

# デジタル工房における手の動きと会話音声の AR 提示による ノウハウ流通支援

根岸 蒼天<sup>1,a)</sup> 角 康之<sup>1,b)</sup>

**概要：**本研究は、デジタル工房における工作機器のノウハウ流通を目的として、過去の作業者の手の動きと会話音声を記録・提示する AR システムを提案する。デジタル工房にある 3D プリンタやレーザーカッター等の工作機器を使用することで人の手ではできない高度な造形が容易に可能である一方、材料の種類や状態、目的に合わせた高度な設定が必要となる。紙のマニュアルや映像では簡単な使い方のみで、高度な設定については説明されないことが多い。また、マニュアルや映像を見て作業することを交互に繰り返すと作業効率の低下や必要な作業を行わずに進めてしまう危険性がある。そこで、過去の作業者の会話音声と手の動きの可視化を行うことでノウハウの流通を支援できると考える。本研究では、Meta Quest 3 のハンドトラッキングとマイクで取得したデータを会話音声を言霊として手の動きと共に空間に埋め込み Unity/Meta XR SDK で可視化し、複数作業の同時提示と近接作業の自動強調を実現した。これにより、作業と並行した学習や、会話中の指示語の理解の促進をすることでノウハウの効率の良い流通を支援する。

## 1. はじめに

本研究は AR 環境で過去の作業者の手の動きと会話音声を提示することで、デジタル工房における工作機器に関するノウハウを効率よく流通させることが目的である。デジタル工房とはレーザーカッターや 3D プリンタなどの工作機器を備えた工房である。工作機器を使用することで人の手では簡単に作れないものを作成することが可能となっている。

工作機器を上手に使うことで高度な造形が可能となる。工作機器を効率よく使いこなすためには単なる使い方だけでなく、材料の大きさや状態、作るものに合わせて高度な設定や調整などのノウハウを身に付ける必要がある。しかし、一人で作業する場合にはノウハウがあまり身につかない可能性がある。ノウハウを含む使い方を身に付けるためには、その場にあるマニュアルや、使い方の映像を見る方法がある。その場のマニュアルには作業の内容は書かれているが、なぜその作業をするかの理由が書かれていないことが多く作業内容に合わせた柔軟な設定方法を身に付けることは難しい。使い方の映像を見る方法では、映像が長くなる場合前半の内容を忘れてしまいやすい。また、マニュアルや映像では基本的な使い方の説明のみで、高度な設定や調整の方法の説明がされないことが多いさらに、マニ

アルや映像を見て作業をするという動作を交互に繰り返すことによって必要な作業が行われなくなる危険性も存在している。

Clemens らは、AR 上で情報を提示することで、手続き的知識を人が教える場合と同程度に教えることができると示した [1]。よって、AR 上で手の動き情報と会話音声を提示することにより、ノウハウの効率の良い流通が可能であると考えられる。また、AR であれば複数の手の動きを同時に提示することが可能である。よって教えながら作業している二人組の手の動きと会話音声を同時に複数組提示することで自分の作業進度や内容に近い情報を得られやすくなる。

Fujikura らは、過去の会話音声を類似状況下の人に対して提示することで知識の共有を促進するシステムを提案した [2]。このシステムをデジタル工房で使用した際、過去の会話音声を聞いたことでより適切に 3D プリンタを使用できたこと報告されている。よって、過去の会話音声を空間に言霊として埋め込み、提示することで工作機器をより適切に使いこなせるようになると考えられる。会話音声であれば、作業の邪魔にならずに聞くことが可能なので作業しながらでも情報を受け取ることができる。しかし、レーザーカッターなどの工作機器のようにいくつかのボタンを押す必要がある場合は「このボタンを押して」などのような指示語が含まれる可能性がある。その場合、会話音声のみでは指示語が何を指示しているかが分かりにくい。そこで、

<sup>1</sup> 公立はこだて未来大学

<sup>a)</sup> s-negishi@sumilab.org

<sup>b)</sup> sumi@acm.org



図 1 記録と提示からなるシステムの流れ

作業者の手の動きを記録し、会話音声とともに可視化することで会話中の指示語が分かりやすくなり、効率よく過去の作業者のノウハウを流通させることができると考えられる。手の動きを AR で表示することで、作業しながら情報を受け取ることが可能である。また、Fujikura らの提案したシステムでは、どの音声を流すべきかの状況は実験者が判断し手動で流していたが、空間トラッキング機能のある HMD の登場により、会話音声を言霊として空間に埋め込むことが可能になった。

以上より本研究では、ヘッドマウントディスプレイ (HMD) を被り過去の作業者の手の動きと会話音声を記録し言霊として空間に埋め込み、新たな作業者に提示することによるノウハウ流通法を提案する。

図 1 はシステムの流れである。左側が記録するときの例で右側が提示するときの例である。左側の記録では、作業者の会話音声と手の動きを HMD を利用して記録する。右側の提示では、記録された手の動きと会話音声を AR 環境で提示する。

## 2. 関連研究

AR によるノウハウ流通支援に関する研究として、Hoffmann らは製造現場の機械作業ノウハウを AR で可視化、共有した [3]。AR を使用し、熟練工のノウハウを可視化・共有できるようにした研究で、AR 上で機械操作の際に圧力や速度などのパラメータの基準値や現在の値を表示することで、ノウハウを共有した。しかし、情報過多が指摘されており、作業中の認知負荷の増大が問題となっている。また、Nils らは、AR 上でテキストや音声、画像、映像を提示することでノウハウを共有し、機械操作作業の準備作業の習熟支援を目指した [4]。作業内容として、工具や機械部品の位置の確認や修正などのような機械操作作業の準備段階における作業を支援している。これらの研究では、事前に熟練工や専門家が AR 上に提示するテキストや、音声などを作成している。本研究では、事前に提示するデータの作成として、2 人で作業している手の動きと会話音声をそ

のまま利用するので AR 上に提示するデータの作成は容易である。

手の動きの可視化を含む作業支援の研究として、Yuanzhi らは、機械操作作業において、半身や全身のアバターを AR 上で提示することによる効果を検証した [5]。AR で提示する情報として、ビデオ、AR のシンプルな視覚ガイド、半身のアバター、全身のアバターの 4 つを比較した。結果として、ボタンを操作するなどの空間的な移動や対象物の位置の特定を含まれる動作において半身のアバターが最も効果的であった。また、Alexander らは、複雑で長時間の組立作業に対して、手の動きを含む AR での情報提示の効果を検証した [6]。8 時間の作業において、HMD で AR 上に情報を提示し支援したところ、紙のマニュアルに比べてエラー率が約 60% 減少した。しかし、エラー率が低下した一方、ユーザに対する認知的負荷が上昇するという結果になった。これらの研究は、手の動きなどを AR で提示することで作業の効率が上がる可能性を示している。一方 AR により認知負荷が上昇する可能性が示唆されている。本研究では、AR で表示する情報を手のみの 3D モデルと音声にすることで認知負荷を抑えつつノウハウの流通を目指す。

AR による作業支援に関する研究として Arthur らは手作業の支援において AR での情報提示による有効性を検証した [7]。AR により空間上に情報を提示することで、紙マニュアルや画面表示に比べて誤り率や認知負荷が低減された。また、Andrew らはスプリンクラー室において行う作業を AR で提示することによる効果を検証した [8]。結果として、作業手順のスキップ率が、20 分の 1 にまで減少した。また、AR システムの利用者は作業の順番を間違えることもなかった。しかし、作業の時間は増加した。よってこの研究では AR を利用することで、作業時間は増加するが作業をより正確に行うことができるようになる可能性が示された。また、Tadeja らは 3D プリンタのノズル詰まり修理を対象に、AR ベースのマルチモーダルガイドシステムの有効性を検討した [9]。ガイドとして、テキスト、音声、画像、3D モデルの 4 種類の案内を組み合わせられており

ユーザがUIを自由に移動、回転、拡大縮小ができるようになっている。ユーザがUIを自分で調節してユーザにあったUIにすることにより認知負荷が低減される可能性を示唆している。これらの研究は作業手順をAR環境で情報を提示することで、エラー率や認知負荷を下げる可能性を示した。本研究では、デジタル工房における実作業の二人での会話と手の動きを記録しAR上で提示することで、紙や画面のマニュアルではあいまいになりがちな指示語の解消を行いより正確な情報提示を行う。同時に、視界の占有率を抑え認知負荷を抑えて実現することを目指している。

### 3. 提案するシステム

本研究では、作業者の会話音声と手の動きを記録するシステムと、記録された手の動きと会話音声を言霊として空間に埋め込み、AR環境で提示するシステムを提案する。過去の作業者の手の動きと音声を記録し音声と手の動きを同期させて、記録時と同じように提示することができるシステムである。このシステムの目的は効率の良いノウハウ流通を支援することである。以下では提案するシステムについて説明を行う。

#### 3.1 Meta Quest 3について

本システムでは、手の動きと音声の記録、提示を行うHMDとしてMeta社のMeta Quest 3を使用した。Meta Quest 3はVRアプリの実行の他にフルカラーのビデオパススルーを利用したAR、MRアプリを実行することが可能である。手の動きと音声の記録・提示はこのビデオパススルーを利用して、AR環境で動作する。ビデオパススルーを利用することによって、HMDを被りながらも現実世界を見て作業することが可能となる。記録に関してはMeta Quest 3のハンドトラッキング機能とマイク機能を用いることで実現している。よって、図2のように手の動きと会話音声の記録と提示をMeta Quest 3のみで実現することが可能となっている。

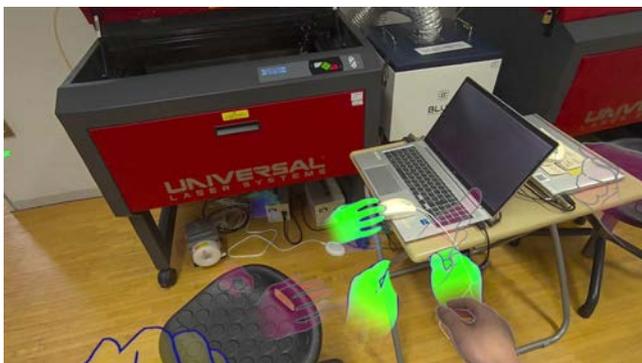


図2 動作イメージ

#### 3.2 手の動きと会話音声の記録

作業者の手の動きと会話音声を記録するシステムを紹介する。このシステムの目的はノウハウを効率良く流通させるための準備として過去の作業者の手の動きとそれに付随する会話音声を記録することである。図3のように2人でHMDを被りシステムを使用しながら作業することで手の動きと、会話音声を記録することができる。

手の動きはMeta Quest 3のハンドトラッキング機能を利用し取得できる手の各関節の座標と姿勢を保存した。また、会話音声は同デバイスのマイクを利用し保存した。手の動きを保存する際の手の各関節の座標や姿勢はシステム内の座標空間に準じている。なので、記録を開始する前に事前に基準の位置を定めその位置に立ち座標のリセットを行う必要がある。座標のリセットは、Meta Quest 3の右コントローラーにあるMetaのマークを長押しすることで可能である。



図3 作業しながら手の動きと会話を記録する例

#### 3.3 記録された手の動きと音声の提示

前節で説明した手法で記録された過去の作業者の手の動きと会話音声のデータをAR環境で手の動きと会話音声を提示するシステムを紹介する。このシステムの目的は、過去の作業者の手の動きと発話を提示することで、デジタル工房内における工作機器使用時のノウハウを効率よく流通させることである。PrasanthはARの組立訓練において、手のジェスチャは視線単独より有効であることを示し、視線と併用することで認知負荷が下がることを示している[10]。本研究では、会話音声と同期した複数の手の動きを提示することで効率的なノウハウ共有を目指している。

AR環境を使用することで、過去の作業者複数人の手の動きと会話音声を同時に提示できる。なので、教えながら2人1組で作業している際の手の動きと会話音声を複数組同時に提示することが可能である。作業の組ごとに色分け

をすることで、どの手と同じ組同士なのか分かりやすくなるようにした。また、手の動きの可視化のために手首から上のみの3Dモデルを使用した。そして、提示する音声については空間音響を用いて各作業者の右手の位置から聞こえるようにした。

同時に複数組提示される際、作業をしている人数分の作業会話が同時に聞こえることになる。全ての作業会話音声と同じような大きさで聞こえると、手の動きとの対応付けが難しくなりユーザが情報を正しく受け取れなくなってしまうと考えられる。そこで、ユーザから一番距離の近い作業者の組が強調されるようにした。図4は、ユーザからの距離により強調度を変化させた例で、青い丸で囲ったピンクと、青色の手の色を薄くすることで、赤い丸で囲った黄色の手の強調度を上げている。このように一番近い組以外の手の色を薄くし、音も小さくすることで一番距離が近い組が相対的に強調されるようになっている。このようにすることで、ユーザから一番距離が近い作業の組の情報を正しく集中して受け取れるようになると思う。さらに、一番近い作業者の組なのでユーザが行っている作業と関連している可能性が高い作業の情報を無意識に選択して受け取れるようになると思う。

前節と同様に手の動きの可視化のための記録はシステム空間内のものに準拠している。よって記録した際に使用した位置で座標のリセットを行う必要がある。

手の動きを可視化することで会話音声からは分からないどのような行動をしたのかが明確にわかるようになる。また、会話音声からは作業の説明により、なぜその行動をしたのかが理由がわかるようになる。よってこのシステムを利用することで、過去の作業者が何をしたのか、なぜしたのかがわかるようになると思う。



図4 強調度を変化させた例

## 4. システムの実装

### 4.1 実装環境について

本研究ではMeta Quest 3で実行することができる、手の動きと音声を記録する機能を備えたアプリケーションと提示する機能を備えたアプリケーションを開発した。開発

環境として、Unity2022.3.30f1を使用した。以下では開発に使用した主要なSDKであるMeta XR SDKと記録、提示システムそれぞれについて説明を行う。

### 4.2 Meta XR SDKについて

Meta Quest 3で動作するアプリケーションを使用するために、Meta社が提供しているMeta XR SDKを利用した。このSDKはMeta社のHMDで動作するアプリケーションを効率的に作成できるようにするために設計されている、このSDKを利用することでパススルーや空間アンカーといったMR、AR機能に加え、手や頭、コントローラのトラッキングや空間オーディオなどの機能を実装することができる。手の動きと会話音声の記録、提示それぞれのアプリケーションがARで動作するためにパススルー機能を利用した。

### 4.3 手の動きと会話音声の記録について

本節では手の動きと会話音声を記録するシステムの実装について説明を行う。手の動きの記録にはMeta Quest 3に内蔵されているハンドトラッキングの機能を利用した。ハンドトラッキング機能で得ることができる、手の24個の関節点それぞれの座標と姿勢の値に加えて記録開始時からの時間情報を毎フレーム記録しCSVファイルとして保存するようにした。座標はx,y,zの3軸で記録され、姿勢はクォータニオンの4軸で記録される。よってCSVファイルに記録されるデータの1行には24個それぞれの関節点につき7種類の値と時間情報を加えた169列のデータが記録される。この169列のデータを毎フレーム記録するので、行は記録したフレーム数となる。本システムは約60FPSで動作するので、例えば10分の記録では169列で約36000行のCSVファイルで保存される。座標や姿勢の値は追加の座標変換などを行わずUnity内で取得できる値をそのまま利用した。手の動きの記録は右手、左手のそれぞれのCSVファイルが出力されるようにした。

保存される各関節点の座標と姿勢の値の基準はUnityアプリケーション内の基準に準じている。この座標の値にはUnityアプリケーション内の原点からの相対位置が記録される。また、手の動きの記録に時間情報を付与することにより記録された会話音声との同期ずれが発生しないようにした。

会話音声の記録について同デバイスのマイクを利用しWAVファイル形式で出力するようにした。

### 4.4 記録されたデータの提示について

本節では記録された手の動きと会話音声を提示するシステムの実装について説明を行う。記録された手の動きのCSVファイルと、会話音声のWAVファイルから会話音声を言霊として手の動きと共に空間に埋め込み、提示をする

機能を実装した。手の動きを可視化するための 3D モデルとして、Meta XR SDK 内にある手の 3D モデルを使用した。この手の 3D モデルの関節点は、手の動きを記録する際に使用したハンドトラッキングで取得することができる関節点と一致している。よって、記録された手の各関節点の座標と姿勢の値を対応している関節点に直接入れるだけで手の動きを再現することが可能になる。

ユーザから一番距離が近い作業者の組を強調するために、ユーザの右手から一番距離が近い作業の組以外の手の透明度を上げることで色を薄くしました、音を小さくした。色を薄くするのは、手の 3D モデルについている色を管理するマテリアルオブジェクトの透明度を直接変更した。そのようにすることで、同じマテリアルオブジェクトがついている 3D モデルの色がまとめて変わるようになる。なので、同じ作業の組には同じマテリアルオブジェクトを割り当てた。

本アプリケーションでは手の動きや会話音声のデータをローカルで管理している。なので、新しいデータの提示をする場合は、手のモデルや割り当てるマテリアルオブジェクトなどを適切に設定をして再度アプリケーションのビルドをする必要がある。

手の動きを可視化するために手の動きが保存された CSV ファイルを使用している。この CSV ファイルは 1 行あたり、手の関節点 24 個×座標と姿勢の数である 7 個に時間情報を加えた 169 列となっている。行の数は記録したフレーム数であり、手の動きを記録するアプリケーションは約 60fps で動作するので、例えば約 10 分の記録の場合約 36000 行が書き込まれた CSV ファイルが出力されることになる。この CSV ファイルを提示する手の数だけ使用するので、ロード時のメモリ不足が懸念される。その対策として、Unity の Stream Reader を使用した。Stream Reader を使用することで、CSV ファイルを一括でロードすることなく逐次的にロードすることが可能である。逐次的にロードすることで、手の動きが保存された CSV ファイルのサイズが大きいものであってもメモリの消費を抑え、メモリ不足のエラーが発生しづらくなるようにした。

手の動きを記録する際に、手が HMD の前になく、横や下に大きく離れたタイミングでハンドトラッキングができなくなることがある。その場合、手の動きを記録している CSV ファイルには手がアプリケーション内の原点にいる場合の座標、姿勢の値が記録される。そのデータをそのまま手の 3D モデルに適用させると記録した際にハンドトラッキングができなかったタイミングで手の 3D モデルがアプリケーション内の原点に瞬間移動してしまう。その対策として、手の座標、姿勢のデータの値が原点にいる場合と同じであった場合、値の適用をスキップするようにした。スキップすることで、手の 3D モデルが原点に移動せずにハンドトラッキングができなかったタイミングの直前の

座標、位置で手が停止するようにした。

## 5. 期待される効果

本システムでは過去の作業者の手の動きの可視化と発話の提示を可能にした。本章では、本システムの期待される効果や利点について考察する。

### 5.1 作業と並行した使用

本システムでは、AR 環境を使用し情報を提示することにより視界を遮らずに作業しながらシステムの利用が可能である。これにより現在の提示された情報の内容をすぐに実践できるので工作機器使用時のノウハウが身につけやすくなると考えられる。また、提示する情報は手の動きと音声のみであるため視界の占有率が高くなりづらいので、作業を続けやすくなる。映像を提示する場合、AR では視界の占有率が高くなるので作業しながら情報を受け取るとは難しいが、視界に提示する情報を手の動きのみにすることで作業しながらでも問題なく情報を受け取れると考えられる。

### 5.2 手の動きの可視化による効果

本システムでは、発話の提示と同時に手の動きを可視化することが可能である。これにより、過去の作業者同士の会話の中に「ここ」などの指示語が含まれている場合に指さしの手の向きなどからわかるようになると考えられる。よって指示語を含む会話内容が理解しやすくなると考えられる。また、過去の作業者が押したボタンの特定など、実際の作業内容が分かるようになる。よって、会話音声のみからでは分かりづらい作業内容が理解しやすくなると考えられる。

### 5.3 提示する情報の強調度の変化について

本システムでは、ユーザから一番距離が近い手の組を強調して表示することが可能である。これにより、複数の作業の組の情報を同時に提示する際にユーザが行っている作業と関連の深い可能性が高い情報を集中して受け取ることができると考えられる。

### 5.4 デバイスについて

本システムは手の動きと会話音声の記録から提示まで、HMD のみで動作する。これにより手の動きを記録するための専用のハンドトラッキングの機材を導入する必要がなくなる。デジタル工房では、工作機器が置かれており狭い可能性があるためハンドトラッキングの機材を導入することは難しい場合があるが、HMD は頭に装着するデバイスなので狭い環境でも使用が可能である。

## 6. システムの課題点

### 6.1 記録する手の座標

本システムでは、記録、提示する際の手の座標の基準は Unity アプリケーション内の基準に準じている。そのため、システム使用時には事前に決めた特定の位置で毎回座標のリセットを行う必要がある。リセットを行う位置がずれてしまうと、手の位置がずれてしまい工作機器のどのボタンを押したのかなどの情報が分かりづらくなってしまう。これに対して Meta Quest 3 の空間アンカーの機能を使用することで対応が可能と考えている。空間アンカー機能では、仮想的なオブジェクトを現実世界基準で配置することができる。よって、空間アンカー機能を使用することで手の動きの座標を現実世界基準にすることで安定して提示することができるようになると考えられる。また、毎回の座標のリセットが必要なくなることでより簡単に使用することが可能である。

## 7. おわりに

本研究では、デジタル工房における工作機器使用時のノウハウの流通を支援するために、AR 環境で過去の作業者の会話音声と言霊として手の動きと共に空間に埋め込み、提示するシステムを提案した。過去の作業者の手の動きと発話を記録、提示するデバイスとして Meta Quest 3 を使用した。Meta Quest 3 のハンドトラッキング機能とマイク、ビデオパススルー機能を使用することで、手の動きと会話音声の記録、AR 環境での提示を可能にした。AR 環境を使用することにより、会話音声の空間への埋め込み、複数人の作業者の手の動きの提示を可能にした。また、使用者に最も近い作業者の組を強調して提示する機能を備えている。

期待される効果として、作業と並行したシステムの利用による学習の促進、手の動きの可視化による指示語を含む会話の理解促進が挙げられた。その一方で、手の位置を記録、提示するための座標がアプリケーション基準であることによって使用時に毎回の座標リセットが必須である点が課題として残っている。

今後は、空間オーディオによる発話位置の提示と、空間アンカーによる現実世界基準の座標を利用した安定化を実装をする。以上により、デジタル工房における効率の良いノウハウ流通手法の確立を目指す。

**謝辞** 本論文の執筆にあたり、文章の作成において一部 ChatGPT を利用した。生成された結果の責任は筆者に帰属する。

## 参考文献

- [1] Clemens Hoffmann, Sebastian Büttner, and Michael Prilla: *Conveying Procedural and Descriptive Knowledge with Augmented Reality*, Proceedings of the 15th international conference on pervasive technologies related to assistive environments. 2022.
- [2] Fujikura Ryo, Yasuyuki Sumi: *Facilitating Experiential Knowledge Sharing through Situated Conversations.*, Proceedings of the Augmented Humans International Conference. 2020.
- [3] Sven Hoffmann, Thomas Ludwig, Florian Jasche, Volker Wulf, and David Randall: *RetrofittAR: Supporting Hardware-Centered Expertise Sharing in Manufacturing Settings through Augmented Reality*, Computer Supported Cooperative Work (CSCW) 32.1 (2023): 93-139
- [4] Nils Darwin Abele, Sven Hoffmann, Aparecido Fabiano Pinatti De Carvalho, Marcus Schweitzer, Volker Wulf, and Karsten Kluth: *Knowledge and expertise sharing—designing an AR mediated cyber physical production system for industrial set up processes*, Congress of the International Ergonomics Association. Cham: Springer International Publishing, 2021.
- [5] Yuanzhi Cao, Xun Qian, Tianyi Wang, Rachel Lee, Ke Huo, and Karthik Ramani: *An Exploratory Study of Augmented Reality Presence for Tutoring Machine Tasks*, Proceedings of the 2020 CHI conference on human factors in computing systems. 2020.
- [6] Alexander Riedel, Johanna Gerlach, Maximilian Dietsch, Frank Engelmann, Nico Brehm, and Tobias Pfeifroth: *Evaluating augmented reality, deep learning and paper based assistance systems in industrial manual assembly*, IFIP International Conference on Advances in Production Management Systems. Cham: Springer Nature Switzerland, 2023.
- [7] Arthur Tang, Charles Owen, Frank Biocca, and Weimin Mou: *Comparative effectiveness of augmented reality in object assembly*, Proceedings of the SIGCHI conference on Human factors in computing systems. 2003.
- [8] Andrew Pringle, Stefanie Hutka, Jesse Mom, Robin van Esch, Niall Heffernan, and Paul: *Ethnographic study of a commercially available augmented reality HMD app for industry work instruction*, Proceedings of the 12th ACM International Conference on Pervasive Technologies Related to Assistive Environments. 2019.
- [9] Tadeja Sławomir K., Solari Bozzi Luca O., Samson Kerr D.G., Pattinson Sebastian W., and Thomas Bohné: *Exploring the repair process of a 3D printer using augmented reality-based guidance*, Computers & Graphics 117 (2023): 134-144.
- [10] Prasanth Sasikumar, Soumith Chittajallu, Navindd Raj, Huidong Bai, Mark Billingham: *Spatial perception enhancement in assembly training using augmented volumetric playback.*, Frontiers in Virtual Reality 2 (2021): 698523.