This article is a technical report without peer review, and its polished and/or extended version may be published elsewhere.

第 30 回日本バーチャルリアリティ学会大会論文集（2025 年 9 月）

メタバース環境における手書きスケッチの 3 次元オブジェクト変換システム

倉井龍太郎 [1)], 畑田裕二 [2)], 廣井裕一 [3)], 平木剛史 [3,4)]

1) 奈良先端科学技術大学院大学 先端科学技術研究科 （〒 630-0192 奈良県生駒市高山町 8916-5, ryu.kurai@gmail.com）

2) 東京大学 大学院情報学環（〒 113-8654 東京都文京区本郷 7-3-1, hatada@cyber.t.u-tokyo.ac.jp）

3) クラスター メタバース研究所（〒 141-0031 東京都品川区西五反田 8-9-5, y.hiroi@cluster.mu）

4) 筑波大学 図書館情報メディア系（〒 305-8550 茨城県つくば市春日 1-2, hiraki@slis.tsukuba.ac.jp）

概要: 本研究では、メタバース空間でのインタラクティブな 3D コンテンツ制作を実現する手書きスケッチ変換システムを提案する。このシステムにおいてユーザーは空間内でペン型ツールを用いて 3D 軌跡を描く。システムは生成 AI を用いて描かれた軌跡から 3D オブジェクトを生成しメタバース内に配置・共有することができる。このシステムの利用により、複数ユーザーが共有空間で手軽に 3D オブジェクトを創造・共有できる環境を実現した。

キーワード： メタバース, スケッチベースモデリング, 3D 生成 AI, VR インタラクション

1. はじめに

VRChat[1], cluster[2], Resonite[3], Rec Room[4]など、アバターを介して他者と社会的なコミュニケーションを行うことができるバーチャル環境であるソーシャル VR プラットフォームは、現在に至るまでユーザー数を大きく伸ばし続けている。特に近年では、こうしたソーシャル VR プラットフォームは、VR ヘッドマウントディスプレイ（HMD）を用いた没入体験のみならず、PC やモバイル端末からアクセスできる非没入型の参加形態をも包括した「メタバース」へと進化しつつある [1]。メタバースは今後、多様なデバイスのユーザーが一同に介し、日常的な交流を行うプラットフォームになることが期待される。

これらのメタバースフォームは、単独で体験するバーチャル環境とは異なり、複数の他者とのコミュニケーションが前提となる創造的・協働的な環境である [1]。特に、User Generated Content（UGC）と呼ばれるユーザー主導の創作物を起点とした文化や活動が、現代のメタバースを従来のバーチャル環境（オンラインゲームなど）と差別化する重要な要素となっている。ユーザーは単なる観察者ではなく、カスタマイズされたオブジェクトや完全な仮想世界を作成し、コミュニティ内で共有・販売することで、メタバースの形成に積極的に参加している。

しかし、これらのプラットフォームにおいて 3D 空間やコンテンツを作成するクリエイターの数は依然として限定的である。この主な要因として、3D オブジェクトの作成には高度なモデリングスキルに加え、各プラットフォーム固有の設定や実装に関する専門知識が必要となることが挙げられる。

他方、近年の生成 AI 技術の急速な発展により、テキストや画像から 3D モデルを生成する手法が民主化されつつあることで、従来は専門的なスキルを要した 3D コンテンツの制作がより直感的で身近なものになりつつある。しかし、こうした生成 AI 技術とメタバースプラットフォームとの統合は未だ充分ではなく、生成された 3D モデルをメタバース空間に持ち込むには依然として技術的なハードルが高い。

このような背景から本研究では、2D のスケッチからインタラクティブに 3D オブジェクトを生成し、それをメタバース空間で他者と共有できるシステムを開発した。本システムの特徴は、(1) メタバース空間内・外でのインタラクティブなスケッチ入力 (図 1)、(2) 生成 AI 技術を活用した高品質な 3D 生成、(3) メタバースへのシームレスなシステム統合、の 3 点である。これにより、3D モデリングの専門知識を持たないユーザでも、メタバースをプレイするデバイスの特性に依らず、3D モデルの作成とメタバースプラットフォームへのアップロードを容易に行うことが可能となる。

2. 関連研究

2.1 スケッチベース 3D モデリング

スケッチベースの 3D モデリング手法は、2D 入力から 3D 形状を生成するインタラクティブなインターフェースとして研究されてきた。特に Igarashi らによる Teddy [2] は、簡単な 2D スケッチから滑らかな 3D 自由形状を生成する先駆的なシステムとして知られている。Teddy では、閉じた 2D 輪郭線から膨らみを持つ 3D 形状を自動生成し、さらなるスケッチ操作により形状の切断や押し出しなどの編集を可能にした。

[1]https://hello.vrchat.com/
[2]https://cluster.mu/
[3]https://resonite.com/
[4]https://recroom.com/

© 2024 日本バーチャルリアリティ学会

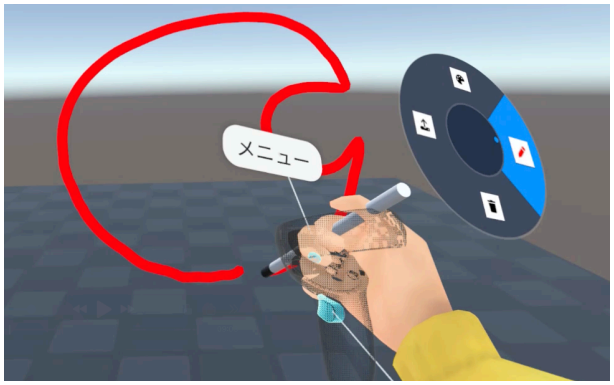

図 1: メタバース空間内でペン型ツールを操作し絵を描く様子

2.2 AI 技術を用いた 2D-3D 変換

近年、深層学習技術の発展により、テキストや画像から 3D モデルを生成する研究が活発化している。DreamFusion [3] は、2D テキストから 3D 表現を生成する手法として、Neural Radiance Field (NeRF) と拡散モデルを組み合わせた Score Distillation Sampling を提案した。Magic3D [4] は、この手法をさらに発展させ、高解像度のテクスチャメッシュ生成を実現している。Stable Fast 3D [5] は、単一画像から高品質な 3D アセットを 0.5 秒で生成する手法を提案し、リアルタイム応用への可能性を示している。著者らは生成 AI を活用し自然言語から 3D モデルを生成しメタバースで利用する Web アプリケーション MagicCraft [6] を提案している。本研究では、これらの生成 AI による 3D モデル生成手法を活用し、さらに 3D 空間内でのスケッチを利用することでメタバースに完結したシステムの構築を行う。

3. 提案システム

3.1 システム概要

提案システムは、VR 空間内での手書きスケッチを 3D オブジェクトに変換し、メタバース空間に自動配置する統合的なシステムである。図 2 にシステムの全体構成を示す。

3.2 VR 空間でのスケッチ入力

ユーザーは VR 空間内で、アバターの手に固定されたペン型ツールを使用して 3 次元的な描画を行う 1。ペンツールには以下の機能を呼び出す UI が付属している。

- 空間への描画モード切替
- 色の選択（パレットから選択可能）
- 描画内容の削除
- 描画完了時のアップロード実行

ユーザがアップロードを選択すると、システムはペンの 3D 軌跡座標とユーザーの視点情報（カメラ座標）を記録する。

3.3 AI を活用した 3D 生成パイプライン

記録された 3D 軌跡は、ユーザーの視点からレンダリングされ、PNG 画像として保存される。この画像は以下で述べる 3 つの処理パイプラインを経て 3D モデルに変換される。

3.3.1 画像の洗練化

画像生成 AI（OpenAI GPT Image API）を使用し、スケッチ画像を 3DCG レンダリング風の画像に変換する。プロンプトの工夫により元のスケッチの形状と雰囲気を保持しながら、詳細なテクスチャと陰影を追加する。画像の変換に以下のプロンプトを用いた。

> 添付された画像は、オブジェクトの概要を示したアウトラインです。
> このアウトラインを元に、オブジェクトを想像し描画してください。
> 描画する画像は、3D の CG のように描画してください。
> オブジェクトの全体像が表示されるように描画してください。
> もし、画像の中に文字列がある場合は、その文字列はオブジェクトの名前や説明です。
> 文字列の説明を画像を描画に利用してください。
> 文字は絶対に描画してはいけません。

3.3.2 3D モデル生成

3D モデル生成 AI（Stability AI Stable Point Aware 3D API）を使用し、洗練化された画像から 3D モデル（GLB 形式）を生成する。この API は単一画像から高品質な 3D メッシュとテクスチャを生成可能である。

3.3.3 メタバースへの統合

提案システムは生成された 3D モデル（GLB ファイル）に、cluster プラットフォーム用のメタデータを付加する。これにより、cluster の空間内でアバターが自由に持ち歩いたり空間内に設置したりできる 3D オブジェクトであるクラフトアイテムとして 3D モデルをアップロードできる。オブジェクトは元のスケッチ位置に配置される。処理中は黄色のプレースホルダーキューブが表示され、ユーザに生成状況を視覚的にフィードバックする。

3.4 Web ベース入力インターフェース

HMD を使用しないユーザ向けに、Web アプリケーションからの画像アップロード機能も提供する (図 3)。この場合、2D 描画ツールで作成した画像を直接システムに入力でき、以降の処理は VR 空間での入力と同様のパイプラインで実行される。

4. 評価と考察

提案システムの性能評価として、68 件の入力データに対する処理時間を測定した結果、軌跡データの受信完了から 3D モデル生成完了まで平均 35.5 秒（標準偏差 10.2 秒）を要した。ただし計測データの中に完了までの時間が 600 秒を超えるものが 1 件あったので外れ値として除去した。この処理時間は、リアルタイムインタラクションには課題が残るものの、創造的な活動においては許容範囲内と考えられる。生成された 3D モデルの品質については、単純な形状では高い再現性を示したが、複雑な形状や細かいディテールの表現には限界が見られた。この改善に向けては、より

© 2024 日本バーチャルリアリティ学会

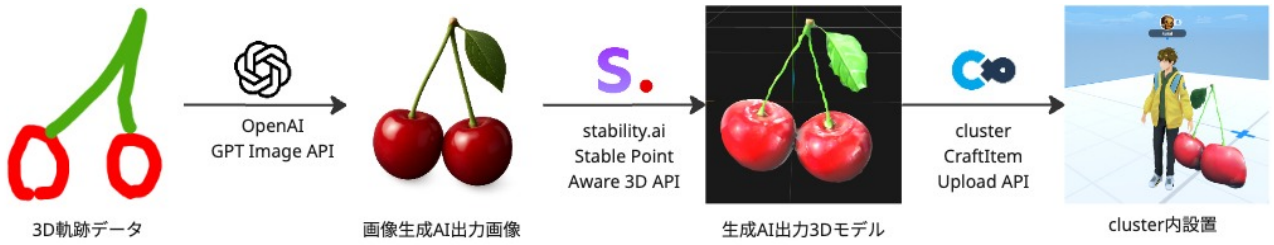

図 2: 3D 軌跡データから cluster 内で表示されるまでのデータの流れ

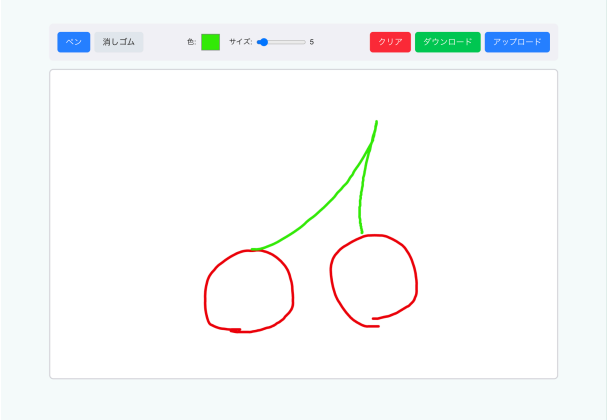

図 3: Web からスケッチを行い本システムに upload する画面のスクリーンショット

高度な 3D 生成モデルの採用や、複数視点からのスケッチ入力への対応が必要であると考えられる。

5. 応用展開

本システムは教育分野での活用が期待される。特に、3D 空間認識や創造性教育において、専門知識なしに 3D コンテンツを作成できる点は大きな利点となる。また、メタバース環境の特性を活かし、複数ユーザーによる同時利用が可能であることから、協調的な 3D コンテンツ作成やブレインストーミングツールとしての応用も想定される。スケッチ入力をベースとする本システムは、創作者の意図をリアルタイムで可視化し、周囲のユーザーがその創作プロセスを目撃し、即座にフィードバックを提供できる環境を作り出す。これは単独での制作作業とは対照的に、創造行為そのものを社会的な活動へと変換する可能性を持つ。

6. おわりに

本研究では、メタバース環境における直感的な 3D コンテンツ作成を実現するため、手書きスケッチから 3D オブジェクトを自動生成するシステムを提案した。VR 空間での自然なスケッチ入力と生成 AI 技術を組み合わせることで、専門知識を持たないユーザーでも容易に 3D コンテンツを作成できることを示した。今後は、生成品質の向上と処理速度の改善を進め、より実用的なシステムへと発展させていく予定である。

謝辞 本研究は、JST ムーンショット型研究開発事業 JPMJMS2013 の助成を受けたものです。

参考文献

[1] Georg David Ritterbusch and Malte Rolf Teichmann. Defining the metaverse: A systematic literature review. *IEEE Access*, Vol. 11, pp. 12368–12377, 2023.

[2] Takeo Igarashi, Satoshi Matsuoka, and Hidehiko Tanaka. Teddy: a sketching interface for 3D freeform design. In *ACM SIGGRAPH 2006 Courses on - SIGGRAPH '06*, New York, New York, USA, 2006. ACM Press.

[3] Ben Poole, Ajay Jain, Jonathan T Barron, and Ben Mildenhall. DreamFusion: Text-to-3D using 2D diffusion. *arXiv [cs.CV]*, September 2022.

[4] Chen-Hsuan Lin, Jun Gao, Luming Tang, Towaki Takikawa, Xiaohui Zeng, Xun Huang, Karsten Kreis, Sanja Fidler, Ming-Yu Liu, and Tsung-Yi Lin. Magic3D: High-resolution text-to-3D content creation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 300–309, 2023.

[5] Stable fast 3D. `https://www.stablefast3d.com/`. Accessed: 2025-7-14.

[6] Ryutaro Kurai, Takefumi Hiraki, Yuichi Hiroi, Yutaro Hirao, Monica Perusquía-Hernández, Hideaki Uchiyama, and Kiyoshi Kiyokawa. MagicCraft: Natural language-driven generation of dynamic and interactive 3D objects for commercial metaverse platforms. *IEEE Access*, Vol. 13, pp. 1–16, 2025.

© 2024 日本バーチャルリアリティ学会