

音声操作で見たい番組をすぐに見つける対話型意図理解技術

音声による対話から意図を理解し、膨大なコンテンツをすばやく検索

東芝はテレビ（TV）視聴の新たなスタイルとして、ユーザーが設定した録画チャンネルと時間帯の番組を常時録画する、タイムシフトマシンを提案しています。これによって、いつでも時間を遡ってTV番組を見ることができるようになりました。

これまで見たい番組は、キーワードやジャンルなどを指定して検索していましたが、今回、自然言語を用いた音声による対話と番組情報のクラウドコンピューティングを組み合わせることで、簡単に、かつスピーディに見たい番組やシーンを見つけ出すことができる“対話型意図理解技術”を開発しました。

タイムシフトマシンをもっと便利に

液晶TV レグザシリーズのタイムシフトマシン機能は、ユーザーが指定したチャンネル数と録画時間に応じて複数チャンネルの番組をまるごと録画するため、これらの番組は過去番組表などを使って放送時間に関係なく視聴できます。

またレグザクラウドサービス「Time-On」では、検索キーワードを指定してシーンを検索することで、見たいポイントから視聴することができます。しかし、録画された番組数が膨大になると、見たい番組やシーンを見落としたり、検索の指示に手間と時間が掛かったりするようになってしまいます。

そこで東芝は、番組を見つけたい、番組を絞り込みたい、シーンを検索したいといったユーザーの要求に対して、自

然言語による音声対話からの確に判断して結果を提示することができる“対話型意図理解技術”を開発しました。

対話型意図理解技術の特長

一般に、ユーザーが何か物事を指示する場合、唯一の表現というものではなく、多種多様な表現があり、また曖昧な表現もあります。特に自由発話を許容するシステムの場合は、これらが顕著になります。そのため表現それぞれについて、処理方法を全て規定することは極めて難しくなります。

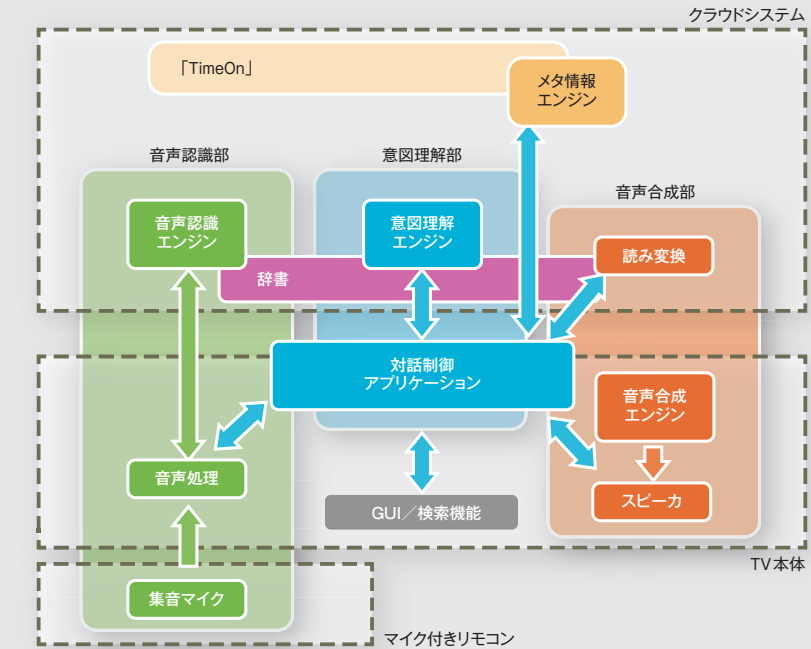
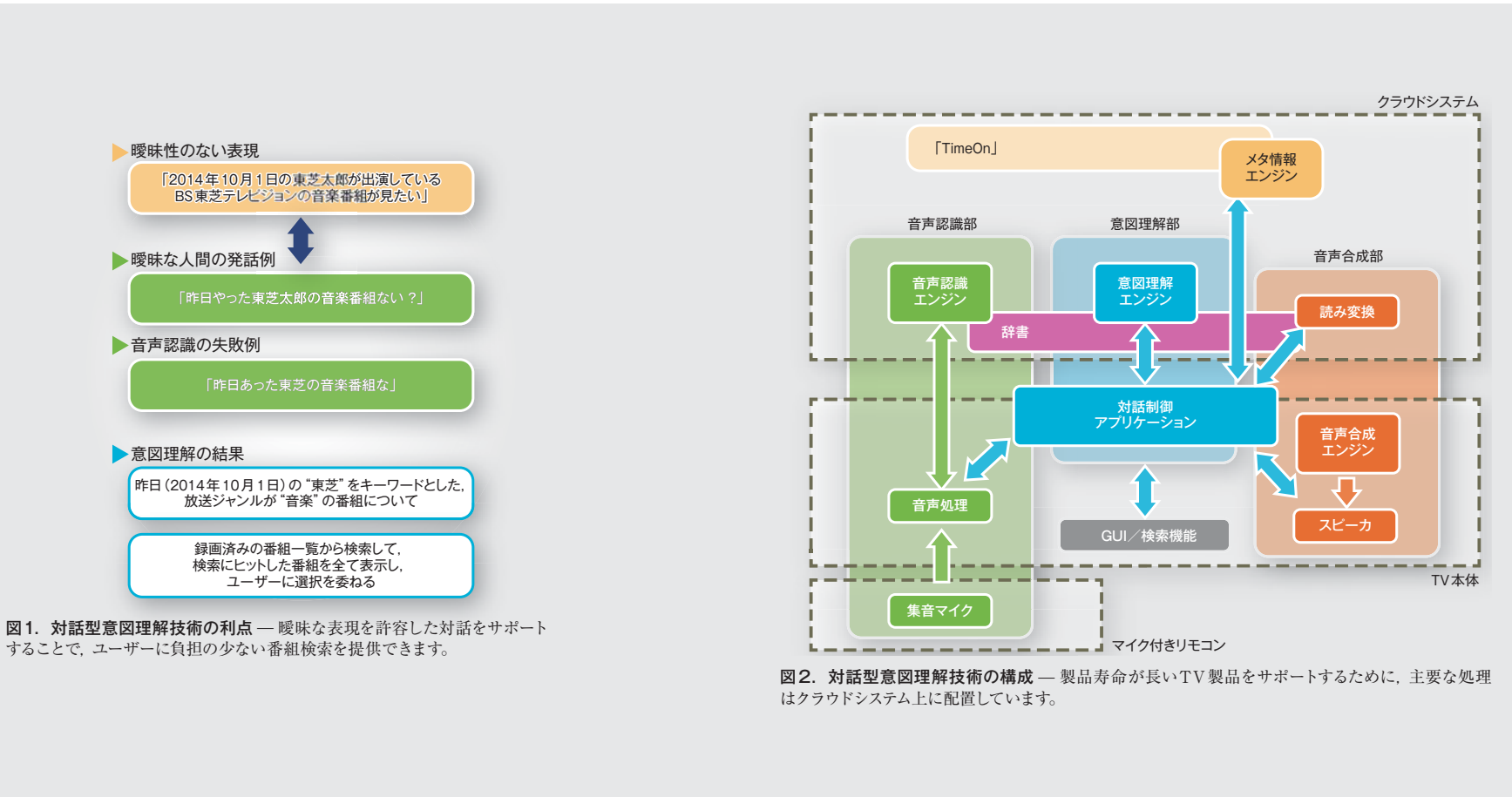
そこで今回、ユーザーの発話を音声認識した表現からその意図を推定するために、事前に様々な表現を大量に学習させた意図理解エンジンを用いることにしました。このとき、現在のTV側の状態を適宜意図理解エンジンに提供することで、正確な意図推定を実現して

います。また曖昧な表現については、クラウドサービス側のメタ情報エンジンと連携することで、ユーザーが検索したいキーワードの検出精度を高める仕組みを採用しています。最終的に検索処理などで得られた結果は、GUI（グラフィカルユーザーインターフェース）だけでなく合成音声でも知らせることで、ユーザーはTVとの対話を楽しむことができます。

このように対話型意図理解技術は、ユーザーとの間で音声発話と音声合成、及びGUIによって対話しながら、TV機能やクラウドサービスとも連携してユーザーの意図を的確に把握し、図1に示すように意図どおりの結果を返すことができる点が、大きな特長になっています。

対話型意図理解技術の構成

対話型意図理解技術は、図2に示す



めクラウドシステム上に、一貫性を持たせた常に最新の辞書を配置する構成になっています。

このように、それぞれの処理部が密接に連携することで、対話型意図理解技術を実現しています。

今後の展望

対話型意図理解技術は、当社製UHDTV（超高精細度テレビ：3,840×2,160画素）のレグザZ10Xに、初めて導入されました。ユーザーは、音声入力可能なマイク付きリモコンを用いて自由に発話し、発話意図に対応した合成音声やGUIを通して、TVと対話することが可能になりました。これによって、TV視聴をより簡単な音声操作で実現できるようになりました。

対話型意図理解技術は、これからも進化していきます。例えば、対話の履歴を活用した推薦番組や操作の提案、よりの確な番組の絞り込みなどが考えられます。

今後、この対話型意図理解技術の改良を進め、TVを更に便利に使えるように努めるとともに、TV以外の製品への応用にも取り組んでいきます。

伊藤 能康
研究開発センター
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