

1. はじめに

近年、Vtuber 市場は、2023 年度に 800 億円程度の規模となる見込みである、2020 年比で約 5 倍になるとみられています。また、国内におけるメタバース市場も急速な拡大が予測されており、2027 年には 2 兆円にも上る規模になると予測されています。そのため、これらの市場の拡大によってアバターの需要が高まっていると考えられます。

現状、Live2D アバターの表現方法には 2 種類あり、内部の人間の表情をトラッキングしたものを Live2D モデルに反映させる方法と、あらかじめ用意した表情や動きを内部の人間が手動で選択する方法があります。これらの方法ではゲーム配信などの両手を離せない状況で、「涙を流す」のような表情に手動で切り替えることは困難です。そこで、私たちは機械学習による豊かな感情表現を可能にする新しいトラッキングソフトウェア、「エモアンプ」を提案します。

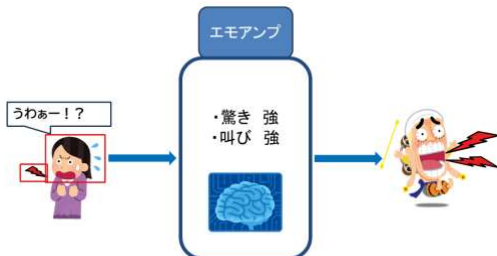


図 1 エモアンプのイメージ

2. 概要

エモアンプは従来のトラッキングに加えて、感情推定等によって独自の表現を可能にするソフトウェアです。ウェブカメラから顔表情、マイクから音声を受け取り、顔表情、声色から感情を推定して、自動かつ連続的に表情を変化させていきます。また、音声から母音を推定し、母音に対応した口のの形にしていきます。

2. 感情推定について

エモアンプでは、ウェブカメラから顔表情、マイクから音声を受け取り、顔表情、声色から感情を推定します。その結果を利用して、パーティクルを追加したり、怒っている表情に変化させたりします。これによって、トラッキングのみでは困難な、アニメ風の表現を自動的に行えるようにします。

3. システム構成

エモアンプのシステム構成を図 2 に示します。PC のマイクおよびウェブカメラで映像と音声を受け取り、Python 上で表情のトラッキングや各種推定を行います。その後、その結果を Unity へ出力し、Unity 上で Live2D モデルの操作を行います。

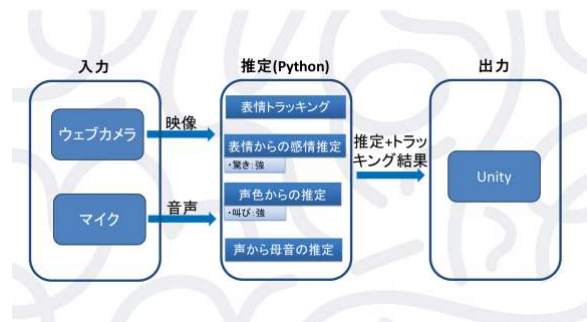


図 2 システム構成図

4. おわりに、まとめ

この「エモアンプ」を使うことによって従来のトラッキングに比べてより豊かな感情表現を行うことができ、さらに、キャラクターの操作に割くリソースを減らすことができるため、より充実した配信を実現することができます。