

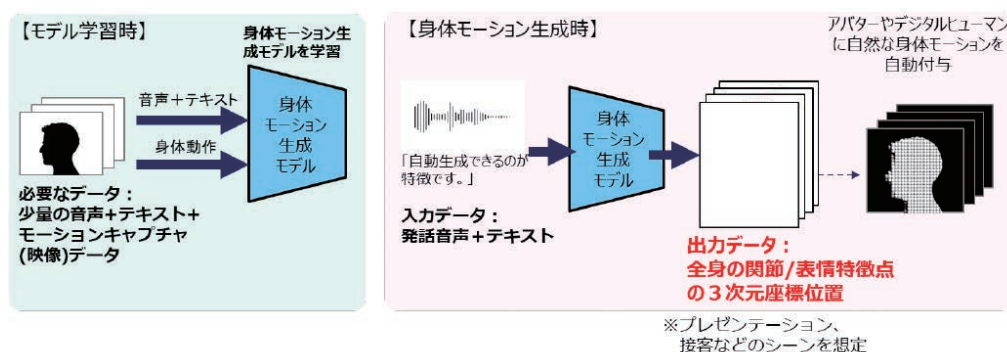
発話内容に対応した自然な動作を生成可能な 身体モーション生成技術

発話音声・テキストに対応する自然な表情や身体の動き（身体モーション）を自動的に生成する技術です。

本技術は、事前に発話音声とその発話内容のテキスト、および身体モーションデータを元に、深層学習モデルを用いた身体モーション生成モデルを構築します。利用時には発話音声とその発話内容のテキスト情報を入力することで、対応する身体モーションを自動生成することができます。身体モーションは個人やシーンごとに様々な特徴を持ち差が大きいいため、特定の人物やシーンごとの特徴を考慮した生成モデルを構築するためには、従来、その人物やシーンごとにそれぞれ大量のデータが必要でした。これに対して、これらのデータが少量でも効率よく生成できる独自技術を開発しました。本技術によって、大規模なデータ収集とモデル学習のコスト

を削減できます。

本技術を利用すれば、対面で会話するために互いに異なる地点から特定の場所へ移動しなくとも、発話音声から身体モーションを生成した結果をアバタに反映することで、対面と同等に音声や身体モーションが可能なアバタを介した自然なリモート対話を実現できます。また、デジタルヒューマンを用いた対話システムを構築する際には、専門家が手作業で膨大なコストを使用して身体モーションを作成していましたが、本技術を用いることで、対話システムの発話に合わせた自然な身体モーションを自動付与できます。このように、本技術によって、移動コストの削減や専門家の作業稼働等の削減が大いに期待されます。



図：発話内容に対応した自然な動作を生成可能な身体モーション生成技術の概要

環境貢献度評価

●評価条件

お客様受付等の接客対応の工程において、開発技術を用い、接客対応音声に合わせた自然な動作を自動付与したデジタルヒューマンで遠隔対応する場合と、従来通りの現場へ移動し対面で接客対応を行う場合でのCO₂排出量を比較することで、開発技術の環境貢献度を定量化しました。

●評価結果

本技術を適用した場合の環境貢献度は、年間960回の接客コミュニケーションあたり1,083kg-CO₂（削減率：25%）でした。

主な削減要因は、自動で自然な動作をするデジタルヒューマンを用いた遠隔からの接客コミュニケーションにより、従来と比べて現場へ移動せずに自然な接客対応ができるため、人移動の削減ができます。

