

2021年3月30日
国立大学法人電気通信大学
ソフトバンク株式会社

手話と音声による双方向コミュニケーションシステムを開発

～聴覚障がい者と健聴者を結ぶ「SureTalk」で新たな社会基盤の構築へ～

国立大学法人電気通信大学（東京都調布市、学長 田野 俊一、以下「電気通信大学」）とソフトバンク株式会社（東京都港区、代表取締役 社長執行役員 兼 CEO 宮内 謙、以下「ソフトバンク」）は、聴覚障がい者と健聴者のより円滑なコミュニケーションを実現する新たな社会基盤の構築に向けて、手話と音声による双方向コミュニケーションシステム「SureTalk（シュアトーク）」を共同開発し、茨城県水戸市などで試験提供を行ってきました。4月からは東京都調布市や福島県聴覚障害者協会などに提供先を拡大し、今後、認識率のさらなる精度向上を図っていきます。

【「SureTalk」の概要】

「SureTalk」は、AI（人工知能）を使用し、手話と音声相互にコミュニケーションをしたり、AIに手話を学習させたりすることができるウェブツールです。このシステムでは、AIが端末のビデオ通話から身体動作を追跡して、手話の特徴を抽出し、手話を認識してテキストへと変換します。健聴者から聴覚障がい者へは、音声を自動でテキスト化することでコミュニケーションを行います。また、AIの手話認識には多くのデータを必要とするため、今後効率的にデータを蓄積できる環境を整備することで、より多くの情報を蓄積し、AIの認識精度を高めることができます。利用者は「SureTalk」にログインして「会話を開始する」ボタンをクリックし、自分が使用する言語モード（手話／音声）を選択した上でトークルームに入室するだけで、簡単に1対1の会話を始めることができます。

■ トークルームイメージ



【社会的な背景】

聴覚障がい者が、日常生活や災害時などに必要な情報を迅速に入手し、安心・安全な生活を送るために、聴覚障がい者と健聴者が十分な情報を共有する新たな社会基盤の構築が不可欠です。具体的には、聴覚障がい者の言語である手話を健聴者が理解するとともに、健聴者の言語である音声を聴覚障がい者に伝える双方向のコミュニケーションシステムの開発と実用化が急務です。

この課題を解決するために、電気通信大学とソフトバンクの2者で「SureTalk」のコアエンジンである手話認識、自然言語処理、音声認識に関する共同研究を国立大学法人名古屋工業大学の技術支援も受けて、約3年半にわたって実施し、加えて双方向コミュニケーションを実現するインフラ、そしてユーザーインターフェースなど多くの技術を各企業と共同開発しました。そして、手話の表現にラベル付けを行うアノテーション^{※1}、個人情報保護を担保するための匿名化技術など革新的な技術開発も行いました。今後、さらに「SureTalk」の利便性を高めていくために、共同開発企業間での連携強化ならびに連携企業の拡大を行い、聴覚障がい者と健聴者のより円滑なコミュニケーションを実現する社会基盤の構築を目指します。

【手話と音声の双方向コミュニケーションシステムの構成と要素技術】



「SureTalk」のシステムは、コアエンジン部と手話データベース部で構成されます。コアエンジン部では聴覚障がい者の手話をテキストに変換して健聴者に通知する手話認識部と自然言語処理部および健聴者の音声をテキストに変換して聴覚障がい者に通知する音声処理部からなり、それぞれの具体的な機能は次のようになります。

まず、手話認識部では、パソコンやスマートフォンなどのカメラを用いて撮影した動画像に対して、深層学習を活用した身体動作の追跡処理を行い、話者の骨格座標を得ます。その後、骨格座標に基づいて話者の姿勢や動作など手話を認識するために重要な特徴を抽出します。最後に、深層学習を活用した時系列認識処理によって話者が表現している手話の単語順を認識します。深層学習による身体動作追跡を用いることで、服装をモノトーンにすることや撮影の背面をグリーンにすること、さらに深度センサーを利用することなどの特殊な条件が不要となり、日常的な環境での撮影が可能になりました。さらに、時系列認識処理に深層学習を適用することで、手話の認識精度を向上させています。

(担当：電気通信大学情報理工学専攻 情報学専攻 高橋 裕樹研究室)

自然言語処理部では、手話単語列と日本語文の対訳例から手話単語と日本語単語との対応関係、助詞の補完、および動詞・助動詞の活用について統計的機械翻訳の手法や深層学習を用いた翻訳モデルを学習します。その翻訳モデルを使用して、手話認識部から出力される手話単語列を変換後の日本語単語列の生起確率が最大になるように翻訳を行います。これにより、手話単語列をより自然な日本語文に変換することが可能になりました。

(担当：電気通信大学情報理工学研究科情報学専攻 内海 彰研究室)

音声処理部では、オープンソースの深層学習に基づく音声認識エンジンを利用して、健聴者が発声した音声情報をリアルタイムに文字情報に変換します。雑音のある環境、反響の多い環境などのさまざまな環境における音声データを独自に収集し、それらを用いて音響モデルを学習させることで、実環境下における認識精度を高めることが可能になりました。また、話者識別機能を実装し、誰が発声したかを判別できるようにすることで、複数話者におけるコミュニケーションを円滑化することを目指しています。

(担当：電気通信大学情報理工学研究科情報・ネットワーク工学専攻 中鹿 亘研究室)

手話データベース部では、東京都調布市のご協力を得て、社会福祉協議会に所属する聴覚障がい者の皆さまから手話言語収集について多くの有益な情報を取得し、それらに基づきソフトバンクの社員による手話動画データ撮影に加え、福島県聴覚障害者協会のご協力を得て、多くの手話動画データを収集する体制を構築しました。そして、撮影した動画を分類するアノテーション作業については、機械学習や深層学習の手法を適用して自動化することに取り組み、作業時間を大幅に短縮して、研究・開発の効率を飛躍的に高めています。また、手話動画はオリジナルの動画の匿名化技術によって個人の特定が困難となる加工を施し、個人情報の安全性を担保しています。

【試験提供中および今後開始予定の自治体・団体】

- ・茨城県水戸市
- ・東京都調布市
- ・福島県聴覚障害者協会

【「SureTalk」の開発協力企業について】

「SureTalk」の開発協力企業および各社からのコメントは、次の通りです。

- ・株式会社ギークフィード：リアルタイムコミュニケーションの確立やアプリケーションの開発など

「このたびは各ジャンルで最先端をリードする企業と名を連ね、大変社会的意義のある新たな価値を創造する産学官連携プロジェクトに参画させていただき、光栄に思います。今後も弊社の培った技術が社会貢献事業の一端を担えるよう社員一同まい進していきます」

- ・株式会社セラク：アプリケーションのユーザーインターフェースの開発など

「聴覚障がい者と健聴者の両者が、このツールを通して十分な情報の共有・コミュニケーションが取れるよう、見やすさ、使いやすさへの配慮と、サービスを通して得られる顧客体験を意識した UI・UX デザインを

設計・実装しました。また、マーケティング視点を持ち、認知拡大～理解深化～ユーザー化を目的としたコミュニケーションデザインとサービスサイト構築を行いました」

・株式会社 Preferred Networks：個人情報保護のための手話動画匿名化の開発など

「個人情報保護の観点から手話動画を匿名化するために、深層学習を用いた姿勢推定技術により、映像上の手話者の手や指先の動きを正確に把握し、それに連動する 3D コンピューターグラフィックス(CG)を自動生成する技術を開発しました。Preferred Networks は今後も深層学習技術を応用して、聴覚障がい者と健聴者のコミュニケーションを円滑にするための技術開発に取り組んでいきます」

・株式会社 ABEJA：手話認識エンジンの前工程であるアノテーション業務など

「長年のモデル開発のノウハウと自社開発のアノテーションツール『ABEJA Platform Annotation』を活用し、38 時間の動画および 5 万手話文書の手話動画のラベリングプロジェクトを担当しました。ソフトバンクとの協業を通じて、良質なモデル開発に欠かせない高品質なデータセットづくりを行うことができました」

【今後の予定】

試験提供にご協力いただいた水戸市役所に加え、調布市社会福祉協議会、福島県聴覚障害者協会などに「SureTalk」を設置して、ご利用いただく場合の課題を抽出し、研究開発にフィードバックすることでコアエンジン部の精度向上を図るとともに、ユーザーインターフェースも含めた利便性、さらなるセキュリティの担保のための施策を講じていく予定です。

また、手話データベースは認識率の向上に非常に重要なため、データの増強にも積極的に取り組んでいます。ソフトバンクの社員によるデータ増強に加え、2021 年 7 月には手話データ登録機能を公開し、手話に興味を持っていただける方に広く訴求していくことで手話データベースの充実を図り、手話認識率の精度向上につなげてユーザーの利便性を高めていきたいと考えています。

なお、現時点で対象とする手話は日本語の文法に準じた「日本語対応手話^{※2}」ですが、将来的には独自の文法体系を有する「日本手話^{※3}」も対象とし、「SureTalk」の利便性をさらに高めていきたいと考えています。

※ 1 手話の情報を集めたデータベースの利便性を高めるために、画像データに、文の区切りや構文などの言語情報を付与すること。

※ 2 日本語文の文法や語順に従って、日本語単語に対応する手話単語で置き換える手話であり、主に難聴者や中途失聴者に利用される。

※ 3 文法的に日本語と対応していない手話で、主にろう者に利用される。日本手話は、それ自体が一つの言語と考えられる。

【「SDGs」の取り組みについて】

「SureTalk」の提供は、「SDGs（持続可能な開発目標）※」の17のゴールのうち、主に下記のゴールに取り組むものです。



ソフトバンクは、「SDGs」の課題解決を重要な経営課題と捉え、国際社会が追求する社会課題の解決に貢献することにより、企業価値の向上と持続的な社会の実現に取り組んでいます。「SureTalk」は、ソフトバンクが「SDGs」で特定する、六つのマテリアリティ（重要課題）のうち、「人・情報をつなぎ新しい感動を創出」というテーマで取り組む活動であり、この活動を通して、障がいの有無にかかわらず「誰もが情報へアクセスできる環境の提供」の実現を目指します。ソフトバンクの「SDGs」の取り組みに関する詳細は、ウェブサイト「サステナビリティ」をご覧ください。

[注]

※ Sustainable Development Goals の略称で、2030 年までの達成を目指して、豊かで活力のある未来を創るための 17 のゴールと 169 のターゲットを定め、2015 年に国連で採択された世界規模の取り組み。

●SoftBank およびソフトバンクの名称、ロゴは、日本国およびその他の国におけるソフトバンクグループ株式会社の登録商標または商標です。

●その他、このプレスリリースに記載されている会社名および製品・サービス名は、各社の登録商標または商標です。

◆ 本リリースに関する問い合わせ先

国立大学法人電気通信大学 総務企画課広報係

TEL : 042-443-5019 MAIL: kouhou-k@office.uec.ac.jp