

WEB ベース MPEG データベースシステムの開発

柳原 広昌 菅野 勝 米山 暁夫 中島 康之

(株) KDDI 研究所

〒356-8502 埼玉県上福岡市大原 2-1-15

{yanap, sugano, yoneyama, nakajima}@kddilabs.jp

近年、DVD の普及やデジタル放送サービスの開始などにより、MPEG 等で圧縮されたデジタル映像コンテンツが広く利用されるようになってきた。また、パソコンの処理能力向上や廉価な MPEG エンコーダの登場によりコンテンツ制作を低コストかつ手軽に行える環境が整いつつあり、放送・映像プロダクション系だけでなく、学校や企業など様々な分野においても利用が進みつつある。今後このように膨大に増えつつけるデジタル映像コンテンツを利活用するには、映像データの管理が重要といえる。本稿では、MPEG コンテンツの管理を効率的かつ簡易に行うことを目的として、独自のカット点検索技術、オーディオインデキシング技術、部分トランスコードによる高速編集技術を用いて、WEB ベースの MPEG データベースシステムを開発したので報告する。

Development of Web-based MPEG database system

Hiromasa Yanagihara Masaru Sugano Akio Yoneyama Yasuyuki Nakajima

KDDI R&D Laboratories Inc.

2-1-15 Ohara, Kamifukuoka, Saitama 356-8502, JAPAN

{yanap, sugano, yoneyama, nakajima}@kddilabs.jp

Recently, vast of digital audio-video contents such as MPEG compressed audio-visual data can be found in, for example, DVD and a digital broadcasting. And MPEG encoding is coming into wide use for not only TV broadcasting and movie productions but also schools and enterprises, owing to the recent progress of PC's processing power and the emergence of a low cost MPEG encoder system. It will bring the huge number of digital audio-video contents. Therefore, it is very useful to archive the contents of the audio-visual data, in case of handling or re-using these contents. In this paper, for the purpose of effective management of MPEG contents, we have developed web-based MPEG database system with our proprietary cut detection technology, audio indexing technology, and fast MPEG editing technology using partial decoding.

1. はじめに

パソコンや映像機器などの普及により多くの機器で MPEG 圧縮された映像コンテンツを再生できるようになってきた。また、パソコンの処理能力向上、蓄積装置の大容量低価格化、廉価な MPEG エンコーダの登場などにより、手軽に映像コンテンツ制作が行える環境が整いつつあり、放送・映像プロダクション系だけでなく、学校や企業など様々な分野においても利用が進みつつある。

このようにマルチメディアコンテンツが大量にデジタル化されて蓄積されるようになるとコンテンツの効率的な管理が重要[1]となり、そのための映像データベース化が行われる。ただし、これまで映像データベースについてはCM送出装置や番組録画など放送局用の大規模なデータベースシステム[2]はあったが、企業や学内などで簡単に利用できるような小規模の映像データベースシステムはあまりなかったといえる。また、後者のようなシステムも専用のクライアントソフトが必要となるなど、システム開発コストやクライアントソフトのインストールコストなどが大きくなる問題があった。

さらにネットワーク環境で映像をハンドリングする場合、MPEGなどで圧縮したとしてもコンテンツを自由にやりとりするにはファイルサイズが大きすぎるため、実質的にネットワーク上の映像コンテンツを自由に編集加工できる環境にはなかったといえる。また、MPEGファイルの場合、フレーム間相関を利用してパツファモデルに従って圧縮している関係上、Motion JPEGなどと異なりフレーム精度の加工が困難なため、MPEGコンテンツをネットワーク上で共有して閲覧することはできても、実際に編集加工にも利用することは困難な点が多かった。

本稿では、これら問題点を解決し、MPEGコンテンツの管理を効率的かつ簡易に行うことを目的として、独自のカット点検索技術[3]、オーディオインデキシング技術[4]、部分トランスコードによる高速編集技術[5]を用いて、WEBベースのMPEGデータベースシステムを開発したので報告する。

2. 開発システムに対する要求条件

筆者らは、1.で述べた問題点を解決するため以下の要求条件を挙げた。システム構築に際してはそれら条件項目を主眼に入れ設計ならびに開発を行った。

- (1) クライアント側の導入コスト削減と簡易操作を実現するため、WEBベースのシステムとし、サーバの最小構成としてPC1台で実現可能なこと。
- (2) 効率的な映像検索を目的として、テキストベース検索だけでなく、映像・音声のメタ情報を用いたコンテンツベース検索機能を有すること。
- (3) 映像管理で必要となる同時アクセスによ

るトランザクション処理の排他制御を実現するためRDBMSを用いること。

- (4) 様々なネットワーク環境に接続されたクライアント上で映像のプレビュー再生を可能とすること。
- (5) ネットワーク上に存在するMPEGストリームを共有しながら、ローカルにダウンロードすることなく高速にリモート編集可能なこと。

3. システム構成

2.で述べた要求条件に従い開発したシステムのシステム構成図を図1に示す。図のようにサーバ・クライアント構成をとっており、サーバではクライアントからのインタラクティブな要求に応えるWEBサーバ(Microsoft IIS)、MPEGデータを管理するための映像DB(Microsoft SQLServer)、独自のMPEG変換・編集・検索処理を行うためのMPEGモジュール、映像ストリーミングを行うVODサーバ(RealServer)の4つの機能を1台のWindows2000Serverで実現した(図2)。

クライアント側についてはInternet ExplorerをWebブラウザとして用い、Webブラウザ上で映像の登録、検索・閲覧、編集、再生、管理すべての作業を行えるようにした。また映像再生には代理映像としてRealVideoを用いて再生を行っている。これらソフトは汎用ソフトとして広く普及しており、クライアント側には特に専用ソフトを実装する必要がないため、サーバ設置のみですぐにMPEGデータベースを運用開始することができるほか、クライアントソフトのインストールコストを大きく削減することができる。また、本サーバをインターネットに接続することによりこれらの代理映像はストリーミングコンテンツとしてそのまま利用できるため、インターネット映像配信に利用することも可能である。

Webブラウザを用いたコンテンツ検索では階層的に画面を構成して検索していくことが多いが、映像検索、閲覧、再生、編集機能については互いに参照しながら利用することが多いため階層構造はなるべくとらず、図3のように主な機能は同一階層で実現しノンリニア編集の操作性になるべく近づけることとした。また、画面左上から順に処理を進められるようなレイアウトを用いた。

4. 登録、検索、閲覧、再生、編集、管理

本システムにおけるMPEGコンテンツの登録、検索・閲覧、再生、編集、管理機能について以下に述べる。

4.1 登録機能

MPEGデータベースを構築するには最初にコンテンツの登録作業が必要になる。登録作業に多くの時間がかかるとデータベースの運用コストを増

大させることとなるため、手元のパソコンから Web ベースでサーバにコンテンツ登録を行うことにより作業の簡便化を図った。(図 4)

通常の MPEG 画像の登録処理については、まず MPEG エンコーダなどで生成されローカルディスクに保存されている MPEG コンテンツを WEB ブラウザ上にドラッグ&ドロップすることにより、サーバの一時蓄積エリアに自動的に FTP アップロードする。

次にコンテンツのタイトル・サブタイトル・キャプションなどのメタ情報を入力して DB 登録を行う。また、カテゴリや公開期間も登録する。公開期間の指定は、CM 映像など閲覧期間に制限があるコンテンツをインターネット上で公開する場合などに有効である。DB 登録時には、AV コンテンツのインデックス情報抽出や代理映像抽出も同時に行う。なお、大量コンテンツの登録に有効な一括 DB 登録や処理負荷のかかる変換プロセスのバッチ処理も可能とした。

4.2 検索・閲覧機能

映像情報は、情報の内容そのものを明示的に表現しているテキスト情報と異なり、再生表示して初めて内容が把握できるという特徴をもち、検索のため早送り再生したとしても内容把握に時間と手間がかかるだけでなく重要なシーンを見逃してしまう恐れもある。また、映像やシーンに付加したテキスト情報で検索を行う場合、映像内容を文字や言葉で 100%表現することは難しいばかりか、キーワードを付加する人と検索キーを入力する人が異なるために同一映像に違うキーワードを連想してしまうケースもあり十分な検索ができない。

そのため、本システムでは DB に登録されたアーカイブ映像の内容を高速かつ分かり易く検索するため、画像の絞り込みを行うテキストベース検索と、さらに結果画像の内容把握を行うコンテンツベース検索の 2 段階で検索を行った(図 5)。

まず、テキストベースの検索については、DB 登録されたコンテンツのメタデータから登録日時、ID、キーワード、及びカテゴリ分類などにより絞り込み検索を行う。テキストベースの検索結果は代表サムネール形式(図 5)やリスト形式で一覧表示される。なお、サムネール上下にコンテンツの属性(MPEG-1, MPEG-2, JPEG など)や再生時間などが一目でわかるよう表示した。これらのサムネールの中からコンテンツを選択すると、コンテンツのタイトルやキャプションなどのメタ情報が表示され(図 6)、同一ウィンドウ内での再生確認も可能となっている。また、同時に映像のカット点等も表示されるため、以下で述べるコンテンツベース検索を用いてコンテンツ内を閲覧して目的のシーンを短時間で検索したり概要把握が可能である。

コンテンツベース検索に関しては、映像とオーディオトラックを時間軸上で効率的に検索する手段としてカット点ブラウズ表示(ショットブラウ

ジング)およびオーディオ分類表示を行った(図 7)。人物・顔画像検出[6]などの高度な映像検索もニュース検索など特定の用途には有効と考えられるが、カット点は各シーンを検索する上で基本であるとともに各シーンをブラウズするだけでも早送り再生などに比べて映像の検索効率はかなり向上すると考えられる。

カット点の検索については、瞬時に画面が切り替わる瞬時カットの他、ワイプやディゾルブ効果を用いながら画面が変わるトランジションカットにも対応し、MPEG の部分復号を用いて圧縮ドメイン上で検出処理を行う方式を用いた[3]。また、カット点位置など指定位置からさらに細かな内容をブラウズできるよう分や秒単位、フレームや GOP 単位などの周期時間インデックス表示を行うテンポラルブラウジングを可能とした。これによりマクロ検索とミクロ検索を複合したコンテンツベースの階層検索を可能とした。

オーディオの分類については、無音、音声、音楽、歓声雑音に分類し、各サムネール位置の分類情報としてシンボルマークで表示することにより、オーディオシーンの概要把握を可能とした。なお、音楽の上にナレーションが入るなど同時に複数の音が重畳された場合は最も支配的な音について分類される。本分類については、MPEG オーディオ圧縮データ上のサブバンドデータの統計的特徴量から、Bayes 決定における正規分布に対する最適識別関数を用いて分類する方式を用いた[4]。この分類は編集などの用途において目的の音声区間などをおおまかに検索するために利用できると考えられる。

カット点検出やオーディオ分類処理はコンテンツ登録時に自動的にを行い、検索閲覧時にはあらかじめ特徴抽出されて DB 上に保存されたメタデータ(図 8)を元に検索を行っているが、ブラウズ表示用サムネールの生成は検索時にリアルタイムで抽出処理を行っている。これにより、あらかじめ抽出して保存しておく場合に比べ、ディスク容量を節減している。

4.3 再生機能

MPEG-1, MPEG-2 などの数 Mbps 級の映像ストリーミングを既存のイントラネットに導入しようとした場合、同時接続ユーザ数に比例した帯域占有による他業務への支障が懸念されるばかりではなく、高価なビデオサーバやデコーダボードが必要となるケースもある。また、インターネットにおいては、ADSL や CATV インターネット等ブロードバンド化が進みつつあるが、上流ネットワークへの負荷や占有帯域を考えた場合、自由にストリーミングできる環境はまだ限られているといえる。このため本システムでは MPEG コンテンツを低解像度の代理映像(Proxy 映像)に変換して低コストなストリーミング配信によるイントラネット・インターネット上でのコンテンツプレビュー再生を

実現している。

代理映像としては RealVideo を用い、MPEG ファイルを登録時に PSTN(-56kbps)、ISDN(-128kbps)、LAN レベル(-500kbps)の3種類のビットレートに変換している。再生の種類としては、先頭タイムコードからの通常再生の他、IN点/OUT点をブラウザ上から指定した区間再生(図6下)、カット点検出で分割された各シーンの最初の数秒のみを集めた映像サマリを用いたダイジェスト映像再生(図9)がある。ダイジェスト映像再生を用いることにより、各カットを連続的にブラウジングできるほか、映像全体のおおまかな流れを動画で確認することができる。さらに、MPEG コンテンツをクライアント環境で直接再生、加工する用途を想定し、権限を持ったユーザがサーバから MPEG コンテンツをクライアント側にダウンロードできる機能を付加した。また、CM 等期限付きコンテンツの管理を目的として、管理者により設定された解禁日・有効期限内のコンテンツのみユーザからアクセスできるようにした。

4.4 リモート編集機能

これまで映像コンテンツを編集する場合、専用の編集装置で一元的に編集するか、サーバ上で共有管理しているコンテンツを一旦ローカル PC などにダウンロードして編集し、編集後のコンテンツを再度サーバにアップロードするような手段がとられていた。前者の場合、編集装置が高価なため時間的にシェアする必要があるなどの問題点がある。また、後者の場合、複数のコンテンツや長時間のコンテンツになるとファイルのダウンロードやアップロードなどで多くの伝送時間が必要になり、編集処理効率が大幅に低下する恐れがある。

本システムでは、このような問題を解決するため、ネットワーク上の DB で一元管理された MPEG 映像に対してクライアント側から WEB ブラウザ上で編集区間の指定を行い、サーバ側に搭載された編集エンジンを用いてリモート編集を行えるようにした。図10に具体的な編集プロセスを示す。

まず、DB 上で編集素材となる映像を絞り込み、(a)マクロ検索(ショットブラウザ)や(b)ミクロ検索(テンポラルブラウザ)により編集区間の設定を行い、必要に応じて編集区間の再生確認(c)後、編集リストに追加し(d)、編集登録処理を行う。複数の区間を接続する場合は同様の処理を該当の映像ファイルに対して行う。各編集区間の変更・削除、挿入も行うことができる。また、スプライスビュー機能により編集前に各編集区間の継ぎ目部分の繋がりが具合を代理映像を用いてプレビュー確認可能となっている。

なお、編集処理を行った履歴は SMPTE 準拠の EDL(Edit Decision List)として保存可能なため、他の編集システムとの相互運用時や編集作業中断・再開時に有効である。

編集処理アルゴリズムに関しては、MPEG ストリ

ームを完全復号せずにフレーム単位で高速に編集する部分トランスコード編集方式[5](図11)を用いており、これによりサーバ負荷軽減、画質劣化抑制、プロセス実行用キャッシュスペースの節約を実現している。

4.5 管理機能

登録された MPEG コンテンツは WEB/DB 上で一元管理され、複数のユーザからの同一コンテンツへの同時アクセス時に必要となる排他処理や、情報の更新、削除などのコンテンツ操作の他、ユーザレベルに合わせたアクセス制御(検索再生・編集登録、ダウンロードの各権限)を行っている。

5. まとめ

MPEG 画像の効率的な管理を可能とする映像管理システムを開発し、オーディオビデオ複合検索・閲覧・再生機能・ネットワーク上でのリモート編集・登録管理機能について述べた。これまでネットワーク上での MPEG コンテンツの加工や閲覧については1.で述べた問題点があったが、圧縮データ上での高速編集技術やカット点検索技術、代理映像の利用により、ネットワーク上でも MPEG コンテンツを有効利用することが可能となった。また、本システムはすべて Web 上で処理できるため、映像制作から教育現場まで様々な目的で利用することが可能である。

今後は、処理規模のスケラビリティや検索ツールの拡充、MPEG-4 や MPEG-7 への対応、ならびに編集処理の高度化などを行う予定である。

6. 参考文献

- [1] 倉橋 他、「コンテンツ制作効率化のための映像コンテンツ管理機能」、信学総大 D-11-74,pp.74,(2000.3).
- [2] 長谷川 他、「大容量アーカイブ活用型放送番組制作システム」、信学総大 D-11-89,pp.89,(2000.3).
- [3] 中島 他、「部分復号を用いた MPEG データからのカット点検出」、信学論(D-11),vol.J81-D-11,no.7,pp.1564-1575,(1998.7)
- [4] 中島 他、「MPEG 符号化データからのオーディオインデキシング」、信学論 D-II Vol.J83-D-II No.5 pp.1361-1371,(2000.5)
- [5] 中島、「MPEG ストリームの編集方法」、TECHI Vol.4 画像&音声圧縮技術のすべて CQ 出版社 pp.214-221,(2000.4)
- [6] Zhu Liu, Yao Wang, "Face detection and tracking in video using dynamic programming", in Proc. of IEEE International Conference on Image Processing 2000, pp.1-53-56, Sep. 2000.

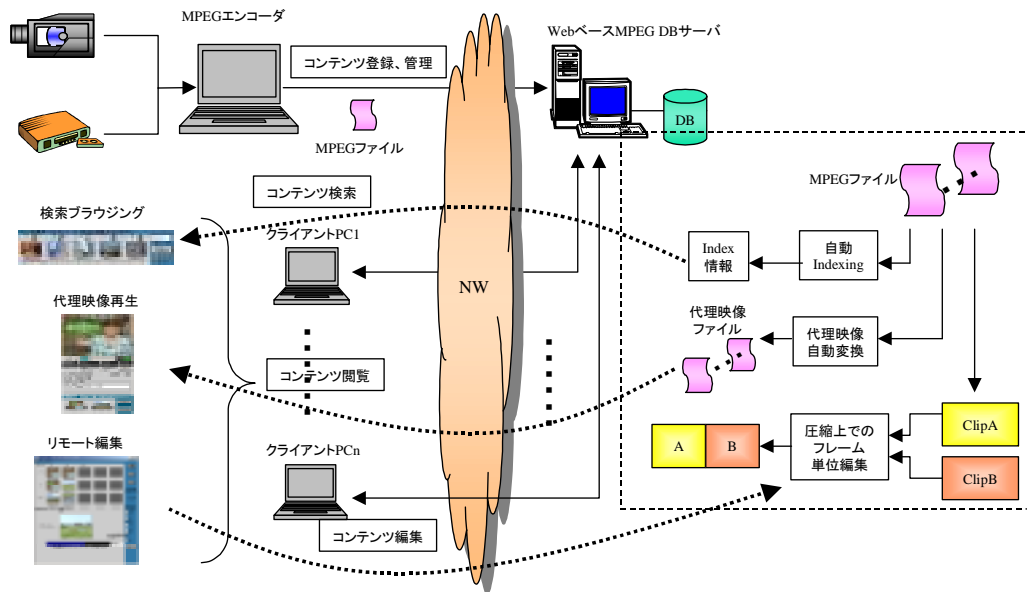


図1 システム構成図

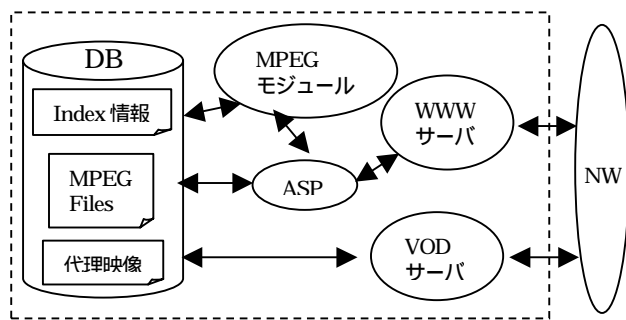


図2 サーバ構成 (最小構成 PC1 台)



図3 クライアントメイン画面



図4 登録機能



図5 検索結果表示

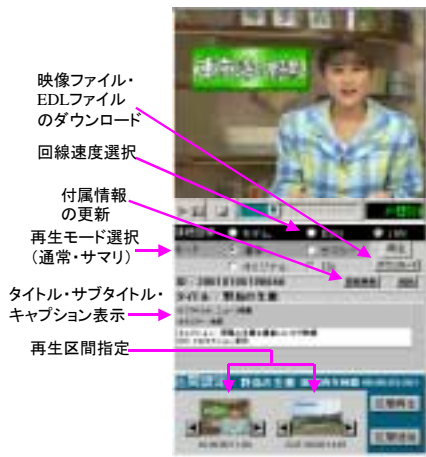


図6 再生画面



図7 コンテンツベース検索

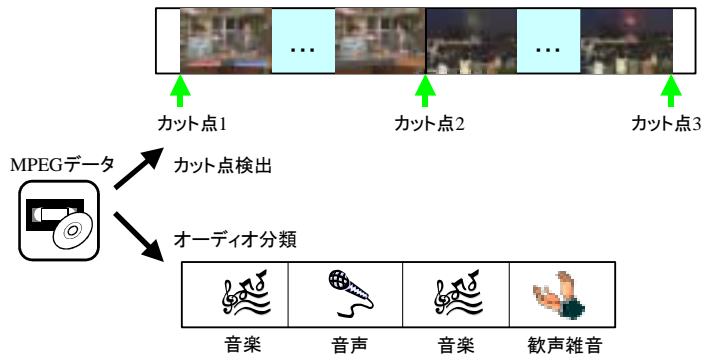


図8 MPEGデータのインデキシング

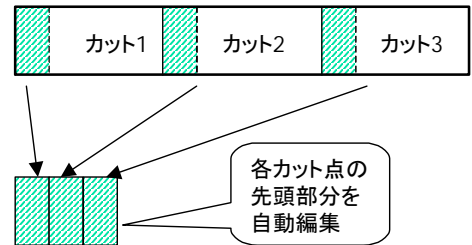


図9 サマリ再生



図10 編集プロセス

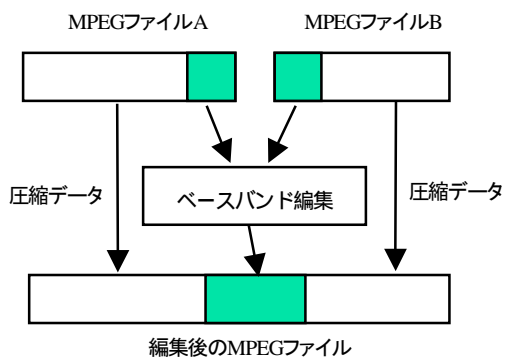


図11 部分トランス符号化によるフレーム単位編集