

10/5 実施

企業：株式会社クリムゾンテクノロジー

テーマ：ユーザーの声をリアルタイムに変換する「AI 声質変換技術」

**Q：AI で学習した音声を対象者がリピート再生する際に、対象者が好む音声にカスタマイズしていくことは可能でしょうか？**

A：AI で学習した声に関してはピッチの微調整のみのカスタマイズになります。  
また、シンセサイザーによる声質変換ではパラメーターで追い込んでいくことが可能です。

**Q：動物や風の音など人間以外の音声変換は可能でしょうか？**

A：AI 声質変換には、ターゲット（キャラクター）の声と、使う人側の声のユニゾンした音声ファイルが必要になりますので、動物や風の音は難しいと思います。シンセサイザー変換を使えば、自然ではない声を作り出すことも可能です。

**Q：「Web 会議や遠隔接客で使える製品を」との記載がありますがこちら以外の一般商材目線での提案の検討余地はありますか？**

A：もちろんです。例えばですが、人の声をデバイスの中で変換して聴きやすくするような、いわば逆補聴器のような、そういった提案の可能性、余地もあるのかなと思います。しかし、現段階で弊社の音声変換は「近接のダイナミックマイク」でないともあまり品質良く変換できないところがあるので、PC・スマホ以外での展開は難しい部分があります。

**Q：スマホや PC は難しいのでは、とのことでしたがラズベリーパイなど活用した方法もあると拝見しました。前もって事前に音声データをインプットさせておくことで、スマホや PC 以外のデバイスで利用する方法もありえる、と認識してよいでしょうか。**

A：ラズベリーパイですと、サンプリング周波数は 16k くらいでしたら実績があります。CD 音質が 44.1k なので、その半分くらいでしょうか。iPhone の電話よりも、少し音質が劣るくらいでしょうか。

**Q：自分の地声は変えずに、トーンだけ変更（マイナーチェンジ）することはできますか？**

A：AI 変換ですとできませんが、シンセサイザー変換ですと声質やトーンのちょっとした変換が可能です。

**Q：AI 変換とシンセサイザー変換の技術的な違いは？**

A：シンセサイザーは声の形を微妙に変化させることができます。AI 変換は、キャラクターなどの声を表現できますが、明確な学習データが必要なので「なんとなくこんな感じの声」ができません。

**Q：アニメキャラに似せた声を生成した場合、著作権の扱いはどうなるのでしょうか？**

A：実は、現在声自体に権利の設定はありません。しかし、一定のキャラクターとして使うのであれば、版元（権利者）の許諾はもちろん必要になります。

**Q：zoom や、オンラインコミュニケーションツールにアドオンで追加できるような仕組みというのは可能でしょうか？**

A：最新のボイドルを使うと、配信の方に自動的ストリーミングが可能です。

例えば Google Meet の音声部分に組み込みたいというニーズがあれば、一つはインストールタイプのプログラムに SDK 利用して展開する。もしくは Web Assembly という形式で JavaScript 的にダウンロードしつつその SDK を動作させるという方法があります。