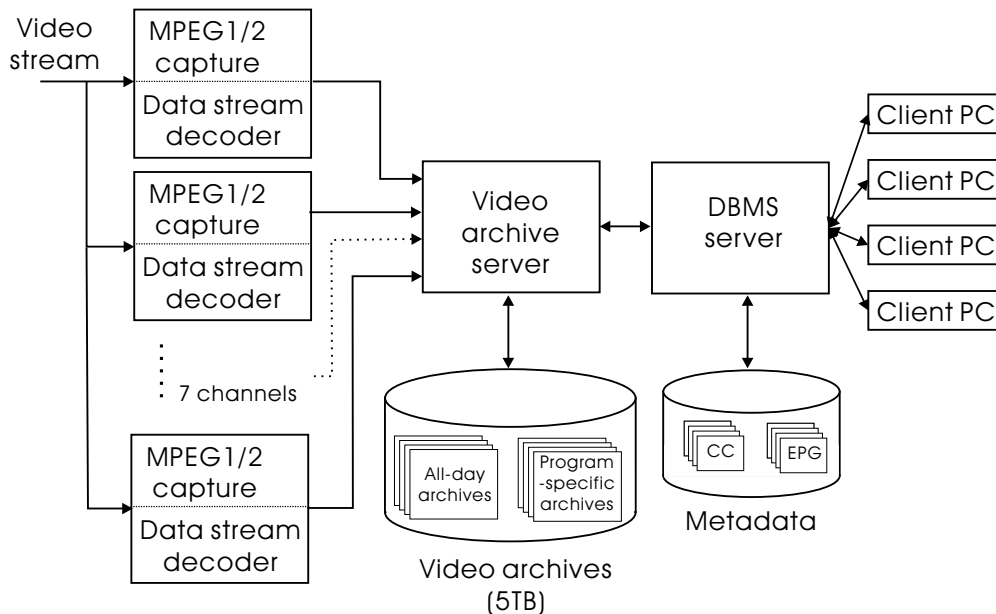


図 1: 映像アーカイブシステムの構成

ることが望ましいと考えられる。しかも、そのようなシステムでは、複数チャンネルで放送されている映像を終日連続して格納するとともに、それらの映像に対するメタデータを収集する必要がある。そのような作業は、人手に頼ることはできないため、必然的に自動運転することが求められるので、システムの自律性も重要な要求事項である。

これまでも映像アーカイブについて様々な研究が行われているが、これらの三つの条件を同時に満たそうとする試みはあまり見られない。特に、自律性については、あまり注目されていなかったと考えられる。カーネギーメロン大学の Informedia プロジェクトでは、特定の種類の番組（ニュースとドキュメンタリ）について、自律性ならびに継続性のあるシステムを構築しているが、多様性という観点では十分ではないと言える。我々の映像アーカイブシステムでは、これら三つの条件を同時に満たすことを設計上の重要な目標としている。

3 システムの実装法

今日では、映像アーカイブシステムの構成要素として利用可能な機器が、市販製品として数多く販売されている。例えば、ビデオキャプチャカードや文字放送デコーダは、PC 用のものが低価格

で普及しており、磁気記憶装置についても大容量化が急速に進んでいる。しかし、我々が求める機能を単独で満足する市販製品はない。そのため、我々は、これらの市販製品を組み合わせ、統合的に利用することで映像アーカイブシステムを実現することにした。そこで、この節では、市販製品の組合せ方など、システム実装上の考察事項について述べる。

図 1 に映像アーカイブシステムのブロック図を、図 2 にその外観を示す。本システムでは、大規模・動的な映像情報ソースとして、実際の放送映像（東京地区の地上波 7 チャンネル）を利用する。7 チャンネルの放送映像を同時に並行して 24 時間終日キャプチャし、各チャンネルの映像ストリームを 1 時間分ずつ MPEG 形式のビデオファイルの形で保存する。また、映像ストリームに加えて、テキスト形式のメタ情報として、文字放送に含まれている文字字幕（クロズドキャプション）およびインターネット上に提供されている電子プログラムガイドを蓄積する。放送映像の MPEG 形式への変換、ならびに、文字放送からのテキストの抽出には、PC にインストールされている MPEG キャプチャカード (Canopus MVR-D2000) および文字放送デコーダカード (Systec 文字ビジョン 550) を用いており、チャンネルごとにこれらのカードを割り当てている。これらのカードを実装した PC

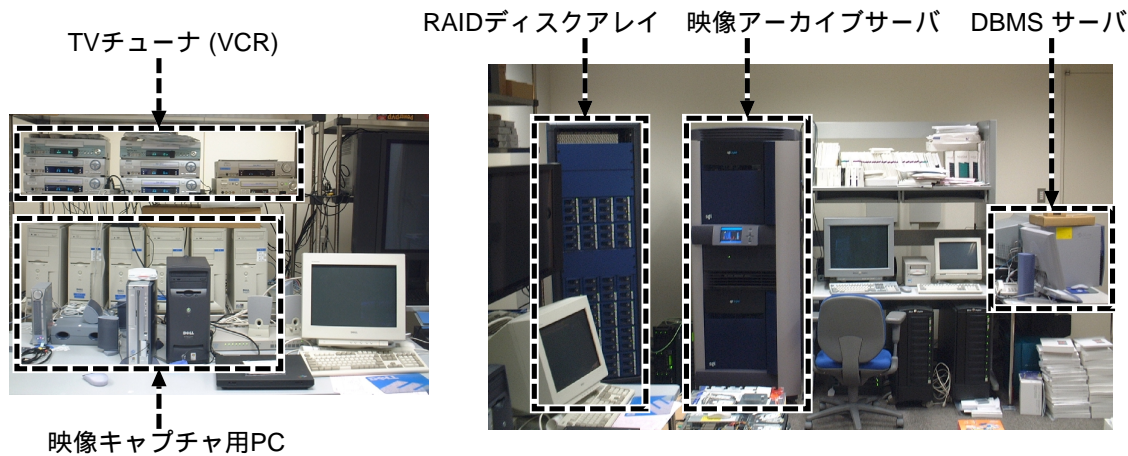


図 2: 試作システムの概観

は、後述の映像アーカイブサーバならびに DBMS サーバとファストイーサネットによって接続されている。これらの PC は、映像アーカイブサーバによって制御されており、映像アーカイブサーバが PC に対して処理を要求する形、つまり、RPC (Remote Procedure Call) を発行する形になっている。RPC を実現する部分には Java RMI を用いている。

獲得されたデータは、映像アーカイブサーバに保存される。映像アーカイブサーバとしては、SGI Origin 2000 を使用しており、ファイバチャネルによって 5TB の RAID ディスクアレイが接続されている。本システムでは、5 TB の容量を 3 TB と 2 TB とに分け、それぞれ「多様性」と「継続性」のためのアーカイブに割り当てている。「多様性」を持つアーカイブとしては、7 チャンネルの放送映像を並行して終日 (24 時間) MPEG-1 の形式で保存しており、容量の制限から、最近の 2 週間分 (約 2,300 時間) を保有し、常時更新し続けている。一方、「継続性」を持つアーカイブとしては、特にニュース映像の重要性に着目し、文字放送による文字字幕放送を行っている NHK のニュース 7 を連日格納し続けている。システムの保守・拡張などのために格納できなかった日もあるため、一貫して連続している訳ではないが、これまでに約 300 日分の映像を蓄積している。このアーカイブについては、顔検出や動き解析などの高度な映像処理を可能とするため高精細の MPEG-2 形式でも格納している。本システムの運転は自動化されて

おり、三つの要求事項のひとつである「自律性」も実現されている。また、現在は、一部のニュース番組のみを継続的に保存しているものの、他の番組を加えることも設定の変更のみで容易に対応できる。

映像アーカイブシステムのもうひとつの構成要素として DBMS サーバがある。ハードウェアとしては Sun Blade 1000 を、DBMS ソフトウェアとしては Oracle 8i を使用している。DBMS サーバは、映像アーカイブの内容に変更があるかどうか定期的にチェックし、変更があった場合には、その内容を DBMS サーバのデータベースに自動的に反映する。現在の実装では、以下のメタデータを DBMS サーバのデータベースに格納している。

- MPEG ビデオファイルの属性 (ファイルパス、放送された時刻、再生時間、チャンネル番号、など)
- 文字放送の文字字幕 (字幕テキスト、放送された時刻、チャンネル番号、など)
- 電子プログラムガイド (番組タイトルなどのテキスト、番組開始時刻、チャンネル番号、など)

文字字幕と電子プログラムガイドは、全文検索できるように、Oracle InterMedia Text によって索引付けされている。これらのメタデータは、膨大な「映像空間」の中から、特定の映像セグメントを選択し取得するための基本的な手段を提供している。

以上の実装によって、前節で要求事項として挙

げた「多様性」「継続性」「自律性」の三点を、かなり満足のいく形で実現できていると考えている。本システムの特徴としては、市販製品を緩やかに組み合わせて統合的に使っている点があり、安定に連携動作させるのに手間がかかるという欠点がある反面、構成要素の置き換えや拡張・増強が容易であるためシステムを柔軟に構成できるという利点がある。

4 ソフトウェアプラットフォーム

この映像アーカイブシステムは、映像インデクシング研究のために研究者によって使われることを想定している。基本的には、RAID ディスクアレイに保存されている MPEG ビデオファイルと、Oracle データベースに保存されているメタデータにアクセスすることさえできれば、このシステムを利用することが可能である。しかし、これらのデータを使用するプロトタイプシステムや実験用のプログラムをゼロから作成するのは容易なことではない。そのため、映像インデクシング研究に適した API (Application Programming Interface) を持つソフトウェアプラットフォームを提供することが重要である。ソフトウェアプラットフォームは、開発コストを下げるだけでなく、映像アーカイブシステムの管理コストも下げることになる。というのは、API を決めることでシステムの内部構造が隠蔽されるため、システム管理者は、API の利用者に大きな影響を与えることなく、システムの内部構造を変更・改良することが可能になるためである。

ソフトウェアプラットフォームの設計に際しては、開発効率の向上を目的として、以下の点に配慮した。

- (1) 研究者が、少ない工数で、映像ファイルおよびそれらのメタデータを利用したアプリケーションを構築できること。
- (2) 研究者の手元のハードウェアで、アプリケーションを開発・実行できること。これは、ソフトウェアプラットフォームが、映像アーカイブシステムへのリモートアクセスを提供する必要があることを意味している。

- (3) 映像アーカイブシステム自体が PC と UNIX ワークステーションの複合体である上、研究者のハードウェア環境も様々であるため、複数のハードウェアに対応できること。

これらの観点から、ソフトウェアプラットフォームの骨格として、関係データベースシステムをサーバとし、Java 言語で書かれたアプリケーションをクライアントとする形態を取ることにした。この組合せは、広く普及している形態であるため、開発ツールや情報を入手しやすいという利点がある。また、Java 言語は、多くのハードウェアで使用できるため、開発環境の選択肢を大きく広げることができる。さらに、Java は、関係データベースシステムとの接続を実現するための仕組みとして JDBC (Java Database Connectivity) を提供しており、多くの関係データベースシステムが JDBC によるリモートアクセスをサポートしている。

骨格としてこのような形態を取るとともに、アプリケーションの開発効率を高めるために、映像データを利用するための基本的な機能をサポートする Java アプレットを作成した。我々はこのアプレットをビデオブラウザと呼んでおり、図 3 にそのシステム構成を示す。このアプレットの主な機能は以下のとおりである。

- プロトタイプシステムを開発するための基本的な GUI コンポーネント (例えば、ビデオファイルを再生する機能) を提供する。
- Oracle データベースに格納されているメタデータ (ビデオファイルの情報ならびに文字放送、電子プログラムガイドの情報) にアクセスするための基本機能を提供する。
- 研究者が独自に作成した GUI コンポーネントをアプレットに組み入れる機能を提供する。

ビデオブラウザは映像を検索・閲覧するためのユーザインターフェイスになっており、ビデオ再生用の PlaybackPanel, キャプションの同期表示のための CaptionMonitor, そして、検索用の 1 個以上の VideoSelector から構成されている (図 4)。

VideoSelector はひとつひとつがタブになっているので、適宜、切り替えて使うことが可能であり、自作の VideoSelector を追加することでアプリケーションを開発することが可能になっている。例え

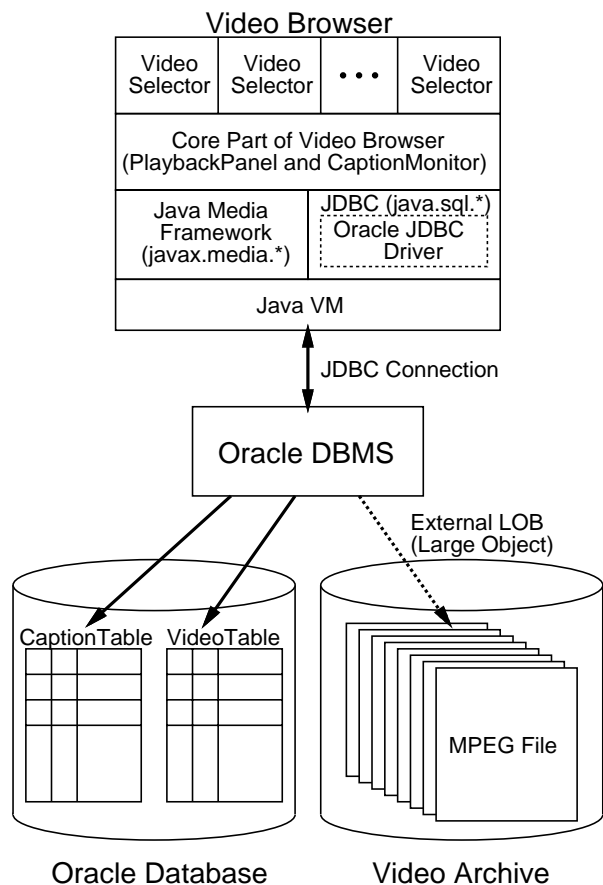


図 3: ビデオブラウザのシステム構成

ば、ビデオブラウザは、デフォルト設定として、以下の VideoSelector を持っている。

- VideoViewer: ビデオファイルの時間順の階層型リスト表示
- CaptionViewer: テキスト情報（文字字幕または電子プログラムガイド）の時間順の階層型リスト表示
- CaptionFinder: テキスト情報（文字字幕または電子プログラムガイド）の全文検索

ビデオブラウザを使ってアプリケーションを開発すれば、ビデオ再生などの基本的な機能は開発する必要がなく、研究者は、独自の機能の実装にのみ専念することができる。そのため、このような共通プラットフォームを提供することで、開発効率の向上が期待できると考えられる。

図 3 に示すとおり、ビデオブラウザは、Oracle DBMS を経由して映像アーカイブ中の MPEG ファ



図 4: ビデオブラウザ

イルにアクセスする。Oracle DBMS には、一般のファイルを外部 LOB (Large Object) として読み出す機能があり、本システムではこの機能を使って MPEG ファイルをデータベースに登録している。これにより、ビデオブラウザはメタデータと MPEG ファイルの両方を JDBC コネクションを通して取得できることになる。読み出された MPEG ストリームは Java Media Framework のプレイヤーに供給され、ランダムアクセス可能な形で再生される。このように、ビデオブラウザの動作には、DBMS サーバへの JDBC コネクションのみが必要であり、NFS などの他の接続を必要としない。そのため、映像アーカイブシステムへのリモートアクセスが容易であり、複数のファストイーサネットセグメントとルータを経由したアクセスであっても、ビデオブラウザが動作することを確認している。

独自の VideoSelector の開発を支援するために、ビデオブラウザは、以下の機能を持つ API を提供している。

- 特定の時刻に特定のチャンネルで放送されていた映像のファイルを探査する機能。
- 特定の映像ファイルを指定された位置から再生する機能。
- 指定されたチャンネルの指定された期間のテキスト情報（文字字幕または電子プログラムガイド）を検索する機能。

- 全文検索によってテキスト情報を検索する機能。
- 映像ファイルから指定された位置のフレーム画像を抜き出す機能。

この API を通して見ると，映像アーカイブは，MPEG ファイルの巨大なジュークボックスのように見えることになる．VideoSelector は，チャンネルと時刻を指定することによって，特定の映像ファイルを特定の位置から再生することが可能になる．映像の再生がファイルの終りに達すると，同じチャンネルの次の映像ファイルに自動的に切り替わるようになっており，利用者は，映像ストリームを連続的に視ることが可能である．

標準装備の VideoSelector である VideoViewer, CaptionViewer, CaptionFinder は，この API を利用して作られている．これらの作成を通じて，この API が映像アーカイブシステムに対する基礎的な機能を提供していることを確認している．今後は，この API の上に，より高度な映像インデクシング手法を実現していく考えである．

5 おわりに

我々は，映像インデクシング研究のための映像アーカイブシステムを試作している．本稿では，まず，映像アーカイブシステムに対する三つの要求事項として，多様性，継続性，自律性について述べた．また，これらの要求を満たす実装のひとつとして，UNIX ワークステーション，RAID ディスクアレイ，複数台の PC を組み合わせたシステムを試作し，その構成について説明した．そして，プロトタイプシステムや映像解析ソフトウェアの開発を支援するためのソフトウェアプラットフォームについても，その役割と機能について説明した．現段階では，まだ，システムの基本的な機能が整備できた状況であるが，このような大規模な映像アーカイブシステムとソフトウェアプラットフォームを利用することにより，実証的な映像インデクシング研究が行えるものと期待している．今後は，これまでに我々が取り組んできた映像インデクシング手法を含め，より高度な映像インデクシング機能を組み込んでいきたいと考えている．

参考文献

- [1] 谷口, 南, 阿久津, 外村, 浜田, 大量映像に対する効率的アクセスインタフェース, 電子情報通信学会技術研究報告, IE-95-533, 1996.
- [2] Nakamura, Y. and Kanade, T., Semantic Analysis for Video Contents Extraction — Spotting by Association in News Video, *Proc. of ACM Multimedia*, 1997.
- [3] Satoh, S., Nakamura, Y. and Kanade, T., Name-It: Naming and Detecting Faces in News Videos, *IEEE MultiMedia*, Vol. 6, No. 1, pp. 22–35, 1999.
- [4] Christel, M. G., Visual Digests for News Video Libraries, *Proc. of ACM Multimedia*, pp. 303–311, 1999.
- [5] Babaguchi, N., Sasamori, S., Kitahashi, T. and Jain, R., Detecting Events from Continuous Media by Intermodal Collaboration and Knowledge Use, *Proc. of ICMCS*, pp. 782–786, 1999.
- [6] 越後, 視覚的な手がかりによる動画像検索, 情報処理学会誌, Vol. 41, No. 4, pp. 351–356, 2000.
- [7] Reiko Hamada, Shin'ichi Satoh, Shuichi Sakai, and Hidehiko Tanaka, “detection of important segments in cooking videos,” in *Proc. of IEEE Workshop on Content-Based Access of Image and Video Libraries*, 2001, pp. 118–123.