

2014年11月11日
NTT アイティ株式会社

認識性能が大きく向上した DNN 型音声認識ソフトウェア「SpeechRec」と 音声認識統合環境「SpeechRec Plus」クラウドサービスを同時に販売開始

～高性能な音声認識の活用がさらに簡単・低コストに～

NTT アイティ株式会社（本社 横浜市中区、代表取締役社長 長谷雅彦）は、深層学習技術により認識性能を大きく向上させた DNN（※1）型音声認識ソフトウェア「SpeechRec」と、音声認識、検索、編集、チューニングまでをワンストップで処理する統合環境「SpeechRec Plus」クラウドサービスを11月17日より販売開始いたします。

コールセンタをはじめ、スマホやタブレットを用いたモバイル環境などでも、NTT アイティの高性能な音声認識製品・サービスをご利用いただくことにより、コスト削減や新たなサービスの創出が可能となります。

本製品・サービスは、「コールセンター/CRM デモ&コンファレンス 2014 in 東京」（11月13日～14日、池袋サンシャインシティ・コンベンションセンター）に出展いたします。

※1 DNN : Deep Neural Network の略。深層学習技術と呼ばれる人間の脳を模した機械学習分野の新技術。

■ 販売開始の経緯

これまで NTT アイティでは、NTT メディアインテリジェンス研究所が開発した最先端の音声認識エンジン「VoiceRex」を搭載した音声認識ソフトウェア「SpeechRec」を提供し、さまざまな分野でご利用いただいております。このたび、新たに開発された高性能な DNN 型リアルタイム音声認識エンジンを搭載した「SpeechRec」を提供開始することとなりました。

また、主にコールセンタでご利用いただいているオンプレミス版の音声認識、検索、編集、チューニング機能を統合したソリューションパッケージ「SpeechRec Plus」を、もっと手軽に導入したいというご要望にお応えして、「SpeechRec Plus」クラウドサービスとして提供開始いたします。

■ DNN 型音声認識ソフトウェア「SpeechRec」の特長

今回新たに開発した音声認識ソフトウェアの特長は以下の通りです。

(1) 音響モデルのモデリング手法として DNN 型を採用

高い認識性能を得られる DNN 型リアルタイム音声認識エンジンの搭載により、従来型に比べ、認識誤りを3割以上削減可能です。

(2) CPU1 コアでリアルタイムの計算が可能

DNN 型は膨大な計算量が必要なため、リアルタイムで動作させるためには、GPGPU(※2)や複数の CPU を使用する必要があり、従来型よりもサーバ台数が増加するという課題がありました。今回、音響モデルのモデリング手法の工夫により、DNN 型においても CPU1 コアでリアルタイムの計算を可能としました。

(3) 「音響モデル」のチューニング無しでも高い認識性能

今回開発した音声認識ソフトウェアでは、ほとんどの用途で「音響モデル」のチューニング無しでも高い認識性能が得られます。これにより、これまで非常に大きなコストがかかっていた「音響モデル」のチューニングコストの低減が可能です。

※2 GPGPU：グラフィックプロセッシングユニットによる汎目的計算技術（General-purpose computing on graphics processing units）。

■ 統合環境「SpeechRec Plus」クラウドサービスの概要

「SpeechRec Plus」は、音声認識、検索、編集、チューニング機能を統合したソリューションパッケージです。これまで、オンプレミス型を提供してまいりましたが、もっと手軽に利用したいというニーズにお応えするため、新たにクラウドサービスとして提供開始いたします。本サービスの特長は以下の通りです。

(1) クラウドサービスのため設備構築の費用や期間が不要

すぐに始めたい、短期で利用したいというお客様にも、手軽に音声認識をご活用いただけます。また、最初は小規模でトライアル的に開始し、段階的に利用拡大していくことも可能です。規模が大きくなった場合には、オンプレミス版への移行も容易です。

(2) クライアント側は Web ブラウザだけで利用可能

音声ファイルのアップロードや検索・修正など、各種操作が Web ブラウザだけで可能です。また、コールセンタの通話の書き起こしをはじめ、多くの用途に利用可能です。例えば、会議の録音音声からの議事録作成や、口述筆記の書き起こしなどにご利用いただけます。これまで、録音音声を再生し、聞きながら行っていたテキスト化作業の効率化が期待できます。

■ NTT アイティの音声認識製品・サービスのラインナップ強化

NTT アイティでは、従来より音声認識製品・サービスを提供してまいりました。今回の製品・サービスの追加により、音声認識の基本的な製品・サービス、および統合製品のラインナップをさらに強化し、音声認識のさまざまなニーズに応えることが可能となります。また、音声認識機能を利用する API 仕様などを統一化しているため、はじめはクラウドサービスでスモールスタートし、規模や期間を拡大する際にオンプレミス型へスムーズに移行することが可能です。

(1) 音声認識クラウドサービス

クライアントのアプリケーションから音声認識を利用するタイプに加え、新たに、HTML5 ブラウザさえあれば音声認識が利用できるサービス「SpeechRec for Browser」を追加いたします。

(2) 統合製品クラウドサービス

「SpeechRec Plus」のクラウドサービスを開始いたします。（本サービスに使用する音声認識エンジンの DNN 型への対応は今後の予定です。）

表 音声認識製品・サービスのラインナップと主な利用例

種別	音声認識	音声認識統合型
クラウドサービス	<input type="checkbox"/> 「SpeechRec Cloud」 <ul style="list-style-type: none"> ・ボイスメモ文字化サービス ・日報音声入力 ・音声コンシェルジュサービス <input checked="" type="checkbox"/> 「SpeechRec for Browser」 <ul style="list-style-type: none"> ・サイト内検索サービス ・対話型サイネージサービス ・ビデオチャットなど WebRTC アプリ連携 	<input checked="" type="checkbox"/> 「SpeechRec Plus」クラウドサービス <ul style="list-style-type: none"> ・通話音声ログ分析向けトライアル ・通話音声ログ分析向けスポットサービス
製品	<input type="checkbox"/> 「SpeechRec Server」 <input type="checkbox"/> 「SpeechRec SDK」 <ul style="list-style-type: none"> ・音声認識システムのオンプレミス構築 ・対話型音声自動応答システム ・会議録作成支援システム 	<input checked="" type="checkbox"/> 「SpeechRec Plus」製品版 <ul style="list-style-type: none"> ・通話音声ログ分析 ・録音音声の書き起こし支援

■：今回新規に追加 □：今回 DNN 版で提供、または DNN 版へバージョンアップ

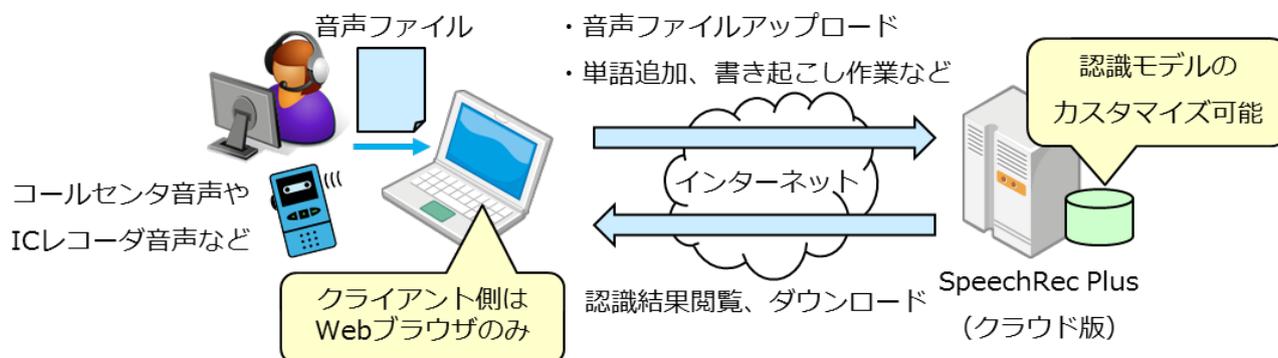


図1 「SpeechRec Plus」(クラウド版) 利用例

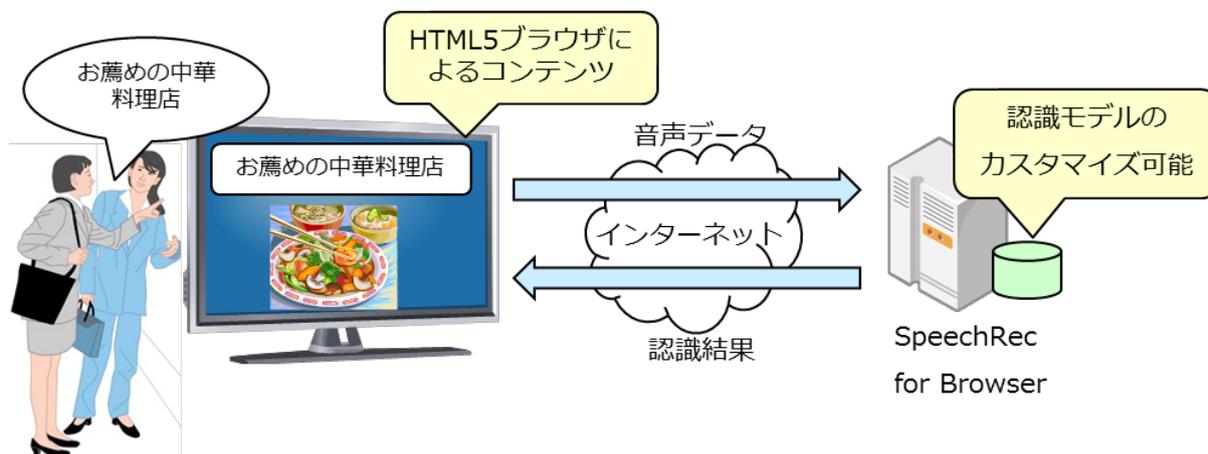


図2 「SpeechRec for Browser」利用例 (対話型サイネージサービス)

■ 販売価格（税別）と販売目標

- ・「SpeechRec Cloud」、「SpeechRec for Browser」
月額 5 万円～
- ・DNN 型音声認識ソフトウェア「SpeechRec Server」
160 万円／サーバ（10 認識プロセスまで）
- ・統合環境「SpeechRec Plus」クラウドサービス
月額 15 万円～
- ・統合環境「SpeechRec Plus」製品版
500 万円／サーバ（10 認識プロセスまで）

※各製品の動作環境、認識モデルチューニング費用などは個別にお問い合わせください。

音声認識製品・サービスの販売目標：年間 1 億円

■ 販売開始日

- ・販売開始：2014 年 11 月 17 日

「SpeechRec」は NTT アイティの登録商標です。
その他の商品名などは、各社の商標または登録商標です。