

番組シーン再生のための 字幕情報を用いた検索技術

Scene Detection and Scene Playback Technologies Using Closed Captions of TV Programs

山下 道生

■ YAMASHITA Michio

テレビやレコーダで複数チャンネルをまるごと録画し、いつでも番組を視聴できる機能や大量に録画された番組の中からユーザーの好みに合う番組を見つけ出す機能が普及してきたが、更に見たいシーンだけを効率的に、かつ簡単に視聴できる機能が望まれるようになってきた。

そこで東芝は、録画番組中の字幕情報を用いてユーザーの好みのシーンを検索し、そのシーンだけを視聴できる機能を開発し、“レグザサーバー”DBR-T460/T450に搭載した。これは、タレント名などユーザーが入力した好みのキーワードが登場するシーンを全録画番組の中からリストアップし、シーンの“頭出し再生”をする機能である。更に、視聴中の番組に関連した話題キーワードを抽出する機能を実現し搭載した。話題キーワードをきっかけにして新しいシーンに出会えるだけでなく、シーンの話題が一目で把握でき、番組をより楽しむことができる。

Toshiba has been supplying REGZA series liquid crystal display (LCD) TVs and REGZA Server recorders to the market incorporating the following functions: (1) the "Time-Shift Machine" function, which enables viewers to watch programs of interest at any time; and (2) the "Zanmai-Play" function, which introduces viewers' preferred programs from among recorded programs using the Time-Shift Machine function. As users have recently tended to spend less time viewing TV programs due to the growing diversity of entertainment resources, demand is also increasing for a function that will allow them to easily access scenes that they really wish to see.

In response to these sophisticated requirements, we have developed the DBR-T460/T450 REGZA Server recorders featuring scene detection and scene playback functions that can create a list of scenes related to a viewer's keywords such as the names of entertainers from the closed captions in all recorded programs and play these scenes, as well as a scene-related keyword detection function that can automatically extract topic keywords related to viewed program scenes using closed captions. These functions offer a new TV program viewing style that saves users' time and effort.

1 まえがき

好きなチャンネルを何日もの長期間にわたり、まるごと録画し続ける“タイムシフトマシン”は、番組を録画し忘れることなく、見たい番組をいつでも見ることができる自由な視聴スタイルを実現した。更に、デジタルハイビジョン液晶テレビ〈レグザ〉Z7シリーズ以降に搭載している、ユーザーの好みなどいろいろな切り口で番組を探し出す“ざんまいプレイ”は、お気に入りの番組はもちろんのこと、今まで見逃していたおもしろい番組をも発見できるようにした⁽¹⁾。

しかし、テレビ視聴時間の制約や視聴スタイルの多様化に伴って、より効率よく、より短い時間で、より簡単に番組を楽しむことが求められるようになった。

これらの要望に応えるため、レグザクラウドサービス「Time-On」(以下、TimeOnと略記)では、録画番組の中の気になるシーンだけを探し出して視聴する“みどころシーン再生”を実現した⁽²⁾。これは、番組のシーン情報と、クラウドサービス上のサーバを利用して実現している。しかしTimeOnを利用するには、インターネットへの接続が必要なことや、全国放送と放送時間や内容が異なる番組では未対応などの課題があった。

そこで新たに、ネットワーク接続なしに、放送波に含まれる字幕情報を利用することで、録画番組中の見たいシーンだけを検索してテレビを楽しめる技術を開発し、“字幕シーンジャンプ”機能として“レグザサーバー”DBR-T460/T450(以下、T460/T450と略記)に搭載した。

字幕情報を用いたシーン検索は、ユーザーの気になる好きなタレント名や気になる用語などをキーワードとして入力すると、大量の録画番組の中から出演者が話す瞬間など、そのキーワードがせりふとして登場するシーンの開始時刻を見つけ出し、“頭出し再生”を可能にする機能である。

更に、リモコンで簡単にシーン検索できるようにするために、今まさに視聴しているシーンに関連する話題キーワードを抽出する機能を、字幕情報などを利用して実現した。抽出した話題キーワードは、シーン検索に用いるキーワードとして使用できる。この機能は、話題キーワードを一目見るだけで番組内容を把握できるなどの効果もあり、字幕シーンジャンプの利便性をより高める機能である。

字幕情報は、出演者のせりふやバックグラウンドミュージック、効果音などを文字や記号で画面上に表示するための情報であり、聴覚障がい者や高齢者への配慮はもちろん、一般視聴者

が番組をより理解するために利用できる。総務省が発表した「平成24年度の字幕放送等の実績」⁽³⁾によると、「視聴覚障害者向け放送普及行政の指針」の普及目標の対象となる放送番組における字幕番組の割合は、NHK（総合）で83.5%（前年比+12.9%）、在京キー5局^(注1)で93.3%（+2.5%）、全国の系列ローカル局^(注2)で66.4%（+2.4%）である。字幕放送は更なる充実に向けた取組みが推進されており、今後字幕放送が増えることでこの技術の有効性は更に増していくと期待できる。

ここでは、前述した機能について、概要と技術的背景を述べる。

2 シーンのリストアップと頭出し再生

字幕シーンジャンプを実現するための、字幕情報を利用したシーンのリストアップと頭出し再生の処理技術について述べる。

2.1 字幕情報の取得と保存

放送波の字幕情報を取得するため、タイムシフトマシンで指定された番組ストリームデータを保存すると同時に、番組ストリームデータから字幕情報を抜き出し、番組情報や字幕情報表示時間などの情報とともにHDD（ハードディスクドライブ）に保存する処理を追加した。

2.2 キーワード入力とシーンのリストアップ

字幕情報の中からシーンを見つけ出すには、ユーザーが入力するキーワードを利用する。HDDに保存している字幕情報の中からキーワードと一致するシーンを探し出し、それをシーンの開始とする。同一番組の短い時間範囲に大量のシーンが見つかることがあるが、これらは内部処理により同一のシーンとみなす。最終的に複数のシーンが見つかった場合は、シーンの開始時刻で並び替え、シーンリストとして表示する。

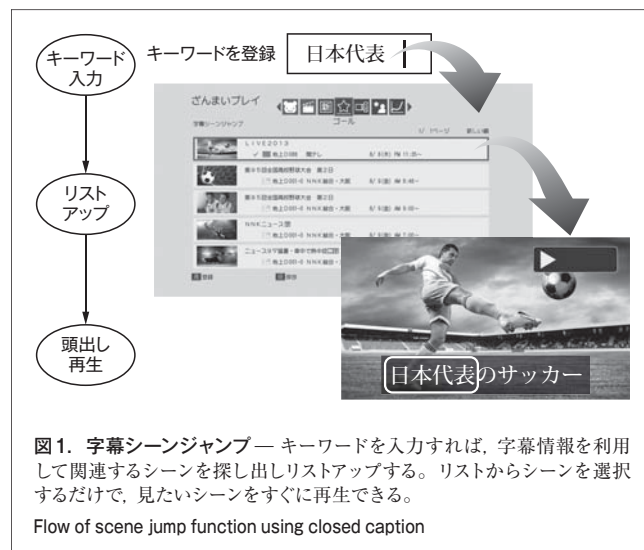
シーンリストには番組タイトルや放送局名、シーン開始時間、サムネイル（字幕が表示される画面の縮小画像）などが表示される。これらの情報を元に、ユーザーはどのリストを選べばどのようなシーンが再生されるかを、ある程度推測できる。

2.3 頭出し再生

リモコンの上下ボタンでシーンリストから好みのシーンを選ぶと、シーン開始時間での頭出し再生が始まる。シーン再生中は通常の再生と変わらない操作が可能である。リストアップされた全てのシーンを効率よくチェックできるようにするため、シーン再生中にリモコンの次スキップボタンを押すことにより、シーンリストの次のシーンを頭出し再生する機能を実装した。これにより、シーンリスト表示に戻ることなくワンボタンで次々と再生できる。

(注1) 日本テレビ放送網(株)、(株)TBSテレビ、(株)テレビ朝日、(株)フジテレビジョン、(株)テレビ東京。

(注2) 在阪準キー4局（(株)毎日放送、朝日放送(株)、讀賣テレビ(株)、関西テレビ放送(株)）及び在名広域4局（中部日本放送(株)、東海テレビ放送(株)、名古屋テレビ放送(株)、中京テレビ放送(株)）を除く101社。



2.4 ざんまいプレイへの組み込み

字幕シーンジャンプをざんまいプレイに組み込んで、今までと変わらないユーザー操作で利用できるようにした。ざんまいプレイは、EPG（電子番組表）情報を元にユーザーの嗜好（しこう）情報を考慮し、様々な切り口で好みの番組を見つけ出す機能である。そこに今回開発した字幕情報を用いたシーン単位での検索を選択できるようにすることで、キーワード入力やシーンのリストアップ、頭出し再生という字幕シーンジャンプの一連の操作の流れを、ざんまいプレイで実現した（図1）。シーンのリストアップには、検索キーワード以外に、除外キーワードや番組ジャンルなども検索条件として指定可能である。

2.5 操作性を高めるデータ処理技術

シーンリストアップ時の操作性を向上させるため、いち早くシーンを検出するデータ処理として、次の三つの技術を実装した。

2.5.1 字幕情報関連データによる高速化 特殊文字変換や単語途中切れ処理、表示タイミング吸収など、字幕情報の特殊性を考慮したデータ処理を行い、更にどこにどのデータがあるのかを索引付けして保存する。これにより、大量の字幕情報へのデータアクセスが高速化できる。

2.5.2 検索処理の最適化 検索処理自体を高速化するのはもちろんのこと、シーン数が多いキーワードほど検索処理に時間が掛かるため、リストの最大シーン数を取得した時点で検索処理を打ち切るようにした。

2.5.3 検索結果の再利用による高速化 一度検索したキーワードのシーンリストを内部メモリに保存しておくことで、同一キーワードでの再検索時には字幕情報の更新分だけを検索処理すればよい。そして、更新分のシーンリストと内部メモリに保存したシーンリストを統合することで、検索処理の大半を省くことができる。

2.6 番組検索とシーン検索の例

従来の番組検索と、今回実現したシーン検索の違いの一例

表1. シーンリストの例
Example of list of scenes

	チャンネル	番組		シーン	
		タイトル	開始/終了時刻	開始時刻	概要
S1	A	NEWS1	19:00 ~ 19:30	19:25	CD売上
S2	C	NEWS2	17:30 ~ 18:00	17:42	イベント参加
S3	A	MUSIC1	15:05 ~ 16:35	15:07	出演者紹介
S4	A	MUSIC1	15:05 ~ 16:35	15:22	曲歌い終わり
S5	A	MUSIC1	15:05 ~ 16:35	15:56	インタビュー
S6	A	NEWS3	13:00 ~ 13:05	13:03	イベント参加
S7	A	NEWS4	12:00 ~ 12:15	12:08	CD売上
S8	B	NEWS5	11:30 ~ 11:45	11:34	イベント参加

を挙げる。三つの地上デジタル放送チャンネルを1日分(24時間)録画した環境で、あるアイドルグループ名を検索したところ、EPG情報による番組リストでは2件しか番組が見つからなかったが、字幕情報によるシーンリスト(表1)では8件のシーン(S1~S8)が見つかった。

S1, S2, S6~S8の5件はニュース番組で見つかったシーンである。そのうちS1, S7の2件はCD売上に関するシーンを検出しており、同じチャンネルの昼と夜に報じられた。残りS2, S6, S8の3件はアイドルグループが参加したイベントに関連するシーンで、三つの放送局がそれぞれ別の視点で同じイベントを報道したものである。

S3~S5の3件は音楽番組で見つかったシーンである。3件とも同一番組内のシーンであったが、それぞれ番組開始直後の出演者紹介シーン、曲の歌い終わりのシーン、インタビューのシーンであった。

シーン検索のメリットは、好みのシーンの開始から番組がすぐに視聴できる点にある。言い換えると、番組検索ではアイドルグループが出演する番組が見つかってもし望のシーンを探すためには、ユーザー自身が早送り操作などで頭出しするという手間が必要であったが、この操作が不要になる。

前述した例では録画環境を1日(24時間)と限定したが、録画時間が長くなればより多くのシーンを見つけることができる。タイムシフトマシンで三つの地上デジタル放送チャンネルを録画する場合、録画可能日数はT460では最大15日分、T450では最大7.5日分で、USB(Universal Serial Bus)-HDDを併用すれば更に長時間の録画が可能である。

3 視聴中の番組に関連した話題キーワードの表示

字幕情報を利用したシーン検索は、ライブ視聴中及び録画番組再生中のシーンに関連した話題キーワードの抽出にも適用した。

3.1 “字幕シーン検索”ボタンと話題キーワード

シーン検索の一連のステップをリモコンの上下ボタンと決定

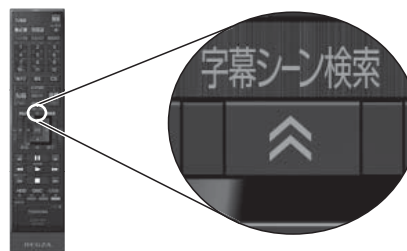


図2. 字幕シーン検索ボタン — 字幕シーンジャンプ機能を簡単なリモコン操作で行うために、字幕シーン検索ボタンを追加した。
Button for scene detection using closed captions



図3. 視聴中の番組シーンの話題キーワード表示例 — 番組視聴中に字幕シーン検索ボタンを押すと、シーンに関連する四つの話題キーワードが表示される。
Example of display of scene-related keywords

ボタンだけの簡単な操作で実現するため、T460/T450のリモコンに字幕シーン検索ボタンを追加した(図2)。このボタンはページアップボタンと併用している。番組視聴中に字幕シーン検索ボタンを押すと、視聴中のシーンに関連する四つの話題キーワードが画面の右側に表示される(図3)。この四つの話題キーワードはシーン検索に用いるキーワードである。好みの話題キーワードをリモコンで選択すれば、選択したキーワードのシーンがリストアップされる。シーンリストから好みのシーンを選択することで、頭出し再生ができる。

このように、簡単なリモコン操作で気になる話題キーワードを元に、シーンを次々と切り替えて楽しむことができる。また、話題キーワードのリストを一目見るだけでも、視聴中の番組の話題が把握でき、番組をより理解し楽しむことができる。

3.2 話題キーワードの抽出

視聴中シーンに関連した話題キーワードを見つけるためには、まず、ライブ番組視聴中なら現在時刻まで、録画番組再生中なら再生時刻前後の字幕情報を取り出す。次に、取り出した字幕情報の文字処理を行い、話題キーワードとしてふさわしいものを候補キーワードとする。それぞれの候補キーワードに対して、話題キーワードらしさをスコアとして算出し、スコア上位を話題キーワードとして画面に表示する。

3.3 実用性と品質を高める処理技術

字幕シーン検索ボタンの実用性向上と話題キーワードの品

質向上のために、次の処理技術を適用した。

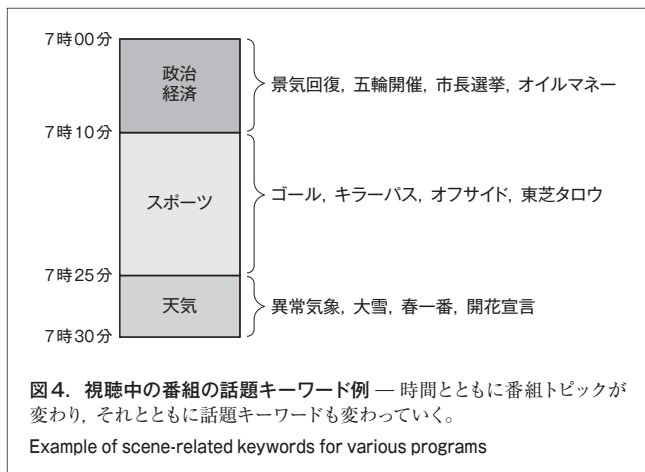
3.3.1 話題キーワードの表示 話題キーワードの出現頻度を増やすために、字幕シーン検索ボタンを押すたびに、次のスコア上位の話題キーワードを表示することにした。話題キーワードが次々と切り替わるため、ユーザーは好みに合うキーワードに遭遇しやすい。

出演者名などユーザーが気になるキーワードを優先して表示するために、話題キーワードのスコア算出式には、人名や用語の区別などキーワードの意味を考慮している。また、字幕シーン検索ボタンを押すタイミングにあったキーワードが表示されるように、ボタンを押したタイミングと話題キーワードの出現タイミングを加味したスコアとしたので、ボタンを押すタイミングで話題キーワードは変わる。

3.3.2 EPG情報の併用 字幕情報がない番組でも話題キーワードが表示できるように、EPG情報からも話題キーワードを取得した。字幕情報とEPG情報を併用することで、より広範囲の情報源から話題キーワードを抽出できる。また、字幕情報はせりふを情報源としているため、ドラマの俳優名やアニメの声優名などを抽出することが難しいが、EPG情報を利用することでそのようなキーワードも話題キーワードとして表示できる。

しかし、EPG情報の候補キーワードは、検索キーワードとして用いてもシーンが見つからない可能性がある。したがって話題キーワードとして決定する前に、シーンが見つかるかを確認する処理が必要である。とはいえ、全ての候補キーワードを確認するには時間が掛かりすぎる。そこで、一つのキーワード当たりの確認処理時間を制限することで、ボタンを押してからキーワードが表示されるまでの時間を短縮し、リモコン操作の応答性を確保した。

3.3.3 検索結果の絞込み 話題キーワードを選択したときのシーンリストは、録画番組全体から検索したシーンリストである。この中からシーン検索結果を絞り込むために、番組ジャンルなどの詳細な検索条件を追加できるようにしている。



3.4 話題キーワードの例

視聴中の番組の話題キーワード例を図4に示す。あるニュース番組を再生したところ、7時から10分までは政治経済、25分まではスポーツ、30分までは天気のトピックを取り上げていた。政治経済トピックを視聴中のとき、景気回復や五輪開催などが話題キーワードとして出現する。スポーツ、天気を視聴中のときも、それぞれのトピックに関連する話題キーワードが出現する。

4 あとがき

テレビやレコーダの高機能化と大容量HDDの普及により、多くの番組を録画可能な製品が一般的となってきた。より効率的に、より短時間で、より簡単に、という視聴スタイルに応えるため、T460/T450では、放送波の字幕情報を利用してシーンを検索する機能を搭載した。この機能では、大量の録画番組の中から出演者が話す瞬間など、人の手では発見できないようなシーンを見つけ出し、そのシーンを頭出し再生ですぐに視聴できるようにした。

それに加え、今見ている場面の話題をキーワードとして画面上に表示する機能を実現した。この機能は話題キーワードを表示するだけでも番組を楽しめ魅力的であるが、更に話題キーワードからシーン検索をするという一連のステップを、リモコンに字幕シーン検索ボタンを追加することで簡単に操作できるようにしている。

今回の字幕情報を用いたシーン検索は、字幕情報がない番組に対応できないことや、出演者のせりふ以外の情報を使用していないことから万能なものではない。今後は、TimeOnのみどころシーン再生で使用しているシーンデータとの併用、また画面上のテロップを自動認識する機能の追加など、シーン検索の精度を向上させていきたい。

テレビの楽しみ方を更に広げるために、使いやすく実用性の高い機能を搭載した製品を、今後も提供していく。

文献

- (1) 中尾雅治 他. 新たな番組視聴スタイルを提供する液晶テレビ〈レグザ〉. 東芝レビュー. 68, 05, 2013, p.7-9.
- (2) 加藤雅也 他. 録画視聴をより楽しむための みどころシーン再生及びタグリスト サービス. 東芝レビュー. 68, 05, 2013, p.14-17.
- (3) 総務省. “平成24年度の字幕放送等の実績”. 総務省ホームページ. <http://www.soumu.go.jp/menu_news/s-news/01ryutsu09_02000071.html>, (参照2014-03-17).



山下 道生 YAMASHITA Michio

パーソナル&クライアントソリューション社 ライフスタイルソリューション開発センター エンベデッドソフトウェア技術開発部主務。デジタルプロダクツのソフトウェア開発に従事。Lifestyle Solutions Development Center