



センサリ社(Sensory,Inc.)製品・技術の説明

2018/3/23

センサリージャパン(株)

1. 製品

● TrulyHandsfree™ <software> : THF

音声トリガー (Trigger) および音声コマンド (Commands) に対応した技術です。多くのスマートフォンなどに採用されています。現在は複数の IC メーカーのプロセッサ、DSP にもポーティングされています。各種 OS 対応の開発用の SDK が準備されています。

● TrulySecure™ <software> : TS

顔認証 (Facial Authentication) + 音声認証 (Speech Verification) の組み合わせ技術です。顔 + 音声認証、または顔認証のみ、音声認証のみなどがアプリケーションで設定できます。現在は OS を必要とするため PC、タブレット、スマートフォン向けになっていません。各種 OS 対応の開発用の SDK が準備されています。

TS 技術は FIDO Certificate (FIDO Alliance) を取得しております。

● TrulyNatural™ <software> : TN

大規模音声認識技術です。1 万語以上のフレーズをテキストから変換して認識辞書を生成し音声認識できる機能があります。また自然言語認識のような技術となります。対話のように自由なフレーズを認識し判断することのできる技術です。組み込みとクラウド環境に対応した技術となっています。各種 OS 対応の開発用の SDK を準備中です。

● VoiceGenie™ <software>

Alexa アクセスに対応した Bluetooth ヘッドセット用の音声認識開発キットです。省電力トリガー待ち受けは前段の DSP で行い、後段の Qualcomm 社 8670/8675 にて音声コマンド認識及び音声合成を行う方式とすることで完全ハンズフリー操作が可能となるリファレンスデザインを提供しています。またスマートフォンアプリのサンプルも提供することで実際に Alexa とトリガー認識すれば AVS クラウドに接続し Alexa 音声アシスタントのサービスを受けることができ、VoiceGenie とトリガー認識すれば音楽操作、Bluetooth 操作、また電話着信操作などに対応しています。

開発用の VG SDK が準備されています。

● BlueGenie™ <software>

Qualcomm 社 BC05/8670/8675 専用の不特定話者音声認識および音声合成 (圧縮身長) 技術、CallerID 再生技術です。Bluetooth Headset や Car Unit に採用されている製品です。

開発用の SDK が準備されています。



- BlueGenieLite™ <software>

BlueGenie の音声合成のみの製品となります。音声再生のみを利用したいという商品にご利用いただけます。

開発用の SDK が準備されています。

- RSC-4x <IC>

RSC-4x はセンサリ社が開発生産している 8bitMPU ベースの音声認識音声合成制御マイコンです。同シリーズには RSC-464 および RSC-4128 があります。安価な IC 製品のため数多く玩具やガジェットに採用されています。開発ソフト、評価ボードなど販売中です。IC 量産対応品です。

- NLP-5x <IC>

NLP-5x はセンサリ社が開発生産している 16bitDSP ベースの音声認識合成制御 DSP です。ハードウェアは数多くの I/O および各種 I/F を装備した IC で連続した単語認識等に対応しているため家電系の製品に採用されています。開発ソフト、評価ボードなど販売中です。IC 量産対応品です。

2. 技術名称と説明

- 不特定話者音声認識 (SI: Speaker Independent)

「不特定話者音声認識」とは不特定多数のユーザーが利用することを前提に辞書を作成したものです。この辞書は事前に録音データまたは既存データから加工し作成するものです。通常はお客様ごとにカスタムに辞書データを作成するため、商品に対応した認識単語を決定し、作成することができます。商品を購入後すぐに使用できるなどの利点があり、また不特定多数で使用する商品などに向いています。

製品: TrulyHandsfree、BlueGenie、RSC-4x、NLP-5x

- 特定話者音声認識 (SD: Speaker Dependent)

「特定話者音声認識」とはユーザーが音声を登録してから認識させる技術となります。ユーザーが自由に単語を決め、登録することができるのでカスタマイズができますが登録語数が多いものは不向きです。

製品: TrulyHandsfree、BlueGenie、RSC-4x、NLP-5x

- 音声認証 (SV: Speaker Verification)

「音声認証」とは話者を認証する技術です。センサリ社の場合はパスワード認証技術となります（声紋認証ではありません）。特定話者音声認識のようにパスワードを登録し、このパスワードの音声と一致しているか否かを認証する技術です。

TrulyHandsfree をベースにした NNSV と TrulySecure をベースにした TSSV があります。



製品: TrulyHandsfree、TrulySecure、RSC-4x、NLP-5x

● 音声トリガー (Voice Trigger)

「音声トリガー」とは押しボタンの代わりに音声で音声操作を開始するための技術です。カーナビゲーションなどでは音声ボタンを押してから会話となりますがこの押しボタンの変わりとなる技術です。スマートフォンなどですとマイクボタンを押してから会話となっていますがこの代わりともなります。センサリ社の音声トリガーは騒音環境に強く、また誤認識が低いことが特徴です。

製品 : TrulyHandfree、BlueGenie、NLP-5x

● 音声コマンド (Voice Commands)

「音声コマンド」は音声トリガーの後に続く操作コマンドとなります。例えば、”Hello Smartphone, Play Music”の場合の Play Music や Stop Music がコマンドに当たります。

製品 : TrulyHandsfree、BlueGenie、NLP-5x

● フィックスドトリガー (FT : Fixed Trigger)

「フィックスドトリガー」は不特定話者音声トリガー認識です。予め商品に対応した音声トリガーフレーズを不特定話者で作成し、使用環境を想定した情報を組み込んだ上で作成した音声トリガーです。各種環境音に強く、誤認識(誤トリガー)が低くすることができる技術です。ポータリング済みのローパワーDSP上で動作させることができます。

辞書作成は一般の方の音声を録音し辞書を作成する必要があります。小型辞書モデル

(DSP等向け)は約140~250名の音声サンプルが必要です。また中型辞書モデル(スマートフォン等)は約250=350名の音声サンプルが必要です。

製品 : TrulyHandsfree、BlueGenie

● ユーザーデファインドトリガー (UDT : User Defined Trigger)

ユーザーがトリガーフレーズを登録できる技術です。登録は特定話者音声認識のように音声登録することで音声トリガー認識させることができます。ポータリング済みのローパワーDSP上で認識動作させることができますが登録処理(Enrollment)はPCやスマートフォンアプリ上で動作させる必要があります、DSP向け認識ではここで生成されたデータをDSPメモリーに転送する必要があります。

OS版ではFTとUDTを同時に動作させることができます(EFTとUDT同時動作はできません)。

製品 : TrulyHandsfree

● ユーザーデファインドコマンド (UDC : User Defined Commands)

ユーザーがコマンドフレーズを登録できる技術です。設定は特定話者音声認識のように音声登録することで音声コマンド認識させることができます。ポータリング済みのローパワーDSP上で認識動作させることができますが登録処理(Enrollment)はPCやスマー



トフォンアプリ上で動作させる必要があります、DSP 向け認識ではここで生成されたデータを DSP メモリーに転送する必要があります。

製品： TrulyHandsfree

- ユーザーデファインドパスフレーズ (UDP: User Defined Passphrase)

UDT の音声認証機能を向上させた技術となります。UDT の場合は登録話者以外のパスワードでも認識しますが UDP の場合はより認証性能を向上し登録話者のみを認識するようにしたものです。ポーティング済みのローパワーDSP 上で認識動作させることができますが登録処理 (Enrollment) は PC やスマートフォンアプリ上で動作させる必要があります、DSP 向け認識ではここで生成されたデータを DSP メモリーに転送する必要があります。

製品： TrulyHandsfree

- 話者識別 (SID: Speaker Identification)

話者識別は複数のユーザーが同じパスワード (または別のパスワード) 登録し、登録されたパスワードの識別する技術です。PC、タブレットや家電製品のように複数人数で使用場合、音声で識別してログインすることができます。また自動車などではドライバーのパスワードでミラー、シート設定を変えるなどの応用が可能です。

閾値設定等で識別率が変化します。例えば、閾値を下げることで同じ男性であれば認識と判断するなどができます。ポーティング済みのローパワーDSP 上で認識動作させることができますが登録処理 (Enrollment) は PC やスマートフォンアプリ上で動作させる必要があります、DSP 向け認識ではここで生成されたデータを DSP メモリーに転送する必要があります。

製品： TrulyHandsfree

- エンロールドフィックスドトリガー(EFT : Enrolled Fixed Trigger)

EFT は FT の拡張技術です。FT ですでに登録された言葉に対し、ユーザーの言葉を登録する方法です。イントネーションやアクセントの違いで FT では認識しづらい場合は EFT 技術を使用してユーザーの発音を認識しやすくする技術となります。登録処理 (Enrollment) は PC やスマートフォンアプリ上で動作させる必要があります、DSP 向け認識ではここで生成されたデータを DSP メモリーに転送する必要があります。

製品： TrulyHandsfree

- ティーエスエスブイ (TSSV : TrulySecure Speaker Verification)

TSSV は TrulySecure の生体認証音声認証部を専用技術ととして商品化したものです。TSSV は DSP 等には実装できないため AP への実装が必要です。通常の話者認証より性能が良く話者の認証する場合にご使用いただく技術となります。THF との組み合わせの場合、THF トリガーを DSP 等にも実装しそこで一度、認識させ一致した後にバッファリングされた音声のを AP に転送し認証処理を行う方法となります (デバイスによっては対応していないものもあります)。

製品： TSSV



- フレーズスポッティング (Phrase Spotting)

フレーズスポッティングは文章に含まれるキーワード(認識すべき言葉)を認識することのできる技術です。例えば、認識キーワードを **Temperature** とし”What is the temperature now”と喋った場合でも **Temperature** が一致したとして認識する技術です。この技術によりキーワードが一致すれば認識することができますので認識フレーズや単語をすべて覚える必要がなくなります。

製品： TrulyHandsfree、NLP-5x

- トリガー・トゥ・サーチ (Trigger to Action)

本技術は音声トリガーの後にクラウドなどの音声検索サービスに音声データを渡し音声検索させる技術です。例えば、”Hello smartphone, what is the temperature outside?” と喋ると “Hello Smartphone”をトリガー認識し、その後続く音声をクラウド転送し音声検索させるものです。スマートフォンなどのアプリケーションに向いています。

製品： TrulyHandsfree

- LVCSR (Large Vocabulary Continues Speech Recognition)

大規模音声認識の技術です。音楽再生の楽曲、アルバム、アーティスト検索やナビゲーションの POI 検索、アドレス検索などに利用できる技術です。

*アドレス検索は現在開発中です。

製品： TrulyNatural

- テキスト音声合成 (TTS: Text-To-Speech)

テキストを音声化して再生する技術です。センサリ社の場合は主に **CallerID** (発信者名)を読み上げることが前提として行っていますので音質はあまり良くはありません。また小型化された技術ですので文書読み上げなどには不向きです。

*現在、US,UK,AU 英語、北米,欧州スペイン語、北米,欧州フランス語、イタリア語、中国語(Mandarin)に対応しています。

製品： TrulyHandsfree、BlueGenie、NLP-5x

- CallerID

CallerID は発信着信者名や番号を意味します。ハンズフリーにて着信した際に相手の名前や番号を。また発信する際に発信者名、番号を読み上げるなどの技術となります。センサリ社の **TTS** を利用し再生を行います。

製品： TrulyHandsfree、BlueGenie

- 音声合成 (Speech Synthesis)

センサリ社の場合、音声合成とは録音音声を圧縮し、再生時に伸張して再生する方式となります。センサリ社独自の圧縮方式で **SX** と呼んでいます。比較的に高音質で圧縮率は 400bps



～28000bps などで設定できます。

製品： BlueGenie、RSC-4x、NLP-5x

3. その他の説明

- 言語

対応言語は現在、約 40 か国（約 20 言語）となります。同じ言語、例えば英語ですと US/UK/Australia/Indian English が準備されており対応した国（地域）に合わせて言語解析アルゴリズムを使用します。

ご不明な点がありましたら下記までご連絡ください。

センサリージャパン株式会社

〒241-0004 横浜市旭区中白根 1-26-28

Tel: 045-958-1730/FAX:-045-958-1731

担当：栗原 kurihara@sensory.co.jp