

クローズドキャプションを利用したテレビ番組の話題分割技術

番組の見た部分だけを選び、効率的に視聴

録画機器の大容量化や多チャンネル化が進むことにより、録(と)りためたテレビ(TV)番組の中から見たい箇所だけを選んで視聴する、“つまみ見”のニーズが高まっています。

東芝は、こうしたニーズに応え、番組の字幕テキスト(クローズドキャプション)を利用して、一つの番組を話題ごとに自動分割し、各話題の内容を簡潔に示すテキスト情報を自動抽出する技術を開発しています。話題分割では、話題の転換を表すことばに着目した方法と、語彙(ごい)の推移に着目した方法の2通りの方法を用いて、番組を高精度に自動分割できます。

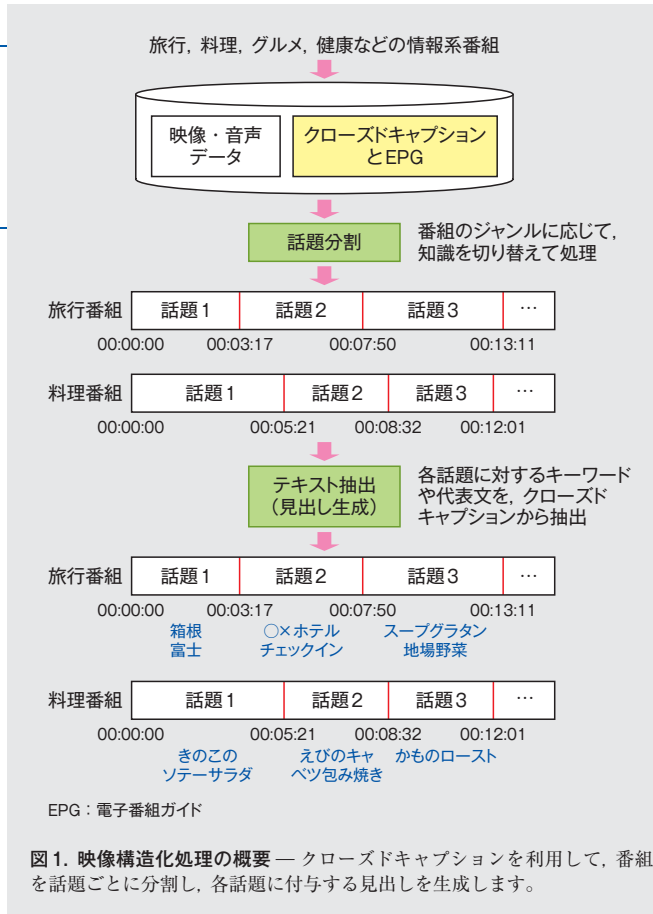


図2. 話題分割とテキスト抽出結果の画面例 — 話題ごとにサムネイルと見出しのテキストを付けて提示しています。選択中の話題に対しては、複数のサムネイルを表示しています。

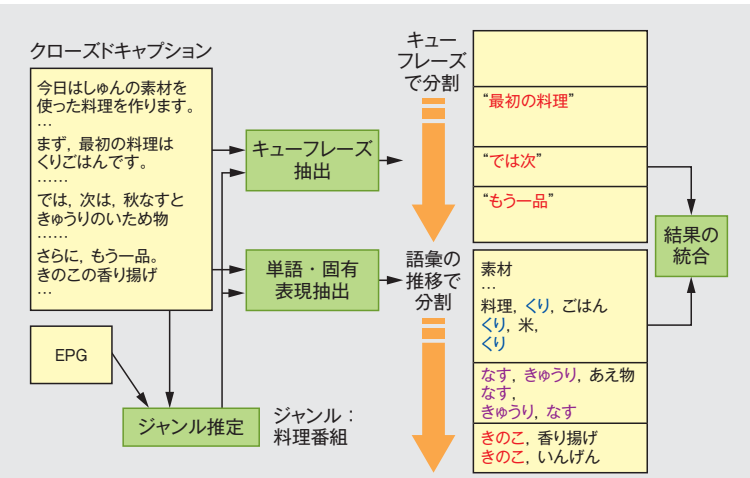


図3. 話題分割の処理フロー — 2通りの方法で番組を分割し、最終的に結果を統合して話題の切れ目を決定します。

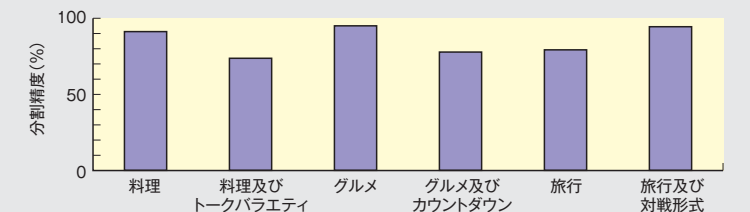


図4. 話題分割の精度 — 自動分割の分割精度は、6ジャンル平均で約80%となっています。

録画したTV番組の“つまみ見”

ハードディスク(HDD)の大容量化やTV放送の多チャンネル化により、パソコンやHDDレコーダなどで録画・蓄積できるTV番組の容量は増加しています。しかし、個人が映像の視聴に費やせる時間は限られています。今後、録画機器の大容量化が更に進むと、録画した番組の中から視聴者が見たい部分をすばやく探すニーズが高まると考えられます。

そこで、番組のクローズドキャプションを利用して、一つの番組を話題ごとに分割し、各話題の内容を簡潔に示すテキストを抽出する技術を開発しています。分割した結果を、各話題の内容を示すテキストとともにユーザーに提示することで、番組の中から見たい話題だけを選んで視聴できる“つまみ

見”の環境を実現します。例えば旅行番組の中から、ユーザーが興味のあるレストラン情報の部分だけを視聴できるようにすることを目指しています。

話題分割に基づく映像構造化

話題分割に基づく映像の構造化処理の概要を図1に示します。番組のクローズドキャプションを利用して、番組を話題ごとに自動分割し、分割された各話題の内容を簡潔に示すキーワードや代表文を抽出して、テキストの見出し情報を自動生成します。

このシステムにより話題分割とテキスト抽出を行った結果の画面例を図2に示します。この例では、旅館での夕食や観光スポットの散策など旅行番組が話題ごとに分割されています。画面には、各話題に対応するサムネイル^(注1)と見出しのテキストが表示され、ユー

ザーは、これらから自分の興味のある話題を見つけて、番組の中を“つまみ見”できます。

話題分割技術の仕組み

キー技術である話題分割処理の概要を図3に示します。番組開始からの経過時間を示すクローズドキャプションのタイムスタンプ(時間情報)を対象とし、二つの方法を用いて話題の切れ目を検出します。

一つは、“ところで”や“では次”のような明示的に話題の切替わりの手がかりとなる語を利用した、キューフレーズに着目する方法です。もう一つは、使われる語彙の推移に着目する方法です。例えば、旅行番組で観光地Aから観光地Bに話題が変わったときに、語

(注1) 多数の画像を一覧表示するために縮小された画像。

彙の推移による方法では、番組の中で使われる地名の変化から、話題の切れ目を検出できます。最終的に、二つの方法の結果を統合して、話題の切れ目を決定します。

評価結果

情報系番組の6ジャンル、25コンテンツに対して実施した、話題分割精度の評価結果を図4に示します。自動で話題分割した結果と、人による分割結果との一致度で評価しています。1番組に対してふたりの人間が行った分割結果を用い、次式で人どうしの分割の一致度に対する相対値を計算します。

$$\text{分割精度} = (\text{自動分割と人による分割の一致度}) / (\text{人どうしの分割の一致度})$$

分割精度は、6ジャンルの平均で約80%という結果でした。放送されて

いる番組の数に対して、評価データ数はまだまだ少ないですが、人による分割と比べて大きく見劣りしない結果となっています。

関連技術

ニュース番組の話題分割をはじめとして、TV番組の構造化については、これまで多くの研究が行われています。画像や音声のメディア処理に基づくアプローチが主流ですが、多くの場合、ニュース番組や特定のスポーツ番組など限られた番組に特化したものでした。

現在開発中の話題分割技術では、クローズドキャプションに言語処理技術を適用することにより、旅行や、料理、グルメ、健康など多様なジャンルの情報系番組に対し、ジャンルごとに言語知識を切り替えて話題分割ができ

ます。更に、テキスト抽出技術により分割した各話題に対して、内容を示すテキストの見出しを自動で付与することもできます。

今後の展望

番組の対象ジャンルを拡大しつつ、いっそうの精度向上に向けて、メディア処理との融合を進めていきます。内容や形式が多岐にわたるTV番組に対応するためには、テキスト処理や画像・音声処理を組み合わせ、適切に処理することが重要だと考えています。

現在、早期の製品化を目指して研究・開発を進めています。

小山 誠

研究開発センター
知識メディアラボラトリー