

D-23 MPEG-7を用いた情報提示システムに関する一検討 A Study on Information Presentation System using MPEG-7

宮崎 勝† 佐野 雅規† 金 淵培† 柴田 正啓†
Masaru Miyazaki Masanori Sano Yeun-Bae Kim Masahiro Shibata

1. まえがき

近年、メディアの多様化に従い、ユーザが欲しい情報を適切な形式で得ることのできる技術が求められている。しかし現存する様々なコンテンツデータベースに関しては、各々に応じたアクセス手段を用いて逐次的に情報を獲得し、その結果からユーザが目的の情報を選択する必要がある。このような手続きは非効率的であり、ユーザのアクセスするコンテンツを制限してしまう大きな要因ともなっている。今回、MPEG-7を共通記述言語として利用することで、複数の異種データベースの内容を統合し、ユーザに合わせたプレゼンテーションを生成できる情報検索・提示システムの試作・検討を行った。

2. システム構成

放送局等のコンテンツプロバイダには、番組制作に利用するための映像、音声および文書情報などを蓄積した数多くのデータベースが存在する。それらの多くはネットワーク経由で利用可能だが、各データベースへのアクセスには専用の端末や特定のポータルサイトの利用が不可欠であり、複数データベースを横断的に検索する機能は提供されていないのが実状である。そこで我々は、番組映像やニュース原稿のデータベース、スポーツの結果データベースといった複数の異種データベースを対象に、統合検索が可能なシステムを試作した。システム概念図を図1に示す。

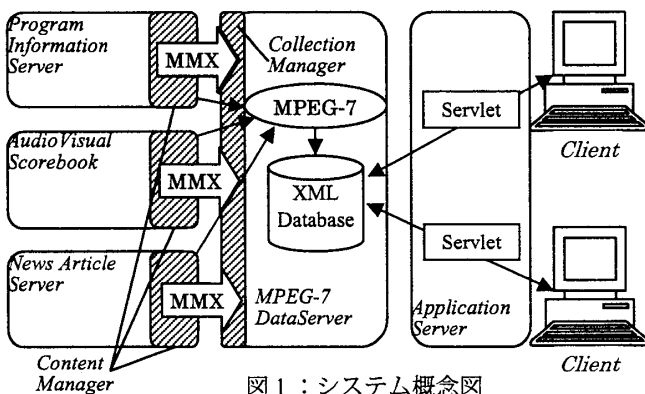


図1：システム概念図

試作システムは、番組情報データベース、映像版スコアブック^[1]、ニュース原稿データベースといったコンテンツサーバ群、MPEG-7 ファイルを収集・蓄積するデータサーバ、各種サービスのための機能を実装するアプリケーションサーバから構成される。

各コンテンツサーバにはコンテンツマネージャと呼ばれる Java アプリケーションがインストールされており、コンテンツサーバがコンテンツメタデータを XML 形式で出力する機能を備えていれば、出力された XML メタデータを

標準的な XSLT を用いて自動変換し、MPEG-7 準拠のメタデータを生成することが可能である。

また、MPEG-7 ファイルの生成と同時に、MPEG-7 メタデータの概要や更新日時といった情報を記述した MMX(MPEG-7 Meta XML)と呼ばれるインデックスファイルを生成し、MPEG-7 データサーバに送信する。データサーバ上の回収マネージャは各コンテンツサーバから受信した MMX ファイルの解析を行い、その内容に基づいて各コンテンツサーバから MPEG-7 メタデータを回収し、XML データベースに蓄積する。MMX ファイルにはメタデータの更新情報だけでなく、コンテンツサーバに関する情報(蓄積するコンテンツの種類、管理情報等)の記述も可能であり、メタデータの送受信だけでなく、様々なサービス提供に利用することができる。MMX ファイル記述例の一部を図2に示す。

```
<DataFile
entity=http://news01.nhk.or.jp/mpeg7/200111010009.xml
retrieval="yes">
<Title>
<FreeTextAnnotation>4日のニュース</FreeTextAnnotation>
</Title>
<Genre href="urn:mpeg:mpeg7:cs:GenreCS:2001:1.3.1">
<Name>DailyNews</Name>
</Genre>
<Version>1.00</Version>
<LastUpdate>2002-07-25T20:20:51</LastUpdate>
</DataFile>
<DataFile>...<DataFile>
<DataFile>...<DataFile>
```

図2：MMX ファイルの記述例

アプリケーションサーバ上ではサービスを提供するサーバレットが複数動作しており、ユーザは汎用的なインターネットブラウザを用いて目的に応じたサービスサイトにアクセスすることで、情報を検索・獲得することができる。

以上の枠組みにより、既存の各種データベースシステムを大幅な仕様の変更なしに本システムに接続することができ、データの自動収集から統合検索までの処理が行えるようになる。

3. 情報の MPEG-7 表現

今回は「大相撲」に関する情報を対象とし、複数データベースから得られる相撲関連情報(ニュース・番組・取り組み結果)をブラウザ上に効果的に表示する手法について検討を行った。

試作システムでは、MPEG-7を共通記述言語として情報統合に利用しているが、ニュース原稿データ、番組データ、取り組み結果データという性質の異なった情報を完全に同一のフォーマットで記述することは困難である。むしろ

† NHK放送技術研究所

様々なサービスにおいて最大限に活用できるよう、各データの性質に合った3種類の MPEG-7 構造を利用して記述を行った。また、コンテンツ内容を自然言語で記述する部分では、重要単語を抽出して記述した KeywordAnnotation タグ、文書の依存構造を記述した DependencyStructure タグ、5W1H の情報を記述した StructuredAnnotation タグという3種類の記述方法を共通して利用し、様々な検索手法への対応を考慮した。

4. 検索処理

本試作システムの検索手法としては、日付範囲を指定したキーワードマッチングを利用している。MPEG-7 データサーバの内部では、MPEG-7 ファイルを作成された月ごとにディレクトリ分けしており、キーワードおよび日付範囲を指定することで、該当ディレクトリ内のファイル群に対してキーワード検索のクエリを適用する。検索対象とするタグは MPEG-7 ファイルの種類毎にあらかじめ決めておき、検索キーワードが存在したファイルを検索結果としてリスト表示することとした。

5. 検索結果表示

システムのメイン画面は、主に検索結果表示部、関連情報表示部、および関連ニュース表示部という3つのエリアから構成される。検索結果の表示例を図3に示す。

検索結果表示部には、目的のキーワードを含んだ3種類の MPEG-7 ファイル（ニュース原稿、番組情報、取り組み結果情報）が時系列で表示される。この際、結果リストの中の最新取り組み情報ファイルの内容が関連情報表示部に星取表の形式で展開され、最新ニュース原稿ファイルの内容が関連ニュース表示部に展開表示される。

関連情報表示部に関しては、MPEG-7 の構造を利用することで、ユーザの入力したキーワードの内容に応じた表示形式の変更が可能である。例えば、ユーザが「琴光喜」「栃東」という2人の力士名で and 検索を行った場合、両キーワード共に取り組み結果 MPEG-7 ファイルの Who タグの内容と一致する。このことから、システムは入力された2つのキーワードの内容が「力士名」だと判断し、2人が対戦した最新の星取表だけでなく、過去の取り組み MPEG-7 ファイルの内容を取得することで、2人の過去の対戦成績を一覧表示することができる。

6. ユーザプロフィールの利用

本試作システムは、ユーザの検索操作やコンテンツ視聴の履歴を、MPEG-7 の UsageHistoryDS を用いて逐次蓄積している。蓄積された履歴情報からはユーザの各データベースやジャンルに関する嗜好が計算され、その結果はログイン時の最新情報提供等に利用できる。ユーザが特定コンテンツサーバの特定ジャンルの映像コンテンツを視聴する機会が多かった場合、次のログイン時には、そのデータベースの最新情報を優先的に検索結果表示部にリスト表示するといったプッシュ型のサービスを実現している。現在利用しているのは主にデータベースアクセスの頻度情報であるが、今後、ユーザが入力した検索キーワード等の解析により、より高度な情報提供が可能になると考えられる。

7. おわりに

MPEG-7 を利用した番組情報提示システムの試作により、MPEG-7 準拠のメタデータに基づいた異種サーバの統合検索が実現可能であることが検証できた。今後は外部のオントロジなどを利用した表示形式のダイナミックな生成などを検討していく予定である。

参考文献

[1]佐野他:”映像版スコアブックの提案”
映像情報メディア学会 マルチメディア情報処理 映像学技報, Vol.25, No.21, pp.7-12 (Mar.2001)

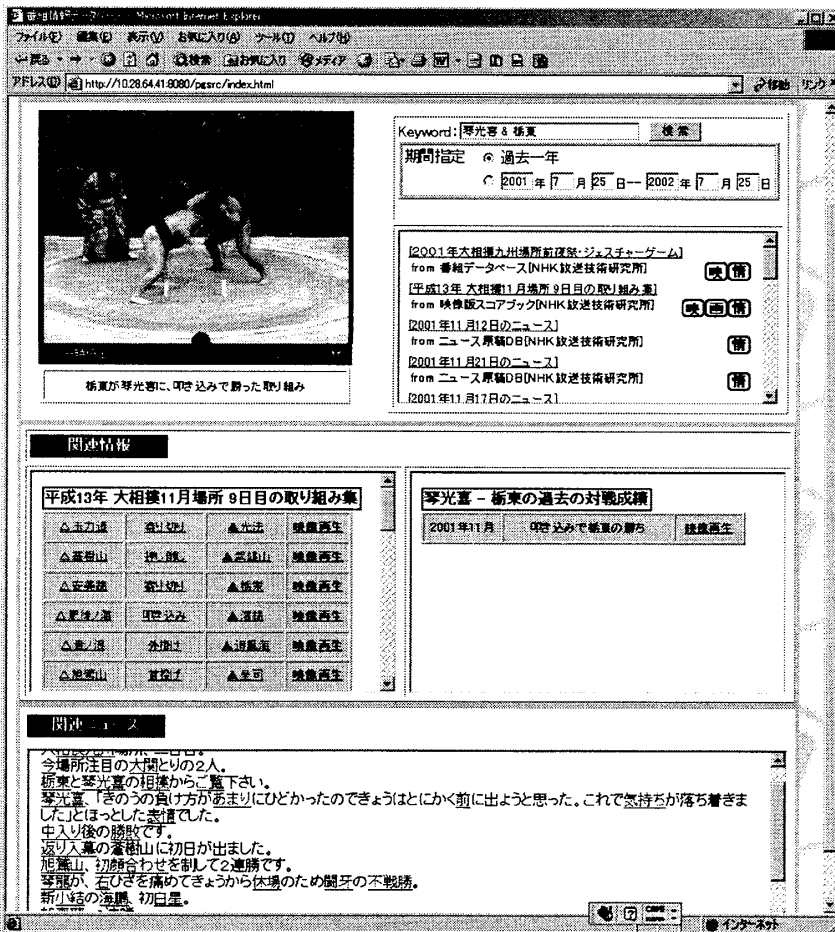


図3：番組情報提示システムの画面表示例