

Title	Sketch2Robpla:スケッチベース関節定義インターフェースシステムの開発
Author(s)	品田, 康平; 謝, 浩然
Citation	情報処理学会第201回コンピュータグラフィックスとビジュアル情報学研究発表会, 2026-CG-201(7): 1-7
Issue Date	2026-03-02
Type	Conference Paper
Text version	publisher
URL	https://hdl.handle.net/10119/20349
Rights	<p>社団法人 情報処理学会, 品田康平, 謝浩然, 情報処理学会第201回コンピュータグラフィックスとビジュアル情報学研究発表会, Vol.2026-CG-201 No.7, 2026, 1-7. ここに掲載した著作物の利用に関する注意: 本著作物の著作権は(社)情報処理学会に帰属します。本著作物は著作権者である情報処理学会の許可のもとに掲載するものです。ご利用に当たっては「著作権法」ならびに「情報処理学会倫理綱領」に従うことをお願いいたします。 Notice for the use of this material The copyright of this material is retained by the Information Processing Society of Japan (IPSJ). This material is published on this web site with the agreement of the author (s) and the IPSJ. Please be complied with Copyright Law of Japan and the Code of Ethics of the IPSJ if any users wish to reproduce, make derivative work, distribute or make available to the public any part or whole thereof. All Rights Reserved, Copyright (C) Information Processing Society of Japan.</p>
Description	情報処理学会第201回コンピュータグラフィックスとビジュアル情報学研究発表会, 芝浦工業大学豊洲キャンパス, 東京, 2026年3月9日-10日

Sketch2Robpla: スケッチベース関節定義 インターフェースシステムの開発

品田 康平^{1,a)} 謝 浩然¹

概要: VLA モデルの研究によって、家庭におけるタスクをロボットが実行することが可能だということが示された。しかし現在のロボットは導入コストが高く、家庭に普及させることは難しい。近年では、家庭用 3D プリンタの普及によって、3D プリンタ製のロボットが安価にロボットを製作することができるようになってきたが、ロボット設計は広範な専門的知識を必要とするため未経験者には難しい。そこで本研究では、非専門家のロボット設計の支援を目的とし、スケッチベースインターフェースと 3D モデル生成 AI を用いたロボット設計システム Sketch2Robpla を提案する。Sketch2Robpla はユーザの入力スケッチからロボットの形状となる 3D モデルを生成し、生成したモデルをスケッチ操作で分割してロボット構造を設計することで、直感的な操作でのロボット設計を可能にし、ユーザ実験によって、専門的知識を持たないユーザのロボット設計の実現可能性を確認した。

Sketch2Robpla: Sketch-Based Robotic Joint Design Interface

KOHEI SHINADA^{1,a)} HAORAN XIE¹

1. はじめに

Physical Intelligence 社の開発した VLA (Vision Language Action) モデル π_0 [1] は自らの有意性を示す実験として、衣類をたたんで見せた。これにより、ロボットが人間の代わりに家事を行う未来の可能性が示された。しかし、一般家庭にロボットが普及するためにはロボットが高価だという問題がある。家庭用 3D プリンタの普及により、自宅や学校でも 3D モデルの立体化が容易になったことから、3D プリンタ製のロボットをがこの問題の解決策になるが、ロボットの設計は広範な専門的知識を必要とするため、参入障壁が高い。加えて、既存のロボット設計手法は CAD 型とモジュール型の二つに分かれており、CAD 型は三次元 CAD と厳密な数値パラメータ入力によって設計する方法で、自由度が高い反面、操作習得に時間がかかる。モジュール型は予め作成された部品のデータベースから部品を配置、接続する形で設計する方法で、習得が容易だが設計空間が限定される都合上、自由度が低い。

本研究では、既存のロボット設計手法の持つ形状自由度の高さと操作の容易さを両立する設計手法が存在しないという課題を解決するため、ユーザのスケッチを入力としてロボットの 3D モデルを生成することで形状設計を行い、生成した 3D モデルを線分で分割して関節を定義することで直感的にロボットの構造設計を行うことのできるシステム Sketch2Robpla (図 1) を開発し、ユーザ実験によって、専門的知識を持たないユーザのロボット設計を支援できることを確認した。また、提案システムを用いて設計したロボットを 3D プリンタで実際に製作、組立を行い、モータが回転することで関節が動作することを確認した。

2. 関連研究

2.1 3D モデル生成

言語や画像の入力を基に 3D モデルを生成する技術は、拡散モデルと共に発展してきた。TRELIS[2], Tripo[3] は単視点画像からの 3D モデル生成を行う研究である。また、OmniPart[4] は単視点画像からの 3D モデル生成に加えて、自動的な分割も行っている。本研究では、Trellis を

¹ 北陸先端科学技術大学院大学

^{a)} s24110225@jaist.ac.jp

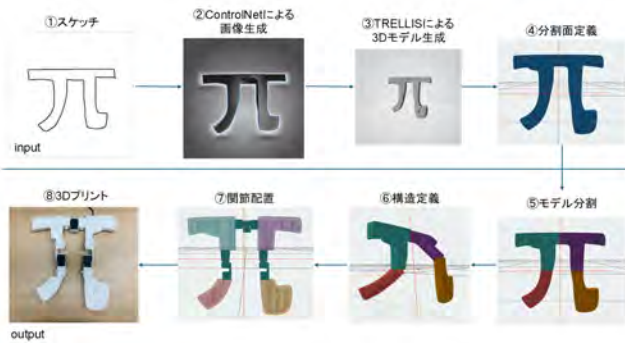


図 1: Sketch2Robpla の概要図

用いて 3D モデルを生成することで、ユーザが 3D モデリングソフトを使わずに形状設計をできるようにした。また、近年では 3D モデルの形状を LLM で理解することで、3D モデル生成の品質を向上させる試みも多く見られる。LLaMa-Mesh[5] は、3D メッシュのプレーンテキスト表現をトークン化して LLM で学習することでテキスト生成の性能を維持しながら 3D メッシュ生成を実現し、3D メッシュの理解や解釈、対話的な生成を可能にした。MeshLLM[6] は LLaMa-Mesh で問題だった学習効率を、プリミティブメッシュを用いることで解決し、分割したプリミティブメッシュから元の形状を再構成する学習プロセスにより、生成品質とメッシュの意味を説明する能力を向上させた。3D-LLM[7] は 3D 点群データを LLM に学習させるというアプローチで 3D 空間の理解を目指し、3D 空間のキャプション生成や質問応答を可能にした。ShapeLLM[8] は 3D 点群から多視点の 2D 画像を蒸留して学習した「Recon++」と呼ばれるエンコーダによって 3D 空間の理解力を向上させた。ShapeLLM-Omni[9] は 3D ボクセルを入力し、トークン化することで言語や視覚と同様に表現する。これによって 3D 生成や説明だけでなく、テキストによる 3D メッシュ編集を可能にした。

2.2 ロボット設計手法

既存のロボット設計手法は 3DCAD を用いる古典的な手法とある程度の役割を持ったモジュールどうしを接続してロボットを設計する手法に分かれている。特に後者の手法によって幅広いユーザがロボット設計を行うことが可能になった。例えば、Interactive Robogami[10] は足、胴、体などのパーツを並べてロボットを設計し、動力的な解析をフィードバックとして設計を支援する。パーツは展開図の形で出力され、折り紙のように折りたたんで組み立ててロボットを製作できる。Interactive Robogami は折り紙構造を用いることで製作を簡略化した。Megaró [11] は 3D プリントによって製作を簡略化した。この研究はユーザがロボットの構造を入力し、各リンクに一つの 3D モデルを割り当て、ロボットの形状と構造を設計する。その

後、設計したロボットで歩行シミュレーションを行い、安定して歩行できるように脚の太さや取付位置を変更し、歩行可能なロボット設計にする。そして、完成した設計を 3D プリンタで印刷して実世界でも歩行可能であることを確認した。RoboGrammar は、生物進化アルゴリズムに基づいて、リンクの長さの変更などを行う変異、歩行による移動距離の低い個体を淘汰する選択、既存の個体の特徴を組み合わせる新たな個体を生み出す交配などによりロボット設計を進化させてきたが、この方法はランダムな設計が生み出されるため、計算コストが高く、最適化に時間がかかった。この問題を解決するべく、ロボット設計の進化に LLM を導入したのが RoboMorph[12] である。RoboMorph はロボット設計と評価の結果を LLM に入力し、設計の変更案を出力することで RoboGrammar に比べて効率的にロボット設計を進化させることに成功した。

3. 提案システム

提案システム Sketch2Robpla はユーザのスケッチから形状生成を行うことで直感的に自由なロボット形状を設計することが可能である。また、生成したロボットを線で分割し、システムが自動的に関節を定義することで直感的にロボットの構造設計を行うことができる。設計した構造は運動学フィードバック機能によって確認でき、運動学の知識が無いユーザもロボットの動きを理解することができる。そして、設計したロボットを自動的に印刷用のデータに変換して出力するため、3D プリンタで部品を印刷、組立を行ってロボットを製作することができる。

3.1 スケッチ入力による 3D モデル生成

ユーザの思い描いた形状をロボットの形状に反映するために、スケッチを入力とした 3D モデル生成を実装した。3D モデル生成には TRELIS[13] を用いる。また、入力スケッチが簡素な場合でも 3D モデルを生成するために、スケッチから生成した画像を TRELIS に入力する。画像生成は ControlNet と Stable Diffusion によって行う。ControlNet によってユーザの入力したスケッチから構図、エッジなどの情報を抽出し、Stable Diffusion によって画像生成を行うことで、スケッチに基づいた画像生成を行うことができる。形状生成を行うインターフェースを図 2 に示す。図内赤枠でスケッチ入力、青枠で画像生成、橙枠で 3D モデルの生成を行う。緑枠の生成設定を図 3 に示す。生成設定を変更することで、生成される画像を変化させることができる。例えば、プロンプトにロボットの種類や質感を入力すれば、生成される画像の品質は向上する。スケッチよりも言葉の扱いが巧みなユーザにとって非常に有用である。生成ステップを増やすことも画像の品質を上げることにつながるがその分時間がかかるため注意する必要がある。スケッチとプロンプトのバランスを調整したい

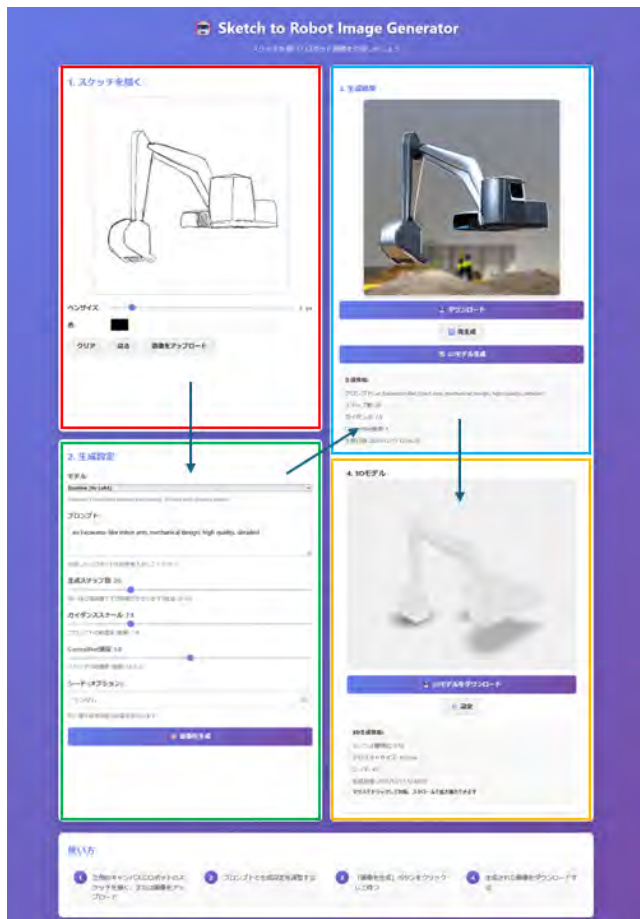


図 2: 形状生成インターフェース

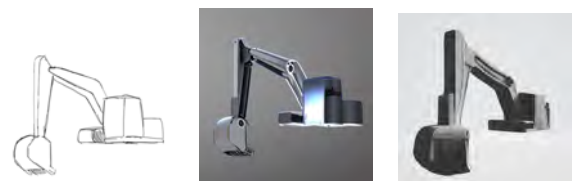
場合、ガイダンススケールと ControlNet 強度を操作する。プロンプトに書いたことがうまく反映されていない場合はガイダンススケールの値を上げることでプロンプトの影響度が上がる。スケッチの線に自信があり、完璧に画像に反映したい場合は ControlNet 強度を上げることでスケッチの影響度が上がる。最後にシードだが、デフォルトはランダムになっているが全く同じ生成設定の場合でもシードを変更することで別の画像が生成される。また、生成候補の画像のシード値を記録しておくことで、生成結果の再現も可能である。画像生成が完了した後は、TRELLIS を用いて 3D モデルを生成する。TRELLIS の場合もシードを変更することで、同一画像から複数の 3D モデルが生成可能であり、シードを控えておくことで再現が可能であるシステムを使って生成した例を図 4 に示す。

3.2 線分によるモデル分割

ロボットの構造入力は 3D モデル上に線分を引くことで行う。この方法は従来手法の数値入力と異なり、ロボットをパーツ分けするように直感的に関節を配置することができる。[Sketch Mode] を押下すると 3D モデルビューのカメラ位置が固定され、ドラッグ操作によって 3D モデルビュー上に線分を描画することが可能になる。そして [Cut



図 3: 生成設定



(a) 入力スケッチ (b) 生成画像 (c) 生成 3D モデル

図 4: 生成例

Model] を押下するとユーザの描画した線分の位置で 3D モデルが分割される。ユーザの線分の両端 2 点とカメラ視点の計 3 点から面を定義して分割する。カメラ視点を参照することで、分割結果がユーザの視覚と一致し、ユーザが直感的にモデルの分割を行うことができる。関節定義の従来手法では、関節位置、回転軸を数値で入力する必要があるが、Sketch2Robpla ではモデル分割時にシステムが自動で関節位置、関節回転軸を計算する。分割時、システムは断面の中心を関節位置として関節を配置する。回転軸に関してはユーザの意図が影響する部分であるため、断面中心を通過する最小断面面積を持つ断面と切断面の角度を比較し、45 度より大きい場合は切断面に垂直なベクトルを回転軸とし、45 度より小さい場合は切断面に水平なベクトルを回転軸とする。どちらの回転軸を採用するか、また切断面に水平なベクトルの平面上の向きの設定は画面右側の図 5 に示すフレームにて編集が可能である。同図下側の [詳細編集] を押下すると、図 6 に示すダイアログが表示され、既存手法である数値入力による関節定義が可能である。



図 5: 関節情報フレーム



図 6: 詳細編集ダイアログ

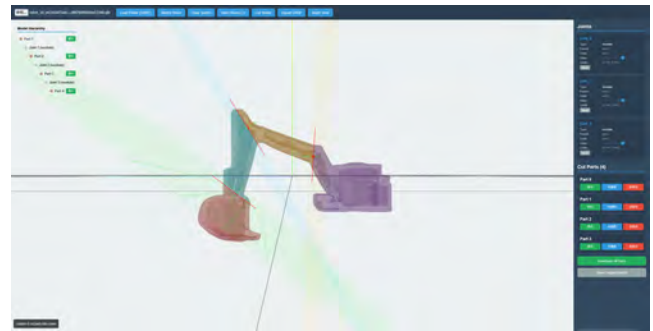
3.3 運動学フィードバック

定義した関節は図 5 に示す関節情報フレーム内にある Value スライダーを操作することで、変位を変更し、3D モデルビューには変位に応じてロボットの姿勢が変化する (図 7)。これによりユーザは運動学計算によるフィードバックを受け取り、現在設計しているロボットの動きと頭の中のロボットの動きを比較することができる。比較した結果、イメージと異なる場合は回転軸の編集を行うことができる。図 5 の回転軸設定/Axis の右の平行と書いてあるボタンは、回転軸が切断面に平行か垂直かを切り替えることができる。その下の回転と書いてあるスライダーは、タイプが平行の場合に表示され、操作することで回転軸が切断面上を回転する。また、関節位置と同様に図 6 に示す詳細編集ダイアログから回転軸を数値入力することもできる。

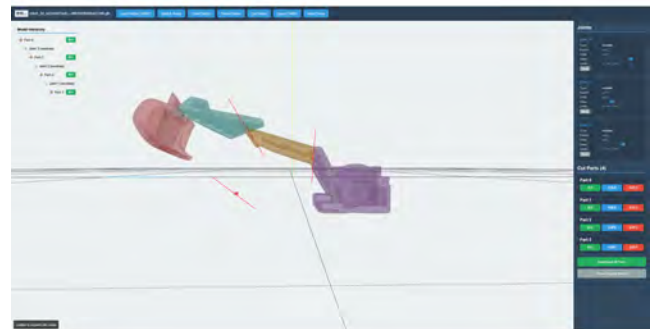
3.4 印刷用データへの変換

分割されたロボットモデルをロボットとして実世界で稼働させるためには、関節として機能する構造が必要である。Sketch2Robpla では関節の機能を持ったメッシュを予め用意し、配置した関節の位置と回転軸を合わせて表示する。このとき、各関節メッシュは回転位置と原点を合わせるように定義しているため、回転軸のみを合わせるためにメッシュの表示位置を幾何変換する。また、配置するメッシュは図 5 に示される関節情報フレームから図 8 に示すプルダウンリストを開いて選択する。

次に、関節メッシュの回転をロボットモデルが妨げないよう、3D モデルの干渉部分を除去する。除去前を図 9(a)、除去後を図 9(b) に示す。除去範囲は切断面と関節メッシュのバウンディングボックスを用いて決定する。除去範囲は切断面を底面、関節メッシュのバウンディングボックスの切断面に垂直な方向の長さを高さとした直方体として定義し、干渉除去の範囲とする。最後に、関節メッシュとロボットの各パーツを結合して、印刷用の 3D モデルが完成する。



(a) 関節回転前



(b) 関節回転後

図 7: 運動学フィードバック

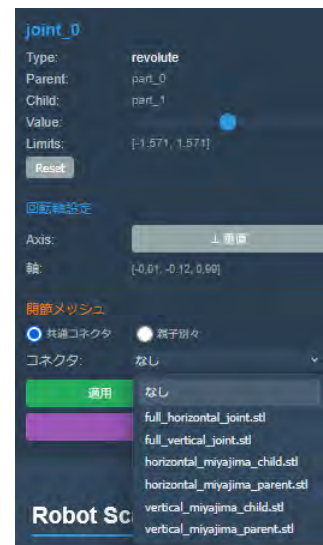
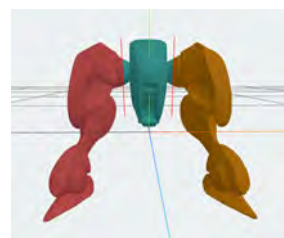
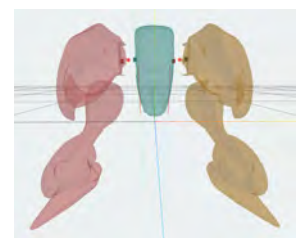


図 8: 関節メッシュプルダウンリスト



(a) 干渉除去前



(b) 干渉除去後

図 9: 干渉除去

4. 結果

4.1 ユーザ実験

ユーザ実験として、本学の 20 代の学生 9 人を被験者として集めた。被験者達は 15 分程度の説明を受けた後、Sketch2Robpla を用いてロボットの形状と構造を設計し、アンケートに答えた。ユーザの設計を図 10 に示す。アンケートの内容は SUS, NASA-TLX, そして以下に示す本研究の課題に関する内容である。

- ロボットを設計した経験について、ある場合は詳細も含む
- 3DCAD を操作した経験、ある場合はどの程度か
- 想像通りの形状を生成できたか
- 想像通りの構造を設計できたか

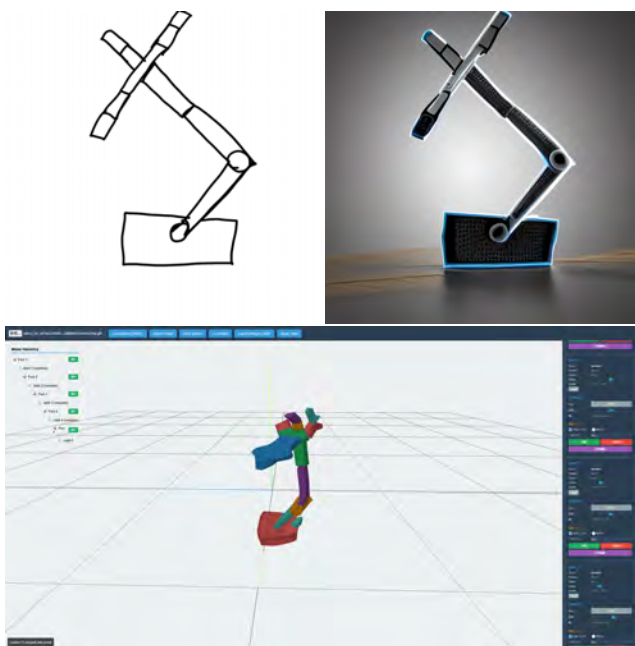


図 10: ユーザによる製作結果

したい」「このシステムは必要以上に複雑だと思う」「このシステムはスムーズに機能し、連携がとれていると思う」「ほとんどの人がこのシステムをすぐに習得できると思う」「このシステムを使っていると自信が持てる」これらの項目に関しては、良い結果が得られた。特に結果が悪かった項目は、「このシステムにはテクニカルサポートが必要」である。ここから、UI 上に操作方法の示唆が足りていないことが考えられる。また、主な操作である線分による分割時にユーザの意図と異なる切断が発生した点や、切断後の接続関係が正常でない場合があった点、その際に接続関係の編集方法が直感的でなく、専門性も高かったことから、テクニカルサポートが必要だったことは間違いない。「このシステムには不規則な点が多いと思う」「このシステムには時間がかかると思う」「このシステムを使う前に学ぶべきことはたくさんあると思う」これらの項目は被験者によって評価が分かれた。また、「このシステムはシンプルで使いやすい」に関してもわずかに悪い評価を受けた。これらの結果は、操作方法の明確な指示、必要な情報の提示ができておらず、また、ロボット設計を作る上で必要になる関節位置、回転軸、接続関係の決定に関して、予期しない動作が多く見られ、それらの編集方法が古典的手法と変わらない数値入力であったことが原因であると考えられる。

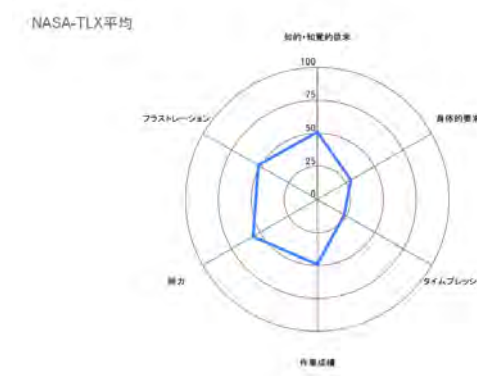


図 12: NASA-TLX の結果 (平均)

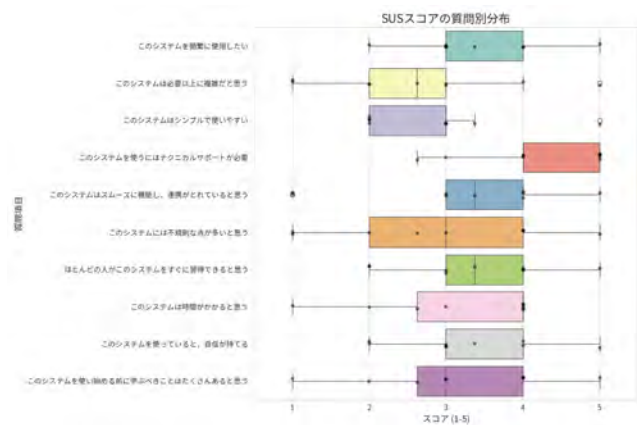


図 11: System Usability Scale の結果

SUS の結果を図 11 に示す。「このシステムを頻繁に使用

次に、NASA-TLX の結果を図 12,13 に示す。作業成績のスコアは既に反転してあるため、スコアが高い方が満足いく成績を残せなかったことを表している。総合スコアの平均値は 48.9 であり、作業負担は中程度であった。項目別に見ると、タイムプレッシャーは特に時間制限を設けなかったことからこの結果になったことが考えられる。身体的要求に関してもシステムにおいて身体に負荷のかかる作業は存在しないため低い結果になったと考えられる。知的・知覚的要求が少し高めの数値であるが、被験者のほとんどが 3DCAD、ロボット設計の未経験者であることが原因の一つであると考えられる。システムの持つ問題は作業成績、努力、フラストレーションの項目に表れている。その中で、作業成績、フラストレーションの項目については回

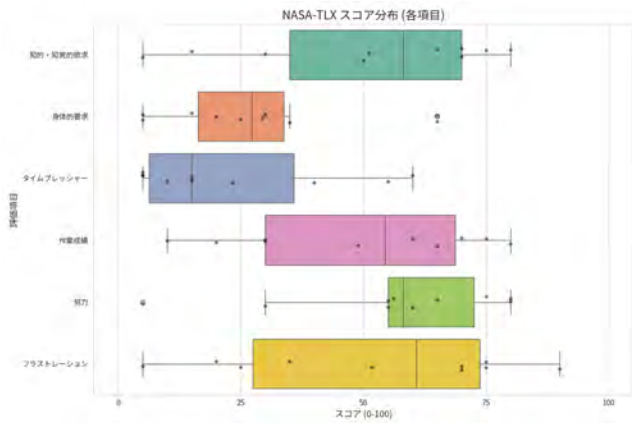
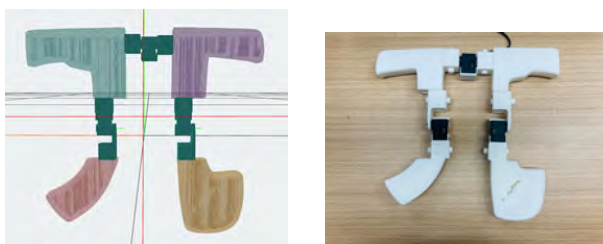


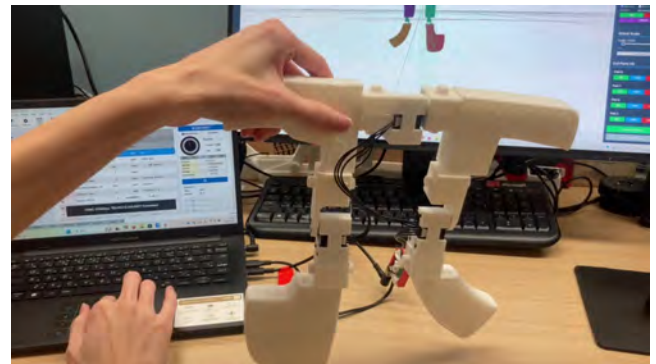
図 13: NASA-TLX の結果 (項目別)



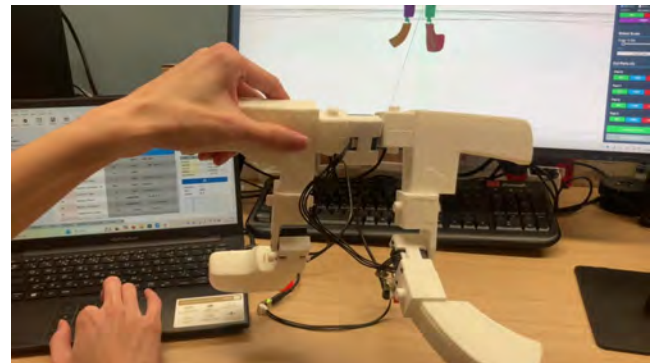
(a) 印刷データ

(b) 印刷結果

図 14: 製作結果



(a) 関節回転前



(b) 関節回転後

図 15: 動作テスト

答の幅が広い。この振れ幅は実験中に不具合がどれだけ発生したかが影響している可能性がある。特に、複雑な 3D モデルを生成した被験者は不具合の発生により満足な関節構造を設計できないことが多かった。最後に努力の項目だが、これが一番高い負荷を示していた。そして、努力スコアが高い被験者は、作業成績、知的・知覚的欲求のスコアも高い傾向にあるため、ユーザがタスクに関する意欲を失っていることや、複雑な計算、判断を強いられている可能性がある。

4.2 製作結果

システムで設計したロボット (図 14(a)) を 3D プリンタ (FlashForge Adventurer 5M) で印刷した結果が図 14(b) である。製作にかかった時間はロボット本体がパーツ 4 つで 8 時間 14 分、関節が水平回転関節 1 つと垂直回転関節 2 つで 1 時間 23 分、計 9 時間 37 分であった。モータ (ROBOTIS DYNAMIXEL XL330-M077-t) を用いた動作テストの結果を図 15 に示す。

5. おわりに

本研究では生成 AI を用いたスケッチ入力 3D モデル生成により高い形状自由度を実現し、線分の描画によるモデル分割及び関節配置によって関節構造入力をより直感的にし、運動学フィードバックによる動作の確認、印刷用データへの変換を行うシステム Sketch2Robpla を開発した。

ユーザ実験により、現状の Sketch2Robpla には、ユーザビリティの欠如や動作の不安定性といった問題が見つかった。これらを改善するために、ユーザへの具体的な操作方法の提示、関節とリンクの接続関係入力、関節位置の編集操作を直感的な操作で実現することにより、これらの問題を解決する必要がある。また、生成した画像、3D モデルの一部編集機能を実装、3D-MMLM による関節推定の実装によってシステムをさらなる改善が見込める。

参考文献

- [1] Kevin Black, Noah Brown, Danny Driess, Adnan Esmail, Michael Equi, Chelsea Finn, Niccolo Fusai, Lachy Groom, Karol Hausman, Brian Ichter, et al. π_0 : A vision-language-action flow model for general robot control. *arXiv preprint arXiv:2410.24164*, 2024.
- [2] Jianfeng Xiang, Zelong Lv, Sicheng Xu, Yu Deng, Ruicheng Wang, Bowen Zhang, Dong Chen, Xin Tong, and Jiaolong Yang. Structured 3d latents for scalable and versatile 3d generation. *arXiv preprint arXiv:2412.01506*, 2024.
- [3] Dmitry Tochilkin, David Pankratz, Zexiang Liu, Zixuan Huang, , Adam Letts, Yangguang Li, Ding Liang, Christian Laforte, Varun Jampani, and Yan-Pei Cao. Tripotr: Fast 3d object reconstruction from a single image. *arXiv preprint arXiv:2403.02151*, 2024.
- [4] Yunhan Yang, Yufan Zhou, Yuan-Chen Guo, Zi-Xin Zou, Yukun Huang, Ying-Tian Liu, Hao Xu, Ding Liang, Yan-Pei Cao, and Xihui Liu. Omnipart: Part-aware 3d generation with semantic decoupling and structural cohesion. In *Proceedings of the SIGGRAPH Asia 2025 Confer-*

- ence Papers, pp. 1–12, 2025.
- [5] Zhengyi Wang, Jonathan Lorraine, Yikai Wang, Hang Su, Jun Zhu, Sanja Fidler, and Xiaohui Zeng. Llama-mesh: Unifying 3d mesh generation with language models. *arXiv preprint arXiv:2411.09595*, 2024.
 - [6] Shuangkang Fang, I Shen, Yufeng Wang, Yi-Hsuan Tsai, Yi Yang, Shuchang Zhou, Wenrui Ding, Takeo Igarashi, Ming-Hsuan Yang, et al. Meshllm: Empowering large language models to progressively understand and generate 3d mesh. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 14061–14072, 2025.
 - [7] Yining Hong, Haoyu Zhen, Peihao Chen, Shuhong Zheng, Yilun Du, Zhenfang Chen, and Chuang Gan. 3d-llm: Injecting the 3d world into large language models. *Advances in Neural Information Processing Systems*, Vol. 36, pp. 20482–20494, 2023.
 - [8] Zekun Qi, Runpei Dong, Shaochen Zhang, Haoran Geng, Chunrui Han, Zheng Ge, Li Yi, and Kaisheng Ma. Shapellm: Universal 3d object understanding for embodied interaction. In *European Conference on Computer Vision*, pp. 214–238. Springer, 2024.
 - [9] Junliang Ye, Zhengyi Wang, Ruowen Zhao, Shenghao Xie, and Jun Zhu. Shapellm-omni: A native multimodal llm for 3d generation and understanding. *arXiv preprint arXiv:2506.01853*, 2025.
 - [10] Adriana Schulz, Cynthia Sung, Andrew Spielberg, Wei Zhao, Robin Cheng, Eitan Grinspun, Daniela Rus, and Wojciech Matusik. Interactive robogami: An end-to-end system for design of robots with ground locomotion. *The International Journal of Robotics Research*, Vol. 36, No. 10, pp. 1131–1147, 2017.
 - [11] Vittorio Megaro, Bernhard Thomaszewski, Maurizio Nitti, Otmar Hilliges, Markus Gross, and Stelian Coros. Interactive design of 3d-printable robotic creatures. *ACM Transactions on Graphics (TOG)*, Vol. 34, No. 6, pp. 1–9, 2015.
 - [12] Kevin Qiu, Władysław Pałucki, Krzysztof Ciebiera, Paweł Fijałkowski, Marek Cygan, and Łukasz Kuciński. Robomorph: Evolving robot morphology using large language models. *arXiv preprint arXiv:2407.08626*, 2024.
 - [13] Jianfeng Xiang, Zelong Lv, Sicheng Xu, Yu Deng, Ruicheng Wang, Bowen Zhang, Dong Chen, Xin Tong, and Jiaolong Yang. Structured 3d latents for scalable and versatile 3d generation. *arXiv preprint arXiv:2412.01506*, 2024.