

# 見たいシーンにすぐアクセスできる番組ナビゲーション技術

Navigation System for News Programs Featuring Direct Access to Desired Scenes

上原 龍也  
UEHARA Tatsuya

堀川 将幸  
HORIKAWA Masayuki

住田 一男  
SUMITA Kazuo

デジタル化の進行により、放送の多チャンネル化が実現している。これに伴い視聴者が見たい番組も増加し、効率よく番組を見る方法が必要となる。この問題を解決するため、ニュースなど情報系番組に対して、視聴者が興味のあるシーンにすぐにアクセス可能な番組ナビゲーションパイロットシステムを試作した。

このシステムでは、番組内の映像情報と音声情報を認識処理し、認識結果を統合することにより、情報系番組をトピック単位に分割し、各々に対し自動的にキーワードを付与する。その結果、番組内容の一覧表示及び検索が可能となり、内容を見ながら番組内をブラウズ(必要なものだけ中身を見ていく)できるようになる。

A digital broadcasting system simultaneously transmits large amounts of program content. Although TV viewers can select their favorite programs, they do not have sufficient time to watch all of them.

To solve this problem, we have developed a navigation system for news programs which allows viewers to directly access desired scenes. The navigation system recognizes both visual and aural information in a news program, and integrates these recognition results. The system then divides the program into news topics and adds keywords, which help viewers in selecting desired scenes, into each topic.

## 1 まえがき

近年、BS(放送衛星)放送及び地上波放送のデジタル化が進行している。デジタル化により、映像の高画質化とともに、チャンネル数が大幅に増加する。既に、デジタルCS(通信衛星)放送では、百数十ものチャンネルを楽しめるようになっている。放送の多チャンネル化により、様々なジャンルの番組を放送できるようになるため、視聴者が好きな番組をいつでも視聴できるようになる。しかし、多チャンネル化によって、以下のような問題も生じている。

- (1) チャンネル数が膨大なため、従来の新聞のテレビ欄のような一覧表示が不可能になり、視聴者が好みの番組を検索するのが困難になった。
- (2) 視聴者が見たい番組も増加するため、従来のテレビ視聴に充てていた時間では、見たい番組を見切れなくなった。

第1の問題点については、視聴者の過去の視聴傾向から好みの番組を推測し、お勧めの番組表を作成する方法が研究されている<sup>(1)</sup>。

一方、第2の問題点を解決するためには、番組を効率的に視聴する方法が必要となる。特に、ニュースなどの情報系番組では、視聴者が必要とするのは、特定の話題の情報であることが多い。シーケンシャル(連続的)に放送される現在のニュースでは、冒頭は政治や事件関連のニュースから始まることが多く、プロ野球の結果を知りたいとしても、

すぐには見ることはできない。

ここでは、第2の問題点を解決するために、録画した番組に対して一覧表示・検索を可能にした番組ナビゲーションシステムについて述べる。

## 2 システムの概要

ニュースなどの情報系番組は、複数のニューストピック(以下、トピックと略記)から構成される。このシステムでは、まず、あらかじめ録画された番組に対しトピックごとに分割し、それぞれのトピックに対しインデックスを付与する。この処理をここでは“番組構造化処理”と呼ぶ。次に、番組構造化処理されたデータに対して、視聴者が自由に番組内容を参照できるよう番組ブラウザを提供する(図1)。

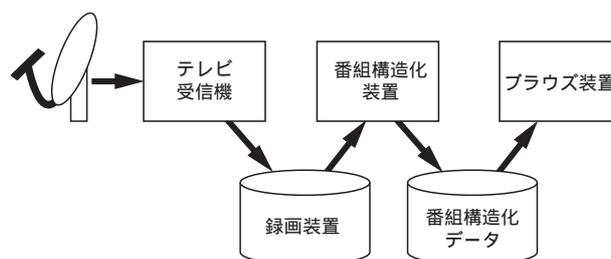


図1. 番組ナビゲーションシステムの概要 ナビゲーションシステムは、番組構造化装置とブラウザ装置から成る。

Schematic diagram of navigation system for news programs

### 3 番組構造化処理

番組構造化処理は、以下の二つのプロセスから成る。

- (1) トピック分割処理
- (2) インデックス付与処理

#### 3.1 トピックの分割

番組をトピックに分割する方法としては、従来アナウンサーが映ったシーンをキーとして分割する方法が提案されてきた<sup>(2)</sup>。しかし、この方法では、フラッシュニュースのようにアナウンサーが登場しない場合や、ワイドショーなどの番組に対しては、ニュースの分割を行うことができない。

この研究では、トピック分割についての手がかりをあらかじめニュース番組内に埋め込むことで、上記の問題を解決した。この手がかりとして、テレビ朝日映像(株)と共同開発したビジュアルコードを利用した。ビジュアルコードは、映像中のテレビの画面としては映らない部分にテロップを使って情報を埋め込む方式で、送信・編集機器など従来の機材で実現可能であることを特長とする。トピックの分野情報をトピックの最初のテロップと同期して、ビジュアルコードを挿入している。

番組製作者の手間を軽減させるため、ビジュアルコードは、トピックの先頭ではなくテロップと同期して挿入されている。したがって、トピックの先頭位置は、映像中に認識されたビジュアルコード出現位置を基準にして推定する必要がある。このため、アナウンサーの音声の切れ目を利用することにより、トピック分割位置を決定している。

#### 3.2 インデックス付与

次に、前節で述べた手法により分割された各々のトピックに対し、視覚的なインデックスとしての代表画面と、言語的なインデックスとしてのキーワードを付与し、視聴者のトピック選択の手助けとする。

**3.2.1 代表画面生成** トピックを代表する画面の選択を行う。一般にニュースでは、トピックの構造は、アナウンサーが登場し、その後で現場の映像が登場する。アナウンサーが登場する画面を選択したのでは、インデックスとして意味を持たないので、まず、トピック中で、アナウンサーが登場する以外の位置を検出し、その位置から数秒後の映像を代表画面とする。この理由は、現場映像の冒頭部分は事件などの場所情報を表すことが多く、この部分を避け内容を表す映像を抽出するためである。

**3.2.2 キーワード抽出** このシステムでは、トピックからキーワードを抽出するため、トピック中のテロップの情報とアナウンサーなどの音声情報からキーワードを抽出する。テロップの情報を用いるために、映像からテロップ位置を検出した後、背景との分離を行い、文字認識を行う<sup>(3)</sup>。また、音声情報についても音声認識技術により、内容を文字にしている。しかし、実際のニュースでは、テロップがスライドな

どの動作を伴う場合や、文字に飾りや色がついているために、認識が完全ではない。また、音声認識に対しても、BGM (BackGround Music) や、言いよどみなどが存在するため、認識率は高くはない。

このような認識誤りに対処するため、テロップ認識の結果と音声認識の結果とビジュアルコードによって示される分野情報に加え、出現頻度などを考慮して確からしい語句をキーワードとして採用する(図2)。

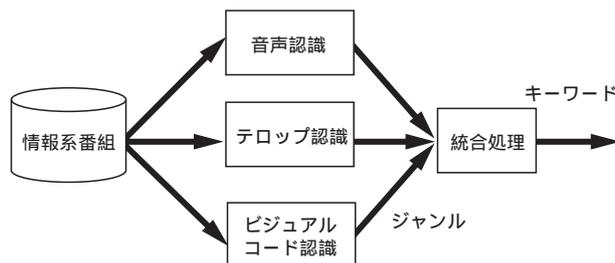


図2. キーワード抽出処理 音声、テロップ、ジャンルの三つの情報を統合してキーワードを抽出する。

Process of keyword extraction

出現頻度を利用することにより、テロップ認識の結果と音声認識のどちらか一方の認識で処理ができない場合でも、認識できた方で複数回出現すればキーワードとして採用される。ただし、ことばによっては、テロップと音声で違うことば(“首相”と“総理大臣”)を使うことがあるので、頻度計算の際には同一のことばとして扱う必要がある。

また、分野情報を利用することにより、認識結果を改善することができる。例えば、音声認識では同音語を正しく認識できない。“とうしゅ”という発音に対しては、野球のトピックなら“投手”、政治のトピックなら“党首”であることが多いが、音声認識ではこれらを区別することはできない。このシステムでは、分野別辞書を用いることにより、分野情報から正しい表記を得ることが可能である。

### 4 番組ブラウザ

3章で述べた番組構造化技術を用いることにより、情報系番組に対して、トピックごとのインデックスを作成することができる。番組ブラウザは、このインデックスを利用して、視聴者に自由に番組内のトピックへのアクセスを実現する。

番組ブラウザは、次の五つのモードを持っている。

- (1) 番組選択モード 録画された番組から表示したい番組を選択する。
- (2) トピックブラウズモード 一つの番組内のニューストピックの代表画面を一覧表示し、自動的に代表画面を



図3．番組ブラウザの表示例 正面に表示したトピックに対して各種情報を表示する。  
Example of navigation system display

順送りにして表示する。

- (3) ヘッドラインモード 一つの番組内のニューストピックの代表キーワードを一覧表示し、選択させるモードである。代表キーワードのカーソル移動に伴い、代表画面も移動する。
- (4) ジャンルソートモード ニューストピックを8分類のジャンルに分類し、一覧表示する。ジャンルのカーソル移動に伴い、代表画面も指定されたジャンルのトピックへ移動する。
- (5) 類似検索モード 各トピックには、キーワードが付与されている。付与されたキーワードどうしの類似度を求めることにより、類似したトピックを検索することが可能となる。このモードでは、指定したトピックに対して、過去に録画した番組中のトピックから類似したものを表示する。この機能を用いることにより、ある事件の過去の経緯を一覧表示したり、更に、リモコンのようなキーワード入力が必要な場合でも、検索可能となる。

操作画面の例を図3に示す。中央に表示された映像が各トピックの代表画面であり、右上には正面に表示されているトピックのキーワードが表示されている。また、左側にはメニューが表示され、ジャンルソートモードでは、各ジャンルを表示する。視聴者は自分の見たいジャンルへジャンプし、キーワードと代表画面を見ながら、検索することができる。好みのトピックが見つければ、そのトピックを全画面で再生することができる。

## 5 あとがき

番組内の映像情報と音声情報を認識処理することにより、視聴者が番組内の内容に自由にアクセスできる番組ナビゲーションシステムについて述べた。この技術は、放送された番組だけに限定されるものでない。今後は、この技術を映

像ライブラリの検索に応用する予定である。

## 謝 辞

この研究は、通信・放送機構の委託研究“ユーザ適応型インタラクティブコンテンツに関する研究開発”として行われたものであり、研究の機会を与えていただいた通信・放送機構、及び共同で研究していただいたテレビ朝日映像(株)に対し、感謝の意を表します。

## 文 献

- (1) 矢川雄一,ほか.個人の嗜好に合ったテレビ番組を自動編成するエージェントの検討.電子情報通信学会 人工知能と知識処理研究会.98,55,1998,p.9-16.
- (2) 有木康雄,ほか.ニュース映像中の記事に対する音声・文字・映像を用いた検索付けと分類.電子情報通信学会 パターン認識・メディア理解報告.96,97,1996,p.31-38.
- (3) 堀 修.テロップ認識のための映像からの文字部抽出法.情報処理学会コンピュータビジョンとイメージメディア.99,114,1999,p.129-136.



上原 龍也 UEHARA Tatsuya

研究開発センター マルチメディアラボラトリー研究主務。自然言語処理、音声対話の研究・開発に従事。人工知能学会会員。  
Multimedia Lab.



堀川 将幸 HORIKAWA Masayuki

研究開発センター ヒューマンインターフェースラボラトリー。ヒューマンインタフェースの研究・開発に従事。  
Human Interface Lab.



住田 一男 SUMITA Kazuo, D.Eng.

研究開発センター ヒューマンインターフェースラボラトリー主任研究員、工博。自然言語処理、ナレッジマネジメントの研究・開発に従事。情報処理学会、電子情報通信学会会員。  
Human Interface Lab.